



**HAL**  
open science

# Formes recombinantes inter-groupes M et O du VIH-1 : Etude de leur potentiel répliatif et de leur émergence in vitro/in vivo

Alice Moisan

## ► To cite this version:

Alice Moisan. Formes recombinantes inter-groupes M et O du VIH-1 : Etude de leur potentiel répliatif et de leur émergence in vitro/in vivo. Virologie. Normandie Université, 2019. Français. NNT : 2019NORMR126 . tel-02509644

**HAL Id: tel-02509644**

**<https://theses.hal.science/tel-02509644>**

Submitted on 17 Mar 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Normandie Université

## THÈSE

**Pour obtenir le diplôme de doctorat**

**Spécialité Sciences de la Vie et de la Santé**

**Préparée au sein de l'Université de Rouen Normandie**

**Formes recombinantes inter-groupes M et O du VIH-1 :  
étude de leur potentiel réplicatif et de leur émergence *in vitro/in vivo***

**Présentée et soutenue par  
Alice MOISAN**

**Thèse soutenue publiquement le 18 décembre 2019  
devant le jury composé de**

Madame Marie-Laure CHAIX	Maître de conférences Université Paris-Diderot	Rapporteur
Monsieur Pierre ROQUES	Chercheur CEA E6 Commissariat à l'Energie Atomique, Fontenay aux roses	Rapporteur
Madame Diane DESCAMPS	Professeur Université Paris-Diderot	Examineur
Monsieur Denys BRAND	Professeur Université de Tours	Examineur
Monsieur Matteo NEGRONI	Directeur de recherche IBMC, Strasbourg	Examineur
Monsieur Jean-Christophe PLANTIER	Professeur Université de Rouen Normandie	Directeur de thèse

**Thèse dirigée par Jean-Christophe PLANTIER, Laboratoire GRAM EA2656**





## **REMERCIEMENTS**

### **Au Professeur Jean-Christophe Plantier**

Pour votre soutien et votre implication à toutes les étapes de ce travail et depuis toutes ces années, pour votre écoute et votre disponibilité au quotidien et pour votre rôle et vos encouragements dans mon orientation professionnelle ; je vous remercie très chaleureusement de la confiance que vous m'accordez.

### **Aux Docteurs Marie-Laure Chaix et Pierre Roques**

Pour l'intérêt que vous avez bien voulu porter à ce travail et pour avoir accepté la charge d'en être les rapporteurs.

### **Aux Professeurs Diane Descamps et Denys Brand**

Pour avoir accepté d'être examinateurs de ce travail et pour le temps que vous avez consacré à en juger le contenu.

### **Au Docteur Matteo Negroni**

Pour tous les échanges constructifs que nous avons depuis plusieurs années au sujet de la recombinaison et surtout des LTRs, pour tout le temps que tu me consacres encore aujourd'hui en te déplaçant jusque Rouen pour juger mon travail et y apporter ton expertise.

### **Au Docteur Thomas Mourez**

Pour les conseils et le soutien que tu m'as apportés. Je n'oublie ni le rôle que tu as joué dans ce travail, ni le gout pour la recherche et pour l'enseignement que tu m'as transmis dès mon arrivée au laboratoire.



## A toute l'équipe du laboratoire de Virologie du CHU de Rouen

Cette Thèse est le fruit d'un long travail d'équipe, qui a pu aboutir grâce la contribution de nombreuses personnes.

Un grand merci, **Fabienne**, pour ta disponibilité, ta gentillesse, ta motivation, ton expertise en la matière et ta précieuse aide. J'espère que nous continuerons cette fructueuse collaboration sur la recombinaison aussi longtemps qu'il y aura des choses à dire sur le sujet !

Merci à tous les étudiants ayant travaillé sur la thématique avec nous, en particulier **Juliette, Manon et Albane**.

Merci **Sandy** pour toute la logistique et pour les longues heures que tu as passées à annoter le document.

Merci à tous mes collègues biologistes, ingénieurs et techniciens pour votre aide, vos conseils, votre soutien et pour contribuer à rendre agréables nos journées au travail.

Un merci particulier, **Adeline**, ma colocataire de bureau, pour avoir supporté mes journées de mutisme devant mon ordinateur, pour m'avoir aidé à me décharger un peu de ma routine hospitalière et pour nos moments extra-hospitaliers aussi.

Et merci, bien évidemment, **Elodie et Guillemette**, mes deux complices de thèse, pour nos nocturnes au laboratoire, pour nos moments de décompression et pour votre amitié luxueusement gallinacée.

## A toute l'équipe de recherche du GRAM 2.0

Pour les bons moments partagés, pour la constante bonne humeur et pour les différentes aides que vous m'avez apportées au cours de la partie bactérienne de ces travaux viraux.

En particulier, merci **Imane** pour ton expertise en clonage et **Stéphanie** pour les autoclaves et les milieux.

Merci **Pierre et Paul-Alain** pour nos échanges constructifs sur les recombinants MOMO.

## **A tous ceux qui permettent un précieux équilibre**

A mes amis de toujours : merci d'avoir participé à mon épanouissement et de m'avoir aidée à relativiser mes périodes de doute et de fatigue. Merci pour tous les bons moments passés et à venir et surtout merci pour votre confiance et votre amitié.

A **Romain** et **Fanny** : merci infiniment pour votre fidélité et votre bonne humeur, pour ces dizaines d'années passées à vos côtés. Merci pour cette grande place que vous occupez dans ma vie et dans mon cœur.

A ma belle-famille : merci pour votre bonne humeur, vos mots doux et vos encouragements.

A ma famille : **Georges, Gisèle, Louis, Thérèse, Martine, André, Camille, Antoine** et **Léa** et à ma « famille bis » : **Christian, Isabelle, Emilie, Fanny** et **Céline**. Merci pour tout ce que vous m'apportez depuis toujours. Merci tous ces moments tous ensemble qui sont pour moi une vraie soupape de décompression et merci pour tout l'amour dans lequel vous m'avez fait grandir.

## **A mes piliers**

**Catherine, Michel, Delphine, Pauline, Alex** et **Anaïde**, ce travail vous doit tellement. Merci pour votre présence dans les bons comme dans les mauvais moments et pour vos encouragements constants malgré la longueur de cette Thèse. Merci aussi pour ces moments ressourçants en famille, pour votre bienveillance, votre générosité, votre tendresse et votre affection, qui m'ont permis d'être ce que je suis aujourd'hui.

**Samir** et **Léna**, merci d'avoir supporté mon rythme infernal des derniers mois. Merci pour le bonheur que vous m'apportez au quotidien, pour votre amour et pour tous les projets à construire ensemble.



## **Formes recombinantes inter-groupes M et O du VIH-1 : étude de leur potentiel répliatif et de leur émergence *in vitro/in vivo***

### **RESUME**

Les VIH-1 sont caractérisés par une forte diversité génétique et la recombinaison génétique constitue un facteur important de leur évolution. Malgré leur grande divergence génétique, les VIH-1 groupe M, pandémique, et groupe O, endémique au Cameroun, peuvent générer des formes recombinantes inter-groupes MO. La description actuellement de 19 formes recombinantes VIH-1/MO (URF\_MO), dont 10 au cours de ces dix dernières années, pose la question d'un éventuel avantage conféré par la recombinaison et des modalités de leur émergence.

Les objectifs de ce travail étaient donc d'étudier le potentiel répliatif ainsi que l'émergence *in vitro* et *in vivo* de formes recombinantes VIH-1/MO.

Le potentiel répliatif a été étudié sur la base d'un profil simple de recombinaison [ $O_{gag/pol}-M_{env}$ ], avec un point de cassure dans Vpr, du fait de l'observation d'un point chaud de recombinaison dans cette région. Après de multiples essais pour construire un clone moléculaire infectieux chimérique (CMIC) à partir de clones moléculaires infectieux parentaux de VIH-1/M sous-type B et de VIH-1/O sous-groupe T, un CMIC pVIH-1/OM a été synthétisé et a permis la génération de virus recombinants par transfection puis co-culture. Un stock a été généré afin de comparer son potentiel répliatif par rapport à celui des souches parentales VIH-1/M et VIH-1/O. Deux marqueurs ont été suivis dans les surnageants de culture : l'activité de la Transcriptase Inverse (TI) et la quantité d'Ag P24. Les résultats ont mis en évidence une supériorité de la souche parentale de groupe M par rapport à celle de groupe O pour les deux marqueurs. En revanche, pour la souche recombinante VIH-1/OM, les données d'activité de la TI ne se superposaient pas à celle de la quantité d'AgP24, suggérant un comportement hybride du recombinant, en termes d'activité enzymatique et de production d'AgP24.

*In vitro*, les modalités d'émergence ont été appréhendées par l'étude de la génération de points de cassure au sein des LTRs initialement différents du recombinant VIH-1/OM généré et dans lesquels un évènement de recombinaison devait avoir lieu pour les rendre identiques. La cinétique d'émergence et les profils de recombinaison ont été analysés par la méthode SGA, afin de caractériser les quasi-espèces ARN et ADN. Nos résultats ont mis en évidence une localisation préférentielle de la recombinaison dans la région R des LTRs, avec trois motifs majoritaires de recombinaison (506-513pb, 513-522pb et 523-547pb), dont le second prédominait en fin de culture. Notre modèle expérimental a été validé en comparant ces résultats aux données *in vivo* des URF\_MO décrites. D'autres motifs de recombinaison ont été observés *in vivo*, suggérant qu'il n'existe pas de motif et de mécanisme uniques de recombinaison.

*In vivo*, nous avons eu l'opportunité de décrire un cas de recombinaison MO, révélant une double infection VIH-1/M+O non diagnostiquée, chez une patiente présentant des résultats de génotypage de résistance discordants par rapport à ses antécédents VIH-1/M. Ce travail a permis de caractériser la dynamique de répliation des populations virales VIH-1/M et VIH-1/O avant recombinaison, puis celle du recombinant VIH-1/MO. Le suivi immuno-virologique de la patiente a permis de dater l'émergence de cette forme recombinante, probablement favorisée par les dysobservances thérapeutiques et la succession d'échecs virologiques. Ce cas souligne la difficulté de prise en charge de telles situations, et, en particulier, la nécessité de traitements antirétroviraux actifs sur les trois populations virales.

En conclusion, nos résultats apportent de nouveaux éléments de compréhension du phénomène de recombinaison entre VIH et nous offrent de nombreuses perspectives de travail.

## **HIV-1/MO intergroup recombinant forms : study of their replicative potential and their emergence *in vitro/in vivo***

### **ABSTRACT**

HIV-1 are characterized by a high genetic diversity and genetic recombination has an strong impact on their evolution. Despite their great genetic divergence, HIV-1 group M, pandemic, and group O, endemic in Cameroon, can generate HIV-1/MO intergroup recombinants. The current description of 19 HIV-1/MO recombinant forms (URF\_MO), including 10 in the last ten years, raises the question of a possible benefit of the recombination and the modalities of their emergence.

Therefore, the objectives of this study were to study the replicative potential and the *in vitro* and *in vivo* emergence of HIV-1/MO recombinant forms.

The replicative potential was analyzed, based on a simple recombination pattern, [ $O_{gag/pol-M_{env}}$ ], harbouring a breakpoint in Vpr, due to a recombination hotspot in this region. After many attempts to generate Chimeric Infectious Molecular Clone (CIMC) from HIV-1/M subtype B and HIV-1/O subgroup T parental infectious molecular clones, a CIMC pVIH-1/OM was synthesized and provided recombinant viruses by transfection and co-culture. A stock was generated to compare the replicative potential of recombinant viruses with HIV-1/M and HIV-1/O parental strains. Two markers were monitored in culture supernatants: Reverse Transcriptase (RT) activity and P24 antigen quantity. The results showed a superiority of the group M parental strain compared to group O for both markers. In contrast, for the HIV-1/OM recombinant strain, RT activity data did not overlap with the amount of P24 antigen, suggesting an hybrid behaviour of the recombinant, in terms of enzyme activity and P24 production.

*In vitro*, the emergence modalities were apprehended by the study of the generation of breakpoints within the LTRs, initially discordant, of the generated HIV-1/OM recombinants and in which a recombination event had to occur to become identical. Emergence kinetics and recombination profiles were analysed, using a Single Genome Amplification strategy, in order to characterize RNA and DNA quasi-species. Our results showed a preferential localization of recombination in the R region of LTRs, with three major recombination patterns (506-513bp, 513-522bp and 523-547bp), the second one predominating at the end of culture. Our experimental model was validated by comparing these results with the *in vivo* data from URF\_MOs. Other recombination patterns were observed *in vivo*, suggesting that patterns and mechanisms of recombination are not unique.

*In vivo*, we had the opportunity to describe a unique case of HIV-1/MO recombination which revealed an undiagnosed HIV-1/M+O dual infection, based on discordant results in a routine drug-resistance test in an individual initially diagnosed as HIV-1/M. This work allowed us to characterize the replication dynamics of the HIV-1/M and HIV-1/O viral populations before recombination, and that of the HIV-1/MO recombinant. The immunovirological follow-up of this patient has allowed to date the emergence of this recombinant form, probably favored by therapeutic dysobservances and the succession of virological failures. This case highlights the complexity of following such situations and the need, in particular, for antiretroviral therapy active on the three viral populations.

In conclusion, our findings provide new insights into the phenomenon of HIV recombination and offer many opportunities to work.

**Keywords :** HIV-1/MO intergroup recombination - HIV-1/M+O dual infections - Replicative potential- Emergence

<b>ANALYSE BIBLIOGRAPHIQUE</b>	<b>21</b>
<b>I- GÉNÉRALITÉS SUR LES VIRUS DE L'IMMUNODÉFICIENCE HUMAINE (VIH)</b>	<b>23</b>
I-1. Circonstances de découverte des VIH	23
I-2. Caractéristiques virologiques	24
I-3. Particule virale	24
I-4. Organisation du génome	25
I-5. Cycle de réplication virale	27
I.5.1. Déroulement de la réplication virale	27
I.5.2. Facteurs de restriction cellulaires et contournement VIH	31
<b>II- ORIGINES, ADAPTATION ET DIVERSITÉ GÉNÉTIQUE DES VIH</b>	<b>35</b>
II-1. Origines simiennes	35
II-2. Adaptation à l'Homme	36
II-3. Diversité génétique	37
II.3.1. Facteurs à l'origine de la diversité génétique	37
II.3.2. Recombinaison au sein des VIH	39
<b>III- CLASSIFICATION DES VIH</b>	<b>47</b>
III-1. Les VIH-1	48
III.1.1. Le groupe M (Major)	48
III.1.2. Les groupes non-M (N, O et P)	50
III-2. Les VIH-2	54
<b>IV- RECOMBINAISON INTER-GROUPES ENTRE LES VIH-1/M ET VIH-1/O</b>	<b>55</b>
III-1. Pré-requis : la double infection VIH-1/M+O	56
III-2. Circonstances de découverte des formes recombinantes VIH-1/MO	58
III-3. Caractéristiques moléculaires des formes recombinantes VIH-1/MO	63
III-3.1. Analyse phylogénétique et détermination du nombre d'URF_MO	63
III-3.2. Analyse des points de cassure des 19 URF_MO	65
III-4. Impacts potentiels de la recombinaison inter-groupes MO	67
<b>RATIONNEL DU TRAVAIL EXPERIMENTAL</b>	<b>69</b>

<b>PARTIE 1.</b>	<b>ETUDE DU POTENTIEL REPLICATIF <i>IN VITRO</i> DE VIRUS RECOMBINANTS INTER-GROUPES M ET O</b>	<b>75</b>
<hr/>		
<b>CONTEXTE ET OBJECTIFS DU TRAVAIL EXPÉRIMENTAL</b>		<b>77</b>
<b>A. GÉNÉRATION DE VIRUS RECOMBINANTS INTER-GROUPES M ET O</b>		<b>80</b>
<b>I- MÉTHODES</b>		<b>80</b>
I-1. Génération de clones moléculaires infectieux chimériques		80
I-1.1. Modalités de génération des clones moléculaires infectieux chimériques		80
I-1.2. Génération des CMIC par synthèse		91
I-2. Production de virus recombinants à partir de clones moléculaires infectieux chimériques		92
I-2.1. Transfection		92
I-2.2. Co-culture		93
I-2.3. Contrôle de la production et de la réplication virales		93
<b>II- RÉSULTATS</b>		<b>95</b>
II-1. Génération de clones moléculaires infectieux chimériques		95
II-1.1. Modalités de génération des clones moléculaires infectieux chimériques		95
II-1.2. Génération des CMIC pVIH-1/OM et pVIH-1/MO		96
II-2. Production de virus recombinants à partir des clones moléculaires infectieux chimériques		101
II-2.1. pVIH-1/OM		101
II-2.2. pVIH-1/MO		102
II-2.3. Bilan de la production de virus recombinants		103
<b>III- DISCUSSION</b>		<b>104</b>
<b>B. ETUDE DE LA CINÉTIQUE DE RÉPLICATION DES VIRUS PARENTAUX ET RECOMBINANT</b>		<b>108</b>
<b>I- MÉTHODES</b>		<b>108</b>
I-1. Production d'un stock de virus parentaux et recombinant		108
I-2. Etude de la cinétique de réplication des virus parentaux et recombinant		110
<b>II- RÉSULTATS</b>		<b>111</b>
II-1. Production du stock de virus parentaux et recombinant		111

II-2.	Cinétique de réplication des virus parentaux et recombinant	112
II-2.1.	Profil des cinétiques de réplication des virus parentaux et recombinant	112
II-2.2.	Comparaison des cinétiques de réplication	117
<b>III-</b>	<b>DISCUSSION</b>	<b>122</b>

## **PARTIE 2. ETUDE DE L'ÉMERGENCE *IN VITRO* ET *IN VIVO* DES FORMES RECOMBINANTES INTER-GROUPES M ET O** **127**

### **A. ETUDE *IN VITRO* DE LA RECOMBINAISON AU SEIN DES RÉGIONS LTR** **129**

#### **CONTEXTE ET OBJECTIFS DU TRAVAIL EXPÉRIMENTAL** **129**

#### **I- MÉTHODES** **132**

I-1.	Etude de l'émergence d'un point de cassure dans les LTRs	133
I-1.1.	Recherche d'un point de cassure dans les LTRs	133
I-1.2.	Etude de la cinétique d'émergence des points de cassure	137
I-2.	Détermination des motifs de recombinaison	138
I-2.1.	Populations ARN et ADN	138
I-2.2.	SGA et séquençage	138
I-2.3.	Recherche de la prédominance d'un profil préférentiel	139
I-3.	Comparaison aux données <i>in vivo</i>	140

#### **II- RÉSULTATS** **140**

II-1.	Etude de l'émergence d'un point de cassure dans les LTRs	140
II-1.1.	Recherche d'un point de cassure dans les LTRs	140
II-1.2.	Etude de la cinétique d'émergence des points de cassure <i>in vitro</i>	142
II-2.	Détermination des motifs de recombinaison	146
II-2.1.	SGA et séquençage	146
II-2.3.	Recherche de la prédominance d'un profil préférentiel	153
II-3.	Comparaison aux données <i>in vivo</i>	154

#### **III- DISCUSSION** **157**

### **B. IDENTIFICATION D'UNE CO-INFECTIION VIH-1/M+O DU FAIT DE L'ÉMERGENCE *IN VIVO* D'UN RECOMBINANT VIH-1/MO** **161**



<b>VALORISATION</b>	<b>161</b>
<b>CONTEXTE ET OBJECTIFS DU TRAVAIL EXPÉRIMENTAL</b>	<b>161</b>
<b>I- MÉTHODES</b>	<b>163</b>
I-1. Analyses sur le prélèvement de novembre 2013	163
I-1.1. Recherche de variants par sérotypage	163
I-1.2. Confirmation moléculaire de double infection avec recombinant	165
I-1.3. Quantification des différentes populations virales	167
I-2. Analyses sur les prélèvements séquentiels et collecte des données immuno-virologiques et thérapeutiques	168
I-2.1 Echantillons et analyses	168
I-2.2 Collecte des données immuno-virologiques et thérapeutiques	168
<b>II- RÉSULTATS</b>	<b>170</b>
II-1. Analyses sur le prélèvement de novembre 2013	170
II-1.1. Recherche de variants par sérotypage	170
II-1.2. Confirmation moléculaire de double infection avec recombinant	170
II-1.3. Quantification des différentes populations virales	176
II-2. Analyses sur les prélèvements séquentiels et évolutions immuno-virologique et thérapeutique	177
II-2.1 Echantillons et analyses	178
II-2.2 Evolutions immuno-virologique et thérapeutique durant toute la durée de suivi	179
<b>III- DISCUSSION</b>	<b>183</b>
<b>CONCLUSION - PERSPECTIVES</b>	<b>187</b>
<b>REFERENCES BIBLIOGRAPHIQUES</b>	<b>195</b>
<b>ANNEXES</b>	<b>211</b>

## Liste des tableaux

TABLEAU 1 : Circonstances de découverte des 20 doubles infections VIH-1/M+O.....	56
TABLEAU 2 : Chronologie de la découverte des 23 formes recombinantes VIH-1/MO.....	58
TABLEAU 3 : Nomenclature et caractéristiques des CMIC à générer .....	86
TABLEAU 4 : Séquences des amorces Gibson Assembly testées selon les CMIC à générer .....	87
TABLEAU 5 : Séquences des amorces utilisées lors de la vérification des constructions .....	90
TABLEAU 6 : Amorces utilisées lors des vérifications sur les virus recombinants VIH-1/OM produits .....	95
TABLEAU 7 : Vérifications réalisées sur les CMIC pVIH-1/OM et pVIH-1/MO générés avec la méthode Gibson Assembly .....	97
TABLEAU 8 : Contrôles sur la progénie virale VIH-1/OM-GC, à J10 de co-culture.....	102
TABLEAU 9 : Contrôles sur les progénies virales VIH-1/MO, produites en fin de transfection .....	102
TABLEAU 10 : Amorces utilisées lors la caractérisation des régions Vpr des virus produits .....	109
TABLEAU 11 : Caractérisation des mélanges de virus parentaux et recombinant .....	111
TABLEAU 12 : Caractéristiques des cinétiques de réplication de chacune des souches étudiées....	113
TABLEAU 13 : Valeurs moyennes et différences calculées pour les différents paramètres étudiés	120
TABLEAU 14 : Nomenclature des extraits étudiés .....	132
TABLEAU 15 : Amorces utilisées pour l'amplification des LTRs recombinants de motif MO .....	135
TABLEAU 16 : Localisation du point de cassure dans les 5'LTR-ARN des 16 extraits étudiés .....	145
TABLEAU 17 : Répartition des 5'LTR-ARN analysés après SGA .....	147
TABLEAU 18 : Proportion des points de cassure recensés dans les 5'LTR-ARN à J6, J16 et J34 après SGA .....	148
TABLEAU 19 : Répartition des LTR-ADN analysés après SGA.....	149
TABLEAU 20 : Proportion des points de cassure recensés dans les 5' et 3'LTR-ADN à J34 après SGA .....	150
TABLEAU 21 : Proportion des points de cassure caractérisés dans les LTR-ADN, 5' et 3'LTRs confondus, à J34 après SGA .....	152
TABLEAU 22 : Evolution de la fréquence des trois motifs de recombinaison majoritaires dans les LTRs au cours de la co-culture.....	153
TABLEAU 23 : Localisation des points de cassure dans les LTRs recombinants de 9 URF_MO.....	154
TABLEAU 24 : Comparaison de la localisation des points de cassure décrits <i>in vitro</i> et <i>in vivo</i> .....	155
TABLEAU 25 : Etude de la diversité des fragments impliqués dans LTRs recombinants des 9 URF_MO .....	156
TABLEAU 26 : Caractéristiques des techniques de charge virale utilisées au cours du suivi de la patiente. ....	169
TABLEAU 27 : Valeurs de charges virales du prélèvement de novembre 2013.....	176

TABLEAU 28 : Synthèse des données immuno-virologiques entre 1999 et 2018.....	181
TABLEAU 29 : Géotypes cumulés de résistance de 2018 .....	183

## Liste des figures

FIGURE 1 : Structure d'une particule virale de VIH-1. ....	25
FIGURE 2 : Organisation génomique des VIH-1 et VIH-2 et fonctions des protéines.....	26
FIGURE 3 : Cycle de réplication du VIH au sein d'une cellule-hôte. ....	28
FIGURE 4 : Mécanisme de la transcription inverse.....	29
FIGURE 5 : Mode d'action antiviral des facteurs de restriction cellulaires et contournement par le VIH. ....	32
FIGURE 6 : Arbre phylogénétique incluant les séquences de VIScpzPts, VIScpzPtt, VISgor et de VIH-1. ....	35
FIGURE 7 : Processus de recombinaison chez les rétrovirus. ....	40
FIGURE 8 : Proposition de mécanismes de recombinaison. ....	41
FIGURE 9 : Répartition globale des points de cassure le long du génome de VIH-1. ....	44
FIGURE 10 : Conséquences possibles de l'évolution virale par recombinaison. ....	46
FIGURE 11 : Organisation génomique de la forme recombinante CRF02_AG. ....	49
FIGURE 12 : Organisation génomique de la forme recombinante CRF03_AB.....	49
FIGURE 13 : Organisation génomique de la forme recombinante CRF56_cpx.....	49
FIGURE 14 : Emergence des variants du VIH-1 : facteurs de restriction et stratégies de contournement. ....	51
FIGURE 15 : Mise en évidence de l'émergence récente du gène <i>asp</i> , par une approche phylogénétique. ....	52
FIGURE 16 : Divergence génétique entre les VIH-1/M et VIH-1/O, en fonction des régions du génome.....	53
FIGURE 17 : Comparaison des arbres phylogénétiques de VIH-1/O et de VIH-1/M .....	54
FIGURE 18 : Organisation génomique de la forme recombinante HIV2_CRF01_AB.....	55
FIGURE 19 : Algorithme séro-moléculaire de dépistage des recombinants <i>vpr</i> .....	60
FIGURE 20 : Diversité génétique des fragments impliqués dans les 19 formes recombinantes identifiées jusqu'à présent. ....	64
FIGURE 21 : Localisation des points de cassure dans le génome des 19 URF_MO décrites actuellement. ....	66
FIGURE 22 : Répartition des profils de recombinaison de 11 des 12 URF_MO complètement séquencés.....	78
FIGURE 23 : Représentation des génomes de VIH recombinants. ....	81
FIGURE 24 : Génome complet de RBF208 .....	82
FIGURE 25 : Principe de la méthode Gibson Assembly, appliqué à la construction pVIH-1/OM.....	83
FIGURE 26 : Mode d'action des différentes enzymes impliquées dans la réaction Gibson Assembly	84
FIGURE 27 : Localisation des PCRs de vérification sur les CMIC générés .....	90

FIGURE 28 : Principe général de la production de virus recombinants VIH-1/OM à partir d'un CMIC .....	92
FIGURE 29 : Profils des CMIC pVIH-1/OM.....	96
FIGURE 30 : Profils des CMIC pVIH-1/MO.....	99
FIGURE 31 : Synthèse générale de la génération de CMIC et de la production de virus recombinants .....	103
FIGURE 32 : Cinétique de répliation des triplicats de VIH-1/M .....	114
FIGURE 33 : Cinétique de répliation des triplicats de VIH-1/O .....	115
FIGURE 34 : Cinétique de répliation des triplicats de VIH-1/OM.....	116
FIGURE 35 : Cinétiques moyennes de répliation des souches parentales et recombinante .....	118
FIGURE 36 : Localisation et nombre de points de cassure retrouvés dans le LTR.....	130
FIGURE 37 : Proposition de mécanismes de génération de provirus recombinants, lors du premier transfert de brin obligatoire et conséquences sur les LTRs. ....	131
FIGURE 38 : Principe général de l'étude des LTRs du recombinant VIH-1/OM généré <i>in vitro</i> . ....	133
FIGURE 39 : Localisation des amorces utilisées pour l'amplification des LTRs recombinants. ....	136
FIGURE 40 : Evolution de la multiplication du virus recombinant VIH-1/OM dans les cellules Jurkat. ....	141
FIGURE 41 : Données de séquençage montrant les zones de recombinaison des LTR-ARN et LTR-ADN.....	142
FIGURE 42 : Profil de migration sur gel d'agarose des amplicons 5' et 3'LTR-ARN avant la co-culture. ....	143
FIGURE 43 : Données de séquençage de la région R des LTRs des surnageants de transfection (J0) .....	143
FIGURE 44 : Profil de migration sur gel d'agarose des amplicons 5'LTR-ARN.....	144
FIGURE 45 : Motif des trois points de cassure mis en évidence dans les régions R des 5'LTR-ARN après SGA. ....	148
FIGURE 46 : Localisation et nombre de points de cassure dans les 5' et 3'LTR-ADN.....	151
FIGURE 47 : Localisation et fréquence des points de cassure dans les LTR-ADN.....	153
FIGURE 48 : Localisation et fréquence des points de cassure décrits <i>in vitro</i> et <i>in vivo</i> .....	156
FIGURE 49 : Discordance observée lors du génotypage de résistance réalisé par le laboratoire de virologie du CHU de Rennes. ....	163
FIGURE 50 : Régions ciblées par les méthodes d'exploration du CNR .....	164
FIGURE 51 : Résultats des PCRs nichées spécifiques de groupes. ....	170
FIGURE 52 : Résultats combinés des PCRs nichées spécifiques de groupes et hétérologues.....	171
FIGURE 53 : Résultats de l'analyse de similarité réalisée sur les trois fragments chevauchants de RBF235.....	172
FIGURE 54 : Simplot de la séquence du génome complet de RBF235. ....	173
FIGURE 55 : Arbres phylogénétiques positionnant RBF235 parmi les URF_MO décrites. ....	174

FIGURE 56 : Synthèse des résultats du séquençage de génome complet de RBF235. ....	175
FIGURE 57 : Exploration des prélèvements séquentiels de RBF235. ....	177
FIGURE 58 : Synthèse des données immuno-virologiques et thérapeutiques entre 1999 et 2018. ....	180
FIGURE 59 : Schéma général représentatif du processus pyramidal sélectif de génération et de sélection des formes recombinantes au sein des VIH.....	193



## Liste des abréviations

3'PPT	<i>3'PolyPurine Tract</i>
AgP24	Antigène P24
ADN	Acide DésoxyriboNucléique
ADN(-)	brin d'ADN négatif
ADN(+)	brin d'ADN positif
ADNc	ADN complémentaire
ARN	Acide RiboNucléique
ARNg	ARN génomique
ARNm	ARN messenger
ARN pré m	ARN pré messenger
ARV	AntiRétroViral
Asp	Protéine anti-sens ( <i>AntiSense Protein</i> )
BHI	<i>Brain Heart Infusion</i>
β-ME	β-Mercapto-Ethanol
CDC	Agence fédérale pour le contrôle et la prévention des maladies ( <i>Center for Disease Control and Prevention</i> )
CFAR	<i>Center For AIDS Reagents</i>
CHU	Centre Hospitalier Universitaire
CMH	Complexe Majeur d'Histocompatibilité
CMI	Clone Moléculaire Infectieux
CMIC	Clone Moléculaire Infectieux Chimérique
CMSP	Cellule Mononuclée du Sang Périphérique
CNR	Centre National de Référence
CPC	Centre Pasteur du Cameroun
cPPT	<i>central PolyPurine Tract</i>
CRF	Forme Recombinante Circulante ( <i>Circulant Recombinant Form</i> )
CV	Charge Virale
DMEM	<i>Dulbecco's minimal essential medium</i>
dNTP	désoxyribonucléotide ( <i>désoxy-N tri-phosphate</i> )
ddNTPs	didésoxyribonucléotide ( <i>didésoxy-N tri-phosphate</i> )
ELISA	Test immuno-enzymatique ( <i>Enzyme Linked ImmunoSorbent Assay</i> )
Env	Enveloppe
GA	Gibson Assembly
GC	Genecust
gp160	Glycoprotéine d'enveloppe précurseur



gp120	Glycoprotéine d'enveloppe de surface
gp41	Glycoprotéine d'enveloppe transmembranaire
GPI	GlycosylPhosphatidyInositol
HTLV	Virus T-Lymphotrope humain ( <i>Human T-Lymphotropic Virus</i> )
IFN	Interféron
II	Inhibiteur de l'Intégrase
IN	Intégrase
INNTI	Inhibiteurs Non Nucléosidiques de la Transcriptase Inverse
IP	Inhibiteur de Protéase
LAV	<i>Lymphadenopathy Associated Virus</i>
LB	Luria-Butani
LTR	Séquence terminale longue répétée ( <i>Long Terminal Repeat</i> )
NGS	Séquençage de nouvelle génération ( <i>Next Generation sequencing</i> )
NIBSC	<i>National Institute for Biological Standards and Control</i>
pb	paire de bases
PBS	<i>Primer Binding Site</i>
PCR	Réaction de polymérisation en chaîne ( <i>Polymerase Chain Reaction</i> )
PR	Protéase
PvVIH	Personnes vivant avec le VIH
RES-O	Réseau national de surveillance des infections à VIH-1/O
RPMI	<i>Roswell Park Memorial Institute</i>
RT-PCR	<i>Reverse Transcriptase-PCR</i>
SGA	Amplification d'un génome unique ( <i>Single Genome Amplification</i> )
SIDA	Syndrome d'Immunodéficience Acquise
SOC	Milieu de croissance hypernutritif ( <i>Super Optimal broth with Catabolite répression</i> )
TCID50	Dose infectant 50 % des cellules de culture ( <i>Tissue Culture Infective Dose 50</i> )
TI	Transcriptase Inverse
URF	Forme Recombinante Unique ( <i>Unique Recombinant Form</i> )
VIH	Virus de l'Immunodéficience Humaine
VIS	Virus de l'Immunodéficience Simienne
VIScpz	Virus de l'Immunodéficience Simienne du chimpanzé
VISgor	Virus de l'Immunodéficience Simienne du gorille
VISsm	Virus de l'Immunodéficience Simienne du mangabey enfumé

# **ANALYSE BIBLIOGRAPHIQUE**



# I- Généralités sur les Virus de l'Immunodéficience Humaine (VIH)

## I-1. Circonstances de découverte des VIH

Il n'existe pas "un" mais "des" VIH, qui possèdent tous une organisation génomique et une structure communes.

En 1981, le CDC (*Center for Disease Control and Prevention*), correspondant à l'agence fédérale des Etats-Unis pour le contrôle et la prévention des maladies, a diffusé une alerte sur une épidémie en cours depuis la fin des années 1970 à New York, San Francisco et Los Angeles, survenant chez des jeunes patients, homosexuels ou usagers de drogues intraveineuses, anormalement immunodéprimés <sup>[1]</sup>. Cette épidémie était caractérisée par des cas de sarcomes de Kaposi, de pneumonies à *Pneumocystis jirovecii* et d'autres infections opportunistes et a d'abord été appelée *syndrome gay* puis Syndrome d'Immunodéficience Acquise (SIDA). Au début des années 1980, des cas similaires surviennent aussi en France, chez des patients hémophiles, et l'hypothèse d'une origine virale a rapidement été privilégiée. Le premier VIH a été isolé en 1983 à l'Institut Pasteur, à Paris, et avait été baptisé LAV (pour *Lymphadenopathy Associated Virus*) et HTLV-III (pour *Human T-Lymphotropic Virus*) <sup>[2, 3]</sup>. La dénomination de VIH est adoptée en 1986 par la communauté scientifique. L'existence et la circulation d'autres variants, génétiquement et/ou antigéniquement différents, ont été mises en évidence par la suite, nécessitant l'établissement d'une classification en types et en groupes. La découverte, la même année, d'un autre VIH présentant une forte divergence génétique, par rapport à la première souche isolée, a ainsi conduit à la différenciation en VIH de type 1 (VIH-1) et de type 2 (VIH-2) <sup>[4]</sup>. La mise en évidence d'autres virus présentant des divergences génétiques moins marquées, mais notables, a conduit à la définition de quatre groupes de variants VIH-1. Le premier correspondait à celui lié à la première souche découverte en 1983, devenu majoritaire, et a été désigné VIH-1 de groupe M (VIH-1/M) pour "*Major*". Les trois autres groupes, tous identifiés chez des personnes d'origine camerounaise, ont été classés en groupe O (VIH-1/O)

pour "*Outlier*" en 1994 <sup>[5, 6]</sup>, en groupe N (VIH-1/N) pour "*Non M, non O*" en 1998 <sup>[6]</sup> et enfin, plus récemment, en groupe P (VIH-1/P) en 2009 <sup>[7]</sup>, afin de suivre la nomenclature existante.

## I-2. Caractéristiques virologiques

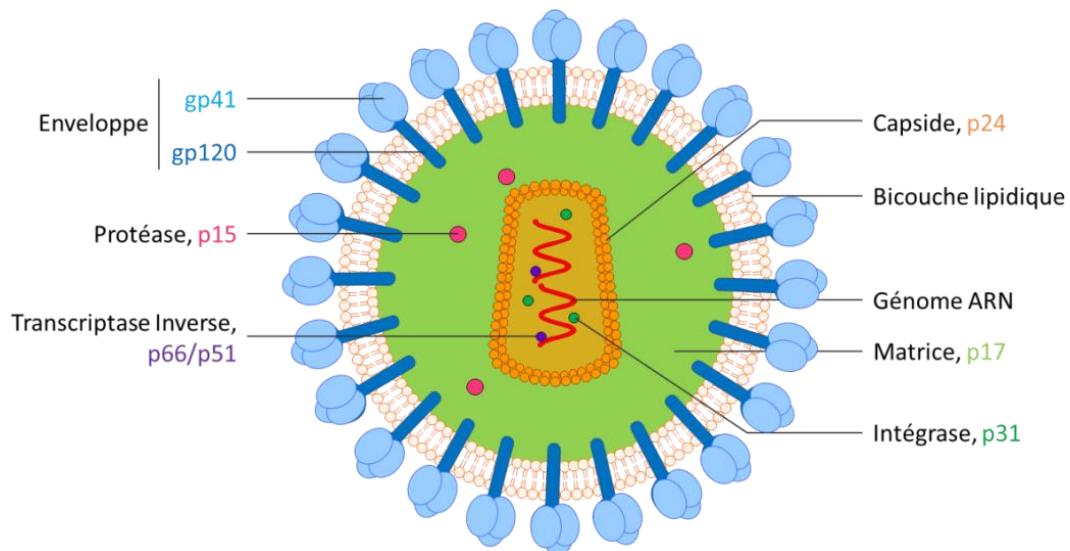
Les VIH ont tous en commun d'être des virus enveloppés, mesurant entre 90 et 120 nanomètres de diamètre. Leur génome est constitué de deux molécules d'ARN monocaténaire structuralement identiques et de polarité positive. Ils appartiennent à la famille des *Retroviridae*, à la sous-famille des *Orthoretrovirinae* et au genre *lentivirus*. Cette grande famille des *Retroviridae* est caractérisée par la présence d'une enzyme spécifique, la transcriptase inverse (TI), permettant la transcription inverse ou rétrotranscription de son génome viral d'ARN monocaténaire en ADN bicaténaire, et par l'intégration de ce dernier dans l'ADN chromosomique de la cellule-hôte <sup>[8, 9]</sup>.

## I-3. Particule virale

La structure de la particule virale de VIH est étroitement liée au cycle de réplication du virus. Chacune de ses étapes est, en effet, gouvernée par une ou plusieurs protéines virales, qui permettent son bon et complet déroulement.

Chaque particule virale contient deux brins d'ARN de structure génomique identique mais de séquence génétique différente. Le virus est donc hétéro-diploïde. De plus, elle comporte trois éléments structuraux : l'enveloppe, la matrice et la capsid (FIGURE 1).

L'enveloppe correspond à une bicouche de phospholipides, issue de membranes cellulaires, associés à deux glycoprotéines virales : la gp120 et la gp41. Ces glycoprotéines d'enveloppe (Env) sont initialement traduites en une glycoprotéine précurseur, la gp160, qui est ensuite traitée et clivée au cours de son transport, via la voie sécrétoire, pour donner la gp120 et la gp41. La gp41, ancrée dans la membrane, est liée de manière non covalente à la sous-unité de surface gp120, qui possède 5 régions hypervariables (de V1 à V5). Les glycoprotéines d'enveloppe ont un rôle essentiel dans l'attachement du virus au récepteur CD4 de la cellule et dans la fusion avec la membrane cellulaire.



**FIGURE 1** : Structure d'une particule virale de VIH-1.  
(Adapté de <https://www.motifolio.com/>)

La matrice, formée par la polymérisation de la protéine p17, tapisse la face interne de l'enveloppe, entoure la capsid. Entre la matrice et la capsid se situe la Protéase (PR) p15, enzyme virale participant à la maturation des virions néoformés.

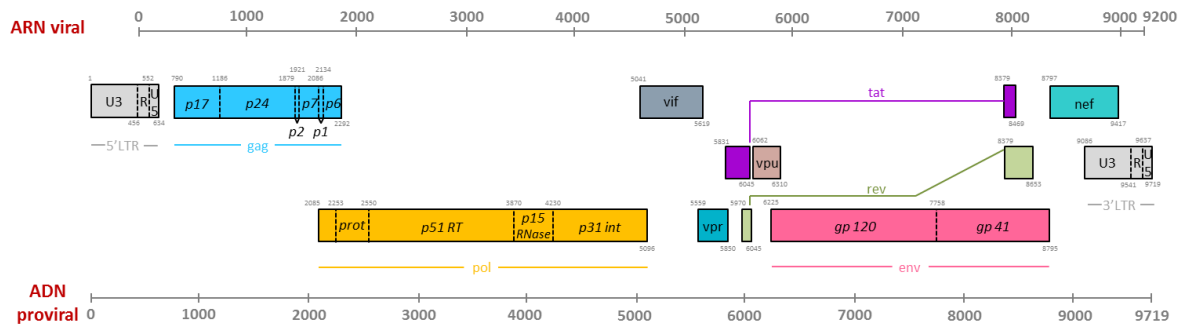
La capsid, en forme de cône tronqué, provient de l'assemblage de la protéine p24. La capsid protège la nucléocapsid, formée de l'association des deux brins d'ARN avec des nucléoprotéines p7. A l'intérieur de la capsid, la TI (p66/p51), enzyme clé intervenant en début de cycle de réplication, est associée à l'Intégrase (IN) p31, enzyme permettant l'intégration de l'ADN proviral au génome de la cellule-hôte.

#### I-4. Organisation du génome

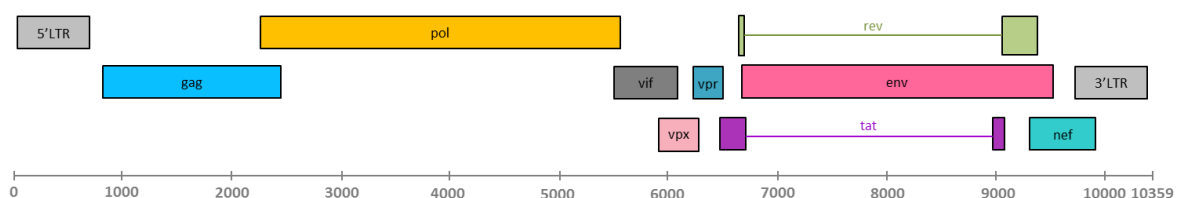
D'une longueur de 9200 paires de bases (pb) environ, le génome ARN du VIH possède trois cadres de lectures ouverts, chevauchants (FIGURE 2). Après transcription inverse de l'ARN viral en un ADN double brin et intégration à l'ADN de la cellule-hôte, le génome proviral obtenu est flanqué de chaque côté, en 5' et en 3', par des séquences terminales longues répétées identiques nommées LTRs (pour Long Terminal Repeat) et mesure environ 9700 pb. Les 5' et 3'LTRs du VIH-1 ont une longueur d'environ 650 pb et sont chacun composés des trois mêmes régions : U3 (unique en 3' de l'ARN), R (répétée aux deux extrémités de l'ARN) et U5 (unique en 5' de l'ARN) (FIGURE 2). Ces séquences participent à l'intégration de l'ADN proviral, à la réplication et à l'expression du génome viral [8]. La séquence promotrice, responsable de la

transcription de l'ARN viral, est celle localisée dans la région LTR à l'extrémité 5' du génome. Etant donné le contexte de séquence, les signaux de terminaison de la transcription et de polyadénylation, quant à eux, se forment exclusivement dans la région LTR à l'extrémité 3'.

## (a) VIH-1



## (b) VIH-2



## (c)

GENES		PROTEINES ET FONCTIONS
<b>LTR</b>	<i>Long Terminal Repeat</i>	Contient le promoteur de la transcription. Intégration du provirus dans le génome de l'hôte et la réplication du génome viral.
<b>gag</b>	<i>group specific antigen</i>	Protéines de structure. (Matrice, Capside, et Nucléocapside)
<b>pol</b>	<i>polymerase</i>	Enzymes virales. (Protéase, Transcriptase Inverse et Intégrase)
<b>env</b>	<i>enveloppe</i>	Glycoprotéines de l'enveloppe (gp120 et gp41). Reconnaissance des récepteur et co-récepteurs de la cellule-hôte.
<b>tat</b>	<i>transcription transactivator</i>	Rétrocontrôle positif de la transcription de l'ADN viral.
<b>rev</b>	<i>regulator of viral expression</i>	Régulation de l'épissage et export des ARNm non épissés et de l'ARN génomique hors du noyau.
<b>vif</b>	<i>viral infectivity factor</i>	Synthèse de l'ADN proviral du VIH-1 et augmentation du pouvoir pathogène du virus.
<b>vpr</b>	<i>viral protein r</i>	Regulation du cycle cellulaire et aide au relargage des virions néoformés. Echappement à la détection par le système immunitaire.
<b>vpu</b>	<i>viral protein u</i>	Bourgeonnement et libération des virions néoformés.
<b>vpx</b>	<i>viral protein x</i>	Nécessaire à la réplication virale efficace et à la pathogénicité.
<b>nef</b>	<i>negative regulatory factor</i>	Nécessaire à la réplication virale efficace. Amélioration de l'infectiosité des particules de VIH-1.

**FIGURE 2 :** Organisation génomique des VIH-1 et VIH-2 et fonctions des protéines.  
(Adapté de Immunobiology <sup>[10]</sup>)

L'organisation des gènes en trois cadres de lecture chevauchants du VIH-1 est représentée en (a). Sous sa forme ARN, le génome viral possède deux LTRs incomplets, constitués en 5' des régions R et U5 et en 3' des régions U3 et R. Une fois rétrotranscrit, le génome devient un ADN double brin d'une longueur de 9719 pb et possède deux LTRs complets, avec les régions U3, R et U5 présentes en 5' et en 3'. L'organisation des gènes en trois cadres de lecture chevauchants des VIH-2 est représentée en (b). Les fonctions des protéines du VIH sont décrites en (c).

Le génome viral comporte trois principaux gènes, *gag*, *pol*, et *env*, codant respectivement pour les protéines de structure interne (Matrice, Capside et Nucléocapside), les enzymes de réplication virale (PR, TI et IN) et les deux glycoprotéines d'enveloppe (gp120 et gp41). Le génome viral comporte également des gènes de régulation, *tat* et *rev*, qui codent pour des protéines ayant un rôle essentiel dans le cycle viral et le pouvoir pathogène du virus, et enfin quatre gènes accessoires, *vif*, *vpr*, *vpu* (VIH-1) ou *vpx* (VIH-2) et *nef* (FIGURE 2). Les protéines accessoires Vif, Vpr, Vpu/Vpx et Nef, qui résultent de l'expression de ces quatre gènes, possèdent diverses fonctions, telles que l'augmentation de l'infectiosité virale, la perturbation des voies cellulaires, l'échappement à la détection immunitaire innée et la neutralisation de l'activité antivirale des facteurs de restriction cellulaires de l'hôte.

En 2016, l'existence d'un dixième gène nommé *asp* (Anti-Sense Protein) a été mise en évidence dans le génome du VIH-1. Ce gène est chevauchant avec celui de l'enveloppe et code pour la protéine anti-sens. Conservé tout au long de l'évolution, il semble spécifique des virus humains pandémiques (VIH-1/M), mais sa fonction demeure inconnue <sup>[11]</sup>.

## I-5. Cycle de réplication virale

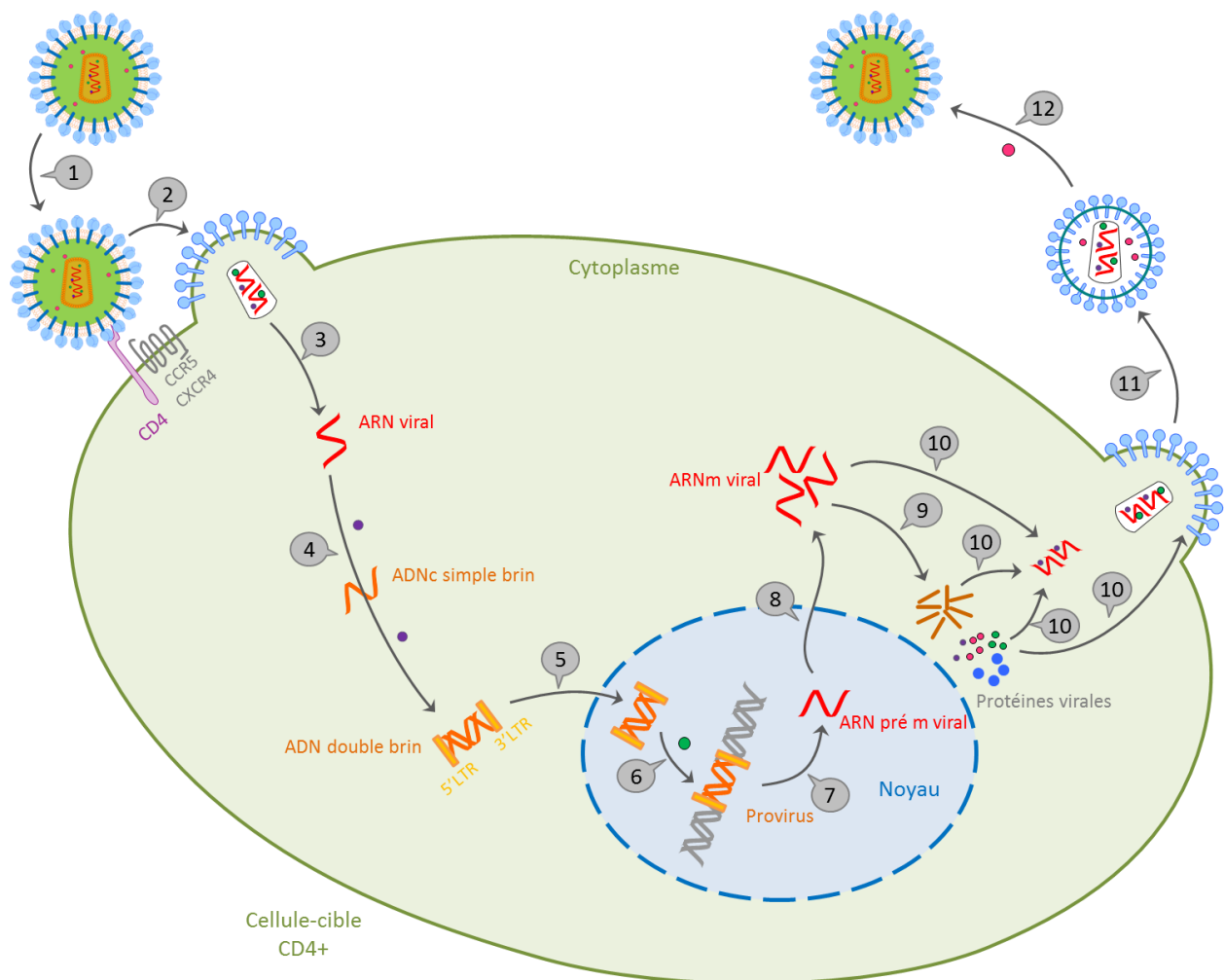
### I.5.1. Déroulement de la réplication virale

Les cellules exprimant à leur surface des récepteurs CD4, tels que les lymphocytes T CD4+, les macrophages, les cellules dendritiques et les cellules de la microglie cérébrale, sont les cibles potentielles d'une infection à VIH. Le cycle de réplication de ce virus au sein de sa cellule-hôte se déroule en plusieurs étapes (FIGURE 3).

Tout d'abord, les glycoprotéines gp120 de la surface virale reconnaissent le récepteur CD4 exprimé à la surface de la cellule cible et s'y attachent. Ce phénomène entraîne un changement conformationnel de la gp120, dont la région variable V3 peut alors se fixer au co-récepteur membranaire (CCR5 et/ou CXCR4), présent à proximité du récepteur CD4. De plus, la gp41, libérée par ce changement de conformation de la gp120, se fixe à son tour à la membrane cellulaire et se replie, ce qui permet la fusion entre l'enveloppe virale et la membrane de la cellule-hôte.

La capside pénètre ensuite dans le cytoplasme et s'y désagrège pour libérer les deux brins d'ARN et les enzymes virales associées.



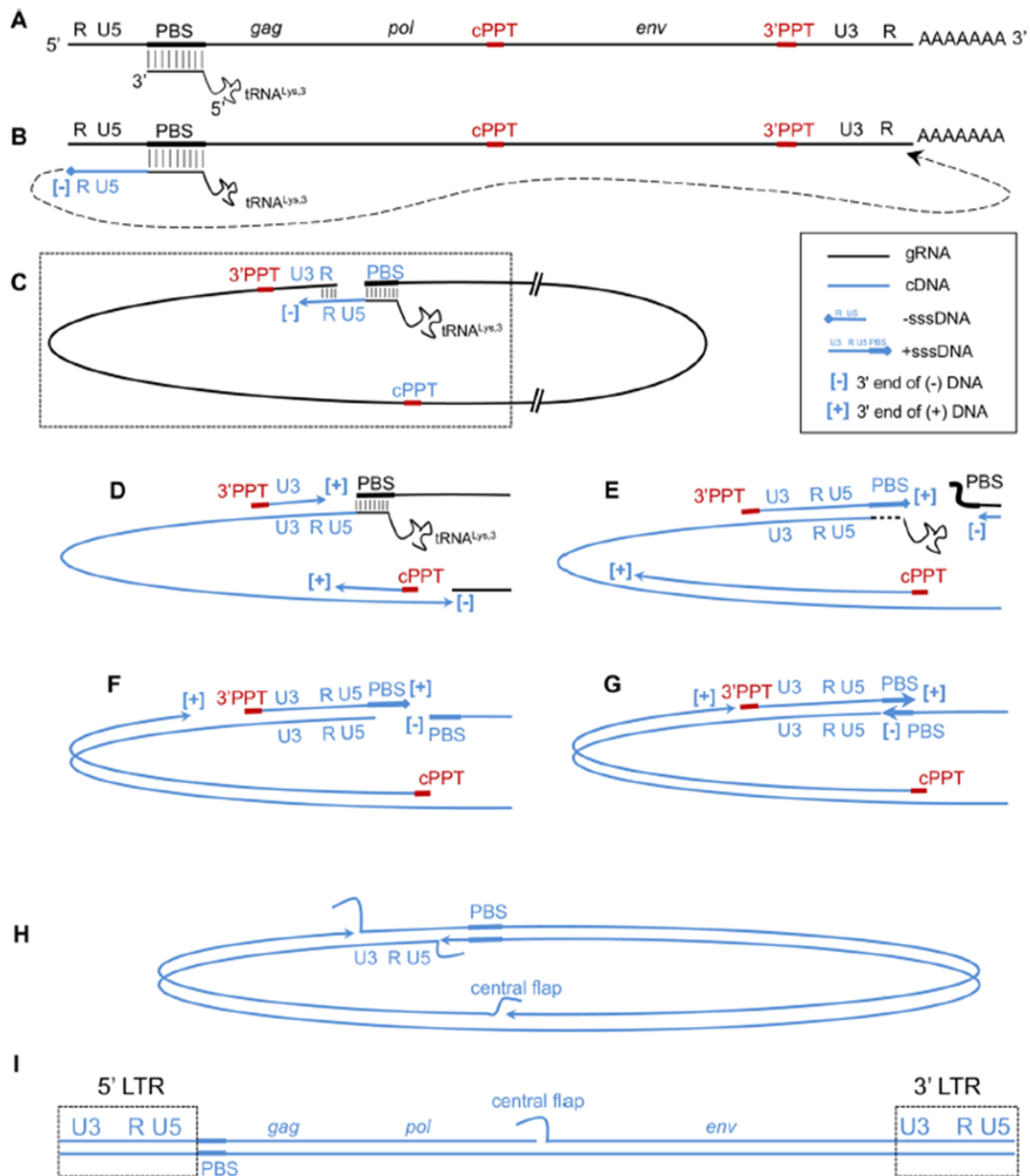


**FIGURE 3** : Cycle de réplication du VIH au sein d'une cellule-hôte.

Les différentes étapes du cycle viral sont représentées par des chiffres.

(1) L'interaction spécifique entre les glycoprotéines d'enveloppe virales et les récepteur et co-récepteurs membranaires permet l'attachement du virus à sa cellule-hôte. (2) La fusion de l'enveloppe virale et de la membrane plasmique entraîne la pénétration de la capside virale dans le cytoplasme cellulaire. (3) Après décapsidation, le génome viral, associé à des enzymes, est libéré dans le cytoplasme. (4) La TI virale (ronds violets) rétrotranscrit l'ARN viral en ADN proviral double brin, flanqué de part et d'autre par les 5' et 3'LTRs complets. (5) Le génome proviral est transporté vers le noyau. (6) L'action de l'intégrase (rond vert) entraîne l'intégration du provirus au génome cellulaire. (7) Le provirus est transcrit en ARN pré-messager par une polymérase cellulaire. (8) Après épissage, l'ARN messager viral est transporté vers le cytoplasme. (9) Le génome joue deux rôles dans la cellule infectée. D'une part, il assure l'expression des protéines virales. Et d'autre part, il est répliqué à l'identique en un grand nombre de copies. (10) Les protéines et les ARN génomiques viraux sont assemblés et encapsidés au niveau de la membrane plasmique. (11) L'acquisition de l'enveloppe a lieu lors du bourgeonnement du virus à la surface de la cellule hôte. (12) Après maturation par la protéase virale activée (rond rose), les virions néoformés peuvent infecter d'autres cellule-hôtes.

Les étapes de décapsidation et de transcription inverse ont lieu en parallèle. Le génome viral subit l'action de la TI, qui possède trois fonctions distinctes : transcriptase inverse, RNase H et ADN polymérase-ADN dépendante. L'initiation de la transcription inverse de l'ARN génomique (ARNg) simple brin viral en un brin d'ADN négatif (ADN (-)) a lieu grâce au tRNA<sup>Lys,3</sup> (FIGURE 4).



**FIGURE 4 : Mécanisme de la transcription inverse.**  
(d'après Cappy et al [12])

(A) L'ARNg viral est représenté de façon schématique avec uniquement les principaux gènes et les régions constitutives du LTR (R et U5 en 5', U3 et R en 3'). L'ARNg est représenté en noir, l'ADN en bleu. L'emplacement du PBS est indiqué par un trait épais. Les séquences d'ARN cPPT et 3'PPT sont indiquées par des traits rouges épais. L'ARN de transfert utilisé pour amorcer la transcription inverse (tRNA<sup>Lys,3</sup>) est représenté hybridé au PBS. (B) La synthèse de l'ADN négatif (ADN (-)) est amorcée par le tRNA<sup>Lys,3</sup> et atteint l'extrémité 5' de l'ARNg avant le premier transfert de brin (flèche pointillée), du fait de la présence de la séquence répétée R. (C) La circularisation de l'ARNg permet la reprise de la synthèse de l'ADN (-). La zone encadrée est celle détaillée en (D-G). (D) La synthèse du brin d'ADN (-) génère le premier LTR complet (U3, R et U5) et continue à travers les régions internes du génome. Les régions 3'PPT et cPPT sont résistantes à la dégradation par l'activité RNase H de la TI et permettent l'initiation de la synthèse du brin d'ADN (+). (E) La synthèse du brin d'ADN (+) commence à partir du 3'PPT pour atteindre le tRNA<sup>Lys,3</sup>, dont seulement une partie est synthétisée. Le PBS de l'ARNg est alors déplacé. (F) La synthèse concomitante de l'ADN (-) permet la copie du PBS de l'ARNg et ensuite sa dégradation ainsi que celle du tRNA<sup>Lys,3</sup>. (G) Les séquences complémentaires des PBS des brins d'ADN (-) et (+) s'hybrident, correspondant au second transfert de brin. Cela permet la reprise de la synthèse de l'ADN. (H) Etant déplacé lui aussi, le LTR du brin d'ADN (-) sert de matrice au brin d'ADN (+). Un 5'LTR double brin est alors généré. La synthèse du brin d'ADN (+) à partir du cPPT déplace celui synthétisé à partir du 3'PPT et génère le 3'LTR double brin. La synthèse de l'ADN (+) initiée à partir du 3'PPT atteint le brin d'ADN (+) synthétisé à partir du cPPT et ne le déplace que partiellement avant de s'arrêter définitivement au site de terminaison central. (I) L'ADN préproviral complet est obtenu, avec les séquences LTR encadrées.

Cet ARN de transfert sert d'amorce à la TI en s'hybridant par complémentarité de séquence au PBS (*Primer Binding Site*), région du génome viral localisée en aval de la région U5 du 5'LTR. Au cours de la synthèse du brin d'ADN (-), l'ARN du segment hybridé ARN/ADN à l'extrémité 5', correspondant aux régions R et U5 du 5'LTR, est détruit par la fonction RNase H de la TI. Le brin d'ADN (-) formé est ensuite déplacé et s'hybride à l'extrémité 3', au niveau de la région R du 3'LTR, du fait de la complémentarité des deux régions R. Il s'agit du premier transfert de brin obligatoire ayant lieu au cours du processus de transcription inverse et conduisant à la circularisation de l'ARNg. La TI, par son activité transcriptase inverse, reprend la synthèse du brin d'ADN (-) à partir de l'ARNg et procède à son élongation jusqu'au PBS. Son activité RNase H permet la dégradation concomitante du brin d'ARNg initial, ainsi que du tRNA<sup>Lys,3</sup>, fixé sur le brin d'ADN (-). La synthèse du brin d'ADN (-) génère le premier LTR complet, constitué des régions U3, R et U5. Du fait de leur résistance à l'activité RNase H de la TI, les 3'PPT (*Polypurine Tract*), région de l'ARNg située en amont de la région U3 du 3'LTR, et cPPT (*central PPT*), situé dans l'IN, subsistent à la fin de l'étape de dégradation de l'ARNg. Ces deux segments d'ARNg servent alors d'amorce à la synthèse du second brin d'ADN, positif (ADN (+)), grâce à l'activité ADN polymérase-ADN dépendante de la TI. La synthèse du brin d'ADN (+) commence à partir du 3'PPT, pour atteindre le PBS de l'ARNg, qui est ensuite dégradé. La complémentarité de séquence des PBS des brins d'ADN (-) et (+) permet leur hybridation au cours du deuxième transfert de brin obligatoire. Le LTR du brin d'ADN (-) sert de matrice à la synthèse de celui du brin d'ADN (+), générant un 5'LTR double brin complet. La synthèse du brin d'ADN (+), initiée à partir du cPPT, entraîne le déplacement du brin d'ADN (+) synthétisé à partir du 3'PPT, et *in fine* la génération du 3'LTR double brin complet. Tout ceci aboutit à la génération d'un ADN proviral double brin, possédant à ses deux extrémités les régions 5' et 3'LTR complètes et dupliquées.

Ainsi, au cours du processus de transcription inverse, les brins d'ADN en cours de synthèse doivent être transférés au moins deux fois d'une matrice à l'autre. Ces deux évènements de transfert de brin obligatoire sont donc nécessaires pour l'obtention d'une molécule d'ADN proviral complète, pouvant par la suite être intégrée au génome de la cellule hôte. Des évènements de transfert de brin facultatif entre les deux copies d'ARNg contenus dans la particule virale peuvent également survenir, lors de la copie des régions internes (entre Gag et Nef) de l'ARNg (choix de copie). Ces évènements de transfert de brin facultatif peuvent être générateurs de recombinaison génétique, dont le phénomène sera détaillé dans le chapitre II-3.2. Recombinaison au sein des VIH (p.39).

L'ADN proviral pénètre ensuite dans le noyau de la cellule-hôte sous la forme d'un complexe de pré-intégration, c'est-à-dire associé à l'IN et à d'autres protéines virales et cellulaires. Ceci permet son intégration dans le génome cellulaire, grâce à l'action de l'IN, enzyme virale liée aux LTRs, et la génération d'un provirus, correspondant à l'ADN viral intégré au chromosome cellulaire.

Le provirus est ensuite transcrit en un ARN pré messager (ARN pré m), par une ARN polymérase cellulaire. Les LTRs possèdent tous les signaux nécessaires au démarrage et à la terminaison de la transcription, qu'ils pilotent <sup>[13]</sup>. Après épissage, l'ARN pré m, devient un ARN messager (ARNm) et est transporté dans le cytoplasme, où se poursuit le cycle viral.

Les ribosomes du réticulum endoplasmique rugueux de la cellule assurent la traduction des ARNm en protéines virales. Les protéines de structure et les enzymes virales sont alors sous la forme de polyprotéines précurseurs Gag et Gag-Pol, respectivement. Ces polyprotéines, ainsi que d'autres protéines virales et l'ARNg viral répliqué, sont transportés vers la membrane plasmique cellulaire, où se déroule l'étape d'assemblage.

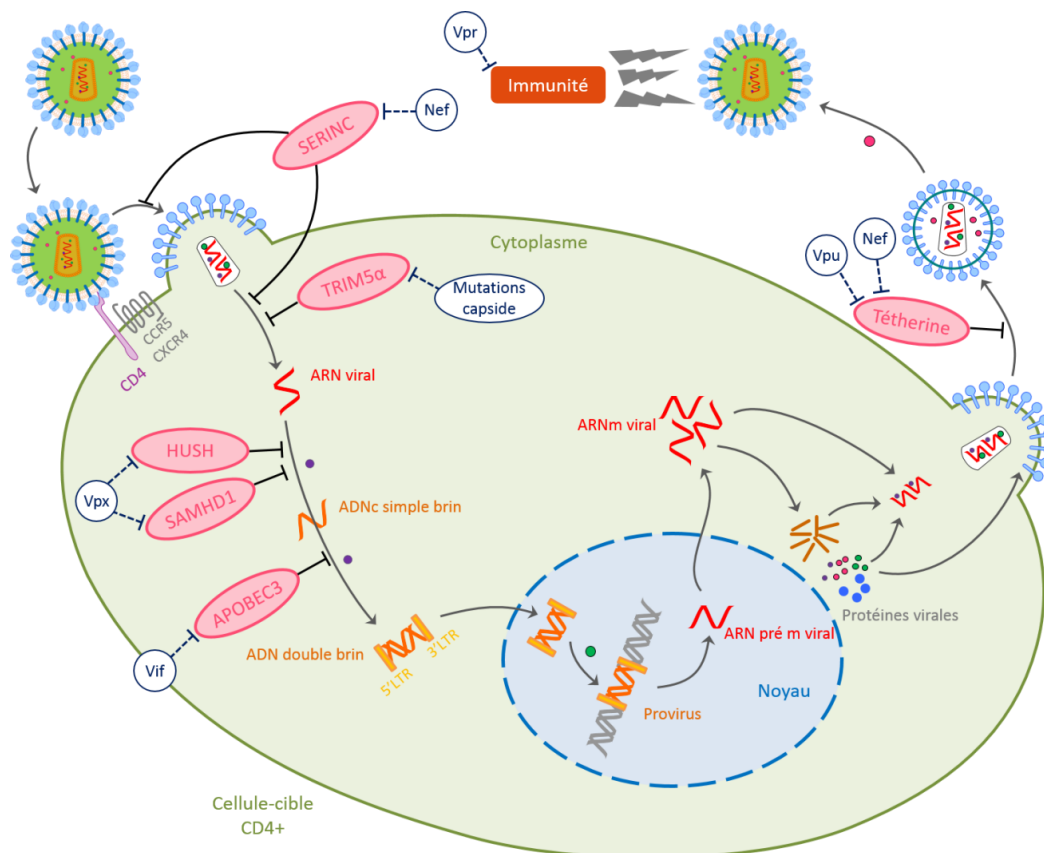
Lors du bourgeonnement, les différents constituants s'entourent d'une bicouche lipidique dérivée de la membrane plasmique. Les glycoprotéines d'enveloppe gp120 et gp41 sont alors incorporées à la surface de la particule virale et y forment des trimères, correspondant aux spicules visibles en microscopie électronique (FIGURE 1, P.25). A ce stade, la particule virale est immature. Une auto-activation de la protéase virale permet le clivage des polyprécurseurs Gag et Gag-Pol. Cette étape de maturation aboutit à la formation de virions, c'est-à-dire de particules virales matures, extracellulaires et pleinement infectieuses. La bonne configuration des particules néoformées est essentielle pour leur assurer une capacité élevée à infecter efficacement de nouvelles cellules cibles.

### **I.5.2. Facteurs de restriction cellulaires et contournement VIH**

Les facteurs de restriction de l'hôte sont définis comme des composants de la défense antirétrovirale innée, qui inhibent la réplication virale et représentent une première ligne de défense contre les virus. Bien que leur expression soit souvent augmentée par les interférons (IFN), beaucoup d'entre eux sont exprimés de manière constitutive, ce qui leur permet d'agir très précocement au cours de l'infection virale <sup>[14]</sup>. Les facteurs de restriction interagissent fréquemment avec les protéines virales, soit pour exercer leurs propriétés antivirales, soit parce qu'ils sont ciblés par des antagonismes viraux, générant en permanence une bataille

moléculaire entre le virus et la cellule [15-17]. Certaines protéines virales jouent le rôle d'antagonistes viraux en facilitant, d'une part, l'échappement du virus à la réponse immunitaire cellulaire de l'hôte et en permettant, d'autre part, sa réplication complète et efficace. Les facteurs de restriction présentent des signatures de pression de sélection caractéristiques de l'évolution et de l'adaptation, témoignant d'une compétition évolutive de longue date entre l'hôte et les virus. Du fait de l'adaptation virus-hôte, les facteurs de restriction sont généralement moins efficaces contre les virus chez leurs hôtes naturels, mais représentent des barrières puissantes contre les transmissions inter-espèces [18]. Enfin, ils sont très sensibles aux mesures de neutralisation ou d'évasion virales, mais lorsqu'ils sont surexprimés, ils entraînent une réduction importante de la production de virus infectieux [19-22].

Les facteurs de restriction anti-VIH contenus dans les cellules infectées montrent une énorme diversité structurale et fonctionnelle et ciblent quasiment toutes les étapes du cycle de réplication virale [19]. En parallèle, le VIH a su développer des stratégies de contournement de cette restriction, grâce notamment à ses protéines accessoires (FIGURE 5).



**FIGURE 5** : Mode d'action antiviral des facteurs de restriction cellulaires et contournement par le VIH.

Seul le mode d'action des facteurs de restriction cellulaires présentant une interaction spécifique avec une protéine accessoire du VIH, pouvant contourner cette activité antivirale, est représenté.

#### a) TRIM5 $\alpha$ / Mutation de la capsid

La protéine **TRIM5 $\alpha$**  bloque l'infection par le VIH en agissant sur plusieurs étapes : la décapsidation, la transcription inverse et le transport nucléaire. Elle agit en modulant la désagrégation de la capsid. Il en résulte une diminution de la quantité d'ADN complémentaire (ADNc) viral synthétisé. Ce facteur de restriction est spécifique d'espèce rétrovirale et semble davantage efficace chez le singe que chez l'Homme <sup>[23, 24]</sup>. Les séquences codant pour la capsid virale influencent la sensibilité virale à la restriction médiée par TRIM5 $\alpha$ . Les particules de VIH contenant des mutations dans leur capsid perturbent l'interaction avec les co-facteurs cellulaires et permettent le contournement de l'activité antivirale de TRIM5 $\alpha$  <sup>[23]</sup>.

#### b) APOBEC3G / Vif

Le facteur **APOBEC3G** participe à la réponse immunitaire innée en agissant sur l'ADN rétroviral simple brin, lors de la transcription inverse <sup>[25-27]</sup>. Il altère la processivité de la TI et entraîne des hyper-mutations non viables pour le virus par une réaction de désamination <sup>[28]</sup>. En inhibant l'incorporation d'APOBEC3G aux particules virales néoformées, la protéine Vif, présente dans les cellules productrices de virus, neutralise son activité antivirale.

