Supplementary A – Further Analysis

Contents

SIAMCAT Results 5
A - Stool 5
B - Saliva 11
C - Biopsy 17
D – Stool + Saliva + Biopsy 23
Concordance Plot and BAS Results 30
E-Biopsy 30
A-Phylum 30
BAS 32
B-Class 38
BAS 40
C-Order 46
BAS 48
D-Family 54
BAS 56
E-Genus 62
BAS 64
F-Stool 70
A-Phylum 70
BAS 72
B-Class 78
BAS 80
C-Order 86
BAS 88
D-Family 94
BAS 96
E-Genus 102
BAS 104
G-Saliva 110
A-Phylum 110
BAS 112
B-Class 118

BAS 120
C-Order 126
BAS 128
D-Family 130
BAS 132
E-Genus 134
BAS 136
GAMLSS-BEZI with Random Effect Results 142
Phylum 142
Class 143
Order 144
Family 146
Genus 150
Machine Learning Results 157
ML1 - Biopsy 157
A- Phylum 157
B- Class 174
C- Order 193
D- Family 220
E- Genus 253
ML2 - Stool 308
A- Phylum 308
B- Class 323
C- Order 341
D- Family 363
E- Genus 394
ML3 - Saliva 439
A- Phylum 439
B- Class 456
C- Order 473
D- Family 496
E- Genus 522
Zero-Inflated Generalized Dirichlet Multinomial (ZIGDM) Tests 555

Biopsy 555
Stool 555
Saliva 555
General 556

SIAMCAT Results

A - Stool

Table A (Stool) .Area under the ROC curve (AUC) for stool samples stratified by algorithm, feature filter, normalization method, and taxonomic level.

Algorithm	Filter	Phylum		Class		Order		Family		Genus	
		RU	LU	RU	LU	RU	LU	RU	LU	RU	LU
LASSO	AB	0.54	0.63	0.74	0.66	0.67	0.69	0.55	0.54	0.49	0.43
	CA	0.59	0.64	0.70	0.62	0.73	0.64	0.65	0.45	0.64	0.51
	PR	0.61	0.64	0.67	0.57	0.65	0.55	0.52	0.56	0.57	0.51
	VA	-	0.27	0.25	0.25	0.25	0.25	0.39	0.39	0.28	0.28
	PA	0.63	0.60	0.71	0.63	0.69	0.54	0.64	0.53	0.64	0.63
RIDGE	AB	0.69	0.69	**0.79**	0.61	0.38	0.41	0.70	0.68	0.64	0.72
	CA	0.66	0.69	0.74	0.63	0.37	0.41	0.81	0.72	0.75	0.74
	PR	0.71	0.67	0.72	0.63	**0.75**	0.66	0.77	0.73	0.76	0.77
	VA	0.26	0.26	0.22	0.22	0.18	0.18	0.40	0.40	0.30	0.30
	PA	**0.76**	0.67	**0.79**	0.68	0.71	0.55	0.79	0.74	**0.77**	0.74
Random Forest	AB	0.70	0.63	0.70	0.60	0.66	0.71	0.67	0.67	0.68	0.68
	CA	0.70	0.69	0.70	0.63	0.64	0.67	0.69	0.60	0.73	0.68
	PR	0.70	0.64	0.76	0.60	0.68	0.65	0.64	0.64	0.67	0.70
	VA	0.46	0.46	0.31	0.31	0.31	0.31	0.43	0.43	0.49	0.49
	PA	0.68	0.68	0.76	0.61	0.66	0.64	0.66	0.70	0.69	0.71
Logit Spike	AB	0.63	0.66	0.68	0.69	0.72	0.65	0.69	0.70	0.68	0.63
	CA	0.70	0.63	0.66	0.67	0.68	0.69	0.64	0.68	0.68	0.63
	PR	**0.76**	0.65	0.68	0.66	0.60	0.67	0.68	0.64	0.72	0.65
	VA	0.65	**0.77**	0.73	**0.79**	**0.75**	**0.76**	**0.86**	**0.82**	0.69	**0.78**
	PA	0.70	0.62	0.69	0.69	0.70	0.68	0.67	0.67	0.64	0.70

*__***Abbreviations***__: AB = Abundance, CA = Cumulative Abundance, PR = Prevalence, VA = Variance, PA = PASS, RU = Rank Unit, LU = Log Unit.

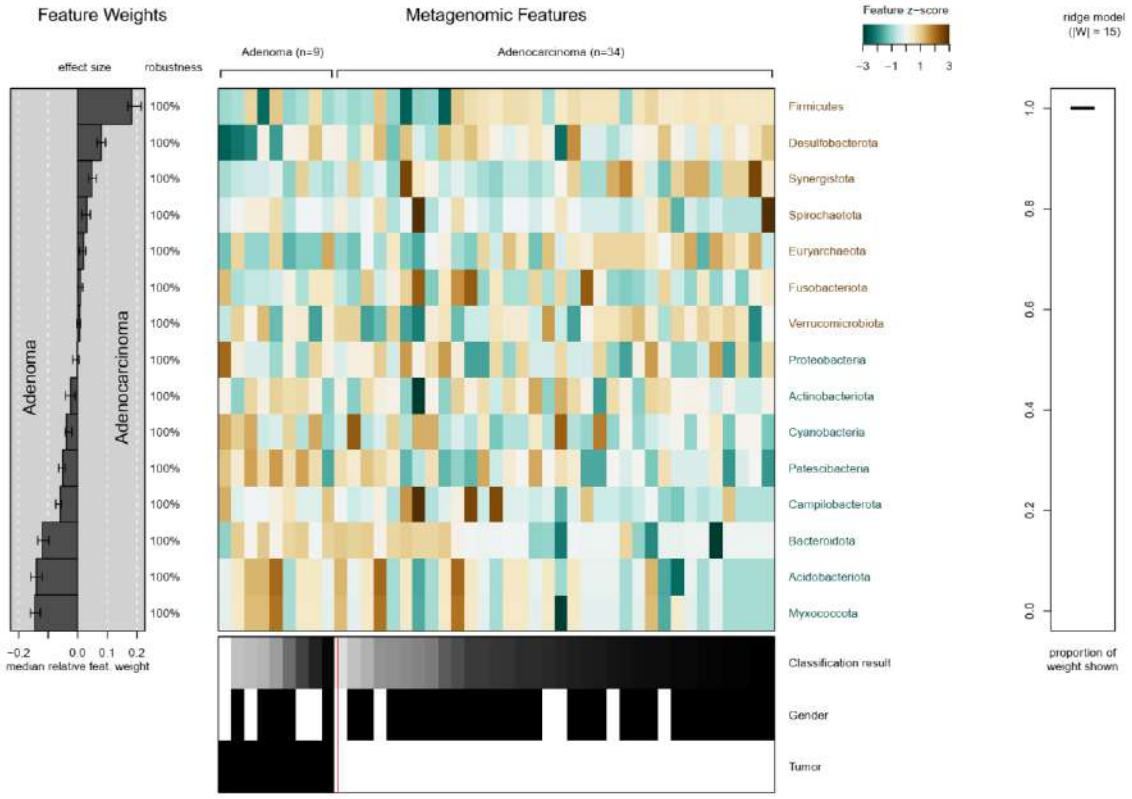


Figure A01: The interpretation results for the **Stool** with **RIDGE** model and **PA** feature selection and **RU** normalization method for **Phylum** category.

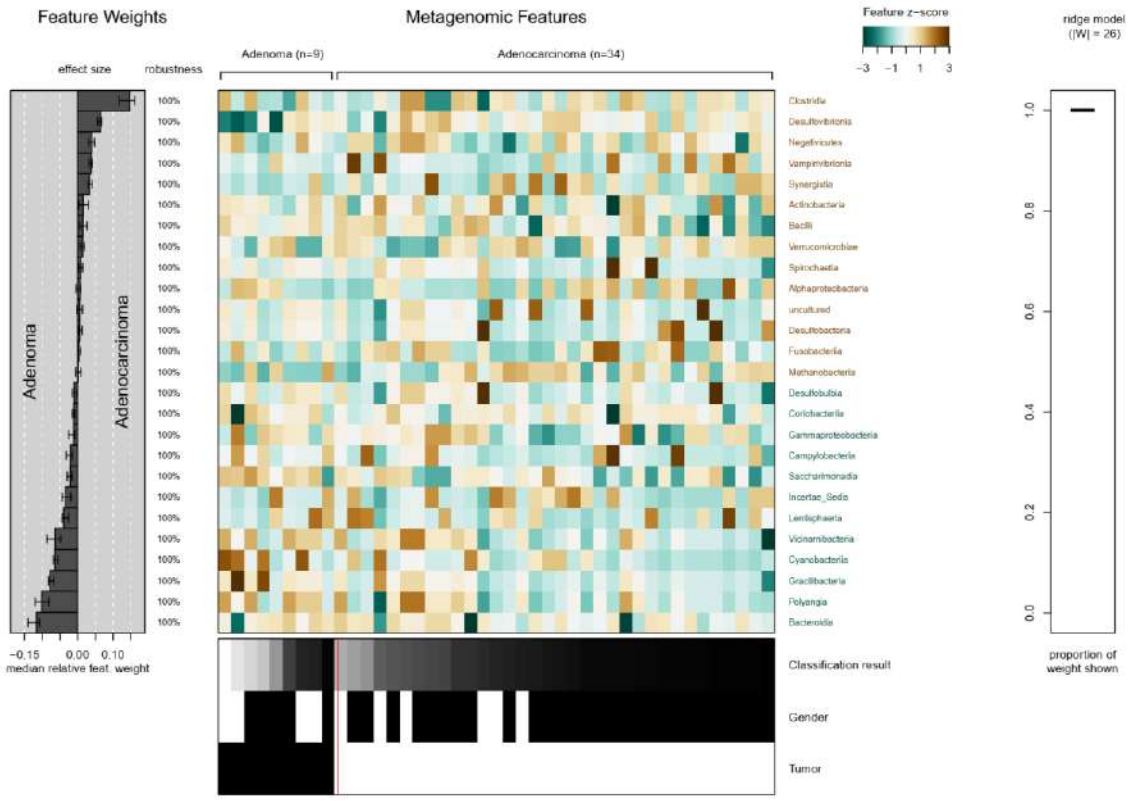


Figure A02: The interpretation results for the **Stool** with **RIDGE** model and **PA** feature selection and **RU** normalization method for **Class** category.

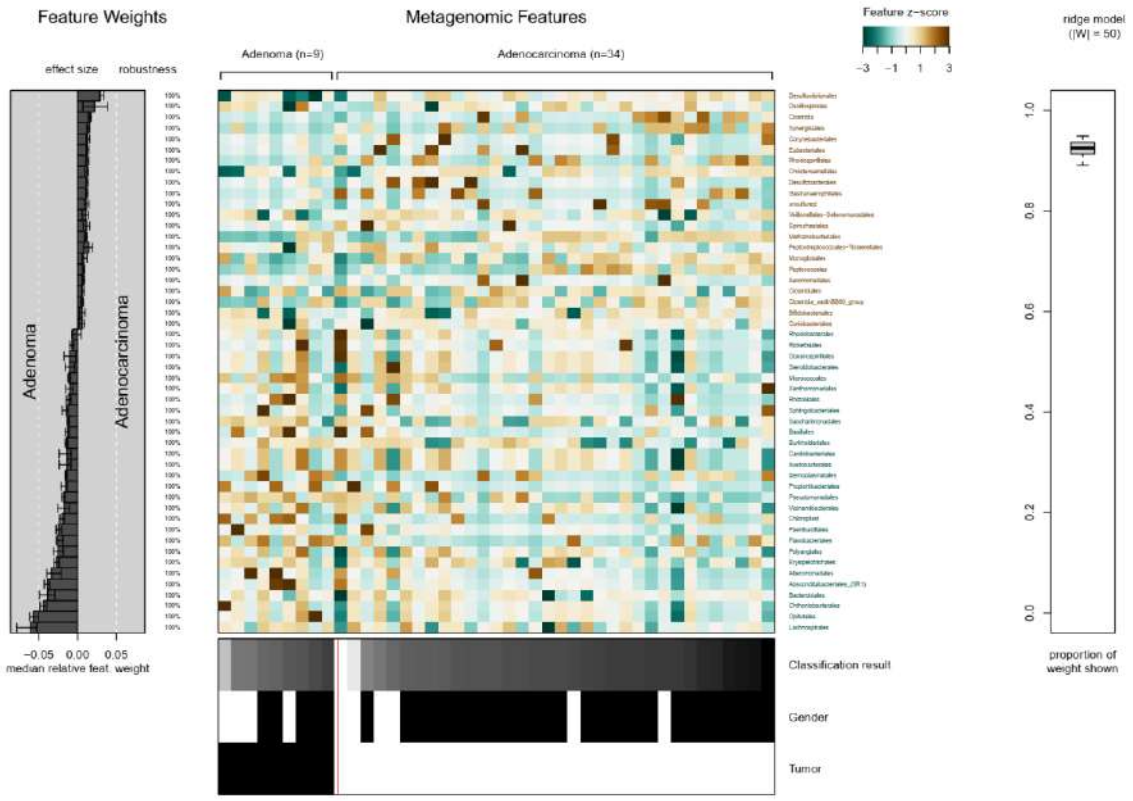


Figure A03: The interpretation results for the **Stool** with **RIDGE** model and **PR** feature selection and **RU** normalization method for **Order** category.

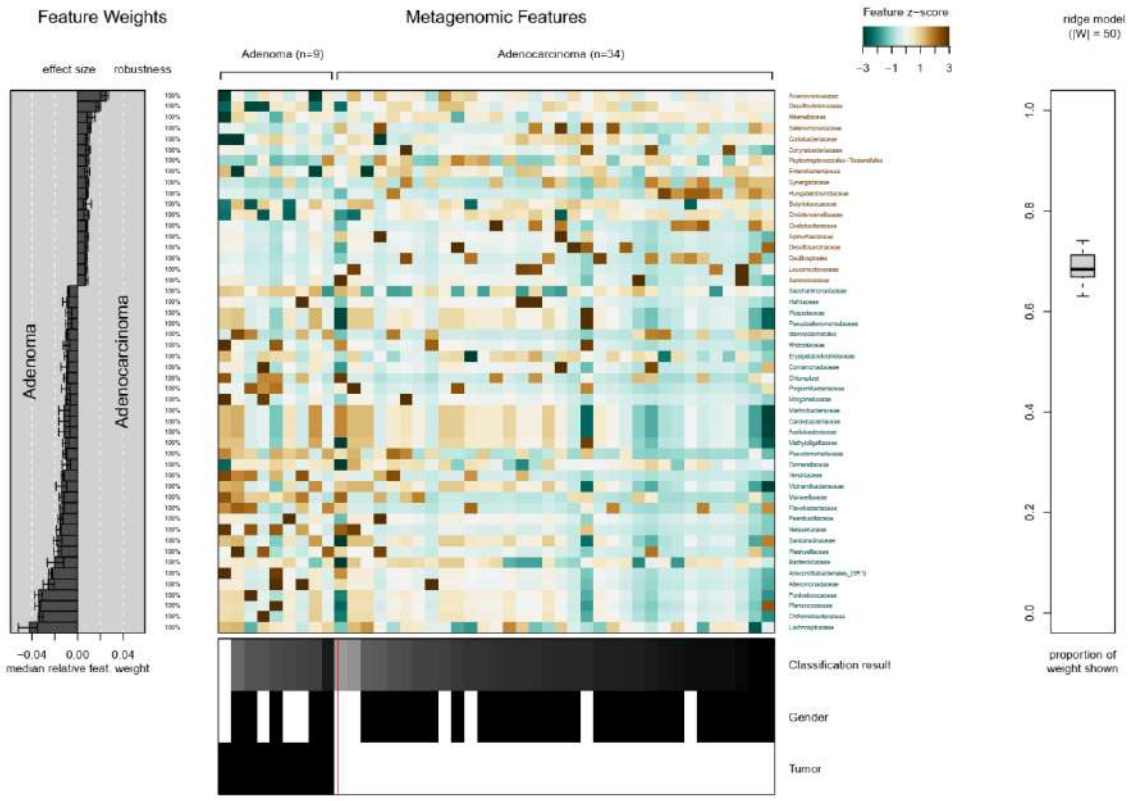


Figure A04: The interpretation results for the **Stool** with **RIDGE** model and **PA** feature selection and **RU** normalization method for **Family** category.

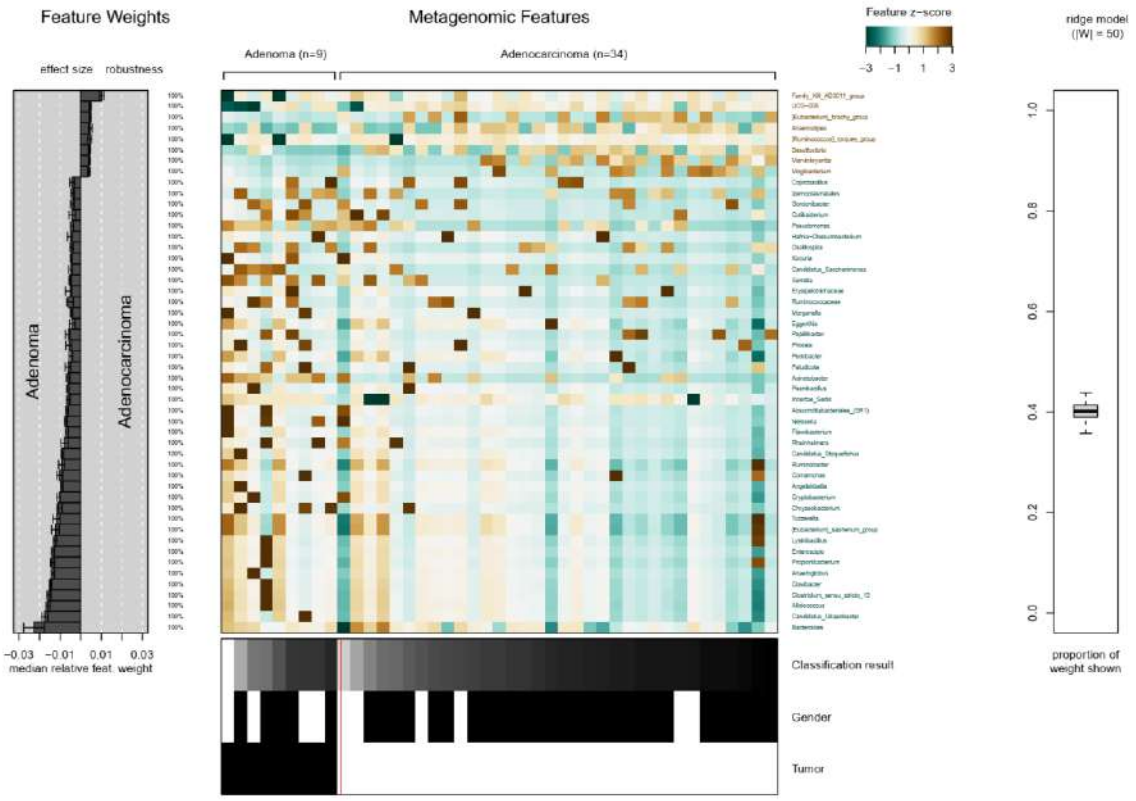


Figure A05: The interpretation results for the **Stool** with **RIDGE** model and **PA** feature selection and **RU** normalization method for **Genus** category.

B - Saliva

Table B (Saliva). Area under the ROC curve (AUC) for saliva samples stratified by algorithm, feature filter, normalization method, and taxonomic level.

Algorithm	Filter	Phylum		Class		Order		Family		Genus	
		RU	LU	RU	LU	RU	LU	RU	LU	RU	LU
LASSO	AB	0.58	0.54	0.62	0.48	0.64	0.65	0.55	0.50	0.62	0.45
	CA	0.54	0.52	0.63	0.50	0.56	0.63	**0.69**	0.64	0.62	0.56
	PR	0.57	0.44	0.43	0.43	0.35	0.56	0.60	**0.72**	0.62	0.56
	VA	-	0.54	0.27	0.27	0.35	0.35	0.31	0.31	0.34	0.34
	PA	0.54	0.47	0.39	0.41	0.39	0.55	0.66	0.65	0.56	0.58
RIDGE	AB	0.58	0.55	0.62	0.54	0.68	0.62	0.35	0.36	0.38	0.40
	CA	0.56	0.52	0.59	0.51	0.34	0.35	0.36	0.37	0.52	0.59
	PR	0.55	0.52	0.40	0.49	0.36	0.37	0.57	0.61	0.52	0.54
	VA	0.40	0.40	0.26	0.26	0.36	0.36	0.34	0.34	0.36	0.36
	PA	0.54	0.50	0.37	0.47	0.37	0.31	0.62	0.62	0.25	0.50
Random Forest	AB	0.48	0.45	0.32	0.41	0.43	0.59	0.25	0.49	0.43	0.49
	CA	0.51	0.45	0.27	0.32	0.43	0.66	0.39	0.54	0.45	0.52
	PR	0.45	0.49	0.32	0.40	0.34	0.66	0.33	0.54	0.31	0.46
	VA	0.48	0.48	0.57	0.57	0.37	0.37	0.24	0.24	0.26	0.26
	PA	0.49	0.39	0.29	0.39	0.40	0.57	0.34	0.57	0.32	0.49
Logit Spike	AB	0.63	0.63	0.61	**0.67**	0.64	0.70	0.61	0.69	0.65	0.66
	CA	**0.64**	0.66	0.65	0.64	**0.68**	0.67	0.66	0.70	**0.67**	**0.69**
	PR	0.62	**0.68**	0.64	0.61	0.65	0.68	0.68	0.67	**0.67**	0.64
	VA	0.60	0.60	0.63	0.66	0.67	0.62	0.66	0.63	0.66	0.65
	PA	0.61	0.63	**0.66**	0.63	0.66	**0.70**	0.65	0.68	0.64	0.65

*__***Abbreviations***__: AB = Abundance, CA = Cumulative Abundance, PR = Prevalence, VA = Variance, PA = PASS, RU = Rank Unit, LU = Log Unit.

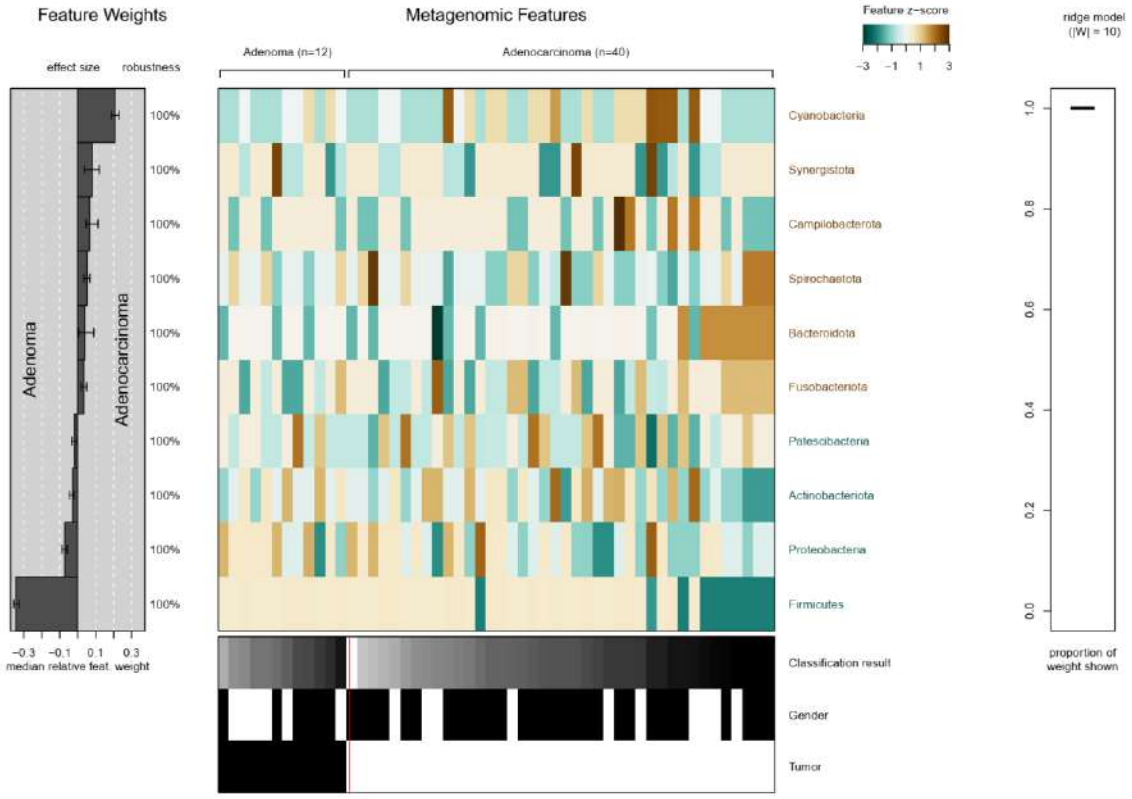


Figure B01: The interpretation results for the **Saliva** with **RIDGE** model and **AB** feature selection and **RU** normalization method for **Phylum** category.

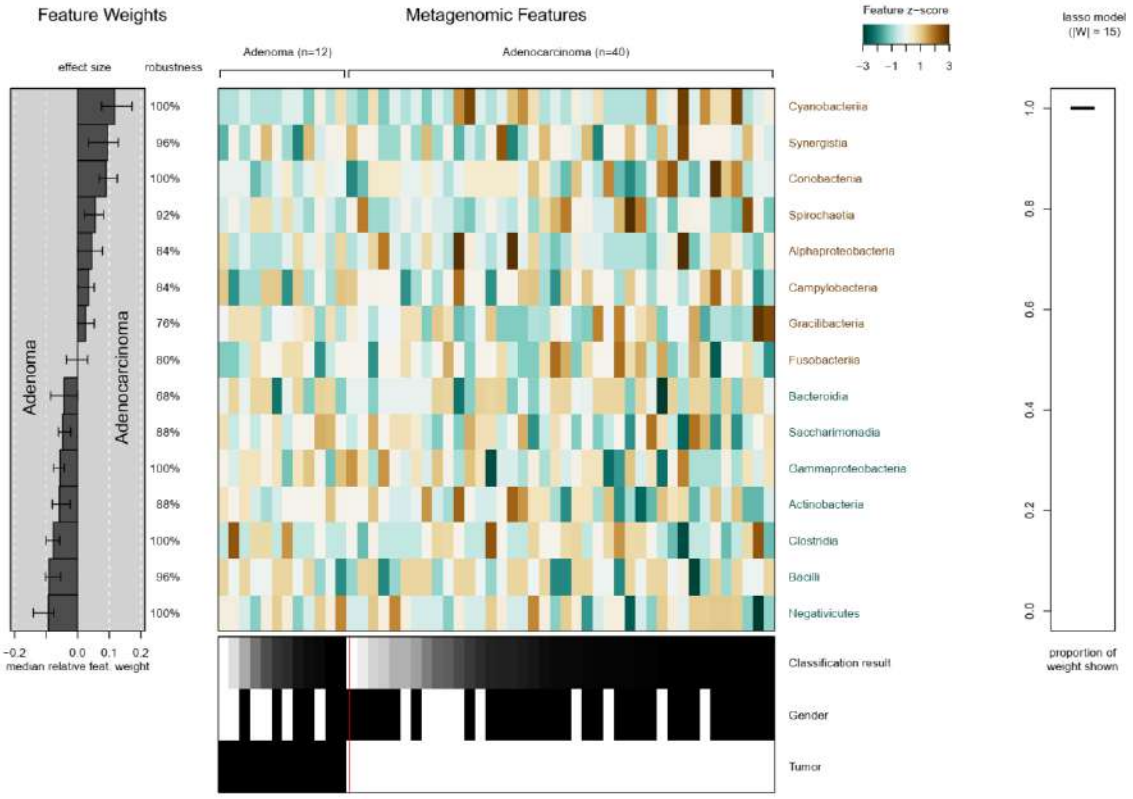


Figure B02: The interpretation results for the **Saliva** with **LASSO** model and **CA** feature selection and **RU** normalization method for **Class** category.

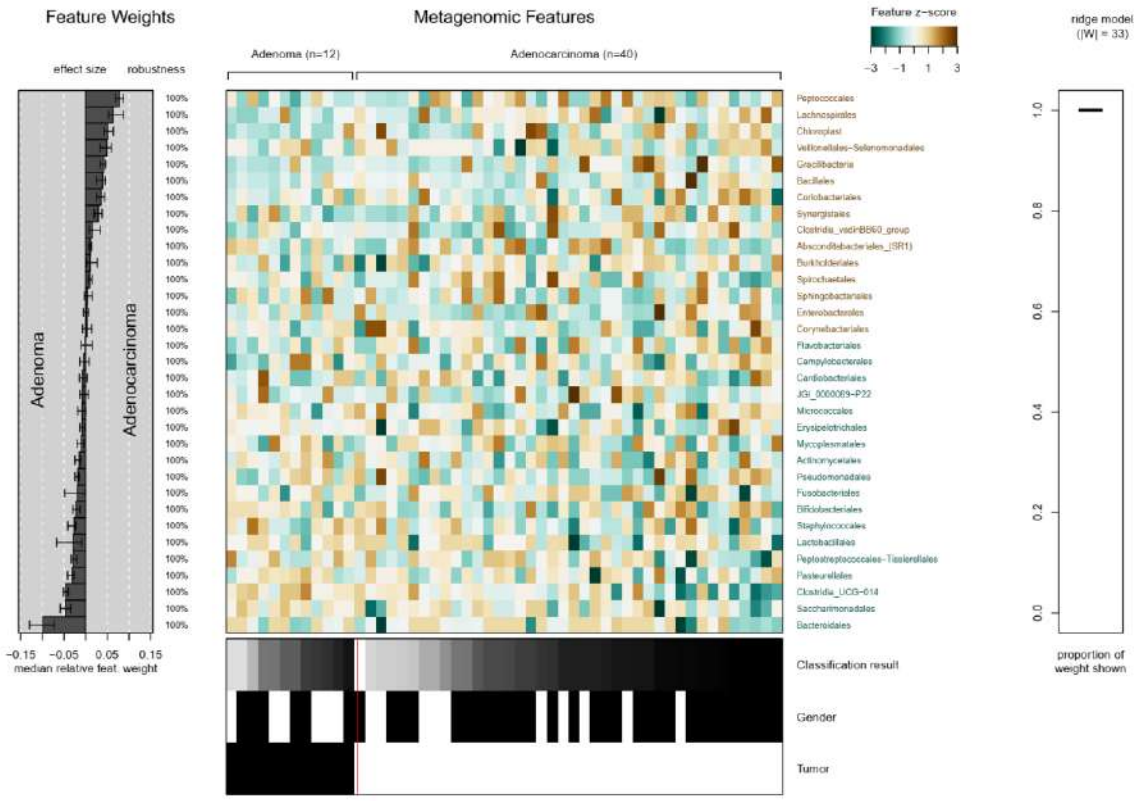


Figure B03: The interpretation results for the **Saliva** with **RIDGE** model and **AB** feature selection and **RU** normalization method for **Order** category.

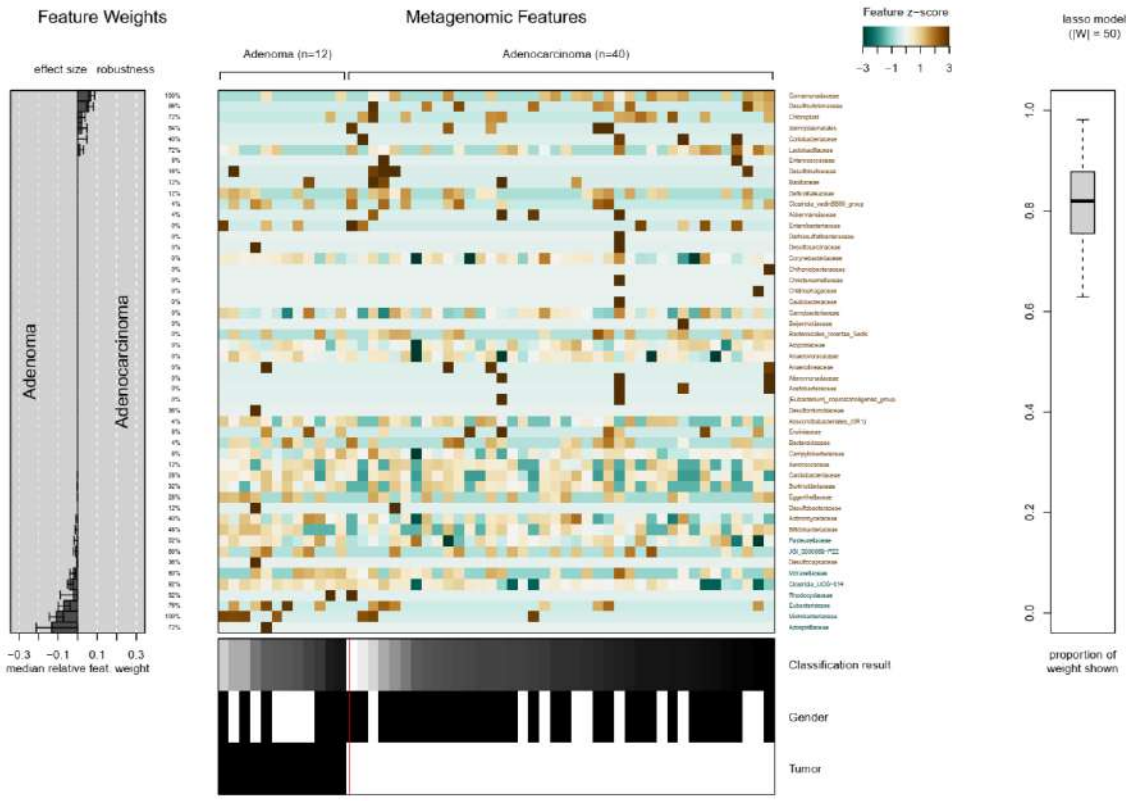


Figure B04: The interpretation results for the **Saliva** with **LASSO** model and **PR** feature selection and **LU** normalization method for **Family** category.

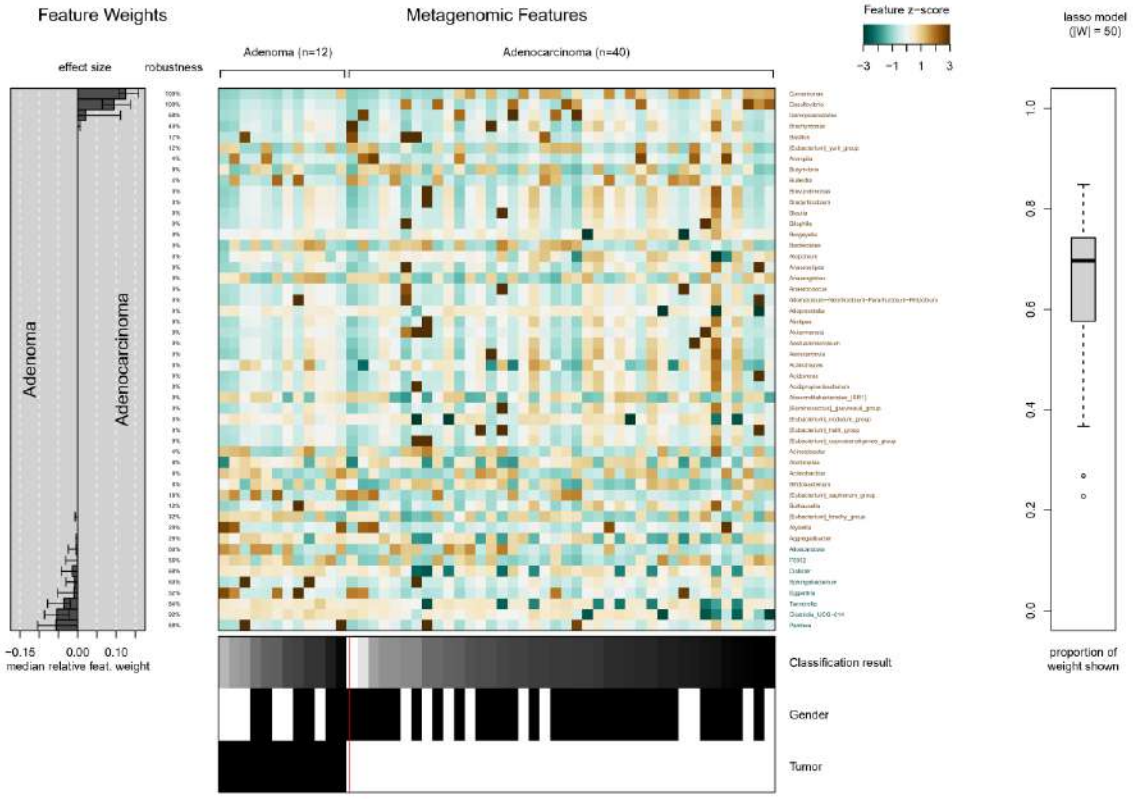


Figure B05: The interpretation results for the **Saliva** with **LASSO** model and **PR** feature selection and **RU** normalization method for **Genus** category.

C - Biopsy

Table C (Biopsy). Area under the ROC curve (AUC) for biopsy samples stratified by algorithm, feature filter, normalization method, and taxonomic level.

Algorithm	Filter	Phylum		Class		Order		Family		Genus	
		RU	LU	RU	LU	RU	LU	RU	LU	RU	LU
LASSO	AB	0.79	0.87	0.79	0.89	0.77	**0.82**	0.72	**0.83**	0.68	0.72
	CA	0.77	0.86	0.70	**0.93**	0.75	**0.82**	0.75	0.76	0.65	0.73
	PR	0.81	0.87	0.76	**0.93**	0.73	0.81	0.70	0.79	0.71	0.67
	VA	-	0.78	-	-	0.71	0.71	0.73	0.73	0.72	0.72
	PA	0.83	0.86	0.72	0.92	0.66	0.81	0.73	0.81	0.71	**0.77**
RIDGE	AB	0.77	0.82	0.86	0.87	0.54	0.55	0.61	0.64	0.56	0.65
	CA	0.78	0.83	0.50	0.36	0.66	0.62	0.66	0.66	0.64	0.63
	PR	0.74	0.83	0.74	0.64	0.64	0.66	0.70	0.64	0.64	0.60
	VA	0.76	0.76	0.73	0.73	0.68	0.68	0.76	0.76	0.69	0.69
	PA	0.77	0.84	0.67	0.61	0.67	0.65	0.68	0.65	0.67	0.62
Random Forest	AB	0.73	0.67	0.78	0.70	0.78	0.70	0.72	0.68	0.66	0.73
	CA	0.69	0.70	0.76	0.77	0.79	0.68	0.69	0.68	0.67	0.71
	PR	0.71	0.70	0.79	0.79	0.75	0.71	0.75	0.68	0.68	0.70
	VA	0.71	0.71	0.70	0.70	0.67	0.67	0.75	0.75	0.77	0.77
	PA	0.77	0.72	0.82	0.76	0.76	0.68	0.77	0.71	0.64	0.76
Logit Spike	AB	0.75	0.78	0.75	0.78	0.80	0.72	**0.80**	0.75	0.72	0.66
	CA	0.77	0.76	0.89	0.80	0.76	0.80	0.76	0.70	0.63	0.68
	PR	**0.94**	**0.89**	**0.93**	0.83	**0.83**	0.77	0.77	0.75	0.73	0.70
	VA	0.76	0.71	0.83	0.69	0.79	0.66	0.77	0.82	**0.84**	**0.86**
	PA	**0.94**	0.86	0.91	0.80	0.80	0.76	**0.80**	0.75	0.68	0.68

*__***Abbreviations***__: AB = Abundance, CA = Cumulative Abundance, PR = Prevalence, VA = Variance, PA = PASS, RU = Rank Unit, LU = Log Unit.



Figure C01: The interpretation results for the **Biopsy** with **RIDGE** model and **PA** feature selection and **RU** normalization method for **Phylum** category.

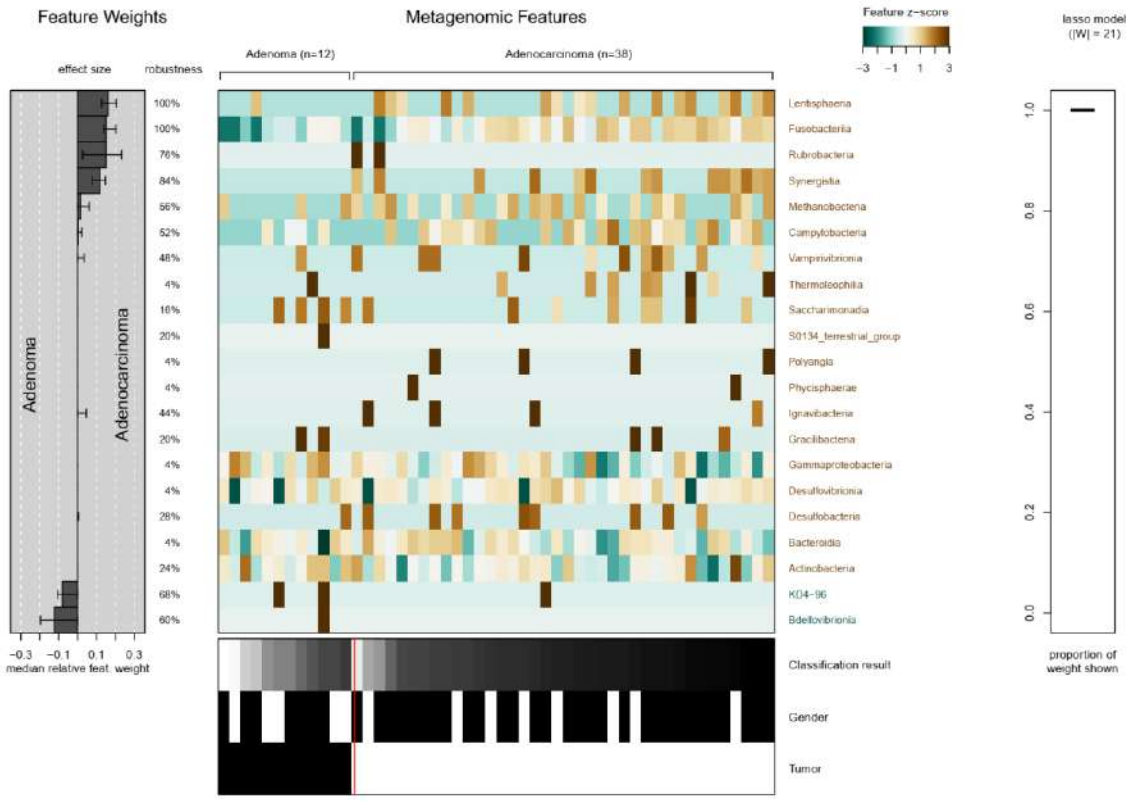


Figure C02: The interpretation results for the **Biopsy** with **LASSO** model and **PR** feature selection and **LU** normalization method for **Class** category.

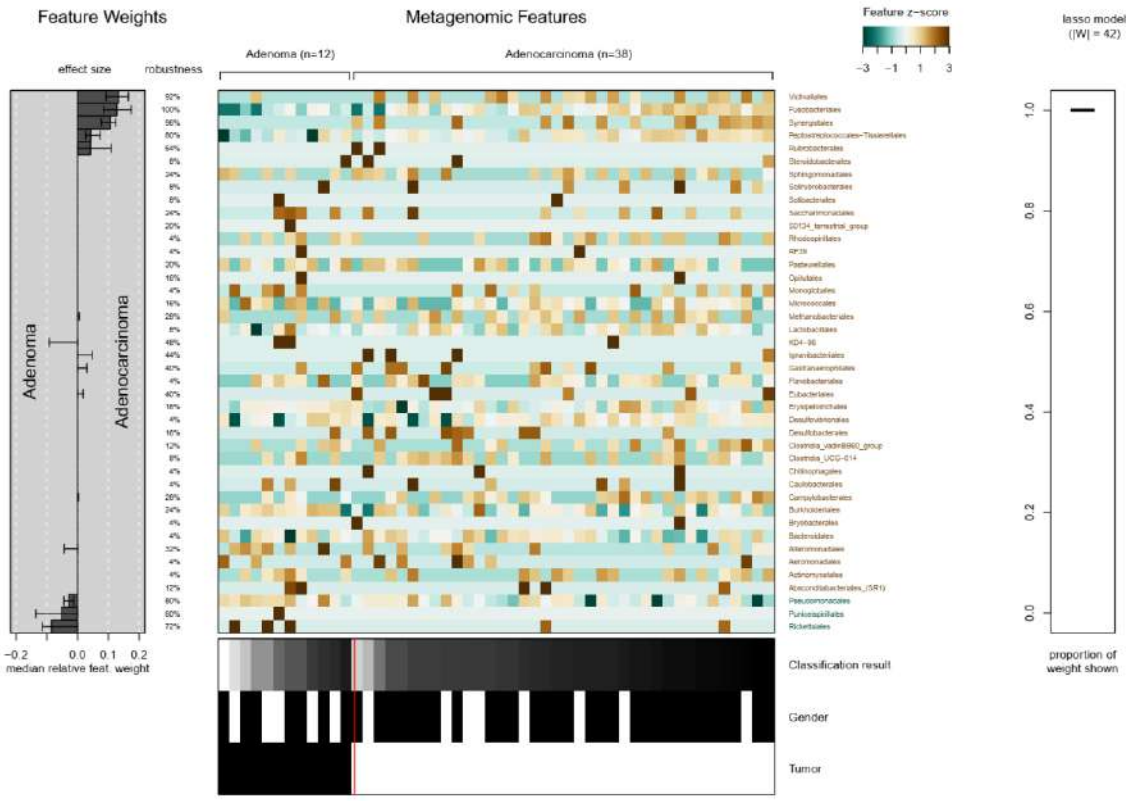


Figure C03: The interpretation results for the **Biopsy** with **LASSO** model and **CA** feature selection and **LU** normalization method for **Order** category.

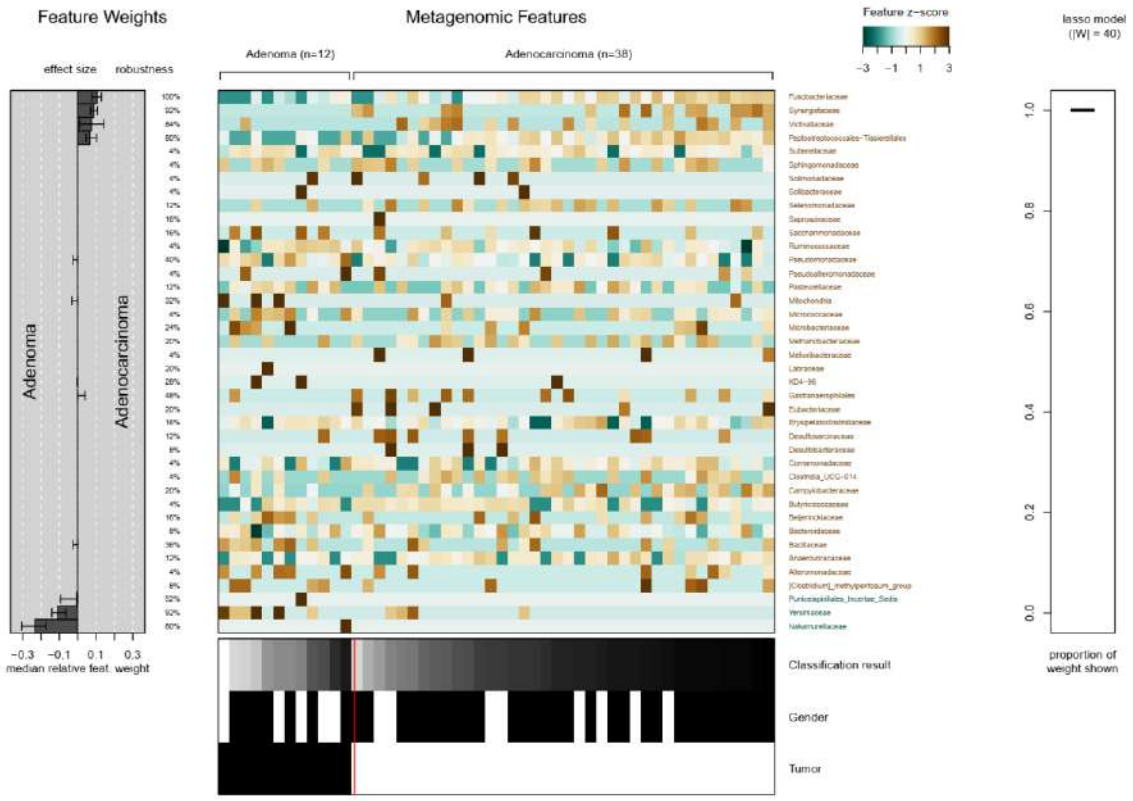


Figure C04: The interpretation results for the **Biopsy** with **LASSO** model and **PA** feature selection and **LU** normalization method for **Family** category.

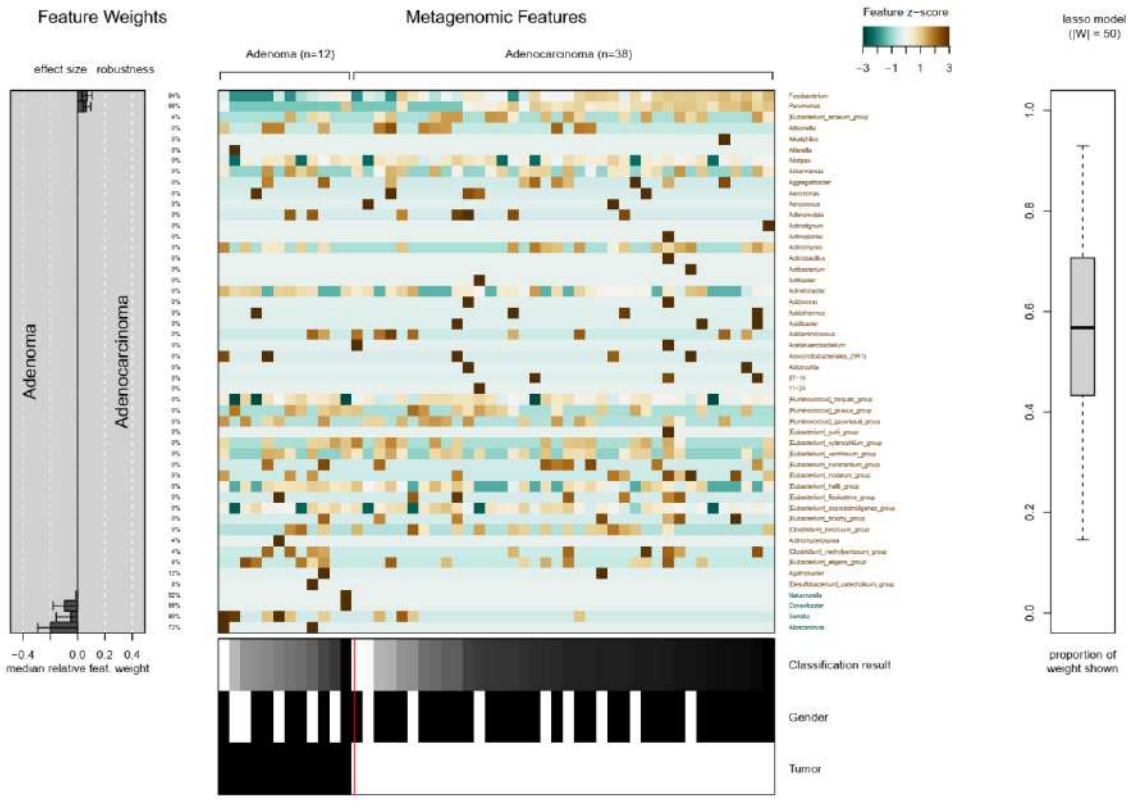


Figure C05: The interpretation results for the **Biopsy** with **LASOO** model and **PA** feature selection and **RU** normalization method for **Genus** category.

D – Stool + Saliva + Biopsy

Stool, Saliva and Biopsy together Analysis

According to the **table A-1**, the **RIDGE** model with **RU** normalization method and **CA** feature selection has the highest **AUC** and it is **0.66** in the **phylum** category for the **Stool, Saliva and Biopsy**. According to the **Figure D01**, three metagenomics including **Firmicutes**, **Myxococcota** and **Fusobacteriota** weights are in direction of **Adenocarcinoma** and three metagenomics including **SAR324_clade(Marine_group_B)**, **Proteobacteria** and **Actinobacteriota** weights are in direction of **Adenoma**.

The **Random Forest** model with **RU** normalization method and **CA** feature selection has the highest **AUC** and it is **0.66** in the **class** category for the **Stool, Saliva and Biopsy**. According to the **Figure D02**, top five metagenomics including **Gracilibacteria**, **Cyanobacteriia**, **Fusobacteriia**, **Actinobacteria** and **Bacilli.**

The **RIDGE** model with **LU** normalization method and **CA** feature selection has the highest **AUC** and it is **0.69** in the **order** category for the **Stool, Saliva and Biopsy**. According to the **Figure D03**, three metagenomics including **Streptomycetales**, **VHS−B4−70** and **Rubrobacterales** weights are in direction of **Adenocarcinoma** and three metagenomics including **S0134_terrestrial_group**, **Opitutales** and **Cyanobacteriales** weights are in direction of Adenoma.

The **RIDGE** model with **RU** normalization method and **CA** feature selection has the highest **AUC** and it is **0.71** in the **family** category for the **Stool, Saliva and Biopsy**. According to the **Figure D04**, three metagenomics including **Oscillospirales**, **Peptostreptococcales−Tissierellales** and **VHS−B4−70** weights are in direction of **Adenocarcinoma** and three metagenomics including **Puniceicoccaceae**, **Thalassospiraceae** and **Hyphomonadaceae** weights are in direction of **Adenoma**.

The **RIDGE** model with **LU** normalization method and **CA** feature selection has the highest **AUC** and it is **0.70** in the **Genus** category for the **Stool, Saliva and Biopsy**. According to the **Figure D05**, three metagenomics including **Prevotellaceae**, **Sinomonas** and **Parascardovia** weights are in direction of **Adenocarcinoma** and three metagenomics including **Clavibacter**, **Sediminispirochaeta** and **Catenisphaera** weights are in direction of **Adenoma**.

Table A-1 – The AUC for **Stool, Saliva and Biopsy together** group by Algorithm, Filter, Normalization Method and Taxonomy.

Algorithm	Filter	Phylum		Class		Order		Family		Genus	
		RU	LU	RU	LU	RU	LU	RU	LU	RU	LU
LASSO	AB	0.63	**0.66**	0.62	**0.67**	0.57	0.63	0.63	0.66	**0.70**	0.66
	CA	0.64	**0.66**	0.62	0.65	0.59	0.66	0.68	**0.70**	0.62	0.64
	PR	0.64	0.65	0.59	0.64	0.63	0.63	0.66	0.61	0.60	0.60
	VA	0.61	0.54	0.56	0.56	0.51	0.51	0.56	0.56	0.55	0.55
	PA	**0.66**	0.64	0.59	0.65	0.57	0.63	0.67	0.63	0.63	0.57
RIDGE	AB	**0.66**	**0.66**	0.66	**0.67**	0.63	0.66	0.67	0.66	0.68	**0.70**
	CA	**0.66**	0.65	0.64	0.66	0.60	**0.69**	**0.71**	**0.70**	0.66	**0.70**
	PR	**0.66**	**0.66**	0.59	0.64	**0.66**	0.64	**0.71**	0.61	0.64	0.67
	VA	0.56	0.56	0.57	0.57	0.53	0.53	0.59	0.59	0.55	0.55
	PA	0.64	0.65	0.63	0.66	0.60	0.63	0.70	0.66	0.64	0.65
Random Forest	AB	0.58	0.58	0.68	0.53	0.59	0.59	0.63	0.64	0.68	0.67
	CA	0.61	0.56	**0.69**	0.51	0.57	0.61	0.60	0.63	0.61	0.64
	PR	0.60	0.59	0.66	0.51	0.59	0.60	0.61	0.56	0.62	0.63
	VA	0.55	0.55	0.49	0.49	0.52	0.52	0.54	0.54	0.49	0.49
	PA	0.60	0.56	**0.69**	0.54	0.57	0.60	0.61	0.60	0.62	0.63

*__***Abbreviations***__: AB = Abundance, CA = Cumulative Abundance, PR = Prevalence, VA = Variance, PA = PASS, RU = Rank Unit, LU = Log Unit.

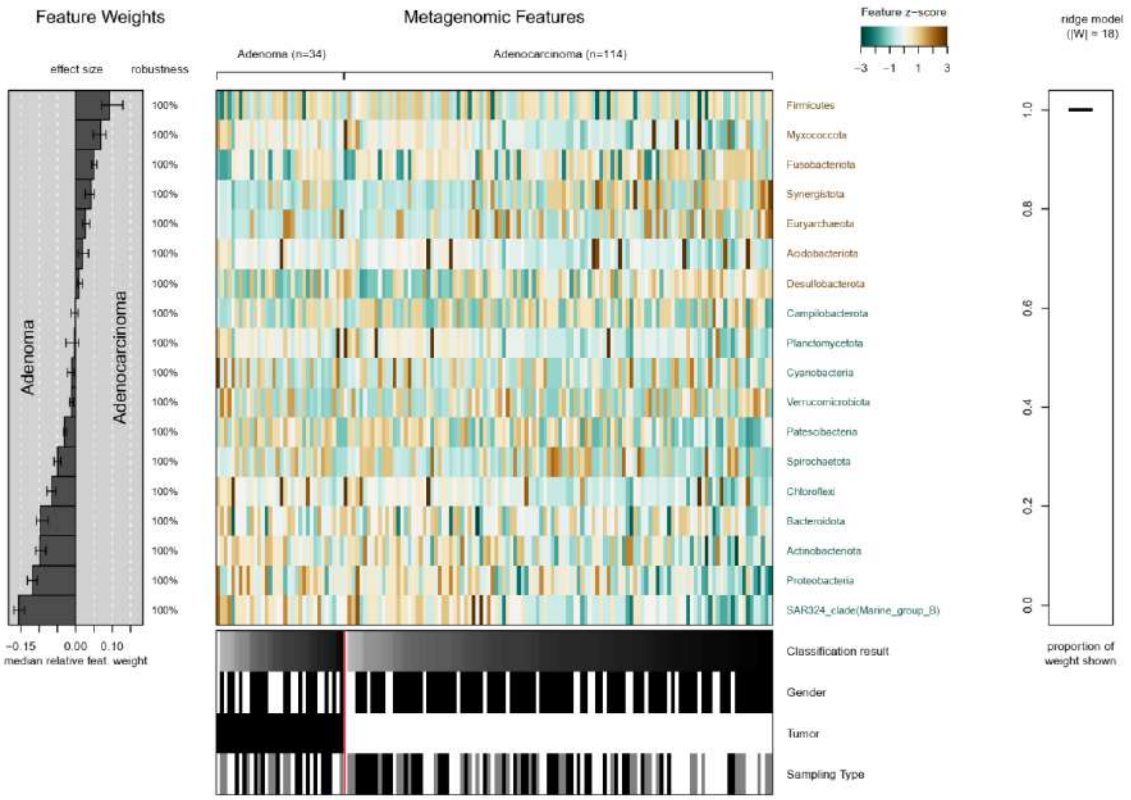


Figure D01: The interpretation results for the **Stool + Saliva + Biopsy** with **RIDGE** model and **CA** feature selection and **RU** normalization method for **Phylum** category.

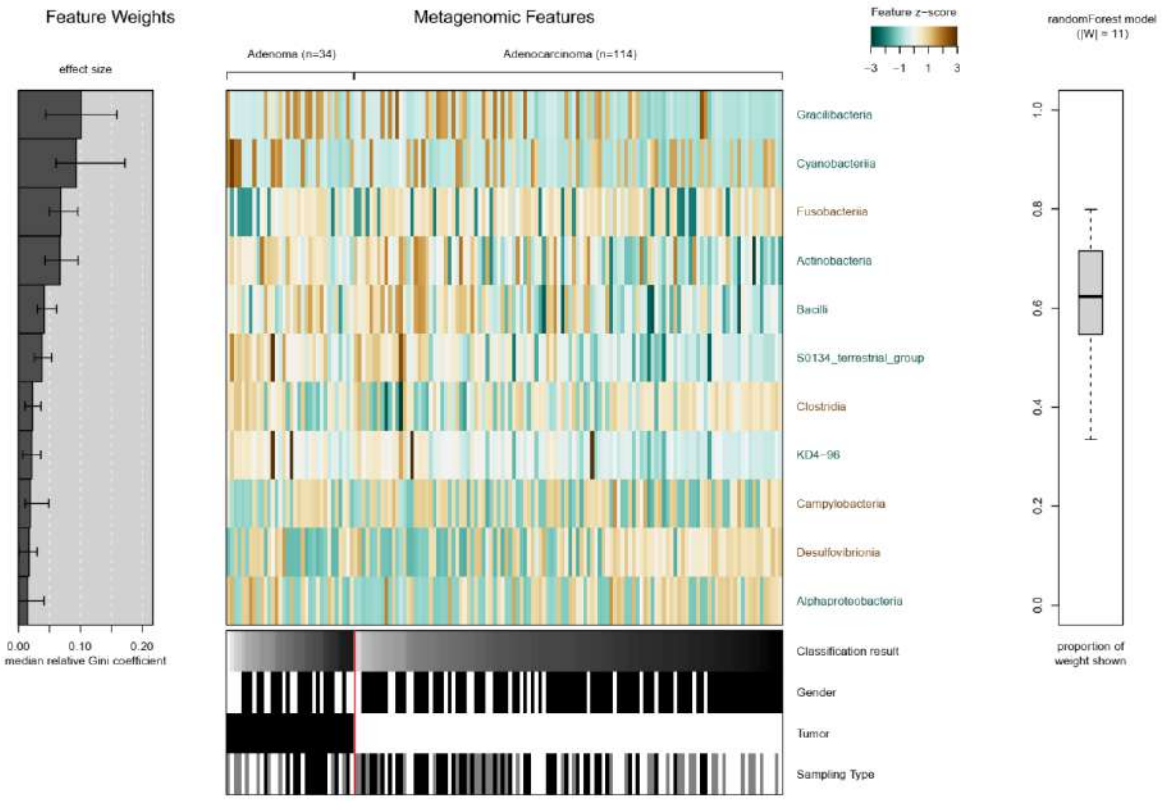


Figure D02: The interpretation results for the **Stool + Saliva + Biopsy** with **Random Forest** model and **CA** feature selection and **RU** normalization method for **Class** category.

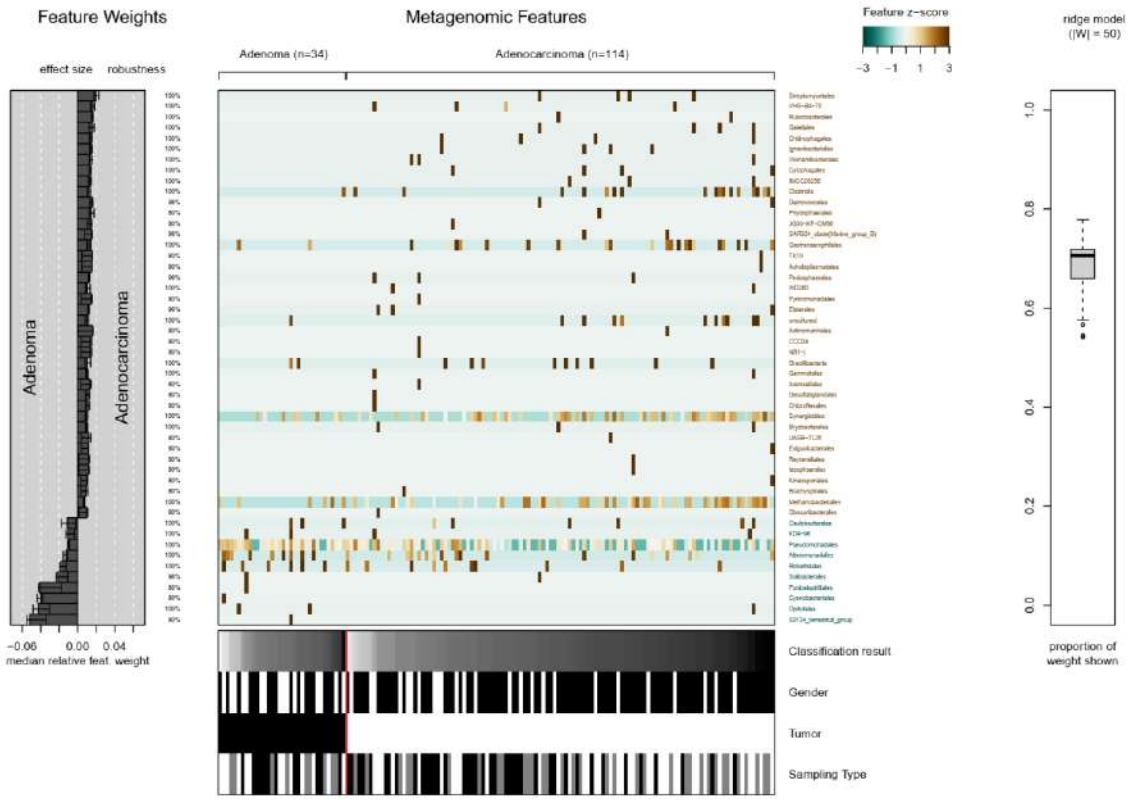


Figure D03: The interpretation results for the **Stool + Saliva + Biopsy** with **RIDGE** model and **CA** feature selection and **LU** normalization method for **Order** category.

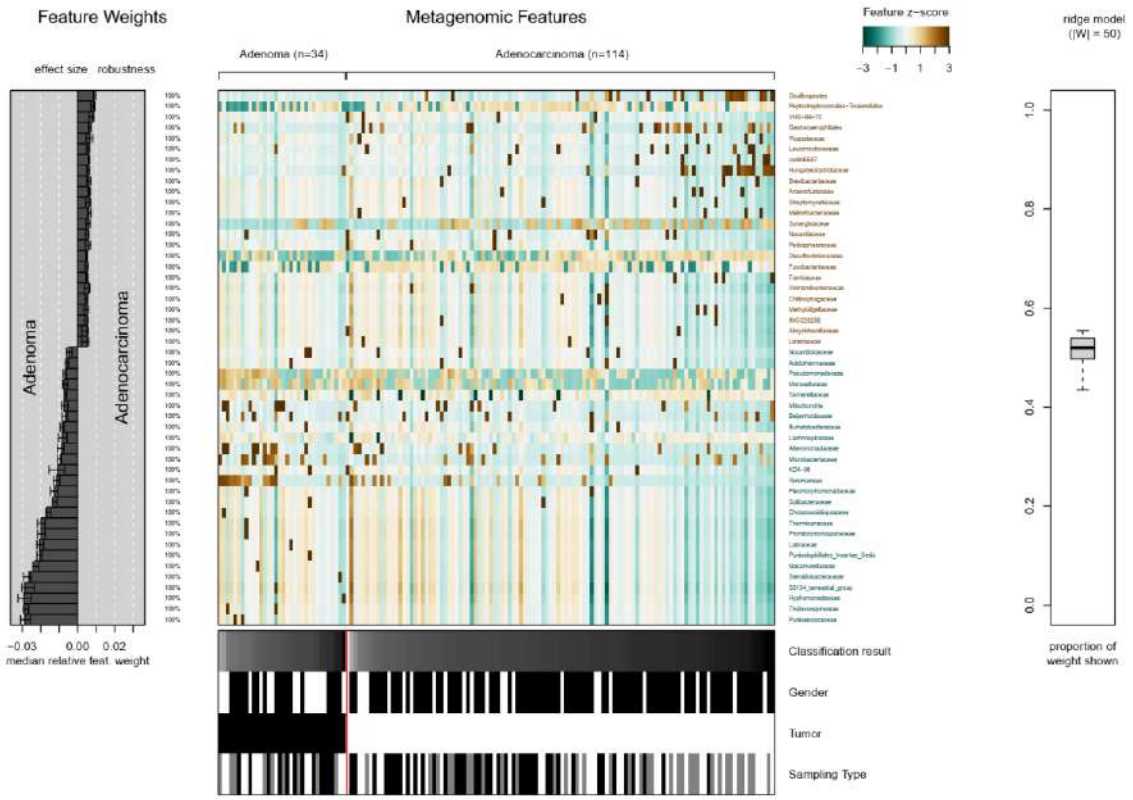


Figure D04: The interpretation results for the **Stool + Saliva + Biopsy** with **RIDGE** model and **CA** feature selection and **RU** normalization method for **Family** category.

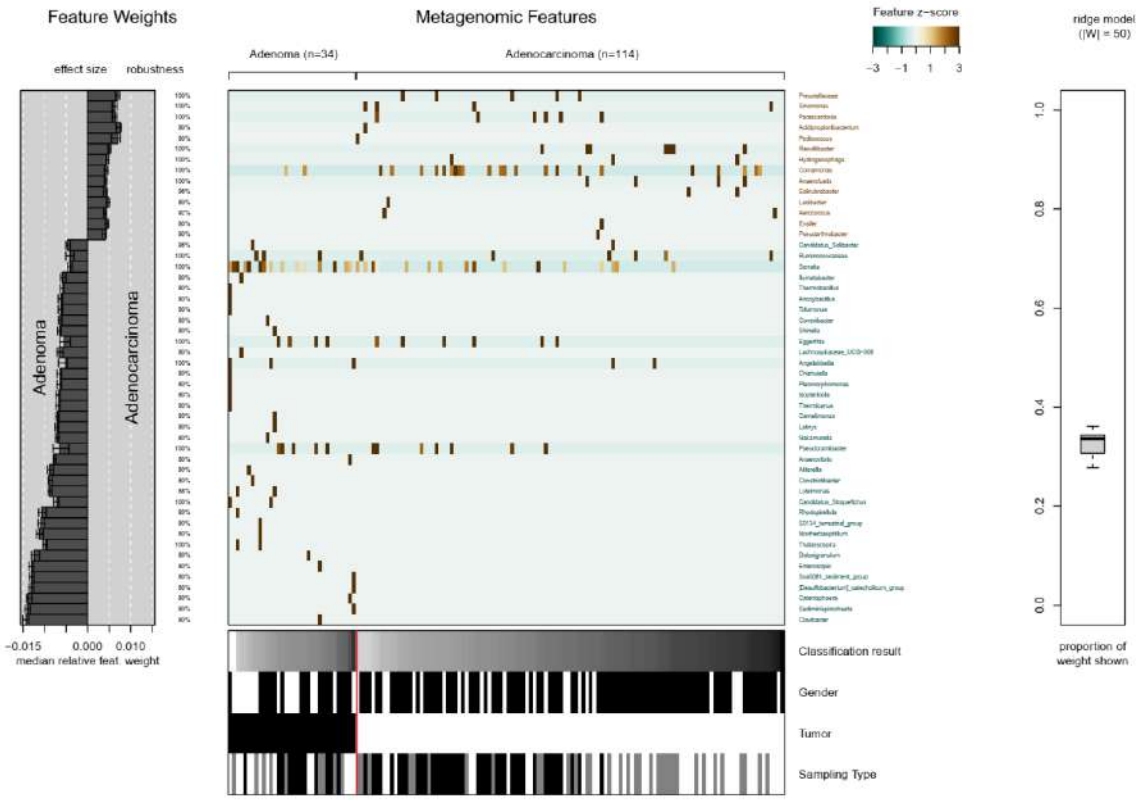


Figure D05: The interpretation results for the **Stool + Saliva + Biopsy** with **RIDGE** model and **CA** feature selection and **LU** normalization method for **Genus** category.

Concordance Plot and BAS Results

The Concordance plot and Bayesian Adaptive Sampling (BAS) are presented for each taxa level and sampling site.

E-Biopsy

A-Phylum

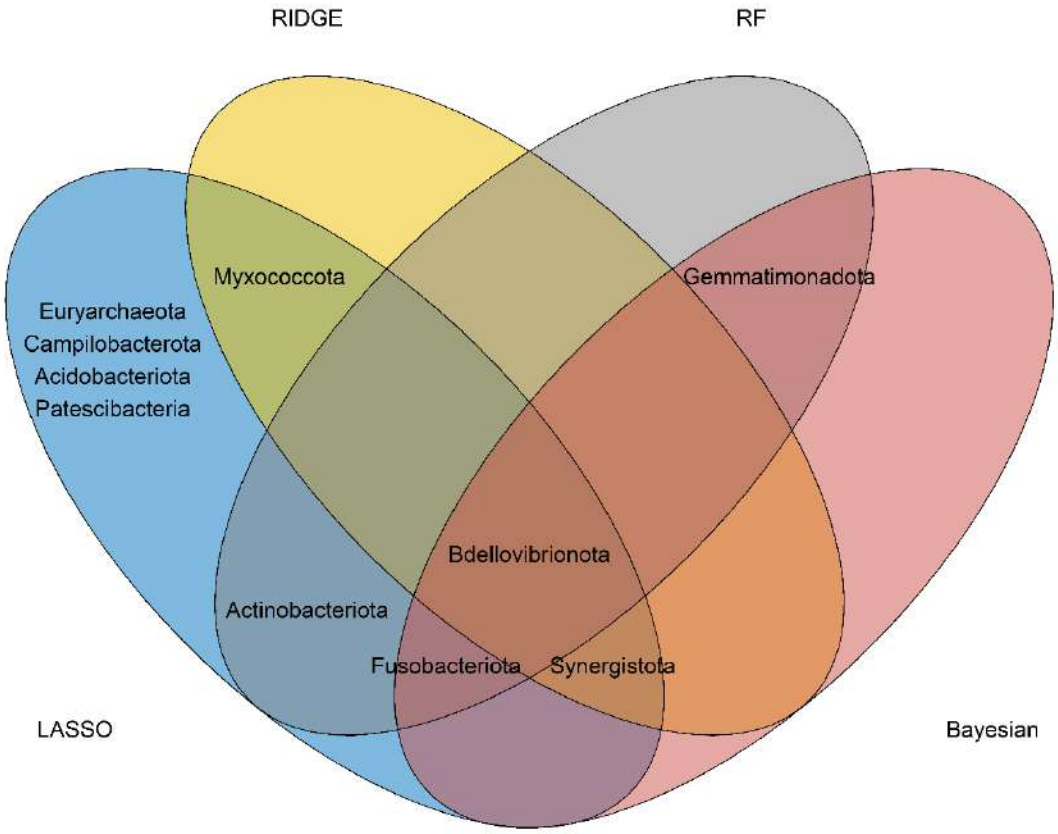


Figure E-A-1 – The Concordance Plot for **Phylum** for **Biopsy**.

Table E-A-1: The Selected method

Row	Method	Filter	Normal	AUC
1	LASSO	PR	LU	0.87
2	RIDGE	PA	LU	0.84
3	RF	PA	RU	0.77
4	**Bayesian**	**PR**	**RU**	**0.94**

Table E-A-2: The common feature names

Common Features for BAS			
Myxococcota	Actinobacteriota	Bdellovibrionota	Fusobacteriota
Synergistota	Gemmatimonadota		

BAS

Log Unit

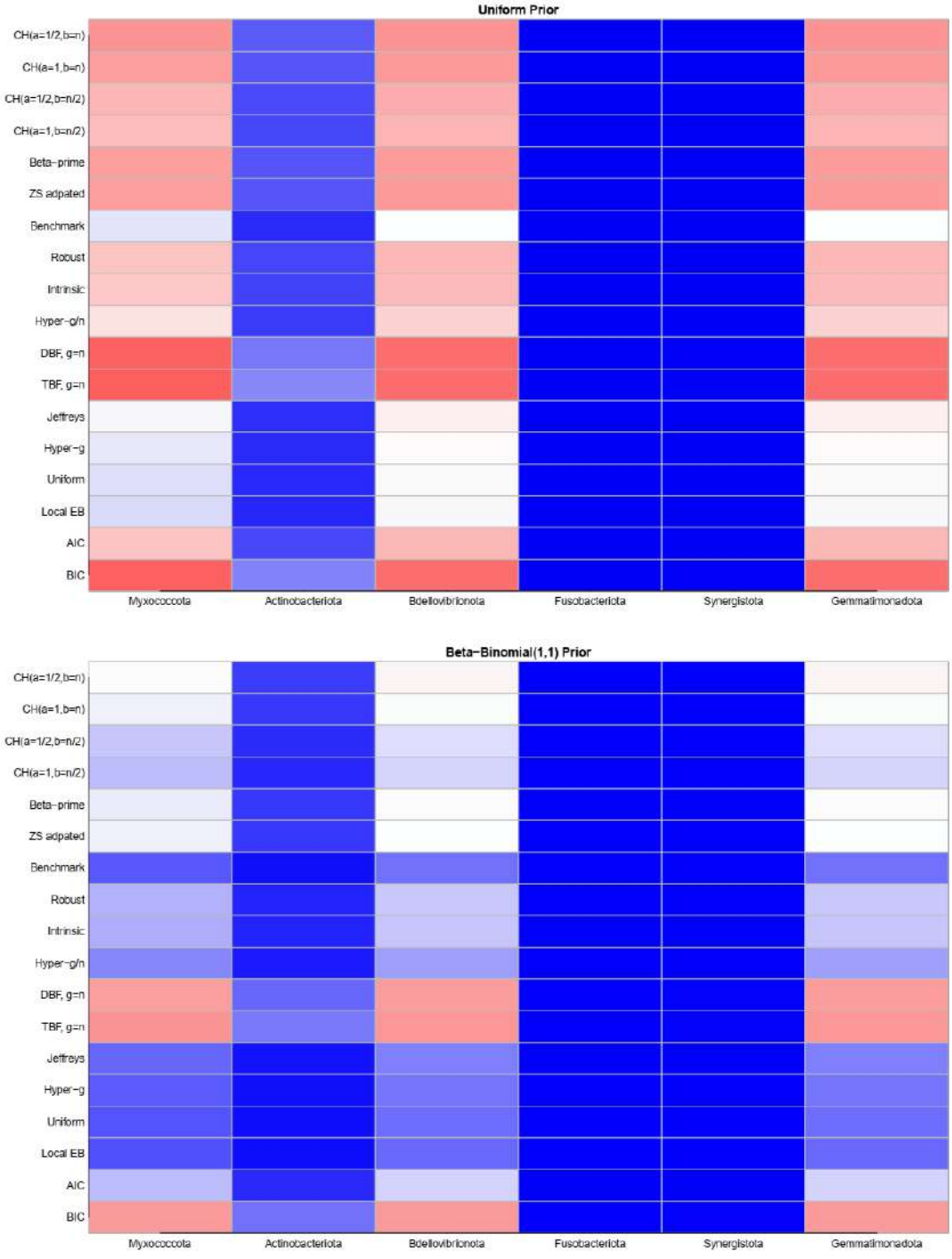


Figure E-A-2 –Marginal posterior inclusion probabilities for the **Phylum** for **Biopsy** with **Log Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

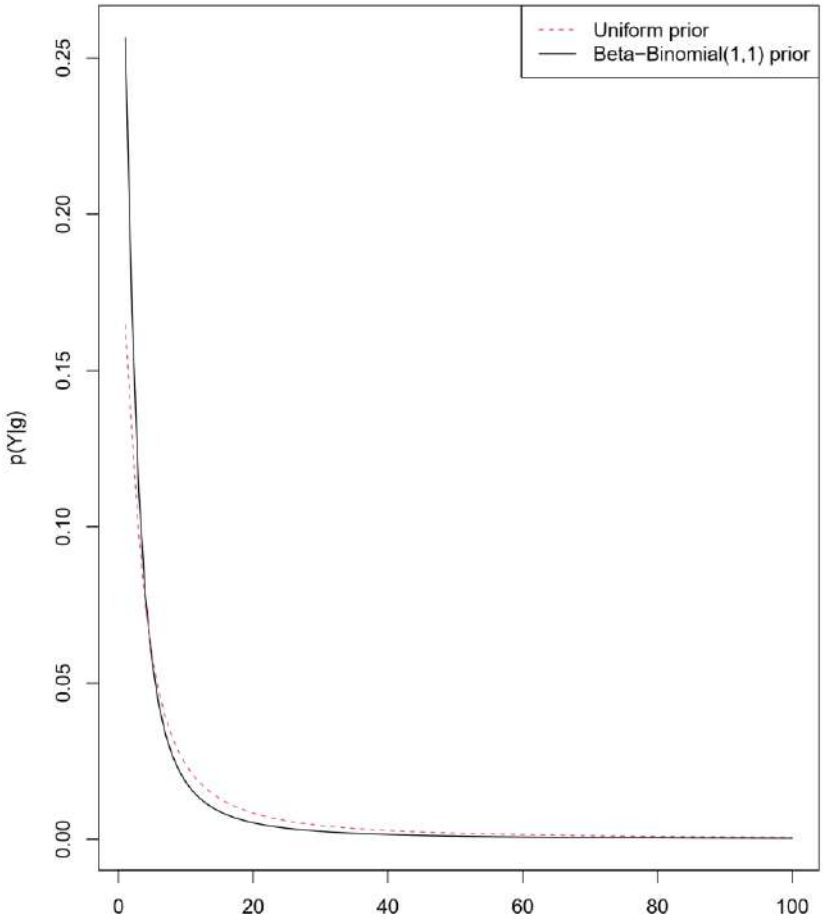


Figure E-A-3: Marginal likelihood of *g* for the **Phylum** for **Biopsy** with **Log Unit** Normalization method.

Table E-A-3: Prediction accuracy for the **Biopsy (Phylum) Log Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.8929	0.8945	11.6591	26.4868	0.3591	0.3620	0.1013	0.1032
02_CH(a=1,b=n)	0.8930	0.8937	41.0239	21.7811	0.3587	0.3596	0.1012	0.1020
03_CH(a=1/2,b=n/2)	0.8928	0.8939	18.6513	23.3163	**0.3573**	0.3645	0.1007	0.1015
04_CH(a=1,b=n/2)	0.8922	0.8940	23.0574	**15.5794**	0.3581	0.3611	0.1005	0.1008
05_Beta-prime	0.8927	0.8941	43.7450	24.6154	0.3595	0.3608	0.1014	0.1027
06_ZS adpated	0.8931	0.8940	41.3775	22.0454	0.3588	0.3596	0.1012	0.1020
07_Robust	0.8934	0.8939	16.8415	26.5488	0.3589	0.3569	0.1000	0.0998
08_Intrinsic	0.8928	0.8937	17.3803	27.6591	0.3577	**0.3553**	**0.0997**	**0.0994**
09_Hyper-g/n	0.8968	0.8970	33.7517	45.8515	0.3603	0.3591	0.1000	0.1004
10_Benchmark	0.9000	0.9009	37.9613	63.1430	0.3724	0.3791	0.1049	0.1065
11_DBF, g=n	0.8924	0.8931	14.9122	43.6751	0.3590	0.3638	0.1021	0.1041
12_TBF, g=n	0.8943	0.8953	20.3479	15.8506	0.3610	0.3642	0.1029	0.1055
13_Jeffreys	0.9036	0.9028	35.4538	36.5628	0.3585	0.3623	0.1014	0.1016
14_Hyper-g	**0.9039**	**0.9039**	30.4249	42.4562	0.3604	0.3646	0.1027	0.1028
15_Uniform	0.9034	0.9030	60.6316	54.3043	0.3647	0.3695	0.1040	0.1040
16_Local EB	0.7231	0.7016	1502.7837	2243.1735	0.6602	0.6837	0.2120	0.2177
17_AIC	0.8923	0.8891	30.7169	34.2283	0.3581	0.3645	0.1008	0.1010
18_BIC	0.8924	0.8943	**10.6832**	21.4936	0.3606	0.3613	0.1018	0.1029

Rank Unit

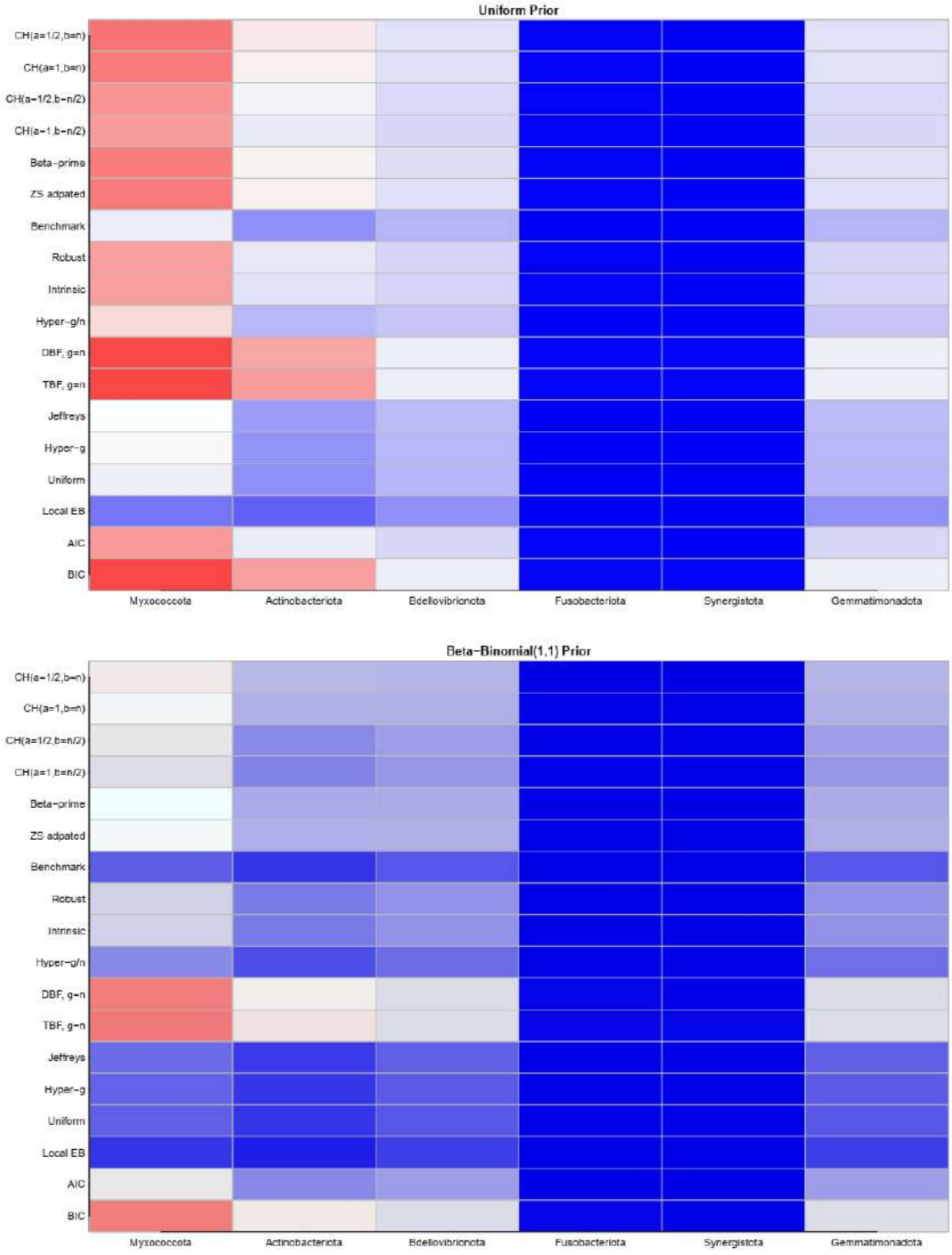


Figure E-A-4 –Marginal posterior inclusion probabilities for the **Phylum** for **Biopsy** with **Rank Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

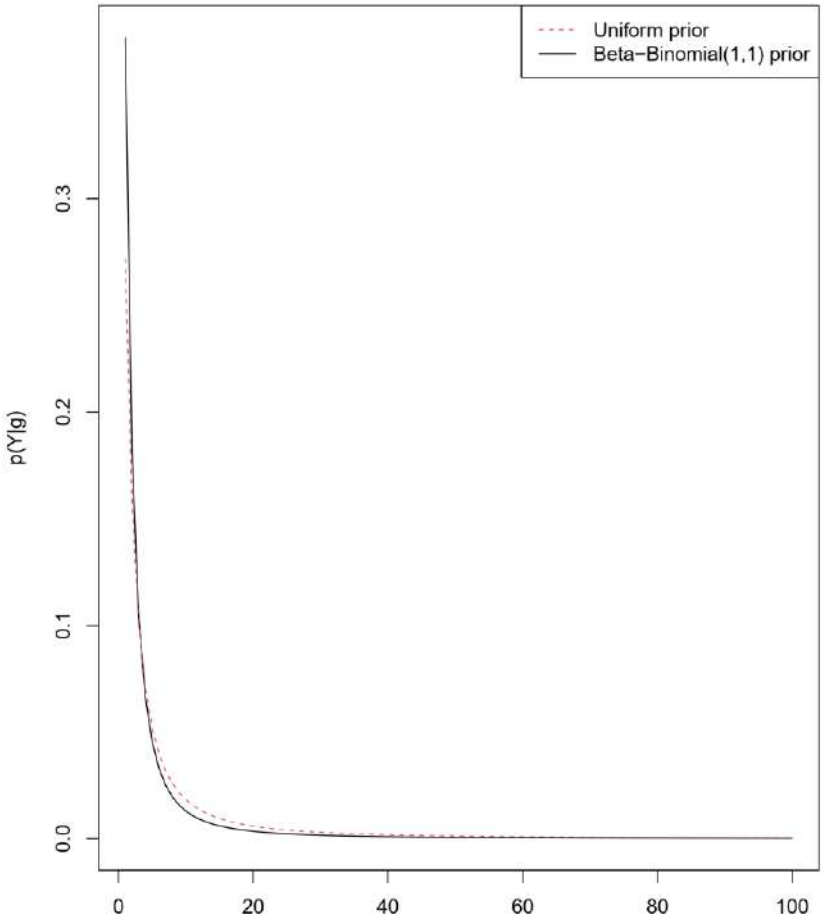


Figure E-A-5: Marginal likelihood of *g* for the **Phylum** for **Biopsy** with **Rank Unit** Normalization method.

Table E-A-4: Prediction accuracy for the **Biopsy (Phylum) Rank Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.8517	0.8516	**4.8273**	13.3200	0.4058	0.4079	0.1102	0.1112
02_CH(a=1,b=n)	0.8522	0.8521	5.4405	6.5281	0.4036	0.4064	0.1097	0.1108
03_CH(a=1/2,b=n/2)	0.8530	0.8516	5.1625	6.5294	0.4021	0.4077	0.1093	0.1104
04_CH(a=1,b=n/2)	0.8527	0.8518	5.1401	11.8442	0.4023	0.4087	0.1091	0.1103
05_Beta-prime	0.8519	0.8517	6.1868	6.5376	0.4042	0.4081	0.1099	0.1111
06_ZS adpated	0.8520	0.8519	5.4783	6.5264	0.4039	0.4069	0.1098	0.1108
07_Robust	0.8540	0.8521	6.4716	11.6332	0.4018	0.4141	0.1089	0.1119
08_Intrinsic	0.8380	0.8364	10.0776	9.3368	**0.4001**	0.4104	0.1084	0.1110
09_Hyper-g/n	0.8381	0.8378	9.4991	18.8122	0.4070	0.4234	**0.1082**	0.1109
10_Benchmark	0.8569	0.8579	30.0272	26.3279	0.4194	0.4218	0.1086	0.1109
11_DBF, g=n	0.8503	0.8505	7.9846	**5.8153**	0.4077	0.4100	0.1109	0.1119
12_TBF, g=n	0.8484	0.8476	7.8468	8.7645	0.4108	0.4152	0.1116	0.1130
13_Jeffreys	0.8580	0.8559	13.3087	25.6798	0.4180	0.4271	0.1087	0.1111
14_Hyper-g	0.8570	0.8579	14.7742	19.9019	0.4219	0.4287	0.1088	0.1113
15_Uniform	**0.8586**	**0.8582**	25.4878	22.0235	0.4202	0.4265	0.1088	0.1113
16_Local EB	0.6682	0.6383	1912.0919	2135.2050	0.7311	0.7642	0.2432	0.2508
17_AIC	0.8511	0.8509	5.1377	7.1992	0.4042	**0.4050**	0.1097	**0.1100**
18_BIC	0.8504	0.8492	6.7248	8.0678	0.4068	0.4105	0.1107	0.1115

B-Class

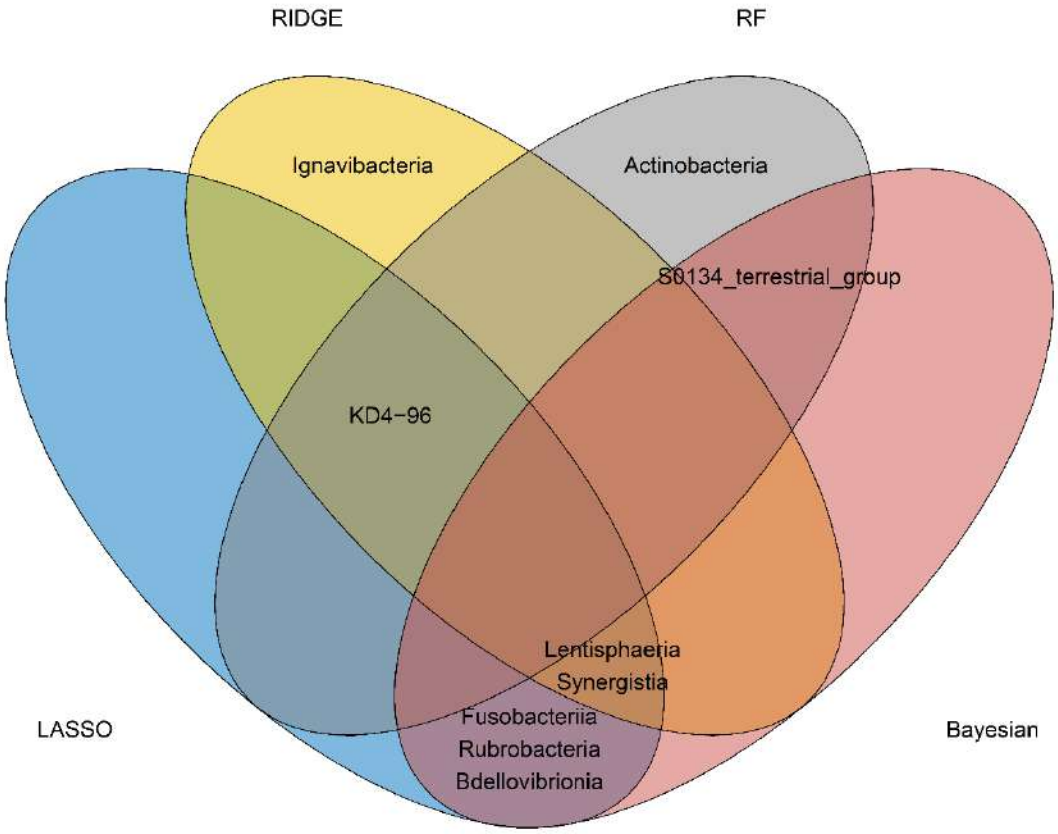


Figure E-B-1 – The Concordance Plot for **Class** for **Biopsy**.

Table E-B-1: The Selected method

Row	Method	Filter	Normal	AUC
1	**LASSO**	**PR**	**LU**	**0.93**
2	RIDGE	AB	LU	0.87
3	RF	PA	RU	0.82
4	**Bayesian**	**PR**	**RU**	**0.93**

Table E-B-2: The common feature names

Common Features for BAS			
KD4−96	S0134_terrestrial_group	Lentisphaeria	Synergistia
Fusobacteriia	Rubrobacteria	Bdellovibrionia	

BAS

Log Unit

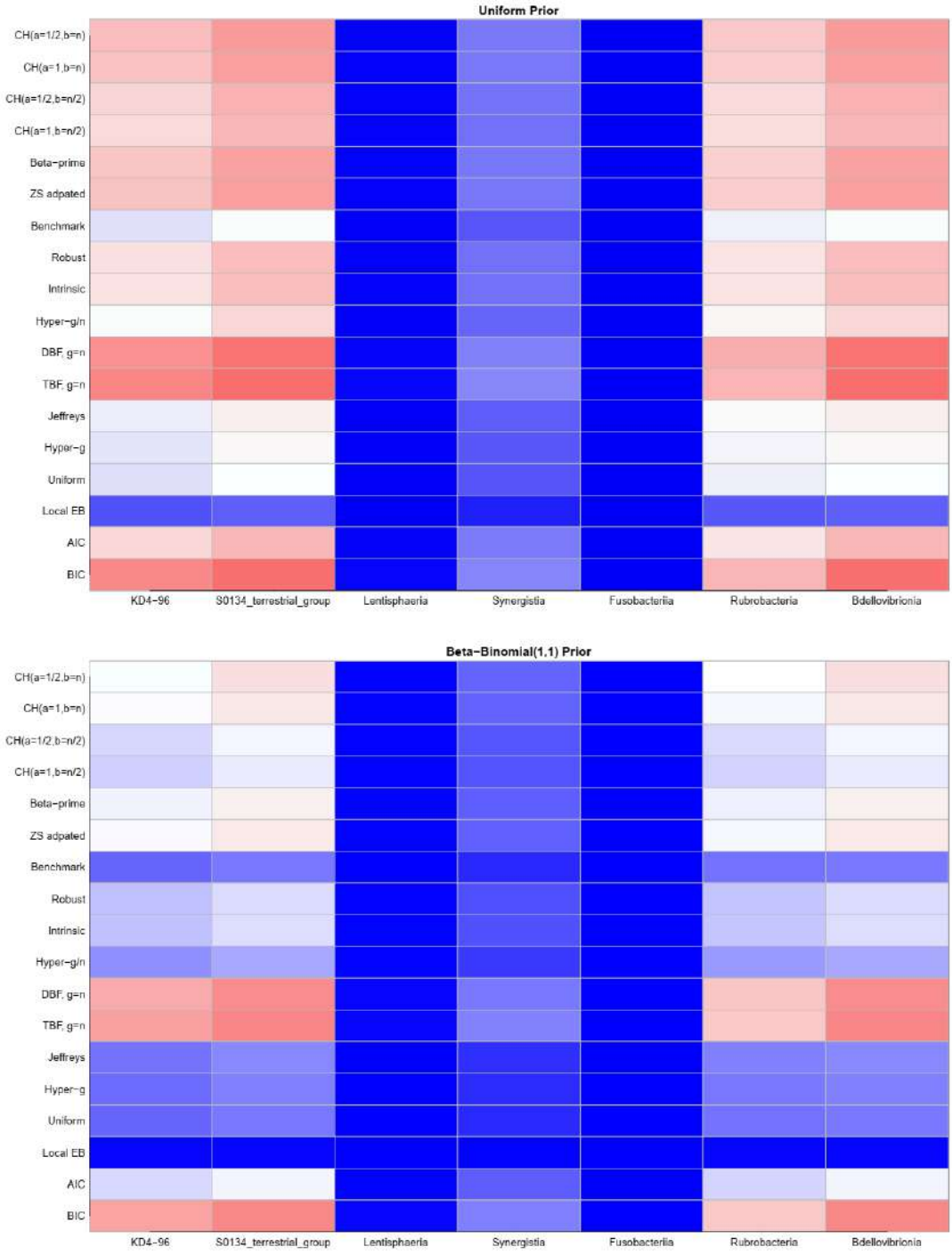


Figure E-B-2 –Marginal posterior inclusion probabilities for the **Class** for **Biopsy** with **Log Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

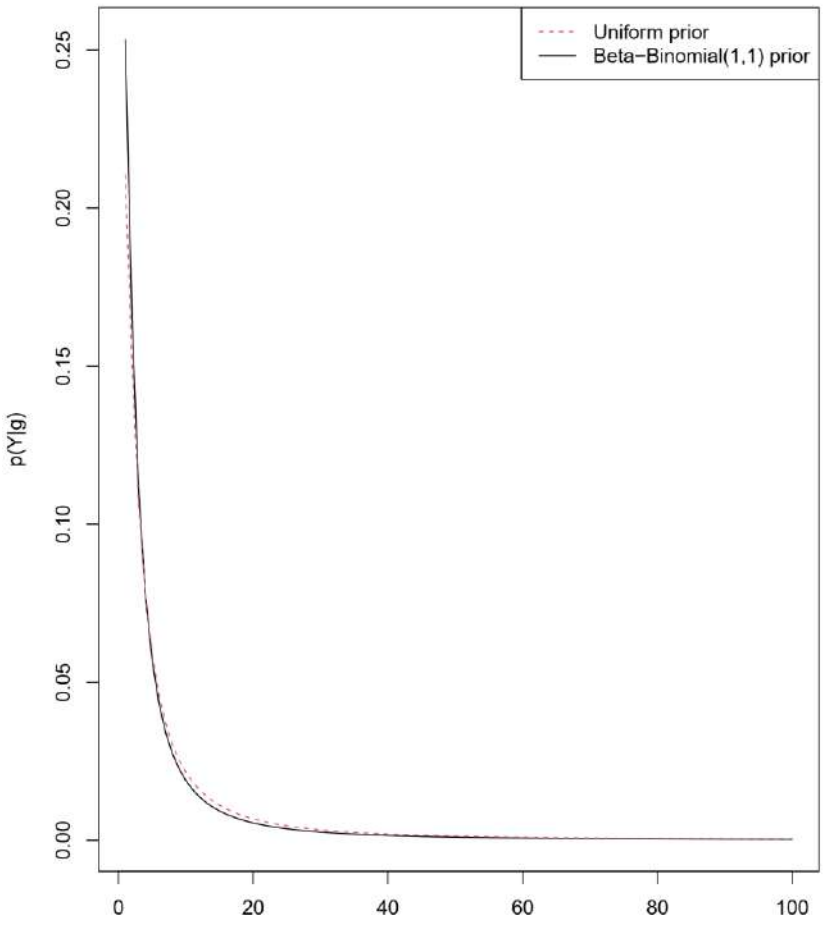


Figure E-B-3: Marginal likelihood of *g* for the **Class** for **Biopsy** with **Log Unit** Normalization method.

Table E-B-3: Prediction accuracy for the **Biopsy (Phylum) Log Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.9118	0.9112	7.6066	10.0749	0.3188	0.3242	0.0847	0.0864
02_CH(a=1,b=n)	0.9116	0.9111	6.9274	14.5335	0.3176	0.3208	0.0845	0.0856
03_CH(a=1/2,b=n/2)	0.9126	0.9115	13.3167	9.0905	0.3169	0.3214	0.0840	0.0849
04_CH(a=1,b=n/2)	0.9131	0.9115	9.4038	8.8892	0.3168	0.3208	0.0838	0.0844
05_Beta-prime	0.9117	0.9122	6.9556	13.1663	0.3179	0.3205	0.0845	0.0859
06_ZS adpated	0.9117	0.9112	6.9290	11.4863	0.3175	0.3208	0.0845	0.0856
07_Robust	0.6165	0.6159	14.1882	11.2544	**0.2657**	**0.2714**	**0.0686**	**0.0709**
08_Intrinsic	0.9114	0.9122	11.5040	8.3028	0.3175	0.3163	0.0831	0.0830
09_Hyper-g/n	0.9146	**0.9136**	12.7182	12.7185	0.3132	0.3175	0.0827	0.0827
10_Benchmark	0.9094	0.9051	19.8779	29.8104	0.3398	0.3550	0.0866	0.0914
11_DBF, g=n	0.9118	0.9125	13.0436	8.4099	0.3234	0.3254	0.0861	0.0878
12_TBF, g=n	0.9106	0.9129	7.7269	11.1605	0.3288	0.3313	0.0877	0.0896
13_Jeffreys	**0.9153**	0.9126	16.7177	16.2126	0.3171	0.3337	0.0824	0.0844
14_Hyper-g	0.9145	0.9096	18.9916	22.9011	0.3274	0.3370	0.0834	0.0859
15_Uniform	0.9128	0.9083	17.7322	21.1141	0.3291	0.3398	0.0846	0.0875
16_Local EB	0.6651	0.6467	2432.8195	3055.5735	0.7347	0.7769	0.2216	0.2273
17_AIC	0.9102	0.9092	**6.6980**	**6.4754**	0.3201	0.3213	0.0845	0.0851
18_BIC	0.9103	0.9093	8.8610	8.3023	0.3257	0.3288	0.0865	0.0877

Rank Unit

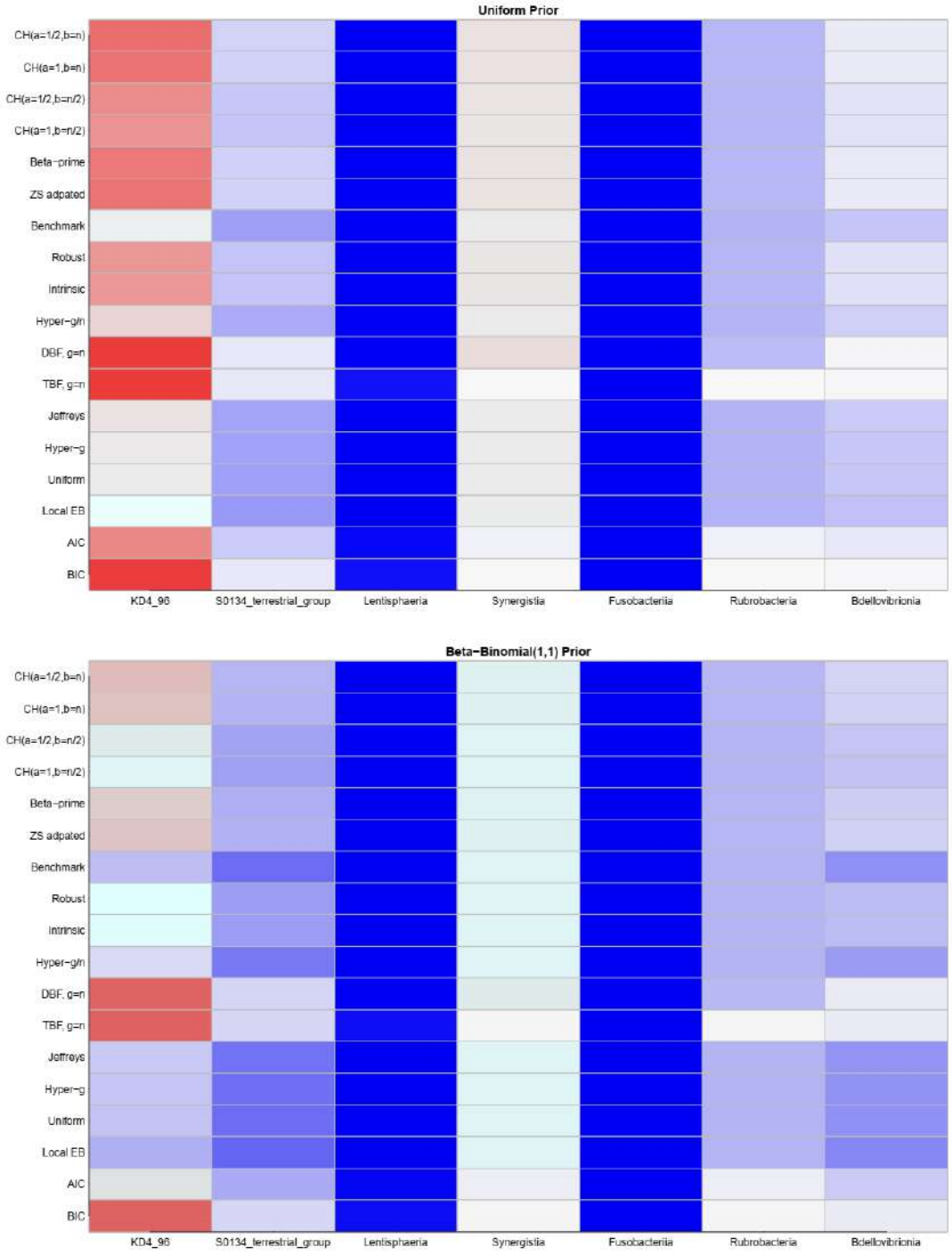


Figure E-B-4 –Marginal posterior inclusion probabilities for the **Class** for **Biopsy** with **Rank Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

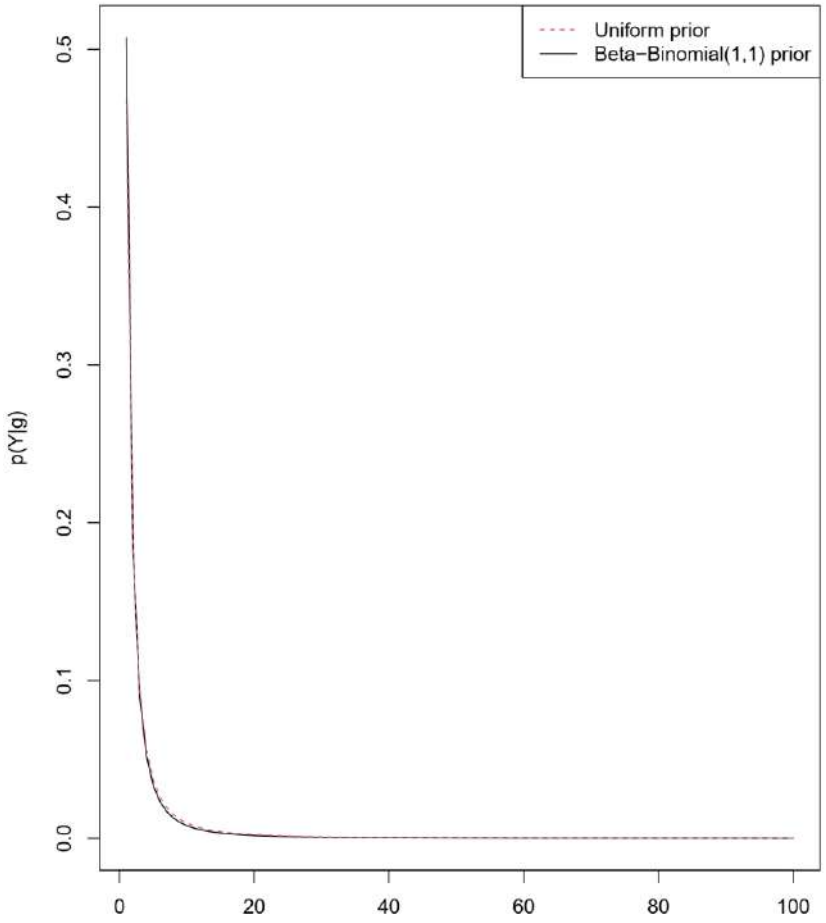


Figure E-B-5: Marginal likelihood of *g* for the **Class** for **Biopsy** with **Rank Unit** Normalization method.

Table E-B-4: Prediction accuracy for the **Biopsy (Phylum) Rank Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.8967	0.8999	4.8763	5.9214	0.2896	0.2785	0.0736	0.0704
02_CH(a=1,b=n)	0.8984	0.9004	7.1117	4.8319	0.2861	0.2776	0.0729	0.0695
03_CH(a=1/2,b=n/2)	0.9003	0.9012	6.6229	5.5943	0.2854	0.2723	0.0710	0.0665
04_CH(a=1,b=n/2)	0.9008	0.9007	12.0676	5.2131	0.2821	0.2735	0.0705	0.0664
05_Beta-prime	0.8988	0.9000	6.8295	**4.3042**	0.2857	0.2792	0.0726	0.0690
06_ZS adpated	0.8984	0.9004	7.1200	6.0890	0.2857	0.2778	0.0729	0.0694
07_Robust	0.9019	0.9041	10.2926	6.2899	0.2727	0.2796	0.0670	0.0669
08_Intrinsic	0.8743	0.8754	5.6000	6.3751	0.2720	0.2728	0.0664	0.0655
09_Hyper-g/n	0.8651	0.8728	8.6206	8.6965	0.2904	0.2841	0.0658	0.0634
10_Benchmark	0.9004	0.8995	11.9185	49.3674	0.2667	0.2744	**0.0619**	0.0662
11_DBF, g=n	0.8924	0.8934	**4.1981**	4.5335	0.2984	0.2946	0.0751	0.0744
12_TBF, g=n	0.8868	0.8877	12.8911	7.0968	0.3079	0.3072	0.0783	0.0778
13_Jeffreys	0.8973	**0.9085**	10.9810	25.8366	0.2738	0.2660	0.0640	0.0625
14_Hyper-g	0.8995	0.9083	10.6283	28.8502	0.2692	0.2594	0.0628	0.0617
15_Uniform	**0.9047**	0.9073	12.8186	30.4937	**0.2620**	**0.2554**	0.0619	**0.0616**
16_Local EB	0.5000	0.5000	-	-	0.9539	1.0743	0.2615	0.2702
17_AIC	0.8916	0.8907	4.2368	5.5101	0.3003	0.2979	0.0767	0.0754
18_BIC	0.8900	0.8905	6.3531	4.5282	0.3037	0.3020	0.0779	0.0770

C-Order

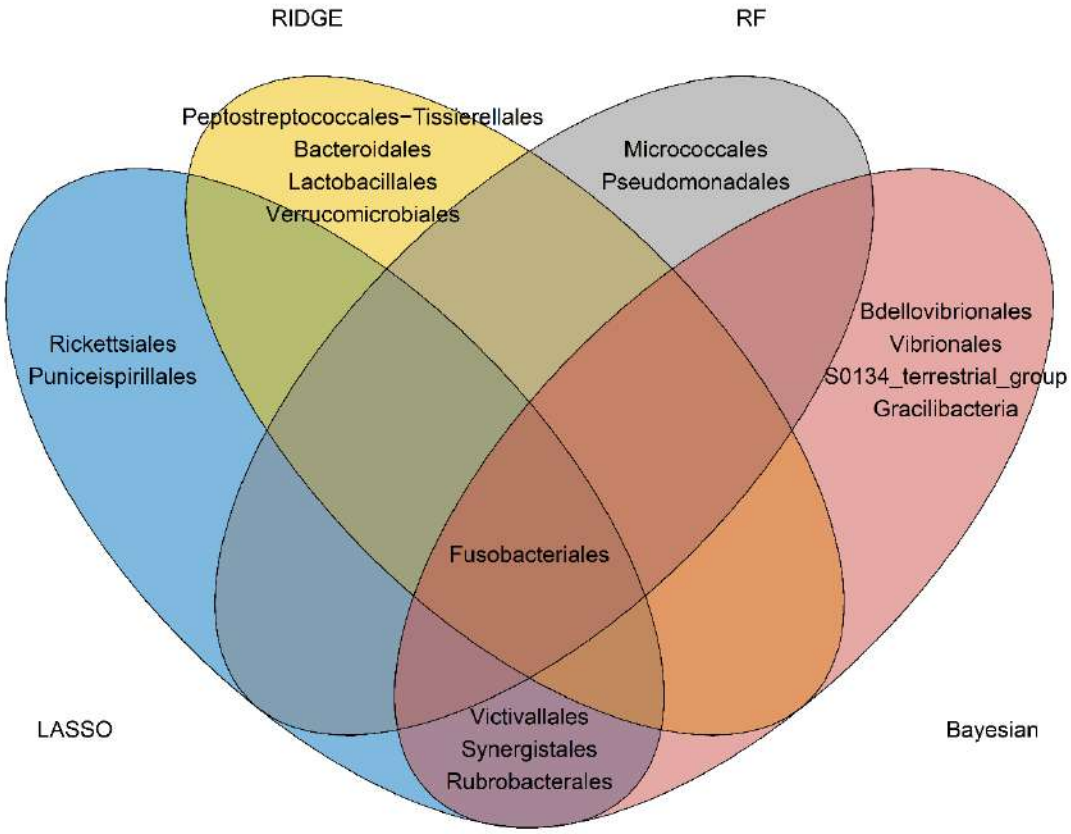


Figure E-C-1 – The Concordance Plot for **Order** for **Biopsy**.

Table E-C-1: The Selected method				
Row	**Method**	**Filter**	**Normal**	**AUC**
1	LASSO	CA	LU	0.82
2	RIDGE	VA	LU	0.68
3	RF	CA	RU	0.79
4	**Bayesian**	**PR**	**RU**	**0.83**

Table E-C-2: The common feature names			
Common Features for BAS			
Fusobacteriales	Victivallales	Synergistales	Rubrobacterales

BAS

Log Unit

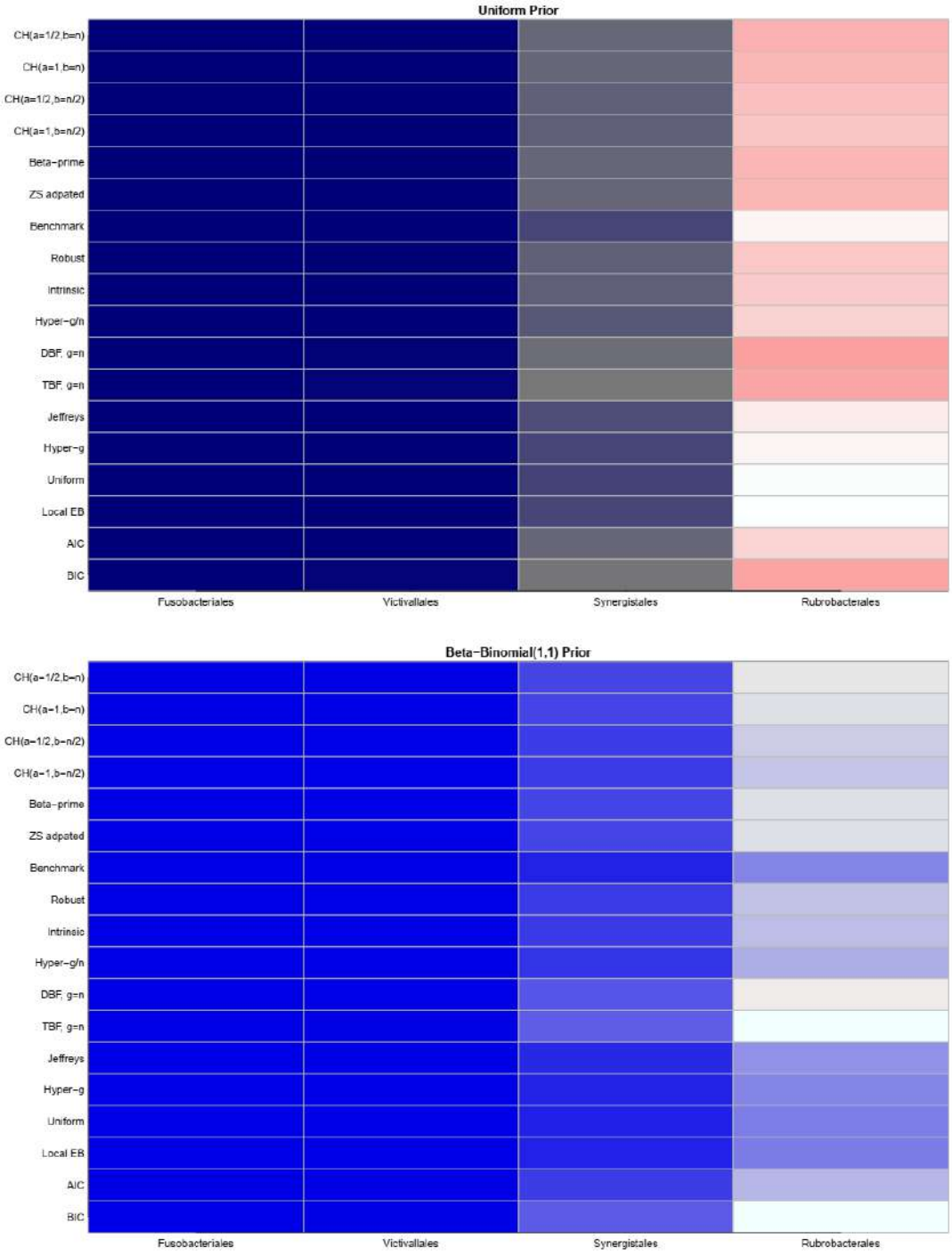


Figure E-C-2 –Marginal posterior inclusion probabilities for the **Order** for **Biopsy** with **Log Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

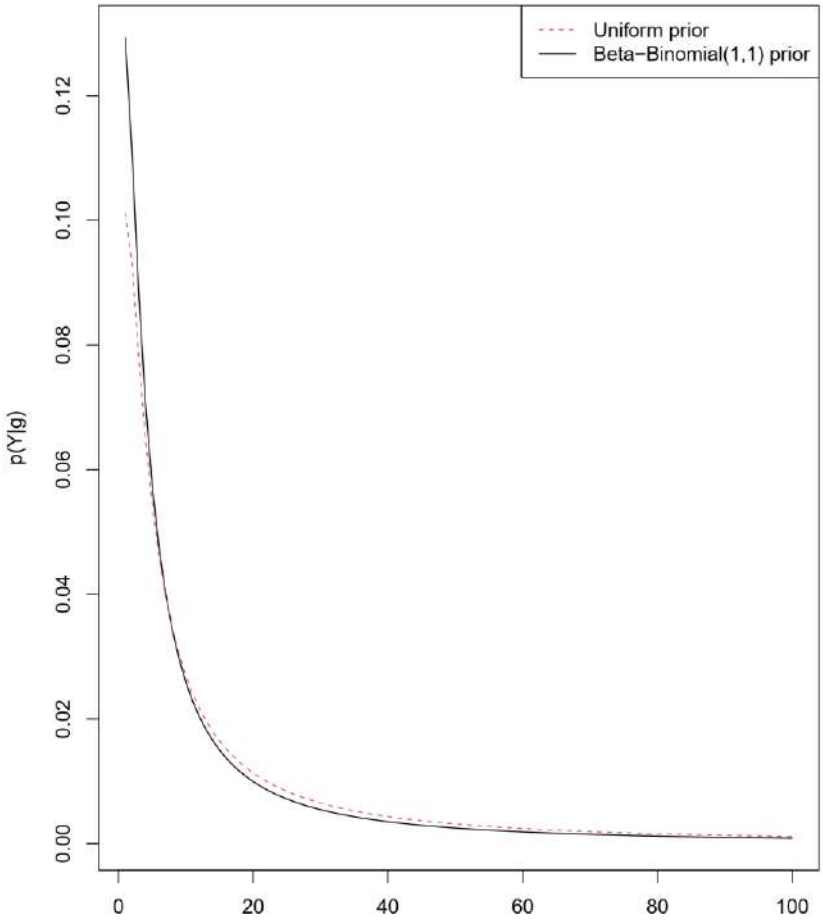


Figure E-C-3: Marginal likelihood of *g* for the **Order** for **Biopsy** with **Log Unit** Normalization method.

Table E-C-3: Prediction accuracy for the **Biopsy (Phylum) Log Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.9243	0.9237	10.3494	11.3926	0.2776	0.2799	0.0741	0.0739
02_CH(a=1,b=n)	0.9256	0.9239	14.0591	11.2729	0.2775	0.2782	0.0738	0.0734
03_CH(a=1/2,b=n/2)	0.9257	0.9269	10.8280	20.8545	0.2740	0.2751	0.0733	0.0730
04_CH(a=1,b=n/2)	0.9254	0.9277	10.4549	13.0273	0.2756	0.2751	0.0731	0.0727
05_Beta-prime	0.9239	0.9238	10.1573	12.3933	0.2801	0.2795	0.0740	0.0737
06_ZS adpated	0.9256	0.9239	14.0915	11.3055	0.2774	0.2779	0.0738	0.0734
07_Robust	0.9239	**0.9290**	9.6129	14.0082	0.2758	**0.2717**	0.0728	0.0727
08_Intrinsic	0.9256	0.9278	12.0275	13.0272	0.2740	0.2722	0.0726	0.0729
09_Hyper-g/n	0.9277	0.9286	15.7991	21.4115	0.2754	0.2789	**0.0725**	**0.0725**
10_Benchmark	0.9277	0.9275	20.3593	28.5359	0.2943	0.3119	0.0755	0.0789
11_DBF, g=n	0.9243	0.9242	12.5259	10.5440	0.2826	0.2816	0.0746	0.0743
12_TBF, g=n	0.9258	0.9233	13.6843	11.4318	0.2790	0.2805	0.0749	0.0750
13_Jeffreys	0.9291	0.9272	16.5459	17.6885	0.2812	0.2920	0.0727	0.0747
14_Hyper-g	0.9283	0.9265	18.9085	22.3340	0.2891	0.3024	0.0740	0.0763
15_Uniform	**0.9298**	0.9286	34.3582	27.4792	0.2945	0.3050	0.0757	0.0782
16_Local EB	0.7503	0.7326	922.2362	1327.4840	0.6068	0.6498	0.1949	0.2088
17_AIC	0.9254	0.9257	**9.2700**	11.4278	**0.2739**	0.2760	0.0734	0.0733
18_BIC	0.9255	0.9256	10.3629	**9.8421**	0.2765	0.2767	0.0741	0.0735

Rank Unit

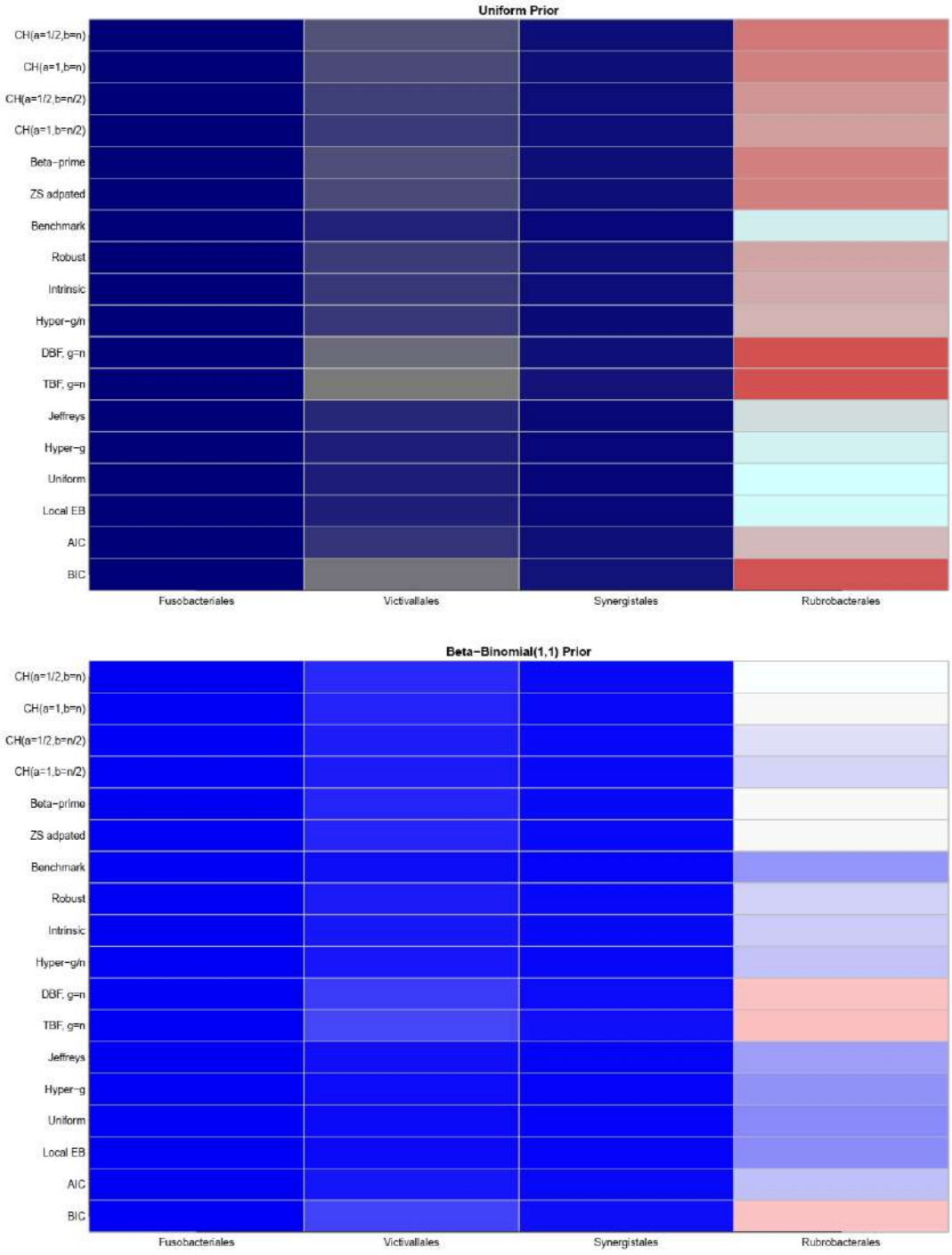


Figure E-C-4 –Marginal posterior inclusion probabilities for the **Order** for **Biopsy** with **Rank Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

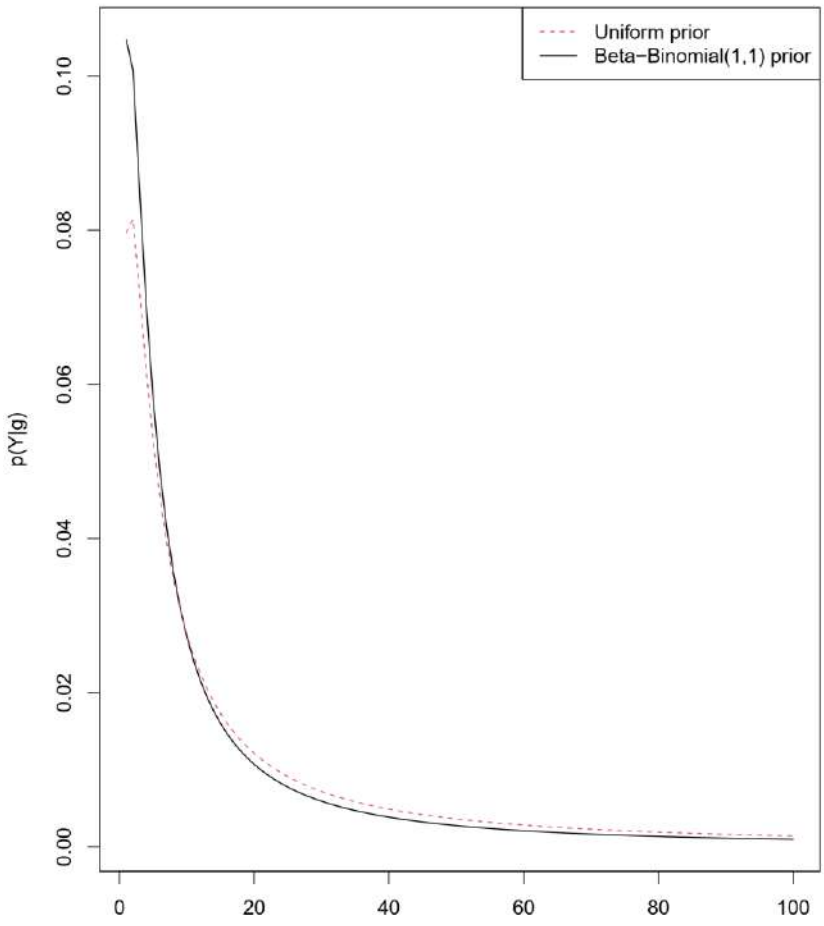


Figure E-C-5: Marginal likelihood of *g* for the **Order** for **Biopsy** with **Rank Unit** Normalization method.

Table E-C-4: Prediction accuracy for the **Biopsy (Phylum) Rank Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.8821	0.8845	14.6700	12.4028	0.3618	0.3606	0.0980	0.0976
02_CH(a=1,b=n)	0.8823	0.8840	10.3203	16.4274	0.3586	0.3546	0.0971	0.0957
03_CH(a=1/2,b=n/2)	0.8829	0.8840	10.7877	13.9581	0.3565	0.3524	0.0964	0.0949
04_CH(a=1,b=n/2)	0.8820	0.8856	23.0055	9.4312	0.3583	0.3494	0.0960	0.0941
05_Beta-prime	0.8824	0.8845	11.6306	10.9669	0.3604	0.3588	0.0977	0.0971
06_ZS adpated	0.8823	0.8839	10.5132	10.5520	0.3587	0.3548	0.0972	0.0958
07_Robust	0.8830	0.8843	7.1817	**8.1840**	0.3550	0.3523	0.0946	0.0937
08_Intrinsic	0.8743	0.8747	7.3477	31.1780	0.3532	0.3522	**0.0938**	0.0933
09_Hyper-g/n	0.8735	0.8762	10.0020	16.5148	**0.3491**	0.3503	0.0949	0.0941
10_Benchmark	0.8835	0.8835	19.7791	59.5459	0.3632	0.3698	0.0962	0.0976
11_DBF, g=n	0.8796	0.8819	9.4355	12.9915	0.3651	0.3631	0.0991	0.0986
12_TBF, g=n	0.8788	0.8802	9.5170	17.3996	0.3654	0.3675	0.0999	0.1001
13_Jeffreys	0.8867	**0.8881**	22.1567	47.4999	0.3516	0.3516	0.0941	0.0934
14_Hyper-g	**0.8877**	**0.8881**	38.4053	41.3551	0.3542	0.3581	0.0944	0.0944
15_Uniform	0.8861	0.8877	29.1776	19.7823	0.3599	0.3610	0.0953	0.0955
16_Local EB	0.6785	0.6690	884.7091	1292.3300	0.7101	0.7391	0.2182	0.2241
17_AIC	0.8815	0.8854	**6.3212**	16.7508	0.3572	**0.3448**	0.0959	**0.0928**
18_BIC	0.8797	0.8813	14.2278	8.5154	0.3620	0.3589	0.0978	0.0971

D-Family

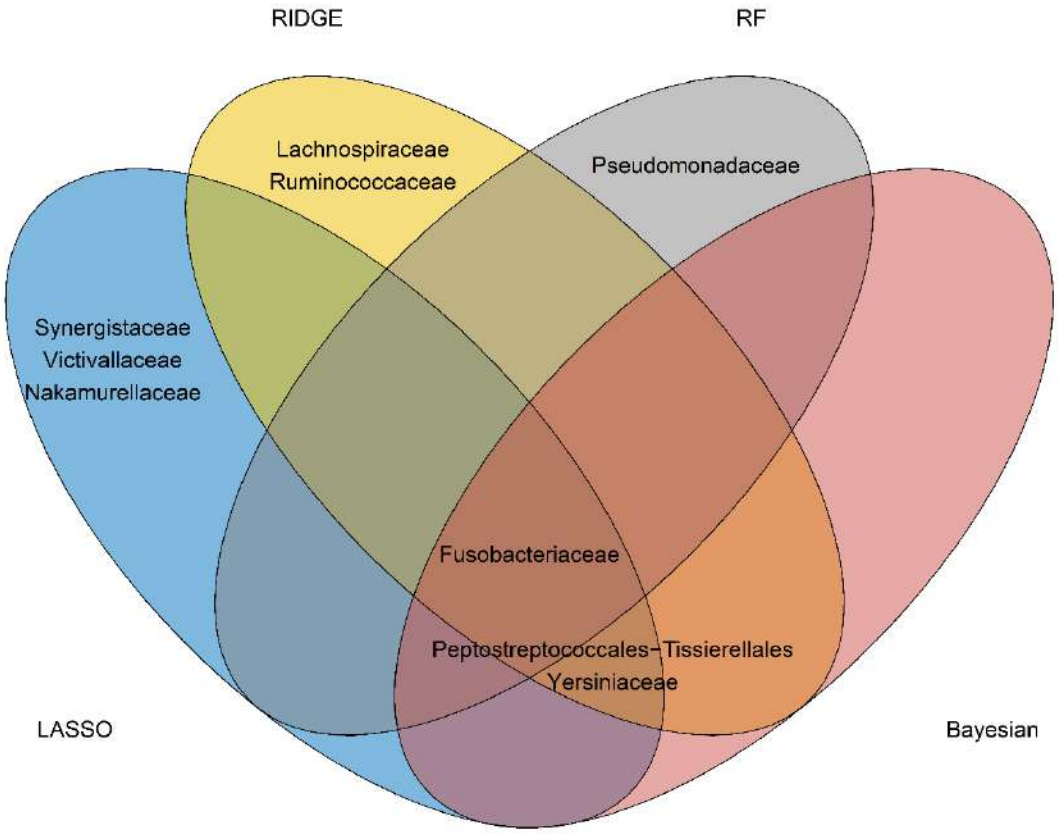


Figure E-D-1 – The Concordance Plot for **Family** for **Biopsy**.

Table E-D-1: The Selected method				
Row	**Method**	**Filter**	**Normal**	**AUC**
1	**LASSO**	**AB**	**LU**	**0.83**
2	RIDGE	VA	LU	0.76
3	RF	PA	RU	0.77
4	Bayesian	VA	LU	0.82

Table E-D-2: The common feature names		
Common Features for BAS		
Fusobacteriaceae	Yersiniaceae	Peptostreptococcales−Tissierellales

BAS

Log Unit

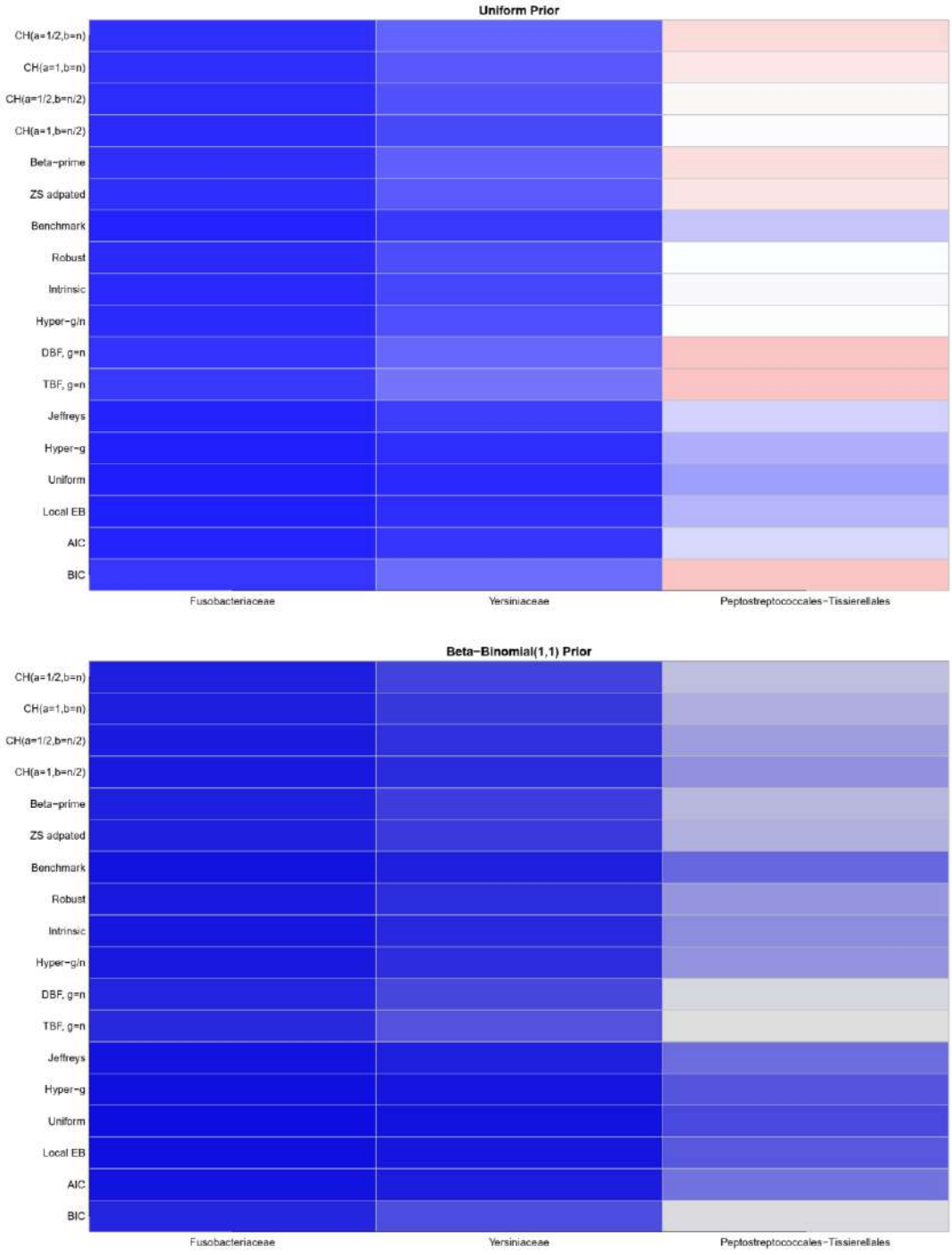


Figure E-D-2 –Marginal posterior inclusion probabilities for the **Family** for **Biopsy** with **Log Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

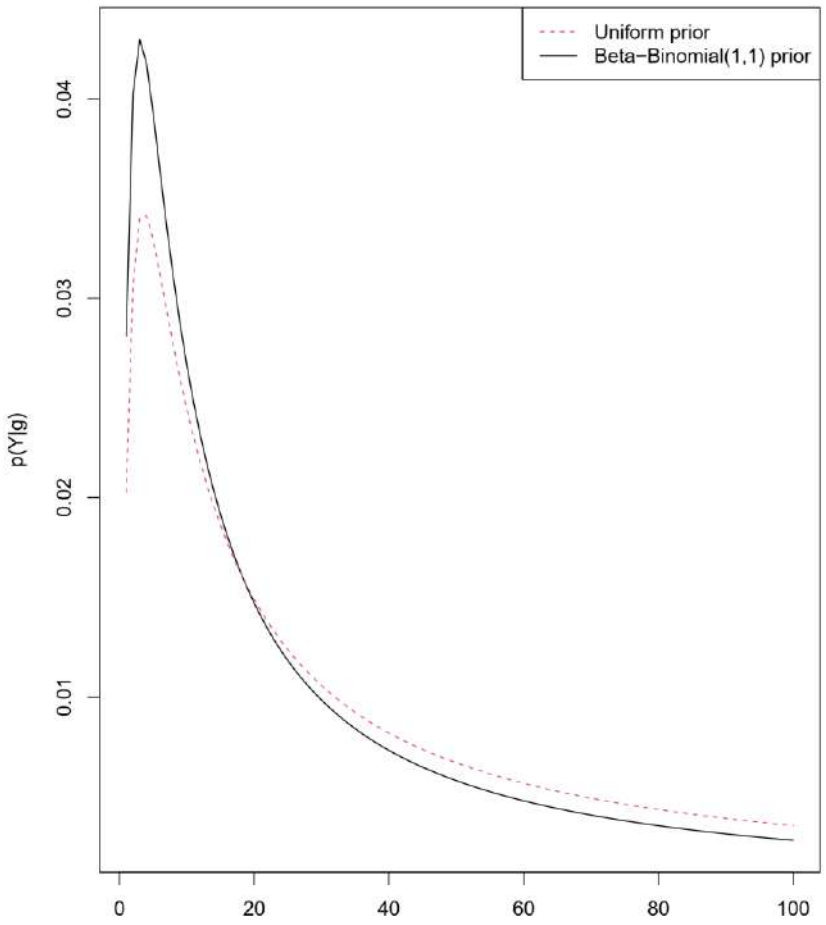


Figure E-D-3: Marginal likelihood of *g* for the **Family** for **Biopsy** with **Log Unit** Normalization method.

Table E-D-3: Prediction accuracy for the **Biopsy (Phylum) Log Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.8328	0.8326	32.9746	35.2933	0.4703	0.4751	0.1361	0.1370
02_CH(a=1,b=n)	0.8333	0.8315	26.7441	10.5182	0.4703	0.4731	0.1355	0.1359
03_CH(a=1/2,b=n/2)	0.8336	0.8315	30.5098	10.0319	0.4693	0.4717	0.1353	0.1357
04_CH(a=1,b=n/2)	0.8335	0.8318	45.6020	10.3940	0.4687	0.4700	0.1348	0.1347
05_Beta-prime	0.8329	0.8321	32.2846	34.7320	0.4699	0.4751	0.1360	0.1368
06_ZS adpated	0.8332	0.8316	30.6133	10.5174	0.4703	0.4730	0.1355	0.1359
07_Robust	0.8306	0.8326	56.0202	7.7983	0.4740	0.4705	0.1348	0.1344
08_Intrinsic	0.8308	0.8322	31.4120	7.4182	0.4736	0.4679	0.1345	0.1339
09_Hyper-g/n	0.8327	0.8334	34.9857	16.4551	0.4664	0.4637	0.1347	0.1345
10_Benchmark	0.8393	0.8369	39.3359	18.8415	0.4651	0.4684	0.1358	0.1372
11_DBF, g=n	0.8325	0.8323	31.9498	58.1955	0.4709	0.4744	0.1357	0.1363
12_TBF, g=n	0.8333	0.8337	23.7260	35.2933	0.4694	0.4707	0.1356	0.1362
13_Jeffreys	0.8346	0.8341	38.8890	9.0660	**0.4595**	**0.4557**	**0.1330**	**0.1317**
14_Hyper-g	0.8383	0.8366	**13.2931**	13.8771	0.4599	0.4570	0.1330	0.1325
15_Uniform	**0.8401**	**0.8375**	14.5180	23.4099	0.4596	0.4607	0.1336	0.1334
16_Local EB	0.7773	0.7683	287.1208	237.3220	0.5941	0.6076	0.1869	0.1888
17_AIC	0.8327	0.8306	48.7311	**6.1724**	0.4680	0.4700	0.1341	0.1339
18_BIC	0.8322	0.8314	21.3743	9.1465	0.4702	0.4734	0.1350	0.1352

Rank Unit

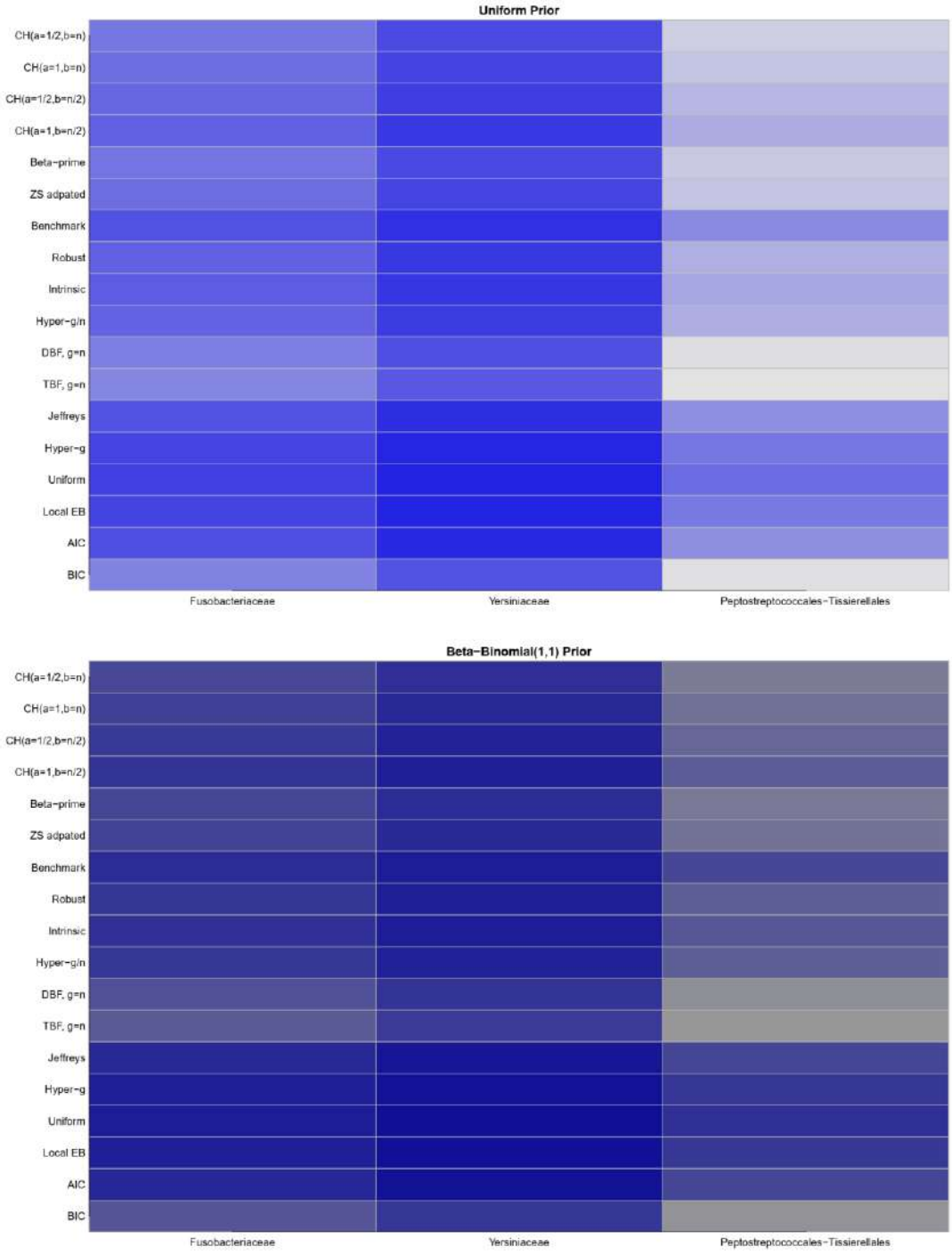


Figure E-D-4 –Marginal posterior inclusion probabilities for the **Family** for **Biopsy** with **Rank Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

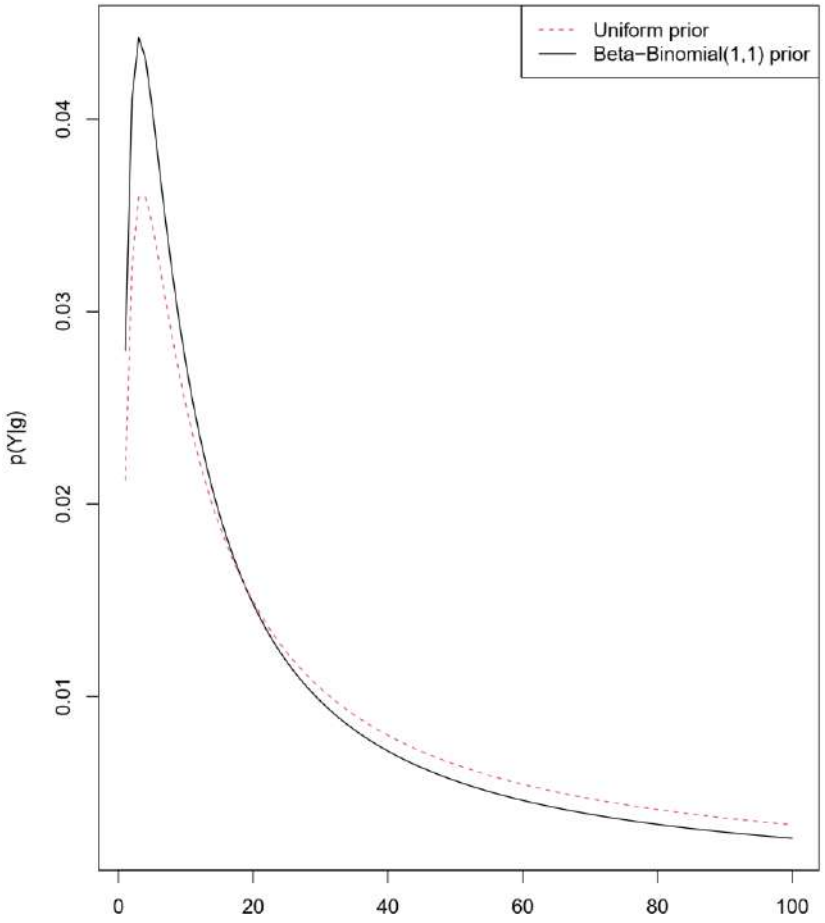


Figure E-D-5: Marginal likelihood of *g* for the **Family** for **Biopsy** with **Rank Unit** Normalization method.

Table E-D-4: Prediction accuracy for the **Biopsy (Phylum) Rank Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.8348	0.8339	2.4918	6.3846	0.4816	0.4875	0.1413	0.1432
02_CH(a=1,b=n)	0.8356	0.8340	2.2959	4.5674	0.4791	0.4826	0.1404	0.1417
03_CH(a=1/2,b=n/2)	0.8358	0.8346	2.4392	7.0031	0.4788	0.4819	0.1404	0.1416
04_CH(a=1,b=n/2)	0.8350	0.8347	9.4908	4.5314	0.4760	0.4792	0.1397	0.1405
05_Beta-prime	0.8349	0.8340	2.4718	5.6854	0.4820	0.4868	0.1412	0.1430
06_ZS adpated	0.8354	0.8343	2.3049	4.8950	0.4793	0.4827	0.1405	0.1417
07_Robust	0.8350	0.8334	4.5319	3.7488	0.4763	0.4803	0.1396	0.1404
08_Intrinsic	0.8356	0.8327	3.4415	3.4685	0.4736	0.4784	0.1392	0.1399
09_Hyper-g/n	0.8353	0.8350	2.5028	5.3438	0.4754	0.4776	0.1397	0.1407
10_Benchmark	0.8395	0.8402	12.7115	8.3263	0.4809	0.4854	0.1424	0.1446
11_DBF, g=n	0.8354	0.8333	2.2550	**2.1997**	0.4803	0.4847	0.1406	0.1420
12_TBF, g=n	0.8342	0.8347	3.5924	2.3439	0.4818	0.4822	0.1407	0.1417
13_Jeffreys	0.8387	0.8365	10.6975	4.0218	**0.4651**	**0.4672**	**0.1377**	**0.1372**
14_Hyper-g	0.8394	0.8395	7.7634	4.7580	0.4720	0.4722	0.1383	0.1383
15_Uniform	**0.8401**	**0.8407**	7.7144	5.3150	0.4728	0.4706	0.1391	0.1388
16_Local EB	0.7662	0.7615	130.6964	165.9737	0.5894	0.5894	0.1818	0.1823
17_AIC	0.8343	0.8323	**1.7267**	3.8620	0.4774	0.4787	0.1393	0.1394
18_BIC	0.8348	0.8340	1.8922	14.8159	0.4786	0.4808	0.1400	0.1406

E-Genus

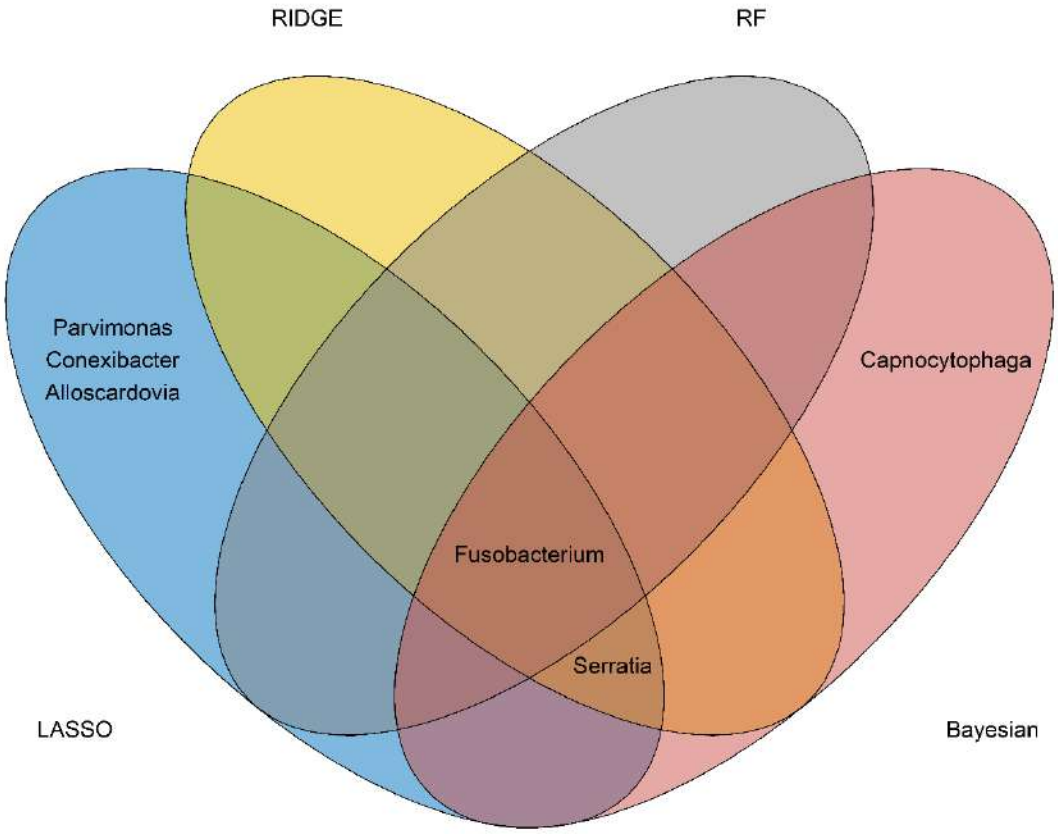


Figure E-E-1 – The Concordance Plot for **Genus** for **Biopsy**.

Table E-E-1: The Selected method

Row	Method	Filter	Normal	AUC
1	LASSO	PA	LU	0.77
2	RIDGE	VA	LU	0.69
3	RF	VA	RU	0.77
4	**Bayesian**	**VA**	**LU**	**0.86**

Table E-E-2: The common feature names

Common Features for BAS	
Fusobacterium	Serratia

BAS

Log Unit

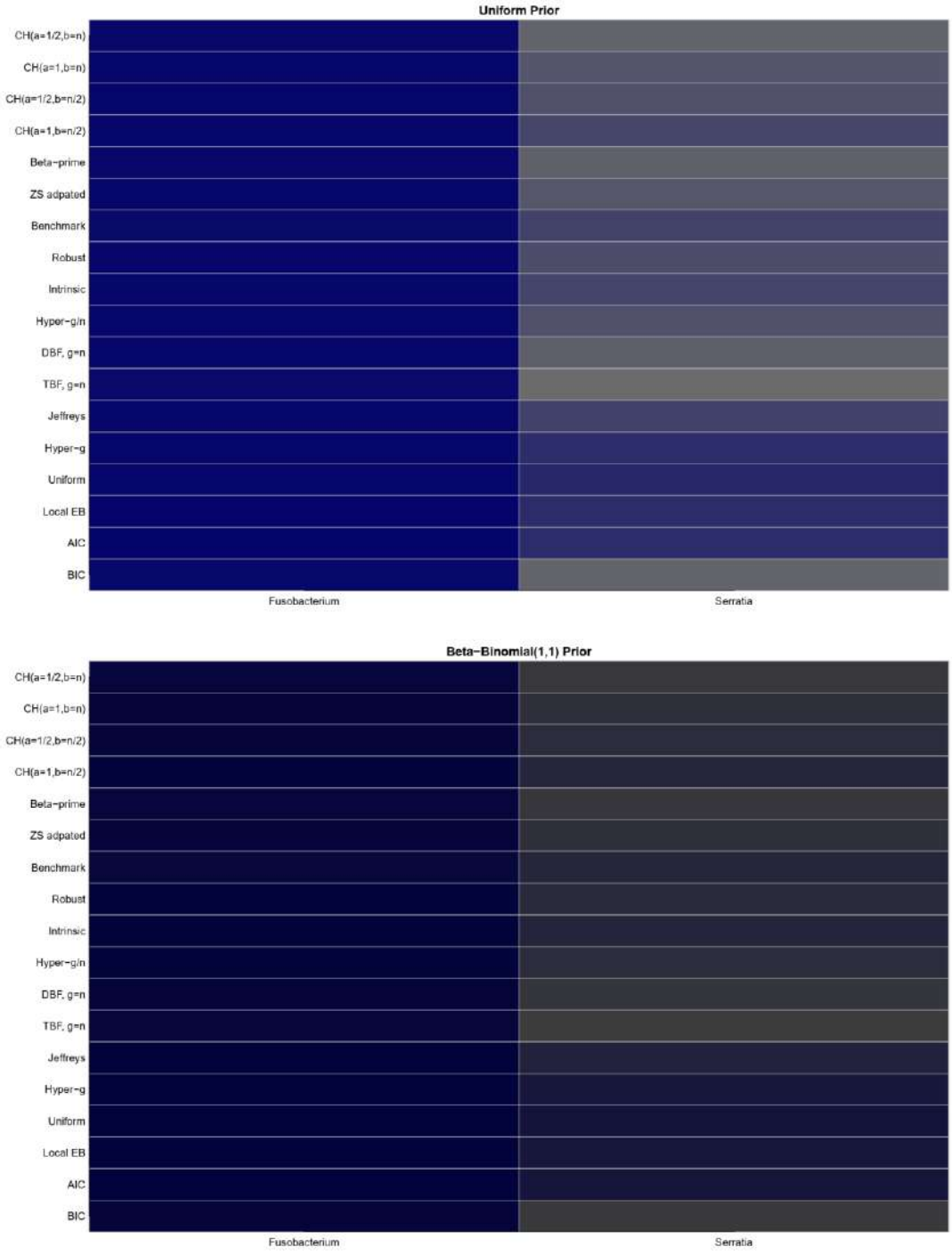


Figure E-E-2 –Marginal posterior inclusion probabilities for the **Genus** for **Biopsy** with **Log Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

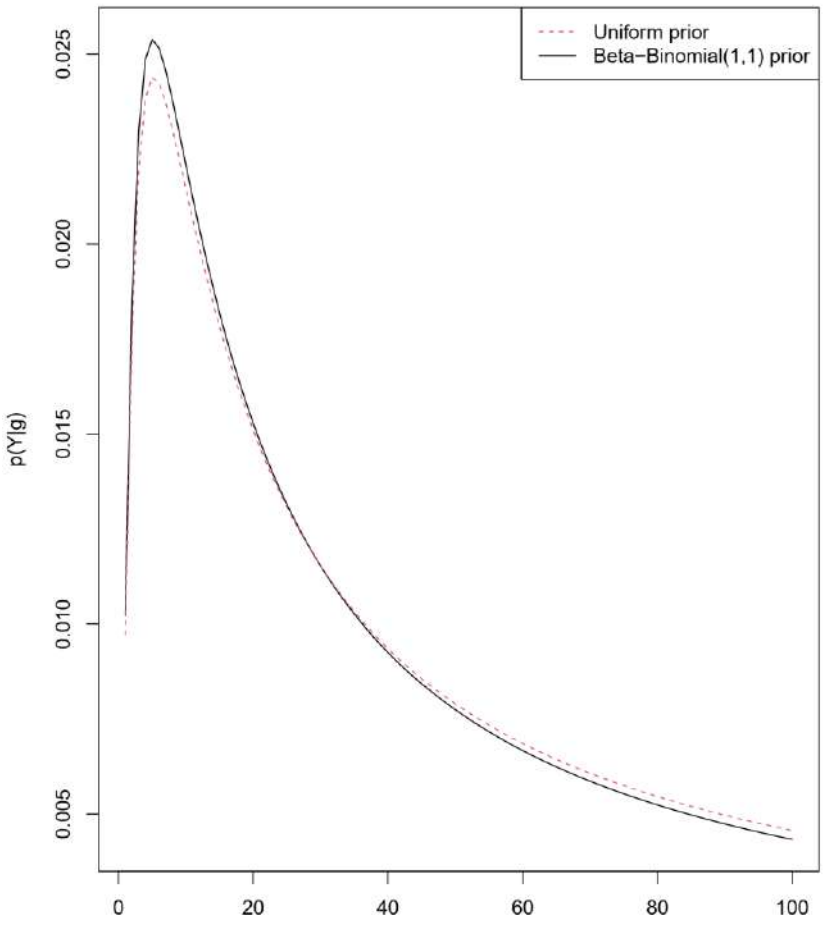


Figure E-E-3: Marginal likelihood of *g* for the **Genus** for **Biopsy** with **Log Unit** Normalization method.

Table E-E-3: Prediction accuracy for the **Biopsy (Phylum) Log Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.8416	**0.8421**	14.8358	35.4932	0.4639	0.4643	0.1391	0.1388
02_CH(a=1,b=n)	0.8415	0.8413	9.9598	9.2878	0.4605	0.4603	0.1377	0.1370
03_CH(a=1/2,b=n/2)	0.8417	0.8416	9.6559	9.2563	0.4602	0.4597	0.1377	0.1371
04_CH(a=1,b=n/2)	0.8414	0.8411	9.3195	7.7024	0.4571	0.4557	0.1365	0.1355
05_Beta-prime	0.8417	0.8419	12.9864	12.6859	0.4635	0.4636	0.1390	0.1386
06_ZS adpated	0.8410	0.8414	10.0972	9.3970	0.4606	0.4605	0.1378	0.1371
07_Robust	0.8415	0.8407	22.7110	7.5998	0.4578	0.4565	0.1363	0.1353
08_Intrinsic	0.8414	0.8401	7.9924	7.1893	0.4553	0.4551	0.1356	0.1346
09_Hyper-g/n	0.8416	0.8415	9.5318	8.6864	0.4562	0.4571	0.1371	0.1363
10_Benchmark	0.8421	0.8413	22.0403	26.2657	0.4802	0.4819	0.1435	0.1439
11_DBF, g=n	0.8411	0.8414	9.3945	8.7528	0.4604	0.4608	0.1377	0.1369
12_TBF, g=n	0.8415	0.8419	9.8450	9.6669	0.4601	0.4600	0.1379	0.1372
13_Jeffreys	0.8421	0.8401	30.1798	7.4994	**0.4472**	**0.4463**	**0.1337**	**0.1324**
14_Hyper-g	**0.8428**	0.8406	9.7876	8.9950	0.4591	0.4589	0.1354	0.1344
15_Uniform	0.8417	0.8407	12.0475	10.4151	0.4685	0.4668	0.1370	0.1359
16_Local EB	0.8111	0.8087	44.8053	49.0847	0.4715	0.4719	0.1433	0.1423
17_AIC	0.8405	0.8408	**6.0714**	**6.0444**	0.4554	0.4515	0.1344	0.1334
18_BIC	0.8412	0.8407	8.5971	6.8597	0.4574	0.4579	0.1363	0.1356

Rank Unit

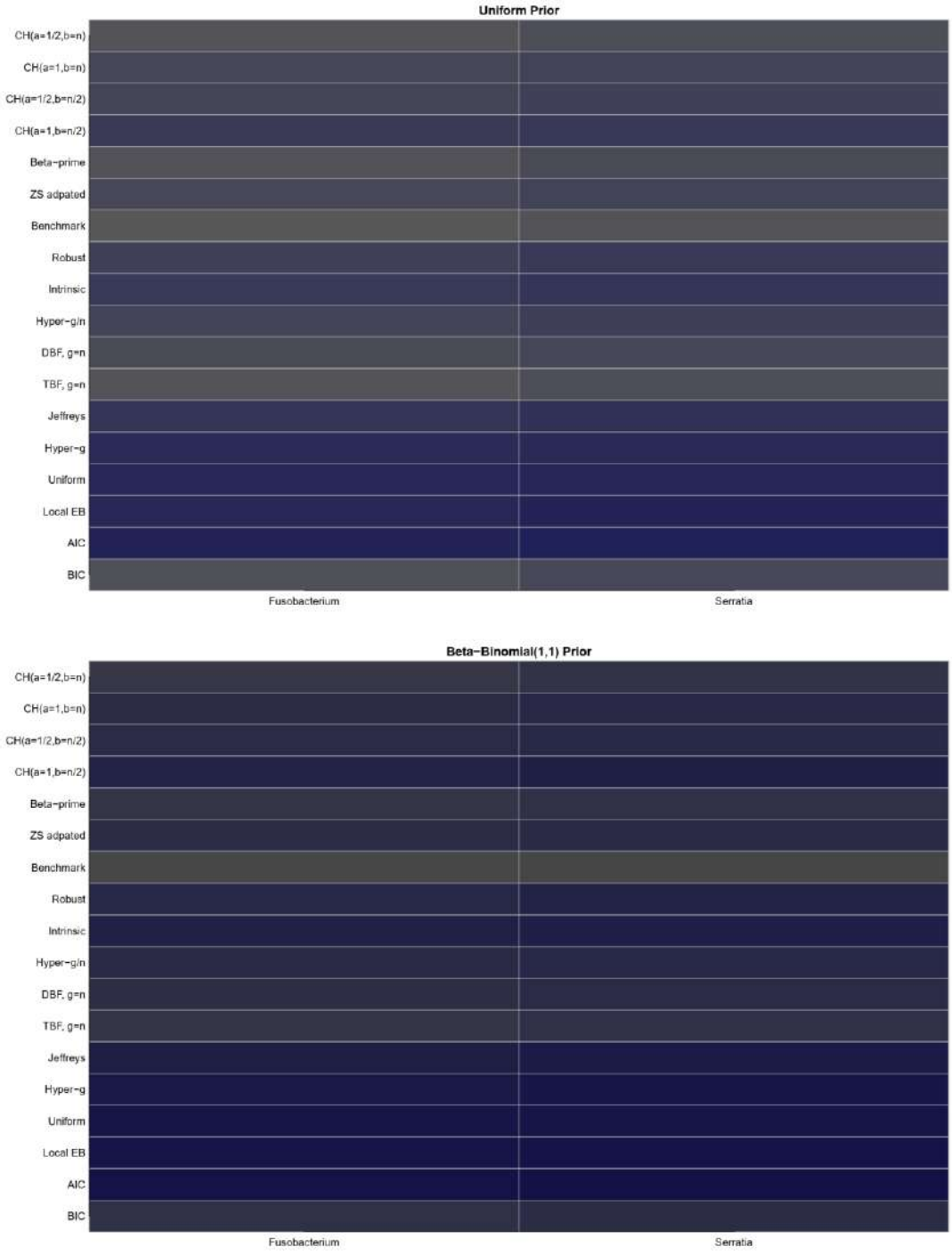


Figure E-E-4 –Marginal posterior inclusion probabilities for the **Genus** for **Biopsy** with **Rank Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

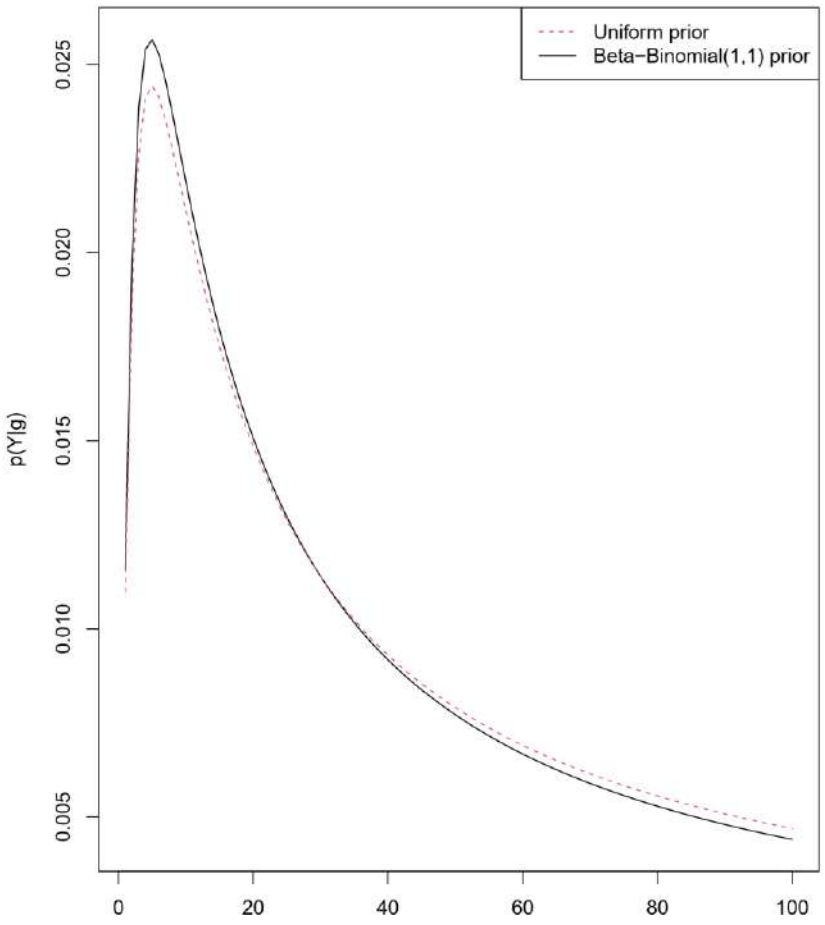


Figure E-E-5: Marginal likelihood of *g* for the **Genus** for **Biopsy** with **Rank Unit** Normalization method.

Table E-E-4: Prediction accuracy for the **Biopsy (Phylum) Rank Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.8483	0.8488	28.9128	20.1607	0.4894	0.4902	0.1475	0.1476
02_CH(a=1,b=n)	0.8467	0.8493	16.8718	14.2302	0.4871	0.4864	0.1459	0.1455
03_CH(a=1/2,b=n/2)	0.8474	0.8500	17.2445	14.8551	0.4870	0.4867	0.1461	0.1458
04_CH(a=1,b=n/2)	0.8481	0.8490	12.7302	16.6745	0.4822	0.4825	0.1446	0.1441
05_Beta-prime	0.8485	0.8487	18.4211	19.5360	0.4893	0.4903	0.1474	0.1475
06_ZS adpated	0.8465	0.8493	16.9857	18.2808	0.4871	0.4867	0.1460	0.1456
07_Robust	0.8484	0.8491	16.1908	13.2808	0.4821	0.4814	0.1443	0.1438
08_Intrinsic	0.8488	0.8492	12.2392	16.1826	0.4805	0.4792	0.1437	0.1430
09_Hyper-g/n	0.8483	0.8490	16.9713	17.8541	0.4853	0.4851	0.1457	0.1453
10_Benchmark	0.8486	0.8513	25.6554	29.0627	0.5010	0.5001	0.1531	0.1533
11_DBF, g=n	0.8483	0.8489	26.8907	17.2525	0.4855	0.4853	0.1455	0.1454
12_TBF, g=n	0.8479	0.8498	17.0644	17.5050	0.4857	0.4843	0.1455	0.1454
13_Jeffreys	0.8492	0.8503	**4.5288**	**3.8968**	0.4764	**0.4718**	0.1423	**0.1405**
14_Hyper-g	0.8506	0.8517	5.5115	5.7340	0.4791	0.4761	0.1443	0.1433
15_Uniform	**0.8517**	**0.8529**	6.1443	8.7157	0.4836	0.4791	0.1456	0.1446
16_Local EB	0.8096	0.8100	25.6455	11.0306	0.5122	0.5086	0.1554	0.1544
17_AIC	0.8502	0.8496	11.4680	11.4086	**0.4750**	0.4754	**0.1419**	0.1414
18_BIC	0.8487	0.8500	15.8776	12.5573	0.4806	0.4814	0.1438	0.1435

F-Stool

A-Phylum

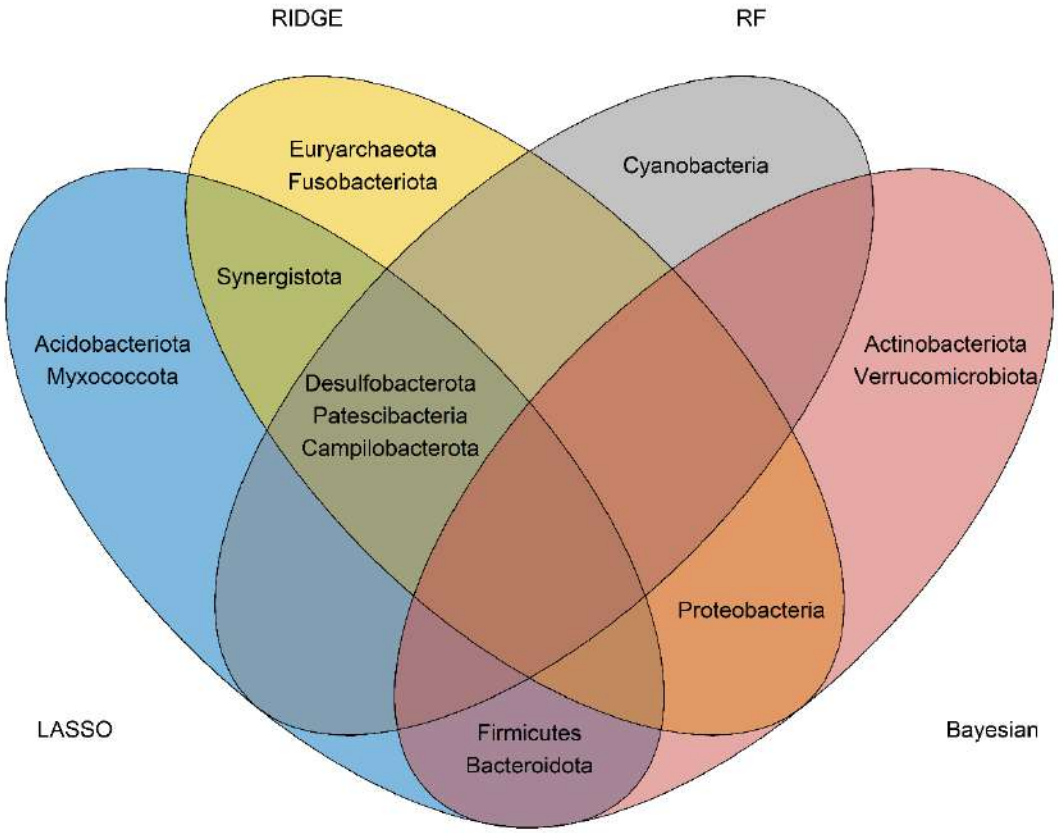


Figure F-A-1 – The Concordance Plot for **Phylum** for **Stool** .

Table F-A-1: The Selected method

Row	Method	Filter	Normal	AUC
1	LASSO	PA	RU	0.63
2	RIDGE	AB	LU	0.69
3	RF	AB	RU	0.7
4	**Bayesian**	**VA**	**LU**	**0.77**

Table F-A-2: The common feature names

Common Features for BAS			
Synergistota	Desulfobacterota	Patescibacteria	Campilobacterota
Firmicutes	Bacteroidota	Proteobacteria	

BAS

Log Unit

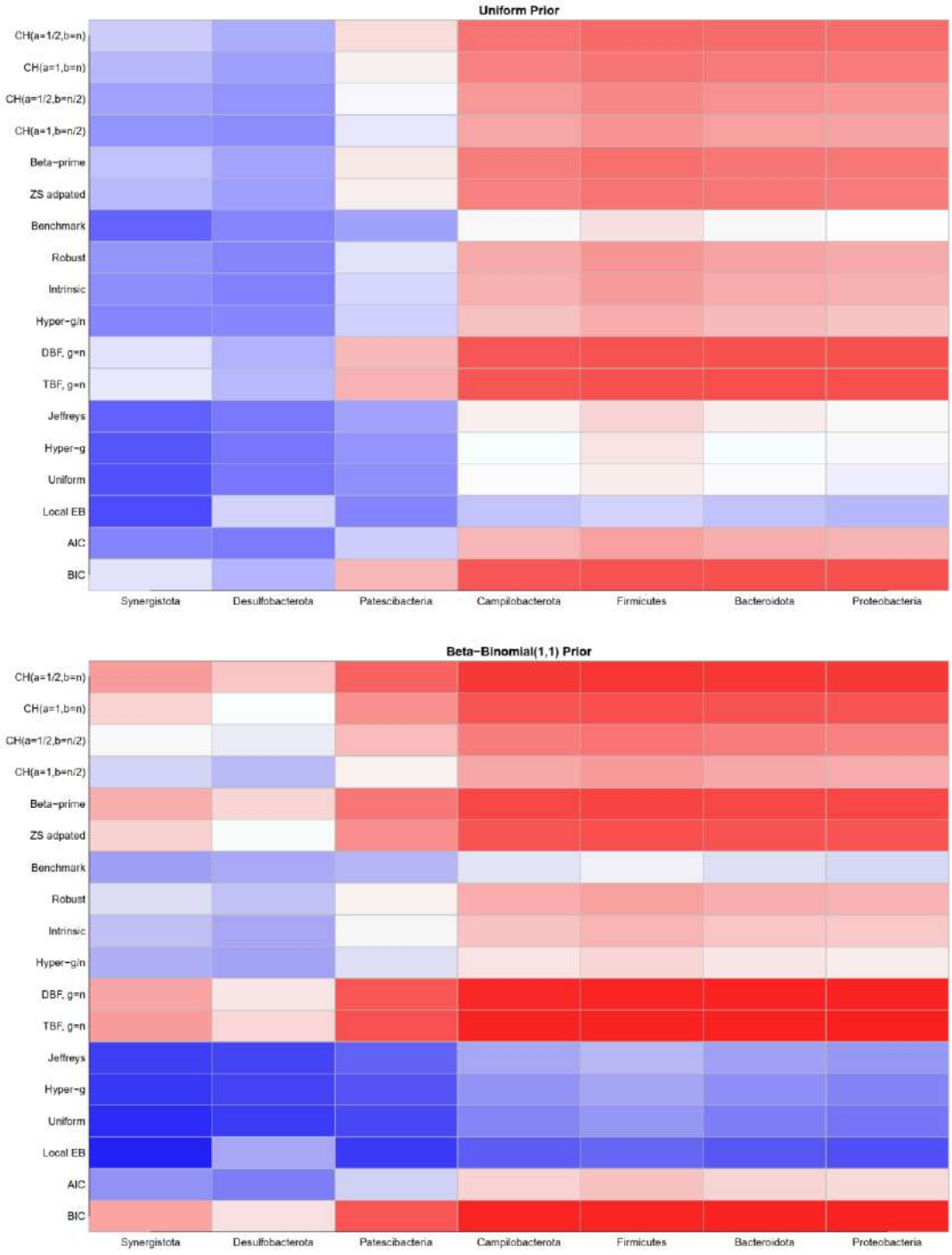


Figure F-A-2 –Marginal posterior inclusion probabilities for the **Phylum** for **Stool** with **Log Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

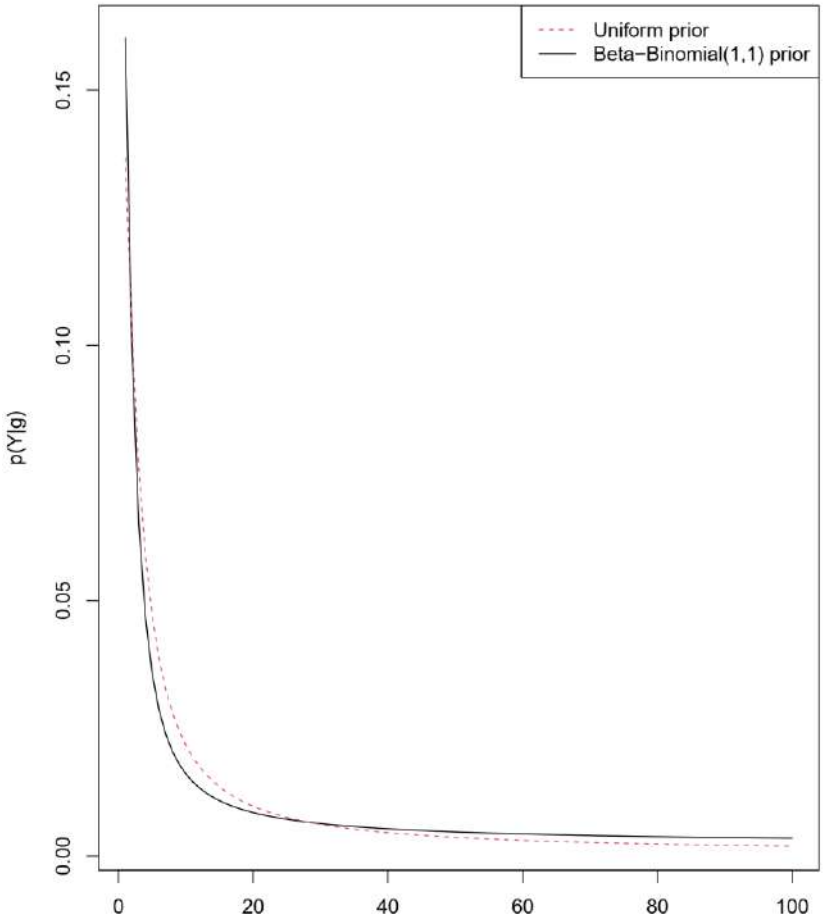


Figure F-A-3: Marginal likelihood of *g* for the **Phylum** for **Stool** with **Log Unit** Normalization method.

Table F-A-3: Prediction accuracy for the **Stool (Phylum) Log Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.7414	0.7401	10.4397	15.0742	0.5738	0.5760	0.1631	0.1640
02_CH(a=1,b=n)	0.7413	0.7404	8.1657	28.6255	0.5750	0.5767	0.1630	0.1642
03_CH(a=1/2,b=n/2)	0.7424	0.7427	8.2404	27.2355	0.5759	0.5815	0.1632	0.1655
04_CH(a=1,b=n/2)	0.7450	0.7424	11.2609	26.6357	0.5757	0.5855	0.1632	0.1658
05_Beta-prime	0.7416	0.7409	8.3129	28.5335	0.5742	0.5768	0.1633	0.1645
06_ZS adpated	0.7414	0.7403	8.1789	27.5259	0.5745	0.5760	0.1630	0.1642
07_Robust	0.7488	0.7411	15.9368	29.3548	0.5763	0.6055	0.1636	0.1703
08_Intrinsic	0.7256	0.7201	13.0135	29.6717	**0.5709**	0.5999	0.1631	0.1693
09_Hyper-g/n	0.7284	0.7253	19.5535	29.5800	0.5772	0.5930	0.1623	0.1664
10_Benchmark	0.7393	0.7300	10.1421	17.8811	0.5889	0.6009	0.1625	0.1668
11_DBF, g=n	0.7397	0.7399	7.2591	16.0313	0.5750	0.5726	0.1633	0.1641
12_TBF, g=n	0.7356	0.7389	8.1463	28.3184	0.5809	**0.5705**	0.1651	0.1642
13_Jeffreys	**0.7562**	**0.7437**	28.8029	30.5801	0.5793	0.5989	0.1618	0.1640
14_Hyper-g	0.7502	0.7383	29.0041	**11.2130**	0.5733	0.6010	**0.1615**	0.1640
15_Uniform	0.7457	0.7335	**6.1584**	29.3156	0.5755	0.6021	0.1615	0.1642
16_Local EB	0.5356	0.5304	578.5094	1153.2322	0.6675	0.7411	0.1819	0.1860
17_AIC	0.7428	0.7412	14.7242	14.1202	0.5714	0.5779	0.1635	0.1650
18_BIC	0.7410	0.7378	6.9214	15.0297	0.5771	0.5720	0.1640	**0.1639**

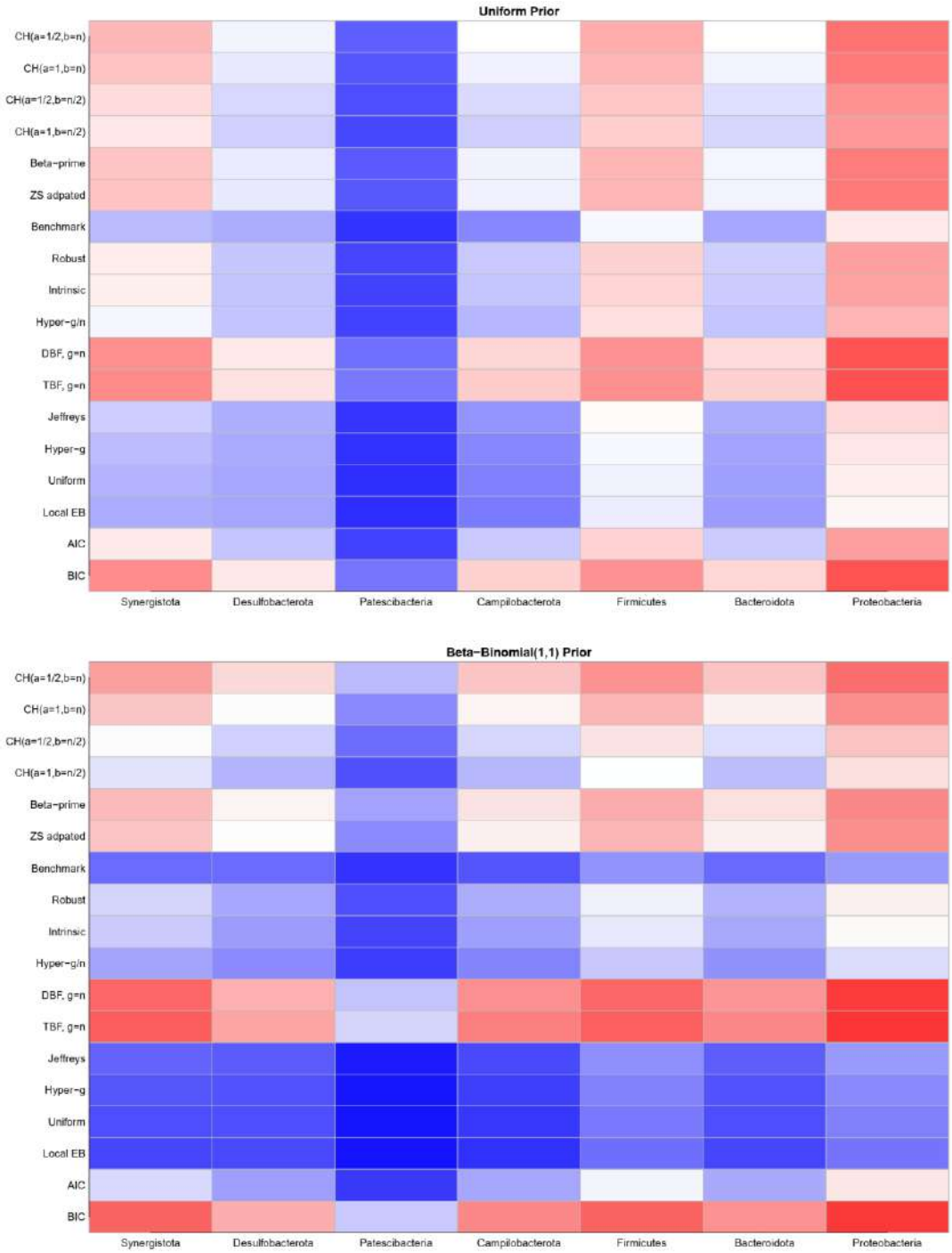


Figure F-A-4 –Marginal posterior inclusion probabilities for the **Phylum** for **Stool** with **Rank Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

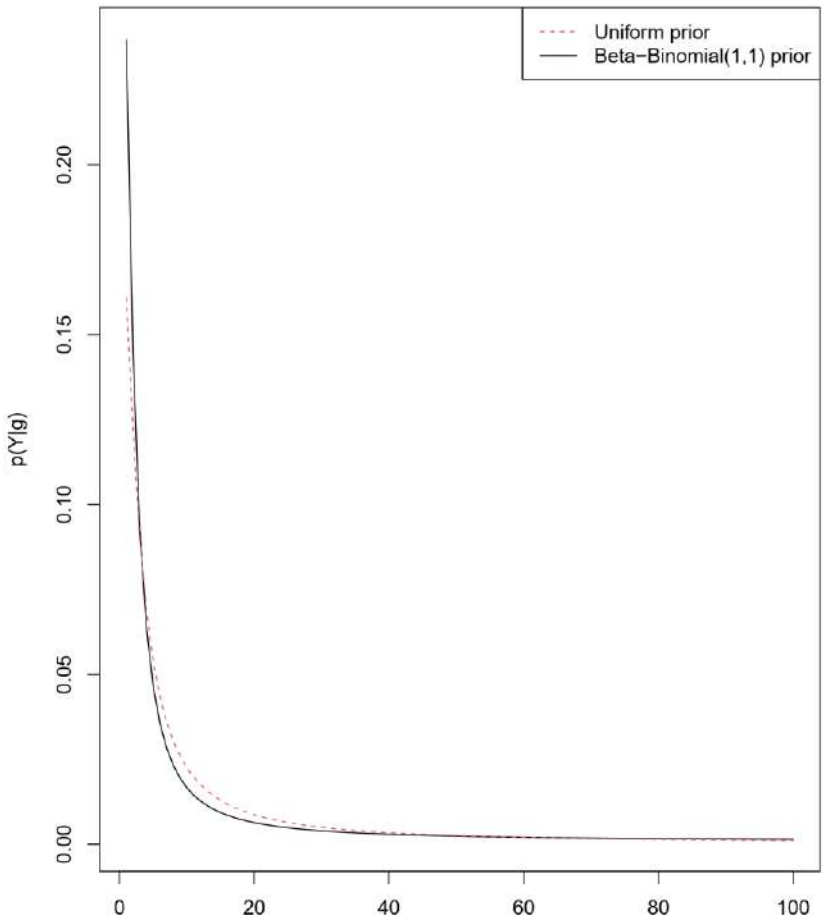


Figure F-A-5: Marginal likelihood of *g* for the **Phylum** for **Stool** with **Rank Unit** Normalization method.

Table F-A-4: Prediction accuracy for the **Stool (Phylum) Rank Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.7782	0.7773	12.2600	5.0236	0.5281	0.5281	0.1475	0.1486
02_CH(a=1,b=n)	0.7767	0.7770	2.5398	3.9355	0.5279	0.5305	0.1472	0.1488
03_CH(a=1/2,b=n/2)	0.7741	0.7788	2.4785	2.8845	0.5320	0.5345	0.1475	0.1500
04_CH(a=1,b=n/2)	0.7745	0.7796	2.5550	2.8046	0.5319	0.5356	0.1475	0.1496
05_Beta-prime	0.7767	0.7764	2.4363	4.7385	0.5277	0.5324	0.1473	0.1493
06_ZS adpated	0.7767	0.7764	2.5209	3.8655	0.5274	0.5311	0.1470	0.1489
07_Robust	0.7784	0.7717	2.8121	2.8576	0.5281	0.5515	0.1473	0.1525
08_Intrinsic	0.7474	0.7391	5.8073	2.2842	0.5122	0.5359	0.1424	0.1475
09_Hyper-g/n	0.7512	0.7454	6.1894	4.0983	**0.5081**	0.5328	**0.1409**	0.1453
10_Benchmark	0.7752	0.7653	8.2151	8.6345	0.5356	0.5489	0.1436	0.1466
11_DBF, g=n	0.7763	0.7774	5.4368	5.3868	0.5280	0.5276	0.1482	0.1488
12_TBF, g=n	0.7789	0.7764	3.9119	6.1403	0.5167	0.5198	0.1452	0.1467
13_Jeffreys	**0.7828**	**0.7801**	5.9327	9.0053	0.5376	0.5410	0.1457	0.1454
14_Hyper-g	0.7819	0.7743	7.8545	6.2137	0.5405	0.5486	0.1452	0.1450
15_Uniform	0.7793	0.7710	16.2184	10.5005	0.5388	0.5460	0.1445	**0.1445**
16_Local EB	0.5293	0.5284	611.6174	1163.7065	0.6589	0.7622	0.1789	0.1837
17_AIC	0.7779	0.7770	**2.0516**	**1.8324**	0.5181	**0.5186**	0.1455	0.1456
18_BIC	0.7815	0.7722	8.0542	3.2282	0.5137	0.5238	0.1438	0.1471

B-Class

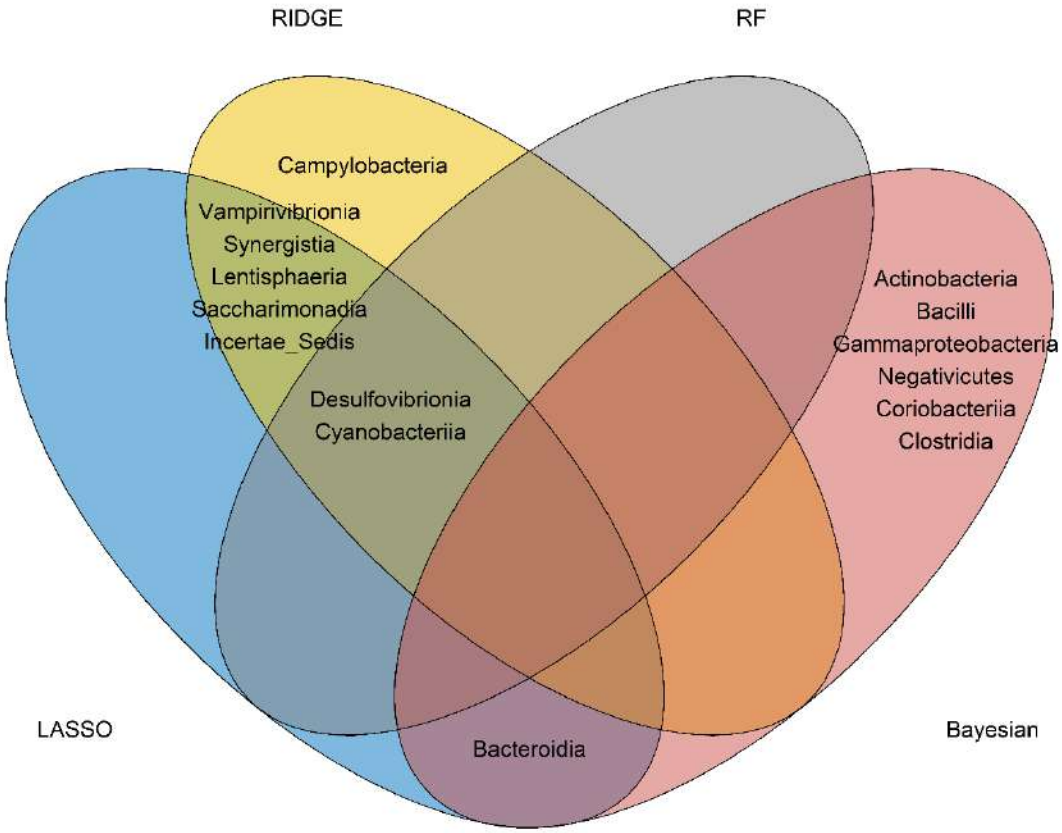


Figure F-B-1 – The Concordance Plot for **Class** for **Stool** .

Table F-B-1: The Selected method				
Row	**Method**	**Filter**	**Normal**	**AUC**
1	LASSO	AB	RU	0.74
2	**RIDGE**	**AB**	**LU**	**0.79**
3	RF	PA	RU	0.76
4	**Bayesian**	**VA**	**LU**	**0.79**

Table F-B-2: The common feature names			
Common Features for BAS			
Vampirivibrionia	Synergistia	Lentisphaeria	Saccharimonadia
Incertae_Sedis	Desulfovibrionia	Cyanobacteriia	Bacteroidia

BAS

Log Unit

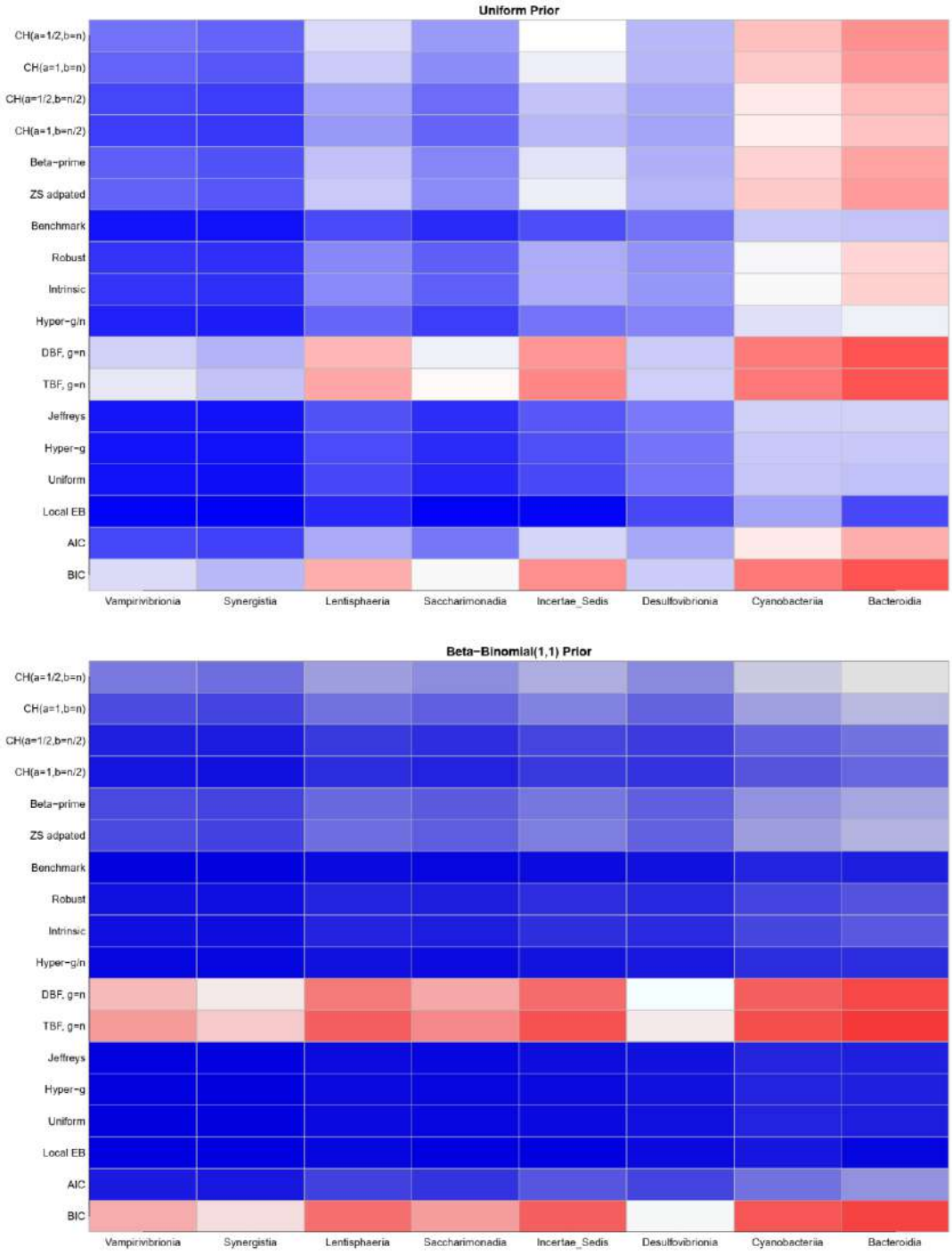


Figure F-B-2 –Marginal posterior inclusion probabilities for the **Class** for **Stool** with **Log Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

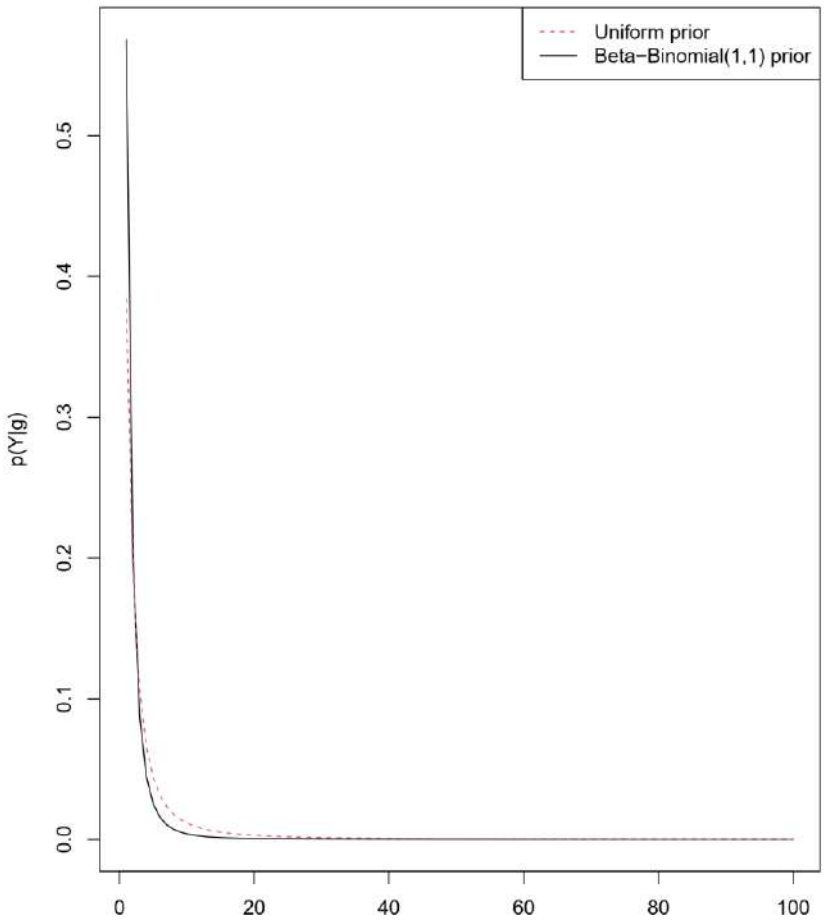


Figure F-B-3: Marginal likelihood of *g* for the **Class** for **Stool** with **Log Unit** Normalization method.

Table F-B-3: Prediction accuracy for the **Stool (Phylum) Log Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.7668	0.7643	1.7415	3.0670	0.5328	0.5420	0.1487	0.1492
02_CH(a=1,b=n)	0.7679	0.7652	2.4137	2.1508	0.5341	0.5403	0.1488	0.1491
03_CH(a=1/2,b=n/2)	0.7652	0.7635	2.3945	2.6566	0.5421	0.5508	0.1485	0.1496
04_CH(a=1,b=n/2)	0.7655	0.7654	2.0967	2.2695	0.5440	0.5523	0.1485	0.1496
05_Beta-prime	0.7675	0.7650	2.2086	2.5784	0.5376	0.5409	0.1490	0.1490
06_ZS adpated	0.7677	0.7641	2.9612	2.3186	0.5340	0.5418	0.1489	0.1491
07_Robust	0.7684	0.7688	2.2070	3.3981	0.5384	0.5415	0.1477	0.1487
08_Intrinsic	0.6627	0.6657	7.4111	3.0712	0.5286	0.5267	0.1450	0.1446
09_Hyper-g/n	0.6669	0.6672	5.2629	5.3168	**0.5136**	**0.5174**	0.1412	0.1372
10_Benchmark	0.7642	0.7622	5.4196	12.1272	0.5316	0.5315	0.1386	0.1405
11_DBF, g=n	0.7682	0.7620	2.0945	4.1183	0.5390	0.5439	0.1490	0.1507
12_TBF, g=n	0.7608	0.7643	**1.2515**	4.6187	0.5488	0.5452	0.1525	0.1514
13_Jeffreys	**0.7746**	**0.7719**	6.0384	9.8922	0.5232	0.5203	0.1408	**0.1354**
14_Hyper-g	0.7717	0.7634	6.9743	6.2434	0.5271	0.5260	0.1395	0.1359
15_Uniform	0.7690	0.7623	9.2903	11.1969	0.5257	0.5314	**0.1385**	0.1377
16_Local EB	0.5009	0.5008	1356.3600	1490.7238	0.7419	0.8489	0.1816	0.1866
17_AIC	0.7633	0.7619	1.9727	**1.8432**	0.5476	0.5492	0.1522	0.1528
18_BIC	0.7644	0.7651	2.7013	8.4868	0.5431	0.5450	0.1518	0.1511

Rank Unit

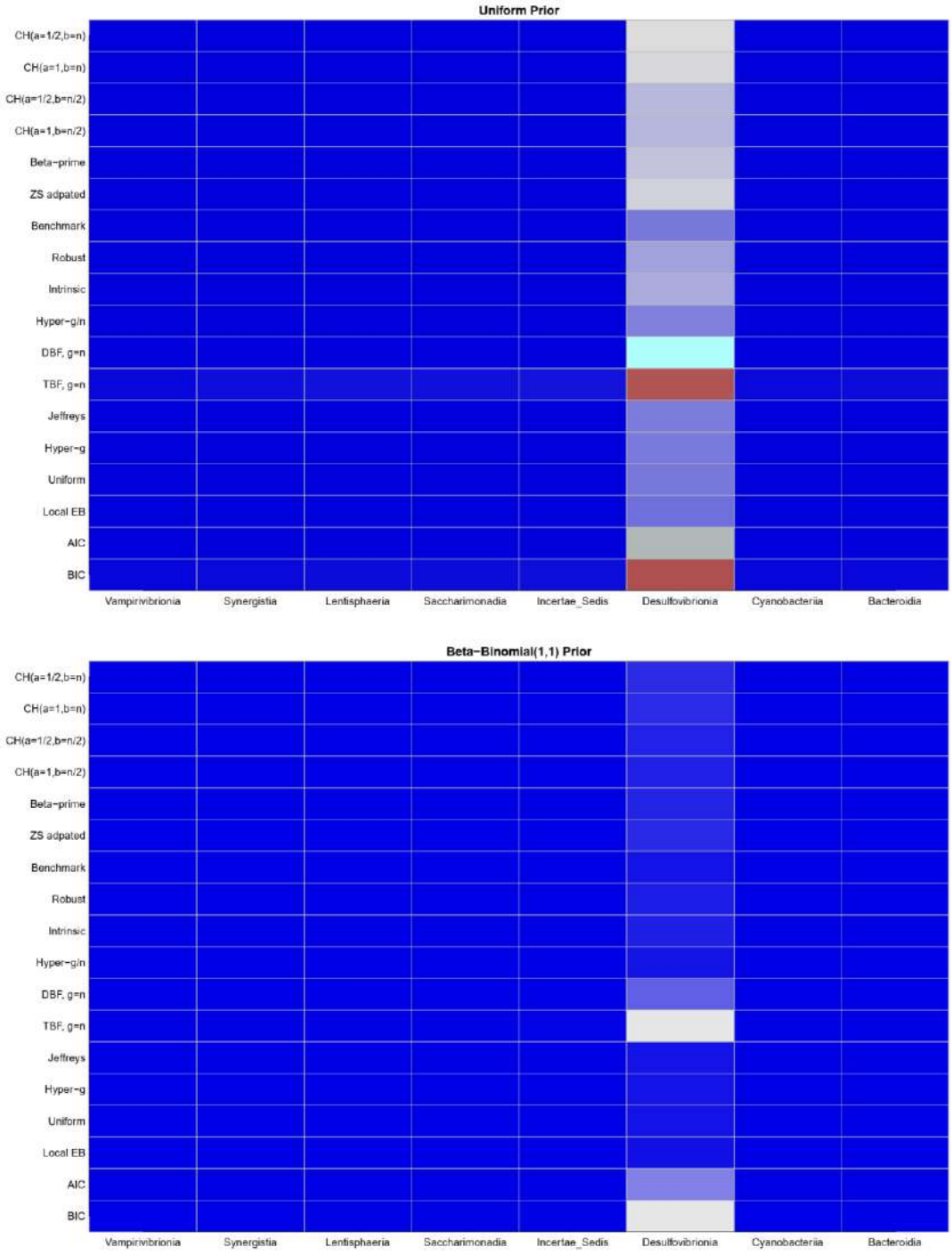


Figure F-B-4 –Marginal posterior inclusion probabilities for the **Class** for **Stool** with **Rank Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

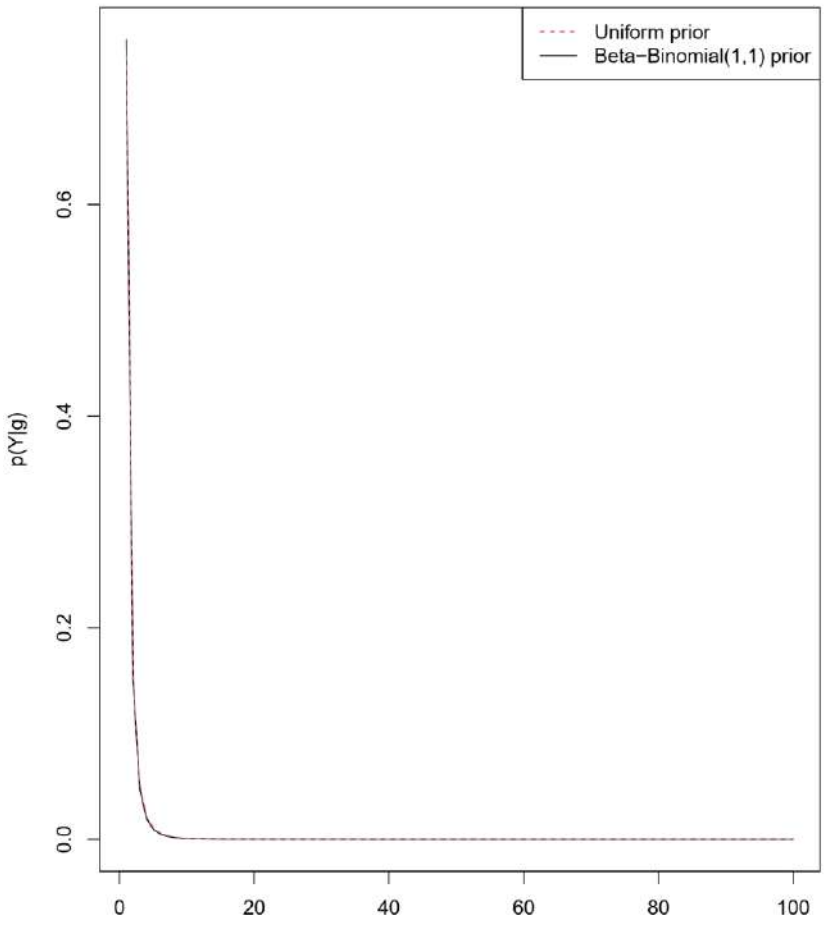


Figure F-B-5: Marginal likelihood of *g* for the **Class** for **Stool** with **Rank Unit** Normalization method.

Table F-B-4: Prediction accuracy for the **Stool (Phylum) Rank Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.7875	0.7885	2.9516	3.8262	0.4498	0.4487	0.1220	0.1218
02_CH(a=1,b=n)	0.7878	0.7885	3.3410	3.1367	0.4487	0.4465	0.1217	0.1211
03_CH(a=1/2,b=n/2)	0.7888	**0.7915**	3.0238	3.1445	0.4431	0.4438	0.1197	0.1199
04_CH(a=1,b=n/2)	0.7889	0.7904	3.6356	3.3972	0.4439	0.4445	0.1194	0.1199
05_Beta-prime	0.7878	0.7897	6.2463	3.6613	0.4482	0.4449	0.1216	0.1207
06_ZS adpated	0.7881	0.7892	3.6827	3.5547	0.4477	0.4457	0.1216	0.1209
07_Robust	**0.7911**	0.7888	4.0798	3.6967	0.4421	0.4562	0.1187	0.1216
08_Intrinsic	0.6338	0.6332	4.0066	3.9338	0.4500	0.4598	0.1212	0.1228
09_Hyper-g/n	0.6333	0.6344	4.7305	6.3888	0.4484	0.4470	0.1183	0.1147
10_Benchmark	0.7853	0.7733	9.1716	8.7132	**0.4335**	0.4436	**0.1110**	0.1146
11_DBF, g=n	0.7847	0.7872	3.0487	2.8719	0.4576	0.4552	0.1247	0.1242
12_TBF, g=n	0.7843	0.7847	3.0709	2.5602	0.4687	0.4699	0.1291	0.1292
13_Jeffreys	0.7897	0.7846	6.9480	6.8023	0.4367	**0.4386**	0.1132	0.1120
14_Hyper-g	0.7876	0.7821	8.0826	7.7279	0.4357	0.4412	0.1121	**0.1118**
15_Uniform	0.7858	0.7780	7.0843	8.5083	0.4374	0.4421	0.1113	0.1127
16_Local EB	0.5000	0.4999	**0.1091**	**-0.5823**	0.8009	0.9286	0.1810	0.1845
17_AIC	0.7862	0.7875	2.6930	2.4334	0.4603	0.4541	0.1258	0.1249
18_BIC	0.7855	0.7865	2.7988	3.4071	0.4635	0.4658	0.1270	0.1278

C-Order

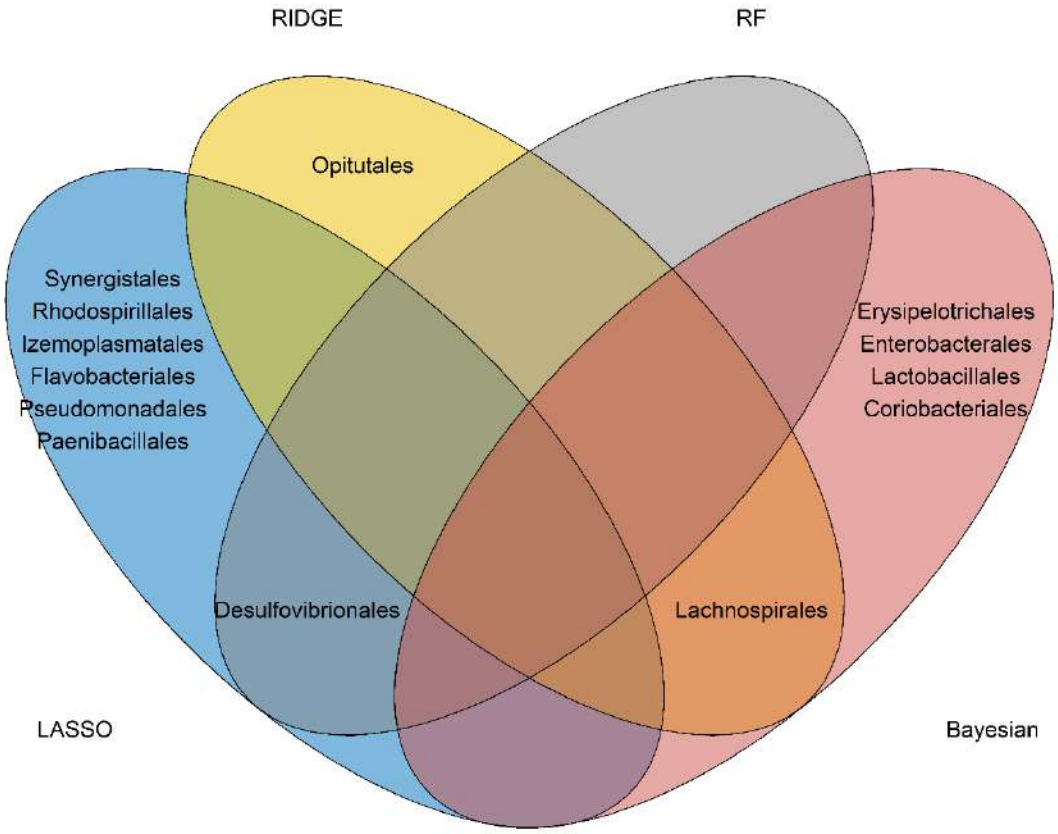


Figure F-C-1 – The Concordance Plot for **Order** for **Stool** .

Table F-C-1: The Selected method				
Row	**Method**	**Filter**	**Normal**	**AUC**
1	LASSO	AB	LU	0.69
2	RIDGE	PR	RU	0.75
3	RF	AB	LU	0.71
4	**Bayesian**	**VA**	**LU**	**0.76**

Table F-C-2: The common feature names	
Common Features for BAS	
Desulfovibrionales	Lachnospirales

BAS

Log Unit

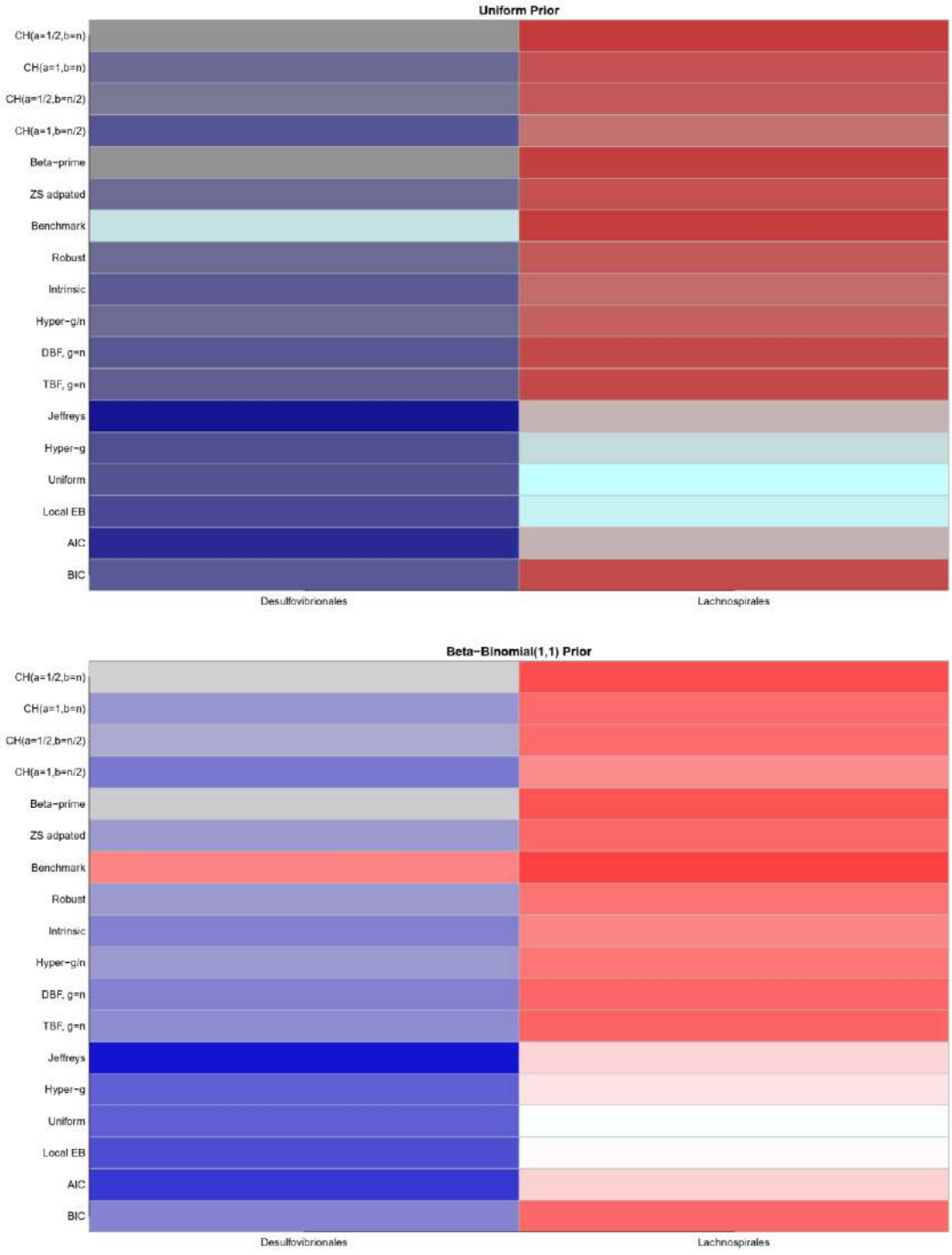


Figure F-C-2 –Marginal posterior inclusion probabilities for the **Order** for **Stool** with **Log Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

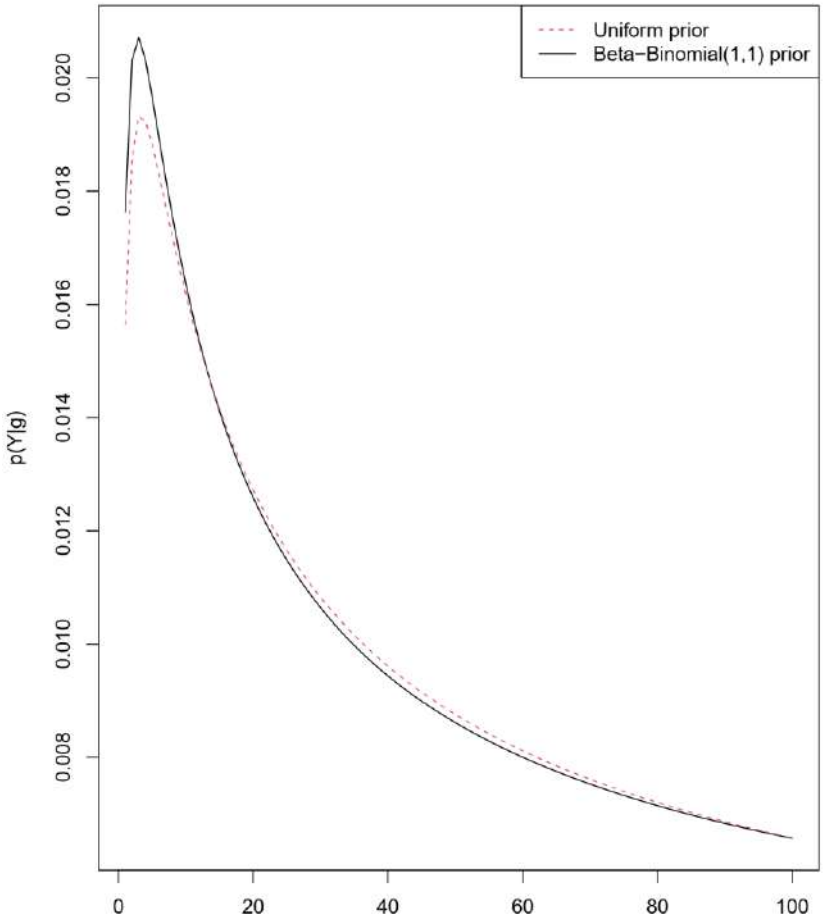


Figure F-C-3: Marginal likelihood of *g* for the **Order** for **Stool** with **Log Unit** Normalization method.

Table F-C-3: Prediction accuracy for the **Stool (Phylum) Log Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.6188	0.6176	13.2880	19.5216	0.4839	0.4878	0.1436	0.1451
02_CH(a=1,b=n)	0.6189	0.6186	10.0186	13.1872	0.4839	0.4871	0.1427	0.1441
03_CH(a=1/2,b=n/2)	0.6180	0.6187	10.1886	17.8765	0.4837	0.4868	0.1431	0.1445
04_CH(a=1,b=n/2)	0.6181	0.6195	8.2218	10.2888	0.4826	0.4842	0.1423	0.1435
05_Beta-prime	0.6183	0.6175	13.1218	19.3247	0.4837	0.4877	0.1435	0.1450
06_ZS adpated	0.6182	0.6191	8.8361	17.4987	0.4840	0.4871	0.1427	0.1441
07_Robust	0.6191	0.6193	8.1483	12.3089	0.4833	0.4861	0.1424	0.1436
08_Intrinsic	0.6199	0.6192	7.1050	8.5855	0.4826	0.4854	0.1421	0.1432
09_Hyper-g/n	0.6191	0.6189	10.1095	13.1603	0.4817	0.4844	0.1427	0.1441
10_Benchmark	0.6203	**0.6233**	26.5130	26.5615	0.4848	0.4864	0.1477	0.1488
11_DBF, g=n	0.6177	0.6182	8.2132	12.6012	0.4837	0.4876	0.1423	0.1436
12_TBF, g=n	0.6177	0.6184	9.5528	12.7796	0.4826	0.4864	0.1423	0.1437
13_Jeffreys	0.6196	0.6211	**4.2827**	**4.1094**	**0.4738**	**0.4730**	**0.1403**	**0.1405**
14_Hyper-g	0.6210	0.6212	7.8531	9.0663	0.4740	0.4753	0.1428	0.1435
15_Uniform	**0.6211**	0.6211	8.5015	9.5721	0.4760	0.4770	0.1438	0.1444
16_Local EB	0.5987	0.5978	11.5541	10.1214	0.4887	0.4890	0.1448	0.1454
17_AIC	0.6199	0.6198	4.5174	5.6052	0.4849	0.4884	0.1414	0.1423
18_BIC	0.6188	0.6193	6.3187	8.4106	0.4837	0.4873	0.1418	0.1430

Rank Unit

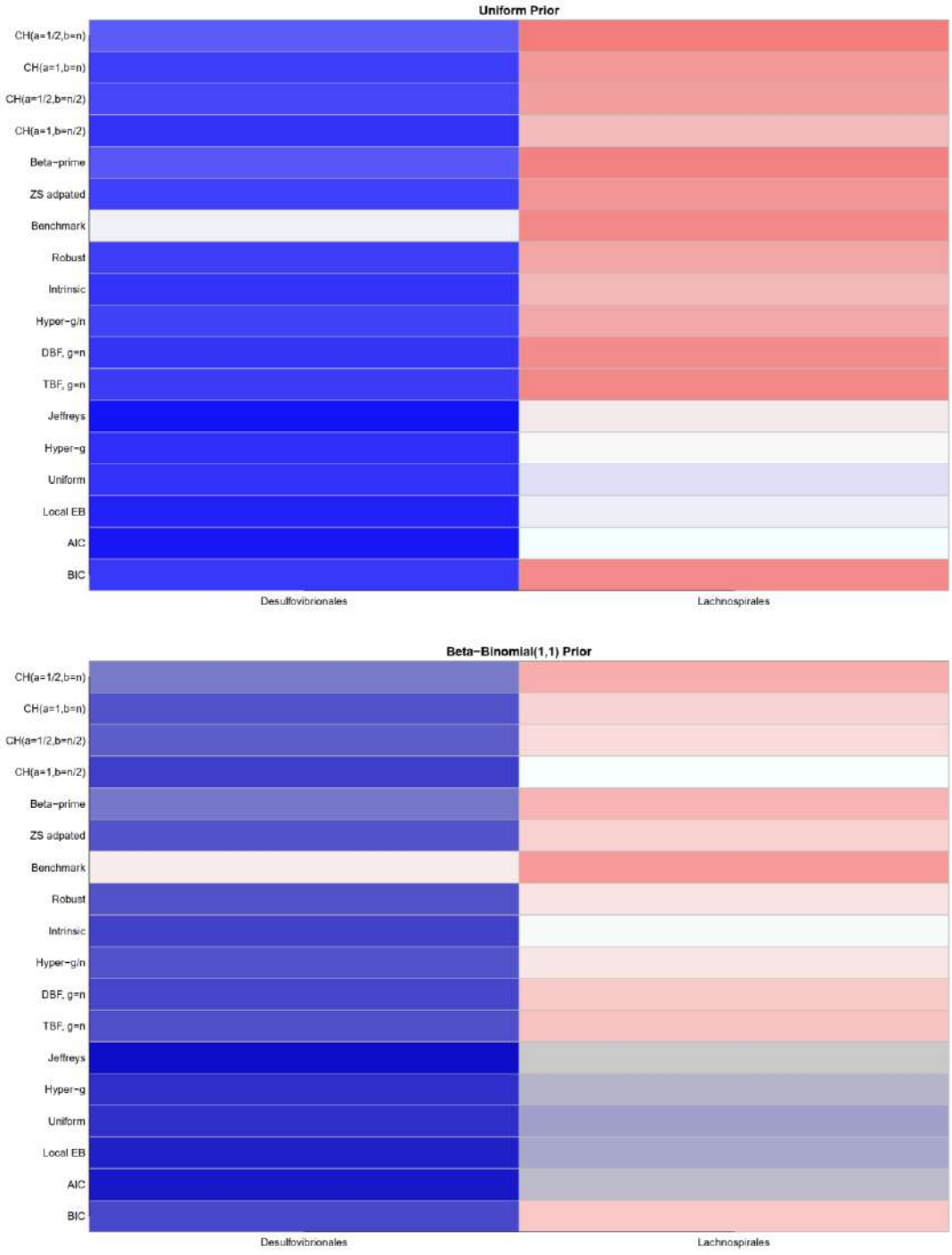


Figure F-C-4 –Marginal posterior inclusion probabilities for the **Order** for **Stool** with **Rank Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

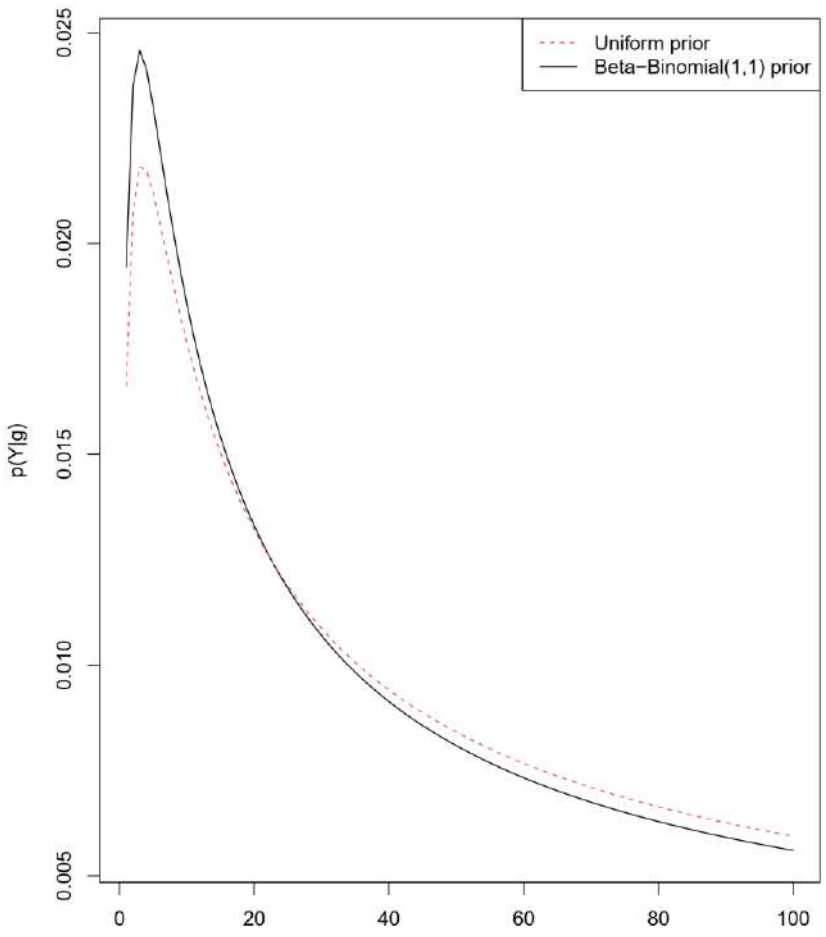


Figure F-C-5: Marginal likelihood of *g* for the **Order** for **Stool** with **Rank Unit** Normalization method.

Table F-C-4: Prediction accuracy for the **Stool (Phylum) Rank Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.6716	0.6761	5.7319	7.9671	0.4653	0.4678	0.1377	0.1382
02_CH(a=1,b=n)	0.6718	0.6783	4.2884	5.4943	0.4635	0.4647	0.1365	0.1368
03_CH(a=1/2,b=n/2)	0.6723	0.6793	4.6941	6.2606	0.4628	0.4641	0.1366	0.1370
04_CH(a=1,b=n/2)	0.6742	0.6810	3.7848	4.5359	0.4605	0.4615	0.1356	0.1357
05_Beta-prime	0.6717	0.6775	5.5077	7.7161	0.4650	0.4675	0.1376	0.1381
06_ZS adpated	0.6716	0.6780	4.3989	5.6590	0.4636	0.4649	0.1366	0.1369
07_Robust	0.6751	0.6822	3.7242	4.4757	0.4625	0.4633	0.1357	0.1358
08_Intrinsic	0.6766	0.6840	3.4510	3.9657	0.4613	0.4620	0.1351	0.1352
09_Hyper-g/n	0.6722	0.6789	4.4751	5.5401	0.4609	0.4621	0.1361	0.1364
10_Benchmark	0.6790	0.6864	17.1984	24.4027	0.4653	0.4667	0.1410	0.1418
11_DBF, g=n	0.6716	0.6764	3.8768	4.8531	0.4641	0.4657	0.1364	0.1367
12_TBF, g=n	0.6708	0.6758	4.0316	5.1019	0.4635	0.4651	0.1365	0.1368
13_Jeffreys	0.6761	0.6838	3.0519	2.9888	0.4515	0.4503	**0.1329**	**0.1322**
14_Hyper-g	0.6799	0.6864	4.6423	4.9035	0.4533	0.4531	0.1347	0.1346
15_Uniform	**0.6832**	**0.6880**	5.4499	5.5700	**0.4505**	**0.4501**	0.1356	0.1354
16_Local EB	0.6610	0.6652	27.7955	14.3523	0.4614	0.4615	0.1386	0.1385
17_AIC	0.6779	0.6848	**2.6400**	**2.8158**	0.4634	0.4641	0.1346	0.1345
18_BIC	0.6725	0.6795	3.2369	3.7942	0.4637	0.4654	0.1357	0.1357

D-Family

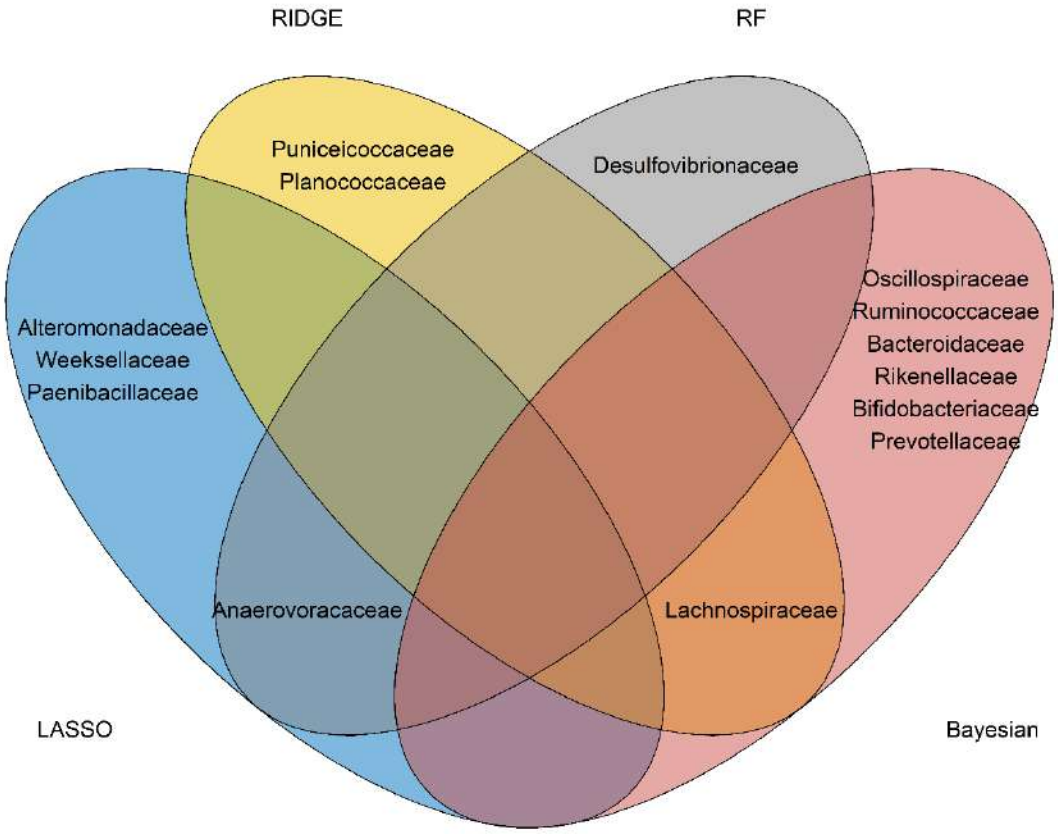


Figure F-D-1 – The Concordance Plot for **Family** for **Stool** .

Table F-D-1: The Selected method				
Row	**Method**	**Filter**	**Normal**	**AUC**
1	LASSO	CA	RU	0.65
2	RIDGE	PA	RU	0.79
3	RF	PA	LU	0.70
4	**Bayesian**	**VA**	**RU**	**0.86**

Table F-D-2: The common feature names	
Common Features for BAS	
Anaerovoracaceae	Lachnospiraceae

BAS

Log Unit

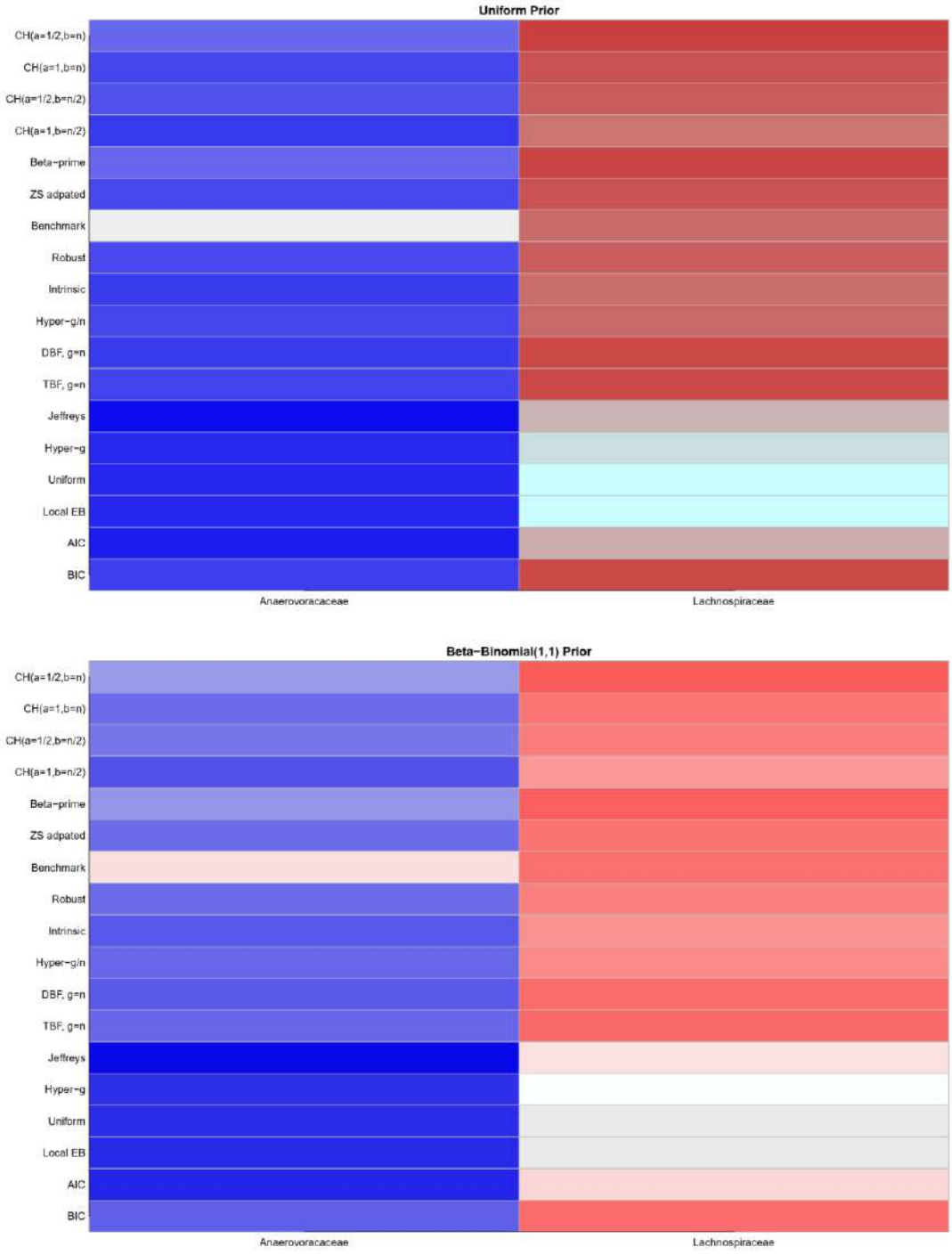


Figure F-D-2 –Marginal posterior inclusion probabilities for the **Family** for **Stool** with **Log Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

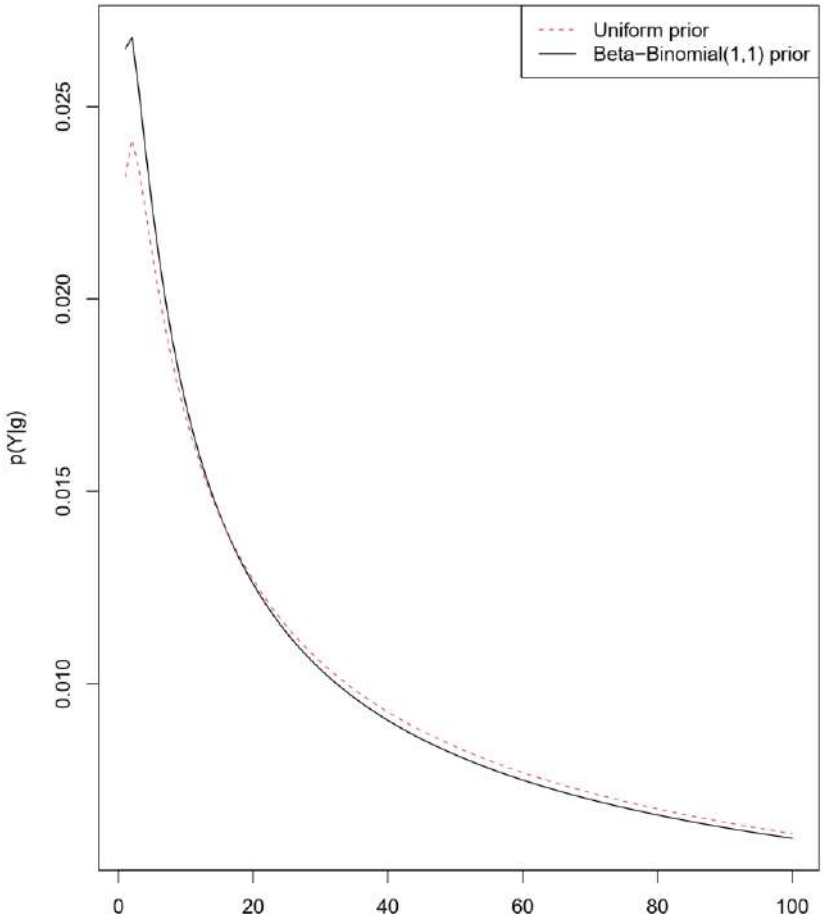


Figure F-D-3: Marginal likelihood of *g* for the **Family** for **Stool** with **Log Unit** Normalization method.

Table F-D-3: Prediction accuracy for the **Stool (Phylum) Log Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.6731	0.6734	-7.7538	-4.9801	0.4530	0.4566	0.1370	0.1386
02_CH(a=1,b=n)	0.6731	0.6739	-0.6320	-8.5094	0.4502	0.4526	0.1356	0.1370
03_CH(a=1/2,b=n/2)	0.6737	0.6735	-0.1860	-17.6677	0.4507	0.4532	0.1360	0.1375
04_CH(a=1,b=n/2)	0.6734	0.6741	2.3411	2.3788	0.4483	0.4520	0.1349	0.1362
05_Beta-prime	0.6732	0.6731	-17.1841	-4.9016	0.4528	0.4563	0.1369	0.1385
06_ZS adpated	0.6735	0.6740	-0.4487	-8.4374	0.4503	0.4528	0.1356	0.1371
07_Robust	0.6733	0.6737	2.1701	2.2630	0.4493	0.4529	0.1350	0.1363
08_Intrinsic	0.6737	0.6742	1.7195	**1.7350**	0.4482	0.4516	0.1346	0.1358
09_Hyper-g/n	0.6737	0.6737	-0.6282	-8.4744	0.4492	0.4531	0.1356	0.1371
10_Benchmark	**0.6755**	0.6733	14.3558	14.4659	0.4630	0.4655	0.1426	0.1438
11_DBF, g=n	0.6738	0.6734	-0.0814	0.0112	0.4490	0.4526	0.1351	0.1365
12_TBF, g=n	0.6733	0.6736	2.9020	4.1857	0.4493	0.4524	0.1352	0.1366
13_Jeffreys	0.6741	0.6752	**0.9718**	2.0896	**0.4419**	**0.4410**	**0.1331**	**0.1333**
14_Hyper-g	0.6744	0.6766	3.5248	4.2037	0.4467	0.4463	0.1356	0.1359
15_Uniform	0.6747	**0.6771**	4.8565	5.5488	0.4482	0.4493	0.1366	0.1368
16_Local EB	0.6543	0.6573	52.6039	111.9692	0.4603	0.4630	0.1418	0.1430
17_AIC	0.6741	0.6747	2.3287	2.1880	0.4480	0.4517	0.1340	0.1349
18_BIC	0.6736	0.6743	1.5680	2.4784	0.4482	0.4506	0.1344	0.1356

Rank Unit

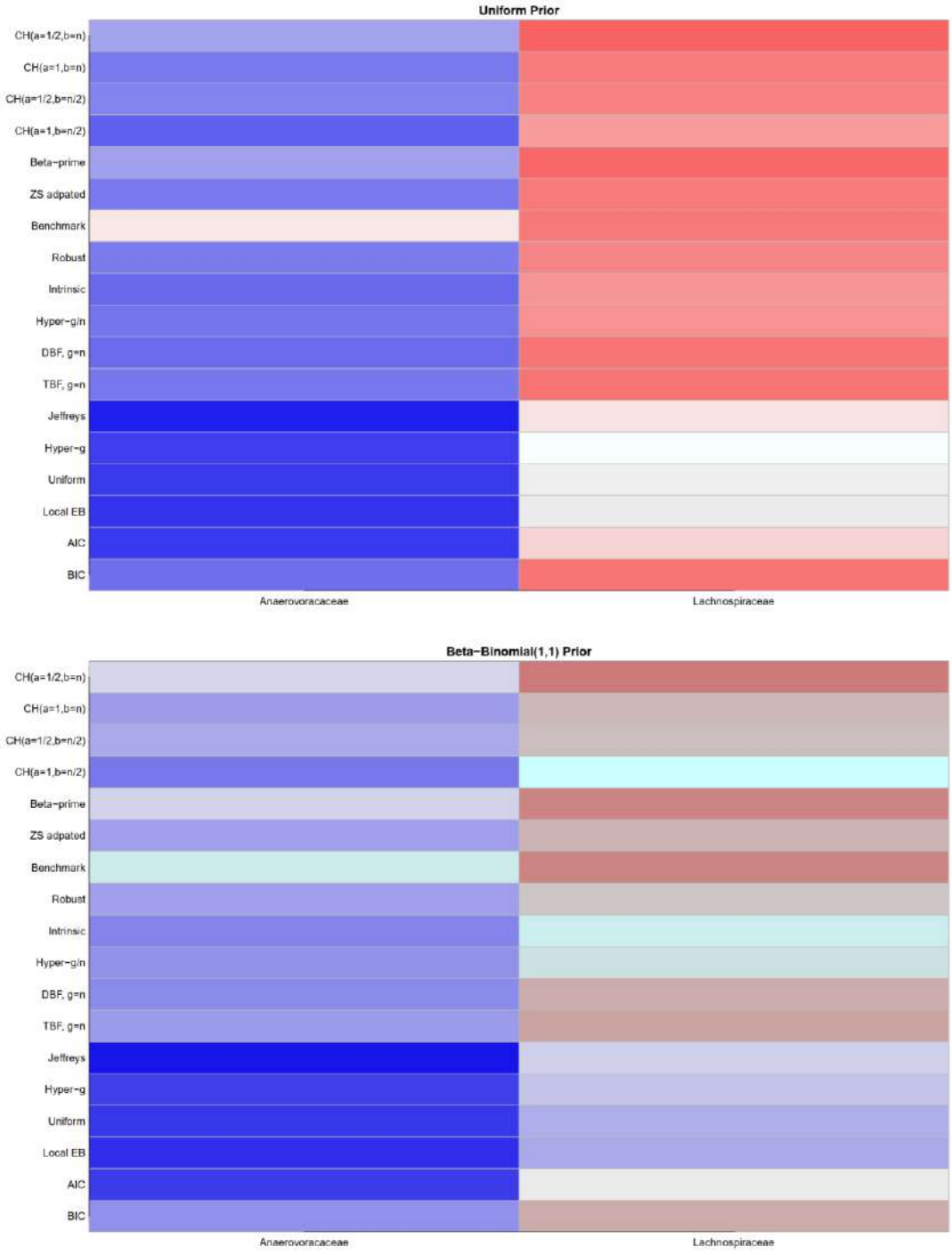


Figure F-D-4 –Marginal posterior inclusion probabilities for the **Family** for **Stool** with **Rank Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

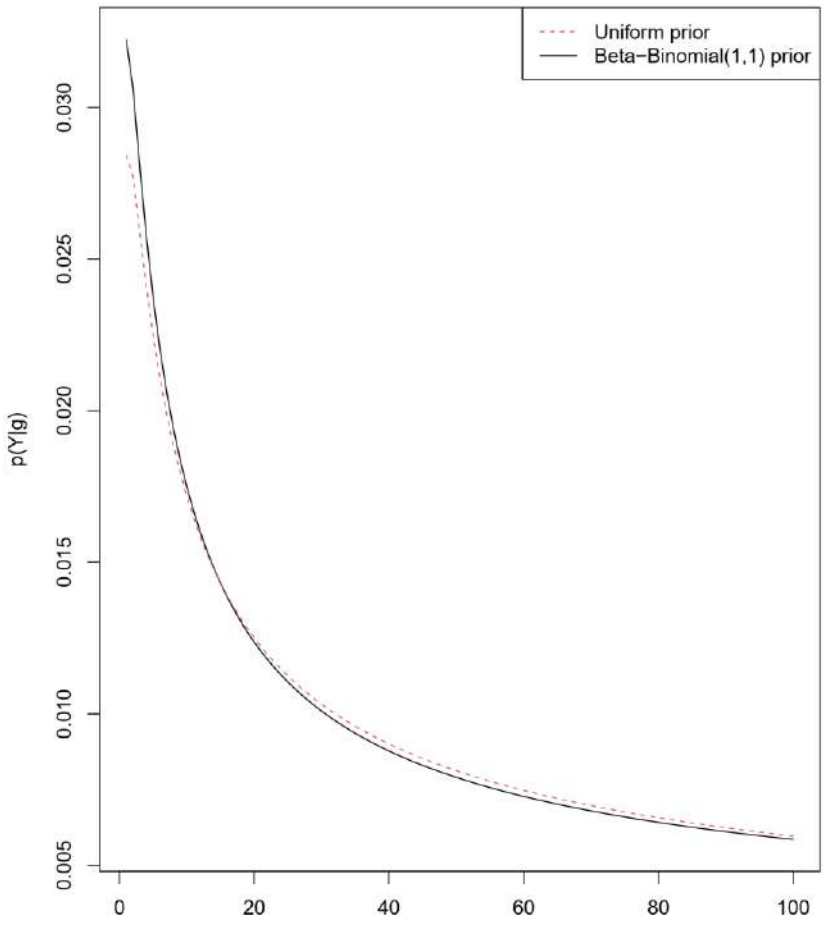


Figure F-D-5: Marginal likelihood of *g* for the **Family** for **Stool** with **Rank Unit** Normalization method.

Table F-D-4: Prediction accuracy for the **Stool (Phylum) Rank Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.6654	0.6668	13.8747	20.2695	0.4687	0.4696	0.1427	0.1437
02_CH(a=1,b=n)	0.6660	0.6679	9.1694	13.8139	0.4657	0.4677	0.1420	0.1429
03_CH(a=1/2,b=n/2)	0.6662	**0.6688**	10.7644	14.4433	0.4661	0.4674	0.1421	0.1430
04_CH(a=1,b=n/2)	0.6663	0.6680	7.8327	18.8767	0.4647	0.4659	0.1415	0.1422
05_Beta-prime	0.6655	0.6675	11.7694	31.6394	0.4674	0.4688	0.1426	0.1436
06_ZS adpated	0.6666	0.6676	10.4470	13.9307	0.4657	0.4678	0.1421	0.1429
07_Robust	0.6676	0.6685	8.6318	12.0511	0.4648	0.4671	0.1418	0.1426
08_Intrinsic	0.6681	0.6685	6.8886	9.7277	0.4647	0.4666	0.1416	0.1423
09_Hyper-g/n	0.6665	0.6675	9.3234	13.8053	0.4651	0.4658	0.1417	0.1424
10_Benchmark	0.6679	0.6680	54.3179	56.9559	0.4696	0.4723	0.1451	0.1463
11_DBF, g=n	0.6664	0.6678	8.7199	11.1646	0.4655	0.4674	0.1419	0.1427
12_TBF, g=n	0.6661	0.6682	8.8480	11.3878	0.4659	0.4668	0.1417	0.1425
13_Jeffreys	0.6667	0.6676	**4.5262**	**5.6595**	0.4573	**0.4582**	**0.1392**	**0.1394**
14_Hyper-g	0.6673	0.6684	8.0023	10.3562	**0.4568**	0.4587	0.1398	0.1403
15_Uniform	0.6674	0.6679	11.6984	12.4500	0.4569	0.4591	0.1402	0.1405
16_Local EB	0.6438	0.6446	17.3519	22.5430	0.4767	0.4790	0.1482	0.1492
17_AIC	**0.6689**	0.6681	4.8354	6.4390	0.4654	0.4692	0.1418	0.1424
18_BIC	0.6677	0.6681	6.2902	8.8140	0.4648	0.4674	0.1417	0.1424

E-Genus

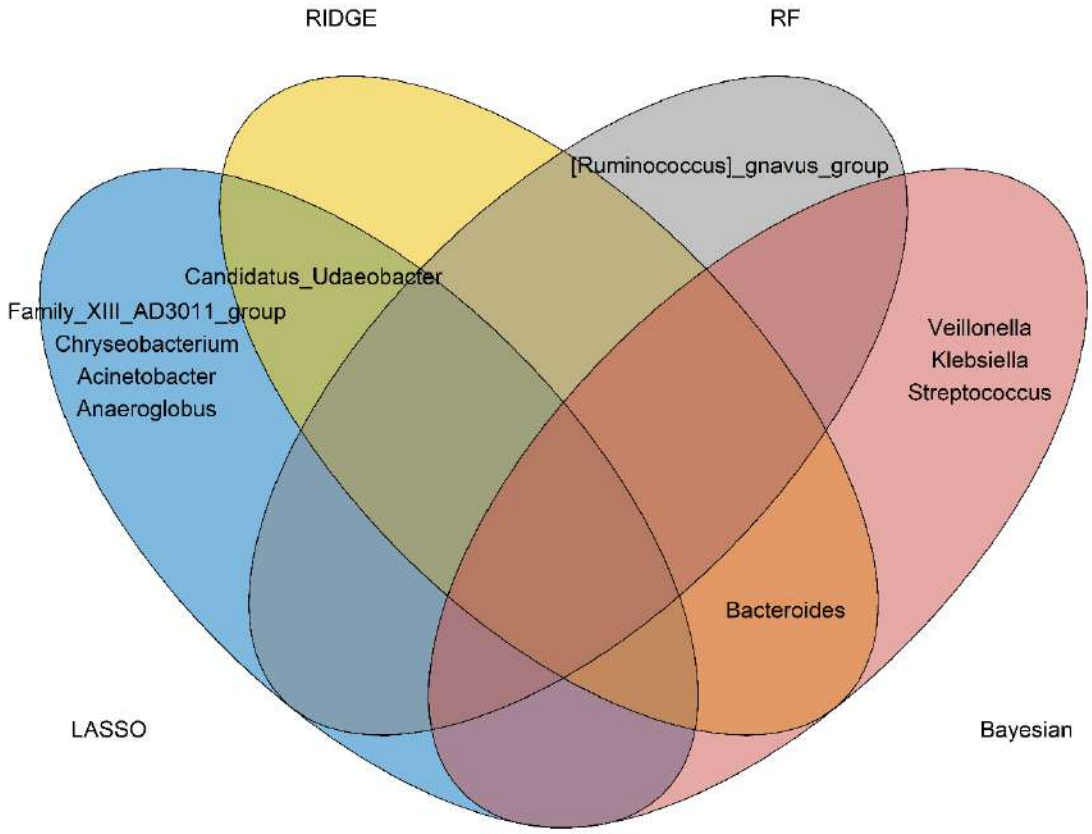


Figure F-E-1 – The Concordance Plot for **Genus** for **Stool** .

