**Supplementary Material**

**Overlapping group screening method**

This study applies overlapping group screening (OGS) to identify gene and interaction biomarkers related to quantitative phenotypes. OGS method has been widely used in genomic data, and the phenotypic features considered include cancer patient survival time (Wang & Chen, 2018), binary classification of cancer and normal tissues (Wang & Chen, 2023), and multivariate classification of cancer subtypes (Wang, Hou & Chen, 2024). Since there are path relationships between genes, different pathways may share the same genes. Therefore, Jacob et al. (2009) proposed the potential effect method, which represents the feature effect vector **β** as a combination of specific group effects, that is,, where represents the potential effect of a specific gene group g, and S is the total number of gene groups (pathways).

For instance, consider six genes organized into five overlapping groups: group 1 includes genes 1 and 2; group 2 includes genes 2 and 3; group 3 includes genes 3 and 4; group 4 includes genes 4 and 5; and group 5 includes genes 5 and 6. Due to overlapping memberships—e.g., gene 2 appears in both groups 1 and 2, gene 3 in groups 2 and 3, and so on—the corresponding effect sizes for the six genes can be expressed as:

In the linear regression model of this study, the steps of overlapping group screening are as follows:

**Step 1:** Use the R package “grpregOverlap” (Zeng & Breheny, 2016) to identify important biological pathways, and assume that P candidate pathways are identified in this step.

**Step 2:** Following the methodology of Wang & Chen (2018), feature variables were constructed based on P candidate pathways, encompassing gene-gene interactions within candidate pathways, between candidate pathways and between candidate and non-candidate pathways. Subsequently, the Sequence Kernel Association Test (SKAT) proposed by Wu et al. (2011) was employed to evaluate group-specific statistical significance.

The definition of SKAT statistic under the linear regression model is:

In step 2, the total number of path combinations is H, ; **m** represents the null model, that is, the linear regression model containing only the intercept term, and its residual vector is estimated by the model; represents the gene-gene interaction information in the candidate path group h, where l is the number of gene-gene interactions in the candidate path group, and represents the j-th gene-gene interaction value of the i-th subject in the candidate path h; is a diagonal weight matrix, and the main diagonal element (j,j) is the weight of the j-th gene-gene interaction in the candidate pathway group h, which is mainly used to improve the effect of statistical analysis. Referring to the method of Wu et al. (2011), an unsupervised weight calculation method is adopted, and the square root of the weight is defined as follows:

and

where is the ratio of the sample variance of the j-th interaction in the pathway group h to the sum of the sample variance of all interactions in the pathway. In other words, these weights are determined based on the relative size of the variance and are modeled using the Beta distribution with parameters (1,25).

Next, using the null model without considering any variables, set the diagonal matrix **V** to:

Whereis the estimated value of under the null model, and I is an n×n unit matrix. Next, set the covariance matrix to:

This assumption assumes that all gene-gene interactions in the candidate pathway interaction group h have no effect.

Under the null hypothesis (i.e., all genes and gene interactions have no effect), the distribution of the SKAT statistic is the sum of weighted chi-square distributions:

are the eigenvalues of ; is an independent random variable that follows a chi-square distribution with 1 degree of freedom. By calculating the mixed chi-square distribution using the Davies method (Davies, 1980), we can obtain the significance level (P-value) of each path group, and use the R package "CompQuadForm" (Duchesne & de Micheaux, 2010) to calculate these P-values as interaction screening indicators. When P-value≦0.05, it can be considered an important interaction feature.

**Step 3:** The selected pathways and the genes and gene interactions in them are used in the subsequent regularized linear regression method (ridge, lasso and elastic net) to build the final prediction model. The regularization step can be performed through the R program package "glmnet" (Simon et al., 2011). In the actual data, the amount of important feature data selected due to P-value ≦ 0.05 is too large, which will affect the model training, so we will adjust the threshold according to different data.

When applying the OGS method, several key assumptions are typically required. First, the overlapping group structure of variables is known in advance, allowing a variable to belong to multiple groups. Second, only a small number of groups are assumed to be truly associated with the response variable, reflecting group sparsity. Third, variables across different groups should have sufficiently low correlation to minimize interference. Additionally, all variables are standardized to have zero mean and unit variance to ensure comparability. Finally, the error term is generally assumed to follow a normal distribution, although other forms of random error may also be acceptable. Together, these assumptions ensure the effectiveness and validity of the OGS method.

**Supplementary Table 1. Advantages and disadvantages of the three regularization functions.**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Lasso | Ridge | Adaptive Lasso |
| ***Advantages*** | **Variable Selection**: By shrinking the coefficients of less important variables to zero, key variables are selected, resulting in a simpler model and reducing the risk of overfitting. **Improved Interpretability**: The simplified model is easier to understand, aiding in the analysis and interpretation of results. **Applicable to High-Dimensional Data**: Lasso is particularly effective when the number of variables exceeds the number of samples, as it can handle high-dimensional data and improve the stability of the model. | **High Stability**: When there is high correlation between variables, Ridge regression retains all variables, preventing the random exclusion of any variable, thus making the model more stable. **Reduced Risk of Overfitting**: By adding a penalty term, Ridge regression limits the complexity of the model, thereby enhancing generalization ability and improving prediction accuracy. | **Consistent Variable Selection:** With a large enough sample, Adaptive Lasso reliably identifies the true non-zero coefficients.  **Reduced Estimation Bias:** Unlike standard Lasso, which applies equal penalties and can bias estimates, Adaptive Lasso adjusts weights to reduce this bias.  **Improved Prediction and Stability:** It typically delivers more accurate variable selection and more stable models, enhancing predictive performance.  **Computationally Efficient:** Compared to complex non-convex methods like SCAD and MCP, Adaptive Lasso is simpler to implement and can be solved using standard Lasso algorithms. |
| ***Disadvantage*** | **Selection Uncertainty**: When variables are highly correlated, Lasso may arbitrarily select one variable while disregarding other equally important variables. **Potential Over-Simplification**: In some cases, Lasso may shrink the coefficients of too many variables to zero, causing the model to lose key information and negatively affecting prediction performance. | **Lack of Variable Selection**: All variables are retained in the model, which can make the model more complex and harder to interpret clearly. **Higher Computational Cost**: When dealing with high-dimensional data, especially with large sample sizes, the computation process may require more resources, resulting in higher computational costs. | **Sensitivity to Initial Estimates:** Relies on OLS for weight calculation, which can be unstable in high-dimensional data or under multicollinearity.  **Extra Hyperparameter:** Requires selecting a tuning parameter; poor choice can affect results, typically addressed via cross-validation.  **Limited Multicollinearity Handling:** Performs better than standard Lasso in reducing bias, but still less effective than Ridge regression in highly correlated settings.  **Higher Computational Cost:** Needs an initial OLS step, making it slower than standard Lasso. |