#### c) Téthérine / Vpu

Grâce à deux points d'ancrage membranaires, la **Téthérine** peut induire l'adhérence des virions néoformés après assemblage complet, les retenir à la surface de la cellule-hôte et bloquer leur libération. De plus, l'association de la Téthérine aux radeaux lipidiques et son interaction à l'actine du cytosquelette favorise son recrutement aux sites de bourgeonnement du virus <sup>[29]</sup>. Ainsi, la Téthérine est un facteur de restriction efficace, qui contribue au contrôle de la réplication virale *in vivo* et empêche la propagation du virus <sup>[30]</sup>. La protéine Vpu interagit directement avec le domaine transmembranaire de la Téthérine et permet le transport du facteur de restriction dans le réseau trans-golgien ou dans des endosomes précoces pour une dégradation protéosomale ou lysosomale, facilitant ainsi la diffusion virale <sup>[22, 31-34]</sup>.

#### d) SERINC / Nef

Les facteurs de restriction **SERINC3** et **SERINC5** altèrent de manière significative la fusogénicité des particules de VIH avec les cellules cibles, la décapsidation et la libération du contenu viral <sup>[22, 35]</sup>. La présence de Nef dans les cellules productrices réduit nettement l'incorporation de SERINC3/C5 dans le virion via une interaction directe, et avec, en parallèle,

une redistribution du facteur de restriction de la membrane plasmique vers un compartiment membranaire intracellulaire [36-38].

#### e) SAMHD1 / Vpx

**SAMHD1** hydrolyse les desoxyribonucléotides (dNTPs) intracellulaires, précurseurs essentiels à la synthèse de l'ADNc par la TI [39, 40], ce qui permet de diminuer considérablement le taux de dNTPs dans les macrophages et de bloquer l'infection de ces derniers par les VIH-1 et VIH-2 [41, 42]. La protéine accessoire Vpx, codée par le génome des VIH-2 (FIGURE 2B, P.26), assure le contournement de l'action antivirale de SAMHD1, en induisant sa dégradation protéosomale. Ceci permet *in fine* d'augmenter le réservoir de dNTPs intracellulaires et d'optimiser l'action de la TI [43]. Les VIH-1 ne possèdent pas le gène *vpx*, mais parviennent quand même à échapper partiellement à l'action de SAMHD1, sans que l'on n'en connaisse le mécanisme [44, 45].

#### f) HUSH / Vpx

Le complexe **HUSH** agit en compactant la chromatine cellulaire. La répression de l'expression des gènes, qui en résulte, entraîne l'inhibition de la transcription du génome proviral intégré à cette chromatine [46, 47]. La protéine accessoire Vpx du VIH-2 permet au virus d'échapper à cette restriction, en entraînant la dégradation du facteur de restriction HUSH, par la même voie que celle de SAMHD1, mais de façon indépendante. Par ce biais, Vpx assure la pérennité de la réplication virale au sein de la cellule-hôte [48, 49].

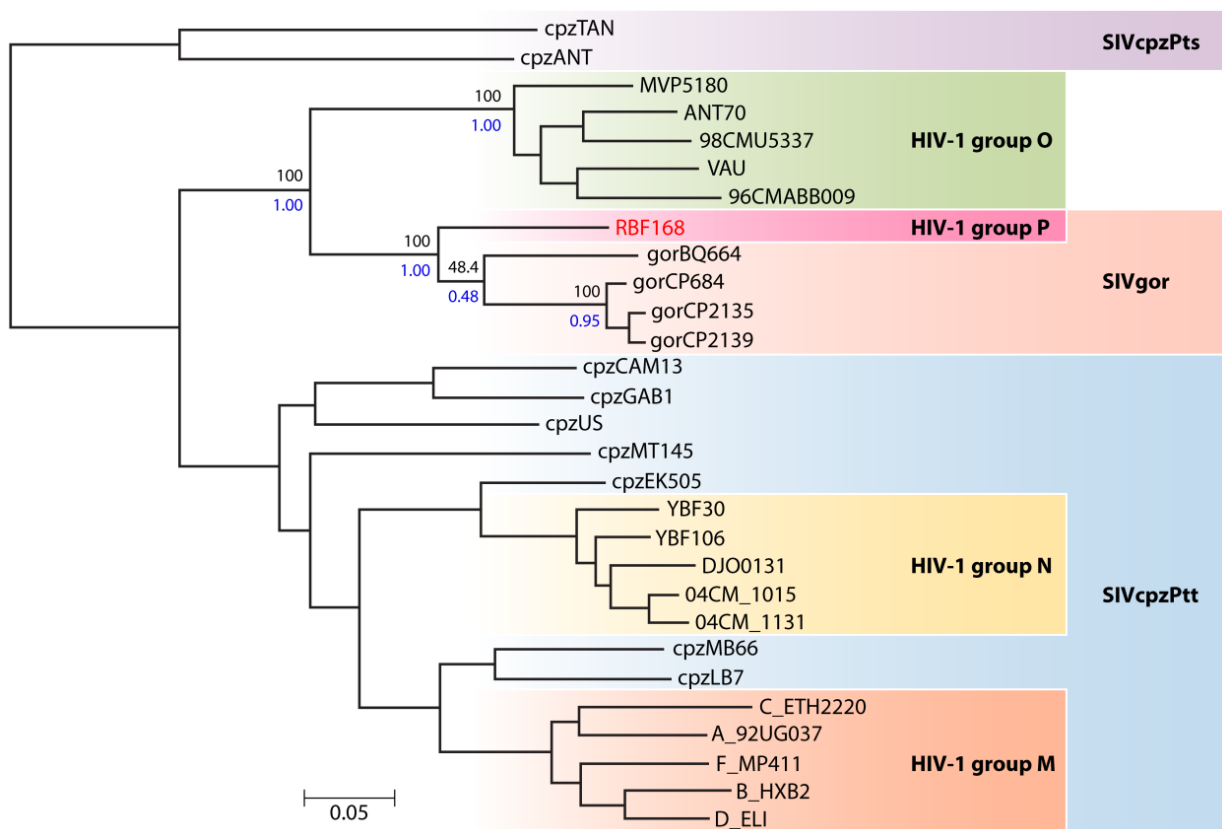
#### g) Immunité / Vpr

La protéine accessoire Vpr est présente chez tous les lentivirus de primates et possède une origine ancestrale commune avec la protéine Vpx [50]. Elle ne cible pas un facteur de restriction particulier. D'une part, elle antagonise la restriction exercée par les macrophages, qui limitent l'expression de Env et la production de descendance infectieuse [51, 52]. D'autre part, elle réduit la production de cytokines antivirales, dans le cadre de la détection immunitaire innée, par l'activation prématurée d'un complexe : le complexe endonucléase SLX4 [22, 53]. Le recrutement de SLX4 par Vpr a lieu pendant la transcription inverse et contribue à éviter que les ADN viraux non productifs soient détectés par les senseurs innés cytosoliques. Ainsi, la protéine accessoire Vpr permet l'échappement à la détection immunitaire innée [53, 54].

## II- Origines, adaptation et diversité génétique des VIH

### II-1. Origines simiennes

L'analyse phylogénétique et l'organisation génomique d'un Virus de l'Immunodéficience Simienne (VIS), identifié chez un chimpanzé en 1989, ont montré que cette souche était proche du VIH-1. Les chimpanzés, notamment l'espèce *Pan troglodytes troglodytes* (VIScpzPtt), ont alors été suspectés d'être la source des VIH-1 [55]. L'arbre phylogénétique de la FIGURE 6, incluant les séquences de VIS de chimpanzé, de gorille (VISgor) et des 4 groupes de VIH-1 (/M, /N, /O et /P), démontre que ces groupes sont le résultat d'épisodes distincts et indépendants de transmission inter-espèces des grands singes à l'Homme [56, 57].



**FIGURE 6** : Arbre phylogénétique incluant les séquences de VIScpzPts, VIScpzPtt, VISgor et de VIH-1. (d'après Plantier et al. [7])

Cet arbre montre que les groupes M, N, O et P du VIH-1 sont la conséquence d'épisodes distincts de transmission inter-espèces de VIScpzPts (*Pan troglodytes schweinfurthii*), VIScpzPtt (*Pan troglodytes troglodytes*), et des VISgor des grands singes à l'Homme.



Les VIS des chimpanzés *Pan troglodytes troglodytes* sauvages ont été identifiés chez des troupes sauvages au sud du Cameroun. Ils présentent une forte diversité génétique, liée à une répartition géographique distincte, ce qui a donné lieu à deux évènements de transmission à l'Homme, correspondant aux VIH-1/M et VIH-1/N [58, 59].

Les VIS des gorilles des plaines occidentales du Cameroun (*Gorilla gorilla gorilla*) sont eux à l'origine des VIH-1/O et VIH-1/P [60]. Or, la position phylogénétique des VISgor parmi les VIScpzPtt implique un ou des passages inter-espèces du chimpanzé au gorille, sans que les modes de transmission soient clairement identifiés. Le gorille serait donc un hôte intermédiaire entre le chimpanzé et l'Homme [61].

Les VIH-2, quant à eux, ont pour origine le VISsmm (pour sooty mangabey monkey) identifié chez un petit singe d'Afrique de l'Ouest, le mangabé enfumé (*Cercocebus atys*) [62].

## II-2. Adaptation à l'Homme

Les premières transmissions du singe à l'Homme ont probablement eu lieu par voie sanguine, lors d'activités humaines à risque de blessure et de contact avec les singes infectés [63], comme la chasse et la préparation de viande de brousse, ou encore les morsures de singes domestiques ou en captivité. Dans de rares occasions, ces infections ont été suivies de l'adaptation du virus à l'Homme et au point de départ de sa diffusion. Une fois la barrière d'espèce franchie avec succès, de multiples facteurs ont permis la progression de l'épidémie dans les populations humaines jusqu'à atteindre la situation dramatique que l'on observe aujourd'hui [64, 65]. Ces facteurs peuvent être biologiques (neutralisation des facteurs de restriction cellulaires, capacité répliquative chez le nouvel hôte), mais aussi environnementaux, sociaux, démographiques, comportementaux et médicaux (injections, transfusions) [66]. Cependant, les différences épidémiologiques entre les différents VIH suggèrent que l'impact de ces différents facteurs n'a probablement pas été le même pour tous les variants.

Ainsi, en comparant les séquences de génomes complets de VIH-1 et de VIScpz, une équipe américaine a identifié le résidu 30R (Arginine, basique) dans la protéine de matrice p17 comme acide aminé signature des VIH-1/M, /N et /O [67] alors que tous les VIScpzPtt et tous les VISgor présentent un résidu 30M (Méthionine, hydrophobe) [68]. Cette convergence évolutive

vers un même résidu M30R après trois évènements indépendants de transmission inter-espèces suggère fortement l'adaptation à une pression de sélection spécifique de l'hôte. Par contre, la souche prototype de VIH-1/P, présente un résidu 30M, comme les souches de VIScpz et VISgor<sup>[7]</sup>. L'absence du résidu 30R n'a donc pas empêché la transmission de ce virus à l'Homme. Cependant, elle pourrait se traduire par une moindre adaptation à cet hôte responsable de sa très faible diffusion, deux souches ayant seulement été découvertes à ce jour, et la seconde présentant, elle, un résidu 30K<sup>[69, 70]</sup>.

Par ailleurs, l'adaptation aux facteurs de restriction cellulaires est également un élément ayant permis la diffusion du VIH-1/M, pandémique, au sein de la population humaine. En effet, une fois la barrière d'espèce franchie, les VIScpz et VISgor se sont heurtés à certains facteurs de restriction cellulaires spécifiques de l'Homme. Certaines de leurs stratégies de contournement, élaborées au sein de leurs hôtes simiens respectifs, sont alors devenues inefficaces. C'est notamment le cas de la Téthérine, la Téthérine humaine contenant une délétion dans son domaine cytoplasmique qui la rend largement résistante à la protéine accessoire Nef, utilisée par les VIScpz et VISgor pour contourner cette restriction. L'activité anti-téthérine, initialement portée par la protéine accessoire Nef des VIScpz et des VISgor, a ainsi été transférée à la protéine Vpu des VIH-1/M<sup>[33, 71]</sup>. Et la perte de l'activité anti-téthérine de Nef, au profit de Vpu, a constitué un élément majeur d'adaptation du singe à l'Homme

Enfin, malgré de possibles limitations adaptatives, en particulier dans l'efficacité de contournement de l'activité des facteurs de restriction, tous les variants de VIH-1, qu'ils soient de groupe M ou non-M, sont pathogènes pour l'Homme et répliquent fortement, contrairement au VIH-2 qui dérive d'un autre variant de VIS. La proximité génétique entre le Chimpanzé, le Gorille et l'Homme pourrait expliquer pourquoi le VIH-1 conserve une haute capacité répliquative (ou *fitness*) chez l'Homme.

## **II-3. Diversité génétique**

### **II.3.1. Facteurs à l'origine de la diversité génétique**

Plusieurs facteurs sont à l'origine de la grande diversité génétique des VIH.

Les différents franchissements des primates non humains à l'Homme, décrits précédemment, constituent l'un de ces facteurs. Puis, après chaque événement de transmission réussi entre le singe et l'Homme, les différents variants ont évolué selon une pression de sélection et des facteurs épidémiologiques propres, contribuant fortement à la diversification des VIH au cours du temps, après leur transmission initiale. Il est difficile de dater ces événements initiaux de transmission inter-espèces, mais des calculs, basés sur une approche moléculaire, suggèrent que les ancêtres communs les plus récents, responsables de l'épidémie actuelle, datent du début du 20<sup>ème</sup> siècle (1909-1930) pour le VIH-1/M [72], des années 1930 (1914-1944) pour le VIH-1/O [73], des années 1940 pour le VIH-2 [74], de 1963 (1948-1977) pour le VIH-1/N [74], et des années 1980 pour le VIH-1/P, même si trop peu de séquences de VIH-1/P et de VISgor sont disponibles pour dater précisément leur ancêtre commun le plus récent [75].

La variabilité naturelle du virus et l'établissement d'une quasi-espèce contribuent également à la diversité des VIH. En effet, après l'infection d'un individu par un VIH, le virus se multiplie de façon intense au sein de son organisme. Ce haut taux de réplication conduit à la production de plus de  $10^{10}$  particules virales par jour [76]. Les VIH étant des virus à ARN, on observe, de plus, de fréquentes mutations liées à la nature de la TI, ADN polymérase ARN-dépendante, qui ne possède pas de fonction de correction [77, 78]. La faible fidélité de la TI introduit en moyenne une mutation non corrigée par cycle infectieux [79]. Ces mécanismes, associés la pression de sélection naturelle de l'hôte humain, conduisent, au cours de la multiplication dans la cellule, à une population virale très hétérogène désignée quasi-espèce. Cet ensemble de mutants évolue comme s'ils ne formaient "quasiment qu'une seule et même espèce". Chez un même individu, on observe alors une évolution du virus au cours du temps, par des phénomènes de sélection et de dérive génétiques.

La recombinaison génétique, consistant en la formation d'un génome chimérique composé de plusieurs fragments génomiques d'origine parentale différente, accentue également la diversification de la quasi-espèce. Elle est décrite classiquement au sein d'un même segment génomique unique, dans le cas où le virus possède un génome non segmenté. Les études d'épidémiologie moléculaire ont permis de mettre en évidence l'importance de la recombinaison dans la génération de la diversité virale de l'ensemble de la pandémie actuelle [80].

## II-3.2. Recombinaison au sein des VIH

### a) Pré-requis

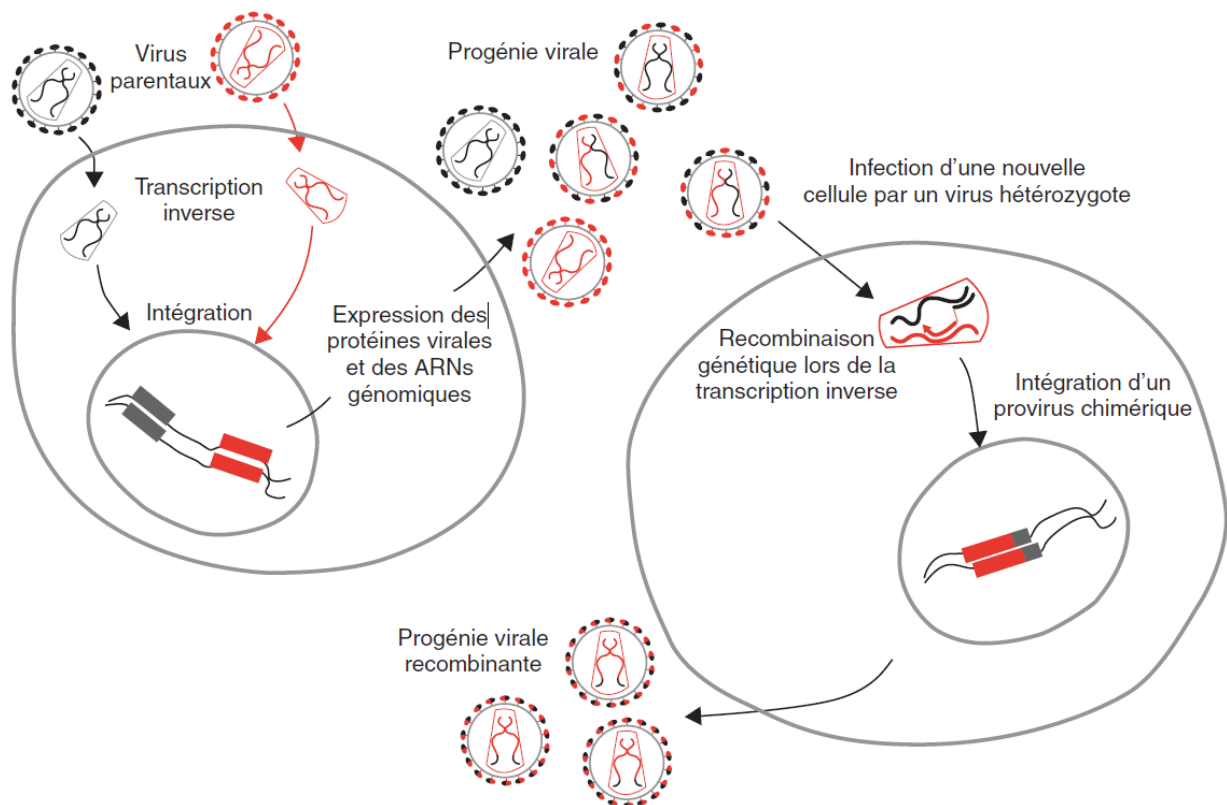
L'infection d'une cellule-hôte est le résultat de l'introduction du virus dans son cytoplasme, de sa décapsidation avec libération du matériel génétique et de l'action de la TI, qui rétrotranscrit l'ARNg viral en un ADN double brin. Chaque particule virale possède, au sein de sa capsid, deux copies hétéro-diploïdes d'ARN de polarité positive, génétiquement différentes [81-83]. Le phénomène de recombinaison, qui a lieu pendant la transcription inverse, conduit donc à la genèse d'un ADN chimérique, portant une partie de l'information génétique contenue dans chacun de ces deux ARN génomiques.

Ainsi, la présence des deux molécules d'ARN génétiquement différentes et l'infection cellulaire par de multiples virus sont le pré-requis essentiel à la recombinaison. A l'échelle même de l'individu, plusieurs variants d'une même quasi-espèce peuvent recombiner entre eux, conduisant à la genèse de formes recombinantes intra-patient. Ces formes recombinantes au sein d'un même patient, sans nouvelle infection préalable, ont un impact sur l'évolution globale des VIH mais n'ont pas de rôle majeur dans la genèse de la diversité génétique, contrairement aux formes recombinantes inter-groupes et inter-sous-types qui conduisent à un "saut" d'évolution. Lorsque les variants proviennent de groupes ou de sous-types différents, les deux virus parentaux doivent être capables d'infecter à la fois le même patient et le même répertoire cellulaire. Cet événement s'observe à l'occasion d'une co-infection, dans le cadre d'une infection simultanée par les deux virus parentaux, ou à l'occasion d'une surinfection, dans le cadre d'infections successives par deux virus parentaux différents.

Dans le cas du VIH, ces situations sont favorisées dans les zones de co-circulation de souches appartenant à différents groupes ou sous-types, ce qui rend importante la probabilité de co-infection d'un même individu par des isolats génétiquement différents. Par ailleurs, le nombre important de recombinants identifiés dans la pandémie prouve, contrairement à ce qui avait été initialement pensé, que l'infection par une souche de VIH chez un individu ne le protège pas de la surinfection [84].

## b) Mécanisme de la recombinaison

La recombinaison des rétrovirus, et donc des VIH, s'effectue en deux phases : une double infection, qui aboutit à la génération de virions hétérozygotes, puis une infection par un virion hétérozygote, qui permet la formation de virions recombinants (FIGURE 7).



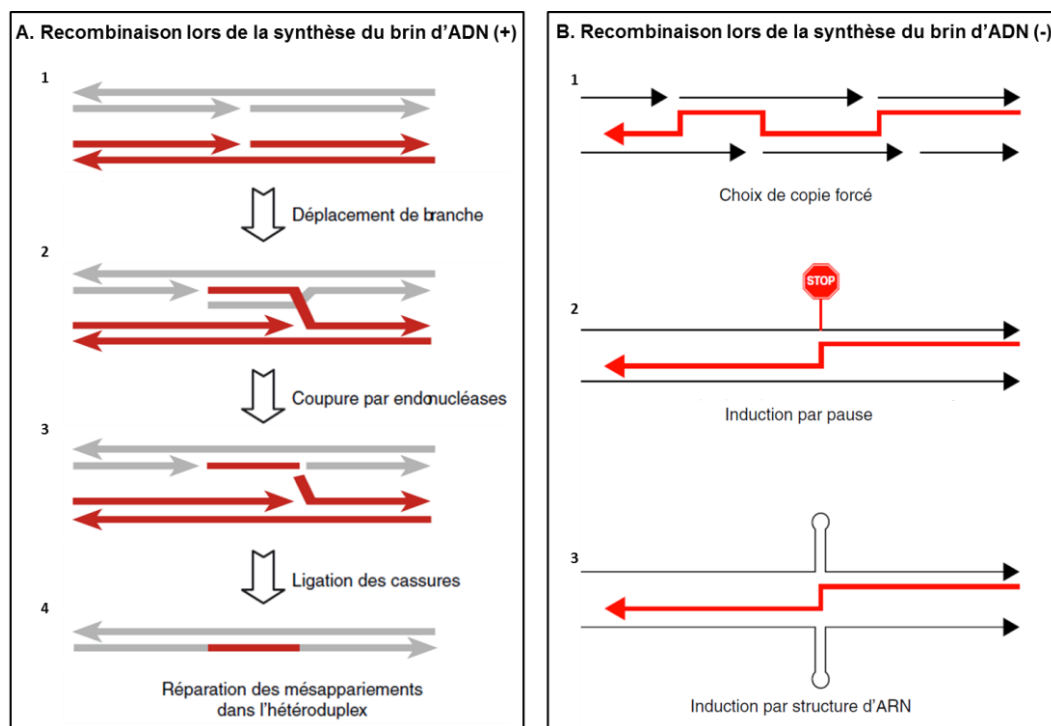
**FIGURE 7 : Processus de recombinaison chez les rétrovirus.**  
(d'après Simon-Lorière et al. [85])

Les virus parentaux A (en noir) et B (en rouge) infectent une même cellule-hôte, au cours d'une double infection. Les étapes de rétrotranscription, d'intégration puis d'expression des protéines virales et des ARNs génomiques ont lieu en parallèle. Au moment de l'assemblage, deux copies d'ARN provenant d'un même virus parental ou de virus parentaux différents sont encapsidées. Après bourgeonnement, la progénie virale contient donc à la fois des virions homozygotes identiques aux virus parentaux A (en noir) ou B (en rouge) et des virions hétérozygotes, avec deux copies d'ARN différentes (en noir et rouge). Un de ces virus hétérozygotes infecte une nouvelle cellule et la présence des deux copies d'ARN différentes entraîne un saut de matrice de la TI, constituant l'évènement de recombinaison génétique. Le provirus au génome chimérique est alors intégré à l'ADN cellulaire. La poursuite du cycle de réplication virale aboutit à la génération d'une progénie virale recombinante et fonctionnelle.

Lors de la première étape, les deux virus parentaux différents infectent la même cellule et subissent l'action de la TI (FIGURE 7). Les ADN résultants sont intégrés au génome cellulaire et les protéines virales et ARNs génomiques sont exprimés. Les génomes, génétiquement différents, doivent alors former un dimère fonctionnel, pouvant être reconnu par la machinerie d'assemblage pour générer un virion hétérozygote. De la même façon, les protéines

structurales virales traduites doivent pouvoir interagir entre elles, afin d'être recrutées correctement par la machinerie lors de l'assemblage des nouvelles particules virales. La progénie virale (c'est-à-dire les virus bourgeonnant à partir de la cellule infectée) est alors une population virale hétérogène, constituée par des virions homozygotes, correspondant aux virus parentaux, et par des virions hétérozygotes, possédant deux brins d'ARN différents (chaque brin correspondant à un parent différent) et devant être fonctionnels pour que la deuxième étape puisse avoir lieu.

Lors de la deuxième étape, une nouvelle cellule est infectée par un virus hétérozygote (FIGURE 7). La recombinaison génétique a lieu pendant la phase de transcription inverse du cycle de réplication. Compte tenu du déroulement de cette phase, la recombinaison génétique peut avoir lieu soit lors de la synthèse du brin d'ADN (-), soit lors de celle du brin d'ADN (+) (FIGURE 8).



**FIGURE 8** : Proposition de mécanismes de recombinaison.  
(d'après Simon-Lorière et al. [85])

Le modèle d'assimilation de branche selon lequel la recombinaison a lieu lors de la synthèse du brin d'ADN (+) est représenté en (A). Les modèles de choix de copie, générant de la recombinaison génétique lors de la synthèse du brin d'ADN (-), sont représentés en (B).

L'hypothèse selon laquelle les molécules d'ADN recombinant sont générées lors de la synthèse du brin d'ADN (+) repose sur des observations en microscope électronique et a abouti

à la proposition du modèle dit "d'assimilation de déplacement de branche" <sup>[86]</sup> (FIGURE 8A). Ce modèle implique que les deux molécules d'ARNg présentes dans la particule virale soient copiées en un ADN (-) entier, ce que de nombreuses études semblent contredire <sup>[87]</sup>. Ainsi, le modèle le plus accepté est désormais celui du "choix de copie" <sup>[88, 89]</sup>, qui suggère que la recombinaison s'opère au moment de la synthèse du brin d'ADN (-) (FIGURE 8B).

La fixation de la TI sur un premier brin d'ARN (ARN donneur du premier parent) initie le processus. Puis ont lieu la transcription de l'ADN et l'expression de l'activité RNase H de la TI. Un arrêt dans la rétrotranscription, suivi du transfert de l'enzyme sur le second brin d'ARN (ARN accepteur du second parent), permet d'entamer la transcription ADN sur le second brin. Il s'agit du phénomène de saut de matrice <sup>[90-94]</sup>.

Plusieurs mécanismes ont été décrits pour expliquer le choix de copie <sup>[85]</sup>. Tout d'abord, le **choix de copie forcé** repose sur l'hypothèse de la discontinuité de l'ARNg. Une cassure de l'ARNg « forcerait » la TI à aller d'une matrice d'ARN à l'autre pour achever la synthèse du brin d'ADN (-)<sup>[95]</sup>. L'**induction par pause** correspond à la présence d'un signal d'arrêt sur l'ARN donneur, ce qui induit le transfert de la TI sur le brin accepteur <sup>[96, 97]</sup>. Enfin, l'**induction par structure d'ARN** repose sur la présence de régions structurées sur la matrice d'ARNg qui ont été décrites comme enclines à favoriser la recombinaison, du fait de la présence de structures secondaires de l'ARN <sup>[98-101]</sup>. La contribution de chacun de ces mécanismes de choix de copie à la recombinaison génétique est avérée, étant donné que les sites fragiles de l'ARNg, les sites de pause de la transcription inverse, ainsi que les structures secondaires de l'ARN constituent des régions où la recombinaison a fréquemment lieu, mais leur rôle dans la variabilité du VIH reste à déterminer.

Selon la proximité génétique des séquences des ARN parentaux donneur et accepteur, la recombinaison sera soit homologue (ARN parentaux de séquences identiques), soit hétérologue (ARN parentaux de séquences différentes). Les génomes en résultant doivent, à leur tour, former un dimère fonctionnel et les protéines structurales doivent pouvoir interagir entre elles afin de permettre l'assemblage des particules chimériques. Le provirus mosaïque obtenu est intégré au génome de la cellule hôte et donne naissance, selon un cycle classique, à une progénie virale recombinante, qui sera ensuite libérée pour infecter de nouvelles cellules (FIGURE 7, p.40). La recombinaison est efficace uniquement lorsque cette progénie virale recombinante s'avère fonctionnelle. Le rôle des interférences possibles entre les protéines

structurales lors de l'assemblage et de la maturation des particules virales complètes, ainsi que celui des facteurs de restriction cellulaires sur la sélection des variants recombinants, restent largement méconnus à ce jour <sup>[102-105]</sup>.

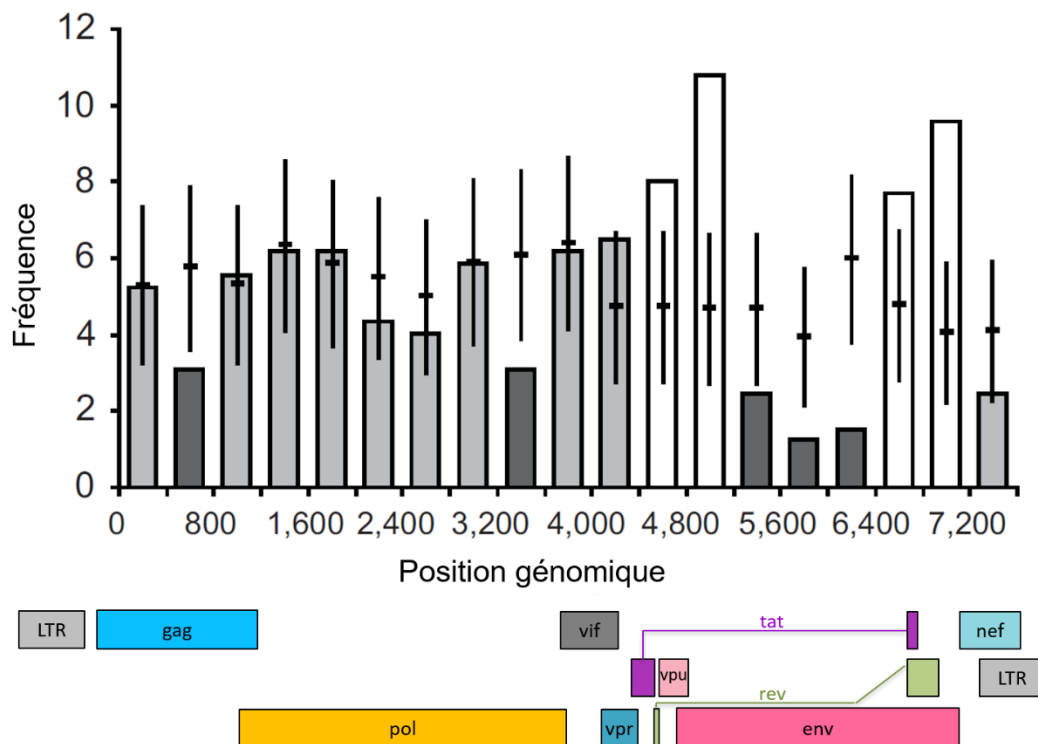
### c) Notion de points chauds de recombinaison

Le saut de la TI d'un ARN à l'autre pourrait s'effectuer à certains points préférentiels du génome, nommés "points chauds" ou "hot spot" de recombinaison.

L'impact de l'homologie de séquence a été très largement étudié par diverses équipes, pour identifier ces points chauds <sup>[106-108]</sup>. Les observations mettent en évidence une corrélation entre l'identité de séquence et la fréquence de recombinaison, bien que des structures secondaires de l'ARN semblent aussi jouer un rôle, dans des régions où l'homologie n'est pas importante. Dans ces études, les zones de recombinaison observées *in vitro* semblent coïncider avec les régions conservées du gène *env* ou avec la présence de régions répétées <sup>[106-108]</sup>. De plus, la fréquence de recombinaison *in vitro* s'avère plus importante, quand la recombinaison a lieu entre deux virus de même sous-type, que lorsqu'elle met en jeu deux virus de sous-types différents <sup>[106, 107]</sup>. Dans une autre étude, cette corrélation entre homologie et fréquence n'est pas retrouvée, mais les points chauds de recombinaison décrits sont situés de part et d'autre du gène de l'enveloppe, laissant suspecter un avantage sélectif probable à transmettre une enveloppe complète d'un virus à l'autre <sup>[108]</sup>.

Archer et *al.* ont analysé un ensemble de génomes recombinants de VIH générés *in vitro* afin de construire un modèle probabiliste de prédiction de la fréquence de points chauds de recombinaison en fonction des régions du génome <sup>[109]</sup>. Ce modèle est basé sur la similarité locale des séquences virales parentales. Ils ont analysé les séquences de nombreuses formes recombinantes circulantes, afin d'établir la distribution théorique des points de cassure le long du génome du VIH-1 et de la comparer à celle obtenue avec leur modèle. Les résultats obtenus indiquent que cette distribution n'est pas aléatoire, faisant apparaître des points chauds de recombinaison au début et à la fin du gène *env* et dans les régions de la TI, de l'IN et des gènes accessoires. Même si dans certaines régions (*gag*, *pol* et gènes accessoires), l'homologie de séquence permet de prédire la répartition des points de cassure, une surreprésentation de points de cassure aux extrémités du gène *env* est retrouvée, associée à une sous-représentation à l'intérieur même du gène (FIGURE 9).





**FIGURE 9** : Répartition globale des points de cassure le long du génome de VIH-1.

Comparaison de la fréquence de points de cassure prédits par le modèle (traits horizontaux) à celle des points de cassure observés naturellement parmi les VIH-1 recombinants extraits de la banque de données de séquences. Les barres blanches indiquent les régions où le nombre de points de cassure naturels est nettement supérieur à celui prédit par le modèle. Les barres gris-claires indiquent les régions où le nombre de points de cassure naturels est superposable à celui prédit par le modèle. Les barres gris-foncées indiquent les régions où le nombre de points de cassure naturels est nettement inférieur à celui prédit par le modèle. Les positions des gènes correspondent à un alignement sur la souche de référence HxB2 (d'après Archer et al. <sup>[109]</sup>).

Une étude plus récente a analysé la répartition des points de cassure dans les séquences publiées de formes recombinantes circulantes, à partir de la base de données Los Alamos <sup>[110]</sup> et de formes recombinantes uniques, et a permis de confirmer les résultats prédits par le modèle d'Archer et al. <sup>[111]</sup>. Cette étude a également montré que les régions avec des probabilités d'appariement élevées semblaient favoriser les événements de recombinaison. Enfin, les points chauds de recombinaison identifiés se situaient préférentiellement dans les régions chevauchantes du génome, partagées par deux gènes, et un point froid de recombinaison ressortait au sein du gène *env*, comme pour les autres études.

Une étude complémentaire visant plus particulièrement la recombinaison intra-gène *env*, l'enveloppe étant une région où les contraintes fonctionnelles sont très importantes, a confirmé l'existence d'une sélection négative des souches générées par des événements de recombinaison délétères <sup>[103]</sup>. Ces résultats ont confirmé les conclusions d'Archer et al. sur la fréquence des points de cassure plus importante aux extrémités d'*env* qu'en son sein <sup>[109]</sup>.

Ces éléments soulignent l'impact important des pressions de sélection *in vivo* sur le profil de recombinaison. En effet, la tendance de la plupart des régions de l'enveloppe (Env) à recombiner en une unité intégrale est probablement directement liée à l'impact d'une protéine Env fonctionnelle sur le *fitness* viral et à sa propension à être soumise à des niveaux élevés de sélection positive face à la pression immunitaire qui lui est imposée [108, 111-114]. La rareté des événements de recombinaison observés au sein même des gènes structuraux *gag*, *pol* et *env* (FIGURE 9) pose l'hypothèse que de telles formes recombinantes seraient soumises à une importante sélection négative, due à la génération d'une protéine dont l'intégrité structurale ou fonctionnelle ne serait pas satisfaisante [115].

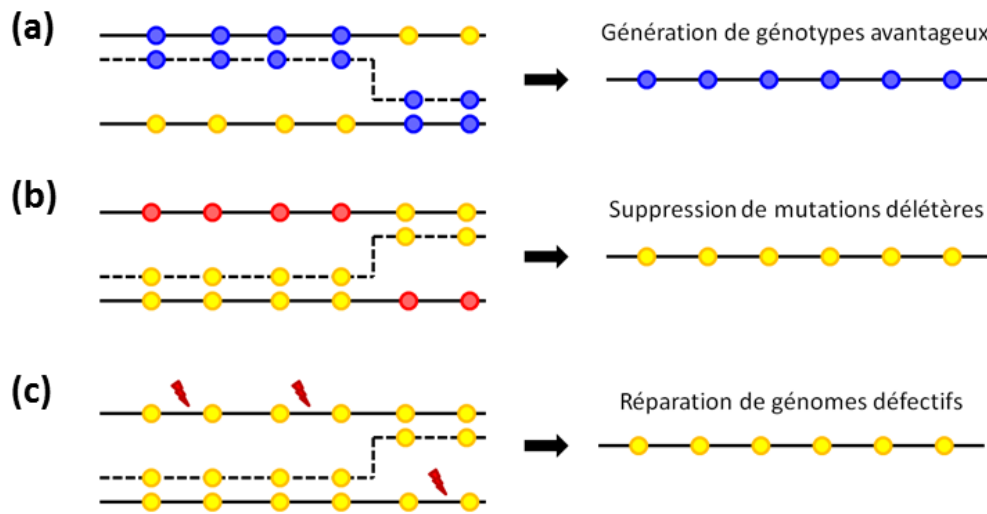
#### d) Conséquences de l'évolution par recombinaison

**Co-évolution génique.** Les gènes et les génomes évoluent généralement par la lente accumulation de mutations ponctuelles, ce qui nécessite souvent l'insertion de mutations compensatoires dans les sites liés. Cette co-évolution permet la préservation des interactions épistatiques [116, 117]. Comme la recombinaison peut impliquer des gènes codant pour des protéines déjà fonctionnelles, la probabilité de produire une descendance viable est plus élevée par rapport à l'insertion d'un nombre équivalent de mutations ponctuelles aléatoires [116]. Mais en introduisant simultanément plusieurs substitutions, une recombinaison a le potentiel de perturber sensiblement des réseaux d'interactions intra-génomiques ayant co-évolué, nuisant ainsi à la fonctionnalité des gènes impliqués [116, 117]. L'équilibre entre les avantages à prendre des raccourcis évolutifs et le risque d'obtenir des chimères non fonctionnelles détermine le rôle joué par la recombinaison dans l'évolution d'un gène ou d'un organisme donné [116].

Deux principaux facteurs ont une grande influence sur la fonctionnalité et l'évolution des protéines recombinantes : la position du point de cassure par rapport à celle des polymorphismes génétiques dans le gène et la position du point de cassure par rapport aux structures secondaires du génome et aux séquences protéiques conditionnant le repliement des protéines [118, 119]. Ainsi, les points de cassure situés à proximité des limites de ces régions auront généralement un impact moindre sur le repliement et donc sur les fonctions des protéines, que ceux qui sont situés en leur sein [117, 120, 121].

**Avantage sélectif.** Une fois les réseaux de co-évolution stabilisés, la recombinaison constitue un phénomène d'évolution génétique qui peut présenter plusieurs avantages pour le virus. Elle permet d'augmenter considérablement la complexité génétique d'une population virale et de

générer un nombre important de nouveaux phénotypes, qui mettraient plus de temps ou auraient moins de chance d'émerger par mutation <sup>[122]</sup> (FIGURE 10).



**FIGURE 10** : Conséquences possibles de l'évolution virale par recombinaison.

En fonction des génotypes donneur et accepteur et de la position du saut de matrice, la recombinaison peut avoir plusieurs effets sur le génome. Les cercles jaunes indiquent les loci de type sauvage. Grâce à la recombinaison, des mutations avantageuses (cercles bleus), présentes chez l'un et/ou l'autre des deux parents, peuvent être additionnées (a) et des mutations délétères (cercles rouges) supprimées (b). La recombinaison permet également de réparer des génomes défectifs chez des virus qui ne seraient plus viables en raison de l'altération des brins d'ADN (c). Les dommages génétiques, tels que la rupture des brins ou des modifications de base oxydantes, sont représentés par des symboles de foudre rouge (d'après Simon-Lorière et al. <sup>[89]</sup>).

Ainsi, la recombinaison peut créer des combinaisons avantageuses de mutations, qui augmentent le taux d'évolution adaptative par rapport à la mutation seule, ou peut dissocier des mutations avantageuses et délétères, permettant aux premières de se propager. Elle peut également supprimer des mutations délétères et rétablir le génotype sauvage, ce qui peut conduire à un avantage sélectif pour la recombinaison si des mutations délétères se produisent assez fréquemment et interagissent de façon synergique <sup>[123-126]</sup>. Enfin, la recombinaison peut régénérer un génome fonctionnel à partir de brins d'ADN parentaux défectifs endommagés, par exemple par une rupture de leur continuité ou par des modifications de bases oxydantes <sup>[127, 128]</sup>

Le large spectre de variants génétiques générés par recombinaison permet aux virus de s'adapter à l'environnement de l'hôte, d'augmenter leur capacité adaptative/répliquative (*fitness*), voire de surpasser les virus parentaux au fil de l'infection. Une étude a, en effet, montré que la recombinaison contribue à façonner la diversité intra-hôte, notamment dans l'enveloppe de VIH-1 <sup>[129]</sup>. La progénie virale recombinante présentait de nouvelles

combinaisons de traits phénotypiques (usage des co-récepteurs, tropisme en fonction des tissus infectés, sensibilité aux inhibiteurs d'entrée et modulation des récepteurs aux anticorps neutralisants), pouvant s'avérer potentiellement avantageuses, par rapport aux souches parentales non recombinantes dont elle était issue <sup>[129]</sup>. D'autres équipes ont mis en évidence que la recombinaison contribue de manière importante à la diversité des quasi-espèces en sélectionnant, d'une part, les portions de génomes présentant un avantage sélectif tout en conservant, d'autre part, la diversité génétique des régions adjacentes.

De plus, la population recombinante générée peut présenter une meilleure adaptation à la pression de sélection immunitaire de l'hôte, entraînant un éventuel échappement à sa réponse immunitaire. Deux équipes ont ainsi mis en évidence un échappement à la réponse T CD8+, médiée par des événements de recombinaison au sein des gènes *gag*, *env* et surtout *nef*, avec pour conséquence une augmentation de la réplication virale <sup>[130, 131]</sup>. Ainsi, la pression immunitaire cellulaire de l'hôte semble jouer un rôle dans la sélection de nouvelles formes recombinantes de VIH.

Par ailleurs, la recombinaison génétique peut participer à l'adaptation du VIH face à la pression de sélection liée aux thérapeutiques antirétrovirales (ARV), en associant des mutations de résistance aux ARV exprimées par des virus divergents, générant de nouvelles combinaisons de mutations et donc des virus à résistance accrue aux ARV <sup>[132, 133]</sup>.

Enfin, la recombinaison pouvait avoir des conséquences sur la progression de la maladie <sup>[134, 135]</sup>. Une étude a notamment mis en évidence l'émergence d'une progénie virale recombinante avec un meilleur *fitness* que la progénie virale parentale, chez un patient hautement immunodéprimé, accélérant sa progression vers le SIDA <sup>[135]</sup>.

### III- Classification des VIH

Les franchissements de barrière d'espèce et la diversification virale importante ont conduit à une classification des VIH en 2 types (VIH-1 et VIH-2), des VIH-1 en quatre groupes (M, N, O et P) et des VIH-2 en 9 groupes (A-I), définie en fonction de l'homologie génétique entre les virus. La diversité liée au fort taux de mutation, associée à une importante production virale et à une évolution dynamique au cours du temps, a conduit à la division des VIH-1/M en

sous-types et celle des VIH-1/O en sous-groupes. Enfin, les évènements de recombinaison décrits au sein des VIH-1/M principalement, mais aussi au sein des VIH-1/O et des VIH-2 ont permis l'identification et la description de nombreuses formes recombinantes.

### III-1. Les VIH-1

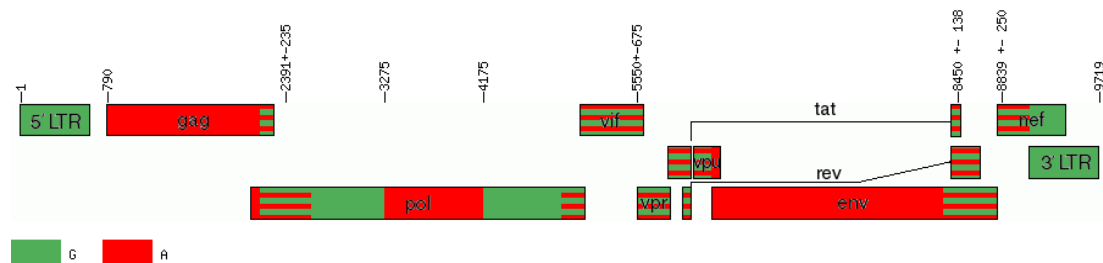
#### III.1.1. Le groupe M (Major)

Ce groupe, pandémique, est responsable de près de 37 millions d'infections dans le monde [Chiffres ONUSIDA 2018] et représente 90% des isolats de VIH-1 caractérisés à ce jour. Au sein de ce groupe, 9 sous-types « purs », non-recombinants, ont été décrits (A-D, F-H, J, K). Ils divergent entre eux en moyenne de 15% et 25% dans les gènes *gag* et *env*, respectivement [136-138]. Chacun d'entre eux est le reflet d'un effet fondateur, lié à l'exportation d'une souche unique dans une population initialement naïve, depuis l'épicentre constitué par la République Démocratique du Congo, ex-Zaïre. Ces effets fondateurs indépendants génèrent des sauts de diversité et entraînent une évolution anormalement discontinue, avec la diffusion de variants aux caractéristiques virologiques communes à des emplacements géographiques dispersés [137]. Ils sont ainsi responsables d'une grande diversité génétique, liée à l'artefact épidémiologique qu'ils constituent.

La globalisation de l'épidémie et le brassage des populations infectées conduisent à une complexification moléculaire de ce groupe, avec la mise en évidence de nombreuses formes recombinantes inter-sous-types circulantes (CRFs) ou uniques (URFs) [137]. Les CRFs comportent des isolats liés génétiquement, présentent la même structure mosaïque et sont décrites chez au moins trois individus non liés épidémiologiquement. A l'inverse, les URFs correspondent à des formes recombinantes identifiées chez des individus liés épidémiologiquement ou pour lesquelles une diffusion plus large dans les populations n'est pas observée. Ces formes constituent des lignages émergents et répondent à une nomenclature bien définie. Les CRFs possèdent toutes un numéro, qui correspond à l'ordre dans lequel elles ont été découvertes et publiées, mais qui ne reflète ni de l'ancienneté de leur émergence, ni de leur histoire évolutive [139]. A la suite du numéro, des lettres indiquent les sous-types présents dans la CRF, dans le cas où ils sont seulement deux (ex : CRF02\_AG, FIGURE 11 ou CRF03\_AB, FIGURE 12). Le terme "cpx", pour complexe, remplace les lettres lorsque trois ou plusieurs sous-types, voire même plusieurs

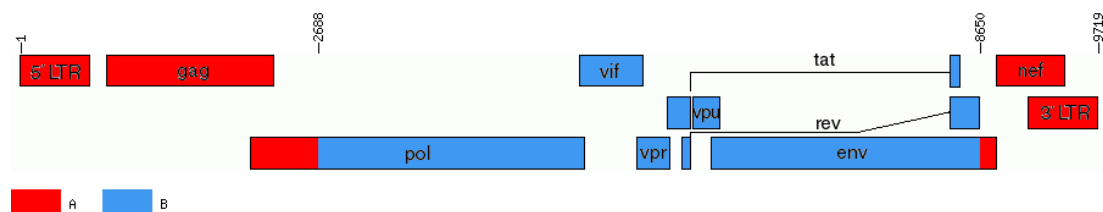
CRFs (les CRFs pouvant recombiner entre elles), sont présents dans la forme mosaïque (ex : CRF56\_cpx, FIGURE 13).

A ce jour, près de 100 CRFs et plusieurs centaines d'URFs ont été rapportées <sup>[110, 140]</sup> et leur nombre ne cesse d'augmenter.



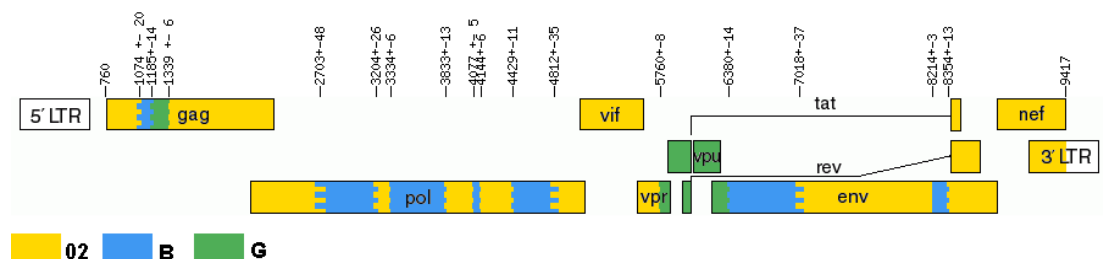
**FIGURE 11 :** Organisation génomique de la forme recombinante CRF02\_AG.

Seules deux souches parentales sont engagées et correspondent à des VIH-1/M de sous-types G (en vert) et A (en rouge). Les portions zébrées représentent des régions conservées entre les sous-types, pouvant correspondre aux deux sous-types adjacents. La localisation des points de cassure est indiquée par rapport à la souche de référence VIH-1/M HxB2 (numéro d'accèsion GenBank K03455). De nombreux points de cassure sont observés le long du génome, indiquant un niveau de complexité intermédiaire dans le profil de recombinaison. <sup>[141-143]</sup>.



**FIGURE 12 :** Organisation génomique de la forme recombinante CRF03\_AB.

Seules deux souches parentales sont engagées dans la forme recombinante CRF03\_AB et correspondent à des VIH-1/M de sous-types A (en rouge) et B (en bleu). La localisation des points de cassure est indiquée par rapport à la souche de référence VIH-1/M HxB2 (numéro d'accèsion GenBank K03455). Deux points de cassure sont situés dans le gène *pol* et dans le gène *env*. Cette souche représente le profil de recombinaison le plus simple <sup>[144-147]</sup>.



**FIGURE 13 :** Organisation génomique de la forme recombinante CRF56\_cpx.

Trois souches parentales sont engagées et correspondent à des VIH-1/M de sous-types CRF02\_AG (en jaune), B (en bleu) et G (en vert). Les portions zébrées représentent des régions conservées entre les sous-types, pouvant correspondre aux deux sous-types adjacents. La localisation des points de cassure est indiquée par rapport à la souche de référence VIH-1/M HxB2 (numéro d'accèsion GenBank K03455). De nombreux points de cassure sont observés le long du génome. La souche CRF56\_cpx montre une organisation génomique complexe, tant par le nombre de points de cassure que par le nombre de sous-types présents décrits (d'après Leoz et al. <sup>[148]</sup>).

A l'échelle mondiale, au moins 90% des souches de groupe M en circulation sont de sous-types non B. Les sous-types prédominants sont les sous-types C (Afrique du Sud, Inde, Chine), puis A (Afrique de l'Est, Europe de l'Est), puis B (Europe de l'Ouest, Amérique du Nord, Australie). En Afrique centrale, on observe une remarquable diversité génétique. En effet, tous les sous-types purs du groupe M et la plupart des nombreuses formes recombinantes complexes sont présents. Ceci reflète l'ancienneté de l'épidémie dans cette région du monde et souligne le rôle crucial de cette région dans la diversification de l'épidémie au cours du temps.

### III.1.2. Les groupes non-M (N, O et P)

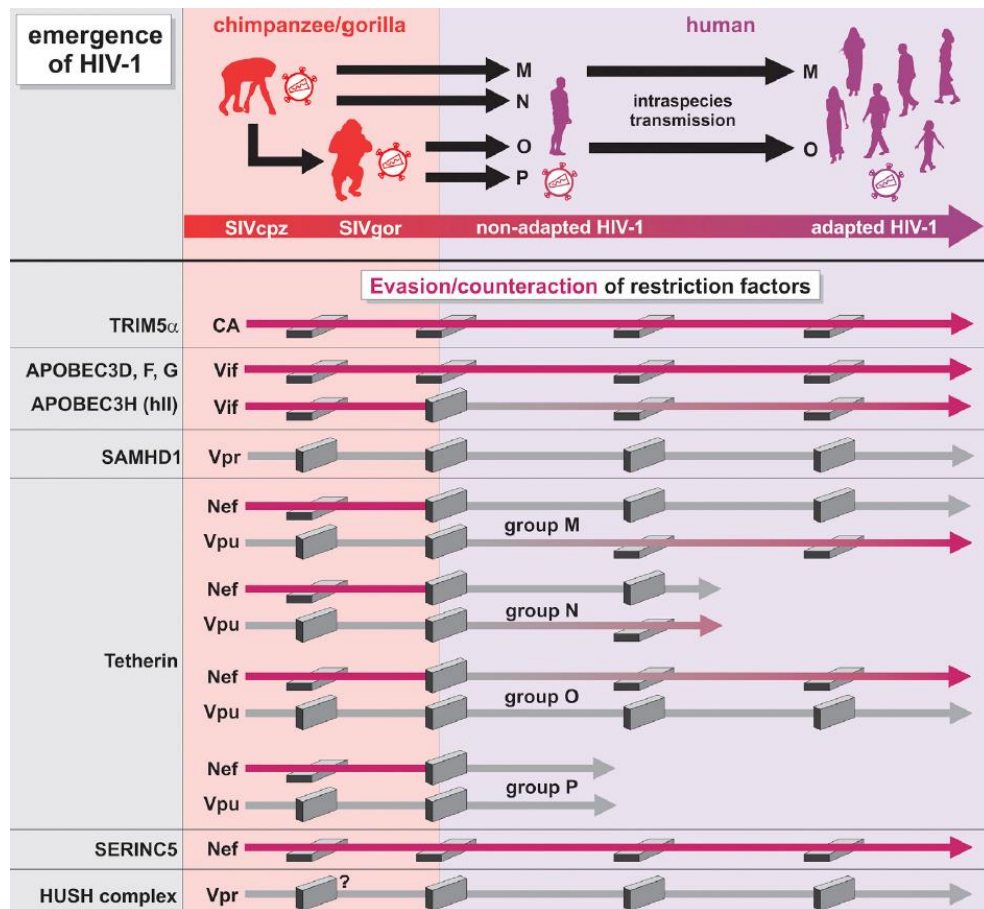
Ces groupes sont endémiques au Cameroun. Cette région correspond au berceau des transmissions inter-espèces du singe à l'Homme, à l'origine de ces différents groupes.

#### a) Le groupe O

Le groupe O (Outlier) est le plus prévalent des virus non-M, infectant 0,6 à 1% des personnes vivant avec le VIH (PvVIH) au Cameroun <sup>[149-151]</sup>. L'épidémie de VIH-1/O est caractérisée par une prévalence qui diminue significativement au cours du temps, avec un nombre de découvertes considéré comme stable depuis une vingtaine d'années, alors que la prévalence du VIH-1/M, pandémique, est en permanente augmentation et progresse de manière exponentielle <sup>[152, 153]</sup>. La faible diffusion du variants de VIH-1/O en dehors du Cameroun est également à soulever, avec seulement quelques cas sporadiques décrits à divers endroits du globe : en Afrique de l'Ouest et de l'Est, en Europe, aux Etats-Unis <sup>[5, 154-164]</sup>. La France, où la file active d'individus infectés par un VIH-1/O atteint 143 patients, apparaît comme le deuxième pays le plus touché, probablement du fait de liens historiquement étroits avec le Cameroun <sup>[165, 166]</sup>. D'ailleurs, chez presque tous les patients ou partenaires de patients VIH-1/O positifs, un lien initial avec le Cameroun est retrouvé.

Ces faibles prévalence et diffusion pourraient être expliquées, en partie, par la différence d'adaptation de ces variants aux facteurs de restriction cellulaires. C'est ce qui a notamment été montré lors de l'étude de l'activité anti-Tétherine. Elle est initialement portée par la protéine Nef des VIScpz et des VISgor et a été conservée chez les VIH-1/O <sup>[167]</sup>, alors que

pour les VIH-1/M, elle a été transférée à la protéine Vpu, après transmission du virus du chimpanzé à l'Homme (FIGURE 14).



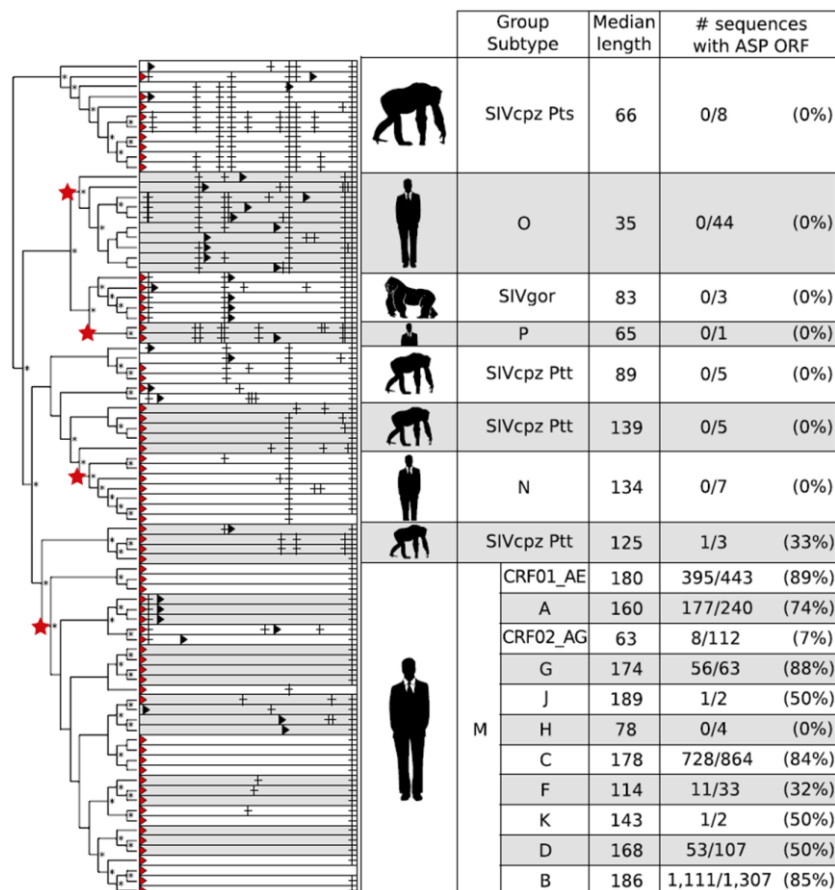
**FIGURE 14 :** Emergence des variants du VIH-1 : facteurs de restriction et stratégies de contournement. (d'après Sauter *et al* <sup>[168]</sup>)

La partie supérieure illustre les événements de transmission inter-espèces du singe à l'Homme, à l'origine des quatre groupes de VIH-1 (M, N, O et P). Les VIH-1/M et /O ont efficacement diffusé au sein de la population humaine, à la différence des VIH-1/N et /P, identifiés chez seulement quelques individus. La partie inférieure montre l'adaptation des protéines accessoires virales (Vif, Nef, Vpu et Vpr) aux facteurs de restriction cellulaires (TRIM5 $\alpha$ , APOBEC3, SAMHD1, Téthérine, SERINC5 et HUSH). Les murs intacts avec des flèches grises symbolisent une barrière d'espèces limitant l'adaptation à l'hôte et donc le contournement du facteur de restriction par la protéine concernée. L'échappement à la neutralisation par les facteurs de restriction est symbolisé par des murs tombés, et des flèches roses.

La Téthérine humaine contient une délétion dans son domaine cytoplasmique, la rendant résistante à la protéine accessoire Nef des VIScpz et des VISgor. Les protéines Nef des VIH-1/O ont conservé leur activité anti-Téthérine, en ciblant une région adjacente à la délétion protectrice de la Téthérine. Le gain de cette activité anti-Téthérine humaine de Nef, en la bloquant dans des compartiments intracellulaire périnucléaires, a probablement facilité la multiplication des VIH-1/O chez l'Homme <sup>[22]</sup>.



Un autre élément opposant le VIH-1/O au VIH-1/M est l'absence de gène *asp* au sein de son génome. En effet, une étude de 2016 a montré, par une approche phylogénétique, une émergence simultanée du cadre de lecture ouvert correspondant au gène *asp* et du VIH-1/M, ainsi qu'une corrélation entre la présence du gène *asp* et l'expansion de la pandémie à VIH-1/M <sup>[11]</sup> (FIGURE 15).



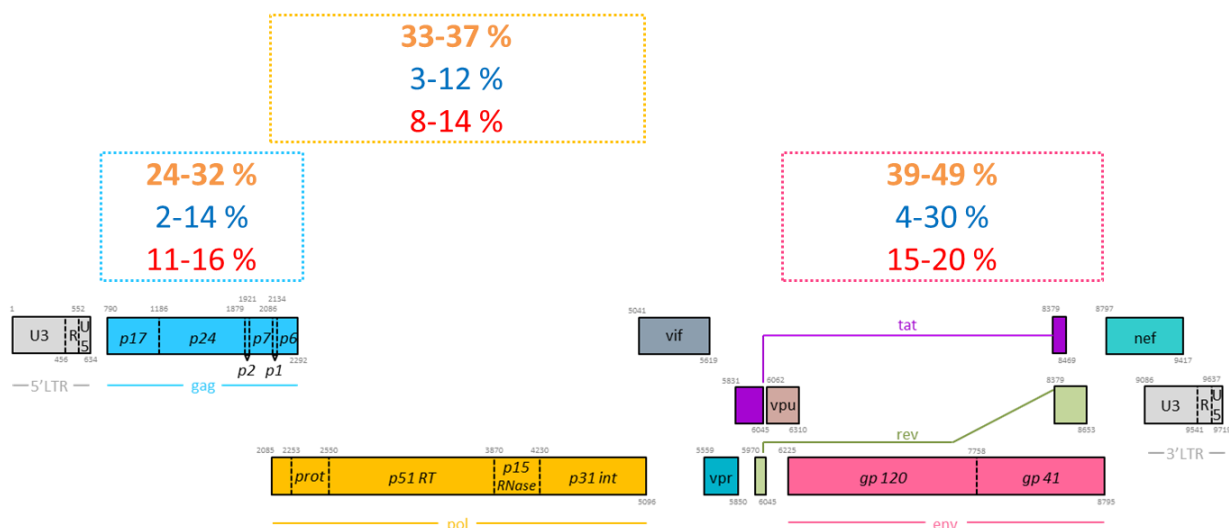
**FIGURE 15 :** Mise en évidence de l'émergence récente du gène *asp*, par une approche phylogénétique. (d'après Cassan *et al* <sup>[11]</sup>)

Cette phylogénie contient des séquences de référence du gène *env* des quatre groupes de VIH-1, des VIScpz, VISgor, et des sous-types et CRFs de VIH-1/M. Les quatre événements indépendants de transmission inter-espèces du singe à l'Homme sont indiquées par des étoiles rouges. Pour chacune des séquences, la distribution des codons d'initiation (triangle rouge ou triangle noir pour un démarrage alternatif) et des codons stop (croix noire) dans la région *asp* sont montrées. Pour chaque groupe, sous-type et CRF, le tableau fournit la longueur moyenne du cadre de lecture *asp*, et la fraction de séquences contenant un cadre de lecture d'une longueur supérieure à 150 codons, correspondant à la présence du gène *asp*.

Des différences de capacité répliquative, ou *fitness* répliquatif, ont également été décrites entre les différents variants de VIH. En effet, les résultats d'essais de compétition *in vitro* ont suggéré des niveaux de répllication variables selon les variants, avec un *fitness* répliquatif du VIH-1/O au moins 100 fois inférieur à celui du VIH-2, lui-même inférieur à celui du VIH-1/M (VIH-

1/M > VIH-2 > VIH-1/O) [169]. Bien que cette "pyramide de *fitness*" corresponde à des différences actuelles de prévalences de ces variants, les VIH-1/O présentent une forte réplication virale *in vivo*, même si celle avant traitement est inférieure de 1 Log<sub>10</sub> cp/mL en moyenne à celle des VIH-1/M [170, 171]. Cette réplication reste cependant bien plus intense que celle observée dans l'infection à VIH-2. Ce groupe de variants est également pathogène pour l'homme, avec une évolution vers le SIDA sans traitement ARV, intermédiaire entre le VIH-1/M et le VIH-2.

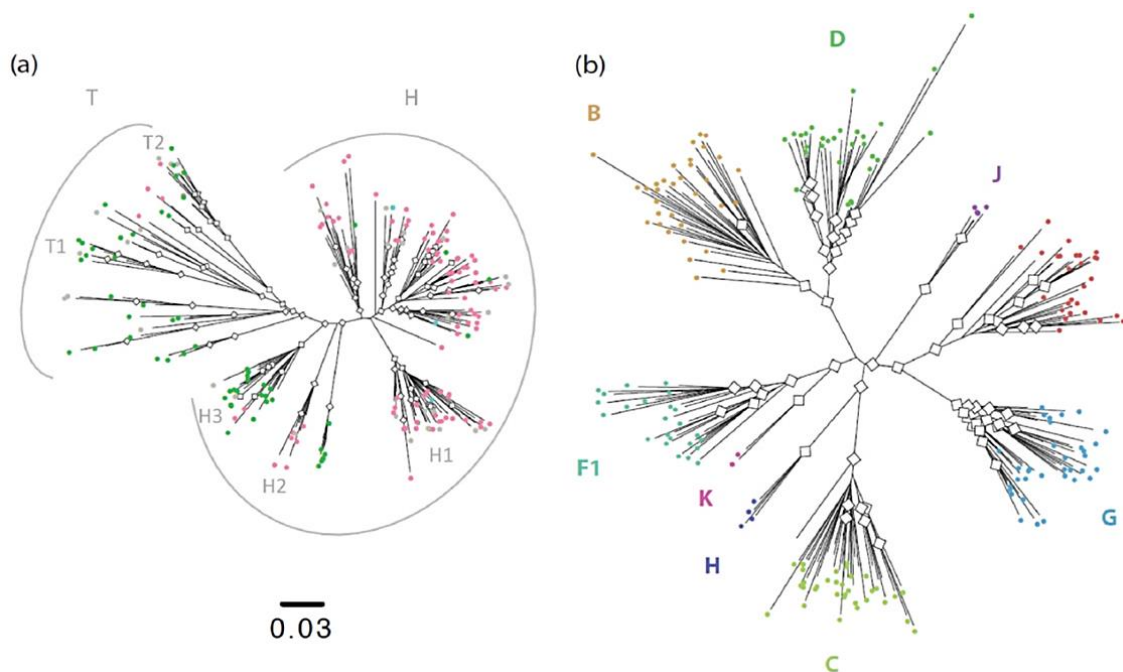
Alors que la diversité génétique intra-groupe O semble globalement superposable à la diversité génétique intra-groupe M (2-14 %, 3-12 % et 4-30 % de divergence des gènes *gag*, *pol* et *env* respectivement, au sein du groupe O), le VIH-1/O se caractérise par une grande divergence génétique par rapport au VIH-1/M (24-32%, 33-37% et 39-49% de différence entre les gènes *gag*, *pol* et *env* respectivement) [137, 172-175] (FIGURE 16).



**FIGURE 16** : Divergence génétique entre les VIH-1/M et VIH-1/O, en fonction des régions du génome.

Les pourcentages de divergence sont positionnés au niveau des gènes concernés. La variabilité génétique entre les groupes M et O figure en orange et en gras. Les pourcentages de variabilités au sein des groupes O et M figurent en bleu et rouge, respectivement.

Sur le plan phylogénétique, les séquences de VIH-1/O se répartissent sous la forme d'un arbre en forme de comète, ce qui diffère fortement de celui du groupe M, en forme de double étoile (FIGURE 17). Cette topologie est due à la diversification différente au sein du VIH-1/O, qui représente un continuum de diversité. Ceci reflète une diversification locale dans une zone géographique et dans le temps, à la différence des sous-types de VIH-1/M, résultats de multiples effets fondateurs.



**FIGURE 17 :** Comparaison des arbres phylogénétiques de VIH-1/O et de VIH-1/M

Les arbres phylogénétiques ont été construits à partir de 190 séquences de VIH-1/O (a) et de 190 séquences de VIH-1/M (b) et montrent les sous-groupes H et T organisés en comète (a) et les sous-types A, B, C, D, F, G, H, J et K organisés en double-étoile (b) (d'après Leoz et al. [73]).

Cette diversification, en deux phases successives de croissance exponentielle au cours du 20<sup>ème</sup> siècle, a conduit à deux populations de VIH-1/O : une population majoritaire (sous-groupe H), regroupant les souches situées dans la tête de la comète et une population minoritaire (sous-groupe T), génétiquement divergente et regroupant les souches situées dans la queue de la comète [73] (FIGURE 17).

#### b) Les groupes N et P

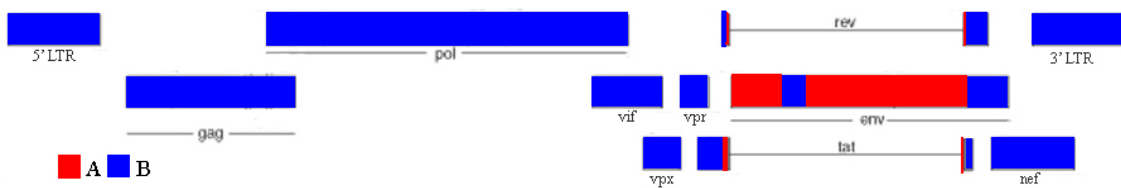
Concernant les groupes N et P, le peu de souches (19 et deux respectivement), provenant tous de patients originaires du Cameroun sauf un et présentant des liens épidémiologiques ou phylogénétiques pour certains, limite l'étude de leur diversification [6, 7, 69, 176-185].

### III-2. Les VIH-2

Contrairement au VIH-1, les VIH-2 sont principalement localisés en Afrique de l'Ouest. Ils sont moins impliqués dans la pandémie de SIDA et ont des histoires naturelle et thérapeutique très différentes du VIH-1. Les VIH-2 sont divisés en 9 groupes phylogénétiques

(A-I), correspondant à autant de transmissions du singe à l'Homme <sup>[186]</sup>. Seulement deux de ces groupes (VIH-2/A et VIH-2/B) sont épidémiques.

Des formes recombinantes ont été également décrites au sein de ce type. Une CRF (CRF01\_AB) impliquant des fragments génomiques dérivant des groupes A et B a été identifiée en 1994 et 2010 <sup>[187, 188]</sup> (FIGURE 18) et une URF, impliquant les mêmes deux groupes, a été découverte chez une patiente camerounaise en 2003 <sup>[189]</sup>.



**FIGURE 18** : Organisation génomique de la forme recombinante HIV2\_CRF01\_AB.

Deux souches parentales sont associées dans cette forme recombinante et correspondent à des VIH-2 de groupes A (en rouge) et B (en bleu). Le profil de recombinaison présente quatre points de cassure, situés pour trois d'entre eux dans le gène *env* <sup>[188]</sup>.

#### IV- Recombinaison inter-groupes entre les VIH-1/M et VIH-1/O

La recombinaison intra-groupe M des VIH-1 occupe une place importante dans l'épidémie, estimée à 18% des infections à VIH dans le monde <sup>[190]</sup>. Cette importante proportion de formes recombinantes circulantes prouve donc que la recombinaison génétique contribue à la grande diversité génétique des VIH.

Ce phénomène n'est cependant pas toujours possible. En effet, malgré la présence non négligeable de co-infections VIH-1+2 en Afrique de l'Ouest, aucun recombinant inter-types n'a été détecté à ce jour. Ceci est vraisemblablement dû à la très forte divergence génétique entre les deux types (40 à 70% en fonction des gènes), combinée à la moindre répliquativité du VIH-2, ce qui conduit à une probabilité de recombinaison jugée très faible <sup>[191]</sup>.

