



**HAL**  
open science

# Recombinaison inter-groupes M et O du VIH-1 : Capacité répliquative, impact du profil de recombinaison et transmission mère-enfant.

Manon Vautrin

► **To cite this version:**

Manon Vautrin. Recombinaison inter-groupes M et O du VIH-1 : Capacité répliquative, impact du profil de recombinaison et transmission mère-enfant.. Médecine humaine et pathologie. Normandie Université, 2023. Français. NNT : 2023NORMR069 . tel-04438100

**HAL Id: tel-04438100**

**<https://theses.hal.science/tel-04438100>**

Submitted on 5 Feb 2024

**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 SANTE

Préparée au sein de l'Université de Rouen Normandie

**Recombinaison inter-groupes M et O du VIH-1 : Capacité répliquative, impact du profil de recombinaison et transmission mère-enfant.**

Présentée et soutenue par  
**MANON VAUTRIN, NÉ-E MABILLOTTE**

**Thèse soutenue le 27/09/2023  
devant le jury composé de**

M. PIERRE ROQUES	DIRECTEUR DE RECHERCHE, Commissariat à l'Energie Atomique	Rapporteur du jury
M. BENOIT VISSEAU	Docteur HDR, Laboratoire Cerba	Rapporteur du jury
M. FABRIZIO MAMMANO	DIRECTEUR DE RECHERCHE, UNIVERSITE DE TOURS	Membre du jury
M. ANNE -GENEVIÈVE MARCELIN	PROFESSEUR DES UNIV - PRATICIEN HOSP., SORBONNE UNIVERSITE	Membre du jury
MME ALICE MOISAN	PRATICIEN HOSPITALIER, CHU de Rouen	Membre du jury
MME ALESSIA ZAMBORLINI	PROFESSEUR DES UNIVERSITES, Université Paris-Saclay	Membre du jury
M. JEAN-CHRISTOPHE PLANTIER	PROFESSEUR DES UNIV - PRATICIEN HOSP., Université de Rouen Normandie	Directeur de thèse

**Thèse dirigée par JEAN-CHRISTOPHE PLANTIER (Dynamique Microbienne associée aux Infections Urinaires et Respiratoires)**



Normande de Biologie Intégrative,  
Santé, Environnement





## Remerciements

J'adresse mes remerciements au Pr. Jean-Christophe Plantier, pour m'avoir accueilli au sein du laboratoire DYNAMICURE, tout d'abord lors d'un stage de master 2 en 2019, puis pour m'avoir permis de continuer au cours de ce travail de doctorat. Merci pour votre confiance, pour avoir partagé votre expérience scientifique et pour vos conseils.

Je voudrais également remercier les Dr. Pierre Roques et Benoît Visseaux, pour avoir accepté d'évaluer mon travail, d'avoir participé à sa correction et pour votre intérêt sur ce sujet. Merci également au Pr. Alessia Zamborlini, d'avoir suivi ce projet au travers du CSI et d'avoir accepté d'en examiner le contenu final. Merci au Dr. Fabrizio Mammano, d'avoir accepté d'évaluer ce travail en tant qu'examineur. Enfin, merci au Pr. Anne-Geneviève Marcelin d'avoir accepté de présider le jury de ce doctorat.

Un immense merci au Dr. Alice Moisan, pour toutes nos conversations autour de la recombinaison et de la diversité génétique des VIH, pour ta pédagogie à l'oral, comme à l'écrit. Merci pour tous les schémas que tu as pu faire et que je vais garder précieusement, ainsi que tes présentations toujours très claires. Merci pour tes conseils, ton soutien et surtout, pour l'optimisme sans failles dont tu as fait preuve tout au long de ce travail.

Merci à toute l'équipe du laboratoire de virologie du CHU de Rouen, qui malgré le contexte et la fatigue, a trouvé le temps de m'aider lorsque j'étais en difficulté, ou lorsque j'avais besoin de conseils. A toi Fanny en particulier, pour ta bienveillance et ton expertise, merci également de m'avoir montré que le chocolat est un précieux allié au quotidien. Merci au Dr. Elodie Alessandri-Gradt, pour ta gentillesse, pour m'avoir fait profiter de ton expérience du L3 et plus généralement, pour tes conseils en culture cellulaire. Merci à toi Fabienne, pour ta bonne humeur, pour tous nos bons moments à parler de mise au point de PCR, de séquençage et de diversité génétique. Merci pour ta réactivité et ta disponibilité depuis mon stage de master 2 au laboratoire. Enfin, merci Sandy pour ton aide précieuse depuis mon arrivée au laboratoire et pour ta bonne humeur au quotidien.

Je n'oublie pas mes collègues de bureau qui se sont succédé au fil des années : Romain, Chloé, Ségolène, Charlène et Lucie. Merci à toi Maria, pour ta douceur et ton sourire, tu es un véritable rayon de soleil. Charlène, merci infiniment pour tous tes conseils, ton organisation et ta grande rigueur scientifique. Ce projet n'aurait jamais autant avancé sans tes travaux de master et avec l'énergie que tu as déployé pour m'aider dans la génération des clones. Vous tous, avez réussi à me motiver lors des moments difficiles et je n'oublierai jamais vos encouragements.

Un grand merci à toute l'équipe du laboratoire de recherche DYNAMICURE, les anciens comme les nouveaux, Sandrine, Vincent, Camille, Laurie, Chervin, Alice, Caroline et Charles. Merci à toi Stéphanie, pour tous tes conseils et pour m'avoir aidé dans la préparation des nombreux milieux de culture. Marie, je te remercie pour ta bonne humeur et ton expertise lors de ces travaux, notamment pour la partie NGS. Merci à tous, pour votre aide et vos conseils, mais également pour tous nos moments passés ensemble à élaborer les calendriers du laboratoire, pour nos discussions légendaires, nos déjeunés et goûtés au top. Une pensée particulière pour Maxime, qui plus qu'un collègue, est devenu un ami et un précieux soutien au quotidien.

À mes parents, Christophe et Isabelle, je suis fière d'être votre fille et je sais que je ne serai jamais arrivée là aujourd'hui sans vous, sans votre amour et vos encouragements. À vous tous, Loïc, Marie-Hélène, Catherine, Morgane, Mathéo et Rafael. À ma belle-famille : Stéphane, Isabelle, David et Marion. Merci pour tous ces bons moments passés ensemble, pour toute cette joie que vous m'avez communiquée.

Enfin, merci à toi Nicolas, ma moitié et collègue doctorant, tu as été mon modèle durant toute cette thèse. Merci pour ton soutien, pour avoir cru en moi et pour tout le bonheur que tu me donnes au quotidien. Je n'oublie pas notre petit Kiwi qui nous a beaucoup aidés avec sa « ronron thérapie ».

*“La théorie, c'est quand on sait tout et que rien ne fonctionne. La pratique, c'est quand tout fonctionne et que personne ne sait pourquoi. Ici, nous avons réuni théorie et pratique : Rien ne fonctionne... et personne ne sait pourquoi !”*

*Albert Einstein*

## **Recombinaison inter-groupes M et O du VIH-1 : Capacité répliquative, impact du profil de recombinaison et transmission mère-enfant**

La forte diversité génétique des VIH-1 et la co-circulation des VIH-1 de groupes M et O au Cameroun a favorisé le phénomène de recombinaison génétique. Ce phénomène de recombinaison est possible entre ces groupes, malgré leur divergence génétique importante. Depuis 1999, 19 formes recombinantes uniques (URF\_MO) ont été décrites. Ces formes sont rares et leurs propriétés virologiques, ainsi que leur impact clinique sont mal compris, en raison de leur profil génétique atypique.

Ces travaux avaient pour objectif, d'une part, d'étudier le potentiel répliquatif et l'impact du profil de recombinaison de ces formes recombinantes VIH-1/MO et d'autre part, de caractériser une nouvelle URF-MO dans un contexte rare de transmission mère-enfant.

Premièrement, des essais de génération de clones moléculaires infectieux (CMI) parentaux VIH-1/M et VIH-1/O et chimériques (CMIC) VIH-1/MO ont été réalisés afin de comparer le potentiel répliquatif des virus parentaux et recombinants correspondants entre eux. Plusieurs essais ont été réalisés au laboratoire et par une société extérieure. Les virus correspondants ont été produits par génétique inverse. Un nouveau CMI VIH-1/M, appelé pM-A-062021 de sous-type A, a pu être produit au cours de ce projet. Aucune autre construction n'a pu aboutir à un clone stable et/ou infectieux après transfection puis culture cellulaire *in vitro*.

Deuxièmement, nous avons comparé les cinétiques de réplication de deux souches recombinantes VIH-1/MO, OM-TB-2016 et RBF208, présentant une diversité génétique et un profil de recombinaison différents. Quatre cinétiques de réplication ont été réalisées après mise en culture *in vitro* de chaque souche. Le suivi de la réplication a été effectué à l'aide de deux marqueurs, la mesure de la quantité d'antigène P24 et de l'activité de la transcriptase inverse, dans chaque surnageant de culture. Les résultats ont montré qu'il n'y avait pas de différence de potentiel répliquatif entre ces deux souches mais, une plus forte mortalité cellulaire a été observée pour la souche OM-TB-2016.

Enfin, ce travail a permis de caractériser une nouvelle forme recombinante VIH-1/MO, au cours de l'année 2022, dans un contexte rare de transmission verticale. Les génomes complets de la mère et de son enfant ont été caractérisés par séquençage Sanger et par le développement d'une technique de séquençage à haut-débit, appliquée à ces formes VIH-1/MO. Notre étude a pu prouver que ces formes recombinantes sont transmissibles de la mère-enfant et qu'elles ne sont pas atténuées *in vivo* en absence de traitement. Deux points de cassure ont été identifiés, comme pour la majorité des formes recombinantes VIH-1/MO déjà identifiées. L'étude phylogénétique a toutefois montré l'absence de lien avec les autres formes recombinantes déjà décrites, portant à 20 le nombre d'URF\_MO décrites *in vivo*.

Ces travaux ont permis de conclure que ces formes possèdent un potentiel répliquatif non atténué *in vitro* et *in vivo*. La faible proportion de ces formes au niveau mondial indique une diffusion limitée, mais les données de ces travaux indiquent qu'elles possèdent néanmoins un potentiel pour émerger et doivent être surveillées.

**Mots-clés** : Recombinaison inter-groupes VIH-1/MO - Potentiel répliquatif – Transmission mère-enfant

## **HIV-1 Intergroup M and O recombination: Replicative capacity, impact of recombination profile and mother-to-child transmission**

The high genetic diversity of HIV-1 and the co-circulation of HIV-1 groups M and O in Cameroon has favored the phenomenon of genetic recombination. Recombination is possible between these groups, despite their significant genetic divergence. Since 1999, 19 unique recombinant forms (URF\_MO) have been described. These forms are rare, their virological properties and clinical impact are poorly understood, due to their atypical genetic profile.

The aim of this work was to study the replicative potential and impact of the recombination profile of these recombinant HIV-1/MO forms, and to characterize a new URF-MO in a rare context of mother-to-child transmission.

Firstly, parental HIV-1/M and HIV-1/O and chimeric HIV-1/MO molecular infectious clone (MIC) generation assays were carried out to compare the replicative potential of the corresponding parental and recombinant viruses. Several tests were carried out in the laboratory and by an external company. The corresponding viruses were produced by reverse genetics. A new HIV-1/M MIC, called pM-A-062021 subtype A, was produced during this project. No other construct was able to produce a stable and/or infectious clone after transfection and subsequent *in vitro* cell culture.

Secondly, we compared the replication kinetics of two recombinant HIV-1/MO strains, OM-TB-2016 and RBF208, with different genetic diversity and recombination profiles. Four replication kinetics were performed after *in vitro* culture of each strain. Replication was monitored using two markers, measurement of the amount of P24 antigen and reverse transcriptase activity, in each culture supernatant. The results showed that there was no difference in replicative potential between these two strains, but higher cell death was observed for strain OM-TB-2016.

Finally, a new recombinant HIV-1/MO form was characterized in a rare context of vertical transmission in 2022. The complete genomes of the mother and her child were characterized, by Sanger sequencing and the development of a high-throughput sequencing technique applied to these HIV-1/MO forms. Our study demonstrated that these recombinant forms are transmissible from mother to child, and that they are not attenuated *in vivo* in the absence of treatment. Two breakpoints were identified, as for the majority of recombinant HIV-1/MO forms already identified. A phylogenetic study showed, however, that there was no link with other recombinant forms already described, bringing to 20 the number of URF\_MOs described *in vivo*.

This work led to the conclusion that these forms have an unattenuated replicative potential *in vitro* and *in vivo*. The low proportion of these forms worldwide indicates their limited distribution, but the data from this work indicates that they nevertheless have the potential to emerge and should be monitored.

**Keywords:** HIV-1/MO inter-group recombination - Replicative potential - Mother-to-child transmission



## Liste des abréviations / sigles

ADN	Acide désoxyribonucléique	
ADNc	ADN complémentaire	
AgP24	Antigène p24	
Amp	Ampicilline	
ARN	Acide ribonucléique	
ARNg	ARN génomique	
ARNm	ARN messenger	
ARNt	ARN de transfert	
ARV	Antirétroviral	
CA	Capside	
CCR5	Récepteur de chimiokine de type 5	<i>Chemokine receptor type 5</i>
CDC	Centre de contrôle des maladies	<i>Center of Disease Control</i>
CFAR	Centre de réactifs pour le SIDA	<i>Center For AIDS Reagents</i>
CHU	Centre Hospitalier Universitaire	
CMI	Clone moléculaire infectieux	
CMIC	Clone moléculaire infectieux chimérique	
CMSPs	Cellules mononuclées du sang périphérique	
CNR	Centre National de Référence	
CPC	Centre Pasteur du Cameroun	
CRF	Formes recombinantes circulantes	<i>Circulating Recombinant Form</i>
CV	Charge virale	
CXCR4	Récepteur de chimiokine de type 4	<i>C-X-C chemokine receptor type 4</i>
DEAE	Diethylethanolamine	
DMEM	<i>Dulbecco's minimal essential medium</i>	
dNTP	Désoxyribonucléotide	Désoxy-N tri-phosphate
ddNTP	Didésoxyribonucléotide	Didésoxy-N tri-phosphate
EFS	Etablissement Français du Sang	
ELISA		<i>Enzyme Linked ImmunoSorbent Assay</i>
GA	Gibson assembly	
GC	Genecust	
GP41	Glycoprotéine d'enveloppe transmembranaire	
HTLV	Virus T-Lymphotropique humain	<i>Human T-lymphotropic virus</i>
ILr	Interleukine II recombinante	
IN	Intégrase	
INNTI	Inhibiteurs non nucléosidiques de la transcriptase inverse	
LAV	Virus associé aux lymphadénopathies	<i>Lymphadenopathy Associated Virus</i>
LB		<i>Lysogeny Broth</i>
LTR		<i>Long Terminal Repeat</i>
MA	Matrice	
MCS	Site de clonage multiple	<i>Multiple cloning site</i>
MOI		<i>Multiplicity of Infection</i>
MRCA	Ancêtre commun le plus récent	<i>Most Recent Common Ancestor</i>
NC	Nucléocapside	
NGS	Séquençage de nouvelle génération	<i>Next Generation sequencing</i>
NIBSC	Institut national des normes et du contrôle biologiques	<i>National Institute for Biological Standards and Control</i>
OMS	Organisation Mondiale de la Santé	<i>World Health Organization</i>
ORF	Cadres de lectures ouverts	<i>Open Reading Frame</i>
pb	Paires de bases	<i>Base pair</i>
PBS	Site de liaison au promoteur	<i>Promotor Binding Site</i>
PCR	Réaction de polymérisation en chaîne	<i>Polymerase Chain Reaction</i>
PPT		<i>Polypurin tract</i>
PR	Protéase	
PvVIH	Personnes vivant avec le VIH	
RES-O	Réseau de surveillance des VIH-1/O en France	
RPMI		<i>Roswell Park Memorial Institute</i>
RLU	Unité de luminescence relative	<i>Relative Light Units</i>
RT		<i>Rétrotranscriptase</i>
RT-PCR		<i>Reverse Transcriptase-PCR</i>

SIDA	Syndrome de l'Immunodéficience Acquis	
SVF	Sérum de veau foetal	
TCID <sub>50</sub>	Dose infectant 50 % des cellules de culture	<i>Tissue Culture Infectious Dose 50</i>
TI	Transcriptase inverse	
TME	Transmission mère-enfant	
URF	Formes recombinantes uniques	<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	

# **ANALYSE BIBLIOGRAPHIQUE**

<b>I.</b>	<b>Généralités sur le Virus de l'Immunodéficience Humaine .....</b>	<b>15</b>
1.	Découverte du VIH .....	15
2.	Données épidémiologiques et modes de transmission du VIH.....	16
3.	Structure et organisation .....	17
a.	Taxonomie .....	17
b.	Particule virale.....	17
c.	Organisation génomique.....	19
4.	Cycle de réplication .....	22
<b>II.</b>	<b>Origines et diversité génétique des VIH .....</b>	<b>26</b>
1.	Origines simiennes .....	26
2.	Adaptation à l'Homme et diffusion .....	28
3.	Diversité génétique .....	29
<b>III.</b>	<b>Classification des VIH .....</b>	<b>30</b>
1.	Les VIH de type 1 (VIH-1).....	30
a.	Les VIH-1 de groupe M (VIH-1/M).....	30
b.	Les VIH-1 de groupe O (VIH-1/O) .....	32
c.	Les VIH-1 de groupe N et P (VIH-1/N et VIH-1/P).....	34
2.	Les VIH de type 2 (VIH-2).....	35
<b>IV.</b>	<b>Recombinaison inter-groupes VIH-1/MO.....</b>	<b>35</b>
1.	La recombinaison génétique .....	35
a.	Prérequis de la recombinaison.....	35
b.	Mécanisme de la recombinaison.....	37
c.	Localisation et fréquence des points de cassure.....	39
2.	Circonstances de découverte des formes recombinantes VIH-1/MO.....	42
3.	Caractéristiques moléculaires des VIH-1/MO .....	49
4.	Conséquences de la recombinaison génétique.....	52
5.	Capacité répliquative et fitness viral.....	54

# **RATIONNEL ET OBJECTIFS**

## **PARTIE EXPERIMENTALE**

<b>I.</b>	<b>Potentiel répliatif et impact du profil de recombinaison des VIH-1/MO</b>	<b>63</b>
1.	Génération <i>in vitro</i> des virus VIH-1/M, VIH-1/O et VIH-1/MO.....	65
a.	Méthodologie.....	65
	Production du stock de surnageants pour chaque virus.....	78
b.	Résultats.....	80
c.	Discussion.....	85
2.	Étude <i>in vitro</i> du potentiel répliatif de deux formes VIH-1/MO.....	89
a.	Méthodologie.....	89
b.	Résultats.....	97
c.	Discussion.....	109
<b>II.</b>	<b>Transmission verticale d'une forme recombinante VIH-1/MO.....</b>	<b>115</b>
1.	Méthodologie.....	116
2.	Résultats.....	125
3.	Discussion.....	137

**Conclusion / perspectives**

**Références bibliographiques**

**Annexes**

## Liste des figures

Figure 1 : Estimation du nombre de personnes vivants avec le VIH dans le monde en 2021 .....	16
Figure 2 : Structure de la particule virale de VIH-1 .....	18
Figure 3 : Organisation du génome du VIH-1 .....	19
Figure 4 : Etape du cycle de réplication du VIH-1 .....	22
Figure 5 : Transcription inverse du génome de l'ARN rétroviral du VIH .....	24
Figure 6 : Relations phylogénétiques et diversité des VIH-1, VIH-2 et SIV.....	27
Figure 7 : Schéma de la forme recombinante CRF02_AG .....	31
Figure 8 : Schéma de la forme recombinante CRF06_cpx.....	31
Figure 9 : Pourcentages de divergences génétiques entre les gènes de structure des VIH-1/M et VIH-1/O.....	33
Figure 10 : Arbre phylogénétique des VIH-1 de groupes O (a) et M (b) .....	34
Figure 11 : Description des 20 doubles infections VIH-1/M+O .....	36
Figure 12 : Mécanisme de la recombinaison génétique .....	37
Figure 13 : Choix de copie ou Assimilation de branches .....	39
Figure 14 : Répartition des points de cassure au sein du VIH-1.....	40
Figure 15 : Schéma de la forme recombinante 97CA.MP645MO .....	43
Figure 16 : Schéma de la forme recombinante 95CM61.56.....	43
Figure 17 : Schéma de la forme recombinante DSC1320.....	44
Figure 18 : Algorithme de dépistage des formes recombinantes vpr .....	47
Figure 19 : Profil et diversité génétique des 19 URF_MO .....	50
Figure 20 : Diversité génétiques de fragments parentaux identifiés au sein des 19 URF_MO.....	51
Figure 21 : Localisation des points de cassures issus des 19 URF_MO .....	52
Figure 22 : Conséquences évolutives de la recombinaison génétique .....	53
Figure 23 : Méthodes de détermination du fitness viral.....	55
Figure 24 : Processus de génération et sélection des VIH-1/MO.....	56
Figure 25 : Principe général de la réaction Gibson Assembly .....	66
Figure 26 : Profil des CMI VIH-1/M et VIH-1/O .....	67
Figure 27 : Méthode Gibson Assembly avec 2 fragments pour la génération des CMI parentaux .....	69
Figure 28 : Etapes de la transformation chimique des produits Gibson Assembly.....	70
Figure 29 : Vue d'ensemble des différentes étapes de production des CMI et CMIC .....	71
Figure 30 : Profil de recombinaison des deux formes VIH-1 « OMOM » et « MOMO ».....	72
Figure 31 : Profil des CMIC VIH-1/MOMO et OMOM.....	73
Figure 32 : Méthode Gibson Assembly pour la génération des CMIC avec 4 fragments.....	75
Figure 33 : Schéma expérimental de production des virus par transfection puis co-culture .....	77
Figure 34 : Schéma expérimental de production des stocks viraux par infection de CMSPs .....	78
Figure 35 : Graphique montrant la quantité d'AgP24 au cours du temps après transfection/co-culture par les CMI.....	80
Figure 36 : Quantité d'AgP24 obtenue au cours du temps après transfection co-culture de pBCF12083	
Figure 37 : Schéma de la forme recombinante RBF208 .....	89
Figure 38 : Schéma de la forme recombinante OM-TB-2016.....	90
Figure 39 : Protocole expérimental de lancement des cinétiques de réplication pour RBF208 et OM-TB-2016 après infection de CMSPs.....	93
Figure 40 : Principe du SG-PERT .....	94
Figure 41 : Paramètres de comparaison entre les souches OM-TB-2016 et RBF208.....	96

Figure 42 : Cinétiques de réplication obtenues pour les souches OM-TB-2016 et 89.6 à MOI $10^{-2}$ et $10^{-4}$ .....	98
Figure 43 : Cinétiques de réplication de la souche OM-TB-2016 .....	102
Figure 44 : Cinétiques de réplication de la souche RBF208 .....	105
Figure 45 : Comparaison des cinétiques moyenne de réplication des souches OM-TB-2016 vs RBF208 .....	108
Figure 46 : Discordance moléculaire du génotypage de résistance effectué au CHU de Lyon pour RBF255.....	116
Figure 47 : Etapes de séquençage des génomes complets par séquençage nouvelle génération et Sanger .....	117
Figure 48 : Position des amorces utilisées pour l'amplification des génomes complets de RBF254 et RBF255.....	118
Figure 49 : Principe de la tagmentation .....	121
Figure 50 : Principe du séquençage par la technologie Illumina.....	122
Figure 51 : Suivi de la CV et du nombre de CD4 pour les échantillons de RBF254 et RBF255 .....	127
Figure 52 : Discordance moléculaire observée pour RBF254 et RBF255 après investigation par le CNR de Rouen .....	128
Figure 53 : Histogramme présentant la qualité des reads obtenus après séquençage NGS .....	130
Figure 54 : Différences de paires de base identifiées entre les séquences Sanger et NGS pour RBF254 et RBF255 .....	131
Figure 55 : Arbres phylogénétiques des régions PR-TI (A) et IN (B) de RBF254 et RBF255.....	133
Figure 56 : Simplot de la séquence du génome complet de RBF254 .....	134
Figure 57 : Génomes complets de RBF254 et RBF255 .....	135
Figure 58 : Arbres phylogénétiques des régions PR-TI (A) et IN (B) de RBF254, RBF255 et des URF_MO .....	136

## **Liste des tableaux**

Tableau 1 : Description des gènes du VIH-1 et de leurs propriétés .....	21
Tableau 2 : Découverte des 23 formes recombinantes VIH-1/MO .....	48
Tableau 3 : Répartition des points de cassure entre les gènes gag et nef du génome des 19 URF_MO .....	51
Tableau 4 : Bilan de la génération des CMI .....	82
Tableau 5 : Caractéristiques des deux CMI parentaux VIH-1/M et VIH-1/O .....	84
Tableau 6 : Caractéristiques des deux formes à comparer OM-TB-2016 et RBF208 .....	97
Tableau 7 : Résultats du comptage cellulaire en fin de culture pour chaque cinétique .....	99
Tableau 8 : Résultats des Innoteests et SG-PERT pour chaque cinétique de la souche OM-TB-2016 ..	101
Tableau 9 : Résultats des Innoteests et SG-PERT pour chaque cinétique de la souche RBF208 .....	104
Tableau 10 : Comparaison des potentiels de réplication de OM-TB-2016 et RBF208 .....	107
Tableau 11 : Amorces utilisées pour l'amplification des génomes complets de RBF254 et RBF255 ..	119
Tableau 12 : Amorces utilisées pour le séquençage Sanger des génomes complets de RBF254 et RBF255 .....	120
Tableau 13 : Données viro-cliniques et thérapeutiques des échantillons pour RBF254 et RBF255 ...	126
Tableau 14 : Résultats de séquençage obtenus pour RBF254 et RBF255 par technologie Illumina ...	130
Tableau 15 : Comparaison des séquences consensus de RBF254 et RBF255 obtenues par séquençage NGS et Sanger .....	132
Tableau 16 : Matrice de distances entre les séquences de RBF254 et RBF255 obtenues par séquençage Sanger et NGS .....	132

## **Liste des annexes**

Annexe 1 : Amorces ayant servi à la génération des CMI .....	159
Annexe 2 : Protocole de transformation chimique des bactéries NEB Stable .....	160
Annexe 3 : Amorces ayant servi à la génération des CMIC .....	161
Annexe 4 : Protocole de génération des fragments de CMIC .....	162
Annexe 5 : Protocole de génétique inverse (transfection / co-culture) .....	162
Annexe 6 : Protocole de séquençage Sanger .....	163
Annexe 7 : Protocole de SG-PERT du laboratoire de virologie du CHU de Rouen .....	164
Annexe 8 : Case report "First evidence of a mother to child transmission of an HIV-1/MO intergroup recombinant form" .....	165

# ANALYSE BIBLIOGRAPHIQUE

## I. Généralités sur le Virus de l'Immunodéficience Humaine

### 1. Découverte du VIH

En 1981, le Centre de Contrôle et de Prévention des Maladies aux Etats-Unis (*CDC Center of Disease Control*), alerte sur une épidémie en cours de pneumonies ou sarcomes de Kaposi, chez des personnes très immunodéprimées (1). Cette épidémie touche alors surtout de jeunes homosexuels, mais également des usagers de drogues intraveineuses. Le nom de *Gay-Related Immune Deficiency*, soit littéralement « Déficit immunitaire lié à l'homosexualité », est adopté aux Etats Unis pour décrire cette maladie. Très stigmatisant et inexact, ce nom est rapidement abandonné au profit de Syndrome de l'Immunodéficience Acquise, ou SIDA. L'épidémie prend une dimension mondiale et de nombreux chercheurs tentent de comprendre l'infection.

En 1983, l'agent étiologique du SIDA est isolé en France à l'Institut Pasteur par l'équipe de Françoise Barré-Sinoussi et Luc Montagné et baptisé à l'occasion LAV pour *Lymphadenopathy Associated Virus* (2). La découverte de ce virus est source de controverse aux Etats-Unis puisqu'une autre équipe ayant collaboré avec l'Institut Pasteur, celle de *Gallo et al*, s'attribue la découverte du virus en 1984 (3, 4). Ils nomment leur virus HTLV-III pour *Human T-lymphotropic virus*, en lien avec le HTLV-I, premier rétrovirus oncogène humain découvert quelques années plus tôt par cette même équipe (5). En 1985, le séquençage montre que le LAV et le HTLV-III sont des virus identiques (6). La communauté scientifique décide donc d'adopter une dénomination commune en 1986 : VIH pour Virus de l'Immunodéficience Humaine. Cette même année, la découverte d'un autre virus proche conduit à la différenciation du VIH en deux types, le VIH de type 1 (VIH-1) et le VIH de type 2 (VIH-2) (7, 8). En 2008, le Prix Nobel de médecine est attribué à l'équipe française de l'Institut Pasteur.

Depuis, la grande diversité de ce virus a conduit à la description de quatre groupes de VIH-1 : le VIH-1 de groupes M (VIH-1/M), identifié en 1983 (2), O (VIH-1/O), identifié en 1990 (9), N (VIH-1/N), identifié en 1998 (10) et P (VIH-1/P), identifié en 2009 (11). Quant au VIH-2, neuf groupes nommés de A à I ont été identifiés à ce jour (12).



## 2. Données épidémiologiques et modes de transmission du VIH

Depuis le début de l'épidémie, on dénombre au total 40,1 millions de décès liés au SIDA (13). Les derniers chiffres publiés par l'Organisation Mondiale de la Santé (OMS) indiquent environ 1,5 millions nouvelles infections et 38,4 millions de personnes vivant avec le VIH (PvVIH) en 2021 (Figure 1) (14). Parmi ces 38,4 millions de personnes, 25,6 millions vivent en Afrique sub-saharienne, soit environ 67% de l'ensemble des PvVIH (Figure 1). Enfin, sur les 1,7 millions d'enfants âgés de 0 à 14 ans contaminés dans le monde, 1,42 millions soit 83,5%, sont des enfants vivant en Afrique sub-saharienne (15).

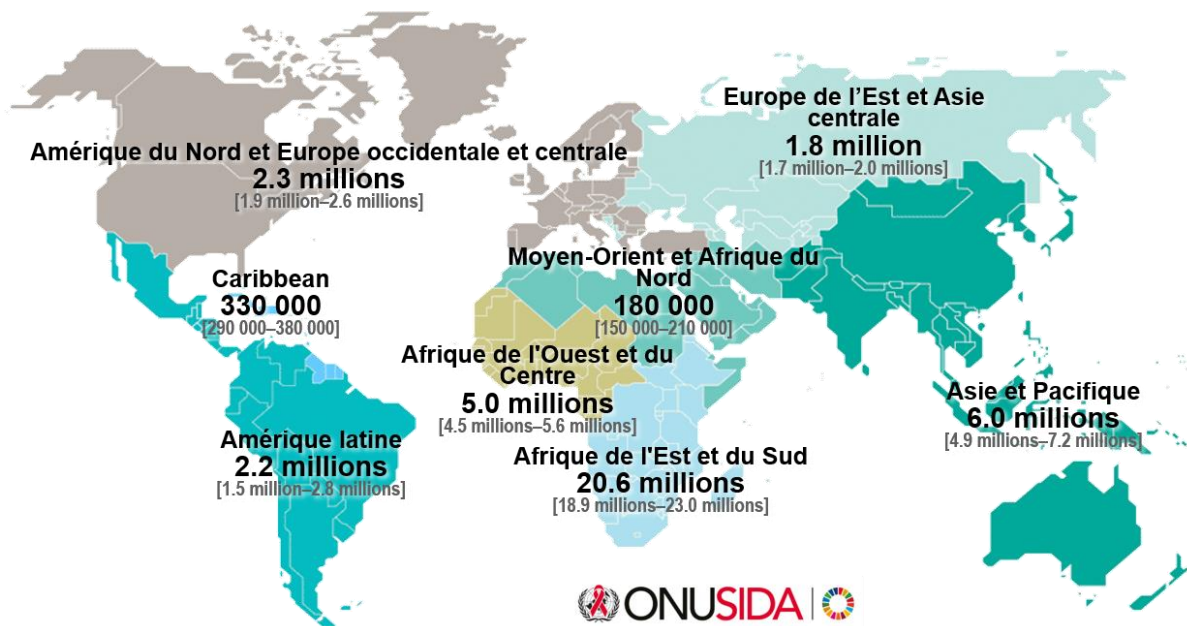


Figure 1 : Estimation du nombre de personnes vivants avec le VIH dans le monde en 2021

Diapositives clés sur les données épidémiologiques du VIH, Chiffres Onusida 2021, publiés le 27 juillet 2022 (15).

En France, les chiffres de 2021 indiquent qu'environ 29% des infections VIH ont été identifiées à un stade avancé de la maladie (16). La plupart des personnes infectées sont asymptomatiques ou présentent un tableau clinique pseudogrippal qui n'évoque pas la contamination. La découverte de l'infection tardive ne permet pas toujours de bien prendre en charge les patients et augmente le risque de transmission, donc de diffusion de l'épidémie, avant mise sous traitement (17).

Il existe trois modes de transmission du VIH : sexuelle, sanguine et enfin, la transmission « mère-enfant » (TME), aussi appelée transmission verticale (18). La transmission sexuelle est le principal mode de transmission du VIH au sein de la population. La transmission sanguine, quant à elle, s'effectue le plus souvent par le biais de l'utilisation d'aiguilles et de seringues partagées, contaminées par du sang infecté. Enfin, le taux de TME est hétérogène au niveau mondial, allant de 6 à 24% selon les chiffres de 2021 (15). En Afrique sub-saharienne, ce taux varie de 9% pour l'Afrique de l'Est/Sud à 21% pour l'Afrique de l'Ouest/Centrale (15). En France, ce taux est passé à environ 0,3% entre 2010 – 2015, contre 15 à 20% en absence de traitement sur cette même période, selon l'enquête périnatale Française (19, 20). Ce constat est lié au fait que l'ensemble des pays ne disposent de façon équivalente des traitements et des moyens de prévention, ou de diagnostic. En France, le taux de TME est très faible car la majorité des femmes enceintes vivant avec le VIH bénéficient d'un accès à un traitement avant ou au cours de la grossesse, afin de limiter le risque de contamination par le contrôle de la charge virale (CV) (21). Plus la CV est élevée et plus le risque de transmettre le virus est important (22). Le mode de contamination par transmission verticale reste toutefois préoccupant, car il est la source de plus de 200 000 nouvelles contaminations en Afrique sub-saharienne pour l'année 2021 (15). Cette transmission verticale du VIH, fera l'objet d'un chapitre dédié page 112 (**II. Transmission verticale d'une forme recombinante VIH-1/MO**).

### **3. Structure et organisation**

#### **a. Taxonomie**

Le VIH appartient à la famille des *Retroviridae* et à la sous-famille des *Orthoretrovirinae*. Comme tous les rétrovirus, le VIH possède une enzyme particulière appelée la transcriptase inverse (TI), ou encore RT pour *reverse transcriptase*. Le VIH appartient au genre *Lentivirus*, puisqu'il est caractérisé par une longue période d'infection et de latence avant d'induire un SIDA.

#### **b. Particule virale**

La particule virale du VIH mesure entre 90 et 120 nanomètres de diamètre, est sphérique et composée de trois éléments structuraux différents : le génome, la capsid et l'enveloppe (Figure 2).

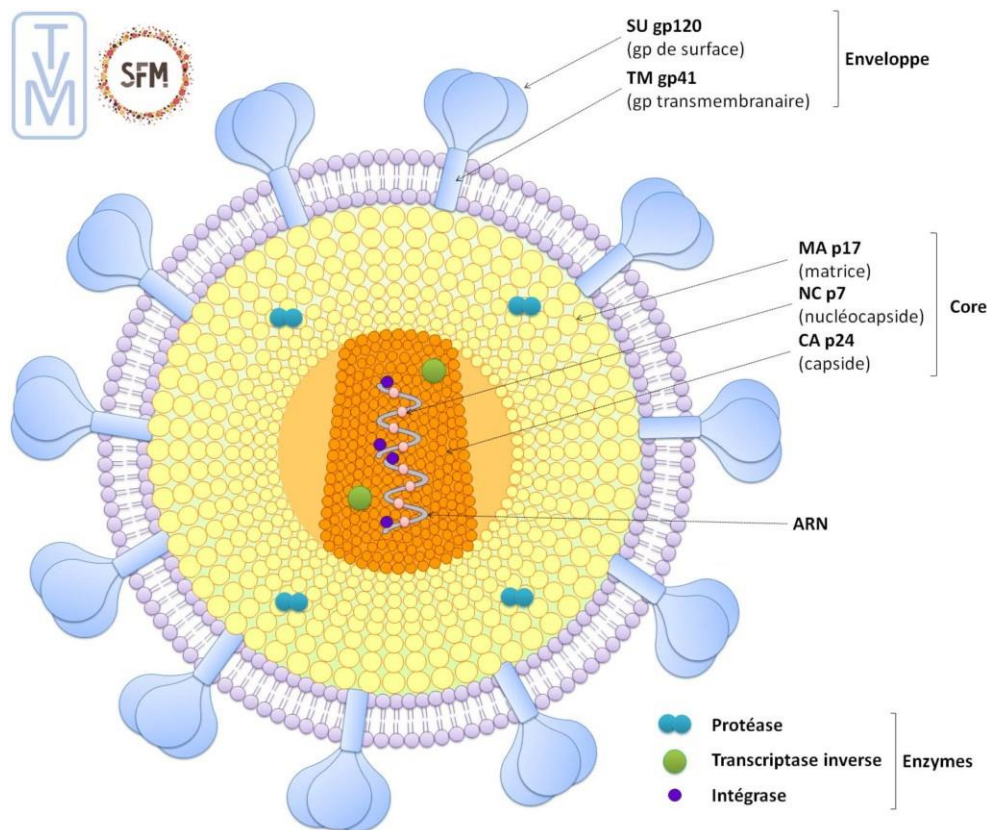


Figure 2 : Structure de la particule virale de VIH-1

Image issue du Traité de virologie médicale, édition 2019 (23).

En violet figure la partie externe de la particule virale avec l'enveloppe et ses constituants. La partie jaune montre la matrice tandis qu'en orange figurent les éléments les plus internes, liés au génome du VIH en gris, associé à la protéine p7 (les deux formant la nucléocapside). Les différentes enzymes clés du cycle réplcatif sont représentés par des ronds bleus, verts ou violets.

Le génome du VIH est composé de deux molécules d'ARN (Acide ribonucléique) associées avec la protéine de nucléocapside NC p7 pour être protégées.

La capsid du VIH est en forme de cône tronqué et résulte de l'assemblage de la protéine CA p24. Elle protège à la fois la nucléocapside et deux des enzymes clés de la réplcation du VIH : la TI p66/p51 ainsi que l'intégrase IN p32. La TI est à l'origine du mécanisme de transcription inverse, ou rétro-transcription, qui consiste à convertir l'ARN du VIH en ADN complémentaire (ADNc). Cet ADN est ensuite intégré dans le génome de la cellule infectée par l'IN par le mécanisme d'intégration.