**Results from Sensitivity Analysis Simulations**

**Supplementary Table 2. The average (standard deviation) of prediction accuracy from 200 simulation analyses under gene structure I with a small magnitude of the true biomarker coefficients.**

| Method | MinMax  Accuracy | MAPE(%) | SMAPE(%) | RMSE | MASE | MAE | MDAE |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Oracle | 0.9584 (0.6787) | 44.3822 (87.6299) | 18.2229 (3.6605) | 2.0492 (0.1506) | 0.0465 (0.0052) | 1.6319 (0.1251) | 1.3789 (0.1550) |
| SIS\_Lasso | -3.8291 (42.3214) | 360.2827 (1058.7723) | 144.9691 (6.5045) | 31.5498 (2.2394) | 0.7176 (0.0569) | 25.2381 (1.9227) | 21.4702 (2.4989) |
| Ordinary\_Lasso | -0.2591 (15.6716) | 217.5971 (451.0018) | 154.1525 (12.0799) | 29.7577 (2.1425) | 0.6764 (0.0488) | 23.8021 (1.8226) | 20.1168 (2.4238) |
| OGS\_Ridge | -0.8168 (61.3900) | 408.3967 (1497.2875) | 137.5026 (12.6376) | 29.0175 (2.9245) | 0.6582 (0.0700) | 23.1543 (2.4597) | 19.5117 (2.7554) |
| OGS\_Lasso | 1.4450 (21.3033) | 483.0494 (1768.3349) | 125.7961 (17.1867) | 27.5165 (3.3677) | 0.6246 (0.0803) | 21.9586 (2.7269) | 18.5099 (2.6316) |
| OGS\_ALasso | 0.6263 (6.5587) | 454.9075 (1192.1604) | 121.6697 (12.6829) | 28.1798 (3.9428) | 0.6410 (0.0991) | 22.5174 (3.2457) | 19.0122 (3.2144) |
| OGS\_G.ridge | 1.4090 (9.8557) | 565.2356 (2303.5608) | 122.0006 (14.3109) | 28.1142 (3.4829) | 0.6380 (0.0905) | 22.4038 (2.8589) | 19.0084 (2.8610) |
| OGS\_G.ridge\_  ALasso | 0.5091 (6.3024) | 462.8222 (1448.6416) | 114.7672 (14.2805) | 26.3068 (4.4508) | 0.5990 (0.1128) | 21.0140 (3.6190) | 17.6727 (3.3996) |
| OGS\_SVR | 2.5321 (8.9954) | 191.8289 (379.7584) | 123.9864 (21.4129) | 24.1813 (3.9237) | 0.5442 (0.0931) | 19.1479 (3.2901) | 15.9412 (3.2786) |
| OGS\_RF | -3.9581 (111.7574) | 217.8561 (498.4318) | 153.0543 (8.5988) | 29.5307 (2.0805) | 0.6705 (0.0471) | 23.5949 (1.7786) | 19.7548 (2.2754) |
| OGS\_KNN | 0.2579 (25.5413) | 242.4793 (640.7222) | 152.5441 (14.4304) | 29.7495 (2.6444) | 0.6745 (0.0614) | 23.7368 (2.2351) | 20.0487 (2.6381) |

**Supplementary Table 3. The average (standard deviation) of prediction accuracy from 200 simulation analyses under gene structure II with a small magnitude of the true biomarker coefficients**

| Method | MinMax  Accuracy | MAPE(%) | SMAPE(%) | RMSE | MASE | MAE | MDAE |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Oracle | 1.0452 (0.5252) | 48.6275 (141.2157) | 17.3159 (10.2152) | 2.2717 (2.5979) | 0.0442 (0.0463) | 1.8117 (2.1299) | 1.5387 (1.9551) |
| SIS\_Lasso | -1.1575 (46.4541) | 332.6069 (517.0225) | 149.5316 (7.1018) | 38.0413 (2.6992) | 0.7480 (0.0557) | 30.4614 (2.3049) | 25.7984 (2.9642) |
| Ordinary\_Lasso | 18.5885 (371.7625) | 177.6429 (244.6892) | 165.1556 (12.0128) | 35.8194 (2.6853) | 0.7027 (0.0464) | 28.6365 (2.2055) | 24.1476 (2.8351) |
| OGS\_Ridge | 1.6044 (24.5463) | 297.3333 (708.0195) | 126.8376 (25.5952) | 30.5233 (6.7541) | 0.6001 (0.1338) | 24.4090 (5.3646) | 20.5941 (4.7858) |
| OGS\_Lasso | 1.6541 (18.0551) | 341.3904 (837.9122) | 123.4776 (29.9574) | 30.3534 (7.0804) | 0.5974 (0.1458) | 24.2525 (5.6600) | 20.3516 (4.9655) |
| OGS\_ALasso | 2.3664 (19.4743) | 427.5722 (951.1080) | 116.2008 (20.1628) | 31.5372 (7.7334) | 0.6192 (0.1578) | 25.1293 (6.0836) | 21.1054 (5.2392) |
| OGS\_G.ridge | 0.8733 (10.6530) | 427.3277 (1101.3571) | 114.9982 (20.5532) | 30.2383 (6.9996) | 0.5942 (0.1440) | 24.1255 (5.5676) | 20.3016 (4.8440) |
| OGS\_G.ridge\_  ALasso | 0.8944 (11.1356) | 435.3688 (1055.6629) | 111.7444 (19.1104) | 30.0367 (7.3106) | 0.5904 (0.1517) | 23.9502 (5.7867) | 20.1657 (5.2009) |
| OGS\_SVR | -4.5813 (61.3677) | 176.2143 (419.5698) | 109.7801 (32.7093) | 24.6430 (7.0751) | 0.4795 (0.1428) | 19.4967 (5.7297) | 16.1519 (5.0252) |
| OGS\_RF | 2.8227 (12.9277) | 202.4689 (345.1017) | 155.1717 (8.5568) | 34.4475 (2.5359) | 0.6756 (0.0451) | 27.5291 (2.0851) | 23.1382 (2.5899) |
| OGS\_KNN | 1.6947 (28.1367) | 220.1376 (360.8229) | 142.6314 (21.4004) | 32.6823 (4.6068) | 0.6412 (0.0939) | 26.0912 (3.7930) | 22.0935 (3.7893) |

