



Methods for optimizing antiviral combination therapies

Niko Beerenwinkel^{1,*}, Thomas Lengauer¹, Martin Däumer², Rolf Kaiser², Hauke Walter³, Klaus Korn³, Daniel Hoffmann⁴ and Joachim Selbig⁵

¹Max Planck Institute for Informatics, Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany, ²Institute of Virology, University of Cologne, Fürst-Pückler-Str. 56, 50935 Köln, Germany, ³Institute of Clinical and Molecular Virology, German National Reference Center for Retroviruses, University of Erlangen–Nürnberg, Schloßgarten 4, 91054 Erlangen, Germany, ⁴Center of Advanced European Studies and Research, Friedensplatz 16, 53111 Bonn, Germany and ⁵Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476 Golm, Germany

Received on January 6, 2003; accepted on February 20, 2003

ABSTRACT

Motivation: Despite some progress with antiretroviral combination therapies, therapeutic success in the management of HIV-infected patients is limited. The evolution of drug-resistant genetic variants in response to therapy plays a key role in treatment failure and finding a new potent drug combination after therapy failure is considered challenging.

Results: To estimate the activity of a drug combination against a particular viral strain, we develop a scoring function whose independent variables describe a set of antiviral agents and viral DNA sequences coding for the molecular targets of the respective drugs. The construction of this activity score involves (1) predicting phenotypic drug resistance from genotypes for each drug individually, (2) probabilistic modeling of predicted resistance values and integration into a score for drug combinations, and (3) searching through the mutational neighborhood of the considered strain in order to estimate activity on nearby mutants. For a clinical data set, we determine the optimal search depth and show that the scoring scheme is predictive of therapeutic outcome. Properties of the activity score and applications are discussed.

Contact: beerenwinkel@mpi-sb.mpg.de

keywords: HIV, antiretroviral therapy, drug resistance, SVM regression, therapy optimization, sequence space search.

INTRODUCTION

Infection with human immunodeficiency virus type 1 (HIV-1) is treated with combinations of drugs from a set of 16 currently approved antiretroviral agents (Jordan *et*

al., 2002, Table 1). Each of these drugs targets one of the two viral enzymes protease (PRO) or reverse transcriptase (RT) and there are three distinct drug classes: protease inhibitors (PI) binding to the protease active site, nucleoside RT inhibitors (NRTI) acting as chain terminating substrates during reverse transcription and nonnucleoside RT inhibitors (NNRTI) that directly bind to the RT molecule. Further compounds are under development, including fusion inhibitors (FI) that block viral cell entry by targeting the HIV envelope protein gp41 (Zollner *et al.*, 2001).

Despite the introduction of highly active antiretroviral therapy (HAART), a combination therapy consisting of three to six different inhibitors from at least two different drug classes, it is still impossible to eradicate the virus from the patients' bodies. Therefore, current treatment strategies aim at maximal suppression of virus load levels (the number of free virus particles in the blood plasma) over long time periods.

Aside from inducing strong side effects, the long-term effectiveness of HAART is also limited by the evolution of drug-resistant variants. Even in patients with viral load levels suppressed below detectable limits (50 copies/ml), ongoing viral replication can be found in a variety of tissues and cell types. Persistent virus production is further facilitated by sub-inhibitory drug levels in infected cells or by host immune failure. Thus, preexisting or newly produced drug resistant mutants can emerge that have a selective advantage under drug pressure. These escape mutants become dominant in the virus population and lead to viral rebound and therapy failure.

The genetic basis of drug resistance is HIV's high mutation rate (estimated about 3×10^{-5} per nucleoside per replication) due to lack of a proof-reading mechanism to-

*To whom correspondence should be addressed.

gether with its high replication rate. Many polymorphisms in the viral genome have been linked to drug resistance. Thus, genetic information can guide the process of selecting antiviral therapies.

Resistance testing

In order to avoid the administration of inactive compounds to the patient, resistance testing has become an important diagnostic tool in clinical practice (Perrin and Telenti, 1998; Vandamme *et al.*, 1999; DeGruttola *et al.*, 2000).

Phenotypic resistance testing measures *in vitro* viral replication of a wild type virus and of the virus isolated from a clinical sample in the presence of increasing drug concentrations (Walter *et al.*, 1999). The resistance factor, defined as the quotient of concentrations needed to inhibit replication of the virus extracted from the clinical sample by 50% (IC_{50}) and the IC_{50} value of the standardized wild type virus,

$$RF := \frac{IC_{50}(\text{clinical sample})}{IC_{50}(\text{wild type})},$$

reports the level of resistance as the fold-change in susceptibility to the drug as compared to a fully susceptible wild type.

Genotypic resistance testing is done by scanning the viral genome for resistance-associated mutations. It is faster and cheaper than phenotypic testing, but results are harder to interpret (Beerenwinkel *et al.*, 2001). Direct sequencing produces genomic data of about 1200 base pairs of the HIV pol gene, which codes for protease and RT. This sequence carries the information on susceptibility or resistance of the patients' virus to each of the drugs. However, it is challenging to retrieve this information from the sequence, because many mutations at different sites can be involved, the effect of a mutation can depend on the presence or absence of other mutations and many polymorphisms confer resistance to more than one drug (cross-resistance).

Using support vector machine (SVM) regression we will derive models that can predict phenotypic drug resistance from genotypic data.

Therapy optimization

A major problem in the management of HIV-infected patients is the selection of a new active regimen after failure of a drug combination, because remaining treatment options are reduced due to accumulated resistance mutations. We approach this problem by constructing a scoring function that estimates the activity of a drug combination against a given viral strain. We will derive this activity score from phenotype predictions of the drugs making up the therapy.

