A target and two drugs for SARS-CoV-2 found by paralog search

Author: Jeremy B. Singer[[1]](#footnote-1)\*

# Abstract

Using a paralog search pipeline, the author searched the ChEMBL 25 database, screening targets in it against the SARS-CoV-2 genome finding a target that has an identical target sequence. The the target, RNA polymerase, was found to have 100% identity with a gene in the viral genome of SARS-CoV-2.

Three known drugs in the ChEMBL 25 database are associated with the target that was identified. Two of those drugs showed high binding affinity in docking simulations, validating them as promising drug candidates to treat SARS-CoV-2.

# Introduction

SARS-CoV-2, also known as COVID-19, is a virus that causes flu like symptoms including respiratory distress, in many cases requiring respirators to maintain oxygenation in patients. It is highly contagious, and is currently causing pandemic infection, with a fatality rate estimated between 2% and 3% [1]. Persons over 60 have may have much higher fatality rates [2].

A target repurposing strategy can provide drugs more quickly and cheaply than creating new drugs and finding new targets [3]. This strategy can produce treatments to ameliorate the disease until a vaccine becomes available, or in addition to the vaccine.

ChEMBL provides a downloadable database that includes drug targets and drug information for those targets, as well as amino acid sequences of the protein targets [4]. Drug targets tend to be proteins that are important enough to the organism to which they belong that they tend to be conserved [5].

Using a paralog search pipeline, the author searched the ChEMBL 25 database, screening targets in it against the SARS-CoV-2 genome and found a high scoring target that has three known drugs [4, 6].

The genome of SARS-CoV-2 (The COVID-19 virus) was downloaded from Genbank via NCBI’s website [7].

Using **jackhmmer** to provide similarity reports and sequence alignments, a pipeline imported scores showing sequence similarity. The target scores were loaded into a PostgreSQL database that also contains the ChEMBL data [8].

# Materials and methods

Using a paralog search pipeline, the author searched the ChEMBL 25 database, screening targets in it against the SARS-CoV-2 genome and found a high scoring target that has three known drugs [4, 6].

# Results and discussion

# Conclusions

# References

1. **Early Release - Case-Fatality Risk Estimates for COVID-19 Calculated by Using a Lag Time for Fatality - Volume 26, Number 6—June 2020 - Emerging Infectious Diseases journal - CDC**. 2020.

2. **Coronavirus Age, Sex, Demographics (COVID-19) - Worldometer**. 2020.

3. Pollastri MP, Campbell RK: **Target repurposing for neglected diseases**. *Future Med Chem* 2011, **3**(10):1307-1315.

4. Gaulton A, Hersey A, Nowotka M, Bento AP, Chambers J, Mendez D, Mutowo P, Atkinson F, Bellis LJ, Cibrián-Uhalte E *et al*: **The ChEMBL database in 2017**. *Nucleic Acids Res* 2017, **45**(Database issue):D945-954.

5. Lv W, Xu Y, Guo Y, Yu Z, Feng G, Liu P, Luan M, Zhu H, Liu G, Zhang M *et al*: **The drug target genes show higher evolutionary conservation than non-target genes**. In: *Oncotarget.* vol. 7; 2016: 4961-4971.

6. Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, Light Y, McGlinchey S, Michalovich D, Al-Lazikani B *et al*: **ChEMBL: a large-scale bioactivity database for drug discovery**. *Nucleic Acids Res* 2012, **40**(Database issue):D1100-1107.

7. **Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, co - Nucleotide - NCBI** [<https://www.ncbi.nlm.nih.gov/pubmed/>]

8. Wheeler TJ, HHMI Janelia Farm Research Campus A, VA 20147, USA, Eddy SR, HHMI Janelia Farm Research Campus A, VA 20147, USA: **nhmmer: DNA homology search with profile HMMs**. *Bioinformatics* 2013, **29**(19):2487-2489.

1. Student in Brandeis Univeristy GPS program

   \* Corresponding author

   E-mail: [jbsing@brandeis.edu](mailto:jbsing@brandeis.edu) (JBS) [↑](#footnote-ref-1)