**Real data application: CESC**

**Supplementary Table 4. Using the GO\_BP gene set database, the TCGA CESC dataset was randomly split 30 times into training/testing sets at an 80:20 ratio.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | MinMax  Accuracy | MAPE(%) | SMAPE(%) | RMSE | MASE | MAE | MDAE |
| SIS\_Lasso | 0.0531 (8.7067) | 248.8911 (131.0282) | 152.1159 (9.6969) | 1.1792 (0.2019) | 0.8678 (0.1180) | 0.8832 (0.0984) | 0.7426 (0.1237) |
| Ordinary\_Lasso | 20.1688 (62.0916) | 105.9586 (7.7862) | 186.7081 (8.1284) | 0.9961 (0.1446) | 0.7593 (0.0821) | 0.7735 (0.0696) | 0.6911 (0.0778) |
| OGS\_Ridge | -3.1264 (17.3200) | 180.3504 (83.9313) | 152.0367 (12.4191) | 1.0359 (0.1743) | 0.7723 (0.0837) | 0.7882 (0.0865) | 0.6669 (0.0990) |
| OGS\_Lasso | 3.7578 (20.6249) | 207.4932 (90.9235) | 148.0256 (15.9701) | 1.0561 (0.1729) | 0.7866 (0.0973) | 0.8013 (0.0878) | 0.6887 (0.0984) |
| OGS\_ALasso | 2.0008 (11.1095) | 251.1212 (110.3807) | 140.9306 (10.3156) | 1.1249 (0.1792) | 0.8286 (0.1243) | 0.8419 (0.0966) | 0.6884 (0.1058) |
| OGS\_G.ridge | -2.1962 (9.3684) | 226.4970 (102.1251) | 143.7313 (11.8310) | 1.0726 (0.1608) | 0.7984 (0.0982) | 0.8131 (0.0848) | 0.6770 (0.0995) |
| OGS\_G.ridge\_ALasso | -0.7741 (14.2250) | 276.2723 (129.7889) | 139.4741 (10.6391) | 1.1579 (0.1855) | 0.8539 (0.1402) | 0.8670 (0.1061) | 0.6964 (0.1134) |
| OGS\_SVR | 0.7270 (5.9202) | 185.9114 (74.8653) | 145.3829 (11.4404) | 0.9909 (0.1398) | 0.7479 (0.0876) | 0.7621 (0.0796) | 0.6516 (0.0876) |
| OGS\_RF | -0.2228 (22.4445) | 176.6443 (65.8339) | 154.0109 (9.2722) | 1.0192 (0.1361) | 0.7748 (0.0943) | 0.7890 (0.0800) | 0.6811 (0.0895) |
| OGS\_KNN | 6.4971 (27.3139) | 191.7745 (72.3600) | 141.5194 (8.2335) | 1.0061 (0.1381) | 0.7416 (0.0780) | 0.7560 (0.0721) | 0.6273 (0.0934) |

The table reports the mean (standard deviation) of the test prediction performance for various prediction methods.

**Supplementary Table 5. Using the GO\_MF gene set database, the TCGA CESC dataset was randomly split 30 times into training/testing sets at an 80:20 ratio.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | MinMax  Accuracy | MAPE(%) | SMAPE(%) | RMSE | MASE | MAE | MDAE |
| SIS\_Lasso | 1.7899 (5.9797) | 249.0458 (99.5361) | 150.0664 (7.8143) | 1.0921 (0.1461) | 0.8493 (0.0966) | 0.8331 (0.0818) | 0.6886 (0.0947) |
| Ordinary\_Lasso | -35.3408 (136.0865) | 110.8463 (30.1299) | 187.6195 (7.6583) | 0.9580 (0.1427) | 0.7481 (0.0731) | 0.7353 (0.0760) | 0.6407 (0.0984) |
| OGS\_Ridge | -4.4512 (23.0845) | 214.1753 (123.9060) | 146.8542 (15.8648) | 0.9977 (0.1313) | 0.7624 (0.0837) | 0.7485 (0.0781) | 0.6252 (0.1076) |
| OGS\_Lasso | 0.8611 (5.0013) | 248.5835 (118.7604) | 142.3790 (13.9041) | 1.0332 (0.1380) | 0.7882 (0.1012) | 0.7720 (0.0773) | 0.6329 (0.1043) |
| OGS\_ALasso | 1.2332 (7.6218) | 340.3063 (430.9317) | 138.9234 (12.4164) | 1.0906 (0.1706) | 0.8211 (0.1222) | 0.8036 (0.0992) | 0.6405 (0.0983) |
| OGS\_G.ridge | 1.4123 (2.9733) | 252.0512 (148.7774) | 140.9324 (14.2011) | 1.0363 (0.1453) | 0.7855 (0.0993) | 0.7698 (0.0813) | 0.6239 (0.1071) |
| OGS\_G.ridge\_ALasso | 3.5368 (11.0685) | 367.4729 (497.3804) | 138.4963 (12.1262) | 1.1189 (0.1965) | 0.8368 (0.1384) | 0.8180 (0.1066) | 0.6472 (0.1007) |
| OGS\_SVR | 0.6653 (8.1395) | 226.3698 (115.7632) | 146.7785 (10.3533) | 0.9603 (0.1390) | 0.7466 (0.0741) | 0.7340 (0.0801) | 0.6351 (0.1085) |
| OGS\_RF | 6.6768 (23.9390) | 212.5159 (95.0107) | 150.7084 (6.8075) | 0.9887 (0.1393) | 0.7664 (0.0803) | 0.7531 (0.0821) | 0.6366 (0.0970) |
| OGS\_KNN | -37.7916 (213.4297) | 208.2011 (79.9046) | 141.6689 (12.6369) | 0.9710 (0.1371) | 0.7378 (0.0768) | 0.7258 (0.0843) | 0.5823 (0.1058) |

The table reports the mean (standard deviation) of the test prediction performance for various prediction methods.

**Supplementary Table 6. Parameter estimates of candidate genes and gene-gene interactions selected by the proposed OGS\_G.ridge\_ALasso method in CESC.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Gene1 | Gene2 | Parameter  estimation | Gene1 | Gene2 | Parameter  Estimation |
| *AKAP5* |  | 0.0025 | *PIF1* | *REC8* | 0.0188 |
| *DEFA1B* |  | -0.1052 | *CSDE1* | *DCTN1* | 0.0165 |
| *DPF3* |  | 0.0913 | *DCTN1* | *SRPX2* | -0.0174 |
| *EZH1* |  | 0.094 | *DCTN1* | *FANCL* | -0.1005 |
| *HAUS1* |  | 0.026 | *CAPZA2* | *DCTN1* | 0.0548 |
| *PABPN1* |  | 0.1344 | *DCTN1* | *RNF6* | 0.0192 |
| *PRR7* |  | -0.0478 | *EXOSC9* | *TDRD6* | 0.0213 |
| *TAF1A* |  | -0.0344 | *ADAM11* | *TDRD6* | 0.095 |
| *TDRD6* |  | 0.0867 | *CKS2* | *TDRD6* | 0.0677 |
| *UXT* |  | 0.1044 | *SPSB3* | *TDRD6* | -0.0388 |
| *WNT1* |  | 0.0838 | *AK2* | *TDRD6* | -0.0082 |
| *ACTR1B* | *CAPN2* | 0.0103 | *DSTN* | *EZH1* | 0.0051 |
| *CAPZA2* | *MYL6B* | 0.0001 | *EZH1* | *STAU1* | 0.1433 |
| *CRK* | *TMOD4* | -0.1473 | *EZH1* | *ITGAV* | 0.0175 |
| *CRK* | *WASF2* | 0.0609 | *EZH1* | *SPEF2* | 0.0929 |
| *CRK* | *YES1* | -0.0443 | *ART1* | *EZH1* | -0.0913 |
| *DSTN* | *GABARAP* | 0.054 | *CCDC66* | *COLQ* | -0.0029 |
| *DSTN* | *NCKAP1* | -0.0165 | *KPNA4* | *WNT1* | 0.0621 |
| *DSTN* | *TMOD4* | 0.0223 | *ITGAV* | *WNT1* | 0.0341 |
| *MYO7A* | *TNS3* | 0.0436 | *SEC23B* | *WNT1* | -0.0031 |
| *AAAS* | *CDC42BPB* | 0.0689 | *SEC31B* | *WNT1* | -0.0919 |
| *CDC42BPB* | *RDH5* | 0.0492 | *CDC73* | *HAUS1* | -0.0128 |
| *CDC42BPB* | *SKAP1* | -0.0526 | *HAUS1* | *RBBP5* | -0.0046 |
| *CDC42BPB* | *REL* | 0.0181 | *HAUS1* | *TAF5L* | -0.018 |
| *CDC42BPB* | *PSME4* | 0.1241 | *HAUS1* | *THRB* | 0.0102 |
| *DTNBP1* | *UNC45B* | -0.1020 | *HAUS1* | *SEC23B* | -0.0002 |
| *RAD50* | *UNC45B* | 0.0669 | *PABPN1* | *TMEM190* | -0.0133 |
| *PSMD2* | *UNC45B* | -0.0774 | *PABPN1* | *SEC23B* | 0.0002 |
| *MFN1* | *UNC45B* | -0.0642 | *CTAGE6* | *PRR7* | -0.042 |
| *TAF5L* | *UNC45B* | 0.051 | *PRR7* | *RAB11FIP5* | -0.0196 |
| *ADAM11* | *UNC45B* | -0.0106 | *PRKAG2* | *TAF1A* | -0.005 |
| *RNF6* | *UNC45B* | 0.0031 | *PLG* | *TAF1A* | -0.1097 |
| *CENPH* | *EZH1* | 0.0001 | *SLC17A8* | *TAF1A* | -0.0664 |
| *CENPH* | *PIF1* | 0.0397 | *SLC17A8* | *UXT* | 0.1643 |
| *CENPH* | *RIF1* | 0.0016 | *UXT* | *VAMP1* | 0.0691 |
| *DCTN1* | *PIF1* | -0.0397 |  |  |  |

