University of Cologne | University Hospital Cologne | Institute of Virology | Laboratory for Viral Resistance Research

Geno2pheno-bNAbs: Interpretable and accurate prediction of HIV bNAb resistance

Martin Pirkl[1,2], Michael Böhm[1,2], Joachim Büch[1,2], Stanley Odidika[2,4,5], Philipp Schommers[2,4,5], Florian Klein[1,2,5], Rolf Kaiser[1,2,6], Thomas Lengauer[1,2,3]

[1]Institute of Virology, Faculty of Medicine and University Hospital Cologne, University of Cologne, [2]German Center for Infection Research (DZIF), Partner Site Cologne-Bonn, [3]Max Planck Institute for Informatics, Saarland Informatics Campus, [4]Department of Internal Medicine, Faculty of Medicine and University Hospital Cologne, University of Cologne, [5]Center for Molecular Medicine Cologne, University of Cologne, [6]EuResist Network GEIE

Background

HIV ist still a serious issue with many deaths, especially in countries with poor access to therapeutic options. Antiretroviral therapy (ART) is a useful option for people living with HIV (PLWH) and effective against many viral strains. However, most ARTs are still a relatively large inconvenience for PLWH. The most common ARTs involve combinations of drugs targeting viral or cellular proteins. Most of these drugs have to be taken every day. An alternative to ARTs with protein inhibitors are broadly neutralizing antibodies (bNAbs, Figure 1). These have been shown to yield long-term viral suppression and have to be administered only once every several weeks or even months. However, bNAbs share the problem of viral resistance with protein inhibitors.

Solutions that are accurate, interpretable and easy to use are necessary to tackle the problem of predicting viral resistance to bNAbs. We developed a web-service g2p-bNAbs that allows users to upload one or more viral genotypes and our service uses trained models to predict resistance to many common bNAbs. The predictions consist of classification (sensitive or resistant) and the IC50 score.

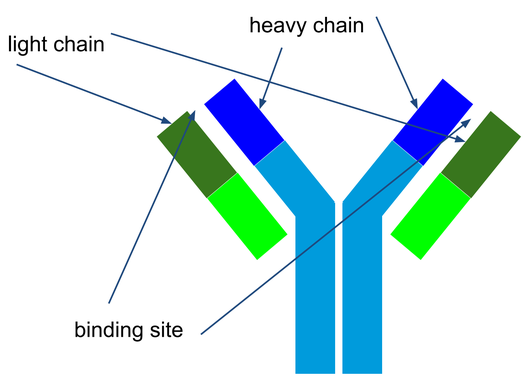

Figure 1: Schematic of an antibody with heavy chain, light chain and binding site highlighted.