L'enveloppe du virus est composée d'une bicouche phospholipidique résultant du bourgeonnement du virus à la surface de la membrane plasmique de la cellule hôte. La glycoprotéine gp160 est un précurseur clivé par une protéase cellulaire, donnant d'une part, la glycoprotéine de surface SU gp120 et d'autre part, la glycoprotéine transmembranaire TM gp41.

Il existe un élément de liaison entre l'enveloppe et la capsid, appelé la matrice, qui résulte de l'assemblage de la protéine MA p17. Cette partie de la particule virale est associée à la troisième enzyme clé du VIH, la protéase PR p15 qui participe à la maturation des nouveaux virions formés.

### c. Organisation génomique

Le génome du VIH mesure environ 9200 paires de bases (pb). Il est composé de deux molécules hétérodimériques d'ARN monocaténaire linéaire de polarité positive, coiffées en 5' et polyadénylées en 3'. Il possède trois cadres de lectures ouverts chevauchants (ORF ; *Open Reading Frame*) (Figure 3).

Trois principaux gènes codent pour les protéines de structure (Tableau 1) :

- Le gène *gag* (*group antigen*), code pour les protéines de structure interne (CA, MA, NC)
- Le gène *pol* (*polymerase*), code pour les enzymes de réplication virale (PR, TI et IN)
- Le gène *env* (*enveloppe*), code pour les deux glycoprotéines d'enveloppe (gp120 et gp41)

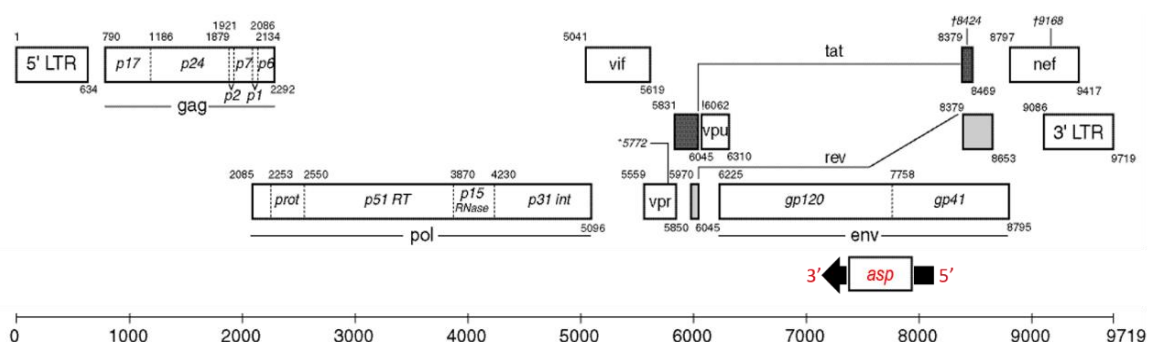


Figure 3 : Organisation du génome du VIH-1

Schéma adapté de *Landmarks of the HIV-1 Genome, HXB2 Strain* de Los Alamos (24).

Les ORFs sont représentées par les rectangles. Le début de chaque gène est indiqué par un nombre en haut à gauche de ces rectangles et la fin du gène en bas à droite. La position de ces gènes est indiquée en paires de bases par l'échelle en bas du schéma selon la souche de référence HxB2 (n° d'accèsion GenBank K03455). La position du gène *asp* a été représentée en rouge pour indiquer le sens inverse de lecture de ce nouvel ORF, entre les positions 7917 et 7342 pb. Le sens de lecture de ce nouvel ORF est indiqué par la flèche noire.

Les gènes *vif* (*viral infectivity factor*), *vpr* (*viral protein r*), *vpu* pour le VIH-1 (*viral protein u*) ou *vpx* pour le VIH-2 (*viral protein x*) et *nef* (*negative regulatory factor*), sont dits « accessoires » et sont impliqués dans différentes fonctions listées dans le Tableau 1, dont l'augmentation de l'infectiosité des virions ou du pouvoir pathogène, l'échappement au système immunitaire, la libération hors de la cellule, ou encore la régulation du cycle de réplication.

Les gènes *tat* (*transcription transactivator*) et *rev* (*regulator of viral expression*) sont appelés gènes de régulation et codent pour des protéines impliquées dans le cycle de réplication viral et le pouvoir pathogène du virus (Tableau 1).

Une nouvelle protéine a été identifiée en 2016, la protéine antisens *Asp* (*antisens protein*) (25), dont les propriétés au cours de la réplication du VIH ne sont pas encore complètement connues (26, 27). Le gène *asp* a été localisé au niveau de l'enveloppe dans le sens opposé à tous les gènes décrits (Figure 3) (28). Il a été décrit parmi des isolats de VIH-1/M et est fortement conservé, mais n'a pas été identifié chez les VIH-2 et VIH-1 non-M (VIH-1/O, VIH-1/N et VIH-1/P) (29).

Comme tous les rétrovirus, le génome ARN de polarité positive du VIH a la particularité d'être rétro-transcrit en ADNc bicaténaire par la TI, au cours du cycle de réplication viral. Cette conversion peut s'effectuer grâce à l'activité enzymatique RNase H de la TI, qui dégrade l'ARN, une fois rétro-transcrit en ADNc. Cet ADNc viral est ensuite intégré à l'ADN cellulaire de l'hôte grâce à l'action de l'IN (23). Cette forme intégrée au sein du noyau cellulaire est appelée ADN proviral, ou provirus et mesure environ 9700 pb. Elle possède des séquences répétées en 5' et 3' à ses extrémités 5' et 3', nommées 5' et 3'LTRs (*Long Terminal Repeat*) (Figure 3). D'une longueur d'environ 650 pb, les LTRs sont composés des trois mêmes régions : U3 (Unique en 3'), R (Répétée aux deux extrémités de l'ARN) et U5 (Unique en 5'). Ils participent à l'intégration du génome viral, à sa réplication et son expression. L'extrémité 5' du génome contient le promoteur d'initiation de la transcription, tandis que l'extrémité 3' contient les signaux de terminaison de la transcription et de polyadénylation. Malgré la séquence identique de part et d'autre du génome, ces deux LTRs ne jouent pas le même rôle au cours du cycle de réplication, les signaux de polyadénylation sont par exemple ignorés au niveau du 5'LTR, mais efficacement utilisés en 3'LTR pour permettre la dimérisation du génome viral (30, 31). De plus, sous forme ARN, il n'y a pas de région U3 en 5'LTR et de région U5 en 3'LTR.

Tableau 1 : Description des gènes du VIH-1 et de leurs propriétés

Gène	Protéine virale	Nom, taille	Fonction(s)
<b>Protéines de structure</b>			
<i>gag</i>	Matrice	MA, p17	Participe en phase précoce du cycle de réplication du VIH, à l'adressage de l'ARN vers la membrane plasmique puis à l'assemblage des virions.
	Capside	CA, p24	Joue un rôle de protection du génome viral et permet le transport de ce dernier vers le noyau cellulaire
	Nucléocapside	NC, p7	Se lie à l'ARN viral pour le protéger, médie l'oligomérisation et l'assemblage de <i>gag</i>
	p6	p6	Protéine requise lors de la phase tardive de bourgeonnement et de libération des virions, qui permet l'incorporation de Vpr au sein de la nouvelle particule virale
<i>pol</i>	Transcriptase inverse	RT, p66, p51	ADN polymérase ARN dépendante qui convertit les ARN viraux en ADNc
	Intégrase	IN, p32	Permet l'intégration de l'ADNc viral dans le génome de la cellule infectée
	Protéase	PR, p15	Entraine le clivage de la polyprotéine Gag-Pol
<i>env</i>	Glycoprotéine d'enveloppe de surface	SU, gp120	Se lie au récepteur CD4 et aux corécepteurs CXCR4 et/ou CCR5 de la cellule cible.
	Glycoprotéine d'enveloppe transmembranaire	TM, gp41	Protéine transmembranaire qui entraine la fusion de l'enveloppe virale avec la membrane cellulaire pour permettre l'entrée du virus dans la cellule
<b>Protéines accessoires</b>			
<i>vif</i>	Facteur d'infectiosité viral	Vif, p23	Participe à la synthèse de l'ADN proviral et augmente l'infectiosité des virions
<i>vpr</i>	Protéine virale R	Vpr, p15	Participe à phénomène d'échappement au système immunitaire, régule la libération des virions néoformés et le cycle cellulaire
<i>vpu</i>	Protéine virale U	Vpu, p16	Participe au bourgeonnement des virions avant libération hors de la cellule
<i>nef</i>	Facteur régulateur négatif	Nef, p27	Améliore le pouvoir pathogène du virus et régule le cycle de réplication
<b>Protéines régulatrices</b>			
<i>tat</i>	Trans-activateur de la transcription	Tat, p14	Permet d'effectuer un rétrocontrôle positif de la transcription de l'ADN viral
<i>rev</i>	Régulateur de l'expression protéique virale	Rev, p19	Permet de réguler l'épissage des ARNm au sein du noyau et l'export des ARNg viraux vers le cytoplasme cellulaire
<b>Protéine asp</b>			
<i>asp</i>	Protéine antisens	Asp	favoriserait la réplication virale

Adapté d'après le Traité de virologie médicale, 2019 (23).

## 4. Cycle de réplication

Le cycle de réplication du VIH est composé de neuf étapes successives (Figure 4).

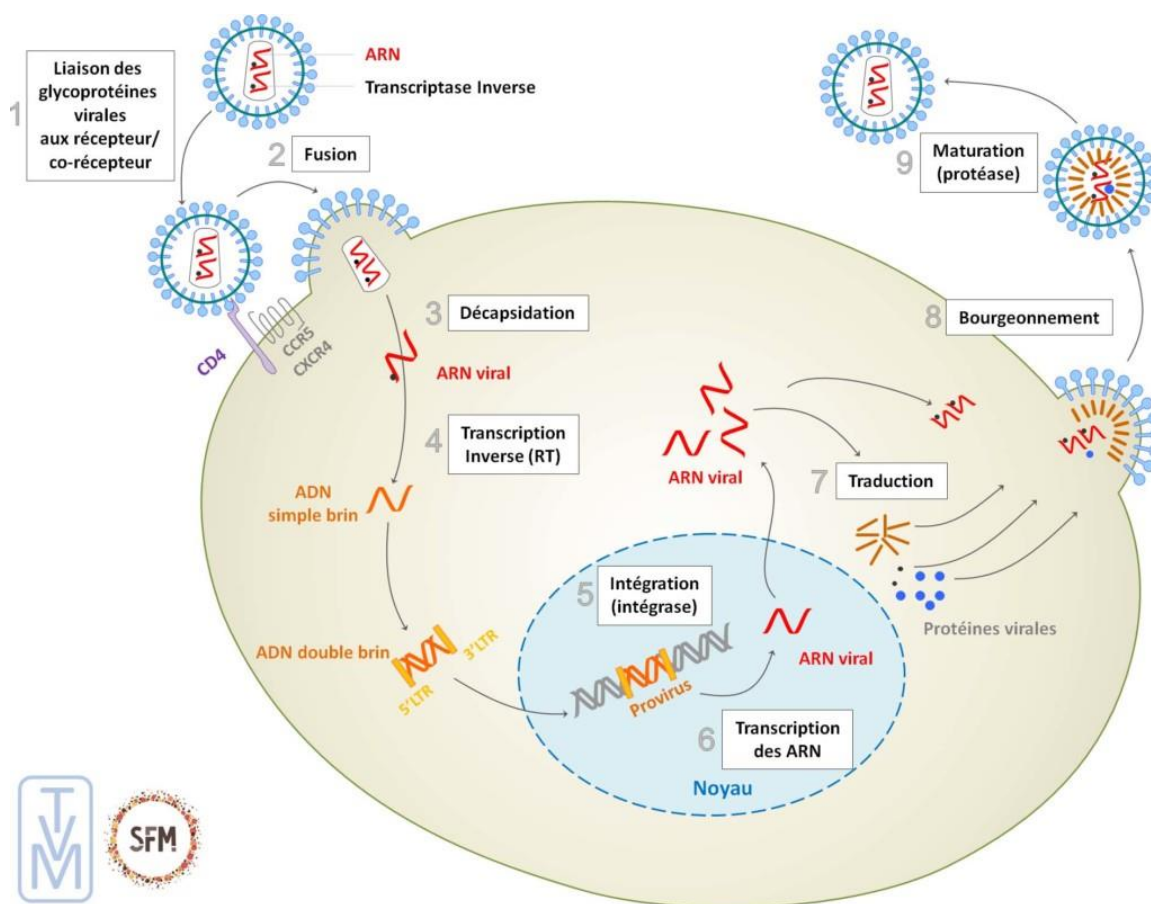


Figure 4 : Etape du cycle de réplication du VIH-1

Image issue du Traité de virologie médicale, édition 2019 (23).

Les différentes étapes du cycle de réplication sont représentées par des chiffres en gris sur la figure. 1) Attachement des glycoprotéines virales à la membrane cellulaire de la cellule cible via les récepteurs et corécepteurs membranaires. 2) Fusion de l'enveloppe virale avec la membrane plasmique de la cellule hôte et pénétration de la capsid dans le cytoplasme cellulaire. 3) Transport de la capsid près du noyau grâce au détournement du réseau de microtubules du cytosquelette cellulaire. 4) Transcription inverse de l'ARN viral au sein de la capsid en parallèle du transport nucléaire. L'ARN viral est alors rétro-transcrit en ADN double brin dans un complexe de pré-intégration comportant les deux LTRs complets 5) Après import nucléaire via la capsid, le complexe de pré-intégration permet à l'intégrase d'incorporer l'ADN double brin viral au génome de la cellule hôte de façon définitive. 6) Transcription des ARN pré-messagers qui seront ensuite épissés puis transportés vers le cytoplasme. 7) Traduction des ARNm viraux en protéines virales et réplication du génome en ARNg. Assemblage de la particule virale, comportant les protéines virales ainsi que les deux copies d'ARNg, près de la membrane plasmique. 8) Bourgeonnement des particules néoformées pour l'acquisition de l'enveloppe virale par bourgeonnement à la surface de la membrane cellulaire. 9) Maturation des virions grâce à l'action de la protéase.

La première étape consiste en l'attachement du virus à sa cellule-hôte. L'enveloppe virale contient plusieurs domaines fonctionnels, notamment au sein de la GP120 qui possède des régions conservées (C1 à C5) et des régions hypervariables (V1 à V5) et qui permet la reconnaissance du virus par la cellule cible (32). Le phénomène d'adsorption se fait grâce à la reconnaissance spécifique du récepteur cellulaire CD4 par la GP120, au niveau du site de fixation. Il s'agit d'une liaison de forte affinité puisqu'une seconde liaison a lieu au niveau d'un corécepteur cellulaire accolé au CD4. L'affinité pour l'un ou l'autre de ces corécepteurs détermine le tropisme viral. Un virus de tropisme X4, est capable de reconnaître uniquement le corécepteur CXCR4 (*chemokine receptor type 4*) alors qu'un virus de tropisme R5 ne reconnaît que des cellules exprimant le corécepteur CCR5 (*chemokine receptor type 5*). Les deux corécepteurs peuvent être reconnus par le même virus, on parle alors de double tropisme X4/R5. Un changement de conformation entraîne l'exposition de la boucle V3 de la GP120, au corécepteur cellulaire, ce qui met en contact la GP41 contenant le peptide de fusion, avec la membrane de la cellule cible (33). La fusion de l'enveloppe virale avec la membrane plasmique permet la pénétration de la capsid virale dans le cytoplasme cellulaire.

Des études récentes ont montré que la décapsidation n'avait pas lieu au sein du cytoplasme, comme décrit initialement (34, 35). Le transport de la capsid près du noyau s'effectuerait suite au détournement du réseau de microtubules du cytosquelette cellulaire, de façon simultanée avec le processus de transcription inverse qui a lieu au sein de la capsid partiellement dégradée (34). La capsid reste quasi-intacte jusqu'à son acheminement au niveau des pores nucléaires de la cellule, afin de protéger le génome viral et éviter la génération de formes défectives du virus. Néanmoins, deux hypothèses restent étudiées à ce jour : soit la décapsidation a lieu au niveau des pores nucléaires, soit elle a lieu au sein du noyau de la cellule (35).

Le mécanisme de transcription inverse est une étape essentielle de la réplication du VIH. L'ensemble des étapes de ce mécanisme est représenté dans la Figure 5.



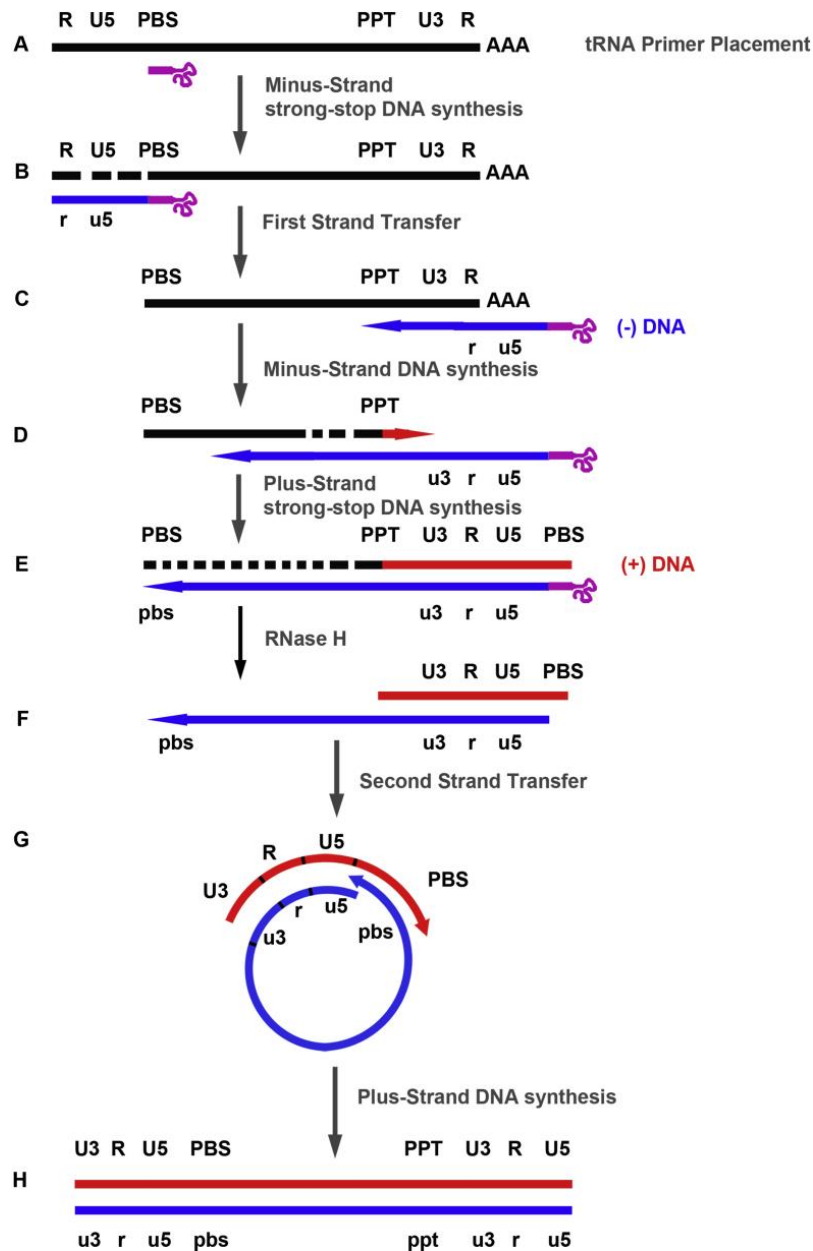


Figure 5 : Transcription inverse du génome de l'ARN rétroviral du VIH

Image issue de René et al, 2018 (36).

Les différentes étapes du mécanisme sont indiquées par des lettres à gauche des figures. L'ARN figure en noir, tandis que l'ADN (-) est représenté en bleu et l'ADN (+) en rouge. A) L'extrémité 3' d'un ARNt cellulaire (violet) est recrutée au niveau de la région PBS de l'ARN rétroviral. B) La synthèse de l'ADN (-) est initiée à partir de l'amorce constituée par l'ARNt jusqu'à l'extrémité 5' du génome ; ce produit est appelé ssDNA ; l'ARN génomique est alors dégradé par l'activité RNase H de la TI. C) L'ADN (-) est transféré à l'extrémité 3' de l'ARN génomique. D) Après le premier transfert de brin, l'élongation du brin d'ADN (-) et la dégradation par la RNase H se poursuivent ; la synthèse de l'ADN (+) est initiée à partir de la séquence PPT (PolyPurin Trackt) qui résiste au clivage par la RNase H. E) Élongation des deux brins d'ADN jusqu'aux séquences PBS respectives (Promotor Binding Site). F) L'ARN rétroviral, l'ARNt et les amorces PPT sont dégradés par l'activité RNase H de la TI. G) Transfert du second brin d'ADN médié par l'appariement des séquences PBS complémentaires aux extrémités 3' de l'ADN (-) et de l'ADN (+). H) Élongation de l'ADN (-) et de l'ADN (+) ce qui permet l'obtention d'un ADN double brin linéaire comprenant deux LTRs complets avec les trois régions U3, R et U5, à chaque extrémité du génome.

Premièrement, la synthèse d'un ADN (-) est initiée grâce à un ARN de transfert (ARNt) qui sert d'amorce au niveau d'un site de liaison de l'ARN viral, appelé PBS pour *Promotor Binding Site*. Cet ADN monocaténaire ainsi synthétisé est transféré au niveau de la région R du 3'LTR au cours d'un premier transfert de brin (Figure 5 : B/C). Ce transfert entraîne la dégradation de l'ARN génomique (ARNg) matrice, via l'activité RNase H de la TI. L'élongation de l'ADN (-) et la dégradation de l'ARNg se poursuivent, permettant la synthèse du second brin d'ADN (+), au niveau du PPT (Polypurin Tract), qui lui résiste à la dégradation effectuée par la TI. L'élongation des deux brins d'ADN jusqu'aux séquences PBS respectives se poursuit tout en supprimant l'ARN viral matrice résiduel, ainsi que l'ARNt et le PPT. L'appariement des séquences PBS à chaque extrémité des deux brins d'ADN, conduit à un deuxième transfert de brin (Figure 5 : F/G), qui permet leur élongation et l'obtention d'un ADN double brin linéaire, composé des deux LTRs complets à chaque extrémité du génome. Ces deux transferts d'une matrice à l'autre sont dits « obligatoires » car ils permettent le bon déroulé du mécanisme de transcription inverse. Du fait de la présence simultanée de deux copies d'ARN viral, des sauts de matrice supplémentaires et facultatifs de la TI peuvent avoir lieu au cours de ce processus et générer des événements de recombinaison génétique. Le phénomène de recombinaison génétique a lieu durant le processus de transcription inverse de la réplication virale et correspond à la formation d'un génome chimérique composé de plusieurs fragments d'ARNg d'origines parentales différentes (37). Ce phénomène sera détaillé dans le paragraphe correspondant page 35 (**1. La recombinaison génétique**).

L'ADN proviral est intégré au génome cellulaire sous une forme d'un provirus, via l'action de l'intégrase. Ce provirus constitue des réservoirs viraux au sein de l'organisme infecté (38, 39). Le mécanisme d'intégration peut avoir lieu partout au sein du génome cellulaire mais semble avoir néanmoins des sites d'intégration préférentiels, au niveau des gènes cellulaires enrichis en chromatine proches des pores nucléaires (40).

L'ADN proviral est transcrit en ARN pré messager par une enzyme cellulaire, l'ARN polymérase II ADN dépendante (41). Cette transcription est sous le contrôle du gène *tat* et de nombreux facteurs de transcription cellulaires qui interagissent avec la région U3 du LTR et permet la synthèse de différents ARNm (ARN messagers) épissés alternativement (42). Les ARNm sont ensuite exportés vers le cytoplasme permettant d'une part, l'expression des protéines virales grâce au mécanisme de traduction des ARNm, après détournement de la machinerie de traduction cellulaire (43, 44). D'autre part, des copies de ces ARNm sont conservées pour servir d'ARNg et sont encapsidées pour constituer le génome viral du nouveau virion (45, 46). Cette encapsidation des protéines et du génome viral s'effectue au niveau de la membrane plasmique cellulaire. La particule virale néoformée va ainsi bourgeonner et acquérir une enveloppe virale à partir de la membrane plasmique de la cellule infectée.

Au cours d'une dernière étape, la protéase virale clive les polyprotéines Gag et Gag-Pol, afin d'assurer la maturation des particules virales après libération hors de la cellule. Cette étape est primordiale car elle assure la viabilité et l'infectiosité des nouveaux virions, qui pourront à leur tour se répliquer dans de nouvelles cellules (47).

## II. Origines et diversité génétique des VIH

### 1. Origines simiennes

En 1989, l'identification du Virus de l'Immunodéficience Simienne (VIS) chez une espèce de chimpanzé *Pan troglodytes troglodytes* (VIScpzPtt) révèle une organisation génomique très similaire entre le VIS de ce singe et le VIH-1 (48, 49). Très vite, une origine simienne du VIH est suspectée et des analyses phylogénétiques révèlent que cette proximité est le reflet de plusieurs événements de transmission inter-espèces du Singe à l'Homme. Des études ont montré que le VIScpzPtt est issu de la recombinaison de trois VIS de petits singes (50). Le VIScpzPtt est à l'origine du VIS du Gorille des plaines occidentales du Cameroun *Gorilla gorilla gorilla* (VISgor) mais aussi des VIH-1 de groupes M et N, après deux événements différents de transmission inter-espèce entre le chimpanzé et l'Homme (51, 52). Le VISgor est quant à lui à l'origine de deux événements indépendants de transmission à l'Homme, ayant abouti aux VIH-1 de groupes O et P (53). Au final, quatre événements de transmission inter-espèces de singes à l'Homme ont conduit à l'émergence des quatre groupes de VIH-1 (54) (Figure 6).

Le VIH-2 provient quant à lui du VISsmm (*sooty mangabey monkey*). Le mangabey enfumé (*Cercocebus atys*) est un petit singe retrouvé au niveau des mêmes zones géographiques que les cas identifiés de VIH-2, il est souvent domestiqué ou chassé pour sa viande (55) (Figure 6). Une étude a d'ailleurs montré que les deux virus possèdent en commun le gène *vpx* (54). Enfin, une autre étude a décrit au total neuf événements de transmission inter-espèce du VISsmm à l'homme ayant donné autant de groupes de VIH-2 nommés de A à I (12).

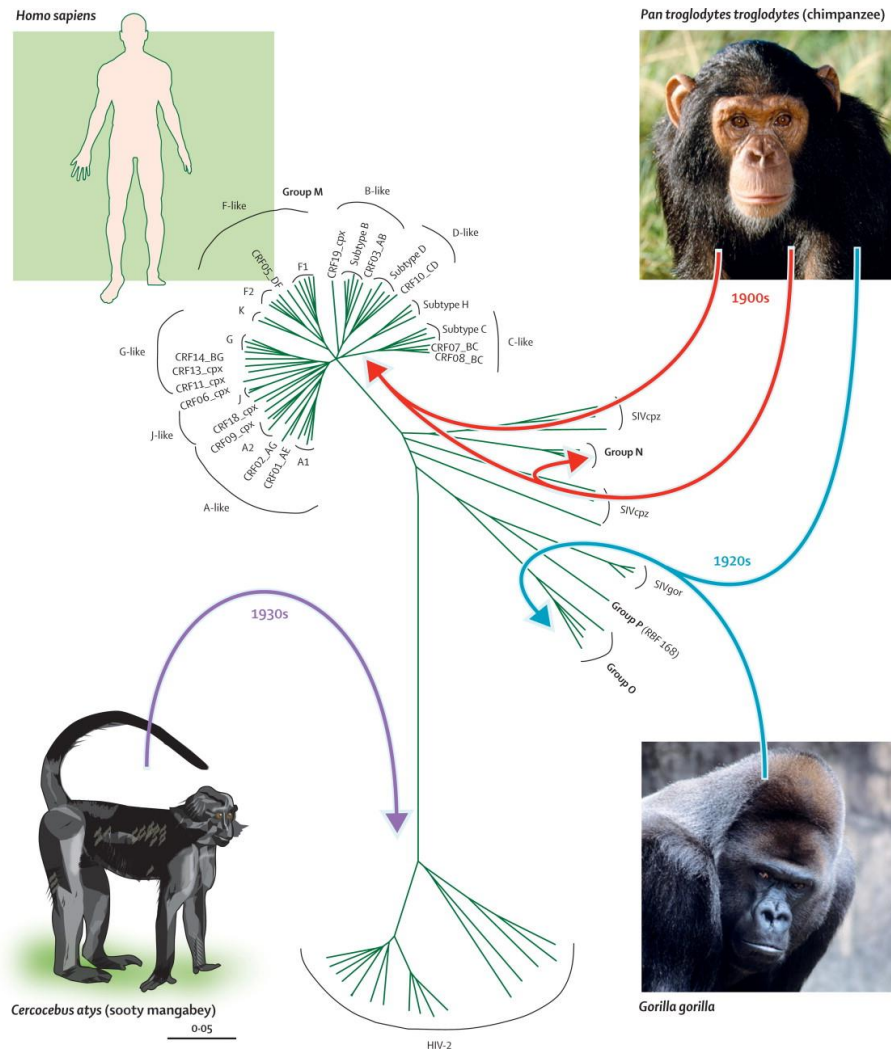


Figure 6 : Relations phylogénétiques et diversité des VIH-1, VIH-2 et SIV

Image issue de Tebit et al, 2011 (56).

Arbre phylogénétique montrant les relations phylogénétiques existantes entre les SIV de plusieurs singes, avec les différents VIH existants à ce jour. CRF = Circulating Recombinant Form ; cpz = Chimpanzé ; gor = Gorille ; cpx = complexe ; SIV = Simian Immunodeficiency Virus. Les événements de transmission inter-espèces figurent en bleu pour le VIH-1/M et VIH-1/N, en vert pour le VIH-1/O et VIH-1/P et en violet pour le VIH-2.

L'origine de la pandémie est localisée en République Démocratique du Congo où l'on a identifié les plus anciennes souches de VIH-1 de groupe M, montrant une diversité moléculaire déjà importante à l'époque. L'émergence du VIH-1 a ainsi pu être estimée vers les années 1920 (57, 58). La datation de l'ancêtre commun le plus récent (*Most Recent Common Ancestor*, MRCA) a mis en évidence un MRCA entre 1909 et 1930 pour le VIH-1/M (57) et entre 1914 et 1944 pour le VIH-1/O (59). Concernant les VIH-1/N et VIH-1/P, cette datation est plus complexe car il existe peu de séquences génétiques disponibles à ce jour, mais on estimerait pour le moment leur apparition entre 1948 et 1977 pour le groupe VIH-1/N (60) et dans les années 1980 pour le VIH-1/P (61). Les neuf groupes de VIH-2 (A - I), proviennent de neuf événements de transmission du SIVsmm à l'Homme (50) qui auraient eu lieu à partir d'un même MRCA datant de 1955 environ (62).

## 2. Adaptation à l'Homme et diffusion

L'hypothèse la plus probable est que la transmission du VIS à l'Homme est liée aux activités humaines de chasse et de préparation de viande de brousse, qui peuvent impliquer un contact du sang contaminé avec une blessure humaine ou des morsures par des singes infectés (54, 63). L'exposition de l'homme au VIS dans certaines régions d'Afrique Centrale et de l'Ouest est très importante, certaines espèces de singes sont domestiquées et parfois fortement contaminées, comme c'est le cas du mangabey enfumé (51). Enfin, la proximité génétique entre Gorille, Chimpanzé et Homme a favorisé la transmission et l'adaptation du VIS chez son nouvel hôte, puisque la barrière d'espèce dépend en partie de la divergence génétique entre les individus (50).

La pression de sélection est le moteur de l'évolution et de l'adaptation, entraînant une compétition entre le virus et l'hôte infecté. L'acquisition de certaines mutations suite à la pression de sélection intra-hôte a permis de favoriser l'émergence de souches plus adaptées, comme par exemple la présence d'un résidu arginine (R) en position 30 de la MA p17, retrouvée chez tous les VIH-1 de groupes M, O et N qui ont diffusé dans la population (64). L'acquisition de cette mutation signature pour trois des quatre groupes de VIH-1 démontre bien la convergence évolutive vers l'acquisition d'un résidu R à cette position. Néanmoins, la première souche de VIH-1/P identifiée en 2009 possédait, un résidu 30M, liée à la présence d'une méthionine à cette même position, comme pour les SIVcpzPtt et le SIVgor. Ces données suggéraient que l'acquisition de cette méthionine en position 30 n'empêchait pas la transmission à l'homme, mais ne favoriserait peut-être pas la diffusion plus large au sein de la population (65). Une deuxième souche de VIH-1/P a été identifiée en 2011 et portait une lysine à cette position (30K), suggérant une adaptation de ce groupe puisque son taux de réplication *in vitro* était très élevé (66).

Récemment, une étude s'est portée sur l'évolution phylogénétique de l'IN des VIH-1/M en comparaison avec celle des VIH-1/O (67). L'IN est une protéine qui possède plusieurs domaines fonctionnels, dont celui nécessaire à l'intégration du génome viral dans la cellule hôte. Cette fonction est assurée par le motif CLA (*C-terminal lysine-amidic*), positionné en C-terminal de l'IN et conservé chez les VIH-1/M, tandis que cette fonction est assurée par le motif NOG (*N-terminal O group*), positionné en N-terminal de l'IN et conservé chez les VIH-1/O. Cette étude a pu mettre en évidence que le motif NOG des VIH-1/O avait été hérité de l'IN du VISgor, possédant le même motif, mais que le motif CLA, assurant cette même fonction pour les VIH-1/M, n'était pas retrouvé à l'identique chez le VIScpz (67). L'acquisition et la conservation du motif CLA pour l'IN de VIH-1/M a donc été nécessaire, après transmission du virus du chimpanzé à l'Homme, montrant bien l'adaptation nécessaire du virus à son nouvel hôte.

Par ailleurs, les mécanismes développés par le VIH pour contourner l'action antivirale des facteurs de restriction cellulaires illustrent également cette adaptation des SIV à l'Homme. La Tetherine est une protéine de la membrane cellulaire capable d'empêcher la libération des virions nouvellement formés de la cellule infectée en les retenant à la surface cellulaire pour dégradation ultérieure (68). Chez les chimpanzés, l'action de la Tetherine est contournée par la protéine accessoire Nef du SIVcpz, mais la région ciblée par Nef est délétée dans la Tetherine humaine, ce qui la rend résistante à Nef. Les VIH-1/M et VIH-1/N ont ainsi développé un autre moyen de contourner l'action antivirale de la Tetherine par l'action de Vpu, ce qui n'est pas le cas des VIH-1/O et VIH-1/P. Ce défaut d'adaptation des VIH-1/O et VIH-1/P serait en partie responsable de leur diffusion restreinte (69).

### 3. Diversité génétique

La diversité génétique des VIH est multi-factorielle. Chaque événement réussi de transmission du Singe à l'Homme a entraîné une évolution virale différente chez l'hôte infecté et, comme mentionné précédemment, a conduit à l'émergence de différents types et groupes de VIH, ayant leurs propres caractéristiques virologiques. Ces caractéristiques ont été acquises grâce à la différence de pression de sélection exercée différemment chez le Singe et l'Homme. Les pressions immunitaires ne sont pas les mêmes chez ces espèces. De plus, l'Homme est soumis à une pression thérapeutique qui est absente chez le Singe. Cette différence de pression, favorise l'émergence et la sélection de mutations différentes au cours du temps, entraînant parfois l'échappement au système immunitaire et la résistance aux traitements antirétroviraux (ARV). C'est le cas de la convergence évolutive observée pour la mutation R30M déjà décrite au paragraphe précédent, qui a favorisé la diffusion chez l'Homme des VIH-1 de groupes M, O et N (64).

Il existe également des facteurs intrinsèques au virus qui conduisent à une variabilité importante d'une même population virale. En effet, la TI est une enzyme peu fidèle car elle est exempte d'activité de correction exonucléasique 3'-5', commettant une erreur tous les 10 000 nucléotides (70–72). Par ailleurs, on constate également la présence de mutations aléatoires qui participe à la diversification du virus. La faible fidélité de la TI couplée à un taux de réplication élevé du virus d'environ  $10^9$  à  $10^{10}$  virions/jours chez un individu non traité (73) entraînent l'établissement d'une quasi-espèce virale. Cette notion fait référence à un ensemble homogène de virus, dynamique au cours du temps et soumis à un processus continu de variation génétique, de pression de sélection et de compétition, entraînant l'apparition de mutations et source de diversité génétique (74, 75). De plus, il a été mis en évidence une évolution génétique intragroupe importante, menant à la description de nombreux sous-types et

sous-groupes de VIH-1 qui seront détaillés page 30 (**III. Classification des VIH**). La diversité des VIH au cours du temps est également liée au phénomène de réactivation des réservoirs constitués par le virus intégré au sein du génome de la cellule, qui permettent la persistance au cours du temps malgré l'instauration de traitement et favorisant l'apparition de mutations lorsque ce traitement est arrêté ou inefficace (75, 76).

Enfin, le phénomène de recombinaison génétique est également source de diversité génétique (77) et fera l'objet d'un paragraphe plus détaillé page 33 (**1. La recombinaison génétique**).

Tous ces facteurs conduisent à l'existence de nombreux variants du VIH qui peuvent entraîner des conséquences physiopathologiques, épidémiologiques, thérapeutiques, diagnostiques et vaccinales.

### **III. Classification des VIH**

De nombreuses souches de VIH ont été décrites depuis sa découverte, il y a maintenant environ 40 ans. Tous ces VIH ont été classés en fonction de leur homologie de séquence. Il existe actuellement deux types de VIH, les VIH-1 et VIH-2 et 4 groupes de VIH-1 (M, O, N et P). Le VIH-1/M a été divisé en différents sous-types et formes recombinantes intra-groupe M et le VIH-1/O, en différents sous-groupes.

#### **1. Les VIH de type 1 (VIH-1)**

##### **a. Les VIH-1 de groupe M (VIH-1/M)**

Les VIH-1 de groupe M, *major* (VIH-1/M), sont à l'origine de la pandémie actuelle et représentent plus de 90% des infections à VIH identifiés aujourd'hui. On compte actuellement 10 sous-types de VIH-1/M, nommés A, B, C, D, F, G, H, J, K et L, le dernier sous-type identifié en 2020 (78). Certains de ces sous-types ont eux-mêmes évolué en sous-sous-types appelés A1, A2, A3, F2, F2. La divergence intra sous-type peut être comprise entre 15 à 20% au sein des gènes *gag* et *env*, tandis que les divergences inter sous-types pour ces mêmes gènes sont plutôt de l'ordre de 25 à 35% (79).

Il existe en plus de ces sous-types « purs », des formes recombinantes qui résultent de la recombinaison entre ces sous-types. La nomenclature de ces formes recombinantes a été décrite dans les années 2000 par Robertson *et al.* (80). D'une part, les formes recombinantes circulantes (CRF,

*Circulating Recombinant Form*) sont définies lorsqu'au moins trois individus sont infectés par une même forme recombinante sans liens épidémiologiques directs. Toutes les CRFs ont un numéro attribué par ordre chronologique de découverte et de publication et à la suite de ce numéro, figurent les sous-types impliqués au sein des différents fragments recombinants (exemple Figure 7). Il existe plus de 130 CRF décrites à ce jour (81).

