HIV-2 Group A in France Displayed Two Clades with Distinct Geographical Origins

Benoit Visseaux 1,2,3, Charlotte Charpentier 1,2,3, Mélanie Bertain 1,2,3, Amel Benseguier 4, Catherine Figaro 4, Florence Damond 1,2,3, Sophie Matheron 1,2,3, Stéphane Hûé 4, Diane Descamps 1,2,3

on the behalf of the French ANRS COS HIV-2 cohort.

BACKGROUND

• To date nine distinct lineages of HIV-2 have been identified and appear to represent independent host transfers. Analysis of phylogenetic data to HIV-1, these lineages have been termed groups A, B, but only groups A and B have spread into humans (Gao et al., 1994; Damond et al., 2004).

• HIV-2 group A is mainly found in Western Africa and seems to have originally spread from Guinea-Bissau during the independence war (1963-1974) (Leroy et al., 2003).

• HIV-2 group B mainly predominates in Ivory Coast (Paparené et al., 1999).

• In France, HIV-2 represents 1 to 2% of all newly diagnosed HIV infections and more than 1948 HIV-2 infected patients have been included in the French ANRS-COS HIV-2 cohort since 1994. Among these patients, the HIV-2 diversity is higher than in other northern countries, with a higher representation of HIV-2 group B (third of all infected viruses) than group A. This is mainly explained by local determinism of patients with a strong representation of patients from Ivory-Coast, Mali, Burkina-Faso or Guinea (Damond et al., 2003).

• HIV-2 group A was originally subdivided in three different clades, based on whole env gene sequences (Breuer 1995). However, these clades were not retrieved when analysing the N-terminus portion of gag1 alone. In another study using the env gene, group A was divided in two distinct clades, reflecting possible geographical divisions (Damond et al., 2001).

OBJECTIVES

• In this study, we aimed to describe HIV-2 diversity observed in France, to confirm the existence of two clades among HIV-2 group A and to explore their geographical locations.

METHODS

• A total of 446 partial pol (99 nt and 87 area, 150 nt), 130 partial env (encompassing the V3 loop, 525 nt and 155 complete of 165 nt) sequences obtained among 586, 118 and 72 patients sampled in Paris, were analysed.

• 207, 374 and 22 publicly available sequences from GenBank were also included for pol, env and rna sequences.

• Available SVMir and SVMir sequences were also retrieved (n=46).

• Maximum likelihood phylogenetic trees were constructed using FastTree 2.2 under the GTR + I + G model. Branch support was estimated with the SH-like test method (wL).

• Possible recombinant sequences inside each genes were checked and removed before phylogenetic analysis using IQP04.

• Patient's country of birth was retrieved for 372 patients sampled in France and included in the French ANRS COS HIV-2 cohort.

RESULTS

• In all trees, HIV-2 group A formed two distinct and strongly supported clusters (branch support > 80%), herein called A1 and A2, in all trees (cf figure 1).

• Overall our three trees, 68.1, 140 and 17.9% sequences belonged to cluster A1, A2 and B, respectively (cf table 1).

• The genetic diversity of HIV-2 sequences was significantly higher in the French cohort than among publically available sequences, as depicted by the higher representation of sequences belonging to the A2 clade (table 2). GTR vs. 73% p<0.001 and B group (34.7 vs. 7.6%, p<0.001).

• Median genetic distance between clades A1 and A2 were 0.12 [IDS=0.11-0.14, 0.12 [0.11-0.13] and 0.15 [0.12-0.18] substitutions/site for pol, env and rna, respectively.

• As expected, all sequences obtained from CRF01_AB clustered together. Only one patient from the French dataset clustered together with CR01_AB sequences. Interestingly this patient was born in Ivory Coast as the patient is strain-contaminated. Among the 85 patients born in Ivory Coast and sampled in France, only 1 seemed to belong to the CRF01_AB.

• Among the 183 sequences with at least two different genes available in this analysis, 17 (9%) presented inconsistent clade assignments across trees. This may suggest the existence of some other recombinant events and need to be explored further by whole genome sequencing.

• When analysing the country of birth of patients from the French ANRS COS HIV-2 cohort, 197 (79%) were born in Western Africa. Patient's country of birth according to the virus's clade, are depicted in figure 2.

• HIV-2 clade A1 was most prevalent amongst patients born in coastal Western African countries (i.e. Senegal, Gabon). HIV-2 clade A2 was found mainly in patients from Ivory-Coast, Mali and Burkina Faso. This distribution was not always observed at the strain level.

• HIV-2 clade A1 was more prevalent amongst patients born in coastal Western African countries (i.e. Senegal, Gabon). HIV-2 clade A2 was found mainly in patients from Ivory-Coast, Mali and Burkina Faso. This distribution was not always observed at the strain level.

DISCUSSION

• This study provides an enhanced understanding of the geographical and genetic diversity of HIV-2. It highlights the co-circulation of two distinct group A clades in France.

• HIV-2 A1 clades should be approached further along the Western coasts of Africa, while A2 clade and B groups seem to have emerged in Mali and Ivory Coast. This is likely explained by an ancient divergent event, followed by founder effects.

• These findings need to be further explored using dated phylogenetic and phylogeographic analysis through molecular inference analyses.

BIBLIOGRAPHY

1. INSERM, IAME, UMR 1137, Paris, France
2. Universitè Paris Diderot, IAME, UMR 1137, Sorbonne Paris Cité, Paris, France
3. AP-HP Hôpital Bichat, Laboratoire de Virologie, Paris, France
4. Université Bordeaux, IRD, Centre INSERM UMR 5597, Epidemiologie-Biostatistique, Bordeaux, France
5. AP-HP Hôpital Bichat, Service de Maladies Infectieuses et Tropicales, Paris, France
6. Department of Infectious Diseases Epidemiology, Faculty of Epidemiology and Public Health, London School of Hygiene and Tropical Medicine, London, United Kingdom

Table 1: Distribution of each HIV-2 clade among French database sequences and publically available sequences.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>A2</td>
<td>B</td>
</tr>
<tr>
<td>(146/497) (29.5%)</td>
<td>(8/497) (1.6%)</td>
<td>(6/497) (1.2%)</td>
</tr>
</tbody>
</table>

Figure 1. HIV-2 gene map. Maximum likelihood phylogenetic trees were reconstructed from HIV-2 partial pol (A1, A2 and B clades) sequences (205/497).

Figure 2. Distribution of country of birth among HIV-2 sequences for patient included in the French ANRS COS HIV-2 cohort.

Figure 3. Distribution of country of birth among HIV-2 sequences for patient included in the French ANRS COS HIV-2 cohort.