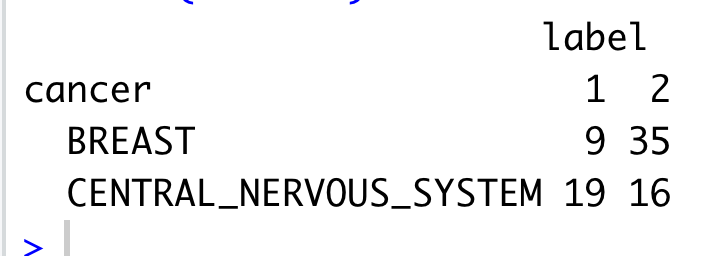
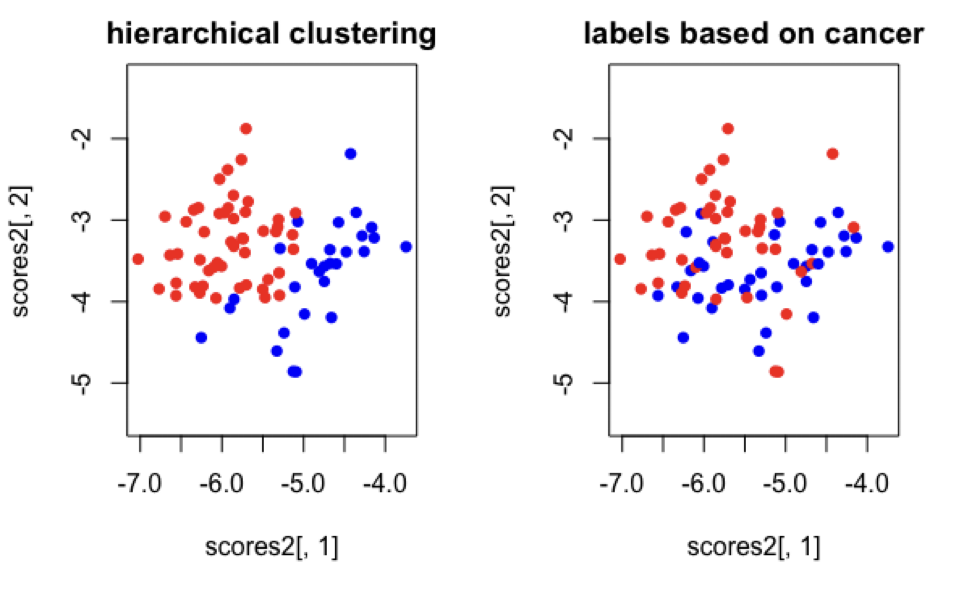
We have seen that the effectiveness of treatment does not always depend on the type of cancer.

In this section, I have considered data related to CENTRAL\_NERVOUS\_SYSTEM AND BREAST CANCER. These cancers show to be particularly different from the AUC point of view. This means that treatment efficacy will depend in some way by the type of cancer.

1. HIERARCHICAL CLUSTER

The labels we obtain are more or less similar to the type of cancers.





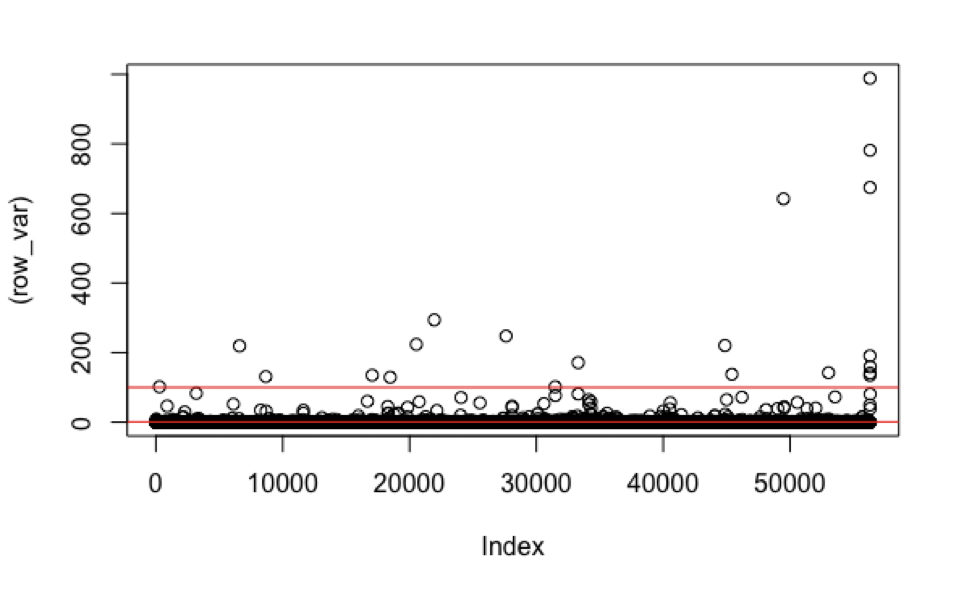
As we can see, the two auc-based clusters are not only based on tumour type, although this seems to have an influence\*.

