Journal of Clinical Virology 65 (2015) 68–73

View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Journal of Clinical Virology

journal homepage: www.elsevier.com/locate/jcv

Effects of sequence alterations on results from genotypic tropism testing

Alejandro Pironti [a,*], Saleta Sierra [b], Rolf Kaiser [b], Thomas Lengauer [a], Nico Pfeifer [a]

[a] Department of Computational Biology and Applied Algorithmics, Max Planck Institute for Informatics, Campus E 14, 66123 Saarbrücken,Germany
[b] Institute of Virology, University of Cologne, Furst- Puckler-Str. 56, 50935 Cologne, Germany

ARTICLE INFO

Article history:
Received 3 November 2014
Received in revised form 28 January 2015
Accepted 6 February 2015

Keywords:
HIV-1
Tropism
Genotypic tropism determination
Geno2pheno
Triplicate
Simulation of Sanger bulk-sequencing errors

ABSTRACT

Background: $\text{geno2pheno}_{[\text{coreceptor}]}$ is a bioinformatic method for genotypic tropism determination (GTD) which has been extensively validated.
Objectives: GTD can be affected by sequencing/base-calling variability and unreliable representation of minority populations in Sanger bulk sequencing. This study aims at quantifying the robustness of $\text{geno2pheno}_{[\text{coreceptor}]}$ with respect to these issues. GTD with a single amplification or in triplicate (henceforth singleton/triplicate) is considered.
Study Design: From a dataset containing 67,997HIV-1 V3 nucleotide sequences, two datasets simulating sequencing variability were created. Further two datasets were created to simulate unreliable representation of minority variants. After interpretation of all sequences with $\text{geno2pheno}_{[\text{coreceptor}]}$, probabilities of change of predicted tropism were calculated.
Results: $\text{geno2pheno}_{[\text{coreceptor}]}$ tends to report reduced false-positive rates (FPRs) when sequence alterations are present. Triplicate FPRs tend to be lower than singleton FPRs, resulting in a bias towards classifying viruses as X4-capable. Alterations introduced into nucleotide sequences by simulation change singleton predicted tropism with a probability $\leq 2\%$. Triplicate prediction lowers this probability for predicted X4 tropism, but raises it for predicted R5 tropism $\leq 6\%$. Simulated limited detection of minority variants in X4 sequences resulted in unchanged predicted tropism with probability above 90% as compared to probability above 98% with triplicate FPRs.
Conclusions: $\text{geno2pheno}_{[\text{coreceptor}]}$ proved to be robust when sequence alterations are present and when detectable minorities are missed by bulk sequencing. Changes in tropism prediction due to sequence alterations as well as triplicate prediction are much more likely to result in false X4-capable predictions than in false R5 predictions.

© 2015 Published by Elsevier B.V.

1. Background

The Human Immunodeficiency Virus type 1 (HIV-1) employs two host molecules in order to enter the host cell: the CD4 receptor and a coreceptor. In vivo, either of two coreceptors can be: CCR5 and CXCR4. The capability to use a certain coreceptor is called viral tropism. HIV-1 so-called R5 strains can only use CCR5. CXCR4-capable strains can use either CXCR4, exclusively (X4 strains), or both coreceptors (dual/mixed tropic viruses) [1]. Maraviroc (MVC) is an antiretroviral that inhibits HIV-1 entry into the cell by binding to CCR5, and is thus ineffective against X4-capable strains. Therefore, viral tropism determination must precede MVC prescription [2].

Tropism can be determined phenotypically or genotypically [3–6]. Phenotypic determination in cell cultures is expensive, time-consuming, and requires specialized labs. Furthermore, samples with viral loads up to 1000 cp/ml often yield indefinite results, although proviral DNA testing is performed also [7]. Genotypic tropism determination (GTD) requires genotyping the third hypervariable loop of the HIV-1 env gene (V3) with subsequent computer-based interpretation. Several methods for interpreting sequences in order to determine HIV-1 tropism have been

Abbreviations: HIV-1, human immunodeficiency virus type 1; MVC, maraviroc; GTD, genotypic tropism determination; FPR, false-positive rate; V3, third hypervariable loop; EP, electropherogram dataset; ME, manually-edited sequence; AS, automatic sequence; SEP, single-electropherogram sequence; indel, insertion or deletion; LA, Los-Alamos dataset; SA, Sanger-alteration dataset; SE, single-error dataset; M, mixture dataset; MS, mixture-sampling dataset; CPT, change in predicted tropism; S, singleton; T, triplicate; SD, standard deviation.

* Corresponding author. Tel.: +49 681 9325 3122; fax: +49 681 9325 3099.
E-mail address: apironti@mpi-inf.mpg.de (A. Pironti).

http://dx.doi.org/10.1016/j.jcv.2015.02.006
1386-6532/© 2015 Published by Elsevier B.V.

developed [6]. $geno2pheno_{[coreceptor]}$ [8] is an extensively validated bioinformatic method for GTD [3–6,9]; its use as an alternative to phenotypic determination is recommended by the European and the Austrian–German HIV-treatment guidelines [12–14].

