A new B/CRF02 circulating recombinant spreading quickly in Paris area, France

Marc Wirden[1], Fabienne De Oliveira[2] ,Alexandre Storto[3], Magali Bouvier-Alias[4], Karine Grenet[5], Marie-Laure Chaix[6], Philippe Simon[5], Eric Froguel[5], Thuy T. Nguyen[1], Cathia Soulié[1], Vincent Calvez[1], Jean-Christophe Plantier[2], Diane Descamps[3], Anne-Geneviève Marcelin[1], Benoit Visseaux[3*] , ANRS AC-11 Resistance Study Group.

1-*Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique (iPLESP); AP-HP Pitié-Salpêtrière Hospital, Paris, France ;* **2-***Normandie Univ, UNIROUEN, EA2656, CHU de Rouen. Laboratoire de virologie associé au Centre National de Référence du VIH, F-76000;* **3-** *IAME, UMR 1137, Univ Paris Diderot, Sorbonne Paris Cité; INSERM; AP-HP, Hôpital Bichat, Virologie, Paris, France ;* **4-** *Hôpital Henri Mondor, Créteil, France ;* **5-** *Grand Hôpital de l'Est Francilien, Jossigny, France ;* **6-** *INSERM U941, Université Paris Diderot, APHP, Service de Virologie, CHU Saint Louis, CNR VIH, Paris, France ; ; * -Presenting author*

POSTER #958

BACKGROUND

- The two majors circulating HIV-1 clades in France, subtype B and CRF02_AG, were originally present in distinct populations, Caucasian and West African population, respectively.
- However, CRF02_AG is increasing among all French HIV populations and several recombinant forms, 7 URFs and the CRF56_cpx, were already identified.
- In this study, we describe a new B/CRF02_AG recombinant spreading quickly and forming a recent transmission cluster among a MSM population.

METHODS

- All French sequence databases from the ANRS AC11 group, coordinating HIV resistance studies across France, were screened for an unusual recombinant pattern: PR and INT clustering with CRF02_AG but RT and env clustering with subtype B.
- For each identified patient, all sequences obtained at their diagnosis time, clinical data, town of residence and demographic data were extracted.
- HIV subtyping was confirmed by phylogenetic analysis for protease (PR) (299bps), reverse transcriptase (RT) (774bps), integrase (696bps) and envelope (328bps) with the Los Alamos National Laboratory reference sequences dataset and using FastTree 2.1.
- A first complete genome sequence has been obtained by bulk sequencing and recombination breakpoints have been estimated using SimPlot.
- The time of the most recent common ancestor was estimated from PR-RT using TempEst 1.5.1.
- ViroSeq system and/or ANRS method were used to sequence the genes. HIV tropisms have been determined with the Geno2Pheno tool.

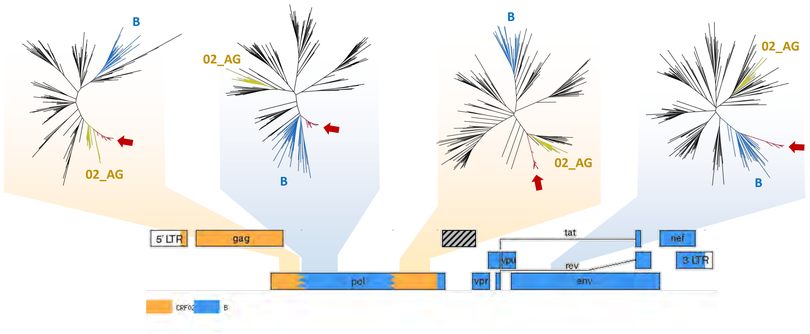
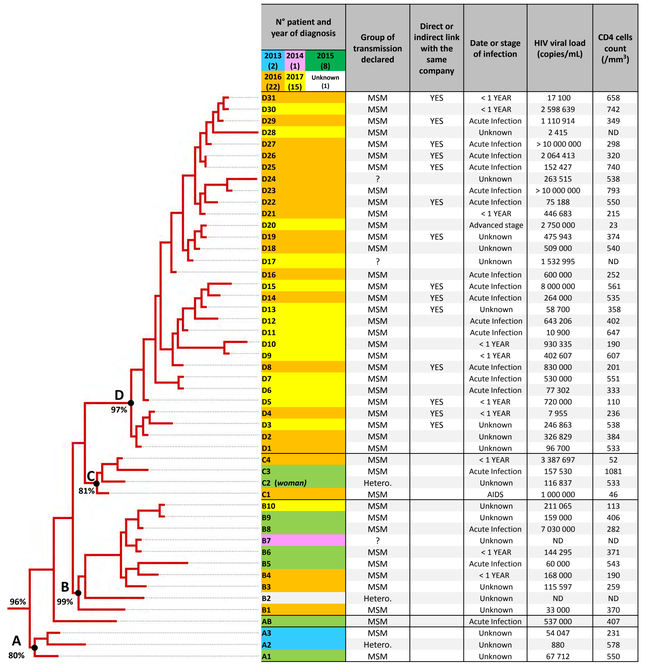


