Weekly Report

# 14/03/2022

BOUDEMIA ALA EDDINE

TEAM CHEN

# **Goal:**

* Extraction of mutational signatures in breast/bladder cancers using SigProfiler.
* Extract mutations by regions type
* Extract mutational signatures based on their regions

# **Procedure:**

* As previously, the mutation signatures are extracted using SigProfiler using the previously downloaded data from TCGA.
* Maftools was used to analyze the MAF files.
* I tried maftools to extract the signatures (very fast but bad results)
* Extract the TSS and UTRs using gtftools.
* Extract the TTS from UCSC.
* Map the mutations from the maf files to their respective regions.
* Run sigprofiler on the obtained files separately.

# **Results:**

## BRCA best signatures:

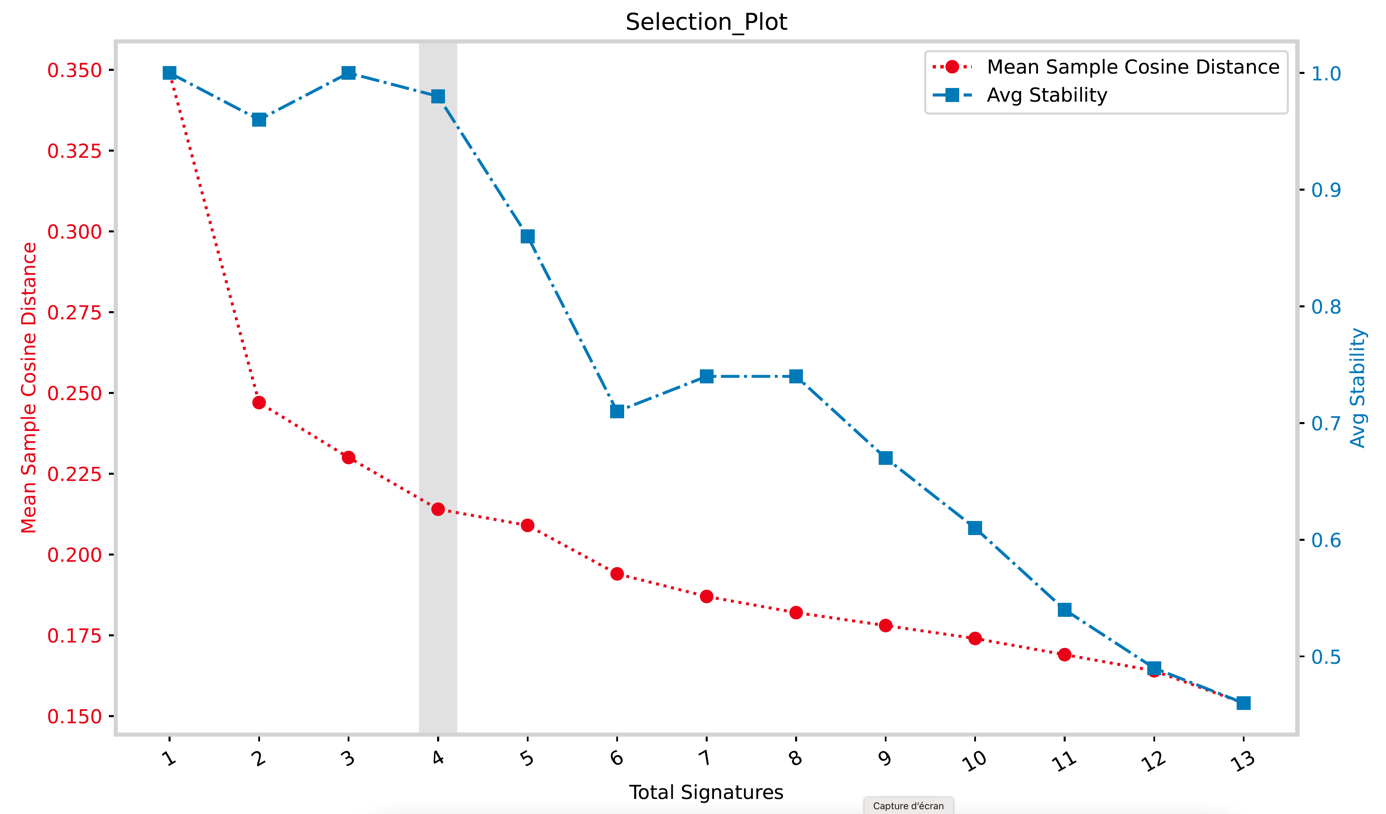


Figure 1 Selection plot of the best number of signatures BRCA

The selection plot from Figure 01 suggests that the best number of signatures to be extracted is 4. Compared to the number of mutational signatures that were reported to be associated with cancer which is 13 sigprofiler failed to deconvolute the signatures from the data at our disposal. Two possible explanations are that WES data do not show some signatures related to other genomic regions and the average mutational load is low for sigprofiler as showed by the authors of the tool. However, when we take a look at the decomposition plot obtained, one can see that most of the significant reference signatures are present, namely: SBS13, SBS2, SBS5, SBS1, and SBS3.

SBS29 as well as SBS6 and SBS10(a&b) are reported in these plots even though they were not reported to contribute to BRCA. SBS6 is associated with defective DNA mismatch repair which could be relevant but SBS29 is related to tobacco, both signatures are not contributing that much to the signatures extracted. However signatures SBS10a and SBS10b which are associated with polymerase exonuclease mutations are present as a distinct signature.

