www.geno2pheno.org

MAX-PLANCK-GESELLSCHAFT

www.euresist.org

97

The Eu*Resist* approach for predicting response to anti HIV-1 therapy

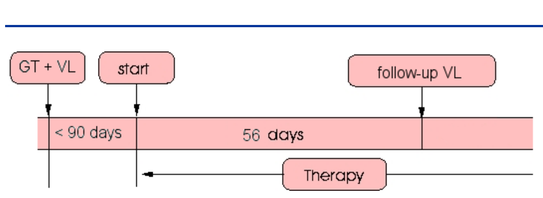
Andre Altmann[1], Michal Rosen-Zvi[2], Mattia Prosperi[3], Ehud Aharoni[2], Hani Neuvirth[2], Eugen Schülter[4], Joachim Büch[1], Yardena Peres[2], Francesca Incardona[5], Anders Sönnerborg[6], Rolf Kaiser[4], Maurizio Zazzi[7], Thomas Lengauer[1]

1 Max Planck Institute for Informatics, Saarbrücken, Germany; 2 IBM Research Laboratory in Haifa, Haifa, Israel;
3 University of Roma TRE, Rome, Italy; 4 University of Cologne, Cologne, Germany; 5 Informa srl, Rome, Italy;
6 Department of Medicine, Karolinksa Institute, Stockholm, Sweden; 7 University of Siena, Siena, Italy

Background

The aims of the Eu*Resist* project are integration of clinical and virologic data from a large cohort of patients and training of a data-driven therapy response prediction system for guiding treatment selection using these data. The Eu*Resist* Integrated Database (EIDB) currently comprises 18,467 patients, 22,006 sequences, 64,864 therapies, and 240,795 viral load (VL) measurements collected from Italy, Sweden, Germany, and Luxembourg.

Methods

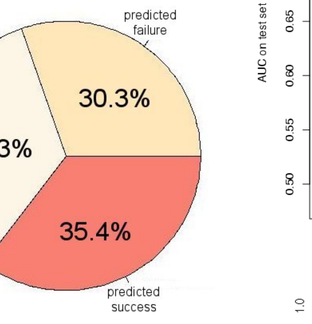


Therapy success (failure) is defined as a VL measurement below (above) 500 cp/ml at 8 [4-12] weeks after therapy start. Baseline VL and genotype must be obtained at most three months before therapy start to allow for inclusion in the training set. Using this Standard Datum definition we extracted 3,023 samples for training three different prediction engines. The engines were designed to work with viral genotype and intended regimen as the only information (minimal feature set) as well as with additional measurements (e.g. baseline VL) and information derived from previous treatments and genotypes (maximal feature set) to enhance prediction performance. All engines use logistic regression as the statistical learning method. Moreover, for achieving a more accurate and robust performance a consensus prediction is computed by taking the mean of the individual predictions. The engines are evaluated on a randomly selected validation set comprising 301 instances that were not used during training.

Generative Discriminative (GD) Engine

A Bayesian network trained on about 20,000 treatments is used to predict the probability of success for a given drug combination. The network is organized in the following way:

The middle layer uses either indicators for drug classes (min) or the number of previously used drugs from that class (max).

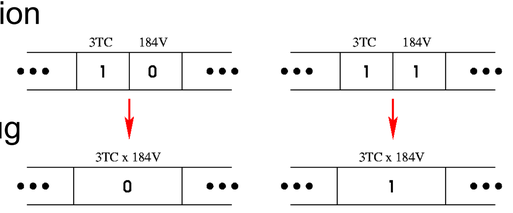


	min	max
Bayesian network prediction	yes	yes
indicators for drugs	yes	yes
indicators for mutations	yes	yes
mutations in past genotypes	no	yes
number of past treatments	no	yes
baseline VL	no	yes

Mixed Effects (ME) Engine

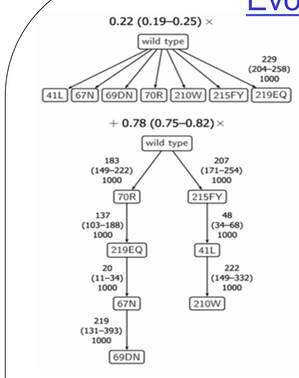
This engine explores the benefit of including second- and third-order variable interactions:

- mutation x mutation
- drug x drug
- drug x drug x drug
- drug x mutation
- drug x drug x mutation



Further features are baseline VL, number of drugs in treatment, number of past treatments, vertical transmission, and NRTI experience. In contrast to the other engines the model is only built for the maximal feature set. Missing values are replaced by standard values.

Evolutionary (EV) Engine



The genetic barrier to drug resistance for every compound in the regimen is computed and used as a feature. Here the genetic barrier is the probability of the virus not to escape from drug pressure by developing further mutations. Moreover, the maximal feature set comprises indicators for use of a drug in a regimen and previous use of a drug.

	min	max
Genetic barrier	yes	yes
indicators for drugs	yes	yes
indicators for mutations	yes	yes
drug x drug	yes	yes
mutation x mutation	yes	yes
drug x mutation	yes	yes
indicators for previous drugs	no	yes
drug x previous drug	no	yes
baseline VL	no	yes