Figure 1. Recombination pattern of the new recombinant form. *This pattern needs to be confirmed with deeper phylogenetic analysis and additional near full genome sequences. Vif gene is of unclear origin on this first analysis. All sequences identified in France with this pattern are depicted in red in phylogenetic trees obtained for the main genes (protease, reverse transcriptase, integrase, envelope).*



N° patient and year of diagnosis	Group of transmission declared	Direct or indirect link with the same company	Date or stage of infection	HIV viral load (copies/mL)	CD4 cells count (/mm³)
2013 (2) 2014 (1) 2015 (8) 2016 (22) 2017 (15) Unknown (1)					
D31	MSM	YES	< 1 YEAR	17 100	658
D30	MSM		< 1 YEAR	2 598 639	742
D29	MSM	YES	Acute Infection	1 110 914	349
D28	MSM		Unknown	2 415	ND
D27	MSM	YES	Acute Infection	> 10 000 000	298
D26	MSM	YES	Acute Infection	2 064 413	320
D25	MSM	YES	Acute Infection	152 427	740
D24	?		Unknown	263 515	538
D23	MSM		Acute Infection	> 10 000 000	793
D22	MSM	YES	Acute Infection	75 188	550
D21	MSM		< 1 YEAR	446 683	215
D20	MSM		Advanced stage	2 750 000	23
D19	MSM	YES	Unknown	475 943	374
D18	MSM		Unknown	509 000	540
D17	?		Unknown	1 532 995	ND
D16	MSM		Acute Infection	600 000	252
D15	MSM	YES	Acute Infection	8 000 000	561
D14	MSM	YES	Acute Infection	264 000	535
D13	MSM	YES	Unknown	58 700	358
D12	MSM		Acute Infection	643 206	402
D11	MSM		Acute Infection	10 900	647
D10	MSM		< 1 YEAR	930 335	190
D9	MSM		< 1 YEAR	402 607	607
D8	MSM	YES	Acute Infection	830 000	201
D7	MSM		Acute Infection	530 000	551
D6	MSM		Acute Infection	77 302	333
D5	MSM	YES	< 1 YEAR	720 000	110
D4	MSM	YES	< 1 YEAR	7 955	236
D3	MSM	YES	Unknown	246 863	538
D2	MSM		Unknown	326 829	384
D1	MSM		Unknown	96 700	533
C4	MSM		< 1 YEAR	3 387 697	52
C3	MSM		Acute Infection	157 530	1081
C2 (*woman*)	Hetero.		Unknown	116 837	533
C1	MSM		AIDS	1 000 000	46
B10	MSM		Unknown	211 065	113
B9	MSM		Unknown	159 000	406
B8	MSM		Acute Infection	7 030 000	282
B7	?		Unknown	ND	ND
B6	MSM		< 1 YEAR	144 295	371
B5	MSM		Acute Infection	60 000	543
B4	MSM		< 1 YEAR	168 000	190
B3	MSM		Unknown	115 597	259
B2	Hetero.		Unknown	ND	ND
B1	MSM		Unknown	33 000	370
AB	MSM		Acute Infection	537 000	407
A3	MSM		Unknown	54 047	231
A2	Hetero.		Unknown	880	578
A1	MSM		Unknown	67 712	550

Figure 2. Phylogenetic tree obtained for the protease + reverse transcriptase region. *The four identified clades are indicated by a letter, from A to D, and the corresponding branch support value (SH-like test). The main clinical data for the corresponding patients are given in the table. Abbreviation: ND - no data*