Aussi, la recombinaison entre les groupes M et O a été jugée impossible et/ou négligeable pendant longtemps, compte tenu des divergences génétiques entre ces virus et de la faible prévalence des VIH-1/O. A l'instar des VIH de types 1 et 2, elle a donc été très peu recherchée. Pourtant, la découverte, en 1998, du VIH-1/N a montré que la recombinaison naturelle entre

deux virus très divergents, mais issus de la même lignée, était possible. En effet, le génome de ce groupe de variants est mosaïque VIScpz/VIH-1/M et occupe donc une position intermédiaire entre les VIScpz, eux-mêmes recombinants, et les VIH-1/M <sup>[6]</sup>.

Finalement, deux formes recombinantes VIH-1/MO ont été décrites pour la première fois, en 1999 <sup>[192-194]</sup>.

### III-1. Pré-requis : la double infection VIH-1/M+O

La co-circulation endémique des VIH-1/M, à prévalence croissante, et des VIH-1/O, en Afrique Centrale, et au Cameroun en particulier, rend possible la co-infection d'un même individu par ces isolats génétiquement différents. Ce postulat a été confirmé lorsque les premières doubles infections par des VIH-1 de groupes M et O (VIH-1/M+O) ont été décrites, d'abord au Bénin en 1996 <sup>[156]</sup>, puis au Cameroun en 1998 <sup>[195]</sup> (TABLEAU 1).

**TABLEAU 1** : Circonstances de découverte des 20 doubles infections VIH-1/M+O

<b>Etude, année de publication</b>	<b>Pays Période d'étude</b>	<b>Nombre de doubles infections M+O décrites</b>
Heyndrickx, 1996 <sup>[156]</sup>	Bénin 1993-1994	1
Takehisa, 1998 <sup>[195]</sup>	Cameroun 1994-1995	1
Vergne, 2003 <sup>[153]</sup>	Cameroun 1996-2001	4
Yamaguchi, 2004 <sup>[193]</sup>	Cameroun 2004	2
De Oliveira, 2017 <sup>[150]</sup>	Cameroun 2006-2009	4
Ngoupo <sup>[185]</sup>	Cameroun 2013-2015	4
Brand, 2004 <sup>[196]</sup>	France 2004	2
Plantier, 2004 <sup>[197]</sup>	France 2003	1
De Oliveira, 2018 <sup>[198]</sup>	France 2004-2016	1

Entre 1993 et 1994, Heyndrickx et *al.* ont collecté, au Bénin, 142 sera de patients connus séropositifs pour le VIH-1, parmi lesquels 20 présentaient une réactivité VIH-1/O en sérotypage, technique recherchant la présence d'anticorps dirigés contre les glycoprotéines d'enveloppe et représentatifs de la diversité des variants humains et simiens. Des PCRs spécifiques de groupes M et O ont alors été réalisées, suivies d'un séquençage et d'une analyse phylogénétique. Le premier cas de double infection VIH-1/M+O a été découvert à cette occasion, chez une patiente née en Côte d'Ivoire, d'origine béninoise <sup>[156]</sup>. Cette étude ne stipulait pas si une forme recombinante VIH-1/MO avait été recherchée.

Lors d'une seconde étude séro-moléculaire (n=211 patients), réalisée au Cameroun, entre 1994 et 1995, des PCRs spécifiques de groupes M et O ont été réalisées sur les ADN intracellulaires extraits de cellules mononuclées du sang périphérique (CMSPs) humaines primaires. Pour une patiente, ces PCRs spécifiques de groupes M et O se sont révélées positives. L'analyse phylogénétique a révélé une triple infection associant des souches VIH-1/M sous-types A et D et une souche VIH-1/O. A ce stade, aucun recombinant VIH-1/MO n'avait encore été décrit <sup>[195]</sup>.

D'autres doubles infections VIH-1/M+O ont depuis été décrites au Cameroun : quatre entre 1996 et 2001 <sup>[153]</sup>, deux en 2004 <sup>[193]</sup>, quatre entre 2006 et 2009 <sup>[150]</sup>, et enfin quatre plus récemment, entre 2013 et 2015 <sup>[185]</sup>.

Le VIH-1/O est retrouvé principalement au Cameroun, mais du fait de liens historiquement étroits avec ce pays, il circule également en Europe, notamment en France. Les deux premiers cas de double infection VIH-1/M+O, hors d'Afrique Centrale, ont ainsi été décrits en France, par Brand et *al.* en 2004 <sup>[196]</sup>. Ces doubles infections M+O ont d'abord été identifiées par sérotypage, dans le cadre de la déclaration obligatoire de l'infection à VIH, puis confirmées par caractérisation moléculaire. La même année, notre équipe décrivait un cas unique de surinfection VIH-1/M, chez une patiente déjà infectée par un VIH-1/O <sup>[197]</sup>. Depuis, un autre cas de double infection M+O a été décrit par notre équipe <sup>[198]</sup>, portant actuellement à 20 le nombre de cas de doubles infections VIH-1/M+O décrits.

Il semble important de noter que toutes les doubles infections mises en évidence hors d'Afrique Centrale ont été décrites chez des patients d'origine camerounaise ou présentant un lien épidémiologique avec le Cameroun. De plus, les outils et les stratégies disponibles au moment de la mise en évidence de ces doubles infections M+O, combiné à l'absence de

prélèvements séquentiels le plus souvent, n'ont pas toujours permis de rechercher les formes recombinantes VIH-1/MO éventuellement associées.

### III-2. Circonstances de découverte des formes recombinantes VIH-1/MO

Les deux premières formes recombinantes inter-groupes M et O ont été décrites chez deux patients camerounais, en 1999, et en 2004, un troisième cas était rapporté <sup>[192-194]</sup> (TABLEAU 2). Tous les trois présentaient une co-infection avec au moins une souche parentale VIH-1/O et/ou VIH-1/M. Ceci était en faveur de doubles infections répliquatives, ayant conduit à la genèse d'une forme recombinante *in vivo*.

**TABLEAU 2** : Chronologie de la découverte des 23 formes recombinantes VIH-1/MO

Etude, année de publication	Pays Période d'étude	Nombre de recombinants VIH-1/MO décrits	Code souche
Peeters, 1999 <sup>[192]</sup>	Cameroun 1999	1	97CA.MP645MO
Takehisa, 1999 <sup>[194]</sup>	Cameroun 1994-1996	1	95CM61.56
Yamaguchi, 2004 <sup>[193]</sup>	Cameroun 1999	1	DSC1320
De Oliveira, 2017 <sup>[150]</sup>	Cameroun 2006-2009	6	YBF205, YBF2011, YBF212, YBF274, YBF282, YBF298
Ngoupo, 2016 <sup>[185, 199]</sup>	Cameroun 2013-2015	3	REC003, REC024, REC107
De Oliveira, 2018 <sup>[198]</sup>	France 2004-2016	11	BCF204, RBF208, BCF174, BCF212, RBF221-1, RBF221-2, RBF222, RBF235, RBF237, RBF240, RBF243

Le premier recombinant MO a été isolé en 1999 chez une femme asymptomatique, originaire du Cameroun <sup>[192]</sup> (TABLEAU 2). La mise en évidence du recombinant a reposé sur une double réactivité M+O en sérotypage, confirmée par des PCRs spécifiques de groupes M et O

sur les ADN intracellulaires issus des CMSPs et par un séquençage de génome complet. Les résultats obtenus ont montré la présence d'une souche recombinante inter-groupes MO associée à une souche parentale VIH-1/M, sans souche parentale VIH-1/O. Après une étape de clonage biologique par dilution limite sur des cultures de CMSPs et la réalisation de PCR spécifiques de groupes sur le produit de co-culture, la souche parentale M n'était plus détectée alors que la souche recombinante était toujours présente. Ce résultat laissait suspecter une capacité répliquative de la forme recombinante MO supérieure à celle de la souche parentale M.

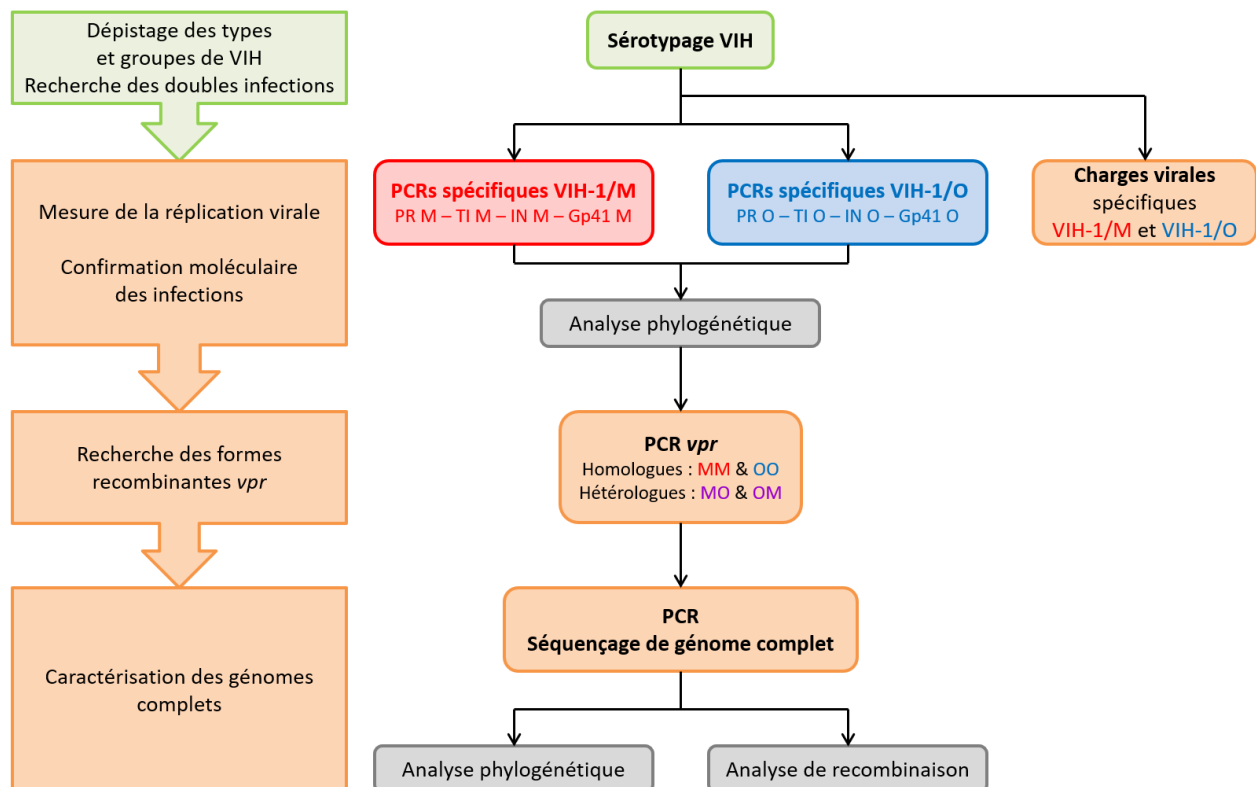
La même année, un autre recombinant a été rapporté à partir d'une triple infection VIH-1/M+M+O décrite en 1998, par Takehisa et *al.* <sup>[194]</sup> (TABLEAU 2). Trois échantillons séquentiels de CMSPs, datant de 1994, 1995 et 1996, ont été caractérisés moléculairement et les résultats obtenus ont permis de retracer l'évolution de la population virale au cours du temps. En effet, en 1994, les gènes *pol* et *env* des trois souches (VIH-1/M sous-types CRF02\_AG et D et VIH-1/O) étaient amplifiés alors qu'en 1995 et 1996, les gènes *env* des trois souches étaient toujours amplifiés, mais seul le VIH-1/O était retrouvé dans le gène *pol*. Des XL-PCR amplifiant des fragments d'environ 3000 pb contenant la fin du gène *pol*, les gènes accessoires et une partie du gène *env* ont été réalisées, et clonées. Cinq profils différents ont alors été mis en évidence, avec un point de cassure identique de motif OM dans la région chevauchante des gènes accessoires *vif* et *vpr*, et de multiples événements de recombinaison intra-groupe M dans *Vpu* et *Env*, régions du génome associées au VIH-1/M. Ces éléments suggèrent que la triple infection a conduit à l'émergence de virus recombinants et que de multiples événements de recombinaison intra-groupe M ont eu lieu à partir d'une forme recombinante VIH/MO initiale.

En 2004, un troisième recombinant VIH-1/MO a été caractérisé chez un patient camerounais, à partir d'un plasma de 1999 <sup>[193]</sup> (TABLEAU 2). Une double réactivité M+O en sérotypage, confirmée moléculairement, a permis de mettre en évidence une double infection par une souche recombinante VIH-1/MO associée à une souche parentale VIH-1/O.

Ces premiers cas décrits d'infection par des virus mosaïques MO ont donc montré que la recombinaison était possible malgré la divergence génétique entre ces deux groupes et que ces formes recombinantes pouvaient circuler au Cameroun, même si leur prévalence n'était pas définie. Par ailleurs, les profils moléculaires de ces trois premiers cas montraient que deux des trois formes recombinantes présentaient un point de cassure dans le gène *vpr*, conduisant à l'hypothèse que cette zone pouvait être un point chaud de recombinaison <sup>[192-194]</sup>. Afin



d'étudier l'étendue de ce phénomène et de vérifier cette hypothèse, deux études ont été réalisées au Cameroun par notre équipe. La première, effectuée entre 2006 et 2009, a consisté à rechercher de manière prospective les doubles infections VIH-1/M+O et les formes recombinantes avec un point de cassure dans *vpr* éventuellement associées, en mettant en place au Centre Pasteur du Cameroun (CPC) des outils sérologique et moléculaire nécessaires à leur dépistage et à leur caractérisation <sup>[150]</sup> (FIGURE 19, TABLEAU 2).



**FIGURE 19** : Algorithme séro-moléculaire de dépistage des recombinants *vpr*

Algorithme utilisé lors des études de dépistage et de caractérisation des recombinants *vpr*. Pour évaluer la répllicativité des souches, différentes mesures de charges virales sont effectuées : une non spécifique, amplifiant à la fois les VIH-1/M et les VIH-1/O et deux, spécifiques du VIH-1/M ou du VIH-1/O. La confirmation moléculaire des infections met en jeu des PCR pour lesquelles quatre régions sont ciblées : trois au sein du gène *pol* (PR, TI et IN) et une au sein du gène *env* (gp41). Ensuite, quatre PCR ciblant le gène *vpr*, et permettant la caractérisation des gènes accessoires, sont réalisées au moyen d'amorces homologues (visant à amplifier des gènes *vpr* de motifs MM et OO) et hétérologues (visant à amplifier des gènes *vpr* de motifs OM et MO). La caractérisation des génomes complets est enfin réalisée par séquençage associé à des analyses phylogénétique et de recombinaison.

Au cours de cette étude, la réactivité VIH-1/O de 6796 sera et de 15000 plasma a été testée par sérotypage. Parmi ces échantillons, 142 étaient réactifs VIH-1/O, dont 53 présentaient une double réactivité M+O. La confirmation moléculaire des résultats de sérotypage a été réalisée pour 39 des 53 prélèvements et une double infection VIH-1/M+O a

été mise en évidence pour 10 patients. Parmi eux, la recherche d'éventuels recombinauts *vpr* a été réalisée, grâce à une PCR amplifiant spécifiquement cette région. Des PCR nichées mettant en jeu des couples d'amorces conçus pour amplifier des gènes *vpr* homologues (*vpr* OO et *vpr* MM) et hétérologues (*vpr* OM et *vpr* MO) ont été réalisées. Après séquençage, cette méthodologie a permis la description de :

- quatre doubles infections sans recombinaut *vpr* associé,
- trois doubles infections associées à un recombinaut *vpr*,
- deux co-infections par un recombinaut *vpr* et sa souche parentale VIH-1/M,
- et d'un cas d'infection par un recombinaut hors *vpr* isolé, c'est-à-dire sans souche parentale associée.

Cependant, cette étude s'est heurtée à des limites méthodologiques. En effet, la sélection des échantillons à étudier se basait sur une double réactivité M+O en sérotypage. Or, la découverte dans le même temps de recombinauts isolés, sans forme parentale associée, rendait nécessaire l'exploration moléculaire d'échantillons présentant des mono-réactivités VIH-1/M ou VIH-1/O. De plus, l'un des trois premiers recombinauts présentait un point de cassure dans la région de l'IN, donc hors *vpr*. Cette étude avait recherché des recombinauts VIH-1/MO avec un point de cassure uniquement dans *vpr*, limitant ainsi la découverte d'autres profils de recombinaison. Les génomes complets des recombinauts *vpr* putatifs mis en évidence au cours des travaux précédents n'avaient pas non plus été caractérisés. Enfin, l'absence de données épidémiologiques disponibles, combinée aux limites rencontrées lors de cette étude, a soulevé la question de la prévalence de ces infections et de l'impact individuel et collectif de l'émergence et de la diffusion de telles formes.

Une seconde étude, portant sur la détermination de la fréquence des doubles infections M+O et des recombinauts MO circulant au Cameroun et la caractérisation des génomes complets des recombinauts, a donc été menée entre 2013 et 2015 <sup>[185, 199]</sup>. Ce travail avait pour objectif secondaire la mise en évidence d'une forme recombinante circulante (CRF\_MO), à l'instar de ce qui est observé avec les VIH-1/M. Il était basé sur le même algorithme séro-moléculaire « *vpr* » mis en place au CPC et utilisé dans le cadre de son activité diagnostique <sup>[199]</sup> (FIGURE 19, P.60). Au cours de cette seconde étude, les réactivités VIH-1/M et VIH-1/O de 6330 sera et de 25862 plasmas, soit 32192 échantillons au total, ont été testés en sérotypage. L'algorithme « *vpr* » a pu être appliqué pour 275 échantillons, parmi lesquels 199 et 47 présentaient une mono-réactivité VIH-1/M et VIH-1/O, respectivement, et 29 étaient

doublement réactifs M+O. Au final, la caractérisation moléculaire de ces prélèvements a permis la mise en évidence de :

- quatre doubles infections VIH-1/M+O sans recombinaison *vpr* associé,
- deux infections par un recombinaison *vpr* isolé,
- et d'une co-infection par un recombinaison *vpu* et sa souche parentale VIH-1/M.

Rapportée à l'effectif analysable de l'étude, la fréquence des doubles infections VIH-1/M+O s'élevait donc à 1,4% et celles des co-infections par une souche parentale VIH-1/M plus un recombinaison VIH-1/MO et infections par un recombinaison *vpr* isolé s'élevaient à 0,36% et 0,72%, respectivement. Aucune CRF\_MO n'a pu être décrite au cours de ce travail.

Au total, ces deux études menées au Cameroun ont mis en évidence la circulation de 9 nouvelles formes recombinantes MO isolées ou associées à des co-infections avec des souches parentales [150, 185, 199].

Des formes recombinantes ont également été détectées en dehors de cette région. En effet, en 2010, l'observation d'une charge virale indétectable chez une patiente non traitée et très immunodéprimée (6 CD4/mm<sup>3</sup>), d'origine camerounaise résidant en France, a nécessité la réalisation d'analyses séro-moléculaires complémentaires au laboratoire associé au Centre National de Référence (CNR) du VIH du CHU de Rouen [200] (TABLEAU 2). L'exploration de cette discordance immuno-virologique a révélé une mono-infection VIH-1/M en sérotypage, des résultats de charges virales spécifiques de groupes M et O respectivement indétectable et quantifiable, et des résultats d'amplification spécifique de groupes positifs ou négatifs selon les régions du génome. Le séquençage du génome complet a finalement permis la description du premier cas de recombinaison inter-groupes MO en dehors du Cameroun, mais cependant associée à cette région, du fait de l'origine de la patiente. Ce fut également la première description de transmission directe d'une forme recombinante MO avec un point de cassure dans *vpr*, sans souche parentale associée, démontrant ainsi la viabilité et le potentiel de diffusion de ces formes recombinantes MO.

Cette circulation de formes recombinantes au Cameroun et la transmission directe de virus recombinants VIH-1/MO isolés, pré-requis essentiel à l'émergence d'une URF et éventuellement d'une CRF, ont conduit à étudier l'étendue de ce phénomène en France, du fait des liens étroits avec le Cameroun [200]. Une recherche rétro-prospective, basée sur une stratégie séro-moléculaire proche de l'algorithme « *vpr* » (FIGURE 19), a donc été menée sur les

prélèvements des 143 patients diagnostiqués VIH-1/O positifs depuis 1992 et sur les prélèvements reçus au CNR VIH de Rouen dans le cadre d'une suspicion de variants divergents du VIH. Ces prélèvements présentaient des résultats discordants ou atypiques lors du dépistage ou du suivi de l'infection. Ces discordances étaient de plusieurs natures <sup>[198]</sup> :

- immuno-virologique (taux de CD4 bas et valeur de charge virale indétectable),
- sérologique (double réactivité M + O lors du sérotypage),
- séro-moléculaire (sérotypage et caractérisation moléculaire discordants),
- moléculaire (génotypage de résistance négatif malgré une charge virale détectable),
- génotypique (présence de séquences de variants de groupes différents).

Au final ce travail a permis d'identifier 11 nouvelles formes recombinantes VIH-1/MO, dont deux chez un même patient <sup>[198]</sup> (TABLEAU 2).

Au total, depuis la description des trois premiers cas en 1999 et 2004, ce sont 20 formes recombinantes supplémentaires, qui ont été identifiées au Cameroun et en France, et pour lesquelles le génome a pu être séquencé partiellement ou complètement (TABLEAU 2).

### III-3. Caractéristiques moléculaires des formes recombinantes VIH-1/MO

Actuellement, 23 formes recombinantes VIH-1/MO sont donc décrites chez 22 patients. Cet effectif, plus important que supposé initialement, a permis d'obtenir des données précises sur ces formes et de procéder à une comparaison de leurs caractéristiques moléculaires <sup>[201]</sup>.

#### III-3.1. Analyse phylogénétique et détermination du nombre d'URF\_MO

La détermination du nombre d'URFs consiste à définir le nombre de formes à prendre en compte. En effet, certaines ont été identifiées chez un même patient, chez deux patients liés épidémiologiquement, ou possiblement chez le même patient à des moments différents. Une analyse phylogénétique des différentes formes pour lesquelles des fragments de *pol* et *env* étaient disponibles (n=12) a donc été effectuée, afin de rechercher leurs liens phylogénétiques et l'existence éventuelle d'une CRF\_MO. Les résultats ont montré l'absence de CRF\_MO et la distinction de 19 URF\_MO, séquencées complètement ou partiellement, soulignant la diversité de ces formes recombinantes et leur complexité moléculaire (FIGURE 20).



**FIGURE 20 :** Diversité génétique des fragments impliqués dans les 19 formes recombinantes identifiées jusqu'à présent.

Sont représentés d'abord les génomes des recombinants complètement séquencés, par ordre croissant du numéro d'identifiant, puis les recombinants partiels, par ordre croissant du numéro d'identifiant. Les fragments appartenant à un VIH-1/O sont représentés en bleu. Les fragments appartenant à un VIH-1/M sont représentés en rouge.

Pour 12 de ces 19 URF\_MO, un génome complet a été obtenu. En revanche, 7 génomes sont restés partiellement séquencés, du fait de la présence de plusieurs populations virales compliquant l'analyse de séquençage, de l'amplification impossible de l'intégralité du génome recombinant ou d'un volume de prélèvement insuffisant. Dans ces cas, la taille des fragments séquencés était variable, allant de 550 à 4000 pb environ.

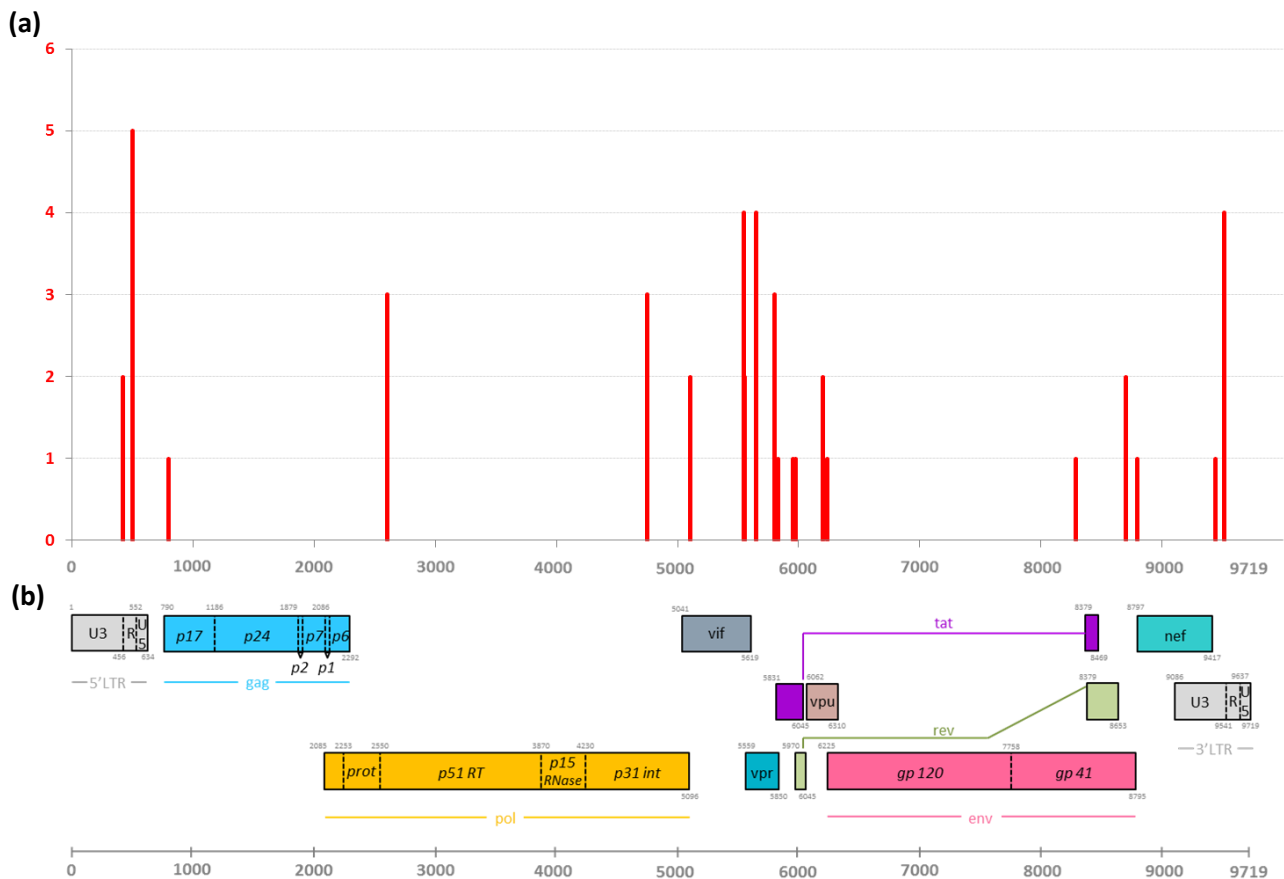
L'analyse phylogénétique a montré que la majorité des fragments de VIH-1/M appartient au sous-type CRF02\_AG (52,6%). Puis, les sous-types impliqués sont le D (15,8%), A et F (10,5% chacun), CRF11\_cpx et une organisation mosaïque intra-M (5,3% chacun). Ces données ne sont pas étonnantes puisque les 22 patients étaient originaires du Cameroun ou présentaient un lien avec ce pays. Elles corrélaient d'ailleurs assez bien avec l'épidémiologie moléculaire du VIH-1/M au Cameroun, qui montre un gradient proche avec d'abord le sous-type CRF02\_AG (50-58%), puis les sous-types G (4,5-11%), D (3-4%), A (2-3,3%) et F (2-4,7%) [202, 203].

Concernant le VIH-1/O, 84% des fragments de groupe O appartiennent au sous-groupe H et seuls 16% au sous-groupe T. Ces résultats sont également comparables à l'épidémiologie moléculaire au Cameroun, caractérisée par un sous-groupe H majoritaire à 77% et un sous-groupe T minoritaire à 23% [73].

### III-3.2. Analyse des points de cassure des 19 URF\_MO

L'analyse des trois premières formes recombinantes intergroupes MO avait laissé suspecter un point chaud de recombinaison dans le gène *vpr*. La découverte de nouveaux recombinants VIH-1/MO donnait l'opportunité d'analyser la fréquence des différents points de cassure détectés, de déterminer un gradient de répartition des points de cassure le long du génome mosaïque et d'identifier d'autres points chauds de recombinaison et d'éventuelles régions génomiques virales sans aucun point de cassure.

L'analyse de la répartition de tous les points de cassure des 19 URF\_MO a permis d'en répertorier 44, dont 32 étaient situés entre les gènes *gag* et *nef* [201]. L'étude de leur fréquence et de leur répartition le long du génome a confirmé la présence d'un point chaud de recombinaison dans le gène *vpr* (18% des points de cassure totaux) et a permis d'en identifier d'autres, notamment dans les LTRs, les gènes *vif* et *vpu* et la région *gp41* (FIGURE 21).



**FIGURE 21** : Localisation des points de cassure dans le génome des 19 URF\_MO décrites actuellement. (d'après Moisan *et al* [201])

La répartition et la fréquence des points de cassure le long du génome est représentée en (a) et l'organisation des gènes en trois cadres de lectures chevauchants des VIH-1 en (b).

Par ailleurs, cette analyse a également révélé l'absence d'événement de recombinaison au sein des régions PR et gp120, et des gènes *rev* et *nef*, suggérant une probable importance de l'homologie de groupe de ces régions dans la viabilité d'une souche (FIGURE 21). Ces résultats sur la fréquence faible ou l'absence de point de cassure dans certaines régions semblent corrélés avec ce qui est décrit pour la recombinaison intra-M. En effet, une étude de la distribution des points "froids" de recombinaison (par opposition aux points chauds) a montré que le gène *gag* et la gp120 présentaient les fréquences de recombinaison intra-M les plus faibles [204]. Dans notre analyse, cette fréquence était de 2,3% (n=1) dans le gène *gag* et nulle dans la gp120 [201].

Finalement, l'étude de la localisation précise des points de recombinaison dans chacune des régions, entre les gènes *gag* et *nef*, a indiqué que 29 des 32 points de cassure (91%) sont situés dans les zones chevauchantes ou aux extrémités des gènes, et seuls 3 (9%) en leur sein.

Ces résultats sont à prendre avec précaution, les génomes n'ayant pas tous été séquencés complètement, mais semblent cohérents avec les études précédentes sur les VIH-1/M, qui suggéraient une distribution préférentielle des points de cassure aux extrémités des gènes [108, 109, 115].

### III-4. Impacts potentiels de la recombinaison inter-groupes MO

Le premier impact potentiel est lié à la présence de fragments de groupe O au sein de la forme recombinante VIH-1/MO et donc à l'impact de la diversité génétique du VIH-1/O lui-même.

Dorénavant, les VIH-1/O ne posent plus de difficultés majeures au niveau du diagnostic ou du suivi virologique de l'infection, du fait de l'adaptation des outils de dépistage sérologique à la diversité des VIH et du caractère non spécifique de groupes des trousse de quantification les plus utilisées. Paradoxalement, sur le plan virologique, le développement de ces trousse non spécifiques de groupes rend, à présent, impossible la détection de ces recombinants VIH-1/MO, puisque, par le passé, c'était l'observation de discordances qui entraînait leur suspicion. Cette absence de spécificité de groupes permet pourtant de les suivre et de les quantifier de façon fiable. Désormais, seule l'absence d'amplification d'une ou de plusieurs régions lors du séquençage de résistance (qui utilise encore des tests d'amplification spécifiques des groupes M et O) peut permettre de suspecter un virus atypique tel qu'un recombinant VIH-1/MO.

Sur le plan thérapeutique, l'observatoire national des variants de groupe O (ORIVAO), en France, et l'étude DYNAMO, au Cameroun, ont montré qu'il n'y avait pas d'impact de la diversité des VIH-1/O sur la réponse immuno-virologique aux traitements ARV actuels, à l'exception des inhibiteurs non nucléosidiques de la transcriptase inverse (INNTI) [170, 205-207]. Les dernières recommandations de l'OMS sur l'utilisation de l'inhibiteur de l'intégrase (II) Dolutégravir (DTG), en première ligne, devrait donc également résoudre cette difficulté, cette molécule étant efficace sur la TI des VIH-1/O [208-210]. Cependant, du fait de l'utilisation des INNTI en première ligne par le passé et jusqu'à la disponibilité du DTG à large échelle, l'absence d'identification d'une infection par un virus recombinant VIH-1/MO, avec une TI appartenant au VIH-1/O, a pu conduire, particulièrement dans la zone d'endémie d'Afrique centrale, à la mise sous traitement par une combinaison thérapeutique incluant un INNTI, avec un risque d'émergence et de transmission de variants résistants.



Sur le plan de la capacité répliquative, l'étude de Peeters et *al.* en 1999 a suggéré une supériorité répliquative du virus recombinant VIH-1/MO quand il était en compétition avec sa souche parentale VIH-1/M, suggérant un avantage répliquatif conféré par la recombinaison MO. Aucune étude depuis n'a été réalisée pour confirmer ou infirmer ces premiers résultats.

Enfin, les différences observées dans les stratégies de contournement des facteurs de restriction cellulaires entre les variants de groupes M et O pourraient, elles aussi, conférer à la forme recombinante VIH-1/MO un avantage potentiel. En effet, selon les régions conservées de chacun des deux groupes, après recombinaison, le contournement de l'activité antivirale de ces facteurs de restriction cellulaire pourrait se révéler plus efficace.

Des hypothèses sur ce potentiel impact des facteurs de restriction cellulaires, des réseaux de co-évolution et des interactions protéiques sur l'émergence réussie d'une forme recombinante VIH-1/MO ont déjà pu être posées, dans le cadre d'une précédente étude, réalisée au laboratoire <sup>[201]</sup>. L'analyse de l'origine parentale des protéines accessoires avait montré que la recombinaison pourrait perturber la fonctionnalité anti-facteurs de restriction de certaines d'entre elles. En effet, dans les URF\_MO analysées, les protéines Vpu et Vif étaient majoritairement issues du groupe M et du groupe O, respectivement, et la protéine Nef, quant à elle, n'était jamais chimérique, suggérant des profils préférentiels de recombinaison potentiellement plus favorables au contournement efficace des facteurs de restriction cellulaires. L'origine parentale des couples de protéines qui interagissent avait, quant à elle, montré une majorité de couples homologues (du même groupe), suggérant l'importance de ces interactions pour la fonctionnalité et la viabilité d'une souche recombinante VIH-1/MO.

# **RATIONNEL DU TRAVAIL EXPERIMENTAL**



L'existence, la circulation et la transmission *in vivo* de formes recombinantes VIH-1 inter-groupes M et O viables et infectieuses suggèrent que ce phénomène de recombinaison entre des variants hautement divergents pourrait conférer à la forme recombinante VIH-1/MO un avantage, par rapport à ses souches parentales de groupes M et O.

L'observation de profils de recombinaison communs et de points de cassure aux mêmes endroits du génome pour différentes URF\_MO soulève, elle, la question de profils préférentiels de recombinaison, pouvant :

- améliorer les stratégies de contournement des facteurs de restriction cellulaires, tout en conservant les réseaux de co-évolution et les interactions protéiques essentiels à la fonctionnalité et à l'infectiosité du virus,
- augmenter le potentiel répliatif de la souche recombinante VIH-1/MO générée par rapport celui de l'une ou de ses deux souches parentales,
- combiner de manière avantageuse les mutations de résistance, naturelles ou sélectionnées, aux traitements ARV des deux souches parentales.

Concernant la capacité répliatif de différents variants de VIH, des premiers éléments de réponse ont été apportés par deux études ayant mis en évidence une supériorité répliatif des variants de groupe M par rapport aux variants de groupe O <sup>[169]</sup> et une apparente supériorité répliatif d'une forme recombinante VIH-1/MO, par rapport à sa souche parentale VIH-1/M <sup>[192]</sup>. De la même façon, une supériorité en terme de niveau de charge virale avant la mise sous ARV des patients et en terme d'évolution de l'infection avait été préalablement rapportée en faveur des variants VIH-1/M, par rapport aux variants VIH-1/O <sup>[170, 205]</sup>. Ces résultats suggèrent que la recombinaison inter-groupes M et O, en fonction de la localisation des points de cassure, pourrait générer des virus recombinants VIH-1/MO davantage agressifs et répliatifs que leurs formes parentales VIH-1/M et/ou VIH-1/O. Ainsi, l'absence de points de cassure relevés au sein de l'enveloppe parmi les recombinants intra-groupe M <sup>[108]</sup> mais aussi parmi les URF\_MO <sup>[201]</sup> suggère un avantage sélectif à transmettre une enveloppe complète d'un virus à l'autre lors d'un évènement de recombinaison. De plus, les VIH-1/O sont caractérisés par des plus faibles prévalence et diffusion que les VIH-1/M, pandémiques. Cette caractéristique est potentiellement liée à leur plus faible capacité à être transmis et à établir une infection chez leur hôte, propriétés impliquant, entre autres, l'enveloppe virale <sup>[169]</sup>. Une forme recombinante, dont l'enveloppe dériverait d'un VIH-1/O, pourrait donc échapper au

système immunitaire de l'hôte et fusionner avec la membrane cellulaire, de manière moins efficace. Ces éléments, combinés à la résistance naturelle aux INNTI, permettent de suggérer qu'il serait plus avantageux pour un recombinant VIH-1/MO de posséder une enveloppe de VIH-1/M et une TI de VIH-1/O.

Par ailleurs, la description de doubles infections VIH-1/M+O et de formes recombinantes VIH-1/MO pose la question des modalités d'émergence de tels recombinants. En effet, différents cas de figure ont été décrits par différentes équipes : surinfection, double infection VIH-1/M+O sans recombinant VIH-1/MO, co-infection VIH-1/MO±M±O, recombinant VIH-1/MO isolé sans souche parentale. Seule une étude, à ce jour, a réalisé une analyse séquentielle de l'évolution des différentes populations virales chez une patiente [194, 195]. Trois prélèvements datant de 1994, 1995, et 1996 ont été analysés. Une triple infection VIH-1/M+M+O avait été mise en évidence en 1994 [195] et une forme recombinante VIH-1/MO avait été retrouvée dans les prélèvements de 1995 et 1996, avec, en parallèle, une apparente disparition des deux populations parentales VIH61/Mau cours du temps [194]. Cependant, la recherche de la souche recombinante n'ayant pas été réalisée sur le prélèvement de 1994, l'équipe n'avait pu conclure quant à la date d'émergence du recombinant. A ce jour, jamais la cinétique d'émergence d'une forme recombinant VIH-1/MO à partir d'une double infection VIH-1/M+O n'a donc été décrite *in vivo*.

Face à ces différentes observations, nous avons cherché à étudier le potentiel répliatif et l'émergence *in vitro*, d'une part, et *in vivo*, d'autre part, de formes recombinantes VIH-1 inter-groupes M et O.

Pour y parvenir, la première étape a été de générer des clones moléculaires infectieux chimériques pVIH-1/MO, à partir de souches parentales de groupes M et O connues et selon les points chauds de recombinaison préalablement rapportés [201]. L'objectif était ensuite de produire les virus recombinants MO correspondants et de comparer leur potentiel répliatif en culture cellulaire *in vitro* à celui de leurs souches parentales VIH-1/M et VIH-1/O, en fonction notamment de l'origine parentale de leur TI et de leur enveloppe.

La seconde étape a été d'étudier les modalités et la dynamique d'émergence des ces virus recombinants.

A cette fin, un premier volet *in vitro* a consisté à analyser d'abord la cinétique d'émergence de points de cassure dans les LTRs des virus recombinants générés *in vitro* lors de l'étape précédente, puisque cette région a été identifiée comme un point chaud de la recombinaison. La localisation précise des points de cassure et leur fréquence ont ensuite été étudiées et comparées à celles décrites *in vivo* pour les LTRs recombinants d'URF\_MO.

Par ailleurs, un second volet *in vivo* a été réalisé, afin de caractériser les différentes populations virales, dans un contexte de discordance moléculaire observée lors d'un génotypage de résistance de routine chez une patiente, et afin de rechercher plus particulièrement la présence d'une forme recombinante VIH-1/MO. A cette occasion, grâce à l'analyse de prélèvements séquentiels répliatifs et à l'utilisation de techniques d'amplification spécifiques de groupes M et O, nous avons déterminé, pour la première fois *in vivo*, la dynamique virale de genèse d'une forme recombinante VIH-1/MO.



**PARTIE 1.**  
**ÉTUDE DU POTENTIEL REPLICATIF *IN VITRO* DE  
VIRUS RECOMBINANTS INTER-GROUPES M ET O**





## Contexte et objectifs du travail expérimental

---

Les données recueillies à partir des URF\_MO actuellement décrites *in vivo* et la présence chez des patients de formes recombinantes VIH-1/MO sans aucune des deux formes parentales d'origine suggèrent l'existence de propriétés phénotypiques conférant un avantage sélectif de ces formes recombinantes VIH-1/MO, par rapport aux formes parentales VIH-1/M et VIH-1/O.

Sur le plan répliatif, une étude de culture en compétition *in vitro* a montré que le *fitness* du VIH-1/O pouvait être 10 à 100 fois inférieur à celui du VIH-1/M, selon les souches étudiées, orientant vers une nette supériorité répliatique du groupe M [169, 211]. Le phénomène de recombinaison inter-groupes M et O pourrait ainsi rééquilibrer cette « pyramide des *fitness* » et donner un avantage répliatif à un recombinant MO par rapport à l'une, l'autre ou les deux souches parentales. A ce jour, seule une étude a étudié le potentiel répliatif d'une forme recombinante VIH-1/MO et les résultats indiquaient une meilleure capacité adaptative (*fitness*) *in vitro* de la souche recombinante par rapport à la souche parentale VIH-1/M, suggérant une potentielle meilleure diffusion et donc un avantage sélectif conféré par la recombinaison, en l'absence de pression de sélection [192].

Par ailleurs, les découvertes de ces multiples virus recombinants VIH-1/MO viables, fonctionnels et répliatifs indiquent leur capacité à contourner efficacement les facteurs de restriction cellulaires et à conserver les interactions et les réseaux de co-évolution protéiques, après les événements de recombinaison. Ceci suggère la présence de motifs particuliers de recombinaison, ayant été sélectionnés car viables et/ou avantageux, face à la pression de sélection exercée par l'hôte. Si les multiples URF\_MO, décrites depuis 1999, présentent des profils et motifs de recombinaison différents, elles partagent des points chauds de recombinaison. Leur étude approfondie a notamment confirmé l'existence d'un tel point chaud dans la région des gènes accessoires centraux (*vif*, *vpr* et *vpu*), et plus particulièrement au sein du gène *vpr*. Ceci a permis d'identifier deux profils majoritaires et simples de recombinaison, avec un unique point de cassure au sein du génome (régions LTR exclues). Le premier, nommé [M<sub>gag/pol</sub>-O<sub>env</sub>], correspond à une première moitié du génome, Gag/Pol, issue d'un VIH-1/M et

une deuxième moitié, Env, issue d'un VIH-1/O. Le second profil présente le motif inverse de recombinaison,  $[O_{gag/pol}-M_{env}]$ , avec la première et la deuxième moitiés du génome issues de VIH-1/O et de VIH-1/M, respectivement. Au total, 11 des 12 URF\_MO complètement séquencées actuellement, soit 92%, présentent l'un de ces deux profils, réparties en 7 profils  $[O_{gag/pol}-M_{env}]$  (64%) contre 4 profils  $[M_{gag/pol}-O_{env}]$  (36%) (FIGURE 22).

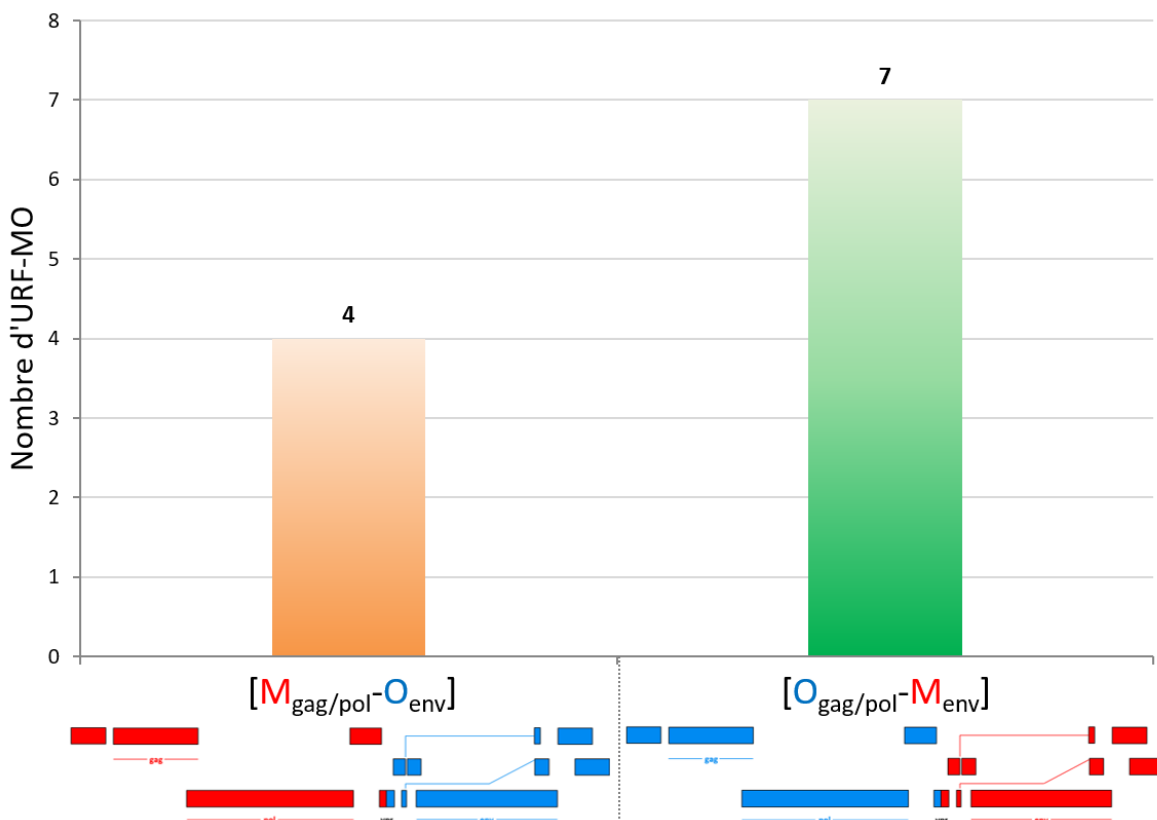


FIGURE 22 : Répartition des profils de recombinaison de 11 des 12 URF\_MO complètement séquencées.

Bien que les effectifs de départ soient faibles et qu'il n'y ait pas d'étude épidémiologique associée, la représentation plus importante du profil  $[O_{gag/pol}-M_{env}]$  pourrait suggérer qu'il est plus avantageux pour le virus d'avoir une TI de VIH-1/O et une enveloppe de VIH-1/M.

Afin de confirmer ces hypothèses d'un avantage réplcatif conféré par la recombinaison et de l'existence de profils préférentiels de recombinaison, notre travail a consisté, dans un premier temps, à générer des virus recombinants selon le profil  $[gag/pol-env]$  et, dans un second temps, à analyser leur potentiel réplcatif *in vitro* par rapport à leurs souches parentales VIH-1/M et VIH-1/O.

Pour ce faire, nous avons utilisé une approche, basée sur la production de virus recombinants VIH-1/MO à partir de clones moléculaires infectieux chimériques (CMIC), grâce à un système de génétique inverse. Les clones moléculaires infectieux (CMI) sont des outils utiles pour étudier diverses propriétés virales. Ils consistent en une séquence complète du génome du VIH insérée dans un plasmide vecteur et peuvent être transfectés dans des cellules sensibles pour produire des virions nouvellement formés. Ces virions découlent tous d'une séquence génétique unique et connue, et proviennent de virus parentaux connus et disponibles, ce qui facilite la comparaison des données. Compte tenu de l'objectif de comparer les deux profils identifiés, nous avons choisi de construire des chimères : [M<sub>gag/pol</sub>-O<sub>env</sub>] ou [O<sub>gag/pol</sub>-M<sub>env</sub>], avec un point de cassure dans *vpr*.

L'éventuel avantage conféré par la recombinaison sur le potentiel répliatif a ensuite été analysé en comparant *in vitro* les cinétiques de répliation de la forme recombinante VIH-1/MO générée par rapport à celles des deux souches parentales VIH-1/M et VIH-1/O.

## A. Génération de virus recombinants inter-groupes M et O

---

### I- Méthodes

#### I-1. Génération de clones moléculaires infectieux chimériques

##### I-1.1. Modalités de génération des clones moléculaires infectieux chimériques

###### a) Sélection des souches parentales et choix des motifs de recombinaison

La première étape de ce travail consistait à mettre au point les CMIC à générer, en sélectionnant d'abord les CMI parentaux à utiliser, en positionnant ensuite les points de cassure dans les génomes recombinants, en déterminant leur motif de recombinaison et, enfin, en choisissant la méthode à appliquer pour la production et le clonage des CMIC.

Compte tenu de l'approche de génétique inverse que nous souhaitions mettre en place, les souches parentales devaient être disponibles sous la forme de CMI, pouvant générer des progénies virales matures et infectieuses. Deux CMI de VIH-1/O étaient disponibles au laboratoire : pRBF206 (sous-groupe T) et pBCF120 (sous-groupe H). Quatre CMI de VIH-1/M ont été sélectionnés dans le catalogue CFAR (Center For AIDS Reagent, section du NIBSC) : p89.6 (sous-type B), p92UG0001 (sous-type D), pNL4-3 (sous-type B) et pCRII-TOPO-SE8131-3 (sous-type A). Les critères de sélection de ces derniers étaient les suivants : ils devaient être facilement amplifiables sur bactéries et correspondre à des souches de VIH-1/M de référence reconnues et circulant en Afrique et/ou en France.

La première construction testée impliquait les deux CMI parentaux pRBF206 pour le groupe O et p89.6 pour le groupe M. Le CMI pRBF206 a été élaboré dans le cadre d'une collaboration avec l'équipe de Kirchoff et *al.* [172, 212]. Il est constitué de la séquence du génome complet de la souche de VIH-1/O RBF206, LTRs compris, insérée dans le plasmide vecteur

pSMART. Ce dernier contient une origine de réplcation, nécessaire à sa réplcation autonome, et un gène de résistance à l'ampicilline, permettant son isolement à partir de bactéries transformées [213]. Le CMI p89.6 est constitué du génome complet de la souche de VIH-1/M de référence 89.6, avec ses deux LTRs, inséré dans le plasmide vecteur pUC19, possédant également un gène de résistance à l'ampicilline [214].

Une fois les CMI parentaux sélectionnés, les motifs de recombinaison ont été déterminés et les points de cassure ont été positionnés. Les motifs choisis lors de la construction des CMIC mimaient les deux profils de génomes recombinants majoritaires, préalablement identifiés *in vivo* : [O<sub>gag/pol</sub>-M<sub>env</sub>] et [M<sub>gag/pol</sub>-O<sub>env</sub>] et devaient permettre de produire des virus recombinants VIH-1/OM et VIH-1/MO, respectivement. Compte tenu du point chaud de recombinaison identifié dans *vpr*, le point de cassure a naturellement été positionné dans ce gène. Lors de travaux préliminaires, basés sur une stratégie de restriction puis ligation, une région conservée dans les génomes de groupes M et O avait été identifiée dans *vpr*. Elle correspondait à l'emplacement du site de restriction de l'enzyme *AvrII*, de séquence 5'-CCTAGG-3'. Une première approche a donc consisté à positionner le point de cassure dans cette région conservée, entre les bases 5661 et 5667 de la souche de référence HxB2 (FIGURE 23).

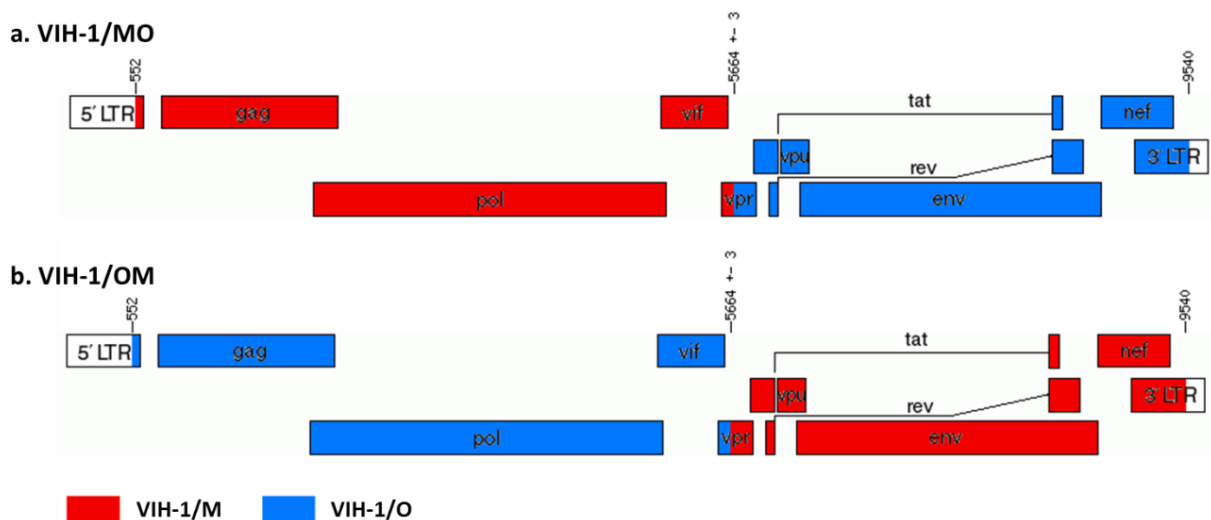
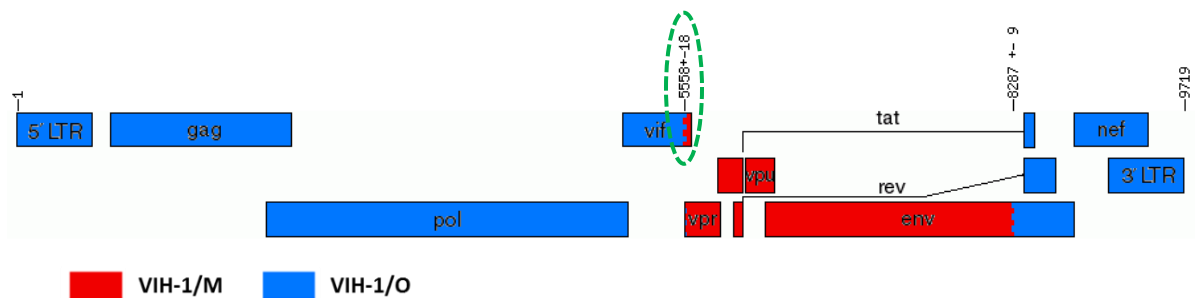


FIGURE 23 : Représentation des génomes de VIH recombinants.

Deux motifs de recombinaison ont été envisagés, avec point de cassure dans le gène *vpr* : le motif [M<sub>gag/pol</sub>-O<sub>env</sub>], en (a) et le motif [O<sub>gag/pol</sub>-M<sub>env</sub>], en (b). La localisation des points de cassure est indiquée par rapport à la souche de référence VIH-1/M HxB2 (numéro d'accèsion GenBank K03455).

Face à l'impossibilité de générer de virus recombinants avec cette première approche, une deuxième construction a été testée, avec le même point de cassure *AvrII* et impliquant pRBF206 et le CMI p92UG0001, constitué du génome complet de la souche de VIH-1/M, inséré dans le plasmide vecteur pCR-Script, avec un gène de résistance à l'ampicilline [215].

Concernant la localisation du point de cassure, une seconde approche a également été testée, reposant sur la reproduction à l'identique du point de cassure retrouvé dans la région chevauchante des gènes accessoires *vif-vpr* du recombinant RBF208, isolé et caractérisé par notre équipe en 2010 [200] (FIGURE 24).



**FIGURE 24** : Génome complet de RBF208  
(d'après Vessière et al. [200])

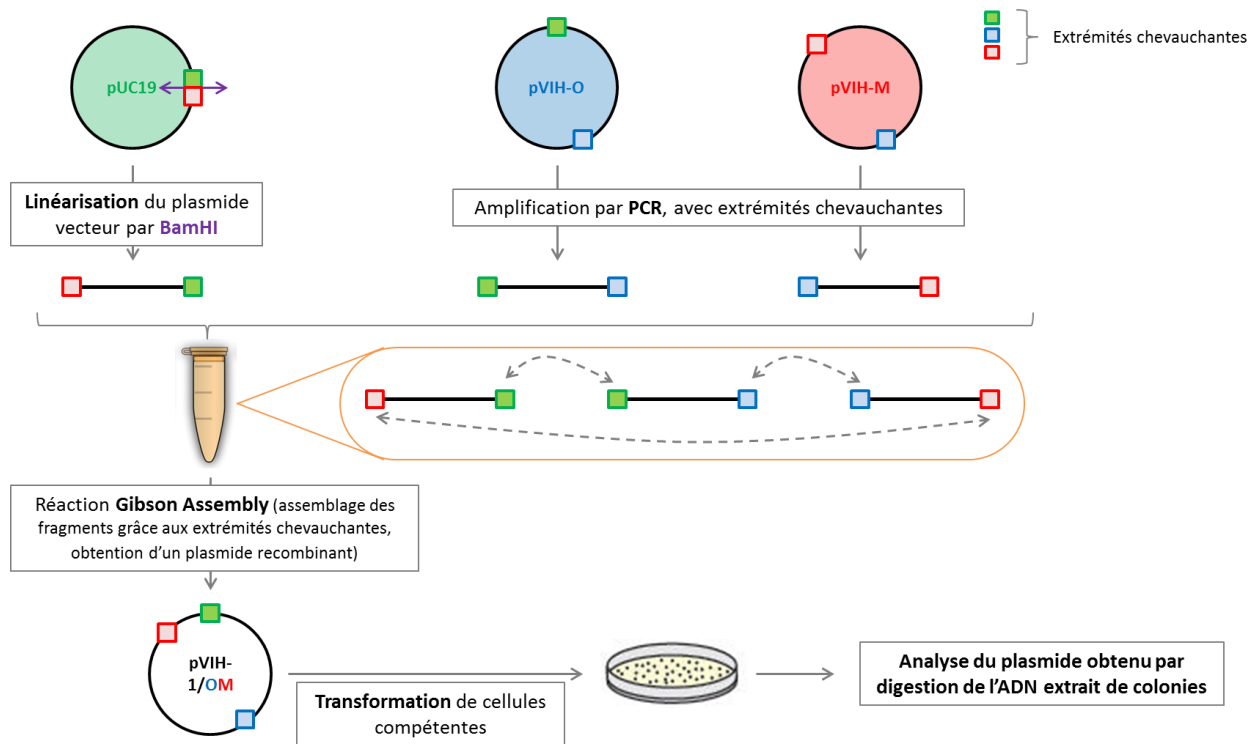
Deux points de cassure ont été identifiés dans cette forme recombinante : *vif-vpr* (motif OM), entouré en vert et reproduit à l'identique dans notre construction, et *env* (motif MO), non pris en compte dans notre construction présentant un unique point de cassure. La localisation des points de cassure est indiquée par rapport à la souche de référence VIH-1/M HxB2 (numéro d'accèsion GenBank K03455).

Les virus attendus devaient avoir les deux mêmes motifs [ $O_{gag/pol}-M_{env}$ ] et [ $M_{gag/pol}-O_{env}$ ] que ceux de la stratégie précédente. Cependant, le point de cassure était positionné entre les bases 5540 et 5576 (souche de référence HxB2) et correspondait à celui de la forme recombinante RBF208.

### b) Stratégie de génération des CMIC

La méthode Gibson Assembly, au moyen du kit *Gibson Assembly™ Cloning Kit* (New England Biolabs – Ipswich, Massachusetts), a été utilisée pour produire les CMIC d'intérêt, au sein de notre laboratoire.

Cette méthode repose sur l'assemblage en une seule étape d'un plasmide vecteur linéarisé et des fragments d'intérêt, amplifiés par une PCR mettant en jeu des amorces avec des queues chevauchantes [216, 217] (FIGURE 25).

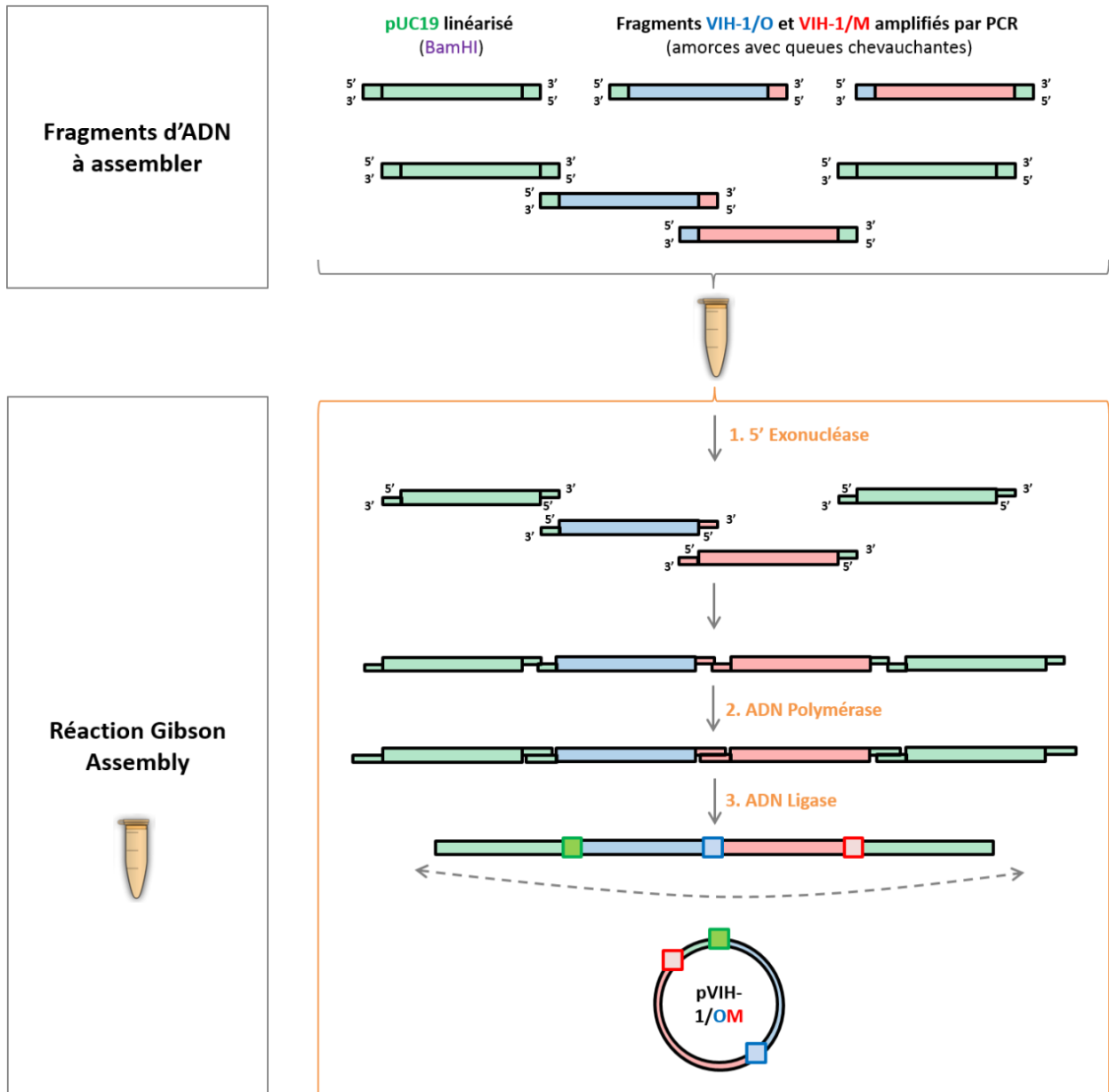


**FIGURE 25 :** Principe de la méthode Gibson Assembly, appliqué à la construction pVIH-1/OM

Les deux amplicons obtenus par PCR à partir des CMI parentaux et le vecteur pUC19 digéré par BamHI sont assemblés deux à deux grâce à leurs extrémités chevauchantes et grâce à l'action simultanée des enzymes engagées dans la réaction Gibson Assembly. Le CMIC obtenu est amplifié sur bactéries puis analysé.

La méthode Gibson Assembly combine l'action de trois enzymes différentes, lors d'une même réaction isothermique, avec un unique tampon. Tout d'abord, une 5' exonucléase génère des extrémités chevauchantes en lysant les régions en 5' de chacun des fragments à assembler, à savoir le plasmide vecteur linéarisé, et les fragments de VIH-1/M et VIH-1/O. Les différents fragments chevauchants s'hybrident alors entre eux. Ensuite, une ADN polymérase incorpore les nucléotides du milieu réactionnel au niveau des zones coupées par la 5' exonucléase, ce qui permet donc la synthèse des brins complémentaires. Enfin, une ADN ligase lie les fragments adjacents hybridés grâce à leur complémentarité et reconstitue ainsi un clone moléculaire circulaire complet et chimérique (FIGURE 26).





**FIGURE 26 :** Mode d'action des différentes enzymes impliquées dans la réaction Gibson Assembly

La réaction Gibson Assembly a lieu grâce à l'action successive de trois enzymes distinctes. Une 5' exonuclease lyse d'abord les extrémités des fragments à amplifier. Ensuite, une ADN polymérase synthétise les brins complémentaires au niveau des fragments chevauchants hybridés. Puis, une ADN ligase circularise la construction obtenue, générant ainsi le CMIC d'intérêt.

**Plasmide vecteur utilisé.** Le choix s'est porté sur le plasmide vecteur pUC19 fourni en petites quantités dans le kit *Gibson Assembly™*. Un stock a donc d'abord été constitué, en transformant par électroporation des bactéries *Escherichia coli*, lignée DH5α, préalablement rendue électrocompétente (*NEB 5-alpha Competent E. coli – New England Biolabs*). Dans la cuve réfrigérée de l'électroporateur, 1μL de la suspension de plasmide a été mélangé à 20μL de bactéries électrocompétentes DH5α conservées dans de la glace. Après décharge du

condensateur de l'électroporateur (2500 Volts), les bactéries électroporées ont été reprises dans 1mL de bouillon SOC (Super Optimal broth with Catabolite repression) stérile et incubées pendant une heure et trente minutes dans une étuve à 30°C sous agitation constante (150rpm), afin de permettre l'expression du gène de résistance à l'ampicilline. Après ces étapes d'électroporation puis d'incubation, les bactéries ont été mises en culture à une température de 30°C dans trois boites de géloses LB (Luria-Bertani) contenant 100µg/mL d'ampicilline, avec 10, 50 et 100µL du bouillon obtenu. Le plasmide possède dans son génome un gène de résistance à l'ampicilline. Pour s'assurer qu'une seule souche bactérienne contenant un plasmide pur avait bien été cultivée, des réisolements à partir d'une unique colonie ont été réalisés, après 24h de culture sur gélose LB + Ampicilline 100µg/mL à 30°C. Dès que des colonies étaient observées sur les réisolements, nous avons procédé à des étapes de purification et d'extraction du clone contenu dans les bactéries cultivées. Une étape préalable de culture des *E. coli* transformées en milieu liquide BHI (Brain Heart Infusion) + Ampicilline 100µg/mL, sous agitation (250rpm), à 30°C et pendant toute une nuit était nécessaire afin d'obtenir une grande quantité de plasmide. Le kit de MINIPrep (*Nucleospin® Plasmid – Macherey-Nagel – Düren, Allemagne*) a été utilisé pour la lyse des bactéries, le dépôt sur colonne de silice, le lavage et l'élution de l'ADN, conformément aux recommandations du fabricant. Une fois l'ADN plasmidique pur récupéré, une quantification de l'éluât a été effectuée grâce à un spectrophotomètre UV (*GE NanoVue Spectrophotometer® – GE Healthcare Life Sciences – Uppsala, Suède*) afin de vérifier la présence d'ADN plasmidique et d'en déterminer la concentration.

**Obtention des fragments à assembler.** La première étape de production des CMIC consistait à générer les différents fragments constitutifs.

Pour ce faire, le vecteur pUC19, initialement circulaire, a été linéarisé au moyen d'une digestion par l'enzyme de restriction BamHI-HF® (*New England Biolabs*). Cette digestion a été réalisée selon un protocole standard de restriction, dans les conditions recommandées par le fabricant, avec une heure d'incubation à 37°C. Les produits de digestion ont ensuite été déposés sur gel d'agarose et SYBRSafe®, afin de vérifier la taille du fragment obtenu après digestion, par comparaison à un marqueur de taille. La concentration du gel d'agarose a été préalablement déterminée en fonction de la taille des fragments attendus.

Les fragments de VIH-1/M et de VIH-1/O ont été générés par PCR. Le TABLEAU 3 récapitule le nom des différents CMIC à générer, leur motif, les CMI parentaux utilisés, et la localisation du point de cassure.

TABLEAU 3 : Nomenclature et caractéristiques des CMIC à générer

	Motif de recombinaison	Souche parentale VIH-1/M (sous-type)	Souche parentale VIH-1/O (sous-groupe)	Localisation du point de cassure
<b>pChimR100</b>	pVIH-1/OM	89.6 (B)	RBF206 (T)	Vpr 5661-5667* AvrII
<b>pChimR101</b>	pVIH-1/MO	89.6 (B)	RBF206 (T)	Vpr 5661-5667* AvrII
<b>pChimR102</b>	pVIH-1/MO	89.6 (B)	RBF206 (T)	Vpr 5540-5576* RBF208
<b>pChimR103</b>	pVIH-1/MO	92UG0001 (D)	RBF206 (T)	Vpr 5661-5667* AvrII
<b>pChimR104</b>	pVIH-1/MO	92UG0001 (D)	RBF206 (T)	Vpr 5540-5576* RBF208

\*La localisation des points de cassure est indiquée par rapport à la souche de référence VIH-1/M HxB2 (numéro d'accèsion GenBank K03455)

Une seule approche a été réalisée pour obtenir le CMIC pVIH-1/OM. En revanche, quatre tentatives ont été réalisées pour générer le CMIC pVIH-1/MO, en faisant varier le sous-type de VIH-1/M impliqué dans le CMI parental de VIH-1/M (sous-type B ou D) et la localisation du point de cassure dans *vpr* (TABLEAU 3).

La méthode Gibson Assembly repose sur l'utilisation d'amorces aux extrémités chevauchantes deux à deux, qui permettent l'hybridation des fragments provenant de groupes différents entre eux et donc de la formation d'une séquence chimérique. La mise au point de ces amorces chevauchantes et des conditions de PCR a été réalisée grâce à l'outil *NEBuilder™ Assembly Tool*, proposé par le fournisseur du kit *Gibson Assembly™* (disponible à l'adresse <http://nebuilder.neb.com/#/>) (TABLEAU 4).

TABLEAU 4 : Séquences des amorces Gibson Assembly testées selon les CMIC à générer

	Fragment	Amorce sens 5'-3'	Amorce antisens 5'-3'	Taille (pb)
pChimR100	VIH-1/M	M-3'Env-Upp gacatttcCTAGGATATGGCTCCATAGTTTAGGG	M-3'LTR-Low caggtcgtactctagagTGCTAGAGATTTTCTACACTAACTAGGATGG	4050
	VIH-1/O	O-5'LTR-Upp agctcggtagccgggTGGATGGGTTAATTTACTCCTATCAGAG	O-5'Pol-Low atatcctagGAAAATGTCTTACTGCTTTTAGC	5724
pChimR101	VIH-1/M	Rec001U agtgaattcgagctcggtagccgggTGGAAGGGCTAGTTCACTCC	Rec002L agccaattcctagGAAAATGTCTCACAGCTTCATTC	5660
	VIH-1/O	Rec003U gtgagacatttcCTAGGAATTGGCTACAAGCC	Rec004L gcatgcctgcaggtcgtactctagagTGCTAGAGAATTTTCTGCTTC	4153
pChimR102	VIH-1/M	Rec001U agtgaattcgagctcggtagccgggTGGAAGGGCTAGTTCACTCC	Rec005L ctgattctctggtGCTTGTTCATCTATCCTCTG	5568
	VIH-1/O	Rec006U agatggaacaagcACCAGAGAATCAGGGGCCAG	Rec007L gcatgcctgcaggtcgtactctagagTGCTAGAGAATTTTCTGCTTCAGTC	4245
pChimR103	VIH-1/M	Rec008U agtgaattcgagctcggtagccgggTGGAAGGGCTAATTTGGTCC	Rec009L agccaattcctagGAAAATGTCTAACAGCCTCAC	5686
	VIH-1/O	Rec010U gtagacatttcCTAGGAATTGGCTACAAGCC	Rec004L gcatgcctgcaggtcgtactctagagTGCTAGAGAATTTTCTGCTTC	4153
pChimR104	VIH-1/M	Rec008U agtgaattcgagctcggtagccgggTGGAAGGGCTAATTTGGTCC	Rec011L ctgattctctggtGCTTGTTCATCTATCTTCTG	5594
	VIH-1/O	Rec006U agatggaacaagcACCAGAGAATCAGGGGCCAG	Rec007L gcatgcctgcaggtcgtactctagagTGCTAGAGAATTTTCTGCTTCAGTC	4245

Les lettres minuscules figurant dans les séquences des amorces représentent les queues devant s'hybrider, par complémentarité, au fragment adjacent, lors de la réaction Gibson Assembly. Les séquences s'hybridant sur le plasmide vecteur pUC19 sont représentées en vert, et celles s'hybridant sur les pVIH-1/M et pVIH-1/O en rouge et bleu, respectivement.