$geno2pheno_{[coreceptor]}$ interprets V3 with a Support Vector Machine (SVM) trained on genotype–phenotype pairs. $Geno2pheno_{[coreceptor]}$ outputs the false-positive rate (FPR)[1] with X4-capable being defined as positive [8]. FPR dichotomization yields a (predicted) viral classification into X4-capable or R5. When the FPR is in a range where MVC antiviral action is considered possible, yet uncertain, the virus is classified as X4-capable. Alternatively, this intermediate FPR range can be explicitly labeled, as is customary for interpretation of drug resistance to other antiretroviral drugs. Thus, MVC administration with an FPR in the intermediate range could be made dependent on whether other therapy options co-exist, rather than excluding it altogether. Furthermore, the use of an intermediate category when determining tropism can also be appropriate for predictions with FPR-decreasing sequence alterations, especially when determining tropism in triplicate.

Establishment of the most suitable cutoff for FPR dichotomization has been a matter of substantial debate. Currently, there is no universally accepted cutoff.

2. Objectives

The input to $geno2pheno_{[coreceptor]}$ is aV3 sequence. Therefore, the quality of the predictions depends on the quality of these sequences. With Sanger bulk sequencing, the measured sequence is a consensus of the dominating strains in the viral population. Here, minorities comprising less than 10%-20% of the viral population are unreliably represented, due to the limits of the experimental technology. X4-capable minorities may render MVC ineffective. Therefore, some labs perform the amplification step of the sequencing procedure in triplicate [15] to increase the chances of detecting minorities. This analysis addresses two related, unresolved questions: (1) How robust is $geno2pheno_{[coreceptor]}$ with respect to sequencing/base-calling variability in terms of change of predicted tropism (CPT)? (2) What is the influence of undetected minority populations on the predictions of $geno2pheno_{[coreceptor]}$? Both issues are of critical importance for assessing the reliability of $geno2pheno_{[coreceptor]}$ for clinical purposes. This article aims at providing answers to both questions.

3. Study design

A dataset of 163,958HIV-1 V3 nucleotide sequences was downloaded from the Los Alamos National Laboratory (http://www.hiv.lanl.gov/) on September, 19th 2013. Nucleotide sequences with duplicate V3 regions were discarded, resulting in the Los-Alamos dataset (LA) comprising 67,997 nucleotide sequences. Subtypes in LA were determined with Comet [16]. LA was used to create further datasets by altering its sequences *in silico*.

The probability of each type of sequence variation (base exchanges, ambiguities, insertions/deletions [indels]) at each sequence position resulting from differential primer specificity, and misincorporation/indel events during the RT-PCR amplification and sequencing reactions was determined using 164 clinical blood samples obtained from the Department of Clinical Virology of the University of Cologne, as described in Section Supplementary methods. This knowledge was used to generate the Sanger-alteration dataset (SA). The dataset contains 10 *in-silico* generated variants of each LA sequence. Alterations in these *in-silico* variants may be present at any of the 105 nucleotide positions.

The single-error dataset (SE) was created by generating sequences from each LA sequence by systematically exchanging every nucleotide in V3 by each of the 15 possible definite and ambiguous bases, independently of their probability of occurrence. Thus, from each sequence in LA, all possible sequences diverging by one definite or ambiguous base were generated.

The mixture dataset (M) was created from sequences in LA containing ambiguities (excluding N). Ambiguities in each of these sequences were combinatorially resolved into all possible sequence alternatives without ambiguities. To avoid combinatorial explosion, sequences that would result in more than 20,000 derived sequences were excluded from this procedure.

The mixture-sampling dataset (MS) was created to simulate a scenario in which sequencing depth is insufficient to resolve all sequence variants in the sample. From each sequence group in M derived from the same LA sequence, a certain proportion of sequences was extracted at random by uniform sampling without replacement, and a new sequence was created by retaining positions that are identical among the sequences in the subset and representing differential positions with the corresponding ambiguities. Proportions represent sequencing depth and ranged from 1% to 100% in steps of 2%; each sequence group was sampled 3×100 times (100 repetitions that allow for triplicate FPRs).

Sequences in LA, SA, SE, M, and MS were interpreted with $geno2pheno_{[coreceptor]}$. For each sequence, the FPR shift was calculated as the difference between the FPR of the altered sequence and that of its unaltered counterpart in LA. When we consider n FPRs from variability-simulation replicates on the same sequence, we call *singleton FPR* the FPR obtained with the first sequence. For all further sequences, we take the minimum FPR among the first n and call it nth replicate FPR. The 3rd replicate FPR is also called triplicate FPR. $geno2pheno_{[coreceptor]}$'s FPR was used to determine coreceptor tropism as X4-capbable or R5. Four different FPR cutoff sets were used for tropism determination:

- {5, 10}: $FPR < 5 \Rightarrow$ X4-capable, $5 \leq FPR < 10 \Rightarrow$ Intermediate, $FPR \geq 10 \Rightarrow$ R5
- {5, 15}: $FPR < 5 \Rightarrow$ X4-capable, $5 \leq FPR < 15 \Rightarrow$ Intermediate, $FPR \geq 15 \Rightarrow$ R5 [14]
- {10}: $FPR < 10 \Rightarrow$ X4-capable, $FPR \geq 10 \Rightarrow$ R5 [13]
- {20}: $FPR < 20 \Rightarrow$ X4-capable, $FPR \geq 20 \Rightarrow$ R5 [13]

According to Austrian–German treatment guidelines, MVC can be effective when a tropism prediction is labeled intermediate albeit with much less certainty than for R5 variants [14].

