BMC
Infectious Diseases

Appraising the performance of genotyping tools in the prediction of coreceptor tropism in HIV-1 subtype C viruses

Saleema Crous, Ram Krishna Shrestha and Simon A Travers[*]

Abstract

Background: In human immunodeficiency virus type 1 (HIV-1) infection, transmitted viruses generally use the CCR5 chemokine receptor as a coreceptor for host cell entry. In more than 50% of subtype B infections, a switch in coreceptor tropism from CCR5- to CXCR4-use occurs during disease progression. Phenotypic or genotypic approaches can be used to test for the presence of CXCR4-using viral variants in an individual's viral population that would result in resistance to treatment with CCR5-antagonists. While genotyping approaches for coreceptor-tropism prediction in subtype B are well established and verified, they are less so for subtype C.

Methods: Here, using a dataset comprising V3 loop sequences from 349 CCR5-using and 56 CXCR4-using HIV-1 subtype C viruses we perform a comparative analysis of the predictive ability of 11 genotypic algorithms in their prediction of coreceptor tropism in subtype C. We calculate the sensitivity and specificity of each of the approaches as well as determining their overall accuracy. By separating the CXCR4-using viruses into CXCR4-exclusive (25 sequences) and dual-tropic (31 sequences) we evaluate the effect of the possible conflicting signal from dual-tropic viruses on the ability of a of the approaches to correctly predict coreceptor phenotype.

Results: We determined that geno2pheno with a false positive rate of 5% is the best approach for predicting CXCR4-usage in subtype C sequences with an accuracy of 94% (89% sensitivity and 99% specificity). Contrary to what has been reported for subtype B, the optimal approaches for prediction of CXCR4-usage in sequence from viruses that use CXCR4 exclusively, also perform best at predicting CXCR4-use in dual-tropic viral variants.

Conclusions: The accuracy of genotyping approaches at correctly predicting the coreceptor usage of V3 sequences from subtype C viruses is very high. We suggest that genotyping approaches can be used to test for coreceptor tropism in HIV-1 group M subtype C with a high degree of confidence that they will identify CXCR4-usage in both CXCR4-exclusive and dual tropic variants.

Keywords: Human immunodeficiency virus, Coreceptor, Chemokine receptors, CXCR4, CCR5, Genotype, Phenotype, Subtype C

Background

To enable cell entry by HIV, the gp120 glycoprotein, present in a trimeric arrangement on the surface of a HIV virion, must first bind to a CD4 receptor on the target cell [1-3]. This binding induces a conformational change in the gp120/gp41 trimer complex [4,5] thereby enabling binding of a chemokine receptor, either CCR5 or CXCR4 [6]. CCR5-tropic viruses are associated with primary transmission and can persist throughout infection [6]. In as many as 50% of HIV-1 subtype B infections, a switch to CXCR4-usage has been observed and this switch is generally regarded as an indicator of disease progression [7-10]. Early studies of HIV-1 subtype C suggested that a switch to CXCR4-usage was less common in subtype C compared to subtype B [11,12], however more recent studies have suggested that between 30-50% of subtype C infected individuals exhibit a change to CXCR4-usage during disease progression [13-18].

* Correspondence: simon@sanbi.ac.za
South African National Bioinformatics Institute, University of the Western Cape, Private Bag X17, Belville 7535, South Africa

Dual-tropic viruses (R5X4) capable of using either CCR5 or CXCR4 for host cell entry have been described [19] as have dual-tropic viruses that, while capable of using either receptor for cell entry, exhibit preferential use of either CCR5 (dual-R) or CXCR4 (dual-X) [20,21]. Detecting the presence of dual-tropic viruses in an individual's viral population is difficult however, as a mixed population of R5 and X4 viruses will be identified as dual in a population-based phenotyping assay.

Determining the coreceptor usage profile of an individual's viral population has been used as an indicator of disease progression and in more recent years as an approach for detecting resistance to CCR5 antagonists such as maraviroc [22-24]. Phenotypic assays, such as Monogram Bioscience's Trofile™ assay [25], are the most effective means of elucidating the coreceptor tropism of a viral population. These approaches, however, are expensive, laborious and unavailable for routine use in all laboratories [26,27]. Thus, genotyping approaches have been suggested to be a viable alternative for routine coreceptor tropism testing [28]. While many amino acid positions throughout gp120 have been suggested to influence coreceptor affinity and tropism [29-35], the V3 loop appears to be the strongest determinant of coreceptor tropism with amino acid mutations affecting V3 net charge, charge at positions 11, 24 and 25 and glycan binding patterns all implicated in causing a switch from CCR5- to CXCR4-usage [36-41].

Early genotypic algorithms predicted the coreceptor tropism of HIV-1V3 sequences using the properties of the amino acids at positions 11 and 25 while later algorithms account for various properties of the entire V3 loop [39,40,42-45]. With the exception of C-PSSM [43] and the Raymond combined 11/25 and net charge rules [46], all of these approaches have been optimised for coreceptor tropism prediction in subtype B and show varying levels of sensitivity at predicting CXCR4-usage in subtype B [47].

