Performance of genotypic algorithms for predicting tropism of HIV-1CRF02_AG subtype

C. Soulié, D.B. Fofana, N. Boukli, S. Sayon, S. Lambert-Niclot, Marc Wirden, A. Simon, C. Katlama, V. Calvez, P.M. Girard, et al.

Performance of genotypic algorithms for predicting tropism of HIV-1CRF02_AG subtype

C. Soulié[a, b], D.B. Fofana[a, c], N. Boukli[c], S. Sayon[a, b], S. Lambert-Niclot[a, b], Marc Wirden[a, b], A. Simon[d], C. Katlama[e], V. Calvez[a, b], P.M. Girard[f], A.G. Marcelin[a, b], L. Morand-Joubert[a, c]

[a] Sorbonne Universités, UPMC Univ Paris 06, INSERM, Institut Pierre Louis d'épidémiologie et de Santé Publique (IPLESP UMRS 1136), F75013 Paris, France

[b] AP-HP, Groupe hospitalier Pitié Salpêtrière, Laboratoire de Virologie, Paris F-75013, France

[c] AP-HP, Hôpital Saint-Antoine, Service de Virologie, Paris F-75013, France

[d] AP-HP, Hôpital Pitié-Salpêtrière, Service de Médecine Interne, Paris F-75013, France

[e] AP-HP, Hôpital Pitié-Salpêtrière, Service de Maladies Infectieuses, Paris F-75013, France

[f] AP-HP, Hôpital Saint Antoine, Service de Maladies Infectieuses, Paris F-75013, France

ABSTRACT

Background: Several genotypic rules for predicting HIV-1 non-B subtypes tropism are commonly used, but there is no consensus about their performances.

Objectives: Three genotypic methods were compared for CRF02_AG HIV-1 tropism determination.

Study design: V3 *env* region of 178 HIV-1 CRF02_AG from Pitié-Salpêtrière and Saint-Antoine Hospitals was sequenced from plasma HIV-1 RNA. HIV-1 tropism was determined by Geno2Pheno algorithm, false positive rate (FPR) 5% or 10%, the 11/25 rule or the combined criteria of the 11/25 and net charge rule.

Results: A concordance of 91.6% was observed between Geno2pheno 5% and the combined criteria. The results were nearly similar for the comparison between Geno2pheno 5% and the 11/25 rule. More mismatches were observed when Geno2pheno was used with the FPR 10%. A lower nadir CD4 cell count was associated with a discordance of tropism prediction between Geno2pheno 5% and the combined criteria or the 11/25 rule (p=0.02 and p=0.03, respectively). A lower HIV-1 viral load was associated with some discordance for the comparison of Geno2pheno 10% and the combined rule (p=0.02).

Conclusion: Geno2pheno FPR 5% or 10% predicted more X4-tropic viruses for this set of CRF02_AG sequences than the combined criteria or the 11/25 rule alone. Furthermore, Geno2pheno FPR 5% was more concordant with the 11/25 rule and the combined rule than Geno2pheno 10% to predict HIV-1 tropism. Overall, Geno2pheno 5% could be used to predict CRF02_AG tropism as well as other genotypic rules.

Keywords: HIV tropism; non B subtype; genotypic prediction.

BACKGROUND

In the natural history of HIV disease, R5-tropic viruses were evidenced in the earlier stages of the disease and X4-tropic viruses emerged in later stages. The presence of X4-tropic HIV may have clinical impact on the disease progression. [1–3] Furthermore, a class of antiretroviral, the CCR5 inhibitors like maraviroc, is only active in patients harbouring exclusively R5-tropic viruses which implies a tropism determination before their prescription. Then, the determination of tropism is useful in clinical practise.

First genotypic methods used very simple rules, such as the 11/25 rule, predicting X4 on the basis of the presence of basic residues 11 or 25 of the V3-loop. The combined criteria of the 11/25 and net charge rules were also use to determine HIV tropism. [4] Several bioinformatics methods for prediction of HIV coreceptor usage have been proposed over the years, several are available as online tools and the most used of them was Geno2pheno [coreceptor]. [5] The system is restricted to using the V3-loop as viral sequence information input and used support vector machines. The values have been achieved on clonal sequences data.