In order to predict long-term viral response we take into account not only the activity of a therapy against the

presently dominating strain, but also its activity against nearby mutants. We use a heuristic search strategy to explore the mutational neighborhood and estimate the activity of a worst case mutant from it.

Finally, we show that our scoring function is predictive of viral response by comparing it to observed virus load changes extracted from a clinical database.

RELATED WORK

Phenotype prediction

The problem of predicting phenotypic drug resistance from the genotype has been dealt with mainly in the form of a classification problem. After defining certain phenotypic cutoff values, sequences are classified into two or more classes ranging from 'susceptible' to 'resistant'.

There is a substantial body of literature linking genetic variations in protease and RT to drug resistance (Shafer *et al.*, 2000). These data result from site directed mutagenesis experiments, from observing the emergence of genetic changes under continuous drug pressure in cell culture, or from clinical samples derived from patients under monotherapy[†]. Several virological expert groups have extracted sets of classification rules from these published data.

Statistical and machine learning methods have been applied to deriving classification models from data sets of matched genotype-phenotype pairs and to identify relevant sequence positions. Successful approaches comprise decision trees (Sevin *et al.*, 2000; Beerenwinkel *et al.*, 2002a) and SVMs (Beerenwinkel *et al.*, 2001). Artificial neural networks have been used to make quantitative phenotype predictions (Wang *et al.*, 2000).

Unlike rule-based systems, data-driven approaches to phenotype prediction are free of publication bias. In particular, they perform better for new drugs for which good results can already be obtained from a moderate number of phenotypic tests, whereas the accumulation of published insights takes much longer (Schmidt *et al.*, 2001).

Knowledge-based therapy optimization

Prediction of clinical outcomes from genotypes is currently dominated by knowledge-based approaches, mainly because data sets of sufficient size for directly learning such relationships are rare. Two computational approaches go beyond the plain application of rules provided by experts.

The CTSHIV (Customized Treatment Strategy for HIV) system is a rule-based expert system designed for finding optimal resistance-avoiding combination therapies (Lathrop *et al.*, 1999b). The system operates on a set of resistance-inferring rules which are applied to a patient's

[†] treatment with just one drug

viral strains and nearby mutants. Drug combinations are scored by identifying the most resistant mutant and the least-resisted drug for this mutant in the respective drug combination. Nearby mutants are generated by a backward-chaining search. The optimal solution is computed by a branch and bound algorithm (Lathrop and Pazzani, 1999a).

Another approach applies fuzzy logic methods to a set of expert rules (De Luca *et al.*, 2002). Rule weights are learned from known clinical outcomes. The resulting system has been shown to improve over the set of rules alone.

SVM REGRESSION MODELS

For a given genomic sequence coding for one of the viral enzymes, protease or RT, and a given drug targeting this molecule, we want to predict the level of phenotypic resistance that the virus exhibits against the drug. For this purpose, we generate SVM regression models from a set of 650 genotype-phenotype pairs derived from patient samples (Beerenwinkel *et al.*, 2001, 2002b).

Let D be the set of antiretroviral drugs. For each drug $d \in D$ denote by $t(d)$ the target molecule of d and by \mathcal{S}_t the set of all sequences coding for a drug target t ($t \in \{\text{PRO}, \text{RT}\}$). We want to learn the resistance function

$$R_d : \begin{array}{l} \mathcal{S}_t \longrightarrow \mathbb{R} \\ s \longmapsto \text{RF of enzyme encoded by } s \text{ to drug } d. \end{array}$$

In order to perform SVM learning we map sequences into a Euclidean vector space X by introducing 20 indicator variables for each amino acid position of the multiple sequence alignment. Unlike other approaches that pre-select certain sequence positions, we consider full protease sequences of length 99 aa and RT sequences up to position 250. Thus, our models can capture resistance phenomena that are not linked to a few prominent sequence positions, but depend on the overall sequence background. The dimension of X is 1980 for protease and 5002 for RT.

While sequence data are well reproducible, there is considerable noise in the experimental determination of the resistance factor, with coefficients of variation ranging from 10 to 60 percent (Walter *et al.*, 1999). The SVM learning strategy (Burges, 1998) is suitable for this type of high-dimensional noisy data. Briefly, SVM learning rests on

- minimizing an upper bound on the generalization error derived from statistical learning theory (that does not depend on the dimension of X) for linear learning functions
- introducing non-linearity by implicitly mapping inputs into a high-dimensional feature space via kernel functions

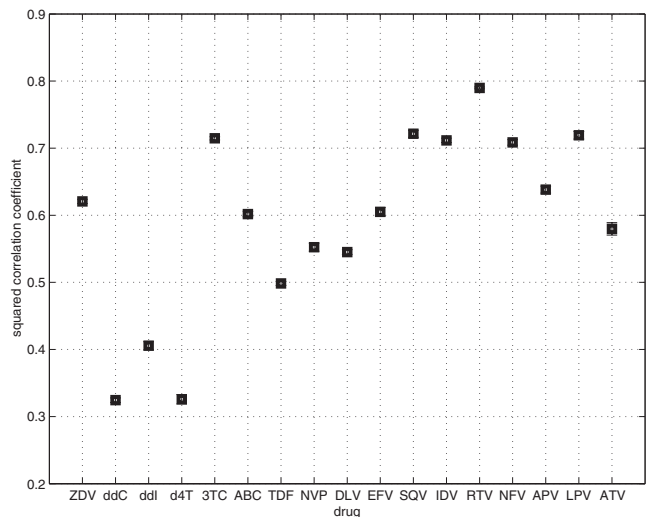


Fig. 1. Performance of SVM regression models. Squared correlation coefficients between observed and predicted resistance factors have been estimated from 10-fold cross-validation. Drug names are encoded by their abbreviations as displayed in Table 1.

