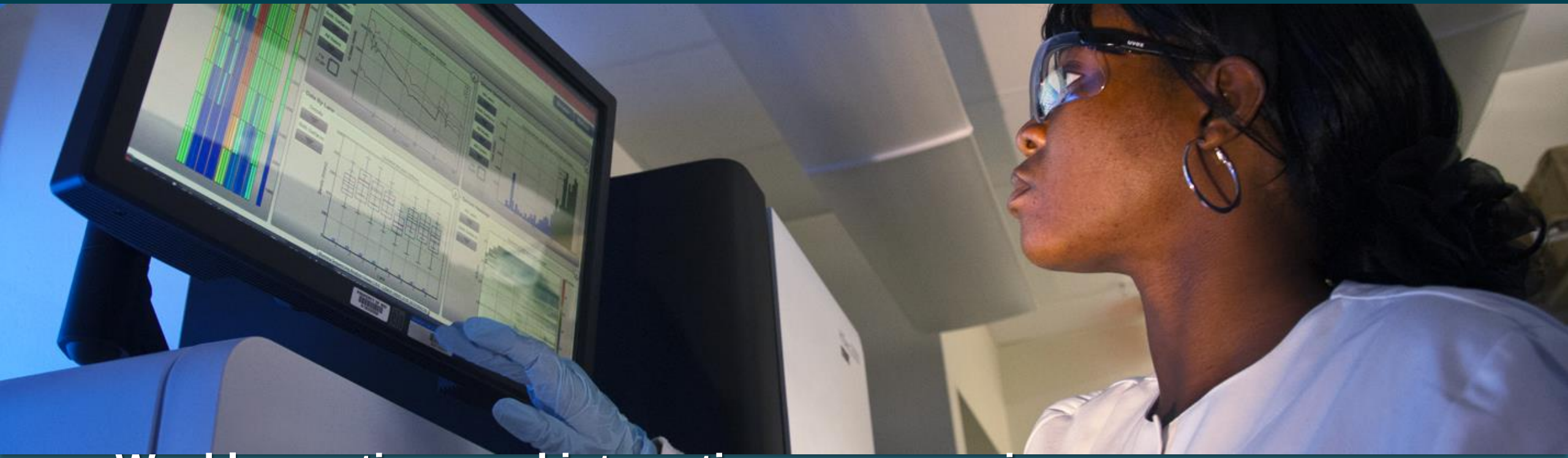
Frederick National Laboratory
for Cancer Research

sponsored by the National Cancer Institute



Weekly meetings and interesting papers review

S. Ravichandran, BIDS
FNLCR

May 15, 2020

Interesting papers

Beyond the Spike: identification of viral targets of the antibody responses to 2 SARS-CoV-2 in COVID-19 patients

What receptor to target?

Conclusion: Our report provides an unbiased characterization of antibody responses to a range of SARS-CoV-2 antigens. The combination of 3 SARS-CoV-2 antibody LIPS assays, i.e. N, ORF3b, and ORF8, is sufficient to identify all COVID-19 patients of our cohort even at early time-points of illness, whilst Spike alone fails to do so. Furthermore, our study highlights the importance of investigating new immunogens NSP1, ORF3b, ORF7a and ORF8 which may mediate immune functions other than neutralization which may be beneficial or harmful to the patient.

PubChem (COVID-19 related data)

- **Compounds**
  - used in SARS-CoV-2 clinical trials; found in COVID-19 PDB structures; Other evidences like Nature, Drugbank, Clinical trials etc.
- **Bioassays**
  - PubChem bioassays with title containing the keyword 'SARS'
- **Genes**
  - Genes for the SARS-CoV-2 virus and SARS-CoV virus and human target genes

PubChem (COVID-19 related data)

Frederick National Laboratory for Cancer Research
sponsored by the National Cancer Institute

- **Compounds (327)**
  - What do we have?
    - ID → SMILES
- **Proteins (71)**
  - Human (37), Severe acute respiratory syndrome coronavirus 2 (14); SARS Coronovirus (14)
- **Genes (246)**
  - human: 231; Severe acute respiratory syndrome coronavirus 2: 11, others: rest)
- **Bioassay (217)**
  - Latest data ; 2018-10-26; Source: ChEMBL (197); Scripps Research Inst (17) and NCATS (3)
  - What receptor?
  - (Protease; 137 entries; Papain-like protease 14 entries; Nsp13; 6 entries)