1. REDUCE THE NUMBER OF GENES

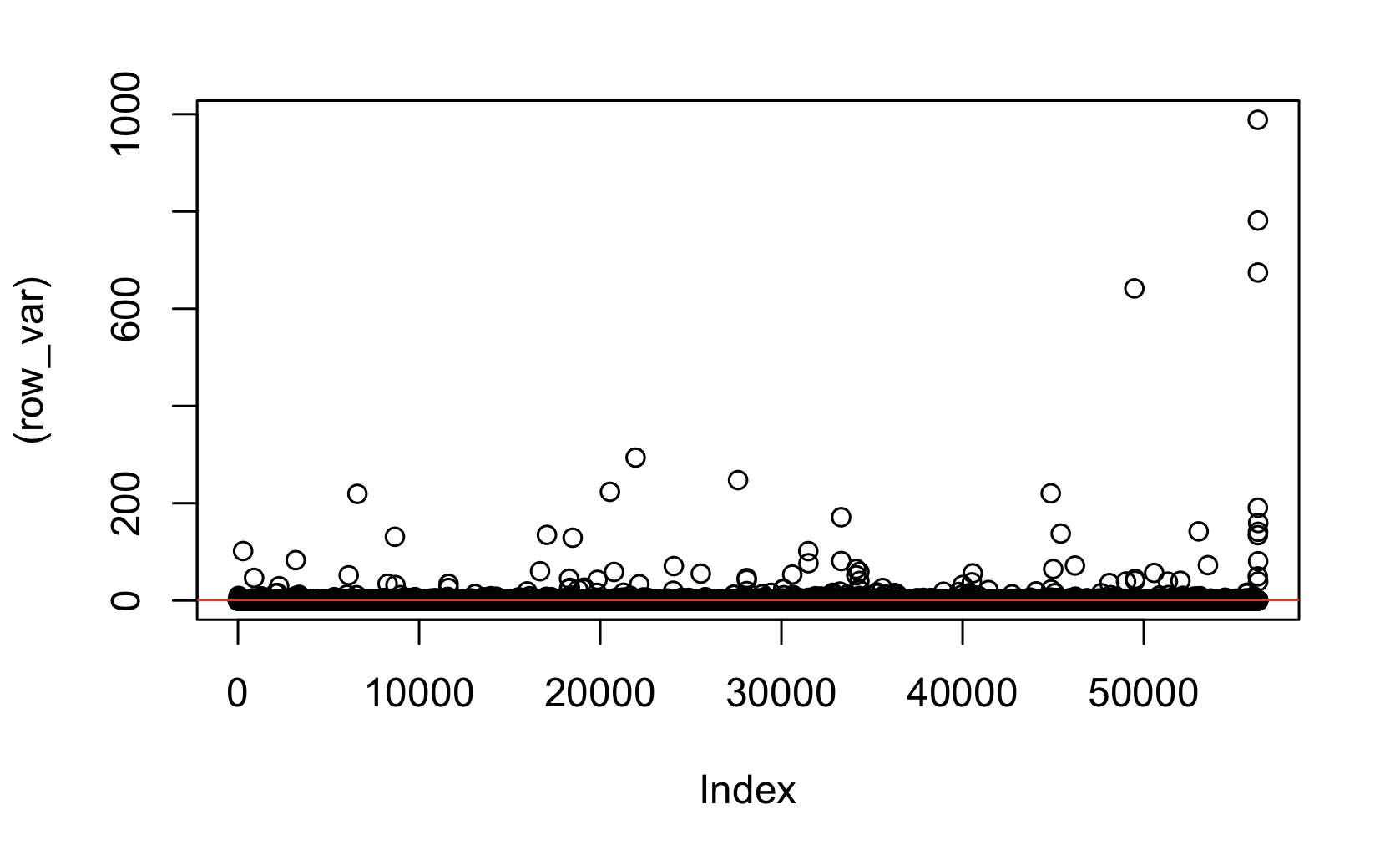
Hyperactive ribosomal biogenesis is widely observed in cancer, which has been partly attributed to the increased rDNA transcription by Pol I in cancer. However, whether **small nucleolar RNAs (snoRNAs), a class of non-coding RNAs crucial in ribosomal RNA (rRNA) maturation and functionality, are involved in cancer remains elusive**.

The crucial point now is how to select genes to be considered in next analysis.

To face this problem, we have first created a separated matrix of normalized data (each cell line will have the same total variance🡪 we are reducing influence of patient). In the following graphs, you can see genes variabilities.



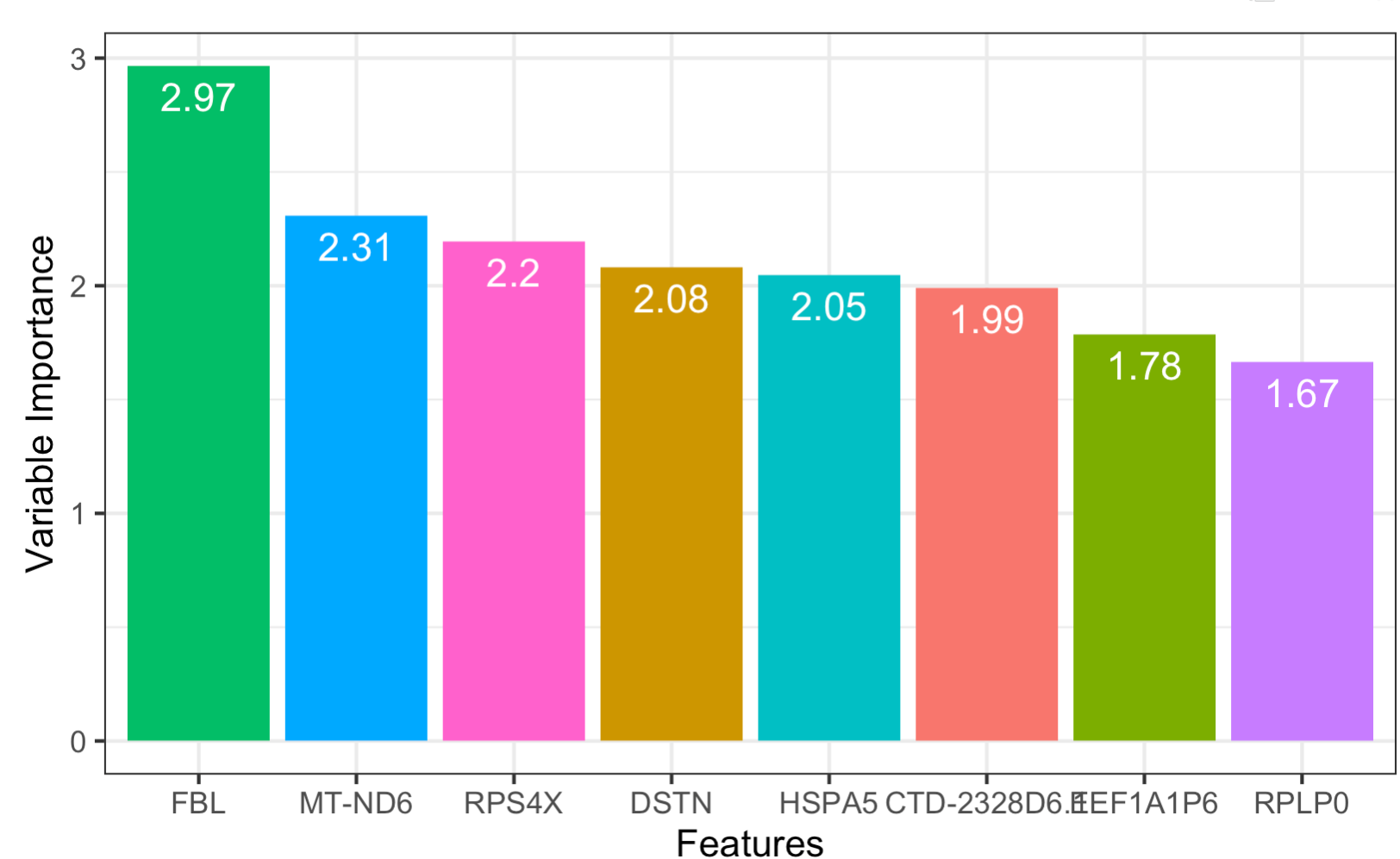
~~Since the variability of some genes between different patients is really high, classification would be based ONLY on these genes. A classification tree using this data suggests some non-coding RNAs as crucial while their importance is not verified in the literature. For this reason, we have removed these “genes” with too high variability~~. We will not consider genes with almost 0 variability. This could be non-expressed genes or 0 could by generated by errors. In any case, low-variability variables won’t have a lot of importance for the classification algorithm 🡪488 genes



