Vandamme 30/4/08 11:52 Page 1

PROOF COPY ONLY

Antiviral Therapy 13:xxxx–xxxx

Original article

Modified *in vivo* HIV fitness under drug selective pressure and estimated genetic barrier towards resistance are predictive for virological response

Koen Deforche[1], Alessandro Cozzi-Lepri[2], Kristof Theys[1], Buenaventura Clotet[3], Ricardo Camacho[4], Jesper Kjaer[5], Kristel Van Laethem[1], Andrew Phillips[2], Yves Moreau[1], Jens D Lundgren[5,6] and Anne-Mieke Vandamme[1] on behalf of the EuroSIDA Study Group*

[1]Katholieke Universiteit Leuven, Leuven, Belgium
[2]Royal Free and University College Medical School, London, UK
[3]University Hospital Germans Trias i Pujol, Barcelona, Spain
[4]Centro Hospitalar de Lisboa, Lisbon, Portugal
[5]University of Copenhagen, Copenhagen, Denmark
[6]Rigshospitalet, Copenhagen, Denmark

*Corresponding author: E-mail: annemie.vandamme@uz.kuleuven.be

Background: A method has been developed to estimate a fitness landscape experienced by HIV-1 under treatment selective pressure as a function of the genotypic sequence thereby also estimating the genetic barrier to resistance.
Methods: We evaluated the performance of two estimated fitness landscapes (nelfinavir [NFV] and zidovudine [AZT] plus lamivudine [3TC]) to predict week 12 viral load (VL) change for 176 treatment change episodes (TCEs) and probability of week 48 virological failure for 90 TCEs, in treatment experienced patients starting these drugs in combination.
Results: A higher genetic barrier for AZT plus 3TC, (quantified per additional mutation required to develop resistance against these drugs) was associated with a 0.54 (95% confidence interval [CI] 0.30–0.77) larger $\log_{10}$ VL reduction at 12 weeks (P<0.0001) and a 0.39 (95% CI 0.23–0.66) lower odds of virological failure at 48 weeks (P=0.0005), in analyses adjusting for the pre-TCE VL and the exact time-lag between the TCE and the date of determining response VL. The strength of these associations was comparable with those seen with expert interpretation systems (Rega, ANRS and HIVDB). A higher genetic barrier to NFV resistance was the only genotypic predictor that tended to be associated with a 0.19 (95% CI 0–0.39) higher $\log_{10}$ VL reduction at 12 weeks (P=0.05) and a 0.63 (95% CI 0.36–1.09) lower odds of virological failure at 48 weeks (P=0.10) per additional mutation.
Conclusions: These results suggest that an estimated genetic barrier derived from fitness landscapes may contribute to an improvement of predicted treatment outcome for NFV and this approach should be explored for other drugs.

AU: OK?

AU: OK?

AU: OK?

Introduction

HIV antiviral drugs interfere with viral proteins resulting in the inhibition of HIV replication. HIV frequently escapes the inhibition of these drugs by selection of drug resistance mutations, leading to treatment failure [1]. To combine drugs in an effective combination, treatment guidelines recommend the use of resistance testing to take into account the presence of resistance mutations [2]. Genotypic resistance assays, which are faster and less expensive, are usually preferred over phenotypic resistance assays, which measure the susceptibility of the virus *in vitro*.

How to best interpret genotypic resistance results, however, remains a challenge. Automated learning of a genotypic prediction system for single drugs directly from clinical outcome data is complicated because drugs are used in combination, the high variability of the HIV genome, the complex interactions between mutations and the many confounding factors including treatment adherence [3]. Thus, the most widely used genotypic interpretation systems (IS) either capture expert knowledge derived from the literature, are based on machine learning techniques that learn correlation between geno-

© 2008 International Medical Press 1359-6535

Vandamme 30/4/08 11:52 Page 2

PROOF COPY ONLY

type and *in vitro* resistance phenotype measurements or use statistically derived mutational scores [4]. The genetic barrier to resistance is the evolutionary distance for the virus to evolve a sufficient number of resistance mutations in its population to overcome the efficacy of the drugs used to suppress viral replication. The genetic barrier relates to how vulnerable the treatment is in the presence of residual replication or minority variants in the viral population. Genetic barriers are not captured by a phenotypic resistance assay [5]. Although there is evidence that antiretrovirals have different genetic barriers, for example, lower for non-nucleoside reverse transcriptase inhibitors (NNRTI) as compared with protease inhibitors (PI), it is still unclear what genetic factors contribute to variation in genetic barriers and to what extent variation in genetic barriers affect durable treatment success.

The evolution of the virus, to overcome drug inhibition, is driven by the selective advantage of mutations in the sequence neighbourhood that it can explore. This defines a fitness landscape and, in a previous work [6], we presented a method to derive this fitness landscape by estimating the selective advantage of single mutations and mutational patterns in agreement with how frequently these are observed in patients after treatment failure. Using computer simulation of HIV evolution from a baseline sequence over this estimated fitness landscape, together with a criterion for resistance, we were able to estimate for each sequence its genetic barrier to resistance to specific drugs or drug combinations.

Predicted susceptibility often varies according to the expert IS used and it has been shown that discordances are more frequent in HIV non-B subtypes, reflecting the fact that resistance in non-B subtypes is less well characterized [7]. The fitness landscapes were trained on sequences from various subtypes (33% non-B subtype), thereby effectively incorporating into the landscape subtype variation. In this study, we wished to evaluate the clinical utility of the genetic barrier concept derived from the estimated fitness landscapes under treatment with nelfinavir (NFV) and zidovudine (AZT) plus lamivudine (3TC) to predict virological outcome for patients starting a treatment with these drugs, compared with that of three expert genotypic IS. We evaluated the association with viral load (VL) reduction from pre-treatment levels at week 12 and the risk of virological failure at week 48.