Table F-E-1: The Selected method				
Row	**Method**	**Filter**	**Normal**	**AUC**
1	LASSO	PA	RU	0.64
2	RIDGE	PA	RU	0.77
3	RF	CA	RU	0.73
4	**Bayesian**	**VA**	**LU**	**0.78**

Table F-E-2: The common feature names	
Common Features for BAS	
Candidatus_Udaeobacter	Bacteroides

BAS

Log Unit

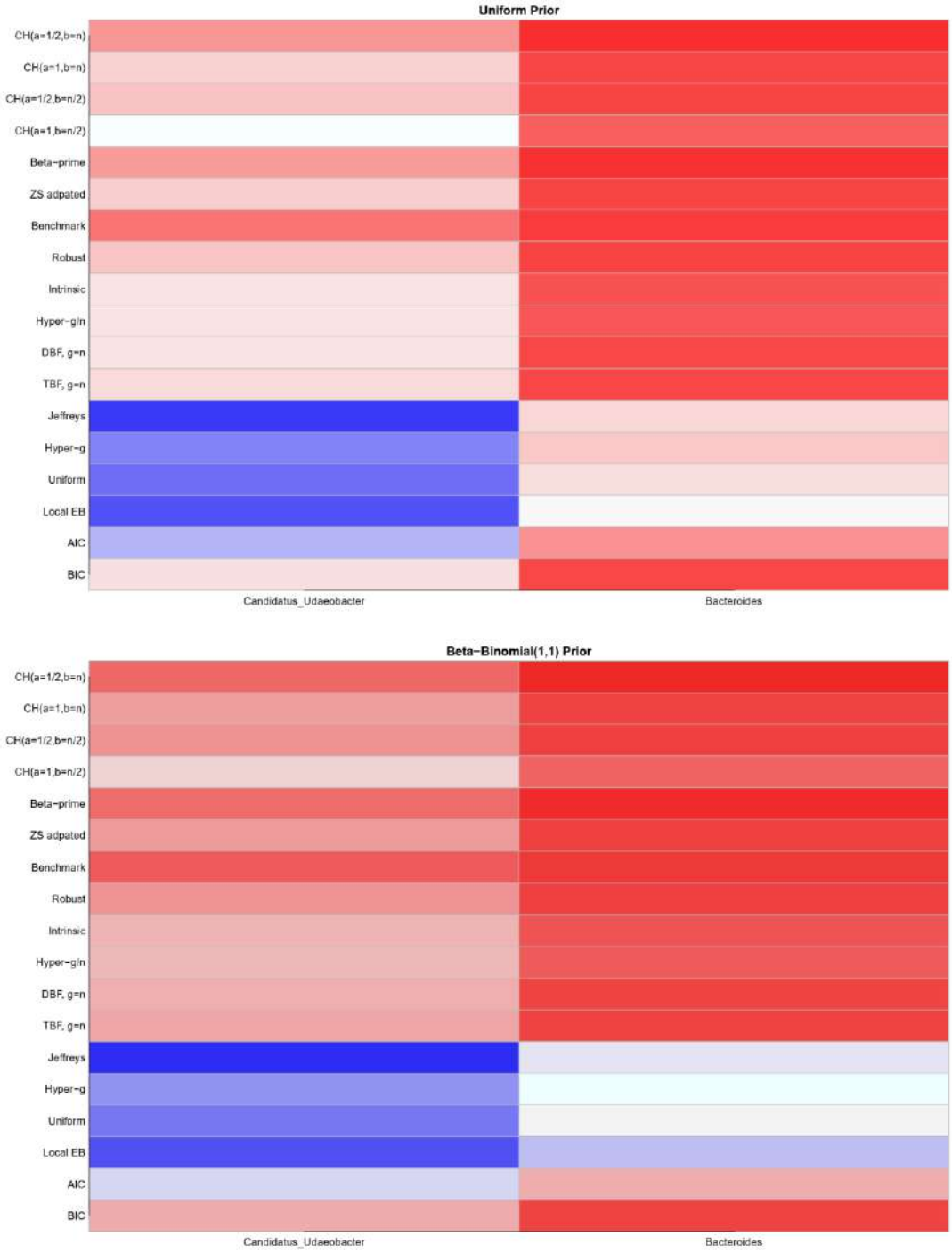


Figure F-E-2 –Marginal posterior inclusion probabilities for the **Genus** for **Stool** with **Log Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

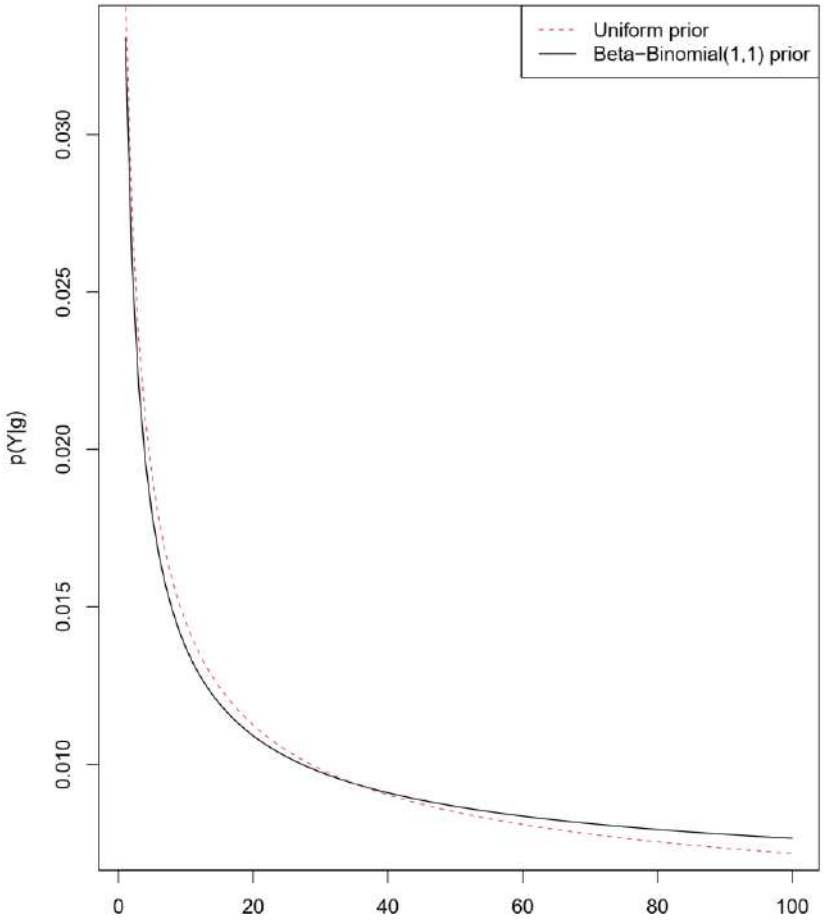


Figure F-E-3: Marginal likelihood of *g* for the **Genus** for **Stool** with **Log Unit** Normalization method.

Table F-E-3: Prediction accuracy for the **Stool (Phylum) Log Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.5486	0.5434	32.7448	10.5596	0.4730	0.4719	0.1445	0.1446
02_CH(a=1,b=n)	0.5444	0.5470	28.6285	9.7689	0.4729	0.4749	0.1449	0.1452
03_CH(a=1/2,b=n/2)	0.5426	0.5450	28.7718	8.5588	0.4736	0.4741	0.1448	0.1450
04_CH(a=1,b=n/2)	0.5467	0.5461	9.5511	7.0875	0.4746	0.4771	0.1451	0.1456
05_Beta-prime	0.5462	0.5476	14.6306	11.5962	0.4733	0.4721	0.1445	0.1447
06_ZS adpated	**0.5518**	0.5471	13.1945	7.6797	0.4725	0.4745	0.1448	0.1451
07_Robust	0.5472	0.5460	7.5702	6.9459	0.4764	0.4789	0.1455	0.1459
08_Intrinsic	0.5442	0.5449	-11.2496	-13.5506	0.4782	0.4805	0.1457	0.1462
09_Hyper-g/n	0.5463	0.5448	28.1393	-11.9558	0.4742	0.4745	0.1447	0.1450
10_Benchmark	0.5511	**0.5488**	24.0406	22.3498	0.4705	0.4735	0.1449	0.1462
11_DBF, g=n	0.5478	0.5461	11.4952	8.8529	0.4735	0.4756	0.1450	0.1453
12_TBF, g=n	0.5502	0.5472	11.9853	29.4328	0.4733	0.4738	0.1448	0.1450
13_Jeffreys	0.5461	0.5418	8.4076	3.7103	0.4773	0.4804	0.1447	0.1453
14_Hyper-g	0.5430	0.5416	10.3609	6.3364	0.4710	0.4728	0.1440	0.1443
15_Uniform	0.5431	0.5416	9.0780	7.0988	**0.4668**	**0.4684**	**0.1435**	**0.1438**
16_Local EB	0.5008	0.5001	-72.3548	98.9171	0.4940	0.4941	0.1543	0.1543
17_AIC	0.5476	0.5455	**5.7837**	**3.1183**	0.4825	0.4867	0.1467	0.1474
18_BIC	0.5484	0.5460	10.5029	6.2752	0.4775	0.4783	0.1456	0.1460

Rank Unit

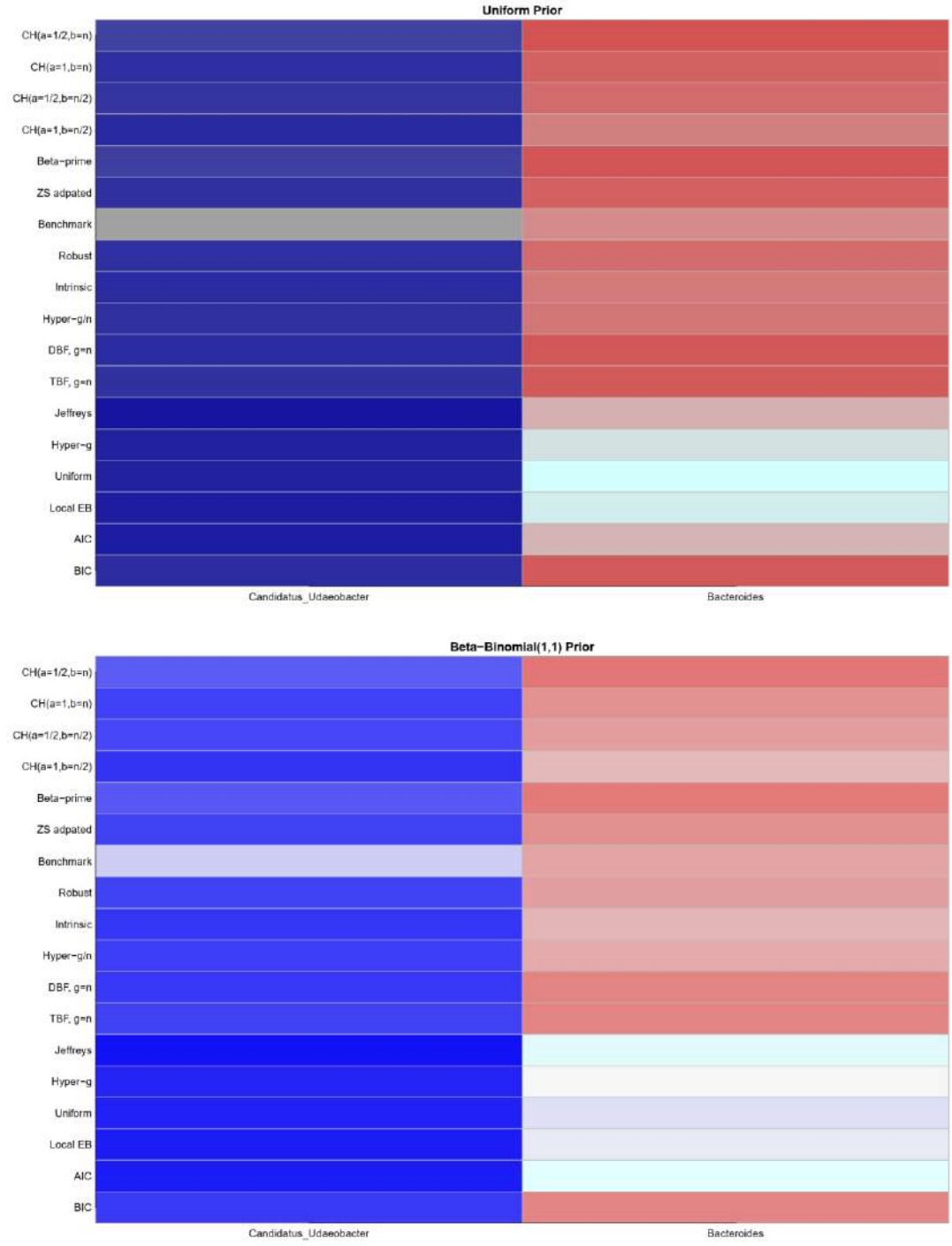


Figure F-E-4 –Marginal posterior inclusion probabilities for the **Genus** for **Stool** with **Rank Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

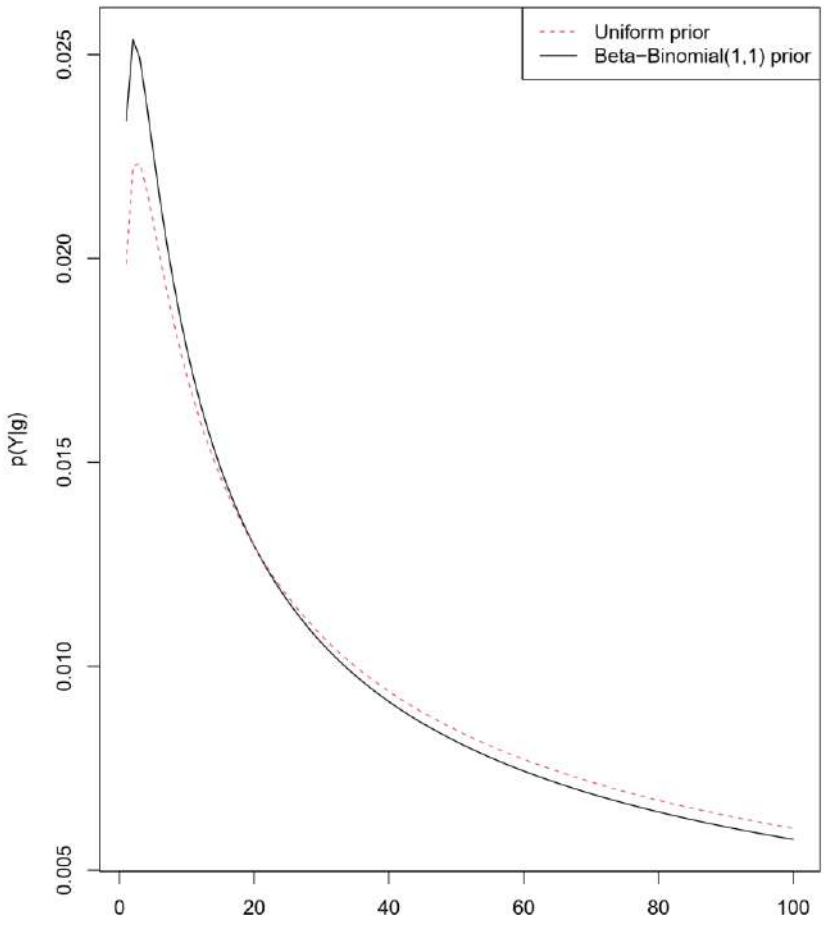


Figure F-E-5: Marginal likelihood of *g* for the **Genus** for **Stool** with **Rank Unit** Normalization method.

Table F-E-4: Prediction accuracy for the **Stool (Phylum) Rank Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.7587	0.7539	6.6450	8.8008	0.4568	0.4620	0.1390	0.1406
02_CH(a=1,b=n)	0.7575	0.7520	5.0122	6.3223	0.4541	0.4598	0.1381	0.1395
03_CH(a=1/2,b=n/2)	0.7566	0.7517	5.4280	7.1196	0.4547	0.4600	0.1384	0.1398
04_CH(a=1,b=n/2)	0.7560	0.7503	4.5462	5.4504	0.4533	0.4583	0.1377	0.1389
05_Beta-prime	0.7580	0.7534	6.2491	8.7275	0.4568	0.4618	0.1390	0.1405
06_ZS adpated	0.7575	0.7524	5.0803	6.7135	0.4548	0.4599	0.1382	0.1396
07_Robust	0.7545	0.7486	4.4251	5.4460	0.4553	0.4589	0.1380	0.1392
08_Intrinsic	0.7540	0.7479	4.1578	5.0633	0.4546	0.4572	0.1377	0.1388
09_Hyper-g/n	0.7578	**0.7541**	5.2007	6.8474	0.4531	0.4570	0.1379	0.1392
10_Benchmark	0.7557	0.7513	22.6274	49.8854	0.4632	0.4657	0.1420	0.1434
11_DBF, g=n	0.7583	0.7536	4.5884	5.6200	0.4536	0.4587	0.1378	0.1391
12_TBF, g=n	0.7589	0.7535	4.7561	5.8788	0.4526	0.4593	0.1377	0.1391
13_Jeffreys	**0.7590**	**0.7541**	3.9546	4.3933	**0.4422**	**0.4449**	**0.1358**	**0.1364**
14_Hyper-g	0.7558	0.7505	6.0813	7.4396	0.4466	0.4504	0.1370	0.1379
15_Uniform	0.7554	0.7497	7.3855	10.4770	0.4482	0.4506	0.1375	0.1383
16_Local EB	0.7413	0.7398	17.3167	30.0022	0.4512	0.4558	0.1407	0.1422
17_AIC	0.7514	0.7463	**3.3902**	**3.9414**	0.4567	0.4600	0.1377	0.1386
18_BIC	0.7555	0.7500	3.9372	4.7640	0.4545	0.4599	0.1376	0.1387

G-Saliva

A-Phylum

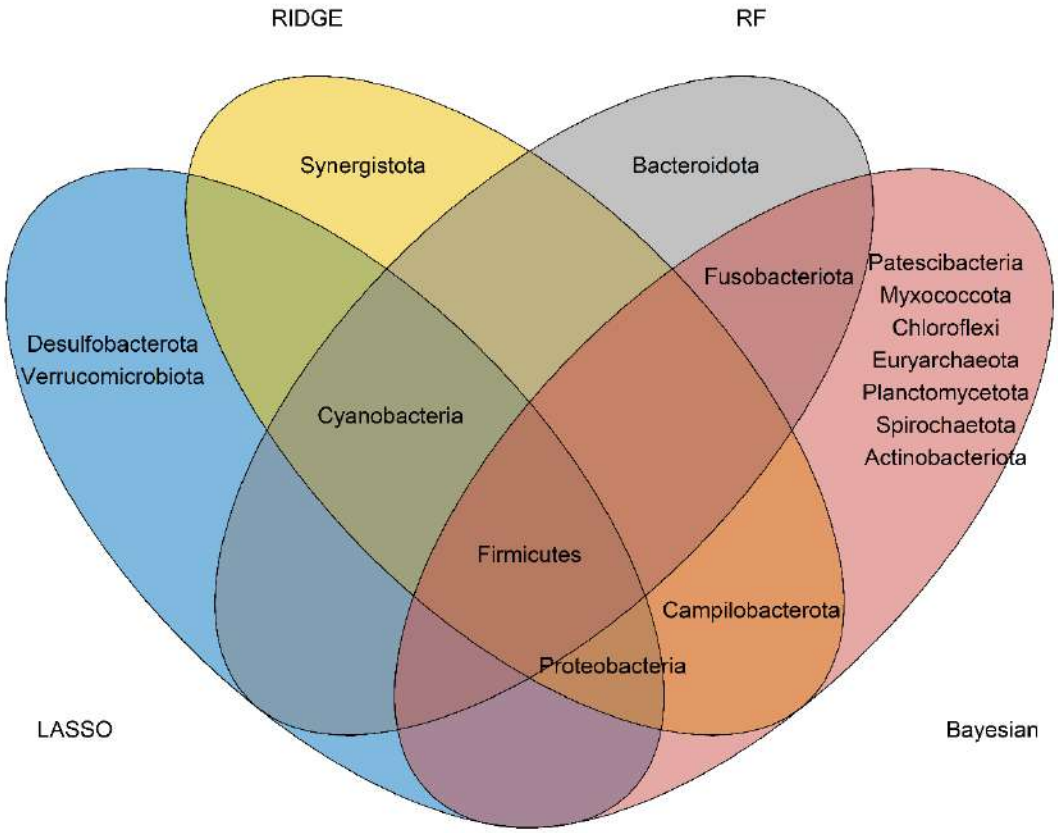


Figure G-A-1 – The Concordance Plot for **Phylum** for **Saliva** .

Table G-A-1: The Selected method				
Row	**Method**	**Filter**	**Normal**	**AUC**
1	LASSO	PR	RU	0.57
2	RIDGE	AB	RU	0.58
3	RF	CA	RU	0.51
4	**Bayesian**	**PR**	**RU**	**0.68**

Table G-A-2: The common feature names			
Common Features for BAS			
Cyanobacteria	Firmicutes	Proteobacteria	Campilobacterota
Fusobacteriota			

BAS

Log Unit

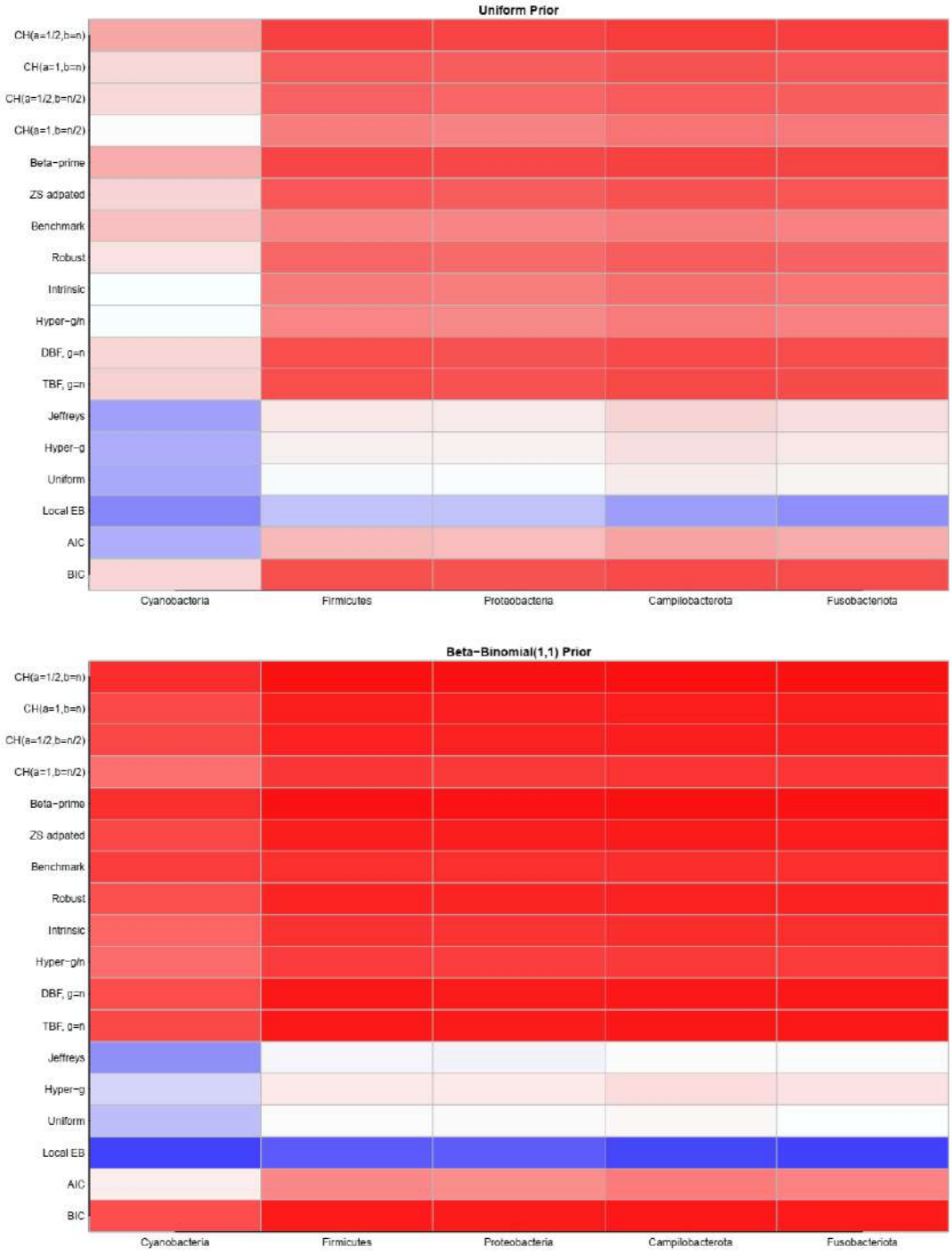


Figure G-A-2 –Marginal posterior inclusion probabilities for the **Phylum** for **Saliva** with **Log Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

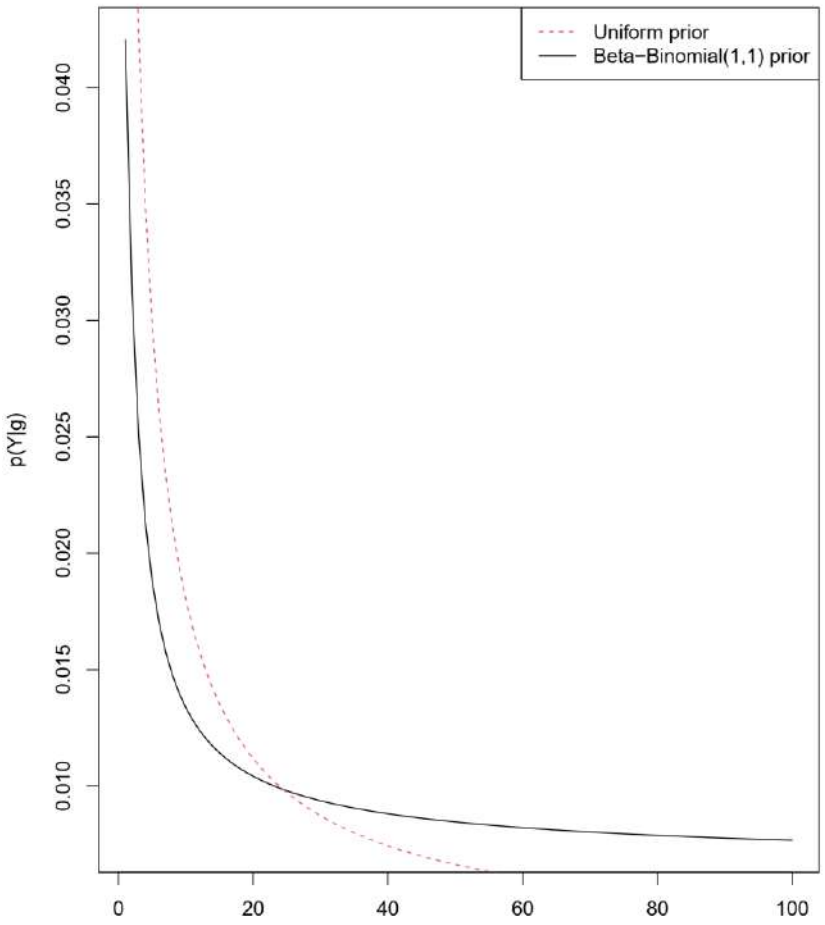


Figure G-A-3: Marginal likelihood of *g* for the **Phylum** for **Saliva** with **Log Unit** Normalization method.

Table G-A-3: Prediction accuracy for the **Saliva (Phylum) Log Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.6304	0.6340	8.6771	34.2235	0.5915	0.5772	0.1866	0.1851
02_CH(a=1,b=n)	0.6300	0.6336	23.6657	17.3424	0.5918	0.5883	0.1872	0.1863
03_CH(a=1/2,b=n/2)	0.6302	0.6338	13.3861	29.2216	0.5920	0.5912	0.1874	0.1870
04_CH(a=1,b=n/2)	0.6296	0.6322	7.5925	15.6097	0.5918	0.5988	0.1879	0.1880
05_Beta-prime	0.6307	0.6350	24.4522	20.5180	0.5914	0.5795	0.1867	0.1855
06_ZS adpated	0.6302	0.6342	13.4588	30.0721	0.5911	0.5876	0.1871	0.1863
07_Robust	0.6297	0.6322	5.0150	14.8715	0.5979	0.6104	0.1893	0.1897
08_Intrinsic	0.6302	0.6312	4.4813	14.1350	0.5971	0.6111	0.1895	0.1902
09_Hyper-g/n	0.6317	0.6355	7.8424	16.5522	0.5886	0.5893	0.1862	0.1864
10_Benchmark	**0.6394**	**0.6435**	15.9503	22.9739	0.5710	**0.5620**	**0.1826**	**0.1827**
11_DBF, g=n	0.6280	0.6301	8.1043	17.9206	0.5872	0.5840	0.1863	0.1852
12_TBF, g=n	0.6283	0.6303	13.5061	18.0719	0.5853	0.5794	0.1856	0.1843
13_Jeffreys	0.6323	0.6331	5.2930	13.3204	0.5906	0.5990	0.1848	0.1859
14_Hyper-g	0.6363	0.6392	8.1117	12.8130	0.5919	0.5919	0.1839	0.1847
15_Uniform	0.6390	0.6410	10.3226	12.9896	0.5835	0.5843	0.1831	0.1839
16_Local EB	0.5814	0.5727	219.5729	213.0072	**0.5633**	0.5705	0.1878	0.1909
17_AIC	0.6304	0.6301	**3.4539**	**6.4716**	0.6021	0.6120	0.1909	0.1914
18_BIC	0.6293	0.6324	13.0144	15.5690	0.5876	0.5898	0.1874	0.1863

Rank Unit

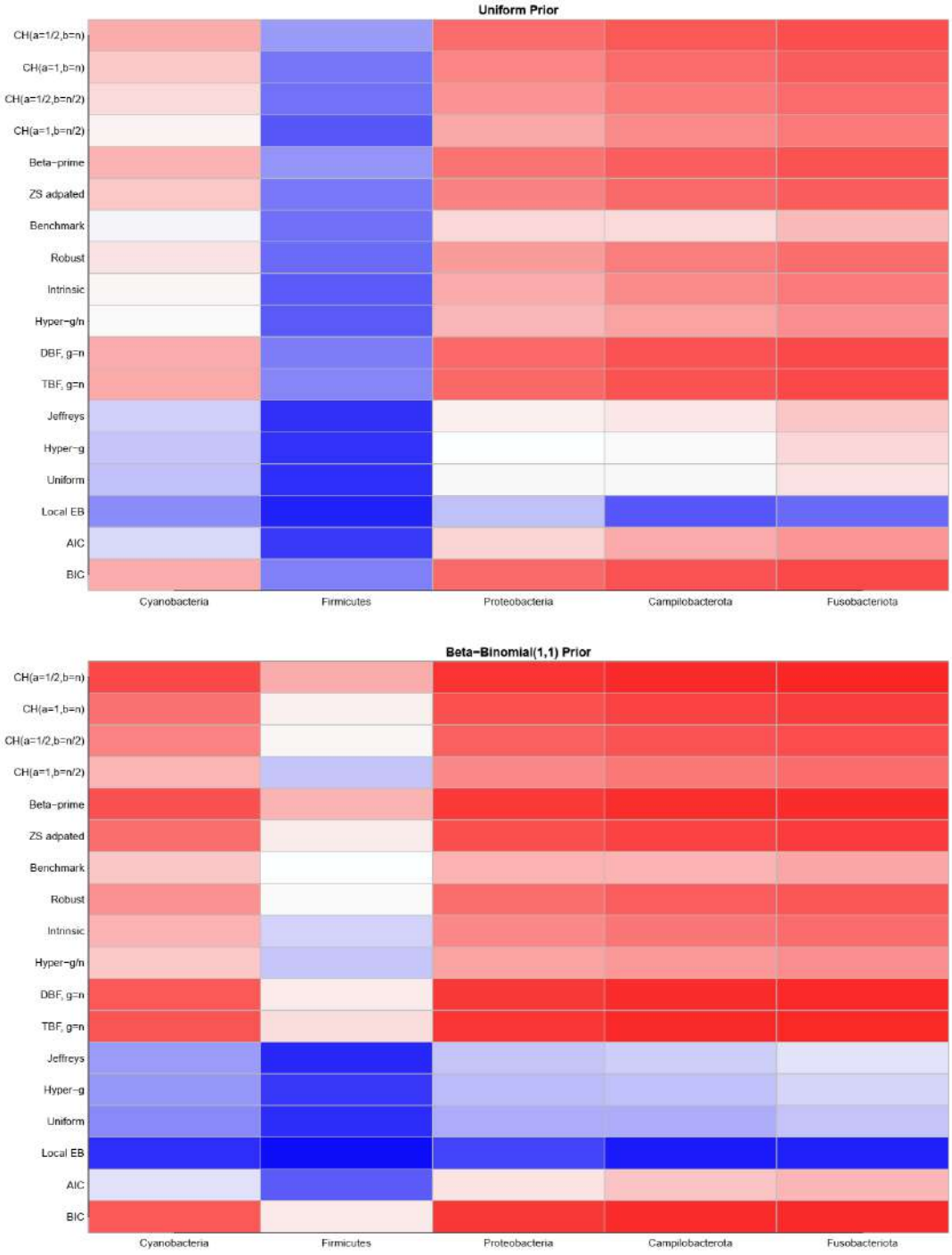


Figure G-A-4 –Marginal posterior inclusion probabilities for the **Phylum** for **Saliva** with **Rank Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

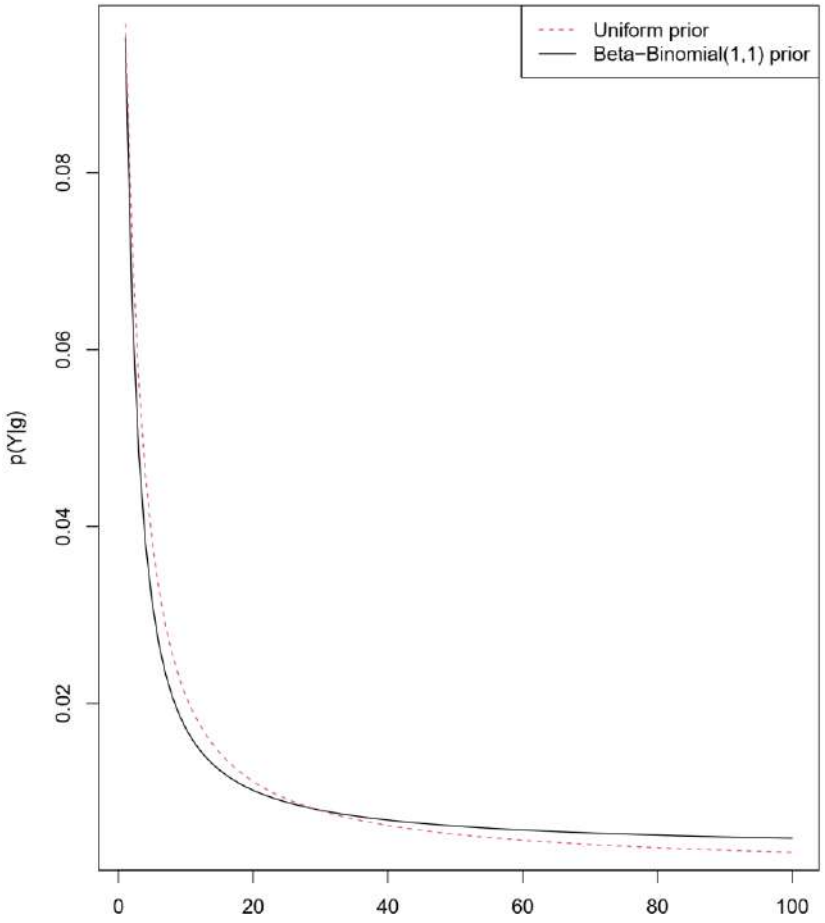


Figure G-A-5: Marginal likelihood of *g* for the **Phylum** for **Saliva** with **Rank Unit** Normalization method.

Table G-A-4: Prediction accuracy for the **Saliva (Phylum) Rank Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.7137	0.7101	2.1349	6.3951	0.5754	0.5804	0.1810	0.1829
02_CH(a=1,b=n)	0.7135	0.7090	1.7594	4.0637	0.5790	0.5821	0.1813	0.1830
03_CH(a=1/2,b=n/2)	0.7138	0.7094	1.6527	3.8268	0.5793	0.5856	0.1814	0.1833
04_CH(a=1,b=n/2)	0.7128	0.7077	1.4355	2.5402	0.5819	0.5906	0.1817	0.1837
05_Beta-prime	0.7141	0.7099	2.1120	6.5858	0.5771	0.5796	0.1811	0.1830
06_ZS adpated	0.7136	0.7095	1.7853	4.1450	0.5787	0.5817	0.1813	0.1830
07_Robust	0.7088	0.7058	1.3056	2.1260	0.5893	0.5991	0.1831	0.1852
08_Intrinsic	0.7077	0.7047	1.2491	1.8262	0.5894	0.5992	0.1829	0.1852
09_Hyper-g/n	0.7137	0.7106	1.7320	2.9540	0.5682	0.5760	0.1794	0.1813
10_Benchmark	0.7126	0.7109	3.7767	9.7627	0.5660	0.5751	0.1766	0.1795
11_DBF, g=n	0.7097	0.7119	1.8283	4.3391	0.5781	0.5746	0.1812	0.1816
12_TBF, g=n	0.7110	0.7097	1.9107	4.7667	0.5789	0.5751	0.1817	0.1816
13_Jeffreys	0.7174	0.7140	1.6656	1.7128	0.5637	0.5715	0.1773	0.1792
14_Hyper-g	**0.7181**	**0.7160**	2.0129	2.2189	**0.5588**	0.5662	0.1762	0.1782
15_Uniform	0.7148	0.7139	2.3274	2.4804	0.5645	**0.5653**	**0.1758**	**0.1776**
16_Local EB	0.6637	0.6531	252.6311	420.8454	0.6129	0.6312	0.2068	0.2136
17_AIC	0.7074	0.7040	**1.1165**	**1.3099**	0.6001	0.6108	0.1847	0.1873
18_BIC	0.7097	0.7115	1.4673	2.8111	0.5874	0.5808	0.1828	0.1824

B-Class

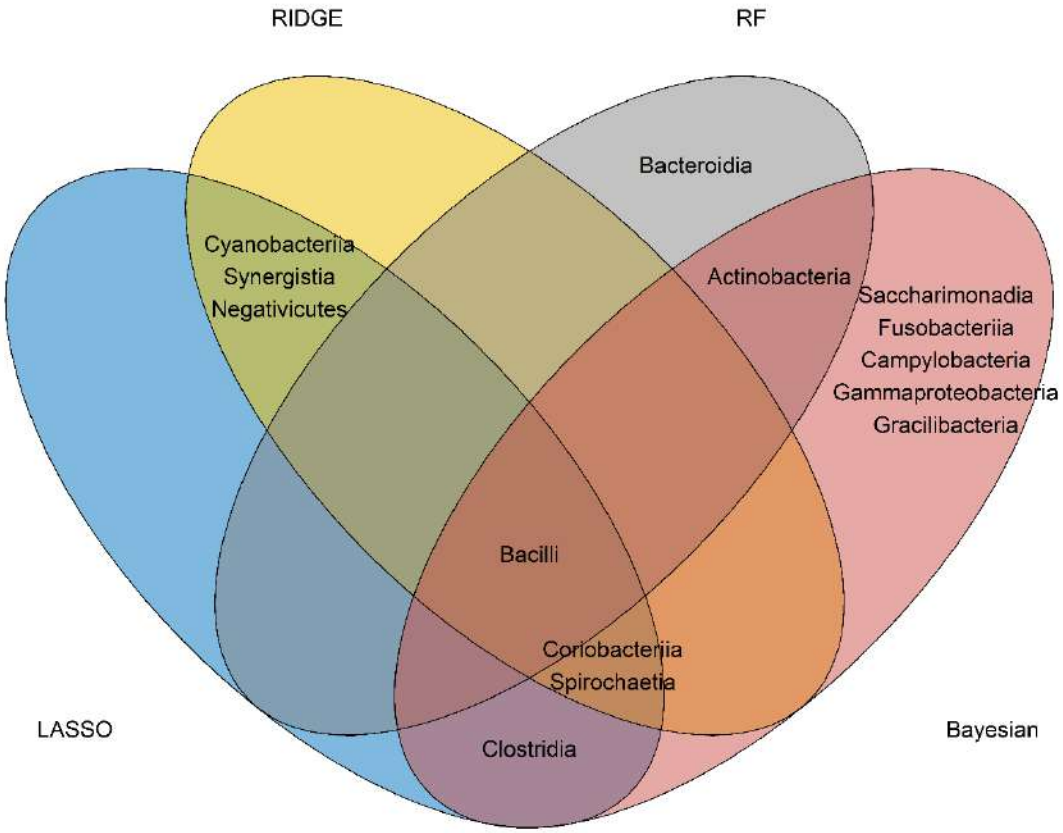


Figure G-B-1 – The Concordance Plot for **Class** for **Saliva** .

Table G-B-1: The Selected method				
Row	**Method**	**Filter**	**Normal**	**AUC**
1	LASSO	CA	RU	0.63
2	RIDGE	AB	RU	0.62
3	RF	VA	LU	0.57
4	**Bayesian**	**AB**	**LU**	**0.67**

Table G-B-2: The common feature names			
Common Features for BAS			
Cyanobacteriia	Synergistia	Negativicutes	Bacilli
Coriobacteriia	Spirochaetia	Clostridia	Actinobacteria

BAS

Log Unit

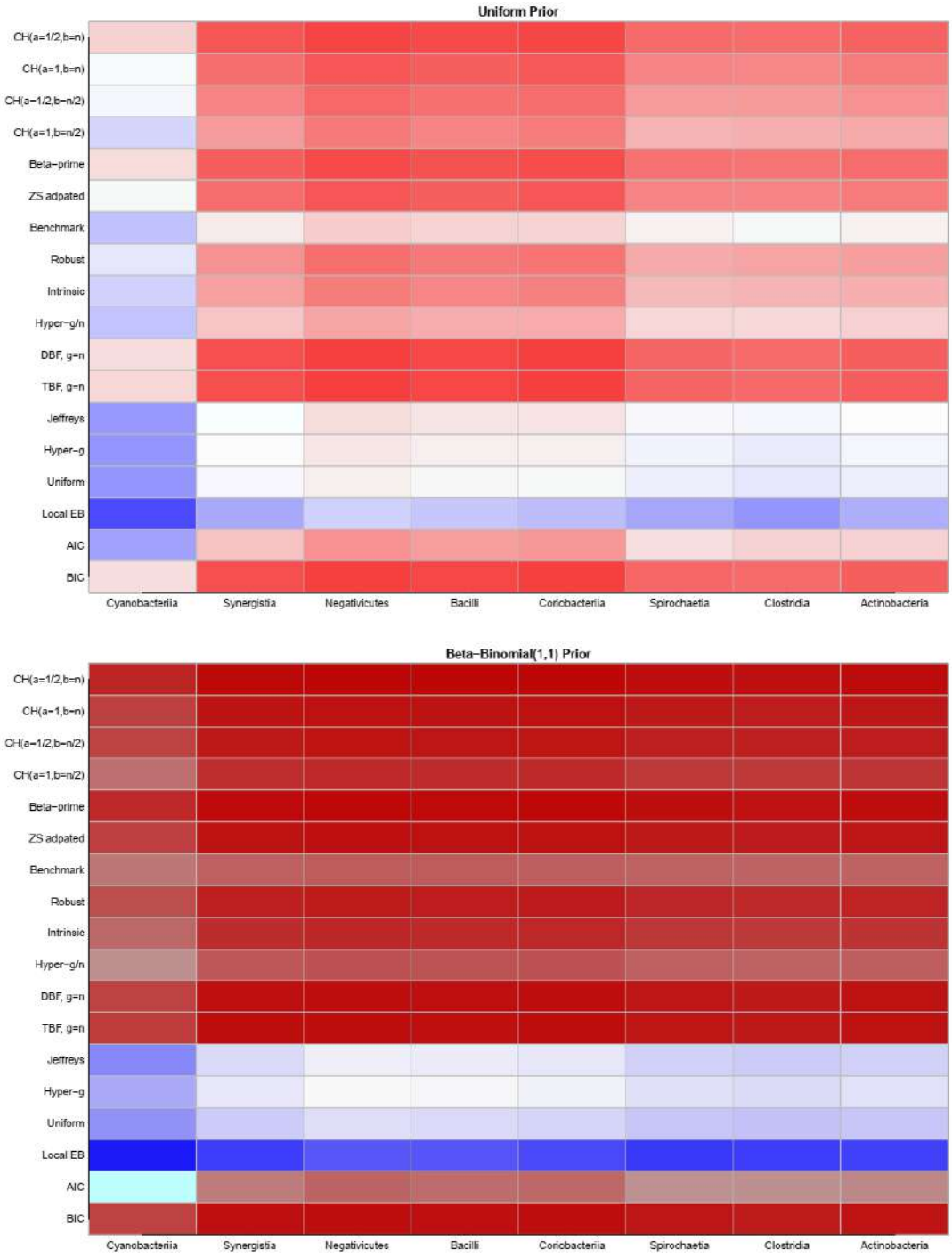


Figure G-B-2 –Marginal posterior inclusion probabilities for the **Class** for **Saliva** with **Log Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

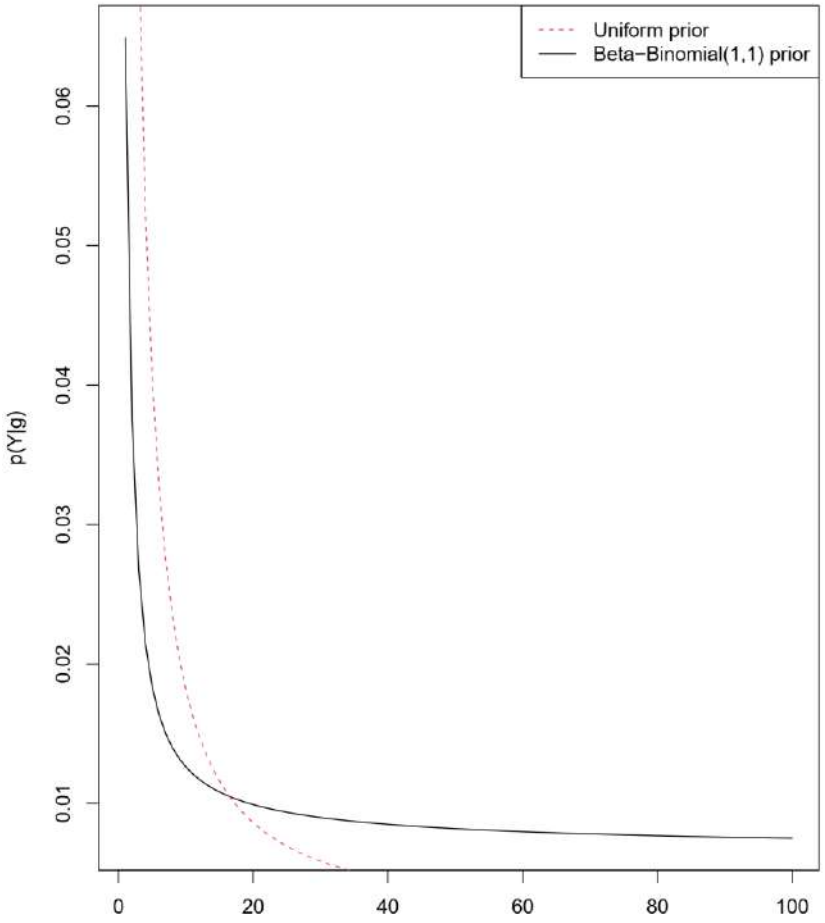


Figure G-B-3: Marginal likelihood of *g* for the **Class** for **Saliva** with **Log Unit** Normalization method.

Table G-B-3: Prediction accuracy for the **Saliva (Phylum) Log Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.6624	0.6650	0.9819	5.7512	0.6711	0.6626	0.2025	0.2016
02_CH(a=1,b=n)	0.6622	0.6650	0.8520	3.3233	0.6718	0.6686	0.2031	0.2024
03_CH(a=1/2,b=n/2)	0.6633	0.6651	0.8120	2.8807	0.6736	0.6815	0.2031	0.2042
04_CH(a=1,b=n/2)	0.6628	0.6624	0.7590	1.8091	0.6757	0.6909	0.2035	0.2051
05_Beta-prime	0.6631	0.6647	0.9371	5.0831	0.6710	0.6671	0.2028	0.2024
06_ZS adpated	0.6621	0.6656	0.8577	3.4091	0.6715	0.6675	0.2031	0.2023
07_Robust	0.6643	0.6661	0.6458	1.0602	0.6888	0.7022	0.2063	0.2084
08_Intrinsic	0.6465	0.6486	0.6935	**1.0359**	0.6818	0.6920	0.2054	0.2073
09_Hyper-g/n	0.6570	0.6618	1.3046	2.2722	0.6662	0.6733	0.2006	0.2014
10_Benchmark	0.6750	**0.6734**	1.4700	3.0196	0.6706	0.6800	0.1995	0.2023
11_DBF, g=n	0.6603	0.6614	**1.0052**	3.8462	0.6702	0.6593	0.2023	**0.2005**
12_TBF, g=n	0.6590	0.6638	1.0568	3.9963	0.6750	**0.6572**	0.2031	0.2009
13_Jeffreys	0.6765	0.6732	0.8146	0.7505	0.6764	0.6907	0.2002	0.2030
14_Hyper-g	**0.6776**	0.6728	1.0124	0.9492	0.6696	0.6827	0.1996	0.2023
15_Uniform	0.6763	0.6726	1.2158	1.0372	0.6705	0.6855	**0.1994**	0.2023
16_Local EB	0.5448	0.5276	276.0288	1225.8204	**0.6620**	0.7162	0.2060	0.2120
17_AIC	0.6608	0.6622	0.6404	0.8855	0.6897	0.6921	0.2062	0.2061
18_BIC	0.6601	0.6648	0.8295	2.4826	0.6713	0.6618	0.2032	0.2007

Rank Unit

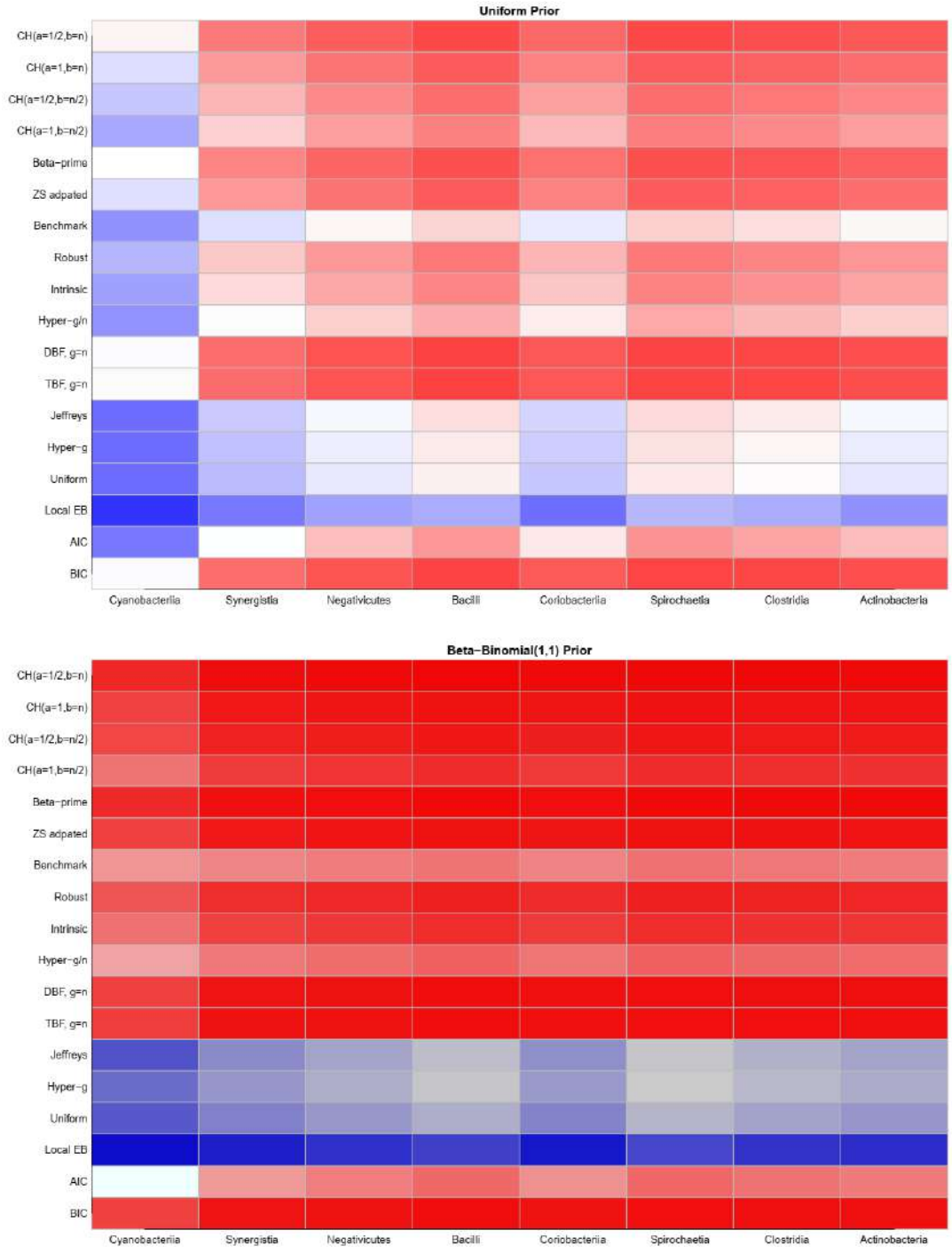


Figure G-B-4 –Marginal posterior inclusion probabilities for the **Class** for **Saliva** with **Rank Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

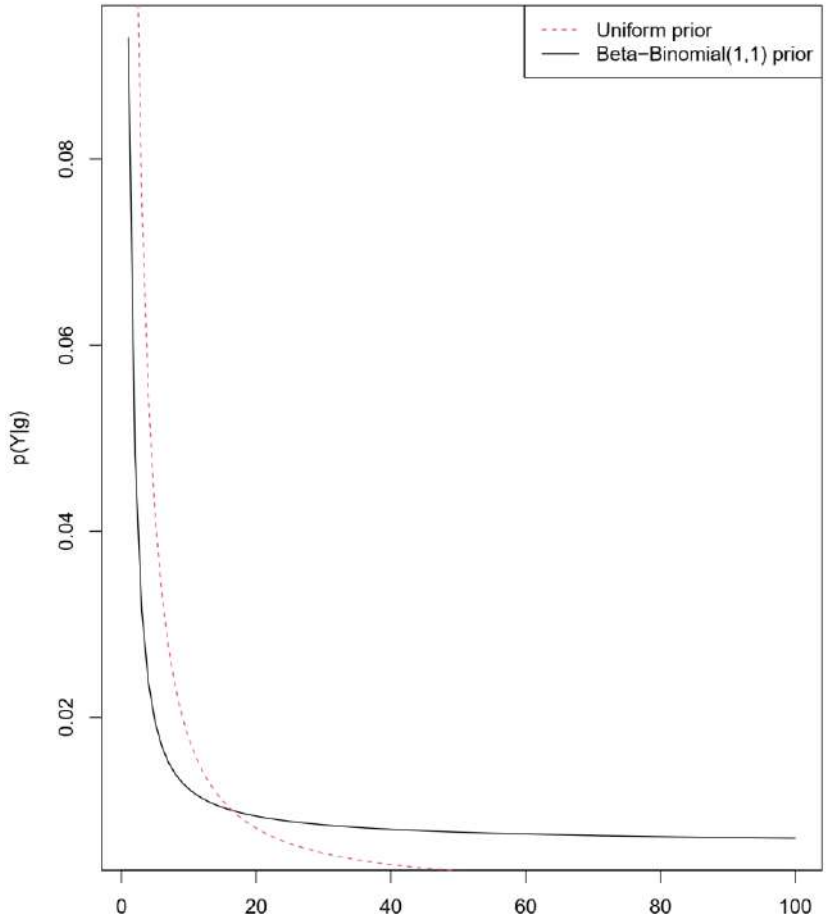


Figure G-B-5: Marginal likelihood of *g* for the **Class** for **Saliva** with **Rank Unit** Normalization method.

Table G-B-4: Prediction accuracy for the **Saliva (Phylum) Rank Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.6879	0.6812	1.1571	5.2066	0.6594	0.6679	0.1977	0.2011
02_CH(a=1,b=n)	0.6876	0.6828	1.0355	3.1847	0.6618	0.6696	0.1982	0.2010
03_CH(a=1/2,b=n/2)	0.6893	0.6886	**0.9997**	2.9408	0.6647	0.6763	0.1987	0.2021
04_CH(a=1,b=n/2)	0.6897	0.6913	1.0409	1.8896	0.6667	0.6778	0.1992	0.2024
05_Beta-prime	0.6887	0.6824	1.1819	4.7439	0.6600	0.6718	0.1980	0.2015
06_ZS adpated	0.6877	0.6817	1.0453	3.2407	0.6616	0.6717	0.1981	0.2013
07_Robust	0.6908	0.6923	0.8185	1.5925	0.6844	0.6931	0.2024	0.2044
08_Intrinsic	0.6637	0.6662	0.8115	1.2343	0.6688	0.6797	0.1998	0.2016
09_Hyper-g/n	0.6696	0.6759	1.5765	2.4573	**0.6536**	0.6657	**0.1956**	0.1984
10_Benchmark	0.6959	0.6903	1.8304	2.7877	0.6617	0.6885	0.1959	0.2001
11_DBF, g=n	0.6852	0.6779	1.1548	3.5320	0.6601	0.6591	0.1986	0.1983
12_TBF, g=n	0.6844	0.6820	1.1449	3.8232	0.6583	0.6546	0.1976	0.1966
13_Jeffreys	**0.6986**	**0.7011**	1.2819	2.5034	0.6681	0.6868	0.1968	0.1997
14_Hyper-g	0.6985	0.6982	2.0620	2.3780	0.6660	0.6838	0.1962	0.1989
15_Uniform	0.6984	0.6962	1.7235	2.2084	0.6601	0.6833	0.1957	0.1987
16_Local EB	0.5424	0.5254	463.1052	875.2205	0.6571	0.7239	0.2021	0.2087
17_AIC	0.6902	0.6940	0.8203	**1.1400**	0.6688	0.6706	0.1996	0.1992
18_BIC	0.6856	0.6857	0.9792	2.4927	0.6580	**0.6496**	0.1975	**0.1961**

C-Order

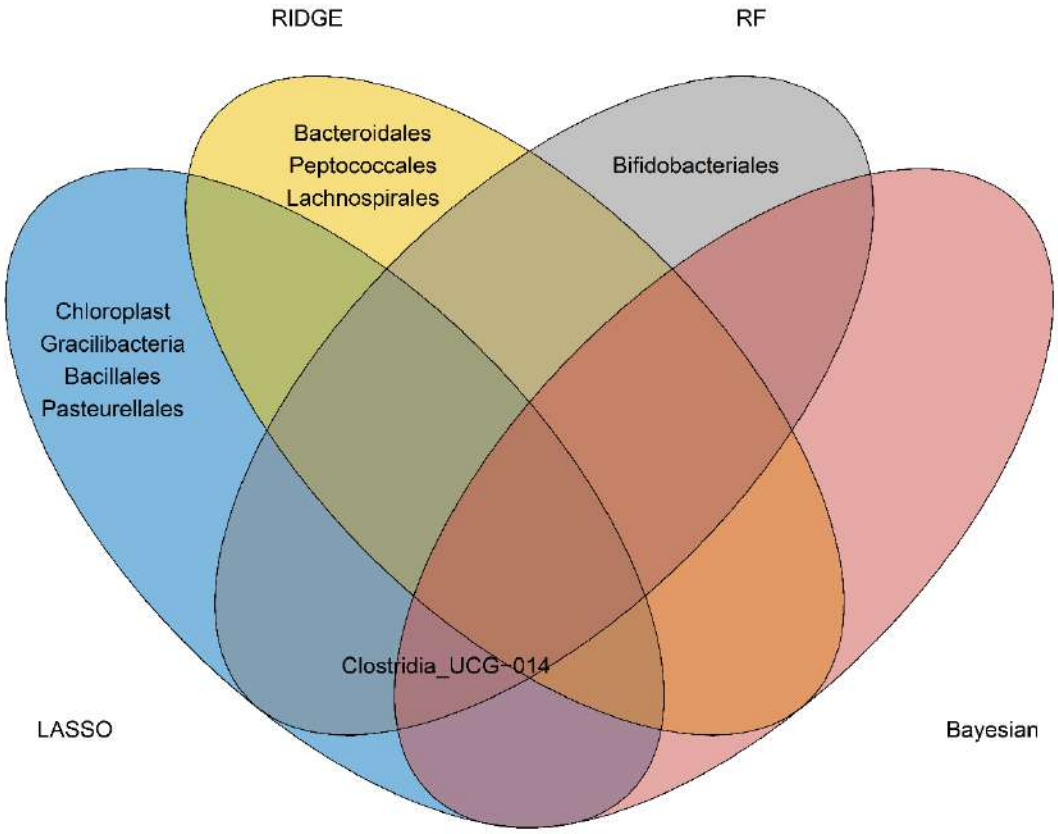


Figure G-C-1 – The Concordance Plot for **Order** for **Saliva** .

Table G-C-1: The Selected method				
Row	**Method**	**Filter**	**Normal**	**AUC**
1	LASSO	AB	LU	0.65
2	RIDGE	AB	RU	0.68
3	RF	PR	LU	0.66
4	**Bayesian**	**PA**	**LU**	**0.7**

Table G-C-2: The common feature names
Common Features for BAS
Clostridia_UCG−014

BAS

Log Unit

It has only one variable remained. Therefore, heatmap is not presented.

Table G-C-3: Prediction accuracy for the **Saliva (Phylum) Log Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.7905	0.7905	19.8486	19.8486	0.4888	0.4888	0.1612	0.1612
02_CH(a=1,b=n)	0.7900	0.7900	15.5464	15.5464	0.4844	0.4844	0.1599	0.1599
03_CH(a=1/2,b=n/2)	0.7901	0.7901	18.0964	18.0964	0.4855	0.4855	0.1603	0.1603
04_CH(a=1,b=n/2)	0.7905	0.7905	14.6602	14.6602	0.4816	0.4816	0.1591	0.1591
05_Beta-prime	0.7903	0.7903	20.4733	20.4733	0.4885	0.4885	0.1612	0.1612
06_ZS adpated	0.7901	0.7901	17.2747	17.2747	0.4846	0.4846	0.1599	0.1599
07_Robust	0.7901	0.7901	15.9120	15.9120	0.4832	0.4832	0.1594	0.1594
08_Intrinsic	0.7903	0.7903	14.1453	14.1453	0.4814	0.4814	0.1589	0.1589
09_Hyper-g/n	0.7901	0.7901	17.5264	17.5264	0.4840	0.4840	0.1598	0.1598
10_Benchmark	0.7903	0.7903	77.2401	77.2401	0.5055	0.5055	0.1668	0.1668
11_DBF, g=n	0.7890	0.7890	15.7964	15.7964	0.4834	0.4834	0.1594	0.1594
12_TBF, g=n	0.7892	0.7892	14.8106	14.8106	0.4836	0.4836	0.1595	0.1595
13_Jeffreys	0.7911	0.7911	12.9279	12.9279	**0.4705**	**0.4705**	**0.1557**	**0.1557**
14_Hyper-g	0.7919	0.7919	17.6325	17.6325	0.4766	0.4766	0.1581	0.1581
15_Uniform	**0.7936**	**0.7936**	21.1524	21.1524	0.4770	0.4770	0.1585	0.1585
16_Local EB	0.7653	0.7653	20.5021	20.5021	0.4870	0.4870	0.1614	0.1614
17_AIC	0.7889	0.7889	**11.8170**	**11.8170**	0.4798	0.4798	0.1579	0.1579
18_BIC	0.7890	0.7890	13.2494	13.2494	0.4824	0.4824	0.1589	0.1589

Rank Unit

It has only one variable remained. Therefore, heatmap is not presented.

Table G-C-4: Prediction accuracy for the **Saliva (Phylum) Rank Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.7491	0.7491	20.8123	20.8123	0.5123	0.5123	0.1703	0.1703
02_CH(a=1,b=n)	0.7499	0.7499	8.4624	8.4624	0.5084	0.5084	0.1692	0.1692
03_CH(a=1/2,b=n/2)	0.7499	0.7499	9.1656	9.1656	0.5092	0.5092	0.1695	0.1695
04_CH(a=1,b=n/2)	0.7499	0.7499	6.4000	6.4000	0.5055	0.5055	0.1685	0.1685
05_Beta-prime	0.7489	0.7489	20.6417	20.6417	0.5121	0.5121	0.1703	0.1703
06_ZS adpated	0.7500	0.7500	8.6073	8.6073	0.5086	0.5086	0.1693	0.1693
07_Robust	0.7498	0.7498	7.3607	7.3607	0.5072	0.5072	0.1689	0.1689
08_Intrinsic	0.7495	0.7495	6.0040	6.0040	0.5054	0.5054	0.1684	0.1684
09_Hyper-g/n	0.7499	0.7499	8.5578	8.5578	0.5077	0.5077	0.1690	0.1690
10_Benchmark	0.7466	0.7466	58.0230	58.0230	0.5214	0.5214	0.1732	0.1732
11_DBF, g=n	0.7497	0.7497	6.6597	6.6597	0.5075	0.5075	0.1689	0.1689
12_TBF, g=n	0.7498	0.7498	6.7734	6.7734	0.5075	0.5075	0.1689	0.1689
13_Jeffreys	0.7504	0.7504	**2.3629**	**2.3629**	**0.4918**	**0.4918**	**0.1646**	**0.1646**
14_Hyper-g	0.7512	0.7512	5.5900	5.5900	0.4970	0.4970	0.1661	0.1661
15_Uniform	**0.7519**	**0.7519**	6.1722	6.1722	0.4965	0.4965	0.1659	0.1659
16_Local EB	0.7262	0.7262	6.9513	6.9513	0.5074	0.5074	0.1689	0.1689
17_AIC	0.7494	0.7494	3.4872	3.4872	0.5036	0.5036	0.1676	0.1676
18_BIC	0.7492	0.7492	4.8274	4.8274	0.5063	0.5063	0.1684	0.1684

D-Family

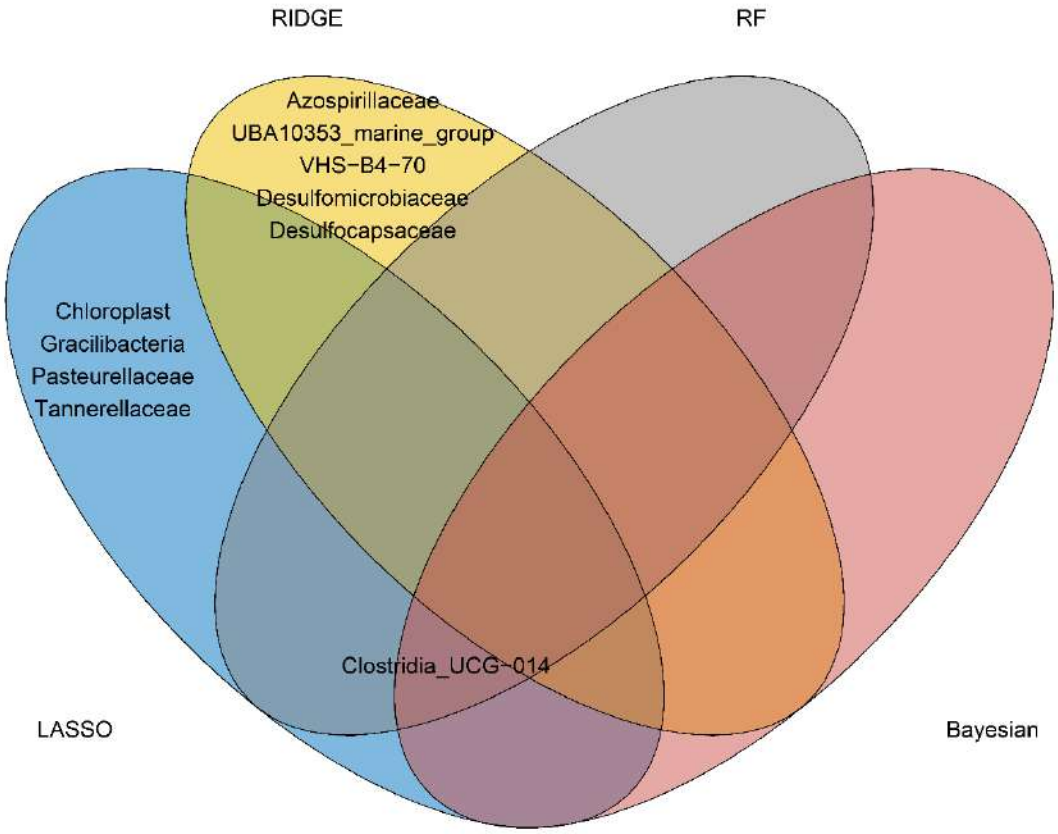


Figure G-D-1 – The Concordance Plot for **Family** for **Saliva** .

Table G-D-1: The Selected method				
Row	**Method**	**Filter**	**Normal**	**AUC**
1	**LASSO**	**PR**	**LU**	**0.72**
2	RIDGE	PA	LU	0.62
3	RF	PA	LU	0.57
4	Bayesian	CA	LU	0.7

Table G-D-2: The common feature names
Common Features for BAS
Clostridia_UCG−014

BAS

Log Unit

It has only one variable remained. Therefore, heatmap is not presented.

Table G-D-3: Prediction accuracy for the **Saliva (Phylum) Log Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.7905	0.7905	18.6963	18.6963	0.4887	0.4887	0.1612	0.1612
02_CH(a=1,b=n)	0.7900	0.7900	15.5405	15.5405	0.4844	0.4844	0.1599	0.1599
03_CH(a=1/2,b=n/2)	0.7901	0.7901	18.0962	18.0962	0.4855	0.4855	0.1603	0.1603
04_CH(a=1,b=n/2)	0.7905	0.7905	14.6600	14.6600	0.4816	0.4816	0.1591	0.1591
05_Beta-prime	0.7903	0.7903	20.4711	20.4711	0.4885	0.4885	0.1612	0.1612
06_ZS adpated	0.7901	0.7901	17.2758	17.2758	0.4846	0.4846	0.1599	0.1599
07_Robust	0.7901	0.7901	15.9133	15.9133	0.4832	0.4832	0.1594	0.1594
08_Intrinsic	0.7903	0.7903	14.1418	14.1418	0.4814	0.4814	0.1589	0.1589
09_Hyper-g/n	0.7901	0.7901	17.5244	17.5244	0.4840	0.4840	0.1598	0.1598
10_Benchmark	0.7903	0.7903	77.2410	77.2410	0.5055	0.5055	0.1668	0.1668
11_DBF, g=n	0.7890	0.7890	15.7959	15.7959	0.4834	0.4834	0.1594	0.1594
12_TBF, g=n	0.7892	0.7892	15.8733	15.8733	0.4836	0.4836	0.1595	0.1595
13_Jeffreys	0.7912	0.7912	12.9286	12.9286	**0.4705**	**0.4705**	**0.1557**	**0.1557**
14_Hyper-g	0.7919	0.7919	19.5998	19.5998	0.4766	0.4766	0.1581	0.1581
15_Uniform	**0.7935**	**0.7935**	21.1528	21.1528	0.4770	0.4770	0.1585	0.1585
16_Local EB	0.7654	0.7654	20.5028	20.5028	0.4869	0.4869	0.1614	0.1614
17_AIC	0.7889	0.7889	**11.8174**	**11.8174**	0.4798	0.4798	0.1579	0.1579
18_BIC	0.7891	0.7891	14.4280	14.4280	0.4824	0.4824	0.1589	0.1589

Rank Unit

It has only one variable remained. Therefore, heatmap is not presented.

Table G-D-4: Prediction accuracy for the **Saliva (Phylum) Rank Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.7816	0.7816	6.2450	6.2450	0.4926	0.4926	0.1642	0.1642
02_CH(a=1,b=n)	0.7810	0.7810	4.4602	4.4602	0.4884	0.4884	0.1629	0.1629
03_CH(a=1/2,b=n/2)	0.7816	0.7816	5.1728	5.1728	0.4893	0.4893	0.1632	0.1632
04_CH(a=1,b=n/2)	0.7819	0.7819	3.9333	3.9333	0.4852	0.4852	0.1620	0.1620
05_Beta-prime	0.7814	0.7814	6.6725	6.6725	0.4924	0.4924	0.1641	0.1641
06_ZS adpated	0.7812	0.7812	4.4934	4.4934	0.4886	0.4886	0.1629	0.1629
07_Robust	0.7812	0.7812	3.9856	3.9856	0.4870	0.4870	0.1624	0.1624
08_Intrinsic	0.7816	0.7816	3.5100	3.5100	0.4851	0.4851	0.1618	0.1618
09_Hyper-g/n	0.7816	0.7816	4.6247	4.6247	0.4870	0.4870	0.1626	0.1626
10_Benchmark	0.7818	0.7818	29.4519	29.4519	0.5039	0.5039	0.1681	0.1681
11_DBF, g=n	0.7818	0.7818	3.8483	3.8483	0.4868	0.4868	0.1624	0.1624
12_TBF, g=n	0.7819	0.7819	3.9739	3.9739	0.4869	0.4869	0.1625	0.1625
13_Jeffreys	0.7838	0.7838	**2.0563**	**2.0563**	**0.4715**	**0.4715**	**0.1580**	**0.1580**
14_Hyper-g	**0.7839**	**0.7839**	4.0770	4.0770	0.4763	0.4763	0.1597	0.1597
15_Uniform	0.7837	0.7837	4.6657	4.6657	0.4763	0.4763	0.1596	0.1596
16_Local EB	0.7619	0.7619	4.8348	4.8348	0.4854	0.4854	0.1623	0.1623
17_AIC	0.7820	0.7820	2.4289	2.4289	0.4828	0.4828	0.1610	0.1610
18_BIC	0.7813	0.7813	3.2189	3.2189	0.4855	0.4855	0.1619	0.1619

E-Genus

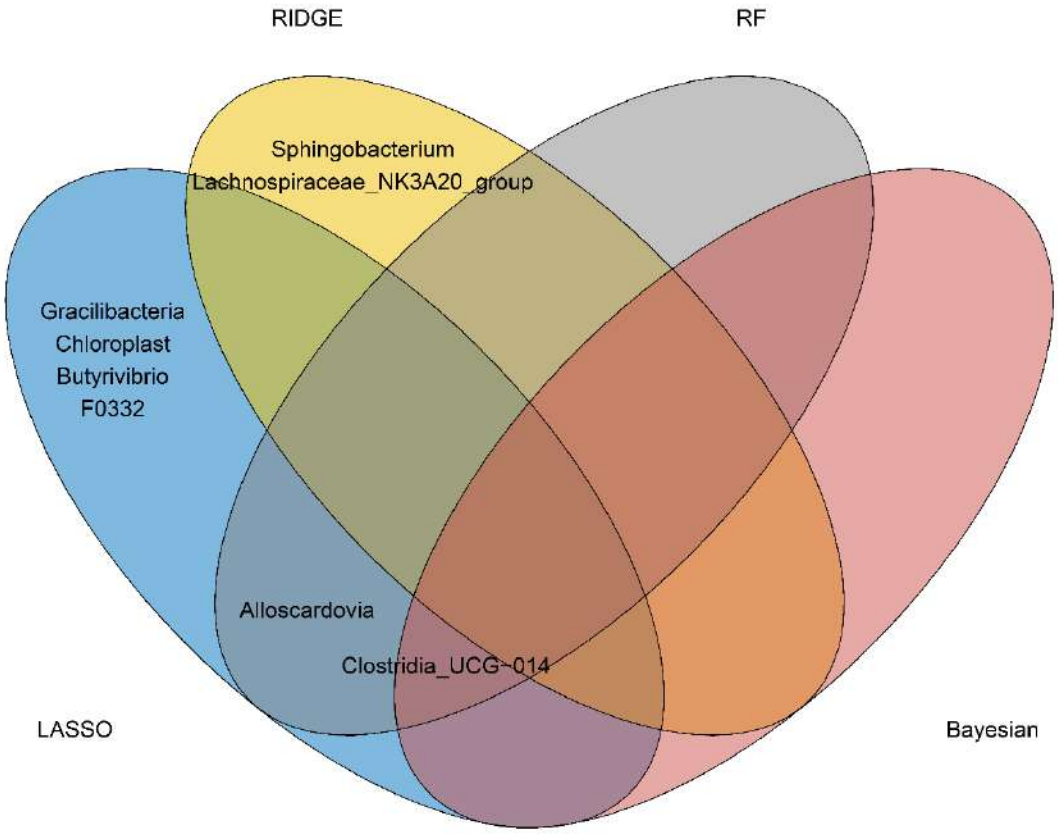


Figure G-E-1 – The Concordance Plot for **Genus** for **Saliva** .

Table G-E-1: The Selected method				
Row	**Method**	**Filter**	**Normal**	**AUC**
1	LASSO	AB	RU	0.62
2	RIDGE	CA	LU	0.59
3	RF	CA	LU	0.52
4	**Bayesian**	**CA**	**LU**	**0.69**

Table G-E-2: The common feature names	
Common Features for BAS	
Alloscardovia	Clostridia_UCG−014

BAS

Log Unit

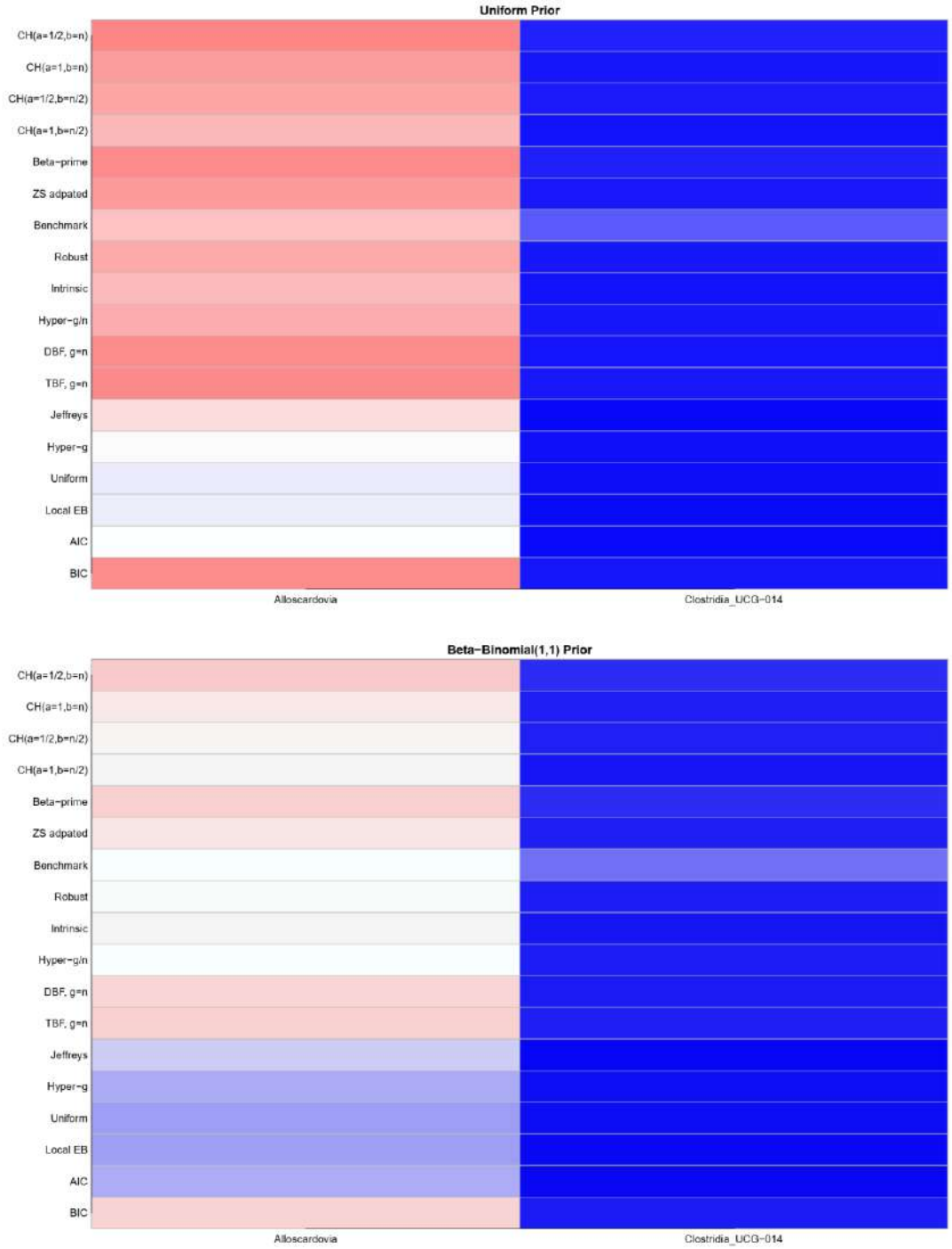


Figure G-E-2 –Marginal posterior inclusion probabilities for the **Genus** for **Saliva** with **Log Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

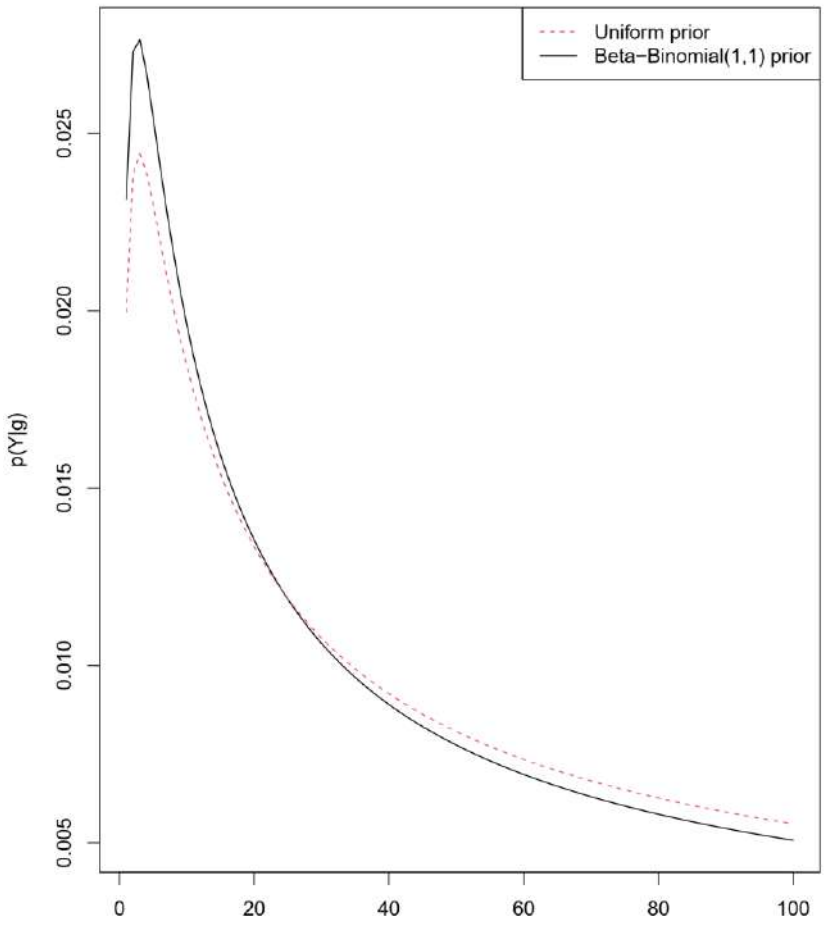


Figure G-E-3: Marginal likelihood of *g* for the **Genus** for **Saliva** with **Log Unit** Normalization method.

Table G-E-3: Prediction accuracy for the **Saliva (Phylum) Log Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.7671	0.7662	7.8023	9.4633	0.4985	0.5023	0.1585	0.1596
02_CH(a=1,b=n)	0.7674	0.7657	6.4491	7.6834	0.4921	0.4969	0.1566	0.1573
03_CH(a=1/2,b=n/2)	0.7675	0.7662	7.2009	8.2850	0.4949	0.4968	0.1570	0.1577
04_CH(a=1,b=n/2)	0.7675	0.7660	6.3709	6.8212	0.4881	0.4901	0.1553	0.1556
05_Beta-prime	0.7670	0.7662	7.7466	9.3265	0.4983	0.5020	0.1584	0.1594
06_ZS adpated	0.7668	0.7659	6.7087	8.0829	0.4938	0.4972	0.1567	0.1575
07_Robust	0.7672	0.7643	6.3106	6.7434	0.4894	0.4914	0.1553	0.1556
08_Intrinsic	0.7674	0.7643	6.1029	6.4198	0.4864	0.4875	0.1545	0.1546
09_Hyper-g/n	0.7683	0.7664	6.7929	7.6676	0.4897	0.4931	0.1563	0.1569
10_Benchmark	0.7688	0.7671	20.4971	28.3342	0.5068	0.5092	0.1632	0.1641
11_DBF, g=n	0.7674	0.7655	6.3814	7.0062	0.4918	0.4970	0.1563	0.1569
12_TBF, g=n	0.7674	0.7660	6.2435	7.4158	0.4922	0.4970	0.1565	0.1573
13_Jeffreys	0.7696	0.7667	6.3832	6.1133	**0.4737**	**0.4715**	**0.1525**	**0.1516**
14_Hyper-g	0.7702	0.7678	8.3998	8.5375	0.4789	0.4787	0.1547	0.1545
15_Uniform	**0.7711**	**0.7680**	9.7869	10.6287	0.4812	0.4808	0.1557	0.1554
16_Local EB	0.7673	0.7645	27.4713	22.7148	0.4889	0.4899	0.1587	0.1588
17_AIC	0.7644	0.7621	**5.0875**	**5.1582**	0.4844	0.4852	0.1527	0.1524
18_BIC	0.7667	0.7646	5.5712	6.1335	0.4896	0.4918	0.1550	0.1552

Rank Unit

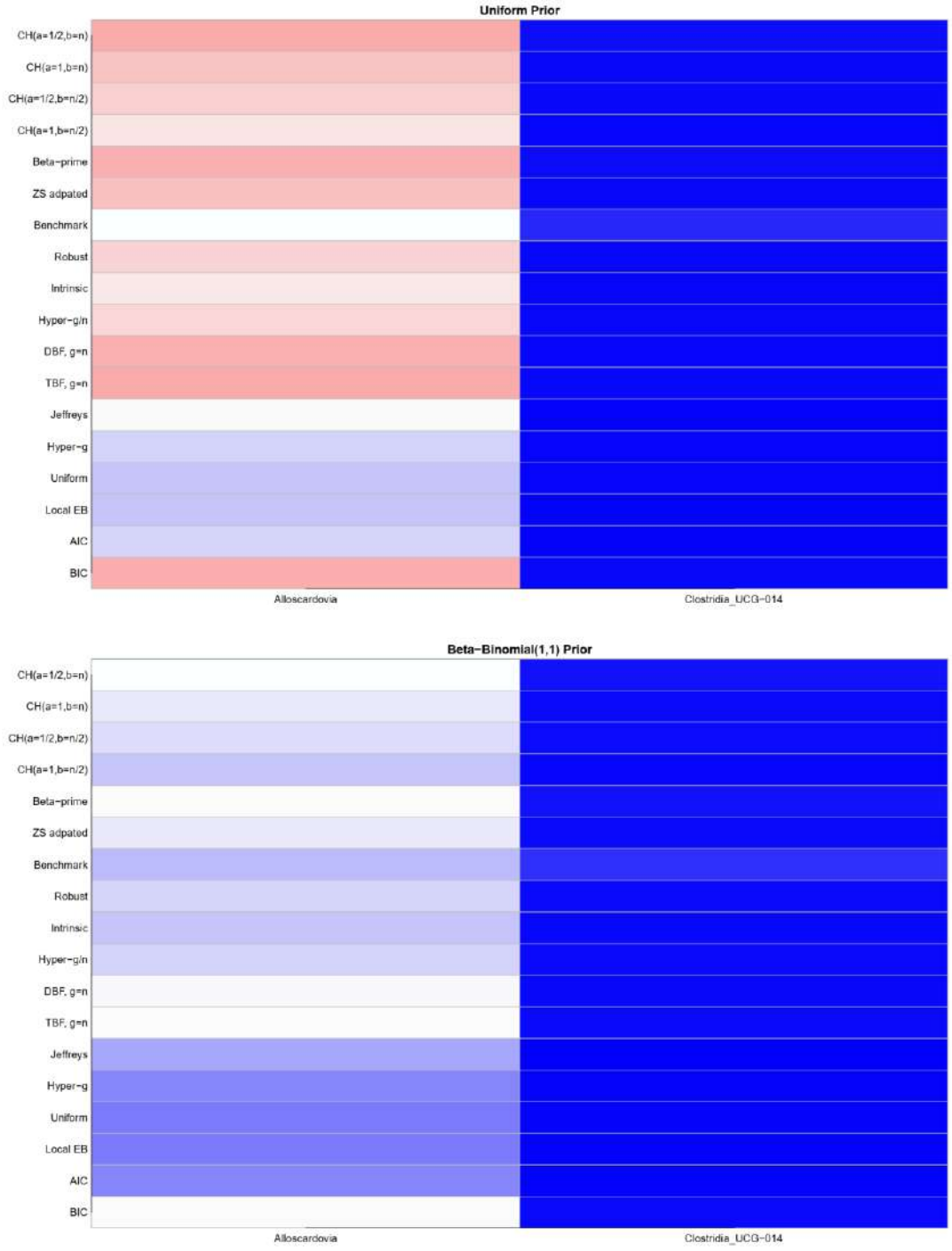


Figure G-E-4 –Marginal posterior inclusion probabilities for the **Genus** for **Saliva** with **Rank Unit** Normalization method. The colors are related to the magnitude of the inclusion probability with darkest blue corresponding to one and red to zero, while 0.5 is shown as white

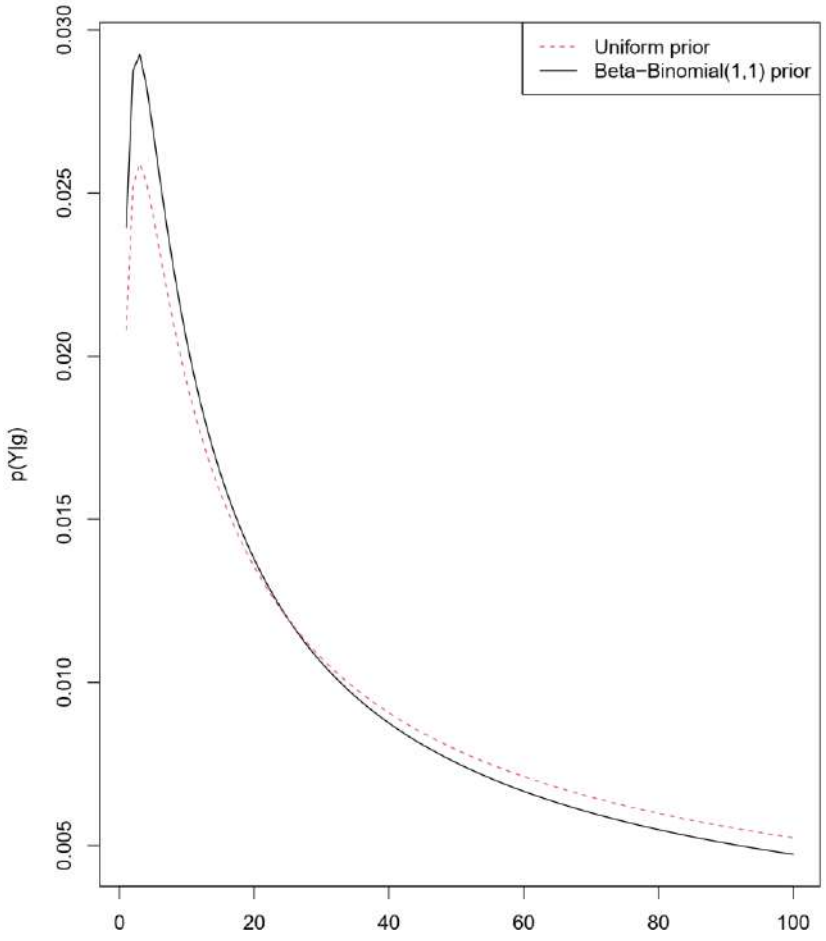


Figure G-E-5: Marginal likelihood of *g* for the **Genus** for **Saliva** with **Rank Unit** Normalization method.

Table G-E-4: Prediction accuracy for the **Saliva (Phylum) Rank Unit**, aggregated from at least 250 bootstrap cross validation sets. Bold font marks the largest AUC, the CS closest to one, and the smallest LS and Brier score

	AUC		CS		LS		Brier	
p(M)	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**	**Unif**	**BB(1,1)**
01_CH(a=1/2,b=n)	0.7928	0.7921	4.9256	7.3426	0.4784	0.4824	0.1521	0.1526
02_CH(a=1,b=n)	0.7930	0.7920	4.0639	5.9344	0.4745	0.4771	0.1503	0.1504
03_CH(a=1/2,b=n/2)	0.7937	0.7929	4.2375	7.8017	0.4743	0.4769	0.1504	0.1506
04_CH(a=1,b=n/2)	0.7936	0.7924	7.4678	5.7835	0.4710	0.4723	0.1489	0.1487
05_Beta-prime	0.7927	0.7920	4.8264	7.3781	0.4781	0.4819	0.1519	0.1524
06_ZS adpated	0.7928	0.7922	4.1418	5.8753	0.4747	0.4773	0.1504	0.1505
07_Robust	0.7915	0.7917	5.3357	5.5568	0.4724	0.4738	0.1488	0.1486
08_Intrinsic	0.7916	0.7913	4.6342	5.1885	0.4708	0.4721	0.1481	0.1477
09_Hyper-g/n	0.7946	0.7950	7.7634	7.3190	0.4698	0.4717	0.1497	0.1497
10_Benchmark	0.7961	0.7961	14.8720	19.2670	0.4846	0.4875	0.1557	0.1566
11_DBF, g=n	0.7918	0.7917	3.9956	5.4698	0.4748	0.4765	0.1502	0.1501
12_TBF, g=n	0.7918	0.7923	4.0025	5.9374	0.4747	0.4768	0.1505	0.1506
13_Jeffreys	0.7965	0.7959	5.0123	5.4615	**0.4612**	**0.4598**	**0.1463**	**0.1452**
14_Hyper-g	0.7966	0.7946	7.4629	7.3893	0.4673	0.4670	0.1482	0.1477
15_Uniform	**0.7979**	**0.7961**	8.3344	8.4066	0.4667	0.4662	0.1492	0.1487
16_Local EB	0.7752	0.7757	41.4101	44.0721	0.4778	0.4791	0.1548	0.1547
17_AIC	0.7906	0.7897	3.9835	**4.2972**	0.4712	0.4702	0.1466	0.1459
18_BIC	0.7906	0.7912	**3.3894**	7.7258	0.4734	0.4749	0.1490	0.1486

GAMLSS-BEZI with Random Effect Results

In the following table, the statistically significant variables are presented for P-Value and FDR under 0.1.

The GAMLSS BEZI with random effect (Subject) + Group (Colorectal Cancer) + Sampling_Site

Phylum

id	Estimate.Tumor1	Std.Error.Tumor1	t value.Tumor1	Pr(>|t|).Tumor1	Estimate.Sampling_Type1	Std.Error.Sampling_Type1	t value.Sampling_Type1	Pr(>|t|).Sampling_Type1	Estimate.Sampling_Type2	Std.Error.Sampling_Type2	t value.Sampling_Type2	Pr(>|t|).Sampling_Type2	pval.adjust.Tumor1	pval.adjust.Sampling_Type1	pval.adjust.Sampling_Type2
Acidobacteriota	0.354	0.001	450.304	0.000	-4.845	0.007	-671.234	0.000	-2.784	13867.505	0.000	1.000	0.000	0.000	1.000
Actinobacteriota	-0.158	0.157	-1.006	0.316	0.171	0.179	0.954	0.342	0.848	0.159	5.333	0.000	0.460	0.455	0.000
Bacteroidota	0.066	0.136	0.482	0.631	0.043	0.139	0.308	0.758	-0.386	0.137	-2.813	0.006	0.721	0.867	0.008
Campilobacterota	0.282	0.217	1.302	0.196	-1.226	0.352	-3.481	0.001	0.043	0.181	0.237	0.813	0.348	0.001	0.867
Chloroflexi	0.966	0.001	1036.597	0.000	-3.084	15249.857	0.000	1.000	-4.625	0.003	-1669.047	0.000	0.000	1.000	0.000
Cyanobacteria	0.609	0.258	2.359	0.020	-0.112	0.182	-0.612	0.542	-1.720	0.272	-6.330	0.000	0.040	0.667	0.000
Desulfobacterota	0.015	0.191	0.077	0.939	-0.937	0.172	-5.433	0.000	-1.773	0.276	-6.425	0.000	0.939	0.000	0.000
Euryarchaeota	-0.748	0.183	-4.088	0.000	1.744	0.183	9.525	0.000	-1.719	0.585	-2.937	0.004	0.000	0.000	0.007
Firmicutes	0.073	0.096	0.760	0.448	0.539	0.101	5.326	0.000	0.484	0.097	4.985	0.000	0.582	0.000	0.000
Fusobacteriota	0.544	0.205	2.658	0.009	-1.293	0.252	-5.128	0.000	0.494	0.173	2.862	0.005	0.021	0.000	0.008
Patescibacteria	-0.390	0.066	-5.871	0.000	0.855	0.280	3.054	0.003	4.382	0.260	16.849	0.000	0.000	0.005	0.000
Planctomycetota	2.460	0.003	807.230	0.000	-2.901	15249.857	0.000	1.000	-4.658	0.009	-511.696	0.000	0.000	1.000	0.000
Proteobacteria	-0.423	0.148	-2.849	0.005	-1.182	0.175	-6.740	0.000	-0.239	0.149	-1.599	0.113	0.014	0.000	0.129
Spirochaetota	0.193	0.268	0.721	0.473	-1.839	0.528	-3.479	0.001	-0.575	0.243	-2.360	0.020	0.582	0.001	0.025
Synergistota	-0.102	0.296	-0.343	0.732	0.373	0.218	1.711	0.090	-0.560	0.238	-2.350	0.020	0.781	0.131	0.025
Verrucomicrobiota	-0.220	0.201	-1.095	0.276	0.641	0.179	3.587	0.001	-1.463	0.466	-3.142	0.002	0.442	0.001	0.004

Class

id	Estimate.Tumor1	Std.Error.Tumor1	t value.Tumor1	Pr(>|t|).Tumor1	Estimate.Sampling_Type1	Std.Error.Sampling_Type1	t value.Sampling_Type1	Pr(>|t|).Sampling_Type1	Estimate.Sampling_Type2	Std.Error.Sampling_Type2	t value.Sampling_Type2	Pr(>|t|).Sampling_Type2	pval.adjust.Tumor1	pval.adjust.Sampling_Type1	pval.adjust.Sampling_Type2
Acidimicrobiia	-1.603	0.000	-3343.339	0.000	-2.752	15249.857	0.000	1.000	-2.789	13867.505	0.000	1.000	0.000	1.000	1.000
Acidobacteriae	0.659	0.001	713.683	0.000	-3.130	15249.857	0.000	1.000	-3.143	13867.505	0.000	1.000	0.000	1.000	1.000
Actinobacteria	-0.132	0.161	-0.819	0.414	-0.133	0.189	-0.706	0.481	1.021	0.161	6.333	0.000	0.555	0.658	0.000
Alphaproteobacteria	0.106	0.133	0.798	0.427	-1.678	0.184	-9.101	0.000	-3.350	0.304	-11.018	0.000	0.555	0.000	0.000
Bacilli	-0.015	0.122	-0.125	0.900	0.082	0.162	0.504	0.615	1.893	0.134	14.142	0.000	0.900	0.762	0.000
Bacteroidia	0.065	0.136	0.481	0.631	0.044	0.139	0.315	0.753	-0.385	0.137	-2.810	0.006	0.684	0.890	0.010
Campylobacteria	0.282	0.217	1.302	0.196	-1.226	0.352	-3.481	0.001	0.043	0.181	0.237	0.813	0.363	0.001	1.000
Clostridia	0.064	0.094	0.679	0.499	0.489	0.084	5.793	0.000	-2.127	0.132	-16.146	0.000	0.602	0.000	0.000
Coriobacteriia	-0.167	0.155	-1.079	0.283	0.979	0.165	5.940	0.000	0.272	0.168	1.615	0.109	0.490	0.000	0.157
Cyanobacteriia	0.340	0.114	2.972	0.004	-2.693	0.104	-25.947	0.000	-4.389	0.166	-26.443	0.000	0.009	0.000	0.000
Desulfobacteria	0.979	0.074	13.238	0.000	-4.432	0.202	-21.985	0.000	-3.990	0.318	-12.545	0.000	0.000	0.000	0.000
Desulfobulbia	0.691	0.002	405.990	0.000	-1.506	0.004	-408.549	0.000	-3.349	0.002	-1416.332	0.000	0.000	0.000	0.000
Desulfovibrionia	-0.153	0.151	-1.012	0.314	-0.901	0.141	-6.373	0.000	-2.361	0.306	-7.707	0.000	0.491	0.000	0.000
Fusobacteriia	0.544	0.205	2.658	0.009	-1.293	0.252	-5.128	0.000	0.494	0.173	2.861	0.005	0.021	0.000	0.009
Gammaproteobacteria	-0.451	0.149	-3.036	0.003	-1.155	0.176	-6.566	0.000	-0.156	0.150	-1.046	0.298	0.009	0.000	0.408
Gracilibacteria	0.654	0.017	37.734	0.000	-0.029	0.256	-0.112	0.911	4.630	0.134	34.550	0.000	0.000	1.000	0.000
Incertae_Sedis	-0.145	0.242	-0.601	0.549	0.268	0.161	1.666	0.098	-1.960	13867.505	0.000	1.000	0.621	0.159	1.000
Lentisphaeria	0.364	0.365	0.997	0.321	1.203	0.221	5.438	0.000	-1.058	13867.505	0.000	1.000	0.491	0.000	1.000
Methanobacteria	-0.748	0.183	-4.086	0.000	1.744	0.183	9.526	0.000	-1.719	0.585	-2.937	0.004	0.000	0.000	0.008
Negativicutes	0.112	0.129	0.866	0.388	-0.122	0.171	-0.710	0.479	1.689	0.139	12.178	0.000	0.555	0.658	0.000
Saccharimonadia	-0.465	0.072	-6.467	0.000	0.828	0.296	2.793	0.006	4.025	0.277	14.514	0.000	0.000	0.012	0.000
Spirochaetia	0.180	0.272	0.662	0.510	-0.652	0.548	-1.188	0.237	0.527	0.297	1.777	0.078	0.602	0.363	0.120
Synergistia	-0.102	0.296	-0.343	0.732	0.373	0.218	1.711	0.090	-0.560	0.238	-2.350	0.020	0.762	0.156	0.033
Thermoleophilia	-1.906	0.001	-3149.867	0.000	-2.854	15249.857	0.000	1.000	-2.940	13867.505	0.000	1.000	0.000	1.000	1.000
Vampirivibrionia	1.369	0.627	2.183	0.031	0.120	0.206	0.586	0.559	-1.712	13867.505	0.000	1.000	0.067	0.727	1.000
Verrucomicrobiae	-0.318	0.207	-1.537	0.127	0.640	0.183	3.491	0.001	-1.585	0.467	-3.393	0.001	0.255	0.001	0.002

Order

id	Estimate.Tumor1	Std.Error.Tumor1	t value.Tumor1	Pr(>|t|).Tumor1	Estimate.Sampling_Type1	Std.Error.Sampling_Type1	t value.Sampling_Type1	Pr(>|t|).Sampling_Type1	Estimate.Sampling_Type2	Std.Error.Sampling_Type2	t value.Sampling_Type2	Pr(>|t|).Sampling_Type2	pval.adjust.Tumor1	pval.adjust.Sampling_Type1	pval.adjust.Sampling_Type2
Absconditabacteriales_.SR1.	0.728	0.015	49.976	0.000	-0.341	0.214	-1.598	0.113	4.354	0.108	40.485	0.000	0.000	0.146	0.000
Acidaminococcales	-0.094	0.180	-0.522	0.602	-0.734	0.156	-4.707	0.000	-1.370	13867.505	0.000	1.000	0.720	0.000	1.000
Actinomycetales	-0.421	0.090	-4.685	0.000	-0.402	0.220	-1.829	0.071	2.418	0.174	13.898	0.000	0.000	0.099	0.000
Aeromonadales	2.458	0.304	8.091	0.000	-4.304	0.272	-15.797	0.000	-3.117	13867.505	0.000	1.000	0.000	0.000	1.000
Alteromonadales	-2.304	0.014	-162.671	0.000	-2.421	0.062	-39.348	0.000	-3.603	0.081	-44.712	0.000	0.000	0.000	0.000
Bacillales	-0.569	0.160	-3.557	0.001	-1.960	0.133	-14.765	0.000	-5.225	0.324	-16.109	0.000	0.001	0.000	0.000
Bacteroidales	0.059	0.136	0.431	0.667	0.087	0.138	0.629	0.530	-0.464	0.138	-3.366	0.001	0.752	0.620	0.002
Bifidobacteriales	-0.074	0.205	-0.364	0.717	0.529	0.195	2.710	0.008	-0.551	0.211	-2.614	0.010	0.783	0.012	0.015
Burkholderiales	0.027	0.153	0.174	0.862	-1.031	0.199	-5.169	0.000	0.844	0.148	5.703	0.000	0.866	0.000	0.000
Campylobacterales	0.282	0.217	1.302	0.196	-1.226	0.352	-3.481	0.001	0.043	0.181	0.237	0.813	0.311	0.001	0.976
Cardiobacteriales	-0.016	0.002	-10.161	0.000	0.505	0.019	25.973	0.000	2.566	0.008	312.637	0.000	0.000	0.000	0.000
Chloroplast	0.346	0.127	2.719	0.008	-2.375	0.108	-21.894	0.000	-4.047	0.170	-23.823	0.000	0.015	0.000	0.000
Christensenellales	0.298	0.223	1.337	0.184	0.509	0.157	3.245	0.002	-1.589	0.990	-1.605	0.111	0.300	0.003	0.145
Chthoniobacterales	-0.004	0.000	-11.302	0.000	-4.503	0.009	-528.281	0.000	-3.218	0.004	-827.267	0.000	0.000	0.000	0.000
Clostridia	0.536	0.399	1.344	0.181	0.750	0.112	6.729	0.000	-1.895	13867.505	0.000	1.000	0.300	0.000	1.000
Clostridia_UCG.014	-0.092	0.177	-0.517	0.606	0.704	0.204	3.447	0.001	0.225	0.194	1.159	0.249	0.720	0.001	0.311
Clostridia_vadinBB60_group	0.164	0.272	0.602	0.549	0.006	0.207	0.029	0.977	-1.430	0.304	-4.707	0.000	0.694	1.000	0.000
Clostridiales	-0.441	0.127	-3.465	0.001	-0.532	0.125	-4.251	0.000	-0.548	13867.505	0.000	1.000	0.002	0.000	1.000
Coriobacteriales	-0.167	0.155	-1.079	0.283	0.979	0.165	5.939	0.000	0.272	0.168	1.615	0.109	0.438	0.000	0.145
Corynebacteriales	0.170	0.208	0.816	0.416	-1.243	0.450	-2.760	0.007	0.619	0.175	3.535	0.001	0.576	0.011	0.001
Desulfobacterales	0.918	0.074	12.397	0.000	-4.431	0.202	-21.971	0.000	-3.989	0.318	-12.543	0.000	0.000	0.000	0.000
Desulfobulbales	0.691	0.002	405.976	0.000	-1.505	0.004	-408.545	0.000	-3.349	0.002	-1416.791	0.000	0.000	0.000	0.000
Desulfovibrionales	-0.153	0.151	-1.013	0.314	-0.901	0.141	-6.373	0.000	-2.361	0.306	-7.707	0.000	0.474	0.000	0.000
DTU014	-0.146	0.242	-0.603	0.548	0.268	0.161	1.666	0.098	-1.960	13867.505	0.000	1.000	0.694	0.129	1.000
Enterobacterales	-0.696	0.157	-4.440	0.000	-1.015	0.162	-6.270	0.000	-2.188	0.245	-8.947	0.000	0.000	0.000	0.000
Erysipelotrichales	-0.134	0.146	-0.912	0.364	0.628	0.145	4.322	0.000	-0.668	0.167	-4.009	0.000	0.524	0.000	0.000
Eubacteriales	-1.289	0.028	-45.643	0.000	-3.172	0.010	-310.851	0.000	-4.244	0.023	-187.059	0.000	0.000	0.000	0.000
Flavobacteriales	-0.032	0.187	-0.169	0.866	-0.507	0.295	-1.717	0.089	1.269	0.199	6.385	0.000	0.866	0.122	0.000
Frankiales	-2.497	0.001	-2073.502	0.000	-2.746	15249.857	0.000	1.000	-2.802	13867.505	0.000	1.000	0.000	1.000	1.000
Fusobacteriales	0.544	0.205	2.658	0.009	-1.293	0.252	-5.128	0.000	0.494	0.173	2.861	0.005	0.017	0.000	0.008
Gastranaerophilales	1.186	0.632	1.877	0.063	0.134	0.232	0.579	0.564	-1.784	13867.505	0.000	1.000	0.115	0.647	1.000

Gracilibacteria	2.049	0.003	795.061	0.000	-1.328	15249.857	0.000	1.000	1.506	0.006	250.592	0.000	0.000	1.000	0.000
Izemoplasmatales	-0.568	0.142	-3.995	0.000	0.345	0.164	2.100	0.038	-0.176	0.300	-0.586	0.559	0.000	0.054	0.684
JGI_0000069.P22	0.259	0.001	372.333	0.000	-2.308	15249.857	0.000	1.000	NA	NA	NA	NA	0.000	1.000	NA
Lachnospirales	-0.036	0.126	-0.288	0.774	-0.290	0.122	-2.372	0.019	-1.676	0.154	-10.909	0.000	0.813	0.029	0.000
Lactobacillales	-0.027	0.126	-0.218	0.828	-0.076	0.186	-0.408	0.684	2.508	0.153	16.361	0.000	0.855	0.771	0.000
Methanobacteriales	-0.748	0.183	-4.088	0.000	1.744	0.183	9.525	0.000	-1.719	0.585	-2.937	0.004	0.000	0.000	0.007
Micrococcales	-0.084	0.166	-0.504	0.615	-0.270	0.261	-1.035	0.303	1.852	0.194	9.569	0.000	0.720	0.361	0.000
Monoglobales	-1.273	0.104	-12.266	0.000	1.113	0.128	8.720	0.000	-0.223	13867.505	0.000	1.000	0.000	0.000	1.000
Mycoplasmatales	-0.188	0.000	-610.160	0.000	-1.679	15249.857	0.000	1.000	NA	NA	NA	NA	0.000	1.000	NA
Oscillospirales	0.089	0.109	0.813	0.418	0.998	0.094	10.628	0.000	-3.354	0.300	-11.190	0.000	0.576	0.000	0.000
Paenibacillales	1.089	0.001	1190.726	0.000	1.416	0.001	1421.761	0.000	-3.198	0.004	-815.413	0.000	0.000	0.000	0.000
Pasteurellales	-0.310	0.121	-2.573	0.012	-0.609	0.265	-2.300	0.023	2.253	0.173	13.051	0.000	0.022	0.035	0.000
Peptococcales	-0.115	0.243	-0.474	0.636	-0.338	0.216	-1.567	0.120	-0.613	0.227	-2.693	0.008	0.730	0.152	0.012
Peptostreptococcales.Tissierellales	0.134	0.186	0.717	0.474	-0.215	0.187	-1.151	0.252	-0.301	0.179	-1.678	0.096	0.639	0.306	0.130
Propionibacteriales	-0.401	0.038	-10.425	0.000	-4.247	0.263	-16.157	0.000	-3.506	0.161	-21.812	0.000	0.000	0.000	0.000
Pseudomonadales	-0.616	0.117	-5.258	0.000	-1.505	0.145	-10.366	0.000	-2.450	0.187	-13.092	0.000	0.000	0.000	0.000
RF39	-0.715	0.117	-6.098	0.000	1.067	0.387	2.760	0.007	-1.103	0.396	-2.783	0.006	0.000	0.011	0.010
Rhizobiales	-0.245	0.026	-9.364	0.000	-1.861	0.108	-17.157	0.000	-5.905	0.350	-16.889	0.000	0.000	0.000	0.000
Rhodobacterales	0.332	0.000	821.107	0.000	-0.857	0.001	-636.999	0.000	-2.122	0.002	-1243.747	0.000	0.000	0.000	0.000
Rhodospirillales	0.618	0.362	1.706	0.091	-0.795	0.276	-2.885	0.005	-2.077	13867.505	0.000	1.000	0.156	0.008	1.000
Rickettsiales	-0.931	0.001	-643.467	0.000	2.784	0.002	1803.464	0.000	0.734	0.002	437.055	0.000	0.000	0.000	0.000
Saccharimonadales	-0.465	0.072	-6.467	0.000	0.828	0.296	2.792	0.006	4.025	0.277	14.514	0.000	0.000	0.011	0.000
Sphingobacteriales	0.595	0.033	18.096	0.000	-6.270	0.089	-70.315	0.000	-2.291	0.030	-76.193	0.000	0.000	0.000	0.000
Sphingomonadales	1.395	0.071	19.677	0.000	-1.364	15249.857	0.000	1.000	-0.362	0.149	-2.428	0.017	0.000	1.000	0.024
Spirochaetales	0.180	0.272	0.661	0.510	-0.651	0.548	-1.188	0.238	0.527	0.297	1.778	0.078	0.673	0.295	0.109
Staphylococcales	0.053	0.148	0.360	0.720	-0.950	0.210	-4.529	0.000	1.155	0.151	7.644	0.000	0.783	0.000	0.000
Synergistales	-0.102	0.296	-0.343	0.732	0.373	0.218	1.709	0.090	-0.560	0.238	-2.350	0.020	0.783	0.122	0.029
Veillonellales.Selenomonadales	0.079	0.136	0.578	0.564	-0.033	0.193	-0.173	0.863	2.076	0.157	13.188	0.000	0.699	0.956	0.000
Verrucomicrobiales	-0.401	0.221	-1.815	0.072	0.750	0.194	3.857	0.000	-0.814	0.522	-1.559	0.122	0.128	0.000	0.156
Victivallales	0.363	0.365	0.997	0.321	1.203	0.221	5.438	0.000	-1.058	13867.505	0.000	1.000	0.474	0.000	1.000
Xanthomonadales	0.854	0.001	1228.836	0.000	-2.402	0.005	-484.748	0.000	-1.202	0.001	-921.595	0.000	0.000	0.000	0.000

Family

id	Estimate.Tumor1	Std. Error.Tumor1	t value.Tumor1	Pr(>|t|).Tumor1	Estimate.Sampling_Type1	Std. Error.Sampling_Type1	t value.Sampling_Type1	Pr(>|t|).Sampling_Type1	Estimate.Sampling_Type2	Std. Error.Sampling_Type2	t value.Sampling_Type2	Pr(>|t|).Sampling_Type2	pval.adjust.Tumor1	pval.adjust.Sampling_Type1	pval.adjust.Sampling_Type2
X.Clostridium._methylpentosum_group	-0.317	0.106	-2.987	0.003	-0.129	0.123	-1.048	0.297	-1.669	13867.505	0.000	1.000	0.008	0.359	1.000
X.Eubacterium._coprostanoligenes_group	-0.074	0.128	-0.580	0.563	1.751	0.125	14.046	0.000	-1.749	0.700	-2.498	0.014	0.723	0.000	0.021
Absconditabacteriales_.SR1.	0.729	0.015	49.977	0.000	-0.341	0.214	-1.598	0.113	4.354	0.108	40.484	0.000	0.000	0.161	0.000
Acidaminococcaceae	-0.094	0.180	-0.522	0.603	-0.734	0.156	-4.707	0.000	-1.370	13867.505	0.000	1.000	0.760	0.000	1.000
Actinomycetaceae	-0.421	0.090	-4.686	0.000	-0.402	0.220	-1.828	0.071	2.418	0.174	13.900	0.000	0.000	0.108	0.000
Aerococcaceae	0.517	0.064	8.120	0.000	2.387	0.278	8.597	0.000	3.308	0.216	15.317	0.000	0.000	0.000	0.000
Akkermansiaceae	-0.401	0.221	-1.815	0.072	0.750	0.194	3.858	0.000	-0.814	0.522	-1.559	0.122	0.134	0.000	0.160
Alteromonadaceae	-0.921	0.017	-53.774	0.000	-2.365	0.063	-37.556	0.000	-2.880	0.085	-33.892	0.000	0.000	0.000	0.000
Anaerovoracaceae	0.204	0.179	1.136	0.258	0.206	0.172	1.195	0.235	-0.179	0.174	-1.029	0.306	0.395	0.301	0.377
Atopobiaceae	0.177	0.150	1.181	0.240	0.320	0.280	1.142	0.256	1.769	0.255	6.936	0.000	0.379	0.321	0.000
Bacillaceae	0.609	0.208	2.928	0.004	0.590	0.153	3.860	0.000	-3.317	0.389	-8.532	0.000	0.009	0.000	0.000
Bacteroidaceae	0.130	0.139	0.937	0.351	-0.173	0.122	-1.414	0.160	-3.219	0.252	-12.793	0.000	0.507	0.214	0.000
Barnesiellaceae	-0.456	0.146	-3.120	0.002	0.770	0.128	5.999	0.000	-0.488	13867.505	0.000	1.000	0.006	0.000	1.000
Beijerinckiaceae	0.400	0.001	635.719	0.000	-2.513	15249.857	0.000	1.000	-4.467	0.009	-500.288	0.000	0.000	1.000	0.000
Bifidobacteriaceae	-0.074	0.205	-0.364	0.717	0.529	0.195	2.711	0.008	-0.551	0.211	-2.614	0.010	0.837	0.014	0.016
Burkholderiaceae	0.622	0.125	4.975	0.000	-0.897	0.544	-1.651	0.102	1.605	0.130	12.389	0.000	0.000	0.147	0.000
Butyricicoccaceae	-0.300	0.174	-1.727	0.087	-0.099	0.144	-0.690	0.492	-1.288	13867.505	0.000	1.000	0.158	0.568	1.000
Campylobacteraceae	0.262	0.218	1.199	0.233	-1.202	0.372	-3.233	0.002	0.099	0.183	0.540	0.590	0.373	0.003	0.701
Cardiobacteriaceae	-0.016	0.002	-10.163	0.000	0.505	0.019	25.930	0.000	2.566	0.008	312.062	0.000	0.000	0.000	0.000
Carnobacteriaceae	0.002	0.081	0.027	0.978	-0.659	0.229	-2.874	0.005	2.723	0.163	16.689	0.000	0.998	0.009	0.000
Chloroplast	0.346	0.127	2.719	0.008	-2.375	0.108	-21.894	0.000	-4.047	0.170	-23.823	0.000	0.016	0.000	0.000
Christensenellaceae	0.298	0.223	1.337	0.184	0.509	0.157	3.248	0.002	-1.589	0.990	-1.605	0.111	0.309	0.003	0.150
Chthoniobacteraceae	-0.060	0.000	-158.034	0.000	-4.503	0.009	-528.276	0.000	-3.218	0.004	-827.249	0.000	0.000	0.000	0.000
Clostridia_UCG.014	-0.091	0.177	-0.516	0.607	0.705	0.204	3.448	0.001	0.225	0.194	1.159	0.249	0.760	0.002	0.314
Clostridia_vadinBB60_group	0.164	0.272	0.602	0.548	0.007	0.207	0.036	0.972	-1.426	0.304	-4.690	0.000	0.723	1.000	0.000
Clostridiaceae	-0.441	0.127	-3.464	0.001	-0.532	0.125	-4.250	0.000	-0.548	13867.505	0.000	1.000	0.002	0.000	1.000
Comamonadaceae	0.159	0.054	2.909	0.005	-4.892	0.445	-11.000	0.000	-3.752	0.190	-19.724	0.000	0.010	0.000	0.000
Coriobacteriaceae	-0.128	0.144	-0.885	0.378	1.236	0.118	10.439	0.000	-0.681	0.496	-1.373	0.173	0.539	0.000	0.224
Coriobacteriales_Incertae_Sedis	-0.135	0.050	-2.679	0.008	1.368	0.061	22.457	0.000	-0.270	13867.505	0.000	1.000	0.017	0.000	1.000
Corynebacteriaceae	0.083	0.198	0.418	0.677	-1.156	0.456	-2.535	0.013	1.260	0.206	6.123	0.000	0.809	0.022	0.000
Defluviitaleaceae	-0.324	0.079	-4.093	0.000	-1.120	0.091	-12.279	0.000	2.010	0.052	39.020	0.000	0.000	0.000	0.000

Desulfobacteraceae	0.949	0.023	40.525	0.000	-1.806	0.015	-123.084	0.000	-4.067	0.014	-293.413	0.000	0.000	0.000	0.000
Desulfocapsaceae	1.335	0.002	589.955	0.000	-1.505	0.005	-327.567	0.000	-2.580	0.008	-308.254	0.000	0.000	0.000	0.000
Desulfosarcinaceae	0.681	0.065	10.510	0.000	-4.791	0.194	-24.664	0.000	-4.316	0.355	-12.160	0.000	0.000	0.000	0.000
Desulfovibrionaceae	-0.194	0.150	-1.293	0.199	-0.894	0.141	-6.364	0.000	-2.180	0.318	-6.857	0.000	0.328	0.000	0.000
DTU014	-0.146	0.242	-0.601	0.549	0.268	0.161	1.666	0.098	-1.960	13867.505	0.000	1.000	0.723	0.144	1.000
Eggerthellaceae	-0.524	0.153	-3.435	0.001	2.069	0.201	10.278	0.000	0.110	0.249	0.441	0.660	0.002	0.000	0.775
Enterobacteriaceae	-0.497	0.169	-2.950	0.004	-1.035	0.165	-6.293	0.000	-2.055	0.360	-5.714	0.000	0.009	0.000	0.000
Enterococcaceae	0.293	0.385	0.761	0.448	-0.233	0.306	-0.761	0.448	-0.984	0.740	-1.331	0.186	0.621	0.523	0.238
Erwiniaceae	-0.064	0.054	-1.171	0.244	-3.557	15249.857	0.000	1.000	-5.128	0.046	-112.279	0.000	0.379	1.000	0.000
Erysipelatoclostridiaceae	-0.278	0.184	-1.511	0.134	0.583	0.166	3.523	0.001	-1.324	0.339	-3.909	0.000	0.232	0.001	0.000
Erysipelotrichaceae	0.014	0.170	0.081	0.935	0.369	0.173	2.132	0.035	-0.186	0.175	-1.061	0.291	0.973	0.055	0.363
Eubacteriaceae	-1.353	0.002	-731.175	0.000	-3.267	0.001	-4755.981	0.000	-4.601	0.002	-2896.528	0.000	0.000	0.000	0.000
Flavobacteriaceae	-0.007	0.208	-0.034	0.973	-0.437	0.344	-1.271	0.207	0.795	0.225	3.537	0.001	0.998	0.272	0.001
Fusobacteriaceae	0.448	0.217	2.061	0.041	-1.279	0.259	-4.930	0.000	0.115	0.182	0.632	0.529	0.080	0.000	0.643
Gastranaerophilales	1.187	0.632	1.878	0.063	0.134	0.232	0.580	0.563	-1.783	13867.505	0.000	1.000	0.118	0.637	1.000
Gemellaceae	-0.028	0.130	-0.215	0.830	-0.788	0.226	-3.490	0.001	1.737	0.165	10.529	0.000	0.909	0.001	0.000
Gracilibacteria	2.049	0.003	795.094	0.000	-1.328	15249.857	0.000	1.000	1.507	0.006	250.598	0.000	0.000	1.000	0.000
Hungateiclostridiaceae	0.536	0.399	1.344	0.181	0.750	0.112	6.729	0.000	-1.895	13867.505	0.000	1.000	0.309	0.000	1.000
Izemoplasmatales	-0.568	0.142	-3.995	0.000	0.345	0.164	2.100	0.038	-0.176	0.300	-0.585	0.560	0.000	0.059	0.673
JGI_0000069.P22	0.259	0.001	372.401	0.000	-2.308	15249.857	0.000	1.000	NA	NA	NA	NA	0.000	1.000	NA
Lachnospiraceae	-0.038	0.127	-0.298	0.766	-0.289	0.122	-2.363	0.020	-1.685	0.154	-10.941	0.000	0.866	0.032	0.000
Lactobacillaceae	-0.505	0.168	-3.001	0.003	-0.365	0.210	-1.732	0.086	-0.648	0.217	-2.987	0.004	0.008	0.130	0.006
Lentimicrobiaceae	0.416	0.001	466.745	0.000	-2.486	15249.857	0.000	1.000	NA	NA	NA	NA	0.000	1.000	NA
Leptotrichiaceae	0.053	0.110	0.481	0.631	-1.349	0.414	-3.261	0.002	1.990	0.163	12.198	0.000	0.778	0.003	0.000
Leuconostocaceae	-0.108	0.002	-55.920	0.000	2.628	0.001	4150.102	0.000	-1.535	0.003	-555.668	0.000	0.000	0.000	0.000
Marinifilaceae	-0.147	0.155	-0.948	0.345	0.144	0.133	1.088	0.279	-1.712	0.446	-3.836	0.000	0.506	0.342	0.000
Methanobacteriaceae	-0.748	0.183	-4.088	0.000	1.744	0.183	9.528	0.000	-1.719	0.585	-2.936	0.004	0.000	0.000	0.006
Microbacteriaceae	-1.537	0.055	-27.736	0.000	-2.658	0.055	-48.774	0.000	-5.431	0.186	-29.156	0.000	0.000	0.000	0.000
Micrococcaceae	-0.056	0.148	-0.382	0.703	-0.448	0.286	-1.566	0.120	2.419	0.208	11.657	0.000	0.831	0.167	0.000
Mitochondria	-0.931	0.001	-643.482	0.000	2.785	0.002	1803.495	0.000	0.734	0.002	437.085	0.000	0.000	0.000	0.000
Monoglobaceae	-1.273	0.104	-12.264	0.000	1.113	0.128	8.720	0.000	-0.223	13867.505	0.000	1.000	0.000	0.000	1.000
Moraxellaceae	-0.381	0.101	-3.787	0.000	-1.271	0.175	-7.274	0.000	-2.309	0.185	-12.496	0.000	0.001	0.000	0.000
Muribaculaceae	-0.380	0.245	-1.549	0.124	0.828	0.168	4.929	0.000	-1.854	0.719	-2.577	0.011	0.219	0.000	0.017

Mycoplasmataceae	-0.188	0.000	-610.327	0.000	-1.679	15249.857	0.000	1.000	NA	NA	NA	NA	0.000	1.000	NA
Neisseriaceae	0.011	0.085	0.128	0.899	0.327	0.380	0.860	0.392	3.420	0.170	20.139	0.000	0.954	0.463	0.000
Nocardiaceae	2.902	0.004	696.458	0.000	-2.315	15249.857	0.000	1.000	-2.216	13867.505	0.000	1.000	0.000	1.000	1.000
Oscillospiraceae	0.383	0.141	2.724	0.008	0.583	0.109	5.359	0.000	-1.493	0.703	-2.124	0.036	0.016	0.000	0.051
Oscillospirales	1.754	1789 4.051	0.000	1.000	0.502	0.341	1.469	0.144	-1.866	14469.264	0.000	1.000	1.000	0.197	1.000
Paenibacillaceae	1.089	0.001	1190.733	0.000	1.416	0.001	1421.758	0.000	-3.197	0.004	-815.511	0.000	0.000	0.000	0.000
Paludibacteraceae	0.171	0.055	3.102	0.002	1.715	15249.857	0.000	1.000	3.061	0.352	8.705	0.000	0.006	1.000	0.000
Pasteurellaceae	-0.310	0.121	-2.573	0.012	-0.609	0.265	-2.300	0.023	2.253	0.173	13.052	0.000	0.023	0.038	0.000
Peptococcaceae	-0.115	0.243	-0.474	0.636	-0.338	0.216	-1.566	0.120	-0.613	0.227	-2.693	0.008	0.778	0.167	0.013
Peptostreptococcaceae	0.000	0.183	0.001	0.999	-0.406	0.186	-2.178	0.031	-0.426	0.178	-2.395	0.018	1.000	0.050	0.026
Peptostreptococcales.Tisierellales	-0.143	0.240	-0.596	0.552	-0.602	0.241	-2.502	0.014	-0.366	0.199	-1.839	0.068	0.723	0.023	0.095
Planococcaceae	-1.040	0.013	-81.911	0.000	-4.224	0.117	-36.070	0.000	-2.243	0.120	-18.620	0.000	0.000	0.000	0.000
Porphyromonadaceae	0.017	0.180	0.093	0.926	-0.312	0.279	-1.118	0.266	1.670	0.213	7.851	0.000	0.973	0.329	0.000
Prevotellaceae	-0.062	0.178	-0.347	0.729	-0.208	0.205	-1.015	0.312	0.994	0.180	5.524	0.000	0.837	0.373	0.000
Propionibacteriaceae	-0.357	0.039	-9.138	0.000	-4.187	0.258	-16.199	0.000	-3.430	0.157	-21.778	0.000	0.000	0.000	0.000
Pseudoalteromonadaceae	-4.576	0.001	-3188.185	0.000	-1.849	0.009	-213.338	0.000	-1.452	0.007	-222.604	0.000	0.000	0.000	0.000
Pseudomonadaceae	-0.876	0.095	-9.200	0.000	-2.298	0.102	-22.450	0.000	-1.948	0.224	-8.684	0.000	0.000	0.000	0.000
RF39	-0.715	0.117	-6.095	0.000	1.067	0.387	2.760	0.007	-1.103	0.396	-2.783	0.006	0.000	0.012	0.010
Rhizobiaceae	-0.902	0.024	-38.303	0.000	-2.565	0.160	-16.074	0.000	-5.940	0.409	-14.537	0.000	0.000	0.000	0.000
Rhodobacteraceae	0.332	0.000	821.142	0.000	-0.857	0.001	-637.009	0.000	-2.122	0.002	-1243.755	0.000	0.000	0.000	0.000
Rikenellaceae	-0.205	0.093	-2.200	0.030	1.244	0.092	13.577	0.000	-2.255	0.229	-9.834	0.000	0.059	0.000	0.000
Ruminococcaceae	-0.087	0.111	-0.785	0.434	0.842	0.099	8.523	0.000	-3.262	0.299	-10.918	0.000	0.611	0.000	0.000
Saccharimonadaceae	-0.515	0.078	-6.586	0.000	0.724	0.305	2.377	0.019	3.750	0.286	13.105	0.000	0.000	0.032	0.000
Saccharimonadales	-0.161	0.046	-3.497	0.001	0.767	0.529	1.450	0.150	4.840	0.502	9.645	0.000	0.002	0.203	0.000
Selenomonadaceae	-0.027	0.182	-0.147	0.883	-2.355	0.329	-7.161	0.000	-0.617	0.140	-4.398	0.000	0.947	0.000	0.000
Sphingomonadaceae	1.395	0.071	19.678	0.000	-1.364	15249.857	0.000	1.000	-0.362	0.149	-2.428	0.017	0.000	1.000	0.024
Spirochaetaceae	0.180	0.272	0.661	0.510	-0.651	0.548	-1.187	0.238	0.528	0.297	1.778	0.078	0.698	0.302	0.107
Staphylococcaceae	0.045	0.261	0.174	0.862	-1.747	0.308	-5.673	0.000	-1.516	0.308	-4.926	0.000	0.934	0.000	0.000
Streptococcaceae	-0.052	0.113	-0.457	0.648	0.232	0.188	1.238	0.219	2.932	0.157	18.678	0.000	0.784	0.284	0.000
Sutterellaceae	-0.183	0.162	-1.125	0.263	-1.549	0.174	-8.883	0.000	-1.634	13867.505	0.000	1.000	0.396	0.000	1.000
Synergistaceae	-0.102	0.296	-0.343	0.732	0.373	0.218	1.709	0.090	-0.560	0.238	-2.350	0.020	0.837	0.134	0.029
Tannerellaceae	0.040	0.147	0.275	0.784	0.551	0.139	3.974	0.000	-1.198	0.179	-6.686	0.000	0.876	0.000	0.000
UCG.010	-0.114	0.193	-0.587	0.558	1.956	0.157	12.427	0.000	0.649	13867.505	0.000	1.000	0.723	0.000	1.000

uncultured	0.471	0.274	1.720	0.088	0.125	0.190	0.660	0.511	-1.591	0.997	-1.596	0.113	0.158	0.584	0.151
Veillonellaceae	0.035	0.134	0.258	0.797	0.098	0.194	0.505	0.614	2.210	0.161	13.735	0.000	0.882	0.687	0.000
Victivallaceae	-0.253	0.237	-1.069	0.287	-0.032	0.176	-0.180	0.857	-1.933	13867.505	0.000	1.000	0.427	0.948	1.000
Weeksellaceae	0.121	0.100	1.211	0.229	-1.701	0.378	-4.495	0.000	1.581	0.165	9.588	0.000	0.372	0.000	0.000
Xanthobacteraceae	-0.324	0.017	-19.045	0.000	-0.818	0.045	-18.011	0.000	-5.571	0.384	-14.517	0.000	0.000	0.000	0.000
Xanthomonadaceae	-0.117	0.001	-152.005	0.000	-2.196	0.005	-440.524	0.000	-1.202	0.001	-931.041	0.000	0.000	0.000	0.000
Yersiniaceae	-1.050	0.148	-7.069	0.000	-0.884	0.121	-7.279	0.000	-4.396	0.324	-13.549	0.000	0.000	0.000	0.000

Genus

id	Estimate.Tumor1	Std.Error.Tumor1	t value.Tumor1	Pr(>|t|).Tumor1	Estimate.Sampling_Type1	Std.Error.Sampling_Type1	t value.Sampling_Type1	Pr(>|t|).Sampling_Type1	Estimate.Sampling_Type2	Std.Error.Sampling_Type2	t value.Sampling_Type2	Pr(>|t|).Sampling_Type2	pval.adjust.Tumor1	pval.adjust.Sampling_Type1	pval.adjust.Sampling_Type2
X.Clostridium._innocuum_group	-0.184	0.245	-0.750	0.455	-0.687	0.200	-3.444	0.001	-1.926	13867.505	0.000	1.000	0.551	0.001	1.000
X.Clostridium._methylpentosum_group	-0.316	0.106	-2.979	0.004	-0.133	0.123	-1.087	0.279	-1.674	13867.505	0.000	1.000	0.007	0.340	1.000
X.Eubacterium._brachy_group	-0.098	0.094	-1.050	0.296	1.064	0.090	11.767	0.000	0.980	0.069	14.126	0.000	0.389	0.000	0.000
X.Eubacterium._coprostanoligenes_group	-0.073	0.129	-0.568	0.571	1.749	0.125	13.999	0.000	-1.746	0.700	-2.494	0.014	0.654	0.000	0.032
X.Eubacterium._eligens_group	-0.506	0.182	-2.780	0.006	-0.141	0.170	-0.827	0.410	-1.148	13867.505	0.000	1.000	0.012	0.483	1.000
X.Eubacterium._hallii_group	-0.115	0.195	-0.590	0.556	-0.745	0.172	-4.336	0.000	-1.885	0.558	-3.376	0.001	0.642	0.000	0.003
X.Eubacterium._nodatum_group	0.081	0.220	0.368	0.714	0.349	0.294	1.187	0.238	0.143	0.235	0.607	0.545	0.782	0.294	0.912
X.Eubacterium._ruminantium_group	-0.762	0.310	-2.461	0.015	-0.030	0.249	-0.120	0.904	-2.502	13867.505	0.000	1.000	0.027	0.985	1.000
X.Eubacterium._saphenum_group	-0.184	0.001	-359.721	0.000	-3.715	0.011	-349.842	0.000	NA	NA	NA	NA	0.000	0.000	NA
X.Eubacterium._siraeum_group	0.300	0.252	1.193	0.235	0.012	0.155	0.075	0.941	-1.218	13867.505	0.000	1.000	0.318	1.000	1.000
X.Eubacterium._ventriosum_group	-0.512	0.131	-3.914	0.000	0.828	0.125	6.645	0.000	-0.763	13867.505	0.000	1.000	0.000	0.000	1.000
X.Eubacterium._xylanophilum_group	-0.296	0.077	-3.842	0.000	-0.851	0.077	-11.045	0.000	-1.865	13867.505	0.000	1.000	0.000	0.000	1.000
X.Eubacterium._yurii_group	1.274	0.001	986.241	0.000	-1.342	0.002	-761.399	0.000	-0.940	0.001	-1081.846	0.000	0.000	0.000	0.000
X.Ruminococcus._gauvreauii_group	-0.588	0.225	-2.616	0.010	1.114	0.185	6.034	0.000	-0.067	0.942	-0.071	0.943	0.019	0.000	1.000
X.Ruminococcus._gnavus_group	-0.761	0.198	-3.849	0.000	1.012	0.198	5.114	0.000	-0.548	13867.505	0.000	1.000	0.000	0.000	1.000
X.Ruminococcus._torques_group	-0.226	0.167	-1.352	0.179	-0.011	0.137	-0.082	0.935	-1.106	13867.505	0.000	1.000	0.258	1.000	1.000
Abiotrophia	0.476	0.066	7.247	0.000	2.506	0.373	6.725	0.000	3.403	0.327	10.391	0.000	0.000	0.000	0.000
Absconditabacteriales_.SR1.	0.728	0.015	49.509	0.000	-0.334	0.214	-1.561	0.122	4.347	0.108	40.186	0.000	0.000	0.158	0.000
Acidaminococcus	0.974	0.273	3.569	0.000	-0.850	0.172	-4.952	0.000	-2.580	13867.505	0.000	1.000	0.001	0.000	1.000
Acinetobacter	-0.648	0.094	-6.901	0.000	-1.239	0.160	-7.757	0.000	-2.679	0.202	-13.260	0.000	0.000	0.000	0.000
Actinobacillus	0.258	0.000	1409.310	0.000	1.523	15249.857	0.000	1.000	3.491	0.003	1366.794	0.000	0.000	1.000	0.000
Actinomyces	-0.421	0.090	-4.656	0.000	-0.392	0.220	-1.784	0.078	2.415	0.174	13.852	0.000	0.000	0.106	0.000
Adlercreutzia	-0.872	0.152	-5.736	0.000	2.087	0.182	11.440	0.000	-0.417	13867.505	0.000	1.000	0.000	0.000	1.000
Aggregatibacter	0.033	0.188	0.176	0.860	2.689	0.856	3.143	0.002	2.018	0.312	6.466	0.000	0.911	0.004	0.000
Akkermansia	-0.403	0.221	-1.819	0.072	0.747	0.195	3.840	0.000	-0.819	0.522	-1.568	0.120	0.114	0.000	0.233
Alistipes	-0.257	0.098	-2.628	0.010	1.205	0.094	12.770	0.000	-2.262	0.985	-2.296	0.024	0.019	0.000	0.052
Allisonella	0.794	0.155	5.123	0.000	-1.203	0.145	-8.289	0.000	-2.649	13867.505	0.000	1.000	0.000	0.000	1.000
Alloprevotella	0.222	0.194	1.142	0.256	-0.913	0.287	-3.177	0.002	-0.846	0.188	-4.487	0.000	0.342	0.003	0.000
Allorhizobium.Neorhizobium.Pararhizobium.Rhizobium	-0.920	0.024	-38.098	0.000	-2.633	0.098	-26.934	0.000	-5.352	0.341	-15.710	0.000	0.000	0.000	0.000

Alloscardovia	-1.646	0.024	-67.667	0.000	2.421	0.086	28.014	0.000	2.220	0.084	26.445	0.000	0.000	0.000	0.000
Anaerococcus	0.401	0.449	0.892	0.374	-1.003	0.319	-3.146	0.002	-2.683	0.604	-4.444	0.000	0.471	0.003	0.000
Anaeroglobus	-0.808	0.001	-1529.436	0.000	-3.116	0.003	-1213.796	0.000	NA	NA	NA	NA	0.000	0.000	NA
Anaerostipes	-1.178	0.156	-7.528	0.000	1.137	0.144	7.921	0.000	-0.166	0.667	-0.249	0.804	0.000	0.000	1.000
Anaerotruncus	-0.581	0.098	-5.945	0.000	-0.169	0.096	-1.763	0.080	-1.656	13867.505	0.000	1.000	0.000	0.109	1.000
Atopobium	0.156	0.145	1.081	0.282	0.065	0.377	0.173	0.863	1.909	0.345	5.536	0.000	0.373	0.945	0.000
Bacillus	0.635	0.217	2.935	0.004	0.632	0.161	3.933	0.000	-3.254	0.398	-8.183	0.000	0.008	0.000	0.000
Bacteroides	0.123	0.138	0.890	0.376	-0.170	0.121	-1.399	0.165	-3.234	0.252	-12.853	0.000	0.471	0.210	0.000
Barnesiella	-0.325	0.146	-2.217	0.029	0.823	0.131	6.291	0.000	-0.574	13867.505	0.000	1.000	0.049	0.000	1.000
Bergeyella	0.259	0.039	6.577	0.000	-1.514	0.474	-3.195	0.002	3.094	0.186	16.604	0.000	0.000	0.003	0.000
Bifidobacterium	0.042	0.210	0.202	0.840	0.539	0.194	2.783	0.006	-0.603	0.222	-2.711	0.008	0.898	0.010	0.018
Bilophila	-0.707	0.137	-5.153	0.000	-1.235	0.146	-8.443	0.000	-0.896	0.565	-1.586	0.116	0.000	0.000	0.227
Blautia	-0.442	0.154	-2.875	0.005	0.189	0.136	1.385	0.169	-1.008	0.964	-1.046	0.298	0.010	0.213	0.541
Bradyrhizobium	-1.077	0.000	-3512.959	0.000	0.983	0.001	1793.969	0.000	-4.348	0.017	-249.653	0.000	0.000	0.000	0.000
Butyricicoccus	-0.334	0.177	-1.888	0.062	-0.064	0.149	-0.431	0.667	-1.315	13867.505	0.000	1.000	0.099	0.757	1.000
Butyricimonas	-0.743	0.154	-4.835	0.000	0.706	0.139	5.079	0.000	-0.630	13867.505	0.000	1.000	0.000	0.000	1.000
Butyrivibrio	0.211	0.312	0.676	0.500	0.085	0.310	0.273	0.786	-1.241	0.305	-4.068	0.000	0.599	0.875	0.000
CAG.56	-0.906	0.145	-6.241	0.000	-0.435	0.153	-2.849	0.005	-1.729	13867.505	0.000	1.000	0.000	0.008	1.000
Campylobacter	0.262	0.218	1.198	0.234	-1.202	0.372	-3.233	0.002	0.097	0.183	0.528	0.598	0.318	0.003	0.992
Candidatus_Saccharimonas	-0.241	0.034	-7.025	0.000	-0.240	0.397	-0.604	0.547	3.983	0.351	11.354	0.000	0.000	0.634	0.000
Candidatus_Soleaferrea	-1.270	0.019	-66.593	0.000	-1.249	0.026	-47.895	0.000	-3.059	13867.505	0.000	1.000	0.000	0.000	1.000
Candidatus_Udaeobacter	-0.068	0.000	-181.554	0.000	-4.491	0.008	-530.788	0.000	-3.222	0.004	-828.219	0.000	0.000	0.000	0.000
Capnocytophaga	0.555	0.116	4.774	0.000	-3.807	0.670	-5.680	0.000	-1.789	0.101	-17.740	0.000	0.000	0.000	0.000
Cardiobacterium	-0.016	0.002	-9.841	0.000	0.516	0.020	25.598	0.000	2.557	0.009	298.742	0.000	0.000	0.000	0.000
Catenibacterium	-0.519	0.498	-1.043	0.299	1.579	0.327	4.834	0.000	-1.400	13867.505	0.000	1.000	0.390	0.000	1.000
Catonella	0.102	0.128	0.801	0.425	-0.205	0.224	-0.917	0.361	-1.731	0.086	-20.025	0.000	0.524	0.430	0.000
Chloroplast	0.346	0.127	2.725	0.007	-2.368	0.108	-21.921	0.000	-4.062	0.170	-23.910	0.000	0.015	0.000	0.000
Christensenella	-1.072	0.030	-35.463	0.000	1.999	0.052	38.241	0.000	-0.342	13867.505	0.000	1.000	0.000	0.000	1.000
Christensenellaceae_R.7_group	0.335	0.226	1.483	0.141	0.485	0.159	3.041	0.003	-1.593	0.991	-1.608	0.111	0.209	0.005	0.220
Chryseobacterium	-0.619	0.234	-2.647	0.009	-3.988	0.423	-9.435	0.000	-4.066	0.255	-15.919	0.000	0.018	0.000	0.000
Citrobacter	0.827	0.209	3.957	0.000	1.194	0.107	11.147	0.000	-2.429	13867.505	0.000	1.000	0.000	0.000	1.000
Cloacibacillus	-0.483	0.612	-0.790	0.431	0.225	0.295	0.763	0.447	-2.589	13867.505	0.000	1.000	0.528	0.523	1.000
Clostridia_UCG.014	-0.091	0.177	-0.516	0.607	0.705	0.204	3.451	0.001	0.221	0.194	1.141	0.256	0.686	0.001	0.469
Clostridia_vadinBB60_group	0.164	0.272	0.605	0.547	0.003	0.207	0.013	0.989	-1.434	0.304	-4.717	0.000	0.637	1.000	0.000

Clostridium_sensu_stricto_1	-0.515	0.121	-4.265	0.000	-0.457	0.119	-3.838	0.000	-0.431	13867.505	0.000	1.000	0.000	0.000	1.000
Colidextribacter	-0.283	0.174	-1.626	0.107	-0.309	0.152	-2.027	0.045	-1.311	13867.505	0.000	1.000	0.165	0.062	1.000
Collinsella	-0.149	0.145	-1.029	0.306	1.232	0.120	10.302	0.000	-0.665	0.496	-1.339	0.184	0.392	0.000	0.351
Comamonas	1.395	0.003	424.108	0.000	-1.865	0.003	-680.444	0.000	-1.017	0.001	-1487.781	0.000	0.000	0.000	0.000
Coprobacillus	-1.178	0.001	-950.856	0.000	0.451	0.002	299.961	0.000	-2.443	13867.505	0.000	1.000	0.000	0.000	1.000
Coprobacter	0.307	0.239	1.287	0.201	0.442	0.160	2.769	0.007	-0.945	13867.505	0.000	1.000	0.281	0.010	1.000
Coprococcus	-0.275	0.170	-1.621	0.108	-0.296	0.148	-2.000	0.048	-1.577	0.972	-1.622	0.108	0.165	0.066	0.215
Corynebacterium	0.113	0.184	0.610	0.543	-1.306	0.455	-2.872	0.005	1.451	0.208	6.978	0.000	0.637	0.008	0.000
Cutibacterium	-0.347	0.037	-9.273	0.000	-4.437	0.271	-16.343	0.000	-3.573	0.250	-14.275	0.000	0.000	0.000	0.000
Defluviitaleaceae_UCG.011	-0.334	0.079	-4.210	0.000	-1.122	0.091	-12.302	0.000	2.000	0.052	38.774	0.000	0.000	0.000	0.000
Delftia	0.021	0.008	2.451	0.016	-0.514	0.224	-2.291	0.024	-3.935	0.106	-37.188	0.000	0.029	0.034	0.000
Desulfosarcina	1.160	0.094	12.357	0.000	-4.781	0.217	-22.005	0.000	-4.302	0.365	-11.791	0.000	0.000	0.000	0.000
Desulfovibrio	0.035	0.269	0.130	0.897	-0.691	0.186	-3.710	0.000	-1.997	0.347	-5.761	0.000	0.938	0.001	0.000
Dialister	0.024	0.192	0.127	0.899	0.703	0.169	4.167	0.000	-0.416	0.190	-2.192	0.031	0.938	0.000	0.066
Dorea	-0.007	0.164	-0.041	0.967	-0.560	0.132	-4.238	0.000	-1.838	0.967	-1.900	0.060	0.982	0.000	0.126
DTU014	-0.143	0.243	-0.590	0.556	0.267	0.161	1.658	0.100	-1.959	13867.505	0.000	1.000	0.642	0.131	1.000
DTU089	-0.436	0.060	-7.271	0.000	1.382	0.070	19.655	0.000	-0.942	13867.505	0.000	1.000	0.000	0.000	1.000
Eggerthella	0.114	0.183	0.623	0.534	0.721	0.189	3.821	0.000	-1.390	13867.505	0.000	1.000	0.633	0.000	1.000
Eikenella	-0.428	0.202	-2.118	0.037	-2.409	15249.857	0.000	1.000	-2.178	0.122	-17.920	0.000	0.061	1.000	0.000
Eisenbergiella	-1.267	0.239	-5.294	0.000	0.405	0.257	1.575	0.118	-1.310	13867.505	0.000	1.000	0.000	0.154	1.000
Enhydrobacter	2.336	0.009	265.781	0.000	0.667	0.001	544.046	0.000	-2.971	0.016	-187.622	0.000	0.000	0.000	0.000
Enterobacter	0.380	0.172	2.213	0.029	1.457	0.122	11.948	0.000	-0.887	13867.505	0.000	1.000	0.049	0.000	1.000
Enterococcus	0.302	0.387	0.782	0.436	-0.220	0.310	-0.708	0.480	-0.975	0.740	-1.318	0.190	0.531	0.559	0.360
Enterorhabdus	-1.066	0.004	-255.191	0.000	1.008	0.017	59.699	0.000	-0.729	13867.505	0.000	1.000	0.000	0.000	1.000
Erysipelatoclostridium	-0.733	0.225	-3.255	0.002	-0.057	0.211	-0.270	0.788	-1.001	0.994	-1.007	0.316	0.003	0.875	0.564
Erysipelotrichaceae	2.172	0.002	921.575	0.000	-0.063	0.002	-32.188	0.000	-3.129	13867.505	0.000	1.000	0.000	0.000	1.000
Erysipelotrichaceae_UCG.003	-0.343	0.143	-2.402	0.018	1.441	0.140	10.304	0.000	-0.811	0.957	-0.847	0.399	0.032	0.000	0.694
Escherichia.Shigella	-0.532	0.183	-2.913	0.004	-1.187	0.184	-6.443	0.000	-1.953	0.380	-5.133	0.000	0.009	0.000	0.000
Eubacterium	NA	NA	NA	NA	-3.267	0.001	-3372.213	0.000	-4.043	13867.505	0.000	1.000	NA	0.000	1.000
F0058	0.241	0.000	758.866	0.000	2.260	15249.857	0.000	1.000	3.592	0.003	1226.644	0.000	0.000	1.000	0.000
F0332	-0.082	0.014	-6.001	0.000	0.271	13736.056	0.000	1.000	2.142	13736.056	0.000	1.000	0.000	1.000	1.000
Faecalibacterium	-0.718	0.119	-6.006	0.000	-0.135	0.111	-1.223	0.224	-2.402	0.576	-4.169	0.000	0.000	0.279	0.000
Family_XIII_AD3011_group	0.300	0.249	1.205	0.231	0.427	0.184	2.316	0.022	-1.203	0.979	-1.230	0.221	0.317	0.032	0.412
Family_XIII_UCG.001	0.074	0.191	0.387	0.700	-0.643	0.130	-4.932	0.000	-1.474	0.226	-6.527	0.000	0.774	0.000	0.000

Filifactor	-0.581	0.157	-3.714	0.000	-3.853	0.466	-8.265	0.000	-3.669	0.097	-37.865	0.000	0.001	0.000	0.000
Flavobacterium	-5.406	0.002	-2859.634	0.000	-7.454	0.008	-901.313	0.000	-4.443	0.005	-952.147	0.000	0.000	0.000	0.000
Flavonifractor	-0.346	0.107	-3.223	0.002	-1.021	0.120	-8.493	0.000	-1.243	13867.505	0.000	1.000	0.004	0.000	1.000
Fretibacterium	-0.238	0.093	-2.572	0.011	-2.229	0.477	-4.675	0.000	-0.123	0.119	-1.036	0.302	0.021	0.000	0.543
Frisingicoccus	1.816	178940.051	0.000	1.000	-0.973	0.284	-3.431	0.001	-2.522	14469.264	0.000	1.000	1.000	0.001	1.000
Fusicatenibacter	-0.154	0.114	-1.351	0.180	-1.729	0.120	-14.383	0.000	-1.848	0.868	-2.130	0.036	0.258	0.000	0.076
Fusobacterium	0.447	0.217	2.057	0.042	-1.282	0.259	-4.940	0.000	0.112	0.182	0.613	0.541	0.069	0.000	0.912
Gastranaerophilales	1.195	0.632	1.891	0.061	0.128	0.232	0.554	0.581	-1.787	13867.505	0.000	1.000	0.099	0.666	1.000
GCA.900066575	-1.011	0.042	-23.854	0.000	-4.055	0.135	-30.110	0.000	-0.970	13867.505	0.000	1.000	0.000	0.000	1.000
Gemella	-0.028	0.130	-0.214	0.831	-0.787	0.226	-3.487	0.001	1.734	0.165	10.515	0.000	0.897	0.001	0.000
Gracilibacteria	2.049	0.003	795.095	0.000	-1.329	15249.857	0.000	1.000	1.505	0.006	250.529	0.000	0.000	1.000	0.000
Granulicatella	-0.007	0.081	-0.083	0.934	-0.735	0.229	-3.202	0.002	2.657	0.162	16.381	0.000	0.956	0.003	0.000
Haemophilus	-0.361	0.096	-3.761	0.000	-0.691	0.270	-2.558	0.012	2.762	0.168	16.420	0.000	0.001	0.018	0.000
Holdemanella	0.384	0.273	1.410	0.161	1.002	0.221	4.541	0.000	-1.136	13867.505	0.000	1.000	0.234	0.000	1.000
Holdemania	-1.104	0.162	-6.813	0.000	-0.418	0.145	-2.893	0.005	-2.124	13867.505	0.000	1.000	0.000	0.007	1.000
Howardella	-0.234	0.227	-1.031	0.305	0.040	0.177	0.224	0.823	-1.668	0.263	-6.336	0.000	0.392	0.909	0.000
Hungatella	-1.623	0.122	-13.334	0.000	-3.036	0.116	-26.074	0.000	-1.924	13867.505	0.000	1.000	0.000	0.000	1.000
Hydrogenoanaerobacterium	1.755	178940.051	0.000	1.000	0.493	0.342	1.443	0.151	-1.871	14469.264	0.000	1.000	1.000	0.194	1.000
Incertae_Sedis	-0.474	0.102	-4.632	0.000	0.399	0.093	4.294	0.000	-0.661	13867.505	0.000	1.000	0.000	0.000	1.000
Intestinibacter	-0.527	0.222	-2.374	0.019	-0.460	0.203	-2.266	0.025	-1.761	13867.505	0.000	1.000	0.033	0.035	1.000
Intestinimonas	0.556	0.215	2.582	0.011	-0.901	0.163	-5.538	0.000	-2.045	13867.505	0.000	1.000	0.020	0.000	1.000
Izemoplasmatales	-0.562	0.142	-3.960	0.000	0.342	0.164	2.089	0.039	-0.183	0.301	-0.608	0.544	0.000	0.054	0.912
JGI_0000069.P22	0.259	0.001	372.384	0.000	-2.308	15249.857	0.000	1.000	NA	NA	NA	NA	0.000	1.000	NA
Johnsonella	-0.338	0.060	-5.627	0.000	-4.240	0.416	-10.196	0.000	-1.601	0.083	-19.327	0.000	0.000	0.000	0.000
Kingella	0.084	0.000	355.694	0.000	3.675	15249.857	0.000	1.000	5.220	0.005	1137.069	0.000	0.000	1.000	0.000
Klebsiella	-0.689	0.265	-2.599	0.010	2.324	0.263	8.839	0.000	-0.204	13867.505	0.000	1.000	0.019	0.000	1.000
Kocuria	3.030	0.001	3657.821	0.000	-0.922	0.001	-1018.519	0.000	-4.274	0.002	-2621.983	0.000	0.000	0.000	0.000
Lachnoanaerobaculum	0.166	0.099	1.671	0.098	-1.465	0.523	-2.800	0.006	0.457	0.121	3.794	0.000	0.154	0.009	0.001
Lachnoclostridium	0.150	0.148	1.014	0.313	-1.667	0.144	-11.554	0.000	-1.176	13867.505	0.000	1.000	0.399	0.000	1.000
Lachnospira	-0.429	0.172	-2.492	0.014	-1.943	0.184	-10.540	0.000	-2.395	0.962	-2.489	0.014	0.026	0.000	0.032
Lachnospiraceae_FCS020_group	-0.454	0.217	-2.098	0.038	0.398	0.173	2.307	0.023	-1.724	13867.505	0.000	1.000	0.063	0.032	1.000
Lachnospiraceae_ND3007_group	-0.215	0.138	-1.558	0.122	-2.880	0.154	-18.647	0.000	-2.310	13867.505	0.000	1.000	0.185	0.000	1.000

Lachnospiraceae_NK4A136_group	-0.324	0.195	-1.663	0.099	-0.072	0.151	-0.479	0.633	-0.879	13867.505	0.000	1.000	0.155	0.722	1.000
Lachnospiraceae_UCG.001	-1.806	0.001	-2147.167	0.000	-0.491	0.001	-640.474	0.000	-2.904	13867.505	0.000	1.000	0.000	0.000	1.000
Lachnospiraceae_UCG.004	0.540	0.103	5.238	0.000	-1.817	0.157	-11.547	0.000	-1.524	13867.505	0.000	1.000	0.000	0.000	1.000
Lachnospiraceae_UCG.008	-2.060	0.001	-2737.011	4.315359607133 64e-311	-0.993	0.001	-1337.184	0.000	-3.125	13867.505	0.000	1.000	0.000	0.000	1.000
Lactobacillus	-0.506	0.168	-3.006	0.003	-0.365	0.211	-1.731	0.086	-0.649	0.217	-2.986	0.004	0.007	0.116	0.009
Lactococcus	0.506	0.212	2.383	0.019	1.021	0.139	7.327	0.000	0.913	0.174	5.259	0.000	0.033	0.000	0.000
Lautropia	0.661	0.033	20.306	0.000	-2.766	0.185	-14.973	0.000	0.096	0.081	1.185	0.239	0.000	0.000	0.440
Lentimicrobium	0.416	0.001	466.657	0.000	-2.486	15249.857	0.000	1.000	NA	NA	NA	NA	0.000	1.000	NA
Leptotrichia	0.052	0.120	0.431	0.667	-1.239	0.414	-2.993	0.003	1.753	0.170	10.296	0.000	0.742	0.006	0.000
Lysinibacillus	-1.069	0.012	-90.631	0.000	-4.236	0.108	-39.273	0.000	-2.248	0.113	-19.844	0.000	0.000	0.000	0.000
Mailhella	0.976	0.173	5.628	0.000	-2.582	0.201	-12.844	0.000	-3.828	13867.505	0.000	1.000	0.000	0.000	1.000
Marvinbryantia	0.653	0.230	2.839	0.005	-0.910	0.117	-7.791	0.000	-1.813	13867.505	0.000	1.000	0.011	0.000	1.000
Megasphaera	-0.147	0.114	-1.292	0.199	-0.836	0.290	-2.882	0.005	1.468	0.190	7.716	0.000	0.280	0.008	0.000
Merdibacter	0.600	0.017	35.216	0.000	-1.089	0.010	-107.436	0.000	-3.016	13867.505	0.000	1.000	0.000	0.000	1.000
Mesorhizobium	-1.826	0.000	-5624.815	0.000	-2.009	15249.857	0.000	1.000	-2.024	13867.505	0.000	1.000	0.000	1.000	1.000
Methanobrevibacter	-0.754	0.182	-4.136	0.000	1.776	0.184	9.664	0.000	-1.718	0.585	-2.935	0.004	0.000	0.000	0.010
Micrococcus	-1.133	0.342	-3.312	0.001	-2.823	15249.857	0.000	1.000	-1.877	0.695	-2.699	0.008	0.003	1.000	0.018
Mitochondria	-0.922	0.001	-638.701	0.000	2.794	0.002	1813.764	0.000	0.723	0.002	431.202	0.000	0.000	0.000	0.000
Mitsuokella	0.242	0.224	1.081	0.282	-0.932	0.277	-3.366	0.001	-3.109	13867.505	0.000	1.000	0.373	0.002	1.000
Mogibacterium	-0.049	0.069	-0.714	0.477	-1.908	0.075	-25.467	0.000	-1.870	0.054	-34.782	0.000	0.574	0.000	0.000
Monoglobus	-1.273	0.104	-12.201	0.000	1.119	0.128	8.712	0.000	-0.219	13867.505	0.000	1.000	0.000	0.000	1.000
Muribaculaceae	-0.751	0.177	-4.243	0.000	1.001	0.136	7.370	0.000	-2.930	0.989	-2.964	0.004	0.000	0.000	0.009
Mycoplasma	-0.189	0.000	-610.591	0.000	-1.679	15249.857	0.000	1.000	NA	NA	NA	NA	0.000	1.000	NA
Negativibacillus	-0.382	0.202	-1.886	0.062	-1.087	0.154	-7.068	0.000	-2.068	13867.505	0.000	1.000	0.099	0.000	1.000
Neisseria	0.114	0.024	4.743	0.000	0.906	0.299	3.036	0.003	4.492	0.125	36.002	0.000	0.000	0.005	0.000
NK4A214_group	-0.497	0.130	-3.817	0.000	0.848	0.128	6.605	0.000	-0.444	13867.505	0.000	1.000	0.001	0.000	1.000
Ochrobactrum	0.325	0.000	701.689	0.000	-3.278	15249.857	0.000	1.000	-4.631	0.007	-687.745	0.000	0.000	1.000	0.000
Odoribacter	-0.231	0.160	-1.438	0.154	-0.402	0.133	-3.011	0.003	-1.784	0.439	-4.065	0.000	0.226	0.005	0.000
Olsenella	0.267	0.271	0.988	0.325	0.803	0.308	2.608	0.010	0.114	0.323	0.354	0.724	0.412	0.015	1.000
Oribacterium	0.276	0.133	2.072	0.041	-0.139	0.401	-0.347	0.730	1.473	0.304	4.850	0.000	0.068	0.824	0.000
Oscillibacter	0.046	0.261	0.177	0.859	0.124	0.216	0.575	0.566	-1.328	13867.505	0.000	1.000	0.911	0.653	1.000
Oscillospira	-0.718	0.177	-4.058	0.000	0.352	0.168	2.094	0.038	-1.585	13867.505	0.000	1.000	0.000	0.053	1.000
Paenibacillus	1.314	0.001	1139.464	0.000	1.414	0.001	1175.064	0.000	-3.198	0.004	-731.736	0.000	0.000	0.000	0.000

Paeniclostridium	-2.056	0.001	-2284.800	2.014460458505546e-315	-1.852	0.002	-1196.741	0.000	-3.160	13867.505	0.000	1.000	8.783047599281796e-314	0.000	1.000
Parabacteroides	-0.005	0.138	-0.038	0.970	0.783	0.112	7.001	0.000	-0.925	0.987	-0.938	0.351	0.982	0.000	0.615
Paracoccus	-0.349	0.001	-606.563	0.000	-0.004	0.001	-2.662	0.009	-3.345	0.008	-436.359	0.000	0.000	0.013	0.000
Paraprevotella	-0.101	0.180	-0.559	0.577	1.045	0.167	6.243	0.000	-0.817	0.988	-0.826	0.410	0.656	0.000	0.705
Parasutterella	-0.457	0.123	-3.701	0.000	-0.934	0.127	-7.360	0.000	-1.537	13867.505	0.000	1.000	0.001	0.000	1.000
Parvimonas	-0.375	0.246	-1.523	0.130	-0.909	0.241	-3.769	0.000	-0.697	0.202	-3.455	0.001	0.195	0.000	0.002
Peptococcus	-0.024	0.273	-0.088	0.930	-0.965	0.292	-3.304	0.001	-0.882	0.226	-3.905	0.000	0.956	0.002	0.000
Peptoniphilus	-1.377	0.140	-9.831	0.000	2.579	0.175	14.776	0.000	-0.762	0.299	-2.548	0.012	0.000	0.000	0.027
Peptostreptococcus	0.210	0.240	0.876	0.383	-0.642	0.270	-2.383	0.019	-0.601	0.193	-3.111	0.002	0.477	0.027	0.006
Phascolarctobacterium	-0.156	0.191	-0.820	0.414	-0.754	0.170	-4.437	0.000	-1.522	13867.505	0.000	1.000	0.513	0.000	1.000
Porphyromonas	0.016	0.180	0.091	0.928	-0.313	0.279	-1.123	0.264	1.665	0.213	7.830	0.000	0.956	0.323	0.000
Prevotella	-0.026	0.188	-0.138	0.891	-0.217	0.229	-0.947	0.345	1.195	0.188	6.344	0.000	0.938	0.413	0.000
Prevotellaceae_NK3B31_group	-0.826	0.169	-4.895	0.000	-0.680	0.170	-3.992	0.000	-0.805	0.977	-0.824	0.412	0.000	0.000	0.705
Prevotellaceae_UCG.001	-0.370	0.211	-1.753	0.082	1.319	0.232	5.690	0.000	-2.079	0.998	-2.083	0.039	0.129	0.000	0.083
Pseudoalteromonas	-4.583	0.001	-32001.124	0.000	-1.845	0.009	-214.039	0.000	-1.455	0.007	-223.081	0.000	0.000	0.000	0.000
Pseudomonas	-0.886	0.095	-9.335	0.000	-2.323	0.102	-22.755	0.000	-1.956	0.224	-8.718	0.000	0.000	0.000	0.000
Pyramidobacter	1.045	0.360	2.899	0.004	1.580	0.116	13.593	0.000	-0.330	0.268	-1.234	0.220	0.009	0.000	0.412
RF39	-0.713	0.117	-6.085	0.000	1.062	0.386	2.754	0.007	-1.110	0.396	-2.805	0.006	0.000	0.010	0.014
Rheinheimera	-0.928	0.015	-61.474	0.000	-2.403	0.055	-43.441	0.000	-2.888	0.075	-38.366	0.000	0.000	0.000	0.000
Rhodococcus	2.904	0.004	697.209	0.000	-2.315	15249.857	0.000	1.000	-2.216	13867.505	0.000	1.000	0.000	1.000	1.000
Rikenellaceae_RC9_gut_group	-0.176	0.134	-1.313	0.192	1.482	0.121	12.201	0.000	-3.100	0.213	-14.531	0.000	0.272	0.000	0.000
Romboutsia	-1.296	0.095	-13.696	0.000	0.328	0.100	3.267	0.001	-0.625	13867.505	0.000	1.000	0.000	0.003	1.000
Roseburia	0.062	0.168	0.371	0.712	-0.847	0.151	-5.602	0.000	-2.371	0.495	-4.785	0.000	0.782	0.000	0.000
Rothia	-0.204	0.057	-3.562	0.001	-1.230	0.279	-4.402	0.000	3.523	0.160	22.084	0.000	0.001	0.000	0.000
Ruminiclostridium	0.548	0.411	1.333	0.185	0.747	0.118	6.304	0.000	-1.931	13867.505	0.000	1.000	0.263	0.000	1.000
Ruminococcus	0.449	0.135	3.333	0.001	1.885	0.122	15.409	0.000	-2.421	0.333	-7.278	0.000	0.003	0.000	0.000
S5.A14a	-0.626	0.276	-2.268	0.025	1.206	0.272	4.427	0.000	-2.519	13867.505	0.000	1.000	0.043	0.000	1.000
Saccharimonadaceae	-0.146	0.141	-1.031	0.305	0.910	0.533	1.707	0.091	1.558	0.522	2.985	0.004	0.392	0.121	0.009
Saccharimonadales	-0.161	0.046	-3.499	0.001	0.776	0.529	1.468	0.145	4.837	0.502	9.635	0.000	0.002	0.187	0.000
Sanguibacteroides	0.425	0.271	1.567	0.119	-0.757	0.227	-3.331	0.001	-3.431	13867.505	0.000	1.000	0.182	0.002	1.000
Selenomonas	0.005	0.160	0.034	0.973	-2.649	0.457	-5.800	0.000	-0.257	0.137	-1.872	0.064	0.982	0.000	0.133
Sellimonas	-1.238	0.263	-4.717	0.000	-0.626	0.245	-2.559	0.012	-3.051	13867.505	0.000	1.000	0.000	0.017	1.000

Senegalimassilia	0.651	0.225	2.891	0.005	0.247	0.149	1.664	0.099	-1.665	13867.505	0.000	1.000	0.009	0.130	1.000
Serratia	-1.054	0.146	-7.225	0.000	-0.882	0.119	-7.409	0.000	-4.433	0.323	-13.705	0.000	0.000	0.000	0.000
Slackia	-0.270	0.191	-1.416	0.160	1.941	0.133	14.566	0.000	-0.816	0.232	-3.521	0.001	0.234	0.000	0.002
Solobacterium	-0.065	0.212	-0.306	0.761	-0.285	0.270	-1.057	0.293	0.211	0.224	0.945	0.346	0.829	0.352	0.613
Sphingomonas	1.763	0.072	24.358	0.000	-1.428	15249.857	0.000	1.000	-1.420	0.190	-7.462	0.000	0.000	1.000	0.000
Staphylococcus	0.029	0.259	0.111	0.912	-1.806	0.317	-5.703	0.000	-1.553	0.308	-5.045	0.000	0.947	0.000	0.000
Stomatobaculum	-0.150	0.015	-10.245	0.000	-0.941	0.112	-8.416	0.000	1.354	0.036	37.254	0.000	0.000	0.000	0.000
Streptococcus	-0.055	0.115	-0.478	0.634	0.205	0.189	1.084	0.281	2.914	0.158	18.394	0.000	0.709	0.340	0.000
Subdoligranulum	-0.337	0.101	-3.336	0.001	1.054	0.102	10.304	0.000	-2.761	0.382	-7.226	0.000	0.003	0.000	0.000
Sutterella	0.083	0.170	0.485	0.628	-2.060	0.209	-9.840	0.000	-1.773	13867.505	0.000	1.000	0.706	0.000	1.000
Tannerella	-0.213	0.061	-3.494	0.001	0.141	0.638	0.220	0.826	2.846	0.384	7.413	0.000	0.002	0.909	0.000
Terrisporobacter	-0.137	0.210	-0.652	0.515	2.369	0.223	10.628	0.000	-0.343	13867.505	0.000	1.000	0.614	0.000	1.000
TM7a	-0.706	0.104	-6.803	0.000	2.321	8388.606	0.000	1.000	1.650	8388.606	0.000	1.000	0.000	1.000	1.000
TM7x	-0.466	0.084	-5.549	0.000	0.909	0.367	2.480	0.015	4.011	0.348	11.528	0.000	0.000	0.022	0.000
Treponema	0.171	0.281	0.607	0.545	-0.649	0.553	-1.173	0.244	0.520	0.309	1.680	0.096	0.637	0.300	0.194
Turicibacter	-0.376	0.312	-1.204	0.231	1.063	0.293	3.632	0.000	-1.189	13867.505	0.000	1.000	0.317	0.001	1.000
Tyzzerella	-0.529	0.203	-2.608	0.010	-0.734	0.182	-4.038	0.000	-1.590	13867.505	0.000	1.000	0.019	0.000	1.000
UBA1819	-0.403	0.114	-3.531	0.001	1.630	0.124	13.097	0.000	-1.289	0.952	-1.354	0.179	0.001	0.000	0.345
UCG.002	0.226	0.179	1.263	0.209	0.831	0.139	5.963	0.000	-1.231	0.707	-1.743	0.084	0.291	0.000	0.172
UCG.003	0.455	0.205	2.223	0.028	-0.192	0.138	-1.384	0.169	-1.364	13867.505	0.000	1.000	0.048	0.213	1.000
UCG.005	0.065	0.220	0.295	0.769	0.265	0.153	1.728	0.087	-1.322	0.708	-1.869	0.064	0.834	0.116	0.133
UCG.010	-0.110	0.194	-0.566	0.573	1.955	0.157	12.417	0.000	0.652	13867.505	0.000	1.000	0.654	0.000	1.000
uncultured	0.181	0.157	1.154	0.251	0.122	0.140	0.874	0.384	-1.461	0.192	-7.630	0.000	0.338	0.455	0.000
Veillonella	0.021	0.101	0.207	0.836	-0.069	0.225	-0.308	0.759	3.143	0.177	17.783	0.000	0.898	0.852	0.000
Victivallis	-0.224	0.147	-1.521	0.131	0.156	0.121	1.294	0.198	-2.089	13867.505	0.000	1.000	0.195	0.248	1.000

Machine Learning Results

The interpretable machine learning (IML) based on the random forest model.

ML1 - Biopsy

A- Phylum

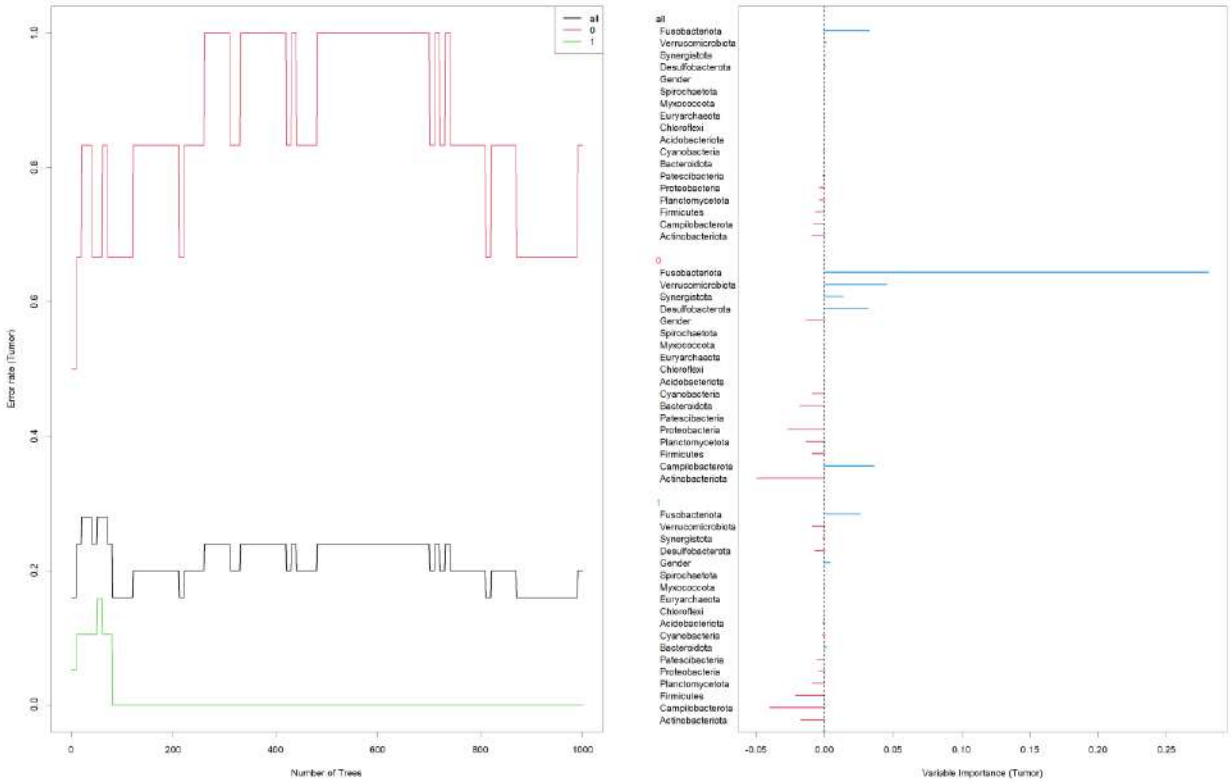


Figure ML1-A-1: The number of trees and Variable Importance Plot for **Phylum** in **Biopsy**

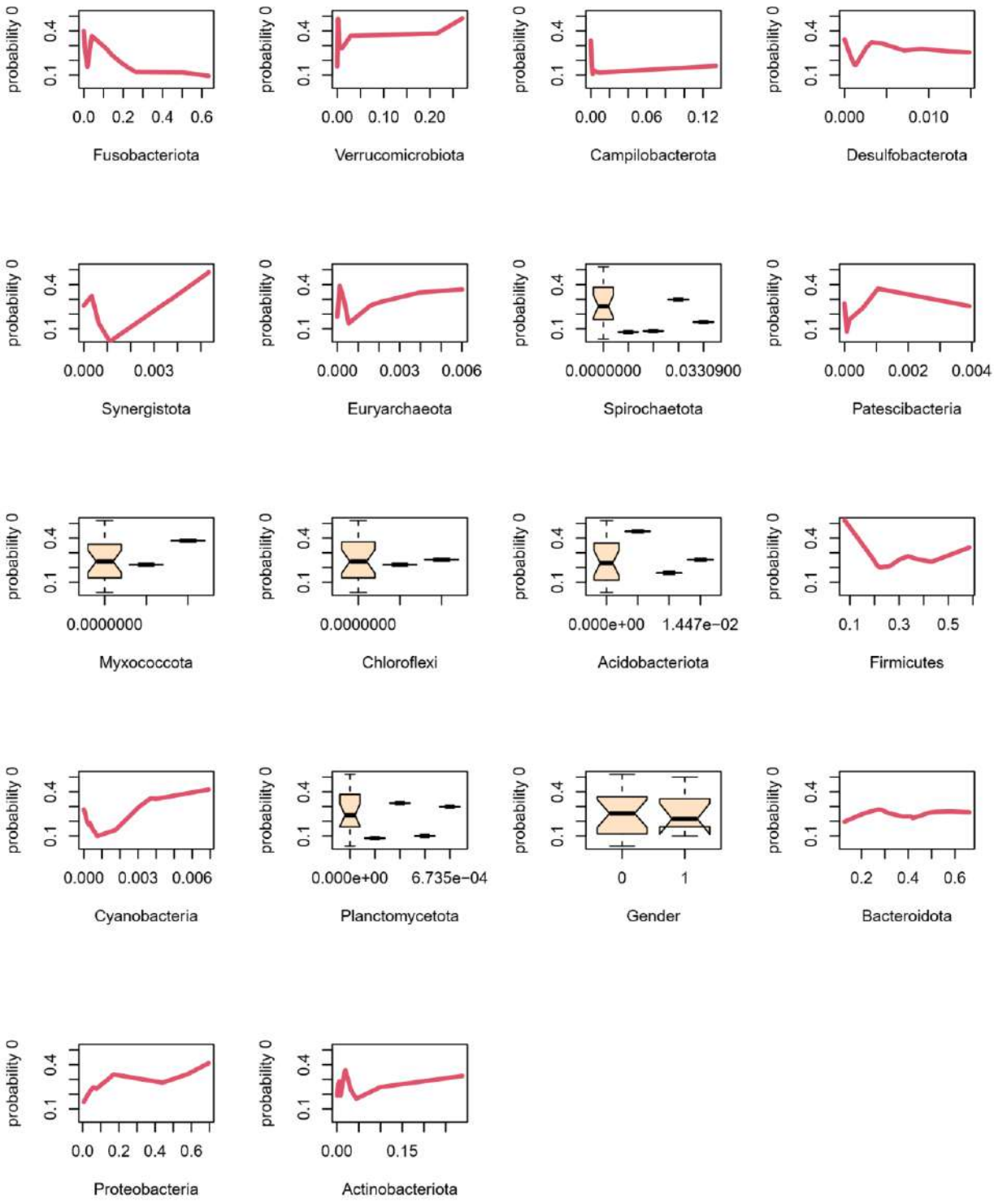


Figure ML1-A-2: Marginal Plot for **Phylum** in **Biopsy**

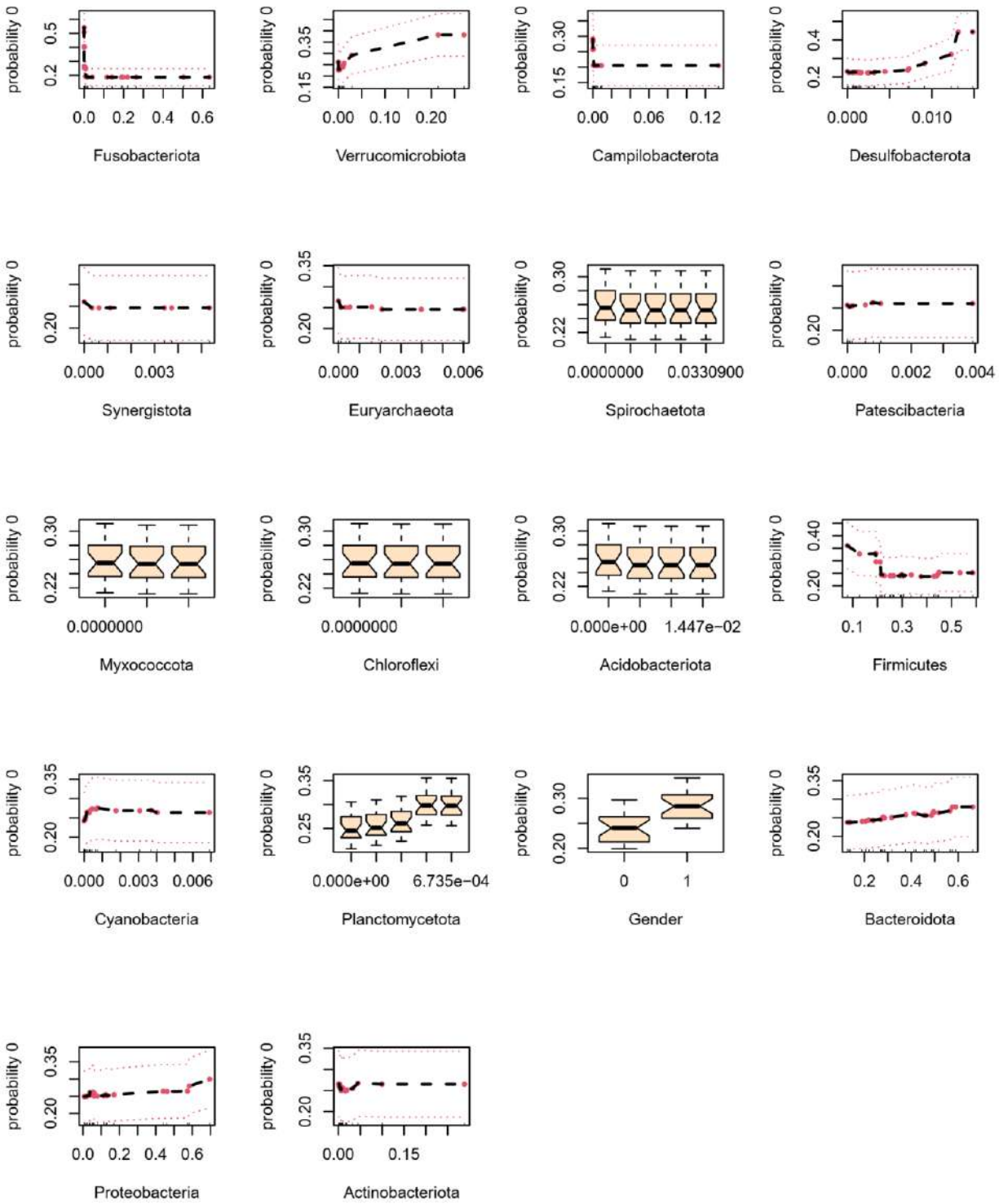


Figure ML1-A-3: Partial Plot for **Phylum** in **Biopsy**

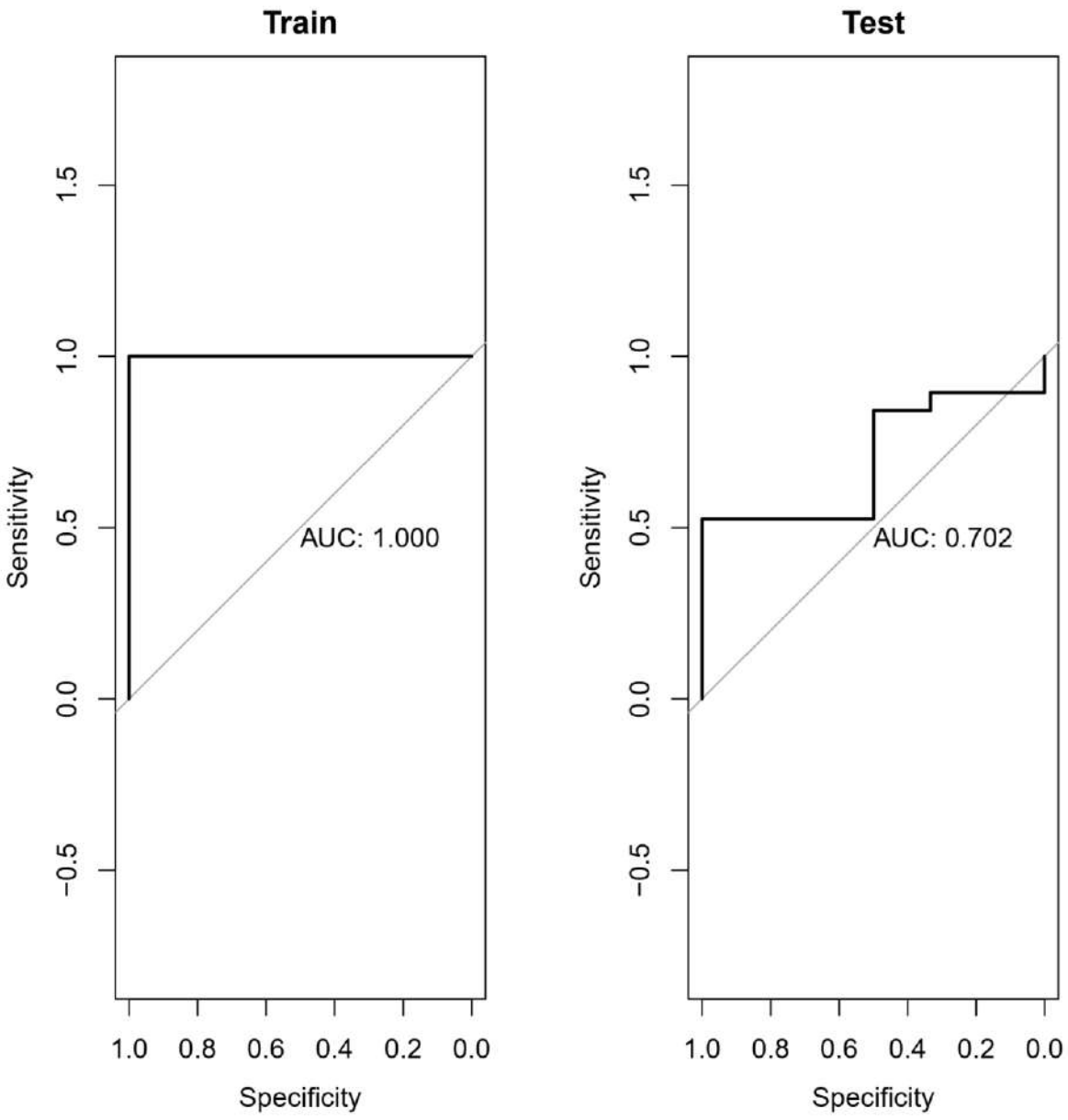


Figure ML1-A-4: The ROC Curve for **Phylum** in **Biopsy**

Table ML1-A-1: The Model Performance			
	Values		
Index	2.50%	50.00%	97.50%
Accuracy	0.52	0.68	0.92
Specificity	0.67	1.00	1.00
Sensitivity	0.37	0.58	0.95

Table ML1-A-2: The Variable Importance Values						
Covariates	Variable Importance			Brier Score		
	all	AP	CRC	all	AP	CRC
Gender	-0.0062	-0.0815	0.0043	-0.0129	-0.0087	-0.0087
Acidobacteriota	-0.0008	0.0091	-0.0057	-0.0036	-0.0025	-0.0025
Actinobacteriota	-0.0082	-0.1223	0.0100	-0.0390	-0.0265	-0.0265
Bacteroidota	-0.0163	-0.0906	-0.0286	-0.0424	-0.0288	-0.0288
Campilobacterota	-0.0119	0.1450	-0.0873	-0.0077	-0.0052	-0.0052
Chloroflexi	0.0000	0.0000	0.0000	-0.0019	-0.0013	-0.0013
Cyanobacteria	-0.0106	-0.0408	-0.0243	-0.0291	-0.0197	-0.0197
Desulfobacterota	0.1097	-0.3851	0.5093	0.3271	0.2223	0.2223
Euryarchaeota	-0.0032	0.0136	-0.0157	-0.0018	-0.0012	-0.0012
Firmicutes	0.0355	-0.2718	0.2117	0.0507	0.0345	0.0345
Fusobacteriota	0.2457	-0.1042	0.9013	0.6591	0.4479	0.4479
Myxococcota	0.0000	0.0000	0.0000	-0.0007	-0.0005	-0.0005
Patescibacteria	-0.0056	-0.0181	-0.0143	-0.0150	-0.0102	-0.0102
Planctomycetota	0.0027	-0.1314	0.0515	0.0026	0.0017	0.0017
Proteobacteria	0.0076	-0.1631	0.0787	-0.0250	-0.0170	-0.0170
Spirochaetota	-0.0028	0.0045	-0.0114	-0.0038	-0.0026	-0.0026
Synergistota	0.0012	0.0453	-0.0100	-0.0055	-0.0037	-0.0037
Verrucomicrobiota	0.0586	-0.2809	0.2961	0.1373	0.0933	0.0933

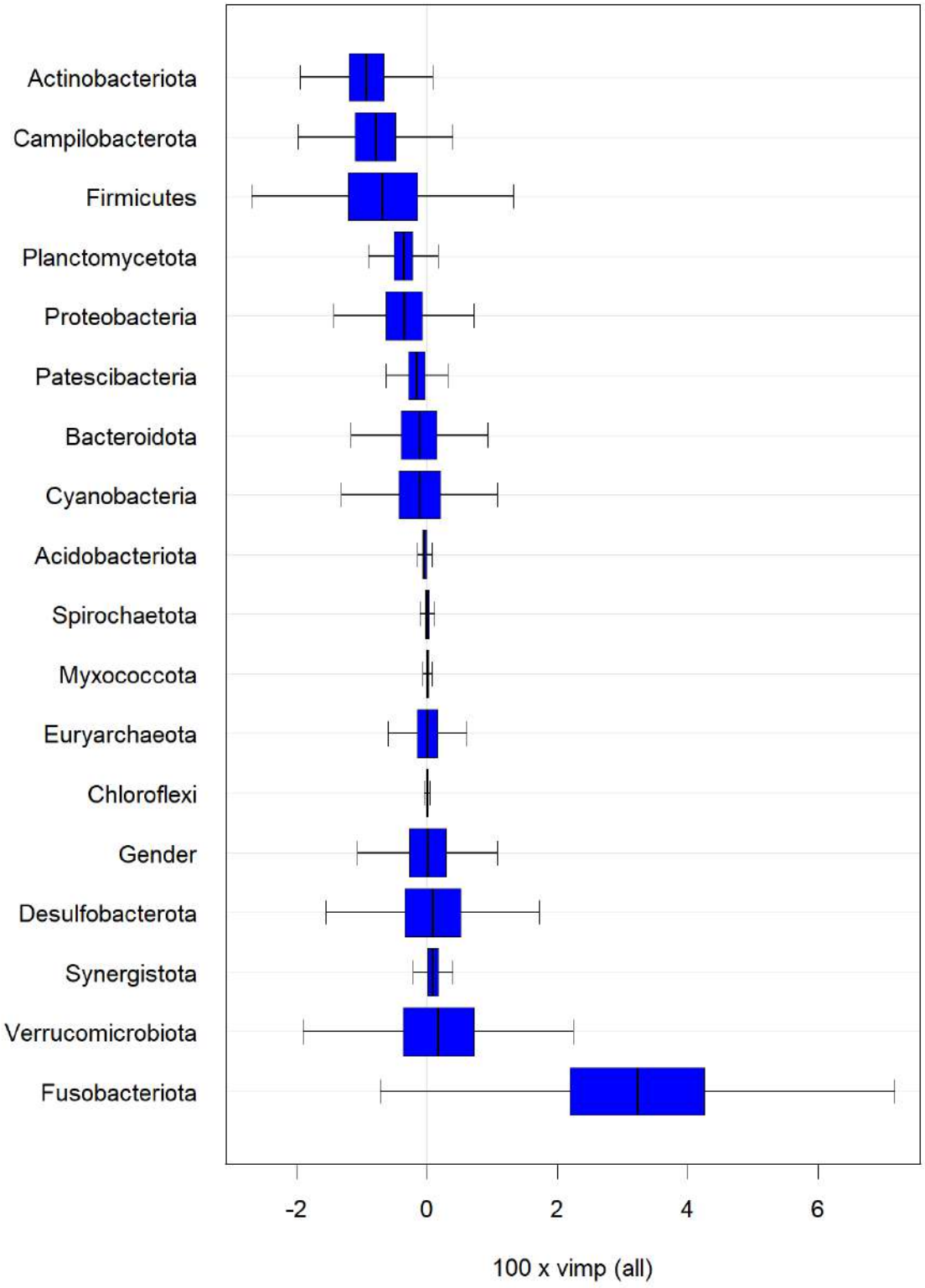


Figure ML1-A-5: The Variable Importance Plot (VIMP) for **Phylum** in **Biopsy**

Best model is selected for LIME results

Table ML1-A-3: The best setting for model and LIME Results							
Setting	**mtry**	**ROC**	**Sens**	**Spec**	**ROCSD**	**SensSD**	**SpecSD**
ntrees: 500 nodesize: 1	10	0.717	0.500	0.900	0.298	0.500	0.137
ntrees: 1000 nodesize: 1	10	0.700	0.300	0.950	0.420	0.447	0.112
ntrees: 500 nodesize: 5	8	0.700	0.300	0.900	0.447	0.447	0.137
ntrees: 500 nodesize: 5	10	0.700	0.500	0.900	0.447	0.500	0.137
ntrees: 1000 nodesize: 5	10	0.733	0.500	0.900	0.384	0.500	0.137

Table ML1-A-4: The LIME Result for Training Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
2	Adenoma	0.866	0.417	0.166	0.446	Fusobacteriota	0.000	0.188	Fusobacteriota <= 0.00111
2	Adenoma	0.866	0.417	0.166	0.446	Firmicutes	0.194	0.072	Firmicutes <= 0.222
2	Adenoma	0.866	0.417	0.166	0.446	Verrucomicrobiota	0.270	0.066	0.006252 < Verrucomicrobiota
2	Adenoma	0.866	0.417	0.166	0.446	Planctomycetota	0.000	-0.076	Planctomycetota <= 0.000168
2	Adenoma	0.866	0.417	0.166	0.446	Campilobacterota	0.000	0.031	Campilobacterota <= 0.00011
3	Adenoma	0.688	0.131	0.220	0.298	Desulfobacterota	0.013	0.085	0.004453 < Desulfobacterota
3	Adenoma	0.688	0.131	0.220	0.298	Firmicutes	0.211	0.065	Firmicutes <= 0.222
3	Adenoma	0.688	0.131	0.220	0.298	Planctomycetota	0.000	-0.070	Planctomycetota <= 0.000168
3	Adenoma	0.688	0.131	0.220	0.298	Campilobacterota	0.000	0.035	Campilobacterota <= 0.00011
3	Adenoma	0.688	0.131	0.220	0.298	Verrucomicrobiota	0.006	-0.036	0.000714 < Verrucomicrobiota <= 0.006252
6	Adenoma	0.718	0.094	0.249	0.231	Desulfobacterota	0.015	0.082	0.004453 < Desulfobacterota
6	Adenoma	0.718	0.094	0.249	0.231	Planctomycetota	0.000	-0.063	Planctomycetota <= 0.000168
6	Adenoma	0.718	0.094	0.249	0.231	Firmicutes	0.337	-0.037	0.292 < Firmicutes <= 0.376
6	Adenoma	0.718	0.094	0.249	0.231	Verrucomicrobiota	0.000	-0.026	Verrucomicrobiota <= 0.000714
6	Adenoma	0.718	0.094	0.249	0.231	Campilobacterota	0.000	0.026	Campilobacterota <= 0.00011
7	Adenoma	0.844	0.367	0.182	0.374	Fusobacteriota	0.000	0.182	Fusobacteriota <= 0.00111
7	Adenoma	0.844	0.367	0.182	0.374	Verrucomicrobiota	0.214	0.072	0.006252 < Verrucomicrobiota
7	Adenoma	0.844	0.367	0.182	0.374	Planctomycetota	0.000	-0.066	Planctomycetota <= 0.000168
7	Adenoma	0.844	0.367	0.182	0.374	Campilobacterota	0.000	0.031	Campilobacterota <= 0.00011
7	Adenoma	0.844	0.367	0.182	0.374	Firmicutes	0.298	-0.028	0.292 < Firmicutes <= 0.376
10	Adenoma	0.690	0.337	0.161	0.390	Fusobacteriota	0.000	0.188	Fusobacteriota <= 0.00111
10	Adenoma	0.690	0.337	0.161	0.390	Planctomycetota	0.000	0.072	0.000168 < Planctomycetota <= 0.000337
10	Adenoma	0.690	0.337	0.161	0.390	Verrucomicrobiota	0.000	-0.025	Verrucomicrobiota <= 0.000714
10	Adenoma	0.690	0.337	0.161	0.390	Desulfobacterota	0.001	-0.029	0.000808 < Desulfobacterota <= 0.001582
10	Adenoma	0.690	0.337	0.161	0.390	Cyanobacteria	0.001	0.024	0.000643 < Cyanobacteria
11	Adenoma	0.846	0.391	0.182	0.385	Fusobacteriota	0.000	0.186	Fusobacteriota <= 0.00111
11	Adenoma	0.846	0.391	0.182	0.385	Firmicutes	0.078	0.075	Firmicutes <= 0.222
11	Adenoma	0.846	0.391	0.182	0.385	Planctomycetota	0.000	-0.064	Planctomycetota <= 0.000168

11	Adenoma	0.846	0.391	0.182	0.385	Campilobacterota	0.000	0.029	Campilobacterota <= 0.00011
11	Adenoma	0.846	0.391	0.182	0.385	Verrucomicrobiota	0.000	-0.024	Verrucomicrobiota <= 0.000714
14	Adenocarcinoma	0.968	0.145	0.736	0.845	Fusobacteriota	0.141	0.102	0.00727 < Fusobacteriota <= 0.14117
14	Adenocarcinoma	0.968	0.145	0.736	0.845	Verrucomicrobiota	0.029	-0.058	0.006252 < Verrucomicrobiota
14	Adenocarcinoma	0.968	0.145	0.736	0.845	Planctomycetota	0.000	0.068	Planctomycetota <= 0.000168
14	Adenocarcinoma	0.968	0.145	0.736	0.845	Campilobacterota	0.000	-0.032	Campilobacterota <= 0.00011
14	Adenocarcinoma	0.968	0.145	0.736	0.845	Firmicutes	0.262	0.030	0.222 < Firmicutes <= 0.292
20	Adenocarcinoma	0.996	0.183	0.691	0.883	Fusobacteriota	0.197	0.104	0.14117 < Fusobacteriota
20	Adenocarcinoma	0.996	0.183	0.691	0.883	Desulfobacterota	0.007	-0.073	0.004453 < Desulfobacterota
20	Adenocarcinoma	0.996	0.183	0.691	0.883	Planctomycetota	0.000	0.082	Planctomycetota <= 0.000168
20	Adenocarcinoma	0.996	0.183	0.691	0.883	Firmicutes	0.295	0.048	0.292 < Firmicutes <= 0.376
20	Adenocarcinoma	0.996	0.183	0.691	0.883	Verrucomicrobiota	0.001	0.031	Verrucomicrobiota <= 0.000714
21	Adenocarcinoma	0.970	0.110	0.680	0.935	Fusobacteriota	0.220	0.093	0.14117 < Fusobacteriota
21	Adenocarcinoma	0.970	0.110	0.680	0.935	Planctomycetota	0.000	0.075	Planctomycetota <= 0.000168
21	Adenocarcinoma	0.970	0.110	0.680	0.935	Verrucomicrobiota	0.001	0.033	Verrucomicrobiota <= 0.000714
21	Adenocarcinoma	0.970	0.110	0.680	0.935	Firmicutes	0.253	0.029	0.222 < Firmicutes <= 0.292
21	Adenocarcinoma	0.970	0.110	0.680	0.935	Campilobacterota	0.008	0.025	0.00180 < Campilobacterota
22	Adenocarcinoma	0.968	0.149	0.717	0.861	Fusobacteriota	0.635	0.094	0.14117 < Fusobacteriota
22	Adenocarcinoma	0.968	0.149	0.717	0.861	Firmicutes	0.215	-0.069	Firmicutes <= 0.222
22	Adenocarcinoma	0.968	0.149	0.717	0.861	Planctomycetota	0.000	0.065	Planctomycetota <= 0.000168
22	Adenocarcinoma	0.968	0.149	0.717	0.861	Verrucomicrobiota	0.001	0.029	Verrucomicrobiota <= 0.000714
22	Adenocarcinoma	0.968	0.149	0.717	0.861	Desulfobacterota	0.001	0.025	0.000808 < Desulfobacterota <= 0.001582
23	Adenocarcinoma	0.818	0.396	0.815	0.627	Fusobacteriota	0.001	-0.187	Fusobacteriota <= 0.00111
23	Adenocarcinoma	0.818	0.396	0.815	0.627	Firmicutes	0.127	-0.074	Firmicutes <= 0.222
23	Adenocarcinoma	0.818	0.396	0.815	0.627	Planctomycetota	0.000	0.071	Planctomycetota <= 0.000168
23	Adenocarcinoma	0.818	0.396	0.815	0.627	Campilobacterota	0.000	-0.032	Campilobacterota <= 0.00011
23	Adenocarcinoma	0.818	0.396	0.815	0.627	Verrucomicrobiota	0.002	0.033	0.000714 < Verrucomicrobiota <= 0.006252
24	Adenocarcinoma	0.914	0.156	0.704	0.854	Fusobacteriota	0.503	0.094	0.14117 < Fusobacteriota
24	Adenocarcinoma	0.914	0.156	0.704	0.854	Firmicutes	0.195	-0.073	Firmicutes <= 0.222
24	Adenocarcinoma	0.914	0.156	0.704	0.854	Planctomycetota	0.000	0.075	Planctomycetota <= 0.000168
24	Adenocarcinoma	0.914	0.156	0.704	0.854	Cyanobacteria	0.000	0.028	Cyanobacteria <= 0.000029
24	Adenocarcinoma	0.914	0.156	0.704	0.854	Verrucomicrobiota	0.000	0.026	Verrucomicrobiota <= 0.000714
25	Adenocarcinoma	0.954	0.037	0.710	0.873	Planctomycetota	0.000	0.057	Planctomycetota <= 0.000168
25	Adenocarcinoma	0.954	0.037	0.710	0.873	Desulfobacterota	0.001	0.031	0.000808 < Desulfobacterota <= 0.001582
25	Adenocarcinoma	0.954	0.037	0.710	0.873	Verrucomicrobiota	0.000	0.025	Verrucomicrobiota <= 0.000714
25	Adenocarcinoma	0.954	0.037	0.710	0.873	Firmicutes	0.251	0.027	0.222 < Firmicutes <= 0.292
25	Adenocarcinoma	0.954	0.037	0.710	0.873	Cyanobacteria	0.000	0.024	Cyanobacteria <= 0.000029
26	Adenocarcinoma	0.942	0.116	0.688	0.926	Fusobacteriota	0.019	0.101	0.00727 < Fusobacteriota <= 0.14117
26	Adenocarcinoma	0.942	0.116	0.688	0.926	Planctomycetota	0.000	0.074	Planctomycetota <= 0.000168
26	Adenocarcinoma	0.942	0.116	0.688	0.926	Verrucomicrobiota	0.000	0.020	Verrucomicrobiota <= 0.000714
26	Adenocarcinoma	0.942	0.116	0.688	0.926	Campilobacterota	0.003	0.022	0.00180 < Campilobacterota
26	Adenocarcinoma	0.942	0.116	0.688	0.926	Firmicutes	0.428	0.021	0.376 < Firmicutes
29	Adenocarcinoma	0.936	0.157	0.729	0.827	Fusobacteriota	0.040	0.095	0.00727 < Fusobacteriota <= 0.14117

29	Adenocarcinoma	0.936	0.157	0.729	0.827	Desulfobacterota	0.007	-0.074	0.004453 < Desulfobacterota
29	Adenocarcinoma	0.936	0.157	0.729	0.827	Planctomycetota	0.000	0.074	Planctomycetota <= 0.000168
29	Adenocarcinoma	0.936	0.157	0.729	0.827	Verrucomicrobiota	0.000	0.030	Verrucomicrobiota <= 0.000714
29	Adenocarcinoma	0.936	0.157	0.729	0.827	Campilobacterota	0.000	-0.028	Campilobacterota <= 0.00011
32	Adenocarcinoma	0.980	0.165	0.706	0.856	Fusobacteriota	0.263	0.100	0.14117 < Fusobacteriota
32	Adenocarcinoma	0.980	0.165	0.706	0.856	Desulfobacterota	0.009	-0.081	0.004453 < Desulfobacterota
32	Adenocarcinoma	0.980	0.165	0.706	0.856	Planctomycetota	0.000	0.069	Planctomycetota <= 0.000168
32	Adenocarcinoma	0.980	0.165	0.706	0.856	Cyanobacteria	0.000	0.033	Cyanobacteria <= 0.000029
32	Adenocarcinoma	0.980	0.165	0.706	0.856	Firmicutes	0.229	0.028	0.222 < Firmicutes <= 0.292
33	Adenocarcinoma	0.878	0.051	0.736	0.861	Planctomycetota	0.000	0.067	Planctomycetota <= 0.000168
33	Adenocarcinoma	0.878	0.051	0.736	0.861	Campilobacterota	0.000	-0.032	Campilobacterota <= 0.00011
33	Adenocarcinoma	0.878	0.051	0.736	0.861	Firmicutes	0.376	0.035	0.292 < Firmicutes <= 0.376
33	Adenocarcinoma	0.878	0.051	0.736	0.861	Verrucomicrobiota	0.001	0.028	0.000714 < Verrucomicrobiota <= 0.006252
33	Adenocarcinoma	0.878	0.051	0.736	0.861	Desulfobacterota	0.003	0.028	0.001582 < Desulfobacterota <= 0.004453
34	Adenocarcinoma	0.932	0.149	0.758	0.786	Fusobacteriota	0.116	0.095	0.00727 < Fusobacteriota <= 0.14117
34	Adenocarcinoma	0.932	0.149	0.758	0.786	Verrucomicrobiota	0.009	-0.069	0.006252 < Verrucomicrobiota
34	Adenocarcinoma	0.932	0.149	0.758	0.786	Campilobacterota	0.000	-0.037	Campilobacterota <= 0.00011
34	Adenocarcinoma	0.932	0.149	0.758	0.786	Planctomycetota	0.000	0.069	Planctomycetota <= 0.000168
34	Adenocarcinoma	0.932	0.149	0.758	0.786	Cyanobacteria	0.003	-0.031	0.000643 < Cyanobacteria
39	Adenocarcinoma	0.828	0.098	0.754	0.755	Desulfobacterota	0.012	-0.090	0.004453 < Desulfobacterota
39	Adenocarcinoma	0.828	0.098	0.754	0.755	Planctomycetota	0.000	0.067	Planctomycetota <= 0.000168
39	Adenocarcinoma	0.828	0.098	0.754	0.755	Campilobacterota	0.000	-0.030	Campilobacterota <= 0.00011
39	Adenocarcinoma	0.828	0.098	0.754	0.755	Cyanobacteria	0.000	0.029	Cyanobacteria <= 0.000029
39	Adenocarcinoma	0.828	0.098	0.754	0.755	Verrucomicrobiota	0.001	0.027	0.000714 < Verrucomicrobiota <= 0.006252
40	Adenocarcinoma	0.960	0.156	0.720	0.860	Fusobacteriota	0.015	0.090	0.00727 < Fusobacteriota <= 0.14117
40	Adenocarcinoma	0.960	0.156	0.720	0.860	Firmicutes	0.222	-0.074	Firmicutes <= 0.222
40	Adenocarcinoma	0.960	0.156	0.720	0.860	Planctomycetota	0.000	0.062	Planctomycetota <= 0.000168
40	Adenocarcinoma	0.960	0.156	0.720	0.860	Desulfobacterota	0.002	0.034	0.001582 < Desulfobacterota <= 0.004453
40	Adenocarcinoma	0.960	0.156	0.720	0.860	Verrucomicrobiota	0.000	0.029	Verrucomicrobiota <= 0.000714
46	Adenocarcinoma	0.776	0.350	0.775	0.741	Fusobacteriota	0.000	-0.187	Fusobacteriota <= 0.00111
46	Adenocarcinoma	0.776	0.350	0.775	0.741	Planctomycetota	0.000	0.058	Planctomycetota <= 0.000168
46	Adenocarcinoma	0.776	0.350	0.775	0.741	Verrucomicrobiota	0.003	0.036	0.000714 < Verrucomicrobiota <= 0.006252
46	Adenocarcinoma	0.776	0.350	0.775	0.741	Firmicutes	0.292	0.031	0.222 < Firmicutes <= 0.292
46	Adenocarcinoma	0.776	0.350	0.775	0.741	Desulfobacterota	0.003	0.028	0.001582 < Desulfobacterota <= 0.004453
47	Adenocarcinoma	0.890	0.130	0.809	0.807	Fusobacteriota	0.015	0.091	0.00727 < Fusobacteriota <= 0.14117
47	Adenocarcinoma	0.890	0.130	0.809	0.807	Verrucomicrobiota	0.009	-0.071	0.006252 < Verrucomicrobiota
47	Adenocarcinoma	0.890	0.130	0.809	0.807	Campilobacterota	0.000	-0.031	Campilobacterota <= 0.00011
47	Adenocarcinoma	0.890	0.130	0.809	0.807	Firmicutes	0.306	0.031	0.292 < Firmicutes <= 0.376
47	Adenocarcinoma	0.890	0.130	0.809	0.807	Bacteroidota	0.589	-0.024	0.494 < Bacteroidota
48	Adenocarcinoma	0.842	0.335	0.793	0.709	Fusobacteriota	0.001	-0.180	Fusobacteriota <= 0.00111
48	Adenocarcinoma	0.842	0.335	0.793	0.709	Planctomycetota	0.000	0.065	Planctomycetota <= 0.000168
48	Adenocarcinoma	0.842	0.335	0.793	0.709	Campilobacterota	0.000	-0.033	Campilobacterota <= 0.00011
48	Adenocarcinoma	0.842	0.335	0.793	0.709	Firmicutes	0.266	0.037	0.222 < Firmicutes <= 0.292

48	Adenocarcinoma	0.842	0.335	0.793	0.709	Verrucomicrobiota	0.004	0.026	0.000714 < Verrucomicrobiota <= 0.006252
49	Adenocarcinoma	0.960	0.116	0.679	0.929	Fusobacteriota	0.189	0.094	0.14117 < Fusobacteriota
49	Adenocarcinoma	0.960	0.116	0.679	0.929	Planctomycetota	0.000	0.069	Planctomycetota <= 0.000168
49	Adenocarcinoma	0.960	0.116	0.679	0.929	Cyanobacteria	0.000	0.032	Cyanobacteria <= 0.000029
49	Adenocarcinoma	0.960	0.116	0.679	0.929	Verrucomicrobiota	0.000	0.026	Verrucomicrobiota <= 0.000714
49	Adenocarcinoma	0.960	0.116	0.679	0.929	Campilobacterota	0.134	0.029	0.00180 < Campilobacterota
50	Adenocarcinoma	0.928	0.069	0.723	0.805	Verrucomicrobiota	0.012	-0.065	0.006252 < Verrucomicrobiota
50	Adenocarcinoma	0.928	0.069	0.723	0.805	Planctomycetota	0.000	0.071	Planctomycetota <= 0.000168
50	Adenocarcinoma	0.928	0.069	0.723	0.805	Cyanobacteria	0.000	0.029	Cyanobacteria <= 0.000029
50	Adenocarcinoma	0.928	0.069	0.723	0.805	Desulfobacterota	0.003	0.026	0.001582 < Desulfobacterota <= 0.004453
50	Adenocarcinoma	0.928	0.069	0.723	0.805	Firmicutes	0.444	0.021	0.376 < Firmicutes

Table ML-A-5: The LIME Result for Test Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
1	Adenocarcinoma	0.766	0.401	0.813	0.603	Fusobacteriota	0.001	-0.183	Fusobacteriota <= 0.00111
1	Adenocarcinoma	0.766	0.401	0.813	0.603	Desulfobacterota	0.007	-0.090	0.004453 < Desulfobacterota
1	Adenocarcinoma	0.766	0.401	0.813	0.603	Planctomycetota	0.000	0.067	Planctomycetota <= 0.000168
1	Adenocarcinoma	0.766	0.401	0.813	0.603	Campilobacterota	0.000	-0.032	Campilobacterota <= 0.00011
1	Adenocarcinoma	0.766	0.401	0.813	0.603	Verrucomicrobiota	0.000	0.028	Verrucomicrobiota <= 0.000714
4	Adenoma	0.594	0.358	0.210	0.287	Fusobacteriota	0.000	0.180	Fusobacteriota <= 0.00111
4	Adenoma	0.594	0.358	0.210	0.287	Planctomycetota	0.000	-0.069	Planctomycetota <= 0.000168
4	Adenoma	0.594	0.358	0.210	0.287	Campilobacterota	0.000	0.031	Campilobacterota <= 0.00011
4	Adenoma	0.594	0.358	0.210	0.287	Firmicutes	0.247	-0.033	0.222 < Firmicutes <= 0.292
4	Adenoma	0.594	0.358	0.210	0.287	Verrucomicrobiota	0.001	-0.032	0.000714 < Verrucomicrobiota <= 0.006252
5	Adenocarcinoma	0.806	0.344	0.784	0.689	Fusobacteriota	0.001	-0.185	Fusobacteriota <= 0.00111
5	Adenocarcinoma	0.806	0.344	0.784	0.689	Planctomycetota	0.000	0.071	Planctomycetota <= 0.000168
5	Adenocarcinoma	0.806	0.344	0.784	0.689	Campilobacterota	0.000	-0.035	Campilobacterota <= 0.00011
5	Adenocarcinoma	0.806	0.344	0.784	0.689	Verrucomicrobiota	0.001	0.030	0.000714 < Verrucomicrobiota <= 0.006252
5	Adenocarcinoma	0.806	0.344	0.784	0.689	Cyanobacteria	0.000	0.024	Cyanobacteria <= 0.000029
8	Adenocarcinoma	0.626	0.414	0.803	0.590	Fusobacteriota	0.000	-0.185	Fusobacteriota <= 0.00111
8	Adenocarcinoma	0.626	0.414	0.803	0.590	Desulfobacterota	0.006	-0.072	0.004453 < Desulfobacterota
8	Adenocarcinoma	0.626	0.414	0.803	0.590	Verrucomicrobiota	0.076	-0.058	0.006252 < Verrucomicrobiota
8	Adenocarcinoma	0.626	0.414	0.803	0.590	Planctomycetota	0.000	0.075	Planctomycetota <= 0.000168
8	Adenocarcinoma	0.626	0.414	0.803	0.590	Cyanobacteria	0.000	0.027	Cyanobacteria <= 0.000029
9	Adenoma	0.636	0.376	0.193	0.377	Fusobacteriota	0.000	0.187	Fusobacteriota <= 0.00111
9	Adenoma	0.636	0.376	0.193	0.377	Verrucomicrobiota	0.009	0.072	0.006252 < Verrucomicrobiota
9	Adenoma	0.636	0.376	0.193	0.377	Planctomycetota	0.000	-0.072	Planctomycetota <= 0.000168
9	Adenoma	0.636	0.376	0.193	0.377	Campilobacterota	0.000	0.028	Campilobacterota <= 0.00011
9	Adenoma	0.636	0.376	0.193	0.377	Firmicutes	0.373	-0.032	0.292 < Firmicutes <= 0.376
12	Adenocarcinoma	0.804	0.042	0.723	0.850	Planctomycetota	0.000	0.065	Planctomycetota <= 0.000168
12	Adenocarcinoma	0.804	0.042	0.723	0.850	Firmicutes	0.330	0.035	0.292 < Firmicutes <= 0.376

12	Adenocarcinoma	0.804	0.042	0.723	0.850	Verrucomicrobiota	0.001	0.030	0.000714 < Verrucomicrobiota <= 0.006252
12	Adenocarcinoma	0.804	0.042	0.723	0.850	Cyanobacteria	0.002	-0.028	0.000643 < Cyanobacteria
12	Adenocarcinoma	0.804	0.042	0.723	0.850	Desulfobacterota	0.000	0.025	Desulfobacterota <= 0.000808
13	Adenocarcinoma	0.988	0.104	0.694	0.919	Fusobacteriota	0.440	0.098	0.14117 < Fusobacteriota
13	Adenocarcinoma	0.988	0.104	0.694	0.919	Planctomycetota	0.000	0.065	Planctomycetota <= 0.000168
13	Adenocarcinoma	0.988	0.104	0.694	0.919	Verrucomicrobiota	0.001	0.021	Verrucomicrobiota <= 0.000714
13	Adenocarcinoma	0.988	0.104	0.694	0.919	Desulfobacterota	0.000	0.021	Desulfobacterota <= 0.000808
13	Adenocarcinoma	0.988	0.104	0.694	0.919	Campilobacterota	0.005	0.022	0.00180 < Campilobacterota
15	Adenocarcinoma	0.898	0.170	0.816	0.759	Fusobacteriota	0.089	0.089	0.00727 < Fusobacteriota <= 0.14117
15	Adenocarcinoma	0.898	0.170	0.816	0.759	Desulfobacterota	0.006	-0.082	0.004453 < Desulfobacterota
15	Adenocarcinoma	0.898	0.170	0.816	0.759	Verrucomicrobiota	0.082	-0.063	0.006252 < Verrucomicrobiota
15	Adenocarcinoma	0.898	0.170	0.816	0.759	Campilobacterota	0.000	-0.024	Campilobacterota <= 0.00011
15	Adenocarcinoma	0.898	0.170	0.816	0.759	Cyanobacteria	0.000	0.023	Cyanobacteria <= 0.000029
16	Adenocarcinoma	0.964	0.200	0.728	0.791	Fusobacteriota	0.121	0.099	0.00727 < Fusobacteriota <= 0.14117
16	Adenocarcinoma	0.964	0.200	0.728	0.791	Desulfobacterota	0.010	-0.087	0.004453 < Desulfobacterota
16	Adenocarcinoma	0.964	0.200	0.728	0.791	Verrucomicrobiota	0.010	-0.062	0.006252 < Verrucomicrobiota
16	Adenocarcinoma	0.964	0.200	0.728	0.791	Planctomycetota	0.000	0.079	Planctomycetota <= 0.000168
16	Adenocarcinoma	0.964	0.200	0.728	0.791	Firmicutes	0.360	0.035	0.292 < Firmicutes <= 0.376
17	Adenoma	0.570	0.408	0.182	0.389	Fusobacteriota	0.000	0.186	Fusobacteriota <= 0.00111
17	Adenoma	0.570	0.408	0.182	0.389	Desulfobacterota	0.014	0.091	0.004453 < Desulfobacterota
17	Adenoma	0.570	0.408	0.182	0.389	Planctomycetota	0.000	-0.072	Planctomycetota <= 0.000168
17	Adenoma	0.570	0.408	0.182	0.389	Campilobacterota	0.000	0.030	Campilobacterota <= 0.00011
17	Adenoma	0.570	0.408	0.182	0.389	Verrucomicrobiota	0.002	-0.029	0.000714 < Verrucomicrobiota <= 0.006252
18	Adenoma	0.572	0.379	0.186	0.368	Fusobacteriota	0.000	0.188	Fusobacteriota <= 0.00111
18	Adenoma	0.572	0.379	0.186	0.368	Verrucomicrobiota	0.058	0.066	0.006252 < Verrucomicrobiota
18	Adenoma	0.572	0.379	0.186	0.368	Planctomycetota	0.000	-0.070	Planctomycetota <= 0.000168
18	Adenoma	0.572	0.379	0.186	0.368	Campilobacterota	0.000	0.036	Campilobacterota <= 0.00011
18	Adenoma	0.572	0.379	0.186	0.368	Desulfobacterota	0.003	-0.037	0.001582 < Desulfobacterota <= 0.004453
19	Adenocarcinoma	0.608	0.433	0.845	0.550	Fusobacteriota	0.001	-0.176	Fusobacteriota <= 0.00111
19	Adenocarcinoma	0.608	0.433	0.845	0.550	Firmicutes	0.162	-0.079	Firmicutes <= 0.222
19	Adenocarcinoma	0.608	0.433	0.845	0.550	Desulfobacterota	0.006	-0.079	0.004453 < Desulfobacterota
19	Adenocarcinoma	0.608	0.433	0.845	0.550	Planctomycetota	0.000	0.068	Planctomycetota <= 0.000168
19	Adenocarcinoma	0.608	0.433	0.845	0.550	Campilobacterota	0.000	-0.030	Campilobacterota <= 0.00011
27	Adenocarcinoma	0.976	0.114	0.702	0.906	Fusobacteriota	0.095	0.094	0.00727 < Fusobacteriota <= 0.14117
27	Adenocarcinoma	0.976	0.114	0.702	0.906	Planctomycetota	0.000	0.082	Planctomycetota <= 0.000168
27	Adenocarcinoma	0.976	0.114	0.702	0.906	Campilobacterota	0.000	-0.032	Campilobacterota <= 0.00011
27	Adenocarcinoma	0.976	0.114	0.702	0.906	Verrucomicrobiota	0.002	0.031	0.000714 < Verrucomicrobiota <= 0.006252
27	Adenocarcinoma	0.976	0.114	0.702	0.906	Desulfobacterota	0.002	0.029	0.001582 < Desulfobacterota <= 0.004453
28	Adenocarcinoma	0.982	0.163	0.732	0.823	Fusobacteriota	0.097	0.100	0.00727 < Fusobacteriota <= 0.14117
28	Adenocarcinoma	0.982	0.163	0.732	0.823	Desulfobacterota	0.005	-0.075	0.004453 < Desulfobacterota
28	Adenocarcinoma	0.982	0.163	0.732	0.823	Planctomycetota	0.000	0.070	Planctomycetota <= 0.000168
28	Adenocarcinoma	0.982	0.163	0.732	0.823	Campilobacterota	0.000	-0.031	Campilobacterota <= 0.00011
28	Adenocarcinoma	0.982	0.163	0.732	0.823	Cyanobacteria	0.000	0.028	Cyanobacteria <= 0.000029

30	Adenocarcinoma	0.898	0.135	0.766	0.881	Fusobacteriota	0.125	0.096	0.00727 < Fusobacteriota <= 0.14117
30	Adenocarcinoma	0.898	0.135	0.766	0.881	Verrucomicrobiota	0.022	-0.067	0.006252 < Verrucomicrobiota
30	Adenocarcinoma	0.898	0.135	0.766	0.881	Firmicutes	0.234	0.036	0.222 < Firmicutes <= 0.292
30	Adenocarcinoma	0.898	0.135	0.766	0.881	Desulfobacterota	0.004	0.028	0.001582 < Desulfobacterota <= 0.004453
30	Adenocarcinoma	0.898	0.135	0.766	0.881	Cyanobacteria	0.000	0.022	Cyanobacteria <= 0.000029
31	Adenocarcinoma	0.836	0.081	0.740	0.758	Desulfobacterota	0.010	-0.082	0.004453 < Desulfobacterota
31	Adenocarcinoma	0.836	0.081	0.740	0.758	Planctomycetota	0.000	0.071	Planctomycetota <= 0.000168
31	Adenocarcinoma	0.836	0.081	0.740	0.758	Verrucomicrobiota	0.000	0.021	Verrucomicrobiota <= 0.000714
31	Adenocarcinoma	0.836	0.081	0.740	0.758	Fusobacteriota	0.004	0.022	0.00111 < Fusobacteriota <= 0.00727
31	Adenocarcinoma	0.836	0.081	0.740	0.758	Cyanobacteria	0.000	-0.014	0.000029 < Cyanobacteria <= 0.000643
35	Adenocarcinoma	0.764	0.113	0.770	0.714	Desulfobacterota	0.009	-0.095	0.004453 < Desulfobacterota
35	Adenocarcinoma	0.764	0.113	0.770	0.714	Planctomycetota	0.000	0.073	Planctomycetota <= 0.000168
35	Adenocarcinoma	0.764	0.113	0.770	0.714	Campilobacterota	0.000	-0.035	Campilobacterota <= 0.00011
35	Adenocarcinoma	0.764	0.113	0.770	0.714	Firmicutes	0.346	0.028	0.292 < Firmicutes <= 0.376
35	Adenocarcinoma	0.764	0.113	0.770	0.714	Cyanobacteria	0.003	-0.026	0.000643 < Cyanobacteria
36	Adenocarcinoma	0.878	0.155	0.713	0.865	Fusobacteriota	0.079	0.098	0.00727 < Fusobacteriota <= 0.14117
36	Adenocarcinoma	0.878	0.155	0.713	0.865	Firmicutes	0.168	-0.070	Firmicutes <= 0.222
36	Adenocarcinoma	0.878	0.155	0.713	0.865	Planctomycetota	0.000	0.067	Planctomycetota <= 0.000168
36	Adenocarcinoma	0.878	0.155	0.713	0.865	Cyanobacteria	0.000	0.027	Cyanobacteria <= 0.000029
36	Adenocarcinoma	0.878	0.155	0.713	0.865	Desulfobacterota	0.004	0.029	0.001582 < Desulfobacterota <= 0.004453
37	Adenocarcinoma	0.832	0.148	0.714	0.859	Fusobacteriota	0.190	0.093	0.14117 < Fusobacteriota
37	Adenocarcinoma	0.832	0.148	0.714	0.859	Firmicutes	0.086	-0.076	Firmicutes <= 0.222
37	Adenocarcinoma	0.832	0.148	0.714	0.859	Planctomycetota	0.000	0.066	Planctomycetota <= 0.000168
37	Adenocarcinoma	0.832	0.148	0.714	0.859	Verrucomicrobiota	0.000	0.032	Verrucomicrobiota <= 0.000714
37	Adenocarcinoma	0.832	0.148	0.714	0.859	Desulfobacterota	0.003	0.030	0.001582 < Desulfobacterota <= 0.004453
38	Adenocarcinoma	0.952	0.199	0.726	0.780	Fusobacteriota	0.186	0.099	0.14117 < Fusobacteriota
38	Adenocarcinoma	0.952	0.199	0.726	0.780	Desulfobacterota	0.005	-0.081	0.004453 < Desulfobacterota
38	Adenocarcinoma	0.952	0.199	0.726	0.780	Verrucomicrobiota	0.026	-0.078	0.006252 < Verrucomicrobiota
38	Adenocarcinoma	0.952	0.199	0.726	0.780	Planctomycetota	0.000	0.084	Planctomycetota <= 0.000168
38	Adenocarcinoma	0.952	0.199	0.726	0.780	Firmicutes	0.335	0.030	0.292 < Firmicutes <= 0.376
41	Adenocarcinoma	0.832	0.142	0.715	0.877	Fusobacteriota	0.193	0.099	0.14117 < Fusobacteriota
41	Adenocarcinoma	0.832	0.142	0.715	0.877	Verrucomicrobiota	0.253	-0.062	0.006252 < Verrucomicrobiota
41	Adenocarcinoma	0.832	0.142	0.715	0.877	Planctomycetota	0.000	0.065	Planctomycetota <= 0.000168
41	Adenocarcinoma	0.832	0.142	0.715	0.877	Desulfobacterota	0.004	0.032	0.001582 < Desulfobacterota <= 0.004453
41	Adenocarcinoma	0.832	0.142	0.715	0.877	Firmicutes	0.296	0.027	0.292 < Firmicutes <= 0.376
42	Adenocarcinoma	0.808	0.053	0.723	0.849	Planctomycetota	0.000	0.071	Planctomycetota <= 0.000168
42	Adenocarcinoma	0.808	0.053	0.723	0.849	Campilobacterota	0.000	-0.032	Campilobacterota <= 0.00011
42	Adenocarcinoma	0.808	0.053	0.723	0.849	Firmicutes	0.339	0.033	0.292 < Firmicutes <= 0.376
42	Adenocarcinoma	0.808	0.053	0.723	0.849	Desulfobacterota	0.002	0.030	0.001582 < Desulfobacterota <= 0.004453
42	Adenocarcinoma	0.808	0.053	0.723	0.849	Verrucomicrobiota	0.000	0.023	Verrucomicrobiota <= 0.000714
43	Adenoma	0.590	0.345	0.206	0.293	Fusobacteriota	0.000	0.185	Fusobacteriota <= 0.00111
43	Adenoma	0.590	0.345	0.206	0.293	Planctomycetota	0.000	-0.063	Planctomycetota <= 0.000168
43	Adenoma	0.590	0.345	0.206	0.293	Desulfobacterota	0.003	-0.030	0.001582 < Desulfobacterota <= 0.004453

43	Adenoma	0.590	0.345	0.206	0.293	Verrucomicrobiota	0.003	-0.029	0.000714 < Verrucomicrobiota <= 0.006252
43	Adenoma	0.590	0.345	0.206	0.293	Campilobacterota	0.000	0.024	Campilobacterota <= 0.00011
44	Adenoma	0.542	0.370	0.206	0.345	Fusobacteriota	0.000	0.181	Fusobacteriota <= 0.00111
44	Adenoma	0.542	0.370	0.206	0.345	Verrucomicrobiota	0.075	0.071	0.006252 < Verrucomicrobiota
44	Adenoma	0.542	0.370	0.206	0.345	Planctomycetota	0.000	-0.061	Planctomycetota <= 0.000168
44	Adenoma	0.542	0.370	0.206	0.345	Desulfobacterota	0.004	-0.032	0.001582 < Desulfobacterota <= 0.004453
44	Adenoma	0.542	0.370	0.206	0.345	Cyanobacteria	0.000	-0.020	Cyanobacteria <= 0.000029
45	Adenoma	0.534	0.065	0.254	0.237	Verrucomicrobiota	0.053	0.067	0.006252 < Verrucomicrobiota
45	Adenoma	0.534	0.065	0.254	0.237	Planctomycetota	0.000	-0.067	Planctomycetota <= 0.000168
45	Adenoma	0.534	0.065	0.254	0.237	Campilobacterota	0.000	0.026	Campilobacterota <= 0.00011
45	Adenoma	0.534	0.065	0.254	0.237	Cyanobacteria	0.000	-0.025	Cyanobacteria <= 0.000029
45	Adenoma	0.534	0.065	0.254	0.237	Euryarchaeota	0.006	-0.017	0.00051 < Euryarchaeota

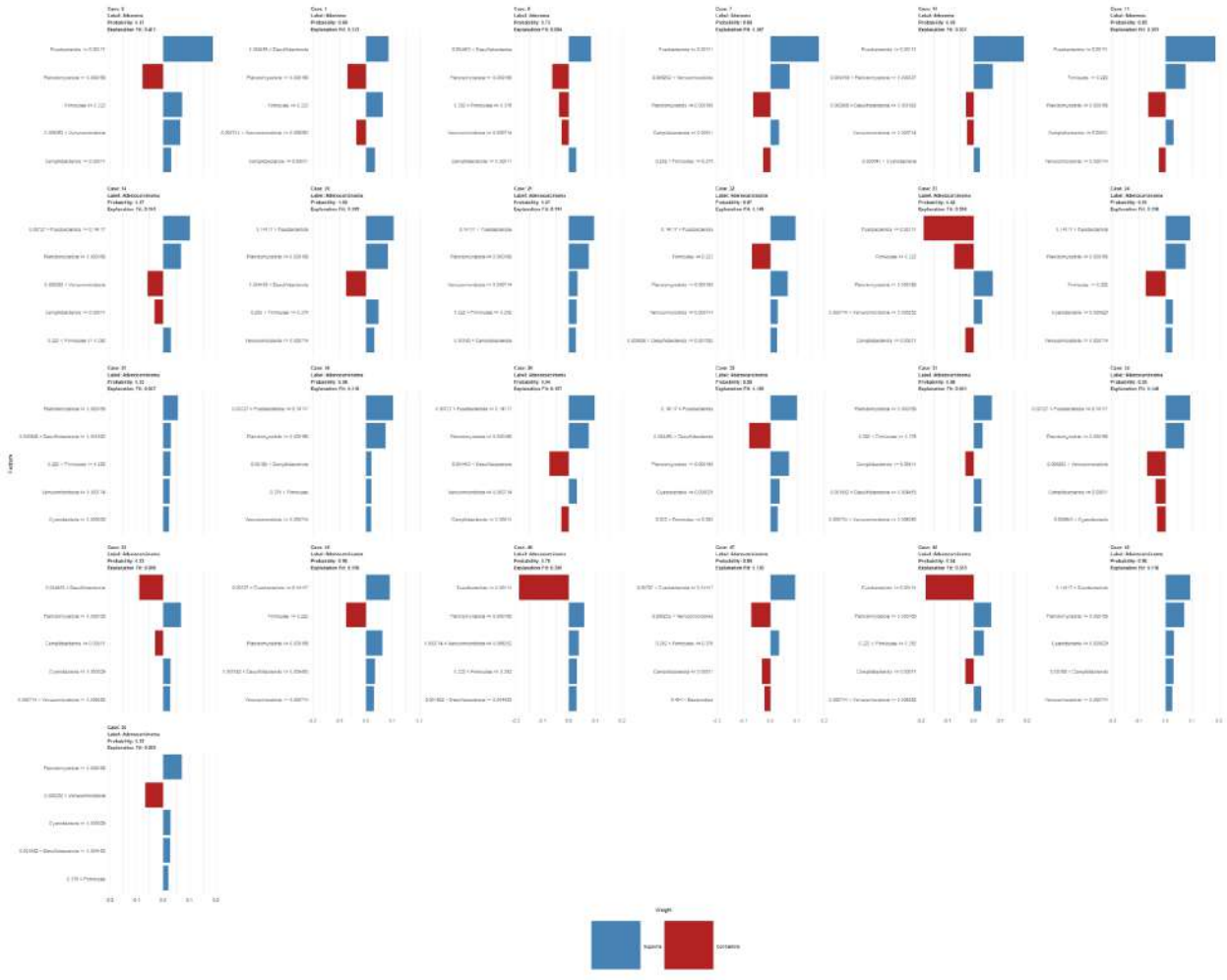


Figure ML1-A-6: The LIME Result in **Training dataset** for **Phylum** in **Biopsy**



Figure ML1-A-7: The LIME Result in **Test dataset** for **Phylum** in **Biopsy**

Table ML1-A-6: The Aggregated SHAP Values								
Variable	Training Dataset				Testing Dataset			
	contribution	position	sign	cumulative	contribution	position	sign	cumulative
intercept	0.7515	21	X	0.752	0.7515	21	X	0.752
Myxococcota	0.0000	20	-1	0.752	0.0000	20	-1	0.751
Firmicutes	0.0000	19	1	0.752	0.0056	19	1	0.757
Cyanobacteria	-0.0001	18	-1	0.751	0.0034	18	1	0.761
Euryarchaeota	-0.0003	17	-1	0.751	0.0000	17	1	0.761
Gender	0.0000	16	1	0.751	0.0021	16	1	0.763
Verrucomicrobiota	0.0000	15	1	0.751	-0.0445	15	-1	0.718
Fusobacteriota	0.0004	14	1	0.752	0.0044	14	1	0.723
Proteobacteria	0.0000	13	-1	0.752	0.0028	13	1	0.725
Spirochaetota	0.0000	12	1	0.752	0.0000	12	1	0.725
Tumor	0.0000	11	1	0.752	-0.0036	11	-1	0.722
Campilobacterota	-0.0006	10	-1	0.751	0.0000	10	-1	0.722
Bacteroidota	0.0000	9	1	0.751	0.0007	9	1	0.722
Chloroflexi	0.0000	8	1	0.751	-0.0001	8	-1	0.722
Planctomycetota	-0.0002	7	-1	0.751	0.0008	7	1	0.723
Desulfobacterota	0.0004	6	1	0.751	0.0054	6	1	0.728
Patescibacteria	0.0000	5	1	0.751	0.0003	5	1	0.729
Synergistota	-0.0001	4	-1	0.751	-0.0007	4	-1	0.728
Acidobacteriota	0.0000	3	1	0.751	-0.0001	3	-1	0.728
Actinobacteriota	0.0003	2	1	0.752	-0.0008	2	-1	0.727
prediction	0.7515	1	X	0.752	0.7272	1	X	0.727

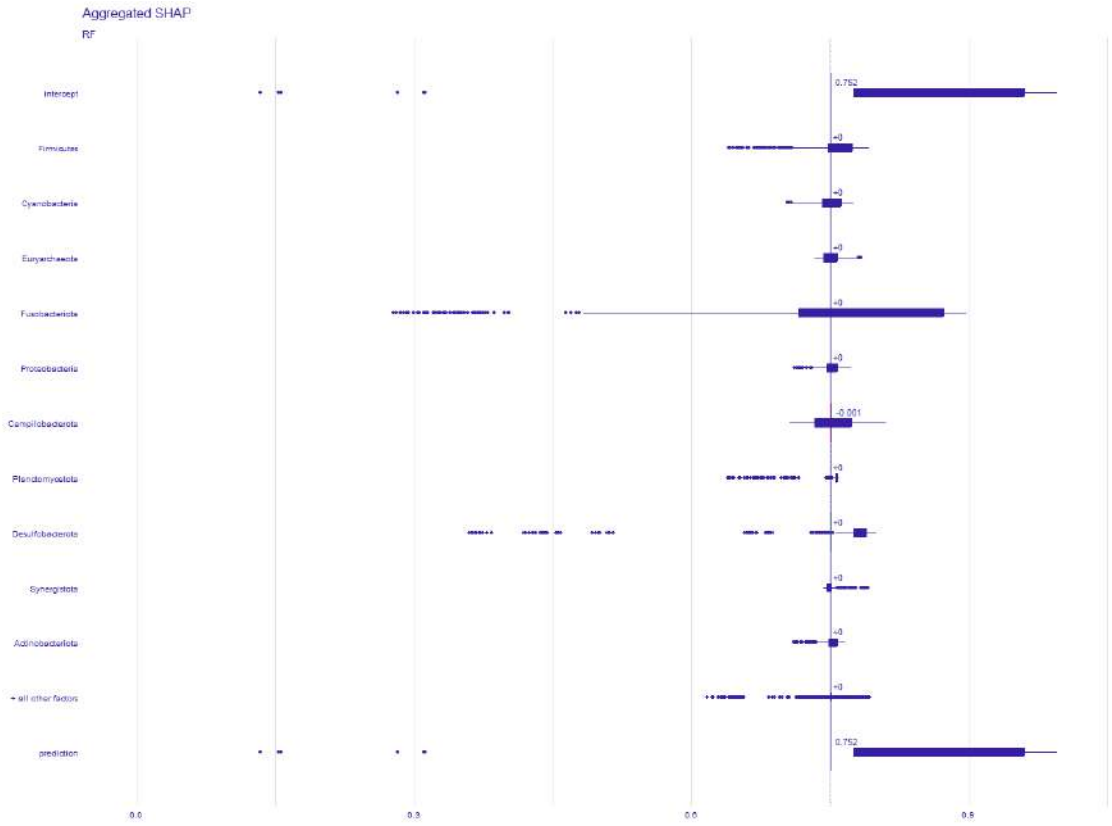


Figure ML1-A-8: The Aggregated SHAP Value in **Training dataset** for **Phylum** in **Biopsy**

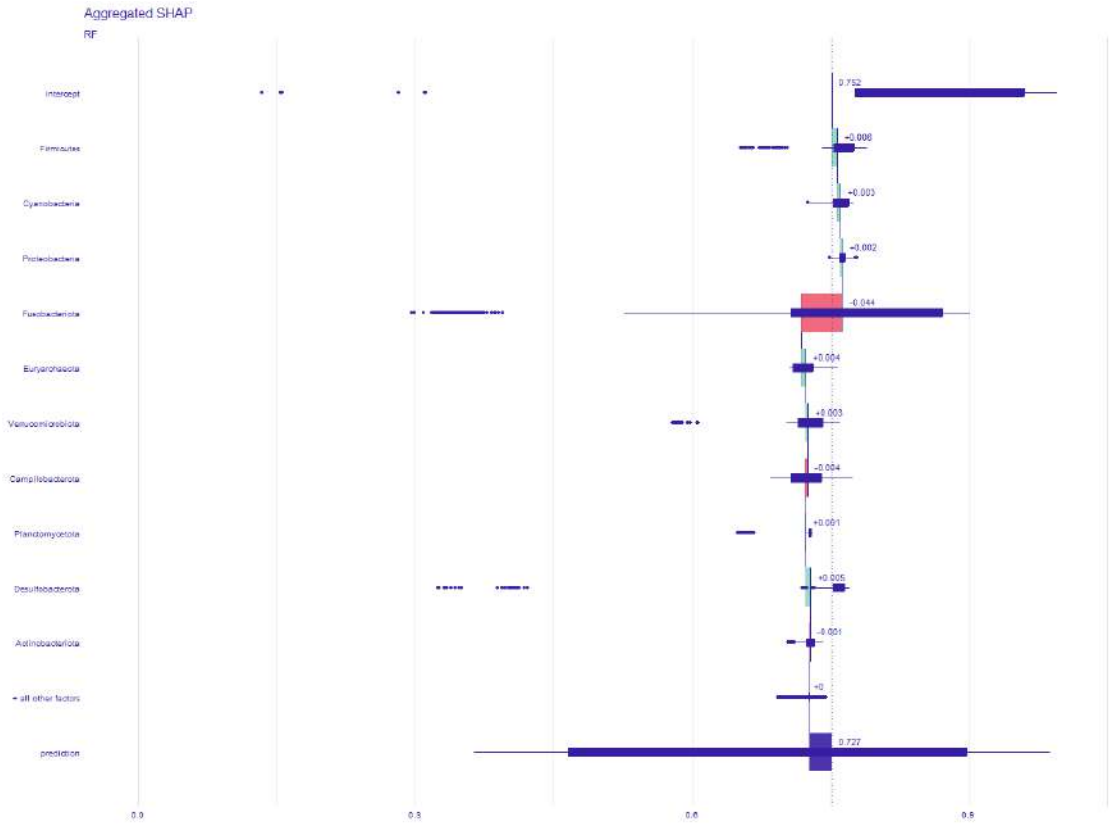


Figure ML1-A-9: The Aggregated SHAP Value in **Test dataset** for **Phylum** in **Biopsy**

B- Class

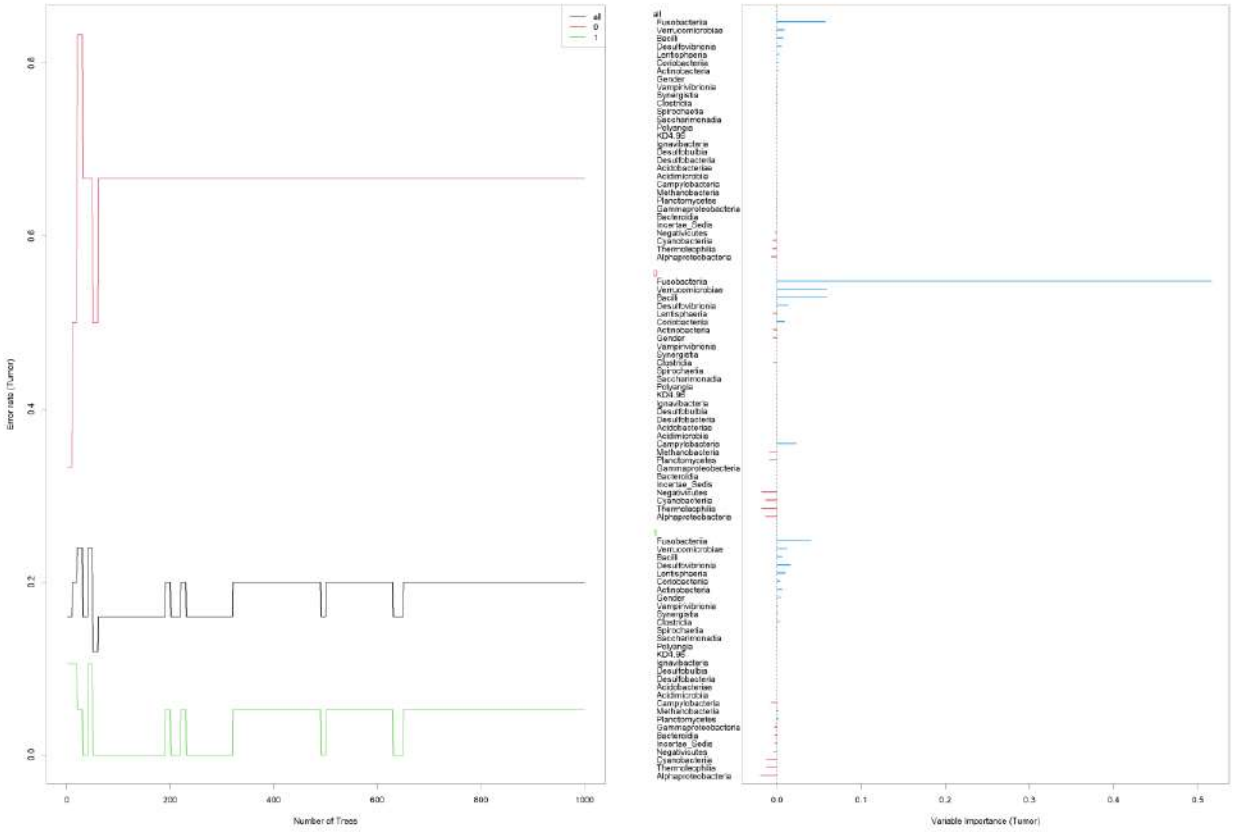
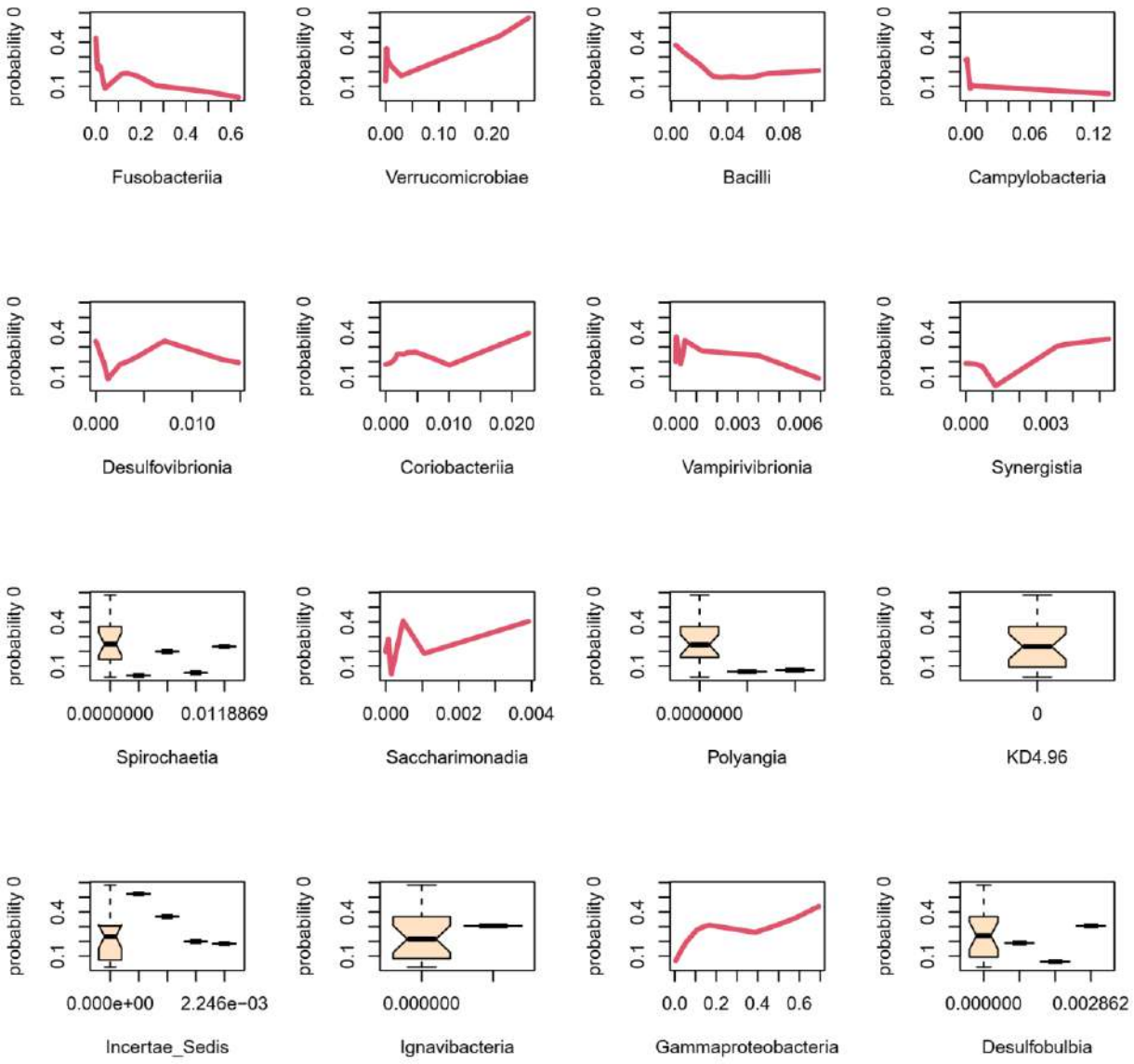


Figure ML1-B-1: The number of trees and Variable Importance Plot for **Class** in **Biopsy**


probability 0
0.1
0.4
0.0 0.2 0.4 0.6
Fusobacteriia
0.00 0.10 0.20
Verrucomicrobiae
0.00 0.04 0.08
Bacilli
0.00 0.06 0.12
Campylobacteria
0.000 0.010
Desulfovibrionia
0.000 0.010 0.020
Coriobacteriia
0.000 0.003 0.006
Vampirivibrionia
0.000 0.003
Synergistia
0.0000000 0.0118869
Spirochaetia
0.000 0.002 0.004
Saccharimonadia
0.0000000
Polyangia
0
KD4.96
0.000e+00 2.246e-03
Incertae_Sedis
0.000000
Ignavibacteria
0.0 0.2 0.4 0.6
Gammaproteobacteria
0.000000 0.002862
Desulfobulbia

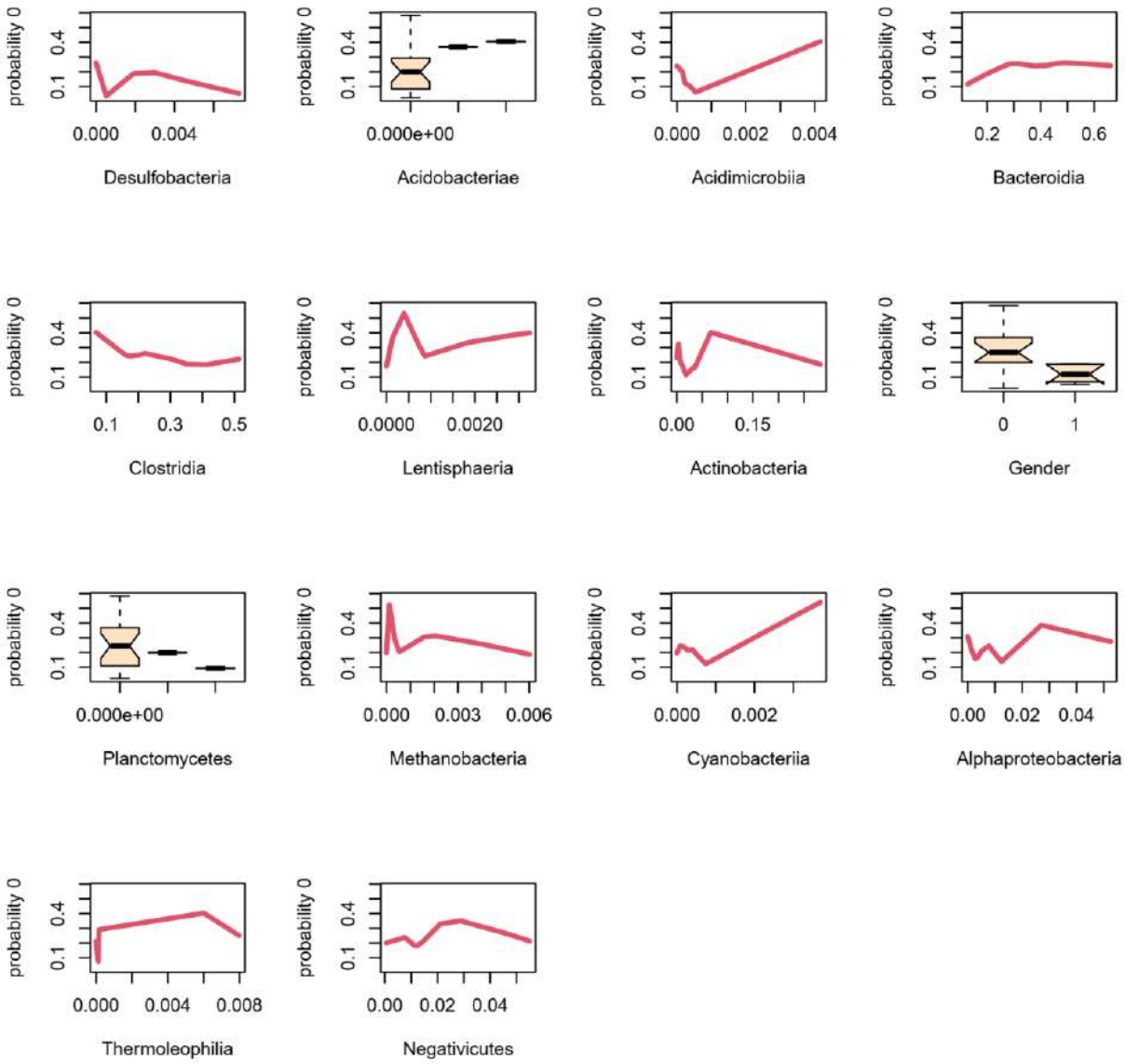
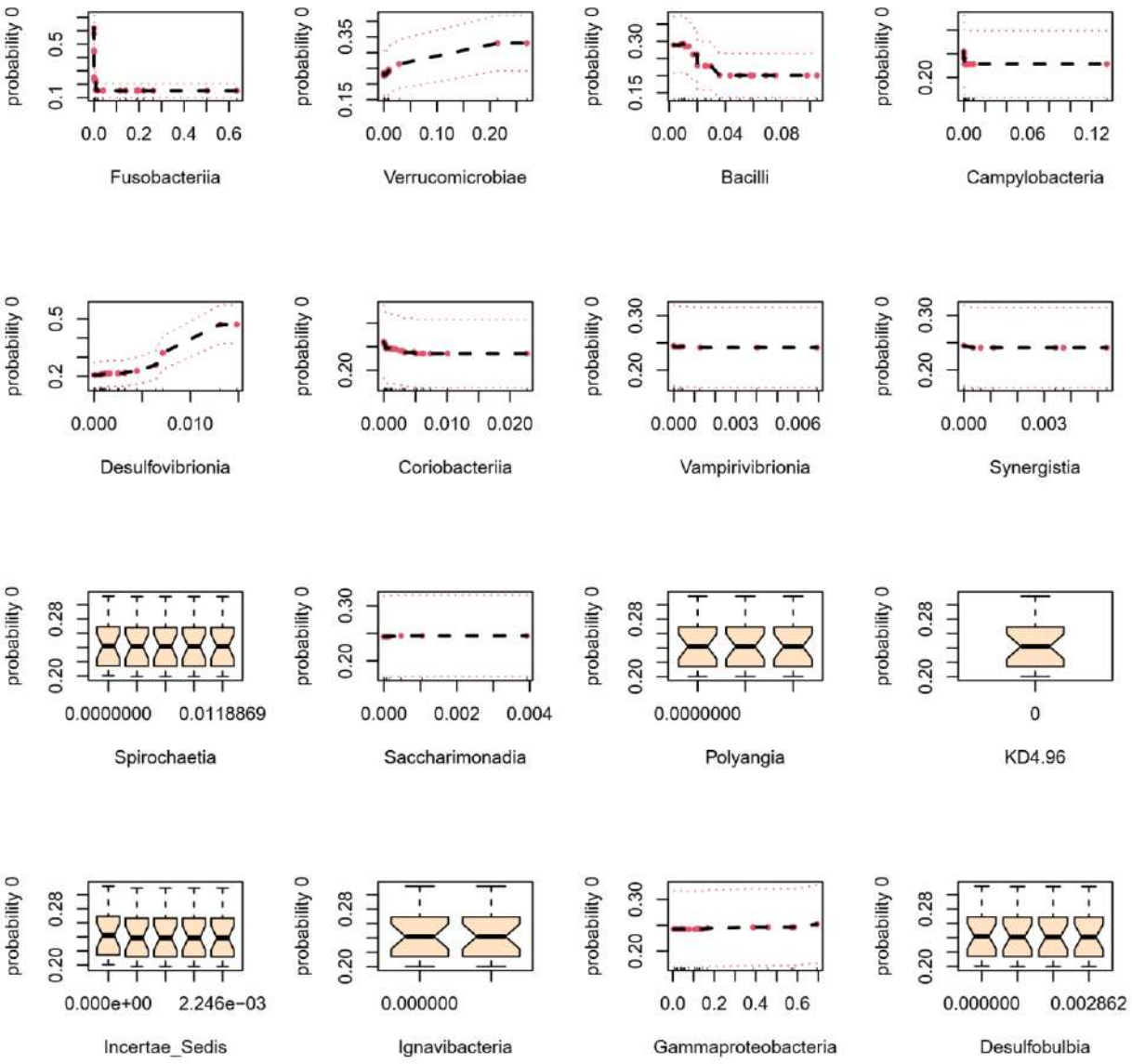


Figure ML1-B-2: Marginal Plot for **Class** in **Biopsy**


probability 0
Fusobacteriia
Verrucomicrobiae
Bacilli
Campylobacteria
Desulfovibrionia
Coriobacteriia
Vampirivibrionia
Synergistia
Spirochaetia
Saccharimonadia
Polyangia
KD4.96
Incertae_Sedis
Ignavibacteria
Gammaproteobacteria
Desulfobulbia