Etant donné la longueur importante des séquences à amplifier (plus de 4000pb), le kit *Long Range PCR*<sup>®</sup> (*Qiagen – Hilden, Allemagne*), contenant une Taq polymérase adaptée à l'amplification de longs fragments d'ADN génomique et de haute fidélité, a été utilisé. Le mélange réactionnel et le protocole d'amplification étaient les mêmes, quelle que soit la construction envisagée (ANNEXE 1). Les amplicons ont ensuite été déposés sur gel d'agarose à 1% puis purifiés sur colonne de silice, grâce au kit *Nucleospin Gel and PCR clean-up* (*Macherey-Nagel*), selon les recommandations du fabricant. Enfin, la quantité d'ADN obtenue après amplification et purification était mesurée grâce au spectrophotomètre UV *BioDrop*  $\mu$ LITE (*Isogen Life Science, Utrecht, Pays-Bas*).

**Réaction Gibson Assembly.** La seconde étape de production des CMIC consistait à réaliser la réaction Gibson Assembly proprement dite. Comme préconisé par le fabricant, 100ng du vecteur linéarisé pUC19 ont été mélangés à 300ng de chacun des deux fragments de VIH à insérer. Les volumes correspondants avaient préalablement été calculés à partir des quantités d'ADN mesurées. A ce mélange, 10 $\mu$ L du tampon de réaction Master Mix 2X ont été ajoutés, ainsi qu'un volume d'eau permettant d'obtenir un volume réactionnel final de 20 $\mu$ L. Après une incubation d'une heure à 50°C, le produit obtenu était dilué au 1/3, en mélangeant 5 $\mu$ L de ce produit à 10 $\mu$ L d'eau, puis conservé dans de la glace.

### c) Amplification des CMIC obtenus par transformation chimique

Les quantités de CMIC obtenues après la réaction Gibson Assembly étant très faibles, la constitution d'un stock était nécessaire. Ainsi, l'amplification des éventuels CMIC générés a été réalisée par transformation chimique de bactéries réceptrices *XL10-Gold*<sup>®</sup> *Ultracompetent Cells* (*Agilent Technologies, Santa Clara, Californie*), rendues compétentes et conçues pour le clonage de plasmide de grande taille ou d'ADN ligué. Pour ce faire, 2 $\mu$ L de  $\beta$ -Mercapto-Ethanol ( $\beta$ -ME) ont été ajoutés à 45 $\mu$ L de bactéries XL-10 Gold préalablement aliquotées. Le mélange a été incubé pendant 10 minutes dans de la glace, en agitant doucement toutes les 2 minutes pour homogénéiser et optimiser le contact entre les bactéries et le  $\beta$ -ME. Ensuite, 2 $\mu$ L d'ADN plasmidique dilué au 1/3 ont été ajoutés à ce mélange, qui a été agité, puis incubé pendant 30 minutes dans de la glace. Un passage brutal à 42°C, pendant 30 secondes, a ensuite été réalisé, constituant la transformation proprement dite. Puis, 500 $\mu$ L de bouillon NZY+, préalablement préchauffé à 42°C, ont été ajoutés au mélange, mis ensuite à incuber à 37°C pendant 1 heure, sous agitation (150rpm), afin de permettre l'expression du gène de résistance à l'ampicilline.

Après leur incubation, les bactéries ont été mises en culture à une température de 37°C dans deux boîtes de gélose LB contenant 100µg/mL d'ampicilline, avec 50 et 100µL du bouillon obtenu. Afin de s'assurer qu'une seule souche bactérienne contenant un plasmide pur avait bien été cultivée, des réisolements à partir de colonies isolées ont été réalisés après 24h de culture sur gélose LB + Ampicilline 100µg/mL à 37°C et incubées dans les mêmes conditions.

#### d) Vérifications sur les produits obtenus

Une fois les stocks de CMIC d'intérêt constitués, des vérifications ont été réalisées, afin de s'assurer que les CMIC générés par Gibson Assembly et amplifiés par transformation bactérienne correspondaient bien aux CMIC initialement définis. Une méthode simple consistait à amplifier et séquencer les régions d'assemblage des fragments d'origine. Ceci permettait de s'assurer que l'intégralité des fragments était bien présente, mais aussi que les motifs de la construction étaient identiques à ceux qui étaient attendus. Pour ce faire, des PCR spécifiques de deux des trois régions d'assemblage ont été conçues (TABLEAU 5, FIGURE 27). Pour la construction pVIH-1/OM, la PCR pUC-O devait amplifier la région d'assemblage du plasmide vecteur pUC19 et du fragment de VIH-1/O issu de pRBF206, situé en 5', alors que la PCR Vpr OM devait amplifier la région autour du point de cassure dans *vpr*. Concernant la construction inverse pVIH-1/MO, la PCR Vpr MO devait amplifier la région autour du point de cassure dans *vpr* et la PCR pUC-O devait amplifier la région d'assemblage du fragment de VIH-1/O issu de pRBF206 et du plasmide vecteur pUC19, située en 3'. Quand cela était nécessaire, deux PCRs supplémentaires, ciblant les régions TI de VIH-1/M (TI M) et C2V3 de VIH-1/O (C2V3 O), ont également été réalisées et séquencées, afin de vérifier la présence des fragments de VIH-1/M et VIH-1/O de part et d'autre du point de cassure dans *vpr* de la construction pVIH-1/MO.

Le séquençage a été réalisé par la méthode de Sanger, sur séquenceur CEQ 8000 (Beckman Coulter – Fullerton, Californie). Dans chaque puits d'une plaque de 96 puits, 4µL de DCTS Quick Start Kit (Beckman Coulter) ont été mélangés avec 11µL d'eau, 3µL d'amorce à 1,6pmol/µL et 2µL de l'amplicon purifié. Une réaction de séquence a été effectuée selon le programme d'amplification suivant : une dénaturation à 96°C pendant 25 secondes, une hybridation à 50°C pendant 25 secondes puis une élongation de 4 minutes à 60°C pendant 30 cycles. Ensuite, des étapes de précipitation à l'éthanol et de purification ont été réalisées. La plaque a été placée dans le séquenceur, dont le principe repose sur l'électrophorèse capillaire.

Une analyse spectrale des fluorescences, émises par les didesoxyribonucléotides (ddNTPs), a permis de déterminer la séquence du fragment étudié.

TABLEAU 5 : Séquences des amorces utilisées lors de la vérification des constructions

CMIC	PCR	Amorce sens 5'-3'	Amorce antisens 5'-3'	Taille (pb)
pVIH-1 /OM	vpr OM	288U ACACAGATTGAYCCTGAAAC	148L TCCAGGGCTCTAGKCTAGGA	538
	pUC-O	M13F GTAAAACGACGGCCAG	229L GCTTTATTGAGCGTTAAGCAG	607
	O-pUC	340U ACTCCTCARAACAATGCAG	M13R CAGGAAACAGCTATGAC	872
pVIH-1 /MO	vpr MO	142U GGGGTCTGCATACAGGAGAA	291L GCCAAGGGGGCAYCTCAGGA	627
	TI M	A35 TTGGTTGCACTTTAAATTTCCATTAGTCTATT	NE1(35) CCTACTAACTTCTGTATGTCATTGACAGTCCAGCT	805
	C2V3 O	V3DURA ATTCCAATACACTATTGTGCTCCA	V3DURR AAAGAATTCTCCATGACAGTTAAA	518

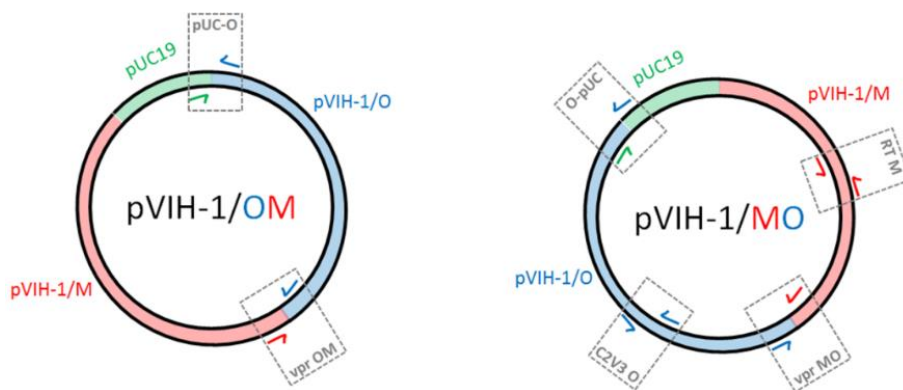


FIGURE 27 : Localisation des PCRs de vérification sur les CMIC générés

Les amorces s'hybridant sur le plasmide vecteur pUC19 sont représentées par des flèches vertes, et celles s'hybridant sur les CMI parentaux pVIH-1/M et pVIH-1/O par des flèches rouges et bleues, respectivement.

A partir de chacun des réisolements réalisés après 24h de culture, une petite quantité de bactéries a été remise en suspension dans 100µL d'eau puis lysée grâce à une incubation de 10 minutes à 100 degrés. L'objectif de la lyse était d'optimiser les PCRs de vérification, en

désagrégeant la membrane bactérienne et libérant ainsi l'ADN plasmidique. Le lysat obtenu a été centrifugé et ajouté à un mélange réactionnel pour amplification (ANNEXE 2).

Quand pour l'une des colonies, les PCRs amplifiaient un fragment visible sur gel d'agarose et de la taille attendue, l'amplicon était purifié sur colonne pour séquençage. En parallèle, la colonie concernée était remise en culture en milieu liquide BHI + Ampicilline 100µg/mL, sous agitation (250rpm), à 37°C et pendant toute une nuit, afin d'obtenir une quantité importante du CMIC. Une étape de purification, d'extraction et de concentration de l'ADN issu de la colonie d'intérêt était réalisée par MiniPrep (*Nucleospin® Plasmid, Macherey-Nagel*), selon les recommandations du fabricant. L'ADN plasmidique obtenu était ensuite quantifié, avec le spectrophotomètre UV *BioDrop µLITE (Isogen Life Science)*.

### I-1.2. Génération des CMIC par synthèse

Face aux difficultés rencontrées au laboratoire pour générer les CMIC, les séquences complètes des génomes mosaïques de motifs OM et MO (correspondant respectivement à la séquence voulue pour les CMIC pChimR100 et pChimR101 TABLEAU 3, P.86) ont été envoyées à la société Genecust (GC), basée au Luxembourg, pour synthèse. La synthèse du CMIC de motif MO n'ayant pas abouti, seul celui de motif OM nous a été envoyé.

Une fois reçu, sous format lyophilisé, et remis en suspension, selon les recommandations du fabricant, le CMIC synthétisé pVIH-1/OM-GC a été amplifié, en transformant par électroporation des bactéries *Escherichia coli*, lignée Stbl4 rendue électrocompétente (*ElectroMAX™ Stbl4™ electrocompetent E. coli – Invitrogen – Carlsbad, Nouveau-Mexique*). Le protocole utilisé était le même que celui utilisé pour le plasmide vecteur pUC19 (Partie I-1.1. Modalités de génération des clones moléculaires infectieux chimériques, P.80). Ainsi, 2µL de plasmide ont été mélangés à 20µL de bactéries Stbl4. Ces dernières, une fois électroporées, ont été reprises dans 1mL de bouillon SOC stérile et incubées pendant une heure et trente minutes à 30°C sous agitation (250rpm), afin de permettre l'expression du gène de résistance à la kanamycine, contenu dans le plasmide synthétisé par Genecust. Puis, 50µL et 100µL du bouillon ont été mis en culture à 30°C dans deux boîtes de géloses LB contenant 50µg/mL de kanamycine. Après 24h, des réisolements sur gélose LB + Kanamycine 50µg/mL à 30°C ont été réalisés, suivis de la purification et de l'extraction du plasmide contenu dans les colonies bactériennes, par MINIPrep, conformément aux recommandations du fabricant. Enfin, une



quantification de l'éluât a été effectuée grâce à un spectrophotomètre UV *BioDrop*  $\mu$ LITE (Isogen Life Science).

Comme recommandé par le fabricant, le produit obtenu a été soumis à une digestion par l'enzyme de restriction *Sma*I, afin de s'assurer que le plasmide issu de la transformation correspondait bien au plasmide attendu, sans excision ni insertion. Le protocole de digestion comportait une heure d'incubation à 25°C et de 30 minutes d'inactivation à 80°C, et était suivi d'une migration sur gel d'agarose pour vérifier l'obtention de deux fragments linéaires, de 10522 et 4836pb.

## I-2. Production de virus recombinants à partir de clones moléculaires infectieux chimériques

Afin de s'assurer de la capacité des CMIC à produire les souches recombinantes correspondantes, nous avons mis en place un protocole reposant sur la transfection de cellules HEK293T à l'aide d'un réactif *JetPRIME*<sup>®</sup> (*Polyplus Transfection, Illkirch, France*) pour produire le virus. Une étape de co-culture avec la lignée cellulaire CEMx174 y était éventuellement associée, afin de permettre la réplication virale (FIGURE 28).

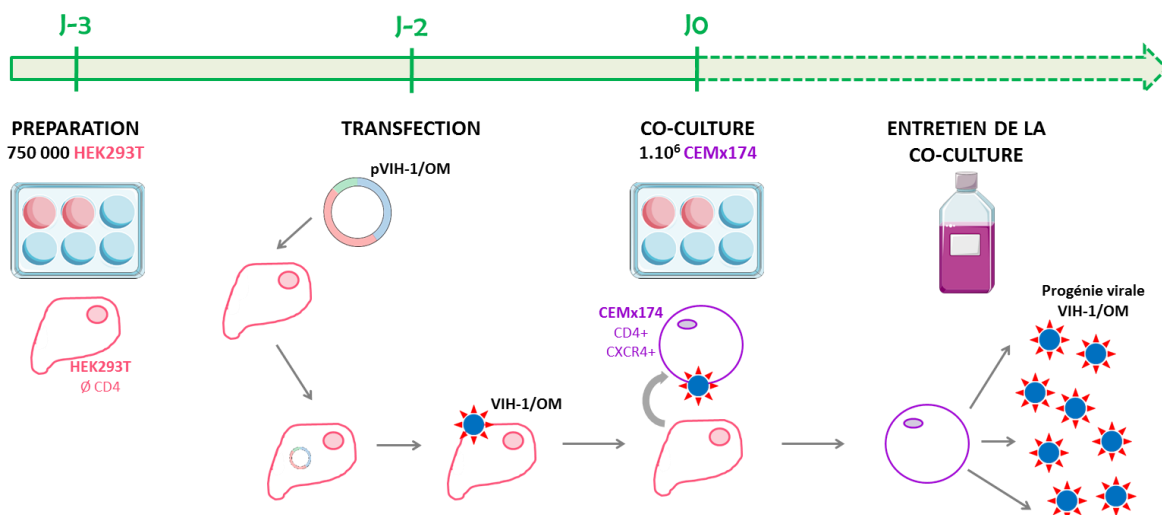


FIGURE 28 : Principe général de la production de virus recombinants VIH-1/OM à partir d'un CMIC

### I-2.1. Transfection

Les cellules HEK293T, adhérentes, ont été maintenues en culture en milieu DMEM (Dulbecco's minimal essential medium) + Glutamax supplémenté par 10% de sérum de veau

foetal décomplémenté (SVFd) et par un antibiotique, la gentamicine, à la concentration de 50µg/mL [218]. Elles ont été entretenues deux fois par semaine avec trypsinisation et réensemencement au tiers.

La veille de la transfection (J-3 de la co-culture), 750 000 cellules HEK293T ont été mises en culture dans une plaque 6 puits dans 2mL de milieu DMEM + 10%SVFd + Gentamicine afin d'obtenir 60 à 80 % de confluence le jour de la transfection. Le lendemain (J-2 de la co-culture), 1µg de CMIC a été dilué dans 200µL de tampon *JetPRIME*<sup>®</sup> (*Polyplus Transfection*) dans un tube Eppendorf. Après agitation et centrifugation, 2µL de réactif *JetPRIME*<sup>®</sup> ont été ajoutés et le mélange a été mis à incuber à température ambiante pendant 10 minutes. Ensuite, 200 µL du mélange réactionnel ont été déposés sur les cellules HEK293T à confluence et 2,5mL de DMEM + 10%SVFd + Gentamicine ont été ajoutés. Après une incubation de quatre heures à 37°C et 5% de CO<sub>2</sub>, le milieu a été renouvelé et la plaque 6 puits a de nouveau été mise à incuber pendant deux jours à 37°C et 5% de CO<sub>2</sub> (FIGURE 28).

## I-2.2. Co-culture

Afin d'amplifier les populations virales ayant bourgeonné à la surface des HEK293T, une étape de co-culture avec la lignée cellulaire CEMx174 a suivi la transfection.

Avant la co-culture, les CEMx174 ont été maintenues en culture dans du milieu RPMI (Roswell Park Memorial Institute) supplémenté par 10% de SVFd et de la gentamicine à la concentration de 50µg/mL [219]. Le milieu a été remplacé et les cellules divisées au 1/5<sup>ème</sup> deux fois par semaine. Trois jours après la transfection, la co-culture a été réalisée (J0) en déposant un million de CEMx174 sur le tapis de HEK293T et en laissant les deux lignées cellulaires en contact pendant 24 heures dans la plaque 6 puits. Les CEMx174 et leur surnageant de culture ont ensuite été transférés en flasque et mis à incuber pendant 15 jours à 37°C et 5% de CO<sub>2</sub>. Le milieu de culture, composé de RPMI + 10%SVFd + Gentamicine, a été renouvelé tous les 3 jours et du surnageant de culture a été prélevé à la même fréquence pour procéder à des contrôles sur la population virale au fil de la culture (FIGURE 28).

## I-2.3. Contrôle de la production et de la réplication virales

Plusieurs tests ont été réalisés pour analyser la bonne production des virus complets par les cellules HEK293T et leur infectiosité, correspondant à leur capacité à infecter des cellules

permissives et sensibles et à y effectuer un cycle de répliation complet, et pour vérifier que les virus infectieux produits avaient bien le profil de départ.

La production virale a été contrôlée par quantification de l'AgP24 dans les surnageants de transfection ou dans les surnageants de co-culture, grâce au kit *VIDAS® HIV P24 II CONFIRMATION* (bioMérieux, Lyon, France), selon les recommandations du fabricant. La charge virale ARN a aussi été mesurée sur le surnageant correspondant au pic d'AgP24, par la méthode *RealTime HIV-Assay* (Abbott Molecular, Chicago, Etats-Unis).

Le caractère infectieux de la progénie virale produite a été contrôlé par le test de TCID50 (Tissue Culture Infective Dose 50), selon la méthode de Montefiori, soit sur le surnageant de transfection, soit sur celui de la co-culture correspondant au pic d'Agp24. Dans les puits d'une plaque 96 puits, 10 000 cellules TZM-bl ont été préalablement ensemencées dans 200µL de milieu DMEM + GlutaMax, supplémenté de 10% de SVFd, de pénicilline (100U/mL), de streptomycine (100µg/mL) et de DEAE/dextran à 15 µg/mL (*Sigma-Aldrich, St Louis, MI, USA*) et ont été infectées par des dilutions en série de virus issus des surnageants. Les TZM-bl ont dans leur génome le gène de la luciférase, sous le contrôle d'un promoteur du VIH-1. En cas d'infection par le virus, l'expression de ce gène est activée, ce qui entraîne la production de luciférase. Après 48 heures d'incubation à 37°C et 5% de CO<sub>2</sub>, 100µL de milieu ont été retirés et remplacés par 100µL de réactif *Luciferase Bright-Glo* (*Promega*), révélant la présence de cette luciférase dans les cellules infectées. La TCID50 a été évaluée par quantification des unités de luminescence relative (RLU) par un lecteur de luminescence *GloMax®-Multi* (*Promega*) pour chaque virus à partir de quatre dilutions en série répliquées.

Enfin, une mesure de l'activité de la TI (*Lenti RT activity, Cavid, Uppsala, Suède*), reflétant la production et l'activité virales, a été réalisée dans le surnageant de co-culture prélevé au pic d'AgP24, selon les recommandations du fabricant.

Le profil de recombinaison des virus infectieux obtenus a été vérifié par PCR. La construction pVIH-1/OM étant la seule à avoir abouti, seul le motif OM a pu être contrôlé, à partir du virus recombinant contenu dans les surnageants de culture. Après extraction de l'ARN avec l'automate EZ1 advanced XL et à l'aide du kit *EZ1® DSP Virus* (*Qiagen*) et selon les recommandations du fabricant, trois régions ont été amplifiées et séquencées par la méthode de Sanger : PR O, Vpr OM (contenant le point de cassure) et C2V3 M (TABLEAU 6). Enfin, quand toutes les vérifications s'étaient avérées positives, un séquençage du génome complet de la souche obtenue a été réalisé, comme précédemment décrit <sup>[199]</sup>.

**TABLEAU 6** : Amorces utilisées lors des vérifications sur les virus recombinants VIH-1/OM produits

	PCR	Amorce sens 5'-3'	Amorce antisens 5'-3'	Taille (pb)
pVIH-1/OM	PR O	Prot 4 CAGCCCCACCRATGGAGG	Nouv L CATTGTTTTACTTTTGGTCCAT	452
	Vpr OM	288U ACACAGATTGAYCCTGAAAC	148L TCCAGGGCTCTAGKCTAGGA	538
	C2V3 M	Env7 AATGGCAGTCTAGCAGAAG	ED33D TTACARTAGAAAAATTCYCCTC	380

## II- Résultats

### II-1. Génération de clones moléculaires infectieux chimériques

#### II-1.1. Modalités de génération des clones moléculaires infectieux chimériques

Afin de sélectionner les CMI parentaux les plus adaptés à nos travaux, les bactéries *E. coli* Stbl4 électrocompétentes ont été transformées avec les deux plasmides de VIH-1/O disponibles, pRBF206 (sous-groupe T) et pBCF120 (sous-groupe H), et les quatre plasmides de VIH-1/M obtenus du CFAR, p89.6 (sous-type B), pNL4-3 (sous-type B), pCRII-TOPO-SE8131-3 (sous-type A), et p92UG0001 (sous-type D).

Lors de sa construction, le plasmide de VIH-1-O pRBF206 avait présenté un meilleur rendement en matière de transformation et d'amplification, ce que nous avons également constaté. C'est donc sur lui que le choix s'est porté.

Pour le VIH-1/M, du fait de la disponibilité de la séquence génétique complète et de difficultés de rendement lors de la transformation avec les trois autres plasmides, nous avons d'abord retenu, le plasmide p89.6 pour la suite des travaux. Le plasmide p92UG0001 a également été utilisé, dans le cadre de l'optimisation de la construction du plasmide pVIH-1/MO, par la méthode Gibson Assembly.

## II-1.2. Génération des CMIC pVIH-1/OM et pVIH-1/MO

### a) Génération du CMIC pVIH-1/OM

La première étape de la méthode Gibson Assembly consistait à amplifier et linéariser le plasmide vecteur pUC19. Après électroporation des bactéries DH5 $\alpha$  par ce vecteur pUC19, 142 ng/ $\mu$ L d'ADN plasmidique ont été extraits puis purifiés. La digestion par BamHI a permis d'obtenir un fragment linéarisé de 2686pb.

**Méthode Gibson Assembly.** Le profil du CMIC pChimR100 correspondait à une partie 5' du génome (des premières bases du 5'LTR au site de restriction AvrII, dans le gène *vpr*) issue du CMI de VIH-1/O pRBF206 et à une partie 3' du génome (du site AvrII aux dernières bases du 3'LTR) issue du CMI de VIH-1/M p89.6 (FIGURE 29).

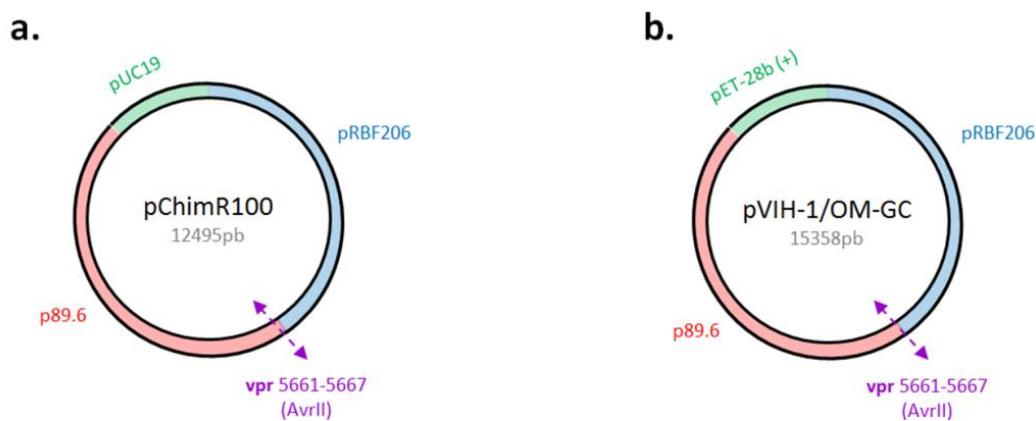


FIGURE 29 : Profils des CMIC pVIH-1/OM

Les portions de génome provenant des CMI pRBF206 (VIH-1/O) et p89.6 (VIH-1/M) figurent en bleu et rouge, respectivement. Les points de cassure correspondent au site de restriction *AvrII*. (a) Le CMIC pChimR100 implique le plasmide vecteur pUC19 (en vert). (b) Le CMIC pVIH-OM-GC implique le plasmide vecteur pET-28b(+) (en vert).

Après PCR, les amplicons obtenus avaient la taille des fragments attendus, 5724pb et 4050pb pour le VIH-1/O et le VIH-1/M, respectivement. Après purification sur colonne, 407 et 363 ng/ $\mu$ L de fragments VIH-1/O et /M, respectivement, ont été obtenus. La réaction Gibson Assembly a été réalisée. Les cultures de bactéries après transformation chimique ont été positives en 24 heures, après avoir procédé à une étape de réisolements en 60 quadrants. Un screening par PCR a été réalisé sur ces 60 colonies. Pour uniquement l'une d'entre elles (colonie n°11), les deux PCRs de vérification Vpr OM et pUC-O étaient positives. Leur séquençage confirmait que la construction plasmidique obtenue était identique au profil attendu (TABLEAU 7).

TABLEAU 7 : Vérifications réalisées sur les CMIC pVIH-1/OM et pVIH-1/MO générés avec la méthode Gibson Assembly

		Transformation chimique			PCR de vérification				CMIC candidat
		n bactéries testées	n colonies positives	numéro colonie	Vpr	pUC	TI	C2V3	
<b>pVIH-1/OM</b>	pChimR100	60	1	n°11	+ <sup>#</sup> Vpr OM	+ pUC-O	NR*	NR	Oui
	pChimR101	60	2	n°17	+ Vpr MO	+ O-pUC	NR	NR	Oui
				n°19	+ Vpr MO	+ O-pUC	NR	NR	Oui
<b>pVIH-1/MO</b>	pChimR102	80	1	n°57	+ Vpr MO	-	-	-	Non
	pChimR103	80	2	n°20	+ Vpr MO	+ O-pUC	-	-	Oui
				n°27	+ Vpr MO	+ O-pUC	+ TI M	+ C2V3 O	Oui
	pChimR104	80	2	n°1	+ Vpr MO	+ O-pUC	NR	NR	Non
				n°80	+ Vpr MO	+ O-pUC	+ TI M	+ C2V3 O	Oui

#PCR positive confirmée par séquençage

\*PCR de vérification non réalisée

Plusieurs MiniPREP ont été réalisées à partir de cette colonie, afin d'obtenir une grande quantité de plasmide (entre 75 et 115 ng/ $\mu$ L).

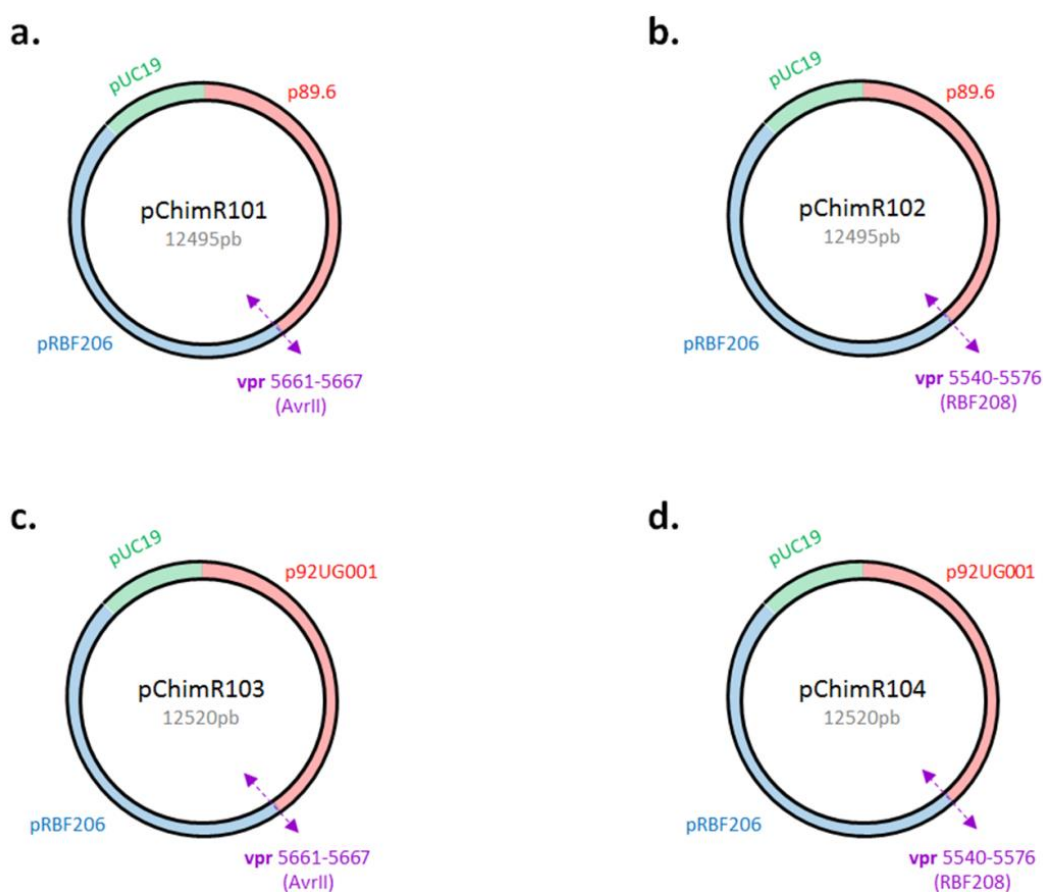
**Synthèse par Genecust.** La société Genecust a synthétisé la construction pVIH-1/OM-GC avec un plasmide vecteur, pET-28b(+), auquel des sites de restriction avaient été ajoutés. L'insert correspondait au même génome mosaïque VIH-1/OM que celui de la construction pChimR100, avec le point de cassure dans le site de restriction *AvrII* contenu dans la région Vpr (TABLEAU 7). Après extraction et purification du CMIC par MiniPrep, 89 ng/ $\mu$ L ont été obtenus. La digestion du produit obtenu a permis l'obtention des deux fragments linéaires attendus.

#### b) Génération des CMIC pVIH-1/MO

**Méthode Gibson Assembly.** Plusieurs tentatives ont été menées pour construire le motif inverse MO (FIGURE 30).

La première, pChimR101, était basée sur la construction strictement inverse à celle du CMIC pChimR100, de motif OM, avec le point de cassure dans la région Vpr localisé au niveau du site de restriction *AvrII* et le CMI parental de VIH-1/M de sous-type B (p89.6). Du fait de difficultés à l'obtenir, une nouvelle construction, pChimR103, a été testée, en remplaçant le CMI parental de VIH-1/M de sous-type B par un CMI parental de VIH-1/M de sous-type D (p92UG0001) (TABLEAU 7, FIGURE 30).

Parallèlement à cela, la localisation du point de cassure dans Vpr a été modifiée. L'hypothèse d'un impact négatif de la localisation du point de cassure dans *AvrII* nous a conduit à privilégier un point de cassure caractérisé *in vivo*, sur un virus recombinant répliatif. Deux autres constructions ont ainsi été testées : pChimR102, basée sur la localisation du point de cassure dans Vpr déterminée chez RBF208 avec le CMI parental de VIH-1/M de sous-type B (p89.6) et pChimR104, basée sur la même localisation du point de cassure dans Vpr mais avec le CMI parental de VIH-1/M de sous-type D (p92UG0001) (TABLEAU 7, FIGURE 30).



**FIGURE 30** : Profils des CMIC pVIH-1/MO

Les portions de génome provenant des CMI parentaux de VIH-1/O pRBF206 et de VIH-1/M figurent en bleu et rouge, respectivement, et celles provenant du plasmide vecteur pUC19 figurent en vert.

(a) Le CMIC pChimR101 implique le CMI parental de VIH-1/M de sous-type B p89.6 et son point de cassure correspond au site de restriction *AvrII*. (b) Le CMIC pChimR102 implique le CMI parental de VIH-1/M de sous-type B p89.6 et son point de cassure reproduit celui du recombinant RBF208. (c) Le CMIC pChimR103 implique le CMI parental de VIH-1/M de sous-type D p92UG001 et son point de cassure correspond au site de restriction *AvrII*. (d) Le CMIC pChimR104 implique le CMI parental de VIH-1/M de sous-type D p92UG001 et son point de cassure reproduit celui du recombinant RBF208.

### Point de cassure défini par le site de restriction *AvrII* (pChimR101 et pChimR103).

Pour pChimR101, l'étape d'amplification par PCR à partir des plasmides parentaux a permis de générer des fragments VIH-1/M et VIH-1/O de taille attendue (5684 et 4176pb, respectivement) avec des concentrations obtenues après purification de 236 et 325 ng/ $\mu$ L, respectivement. La réaction Gibson Assembly a été réalisée et après transformation chimique, 60 réisolements en quadrants ont été réalisés, avec pour chacun un lysat associé. Pour deux des 60 colonies (n°17 et 19), les deux PCRs de vérification Vpr MO et O-pUC ont généré un amplicon visible après migration sur gel d'agarose et de la taille attendue. Les séquences obtenues étaient elles aussi identiques à celles attendues (TABLEAU 7). Des MiniPrep ont alors



été réalisées, permettant d'obtenir 121 ng/μL d'ADN plasmidique pour la colonie n°17 et 134 ng/μL pour la colonie n°19.

Concernant pChimR103, l'étape d'amplification par PCR à partir des plasmides parentaux a permis de générer des fragments VIH-1/M et VIH-1/O de la taille attendue (5686 et 4153pb, respectivement), avec des concentrations obtenues après purification de 599 et 489 ng/μL, respectivement. La réaction Gibson Assembly a été réalisée et après transformation chimique, 80 réisolements en quadrants ont été réalisés, avec pour chacun un lysat associé. Pour les lysats de deux colonies (n°20 et 27), les PCRs de vérification Vpr MO et O-pUC étaient positives et les séquences obtenues après purification des amplicons étaient identiques à celles attendues. Les MiniPrep ont permis d'obtenir 104 ng/μL d'ADN plasmidique pour la colonie n°20 et 99 ng/μL pour la colonie n°27. Sur ces MiniPrep, les quatre PCRs de vérification (Vpr MO, O-pUC, TI M et C2V3 O) ont été réalisées. Elles étaient toutes positives pour la colonie n°27, mais négatives pour la colonie n°20 (TABLEAU 7). Ces deux constructions, pChimR103-20 et pChimR103-27, ont été retenues pour la suite des travaux.

#### **Point de cassure caractérisé *in vivo* (pChimR102 et pChimR104).**

Pour pChimR102, l'étape d'amplification par PCR à partir des plasmides parentaux a permis de générer des fragments de VIH-1/M et de VIH-1/O de la taille attendue (5591 et 4271pb, respectivement), avec des concentrations obtenues après purification de 98 et 272 ng/μL, respectivement. Après, la réaction Gibson Assembly et la transformation chimique des bactéries, un réisolement en quadrant de 80 colonies a été réalisé. Les deux PCRs de vérification Vpr MO et O-pUC ont été effectuées. Pour l'une des 80 colonies (n°57), seule la PCR Vpr MO a généré un amplicon visible sur gel. Une MiniPrep a été réalisée, sur laquelle les mêmes PCRs, ainsi que la PCR supplémentaire C2V3 O ont été réalisées. A nouveau, seule la PCR Vpr MO donnait une bande visible sur gel, de la taille attendue (TABLEAU 7). Ce CMIC n'a donc pas été retenu.

Concernant pChimR104, l'étape d'amplification par PCR à partir des plasmides parentaux a permis de générer des fragments VIH-1/M et VIH-1/O de la taille attendue (5594 et 4245pb, respectivement), avec des concentrations obtenues après purification de 601 et 560 ng/μL, respectivement. La réaction Gibson Assembly et la transformation chimique des bactéries ont été réalisées, et ont permis le réisolement en quadrant de 80 colonies. Pour les lysats de deux colonies (n°1 et n°80), les PCRs de vérification Vpr MO et O-pUC étaient positives et les séquences obtenues après purification des amplicons étaient identiques à celles

attendues. L'extraction et la purification de l'ADN plasmidique par MiniPrep a été réalisée uniquement à partir de la colonie n°80, l'intensité des bandes visibles sur gel étant très faible pour la colonie n°1. La concentration en ADN de cet extrait était de 154 ng/μL. Les quatre PCR de vérification (Vpr MO, O-pUC, TI M et C2V3 O) étaient toutes positives (TABLEAU 7). Cette construction pChimR104-80 a donc été retenue.

**Synthèse par Genecust.** Malgré plusieurs tentatives, la société Genecust n'est pas parvenue à synthétiser le profil inverse (pVIH-1/MO-GC).

## II-2. Production de virus recombinants à partir des clones moléculaires infectieux chimériques

### I-2.1. pVIH-1/OM

#### a) pChimR100

La quantification de l'AgP24 dans les surnageants, après la transfection des cellules HEK293T suivie de la co-culture avec les cellules CEMx174, était négative à J4, J7, J10 et J14 de co-culture. De même, la quantification de l'ARN dans le surnageant de co-culture à J14 était indétectable. La génération de virus, à partir du CMIC pChimR100, n'a donc pas fonctionné.

#### b) pVIH-1/OM-GC

La quantification de l'AgP24 dans les surnageants, après la transfection des cellules HEK293T suivie de la co-culture avec les cellules CEMx174, était négative à J4 mais positive à J7, J10 et J14 de co-culture, avec un pic de production d'AgP24 à J10. A cette date, la charge virale a été quantifiée à 7,28 Log<sub>10</sub> cp/mL, et l'activité de la TI s'élevait à 4433 pg/mL (TABLEAU 8). Ces résultats étaient en faveur de la génération d'une progénie virale viable et réplcative. La vérification du profil recombinant de cette progénie virale a été réalisée avec le séquençage des PCRs PR O, vpr OM et C2V3 M, puis du génome complet. Les résultats obtenus ont confirmé la présence du point de cassure attendu dans la région Vpr (TABLEAU 8).

TABLEAU 8 : Contrôles sur la progénie virale VIH-1/OM-GC, à J10 de co-culture

	<b>AgP24</b> (Log <sub>10</sub> pg/mL)	<b>TCID50</b> (UI/mL)	<b>CV</b> (Log <sub>10</sub> cp/mL)	<b>PR</b> <b>O</b>	<b>Vpr</b> <b>OM</b>	<b>C2V3</b> <b>M</b>
<b>VIH-1/OM</b>	3,7	4433	7,3	+	+	+

#PCR positive confirmée par séquençage

## I-2.2. pVIH-1/MO

La multitude de CMIC pVIH-1/MO générée par la méthode Gibson Assembly a conduit à réaliser uniquement l'étape de transfection sur cellules HEK293T, sans co-culture. L'objectif était de contrôler les profils des virus et leur infectiosité directement en sortie de transfection, et non après plusieurs semaines d'amplification par co-culture, ce qui aurait été, expérimentalement, très lourd.

### a) Point de cassure défini par le site de restriction *AvrII*

Les surnageants issus de la transfection des CMIC pChimR101-17, pChimR101-19, pChimR103-20 et pChimR103-27 ont été récupérés au terme de 72 heures d'incubation. La mesure de l'AgP24 était positive, au-delà de 400 pg/mL, pour les souches ChimR101-17 et ChimR103-20 (TABLEAU 9). En revanche, elle était faiblement positive (72,4 ng/mL) pour ChimR103-27 et équivoque (4,6 ng/mL), considérée comme négative, pour ChimR101-19. L'infectiosité des progénies virales a alors été contrôlée par mesure de la TCID50 sur TZM-bl, mais s'est avérée négative pour les trois souches testées (ChimR101-17, ChimR103-20 et ChimR103-27) (TABLEAU 9). Les trois CMIC candidats correspondants étaient donc capables de générer des virus recombinants, mais étaient non infectieux.

TABLEAU 9 : Contrôles sur les progénies virales VIH-1/MO, produites en fin de transfection

<b>VIH-1/MO</b>	<b>Sous-type</b> <b>VIH-1/M</b>	<b>AgP24</b> (pg/mL)	<b>TCID50</b> (UI/mL)
<b>ChimR101-17</b>	<b>B</b>	> 400	-
<b>ChimR101-19</b>	<b>B</b>	équivoque	-
<b>ChimR103-20</b>	<b>D</b>	> 400	-
<b>ChimR103-27</b>	<b>D</b>	72,4	-

b) Point de cassure caractérisé *in vivo*

Selon le même procédé, les surnageants issus de la transfection du CMIC pChimR104-80 ont été récupérés après 72 heures d'incubation. La mesure de l'AgP24 était positive, supérieure à 400 pg/mL. En revanche, la TCID50 sur TZM-bl était négative. Comme pour les constructions précédentes, le CMIC candidat pChimR104-80 semblait capable de générer des virus recombinants, mais non infectieux (TABLEAU 9).

I-2.3. Bilan de la production de virus recombinants

Un résumé des différentes stratégies de génération des CMIC et des candidats obtenus est présenté dans la FIGURE 31.

Au terme de ces nombreuses expérimentations, seule la construction pVIH-1/OM-GC, synthétisée, a permis de produire des virus recombinants VIH-1/OM infectieux. En revanche, aucun virus recombinant VIH-1/MO infectieux n'a été obtenu (FIGURE 31).

Méthode	pVIH-1/OM		pVIH-1/MO				Synthèse (Genecust)
	Gibson Assembly	Synthèse (Genecust)	Gibson Assembly		<i>In vivo</i> (RBF208)	Site restriction	
Point de cassure	Site restriction	Site restriction	Site restriction		<i>In vivo</i> (RBF208)	Site restriction	
Diversité VIH-1/M	B (89.6)	B (89.6)	B (89.6)	D (92UG0001)	B (89.6)	D (92UG0001)	B (89.6)
CMIC							
Candidats ?	✓ pChimR100	✓ pVIH-1/OM-GC	✓ pChimR101 17 19	✓ pChimR103 20 27	✗	✓ pChimR104 80	
AgP24	-	+	+ -	+ +		+	
TCID50	-	+	- -	- -		-	
Conclusion	✗ Pas de virus recombinant	✓ Virus rec. infectieux	✗ Virus non infectieux	✗ Virus non infectieux	✗ Pas de CMIC	✗ Virus non infectieux	✗ Echec synthèse

FIGURE 31 : Synthèse générale de la génération de CMIC et de la production de virus recombinants

### III- Discussion

Les objectifs de ce travail étaient de générer des clones moléculaires infectieux chimériques, afin d'évaluer les propriétés répliatives *in vitro* de virus recombinants inter-groupes M et O.

Plusieurs stratégies ont été réalisées successivement pour générer de tels CMIC : la génération d'un génome recombinant par un système de PCRs successives, avec des étapes intermédiaires d'hybridation entre les différents fragments, le recours à des digestions enzymatiques suivies d'une étape de ligation, en variant le nombre de fragments devant être ligués (résultats Master 2R non montrés), mais de nombreux obstacles n'ont pas permis d'avoir de CMIC candidats. Dans ce travail, la méthode Gibson Assembly a été développée et a permis de générer 6 CMIC candidats, dont un CMIC de profil OM et cinq de profil MO.

Afin d'optimiser l'amplification des différentes constructions générées, le plasmide vecteur retenu était pUC19, un plasmide de référence « *high copy* », c'est-à-dire présent en un très grand nombre de copies dans les bactéries transformées qui l'hébergent. La transformation chimique a été préférée à l'électroporation, qui est connue pour modifier plus fréquemment les plasmides à amplifier lors de la culture, avec des phénomènes d'excisions, d'insertions, et de mutations. De plus, lors de cette étape de culture bactérienne, la probabilité pour qu'une colonie contienne le plasmide d'intérêt est jugée faible, comme constaté avec un réisolement suivi d'une lyse de 60 à 80 colonies afin d'obtenir un voire deux CMIC candidats. Cette étape de screening semble donc déterminante, et notre stratégie reste encore à optimiser.

Lors de la modélisation des CMIC à générer, le choix des profils de recombinaison a reposé sur les deux profils majoritaires ayant émergé parmi 11 des 12 URF\_MO complètement séquencées actuellement décrites : [O<sub>gag/pol</sub>-M<sub>env</sub>] et [M<sub>gag/pol</sub>-O<sub>env</sub>]. Compte tenu de cette dichotomie et de la description d'un point chaud de la recombinaison dans les gènes accessoires centraux et dans le gène *vpr* en particulier, le point de cassure des CMIC a été positionné dans cette région du génome et en testant plusieurs approches et CMI parentaux. Malgré les différentes constructions et candidats obtenus, les vérifications menées sur les CMIC candidats ont mis en évidence qu'aucun virus recombinant infectieux n'avait pu être généré.

Pour quatre d'entre eux (pChimR101-17, pChimR103-20 pChimR103-27, et pChimR104-80), la transfection a permis de générer des virus recombinants, avec de l'AgP24 quantifiable dans les surnageants de culture, mais ces virus étaient non infectieux, la TCID50 étant négative. Pour les deux autres (pChimR100 et pChimR101-19), la transfection n'a pas permis de produire d'AgP24. Ainsi, malgré la multitude de vérifications menées sur les CMIC candidats, leur capacité à produire des virus viables et infectieux n'était donc pas assurée.

L'hypothèse d'un mauvais positionnement du point de cassure s'est d'abord posée, lorsque le premier candidat pVIH-1/OM a mené à la génération de virus recombinants non infectieux. En effet, la première approche s'était appuyée sur la présence de régions conservées dans le gène *vpr*, correspondant à des sites de restriction et pouvant constituer une zone propice à un saut de matrice de la TI, lors d'une infection naturelle par le VIH. Plusieurs approches alternatives, prenant en compte cette localisation du point de cassure, ont donc été envisagées. Plusieurs interactions fonctionnelles entre des protéines du VIH ont été mises en évidence et sont apparues comme jouant un rôle important dans leur fonctionnalité. C'est notamment le cas entre la gp120 de l'enveloppe et la protéine Nef<sup>f</sup> [22, 37, 220] ou entre la protéine p6 de la région Gag et les protéines Vpr et Vpx [221]. De la même façon, plusieurs études ont identifié la présence de réseaux de co-évolution, avec des interactions entre différents domaines des protéines du VIH, tant au niveau intra- que inter-protéique [222-224]. De telles co-évolutions, impactant la fonctionnalité des protéines impliquées, ont ainsi été décrites entre la protéine p17 et la gp41 [222] et entre la TI et l'IN [223, 224]. L'introduction artificielle de points de cassure dans une zone inadaptée et non retrouvés dans les génomes de souches recombinantes *in vivo* pourrait limiter l'infectiosité et la répliativité des souches générées. Nous avons donc choisi de déplacer le point de cassure dans *vpr* d'une centaine de pb pour le positionner à l'identique du point de cassure d'un recombinant isolé *in vivo*, RBF208 et la méthode Gibson Assembly a été de nouveau appliquée. Dans ces deux approches, un seul point de cassure (dans Vpr) était présent dans la construction. Les LTRs étant différents, cela nécessitait, lors des premiers cycles répliatifs, un phénomène de recombinaison obligatoire au sein de ces régions conservées, pour les rendre identiques. Afin de s'affranchir de cette contrainte et de vérifier si cet élément empêchait la génération de virus recombinants répliatifs et infectieux, nous avons introduit deux points de cassure, l'un dans Vpr et l'autre dans les LTRs, afin de reproduire deux formes décrites *in vivo* : RBF208 et RBF237 (résultats non montrés). Les CMIC correspondants

ont été produits par la méthode Gibson Assembly, mais les étapes suivantes du protocole n'ont pas mieux fonctionné qu'avec les approches précédentes, ne permettant de conclure quant à un possible impact du positionnement du point de cassure.

Une autre explication de nos résultats pouvait résider dans le choix des plasmides parentaux utilisés pour générer les CMIC. Un seul clone de VIH-1/O, produit à partir d'isolats de VIH-1/O de sous-groupe H, avait été publié avant nos travaux [225]. La séquence insérée dans le plasmide ne correspondait pas à un clone viral présent dans une quasi-espèce initiale, mais correspondait à une construction artificielle chimérique de deux isolats distincts. Ce plasmide ne pouvait donc être utilisé dans notre contexte d'étude, basé sur la comparaison de souches décrites *in vivo*. Dans le cadre d'études des propriétés virologiques des VIH-1/O [172, 212, 213], en collaboration avec une équipe nord-américaine, deux clones moléculaires infectieux, pRBF206, et pBCF120, ont été construits à partir de souches isolées dans notre laboratoire. Bien qu'appartenant au sous-groupe T minoritaire des VIH-1/O, pRBF206 a été préféré au CMI pBCF120, représentant du sous-groupe majoritaire H, car il montrait un meilleur rendement d'amplification sur bactéries. Concernant la souche parentale VIH-1/M, aucun clone moléculaire infectieux n'étant disponible au laboratoire, quatre CMI pVIH-1/M ont été sélectionnés dans la liste du CFAR. Après amplification bactérienne, et du fait d'une séquence complète du génome, le choix s'est porté dans un premier temps sur le CMI p89.6 de sous-type B. Or, l'analyse des 19 URF\_MO *in vivo* n'a montré aucune forme recombinante impliquant une souche parentale VIH-1/M de sous-type B, ce sous-type ne circulant pas au Cameroun, région d'émergence de ces formes. D'autres essais utilisant la méthode Gibson Assembly ont donc été réalisés, en ayant recours au CMI p92UG0001 [215], impliquant une souche de sous-type D. Ce sous-type est retrouvé dans plusieurs formes recombinantes *in vivo* (YBF274, RBF208 et RBF235). Ces essais n'ont finalement pas amélioré les résultats obtenus jusqu'alors.

Face aux problématiques soulevées par ces différents résultats, nous avons fait synthétiser les CMIC pVIH-1/OM et pVIH-1/MO par la société Genecust. Seul le CMIC pVIH-1/OM-GC a pu être synthétisé et a généré des virus recombinants et infectieux. Le fait que la société Genecust ne soit pas parvenue à produire le CMIC pVIH-1/MO, malgré des outils industriels de production probablement plus adaptés que les nôtres, montre bien la complexité de notre travail et nous amène à nous questionner sur la viabilité des constructions envisagées, en particulier pour le profil MO.

Mais nos résultats ouvrent ainsi de nombreuses perspectives de travail. En premier lieu, il serait judicieux de réitérer les essais avec des CMI parentaux de VIH-1/M CRF02\_AG et de VIH-1/O sous-groupe H, afin de générer un virus davantage proche, en termes de diversité, des formes recombinantes VIH-1/MO circulant *in vivo*. De plus, le CMI parental de VIH-1/O utilisé dans ces travaux impliquait une souche de VIH-1/O, RBF206, dont le comportement face aux facteurs de restriction cellulaire ne semble pas représentatif des autres variants de groupe O. Il a, en effet, été décrit pour cette souche que la protéine Vpu codée par son génome avait la capacité de contourner efficacement la Tétherine, propriété décrite uniquement pour les variants de groupe M jusqu'alors <sup>[213]</sup>. Des travaux doivent donc être menés afin d'éclaircir ces différentes hypothèses et d'étudier le réel impact de la diversité sur la viabilité des souches recombinantes générées.



## B. Etude de la cinétique de réplcation des virus parentaux et recombinant

---

### I- Méthodes

#### I-1. Production d'un stock de virus parentaux et recombinant

Afin de couvrir l'intégralité des manipulations de mise au point du protocole et de comparaison des souches entre elles, et afin de ne pas introduire de biais lié à plusieurs manipulations consécutives, nous avons produit un stock de souches parentales et recombinante lors d'une seule et même manipulation.

La première étape a consisté en des transfections parallèles de cellules HEK293T par les CMI parentaux, pVIH-1/O (pRBF206) et pVIH-1/M (p89.6), et du CMIC, pVIH-1/OM, synthétisé par Genecust (GC), dans les mêmes conditions et selon la méthodologie déjà décrite (I-2.

Production de virus recombinants à partir de clones moléculaires infectieux) (FIGURE 28, P.92).

La seconde étape a consisté en une co-culture. Les premiers essais de co-culture avaient été réalisés sur lignée CEMx174 (CD4+, CXCR4+), mais compte tenu du tropisme des souches utilisées (CCR5 pour RBF206, CCR5/CXCR4 pour 89.6 et le recombinant produit), la lignée Jurkat X4/R5 a finalement été utilisée pour l'étape d'amplification par co-culture, selon le même protocole que celui précédemment décrit pour la lignée CEMx174. Cette lignée Jurkat X4/R5, obtenue du NIBSC, est une lignée de lymphocytes T CD4+ exprimant de manière endogène le co-récepteur CXCR4 et ayant été transfectée pour exprimer également de manière stable le co-récepteur CCR5 humain. Avant co-culture, les Jurkat non infectées ont été maintenues en culture en milieu RPMI supplémenté par 10% de SVFd, de la gentamicine (50µg/mL) et de la généticine G418 (500µg/mL). Le milieu a été remplacé et les cellules divisées au 1/5<sup>ème</sup> deux fois par semaine.

A partir du quatrième jour de co-culture (J4), 4mL de surnageant de chacune des trois co-cultures (VIH-1/M, VIH-1/O et VIH-1/OM) étaient prélevés deux fois par semaine et congelés. Parmi ces 4mL, 300µL étaient utilisés instantanément pour une quantification de l'AgP24 par VIDAS®, permettant le suivi de la production virale au cours du temps. Afin de conserver le même volume de culture, 4mL de milieu neuf (RPMI + 10% SVFd + Gentamicine + Généticine) étaient ajoutés à chaque co-culture.

Les trois cultures ont été arrêtées à J43. Pour chacune des souches, les surnageants pour lesquels la quantité d'AgP24 était supérieure à 3Log<sub>10</sub> pg/mL ont été décongelés et réunis, afin d'obtenir une concentration importante de virions. Pour caractériser les trois mélanges obtenus, plusieurs analyses ont été réalisées :

- une mesure de l'AgP24 avec le kit VIDAS® HIV P24 II CONFIRMATION (bioMérieux),
- une TCID50 sur cellules TZM-bl, selon la méthode de Montefiori.

Par ailleurs, afin d'exclure toute contamination entre les différentes flasques de la culture, des PCRs nichées amplifiant la région Vpr, où se situe le point de cassure, ont été réalisées sur l'ARN génomique, extrait à partir des surnageants de culture avec l'automate EZ1 advanced XL et le kit EZ1® DSP Virus (Qiagen). Elles mettaient en jeu des couples d'amorces conçus pour amplifier des gènes *vpr* homologues (Vpr MM et Vpr OO n'amplifiant que les souches parentales M et O, respectivement) et hétérologues (Vpr OM n'amplifiant que la souche recombinante OM et Vpr MO n'amplifiant aucune des souches en culture, et considéré comme un témoin négatif) (TABLEAU 10). La nature des amplicons obtenus a été vérifiée par séquençage Sanger.

Une fois caractérisés, les mélanges ont ensuite été aliquotés par volumes de 500µL et congelés à -80°C.

**TABLEAU 10** : Amorces utilisées lors la caractérisation des régions Vpr des virus produits

PCR	Amorce sens 5'-3'	Amorce antisens 5'-3'	Taille (pb)
Vpr MM	142U GGGGTCTGCATACAGGAGAA	148L TCCAGGGCTCTAGKCTAGGA	615
Vpr OO	288U ACACAGATTGAYCCTGAAAC	291L GCCAAGGGGGCAYCTCAGGA	550
Vpr OM	288U ACACAGATTGAYCCTGAAAC	148L TCCAGGGCTCTAGKCTAGGA	538
Vpr MO	142U GGGGTCTGCATACAGGAGAA	291L GCCAAGGGGGCAYCTCAGGA	627

## I-2. Etude de la cinétique de répliation des virus parentaux et recombinant

La cinétique de répliation des souches parentales et recombinante a été réalisée en triplicat sur la lignée Jurkat X4R5. Pour cela, 5 millions de Jurkat ont été infectées par chacune des souches à une MOI de 0,01 UI/cellule (déterminée à partir de la valeur de TCID<sub>50</sub>) et dans 5mL de milieu (RPMI + 10%SVF<sub>d</sub> + Gentamicine + Généticine). Entre le premier jour (J1) et le dernier jour (J30) de culture, 400µL de surnageant ont été prélevés puis congelés 6 jours sur 7 et remplacés par 400µL de milieu neuf, afin de rester à un même volume tout au long de la culture.

Les cinétiques ont été mesurées sur la base de l'activité de la TI (*Lenti RT activity, Cavid*) et la quantité d'AgP24 (*INNOTEST® HIV Antigen mAb, Fujirebio, Belgique*) de chacun des surnageants prélevés, pour chacun des triplicats de chacune des souches, en rétrospectif. Les moyennes des valeurs d'activité de la TI et de la quantité d'AgP24 de chaque triplicat ont été calculées (en Log<sub>10</sub> pg/mL) pour chacun des points de la cinétique et le pic a été défini pour chacune des cinétiques.

Afin de mettre en évidence une éventuelle différence de répliation entre la souche recombinante et les souches parentales, trois paramètres ont été pris en compte : la vitesse d'atteinte du pic (en jours), l'intensité de ce pic (en Log<sub>10</sub> pg/mL) et la vitesse de croissance (en Log<sub>10</sub> /jour). La vitesse de croissance a été définie comme étant la pente de la droite reliant le point suivant le point le plus bas, reflétant le début de la répliation, et le point le plus haut, correspondant au pic de répliation.

## II- Résultats

### II-1. Production du stock de virus parentaux et recombinant

Après transfection de cellules HEK293T par les CMI parentaux, pVIH-1/O et pVIH-1/M, et chimérique, pVIH-1/OM, puis co-culture avec les cellules Jurkat X4/R5, trois pools de surnageants, dont l'AgP24 dépassait les 3 Log<sub>10</sub> pg/mL, ont été obtenus.

Les valeurs d'AgP24 étaient comprises entre 4,5 et 5,9 Log<sub>10</sub> pg/mL et les valeurs de charge virale, comprises entre 8,8 et 9,3 Log<sub>10</sub> cp/mL (TABLEAU 11). Ces valeurs témoignaient d'une production virale très importante, pour les trois souches considérées (VIH-1/M, VIH-1/O et VIH-1/OM). Les valeurs de TCID<sub>50</sub> étaient toutes fortement positives, entre 91 376 et 267 184 UI/mL, reflétant une bonne infectiosité des souches produites. Ces valeurs ne pouvaient cependant pas être comparées entre elles, étant donné que les conditions initiales de culture (nombres de cellules Jurkat X4/R5 infectées, quantité de virus issue de la transfection des cellules HEK293T) avaient varié selon les souches.

**TABLEAU 11** : Caractérisation des mélanges de virus parentaux et recombinant

	AgP24*	TCID <sub>50</sub> <sup>‡</sup>	CV <sup>°</sup>	PCR <i>vpr</i>				Séquençage <i>vpr</i>			
				MM	OO	OM	MO	MM	OO	OM	MO
<b>VIH-1/M</b>	5,9	267184	9,3	+	-	-	-	MM			
<b>VIH-1/O</b>	4,5	226527	8,8	-	+	-	-		OO		
<b>VIH-1/OM</b>	4,5	91376	9,2	-	-	+	-			OM	

\*Log<sub>10</sub> pg/mL

<sup>‡</sup>UI/mL

<sup>°</sup>Charge virale en Log<sub>10</sub> cp/mL

Concernant les PCRs ciblant la région Vpr, les résultats ont confirmé l'absence de contamination. En effet, seuls les motifs Vpr MM, Vpr OO et Vpr OM ont été retrouvés pour les souches parentales (VIH-1/M et VIH-1/O) et recombinante (VIH-1/OM), respectivement. Le motif *vpr* MO n'a été retrouvé pour aucun des surnageants de co-culture, comme attendu (TABLEAU 11).

## II-2. Cinétique de réplcation des virus parentaux et recombinant

### II-2.1. Profil des cinétiques de réplcation des virus parentaux et recombinant

Les cinétiques de réplcation des souches parentales (VIH-1/M et VIH-1/O) et recombinante (VIH-1/OM) ont été réalisées en triplicat et en quantifiant deux marqueurs : l'activité de la TI et la quantité d'AgP24. Pour chacun de ces deux marqueurs et pour chaque triplicat, le délai d'apparition et l'intensité du pic ont été relevés.

Les cultures ont été suivies de manière très rapprochée, avec des prélèvements quotidiens, ce qui a permis d'éviter tout biais dans la détermination de ce délai d'apparition des pics.

#### a) VIH-1/M

Les allures des courbes représentant les cinétiques de l'activité TI et de la quantité d'AgP24 pour les surnageants de culture des trois réplcats de VIH-1/M sont superposables. Globalement, la multiplication virale de la souche parentale VIH-1/M a été rapide, avec une augmentation exponentielle dès le J3 de culture pour l'activité de la TI et dès le J4 pour la quantité d'AgP24. Le pic d'activité de la TI a été atteint entre J12 et J15 et était compris entre 5,1 et 5,4 Log<sub>10</sub> pg/mL. Concernant la quantité d'AgP24, le pic, atteint entre J11 et J12, était d'une valeur de 6,7 Log<sub>10</sub> pg/mL pour les trois réplcats (TABLEAU 12, FIGURE 32).

#### b) VIH-1/O

Les allures des courbes représentant les cinétiques de l'activité TI et de la quantité d'AgP24 pour les surnageants de culture des trois réplcats de VIH-1/O sont également superposables. La courbe de l'activité de la TI a montré d'abord une phase de décroissance, qui n'était pas observée pour la quantité d'AgP24. En revanche, les deux marqueurs ont augmenté de façon exponentielle à partir de J9. Le pic d'activité de la TI a été atteint entre J17 et J18 et était compris entre 4,5 et 4,9 Log<sub>10</sub> pg/mL. Concernant la quantité d'AgP24, le pic était atteint un peu plus tardivement, entre J22 et J25, avec une valeur comprise entre 6,2 et 6,3 Log<sub>10</sub> pg/mL (TABLEAU 12, FIGURE 33).

## c) VIH-1/OM

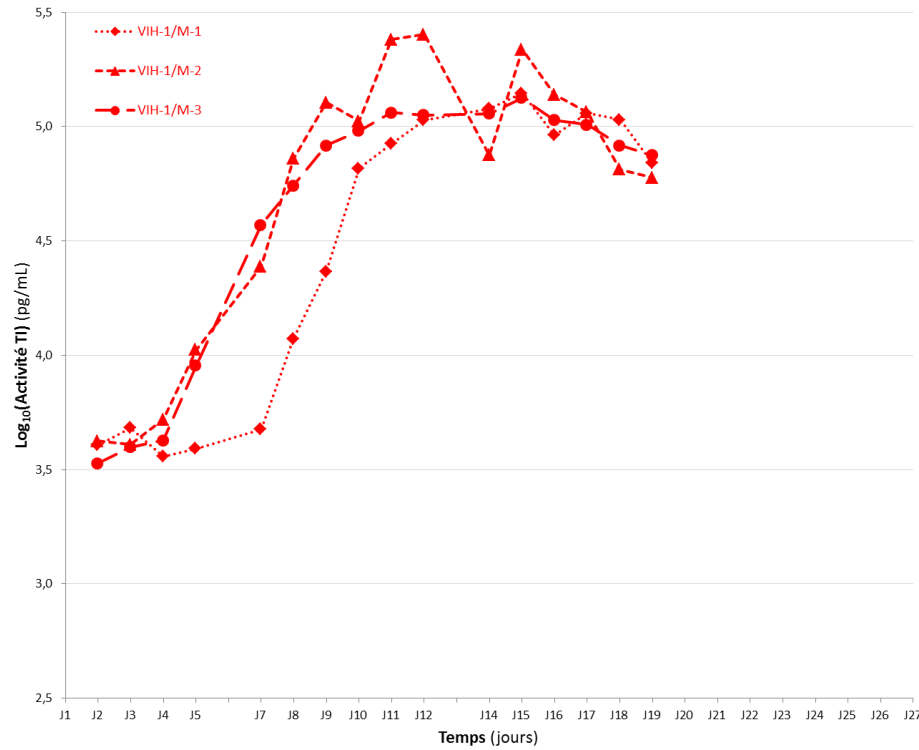
Les allures des courbes représentant les cinétiques de l'activité TI et de la quantité d'AgP24 pour les surnageants de culture des trois réplicats de VIH-1/OM sont elles aussi superposables. La multiplication virale de la souche recombinante a montré une augmentation exponentielle dès J8 pour les deux marqueurs (TI et AgP24). L'activité de la TI a montré un pic compris entre 4,4 et 5,2 Log<sub>10</sub> pg/mL, atteint entre J14 et J19. Le pic de la quantité d'AgP24 est apparu entre J14 et J19, avec une valeur entre 6,2 et 6,5 Log<sub>10</sub> pg/mL (TABLEAU 12, FIGURE 34).

TABLEAU 12 : Caractéristiques des cinétiques de répliation de chacune des souches étudiées

	ACTIVITE TI		AgP24	
	Délai d'apparition du pic	Intensité du pic	Délai d'apparition du pic	Intensité du pic
	(jours)	Log <sub>10</sub> (Activité TI)	(jours)	Log <sub>10</sub> (AgP24)
<b>VIH-1/M-1</b>	15	5,1	12	6,7
<b>VIH-1/M-2</b>	12	5,4	12	6,7
<b>VIH-1/M-3</b>	15	5,1	11	6,7
<b>Δmax-min*</b>	<b>3</b>	<b>0,3</b>	<b>1</b>	<b>0</b>
<b>VIH-1/O-1</b>	18	4,7	22	6,2
<b>VIH-1/O-2</b>	17	4,9	25	6,3
<b>VIH-1/O-3</b>	17	4,5	24	6,2
<b>Δmax-min*</b>	<b>1</b>	<b>0,4</b>	<b>3</b>	<b>0,1</b>
<b>VIH-1/OM-1</b>	19	4,4	19	6,2
<b>VIH-1/OM-2</b>	14	5,2	14	6,5
<b>VIH-1/OM-3</b>	16	4,6	19	6,4
<b>Δmax-min*</b>	<b>5</b>	<b>0,8</b>	<b>5</b>	<b>0,3</b>

\*Différences calculées entre les valeurs maximales et minimales

a. Activité de la TI



b. Quantité d'AgP24

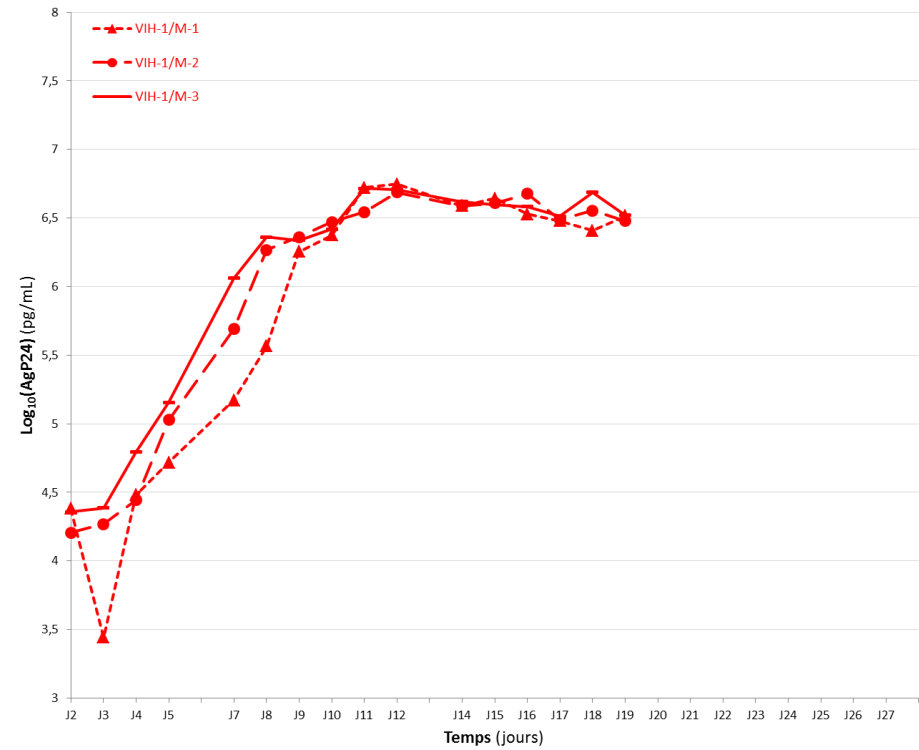
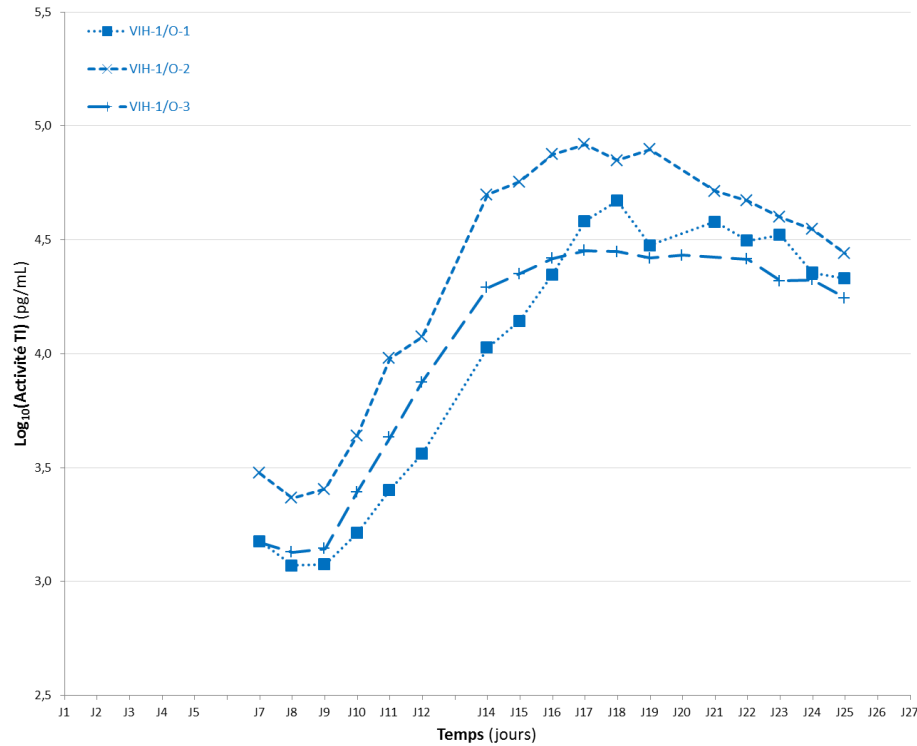


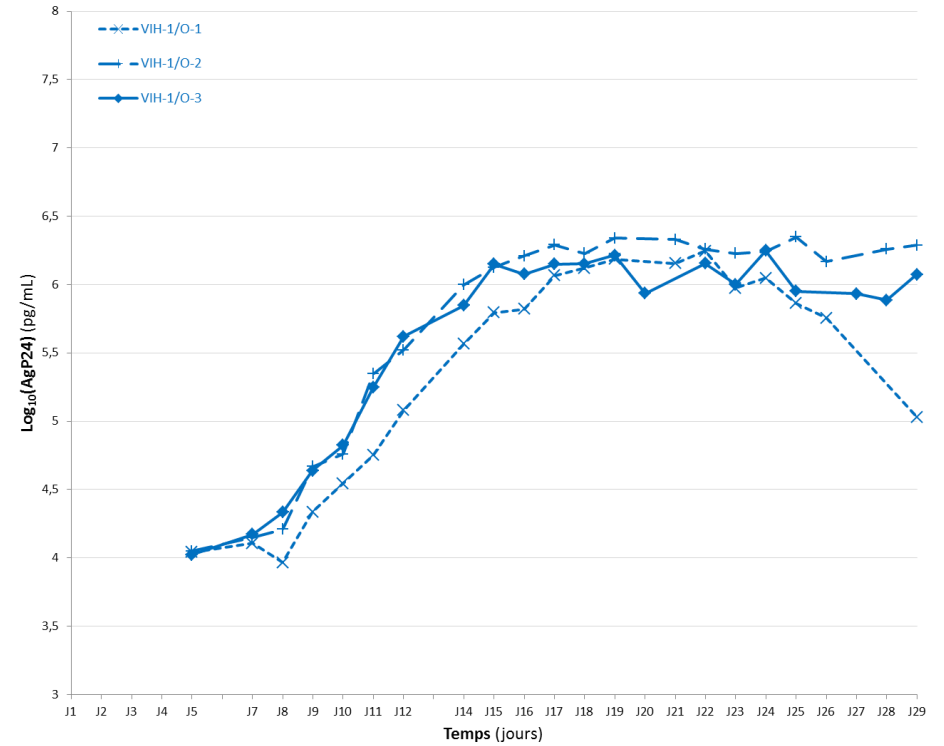
FIGURE 32 : Cinétique de réplcation des triplicats de VIH-1/M

(a) Représentation de l'évolution de l'activité de la TI, en log<sub>10</sub> pg/mL, au cours du temps, en jours, pour les trois triplicats de la souche parentale VIH-1/M. (b) Représentation de l'évolution de la quantité d'AgP24, en log<sub>10</sub> pg/mL, au cours du temps, en jours, pour les trois triplicats de la souche parentale VIH-1/M.