Despite HIV-1 subtype C accounting for almost 60% of worldwide HIV infections [48], the genetic determinants of the switch in coreceptor use are less-well understood than in subtype B. Conflicting reports have been published with some suggesting that these determinants are the same for subtype C as subtype B [46], while others have presented evidence to the contrary [43]. Jensen and colleagues developed the only subtype C specific genotyping tool with a reported sensitivity of 75% [43] while others evaluated the ability of this and other algorithms trained on subtype B data at correctly predicting CXCR4-use in subtype C sequence data [46]. They found that the most appropriate approach for predicting CXCR4-usage in subtype C were C-PSSM and their combined 11/25 and net charge rule [46]. When specificity was considered, however, Raymond and

colleagues approach was significantly better than C-PSSM (96.4% versus 81.8%). The dataset used in this study, however, did not represent the entire spectrum of HIV-1 subtype C diversity in that it had a limited number of phenotyped sequences (55 R5 and 15 X4 sequences) collected from only two countries (Malawi and France).

In this study we have collated a large dataset consisting of all obtainable subtype C sequences with experimentally verified coreceptor tropism and used this to evaluate the performance of various genotyping tools at accurately predicting CXCR4-usage in HIV-1 subtype C. Further, we determine the effect of sequences from dual-tropic viruses on the sensitivity of genotyping methods.

Results and discussion

In total 731 HIV-1 group M subtype C V3 sequences with experimentally verified coreceptor tropism were retrieved. Only one representative sequence for each individual was retained reducing the total number of sequences to 405. The final analysis dataset (available on request) contained sequences from 349 CCR5-using and 56 CXCR4-using viruses. Sequences from CXCR4-using viruses were further separated into R5X4 (dual-tropic) and CXCR4-exclusive viruses with 31 and 25 sequences, respectively, comprising these datasets.

The coreceptor usage of every sequence in each of the datasets was predicted using all of the genotyping approaches. 23 of the sequences tested contained at least one ambiguous nucleotide position. Geno2pheno is the only one of the tools tested that is capable of accounting for ambiguous positions in its genotypic predictions [44]. To assess all of the other approaches, we translated the nucleotide sequences into all the possible combinations of amino acid sequences and if one or more of these translated sequences was predicted as CXCR4-using, the genotyping call for the original sequence was taken as X4. For each of the 23 sequences, all possible translations of the sequence had the same coreceptor tropism prediction for each method. Thus, in this data, ambiguous positions did not affect the genotypic predictions. However, in many cases the presence of ambiguous nucleotide calls, particularly within the codons encoding for amino acid positions 11, 24 and 25, would substantially reduce the ability of approaches to correctly predict coreceptor usage [44]. Thus, the ability to account for ambiguous nucleotide positions in geno2pheno gives it a distinct advantage over all of the other approaches tested here.

The sensitivity of each of the tested approaches at predicting X4 viruses in the CXCR4-using dataset (dual tropic and CXCR4-exclusive combined) varied widely from 40-97% (Table 1 and Figure 1). The method by Raymond and colleagues performed best with 97%

Table 1 Performance of genotyping approaches at predicting CXCR4-usage in viral sequences from individuals infected with HIV-1 group M subtype C

Method	CXCR4-using sensitivity (%)	Specificity
PSSM$_{sinsi}$	76	100
PSSM$_{X4R5}$	75	97
C-PSSM	90	92
Geno2Pheno$_{FPR5}$	89	99
Geno2Pheno$_{FPR10}$	89	94
Geno2Pheno$_{FPR20}$	91	86
WetCat$_{C4.5}$	40	99
WetCat$_{C4.5\ pos.\ 8\&12}$	40	100
WetCat$_{PART}$	55	100
WetCat$_{SVM}$	63	99
11/24/25	68	97
11/25	60	99
Raymond	97	76

Sensitivity corresponds to the ability of the approach to predict CXCR4-use, while specificity corresponds to the ability to correctly predict CCR5-use.

sensitivity while Geno2pheno (FPR20) and C-PSSM exhibited high sensitivities greater than 90%. Two variants of the wetcat package, C4.5 and C4.5 with p8-p12, performed most poorly with sensitivities of 40%,

consistent with previous observations on both subtype B and non-B subtypes [46,47,49].

While predicting CXCR4-usage with high accuracy is important, the ability to correctly identify R5 variants as CCR5-using is equally as important in reducing the amount of false positives that would result in incorrect clinical interpretations. Thus, we also calculated the specificity (proportion of CCR5-tropic viruses correctly predicted as R5) of each approach. All approaches performed well with three having 100% specificity, eight having specificity greater than 90% and geno2pheno (FPR20) and Raymond exhibiting lower specificity of 86% and 76% respectively (Table 1 and Figure 1). These high specificity values are consistent with previous observations in both HIV-1 subtype B and non-B subtypes that all approaches, in general, are better at correctly predicting CCR5-usage than CXCR4-usage [22,43,46,47,49].