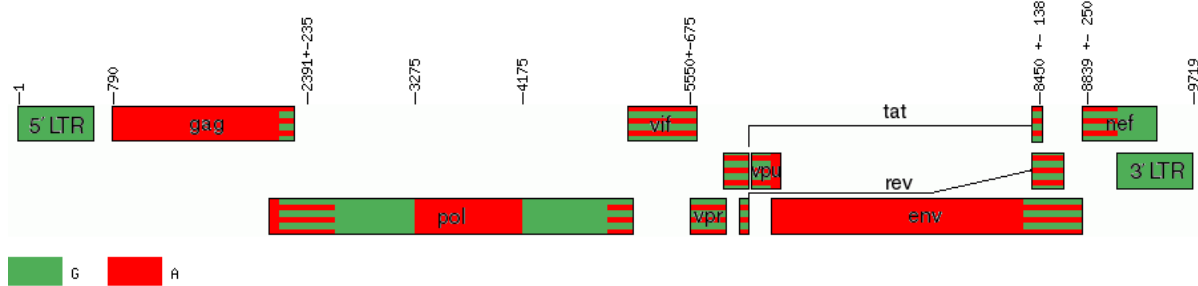


Figure 7 : Schéma de la forme recombinante CRF02\_AG

Deux sous-types sont impliqués dans la structure du génome de la forme recombinante CRF02\_AG. Les portions de VIH-1/M de sous-type G sont représentées en vert, tandis que celles de sous-type A sont en rouge. Les portions zébrées correspondent aux régions conservées entre ces sous-types. La position des points de cassure a été indiquée par rapport à la souche de référence VIH-1/M HxB2 (n° d'accèsion GenBank K03455).

A l'inverse, les formes recombinantes uniques (URF, *Unique Recombinant Form*) ne diffusent pas dans la population, ou sont retrouvées chez des personnes présentant un lien épidémiologique étroit. Il en existe plusieurs centaines décrites aujourd'hui.

Enfin, toutes ces formes recombinantes peuvent elles-mêmes recombiner à nouveau avec un sous-type « purs ou avec une autre forme recombinante, générant ainsi des formes recombinantes complexes. Le suffixe « cpx » est alors ajouté à la fin du nom de cette nouvelle forme (voir exemple Figure 8).

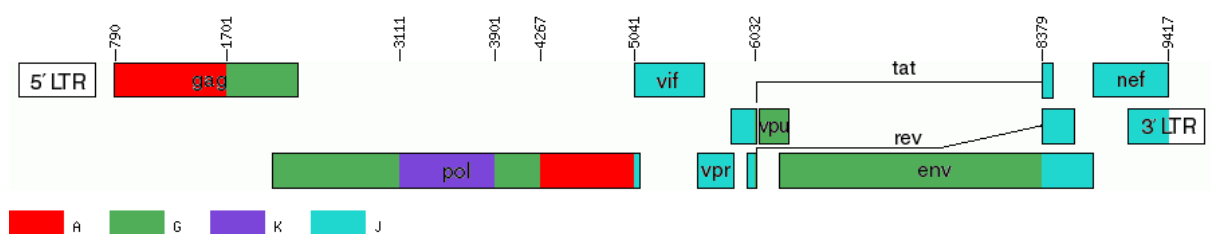


Figure 8 : Schéma de la forme recombinante CRF06\_cpx

Quatre sous-types différents (A, G, K et J) sont impliqués dans la structure du génome mosaïque de CRF06\_cpx (82). Les portions de VIH-1/M de sous-type A sont représentées en rouge, de G en vert, de K en violet et de J en bleu cyan. La position des points de cassure a été indiquée par rapport à la souche de référence VIH-1/M HxB2 (n° d'accèsion GenBank K03455).



La proportion de formes recombinantes intra-M a augmenté au fil du temps, avec pas moins de 57 nouvelles CRFs décrites en seulement cinq ans, entre 2010 et 2015 (83). Le nombre de ces formes recombinantes, incluant à la fois les CRFs et les URFs est sous-estimé et ne cesse de croître comparativement aux autres sous-types non recombinants du VIH-1/M (84, 85). Toutes ces formes recombinantes représentent 18% des infections dans le monde, dont une majorité (33,9%) est causée par un VIH-1/M de sous-type CRF02\_AG et 26,7% par une URF (26,7%) (79, 84). Cette étude soulevait l'hypothèse selon laquelle l'émergence croissante de ces formes était en lien avec l'acquisition d'un avantage réplcatif en terme d'évolution par rapport aux formes parentales dont elles sont issues.

La répartition mondiale de tous les sous-types intra-M purs et recombinants a été décrite récemment par *Bbosa et al* en 2019, indiquant une majorité de sous-type A, C et CRF02\_AG en Afrique, une majorité de sous-type B en Europe et Amérique et enfin, une majorité de CRF01\_AE en Asie. Cette étude a également permis de constater l'augmentation de la prévalence des formes recombinantes au niveau mondial, surtout au niveau des régions où plusieurs sous-types co-circulent, favorisant l'infection d'un même individu par plusieurs souches (85).

#### **b. Les VIH-1 de groupe O (VIH-1/O)**

Le VIH-1 de groupe O (VIH-1/O), ou « *Outlier* », a été découvert en 1990 (ANT-70) chez un couple camerounais (9). Ce groupe est surtout présent dans l'Ouest de l'Afrique Centrale, notamment au Cameroun où il est endémique. Quelques cas sont observés hors de ces régions, en Europe et aux Etats Unis, mais restent très minoritaires et liés aux relations historiques entre ces continents.

Les VIH-1/O présentent une divergence génétique importante par rapport aux VIH-1/M (9), d'environ 50% dans le gène *env* entre les souches de référence VIH-1/M HxB2 et VIH-1/O ANT-70 (59). Les différents pourcentages de divergence retrouvés entre ces deux groupes au niveau des trois gènes de structure du VIH-1, sont de l'ordre de 24 à 32% pour le gène *gag*, de 33 à 37% pour le gène *pol* et de 39 à 49 % pour *env* (Figure 9) (80, 86–89).

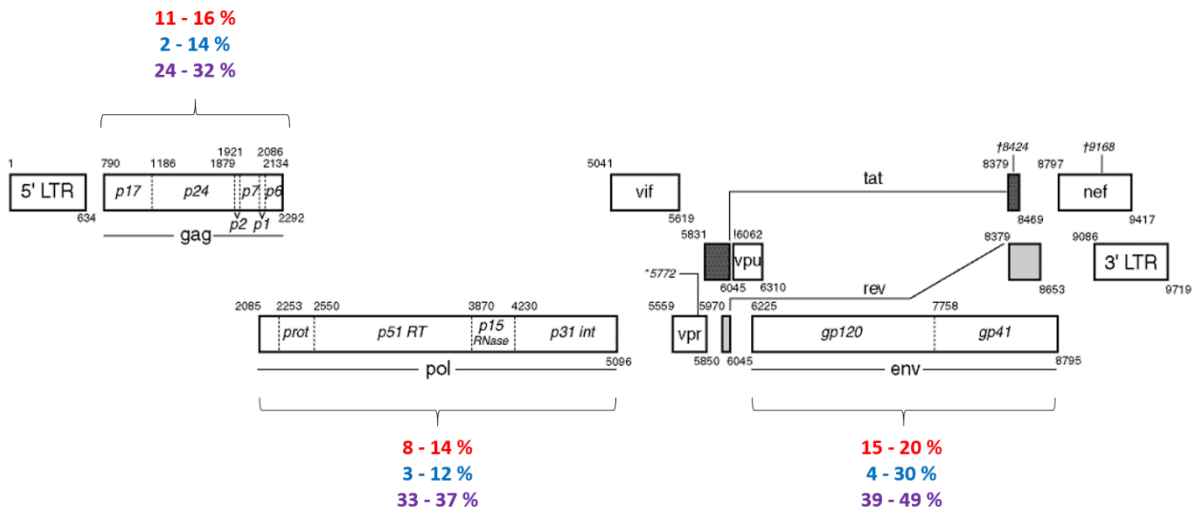


Figure 9 : Pourcentages de divergences génétiques entre les gènes de structure des VIH-1/M et VIH-1/O

Le pourcentage de divergence a été indiqué en gras au niveau des gènes de structure correspondants. Les divergences intra-groupe M sont indiquées en rouge, intra-groupe O en bleu et inter-groupes M et O en violet.

Certains VIH-1/O possèdent un polymorphisme naturel par rapport aux VIH-1/M. C'est le cas par exemple de la mutation Y181C au niveau de la TI des VIH-1/O de sous-groupe H, qui leur confère une résistance aux inhibiteurs non nucléotidiques de la transcriptase inverse INNTI (90–92). Ce polymorphisme est d'autant plus observé au niveau de la GP120, qui est fortement divergente de celle des VIH-1/M (93). Ainsi, les algorithmes de détermination du tropisme tel que geno2pheno, ne sont pas adaptés pour les VIH-1/O (87, 94, 95). Malgré cette absence d'algorithmes, il a été démontré que ces VIH-1/O avaient essentiellement un tropisme R5 et n'étaient pas répliqués au sein de cellules déficientes pour le CCR5 (95, 96).

En analysant 190 séquences de VIH-1/O, issues du RES-O (réseau de surveillance des VIH-1/O en France) et d'autres provenant du Cameroun, un arbre phylogénétique a pu être construit par la méthode du maximum de vraisemblance et confronté à celui obtenu pour 190 VIH-1/M (59) (Figure 10). La topologie de ces arbres a montré que les VIH-1/O s'étaient diversifiés au cours de deux phases évolutives successives (forme de comète) contrairement aux VIH-1/M qui résultent d'évènements évolutifs indépendants (forme de double étoile). Cette étude a donc proposé le regroupement des VIH-1/O en deux sous-groupes, correspondant aux deux phases de croissance successives ayant eu lieu au 20<sup>ème</sup> siècle, avec l'identification du sous-groupe H (*Head*), représentant la tête de la comète et du sous-groupe T (*Tail*), représentant la queue de cette comète (Figure 10).

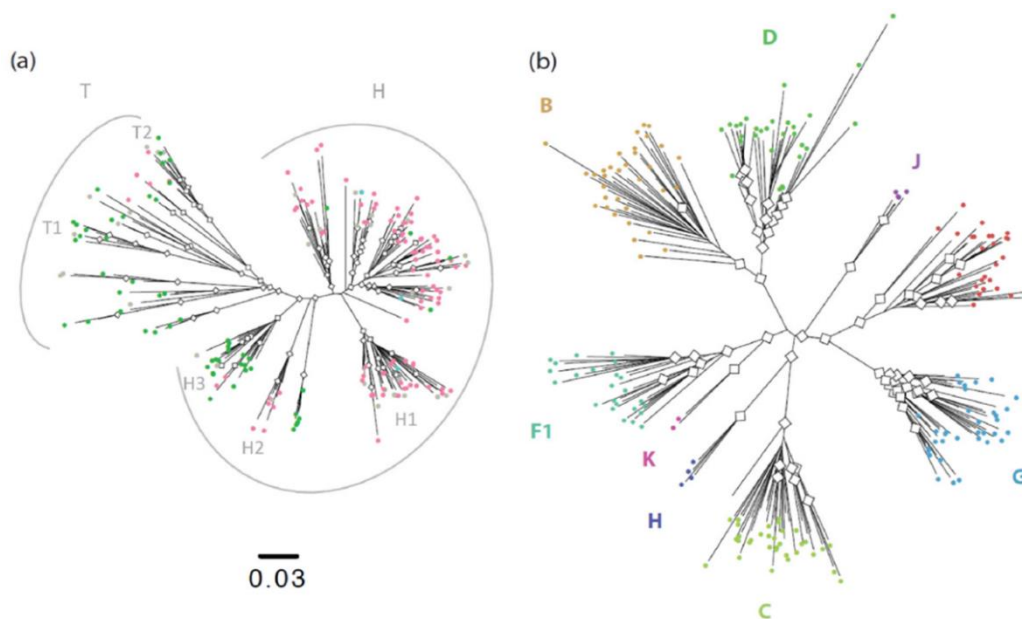


Figure 10 : Arbre phylogénétique des VIH-1 de groupes O (a) et M (b)

D'après Leoz et al. 2015 (59).

- (a) Arbre phylogénétique construit sur la base de 190 séquences de VIH-1/O, montrant une structure en forme de comète avec les VIH-1/O sous-groupe T (Tail) et H (Head).
- (b) Arbre phylogénétique construit sur la base de 190 séquences de VIH-1/M, montrant une structure en forme de double étoile avec les différents sous-types A-D, F-H, J et K du groupe M

Toutes ces différences observées entre les VIH-1/M et VIH-1/O sont liées à plusieurs facteurs, dont en partie leurs origines simiennes à partir des chimpanzés et gorilles respectivement. La pression de sélection exercée par leur hôte respectif, après transmission inter-espèce a également joué un rôle important.

### c. Les VIH-1 de groupe N et P (VIH-1/N et VIH-1/P)

Les VIH-1 de groupes N et P (VIH-1/N et VIH-1/P) ont été découverts plus récemment. C'est en 1998 que le VIH-1/N a été découvert chez un patient camerounais, son nom « N » a été choisi pour « Non M, non O » pour faire suite aux deux groupes VIH-1/M et VIH-1/O précédemment identifiés. Au total, 22 souches sont actuellement décrites chez des patients ayant tous un lien avec le Cameroun sauf pour un ayant un lien avec le Togo (97, 98). Quant au VIH-1/P, seulement deux souches ont été décrites en 2009 et 2011 chez des patients camerounais (11, 66). Trop peu de séquences des groupes N et P sont disponibles pour étudier leur évolution génétique et aucune étude sur le tropisme de ces souches n'a été effectuée à ce jour.

## **2. Les VIH de type 2 (VIH-2)**

Dans le monde, on retrouve le VIH-2 principalement en Afrique de l'Ouest, où il infecte environ 1 à 2 millions de personnes (99). Le VIH-2 est composé de deux principaux groupes épidémiques, appelés A et B et sept autres groupes conduisant à des impasses épidémiologiques, avec seulement un à deux cas décrits. En France, le groupe A est regroupé en deux sous-types A1 (Guinée Bissau, Cap Vert) et A2 (Mali).

Il existe à ce jour seulement deux formes recombinantes du VIH-2, incluant la forme CRF01\_AB (100) qui résulte de la recombinaison entre les groupes A et B et une URF retrouvée chez un seul individu, impliquant les mêmes groupes A et B (101).

## **IV.Recombinaison inter-groupes VIH-1/MO**

### **1. La recombinaison génétique**

#### **a.Prérequis de la recombinaison**

La pandémie de VIH est associée majoritairement aux VIH-1/M, tandis que les VIH-1/O sont endémiques en Afrique centrale et au Cameroun en particulier. La circulation des VIH-1/M et VIH-1/O dans cette région géographique a favorisé l'infection de mêmes patients par ces virus différents. Ce phénomène, appelé double infection VIH-1/M+O, est l'un des prérequis nécessaires à la recombinaison génétique inter-groupes MO. La double infection peut avoir lieu lors d'une infection simultanée par les deux virus (co-infection) ou chez un individu déjà infecté par un autre virus parental (surinfection).

Pour rappel, le phénomène de recombinaison intervient au cours du mécanisme de la transcription inverse et permet la génération d'un génome chimérique à partir de deux ARNg différents. Cette recombinaison ne peut donc avoir lieu que si ces deux ARNg différents sont présents dans une même cellule, pour aboutir à la formation d'un génome chimérique. Concernant la recombinaison inter-groupes MO, cela implique donc que la cellule puisse être infectée par au moins deux virus différents, VIH-1/M et VIH-1/O.

Les deux premières doubles infections VIH-1/M+O ont été décrites au Bénin et au Cameroun, en 1996 et 1998, respectivement (102, 103). Les données de ces études ne permettaient pas de conclure sur la présence de formes recombinantes VIH-1/MO associées. Depuis ces deux premiers cas identifiés, 14 autres doubles infections VIH-1/M+O ont été décrites dans d'autres études réalisées au Cameroun (104–107). Ce n'est qu'à partir de 2003-2004 que des cas de doubles infections M+O ont été décrits hors de cette région, en France (108–110), s'expliquant par les liens historiques étroits entre ces deux pays. Toutes les doubles infections identifiées hors Afrique Centrale ont été décrites chez des patients originaires du Cameroun ou ayant un lien épidémiologique avec ce pays. Dans le cadre de l'activité de surveillance associée au CNR du VIH du CHU de Rouen, un cas de surinfection par un VIH-1/M a d'ailleurs été décrit, chez une patiente camerounaise, déjà infectée par un VIH-1/O (109). Cette étude n'a pas mis en évidence d'émergence d'une forme recombinante VIH-1/MO à partir de cette double infection. Au total, 20 doubles infections VIH-1/M+O ont été décrites depuis 1996 (Figure 11).

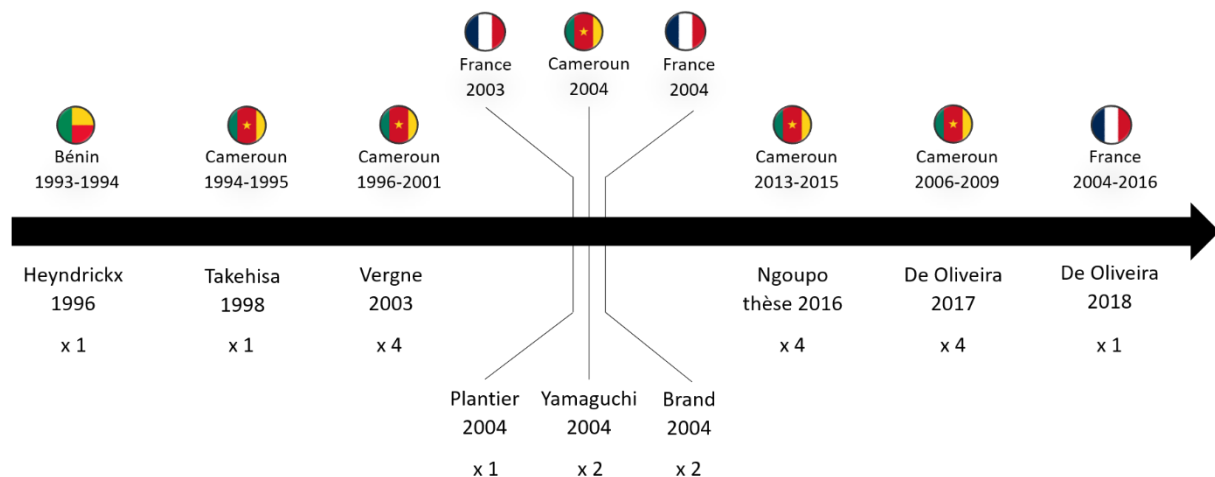


Figure 11 : Description des 20 doubles infections VIH-1/M+O

Malgré la description de ces multiples cas de doubles infections, ce n'est que plus récemment en 2020, qu'une étude a pu démontrer l'émergence *in vivo* d'une forme recombinante VIH-1/MO à partir d'une double infection M+O (111). L'analyse de prélèvements séquentiels d'une patiente initialement diagnostiquée pour une infection à VIH-1/M en 1999, suivie et traitée jusqu'en 2013, a permis d'expliquer l'échec virologique qui était constaté. Une exploration du génotypage de résistance a été réalisée avec des amorces spécifiques de groupes M et O, contrairement à celles utilisées initialement au moment du diagnostic. Une discordance moléculaire avec présence d'une PR de VIH-1/O et d'une TI de VIH-1/M, a permis de finalement montrer, d'une part, que RBF235 avait été doublement infectée par un VIH-1/M et un VIH-1/O et que, d'autre part, cette double infection M+O avait conduit à l'émergence d'une forme recombinante VIH-1/MO *in vivo* (111).

## b.Mécanisme de la recombinaison

Le mécanisme de recombinaison s'effectue en deux phases successives (Figure 12). La première étape consiste en l'infection simultanée d'une même cellule, c'est-à-dire une double infection, par deux virions parentaux différents, permettant la transcription inverse puis l'intégration de chaque génome viral parental au sein du génome cellulaire. Après expression des protéines virales et des ARNg, la progénie virale résultante est composée de virions homozygotes, identiques aux virus parentaux et de virions hétérozygotes, constitués de deux molécules d'ARN différentes (37).

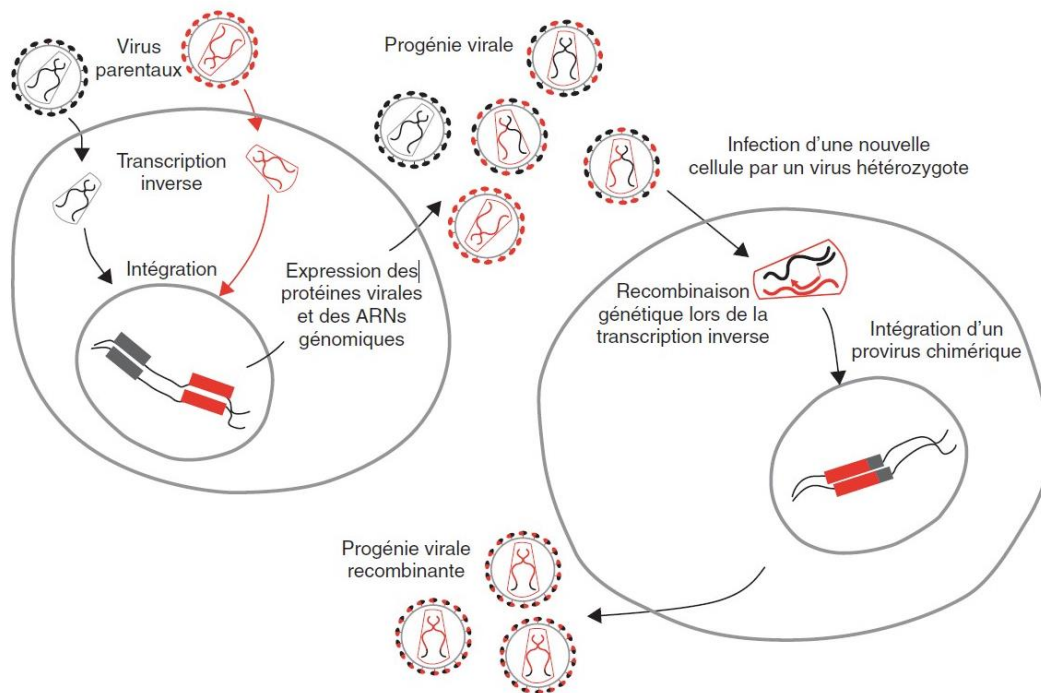


Figure 12 : Mécanisme de la recombinaison génétique

Image issue de *Simon-Lorière et al*, 2006 (37).

La première étape consiste en l'infection d'une cellule cible par deux virions parentaux différents (noir et rouge). Après transcription inverse de leur génome respectif et intégration des provirus correspondants, la progénie virale sera constituée de virions identiques aux virus parentaux, mais également de virions hétérozygotes, avec une copie d'ARNg de chaque virus parental. C'est grâce à la génération de ces virions hétérozygotes qu'une nouvelle cellule infectée par ce dernier pourra alors conduire à la recombinaison génétique. Cette dernière a lieu au cours de la transcription inverse, où la TI pourra effectuer un « saut » d'une matrice à l'autre. Le provirus chimérique correspondant pourra alors être intégré et servira pour l'expression d'une progénie virale recombinante.

Au cours de la deuxième étape, une nouvelle cellule est infectée par un virion hétérozygote précédemment formé. L'étape de recombinaison génétique a lieu au cours de cette deuxième étape, au moment de la transcription inverse. L'arrêt de la transcription, suivi par le transfert de la TI d'un brin à l'autre est appelé « saut de matrice » et provoque l'apparition de points de cassure dans le génome (112).

Selon le modèle du Choix de copie, ou « *copy-choice* », le plus largement reconnu, la recombinaison a lieu durant la synthèse d'ADN (-). La TI « sauterait » ainsi d'un ARN donneur (premier brin, de polarité positive) à un ARN accepteur (deuxième brin, de polarité négative). Ce modèle a été décliné au fil du temps et le modèle du choix de copie forcé, ou « *forced copy-choice* », implique un changement de matrice par la TI lorsque cette dernière rencontre un obstacle sur sa trajectoire, lors de la synthèse du nouveau brin (113).

Ce modèle a été élargi afin d'inclure, par exemple, le phénomène d'Induction de pause, qui induit un changement de matrice lorsque la TI rencontre un codon stop ou un signal d'arrêt anormal sur la séquence (114, 115), ou encore l'induction d'une cassure liée à la structure de l'ARN, notamment lié à la présence d'une structure secondaire de l'ARN (37). Le modèle de choix de copie dynamique « *dynamic copy-choice* », proposé par *Hwang et al*, décrit un équilibre entre l'activité de la polymérase et l'activité RNase H de la TI, influençant le changement de matrice. Lorsque l'activité de la polymérase est élevée et l'activité RNase H de la TI est faible, alors la probabilité d'un changement de matrice augmente. A l'inverse, si cette activité RNase H est faible alors que l'activité polymérase reste élevée, la possibilité de recombinaison sera jugée faible car la TI restera liée au brin néosynthétisé (116). Ces différents mécanismes sont détaillés dans la Figure 13.

Il a été estimé que la TI pouvait se dissocier de sa matrice au moins huit fois par génome ou encore provoquer entre trois et 12 changements de matrice par génome et par cycle de réplication (117). La fréquence de cette recombinaison n'est pas la même d'un virus à un autre, mais semble très importante pour le VIH en particulier car ce taux de recombinaison, d'environ  $1,4 \cdot 10^5$  par base et par cycle, dépasse celui de la mutation qui est d'environ  $1 \cdot 10^5$  mutations spontanées par base et par jour (118, 119).

Le génome chimérique néosynthétisé doit néanmoins pouvoir être reconnu par la machinerie de transcription et d'assemblage de la cellule pour pouvoir générer un virion viable et fonctionnel (120, 121).

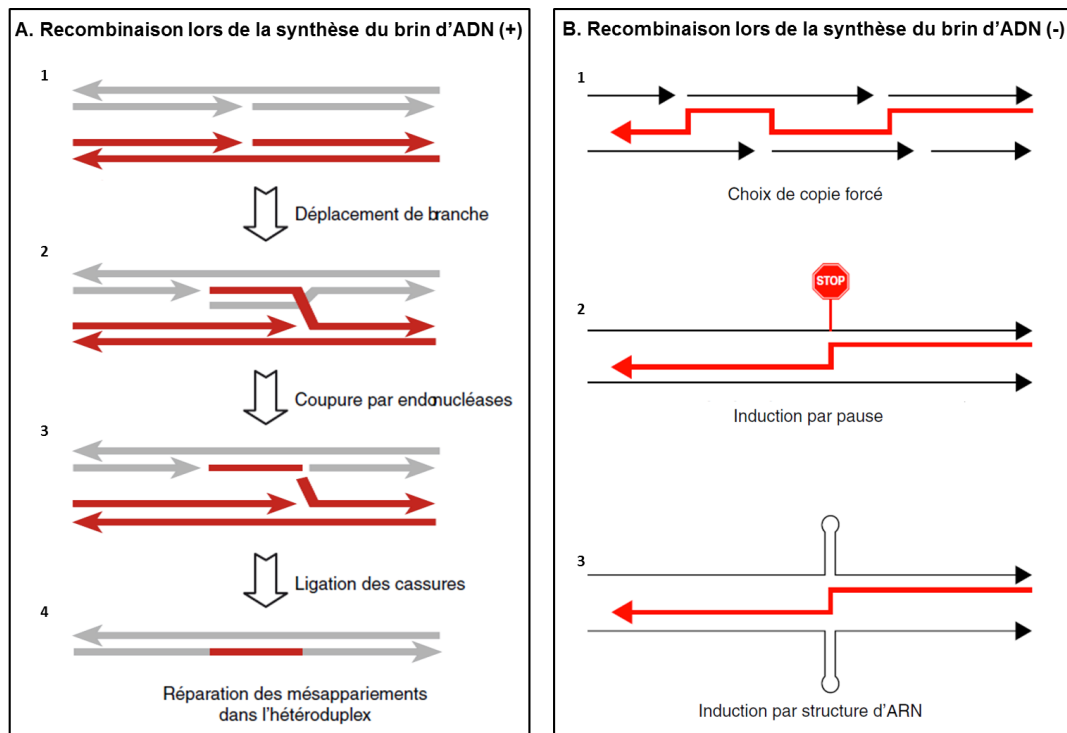


Figure 13 : Choix de copie ou Assimilation de branches

D'après Simon-Lorière et al, 2006 (37)

Le modèle d'assimilation de branche, représenté en (A) indique que la recombinaison a lieu lors de la synthèse du brin d'ADN (+) tandis qu'en (B), sont représentés les différents modèles de choix de copie, impliquant une recombinaison génétique lors de la synthèse d'ADN (-).

### c. Localisation et fréquence des points de cassure

La recombinaison n'a pas lieu partout au sein du génome et on estime que les points de cassure peuvent être prédits en fonction des régions du génome viral du VIH (122, 123). Le positionnement de ces points de cassure au sein du génome n'est donc pas aléatoire et des « points chauds » de la recombinaison ont été décrits, correspondant à des régions génomiques où la recombinaison a lieu de façon préférentielle. A l'inverse des « points froids », correspondant aux régions du génome où l'on observe peu ou pas de recombinaison, ont également été décrits.

Premièrement, le positionnement de ces points de cassure est contraint en lien avec la fonctionnalité des gènes. En effet, la sélection de formes viables à la suite d'un changement au sein du génome a une grande importance pour faire perdurer l'infection. A l'échelle du génome, une étude menée par Archer et al, portant sur l'identification des points chauds de la recombinaison, a montré que les points de cassure étaient majoritairement localisés aux extrémités des gènes (Figure 14) (122). Cette étude a été réalisée à l'aide d'un modèle probabiliste, construit à partir de 162 séquences de formes recombinantes intra-sous-types du VIH-1/M générées *in vitro* et est basé sur la similarité de



leurs séquences parentales décrites sur la base de données du VIH Los Alamos. D'autres séquences ont été utilisées en parallèle, incluant de nombreuses formes recombinantes pour comparer la répartition théorique des points de cassure à celle observées chez ces formes recombinantes *in vivo*. Finalement, ces données ont montré qu'il y avait une bonne corrélation entre le modèle de prédiction et les points de cassure identifiés *in vivo*. Néanmoins, une surreprésentation des points de cassure a été identifiée aux extrémités de l'enveloppe, tandis qu'une sous-représentation était retrouvée aux sein des gènes *gag*, *pol* et *env*. Ces données suggéraient que la recombinaison ne pouvait pas avoir lieu aléatoirement le long du génome, du fait d'une pression de sélection et qu'il serait plus avantageux pour le VIH de transmettre une enveloppe complète, lorsque ce gène est impliqué dans le phénomène de recombinaison.

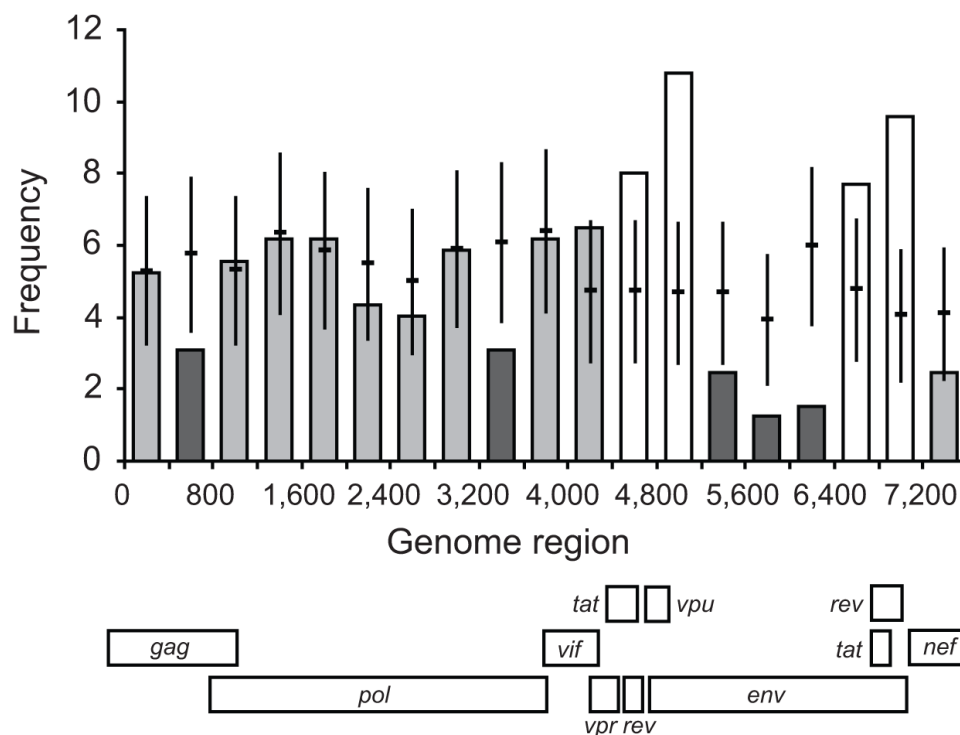


Figure 14 : Répartition des points de cassure au sein du VIH-1

D'après Archer et al, 2008 (122).

Comparaison des points de cassure retrouvés le long du génome selon un modèle probabiliste. Les lignes horizontales indiquent les fréquences prédites par le modèle. Les barres blanches indiquent que le nombre de points de cassure identifiés était significativement plus élevé que la prédiction pour la région, les barres gris-clair indiquent que les données identifiées étaient bien prédites, tandis que les barres gris-foncées indiquent qu'un nombre de points de cassure significativement plus bas que la prédiction du modèle était retrouvé. Sous le graphique sont représentées les régions génomiques correspondantes du VIH-1.

Une autre étude a pu confirmer que les points de cassure au sein du gène *env* étaient surtout localisés au niveau des extrémités des domaines, mais également dans des régions conservées pour ne pas impacter la fonction ou la structure de l'enveloppe (124). La présence d'autres points de

cassures dans *env* avait un effet délétère pour la souche correspondante. A l'échelle du génome, cette étude a également montré, comme pour *Archer et al*, que la fréquence de ces points de cassure était largement inférieure au sein de *env* comparativement aux autres gènes. Ces données montrent que la recombinaison génétique est moins fréquente au sein même de l'enveloppe, en lien avec une plus forte pression de sélection pour ce gène, voire une perte de viabilité selon les clones testés. En effet, ce gène est impliqué dans l'échappement au système immunitaire et permet l'entrée du virus dans la cellule, il est donc soumis à de nombreuses pressions de sélection *in vivo* et régule en partie la capacité répliquative du virus. Ces études ont ainsi montré que le phénomène de recombinaison génétique dépendait en partie de contraintes fonctionnelles et structurelles. Ces données ont d'ailleurs été confirmées dans une étude plus récente réalisée en 2016, basée sur l'analyse de séquences de différentes CRF et URF identifiées en Afrique, Asie et Europe (125).

Deuxièmement, la fréquence de la recombinaison dépend également de l'augmentation du pourcentage d'homologie entre les séquences parentales. Lorsque ce pourcentage d'homologie est grand, il entraîne une augmentation de la fréquence de recombinaison. Plusieurs études, dont celle d'*Archer et al*, ont permis de localiser le phénomène le long du génome par ces prédictions basées sur l'homologie des séquences parentales (122, 126). La recombinaison génétique a ainsi été très décrite au sein des VIH-1 intra-groupe M, dont la fréquence de recombinaison est très importante et constitue un facteur très important pour la diversification de la pandémie actuelle. Comme dit précédemment, environ 18% des infections impliquent ces formes recombinantes au niveau mondial. Contrairement à ce qui a été décrit pour les VIH-1 intra-groupe M, possédant une grande homologie de séquence, aucune recombinaison génétique n'a été mise en évidence pour le moment entre VIH-1 et VIH-2, malgré un nombre important de doubles infections VIH-1+2 (79, 126). Les VIH-1 et VIH-2 possèdent une faible identité de séquences, d'environ 50% dans le gène de l'*env*. L'absence de recombinaison entre ces deux types peut s'expliquer également par la faible prévalence et la faible transmissibilité du VIH-2 par rapport au VIH-1 (127). Il a été montré que le phénomène de recombinaison dépendait également de la longueur de cette homologie entre deux fragments parentaux. Une étude a montré que plus la longueur de l'homologie était conséquente, plus il y avait de chances pour que le phénomène de recombinaison intervienne (128). Aucune augmentation du taux de recombinaison n'était observée avec 20 pb d'homologie, contrairement à une longueur d'homologie de 40 à 60 pb. Au-delà de cette longueur, entre 60 pb jusqu'à 830 pb, une diminution du taux de recombinaison était observée. Les auteurs de cette publication ont conclu qu'une certaine longueur d'homologie était nécessaire pour assurer le maintien de la structure du gène impliqué dans la recombinaison et que ce phénomène était défavorisé si cette homologie était localisée entre deux séquences trop divergentes.

En 2009, une équipe s'est intéressée à plusieurs paramètres concernant la localisation de ces zones de recombinaison (129). A partir de la banque de donnée de Los Alamos HIV Database, différents génomes complets de VIH ont été analysés pour étudier la taille des régions recombinantes, la divergence entre fragments recombinants, le nombre de points de recombinaison. Cette étude a permis de mettre en évidence une grande variété de fragments recombinants, allant de 100 à 6000 pb et montrant que la taille de l'homologie de séquence était très différente en fonction des régions étudiées.

Enfin, il a été démontré que la structure de l'ARN viral avait également un impact sur la fréquence de recombinaison, avec un lien entre l'augmentation du taux de recombinaison et la présence de structures secondaires de l'ARN (130, 131). Par exemple, la présence d'une structure en épingle à cheveux au sein de la région C2 de l'enveloppe favoriserait la recombinaison. Ces données avaient d'ailleurs été visualisée par *Simon-Lorière et al*, et permettaient d'expliquer les résultats obtenus au cours de cette étude, précédemment décrite au moment de l'analyse de la répartition des points de cassure au sein du gène *env*.

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

Les deux premières formes recombinantes VIH-1/MO, ont été identifiées par Peeters et *al.* et Takehisa et *al.*, en 1999. Dans l'étude de Peeters et *al.*, une patiente camerounaise, asymptomatique, présentait une double réactivité VIH-1/M+O après sérotypage de la boucle V3 de l'enveloppe (132). Ces résultats ont été confirmés par des PCR (*Polymerase Chain reaction*) spécifiques de groupes M et O, au niveau des gènes *gag*, *pol*, *env* et des gènes accessoires de l'ADN proviral. Ces PCR ont été réalisés sur les isolats primaires de la patiente et après une étape de clonage biologique avec des cellules mononuclées du sang périphérique (CMSPs), par dilution limite des surnageants viraux. L'analyse des isolats primaires a révélé la présence d'une double réactivité VIH-1/M et VIH-1/O au niveau de l'enveloppe, mais seulement VIH-1/M pour les gènes *gag-pol*. Après culture, la présence unique d'un VIH-1/M au sein des gènes *gag* et *pol* était détectée, associée à la présence unique d'une *env* de VIH-1/O. Ces résultats, couplés à l'absence d'amplification de VIH-1/M dans les gènes accessoires, après dilution limite, montraient la présence d'une forme recombinante VIH-1/MO, présentant au moins un point de cassure au sein des gènes accessoires, associée à sa forme parentale VIH-1/M. Cette étude a également pu montrer que la forme recombinante VIH-1/MO identifiée était plus répliquative que sa forme parentale VIH-1/M, car plus la dilution était élevée moins la forme parentale VIH-1/M était

retrouvée dans la culture. Cette forme a ensuite été séquencée entièrement pour permettre une analyse phylogénétique et une analyse de la recombinaison, révélant la présence de deux points de cassure, un dans le gène *vpr* et l'autre au niveau des LTRs (Figure 15).