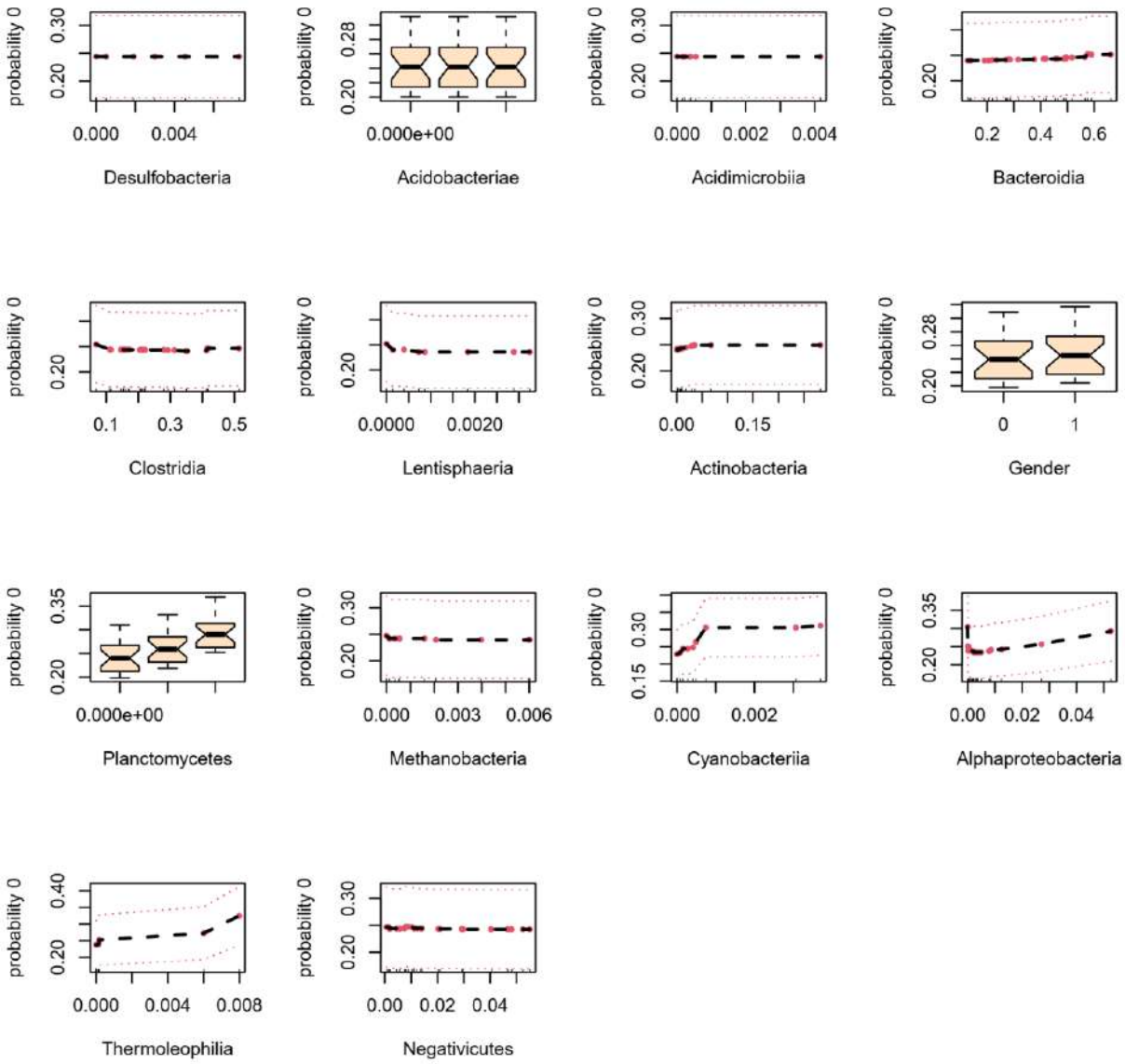


Figure ML1-B-3: Partial Plot for **Class** in **Biopsy**

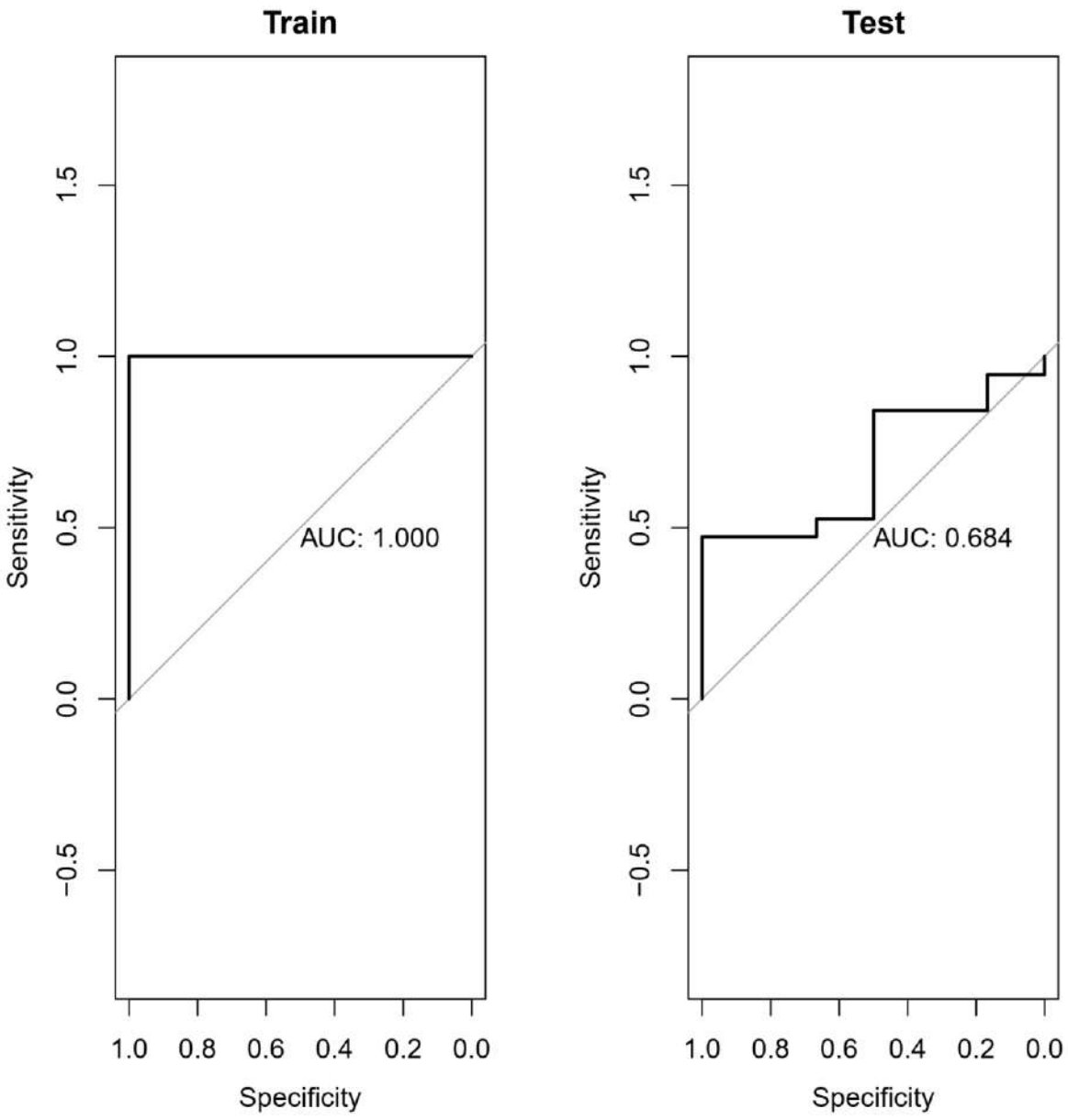


Figure ML1-B-4: The ROC Curve for **Class** in **Biopsy**

Table ML1-B-1: The Model Performance

The Model Performance			
	Values		
Index	2.50%	50.00%	97.50%
Accuracy	0.48	0.68	0.92
Specificity	0.50	1.00	1.00
Sensitivity	0.32	0.58	0.95

Table ML1-B-2: The Variable Importance Values

The Variable Importance Values						
Covariates	Variable Importance			Brier Score		
	all	AP	CRC	all	AP	CRC
Gender	0.0009	-0.0136	0.0072	-0.0082	-0.0056	-0.0056
Acidimicrobiia	-0.0013	0.0000	-0.0043	-0.0011	-0.0007	-0.0007
Acidobacteriae	0.0000	0.0000	0.0000	-0.0017	-0.0011	-0.0011
Actinobacteria	0.0008	-0.0045	0.0043	-0.0098	-0.0067	-0.0067
Alphaproteobacteria	0.0237	-0.3579	0.1974	0.0711	0.0483	0.0483
Bacilli	0.0009	0.0136	0.0000	-0.0065	-0.0044	-0.0044
Bacteroidia	-0.0053	-0.0317	-0.0086	-0.0201	-0.0137	-0.0137
Campylobacteria	-0.0016	0.0136	-0.0100	-0.0222	-0.0151	-0.0151
Clostridia	-0.0036	-0.0544	0.0043	-0.0069	-0.0047	-0.0047
Coriobacteriia	0.0009	-0.0317	0.0129	-0.0068	-0.0046	-0.0046
Cyanobacteriia	0.0075	-0.1767	0.0830	0.0091	0.0062	0.0062
Desulfobacteria	0.0000	0.0000	0.0000	0.0003	0.0002	0.0002
Desulfobulbia	-0.0004	0.0000	-0.0014	-0.0024	-0.0016	-0.0016
Desulfovibrionia	0.1440	-0.4893	0.6610	0.4513	0.3067	0.3067
Fusobacteriia	0.3535	-0.0317	1.2590	0.9053	0.6152	0.6152
Gammaproteobacteria	0.0004	-0.0181	0.0072	-0.0169	-0.0115	-0.0115
Ignavibacteria	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Incertae_Sedis	-0.0016	0.0045	-0.0072	-0.0060	-0.0041	-0.0041
KD4.96	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Lentisphaeria	0.0033	0.0136	0.0072	0.0138	0.0094	0.0094
Methanobacteria	-0.0024	0.0000	-0.0086	-0.0063	-0.0043	-0.0043
Negativicutes	-0.0049	-0.0227	-0.0100	-0.0071	-0.0048	-0.0048
Planctomycetes	0.0049	-0.2356	0.0916	0.0168	0.0114	0.0114

Polyangia	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Saccharimonadia	0.0000	-0.0045	0.0014	-0.0041	-0.0028	-0.0028
Spirochaetia	0.0000	0.0000	0.0000	-0.0009	-0.0006	-0.0006
Synergistia	-0.0022	0.0136	-0.0114	-0.0059	-0.0040	-0.0040
Thermoleophilia	0.0164	-0.2311	0.1302	0.0516	0.0350	0.0350
Vampirivibrionia	-0.0004	0.0091	-0.0043	-0.0045	-0.0030	-0.0030
Verrucomicrobiae	0.0344	-0.1540	0.1703	0.0700	0.0476	0.0476

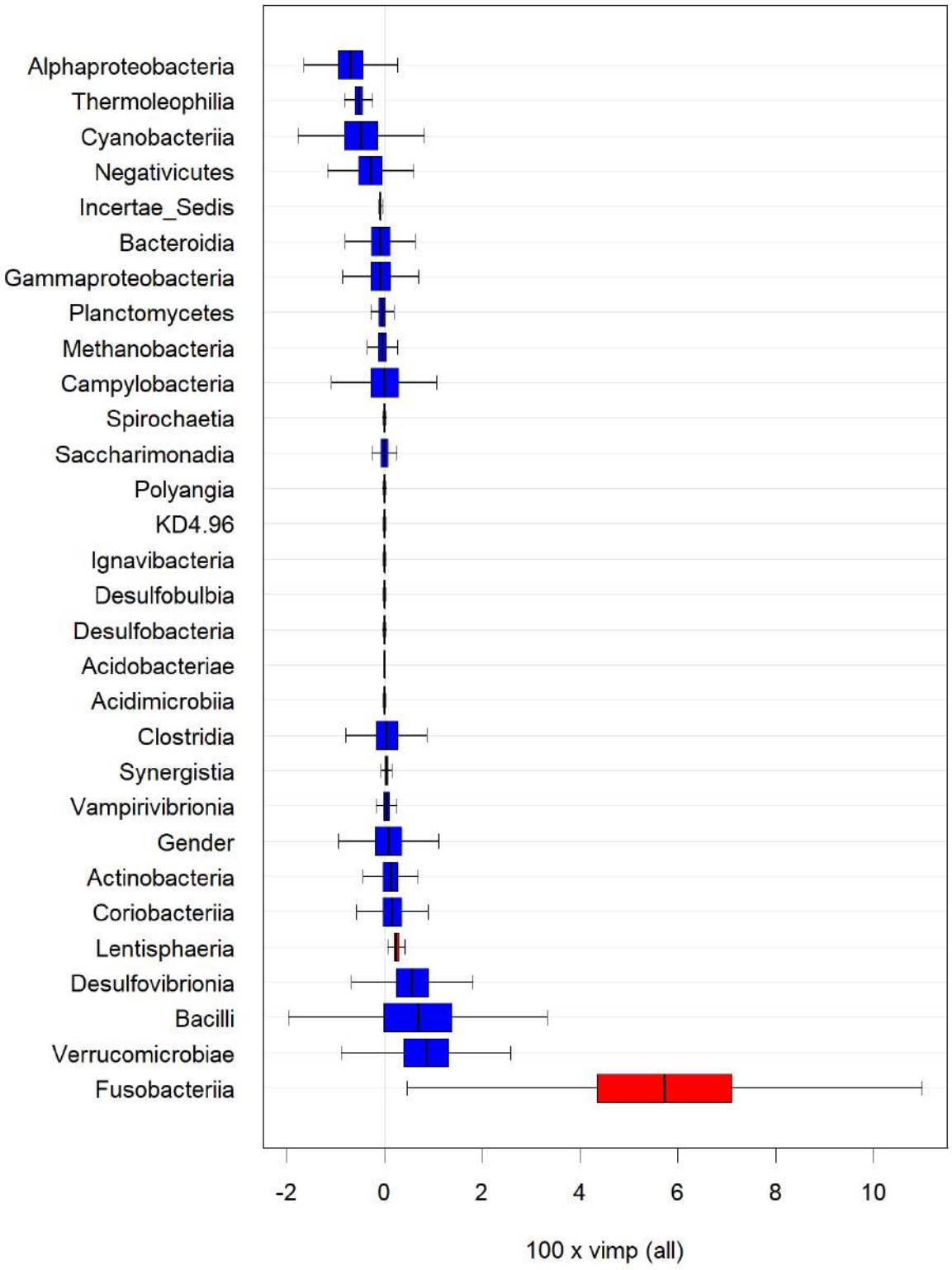


Figure ML1-B-5: The Variable Importance Plot (VIMP) for **Class** in **Biopsy**

Best model is selected for LIME results

Table ML1-B-3: The best setting for model and LIME Results

Setting	mtry	ROC	Sens	Spec	ROCSD	SensSD	SpecSD
ntrees: 500 nodesize: 1	9	0.700	0.300	0.950	0.411	0.447	0.112
ntrees: 1000 nodesize: 1	10	0.642	0.300	0.950	0.356	0.447	0.112
ntrees: 500 nodesize: 5	3	0.667	0.000	1.000	0.358	0.000	0.000
ntrees: 1000 nodesize: 5	8	0.600	0.300	0.950	0.454	0.447	0.112
ntrees: 1000 nodesize: 5	10	0.600	0.500	0.950	0.454	0.500	0.112

Table ML1-B-4: The LIME Result for Training Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
2	Adenoma	0.844	0.456	0.182	0.442	Fusobacteriia	0.000	0.132	Fusobacteriia <= 0.00111
2	Adenoma	0.844	0.456	0.182	0.442	Bacilli	0.010	0.085	Bacilli <= 0.0114
2	Adenoma	0.844	0.456	0.182	0.442	Verrucomicrobiae	0.270	0.080	0.006252 < Verrucomicrobiae
2	Adenoma	0.844	0.456	0.182	0.442	Cyanobacteriia	0.000	-0.071	Cyanobacteriia <= 0.000151
2	Adenoma	0.844	0.456	0.182	0.442	Campylobacteria	0.000	0.035	Campylobacteria <= 0.00011
3	Adenoma	0.728	0.196	0.260	0.239	Desulfovibrionia	0.013	0.088	0.003165 < Desulfovibrionia
3	Adenoma	0.728	0.196	0.260	0.239	Cyanobacteriia	0.000	-0.073	Cyanobacteriia <= 0.000151
3	Adenoma	0.728	0.196	0.260	0.239	Bacilli	0.030	-0.043	0.0265 < Bacilli <= 0.0576
3	Adenoma	0.728	0.196	0.260	0.239	Campylobacteria	0.000	0.030	Campylobacteria <= 0.00011
3	Adenoma	0.728	0.196	0.260	0.239	Verrucomicrobiae	0.006	-0.023	0.000551 < Verrucomicrobiae <= 0.006252
6	Adenoma	0.722	0.187	0.250	0.292	Desulfovibrionia	0.015	0.087	0.003165 < Desulfovibrionia
6	Adenoma	0.722	0.187	0.250	0.292	Cyanobacteriia	0.000	-0.069	Cyanobacteriia <= 0.000151
6	Adenoma	0.722	0.187	0.250	0.292	Campylobacteria	0.000	0.032	Campylobacteria <= 0.00011
6	Adenoma	0.722	0.187	0.250	0.292	Verrucomicrobiae	0.000	-0.031	Verrucomicrobiae <= 0.000551
6	Adenoma	0.722	0.187	0.250	0.292	Negativicutes	0.010	0.024	0.00748 < Negativicutes <= 0.01101
7	Adenoma	0.816	0.441	0.184	0.435	Fusobacteriia	0.000	0.127	Fusobacteriia <= 0.00111
7	Adenoma	0.816	0.441	0.184	0.435	Bacilli	0.011	0.082	Bacilli <= 0.0114
7	Adenoma	0.816	0.441	0.184	0.435	Verrucomicrobiae	0.214	0.079	0.006252 < Verrucomicrobiae
7	Adenoma	0.816	0.441	0.184	0.435	Cyanobacteriia	0.000	-0.072	Cyanobacteriia <= 0.000151
7	Adenoma	0.816	0.441	0.184	0.435	Campylobacteria	0.000	0.035	Campylobacteria <= 0.00011
10	Adenoma	0.688	0.310	0.175	0.368	Fusobacteriia	0.000	0.126	Fusobacteriia <= 0.00111
10	Adenoma	0.688	0.310	0.175	0.368	Cyanobacteriia	0.001	0.073	0.000151 < Cyanobacteriia
10	Adenoma	0.688	0.310	0.175	0.368	Verrucomicrobiae	0.000	-0.041	Verrucomicrobiae <= 0.000551
10	Adenoma	0.688	0.310	0.175	0.368	Bacilli	0.020	0.019	0.0114 < Bacilli <= 0.0265
10	Adenoma	0.688	0.310	0.175	0.368	Lentisphaeria	0.000	0.016	Lentisphaeria <= 0.000135
11	Adenoma	0.788	0.401	0.147	0.435	Fusobacteriia	0.000	0.127	Fusobacteriia <= 0.00111
11	Adenoma	0.788	0.401	0.147	0.435	Bacilli	0.009	0.087	Bacilli <= 0.0114
11	Adenoma	0.788	0.401	0.147	0.435	Cyanobacteriia	0.004	0.072	0.000151 < Cyanobacteriia
11	Adenoma	0.788	0.401	0.147	0.435	Campylobacteria	0.000	0.037	Campylobacteria <= 0.00011
11	Adenoma	0.788	0.401	0.147	0.435	Verrucomicrobiae	0.000	-0.035	Verrucomicrobiae <= 0.000551

14	Adenocarcinoma	0.968	0.310	0.729	0.766	Desulfovibrionia	0.004	-0.084	0.003165 < Desulfovibrionia
14	Adenocarcinoma	0.968	0.310	0.729	0.766	Verrucomicrobiae	0.029	-0.080	0.006252 < Verrucomicrobiae
14	Adenocarcinoma	0.968	0.310	0.729	0.766	Fusobacteriia	0.141	0.073	0.00727 < Fusobacteriia <= 0.14117
14	Adenocarcinoma	0.968	0.310	0.729	0.766	Cyanobacteriia	0.000	0.069	Cyanobacteriia <= 0.000151
14	Adenocarcinoma	0.968	0.310	0.729	0.766	Bacilli	0.069	0.059	0.0576 < Bacilli
20	Adenocarcinoma	0.972	0.230	0.723	0.789	Desulfovibrionia	0.007	-0.085	0.003165 < Desulfovibrionia
20	Adenocarcinoma	0.972	0.230	0.723	0.789	Fusobacteriia	0.197	0.072	0.14117 < Fusobacteriia
20	Adenocarcinoma	0.972	0.230	0.723	0.789	Cyanobacteriia	0.000	0.072	Cyanobacteriia <= 0.000151
20	Adenocarcinoma	0.972	0.230	0.723	0.789	Verrucomicrobiae	0.001	0.026	0.000551 < Verrucomicrobiae <= 0.006252
20	Adenocarcinoma	0.972	0.230	0.723	0.789	Bacilli	0.027	-0.018	0.0114 < Bacilli <= 0.0265
21	Adenocarcinoma	0.948	0.148	0.738	0.837	Fusobacteriia	0.220	0.075	0.14117 < Fusobacteriia
21	Adenocarcinoma	0.948	0.148	0.738	0.837	Cyanobacteriia	0.000	-0.066	0.000151 < Cyanobacteriia
21	Adenocarcinoma	0.948	0.148	0.738	0.837	Verrucomicrobiae	0.001	0.032	Verrucomicrobiae <= 0.000551
21	Adenocarcinoma	0.948	0.148	0.738	0.837	Campylobacteria	0.008	0.032	0.00180 < Campylobacteria
21	Adenocarcinoma	0.948	0.148	0.738	0.837	Desulfovibrionia	0.001	0.026	0.000627 < Desulfovibrionia <= 0.001087
22	Adenocarcinoma	0.970	0.186	0.680	0.934	Fusobacteriia	0.635	0.076	0.14117 < Fusobacteriia
22	Adenocarcinoma	0.970	0.186	0.680	0.934	Cyanobacteriia	0.000	0.069	Cyanobacteriia <= 0.000151
22	Adenocarcinoma	0.970	0.186	0.680	0.934	Bacilli	0.060	0.056	0.0576 < Bacilli
22	Adenocarcinoma	0.970	0.186	0.680	0.934	Verrucomicrobiae	0.001	0.026	0.000551 < Verrucomicrobiae <= 0.006252
22	Adenocarcinoma	0.970	0.186	0.680	0.934	Campylobacteria	0.005	0.026	0.00180 < Campylobacteria
23	Adenocarcinoma	0.818	0.385	0.786	0.649	Fusobacteriia	0.001	-0.128	Fusobacteriia <= 0.00111
23	Adenocarcinoma	0.818	0.385	0.786	0.649	Bacilli	0.007	-0.084	Bacilli <= 0.0114
23	Adenocarcinoma	0.818	0.385	0.786	0.649	Cyanobacteriia	0.000	0.070	Cyanobacteriia <= 0.000151
23	Adenocarcinoma	0.818	0.385	0.786	0.649	Desulfovibrionia	0.001	0.039	Desulfovibrionia <= 0.000627
23	Adenocarcinoma	0.818	0.385	0.786	0.649	Campylobacteria	0.000	-0.035	Campylobacteria <= 0.00011
24	Adenocarcinoma	0.966	0.184	0.675	0.885	Cyanobacteriia	0.000	0.077	Cyanobacteriia <= 0.000151
24	Adenocarcinoma	0.966	0.184	0.675	0.885	Fusobacteriia	0.503	0.076	0.14117 < Fusobacteriia
24	Adenocarcinoma	0.966	0.184	0.675	0.885	Bacilli	0.043	0.050	0.0265 < Bacilli <= 0.0576
24	Adenocarcinoma	0.966	0.184	0.675	0.885	Verrucomicrobiae	0.000	0.030	Verrucomicrobiae <= 0.000551
24	Adenocarcinoma	0.966	0.184	0.675	0.885	Alphaproteobacteria	0.012	-0.023	0.004967 < Alphaproteobacteria
25	Adenocarcinoma	0.948	0.125	0.683	0.846	Cyanobacteriia	0.000	0.074	Cyanobacteriia <= 0.000151
25	Adenocarcinoma	0.948	0.125	0.683	0.846	Verrucomicrobiae	0.000	0.042	Verrucomicrobiae <= 0.000551
25	Adenocarcinoma	0.948	0.125	0.683	0.846	Bacilli	0.036	0.047	0.0265 < Bacilli <= 0.0576
25	Adenocarcinoma	0.948	0.125	0.683	0.846	Coriobacteriia	0.001	-0.019	Coriobacteriia <= 0.000582
25	Adenocarcinoma	0.948	0.125	0.683	0.846	Desulfovibrionia	0.001	0.020	0.000627 < Desulfovibrionia <= 0.001087
26	Adenocarcinoma	0.884	0.240	0.769	0.756	Bacilli	0.003	-0.086	Bacilli <= 0.0114
26	Adenocarcinoma	0.884	0.240	0.769	0.756	Cyanobacteriia	0.000	-0.069	0.000151 < Cyanobacteriia
26	Adenocarcinoma	0.884	0.240	0.769	0.756	Fusobacteriia	0.019	0.067	0.00727 < Fusobacteriia <= 0.14117
26	Adenocarcinoma	0.884	0.240	0.769	0.756	Verrucomicrobiae	0.000	0.036	Verrucomicrobiae <= 0.000551
26	Adenocarcinoma	0.884	0.240	0.769	0.756	Desulfovibrionia	0.001	0.038	Desulfovibrionia <= 0.000627
29	Adenocarcinoma	0.956	0.177	0.691	0.881	Cyanobacteriia	0.000	0.069	Cyanobacteriia <= 0.000151
29	Adenocarcinoma	0.956	0.177	0.691	0.881	Fusobacteriia	0.040	0.071	0.00727 < Fusobacteriia <= 0.14117
29	Adenocarcinoma	0.956	0.177	0.691	0.881	Bacilli	0.036	0.052	0.0265 < Bacilli <= 0.0576

29	Adenocarcinoma	0.956	0.177	0.691	0.881	Campylobacteria	0.000	-0.033	Campylobacteria <= 0.00011
29	Adenocarcinoma	0.956	0.177	0.691	0.881	Verrucomicrobiae	0.000	0.032	Verrucomicrobiae <= 0.000551
32	Adenocarcinoma	0.962	0.204	0.662	0.938	Cyanobacteriia	0.000	0.072	Cyanobacteriia <= 0.000151
32	Adenocarcinoma	0.962	0.204	0.662	0.938	Fusobacteriia	0.263	0.071	0.14117 < Fusobacteriia
32	Adenocarcinoma	0.962	0.204	0.662	0.938	Bacilli	0.059	0.070	0.0576 < Bacilli
32	Adenocarcinoma	0.962	0.204	0.662	0.938	Verrucomicrobiae	0.000	0.032	Verrucomicrobiae <= 0.000551
32	Adenocarcinoma	0.962	0.204	0.662	0.938	Campylobacteria	0.004	0.029	0.00180 < Campylobacteria
33	Adenocarcinoma	0.850	0.121	0.709	0.790	Cyanobacteriia	0.000	0.076	Cyanobacteriia <= 0.000151
33	Adenocarcinoma	0.850	0.121	0.709	0.790	Verrucomicrobiae	0.000	0.044	Verrucomicrobiae <= 0.000551
33	Adenocarcinoma	0.850	0.121	0.709	0.790	Campylobacteria	0.000	-0.034	Campylobacteria <= 0.00011
33	Adenocarcinoma	0.850	0.121	0.709	0.790	Alphaproteobacteria	0.008	-0.025	0.004967 < Alphaproteobacteria
33	Adenocarcinoma	0.850	0.121	0.709	0.790	Desulfovibrionia	0.003	0.019	0.001087 < Desulfovibrionia <= 0.003165
34	Adenocarcinoma	0.868	0.241	0.770	0.786	Verrucomicrobiae	0.007	-0.078	0.006252 < Verrucomicrobiae
34	Adenocarcinoma	0.868	0.241	0.770	0.786	Cyanobacteriia	0.003	-0.070	0.000151 < Cyanobacteriia
34	Adenocarcinoma	0.868	0.241	0.770	0.786	Fusobacteriia	0.116	0.069	0.00727 < Fusobacteriia <= 0.14117
34	Adenocarcinoma	0.868	0.241	0.770	0.786	Bacilli	0.098	0.058	0.0576 < Bacilli
34	Adenocarcinoma	0.868	0.241	0.770	0.786	Desulfovibrionia	0.000	0.038	Desulfovibrionia <= 0.000627
39	Adenocarcinoma	0.928	0.194	0.739	0.764	Desulfovibrionia	0.006	-0.088	0.003165 < Desulfovibrionia
39	Adenocarcinoma	0.928	0.194	0.739	0.764	Cyanobacteriia	0.000	0.067	Cyanobacteriia <= 0.000151
39	Adenocarcinoma	0.928	0.194	0.739	0.764	Bacilli	0.050	0.049	0.0265 < Bacilli <= 0.0576
39	Adenocarcinoma	0.928	0.194	0.739	0.764	Campylobacteria	0.000	-0.030	Campylobacteria <= 0.00011
39	Adenocarcinoma	0.928	0.194	0.739	0.764	Verrucomicrobiae	0.001	0.027	0.000551 < Verrucomicrobiae <= 0.006252
40	Adenocarcinoma	0.924	0.249	0.768	0.760	Bacilli	0.011	-0.085	Bacilli <= 0.0114
40	Adenocarcinoma	0.924	0.249	0.768	0.760	Fusobacteriia	0.015	0.072	0.00727 < Fusobacteriia <= 0.14117
40	Adenocarcinoma	0.924	0.249	0.768	0.760	Cyanobacteriia	0.000	-0.068	0.000151 < Cyanobacteriia
40	Adenocarcinoma	0.924	0.249	0.768	0.760	Desulfovibrionia	0.000	0.039	Desulfovibrionia <= 0.000627
40	Adenocarcinoma	0.924	0.249	0.768	0.760	Verrucomicrobiae	0.000	0.035	Verrucomicrobiae <= 0.000551
46	Adenocarcinoma	0.820	0.322	0.712	0.777	Fusobacteriia	0.000	-0.123	Fusobacteriia <= 0.00111
46	Adenocarcinoma	0.820	0.322	0.712	0.777	Cyanobacteriia	0.000	0.071	Cyanobacteriia <= 0.000151
46	Adenocarcinoma	0.820	0.322	0.712	0.777	Bacilli	0.058	0.052	0.0265 < Bacilli <= 0.0576
46	Adenocarcinoma	0.820	0.322	0.712	0.777	Verrucomicrobiae	0.000	0.041	Verrucomicrobiae <= 0.000551
46	Adenocarcinoma	0.820	0.322	0.712	0.777	Desulfovibrionia	0.003	0.024	0.001087 < Desulfovibrionia <= 0.003165
47	Adenocarcinoma	0.872	0.283	0.734	0.756	Bacilli	0.009	-0.082	Bacilli <= 0.0114
47	Adenocarcinoma	0.872	0.283	0.734	0.756	Cyanobacteriia	0.000	0.070	Cyanobacteriia <= 0.000151
47	Adenocarcinoma	0.872	0.283	0.734	0.756	Verrucomicrobiae	0.009	-0.072	0.006252 < Verrucomicrobiae
47	Adenocarcinoma	0.872	0.283	0.734	0.756	Fusobacteriia	0.015	0.067	0.00727 < Fusobacteriia <= 0.14117
47	Adenocarcinoma	0.872	0.283	0.734	0.756	Desulfovibrionia	0.000	0.039	Desulfovibrionia <= 0.000627
48	Adenocarcinoma	0.852	0.288	0.754	0.733	Fusobacteriia	0.001	-0.126	Fusobacteriia <= 0.00111
48	Adenocarcinoma	0.852	0.288	0.754	0.733	Cyanobacteriia	0.000	0.069	Cyanobacteriia <= 0.000151
48	Adenocarcinoma	0.852	0.288	0.754	0.733	Desulfovibrionia	0.001	0.036	0.000627 < Desulfovibrionia <= 0.001087
48	Adenocarcinoma	0.852	0.288	0.754	0.733	Campylobacteria	0.000	-0.026	Campylobacteria <= 0.00011
48	Adenocarcinoma	0.852	0.288	0.754	0.733	Verrucomicrobiae	0.003	0.026	0.000551 < Verrucomicrobiae <= 0.006252
49	Adenocarcinoma	0.956	0.199	0.666	0.932	Fusobacteriia	0.189	0.078	0.14117 < Fusobacteriia

49	Adenocarcinoma	0.956	0.199	0.666	0.932	Cyanobacteriia	0.000	0.069	Cyanobacteriia <= 0.000151
49	Adenocarcinoma	0.956	0.199	0.666	0.932	Bacilli	0.105	0.059	0.0576 < Bacilli
49	Adenocarcinoma	0.956	0.199	0.666	0.932	Verrucomicrobiae	0.000	0.032	Verrucomicrobiae <= 0.000551
49	Adenocarcinoma	0.956	0.199	0.666	0.932	Campylobacteria	0.134	0.028	0.00180 < Campylobacteria
50	Adenocarcinoma	0.934	0.242	0.736	0.735	Desulfovibrionia	0.003	-0.078	0.003165 < Desulfovibrionia
50	Adenocarcinoma	0.934	0.242	0.736	0.735	Verrucomicrobiae	0.009	-0.077	0.006252 < Verrucomicrobiae
50	Adenocarcinoma	0.934	0.242	0.736	0.735	Cyanobacteriia	0.000	0.073	Cyanobacteriia <= 0.000151
50	Adenocarcinoma	0.934	0.242	0.736	0.735	Bacilli	0.075	0.056	0.0576 < Bacilli
50	Adenocarcinoma	0.934	0.242	0.736	0.735	Coriobacteriia	0.006	0.024	0.005054 < Coriobacteriia

Table ML-B-5: The LIME Result for Test Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
1	Adenocarcinoma	0.788	0.392	0.776	0.631	Fusobacteriia	0.001	-0.130	Fusobacteriia <= 0.00111
1	Adenocarcinoma	0.788	0.392	0.776	0.631	Desulfovibrionia	0.007	-0.087	0.003165 < Desulfovibrionia
1	Adenocarcinoma	0.788	0.392	0.776	0.631	Cyanobacteriia	0.000	0.071	Cyanobacteriia <= 0.000151
1	Adenocarcinoma	0.788	0.392	0.776	0.631	Verrucomicrobiae	0.000	0.033	Verrucomicrobiae <= 0.000551
1	Adenocarcinoma	0.788	0.392	0.776	0.631	Campylobacteria	0.000	-0.032	Campylobacteria <= 0.00011
4	Adenoma	0.552	0.463	0.186	0.438	Fusobacteriia	0.000	0.124	Fusobacteriia <= 0.00111
4	Adenoma	0.552	0.463	0.186	0.438	Bacilli	0.008	0.087	Bacilli <= 0.0114
4	Adenoma	0.552	0.463	0.186	0.438	Desulfovibrionia	0.004	0.086	0.003165 < Desulfovibrionia
4	Adenoma	0.552	0.463	0.186	0.438	Cyanobacteriia	0.000	-0.074	Cyanobacteriia <= 0.000151
4	Adenoma	0.552	0.463	0.186	0.438	Campylobacteria	0.000	0.030	Campylobacteria <= 0.00011
5	Adenocarcinoma	0.832	0.305	0.760	0.733	Fusobacteriia	0.001	-0.125	Fusobacteriia <= 0.00111
5	Adenocarcinoma	0.832	0.305	0.760	0.733	Cyanobacteriia	0.000	0.065	Cyanobacteriia <= 0.000151
5	Adenocarcinoma	0.832	0.305	0.760	0.733	Campylobacteria	0.000	-0.031	Campylobacteria <= 0.00011
5	Adenocarcinoma	0.832	0.305	0.760	0.733	Desulfovibrionia	0.000	0.034	Desulfovibrionia <= 0.000627
5	Adenocarcinoma	0.832	0.305	0.760	0.733	Verrucomicrobiae	0.001	0.029	0.000551 < Verrucomicrobiae <= 0.006252
8	Adenocarcinoma	0.616	0.359	0.786	0.635	Fusobacteriia	0.000	-0.125	Fusobacteriia <= 0.00111
8	Adenocarcinoma	0.616	0.359	0.786	0.635	Verrucomicrobiae	0.076	-0.084	0.006252 < Verrucomicrobiae
8	Adenocarcinoma	0.616	0.359	0.786	0.635	Cyanobacteriia	0.000	0.067	Cyanobacteriia <= 0.000151
8	Adenocarcinoma	0.616	0.359	0.786	0.635	Campylobacteria	0.000	-0.032	Campylobacteria <= 0.00011
8	Adenocarcinoma	0.616	0.359	0.786	0.635	Coriobacteriia	0.009	0.024	0.005054 < Coriobacteriia
9	Adenocarcinoma	0.532	0.392	0.781	0.663	Fusobacteriia	0.000	-0.128	Fusobacteriia <= 0.00111
9	Adenocarcinoma	0.532	0.392	0.781	0.663	Verrucomicrobiae	0.009	-0.081	0.006252 < Verrucomicrobiae
9	Adenocarcinoma	0.532	0.392	0.781	0.663	Cyanobacteriia	0.000	0.070	Cyanobacteriia <= 0.000151
9	Adenocarcinoma	0.532	0.392	0.781	0.663	Bacilli	0.066	0.056	0.0576 < Bacilli
9	Adenocarcinoma	0.532	0.392	0.781	0.663	Campylobacteria	0.000	-0.035	Campylobacteria <= 0.00011
12	Adenocarcinoma	0.744	0.095	0.774	0.760	Cyanobacteriia	0.002	-0.074	0.000151 < Cyanobacteriia
12	Adenocarcinoma	0.744	0.095	0.774	0.760	Desulfovibrionia	0.000	0.030	Desulfovibrionia <= 0.000627
12	Adenocarcinoma	0.744	0.095	0.774	0.760	Verrucomicrobiae	0.001	0.031	0.000551 < Verrucomicrobiae <= 0.006252
12	Adenocarcinoma	0.744	0.095	0.774	0.760	Actinobacteria	0.034	-0.021	0.00838 < Actinobacteria

12	Adenocarcinoma	0.744	0.095	0.774	0.760	Alphaproteobacteria	0.002	0.020	0.000814 < Alphaproteobacteria <= 0.002119
13	Adenocarcinoma	0.980	0.207	0.660	0.938	Fusobacteriia	0.440	0.075	0.14117 < Fusobacteriia
13	Adenocarcinoma	0.980	0.207	0.660	0.938	Cyanobacteriia	0.000	0.068	Cyanobacteriia <= 0.000151
13	Adenocarcinoma	0.980	0.207	0.660	0.938	Bacilli	0.078	0.062	0.0576 < Bacilli
13	Adenocarcinoma	0.980	0.207	0.660	0.938	Verrucomicrobiae	0.001	0.039	Verrucomicrobiae <= 0.000551
13	Adenocarcinoma	0.980	0.207	0.660	0.938	Desulfovibrionia	0.000	0.034	Desulfovibrionia <= 0.000627
15	Adenocarcinoma	0.686	0.327	0.731	0.758	Desulfovibrionia	0.006	-0.092	0.003165 < Desulfovibrionia
15	Adenocarcinoma	0.686	0.327	0.731	0.758	Verrucomicrobiae	0.081	-0.076	0.006252 < Verrucomicrobiae
15	Adenocarcinoma	0.686	0.327	0.731	0.758	Cyanobacteriia	0.000	0.073	Cyanobacteriia <= 0.000151
15	Adenocarcinoma	0.686	0.327	0.731	0.758	Fusobacteriia	0.089	0.064	0.00727 < Fusobacteriia <= 0.14117
15	Adenocarcinoma	0.686	0.327	0.731	0.758	Bacilli	0.071	0.058	0.0576 < Bacilli
16	Adenocarcinoma	0.876	0.292	0.797	0.709	Desulfovibrionia	0.010	-0.080	0.003165 < Desulfovibrionia
16	Adenocarcinoma	0.876	0.292	0.797	0.709	Cyanobacteriia	0.000	-0.069	0.000151 < Cyanobacteriia
16	Adenocarcinoma	0.876	0.292	0.797	0.709	Fusobacteriia	0.121	0.069	0.00727 < Fusobacteriia <= 0.14117
16	Adenocarcinoma	0.876	0.292	0.797	0.709	Verrucomicrobiae	0.009	-0.067	0.006252 < Verrucomicrobiae
16	Adenocarcinoma	0.876	0.292	0.797	0.709	Bacilli	0.074	0.059	0.0576 < Bacilli
17	Adenocarcinoma	0.518	0.389	0.772	0.666	Fusobacteriia	0.000	-0.120	Fusobacteriia <= 0.00111
17	Adenocarcinoma	0.518	0.389	0.772	0.666	Desulfovibrionia	0.014	-0.081	0.003165 < Desulfovibrionia
17	Adenocarcinoma	0.518	0.389	0.772	0.666	Cyanobacteriia	0.000	0.080	Cyanobacteriia <= 0.000151
17	Adenocarcinoma	0.518	0.389	0.772	0.666	Bacilli	0.030	0.047	0.0265 < Bacilli <= 0.0576
17	Adenocarcinoma	0.518	0.389	0.772	0.666	Campylobacteria	0.000	-0.032	Campylobacteria <= 0.00011
18	Adenoma	0.566	0.453	0.188	0.439	Fusobacteriia	0.000	0.131	Fusobacteriia <= 0.00111
18	Adenoma	0.566	0.453	0.188	0.439	Bacilli	0.011	0.087	Bacilli <= 0.0114
18	Adenoma	0.566	0.453	0.188	0.439	Verrucomicrobiae	0.055	0.076	0.006252 < Verrucomicrobiae
18	Adenoma	0.566	0.453	0.188	0.439	Cyanobacteriia	0.000	-0.076	Cyanobacteriia <= 0.000151
18	Adenoma	0.566	0.453	0.188	0.439	Campylobacteria	0.000	0.033	Campylobacteria <= 0.00011
19	Adenocarcinoma	0.616	0.400	0.848	0.574	Fusobacteriia	0.001	-0.126	Fusobacteriia <= 0.00111
19	Adenocarcinoma	0.616	0.400	0.848	0.574	Bacilli	0.010	-0.084	Bacilli <= 0.0114
19	Adenocarcinoma	0.616	0.400	0.848	0.574	Cyanobacteriia	0.000	-0.072	0.000151 < Cyanobacteriia
19	Adenocarcinoma	0.616	0.400	0.848	0.574	Verrucomicrobiae	0.000	0.037	Verrucomicrobiae <= 0.000551
19	Adenocarcinoma	0.616	0.400	0.848	0.574	Campylobacteria	0.000	-0.030	Campylobacteria <= 0.00011
27	Adenocarcinoma	0.918	0.141	0.722	0.833	Fusobacteriia	0.095	0.070	0.00727 < Fusobacteriia <= 0.14117
27	Adenocarcinoma	0.918	0.141	0.722	0.833	Cyanobacteriia	0.000	0.068	Cyanobacteriia <= 0.000151
27	Adenocarcinoma	0.918	0.141	0.722	0.833	Campylobacteria	0.000	-0.030	Campylobacteria <= 0.00011
27	Adenocarcinoma	0.918	0.141	0.722	0.833	Verrucomicrobiae	0.002	0.027	0.000551 < Verrucomicrobiae <= 0.006252
27	Adenocarcinoma	0.918	0.141	0.722	0.833	Alphaproteobacteria	0.014	-0.023	0.004967 < Alphaproteobacteria
28	Adenocarcinoma	0.948	0.254	0.731	0.811	Desulfovibrionia	0.005	-0.084	0.003165 < Desulfovibrionia
28	Adenocarcinoma	0.948	0.254	0.731	0.811	Fusobacteriia	0.097	0.071	0.00727 < Fusobacteriia <= 0.14117
28	Adenocarcinoma	0.948	0.254	0.731	0.811	Cyanobacteriia	0.000	0.065	Cyanobacteriia <= 0.000151
28	Adenocarcinoma	0.948	0.254	0.731	0.811	Bacilli	0.087	0.066	0.0576 < Bacilli
28	Adenocarcinoma	0.948	0.254	0.731	0.811	Campylobacteria	0.000	-0.038	Campylobacteria <= 0.00011
30	Adenocarcinoma	0.868	0.300	0.735	0.756	Desulfovibrionia	0.004	-0.089	0.003165 < Desulfovibrionia
30	Adenocarcinoma	0.868	0.300	0.735	0.756	Verrucomicrobiae	0.021	-0.076	0.006252 < Verrucomicrobiae

30	Adenocarcinoma	0.868	0.300	0.735	0.756	Cyanobacteriia	0.000	0.070	Cyanobacteriia <= 0.000151
30	Adenocarcinoma	0.868	0.300	0.735	0.756	Fusobacteriia	0.125	0.066	0.00727 < Fusobacteriia <= 0.14117
30	Adenocarcinoma	0.868	0.300	0.735	0.756	Bacilli	0.038	0.049	0.0265 < Bacilli <= 0.0576
31	Adenocarcinoma	0.702	0.222	0.696	0.805	Desulfovibrionia	0.010	-0.088	0.003165 < Desulfovibrionia
31	Adenocarcinoma	0.702	0.222	0.696	0.805	Cyanobacteriia	0.000	0.076	Cyanobacteriia <= 0.000151
31	Adenocarcinoma	0.702	0.222	0.696	0.805	Bacilli	0.064	0.061	0.0576 < Bacilli
31	Adenocarcinoma	0.702	0.222	0.696	0.805	Verrucomicrobiae	0.000	0.038	Verrucomicrobiae <= 0.000551
31	Adenocarcinoma	0.702	0.222	0.696	0.805	Campylobacteria	0.000	0.022	0.00011 < Campylobacteria <= 0.00180
35	Adenocarcinoma	0.742	0.226	0.795	0.701	Desulfovibrionia	0.005	-0.085	0.003165 < Desulfovibrionia
35	Adenocarcinoma	0.742	0.226	0.795	0.701	Cyanobacteriia	0.001	-0.075	0.000151 < Cyanobacteriia
35	Adenocarcinoma	0.742	0.226	0.795	0.701	Verrucomicrobiae	0.000	0.045	Verrucomicrobiae <= 0.000551
35	Adenocarcinoma	0.742	0.226	0.795	0.701	Bacilli	0.031	0.050	0.0265 < Bacilli <= 0.0576
35	Adenocarcinoma	0.742	0.226	0.795	0.701	Campylobacteria	0.000	-0.029	Campylobacteria <= 0.00011
36	Adenocarcinoma	0.920	0.245	0.687	0.841	Desulfovibrionia	0.004	-0.080	0.003165 < Desulfovibrionia
36	Adenocarcinoma	0.920	0.245	0.687	0.841	Cyanobacteriia	0.000	0.077	Cyanobacteriia <= 0.000151
36	Adenocarcinoma	0.920	0.245	0.687	0.841	Fusobacteriia	0.079	0.064	0.00727 < Fusobacteriia <= 0.14117
36	Adenocarcinoma	0.920	0.245	0.687	0.841	Bacilli	0.032	0.054	0.0265 < Bacilli <= 0.0576
36	Adenocarcinoma	0.920	0.245	0.687	0.841	Verrucomicrobiae	0.000	0.038	Verrucomicrobiae <= 0.000551
37	Adenocarcinoma	0.702	0.222	0.706	0.823	Bacilli	0.002	-0.080	Bacilli <= 0.0114
37	Adenocarcinoma	0.702	0.222	0.706	0.823	Fusobacteriia	0.190	0.076	0.14117 < Fusobacteriia
37	Adenocarcinoma	0.702	0.222	0.706	0.823	Cyanobacteriia	0.000	0.068	Cyanobacteriia <= 0.000151
37	Adenocarcinoma	0.702	0.222	0.706	0.823	Verrucomicrobiae	0.000	0.031	Verrucomicrobiae <= 0.000551
37	Adenocarcinoma	0.702	0.222	0.706	0.823	Campylobacteria	0.043	0.023	0.00180 < Campylobacteria
38	Adenocarcinoma	0.882	0.254	0.768	0.790	Cyanobacteriia	0.001	-0.075	0.000151 < Cyanobacteriia
38	Adenocarcinoma	0.882	0.254	0.768	0.790	Fusobacteriia	0.186	0.076	0.14117 < Fusobacteriia
38	Adenocarcinoma	0.882	0.254	0.768	0.790	Verrucomicrobiae	0.026	-0.072	0.006252 < Verrucomicrobiae
38	Adenocarcinoma	0.882	0.254	0.768	0.790	Bacilli	0.097	0.057	0.0576 < Bacilli
38	Adenocarcinoma	0.882	0.254	0.768	0.790	Desulfovibrionia	0.000	0.036	Desulfovibrionia <= 0.000627
41	Adenocarcinoma	0.834	0.290	0.731	0.751	Desulfovibrionia	0.004	-0.091	0.003165 < Desulfovibrionia
41	Adenocarcinoma	0.834	0.290	0.731	0.751	Verrucomicrobiae	0.252	-0.073	0.006252 < Verrucomicrobiae
41	Adenocarcinoma	0.834	0.290	0.731	0.751	Cyanobacteriia	0.000	0.070	Cyanobacteriia <= 0.000151
41	Adenocarcinoma	0.834	0.290	0.731	0.751	Fusobacteriia	0.193	0.069	0.14117 < Fusobacteriia
41	Adenocarcinoma	0.834	0.290	0.731	0.751	Bacilli	0.045	0.046	0.0265 < Bacilli <= 0.0576
42	Adenocarcinoma	0.802	0.141	0.700	0.850	Cyanobacteriia	0.000	0.074	Cyanobacteriia <= 0.000151
42	Adenocarcinoma	0.802	0.141	0.700	0.850	Bacilli	0.028	0.048	0.0265 < Bacilli <= 0.0576
42	Adenocarcinoma	0.802	0.141	0.700	0.850	Verrucomicrobiae	0.000	0.036	Verrucomicrobiae <= 0.000551
42	Adenocarcinoma	0.802	0.141	0.700	0.850	Campylobacteria	0.000	-0.033	Campylobacteria <= 0.00011
42	Adenocarcinoma	0.802	0.141	0.700	0.850	Coriobacteriia	0.005	0.026	0.005054 < Coriobacteriia
43	Adenoma	0.506	0.290	0.228	0.309	Fusobacteriia	0.000	0.124	Fusobacteriia <= 0.00111
43	Adenoma	0.506	0.290	0.228	0.309	Cyanobacteriia	0.000	-0.070	Cyanobacteriia <= 0.000151
43	Adenoma	0.506	0.290	0.228	0.309	Campylobacteria	0.000	0.028	Campylobacteria <= 0.00011
43	Adenoma	0.506	0.290	0.228	0.309	Verrucomicrobiae	0.003	-0.025	0.000551 < Verrucomicrobiae <= 0.006252
43	Adenoma	0.506	0.290	0.228	0.309	Alphaproteobacteria	0.006	0.024	0.004967 < Alphaproteobacteria

44	Adenocarcinoma	0.608	0.443	0.778	0.631	Fusobacteriia	0.000	-0.124	Fusobacteriia <= 0.00111
44	Adenocarcinoma	0.608	0.443	0.778	0.631	Desulfovibrionia	0.004	-0.082	0.003165 < Desulfovibrionia
44	Adenocarcinoma	0.608	0.443	0.778	0.631	Verrucomicrobiae	0.074	-0.071	0.006252 < Verrucomicrobiae
44	Adenocarcinoma	0.608	0.443	0.778	0.631	Cyanobacteriia	0.000	0.069	Cyanobacteriia <= 0.000151
44	Adenocarcinoma	0.608	0.443	0.778	0.631	Bacilli	0.067	0.061	0.0576 < Bacilli
45	Adenocarcinoma	0.526	0.228	0.773	0.682	Bacilli	0.006	-0.084	Bacilli <= 0.0114
45	Adenocarcinoma	0.526	0.228	0.773	0.682	Verrucomicrobiae	0.052	-0.072	0.006252 < Verrucomicrobiae
45	Adenocarcinoma	0.526	0.228	0.773	0.682	Cyanobacteriia	0.000	0.063	Cyanobacteriia <= 0.000151
45	Adenocarcinoma	0.526	0.228	0.773	0.682	Campylobacteria	0.000	-0.028	Campylobacteria <= 0.00011
45	Adenocarcinoma	0.526	0.228	0.773	0.682	Coriobacteriia	0.006	0.030	0.005054 < Coriobacteriia

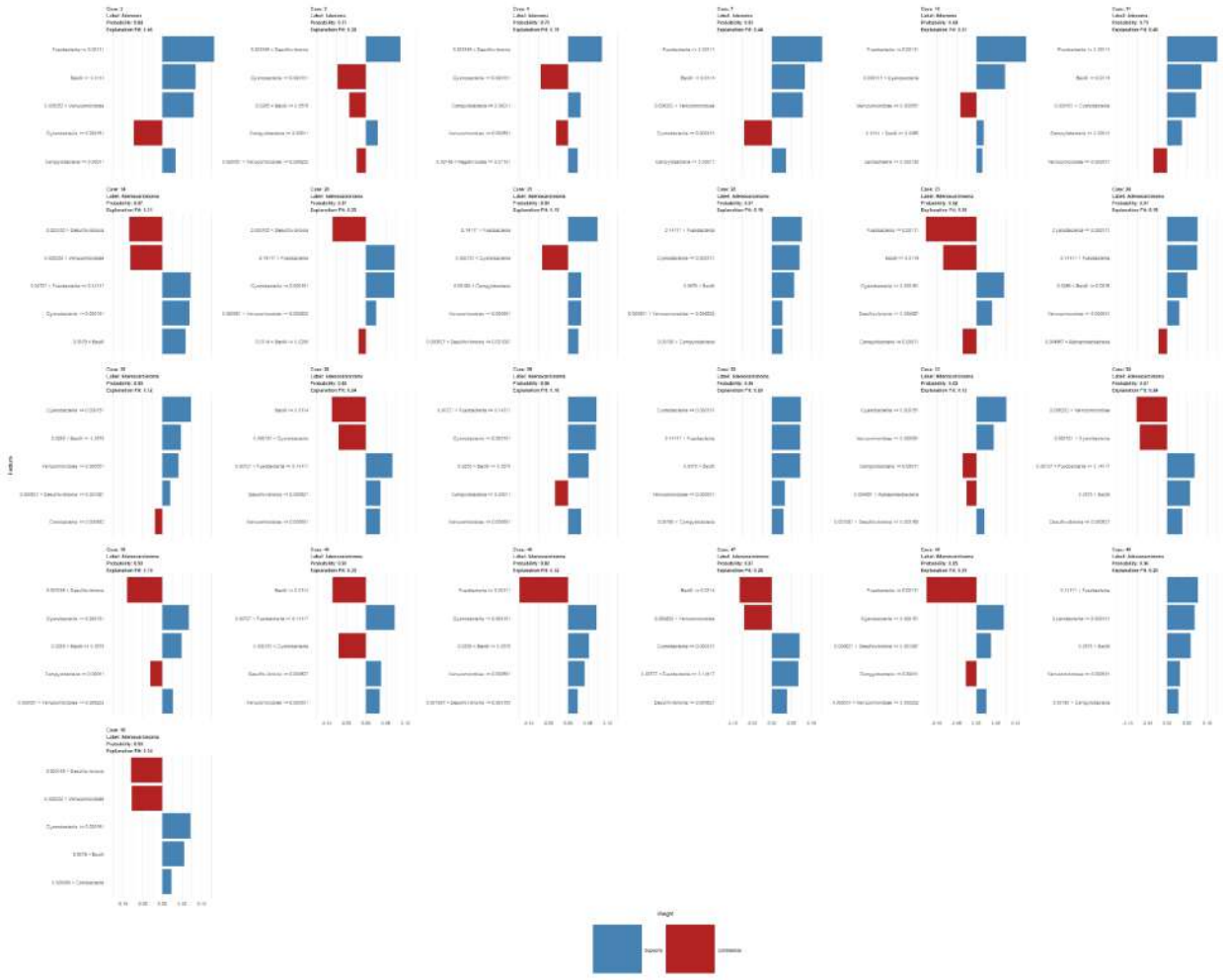


Figure ML1-B-6: The LIME Result in **Training dataset** for **Class** in **Biopsy**

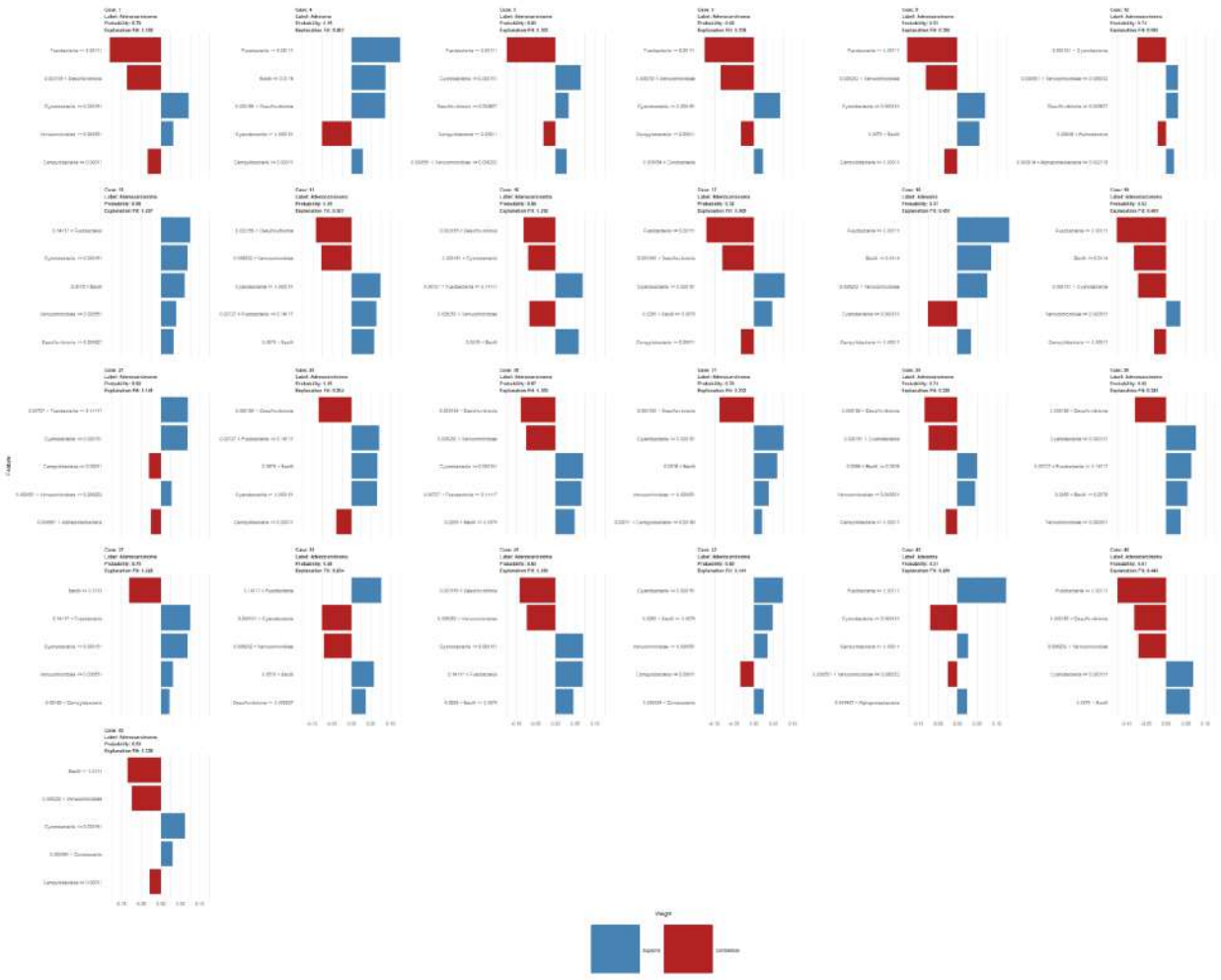


Figure ML1-B-7: The LIME Result in **Test dataset** for **Class** in **Biopsy**

Table ML1-B-6: The Aggregated SHAP Values

Variable	Training Dataset				Testing Dataset			
	contribution	position	sign	cumulative	contribution	position	sign	cumulative
intercept	0.7524	33	X	0.752	0.7524	33	X	0.752
Gender	0.0000	32	1	0.752	0.0000	32	1	0.752
Campylobacteria	0.0002	31	1	0.753	-0.0034	30	-1	0.749
Coriobacteriia	-0.0001	30	-1	0.753	0.0071	29	1	0.756

Bacteroidia	0.0000	29	1	0.753	0.0004	26	1	0.761
Verrucomicrobiae	-0.0001	28	-1	0.752	0.0040	28	1	0.760
Gammaproteobacteria	0.0000	27	1	0.753	0.0002	27	1	0.760
Cyanobacteriia	0.0001	26	1	0.753	0.0001	23	1	0.761
Tumor	0.0000	25	1	0.753	0.0000	25	1	0.761
Negativicutes	0.0001	24	1	0.753	0.0028	22	1	0.764
Desulfobacteria	0.0000	23	1	0.753	0.0001	21	1	0.764
KD4.96	0.0000	22	1	0.753	0.0000	24	1	0.761
Polyangia	0.0000	21	1	0.753	0.0000	31	-1	0.752
Bacilli	0.0006	20	1	0.753	0.0087	19	1	0.772
Actinobacteria	0.0000	19	1	0.753	-0.0014	18	-1	0.770
Alphaproteobacteria	0.0000	18	-1	0.753	-0.0025	17	-1	0.768
Desulfobulbia	0.0000	17	1	0.753	0.0000	16	-1	0.768
Fusobacteriia	-0.0006	16	-1	0.753	-0.0308	9	-1	0.734
Vampirivibrionia	-0.0001	15	-1	0.753	-0.0008	20	-1	0.763
Ignavibacteria	0.0000	14	1	0.753	0.0000	14	1	0.769
Clostridia	-0.0002	13	-1	0.752	-0.0020	12	-1	0.764
Spirochaetia	0.0000	12	-1	0.752	-0.0002	11	-1	0.764
Methanobacteria	0.0002	11	1	0.753	0.0011	15	1	0.769
Thermoleophilia	0.0000	10	1	0.753	0.0008	10	1	0.765
Saccharimonadia	0.0000	9	-1	0.753	-0.0005	8	-1	0.734
Synergistia	0.0000	8	-1	0.753	-0.0008	7	-1	0.733
Planctomycetes	0.0001	7	1	0.753	-0.0022	13	-1	0.766
Desulfovibrionia	-0.0004	6	-1	0.752	-0.0110	6	-1	0.722
Acidimicrobiia	0.0000	5	-1	0.752	-0.0003	5	-1	0.722
Acidobacteriae	0.0000	4	-1	0.752	0.0001	4	1	0.722
Incertae_Sedis	0.0001	3	1	0.752	0.0001	3	1	0.722
Lentisphaeria	0.0001	2	1	0.752	-0.0012	2	-1	0.721
prediction	0.7524	1	X	0.752	0.7206	1	X	0.721

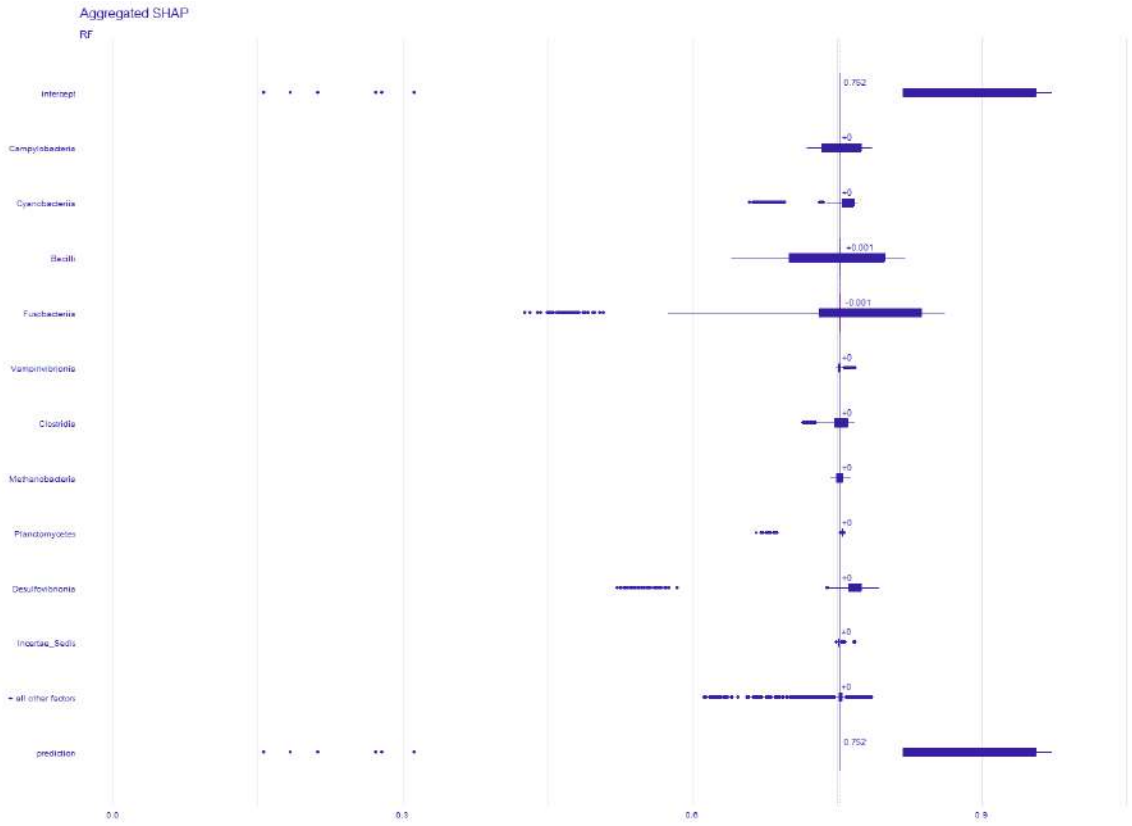


Figure ML1-B-8: The Aggregated SHAP Value in **Training dataset** for **Class** in **Biopsy**

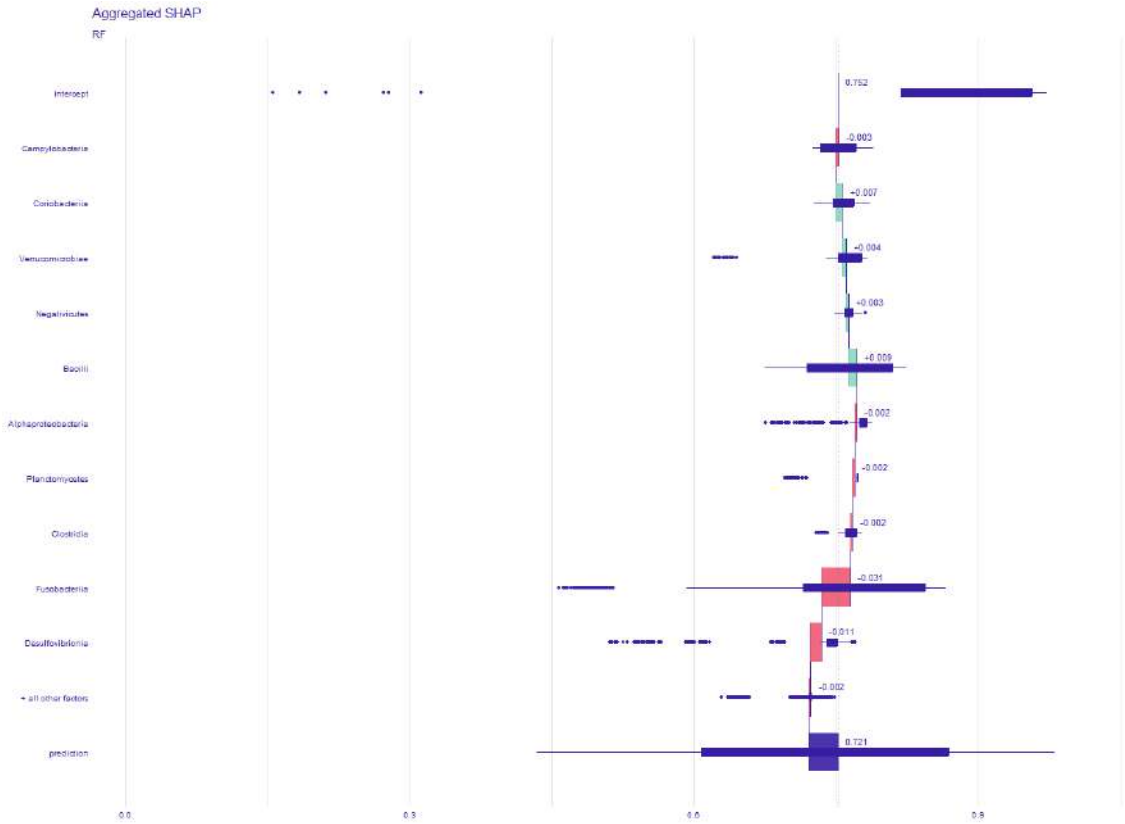


Figure ML1-B-9: The Aggregated SHAP Value in **Test dataset** for **Class** in **Biopsy**

C- Order

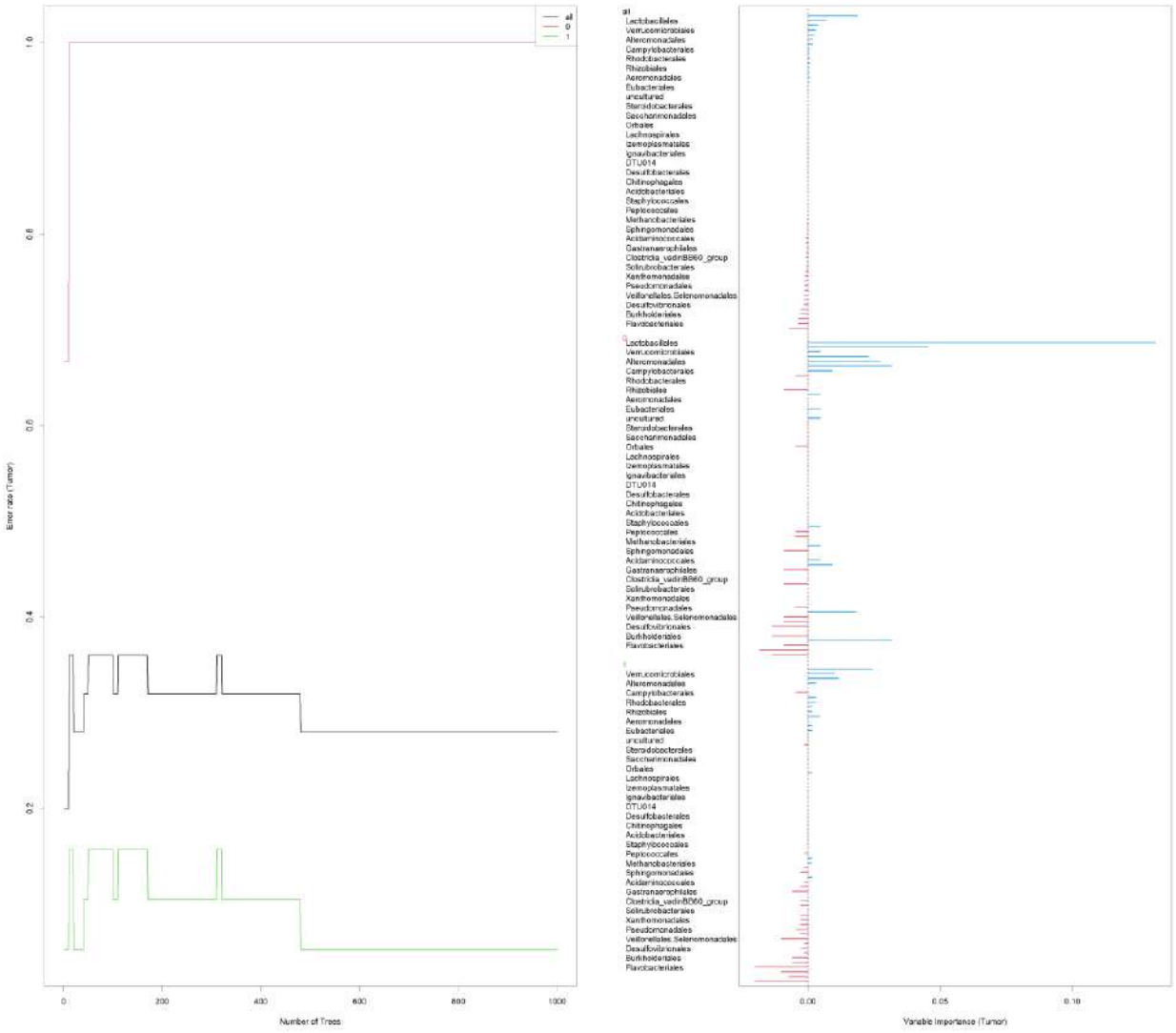
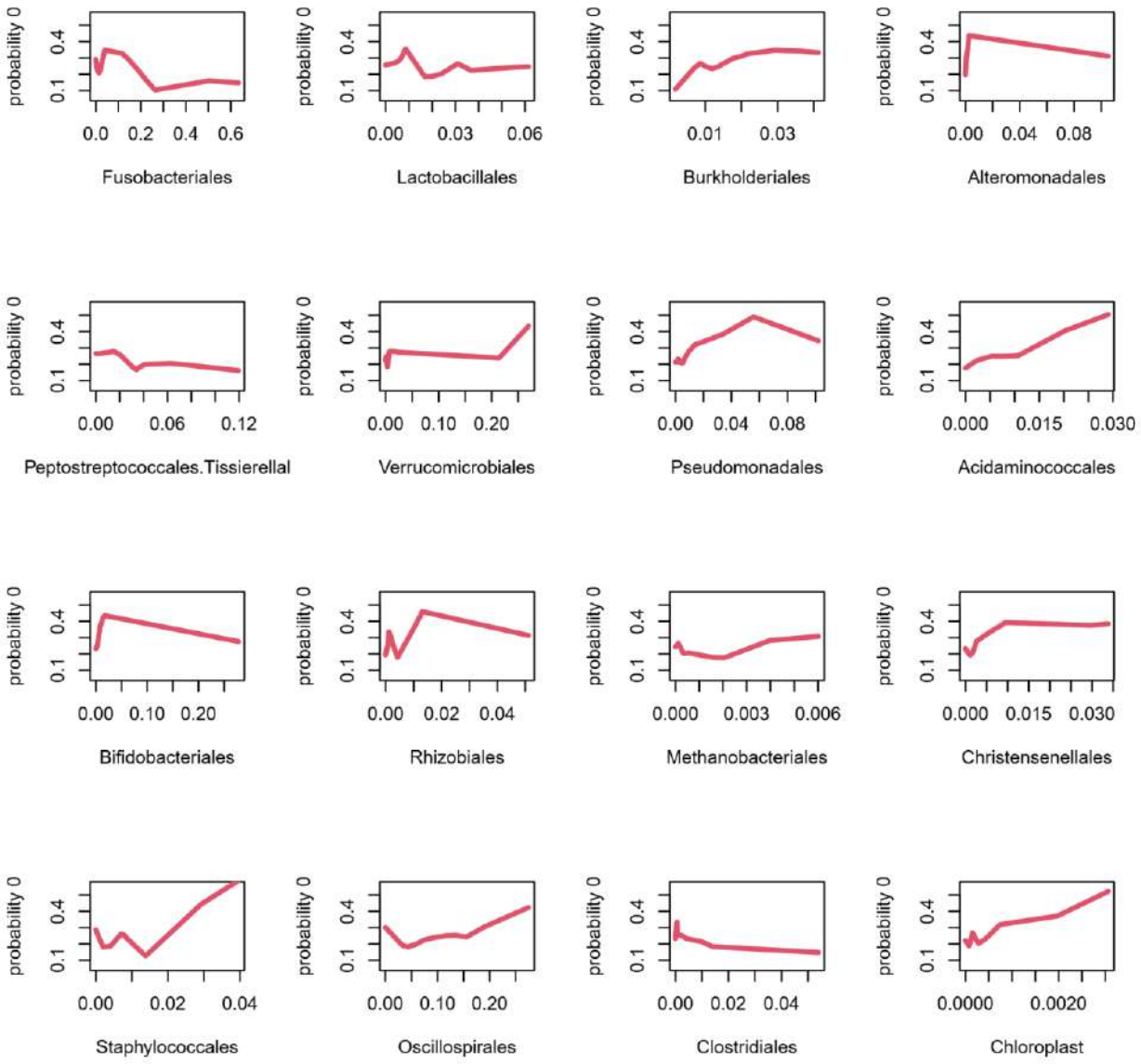
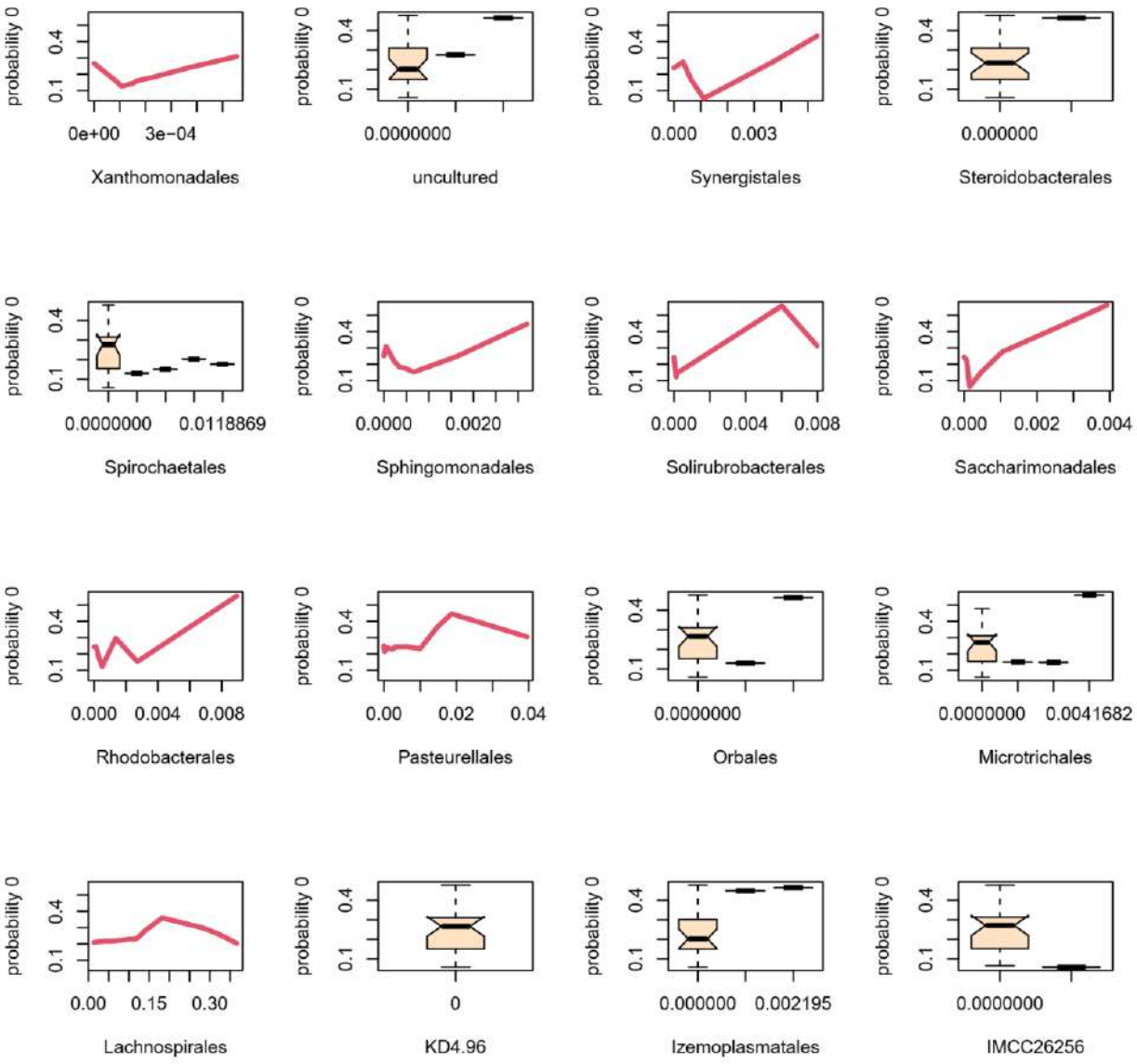
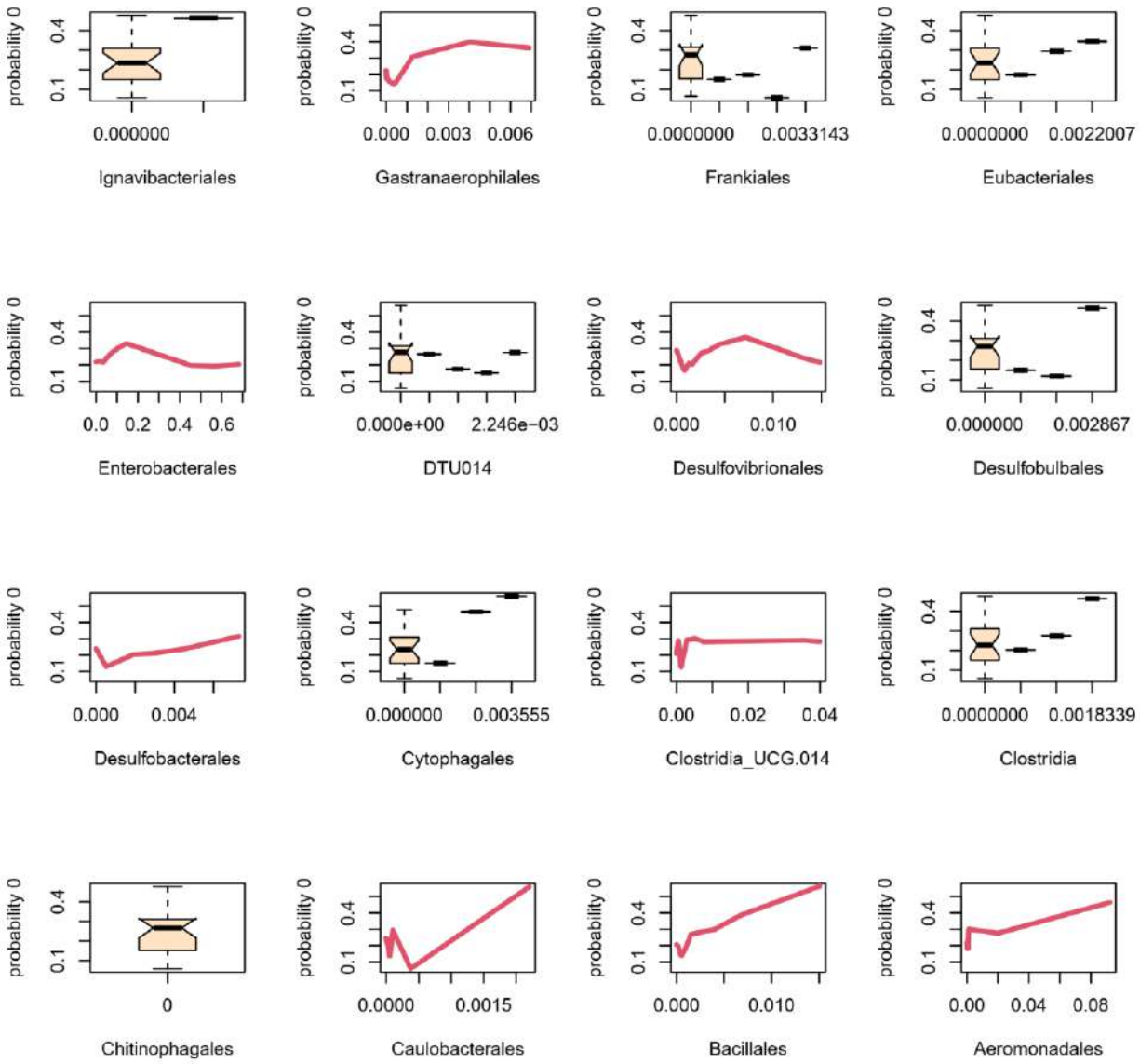


Figure ML1-C-1: The number of trees and Variable Importance Plot for Order in Biopsy


probability 0
0.1
0.4
0.0 0.2 0.4 0.6
Fusobacteriales
0.00 0.03 0.06
Lactobacillales
0.01 0.03
Burkholderiales
0.00 0.04 0.08
Alteromonadales
0.00 0.06 0.12
Peptostreptococcales.Tissierellal
0.00 0.10 0.20
Verrucomicrobiales
0.00 0.04 0.08
Pseudomonadales
0.000 0.015 0.030
Acidaminococcales
0.00 0.10 0.20
Bifidobacteriales
0.00 0.02 0.04
Rhizobiales
0.000 0.003 0.006
Methanobacteriales
0.000 0.015 0.030
Christensenellales
0.00 0.02 0.04
Staphylococcales
0.00 0.10 0.20
Oscillospirales
0.00 0.02 0.04
Clostridiales
0.0000 0.0020
Chloroplast


probability 0
0.1
0.4
0e+00
3e-04
Xanthomonadales
0.0000000
uncultured
0.000
0.003
Synergistales
0.000000
Steroidobacterales
0.0000000
0.0118869
Spirochaetales
0.0000
0.0020
Sphingomonadales
0.000
0.004
0.008
Solirubrobacterales
0.002
Saccharimonadales
Rhodobacterales
0.00
0.02
0.04
Pasteurellales
Orbales
0.0041682
Microtrichales
0.15
0.30
Lachnospirales
0
KD4.96
0.002195
Izemoplasmatales
IMCC26256


probability 0
0.1
0.4
0.000000
Ignavibacteriales
0.000
0.003
0.006
Gastranaerophilales
0.0000000
0.0033143
Frankiales
0.0000000
0.0022007
Eubacteriales
0.0
0.2
0.4
0.6
Enterobacterales
0.000e+00
2.246e−03
DTU014
0.000
0.010
Desulfovibrionales
0.000000
0.002867
Desulfobulbales
0.000
0.004
Desulfobacterales
0.000000
0.003555
Cytophagales
0.00
0.02
0.04
Clostridia_UCG.014
0.0000000
0.0018339
Clostridia
0
Chitinophagales
0.0000
0.0015
Caulobacterales
0.000
0.010
Bacillales
0.00
0.04
0.08
Aeromonadales

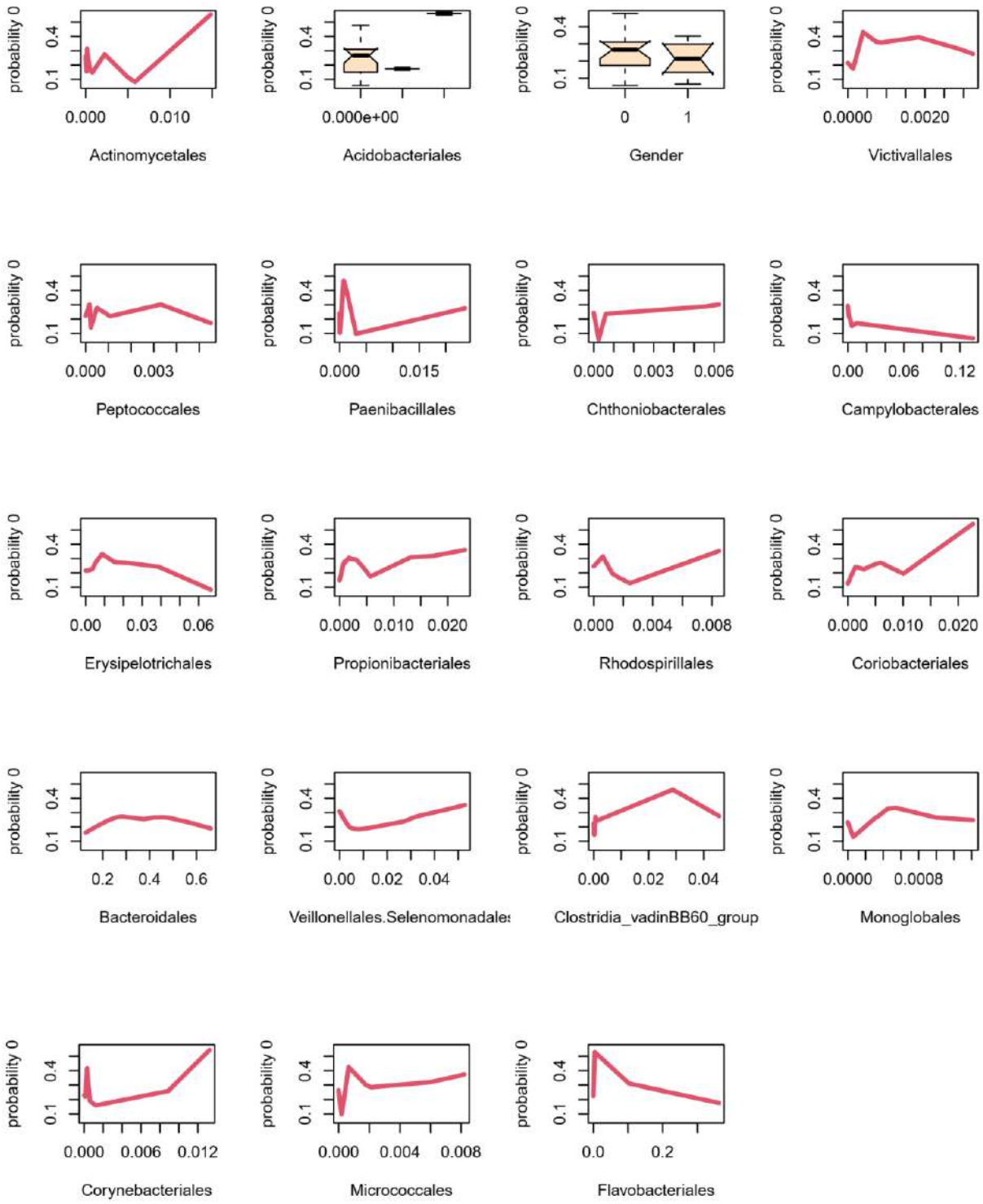
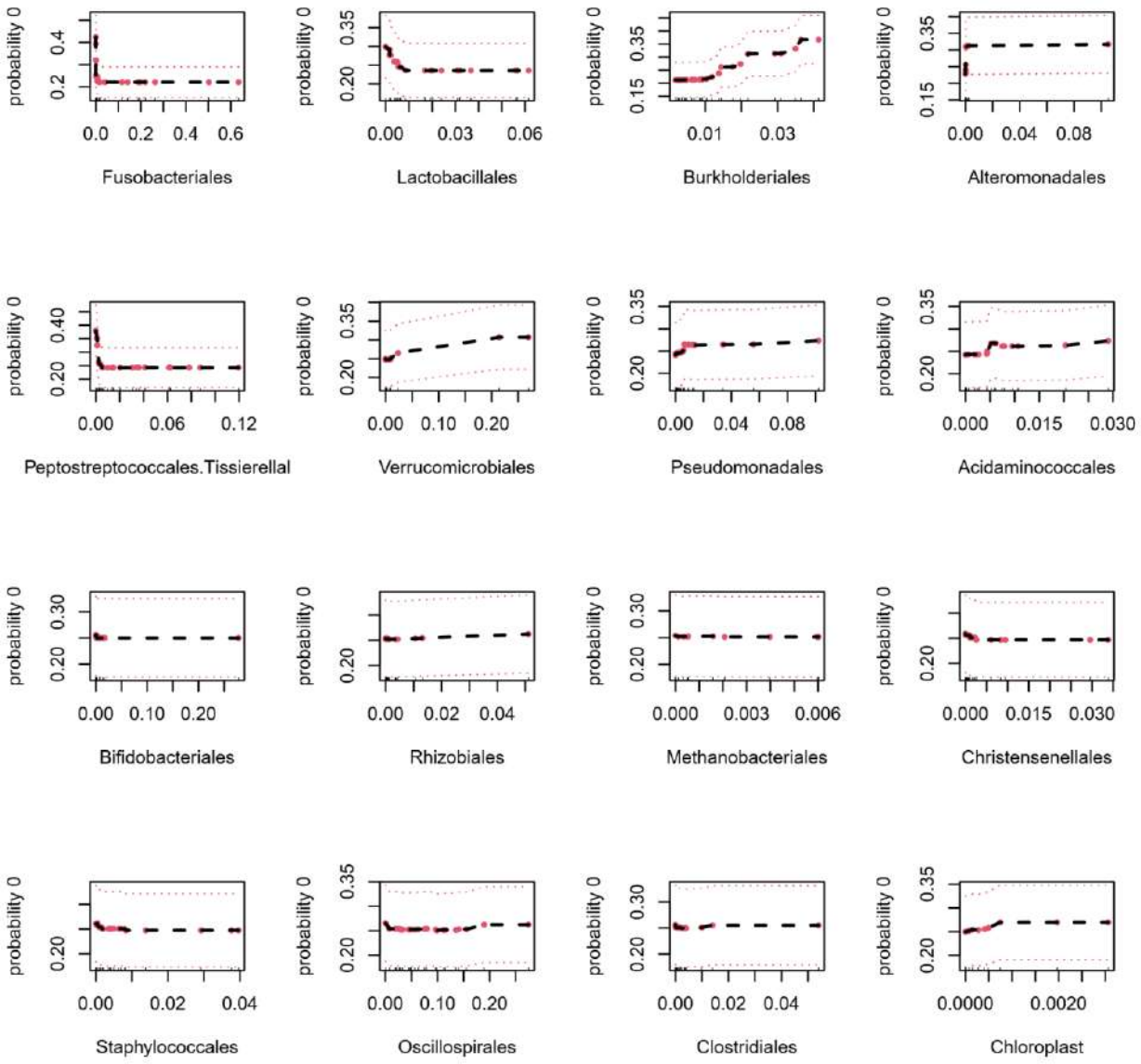
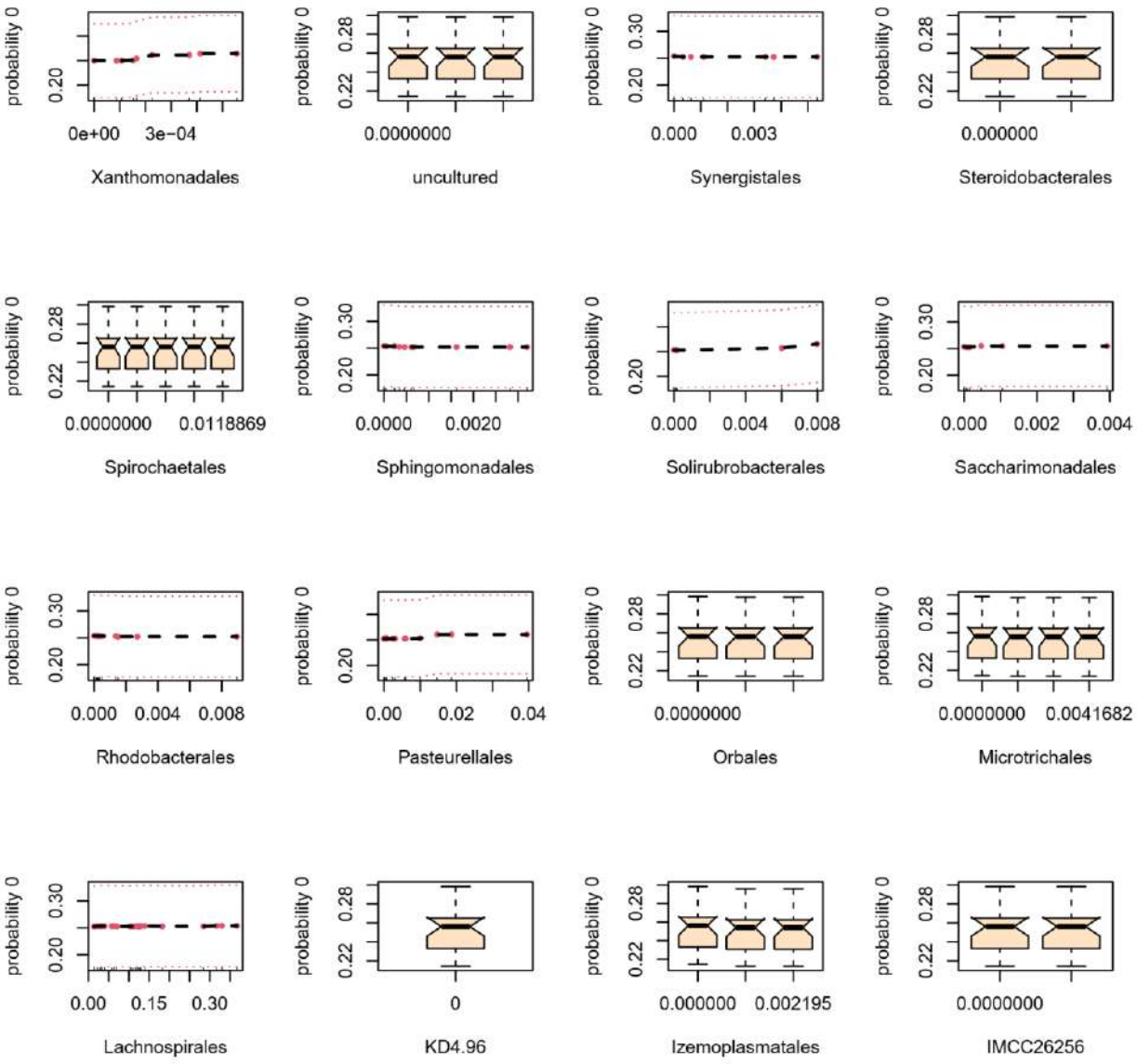
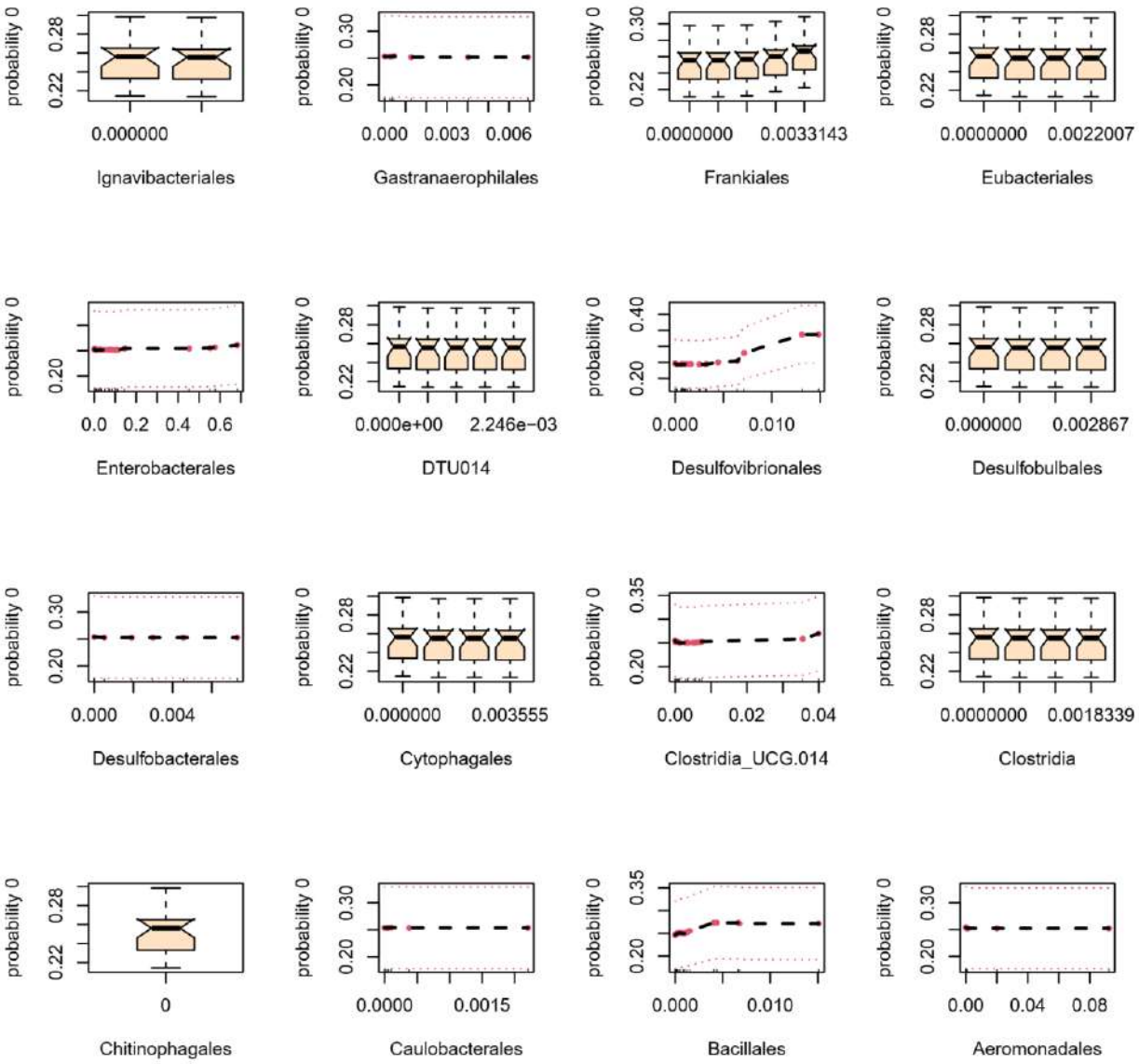


Figure ML1-C-2: Marginal Plot for Order in Biopsy


probability 0
Fusobacteriales
Lactobacillales
Burkholderiales
Alteromonadales
Peptostreptococcales.Tissierellal
Verrucomicrobiales
Pseudomonadales
Acidaminococcales
Bifidobacteriales
Rhizobiales
Methanobacteriales
Christensenellales
Staphylococcales
Oscillospirales
Clostridiales
Chloroplast


probability 0
0.20
0e+00
3e-04
Xanthomonadales
probability 0
0.22
0.28
0.0000000
uncultured
probability 0
0.20
0.30
0.000
0.003
Synergistales
probability 0
0.22
0.28
0.000000
Steroidobacterales
probability 0
0.22
0.28
0.0000000
0.0118869
Spirochaetales
probability 0
0.20
0.30
0.0000
0.0020
Sphingomonadales
probability 0
0.20
0.000
0.004
0.008
Solirubrobacterales
probability 0
0.20
0.30
0.000
0.002
0.004
Saccharimonadales
probability 0
0.20
0.30
0.000
0.004
0.008
Rhodobacterales
probability 0
0.20
0.00
0.02
0.04
Pasteurellales
probability 0
0.22
0.28
0.0000000
Orbales
probability 0
0.22
0.28
0.0000000
0.0041682
Microtrichales
probability 0
0.20
0.30
0.00
0.15
0.30
Lachnospirales
probability 0
0.22
0.28
0
KD4.96
probability 0
0.22
0.28
0.000000
0.002195
Izemoplasmatales
probability 0
0.22
0.28
0.0000000
IMCC26256


probability 0
0.22
0.28
0.000000
Ignavibacteriales
probability 0
0.20
0.30
0.000
0.003
0.006
Gastranaerophilales
probability 0
0.22
0.30
0.0000000
0.0033143
Frankiales
probability 0
0.22
0.28
0.0000000
0.0022007
Eubacteriales
probability 0
0.20
0.0
0.2
0.4
0.6
Enterobacterales
probability 0
0.22
0.28
0.000e+00
2.246e−03
DTU014
probability 0
0.20
0.40
0.000
0.010
Desulfovibrionales
probability 0
0.22
0.28
0.000000
0.002867
Desulfobulbales
probability 0
0.20
0.30
0.000
0.004
Desulfobacterales
probability 0
0.22
0.28
0.000000
0.003555
Cytophagales
probability 0
0.20
0.35
0.00
0.02
0.04
Clostridia_UCG.014
probability 0
0.22
0.28
0.0000000
0.0018339
Clostridia
probability 0
0.22
0.28
0
Chitinophagales
probability 0
0.20
0.30
0.0000
0.0015
Caulobacterales
probability 0
0.20
0.35
0.000
0.010
Bacillales
probability 0
0.20
0.30
0.00
0.04
0.08
Aeromonadales

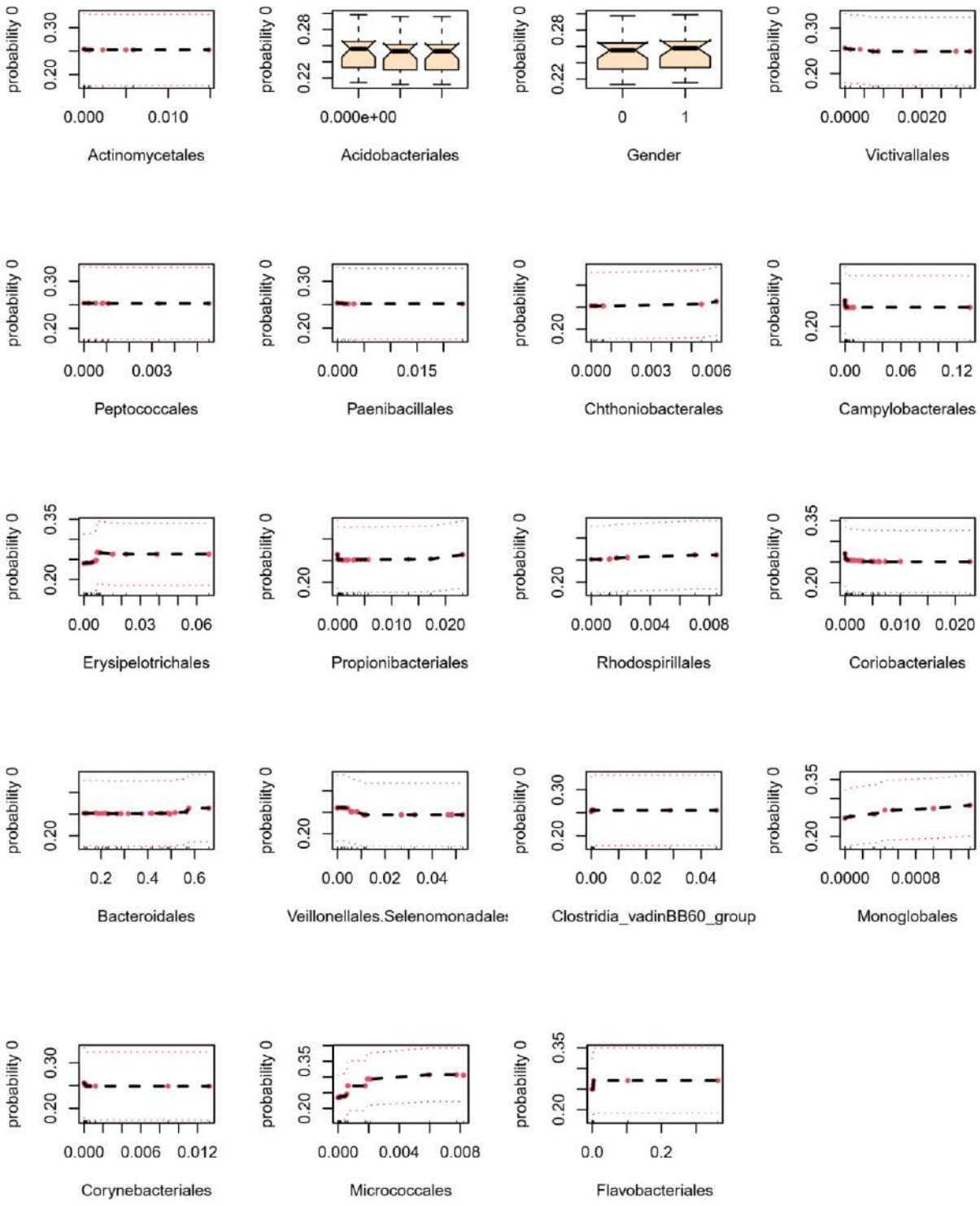


Figure ML1-C-3: Partial Plot for Order in Biopsy

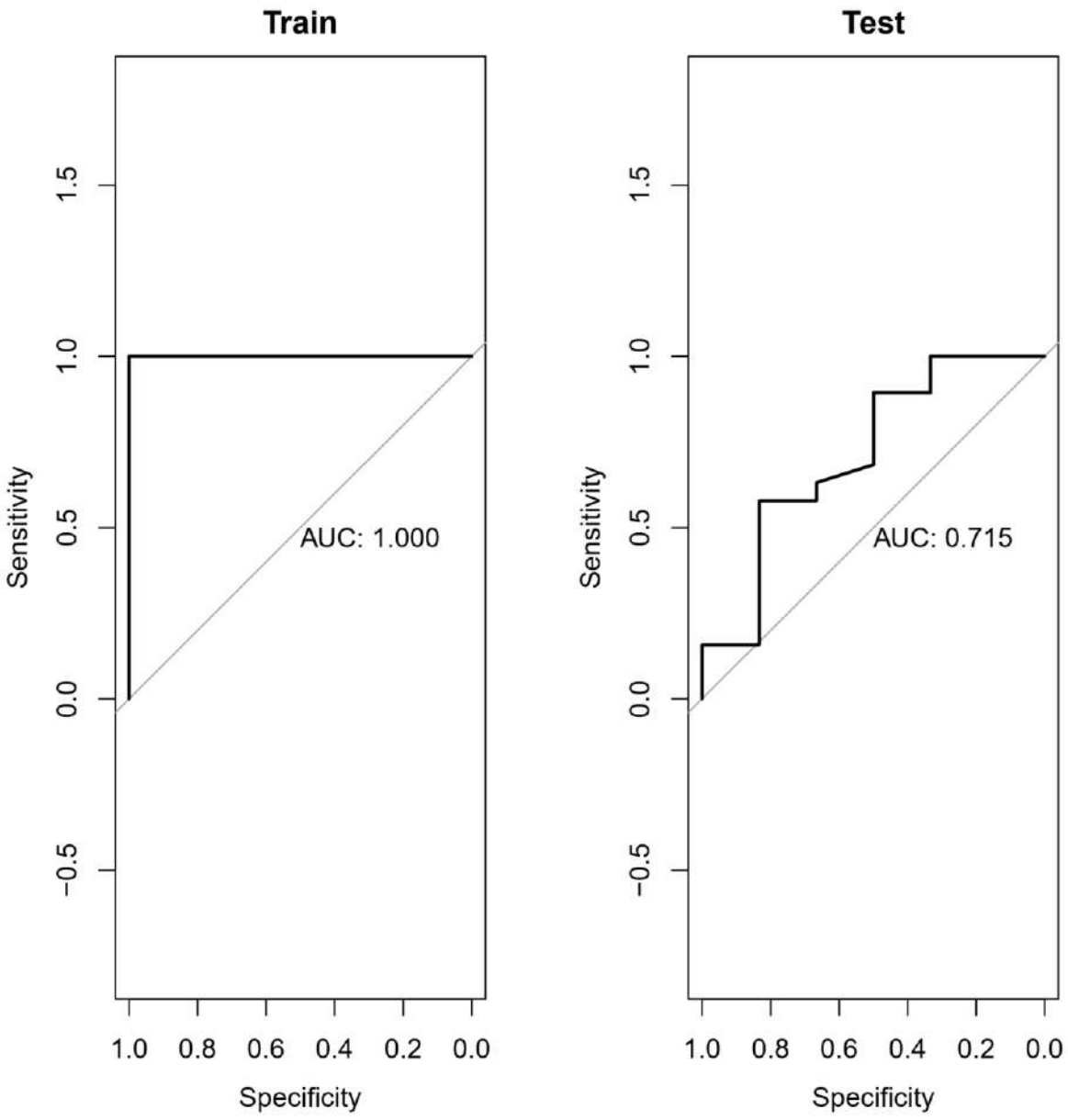


Figure ML1-C-4: The ROC Curve for Order in Biopsy

Table ML1-C-1: The Model Performance

The Model Performance			
	Values		
Index	2.50%	50.00%	97.50%
Accuracy	0.48	0.80	0.96
Specificity	0.33	0.83	1.00
Sensitivity	0.32	0.79	1.00

Table ML1-C-2: The Variable Importance Values

The Variable Importance Values						
Covariates	Variable Importance			Brier Score		
	all	AP	CRC	all	AP	CRC
Gender	-0.0008	-0.0045	-0.0014	-0.0005	-0.0003	-0.0003
Acidaminococcales	-0.0052	-0.0091	-0.0157	-0.0092	-0.0063	-0.0063
Acidobacteriales	-0.0008	0.0136	-0.0072	0.0008	0.0005	0.0005
Actinomycetales	-0.0004	0.0000	-0.0014	-0.0001	-0.0001	-0.0001
Aeromonadales	0.0000	0.0091	-0.0029	0.0035	0.0023	0.0023
Alteromonadales	0.0162	-0.0272	0.0658	0.0475	0.0322	0.0322
Bacillales	-0.0016	-0.0091	-0.0029	-0.0074	-0.0050	-0.0050
Bacteroidales	0.0024	-0.0091	0.0114	-0.0006	-0.0004	-0.0004
Bifidobacteriales	-0.0024	0.0000	-0.0086	0.0001	0.0001	0.0001
Burkholderiales	0.0307	-0.1540	0.1574	0.1493	0.1015	0.1015
Campylobacterales	-0.0044	0.0091	-0.0186	-0.0099	-0.0067	-0.0067
Caulobacterales	-0.0004	0.0000	-0.0014	-0.0006	-0.0004	-0.0004
Chitinophagales	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Chloroplast	0.0013	-0.0453	0.0186	-0.0093	-0.0063	-0.0063
Christensenellales	0.0008	0.0045	0.0014	-0.0057	-0.0039	-0.0039
Chthoniobacterales	0.0024	-0.0136	0.0129	0.0077	0.0052	0.0052
Clostridia	-0.0008	0.0000	-0.0029	-0.0008	-0.0006	-0.0006
Clostridia_UCG.014	0.0028	-0.0544	0.0272	0.0005	0.0003	0.0003
Clostridia_vadinBB60_group	-0.0012	0.0000	-0.0043	-0.0067	-0.0045	-0.0045
Clostridiales	0.0009	-0.0227	0.0100	0.0018	0.0012	0.0012
Coriobacteriales	0.0033	-0.0362	0.0229	0.0092	0.0062	0.0062
Corynebacteriales	-0.0037	-0.0181	-0.0072	-0.0068	-0.0046	-0.0046
Cytophagales	0.0008	0.0091	0.0000	0.0014	0.0009	0.0009
Desulfobacterales	0.0000	0.0045	-0.0014	0.0001	0.0000	0.0000
Desulfobulbales	-0.0004	0.0000	-0.0014	-0.0024	-0.0016	-0.0016
Desulfovibrionales	0.0263	-0.1993	0.1559	0.0939	0.0638	0.0638
DTU014	0.0004	0.0045	0.0000	-0.0022	-0.0015	-0.0015
Enterobacterales	0.0024	-0.0227	0.0157	0.0032	0.0022	0.0022

Erysipelotrichales	-0.0036	0.0045	-0.0143	-0.0069	-0.0047	-0.0047
Eubacteriales	-0.0020	0.0045	-0.0086	-0.0012	-0.0008	-0.0008
Flavobacteriales	-0.0007	-0.0317	0.0072	-0.0046	-0.0031	-0.0031
Frankiales	0.0036	-0.0272	0.0215	0.0030	0.0020	0.0020
Fusobacteriales	0.1081	-0.1948	0.4464	0.3141	0.2135	0.2135
Gastranaerophilales	-0.0012	0.0136	-0.0086	-0.0030	-0.0020	-0.0020
Ignavibacteriales	-0.0004	0.0045	-0.0029	-0.0004	-0.0003	-0.0003
IMCC26256	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Izemoplasmatales	-0.0012	0.0000	-0.0043	-0.0016	-0.0011	-0.0011
KD4.96	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Lachnospirales	-0.0057	-0.0091	-0.0172	-0.0070	-0.0047	-0.0047
Lactobacillales	0.0158	-0.0317	0.0658	0.0325	0.0221	0.0221
Methanobacteriales	0.0004	0.0136	-0.0029	-0.0004	-0.0003	-0.0003
Micrococcales	0.0028	-0.0725	0.0329	0.0183	0.0124	0.0124
Microtrichales	0.0000	0.0045	-0.0014	0.0012	0.0008	0.0008
Monoglobales	0.0020	-0.0997	0.0386	-0.0056	-0.0038	-0.0038
Orbales	0.0000	0.0000	0.0000	0.0003	0.0002	0.0002
Oscillospirales	0.0033	-0.0408	0.0243	0.0053	0.0036	0.0036
Paenibacillales	-0.0012	-0.0045	-0.0029	-0.0011	-0.0007	-0.0007
Pasteurellales	0.0000	-0.0136	0.0043	0.0010	0.0007	0.0007
Peptococcales	-0.0012	-0.0091	-0.0014	0.0010	0.0007	0.0007
Peptostreptococcales.Tissierellales	0.0475	-0.3126	0.2675	0.1771	0.1204	0.1204
Propionibacteriales	-0.0001	-0.0997	0.0315	0.0101	0.0069	0.0069
Pseudomonadales	-0.0012	-0.0272	0.0043	-0.0017	-0.0012	-0.0012
Rhizobiales	0.0016	-0.0227	0.0129	0.0086	0.0059	0.0059
Rhodobacterales	-0.0004	0.0091	-0.0043	-0.0021	-0.0014	-0.0014
Rhodospirillales	0.0008	-0.0181	0.0086	-0.0006	-0.0004	-0.0004
Saccharimonadales	0.0016	0.0091	0.0029	0.0043	0.0029	0.0029
Solirubrobacterales	0.0033	-0.0272	0.0200	0.0004	0.0003	0.0003
Sphingomonadales	-0.0008	0.0000	-0.0029	-0.0020	-0.0014	-0.0014
Spirochaetales	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Staphylococcales	-0.0045	-0.0181	-0.0100	-0.0082	-0.0055	-0.0055
Steroidobacterales	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Synergistales	-0.0004	0.0091	-0.0043	0.0000	0.0000	0.0000
uncultured	-0.0004	0.0000	-0.0014	-0.0016	-0.0011	-0.0011
Veillonellales.Selenomonadales	-0.0057	0.0136	-0.0243	-0.0045	-0.0031	-0.0031
Verrucomicrobiales	0.0169	-0.1042	0.0930	0.0540	0.0367	0.0367
Victivallales	-0.0012	-0.0091	-0.0014	-0.0032	-0.0022	-0.0022
Xanthomonadales	0.0012	-0.0317	0.0143	-0.0053	-0.0036	-0.0036

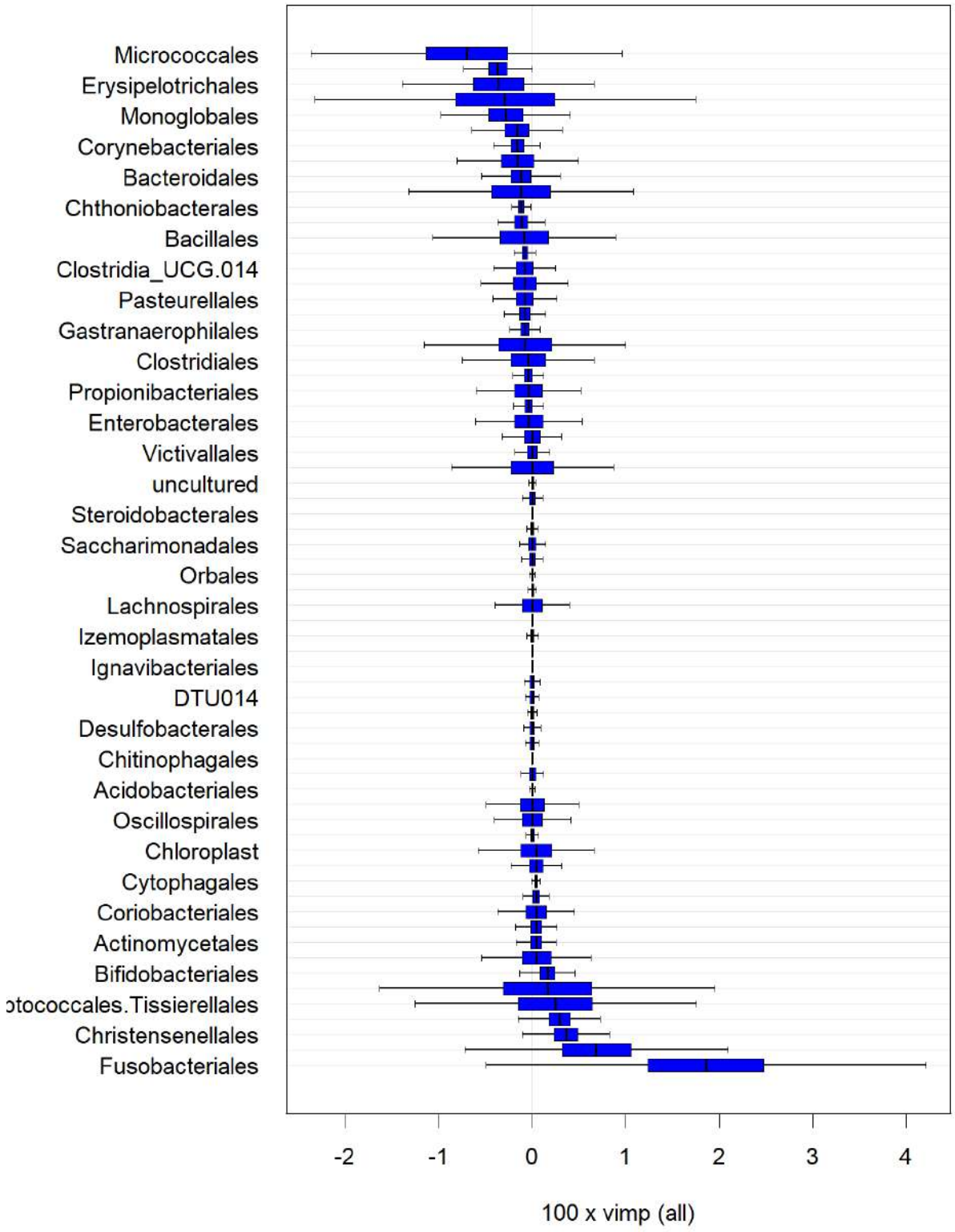


Figure ML1-C-5: The Variable Importance Plot (VIMP) for Order in Biopsy

Best model is selected for LIME results

Table ML1-C-3: The best setting for model and LIME Results

Setting	mtry	ROC	Sens	Spec	ROCSD	SensSD	SpecSD
ntrees: 500 nodesize: 1	1	0.875	0.000	1.000	0.217	0.000	0.000
ntrees: 500 nodesize: 1	5	0.875	0.000	1.000	0.217	0.000	0.000
ntrees: 1000 nodesize: 1	1	0.875	0.000	1.000	0.217	0.000	0.000
ntrees: 1000 nodesize: 1	2	0.875	0.000	1.000	0.217	0.000	0.000
ntrees: 500 nodesize: 5	2	0.900	0.000	1.000	0.224	0.000	0.000
ntrees: 500 nodesize: 5	4	0.900	0.000	1.000	0.224	0.000	0.000
ntrees: 1000 nodesize: 5	2	0.900	0.000	1.000	0.224	0.000	0.000
ntrees: 1000 nodesize: 5	4	0.900	0.000	1.000	0.224	0.000	0.000

Table ML1-C-4: The LIME Result for Training Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
2	Adenoma	0.606	0.303	0.203	0.285	Alteromonadales	0.000	0.031	0.000233 < Alteromonadales
2	Adenoma	0.606	0.303	0.203	0.285	Burkholderiales	0.022	0.029	0.01977 < Burkholderiales
2	Adenoma	0.606	0.303	0.203	0.285	Pseudomonadales	0.012	0.029	0.006297 < Pseudomonadales
2	Adenoma	0.606	0.303	0.203	0.285	Chloroplast	0.000	-0.026	Chloroplast <= 0.000151
2	Adenoma	0.606	0.303	0.203	0.285	Campylobacterales	0.000	0.020	Campylobacterales <= 0.00011
3	Adenoma	0.659	0.339	0.174	0.316	Alteromonadales	0.105	0.031	0.000233 < Alteromonadales
3	Adenoma	0.659	0.339	0.174	0.316	Micrococcales	0.008	0.031	0.000633 < Micrococcales
3	Adenoma	0.659	0.339	0.174	0.316	Burkholderiales	0.036	0.029	0.01977 < Burkholderiales
3	Adenoma	0.659	0.339	0.174	0.316	Pseudomonadales	0.102	0.028	0.006297 < Pseudomonadales
3	Adenoma	0.659	0.339	0.174	0.316	Campylobacterales	0.000	0.022	Campylobacterales <= 0.00011
6	Adenoma	0.530	0.288	0.232	0.250	Alteromonadales	0.000	-0.031	Alteromonadales <= 0.000233
6	Adenoma	0.530	0.288	0.232	0.250	Burkholderiales	0.041	0.030	0.01977 < Burkholderiales
6	Adenoma	0.530	0.288	0.232	0.250	Campylobacterales	0.000	0.020	Campylobacterales <= 0.00011
6	Adenoma	0.530	0.288	0.232	0.250	Peptostreptococcales.Tissierellales	0.000	0.023	Peptostreptococcales.Tissierellales <= 0.00295
6	Adenoma	0.530	0.288	0.232	0.250	Chloroplast	0.000	-0.023	Chloroplast <= 0.000151
7	Adenoma	0.531	0.301	0.233	0.252	Micrococcales	0.006	0.032	0.000633 < Micrococcales
7	Adenoma	0.531	0.301	0.233	0.252	Alteromonadales	0.000	-0.033	Alteromonadales <= 0.000233
7	Adenoma	0.531	0.301	0.233	0.252	Campylobacterales	0.000	0.022	Campylobacterales <= 0.00011
7	Adenoma	0.531	0.301	0.233	0.252	Chloroplast	0.000	-0.025	Chloroplast <= 0.000151
7	Adenoma	0.531	0.301	0.233	0.252	Lactobacillales	0.002	0.023	Lactobacillales <= 0.00204
10	Adenoma	0.620	0.331	0.179	0.321	Micrococcales	0.002	0.032	0.000633 < Micrococcales
10	Adenoma	0.620	0.331	0.179	0.321	Alteromonadales	0.002	0.031	0.000233 < Alteromonadales
10	Adenoma	0.620	0.331	0.179	0.321	Burkholderiales	0.029	0.028	0.01977 < Burkholderiales
10	Adenoma	0.620	0.331	0.179	0.321	Pseudomonadales	0.009	0.027	0.006297 < Pseudomonadales
10	Adenoma	0.620	0.331	0.179	0.321	Peptostreptococcales.Tissierellales	0.001	0.024	Peptostreptococcales.Tissierellales <= 0.00295
11	Adenoma	0.587	0.288	0.176	0.303	Alteromonadales	0.001	0.033	0.000233 < Alteromonadales
11	Adenoma	0.587	0.288	0.176	0.303	Chloroplast	0.002	0.027	0.000151 < Chloroplast

11	Adenoma	0.587	0.288	0.176	0.303	Lactobacillales	0.002	0.025	Lactobacillales <= 0.00204
11	Adenoma	0.587	0.288	0.176	0.303	Campylobacterales	0.000	0.020	Campylobacterales <= 0.00011
11	Adenoma	0.587	0.288	0.176	0.303	Peptostreptococcales.Tissierellales	0.001	0.022	Peptostreptococcales.Tissierellales <= 0.00295
14	Adenocarcinoma	0.877	0.289	0.755	0.779	Alteromonadales	0.000	0.034	Alteromonadales <= 0.000233
14	Adenocarcinoma	0.877	0.289	0.755	0.779	Micrococcales	0.002	-0.031	0.000633 < Micrococcales
14	Adenocarcinoma	0.877	0.289	0.755	0.779	Chloroplast	0.000	0.025	Chloroplast <= 0.000151
14	Adenocarcinoma	0.877	0.289	0.755	0.779	Campylobacterales	0.000	-0.021	Campylobacterales <= 0.00011
14	Adenocarcinoma	0.877	0.289	0.755	0.779	Veillonellales.Selenomonadales	0.048	0.017	0.01189 < Veillonellales.Selenomonadales
20	Adenocarcinoma	0.893	0.255	0.745	0.775	Alteromonadales	0.000	0.032	Alteromonadales <= 0.000233
20	Adenocarcinoma	0.893	0.255	0.745	0.775	Peptostreptococcales.Tissierellales	0.003	-0.026	Peptostreptococcales.Tissierellales <= 0.00295
20	Adenocarcinoma	0.893	0.255	0.745	0.775	Chloroplast	0.000	0.025	Chloroplast <= 0.000151
20	Adenocarcinoma	0.893	0.255	0.745	0.775	Micrococcales	0.000	0.017	Micrococcales <= 0.000125
20	Adenocarcinoma	0.893	0.255	0.745	0.775	Erysipelotrichales	0.022	-0.018	0.00834 < Erysipelotrichales
21	Adenocarcinoma	0.912	0.187	0.745	0.799	Alteromonadales	0.000	0.031	Alteromonadales <= 0.000233
21	Adenocarcinoma	0.912	0.187	0.745	0.799	Chloroplast	0.000	-0.023	0.000151 < Chloroplast
21	Adenocarcinoma	0.912	0.187	0.745	0.799	Campylobacterales	0.008	0.018	0.00180 < Campylobacterales
21	Adenocarcinoma	0.912	0.187	0.745	0.799	Monoglobales	0.000	0.015	Monoglobales <= 6.07e-05
21	Adenocarcinoma	0.912	0.187	0.745	0.799	Acidaminococcales	0.000	0.013	Acidaminococcales <= 0.00272
22	Adenocarcinoma	0.900	0.212	0.724	0.828	Alteromonadales	0.000	0.030	Alteromonadales <= 0.000233
22	Adenocarcinoma	0.900	0.212	0.724	0.828	Chloroplast	0.000	0.025	Chloroplast <= 0.000151
22	Adenocarcinoma	0.900	0.212	0.724	0.828	Micrococcales	0.000	0.017	Micrococcales <= 0.000125
22	Adenocarcinoma	0.900	0.212	0.724	0.828	Acidaminococcales	0.000	0.015	Acidaminococcales <= 0.00272
22	Adenocarcinoma	0.900	0.212	0.724	0.828	Campylobacterales	0.005	0.017	0.00180 < Campylobacterales
23	Adenocarcinoma	0.892	0.247	0.769	0.758	Alteromonadales	0.000	0.030	Alteromonadales <= 0.000233
23	Adenocarcinoma	0.892	0.247	0.769	0.758	Campylobacterales	0.000	-0.021	Campylobacterales <= 0.00011
23	Adenocarcinoma	0.892	0.247	0.769	0.758	Lactobacillales	0.002	-0.023	Lactobacillales <= 0.00204
23	Adenocarcinoma	0.892	0.247	0.769	0.758	Chloroplast	0.000	0.022	Chloroplast <= 0.000151
23	Adenocarcinoma	0.892	0.247	0.769	0.758	Fusobacteriales	0.001	-0.019	Fusobacteriales <= 0.00111
24	Adenocarcinoma	0.859	0.211	0.739	0.807	Alteromonadales	0.000	0.030	Alteromonadales <= 0.000233
24	Adenocarcinoma	0.859	0.211	0.739	0.807	Chloroplast	0.000	0.022	Chloroplast <= 0.000151
24	Adenocarcinoma	0.859	0.211	0.739	0.807	Erysipelotrichales	0.015	-0.019	0.00834 < Erysipelotrichales
24	Adenocarcinoma	0.859	0.211	0.739	0.807	Micrococcales	0.000	0.016	Micrococcales <= 0.000125
24	Adenocarcinoma	0.859	0.211	0.739	0.807	Campylobacterales	0.009	0.019	0.00180 < Campylobacterales
25	Adenocarcinoma	0.888	0.186	0.720	0.818	Alteromonadales	0.000	0.031	Alteromonadales <= 0.000233
25	Adenocarcinoma	0.888	0.186	0.720	0.818	Chloroplast	0.000	0.026	Chloroplast <= 0.000151
25	Adenocarcinoma	0.888	0.186	0.720	0.818	Monoglobales	0.000	0.016	Monoglobales <= 6.07e-05
25	Adenocarcinoma	0.888	0.186	0.720	0.818	Acidaminococcales	0.002	0.012	Acidaminococcales <= 0.00272
25	Adenocarcinoma	0.888	0.186	0.720	0.818	Burkholderiales	0.002	0.012	Burkholderiales <= 0.00448
26	Adenocarcinoma	0.835	0.226	0.791	0.754	Alteromonadales	0.001	-0.028	0.000233 < Alteromonadales
26	Adenocarcinoma	0.835	0.226	0.791	0.754	Lactobacillales	0.002	-0.023	Lactobacillales <= 0.00204
26	Adenocarcinoma	0.835	0.226	0.791	0.754	Chloroplast	0.000	-0.023	0.000151 < Chloroplast
26	Adenocarcinoma	0.835	0.226	0.791	0.754	Micrococcales	0.000	0.018	Micrococcales <= 0.000125
26	Adenocarcinoma	0.835	0.226	0.791	0.754	Campylobacterales	0.003	0.018	0.00180 < Campylobacterales

29	Adenocarcinoma	0.814	0.278	0.766	0.754	Alteromonadales	0.000	0.033	Alteromonadales <= 0.000233
29	Adenocarcinoma	0.814	0.278	0.766	0.754	Pseudomonadales	0.034	-0.025	0.006297 < Pseudomonadales
29	Adenocarcinoma	0.814	0.278	0.766	0.754	Peptostreptococcales.Tissierellales	0.003	-0.023	Peptostreptococcales.Tissierellales <= 0.00295
29	Adenocarcinoma	0.814	0.278	0.766	0.754	Chloroplast	0.000	0.024	Chloroplast <= 0.000151
29	Adenocarcinoma	0.814	0.278	0.766	0.754	Campylobacterales	0.000	-0.020	Campylobacterales <= 0.00011
32	Adenocarcinoma	0.847	0.238	0.736	0.795	Alteromonadales	0.000	0.030	Alteromonadales <= 0.000233
32	Adenocarcinoma	0.847	0.238	0.736	0.795	Pseudomonadales	0.013	-0.028	0.006297 < Pseudomonadales
32	Adenocarcinoma	0.847	0.238	0.736	0.795	Chloroplast	0.000	0.025	Chloroplast <= 0.000151
32	Adenocarcinoma	0.847	0.238	0.736	0.795	Campylobacterales	0.004	0.016	0.00180 < Campylobacterales
32	Adenocarcinoma	0.847	0.238	0.736	0.795	Monoglobales	0.000	0.015	Monoglobales <= 6.07e-05
33	Adenocarcinoma	0.810	0.292	0.768	0.749	Alteromonadales	0.000	0.031	Alteromonadales <= 0.000233
33	Adenocarcinoma	0.810	0.292	0.768	0.749	Burkholderiales	0.031	-0.029	0.01977 < Burkholderiales
33	Adenocarcinoma	0.810	0.292	0.768	0.749	Peptostreptococcales.Tissierellales	0.002	-0.023	Peptostreptococcales.Tissierellales <= 0.00295
33	Adenocarcinoma	0.810	0.292	0.768	0.749	Campylobacterales	0.000	-0.020	Campylobacterales <= 0.00011
33	Adenocarcinoma	0.810	0.292	0.768	0.749	Chloroplast	0.000	0.023	Chloroplast <= 0.000151
34	Adenocarcinoma	0.773	0.328	0.792	0.721	Alteromonadales	0.000	0.032	Alteromonadales <= 0.000233
34	Adenocarcinoma	0.773	0.328	0.792	0.721	Burkholderiales	0.035	-0.028	0.01977 < Burkholderiales
34	Adenocarcinoma	0.773	0.328	0.792	0.721	Micrococcales	0.008	-0.029	0.000633 < Micrococcales
34	Adenocarcinoma	0.773	0.328	0.792	0.721	Pseudomonadales	0.056	-0.026	0.006297 < Pseudomonadales
34	Adenocarcinoma	0.773	0.328	0.792	0.721	Campylobacterales	0.000	-0.020	Campylobacterales <= 0.00011
39	Adenocarcinoma	0.833	0.208	0.794	0.736	Alteromonadales	0.003	-0.031	0.000233 < Alteromonadales
39	Adenocarcinoma	0.833	0.208	0.794	0.736	Campylobacterales	0.000	-0.021	Campylobacterales <= 0.00011
39	Adenocarcinoma	0.833	0.208	0.794	0.736	Chloroplast	0.000	0.022	Chloroplast <= 0.000151
39	Adenocarcinoma	0.833	0.208	0.794	0.736	Acidaminococcales	0.020	-0.014	0.00613 < Acidaminococcales
39	Adenocarcinoma	0.833	0.208	0.794	0.736	Monoglobales	0.001	-0.014	6.07e-05 < Monoglobales
40	Adenocarcinoma	0.879	0.221	0.754	0.771	Alteromonadales	0.000	0.031	Alteromonadales <= 0.000233
40	Adenocarcinoma	0.879	0.221	0.754	0.771	Chloroplast	0.000	-0.024	0.000151 < Chloroplast
40	Adenocarcinoma	0.879	0.221	0.754	0.771	Lactobacillales	0.002	-0.022	Lactobacillales <= 0.00204
40	Adenocarcinoma	0.879	0.221	0.754	0.771	Micrococcales	0.000	0.015	Micrococcales <= 0.000125
40	Adenocarcinoma	0.879	0.221	0.754	0.771	Monoglobales	0.000	0.016	Monoglobales <= 6.07e-05
46	Adenocarcinoma	0.858	0.246	0.745	0.779	Alteromonadales	0.000	0.031	Alteromonadales <= 0.000233
46	Adenocarcinoma	0.858	0.246	0.745	0.779	Chloroplast	0.000	0.025	Chloroplast <= 0.000151
46	Adenocarcinoma	0.858	0.246	0.745	0.779	Peptostreptococcales.Tissierellales	0.003	-0.022	Peptostreptococcales.Tissierellales <= 0.00295
46	Adenocarcinoma	0.858	0.246	0.745	0.779	Micrococcales	0.000	0.018	Micrococcales <= 0.000125
46	Adenocarcinoma	0.858	0.246	0.745	0.779	Fusobacteriales	0.000	-0.019	Fusobacteriales <= 0.00111
47	Adenocarcinoma	0.870	0.232	0.747	0.787	Alteromonadales	0.000	0.033	Alteromonadales <= 0.000233
47	Adenocarcinoma	0.870	0.232	0.747	0.787	Chloroplast	0.000	0.025	Chloroplast <= 0.000151
47	Adenocarcinoma	0.870	0.232	0.747	0.787	Campylobacterales	0.000	-0.020	Campylobacterales <= 0.00011
47	Adenocarcinoma	0.870	0.232	0.747	0.787	Micrococcales	0.000	0.018	Micrococcales <= 0.000125
47	Adenocarcinoma	0.870	0.232	0.747	0.787	Verrucomicrobiales	0.009	-0.015	3.00e-03 < Verrucomicrobiales
48	Adenocarcinoma	0.851	0.226	0.763	0.763	Alteromonadales	0.000	0.031	Alteromonadales <= 0.000233
48	Adenocarcinoma	0.851	0.226	0.763	0.763	Chloroplast	0.000	0.024	Chloroplast <= 0.000151
48	Adenocarcinoma	0.851	0.226	0.763	0.763	Campylobacterales	0.000	-0.020	Campylobacterales <= 0.00011

48	Adenocarcinoma	0.851	0.226	0.763	0.763	Fusobacteriales	0.001	-0.017	Fusobacteriales <= 0.00111
48	Adenocarcinoma	0.851	0.226	0.763	0.763	Acidaminococcales	0.010	-0.018	0.00613 < Acidaminococcales
49	Adenocarcinoma	0.911	0.217	0.736	0.805	Alteromonadales	0.000	0.030	Alteromonadales <= 0.000233
49	Adenocarcinoma	0.911	0.217	0.736	0.805	Chloroplast	0.000	0.025	Chloroplast <= 0.000151
49	Adenocarcinoma	0.911	0.217	0.736	0.805	Erysipelotrichales	0.067	-0.019	0.00834 < Erysipelotrichales
49	Adenocarcinoma	0.911	0.217	0.736	0.805	Micrococcales	0.000	0.015	Micrococcales <= 0.000125
49	Adenocarcinoma	0.911	0.217	0.736	0.805	Campylobacterales	0.134	0.017	0.00180 < Campylobacterales
50	Adenocarcinoma	0.867	0.270	0.761	0.750	Alteromonadales	0.000	0.031	Alteromonadales <= 0.000233
50	Adenocarcinoma	0.867	0.270	0.761	0.750	Micrococcales	0.002	-0.030	0.000633 < Micrococcales
50	Adenocarcinoma	0.867	0.270	0.761	0.750	Chloroplast	0.000	0.025	Chloroplast <= 0.000151
50	Adenocarcinoma	0.867	0.270	0.761	0.750	Erysipelotrichales	0.039	-0.021	0.00834 < Erysipelotrichales
50	Adenocarcinoma	0.867	0.270	0.761	0.750	Monoglobales	0.000	-0.015	6.07e-05 < Monoglobales

Table ML-C-5: The LIME Result for Test Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
1	Adenocarcinoma	0.772	0.255	0.747	0.783	Alteromonadales	0.000	0.033	Alteromonadales <= 0.000233
1	Adenocarcinoma	0.772	0.255	0.747	0.783	Chloroplast	0.000	0.027	Chloroplast <= 0.000151
1	Adenocarcinoma	0.772	0.255	0.747	0.783	Campylobacterales	0.000	-0.021	Campylobacterales <= 0.00011
1	Adenocarcinoma	0.772	0.255	0.747	0.783	Erysipelotrichales	0.018	-0.019	0.00834 < Erysipelotrichales
1	Adenocarcinoma	0.772	0.255	0.747	0.783	Micrococcales	0.000	0.016	Micrococcales <= 0.000125
4	Adenocarcinoma	0.580	0.345	0.793	0.718	Alteromonadales	0.000	0.032	Alteromonadales <= 0.000233
4	Adenocarcinoma	0.580	0.345	0.793	0.718	Micrococcales	0.002	-0.031	0.000633 < Micrococcales
4	Adenocarcinoma	0.580	0.345	0.793	0.718	Burkholderiales	0.028	-0.027	0.01977 < Burkholderiales
4	Adenocarcinoma	0.580	0.345	0.793	0.718	Pseudomonadales	0.011	-0.027	0.006297 < Pseudomonadales
4	Adenocarcinoma	0.580	0.345	0.793	0.718	Campylobacterales	0.000	-0.022	Campylobacterales <= 0.00011
5	Adenocarcinoma	0.716	0.277	0.768	0.749	Alteromonadales	0.000	0.028	Alteromonadales <= 0.000233
5	Adenocarcinoma	0.716	0.277	0.768	0.749	Pseudomonadales	0.009	-0.027	0.006297 < Pseudomonadales
5	Adenocarcinoma	0.716	0.277	0.768	0.749	Burkholderiales	0.020	-0.027	0.01977 < Burkholderiales
5	Adenocarcinoma	0.716	0.277	0.768	0.749	Chloroplast	0.000	0.025	Chloroplast <= 0.000151
5	Adenocarcinoma	0.716	0.277	0.768	0.749	Campylobacterales	0.000	-0.019	Campylobacterales <= 0.00011
8	Adenocarcinoma	0.769	0.273	0.763	0.755	Alteromonadales	0.000	0.033	Alteromonadales <= 0.000233
8	Adenocarcinoma	0.769	0.273	0.763	0.755	Burkholderiales	0.034	-0.028	0.01977 < Burkholderiales
8	Adenocarcinoma	0.769	0.273	0.763	0.755	Chloroplast	0.000	0.025	Chloroplast <= 0.000151
8	Adenocarcinoma	0.769	0.273	0.763	0.755	Campylobacterales	0.000	-0.019	Campylobacterales <= 0.00011
8	Adenocarcinoma	0.769	0.273	0.763	0.755	Erysipelotrichales	0.017	-0.019	0.00834 < Erysipelotrichales
9	Adenocarcinoma	0.561	0.296	0.797	0.719	Alteromonadales	0.002	-0.033	0.000233 < Alteromonadales
9	Adenocarcinoma	0.561	0.296	0.797	0.719	Chloroplast	0.000	0.026	Chloroplast <= 0.000151
9	Adenocarcinoma	0.561	0.296	0.797	0.719	Pseudomonadales	0.023	-0.027	0.006297 < Pseudomonadales
9	Adenocarcinoma	0.561	0.296	0.797	0.719	Peptostreptococcales.Tissierellales	0.000	-0.024	Peptostreptococcales.Tissierellales <= 0.00295
9	Adenocarcinoma	0.561	0.296	0.797	0.719	Campylobacterales	0.000	-0.020	Campylobacterales <= 0.00011
12	Adenocarcinoma	0.706	0.253	0.803	0.717	Alteromonadales	0.001	-0.030	0.000233 < Alteromonadales
12	Adenocarcinoma	0.706	0.253	0.803	0.717	Pseudomonadales	0.033	-0.026	0.006297 < Pseudomonadales

12	Adenocarcinoma	0.706	0.253	0.803	0.717	Peptostreptococcales.Tissierellales	0.003	-0.023	Peptostreptococcales.Tissierellales <= 0.00295
12	Adenocarcinoma	0.706	0.253	0.803	0.717	Chloroplast	0.002	-0.023	0.000151 < Chloroplast
12	Adenocarcinoma	0.706	0.253	0.803	0.717	Acidaminococcales	0.000	0.015	Acidaminococcales <= 0.00272
13	Adenocarcinoma	0.810	0.201	0.740	0.809	Alteromonadales	0.000	0.034	Alteromonadales <= 0.000233
13	Adenocarcinoma	0.810	0.201	0.740	0.809	Chloroplast	0.000	0.022	Chloroplast <= 0.000151
13	Adenocarcinoma	0.810	0.201	0.740	0.809	Campylobacterales	0.005	0.018	0.00180 < Campylobacterales
13	Adenocarcinoma	0.810	0.201	0.740	0.809	Erysipelotrichales	0.013	-0.018	0.00834 < Erysipelotrichales
13	Adenocarcinoma	0.810	0.201	0.740	0.809	Burkholderiales	0.001	0.014	Burkholderiales <= 0.00448
15	Adenocarcinoma	0.799	0.284	0.764	0.754	Alteromonadales	0.000	0.033	Alteromonadales <= 0.000233
15	Adenocarcinoma	0.799	0.284	0.764	0.754	Burkholderiales	0.031	-0.030	0.01977 < Burkholderiales
15	Adenocarcinoma	0.799	0.284	0.764	0.754	Chloroplast	0.000	0.026	Chloroplast <= 0.000151
15	Adenocarcinoma	0.799	0.284	0.764	0.754	Campylobacterales	0.000	-0.020	Campylobacterales <= 0.00011
15	Adenocarcinoma	0.799	0.284	0.764	0.754	Erysipelotrichales	0.032	-0.019	0.00834 < Erysipelotrichales
16	Adenocarcinoma	0.836	0.262	0.775	0.755	Alteromonadales	0.000	0.032	Alteromonadales <= 0.000233
16	Adenocarcinoma	0.836	0.262	0.775	0.755	Micrococcales	0.001	-0.029	0.000633 < Micrococcales
16	Adenocarcinoma	0.836	0.262	0.775	0.755	Chloroplast	0.000	-0.023	0.000151 < Chloroplast
16	Adenocarcinoma	0.836	0.262	0.775	0.755	Campylobacterales	0.003	0.020	0.00180 < Campylobacterales
16	Adenocarcinoma	0.836	0.262	0.775	0.755	Erysipelotrichales	0.011	-0.019	0.00834 < Erysipelotrichales
17	Adenocarcinoma	0.768	0.296	0.770	0.750	Alteromonadales	0.000	0.032	Alteromonadales <= 0.000233
17	Adenocarcinoma	0.768	0.296	0.770	0.750	Micrococcales	0.001	-0.030	0.000633 < Micrococcales
17	Adenocarcinoma	0.768	0.296	0.770	0.750	Peptostreptococcales.Tissierellales	0.003	-0.025	Peptostreptococcales.Tissierellales <= 0.00295
17	Adenocarcinoma	0.768	0.296	0.770	0.750	Campylobacterales	0.000	-0.021	Campylobacterales <= 0.00011
17	Adenocarcinoma	0.768	0.296	0.770	0.750	Chloroplast	0.000	0.023	Chloroplast <= 0.000151
18	Adenocarcinoma	0.780	0.296	0.769	0.751	Alteromonadales	0.000	0.031	Alteromonadales <= 0.000233
18	Adenocarcinoma	0.780	0.296	0.769	0.751	Burkholderiales	0.034	-0.028	0.01977 < Burkholderiales
18	Adenocarcinoma	0.780	0.296	0.769	0.751	Campylobacterales	0.000	-0.023	Campylobacterales <= 0.00011
18	Adenocarcinoma	0.780	0.296	0.769	0.751	Peptostreptococcales.Tissierellales	0.002	-0.024	Peptostreptococcales.Tissierellales <= 0.00295
18	Adenocarcinoma	0.780	0.296	0.769	0.751	Chloroplast	0.000	0.024	Chloroplast <= 0.000151
19	Adenocarcinoma	0.726	0.212	0.804	0.731	Alteromonadales	0.001	-0.029	0.000233 < Alteromonadales
19	Adenocarcinoma	0.726	0.212	0.804	0.731	Chloroplast	0.000	-0.023	0.000151 < Chloroplast
19	Adenocarcinoma	0.726	0.212	0.804	0.731	Campylobacterales	0.000	-0.019	Campylobacterales <= 0.00011
19	Adenocarcinoma	0.726	0.212	0.804	0.731	Fusobacteriales	0.001	-0.017	Fusobacteriales <= 0.00111
19	Adenocarcinoma	0.726	0.212	0.804	0.731	Micrococcales	0.000	0.016	Micrococcales <= 0.000125
27	Adenocarcinoma	0.884	0.221	0.738	0.805	Alteromonadales	0.000	0.032	Alteromonadales <= 0.000233
27	Adenocarcinoma	0.884	0.221	0.738	0.805	Campylobacterales	0.000	-0.022	Campylobacterales <= 0.00011
27	Adenocarcinoma	0.884	0.221	0.738	0.805	Chloroplast	0.000	0.025	Chloroplast <= 0.000151
27	Adenocarcinoma	0.884	0.221	0.738	0.805	Monoglobales	0.000	0.016	Monoglobales <= 6.07e-05
27	Adenocarcinoma	0.884	0.221	0.738	0.805	Veillonellales.Selenomonadales	0.038	0.016	0.01189 < Veillonellales.Selenomonadales
28	Adenocarcinoma	0.818	0.262	0.765	0.758	Alteromonadales	0.000	0.030	Alteromonadales <= 0.000233
28	Adenocarcinoma	0.818	0.262	0.765	0.758	Micrococcales	0.001	-0.028	0.000633 < Micrococcales
28	Adenocarcinoma	0.818	0.262	0.765	0.758	Chloroplast	0.000	0.026	Chloroplast <= 0.000151
28	Adenocarcinoma	0.818	0.262	0.765	0.758	Campylobacterales	0.000	-0.021	Campylobacterales <= 0.00011
28	Adenocarcinoma	0.818	0.262	0.765	0.758	Acidaminococcales	0.007	-0.013	0.00613 < Acidaminococcales

30	Adenocarcinoma	0.787	0.290	0.793	0.713	Alteromonadales	0.002	-0.033	0.000233 < Alteromonadales
30	Adenocarcinoma	0.787	0.290	0.793	0.713	Burkholderiales	0.021	-0.029	0.01977 < Burkholderiales
30	Adenocarcinoma	0.787	0.290	0.793	0.713	Pseudomonadales	0.020	-0.025	0.006297 < Pseudomonadales
30	Adenocarcinoma	0.787	0.290	0.793	0.713	Chloroplast	0.000	0.024	Chloroplast <= 0.000151
30	Adenocarcinoma	0.787	0.290	0.793	0.713	Verrucomicrobiales	0.017	-0.017	3.00e-03 < Verrucomicrobiales
31	Adenocarcinoma	0.746	0.217	0.746	0.779	Alteromonadales	0.000	0.031	Alteromonadales <= 0.000233
31	Adenocarcinoma	0.746	0.217	0.746	0.779	Chloroplast	0.000	0.024	Chloroplast <= 0.000151
31	Adenocarcinoma	0.746	0.217	0.746	0.779	Lactobacillales	0.002	-0.021	Lactobacillales <= 0.00204
31	Adenocarcinoma	0.746	0.217	0.746	0.779	Micrococcales	0.000	0.015	Micrococcales <= 0.000125
31	Adenocarcinoma	0.746	0.217	0.746	0.779	Erysipelotrichales	0.062	-0.018	0.00834 < Erysipelotrichales
35	Adenocarcinoma	0.696	0.320	0.794	0.719	Micrococcales	0.001	-0.032	0.000633 < Micrococcales
35	Adenocarcinoma	0.696	0.320	0.794	0.719	Alteromonadales	0.000	0.030	Alteromonadales <= 0.000233
35	Adenocarcinoma	0.696	0.320	0.794	0.719	Burkholderiales	0.029	-0.027	0.01977 < Burkholderiales
35	Adenocarcinoma	0.696	0.320	0.794	0.719	Chloroplast	0.001	-0.026	0.000151 < Chloroplast
35	Adenocarcinoma	0.696	0.320	0.794	0.719	Campylobacterales	0.000	-0.020	Campylobacterales <= 0.00011
36	Adenocarcinoma	0.728	0.277	0.765	0.748	Micrococcales	0.001	-0.031	0.000633 < Micrococcales
36	Adenocarcinoma	0.728	0.277	0.765	0.748	Alteromonadales	0.000	0.030	Alteromonadales <= 0.000233
36	Adenocarcinoma	0.728	0.277	0.765	0.748	Chloroplast	0.000	0.024	Chloroplast <= 0.000151
36	Adenocarcinoma	0.728	0.277	0.765	0.748	Lactobacillales	0.001	-0.022	Lactobacillales <= 0.00204
36	Adenocarcinoma	0.728	0.277	0.765	0.748	Erysipelotrichales	0.021	-0.017	0.00834 < Erysipelotrichales
37	Adenocarcinoma	0.752	0.268	0.746	0.773	Alteromonadales	0.000	0.032	Alteromonadales <= 0.000233
37	Adenocarcinoma	0.752	0.268	0.746	0.773	Burkholderiales	0.025	-0.027	0.01977 < Burkholderiales
37	Adenocarcinoma	0.752	0.268	0.746	0.773	Chloroplast	0.000	0.023	Chloroplast <= 0.000151
37	Adenocarcinoma	0.752	0.268	0.746	0.773	Micrococcales	0.000	0.018	Micrococcales <= 0.000125
37	Adenocarcinoma	0.752	0.268	0.746	0.773	Lactobacillales	0.001	-0.020	Lactobacillales <= 0.00204
38	Adenocarcinoma	0.697	0.298	0.817	0.685	Alteromonadales	0.003	-0.034	0.000233 < Alteromonadales
38	Adenocarcinoma	0.697	0.298	0.817	0.685	Micrococcales	0.006	-0.033	0.000633 < Micrococcales
38	Adenocarcinoma	0.697	0.298	0.817	0.685	Pseudomonadales	0.011	-0.026	0.006297 < Pseudomonadales
38	Adenocarcinoma	0.697	0.298	0.817	0.685	Chloroplast	0.001	-0.023	0.000151 < Chloroplast
38	Adenocarcinoma	0.697	0.298	0.817	0.685	Verrucomicrobiales	0.026	-0.017	3.00e-03 < Verrucomicrobiales
41	Adenocarcinoma	0.789	0.196	0.747	0.786	Alteromonadales	0.000	0.032	Alteromonadales <= 0.000233
41	Adenocarcinoma	0.789	0.196	0.747	0.786	Chloroplast	0.000	0.025	Chloroplast <= 0.000151
41	Adenocarcinoma	0.789	0.196	0.747	0.786	Erysipelotrichales	0.034	-0.018	0.00834 < Erysipelotrichales
41	Adenocarcinoma	0.789	0.196	0.747	0.786	Campylobacterales	0.010	0.016	0.00180 < Campylobacterales
41	Adenocarcinoma	0.789	0.196	0.747	0.786	Acidaminococcales	0.011	-0.015	0.00613 < Acidaminococcales
42	Adenocarcinoma	0.637	0.278	0.765	0.757	Alteromonadales	0.000	0.032	Alteromonadales <= 0.000233
42	Adenocarcinoma	0.637	0.278	0.765	0.757	Chloroplast	0.000	0.026	Chloroplast <= 0.000151
42	Adenocarcinoma	0.637	0.278	0.765	0.757	Campylobacterales	0.000	-0.021	Campylobacterales <= 0.00011
42	Adenocarcinoma	0.637	0.278	0.765	0.757	Lactobacillales	0.001	-0.023	Lactobacillales <= 0.00204
42	Adenocarcinoma	0.637	0.278	0.765	0.757	Peptostreptococcales.Tissierellales	0.001	-0.023	Peptostreptococcales.Tissierellales <= 0.00295
43	Adenocarcinoma	0.702	0.266	0.766	0.759	Alteromonadales	0.000	0.032	Alteromonadales <= 0.000233
43	Adenocarcinoma	0.702	0.266	0.766	0.759	Peptostreptococcales.Tissierellales	0.001	-0.024	Peptostreptococcales.Tissierellales <= 0.00295
43	Adenocarcinoma	0.702	0.266	0.766	0.759	Chloroplast	0.000	0.024	Chloroplast <= 0.000151

43	Adenocarcinoma	0.702	0.266	0.766	0.759	Campylobacterales	0.000	-0.020	Campylobacterales <= 0.00011
43	Adenocarcinoma	0.702	0.266	0.766	0.759	Lactobacillales	0.001	-0.019	Lactobacillales <= 0.00204
44	Adenocarcinoma	0.807	0.212	0.744	0.781	Alteromonadales	0.000	0.028	Alteromonadales <= 0.000233
44	Adenocarcinoma	0.807	0.212	0.744	0.781	Chloroplast	0.000	0.027	Chloroplast <= 0.000151
44	Adenocarcinoma	0.807	0.212	0.744	0.781	Micrococcales	0.000	0.017	Micrococcales <= 0.000125
44	Adenocarcinoma	0.807	0.212	0.744	0.781	Verrucomicrobiales	0.074	-0.018	3.00e-03 < Verrucomicrobiales
44	Adenocarcinoma	0.807	0.212	0.744	0.781	Erysipelotrichales	0.009	-0.017	0.00834 < Erysipelotrichales
45	Adenocarcinoma	0.759	0.240	0.763	0.763	Alteromonadales	0.000	0.031	Alteromonadales <= 0.000233
45	Adenocarcinoma	0.759	0.240	0.763	0.763	Chloroplast	0.000	0.027	Chloroplast <= 0.000151
45	Adenocarcinoma	0.759	0.240	0.763	0.763	Campylobacterales	0.000	-0.020	Campylobacterales <= 0.00011
45	Adenocarcinoma	0.759	0.240	0.763	0.763	Lactobacillales	0.001	-0.022	Lactobacillales <= 0.00204
45	Adenocarcinoma	0.759	0.240	0.763	0.763	Verrucomicrobiales	0.052	-0.016	3.00e-03 < Verrucomicrobiales



Figure ML1-C-6: The LIME Result in Training dataset for Order in Biopsy

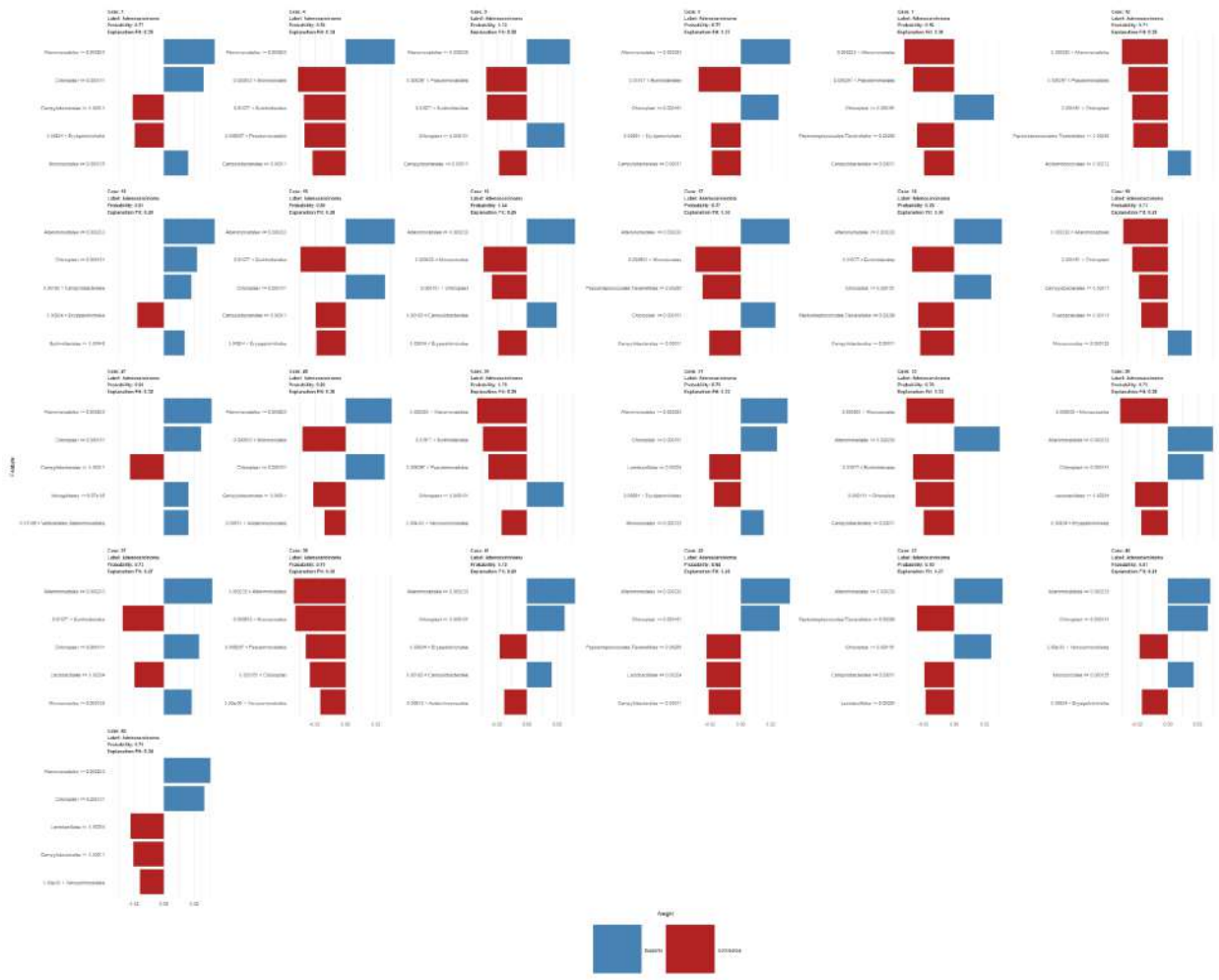


Figure ML1-C-7: The LIME Result in Test dataset for Order in Biopsy

Table ML1-C-6: The Aggregated SHAP Values

Variable	Training Dataset				Testing Dataset			
	contribution	position	sign	cumulative	contribution	position	sign	cumulative
intercept	0.7534	70	X	0.753	0.75344	70	X	0.753
Bacteroidales	0.0000	69	1	0.753	-1.9E-05	68	-1	0.753
Rhizobiales	0.0000	68	1	0.753	0.001291	29	1	0.748
Flavobacteriales	0.0000	67	-1	0.753	-6.1E-05	67	-1	0.753
Chitinophagales	0.0000	66	1	0.753	0	66	1	0.753

Microtrichales	0.0000	65	1	0.753	3.07E-06	69	1	0.753
Gender	0.0000	64	1	0.753	0	64	1	0.755
Veillonellales.Selenomonadales	-0.0001	63	-1	0.753	0.001947	65	1	0.755
Fusobacteriales	0.0000	62	-1	0.753	-0.00365	49	-1	0.760
Methanobacteriales	0.0000	61	1	0.753	0.001824	62	1	0.756
Corynebacteriales	0.0000	60	-1	0.753	0.000582	60	1	0.758
Propionibacteriales	0.0000	59	-1	0.753	0.000528	59	1	0.758
Desulfobacterales	0.0000	58	-1	0.753	-0.00011	43	-1	0.752
Alteromonadales	0.0000	57	-1	0.753	0.001459	57	1	0.759
Coriobacteriales	0.0000	56	1	0.753	0.002282	56	1	0.762
Clostridia	0.0000	55	1	0.753	-0.00021	58	-1	0.758
Clostridiales	0.0000	54	1	0.753	0.001564	52	1	0.763
Chloroplast	0.0000	53	1	0.753	0.000538	54	1	0.760
Verrucomicrobiales	0.0000	52	-1	0.753	0.001842	53	1	0.762
Actinomycetales	0.0000	51	-1	0.753	0.00132	61	1	0.757
Frankiales	0.0000	50	1	0.753	0.000438	50	1	0.764
Gastranaerophilales	0.0000	49	1	0.753	-0.00045	48	-1	0.760
Burkholderiales	0.0001	48	1	0.753	-0.00454	47	-1	0.755
Micrococcales	-0.0001	47	-1	0.753	-0.00261	55	-1	0.759
Staphylococcales	0.0000	46	-1	0.753	-0.00166	45	-1	0.753
Desulfovibrionales	0.0000	45	1	0.753	-0.00122	44	-1	0.752
Campylobacterales	0.0000	44	-1	0.753	-0.00238	35	-1	0.750
Tumor	0.0000	43	1	0.753	0	28	1	0.748
Chthoniobacterales	0.0000	42	-1	0.753	0.001176	42	1	0.753
Rhodobacterales	0.0000	41	1	0.753	2.83E-05	40	1	0.753
Erysipelotrichales	0.0000	40	-1	0.753	-0.00338	32	-1	0.747
Peptococcales	0.0000	39	1	0.753	-0.00011	39	-1	0.753
Clostridia_UCG.014	0.0000	38	-1	0.753	-0.00021	37	-1	0.752
Pasteurellales	0.0000	37	1	0.753	-0.00021	25	-1	0.745
Solirubrobacterales	0.0000	36	-1	0.753	0.000555	36	1	0.752
Caulobacterales	0.0000	35	1	0.753	0.000224	51	1	0.763
Xanthomonadales	0.0000	34	-1	0.753	-0.00132	63	-1	0.754
Enterobacterales	0.0000	33	-1	0.753	0.000781	33	1	0.751
Sphingomonadales	0.0000	32	-1	0.753	-0.00035	31	-1	0.747
Cytophagales	0.0000	31	-1	0.753	-0.00037	20	-1	0.746
Rhodospirillales	0.0000	30	1	0.753	-0.0005	27	-1	0.748
Lactobacillales	0.0000	29	1	0.753	-0.00263	26	-1	0.745
Victivallales	0.0000	28	-1	0.753	-0.00012	21	-1	0.746
Eubacteriales	0.0000	27	1	0.753	0.000175	34	1	0.750
Steroidobacterales	0.0000	26	1	0.753	-8.8E-06	30	-1	0.747
DTU014	0.0000	25	1	0.753	0.000211	12	1	0.747
Peptostreptococcales.Tissierellales	0.0001	24	1	0.753	-0.00126	38	-1	0.752
Saccharimonadales	0.0000	23	-1	0.753	-0.00022	23	-1	0.744

Aeromonadales	0.0000	22	1	0.753	-0.00012	18	-1	0.746
Clostridia_vadinBB60_group	0.0000	21	1	0.753	-0.0002	17	-1	0.746
Ignavibacteriales	0.0000	20	-1	0.753	1.56E-05	41	1	0.753
Orbales	0.0000	19	-1	0.753	-2.1E-05	46	-1	0.755
Monoglobales	0.0000	18	-1	0.753	0.001733	16	1	0.747
Desulfobulbales	0.0000	17	1	0.753	-6E-05	15	-1	0.747
Spirochaetales	0.0000	16	-1	0.753	0.000162	19	1	0.746
KD4.96	0.0000	15	1	0.753	0	14	1	0.747
uncultured	0.0000	14	1	0.753	-0.00018	13	-1	0.747
IMCC26256	0.0000	13	1	0.753	-3.2E-05	11	-1	0.747
Izemoplasmatales	0.0000	12	-1	0.753	-0.00019	10	-1	0.747
Oscillospirales	0.0000	11	-1	0.753	-0.00129	9	-1	0.746
Pseudomonadales	0.0000	10	1	0.753	0.001851	22	1	0.746
Bifidobacteriales	0.0001	9	1	0.753	0.000876	8	1	0.747
Paenibacillales	0.0000	8	-1	0.753	-0.00038	24	-1	0.745
Synergistales	0.0000	7	1	0.753	-0.00012	7	-1	0.747
Acidaminococcales	0.0000	6	1	0.753	-0.00433	6	-1	0.742
Acidobacteriales	0.0000	5	1	0.753	8.25E-05	5	1	0.742
Bacillales	0.0000	4	-1	0.753	0.001412	4	1	0.744
Christensenellales	0.0000	3	-1	0.753	0.001323	3	1	0.745
Lachnospirales	0.0000	2	1	0.753	-0.00013	2	-1	0.745
prediction	0.7534	1	X	0.753	0.745	1	X	0.745

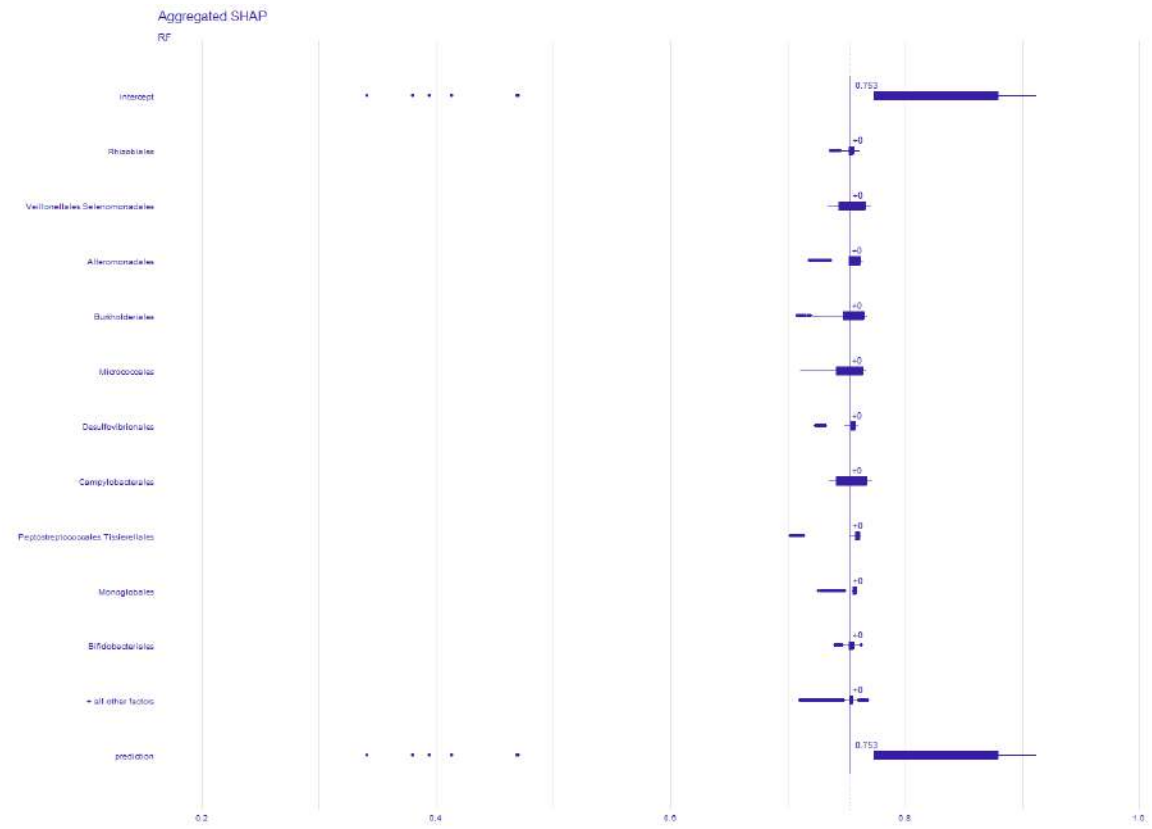


Figure ML1-C-8: The Aggregated SHAP Value in Training dataset for Order in Biopsy

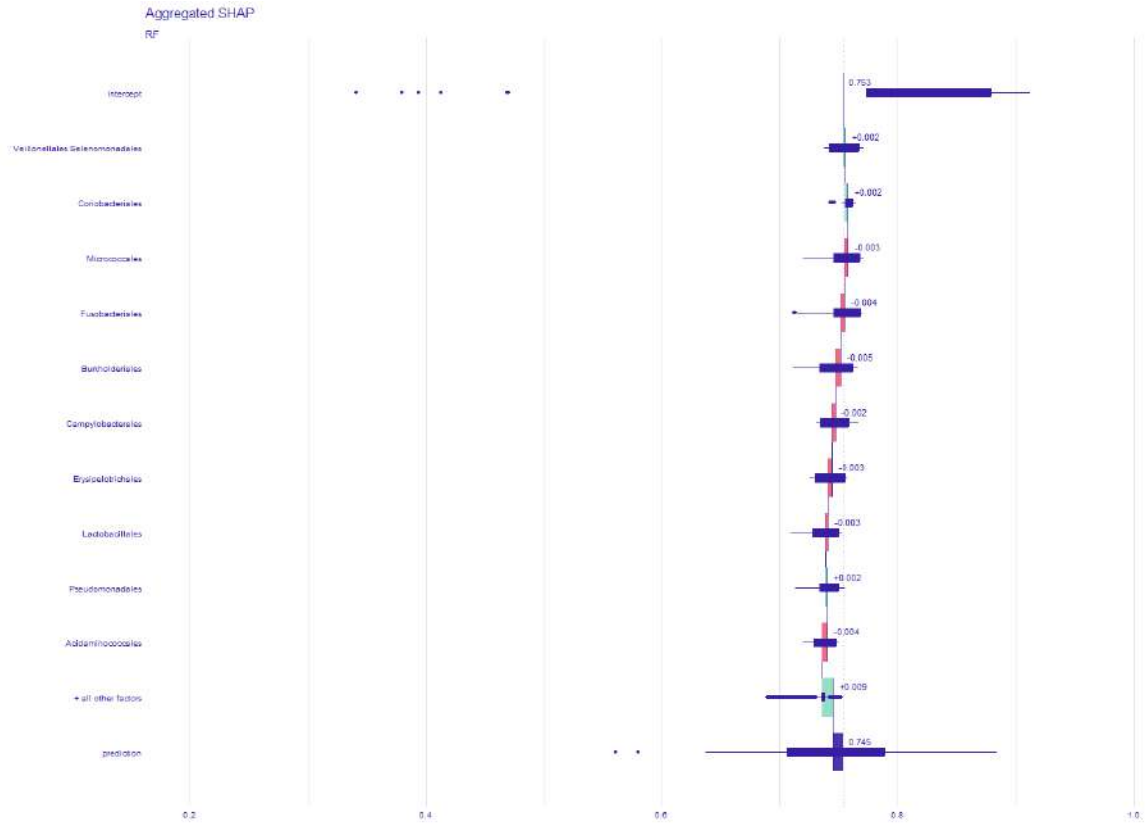


Figure ML1-C-9: The Aggregated SHAP Value in Test dataset for Order in Biopsy

D- Family

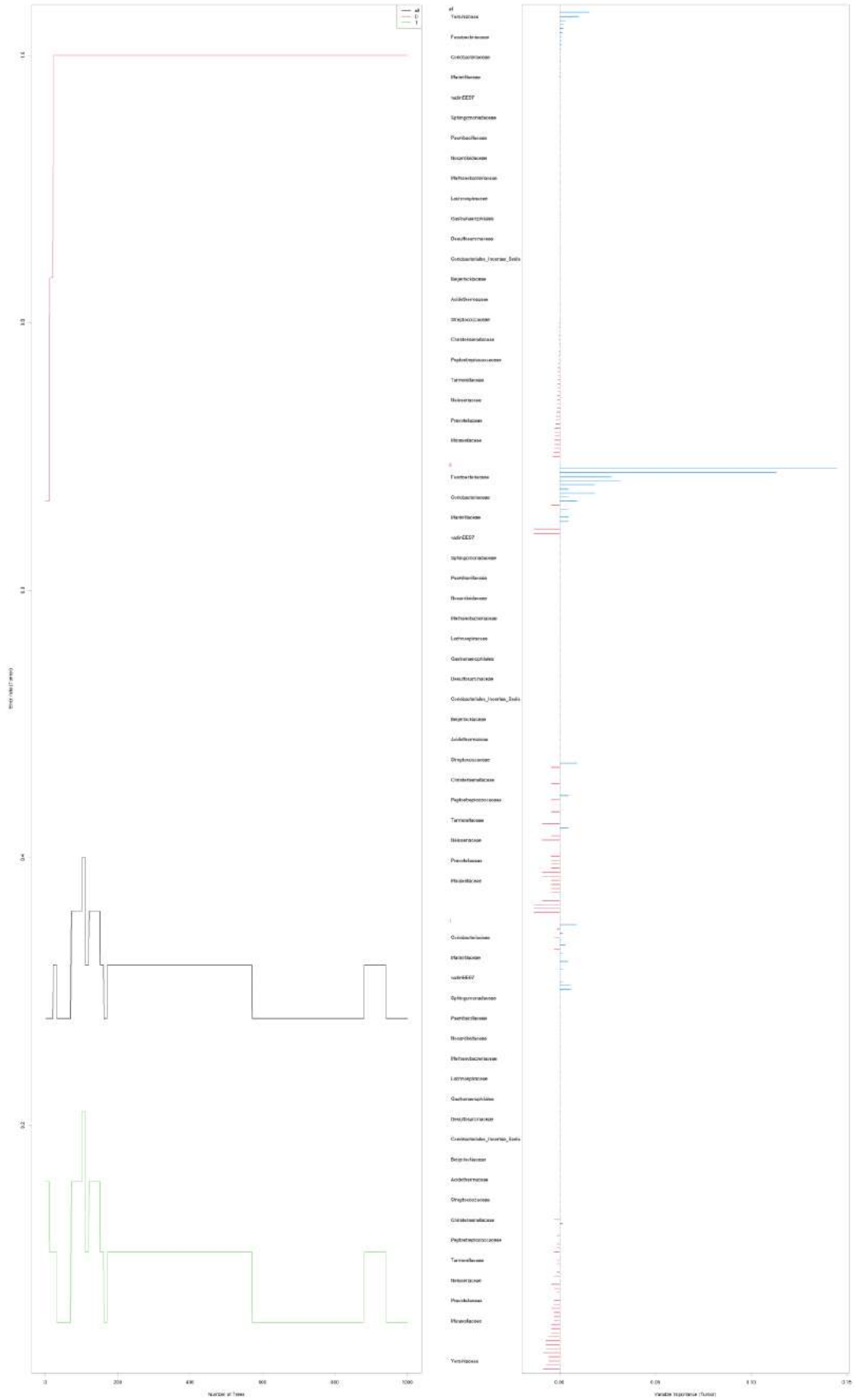
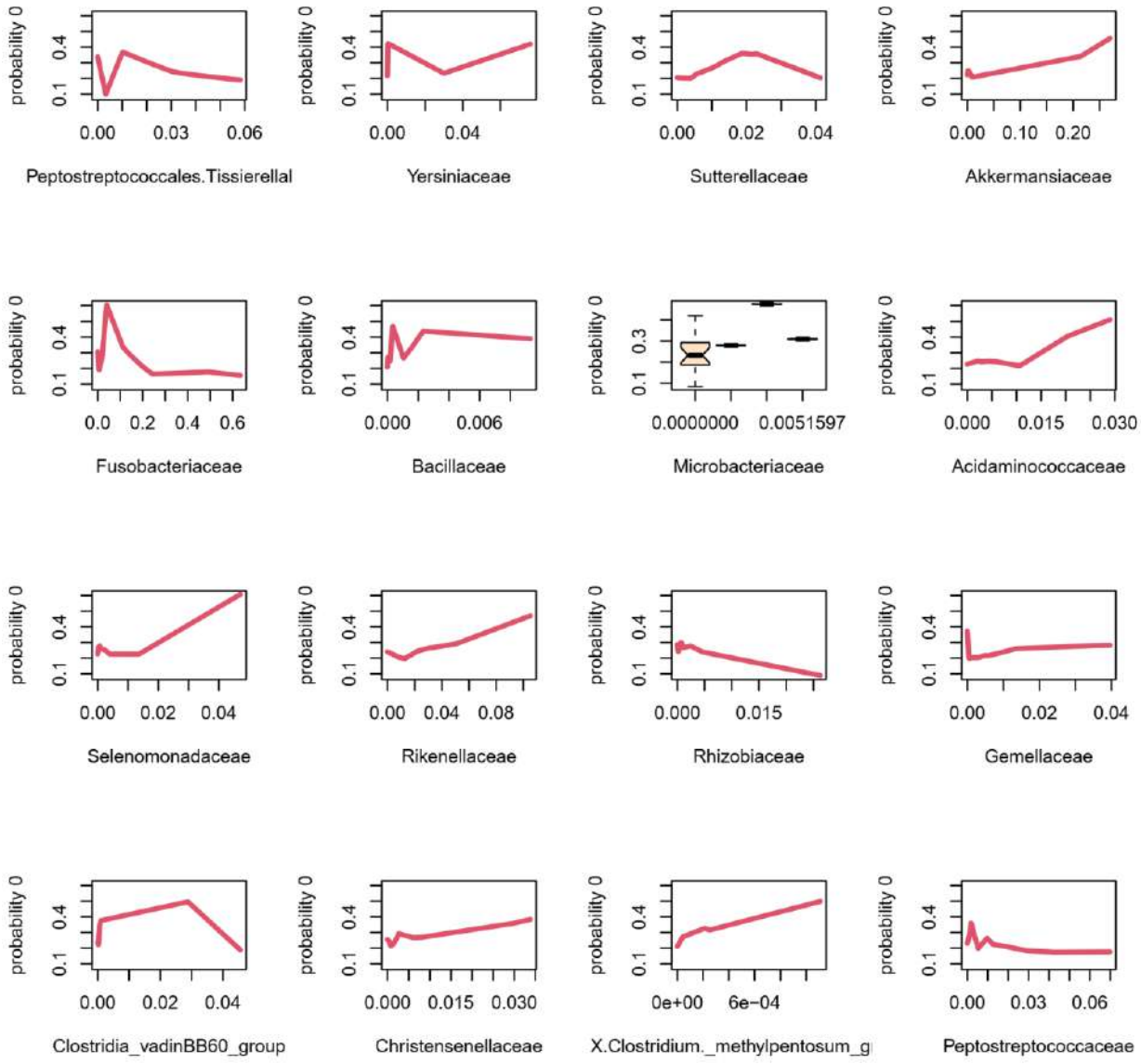
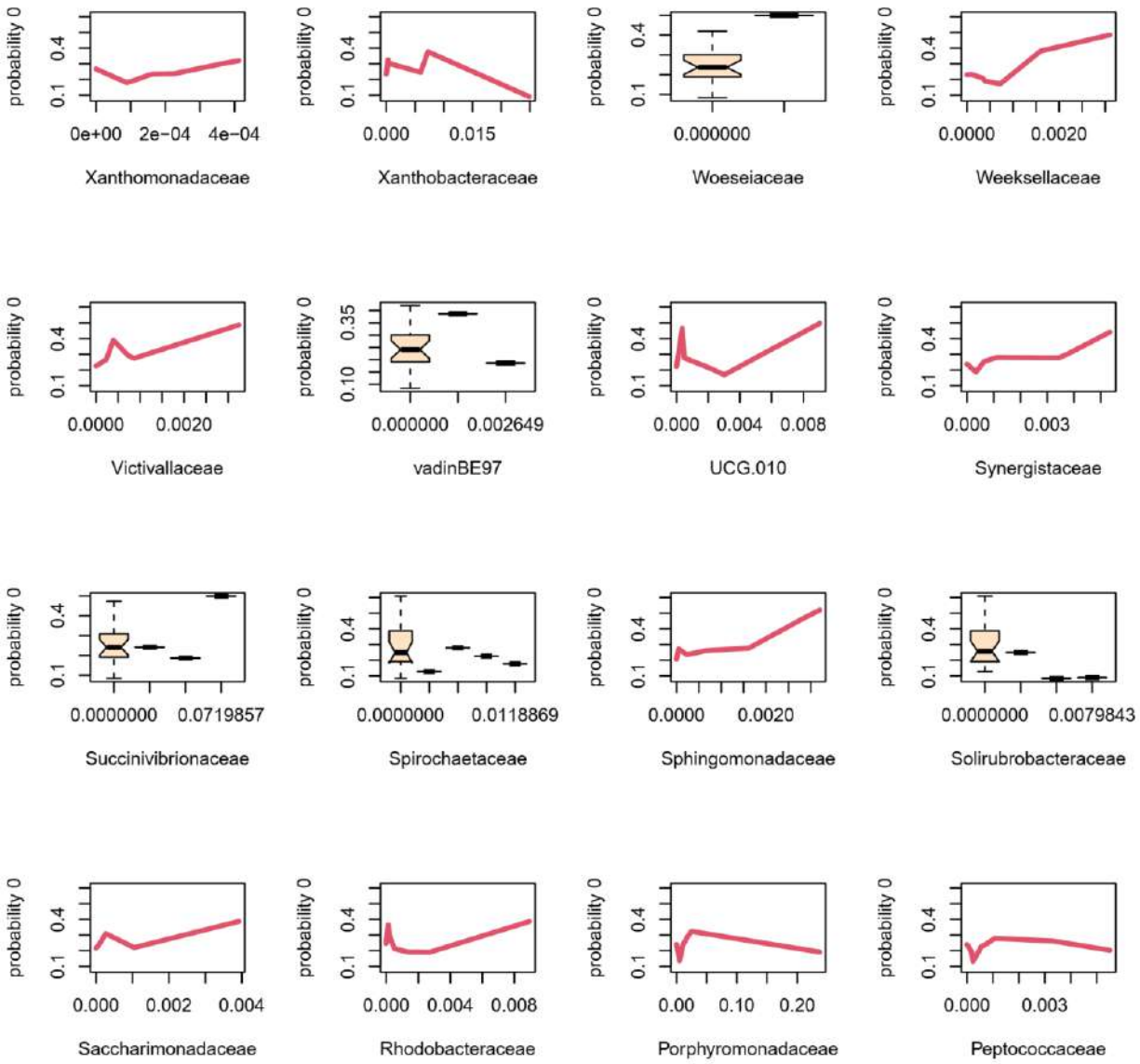
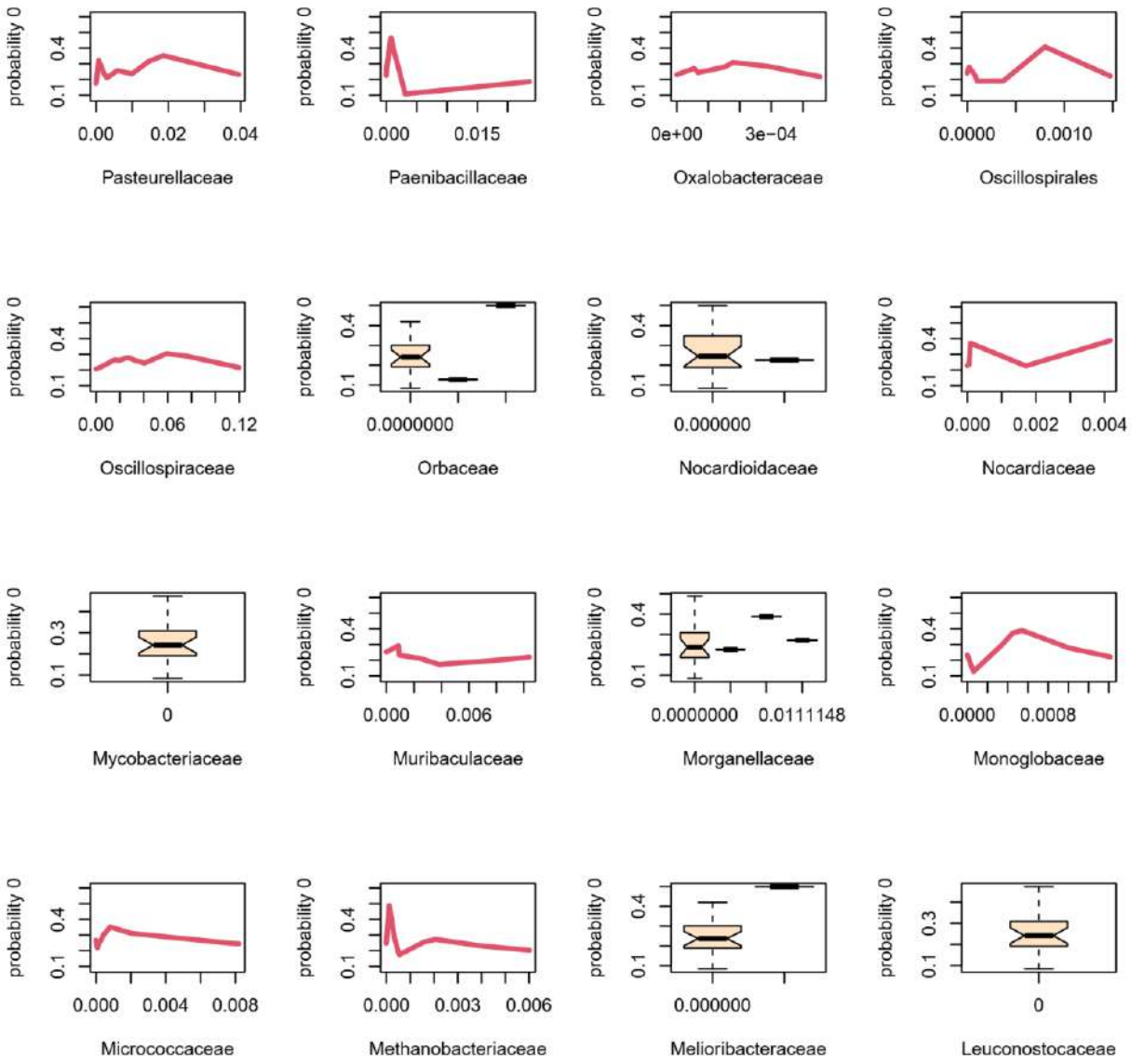
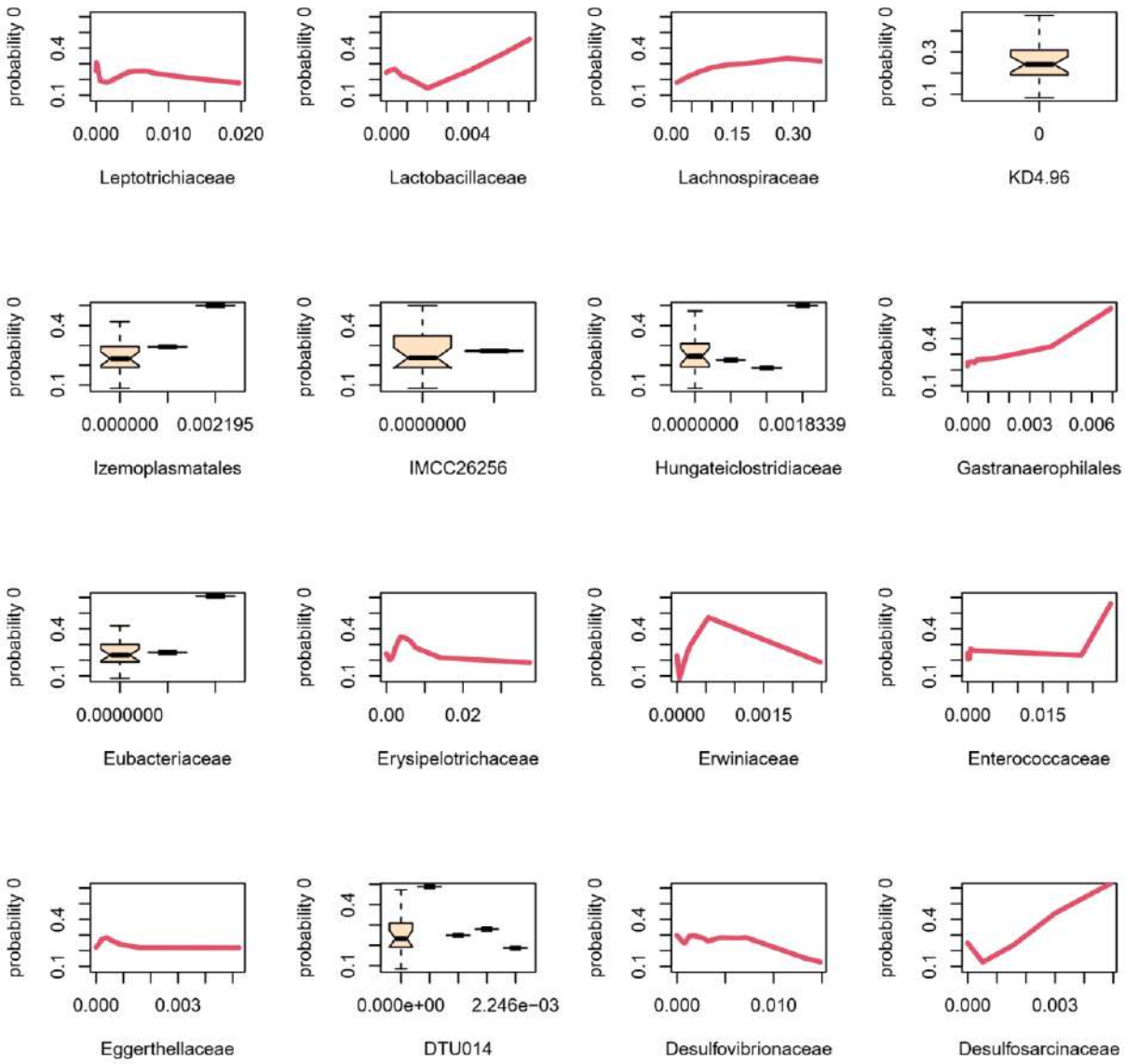


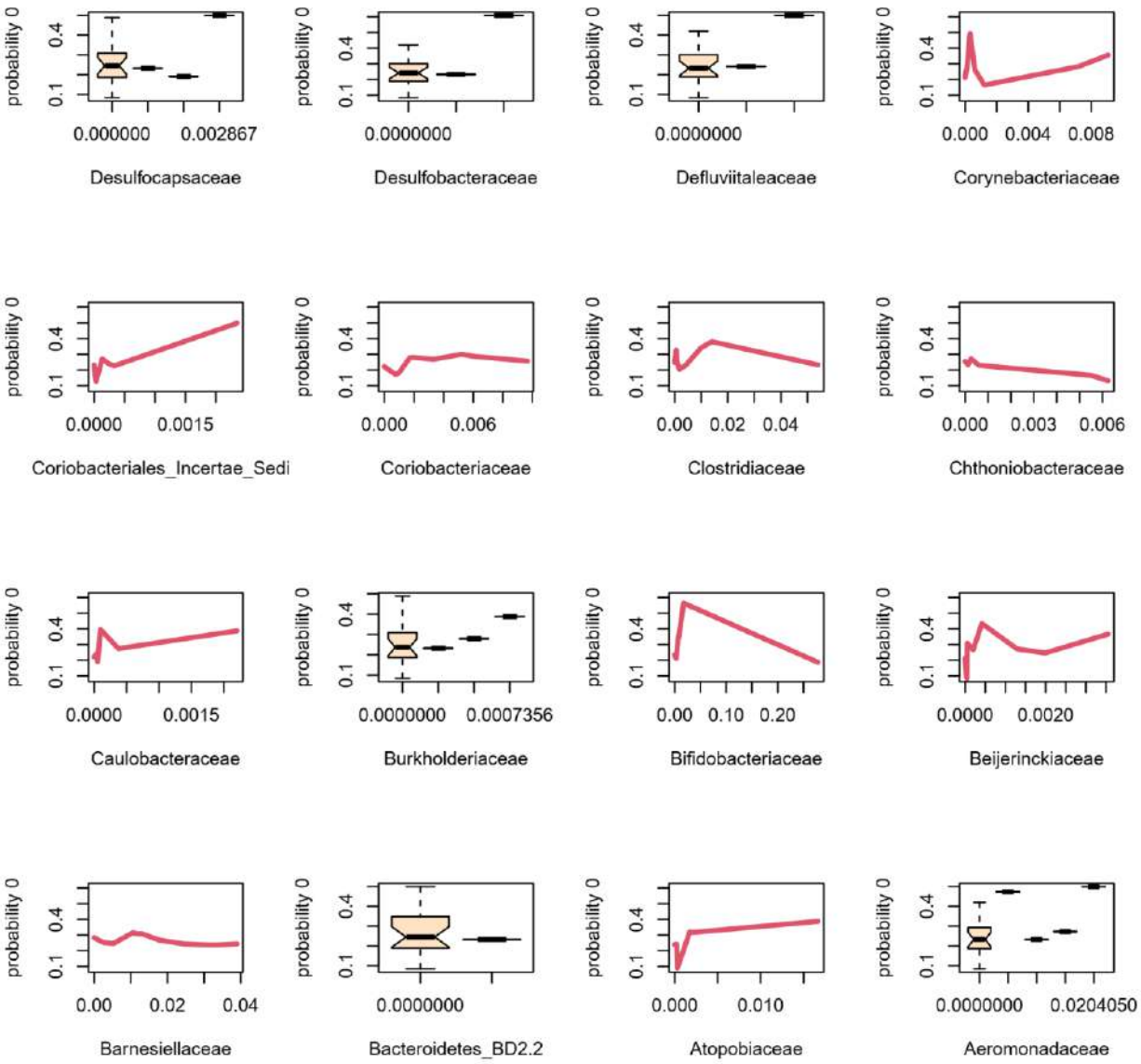
Figure ML1-D-1: The number of trees and Variable Importance Plot for Family in Biopsy

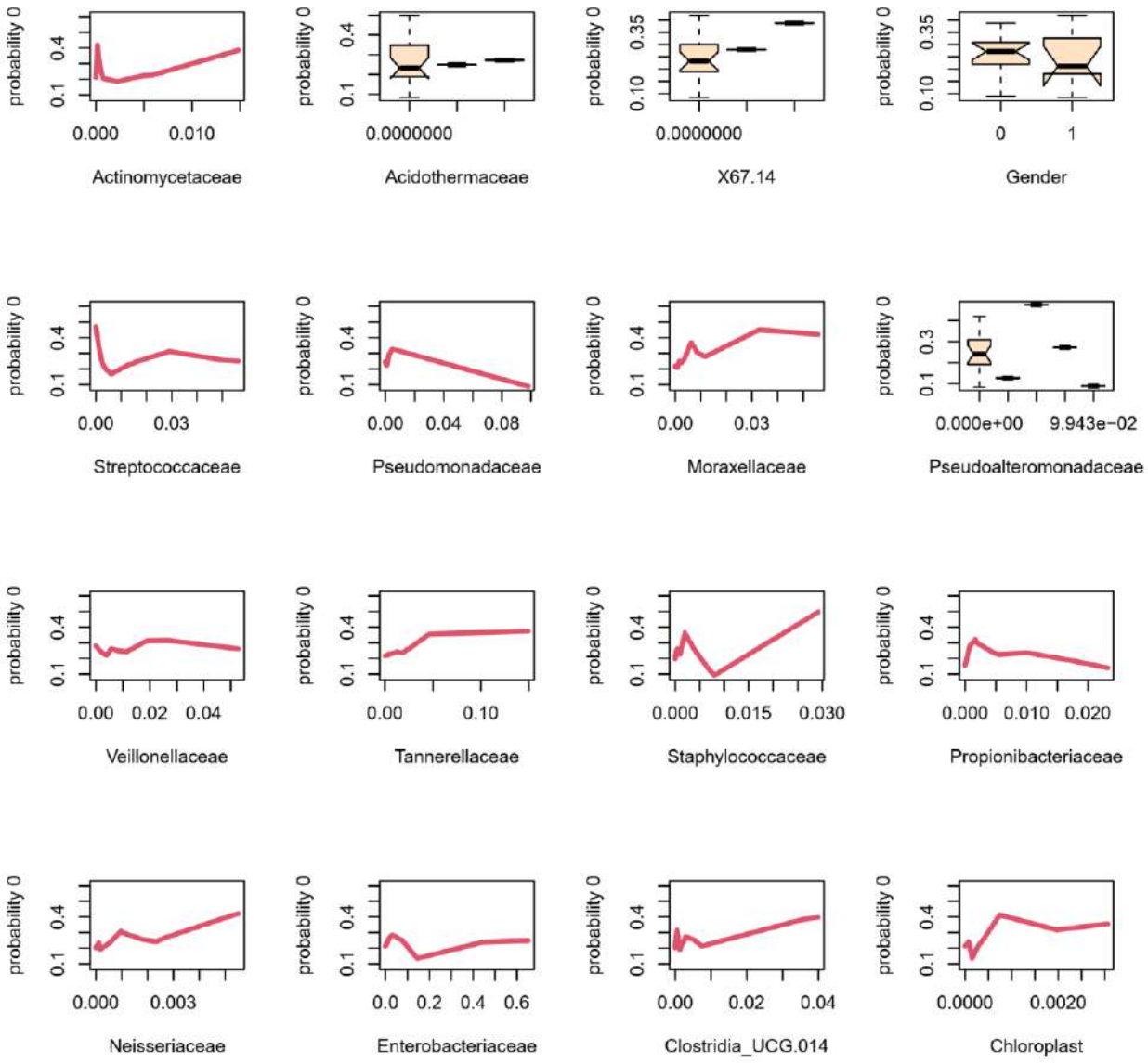



probability 0
0.1
0.4
0e+00 2e-04 4e-04
Xanthomonadaceae
0.000 0.015
Xanthobacteraceae
0.000000
Woeseiaceae
0.0000 0.0020
Weeksellaceae
0.0000 0.0020
Victivallaceae
0.10
0.35
0.000000 0.002649
vadinBE97
0.000 0.004 0.008
UCG.010
0.000 0.003
Synergistaceae
0.0000000 0.0719857
Succinivibrionaceae
0.0000000 0.0118869
Spirochaetaceae
0.0000 0.0020
Sphingomonadaceae
0.0000000 0.0079843
Solirubrobacteraceae
0.000 0.002 0.004
Saccharimonadaceae
0.000 0.004 0.008
Rhodobacteraceae
0.00 0.10 0.20
Porphyromonadaceae
0.000 0.003
Peptococcaceae


probability 0
0.1
0.4
0.00
0.02
0.04
Pasteurellaceae
probability 0
0.1
0.4
0.000
0.015
Paenibacillaceae
probability 0
0.1
0.4
0e+00
3e−04
Oxalobacteraceae
probability 0
0.1
0.4
0.0000
0.0010
Oscillospirales
probability 0
0.1
0.4
0.00
0.06
0.12
Oscillospiraceae
probability 0
0.1
0.4
0.0000000
Orbaceae
probability 0
0.1
0.4
0.000000
Nocardioidaceae
probability 0
0.1
0.4
0.000
0.002
0.004
Nocardiaceae
probability 0
0.1
0.3
0
Mycobacteriaceae
probability 0
0.1
0.4
0.000
0.006
Muribaculaceae
probability 0
0.1
0.4
0.0000000
0.0111148
Morganellaceae
probability 0
0.1
0.4
0.0000
0.0008
Monoglobaceae
probability 0
0.1
0.4
0.000
0.004
0.008
Micrococcaceae
probability 0
0.1
0.4
0.000
0.003
0.006
Methanobacteriaceae
probability 0
0.1
0.4
0.000000
Melioribacteraceae
probability 0
0.1
0.3
0
Leuconostocaceae


probability 0
0.1
0.4
0.000
0.010
0.020
Leptotrichiaceae
probability 0
0.1
0.4
0.000
0.004
Lactobacillaceae
probability 0
0.1
0.4
0.00
0.15
0.30
Lachnospiraceae
probability 0
0.1
0.3
0
KD4.96
probability 0
0.1
0.4
0.000000
0.002195
Izemoplasmatales
probability 0
0.1
0.4
0.0000000
IMCC26256
probability 0
0.1
0.4
0.0000000
0.0018339
Hungateiclostridiaceae
probability 0
0.1
0.4
0.000
0.003
0.006
Gastranaerophilales
probability 0
0.1
0.4
0.0000000
Eubacteriaceae
probability 0
0.1
0.4
0.00
0.02
Erysipelotrichaceae
probability 0
0.1
0.4
0.0000
0.0015
Erwiniaceae
probability 0
0.1
0.4
0.000
0.015
Enterococcaceae
probability 0
0.1
0.4
0.000
0.003
Eggerthellaceae
probability 0
0.1
0.4
0.000e+00
2.246e−03
DTU014
probability 0
0.1
0.4
0.000
0.010
Desulfovibrionaceae
probability 0
0.1
0.4
0.000
0.003
Desulfosarcinaceae


probability 0
0.1
0.4
0.000000
0.002867
Desulfocapsaceae
0.0000000
Desulfobacteraceae
Defluviitaleaceae
0.000
0.004
0.008
Corynebacteriaceae
0.0000
0.0015
Coriobacteriales_Incertae_Sedi
0.006
Coriobacteriaceae
0.00
0.02
0.04
Clostridiaceae
0.003
0.006
Chthoniobacteraceae
Caulobacteraceae
0.0000000
0.0007356
Burkholderiaceae
0.10
0.20
Bifidobacteriaceae
0.0020
Beijerinckiaceae
Barnesiellaceae
Bacteroidetes_BD2.2
0.010
Atopobiaceae
0.0204050
Aeromonadaceae


probability 0
0.1 0.4
0.000 0.010
Actinomycetaceae
probability 0
0.1 0.4
0.0000000
Acidothermaceae
probability 0
0.10 0.35
0.0000000
X67.14
probability 0
0.10 0.35
0 1
Gender
probability 0
0.1 0.4
0.00 0.03
Streptococcaceae
probability 0
0.1 0.4
0.00 0.04 0.08
Pseudomonadaceae
probability 0
0.1 0.4
0.00 0.03
Moraxellaceae
probability 0
0.1 0.3
0.000e+00 9.943e−02
Pseudoalteromonadaceae
probability 0
0.1 0.4
0.00 0.02 0.04
Veillonellaceae
probability 0
0.1 0.4
0.00 0.10
Tannerellaceae
probability 0
0.1 0.4
0.000 0.015 0.030
Staphylococcaceae
probability 0
0.1 0.4
0.000 0.010 0.020
Propionibacteriaceae
probability 0
0.1 0.4
0.000 0.003
Neisseriaceae
probability 0
0.1 0.4
0.0 0.2 0.4 0.6
Enterobacteriaceae
probability 0
0.1 0.4
0.00 0.02 0.04
Clostridia_UCG.014
probability 0
0.1 0.4
0.0000 0.0020
Chloroplast

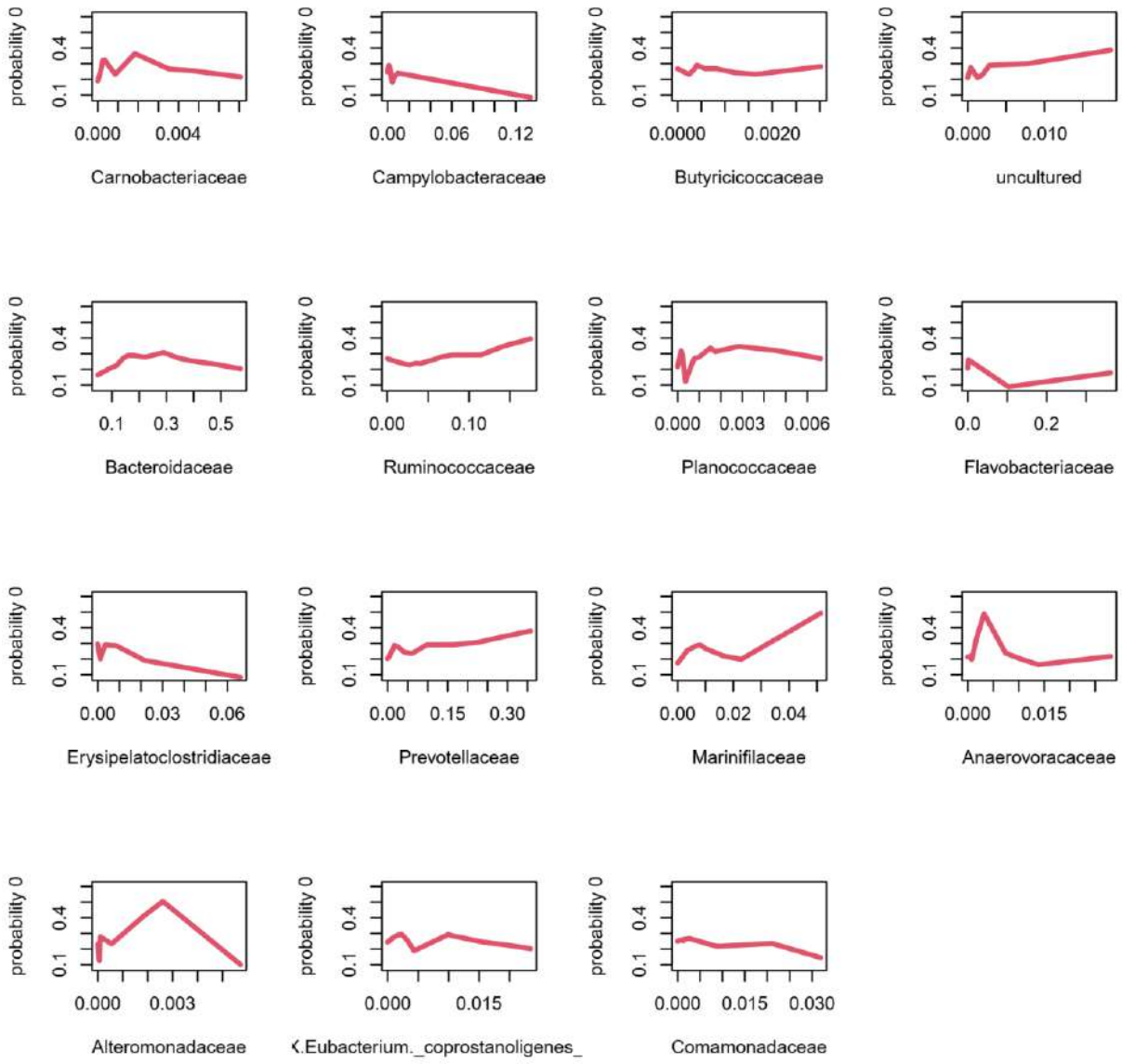
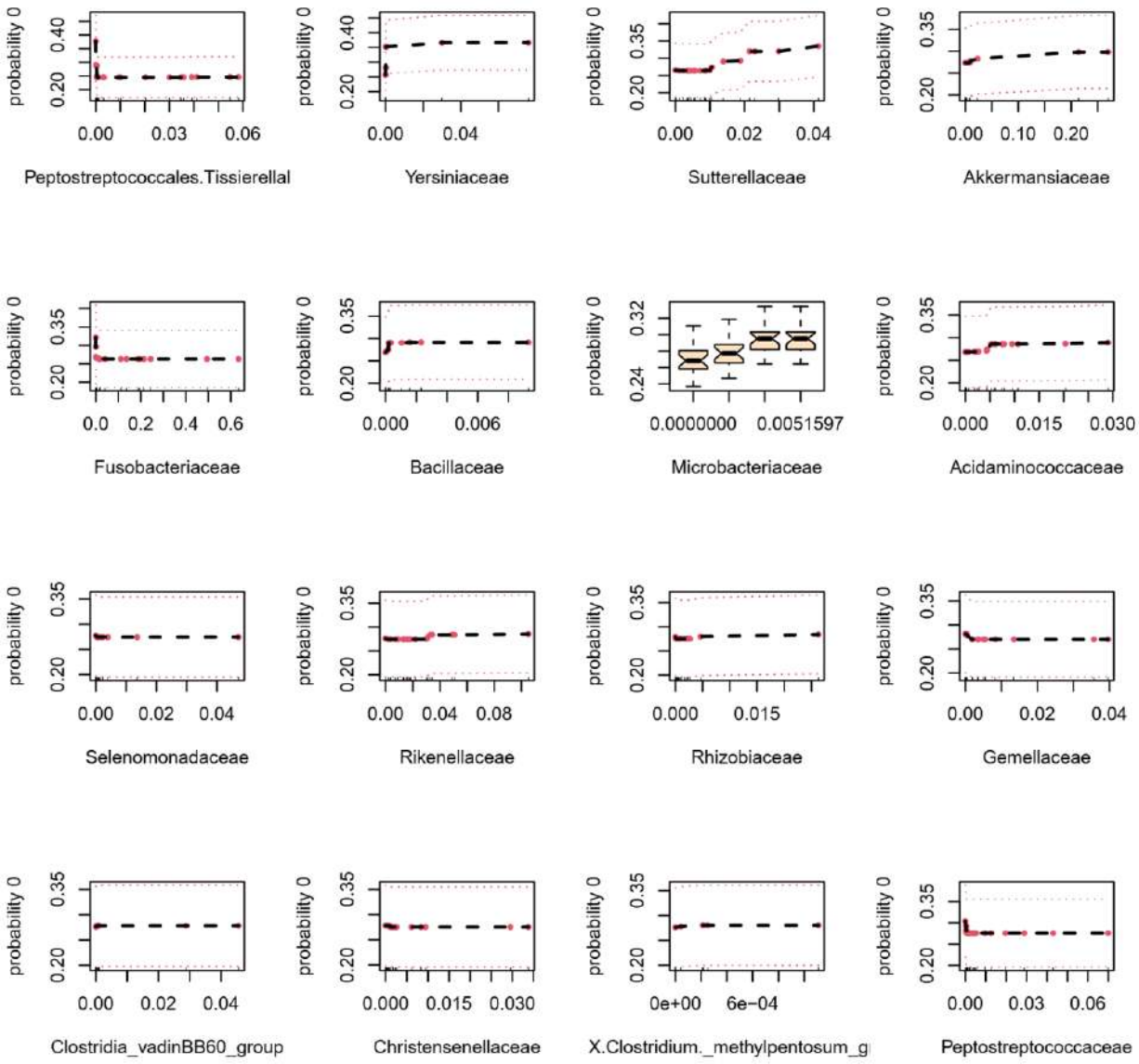
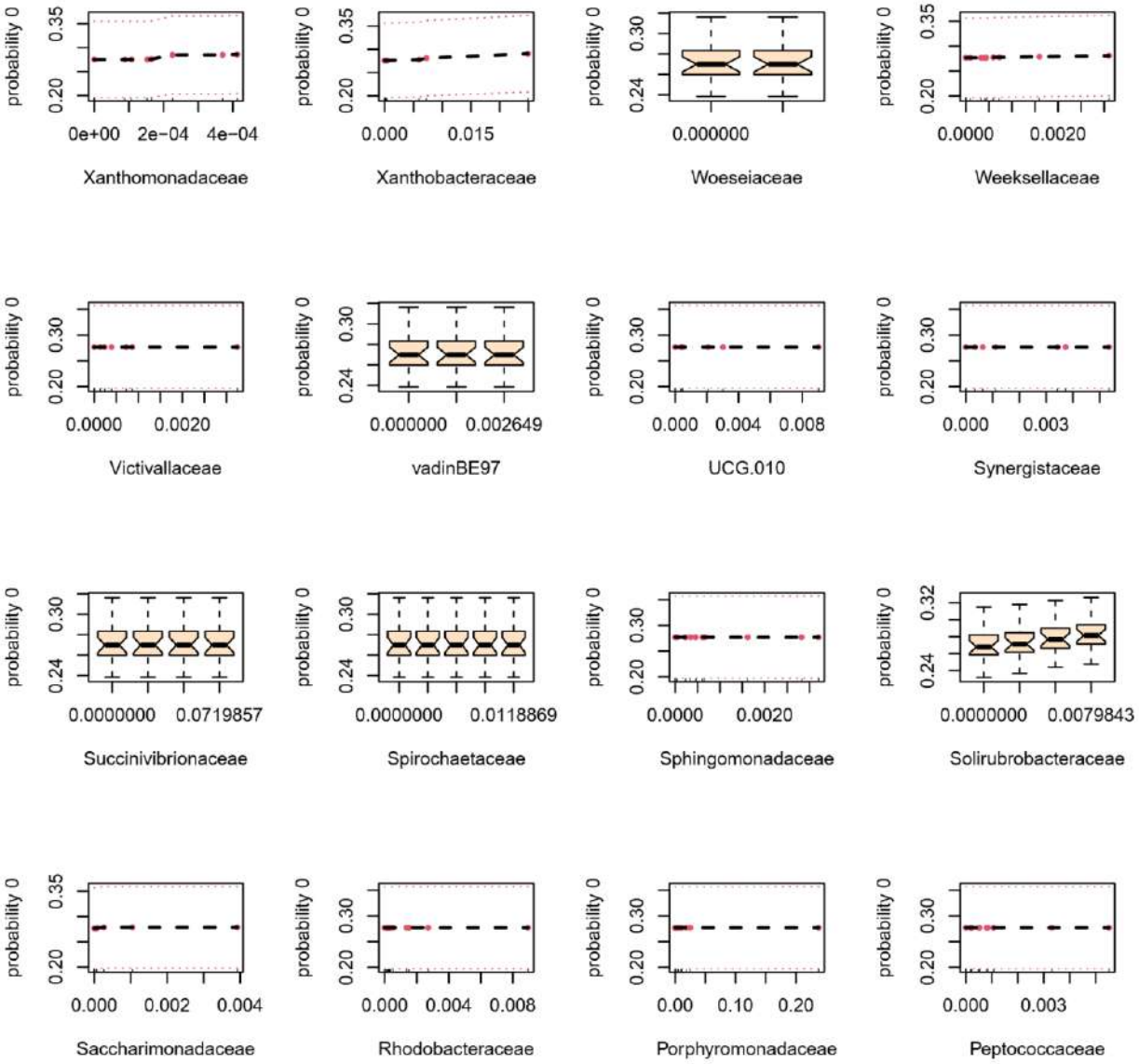
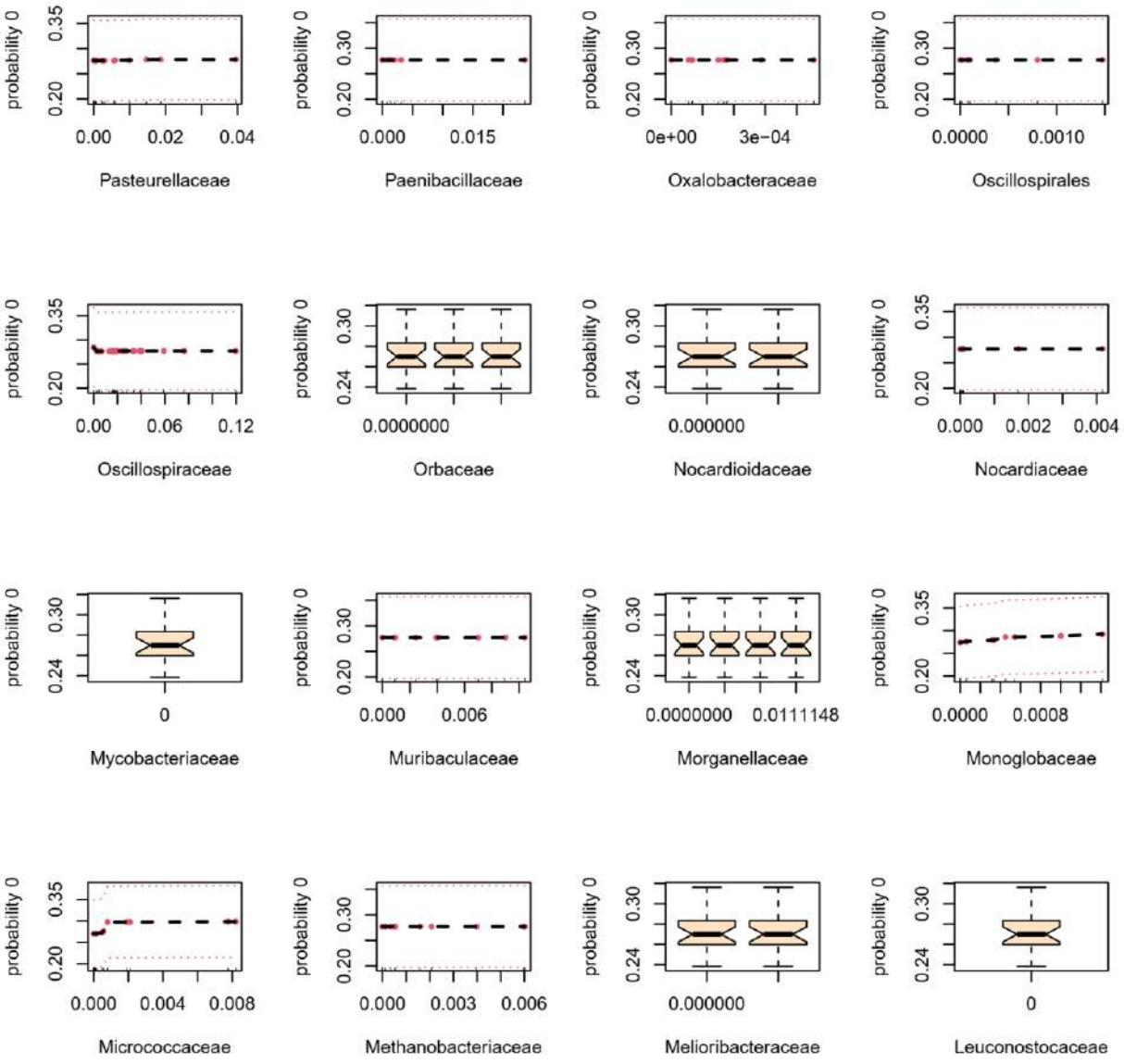
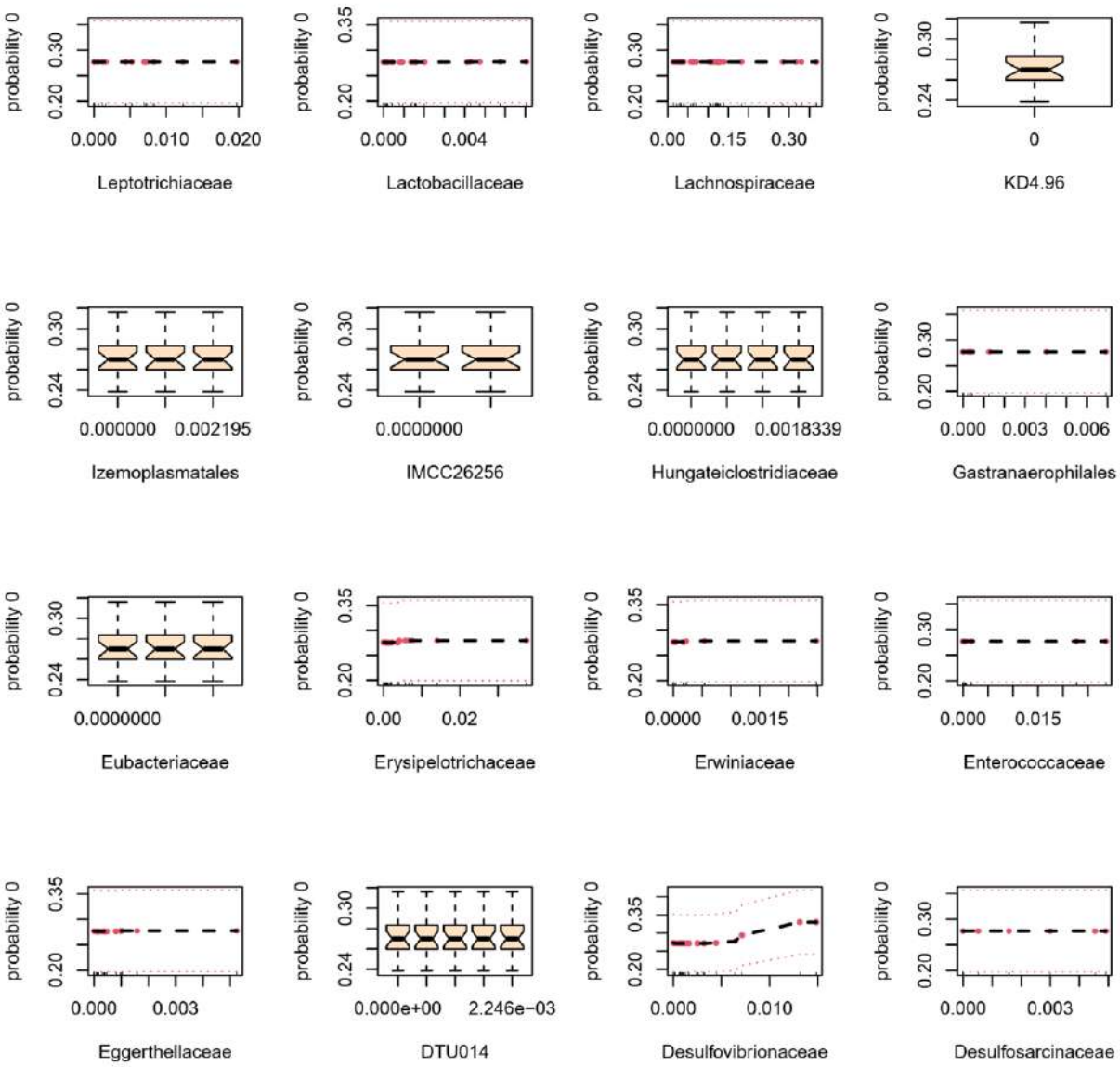


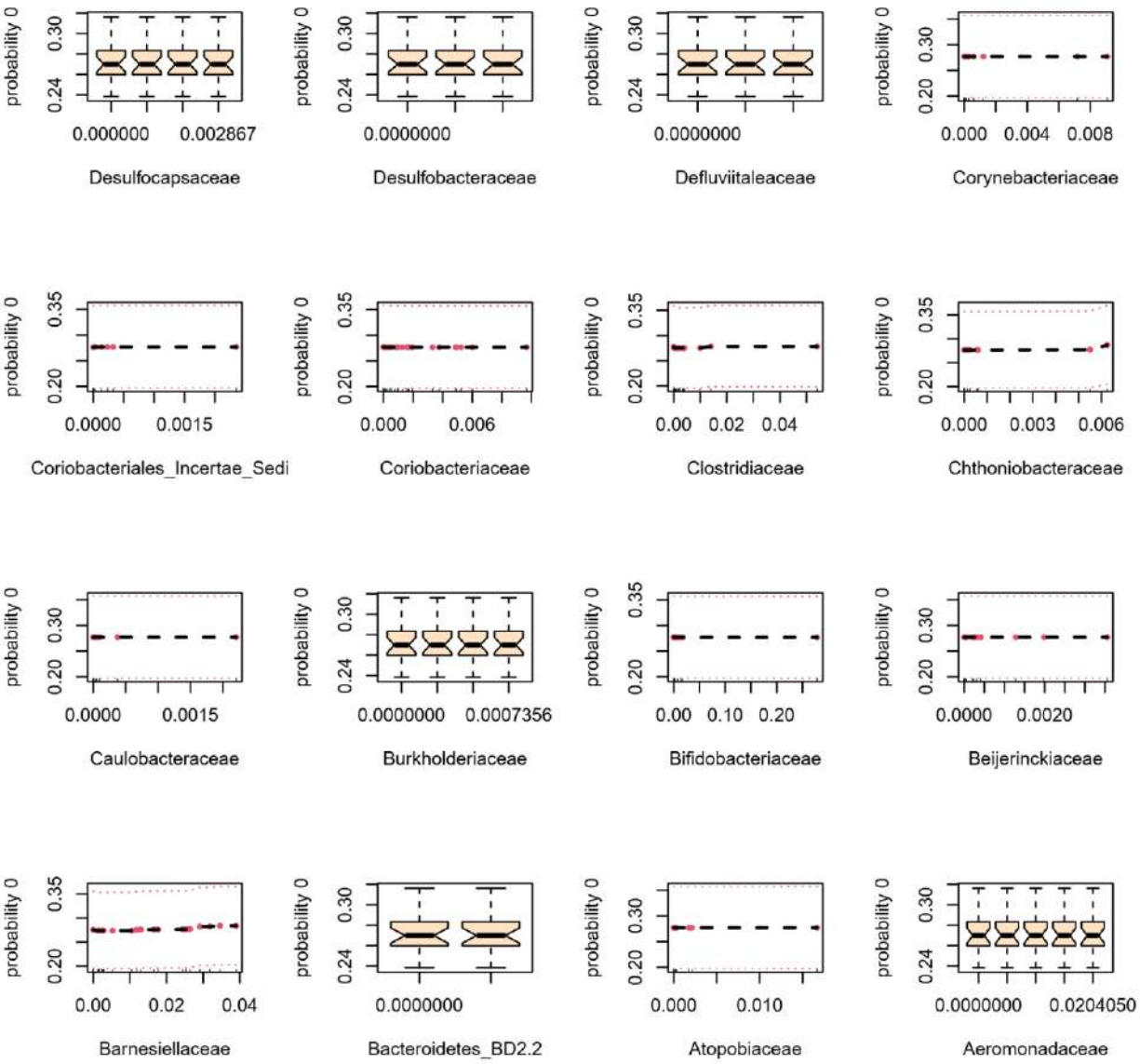
Figure ML1-D-2: Marginal Plot for Family in Biopsy

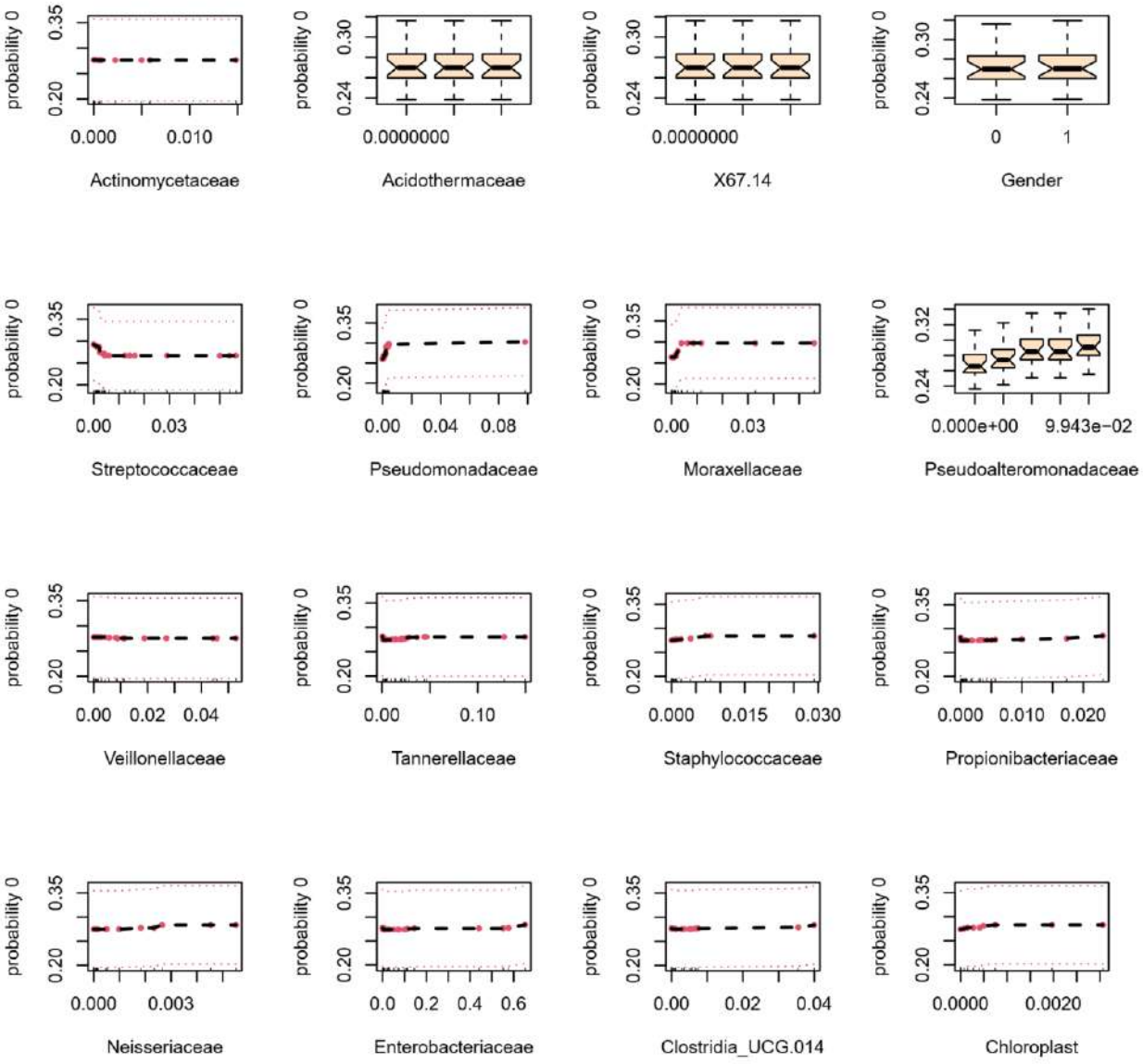

probability 0
Peptostreptococcales.Tissierellal
Yersiniaceae
Sutterellaceae
Akkermansiaceae
Fusobacteriaceae
Bacillaceae
Microbacteriaceae
Acidaminococcaceae
Selenomonadaceae
Rikenellaceae
Rhizobiaceae
Gemellaceae
Clostridia_vadinBB60_group
Christensenellaceae
X.Clostridium._methylpentosum_g
Peptostreptococcaceae


probability 0
0.20
0.35
0e+00 2e-04 4e-04
Xanthomonadaceae
probability 0
0.20
0.35
0.000 0.015
Xanthobacteraceae
probability 0
0.24
0.30
0.000000
Woeseiaceae
probability 0
0.20
0.35
0.0000 0.0020
Weeksellaceae
probability 0
0.20
0.30
0.0000 0.0020
Victivallaceae
probability 0
0.24
0.30
0.000000 0.002649
vadinBE97
probability 0
0.20
0.30
0.000 0.004 0.008
UCG.010
probability 0
0.20
0.30
0.000 0.003
Synergistaceae
probability 0
0.24
0.30
0.0000000 0.0719857
Succinivibrionaceae
probability 0
0.24
0.30
0.0000000 0.0118869
Spirochaetaceae
probability 0
0.20
0.30
0.0000 0.0020
Sphingomonadaceae
probability 0
0.24
0.32
0.0000000 0.0079843
Solirubrobacteraceae
probability 0
0.20
0.35
0.000 0.002 0.004
Saccharimonadaceae
probability 0
0.20
0.30
0.000 0.004 0.008
Rhodobacteraceae
probability 0
0.20
0.30
0.00 0.10 0.20
Porphyromonadaceae
probability 0
0.20
0.30
0.000 0.003
Peptococcaceae


probability 0
Pasteurellaceae
Paenibacillaceae
Oxalobacteraceae
Oscillospirales
Oscillospiraceae
Orbaceae
Nocardioidaceae
Nocardiaceae
Mycobacteriaceae
Muribaculaceae
Morganellaceae
Monoglobaceae
Micrococcaceae
Methanobacteriaceae
Melioribacteraceae
Leuconostocaceae


probability 0
0.20
0.30
0.000
0.010
0.020
Leptotrichiaceae
probability 0
0.20
0.35
0.000
0.004
Lactobacillaceae
probability 0
0.20
0.30
0.00
0.15
0.30
Lachnospiraceae
probability 0
0.24
0.30
0
KD4.96
probability 0
0.24
0.30
0.000000
0.002195
Izemoplasmatales
probability 0
0.24
0.30
0.0000000
IMCC26256
probability 0
0.24
0.30
0.0000000
0.0018339
Hungateiclostridiaceae
probability 0
0.20
0.30
0.000
0.003
0.006
Gastranaerophilales
probability 0
0.24
0.30
0.0000000
Eubacteriaceae
probability 0
0.20
0.35
0.00
0.02
Erysipelotrichaceae
probability 0
0.20
0.35
0.0000
0.0015
Erwiniaceae
probability 0
0.20
0.30
0.000
0.015
Enterococcaceae
probability 0
0.20
0.35
0.000
0.003
Eggerthellaceae
probability 0
0.24
0.30
0.000e+00
2.246e−03
DTU014
probability 0
0.20
0.35
0.000
0.010
Desulfovibrionaceae
probability 0
0.20
0.30
0.000
0.003
Desulfosarcinaceae


probability 0
0.24
0.30
0.000000
0.002867
Desulfocapsaceae
probability 0
0.24
0.30
0.0000000
Desulfobacteraceae
probability 0
0.24
0.30
0.0000000
Defluviitaleaceae
probability 0
0.20
0.30
0.000
0.004
0.008
Corynebacteriaceae
probability 0
0.20
0.35
0.0000
0.0015
Coriobacteriales_Incertae_Sedi
probability 0
0.20
0.35
0.000
0.006
Coriobacteriaceae
probability 0
0.20
0.35
0.00
0.02
0.04
Clostridiaceae
probability 0
0.20
0.35
0.000
0.003
0.006
Chthoniobacteraceae
probability 0
0.20
0.30
0.0000
0.0015
Caulobacteraceae
probability 0
0.24
0.30
0.0000000
0.0007356
Burkholderiaceae
probability 0
0.20
0.35
0.00
0.10
0.20
Bifidobacteriaceae
probability 0
0.20
0.30
0.0000
0.0020
Beijerinckiaceae
probability 0
0.20
0.35
0.00
0.02
0.04
Barnesiellaceae
probability 0
0.24
0.30
0.0000000
Bacteroidetes_BD2.2
probability 0
0.20
0.30
0.000
0.010
Atopobiaceae
probability 0
0.24
0.30
0.0000000
0.0204050
Aeromonadaceae


probability 0
Actinomycetaceae
Acidothermaceae
X67.14
Gender
Streptococcaceae
Pseudomonadaceae
Moraxellaceae
Pseudoalteromonadaceae
Veillonellaceae
Tannerellaceae
Staphylococcaceae
Propionibacteriaceae
Neisseriaceae
Enterobacteriaceae
Clostridia_UCG.014
Chloroplast

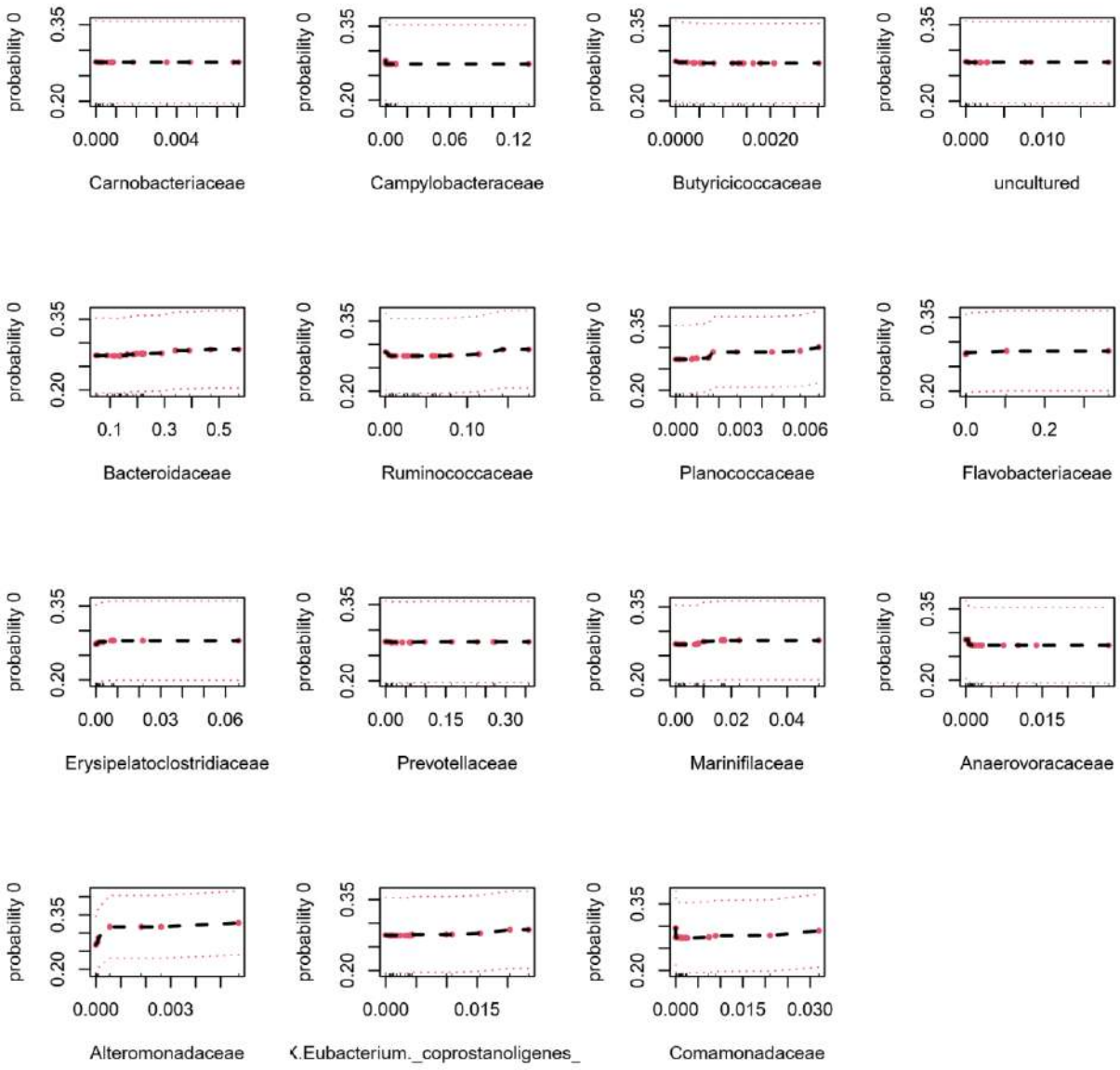


Figure ML1-D-3: Partial Plot for Family in Biopsy

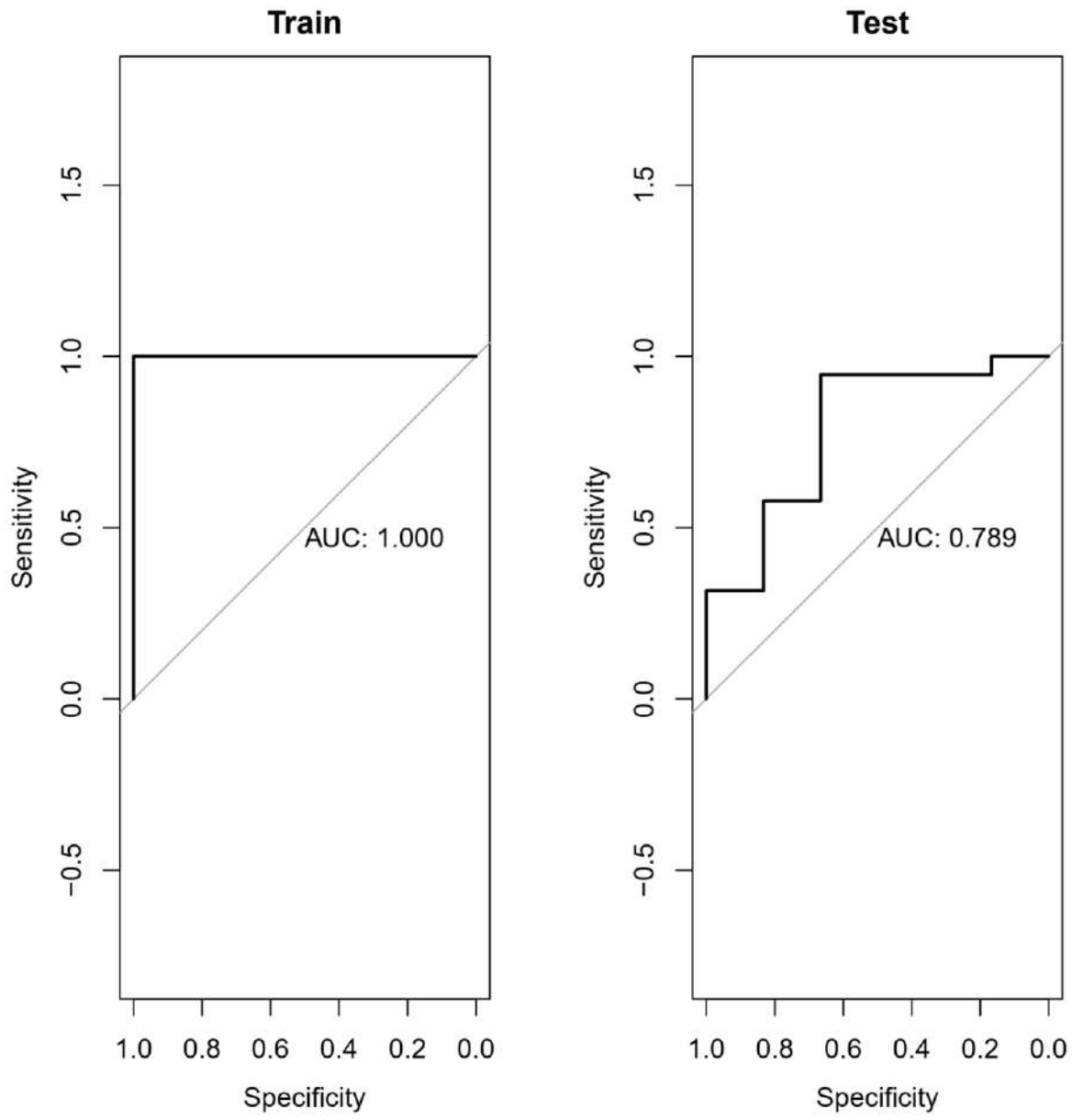


Figure ML1-D-4: The ROC Curve for Family in Biopsy

Table ML1-D-1: The Model Performance

The Model Performance			
	Values		
Index	2.50%	50.00%	97.50%
Accuracy	0.52	0.88	1.00
Specificity	0.50	0.83	1.00
Sensitivity	0.37	0.95	1.00

Table ML1-D-2: The Variable Importance Values

Covariates	Variable Importance			Brier Score		
	all	AP	CRC	all	AP	CRC
Gender	-0.0008	0.0000	-0.0029	-0.0008	-0.0005	-0.0005
X.Clostridium._methylpentosum_group	0.0004	-0.0181	0.0072	0.0015	0.0010	0.0010
X.Eubacterium._coprostanoligenes_group	0.0041	-0.0362	0.0258	0.0046	0.0032	0.0032
X67.14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Acidaminococcaceae	0.0048	0.0272	0.0086	0.0004	0.0003	0.0003
Acidothermaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Actinomycetaceae	0.0000	0.0045	-0.0014	-0.0007	-0.0005	-0.0005
Aeromonadaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Akkermansiaceae	0.0120	-0.0453	0.0572	0.0231	0.0157	0.0157
Alteromonadaceae	0.0203	-0.1269	0.1116	0.0451	0.0306	0.0306
Anaerovoracaceae	-0.0020	-0.0408	0.0057	-0.0019	-0.0013	-0.0013
Atopobiaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Bacillaceae	0.0037	0.0091	0.0100	-0.0031	-0.0021	-0.0021
Bacteroidaceae	0.0028	-0.0181	0.0157	-0.0011	-0.0007	-0.0007
Bacteroidetes_BD2.2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Barnesiellaceae	-0.0045	-0.0498	0.0000	-0.0003	-0.0002	-0.0002
Beijerinckiaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Bifidobacteriaceae	0.0000	0.0000	0.0000	-0.0012	-0.0008	-0.0008
Burkholderiaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Butyricicoccaceae	-0.0016	-0.0181	0.0000	-0.0084	-0.0057	-0.0057
Campylobacteraceae	-0.0021	-0.0045	-0.0057	-0.0052	-0.0035	-0.0035
Carnobacteriaceae	0.0000	0.0000	0.0000	-0.0009	-0.0006	-0.0006
Caulobacteraceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Chloroplast	-0.0032	-0.0272	-0.0029	-0.0116	-0.0079	-0.0079
Christensenellaceae	-0.0021	-0.0045	-0.0057	-0.0049	-0.0034	-0.0034
Chthoniobacteraceae	0.0069	-0.0227	0.0315	0.0088	0.0060	0.0060
Clostridia_UCG.014	0.0012	-0.0272	0.0129	0.0013	0.0009	0.0009
Clostridia_vadinBB60_group	0.0008	0.0000	0.0029	-0.0008	-0.0006	-0.0006

Clostridiaceae	-0.0016	-0.0091	-0.0029	-0.0044	-0.0030	-0.0030
Comamonadaceae	0.0154	-0.0951	0.0844	0.0261	0.0178	0.0178
Coriobacteriaceae	0.0000	0.0000	0.0000	-0.0009	-0.0006	-0.0006
Coriobacteriales_Incertae_Sedis	-0.0004	0.0000	-0.0014	-0.0008	-0.0005	-0.0005
Corynebacteriaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Defluviitaleaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Desulfobacteraceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Desulfocapsaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Desulfosarcinaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Desulfovibrionaceae	0.0350	-0.1404	0.1674	0.0942	0.0640	0.0640
DTU014	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Eggerthellaceae	0.0008	0.0000	0.0029	-0.0012	-0.0008	-0.0008
Enterobacteriaceae	0.0001	-0.0453	0.0143	-0.0036	-0.0025	-0.0025
Enterococcaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Erwiniaceae	0.0000	-0.0136	0.0043	-0.0031	-0.0021	-0.0021
Erysipelatoclostridiaceae	-0.0024	-0.0181	-0.0029	-0.0021	-0.0014	-0.0014
Erysipelotrichaceae	-0.0017	-0.0227	0.0014	-0.0040	-0.0027	-0.0027
Eubacteriaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Flavobacteriaceae	0.0017	-0.0091	0.0086	0.0023	0.0015	0.0015
Fusobacteriaceae	0.0310	-0.0136	0.1145	0.0347	0.0236	0.0236
Gastranaerophilales	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Gemellaceae	-0.0016	0.0045	-0.0072	-0.0050	-0.0034	-0.0034
Hungateiclostridiaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
IMCC26256	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Izemoplasmatales	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
KD4.96	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Lachnospiraceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Lactobacillaceae	0.0004	0.0000	0.0014	-0.0002	-0.0001	-0.0001
Leptotrichiaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Leuconostocaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Marinifilaceae	-0.0004	0.0000	-0.0014	-0.0063	-0.0043	-0.0043
Melioribacteraceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Methanobacteriaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Microbacteriaceae	0.0095	-0.0453	0.0472	0.0104	0.0070	0.0070
Micrococcaceae	0.0078	-0.0317	0.0372	0.0053	0.0036	0.0036
Monoglobaceae	0.0000	-0.0680	0.0215	-0.0063	-0.0043	-0.0043
Moraxellaceae	-0.0029	-0.0317	0.0000	0.0084	0.0057	0.0057
Morganellaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Muribaculaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Mycobacteriaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Neisseriaceae	0.0020	-0.0045	0.0086	0.0011	0.0007	0.0007
Nocardiaceae	0.0000	0.0000	0.0000	-0.0007	-0.0005	-0.0005
Nocardioidaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Orbaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Oscillospiraceae	0.0070	-0.0091	0.0272	0.0081	0.0055	0.0055
Oscillospirales	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Oxalobacteraceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Paenibacillaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Pasteurellaceae	-0.0008	-0.0091	0.0000	-0.0046	-0.0031	-0.0031
Peptococcaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Peptostreptococcaceae	0.0093	-0.0770	0.0572	0.0202	0.0137	0.0137
Peptostreptococcales.Tissierellales	0.0943	0.0770	0.3090	0.1830	0.1243	0.1243
Planococcaceae	0.0061	-0.0589	0.0401	0.0111	0.0076	0.0076
Porphyromonadaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Prevotellaceae	-0.0008	0.0000	-0.0029	-0.0044	-0.0030	-0.0030
Propionibacteriaceae	0.0035	-0.0634	0.0329	0.0039	0.0026	0.0026
Pseudoalteromonadaceae	0.0106	-0.0498	0.0529	0.0112	0.0076	0.0076
Pseudomonadaceae	0.0053	-0.0136	0.0229	0.0122	0.0083	0.0083
Rhizobiaceae	0.0017	-0.0498	0.0215	0.0041	0.0028	0.0028
Rhodobacteraceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Rikenellaceae	0.0025	-0.0227	0.0157	-0.0037	-0.0025	-0.0025
Ruminococcaceae	0.0069	-0.0680	0.0458	0.0128	0.0087	0.0087
Saccharimonadaceae	-0.0004	-0.0091	0.0014	-0.0027	-0.0018	-0.0018
Selenomonadaceae	0.0004	0.0136	-0.0029	-0.0008	-0.0006	-0.0006
Solirubrobacteraceae	0.0045	-0.0408	0.0286	0.0029	0.0020	0.0020
Sphingomonadaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Spirochaetaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Staphylococcaceae	0.0012	-0.0181	0.0100	-0.0015	-0.0010	-0.0010
Streptococcaceae	0.0011	-0.0136	0.0086	0.0007	0.0005	0.0005
Succinivibrionaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Sutterellaceae	0.0352	-0.0680	0.1459	0.0491	0.0334	0.0334
Synergistaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Tannerellaceae	0.0044	-0.0181	0.0215	-0.0009	-0.0006	-0.0006
UCG.010	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
uncultured	-0.0016	-0.0091	-0.0029	-0.0047	-0.0032	-0.0032
vadinBE97	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Veillonellaceae	-0.0020	0.0000	-0.0072	-0.0033	-0.0023	-0.0023
Victivallaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Weeksellaceae	0.0004	-0.0136	0.0057	-0.0049	-0.0033	-0.0033
Woeseiaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Xanthobacteraceae	0.0052	-0.0408	0.0315	0.0104	0.0071	0.0071
Xanthomonadaceae	0.0048	-0.0272	0.0258	0.0030	0.0020	0.0020
Yersiniaceae	0.0747	0.0362	0.2518	0.1497	0.1017	0.1017

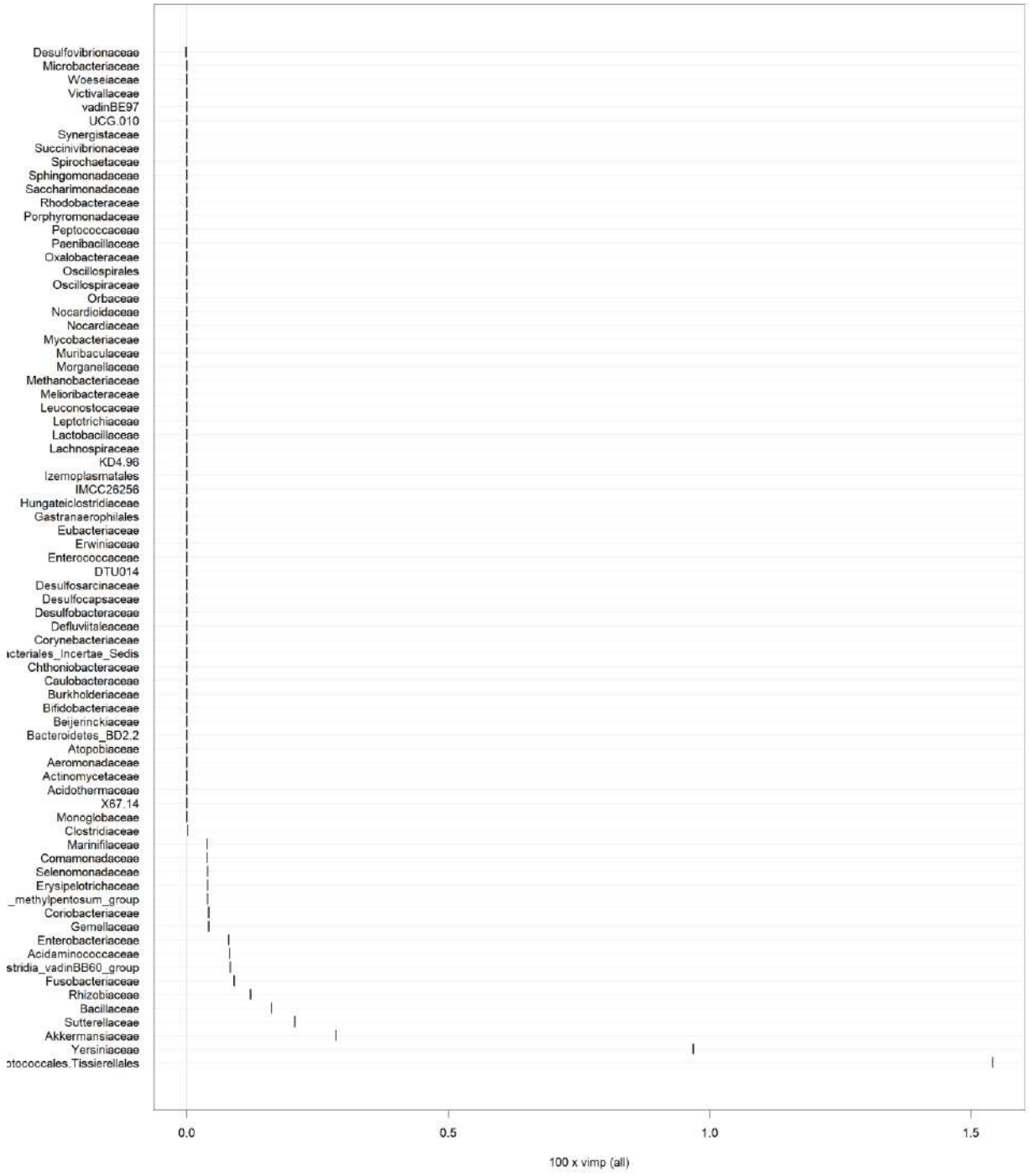


Figure ML1-D-5: The Variable Importance Plot (VIMP) for Family in Biopsy

Best model is selected for LIME results

Table ML1-D-3: The best setting for model and LIME Results

Setting	mtry	ROC	Sens	Spec	ROCSD	SensSD	SpecSD
ntrees: 500 nodesize: 1	8	0.825	0.000	1.000	0.168	0.000	0.000
ntrees: 500 nodesize: 1	9	0.825	0.000	1.000	0.168	0.000	0.000
ntrees: 500 nodesize: 1	10	0.825	0.000	1.000	0.168	0.000	0.000
ntrees: 1000 nodesize: 1	1	0.850	0.000	1.000	0.137	0.000	0.000
ntrees: 1000 nodesize: 1	5	0.850	0.000	1.000	0.137	0.000	0.000
ntrees: 500 nodesize: 5	6	0.875	0.000	1.000	0.125	0.000	0.000
ntrees: 1000 nodesize: 5	1	0.850	0.000	1.000	0.137	0.000	0.000

Table ML1-D-4: The LIME Result for Training Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
2	Adenoma	0.740	0.325	0.300	0.448	Peptostreptococcales.Tissierellales	0.000	0.040	Peptostreptococcales.Tissierellales <= 0.000231
2	Adenoma	0.740	0.325	0.300	0.448	Fusobacteriaceae	0.000	0.029	Fusobacteriaceae <= 0.000517
2	Adenoma	0.740	0.325	0.300	0.448	Sutterellaceae	0.021	0.030	0.01055 < Sutterellaceae
2	Adenoma	0.740	0.325	0.300	0.448	Moraxellaceae	0.009	0.025	0.003850 < Moraxellaceae
2	Adenoma	0.740	0.325	0.300	0.448	Rikenellaceae	0.105	0.024	0.02979 < Rikenellaceae
3	Adenoma	0.656	0.167	0.328	0.395	Moraxellaceae	0.004	0.025	0.003850 < Moraxellaceae
3	Adenoma	0.656	0.167	0.328	0.395	Pseudomonadaceae	0.098	0.024	0.002447 < Pseudomonadaceae
3	Adenoma	0.656	0.167	0.328	0.395	Planococcaceae	0.007	0.022	0.000978 < Planococcaceae
3	Adenoma	0.656	0.167	0.328	0.395	Bacillaceae	0.000	-0.019	Bacillaceae <= 0.000217
3	Adenoma	0.656	0.167	0.328	0.395	Gemellaceae	0.000	0.015	Gemellaceae <= 0.000438
6	Adenoma	0.628	0.291	0.319	0.415	Peptostreptococcales.Tissierellales	0.000	0.040	Peptostreptococcales.Tissierellales <= 0.000231
6	Adenoma	0.628	0.291	0.319	0.415	Sutterellaceae	0.041	0.031	0.01055 < Sutterellaceae
6	Adenoma	0.628	0.291	0.319	0.415	Peptostreptococcaceae	0.000	0.023	Peptostreptococcaceae <= 0.000792
6	Adenoma	0.628	0.291	0.319	0.415	Moraxellaceae	0.000	-0.018	Moraxellaceae <= 0.000927
6	Adenoma	0.628	0.291	0.319	0.415	Rikenellaceae	0.034	0.021	0.02979 < Rikenellaceae
7	Adenoma	0.628	0.306	0.316	0.420	Peptostreptococcales.Tissierellales	0.000	0.041	Peptostreptococcales.Tissierellales <= 0.000231
7	Adenoma	0.628	0.306	0.316	0.420	Sutterellaceae	0.014	0.030	0.01055 < Sutterellaceae
7	Adenoma	0.628	0.306	0.316	0.420	Fusobacteriaceae	0.000	0.028	Fusobacteriaceae <= 0.000517
7	Adenoma	0.628	0.306	0.316	0.420	Rikenellaceae	0.033	0.023	0.02979 < Rikenellaceae
7	Adenoma	0.628	0.306	0.316	0.420	Moraxellaceae	0.000	-0.018	Moraxellaceae <= 0.000927
10	Adenoma	0.726	0.332	0.301	0.447	Peptostreptococcales.Tissierellales	0.000	0.040	Peptostreptococcales.Tissierellales <= 0.000231
10	Adenoma	0.726	0.332	0.301	0.447	Fusobacteriaceae	0.000	0.029	Fusobacteriaceae <= 0.000517
10	Adenoma	0.726	0.332	0.301	0.447	Sutterellaceae	0.023	0.030	0.01055 < Sutterellaceae
10	Adenoma	0.726	0.332	0.301	0.447	Pseudomonadaceae	0.003	0.024	0.002447 < Pseudomonadaceae
10	Adenoma	0.726	0.332	0.301	0.447	Planococcaceae	0.002	0.023	0.000978 < Planococcaceae
11	Adenoma	0.624	0.179	0.322	0.397	Fusobacteriaceae	0.000	0.031	Fusobacteriaceae <= 0.000517