Methods

Fitness landscapes

Fitness landscapes of HIV in the presence of NFV selective pressure as a function of the protease (PRO) sequence, and of AZT+3TC as a function of the reverse transcriptase (RT) sequence, were estimated from cross-sectional data using the method previously described [6]. The fitness functions were estimated on the basis of the difference in prevalence of mutations and mutational patterns in viral sequences from treatment-experienced patients compared with untreated patients. A selective advantage was associated with each mutational pattern to explain the different prevalence as the result of convergent evolution under treatment-selective pressure. For NFV, 7,774 sequences from PI-naive patients and 1,026 sequences from patients treated with NFV as single PI were used, which were of diverse subtypes (subtype B 66%, G 15%, C 7%, and other 13%). For AZT+3TC, 5,528 sequences from nucleoside reverse transcriptase inhibitor (NRTI)-naive patients and 1,420 sequences from patients who were treated with AZT+3TC as the only NRTIs were used (subtype B 55%, G 14%, C 8%, and other 24%). A list of the mutations included in the fitness functions are shown in the Supplementary material. Mutations in RT at positions involved in resistance to NNRTIs (28, 90, 98, 100, 101, 103, 106, 108, 179, 181, 188, 190, 221, 225, 227, 228 and 238) were excluded from the fitness landscape model for AZT+3TC. Next to single nucleotide mutations, double nucleotide mutations at position 215 in RT were considered, as common amino acid mutations selected by AZT require two nucleotide mutations at this residue.

Study population

From the EuroSIDA cohort database, 211 treatment change episodes (TCEs) were considered where patients changed to a treatment containing AZT, 3TC and NFV, when at least one of these was used as a new drug and a genotypic resistance test was performed any time before the date of TCE. The genotype before the TCE was used for evaluating genotypic susceptibility. HIV-1 subtypes were determined using the Rega HIV-1 Subtype Tool version 2.0 [8].

Two endpoints were considered. For 176 patients who had a VL measurement at week 12 (the most recent value falling between week 4 to 16), the change in VL compared with VL before TCE was considered as a first endpoint. Virological outcome at week 48 was considered as a second binary endpoint (a VL>400 copies/ml was defined as failure) in 90 of the 211 patients (42%) who had a VL measurement in the week 48 time window (the most recent value falling between week 16 to 48) and did not have a drug change before the week 48 VL. Thus, overall in the two analyses, out of the 211 originally considered TCEs, 186 (88%) were studied for whom an additional VL was measured in at least one of the two predefined time windows after the date of TCE, as defined above. Of these 186 patients (none of the patients contributed

© 2008 International Medical Press

Vandamme 30/4/08 11:52 Page 3

PROOF COPY ONLY

more than one TCE), 80 (43%) were included in both the week 12 and week 48 analyses, 96 (52%) solely in the week 12 analysis and 10 (5%) solely in the week 48 analysis. The 186 TCEs were not different from the total number of 211 TCEs extracted from the database with respect to age, gender, mode of HIV transmission and the other measured confounders (including pre-TCE VL and percentage with evidence of previous virological failure on 3TC; data not shown).

Genotypic predictions

The susceptibility of the virus to the treatment was estimated on the basis of the estimated fitness in the presence of a specific treatment (F_{NFV} and $F_{AZT+3TC}$). The genetic barrier to resistance was estimated by simulating evolution of the baseline sequence over the fitness landscapes for these drugs until considered 'resistant'. These simulations were repeated 100 times and the average 'time' was used as an estimate for a genetic barrier. As criteria for 'resistance', we considered the development of a major resistance mutation (MRM) or an increase in fitness under treatment that exceeds a fitness threshold, defined as the average fitness of the sequences after appearance of an MRM. For NFV, mutations 30N, 88S or 90M were considered as MRM [9]; for AZT+3TC, mutations 70R, 184V or 215Y/F [10] were considered as MRM. As measures of 'time' to an MRM or to high fitness, the number of mutations and the number of simulated generations were considered. For every baseline PRO and partial RT sequence, the two fitness landscapes were thus used to make five quantitative predictions: *MR* and *GR* (estimated number of mutations and number of generations, respectively, to an MRM); *MF* and *GF* (estimated number of mutations and number of generations, respectively, to a fitness above the chosen fitness threshold); and a parameter indicating the predicted viral fitness (in the sense as explained above, expressed in the $\log_{10}$ scale, which can be considered the quantification of genotypic susceptiblity, $\log_{10} F$).

The performance of these genotypic predictors was compared with that of genotypic susceptibility scores (GSS) computed using the latest versions of three publicly available genotypic resistance IS: Rega 7.1, HIVDB 4.2.9 and ANRS 2006.07. GSS were as defined by the algorithm (for Rega 7.1) or otherwise based on the three-level categorical predictions of these systems (susceptible (S), GSS=1.0, intermediate resistance (I) GSS=0.5 or resistance (R) GSS=0.0). For HIVDB, the scores were grouped as 1,2=S, 3,4=I and 5=R.

Statistical analyses

The genotypic estimates based on the fitness landscapes and the expert systems were evaluated for prediction of VL reduction at week 12 (using a linear regression model accounting for censored outcomes due to the lower limit of detection of the VL assay) or risk of virological failure at week 48 (logistic regression). For the week 12 analysis we included all patients with a week 12 VL and treatment changes were ignored. For the week 48 analysis we included only patients who had a week 48 VL and who did not change treatment before the date of the week 48 VL. The area under the curve from a receiver operating characteristic (ROC) analysis was used to estimate the discriminative power to predict virological failure at week 48 using a univariable logistic regression including only the genotypic predictors. For this analysis, the Rega IS was chosen a priori to be the reference rule-based system for comparison.