**Real data application: BLCA**

**Supplementary Table 7. Using the GO\_BP gene set database, the TCGA BLCA dataset was randomly split 30 times into training/testing sets at an 80:20 ratio.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | MinMax  Accuracy | MAPE(%) | SMAPE(%) | RMSE | MASE | MAE | MDAE |
| SIS\_Lasso | 1.5644 (7.2442) | 262.4814 (136.4544) | 149.0301 (8.0470) | 1.2880 (0.1152) | 0.7767 (0.0772) | 0.8618 (0.0726) | 0.6333 (0.0810) |
| Ordinary\_Lasso | -12.4648 (92.0597) | 136.0988 (65.1091) | 177.3163 (13.0154) | 1.2279 (0.1141) | 0.7369 (0.0798) | 0.8169 (0.0682) | 0.5970 (0.0753) |
| OGS\_Ridge | -2.4329 (16.7077) | 197.6315 (92.4992) | 142.6200 (14.6388) | 1.2078 (0.1155) | 0.7147 (0.0830) | 0.7924 (0.0765) | 0.5768 (0.0927) |
| OGS\_Lasso | 0.6028 (2.4851) | 229.3415 (119.1073) | 137.2149 (11.7991) | 1.2345 (0.1268) | 0.7267 (0.0913) | 0.8053 (0.0833) | 0.5775 (0.1000) |
| OGS\_ALasso | 1.4937 (5.5512) | 250.7560 (126.9351) | 135.1622 (11.4598) | 1.2470 (0.1379) | 0.7397 (0.0984) | 0.8197 (0.0918) | 0.5927 (0.1066) |
| OGS\_G.ridge | 0.6187 (4.0078) | 225.5697 (105.7379) | 136.7069 (11.3242) | 1.2148 (0.1154) | 0.7196 (0.0904) | 0.7970 (0.0777) | 0.5880 (0.0832) |
| OGS\_G.ridge\_ALasso | 1.4920 (2.7965) | 270.2740 (132.9057) | 135.2700 (11.5652) | 1.2647 (0.1496) | 0.7520 (0.1022) | 0.8337 (0.1011) | 0.6036 (0.1121) |
| OGS\_SVR | 0.6389 (3.8483) | 221.3159 (113.3104) | 141.9437 (8.4648) | 1.2036 (0.1155) | 0.7049 (0.0759) | 0.7821 (0.0742) | 0.5583 (0.0775) |
| OGS\_RF | 8.3798 (42.1409) | 197.1931 (67.2098) | 151.0516 (6.2525) | 1.2175 (0.1089) | 0.7258 (0.0741) | 0.8051 (0.0692) | 0.5875 (0.0766) |
| OGS\_KNN | 0.8915 (3.2880) | 190.9952 (113.9106) | 148.9540 (12.1178) | 1.2031 (0.1163) | 0.7087 (0.0745) | 0.7867 (0.0758) | 0.5691 (0.0985) |

The table reports the mean (standard deviation) of the test prediction performance for various prediction methods.

**Supplementary Table 8. Using the GO\_MF gene set database, the TCGA BLCA dataset was randomly split 30 times into training/testing sets at an 80:20 ratio.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | MinMax  Accuracy | MAPE(%) | SMAPE(%) | RMSE | MASE | MAE | MDAE |
| SIS\_Lasso | 2.6739 (5.1272) | 213.3999 (96.4175) | 153.8119 (8.3240) | 1.0617 (0.1881) | 0.7921 (0.0855) | 0.7606 (0.0738) | 0.6110 (0.0836) |
| Ordinary\_Lasso | -5.8013 (99.3215) | 115.1257 (34.6661) | 183.1417 (12.9778) | 1.0068 (0.1750) | 0.7549 (0.0805) | 0.7251 (0.0702) | 0.5716 (0.0645) |
| OGS\_Ridge | -11.6251 (47.1261) | 183.1881 (110.6257) | 163.7035 (21.5812) | 1.0079 (0.2123) | 0.7504 (0.0924) | 0.7208 (0.0871) | 0.5766 (0.0820) |
| OGS\_Lasso | -3.0338 (11.8708) | 256.4838 (108.5379) | 140.8392 (13.6534) | 1.0177 (0.1903) | 0.7578 (0.0808) | 0.7293 (0.0865) | 0.5778 (0.0761) |
| OGS\_ALasso | 15.8059 (129.7423) | 253.5053 (115.5538) | 141.6256 (13.6415) | 1.0126 (0.1902) | 0.7575 (0.0892) | 0.7281 (0.0847) | 0.5829 (0.0923) |
| OGS\_G.ridge | 1.0626 (6.7606) | 278.9754 (101.4271) | 135.6927 (9.4756) | 1.0639 (0.1944) | 0.7679 (0.1262) | 0.7343 (0.0844) | 0.5618 (0.0769) |
| OGS\_G.ridge\_ALasso | 2.8761 (15.1256) | 300.0674 (144.6875) | 138.4194 (13.3292) | 1.1644 (0.3954) | 0.8218 (0.1718) | 0.7842 (0.1188) | 0.6032 (0.0754) |
| OGS\_SVR | 0.6223 (2.8717) | 241.0870 (94.2696) | 138.3407 (9.8613) | 0.9769 (0.1737) | 0.7317 (0.0907) | 0.7029 (0.0798) | 0.5431 (0.0843) |
| OGS\_RF | -1.0985 (31.7244) | 195.7773 (65.6551) | 146.8154 (5.5192) | 0.9767 (0.1715) | 0.7319 (0.0759) | 0.7035 (0.0721) | 0.5599 (0.0595) |
| OGS\_KNN | -2.5397 (11.0732) | 194.4358 (85.1062) | 153.4499 (9.4050) | 0.9902 (0.1674) | 0.7479 (0.0729) | 0.7187 (0.0673) | 0.5763 (0.0563) |

The table reports the mean (standard deviation) of the test prediction performance for various prediction methods.