11	Adenoma	0.624	0.179	0.322	0.397	Peptostreptococcaceae	0.001	0.023	Peptostreptococcaceae <= 0.000792
11	Adenoma	0.624	0.179	0.322	0.397	Planococcaceae	0.003	0.021	0.000978 < Planococcaceae
11	Adenoma	0.624	0.179	0.322	0.397	Acidaminococcaceae	0.000	-0.014	Acidaminococcaceae <= 0.00272
11	Adenoma	0.624	0.179	0.322	0.397	Bacillaceae	0.002	0.014	0.000217 < Bacillaceae
14	Adenocarcinoma	0.902	0.098	0.626	0.703	Moraxellaceae	0.000	0.016	Moraxellaceae <= 0.000927
14	Adenocarcinoma	0.902	0.098	0.626	0.703	Bacillaceae	0.000	0.019	Bacillaceae <= 0.000217
14	Adenocarcinoma	0.902	0.098	0.626	0.703	Acidaminococcaceae	0.000	0.014	Acidaminococcaceae <= 0.00272
14	Adenocarcinoma	0.902	0.098	0.626	0.703	Fusobacteriaceae	0.136	0.014	0.005499 < Fusobacteriaceae <= 0.135935
14	Adenocarcinoma	0.902	0.098	0.626	0.703	Peptostreptococcales.Tissierellales	0.055	0.014	0.030036 < Peptostreptococcales.Tissierellales
20	Adenocarcinoma	0.932	0.141	0.644	0.657	Peptostreptococcaceae	0.001	-0.025	Peptostreptococcaceae <= 0.000792
20	Adenocarcinoma	0.932	0.141	0.644	0.657	Moraxellaceae	0.001	0.019	Moraxellaceae <= 0.000927
20	Adenocarcinoma	0.932	0.141	0.644	0.657	Bacillaceae	0.000	0.019	Bacillaceae <= 0.000217
20	Adenocarcinoma	0.932	0.141	0.644	0.657	Gemellaceae	0.000	-0.014	Gemellaceae <= 0.000438
20	Adenocarcinoma	0.932	0.141	0.644	0.657	Acidaminococcaceae	0.002	0.013	Acidaminococcaceae <= 0.00272
21	Adenocarcinoma	0.926	0.115	0.639	0.679	Pseudomonadaceae	0.004	-0.025	0.002447 < Pseudomonadaceae
21	Adenocarcinoma	0.926	0.115	0.639	0.679	Peptostreptococcales.Tissierellales	0.020	0.018	0.001179 < Peptostreptococcales.Tissierellales <= 0.030036
21	Adenocarcinoma	0.926	0.115	0.639	0.679	Acidaminococcaceae	0.000	0.015	Acidaminococcaceae <= 0.00272
21	Adenocarcinoma	0.926	0.115	0.639	0.679	Bacillaceae	0.000	0.016	Bacillaceae <= 0.000217
21	Adenocarcinoma	0.926	0.115	0.639	0.679	Gemellaceae	0.006	0.015	0.003566 < Gemellaceae
22	Adenocarcinoma	0.920	0.107	0.624	0.707	Moraxellaceae	0.000	0.019	Moraxellaceae <= 0.000927
22	Adenocarcinoma	0.920	0.107	0.624	0.707	Fusobacteriaceae	0.635	0.017	0.135935 < Fusobacteriaceae
22	Adenocarcinoma	0.920	0.107	0.624	0.707	Peptostreptococcales.Tissierellales	0.035	0.017	0.030036 < Peptostreptococcales.Tissierellales
22	Adenocarcinoma	0.920	0.107	0.624	0.707	Acidaminococcaceae	0.000	0.014	Acidaminococcaceae <= 0.00272
22	Adenocarcinoma	0.920	0.107	0.624	0.707	Bacillaceae	0.000	0.016	Bacillaceae <= 0.000217
23	Adenocarcinoma	0.840	0.105	0.658	0.636	Pseudomonadaceae	0.003	-0.023	0.002447 < Pseudomonadaceae
23	Adenocarcinoma	0.840	0.105	0.658	0.636	Bacillaceae	0.000	0.017	Bacillaceae <= 0.000217
23	Adenocarcinoma	0.840	0.105	0.658	0.636	Gemellaceae	0.000	-0.014	Gemellaceae <= 0.000438
23	Adenocarcinoma	0.840	0.105	0.658	0.636	Streptococcaceae	0.002	-0.015	Streptococcaceae <= 0.00166
23	Adenocarcinoma	0.840	0.105	0.658	0.636	Acidaminococcaceae	0.000	0.013	Acidaminococcaceae <= 0.00272
24	Adenocarcinoma	0.878	0.126	0.636	0.682	Pseudomonadaceae	0.003	-0.022	0.002447 < Pseudomonadaceae
24	Adenocarcinoma	0.878	0.126	0.636	0.682	Moraxellaceae	0.000	0.017	Moraxellaceae <= 0.000927
24	Adenocarcinoma	0.878	0.126	0.636	0.682	Bacillaceae	0.000	0.019	Bacillaceae <= 0.000217
24	Adenocarcinoma	0.878	0.126	0.636	0.682	Peptostreptococcales.Tissierellales	0.058	0.018	0.030036 < Peptostreptococcales.Tissierellales
24	Adenocarcinoma	0.878	0.126	0.636	0.682	Fusobacteriaceae	0.494	0.015	0.135935 < Fusobacteriaceae
25	Adenocarcinoma	0.876	0.145	0.642	0.675	Fusobacteriaceae	0.000	-0.031	Fusobacteriaceae <= 0.000517
25	Adenocarcinoma	0.876	0.145	0.642	0.675	Bacillaceae	0.000	0.019	Bacillaceae <= 0.000217
25	Adenocarcinoma	0.876	0.145	0.642	0.675	Peptostreptococcales.Tissierellales	0.039	0.016	0.030036 < Peptostreptococcales.Tissierellales
25	Adenocarcinoma	0.876	0.145	0.642	0.675	Acidaminococcaceae	0.002	0.013	Acidaminococcaceae <= 0.00272
25	Adenocarcinoma	0.876	0.145	0.642	0.675	Gemellaceae	0.013	0.016	0.003566 < Gemellaceae
26	Adenocarcinoma	0.828	0.119	0.646	0.651	Planococcaceae	0.002	-0.022	0.000978 < Planococcaceae
26	Adenocarcinoma	0.828	0.119	0.646	0.651	Tannerellaceae	0.150	-0.020	0.02489 < Tannerellaceae
26	Adenocarcinoma	0.828	0.119	0.646	0.651	Bacillaceae	0.000	0.018	Bacillaceae <= 0.000217
26	Adenocarcinoma	0.828	0.119	0.646	0.651	Acidaminococcaceae	0.000	0.014	Acidaminococcaceae <= 0.00272

26	Adenocarcinoma	0.828	0.119	0.646	0.651	Fusobacteriaceae	0.019	0.016	0.005499 < Fusobacteriaceae <= 0.135935
29	Adenocarcinoma	0.700	0.262	0.695	0.565	Peptostreptococcales.Tissierellales	0.000	-0.038	Peptostreptococcales.Tissierellales <= 0.000231
29	Adenocarcinoma	0.700	0.262	0.695	0.565	Sutterellaceae	0.019	-0.030	0.01055 < Sutterellaceae
29	Adenocarcinoma	0.700	0.262	0.695	0.565	Rikenellaceae	0.049	-0.023	0.02979 < Rikenellaceae
29	Adenocarcinoma	0.700	0.262	0.695	0.565	Moraxellaceae	0.033	-0.023	0.003850 < Moraxellaceae
29	Adenocarcinoma	0.700	0.262	0.695	0.565	Tannerellaceae	0.128	-0.018	0.02489 < Tannerellaceae
32	Adenocarcinoma	0.892	0.104	0.642	0.676	Moraxellaceae	0.012	-0.024	0.003850 < Moraxellaceae
32	Adenocarcinoma	0.892	0.104	0.642	0.676	Peptostreptococcales.Tissierellales	0.036	0.015	0.030036 < Peptostreptococcales.Tissierellales
32	Adenocarcinoma	0.892	0.104	0.642	0.676	Acidaminococcaceae	0.001	0.013	Acidaminococcaceae <= 0.00272
32	Adenocarcinoma	0.892	0.104	0.642	0.676	Fusobacteriaceae	0.243	0.015	0.135935 < Fusobacteriaceae
32	Adenocarcinoma	0.892	0.104	0.642	0.676	Chloroplast	0.000	0.014	Chloroplast <= 0.000151
33	Adenocarcinoma	0.806	0.173	0.654	0.623	Sutterellaceae	0.030	-0.030	0.01055 < Sutterellaceae
33	Adenocarcinoma	0.806	0.173	0.654	0.623	Rikenellaceae	0.051	-0.023	0.02979 < Rikenellaceae
33	Adenocarcinoma	0.806	0.173	0.654	0.623	Moraxellaceae	0.000	0.019	Moraxellaceae <= 0.000927
33	Adenocarcinoma	0.806	0.173	0.654	0.623	Bacillaceae	0.000	0.019	Bacillaceae <= 0.000217
33	Adenocarcinoma	0.806	0.173	0.654	0.623	Acidaminococcaceae	0.007	-0.015	0.00613 < Acidaminococcaceae
34	Adenocarcinoma	0.790	0.141	0.669	0.635	Planococcaceae	0.006	-0.025	0.000978 < Planococcaceae
34	Adenocarcinoma	0.790	0.141	0.669	0.635	Moraxellaceae	0.056	-0.025	0.003850 < Moraxellaceae
34	Adenocarcinoma	0.790	0.141	0.669	0.635	Peptostreptococcales.Tissierellales	0.030	0.018	0.001179 < Peptostreptococcales.Tissierellales <= 0.030036
34	Adenocarcinoma	0.790	0.141	0.669	0.635	Bacillaceae	0.009	-0.015	0.000217 < Bacillaceae
34	Adenocarcinoma	0.790	0.141	0.669	0.635	Gemellaceae	0.036	0.014	0.003566 < Gemellaceae
39	Adenocarcinoma	0.746	0.144	0.676	0.615	Planococcaceae	0.004	-0.024	0.000978 < Planococcaceae
39	Adenocarcinoma	0.746	0.144	0.676	0.615	Peptostreptococcaceae	0.000	-0.021	Peptostreptococcaceae <= 0.000792
39	Adenocarcinoma	0.746	0.144	0.676	0.615	Moraxellaceae	0.000	0.016	Moraxellaceae <= 0.000927
39	Adenocarcinoma	0.746	0.144	0.676	0.615	Gemellaceae	0.000	-0.016	Gemellaceae <= 0.000438
39	Adenocarcinoma	0.746	0.144	0.676	0.615	Acidaminococcaceae	0.020	-0.017	0.00613 < Acidaminococcaceae
40	Adenocarcinoma	0.862	0.106	0.650	0.660	Moraxellaceae	0.000	0.017	Moraxellaceae <= 0.000927
40	Adenocarcinoma	0.862	0.106	0.650	0.660	Tannerellaceae	0.036	-0.020	0.02489 < Tannerellaceae
40	Adenocarcinoma	0.862	0.106	0.650	0.660	Bacillaceae	0.001	-0.016	0.000217 < Bacillaceae
40	Adenocarcinoma	0.862	0.106	0.650	0.660	Acidaminococcaceae	0.000	0.014	Acidaminococcaceae <= 0.00272
40	Adenocarcinoma	0.862	0.106	0.650	0.660	Fusobacteriaceae	0.015	0.015	0.005499 < Fusobacteriaceae <= 0.135935
46	Adenocarcinoma	0.776	0.291	0.701	0.570	Peptostreptococcales.Tissierellales	0.000	-0.041	Peptostreptococcales.Tissierellales <= 0.000231
46	Adenocarcinoma	0.776	0.291	0.701	0.570	Fusobacteriaceae	0.000	-0.028	Fusobacteriaceae <= 0.000517
46	Adenocarcinoma	0.776	0.291	0.701	0.570	Peptostreptococcaceae	0.000	-0.024	Peptostreptococcaceae <= 0.000792
46	Adenocarcinoma	0.776	0.291	0.701	0.570	Rikenellaceae	0.031	-0.022	0.02979 < Rikenellaceae
46	Adenocarcinoma	0.776	0.291	0.701	0.570	Gemellaceae	0.000	-0.016	Gemellaceae <= 0.000438
47	Adenocarcinoma	0.878	0.082	0.644	0.670	Peptostreptococcales.Tissierellales	0.003	0.019	0.001179 < Peptostreptococcales.Tissierellales <= 0.030036
47	Adenocarcinoma	0.878	0.082	0.644	0.670	Bacillaceae	0.000	0.019	Bacillaceae <= 0.000217
47	Adenocarcinoma	0.878	0.082	0.644	0.670	Gemellaceae	0.000	-0.012	Gemellaceae <= 0.000438
47	Adenocarcinoma	0.878	0.082	0.644	0.670	Chloroplast	0.000	0.012	Chloroplast <= 0.000151
47	Adenocarcinoma	0.878	0.082	0.644	0.670	Akkermansiaceae	0.009	-0.012	3.00e-03 < Akkermansiaceae
48	Adenocarcinoma	0.868	0.260	0.664	0.610	Peptostreptococcales.Tissierellales	0.000	-0.042	Peptostreptococcales.Tissierellales <= 0.000231
48	Adenocarcinoma	0.868	0.260	0.664	0.610	Fusobacteriaceae	0.001	-0.029	Fusobacteriaceae <= 0.000517

48	Adenocarcinoma	0.868	0.260	0.664	0.610	Moraxellaceae	0.000	0.018	Moraxellaceae <= 0.000927
48	Adenocarcinoma	0.868	0.260	0.664	0.610	Acidaminococcaceae	0.010	-0.017	0.00613 < Acidaminococcaceae
48	Adenocarcinoma	0.868	0.260	0.664	0.610	Bacillaceae	0.000	0.017	Bacillaceae <= 0.000217
49	Adenocarcinoma	0.946	0.120	0.634	0.679	Peptostreptococcaceae	0.001	-0.020	Peptostreptococcaceae <= 0.000792
49	Adenocarcinoma	0.946	0.120	0.634	0.679	Moraxellaceae	0.000	0.019	Moraxellaceae <= 0.000927
49	Adenocarcinoma	0.946	0.120	0.634	0.679	Peptostreptococcales.Tissierellales	0.003	0.017	0.001179 < Peptostreptococcales.Tissierellales <= 0.030036
49	Adenocarcinoma	0.946	0.120	0.634	0.679	Planococcaceae	0.000	0.014	Planococcaceae <= 0.000176
49	Adenocarcinoma	0.946	0.120	0.634	0.679	Bacillaceae	0.000	0.015	Bacillaceae <= 0.000217
50	Adenocarcinoma	0.908	0.129	0.641	0.664	Moraxellaceae	0.000	0.019	Moraxellaceae <= 0.000927
50	Adenocarcinoma	0.908	0.129	0.641	0.664	Peptostreptococcales.Tissierellales	0.009	0.021	0.001179 < Peptostreptococcales.Tissierellales <= 0.030036
50	Adenocarcinoma	0.908	0.129	0.641	0.664	Bacillaceae	0.000	0.018	Bacillaceae <= 0.000217
50	Adenocarcinoma	0.908	0.129	0.641	0.664	Acidaminococcaceae	0.011	-0.017	0.00613 < Acidaminococcaceae
50	Adenocarcinoma	0.908	0.129	0.641	0.664	Desulfovibrionaceae	0.003	-0.017	0.003165 < Desulfovibrionaceae

Table ML-D-5: The LIME Result for Test Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
1	Adenocarcinoma	0.760	0.111	0.669	0.627	Moraxellaceae	0.000	0.019	Moraxellaceae <= 0.000927
1	Adenocarcinoma	0.760	0.111	0.669	0.627	Gemellaceae	0.000	-0.015	Gemellaceae <= 0.000438
1	Adenocarcinoma	0.760	0.111	0.669	0.627	Bacillaceae	0.000	-0.016	0.000217 < Bacillaceae
1	Adenocarcinoma	0.760	0.111	0.669	0.627	Acidaminococcaceae	0.018	-0.015	0.00613 < Acidaminococcaceae
1	Adenocarcinoma	0.760	0.111	0.669	0.627	Desulfovibrionaceae	0.007	-0.015	0.003165 < Desulfovibrionaceae
4	Adenocarcinoma	0.584	0.311	0.699	0.558	Peptostreptococcales.Tissierellales	0.000	-0.039	Peptostreptococcales.Tissierellales <= 0.000231
4	Adenocarcinoma	0.584	0.311	0.699	0.558	Sutterellaceae	0.023	-0.031	0.01055 < Sutterellaceae
4	Adenocarcinoma	0.584	0.311	0.699	0.558	Fusobacteriaceae	0.000	-0.027	Fusobacteriaceae <= 0.000517
4	Adenocarcinoma	0.584	0.311	0.699	0.558	Moraxellaceae	0.011	-0.024	0.003850 < Moraxellaceae
4	Adenocarcinoma	0.584	0.311	0.699	0.558	Peptostreptococcaceae	0.000	-0.020	Peptostreptococcaceae <= 0.000792
5	Adenocarcinoma	0.698	0.183	0.662	0.624	Fusobacteriaceae	0.000	-0.030	Fusobacteriaceae <= 0.000517
5	Adenocarcinoma	0.698	0.183	0.662	0.624	Moraxellaceae	0.006	-0.025	0.003850 < Moraxellaceae
5	Adenocarcinoma	0.698	0.183	0.662	0.624	Pseudomonadaceae	0.003	-0.021	0.002447 < Pseudomonadaceae
5	Adenocarcinoma	0.698	0.183	0.662	0.624	Peptostreptococcales.Tissierellales	0.001	0.019	0.001179 < Peptostreptococcales.Tissierellales <= 0.030036
5	Adenocarcinoma	0.698	0.183	0.662	0.624	Bacillaceae	0.000	0.018	Bacillaceae <= 0.000217
8	Adenocarcinoma	0.706	0.280	0.680	0.589	Peptostreptococcales.Tissierellales	0.000	-0.039	Peptostreptococcales.Tissierellales <= 0.000231
8	Adenocarcinoma	0.706	0.280	0.680	0.589	Fusobacteriaceae	0.000	-0.028	Fusobacteriaceae <= 0.000517
8	Adenocarcinoma	0.706	0.280	0.680	0.589	Sutterellaceae	0.033	-0.029	0.01055 < Sutterellaceae
8	Adenocarcinoma	0.706	0.280	0.680	0.589	Bacillaceae	0.000	0.019	Bacillaceae <= 0.000217
8	Adenocarcinoma	0.706	0.280	0.680	0.589	Gemellaceae	0.000	-0.015	Gemellaceae <= 0.000438
9	Adenoma	0.542	0.310	0.300	0.445	Peptostreptococcales.Tissierellales	0.000	0.037	Peptostreptococcales.Tissierellales <= 0.000231
9	Adenoma	0.542	0.310	0.300	0.445	Fusobacteriaceae	0.000	0.030	Fusobacteriaceae <= 0.000517
9	Adenoma	0.542	0.310	0.300	0.445	Sutterellaceae	0.016	0.029	0.01055 < Sutterellaceae
9	Adenoma	0.542	0.310	0.300	0.445	Planococcaceae	0.005	0.026	0.000978 < Planococcaceae
9	Adenoma	0.542	0.310	0.300	0.445	Peptostreptococcaceae	0.000	0.022	Peptostreptococcaceae <= 0.000792
12	Adenocarcinoma	0.674	0.121	0.660	0.628	Moraxellaceae	0.027	-0.023	0.003850 < Moraxellaceae

12	Adenocarcinoma	0.674	0.121	0.660	0.628	Pseudomonadaceae	0.007	-0.020	0.002447 < Pseudomonadaceae
12	Adenocarcinoma	0.674	0.121	0.660	0.628	Bacillaceae	0.001	-0.017	0.000217 < Bacillaceae
12	Adenocarcinoma	0.674	0.121	0.660	0.628	Acidaminococcaceae	0.000	0.014	Acidaminococcaceae <= 0.00272
12	Adenocarcinoma	0.674	0.121	0.660	0.628	Planococcaceae	0.000	0.013	Planococcaceae <= 0.000176
13	Adenocarcinoma	0.888	0.111	0.635	0.688	Moraxellaceae	0.000	0.017	Moraxellaceae <= 0.000927
13	Adenocarcinoma	0.888	0.111	0.635	0.688	Bacillaceae	0.000	0.018	Bacillaceae <= 0.000217
13	Adenocarcinoma	0.888	0.111	0.635	0.688	Fusobacteriaceae	0.440	0.018	0.135935 < Fusobacteriaceae
13	Adenocarcinoma	0.888	0.111	0.635	0.688	Acidaminococcaceae	0.001	0.014	Acidaminococcaceae <= 0.00272
13	Adenocarcinoma	0.888	0.111	0.635	0.688	Gemellaceae	0.000	-0.014	Gemellaceae <= 0.000438
15	Adenocarcinoma	0.876	0.104	0.649	0.662	Planococcaceae	0.002	-0.020	0.000978 < Planococcaceae
15	Adenocarcinoma	0.876	0.104	0.649	0.662	Bacillaceae	0.000	0.019	Bacillaceae <= 0.000217
15	Adenocarcinoma	0.876	0.104	0.649	0.662	Desulfovibrionaceae	0.006	-0.017	0.003165 < Desulfovibrionaceae
15	Adenocarcinoma	0.876	0.104	0.649	0.662	Peptostreptococcales.Tissierellales	0.013	0.015	0.001179 < Peptostreptococcales.Tissierellales <= 0.030036
15	Adenocarcinoma	0.876	0.104	0.649	0.662	Fusobacteriaceae	0.057	0.015	0.005499 < Fusobacteriaceae <= 0.135935
16	Adenocarcinoma	0.878	0.105	0.649	0.661	Moraxellaceae	0.000	0.019	Moraxellaceae <= 0.000927
16	Adenocarcinoma	0.878	0.105	0.649	0.661	Desulfovibrionaceae	0.010	-0.021	0.003165 < Desulfovibrionaceae
16	Adenocarcinoma	0.878	0.105	0.649	0.661	Peptostreptococcales.Tissierellales	0.054	0.016	0.030036 < Peptostreptococcales.Tissierellales
16	Adenocarcinoma	0.878	0.105	0.649	0.661	Planococcaceae	0.000	0.012	Planococcaceae <= 0.000176
16	Adenocarcinoma	0.878	0.105	0.649	0.661	Chloroplast	0.000	-0.014	0.000151 < Chloroplast
17	Adenocarcinoma	0.622	0.313	0.699	0.559	Peptostreptococcales.Tissierellales	0.000	-0.038	Peptostreptococcales.Tissierellales <= 0.000231
17	Adenocarcinoma	0.622	0.313	0.699	0.559	Fusobacteriaceae	0.000	-0.030	Fusobacteriaceae <= 0.000517
17	Adenocarcinoma	0.622	0.313	0.699	0.559	Sutterellaceae	0.013	-0.031	0.01055 < Sutterellaceae
17	Adenocarcinoma	0.622	0.313	0.699	0.559	Peptostreptococcaceae	0.001	-0.022	Peptostreptococcaceae <= 0.000792
17	Adenocarcinoma	0.622	0.313	0.699	0.559	Rikenellaceae	0.082	-0.020	0.02979 < Rikenellaceae
18	Adenocarcinoma	0.726	0.190	0.694	0.584	Fusobacteriaceae	0.000	-0.028	Fusobacteriaceae <= 0.000517
18	Adenocarcinoma	0.726	0.190	0.694	0.584	Sutterellaceae	0.024	-0.029	0.01055 < Sutterellaceae
18	Adenocarcinoma	0.726	0.190	0.694	0.584	Tannerellaceae	0.029	-0.019	0.02489 < Tannerellaceae
18	Adenocarcinoma	0.726	0.190	0.694	0.584	Acidaminococcaceae	0.029	-0.019	0.00613 < Acidaminococcaceae
18	Adenocarcinoma	0.726	0.190	0.694	0.584	Gemellaceae	0.000	-0.014	Gemellaceae <= 0.000438
19	Adenocarcinoma	0.664	0.191	0.673	0.608	Peptostreptococcales.Tissierellales	0.000	-0.039	Peptostreptococcales.Tissierellales <= 0.000231
19	Adenocarcinoma	0.664	0.191	0.673	0.608	Bacillaceae	0.000	0.019	Bacillaceae <= 0.000217
19	Adenocarcinoma	0.664	0.191	0.673	0.608	Gemellaceae	0.000	-0.015	Gemellaceae <= 0.000438
19	Adenocarcinoma	0.664	0.191	0.673	0.608	Acidaminococcaceae	0.013	-0.016	0.00613 < Acidaminococcaceae
19	Adenocarcinoma	0.664	0.191	0.673	0.608	Chloroplast	0.000	-0.013	0.000151 < Chloroplast
27	Adenocarcinoma	0.870	0.133	0.635	0.685	Rikenellaceae	0.085	-0.023	0.02979 < Rikenellaceae
27	Adenocarcinoma	0.870	0.133	0.635	0.685	Moraxellaceae	0.000	0.019	Moraxellaceae <= 0.000927
27	Adenocarcinoma	0.870	0.133	0.635	0.685	Bacillaceae	0.000	0.018	Bacillaceae <= 0.000217
27	Adenocarcinoma	0.870	0.133	0.635	0.685	Peptostreptococcales.Tissierellales	0.003	0.018	0.001179 < Peptostreptococcales.Tissierellales <= 0.030036
27	Adenocarcinoma	0.870	0.133	0.635	0.685	Fusobacteriaceae	0.094	0.018	0.005499 < Fusobacteriaceae <= 0.135935
28	Adenocarcinoma	0.776	0.119	0.658	0.635	Moraxellaceae	0.006	-0.024	0.003850 < Moraxellaceae
28	Adenocarcinoma	0.776	0.119	0.658	0.635	Peptostreptococcales.Tissierellales	0.018	0.018	0.001179 < Peptostreptococcales.Tissierellales <= 0.030036
28	Adenocarcinoma	0.776	0.119	0.658	0.635	Desulfovibrionaceae	0.005	-0.019	0.003165 < Desulfovibrionaceae
28	Adenocarcinoma	0.776	0.119	0.658	0.635	Bacillaceae	0.000	0.017	Bacillaceae <= 0.000217

28	Adenocarcinoma	0.776	0.119	0.658	0.635	Acidaminococcaceae	0.007	-0.017	0.00613 < Acidaminococcaceae
30	Adenocarcinoma	0.742	0.165	0.679	0.610	Moraxellaceae	0.018	-0.027	0.003850 < Moraxellaceae
30	Adenocarcinoma	0.742	0.165	0.679	0.610	Rikenellaceae	0.031	-0.024	0.02979 < Rikenellaceae
30	Adenocarcinoma	0.742	0.165	0.679	0.610	Desulfovibrionaceae	0.004	-0.019	0.003165 < Desulfovibrionaceae
30	Adenocarcinoma	0.742	0.165	0.679	0.610	Peptostreptococcales.Tissierellales	0.016	0.018	0.001179 < Peptostreptococcales.Tissierellales <= 0.030036
30	Adenocarcinoma	0.742	0.165	0.679	0.610	Bacillaceae	0.000	-0.017	0.000217 < Bacillaceae
31	Adenocarcinoma	0.750	0.187	0.663	0.626	Peptostreptococcales.Tissierellales	0.000	-0.037	Peptostreptococcales.Tissierellales <= 0.000231
31	Adenocarcinoma	0.750	0.187	0.663	0.626	Gemellaceae	0.000	-0.017	Gemellaceae <= 0.000438
31	Adenocarcinoma	0.750	0.187	0.663	0.626	Moraxellaceae	0.000	0.016	Moraxellaceae <= 0.000927
31	Adenocarcinoma	0.750	0.187	0.663	0.626	Bacillaceae	0.000	0.016	Bacillaceae <= 0.000217
31	Adenocarcinoma	0.750	0.187	0.663	0.626	Desulfovibrionaceae	0.010	-0.016	0.003165 < Desulfovibrionaceae
35	Adenocarcinoma	0.730	0.167	0.661	0.626	Sutterellaceae	0.028	-0.031	0.01055 < Sutterellaceae
35	Adenocarcinoma	0.730	0.167	0.661	0.626	Planococcaceae	0.001	-0.023	0.000978 < Planococcaceae
35	Adenocarcinoma	0.730	0.167	0.661	0.626	Bacillaceae	0.000	0.018	Bacillaceae <= 0.000217
35	Adenocarcinoma	0.730	0.167	0.661	0.626	Peptostreptococcales.Tissierellales	0.004	0.017	0.001179 < Peptostreptococcales.Tissierellales <= 0.030036
35	Adenocarcinoma	0.730	0.167	0.661	0.626	Desulfovibrionaceae	0.005	-0.016	0.003165 < Desulfovibrionaceae
36	Adenocarcinoma	0.796	0.122	0.646	0.660	Desulfovibrionaceae	0.004	-0.021	0.003165 < Desulfovibrionaceae
36	Adenocarcinoma	0.796	0.122	0.646	0.660	Moraxellaceae	0.000	0.017	Moraxellaceae <= 0.000927
36	Adenocarcinoma	0.796	0.122	0.646	0.660	Streptococcaceae	0.001	-0.016	Streptococcaceae <= 0.00166
36	Adenocarcinoma	0.796	0.122	0.646	0.660	Bacillaceae	0.000	0.017	Bacillaceae <= 0.000217
36	Adenocarcinoma	0.796	0.122	0.646	0.660	Peptostreptococcales.Tissierellales	0.007	0.017	0.001179 < Peptostreptococcales.Tissierellales <= 0.030036
37	Adenocarcinoma	0.776	0.126	0.619	0.710	Moraxellaceae	0.000	0.019	Moraxellaceae <= 0.000927
37	Adenocarcinoma	0.776	0.126	0.619	0.710	Bacillaceae	0.000	0.020	Bacillaceae <= 0.000217
37	Adenocarcinoma	0.776	0.126	0.619	0.710	Peptostreptococcales.Tissierellales	0.004	0.019	0.001179 < Peptostreptococcales.Tissierellales <= 0.030036
37	Adenocarcinoma	0.776	0.126	0.619	0.710	Fusobacteriaceae	0.167	0.018	0.135935 < Fusobacteriaceae
37	Adenocarcinoma	0.776	0.126	0.619	0.710	Acidaminococcaceae	0.000	0.015	Acidaminococcaceae <= 0.00272
38	Adenocarcinoma	0.576	0.218	0.692	0.568	Sutterellaceae	0.011	-0.030	0.01055 < Sutterellaceae
38	Adenocarcinoma	0.576	0.218	0.692	0.568	Peptostreptococcaceae	0.000	-0.024	Peptostreptococcaceae <= 0.000792
38	Adenocarcinoma	0.576	0.218	0.692	0.568	Pseudomonadaceae	0.004	-0.024	0.002447 < Pseudomonadaceae
38	Adenocarcinoma	0.576	0.218	0.692	0.568	Planococcaceae	0.002	-0.023	0.000978 < Planococcaceae
38	Adenocarcinoma	0.576	0.218	0.692	0.568	Moraxellaceae	0.007	-0.023	0.003850 < Moraxellaceae
41	Adenocarcinoma	0.782	0.121	0.666	0.645	Desulfovibrionaceae	0.004	-0.022	0.003165 < Desulfovibrionaceae
41	Adenocarcinoma	0.782	0.121	0.666	0.645	Moraxellaceae	0.000	0.018	Moraxellaceae <= 0.000927
41	Adenocarcinoma	0.782	0.121	0.666	0.645	Bacillaceae	0.001	-0.018	0.000217 < Bacillaceae
41	Adenocarcinoma	0.782	0.121	0.666	0.645	Gemellaceae	0.000	-0.014	Gemellaceae <= 0.000438
41	Adenocarcinoma	0.782	0.121	0.666	0.645	Fusobacteriaceae	0.193	0.015	0.135935 < Fusobacteriaceae
42	Adenocarcinoma	0.660	0.272	0.680	0.589	Peptostreptococcales.Tissierellales	0.000	-0.039	Peptostreptococcales.Tissierellales <= 0.000231
42	Adenocarcinoma	0.660	0.272	0.680	0.589	Sutterellaceae	0.016	-0.030	0.01055 < Sutterellaceae
42	Adenocarcinoma	0.660	0.272	0.680	0.589	Peptostreptococcaceae	0.001	-0.021	Peptostreptococcaceae <= 0.000792
42	Adenocarcinoma	0.660	0.272	0.680	0.589	Moraxellaceae	0.000	0.018	Moraxellaceae <= 0.000927
42	Adenocarcinoma	0.660	0.272	0.680	0.589	Bacillaceae	0.000	-0.019	0.000217 < Bacillaceae
43	Adenocarcinoma	0.606	0.264	0.698	0.575	Peptostreptococcales.Tissierellales	0.000	-0.040	Peptostreptococcales.Tissierellales <= 0.000231
43	Adenocarcinoma	0.606	0.264	0.698	0.575	Fusobacteriaceae	0.000	-0.029	Fusobacteriaceae <= 0.000517

43	Adenocarcinoma	0.606	0.264	0.698	0.575	Peptostreptococcaceae	0.000	-0.021	Peptostreptococcaceae <= 0.000792
43	Adenocarcinoma	0.606	0.264	0.698	0.575	Tannerellaceae	0.045	-0.018	0.02489 < Tannerellaceae
43	Adenocarcinoma	0.606	0.264	0.698	0.575	Gemellaceae	0.000	-0.016	Gemellaceae <= 0.000438
44	Adenocarcinoma	0.748	0.186	0.680	0.608	Fusobacteriaceae	0.000	-0.030	Fusobacteriaceae <= 0.000517
44	Adenocarcinoma	0.748	0.186	0.680	0.608	Peptostreptococcaceae	0.000	-0.023	Peptostreptococcaceae <= 0.000792
44	Adenocarcinoma	0.748	0.186	0.680	0.608	Planococcaceae	0.001	-0.021	0.000978 < Planococcaceae
44	Adenocarcinoma	0.748	0.186	0.680	0.608	Moraxellaceae	0.000	0.017	Moraxellaceae <= 0.000927
44	Adenocarcinoma	0.748	0.186	0.680	0.608	Gemellaceae	0.000	-0.015	Gemellaceae <= 0.000438
45	Adenocarcinoma	0.680	0.179	0.647	0.648	Sutterellaceae	0.014	-0.031	0.01055 < Sutterellaceae
45	Adenocarcinoma	0.680	0.179	0.647	0.648	Rikenellaceae	0.051	-0.023	0.02979 < Rikenellaceae
45	Adenocarcinoma	0.680	0.179	0.647	0.648	Moraxellaceae	0.000	0.019	Moraxellaceae <= 0.000927
45	Adenocarcinoma	0.680	0.179	0.647	0.648	Bacillaceae	0.000	0.018	Bacillaceae <= 0.000217
45	Adenocarcinoma	0.680	0.179	0.647	0.648	Peptostreptococcales.Tissierellales	0.008	0.017	0.001179 < Peptostreptococcales.Tissierellales <= 0.030036

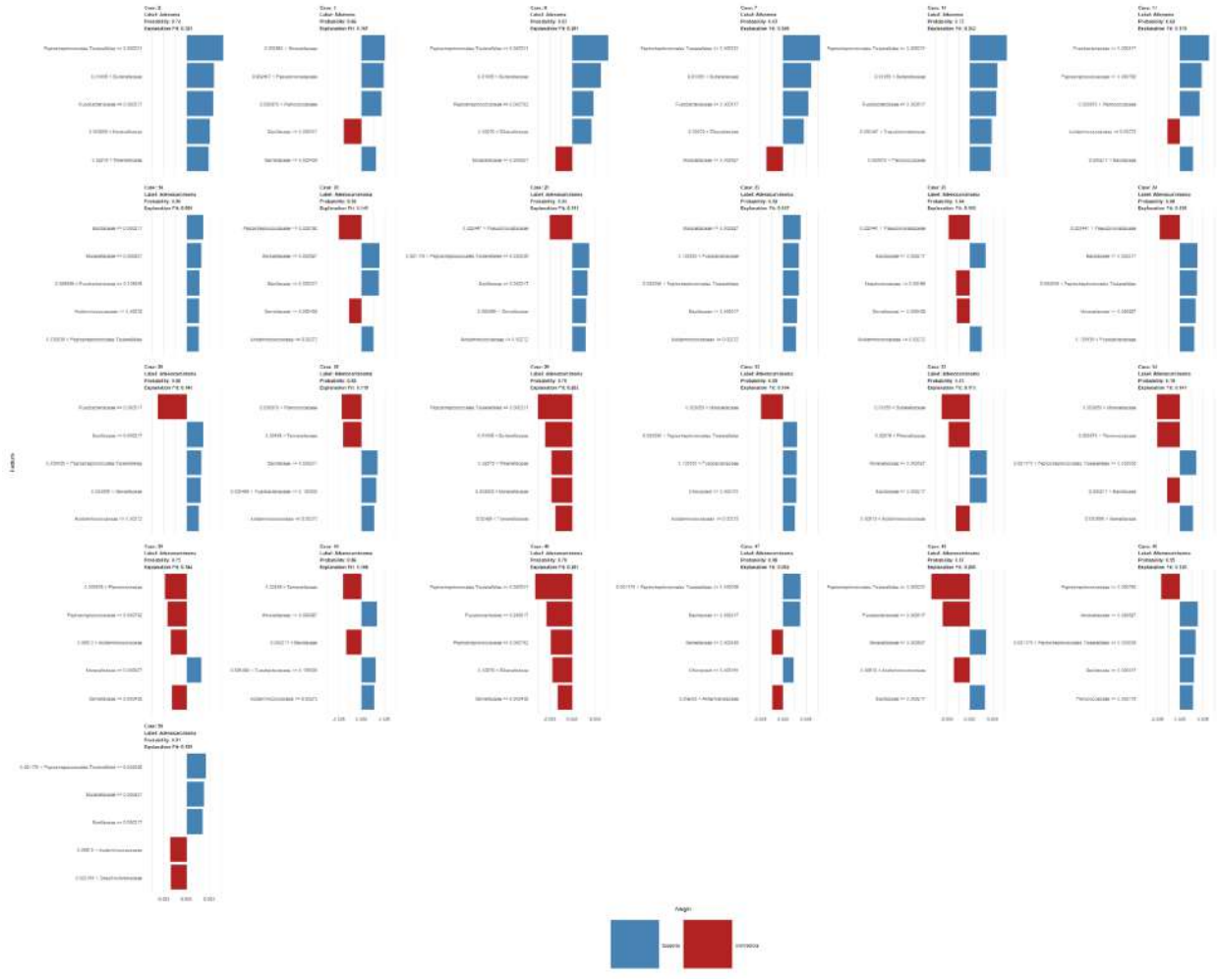


Figure ML1-D-6: The LIME Result in Training dataset for Family in Biopsy

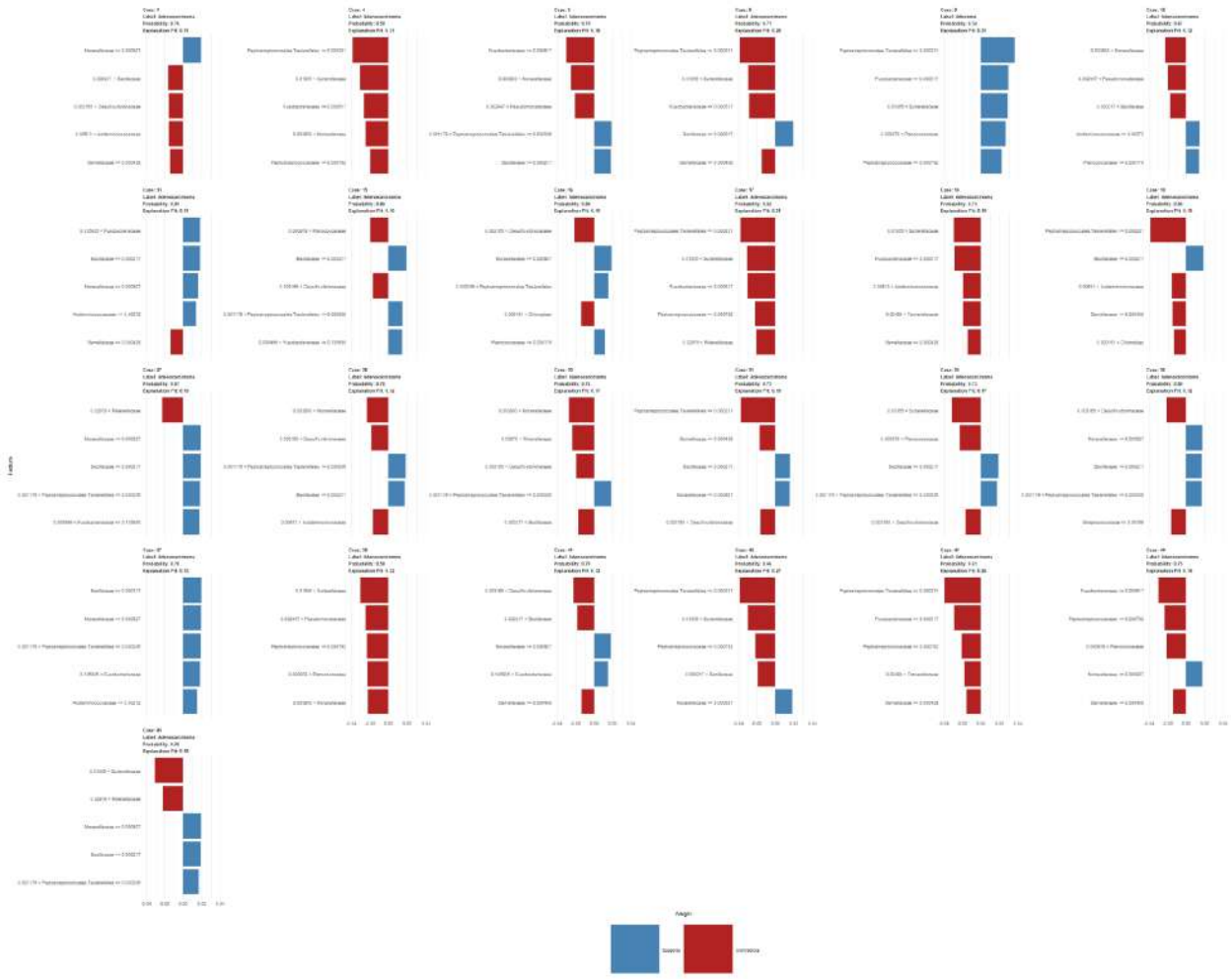


Figure ML1-D-7: The LIME Result in Test dataset for Family in Biopsy

Table ML1-D-6: The Aggregated SHAP Values

Variable	Training Dataset				Testing Dataset			
	contribution	position	sign	cumulative	contribution	position	sign	cumulative
intercept	0.7309	114	X	0.731	0.73088	114	X	0.731
Yersiniaceae	-0.0001	113	-1	0.731	0.002243	113	1	0.733
Bacteroidetes_BD2.2	0.0000	112	1	0.731	0	111	1	0.733
Neisseriaceae	0.0000	111	1	0.731	-0.00048	110	-1	0.733
Ruminococcaceae	0.0000	110	1	0.731	-0.00049	81	-1	0.731
Propionibacteriaceae	0.0000	109	1	0.731	0.000132	109	1	0.733
Rikenellaceae	0.0000	108	1	0.731	0.00054	75	1	0.734
IMCC26256	0.0000	107	1	0.731	0	108	1	0.733
Bifidobacteriaceae	0.0000	106	-1	0.731	0.000333	105	1	0.733
Tannerellaceae	0.0000	105	-1	0.731	0.001939	103	1	0.735
KD4.96	0.0000	104	1	0.731	0	104	1	0.733
UCG.010	0.0000	103	-1	0.731	-0.00014	99	-1	0.733
Muribaculaceae	0.0000	102	-1	0.731	0.000991	106	1	0.733
Morganellaceae	0.0000	101	1	0.731	0	112	1	0.733
Prevotellaceae	0.0000	100	-1	0.731	5.7E-05	101	1	0.733
Chthoniobacteraceae	0.0000	99	1	0.731	0.000304	100	1	0.733
Victivallaceae	0.0000	98	1	0.731	6.64E-05	98	1	0.733
Gastranaerophilales	0.0000	97	1	0.731	1.28E-05	97	1	0.733
Spirochaetaceae	0.0000	96	1	0.731	-3.3E-05	57	-1	0.737
Defluviitaleaceae	0.0000	95	1	0.731	0	96	1	0.733
Woeseiaceae	0.0000	94	1	0.731	0	95	1	0.733
Desulfobacteraceae	0.0000	93	1	0.731	0	92	1	0.731
Alteromonadaceae	0.0000	92	-1	0.731	-0.00094	107	-1	0.732
Melioribacteraceae	0.0000	91	1	0.731	0	32	1	0.722
Pseudoalteromonadaceae	0.0000	90	1	0.731	0.001889	91	1	0.733
Sphingomonadaceae	0.0000	89	-1	0.731	-8.9E-05	33	-1	0.722
Barnesiellaceae	0.0000	88	-1	0.731	0.001142	80	1	0.732
Comamonadaceae	0.0001	87	1	0.731	-0.0019	93	-1	0.731
Oscillospirales	0.0000	86	1	0.731	-0.00017	79	-1	0.732
Staphylococcaceae	0.0001	85	1	0.731	-0.00067	89	-1	0.732
Clostridia_UCG.014	0.0000	84	-1	0.731	-0.00022	88	-1	0.732
Hungateiclostridiaceae	0.0000	83	-1	0.731	-0.00012	67	-1	0.740
Desulfocapsaceae	0.0000	82	-1	0.731	-6E-05	87	-1	0.732
Bacteroidaceae	0.0000	81	-1	0.731	0.000328	85	1	0.731
Micrococcaceae	0.0000	80	1	0.731	-0.00058	54	-1	0.737
Flavobacteriaceae	0.0000	79	-1	0.731	0.001446	83	1	0.732
Nocardioidaceae	0.0000	78	1	0.731	-4.8E-05	78	-1	0.732
Xanthobacteraceae	0.0000	77	1	0.731	0.001585	77	1	0.734
Planococcaceae	0.0000	76	-1	0.731	0.004607	73	1	0.739

Solirubrobacteraceae	0.0000	75	1	0.731	0.000651	72	1	0.740
Caulobacteraceae	0.0000	74	1	0.731	0.000165	71	1	0.740
Peptococcaceae	0.0001	73	1	0.731	-0.00023	70	-1	0.740
Xanthomonadaceae	0.0000	72	1	0.731	-0.00247	102	-1	0.733
Aeromonadaceae	0.0000	71	-1	0.731	-0.00033	82	-1	0.732
Synergistaceae	0.0000	70	-1	0.731	-8.3E-05	94	-1	0.733
Coriobacteriales_Incertae_Sedis	0.0000	69	-1	0.731	0.000214	40	1	0.731
Desulfovibrionaceae	-0.0001	68	-1	0.731	-0.00314	59	-1	0.737
Burkholderiaceae	0.0000	67	1	0.731	0	61	1	0.739
Eubacteriaceae	0.0000	66	-1	0.731	2.78E-05	68	1	0.740
Rhizobiaceae	0.0000	65	1	0.731	0.000651	28	1	0.722
Saccharimonadaceae	0.0000	64	-1	0.731	-0.00043	65	-1	0.739
vadinBE97	0.0000	63	1	0.731	0	90	1	0.733
Atopobiaceae	0.0000	62	-1	0.731	-5.6E-05	64	-1	0.739
Bacillaceae	0.0000	61	1	0.731	0.000311	63	1	0.739
Coriobacteriaceae	0.0000	60	1	0.731	-0.00034	62	-1	0.739
Erysipelotrichaceae	0.0000	59	-1	0.731	-0.00068	49	-1	0.733
uncultured	0.0000	58	1	0.731	0.001361	43	1	0.730
Porphyromonadaceae	0.0000	57	-1	0.731	0.000297	51	1	0.733
Succinivibrionaceae	0.0000	56	1	0.731	-2.8E-05	35	-1	0.722
Christensenellaceae	0.0000	55	-1	0.731	0.000714	74	1	0.734
Clostridiaceae	0.0000	54	-1	0.731	0.001218	60	1	0.740
Mycobacteriaceae	0.0000	53	1	0.731	0	84	1	0.731
Erwiniaceae	0.0000	52	-1	0.731	-0.0001	58	-1	0.737
Clostridia_vadinBB60_group	0.0000	51	1	0.731	-0.00012	56	-1	0.737
Actinomycetaceae	0.0000	50	1	0.731	0.000254	55	1	0.737
Campylobacteraceae	0.0000	49	1	0.731	-0.00043	76	-1	0.733
Selenomonadaceae	0.0000	48	-1	0.731	0.000533	50	1	0.734
Gemellaceae	0.0000	47	1	0.731	-0.00172	86	-1	0.730
Pasteurellaceae	0.0000	46	-1	0.731	0.000232	48	1	0.733
Peptostreptococcaceae	0.0000	45	-1	0.731	-0.00369	53	-1	0.733
Tumor	0.0000	44	1	0.731	0	22	1	0.724
Peptostreptococcales.Tissierellales	0.0001	43	1	0.731	-0.00479	46	-1	0.728
Streptococcaceae	0.0000	42	1	0.731	-7.3E-05	45	-1	0.728
Enterobacteriaceae	0.0000	41	-1	0.731	0.000705	25	1	0.724
Marinifilaceae	0.0000	40	-1	0.731	0.000685	44	1	0.729
Enterococcaceae	0.0000	39	1	0.731	0.00014	42	1	0.730
Veillonellaceae	0.0000	38	1	0.731	0.000377	41	1	0.731
Fusobacteriaceae	0.0000	37	-1	0.731	-0.00336	39	-1	0.728
Moraxellaceae	-0.0001	36	-1	0.731	-0.00232	38	-1	0.725
Sutterellaceae	0.0000	35	1	0.731	-0.00269	37	-1	0.723
Corynebacteriaceae	0.0000	34	1	0.731	-0.00018	36	-1	0.722
Izemoplasmatales	0.0000	33	1	0.731	0	47	1	0.733

Gender	0.0000	32	1	0.731	0	34	1	0.722
X.Eubacterium._coprostanoligenes_group	0.0000	31	-1	0.731	-8.9E-05	52	-1	0.733
Methanobacteriaceae	0.0000	30	-1	0.731	7.07E-05	31	1	0.722
Oscillospiraceae	0.0000	29	1	0.731	8.58E-05	69	1	0.740
Rhodobacteraceae	0.0000	28	-1	0.731	-0.00011	66	-1	0.740
Carnobacteriaceae	0.0000	27	1	0.731	-0.00058	30	-1	0.722
Desulfosarcinaceae	0.0000	26	1	0.731	-4.9E-05	29	-1	0.722
X.Clostridium._methylpentosum_group	0.0000	25	-1	0.731	0.00052	27	1	0.723
Lachnospiraceae	0.0000	24	1	0.731	0.000257	26	1	0.723
Beijerinckiaceae	0.0000	23	1	0.731	0.000152	24	1	0.724
Oxalobacteraceae	0.0000	22	-1	0.731	-8.3E-05	23	-1	0.724
Acidaminococcaceae	0.0000	21	1	0.731	-0.00462	21	-1	0.719
Acidothermaceae	0.0000	20	1	0.731	0	20	1	0.719
Akkermansiaceae	-0.0001	19	-1	0.731	0.001147	19	1	0.721
Anaerovoracaceae	0.0000	18	-1	0.731	-0.00034	18	-1	0.720
Butyricicoccaceae	0.0000	17	1	0.731	0.001554	17	1	0.722
Chloroplast	0.0000	16	-1	0.731	0.000217	16	1	0.722
DTU014	0.0000	15	1	0.731	0	15	1	0.722
Eggerthellaceae	0.0000	14	1	0.731	-0.00051	14	-1	0.721
Erysipelatoclostridiaceae	0.0000	13	-1	0.731	-0.0025	13	-1	0.719
Lactobacillaceae	0.0000	12	1	0.731	-0.00038	12	-1	0.719
Leptotrichiaceae	0.0000	11	-1	0.731	-0.0005	11	-1	0.718
Leuconostocaceae	0.0000	10	1	0.731	0	10	1	0.718
Microbacteriaceae	0.0000	9	1	0.731	-0.00229	9	-1	0.716
Monoglobaceae	0.0000	8	1	0.731	0.000784	8	1	0.717
Nocardiaceae	0.0000	7	1	0.731	8.91E-05	7	1	0.717
Orbaceae	0.0000	6	1	0.731	0	6	1	0.717
Paenibacillaceae	0.0000	5	-1	0.731	-9.3E-05	5	-1	0.717
Pseudomonadaceae	-0.0001	4	-1	0.731	0.004751	4	1	0.721
Weeksellaceae	0.0000	3	1	0.731	-0.00029	3	-1	0.721
X67.14	0.0000	2	1	0.731	0	2	1	0.721
prediction	0.7309	1	X	0.731	0.72104	1	X	0.721

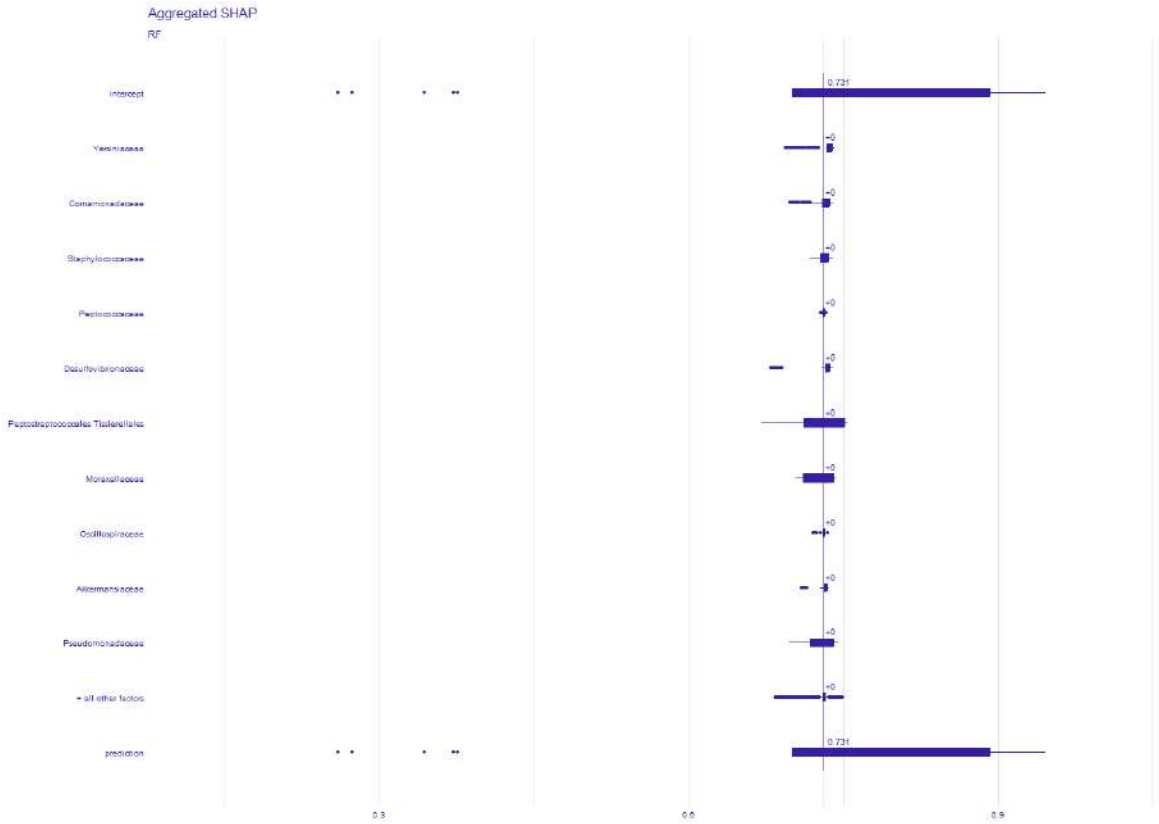


Figure ML1-D-8: The Aggregated SHAP Value in Training dataset for Family in Biopsy

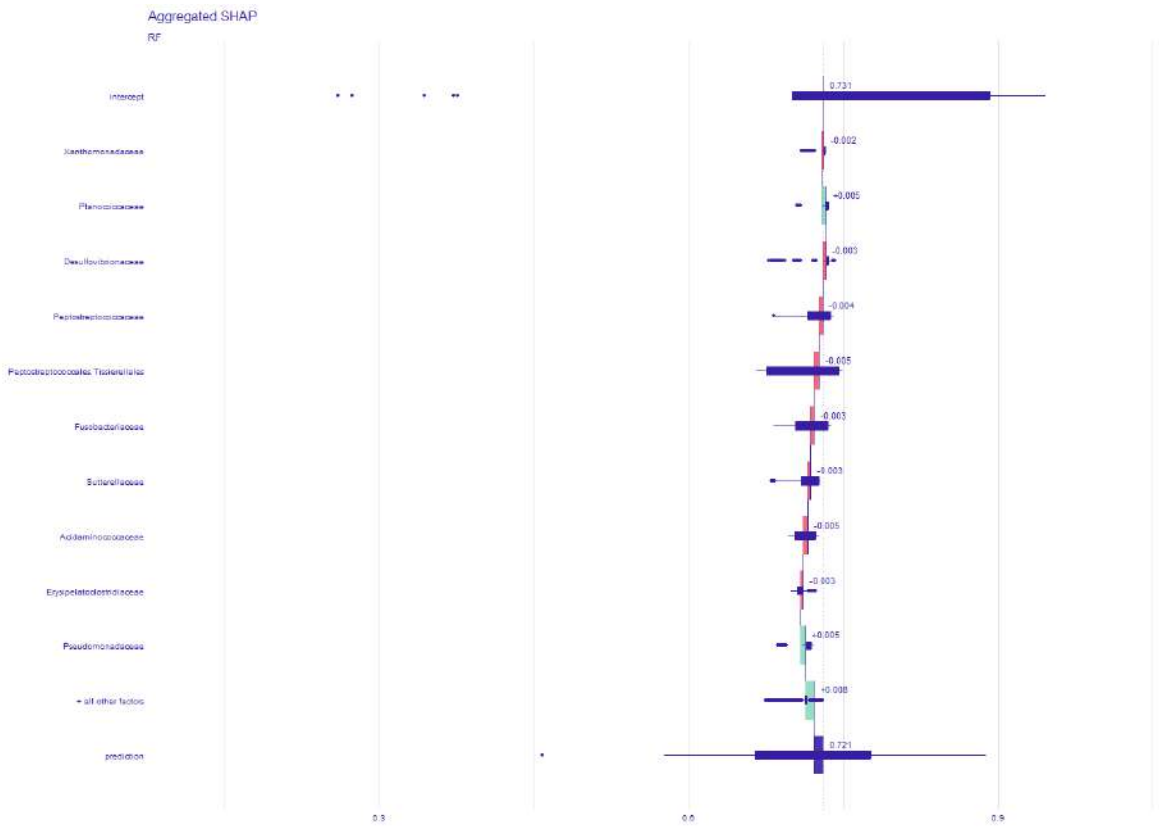


Figure ML1-D-9: The Aggregated SHAP Value in Test dataset for Family in Biopsy

E- Genus

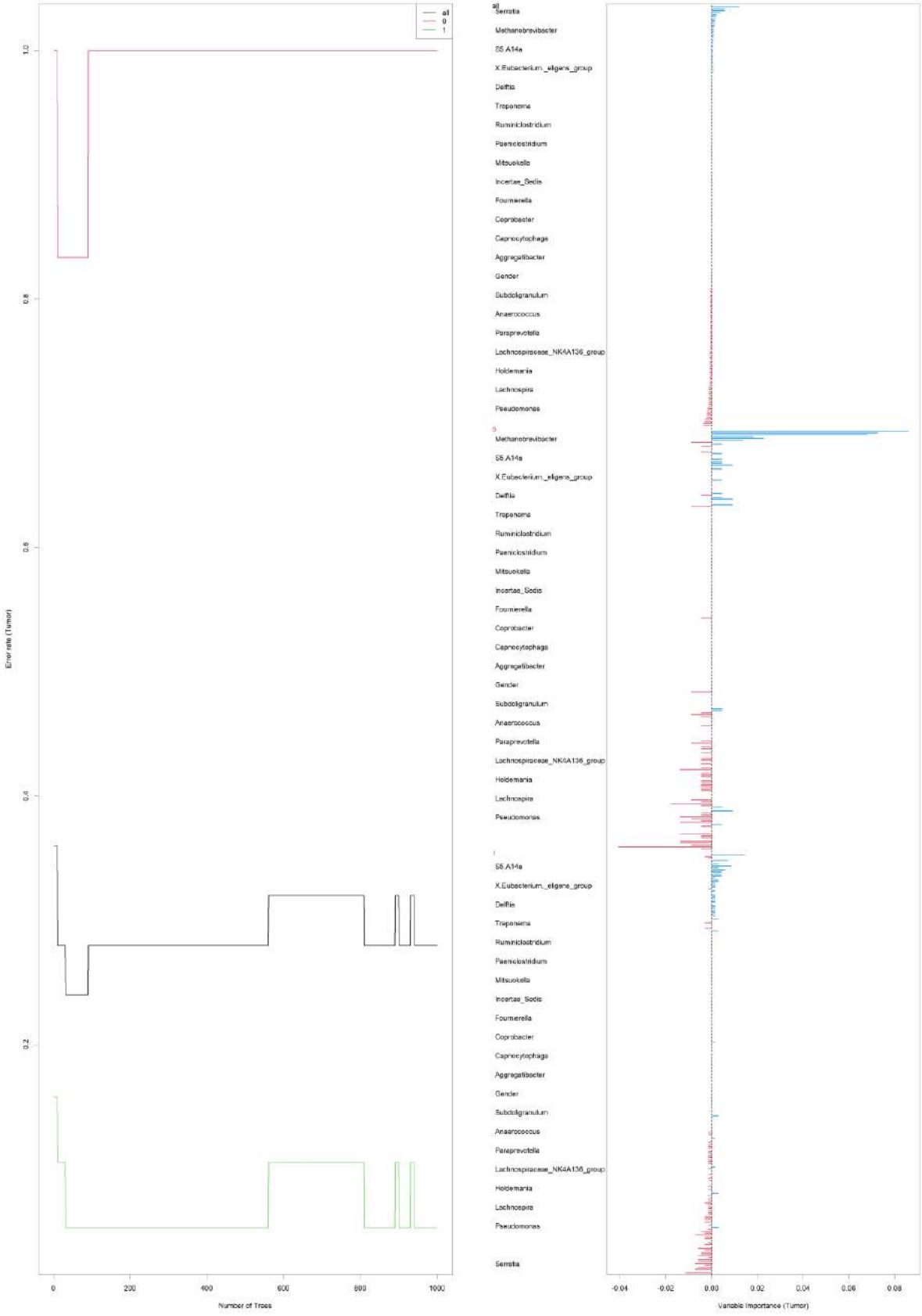
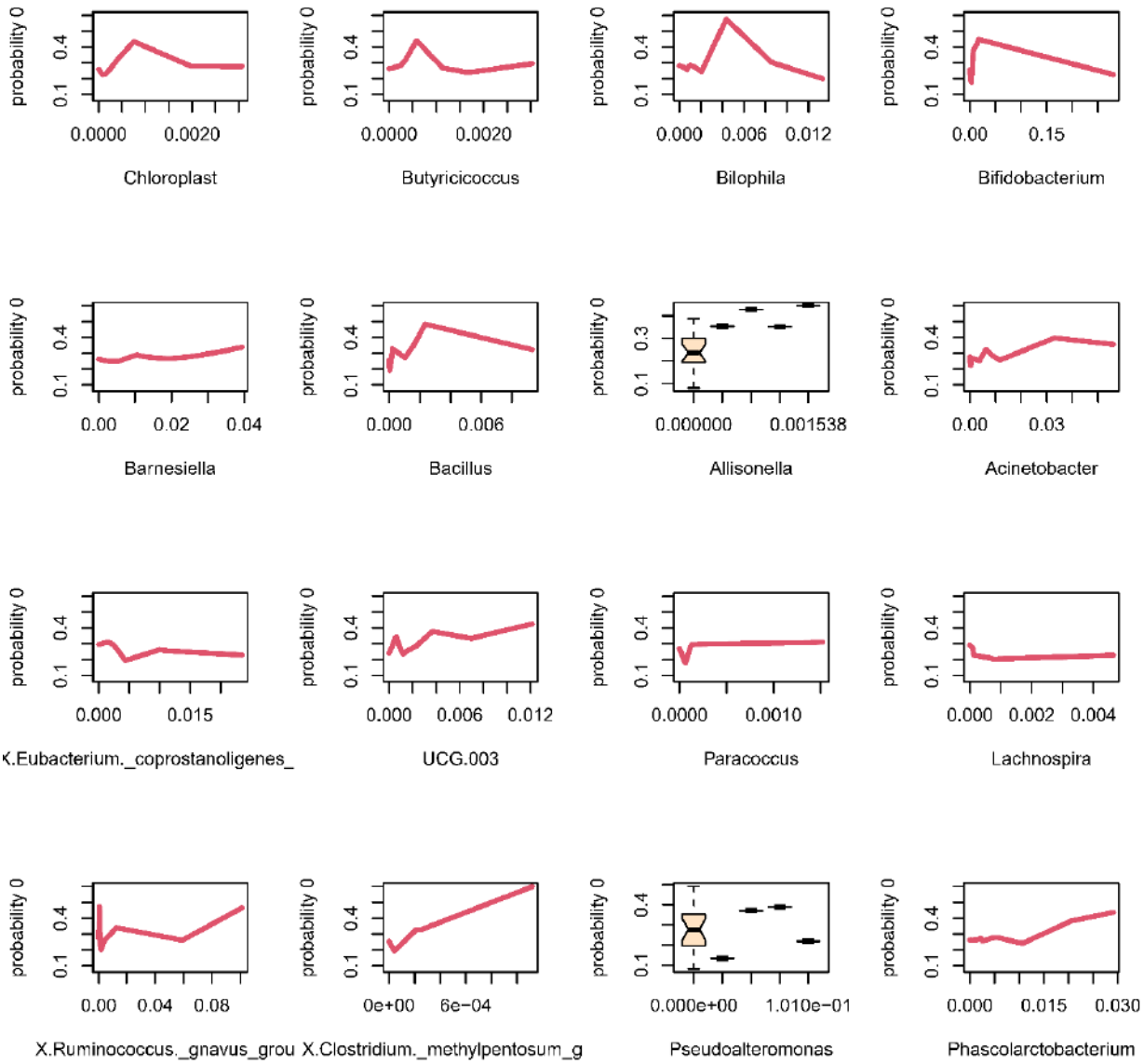
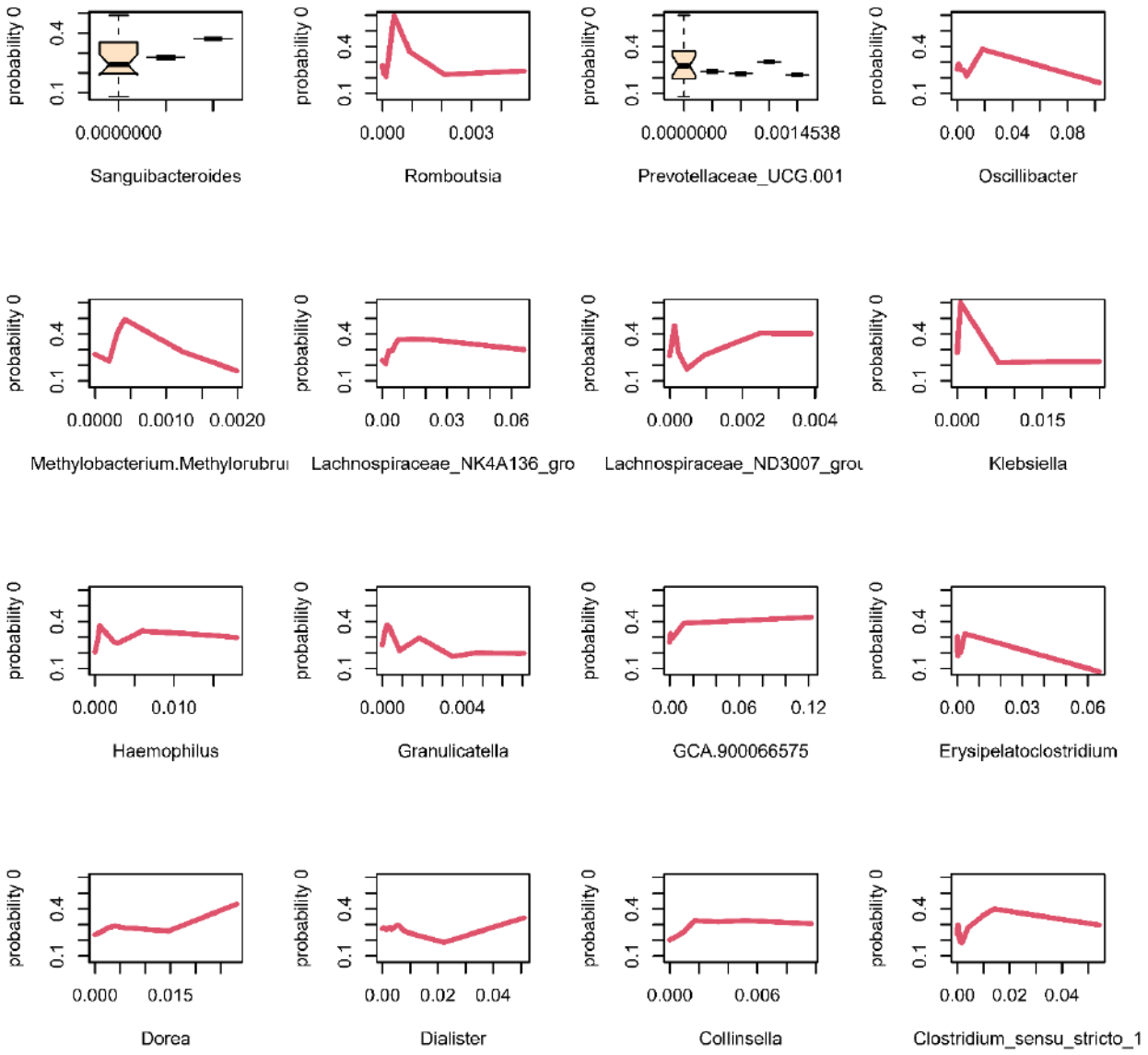
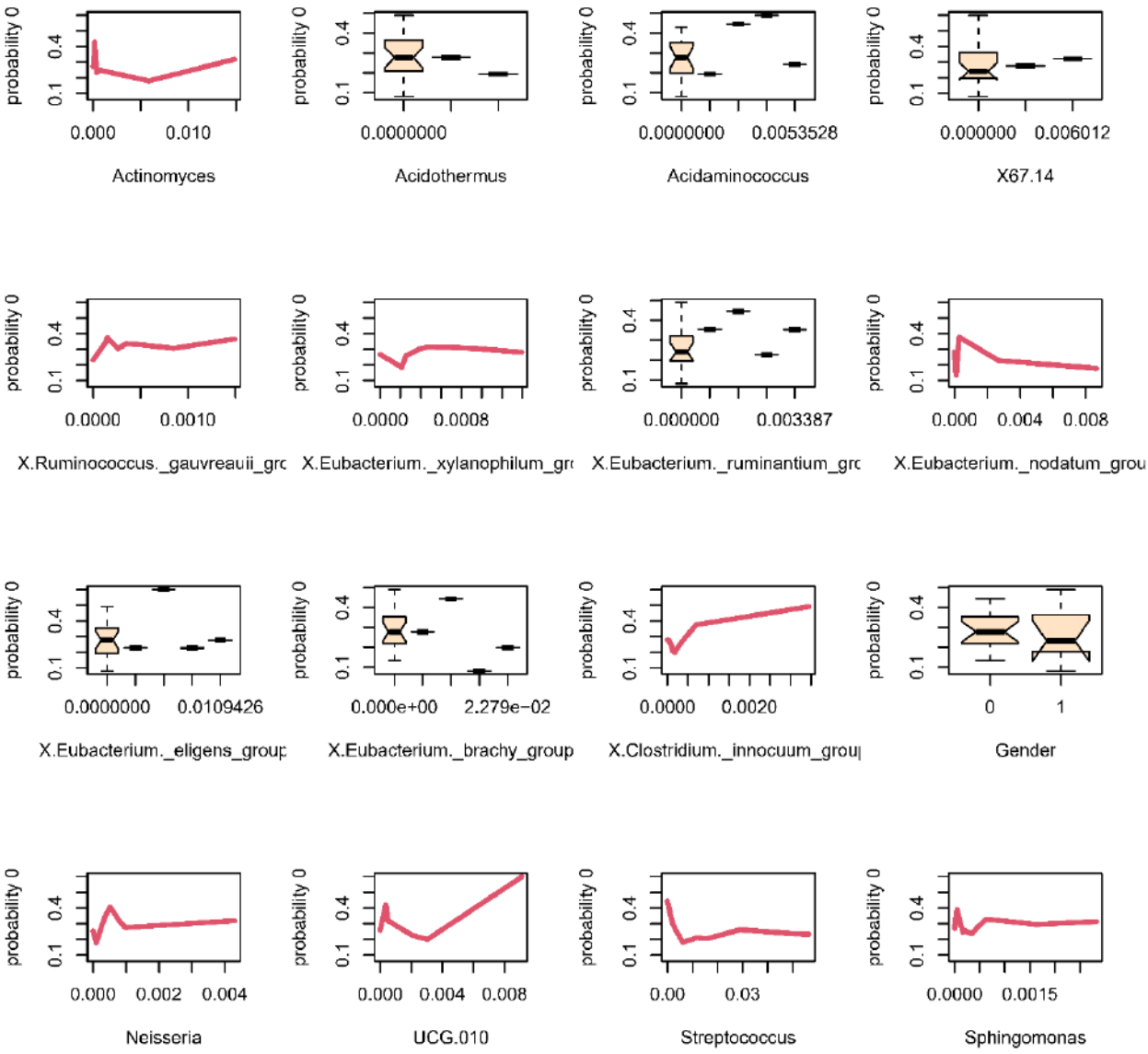
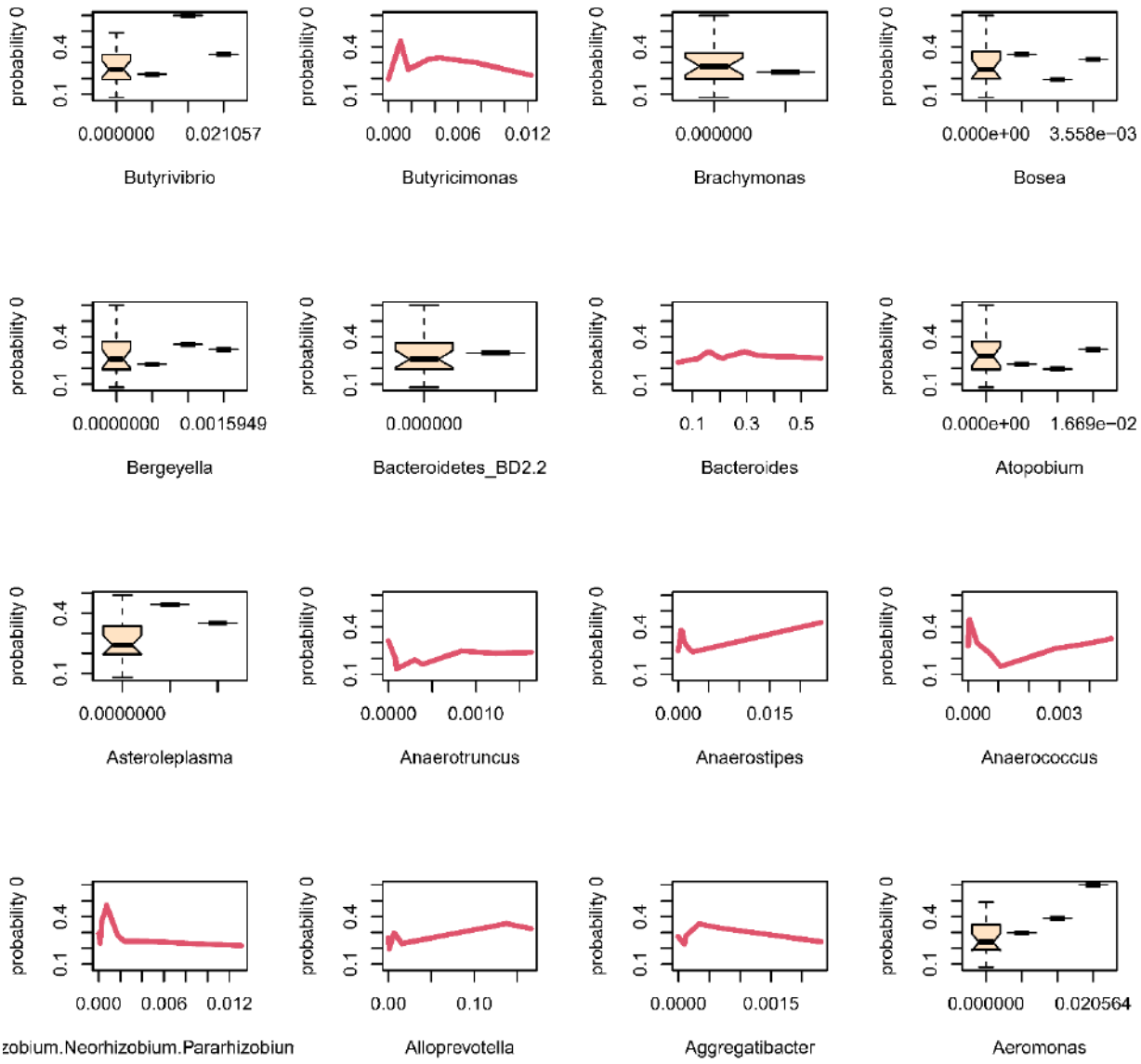

OOB
0
1
Error rate (Tumor)
Number of Trees
0
200
400
600
800
1000
0.2
0.4
0.6
0.8
1.0
Serratia
Methanobrevibacter
SS.A14a
X.Eubacterium._eligens_group
Delftia
Treponema
Ruminiclostridium
Paeniclostridium
Mitsuokella
Incertae_Sedis
Fournierella
Coprobacter
Capnocytophaga
Aggregatibacter
Gender
Subdoligranulum
Anaerococcus
Paraprevotella
Lachnospiraceae_NK4A136_group
Holdemania
Lachnospira
Pseudomonas
Variable Importance (Tumor)
-0.04
-0.02
0.00
0.02
0.04
0.06
0.08

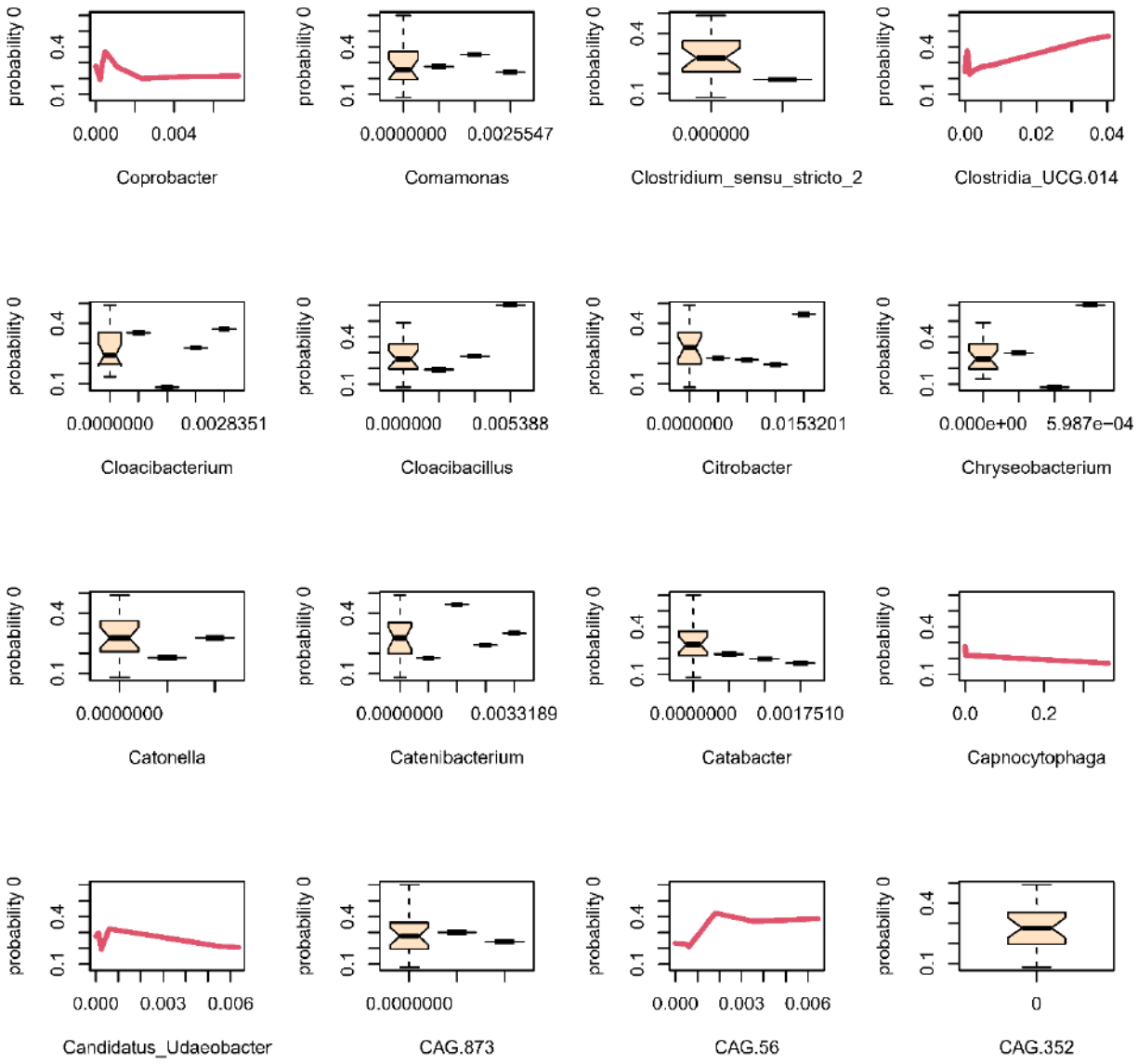
Figure ML1-E-1: The number of trees and Variable Importance Plot for Genus in Biopsy

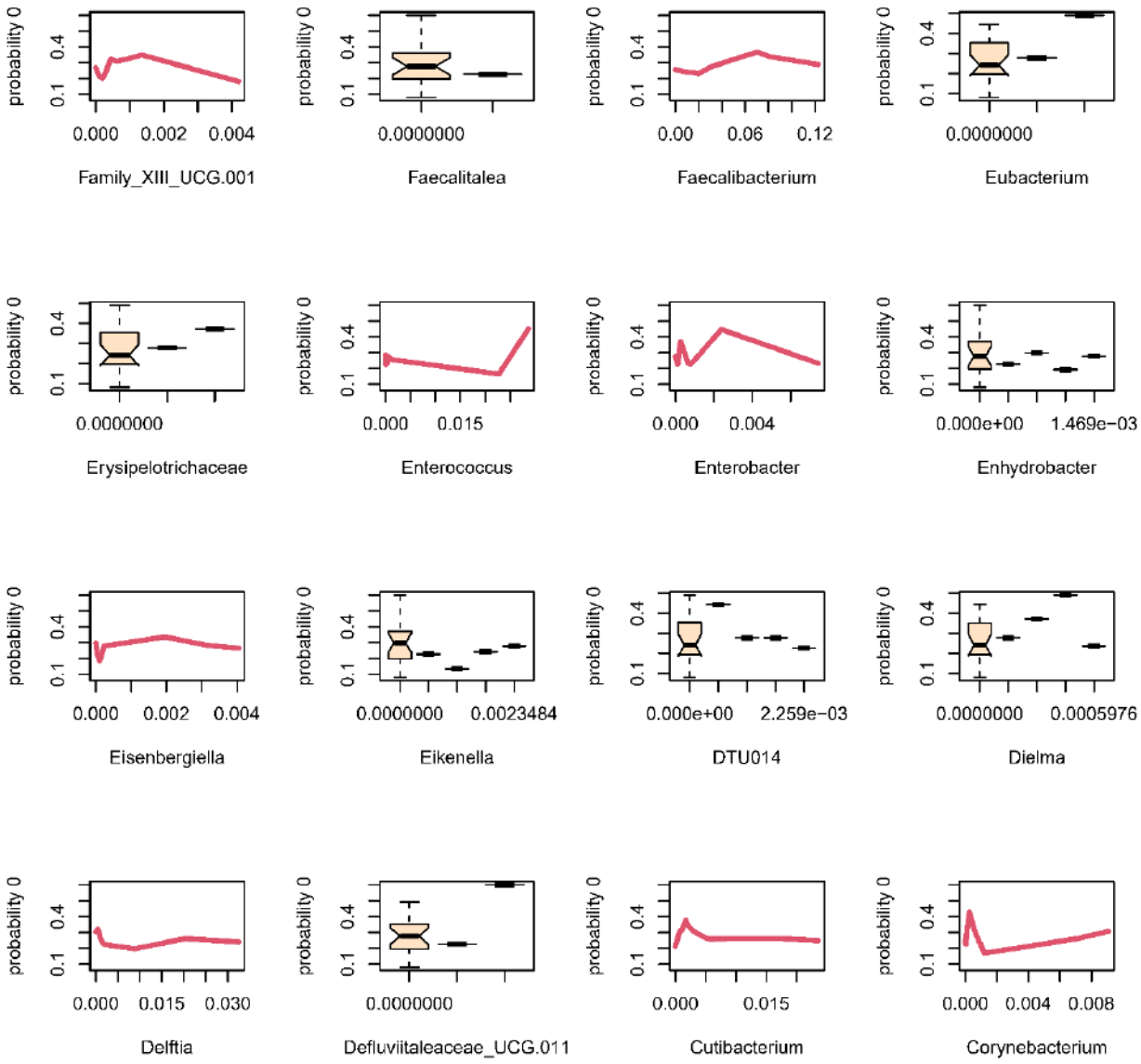

probability 0
0.1
0.4
0.0000
0.0020
Chloroplast
probability 0
0.1
0.4
0.0000
0.0020
Butyricicoccus
probability 0
0.1
0.4
0.000
0.006
0.012
Bilophila
probability 0
0.1
0.4
0.00
0.15
Bifidobacterium
probability 0
0.1
0.4
0.00
0.02
0.04
Barnesiella
probability 0
0.1
0.4
0.000
0.006
Bacillus
probability 0
0.1
0.3
0.000000
0.001538
Allisonella
probability 0
0.1
0.4
0.00
0.03
Acinetobacter
probability 0
0.1
0.4
0.000
0.015
K.Eubacterium._coprostanoligenes_
probability 0
0.1
0.4
0.000
0.006
0.012
UCG.003
probability 0
0.1
0.4
0.0000
0.0010
Paracoccus
probability 0
0.1
0.4
0.000
0.002
0.004
Lachnospira
probability 0
0.1
0.4
0.00
0.04
0.08
X.Ruminococcus._gnavus_grou
probability 0
0.1
0.4
0e+00
6e−04
X.Clostridium._methylpentosum_g
probability 0
0.1
0.4
0.000e+00
1.010e−01
Pseudoalteromonas
probability 0
0.1
0.4
0.000
0.015
0.030
Phascolarctobacterium

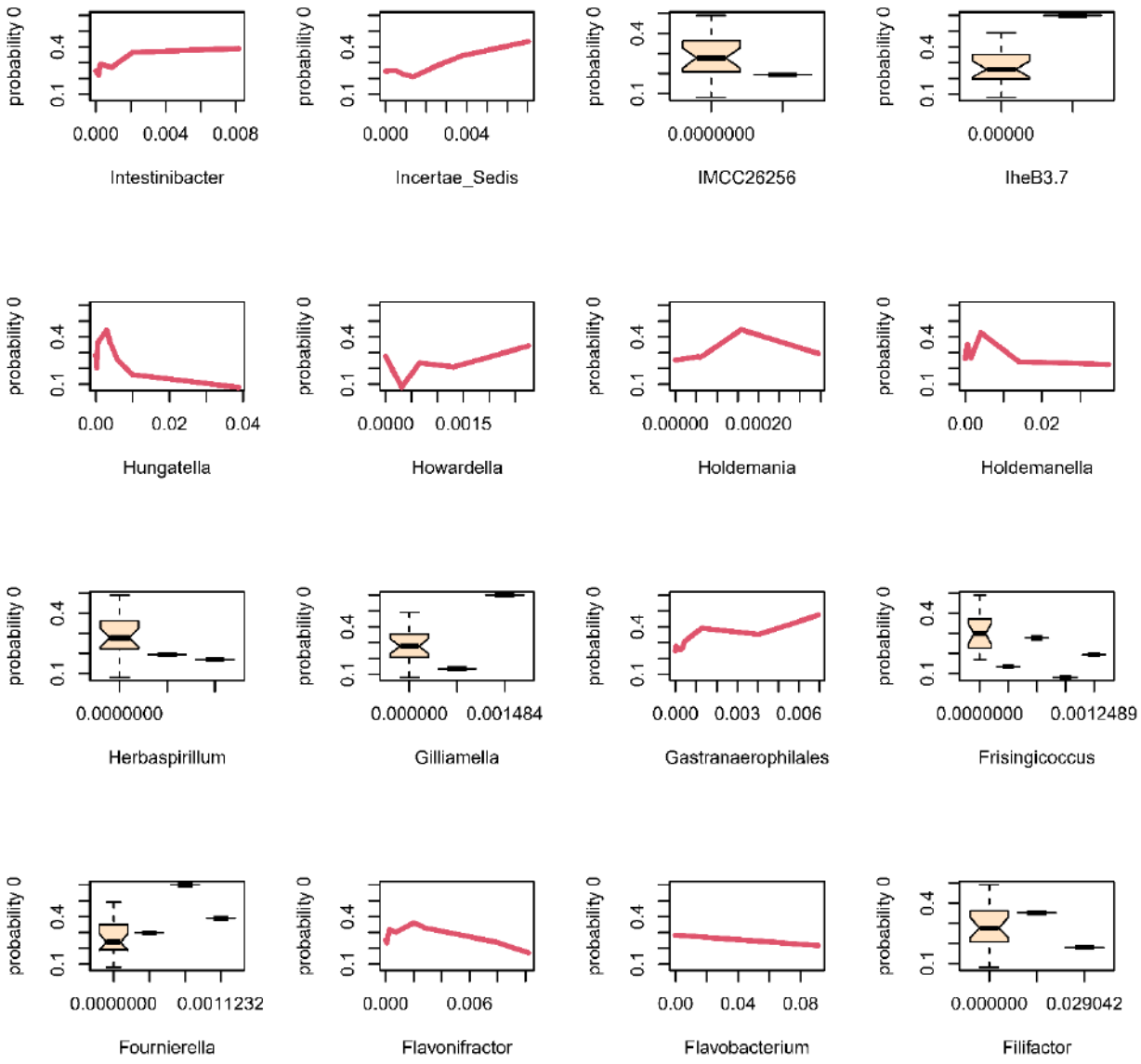

probability 0
0.1
0.4
0.0000000
Sanguibacteroides
0.000
0.003
Romboutsia
0.0000000
0.0014538
Prevotellaceae_UCG.001
0.00
0.04
0.08
Oscillibacter
0.0000
0.0010
0.0020
Methylobacterium.Methylorubru
0.00
0.03
0.06
Lachnospiraceae_NK4A136_gro
0.000
0.002
0.004
Lachnospiraceae_ND3007_grou
0.000
0.015
Klebsiella
0.000
0.010
Haemophilus
0.000
0.004
Granulicatella
0.00
0.06
0.12
GCA.900066575
0.00
0.03
0.06
Erysipelatoclostridium
0.000
0.015
Dorea
0.00
0.02
0.04
Dialister
0.000
0.006
Collinsella
0.00
0.02
0.04
Clostridium_sensu_stricto_1

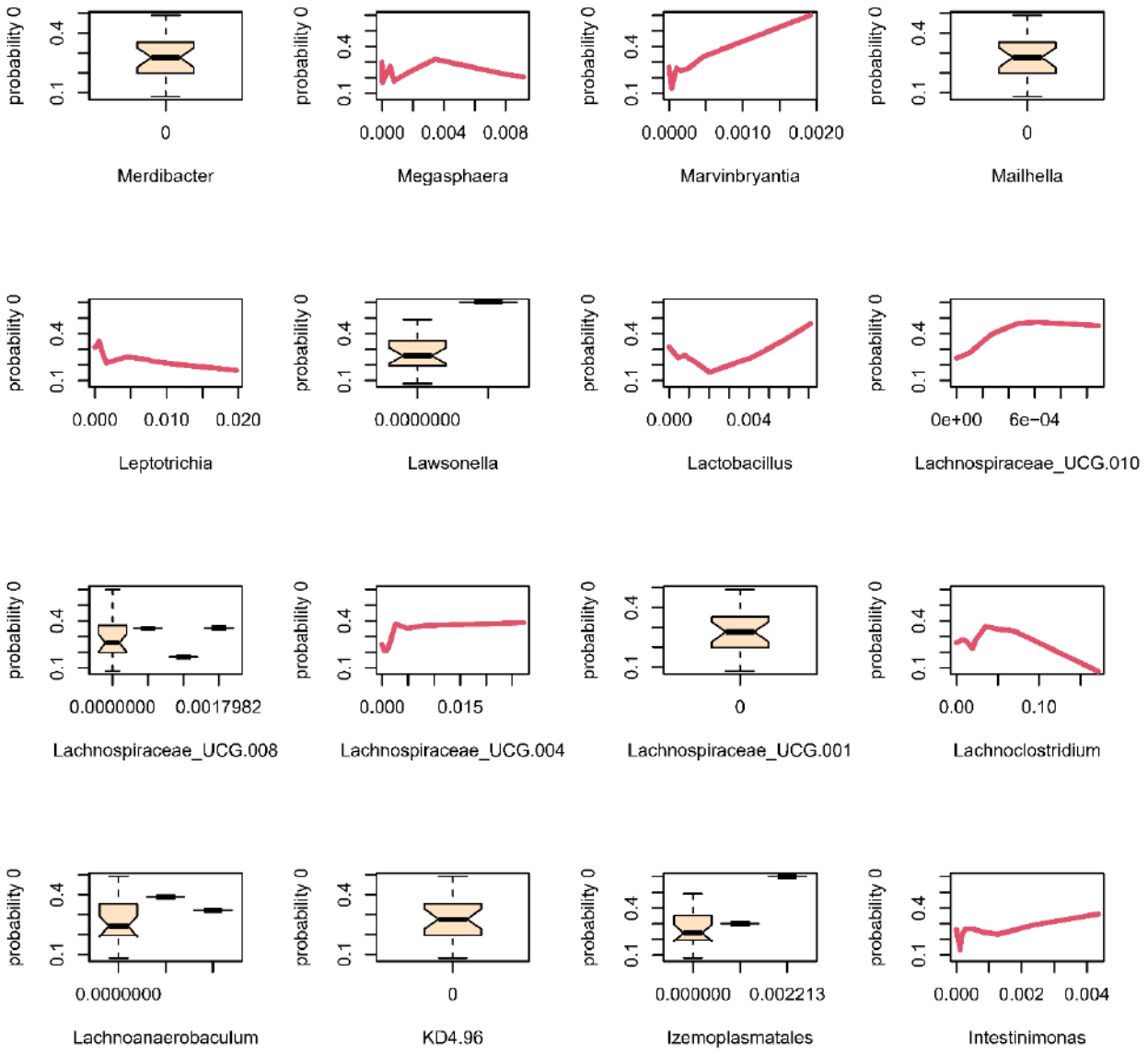

probability 0
0.1
0.4
0.000
0.010
Actinomyces
0.0000000
Acidothermus
0.0000000
0.0053528
Acidaminococcus
0.000000
0.006012
X67.14
0.0000
0.0010
X.Ruminococcus._gauvreauii_grc
0.0000
0.0008
X.Eubacterium._xylanophilum_gr
0.000000
0.003387
X.Eubacterium._ruminantium_grc
0.000
0.004
0.008
X.Eubacterium._nodatum_grou
0.0000000
0.0109426
X.Eubacterium._eligens_group
0.000e+00
2.279e−02
X.Eubacterium._brachy_group
0.0000
0.0020
X.Clostridium._innocuum_grou
0
1
Gender
0.000
0.002
0.004
Neisseria
0.000
0.004
0.008
UCG.010
0.00
0.03
Streptococcus
0.0000
0.0015
Sphingomonas

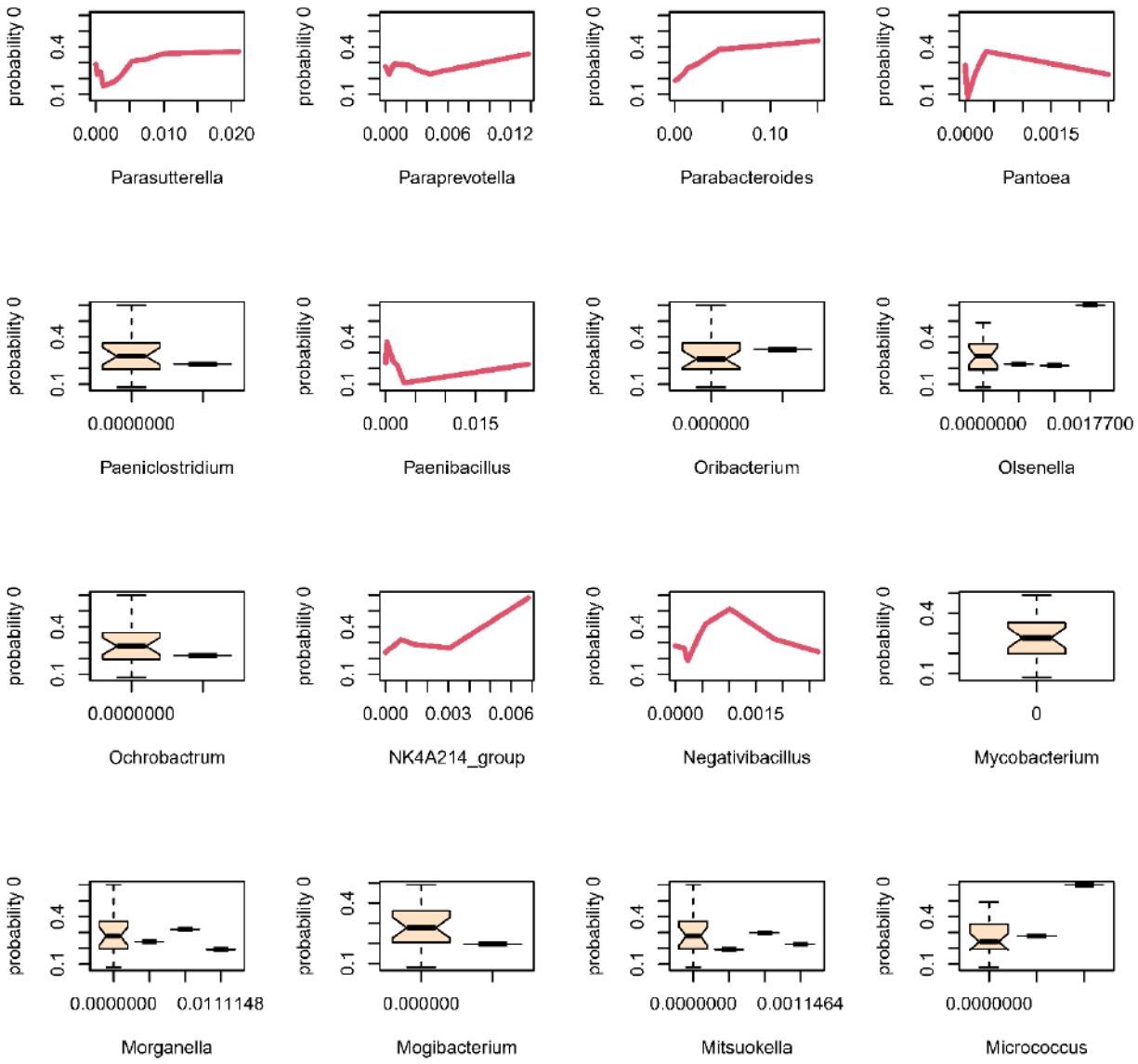

probability 0
0.1
0.4
0.000000
0.021057
Butyrivibrio
0.000
0.006
0.012
Butyricimonas
0.000000
Brachymonas
0.000e+00
3.558e−03
Bosea
0.0000000
0.0015949
Bergeyella
0.000000
Bacteroidetes_BD2.2
0.1
0.3
0.5
Bacteroides
0.000e+00
1.669e−02
Atopobium
0.0000000
Asteroleplasma
probability 0
0.0000
0.0010
Anaerotruncus
0.000
0.015
Anaerostipes
0.003
Anaerococcus
zobium.Neorhizobium.Pararhizobiun
0.00
0.10
Alloprevotella
0.0000
0.0015
Aggregatibacter
0.020564
Aeromonas

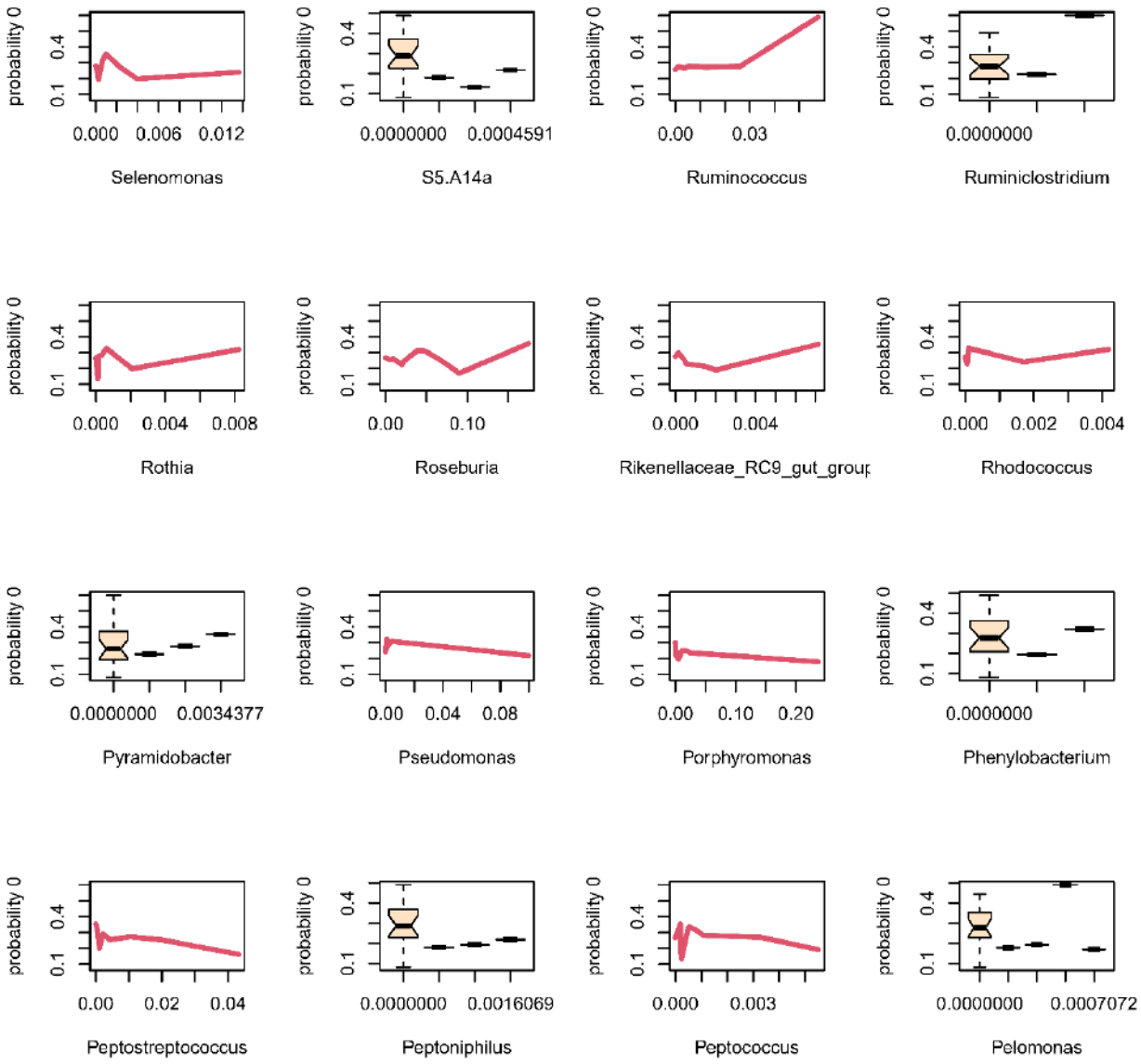

probability 0
0.1
0.4
0.000
0.004
Coprobacter
0.0000000
0.0025547
Comamonas
0.000000
Clostridium_sensu_stricto_2
0.00
0.02
0.04
Clostridia_UCG.014
0.0000000
0.0028351
Cloacibacterium
0.000000
0.005388
Cloacibacillus
0.0000000
0.0153201
Citrobacter
0.000e+00
5.987e−04
Chryseobacterium
Catonella
0.0000000
0.0033189
Catenibacterium
0.0017510
Catabacter
0.0
0.2
Capnocytophaga
0.003
0.006
Candidatus_Udaeobacter
CAG.873
CAG.56
0
CAG.352

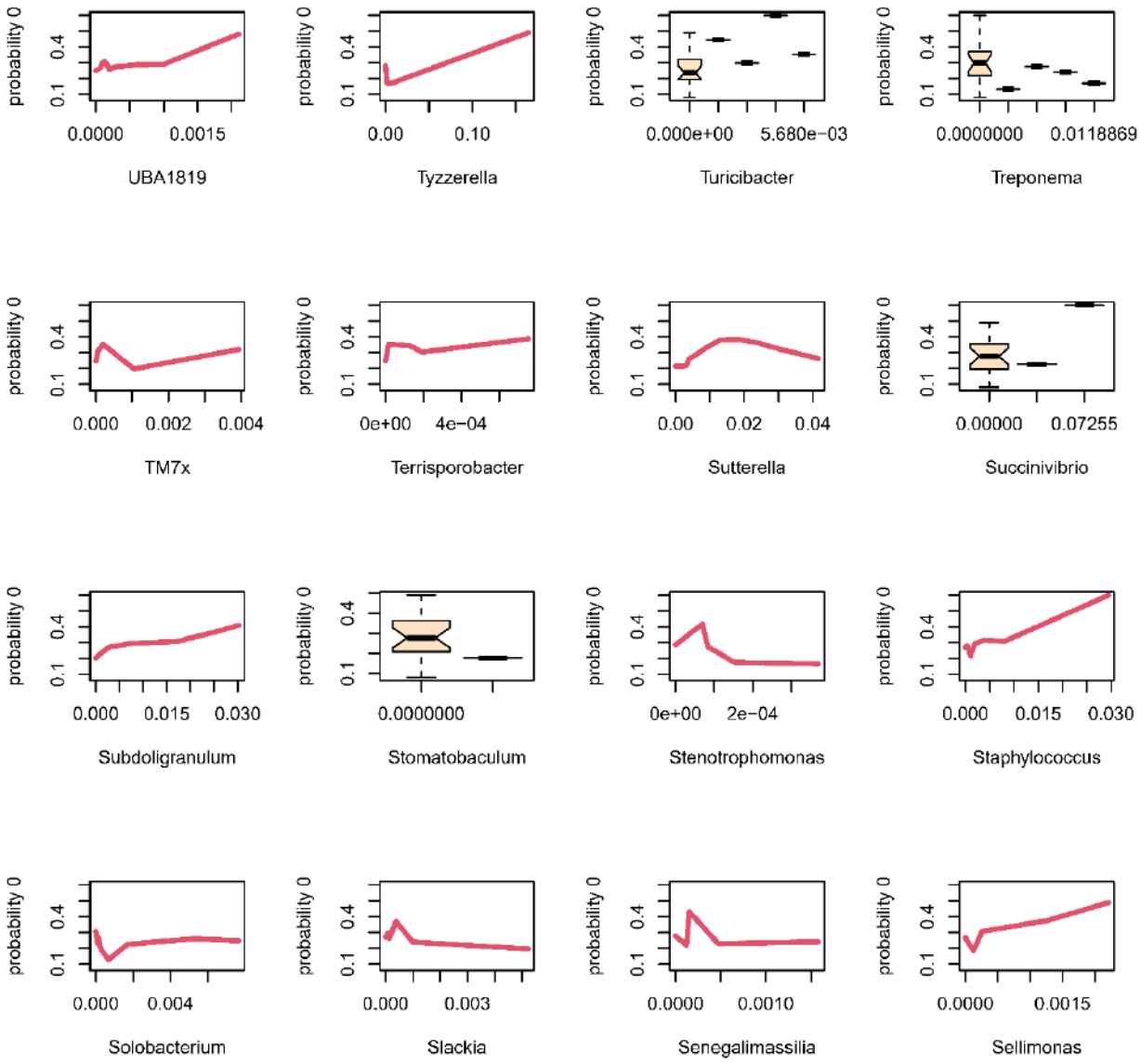

probability 0
0.1
0.4
0.000
0.002
0.004
Family_XIII_UCG.001
0.0000000
Faecalitalea
0.00
0.06
0.12
Faecalibacterium
Eubacterium
Erysipelotrichaceae
0.000
0.015
Enterococcus
0.004
Enterobacter
0.000e+00
1.469e−03
Enhydrobacter
Eisenbergiella
probability
0.0023484
Eikenella
2.259e−03
DTU014
0.0005976
Dielma
0.030
Delftia
Defluviitaleaceae_UCG.011
Cutibacterium
0.008
Corynebacterium

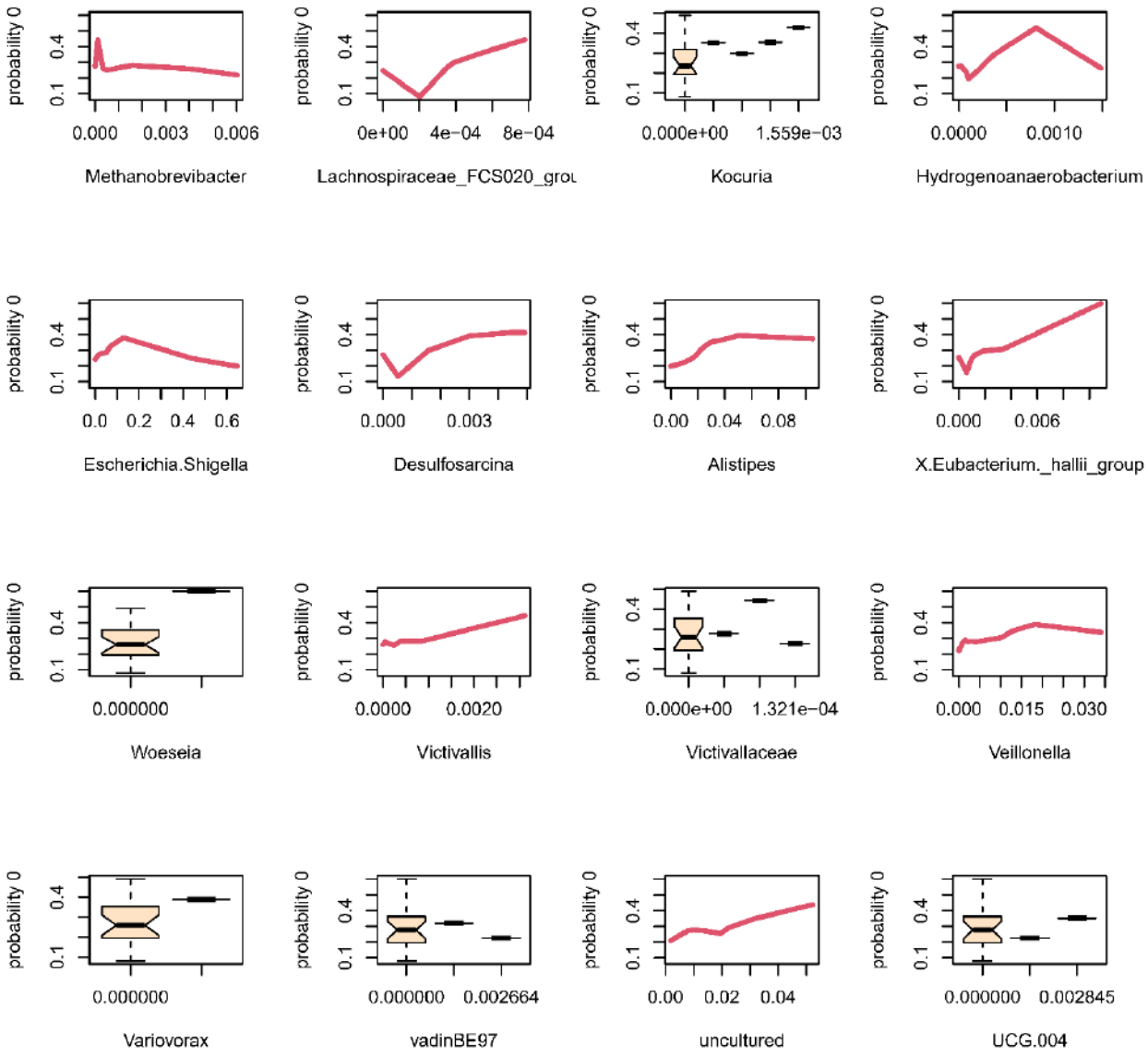

probability 0
0.1
0.4
0.000
0.004
0.008
Intestinibacter
probability 0
0.1
0.4
0.000
0.004
Incertae_Sedis
probability 0
0.1
0.4
0.0000000
IMCC26256
probability 0
0.1
0.4
0.00000
IheB3.7
probability 0
0.1
0.4
0.00
0.02
0.04
Hungatella
probability 0
0.1
0.4
0.0000
0.0015
Howardella
probability 0
0.1
0.4
0.00000
0.00020
Holdemania
probability 0
0.1
0.4
0.00
0.02
Holdemanella
probability 0
0.1
0.4
0.0000000
Herbaspirillum
probability 0
0.1
0.4
0.000000
0.001484
Gilliamella
probability 0
0.1
0.4
0.000
0.003
0.006
Gastranaerophilales
probability 0
0.1
0.4
0.0000000
0.0012489
Frisingicoccus
probability 0
0.1
0.4
0.0000000
0.0011232
Fournierella
probability 0
0.1
0.4
0.000
0.006
Flavonifractor
probability 0
0.1
0.4
0.00
0.04
0.08
Flavobacterium
probability 0
0.1
0.4
0.000000
0.029042
Filifactor

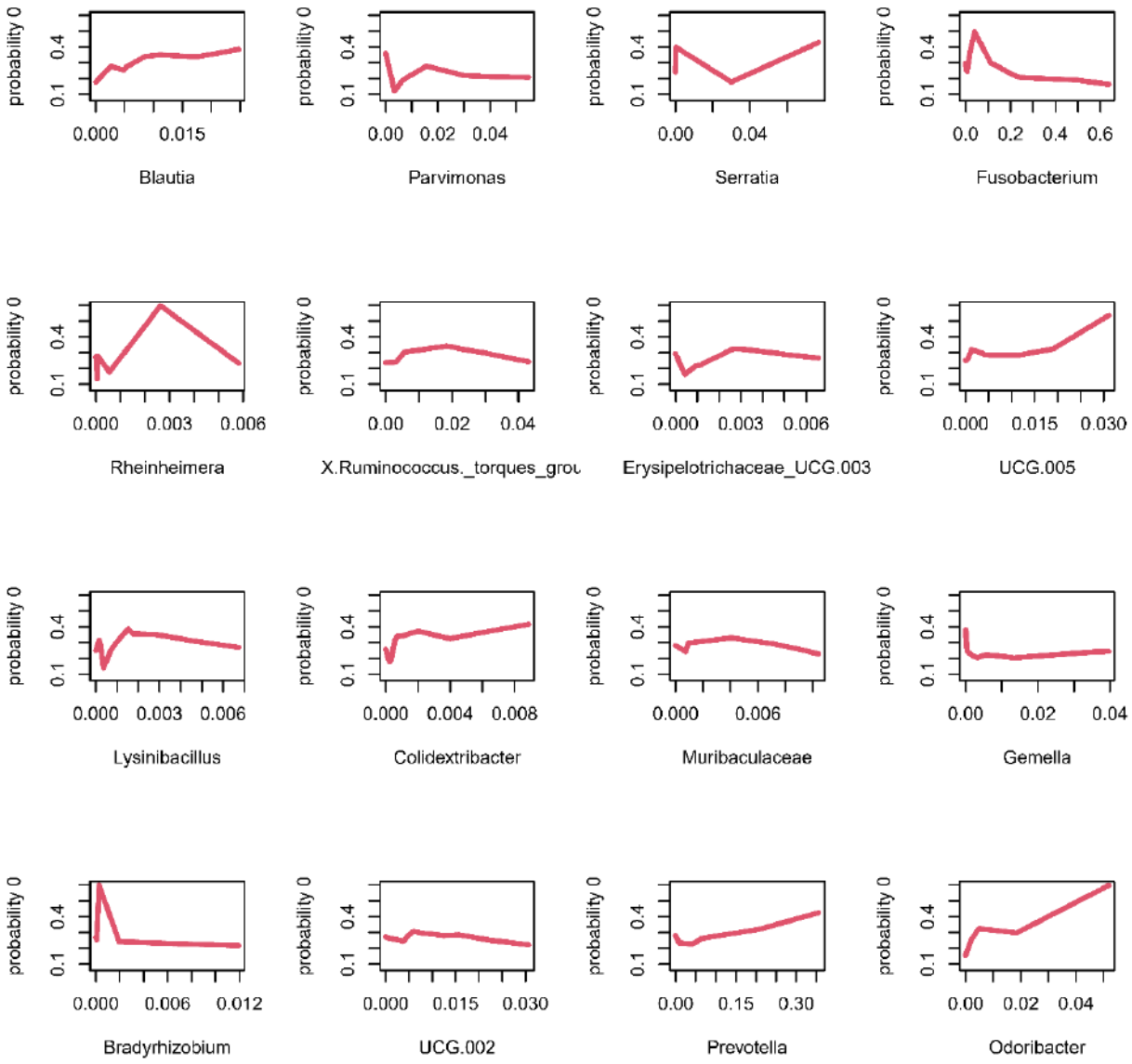

probability 0
0.1
0.4
0
Merdibacter
0.000
0.004
0.008
Megasphaera
0.0000
0.0010
0.0020
Marvinbryantia
Mailhella
0.010
0.020
Leptotrichia
0.0000000
Lawsonella
Lactobacillus
0e+00
6e−04
Lachnospiraceae_UCG.010
0.0017982
Lachnospiraceae_UCG.008
0.015
Lachnospiraceae_UCG.004
Lachnospiraceae_UCG.001
0.00
0.10
Lachnoclostridium
Lachnoanaerobaculum
KD4.96
0.002213
Izemoplasmatales
0.002
Intestinimonas


probability 0
0.1
0.4
0.000
0.010
0.020
Parasutterella
0.000
0.006
0.012
Paraprevotella
0.00
0.10
Parabacteroides
0.0000
0.0015
Pantoea
0.0000000
Paeniclostridium
0.000
0.015
Paenibacillus
0.000000
Oribacterium
0.0000000
0.0017700
Olsenella
0.0000000
Ochrobactrum
probability 0
0.000
0.003
0.006
NK4A214_group
0.0000
0.0015
Negativibacillus
0
Mycobacterium
0.0000000
0.0111148
Morganella
0.000000
Mogibacterium
0.0000000
0.0011464
Mitsuokella
0.0000000
Micrococcus


probability 0
0.1
0.4
0.000 0.006 0.012
Selenomonas
0.0000000 0.0004591
S5.A14a
0.00 0.03
Ruminococcus
0.0000000
Ruminiclostridium
0.000 0.004 0.008
Rothia
0.00 0.10
Roseburia
0.000 0.004
Rikenellaceae_RC9_gut_grou
0.000 0.002 0.004
Rhodococcus
0.0000000 0.0034377
Pyramidobacter
probability 0
0.00 0.04 0.08
Pseudomonas
0.00 0.10 0.20
Porphyromonas
0.0000000
Phenylobacterium
0.00 0.02 0.04
Peptostreptococcus
0.0000000 0.0016069
Peptoniphilus
0.000 0.003
Peptococcus
0.0000000 0.0007072
Pelomonas


probability 0
0.1
0.4
0.0000
0.0015
UBA1819
0.00
0.10
Tyzzerella
0.000e+00
5.680e−03
Turicibacter
0.0000000
0.0118869
Treponema
0.000
0.002
0.004
TM7x
0e+00
4e−04
Terrisporobacter
0.00
0.02
0.04
Sutterella
0.00000
0.07255
Succinivibrio
0.000
0.015
0.030
Subdoligranulum
0.0000000
Stomatobaculum
0e+00
2e−04
Stenotrophomonas
0.000
0.015
0.030
Staphylococcus
0.000
0.004
Solobacterium
0.000
0.003
Slackia
0.0000
0.0010
Senegalimassilia
0.0000
0.0015
Sellimonas


probability 0
0.1
0.4
0.000
0.003
0.006
Methanobrevibacter
0e+00
4e−04
8e−04
Lachnospiraceae_FCS020_grou
0.000e+00
1.559e−03
Kocuria
0.0000
0.0010
Hydrogenoanaerobacterium
0.0
0.2
0.4
0.6
Escherichia.Shigella
0.000
0.003
Desulfosarcina
0.00
0.04
0.08
Alistipes
0.000
0.006
X.Eubacterium._hallii_group
0.000000
Woeseia
0.0000
0.0020
Victivallis
0.000e+00
1.321e−04
Victivallaceae
0.000
0.015
0.030
Veillonella
0.000000
Variovorax
0.000000
0.002664
vadinBE97
0.00
0.02
0.04
uncultured
0.000000
0.002845
UCG.004


probability 0
0.1
0.4
0.000
0.015
Blautia
0.00
0.02
0.04
Parvimonas
0.00
0.04
Serratia
0.0
0.2
0.4
0.6
Fusobacterium
0.000
0.003
0.006
Rheinheimera
X.Ruminococcus._torques_grou
Erysipelotrichaceae_UCG.003
0.030
UCG.005
Lysinibacillus
probability
0.008
Colidextribacter
Muribaculaceae
Gemella
0.012
Bradyrhizobium
UCG.002
0.15
0.30
Prevotella
Odoribacter

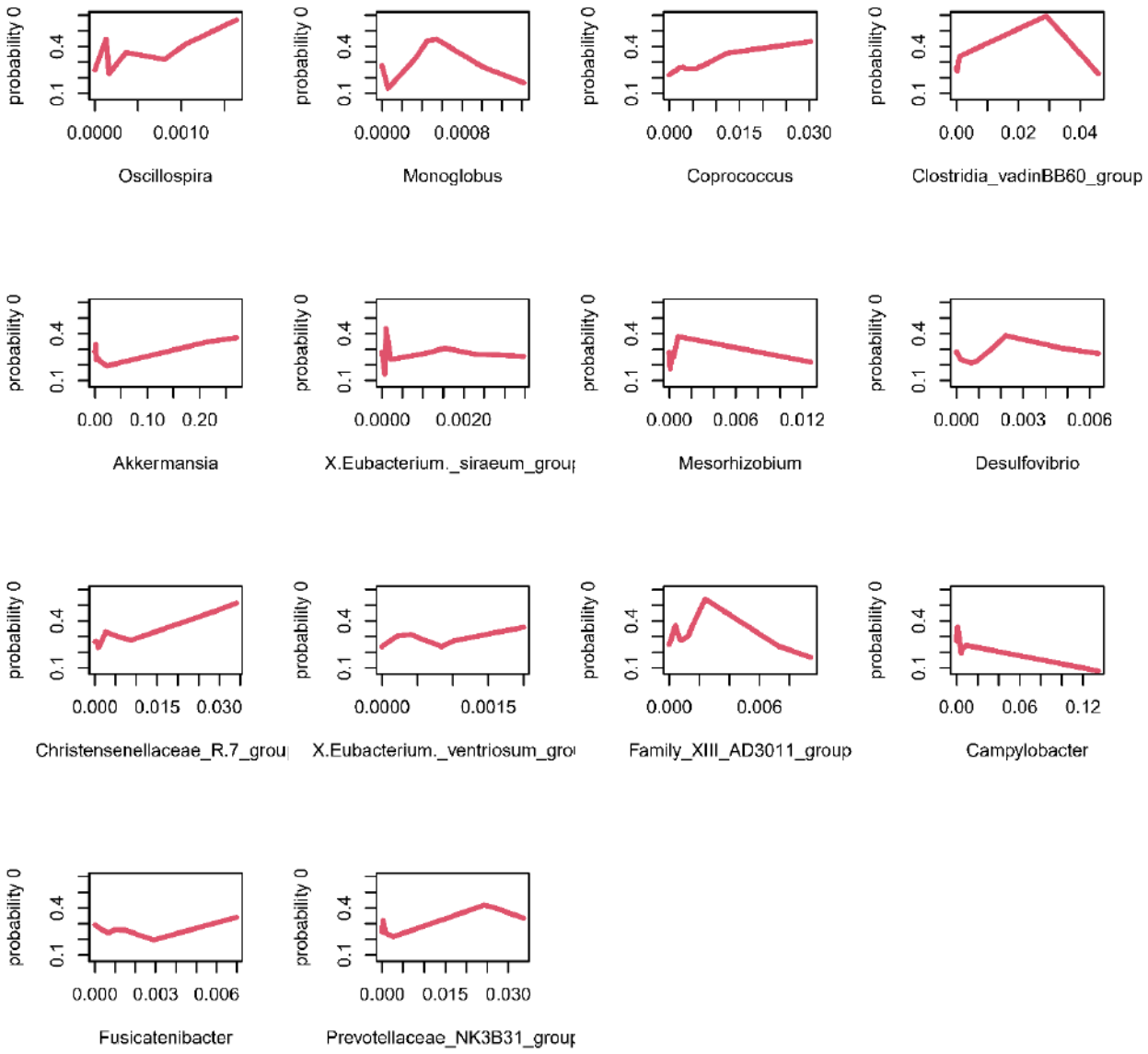
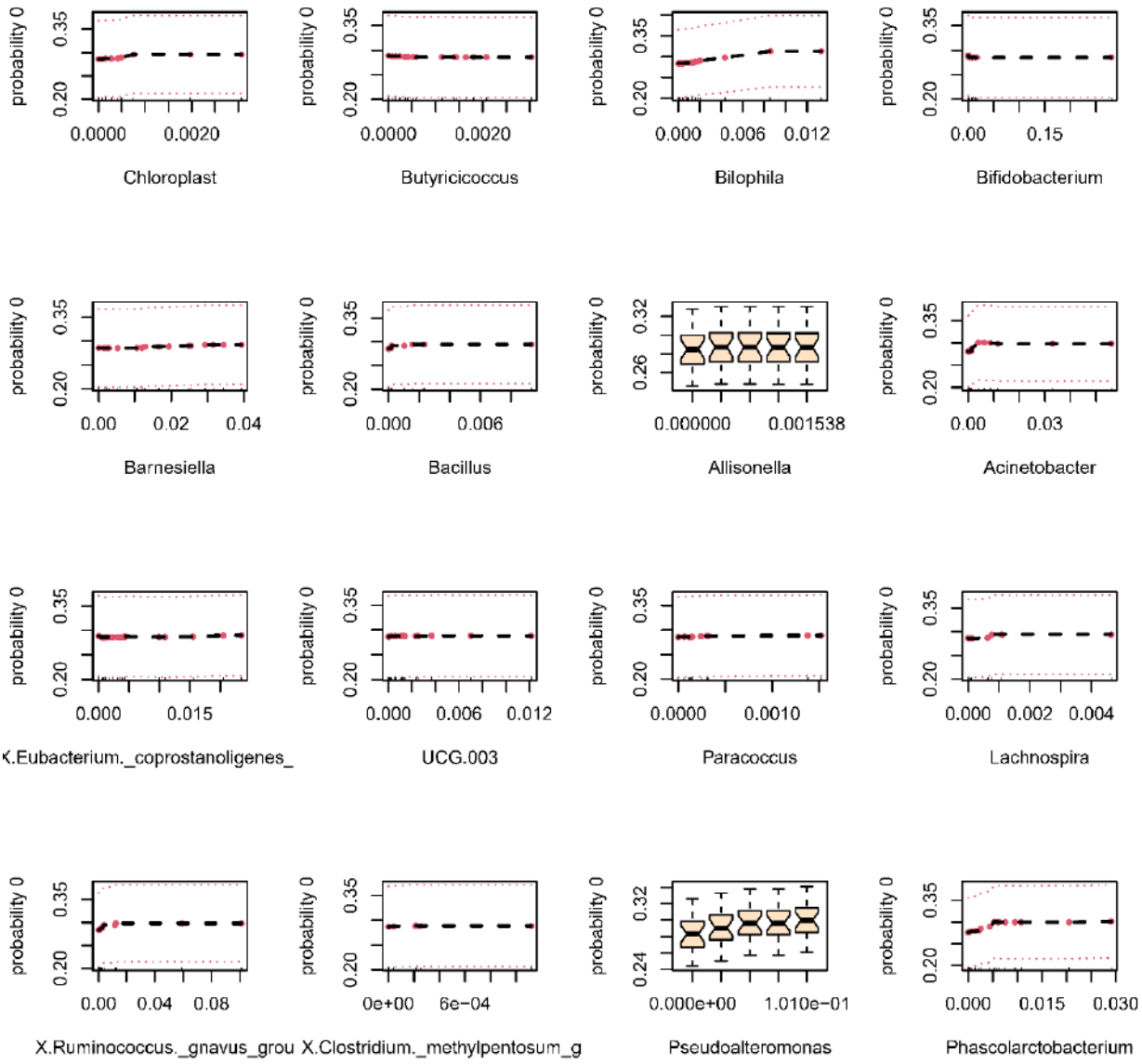
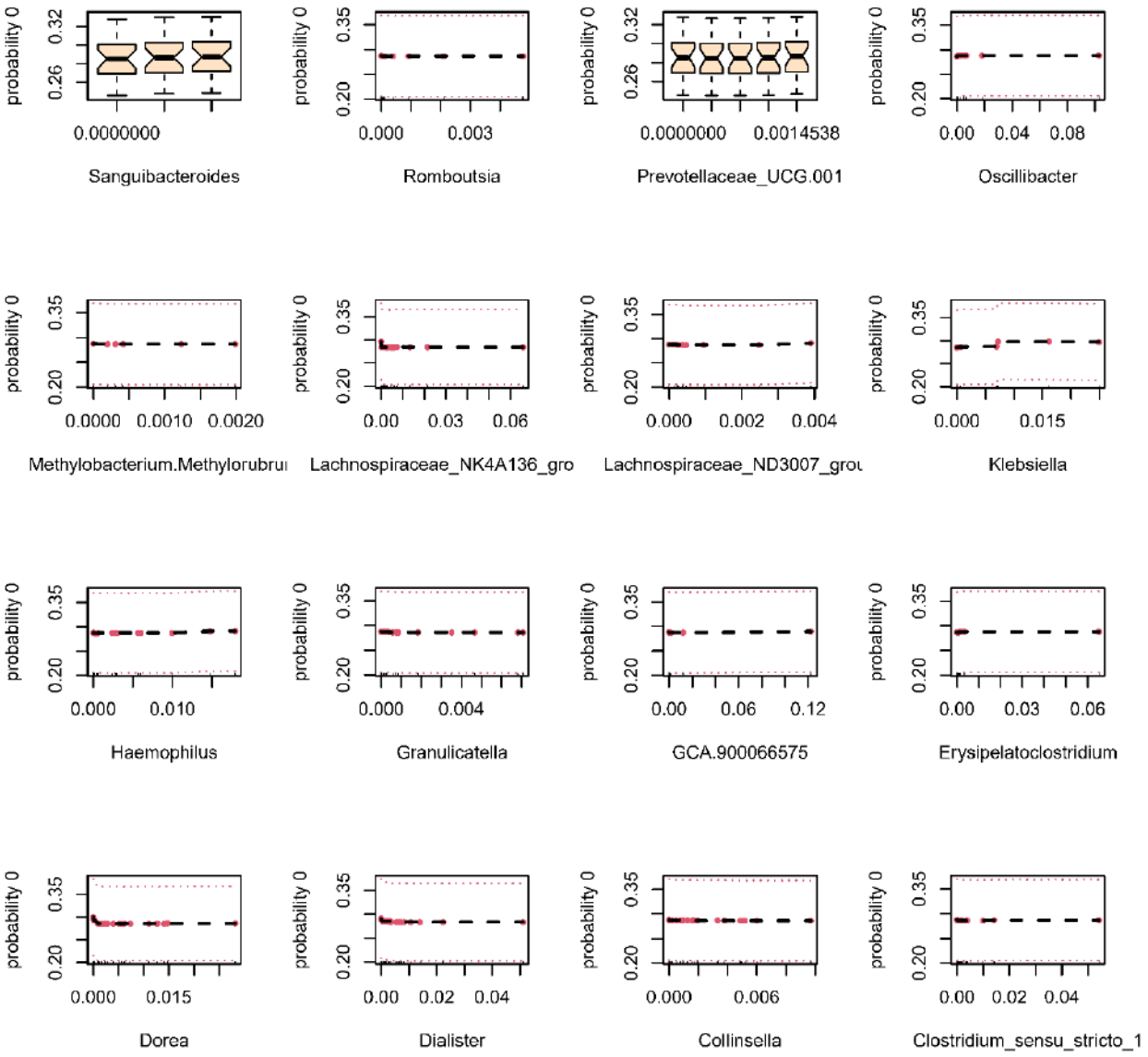
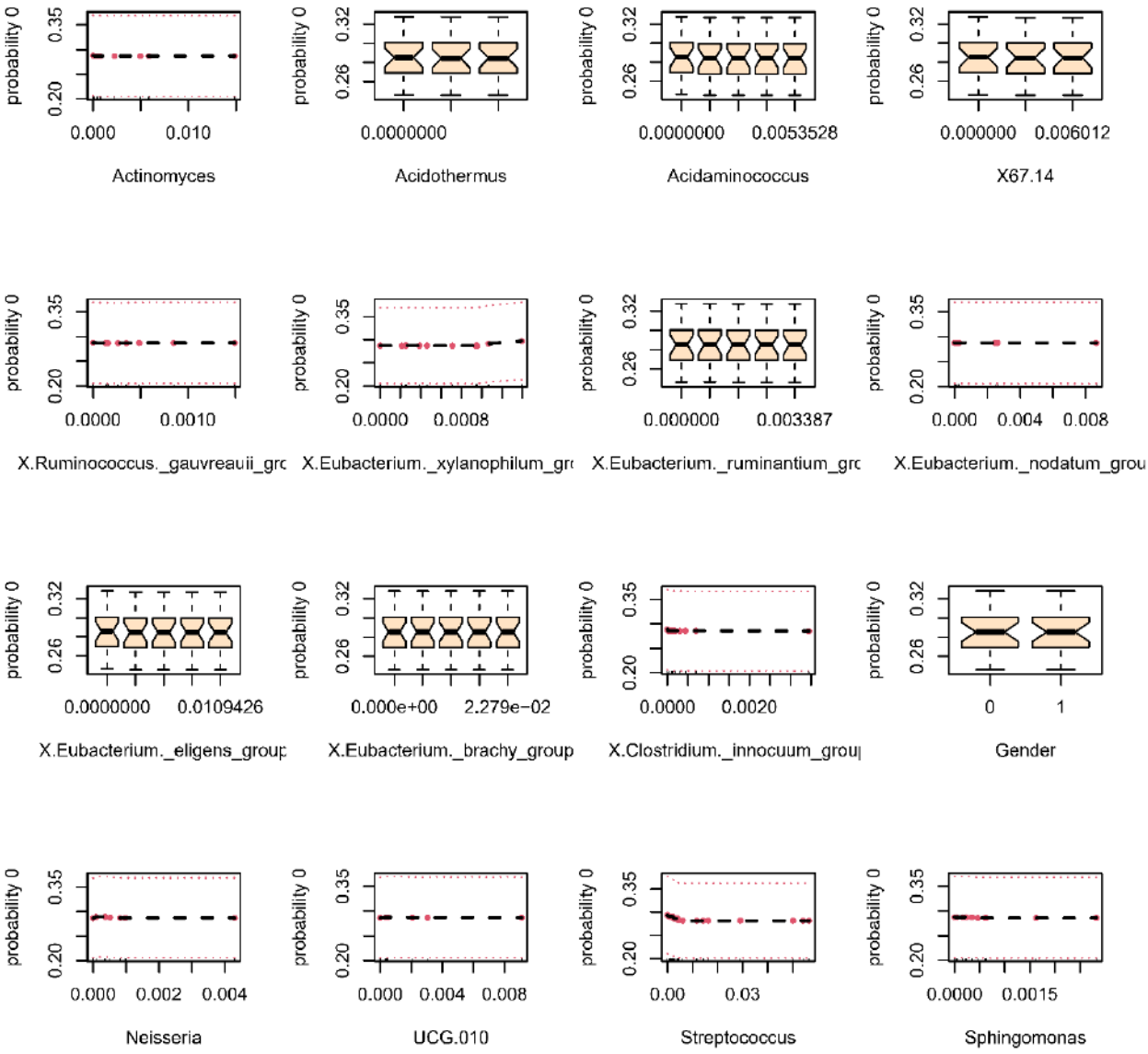
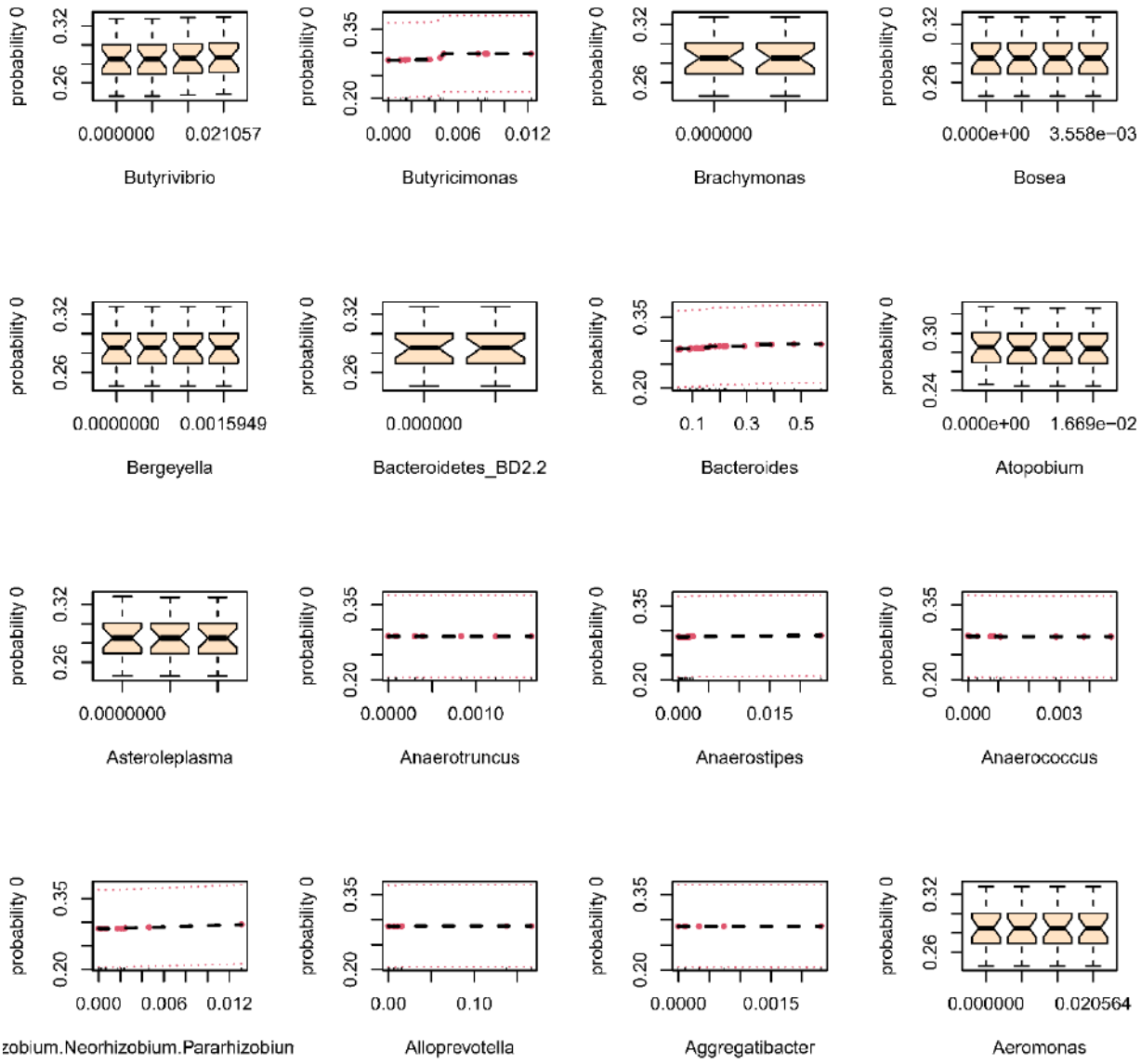


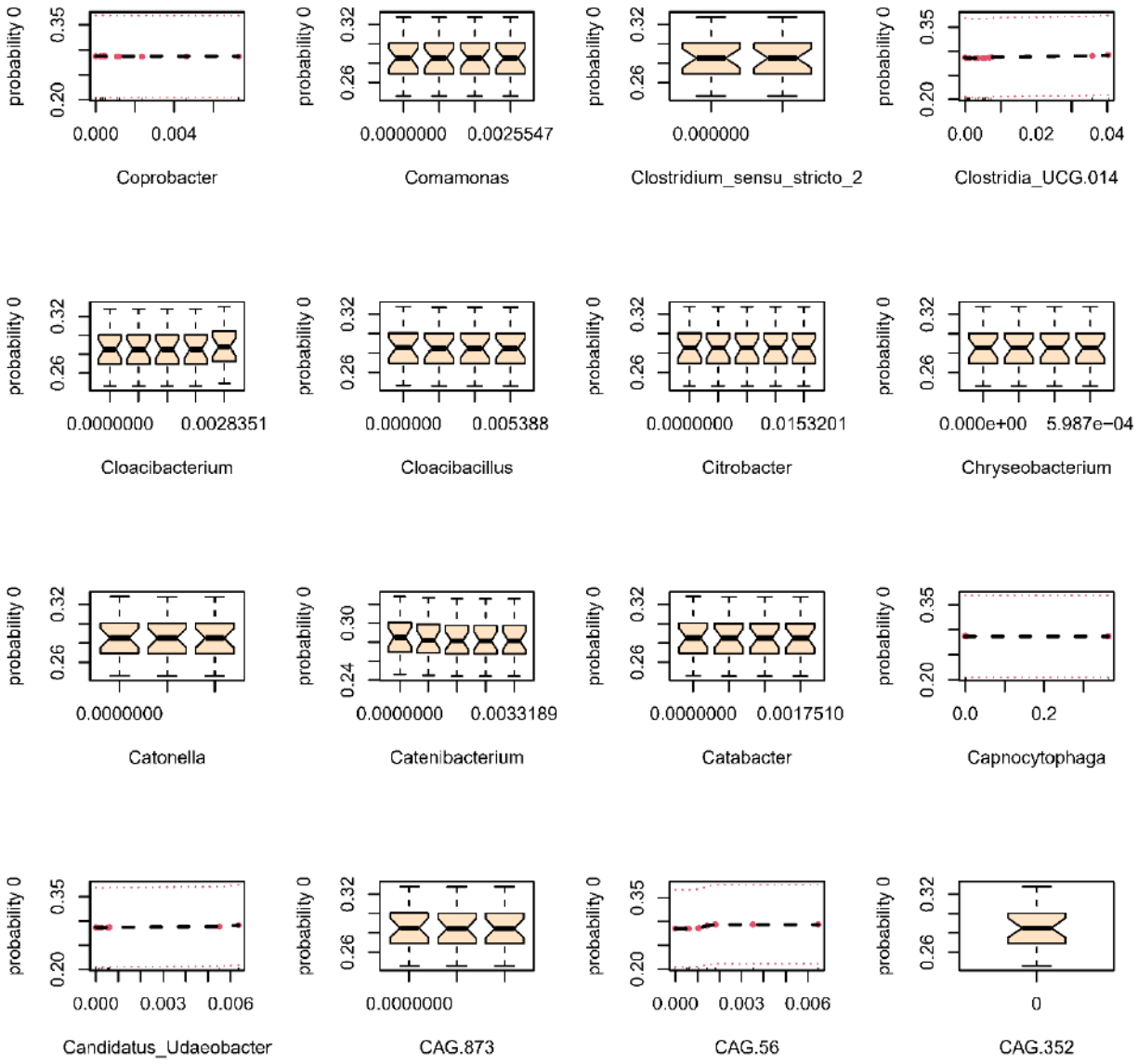
Figure ML1-E-2: Marginal Plot for Genus in Biopsy

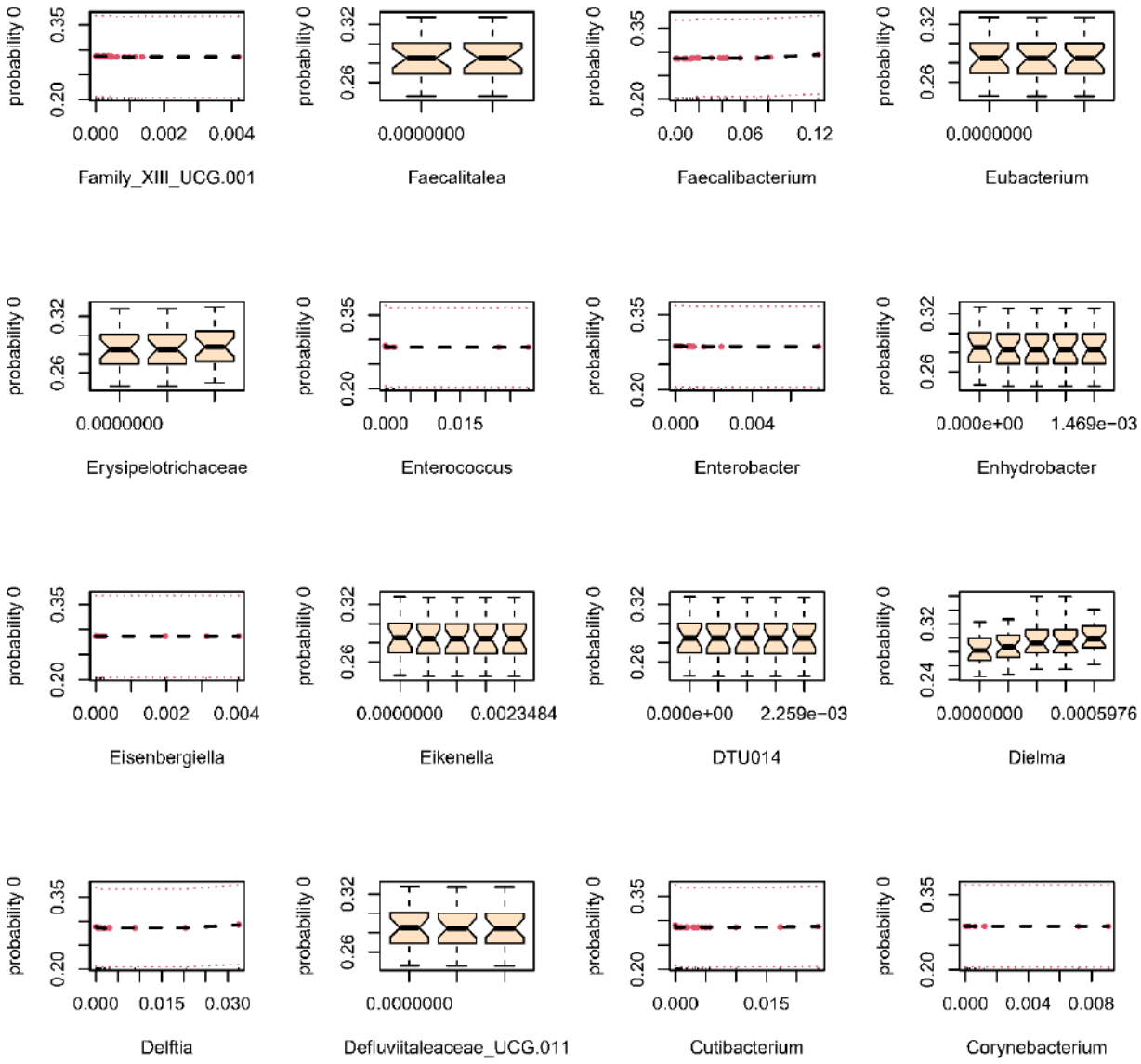

probability 0
0.20
0.35
0.0000
0.0020
Chloroplast
Butyricicoccus
0.000
0.006
0.012
Bilophila
0.00
0.15
Bifidobacterium
0.00
0.02
0.04
Barnesiella
Bacillus
0.26
0.32
0.000000
0.001538
Allisonella
0.03
Acinetobacter
0.015
K.Eubacterium._coprostanoligenes_
probability
UCG.003
Paracoccus
0.002
0.004
Lachnospira
0.08
X.Ruminococcus._gnavus_grou
0e+00
6e-04
X.Clostridium._methylpentosum_g
0.24
0.000e+00
1.010e-01
Pseudoalteromonas
0.030
Phascolarctobacterium

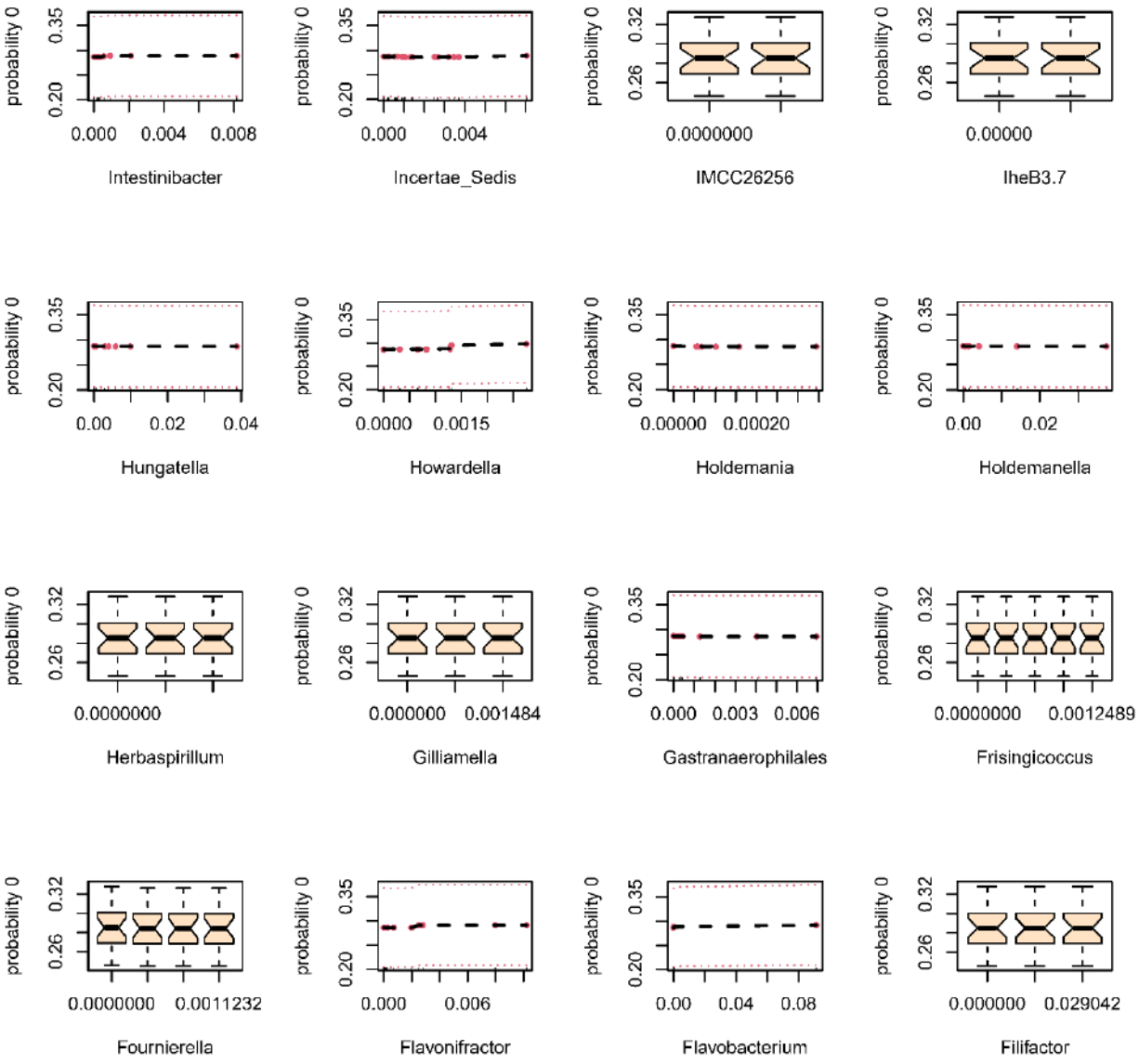

probability 0
0.26
0.32
0.0000000
Sanguibacteroides
probability 0
0.20
0.35
0.000
0.003
Romboutsia
probability 0
0.26
0.32
0.0000000
0.0014538
Prevotellaceae_UCG.001
probability 0
0.20
0.35
0.00
0.04
0.08
Oscillibacter
probability 0
0.20
0.35
0.0000
0.0010
0.0020
Methylobacterium.Methylorubru
probability 0
0.20
0.35
0.00
0.03
0.06
Lachnospiraceae_NK4A136_gro
probability 0
0.20
0.35
0.000
0.002
0.004
Lachnospiraceae_ND3007_grou
probability 0
0.20
0.35
0.000
0.015
Klebsiella
probability 0
0.20
0.35
0.000
0.010
Haemophilus
probability 0
0.20
0.35
0.000
0.004
Granulicatella
probability 0
0.20
0.35
0.00
0.06
0.12
GCA.900066575
probability 0
0.20
0.35
0.00
0.03
0.06
Erysipelatoclostridium
probability 0
0.20
0.35
0.000
0.015
Dorea
probability 0
0.20
0.35
0.00
0.02
0.04
Dialister
probability 0
0.20
0.35
0.000
0.006
Collinsella
probability 0
0.20
0.35
0.00
0.02
0.04
Clostridium_sensu_stricto_1

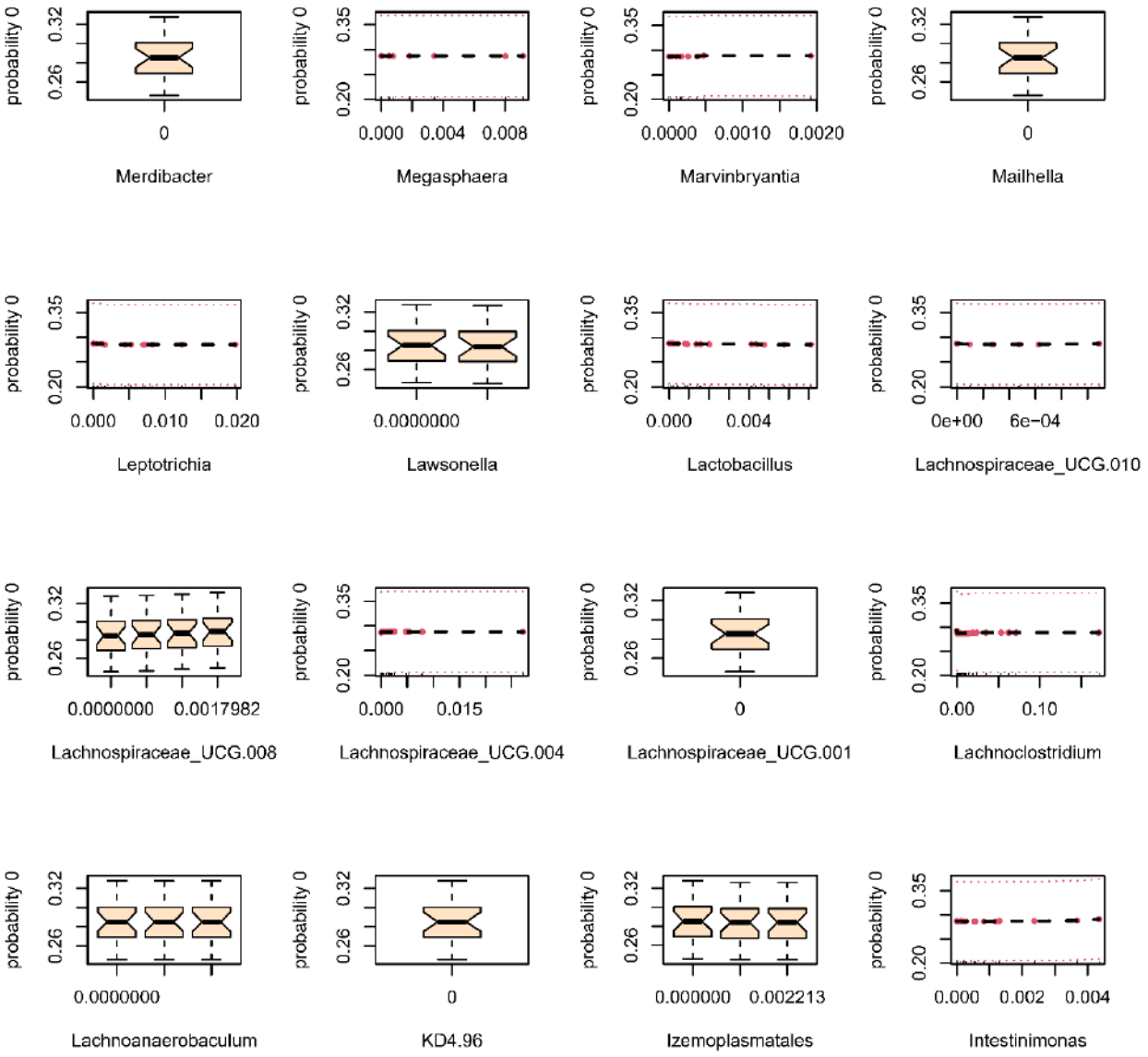

probability 0
0.20
0.35
0.000
0.010
Actinomyces
probability 0
0.26
0.32
0.0000000
Acidothermus
probability 0
0.26
0.32
0.0000000
0.0053528
Acidaminococcus
probability 0
0.26
0.32
0.000000
0.006012
X67.14
probability 0
0.20
0.35
0.0000
0.0010
X.Ruminococcus._gauvreauii_grc
probability 0
0.20
0.35
0.0000
0.0008
X.Eubacterium._xylanophilum_gr
probability 0
0.26
0.32
0.000000
0.003387
X.Eubacterium._ruminantium_grc
probability 0
0.20
0.35
0.000
0.004
0.008
X.Eubacterium._nodatum_grou
probability 0
0.26
0.32
0.0000000
0.0109426
X.Eubacterium._eligens_group
probability 0
0.26
0.32
0.000e+00
2.279e−02
X.Eubacterium._brachy_group
probability 0
0.20
0.35
0.0000
0.0020
X.Clostridium._innocuum_grou
probability 0
0.26
0.32
0
1
Gender
probability 0
0.20
0.35
0.000
0.002
0.004
Neisseria
probability 0
0.20
0.35
0.000
0.004
0.008
UCG.010
probability 0
0.20
0.35
0.00
0.03
Streptococcus
probability 0
0.20
0.35
0.0000
0.0015
Sphingomonas

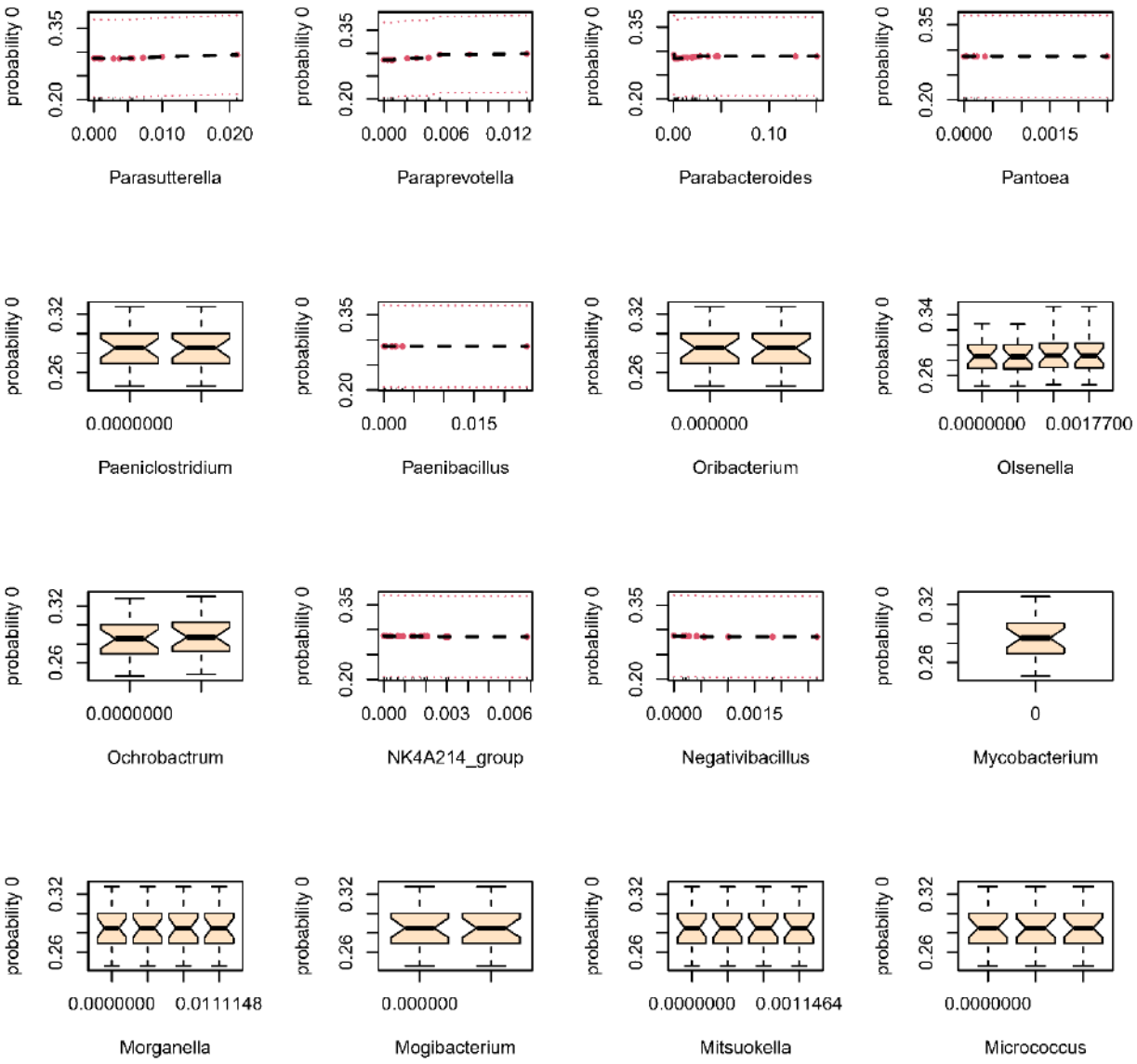

probability 0
0.26
0.32
0.000000
0.021057
Butyrivibrio
probability 0
0.20
0.35
0.000
0.006
0.012
Butyricimonas
probability 0
0.26
0.32
0.000000
Brachymonas
probability 0
0.26
0.32
0.000e+00
3.558e−03
Bosea
probability 0
0.26
0.32
0.0000000
0.0015949
Bergeyella
probability 0
0.26
0.32
0.000000
Bacteroidetes_BD2.2
probability 0
0.20
0.35
0.1
0.3
0.5
Bacteroides
probability 0
0.24
0.30
0.000e+00
1.669e−02
Atopobium
probability 0
0.26
0.32
0.0000000
Asteroleplasma
probability 0
0.20
0.35
0.0000
0.0010
Anaerotruncus
probability 0
0.20
0.35
0.000
0.015
Anaerostipes
probability 0
0.20
0.35
0.000
0.003
Anaerococcus
probability 0
0.20
0.35
0.000
0.006
0.012
zobium.Neorhizobium.Pararhizobiun
probability 0
0.20
0.35
0.00
0.10
Alloprevotella
probability 0
0.20
0.35
0.0000
0.0015
Aggregatibacter
probability 0
0.26
0.32
0.000000
0.020564
Aeromonas

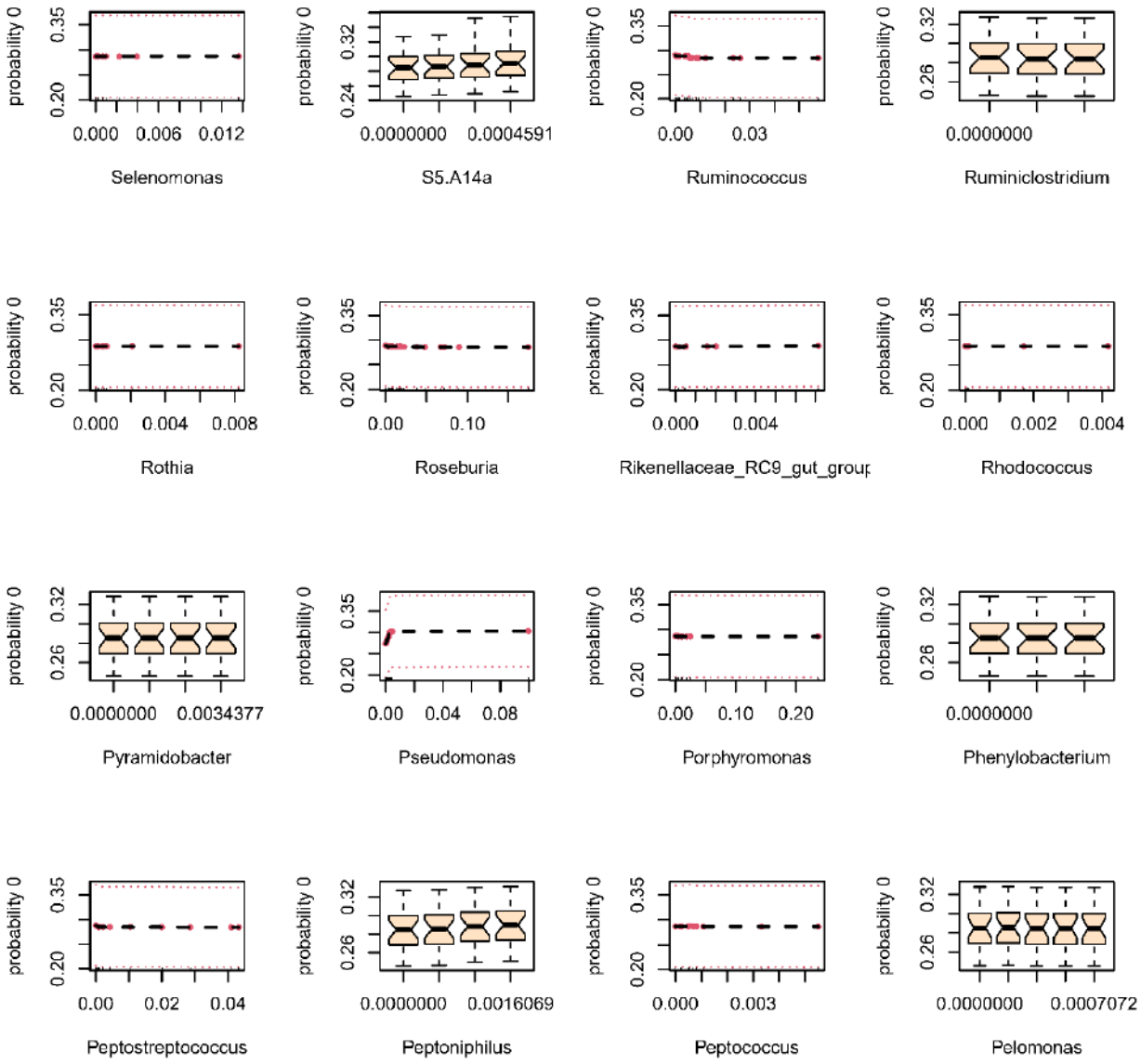

probability 0
0.20
0.35
0.000
0.004
Coprobacter
probability 0
0.26
0.32
0.0000000
0.0025547
Comamonas
probability 0
0.26
0.32
0.000000
Clostridium_sensu_stricto_2
probability 0
0.20
0.35
0.00
0.02
0.04
Clostridia_UCG.014
probability 0
0.26
0.32
0.0000000
0.0028351
Cloacibacterium
probability 0
0.26
0.32
0.000000
0.005388
Cloacibacillus
probability 0
0.26
0.32
0.0000000
0.0153201
Citrobacter
probability 0
0.26
0.32
0.000e+00
5.987e−04
Chryseobacterium
probability 0
0.26
0.32
0.0000000
Catonella
probability 0
0.24
0.30
0.0000000
0.0033189
Catenibacterium
probability 0
0.26
0.32
0.0000000
0.0017510
Catabacter
probability 0
0.20
0.35
0.0
0.2
Capnocytophaga
probability 0
0.20
0.35
0.000
0.003
0.006
Candidatus_Udaeobacter
probability 0
0.26
0.32
0.0000000
CAG.873
probability 0
0.20
0.35
0.000
0.003
0.006
CAG.56
probability 0
0.26
0.32
0
CAG.352

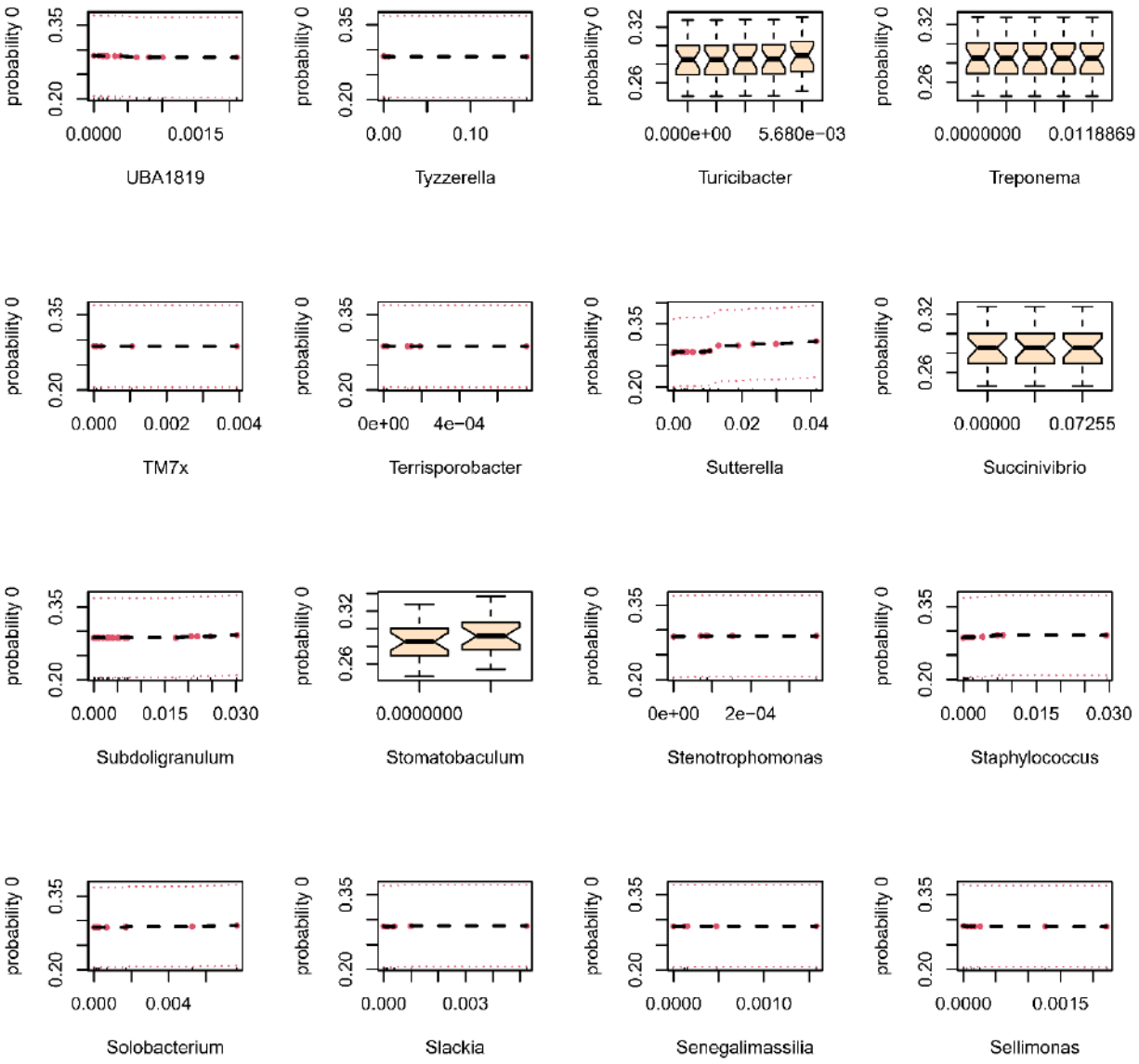

probability 0
Family_XIII_UCG.001
Faecalitalea
Faecalibacterium
Eubacterium
Erysipelotrichaceae
Enterococcus
Enterobacter
Enhydrobacter
Eisenbergiella
Eikenella
DTU014
Dielma
Delftia
Defluviitaleaceae_UCG.011
Cutibacterium
Corynebacterium

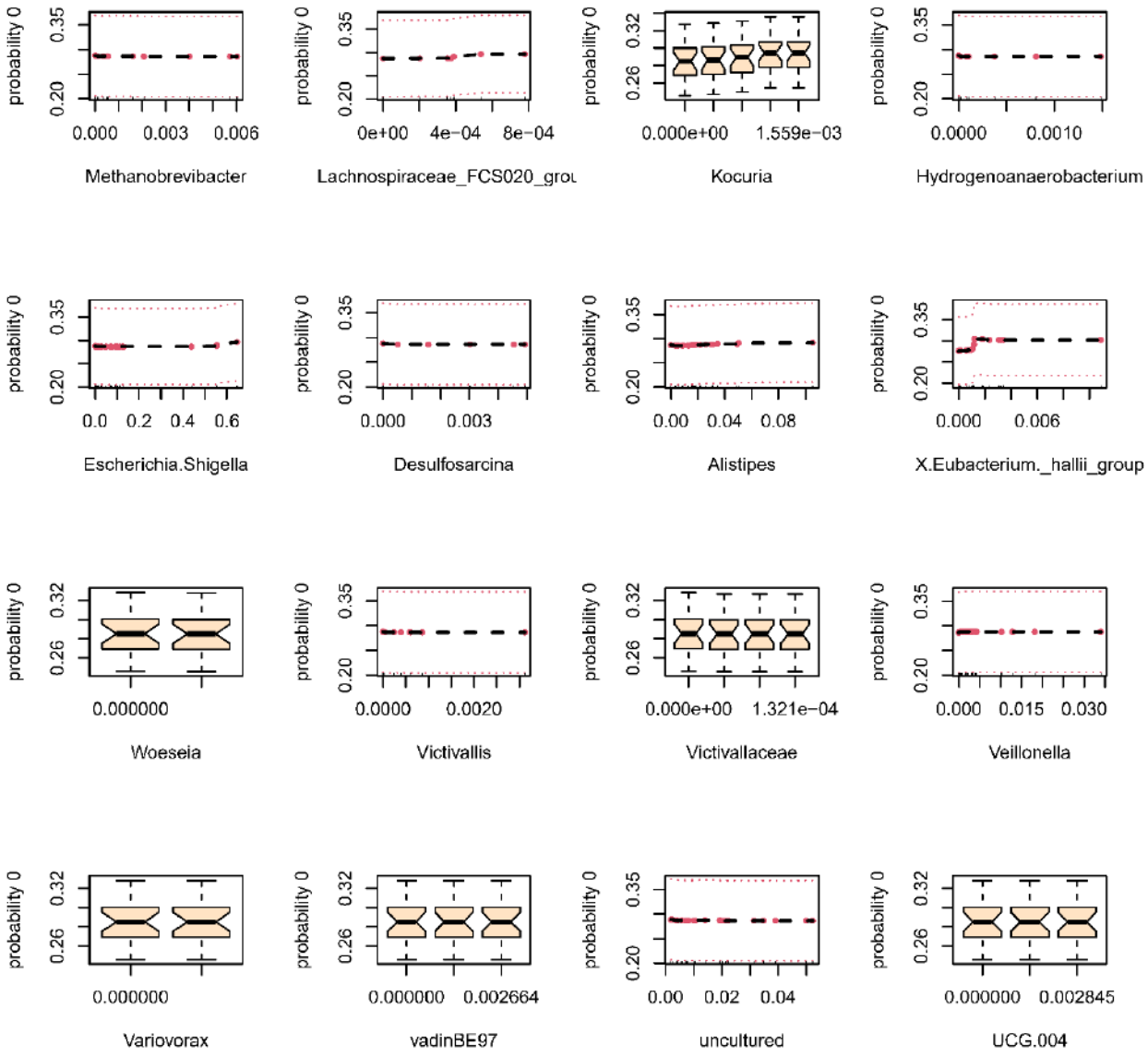

probability 0
0.20
0.35
0.000
0.004
0.008
Intestinibacter
probability 0
0.20
0.35
0.000
0.004
Incertae_Sedis
probability 0
0.26
0.32
0.0000000
IMCC26256
probability 0
0.26
0.32
0.00000
IheB3.7
probability 0
0.20
0.35
0.00
0.02
0.04
Hungatella
probability 0
0.20
0.35
0.0000
0.0015
Howardella
probability 0
0.20
0.35
0.00000
0.00020
Holdemania
probability 0
0.20
0.35
0.00
0.02
Holdemanella
probability 0
0.26
0.32
0.0000000
Herbaspirillum
probability 0
0.26
0.32
0.000000
0.001484
Gilliamella
probability 0
0.20
0.35
0.000
0.003
0.006
Gastranaerophilales
probability 0
0.26
0.32
0.0000000
0.0012489
Frisingicoccus
probability 0
0.26
0.32
0.0000000
0.0011232
Fournierella
probability 0
0.20
0.35
0.000
0.006
Flavonifractor
probability 0
0.20
0.35
0.00
0.04
0.08
Flavobacterium
probability 0
0.26
0.32
0.000000
0.029042
Filifactor

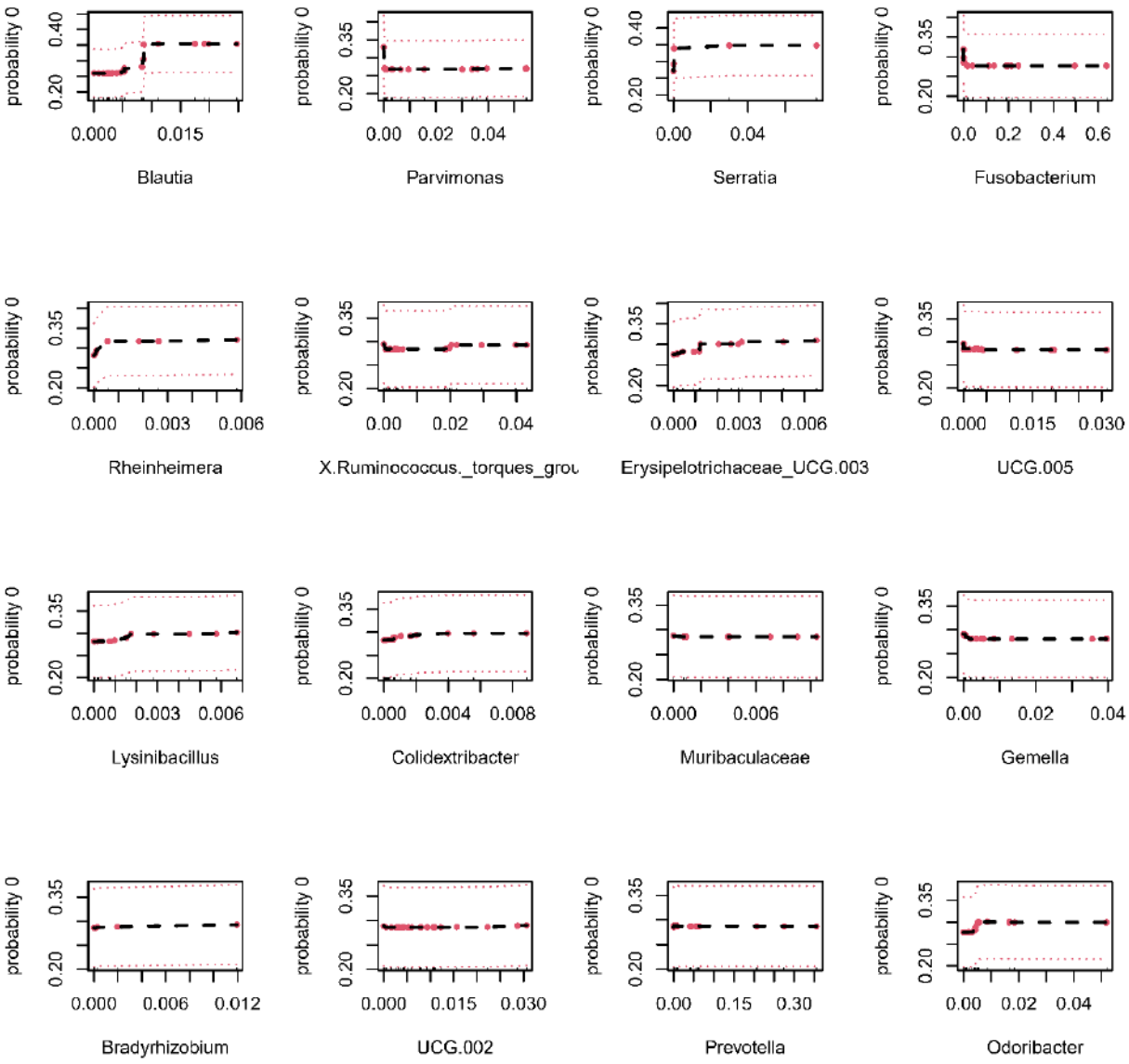

probability 0
0.26
0.32
0
Merdibacter
probability 0
0.20
0.35
0.000
0.004
0.008
Megasphaera
probability 0
0.20
0.35
0.0000
0.0010
0.0020
Marvinbryantia
probability 0
0.26
0.32
0
Mailhella
probability 0
0.20
0.35
0.000
0.010
0.020
Leptotrichia
probability 0
0.26
0.32
0.0000000
Lawsonella
probability 0
0.20
0.35
0.000
0.004
Lactobacillus
probability 0
0.20
0.35
0e+00
6e−04
Lachnospiraceae_UCG.010
probability 0
0.26
0.32
0.0000000
0.0017982
Lachnospiraceae_UCG.008
probability 0
0.20
0.35
0.000
0.015
Lachnospiraceae_UCG.004
probability 0
0.26
0.32
0
Lachnospiraceae_UCG.001
probability 0
0.20
0.35
0.00
0.10
Lachnoclostridium
probability 0
0.26
0.32
0.0000000
Lachnoanaerobaculum
probability 0
0.26
0.32
0
KD4.96
probability 0
0.26
0.32
0.000000
0.002213
Izemoplasmatales
probability 0
0.20
0.35
0.000
0.002
0.004
Intestinimonas


probability 0
0.20
0.35
0.000
0.010
0.020
Parasutterella
probability 0
0.20
0.35
0.000
0.006
0.012
Paraprevotella
probability 0
0.20
0.35
0.00
0.10
Parabacteroides
probability 0
0.20
0.35
0.0000
0.0015
Pantoea
probability 0
0.26
0.32
0.0000000
Paeniclostridium
probability 0
0.20
0.35
0.000
0.015
Paenibacillus
probability 0
0.26
0.32
0.000000
Oribacterium
probability 0
0.26
0.34
0.0000000
0.0017700
Olsenella
probability 0
0.26
0.32
0.0000000
Ochrobactrum
probability 0
0.20
0.35
0.000
0.003
0.006
NK4A214_group
probability 0
0.20
0.35
0.0000
0.0015
Negativibacillus
probability 0
0.26
0.32
0
Mycobacterium
probability 0
0.26
0.32
0.0000000
0.0111148
Morganella
probability 0
0.26
0.32
0.000000
Mogibacterium
probability 0
0.26
0.32
0.0000000
0.0011464
Mitsuokella
probability 0
0.26
0.32
0.0000000
Micrococcus


probability 0
0.20
0.35
0.000
0.006
0.012
Selenomonas
probability 0
0.24
0.32
0.0000000
0.0004591
S5.A14a
probability 0
0.20
0.35
0.00
0.03
Ruminococcus
probability 0
0.26
0.32
0.0000000
Ruminiclostridium
probability 0
0.20
0.35
0.000
0.004
0.008
Rothia
probability 0
0.20
0.35
0.00
0.10
Roseburia
probability 0
0.20
0.35
0.000
0.004
Rikenellaceae_RC9_gut_group
probability 0
0.20
0.35
0.000
0.002
0.004
Rhodococcus
probability 0
0.26
0.32
0.0000000
0.0034377
Pyramidobacter
probability 0
0.20
0.35
0.00
0.04
0.08
Pseudomonas
probability 0
0.20
0.35
0.00
0.10
0.20
Porphyromonas
probability 0
0.26
0.32
0.0000000
Phenylobacterium
probability 0
0.20
0.35
0.00
0.02
0.04
Peptostreptococcus
probability 0
0.26
0.32
0.0000000
0.0016069
Peptoniphilus
probability 0
0.20
0.35
0.000
0.003
Peptococcus
probability 0
0.26
0.32
0.0000000
0.0007072
Pelomonas


probability 0
0.20
0.35
0.0000
0.0015
UBA1819
probability 0
0.20
0.35
0.00
0.10
Tyzzerella
probability 0
0.26
0.32
0.000e+00
5.680e−03
Turicibacter
probability 0
0.26
0.32
0.0000000
0.0118869
Treponema
probability 0
0.20
0.35
0.000
0.002
0.004
TM7x
probability 0
0.20
0.35
0e+00
4e−04
Terrisporobacter
probability 0
0.20
0.35
0.00
0.02
0.04
Sutterella
probability 0
0.26
0.32
0.00000
0.07255
Succinivibrio
probability 0
0.20
0.35
0.000
0.015
0.030
Subdoligranulum
probability 0
0.26
0.32
0.0000000
Stomatobaculum
probability 0
0.20
0.35
0e+00
2e−04
Stenotrophomonas
probability 0
0.20
0.35
0.000
0.015
0.030
Staphylococcus
probability 0
0.20
0.35
0.000
0.004
Solobacterium
probability 0
0.20
0.35
0.000
0.003
Slackia
probability 0
0.20
0.35
0.0000
0.0010
Senegalimassilia
probability 0
0.20
0.35
0.0000
0.0015
Sellimonas


probability 0
0.20
0.35
0.000
0.003
0.006
Methanobrevibacter
probability 0
0.20
0.35
0e+00
4e−04
8e−04
Lachnospiraceae_FCS020_grou
probability 0
0.26
0.32
0.000e+00
1.559e−03
Kocuria
probability 0
0.20
0.35
0.0000
0.0010
Hydrogenoanaerobacterium
probability 0
0.20
0.35
0.0
0.2
0.4
0.6
Escherichia.Shigella
probability 0
0.20
0.35
0.000
0.003
Desulfosarcina
probability 0
0.20
0.35
0.00
0.04
0.08
Alistipes
probability 0
0.20
0.35
0.000
0.006
X.Eubacterium._hallii_group
probability 0
0.26
0.32
0.000000
Woeseia
probability 0
0.20
0.35
0.0000
0.0020
Victivallis
probability 0
0.26
0.32
0.000e+00
1.321e−04
Victivallaceae
probability 0
0.20
0.35
0.000
0.015
0.030
Veillonella
probability 0
0.26
0.32
0.000000
Variovorax
probability 0
0.26
0.32
0.000000
0.002664
vadinBE97
probability 0
0.20
0.35
0.00
0.02
0.04
uncultured
probability 0
0.26
0.32
0.000000
0.002845
UCG.004


probability 0
Blautia
Parvimonas
Serratia
Fusobacterium
Rheinheimera
X.Ruminococcus._torques_grou
Erysipelotrichaceae_UCG.003
UCG.005
Lysinibacillus
Colidextribacter
Muribaculaceae
Gemella
Bradyrhizobium
UCG.002
Prevotella
Odoribacter

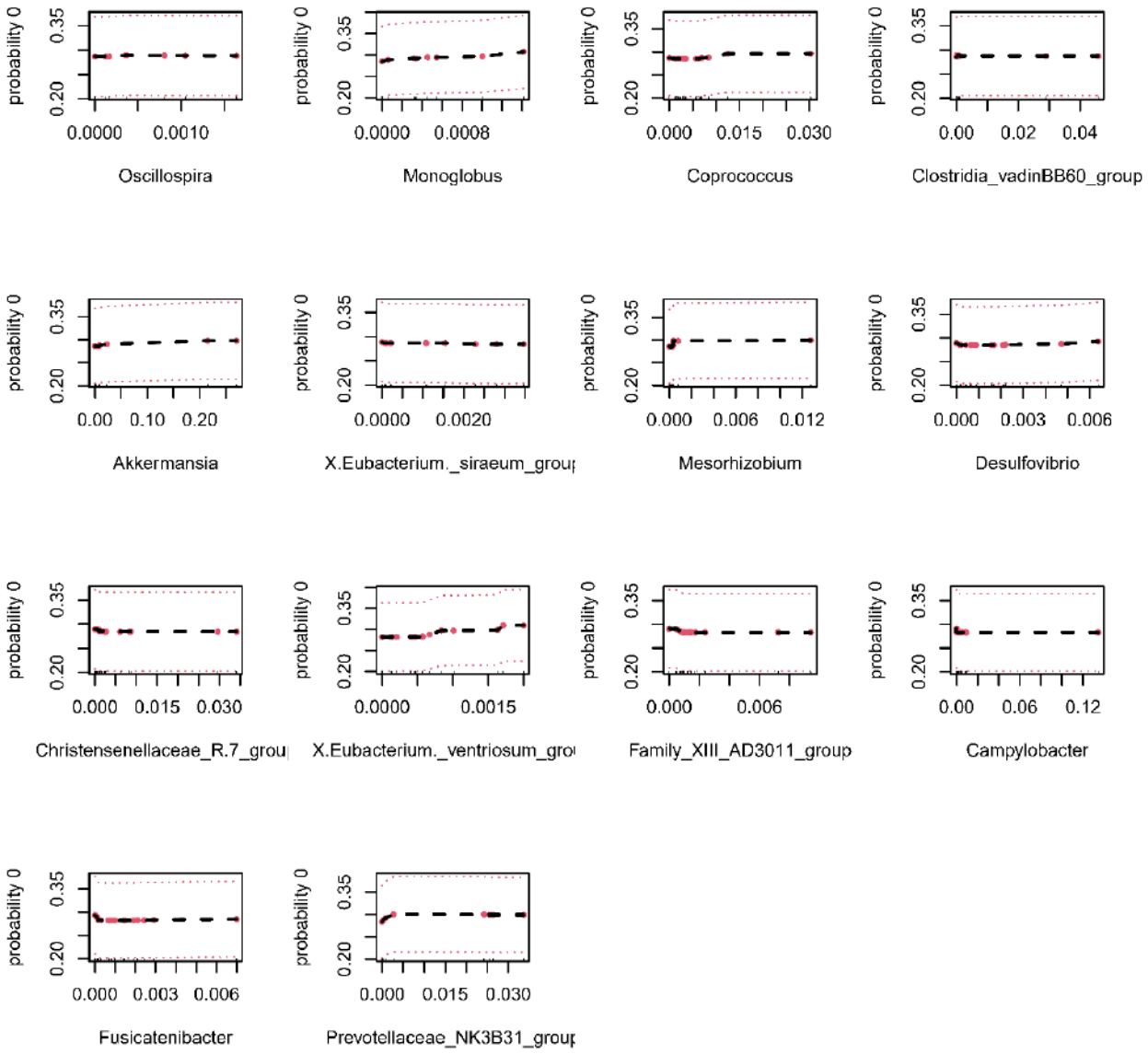


Figure ML1-E-3: Partial Plot for Genus in Biopsy

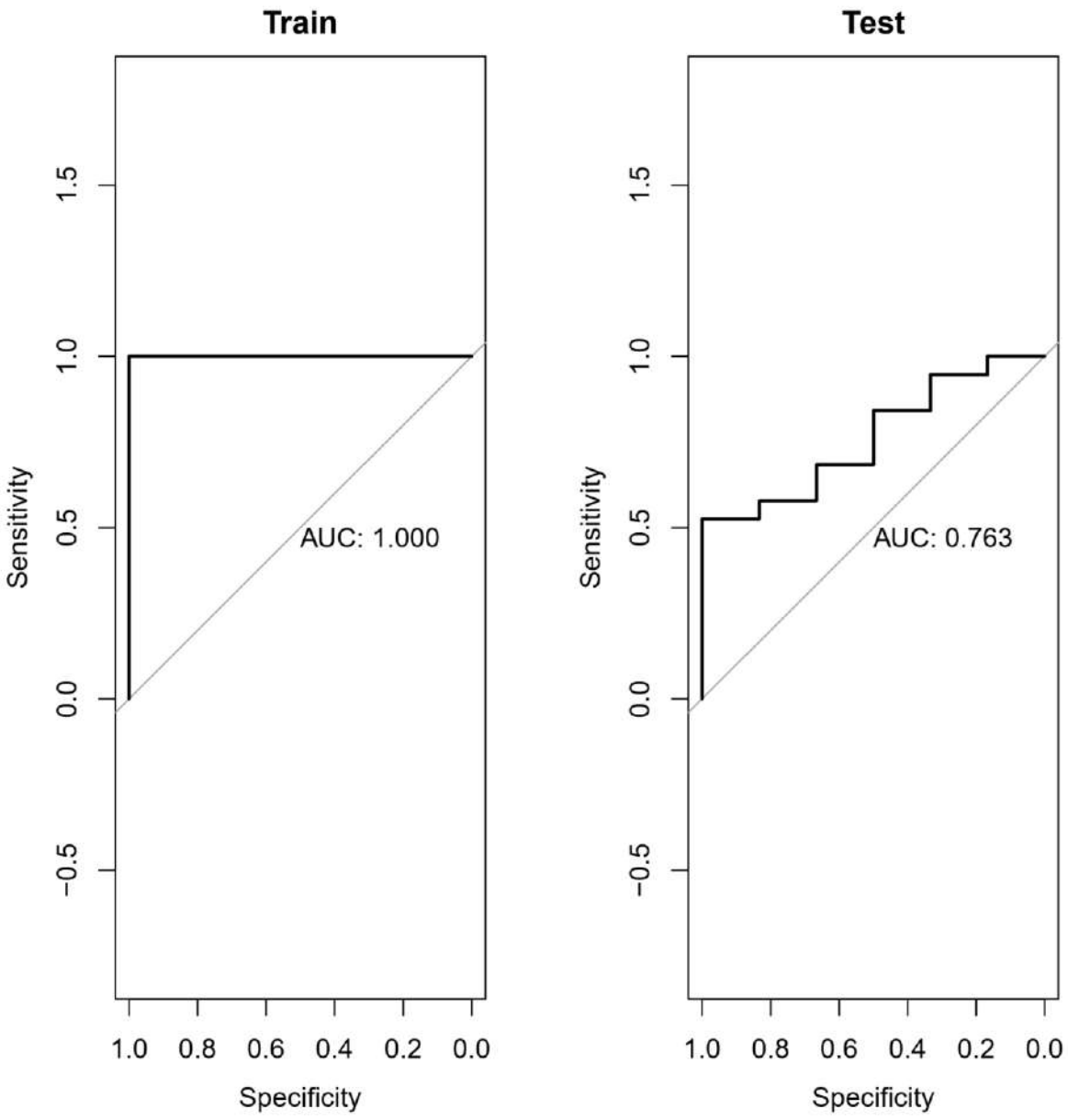


Figure ML1-E-4: The ROC Curve for Genus in Biopsy

Table ML1-E-1: The Model Performance

Index	Values		
	2.50%	50.00%	97.50%
Accuracy	0.52	0.72	0.92
Specificity	0.50	1.00	1.00
Sensitivity	0.37	0.63	1.00

Table ML1-E-2: The Variable Importance Values

Covariates	Variable Importance			Brier Score		
	all	AP	CRC	all	AP	CRC
Gender	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
X.Clostridium._innocuum_group	-0.0004	0.0045	-0.0029	-0.0004	-0.0003	-0.0003
X.Clostridium._methylpentosum_group	0.0004	0.0000	0.0014	-0.0038	-0.0026	-0.0026
X.Eubacterium._brachy_group	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
X.Eubacterium._coprostanoligenes_group	0.0004	-0.0227	0.0086	0.0018	0.0012	0.0012
X.Eubacterium._eligens_group	0.0012	0.0045	0.0029	0.0026	0.0018	0.0018
X.Eubacterium._hallii_group	-0.0001	0.0181	-0.0057	0.0019	0.0013	0.0013
X.Eubacterium._nodatum_group	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
X.Eubacterium._ruminantium_group	0.0000	0.0000	0.0000	0.0010	0.0007	0.0007
X.Eubacterium._siraeum_group	0.0000	0.0045	-0.0014	0.0033	0.0022	0.0022
X.Eubacterium._ventriosum_group	0.0034	-0.0362	0.0229	0.0080	0.0055	0.0055
X.Eubacterium._xylanophilum_group	0.0020	-0.0227	0.0143	0.0088	0.0060	0.0060
X.Ruminococcus._gauvreauii_group	-0.0004	-0.0045	0.0000	-0.0031	-0.0021	-0.0021
X.Ruminococcus._gnavus_group	0.0030	-0.0091	0.0129	0.0084	0.0057	0.0057
X.Ruminococcus._torques_group	0.0036	-0.0453	0.0272	0.0168	0.0114	0.0114
X67.14	0.0000	0.0000	0.0000	-0.0005	-0.0003	-0.0003
Acidaminococcus	0.0000	0.0045	-0.0014	0.0010	0.0007	0.0007
Acidothermus	0.0000	0.0000	0.0000	-0.0001	-0.0001	-0.0001
Acinetobacter	0.0015	0.0045	0.0043	-0.0012	-0.0008	-0.0008
Actinomyces	-0.0004	0.0000	-0.0014	-0.0009	-0.0006	-0.0006
Aeromonas	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Aggregatibacter	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Akkermansia	0.0036	-0.0181	0.0186	0.0116	0.0079	0.0079
Alistipes	-0.0008	-0.0091	0.0000	-0.0060	-0.0041	-0.0041
Allisonella	-0.0004	0.0000	-0.0014	-0.0024	-0.0016	-0.0016
Alloprevotella	0.0000	0.0000	0.0000	-0.0006	-0.0004	-0.0004
Allorhizobium.Neorhizobium.Pararhizobium.Rhizobium	0.0049	-0.0272	0.0258	0.0095	0.0065	0.0065
Anaerococcus	0.0012	0.0136	0.0000	0.0029	0.0020	0.0020
Anaerostipes	0.0008	-0.0136	0.0072	-0.0006	-0.0004	-0.0004
Anaerotruncus	0.0000	0.0000	0.0000	0.0012	0.0008	0.0008
Asteroleplasma	0.0000	0.0000	0.0000	0.0004	0.0002	0.0002

Atopobium	-0.0016	0.0000	-0.0057	-0.0010	-0.0007	-0.0007
Bacillus	0.0016	-0.0136	0.0100	0.0009	0.0006	0.0006
Bacteroides	-0.0024	0.0000	-0.0086	-0.0024	-0.0016	-0.0016
Bacteroidetes_BD2.2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Barnesiella	-0.0024	-0.0317	0.0014	-0.0049	-0.0033	-0.0033
Bergeyella	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Bifidobacterium	-0.0032	-0.0091	-0.0086	-0.0028	-0.0019	-0.0019
Bilophila	0.0107	-0.0680	0.0587	0.0242	0.0164	0.0164
Blautia	0.0610	0.1314	0.1745	0.1832	0.1245	0.1245
Bosea	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Brachymonas	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Bradyrhizobium	0.0012	-0.0181	0.0100	0.0081	0.0055	0.0055
Butyricicoccus	-0.0012	-0.0045	-0.0029	-0.0051	-0.0034	-0.0034
Butyricimonas	0.0028	-0.0317	0.0200	0.0005	0.0003	0.0003
Butyrivibrio	0.0013	0.0000	0.0043	-0.0009	-0.0006	-0.0006
CAG.352	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
CAG.56	-0.0008	-0.0181	0.0029	-0.0040	-0.0027	-0.0027
CAG.873	-0.0004	0.0000	-0.0014	-0.0009	-0.0006	-0.0006
Campylobacter	-0.0020	-0.0136	-0.0029	-0.0077	-0.0052	-0.0052
Candidatus_Udaeobacter	0.0008	-0.0136	0.0072	0.0051	0.0034	0.0034
Capnocytophaga	-0.0008	0.0000	-0.0029	-0.0003	-0.0002	-0.0002
Catabacter	0.0004	0.0045	0.0000	0.0010	0.0007	0.0007
Catenibacterium	-0.0025	0.0045	-0.0100	-0.0026	-0.0018	-0.0018
Catonella	0.0000	0.0000	0.0000	-0.0006	-0.0004	-0.0004
Chloroplast	0.0000	-0.0272	0.0086	0.0029	0.0020	0.0020
Christensenellaceae_R.7_group	-0.0013	-0.0045	-0.0029	-0.0074	-0.0050	-0.0050
Chryseobacterium	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Citrobacter	0.0004	0.0045	0.0000	0.0012	0.0008	0.0008
Cloacibacillus	0.0000	0.0000	0.0000	-0.0003	-0.0002	-0.0002
Cloacibacterium	0.0004	-0.0091	0.0043	-0.0004	-0.0003	-0.0003
Clostridia_UCG.014	0.0004	-0.0091	0.0043	-0.0034	-0.0023	-0.0023
Clostridia_vadinBB60_group	0.0004	0.0000	0.0014	-0.0013	-0.0009	-0.0009
Clostridium_sensu_stricto_1	-0.0012	-0.0045	-0.0029	-0.0010	-0.0007	-0.0007
Clostridium_sensu_stricto_2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Colidextribacter	-0.0024	-0.0272	0.0000	-0.0103	-0.0070	-0.0070
Collinsella	0.0016	0.0045	0.0043	0.0007	0.0005	0.0005
Comamonas	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Coprobacter	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Coprococcus	0.0033	-0.0317	0.0215	0.0031	0.0021	0.0021
Corynebacterium	-0.0004	0.0000	-0.0014	0.0005	0.0004	0.0004
Cutibacterium	0.0020	-0.0272	0.0157	0.0018	0.0012	0.0012
Defluviitaleaceae_UCG.011	0.0004	0.0045	0.0000	0.0026	0.0018	0.0018
Delftia	0.0036	0.0000	0.0129	0.0071	0.0049	0.0049

Desulfosarcina	-0.0016	0.0045	-0.0072	0.0026	0.0018	0.0018
Desulfovibrio	-0.0004	-0.0227	0.0057	-0.0035	-0.0024	-0.0024
Dialister	0.0004	0.0136	-0.0029	0.0033	0.0023	0.0023
Dielma	0.0033	-0.0317	0.0215	0.0049	0.0033	0.0033
Dorea	0.0032	-0.0408	0.0243	0.0078	0.0053	0.0053
DTU014	0.0000	0.0045	-0.0014	0.0014	0.0010	0.0010
Eikenella	-0.0008	0.0000	-0.0029	-0.0006	-0.0004	-0.0004
Eisenbergiella	0.0000	0.0000	0.0000	-0.0017	-0.0012	-0.0012
Enhydrobacter	-0.0029	0.0000	-0.0100	-0.0061	-0.0041	-0.0041
Enterobacter	-0.0016	0.0000	-0.0057	0.0006	0.0004	0.0004
Enterococcus	-0.0020	0.0045	-0.0086	-0.0056	-0.0038	-0.0038
Erysipelatoclostridium	0.0004	0.0000	0.0014	-0.0007	-0.0005	-0.0005
Erysipelotrichaceae	0.0016	-0.0045	0.0072	0.0019	0.0013	0.0013
Erysipelotrichaceae_UCG.003	0.0038	-0.0227	0.0200	0.0150	0.0102	0.0102
Escherichia.Shigella	0.0041	-0.0362	0.0258	0.0128	0.0087	0.0087
Eubacterium	0.0004	0.0045	0.0000	-0.0002	-0.0002	-0.0002
Faecalibacterium	0.0004	-0.0136	0.0057	0.0010	0.0007	0.0007
Faecalitalea	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Family_XIII_AD3011_group	-0.0024	-0.0091	-0.0057	-0.0018	-0.0012	-0.0012
Family_XIII_UCG.001	-0.0012	0.0000	-0.0043	-0.0040	-0.0027	-0.0027
Filifactor	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Flavobacterium	0.0016	-0.0091	0.0086	-0.0002	-0.0002	-0.0002
Flavonifractor	0.0020	-0.0091	0.0100	-0.0035	-0.0024	-0.0024
Fournierella	0.0000	0.0091	-0.0029	0.0013	0.0009	0.0009
Frisingicoccus	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Fusicatenibacter	0.0045	0.0000	0.0157	0.0136	0.0093	0.0093
Fusobacterium	0.0134	-0.0272	0.0558	0.0412	0.0280	0.0280
Gastranaerophilales	0.0000	0.0091	-0.0029	0.0004	0.0003	0.0003
GCA.900066575	0.0008	-0.0045	0.0043	-0.0006	-0.0004	-0.0004
Gemella	-0.0053	0.0091	-0.0215	-0.0075	-0.0051	-0.0051
Gilliamella	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Granulicatella	-0.0017	-0.0045	-0.0043	-0.0042	-0.0029	-0.0029
Haemophilus	-0.0012	-0.0272	0.0043	0.0005	0.0004	0.0004
Herbaspirillum	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Holdemanella	-0.0008	0.0045	-0.0043	-0.0011	-0.0008	-0.0008
Holdemania	0.0000	0.0045	-0.0014	0.0027	0.0019	0.0019
Howardella	0.0009	-0.0498	0.0186	0.0025	0.0017	0.0017
Hungatella	-0.0008	-0.0045	-0.0014	-0.0013	-0.0009	-0.0009
Hydrogenoanaerobacterium	-0.0004	0.0000	-0.0014	0.0021	0.0014	0.0014
IheB3.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
IMCC26256	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Incertae_Sedis	0.0016	-0.0091	0.0086	0.0051	0.0035	0.0035
Intestinibacter	0.0004	0.0000	0.0014	-0.0024	-0.0016	-0.0016

Intestinimonas	0.0000	-0.0091	0.0029	-0.0056	-0.0038	-0.0038
Izemoplasmatales	-0.0008	0.0000	-0.0029	0.0004	0.0003	0.0003
KD4.96	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Klebsiella	0.0012	-0.0272	0.0129	0.0039	0.0027	0.0027
Kocuria	0.0040	-0.0045	0.0157	0.0100	0.0068	0.0068
Lachnoanaerobaculum	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Lachnoclostridium	0.0028	-0.0091	0.0129	0.0076	0.0052	0.0052
Lachnospira	0.0024	-0.0227	0.0157	0.0077	0.0053	0.0053
Lachnospiraceae_FCS020_group	0.0044	-0.0045	0.0172	0.0084	0.0057	0.0057
Lachnospiraceae_ND3007_group	-0.0012	-0.0181	0.0014	-0.0004	-0.0003	-0.0003
Lachnospiraceae_NK4A136_group	-0.0008	-0.0227	0.0043	0.0038	0.0026	0.0026
Lachnospiraceae_UCG.001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Lachnospiraceae_UCG.004	-0.0012	0.0000	-0.0043	-0.0035	-0.0024	-0.0024
Lachnospiraceae_UCG.008	0.0024	-0.0091	0.0114	0.0021	0.0014	0.0014
Lachnospiraceae_UCG.010	0.0000	0.0000	0.0000	-0.0011	-0.0007	-0.0007
Lactobacillus	-0.0008	0.0000	-0.0029	-0.0015	-0.0010	-0.0010
Lawsonella	-0.0016	0.0045	-0.0072	-0.0026	-0.0018	-0.0018
Leptotrichia	-0.0032	0.0000	-0.0114	-0.0074	-0.0050	-0.0050
Lysinibacillus	-0.0012	-0.0136	0.0000	-0.0015	-0.0010	-0.0010
Mailhella	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Marvinbryantia	0.0008	0.0000	0.0029	0.0023	0.0016	0.0016
Megasphaera	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Merdibacter	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Mesorhizobium	0.0037	-0.0272	0.0215	0.0089	0.0060	0.0060
Methanobrevibacter	-0.0004	0.0045	-0.0029	-0.0022	-0.0015	-0.0015
Methylobacterium.Methylorubrum	-0.0004	0.0000	-0.0014	-0.0027	-0.0018	-0.0018
Micrococcus	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Mitsuokella	-0.0004	0.0000	-0.0014	-0.0001	-0.0001	-0.0001
Mogibacterium	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Monoglobus	0.0080	-0.0498	0.0444	0.0105	0.0071	0.0071
Morganella	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Muribaculaceae	0.0001	0.0091	-0.0029	0.0056	0.0038	0.0038
Mycobacterium	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Negativibacillus	-0.0017	0.0091	-0.0086	0.0013	0.0009	0.0009
Neisseria	-0.0016	-0.0045	-0.0043	-0.0033	-0.0023	-0.0023
NK4A214_group	-0.0008	0.0000	-0.0029	-0.0003	-0.0002	-0.0002
Ochrobactrum	0.0024	-0.0045	0.0100	0.0023	0.0016	0.0016
Odoribacter	0.0013	0.0136	0.0000	0.0061	0.0041	0.0041
Olsenella	0.0008	0.0000	0.0029	-0.0034	-0.0023	-0.0023
Oribacterium	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Oscillibacter	-0.0005	-0.0045	0.0000	-0.0023	-0.0016	-0.0016
Oscillospira	0.0004	0.0045	0.0000	-0.0031	-0.0021	-0.0021
Paenibacillus	-0.0004	0.0000	-0.0014	-0.0003	-0.0002	-0.0002

Paeniclostridium	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Pantoea	0.0004	0.0000	0.0014	0.0005	0.0003	0.0003
Parabacteroides	0.0048	-0.0045	0.0186	0.0034	0.0023	0.0023
Paracoccus	-0.0009	0.0045	-0.0043	-0.0024	-0.0017	-0.0017
Paraprevotella	0.0012	-0.0362	0.0157	0.0025	0.0017	0.0017
Parasutterella	0.0004	-0.0317	0.0114	-0.0052	-0.0036	-0.0036
Parvimonas	0.0296	0.0317	0.0944	0.1221	0.0830	0.0830
Pelomonas	-0.0008	0.0000	-0.0029	-0.0008	-0.0005	-0.0005
Peptococcus	-0.0012	-0.0136	0.0000	-0.0049	-0.0033	-0.0033
Peptoniphilus	0.0012	-0.0091	0.0072	0.0006	0.0004	0.0004
Peptostreptococcus	-0.0028	0.0045	-0.0114	0.0002	0.0001	0.0001
Phascolarctobacterium	0.0025	0.0181	0.0029	0.0086	0.0059	0.0059
Phenylobacterium	0.0000	0.0000	0.0000	-0.0003	-0.0002	-0.0002
Porphyromonas	0.0004	0.0045	0.0000	-0.0015	-0.0010	-0.0010
Prevotella	-0.0025	0.0000	-0.0086	-0.0068	-0.0046	-0.0046
Prevotellaceae_NK3B31_group	0.0016	-0.0227	0.0129	-0.0092	-0.0063	-0.0063
Prevotellaceae_UCG.001	0.0008	-0.0045	0.0043	-0.0005	-0.0004	-0.0004
Pseudoalteromonas	-0.0020	-0.0453	0.0072	0.0132	0.0090	0.0090
Pseudomonas	0.0033	0.0181	0.0057	0.0122	0.0083	0.0083
Pyramidobacter	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Rheinheimera	0.0161	-0.0589	0.0758	0.0518	0.0352	0.0352
Rhodococcus	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Rikenellaceae_RC9_gut_group	0.0008	0.0000	0.0029	-0.0004	-0.0003	-0.0003
Romboutsia	-0.0004	0.0000	-0.0014	-0.0005	-0.0003	-0.0003
Roseburia	0.0000	0.0045	-0.0014	0.0003	0.0002	0.0002
Rothia	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Ruminiclostridium	0.0004	0.0136	-0.0029	0.0014	0.0009	0.0009
Ruminococcus	0.0008	0.0091	0.0000	0.0049	0.0033	0.0033
S5.A14a	0.0024	-0.0045	0.0100	0.0093	0.0063	0.0063
Sanguibacteroides	0.0012	0.0000	0.0043	0.0015	0.0010	0.0010
Selenomonas	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Sellimonas	-0.0008	0.0000	-0.0029	-0.0013	-0.0009	-0.0009
Senegalimassilia	0.0004	0.0000	0.0014	0.0005	0.0003	0.0003
Serratia	0.0342	-0.0272	0.1288	0.0947	0.0644	0.0644
Slackia	-0.0012	-0.0136	0.0000	-0.0032	-0.0021	-0.0021
Solobacterium	0.0008	-0.0045	0.0043	0.0056	0.0038	0.0038
Sphingomonas	0.0000	0.0000	0.0000	-0.0041	-0.0028	-0.0028
Staphylococcus	0.0013	0.0000	0.0043	0.0031	0.0021	0.0021
Stenotrophomonas	0.0000	-0.0045	0.0014	0.0012	0.0008	0.0008
Stomatobaculum	0.0041	-0.0045	0.0157	0.0024	0.0016	0.0016
Streptococcus	-0.0024	-0.0091	-0.0057	-0.0037	-0.0025	-0.0025
Subdoligranulum	0.0016	0.0000	0.0057	0.0015	0.0010	0.0010
Succinivibrio	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Sutterella	0.0016	-0.0362	0.0172	0.0000	0.0000	0.0000
Terrisporobacter	-0.0004	0.0000	-0.0014	0.0009	0.0006	0.0006
TM7x	0.0000	0.0000	0.0000	-0.0003	-0.0002	-0.0002
Treponema	-0.0008	0.0000	-0.0029	-0.0008	-0.0005	-0.0005
Turicibacter	0.0004	-0.0181	0.0072	-0.0012	-0.0008	-0.0008
Tyzzerella	0.0000	0.0045	-0.0014	-0.0018	-0.0012	-0.0012
UBA1819	0.0017	0.0136	0.0014	-0.0025	-0.0017	-0.0017
UCG.002	0.0008	-0.0136	0.0072	0.0022	0.0015	0.0015
UCG.003	-0.0004	0.0045	-0.0029	-0.0022	-0.0015	-0.0015
UCG.004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
UCG.005	0.0028	0.0136	0.0057	0.0099	0.0067	0.0067
UCG.010	-0.0008	0.0045	-0.0043	-0.0027	-0.0018	-0.0018
uncultured	-0.0004	-0.0045	0.0000	0.0001	0.0001	0.0001
vadinBE97	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Variovorax	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Veillonella	-0.0004	-0.0045	0.0000	-0.0022	-0.0015	-0.0015
Victivallaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Victivallis	-0.0008	0.0000	-0.0029	-0.0012	-0.0008	-0.0008
Woeseia	0.0000	0.0000	0.0000	-0.0011	-0.0007	-0.0007

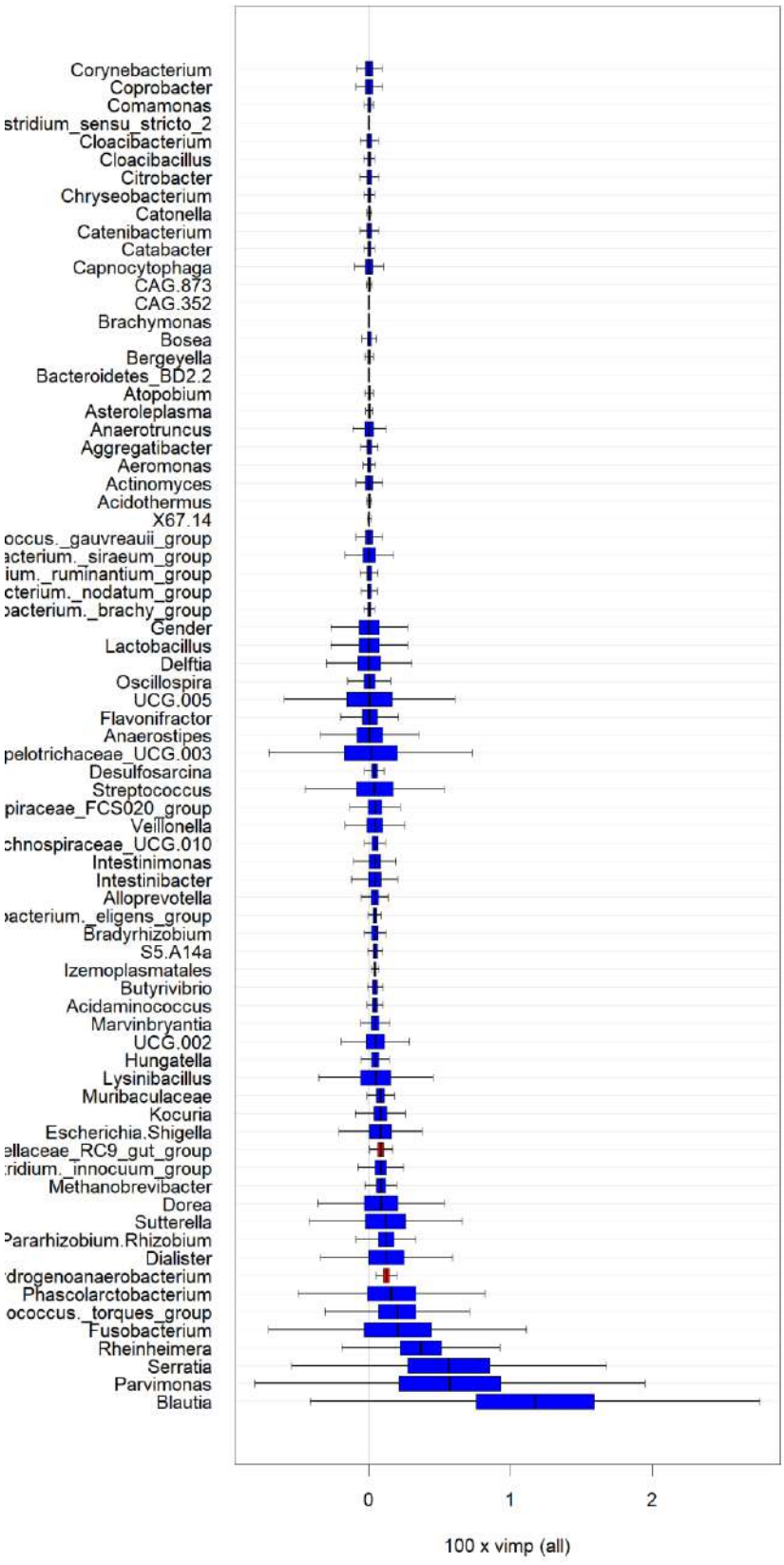

Corynebacterium
Coprobacter
Comamonas
stridium_sensu_stricto_2
Cloacibacterium
Cloacibacillus
Citrobacter
Chryseobacterium
Catonella
Catenibacterium
Catabacter
Capnocytophaga
CAG.873
CAG.352
Brachymonas
Bosea
Bergeyella
Bacteroidetes_BD2.2
Atopobium
Asteroleplasma
Anaerotruncus
Aggregatibacter
Aeromonas
Actinomyces
Acidothermus
X67.14
occus._gauvreauii_group
acterium._siraeum_group
ium._ruminantium_group
cterium._nodatum_group
bacterium._brachy_group
Gender
Lactobacillus
Delftia
Oscillospira
UCG.005
Flavonifractor
Anaerostipes
pelotrichaceae_UCG.003
Desulfosarcina
Streptococcus
piraceae_FCS020_group
Veillonella
chnospiraceae_UCG.010
Intestinimonas
Intestinibacter
Alloprevotella
bacterium._eligens_group
Bradyrhizobium
S5.A14a
Izemoplasmatales
Butyrivibrio
Acidaminococcus
Marvinbryantia
UCG.002
Hungatella
Lysinibacillus
Muribaculaceae
Kocuria
Escherichia.Shigella
ellaceae_RC9_gut_group
ridium._innocuum_group
Methanobrevibacter
Dorea
Sutterella
Pararhizobium.Rhizobium
Dialister
drogenoanaerobacterium
Phascolarctobacterium
ococcus._torques_group
Fusobacterium
Rheinheimera
Serratia
Parvimonas
Blautia
0
1
2
100 x vimp (all)

Figure ML1-E-5: The Variable Importance Plot (VIMP) for Genus in Biopsy

Best model is selected for LIME results

Table ML1-E-3: The best setting for model and LIME Results

Setting	mtry	ROC	Sens	Spec	ROCSD	SensSD	SpecSD
ntrees: 500 nodesize: 1	2	0.850	0.000	1.000	0.137	0.000	0.000
ntrees: 1000 nodesize: 1	2	0.850	0.000	1.000	0.137	0.000	0.000
ntrees: 500 nodesize: 5	4	0.825	0.000	1.000	0.168	0.000	0.000
ntrees: 1000 nodesize: 5	7	0.850	0.000	1.000	0.137	0.000	0.000

Table ML1-E-4: The LIME Result for Training Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
2	Adenoma	0.720	0.291	0.290	0.399	Blautia	0.020	0.039	0.00862 < Blautia
2	Adenoma	0.720	0.291	0.290	0.399	X.Ruminococcus._gnavus_group	0.012	0.018	0.00173 < X.Ruminococcus._gnavus_group
2	Adenoma	0.720	0.291	0.290	0.399	Phascolarctobacterium	0.029	0.018	0.00618 < Phascolarctobacterium
2	Adenoma	0.720	0.291	0.290	0.399	Fusobacterium	0.000	0.017	Fusobacterium <= 0.000518
2	Adenoma	0.720	0.291	0.290	0.399	Acinetobacter	0.009	0.017	0.003851 < Acinetobacter
3	Adenoma	0.625	0.277	0.328	0.366	Blautia	0.019	0.037	0.00862 < Blautia
3	Adenoma	0.625	0.277	0.328	0.366	X.Ruminococcus._gnavus_group	0.002	-0.020	X.Ruminococcus._gnavus_group <= 0.00173
3	Adenoma	0.625	0.277	0.328	0.366	Phascolarctobacterium	0.006	0.019	0.00618 < Phascolarctobacterium
3	Adenoma	0.625	0.277	0.328	0.366	Acinetobacter	0.004	0.019	0.003851 < Acinetobacter
3	Adenoma	0.625	0.277	0.328	0.366	Chloroplast	0.000	-0.016	Chloroplast <= 0.000153
6	Adenoma	0.621	0.256	0.318	0.360	Blautia	0.018	0.038	0.00862 < Blautia
6	Adenoma	0.621	0.256	0.318	0.360	X.Ruminococcus._gnavus_group	0.004	0.017	0.00173 < X.Ruminococcus._gnavus_group
6	Adenoma	0.621	0.256	0.318	0.360	Chloroplast	0.000	-0.014	Chloroplast <= 0.000153
6	Adenoma	0.621	0.256	0.318	0.360	Acinetobacter	0.000	-0.012	Acinetobacter <= 0.000286
6	Adenoma	0.621	0.256	0.318	0.360	X.Ruminococcus._torques_group	0.029	0.013	0.01976 < X.Ruminococcus._torques_group
7	Adenoma	0.685	0.269	0.336	0.343	Blautia	0.011	0.037	0.00862 < Blautia
7	Adenoma	0.685	0.269	0.336	0.343	X.Ruminococcus._gnavus_group	0.000	-0.020	X.Ruminococcus._gnavus_group <= 0.00173
7	Adenoma	0.685	0.269	0.336	0.343	Fusobacterium	0.000	0.017	Fusobacterium <= 0.000518
7	Adenoma	0.685	0.269	0.336	0.343	Chloroplast	0.000	-0.016	Chloroplast <= 0.000153
7	Adenoma	0.685	0.269	0.336	0.343	Acinetobacter	0.000	-0.013	Acinetobacter <= 0.000286
10	Adenoma	0.676	0.288	0.309	0.382	Blautia	0.009	0.039	0.00862 < Blautia
10	Adenoma	0.676	0.288	0.309	0.382	Fusobacterium	0.000	0.019	Fusobacterium <= 0.000518
10	Adenoma	0.676	0.288	0.309	0.382	X.Ruminococcus._gnavus_group	0.000	-0.018	X.Ruminococcus._gnavus_group <= 0.00173
10	Adenoma	0.676	0.288	0.309	0.382	Acinetobacter	0.006	0.018	0.003851 < Acinetobacter
10	Adenoma	0.676	0.288	0.309	0.382	Bacillus	0.002	0.015	0.000169 < Bacillus
11	Adenoma	0.586	0.130	0.322	0.322	Fusobacterium	0.000	0.018	Fusobacterium <= 0.000518
11	Adenoma	0.586	0.130	0.322	0.322	X.Ruminococcus._gnavus_group	0.059	0.017	0.00173 < X.Ruminococcus._gnavus_group

11	Adenoma	0.586	0.130	0.322	0.322	Blautia	0.000	-0.014	Blautia <= 0.00221
11	Adenoma	0.586	0.130	0.322	0.322	Phascolarctobacterium	0.000	-0.012	Phascolarctobacterium <= 0.00140
11	Adenoma	0.586	0.130	0.322	0.322	X.Eubacterium._hallii_group	0.000	-0.011	X.Eubacterium._hallii_group <= 0.000747
14	Adenocarcinoma	0.873	0.133	0.640	0.712	X.Ruminococcus._gnavus_group	0.000	0.020	X.Ruminococcus._gnavus_group <= 0.00173
14	Adenocarcinoma	0.873	0.133	0.640	0.712	Acinetobacter	0.000	0.014	Acinetobacter <= 0.000286
14	Adenocarcinoma	0.873	0.133	0.640	0.712	X.Eubacterium._hallii_group	0.000	0.013	X.Eubacterium._hallii_group <= 0.000747
14	Adenocarcinoma	0.873	0.133	0.640	0.712	Chloroplast	0.000	0.014	Chloroplast <= 0.000153
14	Adenocarcinoma	0.873	0.133	0.640	0.712	Blautia	0.000	0.012	Blautia <= 0.00221
20	Adenocarcinoma	0.884	0.119	0.640	0.688	X.Ruminococcus._gnavus_group	0.000	0.019	X.Ruminococcus._gnavus_group <= 0.00173
20	Adenocarcinoma	0.884	0.119	0.640	0.688	Chloroplast	0.000	0.017	Chloroplast <= 0.000153
20	Adenocarcinoma	0.884	0.119	0.640	0.688	Bacillus	0.000	0.013	Bacillus <= 0.000169
20	Adenocarcinoma	0.884	0.119	0.640	0.688	X.Eubacterium._hallii_group	0.003	-0.013	0.001810 < X.Eubacterium._hallii_group
20	Adenocarcinoma	0.884	0.119	0.640	0.688	Paraprevotella	0.000	0.012	Paraprevotella <= 0.000891
21	Adenocarcinoma	0.877	0.125	0.673	0.658	X.Ruminococcus._gnavus_group	0.000	0.018	X.Ruminococcus._gnavus_group <= 0.00173
21	Adenocarcinoma	0.877	0.125	0.673	0.658	Chloroplast	0.000	-0.017	0.000153 < Chloroplast
21	Adenocarcinoma	0.877	0.125	0.673	0.658	Pseudomonas	0.004	-0.015	0.002448 < Pseudomonas
21	Adenocarcinoma	0.877	0.125	0.673	0.658	Phascolarctobacterium	0.000	0.011	Phascolarctobacterium <= 0.00140
21	Adenocarcinoma	0.877	0.125	0.673	0.658	Mesorhizobium	0.000	-0.013	0.000147 < Mesorhizobium
22	Adenocarcinoma	0.944	0.142	0.637	0.715	X.Ruminococcus._gnavus_group	0.000	0.021	X.Ruminococcus._gnavus_group <= 0.00173
22	Adenocarcinoma	0.944	0.142	0.637	0.715	Phascolarctobacterium	0.000	0.014	Phascolarctobacterium <= 0.00140
22	Adenocarcinoma	0.944	0.142	0.637	0.715	X.Eubacterium._hallii_group	0.001	0.013	X.Eubacterium._hallii_group <= 0.000747
22	Adenocarcinoma	0.944	0.142	0.637	0.715	Blautia	0.001	0.014	Blautia <= 0.00221
22	Adenocarcinoma	0.944	0.142	0.637	0.715	Chloroplast	0.000	0.015	Chloroplast <= 0.000153
23	Adenocarcinoma	0.890	0.132	0.653	0.695	X.Ruminococcus._gnavus_group	0.000	0.019	X.Ruminococcus._gnavus_group <= 0.00173
23	Adenocarcinoma	0.890	0.132	0.653	0.695	Phascolarctobacterium	0.000	0.014	Phascolarctobacterium <= 0.00140
23	Adenocarcinoma	0.890	0.132	0.653	0.695	Pseudomonas	0.003	-0.015	0.002448 < Pseudomonas
23	Adenocarcinoma	0.890	0.132	0.653	0.695	Acinetobacter	0.000	0.012	Acinetobacter <= 0.000286
23	Adenocarcinoma	0.890	0.132	0.653	0.695	X.Eubacterium._hallii_group	0.000	0.012	X.Eubacterium._hallii_group <= 0.000747
24	Adenocarcinoma	0.893	0.123	0.641	0.713	X.Ruminococcus._gnavus_group	0.000	0.019	X.Ruminococcus._gnavus_group <= 0.00173
24	Adenocarcinoma	0.893	0.123	0.641	0.713	Blautia	0.000	0.015	Blautia <= 0.00221
24	Adenocarcinoma	0.893	0.123	0.641	0.713	Acinetobacter	0.000	0.012	Acinetobacter <= 0.000286
24	Adenocarcinoma	0.893	0.123	0.641	0.713	Chloroplast	0.000	0.014	Chloroplast <= 0.000153
24	Adenocarcinoma	0.893	0.123	0.641	0.713	X.Eubacterium._hallii_group	0.000	0.012	X.Eubacterium._hallii_group <= 0.000747
25	Adenocarcinoma	0.906	0.136	0.652	0.689	Fusobacterium	0.000	-0.018	Fusobacterium <= 0.000518
25	Adenocarcinoma	0.906	0.136	0.652	0.689	X.Ruminococcus._gnavus_group	0.000	0.017	X.Ruminococcus._gnavus_group <= 0.00173
25	Adenocarcinoma	0.906	0.136	0.652	0.689	Chloroplast	0.000	0.015	Chloroplast <= 0.000153
25	Adenocarcinoma	0.906	0.136	0.652	0.689	X.Eubacterium._hallii_group	0.000	0.012	X.Eubacterium._hallii_group <= 0.000747
25	Adenocarcinoma	0.906	0.136	0.652	0.689	Phascolarctobacterium	0.001	0.011	Phascolarctobacterium <= 0.00140
26	Adenocarcinoma	0.833	0.128	0.652	0.697	X.Ruminococcus._gnavus_group	0.000	0.021	X.Ruminococcus._gnavus_group <= 0.00173
26	Adenocarcinoma	0.833	0.128	0.652	0.697	Phascolarctobacterium	0.000	0.013	Phascolarctobacterium <= 0.00140
26	Adenocarcinoma	0.833	0.128	0.652	0.697	Blautia	0.004	0.014	0.00221 < Blautia <= 0.00462
26	Adenocarcinoma	0.833	0.128	0.652	0.697	Bacteroides	0.291	-0.014	0.221 < Bacteroides
26	Adenocarcinoma	0.833	0.128	0.652	0.697	Bacillus	0.000	0.012	Bacillus <= 0.000169

29	Adenocarcinoma	0.741	0.142	0.677	0.677	X.Ruminococcus._gnavus_group	0.101	-0.019	0.00173 < X.Ruminococcus._gnavus_group
29	Adenocarcinoma	0.741	0.142	0.677	0.677	Acinetobacter	0.033	-0.019	0.003851 < Acinetobacter
29	Adenocarcinoma	0.741	0.142	0.677	0.677	Phascolarctobacterium	0.000	0.013	Phascolarctobacterium <= 0.00140
29	Adenocarcinoma	0.741	0.142	0.677	0.677	X.Eubacterium._hallii_group	0.000	0.012	X.Eubacterium._hallii_group <= 0.000747
29	Adenocarcinoma	0.741	0.142	0.677	0.677	Blautia	0.003	0.013	0.00221 < Blautia <= 0.00462
32	Adenocarcinoma	0.903	0.141	0.654	0.693	X.Ruminococcus._gnavus_group	0.000	0.019	X.Ruminococcus._gnavus_group <= 0.00173
32	Adenocarcinoma	0.903	0.141	0.654	0.693	Acinetobacter	0.012	-0.019	0.003851 < Acinetobacter
32	Adenocarcinoma	0.903	0.141	0.654	0.693	Blautia	0.001	0.015	Blautia <= 0.00221
32	Adenocarcinoma	0.903	0.141	0.654	0.693	Phascolarctobacterium	0.001	0.011	Phascolarctobacterium <= 0.00140
32	Adenocarcinoma	0.903	0.141	0.654	0.693	Chloroplast	0.000	0.012	Chloroplast <= 0.000153
33	Adenocarcinoma	0.854	0.147	0.668	0.670	X.Ruminococcus._gnavus_group	0.002	-0.019	0.00173 < X.Ruminococcus._gnavus_group
33	Adenocarcinoma	0.854	0.147	0.668	0.670	Phascolarctobacterium	0.008	-0.019	0.00618 < Phascolarctobacterium
33	Adenocarcinoma	0.854	0.147	0.668	0.670	Chloroplast	0.000	0.015	Chloroplast <= 0.000153
33	Adenocarcinoma	0.854	0.147	0.668	0.670	Acinetobacter	0.000	0.012	Acinetobacter <= 0.000286
33	Adenocarcinoma	0.854	0.147	0.668	0.670	X.Eubacterium._hallii_group	0.000	0.012	X.Eubacterium._hallii_group <= 0.000747
34	Adenocarcinoma	0.826	0.126	0.664	0.676	X.Ruminococcus._gnavus_group	0.000	0.019	X.Ruminococcus._gnavus_group <= 0.00173
34	Adenocarcinoma	0.826	0.126	0.664	0.676	Acinetobacter	0.056	-0.016	0.003851 < Acinetobacter
34	Adenocarcinoma	0.826	0.126	0.664	0.676	Phascolarctobacterium	0.000	0.012	Phascolarctobacterium <= 0.00140
34	Adenocarcinoma	0.826	0.126	0.664	0.676	Chloroplast	0.003	-0.014	0.000153 < Chloroplast
34	Adenocarcinoma	0.826	0.126	0.664	0.676	X.Eubacterium._hallii_group	0.000	0.011	X.Eubacterium._hallii_group <= 0.000747
39	Adenocarcinoma	0.728	0.283	0.664	0.655	Blautia	0.025	-0.039	0.00862 < Blautia
39	Adenocarcinoma	0.728	0.283	0.664	0.655	X.Ruminococcus._gnavus_group	0.000	0.019	X.Ruminococcus._gnavus_group <= 0.00173
39	Adenocarcinoma	0.728	0.283	0.664	0.655	Phascolarctobacterium	0.021	-0.018	0.00618 < Phascolarctobacterium
39	Adenocarcinoma	0.728	0.283	0.664	0.655	Chloroplast	0.000	0.015	Chloroplast <= 0.000153
39	Adenocarcinoma	0.728	0.283	0.664	0.655	Acinetobacter	0.000	0.013	Acinetobacter <= 0.000286
40	Adenocarcinoma	0.830	0.127	0.673	0.681	Phascolarctobacterium	0.000	0.014	Phascolarctobacterium <= 0.00140
40	Adenocarcinoma	0.830	0.127	0.673	0.681	X.Ruminococcus._gnavus_group	0.011	-0.016	0.00173 < X.Ruminococcus._gnavus_group
40	Adenocarcinoma	0.830	0.127	0.673	0.681	Chloroplast	0.000	-0.015	0.000153 < Chloroplast
40	Adenocarcinoma	0.830	0.127	0.673	0.681	Blautia	0.003	0.014	0.00221 < Blautia <= 0.00462
40	Adenocarcinoma	0.830	0.127	0.673	0.681	Acinetobacter	0.000	0.012	Acinetobacter <= 0.000286
46	Adenocarcinoma	0.823	0.157	0.654	0.699	X.Ruminococcus._gnavus_group	0.000	0.019	X.Ruminococcus._gnavus_group <= 0.00173
46	Adenocarcinoma	0.823	0.157	0.654	0.699	Fusobacterium	0.000	-0.017	Fusobacterium <= 0.000518
46	Adenocarcinoma	0.823	0.157	0.654	0.699	Phascolarctobacterium	0.000	0.014	Phascolarctobacterium <= 0.00140
46	Adenocarcinoma	0.823	0.157	0.654	0.699	Acinetobacter	0.000	0.014	Acinetobacter <= 0.000286
46	Adenocarcinoma	0.823	0.157	0.654	0.699	Blautia	0.002	0.014	Blautia <= 0.00221
47	Adenocarcinoma	0.885	0.130	0.637	0.710	X.Ruminococcus._gnavus_group	0.001	0.019	X.Ruminococcus._gnavus_group <= 0.00173
47	Adenocarcinoma	0.885	0.130	0.637	0.710	Bacillus	0.000	0.015	Bacillus <= 0.000169
47	Adenocarcinoma	0.885	0.130	0.637	0.710	Blautia	0.002	0.014	Blautia <= 0.00221
47	Adenocarcinoma	0.885	0.130	0.637	0.710	X.Eubacterium._hallii_group	0.000	0.012	X.Eubacterium._hallii_group <= 0.000747
47	Adenocarcinoma	0.885	0.130	0.637	0.710	Chloroplast	0.000	0.013	Chloroplast <= 0.000153
48	Adenocarcinoma	0.879	0.152	0.659	0.674	X.Ruminococcus._gnavus_group	0.000	0.020	X.Ruminococcus._gnavus_group <= 0.00173
48	Adenocarcinoma	0.879	0.152	0.659	0.674	Fusobacterium	0.001	-0.017	Fusobacterium <= 0.000518
48	Adenocarcinoma	0.879	0.152	0.659	0.674	Acinetobacter	0.000	0.013	Acinetobacter <= 0.000286

48	Adenocarcinoma	0.879	0.152	0.659	0.674	Phascolarctobacterium	0.010	-0.015	0.00618 < Phascolarctobacterium
48	Adenocarcinoma	0.879	0.152	0.659	0.674	Chloroplast	0.000	0.014	Chloroplast <= 0.000153
49	Adenocarcinoma	0.922	0.130	0.639	0.708	X.Ruminococcus._gnavus_group	0.002	0.017	X.Ruminococcus._gnavus_group <= 0.00173
49	Adenocarcinoma	0.922	0.130	0.639	0.708	X.Eubacterium._hallii_group	0.001	0.015	X.Eubacterium._hallii_group <= 0.000747
49	Adenocarcinoma	0.922	0.130	0.639	0.708	Phascolarctobacterium	0.000	0.012	Phascolarctobacterium <= 0.00140
49	Adenocarcinoma	0.922	0.130	0.639	0.708	Chloroplast	0.000	0.014	Chloroplast <= 0.000153
49	Adenocarcinoma	0.922	0.130	0.639	0.708	Acinetobacter	0.000	0.011	Acinetobacter <= 0.000286
50	Adenocarcinoma	0.870	0.156	0.673	0.654	Phascolarctobacterium	0.011	-0.020	0.00618 < Phascolarctobacterium
50	Adenocarcinoma	0.870	0.156	0.673	0.654	X.Ruminococcus._gnavus_group	0.000	0.019	X.Ruminococcus._gnavus_group <= 0.00173
50	Adenocarcinoma	0.870	0.156	0.673	0.654	Acinetobacter	0.000	0.014	Acinetobacter <= 0.000286
50	Adenocarcinoma	0.870	0.156	0.673	0.654	X.Eubacterium._hallii_group	0.002	-0.015	0.001810 < X.Eubacterium._hallii_group
50	Adenocarcinoma	0.87	0.155918077	0.672526975	0.654090175	X.Ruminococcus._torques_group	0.043123504	-0.015412481	0.01976 < X.Ruminococcus._torques_group

Table ML-E-5: The LIME Result for Test Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
1	Adenocarcinoma	0.751	0.269	0.662	0.656	Blautia	0.010	-0.037	0.00862 < Blautia
1	Adenocarcinoma	0.751	0.269	0.662	0.656	X.Ruminococcus._gnavus_group	0.002	0.019	X.Ruminococcus._gnavus_group <= 0.00173
1	Adenocarcinoma	0.751	0.269	0.662	0.656	Phascolarctobacterium	0.017	-0.016	0.00618 < Phascolarctobacterium
1	Adenocarcinoma	0.751	0.269	0.662	0.656	Chloroplast	0.000	0.016	Chloroplast <= 0.000153
1	Adenocarcinoma	0.751	0.269	0.662	0.656	Acinetobacter	0.000	0.013	Acinetobacter <= 0.000286
4	Adenocarcinoma	0.691	0.162	0.675	0.654	X.Ruminococcus._gnavus_group	0.000	0.019	X.Ruminococcus._gnavus_group <= 0.00173
4	Adenocarcinoma	0.691	0.162	0.675	0.654	Acinetobacter	0.011	-0.019	0.003851 < Acinetobacter
4	Adenocarcinoma	0.691	0.162	0.675	0.654	Phascolarctobacterium	0.021	-0.018	0.00618 < Phascolarctobacterium
4	Adenocarcinoma	0.691	0.162	0.675	0.654	Fusobacterium	0.000	-0.016	Fusobacterium <= 0.000518
4	Adenocarcinoma	0.691	0.162	0.675	0.654	X.Eubacterium._hallii_group	0.000	0.013	X.Eubacterium._hallii_group <= 0.000747
5	Adenocarcinoma	0.575	0.165	0.686	0.638	Fusobacterium	0.000	-0.019	Fusobacterium <= 0.000518
5	Adenocarcinoma	0.575	0.165	0.686	0.638	X.Ruminococcus._gnavus_group	0.000	0.018	X.Ruminococcus._gnavus_group <= 0.00173
5	Adenocarcinoma	0.575	0.165	0.686	0.638	Phascolarctobacterium	0.009	-0.016	0.00618 < Phascolarctobacterium
5	Adenocarcinoma	0.575	0.165	0.686	0.638	Acinetobacter	0.006	-0.016	0.003851 < Acinetobacter
5	Adenocarcinoma	0.575	0.165	0.686	0.638	X.Ruminococcus._torques_group	0.038	-0.015	0.01976 < X.Ruminococcus._torques_group
8	Adenocarcinoma	0.653	0.281	0.666	0.653	Blautia	0.015	-0.040	0.00862 < Blautia
8	Adenocarcinoma	0.653	0.281	0.666	0.653	X.Ruminococcus._gnavus_group	0.000	0.018	X.Ruminococcus._gnavus_group <= 0.00173
8	Adenocarcinoma	0.653	0.281	0.666	0.653	Fusobacterium	0.000	-0.017	Fusobacterium <= 0.000518
8	Adenocarcinoma	0.653	0.281	0.666	0.653	Chloroplast	0.000	0.015	Chloroplast <= 0.000153
8	Adenocarcinoma	0.653	0.281	0.666	0.653	Phascolarctobacterium	0.000	0.011	Phascolarctobacterium <= 0.00140
9	Adenocarcinoma	0.595	0.155	0.682	0.657	X.Ruminococcus._gnavus_group	0.003	-0.020	0.00173 < X.Ruminococcus._gnavus_group
9	Adenocarcinoma	0.595	0.155	0.682	0.657	Fusobacterium	0.000	-0.018	Fusobacterium <= 0.000518
9	Adenocarcinoma	0.595	0.155	0.682	0.657	Phascolarctobacterium	0.007	-0.016	0.00618 < Phascolarctobacterium
9	Adenocarcinoma	0.595	0.155	0.682	0.657	Blautia	0.003	0.015	0.00221 < Blautia <= 0.00462
9	Adenocarcinoma	0.595	0.155	0.682	0.657	Chloroplast	0.000	0.015	Chloroplast <= 0.000153
12	Adenocarcinoma	0.734	0.124	0.665	0.675	X.Ruminococcus._gnavus_group	0.000	0.017	X.Ruminococcus._gnavus_group <= 0.00173
12	Adenocarcinoma	0.734	0.124	0.665	0.675	Acinetobacter	0.027	-0.016	0.003851 < Acinetobacter

12	Adenocarcinoma	0.734	0.124	0.665	0.675	X.Eubacterium._hallii_group	0.000	0.012	X.Eubacterium._hallii_group <= 0.000747
12	Adenocarcinoma	0.734	0.124	0.665	0.675	Chloroplast	0.002	-0.014	0.000153 < Chloroplast
12	Adenocarcinoma	0.734	0.124	0.665	0.675	Phascolarctobacterium	0.000	0.011	Phascolarctobacterium <= 0.00140
13	Adenocarcinoma	0.865	0.141	0.655	0.692	X.Ruminococcus._gnavus_group	0.004	-0.019	0.00173 < X.Ruminococcus._gnavus_group
13	Adenocarcinoma	0.865	0.141	0.655	0.692	Acinetobacter	0.000	0.015	Acinetobacter <= 0.000286
13	Adenocarcinoma	0.865	0.141	0.655	0.692	Chloroplast	0.000	0.015	Chloroplast <= 0.000153
13	Adenocarcinoma	0.865	0.141	0.655	0.692	Phascolarctobacterium	0.000	0.013	Phascolarctobacterium <= 0.00140
13	Adenocarcinoma	0.865	0.141	0.655	0.692	X.Eubacterium._hallii_group	0.001	0.012	X.Eubacterium._hallii_group <= 0.000747
15	Adenocarcinoma	0.804	0.117	0.636	0.705	X.Ruminococcus._gnavus_group	0.000	0.019	X.Ruminococcus._gnavus_group <= 0.00173
15	Adenocarcinoma	0.804	0.117	0.636	0.705	X.Eubacterium._hallii_group	0.000	0.013	X.Eubacterium._hallii_group <= 0.000747
15	Adenocarcinoma	0.804	0.117	0.636	0.705	Chloroplast	0.000	0.015	Chloroplast <= 0.000153
15	Adenocarcinoma	0.804	0.117	0.636	0.705	Phascolarctobacterium	0.000	0.011	Phascolarctobacterium <= 0.00140
15	Adenocarcinoma	0.804	0.117	0.636	0.705	Bacillus	0.000	0.012	Bacillus <= 0.000169
16	Adenocarcinoma	0.821	0.116	0.649	0.692	X.Ruminococcus._gnavus_group	0.000	0.019	X.Ruminococcus._gnavus_group <= 0.00173
16	Adenocarcinoma	0.821	0.116	0.649	0.692	Acinetobacter	0.000	0.012	Acinetobacter <= 0.000286
16	Adenocarcinoma	0.821	0.116	0.649	0.692	Chloroplast	0.000	-0.012	0.000153 < Chloroplast
16	Adenocarcinoma	0.821	0.116	0.649	0.692	Phascolarctobacterium	0.000	0.011	Phascolarctobacterium <= 0.00140
16	Adenocarcinoma	0.821	0.116	0.649	0.692	Bacillus	0.000	0.013	Bacillus <= 0.000169
17	Adenocarcinoma	0.695	0.287	0.709	0.602	Blautia	0.015	-0.039	0.00862 < Blautia
17	Adenocarcinoma	0.695	0.287	0.709	0.602	X.Ruminococcus._gnavus_group	0.002	-0.020	0.00173 < X.Ruminococcus._gnavus_group
17	Adenocarcinoma	0.695	0.287	0.709	0.602	Fusobacterium	0.000	-0.017	Fusobacterium <= 0.000518
17	Adenocarcinoma	0.695	0.287	0.709	0.602	Phascolarctobacterium	0.016	-0.018	0.00618 < Phascolarctobacterium
17	Adenocarcinoma	0.695	0.287	0.709	0.602	X.Eubacterium._hallii_group	0.007	-0.013	0.001810 < X.Eubacterium._hallii_group
18	Adenocarcinoma	0.617	0.282	0.675	0.639	Blautia	0.010	-0.039	0.00862 < Blautia
18	Adenocarcinoma	0.617	0.282	0.675	0.639	X.Ruminococcus._gnavus_group	0.000	0.020	X.Ruminococcus._gnavus_group <= 0.00173
18	Adenocarcinoma	0.617	0.282	0.675	0.639	Fusobacterium	0.000	-0.017	Fusobacterium <= 0.000518
18	Adenocarcinoma	0.617	0.282	0.675	0.639	Chloroplast	0.000	0.015	Chloroplast <= 0.000153
18	Adenocarcinoma	0.617	0.282	0.675	0.639	X.Eubacterium._hallii_group	0.007	-0.015	0.001810 < X.Eubacterium._hallii_group
19	Adenocarcinoma	0.705	0.139	0.653	0.695	X.Ruminococcus._gnavus_group	0.000	0.019	X.Ruminococcus._gnavus_group <= 0.00173
19	Adenocarcinoma	0.705	0.139	0.653	0.695	Phascolarctobacterium	0.013	-0.017	0.00618 < Phascolarctobacterium
19	Adenocarcinoma	0.705	0.139	0.653	0.695	X.Eubacterium._hallii_group	0.000	0.013	X.Eubacterium._hallii_group <= 0.000747
19	Adenocarcinoma	0.705	0.139	0.653	0.695	Blautia	0.002	0.014	0.00221 < Blautia <= 0.00462
19	Adenocarcinoma	0.705	0.139	0.653	0.695	Mesorhizobium	0.000	0.013	Mesorhizobium <= 0.000147
27	Adenocarcinoma	0.785	0.139	0.639	0.714	X.Ruminococcus._gnavus_group	0.000	0.020	X.Ruminococcus._gnavus_group <= 0.00173
27	Adenocarcinoma	0.785	0.139	0.639	0.714	Acinetobacter	0.000	0.014	Acinetobacter <= 0.000286
27	Adenocarcinoma	0.785	0.139	0.639	0.714	Chloroplast	0.000	0.015	Chloroplast <= 0.000153
27	Adenocarcinoma	0.785	0.139	0.639	0.714	Blautia	0.001	0.014	Blautia <= 0.00221
27	Adenocarcinoma	0.785	0.139	0.639	0.714	X.Eubacterium._hallii_group	0.000	0.013	X.Eubacterium._hallii_group <= 0.000747
28	Adenocarcinoma	0.783	0.283	0.680	0.639	Blautia	0.010	-0.040	0.00862 < Blautia
28	Adenocarcinoma	0.783	0.283	0.680	0.639	X.Ruminococcus._gnavus_group	0.000	0.018	X.Ruminococcus._gnavus_group <= 0.00173
28	Adenocarcinoma	0.783	0.283	0.680	0.639	Acinetobacter	0.006	-0.017	0.003851 < Acinetobacter
28	Adenocarcinoma	0.783	0.283	0.680	0.639	Bacteroides	0.315	-0.014	0.221 < Bacteroides
28	Adenocarcinoma	0.783	0.283	0.680	0.639	Phascolarctobacterium	0.000	0.012	Phascolarctobacterium <= 0.00140

30	Adenocarcinoma	0.775	0.151	0.662	0.673	Acinetobacter	0.018	-0.019	0.003851 < Acinetobacter
30	Adenocarcinoma	0.775	0.151	0.662	0.673	X.Ruminococcus._gnavus_group	0.001	0.019	X.Ruminococcus._gnavus_group <= 0.00173
30	Adenocarcinoma	0.775	0.151	0.662	0.673	Phascolarctobacterium	0.009	-0.017	0.00618 < Phascolarctobacterium
30	Adenocarcinoma	0.775	0.151	0.662	0.673	Blautia	0.000	0.014	Blautia <= 0.00221
30	Adenocarcinoma	0.775	0.151	0.662	0.673	Chloroplast	0.000	0.014	Chloroplast <= 0.000153
31	Adenocarcinoma	0.684	0.123	0.648	0.694	X.Ruminococcus._gnavus_group	0.000	0.018	X.Ruminococcus._gnavus_group <= 0.00173
31	Adenocarcinoma	0.684	0.123	0.648	0.694	Acinetobacter	0.000	0.015	Acinetobacter <= 0.000286
31	Adenocarcinoma	0.684	0.123	0.648	0.694	Phascolarctobacterium	0.000	0.011	Phascolarctobacterium <= 0.00140
31	Adenocarcinoma	0.684	0.123	0.648	0.694	Chloroplast	0.000	0.013	Chloroplast <= 0.000153
31	Adenocarcinoma	0.684	0.123	0.648	0.694	Butyricimonas	0.006	-0.012	0.00356 < Butyricimonas
35	Adenocarcinoma	0.765	0.124	0.654	0.696	X.Ruminococcus._gnavus_group	0.000	0.019	X.Ruminococcus._gnavus_group <= 0.00173
35	Adenocarcinoma	0.765	0.124	0.654	0.696	Chloroplast	0.001	-0.016	0.000153 < Chloroplast
35	Adenocarcinoma	0.765	0.124	0.654	0.696	Blautia	0.004	0.015	0.00221 < Blautia <= 0.00462
35	Adenocarcinoma	0.765	0.124	0.654	0.696	X.Eubacterium._hallii_group	0.000	0.012	X.Eubacterium._hallii_group <= 0.000747
35	Adenocarcinoma	0.765	0.124	0.654	0.696	Bacillus	0.000	0.013	Bacillus <= 0.000169
36	Adenocarcinoma	0.771	0.129	0.670	0.677	X.Ruminococcus._gnavus_group	0.008	-0.020	0.00173 < X.Ruminococcus._gnavus_group
36	Adenocarcinoma	0.771	0.129	0.670	0.677	X.Eubacterium._hallii_group	0.000	0.013	X.Eubacterium._hallii_group <= 0.000747
36	Adenocarcinoma	0.771	0.129	0.670	0.677	Chloroplast	0.000	0.014	Chloroplast <= 0.000153
36	Adenocarcinoma	0.771	0.129	0.670	0.677	Blautia	0.001	0.013	Blautia <= 0.00221
36	Adenocarcinoma	0.771	0.129	0.670	0.677	Bacteroides	0.501	-0.014	0.221 < Bacteroides
37	Adenocarcinoma	0.763	0.125	0.640	0.708	X.Ruminococcus._gnavus_group	0.000	0.018	X.Ruminococcus._gnavus_group <= 0.00173
37	Adenocarcinoma	0.763	0.125	0.640	0.708	Acinetobacter	0.000	0.013	Acinetobacter <= 0.000286
37	Adenocarcinoma	0.763	0.125	0.640	0.708	Phascolarctobacterium	0.000	0.012	Phascolarctobacterium <= 0.00140
37	Adenocarcinoma	0.763	0.125	0.640	0.708	X.Eubacterium._hallii_group	0.000	0.012	X.Eubacterium._hallii_group <= 0.000747
37	Adenocarcinoma	0.763	0.125	0.640	0.708	Chloroplast	0.000	0.013	Chloroplast <= 0.000153
38	Adenocarcinoma	0.689	0.131	0.664	0.678	X.Ruminococcus._gnavus_group	0.000	0.019	X.Ruminococcus._gnavus_group <= 0.00173
38	Adenocarcinoma	0.689	0.131	0.664	0.678	Acinetobacter	0.007	-0.016	0.003851 < Acinetobacter
38	Adenocarcinoma	0.689	0.131	0.664	0.678	X.Eubacterium._hallii_group	0.000	0.012	X.Eubacterium._hallii_group <= 0.000747
38	Adenocarcinoma	0.689	0.131	0.664	0.678	Phascolarctobacterium	0.000	0.012	Phascolarctobacterium <= 0.00140
38	Adenocarcinoma	0.689	0.131	0.664	0.678	Pseudomonas	0.004	-0.014	0.002448 < Pseudomonas
41	Adenocarcinoma	0.806	0.143	0.652	0.695	X.Ruminococcus._gnavus_group	0.000	0.018	X.Ruminococcus._gnavus_group <= 0.00173
41	Adenocarcinoma	0.806	0.143	0.652	0.695	Phascolarctobacterium	0.011	-0.018	0.00618 < Phascolarctobacterium
41	Adenocarcinoma	0.806	0.143	0.652	0.695	Acinetobacter	0.000	0.014	Acinetobacter <= 0.000286
41	Adenocarcinoma	0.806	0.143	0.652	0.695	Blautia	0.001	0.014	Blautia <= 0.00221
41	Adenocarcinoma	0.806	0.143	0.652	0.695	Chloroplast	0.000	0.014	Chloroplast <= 0.000153
42	Adenocarcinoma	0.706	0.132	0.676	0.659	X.Ruminococcus._gnavus_group	0.020	-0.019	0.00173 < X.Ruminococcus._gnavus_group
42	Adenocarcinoma	0.706	0.132	0.676	0.659	Acinetobacter	0.000	0.014	Acinetobacter <= 0.000286
42	Adenocarcinoma	0.706	0.132	0.676	0.659	Chloroplast	0.000	0.015	Chloroplast <= 0.000153
42	Adenocarcinoma	0.706	0.132	0.676	0.659	X.Ruminococcus._torques_group	0.022	-0.013	0.01976 < X.Ruminococcus._torques_group
42	Adenocarcinoma	0.706	0.132	0.676	0.659	Bacillus	0.000	-0.013	0.000169 < Bacillus
43	Adenocarcinoma	0.583	0.146	0.690	0.635	X.Ruminococcus._gnavus_group	0.002	-0.020	0.00173 < X.Ruminococcus._gnavus_group
43	Adenocarcinoma	0.583	0.146	0.690	0.635	Fusobacterium	0.000	-0.017	Fusobacterium <= 0.000518
43	Adenocarcinoma	0.583	0.146	0.690	0.635	Phascolarctobacterium	0.028	-0.018	0.00618 < Phascolarctobacterium

43	Adenocarcinoma	0.583	0.146	0.690	0.635	X.Eubacterium._hallii_group	0.010	-0.014	0.001810 < X.Eubacterium._hallii_group
43	Adenocarcinoma	0.583	0.146	0.690	0.635	Bacillus	0.000	0.013	Bacillus <= 0.000169
44	Adenocarcinoma	0.708	0.158	0.671	0.653	Fusobacterium	0.000	-0.019	Fusobacterium <= 0.000518
44	Adenocarcinoma	0.708	0.158	0.671	0.653	X.Ruminococcus._gnavus_group	0.000	0.019	X.Ruminococcus._gnavus_group <= 0.00173
44	Adenocarcinoma	0.708	0.158	0.671	0.653	Phascolarctobacterium	0.012	-0.018	0.00618 < Phascolarctobacterium
44	Adenocarcinoma	0.708	0.158	0.671	0.653	Bacillus	0.000	0.014	Bacillus <= 0.000169
44	Adenocarcinoma	0.708	0.158	0.671	0.653	X.Ruminococcus._torques_group	0.030	-0.014	0.01976 < X.Ruminococcus._torques_group
45	Adenocarcinoma	0.751	0.145	0.658	0.673	X.Ruminococcus._gnavus_group	0.000	0.020	X.Ruminococcus._gnavus_group <= 0.00173
45	Adenocarcinoma	0.751	0.145	0.658	0.673	Phascolarctobacterium	0.023	-0.017	0.00618 < Phascolarctobacterium
45	Adenocarcinoma	0.751	0.145	0.658	0.673	Acinetobacter	0.000	0.014	Acinetobacter <= 0.000286
45	Adenocarcinoma	0.751	0.145	0.658	0.673	X.Ruminococcus._torques_group	0.023	-0.016	0.01976 < X.Ruminococcus._torques_group
45	Adenocarcinoma	0.751	0.145	0.658	0.673	Chloroplast	0.000	0.015	Chloroplast <= 0.000153



