

Poster Sessions – Abstract P211

Improved therapy-success prediction with GSS estimated from clinical HIV-1 sequences

Pironti, Alejandro¹; Pfeifer, Nico¹; Kaiser, Rolf²; Walter, Hauke³ and Lengauer, Thomas¹

¹Computational Biology and Applied Algorithmics, Max Planck Institute for Informatics, Saarbrücken, Germany. ²Department of Clinical Virology, University of Cologne, Cologne, Germany. ³Diagnostics, Medizinisches Infektiologiezentrum Berlin, Berlin, Germany.

Introduction: Rules-based HIV-1 drug-resistance interpretation (DRI) systems disregard many amino-acid positions of the drug’s target protein. The aims of this study are (1) the development of a drug-resistance interpretation system that is based on HIV-1 sequences from clinical practice rather than hard-to-get phenotypes, and (2) the assessment of the benefit of taking all available amino-acid positions into account for DRI.

Materials and Methods: A dataset containing 34,934 therapy-naïve and 30,520 drug-exposed HIV-1 pol sequences with treatment history was extracted from the EuResist database and the Los Alamos National Laboratory database. 2,550 therapy-change-episode baseline sequences (TCEB) were assigned to test set A. Test set B contains 1,084 TCEB from the HIVdb TCEB repository. Sequences from patients absent in the test sets were used to train three linear support vector machines to produce

Table 1. Performance comparison for therapy-success prediction in test set A, quantified via area under the receiver operating characteristic curve (AUC)

	DE _{full}	DE _{onlyIAS}	DE _{noIAS}	HIVdb	ANRS	GRADE	REGA	geno2pheno _[resistance]
All first-line therapies	0.54	0.52	0.6	0.51	0.51	0.5	0.52	0.53
First-line therapies with TDR	0.64	0.58	0.69	0.48	0.54	0.46	0.53	0.5
First-line therapies without TDR	0.51	0.5	0.58	0.5	0.5	0.5	0.51	0.52
Therapies on pretreated patients	0.68	0.69	0.59	0.68	0.67	0.69	0.69	0.69
All	0.67	0.67	0.65	0.66	0.65	0.66	0.66	0.66

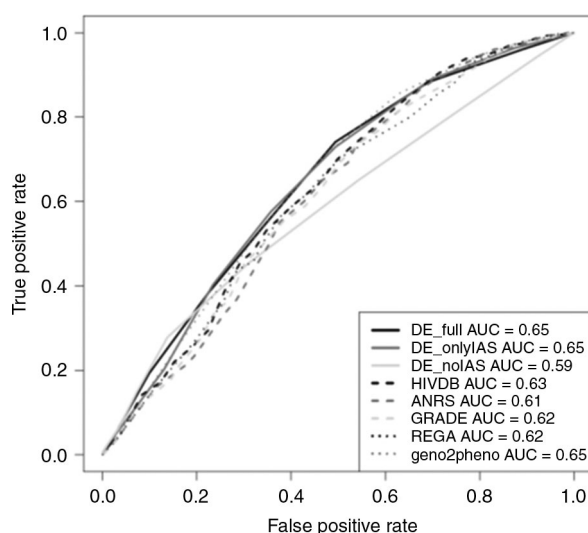


Figure 1. Performance comparison for therapy-success prediction in test set B.

Published 2 November 2014

Copyright: © 2014 Pironti A et al; licensee International AIDS Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

scores that predict drug exposure pertaining to each of 20 antiretrovirals: the first one uses the full amino-acid sequences (DE_{full}), the second one only considers IAS drug-resistance positions ($DE_{onlyIAS}$), and the third one disregards IAS drug-resistance positions (DE_{noIAS}). For performance comparison, test sets A and B were evaluated with DE_{full} , DE_{noIAS} , $DE_{onlyIAS}$, $geno2pheno_{[resistance]}$, HIVdb, ANRS, HIV-GRADE, and REGA. Clinically-validated cut-offs were used to convert the continuous output of the first four methods into susceptible-intermediate-resistant (SIR) predictions. With each method, a genetic susceptibility score (GSS) was calculated for each therapy episode in each test set by converting the SIR prediction for its compounds to integer: $S = 2$, $I = 1$, and $R = 0$. The GSS were used to predict therapy success as defined by the EuResist standard datum definition. Statistical significance was assessed using a Wilcoxon signed-rank test.

Results: A comparison of the therapy-success prediction performances among the different interpretation systems for test set A can be found in Table 1, while those for test set B are found in Figure 1. Therapy-success prediction of first-line therapies with DE_{noIAS} performed better than $DE_{onlyIAS}$ ($p < 10^{-16}$).

Conclusions: Therapy success prediction benefits from the consideration of all available mutations. The increase in performance was largest in first-line therapies with transmitted drug-resistance mutations.