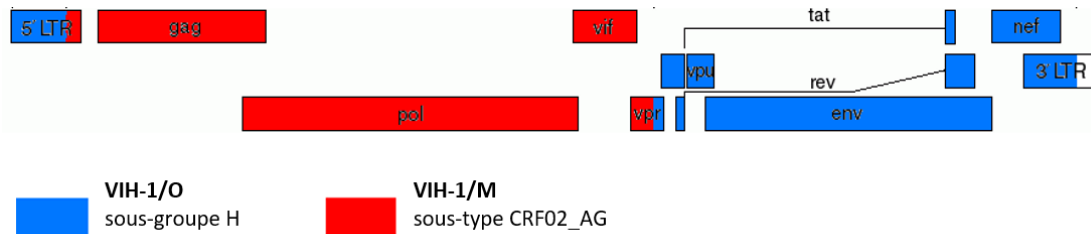


Figure 15 : Schéma de la forme recombinante 97CA.MP645MO

D'après Peeters *et al.* 1999 (132).

La même année, une triple infection par deux souches de VIH-1/M différentes associées à un VIH-1/O a été décrite par Takehisa *et al.*, chez une seconde patiente camerounaise symptomatique. (133). L'évolution de différentes populations virales a été suivie grâce à l'étude de différents prélèvements sanguins pour cette patiente, datant de 1994, 1995 et 1996. En 1994, l'amplification a été réalisée avec succès pour les gènes *pol* et *env* des trois souches de VIH, montrant la présence de fragments de VIH-1/M de deux sous-types différents (CRF02\_AG et D) et d'un VIH-1/O. En revanche, malgré l'amplification du gène *env* pour ces souches en 1995 et 1996, seul le VIH-1/O dans le gène *pol* était amplifiable. Les différents fragments obtenus entre le gène *pol* et le gène *env*, ont été clonés à partir de XL-PCR, pour obtenir les profils correspondants. L'analyse de ces clones a permis de montrer la présence d'une forme recombinante VIH-1/MO possédant des fragments de VIH-1/O, initialement présente et des fragments de VIH-1/M, déjà recombinant entre deux sous-types M différents (D et CRF02\_AG). Un point de cassure a été identifié au sein du gène *vpr* (Figure 16).

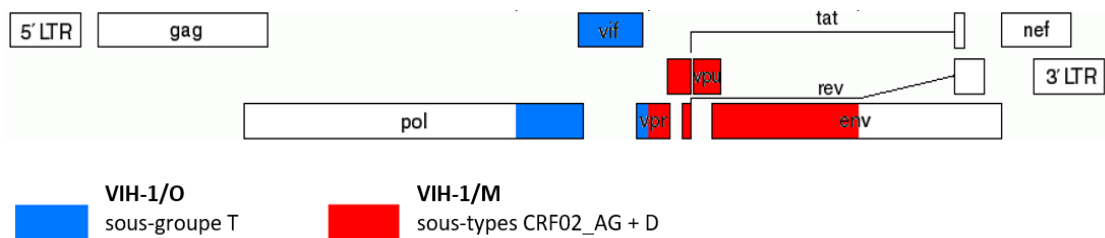


Figure 16 : Schéma de la forme recombinante 95CM61.56

D'après Takehisa *et al.* 1999 (133).

En 2004, une troisième forme recombinante VIH-1/MO a été identifiée par Yamaguchi *et al* après analyse d'un prélèvement de 1999 d'un patient camerounais, présentant une double réactivité sérologique VIH-1/M+O (sérotypage) de la boucle V3 de la GP120. La présence d'une forme recombinante a été confirmée par la réalisation de PCR spécifiques des groupes M et O, au niveau des gènes *pol* et *env* (107). Cette forme recombinante était associée à sa forme parentale VIH-1/O et présentait un point de cassure à la fin du gène *pol*, au sein de l'IN (Figure 17).

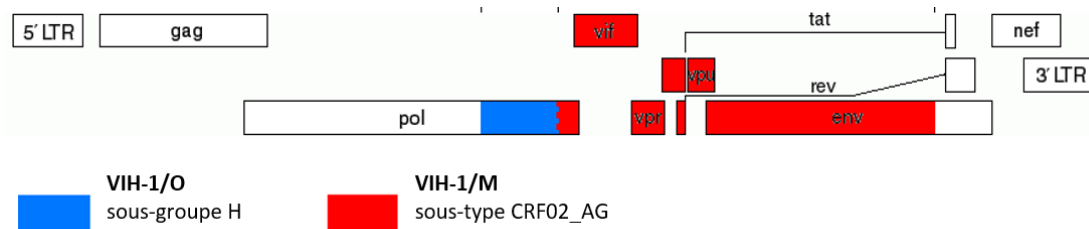


Figure 17 : Schéma de la forme recombinante DSC1320

D'après Yamaguchi *et al.* 2004 (107).

Ces trois premières formes recombinantes VIH-1/MO étaient associées au moins à l'une de leurs formes parentales VIH-1/M ou VIH-1/O, mais rien ne permettait d'indiquer si elles avaient été transmises ou si elles avaient émergé chez un patient doublement infecté M+O. Leur identification a toutefois permis d'affirmer que la recombinaison entre les VIH-1/M et VIH-1/O existait malgré la divergence génétique importante entre ces deux groupes. Ces études ont permis la description de trois formes recombinantes VIH-1/MO, dont une seule avait été entièrement séquencée par Peeters *et al*, mais ne permettaient pas d'analyser la prévalence des formes recombinantes VIH-1/MO, ni d'étudier leurs caractéristiques virologiques plus en détail. L'analyse de ces génomes avait néanmoins permis de suspecter l'existence d'un point chaud de la recombinaison dans le gène *vpr*, puisqu'un point de cassure y était retrouvé pour deux des trois formes VIH-1/MO (Figures 15, 16 et 17) (107, 132, 133).

Dans ce contexte, une première étude a été réalisée au Cameroun entre 2006 et 2009, par Vessière *et al*, afin de mettre en place un algorithme de dépistage des formes recombinantes VIH-1/MO, incluant un point de cassure dans le gène *vpr*, chez des patients présentant une double réactivité M+O au sérotypage. La mise en place de cet algorithme de dépistage a été réalisée au Centre Pasteur au Cameroun (CPC) et reposait sur des outils sérologiques et moléculaires spécifiques de groupe M et O (105, 134). Premièrement, un sérotypage des souches de VIH-1 est effectué avec l'aide d'anticorps spécifiques de la boucle V3 de la GP120, permettant d'identifier le type et le groupe de VIH. Afin de confirmer les éventuelles doubles infections mise en évidence par ce sérotypage, une seconde analyse moléculaire est réalisée, à l'aide de PCR spécifiques des VIH-1/M et VIH-1/O, au

niveau de la PR, RT, IN et GP41. Enfin, des techniques de CV spécifiques de groupe M et O ont également été mises en place, afin d'estimer le niveau de répllication des différents virus. Une fois la double infection mise en évidence, des PCR ont été réalisées afin de déterminer la présence de formes recombinantes VIH-1/MO associées, à partir de la recherche d'un point de cassure dans le gène *vpr*. Ces PCR étaient dites hétérologues, lorsqu'elles permettaient d'amplifier spécifiquement les gènes *vpr* recombinants de motifs OM et MO et, homologues, lorsqu'elles permettaient d'amplifier spécifiquement les gènes *vpr* non recombinants de motifs MM et OO. La mise en place de cet algorithme au cours de cette étude, a permis d'identifier quatre doubles infections sans recombinants associés et six formes recombinantes VIH-1/MO, dont trois étaient associées à une double infection M+O et deux autres à une seule forme parentale. Sur ces six formes, cinq présentaient un point de cassure dans le gène *vpr* et une autre, dans le gène *vpu*. C'est au cours de cette étude qu'une première forme recombinante VIH-1/MO sans formes parentales associées a été identifiée chez une patiente camerounaise vivant en France (135). Cette patiente possédait une CV indétectable malgré l'absence de traitement et une immunodépression sévère de 6 CD4/mm<sup>3</sup>. Après envoi des échantillons au Centre National de Référence du VIH (CNR) au laboratoire de virologie du Centre Hospitalier Universitaire (CHU) de Rouen, cette étude a confirmé la présence d'une forme recombinante avec un point de cassure dans *vpr*. Aucune forme parentale associée n'a été mise en évidence, ce qui suggérait soit une transmission directe de la forme recombinante à la patiente, soit un avantage répllicatif de la forme recombinante par rapport à ses formes parentales.

Les données obtenues au cours de cette première étude ont montré la part importante des formes recombinantes associées à au moins une forme parentale. Ces résultats ont été obtenus sur la base d'une stratégie basée sur la détection des formes recombinantes ayant un point de cassure dans le gène *vpr*, dans des prélèvements doublement réactifs au sérotypage. Cette stratégie a été remise en cause d'une part, par la détection de formes recombinantes associées à une seule deux formes parentales, montrant qu'il fallait explorer également des prélèvements mono-réactifs et d'autre part, par la détection d'une forme recombinante possédant un point de cassure hors *vpr*, dans le gène *vpu*. Cette détection basée sur l'hypothèse d'un point chaud dans *vpr*, ne permettait pas la détection d'autres points de cassure, ou des formes recombinantes plus complexes. De plus, une de ces formes a été identifiée hors du Cameroun, en France, montrait leur potentiel de diffusion et suggérait la possibilité d'une transmission directe des formes recombinantes VIH-1/MO. Enfin, la caractérisation des génomes complets n'avait pas été réalisée au moment de l'étude pour les différentes formes recombinantes identifiées, empêchant l'analyse plus approfondie de ce phénomène de recombinaison.

Fort de ce constat, une seconde étude a été menée au Cameroun entre 2013 et 2015 par *Ngoupo et al.* Cette étude avait pour objectif de compléter la première étude menée par *Vessière et al.*, en étudiant la fréquence des doubles infections VIH-1/M+O dans différentes régions du Cameroun, en incluant les données issues de prélèvements mono-réactifs au sérotypage et en caractérisant les formes recombinantes VIH-1/MO identifiées pour mettre en évidence la circulation d'une éventuelle CRF\_MO. Le même algorithme de dépistage décrit ci-dessus et mis en place par le CPC a été utilisé, avec ajout d'une étape de séquençage des génomes complets (Figure 18). Finalement, cinq doubles infections ont été identifiées, dont quatre sans formes recombinantes VIH-1/MO associées et une, montrant la co-infection d'une forme recombinante ayant un point de cassure dans *vpu*, avec sa forme parentale VIH-1/M. Une même forme recombinante VIH-1/MO, avec point de cassure dans *vpr*, a été identifiée au sein d'un couple Camerounais (REC003/REC024), suggérant la possibilité d'une transmission horizontale (106, 136). Les trois formes recombinantes identifiées ont été séquencées entièrement, afin d'effectuer une comparaison entre elles mais aussi avec les autres formes recombinantes précédemment décrites. La caractérisation moléculaire et phylogénétique de ces formes n'a pas mis en évidence de lien phylogénétique, ce qui ne permettait pas de conclure sur l'existence d'une CRF\_MO. Toutefois, l'analyse effectuée a montré que la transmission directe d'une même forme recombinante avait été possible au sein du couple REC003/REC024. Ces données confirmaient le potentiel de diffusion suggéré par la précédente étude.

Du fait de l'utilisation du même algorithme de dépistage du CPC, cette étude a été confrontée à des limites méthodologiques. L'identification des formes recombinantes selon le positionnement d'un point de cassure dans *vpr* ne favorisait pas l'identification de formes plus complexe, possédant un point de cassure en dehors de cette région. La caractérisation des génomes complets a montré que d'autres points de cassure devaient être explorés le long du génome, notamment au niveau des gènes accessoires centraux et dans les LTRs. Enfin, la mise en évidence de la transmission directe de l'une de ces formes, couplé à l'identification d'une forme recombinante isolée en France, montrait l'importance de suivre l'émergence d'une CRF\_MO, puisque malgré l'identification de rares cas, ces formes avaient diffusé dans la population, dans différentes régions du Cameroun et hors de ce pays.

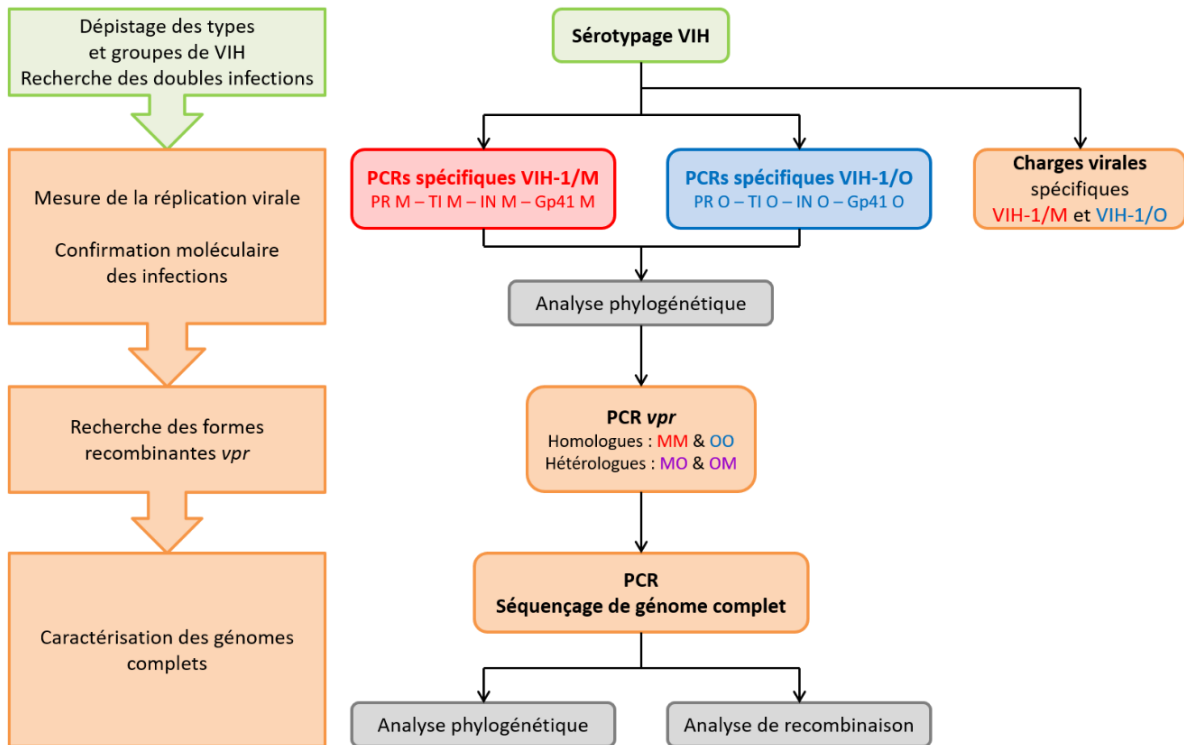


Figure 18 : Algorithme de dépistage des formes recombinantes *vpr*

Issu de Moisan et al, 2019 (137).

Schéma présentant l'algorithme de dépistage utilisé par le CPC pour l'identification des formes recombinantes MO, présentant un point de cassure dans le gène *vpr*.

En plus des trois premières formes identifiées en 1999 et 2004, ces deux études ont permis de décrire neuf nouvelles formes recombinantes, dans différents contextes, montrant le potentiel de diffusion de ces formes et la nécessité d'identifier une CRF\_MO. De plus, ces deux études ont montré que les formes recombinantes VIH-1/MO pouvaient posséder d'autres points de cassure, hors du gène *vpr*, au niveau de *vpu* et des LTRs, montrant l'intérêt de caractériser ces formes plus en détail afin d'identifier l'ensemble des points de cassure le long du génome.

Dans le cadre de la surveillance des nouvelles formes émergentes du VIH en France, une dernière étude a été publiée en 2018 par le CNR du CHU de Rouen. Plusieurs échantillons envoyés pour expertise suite à la mise en évidence d'une discordance soit lors du sérotypage, pour identifier une éventuelle infection atypique (VIH-1/O ou co-infection VIH-1/M+O), soit lors de l'observation d'une CV indétectable chez un patient séropositif non traité ou soit lors du génotypage de résistance pour le suivi du patient. Suite à l'envoi de ces prélèvements au CNR, une stratégie basée sur l'algorithme de dépistage des formes recombinantes du CPC a été entreprise, avec la réalisation de PCR spécifiques de groupes M et O dans les mêmes régions ciblées PR, RT, IN et GP41. Cette étude a permis de décrire un cas de double infection VIH-1/M+O sans recombinant *vpr* et 10 nouvelles formes recombinantes VIH-1/MO, dont trois co-infections impliquant une forme recombinante VIH-1/MO associée à l'une ou ses



deux souches parentales. L'analyse phylogénétique a montré que sur ces 10 formes recombinantes VIH-1/MO, la même forme recombinante était identifiée chez deux paires de patients (RBF222/RBF243 et BCF174/BCF212) ne présentant pas de liens épidémiologiques connus. Enfin, une co-infection par deux formes recombinantes VIH-1/MO chez un même patient (RBF221) a également été décrite. Toutes ces formes ont été identifiées chez des patients Camerounais, ou ayant un lien avec le Cameroun (110). La caractérisation moléculaire et phylogénétique de ces formes n'a pas mis en évidence de CRF\_MO à la suite du séquençage des génomes complets. Toutefois, cette étude a permis l'identification d'une même forme VIH-1/MO chez deux patients non liés épidémiologiquement et à deux reprises, montrant le potentiel de diffusion de ces formes et l'importance de la surveillance de l'émergence d'une forme circulante.

Au total, depuis 1999, ce sont 23 formes recombinantes VIH-1/MO qui ont été décrites chez 22 patients (Tableau 2).

Tableau 2 : Découverte des 23 formes recombinantes VIH-1/MO

<b>Etude, année de publication</b>	<b>Pays Période d'étude</b>	<b>Recombinants VIH-1/MO décrits</b>	<b>Code souche</b>
Peeters, 1999 (132)	Cameroun 1999	1	97CA.MP645MO
Takehisa, 1999 (133)	Cameroun 1994-1996	1	95CM61.56
Yamaguchi, 2004 (107)	Cameroun 1999	1	DSC1320
Vessièrè, 2010 (135)	France 2003-2008	1	RBF208
De Oliveira, 2017 (105)	Cameroun 2006-2009	6	YBF205, YBF2011, YBF212, YBF274, YBF282, YBF298
Ngoupo, 2016 (106, 136)	Cameroun 2013-2015	3	REC003, REC024, REC107
De Oliveira, 2018 (110)	France 2004-2016	10	BCF204, BCF174, BCF212, RBF221-1, RBF221-2, RBF222, RBF235, RBF237, RBF240, RBF243

D'après Moisan et al, 2019 (137).

### 3. Caractéristiques moléculaires des VIH-1/MO

Sur les 23 formes recombinantes VIH-1/MO identifiées, 15 ont été complètement séquencées. Les huit autres ne pouvaient pas l'être, du fait de la présence de formes parentales associées ou d'un volume insuffisant de prélèvement primaire (Figure 19) (138).

La caractérisation de ces 23 formes recombinantes a été entreprise au cours d'une précédente étude, afin de déterminer le nombre précis de formes recombinantes uniques VIH-1/MO (URF\_MO) et de mettre en évidence une éventuelle CRF\_MO (138). Cette étude a permis d'étudier les liens phylogénétiques existants entre ces formes. Une même forme recombinante a été identifiée chez deux patients présentant un lien épidémiologique direct (REC003/REC024) (136). L'identification d'une même forme recombinante chez deux patients non liés épidémiologiquement a également été observée, à deux reprises, chez RBF222/RBF243 et BCF174/BCF212 (110). Néanmoins, la description de seulement deux cas de transmission pour chacune de ces deux formes ne permettait pas de conclure à l'émergence d'une CRF\_MO cette hypothèse. Aucun lien phylogénétique n'a été retrouvé entre les autres formes recombinantes. Finalement, cette étude a permis de montrer qu'il existait 19 URF\_MO différentes, dont 12 avaient été complètement séquencées (Figure 19).

L'analyse des profils moléculaires des 19 URF\_MO au cours de cette précédente étude, a permis de mettre en évidence la diversité génétique des fragments VIH-1/M et VIH-1/O impliqués de part et d'autre des points de cassure identifiés (138). Ainsi, on trouvait pour le VIH-1/M, 52,6% de fragments de sous-types CRF02\_AG (10/19), 15,8% de D (3/19), 10,5% de fragments A et D chacun (2/19 chacun), 5,3% de CRF11\_cpx et 5,3% de forme complexe recombinante intra sous-type CRF02\_AG+D (1/19 chacun) (Figure 20). Concernant les fragments de VIH-1/O, la majorité d'entre eux, soit 84,2% des séquences, appartenait au sous-groupe H (16/19), contre 15,8% qui appartenaient au sous-groupe T (3/19) (Figure 20).

Ces chiffres reflétaient l'épidémiologie moléculaire décrite au Cameroun. En effet, deux études ont montré que 50 à 58% des séquences de VIH-1/M étaient de sous-type CRF02\_AG, 4,5% à 11% de de sous-type G, 3 à 4% de de sous-type D, 2% à 3,3% de de sous-type A et 2% à 4,7% de de sous-type F (139, 140). Quant aux VIH-1/O, une autre étude réalisée dans ce pays a montré que environ 77% et 23% des souches circulantes appartenaient aux sous-groupes H et T, respectivement (59).



Figure 19 : Profil et diversité génétique des 19 URF\_MO

D'après Moisan et al, 2016 (138).

Sur la partie haute figurent les génomes complets séquencés, en dessous, les génomes partiels. Le groupe O est représenté en bleu indiquant le sous-groupe correspondant, tandis que le groupe M est en rouge avec le sous-type correspondant. L'identifiant de l'URF\_MO est indiqué à droite du schéma de chaque forme recombinante.

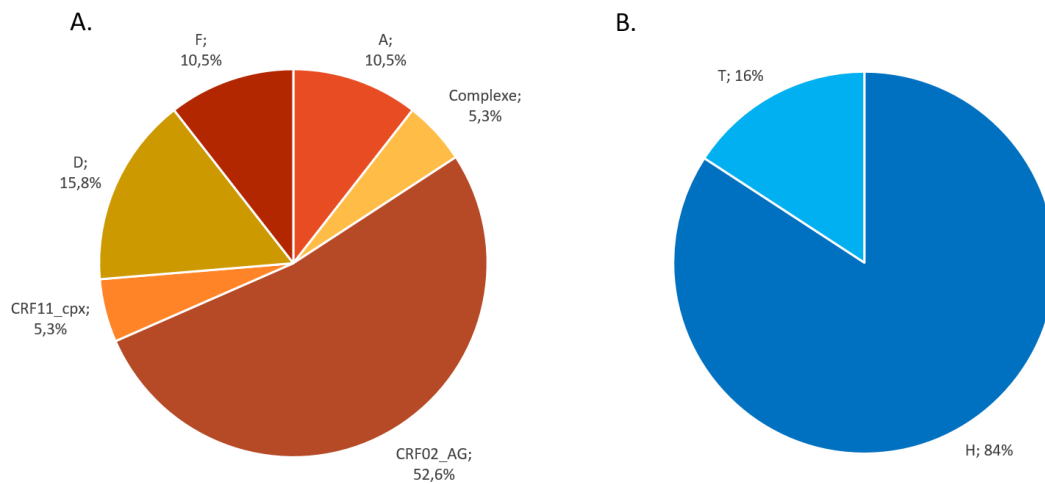


Figure 20 : Diversité génétique de fragments parentaux identifiés au sein des 19 URF\_MO

- A) Diversité des fragments de VIH-1/M retrouvé au sein des 19\_URF\_MO
- B) Diversité des fragments de VIH-1/O retrouvé au sein des 19\_URF\_MO

Enfin, l'analyse de la répartition de tous les points de cassure des 19 URF\_MO a montré que 74,4% étaient localisés entre les gènes *gag* et *nef* ( $n=32/43$ ) (138). Parmi ces points de cassure, la majorité d'entre eux, soit 59,4%, était localisée au sein des gènes accessoires centraux (*vif*, *vpr* et *vpu*) ( $n=19/32$ ) (138) dont 31,2% au sein du gène *vpr* ( $n=10/32$ ), confirmant que ce gène est bien un point chaud de la recombinaison inter-groupes MO.

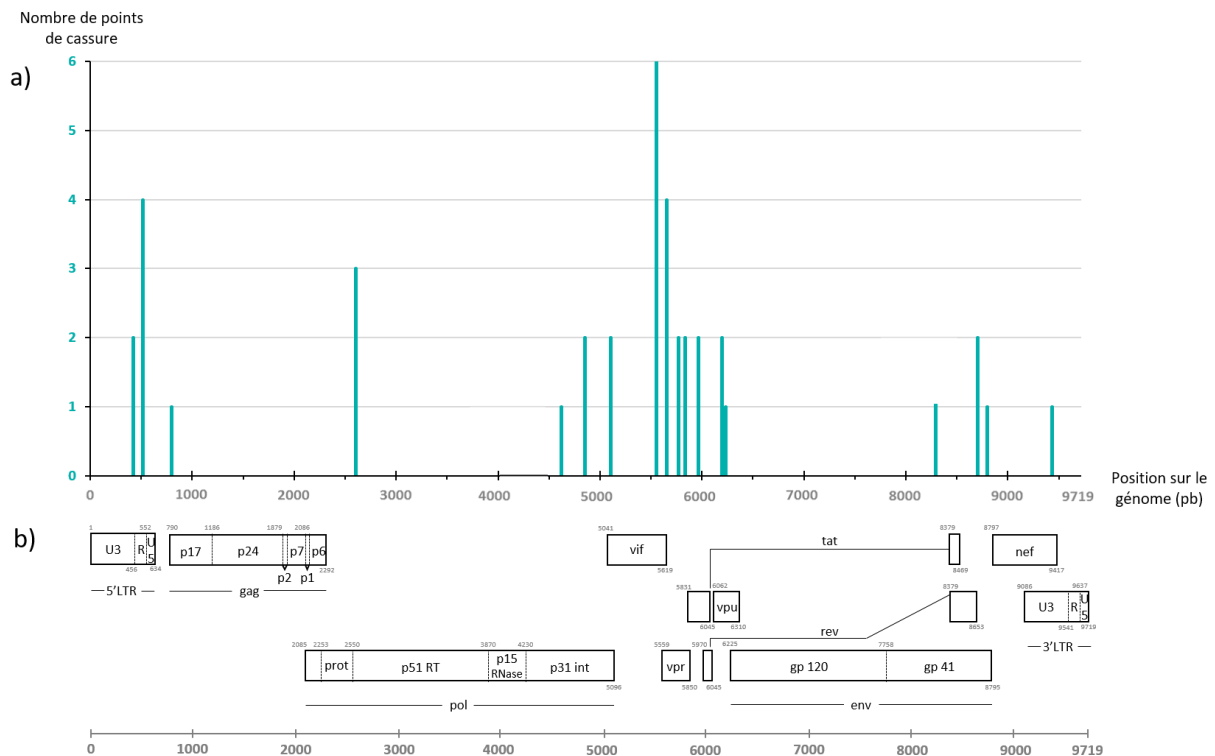
En étudiant la localisation de ces points de cassure plus finement entre *gag* et *nef*, cette précédente étude a également pu montrer que 72% d'entre eux ( $n=22/32$ ) étaient localisés aux extrémités des gènes, contre 19% ( $n=6/32$ ) dans une région chevauchante, partagée par deux gènes et 16% ( $n=4/32$ ) à l'intérieur d'un gène (Tableau 3). Peu, voire aucun, point de cassure n'a été identifié dans *gag*, PR, GP120 et *nef*, suggérant que ces régions constituent des points froids de la recombinaison inter-groupes MO (138). Ces données étaient en accord avec ce qui était décrit dans la littérature pour les formes recombinantes intra VIH-1/M (122).

Tableau 3 : Répartition des points de cassure entre les gènes *gag* et *nef* du génome des 19 URF\_MO

<i>gag</i>		<i>pol</i>		gènes accessoires						<i>env</i>		Total																												
p17	p24	PR	TI	IN	<i>vif</i>	<i>vpr</i>	<i>tat</i>	<i>rev</i>	<i>vpu</i>	GP120	GP41		<i>nef</i>																											
1	0	0	3	3	6	10	1	1	3	0	4	0	32																											
C	E	I	C	E	I	C	E	I	C	E	I	C	E	I	C	E	I																							
0	1	0	0	0	0	0	0	0	3	0	0	0	3	0	6	0	3	7	0	0	1	0	1	0	0	1	2	0	0	0	0	1	2	1	0	0	0	6	22	4

C : Point de cassure retrouvé à l'extrémité chevauchante d'un gène  
 E : Point de cassure retrouvé à l'extrémité d'un gène  
 I : Point de cassure retrouvé à l'intérieur d'un gène

Enfin, d'autres points de cassure ont été décrits hors de la région *gag - nef*, au sein des LTRs, représentant 26% des points de cassure totaux (n=11/43). Une étude menée en 2017 a permis de montrer l'apparition d'un point de cassure obligatoire au sein de ces LTRs, lorsqu'un nombre impair de point de cassure était présent, entre *gag* et *nef* (141). Cette étude a été complétée par notre équipe, qui a permis de démontrer que ces points de cassure étaient majoritairement positionnés au sein de la région R des LTRs aux positions 506-511, 513-520 and 523-546 pb (*en cours de soumission*). Tous les points de cassure identifiés au sein des 19 URF\_MO, ont été schématisés en Figure 21.



**Figure 21** : Localisation des points de cassures issus des 19 URF\_MO

Le nombre de points de cassure identifiés au sein des 19 URF\_MO a été indiqué en fonction de leur position sur le génome, en paires de bases selon la souche de référence HxB2 (n° d'accèsion GenBank K03455).

## 4. Conséquences de la recombinaison génétique

Le phénomène de recombinaison est très répandu chez les virus à ARN, ayant un impact sur la diversification, l'évolution et le suivi des VIH.

Sur le plan évolutif, la variabilité induite par la recombinaison entre deux séquences peut entraîner des modifications plus ou moins bénéfiques pour le virus. La plupart des mutations retrouvées au sein de la quasi-espèce n'affectent pas l'infectiosité, voire rendent le virus moins infectieux ou défectif. Cependant, la recombinaison permet au virus d'accumuler rapidement un grand nombre de mutations et contribue donc grandement à la diversification virale. En fonction de la

position des points de cassure le long du génome, ce phénomène peut altérer la fonction d'un gène important, comme c'est le cas par exemple, pour le gène de l'enveloppe (142). En revanche, d'autres gènes sont moins impactés par la survenue de mutations, comme c'est le cas de la TI (143). Si un gène viral jouant un rôle clé au cours du cycle de réplication, dans le système d'échappement au système immunitaire, ou contre un ARV est altéré, cela peut affecter la capacité du virus à faire face à un environnement donné (144, 145). Ce constat est également valable si ce changement concerne l'interaction avec un facteur de restriction ou la maturation du virion. Comme dit précédemment, la recombinaison ne peut donc pas avoir lieu n'importe où dans le génome, afin de préserver la structure d'un gène et/ou la fonction d'une protéine virale et les virions défectueux ou moins adaptés ne seront pas sélectionnés par les pressions de sélections immunitaire ou thérapeutique (146–149).

D'autres mutations peuvent quant à elles être bénéfiques pour le virus, en le rendant plus infectieux (130, 150). Ces changements liés à la recombinaison peuvent également permettre de diversifier les populations virales intra-hôte en faisant émerger un variant plus adapté que celui/ceux ayant infecté la cellule à l'origine (Figure 22) (130, 145). Ce phénomène de recombinaison permet également de supprimer les mutations délétères, survenue au cours de la diversification du virus, puisque des erreurs sont introduites à chaque cycle de réplication. Enfin, la recombinaison peut permettre de rétablir une fonction altérée par des phénomènes d'oxydation de base, en transformant un virion initialement défectif en virion viable et fonctionnel (Figure 22) (151, 152).

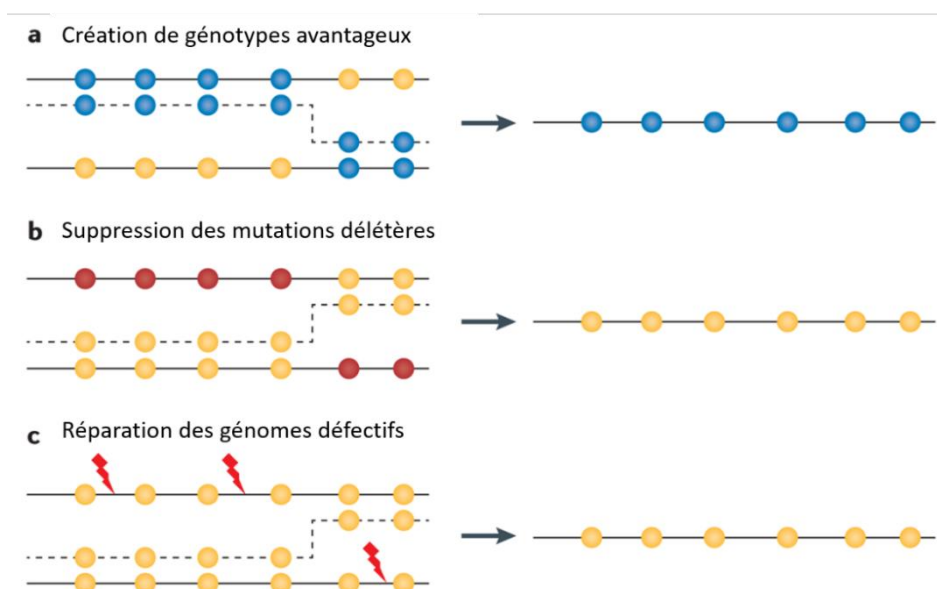


Figure 22 : Conséquences évolutives de la recombinaison génétique

Traduit d'après Simon Loriaire et Holmes, 2011 (130).

Les cercles jaunes indiquent un génotype de type sauvage. En fonction des génotypes donneurs et accepteurs impliqués, de la position du point de cassure et des segments génomiques échangés, le phénomène de recombinaison peut avoir plusieurs conséquences : a) la création d'un avantage répliatif par l'acquisition d'une combinaison de mutations bénéfiques pour le virus (cercles bleus), b) la suppression de mutations délétères pour le virus (cercles rouges) pour restaurer le génotype sauvage, ou encore c) la réparation des génomes défectifs, présentant des cassures ou des modifications de bases (éclairs rouges).

Sur le plan diagnostique, l'adaptation des outils de dépistage au VIH-1/O permet maintenant de diagnostiquer la majorité des VIH en incluant la majorité des variants circulants. Ceci implique paradoxalement que les formes recombinantes ne sont plus mises en évidence au moment du diagnostic, puisque leur détection reposait jusqu'alors sur l'observation de discordances séro-moléculaires ou moléculaires. Désormais, le diagnostic d'une forme recombinante VIH-1/MO repose donc entièrement sur le séquençage de résistance, qui cible toujours plusieurs régions du génome de façon spécifique de groupes (M et O). Lorsque les groupes M et O sont retrouvés de manière concomitante lors du séquençage de résistance, la présence d'une double infection VIH-1/M+O ou d'une forme recombinante VIH-1/MO peut être suspectée.

Sur le plan thérapeutique, certaines études ont montré que la recombinaison intra-groupe M pouvait conférer une résistance aux traitements ARV, donc une meilleure résistance face aux pressions thérapeutiques (153, 154). L'utilisation d'un traitement non adapté pourrait ainsi conduire à la transmission et la diffusion de virus résistants dans la population. Ainsi, un échec virologique a été rapporté récemment chez une patiente infectée par une forme recombinante VIH-1/MO (RBF235), dans un contexte complexe de dysobservance et de diagnostic tardif d'une co-infection VIH-1/M+O ayant favorisé l'émergence d'un recombinant VIH-1/MO. L'étude des prélèvements séquentiels a montré que cet échec perdurait malgré l'adaptation du traitement à six reprises (111). Ces résultats sont toutefois à nuancer, puisque les autres formes recombinantes, étudiées au cours d'un autre travail (BCF204, RBF208, BCF212 et RBF243), ont montré une bonne réponse aux traitements mis en place, sans problème de prise en charge (CV indétectable et CD4 supérieurs à 200/mm<sup>3</sup>) (155). Pour le moment, aucun impact de ces formes recombinantes MO n'a été mis en évidence sur le plan thérapeutique, à l'exception des formes ayant une TI de groupe O, avec donc potentiellement une résistance aux INNTI (155).

## 5. Capacité répliquative et *fitness* viral

Le virus mute de façon constante afin de s'adapter à l'environnement dans lequel il évolue, sous l'effet de pression de sélection. Cette capacité du virus à s'adapter et à se reproduire dans un environnement précis, définit sa capacité répliquative, ou « *fitness* » répliquatif. La notion de *fitness* représente la capacité d'un virus, ou d'une population virale, à devenir prédominant(e) face à d'autres virus en culture, au sein d'un hôte, ou au sein d'un système donné et à s'adapter, se transmettre à la génération suivante (*fitness* adaptatif). Le *fitness* quantifie donc le potentiel d'une souche à évoluer, dans un environnement donné, face à une pression de sélection, qu'elle soit naturelle ou induite (144,

156–158). La capacité répliquative influence grandement le *fitness* viral, en déterminant le potentiel de la souche à se multiplier et produire des virions infectieux au cours du cycle de réplication. Plusieurs méthodes existent pour mesurer le *fitness* répliquatif, tel que l'étude des activités enzymatiques virales (PR ou TI) du virus, l'étude de la cinétique de réplication *in vitro* sur CMSPs, ou encore des essais de compétitions dans des cellules ou tissus (Figure 23) (158). Cette mesure de la capacité répliquative peut être effectuée par la mise en culture de plusieurs souches en parallèle (potentiel répliquatif), cette technique reste sera moins informative que la comparaison de ces souches au sein d'une même culture, pour conclure sur une réelle différence répliquative entre les souches (157). La mesure du *fitness* adaptatif est quant à elle plus complexe car nécessite des études de compétition entre plusieurs souches, parfois dans des systèmes complexes (157, 159, 160).

