We applied random forest (RF) to test the prediction performance of the low-pass WGBD data and we found the overall out-of-bag prediction accuracy was 92.16%. We noticed 1 of 43 non-HCC sample was predicted to HCC while 8 HCC was predicted to non-HCC and 3 mistake classified HCC samples were all belong to early HCC. We also applied five-fold cross-validate and with 100 random resampling in RF approach and find the average sensitivity, specificity and accuracy in test dataset were 62.5%, 97.6% and 91.1% respectively.

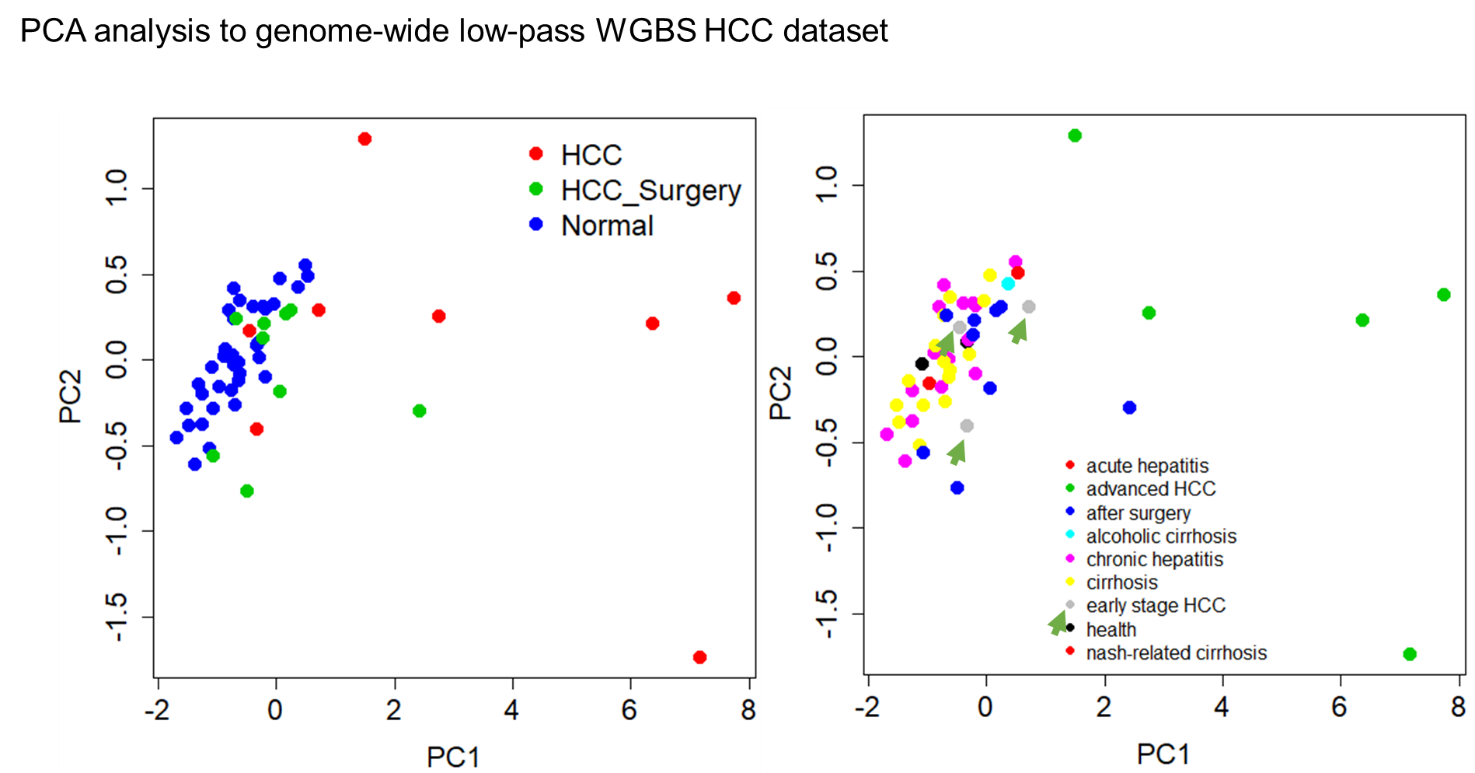


Table: Here, show the first 40 times RF prediction details:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Spe.train | Sen.train | Accu.train | Spe.test | Sen.test | Acc.test |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| 3 | 1 | 1 | 1 | 0.888888889 | 0.5 | 0.818181818 |
| 4 | 1 | 1 | 1 | 1 | 0 | 0.818181818 |
| 5 | 1 | 1 | 1 | 0.833333333 | 0.5 | 0.7 |
| 6 | 1 | 1 | 1 | 1 | 1 | 1 |
| 7 | 1 | 1 | 1 | 1 | 0 | 0.909090909 |
| 8 | 1 | 1 | 1 | 1 | 1 | 1 |
| 9 | 1 | 1 | 1 | 0.875 | 0.5 | 0.8 |
| 10 | 1 | 1 | 1 | 1 | 1 | 1 |
| 11 | 1 | 1 | 1 | 1 | 0 | 0.818181818 |
| 12 | 1 | 1 | 1 | 1 | 1 | 1 |
| 13 | 1 | 1 | 1 | 1 | 1 | 1 |
| 14 | 1 | 1 | 1 | 0.875 | 1 | 0.9 |
| 15 | 1 | 1 | 1 | 1 | 1 | 1 |
| 16 | 1 | 1 | 1 | 1 | 0.666667 | 0.909090909 |
| 17 | 1 | 1 | 1 | 1 | 0 | 0.818181818 |
| 18 | 1 | 1 | 1 | 1 | 0.5 | 0.9 |
| 19 | 1 | 1 | 1 | 1 | 0 | 0.818181818 |
| 20 | 1 | 1 | 1 | 1 | 0.5 | 0.909090909 |