**Supplementary Table 9. Parameter estimates of candidate genes and gene-gene interactions selected by the proposed OGS\_G.ridge\_ALasso method in BLCA.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Gene1 | Gene2 | Parameter  estimation | Gene1 | Gene2 | Parameter  Estimation |
| *ATP6V1B1* |  | 0.0766 | *TRAPPC2* |  | -0.0863 |
| *C1QL2* |  | 0.1271 | *ZFYVE1* |  | 0.0748 |
| *C8A* |  | 0.1287 | *CHMP2A* | *LMOD2* | 0.0154 |
| *CDHR3* |  | 0.0205 | *DYNC2LI1* | *LMOD2* | 0.0006 |
| *CETN2* |  | -0.0345 | *IFT81* | *LMOD2* | -0.0008 |
| *DNAI2* |  | -0.1826 | *LMOD2* | *SPAG9* | -0.0002 |
| *EPHA4* |  | 0.0462 | *LMOD2* | *TTC12* | 0.0258 |
| *GYPB* |  | 0.0975 | *HUS1* | *LMOD2* | 0.0208 |
| *HDAC1* |  | -0.0845 | *LMOD2* | *SYT2* | -0.0643 |
| *HSD3B2* |  | 0.0339 | *LMOD2* | *OGDH* | -0.0917 |
| *HUS1* |  | 0.0575 | *LMOD2* | *TRIOBP* | -0.1217 |
| *IFT81* |  | -0.0630 | *INTS10* | *LMOD2* | 0.0589 |
| *LILRA4* |  | 0.0618 | *LMOD2* | *MAFG* | -0.0771 |
| *LMOD2* |  | 0.1186 | *LMOD2* | *RBM28* | -0.0230 |
| *MIOS* |  | 0.0850 | *LMOD2* | *YAP1* | -0.0656 |
| *MSH3* |  | -0.0331 | *HPR* | *LMOD2* | 0.0623 |
| *NDEL1* |  | 0.0192 | *HUS1* | *UXT* | 0.0492 |
| *NME7* |  | -0.0745 | *CYTH3* | *HUS1* | 0.0235 |
| *PPP1R3A* |  | -0.1414 | *HUS1* | *KSR1* | 0.0060 |
| *TCP11* |  | -0.0585 | *HUS1* | *MACF1* | -0.0532 |
| *THOC3* |  | -0.0318 | *EXOC7* | *HDAC1* | -0.1439 |

**Real data application: LIHC**

**Supplementary Table 10. Using the GO\_BP gene set database, the TCGA LIHC dataset was randomly split 30 times into training/testing sets at an 80:20 ratio.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | MinMax  Accuracy | MAPE(%) | SMAPE(%) | RMSE | MASE | MAE | MDAE |
| SIS\_Lasso | -0.9922 (4.0083) | 342.7927 (188.7815) | 138.6740 (8.7724) | 1.0672 (0.4187) | 0.8841 (0.1637) | 0.6302 (0.0870) | 0.4609 (0.0600) |
| Ordinary\_Lasso | -1.1217 (7.0237) | 306.7290 (202.7388) | 141.4448 (14.2908) | 1.0332 (0.4555) | 0.8318 (0.1294) | 0.5964 (0.0894) | 0.4351 (0.0588) |
| OGS\_Ridge | 4.0547 (37.8266) | 315.8196 (272.9904) | 139.7894 (18.0418) | 1.0189 (0.4305) | 0.8162 (0.1414) | 0.5828 (0.0801) | 0.4258 (0.0643) |
| OGS\_Lasso | -0.6518 (5.1287) | 427.4216 (328.9769) | 133.2293 (15.5566) | 1.0672 (0.3993) | 0.8752 (0.1808) | 0.6200 (0.0807) | 0.4507 (0.0763) |
| OGS\_ALasso | -17.5795 (111.2548) | 429.9222 (324.8058) | 131.9655 (14.9719) | 1.0776 (0.3881) | 0.8831 (0.1834) | 0.6242 (0.0709) | 0.4525 (0.0736) |
| OGS\_G.ridge | -2.8429 (7.3606) | 363.4461 (279.7922) | 134.7505 (15.6782) | 1.0394 (0.4132) | 0.8513 (0.1636) | 0.6034 (0.0674) | 0.4463 (0.0760) |
| OGS\_G.ridge\_ALasso | -1.0157 (5.4305) | 444.7022 (331.0401) | 132.3112 (14.8653) | 1.1022 (0.3741) | 0.9167 (0.2124) | 0.6445 (0.0725) | 0.4699 (0.0831) |
| OGS\_SVR | -0.7410 (7.1773) | 290.5293 (238.9638) | 140.0152 (18.0356) | 0.9824 (0.4700) | 0.7761 (0.1034) | 0.5588 (0.0886) | 0.4193 (0.0566) |
| OGS\_RF | -1.5961 (19.6048) | 335.0581 (220.9181) | 137.8434 (11.3571) | 1.1064 (0.4082) | 0.8923 (0.2551) | 0.6257 (0.1006) | 0.4546 (0.0900) |
| OGS\_KNN | -0.3388 (2.9468) | 287.7770 (200.9595) | 138.8150 (16.5772) | 1.0095 (0.4603) | 0.8114 (0.1359) | 0.5794 (0.0784) | 0.4409 (0.0694) |

The table reports the mean (standard deviation) of the test prediction performance for various prediction methods.

**Supplementary Table 11. Using the GO\_MF gene set database, the TCGA LIHC dataset was randomly split 30 times into training/testing sets at an 80:20 ratio.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | MinMax  Accuracy | MAPE(%) | SMAPE(%) | RMSE | MASE | MAE | MDAE |
| SIS\_Lasso | -1.1177 (4.5382) | 305.9241 (163.0560) | 138.6182 (9.8832) | 0.8941 (0.3488) | 0.8619 (0.1338) | 0.5993 (0.0859) | 0.4490 (0.0702) |
| Ordinary\_Lasso | -1.3510 (12.4165) | 268.0794 (193.8693) | 152.9986 (24.9677) | 0.8639 (0.3736) | 0.8164 (0.0832) | 0.5713 (0.0884) | 0.4351 (0.0507) |
| OGS\_Ridge | -5.8595 (26.4950) | 244.1295 (208.4544) | 149.4075 (19.4835) | 0.8606 (0.3643) | 0.8104 (0.0907) | 0.5660 (0.0817) | 0.4448 (0.0582) |
| OGS\_Lasso | 12.7977 (51.5075) | 321.6163 (243.1693) | 147.4421 (22.2404) | 0.9247 (0.3370) | 0.8851 (0.1306) | 0.6129 (0.0679) | 0.4709 (0.0684) |
| OGS\_ALasso | -1.8984 (15.8613) | 354.6904 (220.7595) | 139.1051 (12.5314) | 0.9375 (0.3250) | 0.9019 (0.1433) | 0.6218 (0.0532) | 0.4783 (0.0449) |
| OGS\_G.ridge | -1.2855 (3.8769) | 303.1838 (192.2798) | 140.9880 (12.6113) | 0.9095 (0.3353) | 0.8772 (0.1244) | 0.6074 (0.0650) | 0.4856 (0.0638) |
| OGS\_G.ridge\_ALasso | -1.1865 (4.4173) | 398.6475 (229.1483) | 138.0139 (11.7827) | 0.9929 (0.3049) | 0.9612 (0.1705) | 0.6599 (0.0500) | 0.5105 (0.0526) |
| OGS\_SVR | -20.1296 (84.8268) | 231.0044 (189.3996) | 152.1234 (15.9736) | 0.8342 (0.3859) | 0.7859 (0.0754) | 0.5499 (0.0815) | 0.4317 (0.0569) |
| OGS\_RF | -1.3956 (5.7879) | 309.1213 (255.9910) | 143.4065 (11.2968) | 0.9495 (0.3926) | 0.8918 (0.2388) | 0.6127 (0.1103) | 0.4696 (0.0872) |
| OGS\_KNN | -8.2010 (53.3963) | 233.9725 (154.0554) | 145.0358 (14.2234) | 0.8505 (0.3740) | 0.8070 (0.0855) | 0.5632 (0.0766) | 0.4438 (0.0578) |