In addition, the prognostic value to predict viral response of the genotypic candidate predictors was also assessed. For the linear and logistic regression analyses, two models were fitted. A first model included, besides the genotypic predictors, only pre-TCE $\log_{10}$ VL and the exact time between the date of TCE start and the endpoint (model 1). In a second, less parsimonious model, we adjusted for a number of additional potential confounders, including the activity of the drugs started besides AZT, 3TC and NFV, past treatment history and others listed in Table 1 (model 2).

Concordance between predicted AZT, 3TC and NFV susceptibility obtained using the three IS was evaluated by calculating the percentage of matching predictions in a descriptive fashion.

Results

Description of study population

The majority of TCEs studied were recorded between 1998 and 1999, when treatment with AZT+3TC+NFV was common (Table 1). The majority of patients (89%) were infected with a HIV-1 subtype B strain.

The TCEs studied are from a patient population that tended to be heavily pretreated: of the 176 patients included in analyses for response at week 12, 148 patients (83%) had previous treatment experience with one or more NRTIs, 134 (76%) with one or more PIs and 50 (28%) with one or more NNRTIs. Nevertheless, for 69 patients (40%) there was no evidence of past virological failure on treatment (defined as having a single VL>400 copies/ml after at least 4 months of starting a regimen and whilst still receiving treatment). Besides AZT, 3TC and NFV, the new treatment contained one other drug in 46 patients (26%) and more than one other drugs in 31 patients (18%).

The 90 patients for whom a TCE was analyzed with a week 48 endpoint had characteristics that were no different from those who were included in the week 12 analysis. There was, however, a slightly lower

Vandamme 30/4/08 11:52 Page 4

PROOF COPY ONLY

percentage of patients with a history of virological failure while on AZT (38%), 3TC (31%) and most notably NFV (4%), and in this group the new treatment included drugs other than AZT, 3TC and NFV for relatively fewer patients (26% versus 46%).

Figure 1 shows the activity of AZT, 3TC and NFV as predicted by the Rega IS at TCE start. Table 2 shows the range of genotypic susceptibility and genetic barrier measures computed from the fitness landscapes. Predictions by Rega and HIVDB were highly concordant (percentage of concordant predictions, for AZT 95%, for 3TC 99% and for NFV 94%), while predictions by ANRS were less in agreement (concordance compared to Rega for AZT 87%, for 3TC 93% and for NFV 91%). All three expert IS and fitness landscape measures indicated a higher activity (and higher barrier to resistance) of AZT, 3TC and NFV at the start of the TCEs included in the week 48 analysis, compared with the week 12 analysis.

The Rega IS estimated the activity of other drugs in the regimen (besides AZT, 3TC and NFV) to have a GSS=1 in 41 patients (23%) and GSS>1 in 11 patients (6%) for the week 12 analysis; for the week 48 analysis, 13 patients (14%) had a GSS=1 and two patients (2%) had a GSS>1.

Prediction of viral load reduction at 12 weeks

For the 176 TCEs studied (of these, 40 patients [23%] had stopped one of the original drugs or started a new drug before the week 12 VL), all genotypic predictors for susceptibility to AZT+3TC, but only the number of major resistance mutations (*MR*) among the genotypic predictors for NFV, were associated with virological response in our linear regression analyses (Table 3, model 1). In the fully adjusted analyses (model 2), in addition to several genotypic predictors for susceptibility to AZT+3TC, a higher VL at TCE start was independently associated with a larger VL reduction (0.7 $\log_{10}$ per $\log_{10}$ higher, 95% confidence interval [CI] 0.4–1.0, $P<0.0001$) and a heterosexual source of HIV, which was independently associated with a 1.3 (95% CI 0.2–2.4) $\log_{10}$ lower VL reduction compared with a homosexual source of HIV ($P=0.02$).

AU: OK?

Table 1. Characteristics of the patients included in the analyses for treatment response

Characteristics	Week 12 (*n*=176)	Week 48 (*n*=90)
Mean age, years	39 (33–45)*	39 (32–45)*
Female gender, *n* (%)	48 (28)	27 (30)
Source of HIV exposure		
Homosexual, *n* (%)	86 (49)	35 (39)
Heterosexual, *n* (%)	40 (23)	27 (30)
Intravenous drug use, *n* (%)	42 (24)	24 (27)
Other, *n* (%)	8 (4.6)	4 (4.4)
Median starting year of treatment change episode, year	1999 (1998–1999)*	1999 (1998–1999)*
Median HIV RNA viral load, $\log_{10}$ copies/ml	4.5 (4.0–5.1)*	4.5 (4.0–5.1)*
Median $CD4^+$ T-cell count, cells/μl	231 (120–383)*	253 (165–424)*
Number of drugs previously failed		
0, *n* (%)	69 (40)	47 (52)
1–3, *n* (%)	51 (29)	26 (29)
4–6, *n* (%)	48 (27)	15 (17)
>6, *n* (%)	8 (4.6)	2 (2.2)
Previous drug failure		
Zidovudine, *n* (%)	72 (41)	34 (38)
Lamivudine, *n* (%)	82 (47)	28 (31)
Nelfinavir, *n* (%)	16 (9.1)	4 (4.4)
Other drugs in treatment		
Didanosine, *n* (%)	18 (10)	3 (3.3)
Abacavir, *n* (%)	12 (6.8)	1 (1.1)
Saquinavir, *n* (%)	10 (5.7)	1 (1.1)
Ritonavir, *n* (%)	15 (8.5)	1 (1.1)
Indinavir, *n* (%)	11 (6.2)	2 (2.2)
Changed therapy before week 12, *n* (%)	40 (22)	0 (0)
Changed therapy before week 48, *n* (%)	90 (51)	0 (0)

Values are expressed as *n* (%) of patients unless stated otherwise. *Values expressed as median (range).