Raymond and colleagues had previously evaluated nine of the 13 approaches studied here using a smaller, geographically limited subtype C dataset comprising 55 R5 and 15 X4 viral sequences sampled from Malawi and France [46]. They reported that the optimal approach for subtype C genotyping was a combination of the 11/25 and net charge rules with sensitivity and specificity for CXCR4-usage prediction in subtype C of 93.3% and 96.4% respectively. Using the larger and more
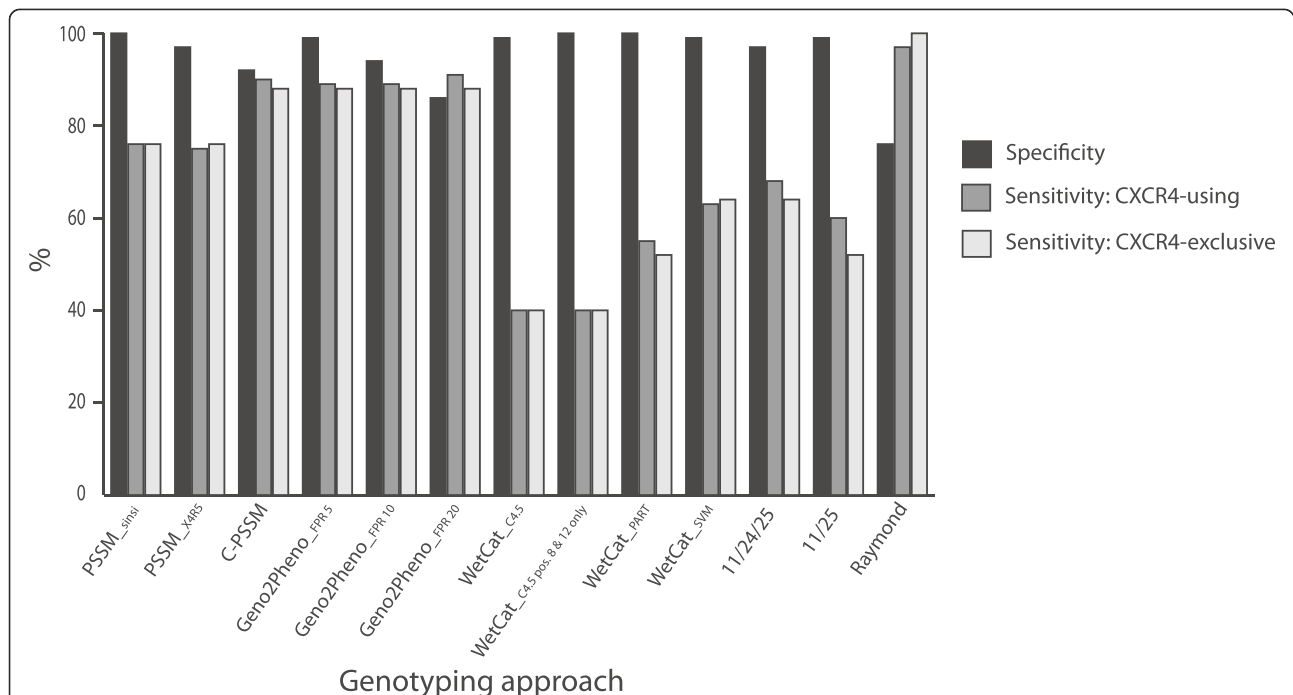

Figure 1 Performance of each of the genotyping algorithms in predicting CXCR4-usage. Sensitivity for both the CXCR4-using and CXCR4-exclusive datasets was calculated as the number of viral sequences predicted as CXCR4-using divided by the total number of CXCR4-using or CXCR4-exclusive sequences tested. Specificity corresponds to the number of CCR5-using viruses predicted as R5 divided by the total number of CCR5-using viral sequences evaluated.

geographically diverse dataset studied here, we estimate sensitivity of 97% and a specificity of 76% for this approach. Compared to the other approaches tested, however, Raymond's method is not the optimal approach. While it does show the highest sensitivity, it also has the lowest specificity of all the approaches tested (Table 1). For the other approaches we find that sensitivity increases by as much as 22% for five of the approaches relative to the Raymond study, while the sensitivity of PSSM$_{sinsi}$, PSSM$_{X4R5}$ and C-PSSM drops by 4%, 5% and 3% respectively. We suggest that the weaker performance on our comprehensive subtype C dataset of the combined 11/25 and net charge rule proposed by Raymond and colleagues is most likely an artifact of the limited sample size/diversity in their dataset that is not present in the larger dataset studied here. In describing C-PSSM, Jensen and colleagues used a dataset consisting 228 R5 sequences and 51 X4 sequences (from 200 and 20 subjects respectively) [43] and reported a sensitivity of 75%, substantially less than the 90% sensitivity reported here, with comparable specificities of 94% and 92%.