The table reports the mean (standard deviation) of the test prediction performance for various prediction methods.

**Supplementary Table 12. Parameter estimates of candidate genes and gene-gene interactions selected by the proposed OGS\_G.ridge\_ALasso method in LIHC.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Gene1 | Gene2 | Parameter  estimation | Gene1 | Gene2 | Parameter  Estimation |
| *BRF2* |  | -0.0261 | *FUT5* | *PLA2G4F* | 0.0952 |
| *DCLRE1C* |  | -0.1113 | *FUT5* | *WASH3P* | 0.0523 |
| *DUOX1* |  | -0.0596 | *KCNB1* | *TMED3* | -0.038 |
| *HYAL3* |  | -0.1263 | *B4GALT2* | *CD28* | -0.0124 |
| *MYO5A* |  | 0.0913 | *B4GALT2* | *CSF2RB* | -0.0037 |
| *NCKAP5* |  | 0.0292 | *B4GALT2* | *GFRA3* | 0.0599 |
| *OPN4* |  | 0.1163 | *B4GALT2* | *HTR3B* | -0.0085 |
| *PIGR* |  | 0.1389 | *B4GALT2* | *PCSK6* | -0.1004 |
| *PLEC* |  | -0.121 | *BTN3A1* | *COG3* | 0.0057 |
| *PPP3R2* |  | 0.0983 | *CCR2* | *FUT5* | -0.087 |
| *RGPD1* |  | 0.0202 | *CD14* | *FUT5* | -0.0211 |
| *RGPD8* |  | -0.0184 | *CD163* | *COG3* | -0.0425 |
| *SCN1A* |  | 0.0081 | *CD226* | *COG3* | -0.0465 |
| *SLC24A4* |  | 0.0307 | *CD28* | *FUT5* | 0.0367 |
| *UQCRHL* |  | -0.0646 | *CD28* | *SAR1B* | 0.0128 |
| *WDR59* |  | 0.0385 | *COG3* | *FOLH1* | 0.01 |
| *ALPK1* | *C3ORF14* | 0.0082 | *COG3* | *IL17RC* | 0.0259 |
| *ALPK1* | *HTR6* | -0.0281 | *COG3* | *PCSK6* | 0.0419 |
| *ALPK1* | *SLC24A4* | 0.0201 | *FUT5* | *HTR3B* | -0.0359 |
| *ALPK1* | *USH2A* | 0.0003 | *FUT5* | *ITGA7* | 0.0165 |
| *ANKS6* | *OPN4* | -0.0195 | *FUT5* | *MR1* | -0.0488 |
| *ANKS6* | *USH2A* | -0.0145 | *FUT5* | *SELL* | -0.0798 |
| *ATP1A1* | *C3ORF14* | 0.0298 | *HTR3B* | *TMED3* | -0.0259 |
| *ATP1A1* | *HYAL3* | -0.0117 | *ITGA7* | *TMED3* | -0.047 |
| *ATP1A1* | *PPP3R2* | -0.0243 | *SELL* | *TMED3* | -0.0017 |
| *C3ORF14* | *CYLD* | 0.0295 | *CCNE1* | *COG3* | 0.013 |
| *C3ORF14* | *FUZ* | 0.0295 | *CNTLN* | *FUT5* | 0.0172 |
| *C3ORF14* | *HTR6* | -0.0131 | *FUT5* | *GPR174* | 0.1549 |
| *C3ORF14* | *KLC3* | 0.0174 | *FUT5* | *TUBE1* | -0.0401 |
| *C3ORF14* | *SLC24A4* | -0.0353 | *GPR174* | *TMED3* | -0.0005 |
| *DRD2* | *PPP3R2* | 0.0319 | *SAR1B* | *TUBE1* | 0.0902 |
| *DRD2* | *TAS2R43* | -0.0146 | *GNAS* | *SAR1B* | -0.0111 |
| *FUZ* | *OPN4* | -0.052 | *COG3* | *SAA1* | -0.0168 |
| *FUZ* | *PPP3R2* | -0.016 | *FUT5* | *HBA2* | 0.0385 |
| *FUZ* | *PRKACG* | 0.0318 | *HBA2* | *SAR1B* | -0.0468 |
| *FUZ* | *RAB34* | 0.0172 | *FPR2* | *FUT5* | 0.0976 |
| *HDAC6* | *HYAL3* | 0.0155 | *FUT5* | *LILRA3* | 0.056 |
| *HDAC6* | *PPP3R2* | -0.1156 | *ADRA1B* | *FUT5* | 0.0563 |
| *HDAC6* | *PRKACG* | 0.0848 | *B4GALT2* | *STOM* | -0.037 |
| *HYAL3* | *SLC24A4* | -0.0972 | *RFTN1* | *SAR1B* | 0.0453 |
| *KLC3* | *OPN4* | -0.0162 | *ABCC9* | *FUT5* | 0.0445 |
| *OPN4* | *PPP3R2* | 0.0623 | *AP3B2* | *B4GALT2* | 0.007 |
| *OPN4* | *SLC24A4* | 0.0692 | *COG3* | *SEL1L* | -0.0162 |
| *PPP3R2* | *RAB34* | -0.0192 | *FUT5* | *OS9* | -0.0996 |
| *SLC24A4* | *TAS2R43* | -0.0205 | *FUT5* | *UGT3A1* | -0.1087 |
| *SLC24A4* | *USH2A* | 0.0043 | *B4GALT2* | *ZFP36* | 0.0018 |
| *CAT* | *SLC24A4* | -0.0116 | *FUT5* | *MEX3A* | -0.0683 |
| *FGL2* | *SLC24A4* | -0.0196 | *FUT5* | *ZFP36* | 0.0253 |
| *FPR2* | *SLC24A4* | -0.1278 | *MEX3A* | *TMED3* | -0.0056 |
| *ITGAX* | *SLC24A4* | 0.1117 | *SAR1B* | *ZFP36* | -0.0183 |
| *LRG1* | *SLC24A4* | -0.0503 | *FUT5* | *PIGR* | 0.119 |
| *AHSP* | *SLC24A4* | 0.0281 | *B4GALT2* | *MMP25* | -0.0213 |
| *CHIT1* | *SLC24A4* | 0.0892 | *B4GALT2* | *PLA1A* | 0.0008 |
| *B4GALT2* | *PDLIM5* | -0.0153 | *FUT5* | *HYAL3* | 0.0294 |
| *FUT5* | *PDLIM5* | -0.0566 | *B4GALT2* | *BDH1* | 0.0506 |
| *GBP1* | *TMED3* | 0.0104 | *COG3* | *GBP5* | -0.017 |
| *PDLIM5* | *TMED3* | -0.0115 | *CHIT1* | *FUT5* | -0.009 |
| *SAR1B* | *TPM3* | -0.0248 | *B4GALT2* | *PRKACG* | 0.0326 |
| *SAR1B* | *WASF2* | -0.0026 | *CHRAC1* | *FUT5* | -0.1093 |
| *B4GALT2* | *PLEC* | 0.0095 | *COG3* | *PHKA1* | 0.0254 |
| *CAT* | *FUT5* | -0.1342 | *FUT5* | *PHKA1* | -0.0279 |
| *COG3* | *DLG2* | 0 | *PRKACG* | *SAR1B* | -0.0026 |
| *COG3* | *EPCAM* | 0.0308 | *ABCC11* | *B4GALT2* | -0.0325 |
| *FUT5* | *GDI2* | -0.0344 | *B4GALT2* | *SNX20* | -0.0413 |
| *FUT5* | *MPDZ* | -0.0057 | *DKK1* | *FUT5* | 0.015 |
| *FUT5* | *ZAP70* | 0.1189 | *FUT5* | *RABEPK* | 0.1031 |
| *PLEC* | *SAR1B* | -0.0009 | *CAT* | *RGPD1* | -0.0419 |
| *RHOB* | *SAR1B* | -0.0325 | *CR1* | *RGPD1* | 0.037 |
| *SAR1B* | *SCN1A* | -0.0377 | *RGPD1* | *SIGLEC14* | 0.0174 |
| *PRPF3* | *TMED3* | 0.0472 | *LCP1* | *RGPD8* | 0.0443 |
| *B4GALT2* | *TESC* | 0 | *LPXN* | *RGPD8* | 0.0952 |
| *COG3* | *KCNB1* | -0.0098 | *CHIT1* | *RGPD8* | 0.0172 |
| *COG3* | *PLA2G4F* | 0.0016 | *HP* | *RGPD1* | -0.0565 |
| *FUT5* | *KCNB1* | 0.1439 | *RGPD1* | *STOM* | 0.0801 |