Figure ML1-E-6: The LIME Result in Training dataset for Genus in Biopsy



Figure ML1-E-7: The LIME Result in Test dataset for Genus in Biopsy

Table ML1-E-6: The Aggregated SHAP Values

Variable	Training Dataset				Testing Dataset			
	contribution	position	sign	cumulative	contribution	position	sign	cumulative
intercept	0.7379	225	X	0.738	0.73792	225	X	0.738
Tyzzerella	0.0000	224	1	0.738	-0.00011	224	-1	0.738
Barnesiella	0.0000	223	-1	0.738	0.001243	223	1	0.739
Sphingomonas	0.0000	222	-1	0.738	-2.2E-05	222	-1	0.739
Asteroleplasma	0.0000	221	1	0.738	0	160	1	0.735
Slackia	0.0000	220	-1	0.738	0.000279	11	1	0.723
Christensenellaceae_R.7_group	0.0000	219	-1	0.738	0.000362	210	1	0.737
Serratia	0.0000	218	-1	0.738	0.002338	152	1	0.736
Peptostreptococcus	0.0000	217	-1	0.738	-0.00078	205	-1	0.737
Intestinibacter	0.0000	216	1	0.738	-0.00077	219	-1	0.738
Oribacterium	0.0000	215	1	0.738	0	218	1	0.738
uncultured	0.0000	214	1	0.738	-1E-05	217	-1	0.738
Rikenellaceae_RC9_gut_group	0.0000	213	1	0.738	-0.00036	215	-1	0.738
Alistipes	0.0000	212	-1	0.738	-0.00074	221	-1	0.738
KD4.96	0.0000	211	1	0.738	0	209	1	0.737
Pelomonas	0.0000	210	1	0.738	6.36E-05	207	1	0.737
Atopobium	0.0000	209	1	0.738	0	202	1	0.736
Collinsella	0.0000	208	1	0.738	0.000339	201	1	0.736
DTU014	0.0000	207	1	0.738	0	200	1	0.736
Sellimonas	0.0000	206	-1	0.738	6.34E-05	199	1	0.736
Sutterella	0.0000	205	-1	0.738	-0.00088	198	-1	0.735
Campylobacter	0.0000	204	1	0.738	-0.00027	216	-1	0.738
Pseudomonas	0.0000	203	1	0.738	0.00246	163	1	0.735
Flavobacterium	0.0000	202	1	0.738	0.000163	126	1	0.733
Gastranaerophilales	0.0000	201	1	0.738	-0.0001	194	-1	0.735
GCA.900066575	0.0000	200	1	0.738	0.000255	149	1	0.737
Micrococcus	0.0000	199	-1	0.738	6.99E-05	193	1	0.735
Subdoligranulum	0.0000	198	1	0.738	0.000338	206	1	0.738
Clostridia_UCG.014	0.0000	197	1	0.738	0.000348	192	1	0.735
Incertae_Sedis	0.0000	196	-1	0.738	0.000424	190	1	0.736
Acinetobacter	0.0000	195	1	0.738	-0.00261	189	-1	0.733
Ruminiclostridium	0.0000	194	1	0.738	-1.9E-05	188	-1	0.733
Catabacter	0.0000	193	1	0.738	0	184	1	0.733
Family_XIII_UCG.001	0.0000	192	-1	0.738	2.52E-05	183	1	0.733
Erysipelotrichaceae	0.0000	191	-1	0.738	-0.0002	182	-1	0.733
CAG.352	0.0000	190	1	0.738	0	168	1	0.733
Mailhella	0.0000	189	1	0.738	0	195	1	0.735
Treponema	0.0000	188	1	0.738	0	178	1	0.733
Kocuria	0.0000	187	1	0.738	-0.00069	177	-1	0.732

Paraprevotella	0.0000	186	-1	0.738	0.000939	176	1	0.733
X.Eubacterium._eligens_group	0.0000	185	1	0.738	0	175	1	0.733
Delftia	0.0000	184	-1	0.738	0.000599	173	1	0.733
Gilliamella	0.0000	183	1	0.738	0	191	1	0.735
Muribaculaceae	0.0000	182	1	0.738	0.000654	172	1	0.734
Neisseria	0.0000	181	1	0.738	0.000101	171	1	0.734
Enhydrobacter	0.0000	180	1	0.738	0	170	1	0.734
X.Clostridium._methylpentosum_group	0.0000	179	1	0.738	0.000504	119	1	0.733
Anaerococcus	0.0000	178	1	0.738	0.000204	167	1	0.734
Butyricimonas	0.0000	177	1	0.738	-0.00091	174	-1	0.732
Family_XIII_AD3011_group	0.0000	176	1	0.738	0.000105	165	1	0.734
Blautia	0.0000	175	-1	0.738	-0.00162	164	-1	0.732
Anaerostipes	0.0000	174	1	0.738	-0.00042	197	-1	0.735
X.Eubacterium._nodatum_group	0.0000	173	1	0.738	-2.8E-05	162	-1	0.735
Dorea	0.0000	172	-1	0.738	0.00045	146	1	0.738
Holdemanella	0.0000	171	1	0.738	-0.00014	169	-1	0.733
X.Eubacterium._ruminantium_group	0.0000	170	1	0.738	-6E-06	158	-1	0.733
Lawsonella	0.0000	169	1	0.738	0	156	1	0.733
vadinBE97	0.0000	168	1	0.738	0	155	1	0.733
Veillonella	0.0000	167	1	0.738	0.000101	154	1	0.734
Chryseobacterium	0.0000	166	1	0.738	0	153	1	0.734
Coprobacter	0.0000	165	1	0.738	5.88E-05	120	1	0.733
Faecalibacterium	0.0000	164	1	0.738	0.000191	150	1	0.736
Mogibacterium	0.0000	163	1	0.738	0	148	1	0.737
Hydrogenoanaerobacterium	0.0000	162	1	0.738	-5.1E-05	145	-1	0.738
Methanobrevibacter	0.0000	161	1	0.738	7.98E-05	53	1	0.721
Filifactor	0.0000	160	-1	0.738	-1.5E-06	144	-1	0.738
Staphylococcus	0.0000	159	-1	0.738	-7.7E-05	95	-1	0.735
X.Ruminococcus._gauvreauii_group	0.0000	158	1	0.738	-0.00011	118	-1	0.733
Lachnospiraceae_ND3007_group	0.0000	157	1	0.738	0.000103	143	1	0.738
Lactobacillus	0.0000	156	-1	0.738	-0.00011	208	-1	0.737
Gender	0.0000	155	1	0.738	0	213	1	0.738
Clostridium_sensu_stricto_2	0.0000	154	1	0.738	0	142	1	0.738
Mesorhizobium	0.0000	153	-1	0.738	-0.00068	211	-1	0.737
Streptococcus	0.0000	152	-1	0.738	-0.0007	140	-1	0.734
Haemophilus	0.0000	151	1	0.738	-9.7E-05	109	-1	0.735
Rheinheimera	0.0000	150	-1	0.738	-0.00068	186	-1	0.733
Roseburia	0.0000	149	-1	0.738	0.000225	139	1	0.734
Actinomyces	0.0000	148	-1	0.738	0.000213	138	1	0.734
Lachnospiraceae_UCG.001	0.0000	147	1	0.738	0	137	1	0.734
Marvinbryantia	0.0000	146	-1	0.738	-0.00099	136	-1	0.733
UCG.005	0.0000	145	1	0.738	-0.00104	204	-1	0.736
UCG.002	0.0000	144	1	0.738	-0.00049	135	-1	0.733

Stenotrophomonas	0.0000	143	-1	0.738	-0.00021	134	-1	0.733
Bradyrhizobium	0.0000	142	1	0.738	0.000443	133	1	0.733
Allorhizobium.Neorhizobium.Pararhizobium.Rhizobium	0.0000	141	-1	0.738	0.000252	151	1	0.736
Bergeyella	0.0000	140	1	0.738	-5.2E-06	132	-1	0.733
Stomatobaculum	0.0000	139	1	0.738	3.71E-06	131	1	0.733
Tumor	0.0000	138	1	0.738	0	130	1	0.733
Corynebacterium	0.0000	137	-1	0.738	-8.5E-05	104	-1	0.734
Ruminococcus	0.0000	136	-1	0.738	0.000949	220	1	0.739
Enterobacter	0.0000	135	1	0.738	-6.9E-05	129	-1	0.733
Lachnospiraceae_FCS020_group	0.0000	134	-1	0.738	-0.00342	141	-1	0.734
X.Eubacterium._coprostanoligenes_group	0.0000	133	1	0.738	-0.00134	159	-1	0.733
Butyrivibrio	0.0000	132	-1	0.738	-3.5E-05	161	-1	0.735
Cloacibacterium	0.0000	131	-1	0.738	0.000237	127	1	0.733
Holdemania	0.0000	130	-1	0.738	1.04E-05	91	1	0.735
Hungatella	0.0000	129	1	0.738	9.2E-05	70	1	0.722
Prevotella	0.0000	128	1	0.738	-0.00015	124	-1	0.733
Candidatus_Udaeobacter	0.0000	127	-1	0.738	0.000192	157	1	0.733
Frisingicoccus	0.0000	126	1	0.738	0	123	1	0.733
UBA1819	0.0000	125	1	0.738	0.000127	122	1	0.733
X.Eubacterium._brachy_group	0.0000	124	1	0.738	0	179	1	0.733
Cloacibacillus	0.0000	123	1	0.738	-0.00011	121	-1	0.733
Lachnoanaerobaculum	0.0000	122	1	0.738	0	117	1	0.733
Merdibacter	0.0000	121	1	0.738	0	196	1	0.735
Methylobacterium.Methylorubrum	0.0000	120	1	0.738	0.000235	116	1	0.734
Monoglobus	0.0000	119	-1	0.738	0.000252	115	1	0.734
Odoribacter	0.0000	118	-1	0.738	0.000167	114	1	0.734
Comamonas	0.0000	117	1	0.738	-3.5E-05	113	-1	0.734
Sanguibacteroides	0.0000	116	-1	0.738	4.24E-05	112	1	0.734
Erysipelatoclostridium	0.0000	115	1	0.738	0.000202	110	1	0.735
Escherichia.Shigella	0.0000	114	-1	0.738	0.000602	187	1	0.734
Peptococcus	0.0000	113	-1	0.738	-7.3E-05	107	-1	0.734
Pyramidobacter	0.0000	112	-1	0.738	-3.6E-05	106	-1	0.734
UCG.003	0.0000	111	-1	0.738	-0.00012	128	-1	0.733
Erysipelotrichaceae_UCG.003	0.0000	110	-1	0.738	-0.00051	84	-1	0.724
Morganella	0.0000	109	1	0.738	0	103	1	0.734
Pseudoalteromonas	0.0000	108	1	0.738	0.001079	102	1	0.736
X.Ruminococcus._torques_group	0.0000	107	-1	0.738	-0.00035	101	-1	0.735
Aggregatibacter	0.0000	106	1	0.738	-2.1E-05	100	-1	0.735
TM7x	0.0000	105	1	0.738	0	99	1	0.735
UCG.010	0.0000	104	1	0.738	-0.00039	98	-1	0.735
Allisonella	0.0000	103	-1	0.738	-7.4E-05	57	-1	0.723
Defluviitaleaceae_UCG.011	0.0000	102	1	0.738	0	97	1	0.735
Izemoplasmatales	0.0000	101	-1	0.738	1.15E-05	96	1	0.735

Porphyromonas	0.0000	100	1	0.738	1.76E-05	74	1	0.721
Bifidobacterium	0.0000	99	-1	0.738	4.65E-05	93	1	0.734
Brachymonas	0.0000	98	1	0.738	0	166	1	0.734
Butyricicoccus	0.0000	97	1	0.738	0.000363	92	1	0.735
Oscillospira	0.0000	96	-1	0.738	0.000151	90	1	0.735
Phascolarctobacterium	0.0000	95	-1	0.738	-0.00394	88	-1	0.725
Bosea	0.0000	94	-1	0.738	-7.9E-05	181	-1	0.733
Rhodococcus	0.0000	93	-1	0.738	-4.6E-06	73	-1	0.721
Selenomonas	0.0000	92	1	0.738	0.000127	87	1	0.725
Senegalimassilia	0.0000	91	-1	0.738	-0.00031	86	-1	0.725
X.Clostridium._innocuum_group	0.0000	90	-1	0.738	-0.00032	85	-1	0.725
Fusobacterium	0.0000	89	1	0.738	-0.00212	83	-1	0.722
Gemella	0.0000	88	-1	0.738	-0.00038	212	-1	0.738
IheB3.7	0.0000	87	1	0.738	0	82	1	0.722
Peptoniphilus	0.0000	86	-1	0.738	0.000319	214	1	0.738
Rothia	0.0000	85	-1	0.738	-6.2E-05	81	-1	0.722
Terrisporobacter	0.0000	84	1	0.738	0	80	1	0.722
Chloroplast	0.0000	83	-1	0.738	0.000339	79	1	0.722
Clostridia_vadinBB60_group	0.0000	82	-1	0.738	-0.00029	78	-1	0.722
Solobacterium	0.0000	81	1	0.738	-0.00077	58	-1	0.723
Dielma	0.0000	80	-1	0.738	0.000681	76	1	0.721
Faecalitalea	0.0000	79	1	0.738	0	75	1	0.721
Lachnospira	0.0000	78	-1	0.738	-0.00036	94	-1	0.734
Lachnospiraceae_UCG.010	0.0000	77	1	0.738	0.000198	105	1	0.735
Parabacteroides	0.0000	76	1	0.738	0.000819	147	1	0.737
Anaerotruncus	0.0000	75	-1	0.738	-0.00017	125	-1	0.733
Dialister	0.0000	74	1	0.738	-4.5E-06	71	-1	0.721
Romboutsia	0.0000	73	1	0.738	0.000476	69	1	0.722
UCG.004	0.0000	72	1	0.738	0	203	1	0.736
Alloprevotella	0.0000	71	1	0.738	-0.00042	68	-1	0.722
Lachnospiraceae_NK4A136_group	0.0000	70	-1	0.738	0.000526	67	1	0.722
Succinivibrio	0.0000	69	1	0.738	0	66	1	0.722
Turicibacter	0.0000	68	1	0.738	0.000364	185	1	0.733
Catonella	0.0000	67	1	0.738	0	65	1	0.722
Bacteroides	0.0000	66	1	0.738	0.000591	64	1	0.723
Ochrobactrum	0.0000	65	1	0.738	-0.00027	180	-1	0.733
Oscillibacter	0.0000	64	-1	0.738	8E-06	63	1	0.723
Prevotellaceae_NK3B31_group	0.0000	63	1	0.738	0.000637	62	1	0.723
X.Eubacterium._siraeum_group	0.0000	62	-1	0.738	0.000217	72	1	0.721
Catenibacterium	0.0000	61	-1	0.738	0.000123	61	1	0.723
Colidextribacter	0.0000	60	1	0.738	-0.00159	77	-1	0.721
Enterococcus	0.0000	59	1	0.738	-1.8E-05	60	-1	0.723
Intestinimonas	0.0000	58	1	0.738	0.00018	59	1	0.724

Desulfovibrio	0.0000	57	1	0.738	-0.00114	56	-1	0.722
Eikenella	0.0000	56	1	0.738	0	55	1	0.722
Howardella	0.0000	55	1	0.738	-0.00061	54	-1	0.721
Parvimonas	0.0000	54	-1	0.738	-0.0057	89	-1	0.729
Acidaminococcus	0.0000	53	1	0.738	-1.5E-05	52	-1	0.721
Acidothermus	0.0000	52	1	0.738	0	51	1	0.721
Aeromonas	0.0000	51	-1	0.738	-2.9E-05	50	-1	0.721
Akkermansia	0.0000	50	1	0.738	0.000716	49	1	0.722
Bacillus	0.0000	49	-1	0.738	0.000385	48	1	0.722
Bacteroidetes_BD2.2	0.0000	48	1	0.738	0	47	1	0.722
Bilophila	0.0000	47	1	0.738	0.000421	111	1	0.734
CAG.56	0.0000	46	1	0.738	-0.00069	46	-1	0.721
CAG.873	0.0000	45	1	0.738	-2.8E-05	45	-1	0.721
Capnocytophaga	0.0000	44	1	0.738	0.000267	44	1	0.722
Citrobacter	0.0000	43	1	0.738	3.37E-05	43	1	0.722
Clostridium_sensu_stricto_1	0.0000	42	1	0.738	0.000115	42	1	0.722
Coprococcus	0.0000	41	1	0.738	-0.00129	41	-1	0.721
Cutibacterium	0.0000	40	1	0.738	0.000194	40	1	0.721
Desulfosarcina	0.0000	39	-1	0.738	-3E-05	39	-1	0.721
Eisenbergiella	0.0000	38	1	0.738	0.000102	38	1	0.721
Eubacterium	0.0000	37	1	0.738	0	37	1	0.721
Flavonifractor	0.0000	36	-1	0.738	0.00094	36	1	0.722
Fournierella	0.0000	35	1	0.738	-0.00021	35	-1	0.722
Fusicatenibacter	0.0000	34	-1	0.738	0.000478	34	1	0.722
Granulicatella	0.0000	33	1	0.738	-0.00048	33	-1	0.722
Herbaspirillum	0.0000	32	1	0.738	0	32	1	0.722
IMCC26256	0.0000	31	1	0.738	0	31	1	0.722
Klebsiella	0.0000	30	-1	0.738	-0.00019	30	-1	0.721
Lachnoclostridium	0.0000	29	1	0.738	0.000739	29	1	0.722
Lachnospiraceae_UCG.004	0.0000	28	-1	0.738	-0.00044	28	-1	0.722
Lachnospiraceae_UCG.008	0.0000	27	1	0.738	0.000592	27	1	0.722
Leptotrichia	0.0000	26	1	0.738	-0.00027	26	-1	0.722
Lysinibacillus	0.0000	25	1	0.738	0.001191	25	1	0.723
Megasphaera	0.0000	24	-1	0.738	0.000214	24	1	0.723
Mitsuokella	0.0000	23	1	0.738	0	23	1	0.723
Mycobacterium	0.0000	22	1	0.738	0	22	1	0.723
Negativibacillus	0.0000	21	-1	0.738	0.000237	21	1	0.724
NK4A214_group	0.0000	20	1	0.738	-0.00012	20	-1	0.724
Olsenella	0.0000	19	-1	0.738	-9.1E-05	19	-1	0.723
Paenibacillus	0.0000	18	1	0.738	-5.9E-05	108	-1	0.734
Paeniclostridium	0.0000	17	1	0.738	0	18	1	0.723
Pantoea	0.0000	16	-1	0.738	-0.00043	17	-1	0.723
Paracoccus	0.0000	15	-1	0.738	0.000392	16	1	0.723

Parasutterella	0.0000	14	1	0.738	-0.00014	15	-1	0.723
Phenylobacterium	0.0000	13	1	0.738	0	14	1	0.723
Prevotellaceae_UCG.001	0.0000	12	-1	0.738	-0.00019	13	-1	0.723
S5.A14a	0.0000	11	-1	0.738	-0.00046	12	-1	0.723
Variovorax	0.0000	10	1	0.738	0	10	1	0.723
Victivallaceae	0.0000	9	1	0.738	0	9	1	0.723
Victivallis	0.0000	8	1	0.738	-6.3E-05	8	-1	0.723
Woeseia	0.0000	7	1	0.738	0	7	1	0.723
X.Eubacterium._hallii_group	-0.0001	6	-1	0.738	-0.00211	6	-1	0.721
X.Eubacterium._ventriosum_group	0.0000	5	-1	0.738	0.001157	5	1	0.722
X.Eubacterium._xylanophilum_group	0.0000	4	1	0.738	-0.00068	4	-1	0.721
X.Ruminococcus._gnavus_group	0.0000	3	-1	0.738	0.001802	3	1	0.723
X67.14	0.0000	2	1	0.738	0	2	1	0.723
prediction	0.7379	1	X	0.738	0.723	1	X	0.723

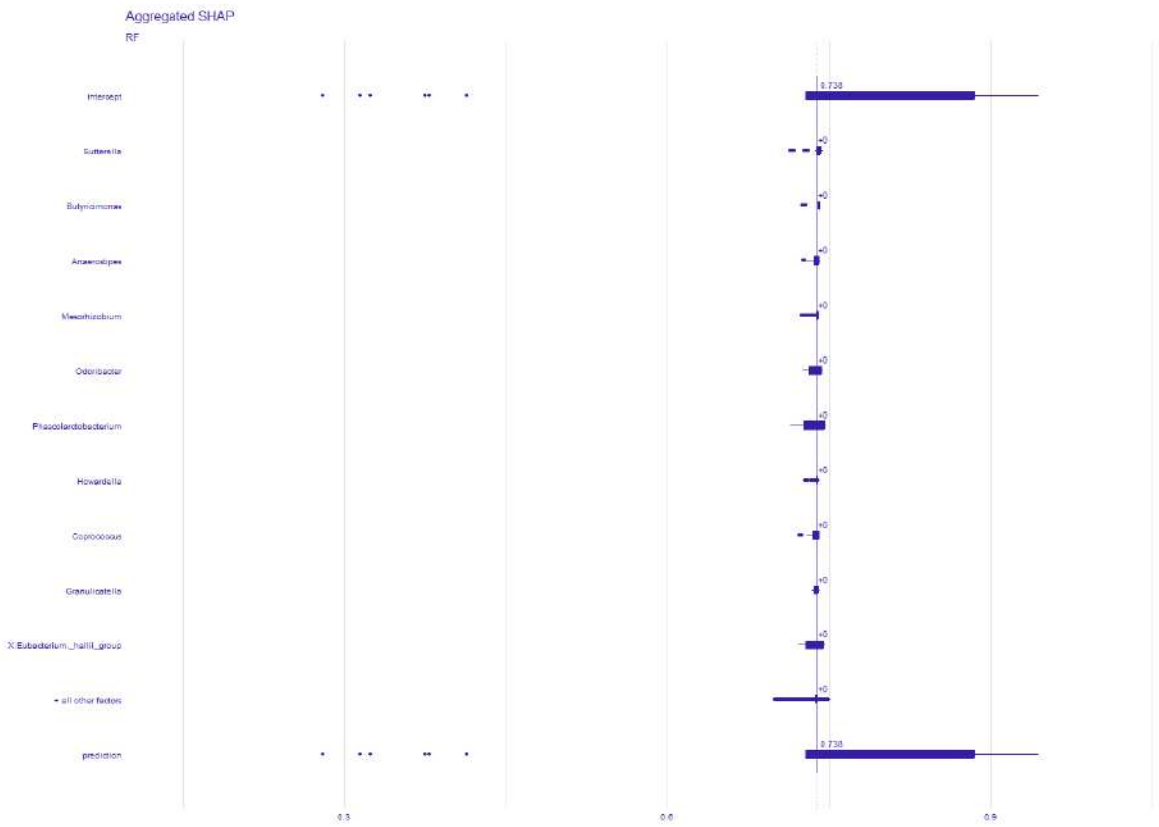


Figure ML1-E-8: The Aggregated SHAP Value in Training dataset for Genus in Biopsy

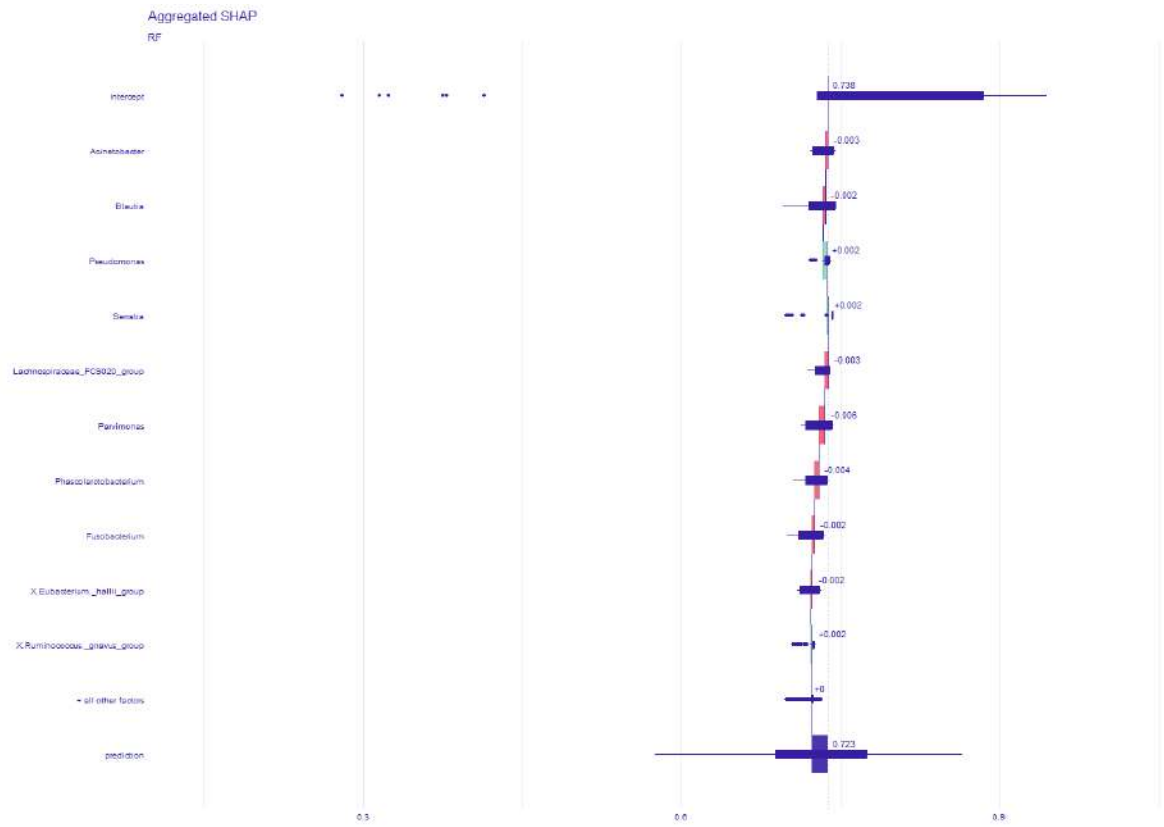


Figure ML1-E-9: The Aggregated SHAP Value in Test dataset for Genus in Biopsy

ML2 - Stool

A- Phylum

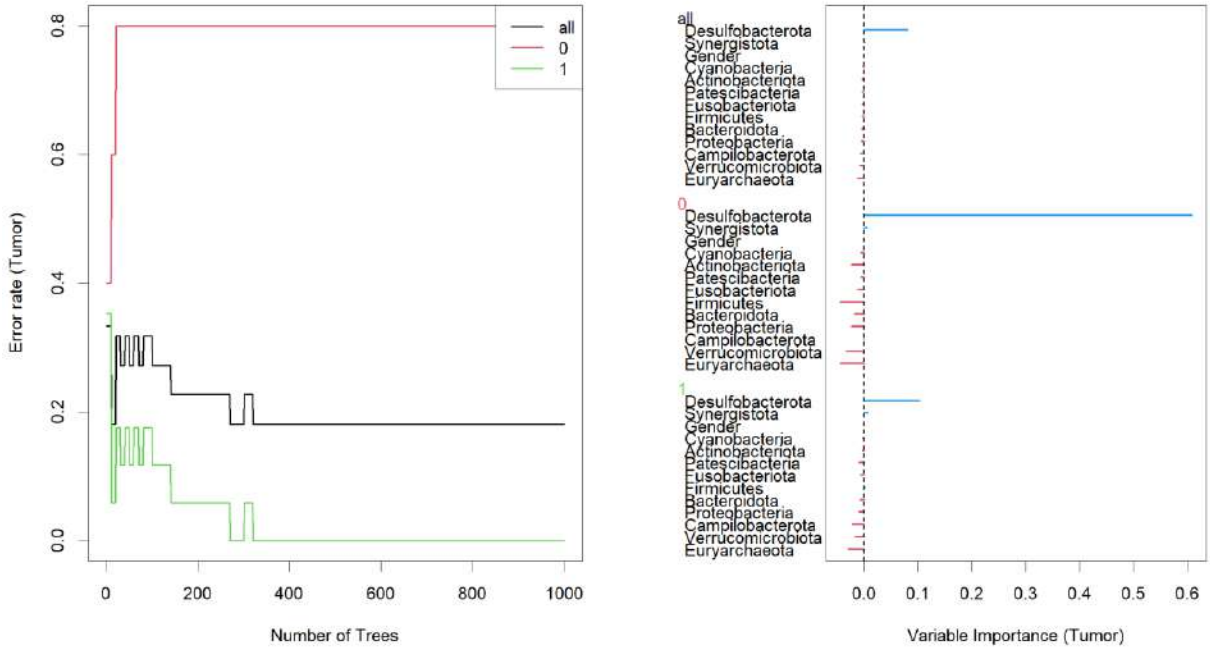


Figure ML2-A-1: The number of trees and Variable Importance Plot for **Phylum** in **Stool**

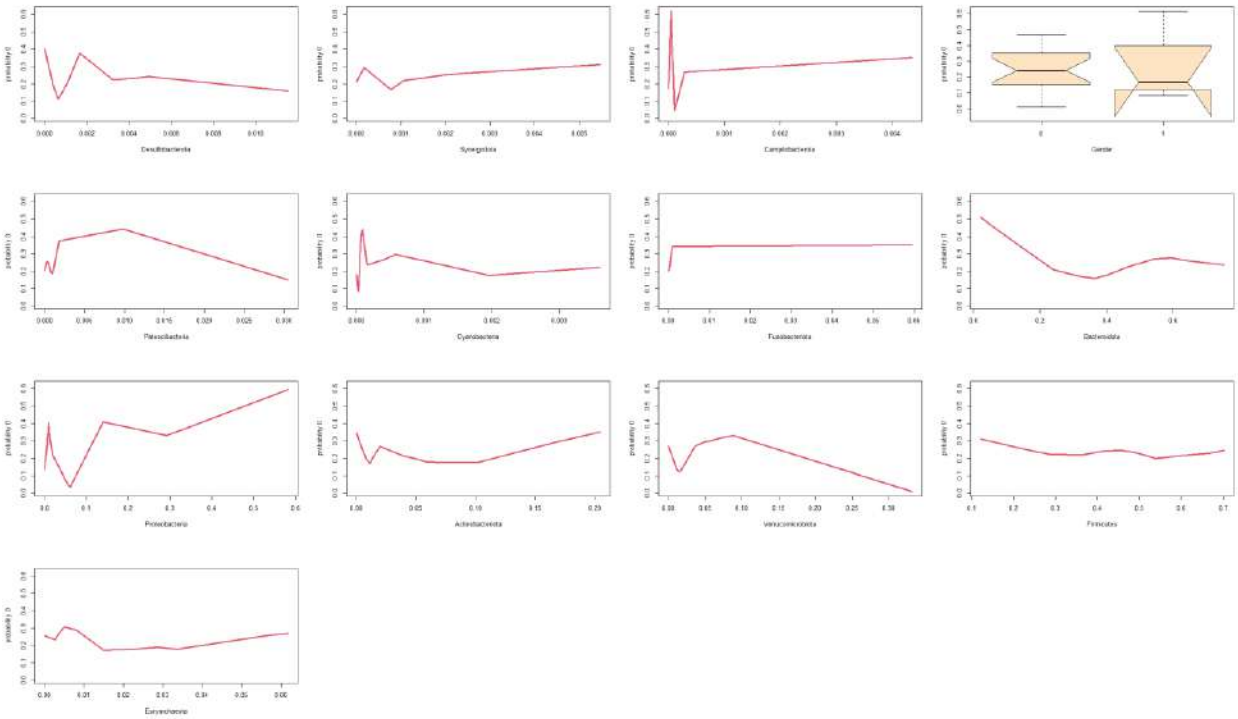


Figure ML2-A-2: Marginal Plot for Phylum in Stool

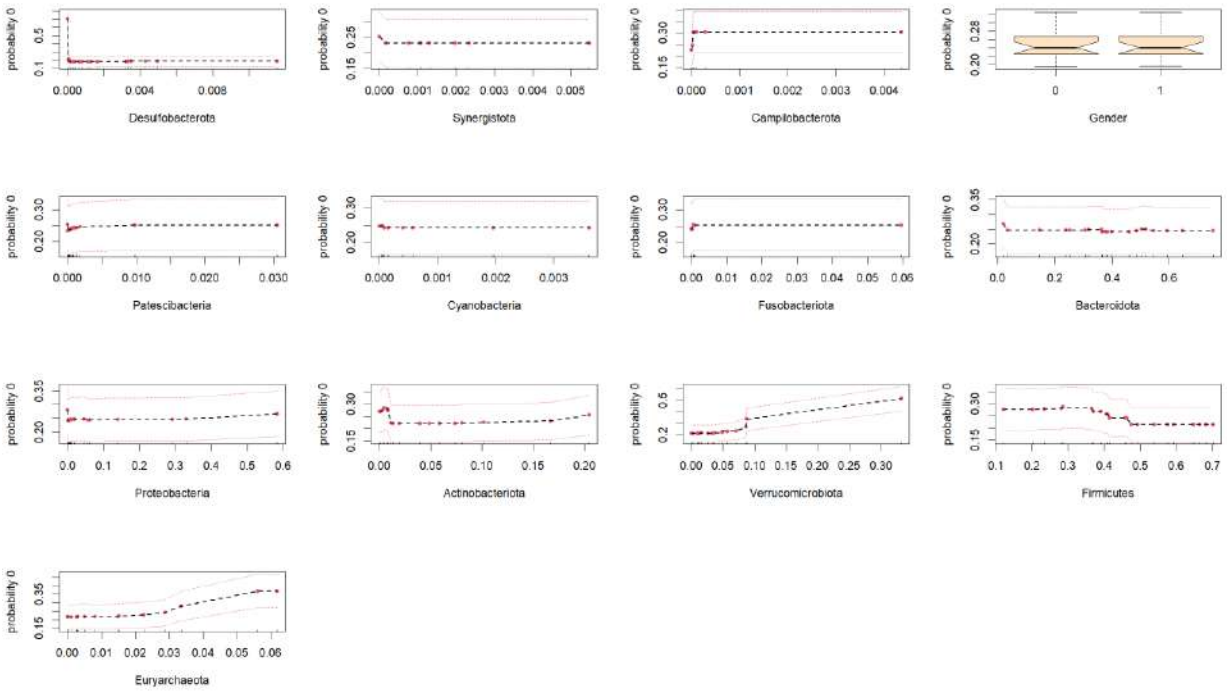


Figure ML2-A-3: Partial Plot for Phylum in Stool

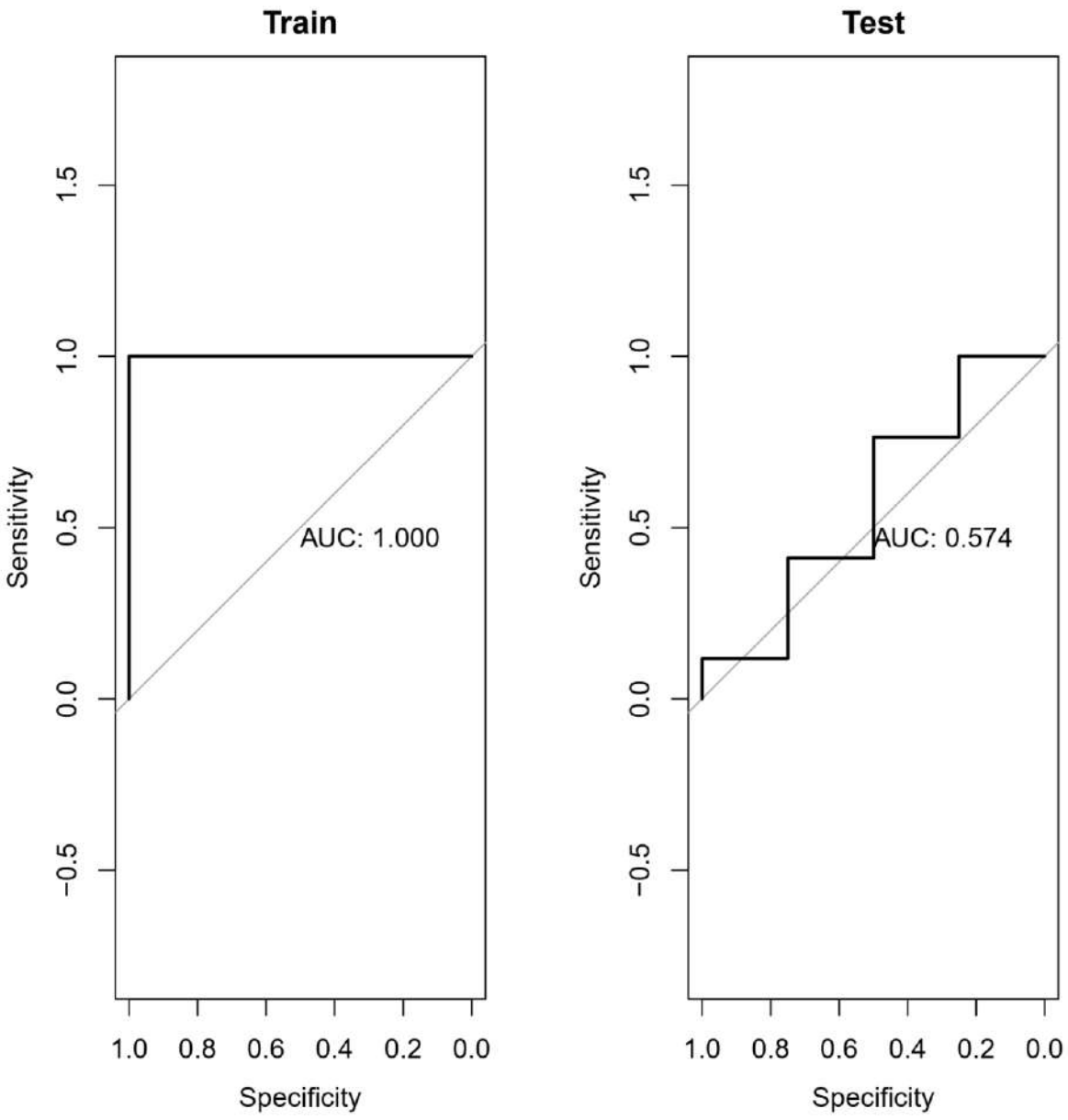


Figure ML2-A-4: The ROC Curve for Phylum in Stool

Table ML2-A-1: The Model Performance

Index	Values		
	2.50%	**50.00%**	**97.50%**
Accuracy	0.24	0.76	0.95
Specificity	0.25	0.75	1.00
Sensitivity	0.06	0.82	1.00

Table ML2-A-2: The Variable Importance Values

Covariates	Varibale Importance			Brier Score		
	all	**AP**	**CRC**	**all**	**AP**	**CRC**
Gender	0.0000	0.0000	0.0000	-0.0062	-0.0042	-0.0042
Actinobacteriota	0.0046	-0.1468	0.0592	-0.0046	-0.0031	-0.0031
Bacteroidota	0.0059	-0.0815	0.0448	-0.0143	-0.0097	-0.0097
Campilobacterota	0.0042	-0.1468	0.0576	-0.0043	-0.0029	-0.0029
Cyanobacteria	-0.0032	-0.0163	-0.0064	-0.0069	-0.0047	-0.0047
Desulfobacterota	0.5226	-0.2175	1.8836	1.5132	1.0284	1.0284
Euryarchaeota	0.0491	-0.2718	0.2510	0.0859	0.0584	0.0584
Firmicutes	0.0068	0.0109	0.0208	-0.0045	-0.0031	-0.0031
Fusobacteriota	-0.0042	-0.0109	-0.0112	-0.0166	-0.0113	-0.0113
Patescibacteria	-0.0001	-0.0217	0.0064	-0.0238	-0.0162	-0.0162
Proteobacteria	0.0160	-0.2120	0.1183	0.0007	0.0005	0.0005
Synergistota	0.0004	0.0489	-0.0128	-0.0021	-0.0014	-0.0014
Verrucomicrobiota	0.1859	-0.7285	0.8619	0.4543	0.3087	0.3087

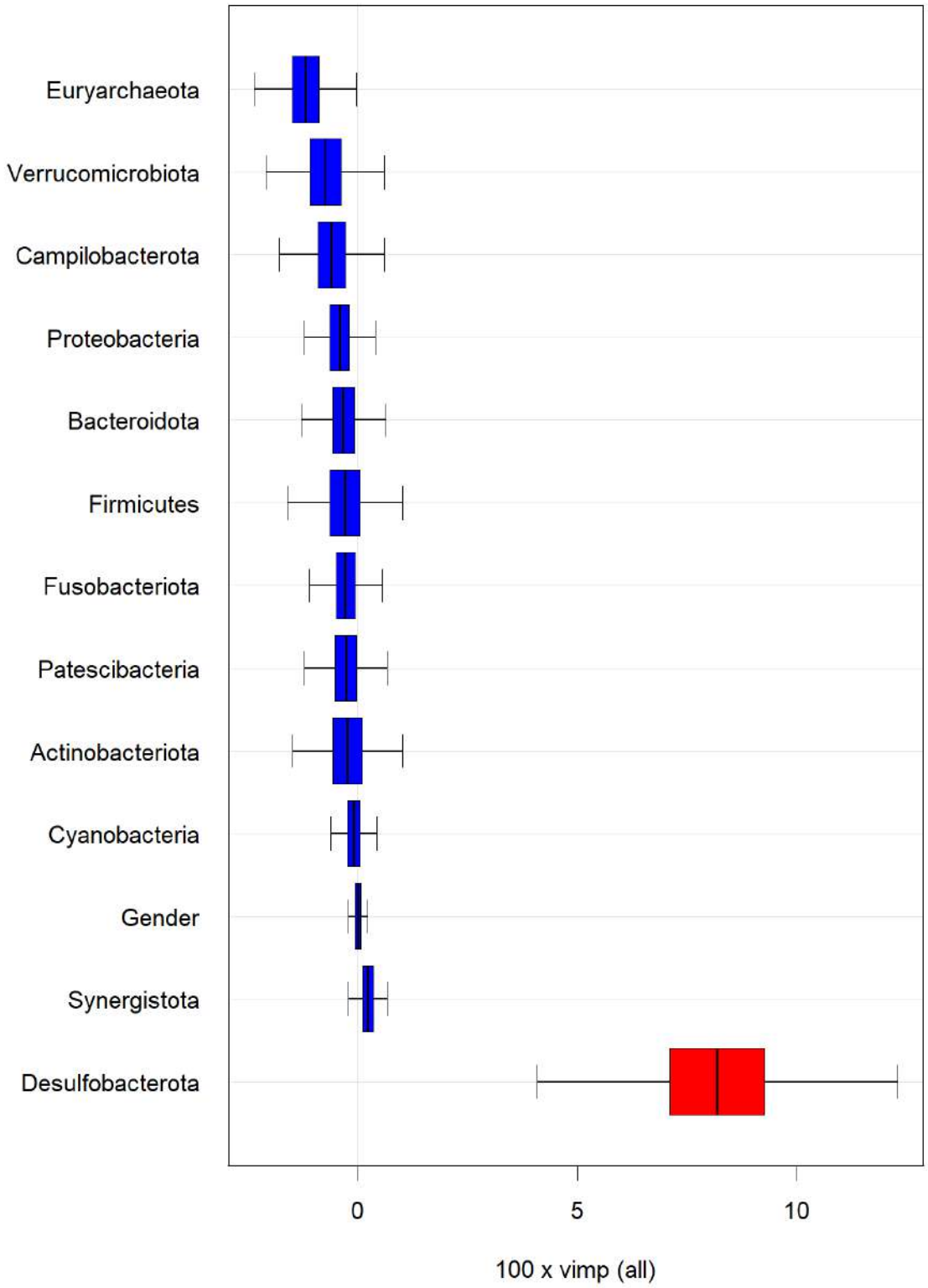

Euryarchaeota
Verrucomicrobiota
Campilobacterota
Proteobacteria
Bacteroidota
Firmicutes
Fusobacteriota
Patescibacteria
Actinobacteriota
Cyanobacteria
Gender
Synergistota
Desulfobacterota
0
5
10
100 x vimp (all)

Figure ML2-A-5: The Variable Importance Plot (VIMP) for Phylum in Stool

Best model is selected for LIME results

Table ML2-S-3: The best setting for model and LIME Results

Setting	mtry	ROC	Sens	Spec	ROCSD	SensSD	SpecSD
ntrees: 500 nodesize: 1	10	0.65	0.4	1	0.48734	0.547723	0
ntrees: 1000 nodesize: 1	7	0.65	0.4	1	0.48734	0.547723	0
ntrees: 1000 nodesize: 1	8	0.65	0.4	1	0.48734	0.547723	0
ntrees: 1000 nodesize: 1	10	0.65	0.4	1	0.48734	0.547723	0
ntrees: 500 nodesize: 5	9	0.65	0.4	1	0.48734	0.547723	0
ntrees: 500 nodesize: 5	10	0.65	0.4	1	0.48734	0.547723	0
ntrees: 1000 nodesize: 5	10	0.65	0.4	1	0.48734	0.547723	0

Table ML2-A-4: The LIME Result for Training Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
1	Adenoma	0.670	0.225	0.217	0.384	Verrucomicrobiota	0.331	0.166	4.92e-02 < Verrucomicrobiota
1	Adenoma	0.670	0.225	0.217	0.384	Firmicutes	0.286	0.053	Firmicutes <= 0.367
1	Adenoma	0.670	0.225	0.217	0.384	Euryarchaeota	0.000	-0.031	Euryarchaeota <= 0.0028
1	Adenoma	0.670	0.225	0.217	0.384	Desulfobacterota	0.001	-0.037	0.000637 < Desulfobacterota <= 0.002835
1	Adenoma	0.670	0.225	0.217	0.384	Actinobacteriota	0.008	0.016	0.00434 < Actinobacteriota <= 0.01308
2	Adenoma	0.662	0.086	0.233	0.288	Euryarchaeota	0.056	0.074	0.0152 < Euryarchaeota
2	Adenoma	0.662	0.086	0.233	0.288	Desulfobacterota	0.003	-0.046	0.002835 < Desulfobacterota
2	Adenoma	0.662	0.086	0.233	0.288	Firmicutes	0.367	0.042	Firmicutes <= 0.367
2	Adenoma	0.662	0.086	0.233	0.288	Verrucomicrobiota	0.015	-0.044	1.41e-02 < Verrucomicrobiota <= 4.92e-02
2	Adenoma	0.662	0.086	0.233	0.288	Fusobacteriota	0.001	0.028	1.87e-04 < Fusobacteriota
3	Adenoma	0.812	0.121	0.271	0.240	Desulfobacterota	0.000	0.109	Desulfobacterota <= 0.000170
3	Adenoma	0.812	0.121	0.271	0.240	Verrucomicrobiota	0.000	-0.055	Verrucomicrobiota <= 4.71e-05
3	Adenoma	0.812	0.121	0.271	0.240	Euryarchaeota	0.003	-0.036	Euryarchaeota <= 0.0028
3	Adenoma	0.812	0.121	0.271	0.240	Firmicutes	0.463	-0.031	0.437 < Firmicutes <= 0.567
3	Adenoma	0.812	0.121	0.271	0.240	Fusobacteriota	0.000	-0.018	Fusobacteriota <= 4.83e-05
6	Adenoma	0.840	0.252	0.176	0.455	Verrucomicrobiota	0.049	0.155	4.92e-02 < Verrucomicrobiota
6	Adenoma	0.840	0.252	0.176	0.455	Desulfobacterota	0.000	0.105	Desulfobacterota <= 0.000170
6	Adenoma	0.840	0.252	0.176	0.455	Euryarchaeota	0.005	-0.022	0.0028 < Euryarchaeota <= 0.0152
6	Adenoma	0.840	0.252	0.176	0.455	Firmicutes	0.411	0.021	0.367 < Firmicutes <= 0.437
6	Adenoma	0.840	0.252	0.176	0.455	Actinobacteriota	0.004	0.020	Actinobacteriota <= 0.00434
9	Adenoma	0.880	0.123	0.248	0.299	Desulfobacterota	0.000	0.108	Desulfobacterota <= 0.000170
9	Adenoma	0.880	0.123	0.248	0.299	Verrucomicrobiota	0.000	-0.059	Verrucomicrobiota <= 4.71e-05
9	Adenoma	0.880	0.123	0.248	0.299	Euryarchaeota	0.000	-0.035	Euryarchaeota <= 0.0028
9	Adenoma	0.880	0.123	0.248	0.299	Fusobacteriota	0.000	0.021	1.87e-04 < Fusobacteriota
9	Adenoma	0.880	0.123	0.248	0.299	Actinobacteriota	0.004	0.016	Actinobacteriota <= 0.00434
11	Adenocarcinoma	0.848	0.069	0.729	0.774	Firmicutes	0.121	-0.051	Firmicutes <= 0.367

11	Adenocarcinoma	0.848	0.069	0.729	0.774	Euryarchaeota	0.001	0.041	Euryarchaeota <= 0.0028
11	Adenocarcinoma	0.848	0.069	0.729	0.774	Verrucomicrobiota	0.033	0.049	1.41e-02 < Verrucomicrobiota <= 4.92e-02
11	Adenocarcinoma	0.848	0.069	0.729	0.774	Desulfobacterota	0.005	0.042	0.002835 < Desulfobacterota
11	Adenocarcinoma	0.848	0.069	0.729	0.774	Fusobacteriota	0.001	-0.035	1.87e-04 < Fusobacteriota
17	Adenocarcinoma	0.864	0.059	0.727	0.786	Verrucomicrobiota	0.000	0.055	Verrucomicrobiota <= 4.71e-05
17	Adenocarcinoma	0.864	0.059	0.727	0.786	Firmicutes	0.281	-0.048	Firmicutes <= 0.367
17	Adenocarcinoma	0.864	0.059	0.727	0.786	Desulfobacterota	0.000	0.044	0.000170 < Desulfobacterota <= 0.000637
17	Adenocarcinoma	0.864	0.059	0.727	0.786	Euryarchaeota	0.000	0.030	Euryarchaeota <= 0.0028
17	Adenocarcinoma	0.864	0.059	0.727	0.786	Actinobacteriota	0.000	-0.021	Actinobacteriota <= 0.00434
18	Adenocarcinoma	0.862	0.097	0.741	0.737	Desulfobacterota	0.000	-0.093	Desulfobacterota <= 0.000170
18	Adenocarcinoma	0.862	0.097	0.741	0.737	Euryarchaeota	0.000	0.042	Euryarchaeota <= 0.0028
18	Adenocarcinoma	0.862	0.097	0.741	0.737	Verrucomicrobiota	0.027	0.047	1.41e-02 < Verrucomicrobiota <= 4.92e-02
18	Adenocarcinoma	0.862	0.097	0.741	0.737	Fusobacteriota	0.000	0.016	Fusobacteriota <= 4.83e-05
18	Adenocarcinoma	0.862	0.097	0.741	0.737	Actinobacteriota	0.001	-0.016	Actinobacteriota <= 0.00434
19	Adenocarcinoma	0.990	0.073	0.727	0.804	Euryarchaeota	0.034	-0.067	0.0152 < Euryarchaeota
19	Adenocarcinoma	0.990	0.073	0.727	0.804	Verrucomicrobiota	0.013	0.059	4.71e-05 < Verrucomicrobiota <= 1.41e-02
19	Adenocarcinoma	0.990	0.073	0.727	0.804	Firmicutes	0.701	0.046	0.567 < Firmicutes
19	Adenocarcinoma	0.990	0.073	0.727	0.804	Desulfobacterota	0.000	0.026	0.000170 < Desulfobacterota <= 0.000637
19	Adenocarcinoma	0.990	0.073	0.727	0.804	Fusobacteriota	0.000	0.012	Fusobacteriota <= 4.83e-05
20	Adenocarcinoma	0.914	0.076	0.725	0.792	Verrucomicrobiota	0.000	0.061	Verrucomicrobiota <= 4.71e-05
20	Adenocarcinoma	0.914	0.076	0.725	0.792	Firmicutes	0.202	-0.049	Firmicutes <= 0.367
20	Adenocarcinoma	0.914	0.076	0.725	0.792	Euryarchaeota	0.000	0.043	Euryarchaeota <= 0.0028
20	Adenocarcinoma	0.914	0.076	0.725	0.792	Desulfobacterota	0.001	0.035	0.000637 < Desulfobacterota <= 0.002835
20	Adenocarcinoma	0.914	0.076	0.725	0.792	Actinobacteriota	0.001	-0.024	Actinobacteriota <= 0.00434
21	Adenocarcinoma	0.964	0.224	0.754	0.704	Verrucomicrobiota	0.088	-0.173	4.92e-02 < Verrucomicrobiota
21	Adenocarcinoma	0.964	0.224	0.754	0.704	Euryarchaeota	0.000	0.032	Euryarchaeota <= 0.0028
21	Adenocarcinoma	0.964	0.224	0.754	0.704	Desulfobacterota	0.002	0.037	0.000637 < Desulfobacterota <= 0.002835
21	Adenocarcinoma	0.964	0.224	0.754	0.704	Firmicutes	0.494	0.030	0.437 < Firmicutes <= 0.567
21	Adenocarcinoma	0.964	0.224	0.754	0.704	Actinobacteriota	0.019	0.024	0.01308 < Actinobacteriota <= 0.05610
22	Adenocarcinoma	0.916	0.128	0.758	0.693	Desulfobacterota	0.000	-0.109	Desulfobacterota <= 0.000170
22	Adenocarcinoma	0.916	0.128	0.758	0.693	Verrucomicrobiota	0.000	0.064	Verrucomicrobiota <= 4.71e-05
22	Adenocarcinoma	0.916	0.128	0.758	0.693	Euryarchaeota	0.000	0.035	Euryarchaeota <= 0.0028
22	Adenocarcinoma	0.916	0.128	0.758	0.693	Firmicutes	0.234	-0.036	Firmicutes <= 0.367
22	Adenocarcinoma	0.916	0.128	0.758	0.693	Actinobacteriota	0.011	-0.019	0.00434 < Actinobacteriota <= 0.01308
23	Adenocarcinoma	0.840	0.206	0.785	0.657	Verrucomicrobiota	0.057	-0.150	4.92e-02 < Verrucomicrobiota
23	Adenocarcinoma	0.840	0.206	0.785	0.657	Euryarchaeota	0.062	-0.069	0.0152 < Euryarchaeota
23	Adenocarcinoma	0.840	0.206	0.785	0.657	Firmicutes	0.665	0.032	0.567 < Firmicutes
23	Adenocarcinoma	0.840	0.206	0.785	0.657	Desulfobacterota	0.000	0.035	0.000170 < Desulfobacterota <= 0.000637
23	Adenocarcinoma	0.840	0.206	0.785	0.657	Actinobacteriota	0.039	0.024	0.01308 < Actinobacteriota <= 0.05610
25	Adenocarcinoma	0.984	0.068	0.729	0.813	Euryarchaeota	0.029	-0.066	0.0152 < Euryarchaeota
25	Adenocarcinoma	0.984	0.068	0.729	0.813	Desulfobacterota	0.012	0.048	0.002835 < Desulfobacterota
25	Adenocarcinoma	0.984	0.068	0.729	0.813	Verrucomicrobiota	0.003	0.049	4.71e-05 < Verrucomicrobiota <= 1.41e-02
25	Adenocarcinoma	0.984	0.068	0.729	0.813	Firmicutes	0.475	0.030	0.437 < Firmicutes <= 0.567

25	Adenocarcinoma	0.984	0.068	0.729	0.813	Actinobacteriota	0.059	0.023	0.05610 < Actinobacteriota
26	Adenocarcinoma	0.990	0.222	0.788	0.645	Verrucomicrobiota	0.086	-0.159	4.92e-02 < Verrucomicrobiota
26	Adenocarcinoma	0.990	0.222	0.788	0.645	Euryarchaeota	0.022	-0.067	0.0152 < Euryarchaeota
26	Adenocarcinoma	0.990	0.222	0.788	0.645	Desulfobacterota	0.003	0.046	0.002835 < Desulfobacterota
26	Adenocarcinoma	0.990	0.222	0.788	0.645	Firmicutes	0.538	0.023	0.437 < Firmicutes <= 0.567
26	Adenocarcinoma	0.990	0.222	0.788	0.645	Fusobacteriota	0.000	0.013	4.83e-05 < Fusobacteriota <= 1.87e-04
29	Adenocarcinoma	0.872	0.251	0.808	0.595	Verrucomicrobiota	0.071	-0.152	4.92e-02 < Verrucomicrobiota
29	Adenocarcinoma	0.872	0.251	0.808	0.595	Desulfobacterota	0.000	-0.109	Desulfobacterota <= 0.000170
29	Adenocarcinoma	0.872	0.251	0.808	0.595	Actinobacteriota	0.020	0.033	0.01308 < Actinobacteriota <= 0.05610
29	Adenocarcinoma	0.872	0.251	0.808	0.595	Euryarchaeota	0.003	0.029	0.0028 < Euryarchaeota <= 0.0152
29	Adenocarcinoma	0.872	0.251	0.808	0.595	Firmicutes	0.416	-0.014	0.367 < Firmicutes <= 0.437
30	Adenocarcinoma	0.990	0.072	0.725	0.826	Euryarchaeota	0.015	-0.065	0.0152 < Euryarchaeota
30	Adenocarcinoma	0.990	0.072	0.725	0.826	Verrucomicrobiota	0.011	0.053	4.71e-05 < Verrucomicrobiota <= 1.41e-02
30	Adenocarcinoma	0.990	0.072	0.725	0.826	Firmicutes	0.648	0.040	0.567 < Firmicutes
30	Adenocarcinoma	0.990	0.072	0.725	0.826	Desulfobacterota	0.001	0.042	0.000637 < Desulfobacterota <= 0.002835
30	Adenocarcinoma	0.990	0.072	0.725	0.826	Actinobacteriota	0.013	0.031	0.01308 < Actinobacteriota <= 0.05610
31	Adenocarcinoma	0.980	0.052	0.701	0.872	Verrucomicrobiota	0.000	0.056	4.71e-05 < Verrucomicrobiota <= 1.41e-02
31	Adenocarcinoma	0.980	0.052	0.701	0.872	Euryarchaeota	0.000	0.044	Euryarchaeota <= 0.0028
31	Adenocarcinoma	0.980	0.052	0.701	0.872	Desulfobacterota	0.001	0.032	0.000637 < Desulfobacterota <= 0.002835
31	Adenocarcinoma	0.980	0.052	0.701	0.872	Actinobacteriota	0.080	0.022	0.05610 < Actinobacteriota
31	Adenocarcinoma	0.980	0.052	0.701	0.872	Firmicutes	0.459	0.017	0.437 < Firmicutes <= 0.567
35	Adenocarcinoma	0.896	0.046	0.711	0.846	Verrucomicrobiota	0.049	0.055	1.41e-02 < Verrucomicrobiota <= 4.92e-02
35	Adenocarcinoma	0.896	0.046	0.711	0.846	Desulfobacterota	0.004	0.043	0.002835 < Desulfobacterota
35	Adenocarcinoma	0.896	0.046	0.711	0.846	Firmicutes	0.683	0.038	0.567 < Firmicutes
35	Adenocarcinoma	0.896	0.046	0.711	0.846	Actinobacteriota	0.013	-0.023	0.00434 < Actinobacteriota <= 0.01308
35	Adenocarcinoma	0.896	0.046	0.711	0.846	Euryarchaeota	0.008	0.021	0.0028 < Euryarchaeota <= 0.0152
41	Adenocarcinoma	0.996	0.055	0.696	0.899	Verrucomicrobiota	0.010	0.061	4.71e-05 < Verrucomicrobiota <= 1.41e-02
41	Adenocarcinoma	0.996	0.055	0.696	0.899	Firmicutes	0.576	0.044	0.567 < Firmicutes
41	Adenocarcinoma	0.996	0.055	0.696	0.899	Desulfobacterota	0.000	0.042	0.000170 < Desulfobacterota <= 0.000637
41	Adenocarcinoma	0.996	0.055	0.696	0.899	Euryarchaeota	0.003	0.033	0.0028 < Euryarchaeota <= 0.0152
41	Adenocarcinoma	0.996	0.055	0.696	0.899	Actinobacteriota	0.049	0.023	0.01308 < Actinobacteriota <= 0.05610
42	Adenocarcinoma	0.948	0.061	0.701	0.852	Verrucomicrobiota	0.000	0.060	Verrucomicrobiota <= 4.71e-05
42	Adenocarcinoma	0.948	0.061	0.701	0.852	Euryarchaeota	0.000	0.040	Euryarchaeota <= 0.0028
42	Adenocarcinoma	0.948	0.061	0.701	0.852	Desulfobacterota	0.004	0.045	0.002835 < Desulfobacterota
42	Adenocarcinoma	0.948	0.061	0.701	0.852	Firmicutes	0.593	0.027	0.567 < Firmicutes
42	Adenocarcinoma	0.948	0.061	0.701	0.852	Actinobacteriota	0.004	-0.020	0.00434 < Actinobacteriota <= 0.01308
43	Adenocarcinoma	0.954	0.033	0.726	0.834	Verrucomicrobiota	0.036	0.051	1.41e-02 < Verrucomicrobiota <= 4.92e-02
43	Adenocarcinoma	0.954	0.033	0.726	0.834	Desulfobacterota	0.001	0.034	0.000170 < Desulfobacterota <= 0.000637
43	Adenocarcinoma	0.954	0.033	0.726	0.834	Euryarchaeota	0.015	0.023	0.0028 < Euryarchaeota <= 0.0152
43	Adenocarcinoma	0.954	0.033	0.726	0.834	Fusobacteriota	0.000	0.016	Fusobacteriota <= 4.83e-05
43	Adenocarcinoma	0.954	0.033	0.726	0.834	Firmicutes	0.408	-0.017	0.367 < Firmicutes <= 0.437

Table ML-A-5: The LIME Result for Test Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
4	Adenoma	0.738	0.118	0.256	0.737	Desulfobacterota	0.000	0.109	Desulfobacterota <= 0.000170
4	Adenoma	0.738	0.118	0.256	0.737	Verrucomicrobiota	0.000	-0.055	4.71e-05 < Verrucomicrobiota <= 1.41e-02
4	Adenoma	0.738	0.118	0.256	0.737	Euryarchaeota	0.000	-0.036	Euryarchaeota <= 0.0028
4	Adenoma	0.738	0.118	0.256	0.737	Actinobacteriota	0.033	-0.023	0.01308 < Actinobacteriota <= 0.05610
4	Adenoma	0.738	0.118	0.256	0.737	Firmicutes	0.754	0.485	0.567 < Firmicutes
5	Adenocarcinoma	0.890	0.067	0.716	0.824	Euryarchaeota	0.000	0.046	Euryarchaeota <= 0.0028
5	Adenocarcinoma	0.890	0.067	0.716	0.824	Firmicutes	0.271	-0.050	Firmicutes <= 0.367
5	Adenocarcinoma	0.890	0.067	0.716	0.824	Verrucomicrobiota	0.002	0.047	4.71e-05 < Verrucomicrobiota <= 1.41e-02
5	Adenocarcinoma	0.890	0.067	0.716	0.824	Desulfobacterota	0.001	0.039	0.000170 < Desulfobacterota <= 0.000637
5	Adenocarcinoma	0.890	0.067	0.716	0.824	Actinobacteriota	0.014	0.026	0.01308 < Actinobacteriota <= 0.05610
7	Adenocarcinoma	0.866	0.056	0.718	0.808	Verrucomicrobiota	0.010	0.057	4.71e-05 < Verrucomicrobiota <= 1.41e-02
7	Adenocarcinoma	0.866	0.056	0.718	0.808	Firmicutes	0.183	-0.051	Firmicutes <= 0.367
7	Adenocarcinoma	0.866	0.056	0.718	0.808	Euryarchaeota	0.000	0.034	Euryarchaeota <= 0.0028
7	Adenocarcinoma	0.866	0.056	0.718	0.808	Desulfobacterota	0.002	0.033	0.000637 < Desulfobacterota <= 0.002835
7	Adenocarcinoma	0.866	0.056	0.718	0.808	Actinobacteriota	0.056	0.018	0.05610 < Actinobacteriota
8	Adenocarcinoma	0.964	0.061	0.696	0.886	Verrucomicrobiota	0.000	0.058	Verrucomicrobiota <= 4.71e-05
8	Adenocarcinoma	0.964	0.061	0.696	0.886	Firmicutes	0.666	0.040	0.567 < Firmicutes
8	Adenocarcinoma	0.964	0.061	0.696	0.886	Desulfobacterota	0.004	0.040	0.002835 < Desulfobacterota
8	Adenocarcinoma	0.964	0.061	0.696	0.886	Euryarchaeota	0.000	0.032	Euryarchaeota <= 0.0028
8	Adenocarcinoma	0.964	0.061	0.696	0.886	Actinobacteriota	0.062	0.020	0.05610 < Actinobacteriota
10	Adenocarcinoma	0.896	0.065	0.700	0.857	Verrucomicrobiota	0.002	0.062	4.71e-05 < Verrucomicrobiota <= 1.41e-02
10	Adenocarcinoma	0.896	0.065	0.700	0.857	Euryarchaeota	0.000	0.044	Euryarchaeota <= 0.0028
10	Adenocarcinoma	0.896	0.065	0.700	0.857	Firmicutes	0.589	0.040	0.567 < Firmicutes
10	Adenocarcinoma	0.896	0.065	0.700	0.857	Desulfobacterota	0.000	0.036	0.000170 < Desulfobacterota <= 0.000637
10	Adenocarcinoma	0.896	0.065	0.700	0.857	Fusobacteriota	0.023	-0.024	1.87e-04 < Fusobacteriota
12	Adenocarcinoma	0.876	0.043	0.705	0.877	Verrucomicrobiota	0.016	0.049	1.41e-02 < Verrucomicrobiota <= 4.92e-02
12	Adenocarcinoma	0.876	0.043	0.705	0.877	Euryarchaeota	0.000	0.033	Euryarchaeota <= 0.0028
12	Adenocarcinoma	0.876	0.043	0.705	0.877	Desulfobacterota	0.001	0.033	0.000637 < Desulfobacterota <= 0.002835
12	Adenocarcinoma	0.876	0.043	0.705	0.877	Firmicutes	0.522	0.031	0.437 < Firmicutes <= 0.567
12	Adenocarcinoma	0.876	0.043	0.705	0.877	Actinobacteriota	0.015	0.025	0.01308 < Actinobacteriota <= 0.05610
13	Adenocarcinoma	0.916	0.031	0.737	0.801	Verrucomicrobiota	0.028	0.047	1.41e-02 < Verrucomicrobiota <= 4.92e-02
13	Adenocarcinoma	0.916	0.031	0.737	0.801	Desulfobacterota	0.001	0.031	0.000637 < Desulfobacterota <= 0.002835
13	Adenocarcinoma	0.916	0.031	0.737	0.801	Firmicutes	0.509	0.028	0.437 < Firmicutes <= 0.567
13	Adenocarcinoma	0.916	0.031	0.737	0.801	Proteobacteria	0.055	-0.021	0.05347 < Proteobacteria
13	Adenocarcinoma	0.916	0.031	0.737	0.801	Fusobacteriota	0.000	-0.021	1.87e-04 < Fusobacteriota
14	Adenocarcinoma	0.992	0.044	0.707	0.888	Verrucomicrobiota	0.029	0.051	1.41e-02 < Verrucomicrobiota <= 4.92e-02
14	Adenocarcinoma	0.992	0.044	0.707	0.888	Desulfobacterota	0.003	0.044	0.002835 < Desulfobacterota
14	Adenocarcinoma	0.992	0.044	0.707	0.888	Euryarchaeota	0.014	0.032	0.0028 < Euryarchaeota <= 0.0152
14	Adenocarcinoma	0.992	0.044	0.707	0.888	Firmicutes	0.542	0.031	0.437 < Firmicutes <= 0.567
14	Adenocarcinoma	0.992	0.044	0.707	0.888	Actinobacteriota	0.027	0.023	0.01308 < Actinobacteriota <= 0.05610
15	Adenocarcinoma	0.940	0.212	0.776	0.651	Verrucomicrobiota	0.144	-0.172	4.92e-02 < Verrucomicrobiota
15	Adenocarcinoma	0.940	0.212	0.776	0.651	Actinobacteriota	0.022	0.027	0.01308 < Actinobacteriota <= 0.05610

15	Adenocarcinoma	0.940	0.212	0.776	0.651	Desulfobacterota	0.002	0.025	0.000637 < Desulfobacterota <= 0.002835
15	Adenocarcinoma	0.940	0.212	0.776	0.651	Firmicutes	0.386	-0.018	0.367 < Firmicutes <= 0.437
15	Adenocarcinoma	0.940	0.212	0.776	0.651	Fusobacteriota	0.000	0.014	Fusobacteriota <= 4.83e-05
16	Adenocarcinoma	0.766	0.088	0.752	0.736	Euryarchaeota	0.046	-0.067	0.0152 < Euryarchaeota
16	Adenocarcinoma	0.766	0.088	0.752	0.736	Verrucomicrobiota	0.000	0.069	Verrucomicrobiota <= 4.71e-05
16	Adenocarcinoma	0.766	0.088	0.752	0.736	Desulfobacterota	0.007	0.040	0.002835 < Desulfobacterota
16	Adenocarcinoma	0.766	0.088	0.752	0.736	Actinobacteriota	0.008	-0.031	0.00434 < Actinobacteriota <= 0.01308
16	Adenocarcinoma	0.766	0.088	0.752	0.736	Fusobacteriota	0.000	-0.027	1.87e-04 < Fusobacteriota
24	Adenocarcinoma	0.986	0.073	0.681	0.895	Verrucomicrobiota	0.000	0.071	Verrucomicrobiota <= 4.71e-05
24	Adenocarcinoma	0.986	0.073	0.681	0.895	Desulfobacterota	0.008	0.046	0.002835 < Desulfobacterota
24	Adenocarcinoma	0.986	0.073	0.681	0.895	Euryarchaeota	0.002	0.038	Euryarchaeota <= 0.0028
24	Adenocarcinoma	0.986	0.073	0.681	0.895	Firmicutes	0.600	0.035	0.567 < Firmicutes
24	Adenocarcinoma	0.986	0.073	0.681	0.895	Actinobacteriota	0.059	0.023	0.05610 < Actinobacteriota
27	Adenocarcinoma	0.872	0.057	0.704	0.864	Verrucomicrobiota	0.000	0.071	Verrucomicrobiota <= 4.71e-05
27	Adenocarcinoma	0.872	0.057	0.704	0.864	Desulfobacterota	0.001	0.040	0.000170 < Desulfobacterota <= 0.000637
27	Adenocarcinoma	0.872	0.057	0.704	0.864	Euryarchaeota	0.000	0.032	Euryarchaeota <= 0.0028
27	Adenocarcinoma	0.872	0.057	0.704	0.864	Firmicutes	0.548	0.023	0.437 < Firmicutes <= 0.567
27	Adenocarcinoma	0.872	0.057	0.704	0.864	Synergistota	0.000	-0.006	Synergistota <= 0.000618
28	Adenocarcinoma	0.956	0.050	0.691	0.885	Verrucomicrobiota	0.000	0.052	4.71e-05 < Verrucomicrobiota <= 1.41e-02
28	Adenocarcinoma	0.956	0.050	0.691	0.885	Desulfobacterota	0.004	0.049	0.002835 < Desulfobacterota
28	Adenocarcinoma	0.956	0.050	0.691	0.885	Euryarchaeota	0.000	0.030	Euryarchaeota <= 0.0028
28	Adenocarcinoma	0.956	0.050	0.691	0.885	Firmicutes	0.515	0.035	0.437 < Firmicutes <= 0.567
28	Adenocarcinoma	0.956	0.050	0.691	0.885	Actinobacteriota	0.019	0.028	0.01308 < Actinobacteriota <= 0.05610
32	Adenocarcinoma	0.798	0.108	0.745	0.722	Desulfobacterota	0.000	-0.089	Desulfobacterota <= 0.000170
32	Adenocarcinoma	0.798	0.108	0.745	0.722	Verrucomicrobiota	0.010	0.060	4.71e-05 < Verrucomicrobiota <= 1.41e-02
32	Adenocarcinoma	0.798	0.108	0.745	0.722	Firmicutes	0.082	-0.045	Firmicutes <= 0.367
32	Adenocarcinoma	0.798	0.108	0.745	0.722	Euryarchaeota	0.000	0.032	Euryarchaeota <= 0.0028
32	Adenocarcinoma	0.798	0.108	0.745	0.722	Fusobacteriota	0.000	0.019	Fusobacteriota <= 4.83e-05
33	Adenocarcinoma	0.800	0.056	0.695	0.881	Verrucomicrobiota	0.001	0.049	4.71e-05 < Verrucomicrobiota <= 1.41e-02
33	Adenocarcinoma	0.800	0.056	0.695	0.881	Euryarchaeota	0.000	0.042	Euryarchaeota <= 0.0028
33	Adenocarcinoma	0.800	0.056	0.695	0.881	Desulfobacterota	0.003	0.042	0.002835 < Desulfobacterota
33	Adenocarcinoma	0.800	0.056	0.695	0.881	Firmicutes	0.579	0.033	0.567 < Firmicutes
33	Adenocarcinoma	0.800	0.056	0.695	0.881	Actinobacteriota	0.015	0.019	0.01308 < Actinobacteriota <= 0.05610
34	Adenocarcinoma	0.884	0.048	0.729	0.827	Verrucomicrobiota	0.030	0.051	1.41e-02 < Verrucomicrobiota <= 4.92e-02
34	Adenocarcinoma	0.884	0.048	0.729	0.827	Firmicutes	0.329	-0.043	Firmicutes <= 0.367
34	Adenocarcinoma	0.884	0.048	0.729	0.827	Desulfobacterota	0.000	0.041	0.000170 < Desulfobacterota <= 0.000637
34	Adenocarcinoma	0.884	0.048	0.729	0.827	Euryarchaeota	0.003	0.025	0.0028 < Euryarchaeota <= 0.0152
34	Adenocarcinoma	0.884	0.048	0.729	0.827	Actinobacteriota	0.043	0.025	0.01308 < Actinobacteriota <= 0.05610
36	Adenoma	0.622	0.058	0.232	0.302	Euryarchaeota	0.051	0.077	0.0152 < Euryarchaeota
36	Adenoma	0.622	0.058	0.232	0.302	Desulfobacterota	0.001	-0.036	0.000637 < Desulfobacterota <= 0.002835
36	Adenoma	0.622	0.058	0.232	0.302	Actinobacteriota	0.002	0.030	Actinobacteriota <= 0.00434
36	Adenoma	0.622	0.058	0.232	0.302	Fusobacteriota	0.000	-0.015	Fusobacteriota <= 4.83e-05
36	Adenoma	0.622	0.058	0.232	0.302	Synergistota	0.000	0.013	Synergistota <= 0.000618

37	Adenocarcinoma	0.864	0.127	0.746	0.737	Desulfobacterota	0.000	-0.108	Desulfobacterota <= 0.000170
37	Adenocarcinoma	0.864	0.127	0.746	0.737	Verrucomicrobiota	0.001	0.049	4.71e-05 < Verrucomicrobiota <= 1.41e-02
37	Adenocarcinoma	0.864	0.127	0.746	0.737	Euryarchaeota	0.001	0.037	Euryarchaeota <= 0.0028
37	Adenocarcinoma	0.864	0.127	0.746	0.737	Firmicutes	0.672	0.039	0.567 < Firmicutes
37	Adenocarcinoma	0.864	0.127	0.746	0.737	Fusobacteriota	0.015	-0.026	1.87e-04 < Fusobacteriota
38	Adenocarcinoma	0.870	0.117	0.724	0.796	Desulfobacterota	0.000	-0.093	Desulfobacterota <= 0.000170
38	Adenocarcinoma	0.870	0.117	0.724	0.796	Verrucomicrobiota	0.000	0.061	Verrucomicrobiota <= 4.71e-05
38	Adenocarcinoma	0.870	0.117	0.724	0.796	Firmicutes	0.654	0.043	0.567 < Firmicutes
38	Adenocarcinoma	0.870	0.117	0.724	0.796	Euryarchaeota	0.000	0.035	Euryarchaeota <= 0.0028
38	Adenocarcinoma	0.870	0.117	0.724	0.796	Actinobacteriota	0.015	0.026	0.01308 < Actinobacteriota <= 0.05610
39	Adenocarcinoma	0.924	0.038	0.721	0.839	Verrucomicrobiota	0.024	0.047	1.41e-02 < Verrucomicrobiota <= 4.92e-02
39	Adenocarcinoma	0.924	0.038	0.721	0.839	Desulfobacterota	0.000	0.042	0.000170 < Desulfobacterota <= 0.000637
39	Adenocarcinoma	0.924	0.038	0.721	0.839	Firmicutes	0.498	0.031	0.437 < Firmicutes <= 0.567
39	Adenocarcinoma	0.924	0.038	0.721	0.839	Proteobacteria	0.283	-0.022	0.05347 < Proteobacteria
39	Adenocarcinoma	0.924	0.038	0.721	0.839	Euryarchaeota	0.009	0.021	0.0028 < Euryarchaeota <= 0.0152
40	Adenoma	0.558	0.233	0.182	0.440	Verrucomicrobiota	0.276	0.160	4.92e-02 < Verrucomicrobiota
40	Adenoma	0.558	0.233	0.182	0.440	Euryarchaeota	0.029	0.065	0.0152 < Euryarchaeota
40	Adenoma	0.558	0.233	0.182	0.440	Firmicutes	0.351	0.050	Firmicutes <= 0.367
40	Adenoma	0.558	0.233	0.182	0.440	Desulfobacterota	0.001	-0.043	0.000170 < Desulfobacterota <= 0.000637
40	Adenoma	0.558	0.233	0.182	0.440	Actinobacteriota	0.002	0.025	Actinobacteriota <= 0.00434

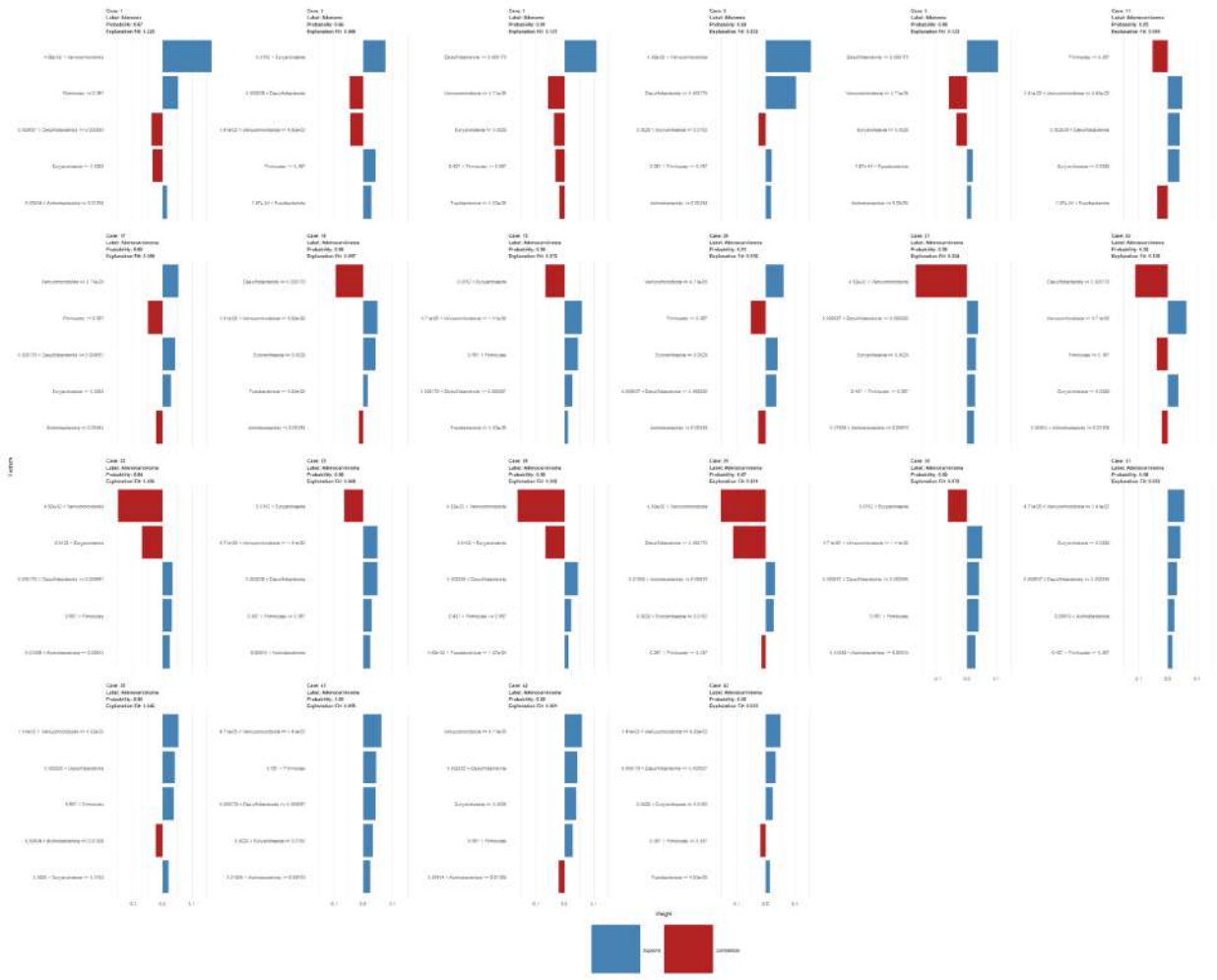


Figure ML2-A-6: The LIME Result in Training dataset for Phylum in Stool

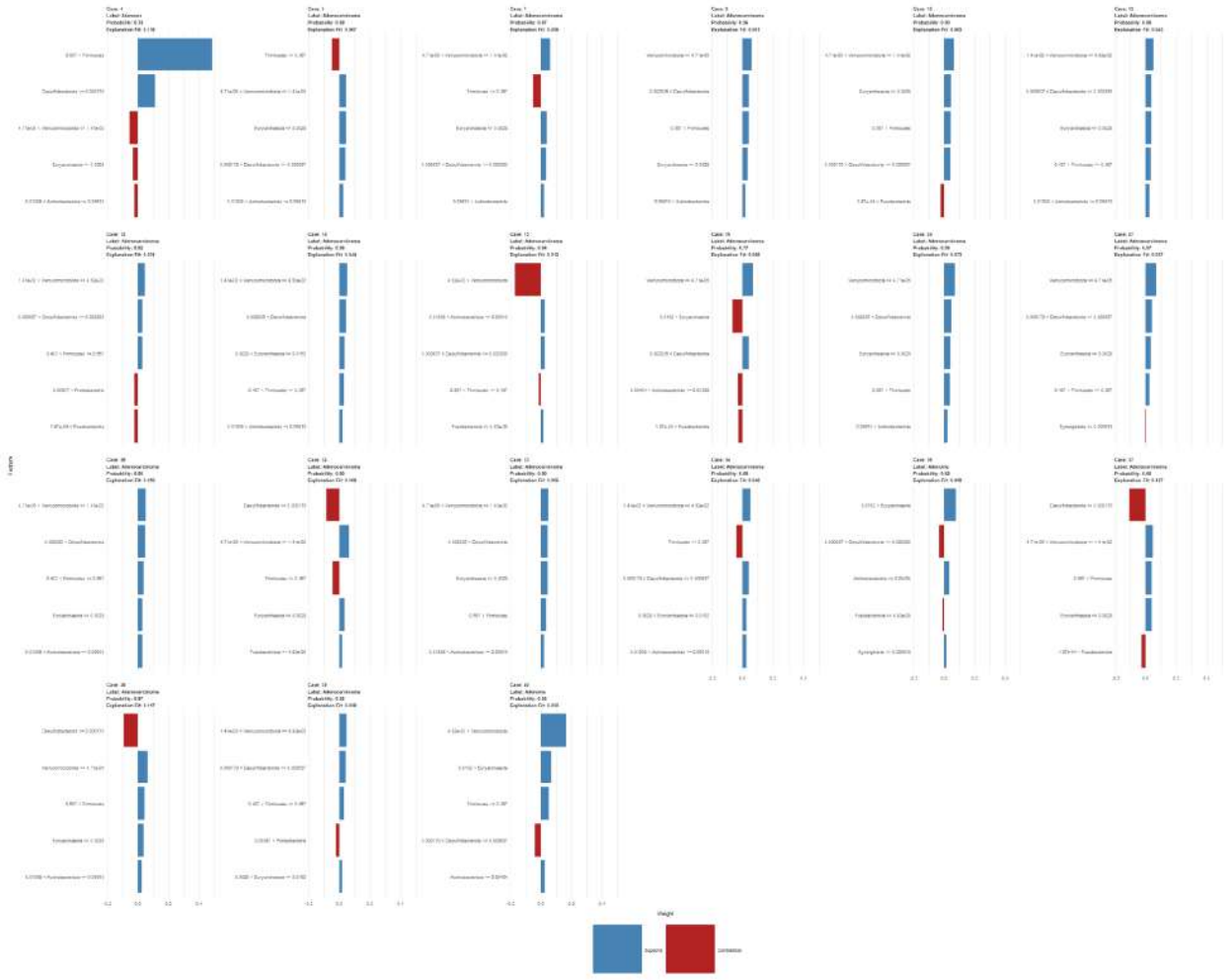


Figure ML2-A-7: The LIME Result in Test dataset for Phylum in Stool

Table ML2-A-6: The Aggregated SHAP Values

Variable	Training Dataset				Testing Dataset			
	contribution	position	sign	cumulative	contribution	position	sign	cumulative
intercept	0.7702	16	X	0.770	0.7702	16	X	0.7702
Firmicutes	-0.0003	15	-1	0.770	0.0082	15	1	0.7784
Desulfobacterota	0.0005	14	1	0.770	0.0505	14	1	0.8289
Verrucomicrobiota	-0.0005	13	-1	0.770	-0.0148	10	-1	0.8168
Fusobacteriota	0.0000	12	-1	0.770	0.0000	11	1	0.8317

Campilobacterota	0.0001	11	1	0.770	0.0028	12	1	0.8317
Tumor	0.0000	10	1	0.770	0.0000	9	1	0.8168
Patescibacteria	-0.0002	9	-1	0.770	-0.0031	8	-1	0.8138
Gender	0.0000	8	1	0.770	0.0000	13	1	0.8289
Cyanobacteria	-0.0001	7	-1	0.770	-0.0001	6	-1	0.8132
Euryarchaeota	0.0002	6	1	0.770	-0.0004	7	-1	0.8133
Synergistota	0.0001	5	1	0.770	0.0000	5	1	0.8133
Proteobacteria	0.0001	4	1	0.770	-0.0041	4	-1	0.8091
Actinobacteriota	0.0001	3	1	0.770	0.0062	3	1	0.8154
Bacteroidota	0.0001	2	1	0.770	0.0009	2	1	0.8163
prediction	0.7702	1	X	0.770	0.8163	1	X	0.8163

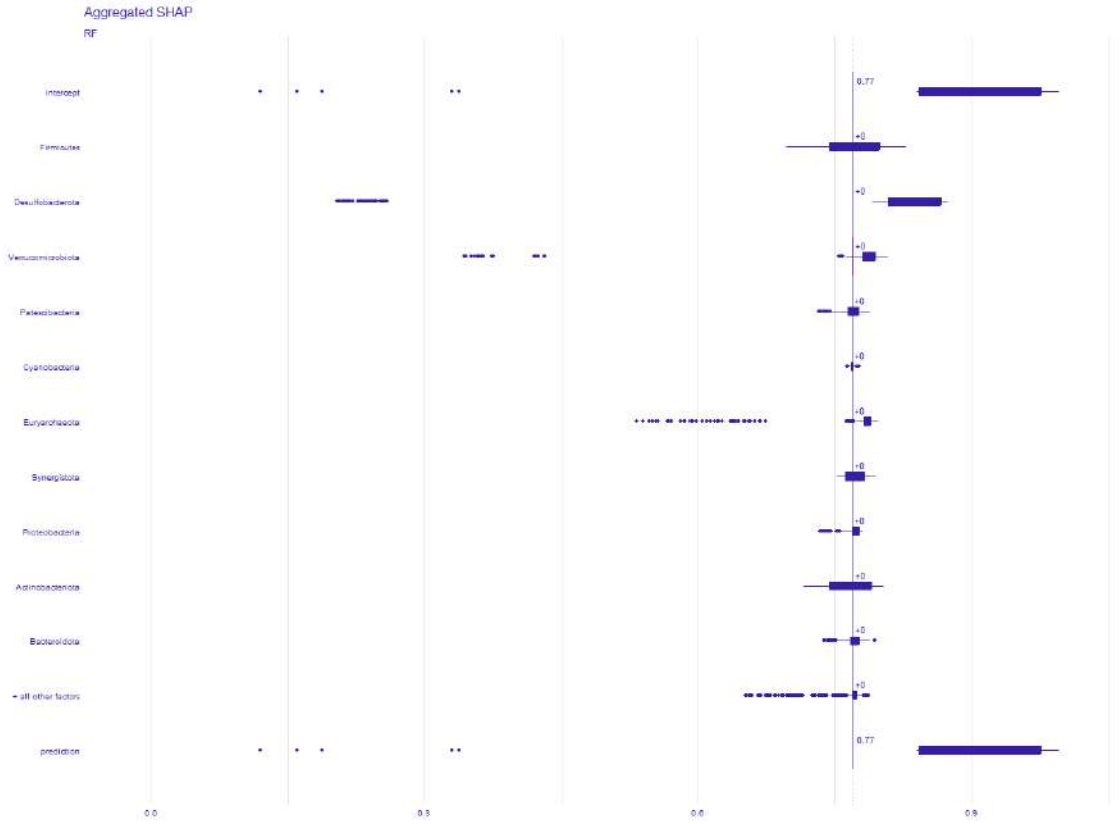


Figure ML2-A-8: The Aggregated SHAP Value in Training dataset for Phylum in Stool

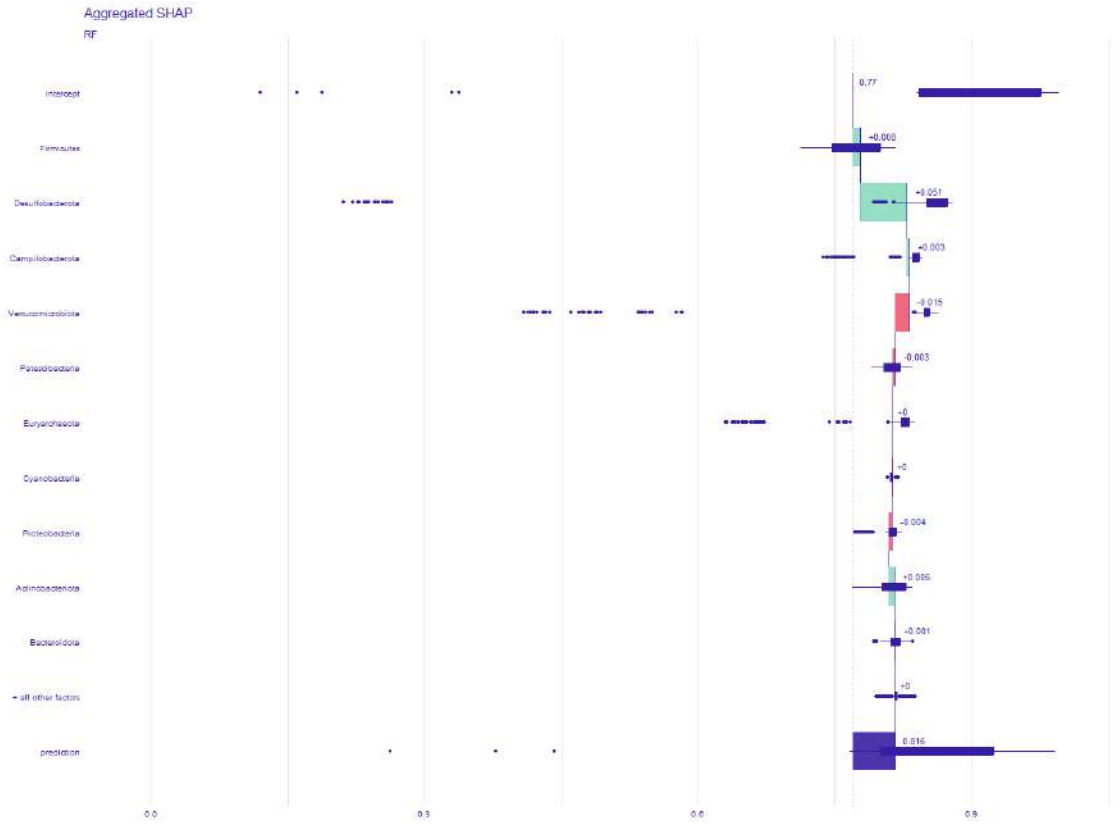


Figure ML2-A-9: The Aggregated SHAP Value in Test dataset for Phylum in Stool

B- Class

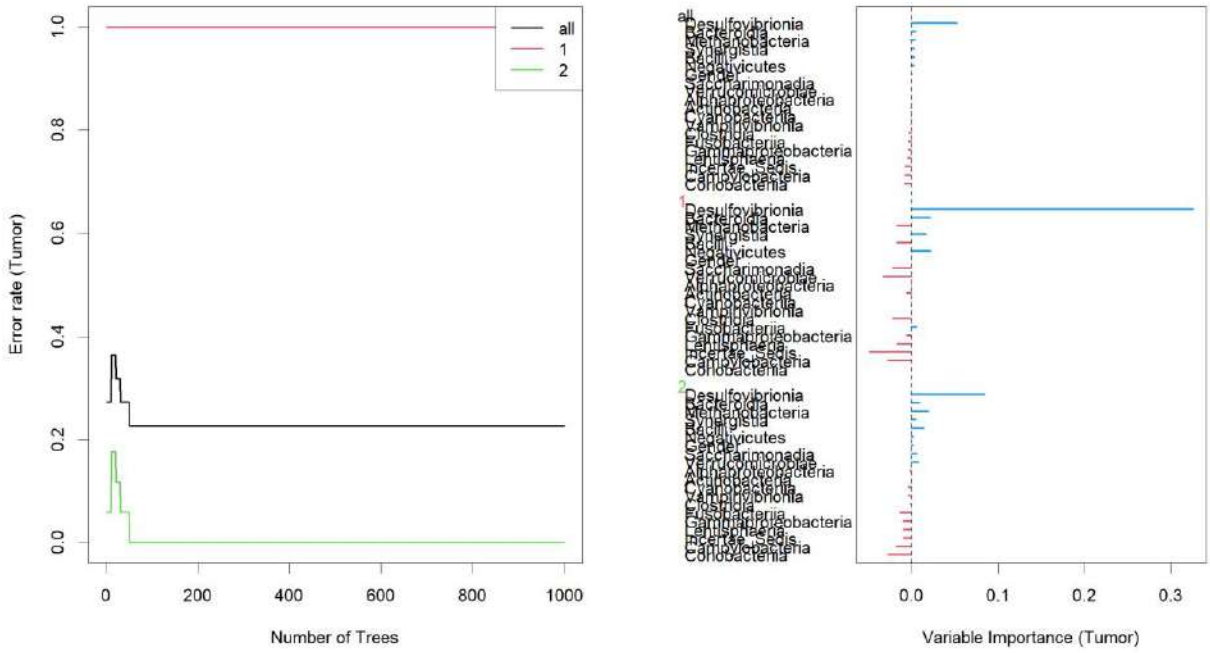
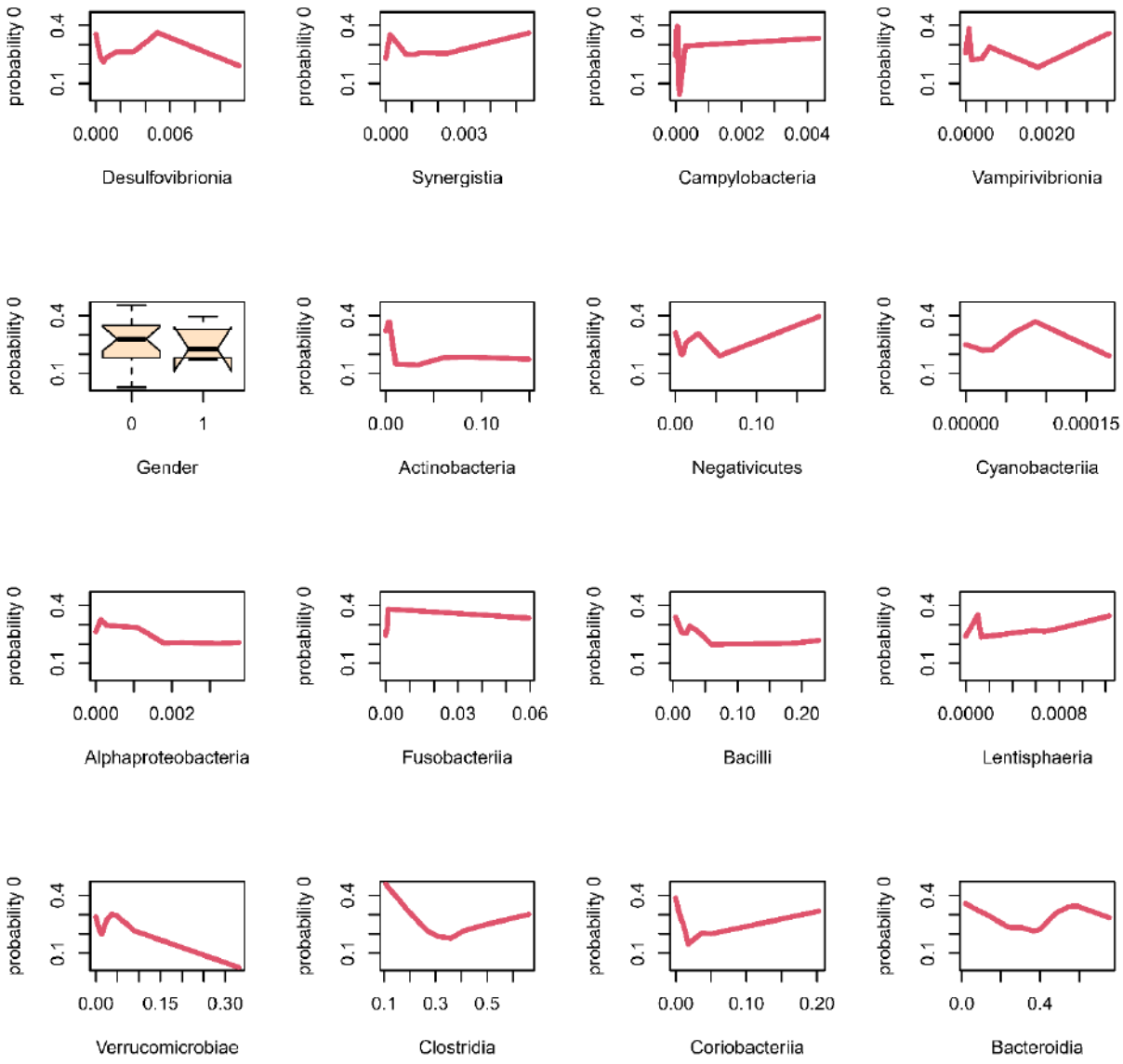


Figure ML2-B-1: The number of trees and Variable Importance Plot for Class in Stool


probability 0
0.1
0.4
0.000
0.006
Desulfovibrionia
0.000
0.003
Synergistia
0.000
0.002
0.004
Campylobacteria
0.0000
0.0020
Vampirivibrionia
0
1
Gender
0.00
0.10
Actinobacteria
0.00
0.10
Negativicutes
0.00000
0.00015
Cyanobacteriia
0.000
0.002
Alphaproteobacteria
probability 0
0.00
0.03
0.06
Fusobacteriia
0.00
0.10
0.20
Bacilli
0.0000
0.0008
Lentisphaeria
0.00
0.15
0.30
Verrucomicrobiae
0.1
0.3
0.5
Clostridia
0.00
0.10
0.20
Coriobacteriia
0.0
0.4
Bacteroidia

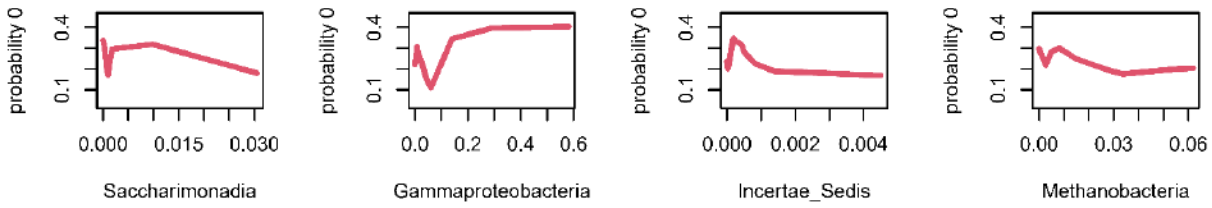
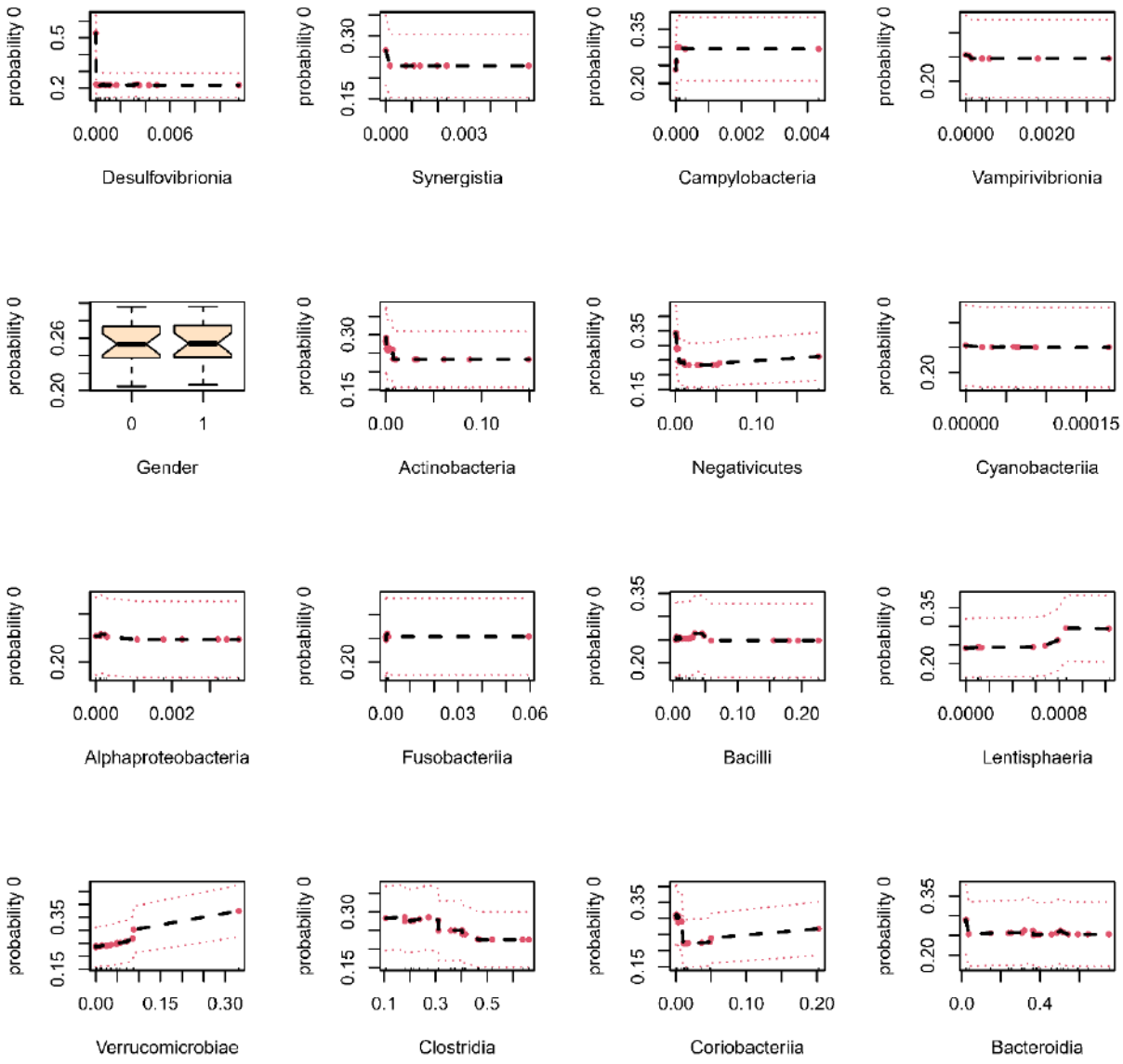


Figure ML2-B-2: Marginal Plot for Class in Stool


probability 0
0.2
0.5
0.000
0.006
Desulfovibrionia
probability 0
0.15
0.30
0.000
0.003
Synergistia
probability 0
0.20
0.35
0.000
0.002
0.004
Campylobacteria
probability 0
0.20
0.0000
0.0020
Vampirivibrionia
probability 0
0.20
0.26
0
1
Gender
probability 0
0.15
0.30
0.00
0.10
Actinobacteria
probability 0
0.15
0.35
0.00
0.10
Negativicutes
probability 0
0.20
0.00000
0.00015
Cyanobacteriia
probability 0
0.20
0.000
0.002
Alphaproteobacteria
probability 0
0.20
0.00
0.03
0.06
Fusobacteriia
probability 0
0.20
0.35
0.00
0.10
0.20
Bacilli
probability 0
0.20
0.35
0.0000
0.0008
Lentisphaeria
probability 0
0.15
0.35
0.00
0.15
0.30
Verrucomicrobiae
probability 0
0.15
0.30
0.1
0.3
0.5
Clostridia
probability 0
0.15
0.35
0.00
0.10
0.20
Coriobacteriia
probability 0
0.20
0.35
0.0
0.4
Bacteroidia

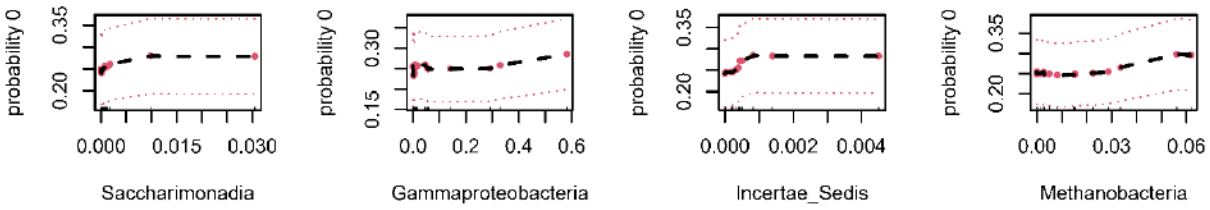


Figure ML2-B-3: Partial Plot for Class in Stool

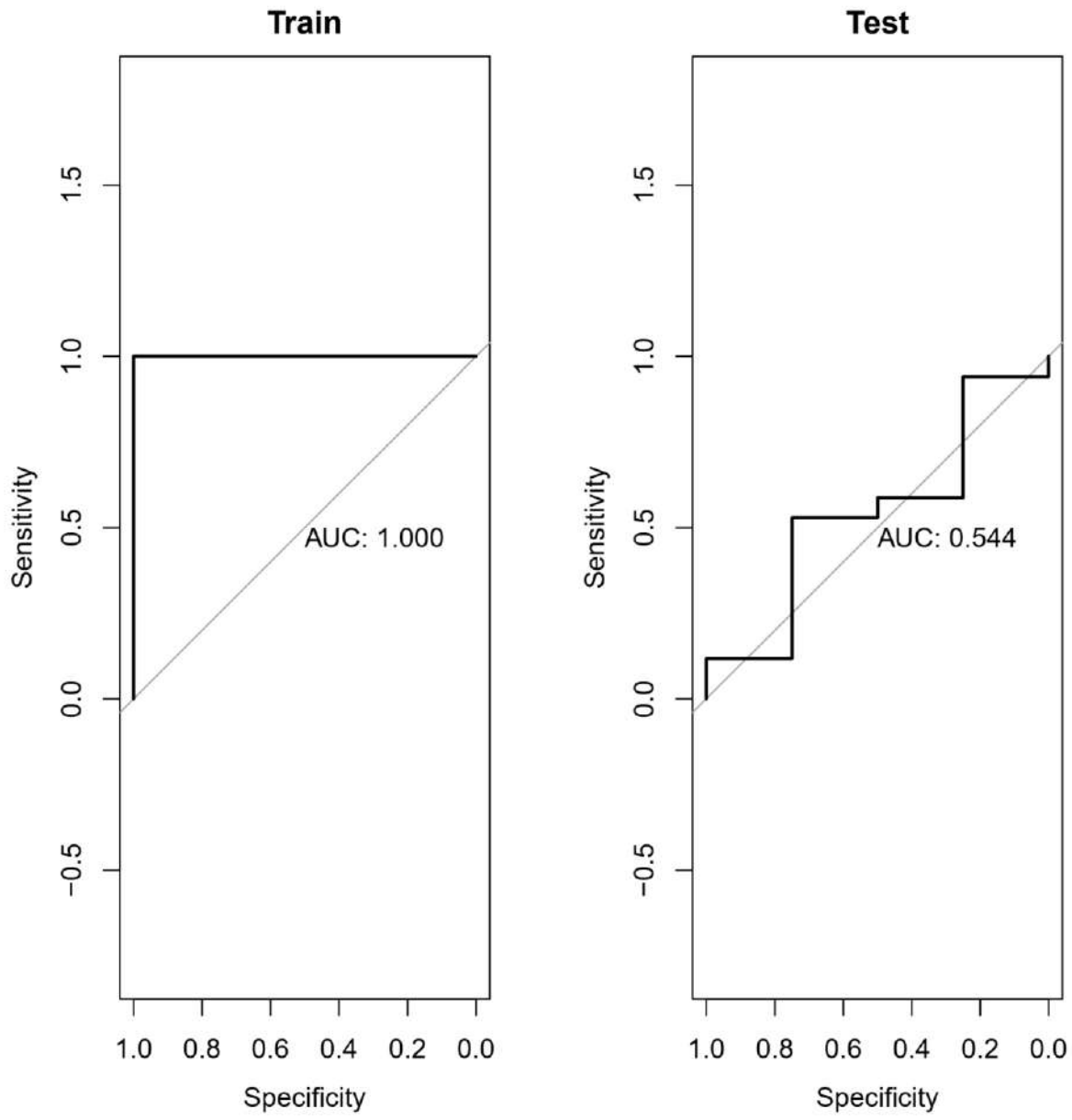


Figure ML2-B-4: The ROC Curve for Class in Stool

Table ML2-B-1: The Model Performance

Index	Values		
	2.50%	**50.00%**	**97.50%**
Accuracy	0.24	0.67	0.90
Specificity	0.25	0.75	1.00
Sensitivity	0.06	0.59	1.00

Table ML2-B-2: The Variable Importance Values

Covariates	Varibale Importance			Brier Score		
	all	**AP**	**CRC**	**all**	**AP**	**CRC**
Gender	-0.0010	-0.0054	-0.0016	-0.0088	-0.0060	-0.0060
Actinobacteria	-0.0199	-0.0544	-0.0528	-0.0158	-0.0107	-0.0107
Alphaproteobacteria	-0.0051	0.0000	-0.0176	-0.0126	-0.0086	-0.0086
Bacilli	-0.0054	-0.0489	-0.0048	-0.0336	-0.0229	-0.0229
Bacteroidia	0.0069	-0.0815	0.0480	0.0090	0.0061	0.0061
Campylobacteria	0.0111	-0.0435	0.0512	0.0185	0.0126	0.0126
Clostridia	-0.0110	-0.0598	-0.0208	-0.0186	-0.0126	-0.0126
Coriobacteriia	0.0199	-0.1848	0.1231	0.0088	0.0060	0.0060
Cyanobacteriia	-0.0073	-0.0054	-0.0240	-0.0189	-0.0129	-0.0129
Desulfovibrionia	0.2770	-0.0163	0.9642	0.7726	0.5250	0.5250
Fusobacteriia	-0.0060	0.0000	-0.0208	-0.0197	-0.0134	-0.0134
Gammaproteobacteria	0.0031	-0.1305	0.0496	-0.0363	-0.0247	-0.0247
Incertae_Sedis	0.0064	-0.0489	0.0368	0.0001	0.0000	0.0000
Lentisphaeria	0.0112	-0.1196	0.0736	0.0181	0.0123	0.0123
Methanobacteria	0.0212	-0.1033	0.1039	0.0264	0.0180	0.0180
Negativicutes	0.0410	-0.2012	0.2015	0.0500	0.0340	0.0340
Saccharimonadia	-0.0002	-0.0870	0.0256	-0.0304	-0.0207	-0.0207
Synergistia	0.0041	0.0272	0.0064	-0.0038	-0.0026	-0.0026
Vampirivibrionia	-0.0028	-0.0054	-0.0080	-0.0072	-0.0049	-0.0049
Verrucomicrobiae	0.0590	-0.3425	0.3038	0.1063	0.0722	0.0722

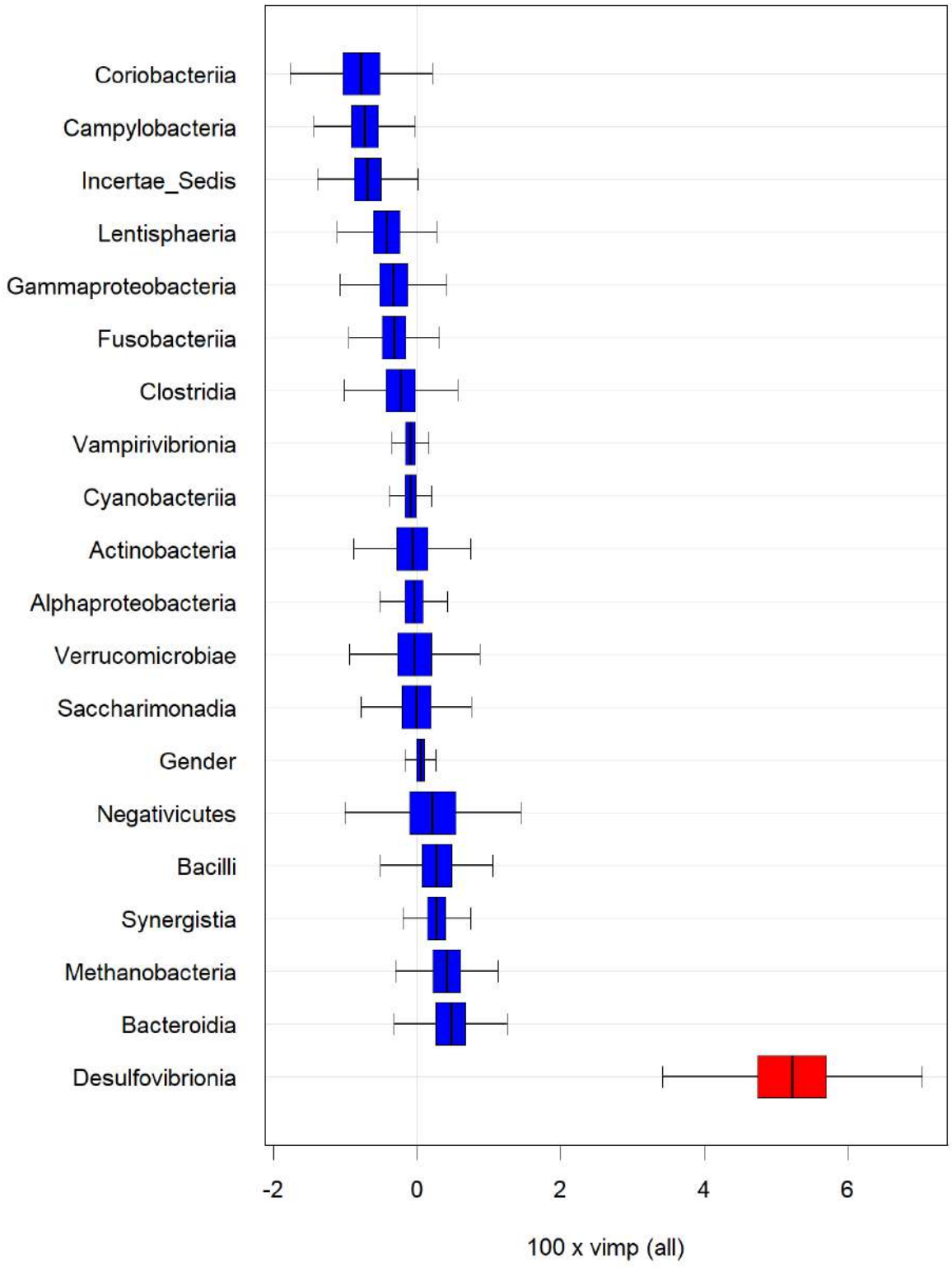


Figure ML2-B-5: The Variable Importance Plot (VIMP) for Class in Stool

Best model is selected for LIME results

Table ML2-B-3: The best setting for model and LIME Results

Setting	mtry	ROC	Sens	Spec	ROCSD	SensSD	SpecSD
ntrees: 500 nodesize: 1	10	0.533333	0	1	0.505525	0	0
ntrees: 1000 nodesize: 1	10	0.533333	0	1	0.505525	0	0
ntrees: 500 nodesize: 5	2	0.466667	0	1	0.447214	0	0
ntrees: 500 nodesize: 5	3	0.466667	0	1	0.447214	0	0
ntrees: 500 nodesize: 5	4	0.466667	0	1	0.447214	0	0
ntrees: 500 nodesize: 5	5	0.466667	0	1	0.447214	0	0
ntrees: 500 nodesize: 5	6	0.466667	0	1	0.447214	0	0
ntrees: 500 nodesize: 5	7	0.466667	0	1	0.447214	0	0
ntrees: 500 nodesize: 5	8	0.466667	0	1	0.447214	0	0
ntrees: 500 nodesize: 5	9	0.466667	0	1	0.447214	0	0
ntrees: 500 nodesize: 5	10	0.466667	0	0.933333	0.447214	0	0.149071
ntrees: 1000 nodesize: 5	2	0.466667	0	1	0.447214	0	0
ntrees: 1000 nodesize: 5	3	0.466667	0	1	0.447214	0	0
ntrees: 1000 nodesize: 5	4	0.466667	0	1	0.447214	0	0
ntrees: 1000 nodesize: 5	5	0.466667	0	1	0.447214	0	0
ntrees: 1000 nodesize: 5	6	0.466667	0	1	0.447214	0	0
ntrees: 1000 nodesize: 5	7	0.466667	0	1	0.447214	0	0
ntrees: 1000 nodesize: 5	8	0.466667	0	1	0.447214	0	0
ntrees: 1000 nodesize: 5	9	0.466667	0	1	0.447214	0	0
ntrees: 1000 nodesize: 5	10	0.466667	0	1	0.447214	0	0

Table ML2-B-4: The LIME Result for Training Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
1	Adenoma	0.708	0.221	0.231	0.362	Verrucomicrobiae	0.331	0.094	4.90e-02 < Verrucomicrobiae
1	Adenoma	0.708	0.221	0.231	0.362	Negativicutes	0.003	0.062	Negativicutes <= 0.00345
1	Adenoma	0.708	0.221	0.231	0.362	Lentisphaeria	0.000	-0.027	Lentisphaeria <= 0.000127
1	Adenoma	0.708	0.221	0.231	0.362	Actinobacteria	0.000	0.024	Actinobacteria <= 0.00105
1	Adenoma	0.708	0.221	0.231	0.362	Desulfovibrionia	0.001	-0.022	0.000637 < Desulfovibrionia <= 0.002729
2	Adenoma	0.636	0.106	0.225	0.338	Coriobacteriia	0.001	0.055	Coriobacteriia <= 0.0016
2	Adenoma	0.636	0.106	0.225	0.338	Negativicutes	0.010	-0.034	0.01011 < Negativicutes <= 0.02789
2	Adenoma	0.636	0.106	0.225	0.338	Incertae_Sedis	0.001	0.032	3.79e-04 < Incertae_Sedis
2	Adenoma	0.636	0.106	0.225	0.338	Lentisphaeria	0.001	0.030	0.000127 < Lentisphaeria
2	Adenoma	0.636	0.106	0.225	0.338	Methanobacteria	0.056	0.029	0.0152 < Methanobacteria
3	Adenoma	0.776	0.148	0.259	0.285	Negativicutes	0.001	0.065	Negativicutes <= 0.00345
3	Adenoma	0.776	0.148	0.259	0.285	Desulfovibrionia	0.000	0.051	Desulfovibrionia <= 0.000170
3	Adenoma	0.776	0.148	0.259	0.285	Verrucomicrobiae	0.000	-0.043	Verrucomicrobiae <= 4.71e-05
3	Adenoma	0.776	0.148	0.259	0.285	Coriobacteriia	0.203	-0.025	0.0156 < Coriobacteriia
3	Adenoma	0.776	0.148	0.259	0.285	Lentisphaeria	0.000	-0.022	Lentisphaeria <= 0.000127

6	Adenoma	0.802	0.258	0.184	0.456	Verrucomicrobiae	0.049	0.093	4.90e-02 < Verrucomicrobiae
6	Adenoma	0.802	0.258	0.184	0.456	Desulfovibrionia	0.000	0.063	Desulfovibrionia <= 0.000170
6	Adenoma	0.802	0.258	0.184	0.456	Negativicutes	0.001	0.058	Negativicutes <= 0.00345
6	Adenoma	0.802	0.258	0.184	0.456	Incertae_Sedis	0.000	0.031	3.79e-04 < Incertae_Sedis
6	Adenoma	0.802	0.258	0.184	0.456	Coriobacteriia	0.003	0.026	0.0016 < Coriobacteriia <= 0.0110
9	Adenoma	0.784	0.137	0.245	0.334	Desulfovibrionia	0.000	0.062	Desulfovibrionia <= 0.000170
9	Adenoma	0.784	0.137	0.245	0.334	Coriobacteriia	0.000	0.049	Coriobacteriia <= 0.0016
9	Adenoma	0.784	0.137	0.245	0.334	Verrucomicrobiae	0.000	-0.039	Verrucomicrobiae <= 4.71e-05
9	Adenoma	0.784	0.137	0.245	0.334	Clostridia	0.180	0.037	Clostridia <= 0.246
9	Adenoma	0.784	0.137	0.245	0.334	Lentisphaeria	0.000	-0.020	Lentisphaeria <= 0.000127
11	Adenocarcinoma	0.844	0.099	0.763	0.700	Coriobacteriia	0.001	-0.057	Coriobacteriia <= 0.0016
11	Adenocarcinoma	0.844	0.099	0.763	0.700	Lentisphaeria	0.001	-0.031	0.000127 < Lentisphaeria
11	Adenocarcinoma	0.844	0.099	0.763	0.700	Clostridia	0.106	-0.032	Clostridia <= 0.246
11	Adenocarcinoma	0.844	0.099	0.763	0.700	Verrucomicrobiae	0.032	0.031	1.36e-02 < Verrucomicrobiae <= 4.90e-02
11	Adenocarcinoma	0.844	0.099	0.763	0.700	Desulfovibrionia	0.005	0.026	0.002729 < Desulfovibrionia
17	Adenocarcinoma	0.898	0.120	0.721	0.726	Coriobacteriia	0.000	-0.057	Coriobacteriia <= 0.0016
17	Adenocarcinoma	0.898	0.120	0.721	0.726	Verrucomicrobiae	0.000	0.042	Verrucomicrobiae <= 4.71e-05
17	Adenocarcinoma	0.898	0.120	0.721	0.726	Clostridia	0.237	-0.034	Clostridia <= 0.246
17	Adenocarcinoma	0.898	0.120	0.721	0.726	Lentisphaeria	0.000	0.032	Lentisphaeria <= 0.000127
17	Adenocarcinoma	0.898	0.120	0.721	0.726	Methanobacteria	0.000	0.023	Methanobacteria <= 0.0028
18	Adenocarcinoma	0.838	0.134	0.779	0.631	Desulfovibrionia	0.000	-0.063	Desulfovibrionia <= 0.000170
18	Adenocarcinoma	0.838	0.134	0.779	0.631	Coriobacteriia	0.001	-0.057	Coriobacteriia <= 0.0016
18	Adenocarcinoma	0.838	0.134	0.779	0.631	Lentisphaeria	0.001	-0.028	0.000127 < Lentisphaeria
18	Adenocarcinoma	0.838	0.134	0.779	0.631	Methanobacteria	0.000	0.021	Methanobacteria <= 0.0028
18	Adenocarcinoma	0.838	0.134	0.779	0.631	Actinobacteria	0.000	-0.021	Actinobacteria <= 0.00105
19	Adenocarcinoma	0.932	0.152	0.726	0.758	Negativicutes	0.002	-0.061	Negativicutes <= 0.00345
19	Adenocarcinoma	0.932	0.152	0.726	0.758	Coriobacteriia	0.015	0.063	0.0110 < Coriobacteriia <= 0.0156
19	Adenocarcinoma	0.932	0.152	0.726	0.758	Methanobacteria	0.034	-0.033	0.0152 < Methanobacteria
19	Adenocarcinoma	0.932	0.152	0.726	0.758	Verrucomicrobiae	0.013	0.034	4.71e-05 < Verrucomicrobiae <= 1.36e-02
19	Adenocarcinoma	0.932	0.152	0.726	0.758	Lentisphaeria	0.000	0.029	Lentisphaeria <= 0.000127
20	Adenocarcinoma	0.870	0.112	0.735	0.738	Coriobacteriia	0.001	-0.055	Coriobacteriia <= 0.0016
20	Adenocarcinoma	0.870	0.112	0.735	0.738	Verrucomicrobiae	0.000	0.040	Verrucomicrobiae <= 4.71e-05
20	Adenocarcinoma	0.870	0.112	0.735	0.738	Clostridia	0.181	-0.036	Clostridia <= 0.246
20	Adenocarcinoma	0.870	0.112	0.735	0.738	Negativicutes	0.017	0.028	0.01011 < Negativicutes <= 0.02789
20	Adenocarcinoma	0.870	0.112	0.735	0.738	Lentisphaeria	0.000	0.026	Lentisphaeria <= 0.000127
21	Adenocarcinoma	0.954	0.185	0.721	0.757	Verrucomicrobiae	0.088	-0.089	4.90e-02 < Verrucomicrobiae
21	Adenocarcinoma	0.954	0.185	0.721	0.757	Coriobacteriia	0.012	0.060	0.0110 < Coriobacteriia <= 0.0156
21	Adenocarcinoma	0.954	0.185	0.721	0.757	Lentisphaeria	0.000	0.025	Lentisphaeria <= 0.000127
21	Adenocarcinoma	0.954	0.185	0.721	0.757	Desulfovibrionia	0.002	0.022	0.000637 < Desulfovibrionia <= 0.002729
21	Adenocarcinoma	0.954	0.185	0.721	0.757	Methanobacteria	0.000	0.018	Methanobacteria <= 0.0028
22	Adenocarcinoma	0.936	0.112	0.746	0.689	Desulfovibrionia	0.000	-0.058	Desulfovibrionia <= 0.000170
22	Adenocarcinoma	0.936	0.112	0.746	0.689	Verrucomicrobiae	0.000	0.037	Verrucomicrobiae <= 4.71e-05
22	Adenocarcinoma	0.936	0.112	0.746	0.689	Clostridia	0.215	-0.033	Clostridia <= 0.246

22	Adenocarcinoma	0.936	0.112	0.746	0.689	Coriobacteriia	0.010	-0.030	0.0016 < Coriobacteriia <= 0.0110
22	Adenocarcinoma	0.936	0.112	0.746	0.689	Lentisphaeria	0.000	0.026	Lentisphaeria <= 0.000127
23	Adenocarcinoma	0.846	0.216	0.766	0.643	Verrucomicrobiae	0.057	-0.087	4.90e-02 < Verrucomicrobiae
23	Adenocarcinoma	0.846	0.216	0.766	0.643	Negativicutes	0.002	-0.061	Negativicutes <= 0.00345
23	Adenocarcinoma	0.846	0.216	0.766	0.643	Methanobacteria	0.062	-0.032	0.0152 < Methanobacteria
23	Adenocarcinoma	0.846	0.216	0.766	0.643	Lentisphaeria	0.000	0.030	Lentisphaeria <= 0.000127
23	Adenocarcinoma	0.846	0.216	0.766	0.643	Clostridia	0.469	0.027	0.468 < Clostridia
25	Adenocarcinoma	0.962	0.069	0.711	0.807	Negativicutes	0.011	0.040	0.01011 < Negativicutes <= 0.02789
25	Adenocarcinoma	0.962	0.069	0.711	0.807	Verrucomicrobiae	0.003	0.031	4.71e-05 < Verrucomicrobiae <= 1.36e-02
25	Adenocarcinoma	0.962	0.069	0.711	0.807	Lentisphaeria	0.000	0.029	Lentisphaeria <= 0.000127
25	Adenocarcinoma	0.962	0.069	0.711	0.807	Methanobacteria	0.029	-0.027	0.0152 < Methanobacteria
25	Adenocarcinoma	0.962	0.069	0.711	0.807	Clostridia	0.404	0.024	0.389 < Clostridia <= 0.468
26	Adenocarcinoma	0.952	0.205	0.765	0.697	Verrucomicrobiae	0.086	-0.089	4.90e-02 < Verrucomicrobiae
26	Adenocarcinoma	0.952	0.205	0.765	0.697	Coriobacteriia	0.012	0.050	0.0110 < Coriobacteriia <= 0.0156
26	Adenocarcinoma	0.952	0.205	0.765	0.697	Incertae_Sedis	0.001	-0.034	3.79e-04 < Incertae_Sedis
26	Adenocarcinoma	0.952	0.205	0.765	0.697	Clostridia	0.468	0.031	0.468 < Clostridia
26	Adenocarcinoma	0.952	0.205	0.765	0.697	Methanobacteria	0.022	-0.026	0.0152 < Methanobacteria
29	Adenocarcinoma	0.924	0.189	0.791	0.622	Verrucomicrobiae	0.071	-0.088	4.90e-02 < Verrucomicrobiae
29	Adenocarcinoma	0.924	0.189	0.791	0.622	Desulfovibrionia	0.000	-0.057	Desulfovibrionia <= 0.000170
29	Adenocarcinoma	0.924	0.189	0.791	0.622	Coriobacteriia	0.010	-0.026	0.0016 < Coriobacteriia <= 0.0110
29	Adenocarcinoma	0.924	0.189	0.791	0.622	Negativicutes	0.051	0.023	0.02789 < Negativicutes
29	Adenocarcinoma	0.924	0.189	0.791	0.622	Clostridia	0.359	-0.021	0.246 < Clostridia <= 0.389
30	Adenocarcinoma	0.894	0.145	0.737	0.774	Negativicutes	0.001	-0.058	Negativicutes <= 0.00345
30	Adenocarcinoma	0.894	0.145	0.737	0.774	Coriobacteriia	0.012	0.058	0.0110 < Coriobacteriia <= 0.0156
30	Adenocarcinoma	0.894	0.145	0.737	0.774	Clostridia	0.637	0.033	0.468 < Clostridia
30	Adenocarcinoma	0.894	0.145	0.737	0.774	Verrucomicrobiae	0.011	0.035	4.71e-05 < Verrucomicrobiae <= 1.36e-02
30	Adenocarcinoma	0.894	0.145	0.737	0.774	Incertae_Sedis	0.000	-0.032	3.79e-04 < Incertae_Sedis
31	Adenocarcinoma	0.950	0.064	0.709	0.786	Negativicutes	0.027	0.037	0.01011 < Negativicutes <= 0.02789
31	Adenocarcinoma	0.950	0.064	0.709	0.786	Clostridia	0.205	-0.033	Clostridia <= 0.246
31	Adenocarcinoma	0.950	0.064	0.709	0.786	Verrucomicrobiae	0.000	0.028	4.71e-05 < Verrucomicrobiae <= 1.36e-02
31	Adenocarcinoma	0.950	0.064	0.709	0.786	Lentisphaeria	0.000	0.026	Lentisphaeria <= 0.000127
31	Adenocarcinoma	0.950	0.064	0.709	0.786	Methanobacteria	0.000	0.019	Methanobacteria <= 0.0028
35	Adenocarcinoma	0.896	0.100	0.732	0.789	Coriobacteriia	0.012	0.057	0.0110 < Coriobacteriia <= 0.0156
35	Adenocarcinoma	0.896	0.100	0.732	0.789	Lentisphaeria	0.001	-0.033	0.000127 < Lentisphaeria
35	Adenocarcinoma	0.896	0.100	0.732	0.789	Clostridia	0.662	0.032	0.468 < Clostridia
35	Adenocarcinoma	0.896	0.100	0.732	0.789	Incertae_Sedis	0.000	-0.029	3.79e-04 < Incertae_Sedis
35	Adenocarcinoma	0.896	0.100	0.732	0.789	Negativicutes	0.012	0.031	0.01011 < Negativicutes <= 0.02789
41	Adenocarcinoma	0.962	0.048	0.695	0.820	Verrucomicrobiae	0.010	0.031	4.71e-05 < Verrucomicrobiae <= 1.36e-02
41	Adenocarcinoma	0.962	0.048	0.695	0.820	Lentisphaeria	0.000	0.029	Lentisphaeria <= 0.000127
41	Adenocarcinoma	0.962	0.048	0.695	0.820	Desulfovibrionia	0.000	0.026	0.000170 < Desulfovibrionia <= 0.000637
41	Adenocarcinoma	0.962	0.048	0.695	0.820	Incertae_Sedis	0.000	0.018	Incertae_Sedis <= 1.44e-05
41	Adenocarcinoma	0.962	0.048	0.695	0.820	Clostridia	0.410	0.021	0.389 < Clostridia <= 0.468
42	Adenocarcinoma	0.910	0.070	0.702	0.846	Verrucomicrobiae	0.000	0.041	Verrucomicrobiae <= 4.71e-05

42	Adenocarcinoma	0.910	0.070	0.702	0.846	Negativicutes	0.047	0.029	0.02789 < Negativicutes
42	Adenocarcinoma	0.910	0.070	0.702	0.846	Clostridia	0.520	0.027	0.468 < Clostridia
42	Adenocarcinoma	0.910	0.070	0.702	0.846	Desulfovibrionia	0.004	0.026	0.002729 < Desulfovibrionia
42	Adenocarcinoma	0.910	0.070	0.702	0.846	Methanobacteria	0.000	0.020	Methanobacteria <= 0.0028
43	Adenocarcinoma	0.964	0.052	0.739	0.758	Incertae_Sedis	0.005	-0.035	3.79e-04 < Incertae_Sedis
43	Adenocarcinoma	0.964	0.052	0.739	0.758	Desulfovibrionia	0.001	0.027	0.000170 < Desulfovibrionia <= 0.000637
43	Adenocarcinoma	0.964	0.052	0.739	0.758	Coriobacteriia	0.017	0.024	0.0156 < Coriobacteriia
43	Adenocarcinoma	0.964	0.052	0.739	0.758	Actinobacteria	0.149	0.023	0.01023 < Actinobacteria
43	Adenocarcinoma	0.964	0.052	0.739	0.758	Lentisphaeria	0.001	-0.020	0.000127 < Lentisphaeria

Table ML-B-5: The LIME Result for Test Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
4	Adenoma	0.512	0.107	0.278	0.224	Desulfovibrionia	0.000	0.063	Desulfovibrionia <= 0.000170
4	Adenoma	0.512	0.107	0.278	0.224	Negativicutes	0.018	-0.036	0.01011 < Negativicutes <= 0.02789
4	Adenoma	0.512	0.107	0.278	0.224	Clostridia	0.643	-0.030	0.468 < Clostridia
4	Adenoma	0.512	0.107	0.278	0.224	Lentisphaeria	0.000	-0.026	Lentisphaeria <= 0.000127
4	Adenoma	0.512	0.107	0.278	0.224	Verrucomicrobiae	0.000	-0.025	4.71e-05 < Verrucomicrobiae <= 1.36e-02
5	Adenocarcinoma	0.722	0.054	0.735	0.754	Verrucomicrobiae	0.000	0.036	Verrucomicrobiae <= 4.71e-05
5	Adenocarcinoma	0.722	0.054	0.735	0.754	Clostridia	0.178	-0.033	Clostridia <= 0.246
5	Adenocarcinoma	0.722	0.054	0.735	0.754	Saccharimonadia	0.003	-0.022	0.000824 < Saccharimonadia
5	Adenocarcinoma	0.722	0.054	0.735	0.754	Desulfovibrionia	0.001	0.021	0.000170 < Desulfovibrionia <= 0.000637
5	Adenocarcinoma	0.722	0.054	0.735	0.754	Incertae_Sedis	0.000	0.017	Incertae_Sedis <= 1.44e-05
7	Adenocarcinoma	0.810	0.057	0.725	0.757	Verrucomicrobiae	0.010	0.038	4.71e-05 < Verrucomicrobiae <= 1.36e-02
7	Adenocarcinoma	0.810	0.057	0.725	0.757	Clostridia	0.157	-0.033	Clostridia <= 0.246
7	Adenocarcinoma	0.810	0.057	0.725	0.757	Lentisphaeria	0.000	0.025	Lentisphaeria <= 0.000127
7	Adenocarcinoma	0.810	0.057	0.725	0.757	Coriobacteriia	0.036	0.023	0.0156 < Coriobacteriia
7	Adenocarcinoma	0.810	0.057	0.725	0.757	Saccharimonadia	0.005	-0.021	0.000824 < Saccharimonadia
8	Adenocarcinoma	0.936	0.069	0.690	0.831	Verrucomicrobiae	0.000	0.040	Verrucomicrobiae <= 4.71e-05
8	Adenocarcinoma	0.936	0.069	0.690	0.831	Coriobacteriia	0.038	0.032	0.0156 < Coriobacteriia
8	Adenocarcinoma	0.936	0.069	0.690	0.831	Lentisphaeria	0.000	0.024	Lentisphaeria <= 0.000127
8	Adenocarcinoma	0.936	0.069	0.690	0.831	Methanobacteria	0.000	0.021	Methanobacteria <= 0.0028
8	Adenocarcinoma	0.936	0.069	0.690	0.831	Clostridia	0.489	0.023	0.468 < Clostridia
10	Adenocarcinoma	0.824	0.057	0.702	0.840	Verrucomicrobiae	0.002	0.037	4.71e-05 < Verrucomicrobiae <= 1.36e-02
10	Adenocarcinoma	0.824	0.057	0.702	0.840	Negativicutes	0.011	0.030	0.01011 < Negativicutes <= 0.02789
10	Adenocarcinoma	0.824	0.057	0.702	0.840	Clostridia	0.565	0.028	0.468 < Clostridia
10	Adenocarcinoma	0.824	0.057	0.702	0.840	Desulfovibrionia	0.000	0.022	0.000170 < Desulfovibrionia <= 0.000637
10	Adenocarcinoma	0.824	0.057	0.702	0.840	Lentisphaeria	0.000	0.021	Lentisphaeria <= 0.000127
12	Adenocarcinoma	0.722	0.040	0.704	0.817	Verrucomicrobiae	0.013	0.029	4.71e-05 < Verrucomicrobiae <= 1.36e-02
12	Adenocarcinoma	0.722	0.040	0.704	0.817	Negativicutes	0.051	0.024	0.02789 < Negativicutes
12	Adenocarcinoma	0.722	0.040	0.704	0.817	Incertae_Sedis	0.000	0.020	Incertae_Sedis <= 1.44e-05
12	Adenocarcinoma	0.722	0.040	0.704	0.817	Methanobacteria	0.000	0.019	Methanobacteria <= 0.0028
12	Adenocarcinoma	0.722	0.040	0.704	0.817	Desulfovibrionia	0.001	0.021	0.000637 < Desulfovibrionia <= 0.002729

13	Adenocarcinoma	0.774	0.098	0.747	0.743	Coriobacteriia	0.001	-0.054	Coriobacteriia <= 0.0016
13	Adenocarcinoma	0.774	0.098	0.747	0.743	Incertae_Sedis	0.000	-0.037	3.79e-04 < Incertae_Sedis
13	Adenocarcinoma	0.774	0.098	0.747	0.743	Clostridia	0.487	0.035	0.468 < Clostridia
13	Adenocarcinoma	0.774	0.098	0.747	0.743	Verrucomicrobiae	0.008	0.033	4.71e-05 < Verrucomicrobiae <= 1.36e-02
13	Adenocarcinoma	0.774	0.098	0.747	0.743	Desulfovibrionia	0.001	0.019	0.000637 < Desulfovibrionia <= 0.002729
14	Adenocarcinoma	0.754	0.099	0.729	0.716	Negativicutes	0.002	-0.065	Negativicutes <= 0.00345
14	Adenocarcinoma	0.754	0.099	0.729	0.716	Lentisphaeria	0.000	0.027	Lentisphaeria <= 0.000127
14	Adenocarcinoma	0.754	0.099	0.729	0.716	Incertae_Sedis	0.000	0.022	Incertae_Sedis <= 1.44e-05
14	Adenocarcinoma	0.754	0.099	0.729	0.716	Verrucomicrobiae	0.029	0.026	1.36e-02 < Verrucomicrobiae <= 4.90e-02
14	Adenocarcinoma	0.754	0.099	0.729	0.716	Coriobacteriia	0.008	-0.022	0.0016 < Coriobacteriia <= 0.0110
15	Adenocarcinoma	0.900	0.184	0.770	0.684	Verrucomicrobiae	0.144	-0.097	4.90e-02 < Verrucomicrobiae
15	Adenocarcinoma	0.900	0.184	0.770	0.684	Negativicutes	0.024	0.038	0.01011 < Negativicutes <= 0.02789
15	Adenocarcinoma	0.900	0.184	0.770	0.684	Lentisphaeria	0.000	-0.025	0.000127 < Lentisphaeria
15	Adenocarcinoma	0.900	0.184	0.770	0.684	Coriobacteriia	0.008	-0.021	0.0016 < Coriobacteriia <= 0.0110
15	Adenocarcinoma	0.900	0.184	0.770	0.684	Actinobacteria	0.014	0.017	0.01023 < Actinobacteria
16	Adenocarcinoma	0.716	0.076	0.741	0.712	Clostridia	0.230	-0.036	Clostridia <= 0.246
16	Adenocarcinoma	0.716	0.076	0.741	0.712	Verrucomicrobiae	0.000	0.036	Verrucomicrobiae <= 4.71e-05
16	Adenocarcinoma	0.716	0.076	0.741	0.712	Methanobacteria	0.046	-0.029	0.0152 < Methanobacteria
16	Adenocarcinoma	0.716	0.076	0.741	0.712	Incertae_Sedis	0.001	-0.027	3.79e-04 < Incertae_Sedis
16	Adenocarcinoma	0.716	0.076	0.741	0.712	Lentisphaeria	0.000	0.027	Lentisphaeria <= 0.000127
24	Adenocarcinoma	0.954	0.081	0.705	0.801	Verrucomicrobiae	0.000	0.045	Verrucomicrobiae <= 4.71e-05
24	Adenocarcinoma	0.954	0.081	0.705	0.801	Lentisphaeria	0.000	0.032	Lentisphaeria <= 0.000127
24	Adenocarcinoma	0.954	0.081	0.705	0.801	Incertae_Sedis	0.001	-0.031	3.79e-04 < Incertae_Sedis
24	Adenocarcinoma	0.954	0.081	0.705	0.801	Negativicutes	0.028	0.026	0.02789 < Negativicutes
24	Adenocarcinoma	0.954	0.081	0.705	0.801	Clostridia	0.496	0.024	0.468 < Clostridia
27	Adenocarcinoma	0.910	0.056	0.691	0.823	Verrucomicrobiae	0.000	0.036	Verrucomicrobiae <= 4.71e-05
27	Adenocarcinoma	0.910	0.056	0.691	0.823	Coriobacteriia	0.100	0.028	0.0156 < Coriobacteriia
27	Adenocarcinoma	0.910	0.056	0.691	0.823	Lentisphaeria	0.000	0.029	Lentisphaeria <= 0.000127
27	Adenocarcinoma	0.910	0.056	0.691	0.823	Desulfovibrionia	0.001	0.023	0.000170 < Desulfovibrionia <= 0.000637
27	Adenocarcinoma	0.910	0.056	0.691	0.823	Incertae_Sedis	0.000	0.016	Incertae_Sedis <= 1.44e-05
28	Adenocarcinoma	0.926	0.058	0.732	0.769	Verrucomicrobiae	0.000	0.039	Verrucomicrobiae <= 4.71e-05
28	Adenocarcinoma	0.926	0.058	0.732	0.769	Lentisphaeria	0.000	-0.027	0.000127 < Lentisphaeria
28	Adenocarcinoma	0.926	0.058	0.732	0.769	Coriobacteriia	0.018	0.026	0.0156 < Coriobacteriia
28	Adenocarcinoma	0.926	0.058	0.732	0.769	Actinobacteria	0.001	-0.022	Actinobacteria <= 0.00105
28	Adenocarcinoma	0.926	0.058	0.732	0.769	Negativicutes	0.087	0.022	0.02789 < Negativicutes
32	Adenocarcinoma	0.682	0.169	0.792	0.626	Negativicutes	0.001	-0.062	Negativicutes <= 0.00345
32	Adenocarcinoma	0.682	0.169	0.792	0.626	Desulfovibrionia	0.000	-0.053	Desulfovibrionia <= 0.000170
32	Adenocarcinoma	0.682	0.169	0.792	0.626	Coriobacteriia	0.001	-0.051	Coriobacteriia <= 0.0016
32	Adenocarcinoma	0.682	0.169	0.792	0.626	Clostridia	0.066	-0.028	Clostridia <= 0.246
32	Adenocarcinoma	0.682	0.169	0.792	0.626	Verrucomicrobiae	0.010	0.027	4.71e-05 < Verrucomicrobiae <= 1.36e-02
33	Adenocarcinoma	0.856	0.093	0.690	0.866	Coriobacteriia	0.013	0.051	0.0110 < Coriobacteriia <= 0.0156
33	Adenocarcinoma	0.856	0.093	0.690	0.866	Verrucomicrobiae	0.001	0.040	4.71e-05 < Verrucomicrobiae <= 1.36e-02
33	Adenocarcinoma	0.856	0.093	0.690	0.866	Negativicutes	0.010	0.033	0.01011 < Negativicutes <= 0.02789

33	Adenocarcinoma	0.856	0.093	0.690	0.866	Lentisphaeria	0.000	0.027	Lentisphaeria <= 0.000127
33	Adenocarcinoma	0.856	0.093	0.690	0.866	Desulfovibrionia	0.003	0.025	0.002729 < Desulfovibrionia
34	Adenocarcinoma	0.908	0.052	0.711	0.794	Verrucomicrobiae	0.030	0.030	1.36e-02 < Verrucomicrobiae <= 4.90e-02
34	Adenocarcinoma	0.908	0.052	0.711	0.794	Lentisphaeria	0.000	0.028	Lentisphaeria <= 0.000127
34	Adenocarcinoma	0.908	0.052	0.711	0.794	Negativicutes	0.077	0.026	0.02789 < Negativicutes
34	Adenocarcinoma	0.908	0.052	0.711	0.794	Clostridia	0.240	-0.024	Clostridia <= 0.246
34	Adenocarcinoma	0.908	0.052	0.711	0.794	Coriobacteriia	0.036	0.023	0.0156 < Coriobacteriia
36	Adenoma	0.614	0.095	0.239	0.358	Negativicutes	0.001	0.063	Negativicutes <= 0.00345
36	Adenoma	0.614	0.095	0.239	0.358	Methanobacteria	0.051	0.030	0.0152 < Methanobacteria
36	Adenoma	0.614	0.095	0.239	0.358	Lentisphaeria	0.000	-0.024	Lentisphaeria <= 0.000127
36	Adenoma	0.614	0.095	0.239	0.358	Actinobacteria	0.000	0.025	Actinobacteria <= 0.00105
36	Adenoma	0.614	0.095	0.239	0.358	Coriobacteriia	0.002	0.025	0.0016 < Coriobacteriia <= 0.0110
37	Adenocarcinoma	0.884	0.098	0.722	0.771	Desulfovibrionia	0.000	-0.061	Desulfovibrionia <= 0.000170
37	Adenocarcinoma	0.884	0.098	0.722	0.771	Verrucomicrobiae	0.001	0.031	4.71e-05 < Verrucomicrobiae <= 1.36e-02
37	Adenocarcinoma	0.884	0.098	0.722	0.771	Negativicutes	0.015	0.029	0.01011 < Negativicutes <= 0.02789
37	Adenocarcinoma	0.884	0.098	0.722	0.771	Lentisphaeria	0.000	0.027	Lentisphaeria <= 0.000127
37	Adenocarcinoma	0.884	0.098	0.722	0.771	Coriobacteriia	0.028	0.023	0.0156 < Coriobacteriia
38	Adenocarcinoma	0.800	0.094	0.721	0.739	Desulfovibrionia	0.000	-0.052	Desulfovibrionia <= 0.000170
38	Adenocarcinoma	0.800	0.094	0.721	0.739	Verrucomicrobiae	0.000	0.042	Verrucomicrobiae <= 4.71e-05
38	Adenocarcinoma	0.800	0.094	0.721	0.739	Lentisphaeria	0.000	0.030	Lentisphaeria <= 0.000127
38	Adenocarcinoma	0.800	0.094	0.721	0.739	Negativicutes	0.149	0.023	0.02789 < Negativicutes
38	Adenocarcinoma	0.800	0.094	0.721	0.739	Coriobacteriia	0.004	-0.024	0.0016 < Coriobacteriia <= 0.0110
39	Adenocarcinoma	0.868	0.040	0.717	0.787	Verrucomicrobiae	0.024	0.028	1.36e-02 < Verrucomicrobiae <= 4.90e-02
39	Adenocarcinoma	0.868	0.040	0.717	0.787	Gammaproteobacteria	0.283	-0.022	0.05347 < Gammaproteobacteria
39	Adenocarcinoma	0.868	0.040	0.717	0.787	Lentisphaeria	0.000	0.023	Lentisphaeria <= 0.000127
39	Adenocarcinoma	0.868	0.040	0.717	0.787	Coriobacteriia	0.101	0.022	0.0156 < Coriobacteriia
39	Adenocarcinoma	0.868	0.040	0.717	0.787	Desulfovibrionia	0.000	0.018	0.000170 < Desulfovibrionia <= 0.000637
40	Adenocarcinoma	0.518	0.209	0.767	0.644	Verrucomicrobiae	0.276	-0.097	4.90e-02 < Verrucomicrobiae
40	Adenocarcinoma	0.518	0.209	0.767	0.644	Coriobacteriia	0.001	-0.047	Coriobacteriia <= 0.0016
40	Adenocarcinoma	0.518	0.209	0.767	0.644	Methanobacteria	0.029	-0.029	0.0152 < Methanobacteria
40	Adenocarcinoma	0.518	0.209	0.767	0.644	Negativicutes	0.050	0.027	0.02789 < Negativicutes
40	Adenocarcinoma	0.518	0.209	0.767	0.644	Lentisphaeria	0.000	0.023	Lentisphaeria <= 0.000127

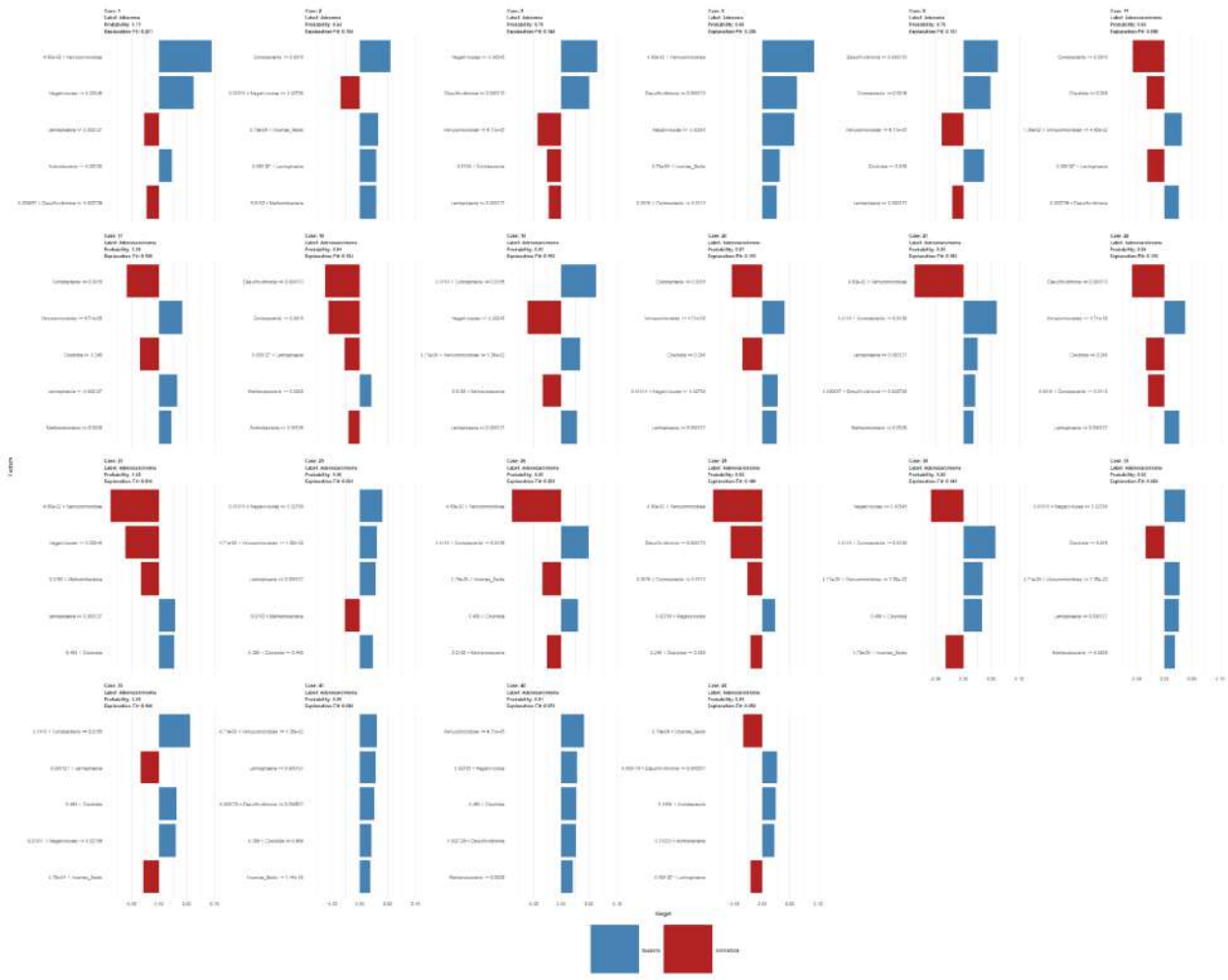


Figure ML2-B-6: The LIME Result in Training dataset for Class in Stool

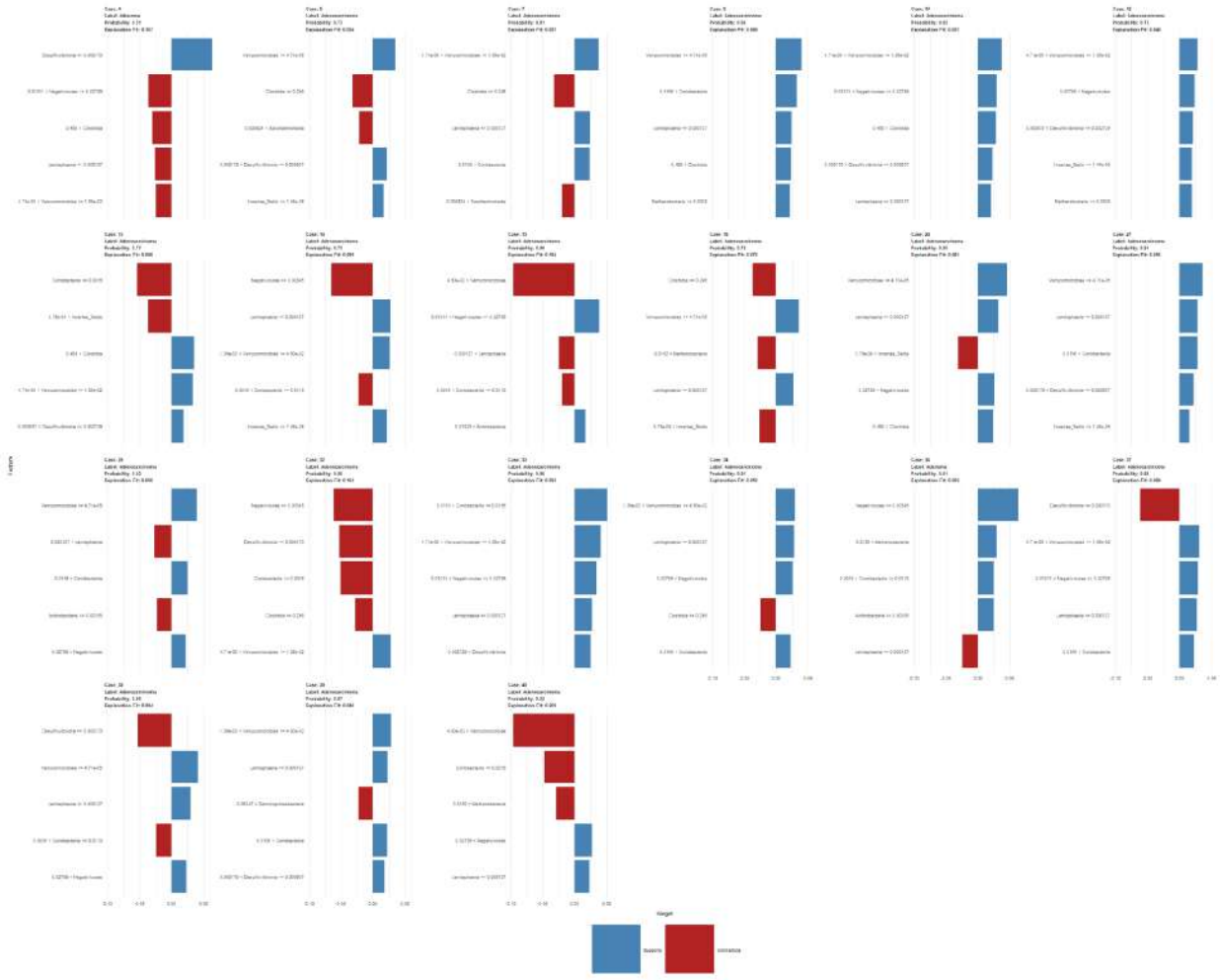


Figure ML2-B-7: The LIME Result in Test dataset for Class in Stool

Table ML2-B-6: The Aggregated SHAP Values

Variable	Training Dataset				Testing Dataset			
	contribution	position	sign	cumulative	contribution	position	sign	cumulative
intercept	0.7648	23	X	0.765	0.7648	23	X	0.7648
Gammaproteobacteria	0.0001	22	1	0.765	0.0014	22	1	0.7662
Coriobacteriia	0.0001	21	1	0.765	-0.0058	20	-1	0.7610
Desulfovibrionia	-0.0002	20	-1	0.765	0.0320	19	1	0.7930
Saccharimonadia	-0.0002	19	-1	0.765	-0.0015	18	-1	0.7914

Tumor	0.0000	18	1	0.765	0.0000	17	1	0.7914
Bacilli	0.0001	17	1	0.765	-0.0016	16	-1	0.7898
Verrucomicrobiae	0.0001	16	1	0.765	-0.0080	10	-1	0.7775
Campylobacteria	-0.0001	15	-1	0.765	0.0005	8	1	0.7766
Synergistia	-0.0001	14	-1	0.765	0.0000	15	-1	0.7898
Methanobacteria	0.0003	13	1	0.765	0.0006	21	1	0.7667
Negativicutes	0.0001	12	1	0.765	0.0008	11	1	0.7855
Incertae_Sedis	0.0000	11	-1	0.765	0.0026	13	1	0.7870
Cyanobacteriia	0.0000	10	-1	0.765	-0.0003	4	-1	0.7751
Lentisphaeria	0.0001	9	1	0.765	-0.0023	12	-1	0.7847
Clostridia	0.000	8	-1	0.765	-0.0054	14	-1	0.7844
Gender	0.000	7	1	0.765	0.0000	5	1	0.7755
Vampirivibrionia	0.000	6	-1	0.765	-0.0003	7	-1	0.7763
Bacteroidia	0.000	5	-1	0.765	-0.0014	9	-1	0.7761
Fusobacteriia	0.000	4	-1	0.765	-0.0008	6	-1	0.7755
Actinobacteria	0.000	3	1	0.765	0.0039	3	1	0.7791
Alphaproteobacteria	0.000	2	-1	0.765	-0.0011	2	-1	0.7780
prediction	0.765	1	X	0.765	0.7780	1	X	0.7780

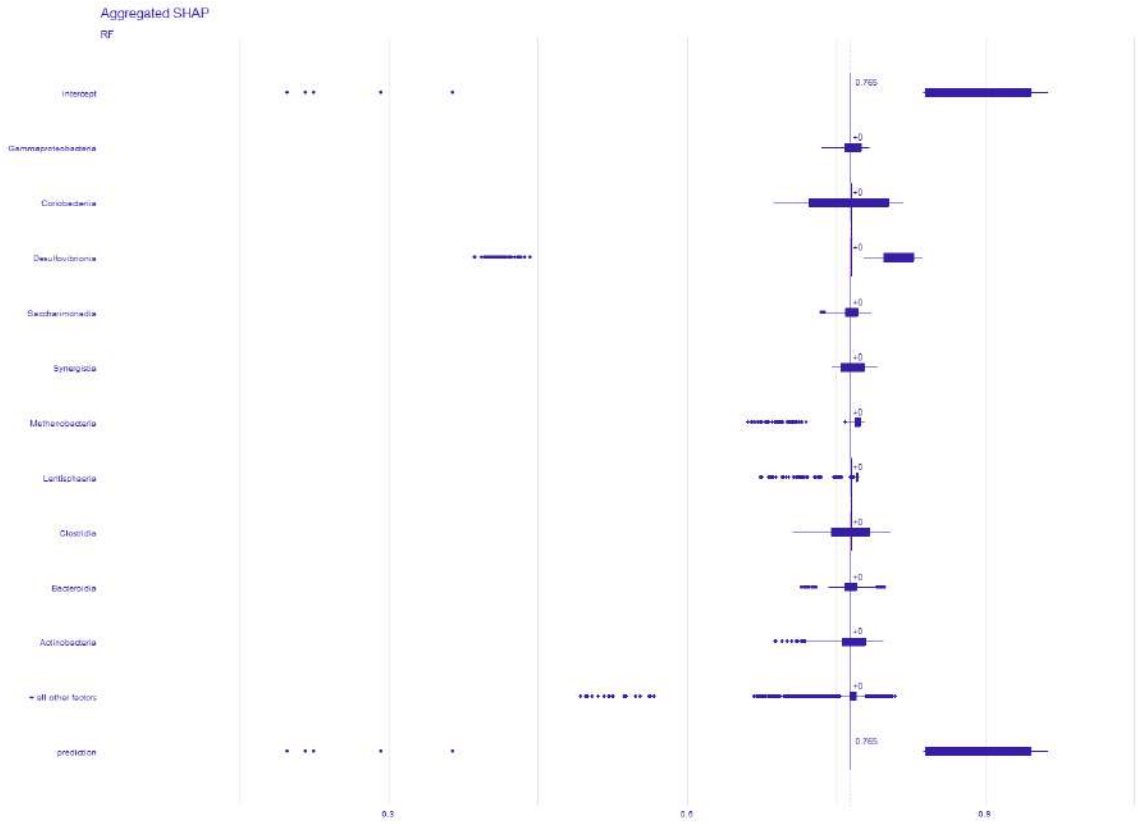


Figure ML2-B-8: The Aggregated SHAP Value in Training dataset for Class in Stool

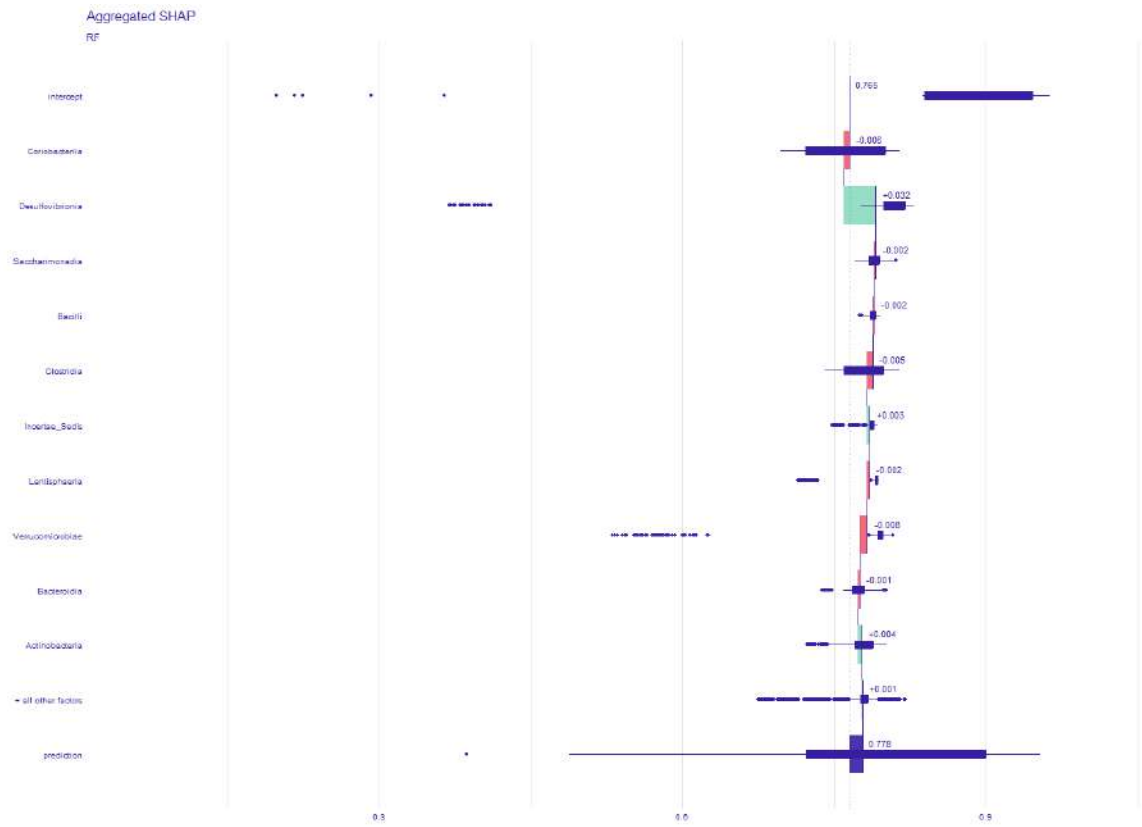


Figure ML2-B-9: The Aggregated SHAP Value in Test dataset for Class in Stool

C- Order

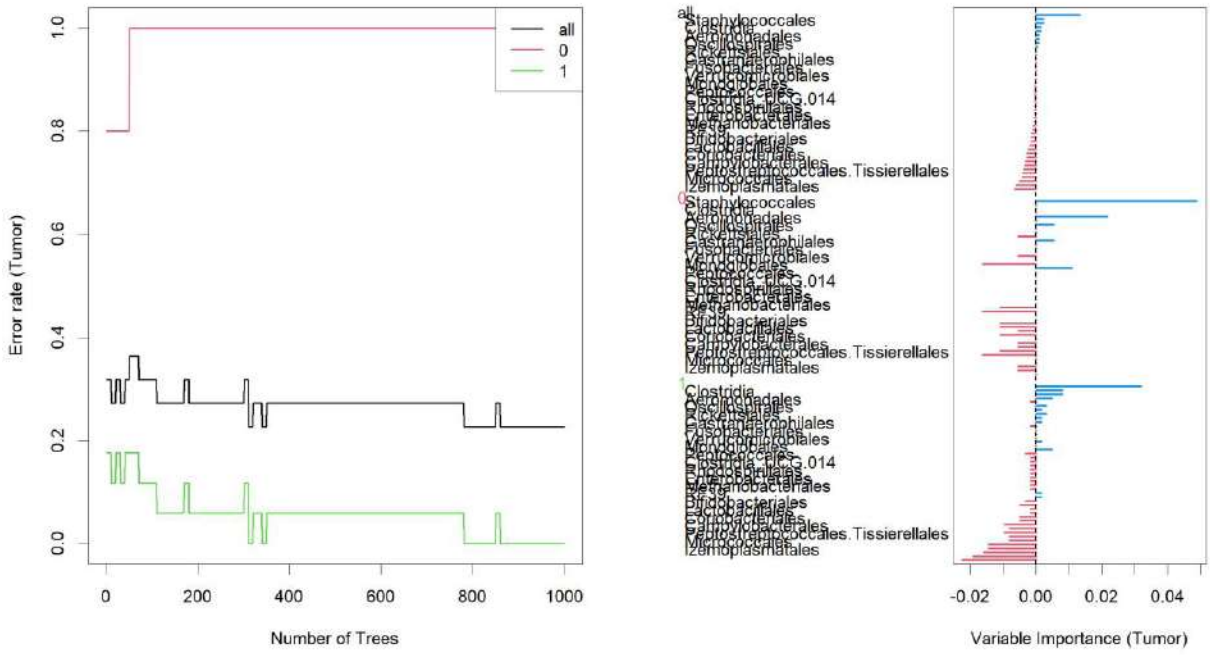
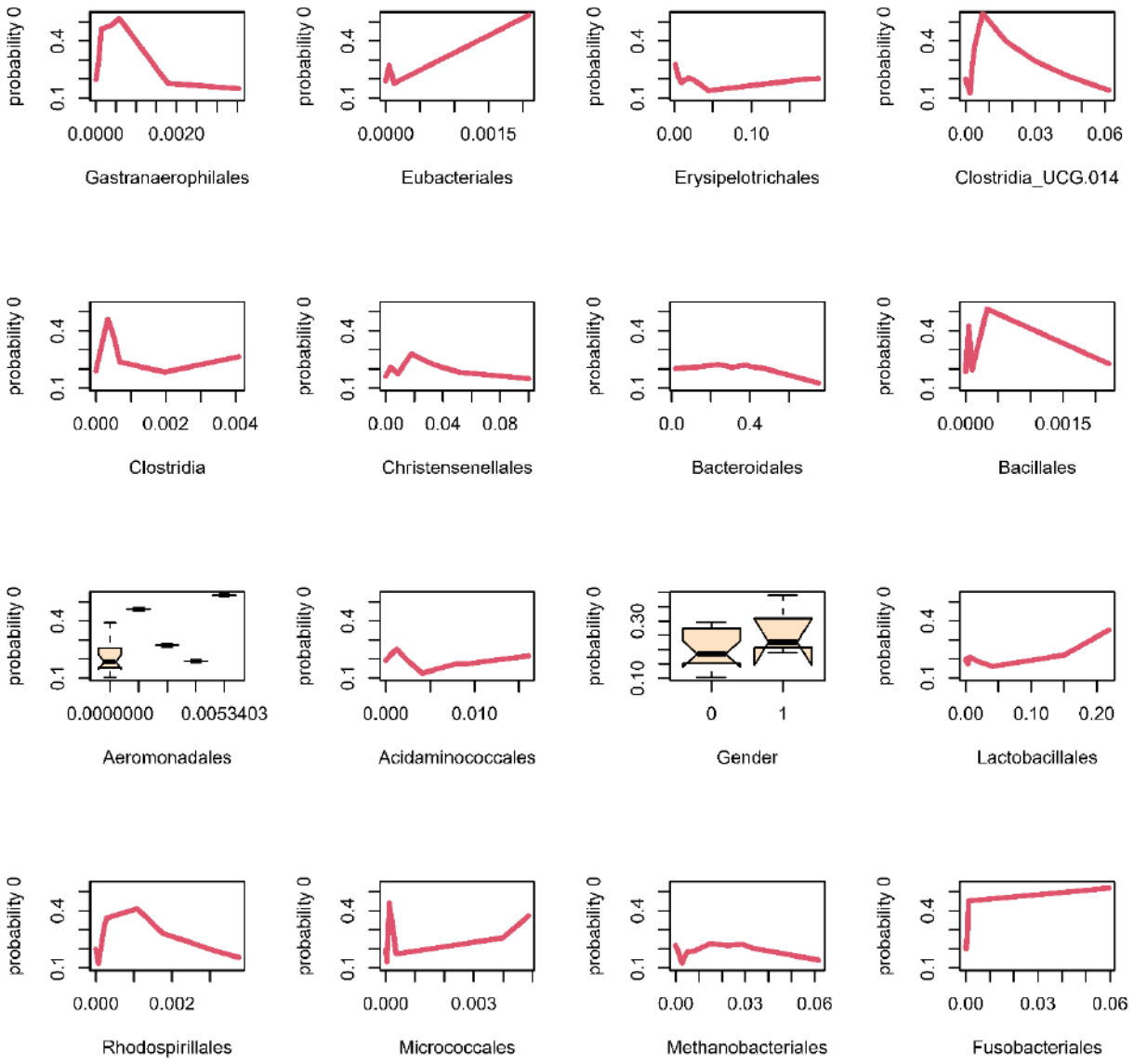
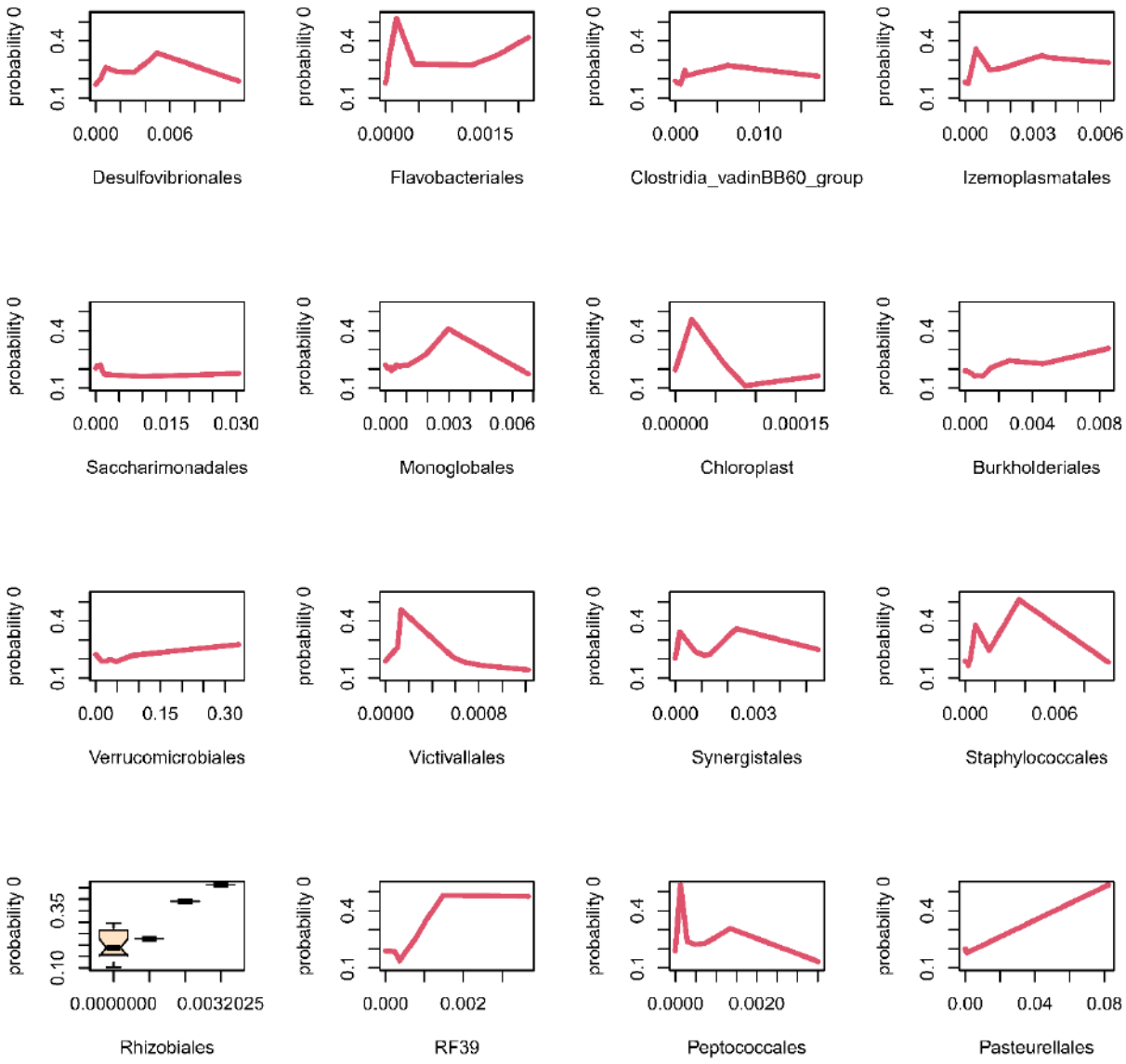


Figure ML2-C-1: The number of trees and Variable Importance Plot for Order in Stool


probability 0
0.1
0.4
0.0000
0.0020
Gastranaerophilales
probability 0
0.1
0.4
0.0000
0.0015
Eubacteriales
probability 0
0.1
0.4
0.00
0.10
Erysipelotrichales
probability 0
0.1
0.4
0.00
0.03
0.06
Clostridia_UCG.014
probability 0
0.1
0.4
0.000
0.002
0.004
Clostridia
probability 0
0.1
0.4
0.00
0.04
0.08
Christensenellales
probability 0
0.1
0.4
0.0
0.4
Bacteroidales
probability 0
0.1
0.4
0.0000
0.0015
Bacillales
probability 0
0.1
0.4
0.0000000
0.0053403
Aeromonadales
probability 0
0.1
0.4
0.000
0.010
Acidaminococcales
probability 0
0.10
0.30
0
1
Gender
probability 0
0.1
0.4
0.00
0.10
0.20
Lactobacillales
probability 0
0.1
0.4
0.000
0.002
Rhodospirillales
probability 0
0.1
0.4
0.000
0.003
Micrococcales
probability 0
0.1
0.4
0.00
0.03
0.06
Methanobacteriales
probability 0
0.1
0.4
0.00
0.03
0.06
Fusobacteriales


probability 0
0.1
0.4
0.000
0.006
Desulfovibrionales
probability 0
0.1
0.4
0.0000
0.0015
Flavobacteriales
probability 0
0.1
0.4
0.000
0.010
Clostridia_vadinBB60_group
probability 0
0.1
0.4
0.000
0.003
0.006
Izemoplasmatales
probability 0
0.1
0.4
0.000
0.015
0.030
Saccharimonadales
probability 0
0.1
0.4
0.000
0.003
0.006
Monoglobales
probability 0
0.1
0.4
0.00000
0.00015
Chloroplast
probability 0
0.1
0.4
0.000
0.004
0.008
Burkholderiales
probability 0
0.1
0.4
0.00
0.15
0.30
Verrucomicrobiales
probability 0
0.1
0.4
0.0000
0.0008
Victivallales
probability 0
0.1
0.4
0.000
0.003
Synergistales
probability 0
0.1
0.4
0.000
0.006
Staphylococcales
probability 0
0.10
0.35
0.0000000
0.0032025
Rhizobiales
probability 0
0.1
0.4
0.000
0.002
RF39
probability 0
0.1
0.4
0.0000
0.0020
Peptococcales
probability 0
0.1
0.4
0.00
0.04
0.08
Pasteurellales

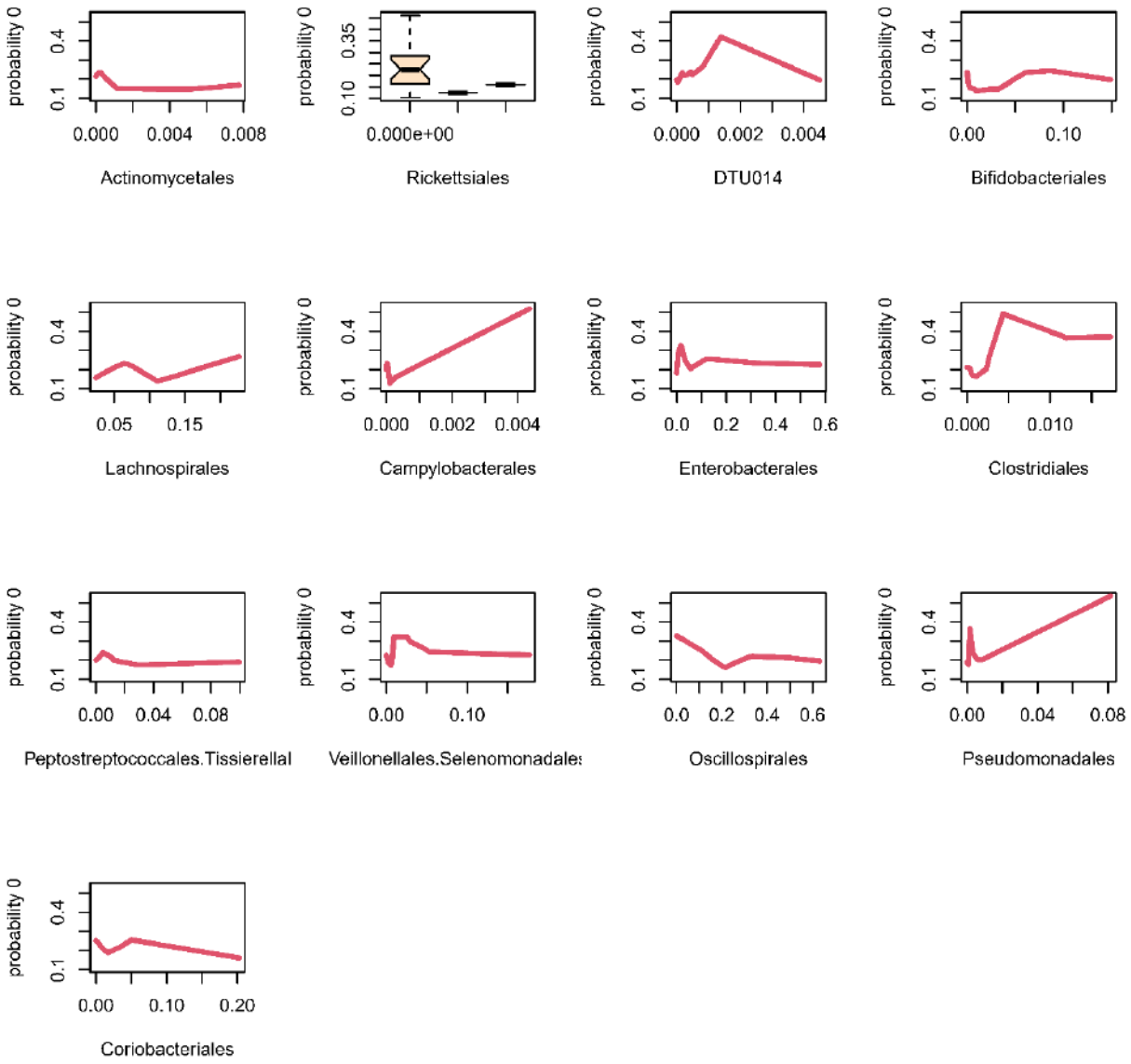
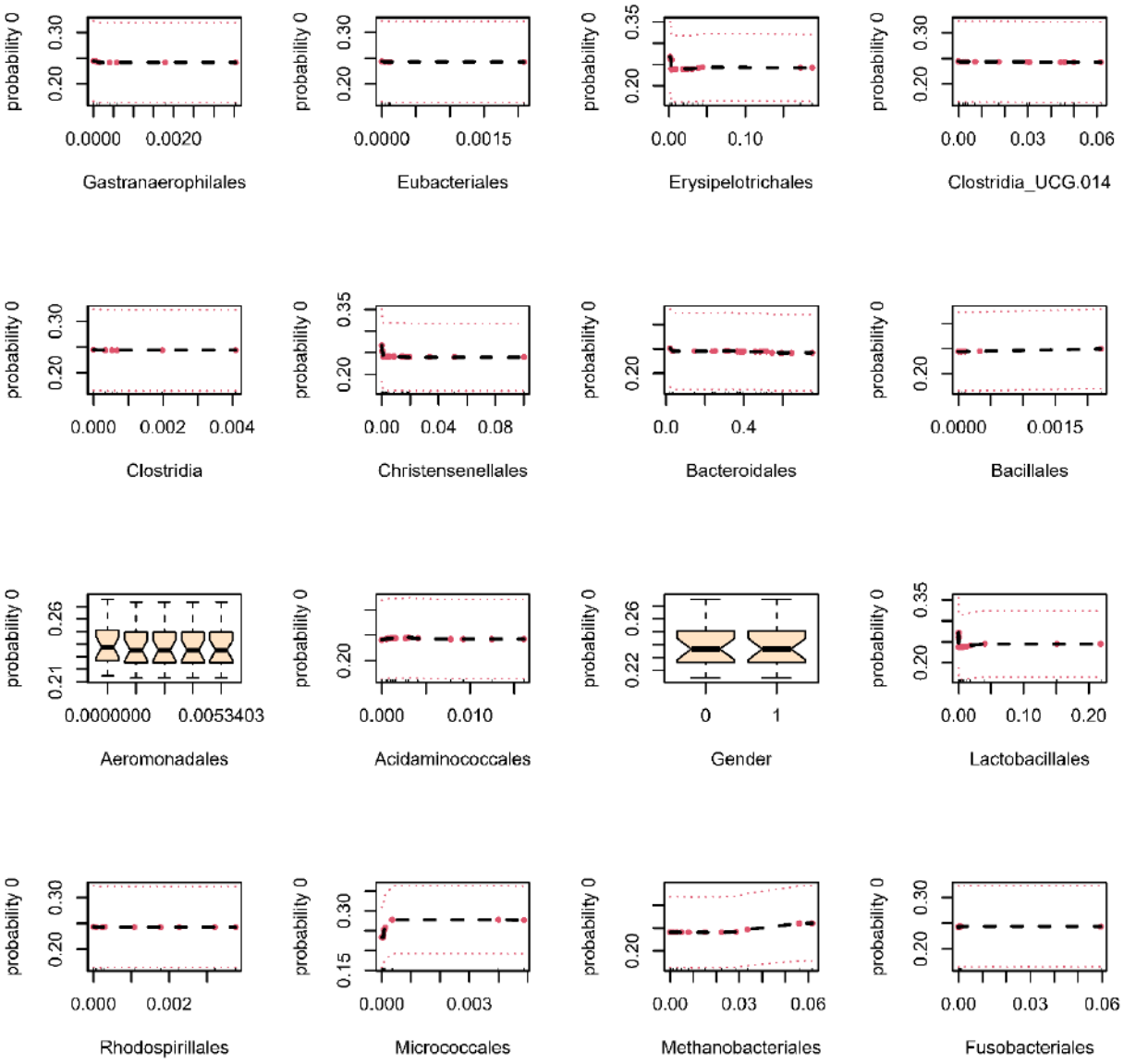
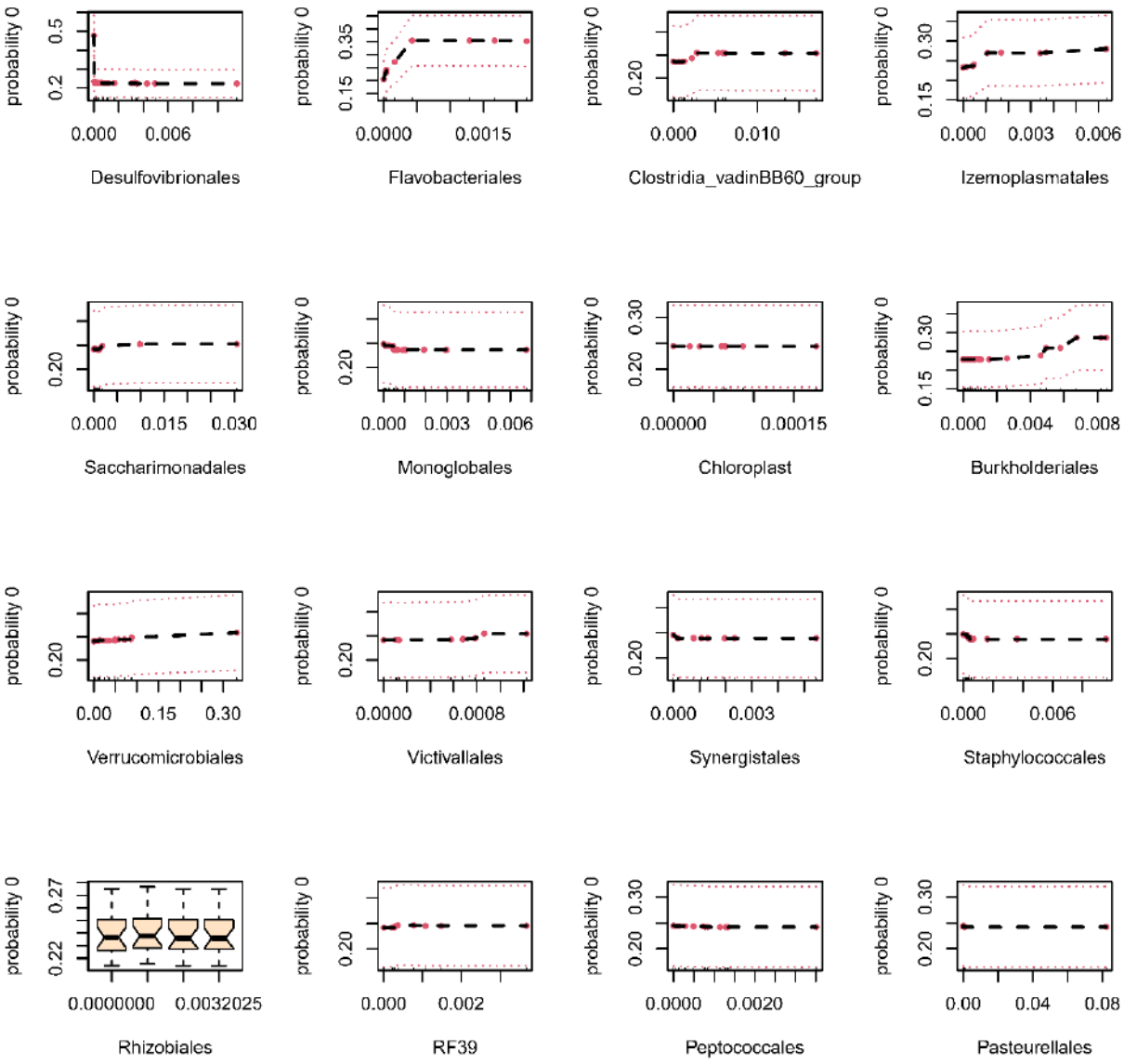


Figure ML2-C-2: Marginal Plot for Order in Stool


probability 0
Gastranaerophilales
Eubacteriales
Erysipelotrichales
Clostridia_UCG.014
Clostridia
Christensenellales
Bacteroidales
Bacillales
Aeromonadales
Acidaminococcales
Gender
Lactobacillales
Rhodospirillales
Micrococcales
Methanobacteriales
Fusobacteriales


probability 0
Desulfovibrionales
Flavobacteriales
Clostridia_vadinBB60_group
Izemoplasmatales
Saccharimonadales
Monoglobales
Chloroplast
Burkholderiales
Verrucomicrobiales
Victivallales
Synergistales
Staphylococcales
Rhizobiales
RF39
Peptococcales
Pasteurellales

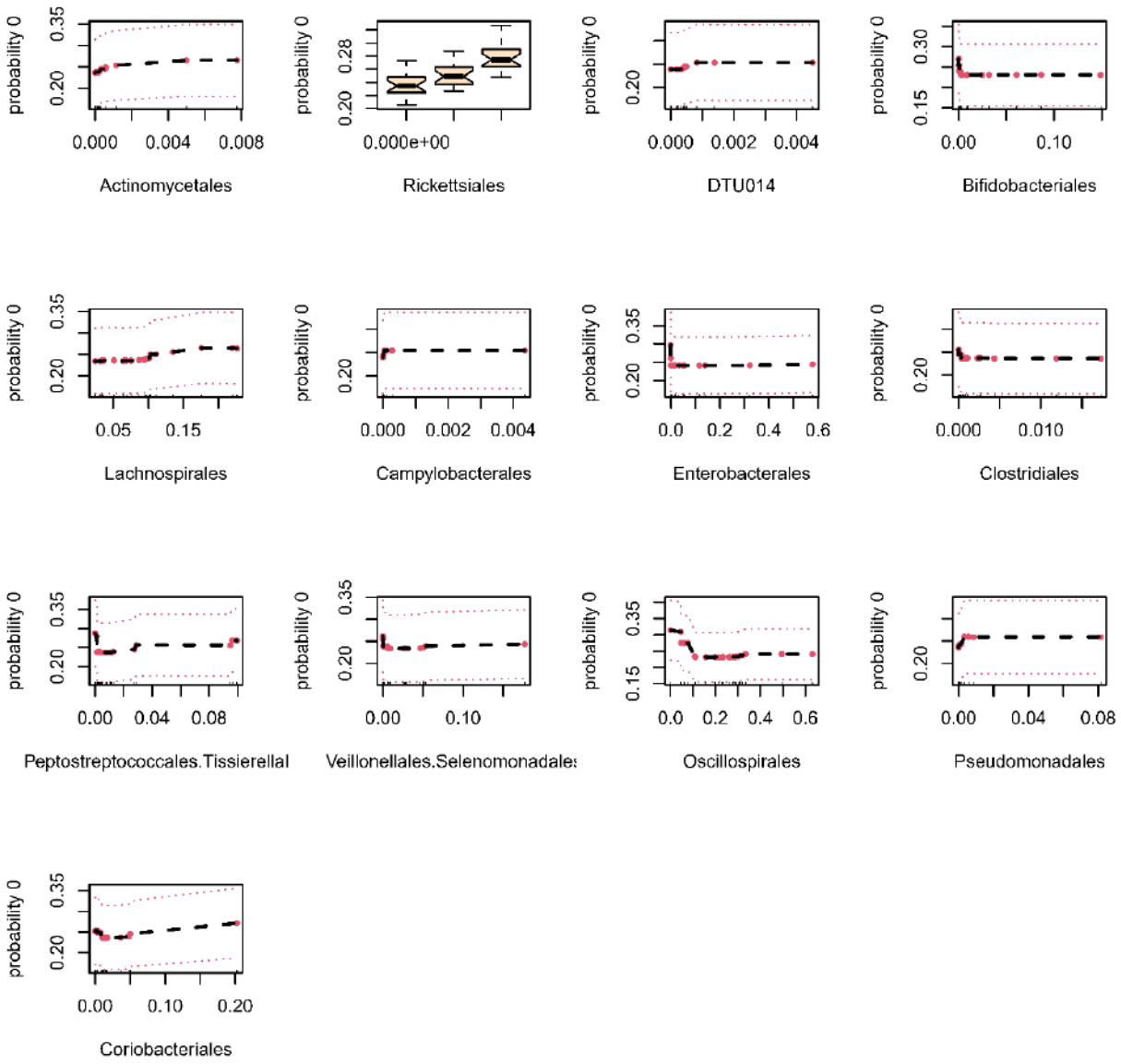


Figure ML2-C-3: Partial Plot for Order in Stool

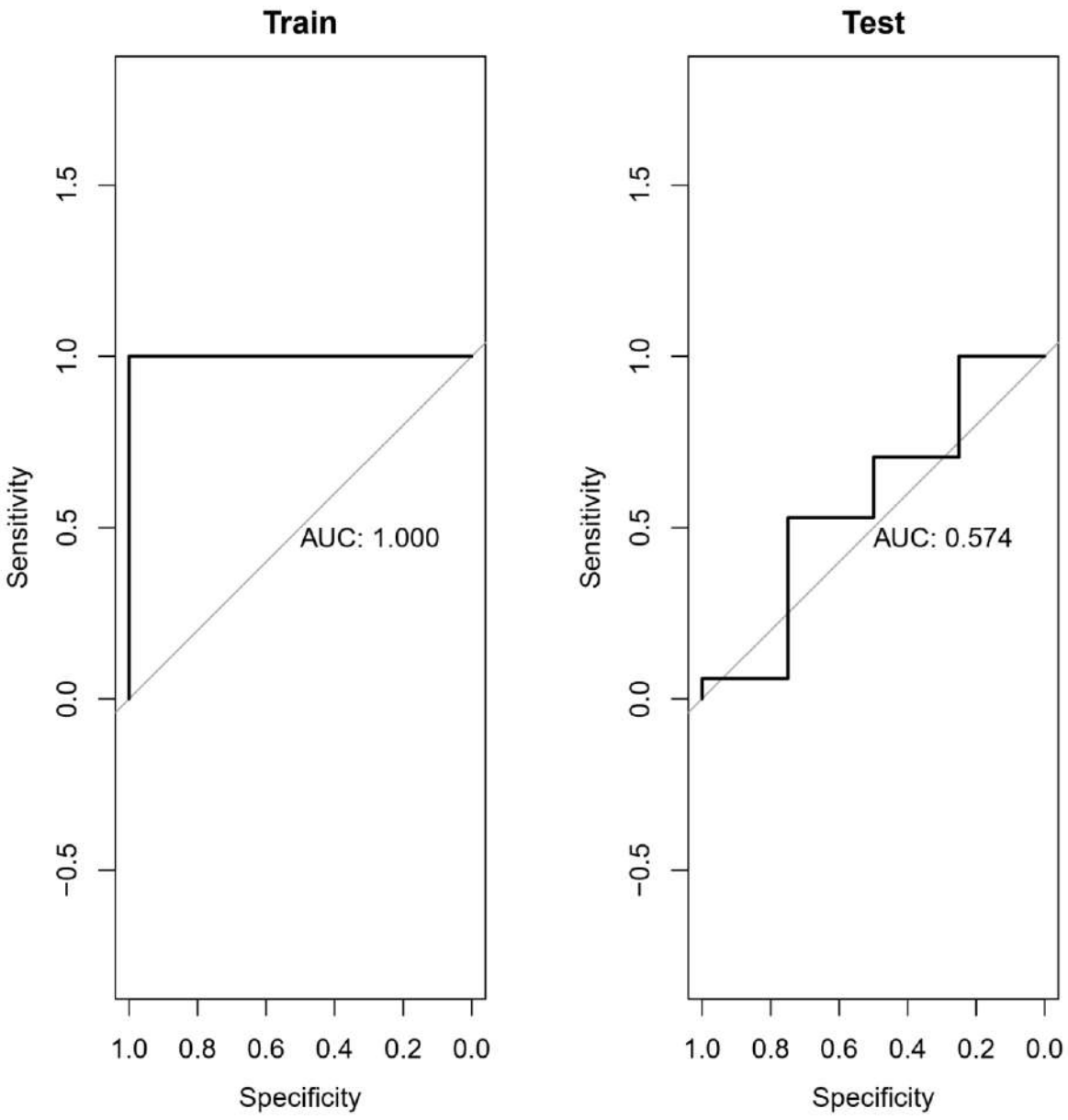


Figure ML2-C-4: The ROC Curve for Order in Stool

Table ML2-C-1: The Model Performance

Index	Values		
	2.50%	50.00%	97.50%
Accuracy	0.24	0.71	0.95
Specificity	0.25	0.75	1.00
Sensitivity	0.06	0.71	1.00

Table ML2-C-2: The Variable Importance Values

The Varibale Importnace Values						
Covariates	Varibale Importance			Brier Score		
	all	AP	CRC	all	AP	CRC
Gender	-0.0005	-0.0054	0.0000	-0.0004	-0.0002	-0.0002
Acidaminococcales	0.0023	0.0054	0.0064	0.0011	0.0008	0.0008
Actinomycetales	-0.0010	-0.0272	0.0048	0.0054	0.0037	0.0037
Aeromonadales	0.0009	0.0000	0.0032	0.0023	0.0016	0.0016
Bacillales	0.0023	-0.0272	0.0160	0.0029	0.0020	0.0020
Bacteroidales	-0.0014	-0.0163	0.0000	0.0075	0.0051	0.0051
Bifidobacteriales	0.0084	-0.0435	0.0416	-0.0093	-0.0063	-0.0063
Burkholderiales	0.0059	-0.0761	0.0432	0.0417	0.0283	0.0283
Campylobacterales	0.0038	-0.0054	0.0144	-0.0056	-0.0038	-0.0038
Chloroplast	-0.0005	0.0054	-0.0032	-0.0017	-0.0012	-0.0012
Christensenellales	0.0046	-0.0109	0.0192	0.0084	0.0057	0.0057
Clostridia	0.0019	0.0054	0.0048	0.0044	0.0030	0.0030
Clostridia_UCG.014	-0.0023	0.0000	-0.0080	-0.0031	-0.0021	-0.0021
Clostridia_vadinBB60_group	-0.0018	-0.0272	0.0016	-0.0128	-0.0087	-0.0087
Clostridiales	-0.0046	-0.0163	-0.0112	-0.0216	-0.0147	-0.0147
Coriobacteriales	0.0126	-0.0924	0.0704	0.0119	0.0081	0.0081
Desulfovibrionales	0.1722	-0.3806	0.7116	0.5754	0.3910	0.3910
DTU014	0.0027	-0.0109	0.0128	-0.0053	-0.0036	-0.0036
Enterobacterales	0.0209	-0.1903	0.1279	0.0509	0.0346	0.0346
Erysipelotrichales	0.0001	-0.0544	0.0160	-0.0006	-0.0004	-0.0004
Eubacteriales	0.0000	0.0054	-0.0016	-0.0031	-0.0021	-0.0021
Flavobacteriales	0.0387	-0.1087	0.1663	0.1827	0.1241	0.1241
Fusobacteriales	-0.0032	-0.0163	-0.0064	-0.0025	-0.0017	-0.0017
Gastranaerophilales	-0.0019	0.0109	-0.0096	-0.0002	-0.0002	-0.0002
Izemoplasmatales	-0.0005	-0.0544	0.0144	0.0047	0.0032	0.0032
Lachnospirales	0.0028	-0.0326	0.0192	0.0008	0.0005	0.0005
Lactobacillales	0.0037	-0.0272	0.0208	0.0165	0.0112	0.0112
Methanobacteriales	0.0010	-0.0544	0.0192	0.0063	0.0043	0.0043
Micrococcales	0.0087	-0.0707	0.0512	0.0268	0.0182	0.0182
Monoglobales	-0.0032	0.0000	-0.0112	-0.0053	-0.0036	-0.0036

Oscillospirales	0.0309	-0.1631	0.1551	0.0681	0.0463	0.0463
Pasteurellales	-0.0014	0.0000	-0.0048	-0.0024	-0.0017	-0.0017
Peptococcales	-0.0014	0.0054	-0.0064	-0.0013	-0.0009	-0.0009
Peptostreptococcales.Tissierellales	0.0291	-0.1740	0.1519	0.0497	0.0338	0.0338
Pseudomonadales	-0.0050	-0.0326	-0.0080	-0.0192	-0.0130	-0.0130
RF39	-0.0009	-0.0054	-0.0016	-0.0047	-0.0032	-0.0032
Rhizobiales	-0.0009	0.0054	-0.0048	0.0001	0.0001	0.0001
Rhodospirillales	-0.0023	0.0000	-0.0080	-0.0067	-0.0046	-0.0046
Rickettsiales	0.0141	-0.1033	0.0799	0.0271	0.0184	0.0184
Saccharimonadales	0.0014	-0.0326	0.0144	0.0014	0.0010	0.0010
Staphylococcales	-0.0004	-0.0109	0.0016	0.0039	0.0027	0.0027
Synergistales	0.0009	0.0054	0.0016	0.0049	0.0034	0.0034
Veillonellales.Selenomonadales	0.0024	-0.0652	0.0272	-0.0047	-0.0032	-0.0032
Verrucomicrobiales	0.0013	-0.0217	0.0112	-0.0017	-0.0011	-0.0011
Victivallales	0.0028	-0.0163	0.0144	0.0078	0.0053	0.0053

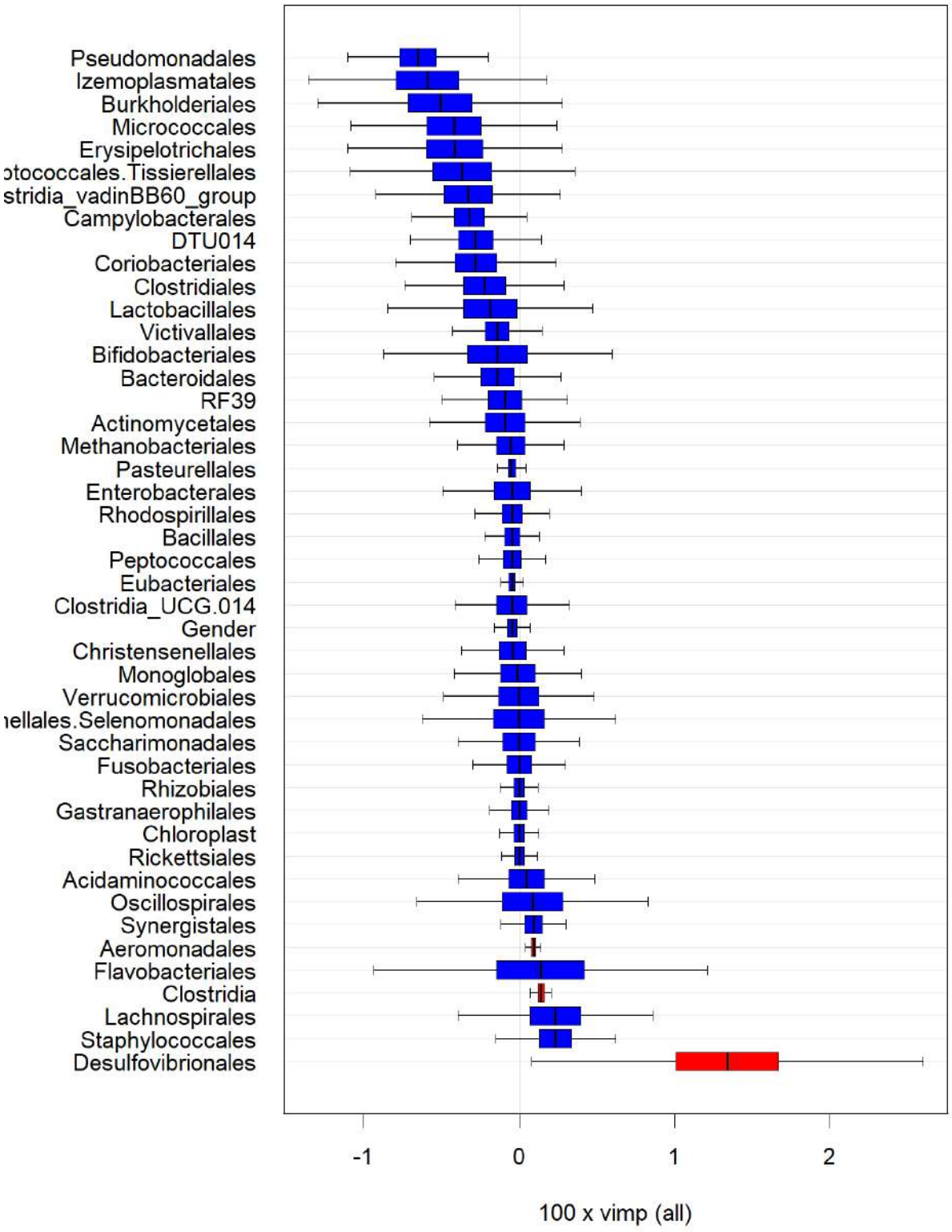


Figure ML2-C-5: The Variable Importance Plot (VIMP) for Order in Stool

Best model is selected for LIME results

Table ML2-C-3: The best setting for model and LIME Results

Setting	mtry	ROC	Sens	Spec	ROCSD	SensSD	SpecSD
ntrees: 500 nodesize: 1	3	0.7	0	1	0.240081	0	0
ntrees: 500 nodesize: 1	5	0.7	0	1	0.240081	0	0
ntrees: 500 nodesize: 1	6	0.7	0	1	0.240081	0	0
ntrees: 500 nodesize: 1	7	0.7	0	0.95	0.240081	0	0.111803
ntrees: 500 nodesize: 1	10	0.7	0	1	0.240081	0	0
ntrees: 1000 nodesize: 1	3	0.766667	0	1	0.27259	0	0
ntrees: 500 nodesize: 5	3	0.766667	0	1	0.27259	0	0
ntrees: 1000 nodesize: 5	8	0.75	0	0.95	0.276385	0	0.111803

Table ML2-C-4: The LIME Result for Training Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
1	Adenoma	0.730	0.339	0.222	0.326	Izemoplasmatales	0.006	0.037	0.000464 < Izemoplasmatales
1	Adenoma	0.730	0.339	0.222	0.326	Flavobacteriales	0.000	0.038	4.36e-05 < Flavobacteriales
1	Adenoma	0.730	0.339	0.222	0.326	Burkholderiales	0.008	0.034	0.004110 < Burkholderiales
1	Adenoma	0.730	0.339	0.222	0.326	Micrococcales	0.000	-0.030	Micrococcales <= 7.24e-05
1	Adenoma	0.730	0.339	0.222	0.326	Bifidobacteriales	0.000	0.026	Bifidobacteriales <= 0.000582
2	Adenoma	0.734	0.335	0.219	0.326	Izemoplasmatales	0.001	0.038	0.000464 < Izemoplasmatales
2	Adenoma	0.734	0.335	0.219	0.326	Flavobacteriales	0.001	0.039	4.36e-05 < Flavobacteriales
2	Adenoma	0.734	0.335	0.219	0.326	Burkholderiales	0.007	0.031	0.004110 < Burkholderiales
2	Adenoma	0.734	0.335	0.219	0.326	Micrococcales	0.000	-0.029	Micrococcales <= 7.24e-05
2	Adenoma	0.734	0.335	0.219	0.326	Actinomycetales	0.005	0.028	0.000341 < Actinomycetales
3	Adenoma	0.688	0.297	0.283	0.249	Izemoplasmatales	0.000	-0.037	Izemoplasmatales <= 0.000464
3	Adenoma	0.688	0.297	0.283	0.249	Flavobacteriales	0.000	-0.037	Flavobacteriales <= 4.36e-05
3	Adenoma	0.688	0.297	0.283	0.249	Micrococcales	0.000	0.034	7.24e-05 < Micrococcales
3	Adenoma	0.688	0.297	0.283	0.249	Veillonellales.Selenomonadales	0.000	0.026	Veillonellales.Selenomonadales <= 0.000246
3	Adenoma	0.688	0.297	0.283	0.249	Actinomycetales	0.000	-0.020	Actinomycetales <= 0.000136
6	Adenoma	0.732	0.336	0.189	0.357	Izemoplasmatales	0.002	0.037	0.000464 < Izemoplasmatales
6	Adenoma	0.732	0.336	0.189	0.357	Flavobacteriales	0.002	0.038	4.36e-05 < Flavobacteriales
6	Adenoma	0.732	0.336	0.189	0.357	Micrococcales	0.000	0.035	7.24e-05 < Micrococcales
6	Adenoma	0.732	0.336	0.189	0.357	Burkholderiales	0.005	0.034	0.004110 < Burkholderiales
6	Adenoma	0.732	0.336	0.189	0.357	Bifidobacteriales	0.001	0.023	Bifidobacteriales <= 0.000582
9	Adenoma	0.754	0.342	0.265	0.282	Izemoplasmatales	0.000	-0.038	Izemoplasmatales <= 0.000464
9	Adenoma	0.754	0.342	0.265	0.282	Flavobacteriales	0.000	-0.036	Flavobacteriales <= 4.36e-05
9	Adenoma	0.754	0.342	0.265	0.282	Micrococcales	0.004	0.032	7.24e-05 < Micrococcales
9	Adenoma	0.754	0.342	0.265	0.282	Oscillospirales	0.001	0.031	Oscillospirales <= 0.108
9	Adenoma	0.754	0.342	0.265	0.282	Actinomycetales	0.000	0.029	0.000341 < Actinomycetales
11	Adenocarcinoma	0.812	0.349	0.740	0.712	Izemoplasmatales	0.003	-0.038	0.000464 < Izemoplasmatales
11	Adenocarcinoma	0.812	0.349	0.740	0.712	Flavobacteriales	0.000	0.037	Flavobacteriales <= 4.36e-05

11	Adenocarcinoma	0.812	0.349	0.740	0.712	Burkholderiales	0.006	-0.032	0.004110 < Burkholderiales
11	Adenocarcinoma	0.812	0.349	0.740	0.712	Micrococcales	0.000	0.032	Micrococcales <= 7.24e-05
11	Adenocarcinoma	0.812	0.349	0.740	0.712	Oscillospirales	0.047	-0.027	Oscillospirales <= 0.108
17	Adenocarcinoma	0.874	0.322	0.741	0.717	Izemoplasmatales	0.000	0.038	Izemoplasmatales <= 0.000464
17	Adenocarcinoma	0.874	0.322	0.741	0.717	Flavobacteriales	0.000	-0.037	4.36e-05 < Flavobacteriales
17	Adenocarcinoma	0.874	0.322	0.741	0.717	Micrococcales	0.000	0.031	Micrococcales <= 7.24e-05
17	Adenocarcinoma	0.874	0.322	0.741	0.717	Oscillospirales	0.063	-0.030	Oscillospirales <= 0.108
17	Adenocarcinoma	0.874	0.322	0.741	0.717	Bifidobacteriales	0.000	-0.026	Bifidobacteriales <= 0.000582
18	Adenocarcinoma	0.880	0.319	0.702	0.754	Izemoplasmatales	0.000	0.039	Izemoplasmatales <= 0.000464
18	Adenocarcinoma	0.880	0.319	0.702	0.754	Flavobacteriales	0.000	0.038	Flavobacteriales <= 4.36e-05
18	Adenocarcinoma	0.880	0.319	0.702	0.754	Burkholderiales	0.005	-0.031	0.004110 < Burkholderiales
18	Adenocarcinoma	0.880	0.319	0.702	0.754	Micrococcales	0.000	0.030	Micrococcales <= 7.24e-05
18	Adenocarcinoma	0.880	0.319	0.702	0.754	Bifidobacteriales	0.000	-0.024	Bifidobacteriales <= 0.000582
19	Adenocarcinoma	0.884	0.296	0.700	0.762	Izemoplasmatales	0.000	0.039	Izemoplasmatales <= 0.000464
19	Adenocarcinoma	0.884	0.296	0.700	0.762	Flavobacteriales	0.000	0.037	Flavobacteriales <= 4.36e-05
19	Adenocarcinoma	0.884	0.296	0.700	0.762	Micrococcales	0.000	0.031	Micrococcales <= 7.24e-05
19	Adenocarcinoma	0.884	0.296	0.700	0.762	Actinomycetales	0.001	-0.026	0.000341 < Actinomycetales
19	Adenocarcinoma	0.884	0.296	0.700	0.762	Lachnospirales	0.227	-0.019	0.1026 < Lachnospirales
20	Adenocarcinoma	0.890	0.298	0.684	0.788	Izemoplasmatales	0.000	0.039	Izemoplasmatales <= 0.000464
20	Adenocarcinoma	0.890	0.298	0.684	0.788	Flavobacteriales	0.000	0.038	Flavobacteriales <= 4.36e-05
20	Adenocarcinoma	0.890	0.298	0.684	0.788	Micrococcales	0.000	0.031	Micrococcales <= 7.24e-05
20	Adenocarcinoma	0.890	0.298	0.684	0.788	Bifidobacteriales	0.000	-0.025	Bifidobacteriales <= 0.000582
20	Adenocarcinoma	0.890	0.298	0.684	0.788	Actinomycetales	0.000	0.022	Actinomycetales <= 0.000136
21	Adenocarcinoma	0.966	0.277	0.709	0.779	Izemoplasmatales	0.000	0.039	Izemoplasmatales <= 0.000464
21	Adenocarcinoma	0.966	0.277	0.709	0.779	Flavobacteriales	0.000	-0.037	4.36e-05 < Flavobacteriales
21	Adenocarcinoma	0.966	0.277	0.709	0.779	Micrococcales	0.000	0.031	Micrococcales <= 7.24e-05
21	Adenocarcinoma	0.966	0.277	0.709	0.779	Actinomycetales	0.000	0.020	Actinomycetales <= 0.000136
21	Adenocarcinoma	0.966	0.277	0.709	0.779	Coriobacteriales	0.012	0.016	0.0110 < Coriobacteriales <= 0.0156
22	Adenocarcinoma	0.936	0.257	0.688	0.795	Izemoplasmatales	0.000	0.038	Izemoplasmatales <= 0.000464
22	Adenocarcinoma	0.936	0.257	0.688	0.795	Flavobacteriales	0.000	0.037	Flavobacteriales <= 4.36e-05
22	Adenocarcinoma	0.936	0.257	0.688	0.795	Micrococcales	0.000	0.031	Micrococcales <= 7.24e-05
22	Adenocarcinoma	0.936	0.257	0.688	0.795	Oscillospirales	0.157	0.017	0.108 < Oscillospirales <= 0.233
22	Adenocarcinoma	0.936	0.257	0.688	0.795	Desulfovibrionales	0.000	-0.015	Desulfovibrionales <= 0.000170
23	Adenocarcinoma	0.900	0.280	0.702	0.761	Izemoplasmatales	0.000	0.038	Izemoplasmatales <= 0.000464
23	Adenocarcinoma	0.900	0.280	0.702	0.761	Flavobacteriales	0.000	0.037	Flavobacteriales <= 4.36e-05
23	Adenocarcinoma	0.900	0.280	0.702	0.761	Micrococcales	0.000	0.028	Micrococcales <= 7.24e-05
23	Adenocarcinoma	0.900	0.280	0.702	0.761	Actinomycetales	0.001	-0.028	0.000341 < Actinomycetales
23	Adenocarcinoma	0.900	0.280	0.702	0.761	Saccharimonadales	0.002	-0.016	0.000824 < Saccharimonadales
25	Adenocarcinoma	0.932	0.303	0.690	0.788	Izemoplasmatales	0.000	0.039	Izemoplasmatales <= 0.000464
25	Adenocarcinoma	0.932	0.303	0.690	0.788	Flavobacteriales	0.000	0.036	Flavobacteriales <= 4.36e-05
25	Adenocarcinoma	0.932	0.303	0.690	0.788	Micrococcales	0.000	0.033	Micrococcales <= 7.24e-05
25	Adenocarcinoma	0.932	0.303	0.690	0.788	Actinomycetales	0.008	-0.028	0.000341 < Actinomycetales
25	Adenocarcinoma	0.932	0.303	0.690	0.788	Oscillospirales	0.232	0.018	0.108 < Oscillospirales <= 0.233

26	Adenocarcinoma	0.910	0.292	0.782	0.709	Flavobacteriales	0.002	-0.040	4.36e-05 < Flavobacteriales
26	Adenocarcinoma	0.910	0.292	0.782	0.709	Izemoplasmatales	0.000	-0.038	0.000464 < Izemoplasmatales
26	Adenocarcinoma	0.910	0.292	0.782	0.709	Micrococcales	0.000	-0.031	7.24e-05 < Micrococcales
26	Adenocarcinoma	0.910	0.292	0.782	0.709	Actinomycetales	0.000	0.019	Actinomycetales <= 0.000136
26	Adenocarcinoma	0.910	0.292	0.782	0.709	Bifidobacteriales	0.061	0.016	0.008930 < Bifidobacteriales
29	Adenocarcinoma	0.906	0.286	0.684	0.792	Flavobacteriales	0.000	0.043	Flavobacteriales <= 4.36e-05
29	Adenocarcinoma	0.906	0.286	0.684	0.792	Izemoplasmatales	0.000	0.038	Izemoplasmatales <= 0.000464
29	Adenocarcinoma	0.906	0.286	0.684	0.792	Micrococcales	0.000	0.031	Micrococcales <= 7.24e-05
29	Adenocarcinoma	0.906	0.286	0.684	0.792	Lachnospirales	0.135	-0.021	0.1026 < Lachnospirales
29	Adenocarcinoma	0.906	0.286	0.684	0.792	Bifidobacteriales	0.009	0.017	0.008930 < Bifidobacteriales
30	Adenocarcinoma	0.888	0.294	0.724	0.747	Flavobacteriales	0.000	0.037	Flavobacteriales <= 4.36e-05
30	Adenocarcinoma	0.888	0.294	0.724	0.747	Izemoplasmatales	0.004	-0.036	0.000464 < Izemoplasmatales
30	Adenocarcinoma	0.888	0.294	0.724	0.747	Micrococcales	0.000	0.032	Micrococcales <= 7.24e-05
30	Adenocarcinoma	0.888	0.294	0.724	0.747	Veillonellales.Selenomonadales	0.000	-0.027	Veillonellales.Selenomonadales <= 0.000246
30	Adenocarcinoma	0.888	0.294	0.724	0.747	Actinomycetales	0.000	0.018	Actinomycetales <= 0.000136
31	Adenocarcinoma	0.900	0.348	0.736	0.718	Flavobacteriales	0.000	0.038	Flavobacteriales <= 4.36e-05
31	Adenocarcinoma	0.900	0.348	0.736	0.718	Izemoplasmatales	0.000	0.037	Izemoplasmatales <= 0.000464
31	Adenocarcinoma	0.900	0.348	0.736	0.718	Micrococcales	0.005	-0.032	7.24e-05 < Micrococcales
31	Adenocarcinoma	0.900	0.348	0.736	0.718	Burkholderiales	0.009	-0.032	0.004110 < Burkholderiales
31	Adenocarcinoma	0.900	0.348	0.736	0.718	Oscillospirales	0.108	-0.029	Oscillospirales <= 0.108
35	Adenocarcinoma	0.886	0.309	0.681	0.790	Izemoplasmatales	0.000	0.041	Izemoplasmatales <= 0.000464
35	Adenocarcinoma	0.886	0.309	0.681	0.790	Flavobacteriales	0.000	0.039	Flavobacteriales <= 4.36e-05
35	Adenocarcinoma	0.886	0.309	0.681	0.790	Micrococcales	0.000	0.033	Micrococcales <= 7.24e-05
35	Adenocarcinoma	0.886	0.309	0.681	0.790	Veillonellales.Selenomonadales	0.000	-0.022	Veillonellales.Selenomonadales <= 0.000246
35	Adenocarcinoma	0.886	0.309	0.681	0.790	Actinomycetales	0.000	0.019	Actinomycetales <= 0.000136
41	Adenocarcinoma	0.942	0.328	0.698	0.759	Flavobacteriales	0.000	0.039	Flavobacteriales <= 4.36e-05
41	Adenocarcinoma	0.942	0.328	0.698	0.759	Izemoplasmatales	0.000	0.040	Izemoplasmatales <= 0.000464
41	Adenocarcinoma	0.942	0.328	0.698	0.759	Micrococcales	0.000	0.031	Micrococcales <= 7.24e-05
41	Adenocarcinoma	0.942	0.328	0.698	0.759	Actinomycetales	0.001	-0.030	0.000341 < Actinomycetales
41	Adenocarcinoma	0.942	0.328	0.698	0.759	Lachnospirales	0.111	-0.020	0.1026 < Lachnospirales
42	Adenocarcinoma	0.916	0.275	0.709	0.764	Flavobacteriales	0.000	0.039	Flavobacteriales <= 4.36e-05
42	Adenocarcinoma	0.916	0.275	0.709	0.764	Izemoplasmatales	0.000	0.039	Izemoplasmatales <= 0.000464
42	Adenocarcinoma	0.916	0.275	0.709	0.764	Micrococcales	0.000	-0.031	7.24e-05 < Micrococcales
42	Adenocarcinoma	0.916	0.275	0.709	0.764	Actinomycetales	0.000	0.021	Actinomycetales <= 0.000136
42	Adenocarcinoma	0.916	0.275	0.709	0.764	Christensenellales	0.001	-0.014	Christensenellales <= 0.00112
43	Adenocarcinoma	0.940	0.275	0.689	0.791	Flavobacteriales	0.000	0.039	Flavobacteriales <= 4.36e-05
43	Adenocarcinoma	0.940	0.275	0.689	0.791	Izemoplasmatales	0.000	0.039	Izemoplasmatales <= 0.000464
43	Adenocarcinoma	0.940	0.275	0.689	0.791	Micrococcales	0.000	0.030	Micrococcales <= 7.24e-05
43	Adenocarcinoma	0.940	0.275	0.689	0.791	Veillonellales.Selenomonadales	0.000	-0.023	Veillonellales.Selenomonadales <= 0.000246
43	Adenocarcinoma	0.940	0.275	0.689	0.791	Oscillospirales	0.234	0.019	0.233 < Oscillospirales <= 0.299

Table ML-C-5: The LIME Result for Test Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
4	Adenocarcinoma	0.696	0.312	0.732	0.728	Izemoplasmatales	0.000	0.038	Izemoplasmatales <= 0.000464
4	Adenocarcinoma	0.696	0.312	0.732	0.728	Flavobacteriales	0.000	0.038	Flavobacteriales <= 4.36e-05
4	Adenocarcinoma	0.696	0.312	0.732	0.728	Micrococcales	0.000	-0.032	7.24e-05 < Micrococcales
4	Adenocarcinoma	0.696	0.312	0.732	0.728	Actinomycetales	0.001	-0.026	0.000341 < Actinomycetales
4	Adenocarcinoma	0.696	0.312	0.732	0.728	Lachnospirales	0.137	-0.023	0.1026 < Lachnospirales
5	Adenocarcinoma	0.790	0.231	0.691	0.793	Izemoplasmatales	0.000	0.038	Izemoplasmatales <= 0.000464
5	Adenocarcinoma	0.790	0.231	0.691	0.793	Flavobacteriales	0.000	0.035	Flavobacteriales <= 4.36e-05
5	Adenocarcinoma	0.790	0.231	0.691	0.793	Micrococcales	0.000	0.029	Micrococcales <= 7.24e-05
5	Adenocarcinoma	0.790	0.231	0.691	0.793	Oscillospirales	0.128	0.015	0.108 < Oscillospirales <= 0.233
5	Adenocarcinoma	0.790	0.231	0.691	0.793	Christensenellales	0.000	-0.014	Christensenellales <= 0.00112
7	Adenocarcinoma	0.816	0.255	0.718	0.767	Izemoplasmatales	0.000	0.041	Izemoplasmatales <= 0.000464
7	Adenocarcinoma	0.816	0.255	0.718	0.767	Flavobacteriales	0.000	0.034	Flavobacteriales <= 4.36e-05
7	Adenocarcinoma	0.816	0.255	0.718	0.767	Oscillospirales	0.055	-0.030	Oscillospirales <= 0.108
7	Adenocarcinoma	0.816	0.255	0.718	0.767	Bifidobacteriales	0.012	0.018	0.008930 < Bifidobacteriales
7	Adenocarcinoma	0.816	0.255	0.718	0.767	Pseudomonadales	0.002	-0.014	8.26e-04 < Pseudomonadales
8	Adenocarcinoma	0.758	0.301	0.760	0.715	Flavobacteriales	0.000	-0.040	4.36e-05 < Flavobacteriales
8	Adenocarcinoma	0.758	0.301	0.760	0.715	Izemoplasmatales	0.000	0.038	Izemoplasmatales <= 0.000464
8	Adenocarcinoma	0.758	0.301	0.760	0.715	Micrococcales	0.000	-0.030	7.24e-05 < Micrococcales
8	Adenocarcinoma	0.758	0.301	0.760	0.715	Actinomycetales	0.001	-0.029	0.000341 < Actinomycetales
8	Adenocarcinoma	0.758	0.301	0.760	0.715	Bifidobacteriales	0.022	0.016	0.008930 < Bifidobacteriales
10	Adenocarcinoma	0.806	0.253	0.691	0.796	Izemoplasmatales	0.000	0.038	Izemoplasmatales <= 0.000464
10	Adenocarcinoma	0.806	0.253	0.691	0.796	Flavobacteriales	0.000	0.037	Flavobacteriales <= 4.36e-05
10	Adenocarcinoma	0.806	0.253	0.691	0.796	Micrococcales	0.000	0.032	Micrococcales <= 7.24e-05
10	Adenocarcinoma	0.806	0.253	0.691	0.796	Staphylococcales	0.000	-0.014	Staphylococcales <= 0.000264
10	Adenocarcinoma	0.806	0.253	0.691	0.796	Veillonellales.Selenomonadales	0.006	0.011	0.004161 < Veillonellales.Selenomonadales <= 0.027607
12	Adenocarcinoma	0.766	0.296	0.684	0.789	Izemoplasmatales	0.000	0.037	Izemoplasmatales <= 0.000464
12	Adenocarcinoma	0.766	0.296	0.684	0.789	Flavobacteriales	0.000	0.036	Flavobacteriales <= 4.36e-05
12	Adenocarcinoma	0.766	0.296	0.684	0.789	Micrococcales	0.000	0.033	Micrococcales <= 7.24e-05
12	Adenocarcinoma	0.766	0.296	0.684	0.789	Actinomycetales	0.000	0.020	Actinomycetales <= 0.000136
12	Adenocarcinoma	0.766	0.296	0.684	0.789	Lachnospirales	0.173	-0.022	0.1026 < Lachnospirales
13	Adenocarcinoma	0.772	0.289	0.682	0.787	Izemoplasmatales	0.000	0.041	Izemoplasmatales <= 0.000464
13	Adenocarcinoma	0.772	0.289	0.682	0.787	Flavobacteriales	0.000	0.037	Flavobacteriales <= 4.36e-05
13	Adenocarcinoma	0.772	0.289	0.682	0.787	Micrococcales	0.000	0.031	Micrococcales <= 7.24e-05
13	Adenocarcinoma	0.772	0.289	0.682	0.787	Lachnospirales	0.118	-0.022	0.1026 < Lachnospirales
13	Adenocarcinoma	0.772	0.289	0.682	0.787	Actinomycetales	0.000	0.019	Actinomycetales <= 0.000136
14	Adenocarcinoma	0.898	0.305	0.704	0.760	Flavobacteriales	0.000	0.037	Flavobacteriales <= 4.36e-05
14	Adenocarcinoma	0.898	0.305	0.704	0.760	Izemoplasmatales	0.000	0.035	Izemoplasmatales <= 0.000464
14	Adenocarcinoma	0.898	0.305	0.704	0.760	Micrococcales	0.000	0.032	Micrococcales <= 7.24e-05
14	Adenocarcinoma	0.898	0.305	0.704	0.760	Actinomycetales	0.001	-0.030	0.000341 < Actinomycetales
14	Adenocarcinoma	0.898	0.305	0.704	0.760	Lachnospirales	0.112	-0.018	0.1026 < Lachnospirales

15	Adenocarcinoma	0.792	0.255	0.726	0.760	Flavobacteriales	0.001	-0.036	4.36e-05 < Flavobacteriales
15	Adenocarcinoma	0.792	0.255	0.726	0.760	Izemoplasmatales	0.000	0.037	Izemoplasmatales <= 0.000464
15	Adenocarcinoma	0.792	0.255	0.726	0.760	Micrococcales	0.000	0.031	Micrococcales <= 7.24e-05
15	Adenocarcinoma	0.792	0.255	0.726	0.760	Bifidobacteriales	0.014	0.018	0.008930 < Bifidobacteriales
15	Adenocarcinoma	0.792	0.255	0.726	0.760	Clostridia_vadinBB60_group	0.005	-0.016	2.65e-03 < Clostridia_vadinBB60_group
16	Adenocarcinoma	0.824	0.251	0.753	0.738	Izemoplasmatales	0.002	-0.038	0.000464 < Izemoplasmatales
16	Adenocarcinoma	0.824	0.251	0.753	0.738	Flavobacteriales	0.000	-0.036	4.36e-05 < Flavobacteriales
16	Adenocarcinoma	0.824	0.251	0.753	0.738	Micrococcales	0.000	0.029	Micrococcales <= 7.24e-05
16	Adenocarcinoma	0.824	0.251	0.753	0.738	Oscillospirales	0.135	0.017	0.108 < Oscillospirales <= 0.233
16	Adenocarcinoma	0.824	0.251	0.753	0.738	Pseudomonadales	0.000	0.013	Pseudomonadales <= 5.27e-05
24	Adenocarcinoma	0.824	0.312	0.737	0.721	Izemoplasmatales	0.000	0.039	Izemoplasmatales <= 0.000464
24	Adenocarcinoma	0.824	0.312	0.737	0.721	Flavobacteriales	0.000	-0.036	4.36e-05 < Flavobacteriales
24	Adenocarcinoma	0.824	0.312	0.737	0.721	Micrococcales	0.000	0.031	Micrococcales <= 7.24e-05
24	Adenocarcinoma	0.824	0.312	0.737	0.721	Actinomycetales	0.000	-0.029	0.000341 < Actinomycetales
24	Adenocarcinoma	0.824	0.312	0.737	0.721	Lachnospirales	0.186	-0.021	0.1026 < Lachnospirales
27	Adenocarcinoma	0.866	0.282	0.695	0.786	Flavobacteriales	0.000	0.038	Flavobacteriales <= 4.36e-05
27	Adenocarcinoma	0.866	0.282	0.695	0.786	Izemoplasmatales	0.000	0.036	Izemoplasmatales <= 0.000464
27	Adenocarcinoma	0.866	0.282	0.695	0.786	Micrococcales	0.000	0.029	Micrococcales <= 7.24e-05
27	Adenocarcinoma	0.866	0.282	0.695	0.786	Actinomycetales	0.000	-0.028	0.000341 < Actinomycetales
27	Adenocarcinoma	0.866	0.282	0.695	0.786	Bifidobacteriales	0.141	0.015	0.008930 < Bifidobacteriales
28	Adenocarcinoma	0.788	0.281	0.671	0.817	Izemoplasmatales	0.000	0.041	Izemoplasmatales <= 0.000464
28	Adenocarcinoma	0.788	0.281	0.671	0.817	Flavobacteriales	0.000	0.036	Flavobacteriales <= 4.36e-05
28	Adenocarcinoma	0.788	0.281	0.671	0.817	Micrococcales	0.000	0.032	Micrococcales <= 7.24e-05
28	Adenocarcinoma	0.788	0.281	0.671	0.817	Actinomycetales	0.000	0.021	Actinomycetales <= 0.000136
28	Adenocarcinoma	0.788	0.281	0.671	0.817	Oscillospirales	0.127	0.017	0.108 < Oscillospirales <= 0.233
32	Adenocarcinoma	0.558	0.315	0.734	0.724	Flavobacteriales	0.000	0.036	Flavobacteriales <= 4.36e-05
32	Adenocarcinoma	0.558	0.315	0.734	0.724	Izemoplasmatales	0.000	0.037	Izemoplasmatales <= 0.000464
32	Adenocarcinoma	0.558	0.315	0.734	0.724	Micrococcales	0.000	-0.028	7.24e-05 < Micrococcales
32	Adenocarcinoma	0.558	0.315	0.734	0.724	Actinomycetales	0.001	-0.028	0.000341 < Actinomycetales
32	Adenocarcinoma	0.558	0.315	0.734	0.724	Oscillospirales	0.003	-0.027	Oscillospirales <= 0.108
33	Adenocarcinoma	0.638	0.328	0.735	0.720	Izemoplasmatales	0.000	0.037	Izemoplasmatales <= 0.000464
33	Adenocarcinoma	0.638	0.328	0.735	0.720	Flavobacteriales	0.000	0.036	Flavobacteriales <= 4.36e-05
33	Adenocarcinoma	0.638	0.328	0.735	0.720	Micrococcales	0.000	-0.030	7.24e-05 < Micrococcales
33	Adenocarcinoma	0.638	0.328	0.735	0.720	Actinomycetales	0.001	-0.030	0.000341 < Actinomycetales
33	Adenocarcinoma	0.638	0.328	0.735	0.720	Oscillospirales	0.004	-0.029	Oscillospirales <= 0.108
34	Adenocarcinoma	0.848	0.334	0.774	0.685	Izemoplasmatales	0.001	-0.038	0.000464 < Izemoplasmatales
34	Adenocarcinoma	0.848	0.334	0.774	0.685	Flavobacteriales	0.000	0.037	Flavobacteriales <= 4.36e-05
34	Adenocarcinoma	0.848	0.334	0.774	0.685	Burkholderiales	0.005	-0.031	0.004110 < Burkholderiales
34	Adenocarcinoma	0.848	0.334	0.774	0.685	Micrococcales	0.001	-0.031	7.24e-05 < Micrococcales
34	Adenocarcinoma	0.848	0.334	0.774	0.685	Actinomycetales	0.001	-0.025	0.000341 < Actinomycetales
36	Adenocarcinoma	0.672	0.287	0.740	0.722	Izemoplasmatales	0.000	0.037	Izemoplasmatales <= 0.000464
36	Adenocarcinoma	0.672	0.287	0.740	0.722	Flavobacteriales	0.001	-0.036	4.36e-05 < Flavobacteriales
36	Adenocarcinoma	0.672	0.287	0.740	0.722	Micrococcales	0.000	0.030	Micrococcales <= 7.24e-05

36	Adenocarcinoma	0.672	0.287	0.740	0.722	Bifidobacteriales	0.000	-0.027	Bifidobacteriales <= 0.000582
36	Adenocarcinoma	0.672	0.287	0.740	0.722	Veillonellales.Selenomonadales	0.000	-0.023	Veillonellales.Selenomonadales <= 0.000246
37	Adenocarcinoma	0.690	0.264	0.756	0.727	Izemoplasmatales	0.000	0.039	Izemoplasmatales <= 0.000464
37	Adenocarcinoma	0.690	0.264	0.756	0.727	Flavobacteriales	0.001	-0.036	4.36e-05 < Flavobacteriales
37	Adenocarcinoma	0.690	0.264	0.756	0.727	Burkholderiales	0.004	-0.031	0.004110 < Burkholderiales
37	Adenocarcinoma	0.690	0.264	0.756	0.727	Desulfovibrionales	0.000	-0.016	Desulfovibrionales <= 0.000170
37	Adenocarcinoma	0.690	0.264	0.756	0.727	Oscillospirales	0.156	0.017	0.108 < Oscillospirales <= 0.233
38	Adenocarcinoma	0.730	0.305	0.701	0.761	Flavobacteriales	0.000	0.037	Flavobacteriales <= 4.36e-05
38	Adenocarcinoma	0.730	0.305	0.701	0.761	Izemoplasmatales	0.000	0.038	Izemoplasmatales <= 0.000464
38	Adenocarcinoma	0.730	0.305	0.701	0.761	Micrococcales	0.000	0.031	Micrococcales <= 7.24e-05
38	Adenocarcinoma	0.730	0.305	0.701	0.761	Actinomycetales	0.001	-0.026	0.000341 < Actinomycetales
38	Adenocarcinoma	0.730	0.305	0.701	0.761	Lachnospirales	0.209	-0.020	0.1026 < Lachnospirales
39	Adenocarcinoma	0.786	0.309	0.734	0.724	Izemoplasmatales	0.000	0.038	Izemoplasmatales <= 0.000464
39	Adenocarcinoma	0.786	0.309	0.734	0.724	Flavobacteriales	0.000	0.036	Flavobacteriales <= 4.36e-05
39	Adenocarcinoma	0.786	0.309	0.734	0.724	Micrococcales	0.000	-0.031	7.24e-05 < Micrococcales
39	Adenocarcinoma	0.786	0.309	0.734	0.724	Actinomycetales	0.001	-0.029	0.000341 < Actinomycetales
39	Adenocarcinoma	0.786	0.309	0.734	0.724	Oscillospirales	0.085	-0.024	Oscillospirales <= 0.108
40	Adenocarcinoma	0.704	0.292	0.702	0.762	Izemoplasmatales	0.000	0.039	Izemoplasmatales <= 0.000464
40	Adenocarcinoma	0.704	0.292	0.702	0.762	Flavobacteriales	0.000	0.035	Flavobacteriales <= 4.36e-05
40	Adenocarcinoma	0.704	0.292	0.702	0.762	Micrococcales	0.000	0.031	Micrococcales <= 7.24e-05
40	Adenocarcinoma	0.704	0.292	0.702	0.762	Bifidobacteriales	0.000	-0.022	Bifidobacteriales <= 0.000582
40	Adenocarcinoma	0.704	0.292	0.702	0.762	Veillonellales.Selenomonadales	0.000	-0.022	Veillonellales.Selenomonadales <= 0.000246



Figure ML2-C-6: The LIME Result in Training dataset for Order in Stool

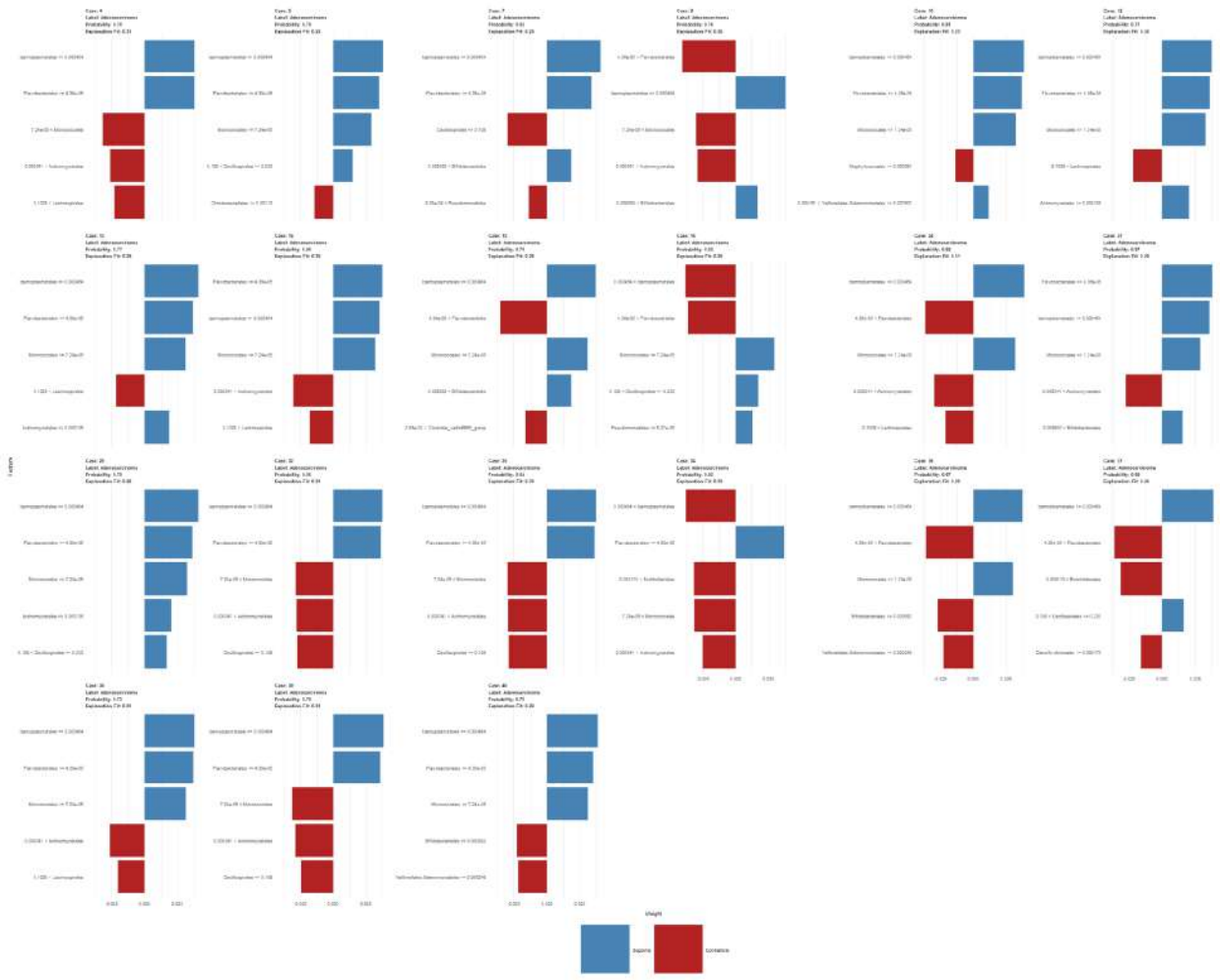


Figure ML2-C-7: The LIME Result in Test dataset for Order in Stool

Table ML2-C-6: The Aggregated SHAP Values

Variable	Training Dataset				Testing Dataset			
	contribution	position	sign	cumulative	contribution	position	sign	cumulative
intercept	0.7602	48	X	0.760	0.7602	48	X	0.7602
Monoglobales	-0.0002	47	-1	0.760	-0.0018	47	-1.0000	0.7584
Lachnospirales	0.0000	46	1	0.760	-0.0034	39	-1.0000	0.7642
Campylobacterales	0.0001	45	1	0.760	-0.0003	43	-1.0000	0.7583
Christensenellales	0.0001	44	1	0.760	-0.0057	44	-1.0000	0.7586

Enterobacterales	0.0000	43	1	0.760	0.0037	41	1.0000	0.7661
Veillonellales.Selenomonadales	-0.0001	42	-1	0.760	0.0041	42	1.0000	0.7624
Pasteurellales	0.0000	41	-1	0.760	0.0014	40	1.0000	0.7675
Chloroplast	0.0000	40	-1	0.760	0.0000	19	1.0000	0.7688
Synergistales	0.0000	39	-1	0.760	-0.0006	38	-1.0000	0.7636
RF39	0.0000	38	1	0.760	0.0000	23	1.0000	0.7685
Bacillales	0.0000	37	1	0.760	0.0000	35	1.0000	0.7634
Tumor	0.0000	36	1	0.760	0.0000	37	1.0000	0.7636
Gastranaerophilales	0.0000	35	-1	0.760	0.0000	28	1.0000	0.7679
DTU014	0.0001	34	1	0.760	0.0010	34	1.0000	0.7645
Desulfovibrionales	0.000	33	-1	0.760	0.0059	32	1.0000	0.7708
Lactobacillales	0.000	32	-1	0.760	-0.0001	30	-1.0000	0.7679
Oscillospirales	0.000	31	-1	0.760	0.0004	33	1.0000	0.7649
Rhizobiales	0.000	30	-1	0.760	0.0000	29	-1.0000	0.7679
Burkholderiales	0.000	29	1	0.760	0.0001	6	1.0000	0.7667
Fusobacteriales	0.000	28	-1	0.760	-0.0003	15	-1.0000	0.7745
Pseudomonadales	0.000	27	1	0.760	0.0011	24	1.0000	0.7685
Peptococcales	0.000	26	-1	0.760	-0.0011	25	-1.0000	0.7673
Aeromonadales	0.000	25	1	0.760	-0.0014	12	-1.0000	0.7704
Eubacteriales	0.000	24	1	0.760	-0.0015	14	-1.0000	0.7731
Rhodospirillales	0.000	23	1	0.760	-0.0008	7	-1.0000	0.7666
Clostridia_vadinBB60_group	0.000	22	-1	0.760	-0.0001	22	-1.0000	0.7684
Bacteroidales	0.000	21	1	0.760	-0.0013	13	-1.0000	0.7718
Bifidobacteriales	0.000	20	1	0.760	0.0056	45	1.0000	0.7643
Verrucomicrobiales	0.000	19	-1	0.760	-0.0017	18	-1.0000	0.7671
Clostridia_UCG.014	0.000	18	1	0.760	-0.0011	16	-1.0000	0.7748
Izemoplasmatales	0.000	17	1	0.760	0.0088	17	1.0000	0.7759
Gender	0.000	16	1	0.760	0.0000	11	1.0000	0.7704
Saccharimonadales	0.000	15	-1	0.760	0.0003	46	1.0000	0.7587
Flavobacteriales	0.000	14	1	0.760	0.0010	10	1.0000	0.7714
Methanobacteriales	0.000	13	-1	0.760	-0.0010	21	-1.0000	0.7674
Clostridiales	0.000	12	1	0.760	0.0016	26	1.0000	0.7685
Staphylococcales	0.000	11	1	0.760	0.0001	9	1.0000	0.7715
Micrococcales	0.000	10	1	0.760	-0.0041	8	-1.0000	0.7674
Clostridia	0.000	9	-1	0.760	-0.0014	5	-1.0000	0.7653
Peptostreptococcales.Tissierellales	0.000	8	1	0.760	0.0014	20	1.0000	0.7688
Victivallales	0.000	7	-1	0.760	-0.0002	36	-1.0000	0.7634
Erysipelotrichales	0.000	6	1	0.760	-0.0028	31	-1.0000	0.7680
Coriobacteriales	0.000	5	-1	0.760	0.0001	4	1.0000	0.7654
Rickettsiales	0.000	4	1	0.760	-0.0011	27	-1.0000	0.7668
Acidaminococcales	0.000	3	-1	0.760	-0.0006	3	-1.0000	0.7648
Actinomycetales	0.000	2	-1	0.760	-0.0018	2	-1.0000	0.7630
prediction	0.760	1	X	0.760	0.7630	1	X	0.7630

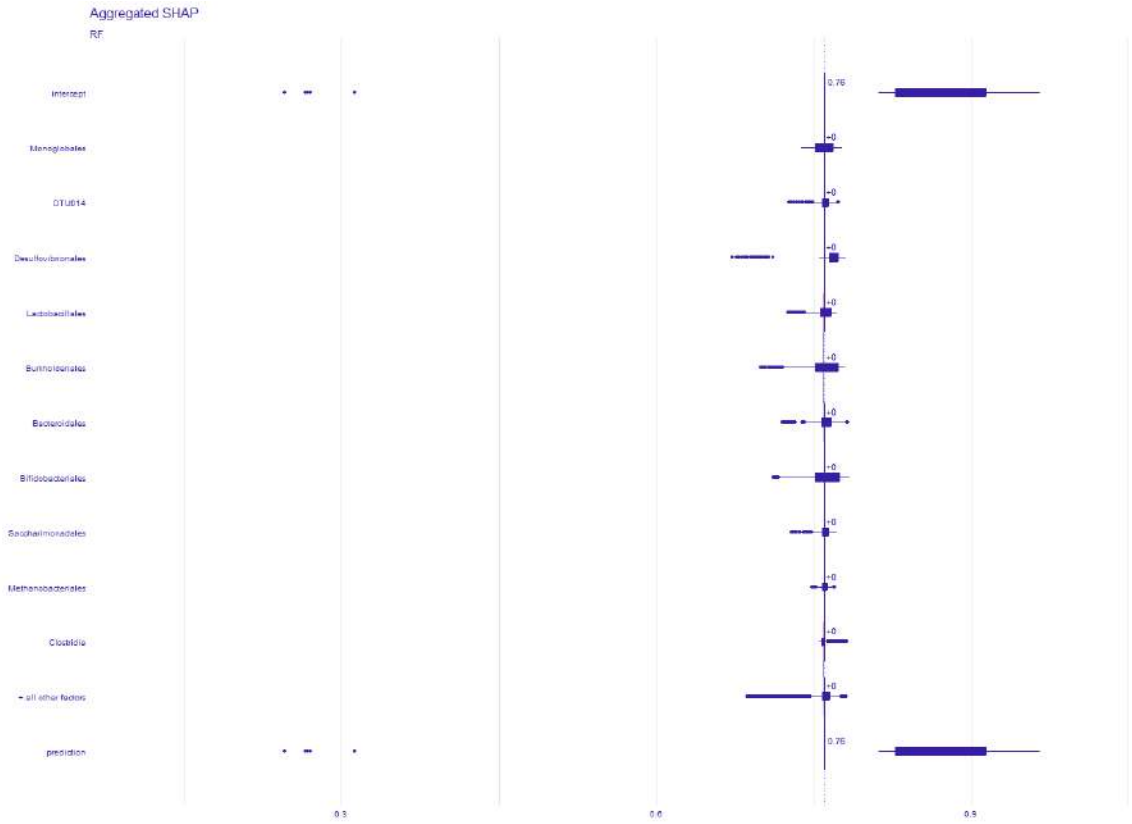


Figure ML2-C-8: The Aggregated SHAP Value in Training dataset for Order in Stool

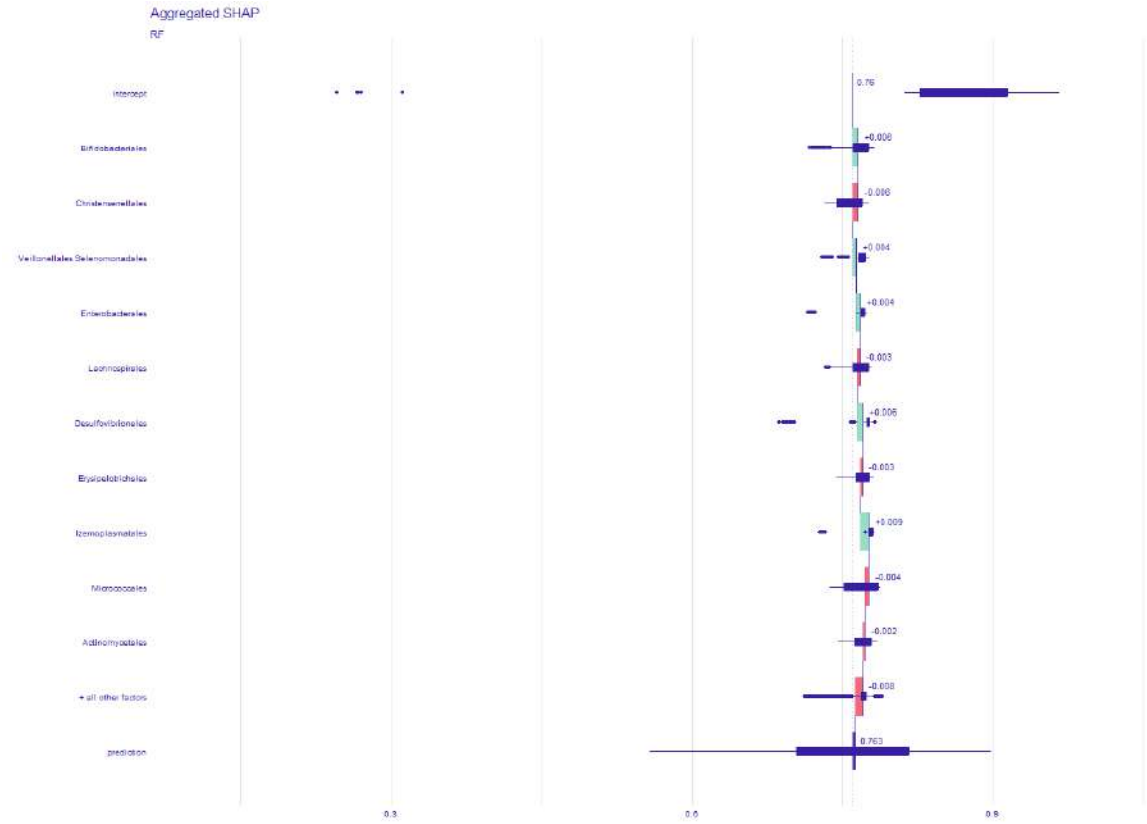


Figure ML2-C-9: The Aggregated SHAP Value in Test dataset for Order in Stool

D- Family

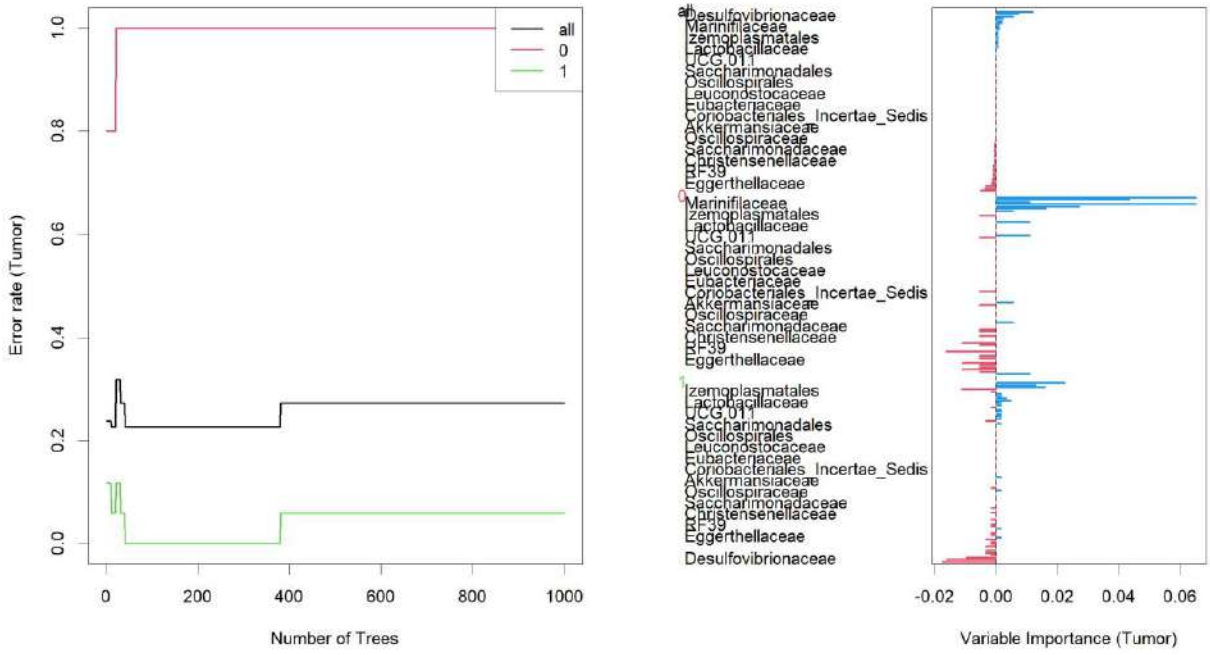
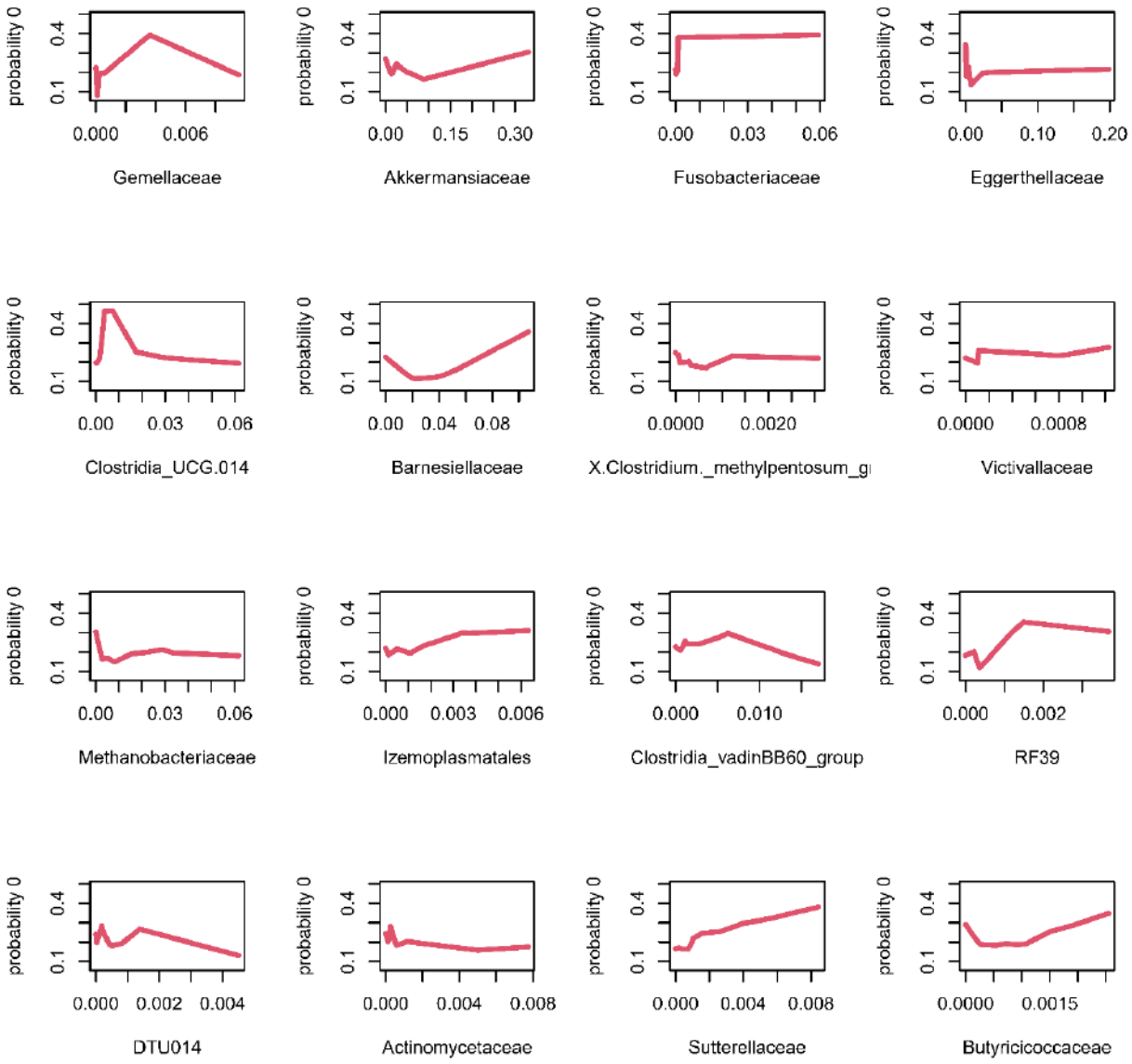
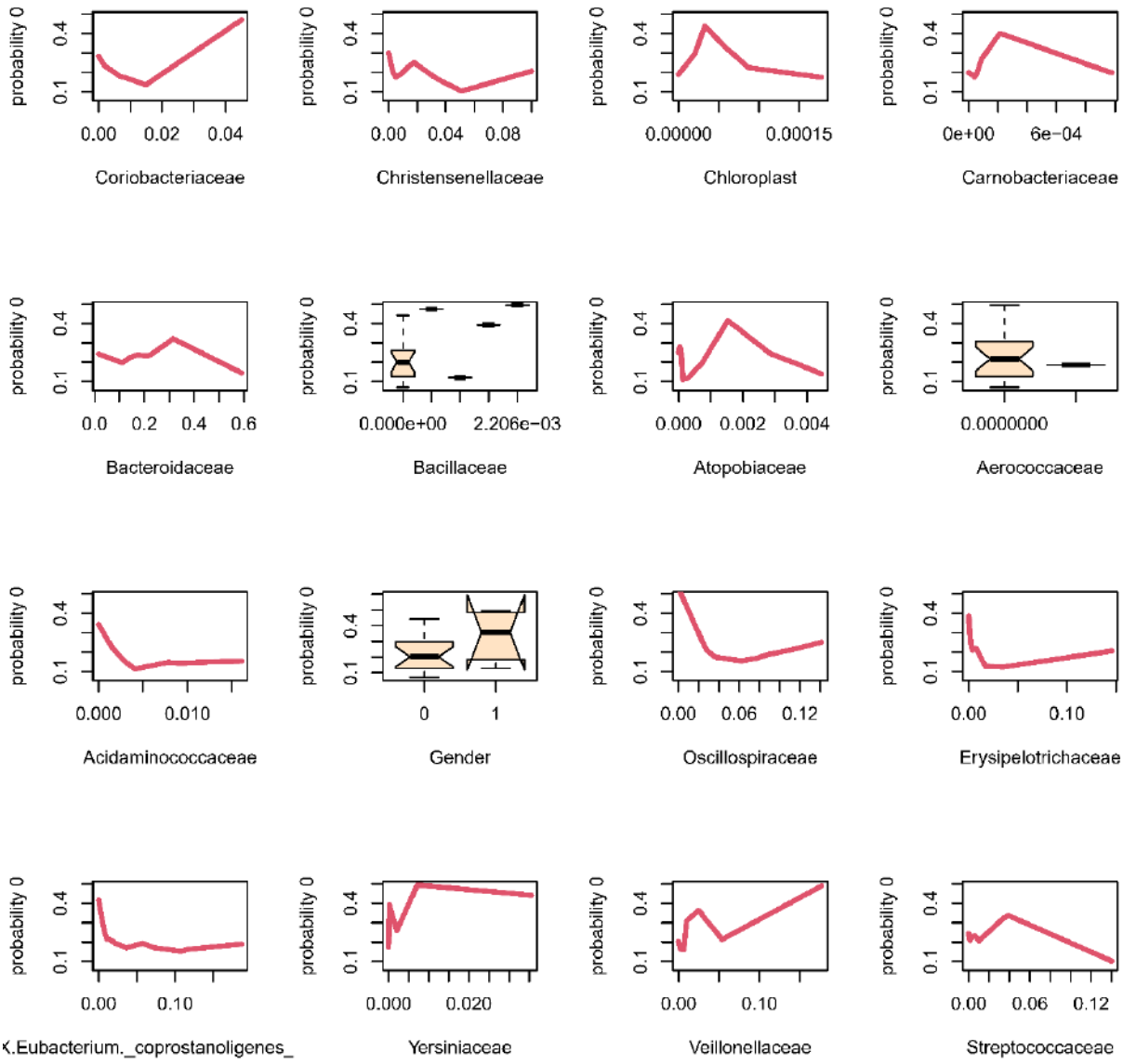
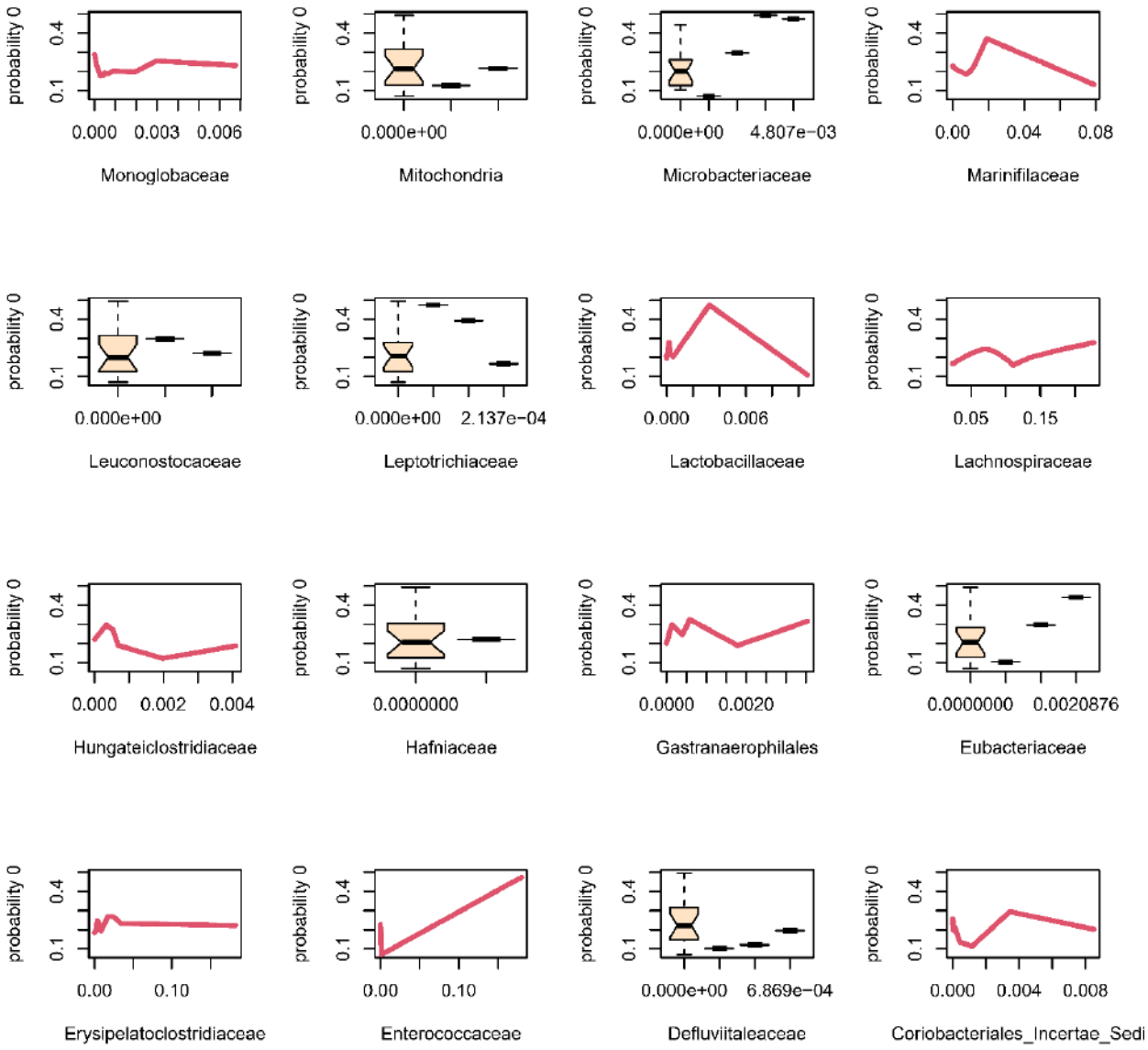
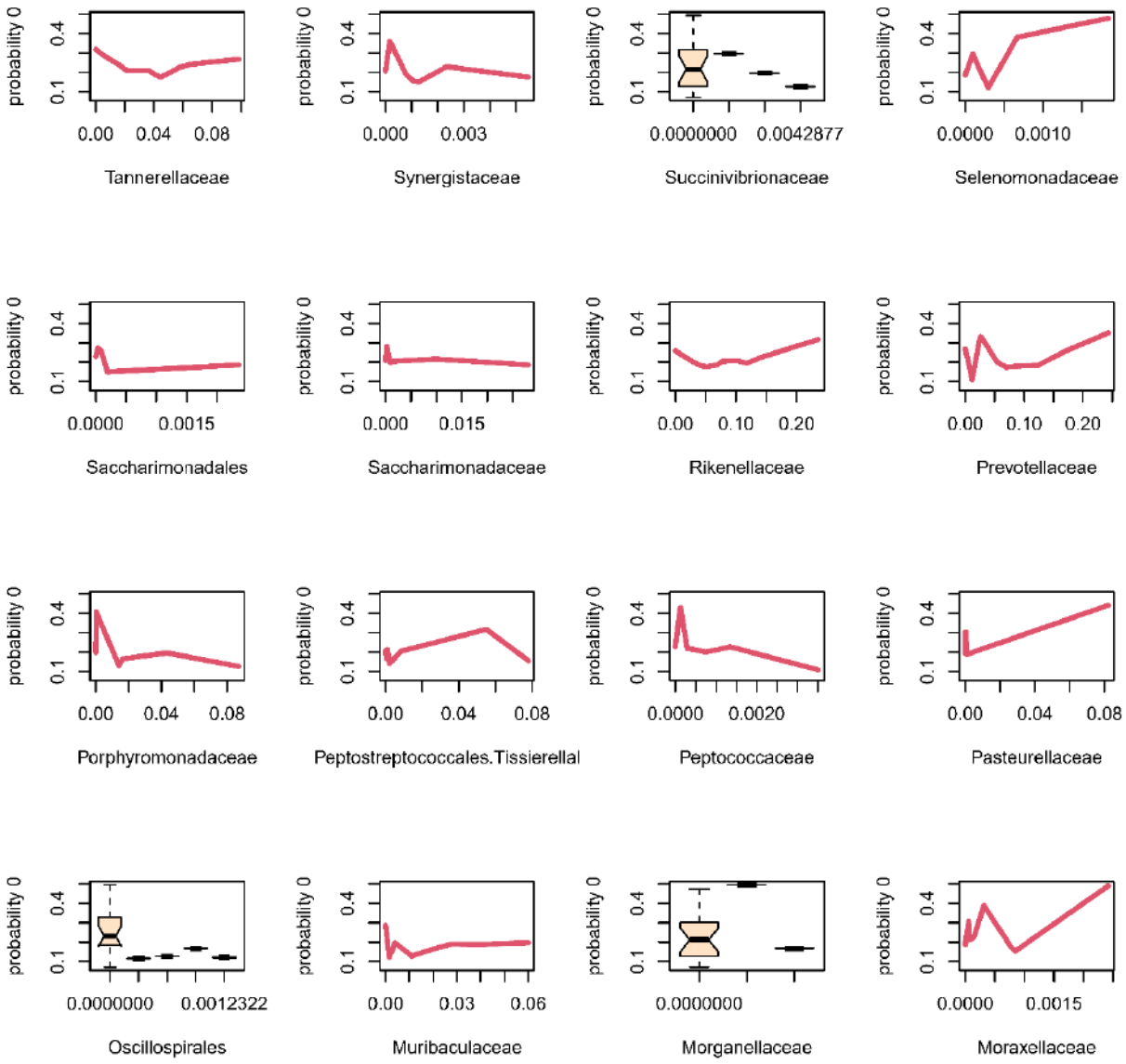