**a. Activité de la TI**



**b. Quantité d'AgP24**

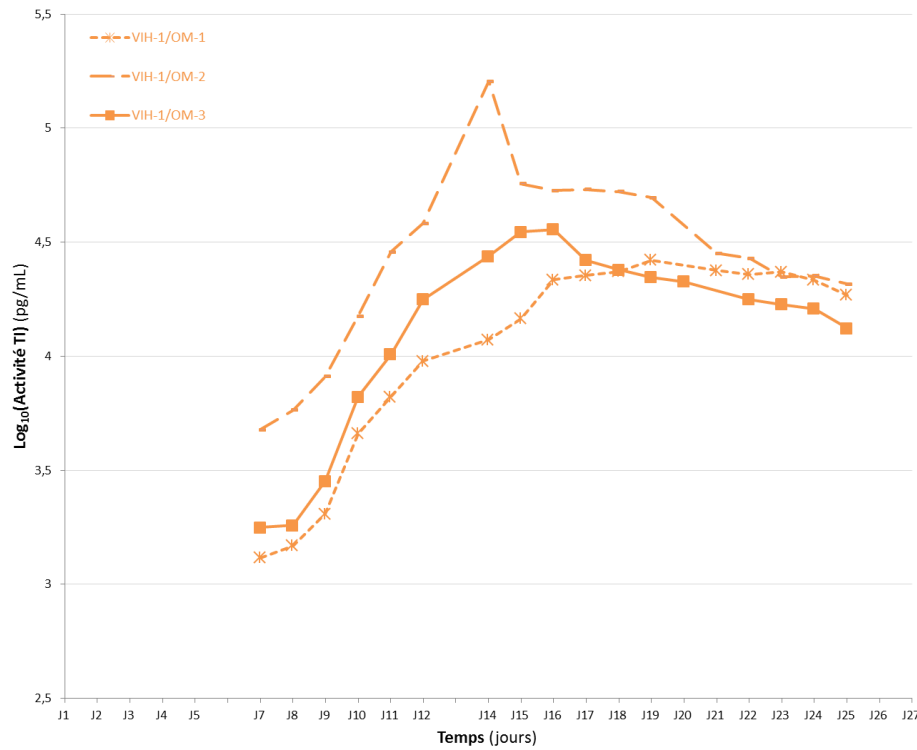


**FIGURE 33 : Cinétique de réplcation des triplicats de VIH-1/O**

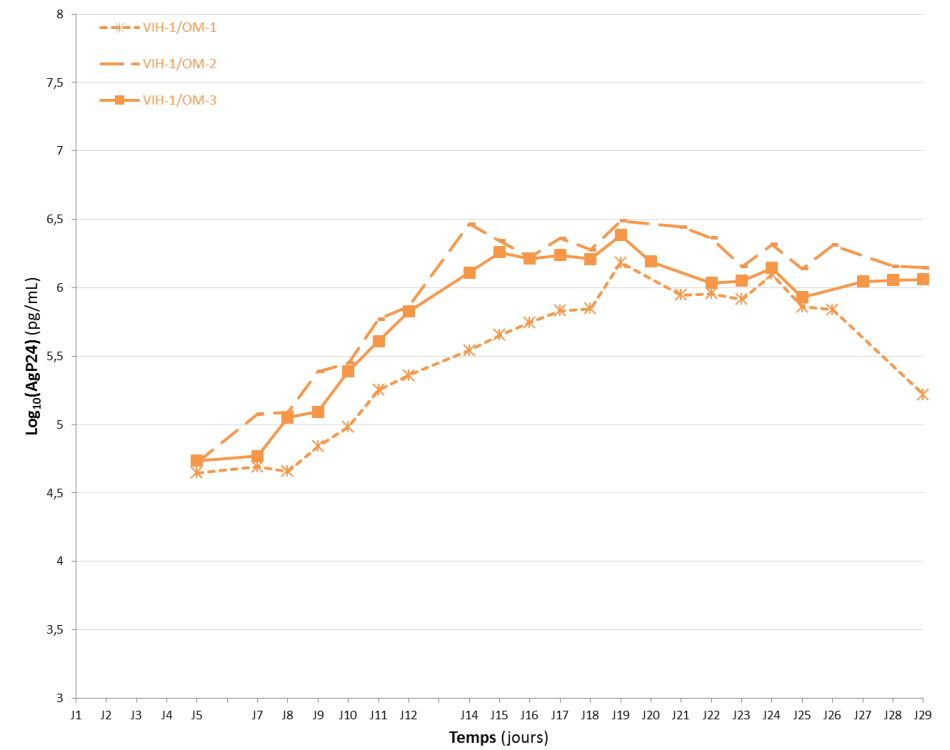
(a) Représentation de l'évolution de l'activité de la TI, en log<sub>10</sub> pg/mL, au cours du temps, en jours, pour les trois triplicats de la souche parentale VIH-1/O. (b) Représentation de l'évolution de la quantité d'AgP24, en log<sub>10</sub> pg/mL, au cours du temps, en jours, pour les trois triplicats de la souche parentale VIH-1/O.



**a. Activité de la TI**



**b. Quantité d'AgP24**



**FIGURE 34 : Cinétique de réplcation des triplicats de VIH-1/OM**

(a) Représentation de l'évolution de l'activité de la TI, en log<sub>10</sub> pg/mL, au cours du temps, en jours, pour les trois triplicats de la souche recombinante VIH-1/OM. (b) Représentation de l'évolution de la quantité d'AgP24, en log<sub>10</sub> pg/mL, au cours du temps, en jours, pour les trois triplicats de la souche recombinante VIH-1/OM.

Le triplicat de chaque souche a toutefois montré une bonne reproductibilité des cultures, avec de faibles différences calculées entre les valeurs maximales et minimales, aussi bien pour le délai d'apparition du pic que pour la production virale (TABLEAU 12).

La souche recombinante a fait apparaître une hétérogénéité inter-triplicats plus importante que pour les souches parentales, avec une différence de cinq jours pour le délai d'apparition des pics, et de 0,3 Log<sub>10</sub> et 0,8 Log<sub>10</sub> pour l'intensité des pics d'activité de la TI et de quantité d'AgP24, respectivement.

Au final, les cinétiques des deux marqueurs suivis, réalisées de manière indépendante en triplicat pour chacune des souches se sont révélées reproductibles et superposables. Quelques valeurs aberrantes ont néanmoins été relevées, dont le point VIH-1/M-2 J14, qui a nécessité de supprimer le point J14 pour toutes les souches, lors de la comparaison des moyennes.

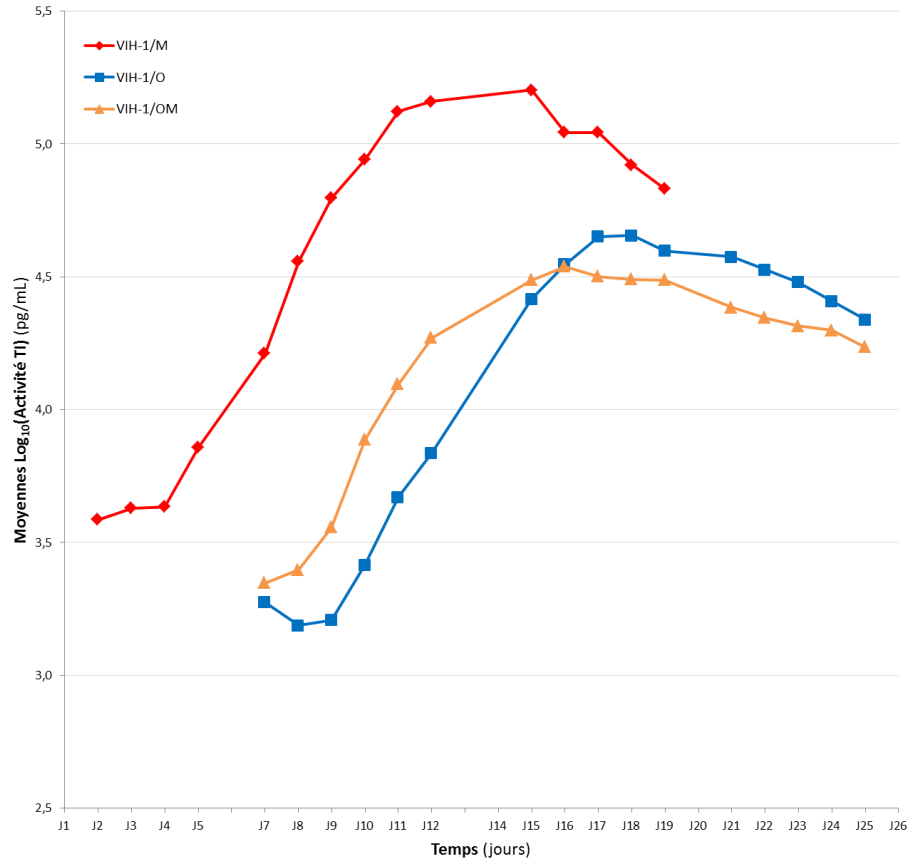
Les différentes données collectées ont été considérées comme fiables et les triplicats de chaque souche ont été pris en compte dans leur intégralité pour la suite des travaux, consistant en la comparaison des cinétiques de réplication des souches parentales et recombinante.

## II-2.2. Comparaison des cinétiques de réplication

Les cinétiques moyennes d'activité de la TI et de quantité d'AgP24 des trois souches sont représentées en FIGURE 35a et 35b, respectivement.

De manière intéressante, les résultats obtenus avec la mesure de l'activité de la TI montrent globalement une cinétique superposable pour les trois souches, avec une phase de croissance exponentielle, jusqu'à l'atteinte d'un pic, suivie d'une phase de décroissance nette. La cinétique de la quantité d'AgP24 des trois souches testées montre, quant à elle, une même phase de croissance exponentielle, atteignant un plateau, avec des valeurs oscillant autour d'une valeur maximale, sans phase de décroissance.

a. Activité de la TI



b. Quantité d'AgP24

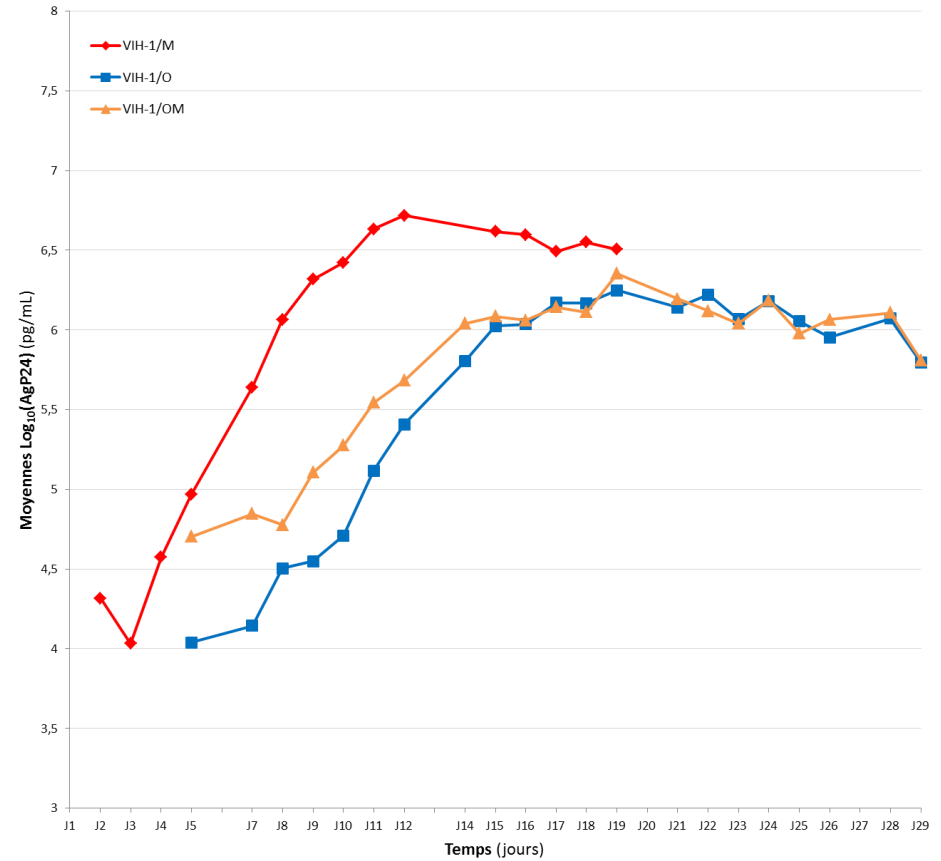


FIGURE 35 : Cinétiques moyennes de réplcation des souches parentales et recombinante

(a) Représentation de l'évolution des moyennes d'activité de la TI des trois triplicats, en  $\log_{10}$  pg/mL, au cours du temps, en jours, pour les souches parentales VIH-1/M (en rouge) et VIH-1/O (en bleu) et recombinante VIH-1/OM (en orange). (b) Représentation de l'évolution des moyennes de la quantité d'AgP24 des trois triplicats, en  $\log_{10}$  pg/mL, au cours du temps, en jours, pour les souches parentales VIH-1/M (en rouge) et VIH-1/O (en bleu) et recombinante VIH-1/OM (en orange).

L'observation de l'allure des cinétiques de l'activité de la TI et de la quantité d'AgP24 fait apparaître que la souche parentale VIH-1/M croît plus rapidement et de manière plus intense que la souche parentale VIH-1/O, mais aussi que la souche recombinante VIH-1/OM. Ces courbes semblent également révéler une cinétique de réplcation de la souche recombinante VIH-1/OM plus proche de la souche parentale VIH-1/O dont elle est issue, que de la parentale VIH-1/M (FIGURE 35).

Afin d'évaluer quantitativement les différences de réplcations entre les souches, nous avons calculé les moyennes des données aux pics d'activité de la TI et de quantité d'AgP24 des triplicats et les vitesses de croissance pour les trois souches (VIH-1/M, VIH-1/O et VIH-1/OM) ainsi que leurs différences (TABLEAU 13).

Concernant d'abord la réplcation des deux souches parentales, il apparaît nettement que la souche VIH-1/M réplque plus fortement que la souche VIH-1/O. En effet, les pics d'activité de la TI et de quantité d'AgP24 ont été atteints plus précocement pour la souche VIH-1/M que pour la souche VIH-1/O (J15 contre J18 pour l'activité de la TI et J12 contre J19, pour la quantité d'AgP24). De la même façon, l'intensité des pics était plus importante pour la souche VIH-1/M que pour la souche VIH-1/O (5,2 Log<sub>10</sub> contre 4,7 Log<sub>10</sub> pg/mL pour l'activité de la TI et 6,7 Log<sub>10</sub> contre 6,2 Log<sub>10</sub> pg/mL pour la quantité d'AgP24), avec, pour les deux marqueurs quantifiés, une différence de 0,5 Log<sub>10</sub> en faveur de la souche VIH-1/M.

Concernant l'estimation de la vitesse de croissance, les résultats étaient différents selon le marqueur étudié. En effet, la vitesse de croissance de la quantité d'AgP24 était plus rapide pour la souche parentale VIH-1/M (0,300 Log<sub>10</sub> contre 0,157 Log<sub>10</sub> pg/mL /jour pour la souche parentale VIH-1/O). L'inverse était observé pour la vitesse de croissance de l'activité de la TI (0,131 Log<sub>10</sub> contre 0,167 Log<sub>10</sub> pg/mL /jour pour les souches VIH-1/M et VIH-1/O, respectivement) (TABLEAU 13).

TABLEAU 13 : Valeurs moyennes et différences calculées pour les différents paramètres étudiés

	ACTIVITE TI				AgP24			
	Délai d'apparition du pic	Durée de répliation <sup>§</sup>	Intensité du pic	Vitesse de croissance	Délai d'apparition du pic	Durée de répliation <sup>§</sup>	Intensité du pic	Vitesse de croissance
	<i>jours</i>	<i>jours</i>	<i>Log<sub>10</sub>(Activité TI)</i>	<i>Log<sub>10</sub>(Activité TI) /jour</i>	<i>jours</i>	<i>jours</i>	<i>Log<sub>10</sub>(AgP24)</i>	<i>Log<sub>10</sub>(AgP24) /jour</i>
VIH-1/OM	16	8	4,5	0,14	19	11	6,4	0,15
VIH-1/M	15	12	5,2	0,13	12	9	6,7	0,30
VIH-1/O	18	9	4,7	0,17	19	14	6,2	0,16
$\Delta$ (OM-M)*	+ 1	- 4	- 0,7	+ 0,01	+ 7	+ 2	- 0,3	- 0,15
$\Delta$ (OM-O)**	- 2	- 1	- 0,2	- 0,03	0	- 3	+ 0,2	- 0,01

<sup>§</sup>Durée entre le démarrage de la répliation et l'atteinte du pic

\*Différences calculées entre les valeurs moyennes de la souche recombinante et de la souche parentale VIH-1/M

\*\*Différences calculées entre les valeurs moyennes de la souche recombinante et de la souche parentale VIH-1/O

Les différences observées en faveur de la souche recombinante apparaissent en orange et celles en faveur des souches parentales VIH-1/M et VIH-1/O apparaissent en rouge et en bleu, respectivement. Si aucune différence apparente n'est relevée, le chiffre est laissé en noir.

Concernant l'activité de la TI, la comparaison des résultats de la souche recombinante VIH-1/OM par rapport aux souches parentales montre que le comportement de la forme recombinante semble se rapprocher de celui de la souche parentale VIH-1/M. En effet, le pic est apparu un jour plus tard pour la souche recombinante VIH-1/OM mais deux jours plus tôt que celui de la souche parentale VIH-1/O et sa vitesse de croissance était similaire à celle de la souche parentale VIH-1/M (TABLEAU 13). En revanche, la durée de réplication active de la souche recombinante VIH-1/OM a semblé plus proche de celle de la souche parentale VIH-1/O, avec 8 et 9 jours de réplication, respectivement, contre 12 jours pour la souche parentale VIH-1/M. Enfin, l'activité de la TI au pic était beaucoup plus faible pour la souche recombinante que pour la souche parentale VIH-1/M et était même inférieure à celle de la souche parentale VIH-1/O. La vitesse de croissance de la souche recombinante VIH-1/OM, associée à une durée de croissance plus faible, explique probablement la moindre intensité de l'activité de la TI au pic, vis-à-vis des souches parentales VIH-1/M ( $\Delta = - 0,7 \text{ Log}_{10} \text{ pg/mL}$ ) et VIH-1/O ( $\Delta = - 0,2 \text{ Log}_{10} \text{ pg/mL}$ ).

Concernant la quantité d'AgP24, la tendance ne semble pas être la même. En effet, la souche recombinante VIH-1/OM a montré un comportement plus proche de la souche parentale VIH-1/O, avec un pic retardé de 7 jours par rapport à la souche parentale VIH-1/M, comme observé pour la souche parentale VIH-1/O (TABLEAU 13). De la même façon, la vitesse de croissance de la souche recombinante VIH-1/OM était très proche de celle de la souche parentale VIH-1/O (0,15 et 0,16  $\text{Log}_{10} \text{ pg/mL /jour}$ , respectivement), alors qu'elle était plus faible de moitié que celle de la souche parentale VIH-1/M (0,30  $\text{Log}_{10} \text{ pg/mL /jour}$ ). Cependant, la souche recombinante VIH-1/OM est parvenue à produire un peu plus d'AgP24 que la souche parentale VIH-1/O ( $\Delta = + 0,2 \text{ Log}_{10} \text{ pg/mL}$ ).

Globalement, tous ces éléments orienteraient vers un comportement hybride de la souche recombinante avec une activité de la TI dont le pic apparaît plus tôt que la souche parentale VIH-1/O, ce qui pourrait être lié aux premières phases du cycle de réplication avec une entrée dans la cellule favorisée par une enveloppe de VIH-1/M. Ensuite, pour l'intensité de l'activité de la TI et pour la production d'AgP24, le comportement de la souche recombinante VIH-1/OM se rapprocherait plutôt de celui de la souche parentale VIH-1/O, ce qui pourrait être, de la même façon, lié au profil de recombinaison de la construction, à savoir [ $O_{\text{gag/pol}}-M_{\text{env}}$ ].

### III- Discussion

L'objectif de ce volet du travail était d'évaluer la répliation *in vitro* du virus recombinant inter-groupes M et O, obtenu par synthèse d'un CMIC, par rapport à ses souches parentales VIH-1/M et VIH-1/O.

En effet, la synthèse par la société Genecust du CMIC pVIH-1/OM, capable de générer un virus recombinant de profil [ $O_{gag/pol-M_{env}}$ ] après transfection et co-culture, nous a permis d'étudier l'éventuel avantage conféré par la recombinaison, par rapport aux souches parentales VIH-1/O et VIH-1/M. Le recours à des CMI parentaux pour générer des CMIC, devant ensuite générer des virus recombinants VIH-1/OM par génétique inverse, nous a permis de nous affranchir de tout biais lié à la diversité, puisque le virus recombinant a été comparé aux deux souches parentales VIH-1/M et VIH-1/O ayant servi à sa génération.

Lors de la vérification de l'infectiosité des CMI parentaux pRBF206 et p89.6, l'étape de co-culture avait d'abord été réalisée avec les cellules CEMx174, qui expriment à leur surface le récepteur CD4 ainsi que le co-récepteur CXCR4. Ce système ne parvenait pas à amplifier la souche RBF206, laissant suspecter un problème de tropisme. Des essais de culture sur lignée cellulaire GHOST-X4 et GHOST-R5 ont été donc réalisés afin de vérifier cette hypothèse et ont confirmé le tropisme uniquement R5 de la souche parentale VIH-1/O RBF206 (résultats non montrés). La génération des stocks a donc été réalisée avec la lignée cellulaire Jurkat X4-R5, présentant à sa surface les deux récepteurs et pouvant ainsi répliquer des souches de tropisme différent. Cependant, les proportions d'expression de chacun des deux co-récepteurs à la surface de cette lignée ne sont pas connues et une surexpression de l'un ou de l'autre pourrait entraîner une sensibilité de ces cellules vis-à-vis d'une l'infection différente en fonction du tropisme des souches. Il pourrait donc être judicieux de mesurer ces pourcentages d'expression en cytométrie en flux, afin de s'assurer de l'absence d'impact de cet élément sur nos données comparatives.

Afin d'éliminer tout biais lié à la variabilité inter-manipulations de la culture cellulaire, chaque cinétique de chacune de souches a été réalisée en triplicat et la reproductibilité des valeurs a permis de procéder aux comparaisons. Globalement, quelques valeurs

exceptionnellement aberrantes ont été observées, mais les valeurs minimales et maximales relevées dans chaque triplicat restaient proches.

Le choix des outils de suivi de la réplication était déterminant pour notre objectif de comparaison du potentiel répliatif. La mesure de la quantité d'AgP24 dans les surnageants de culture est la méthode la plus largement utilisée pour suivre l'évolution de la réplication virale, car elle reflète la production de la protéine de capsid p24 du VIH [226]. Si la vérification de la présence de virus recombinants, générés à partir des CMIC candidats, a été réalisée par la technique VIDAS® HIV P24 II CONFIRMATION (bioMérieux), du fait de sa rapidité et du peu de mesure à réaliser, la comparaison des potentiels répliatifs a été réalisée avec la trousse de quantification de l'AgP24 INNOTEST® HIV Antigen mAb (Fujirebio), pour éviter un possible biais de quantification lié à la diversité antigénique des VIH-1/O, puisque la trousse INNOTEST® avait été précédemment décrite comme étant moins impactée par la diversité que la trousse VIDAS® [226]. Toujours dans le but de nous affranchir de cette problématique liée à la diversité des VIH-1/O, nous avons également réalisé, sur chacun des surnageants, une mesure de l'activité de la TI par la méthode Cavid, habituellement utilisée au laboratoire. Cette méthode présente l'avantage de quantifier une activité enzymatique virale et ne semble pas impactée par la diversité virale [227]. De plus, à la différence de l'AgP24 produit au cours de la réplication qui correspond à une accumulation de la protéine au cours du temps, la mesure de l'activité de la TI apparaît davantage représentative du potentiel répliatif de la souche.

L'utilisation de ces deux marqueurs a mis en évidence des comportements différents en fonction de la souche et du marqueur. Ainsi, la comparaison des cinétiques d'activité de la TI a fait apparaître que la souche parentale VIH-1/M répliquait plus rapidement et plus intensément que les souches recombinante VIH-1/OM et parentale VIH-1/O, mais avec un comportement intermédiaire de la forme recombinante. En effet, les résultats ne se superposaient complètement ni à ceux obtenus avec la souche parentale VIH-1/M, ni avec la souche parentale VIH-1/O. La production d'AgP24 par la souche parentale VIH-1/M était aussi plus importante que celles des deux autres souches, avec toujours un comportement intermédiaire de la souche recombinante VIH-1/OM.

Nos résultats permettent donc de confirmer que la souche parentale VIH-1/M possède un potentiel répliatif significativement supérieur à celui de la souche parentale VIH-1/O, comme décrit *in vitro* par Arien et al. [169]. En revanche, concernant la réplication de la souche



recombinante VIH-1/OM, nos résultats apportent des éléments différents des hypothèses avancées par Peeters et *al.* [192]. En effet, après co-culture des CMSPs d'une patiente co-infectée par une forme recombinante VIH-1/MO et sa souche parentale VIH-1/M, seule la forme recombinante était détectée. L'équipe avait alors suggéré que la capacité répliative de la souche recombinante VIH-1/MO était supérieure à celle de la souche parentale VIH-1/M dont elle était issue [192]. Nos résultats montrent, quel que soit le marqueur, une supériorité répliative nette de la souche parentale VIH-1/M par rapport à la souche recombinante. Cette dernière semble présenter certains avantages par rapport à sa souche parentale VIH-1/O, mais les courbes de suivi de l'activité de la TI et de la quantité d'AgP24 ne sont pas non plus calquées sur celles de la souche parentale VIH-1/M, ce qui ne permet donc pas de démontrer un meilleur potentiel répliatif de la souche recombinante VIH-1/OM.

Le profil de recombinaison [ $O_{gag/pol-M_{env}}$ ] choisi pour cette souche recombinante pourrait expliquer ces résultats. En effet, sa phase de croissance exponentielle débute plus tôt que celle de la souche parentale VIH-1/O, suggérant une infectiosité plus importante de la souche recombinante. L'enveloppe virale joue un rôle clé dans la reconnaissance des récepteurs et co-récepteurs, dans les phases précoces de l'infection ainsi que dans le *fitness* viral [228, 229]. L'enveloppe de la souche recombinante générée provenant de la souche parentale VIH-1/M pourrait donc être favorable à la forme recombinante. La prédominance d'enveloppes de groupe M parmi les recombinants VIH-1/MO décrits *in vivo* (7 profils [ $O_{gag/pol-M_{env}}$ ] contre 4 profils [ $M_{gag/pol-O_{env}}$ ]) est en faveur de cette hypothèse. La production de la souche recombinante VIH-1/OM est, en revanche, apparue moins bonne sur la durée, possiblement du fait que les régions du génome codant pour les protéines de structure et les enzymes de répliation provenaient de la souche parentale VIH-1/O. En effet, le pic d'activité de la TI de la souche recombinante était proche de celui de la souche parentale VIH-1/O (plus faible de 0,2 Log<sub>10</sub> pg/mL), tout comme le pic de la quantité d'AgP24 (plus fort de 0,2 Log<sub>10</sub> pg/mL).

De plus, un avantage de la souche recombinante VIH-1/OM, par rapport à la souche parentale VIH-1/O, peut être attendu, du fait du gène *vpu* appartenant à la souche parentale VIH-1/M. La protéine Vpu des VIH-1/M est en effet connue pour contourner efficacement le facteur de restriction cellulaire Téthérine. Cependant, dans le cadre de notre collaboration sur les plasmides VIH-1/O, les équipes allemande et américaine ont montré, après le démarrage de notre étude, que la souche RBF206, bien qu'appartenant au groupe O, présentait une protéine Vpu capable de contourner efficacement la téthérine, contrairement aux autres variants de ce

groupe <sup>[213]</sup>. Cette souche, n'est donc peut-être pas la plus représentative des VIH-1/O, connus pour avoir une protéine Vpu non adaptée, et présente potentiellement des différences moins marquées, lorsque comparée aux VIH-1/M. Son potentiel répliatif, du fait de cette particularité de Vpu, pourrait donc être supérieur à celui des autres variants de ce groupe et expliquer la faible différence de répliation observée par rapport à la forme recombinante VIH-1/MO.

Les questions soulevées par cette étude sont donc nombreuses et rendent nécessaire de générer de nouveaux CMIC. Différentes approches sont à tester, afin non seulement de s'affranchir des problématiques rencontrées dans cette étude, mais aussi d'étudier l'impact des profils de recombinaison. Des souches parentales de VIH-1/M et de VIH-1/O de même tropisme, et de sous-type et de sous-groupe retrouvés associés dans des formes recombinantes VIH-1/MO *in vivo*, pourront ainsi être utilisées. De plus, l'introduction de deux points de cassure dans le génome du virus recombinant à générer pourrait permettre de contourner le phénomène de recombinaison au sein de LTRs, lors de la production virale. En effet, les différents CMIC pVIH-1/OM avaient été construits avec deux LTRs différents et ont, pour la majorité d'entre eux, produit des virus recombinants complets mais non infectieux. La contrainte de recombinaison dans les LTRs pourrait représenter un frein à la génération d'un progénie virale recombinante infectieuse.



**PARTIE 2.**  
**ETUDE DE L'EMERGENCE *IN VITRO* ET *IN VIVO***  
**DES FORMES RECOMBINANTES INTER-GROUPES**  
**M ET O**



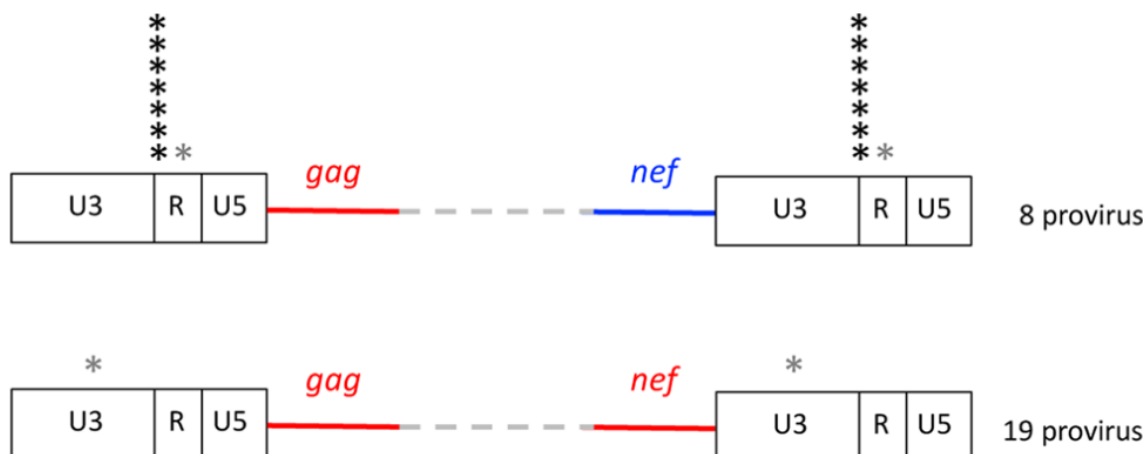
## A. Etude *in vitro* de la recombinaison au sein des régions LTR

---

### Contexte et objectifs du travail expérimental

L'étude de la répartition des points de cassure des recombinants VIH-1/MO a montré la présence d'un point chaud de recombinaison dans la région des LTRs, retrouvé chez 8 des 12 (67%) URF\_MO complètement séquencées [201]. Par ailleurs, lors de l'analyse des caractéristiques moléculaires des 19 URF\_MO, nous avons constaté que lorsqu'un nombre impair de points de cassure était observé au sein des génomes, hors LTRs, un point de cassure était systématiquement retrouvé dans les LTRs [201]. A l'inverse, lorsque le nombre de points de cassure dans le génome, hors LTRs, était pair, aucun point de cassure n'était retrouvé dans les LTRs. Ces différents constats faisaient des LTRs, une région d'intérêt pour l'étude des mécanismes de la recombinaison.

Une étude, portant spécifiquement sur le mécanisme de recombinaison au sein des LTRs, a donc été réalisée en collaboration avec l'équipe de recherche CNRS Architecture et Réactivité de l'ARN de Strasbourg [12]. Ainsi, 27 séquences de LTRs de souches recombinantes intra-groupe M (13 CRFs et 5 URFs) et inter-groupes MO (9 URF\_MO), issues de la base de données Los Alamos et des données du laboratoire de Rouen, ont été analysées. La localisation des points de cassure semblait concentrée à la jonction entre les régions U3 et R, à l'endroit où le premier transfert de brin obligatoire a lieu, lors du cycle de réplication (FIGURE 36) [12].

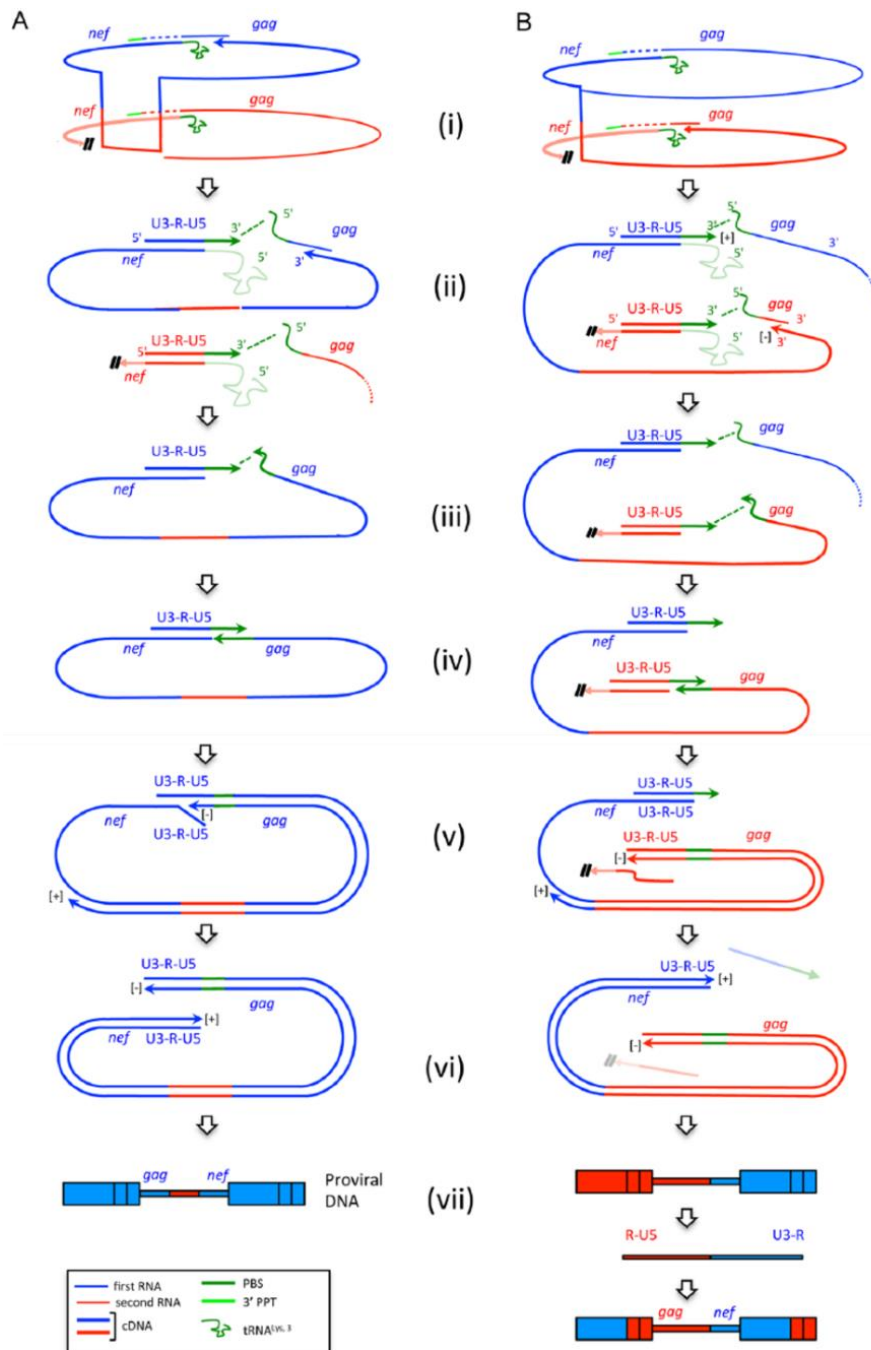


**FIGURE 36** : Localisation et nombre de points de cassure retrouvés dans le LTR.

Les 8 virus (1 CRF et 2 URFs intragroupe M et 5 URF\_MO) dont les régions génomiques Gag et Nef appartenaient à des groupes différents (M/O ou O/M) sont représentées dans la partie supérieure. Les 19 virus (12 CRF et 3 URFs intragroupe M et 4 URF\_MO) dont les régions génomiques Gag et Nef appartenaient au même groupe (soit M, soit O) sont représentées dans la partie inférieure. Les astérisques indiquent l'emplacement des points de cassures identifiés lors des analyses. Les points de cassure attribués au choix de copie sont indiqués en gris.

Afin de déterminer lequel des deux transferts de brin obligatoire était responsable de la génération de points de cassure dans les LTRs de virus recombinants, une analyse approfondie du processus de transcription inverse, avec modélisation de la recombinaison entre un brin d'ARN donneur et un brin d'ARN accepteur, a été menée (FIGURE 37).

Il en est ressorti qu'après le saut de matrice et la fin de la synthèse du brin (-), le brin (+) (soit celui issu de l'ARN donneur, soit celui issu de l'ARN accepteur) doit être transféré une deuxième fois sur le brin (-) pour que sa synthèse s'achève. L'hypothèse avancée était que ce second transfert de brin est guidé par un facteur de proximité spatiale et qu'il a lieu du brin (+) accepteur vers le brin (-) synthétisé. Ce processus aboutit à la génération d'un provirus avec des LTRs d'origine différente et avec un seul point de cassure dans la région interne codante (*gag-nef*). La modélisation de Cappy et *al.* montre alors que la progénie virale générée présente également des LTRs différents. Puisque les deux LTRs d'un génome de VIH sont identiques, à la génération suivante, des provirus avec un point de cassure supplémentaire dans les LTRs, à la jonction entre les régions U3 et R, doivent donc être générés (FIGURE 37).



**FIGURE 37 :** Proposition de mécanismes de génération de provirus recombinants, lors du premier transfert de brin obligatoire et conséquences sur les LTRs. (d'après Cappy et al. [12])

(A) Mécanisme selon lequel le premier transfert de brin pourrait générer des provirus avec des LTRs non recombinants. (B) Mécanisme selon lequel le premier transfert de brin pourrait générer des provirus avec des LTRs ayant un point de cassure à la jonction U3/R. Les flèches indiquent une polymérisation de l'ADN dans le sens 5' à 3'. La synthèse du brin d'ADN (+) à partir du PPT central ne figure pas.

(i) Après initiation de la transcription inverse sur un ARN (bleu), le transfert de brin bloque de manière irréversible (double barre oblique) la potentielle transcription inverse en cours du deuxième ARN (en rouge). (ii) La synthèse de l'ADN (-) se poursuit vers l'extrémité 5' de l'ARNg. L'ADN (+) est synthétisé sur les deux ARN et les ARNt (estompés) ayant servi d'amorce à la TI sont éliminés. L'extrémité 5' des ARN et les ADN (+) correspondants sont maintenus à proximité (ligne pointillée verte). (iii) La synthèse de l'ADN (-) est complétée par la synthèse de son PBS (vert). Les molécules ne participant plus aux étapes suivantes ne figurent plus. (iv) Le deuxième transfert de brin a lieu, guidé par la proximité spatiale des brins (ligne pointillée verte en iii). (v) La synthèse de l'ADN reprend après ce second transfert de brin et remplace les LTRs pré-existants. (vi) La synthèse de l'ADN pré-proviral complet est réalisée (les brins estompés en B sont supprimés). (vii) La structure des ADN proviraux résultants est symbolisée. L'ADN proviral en B va générer un point de cassure à la jonction U3/R lors du cycle suivant.











Cette modélisation implique donc, lorsqu'il y a un nombre impair de points de cassure au sein de la région codante, qu'un évènement de recombinaison ait spontanément lieu dans les LTRs [12].

Afin de valider expérimentalement cette observation et d'étudier de manière précise les mécanismes de recombinaison au sein des LTRs, un travail a été mené *in vitro*, en utilisant le clone moléculaire infectieux chimérique généré précédemment, qui possédait un unique point de cassure dans *vpr* et des LTRs différents. Pour ce faire, nous avons étudié la cinétique d'émergence des points de cassure dans les LTRs ainsi que leur localisation précise, dans des conditions de culture prolongée du virus recombinant VIH-1/OM. Nous avons ensuite comparé ces données de répartition des points de cassure à celles observées *in vivo* dans les 9 URF\_MO présentant un point de cassure dans les LTRs.

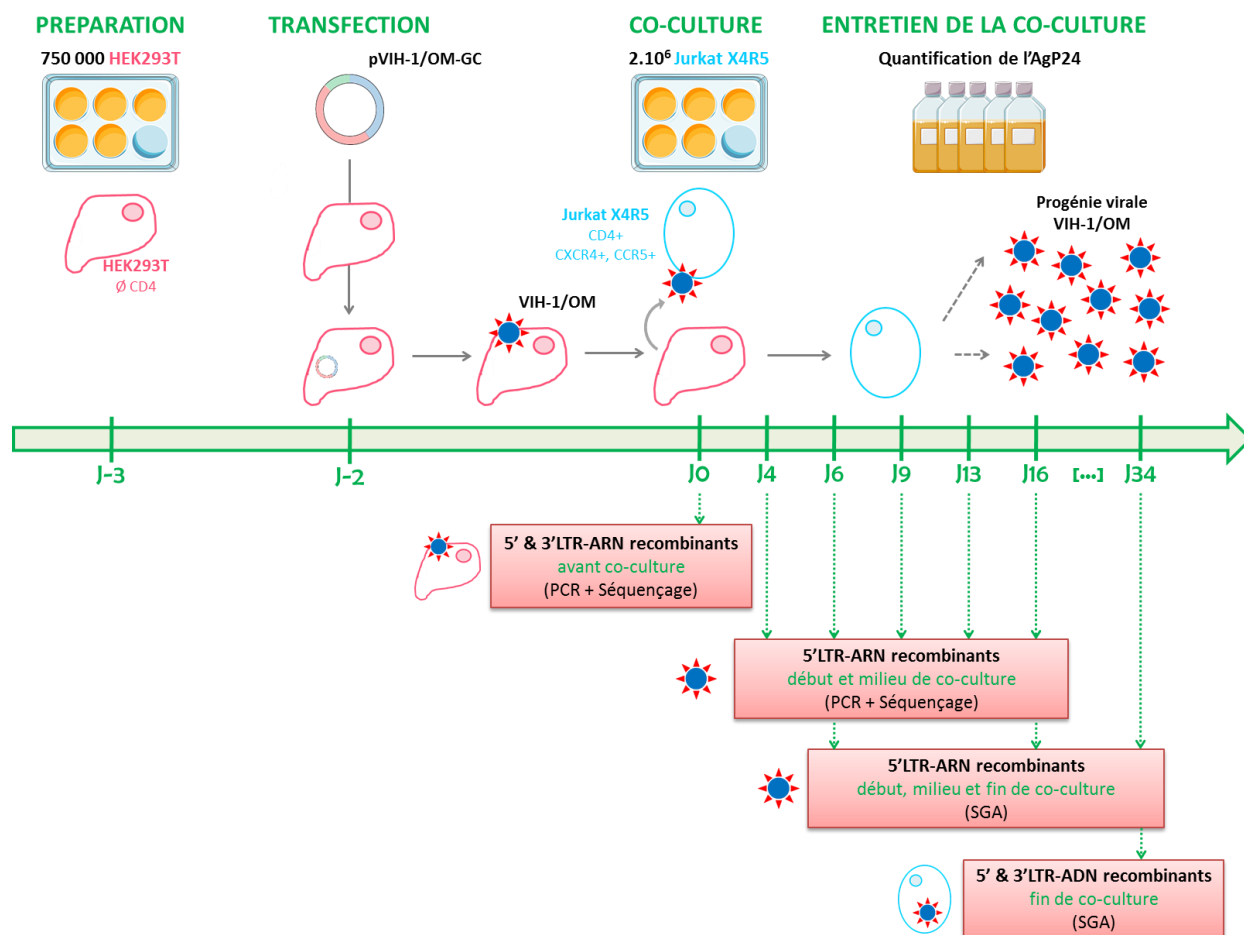
## I- Méthodes

La nomenclature utilisée pour les différents extraits étudiés est indiquée dans le TABLEAU 14. Le principe général de cette étude de la recombinaison dans les LTRs est présenté en FIGURE 38.

TABLEAU 14 : Nomenclature des extraits étudiés

Transfection	HEK293T LTR-ARN	Surnageant co-culture LTR-ARN						Jurkat LTR-ADN
	 J0	 J4	 J6	 J9	 J13	 J16	 J34	 J34
<b>T1</b>	T1-HEK-J0	T1-ARN-J4	T1-ARN-J6	T1-ARN-J9	T1-ARN-J13	T1-ARN-J16	T1-ARN-J34	T1-ADN-J34
<b>T2</b>	T2-HEK-J0	T2-ARN-J4	T2-ARN-J6	T2-ARN-J9	T2-ARN-J13	T2-ARN-J16	T2-ARN-J34	T2-ADN-J34
<b>T3</b>	T3-HEK-J0	T3-ARN-J4	T3-ARN-J6	T3-ARN-J9	T3-ARN-J13	T3-ARN-J16	T3-ARN-J34	T3-ADN-J34
<b>T4</b>	T4-HEK-J0	T4-ARN-J4	T4-ARN-J6	T4-ARN-J9	T4-ARN-J13	T4-ARN-J16	T4-ARN-J34	T4-ADN-J34
<b>T5</b>	T5-HEK-J0	NR*	NR*	NR*	NR*	T5-ARN-J16	NR*	T5-ADN-J34

\*NR : Extraction non réalisée



**FIGURE 38 :** Principe général de l'étude des LTRs du recombinant VIH-1/OM généré *in vitro*.

Après leur dépôt dans les puits d'une plaque 6 puits, les cellules HEK293T ont été transfectées en 5 réplicats par le clone moléculaire infectieux chimérique pVIH-1/OM-GC. Une étape de co-culture avec des cellules Jurkat X4-R5 a été réalisée, afin de permettre la réplication virale. Les LTRs des virus recombinants VIH-1/OM ont été analysés à différents moments de la culture et sous les deux formes génomiques virales (ARN et ADN), afin de rechercher des points cassure et d'étudier leur cinétique d'émergence. Les 5' et 3'LTR-ARN ont été d'abord étudiés en fin de transfection (J0), sur les surnageants de culture des HEK293T, et donc avant toute réplication dans les cellules Jurkat. Les 5'LTR-ARN ont également été étudiés en début, milieu et fin de co-culture, au moyen de la méthode SGA, permettant d'amplifier et de séquencer les multiples populations virales recombinantes présentes dans les surnageants de co-culture. Enfin, les 5' et 3' LTR-ADN, extraits des cellules de fin co-culture à J34, ont également été amplifiés par SGA et séquencés, afin d'étudier toutes les populations virales générées au cours des différents cycles de réplication.

## I-1. Etude de l'émergence d'un point de cassure dans les LTRs

### I-1.1. Recherche d'un point de cassure dans les LTRs

La première étape de ces travaux consistait à vérifier que notre CMIC pVIH-1/OM-GC, avec un unique point de cassure et deux LTRs différents (5'LTR = VIH-1/O et 3'LTR = VIH-1/M), devait recombiner au sein des LTRs pour les avoir identiques et pour ainsi générer un virus recombinant viable et infectieux. A partir de ce CMIC, nous avons choisi de générer, dans les

mêmes conditions, cinq stocks de virus recombinants VIH-1/OM, puis de séquencer leurs 5' et 3'LTR-ARN, afin de déterminer si la recombinaison a toujours lieu dans la même région du LTR et selon les mêmes modalités.

#### a) Génération de virus recombinants

Le virus recombinant a été généré à partir du CMIC pVIH-1/OM-GC, par transfection de cellules HEK293T, suivie d'une étape de co-culture avec la lignée cellulaire Jurkat X4-R5, selon la méthodologie précédemment décrite (I-2. Production de virus recombinants à partir de clones moléculaires infectieux, p.92). Cinq réplicats de transfactions puis co-cultures ont été réalisés (FIGURE 38).

**Transfection.** A J-2 de la co-culture, 2µg de pVIH-1/OM-GC ont été dilués dans 200µL de tampon JetPRIME®, en cinq réplicats. Après centrifugation, 4µL de réactif JetPRIME® ont été ajoutés et l'ensemble a été incubé à température ambiante pendant 10min. La totalité du mix de transfection a été déposée sur les cellules HEK293T et mise à incuber à 37°C et 5% de CO<sub>2</sub> pendant 72h.

**Co-culture.** Au préalable, les cellules Jurkat X4-R5 saines ont été entretenues deux fois par semaine, en renouvelant le milieu et en les divisant au 1/5<sup>ème</sup>.

A J0, les surnageants issus de la transfection ont été retirés de chacun des cinq puits et congelés à -80°C. Ensuite, 2 millions de Jurkat ont été comptées et déposées sur le tapis de HEK293T. L'ensemble a été mis à incuber dans du milieu RPMI + 10%SVFd + Gentamycine + Généticine à 37°C et 5% de CO<sub>2</sub> pendant 3 jours. A J3, les cellules Jurkat et leur surnageant de culture ont été transférés en flasques T25, qui ont été mises à incuber à 37°C pendant 24h, le temps que les cellules sédimentent et se déposent au fond de la flasque.

#### b) Suivi de la multiplication virale

Entre J4 et J34 de la co-culture, la multiplication virale a été suivie tous les 3 à 4 jours par quantification de l'AgP24, présent dans les surnageants de culture, à l'aide du kit VIDAS® HIV P24 II CONFIRMATION (bioMérieux), selon les recommandations du fabricant.

#### c) Caractérisation des LTR-ARN et des LTR-ADN des populations recombinantes

Les deux formes génomiques, ARN viral et ADN intracellulaire, ont été extraites pour ensuite amplifier et séquencer les 5' et 3'LTRs. Par souci de simplification, les régions des LTRs

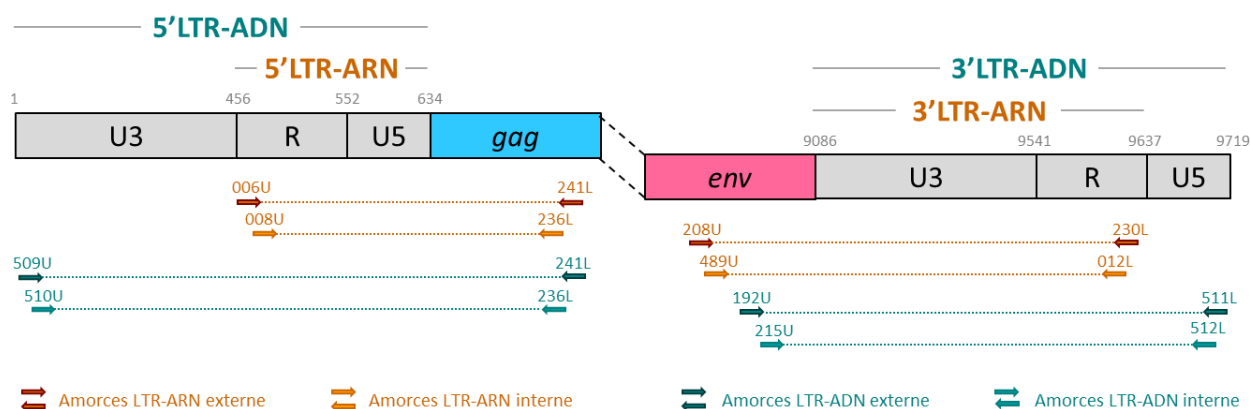
présentes dans l'ARN viral, contenu dans les surnageant de culture, ont été nommées LTR-ARN et les LTRs complets, provenant de l'ADN intracellulaire, intégré ou non au génome de la cellule hôte, ont été nommés LTR-ADN.

**Extraction des ARN et ADN viraux.** Quand la quantité d'AgP24 contenue dans les surnageants de culture dépassait les 3 Log<sub>10</sub> pg/mL, l'ARN viral a été extrait, de manière automatisée, à l'aide du kit *Extraction EZ1 advanced XL (Qiagen)*. A la fin de la culture (J34), les cellules Jurkat ont été remises en suspension, centrifugées et comptées. Elles ont été conservées à -80°C en aliquots de cinq millions de cellules. L'ADN intracellulaire contenu dans ces aliquots a été extrait manuellement sur colonne de silice, avec le kit *Nucleospin Blood – Genomic DNA from Blood (Macherey-Nagel)*, selon les recommandations du fabricant.

**Amplification des 5' et 3'LTRs.** Les régions 5' et 3' LTR recombinantes ont été ensuite amplifiées par un système de PCRs nichées, avec des couples d'amorces différents pour chacun des deux LTRs et pour chacune des deux formes génomiques (ANNEXES 3, 4, 5). L'objectif était d'amplifier spécifiquement le motif MO, généré après recombinaison dans les LTRs, en utilisant des amorces spécifiques du VIH-1/M en 5' et du VIH-1/O en 3' (TABLEAU 15, FIGURE 39).

TABLEAU 15 : Amorces utilisées pour l'amplification des LTRs recombinants de motif MO

Région	PCR	Amorce sens 5'-3'	Amorce antisens 5'-3'	Taille (pb)
5'LTR-ARN	Externe	006U ACCAGRTYTGAGCCYGGGAG	241L TCCATTTTCTATAGATGTCTCCT	1333
	Interne	008U GAGCCYGGGAGCTCTCTG	236L TGTCTTGAGAGCTGGYTCTA	541
3'LTR-ARN	Externe	208U TGGGCAAGTYTGTGGAATTGGTT	230L TCAAGGCAAGCTTTATTGAG	1411
	Interne	489U CCCACAGGCCCGRAGGAA	012L GCAAGCTTTATTGAGSSTTAAGCAG	1223
5'LTR-ADN	Externe	509U TGGAAGGGCTAGTTCACTCC	241L TCCATTTTCTATAGATGTCTCCT	1601
	Interne	510U CCAGAAAAGACAAGATATCC	236L TGTCTTGAGAGCTGGYTCTA	999
3'LTR-ADN	Externe	192U ATGAGGGACAATTGGAGAAGTG	511L GCTAGAGAATTTTCTGCTTCAG	2070
	Interne	215U GARGARGTGGGTTTTCCAGTCA	512L ACGTGATCTGAGGGATCTCTAG	703



**FIGURE 39** : Localisation des amorces utilisées pour l'amplification des LTRs recombinants.

Seules les 1000 premières et les 1000 dernières paires de bases du génome VIH sont représentées. La localisation des différentes régions ciblées est basée sur le génome de la souche de référence VIH-1/M HxB2 (numéro d'accèsion GenBank K03455). Les couples d'amorces de chacune des PCRs nichées (externe et interne) sont symbolisés par des flèches. Les flèches marrons correspondent à ceux utilisés pour l'amplification des LTR-ARN et les flèches vertes à ceux utilisés pour l'amplification des LTR-ADN.

Les amplicons ont été ensuite déposés sur gel d'agarose à 1%, puis purifiés sur colonne, grâce au kit *Nucleospin Gel and PCR clean-up (Macherey-Nagel)*, selon les recommandations du fabricant.

**Séquençage des régions LTRs des virus recombinants.** Les régions 5' et 3'LTR recombinantes ont été séquencées, par la méthode de Sanger, selon la méthodologie précédemment décrite (I-1.1. Modalités de génération des clones moléculaires infectieux chimériques, p.80). Pour l'analyse, le chromatogramme a été aligné par rapport à une séquence de référence, correspondant au LTR recombinant théorique attendu, de motif MO et avec un point de cassure à la jonction U3/R. Les séquences analysées ont ensuite été alignées avec les séquences parentales de référence VIH-1/M (89.6) et VIH-1/O (RBF206), grâce au logiciel MEGA (*Molecular Evolutionary Genetics Analysis, Versions 7 et X*). Les points de cassure ont été délimités de la dernière base séquencée, s'alignant sur la souche parentale VIH-1/M, à la première base séquencée, s'alignant sur la souche parentale VIH-1/O. La localisation du point de cassure dans les régions U3, R ou U5 des LTRs a été ensuite définie, par rapport à la souche de référence VIH-1/M HxB2 (numéro d'accèsion GenBank K03455).

### I-1.2. Etude de la cinétique d'émergence des points de cassure

Une fois la recombinaison dans les LTRs mise en évidence, nous avons cherché à déterminer à quel moment émergeait le point de cassure pendant la culture. En théorie, les cellules HEK293T ne possédant pas tous les constituants nécessaires à la multiplication virale, le point de cassure ne devait pas émerger en sortie de transfection, mais pendant l'étape de co-culture avec les cellules Jurkat, sensibles et permissives au VIH. Nous avons donc d'abord contrôlé l'absence de 5' et 3'LTR-ARN recombinants à l'issue de la transfection dans les surnageants congelés à J0, juste avant la co-culture, et correspondant aux produits de la transfection sur cellules HEK293T. Ensuite, nous avons caractérisé les 5' et 3'LTR-ARN recombinants à différents moments de la co-culture.

#### a) Surnageants HEK2937 (J0)

Les ARNs viraux des cinq surnageants obtenus à l'issue de la transfection ont été extraits de manière automatisée, à l'aide du kit *Extraction EZ1 advanced XL (Qiagen)*. Les 5' et 3'LTR-ARN recombinants ont été recherchés par PCR nichées spécifiques du motif de recombinaison MO, puis purifiés et séquencés, selon la même méthodologie que précédemment (I-1.1. Recherche d'un point de cassure dans les LTRs, p.133) et en utilisant les mêmes amorces (TABLEAU 15, p.135).

#### b) Caractérisation des points de cassure dans les 5'LTR-ARN au cours de la co-culture

La recherche d'un point de cassure dans les 3'LTR-ARN n'ayant pas abouti seuls les 5'LTR-ARN, issus des surnageants de co-culture sur Jurkat, ont été séquencés.

**Echantillons et extraction des ARN viraux.** Seize surnageants de VIH-1/OM-GC ont été analysés, purs ou après dilution au 1/100<sup>ème</sup>. Ils correspondaient aux surnageants prélevés à J4, J6, J9 et J13 de la co-culture (T[1, 2, 3, 4]-ARN-J[4, 6, 9, 13] (TABLEAU 14, p.132)). L'extraction des ARN viraux contenus dans ces 16 surnageants a été effectuée de manière automatisée, avec l'*EZ1 advanced XL (Qiagen)* et le kit *EZ1<sup>®</sup> DSP Virus (Qiagen)*, selon les recommandations du fabricant. Un volume de 200µL de surnageant était prélevé et l'élution était réalisée dans 60µL.

**Amplification et séquençage des 5'LTR-ARN.** Les régions 5'LTR ont été amplifiées par le même système de PCRs nichées que décrit en I-1.1. Recherche d'un point de cassure dans les LTRs (p.133), et en utilisant les mêmes amorces spécifiques du motif de recombinaison MO

(TABLEAU 15, p.135). Les amplicons obtenus ont ensuite été déposés sur gel d'agarose puis purifiés, grâce au kit *Nucleospin Gel and PCR clean-up (Macherey-Nagel)*, selon les recommandations du fabricant. Le séquençage des 5'LTR-ARN recombinants a été réalisé par la méthode de Sanger, selon la méthodologie précédemment décrite (I-1.1. Modalités de génération des clones moléculaires infectieux chimériques, p.80). La position des points de cassure a été définie selon le même procédé que précédemment décrit (I-1. Etude de l'émergence d'un point de cassure dans les LTRs, p.133).

## I-2. Détermination des motifs de recombinaison

L'étape suivante du travail avait pour but de distinguer les différents profils de recombinaison possibles à différents moments de la culture et d'évaluer l'éventuelle prédominance de l'un des profils de recombinaison en fin de culture. Pour cela, nous avons caractérisé les différentes populations de LTRs recombinants produites au cours de la co-culture par la méthode SGA (Single Genome Amplification).

### I-2.1. Populations ARN et ADN

La première analyse a porté sur les 5'LTR-ARN des virus recombinants issus des transfections et avait pour objectif de caractériser les différentes populations virales en début (J6), milieu (J16) et fin (J34) de co-culture. La recherche de ces populations a été réalisée à partir des extraits d'ARN des quatre surnageants collectés au point J6 (T1-ARN-J6, T2-ARN-J6, T3-ARN-J6, et T4-ARN-J6) et à partir des extraits d'ARN d'un seul des quatre surnageants collecté aux points J16 (T1-ARN-J16) et J34 (T1-ARN-J34).

La seconde analyse visait à étudier l'ADN intracellulaire de toutes les populations virales générées à la fin de la co-culture (J34). Les 5' et 3' LTR-ADN ont été analysés pour quatre transfections (T1-ADN-J34, T2-ADN-J34, T3-ADN-J34, et T4-ADN-J34), après extraction de l'ADN des cellules de fin de culture.

### I-2.2. SGA et séquençage

Dans les étapes précédentes, la localisation des points de cassure n'avait pu être déterminée précisément, du fait de la présence de multiples populations virales. Une approche

de SGA a donc été développée. Cette méthode est basée sur la dilution limite des extraits avant leur amplification par PCR et repose sur le fait que les produits de PCR, obtenus à des dilutions donnant des amplifications positives inférieures à 30%, peuvent être considérés comme provenant d'une unique population virale amplifiable, d'après la loi de distribution de Poisson, et peuvent donc être séquencés ensuite [230, 231].

A partir des extraits d'ARN viral, la synthèse de l'ADNc a été réalisée en utilisant le kit *Superscript III First strand synthesis system for RT-PCR (Invitrogen)*, selon les instructions du fabricant (ANNEXE 6). L'ADN (ADNc synthétisé à partir de l'ARN ou ADN proviral extrait à J34) a été dilué en série et chaque dilution a été amplifiée en 10 réplicats pour déterminer la dilution limite, à laquelle au maximum 30% des PCRs étaient positives (soit trois réplicats sur 10). L'amplification des 5' et 3'LTRs recombinants reposait sur la même méthode de PCRs nichées que décrite en I-1.1. Recherche d'un point de cassure dans les LTRs (p.133) avec des amorces spécifiques du motif de recombinaison MO (TABLEAU 15, p.135). Pour chacun des LTRs des quatre transfections, une gamme de différentes dilutions de l'ADN dans de l'eau a été réalisée. Les amplicons obtenus pour chacune des dilutions testées ont été déposés sur gel.

Le point de dilution donnant au maximum 30% de PCR positives a été sélectionné. La même PCR a ensuite été réalisée 96 fois, selon les mêmes conditions opératoires, et à partir de la dilution sélectionnée. Les 96 amplicons obtenus ont été déposés sur gel, et la proportion d'au maximum 30% de PCRs positives a, de nouveau, été vérifiée à cette étape, afin de s'assurer qu'une seule espèce virale serait séquencée par la suite. Les produits des PCRs positives ont été séquencés et la localisation des points de cassure a été déterminée comme décrit en I-1.1. Recherche d'un point de cassure dans les LTRs (p.133).

### I-2.3. Recherche de la prédominance d'un profil préférentiel

Les données de séquençage, obtenues après SGA sur les 5'LTR-ARN à J6, J13 et J34 de la co-culture et sur les 5' et 3'LTR-ADN à J34 de la co-culture, ont ensuite alignées entre elles et analysées de manière séquentielle, afin de mettre en évidence un éventuel profil préférentiel, prédominant en fin de culture.



### I-3. Comparaison aux données *in vivo*

La dernière analyse a consisté à comparer nos données générées *in vitro* aux données des 9 URF\_MO (97CA.MP645, BCF174/212, BCF204, RBF221, RBF235, YBF274, YBF282, REC003/024, et REC107), pour lesquelles un point de cassure dans les LTRs avait été mis en évidence. L'objectif était de faire ressortir des points de cassure communs validant ainsi notre protocole de génération de virus *in vitro*, sans pression de sélection de l'hôte.

Pour cela, nous avons vérifié la localisation de chacun des points de cassure décrits *in vivo*, selon la même méthode que celle utilisée pour les données *in vitro* (I-1.1. Recherche d'un point de cassure dans les LTRs, p.133). Les LTRs étant identiques, nous nous sommes affranchis de l'origine des séquences analysées (5' ou 3'LTR) et avons positionné tous les points de cassure par rapport au 5'LTR du génome de référence VIH-1/M HxB2 (numéro d'accèsion GenBank K03455). Nous avons ensuite comparé la localisation des points de cassure observés *in vivo* à la localisation de ceux générés *in vitro*, en alignant les séquences avec le logiciel MEGA.

## II- Résultats

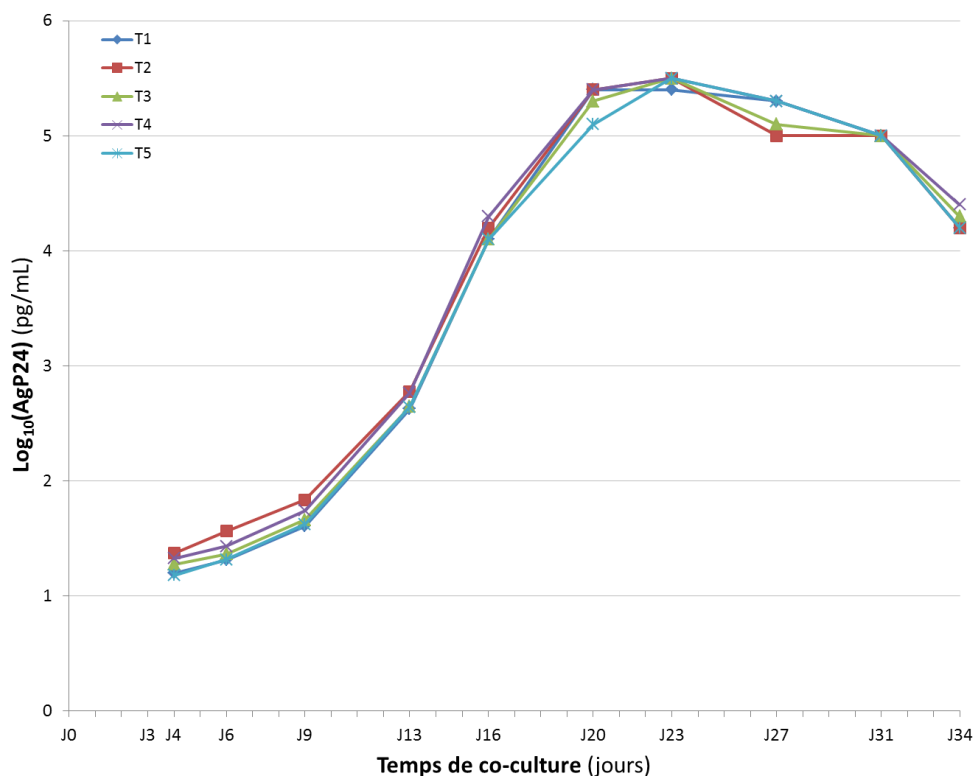
### II-1. Etude de l'émergence d'un point de cassure dans les LTRs

#### II-1.1. Recherche d'un point de cassure dans les LTRs

##### a) Génération de virus recombinants et suivi de la multiplication virale

Après les cinq transfections et co-cultures, l'évolution des titres d'AgP24 des cinq surnageants étudiés a montré des productions virales comparables les unes par rapport aux autres (FIGURE 40).

Globalement, le pic de quantité d'AgP24 a été atteint entre J20 (T1) et J23 (T2, T3, T4 et T5) de co-culture, moment à partir duquel une phase de décroissance a débuté pour chacun des cinq réplicats. La valeur du pic d'AgP24 était considérée comme identique pour les cinq réplicats : 5,4 Log<sub>10</sub> pg/mL pour T1, et 5,5 Log<sub>10</sub> pg/mL pour les quatre autres (T2, T3, T4 et T5), validant la reproductibilité du protocole (FIGURE 40).



**FIGURE 40 :** Evolution de la multiplication du virus recombinant VIH-1/OM dans les cellules Jurkat.

Représentation de l'évolution de la quantité d'AgP24, en  $\log_{10}$  pg/mL, produite au cours de la co-culture dans les cellules Jurkat, en jours, pour les cinq réplicats de transfection du plasmide pVIH-1/OM.

### b) Caractérisation des LTR-ARN et des LTR-ADN des populations recombinantes

Concernant les LTR-ARN, leur caractérisation a été réalisée sur le point J16, date à laquelle tous les AgP24 étaient supérieurs à 3  $\log_{10}$  pg/mL. L'analyse des données de séquençage des 5'LTR-ARN a permis de mettre en évidence une zone de recombinaison, localisée dans la deuxième moitié de la région R entre 506 et 547 pb, pour les cinq réplicats (FIGURE 41).

La position précise des points de cassure n'a pas pu être réalisée à cette étape, compte tenu de l'observation de doubles populations au séquençage, reflétant la présence de populations virales hétérogènes dans chaque surnageant de culture.

De plus, seul un des cinq 3'LTR-ARN amplifiés a pu être séquencé, révélant également une zone de recombinaison dans la deuxième moitié de R, avec présence de doubles populations. En effet, l'étude des points de cassure dans les autres 3'LTR-ARN s'est heurtée à de nombreuses difficultés techniques, lors de l'analyse des séquences ou lors de la mise au point d'amorces séquençant les extrêmes extrémités du génome. En effet, sous forme ARN, le

3'LTR ne possède pas de région U5. L'analyse par PCR et séquençage de la seconde moitié de R, où se situaient les phénomènes de recombinaison, s'est donc révélée difficile.

