IO17 | Large Scale Bioinformatics for Immuno-Oncology

Neoantigens: exercise 4

Francesca Finotello, Federica Eduati, and Pedro L. Fernandes

INSTITUTO
GULBENKIAN
DE CIÊNCIA

To use TIminer, Docker must be running on your computer!

We have predicted a list of candidate neoantigens, i.e. mutated peptides binding to the patients HLA-A molecule: *Patient_1_NetMHCpan_binders.txt*

We also have access to normalized RNA-seq expression data from the same patient (assume the data are already in the format required by TIminer, i.e. log2(TPM+1)): *Patient_1_normalized_expression.txt*

Use TIminer (function TIminerAPI.**filterNeoantigenFile**) to filter only the peptides arising from genes with TPM > 1.

Once you get the results, answer the questions related to this exercise at:

https://b.socrative.com/login/student/

Filtering candidate neoantigens with Timiner

```
from TIminer import TIminerAPI

TIminerAPI.filterNeoantigenFile (…)
```

From TIminer documentation

http://icbi.i-med.ac.at/software/timiner/doc/index.html

TIminer.TIminerAPI.**filterNeoantigenFile**(*neoantigenInputFile*, *geneExpressionInputFile*, *filteredNeoantigenOutputFile=None*, *expressionThreshold=2*) ¶

This function considers a file of candidate neoantigens and a file of normalized gene expression levels, computed as log2(TPM+1), to select neoantigens arising from expressed genes.

Parameters:
- **neoantigenInputFile** (*str*) – Path to the `neoantigen input file` as generated by the Netmhcpan, containing binding peptides as possible neoantigen canditates (see the output files section for details).
- **geneExpressionInputFile** (*str*) – Path to the `gene expression file`: a tab-separated file containing the list of HGNC gene symbols inside the first column and their gene expression, as *log2(TPM+1)*, inside the second column. The output file of Kallisto expression.txt can be used directly as input file (see the format in the output files section).
- **filteredNeoantigenOutputFile** (*str*) – Path to the `output file` (optional, default = *inputfilename_neoantigens.txt*).
- **expressionThreshold** (*int*) – The expression threshold in TPM. Neoantigens arising from genes with an expression lower than the threshold are discarded (optional, default = *2*).

1. Access Socrative with student access at: https://b.socrative.com/login/student/
2. Join the **IO17** room

3. Enter your name and click „done" to start the quiz

I017

Enter your name

Francesca Finotello

DONE