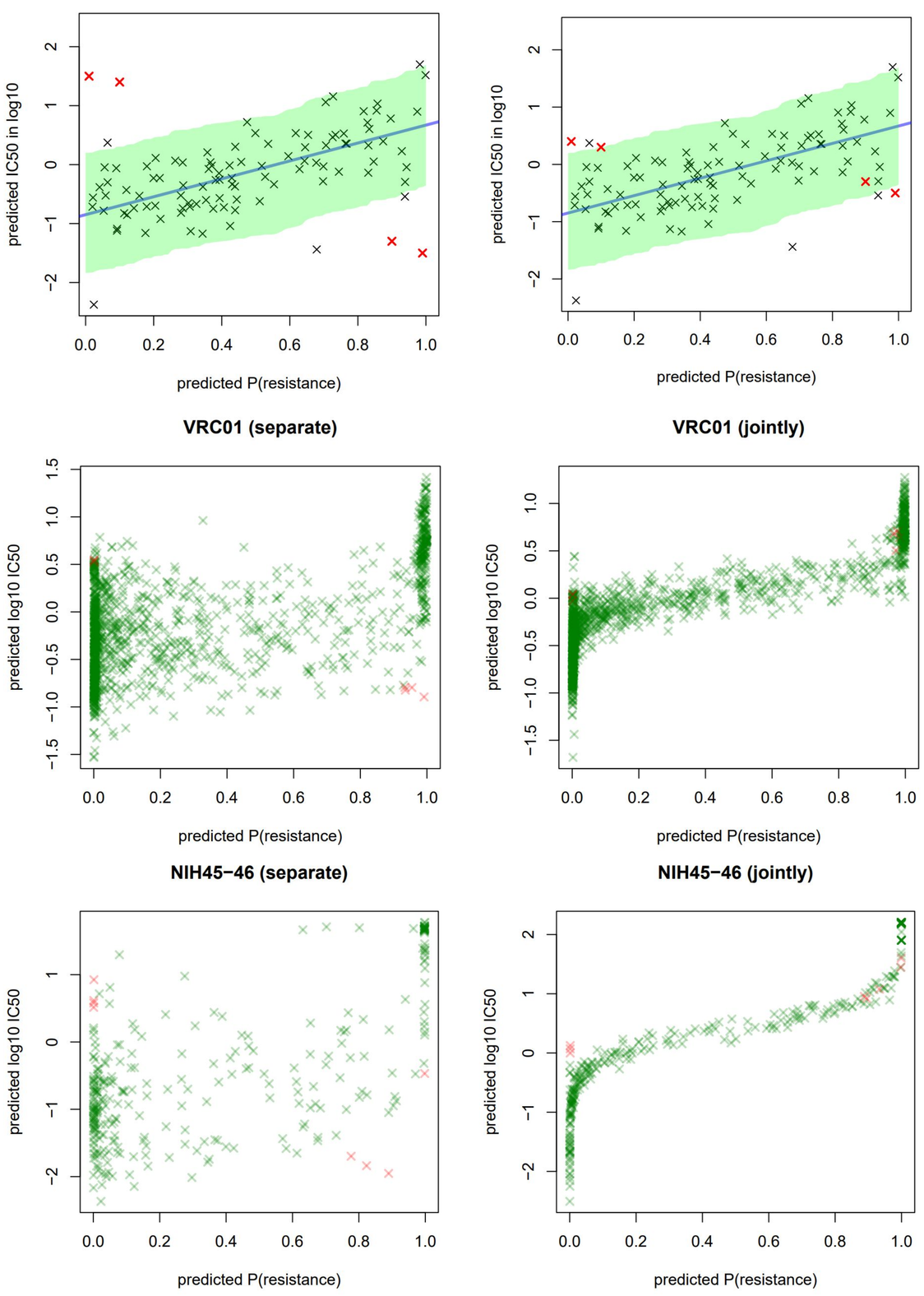

Figure 2: Regularization of discordant outliers with multitask learning (jointly).

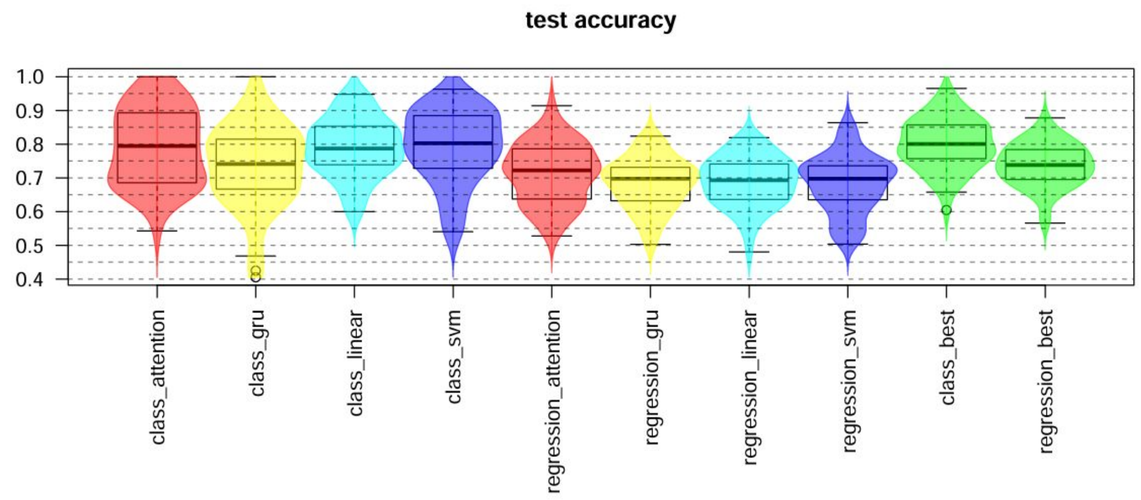

Figure 3: Accuracy of classification (area under the receiver operating characteristic) and regression (Harrell's concordance index[b]) of the different models over 50 bNAbs.

Methods

We used logistic and linear regression as our models for each task. We used multi-task learning to train both models simultaneously. I.e., during training we aim to decrease the sum of the cross-entropy (misclassification), the mean-squared-error (IC50 prediction error) and the negative covariance of class probability and IC50 prediction. The negative covariance penalizes models, which predict a high probability of resistance and a low IC50, and vice versa, for the same sample (Figure 2). In addition this can be viewed as a regularization to prevent overfitting to the training data. The training and test data was downloaded from the CATNAP database[a].

Results

We compared the models of g2p-bNAbs to other state-of-the-art methods like support-vector machines (svm), multihead-attention (attention) and recurrent neural networks with grus (gru), and found them to be competitive in regards to accuracy and have the benefit of being easily interpretable in regards to features, i.e., positions on the envelope (Figure 3).

