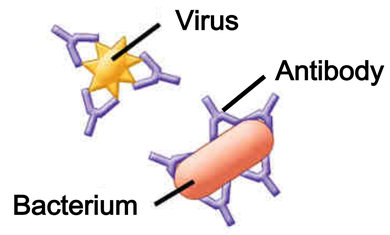
Mixture of Experts for Predicting Antibody-Antigen Binding Affinity from Antigen Sequence

Srivamshi Pittala, Chris Bailey-Kellogg
Department of Computer Science, Dartmouth College, Hanover, NH, USA

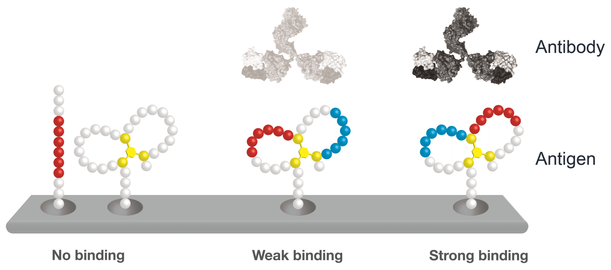
Overview

- ❖ Antibodies provide a key mode of defense employed by the immune system to fight disease causing pathogens
- ❖ Knowledge of antibody-antigen binding affinity has potential to advance vaccine development and antibody-based therapeutics
- ❖ Experimental techniques to determine antibody-antigen binding affinity are difficult to scale-up to large sets of antibodies and antigens
- ❖ We developed a mixture of experts approach for predicting the binding affinity of an antibody against an antigen based on the antigen's sequence
- ❖ Evaluated on a dataset of 52 antibodies and 608 strains of HIV: the predictive accuracy is significantly better than that of individual models

Antibody-Antigen Binding



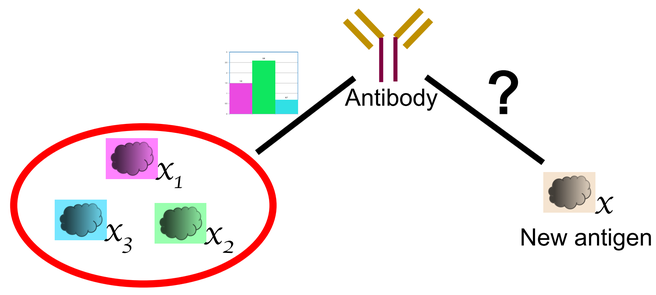
- Antibodies bind to the antigens present on the pathogens, helping the immune system to fight off disease
- Vaccine Design: Binding affinity measurements can be used to identify parts of the antigen which can be used in vaccines
- Antibody-based Therapeutics: Knowledge of antibody-antigen binding affinity can be used to design potent antibodies for therapeutics



- Challenge: Antibody-antigen binding is driven by several chemical and physical factors of both proteins
- Experimental methods are resource intensive and time consuming to scale-up to newly discovered antibodies and antigens, or mutating pathogens
- Learning-based computational methods, though scalable, are not accurate enough for experimental use

Aim: Antibody-specific binding affinity prediction

For an antibody, given its binding affinities to a panel of known antigens, predict its binding affinity to a new antigen based on that antigen's sequence

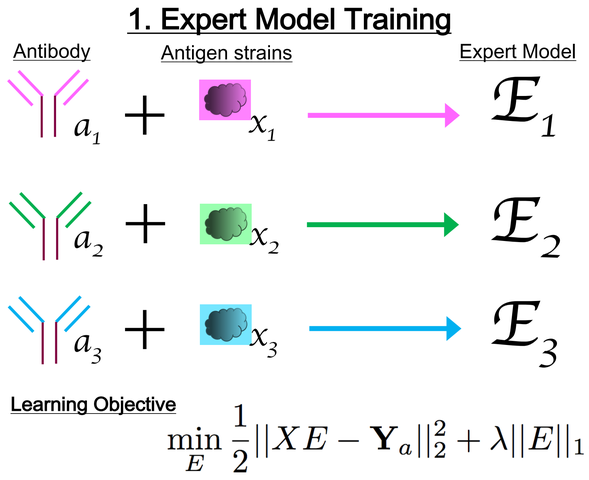


Source : pepscan.com

Method

Antibody-specific binding affinity prediction: 1. Train antibody expert models on antigen sequence [3]
2. Train antibody-specific mixture models on experts' outputs