While some methods are extremely sensitive at correctly predicting CXCR4-use, the optimum approach for clinical implementation also needs to be highly specific in correctly identifying viruses that do not use CXCR4. Thus, we have calculated an accuracy score for each of the approaches tested that takes into account an approach's sensitivity and specificity (Table 2). For the CXCR4-using dataset, we find that three of the 13 approaches tested have an accuracy of 90% or greater at

Table 2 Accuracy of genotyping approaches at correctly predicting coreceptor tropism

Method	CXCR4-using accuracy	CXCR4-exclusive accuracy	R5X4 accuracy
PSSM$_{sinsi}$	88	88	88
PSSM$_{X4R5}$	86	87	86
C-PSSM	91	90	91
Geno2Pheno$_{FPR5}$	94	93	94
Geno2Pheno$_{FPR10}$	92	91	92
Geno2Pheno$_{FPR20}$	88	87	90
WetCat$_{C4.5}$	70	70	70
WetCat$_{C4.5\ pos.\ 8\&12}$	70	70	70
WetCat$_{PART}$	77	76	79
WetCat$_{SVM}$	81	82	81
11/24/25	81	81	82
11/25	79	76	82
Raymond	86	88	85

Accuracy scores are presented for a combined dataset containing CXCR4-using viruses (both CXCR4-exclusive and dual-tropic viruses) as well as separately for the CXCR4-exclusive and dual-tropic viral sequences.
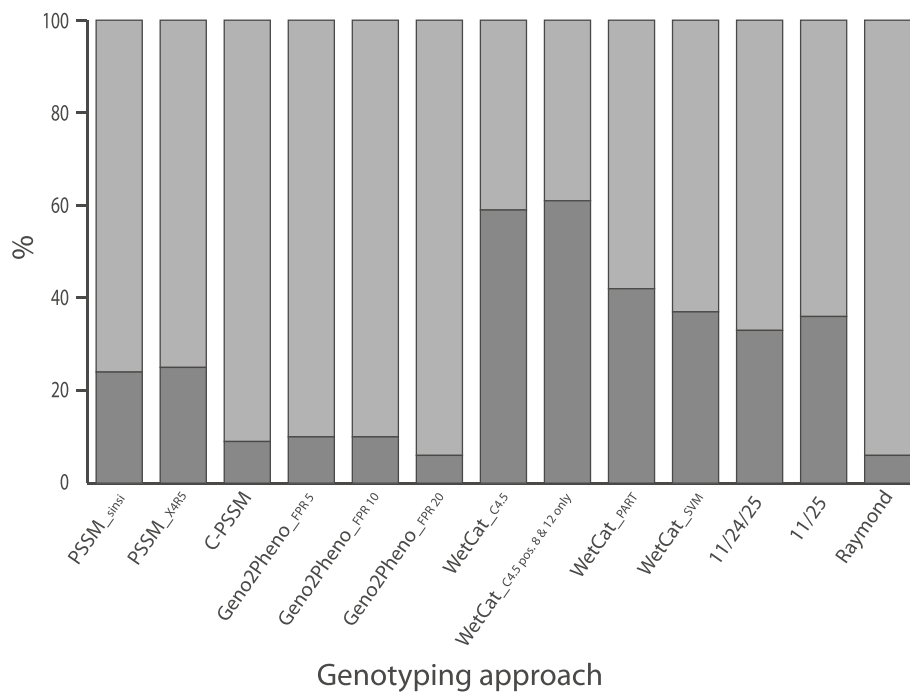
predicting coreceptor usage in HIV-1 group M subtype C viral sequences with geno2pheno (FPR5) being the most accurate of all approaches tested with an accuracy of 94% (89% sensitivity and 99% specificity, Table 2). Two variants of the wetcat package, C4.5 and C4.5 with p8-p12, both perform poorest with accuracy scores of 70% (Table 2).

Dual-tropic viruses are a unique class of viruses in that they can enter host cells using either CCR5 or CXCR4 chemokine receptors, however, some dual-tropic viruses can exhibit preferential use of one of these [19-21]. From a clinical perspective, it is imperative that genotyping approaches correctly identify the CXCR4-using capabilities of dual-tropic viruses. Genotyping algorithms have been shown to vary widely in their predictive ability of CXCR4-usage in subtype B dual-tropic viruses [50]. In general, approaches were observed to underestimate the frequency of CXCR4-usage in dual tropic viruses [50]. Thus, we sought to investigate the effect of dual-tropic viruses on the accuracy of each of the genotyping approaches tested. The CXCR4-using viruses were separated into CXCR4-exclusive and dual-tropic viral sequences and the accuracy of each of the approaches at correctly predicting coreceptor tropism was calculated (Table 2). When dual-tropic sequences are excluded, the accuracy of three of the approaches increases minimally, with four methods showing no change in accuracy and six showing a slight decrease of 1% in accuracy (Table 2). Similarly, when the dual-tropic viruses were studied separately there was minimal effect on the accuracy of each of the approaches (Table 2). There was significant variability in the ability of the approaches to accurately predict CXCR4-usage in dual-tropic viruses, ranging from 40% (wetcat C4.5 with p8-p12) to 94% (Geno2pheno FPR20) of sequences from dual-tropic viruses predicted as CXCR4-using (Figure 2). It appears that, in subtype C at least, the ability of approaches to predict CXCR4-usage in dual tropic viruses directly correlates with their ability to predict CXCR4-usage in CXCR4-exclusive viruses. Such an observation does not appear to hold true in subtype B, however, where some methods with high sensitivity for prediction of CXCR4 viruses in subtype B [47], show low accuracy for the prediction of CXCR4-usage in subtype B dual-tropic viral sequences [50]. Geno2pheno, however, does show high accuracy (90%) for the prediction of CXCR4-usage in subtype B dual-tropic viruses [50].