© 2008 International Medical Press

Vandamme 30/4/08 11:52 Page 5

PROOF COPY ONLY

Prediction of virological failure at 48 weeks

Within the subset of 90 patients for whom a week 48 VL was available and who did not stop or switch any of the drugs included in the original treatment before the date of week 48 VL, 28 patients (31%) had virological failure at week 48.

The logistic regression analyses of the association between the genotypic predictors and the risk of virological failure at 48 weeks revealed the same trends as seen in the week 12 analyses, but the associations were slightly weaker (Table 4). In the fully adjusted analyses (model 2), both previous failure on a regimen containing 3TC (odds ratio [OR]≈25 [95% CI 2–261], P=0.01) and a heterosexual source of HIV infection compared with a homosexual source were independently associated with a higher risk for failure (OR≈50, P=0.02).

AU: OK?

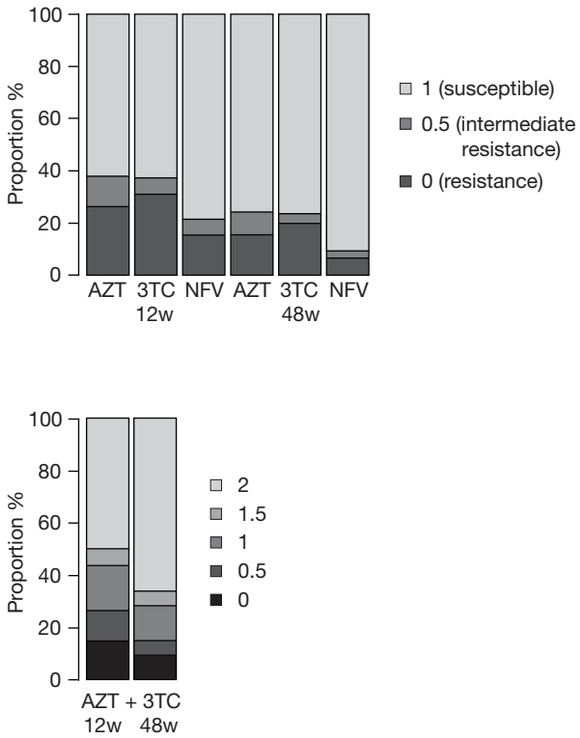
A ROC analysis was performed to compare a joint measure of sensitivity and specificity of the genotypic predictors derived from the fitness landscapes for prediction of virological failure at week 48 (Figure 2) from fitting univariables logistic regression models. The areas under the curve (AUC) were 0.63 (95% CI 0.49–0.76) for *F*, 0.65 (95% CI 0.53–0.78) for *MF* and 0.78 (95% CI 0.68–0.89) for *MR*, compared with 0.72 (95% CI 0.6–0.83) for Rega (P=0.03).

AU: OK?

Discussion

We present results of a retrospective evaluation of the use of two estimated fitness landscapes (for NFV and AZT+3TC) to predict virological outcome to treatment containing a background of AZT+3TC+NFV. The

Figure 1. Distribution of genotypic susceptibility scores



Distribution of genotypic susceptibility scores (GSS) of baseline sequences (as analyzed by Rega 7.1) of treatment change episodes (TCEs) included in the week 12 or week 48 analyses, for zidovudine (AZT), lamivudine (3TC), nelfanavir (NFV) and the combination of AZT+3TC. GSS values for the combination AZT+3TC is the sum of the GSS for AZT and 3TC and indicates the number of active drugs among AZT and 3TC as predicted by Rega.

Table 2. Quantification of genotypic susceptibility and genetic barrier

Factors	Week 12 (n=176)	Week 48 (n=90)
Fitness under treatment, $\log_{10}$ *F*		
NFV	0.51 (0.27–0.78)	0.39 (0.18–0.65)
AZT+3TC	0.18 (-0.04–1.36)	0.06 (-0.07–0.33)
Number of mutations to MRM, *MR*		
NFV	2.94 (1.95–3.55)	3.07 (2.55–3.71)
AZT+3TC	1.36 (0.00–2.10)	1.79 (0.00–2.16)
Number of mutations to high fitness, *MF*		
NFV	3.21 (2.24–4.27)	3.67 (2.67–4.54)
AZT+3TC	4.46 (0.00–5.29)	4.86 (2.88–5.35)
Number of generations to MRM, *GR*		
NFV	122 (77–148)	130 (71–153)
AZT+3TC	50 (0–90)	73 (0–93)
Number of generations to high fitness, *GF*		
NFV	125 (83–170)	140 (102–179)
AZT+3TC	175 (0–209)	194 (105–216)

Quantification of genotypic susceptibility ($\log_{10}$ *F*, estimated fitness under treatment compared with the HXB2 reference sequence) and genetic barrier (*MR*, *MF*, *GR* and *GF*), estimated using the fitness landscapes for 176 treatment change episodes (TCEs) included in the week 12 analysis and 90 TCEs included in the week 48 analysis. Values are expressed as median (interquartile range). AZT, zidovudine; MRM, major resistance mutations; NFV, nelfanavir; 3TC, lamivudine.

