

Supplementary Materials for

COVID-19 Vaccine Candidates: Prediction and Validation of 174 SARS-CoV-2 Epitopes

Marek Prachar^{1,2,3}, Sune Justesen³, Daniel Bisgaard Steen-Jensen³, Stephan Thorgrimsen³, Erik Jurgons⁴, Ole Winther^{1,2,5}, and Frederik Otzen Bagger^{1,6,7,*}

1. Center for Genomic Medicine, Copenhagen University Hospital, Copenhagen, Denmark
2. Bioinformatics Centre, Department of Biology, University of Copenhagen, Copenhagen, Denmark
3. Immunitrack ApS, Copenhagen, Denmark
4. INTAVIS Peptide Services GmbH & Co.KG, Waldhäuser Str. 64, 72076 Tübingen, Germany
5. Department of Applied Mathematics and Computer Science, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
6. UKBB Universitäts-Kinderspital, Department of Biomedicine, Basel, 4031 Basel, Switzerland
7. Swiss Institute of Bioinformatics, Basel, 4053 Basel, Switzerland

* Correspondence to Frederik Otzen Bagger, frederik.otzen.bagger@regionh.dk

This PDF file includes:

Materials and Methods
Table S1
Captions for Data S1-S2

Materials and Methods

Analysis of predictions

To produce the ROC graphs and to calculate their AUC we used RStudio Version 1.2.5033, R version 3.6.2, pROC 1.16.1 package. The same software was used to calculate and plot Spearman correlations, using the ggplot2 3.2.1 package.

PrdX

The software was written in Python Version 3.7.3, NN architecture was built using PyTorch Version 1.3.0.

Table S1. List of reference peptides used for their corresponding allotype .

| Allotype | Sequence | Reference |
|------------------|---------------|---|
| <i>A*0101</i> | VTEHDTLLY | http://www.iedb.org/epitope/71290 |
| <i>A*0201</i> | VLDFAPPGA | Wilms tumor antigen 1 |
| <i>A*0301</i> | AVAHKVHLMYK | https://www.iedb.org/epitope/419554 |
| <i>A*1101</i> | AVFDRKSDAK | https://www.iedb.org/epitope/5316 |
| <i>A*2402</i> | AYAQQIFKIL | https://www.iedb.org/epitope/5731 |
| <i>B*4001</i> | REDQWCGSL | https://www.iedb.org/epitope/53476 |
| <i>C*0102</i> | QYDPVAALF | http://www.iedb.org/epitope/52886 |
| <i>C*0401</i> | QYDPVAALF | http://www.iedb.org/epitope/52886 |
| <i>C*0701</i> | YLHARLREL | Identified by previous screening |
| <i>C*0702</i> | NYFNRMFHF | Identified by previous screening |
| <i>DRB1*0401</i> | AKFVAAWTLKAAA | https://www.iedb.org/epitope/2192 |

Data S1. COVID19-Intavis-Immunitrack-dataset

<https://www.immunitrack.com/wp/wp-content/uploads/Covid19-Intavis-Immunitrack-datasetV2.xlsx>

Excel file containing results of the NeoScreen[®] assay.

Data S2. Full list of predicted and measured peptides used in the benchmark

<https://doi.org/10.5281/zenodo.3739310>

Public datasets of all peptides that were predicted and subsequently measured. List of peptide sequences for each respective allotype. All predictions of the benchmarked tools, not all tested tools use the same measure.