- efficiently solving the resulting optimization problem, which is a quadratic program.

Figure 1 summarizes the ability of the linear models to generalize from the training data, as estimated by cross-validation. Most models can explain 50% to 80% of phenotypic variance given the genotype, with the exception of the three nucleoside analogs zalcitabine (ddC), didanosine (ddI) and stavudine (d4T) with squared correlation coefficients between 0.3 and 0.4. As standard non-linear kernels (polynomial, RBF) did not substantially improve these results (data not shown), and because linearity allows fast evaluations of the regression function R_d , we use these models later for the evaluation on clinical data. However, the algorithm presented in the next section will not be restricted to linear functions and allows for plugging in any non-linear regression function.

SCORING FUNCTION

We want to integrate single predictions of phenotypic resistance for each drug of a combination therapy into a summary activity score for this therapy.

Probabilistic model

Resistance factors scale differently across drugs. In fact, the maximum level of resistance can differ by more than an order of magnitude even within the same drug class. Furthermore, resistance factors from untreated patients are centered around one, but with largely varying dispersion. Nevertheless, the frequency distributions of resistance

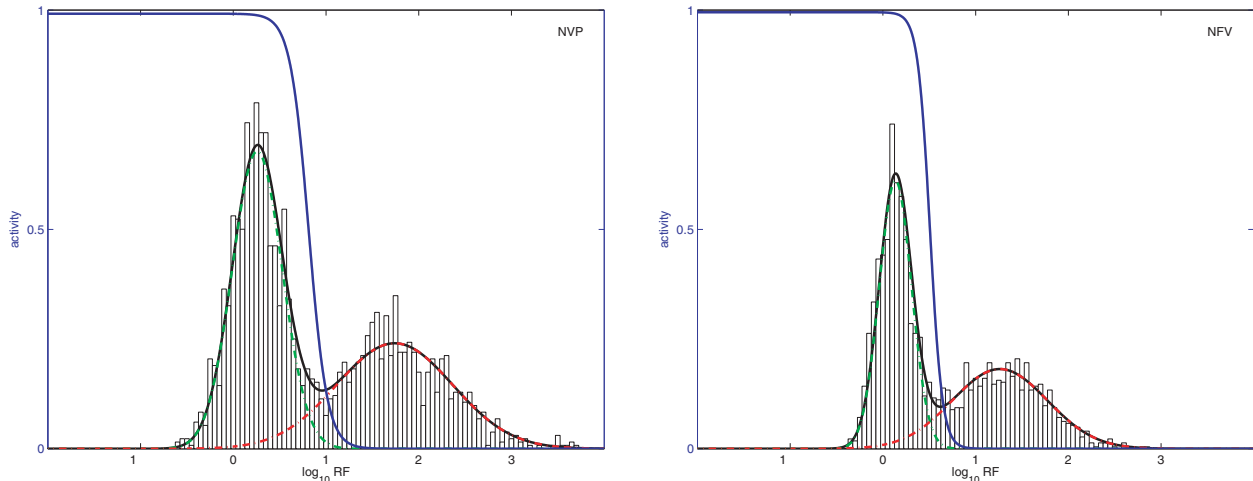


Fig. 2. Frequency distributions of predicted resistance factors for nevirapine (NVP, left) and nelfinavir (NFV, right). Solid lines represent the Gaussian mixture model that was fitted to histogram data. Dashed lines show single Gaussian component densities of each model. The solid logistic curve depicts activity scores.

Table 1. Antiretroviral agents. (See main text for abbreviations)

| Compound | Abbreviation | Target | Class | used since |
|-------------|--------------|--------|-------|------------|
| zidovudine | ZDV | RT | NRTI | 1987 |
| didanosine | ddI | RT | NRTI | 1991 |
| zalcitabine | ddC | RT | NRTI | 1992 |
| stavudine | d4T | RT | NRTI | 1994 |
| lamivudine | 3TC | RT | NRTI | 1995 |
| abacavir | ABC | RT | NRTI | 1999 |
| tenofovir | TDF | RT | NRTI | 2001 |
| nevirapine | NVP | RT | NNRTI | 1996 |
| delavirdine | DLV | RT | NNRTI | 1997 |
| efavirenz | EFV | RT | NNRTI | 1998 |
| saquinavir | SQV | PRO | PI | 1995 |
| indinavir | IDV | PRO | PI | 1996 |
| ritonavir | RTV | PRO | PI | 1996 |
| nelfinavir | NFV | PRO | PI | 1997 |
| amprenavir | APV | PRO | PI | 1999 |
| lopinavir | LPV | PRO | PI | 2000 |
| atazanavir | ATV | PRO | PI | — |
| T-20 | | gp41 | FI | — |
| T-1249 | | gp41 | FI | — |

factors for different drugs have common features. For most drugs they suggest a bimodal density function for the logarithms of resistance factors. Therefore, we assume a ‘two-state model’ for the virus comprising a *susceptible* and a *resistant* state and model the density of predicted resistance factors $x = \log_{10} \text{RF}$ with the Gaussian mixture

$$\alpha \phi(x; \mu_1, \sigma_1) + (1 - \alpha) \phi(x; \mu_2, \sigma_2),$$

where the first and second Gaussian account for the susceptible and resistant subsets, respectively (Fig. 2). Parameters are estimated on 2000 resistance factors predicted from genotypes obtained from routine clinical resistance testing by applying the EM algorithm (Dempster, 1977).