The probability the tropism predicted by $geno2pheno_{[coreceptor]}$ changed due to the introduced sequence alterations was calculated by sample counting as $P(T_{A,C} \mid T_{U,C})$, with $T_{A,C}$ denoting the tropism of the altered sequences as determined with cutoff set C, and $T_{U,C}$ denoting the tropism of the unaltered sequences as determined with cutoff set C. The reference sequence used to number V3 nucleotide positions is consensus B (105 nucleotides), which is the reference used by $geno2pheno_{[coreceptor]}$.

4. Results

Alteration rates estimated for generation of SA (Section Supplementary methods) are shown in Supplementary Tables 1 and 2.

The FPR distribution in LA is illustrated in Supplementary Fig. 1. The numbers of strains by subtype are tabulated in Table 1. In LA, 0.24% of the bases are ambiguous, while 99.76% of the bases are definite.

[1] More accurately: it outputs the smallest FPR, of an SVM-classifier that classifies the sequence under inspection as X4-capable.

Table 1
Number of strains by subtype, circulating recombinant form or group, LA dataset.The numbers of strains of each subtype, circulating recombinant form (CFR) or group in the LA dataset are tabulated above.

Subtype, CRF or Group	Number of strains
B	35312
C	11056
A1	10187
01_AE	4986
D	2691
G	1422
F1	1308
F2	324
O	226
A2	207
H	121
CPZ	68
J	46
K	27
N	11
P	5

Among the FPR shifts between the altered SA sequence and that of its unaltered counterpart in LA, 79% are equal to zero, 11% are below zero and 10% are above zero (see also Supplementary Fig. 2). Fig. 1 shows how the probability of obtaining FPRs lower than the singleton FPR increases with the number of amplifications. Finally, Figs. 2 and 3 depict the change of predicted tropism (CPT) probabilities for different cutoff sets (probabilities labeled *S* for singleton FPRs and *T* for triplicate FPRs). Figs. 2 and 3 show that geno2pheno$_{[coreceptor]}$ is more likely to reduce than to raise FPR when sequence alterations are present, slightly favoring a false prediction of R5 viruses as X4-capable over the reverse misprediction (see also Supplementary Fig. 2). Since the FPR shifts are both negative and positive, FPR determination in triplicate will always reduce FPRs, as the minimum FPR is selected. When determining FPR in triplicate, there is a 27% chance that the triplicate FPR will be lower than the singleton FPR (Fig. 1).

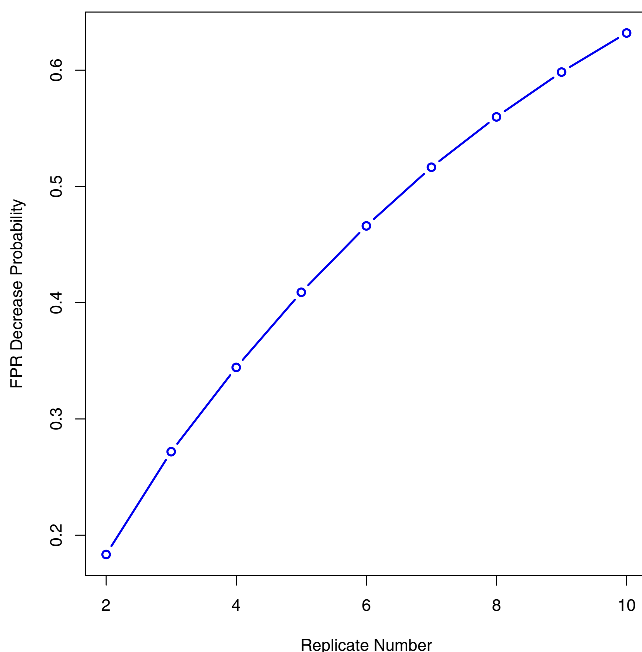

Fig. 1. Probabilities of decreasing FPR as more replicates are performed. FPRs in SA dataset were compared with their unaltered counterparts in the LA dataset. The probabilities of obtaining nth replicate FPRs lower than the singleton FPRs are plotted with increasing values for n (replicate number) in the chart above (Color/BW).