Usually, the bioinformatics tools available to determine HIV tropism were built on phenotypic-genotypic correlation based on HIV B subtype samples. The genotypic

determination is now commonly used to determine tropism for HIV-1 B subtype, but the performances of various genotypic rules or algorithms for predicting tropism of HIV-1 non-B subtypes are still discussed. [4,6,7] It is of importance, because the prevalence of non B subtypes is increasing across Europe and the CRF02_AG recombinant is the predominant non B subtype in France. [5]

OBJECTIVES

The aim of this study was to compare 3 genotypic methods on a set of 178 HIV-1 CRF02_AG recombinants to evaluate the concordance between these measures of HIV-1 tropism.

STUDY DESIGN

This study has enrolled 178 HIV-1 patients infected by CRF02_AG recombinant from two HIV clinical centers (Pitié-Salpêtrière and Saint-Antoine Hospitals) between September 2009 and June 2014. All viruses were identified as HIV-1 CRF02_AG recombinant by Smartgene algorithm (Smartgene®, Switzerland) or by phylogenetic analyses, by estimating the relationships among RT sequences and reference sequences of HIV-1 genetic subtypes and circulating recombinant forms (CRF) obtained from the Los Alamos Database.

For the comparison of sequences, 180 HIV-1 patients infected by B subtype were studied in the same period in the two clinical centers.

The V3 *env* region was sequenced from plasma HIV-1 RNA as previously described and tropism was determined by Geno2pheno algorithm (False Positive Rate, FPR 5% or FPR 10%, http://coreceptor.bioinf.mpi-inf.mpg.de/), the 11/25 rule or the combined criteria of the

11/25 and net charge rule: CXCR4 coreceptor usage if R or K at position 11 of V3 and/or K at position 25, or R at position 25 of V3 and a net charge of ≥5, or a net charge of ≥6. The V3 net charge was calculated by subtracting the number of negatively charged amino acids (D and E) from the number of positively charged ones (K and R). [8] A consensus sequence was built with the predominant amino acid at each position of the V3 loop of gp120.

Comparisons between groups were performed using the non-parametric Mann-Whitney or chi-squared tests. Statview software v5.0 was used. Cohen Kappa coefficient is used to measure the inter-rate agreement for HIV tropism prediction.

RESULTS

The CRF02_AG HIV-1 patients had a median plasma HIV-1 viral load of 4.39 log copies/mL (range 1.58-7), a median CD4 cell count of 366 cells/mm^3 (range 5-1392) and a median of nadir CD4 cell count of 214 cells/mm^3 (range 1-1129). Approximately, half of them were receiving an antiretroviral treatment without CCR5 antagonists. The different genotypic algorithms predicted R5-tropic viruses as follows: 89.3%, 83.1%, 93.2%, and 91.0% for Geno2pheno FPR 5%, Geno2pheno FPR 10%, 11/25 rule and 11/25 rule combined with net charge rule, respectively.

For CRF02_AG recombinant tropism prediction, a concordance of 91.6% (163/178) was observed between Geno2pheno FPR 5% and the combined criteria (Figure 1A). Of the 178 samples, 154 (86.5%) and 9 (5.1%) were identified by both rules as R5- and X4-tropic viruses, respectively. Then, 5 (2.8%) were identified as R5 by Geno2pheno FPR 5% but as X4 by the combined criteria, and 10 (5.6%) were identified as X4 by Geno2pheno FPR 5% but R5 by the combined rule. The results are similar for the comparison between Geno2pheno

FPR 5% and the 11/25 rule in terms of R5 prediction (154 and 5 sequences were identified R5 by both rules and only by Geno2pheno FPR 5%, respectively) (figure 1A). The prediction of X4 tropic viruses slightly differed between the 11/25 rule and the combined criteria as revealed by the specificity score: 42.1% vs 47.2%.

Of the 178 samples, 143 (80.3%) and 11 (6.2%) were identified together by Geno2pheno FPR 10% and the combined criteria rule as R5- and X4-tropic viruses, respectively. Then, 24 sequences were not predicted for the same tropism by these two rules: 19 were identified as R5-tropic viruses by Geno2pheno FPR 10% and X4-tropic viruses by the combined rule; 5 were identified as X4-tropic viruses by Geno2pheno FPR 10% and R5-tropic viruses by the combined rule (figure 1B). For the comparison of Geno2pheno FPR 10% with the 11/25 rule, a discordance was observed 21 and 3 sequences predicted X4- and R5-tropic by Geno2pheno FPR5%, respectively (figure 1B). The specificity score was lower for the comparison with the 11/25 rule than the combined criteria (32.2% vs 36.6%) as demonstrated with Geno2pheno FPR 5% (figure 1). Overall, the Cohen Kappa coefficient evidenced a better correlation between Geno2pheno FPR 5% and the combined criteria (figure 1).