		497	506	513	522	523	547
VIH-1/O		A T C T C T A G C T	G A A C C C	G C T G C T T A A	C G C T C A A T A A A G C T T G C C T T	G A G T G	A - - G A A G
VIH-1/M		G C T A G C T A G G	- - - - -	A - - - - -	- - - - -	- - - - -	C T T C - - -
T1-ARN-J16	5'LTR-ARN	G C T A G C T A G G	- - - - -	R - - - - -	- - - - -	S S - - - - -	A - - G - - -
	3'LTR-ARN	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
T2-ARN-J16	5'LTR-ARN	G C T A G C T A G G	- - - - -	R - - - - -	- - - - -	S S - - - - -	A - - G - - -
	3'LTR-ARN	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
T3-ARN-J16	5'LTR-ARN	G C T A G C T A G G	- - - - -	R - - - - -	- - - - -	S S - - - - -	A - - G - - -
	3'LTR-ARN	G C T A G C T A G G	- - - - -	R - - - - -	- - - - -	S S - - - - -	A - - G - - -
T4-ARN-J16	5'LTR-ARN	G C T A G C T A G G	- - - - -	R - - - - -	- - - - -	S S - - - - -	A - - G - - -
	3'LTR-ARN	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
T5-ARN-J16	5'LTR-ARN	G C T A G C T A G G	- - - - -	R - - - - -	- - - - -	S S - - - - -	A - - G - - -
	3'LTR-ARN	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	- - - - -
T1-ADN-J34	5'LTR-ADN	G C T A G C T A G G	- - - - -	R - - - - -	- - - - -	S S - - - - -	A - - G - - -
	3'LTR-ADN	G C T A G C T A G G	- - - - -	R - - - - -	- - - - -	S S - - - - -	A - - G - - -
T2-ADN-J34	5'LTR-ADN	G C T A G C T A G G	- - - - -	R - - - - -	- - - - -	S S - - - - -	A - - G - - -
	3'LTR-ADN	G C T A G C T A G G	- - - - -	R - - - - -	- - - - -	S S - - - - -	A - - G - - -
T3-ADN-J34	5'LTR-ADN	G C T A G C T A G G	- - - - -	R - - - - -	- - - - -	S S - - - - -	A - - G - - -
	3'LTR-ADN	G C T A G C T A G G	- - - - -	R - - - - -	- - - - -	S S - - - - -	A - - G - - -
T4-ADN-J34	5'LTR-ADN	G C T A G C T A G G	- - - - -	R - - - - -	- - - - -	S S - - - - -	A - - G - - -
	3'LTR-ADN	G C T A G C T A G G	- - - - -	R - - - - -	- - - - -	S S - - - - -	A - - G - - -
T5-ADN-J34	5'LTR-ADN	G C T A G C T A G G	- - - - -	R - - - - -	- - - - -	S S - - - - -	A - - G - - -
	3'LTR-ADN	G C T A G C T A G G	- - - - -	R - - - - -	- - - - -	S S - - - - -	A - - G - - -

FIGURE 41 : Données de séquençage montrant les zones de recombinaison des LTR-ARN et LTR-ADN.

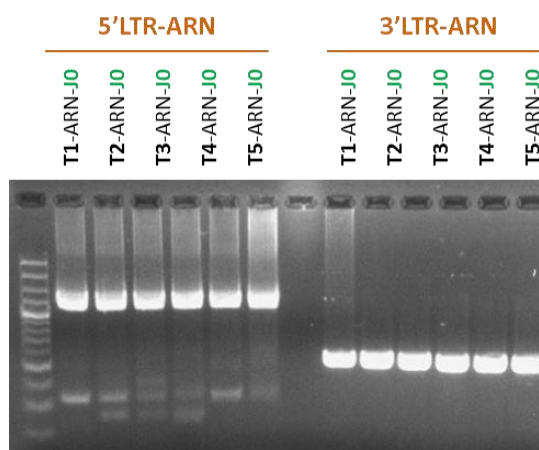
Seule la deuxième moitié de la région R du 5'LTR-ARN est montrée. Les références VIH-1/O et VIH-1/M correspondent aux virus parentaux RBF206 et 89.6, respectivement, et permettent de comparer les différentes formes retrouvées. Les positions des bases correspondent au 5'LTR de la souche de référence HxB2 représentant le VIH-1/M. Les données relatives aux LTR-ARN à J16 de co-culture sont présentées dans la partie supérieure et celles relatives aux LTR-ADN à J34 de co-culture sont présentées dans la partie inférieure. Les tirets représentent les positions conservées entre les VIH-1/M et VIH-1/O. Les lignes grisées correspondent à des échecs de séquençage. Les lettres R et S symbolisent la présence d'une double population G+A et C+G, respectivement, à la position considérée.

Concernant l'analyse des 5' et 3'LTR-ADN à J34 de co-culture, tous présentaient une zone de recombinaison également localisée dans la région R, et avec, là encore, présence de plusieurs populations virales distinctes (FIGURE 41).

## II-1.2. Etude de la cinétique d'émergence des points de cassure *in vitro*

### a) Surnageants HEK2937 (J0)

La migration sur gel d'agarose des amplicons de 5' et 3'LTR-ARN obtenus à partir des cinq surnageants de transfection, a permis d'identifier la présence d'une bande d'environ 541 pb pour le 5'LTR et d'une bande d'environ 1223 pb pour le 3'LTR (FIGURE 42).



**FIGURE 42** : Profil de migration sur gel d'agarose des amplicons 5' et 3'LTR-ARN avant la co-culture.

Les bandes relatives aux 5'LTR-ARN, ayant migré à une hauteur correspondant aux 541 pb attendues, figurent à gauche (puits 2 à 6) et celles relatives aux 3'LTR-ARN, et ayant migré à une hauteur correspondant aux 1223 pb attendues, figurent à droite (puits 7 à 12).

Après séquençage et comparaison des séquences obtenues à celles des souches parentales de VIH-1/M (89.6) et de VIH-1/O (RBF206), aucun point de cassure n'a été observé dans les LTRs étudiés. Les cinq 5'LTRs et les quatre 3'LTRs amplifiés et séquençables correspondaient en intégralité à du VIH-1/O et à du VIH-1/M, respectivement (FIGURE 43).

		497	506	513	522	523	547													
VIH-1/O		A	T	C	T	C	T	A	G	C	T	G	A	A	C	C	G	A	A	G
VIH-1/M		G	C	T	A	G	C	T	A	G	G	-	-	-	-	-	-	-	-	-
		-----																		
T1-HEK-JO	5'LTR-ARN	G	C	T	A	G	C	T	A	G	G	-	-	-	-	-	-	-	-	-
	3'LTR-ARN	A	T	C	T	C	T	A	G	C	T	-	-	-	-	-	-	-	-	-
T2-HEK-JO	5'LTR-ARN	G	C	T	A	G	C	T	A	G	G	-	-	-	-	-	-	-	-	-
	3'LTR-ARN	A	T	C	T	C	T	A	G	C	T	-	-	-	-	-	-	-	-	-
T3-HEK-JO	5'LTR-ARN	G	C	T	A	G	C	T	A	G	G	-	-	-	-	-	-	-	-	-
	3'LTR-ARN	A	T	C	T	C	T	A	G	C	T	-	-	-	-	-	-	-	-	-
T4-HEK-JO	5'LTR-ARN	G	C	T	A	G	C	T	A	G	G	-	-	-	-	-	-	-	-	-
	3'LTR-ARN	-----																		
T5-HEK-JO	5'LTR-ARN	G	C	T	A	G	C	T	A	G	G	-	-	-	-	-	-	-	-	-
	3'LTR-ARN	A	T	C	T	C	T	A	G	C	T	-	-	-	-	-	-	-	-	-

**FIGURE 43** : Données de séquençage de la région R des LTRs des surnageants de transfection (JO)

Seule la deuxième moitié de la région R du 5'LTR-ARN est montrée. Les références VIH-1/O et VIH-1/M correspondent aux virus parentaux RBF206 et 89.6, respectivement. Les positions des bases correspondent au 5'LTR de la souche de référence HxB2 représentant le VIH-1/M (numéro d'accèsion GenBank K03455). La ligne grisée correspond à un séquençage non réalisé. Les tirets représentent les positions conservées entre les VIH-1/M et VIH-1/O.

Les régions LTR des virus issus de la transfection étaient donc identiques à celles du CMIC pVIH-1/OM-GC de départ.

b) Caractérisation des points de cassure dans les 5'LTR-ARN au cours de la co-culture

Les résultats précédents ont donc montré l'absence de recombinaison dans les 5' et 3'LTR-ARN à J0, avant la co-culture, et la présence de points de cassure dans les 5'LTR-ARN à J16. L'étude de l'émergence des points de cassure a donc été focalisée dans le 5'LTR-ARN, entre J0 et J16 de la co-culture (J4, J6, J9 et J13).

La recherche des LTRs recombinants par PCRs nichées spécifiques du motif MO a d'abord été réalisée à partir des surnageants dilués à  $10^{-2}$  et a montré des bandes d'intensité croissante sur gel d'agarose, pouvant correspondre à l'émergence de LTRs recombinants à partir de J6 (FIGURE 44).

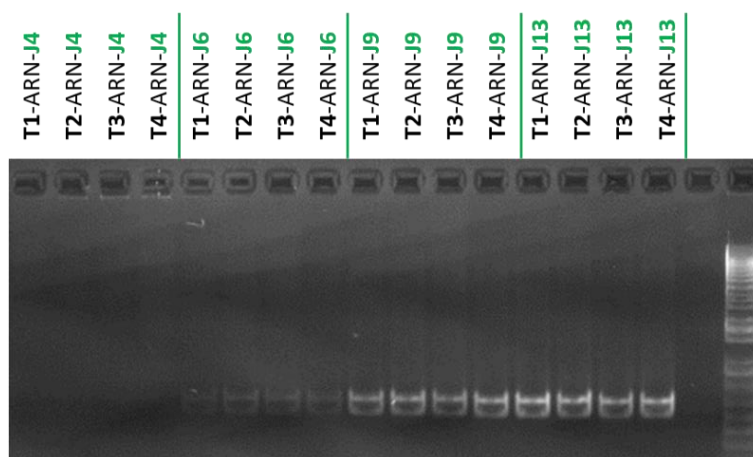


FIGURE 44 : Profil de migration sur gel d'agarose des amplicons 5'LTR-ARN

Les surnageants ont été étudiés pour quatre réplicats de transfection et co-culture à J4, J6, J9 et J13 de la co-culture. Une dilution à  $10^{-2}$  des extraits ARN a été réalisée avant l'étape d'amplification.

L'analyse des surnageants à J4 a alors été répétée, mais sans dilution préalable des extraits, ce qui a finalement permis de constater une positivité de la PCR également à cette date (résultats non montrés).

L'analyse des séquences obtenues a montré que la zone de recombinaison était située entre les positions 506 et 547 pb de la souche de référence VIH-1/M HxB2, pour 10 des 16 extraits étudiés (TABLEAU 16). Cette zone correspondait à la fin de la région R du 5'LTR et corrélait avec les résultats précédemment obtenus, lors de la première étape de recherche des points de cassure dans les LTRs.

TABLEAU 16 : Localisation du point de cassure dans les 5'LTR-ARN des 16 extraits étudiés

Réplicat	Dilution extrait	Localisation point de cassure (pb)	Région du LTR impliquée
T1-ARN-J4	pur	497-546	R
T2-ARN-J4	pur	497-546	R
T3-ARN-J4	pur	NA*	NA*
T4-ARN-J4	pur	497-546	R
T1-ARN-J6	1/100	518-546 <sup>#</sup>	R
T2-ARN-J6	1/100	507-546	R
T3-ARN-J6	1/100	516-546 <sup>#</sup>	R
T4-ARN-J6	1/100	507-546	R
T1-ARN-J9	1/100	507-546	R
T2-ARN-J9	1/100	507-546	R
T3-ARN-J9	1/100	507-546	R
T4-ARN-J9	1/100	507-546	R
T1-ARN-J13	1/100	507-546	R
T2-ARN-J13	1/100	507-546	R
T3-ARN-J13	1/100	507-546	R
T4-ARN-J13	1/100	507-546	R

\*NA : Séquençage non analysable

<sup>#</sup>Zone de recombinaison partiellement séquencée

A J4, l'extrait T3-ARN-J4 n'a pu être séquençé, du fait d'une trop faible quantité de matériel génétique, et les trois autres extraits (T1-ARN-J4, T2-ARN-J4 et T4-ARN-J4) présentaient différents pics pour la même base, en amont de la zone de recombinaison (entre les positions 497 et 507 pb), lié à la difficulté de l'analyse du séquençage dans cette région se situant à l'extrémité 5' du génome ARN et à la plus faible quantité d'ARN viral en début de co-culture (TABLEAU 16).

A J6, la zone de recombinaison a été partiellement caractérisée pour deux des quatre extraits (T1-ARN-J6 et T3-ARN-J6), du fait de la mauvaise qualité des signaux de séquençage à analyser (TABLEAU 16).

La présence de doubles populations dans toutes les séquences obtenues, correspondant à des bases Adénosine et Guanine (A ou G = R), et de Guanine et Cytidine (G ou C = S), indiquait, là encore, la présence de plusieurs populations virales dans les différents surnageants de co-culture étudiés.

Au terme de cette étape de caractérisation des points de cassure dans les 5'LTR-ARN avant et au cours de la co-culture, nous avons pu estimer l'apparition des points de cassure dans les LTRs entre J0 et J4. De plus, la présence de plusieurs populations virales ne nous a permis de définir qu'une zone de recombinaison, localisée entre les positions 507 et 546 pb.

## II-2. Détermination des motifs de recombinaison

Les doubles populations observées lors du séquençage des LTR-ARN recombinants générés au cours de la co-culture ont rendu nécessaire l'analyse d'une seule population virale à la fois. La méthode d'amplification clonale par SGA a donc été utilisée sur les surnageants de début et milieu de co-culture (J6 et J16), mais également à la fin (J34), afin de préciser la localisation des points de cassure et de rechercher un éventuel profil prédominant et préférentiel de recombinaison.

Les LTR-ADN ont aussi été étudiés en fin de co-culture (J34), étant donné que l'ADN viral intracellulaire reflète toutes les populations virales produites et archivées dans les cellules, au fil de la culture et des cycles de réplication virale.

### II-2.1. SGA et séquençage

#### a) 5'LTR-ARN

Afin d'obtenir la dilution limite, les extraits d'ARN ont été dilués de façon croissante en fonction de l'avancée de la co-culture, à J6, J16 puis J34 (TABLEAU 17).

Cette dilution limite était comprise entre 1/100 et 1/190000. Les différentes dilutions limites retenues et testées 96 fois ont permis d'obtenir entre 7,3% et 25% de PCRs positives selon les réplicats testés, pourcentages qui étaient inférieurs à 30% et donc conformes aux conditions requises par le SGA. Au total, 88 séquences de 5'LTR-ARN ont été analysées (TABLEAU 17).

TABLEAU 17 : Répartition des 5'LTR-ARN analysés après SGA

Réplicat	LTR-ARN	Dilution limite	Amplicons positifs	Nb séquences analysées	Total séquences analysées
T1-ARN-J6	5'LTR-ARN	1/300	16,6%	16	88
T2-ARN-J6	5'LTR-ARN	1/300	16,6%	14	
T3-ARN-J6	5'LTR-ARN	1/100	7,3%	7	
T4-ARN-J6	5'LTR-ARN	1/100	25,0%	20	
T1-ARN-J16	5'LTR-ARN	1/175 000	14,6%	11	
T1-ARN-J34	5'LTR-ARN	1/190 000	21,9%	20	

Parmi les 88 séquences analysables, 48 (soit 54,5%) avaient un point de cassure dans le 5'LTR ARN. Parmi elles, 17 avaient été obtenues à J6, 11 à J16 et 20 à J34 (TABLEAU 18). Les 40 séquences restantes correspondaient en intégralité à du VIH-1/O, et donc au CMIC de départ. Ces séquences de VIH-1/O étaient retrouvées uniquement au début de la co-culture (J6) mais ni au milieu, ni à la fin (J16 et J34) (résultats non montrés). Ceci pourrait s'expliquer par la présence en excès du plasmide recombinant lors de la transfection et de la proportion minoritaire de LTRs recombinants à ce stade de la co-culture.

Au final, ces résultats montraient donc qu'à partir de J16, seules des formes recombinantes étaient présentes.

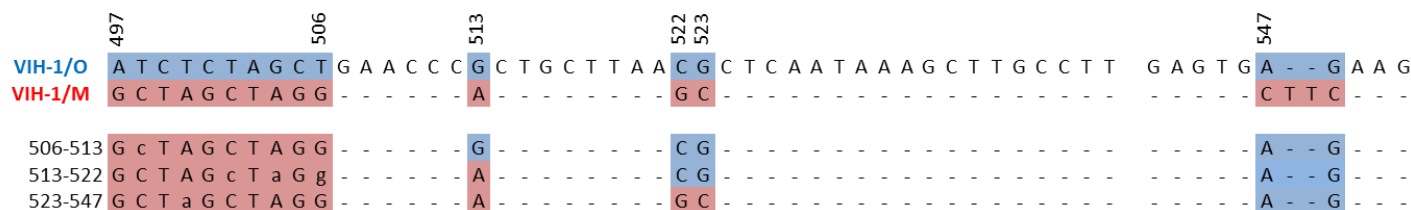
Après alignement, les séquences des 48 5'LTR-ARN recombinants ont montré la présence de trois profils génétiques de recombinaison différents (TABLEAU 18). Les trois points de cassure présents étaient les suivants : 506-513pb, 513-522pb et 523-547pb. Ils étaient tous localisés dans la deuxième moitié de R (FIGURE 45).



**TABLEAU 18** : Proportion des points de cassure recensés dans les 5'LTR-ARN à J6, J16 et J34 après SGA

LTR	Points de cassure (pb)	Localisation	T1-ARN-J6	T2-ARN-J6	T3-ARN-J6	T4-ARN-J6	T1-ARN-J16	T1-ARN-J34	Points de cassure totaux					
									J6	J16	J34			
5'LTR-ARN	506 - 513	R	1	1	-	3	1	1	5	(29,4%)	1	(9,1%)	1	(5,0%)
	513 - 522	R	2	1	-	4	5	14	7	(41,2%)	5	(45,5%)	14	(70,0%)
	523 - 547	R	2	-	-	3	5	5	5	(29,4%)	5	(45,5%)	5	(25,0%)
<b>TOTAL</b>			5	2	0	10	11	20	17		11		20	
						48								

Les tirets indiquent que le point de cassure n'a pas été retrouvé dans l'extrait séquencé.



**FIGURE 45** : Motif des trois points de cassure mis en évidence dans les régions R des 5'LTR-ARN après SGA.

Seule la deuxième moitié de la région R du 5'LTR-ARN est montrée. Les références VIH-1/O et VIH-1/M correspondent aux virus parentaux et permettent de comparer les différentes formes retrouvées. Les positions 506-513, 513-522 et 523-547 pb représentent les formes recombinantes avec les trois points de cassure retrouvés au niveau des 5'LTRs de chaque extrait. Les positions des bases correspondent à la référence HxB2 du VIH-1/M (numéro d'accèsion GenBank K03455). Les tirets représentent les positions conservées entre les VIH-1/M et VIH-1/O.

## b) 5' et 3'LTR-ADN

L'étude de l'ADN nous a permis de nous affranchir des difficultés techniques rencontrées lors de l'étude des 3'LTR-ARN et d'étudier les LTRs dans leur intégralité, puisque complets sous cette forme génomique. Les échantillons d'ADN collectés à J34 ont été dilués de façon importante pour chaque LTR des quatre transfections testées, afin d'obtenir la dilution limite correspondante. Cette dernière était comprise entre 1/130000 et 1/230000. Les différentes dilutions limites retenues et testées 96 fois ont permis d'obtenir entre 11,7% et 24% de PCRs positives selon les LTRs et les transfections testées, pourcentages qui étaient inférieurs à 30% et donc conformes aux conditions requises par le SGA. Au total, 166 séquences dont 91 de 5'LTR-ADN (soit 54,8%) et 75 de 3'LTR-ADN (soit 45,2%), ont été analysées (TABLEAU 19).

TABLEAU 19 : Répartition des LTR-ADN analysés après SGA

Réplicat	LTR-ADN	Dilution limite	Amplicons positifs	Nb séquences analysées		Total séquences analysées
<b>T1-ADN-J34</b>	5'LTR-ADN	1/123 000	24,0%	22	41	<b>166</b>
	3'LTR-ADN	1/130 000	16,7%	19		
<b>T2-ADN-J34</b>	5'LTR-ADN	1/200 000	20,2%	18	38	
	3'LTR-ADN	1/200 000	17,0%	20		
<b>T3-ADN-J34</b>	5'LTR-ADN	1/160 000	22,9%	25	42	
	3'LTR-ADN	1/160 000	16,8%	17		
<b>T4-ADN-J34</b>	5'LTR-ADN	1/180 000	22,9%	26	45	
	3'LTR-ADN	1/140 000	11,7%	19		

Les résultats de séquençage du 5'LTR-ADN ont d'abord montré la présence des trois mêmes profils génétiques de recombinaison dans R que ceux mis en évidence dans les 5'LTR-ARN dans 94.5% des cas. D'autres points de cassure ont également été mis en évidence, de façon minoritaire : 420-443pb (U3), 453-460pb (U3/R), 566-573pb (U5), et 587-604pb (U5), (TABLEAU 20, FIGURE 46).

Concernant le 3'LTR-ADN, les points de cassure 9590-9598pb, 9598-9606pb et 9608-9631pb ont été retrouvés en proportions largement majoritaires (96%). Ils étaient tous localisés dans la deuxième moitié de R. D'autres points de cassure, localisés aux positions 9420-9433pb (*nef*/U3), 9482-9493pb (U3) et 9537-9544pb (U3/R), ont également été mis en évidence (TABLEAU 20, FIGURE 46).

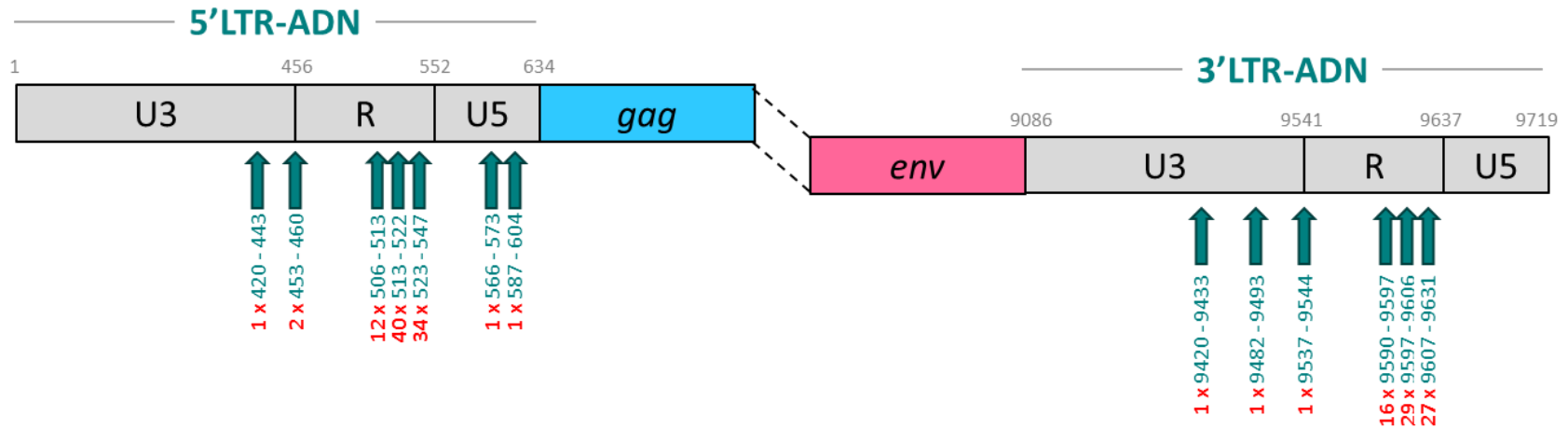
TABLEAU 20 : Proportion des points de cassure recensés dans les 5' et 3'LTR-ADN à J34 après SGA

LTRs	Points de cassure (pb)	Localisation	T1-ADN-J34	T2-ADN-J34	T3-ADN-J34	T4-ADN-J34	Points de cassure totaux	
5'LTR-ADN	420 - 443	U3	-*	-	-	1	1	1,1%
	453 - 460	U3/R	1	-	1	-	2	2,2%
	566 - 573	U5	-	1	-	-	1	1,1%
	587 - 604	U5	-	-	-	1	1	1,1%
	506 - 513	R	4	2	2	4	12	13,2%
	513 - 522	R	11	7	10	12	40	44,0%
	523 - 547	R	6	8	12	8	34	37,4%
3'LTR-ADN	9420 - 9433	<i>nef/U3</i> <sup>§</sup>	-	-	-	1	1	1,3%
	9482 - 9493	U3	-	1	-	-	1	1,3%
	9537 - 9544	U3/R	-	-	-	1	1	1,3%
	9590 - 9597	R	4	5	4	3	16	21,3%
	9597 - 9606	R	12	10	3	4	29	38,7%
	9607 - 9631	R	3	4	10	10	27	36,0%
<b>TOTAL</b>			41	38	42	45		
			166					

Les nouveaux points de cassure identifiés par rapport à l'étude des LTR-ARN figurent en rouge.

<sup>§</sup>Point de cassure identifié dans la région chevauchante entre le gène *nef* et la région U3 du 3'LTR

\*Les tirets indiquent que le point de cassure n'a pas été retrouvé dans l'extrait séquencé.



**FIGURE 46** : Localisation et nombre de points de cassure dans les 5' et 3'LTR-ADN.

Les positions des bases correspondent à la référence HxB2 du VIH-1/M (numéro d'accèsion GenBank K03455). Les flèches vertes correspondent aux positions des points de cassure décrits *in vitro*. Les chiffres en rouge représentent le nombre de points de cassure à la position considérée.

Au final, 7 points de cassure supplémentaires ont été identifiés en travaillant spécifiquement sur les 5' et 3'LTR-ADN à la fin de la co-culture (J34) (TABLEAU 20).

Les LTRs étant identiques, nous avons cherché à déterminer la proportion exacte des différents points de cassure, en nous affranchissant de leur origine (5' ou 3'LTR-ADN). L'alignement de toutes les séquences obtenues après SGA sur le 5'LTR de la souche de référence VIH-1/M HxB2 a confirmé que les trois motifs majoritaires de recombinaison retrouvés dans la région R du 5'LTR-ADN (506-513pb, 513-522pb et 523-547pb) étaient les mêmes que ceux retrouvés majoritairement dans la région R du 3'LTR-ADN (9590-9597pb, 9598-9606pb et 9608-9631pb) et qu'ils correspondaient à 95,2% de toutes les séquences analysées (TABLEAU 21, FIGURE 47).

Les 7 nouveaux points de cassure correspondaient finalement à 6 motifs de recombinaison identifiés en dehors de la région R, et représentaient moins de 5% de toutes les séquences analysées. Parmi eux, quatre se concentraient dans la région U3 ou à la jonction U3/R (3,6%) et deux ont été identifiés dans la région U5 (1,2%) (TABLEAU 21).

**TABLEAU 21** : Proportion des points de cassure caractérisés dans les LTR-ADN, 5' et 3'LTRs confondus, à J34 après SGA

Point de cassure (pb)		Localisation	Points de cassure totaux		
5'LTR-ADN	3'LTR-ADN		n	%	
<b>336 - 349</b>	<b>9420 - 9433</b>	<i>nef</i> /U3 <sup>§</sup>	1	0,6%	<b>0,6%</b>
<b>398 - 409</b>	<b>9482 - 9493</b>	U3	1	0,6%	
<b>420 - 443</b>	<b>9504 - 9527</b>	U3	1	0,6%	<b>3,0%</b>
<b>453 - 460</b>	<b>9537 - 9544</b>	U3/R	3	1,8%	
<b>506 - 513</b>	<b>9590 - 9597</b>	R	28	16,9%	
<b>513 - 522</b>	<b>9597 - 9606</b>	R	69	41,6%	<b>95,2%</b>
<b>523 - 547</b>	<b>9607 - 9631</b>	R	61	36,7%	
<b>566 - 573</b>	<b>9650 - 9657</b>	U5	1	0,6%	
<b>587 - 604</b>	<b>9671 - 9688</b>	U5	1	0,6%	<b>1,2%</b>

Les points de cassure identifiés dans le 5'LTR-ADN et/ou dans le 3'LTR-ADN figurent en rouge dans la colonne correspondante.

<sup>§</sup>Point de cassure identifié dans la région chevauchante entre le gène *nef* et la région U3 du 3'LTR

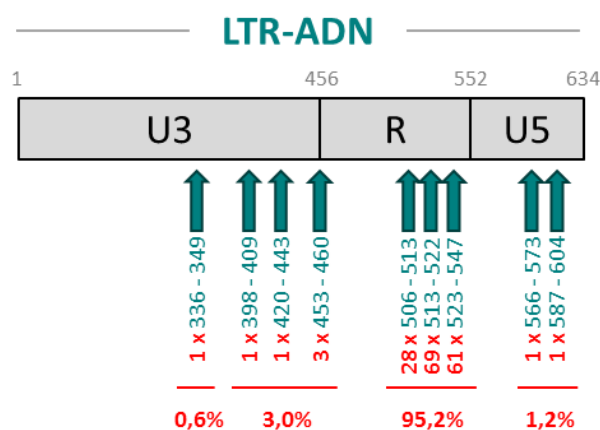


FIGURE 47 : Localisation et fréquence des points de cassure dans les LTR-ADN.

Les positions des bases correspondent à la référence HxB2 du VIH-1/M (numéro d'accèsion GenBank K03455). Les flèches vertes correspondent aux positions des points de cassure décrits *in vitro*. Les chiffres en rouge représentent le nombre de points de cassure à la position considérée, et le pourcentage de points de cassure dans les différentes régions du LTR.

### II-2.3. Recherche de la prédominance d'un profil préférentiel

Les résultats obtenus par SGA ont montré que plus de 95% des séquences analysées avaient un motif de recombinaison dans la zone 506-547pb. Afin d'identifier parmi les trois motifs identifiés (506-513pb, 513-522pb et 523-547pb) celui qui devenait prédominant au cours de la co-culture, nous avons compilé tous les résultats de SGA obtenus sur les 5'LTR-ARN et 5'-3'LTR-ADN.

Ainsi, l'analyse de toutes les données relatives aux 5'LTR-ARN, a montré qu'au début de la co-culture (J6), le motif 513-522pb était présent dans environ 41% des cas, contre 29,5% environ pour chacun des deux autres motifs (506-513pb et 523-547pb) (TABLEAU 22).

TABLEAU 22 : Evolution de la fréquence des trois motifs de recombinaison majoritaires dans les LTRs au cours de la co-culture

Points de cassure (pb)	Localisation	J6	J16	J34	J34
		LTR-ARN	LTR-ARN	LTR-ARN	LTR-ADN
506 - 513	R	29,4%	9,1%	5,0%	16,9%
513 - 522	R	41,2%	45,5%	70,0%	41,6%
523 - 547	R	29,4%	45,5%	25,0%	36,7%

Ensuite, le motif 506-513pb devenait moins fréquent pour atteindre 5% à J34, alors que la fréquence du motif 523-547pb augmentait transitoirement et était retrouvée à 25% à J34. Seule la fréquence du motif 513-522pb progressait continuellement, d'abord légèrement à J16 (45%) puis très nettement en fin de culture (70%)(TABLEAU 22).

De la même façon, l'analyse des données relatives aux LTR-ADN, en fin de co-culture (J34), a montré que le point de cassure localisé aux positions 513-522pb était majoritaire (41,6%), suivi de près par le motif 523-547pb (36,7%). Le dernier motif était nettement moins fréquent (TABLEAU 22).

Ces différents résultats mettent donc en évidence que, quelle que soit la forme génomique (LTR-ARN ou LTR-ADN), le motif 513-522pb semble émerger de façon préférentielle dès le début et devient prédominant à la fin de la co-culture.

### II-3. Comparaison aux données *in vivo*

Nos données de séquençage des LTRs obtenues *in vitro* ont ensuite été comparées à celles des 9 URF\_MO décrites *in vivo*, pour lesquelles un point de cassure avait été mis en évidence dans les LTRs (TABLEAU 23).

**TABLEAU 23** : Localisation des points de cassure dans les LTRs recombinants de 9 URF\_MO

URF_MO	Points de cassure (pb)	Localisation	Fréquence
<b>YBF274</b>	396 - 454	U3/R	
<b>YBF282</b>	396 - 454	U3/R	
<b>REC003/024</b>	396 - 454	U3/R	<b>44,4%</b>
<b>REC107</b>	396 - 454	U3/R	
<b>97CA.MP645</b>	489 - 497	R	<b>11,1%</b>
<b>RBF221</b>	506 - 513	R	
<b>RBF235</b>	506 - 513	R	<b>22,2%</b>
<b>BCF204</b>	513 - 522	R	<b>11,1%</b>
<b>BCF174/212</b>	523 - 547	R	<b>11,1%</b>

La localisation précise des points de cassure dans les URF\_MO à montré l'existence de cinq motifs de recombinaison distincts : 396-454pb (U3/R), 489-497pb (R), 506-513pb (R), 513-522pb (R) et 523-547pb (R) (TABLEAU 23).

Le point de cassure 396-454pb était le plus fréquemment rencontré, puisque présent dans quatre des 9 URF\_MO étudiées (44,4%). Deux URF\_MO présentaient le point de cassure 506-513pb (22,2%). Les trois autres points de cassure étaient chacun présents dans une seule URF\_MO. Ces données montraient donc globalement une relative hétérogénéité de répartition des points de cassure, avec une localisation à la jonction U3/R, mais également dans la première et la deuxième moitié de R. Au total, plus de la moitié des URF\_MO considérées (55,6%) présentaient un point de cassure au sein de la région R (TABLEAU 23).

La comparaison de ces données à nos séquences *in vitro* a montré que les trois motifs, majoritaires en SGA (506-513pb, 513-522pb et 523-547pb), étaient également retrouvés *in vivo* chez deux URF\_MO pour 506-513pb, et une URF\_MO pour 513-522pb et 523-547pb. Inversement, le motif majoritaire 396-454pb retrouvé *in vivo* n'était pas non plus retrouvé *in vitro*. De plus, aucun des points de cassure mis en évidence *in vitro* en dehors de la région R n'était retrouvé *in vivo* (TABLEAU 24, FIGURE 48).

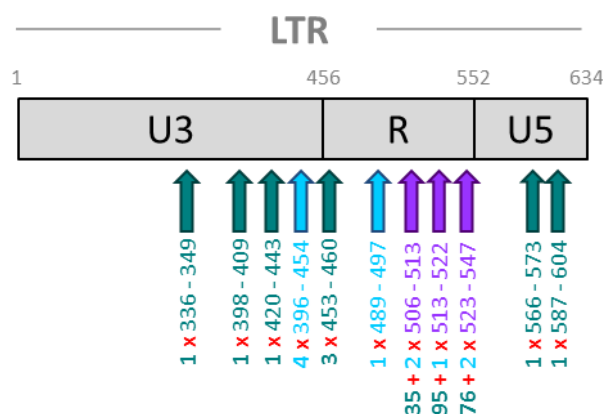
**TABLEAU 24** : Comparaison de la localisation des points de cassure décrits *in vitro* et *in vivo*

Point de cassure (pb)	Localisation	<i>in vitro</i>		<i>in vivo</i>	
		n	%	n	%
336 - 349	<i>nef</i> /U3	1	0,5%	-*	-
398 - 409	U3	1	0,5%	-	-
420 - 443	U3	1	0,5%	-	-
396 - 454	U3/R	-	-	4	44,4%
453 - 460	U3/R	3	1,4%	-	-
489 - 497	R	-	-	1	11,1%
506 - 513	R	35	16,4%	2	22,2%
513 - 522	R	95	44,4%	1	11,1%
523 - 547	R	76	35,5%	1	11,1%
566 - 573	U5	1	0,5%	-	-
587 - 604	U5	1	0,5%	-	-
<b>TOTAL</b>		<b>214</b>		<b>9</b>	

\*Les tirets indiquent que le point de cassure n'a pas été retrouvé.

Les points de cassure communs aux séquences *in vitro* et *in vivo* apparaissent en violet.





**FIGURE 48** : Localisation et fréquence des points de cassure décrits *in vitro* et *in vivo*

Les positions des bases correspondent à la référence HxB2 du VIH-1/M (numéro d'accèsion GenBank K03455). Les flèches vertes correspondent aux points de cassure décrits uniquement *in vitro*. Les flèches bleues correspondent aux points de cassure décrits uniquement *in vivo*. Les flèches violettes correspondent aux points de cassure communs aux séquences *in vitro* et *in vivo*. Les chiffres en vert et en bleu représentent le nombre de points de cassure décrits *in vitro* et *in vivo*, respectivement.

Les fréquences des points de cassure décrits à la fois *in vitro* et *in vivo* n'étaient pas similaires, mais étaient difficilement comparables, du fait de la grande différence des effectifs de LTRs analysés (214 LTRs *in vitro* contre 9 LTRs *in vivo*) (TABLEAU 24).

Etant donné les différences observées *in vitro* et *in vivo*, l'impact de la diversité génétique des souches parentales sur la survenue d'évènements de recombinaison au sein des LTRs peut se poser. L'origine des fragments de VIH-1/M et de VIH-1/O impliqués dans les 9 URF\_MO avec un point de cassure dans les LTRs a été étudiée. Aucun lien n'a été mis en évidence entre la localisation du point de cassure et les sous-types et sous-groupes impliqués (TABLEAU 25).

**TABLEAU 25** : Etude de la diversité des fragments impliqués dans LTRs recombinants des 9 URF\_MO

Points de cassure (pb)	Sous-type VIH-1/M	Sous-groupe VIH-1/O	Effectif
396 - 454	CRF02_AG	H	2
	F	H	1
	D	H	1
489 - 497	CRF02_AG	H	1
506 - 513	CRF02_AG	H	1
	D	H	1
513 - 522	A	H	1
523 - 547	A	H	1

### III- Discussion

Dans notre étude en collaboration avec l'équipe de Strasbourg, portant spécifiquement sur la modélisation du phénomène de recombinaison au sein des LTRs, nous avons émis l'hypothèse d'une recombinaison préférentielle à la jonction entre les régions U3 et R des LTRs [12]. De plus, lors de l'analyse des 19 URF\_MO décrites actuellement, nous avons constaté que, lorsqu'il y a un nombre impair de points de cassure au sein du génome (hors LTRs), les LTRs étaient eux-mêmes recombinants [201].

Le CMIC pVIH-1/OM généré lors de la première partie de ces travaux présentant des LTRs différents, du fait d'un unique point de cassure dans la région codante de son génome (hors LTRs), était l'outil idéal pour étudier et valider expérimentalement ces observations. Nous avons donc recherché *in vitro* l'émergence de points de cassure dans les LTR-ARN à l'issue de la transfection juste avant la co-culture (J0) et entre le début (J4) et le milieu de la co-culture (J16). Nous avons également analysé la localisation précise des points de cassure par SGA à différents temps de culture (J6, J16 et J34), afin de mettre en évidence la région préférentielle de recombinaison et de rechercher un éventuel motif prédominant en fin de culture.

L'ensemble de nos résultats a d'abord permis de confirmer qu'un seul point de cassure présent entre les régions Gag et Nef du génome viral impliquait une recombinaison spontanée, au cours de la réplication virale, entre les LTRs de groupes différents.

L'étude de la cinétique d'émergence a montré que cette recombinaison n'avait pas eu lieu à la fin de la transfection dans les cellules HEK293T. En effet, les génomes des virus VIH-1/OM-GC à la sortie des cellules HEK293T étaient identiques à la construction du CMIC pVIH-1/OM-GC. Ce résultat semblait cohérent, puisque les HEK293T ne possèdent ni le récepteur CD4+, ni les co-récepteurs CXCR4 ou CCR5, nécessaires à rendre sensible une lignée cellulaire à l'infection par le VIH. De plus, ces cellules ne sont pas non plus permissives à l'infection par le VIH, par manque de certains constituants. La recombinaison ne peut avoir lieu sans l'étape de transcription inverse et ce sont donc les cellules Jurkat, sensibles et permissives au VIH, qui permettent de générer des points de cassure dans les LTRs à l'issue d'un cycle de réplication complet du virus.

L'analyse des points suivants de la co-culture (J4 à J13) a montré que les formes recombinantes avec point de cassure dans les LTRs étaient présentes dès J4 et que leur concentration augmentait par la suite.

Les données de séquençage ont alors révélé une co-existence de plusieurs populations virales recombinantes, dont les points de cassure étaient tous localisés dans la seconde moitié de la région R. Les LTR-ARN étant plus courts que les LTR-ADN, et en particulier, les 3'LTR-ARN n'ayant pas de région U5, l'analyse de la seconde moitié de R par PCR s'est avérée difficile pour les 3'LTR-ARN, alors que cette région du LTR semblait principalement concernée par les phénomènes de recombinaison.

L'exploration des différentes populations virales recombinantes était nécessaire. L'outil le plus adaptée aurait été le séquençage à haut débit (Next Generation Sequencing, NGS), mais il n'était pas disponible au laboratoire au moment de l'initiation de ces travaux. Des essais sont actuellement en cours, avec notamment la mise au point d'un outil d'analyse bio-informatique, et devraient permettre de préciser les résultats, tout en allégeant le protocole expérimental. Le recours à la méthode SGA, bien qu'expérimentalement lourde, a constitué une bonne alternative pour préciser la localisation des points de cassure dans les LTRs et y étudier le phénomène de recombinaison spontané. Cette caractérisation a été réalisée à différents points de culture (J6, J16 et J34) pour identifier un éventuel motif préférentiel et caractériser les différentes populations virales présentes. Cette analyse a mis en évidence de multiples motifs de recombinaison dans les différentes régions (U3, R et U5), dont trois (506-513pb, 513-522pb et 523-547pb) représentés très majoritairement dans la région R. L'un d'entre eux en particulier (513-522pb) semble émerger préférentiellement au cours de la co-culture, pour devenir prédominant à la fin de la co-culture représentant 70% et presque 42% des séquences LTR-ARN et LTR-ADN respectivement, à J34. Ce résultat suggère donc un avantage adaptatif et/ou répliatif de cette population, par rapport à celles ayant un autre motif de recombinaison, dans nos conditions expérimentales.

Une fois ces données *in vitro* obtenues, la dernière analyse a consisté à les comparer à nos données *in vivo*, provenant des 9 URF\_MO présentant un point de cassure dans les LTRs. Les trois motifs majoritaires décrits *in vitro* étaient bien retrouvés *in vivo*, validant notre modèle

expérimental. Les fréquences *in vitro/in vivo* de chacun de ces trois points de cassure étaient différentes, mais les effectifs étudiés étaient très déséquilibrés, le SGA ayant généré plus de 200 séquences de LTRs recombinants. Cette comparaison est probablement biaisée par le temps de culture *in vitro* qui est court et la durée d'infection *in vivo* qui est longue et par l'absence de sélection par l'hôte.

Cette comparaison montre également que de nombreux points de cassure ont été retrouvés soit *in vitro*, soit *in vivo*, et dans les trois régions des LTRs (U3, R et U5). L'observation de points de cassure dans toutes les régions des LTRs *in vitro* et dans deux des trois régions (U3 et R) *in vivo* suggère que de multiples profils sont possibles, qu'il n'existe pas de profil préférentiel unique et que la localisation des points de cassure est souche-dépendante. Toutefois nos résultats montrent que parmi les multiples populations générées par recombinaison entre deux souches, l'un des motifs de recombinaison devient quand même préférentiel et prédomine.

Dans notre modèle, impliquant seulement une souche de VIH-1/M sous-type B et une souche de VIH-1/O sous-groupe T, ce sont la région R et le motif 513-522pb qui sont apparus prédominants. Malgré les éléments préliminaires semblant éliminer un lien entre la position du point de cassure et l'origine parentale des fragments, l'impact réel de la diversité des souches VIH-1/M et VIH-1/O impliquées dans une forme recombinante sur la localisation des évènements de recombinaison dans les LTRs reste à explorer.

La mise au point de nos travaux se basait sur l'étude des points de cassure retrouvés dans les LTRs recombinants de virus recombinants (CRFs, URFs et URF\_MO)<sup>[12]</sup>, qui avait mis en évidence une majorité (50%) des LTRs recombinants *in vivo* avec point de cassure à la jonction des régions U3 et R. Selon nos résultats, seuls trois des 166 (1,8%) LTR-ADN recombinants analysés présentaient un point de cassure à la jonction U3/R, sachant que l'analyse de la forme génomique ADN ne permet pas de conclure que les virus recombinants correspondants étaient viables, répliquatifs et infectieux. L'étude précise de la transcription inverse de Cappy et *al.* suggérait pourtant que lorsque les régions Gag et Nef n'impliquaient pas la même souche parentale, comme dans le cadre de nos travaux, la recombinaison devait avoir lieu à cet endroit du LTR, lors du deuxième transfert de brin obligatoire, en intermoléculaire. Cette suggestion reposait également sur le postulat qu'en attribuant au deuxième transfert de brin intermoléculaire la génération des points de cassure en U3/R, le premier transfert de brin avait

lieu en intramoléculaire. Ceci était corroboré par le fait qu'un premier transfert de brin en intermoléculaire (qui pourrait générer les points de cassure en U3/R) se retrouve hautement défavorisé si les séquences de la région R sont divergentes. Dans notre étude, la divergence de 30% entre la région R du VIH-1/M 89.6 et celle du VIH-1/O RBF206 pourrait, en partie, expliquer nos résultats, puisque ce pourcentage est supérieur aux 24,9% de l'étude de Cappy et *al.*.

Ces résultats différents soulèvent donc l'hypothèse de plusieurs modalités de génération des points de cassure au sein des LTRs, qui sont probablement plus diverses que supposé.

En effet, en plus des deux transferts de brin obligatoires lors de la transcription inverse, d'autres transferts de brin facultatifs peuvent avoir lieu et générer également des points de cassure dans toutes les régions du génomes, LTRs compris. Par ailleurs, les quatre URF\_MO, avec points de cassure dans les LTRs, sélectionnées dans l'étude de Cappy et *al.* ne semblent pas représentatives de tous les motifs de recombinaison observés dans cette région. En effet, nos travaux ont montré que parmi toutes les URF\_MO, avec point de cassure dans les LTRs actuellement décrites, plus de 55% montraient un point de cassure au sein de la région R et non à la jonction entre les régions U3 et R. Enfin, les phénomènes de recombinaison sont probablement influencés par la diversité des souches parentales, ainsi que par la pression de sélection de l'hôte *in vivo* qui éliminent toutes les formes générées non adaptées.

## B. Identification d'une co-infection VIH-1/M+O du fait de l'émergence *in vivo* d'un recombinant VIH-1/MO

---

### Valorisation

Ce travail a fait l'objet d'une publication, présentée en ANNEXE 9.

Moisan A., de Oliveira F., Pronier C., Cappy P., Maillard A., Plantier J-C. *In vivo* emergence of an HIV-1/MO recombinant revealed undiagnosed HIV-1/M+O co-infection. *Clin Microbiol Infect.* 2019, *in press.* [doi : 10.1016/j.cmi.2019.08.016](https://doi.org/10.1016/j.cmi.2019.08.016)

### Contexte et objectifs du travail expérimental

Au cours des 20 dernières années, les différentes études épidémiolo-moléculaires menées en France et au Cameroun sur la double infection VIH-1/M+O ont permis la description de cas correspondant à certaines des étapes de la génération d'un recombinant :

- une surinfection par un VIH-1/M chez une patiente VIH-1/O positive, sans détection de recombinant VIH-1/MO associé <sup>[197]</sup>,
- 13 doubles infections VIH-1/M+O sans recherche ou détection de recombinant VIH-1/MO<sup>[150, 153, 156, 185, 196, 198]</sup>,
- 11 co-infections impliquant un recombinant VIH-1/MO et au moins l'une des deux souches parentales VIH-1/M et/ou VIH-1/O <sup>[150, 185, 192-195, 198]</sup>,
- 10 infections par des recombinants VIH-1/MO isolés, sans aucune des souches parentales d'origine, et avec un cas d'infection par deux formes recombinantes différentes chez un patient <sup>[150, 198-200]</sup>.

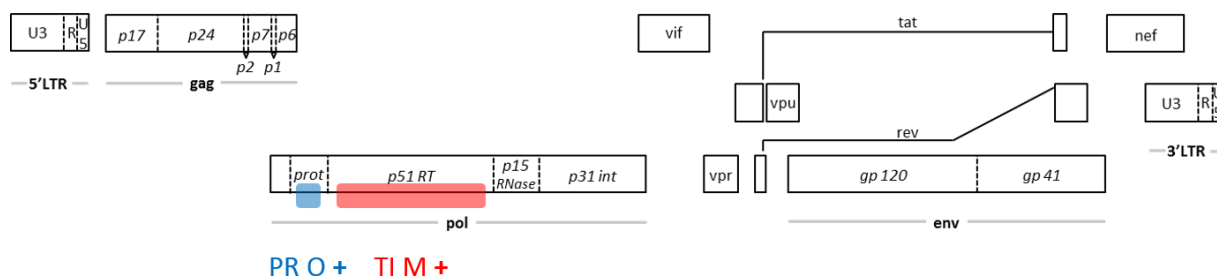
Il est important de noter que, dans la plupart de ces études, la recherche de recombinants VIH-1/MO, lorsqu'elle était réalisée, se basait sur une stratégie ciblant le point chaud de la

recombinaison identifié dans Vpr, avec pour conséquences une sur-représentation de recombinants « Vpr » décrits actuellement et l'existence probable de formes avec d'autres profils de recombinaison.

Jusqu'alors, aucune de ces études n'a décrit la dynamique de genèse *in vivo* d'un recombinant VIH-1/MO et, plus particulièrement, l'émergence d'un recombinant VIH-1/MO chez un patient doublement infecté.

En effet, les données séquentielles concernant ces cas d'infections atypiques sont extrêmement rares, les patients étant peu nombreux et souvent rapidement mis sous traitement ARV efficace ou étant perdus de vue après le dépistage. Une autre difficulté concerne l'identification de ces situations de double infection et/ou leur suivi et la mise en évidence de cette dynamique de recombinaison. En effet, au début de l'épidémie à VIH, les techniques séro-moléculaires commerciales de dépistage et de suivi avaient été conçues à partir de panels de VIH-1/M, conduisant à des faux négatifs, à des sous-quantifications majeures et donc à la suspicion de variants quels qu'ils soient. Désormais, la situation s'est nettement améliorée, avec des outils globalement adaptés à la diversité génétique, grâce au développement de trousses non-spécifiques de groupes <sup>[162, 226, 232-234]</sup>. Mais, paradoxalement, cette situation engendre des difficultés de dépistage de ces variants divergents. En effet, la détection des doubles infections VIH-1/M+O et des formes recombinantes VIH-1/MO, jusqu'alors basée sur l'observation de discordances lors du dépistage et/ou du suivi, en lien avec les défauts de détection des variants de groupe O, est devenue davantage compliquée. A présent, il faut aller jusqu'au génotypage de résistance par séquençage, qui reste spécifique du groupe M, pour suspecter la présence de plusieurs variants et/ou d'une forme recombinante VIH-1/MO.

C'est dans ce contexte que se situe cette étude. En 1999, une femme née au Tchad, dans une région proche de la frontière camerounaise, a été diagnostiquée VIH-1 à l'hôpital de Cochin, à Paris. La cause présumée de sa contamination par le VIH était une transfusion sanguine, lors d'un accouchement au Tchad en 1996. A partir de 2003, son suivi immuno-virologique a été réalisé, de manière régulière, au CHU de Rennes. En novembre 2013, un génotypage de résistance, réalisé avec la trousse ANRS <sup>[235]</sup>, a montré un polymorphisme différent des antériorités, qui correspondaient à un VIH-1/M, en révélant la présence de VIH-1/O dans la PR et de VIH-1/M dans la TI (FIGURE 49).



**FIGURE 49** : Discordance observée lors du génotypage de résistance réalisé par le laboratoire de virologie du CHU de Rennes.

Cette discordance moléculaire suggérait un cas d'infection atypique avec présence d'une forme recombinante VIH-1/MO, conduisant à l'envoi de prélèvements (plasma et culot globulaire) au CNR VIH de Rouen pour exploration.

Les objectifs de ce travail étaient donc, dans un premier temps, de caractériser les différentes populations et en particulier de rechercher la présence d'une forme recombinante VIH-1/MO, et, dans un second temps, de déterminer la dynamique virale chez cette patiente.

## I- Méthodes

### I-1. Analyses sur le prélèvement de novembre 2013

#### I-1.1. Recherche de variants par sérotypage

La première étape de l'exploration de cette discordance reposait sur une technique de sérotypage, ciblant la région de l'enveloppe [236, 237]. Cette technique utilise différents peptides de la gp41, représentatifs des différentes lignées de VIS et de VIH connues, orientant vers un dépistage de lignée (VIH-1/VIScpz/gor et/ou VIH-2/VISmm), et des peptides V3 de la gp120, permettant l'identification spécifique du ou des variants infectants (FIGURE 50). Etant donné la présence de VIH-1/M et de VIH-1/O lors du génotypage de résistance, l'objectif, à ce stade, était de déterminer la réactivité du sérum de novembre 2013, pour identifier une éventuelle double infection VIH-1/M+O.



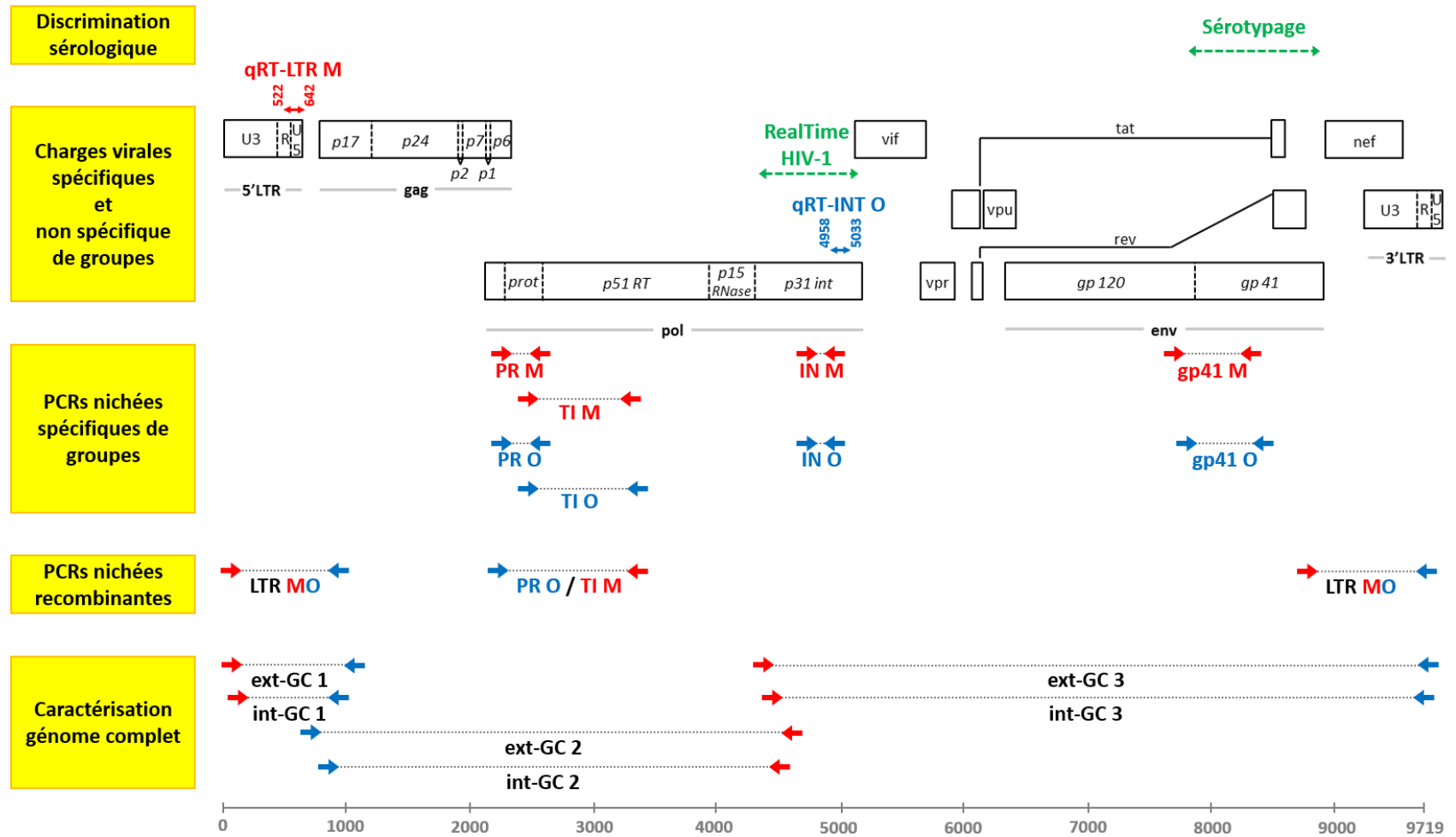


FIGURE 50 : Régions ciblées par les méthodes d'exploration du CNR

Le sérotypage, en vert, cible le gène *env*. La charge virale RealTime HIV-1 non spécifique de groupes est représentée en vert et cible la région de l'*IN*. Les charges virales spécifiques de groupes M (qRT-LTR M) et O (qRT-INT O) figurent en rouge et bleu, respectivement. Seules les amorces des PCRs externes ont été représentées. Les amorces s'hybridant sur du VIH-1/M sont symbolisées par des flèches rouges et celles s'hybridant sur du VIH-1/O par des flèches bleues. Lors de la caractérisation du génome complet, la PCR externe amplifiait le fragment XL-ext-GC et la PCR interne amplifiait deux fragments recombinants chevauchants : int-GC 1 et int-GC 2. La localisation des différentes régions ciblées du génome est basée sur le génome de la souche de référence VIH-1/M HxB2 (numéro d'accèsion GenBank K03455).

## I-1.2. Confirmation moléculaire de double infection avec recombinant

### a) PCRs nichées spécifiques de groupes M et O

La confirmation moléculaire de co-infection a été réalisée sur l'ARN préalablement extrait du prélèvement de novembre 2013 et reposait sur un système de PCRs nichées, utilisant des couples d'amorces spécifiques des groupes M et O et ciblant deux gènes (*pol* et *env*) et quatre régions du génome de VIH-1 : PR, TI, IN et gp41 (FIGURE 50). Cette méthodologie reprenait les couples d'amorces spécifiques du VIH-1/M mis au point par l'ANRS dans le cadre de la recherche de mutations de résistance dans les gènes *pol* (PR M et TI M) et *env* (gp41 M) [238, 239] et les couples d'amorces spécifiques du VIH-1/O (PR O, TI O, gp41 O) mis au point par le CNR VIH de Rouen [149]. Concernant les PCRs nichées IN M et IN O, les couples d'amorces utilisés étaient ceux décrits par Heyndrickx et al. [240].

Après dépôt sur gel d'agarose, tous les amplicons positifs étaient purifiés, puis séquencés par la méthode de Sanger, comme décrit précédemment. Les séquences obtenues étaient analysées et alignées dans le logiciel MEGA, en prenant comme souches de référence des variants de groupes M sous-type D et O sous-groupe H, phylogénétiquement proches de ceux identifiés lors du génotypage de résistance réalisé à Rennes.

### b) PCRs nichées spécifiques des zones de recombinaison

Compte tenu des résultats discordants lors du génotypage de résistance, avec présence de PR O et TI M, la présence d'un point de cassure dans la région PR/TI était suspectée. Afin de rechercher la présence d'un éventuel recombinant VIH-1/MO, une PCR nichée hétérologue, amplifiant une région recombinante PR O/TI M a été mise au point et réalisée. Pour cela, des amorces spécifiques du VIH-1/O en 5' et du VIH-1/M en 3' ont été utilisées (FIGURE 50). De plus, du fait de l'absence d'autres points de cassure dans les régions ciblées par les PCRs nichées spécifiques de groupes M et O, mais aussi du fait des résultats de nos travaux sur la localisation des points de cassure, l'hypothèse d'une recombinaison dans les LTRs a été avancée. Ainsi, afin de confirmer cette hypothèse, les 5' et 3' LTRs ont été amplifiés par le même système de PCRs nichées hétérologues que décrit en I-1.1. Recherche d'un point de cassure dans les LTRs (p.133) et en utilisant les mêmes amorces spécifiques du motif de recombinaison MO (FIGURE 50).

Tous les amplicons générés ont été déposés sur gel d'agarose et purifiés, avant d'être séquencés par la méthode de Sanger. Les séquences obtenues ont été alignées aux séquences de référence utilisées pour le séquençage des PCRs nichées spécifiques de groupes M et O. La

localisation des éventuels points de cassure dans la région PR/TI et dans la région des LTRs a été ensuite définie, par rapport à la souche de référence VIH-1/M HxB2 (numéro d'accèsion GenBank K03455).

### c) Séquençage de génome complet de la forme recombinante MO

Le génome complet de la forme recombinante VIH-1/MO a été séquencé, afin de caractériser le profil de recombinaison et d'identifier d'éventuels autres points de cassure.

La stratégie utilisée, adaptée d'une méthodologie précédemment décrite [7], reposait sur des PCRs nichées chevauchantes hétérologues, réalisées sur l'ADN préalablement extrait du culot globulaire de novembre 2013, avec l'automate *EZ1 advanced XL* et le kit *EZ1® Virus Mini Kit v2.0 (Qiagen)*, selon les recommandations du fabricant. Cette stratégie permettait d'amplifier uniquement la forme recombinante VIH-1/MO sans amplifier les formes parentales VIH-1/M et/ou VIH-1/O.

Trois fragments chevauchants, couvrant l'intégralité du génome ont pour cela d'abord été amplifiés par PCR externe (FIGURE 50). Le premier fragment, ext-GC 1, de 1601 pb, allait du début du 5'LTR au début de *gag* et couvrait le point de cassure identifié dans les LTRs de motif MO. Les amorces utilisées étaient donc spécifiques du groupe M en 5' et du groupe O en 3'. Le second fragment, ext-GC 2, de 4075 pb, allait de début de *gag* au début de l'IN et couvrait le point de cassure situé dans la région PR/TI de motif OM. Les amorces utilisées étaient spécifiques du groupe O en 5' et du groupe M en 3'. Le dernier fragment, ext-GC 3, de 5347 pb, allait du début de l'IN à la fin du 3'LTR et couvrait le point de cassure identifié dans les LTRs de motif MO. Les amorces utilisées étaient donc spécifiques du groupe M en 5' et du groupe O en 3'. Les PCRs internes utilisaient les mêmes types de couples d'amorces hétérologues qu'utilisés lors des PCRs externes et devaient amplifier les fragments chevauchants int-GC 1, int-GC 2 et int-GC 3 de 999 pb, 3880 pb et 5158 pb, respectivement (FIGURE 50). Les PCRs nichées pour amplifier les fragments GC 1 ont été réalisées à l'aide du kit *HotStartaq MasterMix (Qiagen®)* et le mélange réactionnel et le protocole d'amplification sont présentés en ANNEXE 7. Etant donné la longueur importante des fragments GC 2 et GC 3 à amplifier, le kit *Long Range PCR® (Qiagen)* a été utilisé. Les mélanges réactionnels et les protocoles d'amplification étaient les mêmes, quel que soit le fragment amplifié (ANNEXE 8).

Les produits de PCR ont été déposés sur gel, purifiés puis séquencés par la méthode de Sanger. Une fois analysées, les séquences obtenues à partir des amplicons ont été comparées

à des banques de séquences de référence, en procédant à une étude de similarité à l'aide des programmes *HIV BLAST* ([https://www.hiv.lanl.gov/content/sequence/BASIC\\_BLAST/basic\\_blast.html](https://www.hiv.lanl.gov/content/sequence/BASIC_BLAST/basic_blast.html)) et *Retrovirus Genotyping Tool* (<https://www.ncbi.nlm.nih.gov/projects/genotyping/formpage.cgi>), disponibles en ligne. Ces deux programmes reposent sur une méthode utilisant des fenêtres chevauchantes et coulissantes et permettent une analyse par petits fragments, basée sur leur similarité avec les séquences de référence de VIH-1 proposées par l'outil.

La présence des points de cassure suspectés lors de cette analyse a été confirmée par le logiciel SimPlot <sup>[241]</sup>, selon une fenêtre de 200 nucléotides et un déplacement de 10 nucléotides. La localisation de ces points de cassure a été précisée par un alignement dans le logiciel MEGA, en prenant comme souches de référence des variants de groupes M sous-type D et O sous-groupe H, identifiés comme phylogénétiquement proches par l'outil HIV BLAST, ainsi que la souche de référence HxB2. Pour le VIH-1/M sous-type D, les souches suivantes ont été sélectionnées : D.CM.2001.AY371156, D.CM.2001.AY371155, D.TD.1999.AJ488926, D.YE.2002.AY795907 et D.CD.1984.U88822. Pour le VIH-1/O sous-groupe H, les souches suivantes ont été sélectionnées : O.CM.1999.AY169815, O.CM.1998.AY169812, O.SN.1999.AJ302646, O.SN.1998.AJ300450 et O.CM.-.AY618998.

Le profil de recombinaison obtenu pour RBF235 a été comparé à ceux des autres URF\_MO, afin de rechercher d'éventuels points de cassure communs.

Une analyse phylogénétique a ensuite été réalisée, grâce à la construction d'un arbre par la méthode du « plus proche voisin », ou neighbor-joining <sup>[242]</sup>, afin d'exclure tout lien avec une URF\_MO préalablement décrite. Pour cela, le modèle de Kimura à deux paramètres a été utilisé, lors du calcul des distances génétiques <sup>[243]</sup>, et un test de « bootstrap » a été effectué afin de valider la robustesse de la topologie des arbres générés.

Enfin, le schéma du génome complet avec le positionnement des points de cassure du recombinant a été dessiné grâce au programme en ligne *Recombinant HIV-1 Drawing Tool* ([https://www.hiv.lanl.gov/content/sequence/DRAW\\_CRF/recom\\_mapper.html](https://www.hiv.lanl.gov/content/sequence/DRAW_CRF/recom_mapper.html)).

### I-1.3. Quantification des différentes populations virales

Afin d'évaluer le niveau de répllication des différentes populations virales (VIH-1/M et/ou VIH-1/O et/ou VIH-1/MO), des charges virales ont été réalisées sur le plasma avec une

technique non spécifique de groupes (*RealTime HIV-1, Abbott*) et avec des techniques spécifiques de groupe M, ciblant la région LTR (qRT-LTR M, *Generic HIV, Biocentric, Bandol, France*), et de groupe O, ciblant la région de l'IN (qRT-INT O, CNR VIH de Rouen <sup>[244]</sup>) (FIGURE 50, p. 164).

## I-2. Analyses sur les prélèvements séquentiels et collecte des données immuno-virologiques et thérapeutiques

### I-2.1 Echantillons et analyses

Afin de reconstituer séquentiellement la chronologie infectieuse de cette patiente et de caractériser les populations virales au cours du temps, des prélèvements réplicatifs antérieurs à novembre 2013 ont été demandés au CHU de Rennes.

La présence des souches parentales VIH-1/M et VIH-1/O et recombinante VIH-1/MO a été recherchée sur quatre échantillons d'ARN positifs supplémentaires, prélevés entre octobre 2007 et janvier 2013. La même méthodologie que décrite précédemment (sérotypage, quantification et confirmation moléculaire) a été appliquée aux quatre nouveaux échantillons reçus (octobre 2007, novembre 2009, novembre 2011 et janvier 2013).

### I-2.2 Collecte des données immuno-virologiques et thérapeutiques

Pour comprendre l'évolution virologique chez cette patiente, les données de suivi immuno-virologique et de prise en charge thérapeutique (valeurs de charge virale, taux de CD4, génotype de résistance et lignes de traitement) ont été collectées auprès des laboratoires de Virologie de l'Hôpital Cochin, où a été diagnostiquée et suivie la patiente entre 1999 et 2003, et du CHU de Rennes, où est suivie la patiente depuis 2003. Les trousseaux commerciaux de charges virales utilisés durant les 20 ans de suivi biologique de cette patiente, ainsi que leur spécificité de groupes sont présentées dans le TABLEAU 26.

Les données immunologiques de taux de CD4 ont été également été collectées et analysées. Les différentes lignes thérapeutiques successives, ainsi que les périodes de dysobservance de la patiente ont été analysées rétrospectivement. Enfin, les génotypes de résistance cumulés au cours du temps pour chacune des trois populations et le génotype cumulé des trois populations ont été réalisés, en 2018, chez la patiente.

TABLEAU 26 : Caractéristiques des techniques de charge virale utilisées au cours du suivi de la patiente.

Technique de charge virale (Fabricant)	Caractéristiques				Population quantifiable		
	Principe	Seuil Log <sub>10</sub> (cp/mL)	Spécificité de groupes	Région ciblée (gène)	Parentale VIH-1/M	Parentale VIH-1/O	Recombinante VIH-1/OM
<b>RealTime HIV-1</b> (Abbott)	qRT-PCR	1,6	non spécifique	IN ( <i>pol</i> )	+ <sup>§</sup>	+	+
<b>Xpert HIV-1 viral load</b> (Cepheid)	qRT-PCR	1,6	non spécifique	LTR	+	+	+
<b>Veris HIV-1 assay</b> (Beckman Coulter Inc.)	qRT-PCR	1,5	non spécifique	5'LTR	+	+	+
<b>Cobas TaqMan HIV-1 V2</b> (Roche)	qRT-PCR	1,3	non spécifique	MA-CA ( <i>gag</i> ) + 5'LTR	+	+	+
<b>Amplicor Monitor V1.5</b> (Roche)	RT-PCR	1,7	groupe M	<i>gag</i>	+	-	- <sup>§</sup>
<b>Cobas TaqMan HIV-1 V1</b> (Roche)	qRT-PCR	1,6	groupe M	MA-CA ( <i>gag</i> )	+	-	-
<b>Versant HIV-1 RNA 3.0</b> (Bayer Corporation)	bDNA	1,7	groupe M	<i>pol</i>	+	-	+/- <sup>§</sup>
<b>qRT-LTR M</b> (Biocentric)	qRT-PCR	2,3	groupe M	LTR	+	-	-
<b>qRT-INT O</b> (CNR*)	qRT-PCR	1,6	groupe O	IN ( <i>pol</i> )	-	+	-

\*Technique de charge virale mise au point le CNR VIH de Rouen

<sup>§</sup>Le symbole + signifie que la population est quantifiée par la technique

<sup>§</sup>Le symbole - signifie que la population n'est pas quantifiée par la technique

<sup>§</sup>Le symbole +/- signifie que la présence du point de cassure dans la région ciblée par la technique peut avoir un impact sur la quantification.

## II- Résultats

### II-1. Analyses sur le prélèvement de novembre 2013

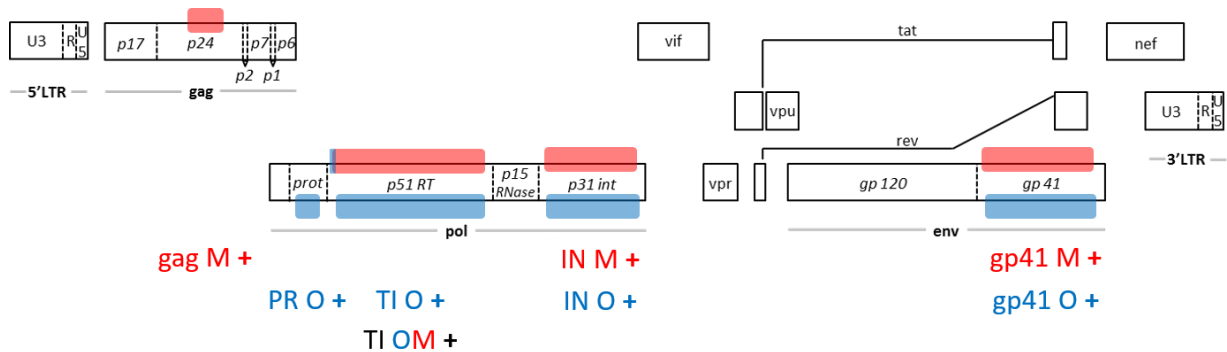
#### II-1.1. Recherche de variants par sérotypage

Le sérotypage montrait une double réactivité M+O, suggérant une co-infection par un VIH-1/O et/ou VIH-1/M et/ou VIH-1/MO. Les analyses moléculaires ont ensuite permis de confirmer une co-infection avec une forme recombinante VIH-1/MO.