Table 3. Linear regression analyses of the association between genotypic predictors of interest and viral load reduction at 12 weeks

	Model 1		Model 2	
Factors	Δlog VL (95% CI)	*P*-value	Δlog VL (95% CI)	*P*-value
GSS $Rega_{AZT+3TC}$	0.90 (0.53–1.27)	<0.0001	0.60 (0.14–1.05)	0.01
GSS $Rega_{NFV}$	0.00 (-0.73–0.74)	0.99	-0.01 (-0.77–0.76)	0.99
GSS $ANRS_{AZT+3TC}$	0.76 (0.43–1.10)	<0.0001	0.45 (0.04–0.86)	0.03
GSS $ANRS_{NFV}$	0.35 (-0.33–1.03)	0.32	0.31 (-0.40–1.03)	0.39
GSS $HIVDB_{AZT+3TC}$	0.86 (0.47–1.24)	<0.0001	0.53 (0.06–1.00)	0.03
GSS $HIVDB_{NFV}$	0.12 (-0.63–0.88)	0.75	0.15 (-0.63–0.92)	0.71
$\log_{10} F_{AZT+3TC}$	-0.38 (-0.58– -0.18)	0.0002	-0.24 (-0.46– -0.01)	0.04
$\log_{10} F_{NFV}$	-0.34 (-0.91–0.23)	0.25	-0.22 (-0.81–0.37)	0.46
$MR_{AZT+3TC}$*	0.54 (0.30–0.77)	<0.0001	0.31 (0.0–0.6)	0.08
MR_{NFV}*	0.19 (0.00–0.39)	0.05	0.15 (-0.05–0.34)	0.13
$MF_{AZT+3TC}$*	0.24 (0.13–0.35)	<0.0001	0.14 (0.00–0.29)	0.05
MF_{NFV}*	0.07 (-0.09–0.23)	0.40	0.04 (0.13–0.21)	0.67
$GR_{AZT+3TC}$†	1.28 (0.72–1.84)	<0.0001	0.66 (-0.08–1.39)	0.08
GR_{NFV}†	0.40 (-0.05–0.86)	0.08	0.30 (-0.16–0.76)	0.20
$GF_{AZT+3TC}$†	0.59 (0.32–0.87)	<0.0001	0.35 (0.00–0.70)	0.05
GF_{NFV}†	0.12 (-0.27–0.51)	0.55	0.04 (-0.37–0.45)	0.85

Linear regression analyses of the association between the genotypic predictors of interest and viral load (VL) reduction at 12 weeks (Δlog VL) using 176 treatment change episodes (TCEs). For each of the three expert systems, two models including different sets of potential confounders were performed (model 1 was adjusted only for baseline VL and time to follow-up sample and model 2 for all factors listed in Table 1). Estimates are expressed as $\log_{10}$ VL reductions from pre-TCE levels per unit higher of the predictor unless stated otherwise. *per additional mutation. †per 100 generations higher. AZT, zidovudine; CI, confidence interval; *F*, fitness under treatment; *GF*, estimated number of generations to a fitness above the chosen fitness threshold; *GR*, estimated number of generations to a major resistance mutation; GSS, genotypic susceptibility scores; *MF*, estimated number of mutations to a fitness above the chosen fitness threshold; *MR*, estimated number of mutations to a major resistance mutation; NFV, nelfinavir; 3TC, lamivudine.

patient population was heavily pretreated and a large group of patients had previously failed on AZT (46%) or 3TC (42%), but fewer had failed on a treatment with NFV (9%).

Overall, predictions of treatment outcome made on the basis of on the estimated fitness landscapes had similar performance compared with expert systems. The results shown in Tables 3 and 4 do not compare the genotypic predictors in terms of their goodness of fit. However, we performed an additional analysis showing that *MR* was superior to Rega in terms of AUC (Figure 2). The fitness landscapes were trained from genotypic data only and all five estimates derived from the fitness landscapes were entirely based on computational analysis. Machine learning has been applied to the problem of genotypic resistance interpretation previously; however, these systems were trained from correlations between genotype and *in vitro* resistance phenotype [11–13] or from correlations between baseline genotype and treatment response [14]. A notable disadvantage of systems trained from *in vitro* resistance phenotypic data is the absence of genetic barrier information and the fact that currently available assays only detect mutations in the dominant virus. Recently [15], it was shown how the estimated genetic barrier may indeed complement phenotypic resistance data towards a better prediction of treatment outcome.

As the results show that the estimated fitness functions have predictive value for treatment response, an evaluation of the effect of several choices and parameters used to estimate the fitness functions, such as effective population size, included mutations and number of interaction terms, on this association with treatment response is warranted. Such a study could help in motivating these choices to improve predictive performance. In addition, further external validations of our genotypic predictors on independent data sources are needed.

Predicted susceptibility to AZT+3TC using all systems seemed to correlate well with both short- and long-term virological response, while predicted susceptibility to NFV using expert IS did not seem to be useful in identifying patients who were more likely to experience poor virological outcomes. The estimated genetic barrier to a major resistance mutation (*MR* and *GR*) were the only genotypic predictors for NFV that tended to be associated with treatment response in our most parsimonious models (with *P*-values in the range 0.05–0.10). The poor association between genotypic interpretations for sensitivity to NFV and virological outcome has been observed previously [16] and a low virological potency of NFV was suggested as one possible explanation. Indeed, the predicted susceptibility of the PI component of regimens containing ritonavir-boosted PI seem to correlate well with the

© 2008 International Medical Press

Vandamme 30/4/08 11:52 Page 7

PROOF COPY ONLY

Table 4. Logistic regression analyses of the assocation between genotypic predictors of interest and the risk of virological failure at 48 weeks