1. CLASSIFICATION TREE

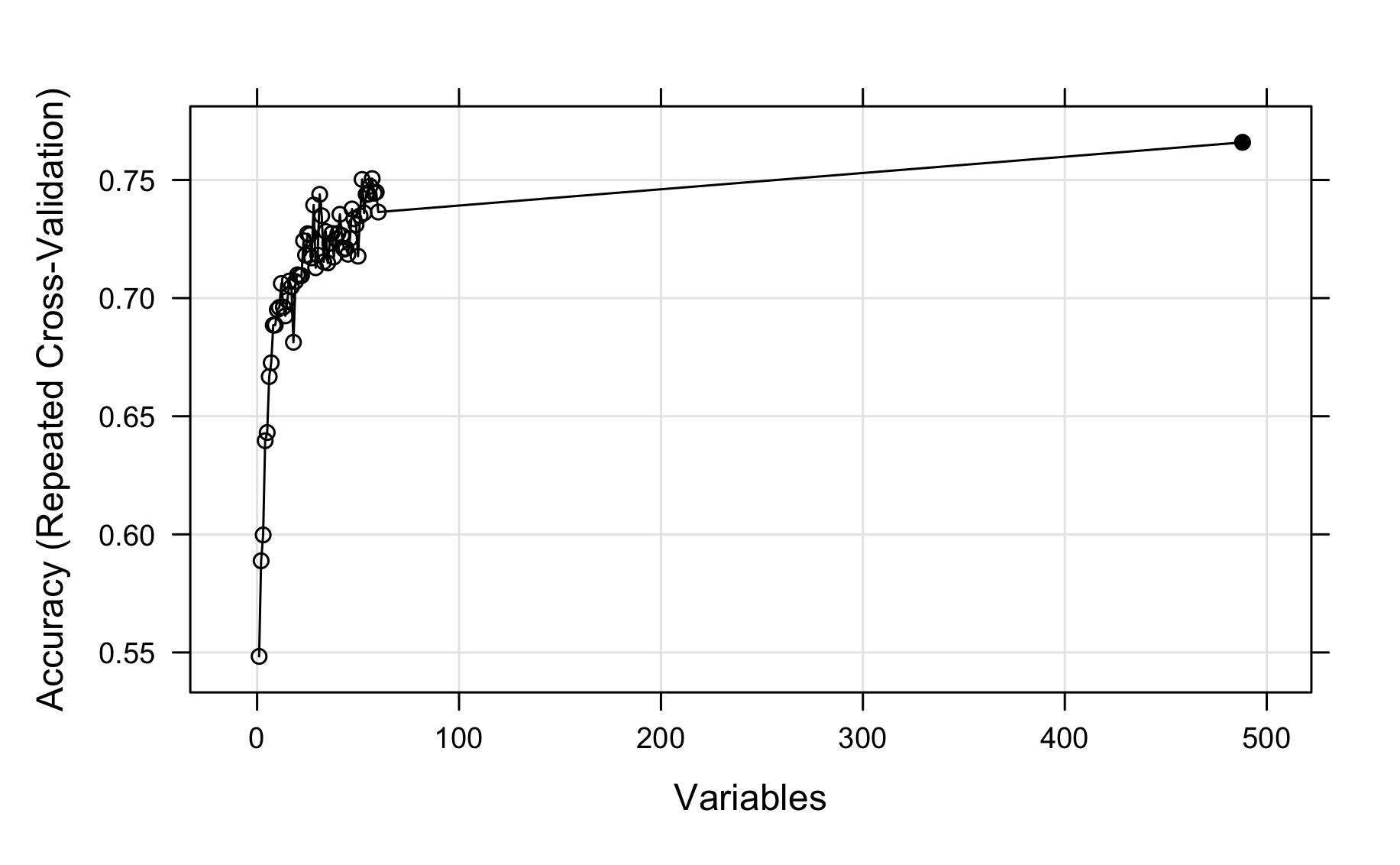
I più importanti classificatori sono i seguenti:

