Algorithm-Based Prediction of HIV-1 Subtype D Coreceptor Use

Julia Dina,a,i Stephanie Raymond,b Anne Maillard,c,i Helene Le Guillou-Guillemette,a,i Audrey Rodalec,e,i Agnes Beby-Defaux,i,i Genevieve Giraudau,i,i Sophie Vallet,i,i Thomas Mourez,h,i Christopher Payan,g,i Astrid Vabret,a,i Annick Ruffault,c,i Virginie Ferre,e,i Jacques Izopet,b Jean-Christophe Plantierh,i

Laboratoire de Virologie, CHU de Caen, Caen, Francea; Laboratoire de Virologie, CHU de Toulouse, Toulouse, Franceb; Laboratoire de Virologie, CHU de Rennes, Rennes, Francec; Laboratoire de Virologie, CHU d’Angers, Angers, Franced; Laboratoire de Virologie, CHU de Nantes, Nantes, Francene; Laboratoire de Virologie, CHU de Poitiers, Poitiers, Francef; Laboratoire de Virologie, CHU de Brest, Brest, Franceg; Laboratoire de Virologie, CHU de Rouen, Rouen, Franceh; Groupe ARGO (Association des Rétrovirologues de l’Ouest)i‡

We compared the coreceptor tropism-predicting performance of a specific genotypic algorithm for HIV-1 subtype D and that of the geno2pheno algorithm with different cutoffs. The D-specific algorithm and geno2pheno with a false-positivity rate cutoff of 2.5% had the same concordance with the phenotypic determination. The geno2pheno algorithm with a false-positivity rate cutoff of 2.5%, more sensitive but slightly less specific, seems to be an appropriate alternative.

The robustness of the genotyping algorithms for predicting coreceptor tropism has been evaluated for HIV-1 subtype B (1, 2). While at least 40% of the patients monitored in Europe are infected with non-B subtypes, few data are available for these non-B genotypes. Recent studies on subtypes C, D, CRF01, and CRF02 (3–6) indicate that the rules could differ, depending on the subtype. The coreceptor tropism distributions of different HIV-1 subtypes also differ considerably (7–10). The disease progresses more rapidly in patients with a subtype D infection, perhaps because of the high prevalence of CXCR4 tropism and dual-mixed virus populations (11,12). It is thus essential to accurately predict tropism both for initiation of therapy targeting the CCR5 coreceptor and for pathophysiological studies. Raymond et al. built a specific rule for predicting CXCR4 usage for subtype D (3) based on one of the following criteria: (i) R (arginine) or K (lysine) at position 11 of hypervariable region V3 of gp120, (ii) R at position 25 of V3 and a net charge of $\frac{1}{2} > 5$, or (iii) a net charge of $\frac{1}{2} > 6$.

This study compared the predictive capacities of the subtype D-specific rule and the geno2pheno (G2P) tool with different false-positivity rate (FPR) cutoffs with a new panel of subtype D samples.

A total of 31 HIV-1 subtype D samples characterized in both the pol and env coding regions were collected from seven French centers. HIV-1 tropism was determined for the clinical management of the patients. The Toulouse tropism test (TTT), a recom-

### TABLE 1 Comparison of genotypic predictions of HIV-1 subtype D tropism with observed phenotypes

<table>
<thead>
<tr>
<th>Genotyping tool and TTT phenotype</th>
<th>TTT phenotype</th>
<th>Performance of the genotyping tool</th>
<th>% Concordance (kappa value)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R5</td>
<td>X4</td>
<td>% Sensitivity</td>
</tr>
<tr>
<td>G2P 10</td>
<td>19</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>6</td>
<td>100</td>
</tr>
<tr>
<td>G2P 5</td>
<td>20</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>6</td>
<td>100</td>
</tr>
<tr>
<td>G2P 2.5</td>
<td>23</td>
<td>1</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5</td>
<td>83</td>
</tr>
<tr>
<td>Combined D rule</td>
<td>24</td>
<td>2</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>4</td>
<td>67</td>
</tr>
</tbody>
</table>

*The kappa coefficient was measured by using MedCalc to assess the agreement between the genotypic algorithms and the phenotypic assay. The correlation between two tests is usually considered good when the kappa coefficient is greater than 0.60.*
binant virus entry phenotypic assay, was used to determine the entry phenotype of each sample as previously described (13). The V3 region of these samples was also sequenced from the bulk env PCR products to determine the V3 genotype (http://www.hivfrenchresistance.org). The data resulting from TTT were compared to the results obtained with the specific D rule and the G2P tool with FPR cutoffs of 10, 5, and 2.5% (G2P 10, G2P 5, and G2P 2.5, respectively). The phenotype-genotype correlations were then determined (Table 1). The phenotypic assay identified 25 virus populations as being R5, 3 as dual-mixed R5-X4, and 3 as X4. The sensitivity with G2P 10 and G2P 5 was 100%, but the specificity was poor (Table 1). G2P 2.5 and the D-specific rule agreed very well with the phenotype (90%). G2P 2.5 predicted 23 R5 virus populations and 5 X4 virus populations, in concordance with the TTT phenotype; the D-specific rule predicted 24 R5 virus populations and 5 X4 virus populations. The best predictors in terms of sensitivity and specificity for detecting X4 variants were G2P 2.5 (83 and 92%, respectively) and the D-specific rule (67 and 96%) (Table 1). Different algorithms were tested for the genotypic prediction of coreceptor tropism with the V3 subtype D sequences whose coreceptor tropism is available in the Los Alamos HIV sequence database (http://www.hiv.lanl.gov/content/index). The data set included 59 R5 viruses and 34 R5-X4 or X4 viruses based on phenotypic assays. Once again, the concordances were best with G2P 2.5 (84%) and the D-specific rule (85%). G2P 2.5 had 76% of the sensitivity and 88% of the specificity of the phenotypic test, while the D-specific rule had 62% of its sensitivity and 98% of its specificity (Table 2).

Virus tropism must be determined before using CCR5 antagonists, to avoid using ineffective drugs, to avoid the selection of strains resistant to the associated drugs, and also for studies on disease progression. Bioinformatic algorithms like G2P were developed as practical alternatives to phenotypic tests, which are expensive and time-consuming. Raymond et al. have already shown the impact of HIV-1 genetic diversity on tropism prediction with such tools (4–6). We have now tested the D-specific rule recently proposed by Raymond et al. on new samples. This D-specific rule and G2P 2.5 performed best in terms of specificity and sensitivity. G2P 10, which is routinely used in France, is clearly inappropriate for subtype D tropism prediction. The correlation between the phenotype and the subtype D-specific rule when tested on our data set (90%) was similar to that reported previously (92%) (3). G2P 2.5 was more sensitive (83%) than the subtype D-specific rule (67%) and only slightly less specific (92%) that the subtype D-specific rule (96%).

Nucleotide sequence accession numbers. The sequences reported here have been assigned GenBank accession numbers HF678992 to HF679022.

REFERENCES