Conclusion

Using a comprehensive, geographically diverse dataset, we find that geno2pheno (FPR5) is the most accurate approach for the prediction of coreceptor tropism in HIV-1 subtype C viral sequences. Coupled with it's high accuracy, the ability of geno2pheno to account for

Figure 2 Ability of each approach at predicting CXCR4-usage in dual-tropic viral sequences. The percentage of dual-tropic sequences predicted as CCR5-using and CXCR4-using is shown with dark and light shaded areas of each bar corresponding to the percentage of sequences predicted as CCR5-using and CXCR4-using respectively.

ambiguous nucleotide calls in V3 sequences gives it a distinct advantage over all other approaches for coreceptor genotyping of sequence data generated from population-based sequencing. We also report that in HIV-1 group M subtype C, sequences from dual-tropic viruses have minimal effect on accuracy calculations and the optimal approaches for prediction of CXCR4-usage in sequence from viruses that use CXCR4 exclusively also perform best at predicting CXCR4-use in dual-tropic viral variants. Based on this work we suggest that viral genotyping of envelope sequences from subtype C infected individuals is feasible with the correct approach and can be undertaken with a high degree of confidence that CXCR4-usage will be accurately identified in both CXCR4-exclusive and dual tropic variants.

Methods
Study data
A dataset consisting of 731 HIV-1 subtype C V3 nucleotide sequences with phenotypically determined coreceptor tropism was sourced. The majority of sequences were retrieved from the HIV LANL Sequence Database (hiv.lanl.gov), with the remainder originating from published literature [43,46]. Multiple sequence alignments were produced manually using MacClade 4.08 [51]. To avoid potential bias in results, multiple samples from the same individuals were excluded with a single representative sequence randomly selected for these individuals.

Genotypic algorithms
The coreceptor tropism of each V3 sequence was predicted using a number of genotyping methods. These comprised $PSSM_{X4R5}$ and $PSSM_{SINSI}$ [42] as well as the subtype C PSSM tool [43], geno2pheno [44] and four variants (C4.5, C4.5 with p8-p12, PART and SVM) of the wetcat package [45]. Tropism was also predicted using the 11/25 [39] and 11/24/25 [40] rules (software implementation available from the corresponding author on request). Raymond and colleagues recently proposed a combination of the 11/25 and charge rules for prediction of CXCR4-use in subtype C sequences [46]. One of the following criteria is required for predicting CXCR4 coreceptor usage: (1) 11 R/K and/or 25K, (2) 25R and a net charge of ≥ +5, or (3) a net charge of ≥ +6 [46]. For geno2pheno, three different false positive rates (5%, 10% and 20%) were used to determine the optimal parameters for the accurate prediction of CXCR4-usage in subtype C viral sequences. Each of the geno2pheno false positive rates used is described as an individual approach throughout the paper for clarity purposes. Sequences duplicated by each of the PSSM tools (as a result of more than one optimal alignment to the reference sequence) were only considered for further analysis when genotypic predictions made by the matrix were the same for all alignment variations. The presence of ambiguous nucleotide calls in a sequence can affect the accuracy of genotyping approaches [44]. Thus, if a tested genotyping

approach was not designed to account for ambiguous nucleotide positions, all possible combinations of amino acid sequences were output and a worst-case scenario approach was employed whereby if one of these translated sequences was predicted as CXCR4-using, the genotyping call for the original sequence was taken as X4.

Viral sequences were separated into three distinct categories (R5, X4 and R5X4) based upon their experimentally verified viral phenotype. Dual-tropic and CXCR4-tropic viruses were studied both separately and together (as CXCR4-using) in order to determine the affect of the conflicting signal of dual-tropic viruses on sensitivity estimates. The sensitivity of each approach for CXCR4 prediction was calculated as the number of predicted X4 viruses in the CXCR4-using dataset divided by the total number of sequences in the CXCR4-using dataset. The specificity of each approach for CXCR4 prediction was calculated as the number of predicted R5 viruses in the CCR5-using dataset divided by the total number of sequences in the CCR5-using dataset. The same method was used to calculate the sensitivity and specificity of each genotyping method on the CXCR4-exclusive and dual-tropic datasets.

Further, an overall accuracy score for each of the approaches used was calculated using:

$$\frac{TP + TN}{TP + TN + FP + FN}$$

where, for the CXCR4-using dataset, TP corresponds to the number of CXCR4-using sequences predicted as CXCR4-using, TN the number of R5 sequences predicted as CCR5-using, FP the number of R5 sequences predicted as CXCR4-using and FN the number of CXCR4-using sequences predicted as CCR5-using. For the CXCR4-exclusive dataset the TP and FN values were calculated only for sequences phenotypically determined to exclusively use CXCR4. For each calculation we normalized the TP and FN values relative to the TN and FP values to account for the disproportionate number of sequences representing the positive (CXCR4-using or CXCR4-exclusive) and negative (CCR5) datasets (see Additional file 1: Table S1 for the uncorrected values).