Most drugs clearly support the two-state model. It is also consistent with the concept of a strong selective pressure acting on the viral population under therapy. Since intermediate resistance states are observed less frequently, these states might be disadvantageous for the virus compared to the fully susceptible state in the untreated patient and to the fully resistant state in the treated patient.

To decide whether a given resistance factor is more likely to belong to the susceptible or the resistant subpopulation we consider the log-likelihood ratio

$$\ell(x) := \log \frac{\text{prob}(\text{sus}|x)}{\text{prob}(\text{res}|x)}.$$

With Bayes’ formula we see that

$$\begin{aligned} \ell(x) &= \log \frac{\text{prob}(x|\text{sus}) \frac{\text{prob}(\text{sus})}{\text{prob}(x)}}{\text{prob}(x|\text{res}) \frac{\text{prob}(\text{res})}{\text{prob}(x)}} \\ &= \log \frac{\phi(x; \mu_1, \sigma_1)}{\phi(x; \mu_2, \sigma_2)} + \log \frac{\text{prob}(\text{sus})}{\text{prob}(\text{res})}. \end{aligned}$$

If $\sigma_1 \neq \sigma_2$, then $\ell(x)$ is a quadratic function with a unique zero at x_0 between μ_1 and μ_2 (provided that $\mu_1 \neq \mu_2$). In particular, if, say $\sigma_1 < \sigma_2$ as in Figure 2, very small (but still observable) resistance factors will

be classified as resistant (because $\ell(x) \rightarrow -\infty$ as $x \rightarrow -\infty$). Since this behavior is biologically unreasonable, we approximate $\ell(x)$ with a monotonic function. Rather than solving the linear regression problem over the given data (as in classical logistic regression) we approximate $\ell(x)$ with its tangent $L(x)$ at x_0 . This linear function behaves like the log-likelihood ratio around x_0 and thus captures its optimal decision behavior in the interesting region between μ_1 and μ_2 that delimits the transition from susceptibility to resistance.

Thus, for each drug d and viral genotype $s \in \mathcal{S}_t(d)$ we first predict the resistance phenotype $x = \log_{10} R_d(s)$ and then define the activity of d against s as the logistic function of $L(x)$,

$$\text{activity}(d, s) := \frac{1}{1 + \exp(-L(x))}.$$

Since $L(x)$ approximates the log-likelihood ratio $\ell(x)$, we have

$$\begin{aligned} \text{activity}(d, s) &\approx \text{prob}(\text{sus}|x) \\ &= \text{prob}(\text{sus} | \log_{10} R_d(s)), \end{aligned}$$

the conditional class probability of s belonging to a susceptible strain.

We identify a virus with the set of sequences coding for its genes and, by slight abuse of notation, extend this definition to viral strains. Formally, let $\text{seq}_t(v)$ be the DNA sequence coding for drug target t of virus v . Then

$$\text{activity}(d, v) := \text{activity}(d, \text{seq}_{t(d)}(v))$$

is the estimated activity of d against v .

Scoring drug combinations

Since phenotypic resistance testing is not performed for drug combinations we cannot directly learn combined effects from data. Thus, we resort to a model assumption based on elementary observations from clinical studies (Jordan *et al.*, 2002).

Drug combinations with drugs from different drug classes benefit from synergistic effects of two or more inhibitors and thus our scoring scheme will be additive between drug classes. By comparison, combinations restricted to a single drug class are generally less potent, because inhibitors with the same drug target and mechanism of action are competing. Thus, we estimate activity within a drug class as the activity of the most active drug.

More precisely, for a drug combination

$$T = \{d_1, \dots, d_n\} \subset D$$

denote by T_c all drugs from T belonging to the same drug class c . Then we define the activity of the combination

therapy T against v as

$$\text{activity}(T, v) := \sum_c \max_{d \in T_c} \text{activity}(d, v),$$

which naturally extends our first definition of activity for a single drug. In particular, if $\text{activity}(d, v) \approx 0$ or 1 for all $d \in T$, then $\text{activity}(T, v)$ is just the number of active drug classes. Note that the activity score is monotonic in the number of drugs. Indeed, adding a further inhibitor cannot decrease the activity of a combination therapy against a single strain.

Maximization over activities within drug classes seems to suggest that combining drugs from the same class is ineffective. However, this is not the case with the extended scoring scheme that is introduced in the following subsection. The within class effect of drug combinations will be illustrated in more detail in the discussion.

Sequence space search

Long-term success of an antiretroviral therapy will not only depend on the current resistance phenotype of the virus, but also on its ability to escape from the selective pressure exerted by the drug combination. Resistant escape variants may be produced by erroneous replication under therapy or may be preexisting in the viral population. The structure of this population is generally thought of as a quasi-species, a dominating strain (or master sequence) surrounded by a large variety of closely related mutants (Holland *et al.*, 1992). In practice, we do not know the distribution of mutants in the quasi-species, but are only given the dominating strain (or a mixture of dominating strains accounting for $\geq 30\%$). Thus, as an estimate of how easily the virus can evade drug pressure, we predict activity against a worst case mutant in different mutational neighborhoods of the given sequence. We do not make any assumption about the preexistence of mutants.