RESULTS

- At January 2018, 49 infected patients were identified as presenting the unusual recombinant pattern. This pattern does not match the previous identified recombinant forms mixing subtype B and CRF02_AG. A first near-full genome sequence has been obtained from patient D8 samples and confirmed the novelty of the recombinant pattern (Figure 1), but this first analysis still needs to be confirmed with additional near-full genome sequences and deeper phylogenetic analysis.
- No drug resistance mutation was identified according the Stanford algorithm. The ANRS resistance algorithm identified the classical polymorphisms of the CRF02_AG in PR (10I, 20I, 36I, 69K and 89M). Interestingly, the ViroSeq genotyping system amplified PR and RT only for half of the tested cases despite high HIV viremia. The tropisms, determined for 19 patients, were X4 in 100% and 25% of cases at the FPR threshold of 10 % or 5 %, respectively.
- All the virus identified with this new recombinant form are included in a recent transmission cluster in the PR-RT (branch support value >95% and maximum genetic distance <3.4%) as well as in all the other tested genes (data not shown). The corresponding patients were diagnosed in 2013 (n=2), 2014 (1), 2015 (8), 2016 (22) and 2017 (15). The tMRCA analysis estimated the emergence of this cluster in July 2012.
- All patients are Caucasians, from French origins, with a median age of 35 [28-43] years and all but one are males. The declared transmission route was MSM for 43/49 patients, heterosexual for 3 and not reported for the 3 others. At the time of diagnosis, 37% were at the acute phase of the HIV infection. Median viral loads were 263 515 [IQR: 65 459-772 583] and 565 000 [IQR: 152 427-2 064 413] copies/mL, and median CD4 cells count were 370 [IQR: 151-538] and 471 [IQR: 320-561] cells/mm³ for non-acute and acute-infections, respectively.
- The phylogenetic analysis showed four distinct clades, herein called clades A, B, C and D, consistent with the shape of the tree and branch support (figure 2). The patients diagnosed between 2013 (depicted in blue) and 2015 (in green) are all included in the first three clades (A, B and C), while the last clade (D) harbors only the patients diagnosed from 2016 to 2017. Interestingly, these clades also presented distinct geographical distribution (figure 3). Thus, patients from clades A and B resided in an area centered on the Val-de-Marne department (green circle – zip code 94), whereas the patients of the clade D, diagnosed later, are mostly localized in the north of the Seine-et-Marne department (red circle) or in more distant French regions.
- For 45% of the patients included in clade D, a common link was identified as they were employed (sometimes seasonally) at the same big company of the area or had had sexual risk behaviors with employees from this company. Three patients, patients D8, D14 and D15, all identified outside the Ile-de-France region, were also seasonal workers in this company. With these information, mixing phylogenetic and epidemiologic analysis, prevention actions are planned to target the workers of this company. As many European seasonal workers, housed in residences, are also recruited, this new recombinant form may also have spread elsewhere in Europe.

CONCLUSION

- A new CRF02_AG/B recombinant, proposed as the CRF94_B02, has been identified and spread quickly in Paris area, particularly among a very targeted population.
- Additional near full genome sequences and phylogenetic analysis are ongoing to deeply characterize this new recombinant form.
- This work shows that combining phylogenetic and demographic analysis can be useful to detect a specific area and/or population and, then, trigger targeted preventive actions allowing to break newly active transmission clusters.

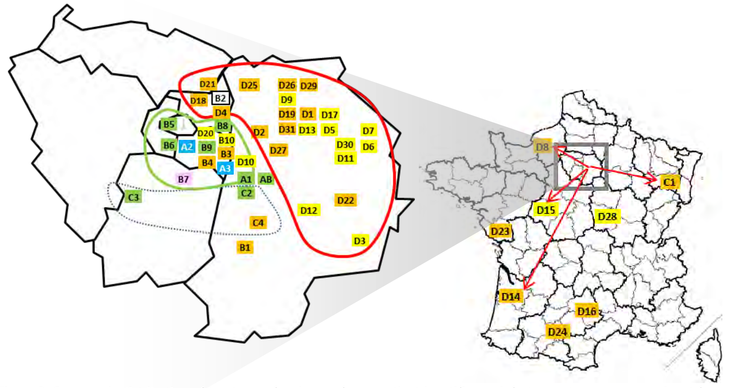


Figure 3. Geographic distribution of the town of residence for all patient identified as infected by the new recombinant form. The color correspond to the year of diagnosis (see Figure 2 for legend). The red arrows indicate a geographical displacement after the probable date of infection.