#### II-1.2. Confirmation moléculaire de double infection avec recombinant

##### a) PCRs nichées spécifiques de groupes M et O

A partir de l'extrait d'ARN de novembre 2013, les quatre PCRs nichées spécifiques de groupe O, PR O, TI O, IN O, et gp41 O, étaient positives et confirmées VIH-1/O par séquençage. En revanche, seules les PCRs nichées IN M et gp41 M se sont révélées positives et confirmées VIH-1/M par séquençage. Les PCRs nichées PR M et TI M étaient bien positives mais, une fois séquencés, les amplicons correspondaient à du VIH-1/O et à du VIH-1/OM, respectivement. Du fait de l'absence de VIH-1/M dans la région PR/TI, une PCR spécifique de groupe M ciblant une petite région de Gag s'est révélée positive, confirmée VIH-1/M par séquençage. (FIGURE 51).



**FIGURE 51** : Résultats des PCRs nichées spécifiques de groupes.

Les régions de VIH-1/M et de VIH-1/O présentes à ce stade figurent en rouge et en bleu, respectivement.

A ce stade, ces résultats, combinés aux résultats du génotypage de résistance réalisé au CHU de Rennes, étaient en faveur d'une co-infection par deux souches parentales VIH-1/O et VIH-1/M associées à une souche recombinante VIH-1/MO.

### b) PCRs nichées spécifiques des zones de recombinaison

Afin de confirmer la présence d'une forme recombinante VIH-MO, les PCRs nichées spécifiques des zones de recombinaison hypothétiques PR O/TI M et LTR MO ont été réalisées sur l'ADN et ont toutes généré un amplicon. L'analyse du séquençage a permis de confirmer la présence d'un point de cassure de motif OM dans la TI et d'en identifier un autre, de motif MO, dans les 5' et 3'LTRs (FIGURE 52).

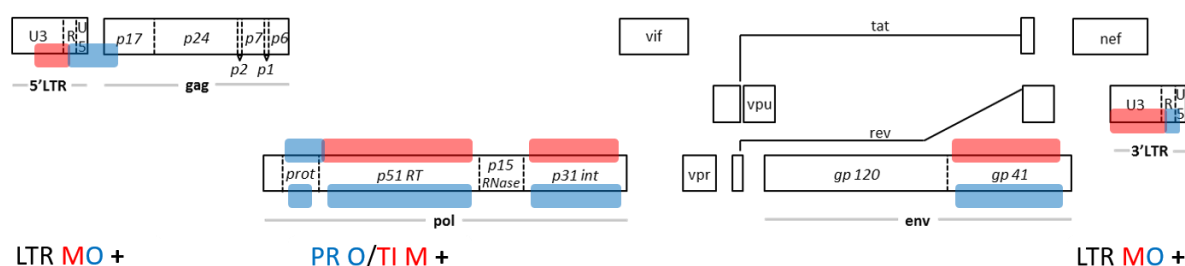


FIGURE 52 : Résultats combinés des PCRs nichées spécifiques de groupes et hétérologues.

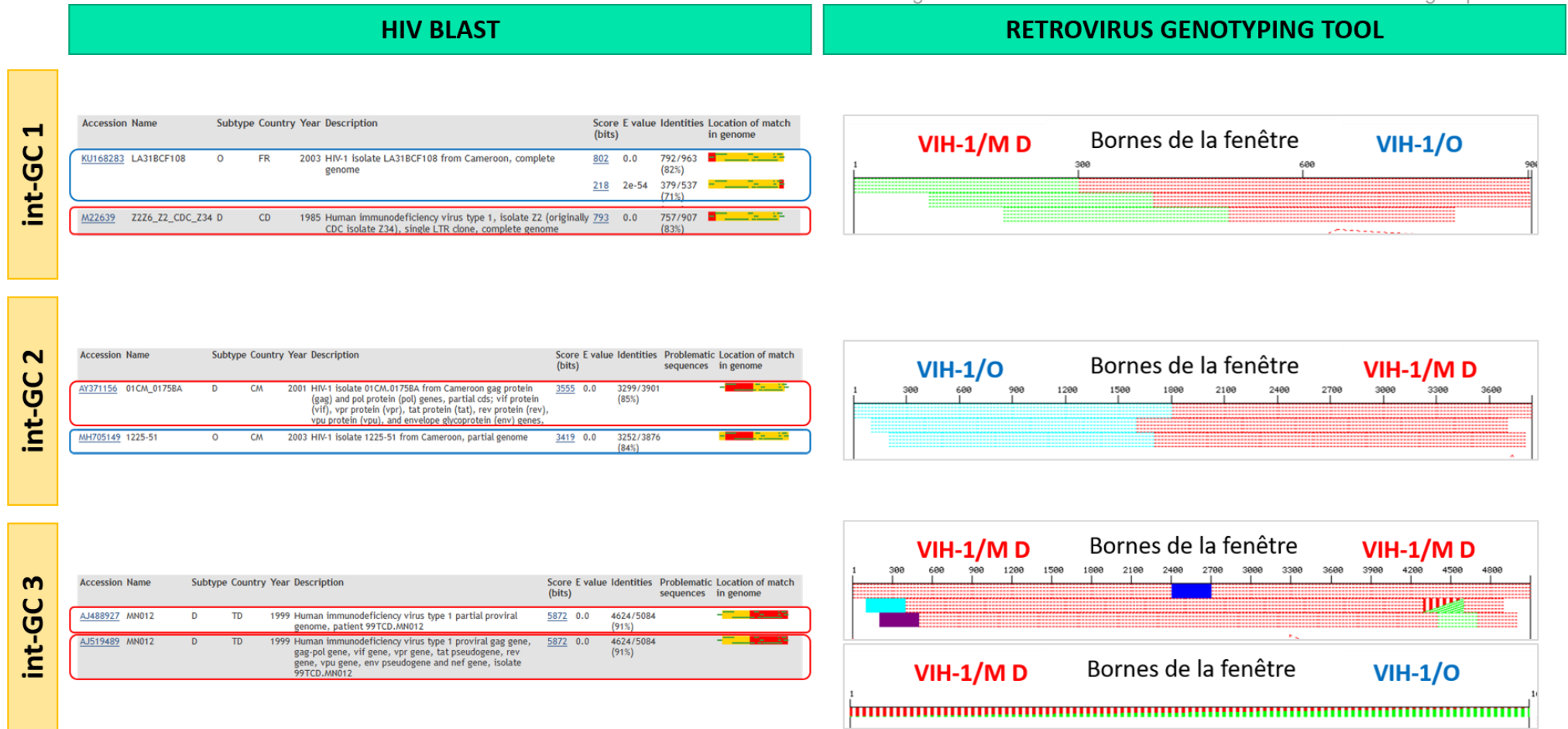
Les régions de VIH-1/M et de VIH-1/O, présentes à ce stade des explorations, figurent en rouge et en bleu, respectivement.

### c) Séquençage de génome complet de la forme recombinante MO

Le séquençage des trois fragments chevauchants, amplifiés à partir de l'ADN intracellulaire et couvrant l'intégralité du génome de la forme recombinante, a révélé la présence de séquences de VIH-1/M sous-type D et de VIH-1/O dans les fragments int-GC 1 et int-GC 2, résultats concordants avec les résultats précédents obtenus à partir des régions (FIGURE 53). Le programme Retrovirus Genotyping Tool a confirmé la présence de deux points de cassure, le premier de motif MO dans le 5'LTR du fragment int-GC 1 et le second, de motif OM dans la région de la TI du fragment int-GC 2 (FIGURE 53).

Du fait de la présence d'un point de cassure dans le 5'LTR, une analyse plus fine a été réalisée pour le fragment int-GC 3, allant de l'intégrase au 3'LTR. Les tailles des fenêtres chevauchantes étudiées avec les deux programmes (HIV BLAST et Retrovirus Genotyping Tool) ont été réduites et ont permis de mettre en évidence la présence de VIH-1/M sous-type D sur la plus grande partie du fragment et de VIH-1/O à la toute fin du 3'LTR, avec présence d'un point de cassure de motif MO dans le 3'LTR, comme attendu (FIGURE 53).





**FIGURE 53** : Résultats de l'analyse de similarité réalisée sur les trois fragments chevauchants de RBF235.

**HIV BLAST.** Pour chacun des fragments, les souches de VIH-1/M et de VIH-1/O, présentant une similarité avec RBF235, sont encadrées en rouge et en bleu, respectivement. **RETROVIRUS GENOTYPING TOOL.** Les parties supérieures des graphes montrent, pour chacune des fenêtres coulissantes, le code couleur correspondant à la séquence de référence présentant le score de similarité le plus élevé. S'il existe deux scores identiques pour une même fenêtre, la barre de la fenêtre est divisée en diagonale. L'analyse de similarité a été répétée plus finement sur la dernière portion du fragment int-GC 3 avec le programme Retrovirus Genotyping Tool, en ne prenant en compte que les 166 dernières pb, afin de mettre en évidence la présence de VIH-1/O à sa toute fin.

La représentation graphique, générée par le logiciel SimPlot, a confirmé les résultats de séquençage des PCRs spécifiques de groupes et hétérologues, d'une part, et ceux des analyses de similarité, d'autre part. Deux points de cassure étaient bien présents dans les LTRs et dans la TI et aucun point de cassure supplémentaire n'a été identifié (FIGURE 54).

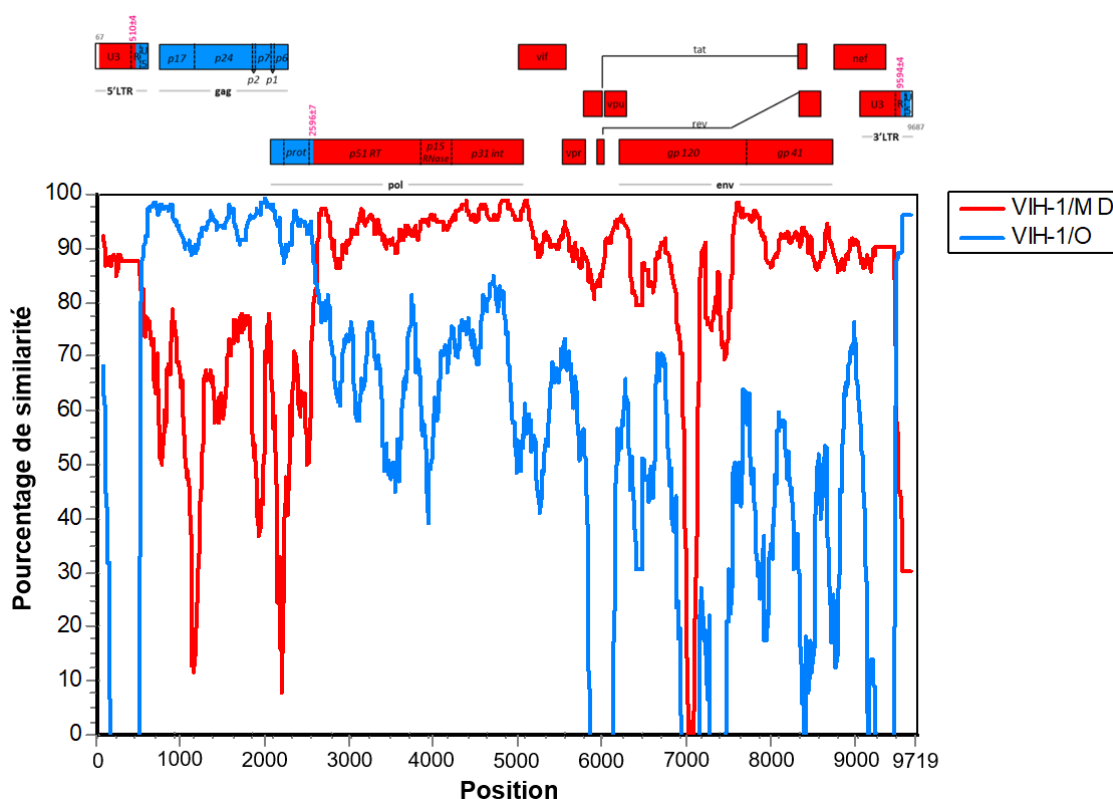


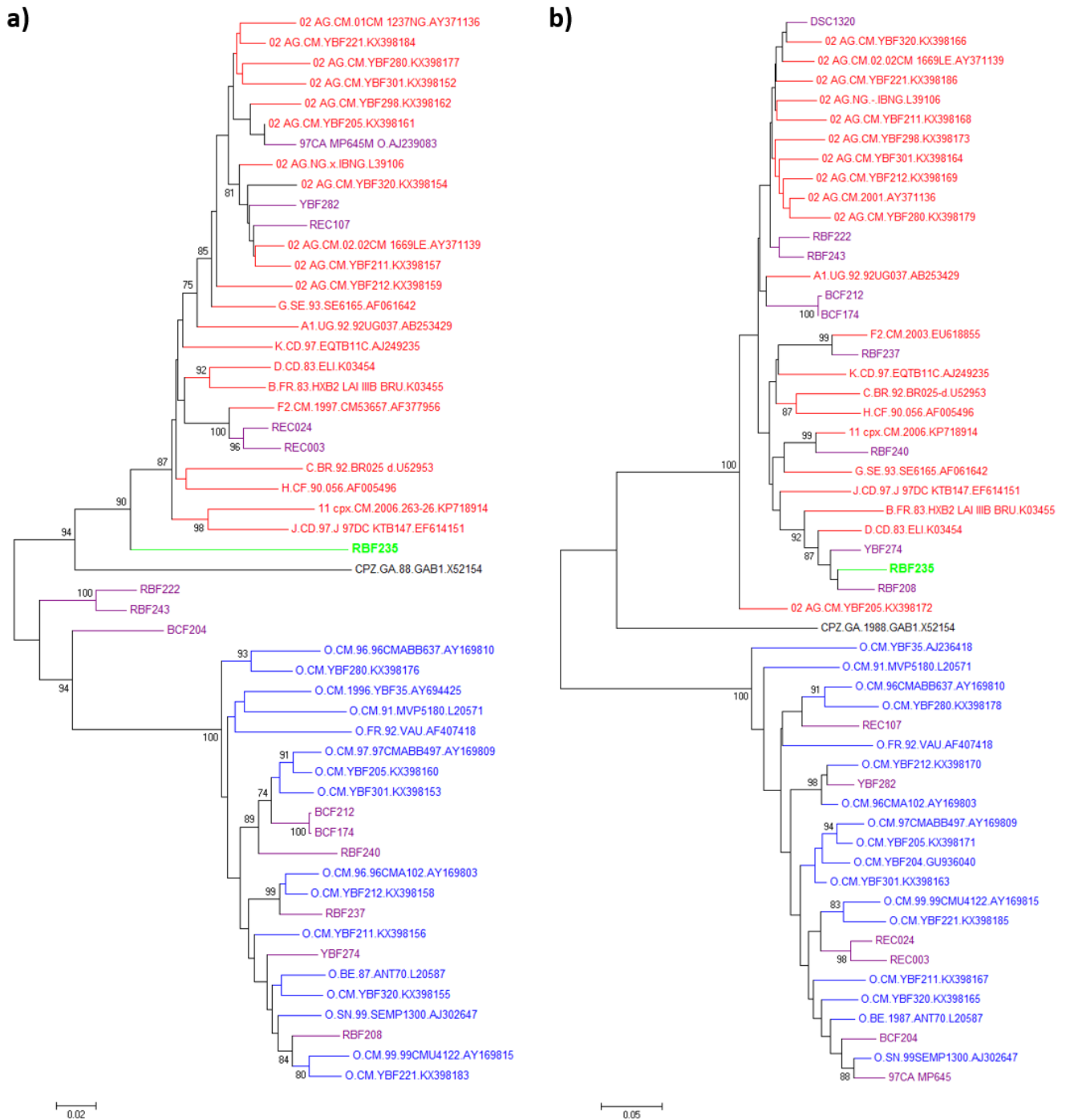
FIGURE 54 : Simplot de la séquence du génome complet de RBF235.

L'analyse a été réalisée avec la souche de VIH-1/M sous-type D de numéro d'accèsion GenBank U88822 et avec la souche de VIH-1/O de numéro d'accèsion GenBank AJ302646. Les régions présentant un pourcentage important de similarité avec le VIH-1/M et le VIH-1/O figurent en rouge et en bleu, respectivement.

Après alignement de la séquence recombinante complète consensus aux séquences des souches phylogénétiquement proches, la localisation des points de cassure a été précisée : 506-513pb et 9590-9597pb pour les 5' et 3'LTRs, respectivement, et 2589-2603 pb pour la TI, en prenant comme référence la souche VIH-1/M HxB2.

L'analyse phylogénétique a, quant à elle, révélé qu'il n'y avait aucun lien entre cette souche recombinante et une URF\_MO déjà décrite, justifiant de la considérer comme une nouvelle URF\_MO (FIGURE 55).

La synthèse des résultats obtenus lors la caractérisation du génome complet de la souche recombinante VIH-1/MO est présentée en FIGURE 56.



**FIGURE 55 :** Arbres phylogénétiques positionnant RBF235 parmi les URF\_MO décrites.

(a) Arbre phylogénétique représentant la région PR-TI (899 pb) de *pol* de la souche recombinante RBF235 (en vert), des URF\_MO complètement séquencées (en violet) et de souches représentatives de VIH-1/M (en rouge) et de VIH-1/O (en bleu). (b) Arbre phylogénétique représentant la région gp41 (443 pb) de *env* de la souche recombinante RBF235 (en vert), des URF\_MO complètement séquencées (en violet) et de souches représentatives de VIH-1/M (en rouge) et de VIH-1/O (en bleu).

Une séquence de VIScpz GAB a été utilisée pour raciner l'arbre. Les analyses phylogénétiques ont été réalisées en utilisant le logiciel MEGA : les distances évolutives ont été calculées à l'aide de la méthode de Kimura à deux paramètres. Mille répétitions des bootstraps ont été réalisées pour évaluer la fiabilité de l'ordre de branchement. Seules les valeurs significatives (>70) de bootstrap figurent.

Partie 2. Etude de l'émergence *in vitro* et *in vivo* des formes recombinantes inter-groupes M et O

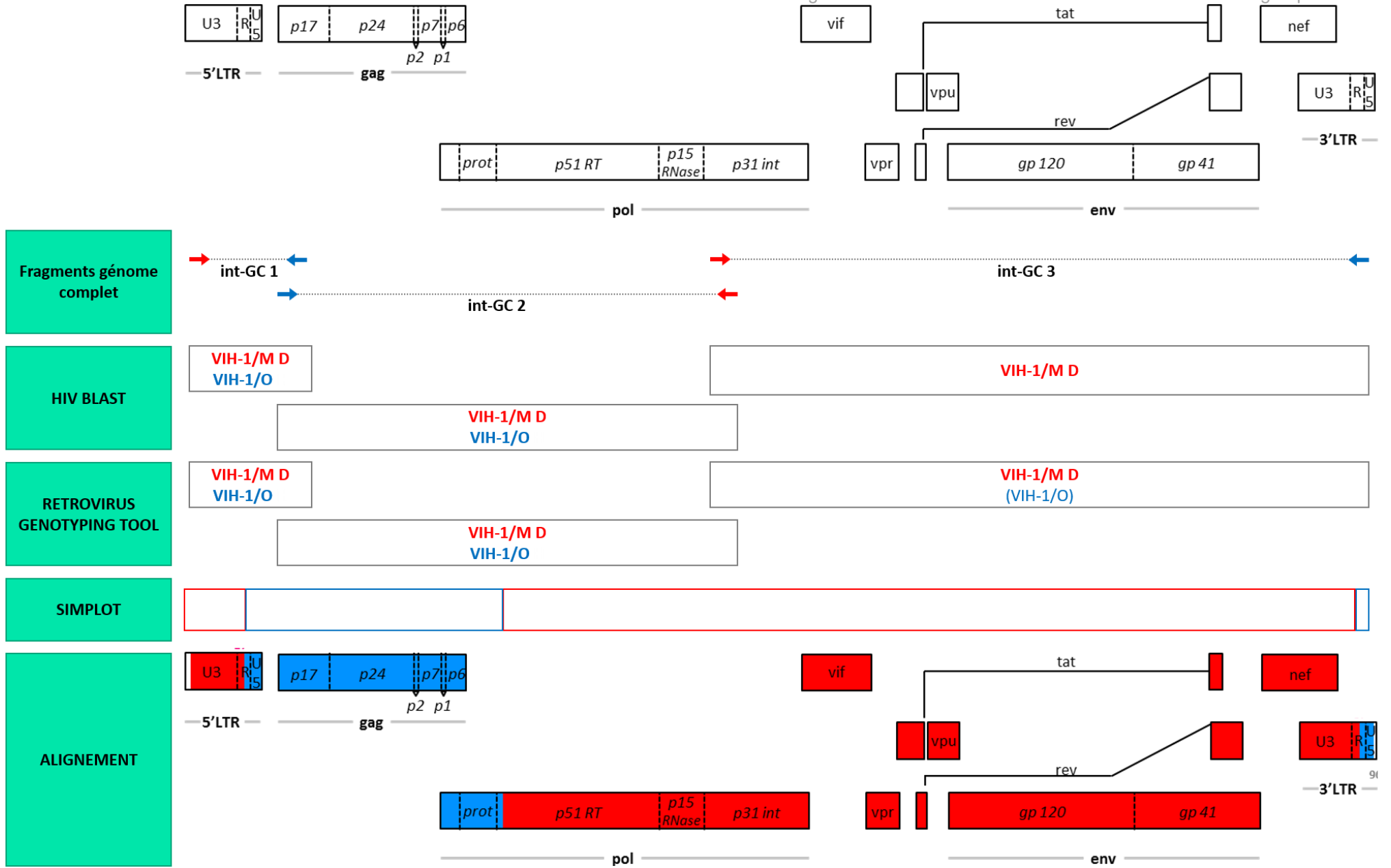


FIGURE 56 : Synthèse des résultats du séquençage de génome complet de RBF235.

### II-1.3. Quantification des différentes populations virales

Les charges virales non spécifique de groupes (RealTime HIV-1) et spécifiques de groupes M (qRT-LTR M) et O (qRT-INT O) étaient toutes quantifiables (TABLEAU 27), sur ce prélèvement de novembre 2013. Associée aux données moléculaires précédentes, l'utilisation combinée de ces trois techniques différentes de charge virale nous a permis de distinguer les espèces répliquatives dans le prélèvement.

En effet, la forme recombinante VIH-1/MO possédant une région IN issue de la souche parentale VIH-1/M, la charge virale spécifique de groupe O positive (3,7 Log<sub>10</sub> cp/mL), ciblant l'IN, témoignait de la répllication de la souche parentale VIH-1/O. Compte tenu du point de cassure de motif MO dans la région des LTRs de la souche recombinante VIH-1/MO, la charge virale spécifique de groupe M, ciblant une région du LTR en aval du point de cassure, ne quantifiait que la souche parentale VIH-1/M. La positivité de la qRT-LTR M (3,2 Log<sub>10</sub> cp/mL) révélait donc la répllication de la souche parentale VIH-1/M. Les charges virales spécifiques de groupes M et O reflétaient donc, au moins, une double infection VIH-1/M+O répllicative. Les PCR spécifiques des zones de recombinaison ayant mis en évidence la présence d'une souche recombinante VIH-1/MO, la charge virale non spécifique de groupes, ciblant l'IN et elle aussi positive (5,5 Log<sub>10</sub> cp/mL), s'expliquait donc par la répllicativité des souches parentales VIH-1/O et VIH-1/M et recombinante VIH-1/MO, en novembre 2013 (TABLEAU 27).

**TABLEAU 27** : Valeurs de charges virales du prélèvement de novembre 2013.

Date de prélèvement	Technique de charge virale		
	RealTime HIV-1 (Abbott)	qRT-INT O (CNR)	qRT-LTR M (Biocentric)
04/11/2013	5,5 Log <sub>10</sub> (cp/mL)	3,7 Log <sub>10</sub> (cp/mL)	3,2 Log <sub>10</sub> (cp/mL)

A ce stade des explorations, nous avons donc mis en évidence la présence de deux souches parentales VIH-1/M et VIH-1/O et d'une souche recombinante VIH-1/MO, chez une patiente connue VIH-1/M depuis 1999. Il n'était, cependant, pas possible de savoir si la patiente avait été surinfectée par la souche de groupe O entre 1999 et 2013 ou si elle était co-infectée VIH-1/M+O depuis sa contamination. Il n'était pas non plus possible de conclure quant au moment de l'émergence de la forme recombinante VIH-1/MO.

## II-2. Analyses sur les prélèvements séquentiels et évolutions immuno-virologique et thérapeutique

Afin d'explorer la chronologie infectieuse et de mieux comprendre les résultats obtenus sur ce prélèvement de novembre 2013, tous les prélèvements répliatifs antérieurs disponibles (n=4) ont été analysés (FIGURE 57).

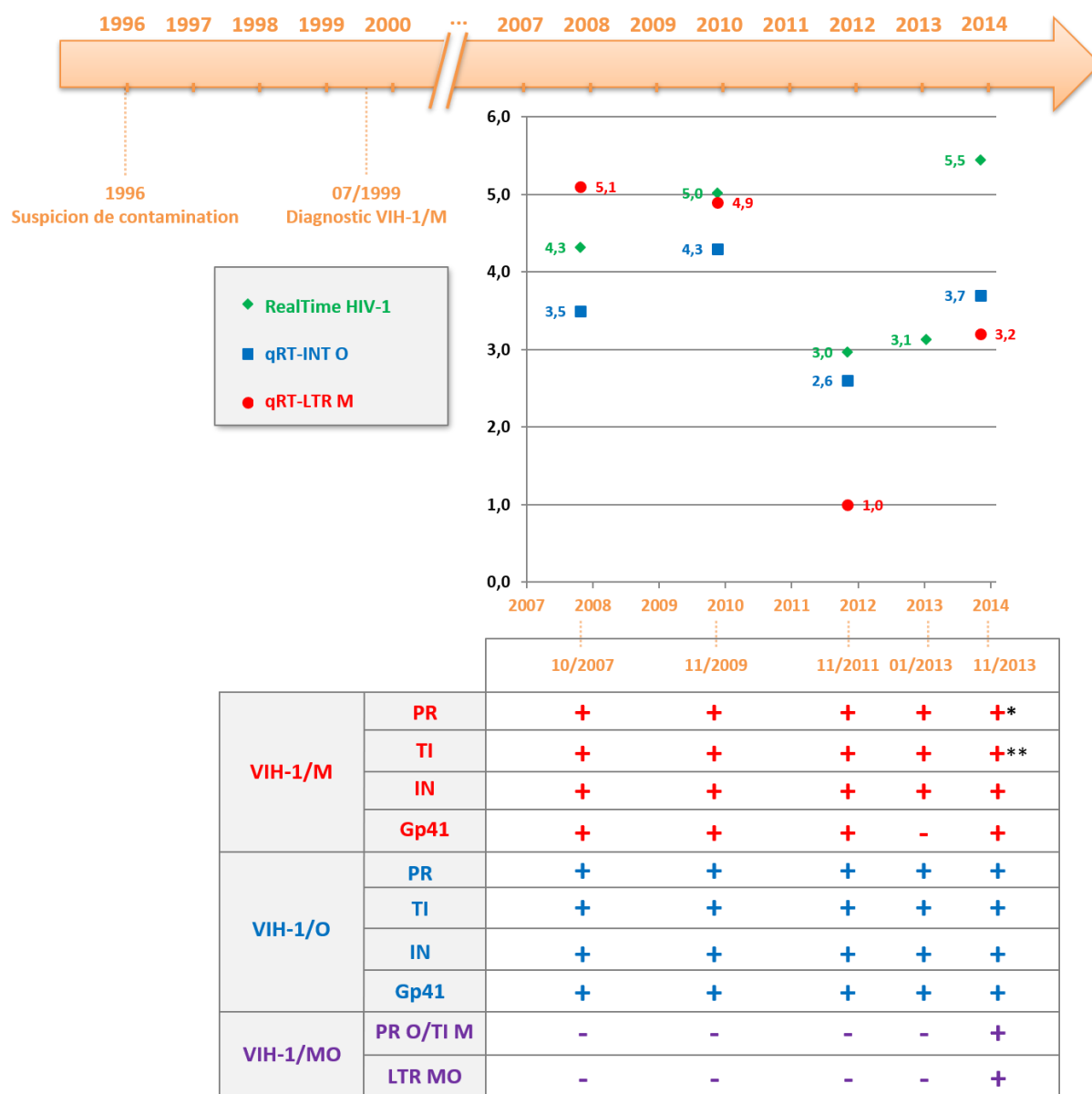


FIGURE 57 : Exploration des prélèvements séquentiels de RBF235.

Les données de charges virales non spécifiques de groupes sont représentées en vert et Celle spécifiques de groupes M et O en rouge et bleu, respectivement. Les valeurs de charges virales, en  $\text{Log}_{10}$  cp/mL, figurent à côté du point correspondant. Les charges virales détectées mais non quantifiées ont été considérées comme égales à  $1,0 \text{ Log}_{10}$  cp/mL. Dans le tableau, les résultats des PCRs nichées spécifiques de groupes M et O sont représentés en rouge et bleu, respectivement. Ceux des PCRs hétérologues sont représentés en violet. Le symbole + correspond à une PCR positive, confirmée par séquençage. Le symbole - correspond à une PCR négative sur gel d'agarose. \*PCR positive mais VIH-1/O au séquençage. \*\*PCR positive mais VIH-1/OM au séquençage.

## II-2.1 Echantillons et analyses

Pour tous les échantillons, le sérotypage de l'enveloppe montrait une double réactivité M+O.

Les charges virales RealTime HIV-1, qRT-LTR M et qRT-INT O étaient toutes quantifiables (à l'exception du prélèvement de novembre 2011 qui n'était que détecté en qRT-LTR M) (FIGURE 57). L'analyse des cibles génomiques des différentes techniques de quantification utilisée nous a permis de suspecter quelles étaient les espèces répliquatives aux différentes dates de prélèvement. La charge virale non spécifique RealTime HIV-1, amplifiant les trois souches, parentales VIH-1/M et VIH-1/O et recombinante VIH-1/MO, était positive aux quatre dates considérées (10/2007, 11/2009, 11/2011 et 01/2013), avec des valeurs comprises entre 3,0 et 5,5 Log<sub>10</sub> cp/mL. Les charges virales spécifiques de groupes M et O étaient également toutes quantifiables sur cette période (à l'exception du prélèvement de janvier 2013, sur lequel seule la charge virale non spécifique de groupes avait été réalisée) avec des valeurs comprises entre 1,0 et 5,1 Log<sub>10</sub> cp/mL et entre 2,6 et 4,3 Log<sub>10</sub> cp/mL, respectivement. Ces résultats indiquaient la réplification des deux souches parentales VIH-1/M et VIH-1/O, entre 2007 et 2013 (FIGURE 57).

Concernant les PCRs nichées spécifiques de groupes, ciblant quatre régions du génome, elles étaient toutes positives pour les quatre échantillons ARN, à la fois pour le VIH-1/M et pour le VIH-1/O, à l'exception de la gp41 M en janvier 2013 (FIGURE 57). L'absence d'amplification de cette région gp41 M était probablement due à la moindre réplification de la souche parentale VIH-1/M à cette date. Le séquençage des amplicons des quatre nouveaux échantillons a confirmé les résultats de PCRs précédents (FIGURE 57). Les mêmes sous-types D de VIH-1/M et sous-groupe H de VIH-1/O ont été retrouvés que sur le prélèvement de novembre 2013.

Par ailleurs, les PCRs nichées spécifiques des zones de recombinaison (PR O/TI M, et LTR MO) étaient toutes les deux positives uniquement pour le prélèvement de novembre 2013, suggérant l'absence de la forme recombinante VIH-1/MO entre octobre 2007 et janvier 2013 (FIGURE 57).

Ainsi, tous ces résultats moléculaires ont permis de conclure à une co-infection VIH-1/M±O±MO répliquative depuis au moins 2007, avec une apparente moindre réplification de la souche parentale VIH-1/M en 2011. Il n'était en revanche pas toujours possible de distinguer précisément les populations virales présentes. L'émergence de la forme recombinante VIH-1/MO avait probablement eu lieu entre janvier et novembre 2013.

## II-2.2 Evolutions immuno-virologique et thérapeutique durant toute la durée de suivi

Une fois les explorations virologiques achevées, nous avons analysé l'histoire immuno-virologique et thérapeutique de cette patiente, afin de mieux comprendre la dynamique virale durant la période et surtout d'adapter sa prise en charge thérapeutique, au regard des nouveaux éléments de son dossier.

### a) Histoire thérapeutique

Dès le diagnostic, en juillet 1999, la patiente a été mise sous traitement ARV, lamivudine (3TC), zidovudine (AZT) et névirapine (NVP). Cette trithérapie ARV était recommandée en première ligne à cette date, sauf pour l'infection à VIH-1/O, mais celle-ci n'était pas connue chez cette patiente (FIGURE 58).

Après de nombreuses périodes d'inobservance rapportées par les médecins et une interruption totale de traitement pendant 6 mois (entre février et septembre 2003), un remplacement de la NVP par l'abacavir (ABC) a été réalisé jusqu'en novembre 2011. Compte tenu de l'inobservance régulière chez cette patiente, de multiples lignes ont ensuite été testées avec une dernière ligne instaurée en octobre 2018 et composée de dolutegravir (DTG) + darunavir (DRV) + ritonavir (RTV) (FIGURE 58).

### b) Cinétique de la charge virale et du taux de CD4

Les données de charges virales collectées entre 1999 et 2018 ont montré une alternance de périodes de réplication et d'indélectabilité (FIGURE 58, TABLEAU 28).

Grâce à la caractérisation des différentes populations virales sur cette période et à la connaissance des techniques de charge virale utilisées, nous pouvions retracer la dynamique virologique chez cette patiente.

A partir de la mise sous traitement ARV de la patiente en septembre 1999 et jusqu'en 2007, année où la réplication a repris de manière importante et durable, les données de charges virales et de CD4 témoignaient d'une efficacité thérapeutique avec une indélectabilité la plupart du temps, mais qui n'était pas mesurable pour la population VIH-1/M du fait des techniques utilisées (FIGURE 58, TABLEAU 28). La cinétique des CD4 montre que, lors du diagnostic de l'infection en juillet 1999, la patiente était immunodéprimée avec 122 CD4/mm<sup>3</sup>. Le taux de CD4 a ensuite présenté une succession de phases de croissance et décroissance jusqu'à atteindre son zénith en mars 2007, avec un taux de 371 CD4/mm<sup>3</sup> (FIGURE 58, TABLEAU 28).



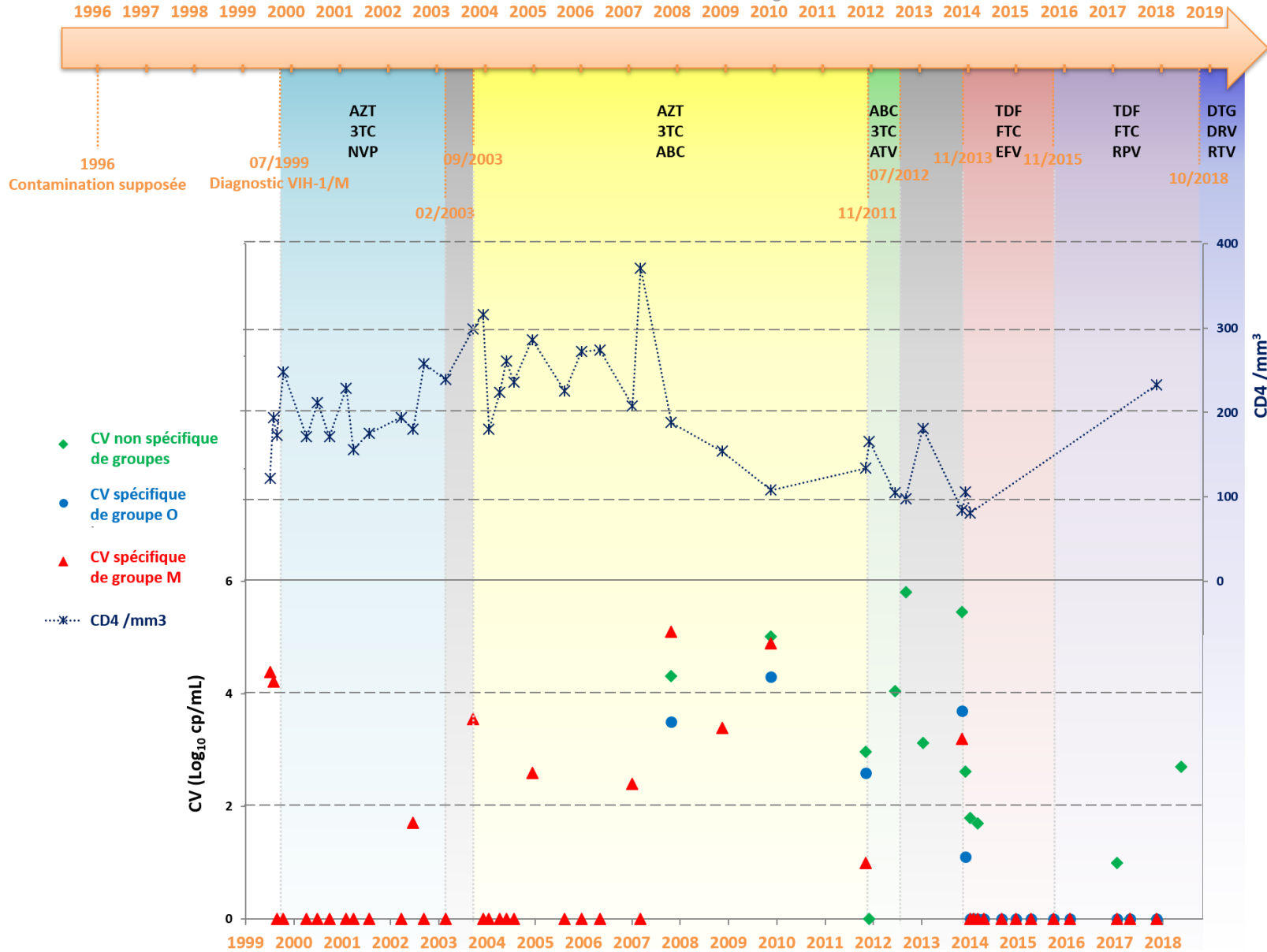


FIGURE 58 : Synthèse des données immuno-virologiques et thérapeutiques entre 1999 et 2018.

Les périodes d'interruption de traitement figurent en gris. INTI = Inhibiteurs Nucléosidiques de la TI : ABC = abacavir, 3TC = lamivudine, AZT = zidovudine, TDF = ténofovir, FTC = emtricitabine. INNTI = Inhibiteurs Non Nucléosidiques de la TI : NVP = névirapine, EFV = efavirenz, RPV = rilpivirine. IP = Inhibiteur de Protéase : ATV = atazanavir.

TABLEAU 28 : Synthèse des données immuno-virologiques entre 1999 et 2018.

Date	Charge virale Log <sub>10</sub> (cp/mL)							CD4 /mm <sup>3</sup>
	non spécifique de groupes			spécifique de groupe M			spécifique de groupe O	
	RealTime HIV-1 (Abbott)	Cobas TaqMan HIV-1 V2 (Roche)	Xpert HIV-1 viral load (Cepheid)	Cobas TaqMan HIV-1 V1 (Roche)	Versant HIV-1 RNA 3.0 (Bayer Corporation)	Amplicor Monitor V1.5 (Roche)	qRT-LTR M (Biocentric)	
05/07/1999					4,4			122
30/07/1999					4,2			194
25/08/1999					0,0			173
13/10/1999					0,0			248
05/04/2000					0,0			171
28/06/2000					0,0			212
27/09/2000					0,0			171
31/01/2001					0,0			229
26/03/2001					0,0			156
23/07/2001					0,0			175
21/03/2002					0,0			194
20/06/2002					1,7			180
12/09/2002					0,0			258
20/02/2003					0,0			239
17/09/2003						3,6		299
01/12/2003						0,0		316
14/01/2004						0,0		180
07/04/2004						0,0		224
26/05/2004						0,0		261
21/07/2004						0,0		236
08/12/2004						2,6		286
10/08/2005						0,0		225
14/12/2005				0,0				272
03/05/2006				0,0				274
03/01/2007				2,4				208
07/03/2007				0,0				371
26/10/2007	4,3						5,1	188
14/11/2008				3,4				154
18/11/2009	5,0						4,9	108
04/11/2011	3,0						1,0	134
02/12/2011		0,0						166
11/06/2012	4,1							105
03/09/2012	5,8							98
11/01/2013	3,1							181
04/11/2013	5,5							84
29/11/2013	2,6							106
03/01/2014	1,8							81
31/01/2014	0,0							
28/02/2014	1,7							
18/04/2014								
27/08/2014								
17/12/2014								
08/04/2015								
23/09/2015								
25/01/2016								
18/01/2017	1,0							
26/04/2017	0,0							
15/11/2017			0,0					233
16/05/2018			2,7					
26/09/2018			2,2					

Le suivi immuno-virologique et thérapeutique de la patiente a été réalisé à l'Hôpital Cochin, à Paris, entre 1999 et 2007, et au CHU de Rennes, depuis 2007. Les charges virales indétectables ont été considérées comme négatives et donc égales à 0,0 Log<sub>10</sub> cp/mL. Les charges virales détectées mais non quantifiées ont été considérées comme égales à 1,0 Log<sub>10</sub> cp/mL.

La période suivante de 2007-2014 était, quant à elle, caractérisée à la fois par des niveaux de réplication importants de toutes les souches (parentales et recombinante) et par une diminution brutale des CD4 de 371 à 148 CD4/mm<sup>3</sup> en 6 mois, atteignant un nadis de 84 CD4/mm<sup>3</sup>, en novembre 2013 (FIGURE 58, TABLEAU 28). Cette période était associée à une dysobservance répétée de la patiente. En particulier, la période d'interruption totale des traitements ARV, entre 2012 et 2013, a probablement favorisé la reprise de réplication des souches parentales VIH-1/M et VIH-1/O, rendant alors possible l'émergence de la forme recombinante VIH-1/MO, le pré-requis de double infection VIH-1/M+O étant là (FIGURE 58, TABLEAU 28).

Entre 2014 et 2018, les charges virales étaient de nouveau indétectables, jusqu'à la reprise d'une réplication en mai 2018. En novembre 2017, les CD4 étaient remontés à 233/mm<sup>3</sup> (FIGURE 58, TABLEAU 28).

La complexité virologique de la situation, associée à une histoire thérapeutique compliquée et mouvementée, nous a conduit à analyser les profils de résistance des différentes populations au cours du temps, afin de suivre l'évolution virale mais surtout d'adapter la prise en charge de la patiente. Des génotypages de résistance ont donc été réalisés spécifiquement sur chacune des populations VIH-1/M, VIH-1/O et VIH-1/OM, en fonction de leur niveau de réplication. Du fait du nombre importants de données, seuls les génotypes cumulés sont présentés (TABLEAU 29).

A ce stade des résultats, seules trois molécules (AZT, DRV, Cabotegravir) restent complètement actives sur les trois populations virales. En effet, on peut constater que la population VIH-1/M apparaît sensible à davantage de molécules (en particulier parmi les INTI), mais celles-ci sont inefficaces sur la population VIH-1/O, risquant ainsi de favoriser la réplication de cette population et de conduire à la génération/genèse d'un recombinant avec une TI de VIH-1/O.

Au final, les génotypes cumulés montrent une relative impasse thérapeutique pour traiter toutes les populations virales présentes chez cette patiente.

TABLEAU 29 : Génotypes cumulés de résistance de 2018

Classe thérapeutique	Molécules	Génotypes cumulés de chaque population*						Génotype cumulé des trois souches	
		VIH-1/M		VIH-1/O		VIH-1/MO			
INTI	zidovudine	S	S	S	S	S	S	S	
	lamivudine	S	R	R	S	S	R	R	
	emtricitabine	S	R	R	S	S	R	R	
	abacavir	S	R	R	S	S	R	R	
	tenofovir	S	R	R	S	S	R	R	
INNTI	nevirapine	S	R	R	S	S	R	R	
	efavirenz	S	R	R	S	S	R	R	
	etravirine	I	R	R	I	I	R	R	
	rilpivirine	S	R	R	S	S	R	R	
IP	lopinavir/r	S	S	S	S	S	I	I	
	atazanavir/r	S	I	I	I	I	I	I	
	darunavir/r	S	S	S	S	S	S	S	
II	raltegravir	NR	S	R	R	R	R	R	
	elvitegravir	NR	S	R	R	R	R	R	
	dolutegravir	NR	S	I	I	I	I	I	
	cabotegravir	NR	S	S	S	S	S	S	

\*Dernier génotype réalisé en 2018 pour les trois populations

Génotypes interprétés avec l'algorithme ANRS version 27 de 2017

**INTI** = Inhibiteurs Nucléosidiques de la TI. **INNTI** = Inhibiteurs Non Nucléosidiques de la TI. **IP** = Inhibiteur de Protéase. **II** = Inhibiteurs d'Intégrase. **IF** = inhibiteur de fusion.

**S** = sensible. **R** = résistant. **I** = intermédiaire. **NR** = non réalisé.

### III- Discussion

L'exploration de ce cas, par l'analyse de prélèvements séquentiels et la collecte des données de suivi, nous a permis de révéler une double infection VIH-1/M+O non diagnostiquée initialement et de décrire, pour la première fois, la dynamique d'émergence *in vivo* d'une forme recombinante VIH-1/MO.

L'exploration de ce cas d'infection atypique, initiée par la vigilance et l'analyse des résultats par les biologistes en charge du dossier, a nécessité le recours à de multiples outils et techniques spécialisés. Ces outils et stratégies « à la carte » ne peuvent être réalisés ailleurs que dans les CNRs, rendant l'exploration des discordances difficiles en routine hospitalière.

Ainsi, très rapidement après la mise en évidence de la discordance, lors du génotypage de résistance de novembre 2013, révélant par ailleurs un manque de spécificité de certaines amorces « ANRS » spécifiques de groupe M, nous avons suspecté une double infection VIH-1/M+O et la présence d'une forme recombinante VIH-1/MO présentant au moins un point de cassure dans la région PR/TI. Cela nous a conduits au développement d'une stratégie et d'outils sur mesure, permettant de caractériser les différentes populations virales. Les différentes PCRs mises au point, spécifiques de groupes M et O et hétérologues, et la stratégie de séquençage du génome complet ont confirmé la présence, d'une souche recombinante VIH-1/MO et d'une souche de VIH-1/O, en plus de la souche VIH-1/M initialement identifiée. La caractérisation précise de la forme recombinante VIH-1/MO a montré, comme la plupart des formes recombinantes décrites jusqu'à présent, un profil de recombinaison relativement simple, avec un point de cassure dans la TI et un second dans la région des LTRs. Cette forme recombinante n'était pas liée phylogénétiquement aux précédentes formes recombinantes décrites et constituait donc une nouvelle URF\_MO.

La disponibilité de prélèvements séquentiels répliatifs nous a permis non seulement de démontrer que la double infection VIH-1/M+O était déjà présente en 2007, mais aussi de quasiment dater l'émergence de la forme recombinante en 2013. En revanche, il n'a pas été possible de suspecter ou de rechercher la double infection avant 2007, du fait de l'absence de prélèvement répliatif à notre disposition et de l'absence de déclaration obligatoire (DO) au moment du diagnostic, en 1999.

Associée aux données immuno-virologiques de suivi sur 20 ans, cette caractérisation des différentes populations virales a permis de comprendre la dynamique virale chez la patiente et de mettre en évidence, pour la première fois, l'émergence d'une forme recombinante. En revanche, il est difficile de décrire une histoire naturelle claire de cette genèse, du fait de l'absence de caractérisation des populations avant 2007 et des nombreuses périodes successives de succès et échecs virologiques. Ces données ont cependant mis en évidence que la dysobservance et les interruptions de traitement, durant la période 2007-2013, ont probablement favorisé l'émergence de la forme recombinante VIH-1/MO. Notre méthodologie ne permettait pas de rechercher les éventuelles formes recombinantes minoritaires sur les prélèvements répliatifs disponibles, puisque les techniques actuelles ne permettent la mise en évidence que des populations virales majoritaires (qu'elles soient

parentales ou recombinante). Il ne peut être exclu que d'autres formes recombinantes aient été générées antérieurement mais que la période de forte réplication observée entre 2007 et 2013 ait favorisé la multiplication de la forme recombinante la plus adaptée. De plus, la présence concomitante de VIH-1/M, de VIH-1/O et de VIH-1/MO chez cette patiente pourrait conduire, en cas d'une reprise de réplication, à de nouvelles formes recombinantes VIH-1/MO, de profils et de motifs davantage complexes, avec une potentielle sélection des mutations de résistance par le phénomène de recombinaison entre les trois populations. Des explorations à plus large échelle, avec notamment le recours au séquençage de haut débit (NGS), deviennent donc nécessaires afin d'étudier toutes ces populations minoritaires.

Bien que désormais décrites à de multiples reprises, les conditions requises à l'émergence d'une forme recombinante à partir d'une double infection VIH-1/M+O demeurent méconnues, du fait de l'absence de prélèvements séquentiels [150, 153, 156, 185, 196-198]. Dans notre cas, les outils de quantification spécifiques de groupes ont révélé que les charges virales des deux souches parentales VIH-1/M et VIH-1/O, entre 2007 et 2013, étaient élevées, permettant d'avoir le pré-requis théorique à une double infection cellulaire. Des travaux portant spécifiquement sur les niveaux de réplication des populations virales parentales requis pour la genèse d'une forme recombinante VIH-1/MO sont maintenant à envisager.

Les cas de surinfections ou de co-infections VIH-1/M+O et la génération de formes recombinantes VIH-1/MO prouvent la dynamique et l'évolution continues des infections à VIH, qui nécessitent une vigilance accrue lorsque des discordances sont observées, que ce soit lors du dépistage et/ou du suivi des patients. A ce titre, le génotypage de résistance reste quasiment la seule analyse permettant de suspecter de telles situations. Du fait de l'absence d'outils de dépistage spécifiques et représentatifs de la diversité génétique importante, ces phénomènes de doubles infections et de recombinaison sont probablement sous-estimés.

Notre cas démontre aussi la difficulté de prise en charge de telles situations, en particulier lorsque la présence de l'infection à VIH-1/O est méconnue. Ainsi, l'utilisation de la classe de INNTI (inefficace sur les VIH-1/O), dès le début, est probablement responsable de la sélection de nombreuses mutations de résistance sur la TI de la population VIH-1/O. La population VIH-1/M est, elle, restée sensible à cette classe, du fait d'une pression thérapeutique efficace, comme en atteste l'indétectabilité retrouvée de manière durable.

Cette situation a largement compliqué la démarche thérapeutique, qui nécessitait d'être efficace sur toutes les populations.

Actuellement, à condition qu'elle soit réalisée dès le diagnostic, la mise sous traitement ARV devrait limiter le risque d'émergence de tels recombinants chez ces patients doublement infectés VIH-1/M+O. Mais ce postulat soulève la question de la couverture du dépistage et du traitement, notamment en zone d'endémie des VIH-1/M et des VIH-1/O, comme le Cameroun, où les INNTI sont encore utilisés, en attendant l'application des nouvelles recommandations préconisant un remplacement par le dolutegravir, d'où la nécessité de bien identifier ces profils d'infections.

Enfin, même si ce cas est anecdotique, il ouvre, avec la quarantaine d'autres cas de double infection VIH-1/M+O et recombinaison VIH-1/MO maintenant répertoriés, un large champ de perspectives de recherche sur les conséquences de telles infections atypiques en termes d'épidémiologie, de physiopathologie et de prise en charge thérapeutique.

## **CONCLUSION - PERSPECTIVES**





Bien que longtemps jugée impossible, la recombinaison entre les VIH-1 de groupes M et O est à l'origine de la circulation d'au moins 19 URF\_MO au Cameroun et en France, décrites depuis 1999. Ce nombre est probablement sous-estimé du fait de la difficulté à mettre en évidence spécifiquement ces formes recombinantes MO, lors du dépistage ou du suivi des patients, et de l'absence de leur recherche systématique. Il soulève néanmoins la question des mécanismes de leur genèse et de leur émergence et celle de l'avantage conféré par la recombinaison en comparaison aux souches parentales VIH-1/M et VIH-1/O. Ainsi, dans ce travail, nous avons cherché à étudier, d'une part, le potentiel répliatif d'une forme recombinante VIH-1/MO en le comparant à celui de ses souches parentales et, d'autre part, l'émergence *in vivo* d'un tel recombinant à partir d'une double infection VIH-1/M+O et *in vitro*, en se focalisant sur le point de chaud de la recombinaison décrit dans les LTRs.

La mise en place de nos travaux portant sur le potentiel répliatif d'une forme recombinante MO a d'abord reposé sur l'hypothèse qu'il serait plus avantageux pour le virus recombinant d'avoir une enveloppe de VIH-1/M, puisque le groupe M est caractérisé par une plus large diffusion au sein de la population humaine, que la région Env est impliquée dans l'entrée du virus dans les cellules-cibles, qu'elle constitue une cible majeure de la réponse immunitaire de l'hôte et que les groupes M et O présentent environ 50% de divergence dans cette région. L'objectif initial était donc de construire des virus recombinants de motifs OM et MO, afin de les comparer entre eux et de déterminer si l'enveloppe de VIH-1/M combinée à une TI de VIH-1/O constituait effectivement un avantage répliatif. Les travaux mis en place afin d'y parvenir n'ont pas abouti, aucun CMIC pVIH-1/MO, capable de générer des virus recombinants de motif MO, n'ayant pu être produit. Les données que nous avons obtenues sont donc partielles mais demeurent intéressantes.

En effet, les résultats de l'étude de Peeters et *al.*, suggérant un *fitness* répliatif en faveur de la forme recombinante VIH-1/MO par rapport à au moins l'une de ses souches parentales <sup>[192]</sup>, associés à ceux d'une étude précédente d'Arien et *al.* suggérant un *fitness* en faveur du groupe M par rapport au groupe O <sup>[169]</sup>, faisaient supposer que le potentiel répliatif de notre souche recombinante VIH-1/MO pourrait être supérieur à celui de ses deux souches parentales. Cependant, les résultats que nous avons obtenus ne sont pas complètement en accord avec cette idée. Bien que les pics de répliation aient été atteints plus précocement, et avec une intensité plus importante pour la souche parentale VIH-1/M que pour la souche

parentale VIH-1/O, nos résultats montrent un comportement, certes hybride, mais non intermédiaire de la forme recombinante VIH-1/MO. Ainsi, selon les marqueurs et les paramètres étudiés, le potentiel répliatif du VIH-1/MO montrait des similitudes soit avec la souche parentale VIH-1/M, soit avec la souche parentale VIH-1/O, et s'avérait même parfois inférieur aux deux. La bonne reproductibilité de ces résultats ayant été contrôlée, validant ainsi notre démarche expérimentale, nous pouvons exclure l'hypothèse de problèmes d'ordre technique, qui aurait pu influencer les conclusions tirées. Ces données restent toutefois préliminaires et nécessitent une vraie étude du *fitness* des différentes souches, en réalisant des doubles infections voire une triple infection VIH-1/M+O+MO, et en les mettant ainsi en compétition entre elles.

De plus, ces résultats sont à mettre en lien avec la diversité inter-groupes M et O, qui laisse supposer que toutes les formes recombinantes ne sont pas forcément possibles et qu'elles n'ont pas forcément toutes le même profil. Du fait de la faible disponibilité de CMI parentaux différents, en termes de diversité génétique, ce travail a été réalisé à partir d'un unique représentant de VIH-1/M, de sous-type B associé à un unique représentant de VIH-1/O, de sous-groupe T. Il n'existe à ce jour aucun recombinant VIH-1/MO décrit *in vivo* montrant cette association. En effet, parmi les 19 URF\_MO répertoriées, 8 (soit 42%) combinent une souche VIH-1/M de sous-type CRF02\_AG avec une souche VIH-1/O de sous-groupe H et 3 (soit 16%) combinent une souche VIH-1/M de sous-type CRF02\_AG avec une souche VIH-1/O de sous-groupe T. De la même façon, il a été décrit, après l'initiation de ce travail, que la souche parentale VIH-1/O RBF206, utilisée lors de ces travaux, possédait une protéine Vpu permettant un contournement efficace de la tétherine, contrairement à toutes les autres souches VIH-1/O. Elle n'apparaît donc pas être la plus représentative du comportement répliatif des variants de groupe O. Il semble donc désormais nécessaire d'étudier le potentiel répliatif d'une souche recombinante provenant d'un CMI parental VIH-1/M de sous-type CRF02\_AG et d'un CMI parental VIH-1/O davantage représentatif de ce groupe. De plus, un seul des deux profils de recombinaison initialement envisagés, à savoir  $[O_{gag/pol}-M_{env}]$  et  $[M_{gag/pol}-O_{env}]$ , n'a été étudié pour le moment, rendant difficile la mise en évidence d'un avantage répliatif conféré par la recombinaison.

Cette question d'une meilleure répliation des formes recombinantes VIH-1/MO est d'importance, car s'il s'avérait qu'elles aient effectivement un meilleur *fitness* que les virus parentaux VIH-1/M ou VIH-1/O, les variants VIH-1/O pourraient à terme être « assimilés » par

les variants VIH-1/M sous la forme de virus mosaïques, permettant une meilleure diffusion. L'épidémie à VIH-1/O pourrait alors, pour subsister, évoluer ou être sélectionnée vers un modèle d'épidémie de variants recombinants VIH-1/MO.

Nos résultats, même s'ils répondent partiellement à cette question, constituent une première étape de faisabilité et offrent de nombreuses perspectives de travail, avec notamment la génération de virus chimériques entre différents variants VIH-1 de groupes M et O, afin d'étudier l'impact de l'origine génétique des fragments M et O impliqués et de rechercher l'existence de profils préférentiels de recombinaison (OM versus MO).

Un autre objectif de ce travail était d'analyser *in vitro* les mécanismes de recombinaison au sein des LTRs, sous leur forme génomique ARN et ADN.

L'étude de l'émergence de points de cassure sur l'ADN intracellulaire archivé a permis d'étudier l'intégralité des LTRs et de montrer que de très nombreux motifs de recombinaison sont possibles, avec une localisation des points de cassure dans quasiment toutes les régions, mais avec néanmoins la mise en évidence de trois motifs largement majoritaires, tous localisés dans la région R. Cette forme génomique ADN ne correspond pas forcément à des virus réplicatifs pouvant devenir prédominants au cours de l'infection, ce qui constitue une limite à son étude. L'étude en parallèle des LTRs sous leur forme ARN, provenant des surnageants de culture, a ainsi présenté l'avantage de s'intéresser uniquement aux virus ayant répliqué et qui étaient donc viables, fonctionnels et infectieux. Bien que cette forme génomique ne soit pas adaptée à l'analyse des 3'LTRs, elle nous a permis de démontrer que, pour cette combinaison VIH-1/M sous-type B et VIH-1/O sous-groupe T, la recombinaison avait eu lieu préférentiellement dans la région R selon trois motifs majoritaires, dont un (513-522pb) devenait prédominant en fin de co-culture.

En parallèle, l'étude des LTRs recombinants d'URF\_MO nous a permis de valider notre modèle expérimental, grâce à l'observation des mêmes points de cassure *in vitro* et *in vivo*. Elle a également mis en évidence que d'autres points de cassure sont possibles, suggérant qu'il n'existe pas un mécanisme unique de recombinaison. Ces points de cassure supplémentaires s'expliquent probablement par la pression de sélection liée à l'hôte et des contraintes de réplication plus importantes *in vivo* qu'*in vitro*. L'hypothèse que la région d'émergence des points de cassure soit dépendante des sous-types de VIH-1/M et des sous-groupes de VIH-1/O

peut également être avancée, bien qu'après avoir comparé la localisation des points de cassure dans les LTRs et la diversité génétique des fragments impliqués dans les URF\_MO, aucun lien n'ait pu être établi, à ce stade. De nombreux profils différents de recombinaison, parfois complexes, sont donc viables et fonctionnels. Si la localisation des points de cassure est déterminée par des contraintes mécanistiques telles que l'homologie de séquence ou la présence de structures secondaires, celle-ci est contrebalancée par la sélection négative des souches générées par des événements de recombinaison délétères.

Enfin, nos travaux ont confirmé que lorsqu'un seul point de cassure est généré au sein du génome codant du virus (c'est-à-dire entre les régions Gag et Nef) au cours d'un cycle de réplication, un événement de recombinaison a obligatoirement lieu dans les LTRs au cours du cycle suivant, générant ainsi une progénie virale avec des régions LTRs identiques et recombinantes.

Si ces travaux à partir du CMIC ont permis de caractériser finement le mécanisme de recombinaison et la localisation des points de cassure, sur un modèle expérimental unique, l'étude de l'émergence d'une forme recombinante VIH-1/MO chez la patiente RBF235 a, elle, permis de caractériser cette dynamique *in vivo*.

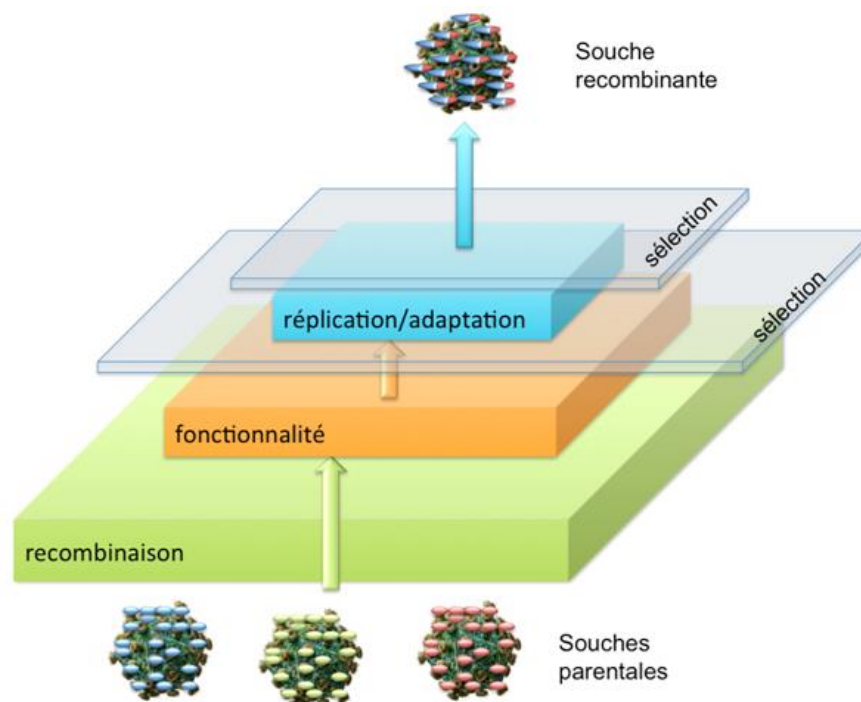
La prise en charge compliquée de ce cas, pour lequel le variant VIH-1/O n'avait pas été identifié au départ, avec l'utilisation d'INNTI, chez une patiente inobservante de surcroît, a abouti à l'émergence d'une forme recombinante VIH-1/MO. L'émergence a été estimée à l'année 2013, associée à une période de réplication intense des deux souches parentales VIH-1/M et VIH-1/O. Cette émergence a possiblement pu avoir eu lieu avant, puisque la double infection VIH-1/M+O était présente depuis au moins 2007, mais être ensuite restée archivée, sans réplication détectable. L'absence de prélèvement réplcatif disponible avant 2007 nous a par ailleurs empêché de déterminer si la double infection était le résultat d'une co-infection VIH-1/M+O dès le départ ou d'une surinfection VIH-1/O plus récente.

La question de l'impact du niveau de réplication sur la survenue d'évènements de recombinaison demeure sans réponse, mais nos données suggèrent que les niveaux de réplication importants des souches parentales ont certainement favorisé leur recombinaison.

Ce cas RBF235 soulève les difficultés de suivi et de prise en charge lorsque de multiples populations virales sont présentes. Il soulève également de nombreuses questions concernant la physiopathologie de ces formes en termes d'agressivité sur les cellules cibles CD4+, de

capacité répliquative et de vitesse d'évolution vers le stade SIDA notamment. La synthèse de toutes les données immuno-virologiques séquentielles disponibles des patients infectés par une forme recombinante VIH-1/MO est en cours, afin d'essayer d'évaluer le comportement de ces souches, en regard de ce qui est déjà décrit pour les variants de groupes M et O.

Actuellement, les techniques permettant la caractérisation des formes recombinantes, qu'elles soient *in vivo* ou *in vitro*, ne permettent d'étudier que les populations majoritaires. Or, la genèse d'une forme recombinante semble s'opérer lentement, avec des étapes de génération populations non répliquatives ou ne devenant jamais prédominante. En effet, l'émergence d'une forme recombinante se présente comme un processus pyramidal sélectif : la génération des génomes recombinants pendant la transcription inverse est suivie par la sélection des formes virales fonctionnelles viables, parmi lesquelles les formes les plus adaptées seront sélectionnées par l'hôte (capacité répliquative, échappement à la réponse immunitaire et aux facteurs de restriction, transmissibilité inter-individuelle) (FIGURE 59).



**FIGURE 59** : Schéma général représentatif du processus pyramidal sélectif de génération et de sélection des formes recombinantes au sein des VIH.

Trois niveaux sont impliqués : le mécanisme de recombinaison avec la possibilité de générer des recombinants au niveau génomique, puis la sélection des génomes chimériques capables de générer des virions viables et enfin les capacités de répliquative et d'adaptation de la progénie virale recombinante (d'après Pierre Cappy <sup>[245]</sup>).

Il apparaît donc désormais intéressant de rechercher et de caractériser ces populations minoritaires, afin d'explorer encore plus en profondeur le mécanisme de recombinaison. La mise au point du séquençage de ces génomes mosaïques par des méthodes de séquençage de nouvelle génération (NGS) devrait permettre à terme de lever cette limite.

Par ailleurs, l'échappement aux facteurs de restriction et au système immunitaire de l'hôte, conféré par un changement antigénique brusque ou l'acquisition d'un motif d'échappement présent sur une quasi-espèce minoritaire, pourra également permettre à la forme recombinante de prédominer par rapport à ses souches parentales <sup>[131]</sup>. Une des perspectives ouvertes par notre travail porte ainsi sur l'impact du facteur de restriction cellulaire APOBEC3G sur les génomes de VIH-1/O et de VIH-1/MO, et de le comparer à ce qui a déjà été rapporté pour les variants VIH-1/M et VIH-2.

En conclusion, nos résultats montrent le processus de recombinaison inter-groupes M et O est probablement plus prolifique qu'imaginé jusqu'alors et conduit à de multiples formes recombinantes VIH-1/MO. Il existe sûrement des contraintes mécanistiques, qui limitent malgré tout le nombre de ces formes viables et fonctionnelles, et auxquelles s'ajoute la dynamique de genèse *in vivo* d'un recombinant. Cette dernière est probablement influencée par de multiples facteurs, tels que la pression de sélection de l'hôte, la compétition entre les variants de groupes M et O pour infecter un même hôte, et le niveau de réplication des souches parentales. Ainsi, nos analyses ont fourni des premiers résultats qui soulèvent de nombreuses questions et perspectives de recherche sur les éléments nécessaires à l'émergence réussie d'une forme recombinante VIH-1/MO.

## **REFERENCES BIBLIOGRAPHIQUES**





1. Control) CCfD. **Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California.** *MMWR Morb Mortal Wkly Rep* 1981; 30(25):305-308.
2. Barre-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, et al. **Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS).** *Science* 1983; 220(4599):868-871.
3. Popovic M, Sarngadharan MG, Read E, Gallo RC. **Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS.** *Science* 1984; 224(4648):497-500.
4. Mboup S, Essex M, Kank P, Ndoye I, Barin F, Samba, et al. **[The HTL IV virus: a new human retrovirus without pathogen protection against the HTLV III/LAV virus?].** *Afr J Sex Transmi Dis* 1986; 2(2):81-84.
5. De Leys R, Vanderborgh B, Vanden Haesevelde M, Heyndrickx L, van Geel A, Wauters C, et al. **Isolation and partial characterization of an unusual human immunodeficiency retrovirus from two persons of west-central African origin.** *J Virol* 1990; 64(3):1207-1216.
6. Simon F, Mauclore P, Roques P, Loussert-Ajaka I, Muller-Trutwin MC, Saragosti S, et al. **Identification of a new human immunodeficiency virus type 1 distinct from group M and group O.** *Nat Med* 1998; 4(9):1032-1037.
7. Plantier JC, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Lemee V, et al. **A new human immunodeficiency virus derived from gorillas.** *Nat Med* 2009; 15(8):871-872.
8. Huraux JM, Nicolas JC, Agut H, Peigue Lafeuille H. **Traité de Virologie Médicale.** Estem; 2003.
9. Girard PM, Katlama C, Pialoux G. **VIH.** Doin; 2011.
10. Janeway CA, Travers P, Walport M, Shlomchik MJ. **Immunobiology.** Garland Science; 2005.
11. Cassan E, Arigon-Chifolleau AM, Mesnard JM, Gross A, Gascuel O. **Concomitant emergence of the antisense protein gene of HIV-1 and of the pandemic.** *Proc Natl Acad Sci U S A* 2016; 113(41):11537-11542.
12. Cappy P, Moisan A, De Oliveira F, Plantier JC, Negroni M. **HIV-1 sequences in the epidemic suggest an alternative pathway for the generation of the Long Terminal Repeats.** *Sci Rep* 2017; 7(1):13715.
13. Fulton AM, Rathjen PD, Kingsman SM, Kingsman AJ. **Upstream and downstream transcriptional control signals in the yeast retrotransposon, TY.** *Nucleic Acids Res* 1988; 16(12):5439-5458.
14. Neil S, Bieniasz P. **Human immunodeficiency virus, restriction factors, and interferon.** *J Interferon Cytokine Res* 2009; 29(9):569-580.
15. Malim MH, Emerman M. **HIV-1 accessory proteins--ensuring viral survival in a hostile environment.** *Cell Host Microbe* 2008; 3(6):388-398.
16. Laguette N, Benkirane M. **How SAMHD1 changes our view of viral restriction.** *Trends Immunol* 2012; 33(1):26-33.
17. Kirchhoff F. **Immune evasion and counteraction of restriction factors by HIV-1 and other primate lentiviruses.** *Cell Host Microbe* 2010; 8(1):55-67.
18. Daugherty MD, Malik HS. **Rules of engagement: molecular insights from host-virus arms races.** *Annu Rev Genet* 2012; 46:677-700.
19. Kluge SF, Sauter D, Kirchhoff F. **SnapShot: antiviral restriction factors.** *Cell* 2015; 163(3):774-774 e771.
20. Sauter D, Kirchhoff F. **HIV replication: a game of hide and sense.** *Curr Opin HIV AIDS* 2016; 11(2):173-181.
21. McLaren PJ, Gawanbacht A, Pyndiah N, Krapp C, Hotter D, Kluge SF, et al. **Identification of potential HIV restriction factors by combining evolutionary genomic signatures with functional analyses.** *Retrovirology* 2015; 12:41.