1. Expert Model Training

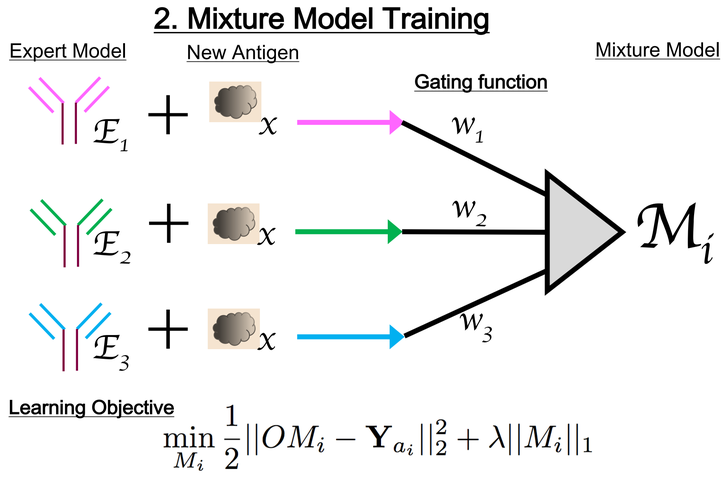


Learning Objective

$$\min_E \frac{1}{2}||XE - \mathbf{Y}_a||_2^2 + \lambda||E||_1$$

X : antigen sequences Y : binding affinity E : model coefficients

2. Mixture Model Training



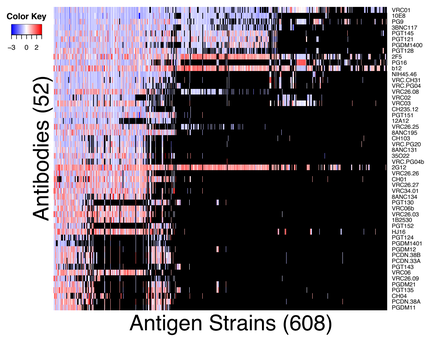
Learning Objective

$$\min_{M_i} \frac{1}{2}||OM_i - \mathbf{Y}_{a_i}||_2^2 + \lambda||M_i||_1$$

O : expert model prediction on antigen x Y : binding affinity M : model coefficients

Dataset

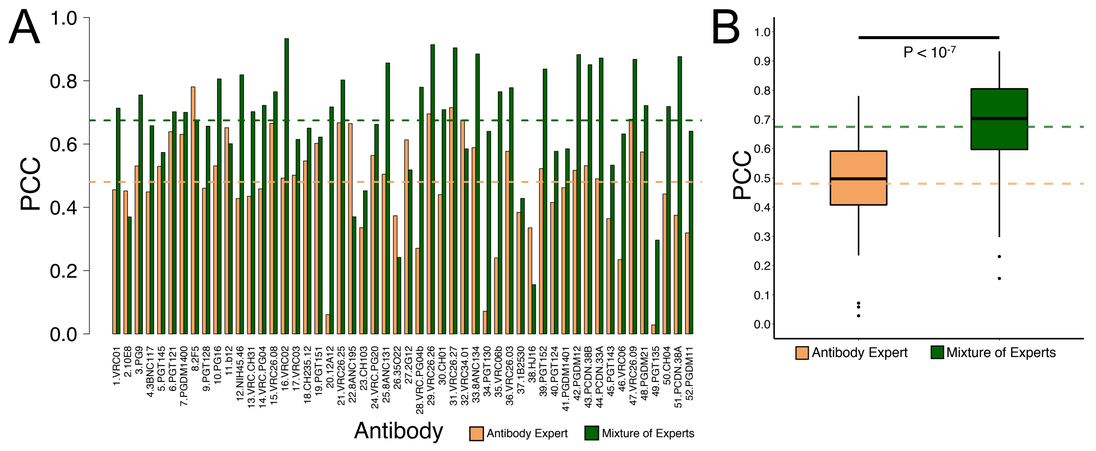
- CATNAP HIV neutralization dataset
- Sub-selected 52 antibodies which have at least 45 antigen strains in their panel
- 65% of the dataset is empty (not determined)

