The supplementary file containing the additional analysis results for the dataset "**GSE55546**". There are twelve samples in the dataset, 4 belongs to the Uvula (control) and 12 HPV active OPSCC. This dataset is the only available match to compare the results with the proposed study. However, the sample size is very less to conduct machine learning interpretations. So, in order to understand the similarity and reproducibility between the datasets, this experiment is conducted. Figure 1 represents the volcano plot generated during the analysis between the Uvula control and HPV active OPSCC samples.

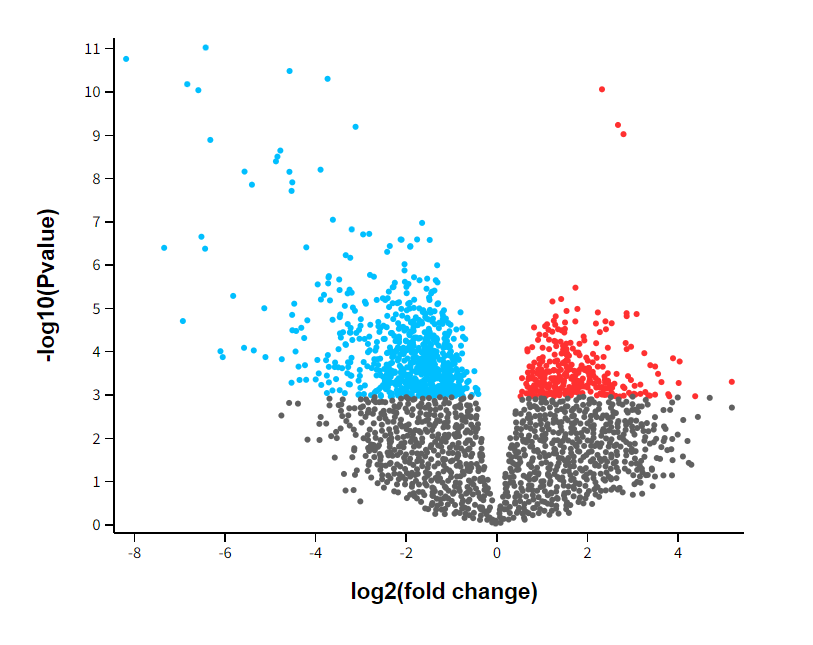


Fig. 1 Volcano Plot representing the differentially expressed genes (Up and Down regulated) in Red and Blue color respectively.

Table 1: The genes *LAMC2*, *ECT2* and *DSG2* identified in the proposed system and its statistical significance on GSE55546 dataset.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **padj** | **pvalue** | **t** | **B** | **log2FoldChange** | **SPOT\_ID** |
| LAMC2 | 4.67E-03 | 8.74E-06 | -6.44281 | 3.815943 | -3.35572083 | A\_23\_P201636 |
| ECT2 | 1.06E-01 | 5.34E-03 | -3.22799 | -2.232359 | -1.45006552 | A\_23\_P44684 |
| DSG2 | 1.38E-01 | 9.70E-03 | -2.94138 | -2.788324 | -1.92251001 | A\_23\_P141730 |

Based on the results from the Table 1, it’s clear that the genes *LAMC2*, *ECT2* and *DSG2* are down-regulated, which thereby confirms the hypothesis deduced in the proposed system. The results evidently proves that the down-regulation of the above genes causes significant impact on the pathogenesis of OPSCC and the survival rate.

Table 2. Comparison of logFC scores between GSE112026 and GSE55546

|  |  |  |  |
| --- | --- | --- | --- |
| **GSE55546** | **padj** | **pvalue** | **log2FoldChange** |
| LAMC2 | 4.67E-03 | 8.74E-06 | -3.35572083 |
| ECT2 | 1.06E-01 | 5.34E-03 | -1.45006552 |
| DSG2 | 1.38E-01 | 9.70E-03 | -1.92251001 |

|  |  |  |  |
| --- | --- | --- | --- |
| **GSE112026** | **padj** | **pvalue** | **log2FoldChange** |
| LAMC2 | 3.17E-34 | 3.24E-37 | -3.27969357 |
| ECT2 | 1.09E-31 | 1.68E-34 | -2.200989687 |
| DSG2 | 7.97E-64 | 1.81E-67 | -3.049925518 |