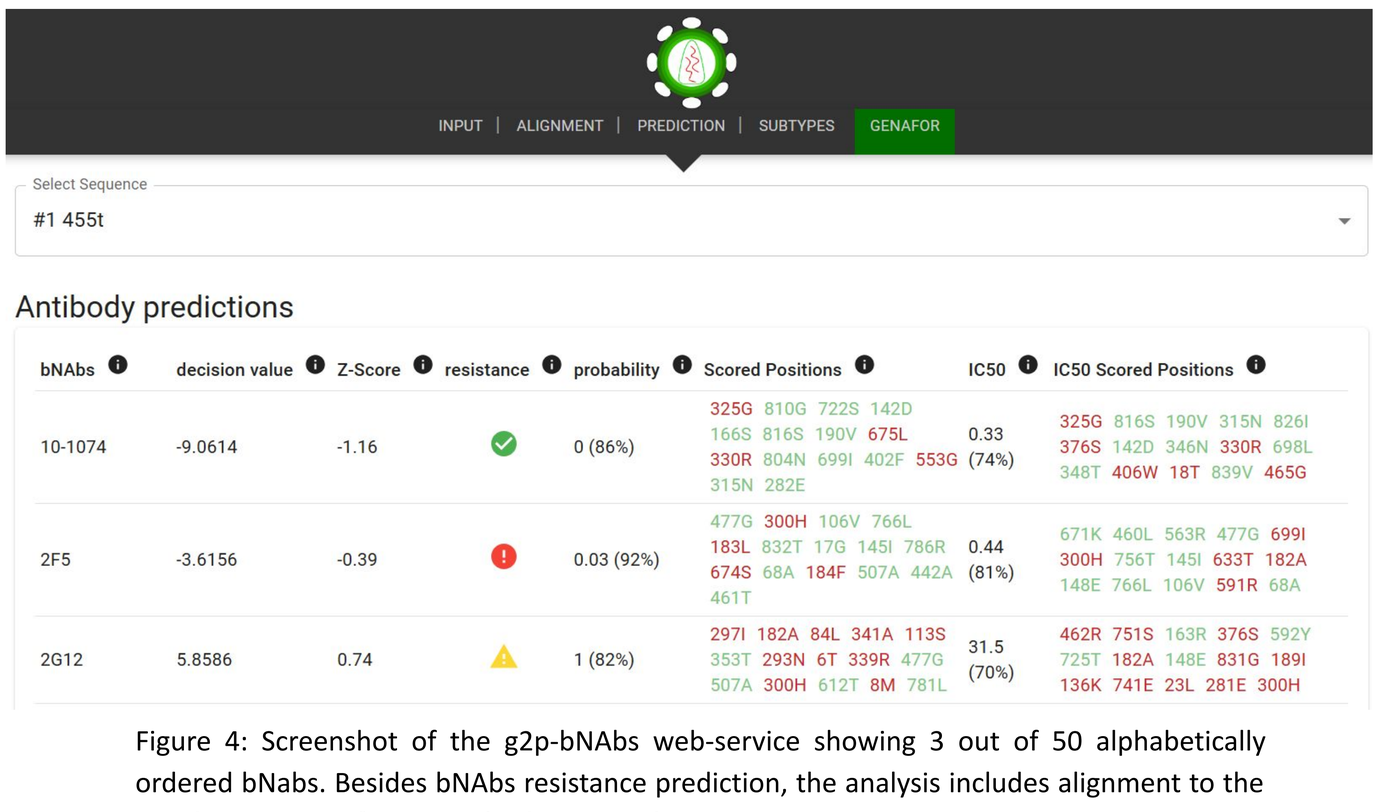

Figure 4: Screenshot of the g2p-bNAbs web-service showing 3 out of 50 alphabetically ordered bNabs. Besides bNAbs resistance prediction, the analysis includes alignment to the envelope region of HIV-1 and subtype prediction.

Conclusion

We developed a web-service for antibody resistance (g2p-bNAbs, Figure 4), which is free to use and can be extended to other viruses, like Sars-Cov2, in the future.

Contact

Martin Pirkl
Institute of Virology
University Hospital Cologne
Fürst-Pückler Str. 56, 50935 Köln
+49 221 478 85809
martin.pirkl@uk-koeln.de

References

a. Yoon H, Macke J, West AP Jr, Foley B, Bjorkman PJ, Korber B, Yusim K. CATNAP: a tool to compile, analyze and tally neutralizing antibody panels. Nucleic Acids Res. 2015 Jul 1;43(W1):W213-9. doi: 10.1093/nar/gkv404. Epub 2015 Jun 4. PMID: 26044712; PMCID: PMC4489231.

b. Harrell FE, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the Yield of Medical Tests. *JAMA.* 1982;247(18):2543–2546. doi:10.1001/jama.1982.03320430047030