A lower nadir CD4 cell count was associated with a discordance of the tropism prediction between geno2pheno FPR5% and the combined criteria or the 11/25 rule ($p=0.0211$ and $p=0.0305$, respectively). A lower HIV-1 viral load was associated with the discordance for the comparison of geno2pheno 10% and the combined rule ($p=0.0235$).

A consensus sequence was established with the main amino acid at each of the 35 positions of the V3 loop for B subtypes and CRF02_AG subtypes (table 1A). There was no difference between both groups of patients (B *versus* CRF02_AG) in terms of CD4 ($p=0.0948$), plasma viral load ($p=0.0696$) and antiretroviral treatment ($p=0.2397$). Overall

polymorphisms were the same for both subtypes. We studied in particular the position 11 and 25 involved in the 11/25 and combined rule; the presence of an arginine codon was statistically different between B and CRF02_AG subtypes (p=0.0385 and p=0.0452 for 11 and 25 position, respectively) (table 1B).

DISCUSSION

The HIV tropism determination can be slightly different according to the algorithm used (Geno2pheno FPR 5%, Geno2pheno FPR 10%, 11/25 rule, 11/25 rule combined to net charge rule) for this set of CRF02_AG recombinant sequences. The Geno2pheno algorithm FPR 5 or 10% predicted more X4-tropic viruses than the 11/25 rule alone or combined to net charge. Then, Geno2pheno seems to be more careful concerning the prediction of tropism in regards to the prescription of CCR5 antagonists.

The bioinformatics tools available to determine HIV tropism were built on phenotypic-genotypic correlation using HIV-1 B subtype samples. Thus, HIV genetic variability between B subtype and CRF02_AG recombinant can have an impact on the tropism determination by genotypic algorithms. Indeed, the V3 region of gp120 gene is known to be highly variable in non B subtypes. [9] It is also known that sensitivity rates of genotypic approaches in non-B subtypes seem to vary between studies, probably reflecting differences in panel of viruses included. [4,8,10,11]

Even the presence of an arginine codon involved in the 11/25 rule was statistically different between B and CRF02_AG subtypes in the present study, the sensitivity of predictive tools was not improved by using simple rules such as 11/25 or 11/25 combined to net charge in comparison with Geno2pheno. Then, the geno2pheno algorithm could be used to predict tropism of CRF02_AG recombinants.

A potential impact of HIV-1 subtype on MVC activity was suggested in the MERIT and MOTIVATE studies. [12–14] It is still unclear if this could be due to different intrinsic maraviroc activity between B and non-B subtypes or lack of sensitivity or specificity in the methods used for the determination of tropism. Whatever, it is necessary to determine tropism in all HIV subtypes with simple and economic methods. It would be interesting to study the clinical response of an antiretroviral treatment including maraviroc regarding the used algorithm to determine HIV tropism.

In summary, Geno2Pheno predicted more X4-tropic viruses for this set of CRF02_AG sequences than the combined criteria or the 11/25 rule alone. Furthermore, Geno2pheno FPR 5% was more concordant with the 11/25 rule and the combined rule than Geno2pheno FPR 10% to predict HIV-1 tropism. Overall, in clinical practice, Geno2pheno FPR 5% could be used to predict CRF02_AG tropism as for B subtype.

ACKNOWLEDGMENT

This study was supported by the ANRS (National French Agency for AIDS Research).

CONFLICTS OF INTEREST STATMENT

None to declare.

BIBLIOGRAPHY

1. Waters L, Mandalia S, Randell P, Wildfire A, Gazzard B, Moyle G. The impact of HIV tropism on decreases in CD4 cell count, clinical progression, and subsequent response to a first antiretroviral therapy regimen. *Clin Infect Dis Off Publ Infect Dis Soc Am* 2008; **46**: 1617–23.

2. Daar ES, Kesler KL, Petropoulos CJ, *et al.* Baseline HIV type 1 coreceptor tropism predicts disease progression. *Clin Infect Dis Off Publ Infect Dis Soc Am* 2007; **45**: 643–9.

3. Poveda E, Briz V, de Mendoza C, *et al.* Prevalence of X4 tropic HIV-1 variants in patients with differences in disease stage and exposure to antiretroviral therapy. *J Med Virol* 2007; **79**: 1040–6.