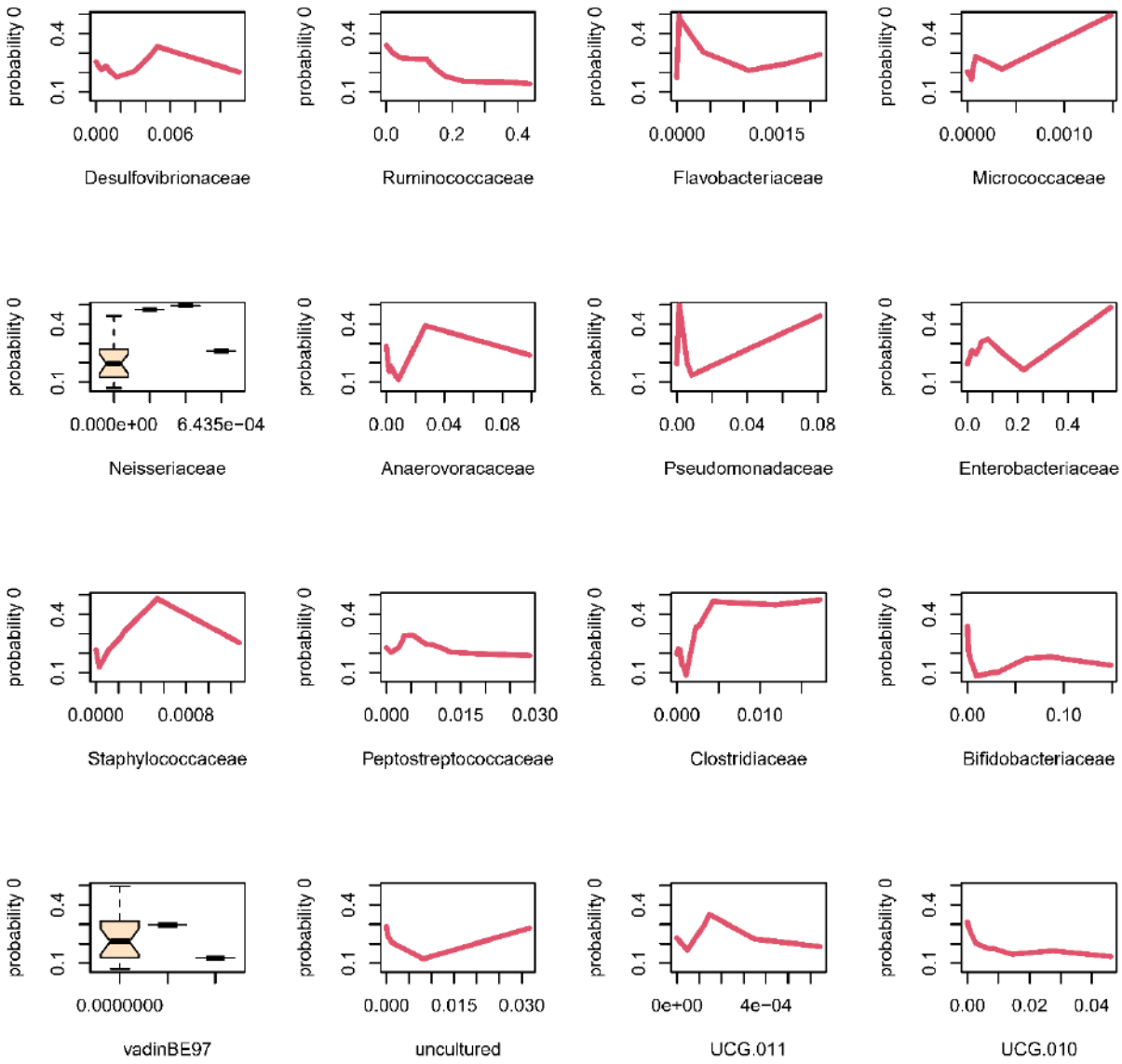
Figure ML2-D-1: The number of trees and Variable Importance Plot for Family in Stool


probability 0
0.1
0.4
0.000
0.006
Gemellaceae
probability 0
0.1
0.4
0.00
0.15
0.30
Akkermansiaceae
probability 0
0.1
0.4
0.00
0.03
0.06
Fusobacteriaceae
probability 0
0.1
0.4
0.00
0.10
0.20
Eggerthellaceae
probability 0
0.1
0.4
0.00
0.03
0.06
Clostridia_UCG.014
probability 0
0.1
0.4
0.00
0.04
0.08
Barnesiellaceae
probability 0
0.1
0.4
0.0000
0.0020
X.Clostridium._methylpentosum_gr
probability 0
0.1
0.4
0.0000
0.0008
Victivallaceae
probability 0
0.1
0.4
0.00
0.03
0.06
Methanobacteriaceae
probability 0
0.1
0.4
0.000
0.003
0.006
Izemoplasmatales
probability 0
0.1
0.4
0.000
0.010
Clostridia_vadinBB60_group
probability 0
0.1
0.4
0.000
0.002
RF39
probability 0
0.1
0.4
0.000
0.002
0.004
DTU014
probability 0
0.1
0.4
0.000
0.004
0.008
Actinomycetaceae
probability 0
0.1
0.4
0.000
0.004
0.008
Sutterellaceae
probability 0
0.1
0.4
0.0000
0.0015
Butyricicoccaceae


probability 0
0.1
0.4
0.00 0.02 0.04
Coriobacteriaceae
0.00 0.04 0.08
Christensenellaceae
0.00000 0.00015
Chloroplast
0e+00 6e−04
Carnobacteriaceae
0.0 0.2 0.4 0.6
Bacteroidaceae
0.000e+00 2.206e−03
Bacillaceae
0.000 0.002 0.004
Atopobiaceae
0.0000000
Aerococcaceae
0.000 0.010
Acidaminococcaceae
0 1
Gender
0.00 0.06 0.12
Oscillospiraceae
0.00 0.10
Erysipelotrichaceae
0.00 0.10
K.Eubacterium._coprostanoligenes_
0.000 0.020
Yersiniaceae
0.00 0.10
Veillonellaceae
0.00 0.06 0.12
Streptococcaceae


probability 0
0.1
0.4
0.000
0.003
0.006
Monoglobaceae
0.000e+00
Mitochondria
0.000e+00
4.807e−03
Microbacteriaceae
0.00
0.04
0.08
Marinifilaceae
0.000e+00
Leuconostocaceae
0.000e+00
2.137e−04
Leptotrichiaceae
0.000
0.006
Lactobacillaceae
0.05
0.15
Lachnospiraceae
0.000
0.002
0.004
Hungateiclostridiaceae
probability 0
0.0000000
Hafniaceae
0.0000
0.0020
Gastranaerophilales
0.0000000
0.0020876
Eubacteriaceae
0.00
0.10
Erysipelatoclostridiaceae
0.00
0.10
Enterococcaceae
0.000e+00
6.869e−04
Defluviitaleaceae
0.000
0.004
0.008
Coriobacteriales_Incertae_Sedi


probability 0
0.1
0.4
0.00 0.04 0.08
Tannerellaceae
0.000 0.003
Synergistaceae
0.0000000 0.0042877
Succinivibrionaceae
0.0000 0.0010
Selenomonadaceae
0.0000 0.0015
Saccharimonadales
0.000 0.015
Saccharimonadaceae
0.00 0.10 0.20
Rikenellaceae
Prevotellaceae
0.00 0.04 0.08
Porphyromonadaceae
probability
Peptostreptococcales.Tissierellal
0.0000 0.0020
Peptococcaceae
Pasteurellaceae
0.0000000 0.0012322
Oscillospirales
0.00 0.03 0.06
Muribaculaceae
0.0000000
Morganellaceae
Moraxellaceae


probability 0
0.1
0.4
0.000
0.006
Desulfovibrionaceae
0.0
0.2
0.4
Ruminococcaceae
0.0000
0.0015
Flavobacteriaceae
0.0000
0.0010
Micrococcaceae
0.000e+00
6.435e-04
Neisseriaceae
0.00
0.04
0.08
Anaerovoracaceae
Pseudomonadaceae
Enterobacteriaceae
0.0000
0.0008
Staphylococcaceae
probability
0.000
0.015
0.030
Peptostreptococcaceae
0.010
Clostridiaceae
0.10
Bifidobacteriaceae
0.0000000
vadinBE97
uncultured
0e+00
4e-04
UCG.011
0.02
0.04
UCG.010

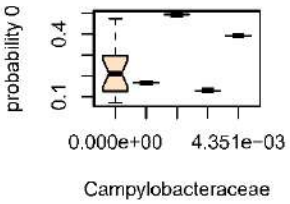
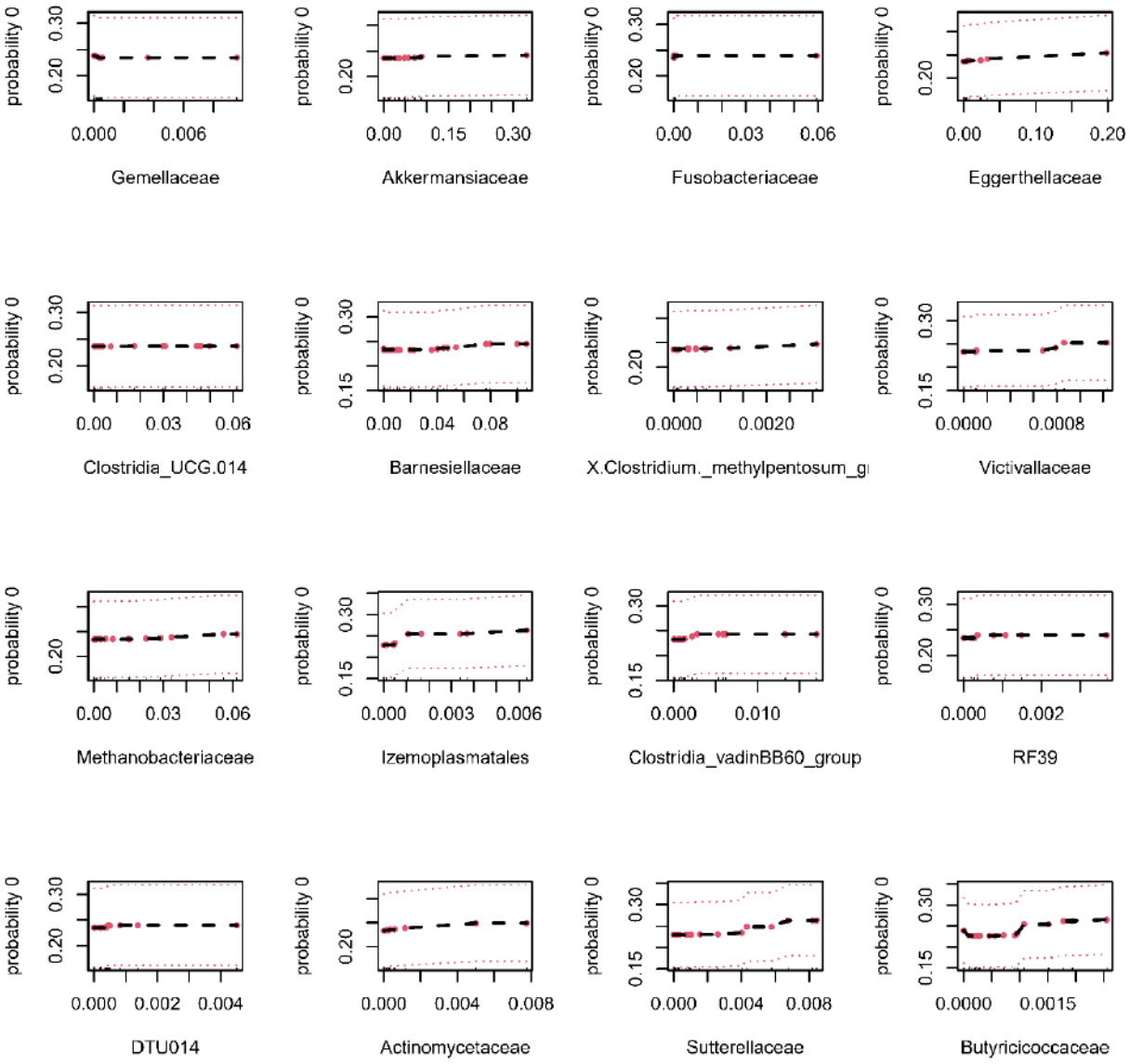
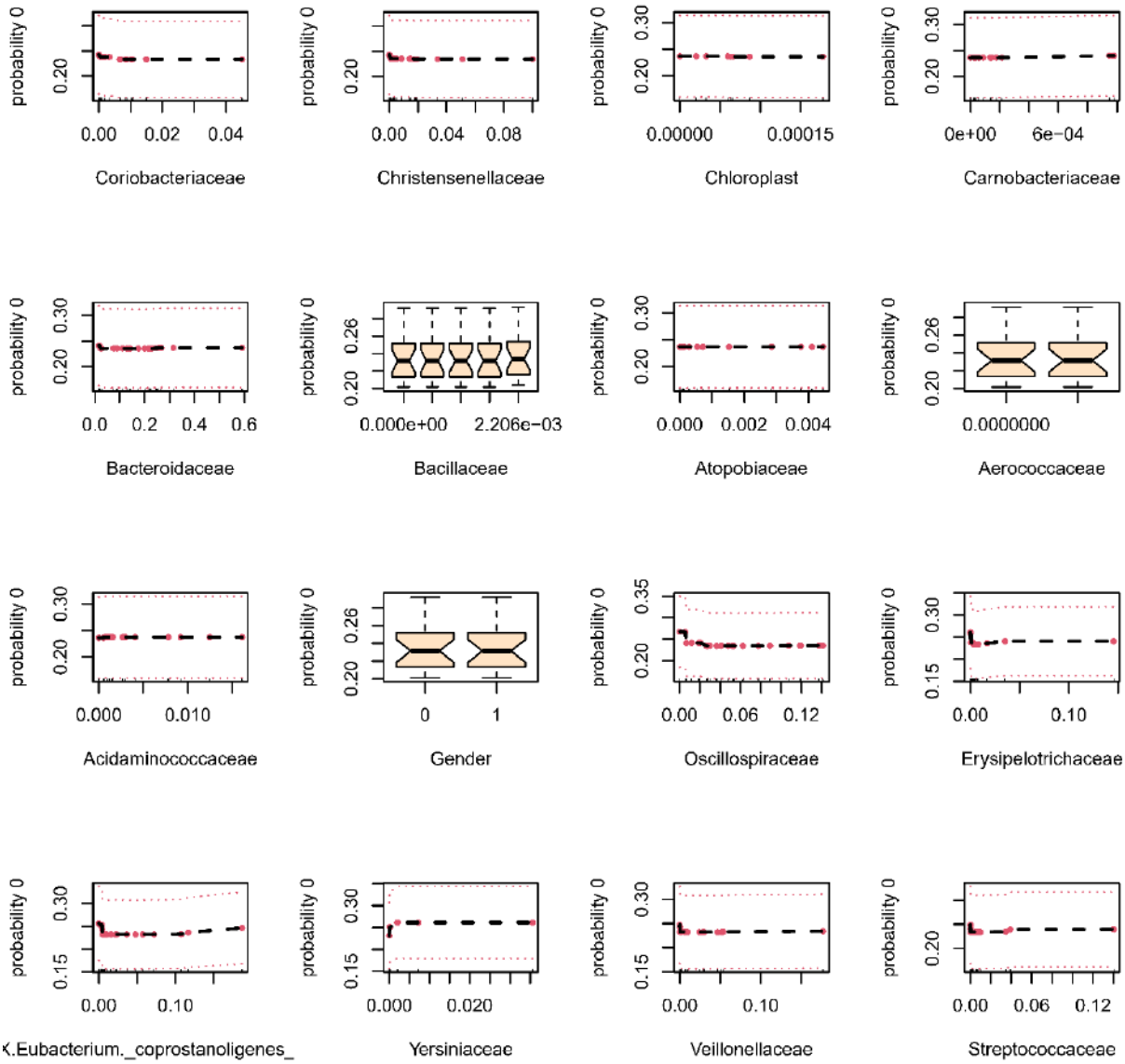
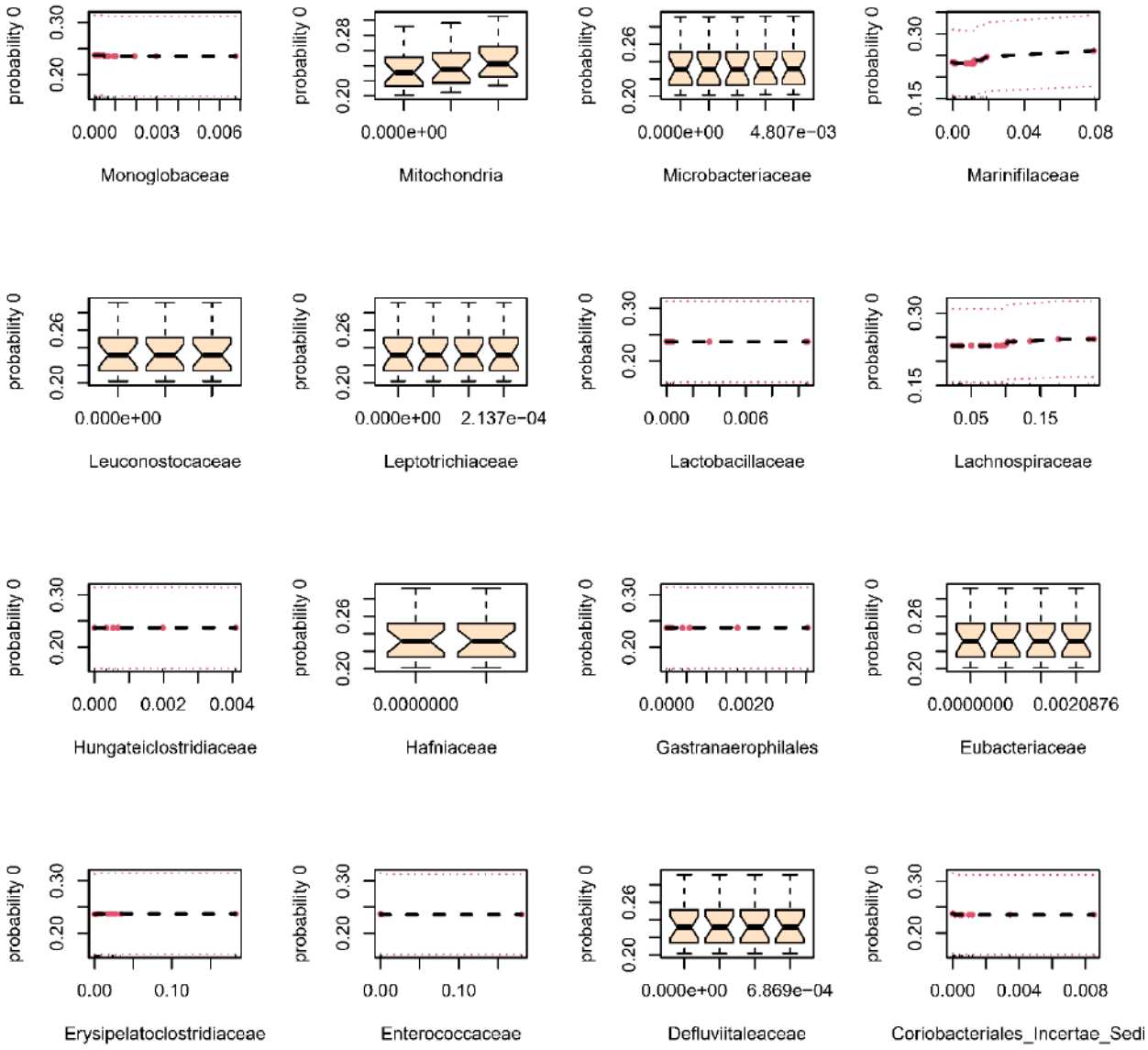
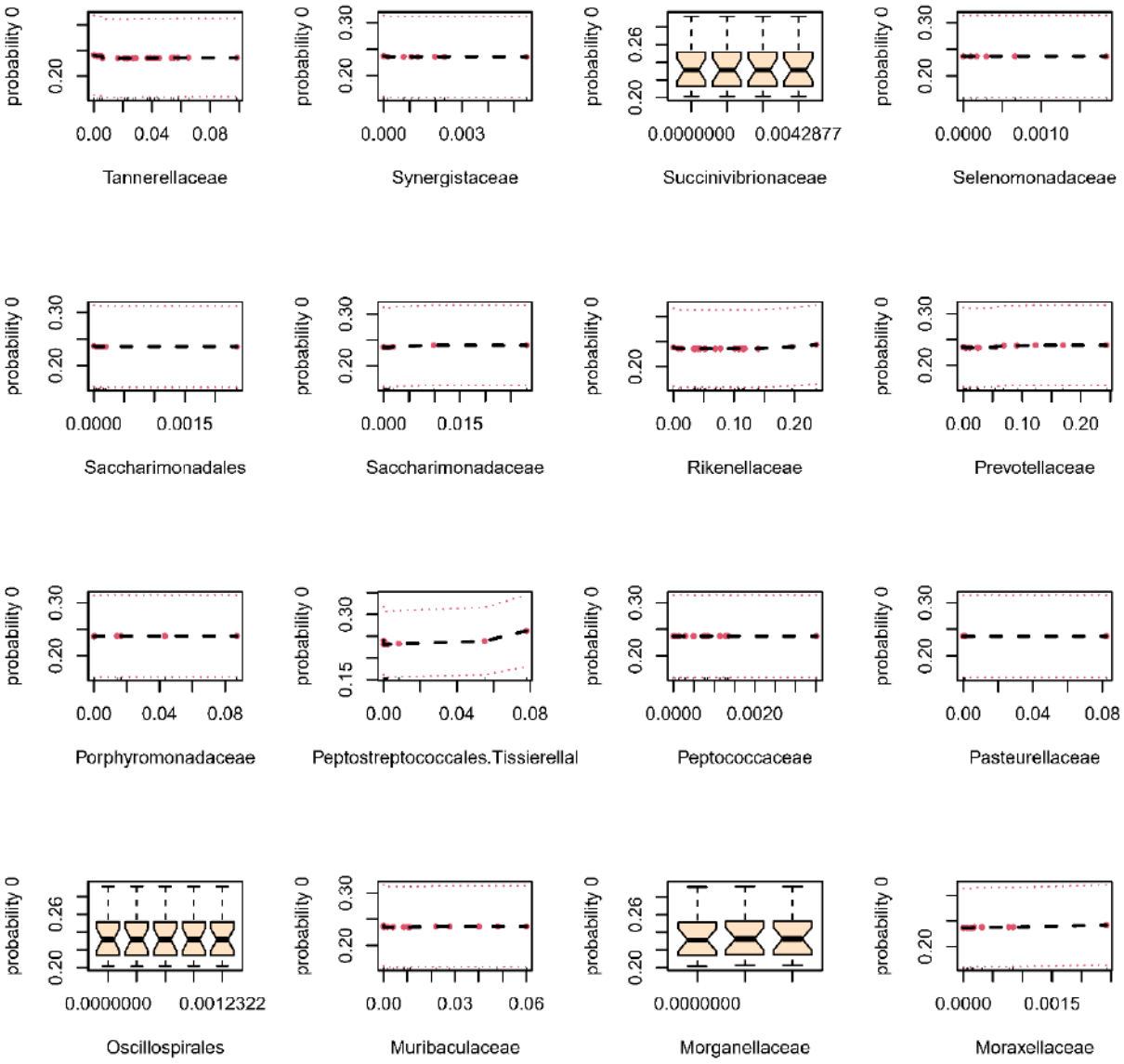


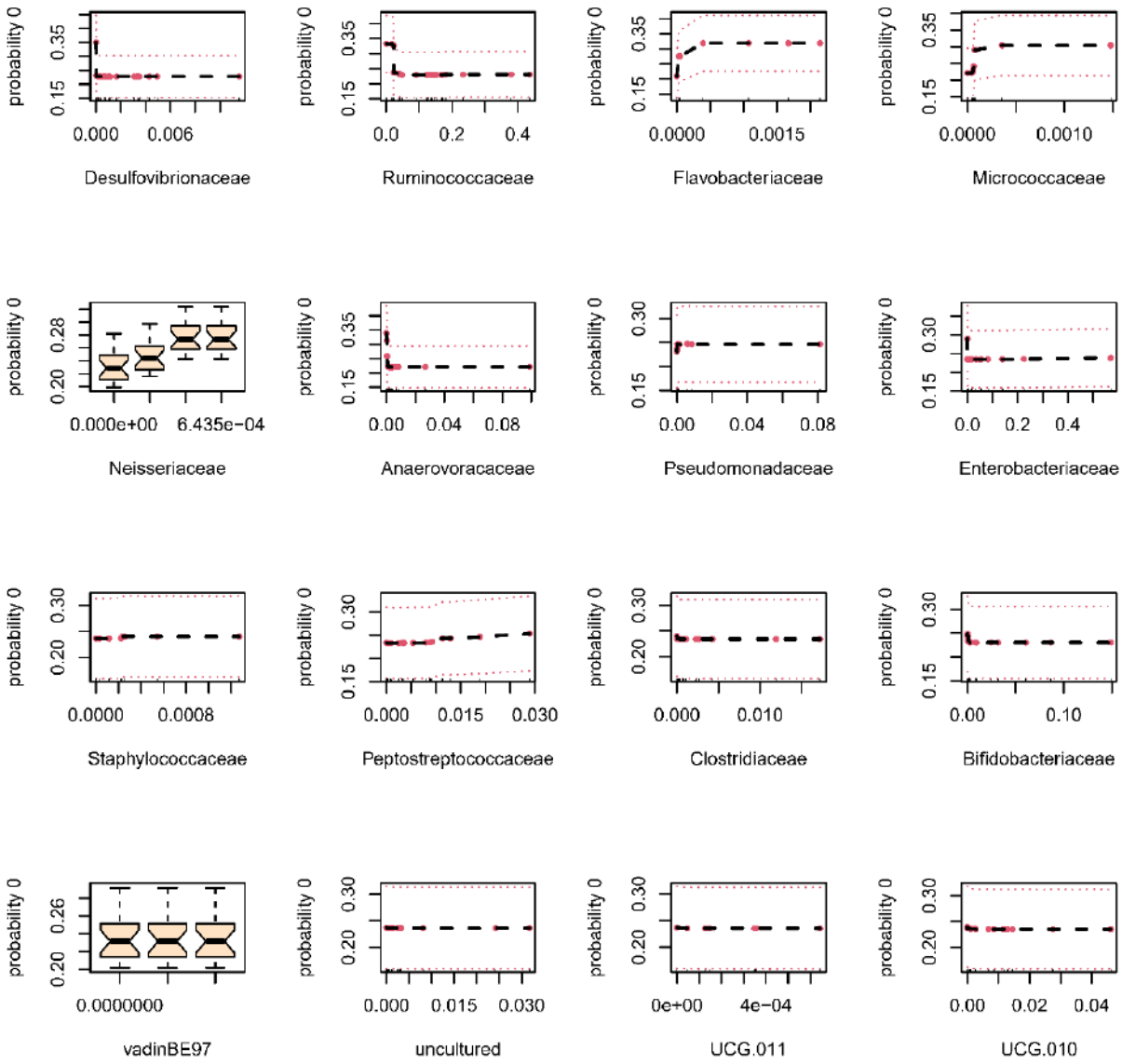
Figure ML2-D-2: Marginal Plot for Family in Stool


probability 0
Gemellaceae
Akkermansiaceae
Fusobacteriaceae
Eggerthellaceae
Clostridia_UCG.014
Barnesiellaceae
X.Clostridium._methylpentosum_gr
Victivallaceae
Methanobacteriaceae
Izemoplasmatales
Clostridia_vadinBB60_group
RF39
DTU014
Actinomycetaceae
Sutterellaceae
Butyricicoccaceae


probability 0
Coriobacteriaceae
Christensenellaceae
Chloroplast
Carnobacteriaceae
Bacteroidaceae
Bacillaceae
Atopobiaceae
Aerococcaceae
Acidaminococcaceae
Gender
Oscillospiraceae
Erysipelotrichaceae
K.Eubacterium._coprostanoligenes_
Yersiniaceae
Veillonellaceae
Streptococcaceae


probability 0
Monoglobaceae
Mitochondria
Microbacteriaceae
Marinifilaceae
Leuconostocaceae
Leptotrichiaceae
Lactobacillaceae
Lachnospiraceae
Hungateiclostridiaceae
Hafniaceae
Gastranaerophilales
Eubacteriaceae
Erysipelatoclostridiaceae
Enterococcaceae
Defluviitaleaceae
Coriobacteriales_Incertae_Sedi


probability 0
0.20
0.00 0.04 0.08
Tannerellaceae
probability 0
0.20 0.30
0.000 0.003
Synergistaceae
probability 0
0.20 0.26
0.0000000 0.0042877
Succinivibrionaceae
probability 0
0.20 0.30
0.0000 0.0010
Selenomonadaceae
probability 0
0.20 0.30
0.0000 0.0015
Saccharimonadales
probability 0
0.20 0.30
0.000 0.015
Saccharimonadaceae
probability 0
0.20
0.00 0.10 0.20
Rikenellaceae
probability 0
0.20 0.30
0.00 0.10 0.20
Prevotellaceae
probability 0
0.20 0.30
0.00 0.04 0.08
Porphyromonadaceae
probability 0
0.15 0.30
0.00 0.04 0.08
Peptostreptococcales.Tissierellal
probability 0
0.20 0.30
0.0000 0.0020
Peptococcaceae
probability 0
0.20 0.30
0.00 0.04 0.08
Pasteurellaceae
probability 0
0.20 0.26
0.0000000 0.0012322
Oscillospirales
probability 0
0.20 0.30
0.00 0.03 0.06
Muribaculaceae
probability 0
0.20 0.26
0.0000000
Morganellaceae
probability 0
0.20
0.0000 0.0015
Moraxellaceae


probability 0
0.15
0.35
0.000
0.006
Desulfovibrionaceae
probability 0
0.15
0.35
0.0
0.2
0.4
Ruminococcaceae
probability 0
0.15
0.35
0.0000
0.0015
Flavobacteriaceae
probability 0
0.15
0.35
0.0000
0.0010
Micrococcaceae
probability 0
0.20
0.28
0.000e+00
6.435e−04
Neisseriaceae
probability 0
0.15
0.35
0.00
0.04
0.08
Anaerovoracaceae
probability 0
0.15
0.30
0.00
0.04
0.08
Pseudomonadaceae
probability 0
0.15
0.30
0.0
0.2
0.4
Enterobacteriaceae
probability 0
0.20
0.30
0.0000
0.0008
Staphylococcaceae
probability 0
0.15
0.30
0.000
0.015
0.030
Peptostreptococcaceae
probability 0
0.20
0.30
0.000
0.010
Clostridiaceae
probability 0
0.15
0.30
0.00
0.10
Bifidobacteriaceae
probability 0
0.20
0.26
0.0000000
vadinBE97
probability 0
0.20
0.30
0.000
0.015
0.030
uncultured
probability 0
0.20
0.30
0e+00
4e−04
UCG.011
probability 0
0.20
0.30
0.00
0.02
0.04
UCG.010

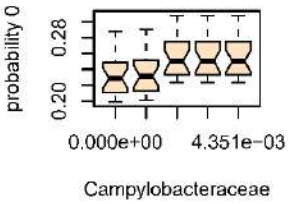


Figure ML2-D-3: Partial Plot for Family in Stool

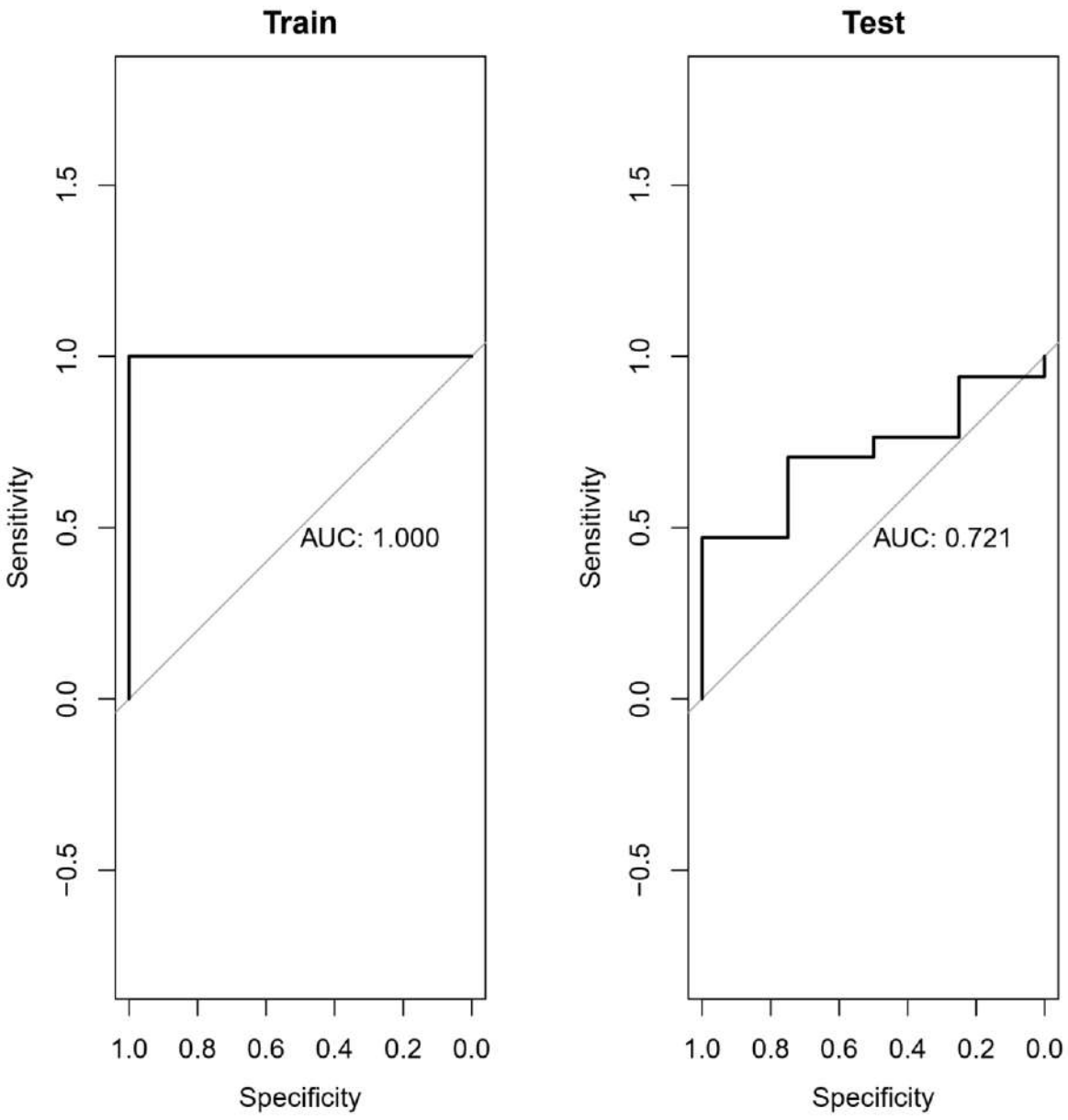


Figure ML2-D-4: The ROC Curve for Family in Stool

Table ML2-D-1: The Model Performance

Index	Values		
	2.50%	50.00%	97.50%
Accuracy	0.43	0.69	0.90
Specificity	0.50	1.00	1.00
Sensitivity	0.29	0.65	0.94

Table ML2-D-2: The Variable Importance Values

Covariates	Variable Importance			Brier Score		
	all	AP	CRC	all	AP	CRC
Gender	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
X.Clostridium._methylpentosum_group	0.0023	-0.0217	0.0144	0.0114	0.0078	0.0078
X.Eubacterium._coprostanoligenes_group	0.0037	-0.0435	0.0256	0.0046	0.0031	0.0031
Acidaminococcaceae	-0.0014	0.0000	-0.0048	-0.0028	-0.0019	-0.0019
Actinomycetaceae	-0.0014	-0.0489	0.0096	-0.0034	-0.0023	-0.0023
Aerococcaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Akkermansiaceae	-0.0005	-0.0109	0.0016	-0.0038	-0.0026	-0.0026
Anaerovoracaceae	0.0790	-0.0326	0.2846	0.1859	0.1264	0.1264
Atopobiaceae	-0.0009	-0.0054	-0.0016	-0.0021	-0.0014	-0.0014
Bacillaceae	0.0000	0.0000	0.0000	-0.0004	-0.0003	-0.0003
Bacteroidaceae	-0.0010	-0.0217	0.0032	-0.0024	-0.0016	-0.0016
Barnesiellaceae	-0.0032	-0.0217	-0.0048	-0.0070	-0.0047	-0.0047
Bifidobacteriaceae	-0.0008	-0.0163	0.0016	-0.0076	-0.0052	-0.0052
Butyricicoccaceae	0.0169	-0.0163	0.0640	0.0230	0.0156	0.0156
Campylobacteraceae	0.0027	-0.0326	0.0192	-0.0127	-0.0087	-0.0087
Carnobacteriaceae	0.0027	-0.0163	0.0144	0.0006	0.0004	0.0004
Chloroplast	0.0000	-0.0054	0.0016	-0.0003	-0.0002	-0.0002
Christensenellaceae	0.0009	-0.0109	0.0064	-0.0014	-0.0010	-0.0010
Clostridia_UCG.014	0.0000	0.0000	0.0000	-0.0003	-0.0002	-0.0002
Clostridia_vadinBB60_group	-0.0010	0.0000	-0.0032	-0.0034	-0.0023	-0.0023
Clostridiaceae	0.0005	0.0000	0.0016	0.0021	0.0014	0.0014
Coriobacteriaceae	-0.0032	-0.0054	-0.0096	-0.0084	-0.0057	-0.0057
Coriobacteriales_Incertae_Sedis	-0.0023	0.0000	-0.0080	-0.0023	-0.0015	-0.0015
Defluviitaleaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Desulfovibrionaceae	0.0942	-0.1414	0.3678	0.2770	0.1883	0.1883
DTU014	-0.0023	-0.0163	-0.0032	-0.0007	-0.0005	-0.0005
Eggerthellaceae	0.0032	-0.0217	0.0176	0.0114	0.0077	0.0077
Enterobacteriaceae	0.0156	-0.0652	0.0720	0.0376	0.0256	0.0256
Enterococcaceae	-0.0009	0.0054	-0.0048	-0.0035	-0.0024	-0.0024
Erysipelatoclostridiaceae	-0.0023	-0.0054	-0.0064	-0.0058	-0.0039	-0.0039
Erysipelotrichaceae	0.0077	-0.0326	0.0368	0.0060	0.0041	0.0041

Eubacteriaceae	-0.0005	0.0000	-0.0016	-0.0016	-0.0011	-0.0011
Flavobacteriaceae	0.0515	-0.0652	0.1983	0.1872	0.1272	0.1272
Fusobacteriaceae	0.0000	0.0000	0.0000	-0.0017	-0.0011	-0.0011
Gastranaerophilales	-0.0005	0.0000	-0.0016	-0.0013	-0.0009	-0.0009
Gemellaceae	0.0000	0.0054	-0.0016	-0.0020	-0.0014	-0.0014
Hafniaceae	0.0000	0.0000	0.0000	-0.0018	-0.0012	-0.0012
Hungateiclostridiaceae	0.0000	0.0000	0.0000	0.0002	0.0002	0.0002
Izemoplasmatales	0.0078	-0.0381	0.0384	0.0089	0.0061	0.0061
Lachnospiraceae	0.0008	-0.0326	0.0128	-0.0069	-0.0047	-0.0047
Lactobacillaceae	0.0010	0.0000	0.0032	0.0011	0.0007	0.0007
Leptotrichiaceae	0.0005	0.0000	0.0016	0.0012	0.0008	0.0008
Leuconostocaceae	0.0000	0.0109	-0.0032	-0.0003	-0.0002	-0.0002
Marinifilaceae	0.0037	-0.0326	0.0224	0.0118	0.0080	0.0080
Methanobacteriaceae	-0.0005	-0.0272	0.0064	0.0045	0.0031	0.0031
Microbacteriaceae	0.0005	0.0000	0.0016	-0.0002	-0.0001	-0.0001
Micrococcaceae	0.0379	-0.0544	0.1487	0.0842	0.0572	0.0572
Mitochondria	0.0010	-0.0109	0.0064	0.0046	0.0031	0.0031
Monoglobaceae	-0.0004	0.0000	-0.0016	-0.0039	-0.0027	-0.0027
Moraxellaceae	-0.0009	-0.0054	-0.0016	-0.0045	-0.0030	-0.0030
Morganellaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Muribaculaceae	0.0014	0.0054	0.0032	0.0009	0.0006	0.0006
Neisseriaceae	0.0037	-0.0598	0.0304	0.0079	0.0054	0.0054
Oscillospiraceae	0.0143	-0.0489	0.0640	0.0186	0.0127	0.0127
Oscillospirales	0.0000	0.0000	0.0000	0.0002	0.0002	0.0002
Pasteurellaceae	-0.0009	0.0000	-0.0032	-0.0022	-0.0015	-0.0015
Peptococcaceae	-0.0010	0.0000	-0.0032	-0.0013	-0.0009	-0.0009
Peptostreptococcaceae	0.0005	-0.0598	0.0192	-0.0027	-0.0018	-0.0018
Peptostreptococcales.Tissierellales	0.0083	-0.0544	0.0448	0.0227	0.0154	0.0154
Porphyromonadaceae	-0.0005	0.0000	-0.0016	-0.0014	-0.0010	-0.0010
Prevotellaceae	0.0000	0.0000	0.0000	0.0004	0.0003	0.0003
Pseudomonadaceae	-0.0041	-0.0326	-0.0048	-0.0163	-0.0111	-0.0111
RF39	-0.0014	0.0000	-0.0048	-0.0005	-0.0004	-0.0004
Rikenellaceae	0.0037	-0.0054	0.0144	-0.0013	-0.0009	-0.0009
Ruminococcaceae	0.0645	-0.0544	0.2398	0.1968	0.1337	0.1337
Saccharimonadaceae	0.0005	0.0000	0.0016	0.0015	0.0010	0.0010
Saccharimonadales	0.0018	0.0000	0.0064	0.0014	0.0009	0.0009
Selenomonadaceae	0.0005	0.0054	0.0000	0.0013	0.0009	0.0009
Staphylococcaceae	0.0000	-0.0054	0.0016	0.0013	0.0009	0.0009
Streptococcaceae	0.0013	-0.0652	0.0240	0.0018	0.0012	0.0012
Succinivibrionaceae	0.0000	0.0054	-0.0016	-0.0008	-0.0005	-0.0005
Sutterellaceae	-0.0015	-0.0652	0.0144	0.0098	0.0067	0.0067
Synergistaceae	-0.0014	0.0000	-0.0048	0.0003	0.0002	0.0002
Tannerellaceae	0.0000	0.0000	0.0000	-0.0009	-0.0006	-0.0006

UCG.010	0.0005	0.0109	-0.0016	-0.0025	-0.0017	-0.0017
UCG.011	-0.0009	0.0000	-0.0032	-0.0018	-0.0012	-0.0012
uncultured	0.0005	0.0000	0.0016	-0.0010	-0.0007	-0.0007
vadinBE97	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Veillonellaceae	-0.0010	-0.0598	0.0144	-0.0049	-0.0033	-0.0033
Victivallaceae	0.0037	-0.0109	0.0160	-0.0017	-0.0011	-0.0011
Yersiniaceae	0.0014	-0.0054	0.0064	-0.0067	-0.0046	-0.0046

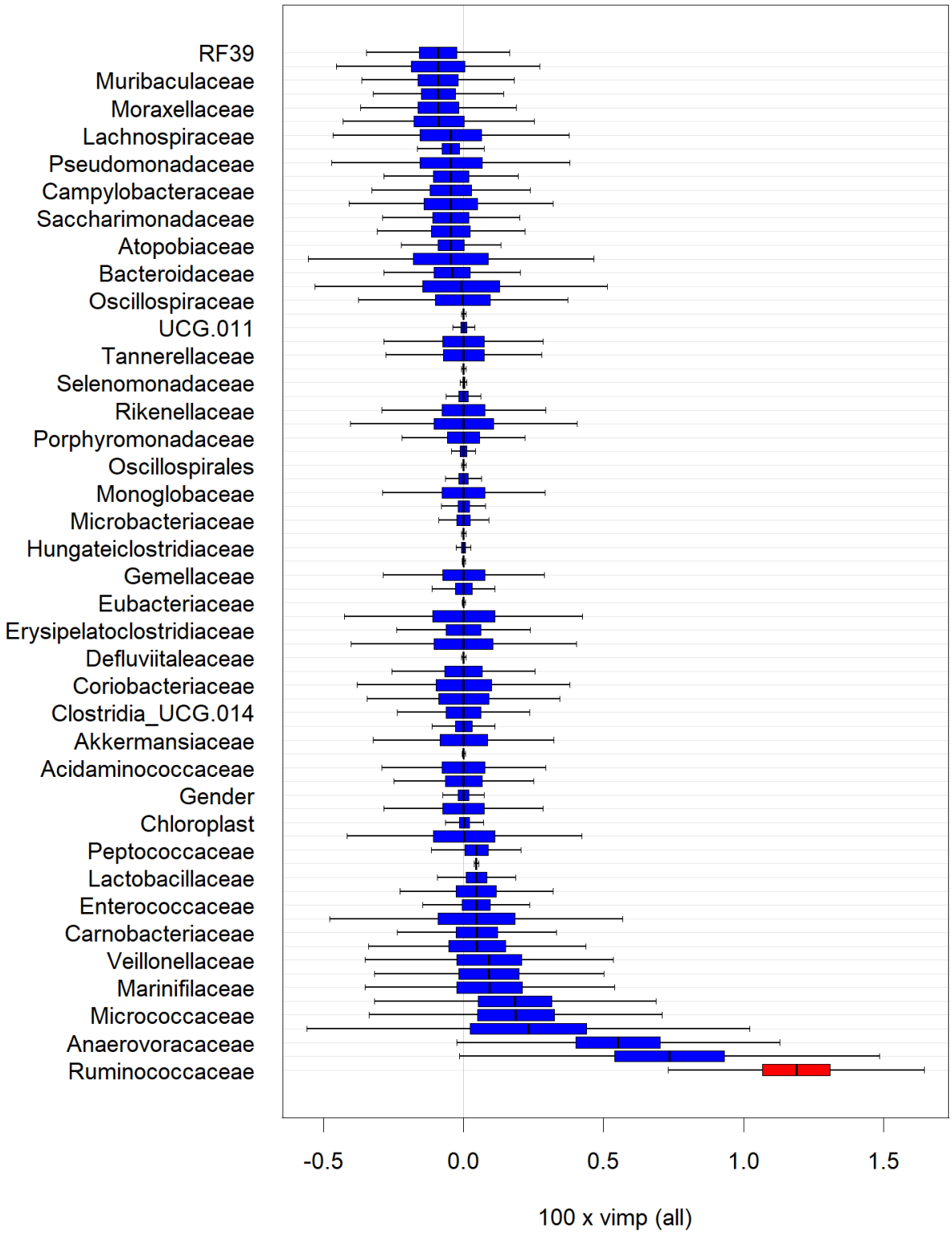


Figure ML2-D-5: The Variable Importance Plot (VIMP) for Family in Stool

Best model is selected for LIME results

Table ML2-D-3: The best setting for model and LIME Results

Setting	mtry	ROC	Sens	Spec	ROCSD	SensSD	SpecSD
ntrees: 500 nodesize: 1	7	0.7	0	1	0.410792	0	0
ntrees: 1000 nodesize: 1	7	0.7	0	1	0.240081	0	0
ntrees: 500 nodesize: 5	8	0.7	0	1	0.410792	0	0
ntrees: 1000 nodesize: 5	8	0.733333	0	1	0.245798	0	0

Table ML2-D-4: The LIME Result for Training Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
1	Adenoma	0.709	0.485	0.326	0.459	Flavobacteriaceae	0.000	0.054	2.24e-05 < Flavobacteriaceae
1	Adenoma	0.709	0.485	0.326	0.459	Anaerovoracaceae	0.000	0.053	Anaerovoracaceae <= 0.000621
1	Adenoma	0.709	0.485	0.326	0.459	Micrococcaceae	0.000	-0.054	Micrococcaceae <= 5.67e-05
1	Adenoma	0.709	0.485	0.326	0.459	Ruminococcaceae	0.018	0.043	Ruminococcaceae <= 0.0432
1	Adenoma	0.709	0.485	0.326	0.459	Izemoplasmatales	0.006	0.037	0.000464 < Izemoplasmatales
2	Adenoma	0.668	0.394	0.332	0.440	Flavobacteriaceae	0.001	0.052	2.24e-05 < Flavobacteriaceae
2	Adenoma	0.668	0.394	0.332	0.440	Micrococcaceae	0.000	-0.054	Micrococcaceae <= 5.67e-05
2	Adenoma	0.668	0.394	0.332	0.440	Ruminococcaceae	0.023	0.046	Ruminococcaceae <= 0.0432
2	Adenoma	0.668	0.394	0.332	0.440	Izemoplasmatales	0.001	0.035	0.000464 < Izemoplasmatales
2	Adenoma	0.668	0.394	0.332	0.440	Sutterellaceae	0.007	0.029	0.003669 < Sutterellaceae
3	Adenoma	0.671	0.438	0.361	0.411	Anaerovoracaceae	0.000	0.057	Anaerovoracaceae <= 0.000621
3	Adenoma	0.671	0.438	0.361	0.411	Flavobacteriaceae	0.000	-0.052	Flavobacteriaceae <= 2.24e-05
3	Adenoma	0.671	0.438	0.361	0.411	Micrococcaceae	0.000	0.052	5.67e-05 < Micrococcaceae
3	Adenoma	0.671	0.438	0.361	0.411	Izemoplasmatales	0.000	-0.033	Izemoplasmatales <= 0.000464
3	Adenoma	0.671	0.438	0.361	0.411	Butyricicoccaceae	0.002	0.027	0.000944 < Butyricicoccaceae
6	Adenoma	0.666	0.434	0.278	0.498	Flavobacteriaceae	0.002	0.053	2.24e-05 < Flavobacteriaceae
6	Adenoma	0.666	0.434	0.278	0.498	Micrococcaceae	0.000	0.054	5.67e-05 < Micrococcaceae
6	Adenoma	0.666	0.434	0.278	0.498	Anaerovoracaceae	0.001	0.053	Anaerovoracaceae <= 0.000621
6	Adenoma	0.666	0.434	0.278	0.498	Izemoplasmatales	0.002	0.033	0.000464 < Izemoplasmatales
6	Adenoma	0.666	0.434	0.278	0.498	Sutterellaceae	0.004	0.027	0.003669 < Sutterellaceae
9	Adenoma	0.780	0.492	0.303	0.479	Micrococcaceae	0.001	0.057	5.67e-05 < Micrococcaceae
9	Adenoma	0.780	0.492	0.303	0.479	Anaerovoracaceae	0.000	0.055	Anaerovoracaceae <= 0.000621
9	Adenoma	0.780	0.492	0.303	0.479	Flavobacteriaceae	0.000	0.053	2.24e-05 < Flavobacteriaceae
9	Adenoma	0.780	0.492	0.303	0.479	Ruminococcaceae	0.001	0.045	Ruminococcaceae <= 0.0432
9	Adenoma	0.780	0.492	0.303	0.479	Izemoplasmatales	0.000	-0.034	Izemoplasmatales <= 0.000464
11	Adenocarcinoma	0.704	0.381	0.611	0.622	Micrococcaceae	0.000	0.056	Micrococcaceae <= 5.67e-05
11	Adenocarcinoma	0.704	0.381	0.611	0.622	Flavobacteriaceae	0.000	0.050	Flavobacteriaceae <= 2.24e-05
11	Adenocarcinoma	0.704	0.381	0.611	0.622	Ruminococcaceae	0.024	-0.040	Ruminococcaceae <= 0.0432
11	Adenocarcinoma	0.704	0.381	0.611	0.622	Izemoplasmatales	0.003	-0.028	0.000464 < Izemoplasmatales
11	Adenocarcinoma	0.704	0.381	0.611	0.622	Sutterellaceae	0.006	-0.027	0.003669 < Sutterellaceae
17	Adenocarcinoma	0.759	0.321	0.562	0.701	Micrococcaceae	0.000	0.053	Micrococcaceae <= 5.67e-05

17	Adenocarcinoma	0.759	0.321	0.562	0.701	Flavobacteriaceae	0.000	0.052	Flavobacteriaceae <= 2.24e-05
17	Adenocarcinoma	0.759	0.321	0.562	0.701	Izemoplasmatales	0.000	0.035	Izemoplasmatales <= 0.000464
17	Adenocarcinoma	0.759	0.321	0.562	0.701	Peptostreptococcaceae	0.013	-0.022	0.008633 < Peptostreptococcaceae
17	Adenocarcinoma	0.759	0.321	0.562	0.701	Anaerovoracaceae	0.027	0.021	0.004251 < Anaerovoracaceae
18	Adenocarcinoma	0.796	0.426	0.583	0.642	Anaerovoracaceae	0.000	-0.055	Anaerovoracaceae <= 0.000621
18	Adenocarcinoma	0.796	0.426	0.583	0.642	Flavobacteriaceae	0.000	0.056	Flavobacteriaceae <= 2.24e-05
18	Adenocarcinoma	0.796	0.426	0.583	0.642	Micrococcaceae	0.000	0.052	Micrococcaceae <= 5.67e-05
18	Adenocarcinoma	0.796	0.426	0.583	0.642	Izemoplasmatales	0.000	0.034	Izemoplasmatales <= 0.000464
18	Adenocarcinoma	0.796	0.426	0.583	0.642	Sutterellaceae	0.004	-0.027	0.003669 < Sutterellaceae
19	Adenocarcinoma	0.859	0.324	0.563	0.697	Flavobacteriaceae	0.000	0.054	Flavobacteriaceae <= 2.24e-05
19	Adenocarcinoma	0.859	0.324	0.563	0.697	Micrococcaceae	0.000	0.051	Micrococcaceae <= 5.67e-05
19	Adenocarcinoma	0.859	0.324	0.563	0.697	Izemoplasmatales	0.000	0.034	Izemoplasmatales <= 0.000464
19	Adenocarcinoma	0.859	0.324	0.563	0.697	Butyricicoccaceae	0.002	-0.027	0.000944 < Butyricicoccaceae
19	Adenocarcinoma	0.859	0.324	0.563	0.697	Anaerovoracaceae	0.004	0.022	0.002449 < Anaerovoracaceae <= 0.004251
20	Adenocarcinoma	0.878	0.349	0.579	0.658	Micrococcaceae	0.000	0.053	Micrococcaceae <= 5.67e-05
20	Adenocarcinoma	0.878	0.349	0.579	0.658	Flavobacteriaceae	0.000	0.051	Flavobacteriaceae <= 2.24e-05
20	Adenocarcinoma	0.878	0.349	0.579	0.658	Ruminococcaceae	0.040	-0.041	Ruminococcaceae <= 0.0432
20	Adenocarcinoma	0.878	0.349	0.579	0.658	Izemoplasmatales	0.000	0.033	Izemoplasmatales <= 0.000464
20	Adenocarcinoma	0.878	0.349	0.579	0.658	Pseudomonadaceae	0.008	-0.018	0.000243 < Pseudomonadaceae
21	Adenocarcinoma	0.920	0.295	0.565	0.699	Micrococcaceae	0.000	0.053	Micrococcaceae <= 5.67e-05
21	Adenocarcinoma	0.920	0.295	0.565	0.699	Flavobacteriaceae	0.000	0.050	Flavobacteriaceae <= 2.24e-05
21	Adenocarcinoma	0.920	0.295	0.565	0.699	Izemoplasmatales	0.000	0.033	Izemoplasmatales <= 0.000464
21	Adenocarcinoma	0.920	0.295	0.565	0.699	Peptostreptococcaceae	0.019	-0.020	0.008633 < Peptostreptococcaceae
21	Adenocarcinoma	0.920	0.295	0.565	0.699	Anaerovoracaceae	0.007	0.019	0.004251 < Anaerovoracaceae
22	Adenocarcinoma	0.923	0.312	0.575	0.667	Micrococcaceae	0.000	0.053	Micrococcaceae <= 5.67e-05
22	Adenocarcinoma	0.923	0.312	0.575	0.667	Flavobacteriaceae	0.000	0.051	Flavobacteriaceae <= 2.24e-05
22	Adenocarcinoma	0.923	0.312	0.575	0.667	Izemoplasmatales	0.000	0.034	Izemoplasmatales <= 0.000464
22	Adenocarcinoma	0.923	0.312	0.575	0.667	Marinifilaceae	0.017	-0.024	0.012128 < Marinifilaceae
22	Adenocarcinoma	0.923	0.312	0.575	0.667	Peptostreptococcaceae	0.009	-0.022	0.008633 < Peptostreptococcaceae
23	Adenocarcinoma	0.864	0.412	0.561	0.668	Micrococcaceae	0.000	0.054	Micrococcaceae <= 5.67e-05
23	Adenocarcinoma	0.864	0.412	0.561	0.668	Anaerovoracaceae	0.000	-0.054	Anaerovoracaceae <= 0.000621
23	Adenocarcinoma	0.864	0.412	0.561	0.668	Flavobacteriaceae	0.000	0.053	Flavobacteriaceae <= 2.24e-05
23	Adenocarcinoma	0.864	0.412	0.561	0.668	Izemoplasmatales	0.000	0.034	Izemoplasmatales <= 0.000464
23	Adenocarcinoma	0.864	0.412	0.561	0.668	Pseudomonadaceae	0.000	0.020	Pseudomonadaceae <= 0.000243
25	Adenocarcinoma	0.891	0.335	0.623	0.613	Flavobacteriaceae	0.000	0.056	Flavobacteriaceae <= 2.24e-05
25	Adenocarcinoma	0.891	0.335	0.623	0.613	Micrococcaceae	0.000	-0.053	5.67e-05 < Micrococcaceae
25	Adenocarcinoma	0.891	0.335	0.623	0.613	Izemoplasmatales	0.000	0.036	Izemoplasmatales <= 0.000464
25	Adenocarcinoma	0.891	0.335	0.623	0.613	Butyricicoccaceae	0.001	-0.027	0.000944 < Butyricicoccaceae
25	Adenocarcinoma	0.891	0.335	0.623	0.613	Peptostreptococcaceae	0.009	-0.024	0.008633 < Peptostreptococcaceae
26	Adenocarcinoma	0.878	0.306	0.629	0.638	Flavobacteriaceae	0.002	-0.057	2.24e-05 < Flavobacteriaceae
26	Adenocarcinoma	0.878	0.306	0.629	0.638	Micrococcaceae	0.000	0.053	Micrococcaceae <= 5.67e-05
26	Adenocarcinoma	0.878	0.306	0.629	0.638	Izemoplasmatales	0.000	-0.028	0.000464 < Izemoplasmatales
26	Adenocarcinoma	0.878	0.306	0.629	0.638	Anaerovoracaceae	0.004	0.022	0.002449 < Anaerovoracaceae <= 0.004251

26	Adenocarcinoma	0.878	0.306	0.629	0.638	Pseudomonadaceae	0.000	0.019	Pseudomonadaceae <= 0.000243
29	Adenocarcinoma	0.937	0.313	0.527	0.718	Micrococcaceae	0.000	0.054	Micrococcaceae <= 5.67e-05
29	Adenocarcinoma	0.937	0.313	0.527	0.718	Flavobacteriaceae	0.000	0.053	Flavobacteriaceae <= 2.24e-05
29	Adenocarcinoma	0.937	0.313	0.527	0.718	Izemoplasmatales	0.000	0.031	Izemoplasmatales <= 0.000464
29	Adenocarcinoma	0.937	0.313	0.527	0.718	Anaerovoracaceae	0.004	0.021	0.004251 < Anaerovoracaceae
29	Adenocarcinoma	0.937	0.313	0.527	0.718	Neisseriaceae	0.000	0.031	Neisseriaceae <= 0.000161
30	Adenocarcinoma	0.859	0.336	0.559	0.684	Flavobacteriaceae	0.000	0.056	Flavobacteriaceae <= 2.24e-05
30	Adenocarcinoma	0.859	0.336	0.559	0.684	Micrococcaceae	0.000	0.056	Micrococcaceae <= 5.67e-05
30	Adenocarcinoma	0.859	0.336	0.559	0.684	Izemoplasmatales	0.004	-0.036	0.000464 < Izemoplasmatales
30	Adenocarcinoma	0.859	0.336	0.559	0.684	Neisseriaceae	0.000	0.030	Neisseriaceae <= 0.000161
30	Adenocarcinoma	0.859	0.336	0.559	0.684	Anaerovoracaceae	0.002	0.018	0.000621 < Anaerovoracaceae <= 0.002449
31	Adenocarcinoma	0.795	0.326	0.683	0.555	Micrococcaceae	0.000	-0.052	5.67e-05 < Micrococcaceae
31	Adenocarcinoma	0.795	0.326	0.683	0.555	Flavobacteriaceae	0.000	-0.052	2.24e-05 < Flavobacteriaceae
31	Adenocarcinoma	0.795	0.326	0.683	0.555	Izemoplasmatales	0.000	0.034	Izemoplasmatales <= 0.000464
31	Adenocarcinoma	0.795	0.326	0.683	0.555	Butyricicoccaceae	0.002	-0.030	0.000944 < Butyricicoccaceae
31	Adenocarcinoma	0.795	0.326	0.683	0.555	Sutterellaceae	0.008	-0.027	0.003669 < Sutterellaceae
35	Adenocarcinoma	0.920	0.302	0.545	0.720	Micrococcaceae	0.000	0.055	Micrococcaceae <= 5.67e-05
35	Adenocarcinoma	0.920	0.302	0.545	0.720	Flavobacteriaceae	0.000	0.053	Flavobacteriaceae <= 2.24e-05
35	Adenocarcinoma	0.920	0.302	0.545	0.720	Izemoplasmatales	0.000	0.029	Izemoplasmatales <= 0.000464
35	Adenocarcinoma	0.920	0.302	0.545	0.720	Pseudomonadaceae	0.000	0.021	Pseudomonadaceae <= 0.000243
35	Adenocarcinoma	0.920	0.302	0.545	0.720	Ruminococcaceae	0.436	0.016	0.1639 < Ruminococcaceae
41	Adenocarcinoma	0.929	0.313	0.544	0.723	Micrococcaceae	0.000	0.054	Micrococcaceae <= 5.67e-05
41	Adenocarcinoma	0.929	0.313	0.544	0.723	Flavobacteriaceae	0.000	0.051	Flavobacteriaceae <= 2.24e-05
41	Adenocarcinoma	0.929	0.313	0.544	0.723	Izemoplasmatales	0.000	0.035	Izemoplasmatales <= 0.000464
41	Adenocarcinoma	0.929	0.313	0.544	0.723	Anaerovoracaceae	0.002	0.021	0.002449 < Anaerovoracaceae <= 0.004251
41	Adenocarcinoma	0.929	0.313	0.544	0.723	Pseudomonadaceae	0.000	0.019	Pseudomonadaceae <= 0.000243
42	Adenocarcinoma	0.874	0.301	0.598	0.666	Micrococcaceae	0.000	-0.054	5.67e-05 < Micrococcaceae
42	Adenocarcinoma	0.874	0.301	0.598	0.666	Flavobacteriaceae	0.000	0.052	Flavobacteriaceae <= 2.24e-05
42	Adenocarcinoma	0.874	0.301	0.598	0.666	Izemoplasmatales	0.000	0.033	Izemoplasmatales <= 0.000464
42	Adenocarcinoma	0.874	0.301	0.598	0.666	Anaerovoracaceae	0.099	0.019	0.004251 < Anaerovoracaceae
42	Adenocarcinoma	0.874	0.301	0.598	0.666	Pseudomonadaceae	0.000	0.018	Pseudomonadaceae <= 0.000243
43	Adenocarcinoma	0.905	0.353	0.568	0.688	Flavobacteriaceae	0.000	0.053	Flavobacteriaceae <= 2.24e-05
43	Adenocarcinoma	0.905	0.353	0.568	0.688	Micrococcaceae	0.000	0.051	Micrococcaceae <= 5.67e-05
43	Adenocarcinoma	0.905	0.353	0.568	0.688	Ruminococcaceae	0.042	-0.039	Ruminococcaceae <= 0.0432
43	Adenocarcinoma	0.905	0.353	0.568	0.688	Izemoplasmatales	0.000	0.035	Izemoplasmatales <= 0.000464
43	Adenocarcinoma	0.905	0.353	0.568	0.688	Anaerovoracaceae	0.008	0.020	0.004251 < Anaerovoracaceae

Table ML-D-5: The LIME Result for Test Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
4	Adenocarcinoma	0.649	0.333	0.612	0.650	Flavobacteriaceae	0.000	0.057	Flavobacteriaceae <= 2.24e-05
4	Adenocarcinoma	0.649	0.333	0.612	0.650	Micrococcaceae	0.000	-0.053	5.67e-05 < Micrococcaceae
4	Adenocarcinoma	0.649	0.333	0.612	0.650	Izemoplasmatales	0.000	0.034	Izemoplasmatales <= 0.000464

4	Adenocarcinoma	0.649	0.333	0.612	0.650	Anaerovoracaceae	0.006	0.023	0.004251 < Anaerovoracaceae
4	Adenocarcinoma	0.649	0.333	0.612	0.650	Peptostreptococcaceae	0.015	-0.022	0.008633 < Peptostreptococcaceae
5	Adenocarcinoma	0.731	0.442	0.579	0.645	Flavobacteriaceae	0.000	0.056	Flavobacteriaceae <= 2.24e-05
5	Adenocarcinoma	0.731	0.442	0.579	0.645	Anaerovoracaceae	0.000	-0.057	Anaerovoracaceae <= 0.000621
5	Adenocarcinoma	0.731	0.442	0.579	0.645	Micrococcaceae	0.000	0.054	Micrococcaceae <= 5.67e-05
5	Adenocarcinoma	0.731	0.442	0.579	0.645	Izemoplasmatales	0.000	0.035	Izemoplasmatales <= 0.000464
5	Adenocarcinoma	0.731	0.442	0.579	0.645	Marinifilaceae	0.034	-0.023	0.012128 < Marinifilaceae
7	Adenocarcinoma	0.683	0.384	0.632	0.600	Flavobacteriaceae	0.000	0.056	Flavobacteriaceae <= 2.24e-05
7	Adenocarcinoma	0.683	0.384	0.632	0.600	Micrococcaceae	0.000	-0.054	5.67e-05 < Micrococcaceae
7	Adenocarcinoma	0.683	0.384	0.632	0.600	Ruminococcaceae	0.035	-0.044	Ruminococcaceae <= 0.0432
7	Adenocarcinoma	0.683	0.384	0.632	0.600	Izemoplasmatales	0.000	0.031	Izemoplasmatales <= 0.000464
7	Adenocarcinoma	0.683	0.384	0.632	0.600	Peptostreptococcaceae	0.012	-0.021	0.008633 < Peptostreptococcaceae
8	Adenocarcinoma	0.791	0.310	0.584	0.670	Micrococcaceae	0.000	0.052	Micrococcaceae <= 5.67e-05
8	Adenocarcinoma	0.791	0.310	0.584	0.670	Flavobacteriaceae	0.000	-0.051	2.24e-05 < Flavobacteriaceae
8	Adenocarcinoma	0.791	0.310	0.584	0.670	Izemoplasmatales	0.000	0.035	Izemoplasmatales <= 0.000464
8	Adenocarcinoma	0.791	0.310	0.584	0.670	Anaerovoracaceae	0.003	0.021	0.002449 < Anaerovoracaceae <= 0.004251
8	Adenocarcinoma	0.791	0.310	0.584	0.670	Neisseriaceae	0.000	0.029	Neisseriaceae <= 0.000161
10	Adenocarcinoma	0.850	0.311	0.551	0.729	Flavobacteriaceae	0.000	0.055	Flavobacteriaceae <= 2.24e-05
10	Adenocarcinoma	0.850	0.311	0.551	0.729	Micrococcaceae	0.000	0.054	Micrococcaceae <= 5.67e-05
10	Adenocarcinoma	0.850	0.311	0.551	0.729	Izemoplasmatales	0.000	0.034	Izemoplasmatales <= 0.000464
10	Adenocarcinoma	0.850	0.311	0.551	0.729	Anaerovoracaceae	0.002	0.021	0.000621 < Anaerovoracaceae <= 0.002449
10	Adenocarcinoma	0.850	0.311	0.551	0.729	Ruminococcaceae	0.249	0.014	0.1639 < Ruminococcaceae
12	Adenocarcinoma	0.770	0.305	0.543	0.721	Flavobacteriaceae	0.000	0.054	Flavobacteriaceae <= 2.24e-05
12	Adenocarcinoma	0.770	0.305	0.543	0.721	Micrococcaceae	0.000	0.052	Micrococcaceae <= 5.67e-05
12	Adenocarcinoma	0.770	0.305	0.543	0.721	Izemoplasmatales	0.000	0.031	Izemoplasmatales <= 0.000464
12	Adenocarcinoma	0.770	0.305	0.543	0.721	Pseudomonadaceae	0.000	0.022	Pseudomonadaceae <= 0.000243
12	Adenocarcinoma	0.770	0.305	0.543	0.721	Anaerovoracaceae	0.001	0.019	0.000621 < Anaerovoracaceae <= 0.002449
13	Adenocarcinoma	0.786	0.321	0.550	0.691	Flavobacteriaceae	0.000	0.053	Flavobacteriaceae <= 2.24e-05
13	Adenocarcinoma	0.786	0.321	0.550	0.691	Micrococcaceae	0.000	0.053	Micrococcaceae <= 5.67e-05
13	Adenocarcinoma	0.786	0.321	0.550	0.691	Izemoplasmatales	0.000	0.035	Izemoplasmatales <= 0.000464
13	Adenocarcinoma	0.786	0.321	0.550	0.691	Marinifilaceae	0.026	-0.021	0.012128 < Marinifilaceae
13	Adenocarcinoma	0.786	0.321	0.550	0.691	Pseudomonadaceae	0.000	0.021	Pseudomonadaceae <= 0.000243
14	Adenocarcinoma	0.848	0.325	0.650	0.576	Flavobacteriaceae	0.000	-0.055	2.24e-05 < Flavobacteriaceae
14	Adenocarcinoma	0.848	0.325	0.650	0.576	Micrococcaceae	0.000	-0.055	5.67e-05 < Micrococcaceae
14	Adenocarcinoma	0.848	0.325	0.650	0.576	Izemoplasmatales	0.000	0.033	Izemoplasmatales <= 0.000464
14	Adenocarcinoma	0.848	0.325	0.650	0.576	Butyricicoccaceae	0.001	-0.026	0.000944 < Butyricicoccaceae
14	Adenocarcinoma	0.848	0.325	0.650	0.576	Neisseriaceae	0.000	0.028	Neisseriaceae <= 0.000161
15	Adenocarcinoma	0.778	0.349	0.622	0.634	Flavobacteriaceae	0.001	-0.053	2.24e-05 < Flavobacteriaceae
15	Adenocarcinoma	0.778	0.349	0.622	0.634	Micrococcaceae	0.000	0.051	Micrococcaceae <= 5.67e-05
15	Adenocarcinoma	0.778	0.349	0.622	0.634	Ruminococcaceae	0.041	-0.039	Ruminococcaceae <= 0.0432
15	Adenocarcinoma	0.778	0.349	0.622	0.634	Izemoplasmatales	0.000	0.032	Izemoplasmatales <= 0.000464
15	Adenocarcinoma	0.778	0.349	0.622	0.634	Anaerovoracaceae	0.001	0.020	0.000621 < Anaerovoracaceae <= 0.002449

16	Adenocarcinoma	0.847	0.334	0.646	0.615	Flavobacteriaceae	0.000	0.056	Flavobacteriaceae <= 2.24e-05
16	Adenocarcinoma	0.847	0.334	0.646	0.615	Micrococcaceae	0.000	-0.054	5.67e-05 < Micrococcaceae
16	Adenocarcinoma	0.847	0.334	0.646	0.615	Izemoplasmatales	0.002	-0.031	0.000464 < Izemoplasmatales
16	Adenocarcinoma	0.847	0.334	0.646	0.615	Marinifilaceae	0.017	-0.024	0.012128 < Marinifilaceae
16	Adenocarcinoma	0.847	0.334	0.646	0.615	Anaerovoracaceae	0.003	0.023	0.002449 < Anaerovoracaceae <= 0.004251
24	Adenocarcinoma	0.811	0.323	0.616	0.645	Flavobacteriaceae	0.000	-0.053	2.24e-05 < Flavobacteriaceae
24	Adenocarcinoma	0.811	0.323	0.616	0.645	Micrococcaceae	0.000	0.051	Micrococcaceae <= 5.67e-05
24	Adenocarcinoma	0.811	0.323	0.616	0.645	Izemoplasmatales	0.000	0.035	Izemoplasmatales <= 0.000464
24	Adenocarcinoma	0.811	0.323	0.616	0.645	Butyricicoccaceae	0.002	-0.028	0.000944 < Butyricicoccaceae
24	Adenocarcinoma	0.811	0.323	0.616	0.645	Anaerovoracaceae	0.004	0.023	0.002449 < Anaerovoracaceae <= 0.004251
27	Adenocarcinoma	0.876	0.302	0.530	0.720	Micrococcaceae	0.000	0.053	Micrococcaceae <= 5.67e-05
27	Adenocarcinoma	0.876	0.302	0.530	0.720	Flavobacteriaceae	0.000	0.052	Flavobacteriaceae <= 2.24e-05
27	Adenocarcinoma	0.876	0.302	0.530	0.720	Izemoplasmatales	0.000	0.032	Izemoplasmatales <= 0.000464
27	Adenocarcinoma	0.876	0.302	0.530	0.720	Anaerovoracaceae	0.011	0.025	0.004251 < Anaerovoracaceae
27	Adenocarcinoma	0.876	0.302	0.530	0.720	Neisseriaceae	0.000	0.028	Neisseriaceae <= 0.000161
28	Adenocarcinoma	0.754	0.425	0.581	0.644	Flavobacteriaceae	0.000	0.055	Flavobacteriaceae <= 2.24e-05
28	Adenocarcinoma	0.754	0.425	0.581	0.644	Micrococcaceae	0.000	0.052	Micrococcaceae <= 5.67e-05
28	Adenocarcinoma	0.754	0.425	0.581	0.644	Anaerovoracaceae	0.001	-0.052	Anaerovoracaceae <= 0.000621
28	Adenocarcinoma	0.754	0.425	0.581	0.644	Izemoplasmatales	0.000	0.035	Izemoplasmatales <= 0.000464
28	Adenocarcinoma	0.754	0.425	0.581	0.644	Butyricicoccaceae	0.001	-0.026	0.000944 < Butyricicoccaceae
32	Adenocarcinoma	0.658	0.373	0.547	0.674	Micrococcaceae	0.000	0.054	Micrococcaceae <= 5.67e-05
32	Adenocarcinoma	0.658	0.373	0.547	0.674	Flavobacteriaceae	0.000	0.052	Flavobacteriaceae <= 2.24e-05
32	Adenocarcinoma	0.658	0.373	0.547	0.674	Ruminococcaceae	0.000	-0.041	Ruminococcaceae <= 0.0432
32	Adenocarcinoma	0.658	0.373	0.547	0.674	Izemoplasmatales	0.000	0.033	Izemoplasmatales <= 0.000464
32	Adenocarcinoma	0.658	0.373	0.547	0.674	Neisseriaceae	0.000	0.029	Neisseriaceae <= 0.000161
33	Adenocarcinoma	0.617	0.370	0.570	0.686	Micrococcaceae	0.000	0.052	Micrococcaceae <= 5.67e-05
33	Adenocarcinoma	0.617	0.370	0.570	0.686	Flavobacteriaceae	0.000	0.051	Flavobacteriaceae <= 2.24e-05
33	Adenocarcinoma	0.617	0.370	0.570	0.686	Ruminococcaceae	0.003	-0.044	Ruminococcaceae <= 0.0432
33	Adenocarcinoma	0.617	0.370	0.570	0.686	Izemoplasmatales	0.000	0.033	Izemoplasmatales <= 0.000464
33	Adenocarcinoma	0.617	0.370	0.570	0.686	Anaerovoracaceae	0.003	0.022	0.002449 < Anaerovoracaceae <= 0.004251
34	Adenocarcinoma	0.731	0.460	0.671	0.551	Anaerovoracaceae	0.001	-0.057	Anaerovoracaceae <= 0.000621
34	Adenocarcinoma	0.731	0.460	0.671	0.551	Micrococcaceae	0.000	-0.053	5.67e-05 < Micrococcaceae
34	Adenocarcinoma	0.731	0.460	0.671	0.551	Flavobacteriaceae	0.000	0.052	Flavobacteriaceae <= 2.24e-05
34	Adenocarcinoma	0.731	0.460	0.671	0.551	Izemoplasmatales	0.001	-0.036	0.000464 < Izemoplasmatales
34	Adenocarcinoma	0.731	0.460	0.671	0.551	Sutterellaceae	0.005	-0.026	0.003669 < Sutterellaceae
36	Adenocarcinoma	0.725	0.308	0.616	0.649	Micrococcaceae	0.000	0.053	Micrococcaceae <= 5.67e-05
36	Adenocarcinoma	0.725	0.308	0.616	0.649	Flavobacteriaceae	0.001	-0.052	2.24e-05 < Flavobacteriaceae
36	Adenocarcinoma	0.725	0.308	0.616	0.649	Izemoplasmatales	0.000	0.033	Izemoplasmatales <= 0.000464
36	Adenocarcinoma	0.725	0.308	0.616	0.649	Anaerovoracaceae	0.006	0.019	0.004251 < Anaerovoracaceae
36	Adenocarcinoma	0.725	0.308	0.616	0.649	Veillonellaceae	0.000	-0.019	Veillonellaceae <= 0.000203
37	Adenocarcinoma	0.501	0.217	0.659	0.618	Flavobacteriaceae	0.001	-0.050	2.24e-05 < Flavobacteriaceae
37	Adenocarcinoma	0.501	0.217	0.659	0.618	Izemoplasmatales	0.000	0.033	Izemoplasmatales <= 0.000464

37	Adenocarcinoma	0.501	0.217	0.659	0.618	Butyricicoccaceae	0.002	-0.027	0.000944 < Butyricicoccaceae
37	Adenocarcinoma	0.501	0.217	0.659	0.618	Anaerovoracaceae	0.005	0.024	0.004251 < Anaerovoracaceae
37	Adenocarcinoma	0.501	0.217	0.659	0.618	Pseudomonadacea e	0.002	-0.020	0.000243 < Pseudomonadaceae
38	Adenocarcinoma	0.696	0.425	0.563	0.666	Anaerovoracaceae	0.000	-0.056	Anaerovoracaceae <= 0.000621
38	Adenocarcinoma	0.696	0.425	0.563	0.666	Micrococcaceae	0.000	0.053	Micrococcaceae <= 5.67e-05
38	Adenocarcinoma	0.696	0.425	0.563	0.666	Flavobacteriaceae	0.000	0.051	Flavobacteriaceae <= 2.24e-05
38	Adenocarcinoma	0.696	0.425	0.563	0.666	Izemoplasmatales	0.000	0.034	Izemoplasmatales <= 0.000464
38	Adenocarcinoma	0.696	0.425	0.563	0.666	Pseudomonadacea e	0.000	0.021	Pseudomonadaceae <= 0.000243
39	Adenocarcinoma	0.671	0.383	0.635	0.597	Micrococcaceae	0.000	-0.053	5.67e-05 < Micrococcaceae
39	Adenocarcinoma	0.671	0.383	0.635	0.597	Flavobacteriaceae	0.000	0.052	Flavobacteriaceae <= 2.24e-05
39	Adenocarcinoma	0.671	0.383	0.635	0.597	Ruminococcaceae	0.030	-0.041	Ruminococcaceae <= 0.0432
39	Adenocarcinoma	0.671	0.383	0.635	0.597	Izemoplasmatales	0.000	0.034	Izemoplasmatales <= 0.000464
39	Adenocarcinoma	0.671	0.383	0.635	0.597	Butyricicoccaceae	0.002	-0.030	0.000944 < Butyricicoccaceae
40	Adenocarcinoma	0.812	0.301	0.545	0.721	Flavobacteriaceae	0.000	0.054	Flavobacteriaceae <= 2.24e-05
40	Adenocarcinoma	0.812	0.301	0.545	0.721	Micrococcaceae	0.000	0.052	Micrococcaceae <= 5.67e-05
40	Adenocarcinoma	0.812	0.301	0.545	0.721	Izemoplasmatales	0.000	0.032	Izemoplasmatales <= 0.000464
40	Adenocarcinoma	0.812	0.301	0.545	0.721	Anaerovoracaceae	0.004	0.020	0.002449 < Anaerovoracaceae <= 0.004251
40	Adenocarcinoma	0.812	0.301	0.545	0.721	Pseudomonadacea e	0.000	0.018	Pseudomonadaceae <= 0.000243

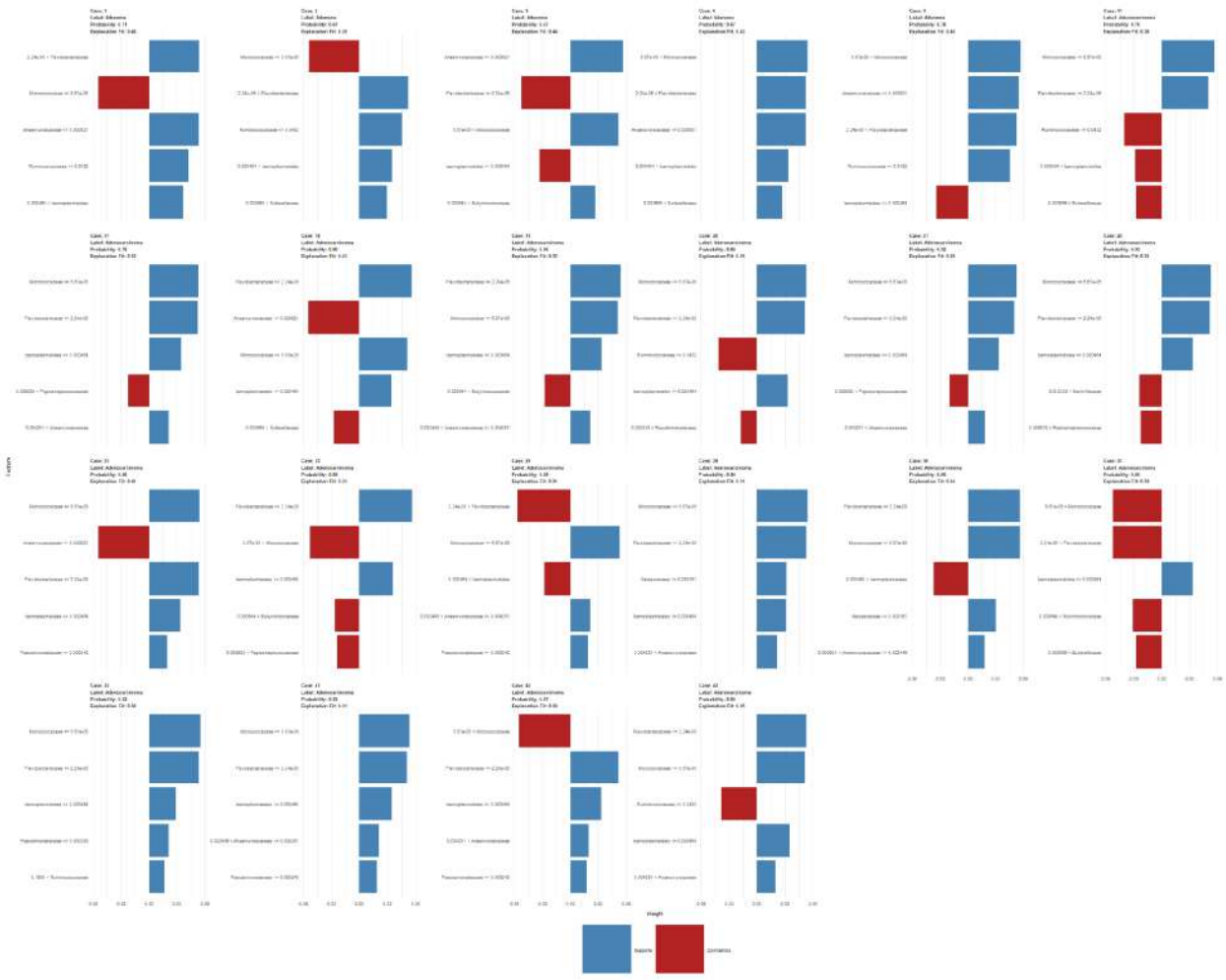


Figure ML2-D-6: The LIME Result in Training dataset for Family in Stool

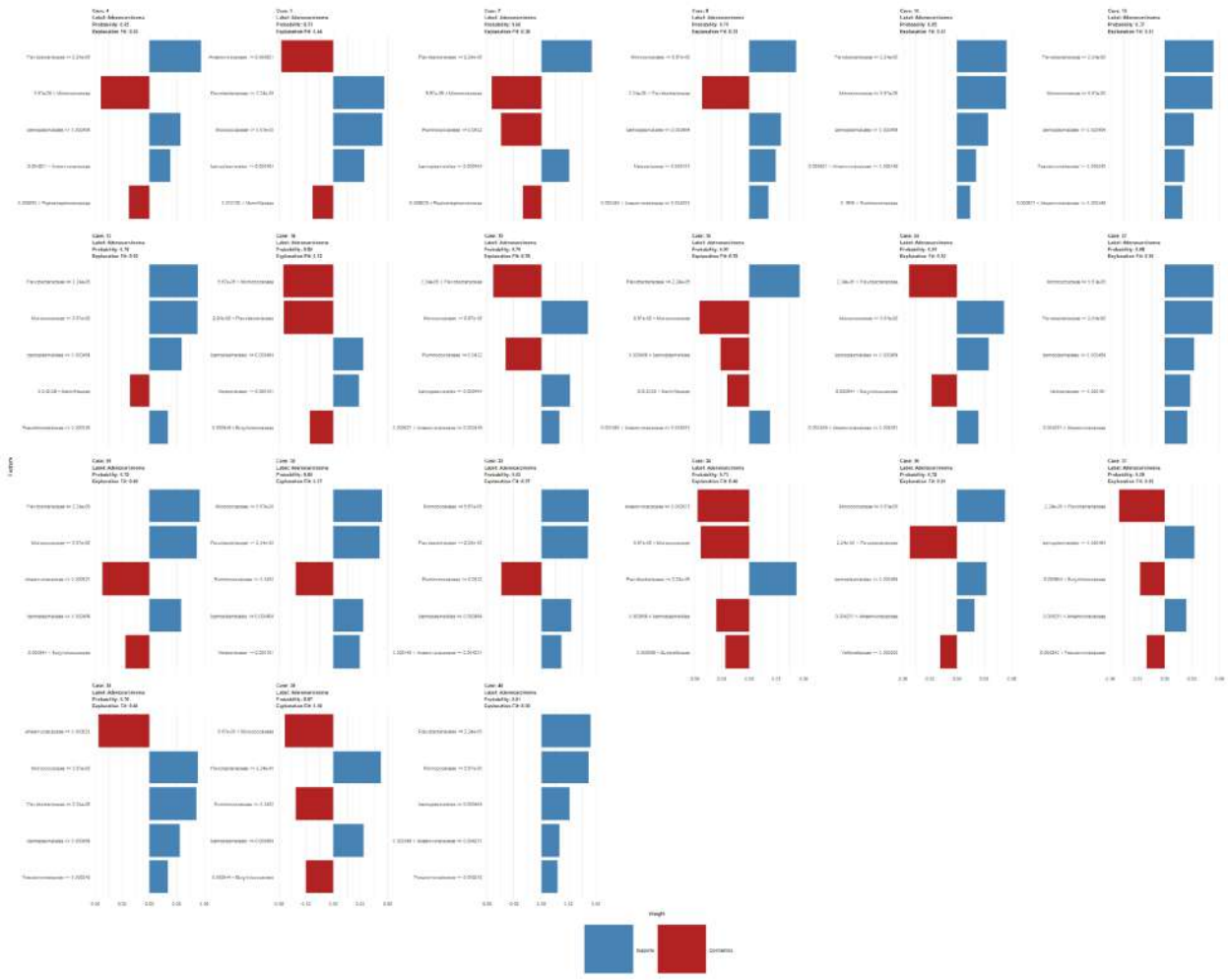


Figure ML2-D-7: The LIME Result in Test dataset for Family in Stool

Table ML2-D-6: The Aggregated SHAP Values

Variable	Training Dataset				Testing Dataset			
	contribution	position	sign	cumulative	contribution	position	sign	cumulative
intercept	0.7362	84	X	0.736	0.7362	84	X	0.7362
Gender	0.0000	83	1	0.736	0.0000	71	1.0000	0.7437
Yersiniaceae	0.0000	82	1	0.736	0.0007	82	1.0000	0.7376
Moraxellaceae	0.0000	81	-1	0.736	0.0006	83	1.0000	0.7369
Ruminococcaceae	0.0001	80	1	0.736	0.0020	75	1.0000	0.7417

Clostridia_UCG.014	0.0000	79	-1	0.736	0.0009	80	1.0000	0.7392
Leptotrichiaceae	0.0000	78	1	0.736	0.0000	68	-1.0000	0.7377
Barnesiellaceae	0.0000	77	-1	0.736	0.0022	78	1.0000	0.7391
Enterococcaceae	0.0000	76	1	0.736	-0.0002	74	-1.0000	0.7415
Selenomonadaceae	0.0000	75	1	0.736	0.0000	73	1.0000	0.7415
Christensenellaceae	0.0000	74	1	0.736	-0.0023	79	-1.0000	0.7369
Defluviitaleaceae	0.0000	73	1	0.736	0.0000	69	1.0000	0.7377
Tumor	0.0000	72	1	0.736	0.0000	14	1.0000	0.7513
Saccharimonadaceae	0.0000	71	-1	0.736	-0.0002	67	-1.0000	0.7375
Gastranaerophilales	0.0000	70	-1	0.736	-0.0001	66	-1.0000	0.7374
Neisseriaceae	0.000	69	1	0.736	-0.0012	65	-1.0000	0.7362
Chloroplast	0.000	68	1	0.736	0.0000	64	-1.0000	0.7362
Micrococcaceae	0.000	67	-1	0.736	-0.0024	63	-1.0000	0.7338
X.Clostridium._methylpentosum_group	0.000	66	1	0.736	0.0006	62	1.0000	0.7345
Akkermansiaceae	0.000	65	1	0.736	-0.0002	61	-1.0000	0.7343
DTU014	0.000	64	-1	0.736	0.0006	60	1.0000	0.7349
Anaerovoracaceae	0.000	63	1	0.736	0.0076	59	1.0000	0.7425
Carnobacteriaceae	0.000	62	-1	0.736	-0.0022	58	-1.0000	0.7403
Peptococcaceae	0.000	61	-1	0.736	0.0001	57	1.0000	0.7404
Leuconostocaceae	0.000	60	1	0.736	0.0000	54	1.0000	0.7408
Pseudomonadaceae	0.000	59	1	0.736	0.0004	56	1.0000	0.7408
Eggerthellaceae	0.000	58	-1	0.736	0.0005	53	1.0000	0.7414
Campylobacteraceae	0.000	57	1	0.736	-0.0009	51	-1.0000	0.7471
Pasteurellaceae	0.000	56	-1	0.736	0.0007	81	1.0000	0.7383
Desulfovibrionaceae	0.000	55	1	0.736	0.0066	52	1.0000	0.7480
Muribaculaceae	0.000	54	-1	0.736	-0.0009	50	-1.0000	0.7462
Synergistaceae	0.000	53	1	0.736	0.0000	55	1.0000	0.7408
RF39	0.000	52	-1	0.736	0.0005	40	1.0000	0.7527
Rikenellaceae	0.000	51	1	0.736	-0.0015	21	-1.0000	0.7567
Coriobacteriales_Incertae_Sedis	0.000	50	-1	0.736	0.0003	41	1.0000	0.7521
Eubacteriaceae	0.000	49	-1	0.736	-0.0001	49	-1.0000	0.7461
Atopobiaceae	0.000	48	-1	0.736	-0.0003	48	-1.0000	0.7458
Flavobacteriaceae	0.000	47	-1	0.736	0.0000	47	-1.0000	0.7458
Izemoplasmatales	0.000	46	-1	0.736	0.0064	46	1.0000	0.7522
UCG.011	0.000	45	1	0.736	0.0000	45	-1.0000	0.7521
Hafniaceae	0.000	44	1	0.736	0.0000	44	1.0000	0.7521
Oscillospirales	0.000	43	1	0.736	0.0000	42	1.0000	0.7518
Oscillospiraceae	0.000	42	-1	0.736	-0.0039	19	-1.0000	0.7536
Clostridia_vadinBB60_group	0.000	41	1	0.736	0.0021	72	1.0000	0.7437
uncultured	0.000	40	-1	0.736	0.0002	36	1.0000	0.7557
Prevotellaceae	0.000	39	-1	0.736	0.0006	76	1.0000	0.7397
Sutterellaceae	0.000	38	1	0.736	0.0008	39	1.0000	0.7535
Bifidobacteriaceae	0.000	37	-1	0.736	0.0023	38	1.0000	0.7558

Fusobacteriaceae	0.000	36	-1	0.736258	-0.0003	37	-1.0000	0.7554
Bacteroidaceae	0.000	35	-1	0.73624	0.0010	26	1.0000	0.7591
Erysipelatoclostridiaceae	0.000	34	-1	0.736238	-0.0004	43	-1.0000	0.7518
UCG.010	0.000	33	-1	0.736237	-0.0013	35	-1.0000	0.7544
Saccharimonadales	0.000	32	1	0.736239	0.0000	18	1.0000	0.7537
Coriobacteriaceae	0.000	31	-1	0.736234	0.0017	34	1.0000	0.7561
Peptostreptococcales.Tissierellales	0.000	30	-1	0.736215	-0.0002	33	-1.0000	0.7559
Aerococcaceae	0.000	29	1	0.736215	0.0000	31	1.0000	0.7552
Mitochondria	0.000	28	1	0.736221	-0.0007	32	-1.0000	0.7552
Morganellaceae	0.000	27	-1	0.736219	0.0001	77	1.0000	0.7392
Marinifilaceae	0.000	26	-1	0.736195	0.0008	30	1.0000	0.7561
Clostridiaceae	0.000	25	-1	0.736187	0.0008	29	1.0000	0.7568
Gemellaceae	0.000	24	-1	0.73617	0.0005	13	1.0000	0.7518
Veillonellaceae	0.000	23	-1	0.736134	0.0023	27	1.0000	0.7581
Hungateiclostridiaceae	0.000	22	-1	0.736131	-0.0001	25	-1.0000	0.7590
Porphyromonadaceae	0.000	21	-1	0.736127	0.0003	24	1.0000	0.7593
Streptococcaceae	0.000	20	1	0.736147	-0.0013	23	-1.0000	0.7581
Butyricicoccaceae	0.000	19	1	0.736203	-0.0060	70	-1.0000	0.7377
Lactobacillaceae	0.000	18	-1	0.736202	-0.0001	12	-1.0000	0.7517
Staphylococcaceae	0.000	17	1	0.736204	0.0009	20	1.0000	0.7576
vadinBE97	0.000	16	1	0.736204	0.0000	17	1.0000	0.7537
Monoglobaceae	0.000	15	-1	0.736202	-0.0004	16	-1.0000	0.7533
X.Eubacterium._coprostanoligenes_group	0.000	14	1	0.736229	-0.0020	15	-1.0000	0.7513
Methanobacteriaceae	0.000	13	1	0.736245	0.0001	22	1.0000	0.7582
Erysipelotrichaceae	0.000	12	-1	0.736202	-0.0081	11	-1.0000	0.7436
Peptostreptococcaceae	0.000	11	-1	0.736183	-0.0010	28	-1.0000	0.7558
Acidaminococcaceae	0.000	10	-1	0.736177	-0.0003	10	-1.0000	0.7432
Actinomycetaceae	0.000	9	1	0.736202	-0.0006	9	-1.0000	0.7427
Bacillaceae	0.000	8	1	0.736204	0.0000	8	1.0000	0.7427
Enterobacteriaceae	0.000	7	1	0.736205	0.0018	7	1.0000	0.7445
Lachnospiraceae	0.000	6	1	0.736219	-0.0024	6	-1.0000	0.7421
Microbacteriaceae	0.000	5	1	0.73622	0.0000	5	-1.0000	0.7421
Succinivibrionaceae	0.000	4	1	0.73622	0.0000	4	1.0000	0.7421
Tannerellaceae	0.000	3	-1	0.736213	-0.0001	3	-1.0000	0.7420
Victivallaceae	0.000	2	1	0.736227	0.0002	2	1.0000	0.7421
prediction	0.736	1	X	0.736227	0.7421	1	X	0.7421

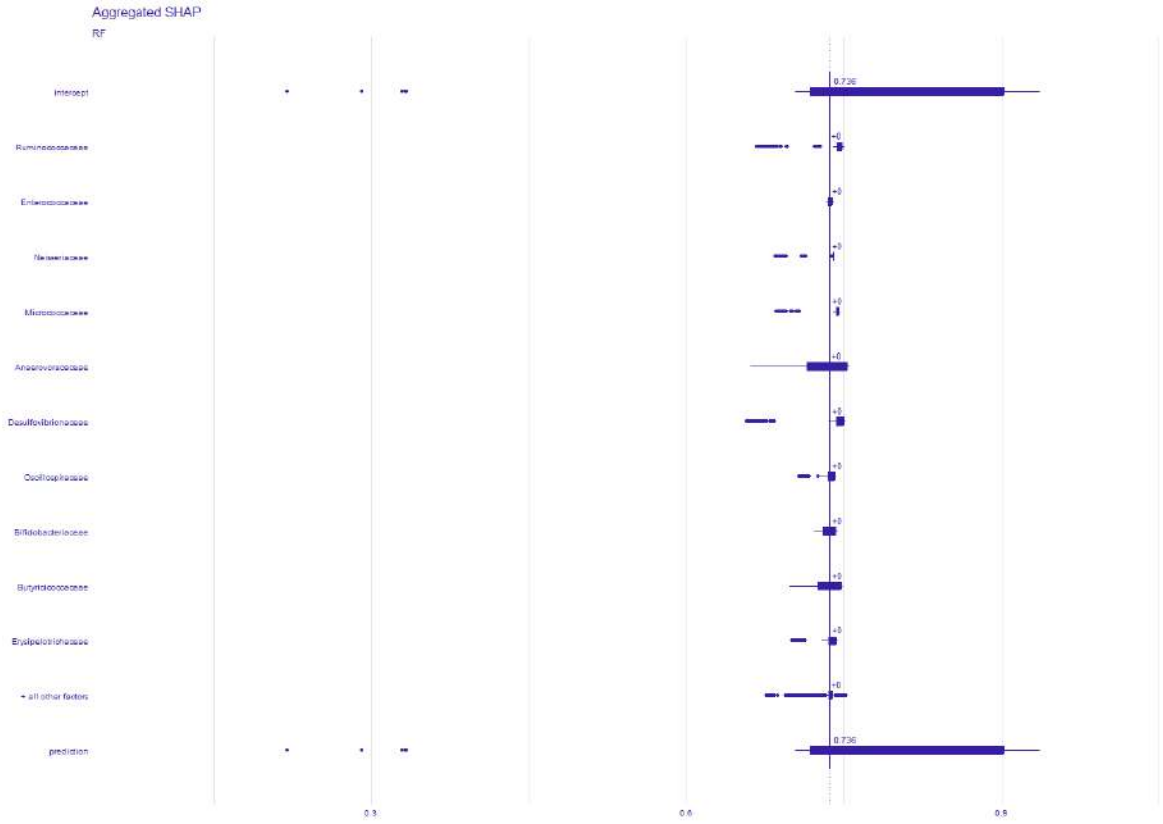


Figure ML2-D-8: The Aggregated SHAP Value in Training dataset for Family in Stool

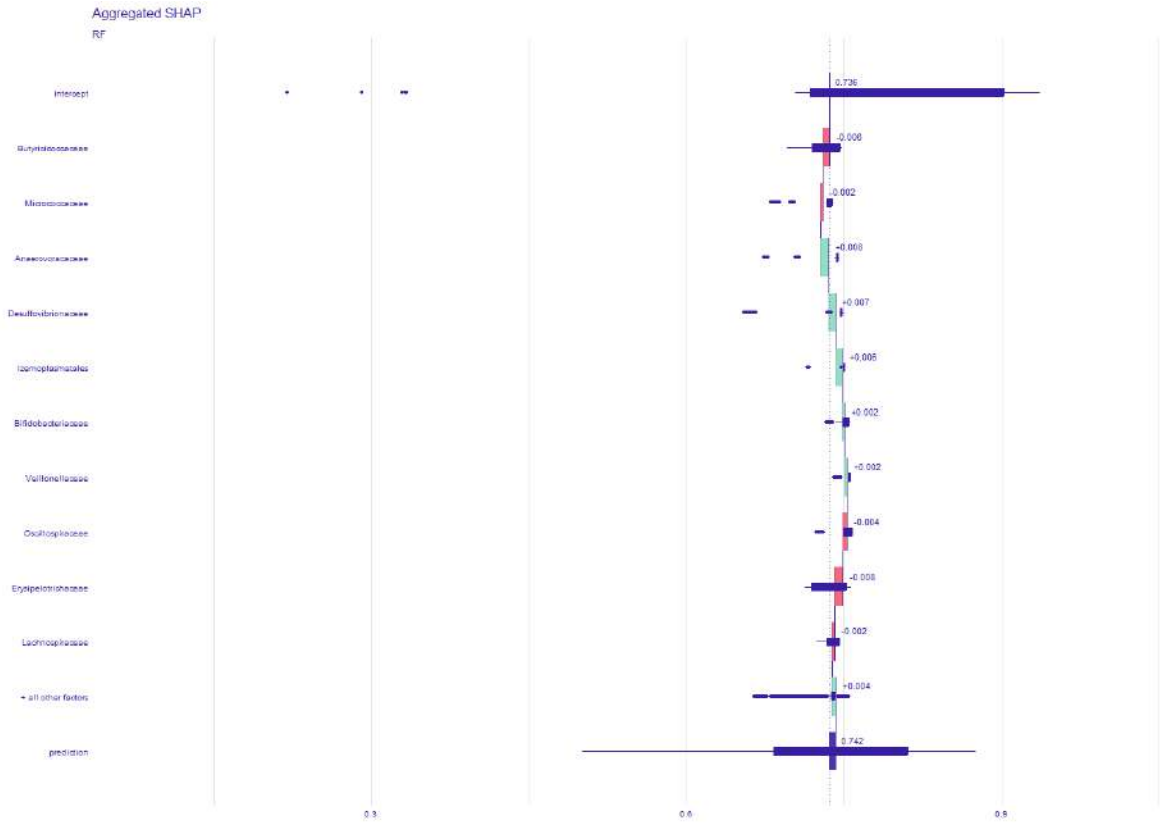


Figure ML2-D-9: The Aggregated SHAP Value in Test dataset for Family in Stool

E- Genus

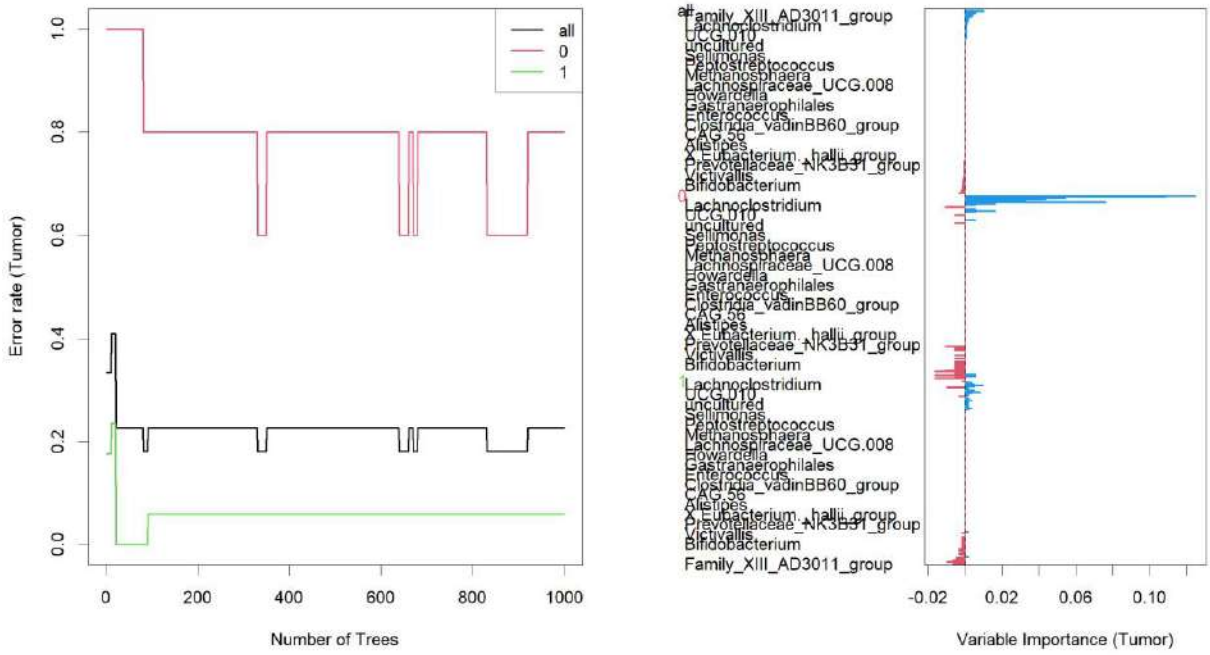
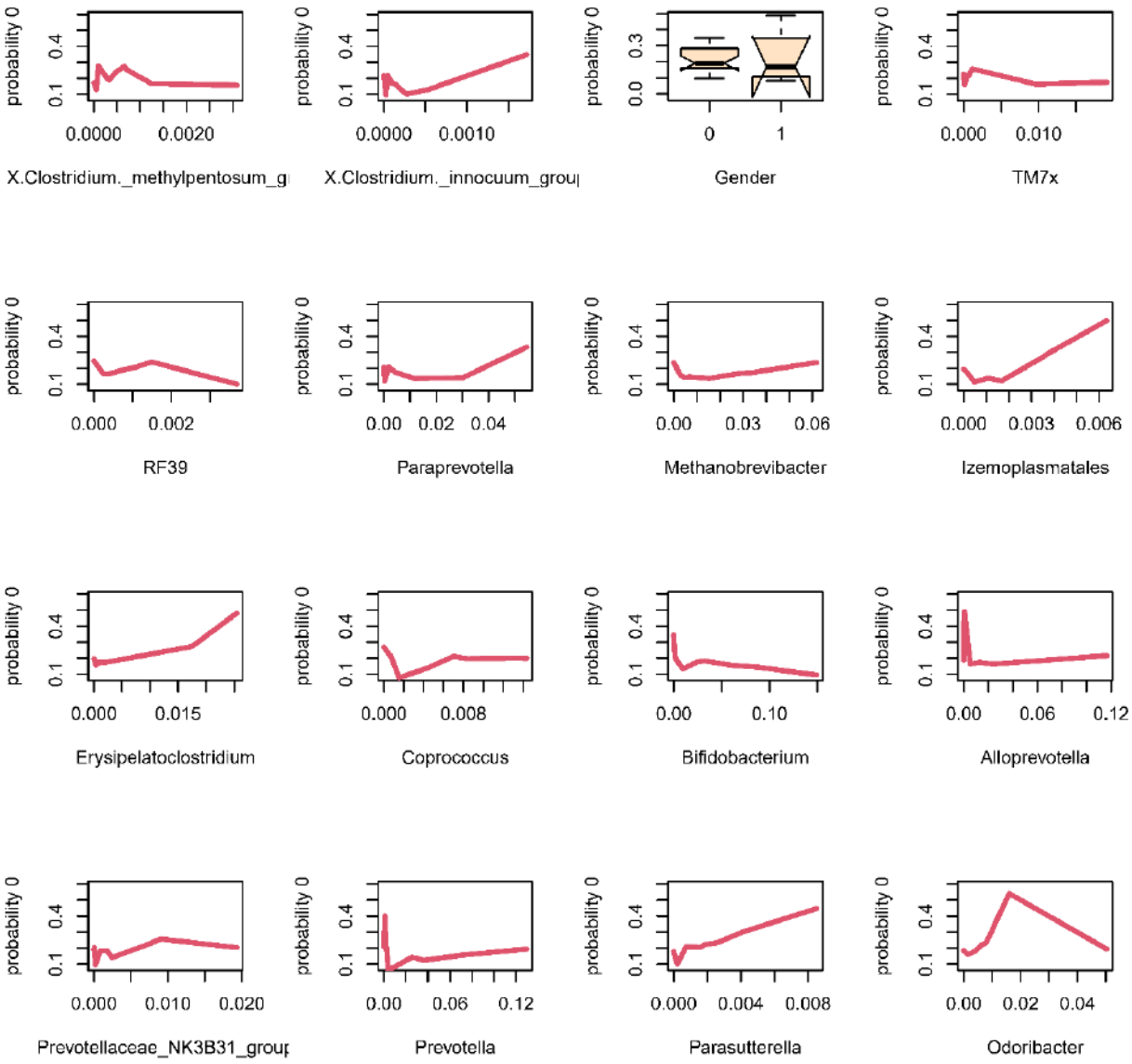
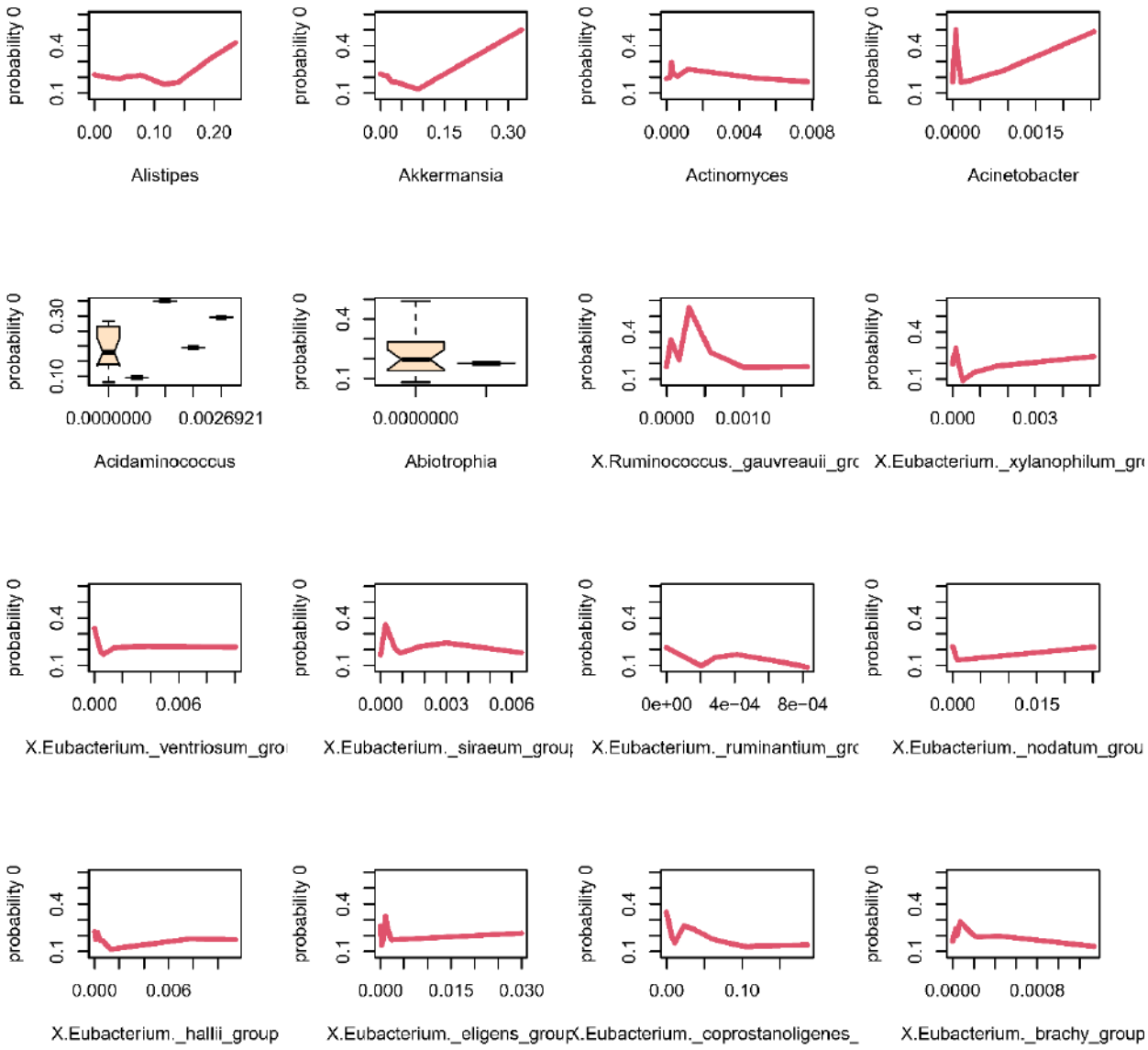
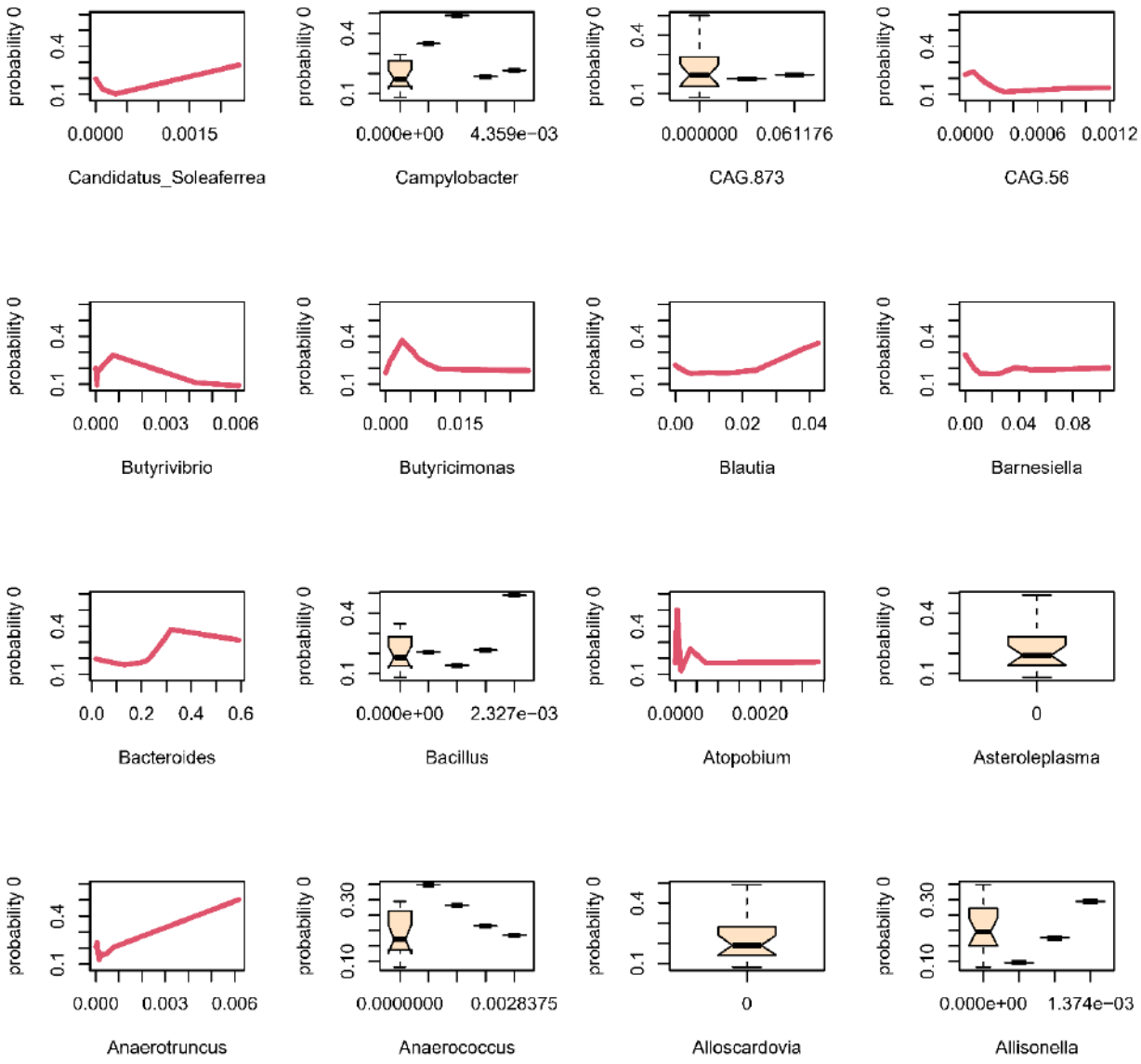
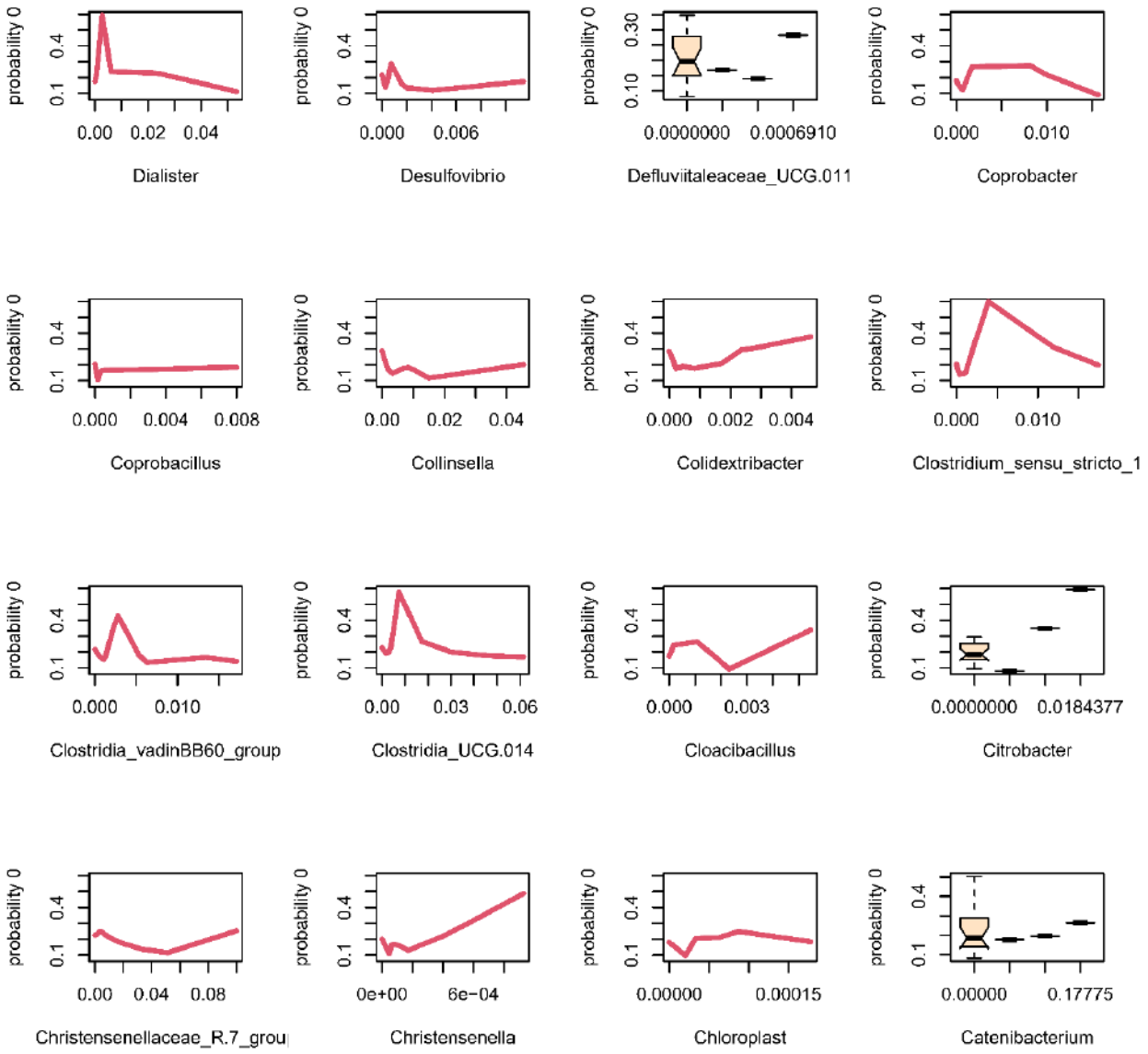


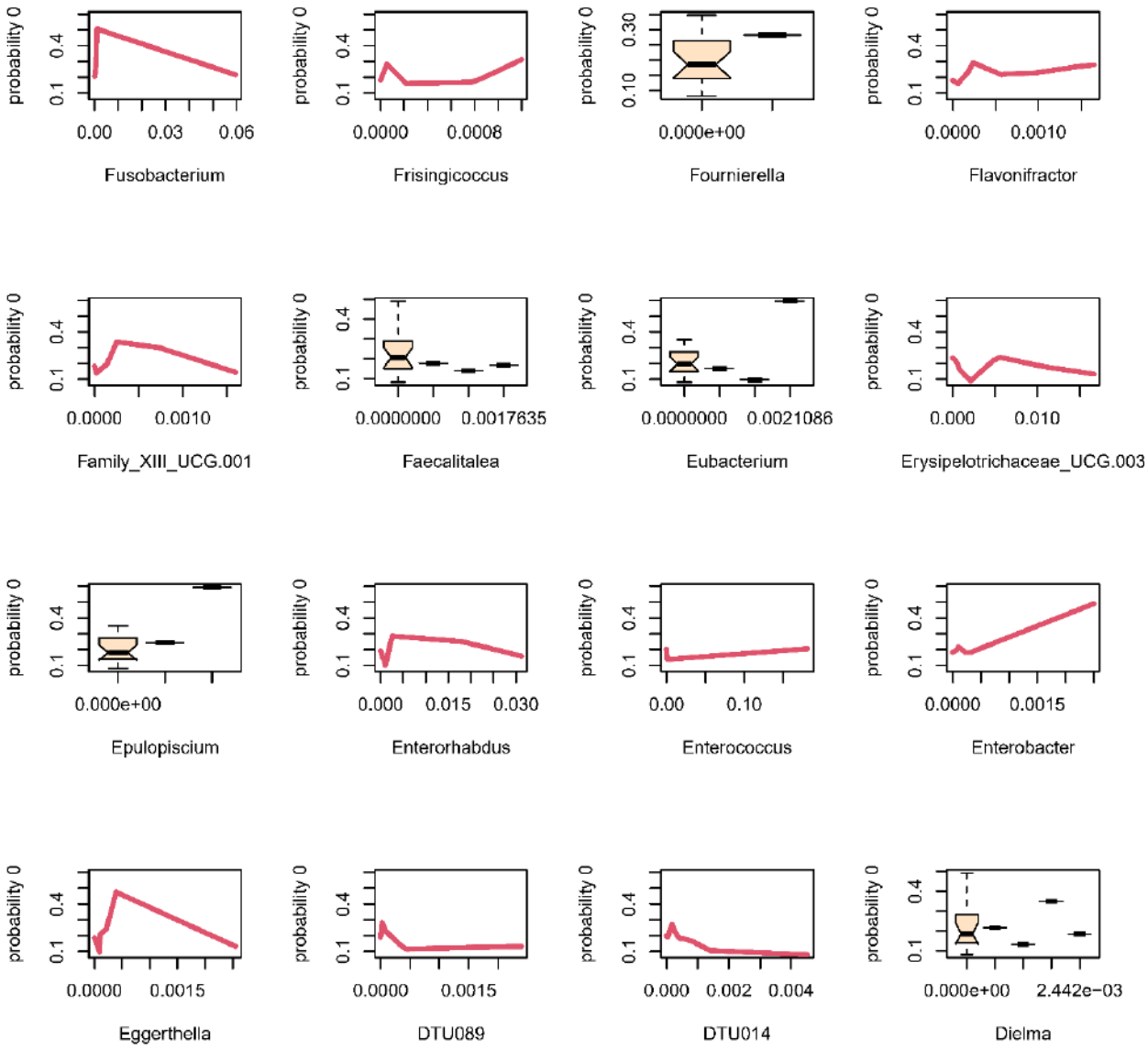
Figure ML2-E-1: The number of trees and Variable Importance Plot for Genus in Stool

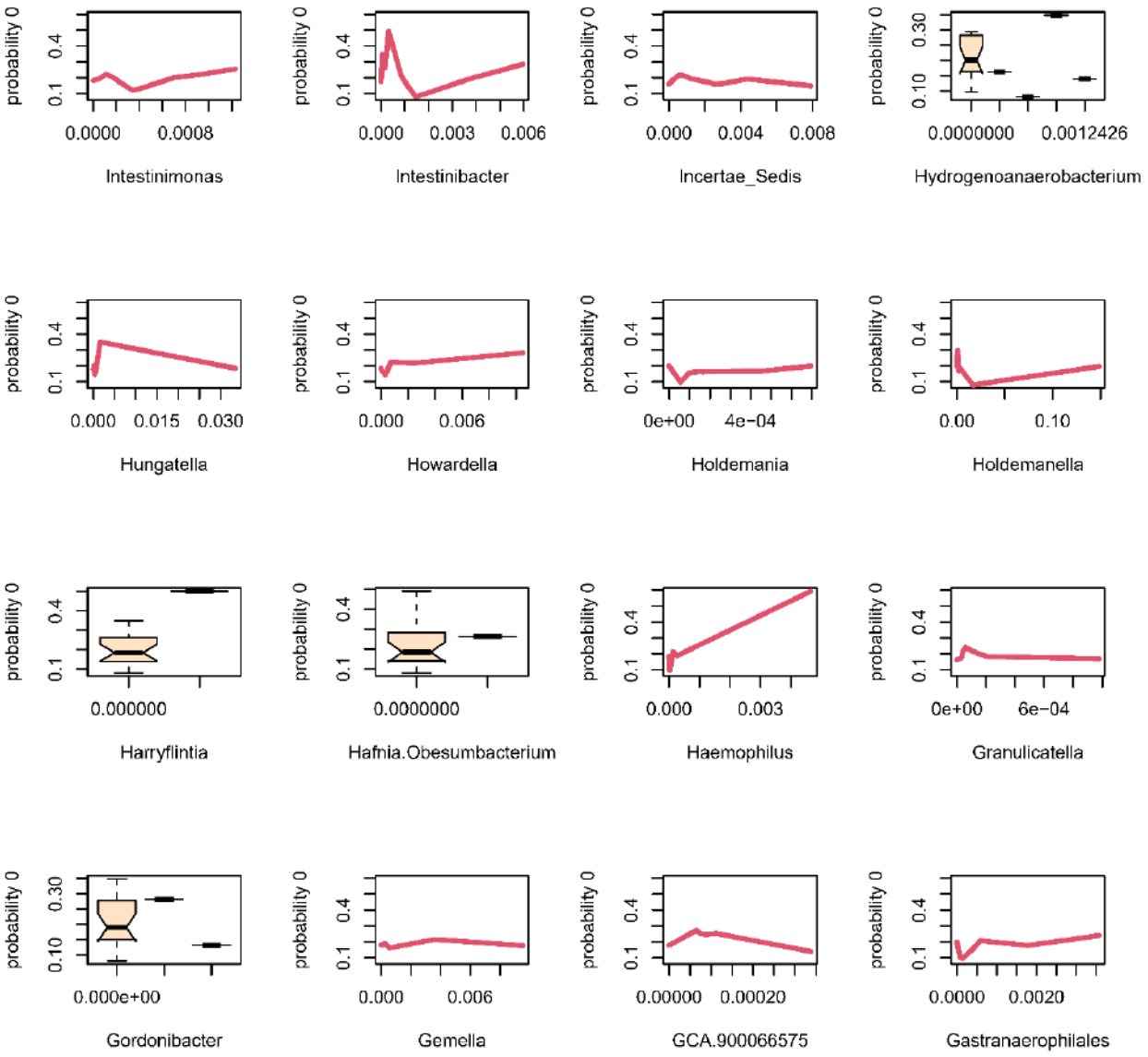

probability 0
0.1
0.4
0.0000
0.0020
X.Clostridium._methylpentosum_g
0.0000
0.0010
X.Clostridium._innocuum_grou
0.0
0.3
0
1
Gender
0.000
0.010
TM7x
0.000
0.002
RF39
0.00
0.02
0.04
Paraprevotella
0.00
0.03
0.06
Methanobrevibacter
0.000
0.003
0.006
Izemoplasmatales
0.000
0.015
Erysipelatoclostridium
0.000
0.008
Coprococcus
0.00
0.10
Bifidobacterium
0.00
0.06
0.12
Alloprevotella
0.000
0.010
0.020
Prevotellaceae_NK3B31_grou
0.00
0.06
0.12
Prevotella
0.000
0.004
0.008
Parasutterella
0.00
0.02
0.04
Odoribacter

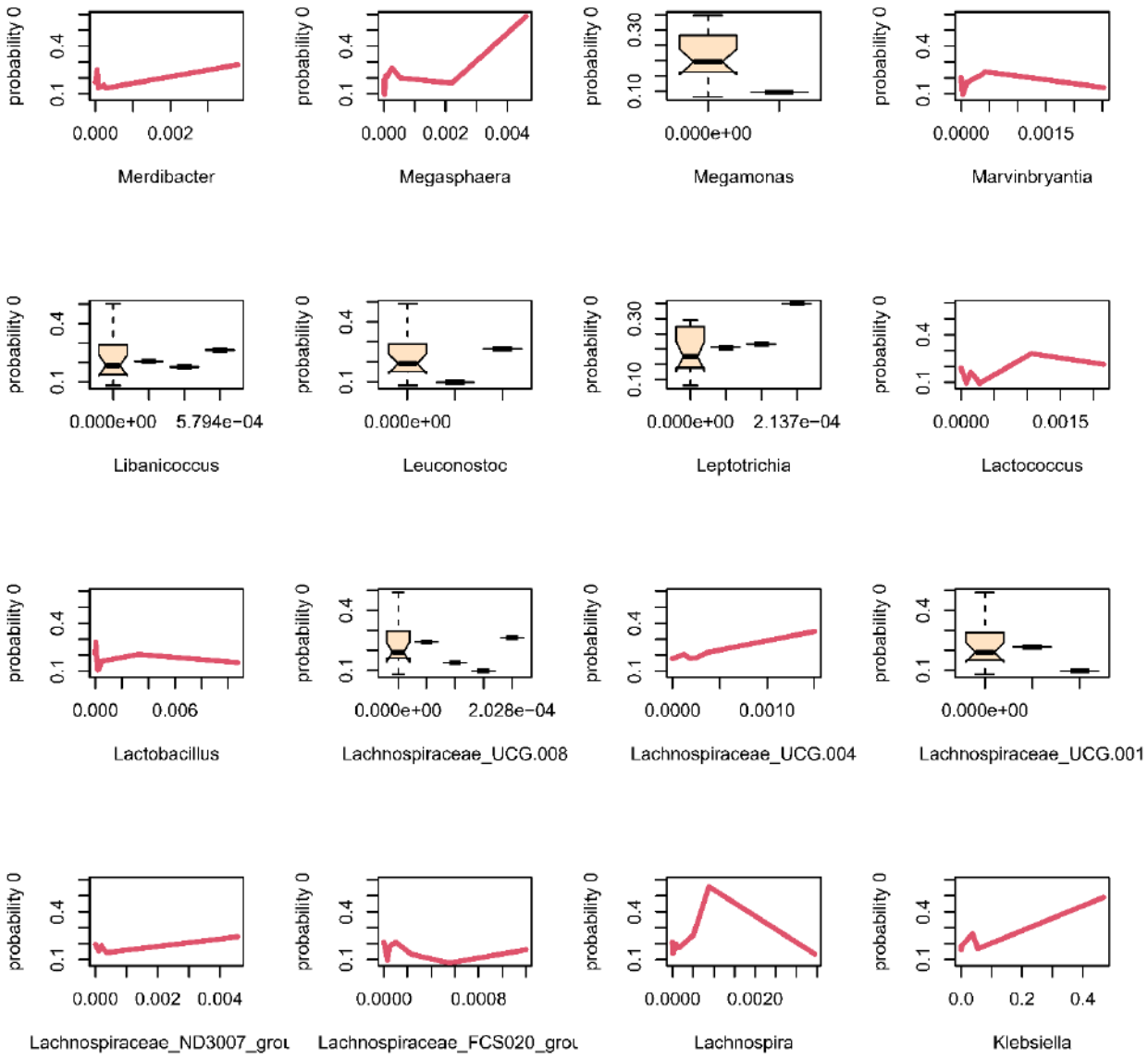

probability 0
0.1
0.4
0.00
0.10
0.20
Alistipes
0.00
0.15
0.30
Akkermansia
0.000
0.004
0.008
Actinomyces
0.0000
0.0015
Acinetobacter
0.10
0.30
0.0000000
0.0026921
Acidaminococcus
0.0000000
Abiotrophia
0.0000
0.0010
X.Ruminococcus._gauvreauii_grc
0.000
0.003
X.Eubacterium._xylanophilum_gr
0.000
0.006
X.Eubacterium._ventriosum_gro
probability
0.000
0.003
0.006
X.Eubacterium._siraeum_grou
0e+00
4e−04
8e−04
X.Eubacterium._ruminantium_grc
0.000
0.015
X.Eubacterium._nodatum_grou
0.000
0.006
X.Eubacterium._hallii_group
0.000
0.015
0.030
X.Eubacterium._eligens_group
0.00
0.10
X.Eubacterium._coprostanoligenes_
0.0000
0.0008
X.Eubacterium._brachy_group

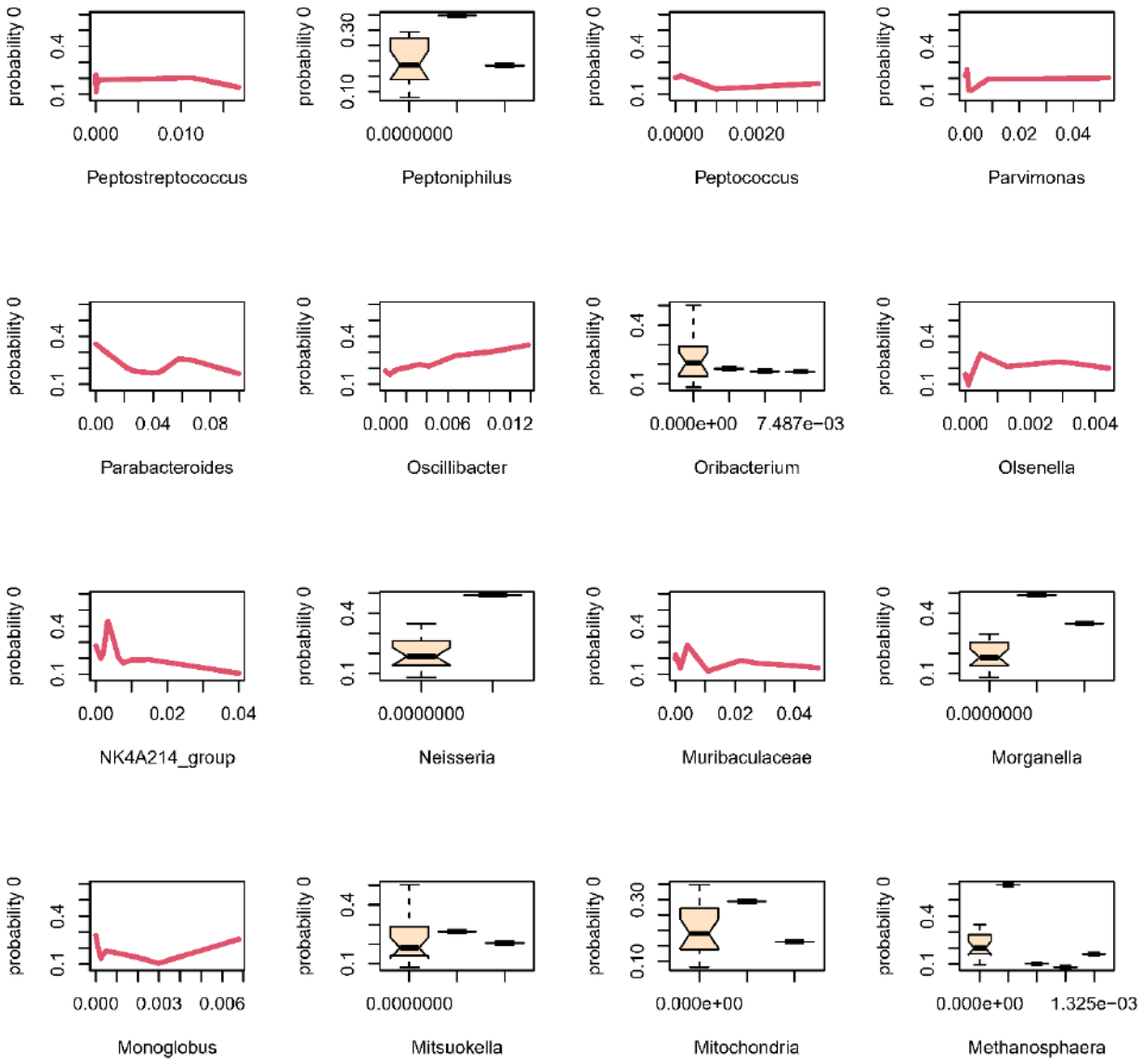

probability 0
0.1
0.4
0.0000
0.0015
Candidatus_Soleaferrea
probability 0
0.1
0.4
0.000e+00
4.359e−03
Campylobacter
probability 0
0.1
0.4
0.000000
0.061176
CAG.873
probability 0
0.1
0.4
0.0000
0.0006
0.0012
CAG.56
probability 0
0.1
0.4
0.000
0.003
0.006
Butyrivibrio
probability 0
0.1
0.4
0.000
0.015
Butyricimonas
probability 0
0.1
0.4
0.00
0.02
0.04
Blautia
probability 0
0.1
0.4
0.00
0.04
0.08
Barnesiella
probability 0
0.1
0.4
0.0
0.2
0.4
0.6
Bacteroides
probability 0
0.1
0.4
0.000e+00
2.327e−03
Bacillus
probability 0
0.1
0.4
0.0000
0.0020
Atopobium
probability 0
0.1
0.4
0
Asteroleplasma
probability 0
0.1
0.4
0.000
0.003
0.006
Anaerotruncus
probability 0
0.10
0.30
0.0000000
0.0028375
Anaerococcus
probability 0
0.1
0.4
0
Alloscardovia
probability 0
0.10
0.30
0.000e+00
1.374e−03
Allisonella

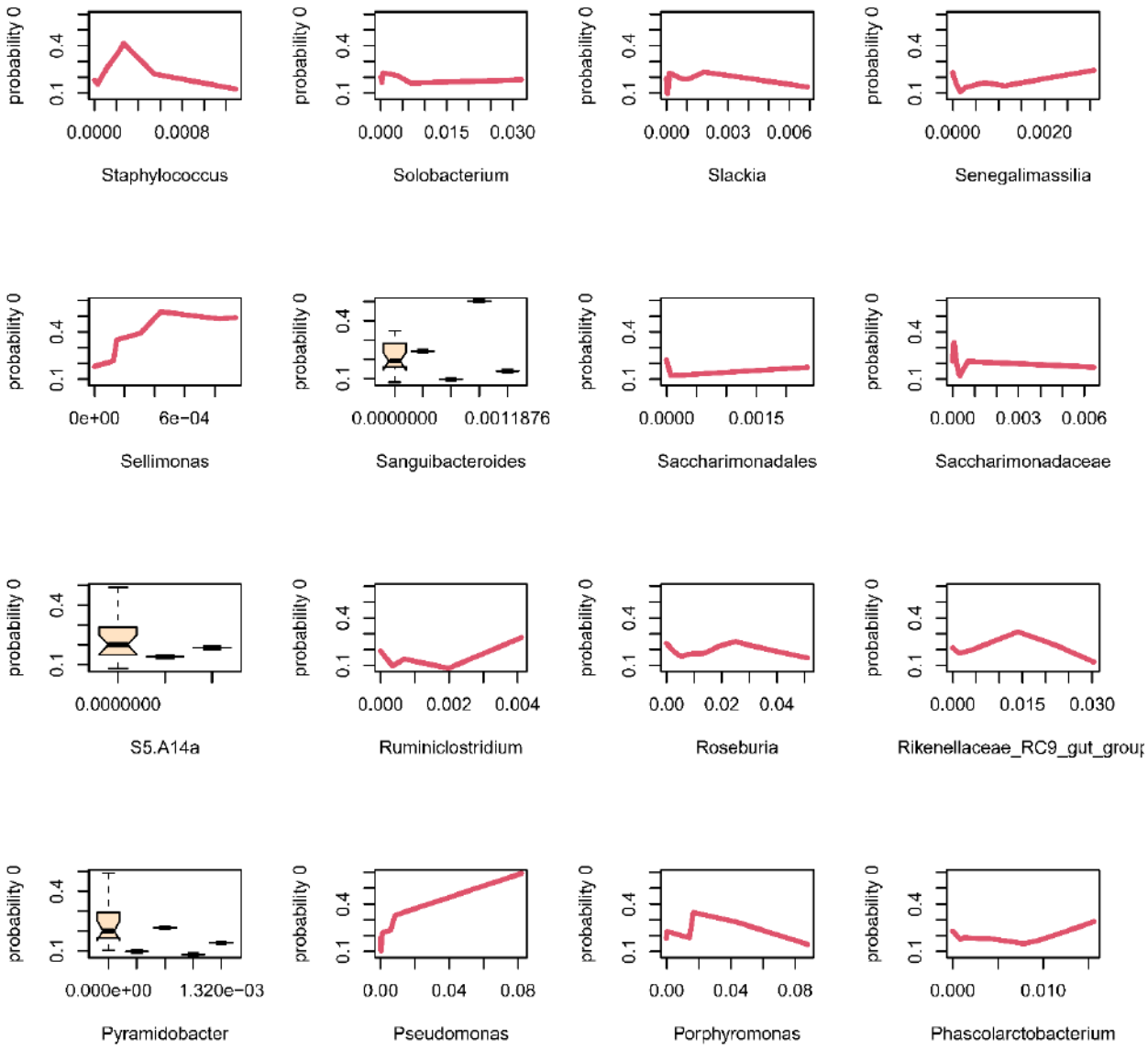

probability 0
0.1
0.4
0.00 0.02 0.04
Dialister
0.000 0.006
Desulfovibrio
0.10 0.30
0.0000000 0.0006910
Defluviitaleaceae_UCG.011
0.000 0.010
Coprobacter
0.000 0.004 0.008
Coprobacillus
0.00 0.02 0.04
Collinsella
0.000 0.002 0.004
Colidextribacter
0.000 0.010
Clostridium_sensu_stricto_1
0.000 0.010
Clostridia_vadinBB60_group
probability 0
0.00 0.03 0.06
Clostridia_UCG.014
0.000 0.003
Cloacibacillus
0.0000000 0.0184377
Citrobacter
0.00 0.04 0.08
Christensenellaceae_R.7_grou
0e+00 6e-04
Christensenella
0.00000 0.00015
Chloroplast
0.00000 0.17775
Catenibacterium

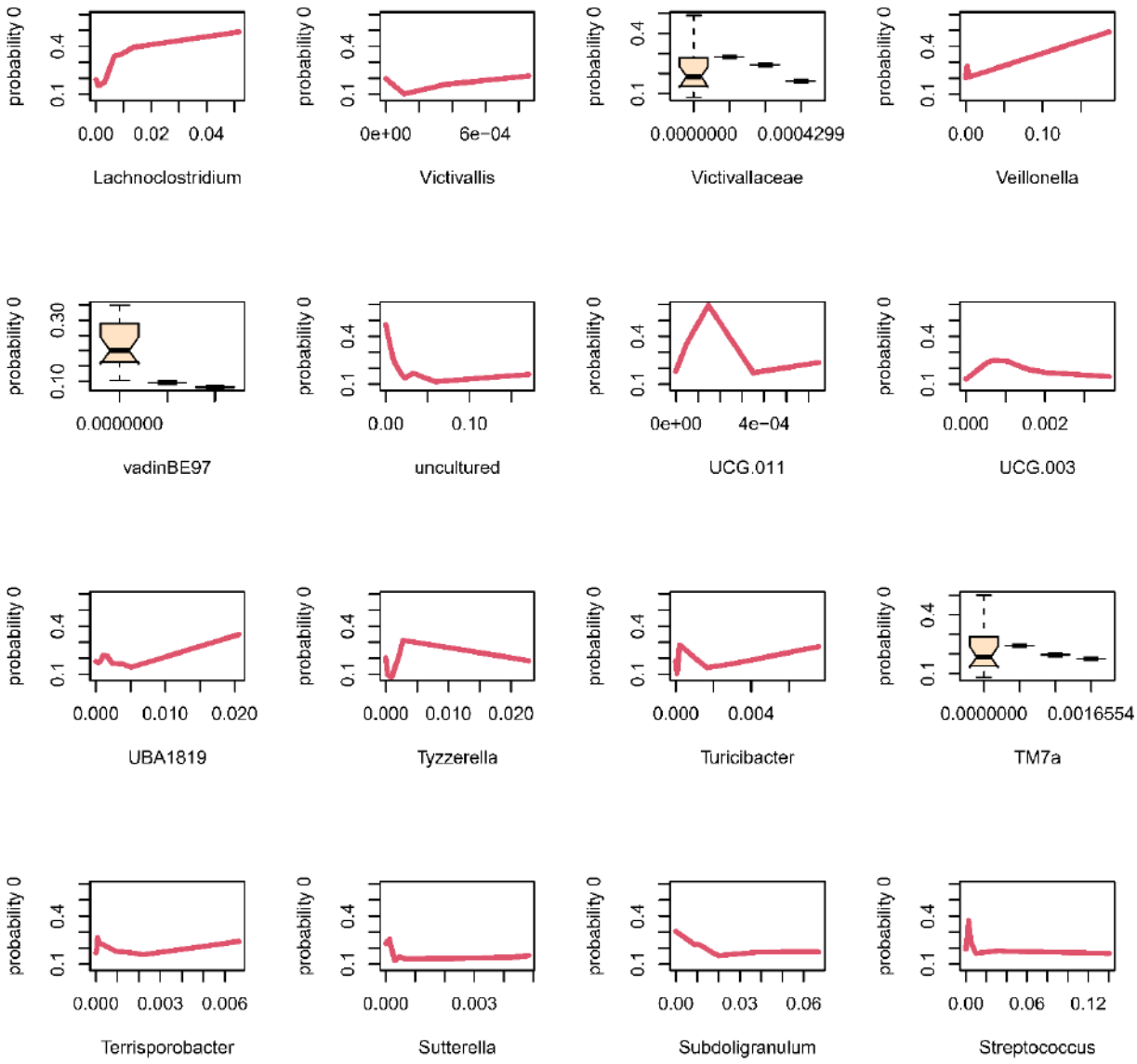

probability 0
0.1
0.4
0.00
0.03
0.06
Fusobacterium
0.0000
0.0008
Frisingicoccus
0.10
0.30
0.000e+00
Fournierella
0.0000
0.0010
Flavonifractor
0.0000
0.0010
Family_XIII_UCG.001
0.0000000
0.0017635
Faecalitalea
0.0000000
0.0021086
Eubacterium
0.000
0.010
Erysipelotrichaceae_UCG.003
0.000e+00
Epulopiscium
probability 0
0.000
0.015
0.030
Enterorhabdus
0.00
0.10
Enterococcus
0.0000
0.0015
Enterobacter
0.0000
0.0015
Eggerthella
0.0000
0.0015
DTU089
0.000
0.002
0.004
DTU014
0.000e+00
2.442e-03
Dielma

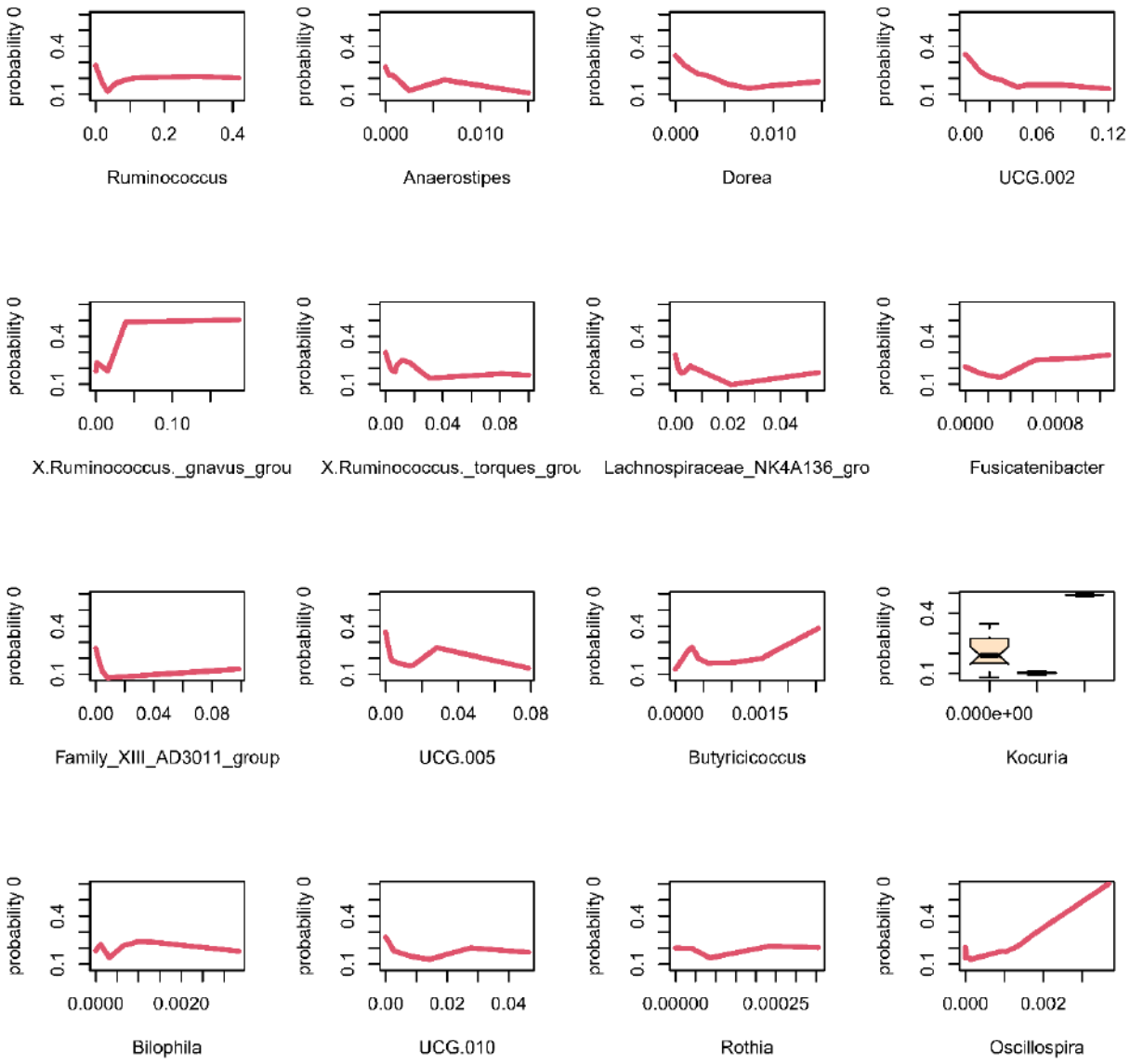

probability 0
0.1
0.4
0.0000
0.0008
Intestinimonas
probability 0
0.1
0.4
0.000
0.003
0.006
Intestinibacter
probability 0
0.1
0.4
0.000
0.004
0.008
Incertae_Sedis
probability 0
0.10
0.30
0.0000000
0.0012426
Hydrogenoanaerobacterium
probability 0
0.1
0.4
0.000
0.015
0.030
Hungatella
probability 0
0.1
0.4
0.000
0.006
Howardella
probability 0
0.1
0.4
0e+00
4e−04
Holdemania
probability 0
0.1
0.4
0.00
0.10
Holdemanella
probability 0
0.1
0.4
0.000000
Harryflintia
probability 0
0.1
0.4
0.0000000
Hafnia.Obesumbacterium
probability 0
0.1
0.4
0.000
0.003
Haemophilus
probability 0
0.1
0.4
0e+00
6e−04
Granulicatella
probability 0
0.10
0.30
0.000e+00
Gordonibacter
probability 0
0.1
0.4
0.000
0.006
Gemella
probability 0
0.1
0.4
0.00000
0.00020
GCA.900066575
probability 0
0.1
0.4
0.0000
0.0020
Gastranaerophilales


probability 0
0.1
0.4
0.000
0.002
Merdibacter
probability 0
0.1
0.4
0.000
0.002
0.004
Megasphaera
probability 0
0.10
0.30
0.000e+00
Megamonas
probability 0
0.1
0.4
0.0000
0.0015
Marvinbryantia
probability 0
0.1
0.4
0.000e+00
5.794e−04
Libanicoccus
probability 0
0.1
0.4
0.000e+00
Leuconostoc
probability 0
0.10
0.30
0.000e+00
2.137e−04
Leptotrichia
probability 0
0.1
0.4
0.0000
0.0015
Lactococcus
probability 0
0.1
0.4
0.000
0.006
Lactobacillus
probability 0
0.1
0.4
0.000e+00
2.028e−04
Lachnospiraceae_UCG.008
probability 0
0.1
0.4
0.0000
0.0010
Lachnospiraceae_UCG.004
probability 0
0.1
0.4
0.000e+00
Lachnospiraceae_UCG.001
probability 0
0.1
0.4
0.000
0.002
0.004
Lachnospiraceae_ND3007_grou
probability 0
0.1
0.4
0.0000
0.0008
Lachnospiraceae_FCS020_grou
probability 0
0.1
0.4
0.0000
0.0020
Lachnospira
probability 0
0.1
0.4
0.0
0.2
0.4
Klebsiella


probability 0
0.1
0.4
0.000
0.010
Peptostreptococcus
probability 0
0.10
0.30
0.0000000
Peptoniphilus
probability 0
0.1
0.4
0.0000
0.0020
Peptococcus
probability 0
0.1
0.4
0.00
0.02
0.04
Parvimonas
probability 0
0.1
0.4
0.00
0.04
0.08
Parabacteroides
probability 0
0.1
0.4
0.000
0.006
0.012
Oscillibacter
probability 0
0.1
0.4
0.000e+00
7.487e−03
Oribacterium
probability 0
0.1
0.4
0.000
0.002
0.004
Olsenella
probability 0
0.1
0.4
0.00
0.02
0.04
NK4A214_group
probability 0
0.1
0.4
0.0000000
Neisseria
probability 0
0.1
0.4
0.00
0.02
0.04
Muribaculaceae
probability 0
0.1
0.4
0.0000000
Morganella
probability 0
0.1
0.4
0.000
0.003
0.006
Monoglobus
probability 0
0.1
0.4
0.0000000
Mitsuokella
probability 0
0.10
0.30
0.000e+00
Mitochondria
probability 0
0.1
0.4
0.000e+00
1.325e−03
Methanosphaera


probability 0
0.1
0.4
0.0000
0.0008
Staphylococcus
0.000
0.015
0.030
Solobacterium
0.000
0.003
0.006
Slackia
0.0000
0.0020
Senegalimassilia
0e+00
6e-04
Sellimonas
0.0000000
0.0011876
Sanguibacteroides
0.0000
0.0015
Saccharimonadales
0.000
0.003
0.006
Saccharimonadaceae
0.0000000
S5.A14a
probability
0.000
0.002
0.004
Ruminiclostridium
0.00
0.02
0.04
Roseburia
0.000
0.015
0.030
Rikenellaceae_RC9_gut_grou
0.000e+00
1.320e-03
Pyramidobacter
0.00
0.04
0.08
Pseudomonas
0.00
0.04
0.08
Porphyromonas
0.000
0.010
Phascolarctobacterium


probability 0
0.1
0.4
0.00
0.02
0.04
Lachnoclostridium
0e+00
6e−04
Victivallis
0.0000000
0.0004299
Victivallaceae
0.00
0.10
Veillonella
0.10
0.30
vadinBE97
uncultured
UCG.011
0.000
0.002
UCG.003
0.010
0.020
UBA1819
Tyzzerella
0.004
Turicibacter
0.0016554
TM7a
0.003
0.006
Terrisporobacter
Sutterella
0.03
0.06
Subdoligranulum
0.12
Streptococcus


probability 0
0.1
0.4
0.0
0.2
0.4
Ruminococcus
0.000
0.010
Anaerostipes
Dorea
0.00
0.06
0.12
UCG.002
0.00
0.10
X.Ruminococcus._gnavus_grou
0.04
0.08
X.Ruminococcus._torques_grou
0.02
Lachnospiraceae_NK4A136_gro
0.0000
0.0008
Fusicatenibacter
Family_XIII_AD3011_group
probability
UCG.005
0.0015
Butyricicoccus
0.000e+00
Kocuria
0.0020
Bilophila
UCG.010
0.00000
0.00025
Rothia
0.002
Oscillospira

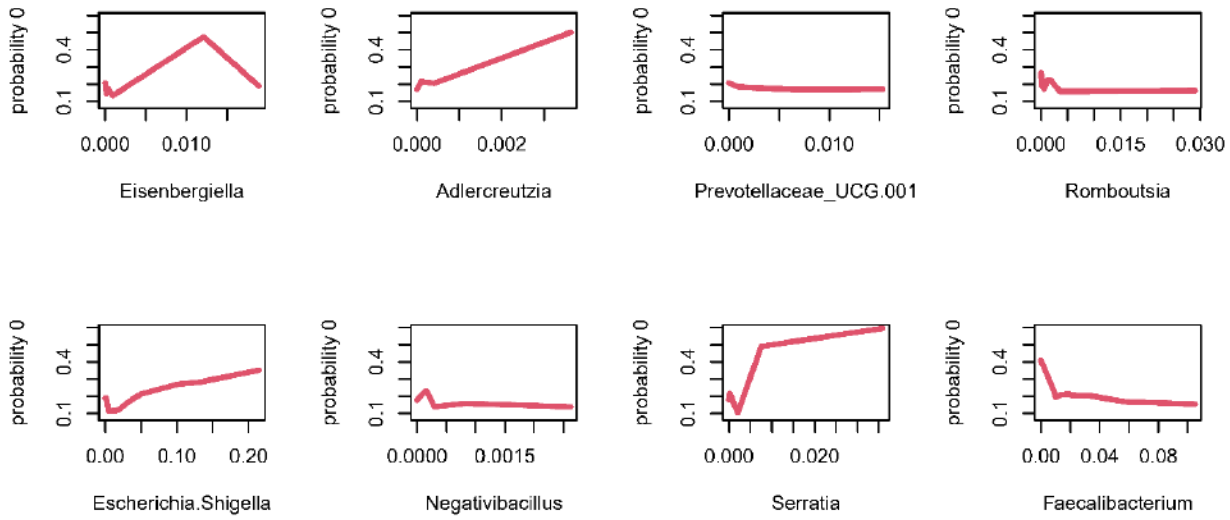
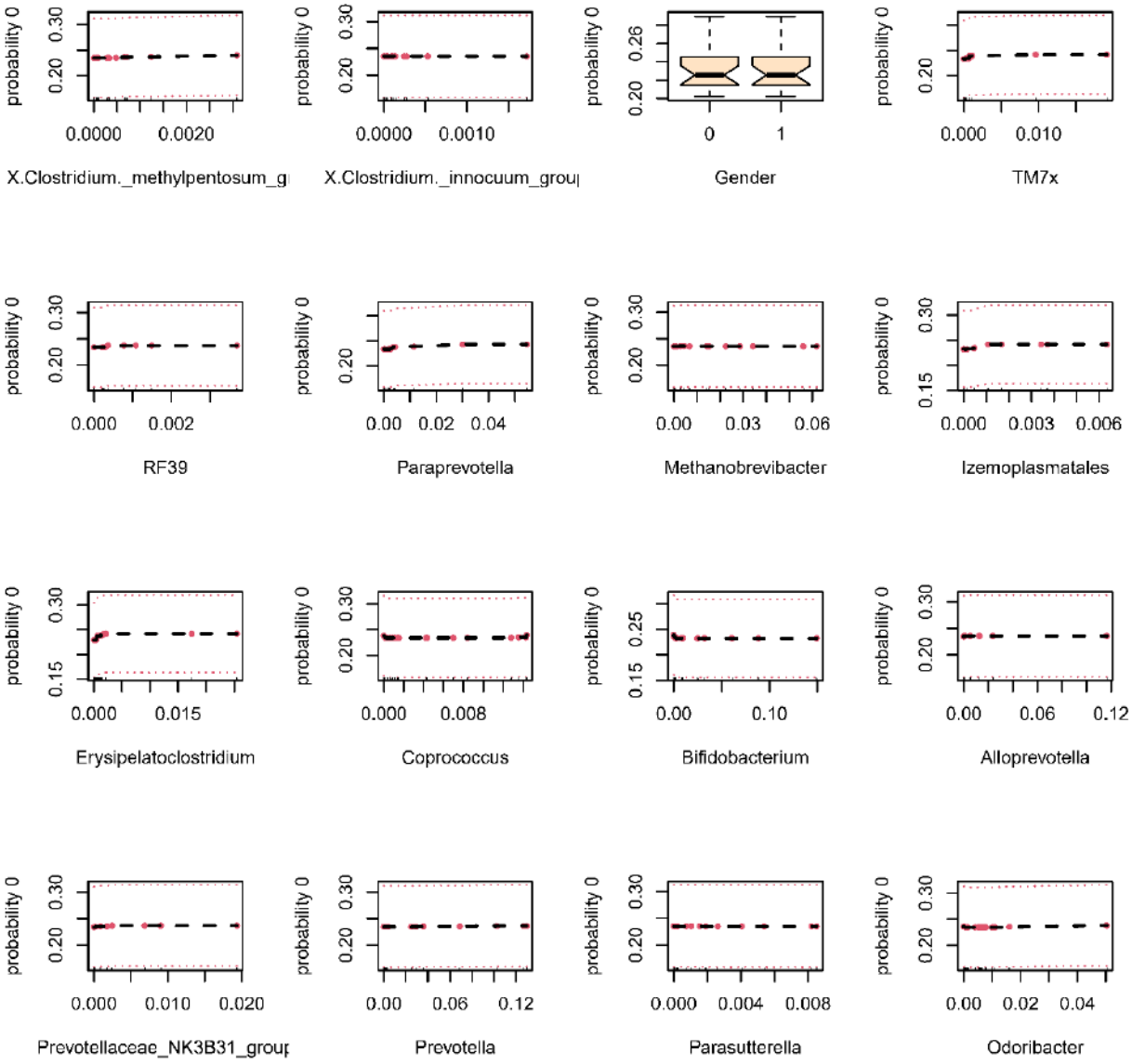
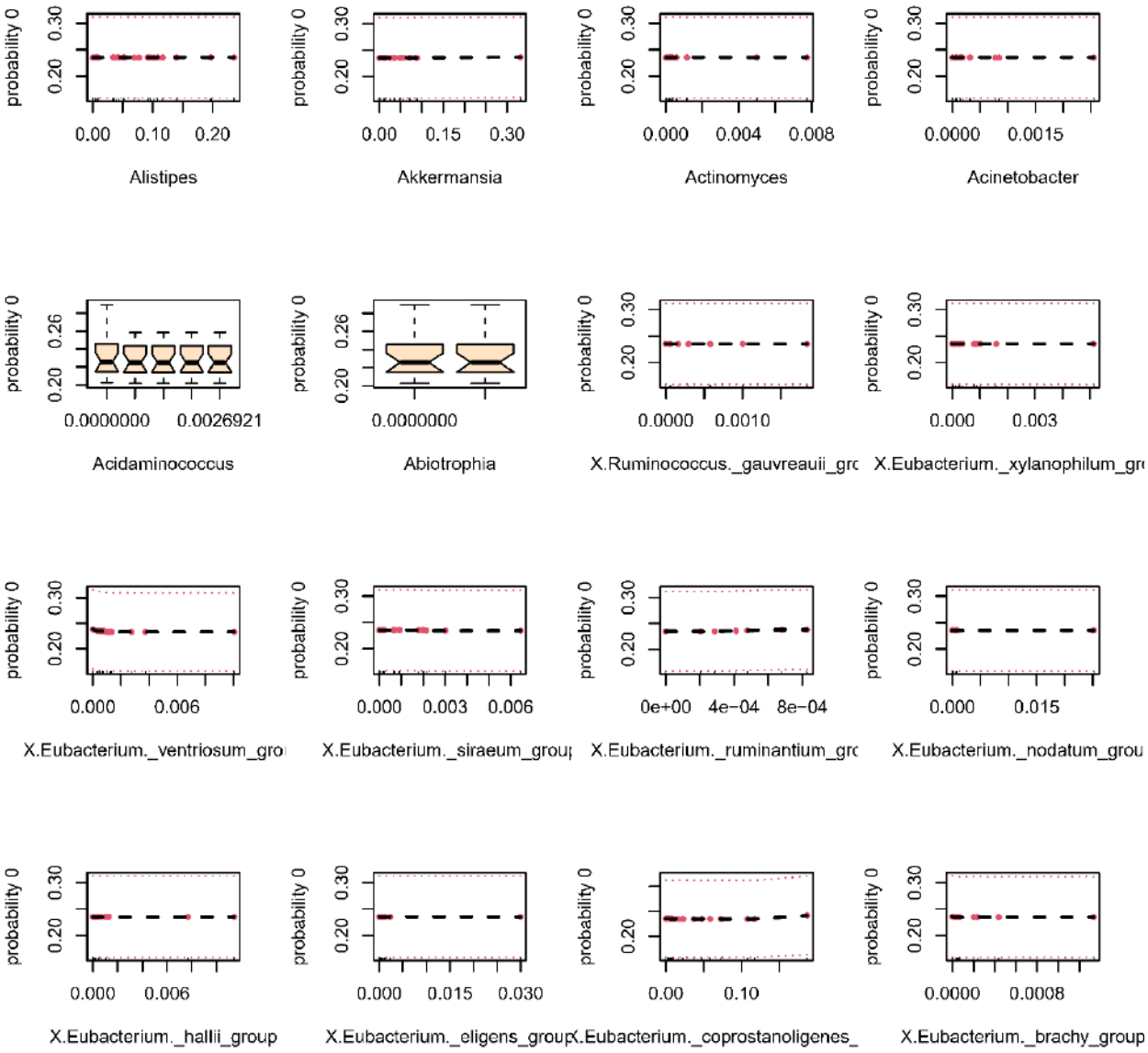
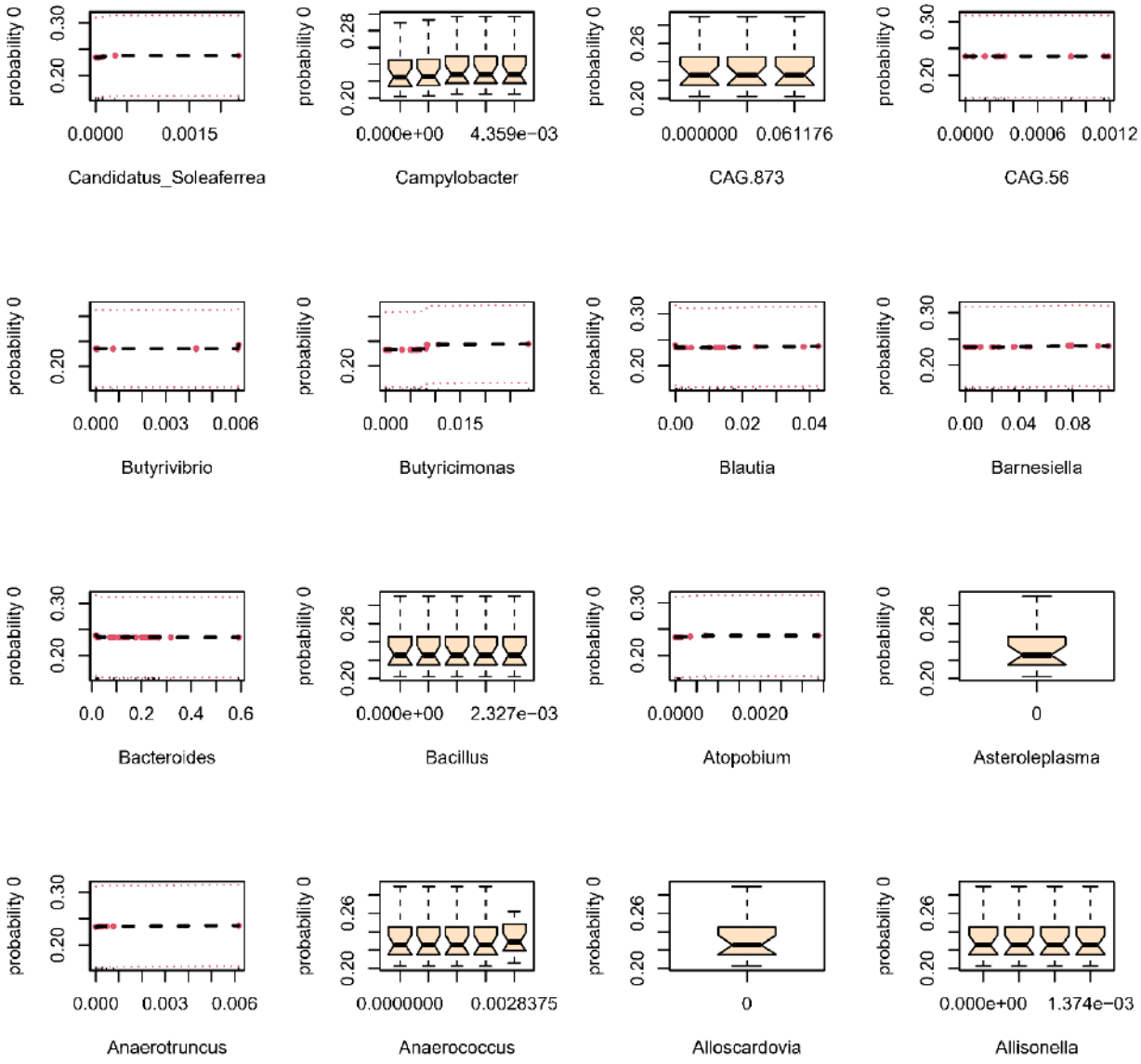
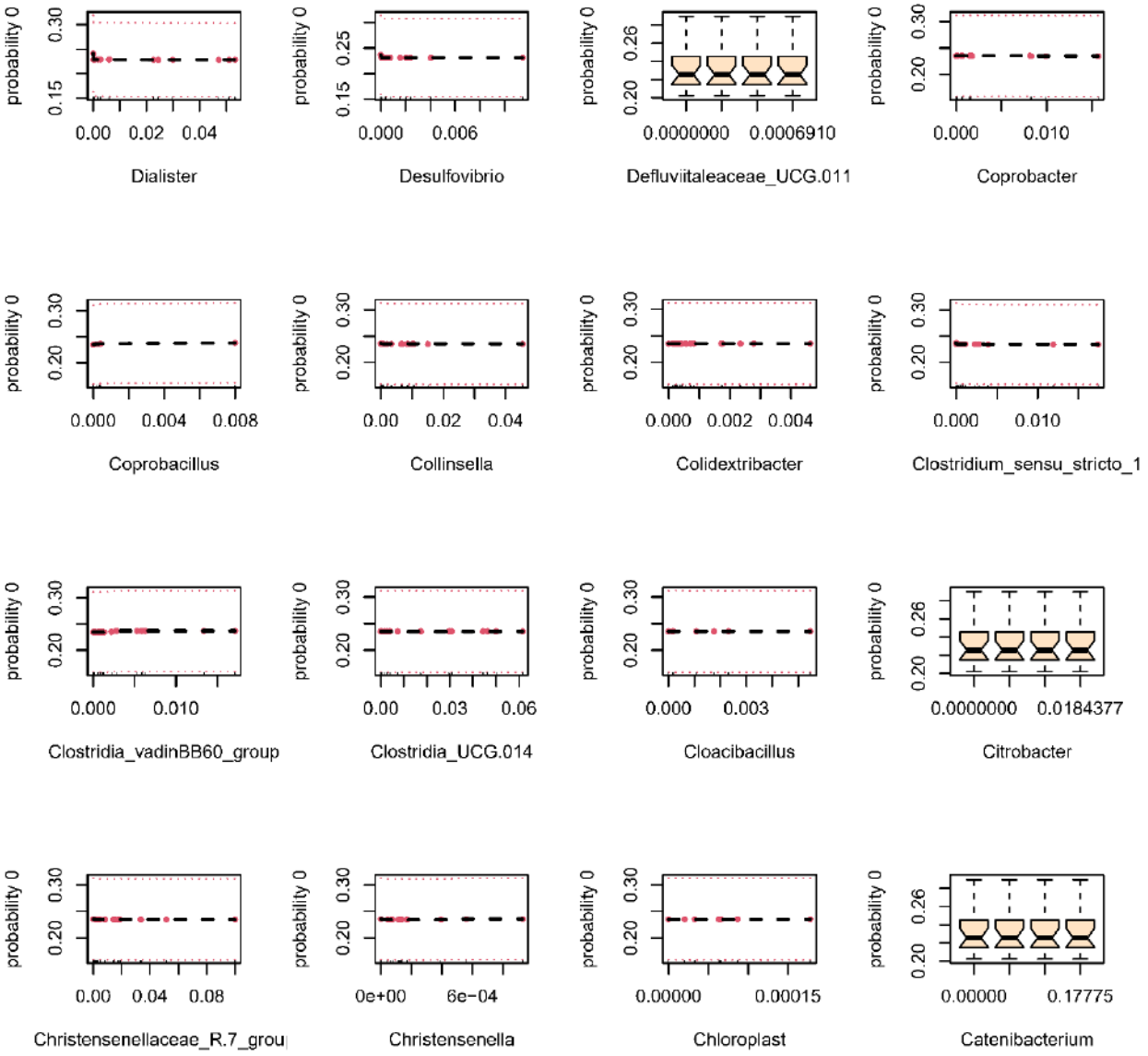


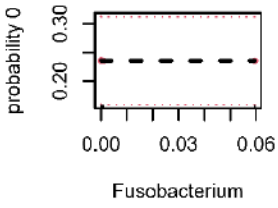
Figure ML2-E-2: Marginal Plot for Genus in Stool


probability 0
X.Clostridium._methylpentosum_g
X.Clostridium._innocuum_grou
Gender
TM7x
RF39
Paraprevotella
Methanobrevibacter
Izemoplasmatales
Erysipelatoclostridium
Coprococcus
Bifidobacterium
Alloprevotella
Prevotellaceae_NK3B31_grou
Prevotella
Parasutterella
Odoribacter

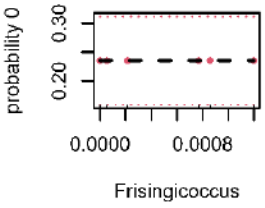

probability 0
0.20
0.30
0.00
0.10
0.20
Alistipes
probability 0
0.20
0.30
0.00
0.15
0.30
Akkermansia
probability 0
0.20
0.30
0.000
0.004
0.008
Actinomyces
probability 0
0.20
0.30
0.0000
0.0015
Acinetobacter
probability 0
0.20
0.26
0.0000000
0.0026921
Acidaminococcus
probability 0
0.20
0.26
0.0000000
Abiotrophia
probability 0
0.20
0.30
0.0000
0.0010
X.Ruminococcus._gauvreauii_grc
probability 0
0.20
0.30
0.000
0.003
X.Eubacterium._xylanophilum_gr
probability 0
0.20
0.30
0.000
0.006
X.Eubacterium._ventriosum_gro
probability 0
0.20
0.30
0.000
0.003
0.006
X.Eubacterium._siraeum_grou
probability 0
0.20
0.30
0e+00
4e−04
8e−04
X.Eubacterium._ruminantium_grc
probability 0
0.20
0.30
0.000
0.015
X.Eubacterium._nodatum_grou
probability 0
0.20
0.30
0.000
0.006
X.Eubacterium._hallii_group
probability 0
0.20
0.30
0.000
0.015
0.030
X.Eubacterium._eligens_group
probability 0
0.20
0.00
0.10
X.Eubacterium._coprostanoligenes_
probability 0
0.20
0.30
0.0000
0.0008
X.Eubacterium._brachy_group


probability 0
0.20
0.30
0.0000
0.0015
Candidatus_Soleaferrea
probability 0
0.20
0.28
0.000e+00
4.359e−03
Campylobacter
probability 0
0.20
0.26
0.000000
0.061176
CAG.873
probability 0
0.20
0.30
0.0000
0.0006
0.0012
CAG.56
probability 0
0.20
0.000
0.003
0.006
Butyrivibrio
probability 0
0.20
0.000
0.015
Butyricimonas
probability 0
0.20
0.30
0.00
0.02
0.04
Blautia
probability 0
0.20
0.30
0.00
0.04
0.08
Barnesiella
probability 0
0.20
0.30
0.0
0.2
0.4
0.6
Bacteroides
probability 0
0.20
0.26
0.000e+00
2.327e−03
Bacillus
probability 0
0.20
0.30
0.0000
0.0020
Atopobium
probability 0
0.20
0.26
0
Asteroleplasma
probability 0
0.20
0.30
0.000
0.003
0.006
Anaerotruncus
probability 0
0.20
0.26
0.0000000
0.0028375
Anaerococcus
probability 0
0.20
0.26
0
Alloscardovia
probability 0
0.20
0.26
0.000e+00
1.374e−03
Allisonella

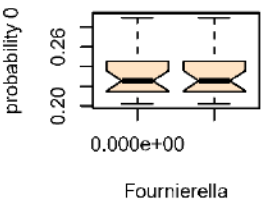

probability 0
Dialister
Desulfovibrio
Defluviitaleaceae_UCG.011
Coprobacter
Coprobacillus
Collinsella
Colidextribacter
Clostridium_sensu_stricto_1
Clostridia_vadinBB60_group
Clostridia_UCG.014
Cloacibacillus
Citrobacter
Christensenellaceae_R.7_grou
Christensenella
Chloroplast
Catenibacterium



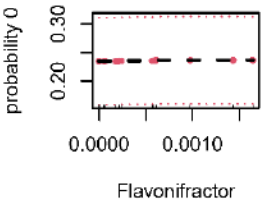
Fusobacterium



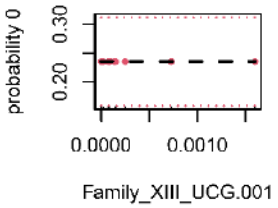
Frisingicoccus



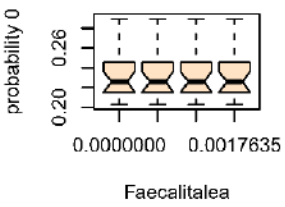
Fournierella



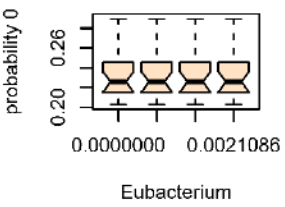
Flavonifractor



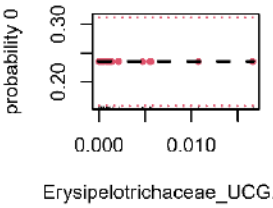
Family_XIII_UCG.001



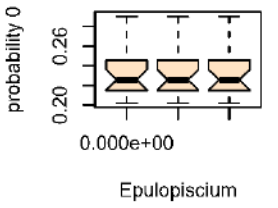
Faecalitalea



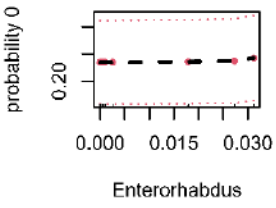
Eubacterium



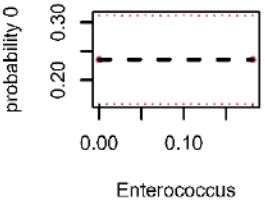
Erysipelotrichaceae_UCG.003



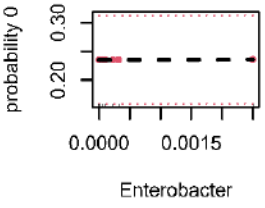
Epulopiscium



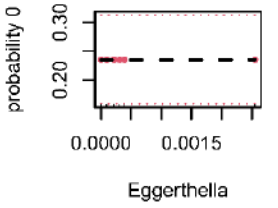
Enterorhabdus



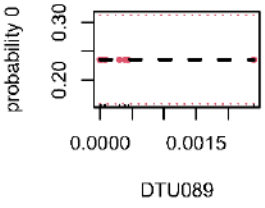
Enterococcus



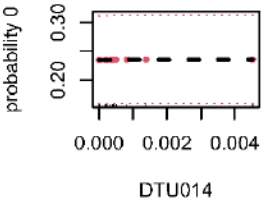
Enterobacter



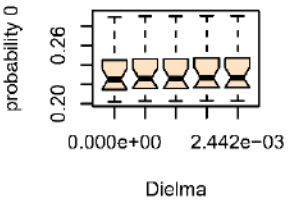
Eggerthella



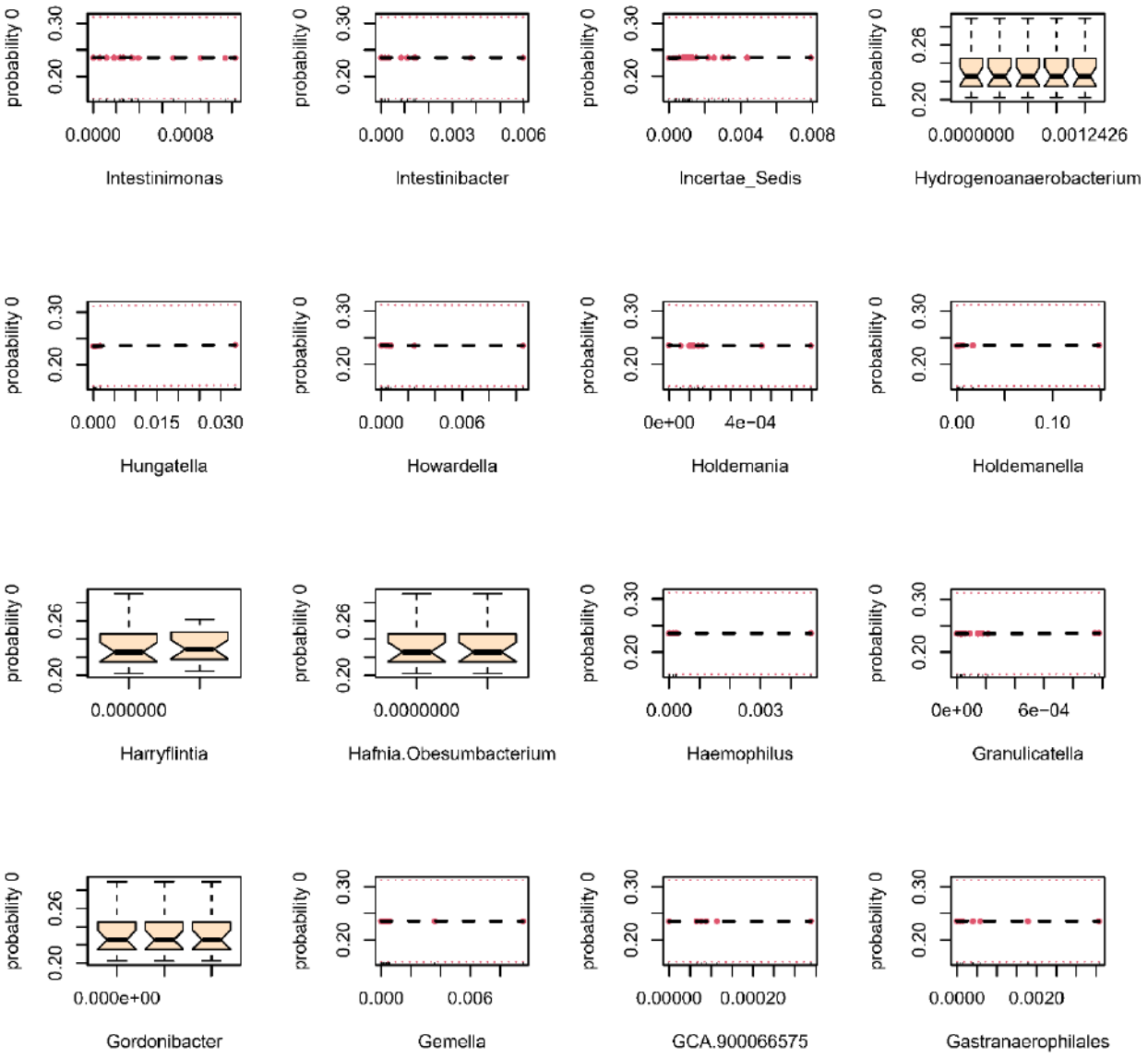
DTU089

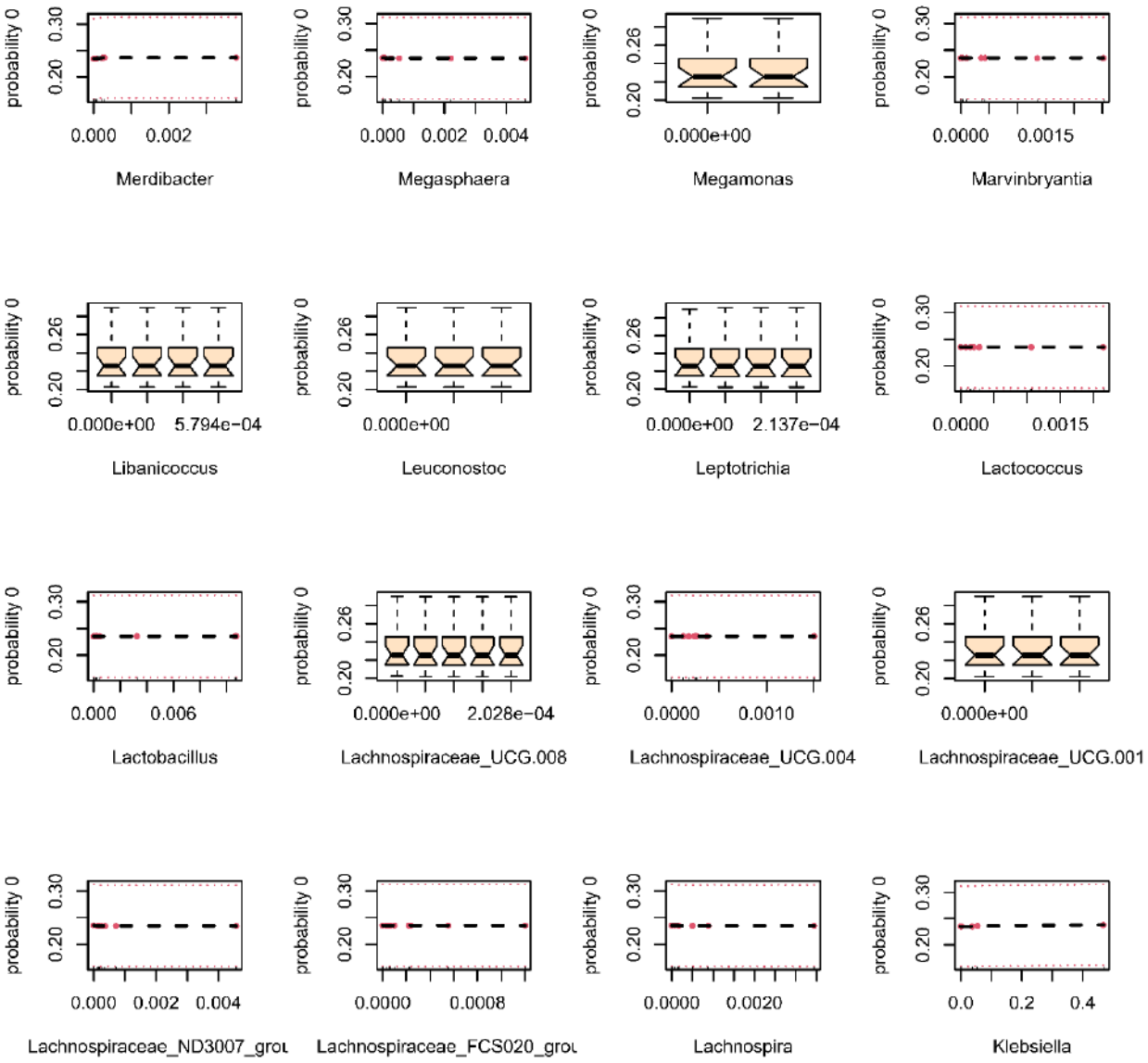


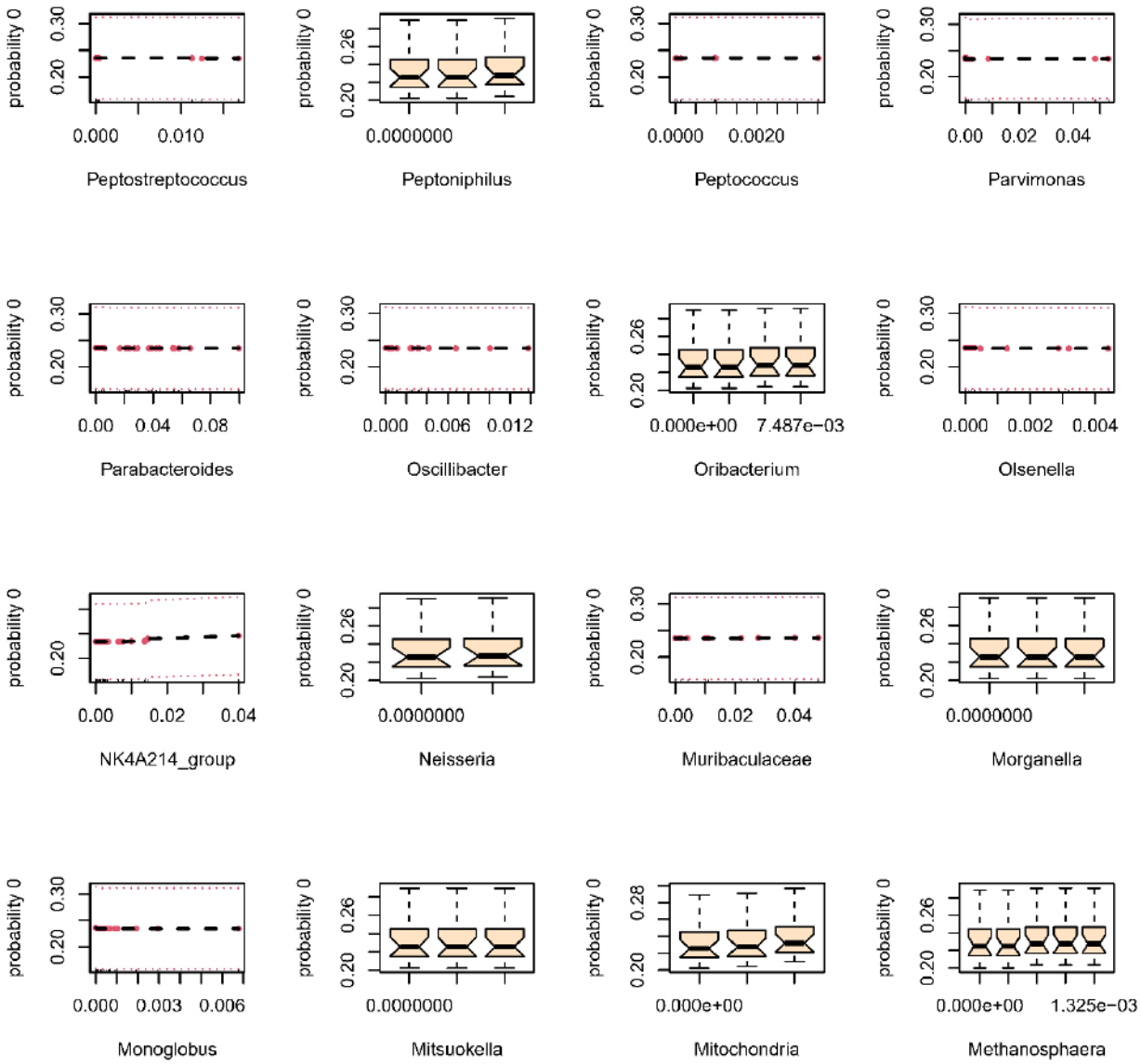
DTU014

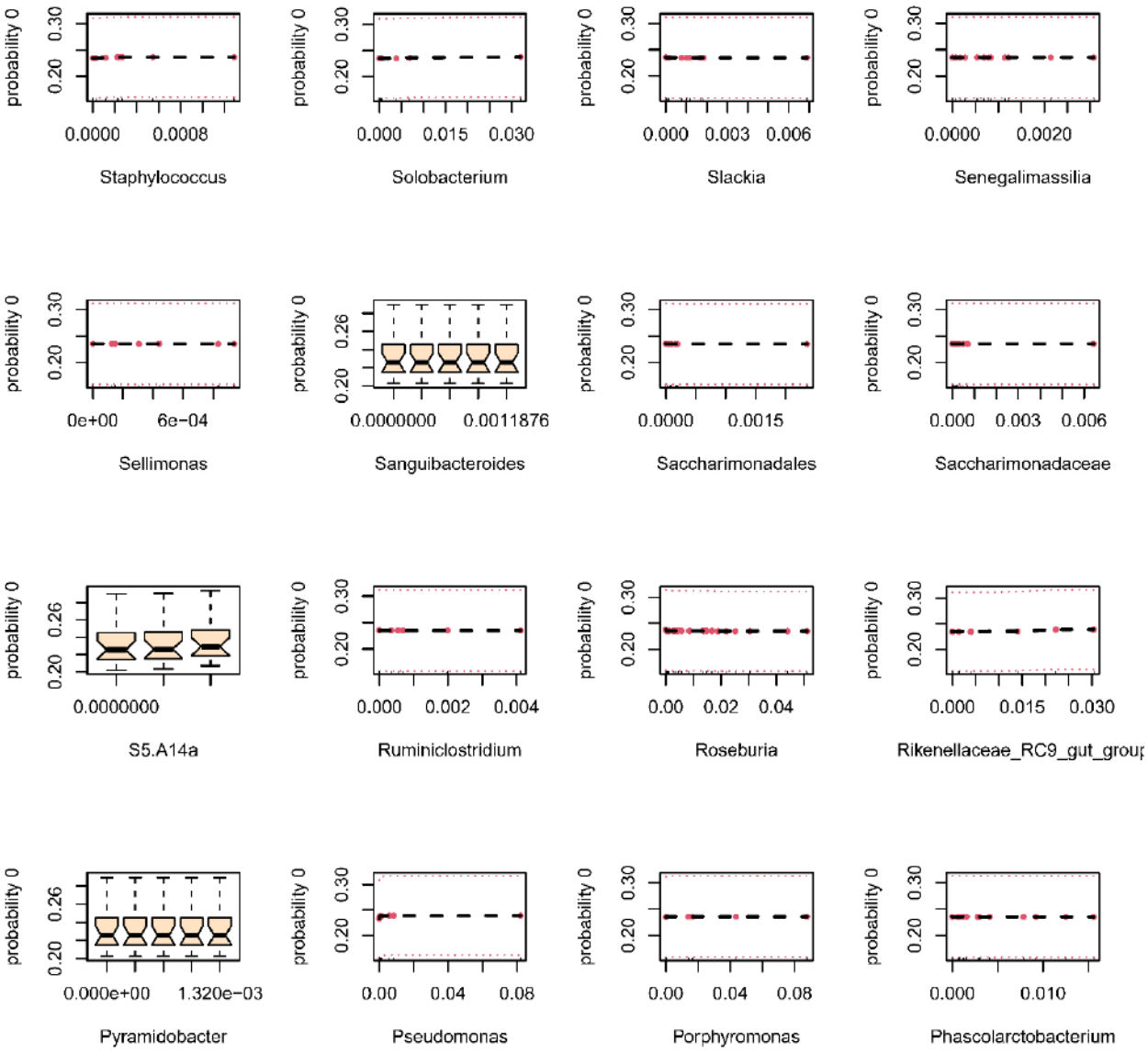


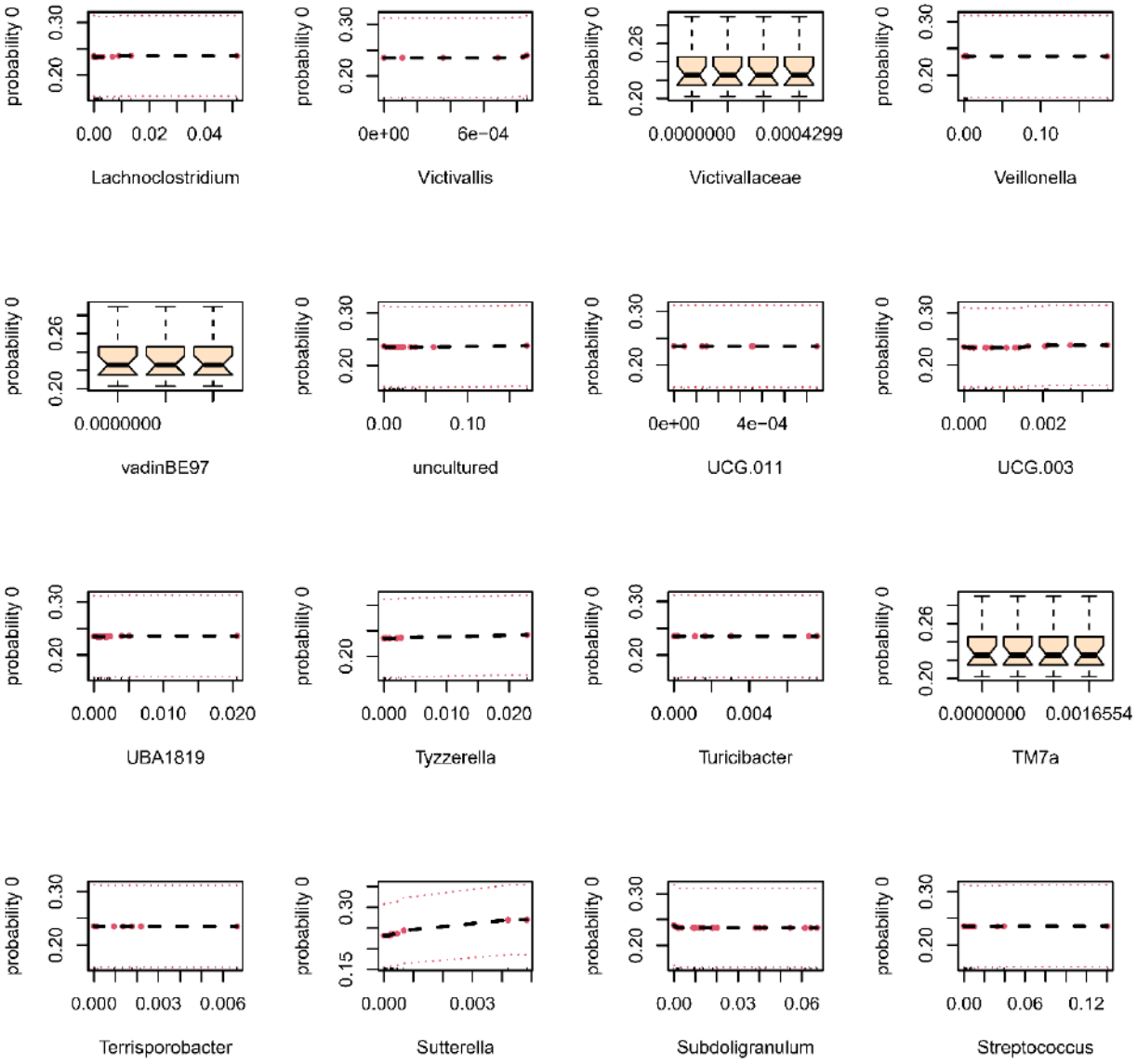
Dielma

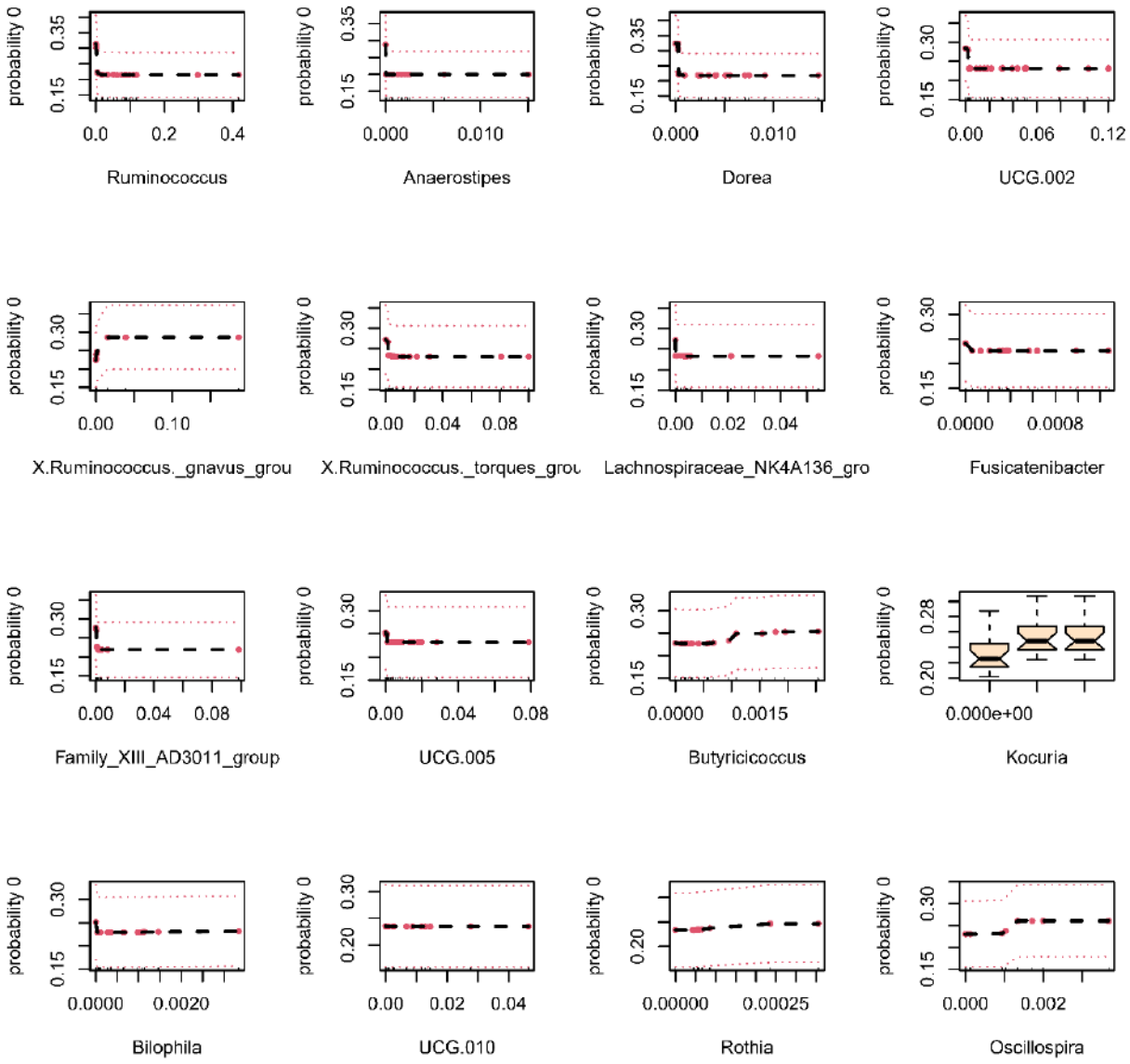

probability 0
0.20
0.30
0.0000
0.0008
Intestinimonas
probability 0
0.20
0.30
0.000
0.003
0.006
Intestinibacter
probability 0
0.20
0.30
0.000
0.004
0.008
Incertae_Sedis
probability 0
0.20
0.26
0.0000000
0.0012426
Hydrogenoanaerobacterium
probability 0
0.20
0.30
0.000
0.015
0.030
Hungatella
probability 0
0.20
0.30
0.000
0.006
Howardella
probability 0
0.20
0.30
0e+00
4e−04
Holdemania
probability 0
0.20
0.30
0.00
0.10
Holdemanella
probability 0
0.20
0.26
0.000000
Harryflintia
probability 0
0.20
0.26
0.0000000
Hafnia.Obesumbacterium
probability 0
0.20
0.30
0.000
0.003
Haemophilus
probability 0
0.20
0.30
0e+00
6e−04
Granulicatella
probability 0
0.20
0.26
0.000e+00
Gordonibacter
probability 0
0.20
0.30
0.000
0.006
Gemella
probability 0
0.20
0.30
0.00000
0.00020
GCA.900066575
probability 0
0.20
0.30
0.0000
0.0020
Gastranaerophilales


probability 0
0.20
0.30
0.000
0.002
Merdibacter
probability 0
0.20
0.30
0.000
0.002
0.004
Megasphaera
probability 0
0.20
0.26
0.000e+00
Megamonas
probability 0
0.20
0.30
0.0000
0.0015
Marvinbryantia
probability 0
0.20
0.26
0.000e+00
5.794e−04
Libanicoccus
probability 0
0.20
0.26
0.000e+00
Leuconostoc
probability 0
0.20
0.26
0.000e+00
2.137e−04
Leptotrichia
probability 0
0.20
0.30
0.0000
0.0015
Lactococcus
probability 0
0.20
0.30
0.000
0.006
Lactobacillus
probability 0
0.20
0.26
0.000e+00
2.028e−04
Lachnospiraceae_UCG.008
probability 0
0.20
0.30
0.0000
0.0010
Lachnospiraceae_UCG.004
probability 0
0.20
0.26
0.000e+00
Lachnospiraceae_UCG.001
probability 0
0.20
0.30
0.000
0.002
0.004
Lachnospiraceae_ND3007_grou
probability 0
0.20
0.30
0.0000
0.0008
Lachnospiraceae_FCS020_grou
probability 0
0.20
0.30
0.0000
0.0020
Lachnospira
probability 0
0.20
0.30
0.0
0.2
0.4
Klebsiella


probability 0
0.20
0.30
0.000
0.010
Peptostreptococcus
probability 0
0.20
0.26
0.0000000
Peptoniphilus
probability 0
0.20
0.30
0.0000
0.0020
Peptococcus
probability 0
0.20
0.30
0.00
0.02
0.04
Parvimonas
probability 0
0.20
0.30
0.00
0.04
0.08
Parabacteroides
probability 0
0.20
0.30
0.000
0.006
0.012
Oscillibacter
probability 0
0.20
0.26
0.000e+00
7.487e−03
Oribacterium
probability 0
0.20
0.30
0.000
0.002
0.004
Olsenella
probability 0
0.20
0.00
0.02
0.04
NK4A214_group
probability 0
0.20
0.26
0.0000000
Neisseria
probability 0
0.20
0.30
0.00
0.02
0.04
Muribaculaceae
probability 0
0.20
0.26
0.0000000
Morganella
probability 0
0.20
0.30
0.000
0.003
0.006
Monoglobus
probability 0
0.20
0.26
0.0000000
Mitsuokella
probability 0
0.20
0.28
0.000e+00
Mitochondria
probability 0
0.20
0.26
0.000e+00
1.325e−03
Methanosphaera


probability 0
0.20
0.30
0.0000
0.0008
Staphylococcus
probability 0
0.20
0.30
0.000
0.015
0.030
Solobacterium
probability 0
0.20
0.30
0.000
0.003
0.006
Slackia
probability 0
0.20
0.30
0.0000
0.0020
Senegalimassilia
probability 0
0.20
0.30
0e+00
6e-04
Sellimonas
probability 0
0.20
0.26
0.0000000
0.0011876
Sanguibacteroides
probability 0
0.20
0.30
0.0000
0.0015
Saccharimonadales
probability 0
0.20
0.30
0.000
0.003
0.006
Saccharimonadaceae
probability 0
0.20
0.26
0.0000000
S5.A14a
probability 0
0.20
0.30
0.000
0.002
0.004
Ruminiclostridium
probability 0
0.20
0.30
0.00
0.02
0.04
Roseburia
probability 0
0.20
0.30
0.000
0.015
0.030
Rikenellaceae_RC9_gut_group
probability 0
0.20
0.26
0.000e+00
1.320e-03
Pyramidobacter
probability 0
0.20
0.30
0.00
0.04
0.08
Pseudomonas
probability 0
0.20
0.30
0.00
0.04
0.08
Porphyromonas
probability 0
0.20
0.30
0.000
0.010
Phascolarctobacterium


probability 0
Lachnoclostridium
Victivallis
Victivallaceae
Veillonella
vadinBE97
uncultured
UCG.011
UCG.003
UBA1819
Tyzzerella
Turicibacter
TM7a
Terrisporobacter
Sutterella
Subdoligranulum
Streptococcus


probability 0
Ruminococcus
Anaerostipes
Dorea
UCG.002
X.Ruminococcus._gnavus_grou
X.Ruminococcus._torques_grou
Lachnospiraceae_NK4A136_gro
Fusicatenibacter
Family_XIII_AD3011_group
UCG.005
Butyricicoccus
Kocuria
Bilophila
UCG.010
Rothia
Oscillospira

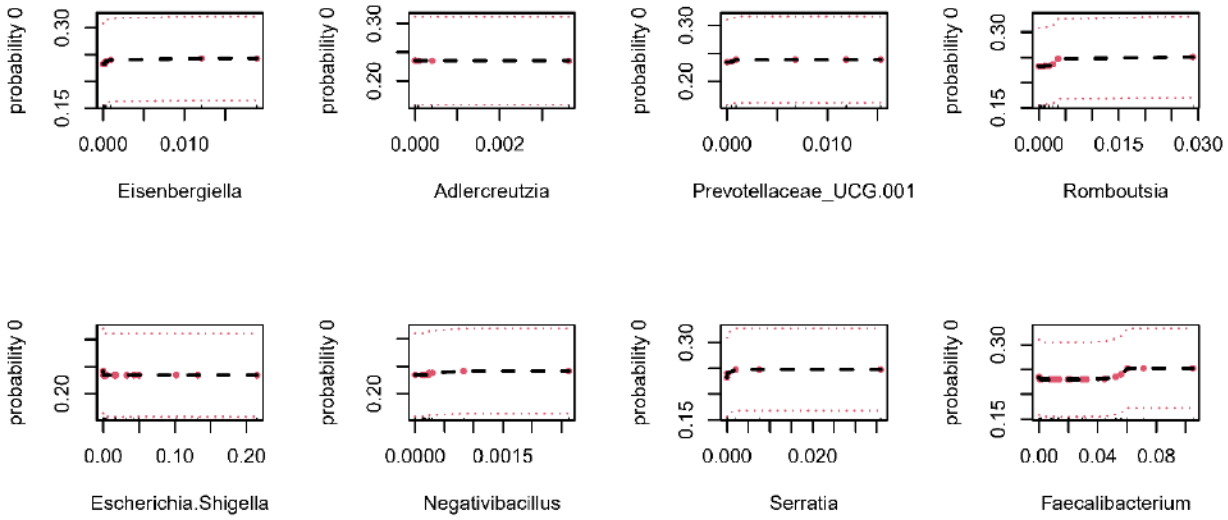


Figure ML2-E-3: Partial Plot for Genus in Stool

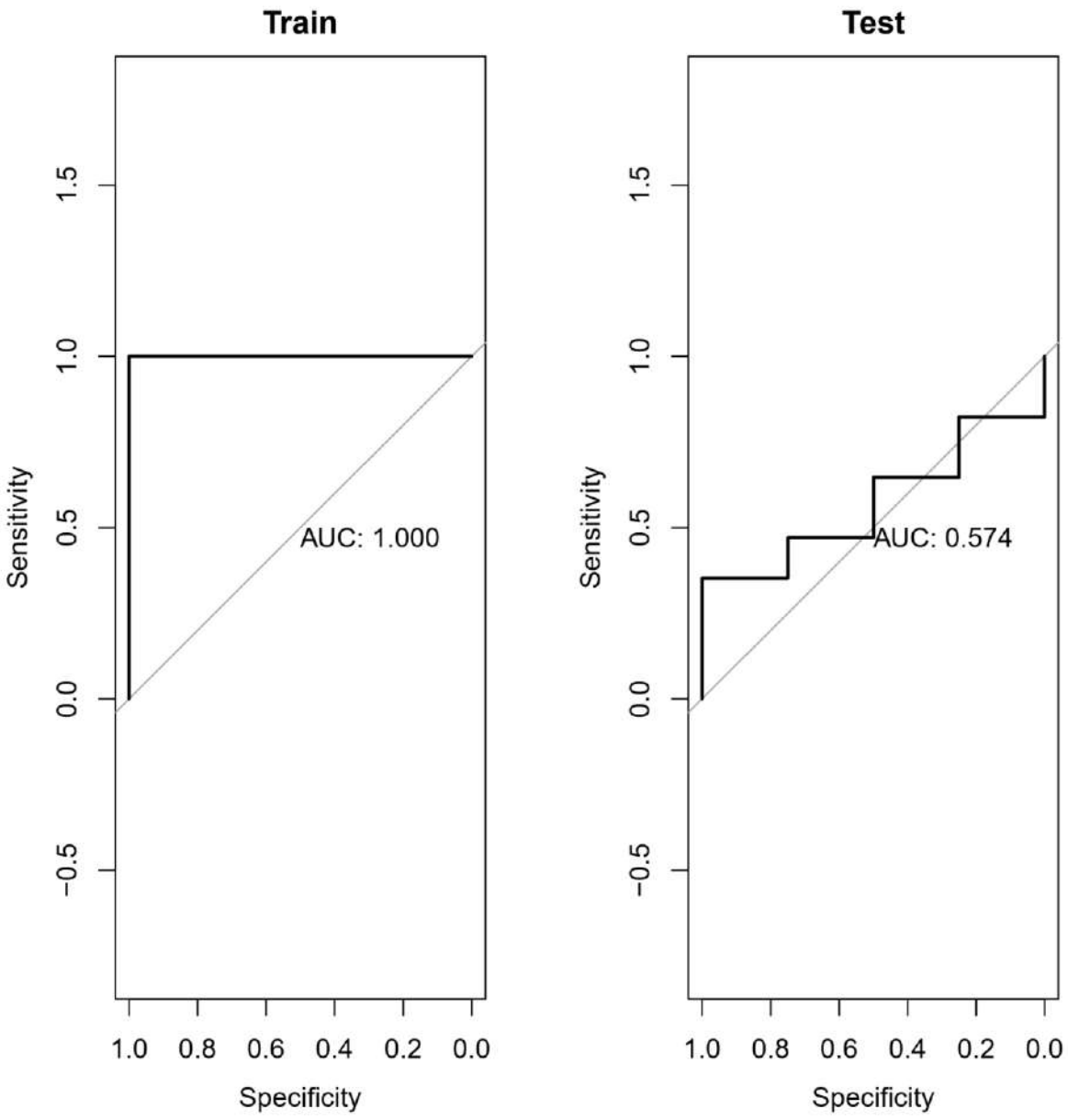


Figure ML2-E-4: The ROC Curve for Genus in Stool

Table ML2-E-1: The Model Performance

Index	Values		
	2.50%	50.00%	97.50%
Accuracy	0.33	0.52	0.86
Specificity	0.50	1.00	1.00
Sensitivity	0.18	0.41	0.88

Table ML2-E-2: The Variable Importance Values

Covariates	Varibale Importance			Brier Score		
	all	AP	CRC	all	AP	CRC
Gender	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
X.Clostridium._innocuum_group	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
X.Clostridium._methylpentosum_group	-0.0004	-0.0217	0.0048	0.0058	0.0040	0.0040
X.Eubacterium._brachy_group	-0.0009	0.0000	-0.0032	-0.0010	-0.0007	-0.0007
X.Eubacterium._coprostanoligenes_group	0.0014	-0.0217	0.0112	0.0084	0.0057	0.0057
X.Eubacterium._eligens_group	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
X.Eubacterium._hallii_group	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
X.Eubacterium._nodatum_group	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
X.Eubacterium._ruminantium_group	0.0005	-0.0054	0.0032	-0.0005	-0.0003	-0.0003
X.Eubacterium._siraeum_group	-0.0005	0.0000	-0.0016	-0.0017	-0.0012	-0.0012
X.Eubacterium._ventriosum_group	-0.0005	0.0054	-0.0032	-0.0019	-0.0013	-0.0013
X.Eubacterium._xylanophilum_group	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
X.Ruminococcus._gauvreauii_group	0.0000	0.0000	0.0000	-0.0004	-0.0003	-0.0003
X.Ruminococcus._gnavus_group	0.0312	0.0054	0.1071	0.0496	0.0337	0.0337
X.Ruminococcus._torques_group	0.0241	-0.0109	0.0863	0.0383	0.0261	0.0261
Abiotrophia	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Acidaminococcus	-0.0005	0.0000	-0.0016	-0.0012	-0.0008	-0.0008
Acinetobacter	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Actinomyces	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Adlercreutzia	-0.0005	0.0000	-0.0016	-0.0009	-0.0006	-0.0006
Akkermansia	0.0000	0.0000	0.0000	-0.0014	-0.0009	-0.0009
Alistipes	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Allisonella	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Alloprevotella	0.0000	-0.0054	0.0016	-0.0009	-0.0006	-0.0006
Alloscardovia	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Anaerococcus	0.0009	-0.0217	0.0096	0.0027	0.0019	0.0019
Anaerostipes	0.0294	0.1577	0.0560	0.1215	0.0825	0.0825
Anaerotruncus	0.0005	0.0000	0.0016	-0.0018	-0.0012	-0.0012
Asteroleplasma	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Atopobium	-0.0005	-0.0163	0.0032	-0.0012	-0.0008	-0.0008
Bacillus	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Bacteroides	0.0015	0.0000	0.0048	0.0036	0.0024	0.0024
Barnesiella	-0.0018	-0.0163	-0.0016	-0.0015	-0.0011	-0.0011
Bifidobacterium	-0.0028	-0.0109	-0.0064	-0.0040	-0.0027	-0.0027
Bilophila	0.0060	-0.0544	0.0368	0.0252	0.0171	0.0171
Blautia	0.0018	-0.0109	0.0096	0.0003	0.0002	0.0002
Butyricicoccus	0.0078	-0.0544	0.0432	0.0043	0.0029	0.0029
Butyricimonas	0.0005	-0.0217	0.0080	0.0032	0.0022	0.0022
Butyrivibrio	0.0019	-0.0217	0.0128	0.0030	0.0020	0.0020
CAG.56	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
CAG.873	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Campylobacter	0.0000	-0.0054	0.0016	-0.0033	-0.0022	-0.0022
Candidatus_Soleaferrea	-0.0009	-0.0054	-0.0016	-0.0004	-0.0003	-0.0003
Catenibacterium	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Chloroplast	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Christensenella	-0.0014	-0.0109	-0.0016	-0.0034	-0.0023	-0.0023
Christensenellaceae_R.7_group	-0.0005	0.0000	-0.0016	-0.0010	-0.0007	-0.0007
Citrobacter	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Cloacibacillus	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Clostridia_UCG.014	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Clostridia_vadinBB60_group	0.0005	-0.0054	0.0032	-0.0028	-0.0019	-0.0019
Clostridium_sensu_stricto_1	0.0014	-0.0054	0.0064	-0.0045	-0.0031	-0.0031
Colidextribacter	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Collinsella	-0.0009	-0.0163	0.0016	-0.0011	-0.0008	-0.0008
Coprobacillus	0.0023	0.0000	0.0080	0.0025	0.0017	0.0017
Coprobacter	0.0005	0.0054	0.0000	-0.0017	-0.0011	-0.0011
Coprococcus	0.0023	-0.0109	0.0112	0.0073	0.0049	0.0049
Defluviitaleaceae_UCG.011	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Desulfovibrio	-0.0023	0.0109	-0.0112	-0.0008	-0.0005	-0.0005
Dialister	0.0014	0.0000	0.0048	-0.0052	-0.0035	-0.0035
Dielma	-0.0005	0.0000	-0.0016	-0.0003	-0.0002	-0.0002
Dorea	0.0495	-0.0163	0.1775	0.1415	0.0962	0.0962
DTU014	-0.0014	-0.0054	-0.0032	-0.0012	-0.0008	-0.0008
DTU089	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Eggerthella	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Eisenbergiella	-0.0004	-0.0217	0.0048	0.0027	0.0018	0.0018
Enterobacter	0.0005	0.0000	0.0016	0.0001	0.0000	0.0000
Enterococcus	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Enterorhabdus	0.0018	-0.0163	0.0112	0.0025	0.0017	0.0017
Epulopiscium	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Erysipelatoclostridium	-0.0037	0.0109	-0.0160	-0.0058	-0.0039	-0.0039
Erysipelotrichaceae_UCG.003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Escherichia.Shigella	0.0009	0.0000	0.0032	-0.0008	-0.0006	-0.0006
Eubacterium	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Faecalibacterium	0.0045	-0.0652	0.0352	0.0104	0.0071	0.0071
Faecalitalea	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Family_XIII_AD3011_group	0.0215	0.0217	0.0688	0.0463	0.0315	0.0315
Family_XIII_UCG.001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Flavonifractor	-0.0009	-0.0054	-0.0016	-0.0022	-0.0015	-0.0015
Fournierella	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Frisingicoccus	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Fusicatenibacter	0.0009	0.0544	-0.0128	0.0020	0.0013	0.0013
Fusobacterium	0.0005	0.0000	0.0016	-0.0011	-0.0008	-0.0008
Gastranaerophilales	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
GCA.900066575	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Gemella	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Gordonibacter	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Granulicatella	0.0009	0.0000	0.0032	0.0019	0.0013	0.0013
Haemophilus	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Hafnia.Obesumbacterium	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Harryflintia	0.0014	-0.0054	0.0064	-0.0010	-0.0007	-0.0007
Holdemanella	-0.0005	-0.0054	0.0000	-0.0013	-0.0009	-0.0009
Holdemania	-0.0005	0.0000	-0.0016	-0.0019	-0.0013	-0.0013
Howardella	-0.0005	0.0000	-0.0016	-0.0025	-0.0017	-0.0017
Hungatella	0.0000	-0.0109	0.0032	0.0018	0.0012	0.0012
Hydrogenoanaerobacterium	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Incertae_Sedis	-0.0005	0.0000	-0.0016	-0.0008	-0.0006	-0.0006
Intestinibacter	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Intestinimonas	0.0005	0.0054	0.0000	0.0000	0.0000	0.0000
Izemoplasmatales	-0.0041	-0.0435	-0.0016	-0.0057	-0.0039	-0.0039
Klebsiella	0.0000	-0.0163	0.0048	-0.0034	-0.0023	-0.0023
Kocuria	0.0036	-0.0761	0.0352	0.0154	0.0104	0.0104
Lachnoclostridium	-0.0001	0.0000	0.0000	-0.0044	-0.0030	-0.0030
Lachnospira	0.0000	0.0054	-0.0016	-0.0004	-0.0003	-0.0003
Lachnospiraceae_FCS020_group	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Lachnospiraceae_ND3007_group	0.0000	0.0000	0.0000	-0.0002	-0.0001	-0.0001
Lachnospiraceae_NK4A136_group	0.0166	-0.0381	0.0688	0.0512	0.0348	0.0348
Lachnospiraceae_UCG.001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Lachnospiraceae_UCG.004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Lachnospiraceae_UCG.008	-0.0005	0.0000	-0.0016	-0.0006	-0.0004	-0.0004
Lactobacillus	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Lactococcus	-0.0009	0.0000	-0.0032	-0.0005	-0.0003	-0.0003
Leptotrichia	-0.0005	0.0000	-0.0016	-0.0009	-0.0006	-0.0006
Leuconostoc	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Libanicoccus	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Marvinbryantia	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Megamonas	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Megasphaera	-0.0005	0.0000	-0.0016	-0.0012	-0.0008	-0.0008
Merdibacter	0.0009	0.0000	0.0032	-0.0007	-0.0005	-0.0005
Methanobrevibacter	0.0000	-0.0054	0.0016	0.0004	0.0003	0.0003
Methanosphaera	0.0005	0.0000	0.0016	-0.0001	-0.0001	-0.0001
Mitochondria	-0.0004	-0.0272	0.0064	0.0044	0.0030	0.0030
Mitsuokella	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Monoglobus	-0.0010	0.0000	-0.0032	-0.0032	-0.0022	-0.0022
Morganella	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Muribaculaceae	-0.0005	-0.0054	0.0000	0.0000	0.0000	0.0000
Negativibacillus	0.0005	-0.0217	0.0080	0.0069	0.0047	0.0047
Neisseria	0.0000	0.0000	0.0000	-0.0021	-0.0014	-0.0014
NK4A214_group	0.0042	-0.0109	0.0176	0.0065	0.0044	0.0044
Odoribacter	-0.0014	-0.0217	0.0016	-0.0030	-0.0021	-0.0021
Olsenella	-0.0005	0.0000	-0.0016	-0.0023	-0.0016	-0.0016
Oribacterium	0.0010	-0.0054	0.0048	0.0005	0.0003	0.0003
Oscillibacter	-0.0005	-0.0054	0.0000	-0.0011	-0.0007	-0.0007
Oscillospira	0.0017	-0.0870	0.0320	0.0172	0.0117	0.0117
Parabacteroides	0.0000	0.0000	0.0000	-0.0017	-0.0011	-0.0011
Paraprevotella	0.0008	-0.0217	0.0096	0.0020	0.0014	0.0014
Parasutterella	0.0000	0.0000	0.0000	-0.0015	-0.0010	-0.0010
Parvimonas	-0.0005	0.0000	-0.0016	-0.0043	-0.0029	-0.0029
Peptococcus	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Peptoniphilus	0.0014	-0.0054	0.0064	0.0022	0.0015	0.0015
Peptostreptococcus	-0.0005	0.0000	-0.0016	-0.0007	-0.0005	-0.0005
Phascolarctobacterium	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Porphyromonas	0.0005	0.0000	0.0016	-0.0022	-0.0015	-0.0015
Prevotella	-0.0005	-0.0109	0.0016	-0.0015	-0.0010	-0.0010
Prevotellaceae_NK3B31_group	0.0019	0.0000	0.0064	-0.0022	-0.0015	-0.0015
Prevotellaceae_UCG.001	-0.0019	-0.0272	0.0016	-0.0059	-0.0040	-0.0040
Pseudomonas	-0.0032	0.0054	-0.0128	-0.0089	-0.0060	-0.0060
Pyramidobacter	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
RF39	-0.0018	-0.0054	-0.0048	0.0005	0.0003	0.0003
Rikenellaceae_RC9_gut_group	0.0023	-0.0109	0.0112	0.0010	0.0007	0.0007
Romboutsia	-0.0001	-0.0544	0.0160	0.0091	0.0062	0.0062
Roseburia	0.0010	0.0054	0.0016	-0.0009	-0.0006	-0.0006
Rothia	0.0041	-0.0163	0.0192	0.0154	0.0105	0.0105
Ruminiclostridium	0.0000	0.0000	0.0000	0.0002	0.0001	0.0001
Ruminococcus	0.0505	0.0979	0.1471	0.1247	0.0848	0.0848
S5.A14a	0.0005	-0.0109	0.0048	0.0056	0.0038	0.0038
Saccharimonadaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Saccharimonadales	0.0000	0.0000	0.0000	-0.0006	-0.0004	-0.0004
Sanguibacteroides	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Sellimonas	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Senegalimassilia	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Serratia	-0.0005	-0.0489	0.0128	-0.0110	-0.0075	-0.0075
Slackia	0.0005	0.0054	0.0000	0.0007	0.0005	0.0005
Solobacterium	0.0009	0.0000	0.0032	-0.0002	-0.0002	-0.0002
Staphylococcus	0.0005	0.0000	0.0016	-0.0041	-0.0028	-0.0028
Streptococcus	-0.0005	0.0000	-0.0016	0.0023	0.0016	0.0016
Subdoligranulum	0.0009	-0.0054	0.0048	-0.0038	-0.0026	-0.0026
Sutterella	0.0063	-0.1196	0.0576	0.0275	0.0187	0.0187
Terrisporobacter	0.0005	0.0000	0.0016	0.0008	0.0005	0.0005
TM7a	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
TM7x	-0.0018	-0.0272	0.0016	-0.0012	-0.0008	-0.0008
Turicibacter	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Tyzzerella	0.0000	-0.0272	0.0080	0.0014	0.0009	0.0009
UBA1819	-0.0009	-0.0054	-0.0016	-0.0007	-0.0005	-0.0005
UCG.002	0.0243	-0.0435	0.0975	0.0693	0.0471	0.0471
UCG.003	-0.0009	-0.0109	0.0000	-0.0007	-0.0005	-0.0005
UCG.005	0.0068	-0.0272	0.0320	0.0085	0.0057	0.0057
UCG.010	-0.0005	0.0000	-0.0016	0.0000	0.0000	0.0000
UCG.011	0.0005	0.0054	0.0000	0.0002	0.0002	0.0002
uncultured	0.0005	-0.0163	0.0064	0.0012	0.0008	0.0008
vadinBE97	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Veillonella	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Victivallaceae	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Victivallis	0.0000	-0.0163	0.0048	0.0053	0.0036	0.0036

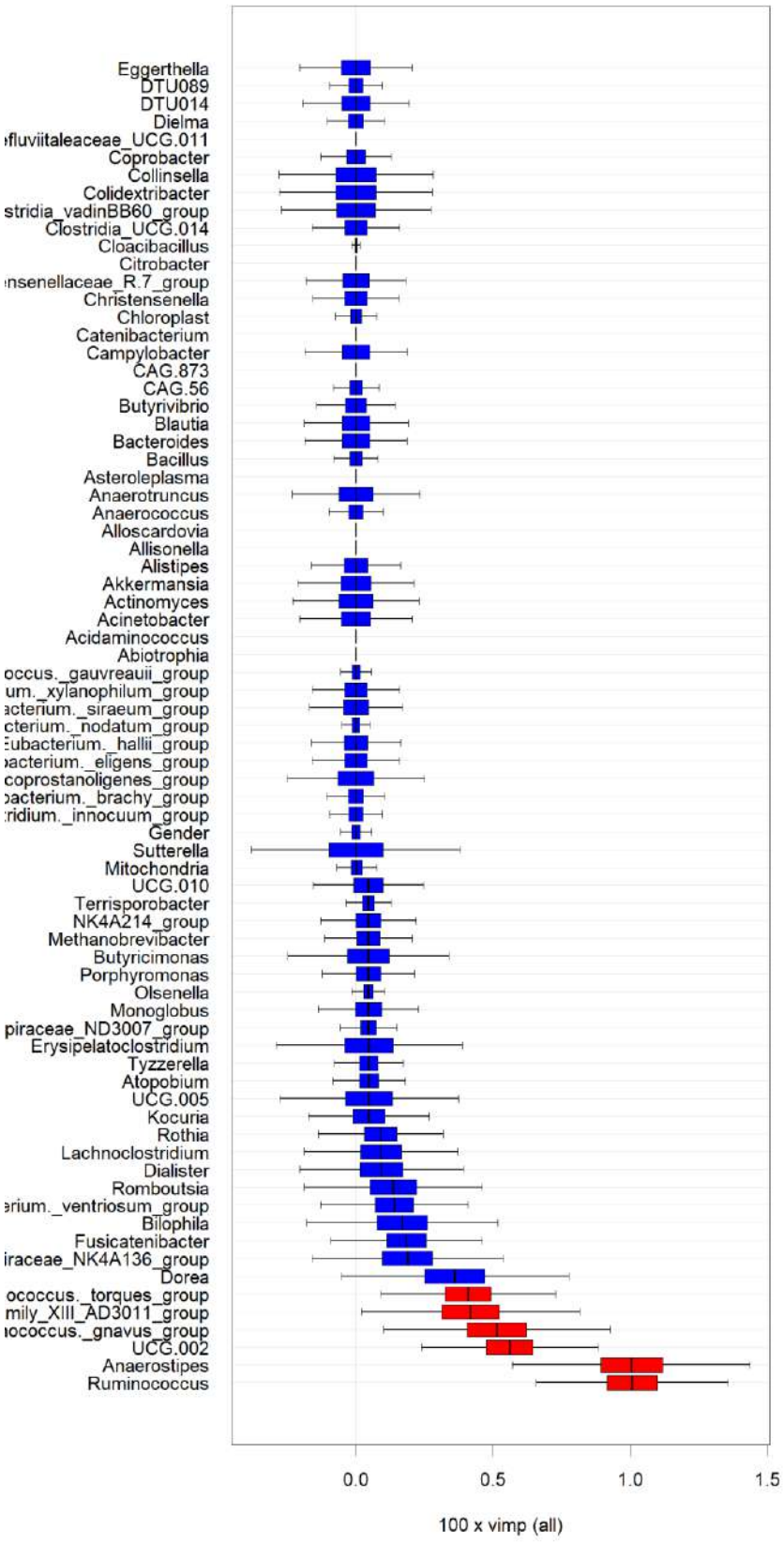

Eggerthella
DTU089
DTU014
Dielma
fluviitaleaceae_UCG.011
Coprobacter
Collinsella
Colidextribacter
stridia_vadinBB60_group
Clostridia_UCG.014
Cloacibacillus
Citrobacter
nsenellaceae_R.7_group
Christensenella
Chloroplast
Catenibacterium
Campylobacter
CAG.873
CAG.56
Butyrivibrio
Blautia
Bacteroides
Bacillus
Asteroleplasma
Anaerotruncus
Anaerococcus
Alloscardovia
Allisonella
Alistipes
Akkermansia
Actinomyces
Acinetobacter
Acidaminococcus
Abiotrophia
occus._gauvreauii_group
um._xylanophilum_group
acterium._siraeum_group
cterium._nodatum_group
Eubacterium._hallii_group
bacterium._eligens_group
coprostanoligenes_group
bacterium._brachy_group
ridium._innocuum_group
Gender
Sutterella
Mitochondria
UCG.010
Terrisporobacter
NK4A214_group
Methanobrevibacter
Butyricimonas
Porphyromonas
Olsenella
Monoglobus
piraceae_ND3007_group
Erysipelatoclostridium
Tyzzerella
Atopobium
UCG.005
Kocuria
Rothia
Lachnoclostridium
Dialister
Romboutsia
erium._ventriosum_group
Bilophila
Fusicatenibacter
iraceae_NK4A136_group
Dorea
ococcus._torques_group
mily_XIII_AD3011_group
ococcus._gnavus_group
UCG.002
Anaerostipes
Ruminococcus
0.0
0.5
1.0
1.5
100 x vimp (all)

Figure ML2-E-5: The Variable Importance Plot (VIMP) for Genus in Stool

Best model is selected for LIME results

Table ML2-E-3: The best setting for model and LIME Results

Setting	mtry	ROC	Sens	Spec	ROCSD	SensSD	SpecSD
ntrees: 500 nodesize: 1	4	0.716667	0	1	0.182574	0	0
ntrees: 1000 nodesize: 1	1	0.766667	0	1	0.136931	0	0
ntrees: 500 nodesize: 5	2	0.716667	0	1	0.389087	0	0
ntrees: 500 nodesize: 5	10	0.716667	0.2	0.95	0.182574	0.447214	0.111803
ntrees: 1000 nodesize: 5	2	0.833333	0	1	0.235702	0	0

Table ML2-E-4: The LIME Result for Training Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
1	Adenoma	0.623	0.235	0.175	0.616	Harryflintia	0.003	0.407	0.001938 < Harryflintia
1	Adenoma	0.623	0.235	0.175	0.616	X.Ruminococcus._gnavus_group	0.186	0.011	0.000763 < X.Ruminococcus._gnavus_group
1	Adenoma	0.623	0.235	0.175	0.616	Eisenbergiella	0.012	0.009	0.000267 < Eisenbergiella
1	Adenoma	0.623	0.235	0.175	0.616	Desulfovibrio	0.000	0.007	Desulfovibrio <= 0.000029
1	Adenoma	0.623	0.235	0.175	0.616	Family_XIII_AD3011_group	0.000	0.007	Family_XIII_AD3011_group <= 0.000433
2	Adenoma	0.626	0.237	0.183	0.618	Coprobacillus	0.008	0.415	0.00598 < Coprobacillus
2	Adenoma	0.626	0.237	0.183	0.618	X.Ruminococcus._gnavus_group	0.016	0.011	0.000763 < X.Ruminococcus._gnavus_group
2	Adenoma	0.626	0.237	0.183	0.618	Eisenbergiella	0.019	0.009	0.000267 < Eisenbergiella
2	Adenoma	0.626	0.237	0.183	0.618	Negativibacillus	0.000	-0.007	Negativibacillus <= 0.000172
2	Adenoma	0.626	0.237	0.183	0.618	Ruminococcus	0.001	0.007	Ruminococcus <= 0.00486
3	Adenoma	0.575	0.215	0.196	0.568	Mitochondria	0.000	0.379	1.78e-04 < Mitochondria
3	Adenoma	0.575	0.215	0.196	0.568	X.Ruminococcus._gnavus_group	0.001	-0.012	X.Ruminococcus._gnavus_group <= 0.000763
3	Adenoma	0.575	0.215	0.196	0.568	Eisenbergiella	0.000	-0.008	Eisenbergiella <= 0.000267
3	Adenoma	0.575	0.215	0.196	0.568	Family_XIII_AD3011_group	0.000	0.007	Family_XIII_AD3011_group <= 0.000433
3	Adenoma	0.575	0.215	0.196	0.568	Negativibacillus	0.001	0.006	0.000172 < Negativibacillus
6	Adenoma	0.573	0.159	0.186	0.204	X.Ruminococcus._gnavus_group	0.000	-0.012	X.Ruminococcus._gnavus_group <= 0.000763
6	Adenoma	0.573	0.159	0.186	0.204	Eisenbergiella	0.000	0.009	0.000267 < Eisenbergiella
6	Adenoma	0.573	0.159	0.186	0.204	Family_XIII_AD3011_group	0.000	0.007	Family_XIII_AD3011_group <= 0.000433
6	Adenoma	0.573	0.159	0.186	0.204	Butyricimonas	0.008	0.007	0.006295 < Butyricimonas
6	Adenoma	0.573	0.159	0.186	0.204	Negativibacillus	0.000	0.006	0.000172 < Negativibacillus
9	Adenoma	0.664	0.267	0.185	0.656	Bacillus	0.002	0.453	0.001746 < Bacillus
9	Adenoma	0.664	0.267	0.185	0.656	X.Ruminococcus._gnavus_group	0.039	0.012	0.000763 < X.Ruminococcus._gnavus_group
9	Adenoma	0.664	0.267	0.185	0.656	Eisenbergiella	0.000	-0.010	Eisenbergiella <= 0.000267
9	Adenoma	0.664	0.267	0.185	0.656	Ruminococcus	0.000	0.007	Ruminococcus <= 0.00486
9	Adenoma	0.664	0.267	0.185	0.656	Family_XIII_AD3011_group	0.000	0.007	Family_XIII_AD3011_group <= 0.000433

11	Adenocarcinoma	0.813	0.159	0.799	0.815	X.Ruminococcus._gnavus_group	0.000	0.012	X.Ruminococcus._gnavus_group <= 0.000763
11	Adenocarcinoma	0.813	0.159	0.799	0.815	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
11	Adenocarcinoma	0.813	0.159	0.799	0.815	Senegalimassilia	0.000	-0.005	Senegalimassilia <= 3.56e-05
11	Adenocarcinoma	0.813	0.159	0.799	0.815	Negativibacillus	0.000	0.006	Negativibacillus <= 0.000172
11	Adenocarcinoma	0.813	0.159	0.799	0.815	Dorea	0.000	-0.006	Dorea <= 0.000459
17	Adenocarcinoma	0.841	0.157	0.808	0.803	X.Ruminococcus._gnavus_group	0.000	0.011	X.Ruminococcus._gnavus_group <= 0.000763
17	Adenocarcinoma	0.841	0.157	0.808	0.803	Eisenbergiella	0.000	-0.010	0.000267 < Eisenbergiella
17	Adenocarcinoma	0.841	0.157	0.808	0.803	Butyricimonas	0.007	-0.006	0.006295 < Butyricimonas
17	Adenocarcinoma	0.841	0.157	0.808	0.803	Ruminococcus	0.005	-0.005	Ruminococcus <= 0.00486
17	Adenocarcinoma	0.841	0.157	0.808	0.803	Candidatus_Soleaferrea	0.000	0.005	Candidatus_Soleaferrea <= 3e-05
18	Adenocarcinoma	0.894	0.177	0.806	0.806	X.Ruminococcus._gnavus_group	0.000	0.012	X.Ruminococcus._gnavus_group <= 0.000763
18	Adenocarcinoma	0.894	0.177	0.806	0.806	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
18	Adenocarcinoma	0.894	0.177	0.806	0.806	Family_XIII_AD3011_group	0.000	-0.008	Family_XIII_AD3011_group <= 0.000433
18	Adenocarcinoma	0.894	0.177	0.806	0.806	Desulfovibrio	0.000	-0.007	Desulfovibrio <= 0.000029
18	Adenocarcinoma	0.894	0.177	0.806	0.806	Butyricimonas	0.008	-0.007	0.006295 < Butyricimonas
19	Adenocarcinoma	0.896	0.148	0.788	0.824	X.Ruminococcus._gnavus_group	0.000	0.013	X.Ruminococcus._gnavus_group <= 0.000763
19	Adenocarcinoma	0.896	0.148	0.788	0.824	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
19	Adenocarcinoma	0.896	0.148	0.788	0.824	Candidatus_Soleaferrea	0.000	0.005	Candidatus_Soleaferrea <= 3e-05
19	Adenocarcinoma	0.896	0.148	0.788	0.824	Negativibacillus	0.000	0.005	Negativibacillus <= 0.000172
19	Adenocarcinoma	0.896	0.148	0.788	0.824	Izemoplasmatales	0.000	0.005	Izemoplasmatales <= 0.000465
20	Adenocarcinoma	0.870	0.172	0.805	0.811	X.Ruminococcus._gnavus_group	0.000	0.013	X.Ruminococcus._gnavus_group <= 0.000763
20	Adenocarcinoma	0.870	0.172	0.805	0.811	Eisenbergiella	0.000	0.010	Eisenbergiella <= 0.000267
20	Adenocarcinoma	0.870	0.172	0.805	0.811	Butyricimonas	0.006	-0.007	0.006295 < Butyricimonas
20	Adenocarcinoma	0.870	0.172	0.805	0.811	Saccharimonadaceae	0.000	-0.005	Saccharimonadaceae <= 8.16e-06
20	Adenocarcinoma	0.870	0.172	0.805	0.811	Desulfovibrio	0.000	-0.005	Desulfovibrio <= 0.000029
21	Adenocarcinoma	0.898	0.146	0.800	0.813	X.Ruminococcus._gnavus_group	0.000	0.011	X.Ruminococcus._gnavus_group <= 0.000763
21	Adenocarcinoma	0.898	0.146	0.800	0.813	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
21	Adenocarcinoma	0.898	0.146	0.800	0.813	Negativibacillus	0.003	-0.006	0.000172 < Negativibacillus
21	Adenocarcinoma	0.898	0.146	0.800	0.813	Rothia	0.000	-0.006	5.27e-05 < Rothia
21	Adenocarcinoma	0.898	0.146	0.800	0.813	Candidatus_Soleaferrea	0.000	0.006	Candidatus_Soleaferrea <= 3e-05
22	Adenocarcinoma	0.857	0.165	0.800	0.814	X.Ruminococcus._gnavus_group	0.000	0.012	X.Ruminococcus._gnavus_group <= 0.000763
22	Adenocarcinoma	0.857	0.165	0.800	0.814	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
22	Adenocarcinoma	0.857	0.165	0.800	0.814	Butyricimonas	0.011	-0.007	0.006295 < Butyricimonas
22	Adenocarcinoma	0.857	0.165	0.800	0.814	Saccharimonadaceae	0.000	-0.006	Saccharimonadaceae <= 8.16e-06
22	Adenocarcinoma	0.857	0.165	0.800	0.814	Izemoplasmatales	0.000	0.005	Izemoplasmatales <= 0.000465
23	Adenocarcinoma	0.904	0.142	0.805	0.806	X.Ruminococcus._gnavus_group	0.001	-0.011	0.000763 < X.Ruminococcus._gnavus_group
23	Adenocarcinoma	0.904	0.142	0.805	0.806	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
23	Adenocarcinoma	0.904	0.142	0.805	0.806	Family_XIII_AD3011_group	0.000	-0.007	Family_XIII_AD3011_group <= 0.000433
23	Adenocarcinoma	0.904	0.142	0.805	0.806	Izemoplasmatales	0.000	0.005	Izemoplasmatales <= 0.000465
23	Adenocarcinoma	0.904	0.142	0.805	0.806	Negativibacillus	0.000	0.005	Negativibacillus <= 0.000172
25	Adenocarcinoma	0.912	0.123	0.806	0.811	X.Ruminococcus._gnavus_group	0.001	-0.010	0.000763 < X.Ruminococcus._gnavus_group
25	Adenocarcinoma	0.912	0.123	0.806	0.811	Eisenbergiella	0.000	0.008	Eisenbergiella <= 0.000267

25	Adenocarcinoma	0.912	0.123	0.806	0.811	Negativibacillus	0.000	0.007	Negativibacillus <= 0.000172
25	Adenocarcinoma	0.912	0.123	0.806	0.811	Holdemania	0.000	-0.005	Holdemania <= 0.000107
25	Adenocarcinoma	0.912	0.123	0.806	0.811	Izemoplasmatales	0.000	0.005	Izemoplasmatales <= 0.000465
26	Adenocarcinoma	0.946	0.134	0.803	0.808	X.Ruminococcus._gnavus_group	0.001	-0.011	0.000763 < X.Ruminococcus._gnavus_group
26	Adenocarcinoma	0.946	0.134	0.803	0.808	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
26	Adenocarcinoma	0.946	0.134	0.803	0.808	Negativibacillus	0.000	0.006	Negativibacillus <= 0.000172
26	Adenocarcinoma	0.946	0.134	0.803	0.808	Rothia	0.000	0.006	Rothia <= 5.27e-05
26	Adenocarcinoma	0.946	0.134	0.803	0.808	Candidatus_Soleaferrea	0.000	-0.005	3e-05 < Candidatus_Soleaferrea
29	Adenocarcinoma	0.903	0.156	0.809	0.806	X.Ruminococcus._gnavus_group	0.000	0.012	X.Ruminococcus._gnavus_group <= 0.000763
29	Adenocarcinoma	0.903	0.156	0.809	0.806	Eisenbergiella	0.000	-0.010	0.000267 < Eisenbergiella
29	Adenocarcinoma	0.903	0.156	0.809	0.806	Negativibacillus	0.000	0.006	Negativibacillus <= 0.000172
29	Adenocarcinoma	0.903	0.156	0.809	0.806	Desulfovibrio	0.000	-0.005	Desulfovibrio <= 0.000029
29	Adenocarcinoma	0.903	0.156	0.809	0.806	Sutterella	0.001	-0.005	2.62e-04 < Sutterella
30	Adenocarcinoma	0.891	0.135	0.795	0.821	X.Ruminococcus._gnavus_group	0.000	0.011	X.Ruminococcus._gnavus_group <= 0.000763
30	Adenocarcinoma	0.891	0.135	0.795	0.821	Eisenbergiella	0.000	0.008	Eisenbergiella <= 0.000267
30	Adenocarcinoma	0.891	0.135	0.795	0.821	Rothia	0.000	0.006	Rothia <= 5.27e-05
30	Adenocarcinoma	0.891	0.135	0.795	0.821	Gemella	0.000	-0.005	Gemella <= 0.000150
30	Adenocarcinoma	0.891	0.135	0.795	0.821	Negativibacillus	0.000	0.005	Negativibacillus <= 0.000172
31	Adenocarcinoma	0.920	0.125	0.794	0.819	X.Ruminococcus._gnavus_group	0.000	0.011	X.Ruminococcus._gnavus_group <= 0.000763
31	Adenocarcinoma	0.920	0.125	0.794	0.819	Eisenbergiella	0.000	0.008	Eisenbergiella <= 0.000267
31	Adenocarcinoma	0.920	0.125	0.794	0.819	Negativibacillus	0.000	0.005	Negativibacillus <= 0.000172
31	Adenocarcinoma	0.920	0.125	0.794	0.819	Izemoplasmatales	0.000	0.005	Izemoplasmatales <= 0.000465
31	Adenocarcinoma	0.920	0.125	0.794	0.819	Candidatus_Soleaferrea	0.000	-0.005	3e-05 < Candidatus_Soleaferrea
35	Adenocarcinoma	0.887	0.155	0.786	0.825	X.Ruminococcus._gnavus_group	0.000	0.012	X.Ruminococcus._gnavus_group <= 0.000763
35	Adenocarcinoma	0.887	0.155	0.786	0.825	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
35	Adenocarcinoma	0.887	0.155	0.786	0.825	Negativibacillus	0.000	0.006	Negativibacillus <= 0.000172
35	Adenocarcinoma	0.887	0.155	0.786	0.825	Candidatus_Soleaferrea	0.000	0.006	Candidatus_Soleaferrea <= 3e-05
35	Adenocarcinoma	0.887	0.155	0.786	0.825	Rothia	0.000	0.006	Rothia <= 5.27e-05
41	Adenocarcinoma	0.909	0.139	0.797	0.821	X.Ruminococcus._gnavus_group	0.000	0.012	X.Ruminococcus._gnavus_group <= 0.000763
41	Adenocarcinoma	0.909	0.139	0.797	0.821	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
41	Adenocarcinoma	0.909	0.139	0.797	0.821	Candidatus_Soleaferrea	0.000	-0.006	3e-05 < Candidatus_Soleaferrea
41	Adenocarcinoma	0.909	0.139	0.797	0.821	Senegalimassilia	0.001	0.005	7.64e-04 < Senegalimassilia
41	Adenocarcinoma	0.909	0.139	0.797	0.821	Oscillospira	0.000	0.005	Oscillospira <= 0.000732
42	Adenocarcinoma	0.900	0.153	0.805	0.813	X.Ruminococcus._gnavus_group	0.000	0.012	X.Ruminococcus._gnavus_group <= 0.000763
42	Adenocarcinoma	0.900	0.153	0.805	0.813	Eisenbergiella	0.001	-0.009	0.000267 < Eisenbergiella
42	Adenocarcinoma	0.900	0.153	0.805	0.813	Saccharimonadaceae	0.000	-0.005	Saccharimonadaceae <= 8.16e-06
42	Adenocarcinoma	0.900	0.153	0.805	0.813	Candidatus_Soleaferrea	0.000	0.006	Candidatus_Soleaferrea <= 3e-05
42	Adenocarcinoma	0.900	0.153	0.805	0.813	Butyricimonas	0.000	0.005	Butyricimonas <= 0.000823
43	Adenocarcinoma	0.912	0.158	0.790	0.828	X.Ruminococcus._gnavus_group	0.000	0.012	X.Ruminococcus._gnavus_group <= 0.000763
43	Adenocarcinoma	0.912	0.158	0.790	0.828	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
43	Adenocarcinoma	0.912	0.158	0.790	0.828	Senegalimassilia	0.001	0.006	7.64e-04 < Senegalimassilia
43	Adenocarcinoma	0.912	0.158	0.790	0.828	Negativibacillus	0.000	0.006	Negativibacillus <= 0.000172