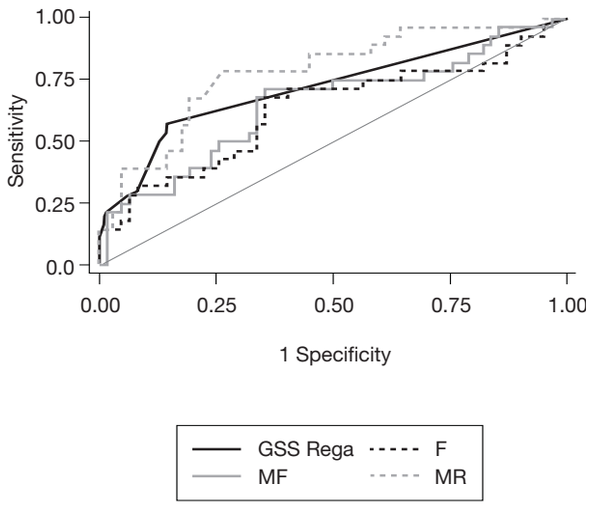
Factors	Model 1		Model 2	
	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
GSS $Rega_{AZT+3TC}$	0.22 (0.09–0.56)	0.001	0.15 (0.03–0.68)	0.01
GSS $Rega_{NFV}$	1.42 (0.17–12.1)	0.75	2.09 (0.13–33.0)	0.60
GSS $ANRS_{AZT+3TC}$	0.35 (0.16–0.78)	0.01	0.28 (0.08–1.02)	0.05
GSS $ANRS_{NFV}$	0.18 (0.02–1.82)	0.15	0.46 (0.01– 4.9)	0.66
GSS $HIVDB_{AZT+3TC}$	0.30 (0.12–0.71)	0.006	0.27 (0.06–1.16)	0.08
GSS $HIVDB_{NFV}$	0.72 (0.08–6.75)	0.77	0.99 (0.05–20.2)	0.99
$\log_{10} F_{AZT+3TC}$	1.63 (0.94–2.83)	0.08	1.46 (0.59–3.63)	0.41
$\log_{10} F_{NFV}$	1.70 (0.37–7.69)	0.49	0.31 (0.03–3.30)	0.33
$MR_{AZT+3TC}$ (per additional mutation)	0.39 (0.23–0.66)	0.0005	0.30 (0.11–0.81)	0.02
MR_{NFV} (per additional mutation)	0.63 (0.36–1.09)	0.10	0.66 (0.32–1.34)	0.25
$MF_{AZT+3TC}$ (per additional mutation)	0.81 (0.64–1.02)	0.07	0.95 (0.62–1.45)	0.80
MF_{NFV} (per additional mutation)	0.90 (0.64–1.26)	0.53	1.09 (0.66–1.81)	0.72
$GR_{AZT+3TC}$ (per 100 generations higher)	0.12 (0.03–0.43)	0.0008	0.06 (0.005–0.64)	0.02
GR_{NFV} (per 100 generations higher)	0.31 (0.09–1.15)	0.08	0.34 (0.06–1.83)	0.21
$GF_{AZT+3TC}$ (per 100 generations higher)	0.62 (0.35–1.08)	0.09	0.91 (0.31–2.64)	0.86
GF_{NFV} (per 100 generations higher)	0.73 (0.31–1.71)	0.47	0.96 (0.28–3.27)	0.95

Logistic regression analyses of the assocation between the genotypic predictors of interest and the risk of virological failure at 48 weeks for 90 treatment change episodes (TCEs). For each system, models 1 and 2 were performed. Unless indicated otherwise, estimates express the risk of week 48 virological failure associated with a unit higher in the predictor. AZT, zidovudine; CI, confidence interval; *F*, fitness under treatment; *GF*, estimated number of generations to a fitness above the chosen fitness threshold; *GR*, estimated number of generations to a major resistance mutation; GSS, genotypic susceptibility scores; *MF*, estimated number of mutations to a fitness above the chosen fitness threshold; *MR*, estimated number of mutations to a major resistance mutation; NFV, nelfinavir; OR, odds ratio; 3TC, lamivudine.

virological response in clinical practice [17]. However, the low prevalence of NFV resistance (the expert systems scored most isolates as susceptible to NFV [Rega 79% for week 12 and 91% for week 48] and very few isolates showing MRMs [no isolates with 30N, one isolate with 90M and four isolates with 82A at the pre-therapy genotypic tests in patients included in the week 48 analysis; data not shown]) may be another possible reason for the poor discriminatory power of the rule-based IS scores. Within the class of PIs, NFV is considered a drug with a relatively low genetic barrier, thus variation in this genetic barrier (quantified in *MR* and *GR*) for isolates without obvious NFV drug resistance may have contributed to better predictions. In this context it is interesting to note that all genotypic estimates, on average, are better for genotypes included in the 48 week analysis (that is, higher predicted susceptibility, lower fitness under drug selective pressure, higher genetic barrier towards resistance), compared with the 12 week analysis. As patients switching treatment were excluded from the 48 week analysis, this provides additional circumstantial evidence that better genotypically predicted virological response is indeed associated with a lower risk of treatment failure.

Among the five estimates derived from fitness landscapes, the estimated genetic barrier to a major resistance mutation (*MR* and *GR*) consistently showed the best predictions and the estimated fitness

Figure 2. ROC curves showing the trade-off between sensitivity and specificity for genotypic predictors



Receiver operating characteristic (ROC) curves showing the trade-off between sensitivity and specificity for genotypic predictors derived from the fitness landscapes compared with Rega 7.1 for prediction of virological failure at week 48 using 90 treatment change episodes. *F*, estimated fitness under treatment; GSS, genotypic susceptibility scores; *MF*, estimated number of mutations to a fitness above the chosen fitness threshold; *MR*, estimated number of mutations to a major resistance mutation.