4. Raymond S, Delobel P, Mavigner M, *et al.* Genotypic prediction of human immunodeficiency virus type 1 CRF02-AG tropism. *J Clin Microbiol* 2009; **47**: 2292–4.

5. Lengauer T, Sander O, Sierra S, Thielen A, Kaiser R. Bioinformatics prediction of HIV coreceptor usage. *Nat Biotechnol* 2007; **25**: 1407–10.

6. Raymond S, Delobel P, Rogez S, *et al.* Genotypic prediction of HIV-1 CRF01-AE tropism. *J Clin Microbiol* 2013; **51**: 564–70.

7. Raymond S, Delobel P, Chaix M-L, *et al.* Genotypic prediction of HIV-1 subtype D tropism. *Retrovirology* 2011; **8**: 56.

8. Recordon-Pinson P, Soulié C, Flandre P, *et al.* Evaluation of the genotypic prediction of HIV-1 coreceptor use versus a phenotypic assay and correlation with the virological response to maraviroc: the ANRS GenoTropism study. *Antimicrob Agents Chemother* 2010; **54**: 3335–40.

9. Cheingsong-Popov R, Lister S, Callow D, Kaleebu P, Beddows S, Weber J. Serotyping HIV type 1 by antibody binding to the V3 loop: relationship to viral genotype. WHO Network for HIV Isolation and Characterization. *AIDS Res Hum Retroviruses* 1994; **10**: 1379–86.

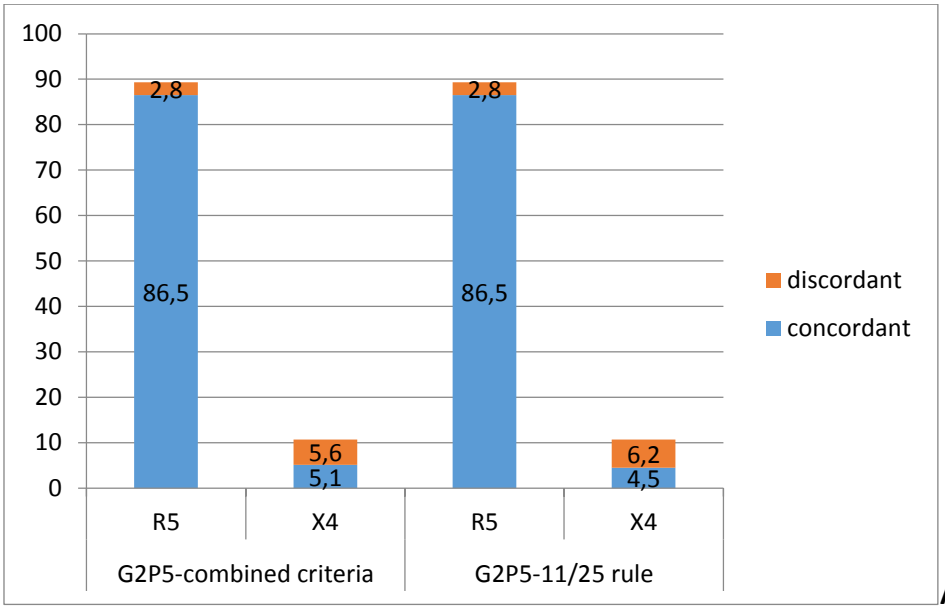
10. Seclén E, Garrido C, González M del M, *et al.* High sensitivity of specific genotypic tools for detection of X4 variants in antiretroviral-experienced patients suitable to be treated with CCR5 antagonists. *J Antimicrob Chemother* 2010; **65**: 1486–92.

11. Delgado E, Fernández-García A, Vega Y, *et al.* Evaluation of genotypic tropism prediction tests compared with in vitro co-receptor usage in HIV-1 primary isolates of diverse subtypes. *J Antimicrob Chemother* 2012; **67**: 25–31.

12. Gulick RM, Lalezari J, Goodrich J, *et al.* Maraviroc for previously treated patients with R5 HIV-1 infection. *N Engl J Med* 2008; **359**: 1429–41.