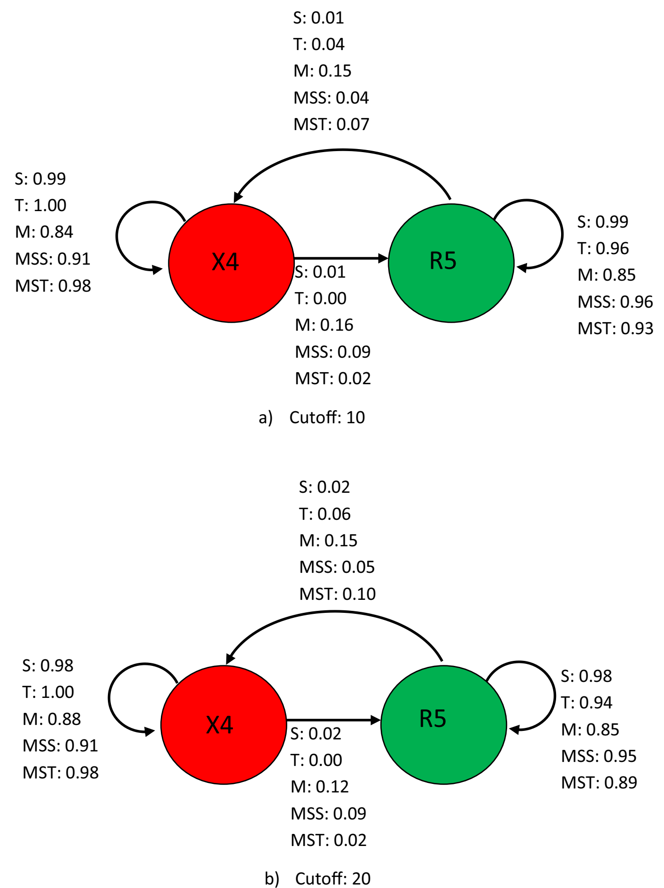

Fig. 2. Conditional probabilities of change in predicted tropism, cutoff sets {10} and {20}: Cutoff sets {10} and {20} were applied to FPRs in LA, SA, M, and MS datasets to calculate the conditional probabilities of CPT. Results are shown for SA dataset singleton (S), SA dataset triplicate (T), M dataset (M), MS dataset singleton (MSS), and MS dataset triplicate (MST) FPRs. There are no replicates on the M dataset since its sequences contain no ambiguities. The circles in the figure represent the predicted tropism. The arrows indicate a change in predicted tropism with the head of the arrow pointing towards the change. The arrow labels contain the probabilities for the respective changes in predicted tropism as calculated with the dataset mentioned above (Color/BW).

The overall average shift in SE is −2.22 (SD = 13.87); 23% of these shifts are zero, 36% are above zero and 40% are below zero (Supplementary Fig. 3). Here again we see the tendency of geno2pheno$_{[coreceptor]}$ to reduce FPR when sequence alterations are present. Fig. 4 shows a plot of FPR shifts averaged by nucleotide position. Alterations in some parts of the V3 loop have a higher propensity for changing coreceptor tropism than others, as expected. The magnitude and sign of the average shifts vary greatly with the nucleotide position. Shifts averaged by nucleotide or ambiguity code can be seen in Table 2.

In LA, 6133 sequences contained ambiguous bases. The sequences in the M dataset were derived from 6118 of these sequences by resolving their ambiguities. Fifteen sequences were excluded to avoid combinatorial explosion. Among the 6133 original sequences, 41% resulted in two derived sequences while 59% resulted in more than two sequences (Supplementary Fig. 4). Supplementary Fig. 5 shows a plot of the average shift in MS against the sampling proportion, for triplicate and for singleton FPRs. The magnitude of the shift decreases as the proportion increases, and was zero for proportions of 85% of the sequences or more. The lowest proportion tested, 1%, yielded a mean singleton FPR shift of 0.61 (SD = 10.91), and a mean triplicate FPR shift of −4.5 (SD = 9.35). Since conditional CPT probabilities decrease as the sampling proportion increases, Figs. 2 and 3 only display those for sampling 1%