22. Fackler OT. **Spotlight on HIV-1 Nef: SERINC3 and SERINC5 Identified as Restriction Factors Antagonized by the Pathogenesis Factor.** *Viruses* 2015; 7(12):6730-6738.
23. Stremlau M, Owens CM, Perron MJ, Kiessling M, Autissier P, Sodroski J. **The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys.** *Nature* 2004; 427(6977):848-853.
24. Strebel K, Luban J, Jeang KT. **Human cellular restriction factors that target HIV-1 replication.** *BMC Med* 2009; 7:48.
25. Chiu YL, Soros VB, Kreisberg JF, Stopak K, Yonemoto W, Greene WC. **Cellular APOBEC3G restricts HIV-1 infection in resting CD4+ T cells.** *Nature* 2005; 435(7038):108-114.
26. Suspene R, Sommer P, Henry M, Ferris S, Guetard D, Pochet S, et al. **APOBEC3G is a single-stranded DNA cytidine deaminase and functions independently of HIV reverse transcriptase.** *Nucleic Acids Res* 2004; 32(8):2421-2429.
27. Yu Q, Konig R, Pillai S, Chiles K, Kearney M, Palmer S, et al. **Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome.** *Nat Struct Mol Biol* 2004; 11(5):435-442.
28. Huthoff H, Towers GJ. **Restriction of retroviral replication by APOBEC3G/F and TRIM5alpha.** *Trends Microbiol* 2008; 16(12):612-619.
29. Kupzig S, Korolchuk V, Rollason R, Sugden A, Wilde A, Banting G. **Bst-2/HM1.24 is a raft-associated apical membrane protein with an unusual topology.** *Traffic* 2003; 4(10):694-709.
30. Sauter D, Specht A, Kirchhoff F. **Tetherin: holding on and letting go.** *Cell* 2010; 141(3):392-398.
31. Douglas JL, Viswanathan K, McCarroll MN, Gustin JK, Fruh K, Moses AV. **Vpu directs the degradation of the human immunodeficiency virus restriction factor BST-2/Tetherin via a {beta}TrCP-dependent mechanism.** *J Virol* 2009; 83(16):7931-7947.
32. Goffinet C, Allespach I, Homann S, Tervo HM, Habermann A, Rupp D, et al. **HIV-1 antagonism of CD317 is species specific and involves Vpu-mediated proteasomal degradation of the restriction factor.** *Cell Host Microbe* 2009; 5(3):285-297.
33. Gupta RK, Hue S, Schaller T, Verschoor E, Pillay D, Towers GJ. **Mutation of a single residue renders human tetherin resistant to HIV-1 Vpu-mediated depletion.** *PLoS Pathog* 2009; 5(5):e1000443.
34. Mangeat B, Gers-Huber G, Lehmann M, Zufferey M, Luban J, Piguet V. **HIV-1 Vpu neutralizes the antiviral factor Tetherin/BST-2 by binding it and directing its beta-TrCP2-dependent degradation.** *PLoS Pathog* 2009; 5(9):e1000574.
35. Usami Y, Wu Y, Gottlinger HG. **SERINC3 and SERINC5 restrict HIV-1 infectivity and are counteracted by Nef.** *Nature* 2015; 526(7572):218-223.
36. Goffinet C. **Cellular Antiviral Factors that Target Particle Infectivity of HIV-1.** *Curr HIV Res* 2016; 14(3):211-216.
37. Usami Y, Gottlinger H. **HIV-1 Nef responsiveness is determined by Env variable regions involved in trimer association and correlates with neutralization sensitivity.** *Cell Rep* 2013; 5(3):802-812.
38. Rosa A, Chande A, Ziglio S, De Sanctis V, Bertorelli R, Goh SL, et al. **HIV-1 Nef promotes infection by excluding SERINC5 from virion incorporation.** *Nature* 2015; 526(7572):212-217.
39. Descours B, Cribier A, Chable-Bessia C, Ayinde D, Rice G, Crow Y, et al. **SAMHD1 restricts HIV-1 reverse transcription in quiescent CD4(+) T-cells.** *Retrovirology* 2012; 9:87.
40. Lahouassa H, Daddacha W, Hofmann H, Ayinde D, Logue EC, Dragin L, et al. **SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates.** *Nat Immunol* 2012; 13(3):223-228.
41. Hrecka K, Hao C, Gierszewska M, Swanson SK, Kesik-Brodacka M, Srivastava S, et al. **Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein.** *Nature* 2011; 474(7353):658-661.

42. Laguette N, Sobhian B, Casartelli N, Ringeard M, Chable-Bessia C, Segeral E, et al. **SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx.** *Nature* 2011; 474(7353):654-657.
43. Brandariz-Nunez A, Valle-Casuso JC, White TE, Laguette N, Benkirane M, Brojatsch J, et al. **Role of SAMHD1 nuclear localization in restriction of HIV-1 and SIVmac.** *Retrovirology* 2012; 9:49.
44. Goujon C, Riviere L, Jarrosson-Wuillaume L, Bernaud J, Rigal D, Darlix JL, et al. **SIVSM/HIV-2 Vpx proteins promote retroviral escape from a proteasome-dependent restriction pathway present in human dendritic cells.** *Retrovirology* 2007; 4:2.
45. Lahouassa H, Daddacha W, Hofmann H, Ayinde D, Logue EC, Dragin L, et al. **SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates.** *Nat Immunol* 2012; 13(3):223-228.
46. Tchasovnikarova IA, Timms RT, Matheson NJ, Wals K, Antrobus R, Gottgens B, et al. **GENE SILENCING. Epigenetic silencing by the HUSH complex mediates position-effect variegation in human cells.** *Science* 2015; 348(6242):1481-1485.
47. Liu N, Lee CH, Swigut T, Grow E, Gu B, Bassik MC, et al. **Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators.** *Nature* 2018; 553(7687):228-232.
48. Chougui G, Munir-Matloob S, Matkovic R, Martin MM, Morel M, Lahouassa H, et al. **HIV-2/SIV viral protein X counteracts HUSH repressor complex.** *Nat Microbiol* 2018; 3(8):891-897.
49. Yurkovetskiy L, Guney MH, Kim K, Goh SL, McCauley S, Dauphin A, et al. **Primate immunodeficiency virus proteins Vpx and Vpr counteract transcriptional repression of proviruses by the HUSH complex.** *Nat Microbiol* 2018; 3(12):1354-1361.
50. Etienne L, Hahn BH, Sharp PM, Matsen FA, Emerman M. **Gene loss and adaptation to hominids underlie the ancient origin of HIV-1.** *Cell Host Microbe* 2013; 14(1):85-92.
51. Collins DR, Lubow J, Lukic Z, Mashiba M, Collins KL. **Vpr Promotes Macrophage-Dependent HIV-1 Infection of CD4+ T Lymphocytes.** *PLoS Pathog* 2015; 11(7):e1005054.
52. Mashiba M, Collins DR, Terry VH, Collins KL. **Vpr overcomes macrophage-specific restriction of HIV-1 Env expression and virion production.** *Cell Host Microbe* 2014; 16(6):722-735.
53. Laguette N, Bregnard C, Hue P, Basbous J, Yatim A, Larroque M, et al. **Premature activation of the SLX4 complex by Vpr promotes G2/M arrest and escape from innate immune sensing.** *Cell* 2014; 156(1-2):134-145.
54. Cohen EA. **From arrest to escape: HIV-1 Vpr cuts a deal.** *Cell Host Microbe* 2014; 15(2):125-127.
55. Vessièrè A, Delaugerre C, Plantier J-C, Simon F. **Passages inter-espèces des virus de l'immunodéficience humaine et simienne chez les primates.** *Revue Francophone des Laboratoires* 2010; 2010(423):43-53.
56. Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, et al. **Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*.** *Nature* 1999; 397(6718):436-441.
57. Van Heuverswyn F, Peeters M. **The origins of HIV and implications for the global epidemic.** *Curr Infect Dis Rep* 2007; 9(4):338-346.
58. Keele BF, Van Heuverswyn F, Li Y, Bailes E, Takehisa J, Santiago ML, et al. **Chimpanzee reservoirs of pandemic and nonpandemic HIV-1.** *Science* 2006; 313(5786):523-526.
59. Van Heuverswyn F, Li Y, Bailes E, Neel C, Lafay B, Keele BF, et al. **Genetic diversity and phylogeographic clustering of SIVcpzPtt in wild chimpanzees in Cameroon.** *Virology* 2007; 368(1):155-171.
60. D'Arc M, Ayoub A, Esteban A, Learn GH, Boue V, Liegeois F, et al. **Origin of the HIV-1 group O epidemic in western lowland gorillas.** *Proc Natl Acad Sci U S A* 2015; 112(11):E1343-1352.
61. Locatelli S, Peeters M. **Cross-species transmission of simian retroviruses: how and why they could lead to the emergence of new diseases in the human population.** *AIDS* 2012; 26(6):659-673.

62. Mourez T, Simon F, Plantier J-C. **Non-M Variants Human Immunodeficiency Virus Type 1.** *Clinical Microbiology Reviews* 2013; 26(3).
63. Sharp PM, Hahn BH. **Origins of HIV and the AIDS pandemic.** *Cold Spring Harb Perspect Med* 2011; 1(1):a006841.
64. Peeters M, Cournaud V, Abela B, Auzel P, Pourrut X, Bibollet-Ruche F, et al. **Risk to human health from a plethora of simian immunodeficiency viruses in primate bushmeat.** *Emerg Infect Dis* 2002; 8(5):451-457.
65. Aghokeng AF, Ayouba A, Mpoudi-Ngole E, Loul S, Liegeois F, Delaporte E, et al. **Extensive survey on the prevalence and genetic diversity of SIVs in primate bushmeat provides insights into risks for potential new cross-species transmissions.** *Infect Genet Evol* 2010; 10(3):386-396.
66. Simon-Lorière, E., GALETTO, R., NEGRONI, M. **Caractéristiques mécanistiques de la recombinaison chez le virus de l'immunodéficience humaine.** Montrouge, FRANCE: John Libbey Eurotext; 2006.
67. Wain LV, Bailes E, Bibollet-Ruche F, Decker JM, Keele BF, Van Heuverswyn F, et al. **Adaptation of HIV-1 to its human host.** *Mol Biol Evol* 2007; 24(8):1853-1860.
68. Takehisa J, Kraus MH, Ayouba A, Bailes E, Van Heuverswyn F, Decker JM, et al. **Origin and biology of simian immunodeficiency virus in wild-living western gorillas.** *J Virol* 2009; 83(4):1635-1648.
69. Vallari A, Holzmayer V, Harris B, Yamaguchi J, Ngansop C, Makamche F, et al. **Confirmation of putative HIV-1 group P in Cameroon.** *J Virol* 2011; 85(3):1403-1407.
70. Alessandri-Gradt E, De Oliveira F, Leoz M, Lemee V, Robertson DL, Feyertag F, et al. **HIV-1 group P infection: towards a dead-end infection?** *Aids* 2018; 32(10):1317-1322.
71. McNatt MW, Zang T, Hatzioannou T, Bartlett M, Fofana IB, Johnson WE, et al. **Species-specific activity of HIV-1 Vpu and positive selection of tetherin transmembrane domain variants.** *PLoS Pathog* 2009; 5(2):e1000300.
72. Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, et al. **HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations.** *Science* 2014; 346(6205):56-61.
73. Leoz M, Feyertag F, Kfutwah A, Mauclore P, Lachenal G, Damond F, et al. **The Two-Phase Emergence of Non Pandemic HIV-1 Group O in Cameroon.** *PLoS Pathog* 2015; 11(8):e1005029.
74. Wertheim JO, Worobey M. **Dating the age of the SIV lineages that gave rise to HIV-1 and HIV-2.** *PLoS Comput Biol* 2009; 5(5):e1000377.
75. Sauter D, Hue S, Petit SJ, Plantier JC, Towers GJ, Kirchhoff F, et al. **HIV-1 Group P is unable to antagonize human tetherin by Vpu, Env or Nef.** *Retrovirology* 2011; 8:103.
76. Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. **Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection.** *Nature* 1995; 373(6510):123-126.
77. Preston BD, Poiesz BJ, Loeb LA. **Fidelity of HIV-1 reverse transcriptase.** *Science* 1988; 242(4882):1168-1171.
78. Hu WS, Rhodes T, Dang Q, Pathak V. **Retroviral recombination: review of genetic analyses.** *Front Biosci* 2003; 8:d143-155.
79. Mansky LM, Temin HM. **Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase.** *J Virol* 1995; 69(8):5087-5094.
80. Rambaut A, Posada D, Crandall KA, Holmes EC. **The causes and consequences of HIV evolution.** *Nat Rev Genet* 2004; 5(1):52-61.
81. Vogt PK. **Genetically stable reassortment of markers during mixed infection with avian tumor viruses.** *Virology* 1971; 46(3):947-952.

82. Hu WS, Temin HM. **Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination.** *Proc Natl Acad Sci U S A* 1990; 87(4):1556-1560.
83. Clavel F, Hoggan MD, Willey RL, Strebel K, Martin MA, Repaske R. **Genetic recombination of human immunodeficiency virus.** *J Virol* 1989; 63(3):1455-1459.
84. Fang G, Weiser B, Kuiken C, Philpott SM, Rowland-Jones S, Plummer F, et al. **Recombination following superinfection by HIV-1.** *AIDS* 2004; 18(2):153-159.
85. Simon-Lorière E, Galetto R, Negroni M. **Caractéristiques mécanistiques de la recombinaison chez le virus de l'immunodéficience humaine.** *Virologie* 2006; 10:431-442.
86. Junghans RP, Boone LR, Skalka AM. **Retroviral DNA H structures: displacement-assimilation model of recombination.** *Cell* 1982; 30(1):53-62.
87. Jetzt AE, Yu H, Klarmann GJ, Ron Y, Preston BD, Dougherty JP. **High rate of recombination throughout the human immunodeficiency virus type 1 genome.** *J Virol* 2000; 74(3):1234-1240.
88. Lai MM. **RNA recombination in animal and plant viruses.** *Microbiol Rev* 1992; 56(1):61-79.
89. Simon-Lorieri E, Holmes EC. **Why do RNA viruses recombine?** *Nat Rev Microbiol* 2011; 9(8):617-626.
90. Telesnitsky A, Goff SP. **Two defective forms of reverse transcriptase can complement to restore retroviral infectivity.** *EMBO J* 1993; 12(11):4433-4438.
91. Oyama F, Kikuchi R, Crouch RJ, Uchida T. **Intrinsic properties of reverse transcriptase in reverse transcription. Associated RNase H is essentially regarded as an endonuclease.** *J Biol Chem* 1989; 264(31):18808-18817.
92. Krug MS, Berger SL. **Ribonuclease H activities associated with viral reverse transcriptases are endonucleases.** *Proc Natl Acad Sci U S A* 1989; 86(10):3539-3543.
93. Wisniewski M, Balakrishnan M, Palaniappan C, Fay PJ, Bambara RA. **The sequential mechanism of HIV reverse transcriptase RNase H.** *J Biol Chem* 2000; 275(48):37664-37671.
94. Schultz SJ, Zhang M, Champoux JJ. **Recognition of internal cleavage sites by retroviral RNases H.** *J Mol Biol* 2004; 344(3):635-652.
95. Coffin JM. **Structure, replication, and recombination of retrovirus genomes: some unifying hypotheses.** *J Gen Virol* 1979; 42(1):1-26.
96. DeStefano JJ, Mallaber LM, Rodriguez-Rodriguez L, Fay PJ, Bambara RA. **Requirements for strand transfer between internal regions of heteropolymer templates by human immunodeficiency virus reverse transcriptase.** *J Virol* 1992; 66(11):6370-6378.
97. DeStefano JJ, Bambara RA, Fay PJ. **The mechanism of human immunodeficiency virus reverse transcriptase-catalyzed strand transfer from internal regions of heteropolymeric RNA templates.** *J Biol Chem* 1994; 269(1):161-168.
98. Kim JK, Palaniappan C, Wu W, Fay PJ, Bambara RA. **Evidence for a unique mechanism of strand transfer from the transactivation response region of HIV-1.** *J Biol Chem* 1997; 272(27):16769-16777.
99. Balakrishnan M, Fay PJ, Bambara RA. **The kissing hairpin sequence promotes recombination within the HIV-1 5' leader region.** *J Biol Chem* 2001; 276(39):36482-36492.
100. Moumen A, Polomack L, Roques B, Buc H, Negroni M. **The HIV-1 repeated sequence R as a robust hot-spot for copy-choice recombination.** *Nucleic Acids Res* 2001; 29(18):3814-3821.
101. Galetto R, Moumen A, Giacomoni V, Veron M, Charneau P, Negroni M. **The structure of HIV-1 genomic RNA in the gp120 gene determines a recombination hot spot in vivo.** *J Biol Chem* 2004; 279(35):36625-36632.
102. Galetto R, Negroni M. **Mechanistic features of recombination in HIV.** *AIDS Rev* 2005; 7(2):92-102.

103. Simon-Loriere E, Galetto R, Hamoudi M, Archer J, Lefevre P, Martin DP, et al. **Molecular mechanisms of recombination restriction in the envelope gene of the human immunodeficiency virus.** *PLoS Pathog* 2009; 5(5):e1000418.
104. Simon-Loriere E, Martin DP, Weeks KM, Negroni M. **RNA structures facilitate recombination-mediated gene swapping in HIV-1.** *J Virol* 2010; 84(24):12675-12682.
105. Hamoudi M. S-LE, Medjahed H., Sodrosky J, Finzi A, and Negroni M.
106. Baird HA, Gao Y, Galetto R, Lalonde M, Anthony RM, Giacomoni V, et al. **Influence of sequence identity and unique breakpoints on the frequency of intersubtype HIV-1 recombination.** *Retrovirology* 2006; 3:91.
107. Baird HA, Galetto R, Gao Y, Simon-Loriere E, Abreha M, Archer J, et al. **Sequence determinants of breakpoint location during HIV-1 intersubtype recombination.** *Nucleic Acids Res* 2006; 34(18):5203-5216.
108. Fan J, Negroni M, Robertson DL. **The distribution of HIV-1 recombination breakpoints.** *Infect Genet Evol* 2007; 7(6):717-723.
109. Archer J, Pinney JW, Fan J, Simon-Loriere E, Arts EJ, Negroni M, et al. **Identifying the important HIV-1 recombination breakpoints.** *PLoS Comput Biol* 2008; 4(9):e1000178.
110. **The circulating recombinant forms (CRFs).** In: <http://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html>.
111. Jia L, Li L, Gui T, Liu S, Li H, Han J, et al. **Analysis of HIV-1 intersubtype recombination breakpoints suggests region with high pairing probability may be a more fundamental factor than sequence similarity affecting HIV-1 recombination.** *Virol J* 2016; 13(1):156.
112. Choisy M, Woelk CH, Guegan JF, Robertson DL. **Comparative study of adaptive molecular evolution in different human immunodeficiency virus groups and subtypes.** *J Virol* 2004; 78(4):1962-1970.
113. Yang W, Bielawski JP, Yang Z. **Widespread adaptive evolution in the human immunodeficiency virus type 1 genome.** *J Mol Evol* 2003; 57(2):212-221.
114. Liu Y, McNevin J, Cao J, Zhao H, Genowati I, Wong K, et al. **Selection on the human immunodeficiency virus type 1 proteome following primary infection.** *J Virol* 2006; 80(19):9519-9529.
115. Martin DP, van der Walt E, Posada D, Rybicki EP. **The evolutionary value of recombination is constrained by genome modularity.** *PLoS Genet* 2005; 1(4):e51.
116. Drummond DA, Silberg JJ, Meyer MM, Wilke CO, Arnold FH. **On the conservative nature of intragenic recombination.** *Proc Natl Acad Sci U S A* 2005; 102(15):5380-5385.
117. Voigt CA, Martinez C, Wang ZG, Mayo SL, Arnold FH. **Protein building blocks preserved by recombination.** *Nat Struct Biol* 2002; 9(7):553-558.
118. Cramer A, Raillard SA, Bermudez E, Stemmer WP. **DNA shuffling of a family of genes from diverse species accelerates directed evolution.** *Nature* 1998; 391(6664):288-291.
119. Stemmer WP. **Rapid evolution of a protein in vitro by DNA shuffling.** *Nature* 1994; 370(6488):389-391.
120. Meyer MM, Silberg JJ, Voigt CA, Endelman JB, Mayo SL, Wang ZG, et al. **Library analysis of SCHEMA-guided protein recombination.** *Protein Sci* 2003; 12(8):1686-1693.
121. Ostermeier M, Benkovic SJ. **Evolution of protein function by domain swapping.** *Adv Protein Chem* 2000; 55:29-77.
122. Song H, Giorgi EE, Ganusov VV, Cai F, Athreya G, Yoon H, et al. **Tracking HIV-1 recombination to resolve its contribution to HIV-1 evolution in natural infection.** *Nat Commun* 2018; 9(1):1928.
123. Lenormand T, Otto SP. **The evolution of recombination in a heterogeneous environment.** *Genetics* 2000; 156(1):423-438.

124. Otto SP, Lenormand T. **Resolving the paradox of sex and recombination.** *Nat Rev Genet* 2002; 3(4):252-261.
125. Barton NH. **Mutation and the evolution of recombination.** *Philos Trans R Soc Lond B Biol Sci* 2010; 365(1544):1281-1294.
126. Rice WR. **Experimental tests of the adaptive significance of sexual recombination.** *Nat Rev Genet* 2002; 3(4):241-251.
127. Iwabu Y, Mizuta H, Kawase M, Kameoka M, Goto T, Ikuta K. **Superinfection of defective human immunodeficiency virus type 1 with different subtypes of wild-type virus efficiently produces infectious variants with the initial viral phenotypes by complementation followed by recombination.** *Microbes Infect* 2008; 10(5):504-513.
128. Quan Y, Xu H, Wainberg MA. **Defective HIV-1 quasispecies in the form of multiply drug-resistant proviral DNA within cells can be rescued by superinfection with different subtype variants of HIV-1 and by HIV-2 and SIV.** *J Antimicrob Chemother* 2014; 69(1):21-27.
129. Brown RJ, Peters PJ, Caron C, Gonzalez-Perez MP, Stones L, Ankghuambom C, et al. **Intercompartmental recombination of HIV-1 contributes to env intrahost diversity and modulates viral tropism and sensitivity to entry inhibitors.** *J Virol* 2011; 85(12):6024-6037.
130. Ritchie AJ, Cai F, Smith NM, Chen S, Song H, Brackenridge S, et al. **Recombination-mediated escape from primary CD8+ T cells in acute HIV-1 infection.** *Retrovirology* 2014; 11:69.
131. Streeck H, Li B, Poon AF, Schneidewind A, Gladden AD, Power KA, et al. **Immune-driven recombination and loss of control after HIV superinfection.** *J Exp Med* 2008; 205(8):1789-1796.
132. Nora T, Charpentier C, Tenaillon O, Hoede C, Clavel F, Hance AJ. **Contribution of recombination to the evolution of human immunodeficiency viruses expressing resistance to antiretroviral treatment.** *J Virol* 2007; 81(14):7620-7628.
133. Moutouh L, Corbeil J, Richman DD. **Recombination leads to the rapid emergence of HIV-1 dually resistant mutants under selective drug pressure.** *Proc Natl Acad Sci U S A* 1996; 93(12):6106-6111.
134. Nishimura Y, Shingai M, Lee WR, Sadjadpour R, Donau OK, Willey R, et al. **Recombination-mediated changes in coreceptor usage confer an augmented pathogenic phenotype in a nonhuman primate model of HIV-1-induced AIDS.** *J Virol* 2011; 85(20):10617-10626.
135. Liu SL, Mittler JE, Nickle DC, Mulvania TM, Shriner D, Rodrigo AG, et al. **Selection for human immunodeficiency virus type 1 recombinants in a patient with rapid progression to AIDS.** *J Virol* 2002; 76(21):10674-10684.
136. Robertson DL AJ, Bradac JA, et al. **Theoretical Biology and Biophysics Group, Los Alamos National Laboratory.** 1999.
137. Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, et al. **HIV-1 nomenclature proposal.** *Science* 2000; 288(5463):55-56.
138. Thomson MM, Perez-Alvarez L, Najera R. **Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy.** *Lancet Infect Dis* 2002; 2(8):461-471.
139. Abecasis AB, Lemey P, Vidal N, de Oliveira T, Peeters M, Camacho R, et al. **Recombination confounds the early evolutionary history of human immunodeficiency virus type 1: subtype G is a circulating recombinant form.** *J Virol* 2007; 81(16):8543-8551.
140. Ozato K, Shin DM, Chang TH, Morse HC, 3rd. **TRIM family proteins and their emerging roles in innate immunity.** *Nat Rev Immunol* 2008; 8(11):849-860.
141. Howard TM, Rasheed S. **Genomic structure and nucleotide sequence analysis of a new HIV type 1 subtype A strain from Nigeria.** *AIDS Res Hum Retroviruses* 1996; 12(15):1413-1425.



142. Carr JK, Salminen MO, Albert J, Sanders-Buell E, Gotte D, Birx DL, et al. **Full genome sequences of human immunodeficiency virus type 1 subtypes G and A/G intersubtype recombinants.** *Virology* 1998; 247(1):22-31.
143. Lee CN, Chen MY, Lin HS, Lee MC, Luo CC, Twu SJ, et al. **HIV type 1 env subtype A variants in Taiwan.** *AIDS Res Hum Retroviruses* 1998; 14(9):807-809.
144. Liitsola K, Tashkinova I, Laukkanen T, Korovina G, Smolskaja T, Momot O, et al. **HIV-1 genetic subtype A/B recombinant strain causing an explosive epidemic in injecting drug users in Kaliningrad.** *AIDS* 1998; 12(14):1907-1919.
145. Lukashov VV, Huismans R, Rakhmanova AG, Lisitsina ZN, Akhtyrskaya NA, Vlasov NN, et al. **Circulation of subtype A and gagA/envB recombinant HIV type 1 strains among injecting drug users in St. Petersburg, Russia, correlates with geographical origin of infections.** *AIDS Res Hum Retroviruses* 1999; 15(17):1577-1583.
146. Liitsola K, Holm K, Bobkov A, Pokrovsky V, Smolskaya T, Leinikki P, et al. **An AB recombinant and its parental HIV type 1 strains in the area of the former Soviet Union: low requirements for sequence identity in recombination.** UNAIDS Virus Isolation Network. *AIDS Res Hum Retroviruses* 2000; 16(11):1047-1053.
147. Zarandia M, Tsertsvadze T, Carr JK, Nadai Y, Sanchez JL, Nelson AK. **HIV-1 genetic diversity and genotypic drug susceptibility in the Republic of Georgia.** *AIDS Res Hum Retroviruses* 2006; 22(5):470-476.
148. Leoz M, Feyertag F, Charpentier C, Delaugerre C, Wirden M, Lemee V, et al. **Characterization of CRF56\_cpx, a new circulating B/CRF02/G recombinant form identified in MSM in France.** *AIDS* 2013; 27(14):2309-2312.
149. Vessiere A, Rousset D, Kfutwah A, Leoz M, Depatureaux A, Simon F, et al. **Diagnosis and monitoring of HIV-1 group O-infected patients in Cameroun.** *J Acquir Immune Defic Syndr* 2010; 53(1):107-110.
150. De Oliveira F, Mourez T, Vessiere A, Ngoupo PA, Alessandri-Gradt E, Simon F, et al. **Multiple HIV-1/M + HIV-1/O dual infections and new HIV-1/MO inter-group recombinant forms detected in Cameroon.** *Retrovirology* 2017; 14(1):1.
151. Villabona-Arenas CJ, Domyeum J, Mouacha F, Butel C, Delaporte E, Peeters M, et al. **HIV-1 group O infection in Cameroon from 2006 to 2013: Prevalence, genetic diversity, evolution and public health challenges.** *Infect Genet Evol* 2015; 36:210-216.
152. Ayouba A, Mauclere P, Martin PM, Cunin P, Mfoupouendoun J, Njinku B, et al. **HIV-1 group O infection in Cameroon, 1986 to 1998.** *Emerg Infect Dis* 2001; 7(3):466-467.
153. Vergne L, Bourgeois A, Mpoudi-Ngole E, Mougnotou R, Mbuagbaw J, Liegeois F, et al. **Biological and genetic characteristics of HIV infections in Cameroon reveals dual group M and O infections and a correlation between SI-inducing phenotype of the predominant CRF02\_AG variant and disease stage.** *Virology* 2003; 310(2):254-266.
154. Kabeya CM, Esu-Williams E, Eni E, Peeters M, Saman E, Delaporte E. **Evidence for HIV-1 group O infection in Nigeria.** *Lancet* 1995; 346(8970):308.
155. Songok EM, Libondo DK, Rotich MC, Oogo SA, Tukei PM. **Surveillance for HIV-1 subtypes O and M in Kenya.** *Lancet* 1996; 347(9016):1700.
156. Heyndrickx L, Alary M, Janssens W, Davo N, van der Groen G. **HIV-1 group O and group M dual infection in Benin.** *Lancet* 1996; 347(9005):902-903.
157. Peeters M, Gaye A, Mboup S, Badombena W, Bassabi K, Prince-David M, et al. **Presence of HIV-1 group O infection in West Africa.** *AIDS* 1996; 10(3):343-344.
158. Peeters M, Gueye A, Mboup S, Bibollet-Ruche F, Ekaza E, Mulanga C, et al. **Geographical distribution of HIV-1 group O viruses in Africa.** *AIDS* 1997; 11(4):493-498.

159. Nkengasong J, Sylla-Koko F, Peeters M, Ellenberger D, Sassan-Morokro M, Ekpini RA, et al. **HIV-1 group O virus infection in Abidjan, Cote d'Ivoire.** *AIDS* 1998; 12(12):1565-1566.
160. Hampl H, Sawitzky D, Stoffler-Meilicke M, Groh A, Schmitt M, Eberle J, et al. **First case of HIV-1 subtype O infection in Germany.** *Infection* 1995; 23(6):369-370.
161. Rayfield MA, Sullivan P, Bandea CI, Britvan L, Otten RA, Pau CP, et al. **HIV-1 group O virus identified for the first time in the United States.** *Emerg Infect Dis* 1996; 2(3):209-212.
162. Soriano V, Gutierrez M, Garcia-Lerma G, Aguilera O, Mas A, Bravo R, et al. **First case of HIV-1 group O infection in Spain.** *Vox Sang* 1996; 71(1):66.
163. Gould K. **Infection with HIV-1 group O.** *AIDS Patient Care STDS* 1997; 11(6):399-405.
164. Sullivan PS, Do AN, Ellenberger D, Pau CP, Paul S, Robbins K, et al. **Human immunodeficiency virus (HIV) subtype surveillance of African-born persons at risk for group O and group N HIV infections in the United States.** *J Infect Dis* 2000; 181(2):463-469.
165. Agut H, Candotti D, Rabanel B, Huraux JM, Remy G, Ingrand D, et al. **Isolation of atypical HIV-1-related retrovirus from AIDS patient.** *Lancet* 1992; 340(8820):681-682.
166. Depatureaux A, Leoz M, De Oliveira F, Gueudin M, Damond F, Descamps D, et al. **[Specific diagnosis and follow-up of HIV-1 group O infection: RES-O data].** *Med Mal Infect* 2010; 40(12):669-676.
167. Kluge SF, Mack K, Iyer SS, Pujol FM, Heigle A, Learn GH, et al. **Nef proteins of epidemic HIV-1 group O strains antagonize human tetherin.** *Cell Host Microbe* 2014; 16(5):639-650.
168. Sauter D, Kirchhoff F. **Key Viral Adaptations Preceding the AIDS Pandemic.** *Cell Host Microbe* 2019; 25(1):27-38.
169. Arien KK, Abraha A, Quinones-Mateu ME, Kestens L, Vanham G, Arts EJ. **The replicative fitness of primary human immunodeficiency virus type 1 (HIV-1) group M, HIV-1 group O, and HIV-2 isolates.** *J Virol* 2005; 79(14):8979-8990.
170. Kouanfack C, Unal G, Schaeffer L, Kfutwah A, Aghokeng A, Mougnotou R, et al. **Comparative immunovirological and clinical responses to antiretroviral therapy between HIV-1 group O- and HIV-1 group M-infected patients (ANRS 12168 DynaMO study).** *Clin Infect Dis* 2019.
171. Unal G. **Infection par le VIH-1 groupe O : Etude des caractéristiques épidémiologiques et de la réponse immuno-virologique aux antirétroviraux:** Paris-Saclay; 2018.
172. Leoz M. **HIV-1 Group O Genetic Diversity: Characterization, Evolution and Associated Viral Properties:** Rouen Normandie; 2016.
173. Roques P, Robertson DL, Souquiere S, Damond F, Ayouba A, Farfara I, et al. **Phylogenetic analysis of 49 newly derived HIV-1 group O strains: high viral diversity but no group M-like subtype structure.** *Virology* 2002; 302(2):259-273.
174. Quinones-Mateu ME, Albright JL, Mas A, Soriano V, Arts EJ. **Analysis of pol gene heterogeneity, viral quasispecies, and drug resistance in individuals infected with group O strains of human immunodeficiency virus type 1.** *J Virol* 1998; 72(11):9002-9015.
175. Leoz M, Depatureaux A, Vessiere A, Roquebert B, Damond F, Rousset D, et al. **Integrase polymorphism and HIV-1 group O diversity.** *AIDS* 2008; 22(10):1239-1243.
176. Ayouba A, Souquieres S, Njinku B, Martin PM, Muller-Trutwin MC, Roques P, et al. **HIV-1 group N among HIV-1-seropositive individuals in Cameroon.** *AIDS* 2000; 14(16):2623-2625.
177. Bodelle P, Vallari A, Coffey R, McArthur CP, Beyeme M, Devare SG, et al. **Identification and genomic sequence of an HIV type 1 group N isolate from Cameroon.** *AIDS Res Hum Retroviruses* 2004; 20(8):902-908.

178. Delaugerre C, De Oliveira F, Lascoux-Combe C, Plantier JC, Simon F. **HIV-1 group N: travelling beyond Cameroon.** *Lancet* 2011; 378(9806):1894.
179. Rodgers MA, Vallari AS, Harris B, Yamaguchi J, Holzmayer V, Forberg K, et al. **Identification of rare HIV-1 Group N, HBV AE, and HTLV-3 strains in rural South Cameroon.** *Virology* 2017; 504:141-151.
180. Rodgers MA, Vallari AS, Yamaguchi J, Holzmayer V, Harris B, Toure-Kane C, et al. **ARCHITECT HIV Combo Ag/Ab and RealTime HIV-1 Assays Detect Diverse HIV Strains in Clinical Specimens.** *AIDS Res Hum Retroviruses* 2018; 34(3):314-318.
181. Roques P, Robertson DL, Souquiere S, Apetrei C, Nerrienet E, Barre-Sinoussi F, et al. **Phylogenetic characteristics of three new HIV-1 N strains and implications for the origin of group N.** *AIDS* 2004; 18(10):1371-1381.
182. Vallari A, Bodelle P, Ngansop C, Makamche F, Ndembi N, Mbanya D, et al. **Four new HIV-1 group N isolates from Cameroon: Prevalence continues to be low.** *AIDS Res Hum Retroviruses* 2010; 26(1):109-115.
183. Yamaguchi J, Coffey R, Vallari A, Ngansop C, Mbanya D, Ndembi N, et al. **Identification of HIV type 1 group N infections in a husband and wife in Cameroon: viral genome sequences provide evidence for horizontal transmission.** *AIDS Res Hum Retroviruses* 2006; 22(1):83-92.
184. Yamaguchi J, McArthur CP, Vallari A, Coffey R, Bodelle P, Beyeme M, et al. **HIV-1 Group N: evidence of ongoing transmission in Cameroon.** *AIDS Res Hum Retroviruses* 2006; 22(5):453-457.
185. Ngoupou PA. **Fréquence et profil génétique des doubles infections VIH-1/M+O et formes recombinantes VIH-1/MO au Cameroun:** Rouen Normandie; 2016.
186. Damond F, Worobey M, Campa P, Farfara I, Colin G, Matheron S, et al. **Identification of a highly divergent HIV type 2 and proposal for a change in HIV type 2 classification.** *AIDS Res Hum Retroviruses* 2004; 20(6):666-672.
187. Gao F, Yue L, Robertson DL, Hill SC, Hui H, Biggar RJ, et al. **Genetic diversity of human immunodeficiency virus type 2: evidence for distinct sequence subtypes with differences in virus biology.** *J Virol* 1994; 68(11):7433-7447.
188. Ibe S, Yokomaku Y, Shiino T, Tanaka R, Hattori J, Fujisaki S, et al. **HIV-2 CRF01\_AB: first circulating recombinant form of HIV-2.** *J Acquir Immune Defic Syndr* 2010; 54(3):241-247.
189. Yamaguchi J, Vallari A, Ndembi N, Coffey R, Ngansop C, Mbanya D, et al. **HIV type 2 intergroup recombinant identified in Cameroon.** *AIDS Res Hum Retroviruses* 2008; 24(1):86-91.
190. Hemelaar J, Gouws E, Ghys PD, Osmanov S. **Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004.** *AIDS* 2006; 20(16):W13-23.
191. Motomura K, Chen J, Hu WS. **Genetic recombination between human immunodeficiency virus type 1 (HIV-1) and HIV-2, two distinct human lentiviruses.** *J Virol* 2008; 82(4):1923-1933.
192. Peeters M, Liegeois F, Torimiro N, Bourgeois A, Mpoudi E, Vergne L, et al. **Characterization of a highly replicative intergroup M/O human immunodeficiency virus type 1 recombinant isolated from a Cameroonian patient.** *J Virol* 1999; 73(9):7368-7375.
193. Yamaguchi J, Bodelle P, Vallari AS, Coffey R, McArthur CP, Schochetman G, et al. **HIV infections in northwestern Cameroon: identification of HIV type 1 group O and dual HIV type 1 group M and group O infections.** *AIDS Res Hum Retroviruses* 2004; 20(9):944-957.
194. Takehisa J, Zekeng L, Ido E, Yamaguchi-Kabata Y, Mboudjeka I, Harada Y, et al. **Human immunodeficiency virus type 1 intergroup (M/O) recombination in cameroon.** *J Virol* 1999; 73(8):6810-6820.
195. Takehisa J, Zekeng L, Ido E, Mboudjeka I, Moriyama H, Miura T, et al. **Various types of HIV mixed infections in Cameroon.** *Virology* 1998; 245(1):1-10.

196. Brand D, Beby-Defaux A, Mace M, Brunet S, Moreau A, Godet C, et al. **First identification of HIV-1 groups M and O dual infections in Europe.** *AIDS* 2004; 18(18):2425-2428.
197. Plantier JC, Lemee V, Dorval I, Gueudin M, Braun J, Hutin P, et al. **HIV-1 group M superinfection in an HIV-1 group O-infected patient.** *AIDS* 2004; 18(18):2444-2446.
198. De Oliveira F, Cappy P, Lemee V, Moisan A, Pronier C, Bocket L, et al. **Detection of numerous HIV-1/MO recombinants in France.** *AIDS* 2018; 32(10):1289-1299.
199. Ngoupo PA, Sadeuh-Mba SA, De Oliveira F, Ngono V, Ngono L, Tchendjou P, et al. **First evidence of transmission of an HIV-1 M/O intergroup recombinant virus.** *AIDS* 2016; 30(1):1-8.
200. Vessiere A, Leoz M, Brodard V, Strady C, Lemee V, Depatureaux A, et al. **First evidence of a HIV-1 M/O recombinant form circulating outside Cameroon.** *AIDS* 2010; 24(7):1079-1082.
201. Moisan A. **Formes recombinantes VIH-1 inter-groupes M et O : profils moléculaires et rôle potentiel des facteurs de restriction cellulaires et des réseaux de co-évolution protéique:** Rouen Normandie; 2016.
202. Brennan CA, Bodelle P, Coffey R, Devare SG, Golden A, Hackett J, Jr., et al. **The prevalence of diverse HIV-1 strains was stable in Cameroonian blood donors from 1996 to 2004.** *J Acquir Immune Defic Syndr* 2008; 49(4):432-439.
203. Tongo M, Martin DP, Zembe L, Mpoudi-Ngole E, Williamson C, Burgers WA. **Characterization of HIV-1 gag and nef in Cameroon: further evidence of extreme diversity at the origin of the HIV-1 group M epidemic.** *Virology* 2013; 10:29.
204. Magiorkinis G, Paraskevis D, Vandamme AM, Magiorkinis E, Sypsa V, Hatzakis A. **In vivo characteristics of human immunodeficiency virus type 1 intersubtype recombination: determination of hot spots and correlation with sequence similarity.** *J Gen Virol* 2003; 84(Pt 10):2715-2722.
205. Unal G, Alessandri-Gradt E, Leoz M, Pavie J, Lefevre C, Panjo H, et al. **Human Immunodeficiency Virus Type 1 Group O Infection in France: Clinical Features and Immunovirological Response to Antiretrovirals.** *Clin Infect Dis* 2018; 66(11):1785-1793.
206. Vallejo A, Heredia A, Mas A, Lee SF, Epstein JS, Soriano V, et al. **Tropism, coreceptor use, and phylogenetic analysis of both the V3 loop and the protease gene of three novel HIV-1 group O isolates.** *J Acquir Immune Defic Syndr Hum Retrovirol* 1998; 18(5):417-425.
207. Depatureaux A, Charpentier C, Leoz M, Unal G, Damond F, Kfutwah A, et al. **Impact of HIV-1 group O genetic diversity on genotypic resistance interpretation by algorithms designed for HIV-1 group M.** *J Acquir Immune Defic Syndr* 2011; 56(2):139-145.
208. Alessandri-Gradt E, Collin G, Tourneroché A, Bertine M, Leoz M, Charpentier C, et al. **HIV-1 non-group M phenotypic susceptibility to integrase strand transfer inhibitors.** *J Antimicrob Chemother* 2017; 72(9):2431-2437.
209. Depatureaux A, Mesplede T, Quashie P, Oliveira M, Moisi D, Plantier JC, et al. **HIV-1 Group O Resistance Against Integrase Inhibitors.** *J Acquir Immune Defic Syndr* 2015; 70(1):9-15.
210. WHO. **WHO recommends dolutegravir as preferred HIV treatment option in all populations.** In; 2019.
211. Geuenich S, Kaderali L, Allespach I, Sertel S, Keppler OT. **Biological signature characteristics of primary isolates from human immunodeficiency virus type 1 group O in ex vivo human tonsil histocultures.** *J Virol* 2009; 83(20):10494-10503.
212. Sauter D, Hotter D, Van Driessche B, Sturzel CM, Kluge SF, Wildum S, et al. **Differential regulation of NF-kappaB-mediated proviral and antiviral host gene expression by primate lentiviral Nef and Vpu proteins.** *Cell Rep* 2015; 10(4):586-599.
213. Mack K, Starz K, Sauter D, Langer S, Bibollet-Ruche F, Learn GH, et al. **Efficient Vpu-Mediated Tetherin Antagonism by an HIV-1 Group O Strain.** *J Virol* 2017; 91(6).

214. Collman R, Balliet JW, Gregory SA, Friedman H, Kolson DL, Nathanson N, et al. **An infectious molecular clone of an unusual macrophage-tropic and highly cytopathic strain of human immunodeficiency virus type 1.** *J Virol* 1992; 66(12):7517-7521.
215. Novelli P, Vella C, Oxford J, Daniels RS. **Construction and characterization of a full-length HIV-1(92UG001) subtype D infectious molecular clone.** *AIDS Res Hum Retroviruses* 2002; 18(1):85-88.
216. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, 3rd, Smith HO. **Enzymatic assembly of DNA molecules up to several hundred kilobases.** *Nat Methods* 2009; 6(5):343-345.
217. Gibson DG, Smith HO, Hutchison CA, 3rd, Venter JC, Merryman C. **Chemical synthesis of the mouse mitochondrial genome.** *Nat Methods* 2010; 7(11):901-903.
218. Tangy F, Charneau P, Jacob Y. **Cells and methodology to generate non-segmented negative strand RNA viruses.** *EP Application, Editor* 2006.
219. Salter RD, Howell DN, Cresswell P. **Genes regulating HLA class I antigen expression in T-B lymphoblast hybrids.** *Immunogenetics* 1985; 21(3):235-246.
220. Zhou J, Aiken C. **Nef enhances human immunodeficiency virus type 1 infectivity resulting from interviral fusion: evidence supporting a role for Nef at the virion envelope.** *J Virol* 2001; 75(13):5851-5859.
221. Selig L, Pages JC, Tanchou V, Preveral S, Berlioz-Torrent C, Liu LX, et al. **Interaction with the p6 domain of the gag precursor mediates incorporation into virions of Vpr and Vpx proteins from primate lentiviruses.** *J Virol* 1999; 73(1):592-600.
222. Beaumont E, Vendrame D, Verrier B, Roch E, Biron F, Barin F, et al. **Matrix and envelope coevolution revealed in a patient monitored since primary infection with human immunodeficiency virus type 1.** *J Virol* 2009; 83(19):9875-9889.
223. Chakraborty A, Sun GQ, Mustavich L, Huang SH, Li BL. **Biochemical interactions between HIV-1 integrase and reverse transcriptase.** *FEBS Lett* 2013; 587(5):425-429.
224. Arachchilage MH, Piontkivska H. **Coevolutionary Analysis Identifies Protein-Protein Interaction Sites between HIV-1 Reverse Transcriptase and Integrase.** *Virus Evol* 2016; 2(1).
225. Tebit DM, Zekeng L, Kaptue L, Gurtler L, Fackler OT, Keppler OT, et al. **Construction and characterization of an HIV-1 group O infectious molecular clone and analysis of vpr- and nef-negative derivatives.** *Virology* 2004; 326(2):329-339.
226. Ly TD, Plantier JC, Leballais L, Gonzalo S, Lemee V, Laperche S. **The variable sensitivity of HIV Ag/Ab combination assays in the detection of p24Ag according to genotype could compromise the diagnosis of early HIV infection.** *J Clin Virol* 2012; 55(2):121-127.
227. Braun J, Plantier JC, Hellot MF, Tuillon E, Gueudin M, Damond F, et al. **A new quantitative HIV load assay based on plasma virion reverse transcriptase activity for the different types, groups and subtypes.** *Aids* 2003; 17(3):331-336.
228. Mohri H, Prada N, Markowitz M. **Viral envelope is a major determinant of enhanced fitness of a multidrug-resistant HIV-1 variant.** *J Acquir Immune Defic Syndr* 2015; 68(5):487-494.
229. Marozsan AJ, Moore DM, Lobritz MA, Fraundorf E, Abrahams A, Reeves JD, et al. **Differences in the fitness of two diverse wild-type human immunodeficiency virus type 1 isolates are related to the efficiency of cell binding and entry.** *J Virol* 2005; 79(11):7121-7134.
230. Palmer S, Kearney M, Maldarelli F, Halvas EK, Bixby CJ, Bazmi H, et al. **Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis.** *J Clin Microbiol* 2005; 43(1):406-413.

231. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, et al. **Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection.** *Proc Natl Acad Sci U S A* 2008; 105(21):7552-7557.
232. Mourez T, Simon F, Plantier JC. **Non-M variants of human immunodeficiency virus type 1.** *Clin Microbiol Rev* 2013; 26(3):448-461.
233. Mourez T, Lemee V, Delbos V, Delaugerre C, Alessandri-Gradt E, Etienne M, et al. **HIV rapid screening tests and self-tests: Be aware of differences in performance and cautious of vendors.** *EBioMedicine* 2018; 37:382-391.
234. Lemee V, Leoz M, Etienne M, De Oliveira F, Plantier JC. **Performance of the Liaison XL Murex HIV Ab/Ag test on clinical samples representing current epidemic HIV variants.** *J Clin Microbiol* 2014; 52(9):3277-3279.
235. **ANRS resistance algorithm.** In: <http://www.hivfrenchresistance.org>.
236. Simon F, Souquiere S, Damond F, Kfutwah A, Makuwa M, Leroy E, et al. **Synthetic peptide strategy for the detection of and discrimination among highly divergent primate lentiviruses.** *AIDS Res Hum Retroviruses* 2001; 17(10):937-952.
237. Barin F, Plantier JC, Brand D, Brunet S, Moreau A, Liandier B, et al. **Human immunodeficiency virus serotyping on dried serum spots as a screening tool for the surveillance of the AIDS epidemic.** *J Med Virol* 2006; 78 Suppl 1:S13-18.
238. Plantier JC, Dachraoui R, Lemee V, Gueudin M, Borsa-Lebas F, Caron F, et al. **HIV-1 resistance genotyping on dried serum spots.** *Aids* 2005; 19(4):391-397.
239. Dachraoui R, Depatureaux A, Chakroun M, Fodha I, Letaief A, Trabelsi A, et al. **Monitoring of HIV-1 resistance in Tunisia (North Africa) with a dried plasma spot strategy.** *J Acquir Immune Defic Syndr* 2008; 47(4):522-525.
240. Heyndrickx L, Janssens W, Gurtler L, Zekeng L, Loussert-Ajaka I, Vereecken K, et al. **Differential diagnosis of HIV type 1 group O and M infection by polymerase chain reaction and PstI restriction analysis of the pol gene fragment.** *AIDS Res Hum Retroviruses* 1998; 14(11):973-977.
241. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, et al. **Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination.** *J Virol* 1999; 73(1):152-160.
242. Saitou N, Nei M. **The neighbor-joining method: a new method for reconstructing phylogenetic trees.** *Mol Biol Evol* 1987; 4(4):406-425.
243. Kimura M. **A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences.** *J Mol Evol* 1980; 16(2):111-120.
244. Gueudin M, Leoz M, Lemee V, De Oliveira F, Vessiere A, Kfutwah A, et al. **A new real-time quantitative PCR for diagnosis and monitoring of HIV-1 group O infection.** *J Clin Microbiol* 2012; 50(3):831-836.
245. Cappy P. **Profil de recombinaison, genèse et étude fonctionnelle des recombinants VIH-MO : à la frontière de nouvelles espèces de VIH.** 2013.



# **ANNEXES**





### **ANNEXE 1 : Mélange réactionnel et programme d'amplification utilisés pour la génération des fragments à assembler par la méthode Gibson Assembly**

Les mélanges réactionnels contenaient 33,1µL d'eau, 2µL de chacune des deux amorces (sens et antisens), 5µL de Reaction Mix, 2,5µL de dNTP Mix, 0,4µL d'Enzyme Mix Long Range et 5µL d'ADN issu des MiniPrep et dilué au 1/10<sup>ème</sup>. Dans un but d'optimisation, le programme d'amplification avait recours à un système d'incrémentation, correspondant à une augmentation progressive de la durée d'élongation au fil des cycles. Ainsi le programme d'amplification comportait un cycle de dénaturation initiale de 3 minutes à 93°C, suivi de 10 premiers cycles constitués par une dénaturation de 30 secondes à 93°C, une hybridation de 30 secondes à 57°C et une élongation de 6 minutes à 68°C. Ensuite, 28 autres cycles avaient lieu, constitués par une dénaturation de 15 secondes à 93°C, une hybridation de 30 secondes à 57°C et une élongation augmentée de 20 secondes à chaque cycle, à température constante de 68°C. Enfin, un cycle d'élongation de 10 minutes à 68°C achevait la réaction d'amplification.

---

### **ANNEXE 2 : Mélange réactionnel et programme d'amplification utilisés lors des vérifications sur les CMIC candidats générés par la méthode Gibson Assembly**

Trois microlitres de surnageant de centrifugation ont été ajoutés à un mélange réactionnel contenant 13µL de *HotStartaq MasterMix (Qiagen®)*, 1µL de MgCl<sub>2</sub> à 25mM, 1µL de chacune des deux amorces sens et anti-sens, et 6µL d'eau. Le programme d'amplification comportait un cycle de dénaturation initiale de 15 minutes à 95°C, suivi de 35 cycles constitués par une dénaturation de 30 secondes à 94°C, une hybridation de 30 secondes à 55°C et une élongation de 1 minute à 72°C. Une étape d'élongation finale de 7 minutes à 72°C achevait l'amplification.

---

### **ANNEXE 3 : Mélange réactionnel et programme d'amplification des LTR-ARN utilisés lors de la PCR externe**

Pour réaliser la RT-PCR externe, 10µL d'extrait ont été ajoutés à un mélange réactionnel de 40µL contenant 25µL de mix provenant du kit *Superscript III (Invitrogen®)*, 8µL de MgSO<sub>4</sub> 5mM, 2µL de chaque amorce, 1µL d'enzyme *RT Platinum TAQ HiFi (Invitrogen®)* et enfin, 2µL d'eau. Le programme d'amplification comportait une étape préalable de rétrotranscription de 30 minutes à 50°C, puis un cycle de dénaturation initiale de 2 minutes à 94°C, suivi de 35 cycles constitués par une dénaturation de 30 secondes à 94°C, une hybridation de 30 secondes à 55°C et une élongation de 2 minute à 68°C. Une étape d'élongation finale de 10 minutes à 68°C achevait l'amplification.

---

### **ANNEXE 4 : Mélange réactionnel et programme d'amplification des LTR-ADN utilisés lors de la PCR externe**

La PCR externe a été réalisée en ajoutant 5µL d'ADN intracellulaire extrait à un mélange réactionnel contenant 25µL de *HotStartaq MasterMix (Qiagen®)*, 2µL de MgCl<sub>2</sub> à 25mM, 2µL de chacune des deux

amorces sens et anti-sens, et 14µL d'eau, portant le volume total à 50µL. Le programme d'amplification comportait un cycle de dénaturation initiale de 15 minutes à 94°C, suivi de 35 cycles constitués par une dénaturation de 30 secondes à 94°C, une hybridation de 30 secondes à 50°C et une élongation de 2 minutes 30 à 72°C. Une étape d'élongation finale de 10 minutes à 72°C achevait l'amplification.

---

#### **ANNEXE 5 : Mélange réactionnel et programme d'amplification des LTRs utilisés lors de la PCR interne**

Lors de la PCR interne, 2µL de chaque amplicon de la PCR externe ont été ajoutés à un mix réactionnel de 48µL contenant 25µL de *HotStartaq MasterMix (Qiagen®)*, 2µL de MgCl<sub>2</sub> à 25mM, 2µL de chaque amorce et 17µL d'eau pure. Le programme d'amplification comportait un cycle de dénaturation initiale de 15 minutes à 95°C, suivi de 35 cycles constitués par une dénaturation de 30 secondes à 94°C, une hybridation de 30 secondes à 50°C et une élongation de 1 minute 30 à 72°C. Une étape d'élongation finale de 10 minutes à 72°C achevait l'amplification.

---

#### **ANNEXE 6 : Mélange réactionnel et programme d'amplification pour la synthèse de l'ADNc à partir des extraits d'ARN viral, en vue du SGA**

Une première étape de dénaturation de l'ARN a été réalisée en présence de 8µL d'ARN, 1µL de dNTPs à 10µM et 1µL de l'amorce 241L à 10µM. Le mélange réactionnel a été incubé pendant 5 minutes à 65°C. La suite de la réaction a été réalisée en complétant le mélange par 10µL d'un mix contenant 4µL de MgCl<sub>2</sub> à 25mM, 2µL de tampon 10X, 2µL de dithiothreitol à 0,1M, 1µL d'inhibiteur de la RNase (RNaseOUT) à 40U/µL et 1µL SuperScript III (SSIII) à 200U/µL. La synthèse d'ADNc a eu lieu à 50°C pendant 50 minutes et a été suivie par une élongation finale à 85°C pendant 5 min, puis d'un arrêt de la réaction à 4°C.

---

#### **ANNEXE 7 : Mélanges réactionnels et programmes d'amplification du fragment GC 1, utilisés lors du séquençage de génome complet**

La PCR externe a été réalisée en ajoutant 5µL d'ADN intracellulaire extrait à un mélange réactionnel contenant 25µL de *HotStartaq MasterMix (Qiagen®)*, 2µL de MgCl<sub>2</sub> à 25mM, 2µL de chacune des deux amorces sens et anti-sens, et 14µL d'eau, portant le volume total à 50µL. Le programme d'amplification comportait un cycle de dénaturation initiale de 15 minutes à 94°C, suivi de 35 cycles constitués par une dénaturation de 30 secondes à 94°C, une hybridation de 30 secondes à 50°C et une élongation de 2 minutes 30 à 72°C. Une étape d'élongation finale de 10 minutes à 72°C achevait l'amplification.

Lors de la PCR interne, 2µL de chaque amplicon de la PCR externe ont été ajoutés à un mix réactionnel de 48µL contenant 25µL de *HotStartaq MasterMix (Qiagen®)*, 2µL de MgCl<sub>2</sub> à 25mM, 2µL de chaque amorce et 17µL d'eau pure. Le programme d'amplification comportait un cycle de dénaturation initiale de 15 minutes à 95°C, suivi de 35 cycles constitués par une dénaturation de 30 secondes à 94°C, une hybridation de 30 secondes à 50°C et une élongation de 1 minute 30 à 72°C. Une étape d'élongation finale de 10 minutes à 72°C achevait l'amplification.

### **ANNEXE 8 : Mélanges réactionnels et programmes d'amplification des fragments GC 2 et GC 3, utilisés lors du séquençage de génome complet**

Lors des PCRs externes, les mélanges réactionnels contenaient 33,1µL d'eau, 2µL de chacune des deux amorces (sens et antisens), 5µL de Reaction Mix, 2,5µL de dNTP Mix, 0,4µL d'Enzyme Mix Long Range et 5µL d'ADN intracellulaire extrait. Le programme d'amplification avait recours à un système d'incrémentatation, correspondant à une augmentation progressive de la durée d'élongation au fil des cycles. Ainsi, il comportait un cycle de dénaturation initiale de 3 minutes à 93°C, suivi de 10 premiers cycles constitués par une dénaturation de 15 secondes à 93°C, une hybridation de 30 secondes à 58°C et une élongation de 6 minutes à 68°C. Ensuite, 28 autres cycles avaient lieu, constitués par une dénaturation de 15 secondes à 93°C, une hybridation de 30 secondes à 58°C et une élongation augmentée de 20 secondes à chaque cycle, à température constante de 68°C. Enfin, un cycle d'élongation de 10 minutes à 68°C achevait la réaction d'amplification.

Lors des PCRs internes, les mélanges réactionnels contenaient 36,1µL d'eau, 2µL de chacune des deux amorces (sens et antisens), 5µL de Reaction Mix, 2,5µL de dNTP Mix, 0,4µL d'Enzyme Mix Long Range et 2µL d'amplicon issu de la PCR externe. Le programme d'amplification avait également recours à un système d'incrémentatation et comportait un cycle de dénaturation initiale de 3 minutes à 93°C, suivi de 35 cycles constitués par une dénaturation de 15 secondes à 93°C, une hybridation de 30 secondes à 50°C et une élongation augmentée de 20 secondes à chaque cycle, à température constante de 68°C. Enfin, un cycle d'élongation de 10 minutes à 68°C achevait la réaction d'amplification.

---

## ANNEXE 9 : Lettre à l'éditeur

*In vivo* emergence of an HIV-1/MO recombinant revealed undiagnosed HI-1/M+O co-infection

## ARTICLE IN PRESS

Clinical Microbiology and Infection xxx (xxxx) xxx



Contents lists available at ScienceDirect

## Clinical Microbiology and Infection

journal homepage: [www.clinicalmicrobiologyandinfection.com](http://www.clinicalmicrobiologyandinfection.com)

## Letter to the Editor

*In vivo* emergence of an HIV-1/MO recombinant revealed undiagnosed HIV-1/M+O co-infectionA. Moisan<sup>1,\*</sup>, F. De Oliveira<sup>1</sup>, C. Pronier<sup>2</sup>, P. Cappy<sup>1,\*</sup>, A. Maillard<sup>2</sup>, J.-C. Plantier<sup>1,\*</sup><sup>1</sup> Normandie Université, UNIROUEN, GRAM 2.0, Rouen University Hospital, Laboratoire de Virologie associé au CNR du VIH, Rouen, France<sup>2</sup> Laboratoire de Virologie, CHU Pontchaillou, Rennes, France

## ARTICLE INFO

## Article history:

Received 7 June 2019  
 Received in revised form  
 21 August 2019  
 Accepted 22 August 2019  
 Available online xxx

Editor: L. Kaiser

## To the Editor,

For the past 20 years, several case reports and epidemiological studies have described human immunodeficiency virus type 1 strain M+O (HIV-1/M+O) dual infections and HIV-1/MO recombinant forms, leading to the definition of 19 unique recombinant forms of HIV-1/MO. These cases corresponded to: (i) one HIV-1/M superinfection in an HIV-1/O-infected individual, (ii) 17 M+O dual infections [1,2], (iii) nine co-infections with an HIV-1/MO recombinant form associated with one or both detectable parental strains [1–5], and (iv) ten MO recombinant forms with no parental strain [1,2]. None of these cases described the viral dynamic during generation of the HIV-1/MO recombinant form *in vivo*, owing to the scarcity of sequential data.

Here, we describe the dynamic of emergence of an HIV-1/MO recombinant form in a context of undiagnosed HIV-1/M+O co-infection.

In 1999, a 32-year-old woman (RBF235), born in Chad and having made many trips to neighbouring Cameroon, was diagnosed as HIV-1/M-positive in Paris, France. In 2003, she moved to Rennes, in western France, where she was followed for an HIV-1/M subtype

D infection. In 2013, in a context of virological failure, a routine drug-resistance test was performed, using the French National Agency for Research on AIDS in-house technique, which revealed the presence of HIV-1/O in the protease (PR) region and of HIV-1/M in the reverse transcriptase (RT) region. To investigate these results suggesting an HIV-1/M+O or HIV-1/MO infection, five RNA-positive samples dating from before 2013 (October 2007, November 2009, November 2011, January 2013 and November 2013) were sent to the French National Reference Centre for HIV in Rouen, France.

First, National Reference Centre (NRC) envelope serotyping showed an M+O dual reactivity for each sample. Molecular confirmation of M+O co-infection was performed, based on HIV-1/M- and HIV-1/O-specific nested (RT-)PCRs targeting four genomic regions (PR, RT, integrase (IN) and gp41), as we previously described. Owing to discordant drug-resistance data, a nested RT-PCR encompassing the PR-RT region was designed to specifically amplify the (O-M) recombinant profile. Sequential sequences obtained from those group-specific and recombinant (RT-)PCRs confirmed the M+O dual infection since 2007 at least, and revealed the emergence of an HIV-1/MO recombinant virus between January and November 2013 (Fig. 1a). The near full-length genome of this recombinant, sequenced from intracellular DNA as previously described [1], confirmed an (O-M) breakpoint in the RT and showed a second breakpoint in the long terminal repeat (LTR) regions (Fig. 1b).

The replication levels of the different viral populations were measured with non-group-specific assay and with group M-specific and group O-specific assays. Nevertheless, group-specific viral load assays failed to specifically quantify the recombinant, due to its genomic profile (Fig. 1b). Global data suggested replication of both M and O parental strains between 2007 and 2013 at least. The absence of amplification of some HIV-1/M regions in 2013 could have been due to the lower replication of the HIV-1/M population at this time (Fig. 1a).

Sequential samples and atypical results allowed us to reveal an M+O dual infection and to track the emergence of an MO recombinant form in an individual initially diagnosed as HIV-1/M. This case highlights the complexity of HIV infection diagnosis when several divergent viral populations are involved. Current diagnostic tools (non-specific screening and viral load assays) now allow the

\* Corresponding author. J.C. Plantier.

\*\* Corresponding author. A. Moisan.

E-mail addresses: [alice.moisan@chu-rouen.fr](mailto:alice.moisan@chu-rouen.fr) (A. Moisan), [jc.plantier@chu-rouen.fr](mailto:jc.plantier@chu-rouen.fr) (J.-C. Plantier).

\* P. Cappy is currently working at Institut National de la Transfusion Sanguine, Paris, France.

<https://doi.org/10.1016/j.cmi.2019.08.016>

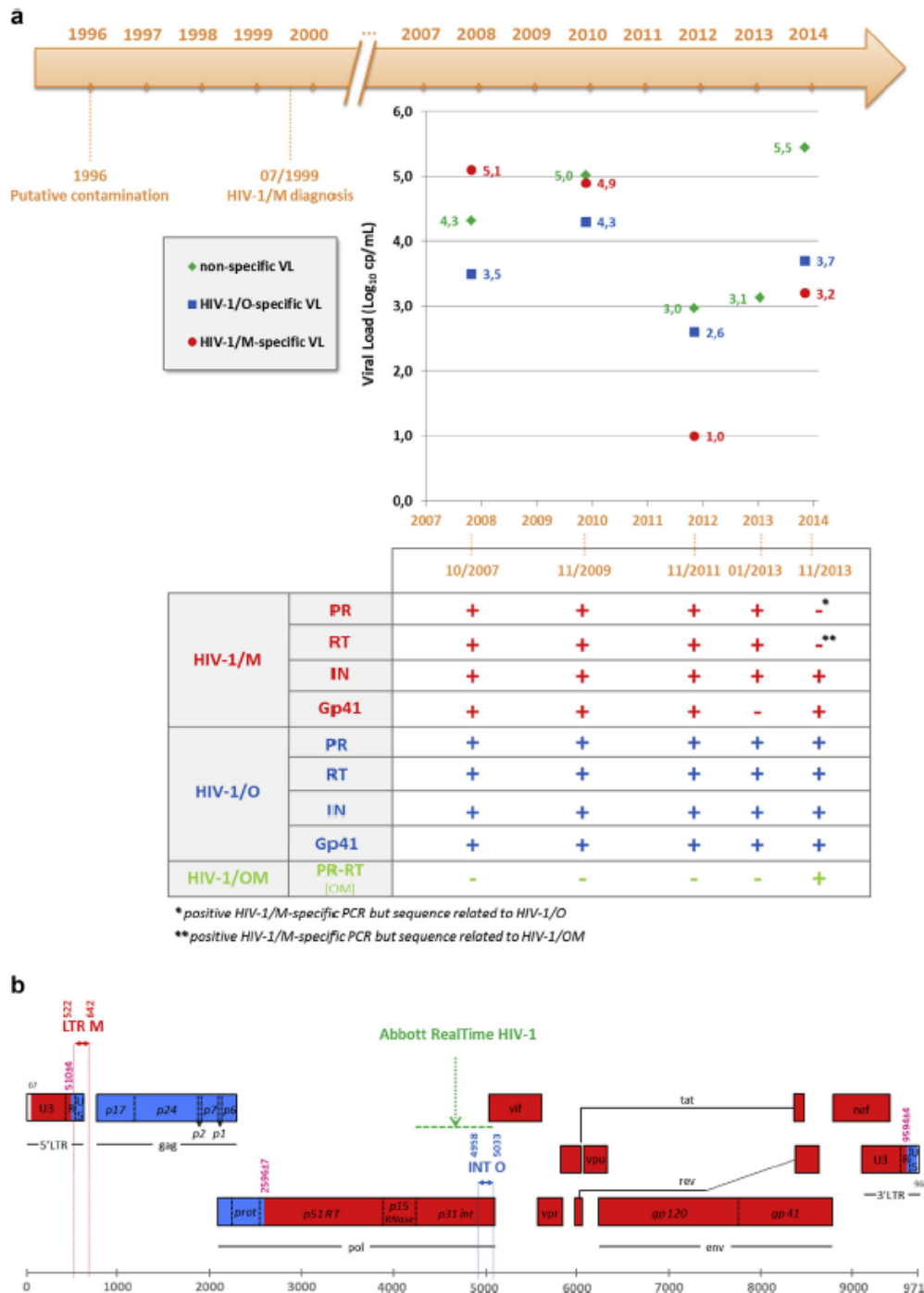
1198-743X/© 2019 European Society of Clinical Microbiology and Infectious Diseases. Published by Elsevier Ltd. All rights reserved.

Please cite this article as: Moisan A et al., *In vivo* emergence of an HIV-1/MO recombinant revealed undiagnosed HIV-1/M+O co-infection, Clinical Microbiology and Infection, <https://doi.org/10.1016/j.cmi.2019.08.016>

## ARTICLE IN PRESS

2

Letter to the Editor / Clinical Microbiology and Infection xxx (xxxx) xxx



**Fig. 1. a: Molecular results obtained for RBF235 samples between 2007 and 2013.** Due to its design, non-group specific viral load (VL) assay RealTime HIV-1 (Abbott, IL; in green) is able to quantify indifferently HIV-1/M, /O and /MO recombinant viruses. Due to the presence of a breakpoint in LTRs, the HIV-1/M-specific VL (Generic HIV-1 viral load assay, Biocentric, Bando, France; in red) only quantifies the M-parental strain. NRC INT O VL assay (in blue) only quantifies the O-parental strain. Limit of quantification of Abbott, Biocentric LTR M and NRC INT O VL assays were 1.6, 2.3 and 1.6  $\text{Log}_{10}$  cp/mL, respectively. Detected but not quantifiable VLs were set at 1.0  $\text{log}_{10}$  copies/mL. Results of group specific (RT)-PCRs are shown in the table. A result was considered positive (+) when the amplification followed by sequencing gave the expected sequence. A result was considered negative when group specific (RT)-PCR was not amplified (-) or when the sequence obtained after sequencing was not as expected (-\*\*). **b: Recombination pattern of RBF235 full-length genome.** Breakpoints in LTR and RT regions are indicated in pink relative to the numbering in HIV-1/M reference strain HxB2 (GenBank accession number K03455). The HIV-1/M and HIV-1/O phylogenetic origin of each genomic region is indicated in red and blue, respectively. The genome map was obtained using the Recombinant HIV Drawing Tool ([https://www.hiv.lanl.gov/content/sequence/DRAW\\_CRF/recom\\_mapper.html](https://www.hiv.lanl.gov/content/sequence/DRAW_CRF/recom_mapper.html)). The target region of each VL assay is indicated above the genome.

Please cite this article as: Moisan A et al., *In vivo* emergence of an HIV-1/MO recombinant revealed undiagnosed HIV-1/M+O co-infection, *Clinical Microbiology and Infection*, <https://doi.org/10.1016/j.cmi.2019.08.016>

## ARTICLE IN PRESS

Letter to the Editor / *Clinical Microbiology and Infection* xxx (xxxx) xxx

3

detection of virtually all strains. Paradoxically, their broad detection range regarding genetic diversity no longer leads to discrepant results, which until now suggested divergent variants or atypical infections. Currently, only resistance genotyping may still reveal discrepancies, as reported for other cases [2]. Here, the lack of specificity of some HIV-1/M-specific primers allowed the amplification of HIV-1/O in the PR-RT region and led us to explore the case; otherwise this patient would, in all likelihood, still be followed for an HIV-1/M infection only. Further exploration with specific tools and a population-targeted strategy showed periods of high replication of parental strains before 2013, which probably favoured the genesis of the recombinant strain. This form, as for most HIV-1/MO recombinant forms described to date [1–5], has a relatively simple pattern of recombination, with one breakpoint in the RT region and another one in the long terminal repeat regions.

In conclusion, cases of HIV-1/M+O super-infection or co-infections and generation of HIV-1/MO recombinants prove the dynamic and continuous evolution of HIV infections, which require increased vigilance when discrepancies are observed in the screening and/or the monitoring of patients. Due to the absence of specific screening, these co-infection and recombination phenomena could be underestimated. The epidemiological impact, physiopathology, impact on antiretroviral treatment strategy and immunovirological response of such atypical infections now need to be further investigated.

#### Transparency declaration

The authors have declared that there are no potential conflicts of interest.

#### Funding

Santé Publique France and Rouen University Hospital.

#### Acknowledgements

We thank Dr J.M. Chaplain, Dr P. Trémeaux and the staff of the virology laboratory of Rouen University Hospital, especially Dr V. Lemée, and are grateful to Nikki Sabourin-Gibbs, Rouen University Hospital, for her help in editing the manuscript.

#### Authors' contribution

AMo, FDO and JCP designed the experiments, analysed the data presented in the work and wrote the paper. AMo collected data, FDO performed experiments, PC contributed to the analysis, and CP and AMa performed the follow up of the HIV infection in Rennes and the first analyses leading to the description of HIV-1/M+O and HIV-1/MO co-infection. All authors have revised and approved the final manuscript.

#### References

- [1] De Oliveira F, Mourez T, Vessiere A, Ngoupo PA, Alessandri-Gradt E, Simon F, et al. Multiple HIV-1/M + HIV-1/O dual infections and new HIV-1/MO intergroup recombinant forms detected in Cameroon. *Retrovirology* 2017;14:1.
- [2] De Oliveira F, Cappy P, Lemee V, Moisan A, Pradier C, Bocket L, et al. Detection of numerous HIV-1/MO recombinants in France. *AIDS* 2018;32:1289–99.
- [3] Peeters M, Liegeois F, Torimiro N, Bourgeois A, Mpoudi E, Vergne L, et al. Characterization of a highly replicative intergroup M/O human immunodeficiency virus type 1 recombinant isolated from a Cameroonian patient. *J Virol* 1999;73:7368–75.
- [4] Takehisa J, Zekeng L, Ido E, Yamaguchi-Kabata Y, Mboudjeka I, Harada Y, et al. Human immunodeficiency virus type 1 intergroup (M/O) recombination in Cameroon. *J Virol* 1999;73:6810–20.
- [5] Yamaguchi J, Bodelle P, Vallan AS, Coffey R, McArthur CP, Schochetman G, et al. HIV infections in northwestern Cameroon: identification of HIV type 1 group O and dual HIV type 1 group M and group O infections. *AIDS Res Hum Retroviruses* 2004;20:944–57.