Vandamme 30/4/08 11:52 Page 8

under treatment showed the worst predictive performance. The fitness functions are likely to be distorted because of approximations and assumptions in the estimation process [6], but these distortions should not necessarily affect the accuracy to predict evolution. The *MR* and *GR* were not using the fitness landscapes as a measure for resistance, but only for evolution towards a separately defined resistance estimate (presence of an MRM). The fact that *MR* and *GR* used the fitness function only to simulate evolution, and that such purely genetic barrier estimates tended to be marginally better than the fitness of mixed estimates (*F*, *MF* and *GF*), leads us to hypothesize that the best use of the fitness functions for prediction of treatment outcome is perhaps to complement a virtual phenotype system for quantifying resistance with an evolutionary dimension to estimate the genetic barrier. This is similar to using a combination of mutagenic trees to model evolution and Geno2pheno to determine resistance [15], which was shown to have good performance compared with other systems. A possible advantage of fitness landscapes over mutagenic tree models is that not only resistance mutations are taken into account, but also information on all variable amino acid positions (including subtype polymorphisms). Also, variation in codon usage is implicitly taken into account through simulation at the nucleotide level. A follow-up of this study in which fitness landscapes are used in combination with phenotypic resistance data to obtain predictions is therefore warranted.

The genetic barrier may not only be beneficial for improved prediction of treatment response because it considers the effect of evolution under treatment, but also because it implicitly captures the risk of preexisting minority variants among the quasispecies variation at baseline [18]. For this reason, the estimated genetic barrier may partially overcome one of the main limitations of the consensus population sequence: the fact that it cannot detect mutations present in a minority of the intrahost virus population.

In our analysis, a simple model including only pre-TCE VL and time between the TCE and the date of the endpoint VL (model 1) and a model including a larger number of potential confounders (model 2) were employed. In the latter analyses, we found that a heterosexual source of infection was a risk factor for virological failure; this association was found in previous analyses of cohort studies [19] and is potentially related to adherence (better in homosexuals than in other risk groups).

Clinical retrospective evaluations of genotypic resistance interpretation systems are continuously needed to validate their performance and ideally these should be conducted in worldwide collaborations including a large number of genotypic sequences and using a standardized analysis plan. In this evaluation, the researchers computing the genotypic predictions were blinded to the virological outcome data, reducing the chance that the observed results were affected by bias.

The fitness functions were estimated from sequences of different subtypes. Because laboratory strains used in the phenotypic resistance assays are subtype B, it is not clear to what extent phenotypic resistance tests are distorted by recombining gene fragments from non-B subtypes isolates into a subtype B vector [20]. Treatment response data is mostly available only for patients infected with a subtype B virus (in this study, less than 11% patients were infected with a non-B subtype strain). This may explain a higher uncertainty in expert opinion for resistance interpretation of genotypes of non-B subtypes, resulting in a higher discordance of expert IS for non-B subtype than for B subtypes [7]. Because the most prevalent subtypes worldwide are HIV-1 subtypes A and C [21], an improved understanding of drug resistance in these and other subtypes is urgently needed as the prevalence of non-B subtypes is increasing, especially in Europe [22–25], and as antiviral treatment becomes increasingly available in developing countries. Unlike expert IS and systems trained using phenotypic resistance data, the fitness functions were estimated without any methodological bias towards subtype B (other than the fact that it was the most prevalent subtype in the training data sets). The results of our analysis suggest a promising role of this new methodology for the interpretation of genotypic data and for the prediction of virological outcomes that could be potentially used also in non-B subtypes.

In conclusion, for a genotypic interpretation aiming to predict the virological response to treatment with AZT, 3TC and NFV irrespective of the subtype, the fitness functions showed performance similar to that of expert IS. If anything, the sensitivity of our method for NFV was slightly greater than the sensitivity of expert IS in the absence of MRM. Because the fitness functions are subtype neutral, and estimated from clinical genotypic data only, these results warrant a larger investigation, ideally in patients infected with other subtypes and for other drugs currently used in clinical practice.

Acknowledgements

KD was funded by a PhD grant of the Institute for the Promotion of Innovation through Sciences and Technology in Flanders (IWT). YM is a post-doctoral researcher with the FWO-Vlaanderen; his research is supported by KULeuven GOA-Mefisto-666 and GOA-

© 2008 International Medical Press

Vandamme 30/4/08 11:52 Page 9

PROOF COPY ONLY

Ambiorics, Belspo IUAP V-22 and EU FP6 NoE Biopattern. This work was supported by FWO-Vlaanderen grant G.0266.04 and by the Katholieke Universiteit Leuven through Grant OT/04/43.

Primary funding support for EuroSIDA is provided by the European Commission BIOMED 1 (CT94-1637), BIOMED 2 (CT97-2713), the 5th Framework (QLK2-2000-00773) and the 6th Framework (LSHP-CT-2006-018632) programs. Current support also includes unrestricted grants by Bristol–Myers Squibb, GlaxoSmithKline, Roche, Gilead, Pfizer, Merck and Co., Tibotec and Boehringer-Ingelheim. The participation of centres from Switzerland was supported by a grant from the Swiss Federal Office for Education and Science.

Additional file

An additional file detailing members of the EuroSIDA Study Group can be accessed via the Volume 13 Issue 3 contents page for Antiviral Therapy, which can be found at www.intmedpress.com (by clicking on 'Antiviral Therapy' then 'Journal PDFs').