* If i consider only a lower bound

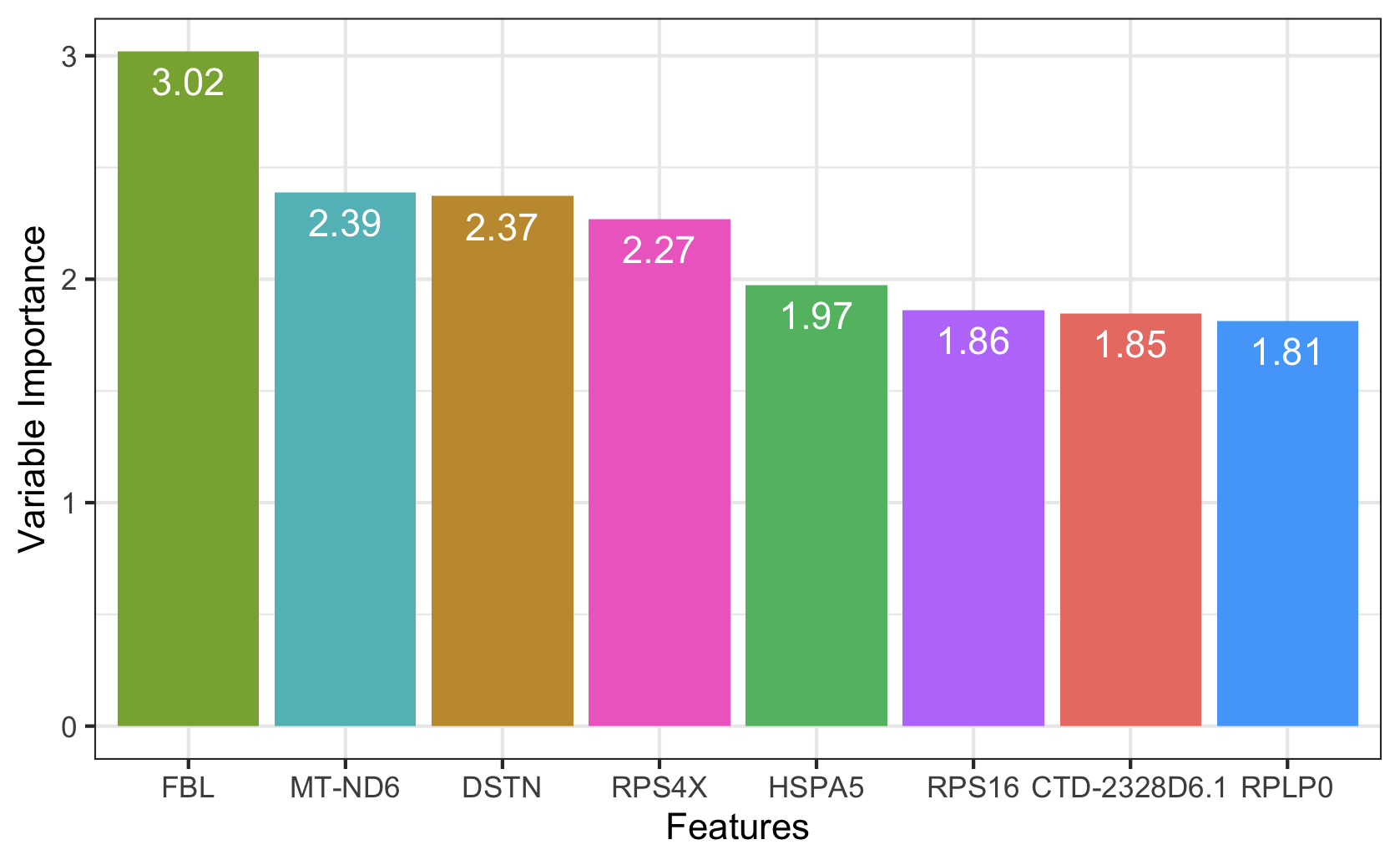


Accuracy Kappa

0.8000000 0.4705882

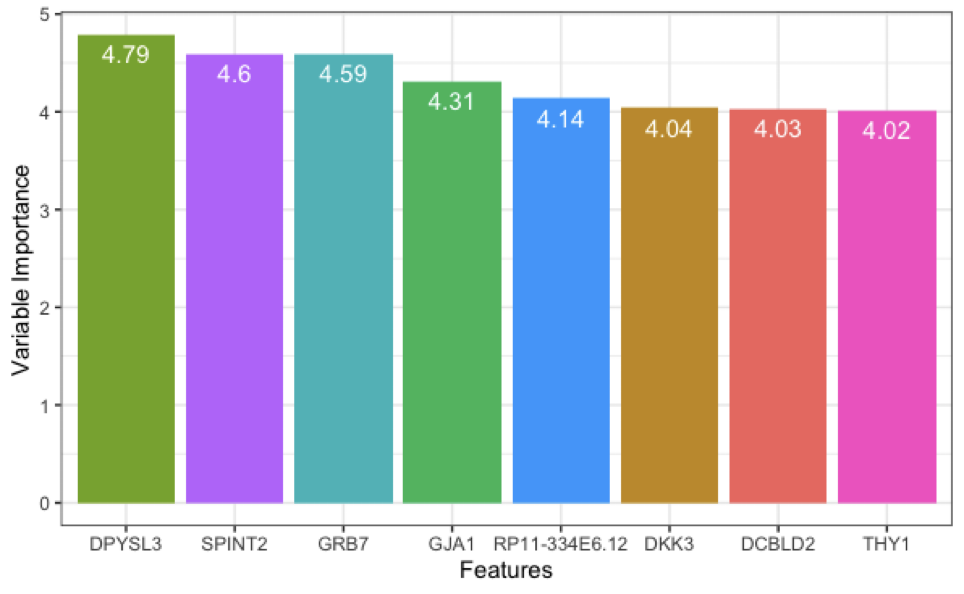


- if I consider also an upper bound



\*What we now wonder is whether these results are simply related to the fact that the two tumors are in different parts of the body and therefore have different gene expression, or whether the groups found earlier also contain in some sense some additional information about the auc.