Une image contenant texte, capture d’écran, intérieur

Description générée automatiquement

Figure 2 Decomposition plot for SBSA

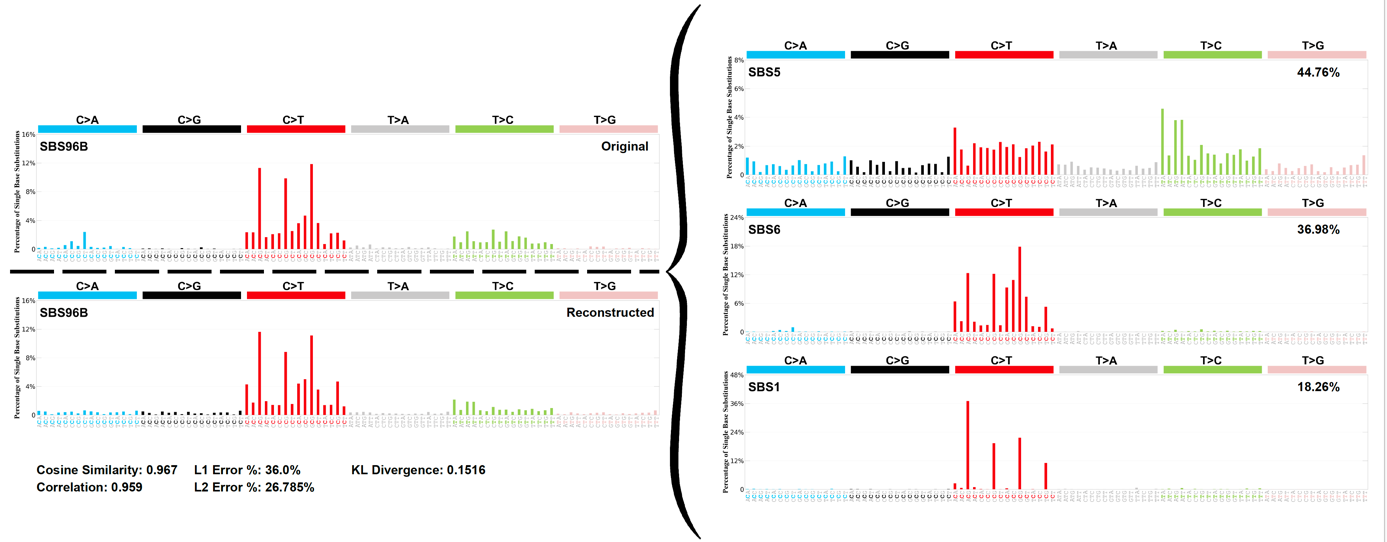


Figure 3 Decomposition plot for SBSB

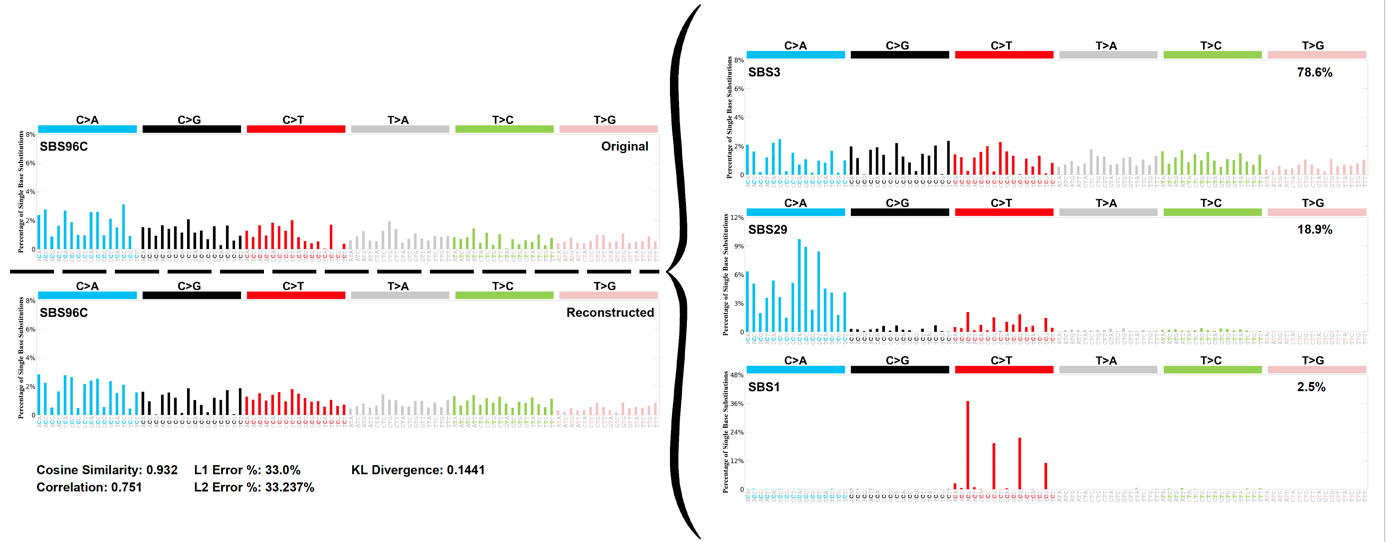


Figure 4 Decomposition plot for SBSC