**Real data application: ESCA**

**Supplementary Table 13. Using the GO\_BP gene set database, the TCGA ESCA dataset was randomly split 30 times into training/testing sets at an 80:20 ratio.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | MinMax  Accuracy | MAPE(%) | SMAPE(%) | RMSE | MASE | MAE | MDAE |
| SIS\_Lasso | 1.7606 (1.4343) | 267.2291 (109.2936) | 110.6799 (11.9051) | 0.9193 (0.1919) | 0.7525 (0.1351) | 0.6367 (0.1030) | 0.4529 (0.1084) |
| Ordinary\_Lasso | -0.2769 (15.4756) | 205.8751 (96.6397) | 117.9105 (13.4214) | 0.8920 (0.2246) | 0.7172 (0.1026) | 0.6117 (0.1085) | 0.4291 (0.0911) |
| OGS\_Ridge | 1.0405 (3.7115) | 231.2552 (106.6499) | 115.6394 (11.4732) | 0.9119 (0.2043) | 0.7582 (0.1176) | 0.6431 (0.1006) | 0.4738 (0.0813) |
| OGS\_Lasso | 2.2096 (7.0094) | 250.8380 (115.0649) | 112.5197 (10.7387) | 0.9126 (0.2007) | 0.7567 (0.1243) | 0.6407 (0.0994) | 0.4644 (0.0864) |
| OGS\_ALasso | -2.8137 (30.1747) | 262.9772 (124.8667) | 112.1385 (9.8781) | 0.9138 (0.2034) | 0.7603 (0.1209) | 0.6446 (0.1019) | 0.4800 (0.0791) |
| OGS\_G.ridge | 1.7169 (2.9358) | 256.3931 (121.5902) | 113.2383 (10.4282) | 0.9060 (0.1970) | 0.7629 (0.1228) | 0.6459 (0.0958) | 0.4917 (0.0797) |
| OGS\_G.ridge\_ALasso | 3.7870 (16.0944) | 271.8954 (125.7532) | 111.6942 (9.9012) | 0.9231 (0.2035) | 0.7659 (0.1260) | 0.6492 (0.1046) | 0.4793 (0.0993) |
| OGS\_SVR | 0.9027 (3.4788) | 226.7182 (118.5072) | 111.6753 (11.3446) | 0.8794 (0.2043) | 0.7156 (0.1190) | 0.6069 (0.0982) | 0.4366 (0.0908) |
| OGS\_RF | 1.5321 (1.3222) | 255.2190 (129.3081) | 113.7589 (12.0656) | 0.9316 (0.1966) | 0.7683 (0.1535) | 0.6492 (0.1082) | 0.4571 (0.1025) |
| OGS\_KNN | 0.5048 (4.8983) | 255.7891 (143.6638) | 112.8689 (12.4645) | 0.9081 (0.2114) | 0.7313 (0.1100) | 0.6218 (0.1029) | 0.4366 (0.0855) |

The table reports the mean (standard deviation) of the test prediction performance for various prediction methods.

**Supplementary Table 14. Using the GO\_MF gene set database, the TCGA ESCA dataset was randomly split 30 times into training/testing sets at an 80:20 ratio.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | MinMax  Accuracy | MAPE(%) | SMAPE(%) | RMSE | MASE | MAE | MDAE |
| SIS\_Lasso | 1.1139 (2.5875) | 245.7376 (126.7428) | 109.3042 (11.3689) | 0.9093 (0.1806) | 0.8312 (0.1702) | 0.6328 (0.0866) | 0.4489 (0.1064) |
| Ordinary\_Lasso | 1.0473 (2.5047) | 197.1554 (94.9034) | 117.1210 (10.0946) | 0.8740 (0.2226) | 0.7891 (0.1295) | 0.6057 (0.0916) | 0.4314 (0.0738) |
| OGS\_Ridge | 1.0033 (2.1484) | 233.5219 (111.9955) | 110.6552 (11.2765) | 0.8609 (0.2031) | 0.7900 (0.1300) | 0.6073 (0.0977) | 0.4463 (0.0955) |
| OGS\_Lasso | 1.1739 (2.1166) | 252.5739 (125.6124) | 107.7254 (12.3926) | 0.8734 (0.2060) | 0.8014 (0.1439) | 0.6158 (0.1074) | 0.4595 (0.1021) |
| OGS\_ALasso | 1.9396 (2.4935) | 270.5448 (128.7228) | 107.7721 (12.6883) | 0.8862 (0.2120) | 0.8143 (0.1519) | 0.6256 (0.1127) | 0.4515 (0.1058) |
| OGS\_G.ridge | 1.1121 (2.1640) | 243.0344 (114.5546) | 109.5380 (10.6710) | 0.8537 (0.1997) | 0.7956 (0.1395) | 0.6105 (0.0979) | 0.4552 (0.0876) |
| OGS\_G.ridge\_ALasso | 3.0075 (5.4905) | 261.8762 (128.7067) | 107.4987 (13.0112) | 0.8788 (0.2098) | 0.8059 (0.1480) | 0.6197 (0.1111) | 0.4484 (0.1022) |
| OGS\_SVR | 2.0019 (5.8171) | 224.1460 (130.1150) | 108.4642 (12.3920) | 0.8228 (0.2160) | 0.7506 (0.1175) | 0.5793 (0.1078) | 0.4252 (0.0872) |
| OGS\_RF | 3.0651 (8.7217) | 232.9411 (129.8664) | 108.3360 (11.4510) | 0.8435 (0.2271) | 0.7707 (0.1187) | 0.5963 (0.1211) | 0.4430 (0.0902) |
| OGS\_KNN | 1.3616 (1.9645) | 230.9537 (129.8228) | 105.0033 (12.3747) | 0.8311 (0.2054) | 0.7480 (0.1119) | 0.5740 (0.0819) | 0.3996 (0.0668) |

The table reports the mean (standard deviation) of the test prediction performance for various prediction methods.

**Supplementary Table 15. Parameter estimates of candidate genes and gene-gene interactions selected by the proposed OGS\_G.ridge\_ALasso method in ESCA.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Gene1** | **Gene2** | **Parameter**  **estimation** | **Gene1** | **Gene2** | **Parameter**  **Estimation** |
| *ACE* |  | 0.0112 | *CYS1* | *HYDIN* | -0.1003 |
| *AQP2* |  | 0.0106 | *CYS1* | *IFT80* | -0.0342 |
| *C15ORF48* |  | 0.1944 | *CYS1* | *TTLL6* | -0.1347 |
| *CALM3* |  | 0.0023 | *DYNLRB2* | *RSPH1* | -0.0834 |
| *FXR1* |  | -0.1141 | *HYDIN* | *RILPL1* | -0.0457 |
| *ITPKA* |  | 0.1207 | *IFT80* | *IQCA1* | -0.0204 |
| *MCM5* |  | -0.0738 | *IFT80* | *NEDD1* | 0.0091 |
| *SLC26A3* |  | 0.0101 | *IQCA1* | *PRKACG* | 0.0788 |
| *AQP2* | *ATP8A1* | 0.1091 | *IQCA1* | *SMO* | 0.1443 |
| *COPZ2* | *ITPKA* | -0.0013 | *KIF3C* | *SLC26A3* | -0.1287 |
| *ITPKA* | *SLC1A1* | 0.048 | *KIF9* | *SLC26A3* | -0.1892 |
| *AQP2* | *SLC26A3* | 0.0314 | *KIF9* | *TTLL6* | -0.0943 |
| *CALM3* | *CD44* | -0.0609 | *MLF1* | *SMO* | 0.02 |
| *CALM3* | *SLC3A2* | 0.048 | *SLC26A3* | *STX3* | 0.0749 |
| *ERBB2* | *SLC26A3* | 0.0241 | *SLC26A8* | *TTLL6* | -0.0187 |
| *NDRG4* | *SLC26A3* | 0.0357 | *ACE* | *DYNC1I1* | 0.1605 |
| *SLC12A2* | *SLC26A3* | -0.0246 | *ACE* | *NCAPH2* | 0.0348 |
| *SLC26A3* | *SLC3A2* | 0.1353 | *CALM3* | *CHMP2B* | 0.0486 |
| *ACE* | *CLUAP1* | 0.0082 | *CALM3* | *DYNC1I1* | 0.0274 |
| *ACE* | *CYS1* | -0.1372 | *CALM3* | *MAPK3* | 0.0181 |
| *ACE* | *PRKACG* | 0.0457 | *CALM3* | *RNF128* | 0.1014 |
| *ACE* | *SLC26A3* | -0.0031 | *CALM3* | *SLC30A3* | 0.1016 |
| *ADCY3* | *GPR161* | 0.0449 | *CALM3* | *VPS37B* | -0.0593 |
| *ADCY3* | *IFT80* | 0.0697 | *CLCN5* | *SLC26A3* | 0.0457 |
| *ADCY3* | *SLC26A3* | 0.023 | *GPC1* | *SLC26A3* | 0.0085 |
| *ADCY3* | *STX3* | 0.0591 | *MAPK3* | *SLC26A3* | 0.0066 |
| *ADCY6* | *CATSPERB* | 0.0188 | *CALM3* | *LGALS3* | 0.1347 |
| *BBS10* | *CALM3* | 0.0949 | *ABCD3* | *ACE* | 0.0294 |
| *BBS10* | *CLUAP1* | -0.0027 | *CALM3* | *CLDND1* | 0.0727 |
| *CALM3* | *SLC26A3* | 0.0177 | *CALM3* | *KIAA1614* | 0.1115 |
| *CATSPERB* | *IQCA1* | 0.06 | *CALM3* | *SLC6A7* | 0.1416 |
| *CATSPERB* | *IQCE* | 0.1674 | *CALM3* | *SPACA4* | 0.029 |
| *CATSPERB* | *RILPL1* | 0.0388 | *GNB4* | *SLC26A3* | -0.018 |
| *CCDC68* | *SLC26A3* | 0.0824 | *ATP8A1* | *SLC26A3* | 0.0516 |
| *CLUAP1* | *KIF9* | 0.0087 | *CALM3* | *OLFM4* | -0.0217 |
| *CLUAP1* | *SLC26A8* | 0.0562 |  |  |  |

**Gene-Specific Kaplan-Meier Survival Curves by BMI Groups in Four Cancers**



**Supplementary Fig. 1. Kaplan-Meier Survival Curves Stratified by *WNT1* Expression (Median Cutoff) Across Different BMI Groups (Normal, Overweight, Obese) in CESC.**



**Supplementary Fig. 2. Kaplan-Meier Survival Curves Stratified by *CRK* Expression (Median Cutoff) Across Different BMI Groups (Normal, Overweight, Obese) in CESC.**



**Supplementary Fig. 3. Kaplan-Meier Survival Curves Stratified by *HUS1* Expression (Median Cutoff) Across Different BMI Groups (Normal, Overweight, Obese) in BLCA.**



**Supplementary Fig. 4. Kaplan-Meier Survival Curves Stratified by *CAT* Expression (Median Cutoff) Across Different BMI Groups (Normal, Overweight, Obese) in LIHC.**

**Supplementary Fig. 5. Kaplan-Meier Survival Curves Stratified by *ZFP36* Expression (Median Cutoff) Across Different BMI Groups (Normal, Overweight, Obese) in LIHC.**



**Supplementary Fig. 6. Kaplan-Meier Survival Curves Stratified by *B4GALT2* Expression (Median Cutoff) Across Different BMI Groups (Normal, Overweight, Obese) in LIHC.**



**Supplementary Fig. 7. Kaplan-Meier Survival Curves Stratified by *ACE* Expression (Median Cutoff) Across Different BMI Groups (Normal, Overweight, Obese) in ESCA.**