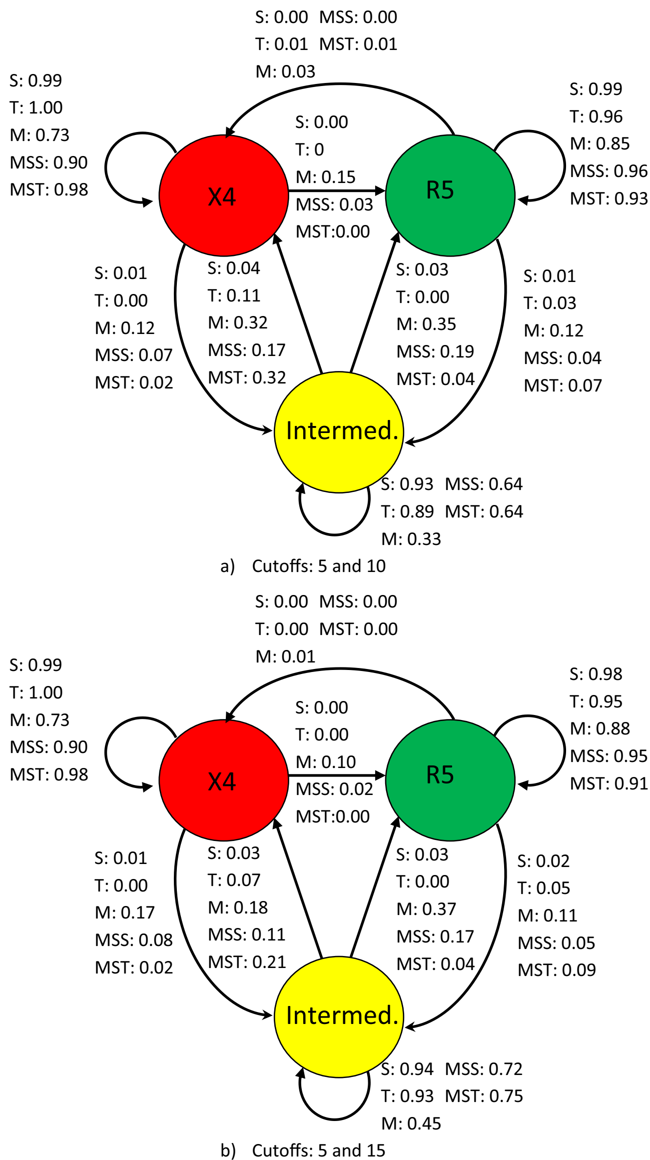

Fig. 3. Conditional probabilities of change in predicted tropism, cutoff sets {5,10} and {5,15}: Cutoff sets {5,10} and {5,15} were applied to FPRs in LA, SA, M, and MS datasets to calculate the conditional probabilities of CPT. Results are shown for SA dataset singleton (S), SA dataset triplicate (T), M dataset (M), MS dataset singleton (MSS), and MS dataset triplicate (MST) FPRs. There are no replicates on the M dataset since its sequences contain no ambiguities. The circles in the figure represent the predicted tropism. The arrows indicate a change in predicted tropism with the head of the arrow pointing towards the change. The arrow labels contain the probabilities for the respective changes in predicted tropism as calculated with the dataset mentioned above (Color/BW).

of the variants in M, labeled *MSS* for singleton FPRs and *MST* for triplicate FPRs. CPT probabilities calculated with all the sequences in M are also displayed in Figs. 2 and 3 (probabilities labeled M).

5. Discussion

This analysis addresses two related questions: the robustness of geno2pheno$_{[coreceptor]}$ with respect to sequence alterations in terms of change of predicted tropism, and the influence of minority populations on the predictions of geno2pheno$_{[coreceptor]}$. Therefore, we have subjected geno2pheno$_{[coreceptor]}$ to a challenge involving the systematic introduction of simulated sequencing variability into a large set of V3 nucleotide sequences in order to generate four datasets: SA simulates alterations due to the Sanger sequencing technique. SE explores the effects of systematically introduced single nucleotide exchanges. With M and MS we investigated the influence of simulated viral minorities undetected by bulk sequencing.

Table 2
Mean FPR shift averaged by nucleotide base, SE dataset.The table shows the mean FPR shift and standard deviation (SD), averaged by substituting nucleotide base after alignment. Shifts were calculated by comparing FPRs in the SE dataset with their unaltered counterparts in the LA dataset. Although gaps were not considered when generating the SE dataset, some nucleotide alterations were transformed to gaps by the alignment program of geno2pheno$_{[coreceptor]}$.

IUPAC base	Mean	SD
A	−3.33	20.01
C	−1.77	17.98
G	−4.85	17.49
T	−2.40	18.33
B	−2.57	12.98
D	−2.18	12.12
H	−1.77	13.26
K	−2.65	12.94
M	−1.88	13.35
N	−2.19	12.50
R	−2.86	12.72
S	−2.89	13.06
V	−2.42	12.26
W	−1.63	13.12
Y	−1.20	14.25
–	0.26	18.07

An accurate study of the effect of sequencing X4-associated codons at amino-acid positions 11, 13, 24, 25, and 32 (nucleotide positions 30–33, 37–39, 69–72, 73–75, 94–96) more accurately by using specific primers has been studied[17]. The primer sets were specifically designed for subtype B. Our study has extended to other subtypes, as about half of the 67,997 sequences were non-Bs (Table 1).

The FPR shifts between the original and the *in-silico* mutated sequences in SE have been analyzed (Fig. 4 and Table 2). 9.1% are less than −20 and 11% are greater than 10. These two shift values were chosen to be multiples of 10, close to the 10th and 90th percentile of the shift distribution. The shift cutoff −20 identifies amino-acid positions 7, 8, 11 (nucleotide positions 19–21, 22–24, 31–33), and insertions after amino-acid positions 21, 22 and 23 (nucleotide positions 61–63, 64–66 and 67–69) as highly relevant for detection of CXCR4-capable viruses with geno2pheno$_{[coreceptor]}$. Using a shift cutoff of 10, positions 9, 10, 17, 28 and 31 (nucleotide positions 25–27, 28–30, 48–51, 82–84, 91–93) are highly relevant for the detection of R5 viruses with geno2pheno$_{[coreceptor]}$. Mean shifts averaged by nucleotide base are negative, unless the sequence alteration is replaced with a gap by the alignment program. The smallest mean shift is −4.85. These results indicate that the position at which the alteration occurs is far more important than the nucleotide base the alteration consists of, as might have been expected.