Additional file

Additional file 1: Table S1. Tables detailing the uncorrected numbers of true positives (CXCR4-usage correctly predicted in CXCR4-using sequences), true negatives (CCR5-usage correctly predicted in CCR5-using sequences), false positives (CXCR4-usage incorrectly predicted in CCR5-using sequences) and false negatives (CCR5-usage incorrectly predicted in CXCR4-using sequences) predicted by each of the approaches. Results are shown for (A) CXCR4-using sequences, (B) CXCR4-exclusive sequences and (C) dual-tropic sequences.

Authors' contributions
SC collated the dataset, tested the various genotyping approaches and wrote the first version of the manuscript. RKS wrote software for the translation of ambiguous nucleotides and for the approach by Raymond and colleagues as well as collating results from a number of the approaches used. SAAT conceived the study, participated in its design and wrote the final manuscript. All authors read and approved the final manuscript.

References
1. Dalgleish AG, Beverley PC, Clapham PR, Crawford DH, Greaves MF, Weiss RA: **The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus.** *Nature* 1984, **312**(5996):763–767.
2. Maddon PJ, Dalgleish AG, McDougal JS, Clapham PR, Weiss RA, Axel R: **The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain.** *Cell* 1986, **47**(3):333–348.
3. McDougal JS, Maddon PJ, Dalgleish AG, Clapham PR, Littman DR, Godfrey M, Maddon DE, Chess L, Weiss RA, Axel R: **The T4 glycoprotein is a cell-surface receptor for the AIDS virus.** *Cold Spring Harb Symp Quant Biol* 1986, **51**(Pt 2):703–711.
4. Sattentau QJ, Moore JP: **Conformational changes induced in the human immunodeficiency virus envelope glycoprotein by soluble CD4 binding.** *J Exp Med* 1991, **174**(2):407–415.
5. Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, Subramaniam S: **Molecular architecture of native HIV-1 gp120 trimers.** *Nature* 2008, **455**(7209):109–113.
6. Dragic T, Litwin V, Allaway GP, Martin SR, Huang Y, Nagashima KA, Cayanan C, Maddon PJ, Koup RA, Moore JP, *et al*: **HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5.** *Nature* 1996, **381**(6584):667–673.
7. Koot M, Keet IP, Vos AH, de Goede RE, Roos MT, Coutinho RA, Miedema F, Schellekens PT, Tersmette M: **Prognostic value of HIV-1 syncytium-inducing phenotype for rate of CD4+ cell depletion and progression to AIDS.** *Ann Intern Med* 1993, **118**(9):681–688.
8. Hazenberg MD, Otto SA, Hamann D, Roos MT, Schuitemaker H, de Boer RJ, Miedema F: **Depletion of naive CD4 T cells by CXCR4-using HIV-1 variants occurs mainly through increased T-cell death and activation.** *AIDS* 2003, **17**(10):1419–1424.
9. Levine B, Sodora DL: **HIV and CXCR4 in a kiss of autophagic death.** *J Clin Invest* 2006, **116**(8):2078–2080.
10. Connor RI, Sheridan KE, Ceradini D, Choe S, Landau NR: **Change in coreceptor use correlates with disease progression in HIV-1–infected individuals.** *J Exp Med* 1997, **185**(4):621–628.
11. Abebe A, Demissie D, Goudsmit J, Brouwer M, Kuiken CL, Pollakis G, Schuitemaker H, Fontanet AL, Rinke de Wit TF: **HIV-1 subtype C syncytium- and non-syncytium-inducing phenotypes and coreceptor usage among Ethiopian patients with AIDS.** *AIDS* 1999, **13**(11):1305–1311.
12. Pollakis G, Abebe A, Kliphuis A, Chalaby MI, Bakker M, Mengistu Y, Brouwer M, Goudsmit J, Schuitemaker H, Paxton WA: **Phenotypic and genotypic comparisons of CCR5- and CXCR4-tropic human immunodeficiency virus type 1 biological clones isolated from subtype C-infected individuals.** *J Virol* 2004, **78**(6):2841–2852.

13. Connell BJ, Michler K, Capovilla A, Venter WD, Stevens WS, Papathanasopoulos MA: **Emergence of X4 usage among HIV-1 subtype C: evidence for an evolving epidemic in South Africa.** *AIDS* 2008, **22**(7):896–899.

14. Kassaye S, Johnston E, McColgan B, Kantor R, Zijenah L, Katzenstein D: **Envelope coreceptor tropism, drug resistance, and viral evolution among subtype C HIV-1-infected individuals receiving nonsuppressive antiretroviral therapy.** *J Acquir Immune Defic Syndr* 2009, **50**(1):9–18.

15. Michler K, Connell BJ, Venter WD, Stevens WS, Capovilla A, Papathanasopoulos MA: **Genotypic characterization and comparison of full-length envelope glycoproteins from South African HIV type 1 subtype C primary isolates that utilize CCR5 and/or CXCR4.** *AIDS Res Hum Retroviruses* 2008, **24**(5):743–751.

16. Cilliers T, Nhlapo J, Coetzer M, Orlovic D, Ketas T, Olson WC, Moore JP, Trkola A, Morris L: **The CCR5 and CXCR4 coreceptors are both used by human immunodeficiency virus type 1 primary isolates from subtype C.** *J Virol* 2003, **77**(7):4449–4456.