We fix a drug target t in this subsection. For a given sequence $s_0 \in \mathcal{S}_t$ let $B_r(s_0)$ be the mutational neighborhood of s_0 at distance r , i.e. the set of all sequences with Hamming distance to s_0 less than or equal to r . A worst case mutant for a fixed combination therapy T is characterized by attaining the minimum

$$\min_{s \in B_r(s_0)} \text{activity}(T_t, s),$$

where $T_t \subset T$ is the subset of drugs with target t .

Since sequence space is so enormous, exhaustive searches are practical only for very restricted neighborhoods. Instead, we use a heuristic search strategy called beam search. We score each visited mutant with the predicted activity. After visiting all one-point mutants of s_0 we maintain only a few of least activity and generate all point mutations of these mutants. Proceeding in this

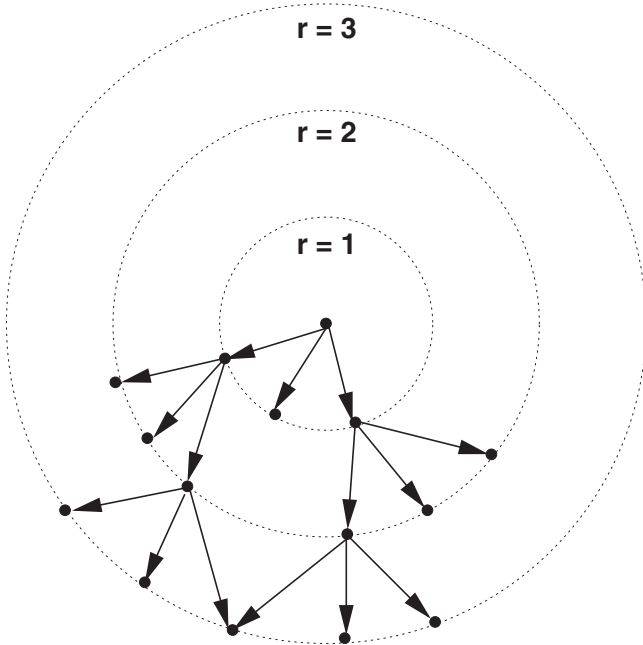


Fig. 3. Exemplified sequence space search in breadth $b = 2$ up to depth $d = 3$. The root of the directed graph represents a query sequence s_0 . An edge (s, s') between two sequences indicates that s' was generated from s by a point substitution in s .

fashion we follow only a fixed number of mutants, say b , at each level (Fig. 3). The search strategy is summarized in the following algorithm.

Input

search depth d , breadth b ,
 sequence $s_0 \in \mathcal{S}_t$,
 set of drugs T with the same target t .

Output

activity scores $a(r)$ for all $r = 0, \dots, d$.

construct priority queues Q, E
 ordered by increasing activity

pq-initialize Q with s_0
 pq-initialize E with \emptyset

set $a(0) := \text{activity}(T, s_0)$

set $r := 1$

while $r \leq d$ and $a(r-1) > 0$

$t_0 := \text{pq-remove}(Q)$
 set $a(r) := \text{activity}(T, t_0)$

set $j := 1$

while $j \leq b$ and $Q \neq \emptyset$

$t_j := \text{pq-remove}(Q)$
 set $j := j + 1$

```

end while
pq-vacate(Q)
set j := 1
while j ≤ b and Q ≠ ∅
    pq-vacate(E)
    expand: pq-insert(E, m) for all
        one-point mutants m of tj
    repeat b times
        pq-insert(Q, pq-remove(E))
    end repeat
    set j := j + 1
end while
set r := r + 1
end while
    
```

Thus, searching the neighborhood of a sequence of length l in breadth b ($l \gg b$) up to depth d is of time complexity $\mathcal{O}(d b l \log l)$.

For a fixed drug combination and search breadth we denote by $a_t(r)$ the search result for drug target t in depth r . Thus, $a_t(r)$ is the estimated activity against drug target t after r point mutations.

The search strategy outlined here is applicable with any (non-linear) model of phenotype prediction. In general, it is not guaranteed to find a global minimum.

Merging search results

Searching is performed for each drug target separately. This is possible because the activity scoring function is additive between drug classes and thus between drug targets. In order to come up with a score for a therapy comprising drugs with different targets, we follow the same greedy strategy as above to merge search results $a_t(r)$ for all t . Presently, $t \in \{\text{PRO}, \text{RT}\}$ and we illustrate the procedure for this case.

At each step we choose between a further substitution in protease or RT by comparing activity scores. Starting with a score of

$$A_0 := a_{\text{PRO}}(0) + a_{\text{RT}}(0) = \text{activity}(T, s_0)$$

we can calculate the worst-case activity for a one-point mutant as the minimum of scores obtained by introducing the mutation in either protease or RT,

$$A_1 := \min \{a_{\text{PRO}}(1) + a_{\text{RT}}(0), a_{\text{PRO}}(0) + a_{\text{RT}}(1)\}.$$

We continue to search from the mutant attaining the minimum and again decide for the second point mutation in protease or RT depending on where the activity is reduced most. Thus, in a greedy fashion, we follow only

the mutant against which the drug combination retains the least activity. The following algorithm implements this strategy:

Input

search depth d ,
 activity scores $a_t(r)$ for all drug targets $t \in \{\text{PRO,RT}\}$,
 and all search depth $r = 0, \dots, d$

Output

scores A_r for all $r = 0, \dots, d$.