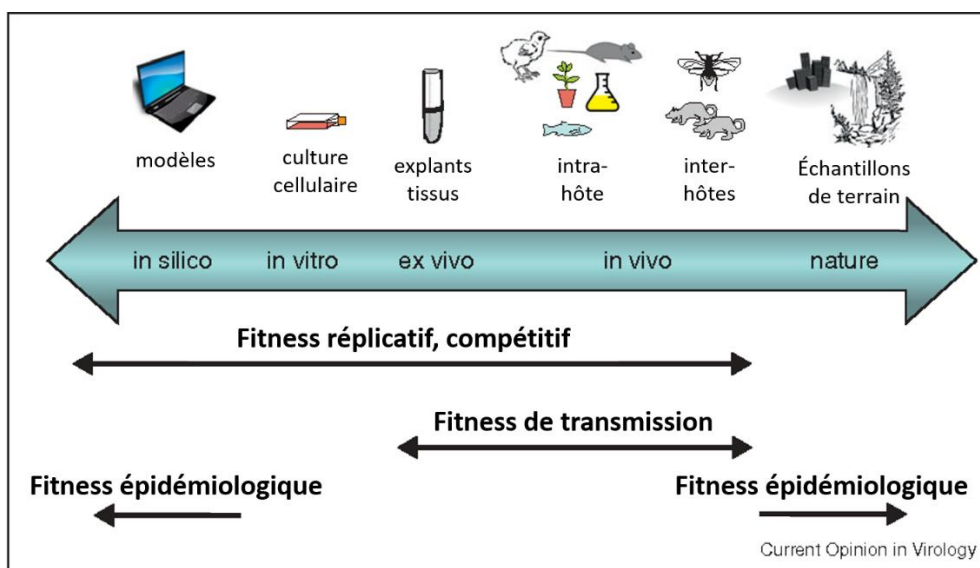


Figure 23 : Méthodes de détermination du *fitness* viral

Traduit d'après Wargo et Kurath, 2012 (144).

Comme vu précédemment, la recombinaison génétique impacte l'évolution et la dynamique des populations virales du VIH. En plus de ce phénomène de recombinaison, les propriétés intrinsèques du virus, incluant un fort taux de réplication couplé à la faible fidélité de la TI, implique la sélection de mutations qui peuvent conduire à un changement de *fitness*. Une corrélation entre augmentation de la diversité virale et augmentation du *fitness* a été retrouvée dans deux études, à partir de population clonales (161, 162). Malgré la description de nombreux points de cassure le long des génomes du VIH, il a été démontré que des contraintes fonctionnelles et structurales limitaient la recombinaison génétique. Ces contraintes ont été étudiées en regard du *fitness* viral, démontrant qu'il existait une interdépendance de la GP120 et GP41. Si des perturbations comme le phénomène de recombinaison entravaient la fonctionnalité du gène de l'enveloppe, alors une diminution du *fitness* était observée



(163, 164). Enfin, un avantage répliatif conféré par l'acquisition d'une mutation de résistance à un traitement, pourra avoir comme conséquence, un gain de *fitness* répliatif, puisqu'le virus résistant sera avantagé par rapport à la population virale initiale ne possédant pas cette résistance (165, 166).

Finalement, la capacité répliative et le *fitness* adaptatif font partie des nombreux processus impliqués dans la sélection des souches de VIH, notamment des formes recombinantes, puisqu'ils permettent la sélection des formes les plus adaptées ou ayant un avantage répliatif, par rapport aux formes parentales dont elles sont issues. Après génération des formes recombinantes par l'action de la TI, de nombreuses pressions de sélection, moteurs de l'évolution génétique, interviennent (Figure 24). Seules les formes infectieuses, non défectives et possédant des protéines et une structure viable pourront être sélectionnées au sein d'une population virale. Un avantage répliatif conféré par une mutation, ou par la recombinaison, peut permettre de faciliter l'émergence d'un virus au sein du patient. Une fois le virus répliatif et infectieux sélectionné, il fait encore face à de multiples pressions de sélection, à la fois exercées par l'hôte (système immunitaire, facteurs de restriction) et par la mise sous traitement ARV si elle a lieu (pression thérapeutique). La diffusion dans la population d'un variant en particulier du VIH, notamment des formes recombinantes, dépend donc grandement de la capacité répliative et du *fitness* adaptatif de ces dernières (Figure 24).

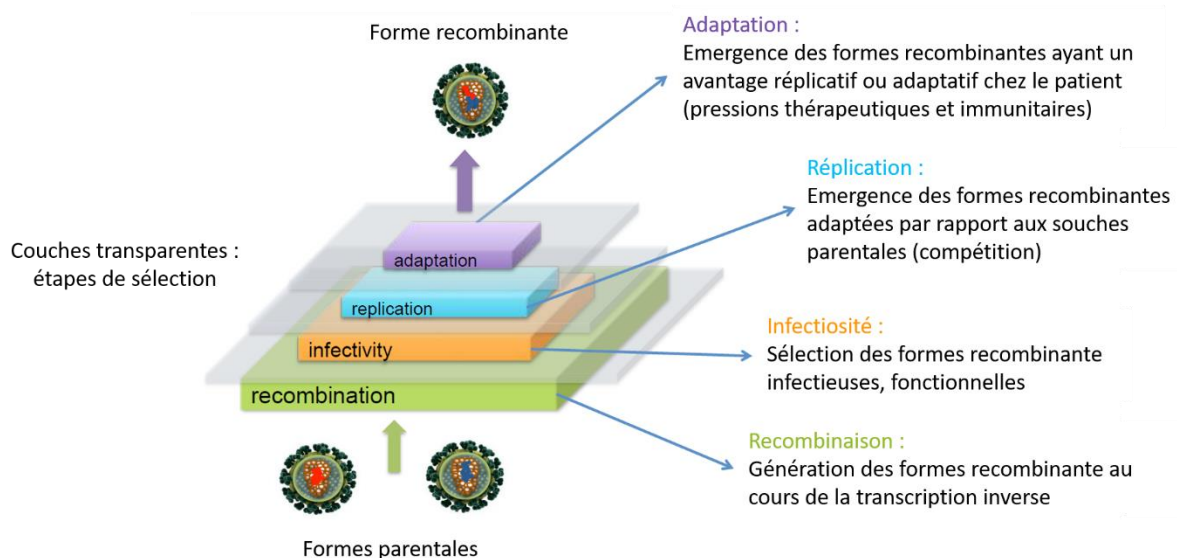


Figure 24 : Processus de génération et sélection des VIH-1/MO

Adapté d'après Cappy et al, 2016 (167).

Depuis sa découverte, le VIH-1 a largement diffusé dans la population et le VIH-1/M est devenu pandémique. Certaines souches sont néanmoins moins adaptées comme c'est le cas du VIH-2. Sa transmission est quatre à dix fois plus faible et entraîne des CV plasmatiques environ 30 fois inférieures à celle du VIH-1 (168, 169). L'étude publiée *Ariën et al* en 2005, a également montré cette différence entre VIH-1 et VIH-2 et avec les VIH-1/O, en établissant une « pyramide de *fitness* » : VIH-1/M > VIH-2 >> VIH-1/O (168). Cette étude a pu montrer, d'une part, que les VIH-1/O étaient 100 fois moins adaptés que tous les isolats de VIH-2 et de VIH-1/M testés et, d'autre part, que la plupart des VIH-2 testés était également moins adaptés que les VIH-1/M. En 1999, l'étude menée par *Peeters et al* avait pu montrer une capacité répliquative supérieure de la forme recombinante VIH-1/MO, après de multiples dilution en culture, donc suggérant un avantage répliquatif par rapport à sa forme parentale VIH-1/M (132). Ces deux études mises en parallèle suggéraient que les formes recombinantes VIH-1/MO pourrait avoir une capacité répliquative supérieure à leurs formes parentales. Cette hypothèse a été étudiée au cours d'une précédente étude, menée en 2019 par notre équipe, sur le potentiel répliquatif d'une forme recombinante *in vitro* par rapport aux deux souches parentales VIH-1/M et VIH-1/O (137). Les résultats de ces travaux ont montré que la souche recombinante produite *in vitro*, possédait un comportement intermédiaire, selon les marqueurs étudiés au cours de la cinétique de répliquative. Cette étude sera détaillée page 62, afin d'introduire les travaux effectués au cours de ce projet de doctorat.



## RATIONNEL ET OBJECTIFS

La mise en évidence de plusieurs formes recombinantes VIH-1/MO depuis 1999, de leur circulation possible en dehors du Cameroun (110, 135), ainsi que de leur transmissibilité au sein d'un couple (136) et une étude de leur réplication *in vitro* (132) suggèrent que ces formes pourraient posséder un avantage répliatif, par rapport aux souches parentales VIH-1/M et VIH-1/O dont elles proviennent.

La description des points de cassure au sein des 19 URF\_MO indique que le phénomène de recombinaison génétique n'intervient pas au hasard. Ces formes MO possèdent, certes, des profils différents mais conservent des similitudes du fait de l'existence de mêmes points de cassure, localisés aux mêmes endroits du génome. Parmi ces URFs, 53% des points de cassure étaient retrouvés le plus fréquemment au niveau au sein du gène *vpr* (n=10/19) (137). Parmi les 12 URF\_MO complètement séquencées, deux profils distincts ont été mis en évidence. Un premier profil a été mis en évidence pour 64% d'entre elles, avec un motif [O<sub>gag/pol</sub>-M<sub>env</sub>] correspondant à une première moitié de génome (*gag/pol*) de VIH-1/O et une seconde (*env*) de VIH-1/M. Le second profil implique le motif inverse [M<sub>gag/pol</sub>-O<sub>env</sub>] pour 36% d'entre elles (137).

Concernant la diversité des fragments impliqués dans ces formes recombinantes, plus de la moitié des fragments de VIH-1/M était de sous-type CRF02\_AG (52,6 %). Les VIH-1 diffusent largement dans la population et le sous-type CRF02\_AG est majoritaire au Cameroun. Des fragments de VIH-1/M au sein d'une forme recombinante pourraient conférer un avantage répliatif, permettant une diffusion facilitée de la nouvelle forme recombinante grâce aux propriétés des VIH-1/M. Concernant les fragments de groupe O, les différents profils de recombinaison comportaient une grande majorité de VIH-1/O de sous-groupe H (84%) (137). Les VIH-1/O possèdent un polymorphisme génétique naturel comparés aux VIH-1/M (90) et, malgré une plus faible prévalence par rapport au groupe M pandémique, il a été démontré que les VIH-1/O de sous-groupe H, possédaient une résistance aux INNTI, basée sur la présence d'une cystéine à la position 181 de la TI (Y181C) (59). Cette propriété, entraîne une résistance naturelle à cette classe de molécules, utilisée jusqu'à récemment dans de nombreuses lignes de traitements ARV (59, 170).

La question du potentiel répliatif de telles formes recombinantes se pose, car les VIH-1/O possèdent des caractéristiques très différentes de celles décrites pour les VIH-1/M. Deux études nous permettent de suspecter que les formes VIH-1/MO pourraient être mieux adaptées que les formes parentales dont elles ont émergé. Une première étude de *Ariën et al.*, a montré une capacité répliatif 100 fois supérieure des VIH-1/M face aux VIH-1/O (168) et une seconde de *Peeters et al.*, a montré une capacité répliatif plus importante de la forme recombinante VIH-1/MO face à sa forme

parentale VIH-1/M en culture (132). Une seule étude portant sur le potentiel répliatif d'une forme recombinante, comparée à ses deux formes parentales, a été menée au laboratoire ; détaillée dans le chapitre suivant. Les résultats de cette étude ont pu montrer que la souche recombinante produite *in vitro*, possédait un comportement différent selon les marqueurs étudiés au cours du suivi des cinétiques de répliation. Cette étude présentait des limites expérimentales, puisque cette forme recombinante générée ne correspondait pas à l'épidémiologie moléculaire observée au sein des 19 URF\_MO. De plus, une seule forme recombinante a pu être générée, ne permettant pas l'étude du potentiel répliatif de plusieurs formes recombinantes. Par ailleurs, aucune étude n'a pu à ce jour comparer le potentiel répliatif entre différentes formes recombinantes. Ces formes pourraient posséder à terme tous les avantages de chaque groupe M et O, combiner des mutations permettant l'échappement au facteur de restriction, voire posséder des *fitness* répliatif et adaptatif plus importants que les formes VIH-1/M et VIH-1/O parentales.

Ces données suggèrent que les formes recombinantes pourraient acquérir une combinaison génétique avantageuse en fonction du positionnement du point de cassure, majoritairement au sein des gènes accessoires centraux. Il pourrait donc être plus avantageux, pour le virus, de combiner une TI de groupe O, impliquant une résistance naturelle à certains INNTI, avec une enveloppe de groupe M, facilitant la diffusion de ce groupe au sein de la population, mais cette hypothèse n'a pas encore pu être prouvée. De plus, la combinaison de fragments de VIH-1/M, pandémique, au sein d'un VIH-1/O, endémique au Cameroun, pourraient conduire à une évolution de l'épidémie à VIH-1/O sous une forme de CRF, potentiellement plus adaptée que le VIH-1/O et diffusant dans la population.

Aujourd'hui, les conditions d'émergence des formes recombinantes inter-groupes MO restent floues. De nombreux contextes impliquant la description d'une ou plusieurs formes recombinantes inter-groupes MO ont été étudiés : des formes recombinantes associées à la présence d'une ou plusieurs formes parentales, des recombinantes isolés sans formes parentales, deux formes recombinantes au sein d'un même patient (105, 107, 110, 111, 132, 133, 135). Différents contextes cliniques ont également été rapportés, mais une seule étude a démontré la transmission directe d'une forme recombinante VIH-1/MO au sein d'un couple (REC003 et REC024) (136). En France, au cours de l'année 2022, la découverte d'une discordance moléculaire chez un nouveau-né au stade SIDA, suite à son admission en réanimation pédiatrique, nous a permis de mettre en évidence la transmissibilité mère-enfant d'une forme recombinante VIH-1/MO. Aucun cas de transmission verticale impliquant l'une de ces formes recombinantes n'avait encore été mis en évidence dans la littérature.

Ce travail de thèse avait donc pour objectif d'approfondir les connaissances sur la recombinaison entre VIH-1 de groupes M et O.

Un premier volet avait pour objectif d'étudier l'impact du profil de recombinaison sur le potentiel répliatif de ces formes recombinantes. Pour y parvenir, nous avons, d'une part, essayé de générer plusieurs clones moléculaires infectieux parentaux VIH-1/M et VIH-1/O et chimériques VIH-1/MO, afin de comparer le potentiel répliatif des formes recombinantes avec celui de ces formes parentales, grâce à un système de génétique inverse. D'autre part, nous avons comparé les cinétiques de répliation de deux souches recombinantes VIH-1/MO différentes, après mise en culture *in vitro*.

Le deuxième volet de cette étude avait pour objectif de caractériser un cas de transmission verticale, par l'étude du profil de recombinaison des deux virus recombinants et du lien phylogénétique entre les virus de la mère et de son enfant. La caractérisation des génomes complets a été entreprise à l'aide de deux technologies de séquençage avec, d'une part, la technologie Sanger déjà maîtrisée et disponible au laboratoire et, d'autre part, le développement d'une technique de séquençage à haut-débit.



## PARTIE EXPERIMENTALE

### **I. Potentiel répliatif et impact du profil de recombinaison des VIH-1/MO**

La génération d'un clone moléculaire infectieux chimérique (CMIC) de VIH-1/MO a été entreprise au cours d'une étude préliminaire au laboratoire, afin de comparer le potentiel répliatif *in vitro* de formes recombinantes MO par rapport à celui des formes parentales VIH-1/M et VIH-1/O dont elles étaient issues (137). Cette étude se basait sur une technique de génétique inverse, impliquant l'utilisation de clones moléculaires infectieux (CMI) pour produire les virus correspondants après culture sur lignée cellulaire sensible et permissive. Les CMI sont des outils de clonage correspondant à une séquence de génome complet de VIH inséré au sein d'un vecteur plasmidique. Ces clones permettent la production de virus dont la séquence génétique est préalablement connue, car ils proviennent de virus parentaux connus. La comparaison des souches entre elles peut donc être réalisée en s'affranchissant de la diversité virale induite par la mise en culture du VIH. Enfin, tous ces clones peuvent être transfectés dans un système cellulaire sensible afin d'y produire les virions correspondants.

Cette étude portait également sur la comparaison des potentiels répliatifs *in vitro* des recombinants entre eux, afin de rechercher un éventuel profil préférentiel de recombinaison (137). Des CMIC ont été envisagés à partir de différentes souches parentales de VIH-1/M et de VIH-1/O et modélisés avec un unique point de cassure dans *vpr*, puisqu'un point chaud de recombinaison a été identifié dans cette région pour 53% des 19 URF\_MO (n=10/19) *in vivo*. L'étude préalable de la répartition des points de cassure au sein de ces formes recombinantes avait montré que sur les 32 points de cassures identifiés au total entre *gag* et *nef*, 44% des points de cassure (n=19/32) étaient localisés dans les gènes accessoires centraux *vif*, *vpr* et *vpu*. Deux profils de recombinaison avaient donc été envisagés selon les profils identifiés comme majoritaires dans les URF\_MO décrites *in vivo* : le motif [O<sub>gag/pol</sub>-M<sub>env</sub>] et le motif inverse [M<sub>gag/pol</sub>-O<sub>env</sub>]. Étant donnée l'hypothèse d'un avantage répliatif conféré par la recombinaison, posée par les travaux de *Peeters et al.*, la comparaison du potentiel répliatif du recombinant à celui de ses formes parentales nécessitait d'avoir à disposition les CMI parentaux de VIH-1/M et VIH-1/O correspondants. Parmi les souches parentales disponibles, un VIH-1/M de sous-type B (89.6) et un VIH-1/O de sous-groupe T (RBF206) ont pu être utilisés (137).

Malgré les multiples stratégies testées, aucun CMIC candidat ne s'est révélé être infectieux après culture *in vitro*. Un CMIC de motif [O<sub>gag/pol</sub>-M<sub>env</sub>] a néanmoins été obtenu par synthèse par la société Genecust (GC, Luxembourg), appelé OM-TB-2016. Afin de comparer les potentiels répliatifs des



souches parentales VIH-1/M et VIH-1/O et de la souche recombinante OM-TB-2016, des cinétiques de réplication ont été réalisées sur lignées cellulaires. Deux marqueurs ont été suivis dans les surnageants de culture : l'activité de la TI et la quantité d'AgP24. Pour procéder à la comparaison, trois paramètres ont été pris en compte pour chacun de ces deux marqueurs : la vitesse d'atteinte du pic, l'intensité de ce pic et la vitesse de réplication. Les résultats de cette étude ont permis de confirmer un potentiel réplicatif supérieur du VIH-1/M par rapport à celle du VIH-1/O, quel que soit le marqueur utilisé en culture, comme ce qui avait été montré par *Ariën et al* (137, 168). Concernant la souche recombinante, un comportement différent entre les deux marqueurs a été observés, montrant un comportement similaire à la souche parentale VIH-1/O pour la production d'AgP24 en termes de délai, d'intensité et de vitesse de production, un comportement similaire à la souche parentale VIH-1/M pour la quantification de l'activité de la TI, en termes de délai et de vitesse, mais une intensité plus proche de celle du VIH-1/O parental. Cette étude n'a pas permis de mettre en évidence un meilleur potentiel réplicatif de la souche recombinante OM-TB-2016, mais a montré que cette forme possédait un comportement intermédiaire selon les marqueurs étudiés.

L'obtention de ce premier CMIC a prouvé qu'il était possible de générer ces clones et d'en étudier leur profil au laboratoire, mais de nombreuses limites ont été soulevées. Premièrement, ces données préliminaires obtenues sur le potentiel réplicatif ne permettaient pas d'indiquer un éventuel avantage réplicatif de la forme recombinante VIH-1/MO. Pour parvenir à une telle conclusion, cela impliquait de mettre en place d'une étude de *fitness*. En effet, la mise en compétition de chaque souche de VIH-1/M, VIH-1/O et VIH-1/MO permettrait de visualiser une différence entre elles, mais nécessite le développement d'outils de révélation appropriés. Ensuite, le choix des CMI parentaux VIH-1/M et VIH-1/O a été conditionné par la disponibilité des souches correspondantes sous forme de clone au laboratoire, limitant le choix pour la génération des CMIC. Ainsi la construction réalisée était complètement artificielle, aucune combinaison VIH-1/M de sous-type B associé à un VIH-1/O de sous-groupe T n'ayant été décrite *in vivo*. Enfin, l'échec de génération du motif de CMIC [ $M_{gag/pol-O_{env}}$ ] rendait impossible la comparaison des deux profils de recombinaison entre eux. Une étude publiée par la suite, a également montré que la souche de VIH-1/O sélectionnée pour ce travail possédait une protéine Vpu semblable à celle des VIH-1/M, permettant de contourner l'action de la Tétherine et qu'elle n'était pas la plus représentative des VIH-1/O (171).

C'est dans ce contexte que se situe ce projet de doctorat, dont le premier objectif était d'étudier et de comparer le potentiel réplicatif de plusieurs formes recombinantes VIH-1/MO, après mise en culture *in vitro*.

## 1. Génération *in vitro* des virus VIH-1/M, VIH-1/O et VIH-1/MO

Afin d'effectuer une comparaison entre formes parentales VIH-1/M et VIH-1/O et formes recombinantes VIH-1/MO, une étape préliminaire de construction de CMI a été effectuée pour ensuite produire les CMIC. Cette partie du travail expérimental avait pour objectifs, d'une part, de comparer les potentiels réplcatifs de deux formes recombinantes de motif inverse, afin d'analyser l'impact du profil de recombinaison et, d'autre part, de comparer le potentiel réplcatif de ces formes à ceux des deux formes parentales dont elles provenaient.

### a. Méthodologie

#### *Choix des souches parentales*

Du fait des travaux réalisés antérieurement, plusieurs CMI de VIH-1/M étaient déjà disponibles : pNL4-3 et p89.6 (sous-type B), pCRII-TOPO-SE8131-3 (sous-type A) et p92UG0001 (sous-type D), tous obtenus du CFAR (*Centre for AIDS Reagents*) du NIBSC (*National Institute for Biological Standards and Control*). Deux CMI de VIH-1/O étaient également disponibles, issus d'une précédente collaboration : pRBF206 (sous-groupe T) et pBCF120 (sous-groupe H) (86). Pourtant, le CMI pBCF120, impliquant le sous-groupe majoritaire H, n'avait pas pu être amplifié au laboratoire et nécessitait une adaptation du protocole de clonage. Etant donné la faible disponibilité de CMI de VIH-1/M sous-type CRF02\_AG (pourtant majoritairement retrouvé dans les URF\_MO) et de CMI de VIH-1/O, la première étape a consisté en la génération de nouveaux CMI parentaux. Pour ce faire, trois souches de VIH-1/O, YBF026, RBF130 et BCF108, toutes de sous-groupe H et trois de VIH-1/M, AAS638, M-A-062021 et M-F-062021 de sous-type CRF02\_AG, A et F, respectivement, ont été testées en plus des CMI déjà disponibles au laboratoire.

#### *Plasmides vecteurs et pBCF120*

Le plasmide vecteur pBR322 (NEB, France) est un plasmide circulaire de 4361 pb (n° d'accèsion J01749.1) présent en 10 à 100 copies par génomes bactérien (172). Ce plasmide possède un gène de résistance à l'ampicilline (Amp) permettant la sélection des bactéries contenant les souches de VIH-1 d'intérêt, après transformation chimique (172, 173).

Le CMI pBCF120 déjà disponible, correspond à un génome complet de VIH-1/O de sous-groupe H, inséré dans le vecteur pSMART-LC-Amp, d'environ 2049 pb, qui possède un gène de résistance à l'Amp. La taille totale de pBCF120 est de 11941 pb.

## Principe général d'assemblage par Gibson Assembly

Afin de générer d'autres CMI, nous avons utilisé la méthode de Gibson Assembly, qui est une technique de biologie moléculaire permettant d'assembler différents fragments au sein d'un plasmide vecteur linéarisé, par l'action de trois enzymes différentes, lors d'une même réaction isothermique (174, 175). Cette technique repose sur l'utilisation de fragments modélisés et obtenus par PCR, à l'aide d'amorces dites chevauchantes, car possédant des régions complémentaires entre les extrémités des différents fragments à assembler (Figure 25). Cette fragmentation par PCR est réalisée en parallèle d'une linéarisation d'un vecteur, dans lequel seront insérés les fragments d'intérêt. Tout d'abord, une 5'exonucléase digère les extrémités 5' de chaque fragment à assembler. Tous ces fragments deviennent alors chevauchants. Une ADN polymérase synthétise alors des ADNc de chaque fragment, puis une ligase lie ces derniers entre eux, permettant la circularisation du plasmide. Le résultat final correspondait à l'assemblage théorique modélisé, appelé produit Gibson Assembly (GA).

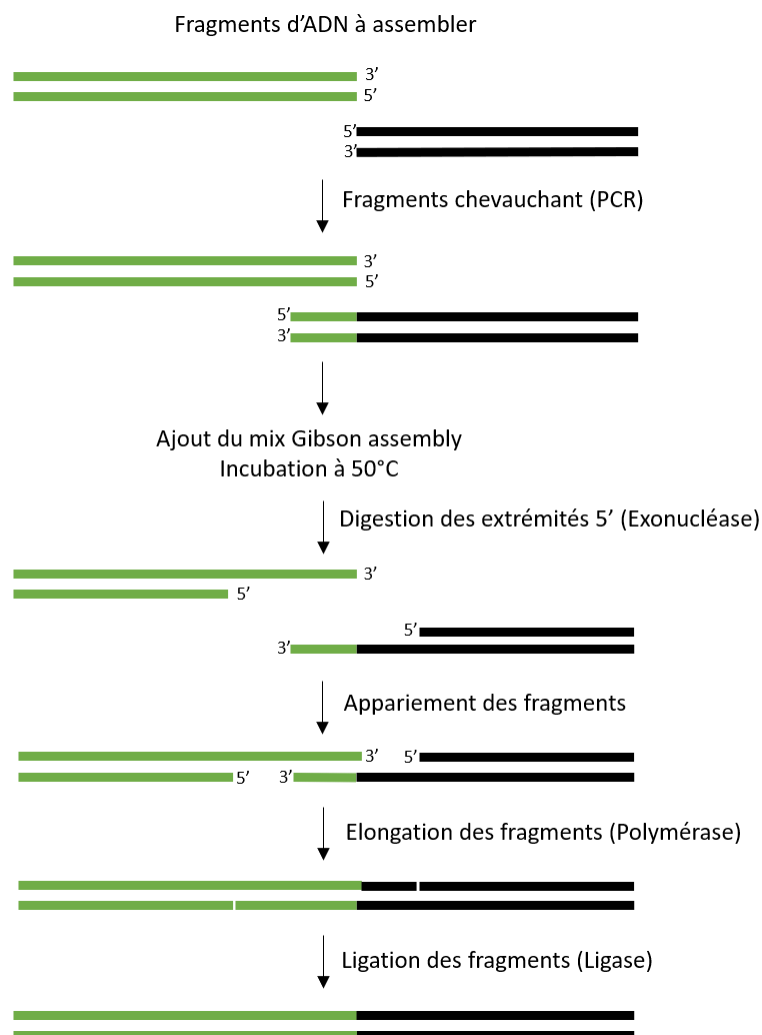


Figure 25 : Principe général de la réaction Gibson Assembly

Image de Tobias Vornholt adaptée en français (CC BY-SA 4.0) (176).

## Génération des CMI

### Modélisation des amorces.

Une première étape de visualisation des clones à l'aide de l'outil NEBuilder® Assembly Tool de NEB a ensuite été réalisée, afin de concevoir les amorces chevauchantes nécessaires à la réaction de Gibson Assembly (Annexe 1).

### Génération des fragments.

Le plasmide pBR322 utilisé au cours de cette étude possède un site de restriction ciblé par l'enzyme BamHI-HF® (NEB) au sein du Multiple Cloning Site (MCS), nous permettant de linéariser au préalable ce plasmide. La digestion enzymatique a été réalisée selon le protocole préconisé par l'outil NEB Cloner® (NEB) en ajoutant 0,5µL d'enzyme à 1µg d'ADN, 5µl de tampon 10X CutSmart et de l'eau RNase free pour un volume réactionnel final de 50µL. Une incubation a ensuite été réalisée pendant 1h à 37°C.

Pour la génération des CMI, la démarche était d'assembler les différents fragments provenant d'un même génome complet, scindé en deux fragments par PCR afin faciliter l'insertion au sein du plasmide pBR322. Le premier fragment, appelé Fragment A, allait du 5'LTR jusqu'au gène *vif* alors que le deuxième fragment, appelé B, s'étendait du début de *vif* jusqu'au 3'LTR (Figure 26). Les amorces ont été positionnées dans la région *vif* de façon à obtenir deux fragments de taille quasi identique, indifféremment de la diversité des souches, soit environ 5000 pb et 4700 pb pour les fragments A et B, respectivement.

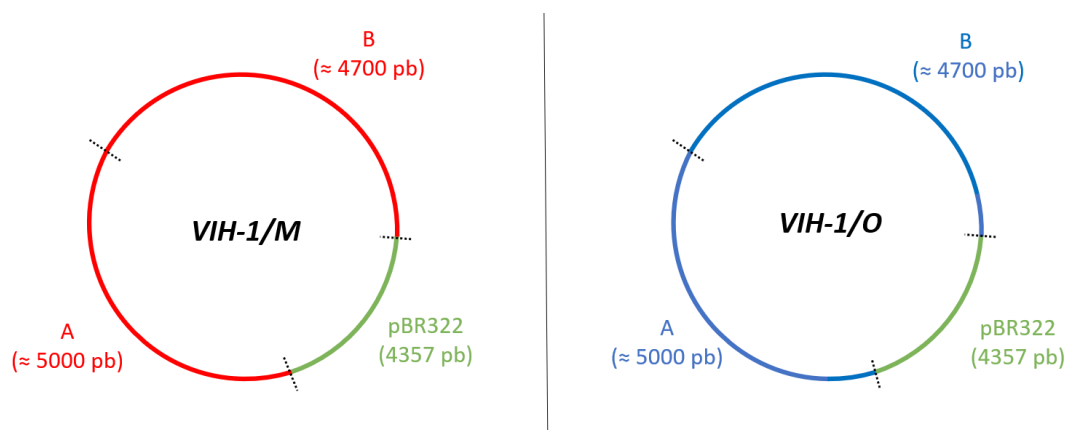


Figure 26 : Profil des CMI VIH-1/M et VIH-1/O

Les profils moléculaires des deux CMI VIH-1/M (à gauche) et VIH-1/O (à droite) sont représentés dans cette figure. Les fragments de VIH-1/O sont représentés en bleu, de VIH-1/M en rouge et le plasmide en vert. Les différents fragments ont été nommés A et B, respectivement, en partant du 5'LTR vers le 3'LTR. Chaque position d'un point de cassure est représentée par des traits en pointillés.

L'ADN intracellulaire des différentes souches a été extrait de façon automatisée sur l'automate EZ1® avec le kit EZ1® advanced XL (Qiagen, Allemagne), à partir d'un culot de cellules de fin de culture précédemment congelées au laboratoire et selon les recommandations du fournisseur. Ces extraits d'ADN ont été quantifiés par Biodrop® (Sciencetec, France) selon les recommandations fournisseurs.

Les deux moitiés de génome ont été amplifiées par PCR à l'aide du kit Long Range® PCR (Qiagen) en complétant 5µL de mix 10X par 2,5µL de dNTP à 10mM chacun, 2µL de chaque amorce sens et antisens à la concentration de 0,4µM (Annexe 1), 500ng d'ADN extrait, 0,4µL d'enzyme et par de l'eau RNase free pour un volume réactionnel final de 50µL. Une première étape de dénaturation initiale a été réalisée 3min à 93 °C, suivie de 10 cycles sans incrémentation de 15s à 93°C, 30s à 58 °C, 6min à 68°C, puis 35 cycles identiques aux précédents ont été réalisés pour l'élongation avec une incrémentation de 20s par cycle ; enfin une élongation finale a été réalisée 10 min à 68 °C.

Chaque fragment a été contrôlé après PCR, par révélation sur gel d'agarose à 1%, puis purifié à l'aide du kit Nucleospin Gel and PCR clean-up (Macherey-Nagel) avant d'être quantifié par Biodrop®.

#### **Assemblage des CMI par réaction de Gibson Assembly.**

Chaque CMI a été produit en utilisant la méthode de GA précédemment décrite, adaptée pour l'assemblage de trois fragments : le génome complet de la souche parentale fragmenté en deux ainsi que le plasmide vecteur linéarisé (Figure 27). Les quantités optimales de chaque fragment ont été calculées selon les recommandations fournisseurs, soit 0,03 à 0,2 pmols dans un volume maximal de 10µL pour 2 à 3 fragments assemblés et 100ng de plasmide pour 200ng de fragments. Ces fragments ont été ajoutés à un mix contenant 10µL de NEBuilder HiFi DNA Assembly Master Mix (NEB) et de l'eau RNase free si besoin pour un volume final de 20µL. Le mélange a été incubé 15min à 50°C.

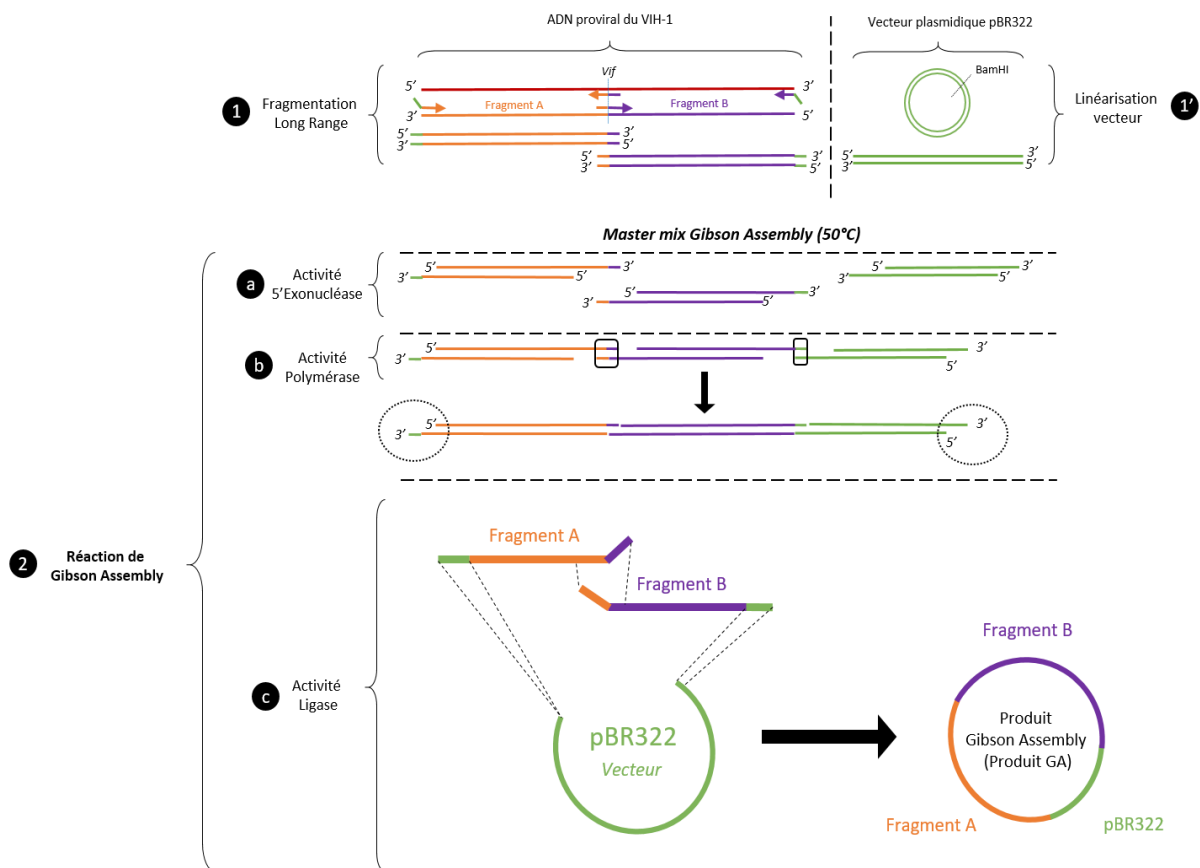


Figure 27 : Méthode Gibson Assembly avec 2 fragments pour la génération des CMI parentaux

Des numéros et lettres minuscules sont indiqués à gauche de chaque étape de l'assemblage par méthode Gibson Assembly, sauf pour la première étape où la fragmentation de la souche parentale et la linéarisation du plasmide vecteur se fait en parallèle. Le passage d'une étape à l'autre est symbolisé par des traits en pointillés en dessous de chaque étape. **1**) L'ADN intracellulaire d'une souche parentale (génomme complet en rouge foncé) est fragmenté en deux parties A et B, représentant la première moitié de génome allant du 5'LTR au point de cassure *vif* (A en orange) et la deuxième moitié de génome allant du point de cassure dans *vif* jusqu'au 3'LTR (B en violet). Ces fragments sont obtenus à l'aide d'amorces modélisées pour obtenir des zones chevauchantes entre fragments à assembler **1'**) En parallèle, le vecteur pBR322 est linéarisé par digestion enzymatique au niveau du site de restriction unique *Bam*HI. **2**) La réaction de GA a lieu en 3 étapes, toutes à 50°C. **a**) Les extrémités 5' de chaque fragment sont digérées par l'activité 5'exonucléase de la polymérase. **b**) Cette dernière peut alors faire correspondre les extrémités 3' et 5' de chaque fragment chevauchant (encadrés noirs) et procéder à l'élongation pour combler les séquences. Les extrémités 3' et 5' restantes (cercles en pointillés noirs), également chevauchants, correspondant aux extrémités du plasmide, permettant la circularisation dans l'étape suivante **c**) La ligase contenue dans le mélange réactionnel assemble les séquences complémentaires par ligation et provoque la circularisation du plasmide. Le génome complet inséré au sein du vecteur, est appelé produit Gibson Assembly (GA).

## Clonage bactérien.

Les différents produits issus de l'assemblage Gibson Assembly ont été, par la suite, clonés dans des bactéries, afin d'être amplifiés (Figure 28). Les bactéries NEB® Stable Competent *E. coli* (New England BioLabs, Etat Unis) ont été choisies car elles sont chimiocompétentes, elles permettent une haute efficacité et un bon rendement de transformation chimique et elles sont idéales pour la purification de clones plasmidiques contenant des éléments répétés ou instables, comme c'est le cas dans les génomes de VIH (177). Le milieu NEB® Outgrowth Medium (NEB), fourni avec ces cellules, permet une bonne survie des bactéries après choc thermique (177). Ces bactéries ont également été choisies car elles sont sensibles à l'ampicilline (Amp) à la concentration de 100 µg/mL (Amp 100). Elles possèdent également le gène codant la β-galactosidase qui permet de visualiser les colonies transformées en présence de X-gal, qui est un substrat chromogène de cette enzyme. Le protocole de clonage complet a été détaillé en Annexe 2.