17. Papathanasopoulos MA, Cilliers T, Morris L, Mokili JL, Dowling W, Birx DL, McCutchan FE: **Full-length genome analysis of HIV-1 subtype C utilizing CXCR4 and intersubtype recombinants isolated in South Africa.** *AIDS Res Hum Retroviruses* 2002, **18**(12):879–886.

18. Johnston ER, Zijenah LS, Mutetwa S, Kantor R, Kittinunvorakoon C, Katzenstein DA: **High frequency of syncytium-inducing and CXCR4-tropic viruses among human immunodeficiency virus type 1 subtype C-infected patients receiving antiretroviral treatment.** *J Virol* 2003, **77** (13):7682–7688.

19. Berger EA, Doms RW, Fenyo EM, Korber BT, Littman DR, Moore JP, Sattentau QJ, Schuitemaker H, Sodroski J, Weiss RA: **A new classification for HIV-1.** *Nature* 1998, **391**(6664):240.

20. Huang W, Eshleman SH, Toma J, Fransen S, Stawiski E, Paxinos EE, Whitcomb JM, Young AM, Donnell D, Mmiro F, *et al*: **Coreceptor tropism in human immunodeficiency virus type 1 subtype D: high prevalence of CXCR4 tropism and heterogeneous composition of viral populations.** *J Virol* 2007, **81**(15):7885–7893.

21. Huang W, Eshleman SH, Toma J, Stawiski E, Whitcomb JM, Jackson JB, Guay L, Musoke P, Parkin N, Petropoulos CJ: **Vertical transmission of X4-tropic and dual-tropic HIV-1 in five Ugandan mother-infant pairs.** *AIDS* 2009, **23** (14):1903–1908.

22. Poveda E, Seclen E, Gonzalez Mdel M, Garcia F, Chueca N, Aguilera A, Rodriguez JJ, Gonzalez-Lahoz J, Soriano V: **Design and validation of new genotypic tools for easy and reliable estimation of HIV tropism before using CCR5 antagonists.** *J Antimicrob Chemother* 2009, **63**(5):1006–1010.

23. Poveda E, Briz V, Quinones-Mateu M, Soriano V: **HIV tropism: diagnostic tools and implications for disease progression and treatment with entry inhibitors.** *AIDS* 2006, **20**(10):1359–1367.

24. Collins S, i-Base H: **Treatment failure and tropism changes in maraviroc trial related to previously undetected CXCR4, rather than a mutational shift from CCR5.** In *XVI International HIV Drug Resistance Workshop*. Barbados: HIV Treatment Bulletin; 2007.

25. Whitcomb JM, Huang W, Fransen S, Limoli K, Toma J, Wrin T, Chappey C, Kiss LD, Paxinos EE, Petropoulos CJ: **Development and characterization of a novel single-cycle recombinant-virus assay to determine human immunodeficiency virus type 1 coreceptor tropism.** *Antimicrob Agents Chemother* 2007, **51**(2):566–575.

26. Prosperi MC, Bracciale L, Fabbiani M, Di Giambenedetto S, Razzolini F, Meini G, Colafigli M, Marzocchetti A, Cauda R, Zazzi M, *et al*: **Comparative determination of HIV-1 co-receptor tropism by enhanced sensitivity trofile, gp120 V3-loop RNA and DNA genotyping.** *Retrovirology* 2010, **7**:56.

27. Sierra S, Kaiser R, Thielen A, Lengauer T: **Genotypic coreceptor analysis.** *Eur J Med Res* 2007, **12**(9):453–462.

28. McGovern RA, Thielen A, Mo T, Dong W, Woods CK, Chapman D, Lewis M, James I, Heera J, Valdez H, *et al*: **Population-based V3 genotypic tropism assay: a retrospective analysis using screening samples from the A4001029 and MOTIVATE studies.** *AIDS* 2010, **24**(16):2517–2525.

29. Rizzuto C, Sodroski J: **Fine definition of a conserved CCR5-binding region on the human immunodeficiency virus type 1 glycoprotein 120.** *AIDS Res Hum Retroviruses* 2000, **16**(8):741–749.

30. Rizzuto CD, Wyatt R, Hernandez-Ramos N, Sun Y, Kwong PD, Hendrickson WA, Sodroski J: **A conserved HIV gp120 glycoprotein structure involved in chemokine receptor binding.** *Science* 1998, **280** (5371):1949–1953.

31. Boyd MT, Simpson GR, Cann AJ, Johnson MA, Weiss RA: **A single amino acid substitution in the V1 loop of human immunodeficiency virus type 1 gp120 alters cellular tropism.** *J Virol* 1993, **67**(6):3649–3652.

32. Bergeron L, Sullivan N, Sodroski J: **Target cell-specific determinants of membrane fusion within the human immunodeficiency virus type 1 gp120 third variable region and gp41 amino terminus.** *J Virol* 1992, **66** (4):2389–2397.

33. Ross TM, Cullen BR: **The ability of HIV type 1 to use CCR-3 as a coreceptor is controlled by envelope V1/V2 sequences acting in conjunction with a CCR-5 tropic V3 loop.** *Proc Natl Acad Sci U S A* 1998, **95**(13):7682–7686.