set $A_0 := h_{00} := a_{\text{PRO}}(0) + a_{\text{RT}}(0)$
 set $i := 0, j := 0$

```

while  $i + j < d$ 
  set  $h_{i+1,j} := a_{\text{PRO}}(i + 1) + a_{\text{RT}}(j)$ 
  set  $h_{i,j+1} := a_{\text{PRO}}(i) + a_{\text{RT}}(j + 1)$ 
  if  $h_{i+1,j} < h_{i,j+1}$ 
    set  $i := i + 1$ 
  else
    set  $j := j + 1$ 
  end if
  set  $A_{i+j} := h_{ij}$ 
end while
    
```

The procedure can be viewed as a directed walk through the matrix

$$H = (h)_{ij} = (a_{\text{PRO}}(i) + a_{\text{RT}}(j))_{ij}.$$

In general, merging search results of depth d from m different drug targets is of time complexity $\mathcal{O}(md)$.

We will refer to the resulting scores A_r as

$$\text{activity}_r(T, v),$$

the estimated activity of T against v after r point mutations.

EVALUATION ON CLINICAL DATA

In order to test whether the constructed scoring function can predict treatment response in a clinical setting, we compare estimated activities to observed virus load changes. We extracted from a clinical database patient data documenting the success of a therapy change that was accompanied by a genotypic resistance test. More precisely, a drug combination was included in the analysis if genotype and virus load were determined before the beginning of the therapy and a virus load follow-up value was available at a time point $90(\pm 10)$ days later. A therapy was considered successful if virus load decreased by 2 or more \log_{10} and failure otherwise.

For a group of 96 patients (28 successful therapies vs. 68 failures) matching these criteria, activity scores were

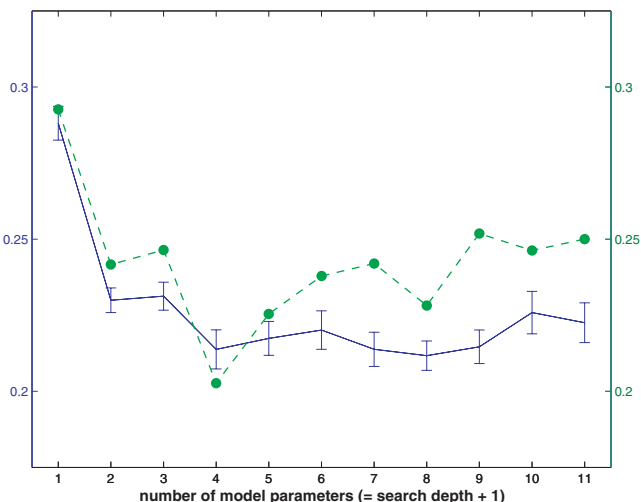


Fig. 4. Determining the optimal search depth. The expected prediction error as estimated by 12-fold cross-validation (solid line, with one-standard-error bars) and AIC (dashed line) is displayed for all models up to search depth 10.

computed up to search depth ten at a fixed search breadth of $b = 10$. To classify therapies as either *successful* or *failure* we performed linear discriminant analysis on the estimated activities.

Model selection

In order to find the optimal search depth for the given classification problem we constructed linear decision models with the $d + 1$ predictive variables

$$\text{activity}_r(T, v), \quad r = 0, \dots, d,$$

for all $d = 0, \dots, 10$. This procedure gives rise to eleven models, each incorporating all search results up to depth d . To find the best among these ordered feature subsets, the expected prediction error of each model was estimated by 12-fold cross-validation and by the Akaike information criterion (AIC) with 0–1 loss (Hastie et al., 2001) (Fig. 4).

Both the ‘one-standard-error rule’ (picking the most parsimonious model within one standard error of the minimum) and AIC argue for a search depth of three point mutations in this case. In particular, both estimates improve considerably at search depth three over depth zero stressing the utility of searching the mutational neighborhood. The best model predicts therapy success from viral genotype with an expected error rate of 21.4% (cross-validation estimate). Out of 96 cross-validation test cases we predict 30 successes (positives) and 66 failures (negatives) including 12 false positives and 10 false negatives, respectively. Thus, predictions are not independent of true class labels ($p < 0.00002$, Fisher’s exact test).

In general, the optimal search depth will depend on the definition of therapeutic success and possibly on characteristics of the patient group. In contrast, variations of the search breadth did not qualitatively alter these results.

DISCUSSION

We have constructed a scoring function that estimates the activity of a combination therapy against a viral strain. Drug resistance is only one of several factors that determine therapeutic outcome, but it plays a key role in treatment failure. Thus, we propose a method for selecting optimal drug combinations with respect to resistance based on viral genomic data. The high genomic variability of HIV strains from different patients forces clinicians to evaluate treatment options individually. The computational approach presented promises to be helpful in balancing an increasing number of drug combinations against the background of complex mutational patterns.

The construction of the activity score was driven by the observation that the viral population will change when exposed to the selective pressure of a new drug combination. This change will be effected by the accumulation of new advantageous escape mutations. We took this phenomenon into account by estimating the minimal activity of any mutant that differs from the strain under consideration by a fixed number of point mutations. This approach captures a fundamental property of combination therapies.

Within-class synergy

The use of more than one drug at the same time can delay the development of drug resistance. For drugs with different drug targets this synergy is reflected in the additivity of the activity score between drug classes. For drugs targeting the same molecule, there is also a synergistic effect albeit of a different kind.