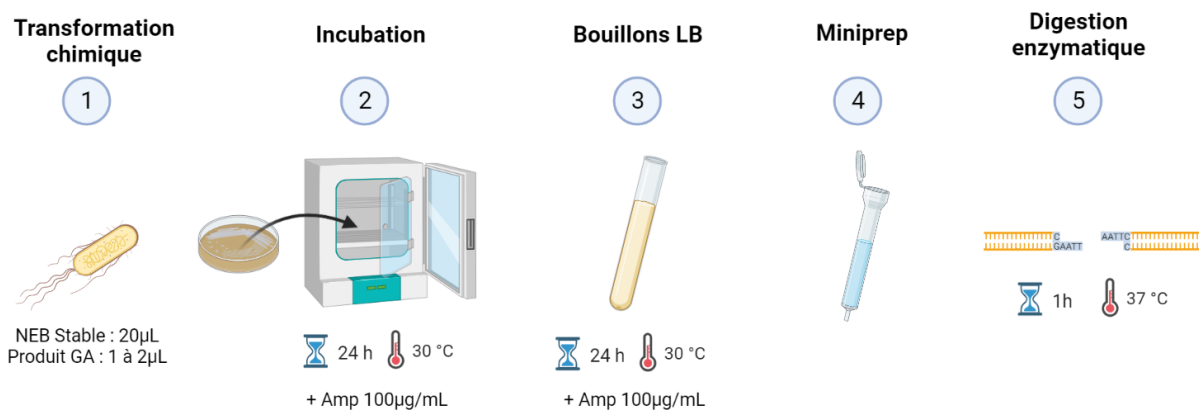


Figure 28 : Etapes de la transformation chimique des produits Gibson Assembly

Créé avec BioRender.com

## Vérification des clones.

Chaque candidat a ensuite été quantifié par Biodrop®, puis une digestion a été réalisée par une ou plusieurs enzymes de restriction afin de valider le profil des candidats. Ces enzymes ont été sélectionnées en fonction de la présence des sites de restriction positionnés sur les clones, afin d'amplifier au maximum trois fragments différents. Chaque clone ayant le bon profil de digestion, ou une taille totale attendue, était ensuite testé en culture cellulaire.

L'ensemble des étapes de production des CMI et CMIC a été schématisé dans la Figure 29.

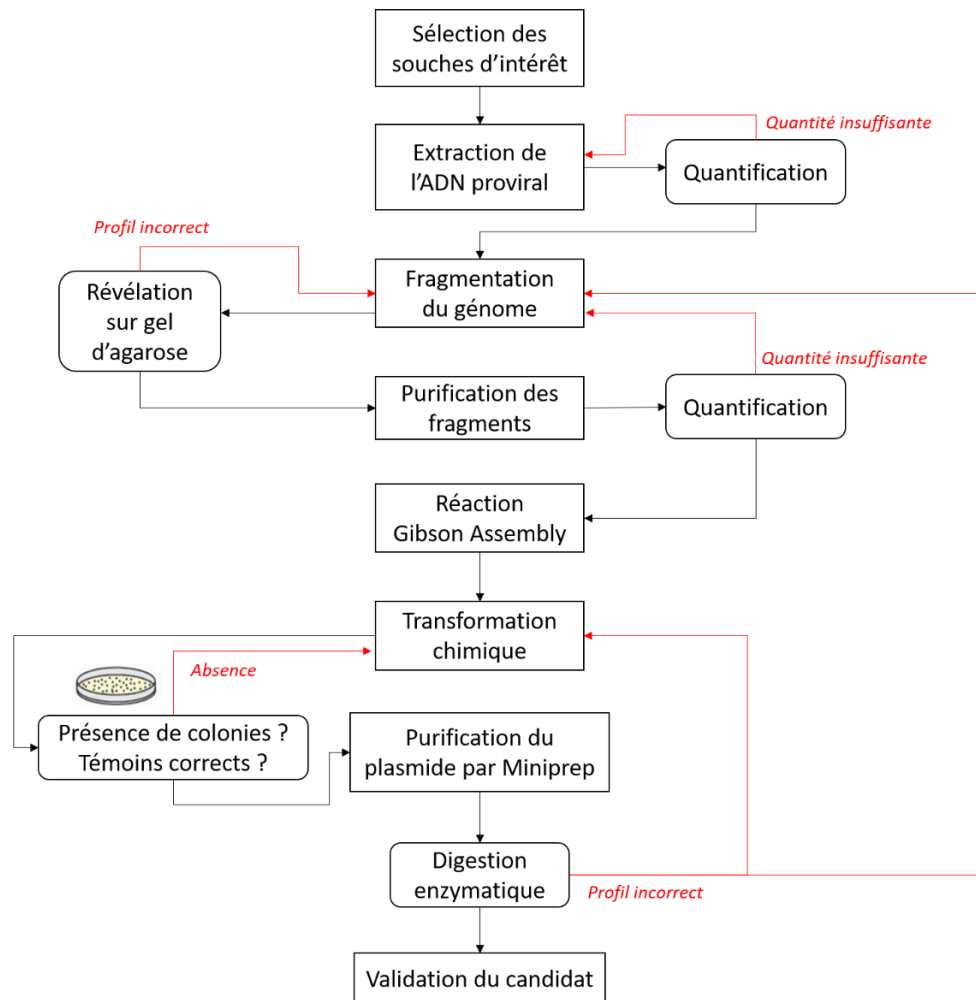


Figure 29 : Vue d'ensemble des différentes étapes de production des CMI et CMIC

Les cadres rectangulaires montrent les étapes clés du protocole, celles impliquant des étapes de vérifications sont représentées par des cadres à coins arrondis. Les flèches noires symbolisent le passage d'une étape à une autre tandis que les flèches rouges impliquent un retour à une étape précédente après échec/impossibilité de passer à l'étape suivante.

### Génération des CMIC :

#### Modélisation des amorces et CMIC.

Après la génération des CMI, nous avons modélisé les différents CMIC à produire. Les points de cassure ont été positionnés au niveau des points chauds de la recombinaison identifiés *in vivo* (137). Un premier point de cassure a donc été positionné dans le gène *vpr*, entre les bases 5624 et 5636. Le deuxième point de cassure a été positionné dans les 5' et 3'LTRs, entre les bases 513 et 522, au niveau du motif de recombinaison décrit comme majoritaire dans une étude précédente (137). Deux profils de recombinaison ont alors été envisagés (Figure 30).



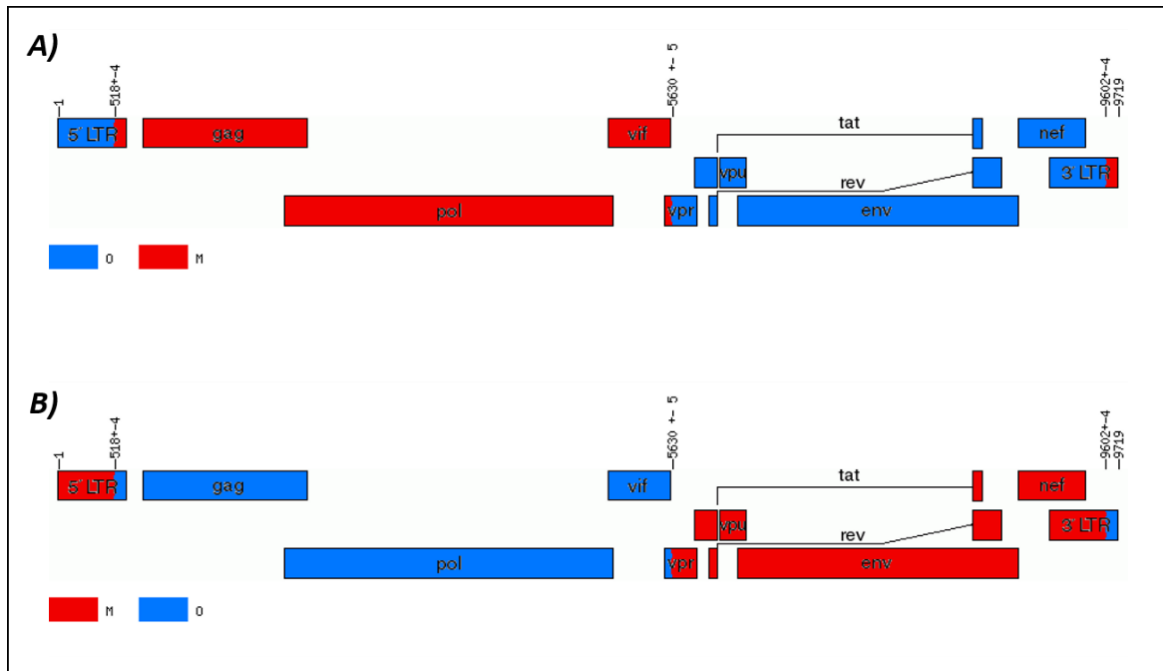


Figure 30 : Profil de recombinaison des deux formes VIH-1 « OMOM » et « MOMO »

Le profil moléculaire des deux formes VIH-1/OMOM et MOMO sont représentés en A et B respectivement. Les fragments de VIH-1/O sont visibles en bleu tandis que les fragments de VIH-1/M sont en rouge. Chaque position a été indiquées par rapport à la souche de référence HxB2 allant du premier codon 1, au dernier 9719. Les deux points de cassure ont été positionné sur les génome dans *vpr* aux positions 5624 - 5636 pb ( $5630_{\pm 5}$ ) et dans les LTRs aux positions 513 - 522 pb ( $518_{\pm 4}$  et  $9602_{\pm 4}$ ).

Le premier profil envisagé était constitué d'une première moitié de génome VIH-1/M entre *gag* et *vpr*, d'un gène *vpr* recombinant de motif MO, d'une deuxième moitié de génome VIH-1/O entre *vpr* et *nef* et de LTRs recombinants, de motif OM. Ce premier profil a été nommé VIH-1/OMOM (Figure 30A). Un profil de recombinaison inverse VIH-1/MOMO a également été modélisé au cours de cette étude, associant un point de cassure de motif MO au sein des LTRs et un point de cassure de motif OM dans le gène *vpr* (Figure 30B).

L'assemblage de ces fragments nécessitait également de modéliser des amorces chevauchantes qui devaient s'hybrider au niveau des régions conservées (Annexe 3). La génération des CMIC implique d'assembler des fragments adjacents de groupes M et O, qui sont assez divergents et rendent l'assemblage difficile, contrairement aux fragments utilisés pour la génération des CMI, qui proviennent d'une même souche de départ.

## Génération des fragments.

Selon notre modélisation avec deux points de cassure, l'assemblage comportait quatre fragments au lieu de deux comme décrit pour les CMI et ces fragments étaient de tailles très différentes. Le positionnement d'un point de cassure dans les gènes accessoires centraux entraîne une coupure du génome complet en deux fragments d'environ 5100 et 4300 pb de part et d'autre de *vpr*, tandis que le point de cassure positionné dans les LTRs, entraîne-lui, la génération de plus petits fragments allant de 150 à 300 pb (Figure 31). Chaque fragment a été nommé de A à D dans le sens 5' vers 3' et en partant de la fin de pBR322. Précédant cette lettre, la diversité du fragment VIH-1/M ou VIH-1/O a été renseignée par la lettre M ou O correspondante.

Le premier fragment (MA ou OA) allait donc du début de la région U3 du 5'LTR jusqu'au premier point de cassure à la fin de la région R du 5'LTR, tandis que le deuxième fragment (MB ou OB), allait de R jusqu'au point de cassure du gène *vpr*. Le troisième fragment (MC ou OC), quant à lui, allait du point de cassure *vpr* jusqu'à celui décrit au sein du 3'LTR, identique au 5'LTR. Enfin, le dernier fragment (MD ou OD), allait de la région R du 3'LTR jusqu'à la fin du 3'LTR (Figure 31). Le protocole d'obtention de ces fragments a été détaillé en Annexe 4.

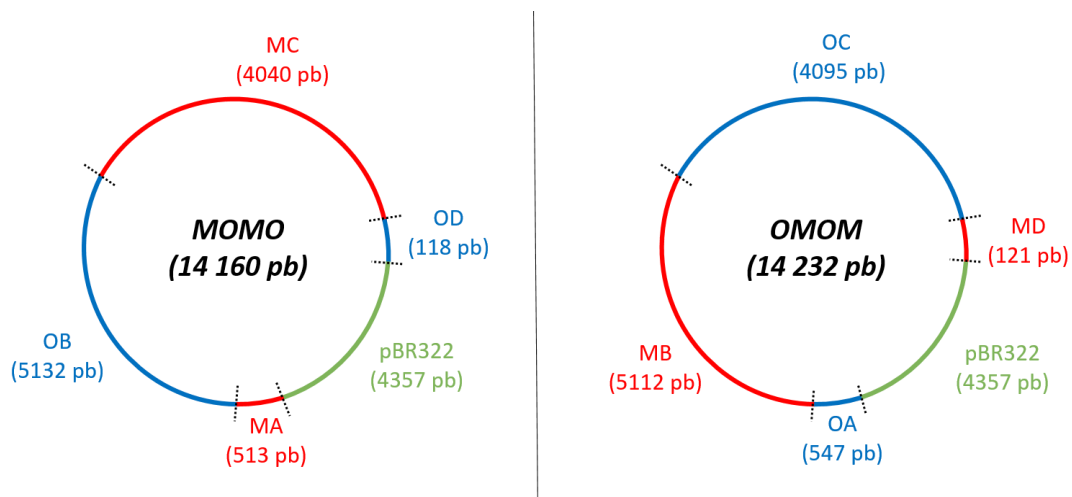


Figure 31 : Profil des CMIC VIH-1/MOMO et OMOM

Les profils moléculaires des deux CMIC MOMO (à gauche) et OMOM (à droite) sont représentés dans cette figure. Les fragments de groupe O sont représentés en bleu, de groupe M en rouge et le plasmide en vert. Les différents fragments ont été nommés A, B, C et D, respectivement, en partant du 5'LTR vers le 3'LTR. A côté de chaque lettre représentative d'un fragment, le groupe M ou O associé a été indiqué et la taille attendue de chaque fragment a été également indiquée entre parenthèse. La position d'un point de cassure a été représentée par des traits en pointillés. La taille totale attendue des CMIC a été indiquée entre parenthèse en noir au centre de la figure.

### **Assemblage des CMIC par réaction de Gibson Assembly.**

Concernant la production des CMIC, les différents fragments ont été assemblés par la même méthodologie que pour les CMI, décrite précédemment : **Génération des CMIC : Assemblage par réaction de Gibson Assembly** (page 66). Néanmoins, le protocole de Gibson Assembly a été adapté au niveau des quantités de fragments utilisées et des tailles des fragments, soit 0,2 à 0,5 pmols dans un volume maximal de 10µL pour 4 à 6 fragments assemblés (Figure 32). Deux fois plus de petits fragments, d'environ 120 et 500 pb, que de fragments moyens d'environ 5100 et 4000 pb, ont été ajoutés pour respecter le ratio de 200 ng de fragments totaux pour 100ng de plasmide.

### **Clonage bactérien et vérification des clones.**

Les CMIC candidats ont été obtenus après clonage bactérien selon le même protocole mis en place pour les CMI (Annexe 2) et vérifiés selon le même protocole mis en place pour les CMI, décrit dans le paragraphe : **Génération des CMI : Vérification des clones** (page 68).

### **Clonage des CMIC par la société Genecust**

En parallèle des différentes expérimentations de clonage au laboratoire, nous avons entrepris la production des CMIC d'intérêt à l'aide de la société Genecust déjà sollicitée au cours de l'étude préliminaire de 2019 (137). Des tests de clonage ont été réalisés à partir des CMI VIH-1/M et VIH-1/O produits au cours de cette étude.

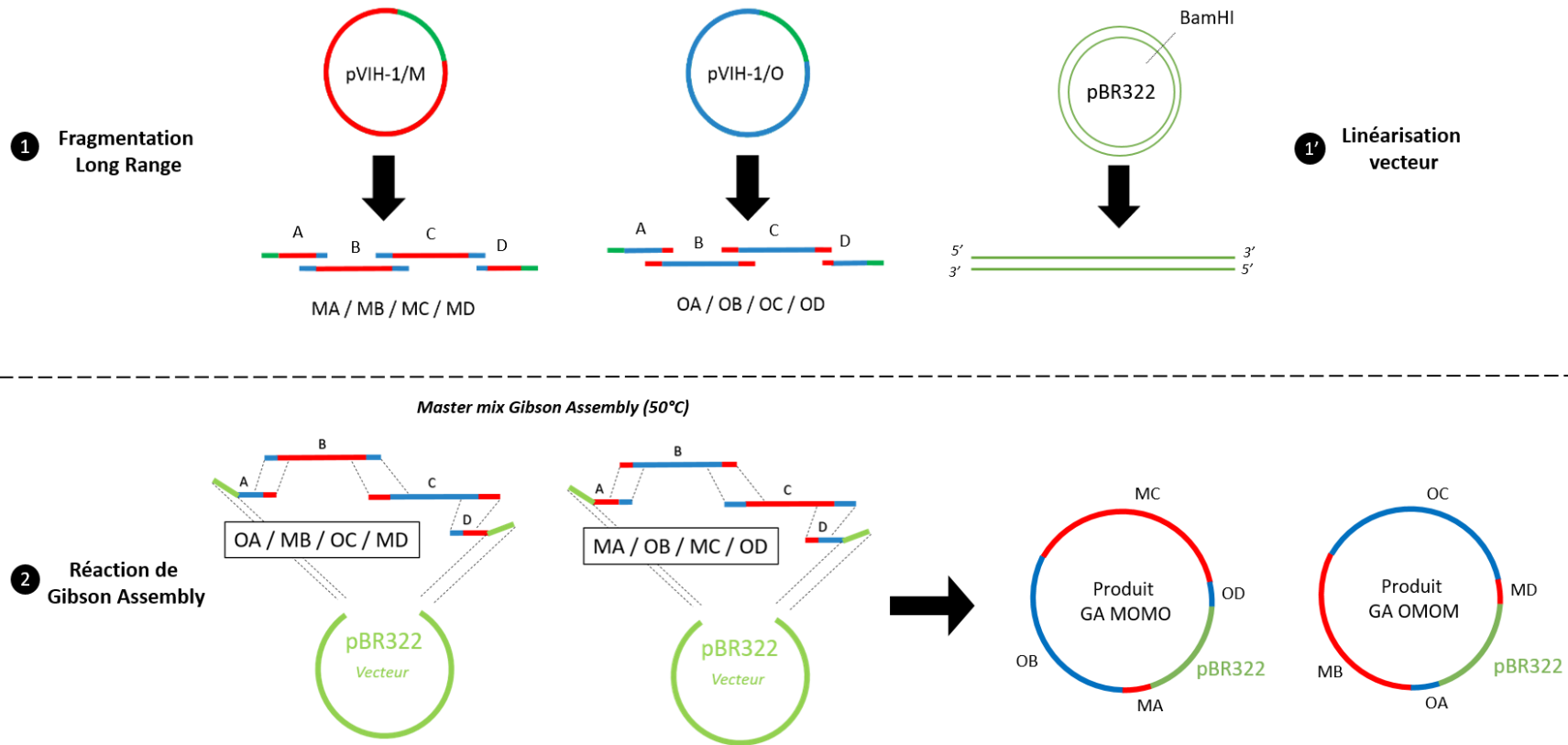


Figure 32 : Méthode Gibson Assembly pour la génération des CMIC avec 4 fragments

**1)** Le génome complet de deux CMI contenant les génomes parentaux VIH-1/M (rouge) et VIH-1/O (bleu) inséré dans pBR322 (en vert), est fragmenté pour générer 4 fragments selon la même méthode que pour la génération des CMI. Les quatre fragments obtenus (A, B, C et D) possèdent des parties chevauchantes en 3' et 5' pour permettre la réaction GA. Ces derniers sont précédés de la lettre M si le fragment est de groupe M (MA, MB, MC, MD) et O si le fragment est de groupe O (OA, OB, OC, OD). **2)** La réaction GA a lieu selon les mêmes conditions que pour les CMI mais incluant 4 fragments. La circularisation du CMIC entraîne la production d'un génome chimérique génome complet inséré au sein du vecteur.

### ***Génération des virus correspondants par génétique inverse***

Les CMI et CMIC candidats, présentant le bon profil de digestion après clonage, ont été mis en culture par un système de génétique inverse, afin de générer la production de particules virales infectieuses. Une première étape de transfection a été réalisée pour générer les virus correspondants à chaque clone, puis une deuxième étape de co-culture a été réalisée afin d'amplifier ces virus.

Une première étape de transfection de cellules HEK293T a été effectuée pour tester chaque clone candidat purifié (Annexe 5). La transfection consiste à insérer des acides nucléiques, tel qu'un plasmide, au sein d'une cellule eucaryote, permettant ainsi la production de protéines ou d'un virus dans notre cas, par cette cellule. Les cellules HEK293T sont des cellules embryonnaires de rein humain largement utilisées pour la culture des rétrovirus. Elles sont hautement transfectables grâce à l'expression endogène de l'antigène T du virus simien 40 (SV40) et sont facilement cultivables *in vitro*. Ces cellules adhérentes ont été divisées au 1/10<sup>ième</sup> deux fois par semaine, dans du milieu DMEM+GlutaMAX™ (*Dulbecco's minimal essential medium*, ThermoFisher Scientific, France) + Gentamicine (50µg/mL) frais, avec 10% de sérum de veau fœtal décomplémenté (SVF<sub>d</sub>, Eurobio scientific, France) 30 min à 56°C.

Comme ces cellules ne possèdent ni CD4, ni corécepteurs, il n'y a pas d'amplification virale possible après l'étape de transfection. La deuxième étape de co-culture, consistant à produire des particules virales infectieuses, a été réalisée par l'infection de cellules Jurkat X4-R5 (CFAR, NIBSC) (Annexe 5). Cette lignée de lymphocytes T immortalisés, exprime de façon stable le corécepteur CR5, en plus du CXCR4 déjà exprimé de façon endogène. Cette caractéristique nous permettait de nous affranchir d'un éventuel problème de tropisme des souches et de faciliter la comparaison entre elles. Les cellules Jurkat X4-R5 ont été entretenues dans du milieu RPMI+GlutaMAX™ 1640 (Roswell Park Memorial Institute, Lonza, Suisse) avec 10% de SVF<sub>d</sub>, de la Gentamicine (50µg/mL) et avec ajout de Généticine G418 à la concentration de 500µg/mL final (Fisher scientific™, France) afin de sélectionner les cellules possédant les deux corécepteurs CXCR4 et CCR5. Ces cellules ont été divisées deux fois par semaine au quart.

L'ensemble des expérimentations de culture ayant permis la production des particules virales est schématisé dans la Figure 33.

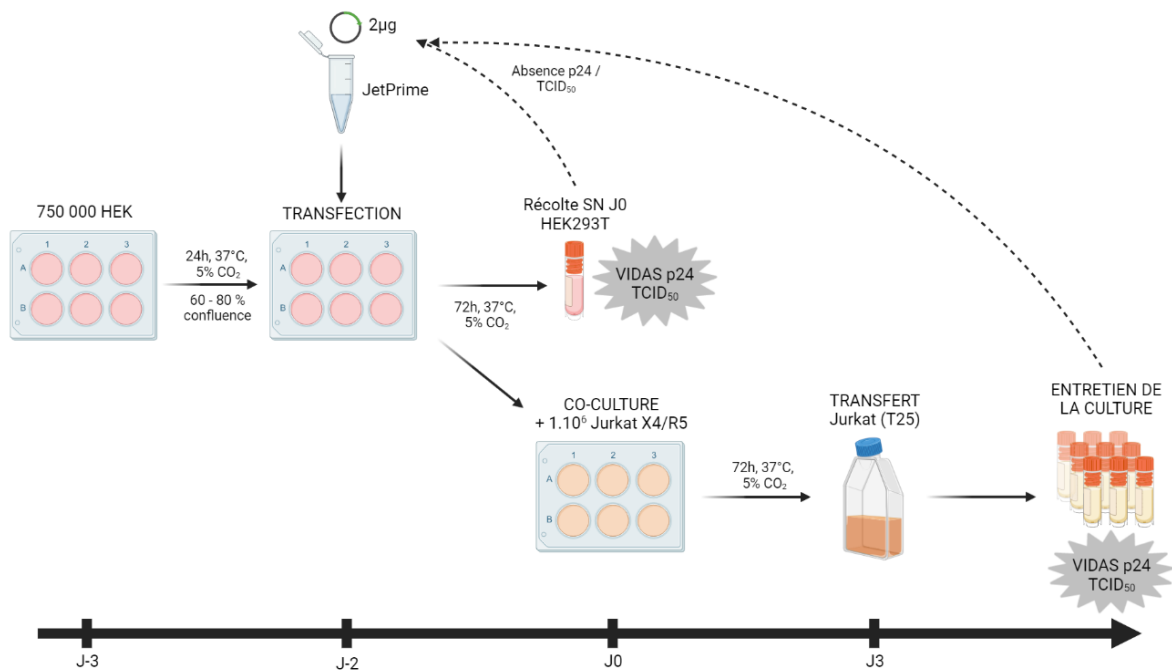


Figure 33 : Schéma expérimental de production des virus par transfection puis co-culture

Créé avec BioRender.com

### Vérification de la production virale

Une mesure de la quantité d'antigène P24 (AgP24) a été effectuée à la fin de la culture à l'aide du kit Innostest® HIV Antigen mAb (Fujirebio, Belgique) selon les recommandations du fournisseur, pour l'ensemble des prélèvements, afin de mettre en évidence la présence du virus.

L'infectiosité a ensuite été déterminée par la mesure de la Tissue Culture Infectious Dose 50 (TCID<sub>50</sub>) en duplicat, selon la technique de Montefiori, pour les candidats ayant produit de l'AgP24 (178). Cette technique a été réalisée en infectant des cellules adhérentes TZM-bl (CFAR, NIBSC) exprimant les corécepteurs CXCR4 et CCR5 de façon constitutive, ainsi que les récepteurs CD4 (179). En présence du promoteur de *tat* du VIH-1, donc s'il y a bien eu infection des cellules, celles-ci expriment le gène rapporteur de la luciférase de luciole, ce qui permet de mesurer la luminescence émise par ces cellules. Les TZM-bl utilisées au cours de cette étude ont été entretenues deux fois par semaine selon les mêmes modalités que les HEK293T. Le jour du test de TCID<sub>50</sub>, une première étape impliquant la dilution en série au 1/5<sup>ième</sup> des surnageants a été effectuée en duplicat dans une plaque 96 puits. Dans chaque puits, 1,5.10<sup>4</sup> cellules ont été déposées pour 100µL de DMEM+SVFd+Gentamicine et complété avec du DEAE dextran (Sigma-Aldrich, Etat Unis) à la concentration finale de 40µg/mL afin de favoriser l'infection des cellules. Après 72h d'incubation de la

plaque à 37°C avec 5% de CO<sub>2</sub>, 100µL de surnageant ont été enlevés de chaque puits, avant ajout de 100µL de réactif Bright-Glo™ Luciferase Assay (Promega, Etats Unis) puis incubation de 2min pour lyser les cellules. Afin d'effectuer la révélation de la plaque par lecture de la luminescence et après une brève homogénéisation, 150µL du mélange obtenu ont été transférés dans une plaque 96 puits blanche opaque. La TCID<sub>50</sub> a ensuite été mesurée pour chaque virus, par la quantification de la luminescence émise pour chaque puit, à l'aide du luminomètre Centro LB 960 (Berthold Technologies, France) et exprimée en unité de luminescence relative (RLU pour *Relative Light Units*).

Si les deux marqueurs (AgP24 et TCID<sub>50</sub>) étaient positifs, le génome de la souche correspondante était alors entièrement séquencé par technique Sanger (Annexe 6) (180).

### Production du stock de surnageants pour chaque virus

L'amplification des souches a été réalisée après infection des CMSPs et schématisée dans la Figure 34.

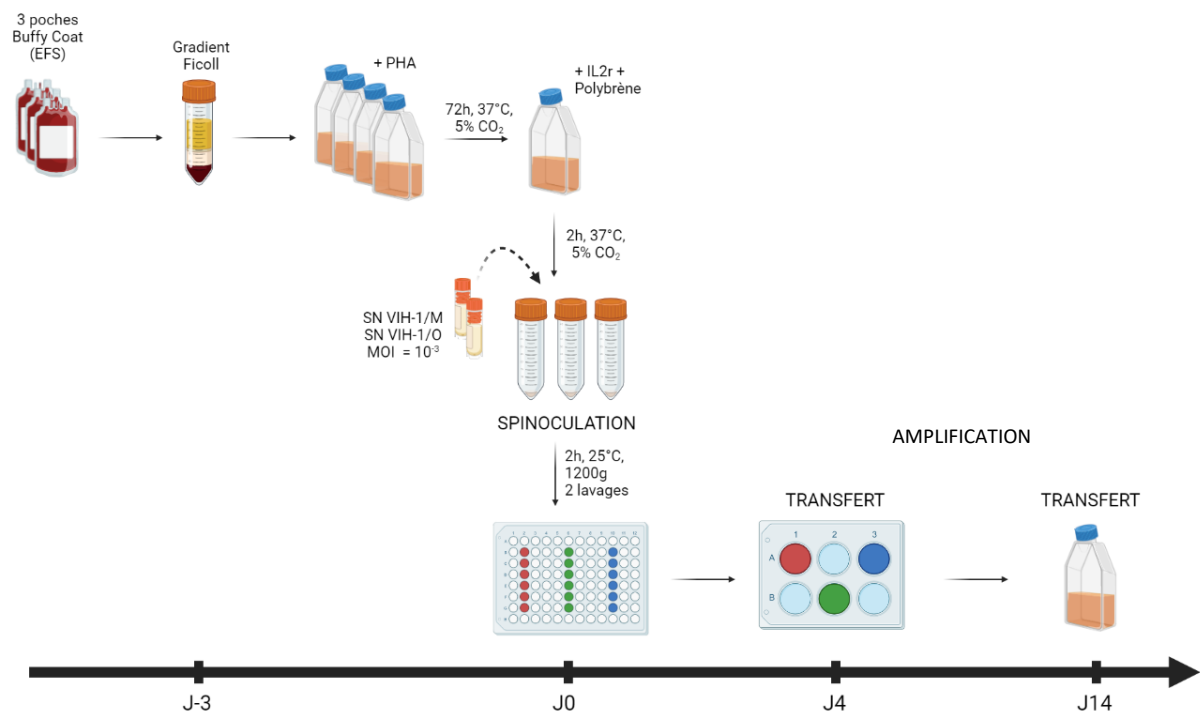


Figure 34 : Schéma expérimental de production des stocks viraux par infection de CMSPs

Créé avec BioRender.com

Les CMSPs ont été récupérées trois jours avant infection (J-3) par centrifugation sur gradient de densité à l'aide du milieu de séparation Ficoll-Hypaque (ThermoFisher Scientific) à partir de couches

leucocytaires. Les CMSPs ont été obtenues de trois poches différentes de donneurs sains récupérées via l'Établissement Français du Sang (EFS) du CHU de Rouen. Après isolement, les CMSPs ont été cultivées dans du RPMI+GlutaMAX™ 1640 (Lonza) avec 10% de SVFd, de la Gentamicine (50µg/mL). Ce milieu a été complété par de la phytohétatoglutinine (0,2µg/mL) (PHA, ThermoFisher Scientific), un agent mitogène utilisé pour activer la multiplication des lymphocytes durant une période de deux jours et provoquer l'agglutination des érythrocytes.

Le jour de l'infection (J0), les cellules activées ont été mélangées et ramenées à une concentration de  $1.10^6$  cellules/mL. Deux millions de cellules ont été centrifugées pendant 2h à 1200xg à température ambiante, avec les souches virales à amplifier, à une MOI (*Multiplicity of Infection*) de 0,001. La MOI représente le ratio de la quantité de virus infectieux (renseignée par la TCID<sub>50</sub>) sur la quantité de cellules réceptrices à infecter (déterminée par comptage). Les volumes de surnageants de culture à utiliser ont été calculés selon la formule suivante :

$$V_{SN} = \frac{MOI \times n_{cellules}}{PFU/mL (=0,7*TCID_{50})}$$

V<sub>SN</sub> : Volume de surnageant (mL)

n<sub>cellules</sub> : Nombre de cellules à infecter (million/mL)

PFU/mL (Plaque-Forming Unit) : Nombre de particules virales capables de former des plaques par unité de volume

TCID<sub>50</sub> : Tissue Culture Infectious Dose

Après centrifugation, les CMSPs ont été lavées deux fois grâce à l'ajout de RPMI + Gentamicine (50µg/mL), centrifugées 5min à 491xg puis remises en suspension avec 2mL RPMI + 10% SVFd + Gentamicine, complété par de l'interleukine II recombinante PROLEUKIN® (540 UI/mL final) (IL2r, Novartis, Bâle, Suisse) pour stimuler la multiplication des CMSPs, ainsi que du polybrène (2µg/mL) (Sigma-Aldrich) pour favoriser la pénétration des virus par modification de la charge électrique des lymphocytes. Après obtention d'une concentration finale de  $1.10^6$  cellules/mL, 300µL de cellules infectées ont été déposées dans une plaque 96 puits en 6 réplicats. Afin d'amplifier les surnageants viraux, l'ensemble des puits d'un même candidat a été mélangé puis transféré dans un puits d'une plaque 6 puits au J7 de la culture, avant d'être transféré dans une T25 au J14. La culture a ensuite été entretenue jusqu'au J28, avec ajout de cellules et renouvellement du milieu une fois par semaine. Des prélèvements ont été effectués tous les 3 à 4 jours, tout au long de la culture, afin de vérifier la présence d'AgP24 par Innostest (Fujirebio). Une mesure de la TCID<sub>50</sub> a également été effectuée à J4, J11, J21 et J25 comme précédemment décrit au paragraphe **Vérification de la production virale** (page 75). L'ensemble des surnageants présentant une quantité d'AgP24 supérieure à 3 Log<sub>10</sub> pg/mL a été mélangé, afin de constituer un stock de surnageant nécessaire à la réalisation des cinétiques de réplication en plusieurs réplicats. Ce stock a été caractérisé par mesure de l'AgP24 en Innostest et de la TCID<sub>50</sub> selon les mêmes modes opératoires déjà décrits.



## b. Résultats

### Génération des CMI parentaux

La quantification des ADN proviraux extraits pour la génération des CMI a été renseignée dans le Tableau 4.

De nombreux essais de clonage ont été réalisés pour chaque souche (Tableau 4). Le bon profil de digestion du clone VIH-1/M (M-A-062021) de sous-type D a été obtenu pour 20% des colonies testées (n=1/5) dès le premier essai de transformation. Les autres clones testés ont été plus difficiles à cloner, avec pour la plupart moins ou environ 10% des colonies ayant le bon profil de digestion (Tableau 4). Tous les clones ayant le bon profil de digestion ont été utilisés pour effectuer une transfection des cellules HEK293T suivi d'une co-culture sur Jurkat X4-R5.

Au total, trois VIH-1/O (BCF120, YBF026 et RBF130) et trois VIH-1/M (AAS638, M-A-062021 et M-F-062021) ont permis d'obtenir une quantification d'AgP24 supérieure au seuil de 400 pg d'AgP24/mL sans dilution au préalable des surnageants de culture (Figure 35). Ces résultats ont pu montrer que les valeurs d'AgP24 pour ces candidats étaient toutes supérieures à 2,6 Log<sub>10</sub> pg/mL entre J4 et J21, sauf pour la souche BCF108 qui n'a donné aucune valeur d'AgP24 et pour RBF130 avec une légère diminution à 2,3 Log<sub>10</sub> d'AgP24. Seules les variations de plus ou moins 0,5 Log<sub>10</sub> d'AgP24 ont été considérées comme significatives.

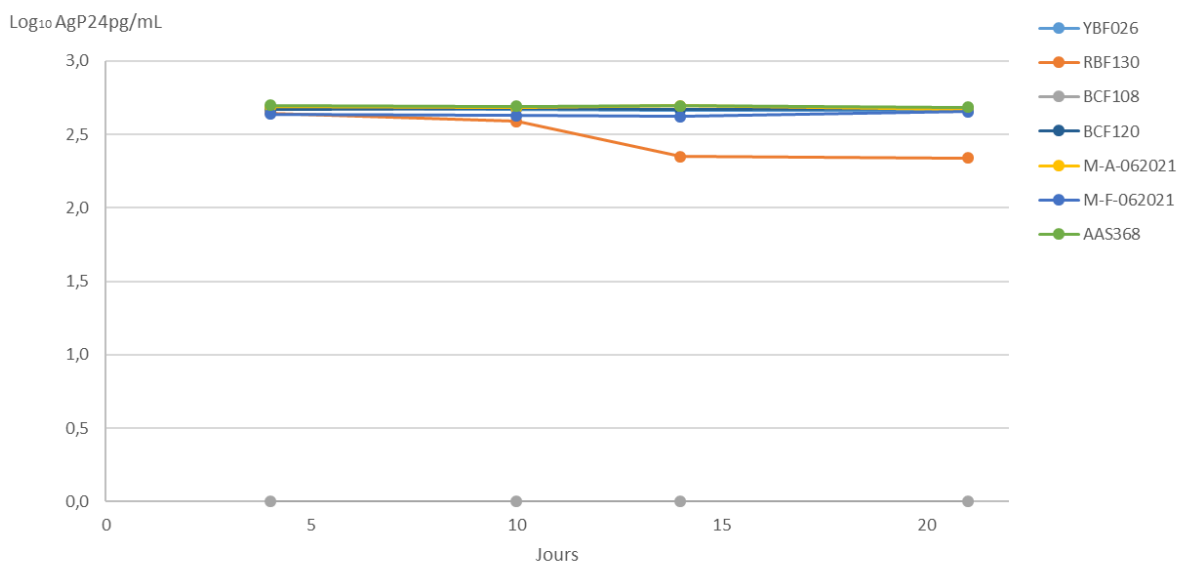


Figure 35 : Graphique montrant la quantité d'AgP24 au cours du temps après transfection/co-culture par les CMI

Plusieurs tentatives de transfection ont été effectuées pour chaque clone, mais seul le candidat VIH-1/M M-A-062021 s'est révélé être infectieux, après une nouvelle transfection, avec l'obtention d'une TCID<sub>50</sub> à J0 de 1 491 TCID<sub>50</sub>/mL (Tableau 4). Cependant, la production des virions sur cellules Jurkat X4-R5 n'a pas permis d'amplifier ce clone car on observait une légère diminution significative ( $\approx 0,6 \text{ Log}_{10} \text{ pg/mL}$ ) de la quantité d'AgP24 au cours du temps (passage de 4,2 à 3,6  $\text{Log}_{10} \text{ pg/mL}$ ). Une quantité insuffisante de surnageant associée à une faible TCID<sub>50</sub> ne nous permettait pas de lancer les cinétiques de répllication en plusieurs réplicats. Une troisième transfection a donc été lancée pour ce clone révélant un titre infectieux de 6 988 TCID<sub>50</sub>/mL dans le surnageant de transfection à J0 et suffisant pour lancer la production d'un stock en amplifiant le virus sur CMSPs.