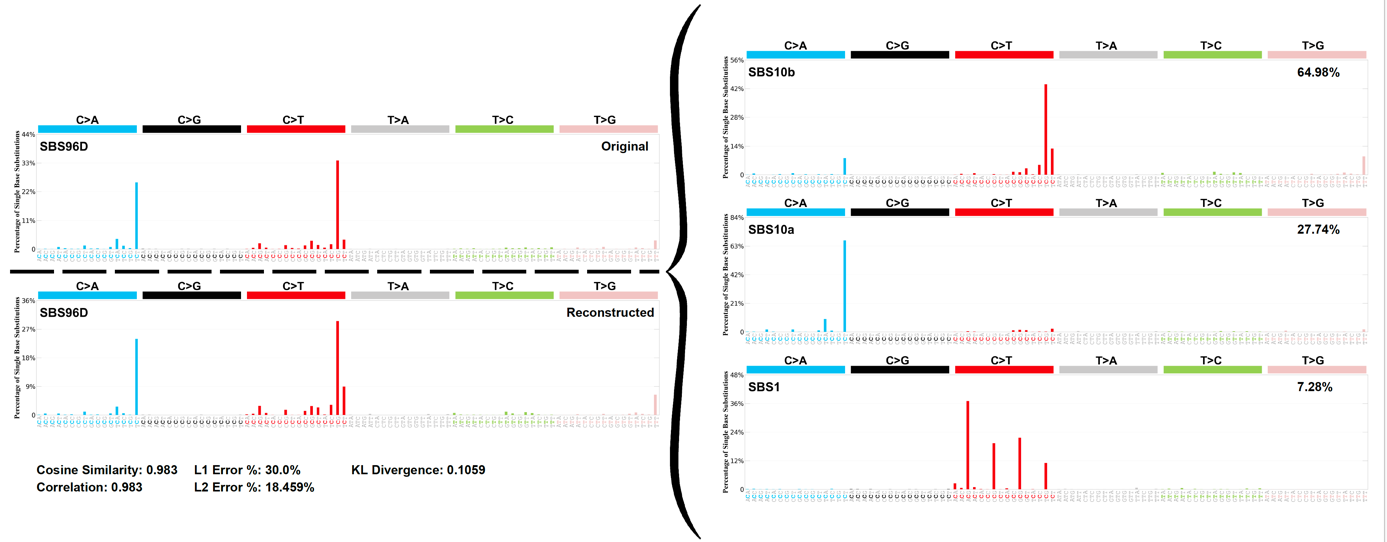


Figure Decomposition plot for SBSD

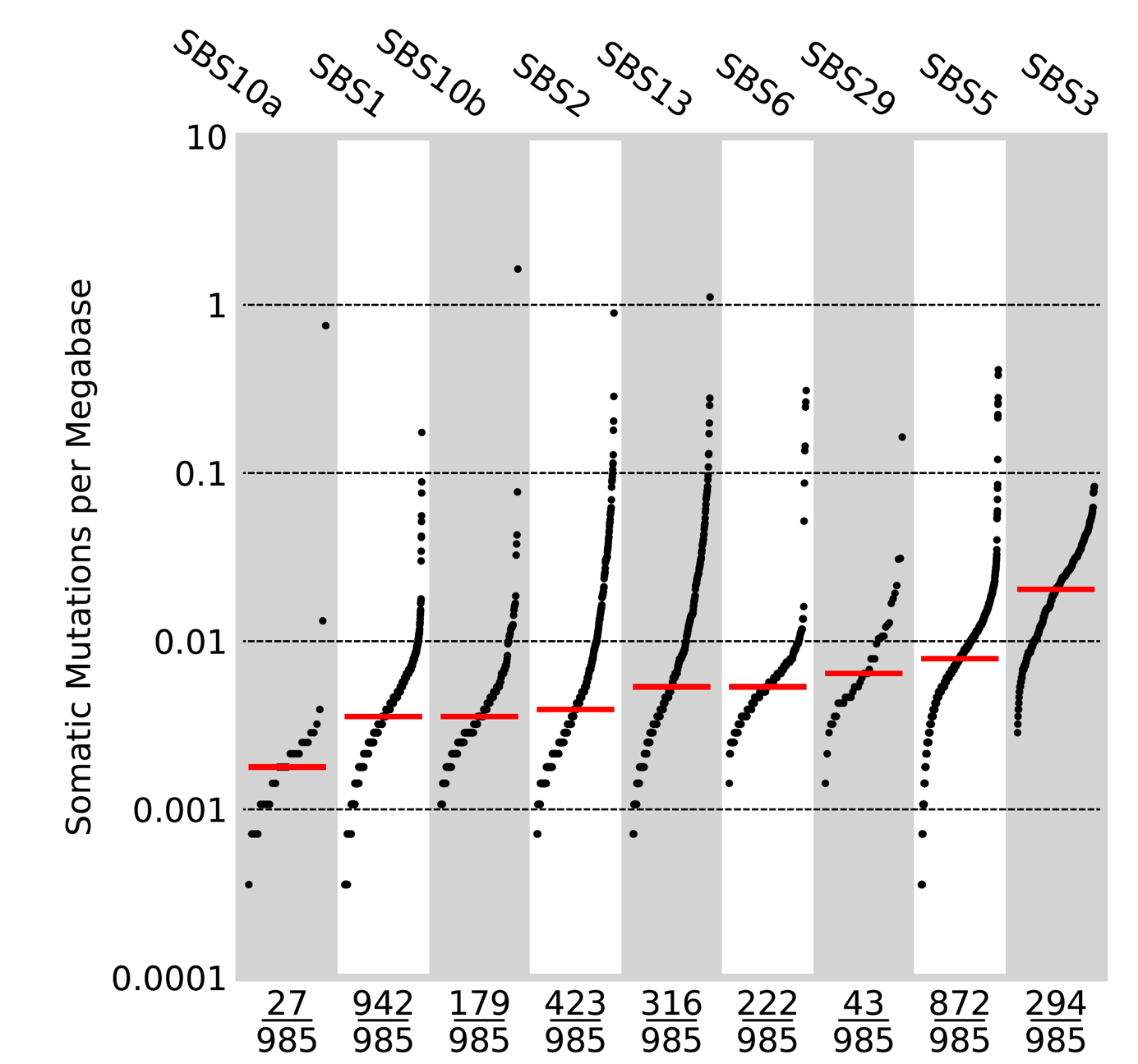


Figure 6 Activity of COSMIC signatures in the BRCA

Figure 06 shows that SBS1 is present in all samples but going back to the decomposition plots we observe that it is present in every extracted signatures mainly in moderate percentages, therefore I am not sure how to interpret this. At the same time SBS1 is ubiquitous signature in all cancer types. Similarly, SBS5 is common in all cancer types and is present most samples here.

## BLCA best signatures: