Antiviral Therapy 9:343-352

Simple linear model provides highly accurate genotypic predictions of HIV-1 drug resistance

*Kai Wang, Ekachai Jenwitheesuk, Ram Samudrala and John E Mittler**

Department of Microbiology, University of Washington, Seattle, Wash., USA

*Corresponding author: Tel: +1 206 732 6160; Fax: +1 206 732 6055; E-mail: jmittler@u.washington.edu

Drug resistance is a major obstacle to the successful treatment of HIV-1 infection. Genotypic assays are used widely to provide indirect evidence of drug resistance, but the performance of these assays has been mixed. We used standard stepwise linear regression to construct drug resistance models for seven protease inhibitors and 10 reverse transcriptase inhibitors using data obtained from the Stanford HIV drug resistance database. We evaluated these models by hold-one-out experiments and by tests on an independent dataset. Our linear model outperformed other publicly available genotypic interpretation algorithms, including decision tree, support vector machine and four rules-based algorithms (HIVdb, VGI, ANRS and Rega) under both tests. Interestingly, our model did well despite the absence of any terms for interactions between different residues in protease or reverse transcriptase. The resulting linear models are easy to understand and can potentially assist in choosing combination therapy regimens.

Introduction

Inhibitors of human immunodeficiency virus type 1 (HIV-1) protease and reverse transcriptase are widely used in the clinical treatment of acute immunodeficiency syndrome (AIDS). However, drug-resistant variants of the virus severely limit the long-term effectiveness of anti-HIV drugs [1]. In recent years, genotypic or phenotypic resistance testing has become an important part of choosing and optimizing combination therapy for treating HIV-infected individuals [2]. Phenotypic assays directly measure viral replication in the presence of drug, but the process is time-consuming and expensive. Genotypic assays do not have such disadvantages, though the resulting sequence information can be hard to interpret because of the large number and the complex patterns of drug-resistance mutations [3].

A variety of algorithms have been developed for HIV genotypic interpretation, including database pattern search method [4], rule-based algorithms [5–8], neural networks/machine learning [9–12], molecular dynamics simulations [13,14], decision trees/recursive partitioning [9,15,16] and linear discriminant analysis [9]. However, current widely used genotypic interpretation systems still do not have satisfactory performance on newly derived datasets [17–19]. We hypothesized that HIV drug resistance might be better predicted using a simple linear model, in which each mutation contributes to drug resistance independently and quantitatively. To select independent variables (mutations) to be included in our model, we relied on standard stepwise regression techniques. We then evaluated this model by various hold-one-out procedures and by a test against an independent dataset. Results from our model were compared with other widely used genotypic interpretation methods, including four rule-based algorithms (HIVdb [5], VGI [8], ANRS [7] and Rega [6]), the decision tree (DT) algorithm [16] and the support vector machine (SVM) algorithm [11]. Although the rule-based algorithms are largely based on clinical trial data and designed to predict clinical outcome instead of phenotypic response, such comparisons are still informative exercises for the evaluation of our linear model.

Methods

Data source

We downloaded genotypic data and corresponding drug resistance data (version 1.2) for seven protease inhibitors and 10 reverse transcriptase inhibitors from the Stanford HIV drug resistance database (http://hivdb.stanford.edu/cgi-bin/GenoPhenoDS.cgi). The genotypic data is solely the sequence information for the HIV protease or reverse transcriptase. The drug resistance data were represented by the IC_{50} (50% inhibitory concentration) fold change over the wild-type virus with subtype B consensus sequence, and were determined by either Virologic's PhenoSense assay or Virco's Antivirogram assay. We divided these data into a Virologic dataset and a Virco dataset. To

reduce the number of independent variables in our regression model, we only used mutations in 'important' positions that are known to influence drug resistance, and are classified as major and minor mutations in the Stanford database. These are at positions 10, 20, 24, 30, 32, 33, 36, 46, 47, 48, 50, 53, 54, 60, 63, 71, 73, 77, 82, 84, 88, 90 and 93 for protease inhibitors, and positions 41, 44, 62, 65, 67, 69, 70, 74, 75, 77, 98, 100, 101, 103, 106, 108, 115, 116, 118, 151, 179, 181, 184, 188, 190, 210, 215, 219, 225, 227, 230, 236 and 238 for reverse transcriptase inhibitors. In cases where the IC_{50} fold change was reported as more or less than a certain value, we treated it as the exact value. When the IC_{50} fold change was reported as zero, we treated it as 0.1 for the purposes of taking natural logarithms.

Full regression model

We built up a full regression model based upon our current knowledge about HIV drug resistance. In the model, the dependent variable is the natural logarithm of the IC_{50} fold change, while the independent variables are indicator variables corresponding to mutations. We used the natural logarithm of IC_{50} fold change as the dependent variable since the binding energy between drug and corresponding HIV target is proportional to the logarithm of inhibition constant K_i based on the Gibbs free energy equation, and K_i is proportional to IC_{50} based on the Cheng-Prusoff equation [20]. Each independent variable takes the value of 1 if the sequence contains the corresponding mutation and 0 otherwise. Due to the small size of the datasets, we did not arbitrarily discard rare mutations from the full model; instead, we used stepwise regression procedures (see below) to decide whether or not to keep an independent variable (mutation). In a few cases where the sequence contained a mixture of two or more amino acids at the same position, we took the value of each of the corresponding indicator variables to be 1. Due to the relatively small number of records compared to the number of variables, sometimes two or more independent variables have the exactly same distribution in all the records. In such cases we simply kept one such variable at random.

Stepwise regression method

Our model treats the natural logarithm of IC_{50} fold change as a linear combination of position- and type-specific mutations with different weights plus a constant. We used a backward stepwise regression method to optimize the parameters for each independent variable and the constant. The stepwise regression begins with the full model, where all independent variables are contained in the model. In each subsequent step, the removal statistic is computed for each independent variable eligible to be removed from the model, and the variable with the highest removal statistic is removed from the current model if it is more than a critical removal value. Then the entry statistic is computed for each independent variable that is not included in the current model, and the variable with the lowest entry statistic is added into the current model if it is lower than the entry statistic. The stepwise regression stops if neither entry nor removal is performed in the step, and the remaining variables comprise the reduced model. We used the *P*-value as the entry and removal statistic, which indicates the possibility of observing such data when such variable is not associated with the dependent variable. The critical *P*-values for removal and entry were set to 0.051 and 0.05, respectively. Regression analysis and data manipulation were done using the statistics software STATA (College Station, Tex., USA).

Evaluation and comparison of genotypic interpretation algorithms

Two validation methods were used for the evaluation and comparison of our linear model and six other algorithms. In the first evaluation method, we used a hold-one-out procedure on the Virologic dataset, where for every genotype–phenotype paired record in the dataset, we trained our model on all the other records and tested our model on this record. In the second evaluation method, we trained our model on the Virologic dataset, and tested the resulting model on all records in the Virco dataset. Since the rule-based algorithms cannot give quantitative prediction of IC_{50} values, for comparison purposes, we categorized each record as either resistant or susceptible, based on manufacturer recommended cut-off values at http://www.phenosense.com/pdf/CLINICAL-CUTOFF.pdf for the Virologic dataset and at http://www.vircolab.com/web/page.asp?id=84 for the Virco dataset. Clinical cut-offs are used for five drugs [abacavir (ABC), didanosine (ddI), stavudine (d4T), tenofovir (TFV) and lopinavir (LPV)] in the Virologic dataset, while all other cut-offs are biological cut-offs. Neither company gave a cut-off value for the relatively new drug atazanavir (ATV), so we arbitrarily set it to be 2.5 for both datasets.

Prediction by other algorithms

The phenotypic prediction by ANRS (version 2002.3), HIVdb (version 2003.8), Rega (version 5.5) and VGI (version 4) was done through the HIValg tool (version 3.6) at http://hivdb.stanford.edu, which gives a 'SIR' interpretation for each genotype, representing sensitive, intermediate resistant and resistant, respectively. We combined 'I' and 'R' predictions into a single resistant category. The most recent version of ANRS and

Rega algorithm is not available through the Stanford database. The prediction by decision tree and support vector machine was done through the geno2pheno tool (version 2.2) at http://www.genafor.org, using the same or rounded (when the same cut-off cannot be handled by the website) cut-off values as our regression model.

A simple pattern-matching algorithm was also used in a control experiment to evaluate data quality. For any given sequence, we compare it with all other sequences in the dataset, and predict its IC_{50} value as that of the sequence with the highest identity (as determined by the number of identical residues). This algorithm is extremely simple to perform, and will generate good results if the dataset contains many very similar records.

Results

Performance of the regression model on Virologic dataset

We collected 5507 genotype–phenotype paired records covering 17 anti-HIV drugs from the Virologic dataset. We constructed a separate regression model for each drug and checked the validity of these models. We checked the distribution of studentized residuals and noted a highly skewed distribution for the drug lamivudine (3TC). We plotted the prediction error versus predicted value for all models. With the exception of 3TC and tenofovir disoproxil fumarate (TDF), we had good homogeneity of error variance. Since multicollinearity between variables can destabilize regression models, we also checked variance inflation factors for independent variables in all models. No multicollinearity problems were detected except for the drug TDF. Due to the violation of regression assumptions for 3TC and TDF, our models for these drugs cannot be considered reliable.

Due to the limited number of records available, we performed hold-one-out experiments to estimate the predictive power of our method. In this cross-validation technique, for each sample, a regression model is generated on all but this sample and the resulting model is then used to make a quantitative prediction for this sample. For the 15 drugs for which we have reliable models, the correlation coefficients between the natural logarithm of experimental and predicted IC_{50} fold change values were generally very high, ranging from 0.761 for ATV to 0.946 for ritonavir (RTV), with a median of 0.887 (Table 1). The scatter plots also demonstrated the good correspondence between the experimental values and our quantitative predictions (Figure 1). The low correlation and poor performance for ATV in the hold-one-out experiments can be explained by the fact that we had only 101 records. For drugs for which we had at least 200 records and reliable models, the correlation coefficients were 0.832 or higher.

A more direct measure of prediction accuracy is the binary prediction of whether the sample is resistant or susceptible to a certain drug, using manufacturer-established IC_{50} cut-off values. We calculated the sensitivity and specificity of our prediction by the hold-one-out procedure (Table 1), and compared the percent correct prediction given by our regression method and other publicly available algorithms on the same dataset (Figure 2). These algorithms included four rules-based algorithms (ANRS, HIVdb, Rega, VGI), the DT algorithm and the SVM algorithm. Among all the seven algorithms, our regression method had the best overall prediction accuracy (88.7% on average), giving the highest prediction rate for 10 out of 15 drugs for which we have reliable models, and the second or third highest prediction rate for the other five drugs. For the two drugs (3TC and TDF) for which our regression assumptions were violated, the performance of our regression model was still comparable to that of other algorithms.

Since our dataset was collected from the literature (13–27 publications for each drug except 3TC and TDF), it is likely to contain phylogenetically related sequences and sequences from patients with similar drug treatment histories. The presence of such 'non-independent' sequences could potentially lead to spuriously high correlations and prediction accuracies in our hold-one-out experiments. To assess the extent of this problem, as a control experiment, we used a simple pattern-matching algorithm to determine IC_{50} values using the same hold-one-out procedure. If the good performance of the regression model were due to the presence of many highly similar sequences, we would expect this simple pattern-matching algorithm to have comparable performance. The average correlation between experimental and predicted IC_{50} values under this simple pattern-matching method was 0.68, which is significantly lower than the average correlation of 0.88 obtained using our regression procedures.

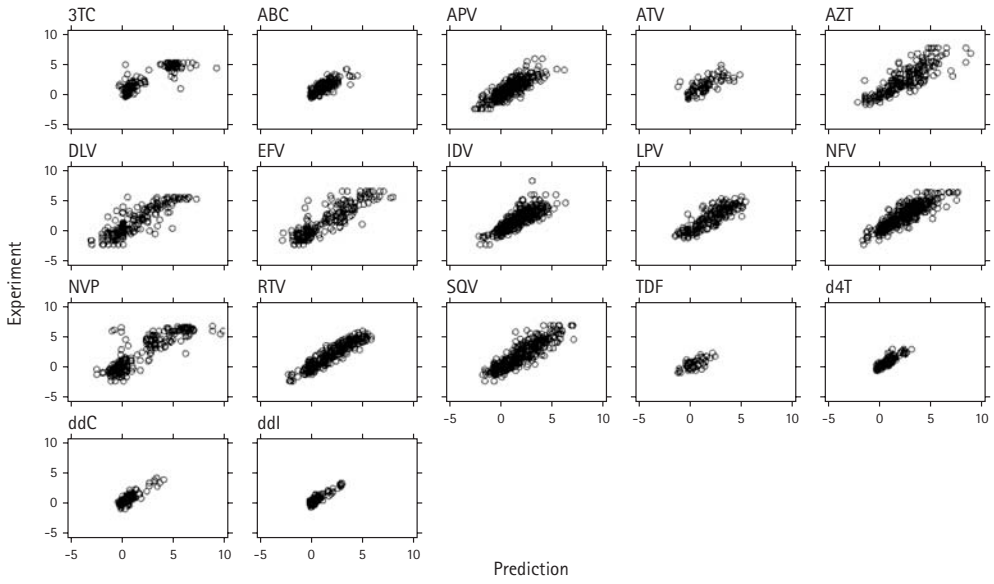
To further address this 'related-sequence' hypothesis, we performed a series of modified hold-one-out regression experiments in which sequences that were the same as the hold-out sequence, sequences that differed from the hold-out sequence at one amino acid position, and sequences that differed from the hold-out sequence at two amino acid positions were removed prior to the hold-one-out experiments. These removal experiments resulted in deletion of 24.2, 34 and 39% of the records from the dataset, respectively (due to limited sample size, for the latter set of deletions we could only estimate correlation coefficients for some of the drugs). Despite these deletions, the average correlation coefficients and the percent prediction rates

Table 1. Performance of regression models on Virologic dataset using the hold-one-out experiments

Drug class	Drug	Number of records	Fold change cut-off	Resistant fraction (%)	Results for hold-one-out experiments on Virologic dataset			
					Correlation coefficient	Rank correlation coefficient	Sensitivity (%)	Specificity (%)
PI	APV	445	2.5	46.5	0.860	0.859	79.2	88.7
	ATV	101	2.5	63.4	0.761	0.784	84.4	81.1
	IDV	510	2.5	60.0	0.872	0.887	91.8	90.7
	LPV	228	10.0	69.7	0.867	0.858	82.8	84.8
	NFV	517	2.5	74.5	0.887	0.876	95.3	85.6
	RTV	438	2.5	55.7	0.946	0.929	95.9	91.8
	SQV	503	2.5	51.5	0.909	0.901	95.0	87.7
NNRTI	DLV	340	2.5	32.1	0.897	0.799	80.7	93.9
	EFV	333	2.5	35.7	0.907	0.832	93.3	97.2
	NVP	348	2.5	41.1	0.887	0.775	86.7	98.0
NRTI	3TC*	282	1.5	80.5	0.928	0.759	97.8	1.8
	ABC	283	4.5	45.9	0.832	0.810	80.0	88.2
	AZT	281	2.2	60.1	0.880	0.898	98.2	83.0
	d4T	284	1.7	51.1	0.906	0.884	86.2	89.2
	ddC	272	1.7	58.8	0.887	0.756	94.4	84.8
	ddI	282	1.7	42.2	0.911	0.762	56.3	89.0
	TDF*	60	1.4	51.7	0.694	0.596	67.7	62.1

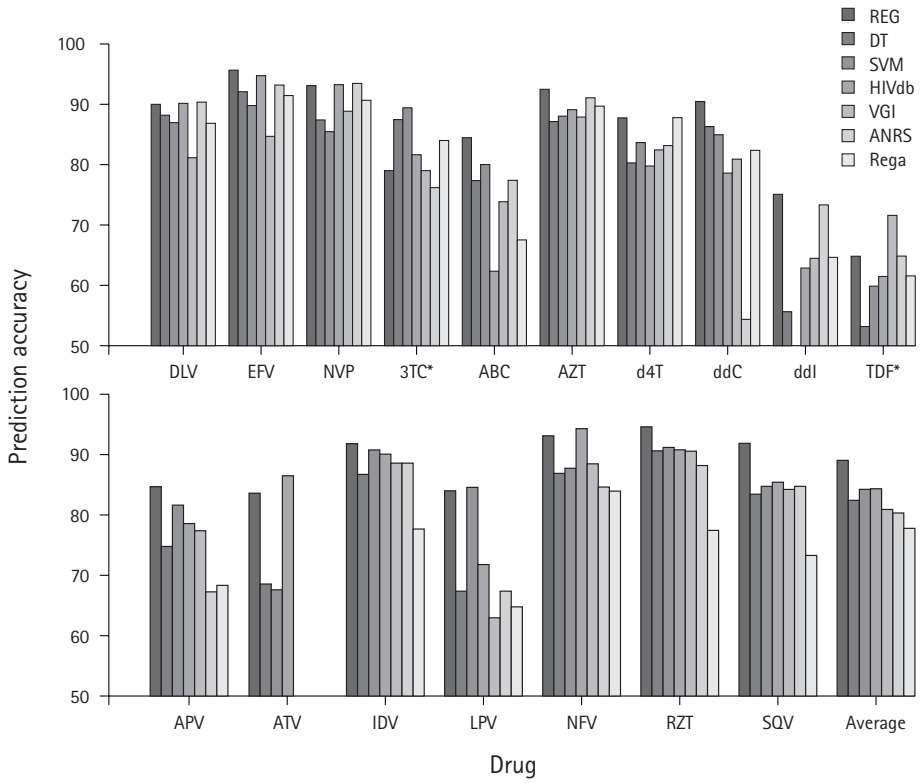
In general the regression models are highly accurate in quantitative prediction.
PI, protease inhibitor; NNRTI, non-nucleoside reverse transcriptase inhibitor; NRTI, nucleoside reverse transcriptase inhibitor; APV, amprenavir; ATV, atazanavir; IDV, indinavir; LPV, lopinavir; NFV, nelfinavir; RTV, ritonavir; SQV, saquinavir; DLV, delavirdine; EFV, efavirenz; NVP, nevirapine; 3TC, lamivudine; ABC, abacavir; AZT, zidovudine; d4T, stavudine; ddC, zalcitabine; ddI, didanosine; TDF, tenofovir.
*Model not reliable due to violation of regression assumptions.

Figure 1. Scatter plot of experimental versus predicted resistance values on the Virologic dataset



These values are represented as natural logarithm of IC_{50} fold change. For most drugs our quantitative predictions correlate well with the experimental values.

Figure 2. Comparison of prediction accuracy given by seven algorithms on Virologic dataset



The prediction accuracy for our regression model (REG) is calculated by the hold-one-out procedure. Three rule-based algorithms (VGI, ANRS and Rega) do not have interpretations for ATV. The average prediction accuracy values are calculated for all records for which the corresponding algorithm makes predictions. Our regression method has the best overall prediction accuracy among the seven algorithms. *Drugs for which regression assumptions were violated.

changed very little (–0.02 and –0.7%, –0.02 and 0.0%, and –0.04 and +0.2%, respectively), indicating that the high correlations and high predictions accuracies by our regression model in the hold-one-out experiments cannot be explained by the presence of very similar sequences in the dataset.

Performance of the regression model on Virco dataset

We also tested the ability of our models, which are based on sequences phenotyped by the Virologic PhenoSense assay, to predict resistance levels in an independent dataset (sequences that were phenotyped by the Virco Antivirogram assay) (Table 2). Both the Virologic and Virco assays are standardized commercial assays with good reproducibility. To help ensure independence, any sequence that was common to the Virologic dataset was removed from the Virco dataset [21]. The correlation coefficients between experimental and predicted values were much lower than those of the Virologic dataset, which was not unexpected because of the discordance of different phenotypic testing strategies [22]. In general the prediction accuracy was lower for the seven algorithms we evaluated, which might be due to the cut-off discordances, or the different performance of two distinct phenotyping assays, or the presence of non-standard mutations in the Virco dataset. Nevertheless, our method still had the highest overall prediction accuracy (86.8% on average) and out-performed all other methods for nine out of 15 drugs for which we had reliable models (Figure 3). For four of the drugs [d4T, ddI, zalcitabine (ddC) and TDF] our regression models have very low sensitivity, which is mainly due to the higher cut-off values provided by the manufacturer. Since our models can give quantitative prediction, in clinical practice we may change the cut-off in our binary prediction to achieve higher sensitivity given required minimum specificity values.

Analysis of the regression model for lopinavir

Since LPV (manufactured in combination with RTV) is currently one of the most administered protease inhibitors, we compared the mutations in our regression model with mutations previously reported to confer LPV resistance. A recent report (revision October 2003) by the IAS-USA panel (http://www.iasusa.org) listed mutations at 16 positions of protease that are associated with LPV

Table 2. Performance of the regression models that were trained on Virologic dataset and tested on Virco dataset

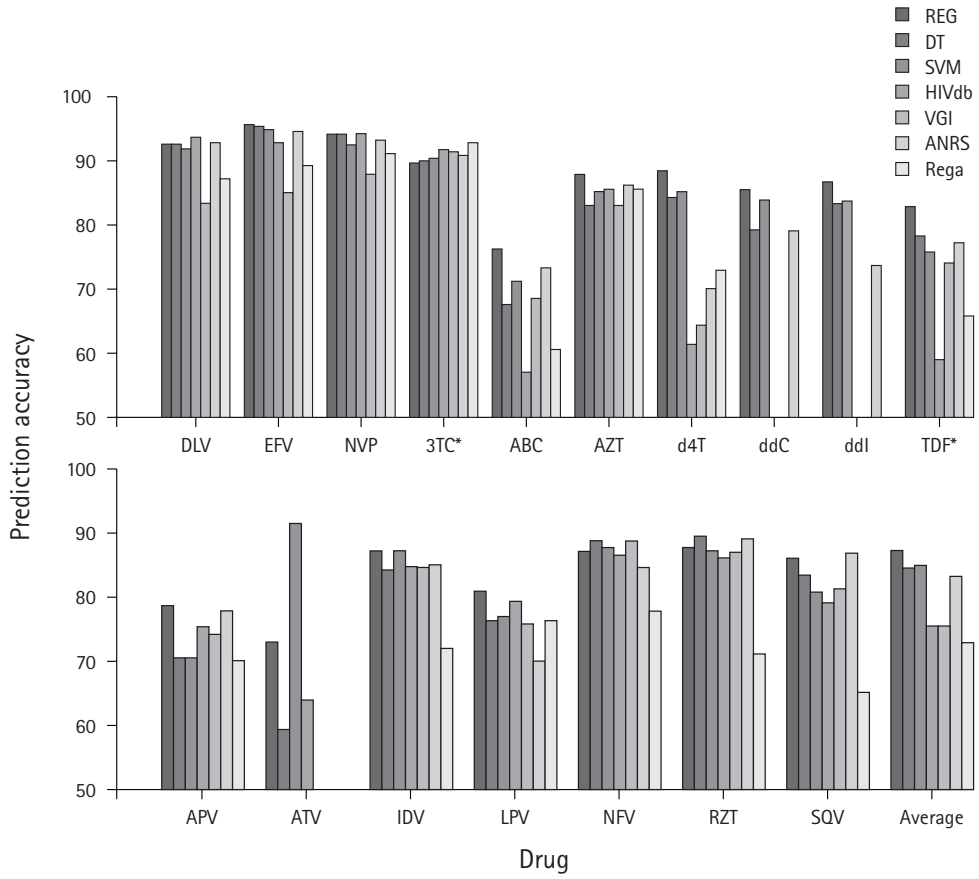
Drug class	Drug	Number of records	Fold change cut-off	Resistant fraction (%)	Results on Virco dataset			
					Correlation coefficient	Rank correlation coefficient	Sensitivity (%)	Specificity (%)
PI	APV	241	2.5	35.7	0.640	0.620	61.6	84.5
	ATV	22	2.5	54.5	0.551	0.564	100.0	40.0
	IDV	322	3.0	49.1	0.747	0.763	82.9	89.0
	LPV	175	2.5	41.7	0.646	0.685	80.8	80.4
	NFV	313	4.0	58.1	0.773	0.796	90.1	80.9
	RTV	325	3.5	50.2	0.816	0.818	89.0	85.8
	SQV	324	2.5	33.0	0.809	0.744	89.7	83.4
NNRTI	DLV	292	10.0	31.5	0.840	0.780	73.9	99.5
	EFV	248	6.0	38.3	0.907	0.829	91.6	97.4
	NVP	371	8.0	34.2	0.777	0.658	82.7	99.2
NRTI	3TC*	417	4.5	67.4	0.854	0.772	87.5	93.4
	ABC	348	3.0	35.6	0.664	0.672	81.5	70.1
	AZT	407	4.0	43.0	0.754	0.807	82.9	89.7
	d4T	382	3.0	16.0	0.528	0.441	60.7	93.1
	ddC	385	3.5	17.1	0.458	0.465	21.2	97.8
	ddI	398	3.5	16.8	0.560	0.487	19.4	98.8
	TDF*	161	3.0	16.1	0.428	0.435	46.2	89.6

The correlation coefficients were lower than those of the hold-one-out experiments on Virologic dataset. Most models have high prediction accuracy on binary prediction. PI, protease inhibitor; NNRTI, non-nucleoside reverse transcriptase inhibitor; NRTI, nucleoside reverse transcriptase inhibitor; APV, amprenavir; ATV, atazanavir; IDV, indinavir; LPV, lopinavir; NFV, nelfinavir; RTV, ritonavir; SQV, saquinavir; DLV, delavirdine; EFV, efavirenz; NVP, nevirapine; 3TC, lamivudine; ABC, abacavir; AZT, zidovudine; d4T, stavudine; ddC, zalcitabine; ddI, didanosine; TDF, tenofovir. *Model not reliable due to violation of regression assumptions.

resistance: amino acids positions 10, 20, 24, 32, 33, 46, 47, 50, 53, 54, 63, 71, 73, 82, 84 and 90. We found significant scores for mutations at 13 of these 16 positions in our reduced regression model trained on the Virologic dataset (Figure 4). Of the three positions (32, 53 and 90) that are not included in the regression model, position 90 is most mysterious, since it is commonly believed that the 90M mutation is associated with reduced susceptibility to all proteases though the mechanism remains unknown [23]. It is, therefore, interesting to investigate why 90M was discarded by stepwise regression. In the Virologic dataset we used, the 90M mutation had a prevalence of 37.84% (42/111) in susceptible patients and a prevalence of 56.41% (66/117) in patients with reduced susceptibility. A Wilcoxon rank-sum test showed a significant difference in distribution of IC_{50} values for LPV between those patients with or without the 90M mutation (P=0.0014), consistent with previous reports [24]. However, when analysing the 108 records that contained the 90M mutation more carefully, we found that 90M always co-occurred with other highly resistant mutations reported by our model. We found that 86.11% (93/108), 47.22% (51/108), 52.78% (57/108), 38.89% (42/108), 45.37% (49/108) and 38.89% (42/108) of the 90M mutations co-occurred with the mutations 10F/I/V/X, 46I/X, 54T/V/X, 71V, 82A/F/X and 84V, respectively. In fact, in our dataset the 90M mutation never occurred without a mutation at one of the six positions above. The same phenomenon happens for the 32I and 53L mutations (mutations that have much lower prevalence than 90M). Multivariate analyses conducted by other researchers have also failed to associate the above three mutations with LPV resistance [24,25]. These observations suggest that these three mutations per se may not contribute to drug resistance; they may merely tend to co-occur with other highly resistant mutations so patients with this mutation tend to have higher IC_{50} values.

We also noticed that mutations at six additional positions not in the IAS-USA report were included in the regression model: 30N, 36X, 48M/V, 77T, 88D/G/S and 93L. A recent report has shown that the 48M/V mutation is associated with LPV resistance [26]. Both the IAS-USA panel and other reviewers [1] reported that mutations at position 36 are commonly associated with resistance to IDV, RTV, NFV and AZT, while the 30N and 88D/S mutations are associated with resistance to NFV. Our model suggests possible cross resistance of these mutations. Although the 48M, 77T and 88G mutations have high scores in the regression model, they are rare and insignificant mutations that occur only once or twice in the dataset.

Figure 3. Comparison of prediction accuracy given by seven algorithms on Virco dataset



Our regression models (REG) used here were trained on Virologic dataset. Three rule-based algorithms (VGI, ANRS and Rega) do not have interpretations for ATV. The average prediction accuracy values are calculated for all records for which the corresponding algorithm makes predictions. Our regression method has the best overall prediction accuracy among the seven algorithms. *Drugs for which regression assumptions were violated.

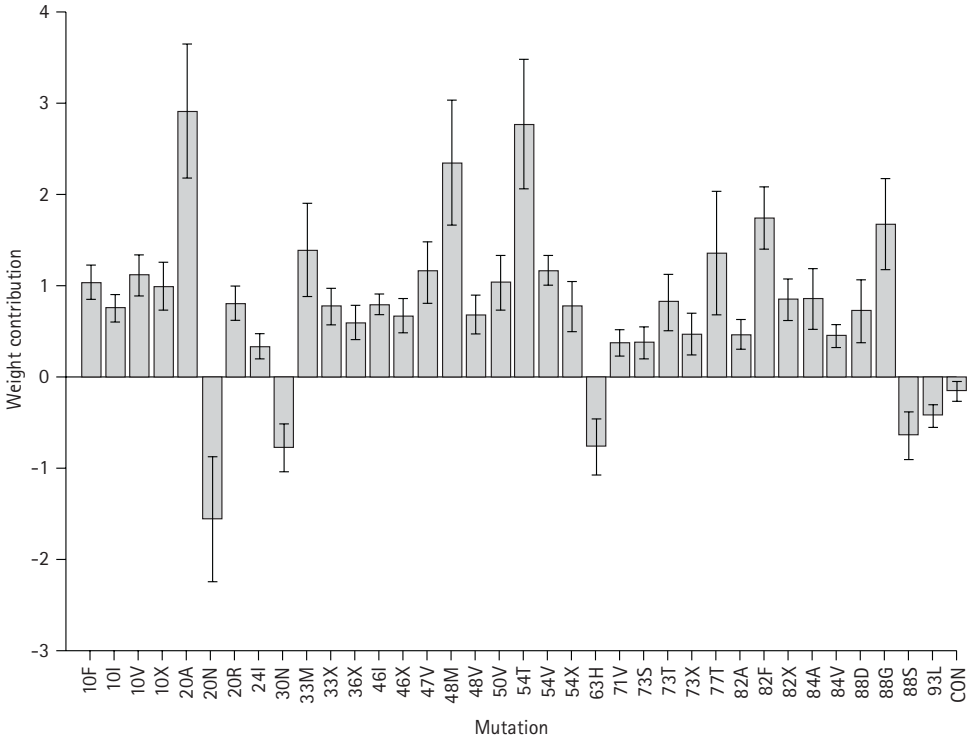
Interestingly, we also noted that five mutations (20N, 30N, 63H, 88S and 93L) had negative scores in our regression model (–1.54, –0.75, –0.75, –0.63 and –0.41, respectively), indicating enhanced susceptibility to LPV. Among them, 93L is especially interesting because it has a high prevalence of 26.32% (60/208) in the dataset. A Wilcoxon rank-sum test confirmed that patients carrying this mutation tend to have lower IC_{50} values (P=0.0033). In the rule-based algorithm HIVdb [5], all of these five mutations had very low or zero scores, though HIVdb gives no hint that these mutations may enhance susceptibility to LPV in a statistical sense. This analysis of the LPV model illustrates the fact that our method largely corresponds to current biological knowledge from the literature, and that it can identify new mutations that may be of importance to understanding drug resistance. All regression models and executable programs are available at http://software.compbio.washington.edu/misc/downloads/hiv_lr/.

Discussion

We used a purely statistical approach to investigate the relationship between HIV genotype and drug resistance. Based on both hold-one-out experiments and tests on an independent dataset, our method outperformed six other publicly available HIV genotypic interpretation algorithms for most drugs at predicting drug resistance in published datasets. This good performance may be due to the quantitative nature of our statistical model. Unlike some of the other methods, which classify sequences into distinct categories (for example, sensitive or resistant), our method employs IC_{50} values in a quantitative fashion. Since our model is both accurate and easy to understand by clinicians, we believe that it has great potential for use in selecting combination therapies when combined with other pertinent clinical and pharmacological information.

One surprise was that our model performed well despite the absence of statistical terms for interaction

Figure 4. Regression model for lopinavir trained on the entire Virologic dataset



Independent variables (mutations) and the constant in the linear model are shown in the x-axis, while the corresponding weight contribution and the standard error of the weights are shown in the y-axis. For a given sequence, the natural logarithm of IC_{50} fold change for LPV can be estimated by sum of the weights of corresponding mutations and the weight of the constant. This model can also be used to interpret effect of individual mutations in a statistical sense.
X, unknown mutations that are not explicitly represented in the Stanford database (such mutations are coded as 'Z' in the Stanford database).

effects. Experimental evidence has shown that interactions between different drug-resistance mutations can occur [27]; however, it is not always possible to obtain data for all of the genotypes (WT-WT, WT-MUT, MUT-WT and MUT-MUT) needed to quantify an interaction effect. A compensatory mutation, for example, may only appear after the primary resistance mutations have appeared, while incompatible mutations may never co-occur *in vivo*. Without a large number of genotype–phenotype records, it is hard to quantify such interactions from clinical datasets. Thus, while our model is good at predicting resistance from typical clinical isolates, it may not work well in predicting resistance for specially made laboratory strains. For the same reason, we caution that the weights for each mutation in our model are purely statistical entities, and should not be interpreted in a physical or biological sense as to the actual binding of drugs and HIV targets. Another potential problem with our regression models is that they cannot be used for novel rare mutations that are not included in our training process. However, with the increasing size of public databases, this problem will be less severe in the future.

We note that our model, like most other genotypic models, is optimized to predict fold changes in IC_{50} values instead of clinical outcome. Although other factors, such as viral fitness in the absence of drug, the shape and slope of the viral inhibition curve, pharmocokinetics, patient treatment history, adherence to drug regimens and stochastic mutational events, influence clinical response to drug therapy, IC_{50} values are the most obvious and widely used predictors of drug resistance. In clinical settings, IC_{50} values have high correlation with response to antiretroviral therapy [28], demonstrating that IC_{50} values can be used to predict clinical outcome. In addition, constantly monitoring IC_{50} values by genotypic tests during treatment may provide further insight to patient prognosis, since it has been shown that resistance is generated mainly by conversion of drug-sensitive cases to drug-resistant cases, not by transmission of resistant strains [29]. Our method has the advantage that it gives highly accurate quantitative predictions with statistically meaningful confidence intervals. When combined with other quantitative models that account for mutation rates, pharmacodynamics [30] and viral fitness in the absence of drug [31], our results suggest a future in which genotypic assays could become highly reliable predictors of drug resistance.

Acknowledgements

This work was supported by NIH grant R21-AI52063-01 (JEM) and a Searle Scholar's Award (RS). We thank the authors of the published HIV resistance datasets for making their data publicly available. We thank the curators of the Stanford HIV drug resistance database for collecting and compiling the drug resistance data used in our study.

References

1. Menendez-Arias L. Targeting HIV: antiretroviral therapy and development of drug resistance. *Trends in Pharmacological Sciences* 2002; **23:**381–388.
2. Aslanzadeh J. HIV resistance testing: an update. *Annals of Clinical & Laboratory Science* 2002; **32:**406–413.
3. Shafer RW, Kantor R & Gonzales MJ. The genetic basis of HIV-1 resistance to reverse transcriptase and protease inhibitors. *AIDS Reviews* 2000; **2:**211–228.
4. Shafer RW, Jung DR & Betts BJ. Immunodeficiency virus type 1 reverse transcriptase and protease mutation search engine for queries. *Nature Medicin*e 2000; **6:**1290–1292.
5. Kantor R, Machekano R, Gonzales MJ, Dupnik K, Shapiro JM & Shafer RW. Human immunodeficiency virus reverse transcriptase and protease sequence database: an expanded data model integrating natural language text and sequence analysis programs. *Nucleic Acids Research* 2001; **29:**296–299.
6. Van Laethem K, De Luca A, Antinori A, Cingolani A, Perna CF & Vandamme AM. A genotypic drug resistance interpretation algorithm that significantly predicts therapy response in HIV-1-infected patients. *Antiviral Therapy* 2002; **7:**123–129.
7. Rousseau M-N, Vergne L, Montes B, Peeters M, Reynes J, Delaporte E & Segondy M. Patterns of resistance mutations to antiretroviral drugs in extensively treated HIV-1-infected patients with failure of highly active antiretroviral therapy. *Journal of Acquired Immune Deficiency Syndromes* 2001; **26:**36–43.
8. Reid C, Bassett R, Day S, Larder B, DeGruttola V & Winslow D. A dynamic rules-based interpretation system derived by an expert panel is predictive of virological failure. *Antiviral Therapy* 2002; **7:**S91.
9. Sevin AD, DeGruttola V, Nijhuis M, Schapiro JM, Foulkes AS, Para MF & Boucher CAB. Methods for investigation of the relationship between drug-susceptibility phenotype and human immunodeficiency virus type 1 genotype with applications to AIDS clinical trials group. *Journal of Infectious Diseases* 2000; **182:**59–67.
10 Wang D, DeGruttola V, Hammer S, Harrigan R, Larder B, Wegner S, Winslow D & Zazzi M. A collaborative HIV resistance response database initiative: predicting virological response using neural network models. *Antiviral Therapy* 2002; **7:**S96.
11. Beerenwinkel N, Schmidt B, Walter H, Kaiser R, Lengauer T, Hoffmann D, Korn K & Selbig J. Quantitative phenotype prediction by support vector machines. *Antiviral Therapy* 2002; **7:**S74.
12. Draghici S & Potter RB. Predicting HIV drug resistance with neural networks. *Bioinformatics* 2003; **19:**98–107.
13. Wang W & Kollman PA. Computational study of protein specificity: The molecular basis of HIV-1 protease drug resistance. *Proceedings of the National Academy of Sciences, USA* 2001; **98:**14937–14942.
14. Jenwitheesuk E & Samudrala R. Improved prediction of HIV-1 protease-inhibitor binding energies by molecular dynamics simulations. *BMC Structural Biology* 2003; **3:**2.
15. King M, Kempf D, Isaacson J, Rode R, Brun S, Bernstein B, Calvez V, Cohen-Codar I, Guillevic E, Chauvin J & Sun E. Using classification trees to explore relationships between viral genotype and response to lopinavir/ritonavir-based regimens. *Antiviral Therapy* 2002; **7:**S82.
16. Beerenwinkel N, Schmidt B, Walter H, Kaiser R, Lengauer T, Hoffmann D, Korn K & Selbig J. Diversity and complexity of HIV-1 drug resistance: a bioinformatics approach to predicting phenotype from genotype. *Proceedings of the National Academy of Sciences, USA* 2002; **99:**8271–8276.
17. Braun P, Helm M, Ehret R, Schmidt B, Stürner K, Walter H, Hoehn C, Korn K & Knechten H. Predictive value of different drug resistance interpretation systems in therapy management of HIV-infected patients in daily routine. *Antiviral Therapy* 2002; **7:**S77.
18. Zolopa A, Lazzeroni L, Rinehart A & Kuritzkes D. Accuracy, precision and consistency of expert HIV-1 genotype interpretation: an international comparison (The GUESS study). *Antiviral Therapy* 2002; **7:**S97.
19. De Luca A, Cingolani A, Giambenedetto SD, Trotta MP, Baldini F, Rizzo MG, Bertoli A, Liuzzi G, Narciso P, Murri R, Ammassari A, Perno CF & Antinori A. Variable prediction of antiretroviral treatment outcome by different systems for interpreting genotypic human immunodeficiency virus type 1 drug resistance. *Journal of Infectious Diseases* 2003; **187:**1934–1943.
20. Cheng Y-C & Prusoff WH. Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochemical Pharmacology* 1973; **22:**3099–3108.
21. Qari SH, Respess R, Weinstock H, Beltrami EM, Hertogs K, Larder BA, Petropoulos CJ, Hellmann N & Heneine W. Comparative analysis of two commercial phenotypic assays for drug susceptibility testing of human immunodeficiency virus type 1. *Journal of Clinical Microbiology* 2002; **40:**31–35.
22. Wang K, Samudrala R & Mittler JE. Weak agreement between predictions of 'reduced susceptibility' from antivirogram and phenosense assays. *Journal of Clinical Microbiology* 2004; **42:**2353–2354.
23. Shafer RW. Genotypic testing for human immunodeficiency virus type 1 drug resistance. *Clinical Microbiology Reviews* 2002; **15:**247–277.
24. Kempf DJ, Isaacson JD, King MS, Brun SC, Xu Y, Real K, Bernstein BM, Japour AJ, Sun E & Rode RA. Identification of genotypic changes in human immunodeficiency virus protease that correlate with reduced susceptibility to the protease inhibitor lopinavir among viral isolates from protease inhibitor-experienced patients. *Journal of Virology* 2001; **75:**7462–7469.
25. Paulsen D, Liao Q, Fusco G, St Clair M, Shaefer M & Ross L. Genotypic and phenotypic cross-resistance patterns to lopinavir and amprenavir in protease inhibitor-experienced patients with HIV viremia. *AIDS Research & Human Retroviruses* 2002; **18:**1011–1019.
26. Parkin NT, Chappey C & Petropoulos CJ. Improving lopinavir genotype algorithm through phenotype correlations: novel mutation patterns and amprenavir cross-resistance. *AIDS* 2003; **17:**955–961.
27. Tisdale M, Kemp S, Parry N & Larder B. Rapid *in vitro* selection of human immunodeficiency virus type 1 resistant to 3'-thiacytidine inhibitors due to a mutation in the YMDD region of reverse transcriptase. *Proceedings of the National Academy of Sciences, USA* 1993; **90:**5653–5656.
28. DeGruttola V, Dix L, D'Aquila R, Holder D, Phillips A, Ait-Khaled M, Baxter J, Clevenbergh P, Hammer S, Harrigan R, Katzenstein D, Lanier R, Miller M, Para M, Yerly S, Zolopa A, Murray J, Patick A, Miller V, Castillo S, Pedneault L & Mellors J. The relation between baseline HIV drug resistance and response to antiretroviral therapy: re-analysis of retrospective and prospective studies using a standardized data analysis plan. *Antiviral Therapy* 2000; **5:**41–48.
29. Blower SM, Aschenbach AN, Gershengorn HB & Kahn JO. Predicting the unpredictable: transmission of drug-resistant HIV. *Nature Medicine* 2001; **7:**1016–1020.

30. Wahl LM & Nowak MA. Adherence and drug resistance: predictions for therapy outcome. *Proceedings of the Royal Society of London* 2000; **267:**835–843.

31. Bates M, Wrin T, Huang W, Petropoulos CJ & Hellmann N. Practical applications of viral fitness in clinical practice. *Current Opinion on Infectious Diseases* 2003; **16:**11–18.

Received 12 October 2003, accepted 5 January 2004