Tableau 4 : Bilan de la génération des CMI

	YBF026			RBF130			BCF108			BCF120			AAS638				M-A-062021			M-F-062021								
<b>Diversité</b>	VIH-1/O			VIH-1/O			VIH-1/O			VIH-1/O			VIH-1/M				VIH-1/M			VIH-1/M								
<b>Quantification de l'ADN proviral (µg/mL)</b>	159,2			203,2			280,3			NA			-				418,8			171,1								
<b>Taille Fragments A / B (PCR Long range)</b>	5 160 pb / 4 786 pb			5 038 pb / 4 899 pb			5 154 pb / 4 656 pb			NA			5 078 pb / 4 723 pb				5 077 pb / 4 774 pb			5 103 pb / 4 680 pb								
<b>Réaction GA</b>	x1			x5			x2			NA			x1				x1			x2								
<b>Transformation chimique</b>	x2			x6			x3			x5			x2				x1			x3								
<b>Profils de digestion attendu</b>	n = 3/17 (18%)			délétions n = 1/105 (0,9%)			n = 1/59 (1,7%)			n = 3/70 (4,3%)			n = 4/39 (10,3%)				n = 1/5 (20%)			n = 5/44 (11%)								
<b>Enzyme de restriction</b>	AvrII ; SmaI			AflIII ; BstEII ; SmaI ; NdeI			NdeI ; MscI/SalI ; SalI/NotI ; EcoRI			MluI-HF/NotI-HF			BamHI-HF				AvrII			NheI								
<b>Clones candidats</b>	1	11	16	20	98		12			1	2	3	6	8	10	27	5	19	21	34	43	44						
<b>Transfection (Quantification Log<sub>10</sub> AgP24 (pg/mL))</b>	J0	-	> 2,6	0,9	0,7	1,5	0,8	3,2	2,9	2,6	-	-	-	-	4,2	0,8	2,7	> 2,6	> 2,6	0,7								
<b>Co-culture Jurkat X4/R5 (Quantification Log<sub>10</sub> AgP24 (pg/mL))</b>	J4	2,7	NA	NA	-	2,6	0,7	4,3	4,0	3,7	> 2,6	> 2,6	0,7	2,4	4,1	-	2,7	NA	NA	NA								
	J7	-	NA	NA	0,7	-	-	4,4	4,4	3,9	> 2,6	> 2,6	0,7	1,8	3,9	1,5	-	NA	NA	NA								
	J11	2,7	NA	NA	-	2,6	0,8	4,8	4,4	4,3	> 2,6	> 2,6	0,7	1,2	3,8	1,2	2,7	NA	NA	NA								
	J14	2,7	NA	NA	0,8	2,3	0,0	5,3	5,0	4,6	> 2,6	> 2,6	0,7	0,9	3,6	-	2,7	NA	NA	NA								
	J17	NA	NA	NA	NA	-	-	> 5,5	5,2	5,3	NA	NA	NA	NA	NA	NA	NA	-	NA	NA	NA							
	J21	2,7	NA	NA	NA	2,3	0,0	> 5,5	5,4	> 5,5	NA	NA	NA	NA	NA	NA	NA	2,7	NA	NA	NA							
<b>Infectiosité (TCID<sub>50</sub>/mL)</b>	< 50			< 50			NA			78 125			15 625				78 125			< 50			6 988			< 50		

GA : Gibson Assembly

- : les tirets représentent les valeurs non disponibles en raison de l'absence de quantification du surnageant correspondant

NA : non applicable ou indique les prélèvements inexistant (absence de réalisation de co-culture ou quantification du surnageant de transfection uniquement)

Les différentes enzymes de restriction utilisées pour la digestion de contrôle des candidats ont été représentés séparés par des points virgules tandis que les doubles restrictions ont été représentées par un slash.

Chaque date de prélèvement a été indiquée par la lettre J suivie du nombre correspondant au jour où le prélèvement a été effectuée après infection (J0)

> : Indique la limite de quantification des Agp24

Seules les tentatives effectuées avec les bactéries NEB Stable ont été répertoriées dans ce tableau

Concernant les VIH-1/O, les candidats YBF026, RBF130 et BCF120, ayant tous produit de l'AgP24 après transfection, ne se sont pas révélés infectieux après révélation par TCID<sub>50</sub>. Une nouvelle transfection a été réalisée avec tous ces plasmides mais seuls trois candidats VIH-1/O BCF120 ont permis d'obtenir un pic de production d'AgP24 à J21, d'une intensité supérieure à 5 Log<sub>10</sub> pg/mL, associé à une TCID<sub>50</sub> de 78 125 TCID<sub>50</sub>/mL (clone 1 et 3) et 15 625 TCID<sub>50</sub>/mL (clone 2). Contrairement au clone VIH-1/M M-A-062021, nous avons pu observer une augmentation de la quantité d'AgP24 après transfection puis infection des cellules Jurkat X4-R5 (Tableau 4, Figure 36).

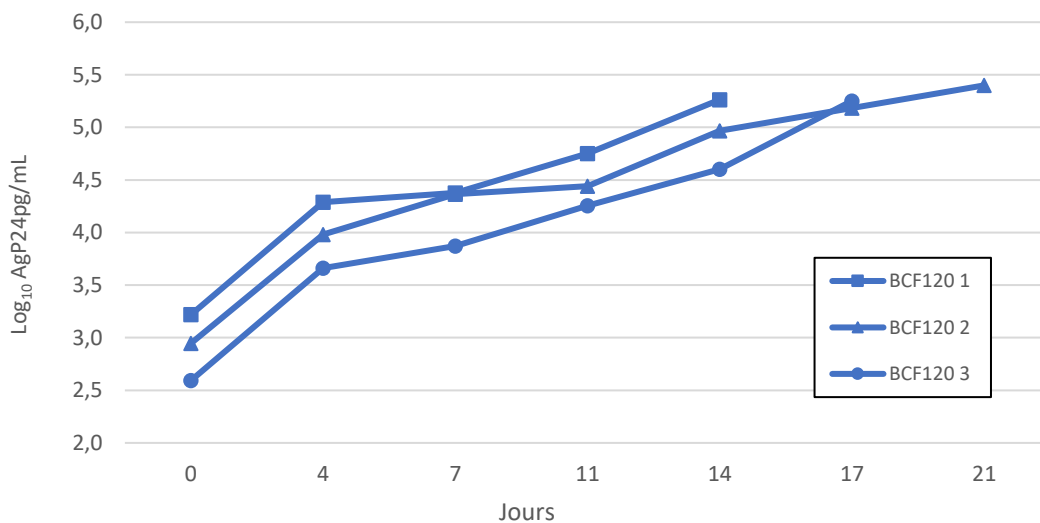


Figure 36 : Quantité d'AgP24 obtenue au cours du temps après transfection co-culture de pBCF120

Les trois candidats obtenus après transformation chimique des bactéries avec le CMI pBCF120 ont été mis en culture après transfection de cellules HEK293T et co-culture avec des cellules Jurkat X4-R5. Les points nécessitant une dilution plus importante du surnageant en Innostest n'ont pas été réalisés (J17 et J21 pour le candidat 1 et J21 pour le candidat 3) et dépassait le point le plus haut de la gamme d'étalonnage.

Les deux souches VIH-1/M (M-A-062021) et VIH-1/O (BCF120) ont été amplifiées à l'aide de CMSPs, avec ajout de cellules une fois par semaine pendant 28 jours, afin de constituer un stock de surnageant. Les deux souches présentaient une TCID<sub>50</sub> positive suffisante pour infecter les cellules à une MOI de 10<sup>-3</sup>. Après amplification, chaque prélèvement a été quantifié par Innostest et ceux présentant une quantité d'AgP24 supérieure à 3 Log<sub>10</sub> pg/mL ont été mélangés pour constituer un stock. Au total trois surnageants ont été sélectionnés pour M-A-062021, avec des quantités d'AgP24 comprises entre 3,9 et 4,5 Log<sub>10</sub> pg/mL et sept pour BCF120, avec des quantités d'AgP24 comprises entre 5,2 et 6,1 Log<sub>10</sub> pg/mL. Ces mélanges ont ensuite été caractérisés par quantification de l'AgP24 toujours en Innostest et par la mesure de la TCID<sub>50</sub> (Tableau 5).

Tableau 5 : Caractéristiques des deux CMI parentaux VIH-1/M et VIH-1/O

	Profil	Diversité	Quantité d'AgP24 Innotest (pg/mL)	TCID <sub>50</sub> sur TZM-bl (TCID <sub>50</sub> /mL)
<b>M-A-062021</b>	VIH-1/M	Sous-type A	250 783 (5,4 Log <sub>10</sub> )	3 125
<b>BCF120</b>	VIH-1/O	Sous-groupe H	27 254 (4,4 Log <sub>10</sub> )	174 693

Le génome complet des deux souches a été séquencé à partir des surnageants de culture issus de la production des stocks et a permis de montrer qu'il s'agissait de la même séquence que le CMI de départ ayant servi pour la transfection des cellules HEK293T, validant la production et l'amplification de ces deux clones.

### Génération des CMIC

Une fois l'obtention de deux CMI parentaux VIH-1/M (pM-A-062021) et VIH-1/O (BCF120), nous avons entrepris des tests de clonage des deux profils de CMIC OMOM et MOMO, toujours à l'aide de la méthodologie Gibson Assembly. Au total, 48 colonies ont été obtenues pour le CMIC de profil OMOM et 120 pour le profil inverse MOMO. Malgré l'absence du profil de digestion attendu, 22,9% et 9,2% des colonies pour OMOM et MOMO respectivement, présentaient une taille totale attendue pour le CMIC correspondant. Etant donné ce résultat, ces CMIC ont été considérés comme des candidats et ont été testés en culture. La mesure d'AgP24 après transfection et co-culture a révélé une production d'AgP24 supérieure à 2,6 Log<sub>10</sub> d'AgP24/mL. Après vérification par la mise en culture sur CMSPs comme pour les souches parentales VIH-1/M et VIH-1/O, ces clones ont permis l'obtention d'une quantité d'AgP24 de 4 et 3,6 Log<sub>10</sub> pg/mL ainsi qu'une TCID<sub>50</sub> de 625 et 280 TCID<sub>50</sub>/mL pour OMOM et MOMO respectivement. Cependant, le séquençage des génomes complets a révélé l'absence de point de cassure aux positions attendues, montrant l'amplification d'une souche de VIH-1/M, représentative de la souche parentale VIH-1/M M-A-062021, sans point de cassure dans le gène *vpr*.

### Clonage des CMIC par la société Genecust

Plusieurs combinaisons de CMI VIH-1/M et VIH-1/O ont été testées par la société extérieure Genecust, sans permettre d'aboutir à un clone stable.

### c. Discussion

Notre objectif initial de comparaison de potentiels réplicatifs de deux formes recombinantes de profils inverses entre elles et avec leurs formes parentales n'a pas pu être atteint, malgré de multiples tentatives de clonage et/ou d'amplification en culture cellulaire. De nombreuses étapes critiques ont été identifiées a posteriori dans le protocole d'obtention de ces clones.

Concernant les étapes du clonage, tout d'abord, trois souches d'*E. coli* (*XL-10 gold*, *Dh5 $\alpha$*  et *108*) ont été préalablement testées pour l'étape de transformation chimique, avant la sélection finale des bactéries NEB® Stable Competent *E. coli*. Une amélioration du rendement a été observée avec ces bactéries, d'ailleurs recommandées pour le clonage d'inserts lentiviraux, puisque peu stables et possédant des éléments répétés, tels que les LTRs du VIH (177). Cependant, un manque de stabilité des candidats testés a été observé, conduisant à l'hypothèse d'une pression de sélection négative dans la bactérie. En effet, des profils de restriction incorrects ont été obtenus pour la majorité des clones testés avec seulement 5% de candidats ayant le bon profil attendu ( $n = 3/339$ ). Ces profils incorrects incluaient à la fois des clones présentant quelques bases délétées ou de multiples délétions, mais aussi des clones avec une excision complète du génome de la souche pour ne laisser que le plasmide vecteur vide.

Au total, sur 339 tentatives de clonage réalisées pour les CMI, seulement 1,2% des clones ont permis d'aboutir à la génération d'un virus infectieux. Sur les 88 tentatives de génération des CMI VIH-1/M, 11,4% des clones ont montré le bon profil de digestion ( $n=10/88$ ), dont 10% étaient infectieux au final ( $n=1/10$ ). Sur les 251 tentatives de génération des CMI VIH-1/O, seuls 3,2% des clones ont montré le bon profil de digestion ( $n=8/251$ ), mais dont 50% étaient infectieux au final ( $n=4/8$ ). Ce résultat suggère que la réussite du clonage dépend probablement de la séquence de la souche de VIH insérée au sein du plasmide et est donc liée à la diversité génétique des souches testées. Notre objectif de démontrer un impact de la diversité des souches par la génération des clones correspondant n'a pas pu être mis en évidence avec ces résultats, même si ces derniers suggéraient que toutes les souches des VIH ne peuvent pas être clonées par la méthodologie testée au cours de ce travail.

La fragmentation des génomes parentaux VIH-1/M et VIH-1/O, au début de la région Vif, avait pour objectif d'obtenir deux fragments de taille équivalente, comme ce qui avait déjà été montré dans une autre étude et devait faciliter l'assemblage des CMI selon la méthode de Gibson Assembly (175, 181). Malgré une modélisation de ces CMI, représentative de la littérature, avec une localisation du point de cassure à l'extrémité du gène *vif* et dans une région conservée assurant une homologie de séquence

de part et d'autre du point de cassure, cette cassure artificielle, non décrite *in vivo*, ne favorisait peut-être pas le réassemblage final (122, 126, 142).

Ce projet a néanmoins permis de générer un nouveau CMI VIH-1/M, appelé pM-A-062021 et correspondant à un sous-type A. A notre connaissance, ce CMI s'ajoute à quelques autres clones de sous-type A produits dans la littérature (182–184), il sera donc partagé avec la communauté scientifique via le CFAR du NIBSC. La souche VIH-1/O YBF026 a, quant à elle, permis d'obtenir 18% de clones ayant le bon profil de digestion, mais tous se sont révélés non infectieux, même après infection de CMSPs. A ce jour, seulement trois CMI VIH-1/O sont disponibles : BCF120 et RBF206 déjà testés au cours de cette étude et d'une précédente au laboratoire, ainsi qu'un autre clone construit par l'équipe de *Tebit et al*, correspondant à un mélange de deux souches de VIH-1/O et donc non utilisable dans notre contexte d'étude (86, 185).

Concernant la génération des CMIC, les deux profils OMOM et MOMO ont été assemblés à partir des CMI parentaux VIH-1/M M-A-062021 et VIH-1/O BCF120. Aucun bon profil de digestion n'a été observé pour ces deux CMIC OMOM et MOMO. Pour la totalité des colonies obtenues après transformation chimique, un profil de digestion identique à la souche parentale VIH-1/M (M-A-062021) a été observé. L'absence de point de cassure dans le gène *vpr*, identifié lors du séquençage des génomes complets après culture cellulaire a démontré la complexité de l'étape d'assemblage des candidats et a remis en cause l'efficacité du protocole mis en place au cours de cette étude. L'hypothèse formulée à cette étape était qu'il y avait persistance des génomes complets des CMI VIH-1/M et VIH-1/O en plus de la forme recombinantes VIH-1/MO au moment de l'assemblage du produit GA. Une sélection plus importante de la forme parentale VIH-1/M dans la bactérie, pouvait s'expliquer par la visualisation de 20% de bons profils de digestion pour la souche M-A-062021 contre seulement 4,3% pour la souche VIH-1/O BCF120, au cours des essais de clonage des CMI. Après transformation chimique, le clone le plus stable est sûrement sélectionné par les bactéries, au détriment des autres présents dans l'assemblage Gibson Assembly.

Après de nombreuses tentatives et plusieurs adaptations des protocoles de clonage au laboratoire, nous avons fait appel à une société extérieure afin de générer les CMIC d'intérêt à partir des CMI parentaux disponibles. L'absence d'amplification possible pour certains CMI et un très faible pourcentage de bon profil de digestion pour d'autres ont rendu la génération des CMIC impossible à réaliser par cette société. Ces résultats démontraient la non faisabilité du projet associé à un manque de stabilité de nos souches dans les inserts et bactéries testées.

Concernant l'étape de la culture cellulaire, l'utilisation de cellules Jurkat X4-R5, modifiées et difficilement cultivables, a rendu les expérimentations difficiles à reproduire. En effet, ces cellules Jurkat X4-R5 ont été particulièrement difficiles à décongeler après passage dans de l'azote liquide, avec plus de 70% de mortalité cellulaire estimée au moment de la décongélation. L'utilisation de ces cellules a été remise en cause, du fait d'un manque de reproductibilité entre expérimentations successives, à partir d'un même candidat testé plusieurs fois et du fait de l'absence d'amplification des souches après infection, pourtant bien amplifiées sur CMSPs. L'hypothèse de l'impact, sur l'expression de nos souches, d'une différence d'expression entre les corécepteurs pour ces cellules avait déjà été émise dans une précédente étude (137). Un test de cytométrie en flux a été réalisé afin de vérifier la présence des deux corécepteurs sur cette lignée Jurkat X4-R5, à l'aide d'une collaboration avec le laboratoire Inserm UMR 1234 PANTHER de l'Université de Rouen Normandie. Cette analyse a révélé la présence de seulement 62,5% de cellules Jurkat ayant réellement le double tropisme X4-R5 (résultats non montrés). L'utilisation des CMSPs ont donc été justifiée par ces résultats. Concernant la technique de TCID<sub>50</sub> utilisée pour quantifier l'infectiosité de ces souches, celle-ci a été adaptée à partir d'un protocole de l'unité Inserm UMR 1259 MAVIVH de l'Université de Tours. Cette technique avait été testée au préalable sur un panel de VIH-1/O au laboratoire, afin de vérifier son efficacité sur une large diversité de VIH. La présence des deux corécepteurs à la surface de ces cellules permettait de s'affranchir de la diversité de nos souches et nous permettait de les comparer entre elles à l'aide de ce marqueur.

Depuis la construction du premier CMI de VIH-1 complet pNL4-3 (sous-type B) en 1986 (186), ces outils se sont révélés très précieux car ils permettent l'étude de la diversité, à partir de souches clonales. De nombreux clones de différents sous-types ont été produits depuis, adaptés pour diverses applications telle que l'étude des propriétés de tropisme ou de réplication (181–183, 187–189). Malgré la complexité de génération de ces clones, il est important de développer des outils nous permettant d'étudier l'impact de la diversité des souches, car peu de CMI parentaux sont disponibles selon la littérature.





## 2. Étude *in vitro* du potentiel répliatif de deux formes VIH-1/MO

Deux souches recombinantes VIH-1/MO avaient été isolées en culture lors de travaux précédents : RBF208 (110) et OM-TB-2016 (137). Ces souches ne possédaient ni les mêmes points de cassure, ni les mêmes virus parentaux de groupes M et O. Ainsi, l'objectif de comparer le potentiel répliatif *in vitro* de deux formes recombinantes VIH-1/MO possédant des profils différents, permettait de conclure sur une différence de potentiel répliatif en fonction du profil de recombinaison, que ce soit en terme de diversité génétique ou en terme de positionnement du point de cassure, ce qu'aucun travail n'avait encore étudié jusqu'alors.

### a. Méthodologie

#### *Choix et caractéristiques des souches recombinantes VIH-1/MO*

La forme recombinante RBF208 (numéro d'accèsion GenBank GQ351296) possède deux points de cassure. Le premier, de motif OM, est localisé au niveau de la région chevauchante entre les gènes accessoires *vif-vpr*, entre les bases 5540 et 5576 et le second, de motif MO, dans la région GP41 de l'enveloppe, entre les bases 8278 et 8296 par rapport à la souche de référence HxB2. Le profil de recombinaison de type OMO de RBF208 est représenté dans la Figure 37 (135). Les fragments de VIH-1/O appartiennent au sous-groupe H, alors que celui de VIH-1/M appartient au sous-type D. Un tropisme R5 a été prédit pour cette souche. Un stock de surnageant de culture avait préalablement été constitué au laboratoire.

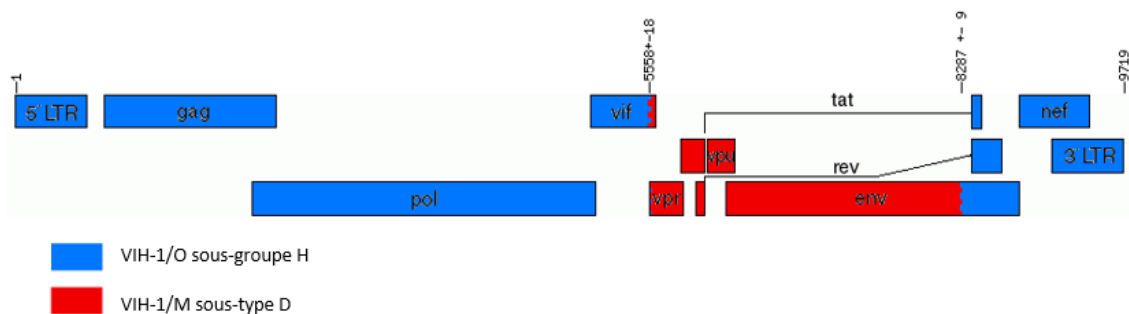


Figure 37 : Schéma de la forme recombinante RBF208

Profil moléculaire de la forme recombinante RBF208. Les fragments de VIH-1/O sont représentés en bleu tandis que les fragments de VIH-1/M sont en rouge. Chaque position a été indiquées par rapport à la souche de référence HxB2 (n° d'accèsion GenBank K03455).

La seconde forme recombinante étudiée était OM-TB-2016, produite à partir d'un CMIC synthétisé en 2017, au cours des travaux préliminaires (137). Cette souche possédait un point de cassure de motif OM dans le gène *vpr*, entre les bases 5661 et 5667 et un second de motif MO dans les LTRs (Figure 38). Les fragments de VIH-1/O appartenaient au sous-groupe T et ceux de VIH-1/M au sous-type B. Un tropisme X4 a été prédit pour cette souche. Un stock a également été constitué et caractérisé pour cette souche avant notre étude par amplification sur cellules Jurkat X4-R5 (137).

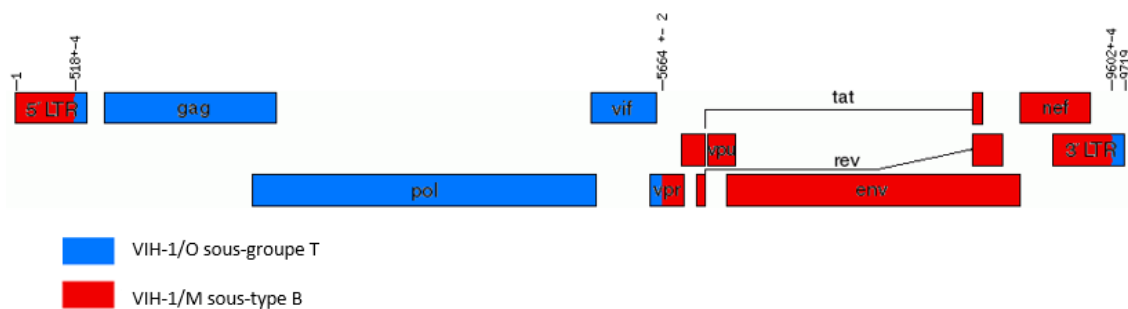


Figure 38 : Schéma de la forme recombinante OM-TB-2016

Profil moléculaire de la forme recombinante OM-TB-2016. Les fragments de VIH-1/O sont représentés en bleu tandis que les fragments de VIH-1/M sont en rouge. Chaque position a été indiquée par rapport à la souche de référence HxB2 (n° d'accèsion GenBank K03455).

Le volume de surnageant pour ces deux souches devait être suffisant pour lancer des cinétiques à une MOI définie (page 77). Pour cela, nous avons re-caractérisé les stocks initialement produits pour chacune des deux souches par quantification de l'AgP24 en Innostest selon les recommandations fournisseur et de la TCID<sub>50</sub>, selon la technique de Montefiori déjà décrite précédemment page 75 (**Vérification de la production virale**).

### **Conditions d'amplification des deux VIH-1/MO**

Un test préliminaire a été réalisé, afin de déterminer les meilleures conditions de cultures pour le lancement des cinétiques. Deux MOI, une MOI de 10<sup>-2</sup> (0,001 UI/cellule) versus une MOI de 10<sup>-4</sup> (0,0001 UI/cellule), ont été testées pour la souche OM-TB-2016, par comparaison avec la cinétique de répllication de sa souche parentale VIH-1/M 89.6. Ce choix se basait sur la quantité de surnageant à prélever pour réaliser nos tests et selon la réalisation de cinétiques antérieures à une MOI de 10<sup>-1</sup> (0,01 UI/cellule) pour ces souches (137). Ces MOI ont été calculées avec les valeurs initiales de TCID<sub>50</sub> obtenues pour chaque stock des deux souches à tester selon les calculs suivants :

**Calculs de MOI pour les deux souches à tester :**

$$V_{SN} OM_{GC} = \frac{(1.10^{-2} \times 5.10^6)}{0,7 \times 45\ 688} = 1,6 \text{ mL}$$

$$V_{SN} 89.6 = \frac{(1.10^{-2} \times 5.10^6)}{0,7 \times 133\ 592} = 267,3 \mu\text{L}$$

$$V_{SN} OM_{GC} = \frac{(1.10^{-4} \times 5.10^6)}{0,7 \times 45\ 688} = 15,6 \mu\text{L}$$

$$V_{SN} 89.6 = \frac{(1.10^{-4} \times 5.10^6)}{0,7 \times 133\ 192} = 2,7 \mu\text{L}$$

Le choix de prendre ces deux souches OM-TB-2016 et 89.6 a reposé sur la quantité d'échantillons disponibles pour effectuer des tests, ainsi que sur la base des travaux de culture précédemment réalisés (137). Le choix des cellules s'est néanmoins porté sur les CMSPs après avoir mis en évidence une différence d'expression du corécepteur R5 chez la lignée Jurkat X4-R5 par cytométrie en flux.

Les CMSPs ont été récupérées puis activées selon les mêmes modalités que décrites précédemment page 76 (**Production du stock de surnageants pour chaque virus**). Le jour de l'infection (J0), les cellules ont été centrifugées puis remises en suspension avec du RPMI+Glutamax + 10% SVFd + Gentamicine (50µg/mL) + IL2 (540 UI/mL final) + polybrène (2µg/mL) afin de stimuler la multiplication des CMSPs pendant 2h. Après incubation, des surnageants de culture de OM-TB-2016 et 89.6 ont été décongelés, afin de procéder à l'infection de  $5.10^6$  CMSPs préalablement déposées dans un Falcon, aux deux MOI à tester. Les cellules ont été incubées avec les surnageants viraux 4h à 37°C avec 5% de CO<sub>2</sub>. Le surnageant viral résiduel de chaque Falcon a été éliminé par deux lavages du culot cellulaire avec du RPMI+Glutamax + Gentamicine (50µg/mL), avant remise en suspension.

Les cellules infectées ont ensuite été transférées dans une plaque 6 puits, dans un volume final de 5mL de RPMI+Glutamax + 10% SVFd + Gentamicine (40µg/mL) + IL2 (540 UI/mL final) + polybrène (2µg/mL) pour une concentration finale de  $1.10^6$  cellules/mL. La plaque a ensuite été incubée à 37°C, 5% de CO<sub>2</sub> pendant 25 jours. Des prélèvements ont été effectués trois fois par semaine pendant ces 25 jours afin de suivre la production d'AgP24, par Innostest. Le nombre de prélèvement à effectuer pour nos cinétiques a été déterminé après révélation de cette culture test, afin d'estimer le plus précisément possible le délai d'obtention du pic de réplication de chaque souche. Le nombre de réplicat lancés a été déterminé pour permettre de conclure sur la reproductibilité des cinétiques à partir de mêmes conditions de culture.

Le taux de mortalité cellulaire des CMSPs a été calculé avant l'infection (J-3), le jour de l'infection (J0) et le dernier jour de culture (J25). Ce taux de mortalité a été estimé après dilution de 25µL de surnageant avec 25µL de bleu Trypan (Sigma-Aldrich) 0,4% dilué au 1/10<sup>ème</sup>, puis comptage des cellules sur lame de Kova (ThermoFisher Scientific, France). Les cellules apparaissant en bleu

indiquaient les cellules mortes. Le rendement de la production des CMSPs après stimulation par PHA + IL2 a été estimé par le calcul suivant :

**Calculs du rendement de production des CMSPs après stimulation par la PHA + IL2 :**

$$\text{Rendement} = \frac{(\text{nombre de cellules totales mises en PHA avant infection})}{(\text{nombre de cellules comptées après incubation 2h dans PHA + IL2})} * 100$$

La MOI retenue pour notre étude et donc le volume requis pour le lancement des cinétiques ont été calculés, après re-caractérisation des TCID<sub>50</sub> obtenues pour chaque stock des deux souches à tester et selon les résultats de cette culture test, selon les calculs suivants :

**Calculs de MOI pour les deux souches à tester :**

$$V_{SN} \text{ RBF208} = \frac{(1.10^{-4} \times 5.10^6)}{0,7 \times 78 \ 125} = 9,1\mu L$$

$$V_{SN} \text{ OM\_GC} = \frac{(1.10^{-4} \times 5.10^6)}{0,7 \times 26 \ 362} = 27,1\mu L$$

### **Cinétiques de réplication**

Chaque cinétique a été lancée avec un lot de trois poches de sang différentes et les CMSPs ont été récupérées puis activées selon les mêmes modalités décrites précédemment page 76 (**Production du stock de surnageants pour chaque virus**).

Le jour du lancement des cinétiques de réplication (J0), les CMSPs ont été centrifugées puis remises en suspension avec du RPMI+Glutamax + 10% SVFd + Gentamicine (50µg/mL) + IL2 (540 UI/mL final) + polybrène (2µg/mL) afin de stimuler la multiplication des CMSPs pendant 2h. Après incubation, des surnageants de culture de RBF208 et OM-TB-2016 ont été décongelés, afin de procéder à l'infection de 5.10<sup>6</sup> CMSPs préalablement déposées dans un Falcon, à une MOI de 10<sup>-4</sup> pour les deux souches d'intérêt. Les cellules ont été incubées avec les surnageants 4h à 37°C avec 5% de CO<sub>2</sub>. Le surnageant viral résiduel de chaque Falcon a été éliminé par deux lavages du culot cellulaire avec du RPMI+Glutamax + Gentamicine (50µg/mL), avant remise en suspension. Un premier prélèvement a alors été réalisé après ces lavages pour constituer le J0 de la cinétique. Les cellules infectées ont ensuite été transférées dans une plaque 6 puits, dans un volume final de 5mL de RPMI+Glutamax + 10% SVFd + Gentamicine (40µg/mL) + IL2 (540 UI/mL final) + polybrène (2µg/mL) pour une concentration finale de 1.10<sup>6</sup> cellules/mL. La plaque a ensuite été incubée à 37°C, 5% de CO<sub>2</sub> pendant une dizaine de jours.

Quatre mono-infections ont été réalisées et appelées M1, M2, M3 et M4 dans l'ordre chronologique de réalisation. Les cultures ont été stoppées à J10 pour M1 et J11 pour les autres cinétiques. Un témoin cellulaire a été réalisé en suivant le même protocole d'infection, mais sans ajout de surnageant viral et déposé sur la plaque 6 puits. Le surnageant a été récolté après 5min de centrifugation à 6000rpm, afin d'enlever les cellules, deux fois par jour vers 9h30 et 17h30. Pour chaque puits, ces prélèvements ont été effectués 6 jours sur 7 pour M1, puis 7 jours sur 7 pour les autres cinétiques après analyse des résultats de M1 nécessitant une adaptation de la fréquence des prélèvements. Les volumes récoltés ont été congelés pour révélation ultérieure.

Pour chaque cinétique, le taux de mortalité cellulaire a été calculé avant l'infection (J-3), le jour de l'infection (J0) et le dernier jour de culture (J10/J11) selon le même protocole que décrit précédemment page 87 (**Conditions d'amplification des deux VIH-1/MO**). L'ensemble du protocole utilisé pour effectuer nos quatre cinétiques a été schématisé dans la Figure 39.

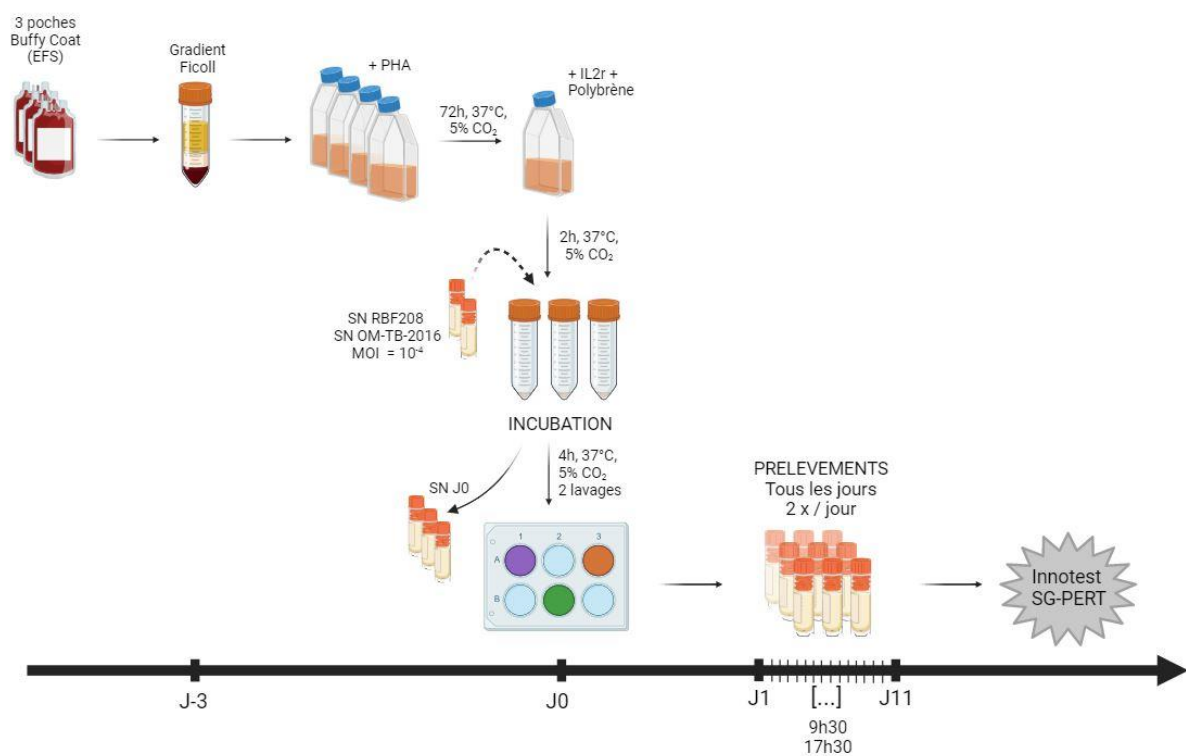


Figure 39 : Protocole expérimental de lancement des cinétiques de réplication pour RBF208 et OM-TB-2016 après infection de CMSPs

Créé avec BioRender.com

## Révélation des cinétiques

Chaque prélèvement, issu de chacune des quatre cinétiques, a été révélé au moyen de deux marqueurs : la mesure de la quantité d'AgP24, reflet de la production de virions et l'activité de la TI, mesurant une activité enzymatique virale et donc reflet de la viabilité virions produits.

La quantité d'AgP24 dans ces surnageants a été déterminée par la méthode Innostest (Fujirebio) selon les recommandations du fournisseur.

L'activité de la TI, a, quant à elle, été déterminée par la technique de SG-PERT, après adaptation du protocole de *Vermeire et al.* (190, 191) (Annexe 7). Le surnageant acellulaire non dilué a été lysé, puis ajouté à un mélange réactionnel contenant une matrice ARN, des amorces complémentaires de cette matrice ainsi qu'un mélange SYBR Green I. Au cours d'une unique réaction, les molécules de TI contenues dans les virions sont relarguées dans le surnageant et peuvent alors convertir un ARN matrice en ADNc. Ces ADNc sont ensuite quantifiés par une amplification de PCR quantitative (qPCR), en utilisant des sondes spécifiques de la matrice, marquées par des fluorophores pour la détection de signal. Le niveau d'activité de la transcriptase inverse est donc représenté par la quantité d'ADNc synthétisée au cours du temps dans le surnageant viral et permet de mesurer la quantité de particules rétrovirales. Le principe de cette technique est schématisé dans la Figure 40 ci-dessous.

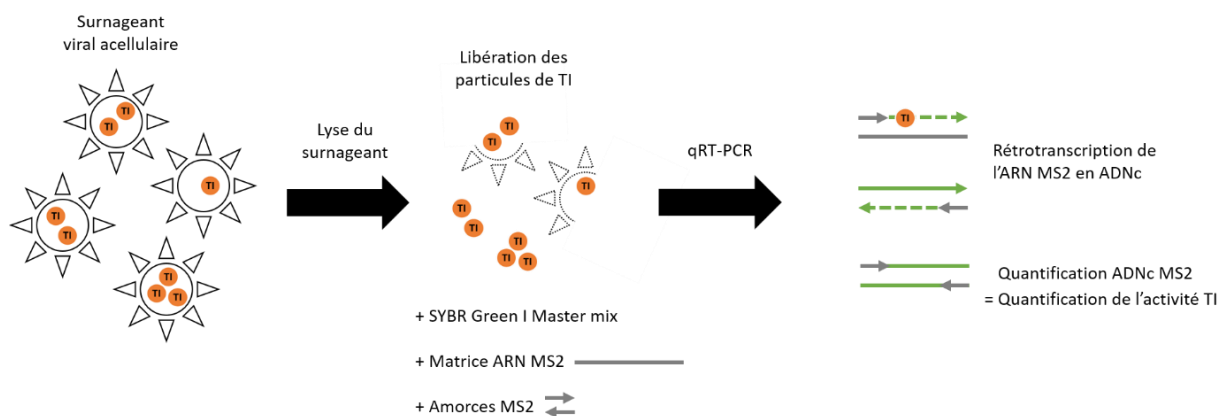


Figure 40 : Principe du SG-PERT

D'après *Vermeire et al* (191).

Le surnageant contenant le rétrovirus acellulaire est lysé et ajouté à un mélange réactionnel contenant la matrice d'ARN MS2, des amorces complémentaires de cette matrice MS2 et un mix de qPCR SYBR Green I. Au cours d'une réaction en une étape, les molécules de transcriptase inverse (TI) dérivées des particules rétrovirales peuvent alors convertir l'ARN MS2 matrice en ADNc qui sera ensuite quantifié par amplification qPCR de l'ADNc MS2. La quantité d'ADNc synthétisé représente le niveau d'activité de la TI dans le surnageant viral.

### *Comparaison des cinétiques de réplifications des deux formes recombinantes*

La reproductibilité entre les différentes cinétiques d'une même souche a été déterminée par la mesure de l'écart entre chaque point des réplicats d'une même souche. Cet écart, appelé  $\Delta_{\max-\min}$ , a été calculé en réalisant la soustraction de la valeur la plus haute par celle la plus basse. Seules les variations de plus ou moins 0,5  $\text{Log}_{10}$  ont été considérées comme significatives. L'écart obtenu entre réplikat d'une même souche a été mis en regard avec la mortalité cellulaire observée avant et après infection des CMSPs.