As an example we consider the NRTI zidovudine (ZDV) and lamivudine (3TC). Both drugs act as chain terminators, but HIV develops different mutations under ZDV and 3TC mono-therapy. Therefore, ZDV + 3TC is a frequent drug combination in clinical practice. Figure 5 compares activity scores at different search depths for the two mono-therapies and the double therapy on a susceptible strain. Resistance to 3TC can be accomplished by a single point mutation that leads to the amino acid change M184V in the RT. High level resistance to ZDV can be attained with two different point mutations the first reducing activity to less than 50%.

Under double therapy the situation is different: As expected, the search process failed to identify a single point mutation that significantly reduces the activity of both inhibitors at the same time. The estimated activity is almost 80%. Overall, the activity curve for the double

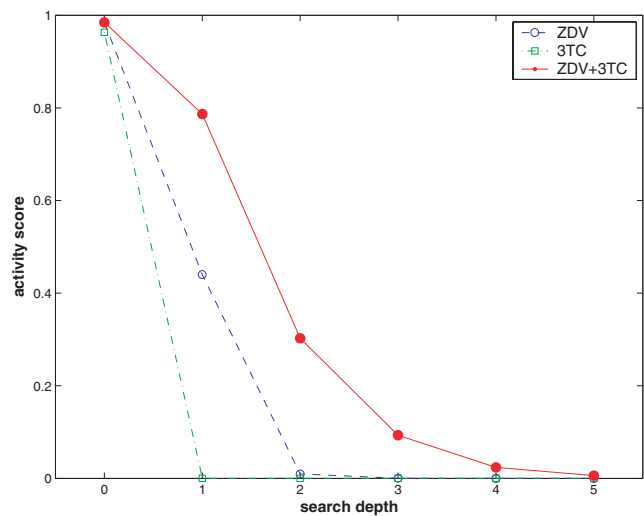


Fig. 5. Activity scores for ZDV (dashed line), 3TC (dash-dot line) and ZDV+3TC (solid line) at search depth 0, . . . , 5 on a susceptible strain. Drugs with orthogonal resistance profiles can delay the development of drug resistance by increasing the genetic barrier for the virus.

therapy is shifted to the right. It is the maximum of the two mono-therapy curves only before the introduction of point mutations. Thus, the double therapy is estimated to maintain activity over a longer time period (or more likely to be active against a broad quasi-species). Sequence space search reveals this increased genetic barrier for the virus. In general, the scoring function will favor combinations of drugs with orthogonal genetic resistance profiles relative to the given viral strain.

Another consequence of the search process is that the overall scoring function is (in principle) no longer monotonic in the number of drugs, as was the case without searching. It may happen that adding a drug results in an estimated mutational path that leads to complete resistance earlier than without that drug. This may be the case for adding a drug with an intermediate resistance profile of two other drugs to the double therapy comprising these two drugs. The new drug could ‘guide’ the sequence space search and make it more effective. However, this issue is more relevant if we try to optimize a sequence of therapies. Using up therapeutic options too lavishly will then be penalized.

Improving activity estimation

Although our scoring scheme appears to capture key properties of combination therapies and viral evolution under such drug pressures, it may benefit from further development. One simplification of our approach is that we represent the intra-patient viral population by a single

strain, the dominating virus. We meet here practical needs, because current routine genotyping is based on population sequencing and only detects a mixture of those variants that represent at least 30% of the virus population. However, new experimental and computational techniques may allow for determining more accurate viral population frequencies in the future (Wildenberg *et al.*, 2002).

Dealing with more than one sequence raises another issue, namely that of recombination. The recombination rate of HIV-1 has been estimated to be about tenfold greater than its point substitution rate (Jetzt *et al.*, 2000; Jung *et al.*, 2002). Since our strategy for searching the mutational neighborhood of a sequence is based on generating point mutations, this process may overestimate the distance of genetic escape variants.

Decision support

The proposed method intends to support clinical decision making. Indeed, the previous section has revealed that the management of HIV infected patients can benefit from exploiting the information encoded in the viral genome. For a given genotype, the scoring function provides a means of ranking combination therapies by their success probabilities. Thus, activity scores can help in the individualized design of therapeutic protocols.

ACKNOWLEDGEMENTS

We thank Jörg Rahnenführer for helpful discussions.

Patient data on clinical outcomes were kindly provided by Jürgen Rockstroh (Medizinische Klinik und Poliklinik I, Universitätsklinikum Bonn, Germany), Gerd Fätkenheuer (Klinik I für Innere Medizin, Universität zu Köln, Germany), Mark Oette (Klinik für Gastroenterologie, Hepatologie und Infektiologie, Heinrich-Heine-Universität Düsseldorf, Germany) and Thomas Berg (Medizinisches Labor Dr. Berg, Berlin, Germany).

Part of this research has been funded by Deutsche Forschungsgemeinschaft under Grant No. HO 1582/1-2.