34. Hoffman NG, Seillier-Moiseiwitsch F, Ahn J, Walker JM, Swanstrom R: **Variability in the human immunodeficiency virus type 1 gp120 Env protein linked to phenotype-associated changes in the V3 loop.** *J Virol* 2002, **76**(8):3852–3864.

35. Nabatov AA, Pollakis G, Linnemann T, Kliphius A, Chalaby MI, Paxton WA: **Intrapatient alterations in the human immunodeficiency virus type 1 gp120 V1V2 and V3 regions differentially modulate coreceptor usage, virus inhibition by CC/CXC chemokines, soluble CD4, and the b12 and 2 G12 monoclonal antibodies.** *J Virol* 2004, **78**(1):524–530.

36. Clevestig P, Pramanik L, Leitner T, Ehrnst A: **CCR5 use by human immunodeficiency virus type 1 is associated closely with the gp120 V3 loop N-linked glycosylation site.** *J Gen Virol* 2006, **87**(Pt 3):607–612.

37. Pollakis G, Kang S, Kliphuis A, Chalaby MI, Goudsmit J, Paxton WA: **N-linked glycosylation of the HIV type-1 gp120 envelope glycoprotein as a major determinant of CCR5 and CXCR4 coreceptor utilization.** *J Biol Chem* 2001, **276**(16):13433–13441.

38. Polzer S, Dittmar MT, Schmitz H, Schreiber M: **The N-linked glycan g15 within the V3 loop of the HIV-1 external glycoprotein gp120 affects coreceptor usage, cellular tropism, and neutralization.** *Virology* 2002, **304** (1):70–80.

39. Fouchier RA, Groenink M, Kootstra NA, Tersmette M, Huisman HG, Miedema F, Schuitemaker H: **Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule.** *J Virol* 1992, **66**(5):3183–3187.

40. Cardozo T, Kimura T, Philpott S, Weiser B, Burger H, Zolla-Pazner S: **Structural basis for coreceptor selectivity by the HIV type 1 V3 loop.** *AIDS Res Hum Retroviruses* 2007, **23**(3):415–426.

41. Resch W, Hoffman N, Swanstrom R: **Improved success of phenotype prediction of the human immunodeficiency virus type 1 from envelope variable loop 3 sequence using neural networks.** *Virology* 2001, **288**(1):51–62.

42. Jensen MA, Li FS, Van't Wout AB, Nickle DC, Shriner D, He HX, McLaughlin S, Shankarappa R, Margolick JB, Mullins JI: **Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of human immunodeficiency virus type 1 env V3 loop sequences.** *J Virol* 2003, **77**(24)):13376–13388.

43. Jensen MA, Coetzer M, Van'tWout AB, Morris L, Mullins JI: **A reliable phenotype predictor for human immunodeficiency virus type 1 subtype C based on envelope V3 sequences.** *J Virol* 2006, **80**(10):4698–4704.

44. Sing T, Low AJ, Beerenwinkel N, Sander O, Cheung PK, Domingues FS, Buch J, Daumer M, Kaiser R, Lengauer T, *et al*: **Predicting HIV coreceptor usage on the basis of genetic and clinical covariates.** *Antivir Ther* 2007, **12** (7):1097–1106.

45. Pillai S, Good B, Richman D, Corbeil J: **A new perspective on V3 phenotype prediction.** *AIDS Res Hum Retroviruses* 2003, **19**(2):145–149.

46. Raymond S, Delobel P, Mavigner M, Ferradini L, Cazabat M, Souyris C, Sandres-Saune K, Pasquier C, Marchou B, Massip P, *et al*: **Prediction of HIV type 1 subtype C tropism by genotypic algorithms built from subtype B viruses.** *J Acquir Immune Defic Syndr* 2010, **53**(2):167–175.

47. Garrido C, Roulet V, Chueca N, Poveda E, Aguilera A, Skrabal K, Zahonero N, Carlos S, Garcia F, Faudon JL, *et al*: **Evaluation of eight different bioinformatics tools to predict viral tropism in different human immunodeficiency virus type 1 subtypes.** *J Clin Microbiol* 2008, **46**(3):887–891.

48. Requejo HI: **Worldwide molecular epidemiology of HIV.** *Rev Saude Publica* 2006, **40**(2):331–345.

49. Seclen E, Garrido C, Gonzalez Mdel M, Gonzalez-Lahoz J, de Mendoza C, Soriano V, Poveda E: **High sensitivity of specific genotypic tools for detection of X4 variants in antiretroviral-experienced patients suitable to be treated with CCR5 antagonists.** *J Antimicrob Chemother* 2010, **65** (7)):1486–1492.

50. Mefford ME, Gorry PR, Kunstman K, Wolinsky SM, Gabuzda D: **Bioinformatic prediction programs underestimate the frequency of CXCR4 usage by R5X4 HIV type 1 in brain and other tissues.** *AIDS Res Hum Retroviruses* 2008, **24**(9):1215–1220.
51. Maddison WP, Maddison DR: **MacClade.**. 45th editionSunderland: Sinauer; 1992.