Disclosure statement

None of the authors of this report have financial or personal relationships with people or organizations that could influence this work. Although most authors have, at some stage in the past, received funding from a variety of pharmaceutical companies for research, travel grants, speaking engagements or consultancy fees.

References

1. Vandamme AM. Managing resistance to anti-HIV drugs: an important consideration for effective disease management. *Drugs* 1999; **57:**337–361.
2. Vandamme AM, Sonnerborg A, Ait-Khaled M, *et al.* Updated European recommendations for the clinical use of HIV drug resistance testing. *Antivir Ther* 2004; **9:**829–848.
3. DiRienzo G, DeGruttola V. Collaborative HIV resistance-response database initiatives: sample size for detection of relationships between HIV-1 genotype and HIV-1 RNA response using a non-parametric approach. *Antivir Ther* 2002; **7:**S71.
4. Marcelin AG, Flandre P, Pavie J, *et al.* Clinically relevant genotype interpretation of resistance to didanosine. *Antimicrob Agents Chemother* 2005; **49:**1739–1744.
5. Van Laethem K, Vandamme AM. Interpreting resistance data for HIV-1 therapy management – know the limitations. *AIDS Reviews* 2006; **8:**37–43.
6. Deforche K, Camacho R, Van Laethem K, *et al.* Estimation of an *in vivo* fitness landscape experienced by HIV-1 under drug selective pressure useful for prediction of drug resistance evolution during treatment. *Bioinformatics* 2007; **24:**34–41.
7. Snoeck J, Kantor R, Shafer RW, *et al.* Discordances between interpretation algorithms for genotypic resistance to protease and reverse transcriptase inhibitors of human immunodeficiency virus are subtype dependent. *Antimicrob Agents Chemother* 2006; **50:**694–701.
8. de Oliveira T, Deforche K, Cassol S, *et al.* An automated genotyping system for analysis of HIV-1 and other microbial sequences. *Bioinformatics* 2005; **21:**3797–3800.
9. Deforche K, Silander T, Camacho R, *et al.* Analysis of HIV-1 pol sequences using Bayesian Networks: implications for drug resistance. *Bioinformatics* 2006; **22:**2975–2979.
10. Beerenwinkel N, Rahnenführer J, Däumer M, *et al.* Learning multiple evolutionary pathways from cross-sectional data. *J Compu Biol* 2005; **12:**584–598.
11. Beerenwinkel N, Daumer M, Oette M, *et al.* Geno2pheno: estimating phenotypic drug resistance from HIV-1 genotypes. *Nucl Acids Res* 2003; **31:**3850–3855.
12. Wang D, Larder B. Enhanced prediction of lopinavir resistance from genotype by use of artificial neural networks. *J Infect Dis* 2003; **188:**653–660.
13. Loutfy MR, Raboud JM, Walmsley SL, *et al.* Predictive value of HIV-1 protease genotype and virtual phenotype on the virological response to lopinavir/ritonavir-containing salvage regimens. *Antivir Ther* 2004; **9:**595–602.
14. Larder B, Wang D, Revell A, *et al.* The development of artificial neural networks to predict virological response to combination HIV therapy. *Antivir Ther* 2007; **12:**15–24.
15. Altmann A, Beerenwinkel N, Sing T, *et al.* Improved prediction of response to antiretroviral combination therapy using the genetic barrier to drug resistance. *Antivir Ther* 2007; **12:**169–178.
16. De Luca A, Cingolani A, Giambenedetto SD, *et al.* Variable prediction of antiretroviral treatment outcome by different systems for interpreting genotypic human immunodeficiency virus type 1 drug resistance. *J Infect Dis* 2003; **187:**1934–1943.
17. Fox ZV, Geretti AM, Kjaer J, *et al.* The ability of four genotypic interpretation systems to predict virological response to ritonavir-boosted protease inhibitors. *AIDS* 2007; **21:**2033–2042.
18. Roquebert B, Malet I, Wirden M, *et al.* Role of HIV-1 minority populations on resistance mutational pattern evolution and susceptibility to protease inhibitors. *AIDS* 2006; **20:**287–289.
19. Fardet L, Mary-Krause M, Heard I, *et al.* On HIV FHD, influence of gender and HIV transmission group on initial highly active antiretroviral therapy prescription and treatment response. *HIV Med* 2006; **7:**520–529.
20. Kantor R, Katzenstein D. Drug resistance in non-subtype B HIV-1. *J Clin Virol* 2004; **29:**152–159.
21. Osmanov S, Pattout C, Walker N, *et al.* Estimated global distribution and regional spread of HIV-1 genetic subtypes in the year 2000. *J Acquir Immune Defic Syndr* 2002; **29:**184–190.
22. Böni J, Pyra H, Gebhardt M, *et al.* High frequency of non-B-subtypes in newly diagnosed HIV-1 infections in Switzerland. *J Acquir Immune Defic Syndr* 1999; **22:**174–179.
23. Balotta C, Facchi G, Violin M, *et al.* Increasing prevalence of non-clade B HIV-1 strains in heterosexual men and women, as monitored by analysis of reverse transcriptase and protease sequences. *J Acquir Immune Defic Syndr* 2001; **27:**499–505.
24. Snoeck J, Laethem KV, Hermans P, *et al.* Rising prevalence of HIV-1 non-B -subtypes in Belgium: 1983–2001. *J Acquir Immune Defic Syndr* 2004; **35:**279–285.
25. Aggarwal I, Smith M, Tatt ID, *et al.* Evidence for onward transmission of HIV-1 -subtype strains in the United Kingdom. *J Acquir Immune Defic Syndr* 2006; **41:**201–209.

Accepted for publication 8 January 2008

Vandamme 30/4/08 11:52 Page 10