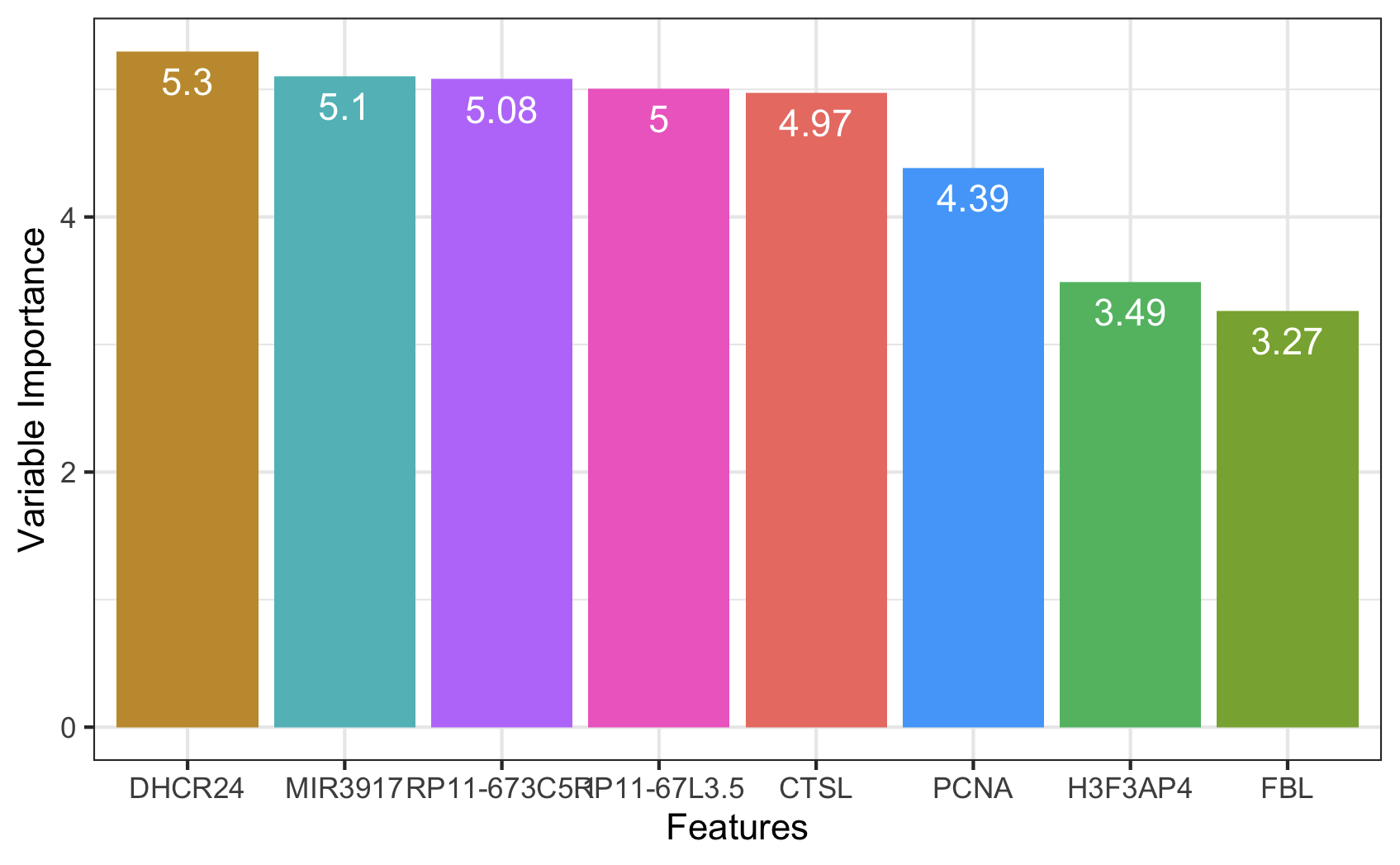
Using simply tumor as a group label, we would obtain the following influential genes:



Accuracy Kappa

0.9333333 0.8648649

**CONSIDERING ONLY BREAST**



RMSE Rsquared MAE

0.6958371 0.1768890 0.6607458

**DHCR24**

These results indicate that DHCR24 promotes the growth of breast cancer stem-like cells in part through enhancing the Hedgehog signaling pathway. Our data suggest that cholesterol contribute to breast carcinogenesis by enhancing Hedgehog signaling and cancer stem-like cell populations. Enzymes including DHCR24 involved in cholesterol biosynthesis should be considered as potential treatment targets for breast cancer. <https://pubmed.ncbi.nlm.nih.gov/32713162/#:~:text=These%20results%20indicate%20that%20DHCR24,cancer%20stem%2Dlike%20cell%20populations>.