43	Adenocarcinoma	0.912	0.158	0.790	0.828	Rothia	0.000	0.005	Rothia <= 5.27e-05

Table ML-E-5: The LIME Result for Test Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
4	Adenocarcinoma	0.799	0.171	0.818	0.797	X.Ruminococcus._gnavus_group	0.001	-0.013	0.000763 < X.Ruminococcus._gnavus_group
4	Adenocarcinoma	0.799	0.171	0.818	0.797	Eisenbergiella	0.000	0.010	Eisenbergiella <= 0.000267
4	Adenocarcinoma	0.799	0.171	0.818	0.797	Desulfovibrio	0.000	-0.006	Desulfovibrio <= 0.000029
4	Adenocarcinoma	0.799	0.171	0.818	0.797	Senegalimassilia	0.000	-0.005	Senegalimassilia <= 3.56e-05
4	Adenocarcinoma	0.799	0.171	0.818	0.797	Candidatus_Soleaferrea	0.002	-0.006	3e-05 < Candidatus_Soleaferrea
5	Adenocarcinoma	0.743	0.156	0.826	0.787	X.Ruminococcus._gnavus_group	0.012	-0.011	0.000763 < X.Ruminococcus._gnavus_group
5	Adenocarcinoma	0.743	0.156	0.826	0.787	Eisenbergiella	0.000	-0.008	0.000267 < Eisenbergiella
5	Adenocarcinoma	0.743	0.156	0.826	0.787	Butyricimonas	0.026	-0.008	0.006295 < Butyricimonas
5	Adenocarcinoma	0.743	0.156	0.826	0.787	Family_XIII_AD3011_group	0.000	-0.006	Family_XIII_AD3011_group <= 0.000433
5	Adenocarcinoma	0.743	0.156	0.826	0.787	Senegalimassilia	0.000	-0.005	Senegalimassilia <= 3.56e-05
7	Adenocarcinoma	0.779	0.155	0.793	0.818	X.Ruminococcus._gnavus_group	0.000	0.013	X.Ruminococcus._gnavus_group <= 0.000763
7	Adenocarcinoma	0.779	0.155	0.793	0.818	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
7	Adenocarcinoma	0.779	0.155	0.793	0.818	Ruminococcus	0.002	-0.006	Ruminococcus <= 0.00486
7	Adenocarcinoma	0.779	0.155	0.793	0.818	Negativibacillus	0.000	0.006	Negativibacillus <= 0.000172
7	Adenocarcinoma	0.779	0.155	0.793	0.818	Izemoplasmatales	0.000	0.005	Izemoplasmatales <= 0.000465
8	Adenocarcinoma	0.790	0.145	0.795	0.821	X.Ruminococcus._gnavus_group	0.000	0.012	X.Ruminococcus._gnavus_group <= 0.000763
8	Adenocarcinoma	0.790	0.145	0.795	0.821	Eisenbergiella	0.000	0.008	Eisenbergiella <= 0.000267
8	Adenocarcinoma	0.790	0.145	0.795	0.821	Desulfovibrio	0.000	-0.005	Desulfovibrio <= 0.000029
8	Adenocarcinoma	0.790	0.145	0.795	0.821	Rothia	0.000	0.005	Rothia <= 5.27e-05
8	Adenocarcinoma	0.790	0.145	0.795	0.821	Oscillospira	0.000	0.005	Oscillospira <= 0.000732
10	Adenocarcinoma	0.806	0.149	0.800	0.814	X.Ruminococcus._gnavus_group	0.000	0.011	X.Ruminococcus._gnavus_group <= 0.000763
10	Adenocarcinoma	0.806	0.149	0.800	0.814	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
10	Adenocarcinoma	0.806	0.149	0.800	0.814	Desulfovibrio	0.000	-0.006	Desulfovibrio <= 0.000029
10	Adenocarcinoma	0.806	0.149	0.800	0.814	Negativibacillus	0.000	-0.006	0.000172 < Negativibacillus
10	Adenocarcinoma	0.806	0.149	0.800	0.814	Izemoplasmatales	0.000	0.006	Izemoplasmatales <= 0.000465
12	Adenocarcinoma	0.805	0.146	0.803	0.817	X.Ruminococcus._gnavus_group	0.001	0.012	X.Ruminococcus._gnavus_group <= 0.000763
12	Adenocarcinoma	0.805	0.146	0.803	0.817	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
12	Adenocarcinoma	0.805	0.146	0.803	0.817	Saccharimonadaceae	0.000	-0.006	Saccharimonadaceae <= 8.16e-06
12	Adenocarcinoma	0.805	0.146	0.803	0.817	Negativibacillus	0.001	-0.006	0.000172 < Negativibacillus
12	Adenocarcinoma	0.805	0.146	0.803	0.817	Gemella	0.000	0.005	0.000378 < Gemella
13	Adenocarcinoma	0.799	0.138	0.800	0.814	X.Ruminococcus._gnavus_group	0.000	0.011	X.Ruminococcus._gnavus_group <= 0.000763
13	Adenocarcinoma	0.799	0.138	0.800	0.814	Eisenbergiella	0.000	0.008	Eisenbergiella <= 0.000267
13	Adenocarcinoma	0.799	0.138	0.800	0.814	Butyricimonas	0.009	-0.006	0.006295 < Butyricimonas
13	Adenocarcinoma	0.799	0.138	0.800	0.814	Rothia	0.000	0.006	Rothia <= 5.27e-05
13	Adenocarcinoma	0.799	0.138	0.800	0.814	Senegalimassilia	0.000	-0.005	Senegalimassilia <= 3.56e-05
14	Adenocarcinoma	0.855	0.142	0.809	0.811	X.Ruminococcus._gnavus_group	0.003	-0.012	0.000763 < X.Ruminococcus._gnavus_group
14	Adenocarcinoma	0.855	0.142	0.809	0.811	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267

14	Adenocarcinoma	0.855	0.142	0.809	0.811	Candidatus_Soleaferrea	0.000	0.005	Candidatus_Soleaferrea <= 3e-05
14	Adenocarcinoma	0.855	0.142	0.809	0.811	Senegalimassilia	0.000	-0.005	Senegalimassilia <= 3.56e-05
14	Adenocarcinoma	0.855	0.142	0.809	0.811	Gemella	0.004	0.005	0.000378 < Gemella
15	Adenocarcinoma	0.868	0.146	0.797	0.820	X.Ruminococcus._gnavus_group	0.000	0.011	X.Ruminococcus._gnavus_group <= 0.000763
15	Adenocarcinoma	0.868	0.146	0.797	0.820	Eisenbergiella	0.000	0.008	Eisenbergiella <= 0.000267
15	Adenocarcinoma	0.868	0.146	0.797	0.820	Butyricimonas	0.009	-0.007	0.006295 < Butyricimonas
15	Adenocarcinoma	0.868	0.146	0.797	0.820	Negativibacillus	0.000	0.006	Negativibacillus <= 0.000172
15	Adenocarcinoma	0.868	0.146	0.797	0.820	Senegalimassilia	0.002	0.006	7.64e-04 < Senegalimassilia
16	Adenocarcinoma	0.871	0.169	0.793	0.818	X.Ruminococcus._gnavus_group	0.000	0.012	X.Ruminococcus._gnavus_group <= 0.000763
16	Adenocarcinoma	0.871	0.169	0.793	0.818	Eisenbergiella	0.000	0.010	Eisenbergiella <= 0.000267
16	Adenocarcinoma	0.871	0.169	0.793	0.818	Butyricimonas	0.012	-0.008	0.006295 < Butyricimonas
16	Adenocarcinoma	0.871	0.169	0.793	0.818	Negativibacillus	0.000	0.005	Negativibacillus <= 0.000172
16	Adenocarcinoma	0.871	0.169	0.793	0.818	Oscillospira	0.000	0.005	Oscillospira <= 0.000732
24	Adenocarcinoma	0.872	0.137	0.806	0.806	X.Ruminococcus._gnavus_group	0.000	0.010	X.Ruminococcus._gnavus_group <= 0.000763
24	Adenocarcinoma	0.872	0.137	0.806	0.806	Eisenbergiella	0.000	0.008	Eisenbergiella <= 0.000267
24	Adenocarcinoma	0.872	0.137	0.806	0.806	Butyricimonas	0.008	-0.006	0.006295 < Butyricimonas
24	Adenocarcinoma	0.872	0.137	0.806	0.806	Candidatus_Soleaferrea	0.000	-0.006	3e-05 < Candidatus_Soleaferrea
24	Adenocarcinoma	0.872	0.137	0.806	0.806	Negativibacillus	0.000	-0.006	0.000172 < Negativibacillus
27	Adenocarcinoma	0.817	0.161	0.818	0.796	X.Ruminococcus._gnavus_group	0.001	-0.012	0.000763 < X.Ruminococcus._gnavus_group
27	Adenocarcinoma	0.817	0.161	0.818	0.796	Eisenbergiella	0.000	0.008	Eisenbergiella <= 0.000267
27	Adenocarcinoma	0.817	0.161	0.818	0.796	Ruminococcus	0.001	-0.007	Ruminococcus <= 0.00486
27	Adenocarcinoma	0.817	0.161	0.818	0.796	Negativibacillus	0.002	-0.006	0.000172 < Negativibacillus
27	Adenocarcinoma	0.817	0.161	0.818	0.796	Desulfovibrio	0.000	-0.005	Desulfovibrio <= 0.000029
28	Adenocarcinoma	0.800	0.157	0.795	0.822	X.Ruminococcus._gnavus_group	0.000	0.011	X.Ruminococcus._gnavus_group <= 0.000763
28	Adenocarcinoma	0.800	0.157	0.795	0.822	Eisenbergiella	0.000	0.010	Eisenbergiella <= 0.000267
28	Adenocarcinoma	0.800	0.157	0.795	0.822	Rothia	0.000	0.007	Rothia <= 5.27e-05
28	Adenocarcinoma	0.800	0.157	0.795	0.822	Negativibacillus	0.000	0.006	Negativibacillus <= 0.000172
28	Adenocarcinoma	0.800	0.157	0.795	0.822	UCG.011	0.000	-0.006	UCG.011 <= 3.48e-05
32	Adenocarcinoma	0.663	0.164	0.807	0.810	X.Ruminococcus._gnavus_group	0.000	0.011	X.Ruminococcus._gnavus_group <= 0.000763
32	Adenocarcinoma	0.663	0.164	0.807	0.810	Eisenbergiella	0.000	0.010	Eisenbergiella <= 0.000267
32	Adenocarcinoma	0.663	0.164	0.807	0.810	Ruminococcus	0.000	-0.007	Ruminococcus <= 0.00486
32	Adenocarcinoma	0.663	0.164	0.807	0.810	Desulfovibrio	0.000	-0.006	Desulfovibrio <= 0.000029
32	Adenocarcinoma	0.663	0.164	0.807	0.810	Saccharimonadaceae	0.000	-0.005	Saccharimonadaceae <= 8.16e-06
33	Adenocarcinoma	0.579	0.176	0.812	0.583	X.Ruminococcus._gnavus_group	0.000	0.012	X.Ruminococcus._gnavus_group <= 0.000763
33	Adenocarcinoma	0.579	0.176	0.812	0.583	Eisenbergiella	0.000	-0.009	0.000267 < Eisenbergiella
33	Adenocarcinoma	0.579	0.176	0.812	0.583	Anaerococcus	0.076	-0.218	0.002128 < Anaerococcus
33	Adenocarcinoma	0.579	0.176	0.812	0.583	Family_XIII_AD3011_group	0.000	-0.007	Family_XIII_AD3011_group <= 0.000433
33	Adenocarcinoma	0.579	0.176	0.812	0.583	Dorea	0.000	-0.006	Dorea <= 0.000459
34	Adenocarcinoma	0.760	0.154	0.805	0.811	X.Ruminococcus._gnavus_group	0.000	0.013	X.Ruminococcus._gnavus_group <= 0.000763
34	Adenocarcinoma	0.760	0.154	0.805	0.811	Eisenbergiella	0.000	0.008	Eisenbergiella <= 0.000267
34	Adenocarcinoma	0.760	0.154	0.805	0.811	Family_XIII_AD3011_group	0.000	-0.006	Family_XIII_AD3011_group <= 0.000433
34	Adenocarcinoma	0.760	0.154	0.805	0.811	Gemella	0.000	-0.005	Gemella <= 0.000150

34	Adenocarcinoma	0.760	0.154	0.805	0.811	Senegalimassilia	0.000	-0.005	Senegalimassilia <= 3.56e-05
36	Adenocarcinoma	0.781	0.162	0.793	0.821	X.Ruminococcus._gnavus_group	0.000	0.013	X.Ruminococcus._gnavus_group <= 0.000763
36	Adenocarcinoma	0.781	0.162	0.793	0.821	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
36	Adenocarcinoma	0.781	0.162	0.793	0.821	Senegalimassilia	0.000	-0.005	Senegalimassilia <= 3.56e-05
36	Adenocarcinoma	0.781	0.162	0.793	0.821	Candidatus_Soleaferrea	0.000	0.005	Candidatus_Soleaferrea <= 3e-05
36	Adenocarcinoma	0.781	0.162	0.793	0.821	Negativibacillus	0.000	0.005	Negativibacillus <= 0.000172
37	Adenocarcinoma	0.718	0.155	0.818	0.797	X.Ruminococcus._gnavus_group	0.009	-0.012	0.000763 < X.Ruminococcus._gnavus_group
37	Adenocarcinoma	0.718	0.155	0.818	0.797	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
37	Adenocarcinoma	0.718	0.155	0.818	0.797	Desulfovibrio	0.000	-0.006	Desulfovibrio <= 0.000029
37	Adenocarcinoma	0.718	0.155	0.818	0.797	Negativibacillus	0.000	-0.006	0.000172 < Negativibacillus
37	Adenocarcinoma	0.718	0.155	0.818	0.797	Ruminococcus	0.001	-0.006	Ruminococcus <= 0.00486
38	Adenocarcinoma	0.755	0.163	0.799	0.812	X.Ruminococcus._gnavus_group	0.000	0.012	X.Ruminococcus._gnavus_group <= 0.000763
38	Adenocarcinoma	0.755	0.163	0.799	0.812	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
38	Adenocarcinoma	0.755	0.163	0.799	0.812	Ruminococcus	0.000	-0.006	Ruminococcus <= 0.00486
38	Adenocarcinoma	0.755	0.163	0.799	0.812	Sutterella	0.001	-0.006	2.62e-04 < Sutterella
38	Adenocarcinoma	0.755	0.163	0.799	0.812	Izemoplasmatales	0.000	0.005	Izemoplasmatales <= 0.000465
39	Adenocarcinoma	0.771	0.161	0.807	0.809	X.Ruminococcus._gnavus_group	0.000	0.011	X.Ruminococcus._gnavus_group <= 0.000763
39	Adenocarcinoma	0.771	0.161	0.807	0.809	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
39	Adenocarcinoma	0.771	0.161	0.807	0.809	Ruminococcus	0.002	-0.007	Ruminococcus <= 0.00486
39	Adenocarcinoma	0.771	0.161	0.807	0.809	Negativibacillus	0.001	-0.006	0.000172 < Negativibacillus
39	Adenocarcinoma	0.771	0.161	0.807	0.809	UCG.011	0.000	-0.005	UCG.011 <= 3.48e-05
40	Adenocarcinoma	0.764	0.164	0.811	0.803	X.Ruminococcus._gnavus_group	0.010	-0.012	0.000763 < X.Ruminococcus._gnavus_group
40	Adenocarcinoma	0.764	0.164	0.811	0.803	Eisenbergiella	0.000	0.009	Eisenbergiella <= 0.000267
40	Adenocarcinoma	0.764	0.164	0.811	0.803	Butyricimonas	0.008	-0.007	0.006295 < Butyricimonas
40	Adenocarcinoma	0.764	0.164	0.811	0.803	Negativibacillus	0.000	0.006	Negativibacillus <= 0.000172
40	Adenocarcinoma	0.764	0.164	0.811	0.803	Saccharimonadaceae	0.000	-0.005	Saccharimonadaceae <= 8.16e-06

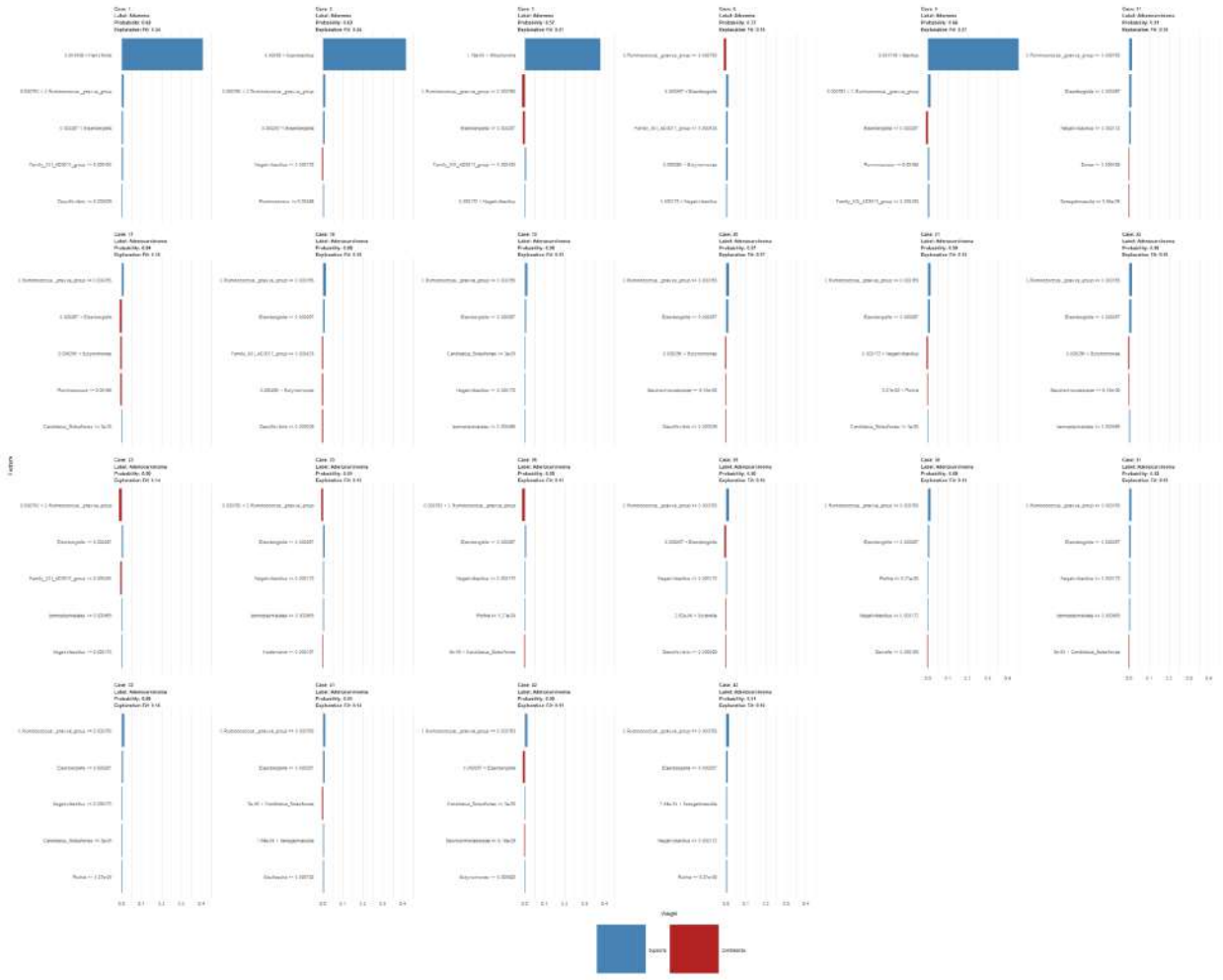


Figure ML2-E-6: The LIME Result in Training dataset for Genus in Stool

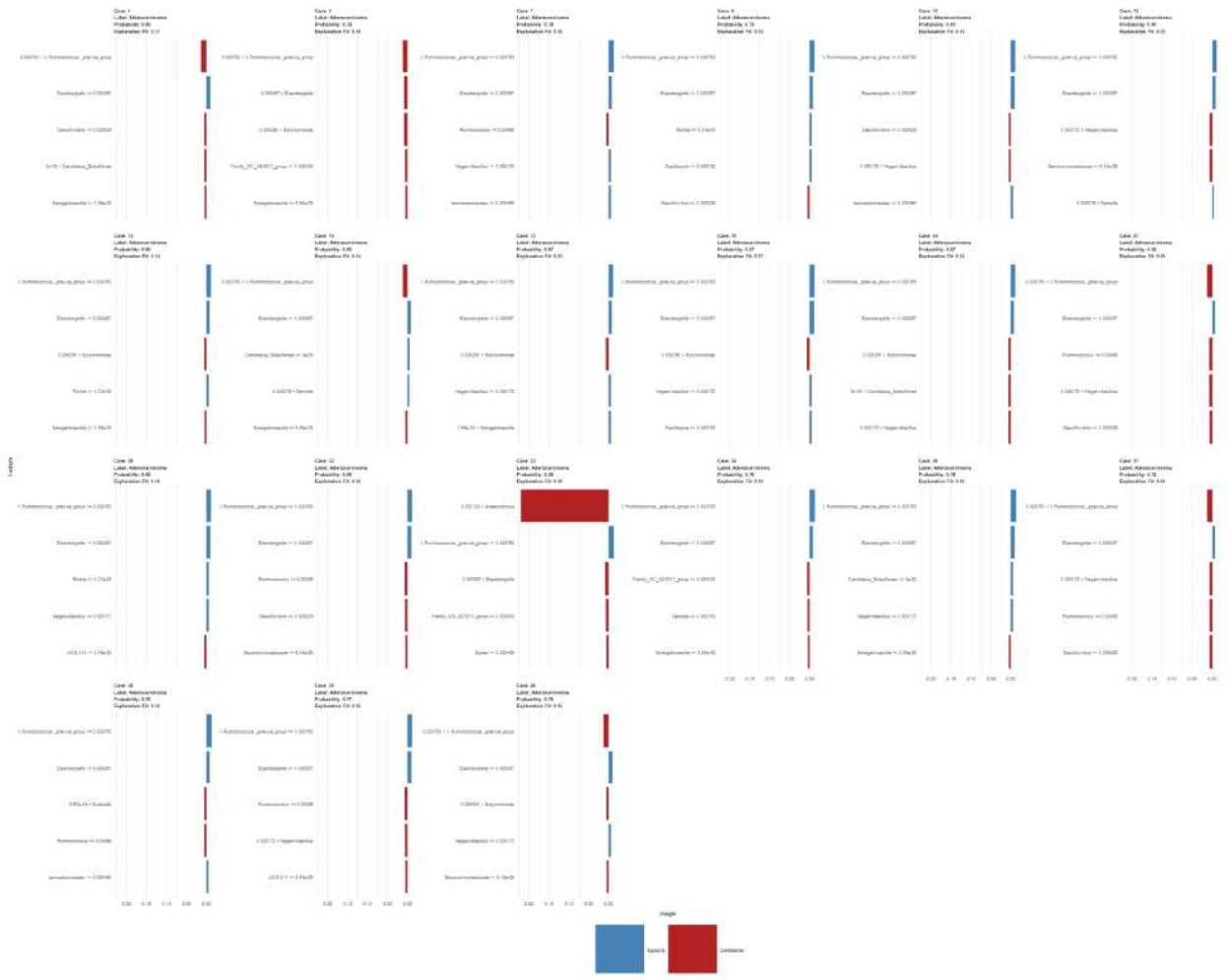


Figure ML2-E-7: The LIME Result in Test dataset for Genus in Stool

Table ML2-E-6: The Aggregated SHAP Values

Variable	Training Dataset				Testing Dataset			
	contribution	position	sign	cumulative	contribution	position	sign	cumulative
intercept	0.7769	187	X	0.777	0.7769	187	X	0.777
UCG.003	0.0000	186	1	0.777	-0.0005	185	-1	0.777
Atopobium	0.0000	185	1	0.777	-0.0002	26	-1	0.781
Defluviitaleaceae_UCG.011	0.0000	184	-1	0.777	-0.0001	183	-1	0.777
Lachnospiraceae_UCG.001	0.0000	183	-1	0.777	0.0005	180	1	0.777

Klebsiella	0.0000	182	-1	0.777	0.0003	176	1	0.776
Tumor	0.0000	181	1	0.777	0.0000	174	1	0.777
Gender	0.0000	180	1	0.777	0.0000	173	1	0.777
Adlercreutzia	0.0000	179	-1	0.777	0.0005	171	1	0.777
Megamonas	0.0000	178	-1	0.777	0.0001	184	1	0.777
Lachnospira	0.0000	177	-1	0.777	-0.0002	170	-1	0.777
Candidatus_Soleaferrea	0.0000	176	1	0.777	-0.0012	161	-1	0.775
Peptoniphilus	0.0000	175	-1	0.777	0.0002	14	1	0.782
Turicibacter	0.0000	174	-1	0.777	0.0001	168	1	0.777
UCG.002	0.0000	173	1	0.777	-0.0021	164	-1	0.776
CAG.56	0.000	172	1	0.777	0.0002	169	1	0.777
UCG.005	0.000	171	-1	0.777	-0.0016	163	-1	0.774
Serratia	0.000	170	-1	0.777	0.0002	155	1	0.776
Senegalimassilia	0.000	169	1	0.777	0.0001	55	1	0.782
Escherichia.Shigella	0.000	168	-1	0.777	0.0008	166	1	0.778
Faecalibacterium	0.000	167	-1	0.777	-0.0007	134	-1	0.777
Holdemanella	0.000	166	1	0.777	0.0001	132	1	0.777
Victivallaceae	0.000	165	-1	0.777	-0.0001	103	-1	0.783
Ruminiclostridium	0.000	164	1	0.777	-0.0007	165	-1	0.778
Terrisporobacter	0.000	163	-1	0.777	-0.0007	79	-1	0.785
X.Eubacterium._siraeum_group	0.000	162	-1	0.777	-0.0002	40	-1	0.783
Incertae_Sedis	0.000	161	-1	0.777	0.0003	167	1	0.778
TM7a	0.000	160	1	0.777	-0.0003	154	-1	0.776
Phascolarctobacterium	0.000	159	1	0.777	0.0000	179	-1	0.777
Saccharimonadales	0.000	158	-1	0.777	0.0005	147	1	0.777
Family_XIII_UCG.001	0.000	157	1	0.777	0.0005	71	1	0.783
Lachnospiraceae_NK4A136_group	0.000	156	-1	0.777	0.0006	151	1	0.778
Oscillibacter	0.000	155	-1	0.777	-0.0004	146	-1	0.777
Bacillus	0.000	154	1	0.777	0.0000	141	1	0.776
UCG.010	0.000	153	1	0.777	-0.0004	144	-1	0.777
Methanobrevibacter	0.000	152	1	0.777	-0.0003	159	-1	0.774
Rothia	0.000	151	-1	0.777	-0.0005	142	-1	0.776
Dorea	0.000	150	-1	0.777	0.0007	140	1	0.777
Peptococcus	0.000	149	1	0.777	-0.0001	139	-1	0.776
X.Eubacterium._hallii_group	0.000	148	1	0.777	0.0004	45	1	0.783
vadinBE97	0.000	147	1	0.777	0.0001	129	1	0.779
Colidextribacter	0.000	146	-1	0.777	0.0004	148	1	0.777
Methanosphaera	0.000	145	1	0.777	0.0001	135	1	0.778
Alloprevotella	0.000	144	1	0.777	0.0000	84	1	0.784
Asteroleplasma	0.000	143	1	0.777	0.0000	172	1	0.777
Desulfovibrio	0.000	142	1	0.777	0.0006	35	1	0.782
Solobacterium	0.000	141	-1	0.777	0.0001	133	1	0.777
Monoglobus	0.000	140	1	0.777	-0.0004	150	-1	0.777

Slackia	0.000	139	-1	0.776887	0.0007	153	1	0.777
Pseudomonas	0.000	138	1	0.776892	0.0000	52	1	0.782
Peptostreptococcus	0.000	137	1	0.776895	-0.0001	127	-1	0.778
Bifidobacterium	0.000	136	-1	0.776838	0.0009	126	1	0.779
Blautia	0.000	135	1	0.776841	0.0001	125	1	0.780
Gordonibacter	0.000	134	-1	0.776811	0.0011	130	1	0.778
RF39	0.000	133	-1	0.776793	0.0002	88	1	0.785
Enterococcus	0.000	132	-1	0.776789	0.0000	145	1	0.777
X.Eubacterium._brachy_group	0.000	131	1	0.776811	-0.0016	72	-1	0.783
X.Eubacterium._ventriosum_group	0.000	130	-1	0.776778	-0.0004	61	-1	0.781
Alistipes	0.000	129	1	0.776789	0.0002	112	1	0.783
Clostridia_UCG.014	0.000	128	-1	0.776752	-0.0001	121	-1	0.780
Campylobacter	0.000	127	-1	0.776745	-0.0005	178	-1	0.776
Oscillospira	0.000	126	-1	0.776738	0.0018	162	1	0.776
X.Ruminococcus._gauvreauii_group	0.000	125	-1	0.776726	0.0009	119	1	0.782
Eisenbergiella	0.000	124	1	0.776798	0.0014	83	1	0.786
Allisonella	0.000	123	1	0.776814	0.0004	118	1	0.783
Clostridia_vadinBB60_group	0.000	122	-1	0.776751	-0.0004	98	-1	0.785
Negativibacillus	0.000	121	-1	0.77674	-0.0006	116	-1	0.782
Enterorhabdus	0.000	120	1	0.77676	0.0001	104	1	0.783
Akkermansia	0.000	119	-1	0.776757	-0.0001	113	-1	0.782
Lachnospiraceae_UCG.004	0.000	118	1	0.776777	0.0000	77	-1	0.784
Acinetobacter	0.000	117	-1	0.776753	0.0003	66	1	0.783
Hafnia.Obesumbacterium	0.000	116	-1	0.776748	0.0000	122	1	0.780
Hydrogenoanaerobacterium	0.000	115	-1	0.776726	0.0000	16	1	0.782
Lachnospiraceae_UCG.008	0.000	114	-1	0.7767	0.0001	105	1	0.783
DTU089	0.000	113	1	0.7767	-0.0001	54	-1	0.782
Gastranaerophilales	0.000	112	1	0.776703	0.0000	106	-1	0.783
Sutterella	0.000	111	-1	0.776684	-0.0005	65	-1	0.782
Holdemania	0.000	110	-1	0.776668	0.0021	120	1	0.782
Sellimonas	0.000	109	1	0.7767	-0.0003	89	-1	0.784
Barnesiella	0.000	108	1	0.776743	0.0002	138	1	0.777
Libanicoccus	0.000	107	1	0.776751	0.0000	110	-1	0.783
Bilophila	0.000	106	1	0.777	0.0006	102	1	0.784
Prevotellaceae_UCG.001	0.000	105	-1	0.777	0.0008	100	1	0.785
UCG.011	0.000	104	1	0.777	-0.0005	46	-1	0.783
Butyrivibrio	0.000	103	1	0.777	-0.0001	117	-1	0.783
S5.A14a	0.000	102	1	0.777	0.0000	44	-1	0.783
Tyzzerella	0.000	101	1	0.777	0.0002	96	1	0.785
Victivallis	0.000	100	-1	0.777	0.0001	27	1	0.781
Mitsuokella	0.000	99	1	0.777	-0.0001	137	-1	0.777
X.Eubacterium._eligens_group	0.000	98	-1	0.777	0.0003	91	1	0.785
Muribaculaceae	0.000	97	-1	0.777	-0.0001	62	-1	0.782

X.Eubacterium._nodatum_group	0.000	96	1	0.777	-0.0002	109	-1	0.783
Erysipelotrichaceae_UCG.003	0.000	95	-1	0.777	-0.0001	86	-1	0.784
Fournierella	0.000	94	1	0.777	0.0001	85	1	0.784
Clostridium_sensu_stricto_1	0.000	93	1	0.777	0.0004	108	1	0.784
Prevotellaceae_NK3B31_group	0.000	92	-1	0.777	0.0000	7	-1	0.781
X.Eubacterium._ruminantium_group	0.000	91	1	0.777	-0.0002	78	-1	0.784
Anaerostipes	0.000	90	1	0.777	0.0000	4	-1	0.781
Cloacibacillus	0.000	89	-1	0.777	-0.0003	95	-1	0.784
Olsenella	0.000	88	-1	0.777	-0.0006	82	-1	0.785
Enterobacter	0.000	87	-1	0.777	-0.0001	80	-1	0.785
Lactobacillus	0.000	86	1	0.777	-0.0001	160	-1	0.774
Sanguibacteroides	0.000	85	1	0.777	0.0000	42	1	0.784
Saccharimonadaceae	0.000	84	1	0.777	0.0004	28	1	0.781
Mitochondria	0.000	83	1	0.777	-0.0003	74	-1	0.784
Coprobacter	0.000	82	-1	0.777	0.0000	101	-1	0.784
Christensenellaceae_R.7_group	0.000	81	-1	0.777	-0.0010	181	-1	0.776
Coprobacillus	0.000	80	1	0.777	0.0002	158	1	0.774
Dialister	0.000	79	-1	0.777	0.0004	69	1	0.783
Family_XIII_AD3011_group	0.000	78	-1	0.777	0.0010	48	1	0.784
Intestinimonas	0.000	77	1	0.777	-0.0002	107	-1	0.783
Merdibacter	0.000	76	-1	0.777	-0.0004	67	-1	0.782
Anaerococcus	0.000	75	-1	0.777	-0.0004	17	-1	0.782
Parvimonas	0.000	74	-1	0.777	-0.0001	15	-1	0.782
Catenibacterium	0.000	73	-1	0.777	0.0001	81	1	0.785
Haemophilus	0.000	72	1	0.777	0.0020	24	1	0.783
Staphylococcus	0.000	71	-1	0.777	0.0005	60	1	0.782
Anaerotruncus	0.000	70	1	0.777	0.0003	99	1	0.785
Howardella	0.000	69	1	0.777	0.0000	59	1	0.782
Izemoplasmatales	0.000	68	-1	0.777	0.0012	157	1	0.775
Eggerthella	0.000	67	1	0.777	-0.0002	58	-1	0.782
Megasphaera	0.000	66	-1	0.777	0.0001	56	1	0.782
Parabacteroides	0.000	65	1	0.777	0.0000	25	-1	0.781
Rikenellaceae_RC9_gut_group	0.000	64	1	0.777	0.0002	57	1	0.782
Butyricicoccus	0.000	63	1	0.777	-0.0010	9	-1	0.781
Fusicatenibacter	0.000	62	-1	0.777	0.0008	111	1	0.783
Veillonella	0.000	61	-1	0.777	0.0009	186	1	0.778
Epulopiscium	0.000	60	1	0.777	-0.0001	75	-1	0.784
Frisingicoccus	0.000	59	-1	0.777	-0.0004	21	-1	0.783
Lactococcus	0.000	58	1	0.777	-0.0002	68	-1	0.783
Odoribacter	0.000	57	1	0.777	-0.0009	149	-1	0.776
Streptococcus	0.000	56	-1	0.777	-0.0008	70	-1	0.782
UBA1819	0.000	55	1	0.777	-0.0005	47	-1	0.783
Christensenella	0.000	54	1	0.777	0.0000	90	1	0.785

Collinsella	0.000	53	1	0.777	0.0006	175	1	0.777
Subdoligranulum	0.000	52	1	0.777	0.0002	73	1	0.784
Kocuria	0.000	51	1	0.777	0.0003	131	1	0.777
Leuconostoc	0.000	50	-1	0.777	0.0001	76	1	0.784
Pyramidobacter	0.000	49	1	0.777	0.0002	92	1	0.784
uncultured	0.000	48	1	0.777	0.0006	156	1	0.776
Dielma	0.000	47	1	0.777	0.0004	152	1	0.777
Fusobacterium	0.000	46	-1	0.777	-0.0003	143	-1	0.776
Citrobacter	0.000	45	-1	0.777	-0.0005	32	-1	0.780
Flavonifractor	0.000	44	1	0.777	0.0003	124	1	0.780
Lachnospiraceae_ND3007_group	0.000	43	1	0.777	-0.0010	38	-1	0.782
Leptotrichia	0.000	42	-1	0.777	-0.0001	93	-1	0.784
NK4A214_group	0.000	41	-1	0.777	-0.0002	37	-1	0.782
Roseburia	0.000	40	-1	0.777	0.0001	51	1	0.782
Actinomyces	0.000	39	1	0.777	0.0002	20	1	0.784
Intestinibacter	0.000	38	-1	0.777	0.0002	36	1	0.782
X.Ruminococcus._gnavus_group	0.000	37	1	0.777	0.0003	114	1	0.782
Ruminococcus	0.000	36	-1	0.777	-0.0019	34	-1	0.781
TM7x	0.000	35	-1	0.777	0.0005	50	1	0.783
Alloscardovia	0.000	34	1	0.777	0.0000	64	1	0.782
GCA.900066575	0.000	33	-1	0.777	0.0010	31	1	0.781
Romboutsia	0.000	32	-1	0.777	-0.0005	29	-1	0.781
Porphyromonas	0.000	31	-1	0.777	-0.0004	177	-1	0.776
X.Ruminococcus._torques_group	0.000	30	-1	0.777	0.0005	49	1	0.783
Gemella	0.000	29	1	0.777	0.0009	136	1	0.777
Parasutterella	0.000	28	1	0.777	0.0000	182	1	0.777
Erysipelatoclostridium	0.000	27	-1	0.777	-0.0004	18	-1	0.783
Marvinbryantia	0.000	26	-1	0.777	0.0000	23	1	0.783
Coprococcus	0.000	25	1	0.777	0.0009	22	1	0.784
Eubacterium	0.000	24	-1	0.777	-0.0003	123	-1	0.780
Morganella	0.000	23	-1	0.777	0.0001	128	1	0.779
Harryflintia	0.000	22	-1	0.777	0.0004	30	1	0.781
Neisseria	0.000	21	-1	0.777	-0.0002	97	-1	0.784
Faecalitalea	0.000	20	-1	0.777	-0.0002	13	-1	0.782
Lachnoclostridium	0.000	19	1	0.777	0.0002	43	1	0.784
Paraprevotella	0.000	18	-1	0.777	0.0001	8	1	0.781
X.Clostridium._methylpentosum_group	0.000	17	1	0.777	-0.0001	12	-1	0.782
Hungatella	0.000	16	1	0.777	0.0001	94	1	0.784
X.Clostridium._innocuum_group	0.000	15	1	0.777	-0.0003	11	-1	0.782
Bacteroides	0.000	14	-1	0.777	0.0001	10	1	0.782
Granulicatella	0.000	13	-1	0.777	-0.0004	63	-1	0.782
X.Eubacterium._xylanophilum_group	0.000	12	-1	0.777	0.0000	115	1	0.782
Butyricimonas	0.000	11	-1	0.777	-0.0005	33	-1	0.780

Oribacterium	0.000	10	-1	0.777	0.0002	53	1	0.782
Prevotella	0.000	9	1	0.777	-0.0002	87	-1	0.784
X.Eubacterium._coprostanoligenes_group	0.000	8	1	0.777	-0.0006	41	-1	0.783
Abiotrophia	0.000	7	1	0.777	0.0000	6	1	0.781
Acidaminococcus	0.000	6	1	0.777	0.0000	5	1	0.781
CAG.873	0.000	5	-1	0.777	-0.0001	3	-1	0.781
Chloroplast	0.000	4	-1	0.777	-0.0001	2	-1	0.781
DTU014	0.000	3	-1	0.777	-0.0005	19	-1	0.783
Lachnospiraceae_FCS020_group	0.000	2	1	0.777	0.0001	39	1	0.783
prediction	0.777	1	X	0.777	0.7807	1	X	0.781

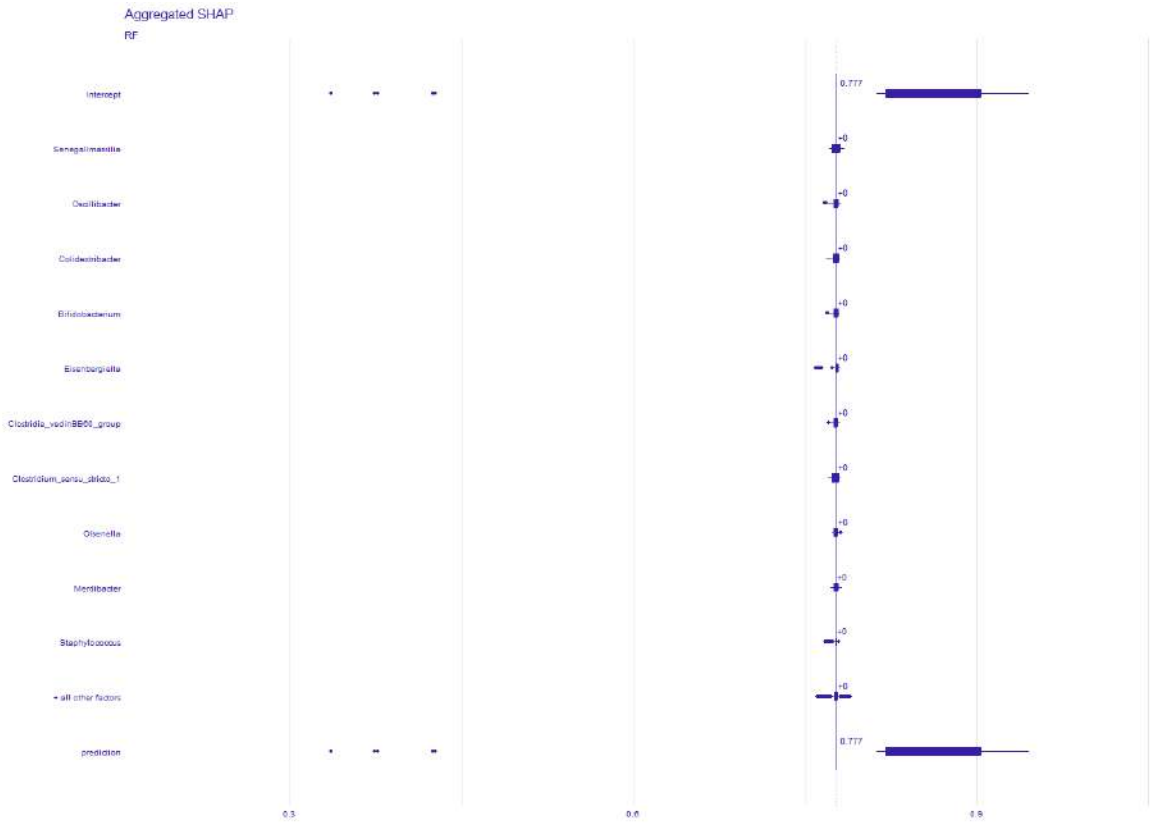


Figure ML2-E-8: The Aggregated SHAP Value in Training dataset for Genus in Stool

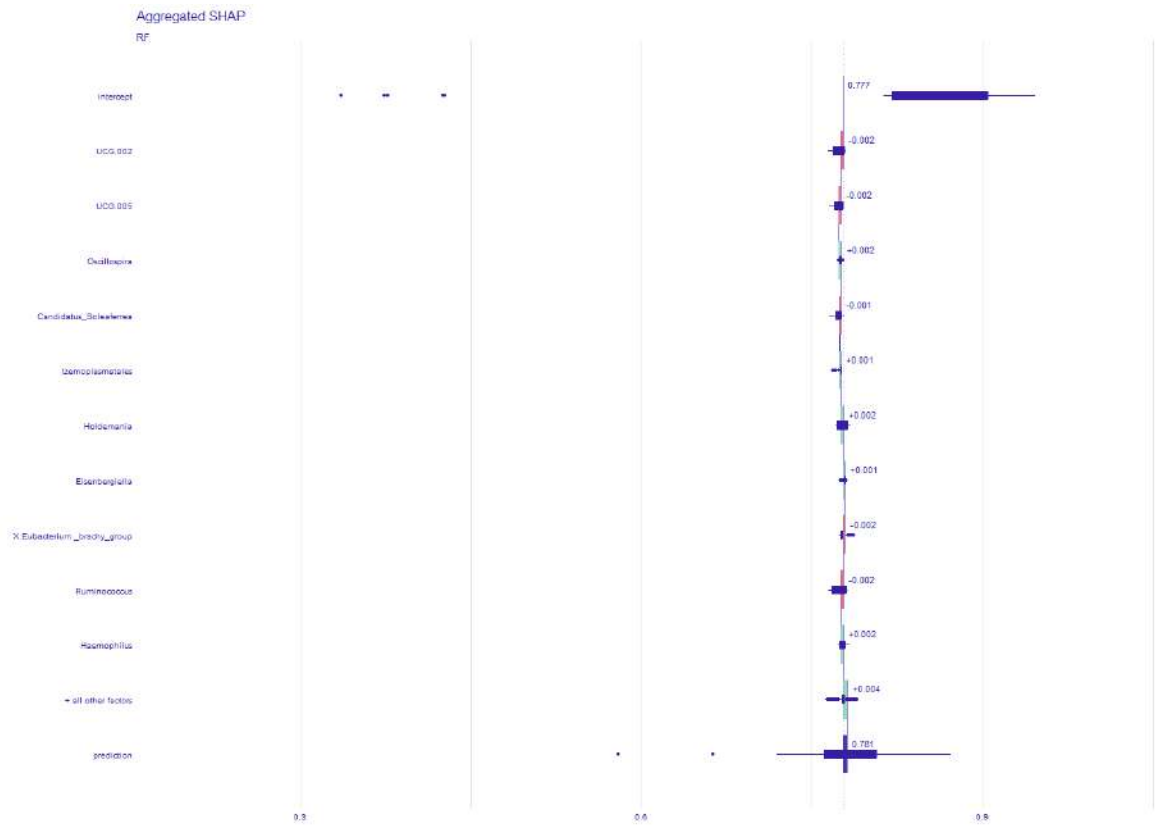


Figure ML2-E-9: The Aggregated SHAP Value in Test dataset for Genus in Stool

ML3 - Saliva

A- Phylum

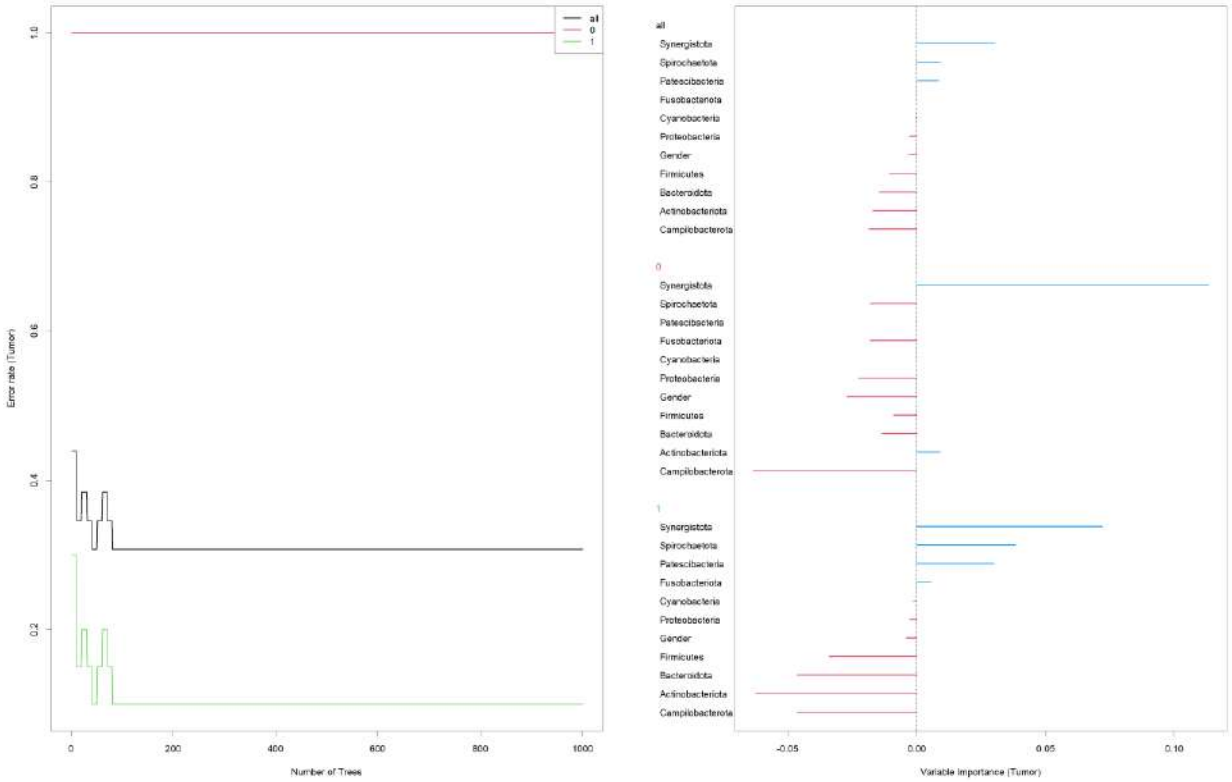


Figure ML3-A-1: The number of trees and Variable Importance Plot for Phylum in Saliva

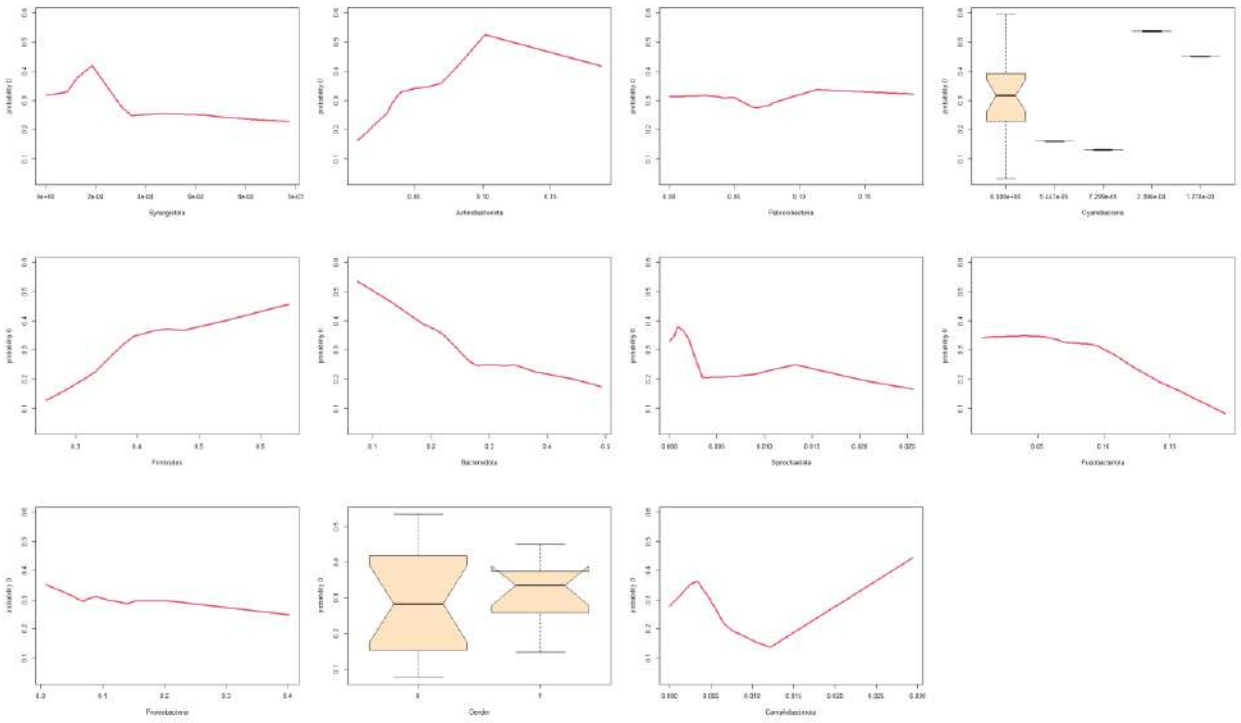


Figure ML3-A-2: Marginal Plot for Phylum in Saliva

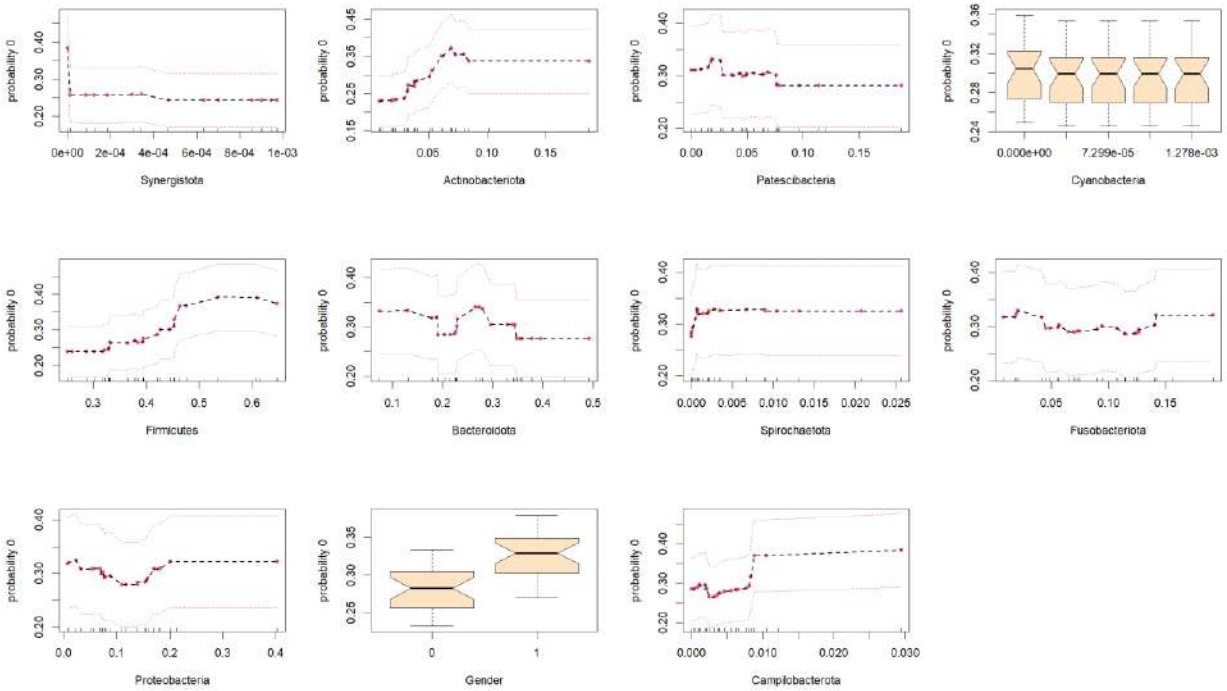


Figure ML3-A-3: Partial Plot for Phylum in Saliva

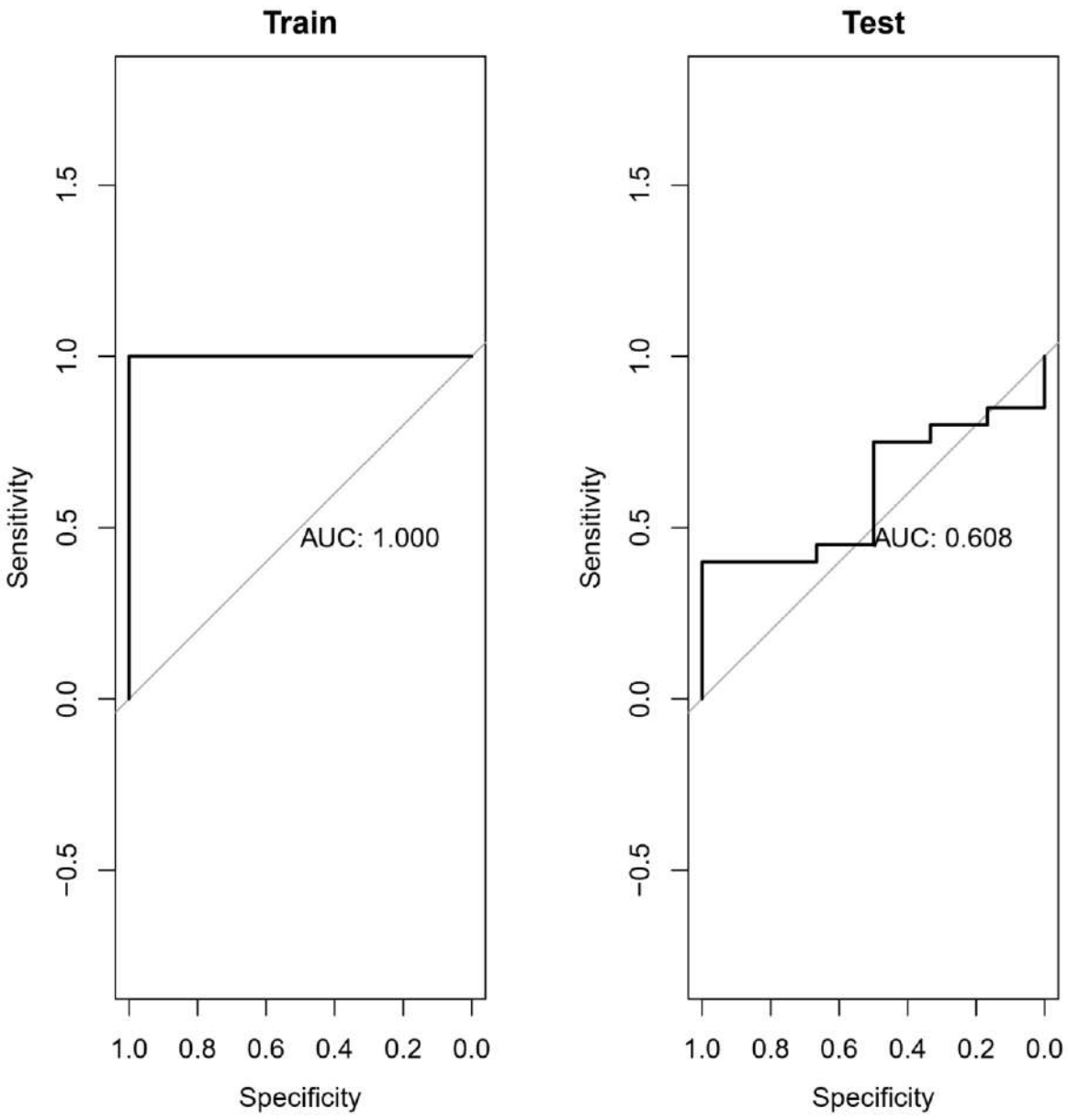


Figure ML3-A-4: The ROC Curve for Phylum in Saliva

Table ML3-A-1: The Model Performance

Index	Values		
	2.50%	**50.00%**	**97.50%**
Accuracy	0.38	0.62	0.85
Specificity	0.50	1.00	1.00
Sensitivity	0.20	0.50	0.90

Table ML3-A-2: The Variable Importance Values

Covariates	Varibale Importance			Brier Score		
	all	**AP**	**CRC**	**all**	**AP**	**CRC**
Gender	0.0001	-0.0091	0.0027	-0.0067	-0.0045	-0.0045
Actinobacteriota	-0.0427	0.0544	-0.1658	-0.0801	-0.0545	-0.0545
Bacteroidota	-0.0160	-0.0544	-0.0394	-0.0759	-0.0516	-0.0516
Campilobacterota	0.0136	-0.3715	0.1590	0.0084	0.0057	0.0057
Cyanobacteria	-0.0054	0.0000	-0.0190	-0.0074	-0.0050	-0.0050
Firmicutes	-0.0063	-0.0498	-0.0068	-0.0279	-0.0190	-0.0190
Fusobacteriota	0.0061	-0.1404	0.0639	-0.0326	-0.0222	-0.0222
Patescibacteria	-0.0108	0.0227	-0.0449	-0.0379	-0.0258	-0.0258
Proteobacteria	-0.0013	-0.1314	0.0353	-0.0191	-0.0130	-0.0130
Spirochaetota	-0.0053	0.0453	-0.0326	-0.0077	-0.0052	-0.0052
Synergistota	0.0664	0.1948	0.1740	0.2263	0.1538	0.1538

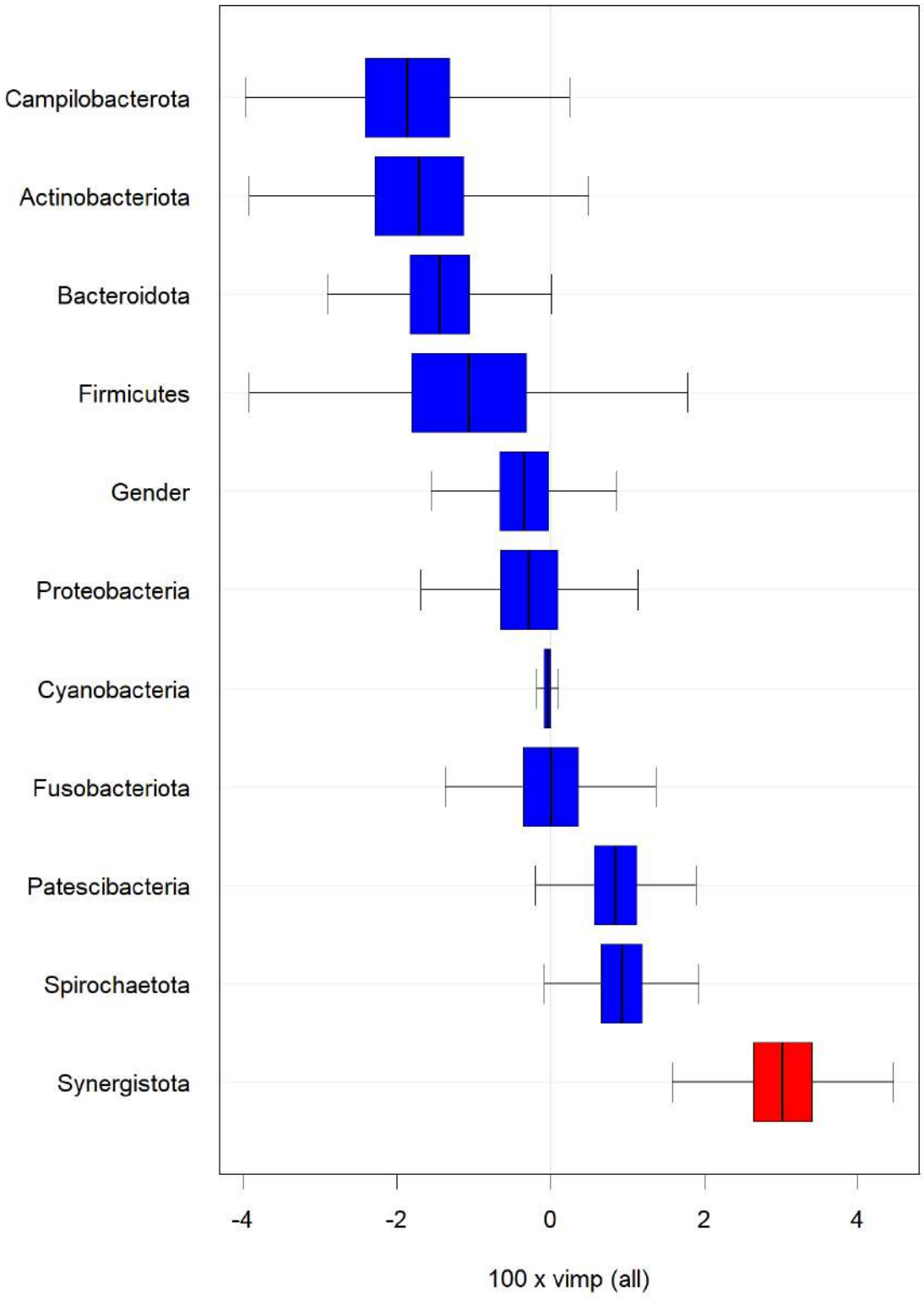


Figure ML3-A-5: The Variable Importance Plot (VIMP) for Phylum in Saliva

Best model is selected for LIME results

Table ML3-S-3: The best setting for model and LIME Results

Setting	mtry	ROC	Sens	Spec	ROCSD	SensSD	SpecSD
ntrees: 500 nodesize: 1	1	0.5	0	1	0.306186	0	0
ntrees: 500 nodesize: 1	2	0.5	0	0.95	0.306186	0	0.111803
ntrees: 1000 nodesize: 1	1	0.5	0	1	0.306186	0	0
ntrees: 1000 nodesize: 1	2	0.5	0	0.95	0.306186	0	0.111803
ntrees: 500 nodesize: 5	1	0.4	0	1	0.285044	0	0
ntrees: 1000 nodesize: 5	1	0.425	0	1	0.273861	0	0

Table ML3-A-4: The LIME Result for Training Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
2	Adenoma	0.580	0.121	0.198	0.239	Campilobacterota	0.009	0.044	0.00743 < Campilobacterota
2	Adenoma	0.580	0.121	0.198	0.239	Patescibacteria	0.026	0.019	Patescibacteria <= 0.0263
2	Adenoma	0.580	0.121	0.198	0.239	Actinobacteriota	0.033	-0.019	0.0240 < Actinobacteriota <= 0.0392
2	Adenoma	0.580	0.121	0.198	0.239	Synergistota	0.000	-0.016	4.86e-05 < Synergistota <= 4.37e-04
2	Adenoma	0.580	0.121	0.198	0.239	Proteobacteria	0.170	0.013	0.1579 < Proteobacteria
3	Adenoma	0.654	0.295	0.164	0.299	Synergistota	0.000	0.044	Synergistota <= 4.86e-05
3	Adenoma	0.654	0.295	0.164	0.299	Campilobacterota	0.012	0.042	0.00743 < Campilobacterota
3	Adenoma	0.654	0.295	0.164	0.299	Actinobacteriota	0.069	0.038	0.0670 < Actinobacteriota
3	Adenoma	0.654	0.295	0.164	0.299	Bacteroidota	0.188	0.034	Bacteroidota <= 0.207
3	Adenoma	0.654	0.295	0.164	0.299	Spirochaetota	0.000	-0.022	Spirochaetota <= 4.14e-05
6	Adenoma	0.716	0.348	0.162	0.308	Firmicutes	0.608	0.060	0.453 < Firmicutes
6	Adenoma	0.716	0.348	0.162	0.308	Synergistota	0.000	0.039	Synergistota <= 4.86e-05
6	Adenoma	0.716	0.348	0.162	0.308	Actinobacteriota	0.080	0.039	0.0670 < Actinobacteriota
6	Adenoma	0.716	0.348	0.162	0.308	Bacteroidota	0.131	0.033	Bacteroidota <= 0.207
6	Adenoma	0.716	0.348	0.162	0.308	Campilobacterota	0.006	-0.026	0.00371 < Campilobacterota <= 0.00743
7	Adenoma	0.678	0.247	0.174	0.281	Firmicutes	0.536	0.060	0.453 < Firmicutes
7	Adenoma	0.678	0.247	0.174	0.281	Synergistota	0.000	0.040	Synergistota <= 4.86e-05
7	Adenoma	0.678	0.247	0.174	0.281	Actinobacteriota	0.038	-0.018	0.0240 < Actinobacteriota <= 0.0392
7	Adenoma	0.678	0.247	0.174	0.281	Patescibacteria	0.018	0.013	Patescibacteria <= 0.0263
7	Adenoma	0.678	0.247	0.174	0.281	Spirochaetota	0.003	0.012	8.31e-04 < Spirochaetota <= 5.88e-03
10	Adenoma	0.726	0.277	0.167	0.301	Firmicutes	0.464	0.062	0.453 < Firmicutes
10	Adenoma	0.726	0.277	0.167	0.301	Synergistota	0.000	0.038	Synergistota <= 4.86e-05
10	Adenoma	0.726	0.277	0.167	0.301	Actinobacteriota	0.061	0.033	0.0392 < Actinobacteriota <= 0.0670
10	Adenoma	0.726	0.277	0.167	0.301	Spirochaetota	0.001	0.014	4.14e-05 < Spirochaetota <= 8.31e-04
10	Adenoma	0.726	0.277	0.167	0.301	Fusobacteriota	0.057	-0.013	0.0502 < Fusobacteriota <= 0.0823
11	Adenoma	0.696	0.166	0.178	0.265	Synergistota	0.000	0.040	Synergistota <= 4.86e-05
11	Adenoma	0.696	0.166	0.178	0.265	Actinobacteriota	0.070	0.037	0.0670 < Actinobacteriota
11	Adenoma	0.696	0.166	0.178	0.265	Campilobacterota	0.002	-0.021	0.00197 < Campilobacterota <= 0.00371
11	Adenoma	0.696	0.166	0.178	0.265	Fusobacteriota	0.021	0.016	Fusobacteriota <= 0.0502

11	Adenoma	0.696	0.166	0.178	0.265	Spirochaetota	0.001	0.015	4.14e-05 < Spirochaetota <= 8.31e-04
14	Adenocarcinoma	0.848	0.275	0.811	0.765	Synergistota	0.000	-0.043	Synergistota <= 4.86e-05
14	Adenocarcinoma	0.848	0.275	0.811	0.765	Patescibacteria	0.114	0.039	0.0761 < Patescibacteria
14	Adenocarcinoma	0.848	0.275	0.811	0.765	Actinobacteriota	0.073	-0.036	0.0670 < Actinobacteriota
14	Adenocarcinoma	0.848	0.275	0.811	0.765	Bacteroidota	0.179	-0.030	Bacteroidota <= 0.207
14	Adenocarcinoma	0.848	0.275	0.811	0.765	Spirochaetota	0.000	0.024	Spirochaetota <= 4.14e-05
20	Adenocarcinoma	0.902	0.281	0.790	0.819	Firmicutes	0.329	0.046	Firmicutes <= 0.329
20	Adenocarcinoma	0.902	0.281	0.790	0.819	Synergistota	0.000	-0.038	Synergistota <= 4.86e-05
20	Adenocarcinoma	0.902	0.281	0.790	0.819	Patescibacteria	0.187	0.034	0.0761 < Patescibacteria
20	Adenocarcinoma	0.902	0.281	0.790	0.819	Actinobacteriota	0.051	-0.034	0.0392 < Actinobacteriota <= 0.0670
20	Adenocarcinoma	0.902	0.281	0.790	0.819	Spirochaetota	0.000	0.022	Spirochaetota <= 4.14e-05
21	Adenocarcinoma	0.930	0.379	0.757	0.878	Actinobacteriota	0.010	0.050	Actinobacteriota <= 0.0240
21	Adenocarcinoma	0.930	0.379	0.757	0.878	Firmicutes	0.321	0.048	Firmicutes <= 0.329
21	Adenocarcinoma	0.930	0.379	0.757	0.878	Campilobacterota	0.008	-0.045	0.00743 < Campilobacterota
21	Adenocarcinoma	0.930	0.379	0.757	0.878	Synergistota	0.001	0.036	4.37e-04 < Synergistota
21	Adenocarcinoma	0.930	0.379	0.757	0.878	Bacteroidota	0.356	0.031	0.347 < Bacteroidota
22	Adenocarcinoma	0.848	0.341	0.810	0.732	Firmicutes	0.476	-0.061	0.453 < Firmicutes
22	Adenocarcinoma	0.848	0.341	0.810	0.732	Campilobacterota	0.008	-0.042	0.00743 < Campilobacterota
22	Adenocarcinoma	0.848	0.341	0.810	0.732	Synergistota	0.001	0.042	4.37e-04 < Synergistota
22	Adenocarcinoma	0.848	0.341	0.810	0.732	Bacteroidota	0.191	-0.032	Bacteroidota <= 0.207
22	Adenocarcinoma	0.848	0.341	0.810	0.732	Actinobacteriota	0.038	0.016	0.0240 < Actinobacteriota <= 0.0392
23	Adenocarcinoma	0.872	0.149	0.812	0.728	Campilobacterota	0.010	-0.042	0.00743 < Campilobacterota
23	Adenocarcinoma	0.872	0.149	0.812	0.728	Actinobacteriota	0.084	-0.035	0.0670 < Actinobacteriota
23	Adenocarcinoma	0.872	0.149	0.812	0.728	Synergistota	0.000	0.019	4.86e-05 < Synergistota <= 4.37e-04
23	Adenocarcinoma	0.872	0.149	0.812	0.728	Patescibacteria	0.038	-0.014	0.0263 < Patescibacteria <= 0.0492
23	Adenocarcinoma	0.872	0.149	0.812	0.728	Spirochaetota	0.000	-0.012	4.14e-05 < Spirochaetota <= 8.31e-04
24	Adenocarcinoma	0.906	0.251	0.748	0.913	Firmicutes	0.288	0.049	Firmicutes <= 0.329
24	Adenocarcinoma	0.906	0.251	0.748	0.913	Synergistota	0.000	0.038	4.37e-04 < Synergistota
24	Adenocarcinoma	0.906	0.251	0.748	0.913	Patescibacteria	0.076	0.038	0.0761 < Patescibacteria
24	Adenocarcinoma	0.906	0.251	0.748	0.913	Campilobacterota	0.003	0.022	0.00197 < Campilobacterota <= 0.00371
24	Adenocarcinoma	0.906	0.251	0.748	0.913	Proteobacteria	0.139	0.018	0.0993 < Proteobacteria <= 0.1579
25	Adenocarcinoma	0.844	0.220	0.796	0.801	Synergistota	0.000	-0.041	Synergistota <= 4.86e-05
25	Adenocarcinoma	0.844	0.220	0.796	0.801	Bacteroidota	0.359	0.035	0.347 < Bacteroidota
25	Adenocarcinoma	0.844	0.220	0.796	0.801	Actinobacteriota	0.062	-0.035	0.0392 < Actinobacteriota <= 0.0670
25	Adenocarcinoma	0.844	0.220	0.796	0.801	Spirochaetota	0.000	0.023	Spirochaetota <= 4.14e-05
25	Adenocarcinoma	0.844	0.220	0.796	0.801	Campilobacterota	0.003	0.023	0.00197 < Campilobacterota <= 0.00371
26	Adenocarcinoma	0.958	0.180	0.757	0.890	Actinobacteriota	0.020	0.048	Actinobacteriota <= 0.0240
26	Adenocarcinoma	0.958	0.180	0.757	0.890	Patescibacteria	0.119	0.035	0.0761 < Patescibacteria
26	Adenocarcinoma	0.958	0.180	0.757	0.890	Campilobacterota	0.006	0.025	0.00371 < Campilobacterota <= 0.00743
26	Adenocarcinoma	0.958	0.180	0.757	0.890	Synergistota	0.000	0.014	4.86e-05 < Synergistota <= 4.37e-04
26	Adenocarcinoma	0.958	0.180	0.757	0.890	Firmicutes	0.330	0.011	0.329 < Firmicutes <= 0.389
29	Adenocarcinoma	0.938	0.374	0.768	0.886	Firmicutes	0.260	0.050	Firmicutes <= 0.329

29	Adenocarcinoma	0.938	0.374	0.768	0.886	Actinobacteriota	0.009	0.048	Actinobacteriota <= 0.0240
29	Adenocarcinoma	0.938	0.374	0.768	0.886	Synergistota	0.000	-0.042	Synergistota <= 4.86e-05
29	Adenocarcinoma	0.938	0.374	0.768	0.886	Bacteroidota	0.492	0.038	0.347 < Bacteroidota
29	Adenocarcinoma	0.938	0.374	0.768	0.886	Spirochaetota	0.000	0.023	Spirochaetota <= 4.14e-05
32	Adenocarcinoma	0.858	0.151	0.798	0.804	Synergistota	0.000	-0.040	Synergistota <= 4.86e-05
32	Adenocarcinoma	0.858	0.151	0.798	0.804	Patescibacteria	0.077	0.036	0.0761 < Patescibacteria
32	Adenocarcinoma	0.858	0.151	0.798	0.804	Actinobacteriota	0.038	0.018	0.0240 < Actinobacteriota <= 0.0392
32	Adenocarcinoma	0.858	0.151	0.798	0.804	Spirochaetota	0.001	-0.014	4.14e-05 < Spirochaetota <= 8.31e-04
32	Adenocarcinoma	0.858	0.151	0.798	0.804	Fusobacteriota	0.123	0.006	0.0823 < Fusobacteriota <= 0.1242
33	Adenocarcinoma	0.844	0.295	0.796	0.775	Firmicutes	0.453	-0.063	0.453 < Firmicutes
33	Adenocarcinoma	0.844	0.295	0.796	0.775	Bacteroidota	0.378	0.036	0.347 < Bacteroidota
33	Adenocarcinoma	0.844	0.295	0.796	0.775	Actinobacteriota	0.040	-0.035	0.0392 < Actinobacteriota <= 0.0670
33	Adenocarcinoma	0.844	0.295	0.796	0.775	Campilobacterota	0.004	0.025	0.00371 < Campilobacterota <= 0.00743
33	Adenocarcinoma	0.844	0.295	0.796	0.775	Synergistota	0.000	0.016	4.86e-05 < Synergistota <= 4.37e-04
34	Adenocarcinoma	0.910	0.192	0.772	0.843	Actinobacteriota	0.008	0.045	Actinobacteriota <= 0.0240
34	Adenocarcinoma	0.910	0.192	0.772	0.843	Synergistota	0.001	0.040	4.37e-04 < Synergistota
34	Adenocarcinoma	0.910	0.192	0.772	0.843	Campilobacterota	0.005	0.021	0.00371 < Campilobacterota <= 0.00743
34	Adenocarcinoma	0.910	0.192	0.772	0.843	Patescibacteria	0.017	-0.019	Patescibacteria <= 0.0263
34	Adenocarcinoma	0.910	0.192	0.772	0.843	Proteobacteria	0.214	-0.016	0.1579 < Proteobacteria
38	Adenocarcinoma	0.966	0.389	0.758	0.882	Actinobacteriota	0.018	0.047	Actinobacteriota <= 0.0240
38	Adenocarcinoma	0.966	0.389	0.758	0.882	Campilobacterota	0.008	-0.045	0.00743 < Campilobacterota
38	Adenocarcinoma	0.966	0.389	0.758	0.882	Firmicutes	0.252	0.045	Firmicutes <= 0.329
38	Adenocarcinoma	0.966	0.389	0.758	0.882	Patescibacteria	0.079	0.039	0.0761 < Patescibacteria
38	Adenocarcinoma	0.966	0.389	0.758	0.882	Synergistota	0.001	0.037	4.37e-04 < Synergistota
39	Adenocarcinoma	0.870	0.303	0.810	0.768	Firmicutes	0.613	-0.064	0.453 < Firmicutes
39	Adenocarcinoma	0.870	0.303	0.810	0.768	Synergistota	0.000	-0.039	Synergistota <= 4.86e-05
39	Adenocarcinoma	0.870	0.303	0.810	0.768	Spirochaetota	0.000	0.023	Spirochaetota <= 4.14e-05
39	Adenocarcinoma	0.870	0.303	0.810	0.768	Campilobacterota	0.004	0.021	0.00197 < Campilobacterota <= 0.00371
39	Adenocarcinoma	0.870	0.303	0.810	0.768	Actinobacteriota	0.030	0.018	0.0240 < Actinobacteriota <= 0.0392
41	Adenocarcinoma	0.838	0.090	0.800	0.768	Actinobacteriota	0.187	-0.039	0.0670 < Actinobacteriota
41	Adenocarcinoma	0.838	0.090	0.800	0.768	Synergistota	0.000	0.020	4.86e-05 < Synergistota <= 4.37e-04
41	Adenocarcinoma	0.838	0.090	0.800	0.768	Patescibacteria	0.046	-0.012	0.0263 < Patescibacteria <= 0.0492
41	Adenocarcinoma	0.838	0.090	0.800	0.768	Spirochaetota	0.001	-0.010	8.31e-04 < Spirochaetota <= 5.88e-03
41	Adenocarcinoma	0.838	0.090	0.800	0.768	Fusobacteriota	0.070	0.009	0.0502 < Fusobacteriota <= 0.0823
46	Adenocarcinoma	0.870	0.286	0.788	0.826	Firmicutes	0.318	0.048	Firmicutes <= 0.329
46	Adenocarcinoma	0.870	0.286	0.788	0.826	Synergistota	0.000	-0.038	Synergistota <= 4.86e-05
46	Adenocarcinoma	0.870	0.286	0.788	0.826	Bacteroidota	0.348	0.038	0.347 < Bacteroidota
46	Adenocarcinoma	0.870	0.286	0.788	0.826	Actinobacteriota	0.040	-0.032	0.0392 < Actinobacteriota <= 0.0670
46	Adenocarcinoma	0.870	0.286	0.788	0.826	Campilobacterota	0.002	0.022	0.00197 < Campilobacterota <= 0.00371
47	Adenocarcinoma	0.922	0.267	0.797	0.768	Campilobacterota	0.029	-0.040	0.00743 < Campilobacterota
47	Adenocarcinoma	0.922	0.267	0.797	0.768	Synergistota	0.001	0.037	4.37e-04 < Synergistota
47	Adenocarcinoma	0.922	0.267	0.797	0.768	Patescibacteria	0.096	0.036	0.0761 < Patescibacteria

47	Adenocarcinoma	0.922	0.267	0.797	0.768	Actinobacteriota	0.053	-0.033	0.0392 < Actinobacteriota <= 0.0670
47	Adenocarcinoma	0.922	0.267	0.797	0.768	Bacteroidota	0.204	-0.029	Bacteroidota <= 0.207
49	Adenocarcinoma	0.906	0.234	0.793	0.811	Actinobacteriota	0.022	0.048	Actinobacteriota <= 0.0240
49	Adenocarcinoma	0.906	0.234	0.793	0.811	Synergistota	0.000	-0.039	Synergistota <= 4.86e-05
49	Adenocarcinoma	0.906	0.234	0.793	0.811	Bacteroidota	0.204	-0.029	Bacteroidota <= 0.207
49	Adenocarcinoma	0.906	0.234	0.793	0.811	Spirochaetota	0.000	0.021	Spirochaetota <= 4.14e-05
49	Adenocarcinoma	0.906	0.234	0.793	0.811	Firmicutes	0.364	0.017	0.329 < Firmicutes <= 0.389
51	Adenocarcinoma	0.832	0.287	0.812	0.728	Firmicutes	0.646	-0.059	0.453 < Firmicutes
51	Adenocarcinoma	0.832	0.287	0.812	0.728	Actinobacteriota	0.102	-0.040	0.0670 < Actinobacteriota
51	Adenocarcinoma	0.832	0.287	0.812	0.728	Bacteroidota	0.074	-0.028	Bacteroidota <= 0.207
51	Adenocarcinoma	0.832	0.287	0.812	0.728	Campilobacterota	0.004	0.027	0.00371 < Campilobacterota <= 0.00743
51	Adenocarcinoma	0.832	0.287	0.812	0.728	Synergistota	0.000	0.017	4.86e-05 < Synergistota <= 4.37e-04
52	Adenocarcinoma	0.960	0.321	0.739	0.929	Firmicutes	0.302	0.047	Firmicutes <= 0.329
52	Adenocarcinoma	0.960	0.321	0.739	0.929	Actinobacteriota	0.020	0.046	Actinobacteriota <= 0.0240
52	Adenocarcinoma	0.960	0.321	0.739	0.929	Synergistota	0.001	0.038	4.37e-04 < Synergistota
52	Adenocarcinoma	0.960	0.321	0.739	0.929	Bacteroidota	0.396	0.035	0.347 < Bacteroidota
52	Adenocarcinoma	0.960	0.321	0.739	0.929	Campilobacterota	0.007	0.024	0.00371 < Campilobacterota <= 0.00743

Table ML-A-5: The LIME Result for Test Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
1	Adenocarcinoma	0.752	0.172	0.813	0.760	Synergistota	0.000	-0.040	Synergistota <= 4.86e-05
1	Adenocarcinoma	0.752	0.172	0.813	0.760	Actinobacteriota	0.073	-0.037	0.0670 < Actinobacteriota
1	Adenocarcinoma	0.752	0.172	0.813	0.760	Campilobacterota	0.004	0.026	0.00371 < Campilobacterota <= 0.00743
1	Adenocarcinoma	0.752	0.172	0.813	0.760	Proteobacteria	0.169	-0.014	0.1579 < Proteobacteria
1	Adenocarcinoma	0.752	0.172	0.813	0.760	Firmicutes	0.331	0.012	0.329 < Firmicutes <= 0.389
4	Adenocarcinoma	0.624	0.293	0.826	0.695	Firmicutes	0.474	-0.061	0.453 < Firmicutes
4	Adenocarcinoma	0.624	0.293	0.826	0.695	Campilobacterota	0.019	-0.044	0.00743 < Campilobacterota
4	Adenocarcinoma	0.624	0.293	0.826	0.695	Bacteroidota	0.140	-0.026	Bacteroidota <= 0.207
4	Adenocarcinoma	0.624	0.293	0.826	0.695	Proteobacteria	0.167	-0.017	0.1579 < Proteobacteria
4	Adenocarcinoma	0.624	0.293	0.826	0.695	Synergistota	0.000	0.017	4.86e-05 < Synergistota <= 4.37e-04
5	Adenocarcinoma	0.748	0.283	0.803	0.753	Firmicutes	0.627	-0.063	0.453 < Firmicutes
5	Adenocarcinoma	0.748	0.283	0.803	0.753	Synergistota	0.001	0.037	4.37e-04 < Synergistota
5	Adenocarcinoma	0.748	0.283	0.803	0.753	Bacteroidota	0.100	-0.030	Bacteroidota <= 0.207
5	Adenocarcinoma	0.748	0.283	0.803	0.753	Actinobacteriota	0.030	0.021	0.0240 < Actinobacteriota <= 0.0392
5	Adenocarcinoma	0.748	0.283	0.803	0.753	Spirochaetota	0.000	-0.015	4.14e-05 < Spirochaetota <= 8.31e-04
8	Adenocarcinoma	0.782	0.208	0.779	0.825	Actinobacteriota	0.102	-0.041	0.0670 < Actinobacteriota
8	Adenocarcinoma	0.782	0.208	0.779	0.825	Patescibacteria	0.100	0.038	0.0761 < Patescibacteria
8	Adenocarcinoma	0.782	0.208	0.779	0.825	Synergistota	0.001	0.037	4.37e-04 < Synergistota
8	Adenocarcinoma	0.782	0.208	0.779	0.825	Campilobacterota	0.004	0.026	0.00371 < Campilobacterota <= 0.00743

8	Adenocarcinoma	0.782	0.208	0.779	0.825	Spirochaetota	0.003	-0.013	8.31e-04 < Spirochaetota <= 5.88e-03
9	Adenocarcinoma	0.684	0.125	0.790	0.782	Bacteroidota	0.392	0.036	0.347 < Bacteroidota
9	Adenocarcinoma	0.684	0.125	0.790	0.782	Actinobacteriota	0.051	-0.032	0.0392 < Actinobacteriota <= 0.0670
9	Adenocarcinoma	0.684	0.125	0.790	0.782	Synergistota	0.000	0.019	4.86e-05 < Synergistota <= 4.37e-04
9	Adenocarcinoma	0.684	0.125	0.790	0.782	Fusobacteriota	0.045	-0.017	Fusobacteriota <= 0.0502
9	Adenocarcinoma	0.684	0.125	0.790	0.782	Patescibacteria	0.044	-0.015	0.0263 < Patescibacteria <= 0.0492
12	Adenocarcinoma	0.804	0.064	0.775	0.838	Campilobacterota	0.006	0.026	0.00371 < Campilobacterota <= 0.00743
12	Adenocarcinoma	0.804	0.064	0.775	0.838	Actinobacteriota	0.035	0.022	0.0240 < Actinobacteriota <= 0.0392
12	Adenocarcinoma	0.804	0.064	0.775	0.838	Proteobacteria	0.121	0.016	0.0993 < Proteobacteria <= 0.1579
12	Adenocarcinoma	0.804	0.064	0.775	0.838	Patescibacteria	0.036	-0.010	0.0263 < Patescibacteria <= 0.0492
12	Adenocarcinoma	0.804	0.064	0.775	0.838	Fusobacteriota	0.101	0.010	0.0823 < Fusobacteriota <= 0.1242
13	Adenocarcinoma	0.670	0.350	0.835	0.693	Firmicutes	0.616	-0.062	0.453 < Firmicutes
13	Adenocarcinoma	0.670	0.350	0.835	0.693	Synergistota	0.000	-0.040	Synergistota <= 4.86e-05
13	Adenocarcinoma	0.670	0.350	0.835	0.693	Actinobacteriota	0.127	-0.035	0.0670 < Actinobacteriota
13	Adenocarcinoma	0.670	0.350	0.835	0.693	Bacteroidota	0.136	-0.030	Bacteroidota <= 0.207
13	Adenocarcinoma	0.670	0.350	0.835	0.693	Spirochaetota	0.000	0.025	Spirochaetota <= 4.14e-05
15	Adenocarcinoma	0.700	0.215	0.826	0.725	Synergistota	0.000	-0.040	Synergistota <= 4.86e-05
15	Adenocarcinoma	0.700	0.215	0.826	0.725	Campilobacterota	0.009	-0.041	0.00743 < Campilobacterota
15	Adenocarcinoma	0.700	0.215	0.826	0.725	Actinobacteriota	0.044	-0.034	0.0392 < Actinobacteriota <= 0.0670
15	Adenocarcinoma	0.700	0.215	0.826	0.725	Proteobacteria	0.110	0.023	0.0993 < Proteobacteria <= 0.1579
15	Adenocarcinoma	0.700	0.215	0.826	0.725	Spirochaetota	0.001	-0.009	8.31e-04 < Spirochaetota <= 5.88e-03
16	Adenocarcinoma	0.684	0.316	0.806	0.745	Firmicutes	0.565	-0.065	0.453 < Firmicutes
16	Adenocarcinoma	0.684	0.316	0.806	0.745	Synergistota	0.001	0.037	4.37e-04 < Synergistota
16	Adenocarcinoma	0.684	0.316	0.806	0.745	Actinobacteriota	0.073	-0.036	0.0670 < Actinobacteriota
16	Adenocarcinoma	0.684	0.316	0.806	0.745	Campilobacterota	0.003	0.019	0.00197 < Campilobacterota <= 0.00371
16	Adenocarcinoma	0.684	0.316	0.806	0.745	Patescibacteria	0.019	-0.017	Patescibacteria <= 0.0263
17	Adenocarcinoma	0.852	0.168	0.774	0.840	Actinobacteriota	0.015	0.044	Actinobacteriota <= 0.0240
17	Adenocarcinoma	0.852	0.168	0.774	0.840	Patescibacteria	0.094	0.035	0.0761 < Patescibacteria
17	Adenocarcinoma	0.852	0.168	0.774	0.840	Synergistota	0.000	0.015	4.86e-05 < Synergistota <= 4.37e-04
17	Adenocarcinoma	0.852	0.168	0.774	0.840	Proteobacteria	0.162	-0.014	0.1579 < Proteobacteria
17	Adenocarcinoma	0.852	0.168	0.774	0.840	Fusobacteriota	0.042	-0.014	Fusobacteriota <= 0.0502
18	Adenocarcinoma	0.822	0.148	0.782	0.809	Actinobacteriota	0.021	0.047	Actinobacteriota <= 0.0240
18	Adenocarcinoma	0.822	0.148	0.782	0.809	Campilobacterota	0.005	0.026	0.00371 < Campilobacterota <= 0.00743
18	Adenocarcinoma	0.822	0.148	0.782	0.809	Patescibacteria	0.013	-0.017	Patescibacteria <= 0.0263
18	Adenocarcinoma	0.822	0.148	0.782	0.809	Spirochaetota	0.001	-0.015	4.14e-05 < Spirochaetota <= 8.31e-04
18	Adenocarcinoma	0.822	0.148	0.782	0.809	Fusobacteriota	0.039	-0.014	Fusobacteriota <= 0.0502
19	Adenocarcinoma	0.656	0.362	0.848	0.659	Firmicutes	0.592	-0.059	0.453 < Firmicutes
19	Adenocarcinoma	0.656	0.362	0.848	0.659	Synergistota	0.000	-0.041	Synergistota <= 4.86e-05

19	Adenocarcinoma	0.656	0.362	0.848	0.659	Campilobacterota	0.008	-0.044	0.00743 < Campilobacterota
19	Adenocarcinoma	0.656	0.362	0.848	0.659	Bacteroidota	0.089	-0.029	Bacteroidota <= 0.207
19	Adenocarcinoma	0.656	0.362	0.848	0.659	Patescibacteria	0.004	-0.015	Patescibacteria <= 0.0263
27	Adenocarcinoma	0.960	0.321	0.738	0.933	Firmicutes	0.326	0.048	Firmicutes <= 0.329
27	Adenocarcinoma	0.960	0.321	0.738	0.933	Actinobacteriota	0.008	0.047	Actinobacteriota <= 0.0240
27	Adenocarcinoma	0.960	0.321	0.738	0.933	Synergistota	0.001	0.038	4.37e-04 < Synergistota
27	Adenocarcinoma	0.960	0.321	0.738	0.933	Bacteroidota	0.378	0.033	0.347 < Bacteroidota
27	Adenocarcinoma	0.960	0.321	0.738	0.933	Campilobacterota	0.007	0.028	0.00371 < Campilobacterota <= 0.00743
28	Adenocarcinoma	0.650	0.326	0.825	0.735	Firmicutes	0.484	-0.060	0.453 < Firmicutes
28	Adenocarcinoma	0.650	0.326	0.825	0.735	Synergistota	0.000	-0.041	Synergistota <= 4.86e-05
28	Adenocarcinoma	0.650	0.326	0.825	0.735	Actinobacteriota	0.044	-0.035	0.0392 < Actinobacteriota <= 0.0670
28	Adenocarcinoma	0.650	0.326	0.825	0.735	Campilobacterota	0.004	0.026	0.00371 < Campilobacterota <= 0.00743
28	Adenocarcinoma	0.650	0.326	0.825	0.735	Proteobacteria	0.117	0.021	0.0993 < Proteobacteria <= 0.1579
30	Adenocarcinoma	0.858	0.123	0.777	0.825	Patescibacteria	0.149	0.035	0.0761 < Patescibacteria
30	Adenocarcinoma	0.858	0.123	0.777	0.825	Actinobacteriota	0.042	-0.035	0.0392 < Actinobacteriota <= 0.0670
30	Adenocarcinoma	0.858	0.123	0.777	0.825	Campilobacterota	0.003	0.019	0.00197 < Campilobacterota <= 0.00371
30	Adenocarcinoma	0.858	0.123	0.777	0.825	Synergistota	0.000	0.018	4.86e-05 < Synergistota <= 4.37e-04
30	Adenocarcinoma	0.858	0.123	0.777	0.825	Firmicutes	0.349	0.010	0.329 < Firmicutes <= 0.389
31	Adenocarcinoma	0.784	0.173	0.818	0.723	Campilobacterota	0.008	-0.046	0.00743 < Campilobacterota
31	Adenocarcinoma	0.784	0.173	0.818	0.723	Bacteroidota	0.187	-0.035	Bacteroidota <= 0.207
31	Adenocarcinoma	0.784	0.173	0.818	0.723	Patescibacteria	0.010	-0.018	Patescibacteria <= 0.0263
31	Adenocarcinoma	0.784	0.173	0.818	0.723	Actinobacteriota	0.027	0.018	0.0240 < Actinobacteriota <= 0.0392
31	Adenocarcinoma	0.784	0.173	0.818	0.723	Proteobacteria	0.198	-0.014	0.1579 < Proteobacteria
35	Adenocarcinoma	0.726	0.207	0.800	0.795	Synergistota	0.000	-0.042	Synergistota <= 4.86e-05
35	Adenocarcinoma	0.726	0.207	0.800	0.795	Actinobacteriota	0.045	-0.033	0.0392 < Actinobacteriota <= 0.0670
35	Adenocarcinoma	0.726	0.207	0.800	0.795	Bacteroidota	0.381	0.029	0.347 < Bacteroidota
35	Adenocarcinoma	0.726	0.207	0.800	0.795	Spirochaetota	0.000	0.023	Spirochaetota <= 4.14e-05
35	Adenocarcinoma	0.726	0.207	0.800	0.795	Campilobacterota	0.003	0.019	0.00197 < Campilobacterota <= 0.00371
36	Adenocarcinoma	0.710	0.161	0.806	0.742	Campilobacterota	0.008	-0.042	0.00743 < Campilobacterota
36	Adenocarcinoma	0.710	0.161	0.806	0.742	Actinobacteriota	0.133	-0.039	0.0670 < Actinobacteriota
36	Adenocarcinoma	0.710	0.161	0.806	0.742	Synergistota	0.000	0.021	4.86e-05 < Synergistota <= 4.37e-04
36	Adenocarcinoma	0.710	0.161	0.806	0.742	Spirochaetota	0.000	-0.015	4.14e-05 < Spirochaetota <= 8.31e-04
36	Adenocarcinoma	0.710	0.161	0.806	0.742	Fusobacteriota	0.071	0.011	0.0502 < Fusobacteriota <= 0.0823
37	Adenocarcinoma	0.824	0.247	0.763	0.867	Firmicutes	0.241	0.051	Firmicutes <= 0.329
37	Adenocarcinoma	0.824	0.247	0.763	0.867	Synergistota	0.001	0.038	4.37e-04 < Synergistota
37	Adenocarcinoma	0.824	0.247	0.763	0.867	Bacteroidota	0.171	-0.031	Bacteroidota <= 0.207
37	Adenocarcinoma	0.824	0.247	0.763	0.867	Campilobacterota	0.004	0.026	0.00371 < Campilobacterota <= 0.00743

37	Adenocarcinoma	0.824	0.247	0.763	0.867	Actinobacteriota	0.025	0.021	0.0240 < Actinobacteriota <= 0.0392
40	Adenocarcinoma	0.578	0.312	0.841	0.675	Firmicutes	0.510	-0.064	0.453 < Firmicutes
40	Adenocarcinoma	0.578	0.312	0.841	0.675	Synergistota	0.000	-0.040	Synergistota <= 4.86e-05
40	Adenocarcinoma	0.578	0.312	0.841	0.675	Actinobacteriota	0.096	-0.035	0.0670 < Actinobacteriota
40	Adenocarcinoma	0.578	0.312	0.841	0.675	Patescibacteria	0.002	-0.014	Patescibacteria <= 0.0263
40	Adenocarcinoma	0.578	0.312	0.841	0.675	Spirochaetota	0.000	-0.013	4.14e-05 < Spirochaetota <= 8.31e-04
42	Adenocarcinoma	0.692	0.203	0.812	0.752	Synergistota	0.000	-0.041	Synergistota <= 4.86e-05
42	Adenocarcinoma	0.692	0.203	0.812	0.752	Actinobacteriota	0.047	-0.036	0.0392 < Actinobacteriota <= 0.0670
42	Adenocarcinoma	0.692	0.203	0.812	0.752	Bacteroidota	0.067	-0.028	Bacteroidota <= 0.207
42	Adenocarcinoma	0.692	0.203	0.812	0.752	Spirochaetota	0.000	0.022	Spirochaetota <= 4.14e-05
42	Adenocarcinoma	0.692	0.203	0.812	0.752	Campilobacterota	0.003	0.023	0.00197 < Campilobacterota <= 0.00371
43	Adenocarcinoma	0.734	0.280253481	0.839489881	0.680346299	Firmicutes	0.635896693	-0.058007754	0.453 < Firmicutes
43	Adenocarcinoma	0.734	0.280253481	0.839489881	0.680346299	Synergistota	0	-0.03969187	Synergistota <= 4.86e-05
43	Adenocarcinoma	0.734	0.280253481	0.839489881	0.680346299	Bacteroidota	0.102701208	-0.029412039	Bacteroidota <= 0.207
43	Adenocarcinoma	0.734	0.280253481	0.839489881	0.680346299	Fusobacteriota	0.007875779	-0.01637176	Fusobacteriota <= 0.0502
43	Adenocarcinoma	0.734	0.280253481	0.839489881	0.680346299	Patescibacteria	0.007809928	-0.015660159	Patescibacteria <= 0.0263
44	Adenocarcinoma	0.718	0.328265134	0.835665924	0.697876777	Firmicutes	0.454773196	-0.060170507	0.453 < Firmicutes
44	Adenocarcinoma	0.718	0.328265134	0.835665924	0.697876777	Synergistota	3.15114E-05	-0.038547389	Synergistota <= 4.86e-05
44	Adenocarcinoma	0.718	0.328265134	0.835665924	0.697876777	Actinobacteriota	0.102207377	-0.033554383	0.0670 < Actinobacteriota
44	Adenocarcinoma	0.718	0.328265134	0.835665924	0.697876777	Bacteroidota	0.202524067	-0.031119637	Bacteroidota <= 0.207
44	Adenocarcinoma	0.718	0.328265134	0.835665924	0.697876777	Campilobacterota	0.005167877	0.025602768	0.00371 < Campilobacterota <= 0.00743
45	Adenocarcinoma	0.654	0.302240304	0.824973523	0.73264672	Firmicutes	0.546379045	-0.061660871	0.453 < Firmicutes
45	Adenocarcinoma	0.654	0.302240304	0.824973523	0.73264672	Synergistota	0	-0.041484544	Synergistota <= 4.86e-05
45	Adenocarcinoma	0.654	0.302240304	0.824973523	0.73264672	Bacteroidota	0.186484702	-0.030424022	Bacteroidota <= 0.207
45	Adenocarcinoma	0.654	0.302240304	0.824973523	0.73264672	Actinobacteriota	0.036891922	0.021223806	0.0240 < Actinobacteriota <= 0.0392
45	Adenocarcinoma	0.654	0.302240304	0.824973523	0.73264672	Proteobacteria	0.152058111	0.020018828	0.0993 < Proteobacteria <= 0.1579
48	Adenocarcinoma	0.568	0.35653072	0.837078653	0.68702766	Firmicutes	0.634752495	-0.064385857	0.453 < Firmicutes
48	Adenocarcinoma	0.568	0.35653072	0.837078653	0.68702766	Synergistota	0	-0.040870034	Synergistota <= 4.86e-05
48	Adenocarcinoma	0.568	0.35653072	0.837078653	0.68702766	Actinobacteriota	0.108535343	-0.035790337	0.0670 < Actinobacteriota
48	Adenocarcinoma	0.568	0.35653072	0.837078653	0.68702766	Bacteroidota	0.169036464	-0.031151622	Bacteroidota <= 0.207
48	Adenocarcinoma	0.568	0.35653072	0.837078653	0.68702766	Spirochaetota	0	0.022146856	Spirochaetota <= 4.14e-05
50	Adenocarcinoma	0.7	0.212057894	0.808375782	0.775589049	Synergistota	0	-0.040550363	Synergistota <= 4.86e-05
50	Adenocarcinoma	0.7	0.212057894	0.808375782	0.775589049	Patescibacteria	0.084098063	0.038724956	0.0761 < Patescibacteria
50	Adenocarcinoma	0.7	0.212057894	0.808375782	0.775589049	Actinobacteriota	0.152142683	-0.035357965	0.0670 < Actinobacteriota
50	Adenocarcinoma	0.7	0.212057894	0.808375782	0.775589049	Campilobacterota	0.00331274	0.018633162	0.00197 < Campilobacterota <= 0.00371
50	Adenocarcinoma	0.7	0.212057894	0.808375782	0.775589049	Spirochaetota	0.000596607	-0.014236523	4.14e-05 < Spirochaetota <= 8.31e-04

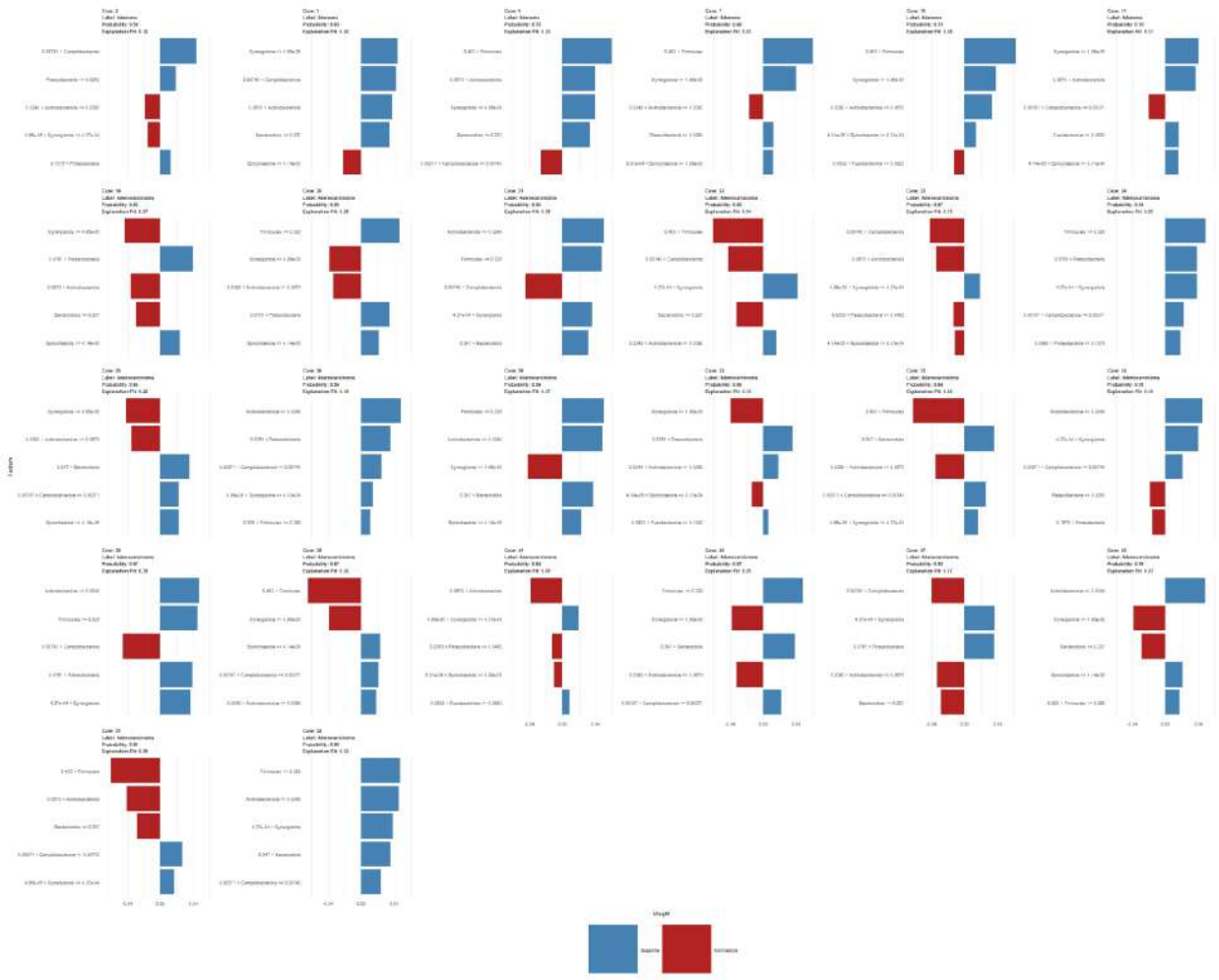


Figure ML3-A-6: The LIME Result in Training dataset for Phylum in Saliva

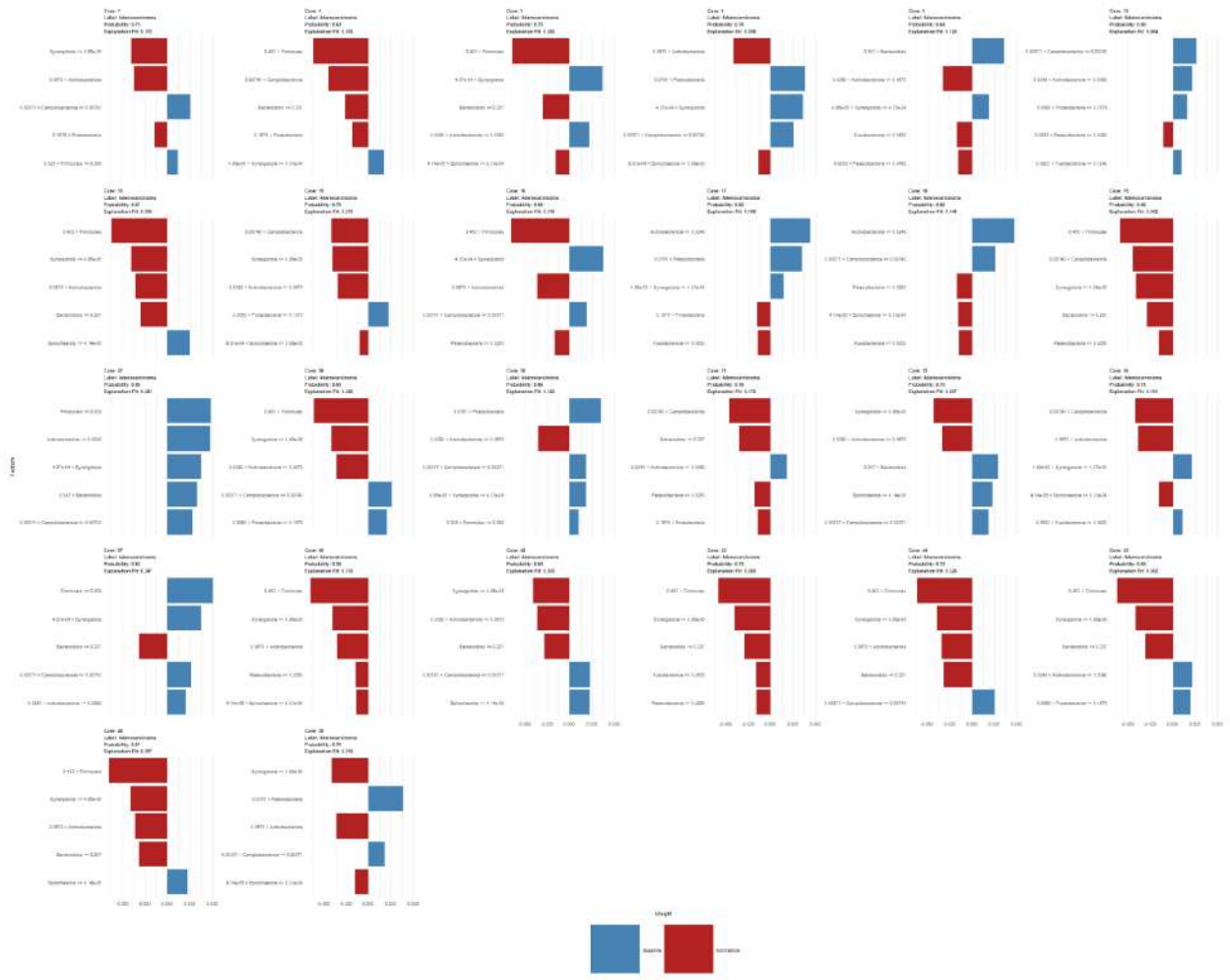


Figure ML3-A-7: The LIME Result in Test dataset for Phylum in Saliva

Table ML3-A-6: The Aggregated SHAP Values

Variable	Training Dataset				Testing Dataset			
	contribution	position	sign	cumulative	contribution	position	sign	cumulative
intercept	0.7605	14	X	0.760	0.7605	14	X	0.7605
Cyanobacteria	-0.0001	13	-1	0.760	0.0052	5	1	0.7503
Tumor	0.0000	12	1	0.760	0.0000	11	1	0.7665
Fusobacteriota	0.0005	11	1	0.761	-0.0040	9	-1	0.7455
Spirochaetota	0.0002	10	1	0.761	0.0050	7	1	0.7496
Synergistota	0.0000	9	-1	0.761	-0.0009	8	-1	0.7446
Firmicutes	0.0003	8	1	0.761	-0.0170	10	-1	0.7495
Proteobacteria	-0.0002	7	-1	0.761	-0.0044	6	-1	0.7452
Patescibacteria	-0.0005	6	-1	0.761	-0.0035	4	-1	0.7468
Gender	0.0000	5	1	0.761	0.0000	12	1	0.7665
Campilobacterota	-0.0002	4	-1	0.760	0.0061	13	1	0.7665
Actinobacteriota	-0.0004	3	-1	0.760	-0.0016	3	-1	0.7453
Bacteroidota	0.0004	2	1	0.760	-0.0170	2	-1	0.7282
prediction	0.7605	1	X	0.760	0.7282	1	X	0.7282

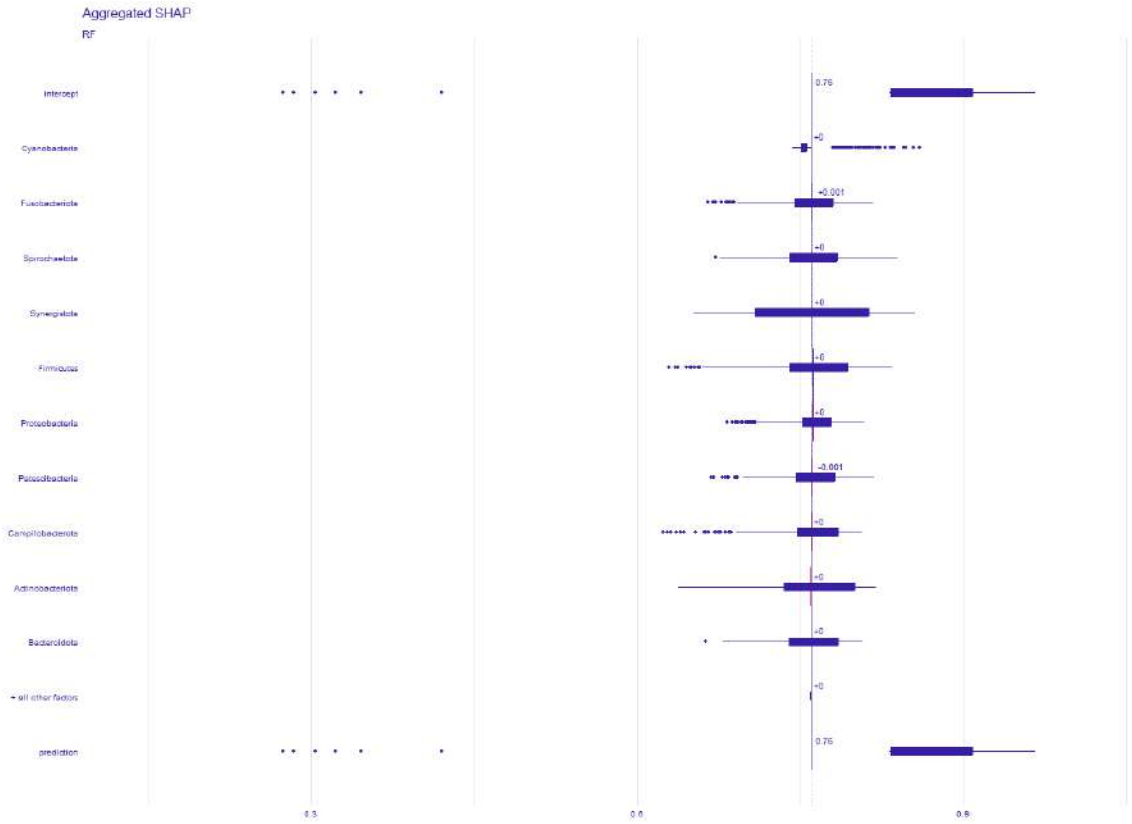


Figure ML3-A-8: The Aggregated SHAP Value in Training dataset for Phylum in Saliva

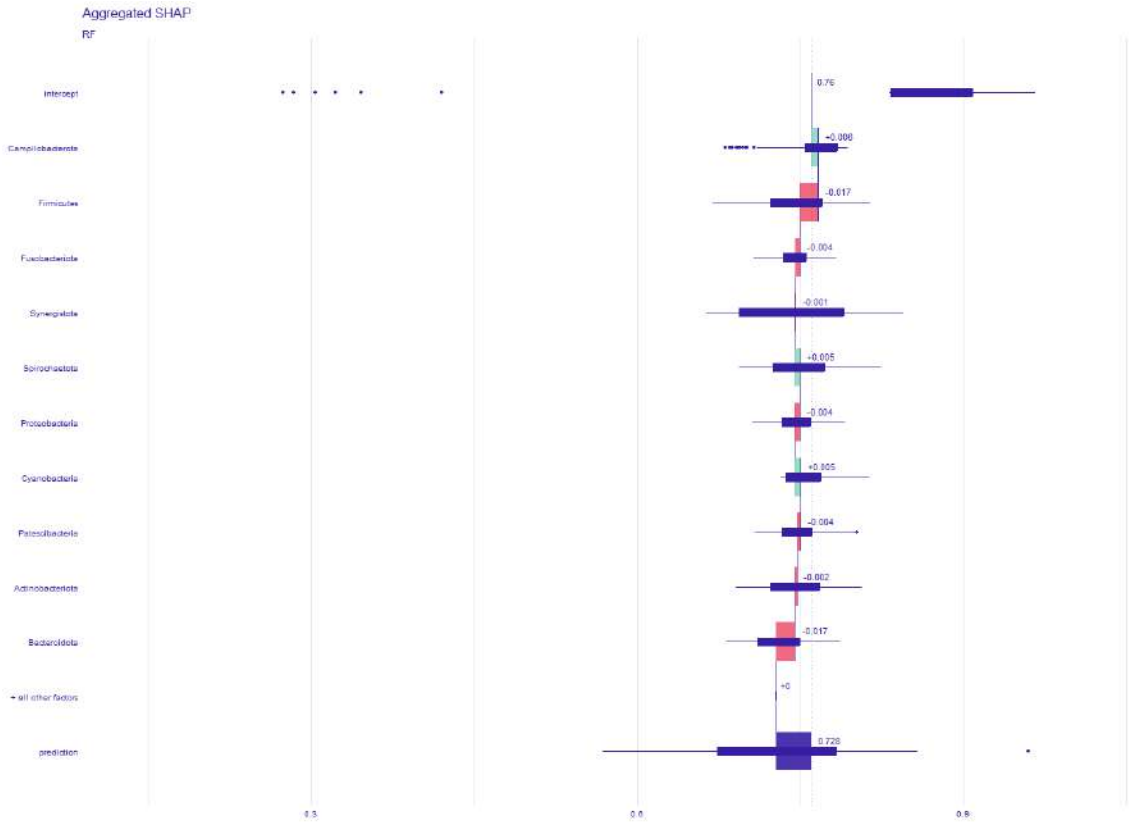


Figure ML3-A-9: The Aggregated SHAP Value in Test dataset for Phylum in Saliva

B- Class

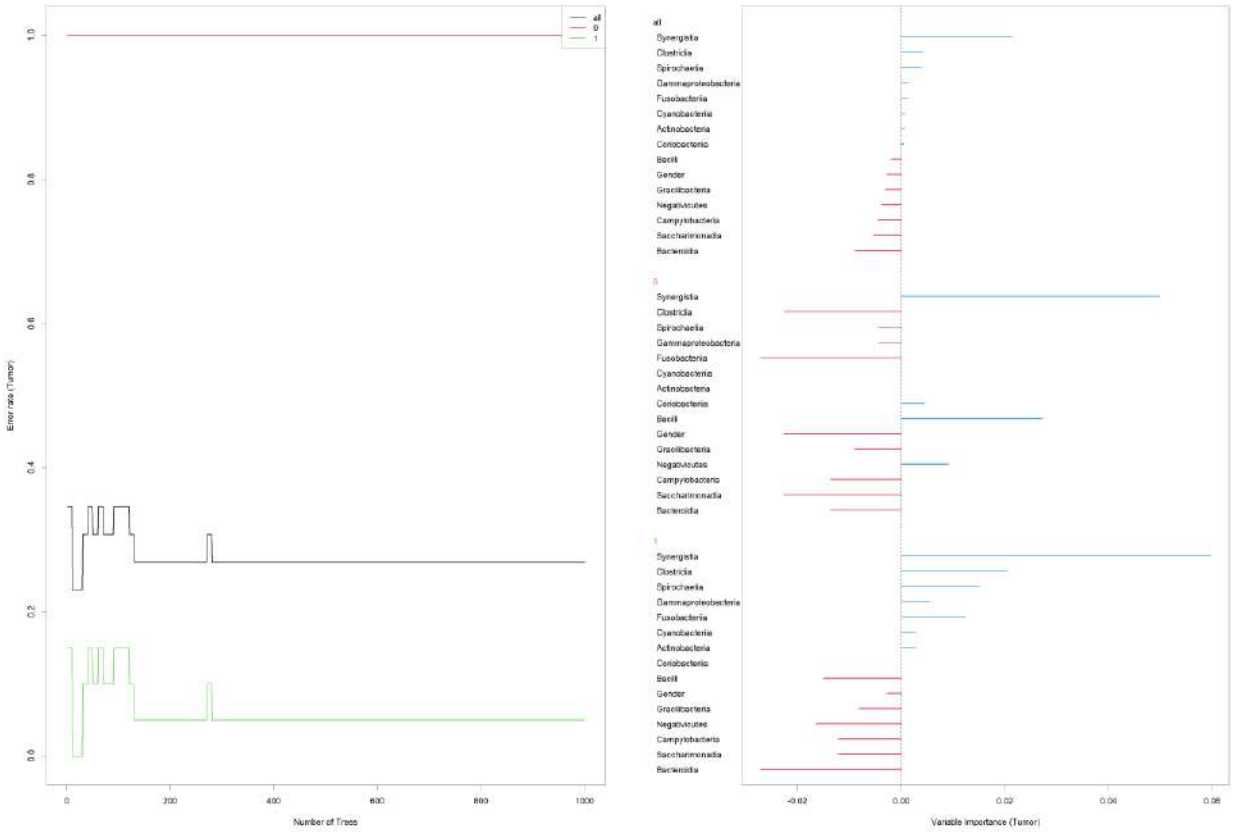


Figure ML3-B-1: The number of trees and Variable Importance Plot for Class in Saliva

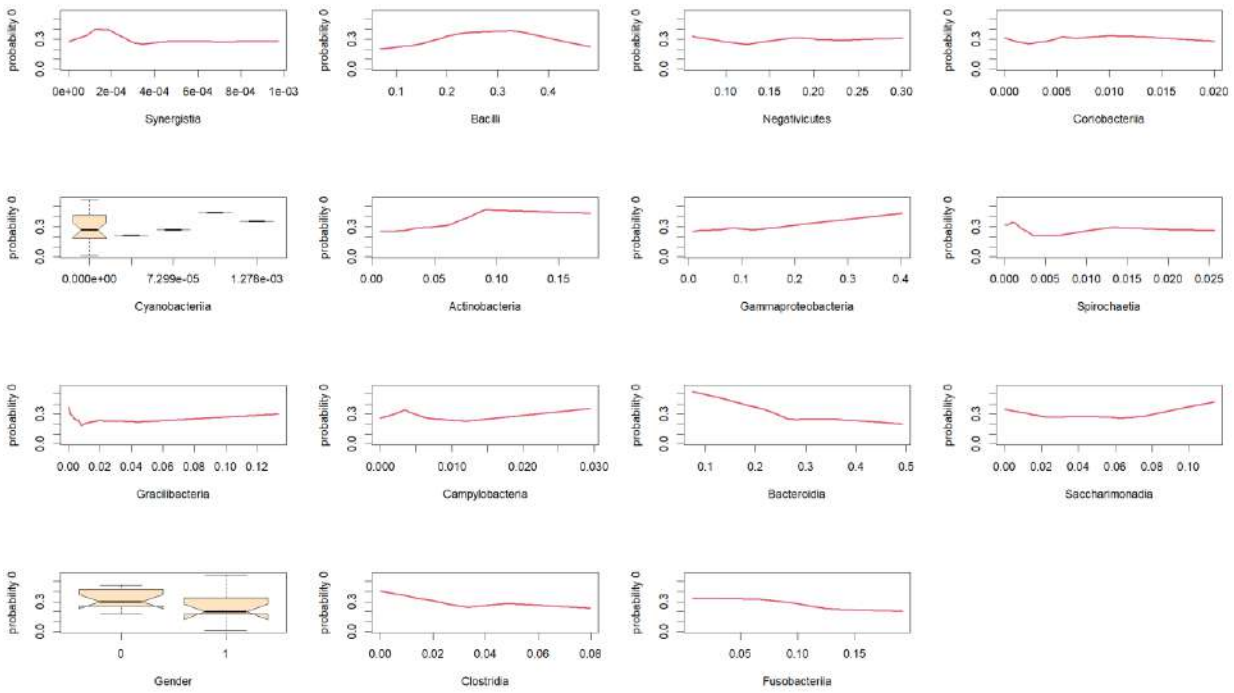


Figure ML3-B-2: Marginal Plot for Class in Saliva

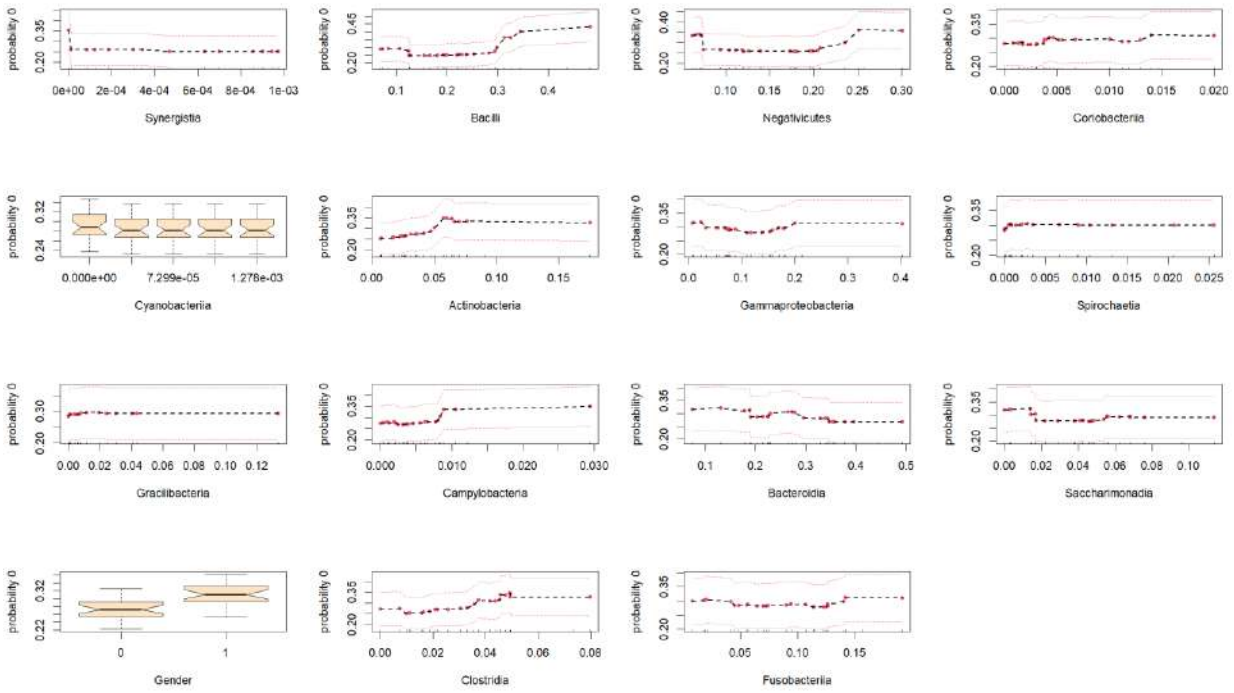


Figure ML3-B-3: Partial Plot for Class in Saliva

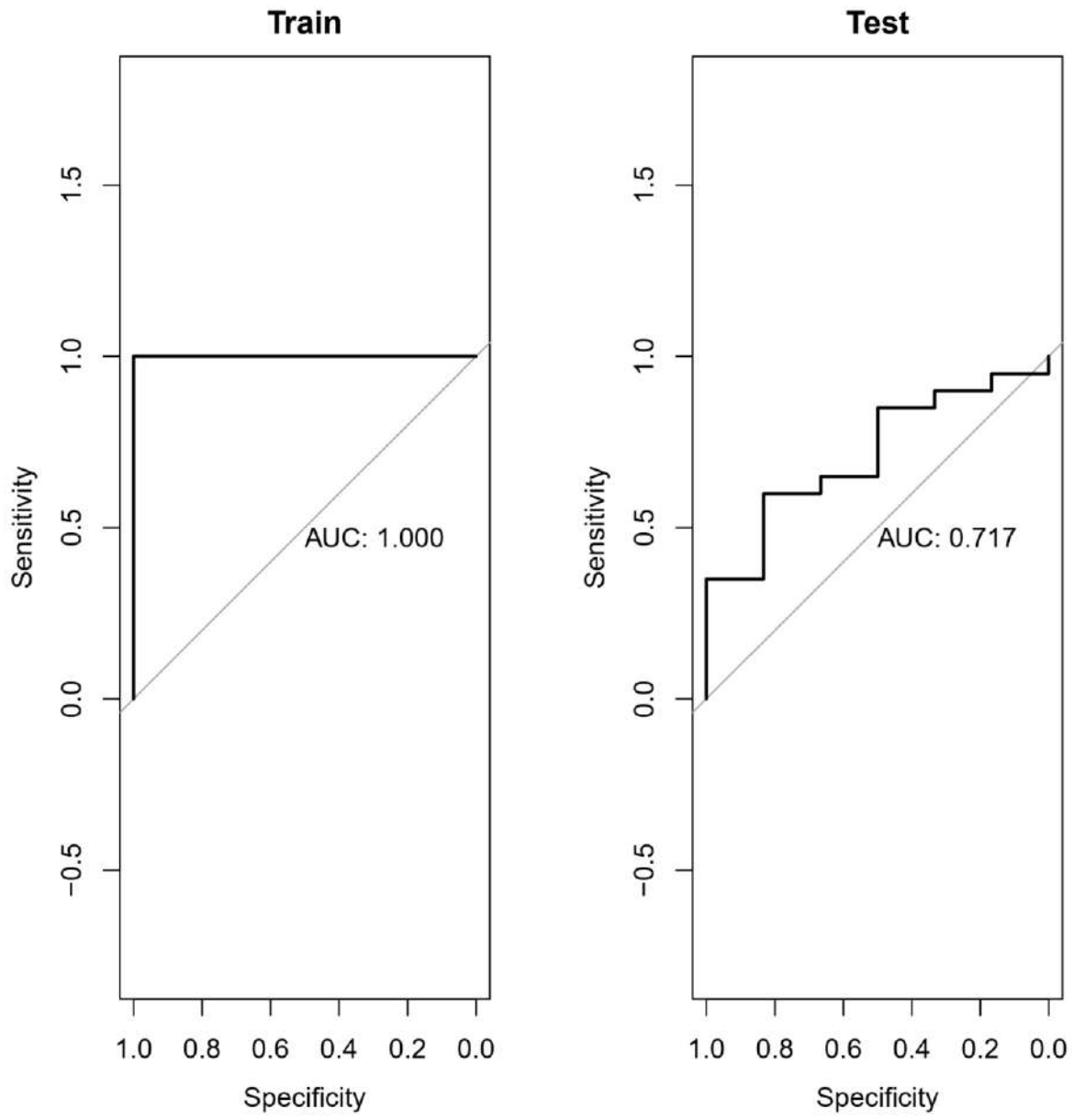


Figure ML3-B-4: The ROC Curve for Class in Saliva

Table ML3-B-1: The Model Performance

Index	Values		
	2.50%	**50.00%**	**97.50%**
Accuracy	0.42	0.73	0.92
Specificity	0.50	0.83	1.00
Sensitivity	0.25	0.65	0.95

Table ML3-B-2: The Variable Importance Values

Covariates	Varibale Importance			Brier Score		
	all	**AP**	**CRC**	**all**	**AP**	**CRC**
Gender	-0.0074	-0.0815	-0.0014	-0.0157	-0.0107	-0.0107
Actinobacteria	-0.0039	-0.0045	-0.0122	0.0025	0.0017	0.0017
Bacilli	0.0848	-0.3443	0.3996	0.1345	0.0914	0.0914
Bacteroidia	-0.0222	-0.0317	-0.0680	-0.0469	-0.0319	-0.0319
Campylobacteria	0.0046	-0.1676	0.0666	-0.0130	-0.0088	-0.0088
Clostridia	-0.0007	-0.1586	0.0449	-0.0274	-0.0186	-0.0186
Coriobacteriia	0.0012	-0.0544	0.0204	-0.0107	-0.0073	-0.0073
Cyanobacteriia	-0.0054	0.0091	-0.0217	-0.0085	-0.0058	-0.0058
Fusobacteriia	0.0085	-0.0997	0.0598	-0.0034	-0.0023	-0.0023
Gammaproteobacteria	0.0156	-0.1133	0.0883	0.0192	0.0131	0.0131
Gracilibacteria	-0.0038	0.0000	-0.0136	-0.0187	-0.0127	-0.0127
Negativicutes	0.0541	-0.3171	0.2841	0.0731	0.0497	0.0497
Saccharimonadia	0.0039	-0.1178	0.0489	-0.0362	-0.0246	-0.0246
Spirochaetia	0.0036	0.0181	0.0068	0.0016	0.0011	0.0011
Synergistia	0.0392	0.0725	0.1155	0.1558	0.1059	0.1059

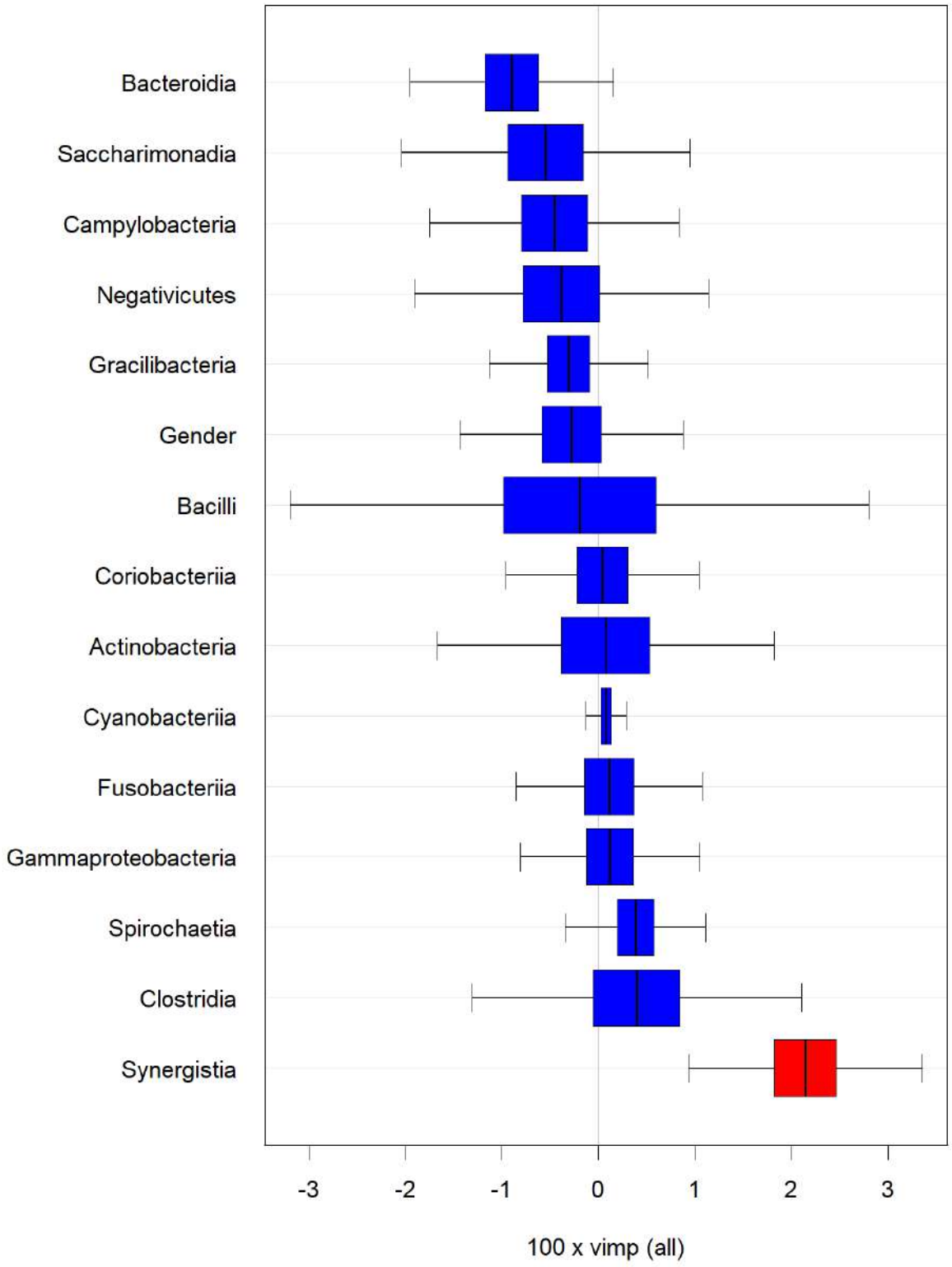


Figure ML3-B-5: The Variable Importance Plot (VIMP) for Class in Saliva

Best model is selected for LIME results

Table ML3-B-3: The best setting for model and LIME Results

Setting	mtry	ROC	Sens	Spec	ROCSD	SensSD	SpecSD
ntrees: 500 nodesize: 1	10	0.375	0	0.9	0.176777	0	0.223607
ntrees: 1000 nodesize: 1	1	0.375	0	1	0.25	0	0
ntrees: 1000 nodesize: 1	7	0.375	0	0.9	0.176777	0	0.223607
ntrees: 1000 nodesize: 1	10	0.375	0	0.9	0.176777	0	0.223607
ntrees: 500 nodesize: 5	1	0.375	0	1	0.176777	0	0
ntrees: 1000 nodesize: 5	1	0.325	0	1	0.30104	0	0

Table ML3-B-4: The LIME Result for Training Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
2	Adenoma	0.678	0.222	0.260	0.389	Campylobacteria	0.009	0.101	0.00743 < Campylobacteria
2	Adenoma	0.678	0.222	0.260	0.389	Bacilli	0.223	-0.074	0.210 < Bacilli <= 0.291
2	Adenoma	0.678	0.222	0.260	0.389	Fusobacteriia	0.141	0.039	0.1242 < Fusobacteriia
2	Adenoma	0.678	0.222	0.260	0.389	Saccharimonadia	0.014	0.039	Saccharimonadia <= 0.0183
2	Adenoma	0.678	0.222	0.260	0.389	Gammaproteobacteria	0.170	0.024	0.1579 < Gammaproteobacteria
3	Adenoma	0.734	0.308	0.207	0.479	Campylobacteria	0.012	0.106	0.00743 < Campylobacteria
3	Adenoma	0.734	0.308	0.207	0.479	Negativicutes	0.251	0.075	0.198 < Negativicutes
3	Adenoma	0.734	0.308	0.207	0.479	Actinobacteria	0.064	0.066	0.0589 < Actinobacteria
3	Adenoma	0.734	0.308	0.207	0.479	Synergistia	0.000	0.055	Synergistia <= 4.86e-05
3	Adenoma	0.734	0.308	0.207	0.479	Bacilli	0.107	-0.030	Bacilli <= 0.135
6	Adenoma	0.732	0.469	0.226	0.440	Bacilli	0.482	0.175	0.291 < Bacilli
6	Adenoma	0.732	0.469	0.226	0.440	Actinobacteria	0.076	0.064	0.0589 < Actinobacteria
6	Adenoma	0.732	0.469	0.226	0.440	Synergistia	0.000	0.052	Synergistia <= 4.86e-05
6	Adenoma	0.732	0.469	0.226	0.440	Negativicutes	0.119	-0.042	0.096 < Negativicutes <= 0.139
6	Adenoma	0.732	0.469	0.226	0.440	Campylobacteria	0.006	-0.036	0.00371 < Campylobacteria <= 0.00743
7	Adenoma	0.664	0.478	0.209	0.484	Bacilli	0.313	0.176	0.291 < Bacilli
7	Adenoma	0.664	0.478	0.209	0.484	Synergistia	0.000	0.051	Synergistia <= 4.86e-05
7	Adenoma	0.664	0.478	0.209	0.484	Clostridia	0.049	0.050	0.0455 < Clostridia
7	Adenoma	0.664	0.478	0.209	0.484	Saccharimonadia	0.017	0.037	Saccharimonadia <= 0.0183
7	Adenoma	0.664	0.478	0.209	0.484	Negativicutes	0.173	-0.038	0.139 < Negativicutes <= 0.198
10	Adenoma	0.790	0.438	0.221	0.445	Bacilli	0.345	0.177	0.291 < Bacilli
10	Adenoma	0.790	0.438	0.221	0.445	Synergistia	0.000	0.053	Synergistia <= 4.86e-05
10	Adenoma	0.790	0.438	0.221	0.445	Clostridia	0.049	0.044	0.0455 < Clostridia
10	Adenoma	0.790	0.438	0.221	0.445	Campylobacteria	0.002	-0.030	Campylobacteria <= 0.00197
10	Adenoma	0.790	0.438	0.221	0.445	Fusobacteriia	0.057	-0.019	0.0502 < Fusobacteriia <= 0.0823
11	Adenoma	0.706	0.192	0.227	0.428	Negativicutes	0.256	0.073	0.198 < Negativicutes
11	Adenoma	0.706	0.192	0.227	0.428	Synergistia	0.000	0.058	Synergistia <= 4.86e-05
11	Adenoma	0.706	0.192	0.227	0.428	Actinobacteria	0.059	0.057	0.0589 < Actinobacteria
11	Adenoma	0.706	0.192	0.227	0.428	Clostridia	0.046	0.050	0.0455 < Clostridia

11	Adenoma	0.706	0.192	0.227	0.428	Campylobacteria	0.002	-0.037	0.00197 < Campylobacteria <= 0.00371
14	Adenocarcinoma	0.838	0.202	0.715	0.759	Bacilli	0.200	0.075	0.135 < Bacilli <= 0.210
14	Adenocarcinoma	0.838	0.202	0.715	0.759	Actinobacteria	0.066	-0.066	0.0589 < Actinobacteria
14	Adenocarcinoma	0.838	0.202	0.715	0.759	Synergistia	0.000	-0.054	Synergistia <= 4.86e-05
14	Adenocarcinoma	0.838	0.202	0.715	0.759	Negativicutes	0.178	0.050	0.139 < Negativicutes <= 0.198
14	Adenocarcinoma	0.838	0.202	0.715	0.759	Campylobacteria	0.003	0.040	0.00197 < Campylobacteria <= 0.00371
20	Adenocarcinoma	0.910	0.123	0.690	0.797	Bacilli	0.238	0.076	0.210 < Bacilli <= 0.291
20	Adenocarcinoma	0.910	0.123	0.690	0.797	Synergistia	0.000	-0.050	Synergistia <= 4.86e-05
20	Adenocarcinoma	0.910	0.123	0.690	0.797	Clostridia	0.016	0.032	0.0160 < Clostridia <= 0.0317
20	Adenocarcinoma	0.910	0.123	0.690	0.797	Campylobacteria	0.000	0.029	Campylobacteria <= 0.00197
20	Adenocarcinoma	0.910	0.123	0.690	0.797	Gammaproteobacteria	0.155	0.019	0.0992 < Gammaproteobacteria <= 0.1579
21	Adenocarcinoma	0.912	0.281	0.768	0.533	Campylobacteria	0.008	-0.105	0.00743 < Campylobacteria
21	Adenocarcinoma	0.912	0.281	0.768	0.533	Negativicutes	0.205	-0.080	0.198 < Negativicutes
21	Adenocarcinoma	0.912	0.281	0.768	0.533	Clostridia	0.047	-0.051	0.0455 < Clostridia
21	Adenocarcinoma	0.912	0.281	0.768	0.533	Synergistia	0.001	0.044	4.37e-04 < Synergistia
21	Adenocarcinoma	0.912	0.281	0.768	0.533	Fusobacteriia	0.191	-0.042	0.1242 < Fusobacteriia
22	Adenocarcinoma	0.806	0.524	0.748	0.590	Bacilli	0.326	-0.170	0.291 < Bacilli
22	Adenocarcinoma	0.806	0.524	0.748	0.590	Campylobacteria	0.008	-0.103	0.00743 < Campylobacteria
22	Adenocarcinoma	0.806	0.524	0.748	0.590	Synergistia	0.001	0.046	4.37e-04 < Synergistia
22	Adenocarcinoma	0.806	0.524	0.748	0.590	Negativicutes	0.139	0.043	0.096 < Negativicutes <= 0.139
22	Adenocarcinoma	0.806	0.524	0.748	0.590	Actinobacteria	0.035	0.026	0.0227 < Actinobacteria <= 0.0351
23	Adenocarcinoma	0.828	0.249	0.725	0.653	Campylobacteria	0.010	-0.098	0.00743 < Campylobacteria
23	Adenocarcinoma	0.828	0.249	0.725	0.653	Bacilli	0.219	0.069	0.210 < Bacilli <= 0.291
23	Adenocarcinoma	0.828	0.249	0.725	0.653	Actinobacteria	0.070	-0.061	0.0589 < Actinobacteria
23	Adenocarcinoma	0.828	0.249	0.725	0.653	Negativicutes	0.180	0.049	0.139 < Negativicutes <= 0.198
23	Adenocarcinoma	0.828	0.249	0.725	0.653	Fusobacteriia	0.126	-0.030	0.1242 < Fusobacteriia
24	Adenocarcinoma	0.904	0.085	0.677	0.798	Negativicutes	0.126	0.045	0.096 < Negativicutes <= 0.139
24	Adenocarcinoma	0.904	0.085	0.677	0.798	Synergistia	0.000	0.039	4.37e-04 < Synergistia
24	Adenocarcinoma	0.904	0.085	0.677	0.798	Campylobacteria	0.003	0.041	0.00197 < Campylobacteria <= 0.00371
24	Adenocarcinoma	0.904	0.085	0.677	0.798	Fusobacteriia	0.156	-0.037	0.1242 < Fusobacteriia
24	Adenocarcinoma	0.904	0.085	0.677	0.798	Bacilli	0.126	0.034	Bacilli <= 0.135
25	Adenocarcinoma	0.892	0.149	0.706	0.773	Bacilli	0.192	0.076	0.135 < Bacilli <= 0.210
25	Adenocarcinoma	0.892	0.149	0.706	0.773	Synergistia	0.000	-0.048	Synergistia <= 4.86e-05
25	Adenocarcinoma	0.892	0.149	0.706	0.773	Saccharimonadia	0.014	-0.038	Saccharimonadia <= 0.0183
25	Adenocarcinoma	0.892	0.149	0.706	0.773	Negativicutes	0.181	0.041	0.139 < Negativicutes <= 0.198
25	Adenocarcinoma	0.892	0.149	0.706	0.773	Campylobacteria	0.003	0.036	0.00197 < Campylobacteria <= 0.00371
26	Adenocarcinoma	0.862	0.127	0.701	0.724	Negativicutes	0.235	-0.065	0.198 < Negativicutes
26	Adenocarcinoma	0.862	0.127	0.701	0.724	Campylobacteria	0.006	0.043	0.00371 < Campylobacteria <= 0.00743
26	Adenocarcinoma	0.862	0.127	0.701	0.724	Actinobacteria	0.018	0.040	Actinobacteria <= 0.0227
26	Adenocarcinoma	0.862	0.127	0.701	0.724	Bacilli	0.073	0.039	Bacilli <= 0.135
26	Adenocarcinoma	0.862	0.127	0.701	0.724	Fusobacteriia	0.125	-0.035	0.1242 < Fusobacteriia
29	Adenocarcinoma	0.940	0.103	0.709	0.746	Synergistia	0.000	-0.046	Synergistia <= 4.86e-05

29	Adenocarcinoma	0.940	0.103	0.709	0.746	Negativicutes	0.120	0.042	0.096 < Negativicutes <= 0.139
29	Adenocarcinoma	0.940	0.103	0.709	0.746	Actinobacteria	0.007	0.041	Actinobacteria <= 0.0227
29	Adenocarcinoma	0.940	0.103	0.709	0.746	Saccharimonadia	0.003	-0.037	Saccharimonadia <= 0.0183
29	Adenocarcinoma	0.940	0.103	0.709	0.746	Bacilli	0.128	0.038	Bacilli <= 0.135
32	Adenocarcinoma	0.930	0.146	0.687	0.808	Bacilli	0.282	0.073	0.210 < Bacilli <= 0.291
32	Adenocarcinoma	0.930	0.146	0.687	0.808	Synergistia	0.000	-0.055	Synergistia <= 4.86e-05
32	Adenocarcinoma	0.930	0.146	0.687	0.808	Negativicutes	0.140	0.048	0.139 < Negativicutes <= 0.198
32	Adenocarcinoma	0.930	0.146	0.687	0.808	Clostridia	0.031	0.033	0.0160 < Clostridia <= 0.0317
32	Adenocarcinoma	0.930	0.146	0.687	0.808	Campylobacteria	0.000	0.022	Campylobacteria <= 0.00197
33	Adenocarcinoma	0.882	0.170	0.705	0.722	Negativicutes	0.207	-0.074	0.198 < Negativicutes
33	Adenocarcinoma	0.882	0.170	0.705	0.722	Bacilli	0.166	0.069	0.135 < Bacilli <= 0.210
33	Adenocarcinoma	0.882	0.170	0.705	0.722	Clostridia	0.080	-0.050	0.0455 < Clostridia
33	Adenocarcinoma	0.882	0.170	0.705	0.722	Campylobacteria	0.004	0.040	0.00371 < Campylobacteria <= 0.00743
33	Adenocarcinoma	0.882	0.170	0.705	0.722	Synergistia	0.000	0.030	4.86e-05 < Synergistia <= 4.37e-04
34	Adenocarcinoma	0.862	0.132	0.669	0.827	Bacilli	0.229	0.067	0.210 < Bacilli <= 0.291
34	Adenocarcinoma	0.862	0.132	0.669	0.827	Campylobacteria	0.005	0.046	0.00371 < Campylobacteria <= 0.00743
34	Adenocarcinoma	0.862	0.132	0.669	0.827	Synergistia	0.001	0.043	4.37e-04 < Synergistia
34	Adenocarcinoma	0.862	0.132	0.669	0.827	Actinobacteria	0.007	0.038	Actinobacteria <= 0.0227
34	Adenocarcinoma	0.862	0.132	0.669	0.827	Saccharimonadia	0.017	-0.037	Saccharimonadia <= 0.0183
38	Adenocarcinoma	0.922	0.229	0.714	0.684	Campylobacteria	0.008	-0.109	0.00743 < Campylobacteria
38	Adenocarcinoma	0.922	0.229	0.714	0.684	Clostridia	0.073	-0.045	0.0455 < Clostridia
38	Adenocarcinoma	0.922	0.229	0.714	0.684	Synergistia	0.001	0.045	4.37e-04 < Synergistia
38	Adenocarcinoma	0.922	0.229	0.714	0.684	Actinobacteria	0.016	0.041	Actinobacteria <= 0.0227
38	Adenocarcinoma	0.922	0.229	0.714	0.684	Bacilli	0.085	0.038	Bacilli <= 0.135
39	Adenocarcinoma	0.822	0.509	0.795	0.506	Bacilli	0.294	-0.173	0.291 < Bacilli
39	Adenocarcinoma	0.822	0.509	0.795	0.506	Negativicutes	0.300	-0.072	0.198 < Negativicutes
39	Adenocarcinoma	0.822	0.509	0.795	0.506	Synergistia	0.000	-0.052	Synergistia <= 4.86e-05
39	Adenocarcinoma	0.822	0.509	0.795	0.506	Campylobacteria	0.004	0.044	0.00197 < Campylobacteria <= 0.00371
39	Adenocarcinoma	0.822	0.509	0.795	0.506	Saccharimonadia	0.004	-0.036	Saccharimonadia <= 0.0183
41	Adenocarcinoma	0.862	0.153	0.696	0.730	Bacilli	0.253	0.077	0.210 < Bacilli <= 0.291
41	Adenocarcinoma	0.862	0.153	0.696	0.730	Actinobacteria	0.175	-0.061	0.0589 < Actinobacteria
41	Adenocarcinoma	0.862	0.153	0.696	0.730	Clostridia	0.049	-0.043	0.0455 < Clostridia
41	Adenocarcinoma	0.862	0.153	0.696	0.730	Campylobacteria	0.001	0.032	Campylobacteria <= 0.00197
41	Adenocarcinoma	0.862	0.153	0.696	0.730	Synergistia	0.000	0.028	4.86e-05 < Synergistia <= 4.37e-04
46	Adenocarcinoma	0.948	0.142	0.688	0.826	Bacilli	0.181	0.075	0.135 < Bacilli <= 0.210
46	Adenocarcinoma	0.948	0.142	0.688	0.826	Synergistia	0.000	-0.050	Synergistia <= 4.86e-05
46	Adenocarcinoma	0.948	0.142	0.688	0.826	Campylobacteria	0.002	0.042	0.00197 < Campylobacteria <= 0.00371
46	Adenocarcinoma	0.948	0.142	0.688	0.826	Negativicutes	0.111	0.042	0.096 < Negativicutes <= 0.139
46	Adenocarcinoma	0.948	0.142	0.688	0.826	Clostridia	0.026	0.029	0.0160 < Clostridia <= 0.0317
47	Adenocarcinoma	0.866	0.241	0.704	0.729	Campylobacteria	0.029	-0.104	0.00743 < Campylobacteria
47	Adenocarcinoma	0.866	0.241	0.704	0.729	Bacilli	0.193	0.070	0.135 < Bacilli <= 0.210
47	Adenocarcinoma	0.866	0.241	0.704	0.729	Negativicutes	0.195	0.046	0.139 < Negativicutes <= 0.198

47	Adenocarcinoma	0.866	0.241	0.704	0.729	Synergistia	0.001	0.040	4.37e-04 < Synergistia
47	Adenocarcinoma	0.866	0.241	0.704	0.729	Bacteroidia	0.204	-0.027	Bacteroidia <= 0.207
49	Adenocarcinoma	0.846	0.451	0.771	0.569	Bacilli	0.298	-0.175	0.291 < Bacilli
49	Adenocarcinoma	0.846	0.451	0.771	0.569	Synergistia	0.000	-0.057	Synergistia <= 4.86e-05
49	Adenocarcinoma	0.846	0.451	0.771	0.569	Saccharimonadia	0.000	-0.041	Saccharimonadia <= 0.0183
49	Adenocarcinoma	0.846	0.451	0.771	0.569	Actinobacteria	0.022	0.039	Actinobacteria <= 0.0227
49	Adenocarcinoma	0.846	0.451	0.771	0.569	Campylobacteria	0.000	0.031	Campylobacteria <= 0.00197
51	Adenocarcinoma	0.760	0.498	0.769	0.525	Bacilli	0.438	-0.176	0.291 < Bacilli
51	Adenocarcinoma	0.760	0.498	0.769	0.525	Negativicutes	0.199	-0.071	0.198 < Negativicutes
51	Adenocarcinoma	0.760	0.498	0.769	0.525	Actinobacteria	0.090	-0.065	0.0589 < Actinobacteria
51	Adenocarcinoma	0.760	0.498	0.769	0.525	Campylobacteria	0.004	0.041	0.00371 < Campylobacteria <= 0.00743
51	Adenocarcinoma	0.760	0.498	0.769	0.525	Synergistia	0.000	0.027	4.86e-05 < Synergistia <= 4.37e-04
52	Adenocarcinoma	0.986	0.133	0.664	0.826	Bacilli	0.155	0.075	0.135 < Bacilli <= 0.210
52	Adenocarcinoma	0.986	0.133	0.664	0.826	Synergistia	0.001	0.042	4.37e-04 < Synergistia
52	Adenocarcinoma	0.986	0.133	0.664	0.826	Actinobacteria	0.017	0.042	Actinobacteria <= 0.0227
52	Adenocarcinoma	0.986	0.133	0.664	0.826	Campylobacteria	0.007	0.037	0.00371 < Campylobacteria <= 0.00743
52	Adenocarcinoma	0.986	0.133	0.664	0.826	Fusobacteriia	0.126	-0.034	0.1242 < Fusobacteriia

Table ML-B-5: The LIME Result for Test Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
1	Adenocarcinoma	0.766	0.177	0.711	0.737	Bacilli	0.223	0.075	0.210 < Bacilli <= 0.291
1	Adenocarcinoma	0.766	0.177	0.711	0.737	Actinobacteria	0.073	-0.062	0.0589 < Actinobacteria
1	Adenocarcinoma	0.766	0.177	0.711	0.737	Synergistia	0.000	-0.053	Synergistia <= 4.86e-05
1	Adenocarcinoma	0.766	0.177	0.711	0.737	Campylobacteria	0.004	0.045	0.00371 < Campylobacteria <= 0.00743
1	Adenocarcinoma	0.766	0.177	0.711	0.737	Saccharimonadia	0.040	0.022	0.0183 < Saccharimonadia <= 0.0442
4	Adenocarcinoma	0.728	0.223	0.687	0.777	Campylobacteria	0.019	-0.096	0.00743 < Campylobacteria
4	Adenocarcinoma	0.728	0.223	0.687	0.777	Bacilli	0.260	0.074	0.210 < Bacilli <= 0.291
4	Adenocarcinoma	0.728	0.223	0.687	0.777	Negativicutes	0.191	0.046	0.139 < Negativicutes <= 0.198
4	Adenocarcinoma	0.728	0.223	0.687	0.777	Clostridia	0.024	0.036	0.0160 < Clostridia <= 0.0317
4	Adenocarcinoma	0.728	0.223	0.687	0.777	Actinobacteria	0.027	0.031	0.0227 < Actinobacteria <= 0.0351
5	Adenocarcinoma	0.576	0.065	0.687	0.753	Synergistia	0.001	0.043	4.37e-04 < Synergistia
5	Adenocarcinoma	0.576	0.065	0.687	0.753	Campylobacteria	0.001	0.036	Campylobacteria <= 0.00197
5	Adenocarcinoma	0.576	0.065	0.687	0.753	Actinobacteria	0.028	0.034	0.0227 < Actinobacteria <= 0.0351
5	Adenocarcinoma	0.576	0.065	0.687	0.753	Bacteroidia	0.100	-0.029	Bacteroidia <= 0.207
5	Adenocarcinoma	0.576	0.065	0.687	0.753	Gammaproteobacteria	0.190	-0.019	0.1579 < Gammaproteobacteria
8	Adenocarcinoma	0.862	0.157	0.676	0.808	Actinobacteria	0.092	-0.064	0.0589 < Actinobacteria
8	Adenocarcinoma	0.862	0.157	0.676	0.808	Bacilli	0.237	0.067	0.210 < Bacilli <= 0.291
8	Adenocarcinoma	0.862	0.157	0.676	0.808	Synergistia	0.001	0.045	4.37e-04 < Synergistia
8	Adenocarcinoma	0.862	0.157	0.676	0.808	Negativicutes	0.110	0.046	0.096 < Negativicutes <= 0.139

8	Adenocarcinoma	0.862	0.157	0.676	0.808	Campylobacteria	0.004	0.039	0.00371 < Campylobacteria <= 0.00743
9	Adenocarcinoma	0.796	0.127	0.678	0.805	Bacilli	0.238	0.075	0.210 < Bacilli <= 0.291
9	Adenocarcinoma	0.796	0.127	0.678	0.805	Clostridia	0.065	-0.050	0.0455 < Clostridia
9	Adenocarcinoma	0.796	0.127	0.678	0.805	Negativicutes	0.109	0.044	0.096 < Negativicutes <= 0.139
9	Adenocarcinoma	0.796	0.127	0.678	0.805	Synergistia	0.000	0.031	4.86e-05 < Synergistia <= 4.37e-04
9	Adenocarcinoma	0.796	0.127	0.678	0.805	Campylobacteria	0.001	0.027	Campylobacteria <= 0.00197
12	Adenocarcinoma	0.942	0.094	0.659	0.851	Bacilli	0.237	0.067	0.210 < Bacilli <= 0.291
12	Adenocarcinoma	0.942	0.094	0.659	0.851	Negativicutes	0.136	0.044	0.096 < Negativicutes <= 0.139
12	Adenocarcinoma	0.942	0.094	0.659	0.851	Campylobacteria	0.006	0.042	0.00371 < Campylobacteria <= 0.00743
12	Adenocarcinoma	0.942	0.094	0.659	0.851	Saccharimonadia	0.035	0.020	0.0183 < Saccharimonadia <= 0.0442
12	Adenocarcinoma	0.942	0.094	0.659	0.851	Gammaproteobacteria	0.121	0.020	0.0992 < Gammaproteobacteria <= 0.1579
13	Adenocarcinoma	0.506	0.538	0.825	0.422	Bacilli	0.366	-0.177	0.291 < Bacilli
13	Adenocarcinoma	0.506	0.538	0.825	0.422	Negativicutes	0.238	-0.073	0.198 < Negativicutes
13	Adenocarcinoma	0.506	0.538	0.825	0.422	Actinobacteria	0.112	-0.062	0.0589 < Actinobacteria
13	Adenocarcinoma	0.506	0.538	0.825	0.422	Synergistia	0.000	-0.054	Synergistia <= 4.86e-05
13	Adenocarcinoma	0.506	0.538	0.825	0.422	Saccharimonadia	0.001	-0.038	Saccharimonadia <= 0.0183
15	Adenocarcinoma	0.756	0.253	0.729	0.699	Campylobacteria	0.009	-0.109	0.00743 < Campylobacteria
15	Adenocarcinoma	0.756	0.253	0.729	0.699	Bacilli	0.245	0.072	0.210 < Bacilli <= 0.291
15	Adenocarcinoma	0.756	0.253	0.729	0.699	Synergistia	0.000	-0.056	Synergistia <= 4.86e-05
15	Adenocarcinoma	0.756	0.253	0.729	0.699	Negativicutes	0.107	0.038	0.096 < Negativicutes <= 0.139
15	Adenocarcinoma	0.756	0.253	0.729	0.699	Fusobacteriia	0.069	0.025	0.0502 < Fusobacteriia <= 0.0823
16	Adenocarcinoma	0.628	0.482	0.736	0.620	Bacilli	0.416	-0.178	0.291 < Bacilli
16	Adenocarcinoma	0.628	0.482	0.736	0.620	Actinobacteria	0.066	-0.066	0.0589 < Actinobacteria
16	Adenocarcinoma	0.628	0.482	0.736	0.620	Synergistia	0.001	0.044	4.37e-04 < Synergistia
16	Adenocarcinoma	0.628	0.482	0.736	0.620	Campylobacteria	0.003	0.044	0.00197 < Campylobacteria <= 0.00371
16	Adenocarcinoma	0.628	0.482	0.736	0.620	Negativicutes	0.108	0.040	0.096 < Negativicutes <= 0.139
17	Adenocarcinoma	0.688	0.405	0.720	0.670	Bacilli	0.311	-0.175	0.291 < Bacilli
17	Adenocarcinoma	0.688	0.405	0.720	0.670	Actinobacteria	0.009	0.037	Actinobacteria <= 0.0227
17	Adenocarcinoma	0.688	0.405	0.720	0.670	Campylobacteria	0.001	0.032	Campylobacteria <= 0.00197
17	Adenocarcinoma	0.688	0.405	0.720	0.670	Clostridia	0.025	0.033	0.0160 < Clostridia <= 0.0317
17	Adenocarcinoma	0.688	0.405	0.720	0.670	Synergistia	0.000	0.024	4.86e-05 < Synergistia <= 4.37e-04
18	Adenocarcinoma	0.792	0.139	0.700	0.718	Negativicutes	0.232	-0.071	0.198 < Negativicutes
18	Adenocarcinoma	0.792	0.139	0.700	0.718	Campylobacteria	0.005	0.045	0.00371 < Campylobacteria <= 0.00743
18	Adenocarcinoma	0.792	0.139	0.700	0.718	Bacilli	0.134	0.042	Bacilli <= 0.135
18	Adenocarcinoma	0.792	0.139	0.700	0.718	Actinobacteria	0.014	0.040	Actinobacteria <= 0.0227
18	Adenocarcinoma	0.792	0.139	0.700	0.718	Saccharimonadia	0.010	-0.038	Saccharimonadia <= 0.0183
19	Adenocarcinoma	0.510	0.573	0.807	0.467	Bacilli	0.444	-0.182	0.291 < Bacilli
19	Adenocarcinoma	0.510	0.573	0.807	0.467	Campylobacteria	0.008	-0.100	0.00743 < Campylobacteria
19	Adenocarcinoma	0.510	0.573	0.807	0.467	Synergistia	0.000	-0.052	Synergistia <= 4.86e-05

19	Adenocarcinoma	0.510	0.573	0.807	0.467	Saccharimonadia	0.004	-0.044	Saccharimonadia <= 0.0183
19	Adenocarcinoma	0.510	0.573	0.807	0.467	Negativicutes	0.136	0.038	0.096 < Negativicutes <= 0.139
27	Adenocarcinoma	0.946	0.159	0.663	0.830	Bacilli	0.157	0.076	0.135 < Bacilli <= 0.210
27	Adenocarcinoma	0.946	0.159	0.663	0.830	Synergistia	0.001	0.053	4.37e-04 < Synergistia
27	Adenocarcinoma	0.946	0.159	0.663	0.830	Clostridia	0.059	-0.048	0.0455 < Clostridia
27	Adenocarcinoma	0.946	0.159	0.663	0.830	Actinobacteria	0.006	0.042	Actinobacteria <= 0.0227
27	Adenocarcinoma	0.946	0.159	0.663	0.830	Negativicutes	0.110	0.043	0.096 < Negativicutes <= 0.139
28	Adenocarcinoma	0.706	0.187	0.714	0.737	Bacilli	0.182	0.076	0.135 < Bacilli <= 0.210
28	Adenocarcinoma	0.706	0.187	0.714	0.737	Negativicutes	0.288	-0.071	0.198 < Negativicutes
28	Adenocarcinoma	0.706	0.187	0.714	0.737	Synergistia	0.000	-0.050	Synergistia <= 4.86e-05
28	Adenocarcinoma	0.706	0.187	0.714	0.737	Campylobacteria	0.004	0.047	0.00371 < Campylobacteria <= 0.00743
28	Adenocarcinoma	0.706	0.187	0.714	0.737	Fusobacteriia	0.077	0.021	0.0502 < Fusobacteriia <= 0.0823
30	Adenocarcinoma	0.778	0.082	0.666	0.833	Bacilli	0.209	0.066	0.135 < Bacilli <= 0.210
30	Adenocarcinoma	0.778	0.082	0.666	0.833	Campylobacteria	0.003	0.037	0.00197 < Campylobacteria <= 0.00371
30	Adenocarcinoma	0.778	0.082	0.666	0.833	Synergistia	0.000	0.027	4.86e-05 < Synergistia <= 4.37e-04
30	Adenocarcinoma	0.778	0.082	0.666	0.833	Fusobacteriia	0.058	0.021	0.0502 < Fusobacteriia <= 0.0823
30	Adenocarcinoma	0.778	0.082	0.666	0.833	Saccharimonadia	0.057	0.016	0.0561 < Saccharimonadia
31	Adenocarcinoma	0.564	0.532	0.792	0.465	Bacilli	0.338	-0.176	0.291 < Bacilli
31	Adenocarcinoma	0.564	0.532	0.792	0.465	Campylobacteria	0.008	-0.103	0.00743 < Campylobacteria
31	Adenocarcinoma	0.564	0.532	0.792	0.465	Saccharimonadia	0.007	-0.046	Saccharimonadia <= 0.0183
31	Adenocarcinoma	0.564	0.532	0.792	0.465	Fusobacteriia	0.125	-0.034	0.1242 < Fusobacteriia
31	Adenocarcinoma	0.564	0.532	0.792	0.465	Actinobacteria	0.026	0.030	0.0227 < Actinobacteria <= 0.0351
35	Adenocarcinoma	0.846	0.188	0.716	0.732	Negativicutes	0.202	-0.075	0.198 < Negativicutes
35	Adenocarcinoma	0.846	0.188	0.716	0.732	Bacilli	0.195	0.069	0.135 < Bacilli <= 0.210
35	Adenocarcinoma	0.846	0.188	0.716	0.732	Synergistia	0.000	-0.053	Synergistia <= 4.86e-05
35	Adenocarcinoma	0.846	0.188	0.716	0.732	Campylobacteria	0.003	0.042	0.00197 < Campylobacteria <= 0.00371
35	Adenocarcinoma	0.846	0.188	0.716	0.732	Actinobacteria	0.028	0.032	0.0227 < Actinobacteria <= 0.0351
36	Adenocarcinoma	0.774	0.274	0.734	0.631	Campylobacteria	0.008	-0.101	0.00743 < Campylobacteria
36	Adenocarcinoma	0.774	0.274	0.734	0.631	Bacilli	0.230	0.074	0.210 < Bacilli <= 0.291
36	Adenocarcinoma	0.774	0.274	0.734	0.631	Actinobacteria	0.104	-0.071	0.0589 < Actinobacteria
36	Adenocarcinoma	0.774	0.274	0.734	0.631	Clostridia	0.061	-0.048	0.0455 < Clostridia
36	Adenocarcinoma	0.774	0.274	0.734	0.631	Negativicutes	0.157	0.044	0.139 < Negativicutes <= 0.198
37	Adenocarcinoma	0.820	0.099	0.675	0.800	Synergistia	0.001	0.048	4.37e-04 < Synergistia
37	Adenocarcinoma	0.820	0.099	0.675	0.800	Saccharimonadia	0.015	-0.039	Saccharimonadia <= 0.0183
37	Adenocarcinoma	0.820	0.099	0.675	0.800	Negativicutes	0.107	0.039	0.096 < Negativicutes <= 0.139
37	Adenocarcinoma	0.820	0.099	0.675	0.800	Bacilli	0.116	0.038	Bacilli <= 0.135
37	Adenocarcinoma	0.820	0.099	0.675	0.800	Campylobacteria	0.004	0.038	0.00371 < Campylobacteria <= 0.00743
40	Adenoma	0.616	0.526	0.175	0.573	Bacilli	0.354	0.181	0.291 < Bacilli
40	Adenoma	0.616	0.526	0.175	0.573	Actinobacteria	0.092	0.068	0.0589 < Actinobacteria
40	Adenoma	0.616	0.526	0.175	0.573	Synergistia	0.000	0.058	Synergistia <= 4.86e-05

40	Adenoma	0.616	0.526	0.175	0.573	Clostridia	0.056	0.051	0.0455 < Clostridia
40	Adenoma	0.616	0.526	0.175	0.573	Saccharimonadia	0.002	0.041	Saccharimonadia <= 0.0183
42	Adenoma	0.508	0.111	0.271	0.309	Synergistia	0.000	0.055	Synergistia <= 4.86e-05
42	Adenoma	0.508	0.111	0.271	0.309	Negativicutes	0.183	-0.043	0.139 < Negativicutes <= 0.198
42	Adenoma	0.508	0.111	0.271	0.309	Saccharimonadia	0.003	0.041	Saccharimonadia <= 0.0183
42	Adenoma	0.508	0.111	0.271	0.309	Campylobacteria	0.003	-0.041	0.00197 < Campylobacteria <= 0.00371
42	Adenoma	0.508	0.111	0.271	0.309	Bacteroidia	0.067	0.026	Bacteroidia <= 0.207
43	Adenoma	0.566	0.078956971	0.279447952	0.304892561	Synergistia	0	0.053531409	Synergistia <= 4.86e-05
43	Adenoma	0.566	0.078956971	0.279447952	0.304892561	Saccharimonadia	0.007691396	0.042835316	Saccharimonadia <= 0.0183
43	Adenoma	0.566	0.078956971	0.279447952	0.304892561	Clostridia	0.007243609	-0.02540001	Clostridia <= 0.0160
43	Adenoma	0.566	0.078956971	0.279447952	0.304892561	Campylobacteria	0.000302915	-0.025309394	Campylobacteria <= 0.00197
43	Adenoma	0.566	0.078956971	0.279447952	0.304892561	Coriobacteriia	0.002186253	-0.020212712	Coriobacteriia <= 0.00231
44	Adenocarcinoma	0.598	0.469960198	0.789796356	0.516171418	Bacilli	0.370826703	-0.175608669	0.291 < Bacilli
44	Adenocarcinoma	0.598	0.469960198	0.789796356	0.516171418	Actinobacteria	0.100048843	-0.062527395	0.0589 < Actinobacteria
44	Adenocarcinoma	0.598	0.469960198	0.789796356	0.516171418	Synergistia	3.15114E-05	-0.051224124	Synergistia <= 4.86e-05
44	Adenocarcinoma	0.598	0.469960198	0.789796356	0.516171418	Campylobacteria	0.005167877	0.039909672	0.00371 < Campylobacteria <= 0.00743
44	Adenocarcinoma	0.598	0.469960198	0.789796356	0.516171418	Bacteroidia	0.202524067	-0.024174422	Bacteroidia <= 0.207
45	Adenocarcinoma	0.59	0.438831304	0.751164525	0.621000601	Bacilli	0.435791327	-0.173422961	0.291 < Bacilli
45	Adenocarcinoma	0.59	0.438831304	0.751164525	0.621000601	Synergistia	0	-0.055024064	Synergistia <= 4.86e-05
45	Adenocarcinoma	0.59	0.438831304	0.751164525	0.621000601	Negativicutes	0.097380586	0.040354367	0.096 < Negativicutes <= 0.139
45	Adenocarcinoma	0.59	0.438831304	0.751164525	0.621000601	Campylobacteria	0.001893022	0.029377816	Campylobacteria <= 0.00197
45	Adenocarcinoma	0.59	0.438831304	0.751164525	0.621000601	Actinobacteria	0.031609069	0.028550918	0.0227 < Actinobacteria <= 0.0351
48	Adenoma	0.508	0.504850087	0.197679894	0.521389527	Bacilli	0.391159096	0.171736422	0.291 < Bacilli
48	Adenoma	0.508	0.504850087	0.197679894	0.521389527	Negativicutes	0.224933795	0.071946513	0.198 < Negativicutes
48	Adenoma	0.508	0.504850087	0.197679894	0.521389527	Actinobacteria	0.101283357	0.063895886	0.0589 < Actinobacteria
48	Adenoma	0.508	0.504850087	0.197679894	0.521389527	Synergistia	0	0.051499752	Synergistia <= 4.86e-05
48	Adenoma	0.508	0.504850087	0.197679894	0.521389527	Campylobacteria	0.002037075	-0.03536894	0.00197 < Campylobacteria <= 0.00371
50	Adenocarcinoma	0.692	0.477183779	0.776752559	0.560165526	Bacilli	0.305054676	-0.173374441	0.291 < Bacilli
50	Adenocarcinoma	0.692	0.477183779	0.776752559	0.560165526	Actinobacteria	0.150965169	-0.065882858	0.0589 < Actinobacteria
50	Adenocarcinoma	0.692	0.477183779	0.776752559	0.560165526	Synergistia	0	-0.048454586	Synergistia <= 4.86e-05
50	Adenocarcinoma	0.692	0.477183779	0.776752559	0.560165526	Campylobacteria	0.00331274	0.042030544	0.00197 < Campylobacteria <= 0.00371
50	Adenocarcinoma	0.692	0.477183779	0.776752559	0.560165526	Clostridia	0.024013439	0.029094308	0.0160 < Clostridia <= 0.0317



Figure ML3-B-6: The LIME Result in Training dataset for Class in Saliva

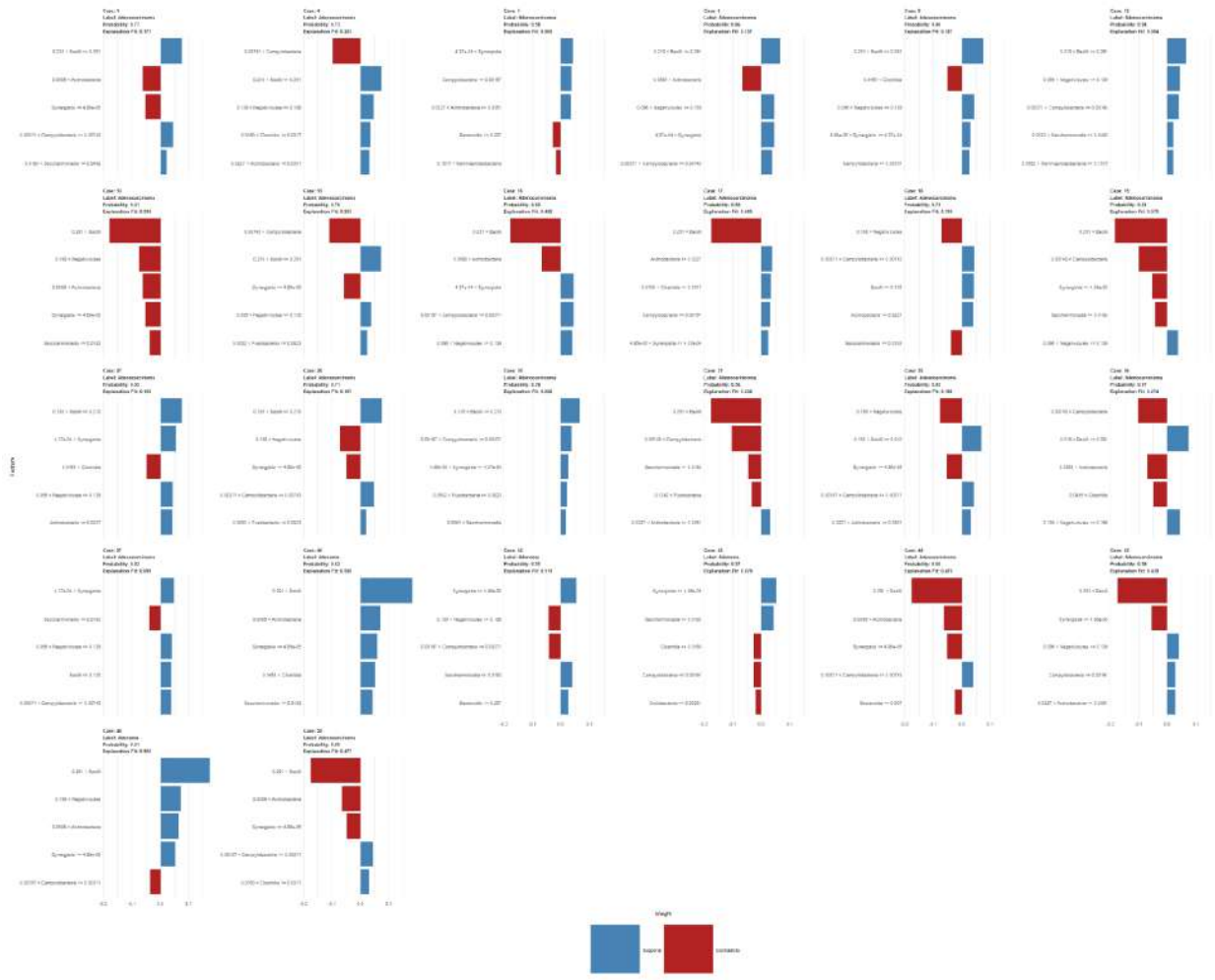


Figure ML3-B-7: The LIME Result in Test dataset for Class in Saliva

Table ML3-B-6: The Aggregated SHAP Values

Variable	Training Dataset				Testing Dataset			
	contribution	position	sign	cumulative	contribution	position	sign	cumulative
intercept	0.7413	18	X	0.741	0.7413	18	X	0.741
Campylobacteria	0.0008	17	1	0.742	0.0075	17	1	0.749
Saccharimonadia	0.0002	16	1	0.742	-0.0015	15	-1	0.741
Tumor	0.0000	15	1	0.742	0.0000	14	1	0.741
Coriobacteriia	-0.0001	14	-1	0.742	-0.0025	13	-1	0.738

Bacteroidia	0.0000	13	-1	0.742	-0.0066	16	-1	0.742
Cyanobacteriia	-0.0001	12	-1	0.742	-0.0003	12	-1	0.738
Negativicutes	-0.0003	11	-1	0.742	0.0064	10	1	0.744
Gracilibacteria	-0.0001	10	-1	0.742	0.0005	9	1	0.745
Synergistia	-0.0005	9	-1	0.741	-0.0042	8	-1	0.741
Spirochaetia	0.0001	8	1	0.741	0.0003	7	1	0.741
Gender	0.0000	7	1	0.741	0.0000	11	1	0.738
Clostridia	0.0000	6	1	0.741	0.0060	6	1	0.747
Gammaproteobacteria	0.0001	5	1	0.741	-0.0045	4	-1	0.747
Fusobacteriia	0.0004	4	1	0.742	0.0043	5	1	0.751
Actinobacteria	-0.0004	3	-1	0.741	-0.0075	3	-1	0.739
Bacilli	-0.0001	2	-1	0.741	-0.0597	2	-1	0.679
prediction	0.7413	1	X	0.741	0.6795	1	X	0.679

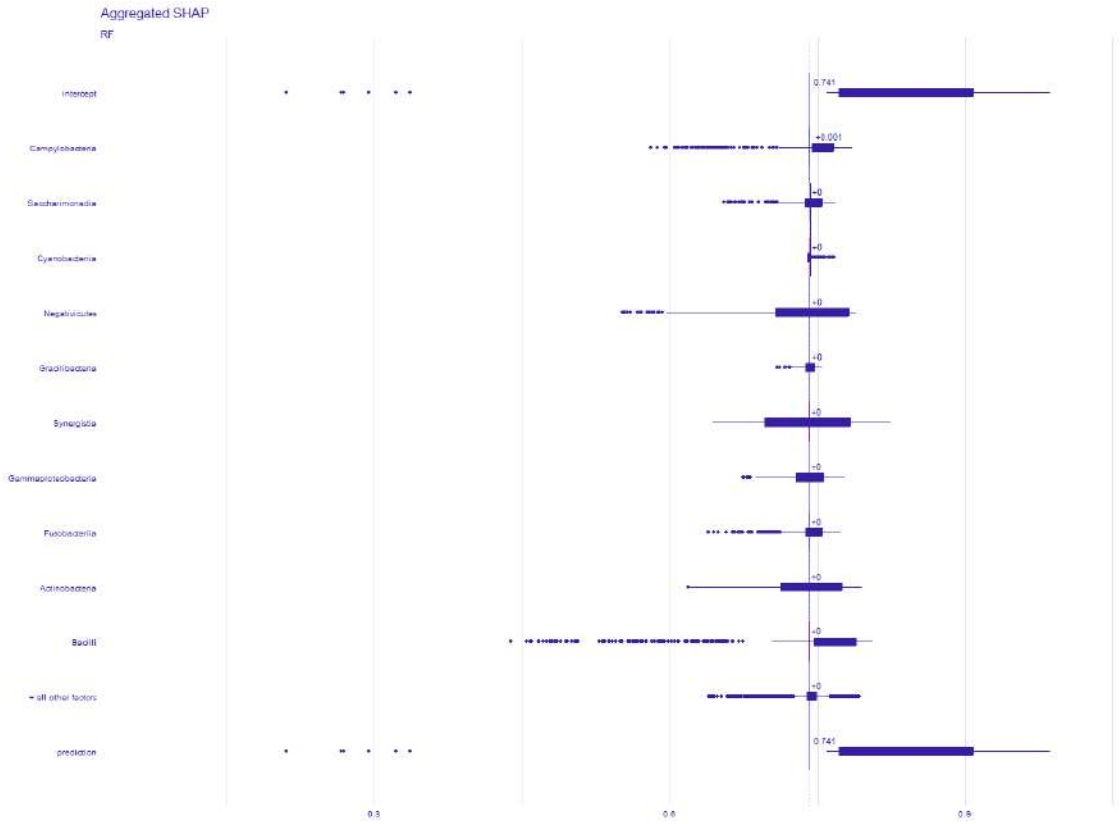


Figure ML3-B-8: The Aggregated SHAP Value in Training dataset for Class in Saliva

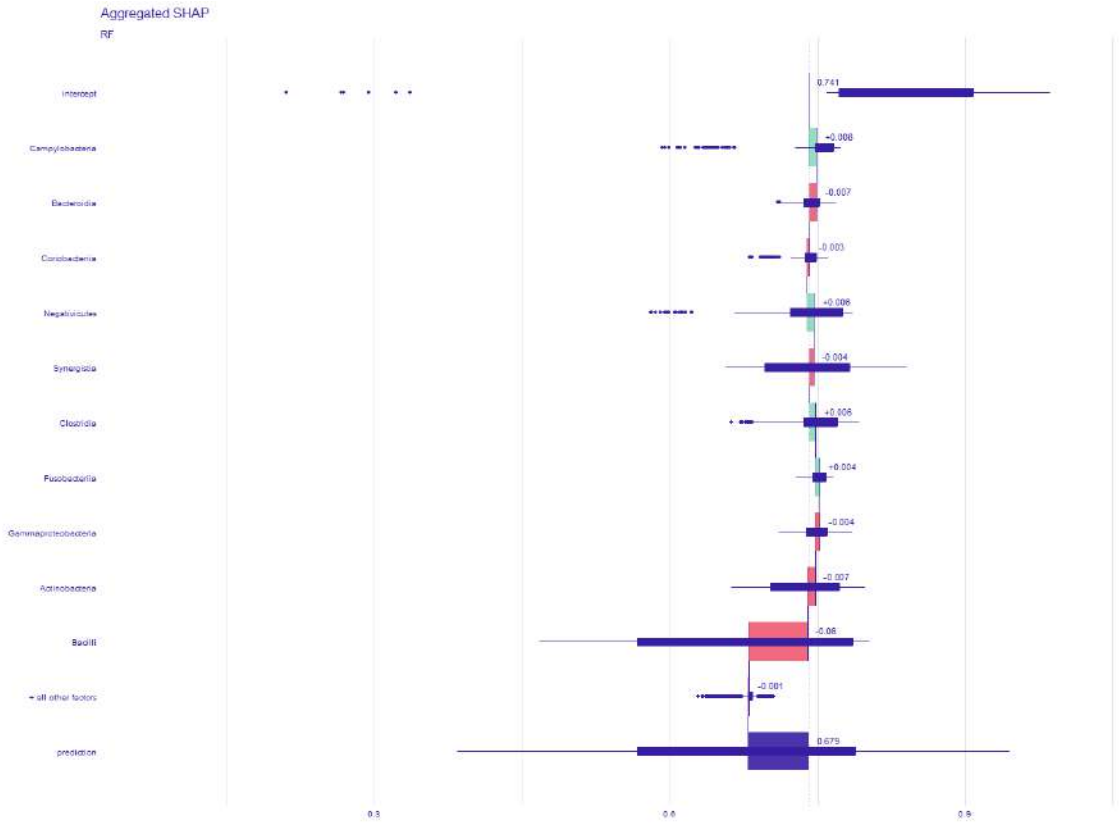


Figure ML3-B-9: The Aggregated SHAP Value in Test dataset for Class in Saliva

C- Order

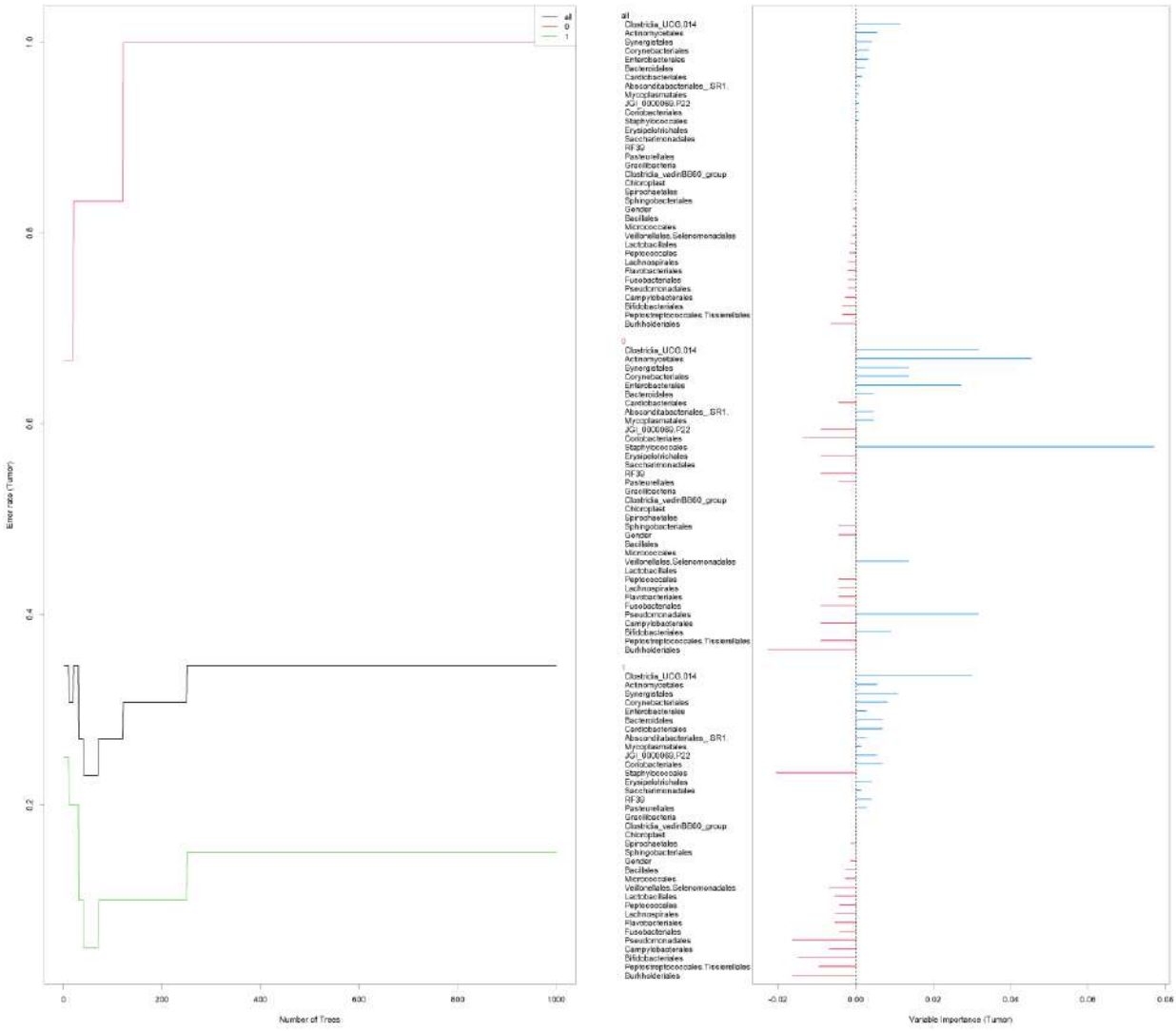
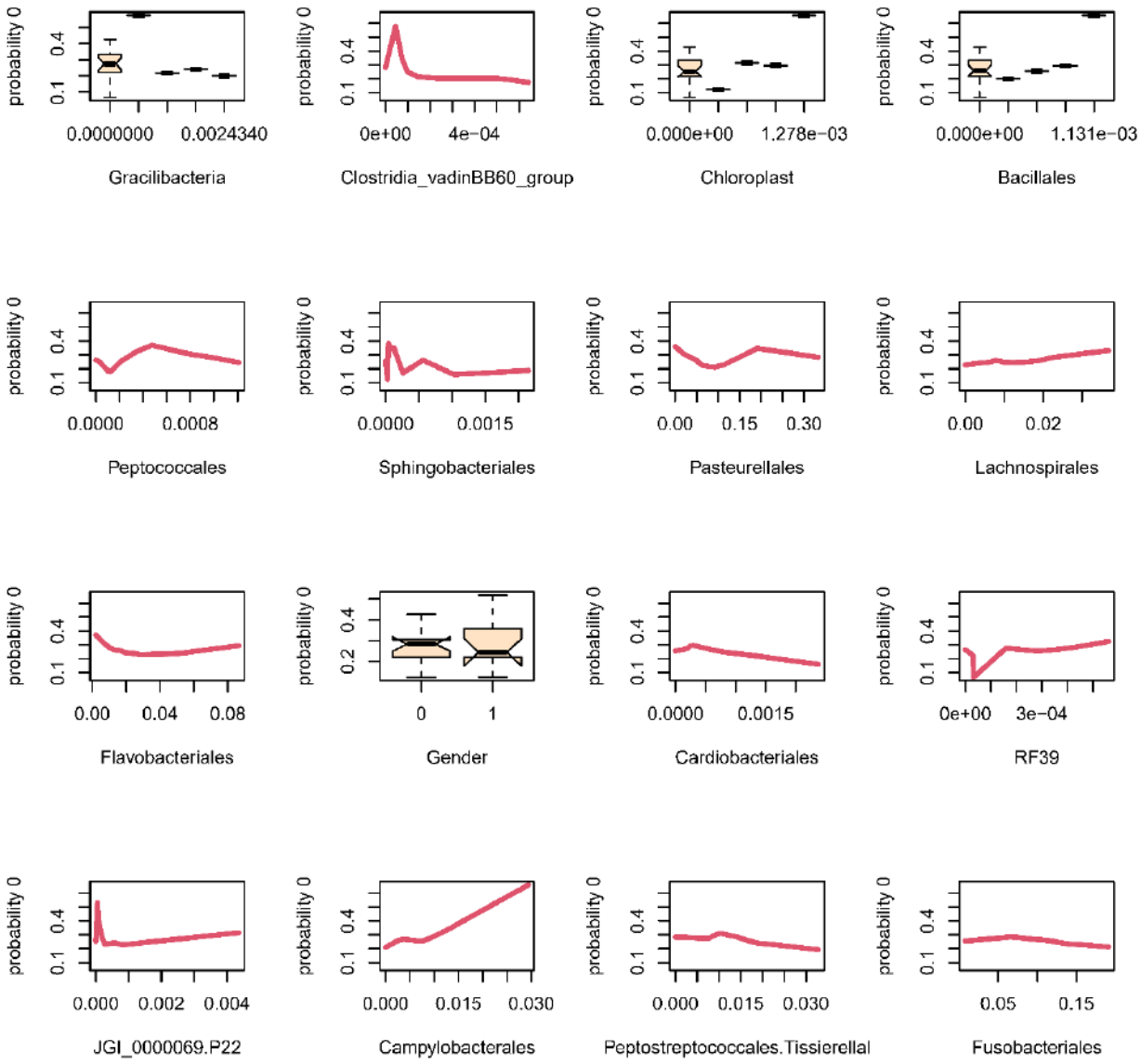
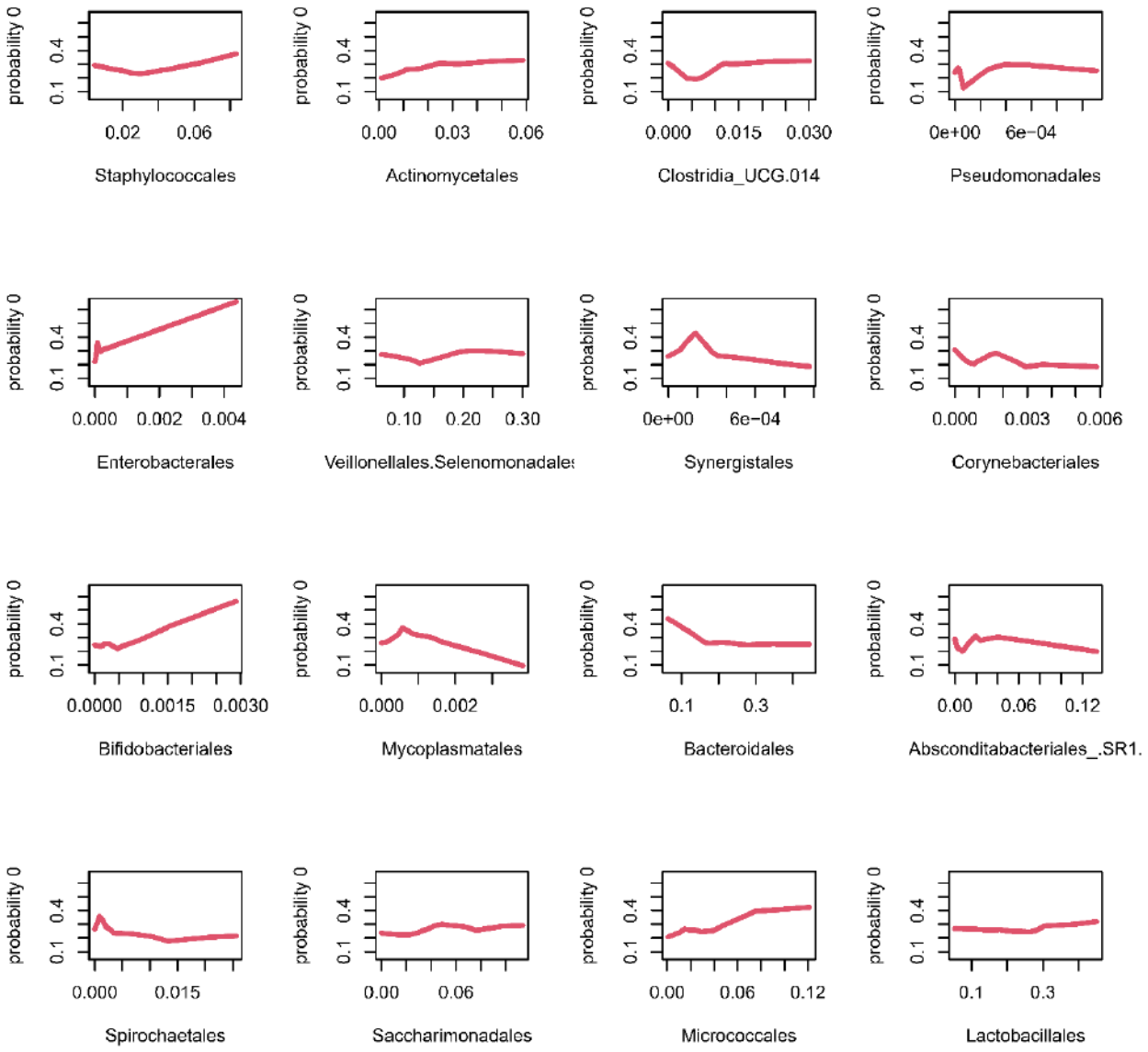


Figure ML3-C-1: The number of trees and Variable Importance Plot for Order in Saliva


probability 0
0.1 0.4
0.0000000 0.0024340
Gracilibacteria
0e+00 4e−04
Clostridia_vadinBB60_group
0.000e+00 1.278e−03
Chloroplast
0.000e+00 1.131e−03
Bacillales
0.0000 0.0008
Peptococcales
0.0000 0.0015
Sphingobacteriales
0.00 0.15 0.30
Pasteurellales
0.00 0.02
Lachnospirales
0.00 0.04 0.08
Flavobacteriales
probability 0
0.2 0.4
0 1
Gender
0.0000 0.0015
Cardiobacteriales
0e+00 3e−04
RF39
0.000 0.002 0.004
JGI_0000069.P22
0.000 0.015 0.030
Campylobacterales
0.000 0.015 0.030
Peptostreptococcales.Tissierellal
0.05 0.15
Fusobacteriales


probability 0
0.1 0.4
0.02 0.06
Staphylococcales
probability 0
0.1 0.4
0.00 0.03 0.06
Actinomycetales
probability 0
0.1 0.4
0.000 0.015 0.030
Clostridia_UCG.014
probability 0
0.1 0.4
0e+00 6e−04
Pseudomonadales
probability 0
0.1 0.4
0.000 0.002 0.004
Enterobacterales
probability 0
0.1 0.4
0.10 0.20 0.30
Veillonellales.Selenomonadales
probability 0
0.1 0.4
0e+00 6e−04
Synergistales
probability 0
0.1 0.4
0.000 0.003 0.006
Corynebacteriales
probability 0
0.1 0.4
0.0000 0.0015 0.0030
Bifidobacteriales
probability 0
0.1 0.4
0.000 0.002
Mycoplasmatales
probability 0
0.1 0.4
0.1 0.3
Bacteroidales
probability 0
0.1 0.4
0.00 0.06 0.12
Absconditabacteriales_.SR1.
probability 0
0.1 0.4
0.000 0.015
Spirochaetales
probability 0
0.1 0.4
0.00 0.06
Saccharimonadales
probability 0
0.1 0.4
0.00 0.06 0.12
Micrococcales
probability 0
0.1 0.4
0.1 0.3
Lactobacillales

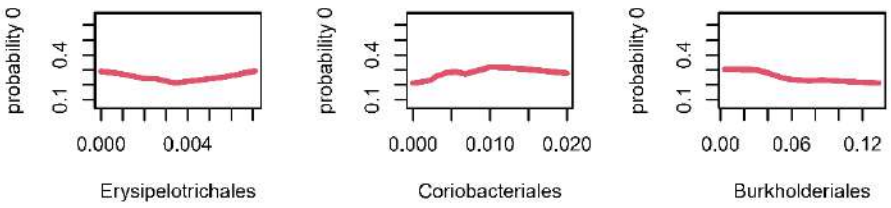
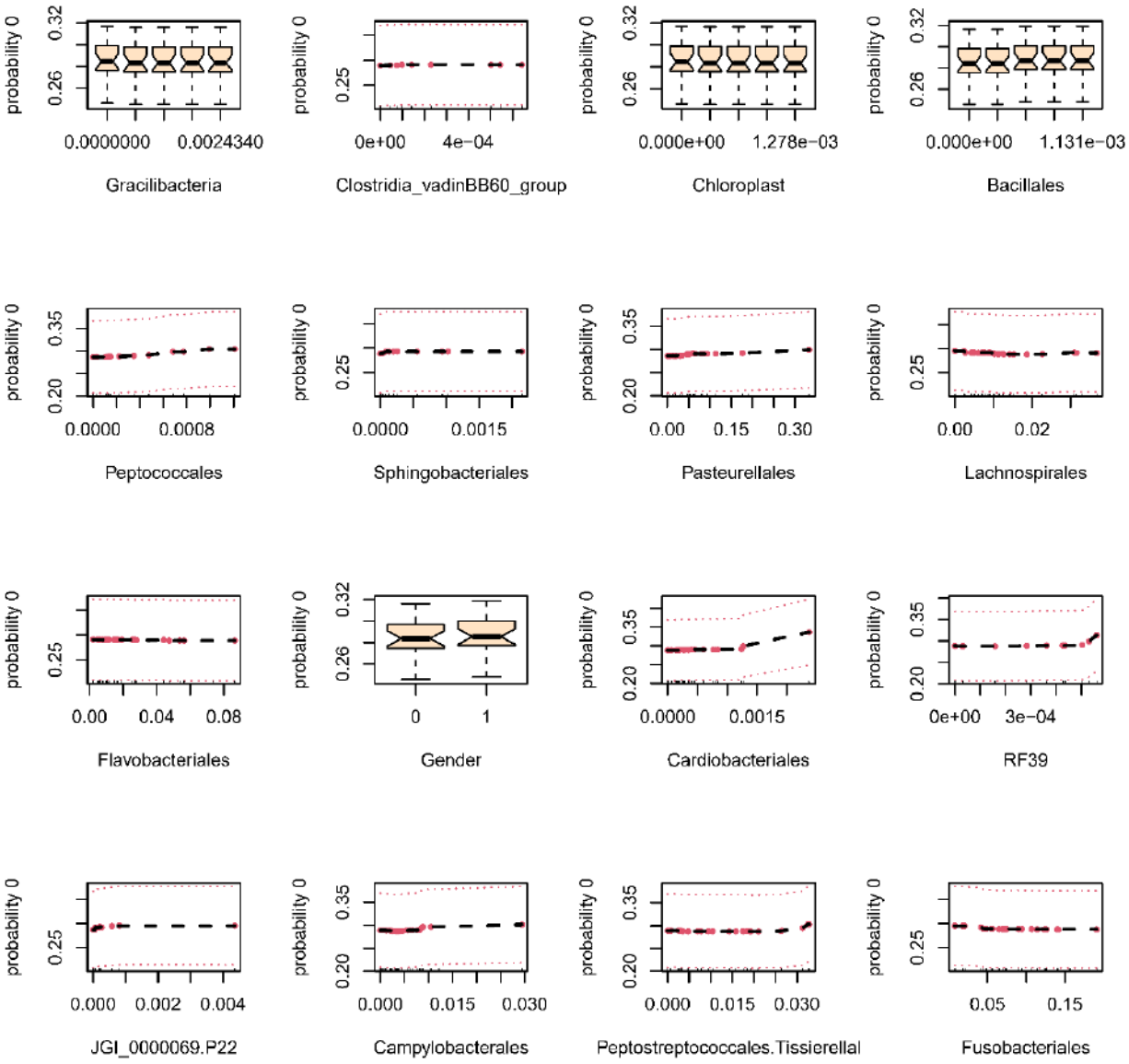
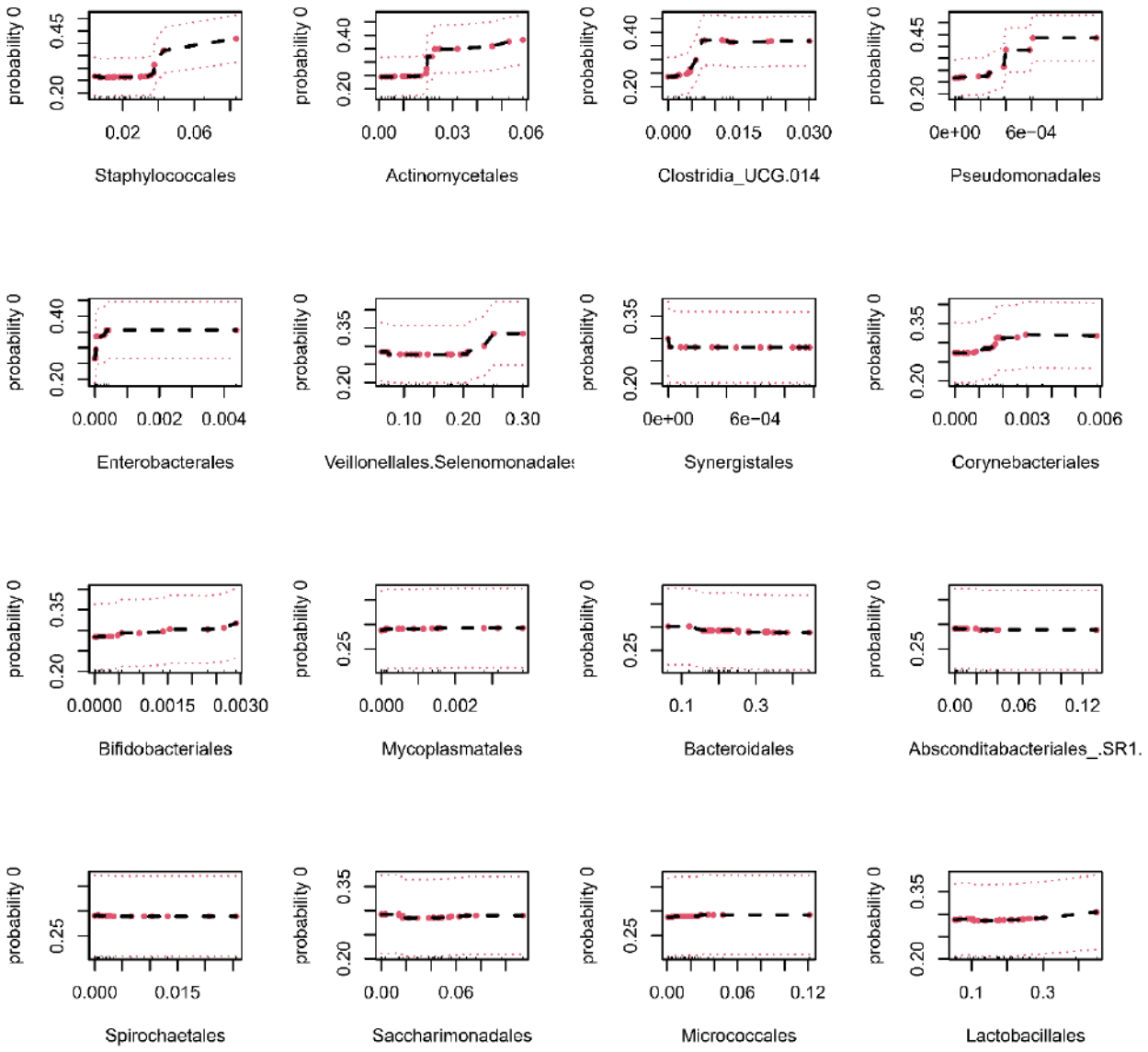


Figure ML3-C-2: Marginal Plot for Order in Saliva


probability 0
0.26
0.32
0.0000000
0.0024340
Gracilibacteria
probability 0
0.25
0e+00
4e−04
Clostridia_vadinBB60_group
probability 0
0.26
0.32
0.000e+00
1.278e−03
Chloroplast
probability 0
0.26
0.32
0.000e+00
1.131e−03
Bacillales
probability 0
0.20
0.35
0.0000
0.0008
Peptococcales
probability 0
0.25
0.0000
0.0015
Sphingobacteriales
probability 0
0.20
0.35
0.00
0.15
0.30
Pasteurellales
probability 0
0.25
0.00
0.02
Lachnospirales
probability 0
0.25
0.00
0.04
0.08
Flavobacteriales
probability 0
0.26
0.32
0
1
Gender
probability 0
0.20
0.35
0.0000
0.0015
Cardiobacteriales
probability 0
0.20
0.35
0e+00
3e−04
RF39
probability 0
0.25
0.000
0.002
0.004
JGI_0000069.P22
probability 0
0.20
0.35
0.000
0.015
0.030
Campylobacterales
probability 0
0.20
0.35
0.000
0.015
0.030
Peptostreptococcales.Tissierellal
probability 0
0.25
0.05
0.15
Fusobacteriales


probability 0
Staphylococcales
Actinomycetales
Clostridia_UCG.014
Pseudomonadales
Enterobacterales
Veillonellales.Selenomonadales
Synergistales
Corynebacteriales
Bifidobacteriales
Mycoplasmatales
Bacteroidales
Absconditabacteriales_.SR1.
Spirochaetales
Saccharimonadales
Micrococcales
Lactobacillales

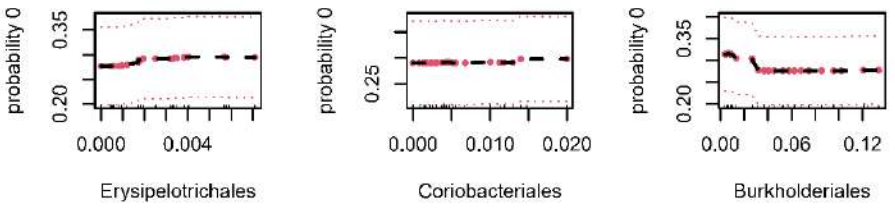


Figure ML3-C-3: Partial Plot for Order in Saliva

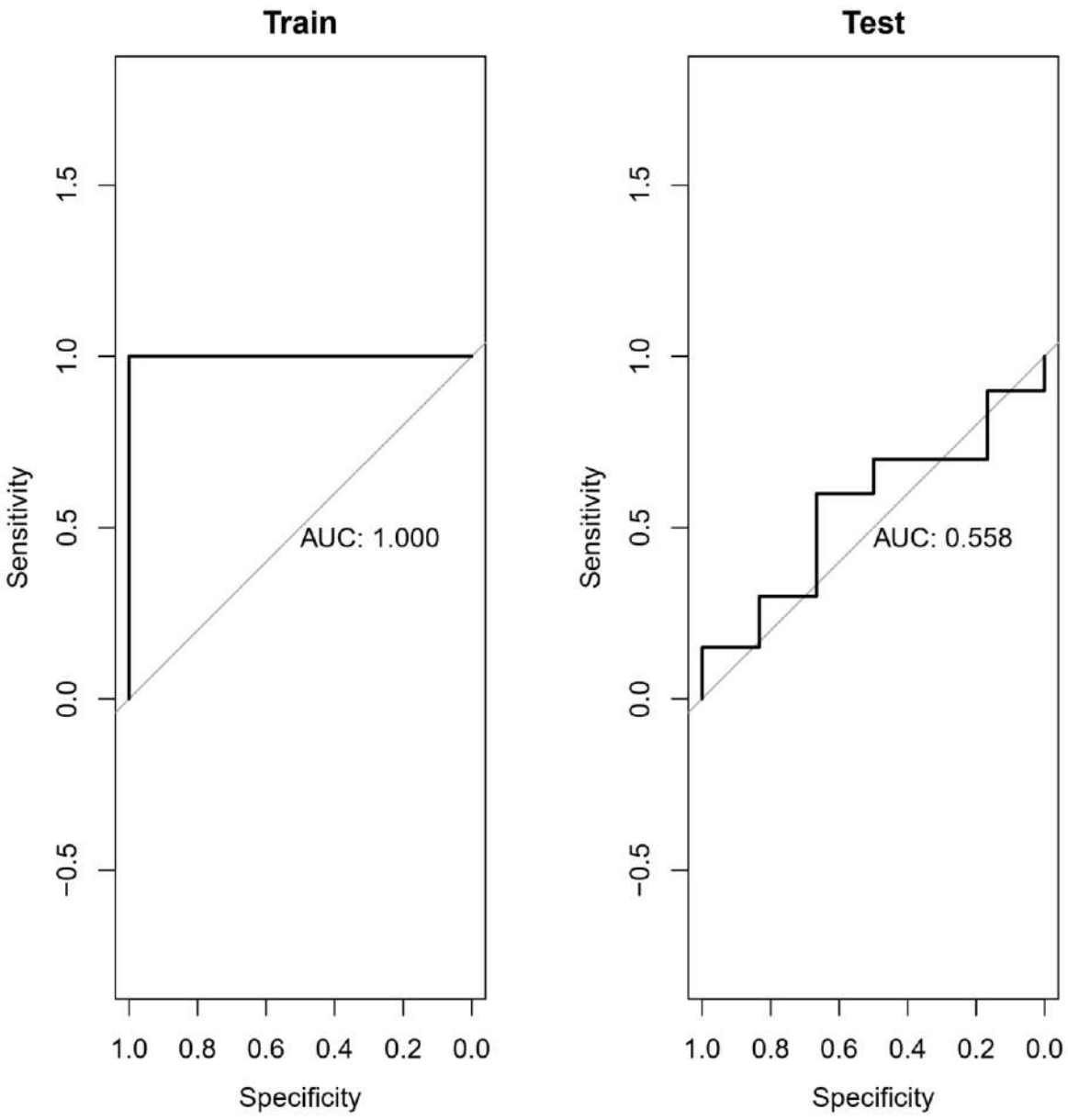


Figure ML3-C-4: The ROC Curve for Order in Saliva

Table ML3-C-1: The Model Performance

Index	Values		
	2.50%	50.00%	97.50%
Accuracy	0.31	0.65	0.85
Specificity	0.33	0.83	1.00
Sensitivity	0.10	0.60	0.95

Table ML3-C-2: The Variable Importance Values

Covariates	Varibale Importance			Brier Score		
	all	AP	CRC	all	AP	CRC
Gender	-0.0016	-0.0136	-0.0014	-0.0032	-0.0022	-0.0022
Absconditabacteriales_.SR1.	0.0000	-0.0045	0.0014	0.0006	0.0004	0.0004
Actinomycetales	0.0483	-0.0815	0.1930	0.1249	0.0849	0.0849
Bacillales	0.0004	-0.0091	0.0041	-0.0006	-0.0004	-0.0004
Bacteroidales	0.0004	-0.0544	0.0177	0.0045	0.0030	0.0030
Bifidobacteriales	0.0047	-0.0861	0.0421	-0.0014	-0.0010	-0.0010
Burkholderiales	0.0001	-0.0906	0.0272	-0.0006	-0.0004	-0.0004
Campylobacterales	0.0023	-0.0544	0.0245	-0.0045	-0.0031	-0.0031
Cardiobacteriales	0.0097	-0.2129	0.0979	0.0366	0.0249	0.0249
Chloroplast	-0.0004	0.0000	-0.0014	-0.0016	-0.0011	-0.0011
Clostridia_UCG.014	0.0382	0.0136	0.1291	0.1172	0.0797	0.0797
Clostridia_vadinBB60_group	-0.0023	-0.0136	-0.0041	-0.0024	-0.0016	-0.0016
Coriobacteriales	0.0016	-0.0091	0.0082	0.0046	0.0031	0.0031
Corynebacteriales	0.0008	-0.0634	0.0217	0.0055	0.0037	0.0037
Enterobacterales	0.0234	-0.1178	0.1169	0.0452	0.0307	0.0307
Erysipelotrichales	-0.0070	0.0181	-0.0299	-0.0099	-0.0067	-0.0067
Flavobacteriales	-0.0028	-0.0181	-0.0041	-0.0069	-0.0047	-0.0047
Fusobacteriales	-0.0012	-0.0091	-0.0014	-0.0041	-0.0028	-0.0028
Gracilibacteria	-0.0016	0.0000	-0.0054	-0.0006	-0.0004	-0.0004
JGI_0000069.P22	0.0000	-0.0362	0.0109	-0.0082	-0.0056	-0.0056
Lachnospirales	-0.0011	-0.0272	0.0041	-0.0103	-0.0070	-0.0070
Lactobacillales	0.0052	-0.0498	0.0326	-0.0031	-0.0021	-0.0021
Micrococcales	-0.0011	0.0000	-0.0041	-0.0020	-0.0014	-0.0014
Mycoplasmatales	-0.0016	-0.0091	-0.0027	-0.0041	-0.0028	-0.0028
Pasteurellales	-0.0007	-0.0272	0.0054	-0.0018	-0.0012	-0.0012
Peptococcales	-0.0008	-0.0498	0.0122	-0.0033	-0.0023	-0.0023
Peptostreptococcales.Tissierellales	-0.0005	-0.0680	0.0190	-0.0059	-0.0040	-0.0040
Pseudomonadales	0.0762	-0.4938	0.4132	0.1869	0.1270	0.1270
RF39	0.0010	-0.1223	0.0408	-0.0051	-0.0034	-0.0034
Saccharimonadales	-0.0028	-0.0362	0.0014	-0.0105	-0.0071	-0.0071

Sphingobacteriales	-0.0012	-0.0136	0.0000	-0.0081	-0.0055	-0.0055
Spirochaetales	0.0008	-0.0045	0.0041	-0.0025	-0.0017	-0.0017
Staphylococcales	0.0878	-0.2718	0.3874	0.2061	0.1401	0.1401
Synergistales	0.0023	0.0136	0.0041	0.0124	0.0084	0.0084
Veillonellales.Selenomonadales	0.0233	-0.1178	0.1169	0.0322	0.0219	0.0219

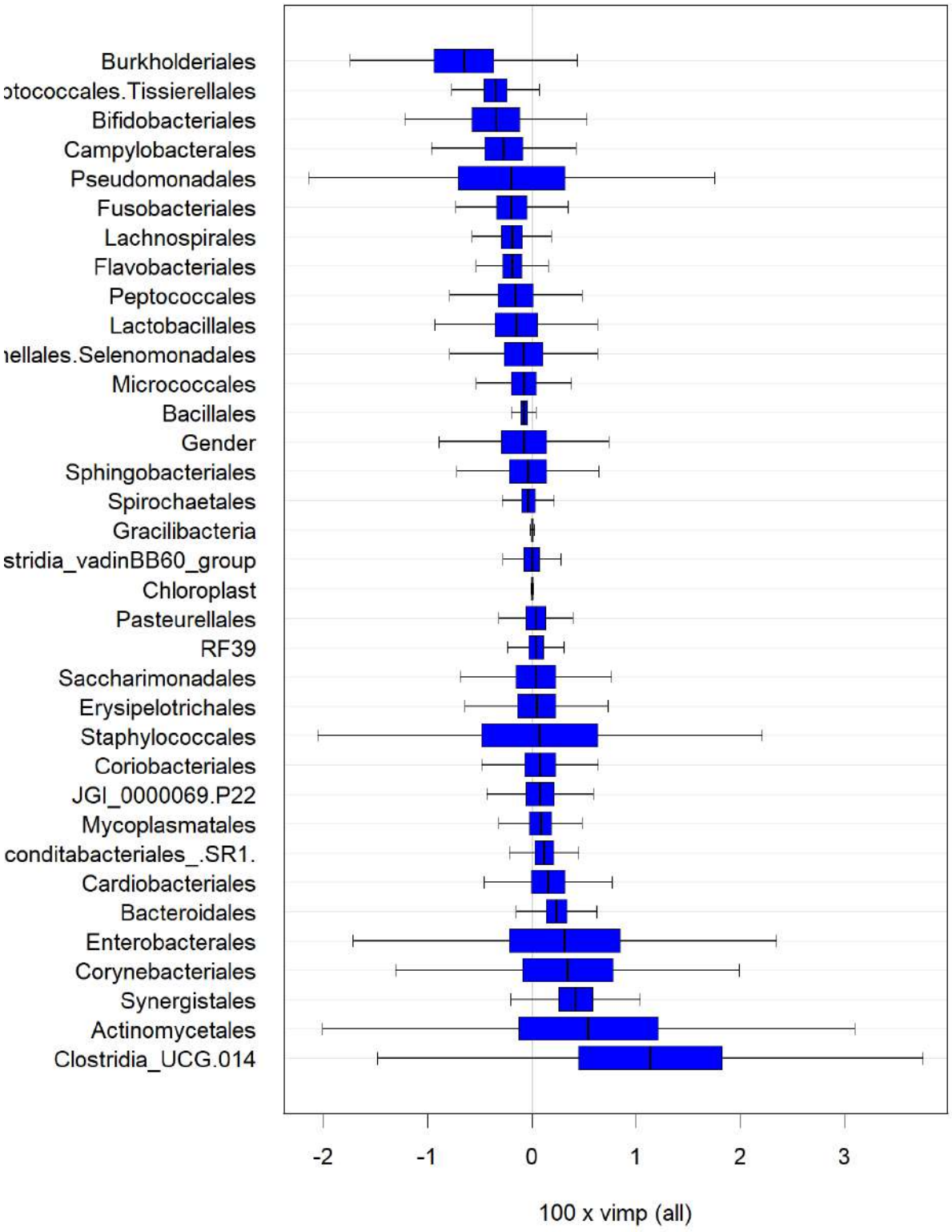


Figure ML3-C-5: The Variable Importance Plot (VIMP) for Order in Saliva

Best model is selected for LIME results

Table ML3-C-3: The best setting for model and LIME Results

Setting	mtry	ROC	Sens	Spec	ROCSD	SensSD	SpecSD
ntrees: 500 nodesize: 1	7	0.8	0	0.95	0.273861	0	0.111803
ntrees: 500 nodesize: 1	8	0.8	0	0.95	0.273861	0	0.111803
ntrees: 500 nodesize: 1	10	0.8	0.2	0.95	0.273861	0.447214	0.111803
ntrees: 1000 nodesize: 1	5	0.8	0	0.95	0.273861	0	0.111803
ntrees: 1000 nodesize: 1	6	0.8	0	0.95	0.273861	0	0.111803
ntrees: 1000 nodesize: 1	7	0.8	0	0.95	0.273861	0	0.111803
ntrees: 1000 nodesize: 1	8	0.8	0	0.95	0.273861	0	0.111803
ntrees: 1000 nodesize: 1	9	0.8	0	0.95	0.273861	0	0.111803
ntrees: 1000 nodesize: 1	10	0.8	0.2	0.95	0.273861	0.447214	0.111803
ntrees: 500 nodesize: 5	9	0.775	0.2	0.9	0.256174	0.447214	0.136931
ntrees: 1000 nodesize: 5	9	0.8	0.2	0.9	0.273861	0.447214	0.136931

Table ML3-C-4: The LIME Result for Training Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
2	Adenoma	0.642	0.409	0.293	0.491	Staphylococcales	0.043	0.108	0.0352 < Staphylococcales
2	Adenoma	0.642	0.409	0.293	0.491	Pseudomonadales	0.001	0.099	2.38e-04 < Pseudomonadales
2	Adenoma	0.642	0.409	0.293	0.491	Enterobacterales	0.000	-0.069	Enterobacterales <= 4.03e-05
2	Adenoma	0.642	0.409	0.293	0.491	Clostridia_UCG.014	0.008	0.033	0.00404 < Clostridia_UCG.014 <= 0.01058
2	Adenoma	0.642	0.409	0.293	0.491	Bifidobacteriales	0.003	0.027	0.000537 < Bifidobacteriales
3	Adenoma	0.691	0.457	0.213	0.586	Actinomycetales	0.053	0.102	0.02245 < Actinomycetales
3	Adenoma	0.691	0.457	0.213	0.586	Pseudomonadales	0.000	0.097	2.38e-04 < Pseudomonadales
3	Adenoma	0.691	0.457	0.213	0.586	Clostridia_UCG.014	0.012	0.071	0.01058 < Clostridia_UCG.014
3	Adenoma	0.691	0.457	0.213	0.586	Enterobacterales	0.000	0.064	4.03e-05 < Enterobacterales
3	Adenoma	0.691	0.457	0.213	0.586	Veillonellales.Selenomonadales	0.251	0.039	0.198 < Veillonellales.Selenomonadales
6	Adenoma	0.693	0.433	0.255	0.464	Actinomycetales	0.059	0.110	0.02245 < Actinomycetales
6	Adenoma	0.693	0.433	0.255	0.464	Pseudomonadales	0.001	0.105	2.38e-04 < Pseudomonadales
6	Adenoma	0.693	0.433	0.255	0.464	Enterobacterales	0.000	0.071	4.03e-05 < Enterobacterales
6	Adenoma	0.693	0.433	0.255	0.464	Clostridia_UCG.014	0.002	-0.039	0.00172 < Clostridia_UCG.014 <= 0.00404
6	Adenoma	0.693	0.433	0.255	0.464	Staphylococcales	0.030	-0.038	0.0215 < Staphylococcales <= 0.0352
7	Adenoma	0.667	0.338	0.357	0.345	Staphylococcales	0.065	0.104	0.0352 < Staphylococcales
7	Adenoma	0.667	0.338	0.357	0.345	Enterobacterales	0.000	-0.071	Enterobacterales <= 4.03e-05
7	Adenoma	0.667	0.338	0.357	0.345	Pseudomonadales	0.000	-0.048	Pseudomonadales <= 2.25e-05
7	Adenoma	0.667	0.338	0.357	0.345	Corynebacteriales	0.003	0.046	0.001784 < Corynebacteriales
7	Adenoma	0.667	0.338	0.357	0.345	Actinomycetales	0.006	-0.042	0.00601 < Actinomycetales <= 0.01363
10	Adenoma	0.777	0.498	0.248	0.540	Staphylococcales	0.083	0.110	0.0352 < Staphylococcales
10	Adenoma	0.777	0.498	0.248	0.540	Actinomycetales	0.025	0.096	0.02245 < Actinomycetales
10	Adenoma	0.777	0.498	0.248	0.540	Enterobacterales	0.000	0.070	4.03e-05 < Enterobacterales
10	Adenoma	0.777	0.498	0.248	0.540	Clostridia_UCG.014	0.022	0.068	0.01058 < Clostridia_UCG.014

10	Adenoma	0.777	0.498	0.248	0.540	Pseudomonadales	0.000	-0.051	Pseudomonadales <= 2.25e-05
11	Adenoma	0.597	0.370	0.243	0.495	Actinomycetales	0.023	0.110	0.02245 < Actinomycetales
11	Adenoma	0.597	0.370	0.243	0.495	Clostridia_UCG.014	0.012	0.068	0.01058 < Clostridia_UCG.014
11	Adenoma	0.597	0.370	0.243	0.495	Enterobacterales	0.000	0.065	4.03e-05 < Enterobacterales
11	Adenoma	0.597	0.370	0.243	0.495	Corynebacteriales	0.002	0.050	0.001784 < Corynebacteriales
11	Adenoma	0.597	0.370	0.243	0.495	Staphylococcales	0.016	-0.042	0.0129 < Staphylococcales <= 0.0215
14	Adenocarcinoma	0.840	0.194	0.664	0.708	Enterobacterales	0.000	-0.068	4.03e-05 < Enterobacterales
14	Adenocarcinoma	0.840	0.194	0.664	0.708	Pseudomonadales	0.000	0.049	Pseudomonadales <= 2.25e-05
14	Adenocarcinoma	0.840	0.194	0.664	0.708	Clostridia_UCG.014	0.001	0.055	Clostridia_UCG.014 <= 0.00172
14	Adenocarcinoma	0.840	0.194	0.664	0.708	Staphylococcales	0.020	0.039	0.0129 < Staphylococcales <= 0.0215
14	Adenocarcinoma	0.840	0.194	0.664	0.708	Burkholderiales	0.005	-0.032	Burkholderiales <= 0.0108
20	Adenocarcinoma	0.893	0.202	0.608	0.752	Enterobacterales	0.000	0.066	Enterobacterales <= 4.03e-05
20	Adenocarcinoma	0.893	0.202	0.608	0.752	Pseudomonadales	0.000	0.051	Pseudomonadales <= 2.25e-05
20	Adenocarcinoma	0.893	0.202	0.608	0.752	Corynebacteriales	0.006	-0.051	0.001784 < Corynebacteriales
20	Adenocarcinoma	0.893	0.202	0.608	0.752	Actinomycetales	0.005	0.045	Actinomycetales <= 0.00601
20	Adenocarcinoma	0.893	0.202	0.608	0.752	Staphylococcales	0.021	0.035	0.0129 < Staphylococcales <= 0.0215
21	Adenocarcinoma	0.891	0.206	0.643	0.687	Clostridia_UCG.014	0.021	-0.072	0.01058 < Clostridia_UCG.014
21	Adenocarcinoma	0.891	0.206	0.643	0.687	Enterobacterales	0.000	0.065	Enterobacterales <= 4.03e-05
21	Adenocarcinoma	0.891	0.206	0.643	0.687	Actinomycetales	0.006	0.045	Actinomycetales <= 0.00601
21	Adenocarcinoma	0.891	0.206	0.643	0.687	Staphylococcales	0.007	0.041	Staphylococcales <= 0.0129
21	Adenocarcinoma	0.891	0.206	0.643	0.687	Veillonellales.Selenomonadales	0.205	-0.035	0.198 < Veillonellales.Selenomonadales
22	Adenocarcinoma	0.895	0.300	0.652	0.672	Pseudomonadales	0.000	-0.099	2.38e-04 < Pseudomonadales
22	Adenocarcinoma	0.895	0.300	0.652	0.672	Enterobacterales	0.000	0.068	Enterobacterales <= 4.03e-05
22	Adenocarcinoma	0.895	0.300	0.652	0.672	Actinomycetales	0.005	0.047	Actinomycetales <= 0.00601
22	Adenocarcinoma	0.895	0.300	0.652	0.672	Clostridia_UCG.014	0.004	0.045	0.00172 < Clostridia_UCG.014 <= 0.00404
22	Adenocarcinoma	0.895	0.300	0.652	0.672	Corynebacteriales	0.003	-0.042	0.001784 < Corynebacteriales
23	Adenocarcinoma	0.823	0.338	0.616	0.720	Actinomycetales	0.046	-0.104	0.02245 < Actinomycetales
23	Adenocarcinoma	0.823	0.338	0.616	0.720	Enterobacterales	0.000	0.070	Enterobacterales <= 4.03e-05
23	Adenocarcinoma	0.823	0.338	0.616	0.720	Pseudomonadales	0.000	0.054	Pseudomonadales <= 2.25e-05
23	Adenocarcinoma	0.823	0.338	0.616	0.720	Clostridia_UCG.014	0.002	0.046	0.00172 < Clostridia_UCG.014 <= 0.00404
23	Adenocarcinoma	0.823	0.338	0.616	0.720	Staphylococcales	0.008	0.037	Staphylococcales <= 0.0129
24	Adenocarcinoma	0.942	0.118	0.622	0.749	Enterobacterales	0.000	0.067	Enterobacterales <= 4.03e-05
24	Adenocarcinoma	0.942	0.118	0.622	0.749	Staphylococcales	0.011	0.037	Staphylococcales <= 0.0129
24	Adenocarcinoma	0.942	0.118	0.622	0.749	Pseudomonadales	0.000	0.027	2.25e-05 < Pseudomonadales <= 2.38e-04
24	Adenocarcinoma	0.942	0.118	0.622	0.749	Clostridia_UCG.014	0.005	-0.027	0.00404 < Clostridia_UCG.014 <= 0.01058
24	Adenocarcinoma	0.942	0.118	0.622	0.749	Bifidobacteriales	0.000	0.022	Bifidobacteriales <= 0.000147
25	Adenocarcinoma	0.868	0.216	0.570	0.817	Enterobacterales	0.000	0.069	Enterobacterales <= 4.03e-05
25	Adenocarcinoma	0.868	0.216	0.570	0.817	Pseudomonadales	0.000	0.057	Pseudomonadales <= 2.25e-05
25	Adenocarcinoma	0.868	0.216	0.570	0.817	Clostridia_UCG.014	0.000	0.061	Clostridia_UCG.014 <= 0.00172
25	Adenocarcinoma	0.868	0.216	0.570	0.817	Staphylococcales	0.014	0.040	0.0129 < Staphylococcales <= 0.0215
25	Adenocarcinoma	0.868	0.216	0.570	0.817	Bifidobacteriales	0.000	0.018	Bifidobacteriales <= 0.000147
26	Adenocarcinoma	0.852	0.265	0.604	0.751	Clostridia_UCG.014	0.014	-0.070	0.01058 < Clostridia_UCG.014

26	Adenocarcinoma	0.852	0.265	0.604	0.751	Pseudomonadales	0.000	0.061	Pseudomonadales <= 2.25e-05
26	Adenocarcinoma	0.852	0.265	0.604	0.751	Enterobacterales	0.000	0.065	Enterobacterales <= 4.03e-05
26	Adenocarcinoma	0.852	0.265	0.604	0.751	Actinomycetales	0.003	0.049	Actinomycetales <= 0.00601
26	Adenocarcinoma	0.852	0.265	0.604	0.751	Staphylococcales	0.004	0.042	Staphylococcales <= 0.0129
29	Adenocarcinoma	0.946	0.372	0.623	0.726	Staphylococcales	0.038	-0.112	0.0352 < Staphylococcales
29	Adenocarcinoma	0.946	0.372	0.623	0.726	Enterobacterales	0.000	0.061	Enterobacterales <= 4.03e-05
29	Adenocarcinoma	0.946	0.372	0.623	0.726	Pseudomonadales	0.000	0.051	Pseudomonadales <= 2.25e-05
29	Adenocarcinoma	0.946	0.372	0.623	0.726	Clostridia_UCG.014	0.001	0.053	Clostridia_UCG.014 <= 0.00172
29	Adenocarcinoma	0.946	0.372	0.623	0.726	Actinomycetales	0.006	0.049	Actinomycetales <= 0.00601
32	Adenocarcinoma	0.948	0.172	0.598	0.770	Enterobacterales	0.000	0.072	Enterobacterales <= 4.03e-05
32	Adenocarcinoma	0.948	0.172	0.598	0.770	Pseudomonadales	0.000	0.055	Pseudomonadales <= 2.25e-05
32	Adenocarcinoma	0.948	0.172	0.598	0.770	Actinomycetales	0.011	0.039	0.00601 < Actinomycetales <= 0.01363
32	Adenocarcinoma	0.948	0.172	0.598	0.770	Staphylococcales	0.033	0.035	0.0215 < Staphylococcales <= 0.0352
32	Adenocarcinoma	0.948	0.172	0.598	0.770	Clostridia_UCG.014	0.005	-0.029	0.00404 < Clostridia_UCG.014 <= 0.01058
33	Adenocarcinoma	0.753	0.324	0.681	0.590	Pseudomonadales	0.000	-0.094	2.38e-04 < Pseudomonadales
33	Adenocarcinoma	0.753	0.324	0.681	0.590	Enterobacterales	0.000	0.070	Enterobacterales <= 4.03e-05
33	Adenocarcinoma	0.753	0.324	0.681	0.590	Clostridia_UCG.014	0.013	-0.071	0.01058 < Clostridia_UCG.014
33	Adenocarcinoma	0.753	0.324	0.681	0.590	Staphylococcales	0.021	0.041	0.0129 < Staphylococcales <= 0.0215
33	Adenocarcinoma	0.753	0.324	0.681	0.590	Veillonellales.Selenomonadales	0.207	-0.037	0.198 < Veillonellales.Selenomonadales
34	Adenocarcinoma	0.937	0.204	0.577	0.835	Enterobacterales	0.000	0.070	Enterobacterales <= 4.03e-05
34	Adenocarcinoma	0.937	0.204	0.577	0.835	Pseudomonadales	0.000	0.053	Pseudomonadales <= 2.25e-05
34	Adenocarcinoma	0.937	0.204	0.577	0.835	Actinomycetales	0.003	0.051	Actinomycetales <= 0.00601
34	Adenocarcinoma	0.937	0.204	0.577	0.835	Clostridia_UCG.014	0.002	0.043	0.00172 < Clostridia_UCG.014 <= 0.00404
34	Adenocarcinoma	0.937	0.204	0.577	0.835	Staphylococcales	0.030	0.040	0.0215 < Staphylococcales <= 0.0352
38	Adenocarcinoma	0.866	0.231	0.613	0.742	Clostridia_UCG.014	0.030	-0.070	0.01058 < Clostridia_UCG.014
38	Adenocarcinoma	0.866	0.231	0.613	0.742	Enterobacterales	0.000	0.067	Enterobacterales <= 4.03e-05
38	Adenocarcinoma	0.866	0.231	0.613	0.742	Pseudomonadales	0.000	0.053	Pseudomonadales <= 2.25e-05
38	Adenocarcinoma	0.866	0.231	0.613	0.742	Actinomycetales	0.013	0.044	0.00601 < Actinomycetales <= 0.01363
38	Adenocarcinoma	0.866	0.231	0.613	0.742	Staphylococcales	0.024	0.035	0.0215 < Staphylococcales <= 0.0352
39	Adenocarcinoma	0.833	0.216	0.595	0.785	Enterobacterales	0.000	0.070	Enterobacterales <= 4.03e-05
39	Adenocarcinoma	0.833	0.216	0.595	0.785	Clostridia_UCG.014	0.000	0.059	Clostridia_UCG.014 <= 0.00172
39	Adenocarcinoma	0.833	0.216	0.595	0.785	Pseudomonadales	0.000	0.052	Pseudomonadales <= 2.25e-05
39	Adenocarcinoma	0.833	0.216	0.595	0.785	Actinomycetales	0.010	0.047	0.00601 < Actinomycetales <= 0.01363
39	Adenocarcinoma	0.833	0.216	0.595	0.785	Veillonellales.Selenomonadales	0.300	-0.038	0.198 < Veillonellales.Selenomonadales
41	Adenocarcinoma	0.830	0.316	0.636	0.726	Actinomycetales	0.053	-0.106	0.02245 < Actinomycetales
41	Adenocarcinoma	0.830	0.316	0.636	0.726	Enterobacterales	0.000	0.070	Enterobacterales <= 4.03e-05
41	Adenocarcinoma	0.830	0.316	0.636	0.726	Clostridia_UCG.014	0.002	0.054	Clostridia_UCG.014 <= 0.00172
41	Adenocarcinoma	0.830	0.316	0.636	0.726	Staphylococcales	0.018	0.042	0.0129 < Staphylococcales <= 0.0215
41	Adenocarcinoma	0.830	0.316	0.636	0.726	Pseudomonadales	0.000	0.031	2.25e-05 < Pseudomonadales <= 2.38e-04
46	Adenocarcinoma	0.894	0.431	0.693	0.560	Staphylococcales	0.037	-0.104	0.0352 < Staphylococcales
46	Adenocarcinoma	0.894	0.431	0.693	0.560	Pseudomonadales	0.000	-0.093	2.38e-04 < Pseudomonadales
46	Adenocarcinoma	0.894	0.431	0.693	0.560	Enterobacterales	0.000	0.069	Enterobacterales <= 4.03e-05

46	Adenocarcinoma	0.894	0.431	0.693	0.560	Corynebacteriales	0.002	-0.047	0.001784 < Corynebacteriales
46	Adenocarcinoma	0.894	0.431	0.693	0.560	Clostridia_UCG.014	0.002	0.043	0.00172 < Clostridia_UCG.014 <= 0.00404
47	Adenocarcinoma	0.655	0.440	0.763	0.485	Actinomycetales	0.032	-0.104	0.02245 < Actinomycetales
47	Adenocarcinoma	0.655	0.440	0.763	0.485	Pseudomonadales	0.001	-0.102	2.38e-04 < Pseudomonadales
47	Adenocarcinoma	0.655	0.440	0.763	0.485	Enterobacterales	0.004	-0.073	4.03e-05 < Enterobacterales
47	Adenocarcinoma	0.655	0.440	0.763	0.485	Corynebacteriales	0.004	-0.043	0.001784 < Corynebacteriales
47	Adenocarcinoma	0.655	0.440	0.763	0.485	Clostridia_UCG.014	0.002	0.045	0.00172 < Clostridia_UCG.014 <= 0.00404
49	Adenocarcinoma	0.858	0.343	0.699	0.659	Staphylococcales	0.035	-0.105	0.0352 < Staphylococcales
49	Adenocarcinoma	0.858	0.343	0.699	0.659	Enterobacterales	0.000	-0.068	4.03e-05 < Enterobacterales
49	Adenocarcinoma	0.858	0.343	0.699	0.659	Clostridia_UCG.014	0.000	0.059	Clostridia_UCG.014 <= 0.00172
49	Adenocarcinoma	0.858	0.343	0.699	0.659	Actinomycetales	0.001	0.046	Actinomycetales <= 0.00601
49	Adenocarcinoma	0.858	0.343	0.699	0.659	Pseudomonadales	0.000	0.029	2.25e-05 < Pseudomonadales <= 2.38e-04
51	Adenocarcinoma	0.805	0.343	0.651	0.675	Staphylococcales	0.043	-0.106	0.0352 < Staphylococcales
51	Adenocarcinoma	0.805	0.343	0.651	0.675	Enterobacterales	0.000	0.067	Enterobacterales <= 4.03e-05
51	Adenocarcinoma	0.805	0.343	0.651	0.675	Clostridia_UCG.014	0.000	0.063	Clostridia_UCG.014 <= 0.00172
51	Adenocarcinoma	0.805	0.343	0.651	0.675	Veillonellales.Selenomonadales	0.199	-0.042	0.198 < Veillonellales.Selenomonadales
51	Adenocarcinoma	0.805	0.343	0.651	0.675	Actinomycetales	0.010	0.042	0.00601 < Actinomycetales <= 0.01363
52	Adenocarcinoma	0.930	0.187	0.597	0.774	Enterobacterales	0.000	0.068	Enterobacterales <= 4.03e-05
52	Adenocarcinoma	0.930	0.187	0.597	0.774	Pseudomonadales	0.000	0.058	Pseudomonadales <= 2.25e-05
52	Adenocarcinoma	0.930	0.187	0.597	0.774	Actinomycetales	0.006	0.045	0.00601 < Actinomycetales <= 0.01363
52	Adenocarcinoma	0.930	0.187	0.597	0.774	Staphylococcales	0.034	0.041	0.0215 < Staphylococcales <= 0.0352
52	Adenocarcinoma	0.930	0.187	0.597	0.774	Clostridia_UCG.014	0.006	-0.034	0.00404 < Clostridia_UCG.014 <= 0.01058

Table ML-C-5: The LIME Result for Test Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
1	Adenocarcinoma	0.778	0.426	0.694	0.561	Staphylococcales	0.049	-0.107	0.0352 < Staphylococcales
1	Adenocarcinoma	0.778	0.426	0.694	0.561	Pseudomonadales	0.000	-0.098	2.38e-04 < Pseudomonadales
1	Adenocarcinoma	0.778	0.426	0.694	0.561	Enterobacterales	0.000	0.067	Enterobacterales <= 4.03e-05
1	Adenocarcinoma	0.778	0.426	0.694	0.561	Actinomycetales	0.013	0.048	0.00601 < Actinomycetales <= 0.01363
1	Adenocarcinoma	0.778	0.426	0.694	0.561	Corynebacteriales	0.004	-0.043	0.001784 < Corynebacteriales
4	Adenocarcinoma	0.645	0.319	0.748	0.529	Pseudomonadales	0.001	-0.101	2.38e-04 < Pseudomonadales
4	Adenocarcinoma	0.645	0.319	0.748	0.529	Clostridia_UCG.014	0.014	-0.069	0.01058 < Clostridia_UCG.014
4	Adenocarcinoma	0.645	0.319	0.748	0.529	Enterobacterales	0.000	-0.067	4.03e-05 < Enterobacterales
4	Adenocarcinoma	0.645	0.319	0.748	0.529	Staphylococcales	0.013	0.045	0.0129 < Staphylococcales <= 0.0215
4	Adenocarcinoma	0.645	0.319	0.748	0.529	Campylobacterales	0.019	-0.027	0.00743 < Campylobacterales
5	Adenocarcinoma	0.718	0.165	0.663	0.767	Enterobacterales	0.000	-0.067	4.03e-05 < Enterobacterales
5	Adenocarcinoma	0.718	0.165	0.663	0.767	Actinomycetales	0.006	0.048	Actinomycetales <= 0.00601
5	Adenocarcinoma	0.718	0.165	0.663	0.767	Clostridia_UCG.014	0.002	0.048	0.00172 < Clostridia_UCG.014 <= 0.00404
5	Adenocarcinoma	0.718	0.165	0.663	0.767	Pseudomonadales	0.000	0.038	2.25e-05 < Pseudomonadales <= 2.38e-04
5	Adenocarcinoma	0.718	0.165	0.663	0.767	Staphylococcales	0.035	0.037	0.0215 < Staphylococcales <= 0.0352
8	Adenocarcinoma	0.667	0.385	0.652	0.636	Actinomycetales	0.036	-0.107	0.02245 < Actinomycetales

8	Adenocarcinoma	0.667	0.385	0.652	0.636	Clostridia_UCG.014	0.016	-0.069	0.01058 < Clostridia_UCG.014
8	Adenocarcinoma	0.667	0.385	0.652	0.636	Enterobacterales	0.000	0.066	Enterobacterales <= 4.03e-05
8	Adenocarcinoma	0.667	0.385	0.652	0.636	Pseudomonadales	0.000	0.054	Pseudomonadales <= 2.25e-05
8	Adenocarcinoma	0.667	0.385	0.652	0.636	Staphylococcales	0.019	0.039	0.0129 < Staphylococcales <= 0.0215
9	Adenocarcinoma	0.541	0.283	0.677	0.603	Pseudomonadales	0.001	-0.094	2.38e-04 < Pseudomonadales
9	Adenocarcinoma	0.541	0.283	0.677	0.603	Enterobacterales	0.000	0.071	Enterobacterales <= 4.03e-05
9	Adenocarcinoma	0.541	0.283	0.677	0.603	Corynebacteriales	0.005	-0.051	0.001784 < Corynebacteriales
9	Adenocarcinoma	0.541	0.283	0.677	0.603	Staphylococcales	0.034	0.036	0.0215 < Staphylococcales <= 0.0352
9	Adenocarcinoma	0.541	0.283	0.677	0.603	Clostridia_UCG.014	0.010	-0.036	0.00404 < Clostridia_UCG.014 <= 0.01058
12	Adenocarcinoma	0.863	0.175	0.618	0.778	Enterobacterales	0.000	0.064	Enterobacterales <= 4.03e-05
12	Adenocarcinoma	0.863	0.175	0.618	0.778	Actinomycetales	0.009	0.049	0.00601 < Actinomycetales <= 0.01363
12	Adenocarcinoma	0.863	0.175	0.618	0.778	Clostridia_UCG.014	0.003	0.045	0.00172 < Clostridia_UCG.014 <= 0.00404
12	Adenocarcinoma	0.863	0.175	0.618	0.778	Corynebacteriales	0.002	-0.042	0.001784 < Corynebacteriales
12	Adenocarcinoma	0.863	0.175	0.618	0.778	Staphylococcales	0.020	0.044	0.0129 < Staphylococcales <= 0.0215
13	Adenocarcinoma	0.546	0.292	0.714	0.620	Pseudomonadales	0.001	-0.094	2.38e-04 < Pseudomonadales
13	Adenocarcinoma	0.546	0.292	0.714	0.620	Enterobacterales	0.000	-0.064	4.03e-05 < Enterobacterales
13	Adenocarcinoma	0.546	0.292	0.714	0.620	Clostridia_UCG.014	0.000	0.058	Clostridia_UCG.014 <= 0.00172
13	Adenocarcinoma	0.546	0.292	0.714	0.620	Actinomycetales	0.010	0.044	0.00601 < Actinomycetales <= 0.01363
13	Adenocarcinoma	0.546	0.292	0.714	0.620	Veillonellales.Selenomonadales	0.238	-0.036	0.198 < Veillonellales.Selenomonadales
15	Adenocarcinoma	0.670	0.402	0.669	0.614	Staphylococcales	0.039	-0.101	0.0352 < Staphylococcales
15	Adenocarcinoma	0.670	0.402	0.669	0.614	Pseudomonadales	0.001	-0.098	2.38e-04 < Pseudomonadales
15	Adenocarcinoma	0.670	0.402	0.669	0.614	Enterobacterales	0.000	0.073	Enterobacterales <= 4.03e-05
15	Adenocarcinoma	0.670	0.402	0.669	0.614	Clostridia_UCG.014	0.003	0.045	0.00172 < Clostridia_UCG.014 <= 0.00404
15	Adenocarcinoma	0.670	0.402	0.669	0.614	Corynebacteriales	0.001	0.028	0.000195 < Corynebacteriales <= 0.000822
16	Adenocarcinoma	0.605	0.479	0.698	0.547	Staphylococcales	0.056	-0.108	0.0352 < Staphylococcales
16	Adenocarcinoma	0.605	0.479	0.698	0.547	Actinomycetales	0.040	-0.104	0.02245 < Actinomycetales
16	Adenocarcinoma	0.605	0.479	0.698	0.547	Enterobacterales	0.000	0.069	Enterobacterales <= 4.03e-05
16	Adenocarcinoma	0.605	0.479	0.698	0.547	Corynebacteriales	0.002	-0.053	0.001784 < Corynebacteriales
16	Adenocarcinoma	0.605	0.479	0.698	0.547	Clostridia_UCG.014	0.003	0.045	0.00172 < Clostridia_UCG.014 <= 0.00404
17	Adenocarcinoma	0.839	0.396	0.608	0.729	Staphylococcales	0.039	-0.111	0.0352 < Staphylococcales
17	Adenocarcinoma	0.839	0.396	0.608	0.729	Enterobacterales	0.000	0.068	Enterobacterales <= 4.03e-05
17	Adenocarcinoma	0.839	0.396	0.608	0.729	Pseudomonadales	0.000	0.058	Pseudomonadales <= 2.25e-05
17	Adenocarcinoma	0.839	0.396	0.608	0.729	Clostridia_UCG.014	0.001	0.058	Clostridia_UCG.014 <= 0.00172
17	Adenocarcinoma	0.839	0.396	0.608	0.729	Actinomycetales	0.002	0.048	Actinomycetales <= 0.00601
18	Adenocarcinoma	0.658	0.190	0.680	0.709	Enterobacterales	0.000	-0.067	4.03e-05 < Enterobacterales
18	Adenocarcinoma	0.658	0.190	0.680	0.709	Clostridia_UCG.014	0.000	0.057	Clostridia_UCG.014 <= 0.00172
18	Adenocarcinoma	0.658	0.190	0.680	0.709	Actinomycetales	0.003	0.044	Actinomycetales <= 0.00601
18	Adenocarcinoma	0.658	0.190	0.680	0.709	Veillonellales.Selenomonadales	0.232	-0.039	0.198 < Veillonellales.Selenomonadales
18	Adenocarcinoma	0.658	0.190	0.680	0.709	Staphylococcales	0.002	0.034	Staphylococcales <= 0.0129
19	Adenocarcinoma	0.535	0.117	0.636	0.727	Enterobacterales	0.000	0.067	Enterobacterales <= 4.03e-05
19	Adenocarcinoma	0.535	0.117	0.636	0.727	Clostridia_UCG.014	0.004	0.049	0.00172 < Clostridia_UCG.014 <= 0.00404
19	Adenocarcinoma	0.535	0.117	0.636	0.727	Burkholderiales	0.000	-0.029	Burkholderiales <= 0.0108

19	Adenocarcinoma	0.535	0.117	0.636	0.727	Actinomycetales	0.016	0.019	0.01363 < Actinomycetales <= 0.02245
19	Adenocarcinoma	0.535	0.117	0.636	0.727	Campylobacterales	0.008	-0.017	0.00743 < Campylobacterales
27	Adenocarcinoma	0.759	0.230	0.697	0.676	Enterobacterales	0.001	-0.078	4.03e-05 < Enterobacterales
27	Adenocarcinoma	0.759	0.230	0.697	0.676	Clostridia_UCG.014	0.016	-0.068	0.01058 < Clostridia_UCG.014
27	Adenocarcinoma	0.759	0.230	0.697	0.676	Actinomycetales	0.003	0.049	Actinomycetales <= 0.00601
27	Adenocarcinoma	0.759	0.230	0.697	0.676	Staphylococcales	0.030	0.044	0.0215 < Staphylococcales <= 0.0352
27	Adenocarcinoma	0.759	0.230	0.697	0.676	Pseudomonadales	0.000	0.032	2.25e-05 < Pseudomonadales <= 2.38e-04
28	Adenocarcinoma	0.837	0.150	0.637	0.717	Enterobacterales	0.000	0.069	Enterobacterales <= 4.03e-05
28	Adenocarcinoma	0.837	0.150	0.637	0.717	Staphylococcales	0.014	0.043	0.0129 < Staphylococcales <= 0.0215
28	Adenocarcinoma	0.837	0.150	0.637	0.717	Veillonellales.Selenomonadales	0.288	-0.038	0.198 < Veillonellales.Selenomonadales
28	Adenocarcinoma	0.837	0.150	0.637	0.717	Pseudomonadales	0.000	0.038	2.25e-05 < Pseudomonadales <= 2.38e-04
28	Adenocarcinoma	0.837	0.150	0.637	0.717	Clostridia_UCG.014	0.004	-0.032	0.00404 < Clostridia_UCG.014 <= 0.01058
30	Adenocarcinoma	0.682	0.267	0.654	0.672	Pseudomonadales	0.001	-0.099	2.38e-04 < Pseudomonadales
30	Adenocarcinoma	0.682	0.267	0.654	0.672	Enterobacterales	0.000	0.069	Enterobacterales <= 4.03e-05
30	Adenocarcinoma	0.682	0.267	0.654	0.672	Actinomycetales	0.011	0.041	0.00601 < Actinomycetales <= 0.01363
30	Adenocarcinoma	0.682	0.267	0.654	0.672	Staphylococcales	0.019	0.041	0.0129 < Staphylococcales <= 0.0215
30	Adenocarcinoma	0.682	0.267	0.654	0.672	Clostridia_UCG.014	0.007	-0.034	0.00404 < Clostridia_UCG.014 <= 0.01058
31	Adenocarcinoma	0.817	0.206	0.684	0.703	Enterobacterales	0.000	-0.070	4.03e-05 < Enterobacterales
31	Adenocarcinoma	0.817	0.206	0.684	0.703	Clostridia_UCG.014	0.001	0.058	Clostridia_UCG.014 <= 0.00172
31	Adenocarcinoma	0.817	0.206	0.684	0.703	Corynebacteriales	0.004	-0.054	0.001784 < Corynebacteriales
31	Adenocarcinoma	0.817	0.206	0.684	0.703	Actinomycetales	0.005	0.047	Actinomycetales <= 0.00601
31	Adenocarcinoma	0.817	0.206	0.684	0.703	Staphylococcales	0.024	0.037	0.0215 < Staphylococcales <= 0.0352
35	Adenocarcinoma	0.925	0.223	0.574	0.843	Enterobacterales	0.000	0.067	Enterobacterales <= 4.03e-05
35	Adenocarcinoma	0.925	0.223	0.574	0.843	Pseudomonadales	0.000	0.054	Pseudomonadales <= 2.25e-05
35	Adenocarcinoma	0.925	0.223	0.574	0.843	Clostridia_UCG.014	0.002	0.052	Clostridia_UCG.014 <= 0.00172
35	Adenocarcinoma	0.925	0.223	0.574	0.843	Actinomycetales	0.004	0.049	Actinomycetales <= 0.00601
35	Adenocarcinoma	0.925	0.223	0.574	0.843	Staphylococcales	0.014	0.046	0.0129 < Staphylococcales <= 0.0215
36	Adenocarcinoma	0.579	0.192	0.691	0.633	Enterobacterales	0.000	-0.064	4.03e-05 < Enterobacterales
36	Adenocarcinoma	0.579	0.192	0.691	0.633	Pseudomonadales	0.000	0.052	Pseudomonadales <= 2.25e-05
36	Adenocarcinoma	0.579	0.192	0.691	0.633	Corynebacteriales	0.002	-0.048	0.001784 < Corynebacteriales
36	Adenocarcinoma	0.579	0.192	0.691	0.633	Burkholderiales	0.008	-0.038	Burkholderiales <= 0.0108
36	Adenocarcinoma	0.579	0.192	0.691	0.633	Staphylococcales	0.018	0.039	0.0129 < Staphylococcales <= 0.0215
37	Adenocarcinoma	0.954	0.188	0.591	0.833	Enterobacterales	0.000	0.069	Enterobacterales <= 4.03e-05
37	Adenocarcinoma	0.954	0.188	0.591	0.833	Clostridia_UCG.014	0.000	0.054	Clostridia_UCG.014 <= 0.00172
37	Adenocarcinoma	0.954	0.188	0.591	0.833	Actinomycetales	0.005	0.051	Actinomycetales <= 0.00601
37	Adenocarcinoma	0.954	0.188	0.591	0.833	Staphylococcales	0.008	0.037	Staphylococcales <= 0.0129
37	Adenocarcinoma	0.954	0.188	0.591	0.833	Pseudomonadales	0.000	0.030	2.25e-05 < Pseudomonadales <= 2.38e-04
40	Adenocarcinoma	0.620	0.457	0.654	0.615	Staphylococcales	0.043	-0.107	0.0352 < Staphylococcales
40	Adenocarcinoma	0.620	0.457	0.654	0.615	Actinomycetales	0.025	-0.099	0.02245 < Actinomycetales
40	Adenocarcinoma	0.620	0.457	0.654	0.615	Enterobacterales	0.000	0.072	Enterobacterales <= 4.03e-05
40	Adenocarcinoma	0.620	0.457	0.654	0.615	Pseudomonadales	0.000	0.052	Pseudomonadales <= 2.25e-05
40	Adenocarcinoma	0.620	0.457	0.654	0.615	Clostridia_UCG.014	0.002	0.044	0.00172 < Clostridia_UCG.014 <= 0.00404