Triplicate sequencing is performed with the intention of improving the detection of X4-capable minority variants. In triplicate testing, the lowest FPR is considered the correct one, without comparing the obtained nucleotide sequences. Thus, if the lowest FPR is caused by a nucleotide sequencing error, triplicate testing will bias the prediction towards favoring X4-capable. If the obtained sequences are not manually inspected, errors in the individual sequences remain undetected. Thus, the more replications are sequenced, the higher the chance you get an FPR lower than the first one, even if X4-capable minority variants are absent. In this case, triplicate testing can result in an exclusion of MVC-eligible patients, without increasing the safety of the prediction.

Figs. 2 and 3 show the influence of introduced variability for change of predicted tropism depending on the FPR cutoffs, for both

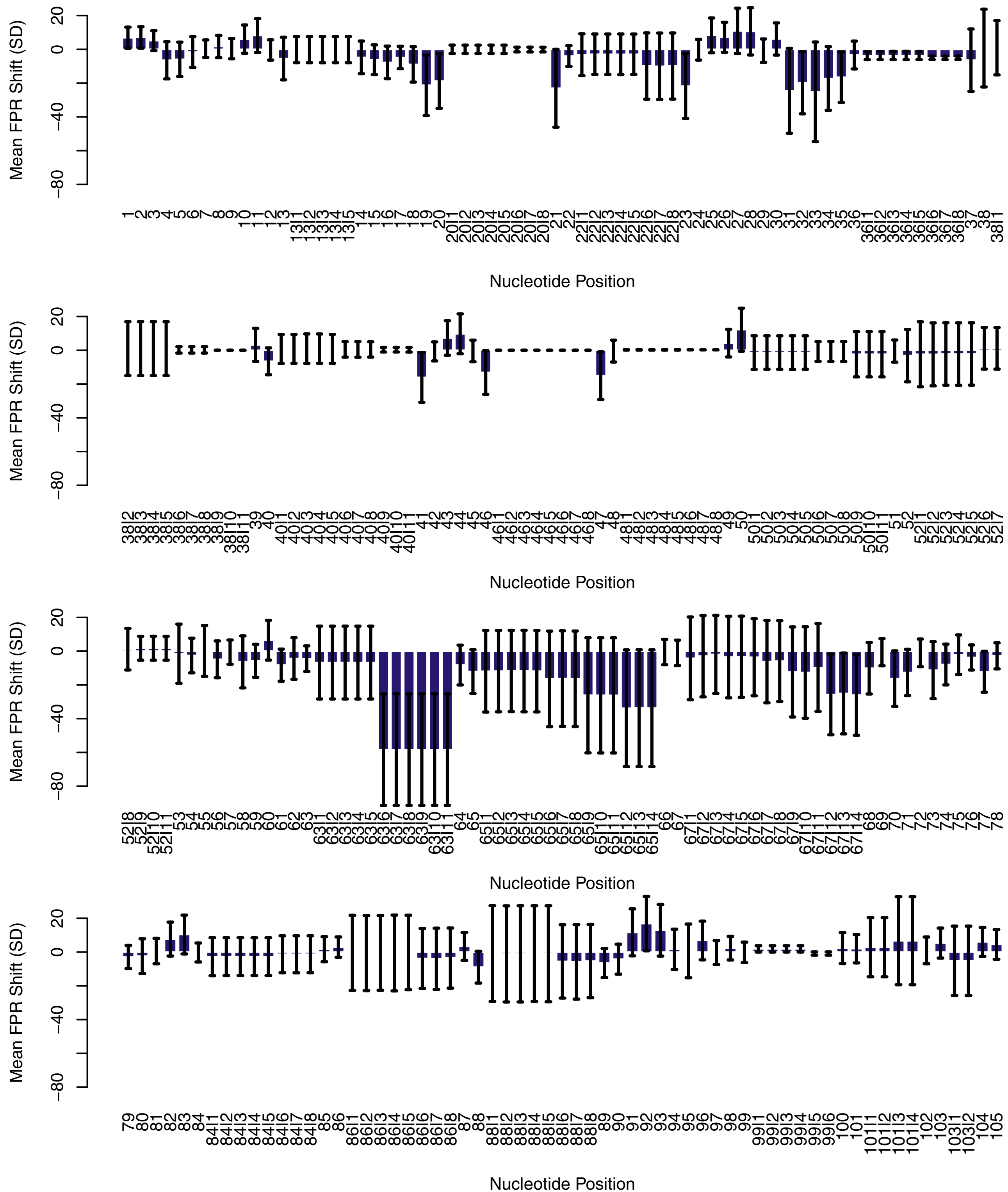

Fig. 4. Mean FPR shifts averaged by nucleotide position, SE Dataset. The graphic above shows the mean shift in the FPR averaged by nucleotide position. Shifts were calculated by comparing FPRs in the SE dataset with their unaltered counterparts in the LA dataset. Each bar represents a nucleotide position; its height indicates the mean FPR shift resulting from an alteration at that position. Error bars show standard deviation. Insertions are labeled with the preceding nucleotide position, the character "I" and an index (Color/BW).