**CSTL is influential**

OTHERS not clear if they are influential

Using breast

The confusion matrix is

class.assigned

class.true 1 2 3

1 9 5 0

2 2 21 0

3 1 2 4

The APER of LDA is 0.2272727

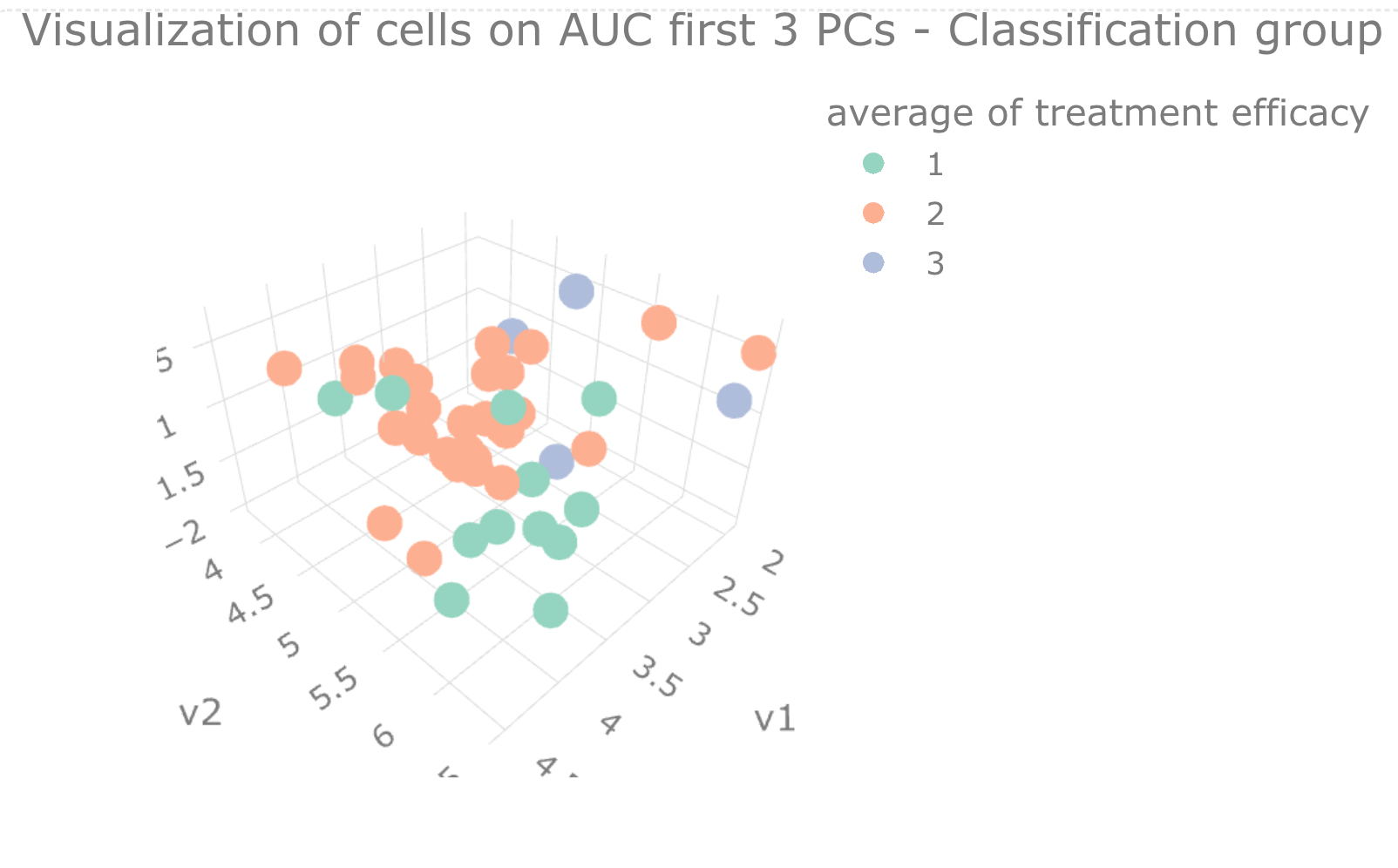
class.assigned

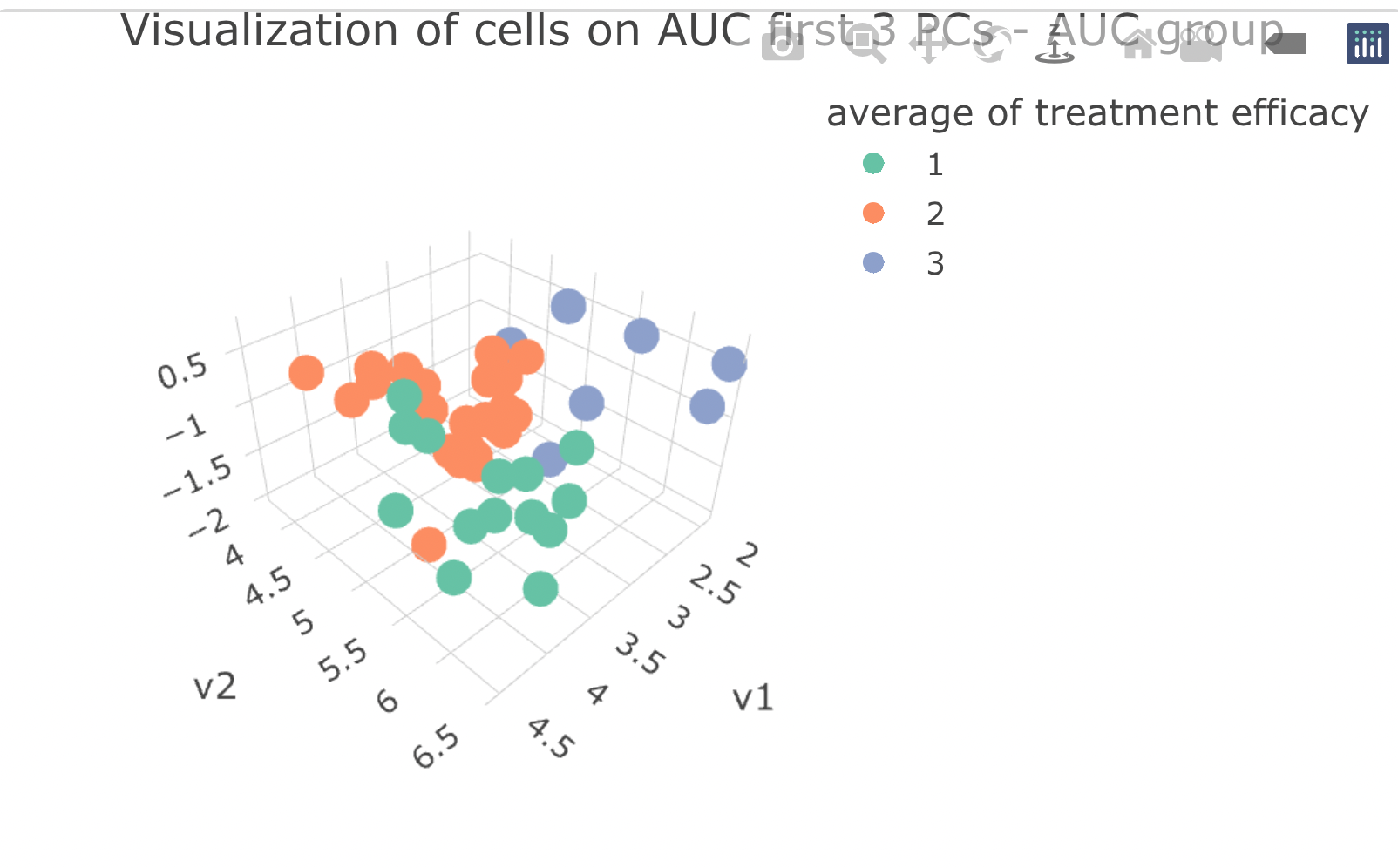
class.true 1 2 3

1 6 7 1

2 5 14 4

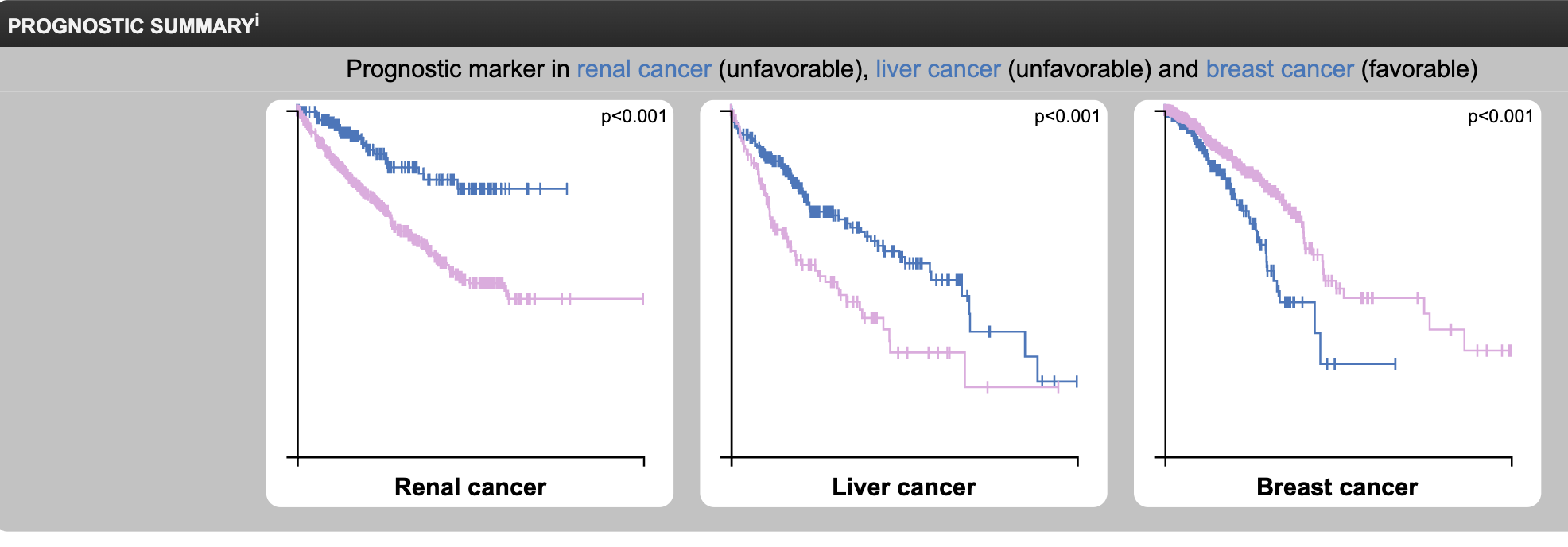
3 2 3 2





ANALYSIS OF INFLUENTIAL GENES

* **FBL (ferritin bearing lymphocytes)**

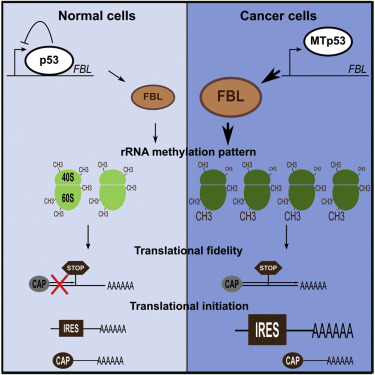


* FBL was found to be a significant predictor for early breast cancer. Positive FBL is associated with early manifestation of breast cancer and may be considered as a tool for the screening of breast cancer in high risk women.[[1]](#footnote-1)
* Flora Nguyen Van Long discovered in 2022 that in addition to breast tumours expressing high level of FBL, about 10% of the breast tumors express low level of FBL. *A correlation* ***between low FBL mRNA level and lack of FBL detection at protein level using immunohistochemistry was observed****. Interestingly, multivariate analyses revealed that these low FBL tumors displayed poor outcome compared to current clinical gold standards. Transcriptomic data revealed that FBL expression is proportionally associated with distinct amount of ribosomes, low FBL level being associated with low amount of ribosomes. Moreover, the molecular programs supported by low and high FBL expressing tumors were distinct.*

