

Statistical Analysis of the Genome Trend for SARS-1 and MERS Family with respect to the novel COVID-19 virus

Overview:

This is a small initial analysis performed on the SARS-1 and MERS gene family to understand the nature of potential antidotes for the current COVID-19 virus, from the existing FDA approved drugs, based on the connectivity scores and the p-value of the drugs to the current COVID-19 genome sequence.

Dataset:

For this purpose, the [‘Harmonizome’](#) dataset has been used, which contains more than 70 million functional associations between differentially expressed genes. Moreover, it contains data from the lungs of mice and lung cell lines infected with 2 earlier versions of Coronavirus, which are highly similar to the novel COVIS-19, coming from the gene family of SARS-1 and MERS.

Methodology and Understanding:

This dataset was used to prepare a connectivity map from all the available genome sequences. A package by the name of [PharmacoGX](#) helps in generating a genome-scale connectivity map, with the functional associations of the COVID-19 as a starting point. Each connection can be traced back to the starting point to see how good a dock it is for the COVID-19 as a solution. To quantify the *goodness* of the drug in the database with the COVID-19 virus, each connection has been quantified with a connectivity score and a p-value. The quantification terminologies have been defined below:

Connectivity: This is a score that defines the docking *goodness*. The more the negative magnitude, the more its capability to reverse the effect of the virus (in this case COVID-19).

P-value: This is basically a measure of statistical significance. Considering each connection between the drugs from the dataset and COVID-19 virus, the p-value is the probability of obtaining favourable results in an extremely opposite case. This basically means that the smaller the p-value for each connection, the stronger the chances of that drug being an antidote.

Hence, the subset of the '[Harmonizome](#)' dataset particular to the SARS-1 and the MERS family was analyzed for all possible connections to the COVID-19 virus and the results were sorted based on the connectivity score and the p-value. As a result of this exercise, the FDA approved drug, 'Valproic Acid', which is originally used for migraine headache, epilepsy and bipolar disorder, was identified as a potential treatment for the COVID-19 virus.

Credits and References:

Majority of this idea has been inspired by the work done by Ms Tinka Vidovic, a PhD student at the Mediterranean Institute of Life Sciences.

Vidovic, T. (2020). *tinkavidovic/competition*. GITHUB.

[tinkavidovic/competition](https://github.com/tinkavidovic/competition)