REFERENCES

- Beerenwinkel,N., Schmidt,B., Walter,H., Kaiser,R., Lengauer,T., Hoffmann,D., Korn,K. and Selbig,J. (2001) Geno2pheno: interpreting genotypic HIV drug resistance tests. *IEEE Intelligent Systems*, **16**, 35–41.
- Beerenwinkel,N., Schmidt,B., Walter,H., Kaiser,R., Lengauer,T., Hoffmann,D., Korn,K. and Selbig,J. (2002a) Diversity and complexity of HIV-1 drug resistance: a bioinformatics approach to predicting phenotype from genotype. *Proc. Natl Acad. Sci. USA*, **99**, 8271–8276.
- Beerenwinkel,N., Schmidt,B., Walter,H., Kaiser,R., Lengauer,T., Hoffmann,D., Korn,K. and Selbig,J. (2002b) Quantitative phenotype prediction by support vector machines. *Antivir. Ther.*, **7**, S74.
- Beerenwinkel,N., Däumer,M., Sierra-Aragon,S., Schmidt,B., Walter,H., Korn,K., Oette,M., Fätkenheuer,G., Rockstroh,J., Lengauer,T. *et al.* (2002) Geno2pheno is predictive of short-term virological response. *Antivir. Ther.*, **7**, S74.
- Burges,C.J.C. (1998) A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, **2**, 121–167.
- DeGruttola,V., Dix,L., D'Aquila,R., Holder,D., Phillips,A., Ait-Khaled,M., Baxter,J., Clevenbergh,P., Hammer,S. and Harrigan,R., *et al.* (2000) 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. *Antivir. Ther.*, **5**, 41–48.
- De Luca,A., Vendittoli,M., Baldini,F., Di Giambenedetto,S., Rizzo,M.G., Trotta,M.P., Cingolani,A., Forbici,F., Perno,C.F., Antinori,A. and Ulivi,G. (2002) Construction, training and clinical validation of an inferential interpretation system for genotypic HIV-1 drug resistance based on fuzzy rules learning from virological outcomes. *Antivir. Ther.*, **7**, S76.
- Dempster,A., Laird,N. and Rubin,D. (1977) Maximum likelihood from incomplete data via the EM algorithm (with discussions). *J. R. Statist. Soc. B*, **39**, 1–38.
- Hastie,T., Tibshirani,T. and Friedman,J. (2001) *The Elements of Statistical Learning*. Springer, New York.
- Holland,J.J., De La Torre,J.C. and Steinhauer,D.A. (1992) RNA virus populations as quasispecies. *Curr. Top. Microbiol. Immunol.*, **176**, 1–20.
- Jetzt,A.E., Yu,H., Klarmann,G.J., Ron,Y., Preston,B.D. and Dougherty,J.P. (2000) High rate of recombination throughout the human immunodeficiency virus type 1 genome. *J. Virol.*, **74**, 1234–1240.
- Jordan,R., Gold,L., Cummins,C. and Hyde,C. (2002) Systematic review and meta-analysis of evidence for increasing numbers of drugs in antiretroviral combination therapy. *British Med. J.*, **324**, 1–10.
- Jung,A., Maier,R., Vartanian,J.P., Bocharov,G., Jung,V., Fischer,U., Meese,E., Wain-Hobson,S. and Meyerhans,A. (2002) Multiply infected spleen cells in HIV patients. *Nature*, **418**, 144.
- Lathrop,R.H. and Pazzani,M.J. (1999a) Combinatorial optimization in rapidly mutating drug-resistant viruses. *J. Comb. Optim.*, **3**, 301–320.
- Lathrop,R.H., Steffen,N.R., Raphael,M., Deeds-Rubin,S., Pazzani,M.J., Cimoch,P.J., See,D.M. and Tilles,J.G. (1999b) Knowledge-based avoidance of drug-resistant HIV mutants. *AI Magazine*, **20**, 13–25.
- Perrin,L. and Telenti,A. (1998) HIV treatment failure: testing for HIV resistance in clinical practice. *Science*, **280**, 1871–1873.
- Schmidt,B., Walter,H., Schwingel,E., Beerenwinkel,N., Selbig,J., Kaiser,R., Hoffmann,D., Shafer,R., Keulen,W., Been-Tiktak,A.M., Boucher,C.A. and Korn,K. (2001) Comparison of different interpretation systems for genotypic HIV-1 drug resistance data. *Antivir. Ther.*, **6**, 102.
- Sevin,A.D., DeGruttola,V., Nijhuis,M., Schapiro,J.M., Foulkes,A.S., Para,M.F. and Boucher,C.A. (2000) Methods for investigation of the relationship between drug-susceptibility phenotype and human immunodeficiency virus type 1 genotype with applications to AIDS clinical trials group 333. *J. Infect. Dis.*, **182**, 59–67.
- Shafer,R.W., Kantor,R. and Gonzales,M.J. (2000) The genetic

- basis of HIV-1 resistance to reverse transcriptase and protease inhibitors. *AIDS Reviews*, **2**, 211–228.
- Vandamme,A.M., Van Laethem,K. and De Clerq,E. (1999) Managing resistance to anti HIV drugs: an important consideration for effective disease management. *Drugs*, **57**, 337–361.
- Walter,H., Schmidt,B., Korn,K., Vandamme,A.M., Harrer,T. and Überla,K. (1999) Rapid, phenotypic HIV-1 drug sensitivity assay for protease and reverse transcriptase inhibitors. *J. Clin. Virol.*, **13**, 71–80.
- Wang,D., Bloor,S. and Larder,B.A. (2000) The application of neural networks in predicting phenotypic resistance from genotypes for HIV-1 protease inhibitors. *Antivir. Ther.*, **5**, 51–52.
- Wildenberg,A., Skiena,S. and Sumazin,P. (2002) Deconvolving sequence variation in mixed DNA populations. *Proceedings of the Sixth Annual International Conference on Computational Biology*. pp. 311–320.
- Zollner,B., Feucht,H.H., Schroter,M., Schafer,P., Plettenberg,A., Stoehr,A. and Laufs,R. (2001) Primary genotypic resistance of HIV-1 to the fusion inhibitor T-20 in long-term infected patients. *AIDS*, **15**, 935–936.