*ConclusionAltogether, we identified FBL as a powerful ribosome biogenesis-related independent marker of breast cancer outcome. Surprisingly we unveil a dual association of the ribosome biogenesis FBL factor with prognosis. These data suggest that hyper- but also hypo-activation of ribosome biogenesis are molecular traits of distinct tumors.[[2]](#footnote-2)*

* **Hyperactive ribosomal biogenesis is widely observed in cancer***, which has been partly attributed to the increased rDNA transcription by Pol I in cancer. However, whether small nucleolar RNAs (snoRNAs), a class of non-coding RNAs crucial in ribosomal RNA (rRNA) maturation and functionality, are involved in cancer remains elusive. We report that snoRNAs and fibrillarin (FBL, an enzymatic small nucleolar ribonucleoprotein, snoRNP)* ***are frequently overexpressed in both murine and human breast cancer*** *as well as in prostate cancers, and significantly, that* ***this overexpression is essential for tumorigenicity in vitro and in vivo****. We demonstrate that when the elevated snoRNA pathway is suppressed, the tumor suppressor p53\* can act as a sentinel of snoRNP perturbation, the activation of which mediates the growth inhibitory effect. On the other hand, high level of FBL interferes with the activation of p53 by stress. We further show that p53 activation by FBL knockdown is not only regulated by the ribosomal protein-MDM2-mediated protein stabilization pathway, but also by enhanced PTB-dependent, cap-independent translation. Together, our data uncover an essential role of deregulated snoRNA biogenesis in tumors and a new mechanism of nucleolar modulation of p53.*

**\*NOTE: p53** is crucial in vertebrates, where it prevents cancer formation. As such, p53 has been described as "the guardian of the genome" because of its role in conserving stability by preventing genome mutation.*[[3]](#footnote-3)*



**MT-ND6c??**

**RPS4X : X-linked ribosomal protein S4**

studies have suggested that RPS4X may be important in tumor progression, and demonstrated that RPS4X physically interacts with Y-box binding protein-1 (YB-1) in breast and ovarian cancer cell lines. The RPS4X/YB-1 complex is critical in counteracting cisplatin resistance in MCF7 and MDA-MB-231 breast cancer cells. Immunohistochemistry studies indicated that high expression of RPS4X was associated with a lower risk of death and later disease progression as compared to low expression of RPS4X. [[4]](#footnote-4)

**HSPA5 Regulates Ferroptotic Cell Death in Cancer Cells**

echanistically, activating transcription factor 4 (ATF4) resulted in the induction of HSPA5, which in turn bound glutathione peroxidase 4 (GPX4) and protected against GPX4 protein degradation and subsequent lipid peroxidation. Importantly, the HSPA5-GPX4 pathway mediated ferroptosis resistance, limiting the anticancer activity of gemcitabine.

**REG1A**

High levels of REG1A expression within tumors are an independent predictor of poor prognosis in patients with breast cancer.

**Only 1 patient has high level of reg1a: 'HCC1569\_BREAST'**

HER2-positive breast cancer is a breast cancer that tests positive for a protein called human epidermal growth factor receptor 2 (HER2). This protein promotes the growth of cancer cells. In about 1 of every 5 breast cancers, the cancer cells have extra copies of the gene that makes the HER2 protein. Breast cancer cells with higher than normal levels of HER2 are called HER2-positive. These cancers tend to grow and spread faster than breast cancers that are ..

ALTRO [TO BE REMOVED]



Immagine che contiene testo

Descrizione generata automaticamente

Immagine che contiene testo

Descrizione generata automaticamente

~~DOUBTS: Without taking out data, I receive these as classifiers~~

~~[1] "X169" "X189" "X33" "X377" "X468" "X87" "X192" "X95" "X163" "X162"~~

~~[11] "X128" "X451" "X58" "X210" "X49" "X243" "X413~~"

While the highest variabilities are associated with:

1] "ZNF143" "CACTIN-AS1" "HELQ"

[4] "SCLT1" "SMIM12" "CEP57L1"

[7] "METTL6" "RP11-288I21.1" "DCTN1-AS1"

[10] "AC068831.6" "C2orf42" "METTL14"

[13] "RP1-92O14.6" "PCBP1-AS1" "C1orf50"

[16] "ANAPC10" "C1orf86" "FBXO42"

[19] "CTC-512J14.5" "DNAJC16" "RP11-159G9.5"

[22] "AC005037.3" "CCDC66" "VTI1A"

[25] "RP11-134E15.2" "ZBTB14" "RP11-340I6.6"

[28] "GANC" "ASTE1" "CREB1"

[31] "APOPT1" "C9orf85" "C10orf88"

[34] "AC007040.6" "SPATA5" "FOXO3B"

[37] "CEP128" "LYRM4" "RP11-23P13.4"

[40] "FBXL4" "GMEB1" "AC093391.2"

[43] "SCAPER" "WDR7" "RBM45"

[46] "ZNF557" "DENND4A" "C9orf156"

[49] "AC083899.3" "RP11-840I19.3" "ZNF236"

[52] "SRBD1" "AC055764.1" "ERCC4"

[55] "VPS53" "CTD-3113P16.5" "RPL23AP79"

[58] "DCUN1D3" "PARP16" "CTB-31O20.3"

[61] "CNTLN" "ZNF174" "GID4"

[64] "FAM76A" "PAPOLG" "TAMM41"

[67] "SZT2" "SEC22A" "RP11-603J24.5"

[70] "MTO1" "ZNF227" "RWDD3"

[73] "C10orf12" "ZNF35" "FARS2"

[76] "RP5-862P8.3" "GNG10" "CTD-2013N24.2"

[79] "ST7L" "ZKSCAN2" "AC018463.4"

[82] "ZNF526" "ZNF430" "NOL9"

[85] "TCAIM" "ZKSCAN4" "DET1"

[88] "SZT2-AS1" "PHF7" "RP11-352G18.2"

[91] "MYO9A" "RP11-362K14.5" "RP11-375N15.2"

[94] "S100PBP" "RP11-815I9.4" "INO80C"

[97] "RP5-1021I20.5" "ZBTB11-AS1" "RP4-816N1.6"

[100] "CNOT4"

Take as a “backup result”: a set of genes associated with efficacy

We can try to remove all the mitocondres

JUST CLINICAL DATA seem not to be enough

Try to remove genes with a lot of variance

NOTE:

"ZR7530\_BREAST" and BT549\_BREAST" were not considered

1. Moroz, Chaya, et al. "FBL blood test as a predictive marker of breast cancer in high risk women." Medical Oncology 14.1 (1997): 39-42.

   <https://pubmed.ncbi.nlm.nih.gov/9232610/> [↑](#footnote-ref-1)
2. Nguyen Van Long, Flora, et al. "Low level of Fibrillarin, a ribosome biogenesis factor, is a new independent marker of poor outcome in breast cancer." BMC cancer 22.1 (2022): 1-12. [↑](#footnote-ref-2)
3. Marcel, Virginie, et al. "p53 acts as a safeguard of translational control by regulating fibrillarin and rRNA methylation in cancer." Cancer cell 24.3 (2013): 318-330. <https://www.sciencedirect.com/science/article/pii/S1535610813003590> [↑](#footnote-ref-3)
4. Tsofack, S.P., Meunier, L., Sanchez, L. et al. Low expression of the X-linked ribosomal protein S4 in human serous epithelial ovarian cancer is associated with a poor prognosis. BMC Cancer 13, 303 (2013). <https://doi.org/10.1186/1471-2407-13-303> [↑](#footnote-ref-4)