single and triplicate testing. geno2pheno$_{[coreceptor]}$ is robust when introduced variability is present: for all cutoff sets, the probability of no CPT is above 98% (or above 93% if the intermediate tropism prediction is considered) for singleton FPRs. In this scenario, triplicate FPRs raise the probability that predicted X4 capability will not change by up to 2%, but reduce the probability that predicted R5 tropism will not change by up to 4%.

To address the study of minorities, we created the M and MS datasets. M contains the variants from which the sequence with ambiguities could have arisen, but might also contain variants

absent in the sample. In contrast, MS presents a more realistic picture, since a sequence that may contain ambiguities is constructed from a sample of the sequences in M representing the limited sequencing depth of an experiment. For the smallest fraction tested, 1%, the probability predicted X4 tropism does not change is ≥90% with singleton FPRs, and 98% with triplicate FPRs. Thus, triplicate FPRs raise the probability predicted X4 tropism does not change by 9%, at most, for the lowest tested proportion (Figs. 2 and 3).

The lowest value obtained for the probability that there is no CPT is 90% with $\text{geno2pheno}_{[\text{coreceptor}]}$. In order to put this number into context, we mention Trofile's reported sensitivity to detect MVC responders, 92% [11].

$\text{geno2pheno}_{[\text{coreceptor}]}$ proved to be robust in the presence of sequence alterations and when detectable minorities are missed by bulk sequencing. False R5 predictions were either rare or absent in our analysis, depending on the selected cutoffs. If a CPT occurs, it is much more likely that sequence alterations result in false X4-capable predictions than false R5 predictions. This speaks for the safety of MVC prescription to patients with predicted R5 viruses.

Competing interest

The Max Planck Institute for Informatics (MPI) collaborates with a company in the area of medical diagnostics on the topic of viral tropism of HIV. In its context, version-controlled implementation of $\text{geno2pheno}_{[\text{coreceptor}]}$ with secure access is maintained by MPI under a paid service agreement.

Ethical approval

Not required.

Acknowledgements

This work was partially funded by CHAIN (EU grant number 223131, Health-F3-2009-223131), and performed in the context of MASTER/HIV-HEP (German Health Ministry), EUCOHIV (www.eucohiv.org) and RESINA (German Health Ministry Grant number 310/4476).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jcv.2015.02.006.