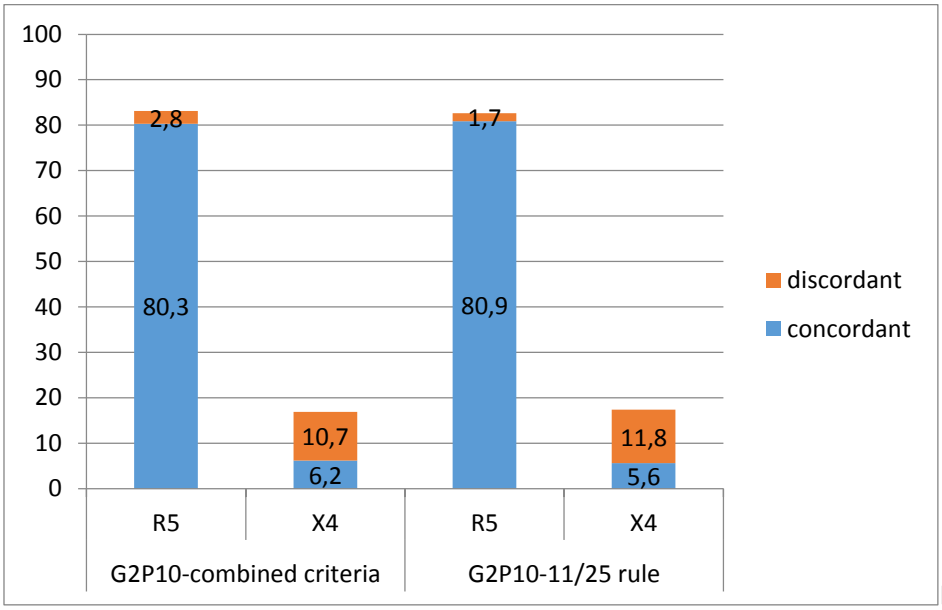
13. Cooper DA, Heera J, Goodrich J, *et al.* Maraviroc versus efavirenz, both in combination with zidovudine-lamivudine, for the treatment of antiretroviral-naive subjects with CCR5-tropic HIV-1 infection. *J Infect Dis* 2010; **201**: 803–13.

14. Fätkenheuer G, Nelson M, Lazzarin A, *et al.* Subgroup analyses of maraviroc in previously treated R5 HIV-1 infection. *N Engl J Med* 2008; **359**: 1442–55.



A

Sensitivity: 96.8%	Sensitivity: 96.8%
Specificity: 47.4%	Specificity: 42.1%
Kappa=0.5002	Kappa=0.4525



B

Sensitivity: 96.6%	Sensitivity: 97.9%
Specificity: 36.6%	Specificity: 32.2%
Kappa=0.4090	Kappa=0.3920

Figure 1: Representation of tropism prediction by different algorithm for CRF02_AG recombinant

A: comparison between G2P5 (Geno2pheno FPR 5%) and combined criteria or 11/25 rule

B: comparison between G2P10 (Geno2pheno FPR 10%) and combined criteria or 11/25 rule

Gp120	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330
V3 loop	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
HXB2	C	T	R	P	N	N	N	T	R	K	S	I	R	I	Q	R	G	P	G	R	A	F	V	T	I	G	I	G	D	M	R	Q	A	H	C
B	C	T	R	P	**N**	N	N	T	R	K	S	**I**	**H**	I	G	P	G	**R**	**A**	F	Y	A	T	G	**D/E**	I	I	G	D	I	R	Q	A	H	C
CRF02_AG	C	T	R	P	**D/N**	N	N	T	R	K	S	**V**	**R**	I	G	P	G	**Q**	**A/T**	F	Y	A	T	G	**D**	I	I	G	D	I	R	Q	A	H	C

Table 1A: Amino acid consensus sequences of V3 loop-gp120 HIV B subtype and CRF02_AG recombinants.

		B subtype	CRF02_AG recombinant
	A	0.5	
	D	1.5	1.1
	G	29.0	18.2
	K		0.5
	L	0.5	0.5
	N	0.5	0.5
	Q		0.5
	R	7.5	2.7
	S	**59.5**	**75.4**
	T	0.5	
11 S	Y	0.5	0.5
	A	5.7	6.9
	D	**33.0**	**51.0**
	E	22.2	13.7
	G	5.7	6.9
	H	1.4	
	I	7.1	2.4
	K	2.4	2.9
	N	3.8	4.9
	Q	9.4	2.9
	R	6.1	1.9
	S	2.8	1.5
	T	0.5	2.4
	V		0.5
	Y		0.5
25 I	*		1.5

Table 1B: Distribution of amino acid at position 11 and 25 of V3 loop-gp120.

In bold, the major amino acid