AMPL

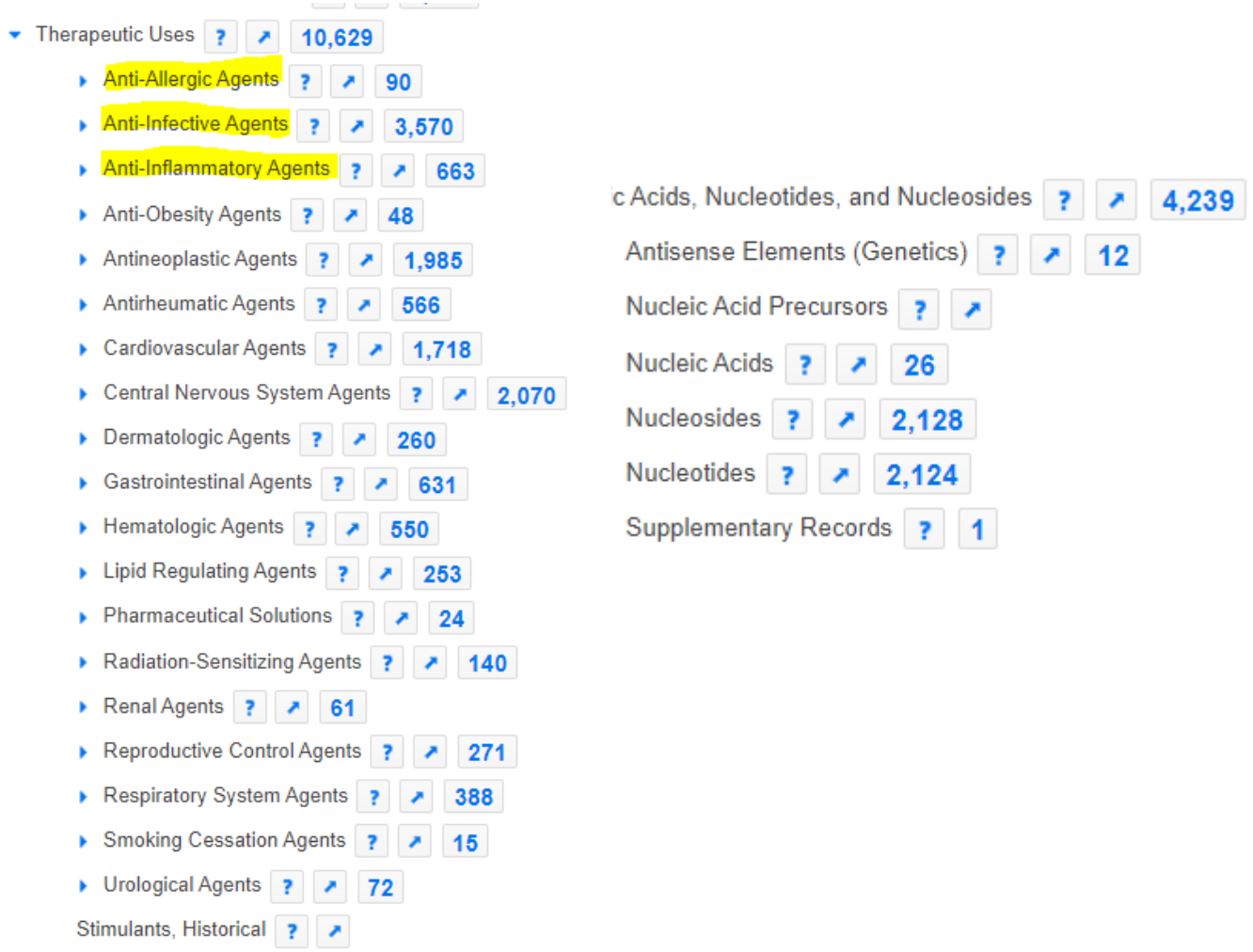
PubChem AID:	1706
Protein Target:	3C-like protease [Avian infectious bronchitis virus]
Source:	The Scripps Research Institute Molecular Screening Center
External ID:	3CLPRO_INH_QFRET_1536_%INH
BioAssay Type:	Primary Screening
Tested Substances:	All (290,893) Active (405) Inactive (290,488) Data Table
Tested Compounds:	All (290,726) Active (405) Inactive (290,321)
Version:	1.2 Revision History
Status:	Live
Dates:	Modify: 2010-06-15 Deposit: 2009-05-01

Take the compounds
Convert them into 2 class
And do classification

UQ

We can do this now

AMPL



Classification experiments

A set of classification model experiments were also conducted for a panel of 28 bioactivity datasets, without any hyperparameter tuning. In total 2,130 neural network and random forest models were generated. A dose concentration threshold was used to label active and inactive compounds on a per-dataset basis using thresholds provided by domain experts at GlaxoSmithKline. The classes were extremely unbalanced, which partially explains the high ROC-AUC scores.

Project Idea: Receptor based modeling and choosing the right targets

- **Spike Protein**
  - Hyeryun Choe and Michael Farzan at the Scripps Research Institute immunized rats with Spike protein RBD and saw rats created antibodies that compromised the virus
  - http://doi.org/ggrs5t

- **M-Pro**
  - Z. Jin *et al*. *Nature* https://doi.org/10.1038/s41586-020-2223-y; 2020
  - Identified compounds that inhibit the protease
  - Antiviral assay and HTS (Experimental)
  - Identified power inhibitors like ebselen
    - safety has already been tested (??) in people

Coronavirus SARS-Cov-2

- **Important Genes and Proteins**
- **Virus entry**
  - Attachment and entry to Type-II Pneumocyte
    - **<u>Key interactions</u>**
    - Spike Protein (a.k.a S-Protein, S)
      - Interaction with host ACE2 receptor
      - Interaction with host serine Protease TMPRSS2 cleaves viral Spike protein and allows for interaction with host cell-membrane
  - Endocytocis → Fusion → Translation -→ nsp poly-proteins → Proteolysis → Replication → Assembly → Virion release