References

[1] J.A. Levy, HIV the Pathogenesis of AIDS, 3rd ed, ASM. Press, D.C. Washington, 2007.
[2] C.M. Perry, Maraviroc: a review of its use in the management of CCR5-tropic HIV-1 infection, Drugs 118 (2010) 9–1213, http://dx.doi.org/10.2165/11203940-000000000-00000.
[3] M.C. Prosperi, L. Bracciale, M. Fabbiani, S. Di Giambenedetto, F. Razzolini, G. Meini, et al., Comparative determination of HIV-1 co-receptor tropism by enhanced sensitivity trofile, gp120 V3-loop RNA and DNA genotyping, Retrovirology 7 (2010) 56, http://dx.doi.org/10.1186/1742-4690-7-56.
[4] P. Recordon-Pinson, C. Soulie, P. Flandre, D. Descamps, M. Lazrek, C. Charpentier, et al., Evaluation of the genotypicprediction of HIV-1coreceptor use versus a phenotypic assay and correlation with the virological response to Maraviroc: the ANRS GenoTropism Study, Antimicrob. Agents Chemother. (2010) 3335–3340, http://dx.doi.org/10.1128/AAC.10.148.
[5] V. Svicher, R. D'Arrigo, C. Alteri, M. Andreoni, G. Angarano, A. Antinori, et al., Performance of genotypic tropism testing in clinical practice using the enhanced sensitivity version of trofile as reference assay: results from the OSCAR Study Group, New Microbiol. 33 (2010) 195–206.
[6] E. Poveda, R. Paredes, S. Moreno, J. Alcamí, J. Córdoba, R. Delgado, et al., Update on clinical and methodological recommendations for genotypic determination of HIV tropism to guide the usage of CCR5 antagonists, AIDS Rev. 14 (2012) 208–217.
[7] D.N.A. Trofile®, HIV Tropism Test), (n.d); http://www.monogrambio.com/hiv/tropism/trofile-dna/ (accessed 20.11.13.).
[8] T. Lengauer, O. Sander, S. Sierra, A. Thielen, R. Kaiser, Bioinformatics prediction of HIV coreceptor usage, Nat. Biotechnol. 25 (2007) 1407–1410, http://dx.doi.org/10.1038/nbt1371.
[9] J. Symons, L. Vandekerckhove, R. Paredes, C. Verhofstede, R. Bellido, E. Demecheleer, et al., Impact of triplicate testing on HIV genotypic tropism prediction in routine clinical practice, Clin. Microbiol. Infect. Off. Publ. Eur. Soc. Clin. Microbiol. Infect. Dis. 18 (2012) 606–612, http://dx.doi.org/10.1111/j. 1469-0691.2011.03631.
[10] R.A. McGovern, A. Thielen, S. Portsmouth, T. Mo, W. Dong, C.K. Woods, et al., Population-based sequencing of the V3-loop can predict the virological response to maraviroc in treatment-naive patients of the MERIT trial, J. Acquir. Immune Defic. Syndr. 61 (2012) 9–286, http://dx.doi.org/10.1097/QAI.0b013e31826249cf.
[11] R.A. McGovern, A. Thielen, T. Mo, W. Dong, C.K. Woods, D. Chapman, et al., Population-based V3 genotypic tropism assay: a retrospective analysis using screening samples from the A4001029 and MOTIVATE studies, AIDS Lond. Engl Engl. 24 (2010) 2517–2525, http://dx.doi.org/10.1097/QAD.0b013e32833e6cfb.
[12] A. Vandekerckhove, R. Wensing, B. Brun-Vezinet, A. De Luca, et al., Consensus statement of the European guidelines on clinical management of HIV-1 tropism testing, J. Int. AIDS Soc. 13 (2010) O7, http://dx.doi.org/10.1186/1758-2652-13-S4-O7.
[13] L.P.R. Vandekerckhove, A.M.J. Wensing, R. Kaiser, F. Brun-Vézinet, B. Clotet, A. De Luca, et al., European guidelines on the clinical management of HIV-1 tropism testing, Lancet Infect. Dis. 11 (2011) 394–407, http://dx.doi.org/10.1016/S1473-3099(10).
[14] H. Walter, J., Eberle, H. Müller, Empfehlung zur Bestimmung des HIV-1-Korezeptor-Gebrauchs (DAIG Recommendations for the HIV-1 Tropism Testing), (n.d.). http://www.daignet.de/site-content/hiv-therapie/leitlinien-1/Leitlinien%20zur%20Topismus_Testung%20Stand%20Juni%202009.pdf. (accessed 10. 8. 2013).
[15] R.A. McGovern, P.R. Harrigan, L.C. Swenson, Genotypic inference of HIV-1 tropism using population-based sequencing of V3, J. Vis. Exp. JoVE (2010), http://dx.doi.org/10.3791/2531.
[16] N.R. Pineda-Peña, S. Imbrechts, P. Libin, A.B. Abecasis, K. Deforche, et al., Automated subtyping of HIV-1 genetic sequences for clinical and surveillance purposes: performance evaluation of the new REGA version 3 and seven other tools, Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect Dis. (2013), http://dx.doi.org/10.1016/j.meegid.2013.04.032.
[17] S. McLaughlin, L.C. Swenson, S. Hu, P. Hughes, P.R. Harrigan, R.W. Coombs, et al., Development of a novel codon-specific polymerase chain reaction for the detection of CXCR4-Utilizing HIV Type 1 Subtype B, AIDS Res. Hum. Retroviruses. 29 (2013) 814–825, http://dx.doi.org/10.1089/aid.2012.0024.
[18] S. Sierra, R. Kaiser, N. Lübke, A. Thielen, E. Schuelter, E. Heger, et al., Prediction of HIV -1 coreceptor usage (tropism) by sequence analysis using a genotypic approach, J. Vis. Exp. JoVE (2011), http://dx.doi.org/10.3791/3264.
[19] C.K. Woods, C.J. Brumme, T.F. Liu, C.K.S. Chui, A.L. Chu, B. Wynhoven, et al., Automating HIV drug resistance genotyping with RECall a freely accessible sequence analysis tool, J. Clin. Microbiol. 50 (2012) 1936–1942, http://dx.doi.org/10.1128/JCM.6689.11.
[20] B. Ewing, P. Green, Base-calling of automated sequencer traces using phred. II. Error probabilities, Genome Res. 8 (1998) 186–194.
[21] A.D. Johnson, An extended IUPAC nomenclature code for polymorphic nucleic acids, Bioinformatics 138 (2010) 1386–1389, http://dx.doi.org/10.1093/bioinformatics/btq098.
[22] G. Myers, A dataset generator for whole genome shotgun sequencing, Proc: Int. Conf. Intell. Syst. Mol. Biol. ISMB Int. Conf. Intell. Syst. Mol. Biol. 20 (1999) 202–210.
[23] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. R. Stat. Soc. Ser. B Stat. Methodol. 57 (1995) 289–300.