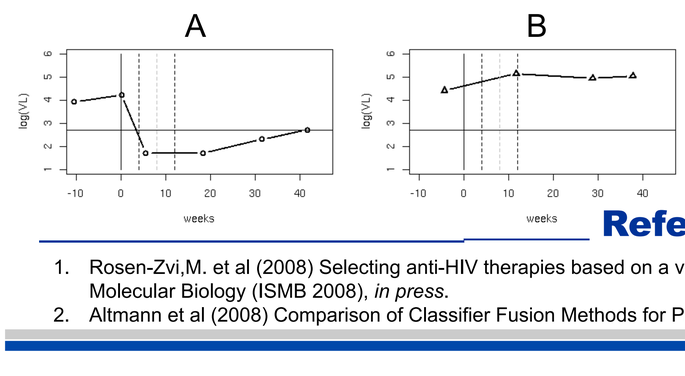
Results

Agreement among prediction engines

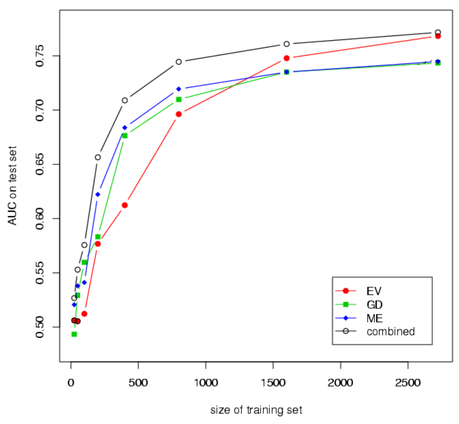
In over 80% of cases labeled as *success* the engines agree on the correct label, while this occurs only in 30% of cases labeled as *failure*:

Agreement on 202 successful (left) and 99 failing (right) treatments of the validation set

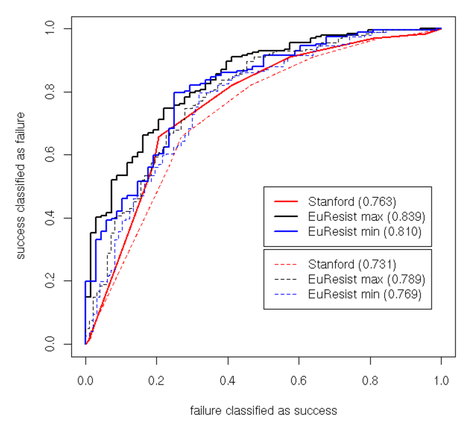
An analysis of the 35 failing treatments that were incorrectly predicted by all engines revealed that 16 cases had a VL measure below 500 cp/ml once during the course of the therapy (VL trajectories C and D). In the remaining 64 cases this occurred only 13 times (p=0.011, Fisher's Exact). Thus the combination of the engines could foresee the trend of the treatment. Remaining cases are likely to be explained with adherence problems.



Performance of the combined EuResist system



The combination of the three engines by using their mean results in a more robust engine. The mean combiner learns faster and significantly outperforms the single best engine with a training set size of 200 samples up to a training set size of 1600 samples.



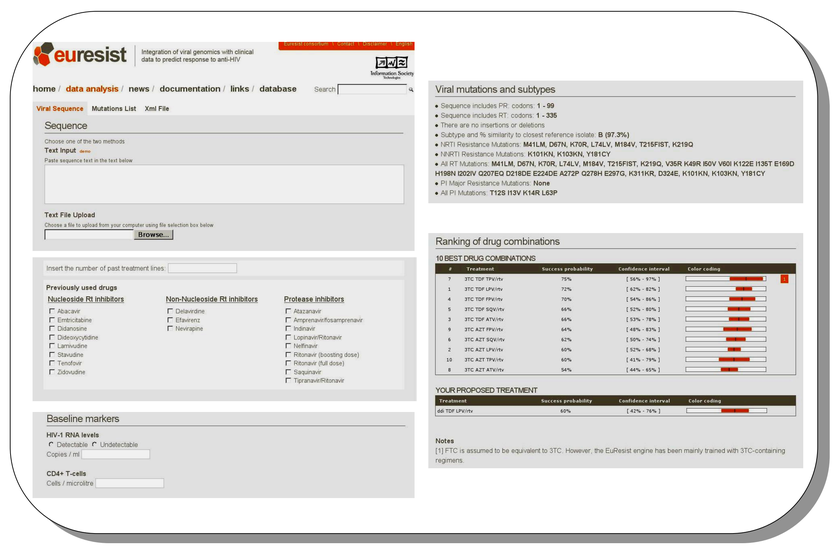
Using the maximal feature set consistently improves over the minimal feature set. The improvement is significant (p=0.048; paired t-test).

The combined Eu*Resist* engine (using either the minimal or maximal features set) outperforms a prediction based on genotypic susceptibility scores (GSS) computed using the Stanford system. After removal of ambiguous cases (VL trajectories C and D) the performance increases for all three systems: Stanford 0.763; EuResist max 0.839; EuResist min 0.810.

Remarkably, the increase is less pronounced for the GSS based prediction.

Conclusions

- Additional features like information on past treatments and viral load significantly improve the performance.
- Combination of the three engines yields a further increase in performance and robustness of the system.
- A shortcoming in the current definition of success and failure could be detected.
- Despite that shortcoming the combined engine could detect the trend correctly. However, the performance measure is biased and underestimates the true performance.
- The Eu*Resist* web service will be freely available at end of June 2008: http://www.euresist.org



References

1. Rosen-Zvi,M. et al (2008) Selecting anti-HIV therapies based on a variety of genomic and clinical factors, The 16th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB 2008), *in press*.
2. Altmann et al (2008) Comparison of Classifier Fusion Methods for Predicting Response to Anti HIV-1 Therapy. 7th European Conference on Computational Biology, *submitted*.

Acknowledgements

The work was supported by the Eu*Resist* project (IST -4- 027173-STP).