42	Adenocarcinoma	0.754	0.170	0.680	0.719	Enterobacterales	0.000	-0.067	4.03e-05 < Enterobacterales
42	Adenocarcinoma	0.754	0.170	0.680	0.719	Clostridia_UCG.014	0.000	0.058	Clostridia_UCG.014 <= 0.00172
42	Adenocarcinoma	0.754	0.170	0.680	0.719	Staphylococcales	0.021	0.045	0.0129 < Staphylococcales <= 0.0215
42	Adenocarcinoma	0.754	0.170	0.680	0.719	Pseudomonadales	0.000	0.031	2.25e-05 < Pseudomonadales <= 2.38e-04
42	Adenocarcinoma	0.754	0.170	0.680	0.719	Burkholderiales	0.006	-0.028	Burkholderiales <= 0.0108
43	Adenoma	0.631	0.29205459	0.262818111	0.459204477	Actinomycetales	0.023798549	0.101038537	0.02245 < Actinomycetales
43	Adenoma	0.631	0.29205459	0.262818111	0.459204477	Enterobacterales	0.000790212	0.066266407	4.03e-05 < Enterobacterales
43	Adenoma	0.631	0.29205459	0.262818111	0.459204477	Clostridia_UCG.014	6.5851E-05	-0.049125729	Clostridia_UCG.014 <= 0.00172
43	Adenoma	0.631	0.29205459	0.262818111	0.459204477	Corynebacteriales	0.003555954	0.047728437	0.001784 < Corynebacteriales
43	Adenoma	0.631	0.29205459	0.262818111	0.459204477	Bifidobacteriales	0.00134336	0.030478713	0.000537 < Bifidobacteriales
44	Adenocarcinoma	0.847	0.355531127	0.610741104	0.736160156	Actinomycetales	0.025823633	-0.103956692	0.02245 < Actinomycetales
44	Adenocarcinoma	0.847	0.355531127	0.610741104	0.736160156	Enterobacterales	0	0.070509463	Enterobacterales <= 4.03e-05
44	Adenocarcinoma	0.847	0.355531127	0.610741104	0.736160156	Pseudomonadales	0	0.055083718	Pseudomonadales <= 2.25e-05
44	Adenocarcinoma	0.847	0.355531127	0.610741104	0.736160156	Clostridia_UCG.014	0	0.056820402	Clostridia_UCG.014 <= 0.00172
44	Adenocarcinoma	0.847	0.355531127	0.610741104	0.736160156	Staphylococcales	0.013014227	0.046962159	0.0129 < Staphylococcales <= 0.0215
45	Adenocarcinoma	0.837	0.320877792	0.657341288	0.658760064	Staphylococcales	0.039885538	-0.105988064	0.0352 < Staphylococcales
45	Adenocarcinoma	0.837	0.320877792	0.657341288	0.658760064	Enterobacterales	0	0.067459746	Enterobacterales <= 4.03e-05
45	Adenocarcinoma	0.837	0.320877792	0.657341288	0.658760064	Actinomycetales	0.006911732	0.04996396	0.00601 < Actinomycetales <= 0.01363
45	Adenocarcinoma	0.837	0.320877792	0.657341288	0.658760064	Corynebacteriales	0.003918116	-0.047301551	0.001784 < Corynebacteriales
45	Adenocarcinoma	0.837	0.320877792	0.657341288	0.658760064	Pseudomonadales	0.000220119	0.037284686	2.25e-05 < Pseudomonadales <= 2.38e-04
48	Adenocarcinoma	0.644	0.547752678	0.800556678	0.386446259	Staphylococcales	0.044000815	-0.106943928	0.0352 < Staphylococcales
48	Adenocarcinoma	0.644	0.547752678	0.800556678	0.386446259	Actinomycetales	0.025178244	-0.102539938	0.02245 < Actinomycetales
48	Adenocarcinoma	0.644	0.547752678	0.800556678	0.386446259	Pseudomonadales	0.000325932	-0.097666558	2.38e-04 < Pseudomonadales
48	Adenocarcinoma	0.644	0.547752678	0.800556678	0.386446259	Enterobacterales	0.000325932	-0.067468493	4.03e-05 < Enterobacterales
48	Adenocarcinoma	0.644	0.547752678	0.800556678	0.386446259	Veillonellales.Selenomonadales	0.224933795	-0.039491501	0.198 < Veillonellales.Selenomonadales
50	Adenocarcinoma	0.843	0.128703628	0.671721913	0.74078637	Enterobacterales	0.000266903	-0.067551975	4.03e-05 < Enterobacterales
50	Adenocarcinoma	0.843	0.128703628	0.671721913	0.74078637	Clostridia_UCG.014	0.003775895	0.044308168	0.00172 < Clostridia_UCG.014 <= 0.00404
50	Adenocarcinoma	0.843	0.128703628	0.671721913	0.74078637	Staphylococcales	0.025143853	0.036368003	0.0215 < Staphylococcales <= 0.0352
50	Adenocarcinoma	0.843	0.128703628	0.671721913	0.74078637	Pseudomonadales	7.85009E-05	0.032156471	2.25e-05 < Pseudomonadales <= 2.38e-04
50	Adenocarcinoma	0.843	0.128703628	0.671721913	0.74078637	Burkholderiales	0.050743011	0.02378379	0.0388 < Burkholderiales <= 0.0729



Figure ML3-C-6: The LIME Result in Training dataset for Order in Saliva

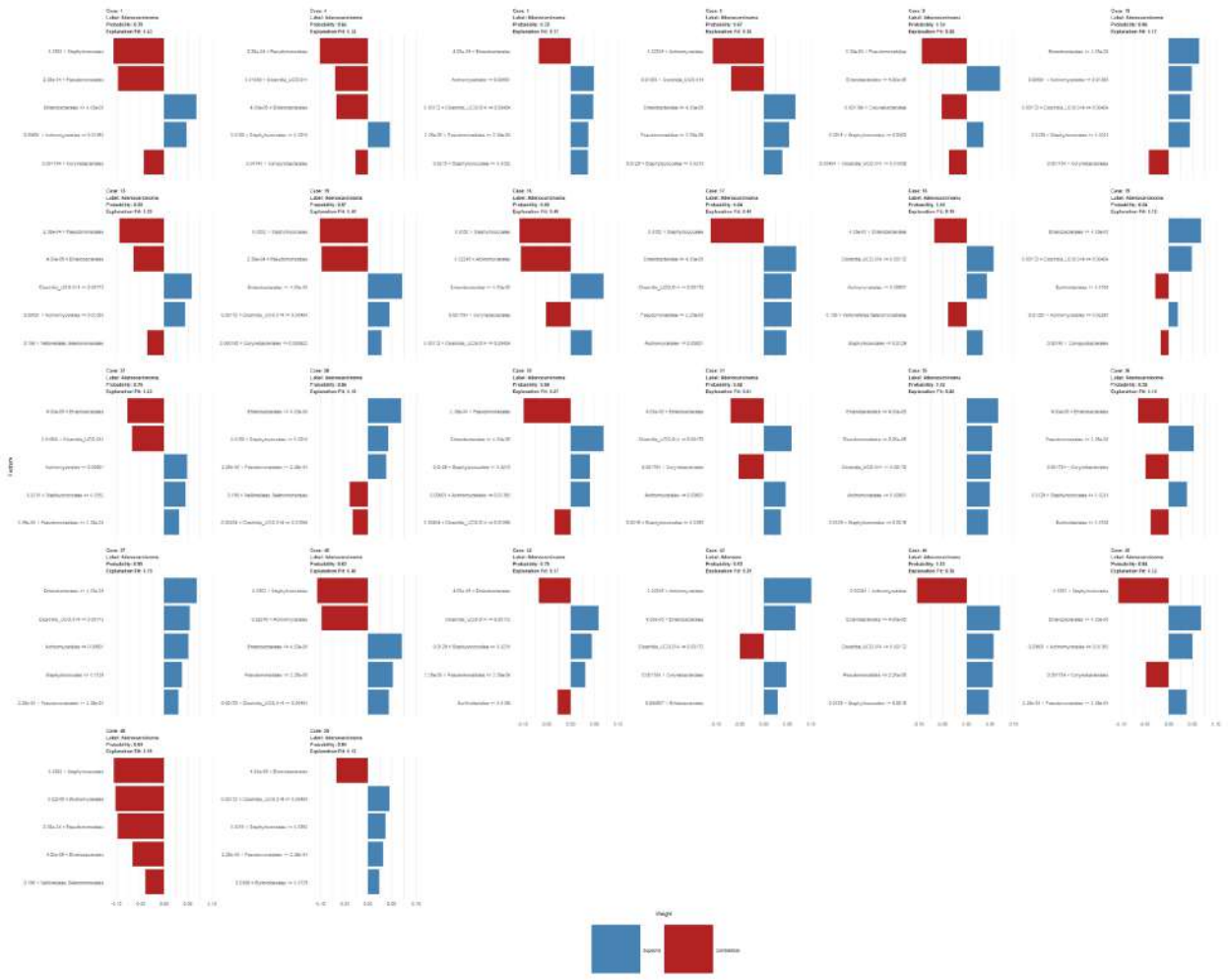


Figure ML3-C-7: The LIME Result in Test dataset for Order in Saliva

Table ML3-C-6: The Aggregated SHAP Values

Variable	Training Dataset				Testing Dataset			
	contribution	position	sign	cumulative	contribution	position	sign	cumulative
intercept	0.7382	38	X	0.738	0.7382	38	X	0.738
Coriobacteriales	0.0001	37	1	0.738	-0.0006	33	-1	0.736
Flavobacteriales	0.0000	36	1	0.738	0.0001	35	1	0.739
Bacteroidales	0.0000	35	-1	0.738	-0.0025	34	-1	0.737
Synergistales	0.0000	34	1	0.738	0.0008	31	1	0.733

Cardiobacteriales	0.0000	33	-1	0.738	-0.0042	32	-1	0.732
RF39	0.0000	32	1	0.738	-0.0004	37	-1	0.738
Campylobacterales	0.0000	31	1	0.738	0.0026	30	1	0.735
Veillonellales.Selenomonadales	0.0000	30	-1	0.738	0.0036	29	1	0.739
Tumor	0.0000	29	1	0.738	0.0000	28	1	0.739
Fusobacteriales	0.0000	28	1	0.738	-0.0010	27	-1	0.738
Chloroplast	0.0000	27	-1	0.738	0.0001	25	1	0.737
Spirochaetales	0.0000	26	-1	0.738	-0.0004	24	-1	0.737
Erysipelotrichales	0.0000	25	1	0.738	0.0016	36	1	0.739
Micrococcales	0.0000	24	1	0.738	0.0002	23	1	0.737
Sphingobacteriales	-0.0001	23	-1	0.738	-0.0004	22	-1	0.736
Lachnospirales	0.0000	22	1	0.738	0.0007	16	1	0.700
Corynebacteriales	-0.0001	21	-1	0.738	-0.0063	12	-1	0.699
Pseudomonadales	0.0001	20	1	0.738	-0.0220	20	-1	0.715
Burkholderiales	-0.0001	19	-1	0.738	0.0027	14	1	0.705
Enterobacterales	0.0000	18	1	0.738	-0.0078	19	-1	0.707
Gracilibacteria	0.0000	17	1	0.738	0.0000	9	1	0.697
Saccharimonadales	-0.0001	16	-1	0.738	-0.0007	18	-1	0.707
Staphylococcales	0.0001	15	1	0.738	-0.0070	17	-1	0.700
Mycoplasmatales	0.0001	14	1	0.738	0.0007	21	1	0.737
Pasteurellales	0.0000	13	-1	0.738	0.0021	15	1	0.702
Clostridia_UCG.014	-0.0001	12	-1	0.738	0.0138	6	1	0.712
Bacillales	0.0000	11	1	0.738	0.0008	13	1	0.706
Bifidobacteriales	0.0001	10	1	0.738	-0.0011	26	-1	0.737
Peptococcales	0.0000	9	1	0.738	0.0019	10	1	0.697
Clostridia_vadinBB60_group	0.0000	8	1	0.738	0.0005	8	1	0.698
Lactobacillales	-0.0001	7	-1	0.738	-0.0041	11	-1	0.695
Gender	0.0000	6	1	0.738	0.0000	7	1	0.698
Absconditabacteriales_.SR1.	0.0000	5	-1	0.738	-0.0005	5	-1	0.711
Actinomycetales	0.0000	4	-1	0.738	0.0025	4	1	0.714
JGI_0000069.P22	0.0000	3	1	0.738	0.0006	3	1	0.714
Peptostreptococcales.Tissierellales	0.0000	2	-1	0.738	-0.0014	2	-1	0.713
prediction	0.7382	1	X	0.738	0.7128	1	X	0.713

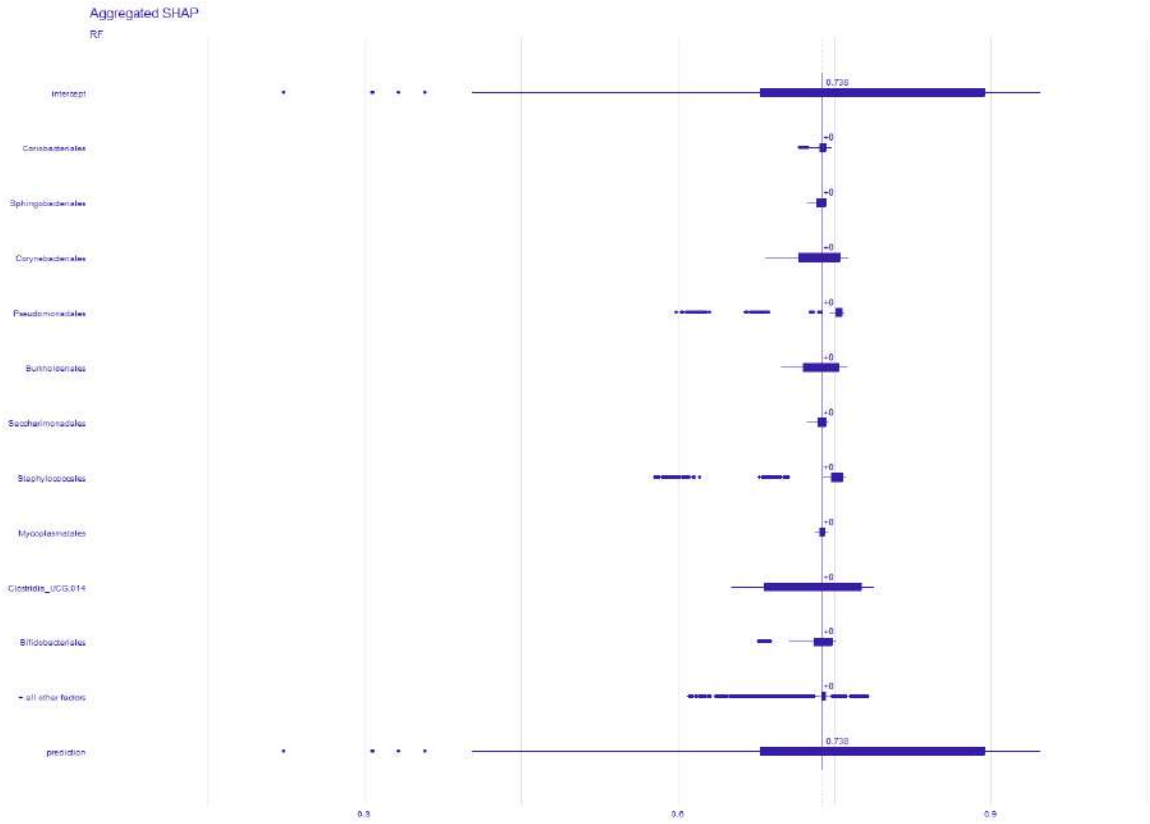


Figure ML3-C-8: The Aggregated SHAP Value in Training dataset for Order in Saliva

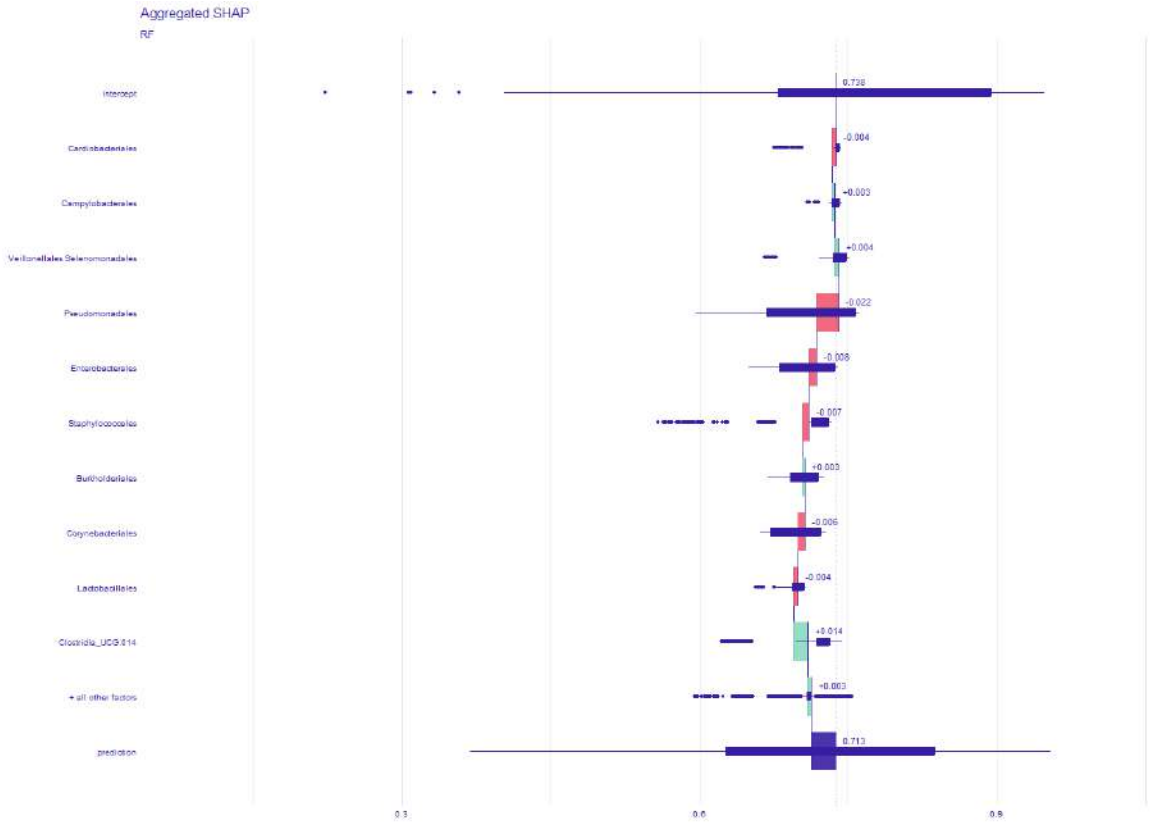


Figure ML3-C-9: The Aggregated SHAP Value in Test dataset for Order in Saliva

D- Family

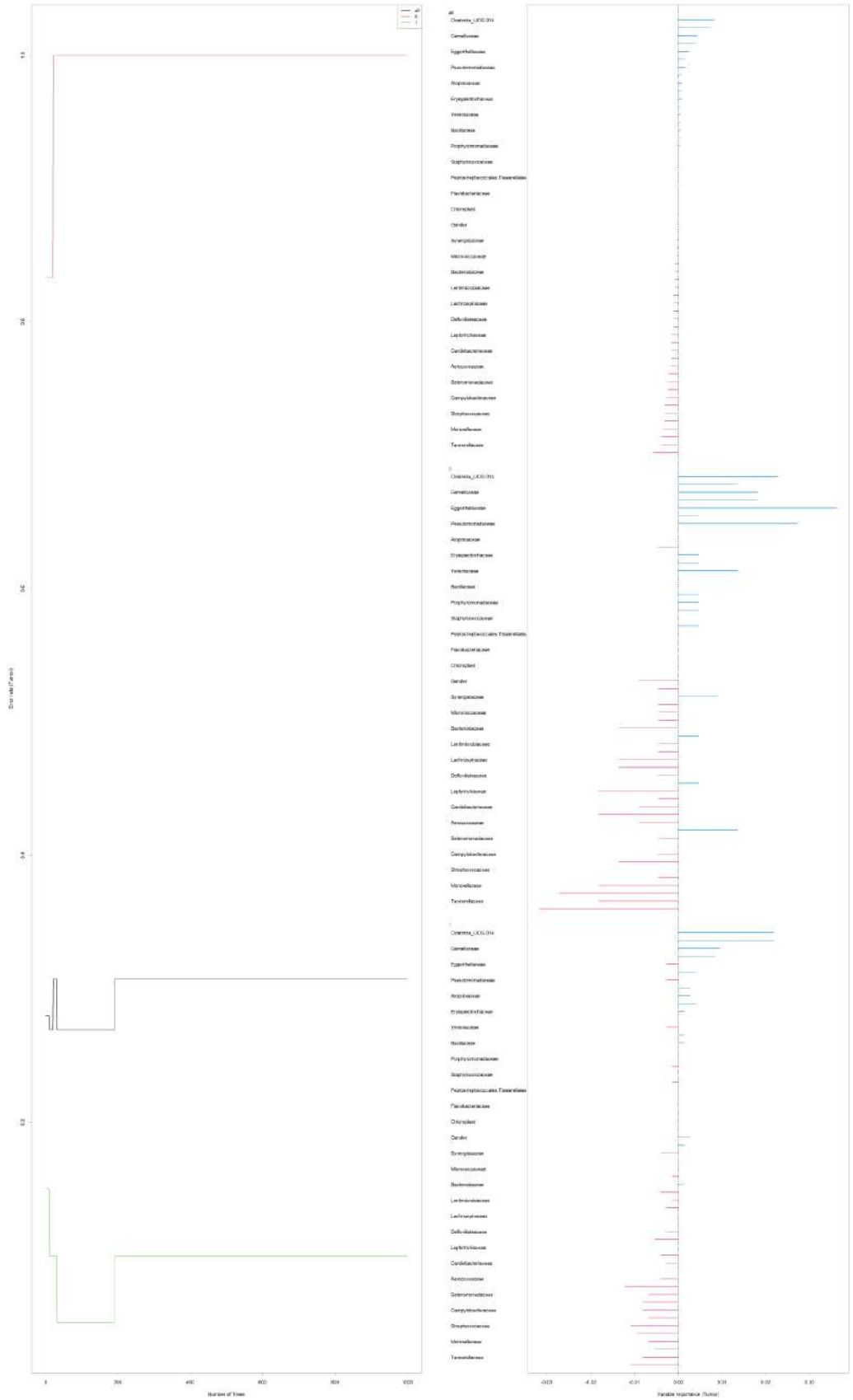
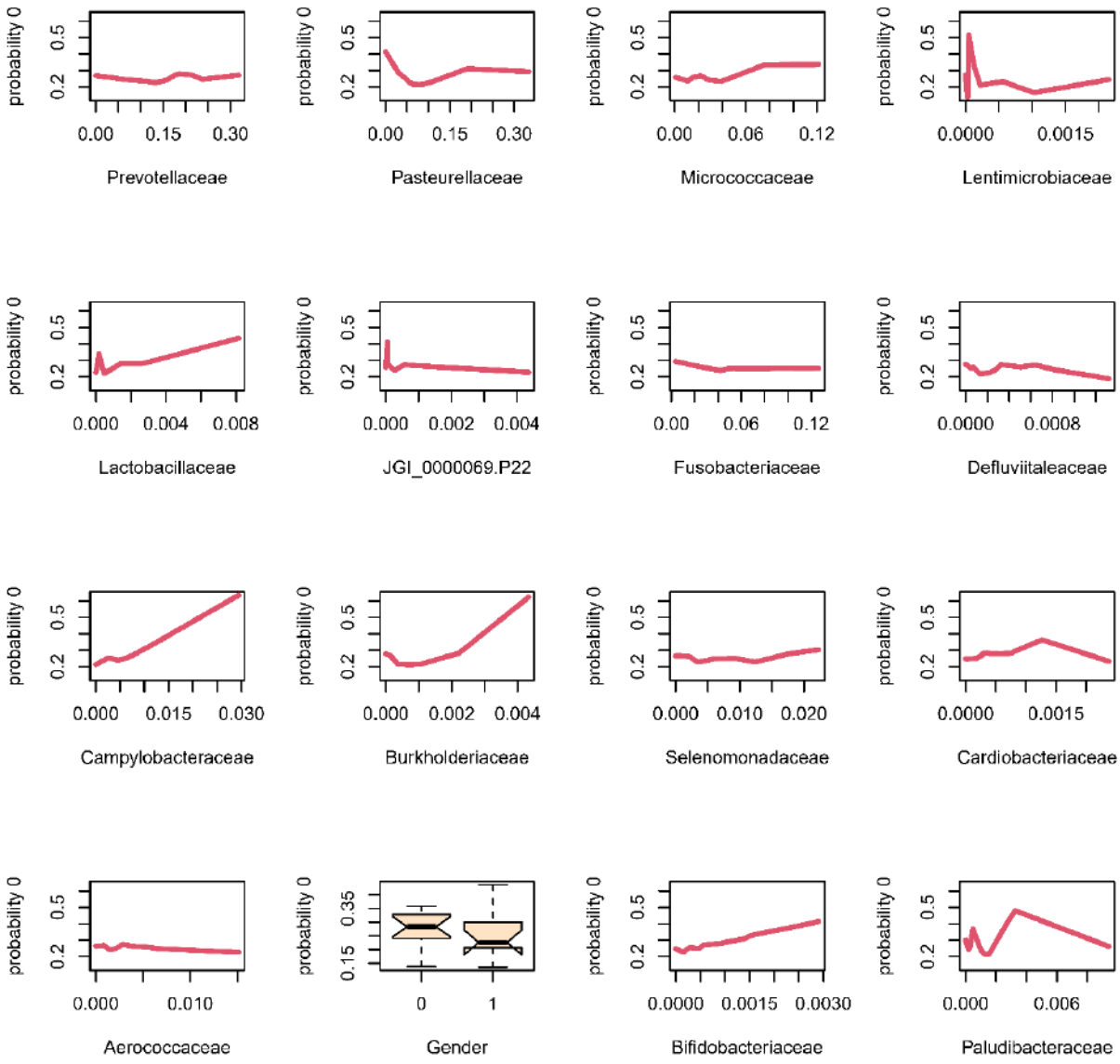
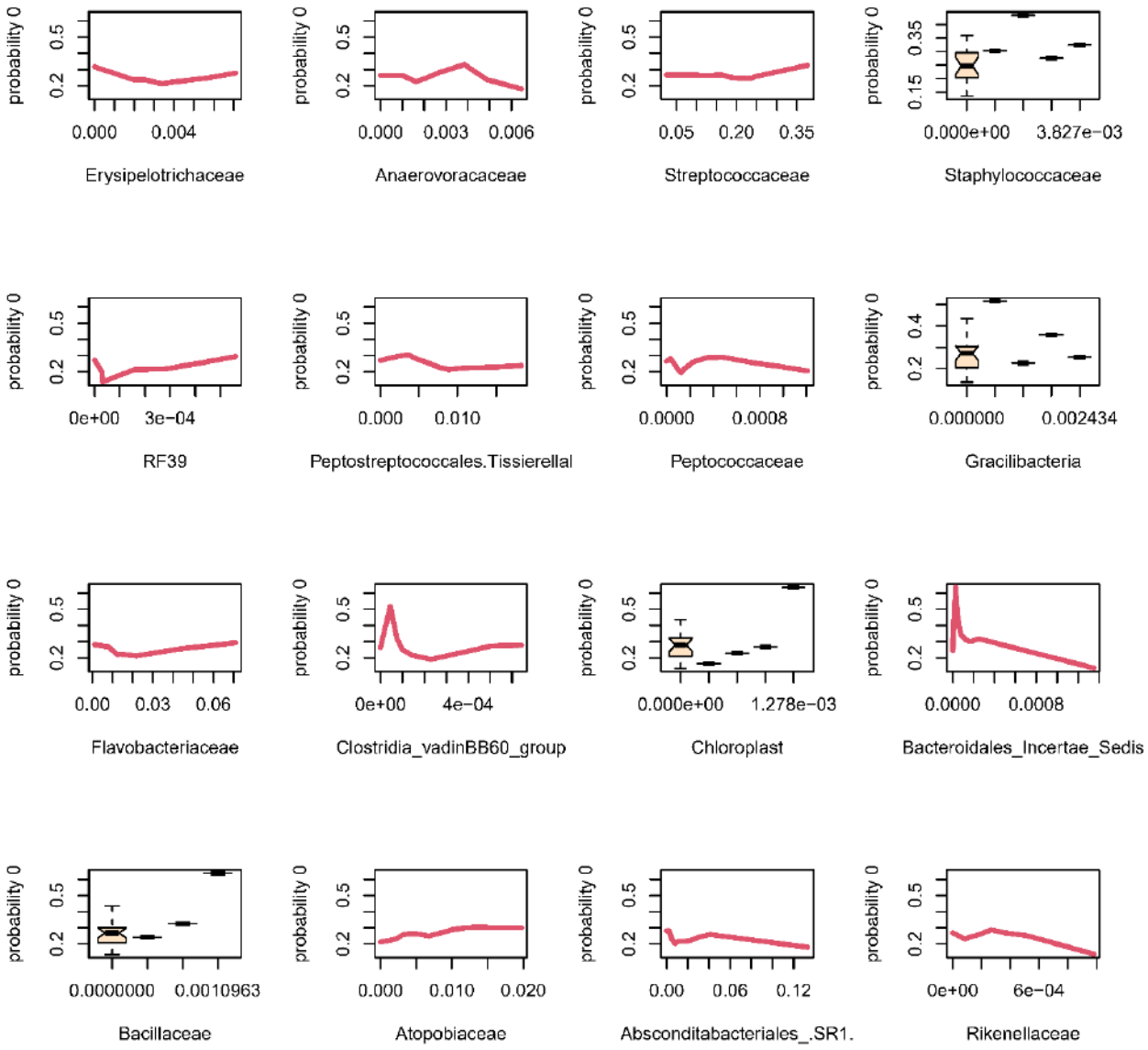
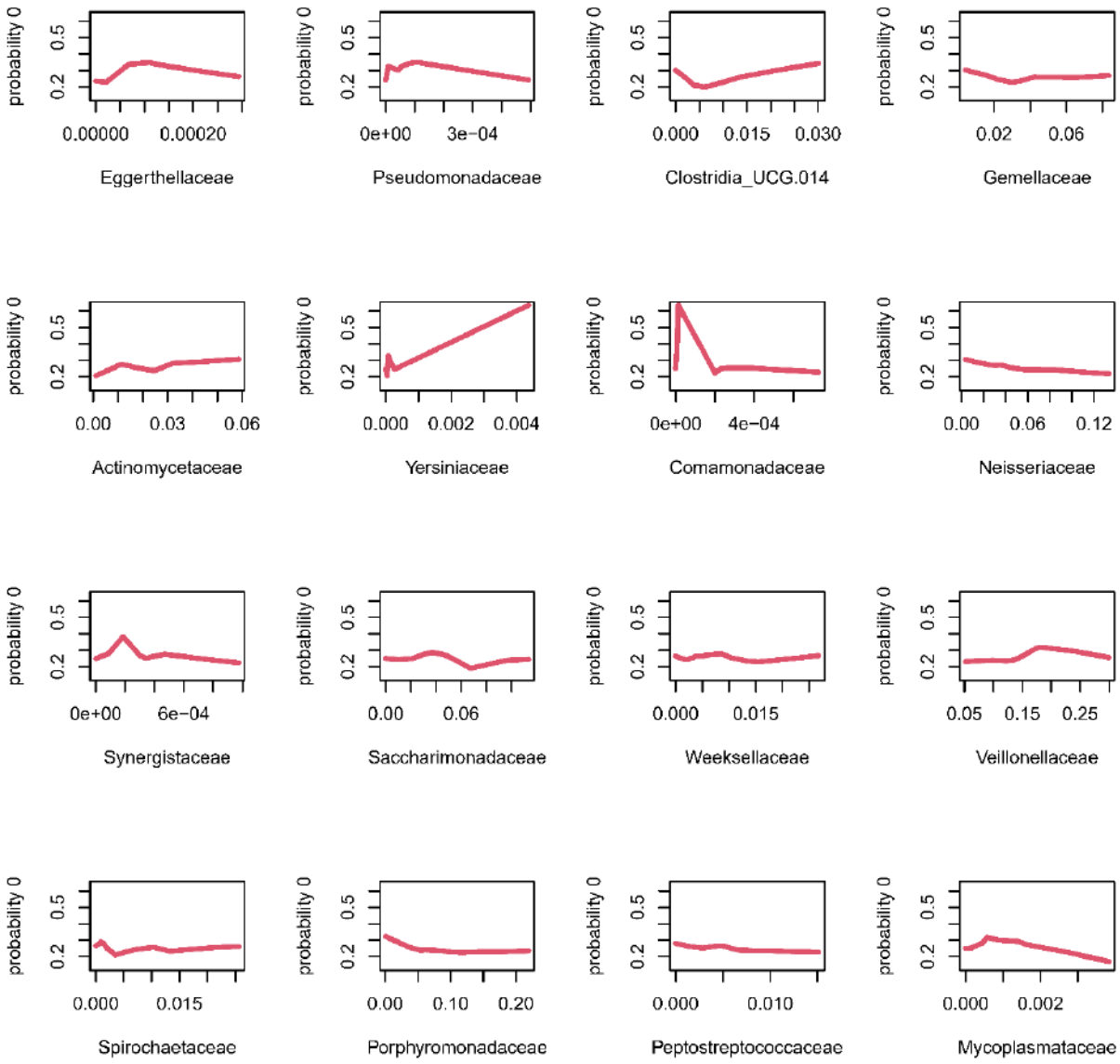


Figure ML3-D-1: The number of trees and Variable Importance Plot for Family in Saliva


probability 0
0.2
0.5
0.00
0.15
0.30
Prevotellaceae
probability 0
0.2
0.5
0.00
0.15
0.30
Pasteurellaceae
probability 0
0.2
0.5
0.00
0.06
0.12
Micrococcaceae
probability 0
0.2
0.5
0.0000
0.0015
Lentimicrobiaceae
probability 0
0.2
0.5
0.000
0.004
0.008
Lactobacillaceae
probability 0
0.2
0.5
0.000
0.002
0.004
JGI_0000069.P22
probability 0
0.2
0.5
0.00
0.06
0.12
Fusobacteriaceae
probability 0
0.2
0.5
0.0000
0.0008
Defluviitaleaceae
probability 0
0.2
0.5
0.000
0.015
0.030
Campylobacteraceae
probability 0
0.2
0.5
0.000
0.002
0.004
Burkholderiaceae
probability 0
0.2
0.5
0.000
0.010
0.020
Selenomonadaceae
probability 0
0.2
0.5
0.0000
0.0015
Cardiobacteriaceae
probability 0
0.2
0.5
0.000
0.010
Aerococcaceae
probability 0
0.15
0.35
0
1
Gender
probability 0
0.2
0.5
0.0000
0.0015
0.0030
Bifidobacteriaceae
probability 0
0.2
0.5
0.000
0.006
Paludibacteraceae


probability 0
0.2
0.5
0.000
0.004
Erysipelotrichaceae
0.000
0.003
0.006
Anaerovoracaceae
0.05
0.20
0.35
Streptococcaceae
0.15
0.35
0.000e+00
3.827e−03
Staphylococcaceae
0e+00
3e−04
RF39
0.000
0.010
Peptostreptococcales.Tissierellal
0.0000
0.0008
Peptococcaceae
0.4
0.000000
0.002434
Gracilibacteria
0.00
0.03
0.06
Flavobacteriaceae
probability
0e+00
4e−04
Clostridia_vadinBB60_group
0.000e+00
1.278e−03
Chloroplast
Bacteroidales_Incertae_Sedis
0.0000000
0.0010963
Bacillaceae
0.020
Atopobiaceae
0.12
Absconditabacteriales_.SR1.
6e−04
Rikenellaceae


probability 0
0.2
0.5
0.00000
0.00020
Eggerthellaceae
0e+00
3e−04
Pseudomonadaceae
0.000
0.015
0.030
Clostridia_UCG.014
0.02
0.06
Gemellaceae
0.00
0.03
0.06
Actinomycetaceae
0.000
0.002
0.004
Yersiniaceae
0e+00
4e−04
Comamonadaceae
0.00
0.06
0.12
Neisseriaceae
0e+00
6e−04
Synergistaceae
probability 0
0.00
0.06
Saccharimonadaceae
0.000
0.015
Weeksellaceae
0.05
0.15
0.25
Veillonellaceae
0.000
0.015
Spirochaetaceae
0.00
0.10
0.20
Porphyromonadaceae
0.000
0.010
Peptostreptococcaceae
0.000
0.002
Mycoplasmataceae

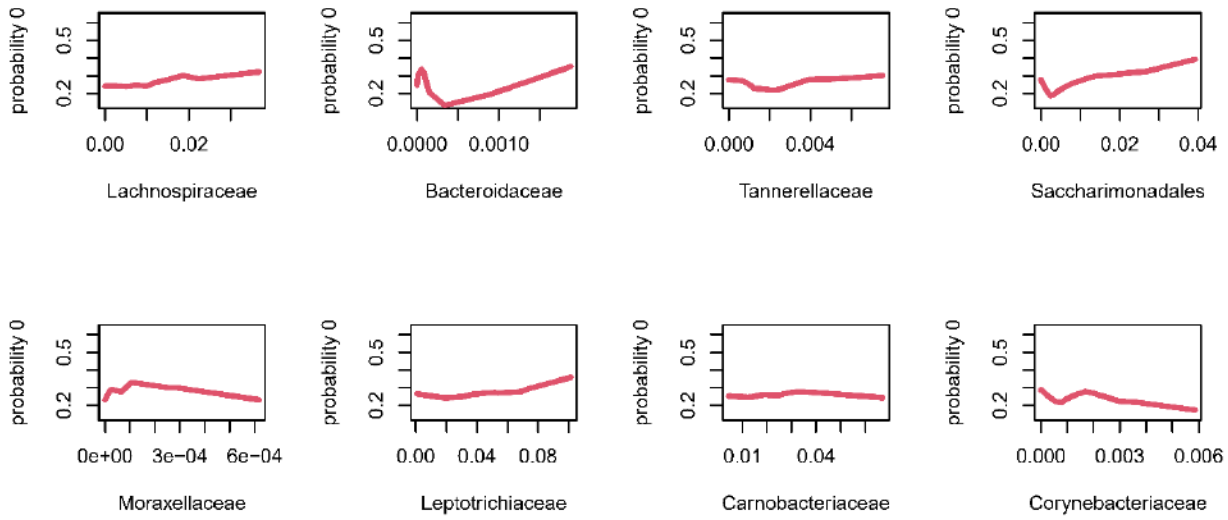
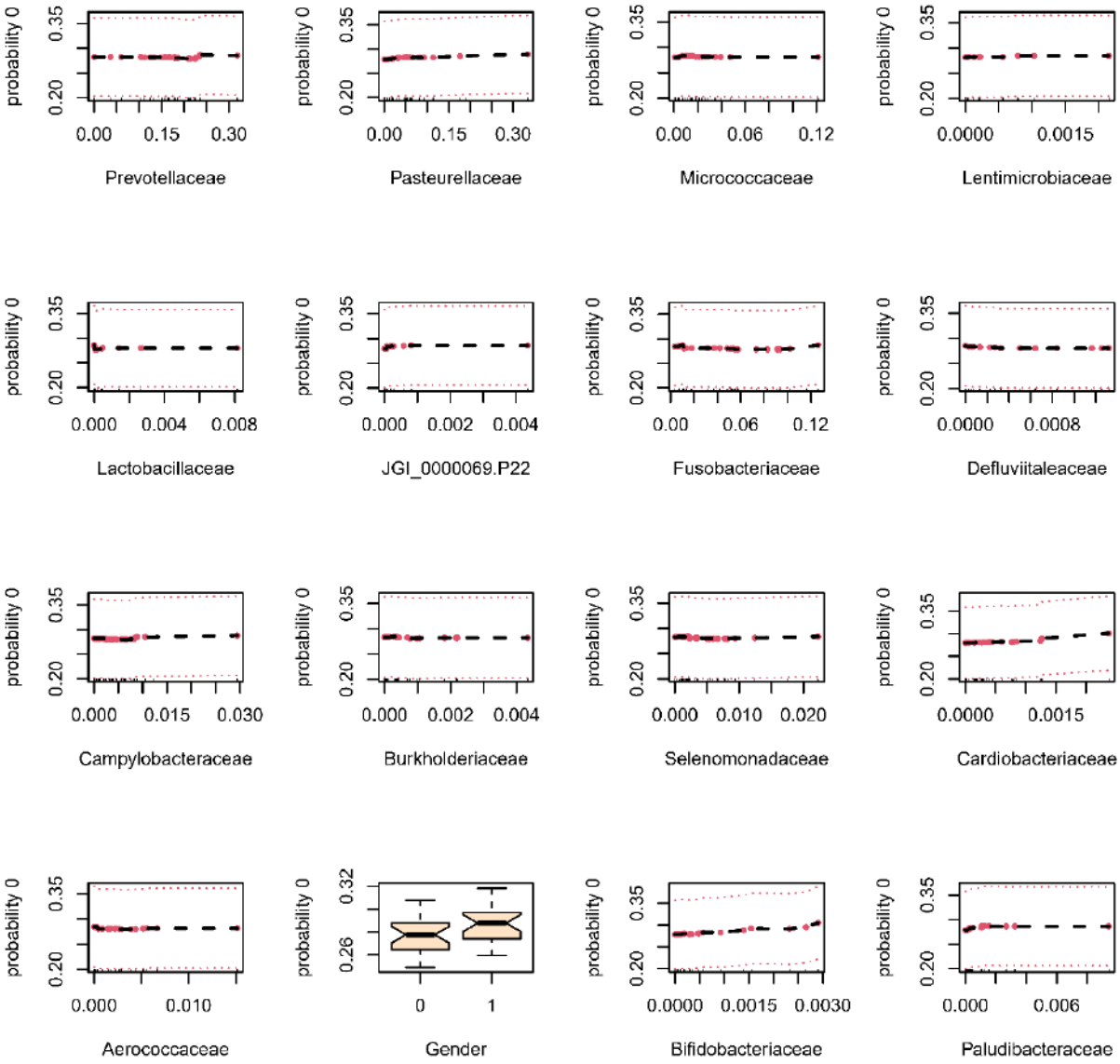
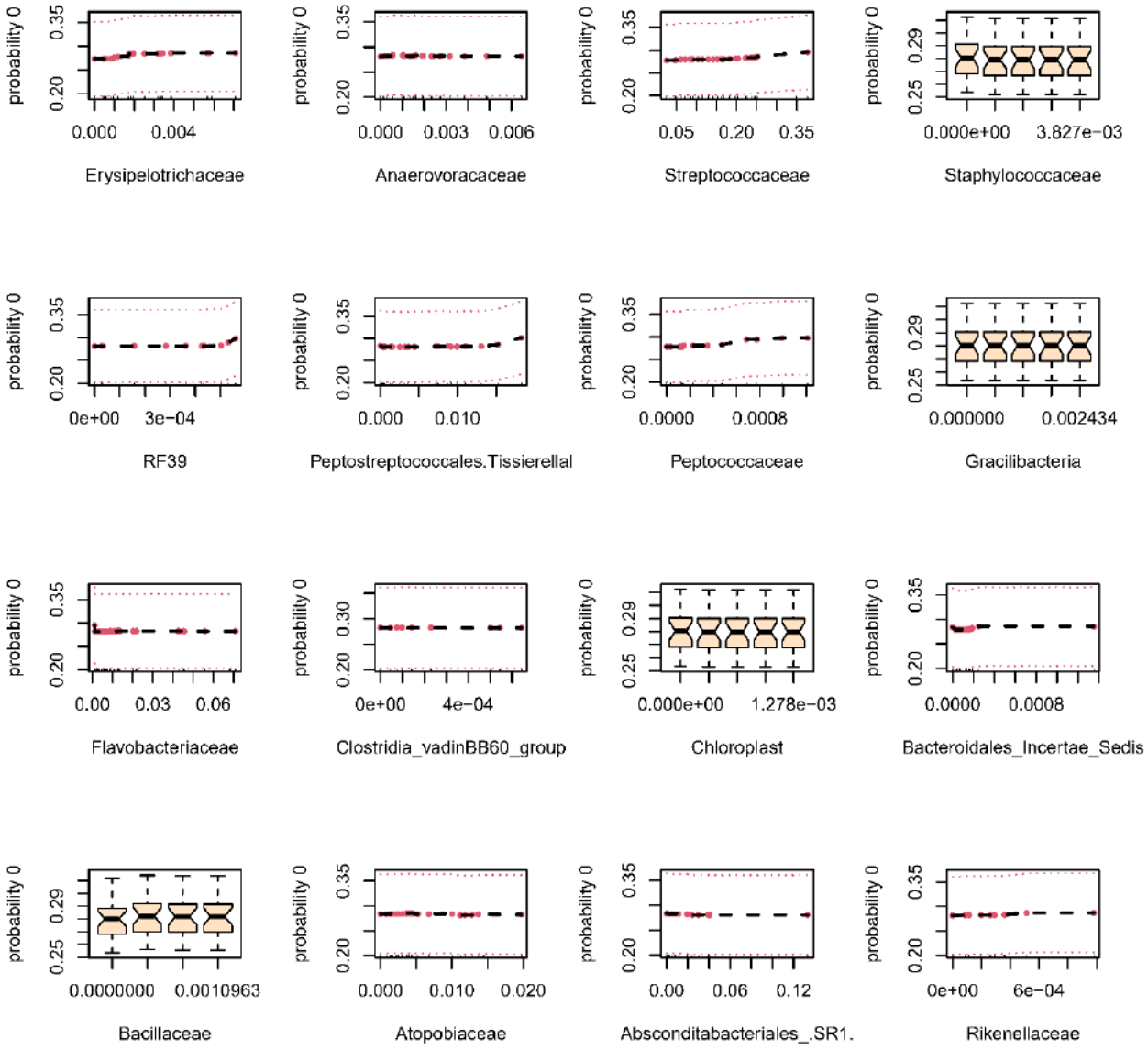
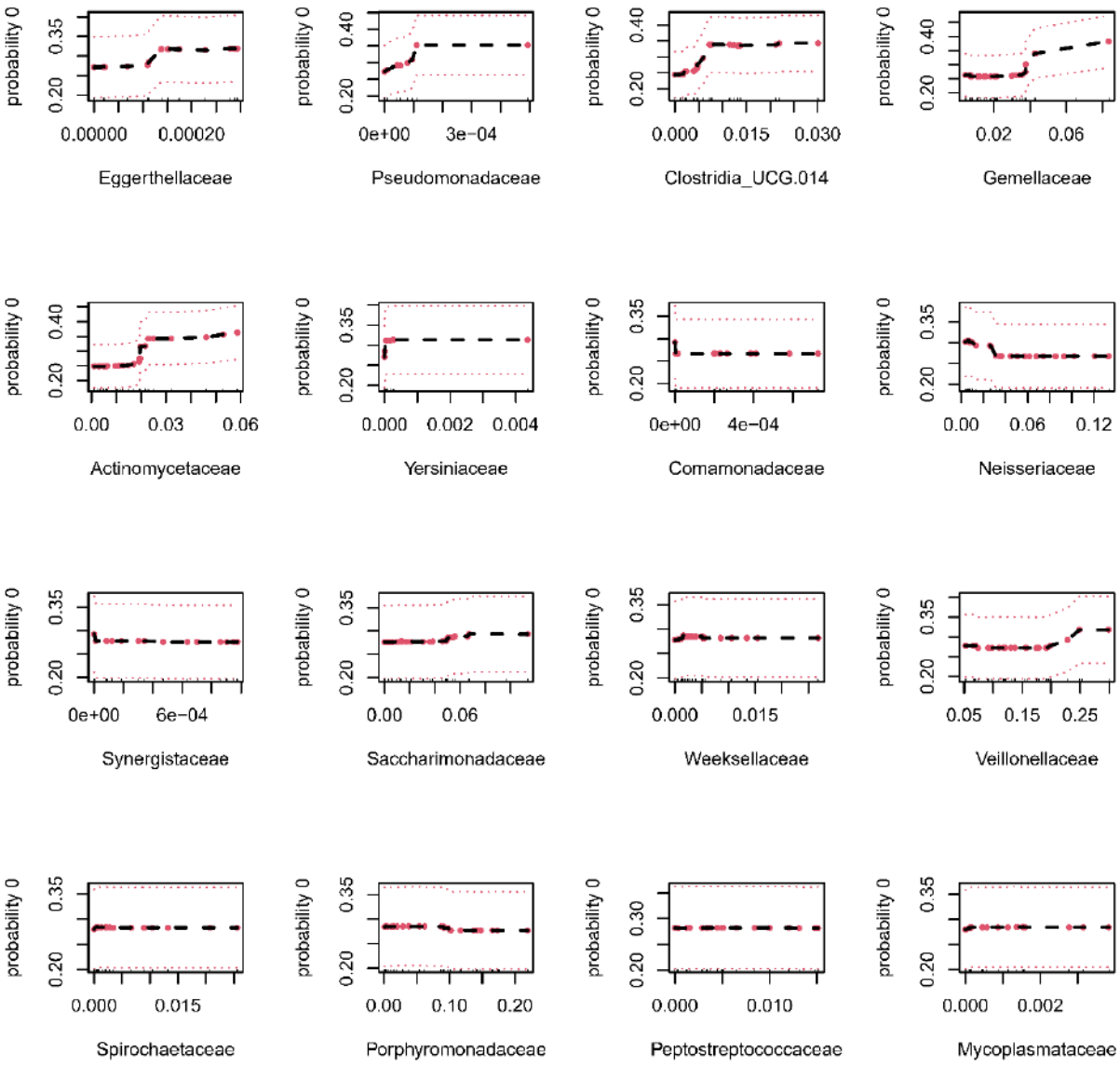


Figure ML3-D-2: Marginal Plot for Family in Saliva


probability 0
0.20
0.35
0.00
0.15
0.30
Prevotellaceae
probability 0
0.20
0.35
0.00
0.15
0.30
Pasteurellaceae
probability 0
0.20
0.35
0.00
0.06
0.12
Micrococcaceae
probability 0
0.20
0.35
0.0000
0.0015
Lentimicrobiaceae
probability 0
0.20
0.35
0.000
0.004
0.008
Lactobacillaceae
probability 0
0.20
0.35
0.000
0.002
0.004
JGI_0000069.P22
probability 0
0.20
0.35
0.00
0.06
0.12
Fusobacteriaceae
probability 0
0.20
0.35
0.0000
0.0008
Defluviitaleaceae
probability 0
0.20
0.35
0.000
0.015
0.030
Campylobacteraceae
probability 0
0.20
0.35
0.000
0.002
0.004
Burkholderiaceae
probability 0
0.20
0.35
0.000
0.010
0.020
Selenomonadaceae
probability 0
0.20
0.35
0.0000
0.0015
Cardiobacteriaceae
probability 0
0.20
0.35
0.000
0.010
Aerococcaceae
probability 0
0.26
0.32
0
1
Gender
probability 0
0.20
0.35
0.0000
0.0015
0.0030
Bifidobacteriaceae
probability 0
0.20
0.35
0.000
0.006
Paludibacteraceae


probability 0
Erysipelotrichaceae
Anaerovoracaceae
Streptococcaceae
Staphylococcaceae
RF39
Peptostreptococcales.Tissierellal
Peptococcaceae
Gracilibacteria
Flavobacteriaceae
Clostridia_vadinBB60_group
Chloroplast
Bacteroidales_Incertae_Sedis
Bacillaceae
Atopobiaceae
Absconditabacteriales_.SR1.
Rikenellaceae


probability 0
0.20
0.35
0.00000
0.00020
Eggerthellaceae
probability 0
0.20
0.40
0e+00
3e−04
Pseudomonadaceae
probability 0
0.20
0.40
0.000
0.015
0.030
Clostridia_UCG.014
probability 0
0.20
0.40
0.02
0.06
Gemellaceae
probability 0
0.20
0.40
0.00
0.03
0.06
Actinomycetaceae
probability 0
0.20
0.35
0.000
0.002
0.004
Yersiniaceae
probability 0
0.20
0.35
0e+00
4e−04
Comamonadaceae
probability 0
0.20
0.35
0.00
0.06
0.12
Neisseriaceae
probability 0
0.20
0.35
0e+00
6e−04
Synergistaceae
probability 0
0.20
0.35
0.00
0.06
Saccharimonadaceae
probability 0
0.20
0.35
0.000
0.015
Weeksellaceae
probability 0
0.20
0.35
0.05
0.15
0.25
Veillonellaceae
probability 0
0.20
0.35
0.000
0.015
Spirochaetaceae
probability 0
0.20
0.35
0.00
0.10
0.20
Porphyromonadaceae
probability 0
0.20
0.30
0.000
0.010
Peptostreptococcaceae
probability 0
0.20
0.35
0.000
0.002
Mycoplasmataceae

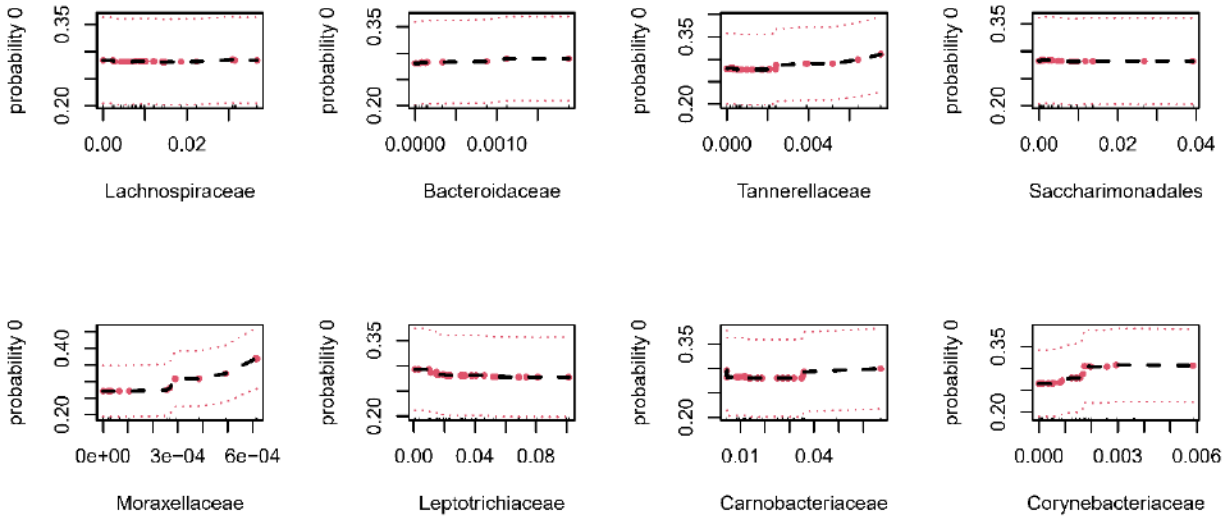


Figure ML3-D-3: Partial Plot for Family in Saliva

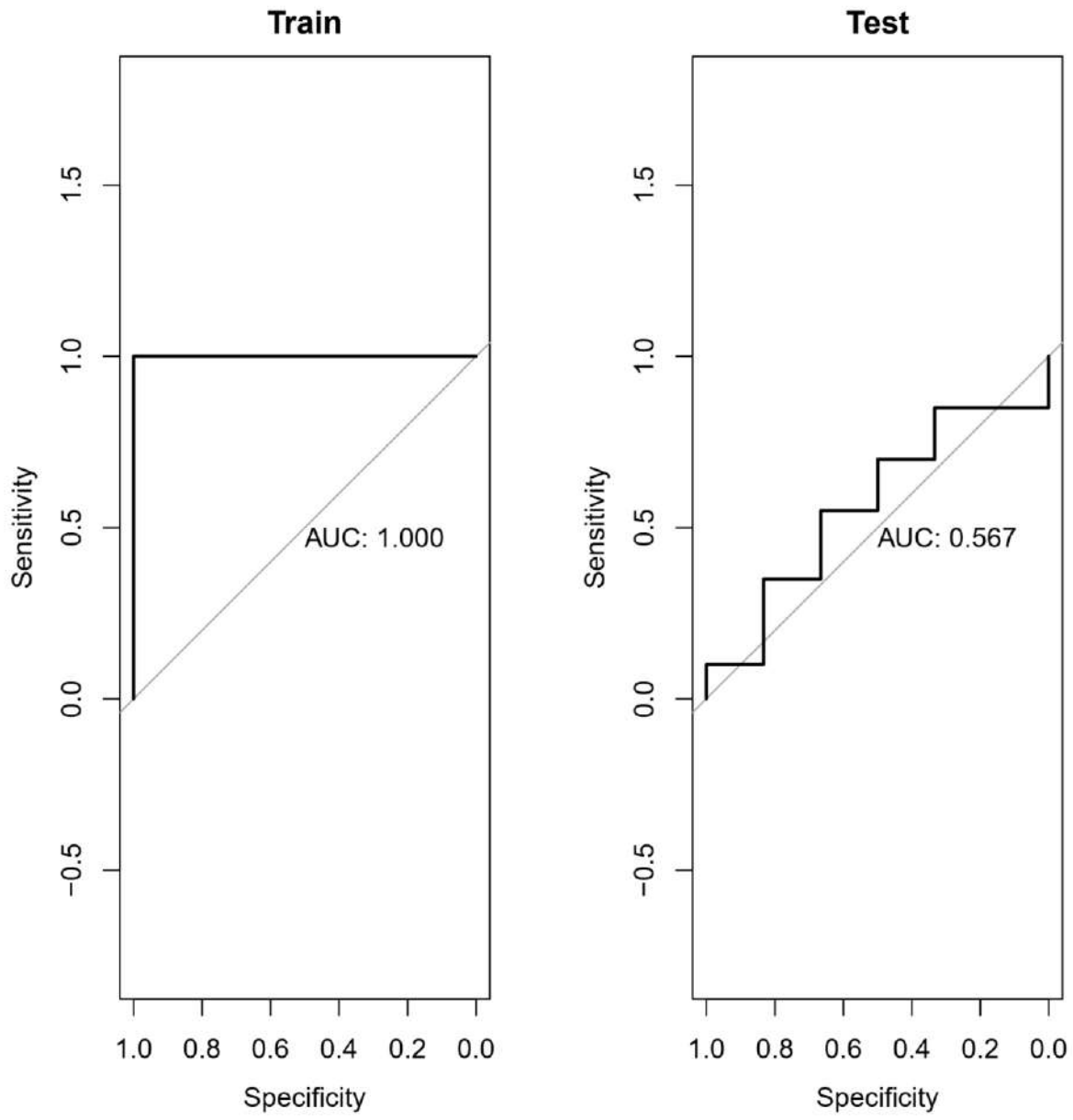


Figure ML3-D-4: The ROC Curve for Family in Saliva

Table ML3-D-1: The Model Performance

Index	Values		
	2.50%	50.00%	97.50%
Accuracy	0.31	0.65	0.85
Specificity	0.33	0.83	1.00
Sensitivity	0.10	0.65	0.95

Table ML3-D-2: The Variable Importance Values

Covariates	Variable Importance			Brier Score		
	all	AP	CRC	all	AP	CRC
Gender	-0.0004	-0.0181	0.0041	-0.0074	-0.0050	-0.0050
Absconditabacteriales_.SR1.	-0.0027	0.0000	-0.0095	-0.0034	-0.0023	-0.0023
Actinomycetaceae	0.0405	-0.0544	0.1577	0.1234	0.0838	0.0838
Aerococcaceae	0.0008	-0.0227	0.0095	-0.0088	-0.0060	-0.0060
Anaerovoracaceae	-0.0028	0.0136	-0.0136	-0.0089	-0.0061	-0.0061
Atopobiaceae	0.0000	-0.0045	0.0014	0.0026	0.0017	0.0017
Bacillaceae	-0.0031	-0.0091	-0.0082	-0.0037	-0.0025	-0.0025
Bacteroidaceae	-0.0016	-0.0181	0.0000	-0.0082	-0.0055	-0.0055
Bacteroidales_Incertae_Sedis	0.0012	-0.0091	0.0068	-0.0019	-0.0013	-0.0013
Bifidobacteriaceae	0.0019	-0.0634	0.0258	-0.0087	-0.0059	-0.0059
Burkholderiaceae	0.0012	0.0000	0.0041	0.0032	0.0022	0.0022
Campylobacteraceae	-0.0027	-0.0227	-0.0027	-0.0038	-0.0026	-0.0026
Cardiobacteriaceae	0.0028	-0.0498	0.0245	-0.0039	-0.0027	-0.0027
Carnobacteriaceae	0.0142	-0.1269	0.0870	0.0092	0.0062	0.0062
Chloroplast	0.0000	0.0000	0.0000	-0.0007	-0.0005	-0.0005
Clostridia_UCG.014	0.0186	-0.0453	0.0788	0.1007	0.0684	0.0684
Clostridia_vadinBB60_group	-0.0015	-0.0045	-0.0041	-0.0009	-0.0006	-0.0006
Comamonadaceae	0.0023	0.0272	0.0000	0.0145	0.0099	0.0099
Corynebacteriaceae	0.0019	-0.0362	0.0177	0.0004	0.0002	0.0002
Defluviitaleaceae	0.0000	-0.0045	0.0014	-0.0051	-0.0035	-0.0035
Eggerthellaceae	0.0116	-0.0453	0.0544	0.0311	0.0212	0.0212
Erysipelotrichaceae	-0.0028	0.0045	-0.0109	-0.0144	-0.0098	-0.0098
Flavobacteriaceae	0.0038	-0.0589	0.0313	0.0000	0.0000	0.0000
Fusobacteriaceae	-0.0027	-0.0544	0.0068	-0.0125	-0.0085	-0.0085
Gemellaceae	0.0498	-0.2265	0.2419	0.1211	0.0823	0.0823
Gracilibacteria	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
JGI_0000069.P22	0.0000	0.0045	-0.0014	-0.0120	-0.0081	-0.0081
Lachnospiraceae	0.0004	-0.0317	0.0109	-0.0023	-0.0016	-0.0016
Lactobacillaceae	-0.0019	-0.0136	-0.0027	-0.0086	-0.0059	-0.0059
Lentimicrobiaceae	0.0004	-0.0045	0.0027	0.0027	0.0019	0.0019
Leptotrichiaceae	0.0015	-0.0272	0.0136	-0.0141	-0.0096	-0.0096

Micrococcaceae	-0.0015	0.0000	-0.0054	-0.0081	-0.0055	-0.0055
Moraxellaceae	0.0347	-0.2492	0.1957	0.0695	0.0472	0.0472
Mycoplasmataceae	0.0004	0.0045	0.0000	-0.0022	-0.0015	-0.0015
Neisseriaceae	-0.0077	-0.0045	-0.0258	-0.0060	-0.0041	-0.0041
Paludibacteraceae	-0.0035	-0.0227	-0.0054	-0.0179	-0.0122	-0.0122
Pasteurellaceae	0.0027	-0.0091	0.0122	-0.0051	-0.0035	-0.0035
Peptococcaceae	-0.0027	-0.0680	0.0109	-0.0053	-0.0036	-0.0036
Peptostreptococcaceae	-0.0015	0.0000	-0.0054	-0.0016	-0.0011	-0.0011
Peptostreptococcales.Tissierellales	0.0040	-0.0997	0.0435	0.0062	0.0042	0.0042
Porphyromonadaceae	-0.0031	-0.0181	-0.0054	-0.0097	-0.0066	-0.0066
Prevotellaceae	-0.0023	-0.0181	-0.0027	-0.0108	-0.0074	-0.0074
Pseudomonadaceae	0.0257	-0.1767	0.1427	0.0516	0.0351	0.0351
RF39	0.0051	-0.0362	0.0285	0.0004	0.0002	0.0002
Rikenellaceae	-0.0016	-0.0408	0.0068	-0.0049	-0.0033	-0.0033
Saccharimonadaceae	-0.0023	-0.0136	-0.0041	-0.0185	-0.0126	-0.0126
Saccharimonadales	-0.0039	-0.0091	-0.0109	-0.0082	-0.0056	-0.0056
Selenomonadaceae	-0.0028	-0.0272	-0.0014	-0.0113	-0.0077	-0.0077
Spirochaetaceae	-0.0031	0.0045	-0.0122	-0.0039	-0.0027	-0.0027
Staphylococcaceae	0.0000	0.0045	-0.0014	0.0003	0.0002	0.0002
Streptococcaceae	-0.0039	-0.0498	0.0014	-0.0046	-0.0031	-0.0031
Synergistaceae	0.0058	0.0136	0.0163	0.0163	0.0111	0.0111
Tannerellaceae	0.0094	-0.1042	0.0639	0.0012	0.0008	0.0008
Veillonellaceae	0.0233	-0.0815	0.1060	0.0248	0.0168	0.0168
Weeksellaceae	-0.0050	0.0181	-0.0231	-0.0022	-0.0015	-0.0015
Yersiniaceae	0.0116	-0.0362	0.0516	0.0251	0.0171	0.0171

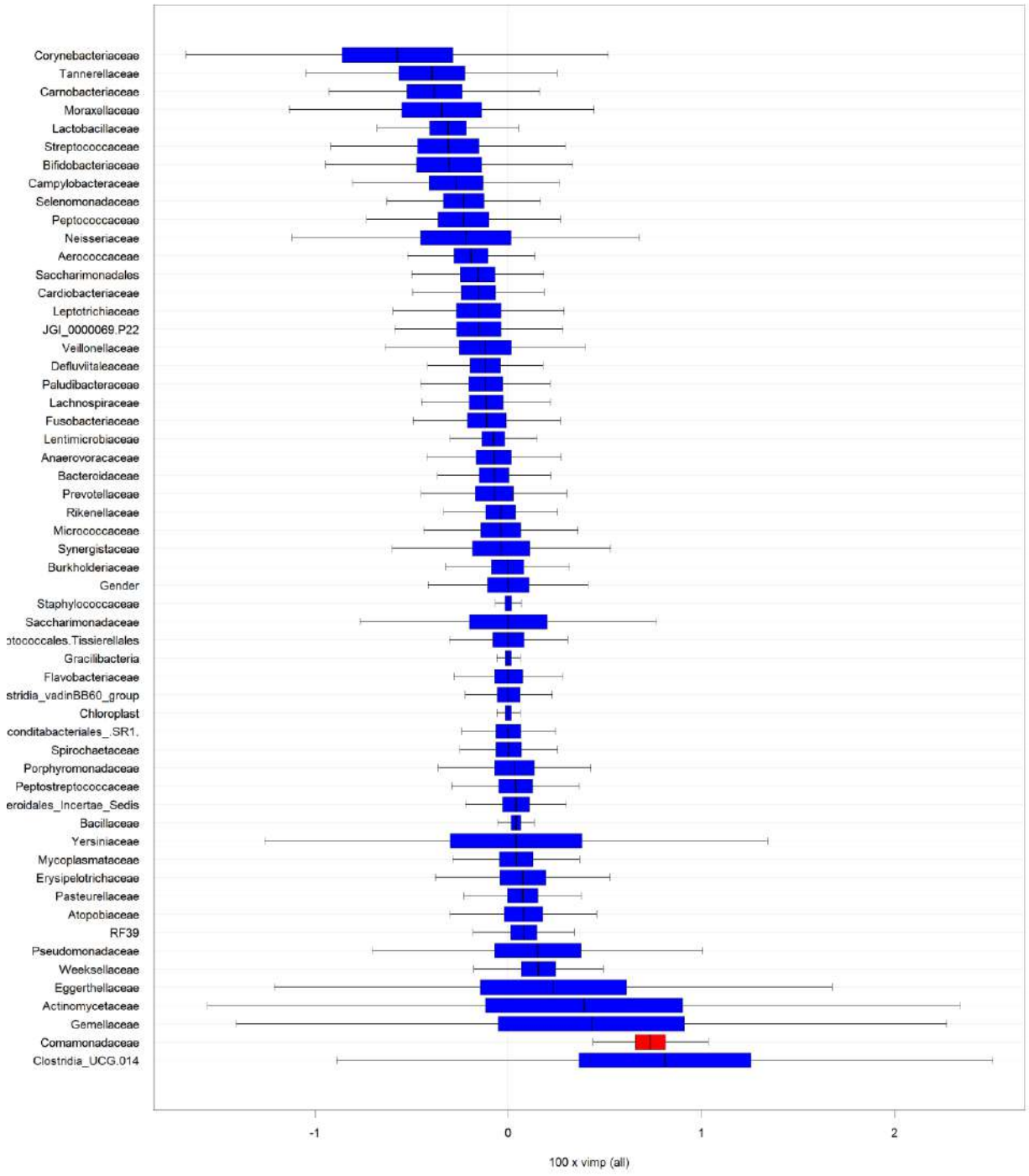


Figure ML3-D-5: The Variable Importance Plot (VIMP) for Family in Saliva

Best model is selected for LIME results

Table ML3-D-3: The best setting for model and LIME Results

Setting	mtry	ROC	Sens	Spec	ROCSD	SensSD	SpecSD

ntrees: 500 nodesize: 1	10	0.4625	0	0.95	0.214695	0	0.111803
ntrees: 1000 nodesize: 1	7	0.45	0	1	0.111803	0	0
ntrees: 1000 nodesize: 1	9	0.45	0	0.95	0.209165	0	0.111803
ntrees: 500 nodesize: 5	5	0.45	0	0.95	0.189572	0	0.111803
ntrees: 500 nodesize: 5	9	0.45	0	0.95	0.209165	0	0.111803
ntrees: 1000 nodesize: 5	10	0.5	0	0.95	0.306186	0	0.111803

Table ML3-D-4: The LIME Result for Training Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
2	Adenoma	0.706	0.342	0.333	0.412	Pseudomonadaceae	0.000	-0.070	Pseudomonadaceae <= 6.6e-06
2	Adenoma	0.706	0.342	0.333	0.412	Gemellaceae	0.043	0.069	0.0352 < Gemellaceae
2	Adenoma	0.706	0.342	0.333	0.412	Eggerthellaceae	0.000	0.041	0.000131 < Eggerthellaceae
2	Adenoma	0.706	0.342	0.333	0.412	Moraxellaceae	0.001	0.021	0.000218 < Moraxellaceae
2	Adenoma	0.706	0.342	0.333	0.412	Clostridia_UCG.014	0.008	0.019	0.00404 < Clostridia_UCG.014 <= 0.01058
3	Adenoma	0.750	0.450	0.249	0.519	Pseudomonadaceae	0.000	0.072	6.6e-06 < Pseudomonadaceae
3	Adenoma	0.750	0.450	0.249	0.519	Clostridia_UCG.014	0.012	0.069	0.01058 < Clostridia_UCG.014
3	Adenoma	0.750	0.450	0.249	0.519	Actinomycetaceae	0.053	0.060	0.02245 < Actinomycetaceae
3	Adenoma	0.750	0.450	0.249	0.519	Eggerthellaceae	0.000	0.041	0.000131 < Eggerthellaceae
3	Adenoma	0.750	0.450	0.249	0.519	Veillonellaceae	0.249	0.029	0.1918 < Veillonellaceae
6	Adenoma	0.724	0.359	0.338	0.352	Pseudomonadaceae	0.000	0.072	6.6e-06 < Pseudomonadaceae
6	Adenoma	0.724	0.359	0.338	0.352	Actinomycetaceae	0.059	0.059	0.02245 < Actinomycetaceae
6	Adenoma	0.724	0.359	0.338	0.352	Eggerthellaceae	0.000	-0.045	Eggerthellaceae <= 0.000131
6	Adenoma	0.724	0.359	0.338	0.352	Clostridia_UCG.014	0.002	-0.042	0.00172 < Clostridia_UCG.014 <= 0.00404
6	Adenoma	0.724	0.359	0.338	0.352	Gemellaceae	0.030	-0.030	0.0211 < Gemellaceae <= 0.0352
7	Adenoma	0.726	0.363	0.339	0.383	Gemellaceae	0.065	0.072	0.0352 < Gemellaceae
7	Adenoma	0.726	0.363	0.339	0.383	Pseudomonadaceae	0.000	-0.068	Pseudomonadaceae <= 6.6e-06
7	Adenoma	0.726	0.363	0.339	0.383	Eggerthellaceae	0.000	0.044	0.000131 < Eggerthellaceae
7	Adenoma	0.726	0.363	0.339	0.383	Actinomycetaceae	0.006	-0.028	0.00601 < Actinomycetaceae <= 0.01363
7	Adenoma	0.726	0.363	0.339	0.383	Neisseriaceae	0.007	0.023	Neisseriaceae <= 0.0108
10	Adenoma	0.760	0.546	0.310	0.472	Pseudomonadaceae	0.000	-0.071	Pseudomonadaceae <= 6.6e-06
10	Adenoma	0.760	0.546	0.310	0.472	Gemellaceae	0.083	0.064	0.0352 < Gemellaceae
10	Adenoma	0.760	0.546	0.310	0.472	Clostridia_UCG.014	0.022	0.068	0.01058 < Clostridia_UCG.014
10	Adenoma	0.760	0.546	0.310	0.472	Actinomycetaceae	0.025	0.057	0.02245 < Actinomycetaceae
10	Adenoma	0.760	0.546	0.310	0.472	Eggerthellaceae	0.000	0.044	0.000131 < Eggerthellaceae
11	Adenoma	0.670	0.443	0.289	0.475	Pseudomonadaceae	0.000	0.074	6.6e-06 < Pseudomonadaceae
11	Adenoma	0.670	0.443	0.289	0.475	Clostridia_UCG.014	0.012	0.070	0.01058 < Clostridia_UCG.014
11	Adenoma	0.670	0.443	0.289	0.475	Actinomycetaceae	0.023	0.057	0.02245 < Actinomycetaceae

11	Adenoma	0.670	0.443	0.289	0.475	Eggerthellaceae	0.000	-0.042	Eggerthellaceae <= 0.000131
11	Adenoma	0.670	0.443	0.289	0.475	Neisseriaceae	0.006	0.027	Neisseriaceae <= 0.0108
14	Adenocarcinoma	0.914	0.285	0.584	0.754	Pseudomonadaceae	0.000	0.068	Pseudomonadaceae <= 6.6e-06
14	Adenocarcinoma	0.914	0.285	0.584	0.754	Clostridia_UCG.014	0.001	0.050	Clostridia_UCG.014 <= 0.00172
14	Adenocarcinoma	0.914	0.285	0.584	0.754	Eggerthellaceae	0.000	0.044	Eggerthellaceae <= 0.000131
14	Adenocarcinoma	0.914	0.285	0.584	0.754	Gemellaceae	0.020	0.030	0.0129 < Gemellaceae <= 0.0211
14	Adenocarcinoma	0.914	0.285	0.584	0.754	Saccharimonadaceae	0.114	-0.022	0.0502 < Saccharimonadaceae
20	Adenocarcinoma	0.924	0.251	0.602	0.696	Pseudomonadaceae	0.000	0.069	Pseudomonadaceae <= 6.6e-06
20	Adenocarcinoma	0.924	0.251	0.602	0.696	Eggerthellaceae	0.000	0.047	Eggerthellaceae <= 0.000131
20	Adenocarcinoma	0.924	0.251	0.602	0.696	Actinomycetaceae	0.005	0.027	Actinomycetaceae <= 0.00601
20	Adenocarcinoma	0.924	0.251	0.602	0.696	Saccharimonadaceae	0.054	-0.024	0.0502 < Saccharimonadaceae
20	Adenocarcinoma	0.924	0.251	0.602	0.696	Corynebacteriaceae	0.006	-0.024	0.001784 < Corynebacteriaceae
21	Adenocarcinoma	0.878	0.368	0.681	0.593	Clostridia_UCG.014	0.021	-0.072	0.01058 < Clostridia_UCG.014
21	Adenocarcinoma	0.878	0.368	0.681	0.593	Pseudomonadaceae	0.000	-0.068	6.6e-06 < Pseudomonadaceae
21	Adenocarcinoma	0.878	0.368	0.681	0.593	Eggerthellaceae	0.000	0.046	Eggerthellaceae <= 0.000131
21	Adenocarcinoma	0.878	0.368	0.681	0.593	Actinomycetaceae	0.006	0.032	Actinomycetaceae <= 0.00601
21	Adenocarcinoma	0.878	0.368	0.681	0.593	Veillonellaceae	0.199	-0.025	0.1918 < Veillonellaceae
22	Adenocarcinoma	0.916	0.270	0.588	0.745	Pseudomonadaceae	0.000	0.069	Pseudomonadaceae <= 6.6e-06
22	Adenocarcinoma	0.916	0.270	0.588	0.745	Eggerthellaceae	0.000	0.043	Eggerthellaceae <= 0.000131
22	Adenocarcinoma	0.916	0.270	0.588	0.745	Clostridia_UCG.014	0.004	0.042	0.00172 < Clostridia_UCG.014 <= 0.00404
22	Adenocarcinoma	0.916	0.270	0.588	0.745	Actinomycetaceae	0.005	0.027	Actinomycetaceae <= 0.00601
22	Adenocarcinoma	0.916	0.270	0.588	0.745	Corynebacteriaceae	0.003	-0.025	0.001784 < Corynebacteriaceae
23	Adenocarcinoma	0.898	0.347	0.638	0.681	Pseudomonadaceae	0.000	0.071	Pseudomonadaceae <= 6.6e-06
23	Adenocarcinoma	0.898	0.347	0.638	0.681	Actinomycetaceae	0.046	-0.056	0.02245 < Actinomycetaceae
23	Adenocarcinoma	0.898	0.347	0.638	0.681	Clostridia_UCG.014	0.002	0.044	0.00172 < Clostridia_UCG.014 <= 0.00404
23	Adenocarcinoma	0.898	0.347	0.638	0.681	Eggerthellaceae	0.000	-0.039	0.000131 < Eggerthellaceae
23	Adenocarcinoma	0.898	0.347	0.638	0.681	Gemellaceae	0.008	0.023	Gemellaceae <= 0.0129
24	Adenocarcinoma	0.928	0.217	0.586	0.712	Pseudomonadaceae	0.000	0.068	Pseudomonadaceae <= 6.6e-06
24	Adenocarcinoma	0.928	0.217	0.586	0.712	Eggerthellaceae	0.000	0.038	Eggerthellaceae <= 0.000131
24	Adenocarcinoma	0.928	0.217	0.586	0.712	Saccharimonadaceae	0.066	-0.023	0.0502 < Saccharimonadaceae
24	Adenocarcinoma	0.928	0.217	0.586	0.712	Gemellaceae	0.011	0.022	Gemellaceae <= 0.0129
24	Adenocarcinoma	0.928	0.217	0.586	0.712	Moraxellaceae	0.000	0.021	Moraxellaceae <= 0.000218
25	Adenocarcinoma	0.888	0.304	0.607	0.727	Pseudomonadaceae	0.000	0.071	Pseudomonadaceae <= 6.6e-06
25	Adenocarcinoma	0.888	0.304	0.607	0.727	Clostridia_UCG.014	0.000	0.052	Clostridia_UCG.014 <= 0.00172
25	Adenocarcinoma	0.888	0.304	0.607	0.727	Eggerthellaceae	0.000	-0.044	0.000131 < Eggerthellaceae
25	Adenocarcinoma	0.888	0.304	0.607	0.727	Moraxellaceae	0.000	0.021	Moraxellaceae <= 0.000218
25	Adenocarcinoma	0.888	0.304	0.607	0.727	Gemellaceae	0.014	0.022	0.0129 < Gemellaceae <= 0.0211
26	Adenocarcinoma	0.898	0.356	0.621	0.660	Clostridia_UCG.014	0.014	-0.068	0.01058 < Clostridia_UCG.014

26	Adenocarcinoma	0.898	0.356	0.621	0.660	Pseudomonadaceae	0.000	0.067	Pseudomonadaceae <= 6.6e-06
26	Adenocarcinoma	0.898	0.356	0.621	0.660	Eggerthellaceae	0.000	0.042	Eggerthellaceae <= 0.000131
26	Adenocarcinoma	0.898	0.356	0.621	0.660	Veillonellaceae	0.228	-0.031	0.1918 < Veillonellaceae
26	Adenocarcinoma	0.898	0.356	0.621	0.660	Actinomycetaceae	0.003	0.029	Actinomycetaceae <= 0.00601
29	Adenocarcinoma	0.952	0.394	0.597	0.717	Pseudomonadaceae	0.000	0.068	Pseudomonadaceae <= 6.6e-06
29	Adenocarcinoma	0.952	0.394	0.597	0.717	Gemellaceae	0.038	-0.066	0.0352 < Gemellaceae
29	Adenocarcinoma	0.952	0.394	0.597	0.717	Clostridia_UCG.014	0.001	0.048	Clostridia_UCG.014 <= 0.00172
29	Adenocarcinoma	0.952	0.394	0.597	0.717	Eggerthellaceae	0.000	0.041	Eggerthellaceae <= 0.000131
29	Adenocarcinoma	0.952	0.394	0.597	0.717	Actinomycetaceae	0.006	0.030	Actinomycetaceae <= 0.00601
32	Adenocarcinoma	0.946	0.230	0.573	0.757	Pseudomonadaceae	0.000	0.070	Pseudomonadaceae <= 6.6e-06
32	Adenocarcinoma	0.946	0.230	0.573	0.757	Eggerthellaceae	0.000	0.043	Eggerthellaceae <= 0.000131
32	Adenocarcinoma	0.946	0.230	0.573	0.757	Actinomycetaceae	0.011	0.027	0.00601 < Actinomycetaceae <= 0.01363
32	Adenocarcinoma	0.946	0.230	0.573	0.757	Gemellaceae	0.033	0.025	0.0211 < Gemellaceae <= 0.0352
32	Adenocarcinoma	0.946	0.230	0.573	0.757	Moraxellaceae	0.000	0.018	Moraxellaceae <= 0.000218
33	Adenocarcinoma	0.830	0.353	0.629	0.625	Pseudomonadaceae	0.000	0.070	Pseudomonadaceae <= 6.6e-06
33	Adenocarcinoma	0.830	0.353	0.629	0.625	Clostridia_UCG.014	0.013	-0.066	0.01058 < Clostridia_UCG.014
33	Adenocarcinoma	0.830	0.353	0.629	0.625	Eggerthellaceae	0.000	0.040	Eggerthellaceae <= 0.000131
33	Adenocarcinoma	0.830	0.353	0.629	0.625	Veillonellaceae	0.195	-0.026	0.1918 < Veillonellaceae
33	Adenocarcinoma	0.830	0.353	0.629	0.625	Neisseriaceae	0.004	-0.024	Neisseriaceae <= 0.0108
34	Adenocarcinoma	0.936	0.243	0.581	0.776	Pseudomonadaceae	0.000	0.065	Pseudomonadaceae <= 6.6e-06
34	Adenocarcinoma	0.936	0.243	0.581	0.776	Eggerthellaceae	0.000	0.041	Eggerthellaceae <= 0.000131
34	Adenocarcinoma	0.936	0.243	0.581	0.776	Clostridia_UCG.014	0.002	0.039	0.00172 < Clostridia_UCG.014 <= 0.00404
34	Adenocarcinoma	0.936	0.243	0.581	0.776	Actinomycetaceae	0.003	0.028	Actinomycetaceae <= 0.00601
34	Adenocarcinoma	0.936	0.243	0.581	0.776	Gemellaceae	0.030	0.023	0.0211 < Gemellaceae <= 0.0352
38	Adenocarcinoma	0.884	0.353	0.674	0.632	Pseudomonadaceae	0.000	-0.069	6.6e-06 < Pseudomonadaceae
38	Adenocarcinoma	0.884	0.353	0.674	0.632	Clostridia_UCG.014	0.030	-0.067	0.01058 < Clostridia_UCG.014
38	Adenocarcinoma	0.884	0.353	0.674	0.632	Eggerthellaceae	0.000	0.042	Eggerthellaceae <= 0.000131
38	Adenocarcinoma	0.884	0.353	0.674	0.632	Gemellaceae	0.020	0.029	0.0129 < Gemellaceae <= 0.0211
38	Adenocarcinoma	0.884	0.353	0.674	0.632	Actinomycetaceae	0.013	0.025	0.00601 < Actinomycetaceae <= 0.01363
39	Adenocarcinoma	0.918	0.296	0.587	0.739	Pseudomonadaceae	0.000	0.073	Pseudomonadaceae <= 6.6e-06
39	Adenocarcinoma	0.918	0.296	0.587	0.739	Clostridia_UCG.014	0.000	0.046	Clostridia_UCG.014 <= 0.00172
39	Adenocarcinoma	0.918	0.296	0.587	0.739	Eggerthellaceae	0.000	0.038	Eggerthellaceae <= 0.000131
39	Adenocarcinoma	0.918	0.296	0.587	0.739	Veillonellaceae	0.299	-0.032	0.1918 < Veillonellaceae
39	Adenocarcinoma	0.918	0.296	0.587	0.739	Actinomycetaceae	0.010	0.026	0.00601 < Actinomycetaceae <= 0.01363
41	Adenocarcinoma	0.910	0.359	0.610	0.685	Pseudomonadaceae	0.000	0.068	Pseudomonadaceae <= 6.6e-06
41	Adenocarcinoma	0.910	0.359	0.610	0.685	Actinomycetaceae	0.053	-0.057	0.02245 < Actinomycetaceae
41	Adenocarcinoma	0.910	0.359	0.610	0.685	Clostridia_UCG.014	0.002	0.049	Clostridia_UCG.014 <= 0.00172
41	Adenocarcinoma	0.910	0.359	0.610	0.685	Eggerthellaceae	0.000	0.041	Eggerthellaceae <= 0.000131

41	Adenocarcinoma	0.910	0.359	0.610	0.685	Carnobacteriaceae	0.035	-0.025	0.0311 < Carnobacteriaceae
46	Adenocarcinoma	0.942	0.360	0.618	0.673	Gemellaceae	0.037	-0.069	0.0352 < Gemellaceae
46	Adenocarcinoma	0.942	0.360	0.618	0.673	Pseudomonadaceae	0.000	0.065	Pseudomonadaceae <= 6.6e-06
46	Adenocarcinoma	0.942	0.360	0.618	0.673	Eggerthellaceae	0.000	0.040	Eggerthellaceae <= 0.000131
46	Adenocarcinoma	0.942	0.360	0.618	0.673	Clostridia_UCG.014	0.002	0.039	0.00172 < Clostridia_UCG.014 <= 0.00404
46	Adenocarcinoma	0.942	0.360	0.618	0.673	Corynebacteriaceae	0.002	-0.020	0.001784 < Corynebacteriaceae
47	Adenocarcinoma	0.804	0.352	0.722	0.570	Pseudomonadaceae	0.000	-0.072	6.6e-06 < Pseudomonadaceae
47	Adenocarcinoma	0.804	0.352	0.722	0.570	Actinomycetaceae	0.032	-0.057	0.02245 < Actinomycetaceae
47	Adenocarcinoma	0.804	0.352	0.722	0.570	Clostridia_UCG.014	0.002	0.043	0.00172 < Clostridia_UCG.014 <= 0.00404
47	Adenocarcinoma	0.804	0.352	0.722	0.570	Eggerthellaceae	0.000	-0.039	0.000131 < Eggerthellaceae
47	Adenocarcinoma	0.804	0.352	0.722	0.570	Carnobacteriaceae	0.035	-0.026	0.0311 < Carnobacteriaceae
49	Adenocarcinoma	0.888	0.405	0.599	0.718	Pseudomonadaceae	0.000	0.068	Pseudomonadaceae <= 6.6e-06
49	Adenocarcinoma	0.888	0.405	0.599	0.718	Gemellaceae	0.035	-0.069	0.0352 < Gemellaceae
49	Adenocarcinoma	0.888	0.405	0.599	0.718	Clostridia_UCG.014	0.000	0.051	Clostridia_UCG.014 <= 0.00172
49	Adenocarcinoma	0.888	0.405	0.599	0.718	Eggerthellaceae	0.000	0.040	Eggerthellaceae <= 0.000131
49	Adenocarcinoma	0.888	0.405	0.599	0.718	Actinomycetaceae	0.001	0.030	Actinomycetaceae <= 0.00601
51	Adenocarcinoma	0.874	0.426	0.680	0.606	Gemellaceae	0.043	-0.071	0.0352 < Gemellaceae
51	Adenocarcinoma	0.874	0.426	0.680	0.606	Pseudomonadaceae	0.000	-0.072	6.6e-06 < Pseudomonadaceae
51	Adenocarcinoma	0.874	0.426	0.680	0.606	Clostridia_UCG.014	0.000	0.049	Clostridia_UCG.014 <= 0.00172
51	Adenocarcinoma	0.874	0.426	0.680	0.606	Eggerthellaceae	0.000	0.044	Eggerthellaceae <= 0.000131
51	Adenocarcinoma	0.874	0.426	0.680	0.606	Neisseriaceae	0.008	-0.025	Neisseriaceae <= 0.0108
52	Adenocarcinoma	0.952	0.225	0.583	0.761	Pseudomonadaceae	0.000	0.070	Pseudomonadaceae <= 6.6e-06
52	Adenocarcinoma	0.952	0.225	0.583	0.761	Eggerthellaceae	0.000	0.042	Eggerthellaceae <= 0.000131
52	Adenocarcinoma	0.952	0.225	0.583	0.761	Gemellaceae	0.034	0.025	0.0211 < Gemellaceae <= 0.0352
52	Adenocarcinoma	0.952	0.225	0.583	0.761	Actinomycetaceae	0.006	0.024	0.00601 < Actinomycetaceae <= 0.01363
52	Adenocarcinoma	0.952	0.225	0.583	0.761	Veillonellaceae	0.098	0.017	0.0937 < Veillonellaceae <= 0.1343

Table ML-D-5: The LIME Result for Test Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
1	Adenocarcinoma	0.838	0.366	0.618	0.661	Pseudomonadaceae	0.000	0.071	Pseudomonadaceae <= 6.6e-06
1	Adenocarcinoma	0.838	0.366	0.618	0.661	Gemellaceae	0.049	-0.069	0.0352 < Gemellaceae
1	Adenocarcinoma	0.838	0.366	0.618	0.661	Eggerthellaceae	0.000	0.041	Eggerthellaceae <= 0.000131
1	Adenocarcinoma	0.838	0.366	0.618	0.661	Corynebacteriaceae	0.004	-0.023	0.001784 < Corynebacteriaceae
1	Adenocarcinoma	0.838	0.366	0.618	0.661	Actinomycetaceae	0.013	0.025	0.00601 < Actinomycetaceae <= 0.01363
4	Adenocarcinoma	0.728	0.348	0.618	0.659	Clostridia_UCG.014	0.014	-0.069	0.01058 < Clostridia_UCG.014
4	Adenocarcinoma	0.728	0.348	0.618	0.659	Pseudomonadaceae	0.000	0.067	Pseudomonadaceae <= 6.6e-06
4	Adenocarcinoma	0.728	0.348	0.618	0.659	Eggerthellaceae	0.000	0.044	Eggerthellaceae <= 0.000131
4	Adenocarcinoma	0.728	0.348	0.618	0.659	Carnobacteriaceae	0.038	-0.027	0.0311 < Carnobacteriaceae

4	Adenocarcinoma	0.728	0.348	0.618	0.659	Gemellaceae	0.013	0.026	0.0129 < Gemellaceae <= 0.0211
5	Adenocarcinoma	0.814	0.254	0.648	0.704	Pseudomonadaceae	0.000	-0.071	6.6e-06 < Pseudomonadaceae
5	Adenocarcinoma	0.814	0.254	0.648	0.704	Eggerthellaceae	0.000	0.039	Eggerthellaceae <= 0.000131
5	Adenocarcinoma	0.814	0.254	0.648	0.704	Clostridia_UCG.014	0.002	0.038	0.00172 < Clostridia_UCG.014 <= 0.00404
5	Adenocarcinoma	0.814	0.254	0.648	0.704	Actinomycetaceae	0.006	0.026	Actinomycetaceae <= 0.00601
5	Adenocarcinoma	0.814	0.254	0.648	0.704	Gemellaceae	0.035	0.024	0.0211 < Gemellaceae <= 0.0352
8	Adenocarcinoma	0.596	0.397	0.663	0.608	Pseudomonadaceae	0.000	0.068	Pseudomonadaceae <= 6.6e-06
8	Adenocarcinoma	0.596	0.397	0.663	0.608	Clostridia_UCG.014	0.016	-0.068	0.01058 < Clostridia_UCG.014
8	Adenocarcinoma	0.596	0.397	0.663	0.608	Actinomycetaceae	0.036	-0.055	0.02245 < Actinomycetaceae
8	Adenocarcinoma	0.596	0.397	0.663	0.608	Saccharimonadaceae	0.082	-0.026	0.0502 < Saccharimonadaceae
8	Adenocarcinoma	0.596	0.397	0.663	0.608	Gemellaceae	0.019	0.025	0.0129 < Gemellaceae <= 0.0211
9	Adenocarcinoma	0.626	0.197	0.605	0.693	Pseudomonadaceae	0.000	0.065	Pseudomonadaceae <= 6.6e-06
9	Adenocarcinoma	0.626	0.197	0.605	0.693	Eggerthellaceae	0.000	0.039	Eggerthellaceae <= 0.000131
9	Adenocarcinoma	0.626	0.197	0.605	0.693	Gemellaceae	0.034	0.023	0.0211 < Gemellaceae <= 0.0352
9	Adenocarcinoma	0.626	0.197	0.605	0.693	Corynebacteriaceae	0.005	-0.020	0.001784 < Corynebacteriaceae
9	Adenocarcinoma	0.626	0.197	0.605	0.693	Clostridia_UCG.014	0.010	-0.019	0.00404 < Clostridia_UCG.014 <= 0.01058
12	Adenocarcinoma	0.788	0.279	0.688	0.671	Pseudomonadaceae	0.000	-0.071	6.6e-06 < Pseudomonadaceae
12	Adenocarcinoma	0.788	0.279	0.688	0.671	Eggerthellaceae	0.000	-0.042	0.000131 < Eggerthellaceae
12	Adenocarcinoma	0.788	0.279	0.688	0.671	Clostridia_UCG.014	0.003	0.041	0.00172 < Clostridia_UCG.014 <= 0.00404
12	Adenocarcinoma	0.788	0.279	0.688	0.671	Gemellaceae	0.020	0.029	0.0129 < Gemellaceae <= 0.0211
12	Adenocarcinoma	0.788	0.279	0.688	0.671	Actinomycetaceae	0.009	0.027	0.00601 < Actinomycetaceae <= 0.01363
13	Adenocarcinoma	0.678	0.184	0.649	0.703	Clostridia_UCG.014	0.000	0.051	Clostridia_UCG.014 <= 0.00172
13	Adenocarcinoma	0.678	0.184	0.649	0.703	Eggerthellaceae	0.000	0.039	Eggerthellaceae <= 0.000131
13	Adenocarcinoma	0.678	0.184	0.649	0.703	Veillonellaceae	0.238	-0.032	0.1918 < Veillonellaceae
13	Adenocarcinoma	0.678	0.184	0.649	0.703	Carnobacteriaceae	0.056	-0.030	0.0311 < Carnobacteriaceae
13	Adenocarcinoma	0.678	0.184	0.649	0.703	Actinomycetaceae	0.010	0.026	0.00601 < Actinomycetaceae <= 0.01363
15	Adenocarcinoma	0.774	0.362	0.678	0.626	Pseudomonadaceae	0.000	-0.072	6.6e-06 < Pseudomonadaceae
15	Adenocarcinoma	0.774	0.362	0.678	0.626	Gemellaceae	0.039	-0.070	0.0352 < Gemellaceae
15	Adenocarcinoma	0.774	0.362	0.678	0.626	Clostridia_UCG.014	0.003	0.040	0.00172 < Clostridia_UCG.014 <= 0.00404
15	Adenocarcinoma	0.774	0.362	0.678	0.626	Eggerthellaceae	0.000	0.037	Eggerthellaceae <= 0.000131
15	Adenocarcinoma	0.774	0.362	0.678	0.626	Carnobacteriaceae	0.020	0.012	0.0154 < Carnobacteriaceae <= 0.0311
16	Adenocarcinoma	0.594	0.459	0.736	0.534	Gemellaceae	0.056	-0.069	0.0352 < Gemellaceae
16	Adenocarcinoma	0.594	0.459	0.736	0.534	Pseudomonadaceae	0.000	-0.070	6.6e-06 < Pseudomonadaceae
16	Adenocarcinoma	0.594	0.459	0.736	0.534	Actinomycetaceae	0.040	-0.059	0.02245 < Actinomycetaceae
16	Adenocarcinoma	0.594	0.459	0.736	0.534	Eggerthellaceae	0.000	-0.044	0.000131 < Eggerthellaceae

16	Adenocarcinoma	0.594	0.459	0.736	0.534	Clostridia_UCG.014	0.003	0.040	0.00172 < Clostridia_UCG.014 <= 0.00404
17	Adenocarcinoma	0.758	0.421	0.638	0.672	Pseudomonadaceae	0.000	0.070	Pseudomonadaceae <= 6.6e-06
17	Adenocarcinoma	0.758	0.421	0.638	0.672	Gemellaceae	0.039	-0.069	0.0352 < Gemellaceae
17	Adenocarcinoma	0.758	0.421	0.638	0.672	Clostridia_UCG.014	0.001	0.050	Clostridia_UCG.014 <= 0.00172
17	Adenocarcinoma	0.758	0.421	0.638	0.672	Eggerthellaceae	0.000	-0.043	0.000131 < Eggerthellaceae
17	Adenocarcinoma	0.758	0.421	0.638	0.672	Actinomycetaceae	0.002	0.025	Actinomycetaceae <= 0.00601
18	Adenocarcinoma	0.682	0.184	0.672	0.701	Clostridia_UCG.014	0.000	0.049	Clostridia_UCG.014 <= 0.00172
18	Adenocarcinoma	0.682	0.184	0.672	0.701	Eggerthellaceae	0.000	-0.046	0.000131 < Eggerthellaceae
18	Adenocarcinoma	0.682	0.184	0.672	0.701	Actinomycetaceae	0.003	0.031	Actinomycetaceae <= 0.00601
18	Adenocarcinoma	0.682	0.184	0.672	0.701	Veillonellaceae	0.220	-0.027	0.1918 < Veillonellaceae
18	Adenocarcinoma	0.682	0.184	0.672	0.701	Gemellaceae	0.001	0.023	Gemellaceae <= 0.0129
19	Adenocarcinoma	0.642	0.251	0.606	0.700	Pseudomonadaceae	0.000	0.065	Pseudomonadaceae <= 6.6e-06
19	Adenocarcinoma	0.642	0.251	0.606	0.700	Eggerthellaceae	0.000	0.040	Eggerthellaceae <= 0.000131
19	Adenocarcinoma	0.642	0.251	0.606	0.700	Clostridia_UCG.014	0.004	0.040	0.00172 < Clostridia_UCG.014 <= 0.00404
19	Adenocarcinoma	0.642	0.251	0.606	0.700	Neisseriaceae	0.000	-0.026	Neisseriaceae <= 0.0108
19	Adenocarcinoma	0.642	0.251	0.606	0.700	Carnobacteriaceae	0.062	-0.024	0.0311 < Carnobacteriaceae
27	Adenocarcinoma	0.860	0.367	0.601	0.700	Clostridia_UCG.014	0.016	-0.069	0.01058 < Clostridia_UCG.014
27	Adenocarcinoma	0.860	0.367	0.601	0.700	Pseudomonadaceae	0.000	0.069	Pseudomonadaceae <= 6.6e-06
27	Adenocarcinoma	0.860	0.367	0.601	0.700	Eggerthellaceae	0.000	0.046	Eggerthellaceae <= 0.000131
27	Adenocarcinoma	0.860	0.367	0.601	0.700	Actinomycetaceae	0.003	0.029	Actinomycetaceae <= 0.00601
27	Adenocarcinoma	0.860	0.367	0.601	0.700	Gemellaceae	0.030	0.025	0.0211 < Gemellaceae <= 0.0352
28	Adenocarcinoma	0.868	0.224	0.586	0.715	Pseudomonadaceae	0.000	0.067	Pseudomonadaceae <= 6.6e-06
28	Adenocarcinoma	0.868	0.224	0.586	0.715	Eggerthellaceae	0.000	0.041	Eggerthellaceae <= 0.000131
28	Adenocarcinoma	0.868	0.224	0.586	0.715	Gemellaceae	0.014	0.030	0.0129 < Gemellaceae <= 0.0211
28	Adenocarcinoma	0.868	0.224	0.586	0.715	Veillonellaceae	0.284	-0.027	0.1918 < Veillonellaceae
28	Adenocarcinoma	0.868	0.224	0.586	0.715	Moraxellaceae	0.000	0.018	Moraxellaceae <= 0.000218
30	Adenocarcinoma	0.766	0.237	0.591	0.733	Pseudomonadaceae	0.000	0.068	Pseudomonadaceae <= 6.6e-06
30	Adenocarcinoma	0.766	0.237	0.591	0.733	Eggerthellaceae	0.000	0.043	Eggerthellaceae <= 0.000131
30	Adenocarcinoma	0.766	0.237	0.591	0.733	Gemellaceae	0.019	0.030	0.0129 < Gemellaceae <= 0.0211
30	Adenocarcinoma	0.766	0.237	0.591	0.733	Actinomycetaceae	0.011	0.024	0.00601 < Actinomycetaceae <= 0.01363
30	Adenocarcinoma	0.766	0.237	0.591	0.733	Clostridia_UCG.014	0.007	-0.022	0.00404 < Clostridia_UCG.014 <= 0.01058
31	Adenocarcinoma	0.812	0.290	0.638	0.722	Pseudomonadaceae	0.000	-0.066	6.6e-06 < Pseudomonadaceae
31	Adenocarcinoma	0.812	0.290	0.638	0.722	Clostridia_UCG.014	0.001	0.050	Clostridia_UCG.014 <= 0.00172
31	Adenocarcinoma	0.812	0.290	0.638	0.722	Eggerthellaceae	0.000	0.045	Eggerthellaceae <= 0.000131
31	Adenocarcinoma	0.812	0.290	0.638	0.722	Actinomycetaceae	0.005	0.030	Actinomycetaceae <= 0.00601
31	Adenocarcinoma	0.812	0.290	0.638	0.722	Gemellaceae	0.024	0.025	0.0211 < Gemellaceae <= 0.0352
35	Adenocarcinoma	0.926	0.297	0.585	0.746	Pseudomonadaceae	0.000	0.071	Pseudomonadaceae <= 6.6e-06
35	Adenocarcinoma	0.926	0.297	0.585	0.746	Clostridia_UCG.014	0.002	0.051	Clostridia_UCG.014 <= 0.00172

35	Adenocarcinoma	0.926	0.297	0.585	0.746	Eggerthellaceae	0.000	0.039	Eggerthellaceae <= 0.000131
35	Adenocarcinoma	0.926	0.297	0.585	0.746	Veillonellaceae	0.198	-0.027	0.1918 < Veillonellaceae
35	Adenocarcinoma	0.926	0.297	0.585	0.746	Actinomycetaceae	0.004	0.026	Actinomycetaceae <= 0.00601
36	Adenocarcinoma	0.700	0.227	0.604	0.695	Pseudomonadaceae	0.000	0.070	Pseudomonadaceae <= 6.6e-06
36	Adenocarcinoma	0.700	0.227	0.604	0.695	Eggerthellaceae	0.000	0.040	Eggerthellaceae <= 0.000131
36	Adenocarcinoma	0.700	0.227	0.604	0.695	Gemellaceae	0.018	0.027	0.0129 < Gemellaceae <= 0.0211
36	Adenocarcinoma	0.700	0.227	0.604	0.695	Neisseriaceae	0.007	-0.024	Neisseriaceae <= 0.0108
36	Adenocarcinoma	0.700	0.227	0.604	0.695	Corynebacteriaceae	0.002	-0.022	0.001784 < Corynebacteriaceae
37	Adenocarcinoma	0.932	0.280	0.639	0.717	Pseudomonadaceae	0.000	-0.068	6.6e-06 < Pseudomonadaceae
37	Adenocarcinoma	0.932	0.280	0.639	0.717	Clostridia_UCG.014	0.000	0.049	Clostridia_UCG.014 <= 0.00172
37	Adenocarcinoma	0.932	0.280	0.639	0.717	Eggerthellaceae	0.000	0.043	Eggerthellaceae <= 0.000131
37	Adenocarcinoma	0.932	0.280	0.639	0.717	Actinomycetaceae	0.005	0.029	Actinomycetaceae <= 0.00601
37	Adenocarcinoma	0.932	0.280	0.639	0.717	Gemellaceae	0.008	0.025	Gemellaceae <= 0.0129
40	Adenocarcinoma	0.690	0.441	0.626	0.647	Pseudomonadaceae	0.000	0.066	Pseudomonadaceae <= 6.6e-06
40	Adenocarcinoma	0.690	0.441	0.626	0.647	Gemellaceae	0.043	-0.068	0.0352 < Gemellaceae
40	Adenocarcinoma	0.690	0.441	0.626	0.647	Actinomycetaceae	0.025	-0.060	0.02245 < Actinomycetaceae
40	Adenocarcinoma	0.690	0.441	0.626	0.647	Eggerthellaceae	0.000	0.041	Eggerthellaceae <= 0.000131
40	Adenocarcinoma	0.690	0.441	0.626	0.647	Clostridia_UCG.014	0.002	0.042	0.00172 < Clostridia_UCG.014 <= 0.00404
42	Adenocarcinoma	0.848	0.286	0.587	0.745	Pseudomonadaceae	0.000	0.067	Pseudomonadaceae <= 6.6e-06
42	Adenocarcinoma	0.848	0.286	0.587	0.745	Clostridia_UCG.014	0.000	0.050	Clostridia_UCG.014 <= 0.00172
42	Adenocarcinoma	0.848	0.286	0.587	0.745	Eggerthellaceae	0.000	0.041	Eggerthellaceae <= 0.000131
42	Adenocarcinoma	0.848	0.286	0.587	0.745	Neisseriaceae	0.006	-0.026	Neisseriaceae <= 0.0108
42	Adenocarcinoma	0.848	0.286	0.587	0.745	Gemellaceae	0.020	0.026	0.0129 < Gemellaceae <= 0.0211
43	Adenoma	0.532	0.34312123	0.324027822	0.380276176	Gemellaceae	0.067207523	0.069304936	0.0352 < Gemellaceae
43	Adenoma	0.532	0.34312123	0.324027822	0.380276176	Actinomycetaceae	0.023798549	0.057279719	0.02245 < Actinomycetaceae
43	Adenoma	0.532	0.34312123	0.324027822	0.380276176	Clostridia_UCG.014	6.5851E-05	-0.049673528	Clostridia_UCG.014 <= 0.00172
43	Adenoma	0.532	0.34312123	0.324027822	0.380276176	Eggerthellaceae	0	-0.03968751	Eggerthellaceae <= 0.000131
43	Adenoma	0.532	0.34312123	0.324027822	0.380276176	Corynebacteriaceae	0.003555954	0.019024738	0.001784 < Corynebacteriaceae
44	Adenocarcinoma	0.876	0.368491019	0.592823309	0.724221991	Pseudomonadaceae	0	0.071053134	Pseudomonadaceae <= 6.6e-06
44	Adenocarcinoma	0.876	0.368491019	0.592823309	0.724221991	Actinomycetaceae	0.02582363	-0.057827977	0.02245 < Actinomycetaceae
44	Adenocarcinoma	0.876	0.368491019	0.592823309	0.724221991	Clostridia_UCG.014	0	0.048033803	Clostridia_UCG.014 <= 0.00172
44	Adenocarcinoma	0.876	0.368491019	0.592823309	0.724221991	Eggerthellaceae	0	0.041353546	Eggerthellaceae <= 0.000131
44	Adenocarcinoma	0.876	0.368491019	0.592823309	0.724221991	Gemellaceae	0.013014227	0.028786177	0.0129 < Gemellaceae <= 0.0211
45	Adenocarcinoma	0.734	0.368859578	0.687096048	0.588189755	Pseudomonadaceae	0.000220119	-0.071356449	6.6e-06 < Pseudomonadaceae
45	Adenocarcinoma	0.734	0.368859578	0.687096048	0.588189755	Gemellaceae	0.039885538	-0.069614308	0.0352 < Gemellaceae
45	Adenocarcinoma	0.734	0.368859578	0.687096048	0.588189755	Eggerthellaceae	0	0.044119506	Eggerthellaceae <= 0.000131
45	Adenocarcinoma	0.734	0.368859578	0.687096048	0.588189755	Corynebacteriaceae	0.003918116	-0.024202313	0.001784 < Corynebacteriaceae
45	Adenocarcinoma	0.734	0.368859578	0.687096048	0.588189755	Actinomycetaceae	0.006911732	0.022147271	0.00601 < Actinomycetaceae <= 0.01363
48	Adenocarcinoma	0.63	0.433212125	0.706663918	0.526967119	Gemellaceae	0.044000815	-0.069211058	0.0352 < Gemellaceae

48	Adenocarcinoma	0.63	0.433212125	0.706663918	0.526967119	Pseudomonadaceae	0.000325932	-0.067572024	6.6e-06 < Pseudomonadaceae
48	Adenocarcinoma	0.63	0.433212125	0.706663918	0.526967119	Actinomycetaceae	0.025178244	-0.058124799	0.02245 < Actinomycetaceae
48	Adenocarcinoma	0.63	0.433212125	0.706663918	0.526967119	Eggerthellaceae	0	0.042359048	Eggerthellaceae <= 0.000131
48	Adenocarcinoma	0.63	0.433212125	0.706663918	0.526967119	Carnobacteriaceae	0.046404563	-0.027147966	0.0311 < Carnobacteriaceae
50	Adenocarcinoma	0.792	0.261607589	0.659722911	0.671443039	Pseudomonadaceae	6.28008E-05	-0.069790657	6.6e-06 < Pseudomonadaceae
50	Adenocarcinoma	0.792	0.261607589	0.659722911	0.671443039	Clostridia_UCG.014	0.003775895	0.044518222	0.00172 < Clostridia_UCG.014 <= 0.00404
50	Adenocarcinoma	0.792	0.261607589	0.659722911	0.671443039	Eggerthellaceae	0	0.040259098	Eggerthellaceae <= 0.000131
50	Adenocarcinoma	0.792	0.261607589	0.659722911	0.671443039	Saccharimonadaceae	0.083760509	-0.025003019	0.0502 < Saccharimonadaceae
50	Adenocarcinoma	0.792	0.261607589	0.659722911	0.671443039	Gemellaceae	0.025026102	0.021736484	0.0211 < Gemellaceae <= 0.0352



Figure ML3-D-6: The LIME Result in Training dataset for Family in Saliva

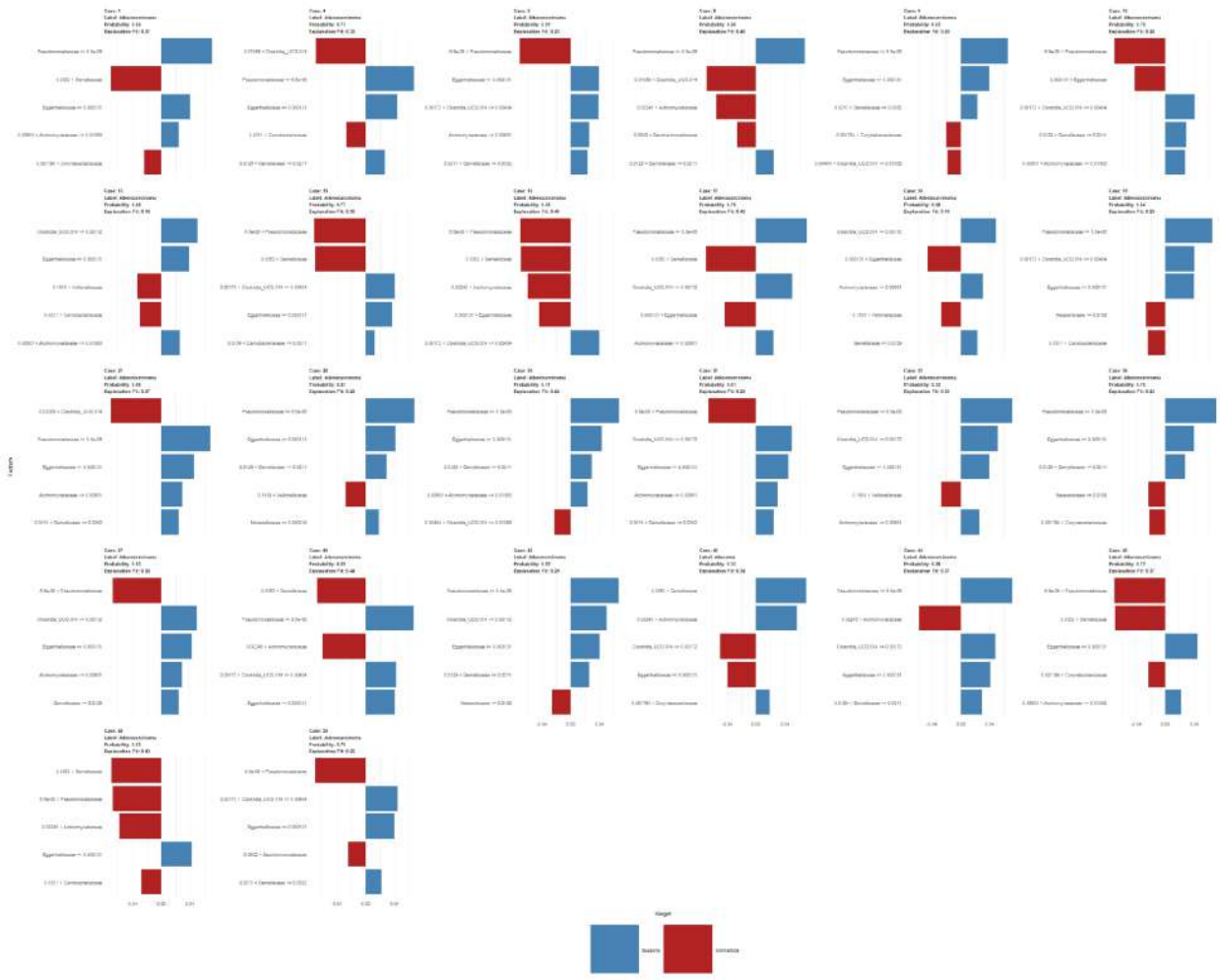


Figure ML3-D-7: The LIME Result in Test dataset for Family in Saliva

Table ML3-D-6: The Aggregated SHAP Values

Variable	Training Dataset				Testing Dataset			
	contribution	position	sign	cumulative	contribution	position	sign	cumulative
intercept	0.7594	59	X	0.759	0.7594	59	X	0.759
Yersiniaceae	0.0000	58	-1	0.759	0.0009	49	1	0.765
Comamonadaceae	0.0000	57	-1	0.759	0.0002	57	1	0.759
Gracilibacteria	0.0000	56	1	0.759	0.0000	55	1	0.762
Anaerovoracaceae	0.0001	55	1	0.759	0.0002	54	1	0.763

Flavobacteriaceae	-0.0001	54	-1	0.759	-0.0013	53	-1	0.761
Paludibacteraceae	0.0000	53	1	0.759	0.0001	52	1	0.761
Veillonellaceae	0.0000	52	1	0.759	0.0019	50	1	0.764
Clostridia_UCG.014	0.0004	51	1	0.760	0.0138	47	1	0.779
RF39	0.0000	50	1	0.760	-0.0006	58	-1	0.759
Fusobacteriaceae	0.0000	49	-1	0.760	0.0008	48	1	0.766
Mycoplasmataceae	-0.0001	48	-1	0.760	0.0008	51	1	0.762
Saccharimonadales	-0.0001	47	-1	0.760	0.0002	39	1	0.785
Tumor	0.0000	46	1	0.760	0.0000	44	1	0.779
Defluviitaleaceae	0.0000	45	-1	0.760	-0.0004	36	-1	0.782
Eggerthellaceae	0.0001	44	1	0.760	0.0033	56	1	0.762
Corynebacteriaceae	0.0000	43	-1	0.760	-0.0033	18	-1	0.753
Tannerellaceae	0.0001	42	1	0.760	0.0018	43	1	0.781
Pasteurellaceae	0.0000	41	1	0.760	0.0009	16	1	0.755
Neisseriaceae	-0.0001	40	-1	0.760	0.0029	42	1	0.784
Carnobacteriaceae	0.0000	39	1	0.760	-0.0028	37	-1	0.782
Prevotellaceae	0.0000	38	-1	0.760	-0.0003	45	-1	0.779
Bacteroidaceae	0.0000	37	1	0.760	0.0002	41	1	0.784
Campylobacteraceae	0.0000	36	-1	0.760	0.0009	40	1	0.785
Burkholderiaceae	0.0000	35	1	0.760	0.0002	23	1	0.770
Staphylococcaceae	0.0000	34	1	0.760	0.0001	38	1	0.785
Weeksellaceae	0.0000	33	1	0.760	-0.0006	20	-1	0.769
Spirochaetaceae	-0.0001	32	-1	0.760	-0.0004	32	-1	0.774
Porphyromonadaceae	-0.0002	31	-1	0.760	-0.0015	27	-1	0.771
Peptostreptococcaceae	0.0001	30	1	0.760	-0.0006	34	-1	0.782
Lentimicrobiaceae	0.0000	29	1	0.760	0.0003	35	1	0.782
Cardiobacteriaceae	-0.0001	28	-1	0.760	-0.0021	24	-1	0.770
Lachnospiraceae	0.0000	27	1	0.760	0.0007	26	1	0.771
Rikenellaceae	0.0000	26	-1	0.760	-0.0002	29	-1	0.772
Peptococcaceae	0.0000	25	-1	0.760	0.0014	7	1	0.752
Gender	0.0000	24	1	0.760	0.0000	22	1	0.770
Synergistaceae	-0.0001	23	-1	0.760	0.0009	25	1	0.772
Aerococcaceae	0.0000	22	-1	0.759	0.0003	5	1	0.747
Micrococcaceae	0.0000	21	-1	0.759	-0.0003	28	-1	0.772
Lactobacillaceae	0.0001	20	1	0.760	-0.0001	21	-1	0.770
Bacteroidales_Incertae_Sedis	0.0000	19	1	0.760	0.0000	46	-1	0.779
Peptostreptococcales.Tissierellales	0.0001	18	1	0.760	-0.0032	9	-1	0.751
Erysipelotrichaceae	0.0000	17	-1	0.760	0.0011	17	1	0.754
Atopobiaceae	0.0000	16	-1	0.760	-0.0009	8	-1	0.750
Streptococcaceae	-0.0001	15	-1	0.760	-0.0050	15	-1	0.750
Bacillaceae	0.0000	14	1	0.760	0.0006	14	1	0.750
Saccharimonadaceae	-0.0001	13	-1	0.759	0.0027	13	1	0.753
Chloroplast	0.0000	12	-1	0.759	-0.0003	31	-1	0.774

Clostridia_vadinBB60_group	0.0000	11	1	0.759	0.0008	12	1	0.754
JGI_0000069.P22	0.0000	10	-1	0.759	0.0002	11	1	0.754
Bifidobacteriaceae	0.0001	9	1	0.759	0.0003	10	1	0.754
Gemellaceae	0.0001	8	1	0.760	-0.0047	6	-1	0.747
Leptotrichiaceae	0.0000	7	1	0.760	-0.0016	30	-1	0.772
Moraxellaceae	-0.0001	6	-1	0.760	-0.0070	33	-1	0.775
Absconditabacteriales_.SR1.	-0.0001	5	-1	0.759	-0.0004	4	-1	0.747
Actinomycetaceae	-0.0003	4	-1	0.759	0.0003	3	1	0.747
Pseudomonadaceae	0.0001	3	1	0.759	-0.0133	19	-1	0.756
Selenomonadaceae	0.0001	2	1	0.759	-0.0001	2	-1	0.747
prediction	0.7594	1	X	0.759	0.7469	1	X	0.747

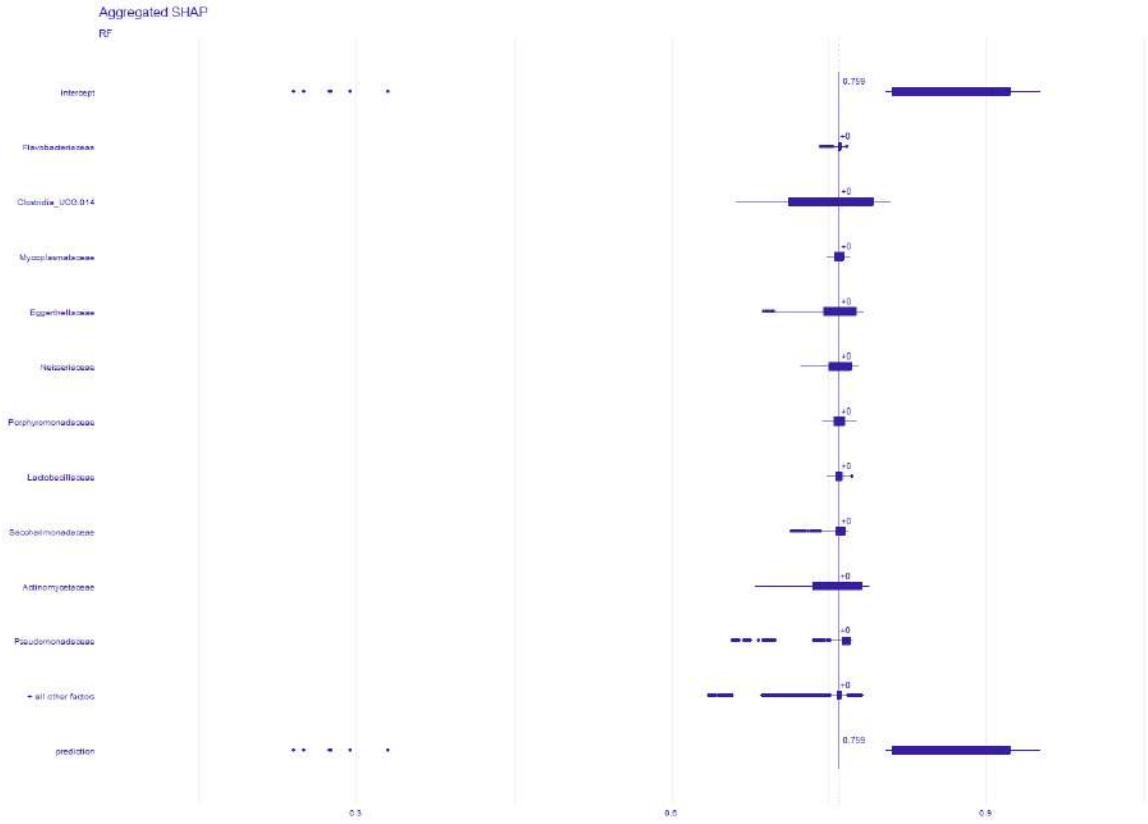


Figure ML3-D-8: The Aggregated SHAP Value in Training dataset for Family in Saliva

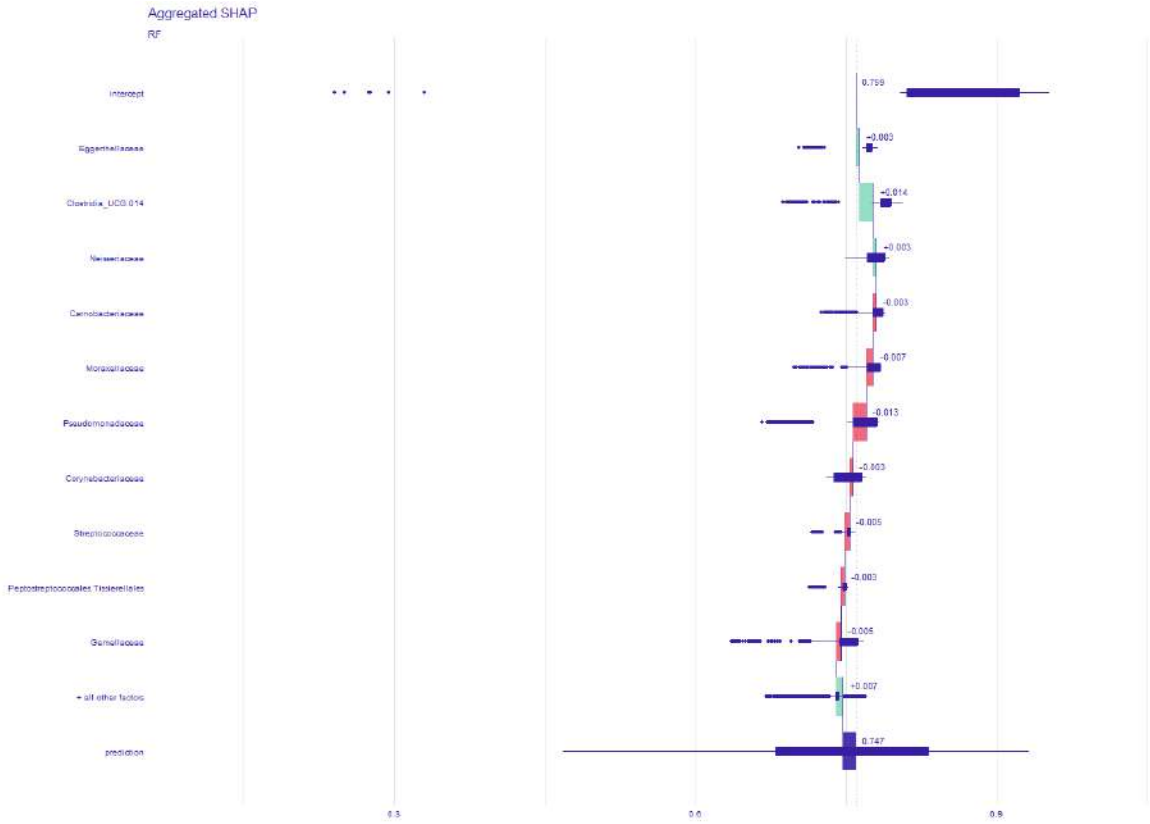
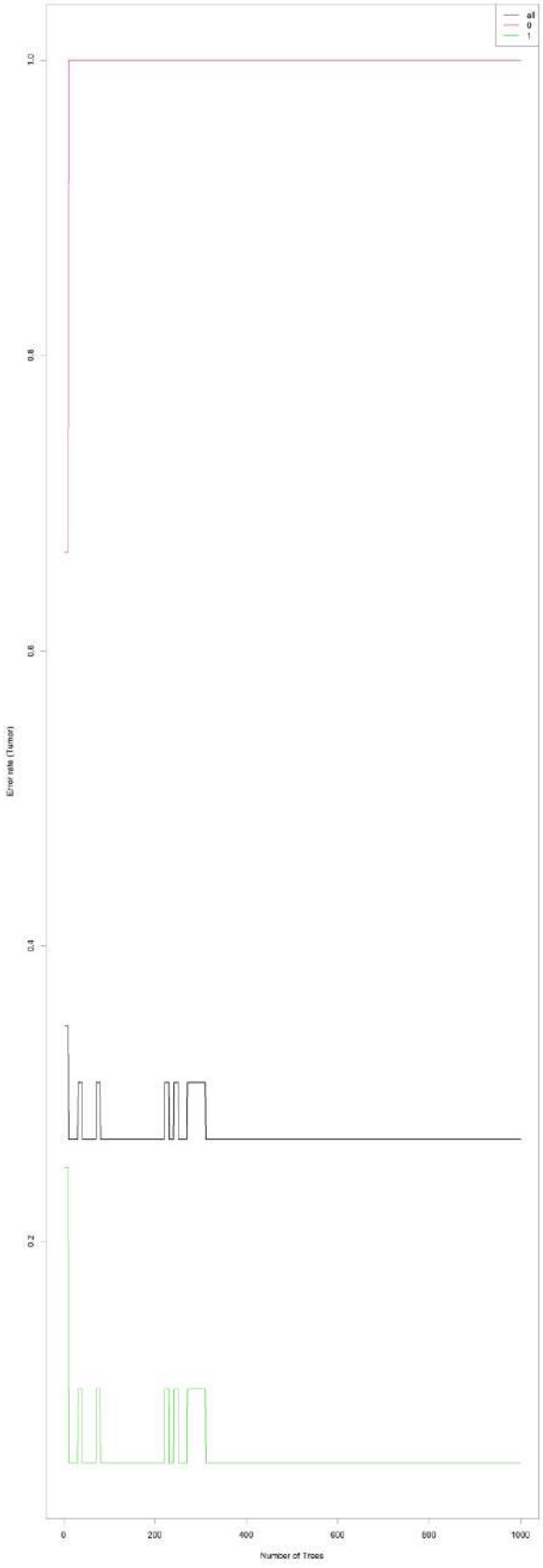


Figure ML3-D-9: The Aggregated SHAP Value in Test dataset for Family in Saliva

E- Genus


all
0
1
Error rate (Tumor)
0.2
0.4
0.6
0.8
1.0
0
200
400
600
800
1000
Number of Trees

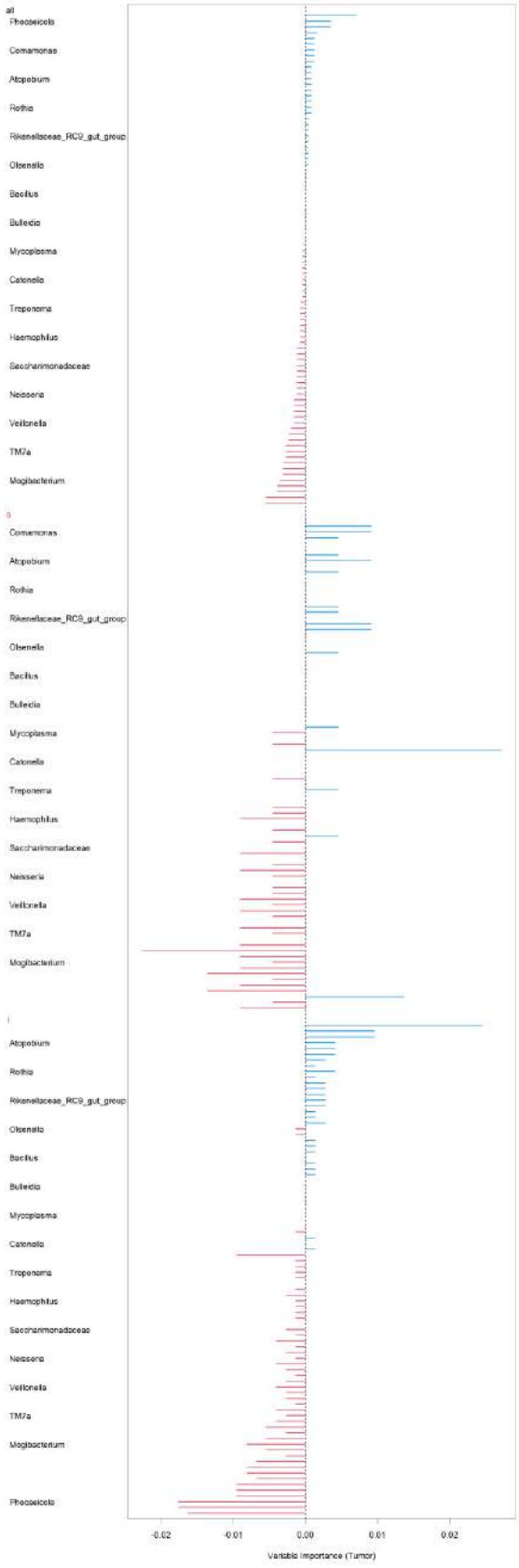
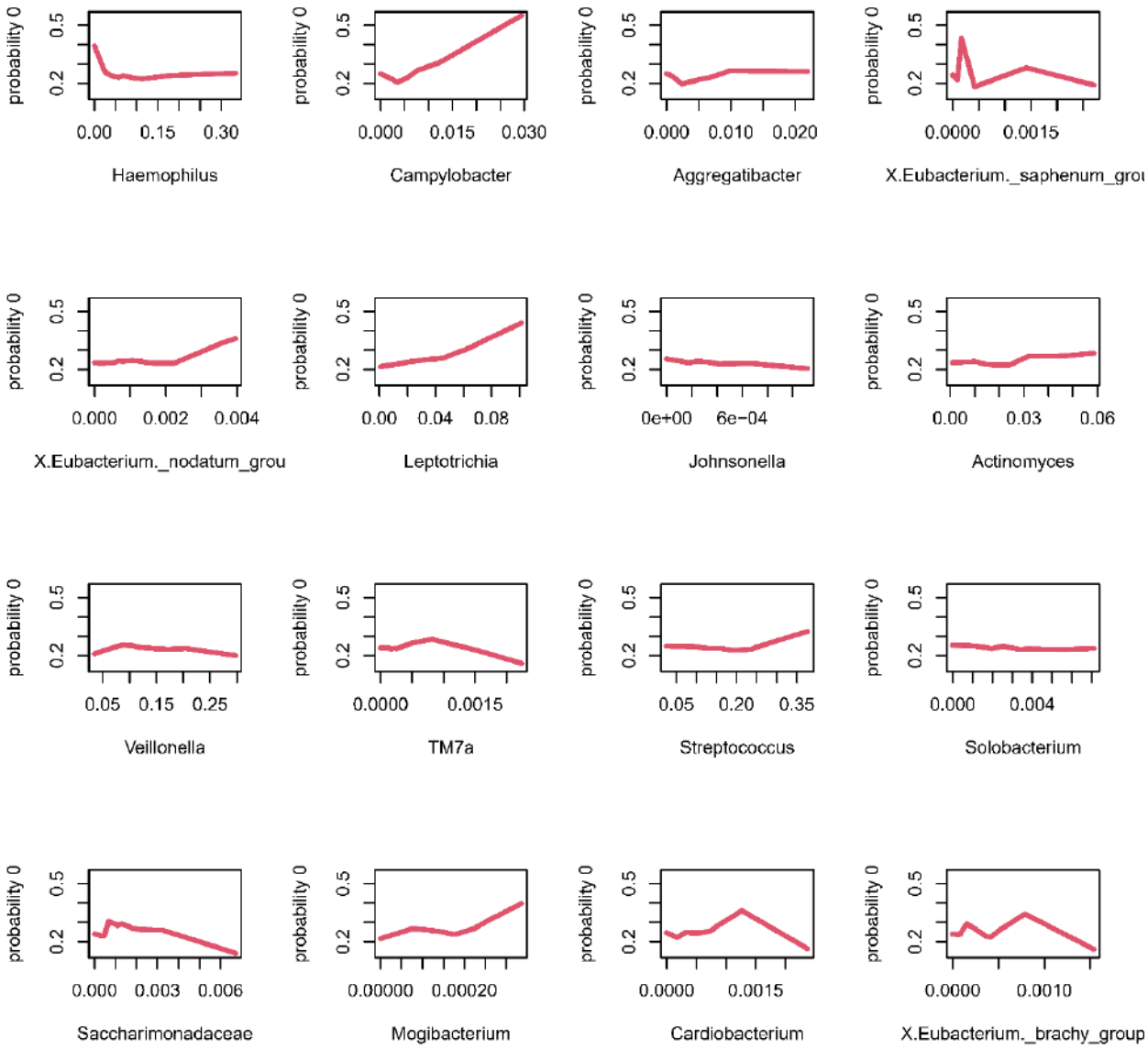
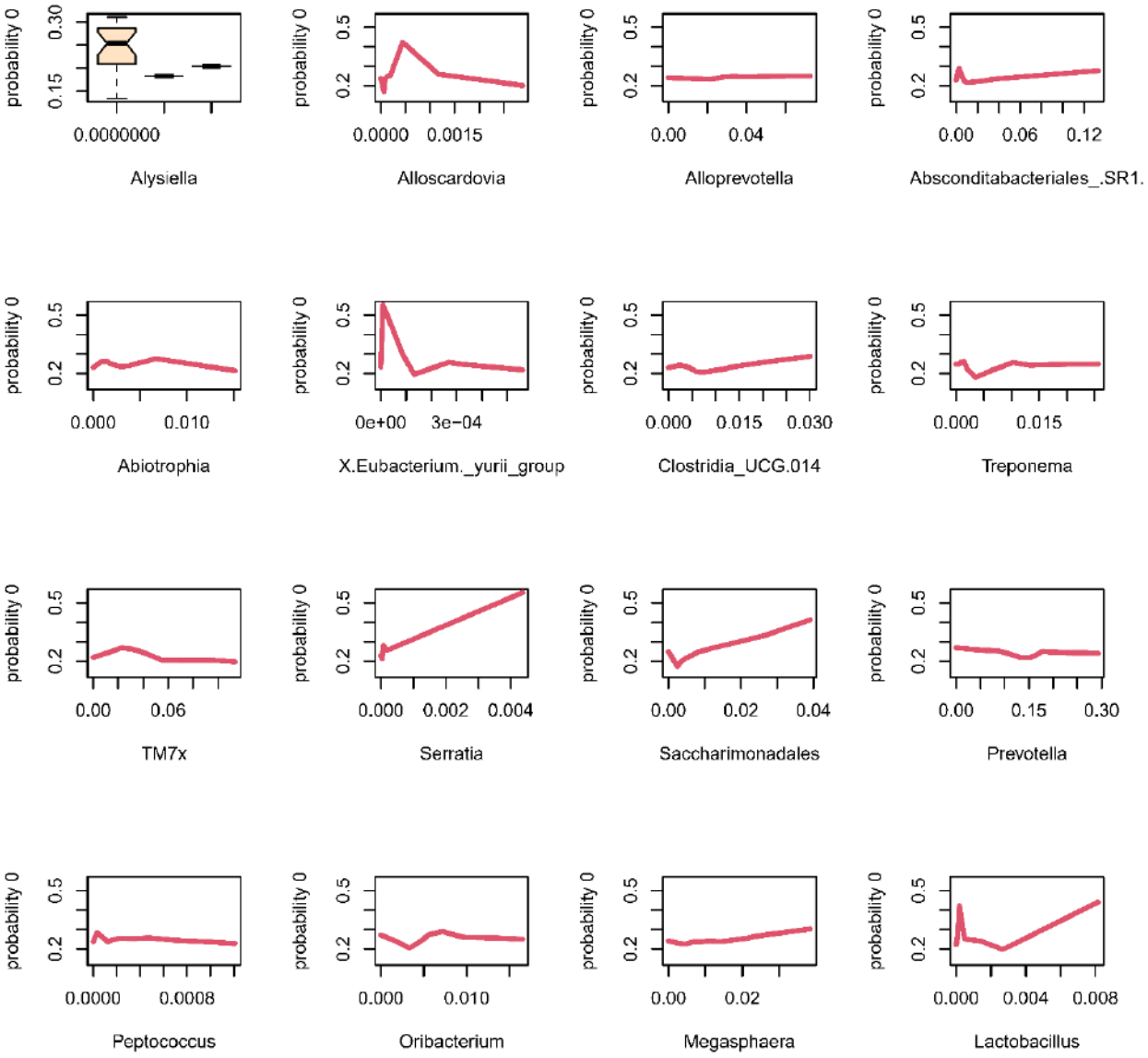
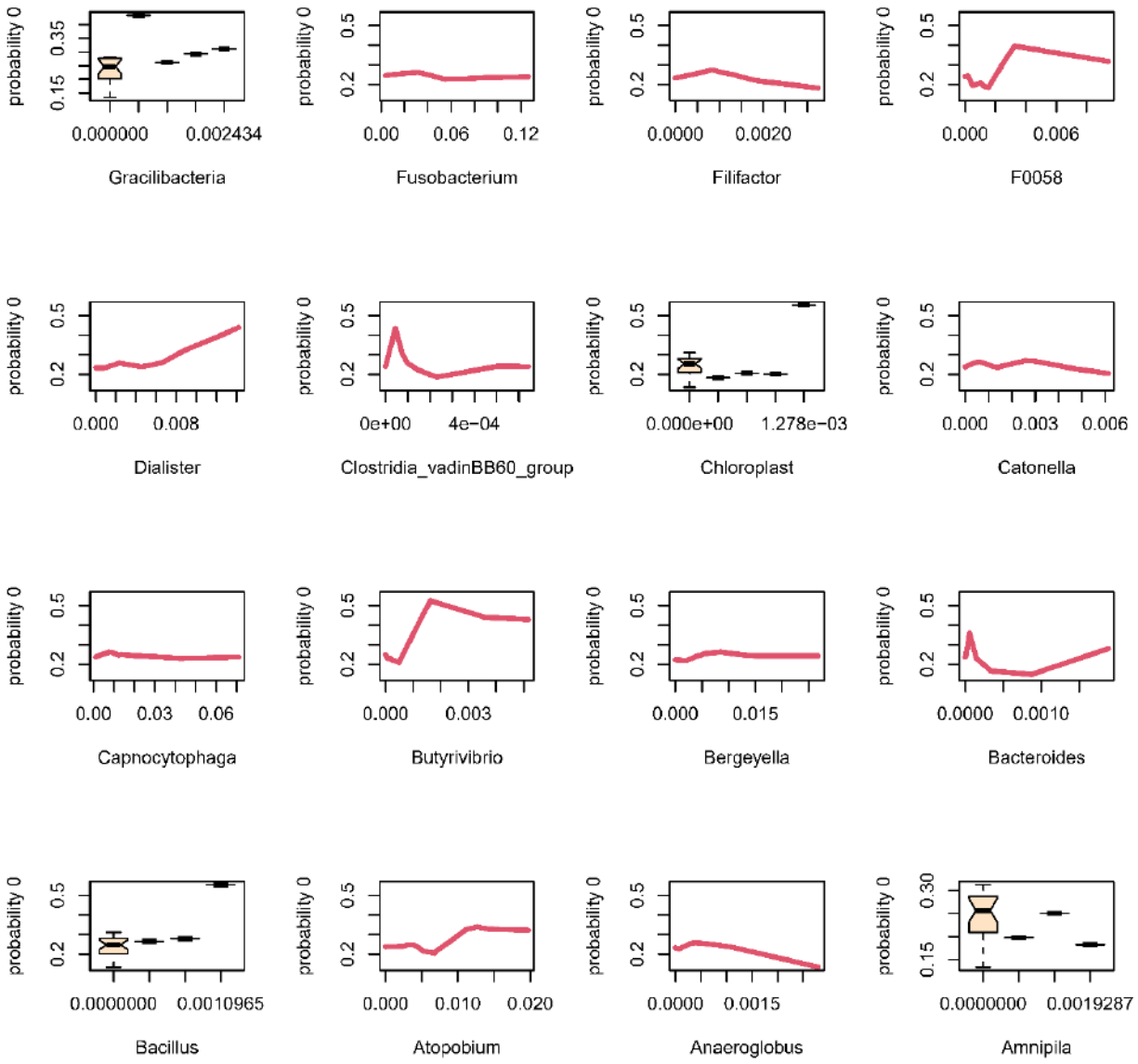
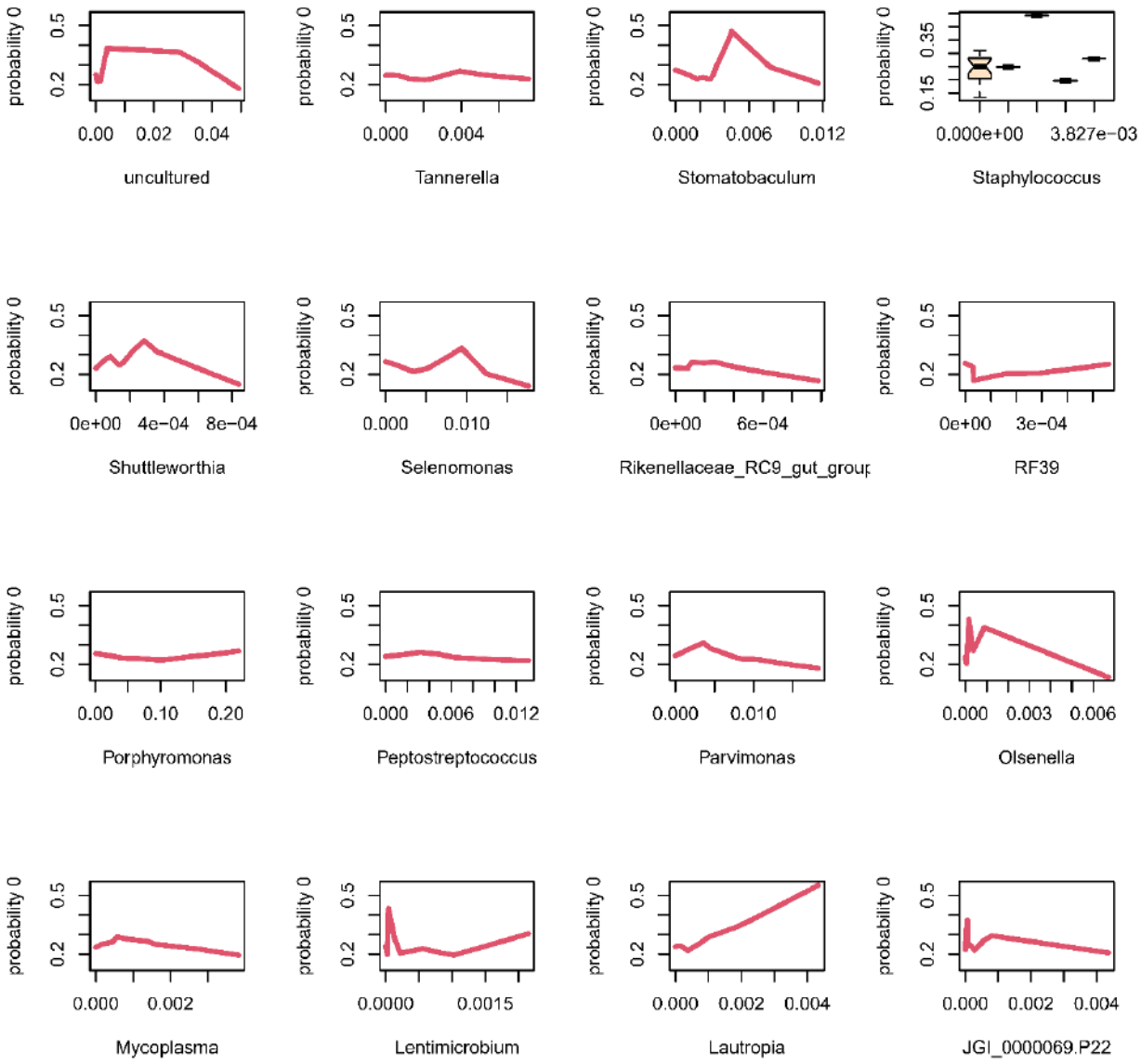

all
Phocaeicola
Comamonas
Atopobium
Rothia
Rikenellaceae_RC9_gut_group
Olsenella
Bacillus
Bulleidia
Mycoplasma
Catonella
Treponema
Haemophilus
Saccharimonadaceae
Neisseria
Veillonella
TM7x
Mogibacterium
0
Comamonas
Atopobium
Rothia
Rikenellaceae_RC9_gut_group
Olsenella
Bacillus
Bulleidia
Mycoplasma
Catonella
Treponema
Haemophilus
Saccharimonadaceae
Neisseria
Veillonella
TM7x
Mogibacterium
1
Atopobium
Rothia
Rikenellaceae_RC9_gut_group
Olsenella
Bacillus
Bulleidia
Mycoplasma
Catonella
Treponema
Haemophilus
Saccharimonadaceae
Neisseria
Veillonella
TM7x
Mogibacterium
Phocaeicola
-0.02
-0.01
0.00
0.01
0.02
Variable Importance (Tumor)

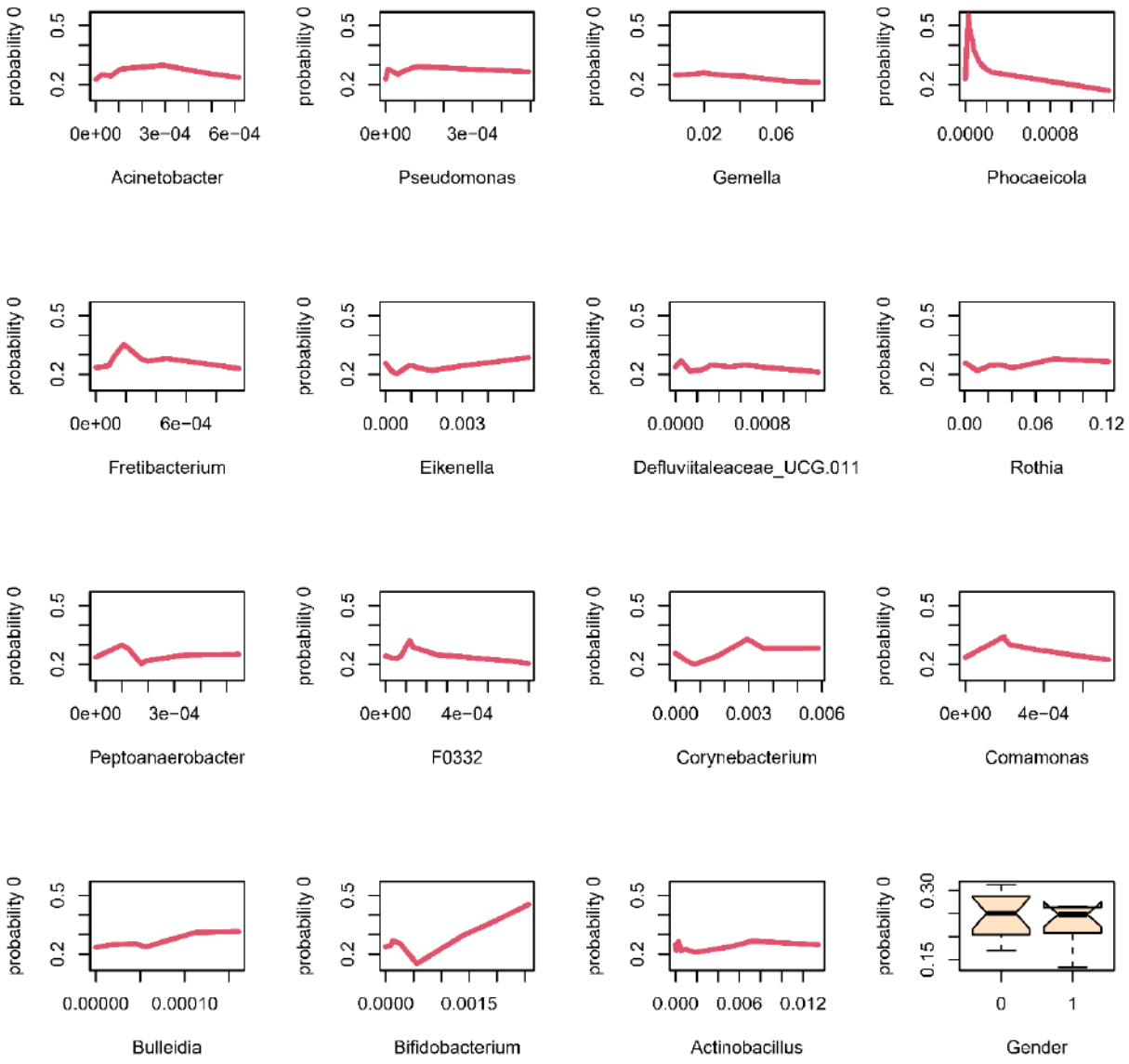
Figure ML3-E-1: The number of trees and Variable Importance Plot for Genus in Saliva


probability 0
0.2
0.5
0.00 0.15 0.30
Haemophilus
0.000 0.015 0.030
Campylobacter
0.000 0.010 0.020
Aggregatibacter
0.0000 0.0015
X.Eubacterium._saphenum_grou
0.000 0.002 0.004
X.Eubacterium._nodatum_grou
0.00 0.04 0.08
Leptotrichia
0e+00 6e-04
Johnsonella
0.00 0.03 0.06
Actinomyces
0.05 0.15 0.25
Veillonella
0.0000 0.0015
TM7a
0.05 0.20 0.35
Streptococcus
0.000 0.004
Solobacterium
0.000 0.003 0.006
Saccharimonadaceae
0.00000 0.00020
Mogibacterium
0.0000 0.0015
Cardiobacterium
0.0000 0.0010
X.Eubacterium._brachy_group


probability 0
0.15
0.30
0.0000000
Alysiella
probability 0
0.2
0.5
0.0000
0.0015
Alloscardovia
probability 0
0.2
0.5
0.00
0.04
Alloprevotella
probability 0
0.2
0.5
0.00
0.06
0.12
Absconditabacteriales_.SR1.
probability 0
0.2
0.5
0.000
0.010
Abiotrophia
probability 0
0.2
0.5
0e+00
3e−04
X.Eubacterium._yurii_group
probability 0
0.2
0.5
0.000
0.015
0.030
Clostridia_UCG.014
probability 0
0.2
0.5
0.000
0.015
Treponema
probability 0
0.2
0.5
0.00
0.06
TM7x
probability 0
0.2
0.5
0.000
0.002
0.004
Serratia
probability 0
0.2
0.5
0.00
0.02
0.04
Saccharimonadales
probability 0
0.2
0.5
0.00
0.15
0.30
Prevotella
probability 0
0.2
0.5
0.0000
0.0008
Peptococcus
probability 0
0.2
0.5
0.000
0.010
Oribacterium
probability 0
0.2
0.5
0.00
0.02
Megasphaera
probability 0
0.2
0.5
0.000
0.004
0.008
Lactobacillus


probability 0
0.15 0.35
0.000000 0.002434
Gracilibacteria
probability 0
0.2 0.5
0.00 0.06 0.12
Fusobacterium
probability 0
0.2 0.5
0.0000 0.0020
Filifactor
probability 0
0.2 0.5
0.000 0.006
F0058
probability 0
0.2 0.5
0.000 0.008
Dialister
probability 0
0.2 0.5
0e+00 4e−04
Clostridia_vadinBB60_group
probability 0
0.2 0.5
0.000e+00 1.278e−03
Chloroplast
probability 0
0.2 0.5
0.000 0.003 0.006
Catonella
probability 0
0.2 0.5
0.00 0.03 0.06
Capnocytophaga
probability 0
0.2 0.5
0.000 0.003
Butyrivibrio
probability 0
0.2 0.5
0.000 0.015
Bergeyella
probability 0
0.2 0.5
0.0000 0.0010
Bacteroides
probability 0
0.2 0.5
0.0000000 0.0010965
Bacillus
probability 0
0.2 0.5
0.000 0.010 0.020
Atopobium
probability 0
0.2 0.5
0.0000 0.0015
Anaeroglobus
probability 0
0.15 0.30
0.0000000 0.0019287
Amnipila


probability 0
0.2
0.5
0.00
0.02
0.04
uncultured
0.000
0.004
Tannerella
0.000
0.006
0.012
Stomatobaculum
0.15
0.35
0.000e+00
3.827e−03
Staphylococcus
0e+00
4e−04
8e−04
Shuttleworthia
0.000
0.010
Selenomonas
0e+00
6e−04
Rikenellaceae_RC9_gut_grou
0e+00
3e−04
RF39
0.00
0.10
0.20
Porphyromonas
0.000
0.006
0.012
Peptostreptococcus
0.000
0.010
Parvimonas
0.000
0.003
0.006
Olsenella
0.000
0.002
Mycoplasma
0.0000
0.0015
Lentimicrobium
0.000
0.002
0.004
Lautropia
0.000
0.002
0.004
JGI_0000069.P22


probability 0
0.2
0.5
0e+00
3e−04
6e−04
Acinetobacter
0e+00
3e−04
Pseudomonas
0.02
0.06
Gemella
0.0000
0.0008
Phocaeicola
0e+00
6e−04
Fretibacterium
0.000
0.003
Eikenella
0.0000
0.0008
Defluviitaleaceae_UCG.011
0.00
0.06
0.12
Rothia
0e+00
3e−04
Peptoanaerobacter
0e+00
4e−04
F0332
0.000
0.003
0.006
Corynebacterium
0e+00
4e−04
Comamonas
0.00000
0.00010
Bulleidia
0.0000
0.0015
Bifidobacterium
0.000
0.006
0.012
Actinobacillus
0.15
0.30
0
1
Gender

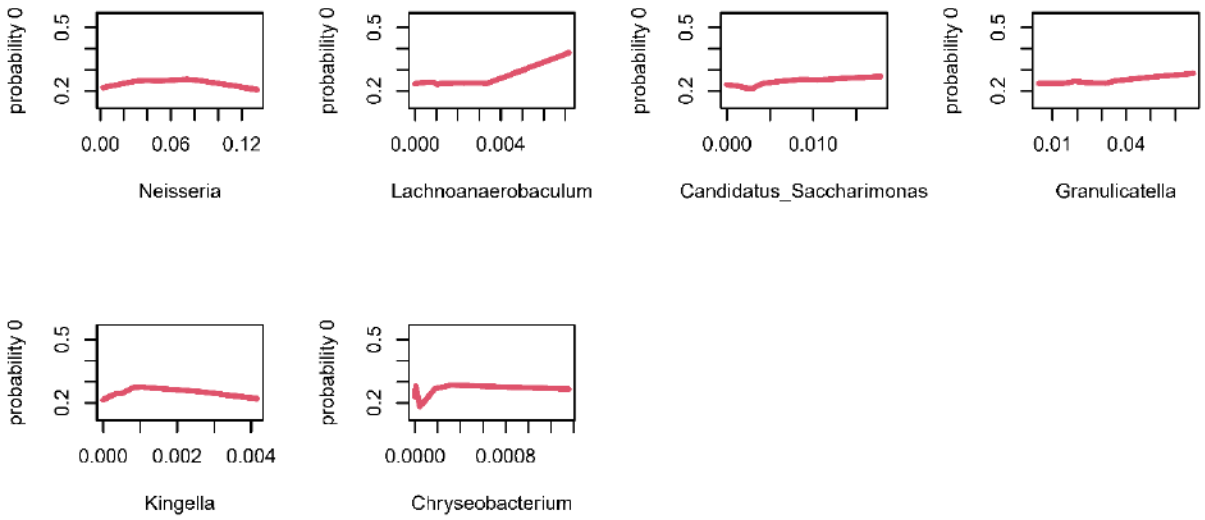
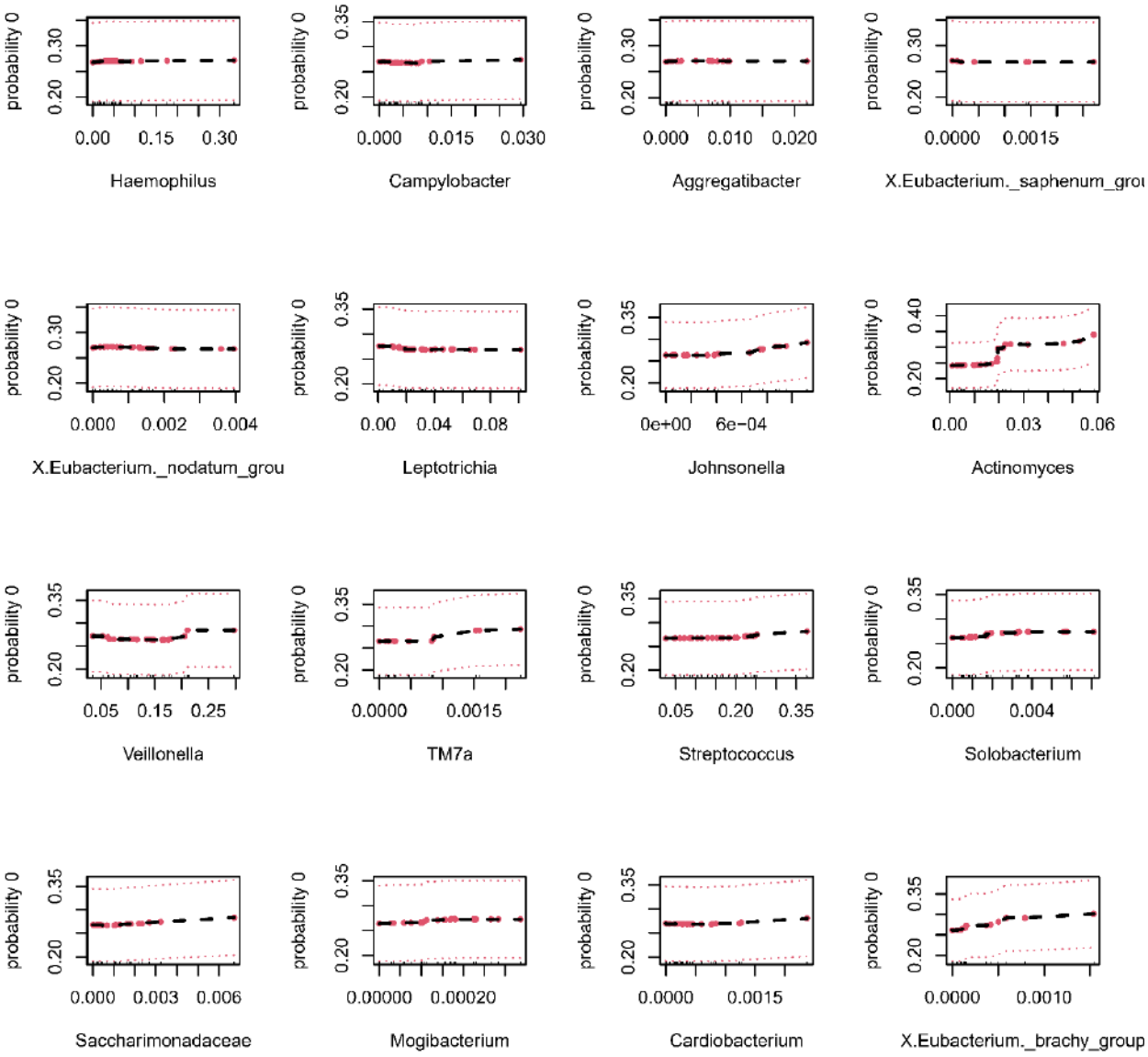
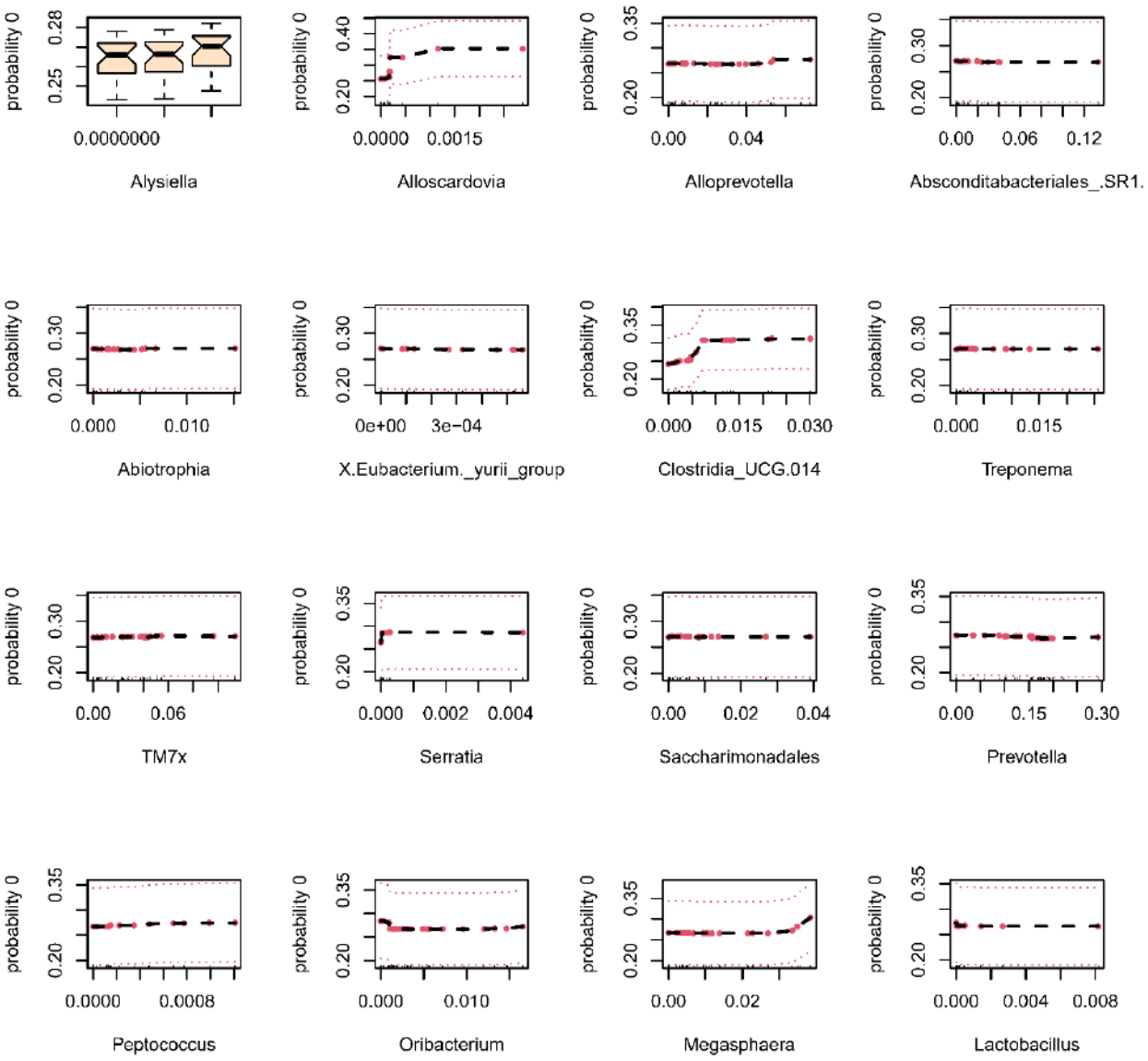
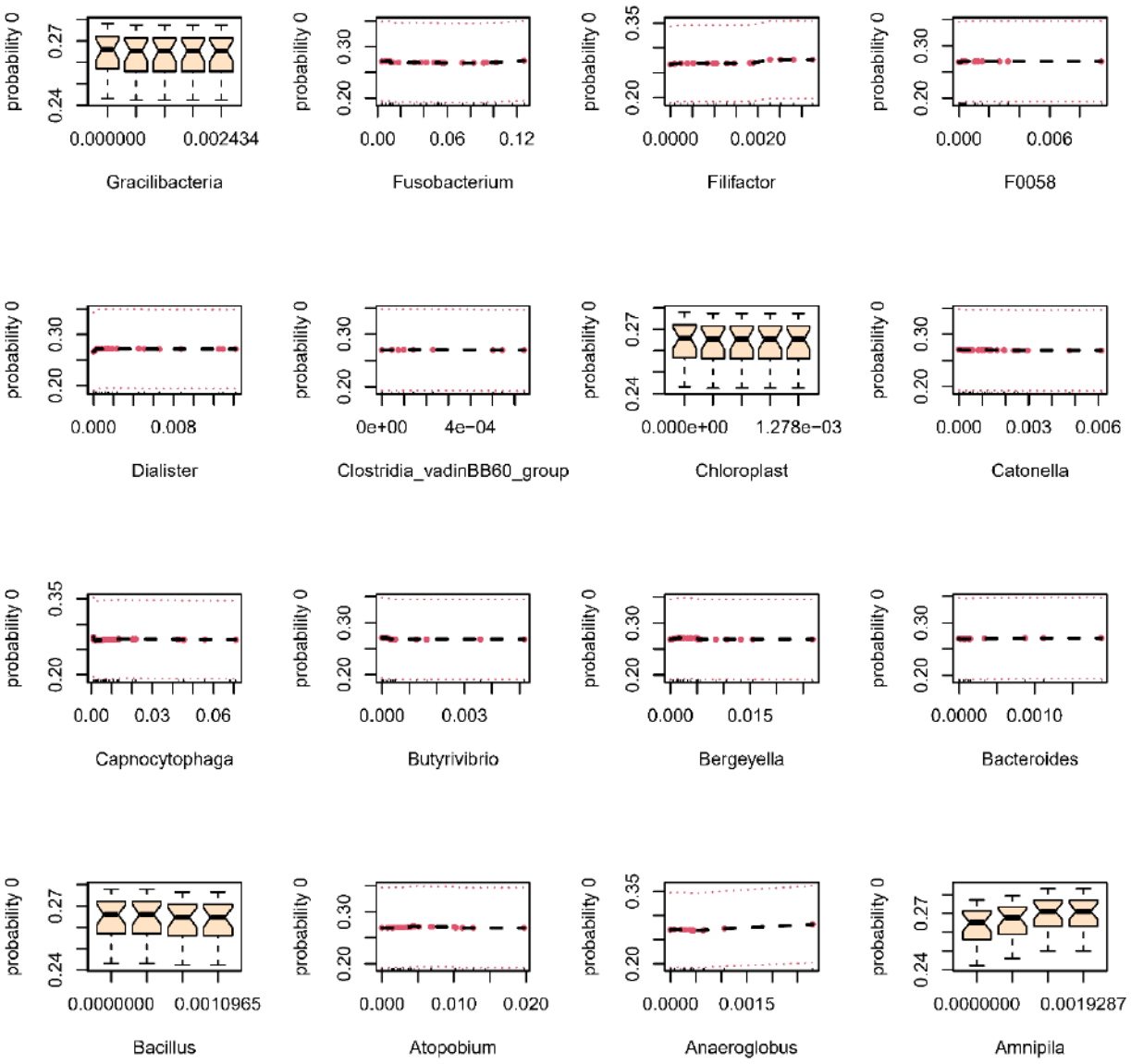
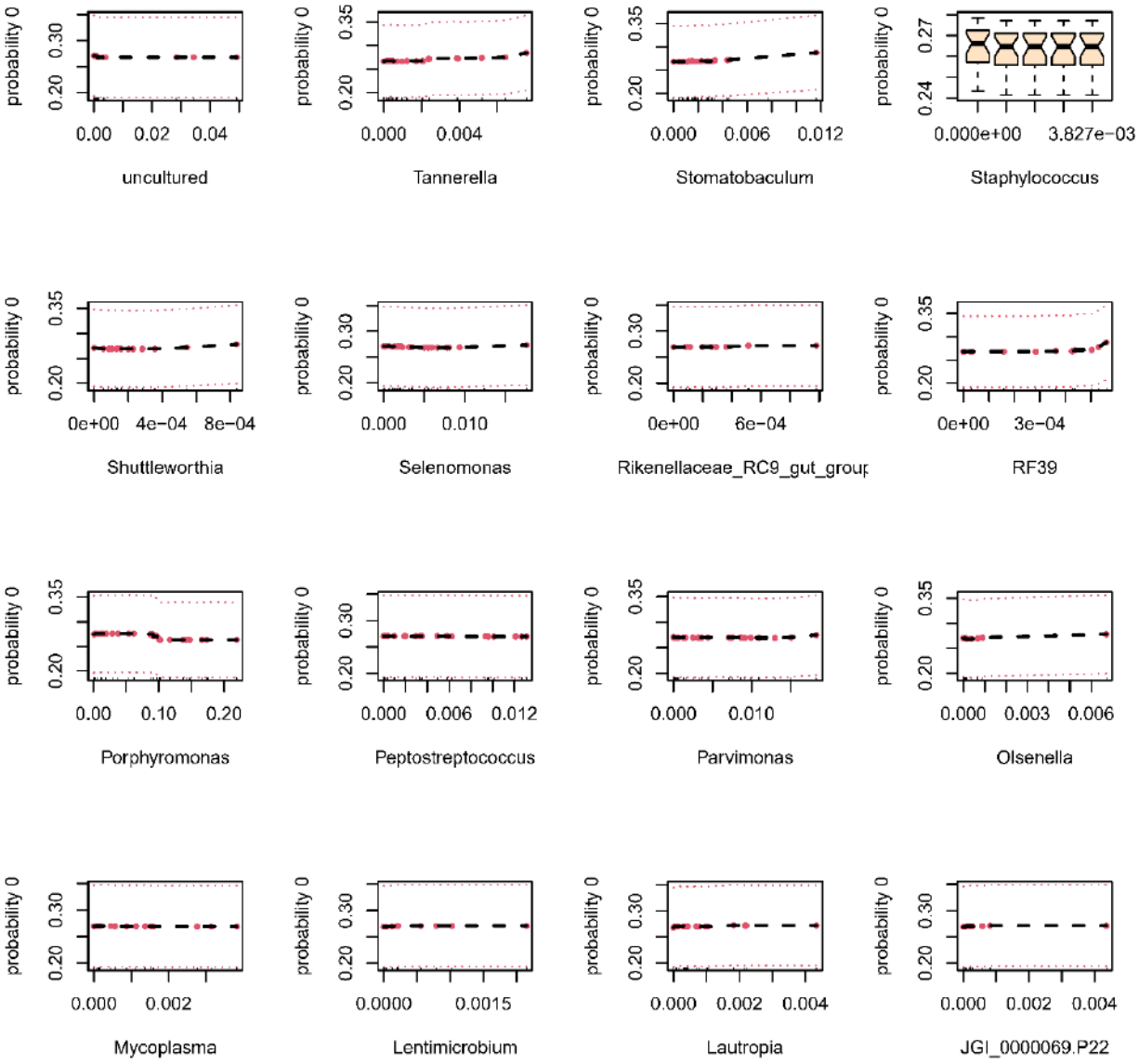


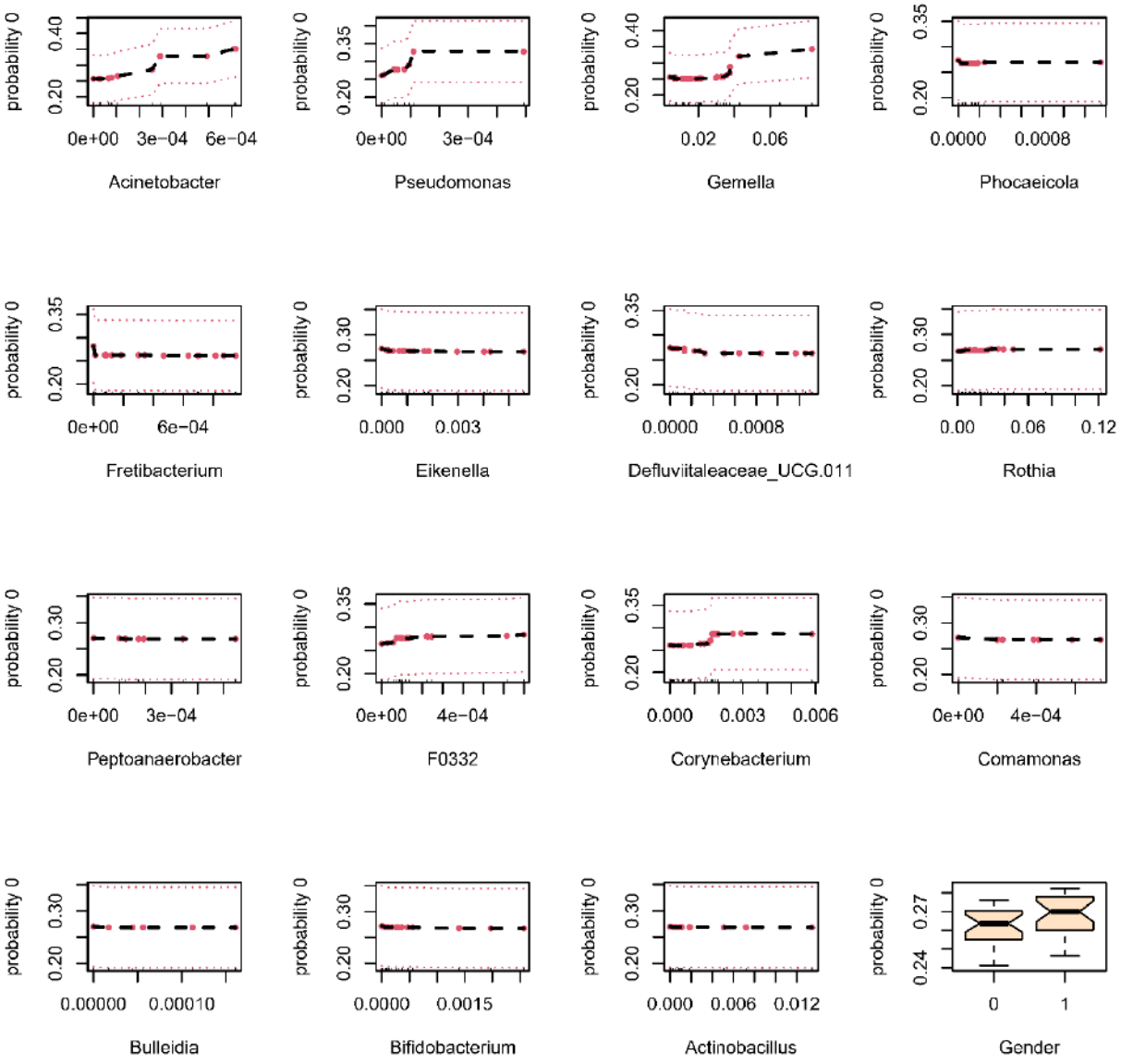
Figure ML3-E-2: Marginal Plot for Genus in Saliva


probability 0
Haemophilus
Campylobacter
Aggregatibacter
X.Eubacterium._saphenum_grou
X.Eubacterium._nodatum_grou
Leptotrichia
Johnsonella
Actinomyces
Veillonella
TM7a
Streptococcus
Solobacterium
Saccharimonadaceae
Mogibacterium
Cardiobacterium
X.Eubacterium._brachy_group


probability 0
0.25
0.28
0.0000000
Alysiella
probability 0
0.20
0.40
0.0000
0.0015
Alloscardovia
probability 0
0.20
0.35
0.00
0.04
Alloprevotella
probability 0
0.20
0.30
0.00
0.06
0.12
Absconditabacteriales_.SR1.
probability 0
0.20
0.30
0.000
0.010
Abiotrophia
probability 0
0.20
0.30
0e+00
3e−04
X.Eubacterium._yurii_group
probability 0
0.20
0.35
0.000
0.015
0.030
Clostridia_UCG.014
probability 0
0.20
0.30
0.000
0.015
Treponema
probability 0
0.20
0.30
0.00
0.06
TM7x
probability 0
0.20
0.35
0.000
0.002
0.004
Serratia
probability 0
0.20
0.30
0.00
0.02
0.04
Saccharimonadales
probability 0
0.20
0.35
0.00
0.15
0.30
Prevotella
probability 0
0.20
0.35
0.0000
0.0008
Peptococcus
probability 0
0.20
0.35
0.000
0.010
Oribacterium
probability 0
0.20
0.35
0.00
0.02
Megasphaera
probability 0
0.20
0.35
0.000
0.004
0.008
Lactobacillus


probability 0
0.24
0.27
0.000000
0.002434
Gracilibacteria
probability 0
0.20
0.30
0.00
0.06
0.12
Fusobacterium
probability 0
0.20
0.35
0.0000
0.0020
Filifactor
probability 0
0.20
0.30
0.000
0.006
F0058
probability 0
0.20
0.30
0.000
0.008
Dialister
probability 0
0.20
0.30
0e+00
4e−04
Clostridia_vadinBB60_group
probability 0
0.24
0.27
0.000e+00
1.278e−03
Chloroplast
probability 0
0.20
0.30
0.000
0.003
0.006
Catonella
probability 0
0.20
0.35
0.00
0.03
0.06
Capnocytophaga
probability 0
0.20
0.30
0.000
0.003
Butyrivibrio
probability 0
0.20
0.30
0.000
0.015
Bergeyella
probability 0
0.20
0.30
0.0000
0.0010
Bacteroides
probability 0
0.24
0.27
0.0000000
0.0010965
Bacillus
probability 0
0.20
0.30
0.000
0.010
0.020
Atopobium
probability 0
0.20
0.35
0.0000
0.0015
Anaeroglobus
probability 0
0.24
0.27
0.0000000
0.0019287
Amnipila


probability 0
uncultured
Tannerella
Stomatobaculum
Staphylococcus
Shuttleworthia
Selenomonas
Rikenellaceae_RC9_gut_group
RF39
Porphyromonas
Peptostreptococcus
Parvimonas
Olsenella
Mycoplasma
Lentimicrobium
Lautropia
JGI_0000069.P22


probability 0
0.20
0.40
0e+00
3e−04
6e−04
Acinetobacter
probability 0
0.20
0.35
0e+00
3e−04
Pseudomonas
probability 0
0.20
0.40
0.02
0.06
Gemella
probability 0
0.20
0.35
0.0000
0.0008
Phocaeicola
probability 0
0.20
0.35
0e+00
6e−04
Fretibacterium
probability 0
0.20
0.30
0.000
0.003
Eikenella
probability 0
0.20
0.35
0.0000
0.0008
Defluviitaleaceae_UCG.011
probability 0
0.20
0.30
0.00
0.06
0.12
Rothia
probability 0
0.20
0.30
0e+00
3e−04
Peptoanaerobacter
probability 0
0.20
0.35
0e+00
4e−04
F0332
probability 0
0.20
0.35
0.000
0.003
0.006
Corynebacterium
probability 0
0.20
0.30
0e+00
4e−04
Comamonas
probability 0
0.20
0.30
0.00000
0.00010
Bulleidia
probability 0
0.20
0.30
0.0000
0.0015
Bifidobacterium
probability 0
0.20
0.30
0.000
0.006
0.012
Actinobacillus
probability 0
0.24
0.27
0
1
Gender

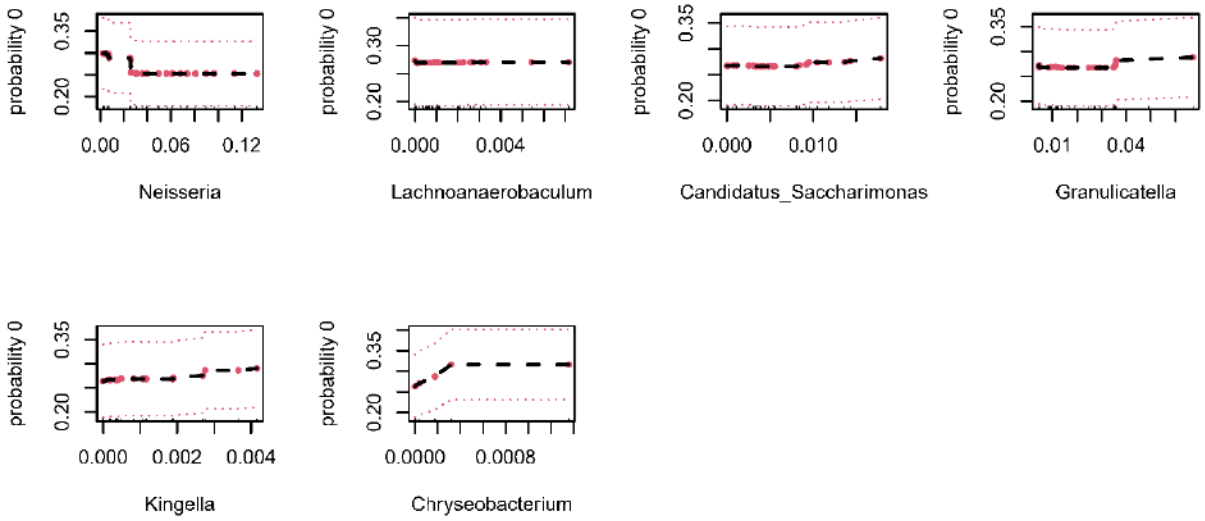


Figure ML3-E-3: Partial Plot for Genus in Saliva

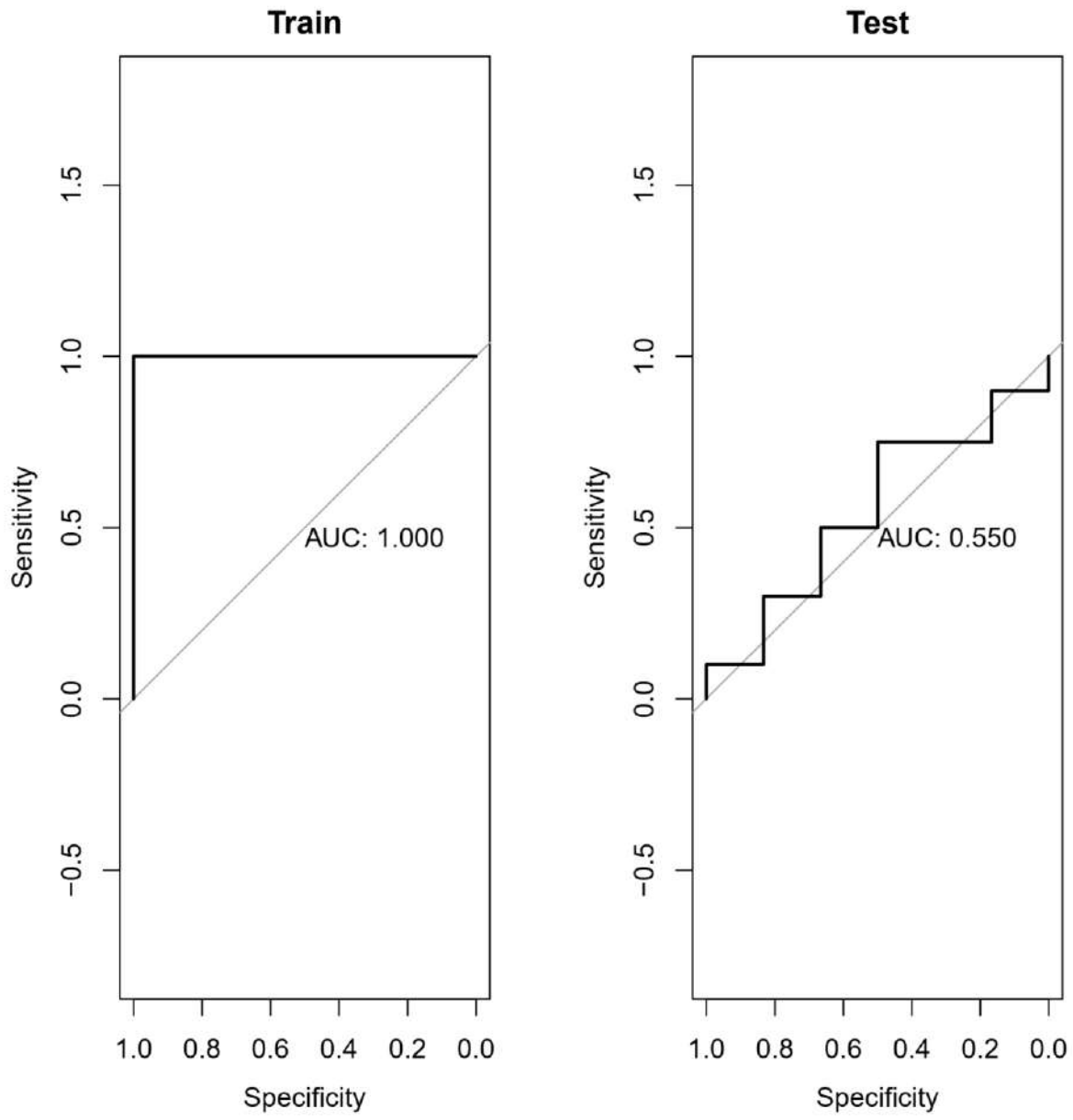


Figure ML3-E-4: The ROC Curve for Genus in Saliva

Table ML3-E-1: The Model Performance

Index	Values		
	2.50%	50.00%	97.50%
Accuracy	0.27	0.69	0.88
Specificity	0.33	0.67	1.00
Sensitivity	0.05	0.70	0.95

Table ML3-E-2: The Variable Importance Values

Covariates	Varibale Importance			Brier Score		
	all	AP	CRC	all	AP	CRC
Gender	-0.0019	-0.0181	-0.0014	-0.0047	-0.0032	-0.0032
X.Eubacterium._brachy_group	0.0023	-0.1087	0.0408	-0.0011	-0.0008	-0.0008
X.Eubacterium._nodatum_group	0.0004	0.0000	0.0014	-0.0075	-0.0051	-0.0051
X.Eubacterium._saphenum_group	-0.0004	0.0000	-0.0014	-0.0004	-0.0003	-0.0003
X.Eubacterium._yurii_group	-0.0012	0.0000	-0.0041	-0.0034	-0.0023	-0.0023
Abiotrophia	0.0012	-0.0045	0.0054	-0.0050	-0.0034	-0.0034
Absconditabacteriales_.SR1.	0.0000	0.0000	0.0000	-0.0008	-0.0006	-0.0006
Acinetobacter	0.0329	-0.1722	0.1672	0.0674	0.0458	0.0458
Actinobacillus	-0.0012	-0.0045	-0.0027	-0.0012	-0.0008	-0.0008
Actinomyces	0.0318	-0.0362	0.1223	0.0857	0.0582	0.0582
Aggregatibacter	0.0004	0.0000	0.0014	0.0002	0.0001	0.0001
Alloprevotella	0.0000	-0.0045	0.0014	0.0011	0.0007	0.0007
Alloscardovia	0.0499	-0.1314	0.2134	0.1107	0.0752	0.0752
Alysiella	0.0020	0.0000	0.0068	0.0029	0.0020	0.0020
Amnipila	-0.0020	-0.0091	-0.0041	-0.0040	-0.0027	-0.0027
Anaeroglobus	0.0027	-0.0362	0.0204	-0.0013	-0.0009	-0.0009
Atopobium	0.0004	0.0000	0.0014	-0.0010	-0.0007	-0.0007
Bacillus	-0.0008	0.0000	-0.0027	-0.0008	-0.0005	-0.0005
Bacteroides	0.0000	-0.0045	0.0014	-0.0007	-0.0005	-0.0005
Bergeyella	-0.0012	0.0045	-0.0054	-0.0054	-0.0037	-0.0037
Bifidobacterium	0.0012	0.0000	0.0041	0.0046	0.0031	0.0031
Bulleidia	0.0008	0.0045	0.0014	0.0002	0.0001	0.0001
Butyrivibrio	-0.0004	0.0000	-0.0014	0.0010	0.0007	0.0007
Campylobacter	-0.0012	-0.0227	0.0027	-0.0033	-0.0022	-0.0022
Candidatus_Saccharimonas	0.0015	-0.0272	0.0136	-0.0074	-0.0050	-0.0050
Capnocytophaga	0.0019	-0.0181	0.0122	-0.0005	-0.0003	-0.0003
Cardiobacterium	-0.0023	-0.0408	0.0041	0.0015	0.0010	0.0010
Catonella	-0.0008	-0.0045	-0.0014	-0.0043	-0.0029	-0.0029
Chloroplast	-0.0004	0.0000	-0.0014	-0.0020	-0.0014	-0.0014
Chryseobacterium	0.0206	-0.0951	0.1006	0.0502	0.0341	0.0341
Clostridia_UCG.014	0.0069	-0.0272	0.0326	0.0648	0.0440	0.0440

Clostridia_vadinBB60_group	0.0000	0.0000	0.0000	-0.0016	-0.0011	-0.0011
Comamonas	-0.0008	0.0000	-0.0027	0.0020	0.0014	0.0014
Corynebacterium	0.0000	-0.0227	0.0068	0.0059	0.0040	0.0040
Defluviitaleaceae_UCG.011	-0.0016	-0.0045	-0.0041	-0.0020	-0.0014	-0.0014
Dialister	-0.0035	0.0045	-0.0136	0.0003	0.0002	0.0002
Eikenella	-0.0004	0.0000	-0.0014	-0.0010	-0.0007	-0.0007
F0058	-0.0008	0.0000	-0.0027	-0.0043	-0.0029	-0.0029
F0332	0.0062	-0.0045	0.0231	0.0209	0.0142	0.0142
Filifactor	-0.0008	-0.0091	0.0000	-0.0001	-0.0001	-0.0001
Fretibacterium	0.0070	0.0136	0.0204	0.0211	0.0144	0.0144
Fusobacterium	0.0023	-0.0136	0.0122	0.0002	0.0002	0.0002
Gemella	0.0307	-0.1404	0.1495	0.1077	0.0732	0.0732
Gracilibacteria	-0.0004	0.0000	-0.0014	-0.0016	-0.0011	-0.0011
Granulicatella	0.0081	-0.0272	0.0367	0.0150	0.0102	0.0102
Haemophilus	-0.0004	-0.0045	0.0000	-0.0055	-0.0037	-0.0037
JGI_0000069.P22	0.0000	-0.0045	0.0014	-0.0018	-0.0012	-0.0012
Johnsonella	0.0058	-0.0544	0.0367	0.0118	0.0080	0.0080
Kingella	0.0060	-0.0453	0.0340	0.0071	0.0048	0.0048
Lachnoanaerobaculum	0.0012	-0.0091	0.0068	-0.0031	-0.0021	-0.0021
Lactobacillus	0.0000	-0.0091	0.0027	-0.0060	-0.0041	-0.0041
Lautropia	0.0012	0.0000	0.0041	0.0025	0.0017	0.0017
Lentimicrobium	0.0004	0.0000	0.0014	-0.0018	-0.0012	-0.0012
Leptotrichia	0.0019	-0.0136	0.0109	-0.0024	-0.0017	-0.0017
Megasphaera	0.0155	-0.0680	0.0748	0.0218	0.0148	0.0148
Mogibacterium	-0.0043	-0.0091	-0.0122	-0.0166	-0.0113	-0.0113
Mycoplasma	-0.0012	0.0000	-0.0041	-0.0022	-0.0015	-0.0015
Neisseria	0.0085	-0.0272	0.0381	0.0093	0.0063	0.0063
Olsenella	0.0023	-0.0272	0.0163	0.0041	0.0028	0.0028
Oribacterium	0.0082	-0.0634	0.0476	0.0001	0.0001	0.0001
Parvimonas	0.0028	-0.0091	0.0122	0.0000	0.0000	0.0000
Peptoanaerobacter	-0.0008	-0.0045	-0.0014	0.0005	0.0004	0.0004
Peptococcus	-0.0004	-0.0091	0.0014	-0.0067	-0.0045	-0.0045
Peptostreptococcus	0.0000	0.0000	0.0000	0.0009	0.0006	0.0006
Phocaeicola	0.0066	-0.0091	0.0258	0.0153	0.0104	0.0104
Porphyromonas	-0.0034	-0.0045	-0.0109	0.0038	0.0026	0.0026
Prevotella	-0.0008	-0.0272	0.0054	-0.0040	-0.0027	-0.0027
Pseudomonas	0.0244	-0.0815	0.1101	0.0605	0.0411	0.0411
RF39	0.0058	-0.0408	0.0326	-0.0013	-0.0009	-0.0009
Rikenellaceae_RC9_gut_group	0.0012	-0.0045	0.0054	-0.0018	-0.0012	-0.0012
Rothia	-0.0027	0.0091	-0.0122	-0.0085	-0.0058	-0.0058
Saccharimonadaceae	0.0016	-0.0634	0.0245	0.0109	0.0074	0.0074
Saccharimonadales	-0.0012	-0.0045	-0.0027	-0.0051	-0.0035	-0.0035
Selenomonas	0.0011	-0.0181	0.0095	-0.0003	-0.0002	-0.0002

Serratia	0.0034	-0.0136	0.0163	0.0086	0.0058	0.0058
Shuttleworthia	0.0031	-0.0362	0.0217	-0.0027	-0.0018	-0.0018
Solobacterium	-0.0032	0.0091	-0.0136	-0.0055	-0.0038	-0.0038
Staphylococcus	-0.0004	0.0045	-0.0027	-0.0011	-0.0008	-0.0008
Stomatobaculum	0.0001	-0.0770	0.0231	-0.0019	-0.0013	-0.0013
Streptococcus	-0.0004	-0.0408	0.0109	-0.0054	-0.0037	-0.0037
Tannerella	0.0036	-0.0272	0.0204	0.0011	0.0007	0.0007
TM7a	0.0067	-0.0634	0.0421	0.0050	0.0034	0.0034
TM7x	-0.0019	-0.0091	-0.0041	-0.0093	-0.0063	-0.0063
Treponema	0.0000	0.0045	-0.0014	-0.0042	-0.0029	-0.0029
uncultured	-0.0016	0.0045	-0.0068	-0.0021	-0.0014	-0.0014
Veillonella	0.0015	-0.0680	0.0258	-0.0015	-0.0010	-0.0010

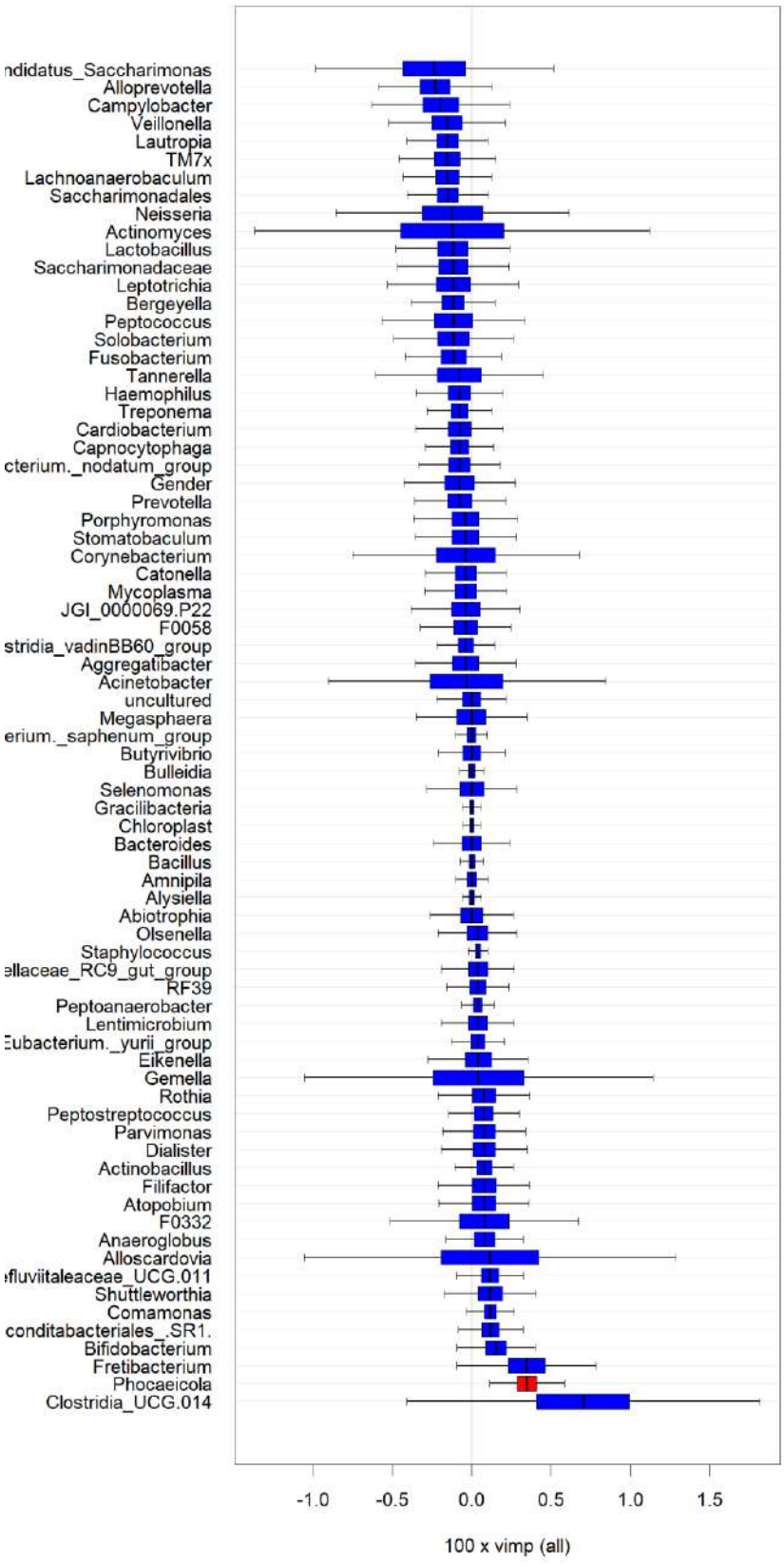

ndidatus_Saccharimonas
Alloprevotella
Campylobacter
Veillonella
Lautropia
TM7x
Lachnoanaerobaculum
Saccharimonadales
Neisseria
Actinomyces
Lactobacillus
Saccharimonadaceae
Leptotrichia
Bergeyella
Peptococcus
Solobacterium
Fusobacterium
Tannerella
Haemophilus
Treponema
Cardiobacterium
Capnocytophaga
cterium._nodatum_group
Gender
Prevotella
Porphyromonas
Stomatobaculum
Corynebacterium
Catonella
Mycoplasma
JGI_0000069.P22
F0058
stridia_vadinBB60_group
Aggregatibacter
Acinetobacter
uncultured
Megasphaera
erium._saphenum_group
Butyrivibrio
Bulleidia
Selenomonas
Gracilibacteria
Chloroplast
Bacteroides
Bacillus
Amnipila
Alysiella
Abiotrophia
Olsenella
Staphylococcus
ellaceae_RC9_gut_group
RF39
Peptoanaerobacter
Lentimicrobium
Eubacterium._yurii_group
Eikenella
Gemella
Rothia
Peptostreptococcus
Parvimonas
Dialister
Actinobacillus
Filifactor
Atopobium
F0332
Anaeroglobus
Alloscardovia
efluviitaleaceae_UCG.011
Shuttleworthia
Comamonas
conditabacteriales_.SR1.
Bifidobacterium
Fretibacterium
Phocaeicola
Clostridia_UCG.014
-1.0
-0.5
0.0
0.5
1.0
1.5
100 x vimp (all)

Figure ML3-E-5: The Variable Importance Plot (VIMP) for Genus in Saliva

Best model is selected for LIME results

Table ML3-E-3: The best setting for model and LIME Results

Setting	mtry	ROC	Sens	Spec	ROCSD	SensSD	SpecSD
ntrees: 500 nodesize: 1	8	0.35	0	1	0.285044	0	0
ntrees: 1000 nodesize: 1	9	0.3	0	0.95	0.32596	0	0.111803
ntrees: 500 nodesize: 5	10	0.35	0	0.95	0.285044	0	0.111803
ntrees: 1000 nodesize: 5	8	0.3	0	0.95	0.32596	0	0.111803
ntrees: 1000 nodesize: 5	10	0.3	0	0.95	0.32596	0	0.111803

Table ML3-E-4: The LIME Result for Training Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
2	Adenoma	0.678	0.329	0.312	0.440	Alloscardovia	0.003	0.062	0.000122 < Alloscardovia
2	Adenoma	0.678	0.329	0.312	0.440	Gemella	0.043	0.052	0.0352 < Gemella
2	Adenoma	0.678	0.329	0.312	0.440	Pseudomonas	0.000	-0.032	Pseudomonas <= 6.6e-06
2	Adenoma	0.678	0.329	0.312	0.440	Acinetobacter	0.001	0.026	7.52e-05 < Acinetobacter
2	Adenoma	0.678	0.329	0.312	0.440	X.Eubacterium._brachy_group	0.001	0.020	1.51e-04 < X.Eubacterium._brachy_group
3	Adenoma	0.660	0.403	0.333	0.433	Alloscardovia	0.000	-0.059	Alloscardovia <= 0.000122
3	Adenoma	0.660	0.403	0.333	0.433	Actinomyces	0.053	0.055	0.02229 < Actinomyces
3	Adenoma	0.660	0.403	0.333	0.433	Clostridia_UCG.014	0.012	0.042	0.01058 < Clostridia_UCG.014
3	Adenoma	0.660	0.403	0.333	0.433	Pseudomonas	0.000	0.034	6.6e-06 < Pseudomonas
3	Adenoma	0.660	0.403	0.333	0.433	Acinetobacter	0.000	0.027	7.52e-05 < Acinetobacter
6	Adenoma	0.690	0.343	0.348	0.387	Alloscardovia	0.000	-0.059	Alloscardovia <= 0.000122
6	Adenoma	0.690	0.343	0.348	0.387	Actinomyces	0.058	0.055	0.02229 < Actinomyces
6	Adenoma	0.690	0.343	0.348	0.387	Pseudomonas	0.000	0.033	6.6e-06 < Pseudomonas
6	Adenoma	0.690	0.343	0.348	0.387	Acinetobacter	0.001	0.031	7.52e-05 < Acinetobacter
6	Adenoma	0.690	0.343	0.348	0.387	Clostridia_UCG.014	0.002	-0.022	0.00172 < Clostridia_UCG.014 <= 0.00404
7	Adenoma	0.686	0.358	0.400	0.361	Alloscardovia	0.000	-0.062	Alloscardovia <= 0.000122
7	Adenoma	0.686	0.358	0.400	0.361	Gemella	0.065	0.054	0.0352 < Gemella
7	Adenoma	0.686	0.358	0.400	0.361	Pseudomonas	0.000	-0.031	Pseudomonas <= 6.6e-06
7	Adenoma	0.686	0.358	0.400	0.361	Corynebacterium	0.003	0.030	0.001785 < Corynebacterium
7	Adenoma	0.686	0.358	0.400	0.361	Acinetobacter	0.000	-0.029	Acinetobacter <= 7.52e-05
10	Adenoma	0.704	0.475	0.300	0.477	Alloscardovia	0.001	0.061	0.000122 < Alloscardovia
10	Adenoma	0.704	0.475	0.300	0.477	Actinomyces	0.025	0.055	0.02229 < Actinomyces
10	Adenoma	0.704	0.475	0.300	0.477	Gemella	0.083	0.052	0.0352 < Gemella
10	Adenoma	0.704	0.475	0.300	0.477	Clostridia_UCG.014	0.022	0.042	0.01058 < Clostridia_UCG.014
10	Adenoma	0.704	0.475	0.300	0.477	Pseudomonas	0.000	-0.033	Pseudomonas <= 6.6e-06
11	Adenoma	0.622	0.411	0.273	0.496	Alloscardovia	0.000	0.063	0.000122 < Alloscardovia
11	Adenoma	0.622	0.411	0.273	0.496	Actinomyces	0.023	0.058	0.02229 < Actinomyces
11	Adenoma	0.622	0.411	0.273	0.496	Clostridia_UCG.014	0.012	0.043	0.01058 < Clostridia_UCG.014

11	Adenoma	0.622	0.411	0.273	0.496	Pseudomonas	0.000	0.032	6.6e-06 < Pseudomonas
11	Adenoma	0.622	0.411	0.273	0.496	Corynebacterium	0.002	0.028	0.001785 < Corynebacterium
14	Adenocarcinoma	0.894	0.267	0.559	0.734	Alloscardovia	0.000	0.059	Alloscardovia <= 0.000122
14	Adenocarcinoma	0.894	0.267	0.559	0.734	Clostridia_UCG.014	0.001	0.034	Clostridia_UCG.014 <= 0.00172
14	Adenocarcinoma	0.894	0.267	0.559	0.734	Acinetobacter	0.000	0.031	Acinetobacter <= 7.52e-05
14	Adenocarcinoma	0.894	0.267	0.559	0.734	Pseudomonas	0.000	0.033	Pseudomonas <= 6.6e-06
14	Adenocarcinoma	0.894	0.267	0.559	0.734	F0332	0.000	0.018	F0332 <= 1.43e-05
20	Adenocarcinoma	0.866	0.257	0.579	0.699	Alloscardovia	0.000	0.060	Alloscardovia <= 0.000122
20	Adenocarcinoma	0.866	0.257	0.579	0.699	Pseudomonas	0.000	0.034	Pseudomonas <= 6.6e-06
20	Adenocarcinoma	0.866	0.257	0.579	0.699	Acinetobacter	0.000	0.028	Acinetobacter <= 7.52e-05
20	Adenocarcinoma	0.866	0.257	0.579	0.699	Corynebacterium	0.006	-0.026	0.001785 < Corynebacterium
20	Adenocarcinoma	0.866	0.257	0.579	0.699	Actinomyces	0.005	0.025	Actinomyces <= 0.00585
21	Adenocarcinoma	0.852	0.296	0.647	0.627	Alloscardovia	0.000	0.059	Alloscardovia <= 0.000122
21	Adenocarcinoma	0.852	0.296	0.647	0.627	Clostridia_UCG.014	0.021	-0.043	0.01058 < Clostridia_UCG.014
21	Adenocarcinoma	0.852	0.296	0.647	0.627	Pseudomonas	0.000	-0.032	6.6e-06 < Pseudomonas
21	Adenocarcinoma	0.852	0.296	0.647	0.627	Acinetobacter	0.000	-0.031	7.52e-05 < Acinetobacter
21	Adenocarcinoma	0.852	0.296	0.647	0.627	Actinomyces	0.006	0.026	0.00585 < Actinomyces <= 0.01362
22	Adenocarcinoma	0.838	0.276	0.669	0.609	Alloscardovia	0.000	-0.061	0.000122 < Alloscardovia
22	Adenocarcinoma	0.838	0.276	0.669	0.609	Pseudomonas	0.000	0.032	Pseudomonas <= 6.6e-06
22	Adenocarcinoma	0.838	0.276	0.669	0.609	Corynebacterium	0.003	-0.029	0.001785 < Corynebacterium
22	Adenocarcinoma	0.838	0.276	0.669	0.609	Acinetobacter	0.000	-0.028	7.52e-05 < Acinetobacter
22	Adenocarcinoma	0.838	0.276	0.669	0.609	Actinomyces	0.004	0.025	Actinomyces <= 0.00585
23	Adenocarcinoma	0.836	0.346	0.649	0.618	Alloscardovia	0.000	-0.061	0.000122 < Alloscardovia
23	Adenocarcinoma	0.836	0.346	0.649	0.618	Actinomyces	0.046	-0.054	0.02229 < Actinomyces
23	Adenocarcinoma	0.836	0.346	0.649	0.618	Pseudomonas	0.000	0.033	Pseudomonas <= 6.6e-06
23	Adenocarcinoma	0.836	0.346	0.649	0.618	Acinetobacter	0.000	0.028	Acinetobacter <= 7.52e-05
23	Adenocarcinoma	0.836	0.346	0.649	0.618	Clostridia_UCG.014	0.002	0.023	0.00172 < Clostridia_UCG.014 <= 0.00404
24	Adenocarcinoma	0.922	0.228	0.579	0.697	Alloscardovia	0.000	0.057	Alloscardovia <= 0.000122
24	Adenocarcinoma	0.922	0.228	0.579	0.697	Pseudomonas	0.000	0.032	Pseudomonas <= 6.6e-06
24	Adenocarcinoma	0.922	0.228	0.579	0.697	Acinetobacter	0.000	0.027	Acinetobacter <= 7.52e-05
24	Adenocarcinoma	0.922	0.228	0.579	0.697	F0332	0.000	0.019	F0332 <= 1.43e-05
24	Adenocarcinoma	0.922	0.228	0.579	0.697	Johnsonella	0.001	-0.016	0.000603 < Johnsonella
25	Adenocarcinoma	0.876	0.252	0.564	0.729	Alloscardovia	0.000	0.060	Alloscardovia <= 0.000122
25	Adenocarcinoma	0.876	0.252	0.564	0.729	Pseudomonas	0.000	0.032	Pseudomonas <= 6.6e-06
25	Adenocarcinoma	0.876	0.252	0.564	0.729	Clostridia_UCG.014	0.000	0.030	Clostridia_UCG.014 <= 0.00172
25	Adenocarcinoma	0.876	0.252	0.564	0.729	Acinetobacter	0.000	0.025	Acinetobacter <= 7.52e-05
25	Adenocarcinoma	0.876	0.252	0.564	0.729	F0332	0.000	0.018	F0332 <= 1.43e-05
26	Adenocarcinoma	0.904	0.299	0.583	0.683	Alloscardovia	0.000	0.059	Alloscardovia <= 0.000122
26	Adenocarcinoma	0.904	0.299	0.583	0.683	Clostridia_UCG.014	0.014	-0.042	0.01058 < Clostridia_UCG.014
26	Adenocarcinoma	0.904	0.299	0.583	0.683	Pseudomonas	0.000	0.035	Pseudomonas <= 6.6e-06
26	Adenocarcinoma	0.904	0.299	0.583	0.683	Acinetobacter	0.000	0.028	Acinetobacter <= 7.52e-05
26	Adenocarcinoma	0.904	0.299	0.583	0.683	Actinomyces	0.003	0.021	Actinomyces <= 0.00585

29	Adenocarcinoma	0.912	0.353	0.645	0.626	Alloscardovia	0.000	-0.059	0.000122 < Alloscardovia
29	Adenocarcinoma	0.912	0.353	0.645	0.626	Gemella	0.038	-0.052	0.0352 < Gemella
29	Adenocarcinoma	0.912	0.353	0.645	0.626	Pseudomonas	0.000	0.033	Pseudomonas <= 6.6e-06
29	Adenocarcinoma	0.912	0.353	0.645	0.626	Clostridia_UCG.014	0.001	0.031	Clostridia_UCG.014 <= 0.00172
29	Adenocarcinoma	0.912	0.353	0.645	0.626	Acinetobacter	0.000	0.028	Acinetobacter <= 7.52e-05
32	Adenocarcinoma	0.922	0.240	0.570	0.734	Alloscardovia	0.000	0.060	Alloscardovia <= 0.000122
32	Adenocarcinoma	0.922	0.240	0.570	0.734	Pseudomonas	0.000	0.033	Pseudomonas <= 6.6e-06
32	Adenocarcinoma	0.922	0.240	0.570	0.734	Acinetobacter	0.000	0.026	Acinetobacter <= 7.52e-05
32	Adenocarcinoma	0.922	0.240	0.570	0.734	Actinomyces	0.011	0.027	0.00585 < Actinomyces <= 0.01362
32	Adenocarcinoma	0.922	0.240	0.570	0.734	Gemella	0.033	0.019	0.0211 < Gemella <= 0.0352
33	Adenocarcinoma	0.816	0.290	0.616	0.657	Alloscardovia	0.000	0.059	Alloscardovia <= 0.000122
33	Adenocarcinoma	0.816	0.290	0.616	0.657	Clostridia_UCG.014	0.013	-0.040	0.01058 < Clostridia_UCG.014
33	Adenocarcinoma	0.816	0.290	0.616	0.657	Pseudomonas	0.000	0.033	Pseudomonas <= 6.6e-06
33	Adenocarcinoma	0.816	0.290	0.616	0.657	Acinetobacter	0.000	-0.030	7.52e-05 < Acinetobacter
33	Adenocarcinoma	0.816	0.290	0.616	0.657	Gemella	0.021	0.020	0.0129 < Gemella <= 0.0211
34	Adenocarcinoma	0.864	0.229	0.574	0.734	Alloscardovia	0.000	0.057	Alloscardovia <= 0.000122
34	Adenocarcinoma	0.864	0.229	0.574	0.734	Pseudomonas	0.000	0.031	Pseudomonas <= 6.6e-06
34	Adenocarcinoma	0.864	0.229	0.574	0.734	Acinetobacter	0.000	0.028	Acinetobacter <= 7.52e-05
34	Adenocarcinoma	0.864	0.229	0.574	0.734	Actinomyces	0.003	0.022	Actinomyces <= 0.00585
34	Adenocarcinoma	0.864	0.229	0.574	0.734	Clostridia_UCG.014	0.002	0.023	0.00172 < Clostridia_UCG.014 <= 0.00404
38	Adenocarcinoma	0.834	0.291	0.620	0.658	Alloscardovia	0.000	0.059	Alloscardovia <= 0.000122
38	Adenocarcinoma	0.834	0.291	0.620	0.658	Clostridia_UCG.014	0.030	-0.039	0.01058 < Clostridia_UCG.014
38	Adenocarcinoma	0.834	0.291	0.620	0.658	Pseudomonas	0.000	-0.033	6.6e-06 < Pseudomonas
38	Adenocarcinoma	0.834	0.291	0.620	0.658	Acinetobacter	0.000	0.026	Acinetobacter <= 7.52e-05
38	Adenocarcinoma	0.834	0.291	0.620	0.658	Actinomyces	0.013	0.025	0.00585 < Actinomyces <= 0.01362
39	Adenocarcinoma	0.882	0.266	0.568	0.744	Alloscardovia	0.000	0.060	Alloscardovia <= 0.000122
39	Adenocarcinoma	0.882	0.266	0.568	0.744	Pseudomonas	0.000	0.033	Pseudomonas <= 6.6e-06
39	Adenocarcinoma	0.882	0.266	0.568	0.744	Clostridia_UCG.014	0.000	0.031	Clostridia_UCG.014 <= 0.00172
39	Adenocarcinoma	0.882	0.266	0.568	0.744	Acinetobacter	0.000	0.026	Acinetobacter <= 7.52e-05
39	Adenocarcinoma	0.882	0.266	0.568	0.744	Actinomyces	0.010	0.025	0.00585 < Actinomyces <= 0.01362
41	Adenocarcinoma	0.846	0.371	0.586	0.684	Alloscardovia	0.000	0.061	Alloscardovia <= 0.000122
41	Adenocarcinoma	0.846	0.371	0.586	0.684	Actinomyces	0.053	-0.056	0.02229 < Actinomyces
41	Adenocarcinoma	0.846	0.371	0.586	0.684	Clostridia_UCG.014	0.002	0.032	Clostridia_UCG.014 <= 0.00172
41	Adenocarcinoma	0.846	0.371	0.586	0.684	Pseudomonas	0.000	0.033	Pseudomonas <= 6.6e-06
41	Adenocarcinoma	0.846	0.371	0.586	0.684	Acinetobacter	0.000	0.027	Acinetobacter <= 7.52e-05
46	Adenocarcinoma	0.878	0.352	0.602	0.643	Alloscardovia	0.000	0.062	Alloscardovia <= 0.000122
46	Adenocarcinoma	0.878	0.352	0.602	0.643	Gemella	0.037	-0.052	0.0352 < Gemella
46	Adenocarcinoma	0.878	0.352	0.602	0.643	Pseudomonas	0.000	0.033	Pseudomonas <= 6.6e-06
46	Adenocarcinoma	0.878	0.352	0.602	0.643	Corynebacterium	0.002	-0.028	0.001785 < Corynebacterium
46	Adenocarcinoma	0.878	0.352	0.602	0.643	Acinetobacter	0.000	0.026	Acinetobacter <= 7.52e-05
47	Adenocarcinoma	0.724	0.355	0.722	0.526	Alloscardovia	0.000	-0.060	0.000122 < Alloscardovia
47	Adenocarcinoma	0.724	0.355	0.722	0.526	Actinomyces	0.032	-0.055	0.02229 < Actinomyces

47	Adenocarcinoma	0.724	0.355	0.722	0.526	Pseudomonas	0.000	-0.032	6.6e-06 < Pseudomonas
47	Adenocarcinoma	0.724	0.355	0.722	0.526	Corynebacterium	0.004	-0.024	0.001785 < Corynebacterium
47	Adenocarcinoma	0.724	0.355	0.722	0.526	Acinetobacter	0.000	-0.024	7.52e-05 < Acinetobacter
49	Adenocarcinoma	0.814	0.353	0.588	0.685	Alloscardovia	0.000	0.057	Alloscardovia <= 0.000122
49	Adenocarcinoma	0.814	0.353	0.588	0.685	Gemella	0.035	-0.052	0.0352 < Gemella
49	Adenocarcinoma	0.814	0.353	0.588	0.685	Pseudomonas	0.000	0.034	Pseudomonas <= 6.6e-06
49	Adenocarcinoma	0.814	0.353	0.588	0.685	Clostridia_UCG.014	0.000	0.030	Clostridia_UCG.014 <= 0.00172
49	Adenocarcinoma	0.814	0.353	0.588	0.685	Acinetobacter	0.000	0.027	Acinetobacter <= 7.52e-05
51	Adenocarcinoma	0.818	0.359	0.618	0.651	Alloscardovia	0.000	0.060	Alloscardovia <= 0.000122
51	Adenocarcinoma	0.818	0.359	0.618	0.651	Gemella	0.043	-0.049	0.0352 < Gemella
51	Adenocarcinoma	0.818	0.359	0.618	0.651	Pseudomonas	0.000	-0.039	6.6e-06 < Pseudomonas
51	Adenocarcinoma	0.818	0.359	0.618	0.651	Clostridia_UCG.014	0.000	0.032	Clostridia_UCG.014 <= 0.00172
51	Adenocarcinoma	0.818	0.359	0.618	0.651	Acinetobacter	0.000	0.028	Acinetobacter <= 7.52e-05
52	Adenocarcinoma	0.912	0.232	0.576	0.732	Alloscardovia	0.000	0.057	Alloscardovia <= 0.000122
52	Adenocarcinoma	0.912	0.232	0.576	0.732	Pseudomonas	0.000	0.030	Pseudomonas <= 6.6e-06
52	Adenocarcinoma	0.912	0.232	0.576	0.732	Acinetobacter	0.000	0.025	Acinetobacter <= 7.52e-05
52	Adenocarcinoma	0.912	0.232	0.576	0.732	Actinomyces	0.006	0.024	0.00585 < Actinomyces <= 0.01362
52	Adenocarcinoma	0.912	0.232	0.576	0.732	Gemella	0.034	0.020	0.0211 < Gemella <= 0.0352

Table ML-E-5: The LIME Result for Test Dataset

case	label	label_prob	model_r2	model_intercept	model_prediction	feature	feature_value	feature_weight	feature_desc
1	Adenocarcinoma	0.748	0.368	0.600	0.639	Alloscardovia	0.000	0.060	Alloscardovia <= 0.000122
1	Adenocarcinoma	0.748	0.368	0.600	0.639	Gemella	0.049	-0.055	0.0352 < Gemella
1	Adenocarcinoma	0.748	0.368	0.600	0.639	Pseudomonas	0.000	0.033	Pseudomonas <= 6.6e-06
1	Adenocarcinoma	0.748	0.368	0.600	0.639	Acinetobacter	0.000	0.030	Acinetobacter <= 7.52e-05
1	Adenocarcinoma	0.748	0.368	0.600	0.639	Corynebacterium	0.004	-0.028	0.001785 < Corynebacterium
4	Adenocarcinoma	0.758	0.299	0.670	0.590	Alloscardovia	0.000	-0.062	0.000122 < Alloscardovia
4	Adenocarcinoma	0.758	0.299	0.670	0.590	Clostridia_UCG.014	0.014	-0.041	0.01058 < Clostridia_UCG.014
4	Adenocarcinoma	0.758	0.299	0.670	0.590	Pseudomonas	0.000	0.032	Pseudomonas <= 6.6e-06
4	Adenocarcinoma	0.758	0.299	0.670	0.590	Acinetobacter	0.000	-0.026	7.52e-05 < Acinetobacter
4	Adenocarcinoma	0.758	0.299	0.670	0.590	F0332	0.000	0.017	F0332 <= 1.43e-05
5	Adenocarcinoma	0.846	0.263	0.628	0.682	Alloscardovia	0.000	0.061	Alloscardovia <= 0.000122
5	Adenocarcinoma	0.846	0.263	0.628	0.682	Pseudomonas	0.000	-0.031	6.6e-06 < Pseudomonas
5	Adenocarcinoma	0.846	0.263	0.628	0.682	Actinomyces	0.006	0.027	Actinomyces <= 0.00585
5	Adenocarcinoma	0.846	0.263	0.628	0.682	Acinetobacter	0.000	-0.026	7.52e-05 < Acinetobacter
5	Adenocarcinoma	0.846	0.263	0.628	0.682	Clostridia_UCG.014	0.002	0.023	0.00172 < Clostridia_UCG.014 <= 0.00404
8	Adenocarcinoma	0.710	0.391	0.665	0.567	Alloscardovia	0.001	-0.061	0.000122 < Alloscardovia
8	Adenocarcinoma	0.710	0.391	0.665	0.567	Actinomyces	0.036	-0.057	0.02229 < Actinomyces
8	Adenocarcinoma	0.710	0.391	0.665	0.567	Clostridia_UCG.014	0.016	-0.039	0.01058 < Clostridia_UCG.014
8	Adenocarcinoma	0.710	0.391	0.665	0.567	Pseudomonas	0.000	0.031	Pseudomonas <= 6.6e-06
8	Adenocarcinoma	0.710	0.391	0.665	0.567	Acinetobacter	0.000	0.030	Acinetobacter <= 7.52e-05

9	Adenocarcinoma	0.584	0.246	0.680	0.582	Alloscardovia	0.001	-0.061	0.000122 < Alloscardovia
9	Adenocarcinoma	0.584	0.246	0.680	0.582	Pseudomonas	0.000	0.031	Pseudomonas <= 6.6e-06
9	Adenocarcinoma	0.584	0.246	0.680	0.582	Corynebacterium	0.005	-0.028	0.001785 < Corynebacterium
9	Adenocarcinoma	0.584	0.246	0.680	0.582	X.Eubacterium._brachy_group	0.000	-0.022	1.51e-04 < X.Eubacterium._brachy_group
9	Adenocarcinoma	0.584	0.246	0.680	0.582	F0332	0.000	-0.019	1.20e-04 < F0332
12	Adenocarcinoma	0.782	0.279	0.610	0.667	Alloscardovia	0.000	0.061	Alloscardovia <= 0.000122
12	Adenocarcinoma	0.782	0.279	0.610	0.667	Pseudomonas	0.000	-0.030	6.6e-06 < Pseudomonas
12	Adenocarcinoma	0.782	0.279	0.610	0.667	Acinetobacter	0.000	0.031	Acinetobacter <= 7.52e-05
12	Adenocarcinoma	0.782	0.279	0.610	0.667	Corynebacterium	0.002	-0.031	0.001785 < Corynebacterium
12	Adenocarcinoma	0.782	0.279	0.610	0.667	Clostridia_UCG.014	0.003	0.026	0.00172 < Clostridia_UCG.014 <= 0.00404
13	Adenocarcinoma	0.672	0.249	0.623	0.691	Alloscardovia	0.000	0.061	Alloscardovia <= 0.000122
13	Adenocarcinoma	0.672	0.249	0.623	0.691	Clostridia_UCG.014	0.000	0.032	Clostridia_UCG.014 <= 0.00172
13	Adenocarcinoma	0.672	0.249	0.623	0.691	Acinetobacter	0.000	-0.028	7.52e-05 < Acinetobacter
13	Adenocarcinoma	0.672	0.249	0.623	0.691	Actinomyces	0.010	0.021	0.00585 < Actinomyces <= 0.01362
13	Adenocarcinoma	0.672	0.249	0.623	0.691	F0332	0.000	-0.018	1.20e-04 < F0332
15	Adenocarcinoma	0.788	0.347	0.675	0.595	Alloscardovia	0.000	-0.061	0.000122 < Alloscardovia
15	Adenocarcinoma	0.788	0.347	0.675	0.595	Gemella	0.039	-0.051	0.0352 < Gemella
15	Adenocarcinoma	0.788	0.347	0.675	0.595	Acinetobacter	0.000	0.035	Acinetobacter <= 7.52e-05
15	Adenocarcinoma	0.788	0.347	0.675	0.595	Pseudomonas	0.000	-0.030	6.6e-06 < Pseudomonas
15	Adenocarcinoma	0.788	0.347	0.675	0.595	Clostridia_UCG.014	0.003	0.027	0.00172 < Clostridia_UCG.014 <= 0.00404
16	Adenocarcinoma	0.608	0.456	0.670	0.555	Alloscardovia	0.000	0.060	Alloscardovia <= 0.000122
16	Adenocarcinoma	0.608	0.456	0.670	0.555	Actinomyces	0.038	-0.054	0.02229 < Actinomyces
16	Adenocarcinoma	0.608	0.456	0.670	0.555	Gemella	0.056	-0.056	0.0352 < Gemella
16	Adenocarcinoma	0.608	0.456	0.670	0.555	Corynebacterium	0.002	-0.033	0.001785 < Corynebacterium
16	Adenocarcinoma	0.608	0.456	0.670	0.555	Pseudomonas	0.000	-0.032	6.6e-06 < Pseudomonas
17	Adenocarcinoma	0.788	0.380	0.646	0.622	Alloscardovia	0.000	-0.062	0.000122 < Alloscardovia
17	Adenocarcinoma	0.788	0.380	0.646	0.622	Gemella	0.039	-0.056	0.0352 < Gemella
17	Adenocarcinoma	0.788	0.380	0.646	0.622	Clostridia_UCG.014	0.001	0.031	Clostridia_UCG.014 <= 0.00172
17	Adenocarcinoma	0.788	0.380	0.646	0.622	Pseudomonas	0.000	0.032	Pseudomonas <= 6.6e-06
17	Adenocarcinoma	0.788	0.380	0.646	0.622	Acinetobacter	0.000	0.030	Acinetobacter <= 7.52e-05
18	Adenocarcinoma	0.756	0.241	0.611	0.721	Alloscardovia	0.000	0.062	Alloscardovia <= 0.000122
18	Adenocarcinoma	0.756	0.241	0.611	0.721	Clostridia_UCG.014	0.000	0.032	Clostridia_UCG.014 <= 0.00172
18	Adenocarcinoma	0.756	0.241	0.611	0.721	Acinetobacter	0.000	-0.026	7.52e-05 < Acinetobacter
18	Adenocarcinoma	0.756	0.241	0.611	0.721	Actinomyces	0.003	0.024	Actinomyces <= 0.00585
18	Adenocarcinoma	0.756	0.241	0.611	0.721	Corynebacterium	0.001	0.017	0.000195 < Corynebacterium <= 0.000822
19	Adenocarcinoma	0.648	0.228	0.591	0.704	Alloscardovia	0.000	0.063	Alloscardovia <= 0.000122
19	Adenocarcinoma	0.648	0.228	0.591	0.704	Pseudomonas	0.000	0.031	Pseudomonas <= 6.6e-06
19	Adenocarcinoma	0.648	0.228	0.591	0.704	Clostridia_UCG.014	0.004	0.021	0.00172 < Clostridia_UCG.014 <= 0.00404
19	Adenocarcinoma	0.648	0.228	0.591	0.704	F0332	0.000	0.016	F0332 <= 1.43e-05
19	Adenocarcinoma	0.648	0.228	0.591	0.704	Neisseria	0.000	-0.018	Neisseria <= 0.00723
27	Adenocarcinoma	0.774	0.286	0.587	0.690	Alloscardovia	0.000	0.058	Alloscardovia <= 0.000122
27	Adenocarcinoma	0.774	0.286	0.587	0.690	Clostridia_UCG.014	0.016	-0.039	0.01058 < Clostridia_UCG.014

27	Adenocarcinoma	0.774	0.286	0.587	0.690	Pseudomonas	0.000	0.031	Pseudomonas <= 6.6e-06
27	Adenocarcinoma	0.774	0.286	0.587	0.690	Acinetobacter	0.000	0.028	Acinetobacter <= 7.52e-05
27	Adenocarcinoma	0.774	0.286	0.587	0.690	Actinomyces	0.003	0.026	Actinomyces <= 0.00585
28	Adenocarcinoma	0.844	0.229	0.580	0.694	Alloscardovia	0.000	0.061	Alloscardovia <= 0.000122
28	Adenocarcinoma	0.844	0.229	0.580	0.694	Pseudomonas	0.000	0.032	Pseudomonas <= 6.6e-06
28	Adenocarcinoma	0.844	0.229	0.580	0.694	Acinetobacter	0.000	0.022	Acinetobacter <= 7.52e-05
28	Adenocarcinoma	0.844	0.229	0.580	0.694	F0332	0.000	0.016	F0332 <= 1.43e-05
28	Adenocarcinoma	0.844	0.229	0.580	0.694	Veillonella	0.258	-0.016	0.1662 < Veillonella
30	Adenocarcinoma	0.772	0.262	0.595	0.709	Alloscardovia	0.000	0.063	Alloscardovia <= 0.000122
30	Adenocarcinoma	0.772	0.262	0.595	0.709	Pseudomonas	0.000	0.031	Pseudomonas <= 6.6e-06
30	Adenocarcinoma	0.772	0.262	0.595	0.709	Acinetobacter	0.000	-0.028	7.52e-05 < Acinetobacter
30	Adenocarcinoma	0.772	0.262	0.595	0.709	Gemella	0.019	0.025	0.0129 < Gemella <= 0.0211
30	Adenocarcinoma	0.772	0.262	0.595	0.709	Actinomyces	0.011	0.024	0.00585 < Actinomyces <= 0.01362
31	Adenocarcinoma	0.672	0.277	0.615	0.666	Alloscardovia	0.000	0.060	Alloscardovia <= 0.000122
31	Adenocarcinoma	0.672	0.277	0.615	0.666	Pseudomonas	0.000	-0.034	6.6e-06 < Pseudomonas
31	Adenocarcinoma	0.672	0.277	0.615	0.666	Corynebacterium	0.003	-0.028	0.001785 < Corynebacterium
31	Adenocarcinoma	0.672	0.277	0.615	0.666	Clostridia_UCG.014	0.001	0.028	Clostridia_UCG.014 <= 0.00172
31	Adenocarcinoma	0.672	0.277	0.615	0.666	Acinetobacter	0.000	0.026	Acinetobacter <= 7.52e-05
35	Adenocarcinoma	0.880	0.276	0.566	0.743	Alloscardovia	0.000	0.059	Alloscardovia <= 0.000122
35	Adenocarcinoma	0.880	0.276	0.566	0.743	Clostridia_UCG.014	0.002	0.033	Clostridia_UCG.014 <= 0.00172
35	Adenocarcinoma	0.880	0.276	0.566	0.743	Pseudomonas	0.000	0.032	Pseudomonas <= 6.6e-06
35	Adenocarcinoma	0.880	0.276	0.566	0.743	Acinetobacter	0.000	0.027	Acinetobacter <= 7.52e-05
35	Adenocarcinoma	0.880	0.276	0.566	0.743	Actinomyces	0.004	0.026	Actinomyces <= 0.00585
36	Adenocarcinoma	0.672	0.264	0.576	0.689	Alloscardovia	0.000	0.060	Alloscardovia <= 0.000122
36	Adenocarcinoma	0.672	0.264	0.576	0.689	Pseudomonas	0.000	0.035	Pseudomonas <= 6.6e-06
36	Adenocarcinoma	0.672	0.264	0.576	0.689	Acinetobacter	0.000	0.029	Acinetobacter <= 7.52e-05
36	Adenocarcinoma	0.672	0.264	0.576	0.689	Corynebacterium	0.002	-0.028	0.001785 < Corynebacterium
36	Adenocarcinoma	0.672	0.264	0.576	0.689	F0332	0.000	0.018	F0332 <= 1.43e-05
37	Adenocarcinoma	0.828	0.295	0.592	0.713	Alloscardovia	0.000	0.065	Alloscardovia <= 0.000122
37	Adenocarcinoma	0.828	0.295	0.592	0.713	Clostridia_UCG.014	0.000	0.031	Clostridia_UCG.014 <= 0.00172
37	Adenocarcinoma	0.828	0.295	0.592	0.713	Pseudomonas	0.000	-0.032	6.6e-06 < Pseudomonas
37	Adenocarcinoma	0.828	0.295	0.592	0.713	Acinetobacter	0.000	0.030	Acinetobacter <= 7.52e-05
37	Adenocarcinoma	0.828	0.295	0.592	0.713	Actinomyces	0.005	0.027	Actinomyces <= 0.00585
40	Adenocarcinoma	0.576	0.423	0.667	0.565	Alloscardovia	0.001	-0.060	0.000122 < Alloscardovia
40	Adenocarcinoma	0.576	0.423	0.667	0.565	Actinomyces	0.025	-0.055	0.02229 < Actinomyces
40	Adenocarcinoma	0.576	0.423	0.667	0.565	Gemella	0.043	-0.049	0.0352 < Gemella
40	Adenocarcinoma	0.576	0.423	0.667	0.565	Pseudomonas	0.000	0.033	Pseudomonas <= 6.6e-06
40	Adenocarcinoma	0.576	0.423	0.667	0.565	Acinetobacter	0.000	0.029	Acinetobacter <= 7.52e-05
42	Adenocarcinoma	0.750	0.259	0.663	0.623	Alloscardovia	0.000	-0.060	0.000122 < Alloscardovia
42	Adenocarcinoma	0.750	0.259	0.663	0.623	Clostridia_UCG.014	0.000	0.033	Clostridia_UCG.014 <= 0.00172
42	Adenocarcinoma	0.750	0.259	0.663	0.623	Pseudomonas	0.000	0.032	Pseudomonas <= 6.6e-06
42	Adenocarcinoma	0.750	0.259	0.663	0.623	Acinetobacter	0.000	-0.028	7.52e-05 < Acinetobacter

42	Adenocarcinoma	0.750	0.259	0.663	0.623	Neisseria	0.006	-0.018	Neisseria <= 0.00723
43	Adenocarcinoma	0.504	0.440566751	0.65229591	0.602913265	Alloscardovia	7.90212E-05	0.06135943	Alloscardovia <= 0.000122
43	Adenocarcinoma	0.504	0.440566751	0.65229591	0.602913265	Gemella	0.067207523	-0.053519925	0.0352 < Gemella
43	Adenocarcinoma	0.504	0.440566751	0.65229591	0.602913265	Actinomyces	0.023798549	-0.0547971	0.02229 < Actinomyces
43	Adenocarcinoma	0.504	0.440566751	0.65229591	0.602913265	Corynebacterium	0.003555954	-0.033690641	0.001785 < Corynebacterium
43	Adenocarcinoma	0.504	0.440566751	0.65229591	0.602913265	Clostridia_UCG.014	6.5851E-05	0.031265591	Clostridia_UCG.014 <= 0.00172
44	Adenocarcinoma	0.82	0.355581987	0.587124504	0.681583804	Alloscardovia	0	0.059395844	Alloscardovia <= 0.000122
44	Adenocarcinoma	0.82	0.355581987	0.587124504	0.681583804	Actinomyces	0.025767509	-0.052802095	0.02229 < Actinomyces
44	Adenocarcinoma	0.82	0.355581987	0.587124504	0.681583804	Pseudomonas	0	0.032487642	Pseudomonas <= 6.6e-06
44	Adenocarcinoma	0.82	0.355581987	0.587124504	0.681583804	Clostridia_UCG.014	0	0.029389079	Clostridia_UCG.014 <= 0.00172
44	Adenocarcinoma	0.82	0.355581987	0.587124504	0.681583804	Acinetobacter	0	0.02598883	Acinetobacter <= 7.52e-05
45	Adenocarcinoma	0.716	0.355616957	0.634423664	0.607987059	Alloscardovia	0	0.062994456	Alloscardovia <= 0.000122
45	Adenocarcinoma	0.716	0.355616957	0.634423664	0.607987059	Gemella	0.039885538	-0.051074581	0.0352 < Gemella
45	Adenocarcinoma	0.716	0.355616957	0.634423664	0.607987059	Pseudomonas	0.000220119	-0.033217185	6.6e-06 < Pseudomonas
45	Adenocarcinoma	0.716	0.355616957	0.634423664	0.607987059	Corynebacterium	0.003918116	-0.029589038	0.001785 < Corynebacterium
45	Adenocarcinoma	0.716	0.355616957	0.634423664	0.607987059	Acinetobacter	0	0.024449743	Acinetobacter <= 7.52e-05
48	Adenocarcinoma	0.64	0.443018787	0.640326221	0.589104924	Alloscardovia	0	0.060359976	Alloscardovia <= 0.000122
48	Adenocarcinoma	0.64	0.443018787	0.640326221	0.589104924	Actinomyces	0.025178244	-0.054204202	0.02229 < Actinomyces
48	Adenocarcinoma	0.64	0.443018787	0.640326221	0.589104924	Gemella	0.044000815	-0.051518641	0.0352 < Gemella
48	Adenocarcinoma	0.64	0.443018787	0.640326221	0.589104924	Pseudomonas	0.000325932	-0.034200821	6.6e-06 < Pseudomonas
48	Adenocarcinoma	0.64	0.443018787	0.640326221	0.589104924	Acinetobacter	0	0.02834239	Acinetobacter <= 7.52e-05
50	Adenocarcinoma	0.824	0.247482051	0.598943664	0.693988814	Alloscardovia	0	0.057843038	Alloscardovia <= 0.000122
50	Adenocarcinoma	0.824	0.247482051	0.598943664	0.693988814	Pseudomonas	6.28101E-05	-0.033811731	6.6e-06 < Pseudomonas
50	Adenocarcinoma	0.824	0.247482051	0.598943664	0.693988814	Acinetobacter	1.57025E-05	0.03024545	Acinetobacter <= 7.52e-05
50	Adenocarcinoma	0.824	0.247482051	0.598943664	0.693988814	Clostridia_UCG.014	0.003776459	0.023869803	0.00172 < Clostridia_UCG.014 <= 0.00404
50	Adenocarcinoma	0.824	0.247482051	0.598943664	0.693988814	F0332	0	0.016898589	F0332 <= 1.43e-05



Figure ML3-E-6: The LIME Result in Training dataset for Genus in Saliva

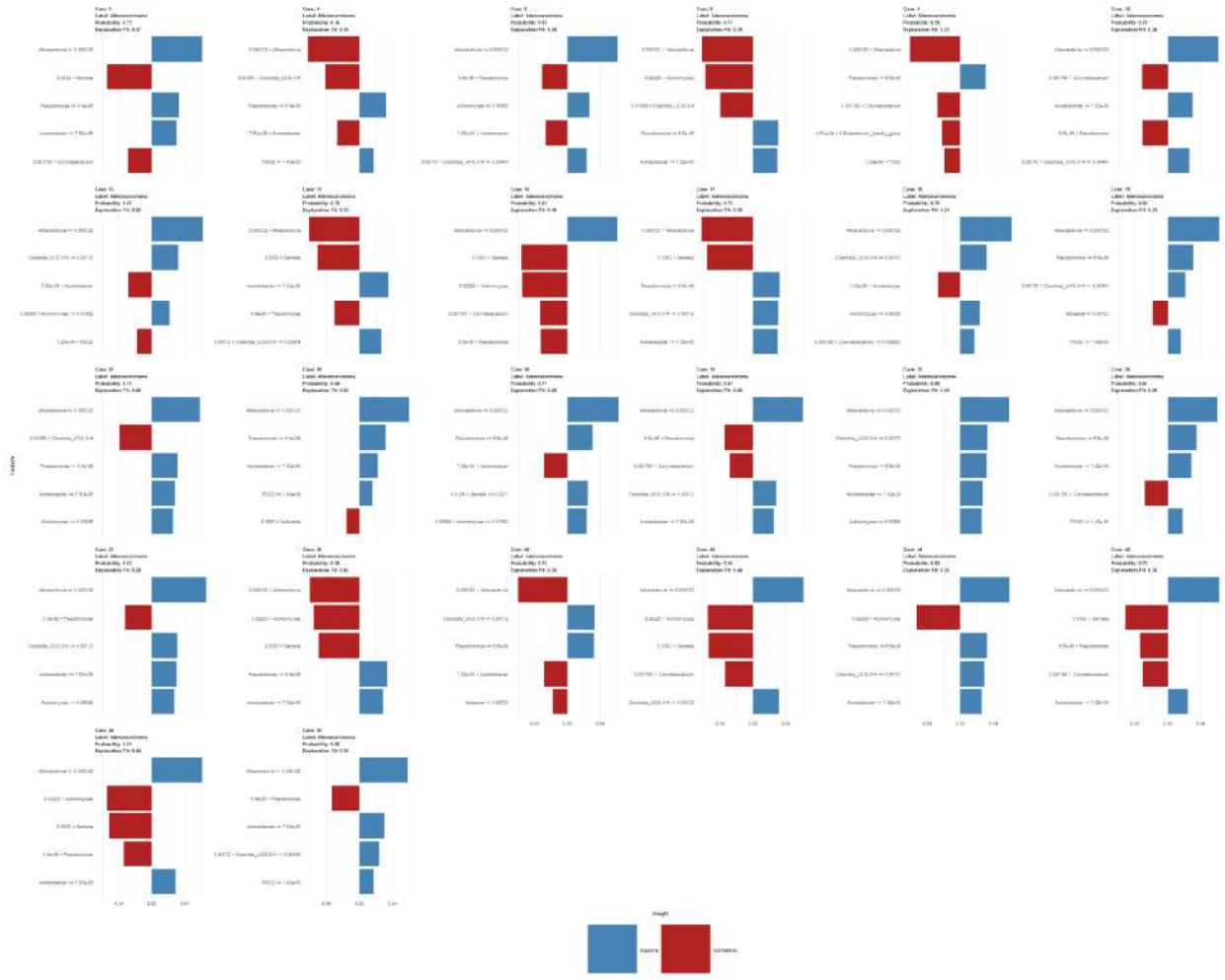


Figure ML3-E-7: The LIME Result in Test dataset for Genus in Saliva

Table ML3-E-6: The Aggregated SHAP Values

Variable	Training Dataset				Testing Dataset			
	contribution	position	sign	cumulative	contribution	position	sign	cumulative
intercept	0.7373	89	X	0.737	0.7373	89	X	0.737
Streptococcus	0.0000	88	1	0.737	-0.0045	76	-1	0.734
Bacteroides	0.0000	87	1	0.737	-0.0001	78	-1	0.738
Eikenella	0.0000	86	1	0.737	-0.0015	82	-1	0.738

Anaeroglobus	0.0000	85	-1	0.737	0.0005	85	1	0.738
Gracilibacteria	0.0000	84	1	0.737	0.0000	84	1	0.738
Haemophilus	0.0000	83	-1	0.737	0.0019	83	1	0.740
Johnsonella	0.0000	82	1	0.737	-0.0002	81	-1	0.738
Alysiella	0.0000	81	-1	0.737	-0.0004	87	-1	0.737
Peptococcus	0.0000	80	-1	0.737	0.0005	79	1	0.738
Filifactor	0.0000	79	-1	0.737	0.0003	86	1	0.737
X.Eubacterium._yurii_group	0.0000	78	-1	0.737	0.0000	77	-1	0.738
Saccharimonadales	0.0000	77	-1	0.737	0.0000	88	1	0.737
Lentimicrobium	0.0000	76	-1	0.737	0.0001	75	1	0.734
Chryseobacterium	0.0000	75	-1	0.737	0.0010	74	1	0.735
Lachnoanaerobaculum	0.0000	74	-1	0.737	-0.0004	73	-1	0.735
Bacillus	0.0000	73	1	0.737	0.0001	72	1	0.735
Mogibacterium	0.0000	72	-1	0.737	0.0005	70	1	0.735
Rikenellaceae_RC9_gut_group	0.0000	71	-1	0.737	-0.0003	56	-1	0.744
Rothia	0.0000	70	1	0.737	-0.0004	68	-1	0.735
Dialister	0.0000	69	-1	0.737	0.0004	67	1	0.735
Tumor	0.0000	68	1	0.737	0.0000	80	1	0.738
uncultured	0.0000	67	-1	0.737	-0.0002	69	-1	0.735
TM7a	0.0000	66	1	0.737	-0.0008	66	-1	0.734
Veillonella	0.0000	65	1	0.737	0.0001	65	1	0.734
JGI_0000069.P22	0.0000	64	1	0.737	0.0001	64	1	0.734
Lautropia	0.0000	63	1	0.737	0.0003	62	1	0.735
Parvimonas	0.0000	62	1	0.737	-0.0022	61	-1	0.733
Clostridia_UCG.014	0.0000	61	-1	0.737	0.0079	60	1	0.741
Chloroplast	0.0000	60	1	0.737	0.0000	59	1	0.741
Megasphaera	0.0000	59	1	0.737	0.0039	58	1	0.745
Saccharimonadaceae	0.0000	58	1	0.737	0.0005	32	1	0.741
Porphyromonas	0.0000	57	1	0.737	-0.0010	57	-1	0.744
Capnocytophaga	0.0000	56	-1	0.737	-0.0013	43	-1	0.737
Stomatobaculum	0.0000	55	-1	0.737	0.0000	39	1	0.739
X.Eubacterium._saphenum_group	0.0000	54	1	0.737	0.0002	55	1	0.744
Solobacterium	0.0001	53	1	0.737	0.0008	54	1	0.745
Atopobium	0.0000	52	-1	0.737	0.0000	52	-1	0.744
Gemella	0.0000	51	-1	0.737	-0.0034	51	-1	0.740
Olsenella	0.0000	50	-1	0.737	0.0004	50	1	0.741
Oribacterium	0.0000	49	1	0.737	0.0006	63	1	0.735
Tannerella	0.0000	48	-1	0.737	0.0010	49	1	0.742
Acinetobacter	0.0000	47	1	0.737	-0.0003	48	-1	0.741
Leptotrichia	0.0000	46	1	0.737	-0.0022	45	-1	0.739
Peptoanaerobacter	0.0000	45	-1	0.737	-0.0003	44	-1	0.738
Bifidobacterium	0.0000	44	-1	0.737	0.0002	42	1	0.737

Gender	0.0000	43	1	0.737	0.0000	40	1	0.739
Cardiobacterium	0.0000	42	-1	0.737	-0.0009	53	-1	0.744
Bergeyella	0.0000	41	1	0.737	0.0006	37	1	0.739
F0332	-0.0001	40	-1	0.737	0.0018	30	1	0.743
Phocaeicola	0.0001	39	1	0.737	-0.0001	36	-1	0.739
Aggregatibacter	0.0000	38	1	0.737	0.0003	35	1	0.740
Mycoplasma	0.0000	37	1	0.737	0.0006	9	1	0.732
Actinomyces	0.0001	36	1	0.737	0.0011	34	1	0.741
Clostridia_vadinBB60_group	0.0000	35	-1	0.737	0.0001	33	1	0.741
Catonella	0.0000	34	1	0.737	-0.0003	31	-1	0.741
Fusobacterium	0.0000	33	1	0.737	0.0010	29	1	0.744
Peptostreptococcus	0.0000	32	-1	0.737	-0.0004	28	-1	0.743
Selenomonas	0.0000	31	-1	0.737	-0.0005	27	-1	0.743
Butyrivibrio	0.0000	30	1	0.737	-0.0001	26	-1	0.743
RF39	0.0000	29	1	0.737	-0.0004	8	-1	0.731
F0058	0.0000	28	1	0.737	-0.0001	38	-1	0.739
X.Eubacterium._nodatum_group	0.0000	27	1	0.737	0.0001	25	1	0.743
Serratia	0.0000	26	1	0.737	0.0002	24	1	0.743
Candidatus_Saccharimonas	0.0000	25	1	0.737	0.0012	23	1	0.744
Fretibacterium	0.0002	24	1	0.738	0.0006	22	1	0.745
TM7x	0.0000	23	-1	0.738	0.0005	21	1	0.745
Prevotella	0.0000	22	-1	0.738	-0.0023	20	-1	0.743
Comamonas	0.0000	21	1	0.738	-0.0001	19	-1	0.743
Granulicatella	-0.0001	20	-1	0.738	-0.0010	18	-1	0.742
Pseudomonas	0.0000	19	-1	0.738	-0.0062	17	-1	0.736
Actinobacillus	0.0000	18	-1	0.738	0.0000	16	1	0.736
Corynebacterium	0.0000	17	-1	0.738	-0.0030	15	-1	0.733
Alloscardovia	0.0000	16	-1	0.738	-0.0031	14	-1	0.730
Amnipila	0.0000	15	-1	0.738	-0.0002	47	-1	0.741
Neisseria	-0.0001	14	-1	0.737	0.0015	13	1	0.731
X.Eubacterium._brachy_group	-0.0001	13	-1	0.737	-0.0006	12	-1	0.731
Campylobacter	0.0000	12	1	0.737	0.0015	41	1	0.739
Lactobacillus	0.0000	11	-1	0.737	-0.0005	46	-1	0.741
Alloprevotella	0.0000	10	-1	0.737	0.0005	11	1	0.731
Bulleidia	0.0000	9	1	0.737	0.0002	10	1	0.731
Shuttleworthia	0.0000	8	-1	0.737	0.0001	71	1	0.735
Staphylococcus	0.0000	7	1	0.737	0.0001	7	1	0.732
Treponema	0.0000	6	-1	0.737	0.0001	6	1	0.732
Abiotrophia	0.0000	5	1	0.737	0.0001	5	1	0.732
Absconditabacteriales_.SR1.	0.0000	4	1	0.737	0.0000	4	-1	0.732
Defluviitaleaceae_UCG.011	0.0000	3	-1	0.737	-0.0002	3	-1	0.732
Kingella	0.0000	2	1	0.737	-0.0024	2	-1	0.729

prediction	0.7373	1	X	0.737	0.7292	1	X	0.729

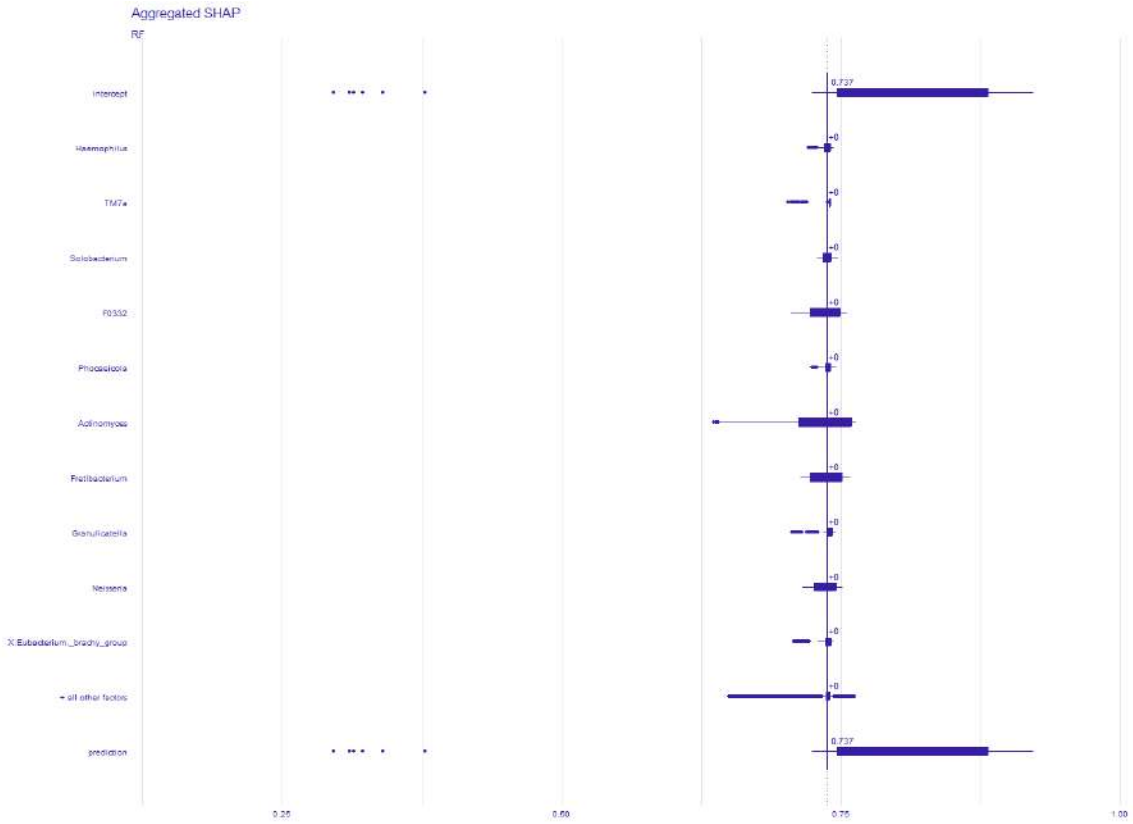


Figure ML3-E-8: The Aggregated SHAP Value in Training dataset for Genus in Saliva

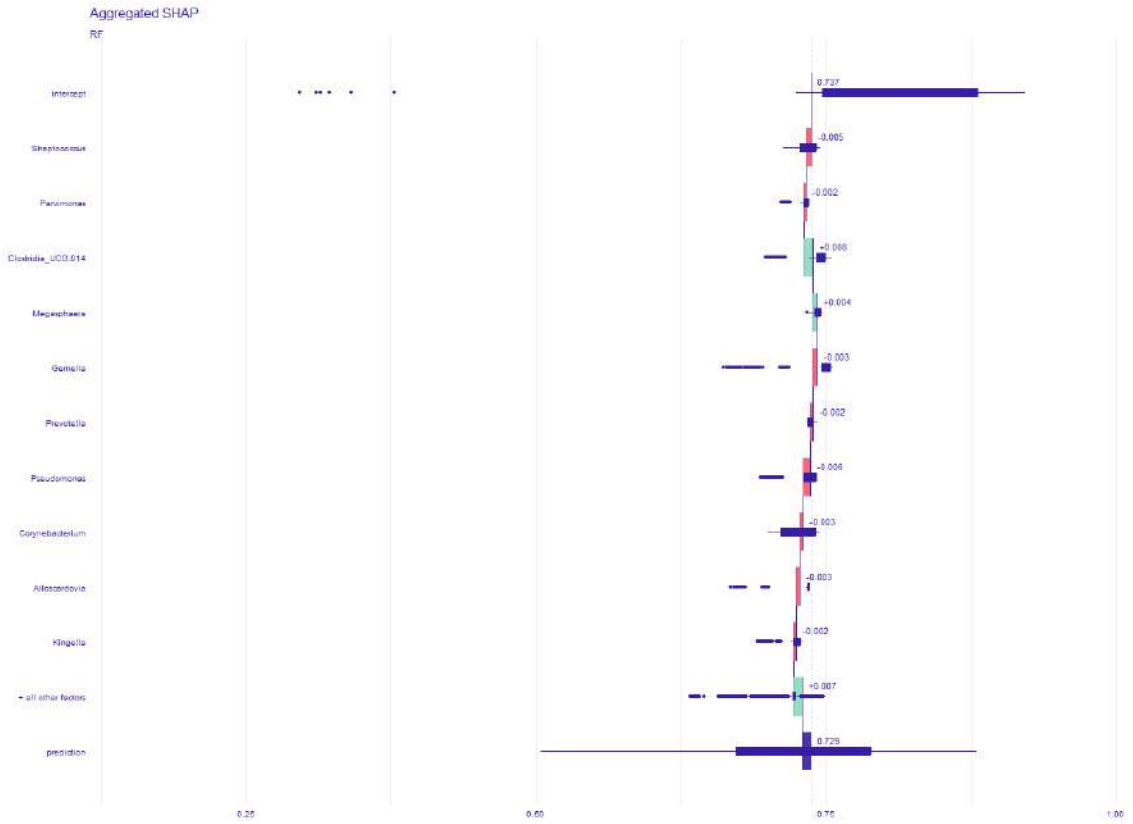


Figure ML3-E-9: The Aggregated SHAP Value in Test dataset for Genus in Saliva

Zero-Inflated Generalized Dirichlet Multinomial (ZIGDM) Tests

The Zero-Inflated Generalized Dirichlet Multinomial (ZIGDM) Tests are provided for each sampling sites and taxa levels with desired false discovery rate equals to 0.05. The lower bound of zero counts for the taxon that needs zero-inflated model is considered 10.

Biopsy

Table ZIGDM 1 - The P-Value for ZIGDM Tests for Biopsy					
Taxa Level	Non-Zero-Inflated ZIGDM		ZIGDM		
	Differential Mean	Differential Dispersion	Differential Presence-Absence Frequency	Differential Mean	Differential Dispersion
Phylum	**0.02**	0.09	**0.00**	0.22	0.33
Class	0.12	0.14	**0.00**	**0.00**	1.00
Order	0.20	**0.00**	**0.00**	1.00	**0.00**
Family	**0.03**	**0.00**	**0.00**	1.00	**0.00**
Genus	**0.00**	**0.00**	**0.00**	1.00	1.00

Bold numbers: Statistically significant with FDR = 0.05

Stool

Table ZIGDM 2- The P-Value for ZIGDM Tests for Stool					
Taxa Level	non-zero-inflated ZIGDM		ZIGDM		
	differential mean	differential dispersion	differential presence-absence frequency	differential mean	differential dispersion
Phylum	0.14	0.09	**0.05**	1.00	0.94
Class	0.22	0.02	0.14	1.00	0.26
Order	**0.02**	0.38	0.24	0.92	1.00
Family	**0.00**	0.76	1.00	0.49	0.75
Genus	**0.00**	**0.00**	**0.00**	1.00	1.00

Bold numbers: Statistically significant with FDR = 0.05

Saliva

Table ZIGDM 3 - The P-Value for ZIGDM Tests for Saliva					
Taxa Level	non-zero-inflated ZIGDM		ZIGDM		
	differential mean	differential dispersion	differential presence-absence frequency	differential mean	differential dispersion
Phylum	0.77	**0.00**	0.86	0.87	**0.00**
Class	0.88	0.08	0.56	1.00	0.27
Order	0.30	**0.01**	0.13	0.35	0.57
Family	0.11	**0.00**	**0.00**	1.00	**0.00**
Genus	0.40	**0.00**	**0.00**	1.00	1.00

Bold numbers: Statistically significant with FDR = 0.05

General

Table ZIGDM 4 - The P-Value for ZIGDM Tests for General					
	non-zero-inflated ZIGDM		ZIGDM		
Taxa Level	differential mean	differential dispersion	differential presence-absence frequency	differential mean	differential dispersion
Phylum	**0.05**	**0.00**	**0.02**	0.62	0.56
Class	0.23	0.17	0.06	0.93	0.41
Order	0.53	0.63	0.54	1.00	1.00
Family	**0.04**	**0.00**	1.00	1.00	1.00
Genus	**0.00**	**0.00**	1.00	1.00	1.00

Bold numbers: Statistically significant with FDR = 0.05