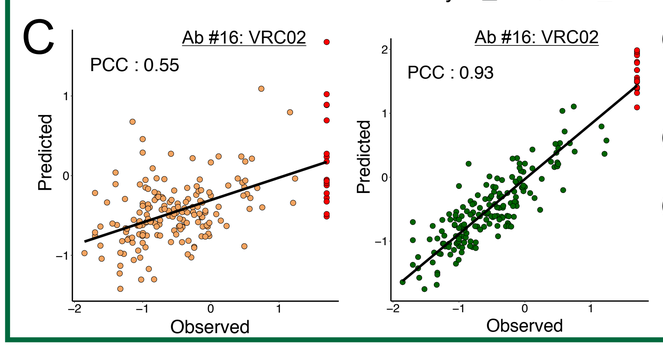


- Evaluation via 10-fold cross-validation; metric Pearson Correlation Coefficient (PCC) between observed and predicted binding affinity

Results

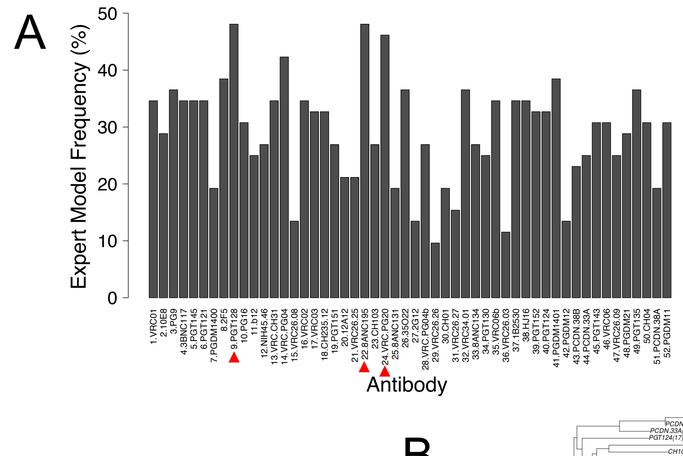
Mixture Models Yield Significant Performance Gains Over Individual Expert Models

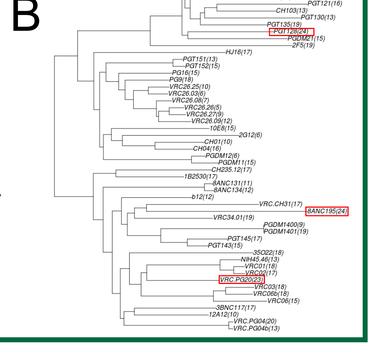




(A) Per-antibody PCC between the observed and predicted binding affinities of the expert (yellow) and mixture (green) models.
(B) Comparison of performance across 52 antibodies (Wilcoxon-Mann-Whitney test).
(C) Scatterplots of observed and predicted binding affinities from expert (left) and mixture (right) models for antibody VRC02.

Mixture Models Are Not Dominated by a Few Expert Models





(A) Per-antibody frequency of appearance of the antibody's expert model in 52 mixture models. The top three most frequent antibodies are highlighted in red.
(B) Antibodies clustered based on their heavy chain protein sequences.

Future Directions

- Establishing the biological relevance of the trained expert and mixture models
- Exploring the impact of adding physicochemical properties of amino acids to the feature sets
- Considering structural elements of antibodies as features to predict interactions

References

1. Charles A. Janeway. How the immune system works to protect the host from infection: A personal view. *Proceedings of the National Academy of Sciences*, 98(13):7461–7468, 2001.
2. Florian Klein, Hugo Mouquet, Pia Dosenovic, Johannes F. Scheid, Louise Scharf, and Michel C. Nussenzweig. Antibodies in hiv-1 vaccine development and therapy. *Science*, 341(6151):1199– 1204, 2013.
3. Andrew L. Ferguson, Emilia Falkowska, Laura M. Walker, Michael S. Seaman, Dennis R. Burton, and Arup K. Chakraborty. Computational prediction of broadly neutralizing hiv-1 antibody epitopes from neutralization activity data. *PLOS ONE*, 8(12):1–16, 12 2013.
4. Zhao Kang, Chong Peng, and Qiang Cheng. Top-n recommender system via matrix completion. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, AAAI'16, pages 179–184. AAAI Press, 2016.
5. Hyejin Yoon, Jennifer Macke, Anthony P. West, Jr, Brian Foley, Pamela J. Bjorkman, Bette Korber, and Karina Yusim. Catnap: a tool to compile, analyze and tally neutralizing antibody panels. *Nucleic Acids Research*, 43(W1):W213–W219, 2015.