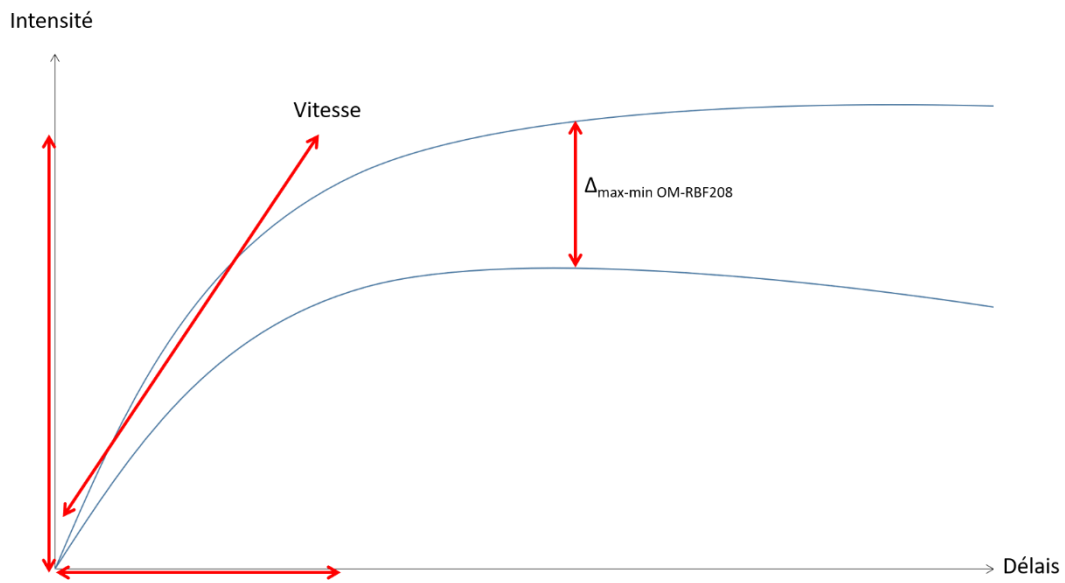
Une moyenne a été effectuée entre chaque point correspondant de chaque cinétique, pour chaque souche et appelée « moyenne réplikat ». Afin de prendre en compte l'obtention de deux profils de cinétique différents, une seconde moyenne a été calculée, ne prenant en compte que la valeur correspondant aux pics de ces différentes cinétiques et a été appelée « moyenne pic ». Ces pics de réplification ont été définis comme étant le point le plus haut, séparant la phase exponentielle du plateau de réplification.

Afin de comparer nos souches entre elles, trois paramètres ont été étudiés pour les deux marqueurs (Figure 41) :

- ⇒ Le délai d'apparition du pic de réplification (en heures) de la souche a été déterminé par le point où la phase exponentielle de réplification atteignait un plateau.
- ⇒ L'intensité du pic a été calculée par la production d'AgP24 (en  $\text{Log}_{10}$  pg/mL) et l'activité de la TI (en  $\text{Log}_{10}$  mU TI/mL).
- ⇒ La vitesse de réplification (en  $\text{Log}_{10}$  pg/mL/heure pour l'AgP24 ou en  $\text{Log}_{10}$  mU TI/mL/heure pour l'activité de la TI), correspondant au coefficient directeur de la pente commençant au début de la phase exponentielle jusqu'au du pic de réplification, a également été calculée.

La comparaison des cinétiques entre les souches OM-TB-2016 et RBF208 a été réalisée en prenant en compte la moyenne réplikat. Les écarts calculés entre les valeurs ont été déterminés par la soustraction de la valeur obtenue la plus haute, par celle la plus basse entre les différentes cinétiques des deux souches, appelée  $\Delta_{\max-\min \text{ OM-RBF208}}$  (Figure 41).





**Figure 41** : Paramètres de comparaison entre les souches OM-TB-2016 et RBF208

En bleu sont représentées les courbes moyennes théoriques des deux souches OM-TB-2016 et RBF208 après avoir effectué la moyenne de chaque réplicat pour chaque souche. Les paramètres de comparaison pour les deux marqueurs testés (AgP24 et TI) ont été mis en rouge.

## b. Résultats

### Caractérisation des souches recombinantes VIH-1/MO

Pour la souche OM-TB-2016, la caractérisation du stock de surnageant nécessaire à la réalisation des différentes cinétiques a montré une quantité d'AgP24 de 532 977 pg/mL (5,7 Log<sub>10</sub>), associée à une TCID<sub>50</sub> de 26 362 TCID<sub>50</sub>/mL. Pour la souche RBF208, la caractérisation de la quantité d'AgP24 était de 784 991 pg/mL (5,9 Log<sub>10</sub>), associée à une TCID<sub>50</sub> de 78 125 TCID<sub>50</sub>/mL. L'ensemble des données de caractérisation du stock pour lancement des cinétiques de réplication est représenté dans le Tableau 6.

Tableau 6 : Caractéristiques des deux formes à comparer OM-TB-2016 et RBF208

	Profil de recombinaison	Diversité des fragments parentaux	Prédiction tropisme	AgP24 (pg/mL)	TCID <sub>50</sub> sur TZM-bl (TCID <sub>50</sub> /mL)	Localisation des points de cassure
<b>RBF208</b>	OMO	VIH-1/O-H VIH-1/M-D	R5	784 991 (5,9 Log <sub>10</sub> )	78 125	<i>vif-vpr</i> / 5540-5576 pb <i>gp41</i> / 8278-8296 pb
<b>OM-TB-2016</b>	MOMO	VIH-1/O-T VIH-1/M-B	X4	532 977 (5,7 Log <sub>10</sub> )	26 362	<i>vpr</i> / 5661 – 5667 pb <i>LTRs</i> / 513 – 522 pb

Pour la même quantité de particule virale, représentée par la quantification en AgP24 d'environ 5,8 Log<sub>10</sub> pg/mL d'AgP24 pour les deux souches, une différence d'infectiosité a été mise en évidence par la technique de TCID<sub>50</sub> sur TZM-bl.

### Conditions d'amplification des deux virus recombinants VIH-1/MO

Le taux de mortalité cellulaire avant infection (J-3) était de 13% contre 6,5% le jour de l'infection (J0), soit un rendement fort de 144% après stimulation. Le dernier jour de culture (J25), une mortalité cellulaire de 45% et 60% a été observée pour la souche OM-TB-2016 et de 55% et 61% pour la souche RBF208, à MOI 10<sup>-2</sup> et 10<sup>-4</sup> respectivement. Ces résultats montraient un faible écart de mortalité cellulaire entre les deux souches, avec une mortalité plus élevée à la MOI de 10<sup>-2</sup> comparativement à la MOI de 10<sup>-4</sup>.

Le test préliminaire d'infection de CMSPs par les souches OM-TB-2016 et de 89,6, à MOI 10<sup>-2</sup> versus 10<sup>-4</sup> a permis de fixer les conditions de lancement des cinétiques à la MOI de 10<sup>-4</sup> (Figure 42). Peu de différences ont en effet été observées entre les deux MOI testées, avec l'obtention d'un pic de production d'AgP24 vers J4 pour les deux souches, à une intensité de 6 Log<sub>10</sub> pg/mL d'AgP24, en faveur

de l'utilisation de la MOI la plus faible consommant moins de surnageant. La réalisation de trois prélèvements par semaine a montré que cette fréquence d'analyse n'était pas suffisante pour l'interprétation du délai d'apparition du pic de répliation, avec un plateau atteint très rapidement entre J2 et J4 dans ces conditions de culture, en faveur de la réalisation de prélèvements plus rapprochés dans le temps. Enfin, la réalisation de ce test préliminaire a montré que le temps de culture pouvait être diminué à une dizaine de jours, avec l'obtention d'un plateau de J4 à J25 pour les deux souches testées.

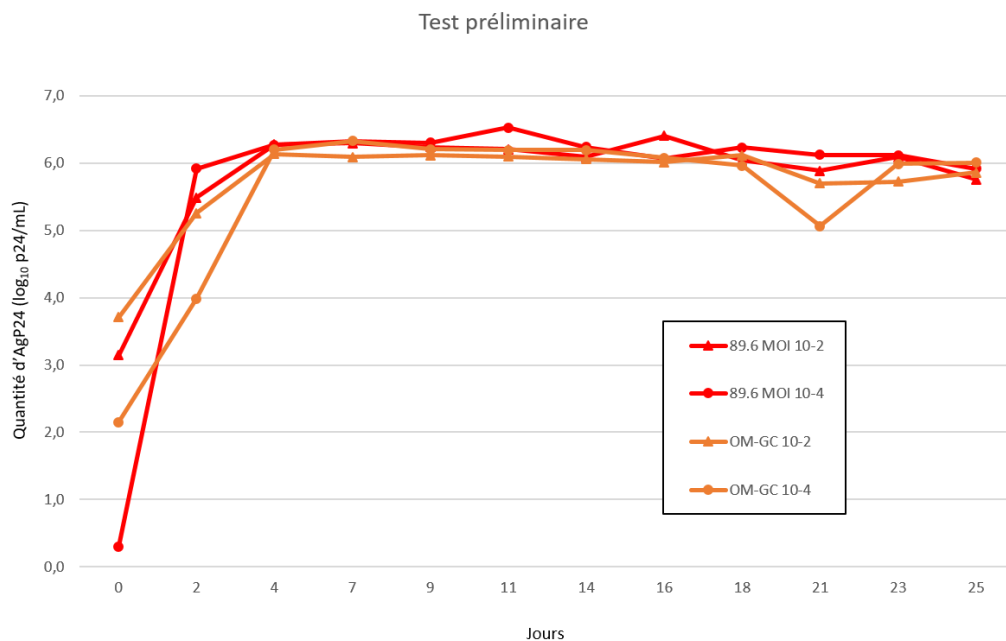


Figure 42 : Cinétiques de répliation obtenues pour les souches OM-TB-2016 et 89.6 à MOI 10<sup>-2</sup> et 10<sup>-4</sup>

Graphique représentant les résultats d'Innotest après mise en culture de la souche OM-TB-2016, représentée en orange et la souche 89.6 en rouge. Les cinétiques lancées à une MOI de 10<sup>-2</sup> sont représentées par un triangle tandis que les cinétiques lancées à une MOI de 10<sup>-4</sup> sont représentées par un rond.

### Cinétique de mortalité cellulaire

Pour chaque cinétique, la mortalité cellulaire a été déterminée avant activation avec la PHA (J-3), le jour de l'infection (J0) et le dernier jour de la culture (Tableau 7) afin de voir l'impact de ce paramètre sur le déroulé de nos cinétiques.

Au J-3, les résultats étaient semblables entre cinétiques, avec une moyenne de 12,7% de mortalité cellulaire entre cinétiques et un faible écart de 3,6% entre la M1 et la M2 (Tableau 7). Juste avant l'infection des cellules au J0, on comptait 7,7% de mortalité pour M1, contre 12,4% ; 13,7% et 15,7% pour M3, M4 et M2 respectivement, soit un écart de 8% entre la M1 et la M2. Ces résultats montraient un niveau basal de mortalité cellulaire avant infection des cellules, de 12% en moyenne et

donc, traduisant une bonne reproductibilité du protocole de récupération et activation des CMSPs entre les différents répliquats.

Tableau 7 : Résultats du comptage cellulaire en fin de culture pour chaque cinétique

N° Mono-infection	Souche testée	J-3		J0	Fin de culture M1 (J10) / M2-M4 (J11)			
		Mortalité cellulaire	Mortalité cellulaire	Rendement	Nombre de cellules vivantes / mL	Nombre total de cellules vivantes	Mortalité cellulaire	Nombre total de cellules vivantes + mortes
<b>M1</b>	OM-TB-2016				1,2E+06	4,7E+06	69,4%	1,5E+07
	RBF208	14,6%	7,7%	54,0%	2,7E+06	1,1E+07	41,2%	1,9E+07
	cellules seules				5,5E+06	2,3E+07	35,2%	3,6E+07
<b>M2</b>	OM-TB-2016				3,6E+05	1,5E+06	87,1%	1,1E+07
	RBF208	11,0%	15,7%	78,0%	8,2E+05	3,4E+06	76,7%	1,5E+07
	cellules seules				4,0E+05	1,6E+06	80,4%	8,2E+06
<b>M3</b>	OM-TB-2016				4,4E+05	1,7E+06	80,5%	8,6E+06
	RBF208	11,2%	12,4%	40,8%	3,1E+06	1,3E+07	23,3%	1,7E+07
	cellules seules				6,5E+06	2,6E+07	10,9%	2,9E+07
<b>M4</b>	OM-TB-2016				1,0E+05	4,1E+05	91,9%	5,1E+06
	RBF208	13,8%	13,7%	163,3%	2,6E+05	1,1E+06	75,5%	4,5E+06
	cellules seules				3,6E+05	1,5E+06	75,7%	6,2E+06
<b>Moyennes</b>	OM-TB-2016				5,2E+05	2,1E+06	82,3%	1,0E+07
	RBF208	12,7%	12,4%	84,0%	1,7E+06	7,1E+06	54,2%	1,4E+07
	cellules seules				3,2E+06	1,3E+07	50,5%	2,0E+07
<b>Différence Δmax-min</b>	OM-TB-2016				1,1E+06	4,3E+06	11,4%	1,0E+07
	RBF208	3,6%	8,0%	122,5%	2,8E+06	1,2E+07	53,4%	1,4E+07
	cellules seules				6,2E+06	2,5E+07	69,5%	3,0E+07

Tableau représentant les résultats de comptage cellulaire après coloration au bleu trypan des cellules au J-3, J0 et J10/11 de culture. Chaque moyenne a été calculée en prenant les valeurs de chaque répliquats de cinétique, tandis que les écarts ont été calculés en réalisant la soustraction de la valeur de la cinétique la plus haute par celle la plus basse.

A la fin de la culture, la mortalité cellulaire était globalement plus élevée pour la souche OM-TB-2016 avec une moyenne de 82,3% entre les répliquats, contre 54,2% et 50,5% pour RBF208 et le témoin cellulaire non infecté, respectivement (Tableau 7). En analysant les cellules témoins non infectées, il y avait également une plus forte mortalité à la fin de la M2 et la M4, avec des valeurs dépassant 75% de mortalité. Les cinétiques M3 et M1 pour ce même témoin, montraient une mortalité cellulaire moyenne moins importante de 10,9% et 35,2% en moyenne, soit un écart maximal de 69,5% entre la M2 et la M3. Enfin, nous avons pu constater un écart de cellules viables résiduelles après infection entre les souches, avec une moyenne de  $2,1 \cdot 10^6$  cellules pour la souche OM-TB-2016 ; de  $7,1 \cdot 10^6$  cellules pour RBF208 et enfin de  $1,3 \cdot 10^7$  cellules pour le témoin cellulaire non infecté (Tableau 7). Ces résultats montraient plus de cellules viables résiduelles pour le témoin cellulaire, en comparaison avec nos deux souches et également, plus de cellules viables résiduelles pour la souche RBF208, en comparaison avec OM-TB-2016.

### *Cinétique de répllication de la souche OM-TB-2016*

Concernant les résultats d'Innotest, les différentes cinétiques ont montré deux tendances : les courbes obtenues pour M1 et M3 d'une part, avec un plateau vers 6 Log<sub>10</sub> d'AgP24/mL et M2 et M4, d'autre part, avec un plateau vers 5 Log<sub>10</sub> d'AgP24/mL (Figure 43). Le pic de production d'AgP24 a été obtenu entre 112 et 168 heures post-infection, avec une intensité comprise entre 5,19 (M1) et 6,27 (M2) Log<sub>10</sub> pg/mL d'AgP24, soit un écart maximal entre cinétique de 1,08 Log<sub>10</sub> pg/mL d'AgP24 (Tableau 8). Cet écart indiquait une faible reproductibilité, étant supérieur au seuil de 0,5 Log<sub>10</sub> fixé comme indiquant des valeurs significatives. Sur l'ensemble des points révélés pour les deux marqueurs, une seule valeur aberrante a été exclue de l'analyse pour OM-TB-2016 au J7 (160h) lors de la révélation par Innotest.

Concernant l'activité de la TI, les différentes courbes obtenues pour OM-TB-2016 étaient également superposables entre M1 et M3, avec un plateau atteint vers 5 Log<sub>10</sub> mU TI/mL (Figure 43). Un écart a été observé entre ces cinétiques M1/M3 et la cinétique M2. En effet, le pic d'activité de la TI a été obtenu entre 120 et 160 heures post-infection, avec une intensité comprise entre 3,84 (M2) et 4,96 (M1) Log<sub>10</sub> mU TI/mL, soit un écart maximal entre cinétique de 1,12 Log<sub>10</sub> mU TI/mL entre M1 et M2 (Tableau 8). Cet écart indiquait également une faible reproductibilité, selon le seuil d'écart de 0,5 Log<sub>10</sub> fixé. Enfin, la cinétique M4 a révélé un comportement intermédiaire entre M1/M3 et M2 avec un début de répllication semblable à M2 mais un plateau se rapprochant de celui obtenu pour M3 (Figure 43).

La vitesse de répllication était en moyenne d'environ 0,04 Log<sub>10</sub> pg/mL d'AgP24 et 0,04 Log<sub>10</sub> mU TI/mL entre répllicats, pour les deux paramètres, montrant une répllication virale similaire pour ces deux paramètres (Tableau 10). Peu d'écart a été observé entre la vitesse de répllication de chaque cinétique pour les deux paramètres (0,0183 Log<sub>10</sub> pg/mL d'AgP24 et 0,0195 Log<sub>10</sub> mU TI/mL).

La reproductibilité de ces cinétiques a été validée malgré l'obtention de deux profils différents et des écarts légèrement supérieurs à 0,5 Log<sub>10</sub>. En effet, un lien avec la mortalité cellulaire au moment de l'infection a été observé. Plus la mortalité des CMSPs était élevée et plus l'intensité du pic pour ces deux paramètres diminuait (Tableau 7).

Tableau 8 : Résultats des Innotests et SG-PERT pour chaque cinétique de la souche OM-TB-2016

jour	heures	OM-TB-2016 Log <sub>10</sub> Quantité AgP24 (pg/mL)					OM-TB-2016 Log <sub>10</sub> Activité TI (mU TI/mL)				
		M1	M2	M3	M4	Moyenne M1 à M4	M1	M2	M3	M4	Moyenne M1 à M4
0	0	1,26	0,87	0,68	0	0,7	0,04	0,00	0,27	0,00	0,01
1.1	16	2,02	1,47	1,5	0	1,25	0,22	0,02	0,11	0,00	0,03
1.2	24	1,86	1,5	1,63	0	1,25	0,12	0,00	0,00	0,00	0,00
2.1	40	3,08	1,53	2,45	0,76	1,96	1,48	0,14	1,36	0,13	1,15
2.2	48	3,19	1,69	3,01	1,38	2,32	2,05	0,22	1,65	0,46	1,62
3.1	64	4,31	2,17	4,15	1,84	3,12	3,14	1,35	2,79	1,48	2,73
3.2	72	4,65	2,39	4,26	2,63	3,48	3,38	1,23	3,07	1,66	2,96
4.1	88	5,27	3,15	5,15	3,49	4,26	3,86	2,31	4,05	2,63	3,69
4.2	96	5,53	3,41	5,21	3,84	4,5	4,15	2,33	3,93	2,98	3,79
5.1	112	<b>6,27</b>	4,23	5,61	4,68	5,2	4,47	3,12	4,40	3,81	4,21
5.2	120	6,01	4,46	5,95	4,93	5,34	<b>4,96</b>	3,35	4,32	4,13	<b>4,51</b>
6.1	136	-*	4,75	<b>5,97</b>	5,2	5,31	-	3,47	4,38	<b>4,23</b>	4,17
6.2	144	-	4,78	5,89	<b>5,43</b>	5,36	-	3,63	4,38	4,02	4,12
7.1	160	6,23	4,99	6,05	<del>5,33</del>	5,76	4,71	<b>3,84</b>	<b>4,75</b>	4,18	4,53
7.2	168	6,47	<b>5,19</b>	6,2	5,33	<b>5,80</b>	4,58	3,81	4,37	4,23	4,33
8.1	184	6,37	5,01	6,31	5,51	5,80	4,47	3,61	4,28	4,20	4,25
8.2	192	6,38	5,04	6,13	5,53	5,77	4,49	3,58	4,17	4,05	4,18
9.1	208	6,23	5,09	5,8	5,16	5,57	4,07	3,73	3,93	3,97	3,97
9.2	216	6,37	4,95	5,86	5,18	5,59	4,27	3,33	4,06	3,64	3,99
10.1	232	6,22	4,87	6,12	5,19	5,6	4,27	3,46	4,25	3,69	4,06
10.2	240	6,18	4,68	6,27	5,1	5,56	4,45	3,43	3,93	3,70	4,07
11.1	256	-	4,67	6,05	5,37	5,36	-	3,33	4,14	3,66	3,88
11.2	264	-	4,82	5,55	5,24	5,2	-	3,18	3,77	3,73	3,65

Le pic est représenté en gras (valeur la plus haute obtenue avant le plateau de réplication). Les différents prélèvements effectués au cours de différentes cinétiques ont été indiqués dans la première colonne par le jour (1 à 11) et l'heure correspondante. Deux prélèvements ont été effectués par jour, représentés par le numéro 1 ou 2.

Moyenne M1 à M4 = « Moyenne réplicats », calculée entre chaque prélèvement des différentes cinétiques.

\* un tiret correspond à un prélèvement non réalisé

~~5,33~~ : valeur aberrante exclue de l'analyse

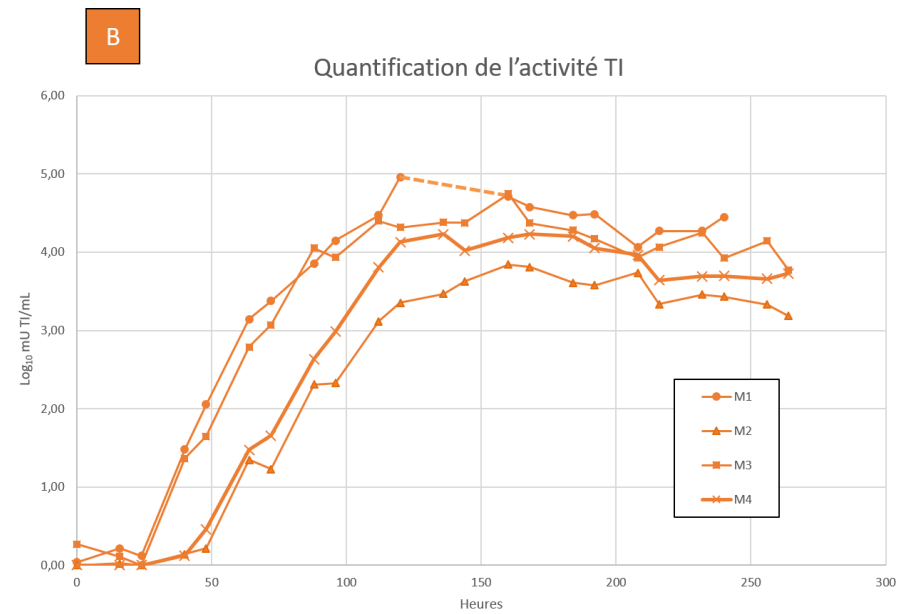
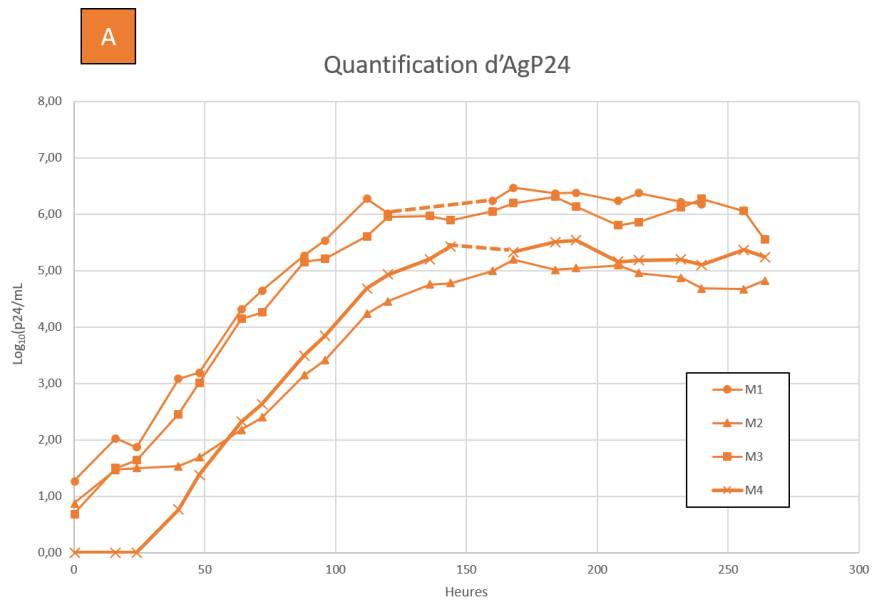


Figure 43 : Cinétiques de répllication de la souche OM-TB-2016

A) Cinétique de la quantité d'AgP24 de quatre réplicats pour la souche OM-TB-2016

L'axe des abscisses représente le temps écoulé en heures post infection, tandis que l'axe des ordonnées représente la quantité d'AgP24 obtenue en Log<sub>10</sub> pg/mL

B) Cinétique de l'activité de la TI de quatre réplicats pour la souche OM-TB-2016

L'axe des abscisses représente le temps écoulé en heures post infection, tandis que l'axe des ordonnées représente la quantité d'activité TI obtenue en Log<sub>10</sub> mU TI/mL

Les pointillés indiquent que la courbe a été extrapolée à cause de données manquantes entre les deux points reliés.

### *Cinétique de réplication de la souche RBF208*

Concernant les résultats d'Innotest pour RBF208, les différentes cinétiques ont montré deux tendances de courbes également avec celles obtenues pour M1 et M3 d'un côté, avec un plateau vers 6-7 Log<sub>10</sub> d'Agp24/mL et M2 et M4 d'autre part, avec un plateau vers 4,5-5 Log<sub>10</sub> pg/mL (Figure 44, Tableau 9). Le pic de production d'AgP24 a été obtenu entre 168 et 208 heures post-infection, avec une intensité comprise entre 4,76 (M2) et 7,36 (M1) Log<sub>10</sub> pg/mL d'AgP24 (Tableau 9), soit un écart entre ces cinétiques de 2,15 Log<sub>10</sub> pg/mL d'AgP24 (Tableau 9). La vitesse de réplication allait, quant à elle, de 0,0268 (M4) à 0,0412 (M1) Log<sub>10</sub> pg/mL d'AgP24/h, avec une moyenne d'environ 0,03 Log<sub>10</sub> pg/mL d'AgP24/h entre réplicats (Tableau 9).

Concernant l'activité de la TI, les différentes courbes obtenues pour RBF208 étaient également superposables entre M1 et M3, avec un plateau atteint vers 5 Log<sub>10</sub> mU TI/mL et M2 et M4 d'autre part, avec un plateau vers 3,5-4 Log<sub>10</sub> mU TI/mL (Figure 44). Le pic d'activité de la TI a été obtenu entre 168 et 208 heures post-infection, avec une intensité comprise entre 3,45 (M2) et 5,08 (M3) Log<sub>10</sub> mU TI/mL soit un écart maximal entre cinétique de 1,63 Log<sub>10</sub> mU TI/mL entre M1 et M2 (Tableau 9). La vitesse de réplication était en moyenne d'environ 0,04 Log<sub>10</sub> mU TI/mL entre réplicats, pour les deux paramètres (Tableau 10).

La reproductibilité de ces cinétiques a été validée comme pour la souche OM-TB-2016, malgré l'obtention de deux profils également, toujours en lien avec la mortalité cellulaire au moment de l'infection (Tableau 7).



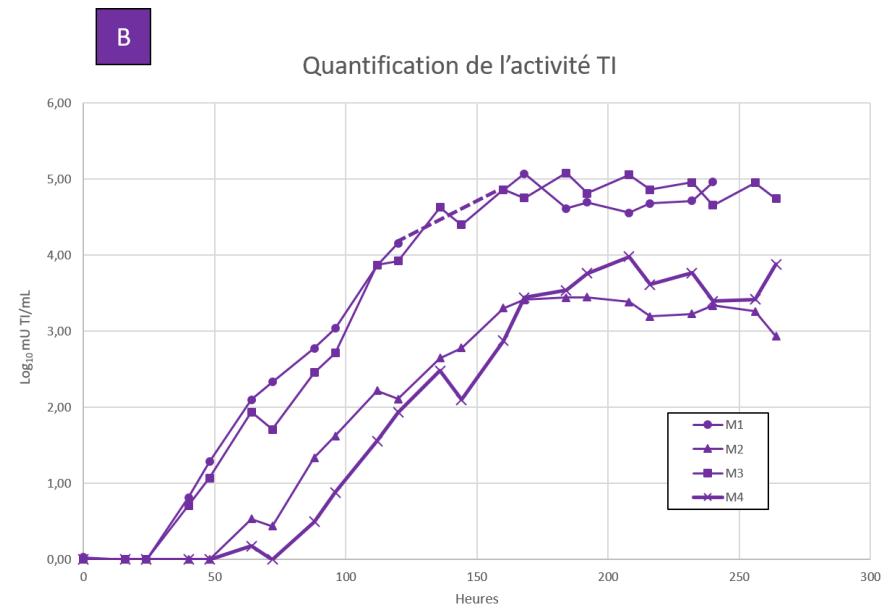
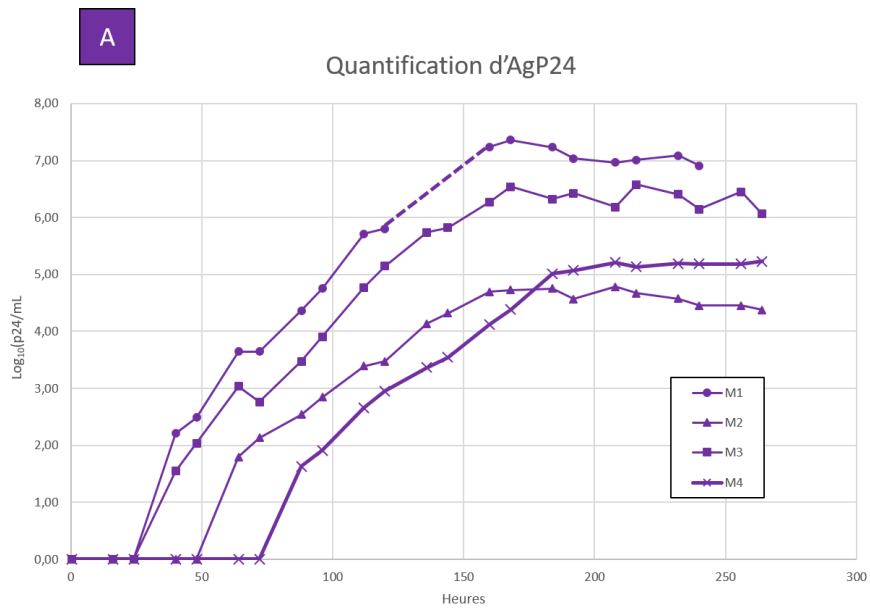
Tableau 9 : Résultats des Innotests et SG-PERT pour chaque cinétique de la souche RBF208

jour	heures	RBF208 Log <sub>10</sub> Quantité AgP24 (pg/mL)					RBF208 Log <sub>10</sub> Activité TI (mU TI/mL)				
		M1	M2	M3	M4	Moyenne M1 à M4	M1	M2	M3	M4	Moyenne M1 à M4
0	0	0	0	0	0	0	0,03	0,00	0,00	0,00	0,00
1.1	16	0	0	0	0	0	0,00	0,00	0,00	0,00	0,00
1.2	24	0	0	0	0	0	0,00	0,00	0,00	0,00	0,00
2.1	40	2,21	0	1,55	0	0,94	0,81	0,00	0,71	0,00	0,49
2.2	48	2,49	0	2,03	0	1,13	1,28	0,00	1,07	0,00	0,90
3.1	64	3,65	1,8	3,04	0	2,12	2,10	0,53	1,94	0,18	1,73
3.2	72	3,65	2,13	2,76	0	2,14	2,33	0,44	1,71	0,00	1,83
4.1	88	4,37	2,54	3,48	1,63	3	2,78	1,34	2,45	0,50	2,36
4.2	96	4,75	2,84	3,91	1,91	3,35	3,04	1,62	2,71	0,88	2,62
5.1	112	5,71	3,39	4,78	2,66	4,14	3,86	2,22	3,87	1,56	3,57
5.2	120	5,8	3,48	5,15	2,95	4,35	4,16	2,11	3,92	1,94	3,76
6.1	136	6,47	4,14	5,74	3,37	4,41	-	2,65	4,63	2,48	4,16
6.2	144	6,7	4,32	5,82	3,55	4,56	-	2,78	4,40	2,10	3,94
7.1	160	7,24	4,70	6,28	4,12	5,58	4,86	3,31	4,86	2,87	4,57
7.2	168	<b>7,36</b>	4,73	<b>6,54</b>	4,39	5,75	<b>5,07</b>	3,41	4,75	3,44	<b>4,65</b>
8.1	184	7,23	<b>4,76</b>	6,32	5,02	<b>5,83</b>	4,61	3,44	<b>5,08</b>	3,54	4,62
8.2	192	7,03	4,57	6,43	5,07	5,77	4,70	<b>3,45</b>	4,81	3,76	4,49
9.1	208	6,97	4,79	6,18	<b>5,21</b>	5,79	4,56	3,38	5,06	<b>3,98</b>	4,61
9.2	216	7,01	4,67	6,58	5,13	5,85	4,68	3,19	4,86	3,61	4,50
10.1	232	7,09	4,58	6,41	5,19	5,82	4,71	3,23	4,96	3,76	4,57
10.2	240	6,91	4,46	6,14	5,19	5,67	4,96	3,34	4,66	3,39	4,55
11.1	256	- *	4,45	6,45	5,19	5,37	-	3,26	4,95	3,42	4,50
11.2	264	-	4,38	6,06	5,23	5,22	-	2,93	4,74	3,88	4,33

Le pic est représenté en gras (valeur la plus haute obtenue avant le plateau de réplication). Les différents prélèvements effectués au cours de différentes cinétiques ont été indiqués dans la première colonne par le jour (1 à 11) et l'heure correspondante. Deux prélèvements ont été effectués par jour, représentés par le numéro 1 ou 2.

Moyenne M1 à M4 = « Moyenne réplicats », calculée entre chaque prélèvement des différentes cinétiques.

\* un tiret correspond à un prélèvement non réalisé



**Figure 44 : Cinétiques de réplication de la souche RBF208**

- A) Représentation des courbes de cinétiques de quatre réplicats pour la souche RBF208 après révélation en Innostest de la quantité d'AgP24  
L'axe des abscisses représente le temps écoulé en heures post infection, tandis que l'axe des ordonnées représente la quantité d'AgP24 obtenue en Log<sub>10</sub> pg/mL
- B) Représentation des courbes de cinétiques de quatre réplicats pour la souche RBF208 après révélation SG-PERT de l'activité de la TI  
L'axe des abscisses représente le temps écoulé en heures post infection, tandis que l'axe des ordonnées représente la quantité d'activité TI obtenue en Log<sub>10</sub> mU TI/mL

Les pointillés indiquent que la courbe a été extrapolée à cause de données manquantes entre les deux points reliés.

### *Comparaison des cinétiques de réplifications des deux formes recombinantes*

L'obtention de résultats similaires, avec deux profils de cinétique distincts pour les deux souches, permettait d'effectuer la moyenne de chacun des deux marqueurs des quatre cinétiques de chacune des deux souches recombinantes étudiées.

En prenant en compte la moyenne obtenue pour chaque souche entre réplikat (moyenne réplikat), le délai d'obtention du pic d'AgP24 était de 168h en moyenne pour la souche OM-TB-2016, à une intensité de 5,80 Log<sub>10</sub> pg/mL et de 184h pour RBF208 avec une intensité de 5,83 Log<sub>10</sub> pg/mL (Tableau 10). Ces résultats ont montré un faible écart de réplification de 16h entre ces souches, pour atteindre une même intensité du pic de production d'AgP24 (différence inférieure à 0,05 Log<sub>10</sub> d'AgP24/mL entre les souches pour ce marqueur) (Tableau 10). Les données d'analyse de la pente au niveau de la phase exponentielle de culture ont montré une vitesse de réplification de 0,0347 Log<sub>10</sub> pg/mL contre 0,0375 Log<sub>10</sub> pg/mL d'AgP24 par heure pour les souches OM-TB-2016 et RBF208 respectivement, traduisant une même vitesse de réplification (Tableau 10). Selon ces résultats, aucune différence en termes de production d'AgP24 n'a été mise en évidence par cette étude.

Le délai d'obtention du pic d'activité de la TI était de 120h en moyenne pour la souche OM-TB-2016, avec une intensité de 4,51 Log<sub>10</sub> d'AgP24/mL et de 168h pour RBF208 avec une intensité de 4,65 Log<sub>10</sub> d'AgP24/mL (Tableaux 10). Ces résultats traduisaient un faible écart de production de 0,14 Log<sub>10</sub> pg/mL et un écart de réplification de 48h entre ces souches. Les vitesses de réplification calculées à partir des données de l'activité de la TI, ont montré quant à elles, des valeurs de 0,0454 Log<sub>10</sub> mU TI/mL/h pour OM-TB-2016 et 0,0327 Log<sub>10</sub> mU TI/mL/h pour RBF208, traduisant un léger écart de réplification de 0,0127 Log<sub>10</sub> mU TI/mL/h en faveur de la souche OM-TB-2016 (Tableau 10). Ces résultats ne mettaient en évidence aucune différence significative, c'est-à-dire supérieure à 0,5 Log<sub>10</sub> mU TI/mL/h, en termes d'activité de la TI.

Les comparaisons des trois paramètres de chacun des deux marqueurs ont montré que les souches OM-TB-2016 et RBF208 avaient le même potentiel réplikatif dans nos conditions d'étude.

Les deux souches RBF208 et OM-TB-2016 possèdent des motifs de recombinaison et une diversité génétique qui sont différents (Figure 37 et 38), cette différence de diversité génétique n'a pas eu d'impact sur les potentiels réplikatifs de ces deux souches, qui apparaissaient similaires quels que soient le marqueur et le paramètre suivis.

Tableau 10 : Comparaison des potentiels de réplication de OM-TB-2016 et RBF208

	AgP24 (Innotests)			ACTIVITE TI (SG-PERT)		
	Intensité du pic	Délais d'apparition du pic	Vitesse d'atteinte du pic	Intensité du pic	Délais d'apparition du pic	Vitesse d'atteinte du pic
	$\text{Log}_{10} (\text{AgP24/mL})$	(heures)	$\text{Log}_{10} (\text{AgP24/mL}) / \text{heures}$	$\text{Log}_{10} (\text{activité mU TI/mL})$	(heures)	$\text{Log}_{10} (\text{activité mU TI/mL}) / \text{heures}$
VIH-1/OM_GC_M1	6,27	112	0,0463	4,96	120	0,0463
VIH-1/OM_GC_M2	5,19	168	0,0280	3,84	160	0,0331
VIH-1/OM_GC_M3	5,97	136	0,0424	4,75	160	0,0269
VIH-1/OM_GC_M4	5,43	144	0,0442	4,23	136	0,0464
$\Delta_{\text{max-min}}$	1,08	56	0,0183	1,12	40	0,0195
<b>moyenne pic</b>	5,72	140	0,0402	4,45	144	0,0382
<b>moyenne réplcats</b>	<b>5,80</b>	<b>168</b>	<b>0,0347</b>	<b>4,51</b>	<b>120</b>	<b>0,0454</b>
VIH-1/RBF208_M1	7,36	168	0,0412	5,07	168	0,0326
VIH-1/RBF208_M2	4,76	160	0,0312	3,45	192	0,0248
VIH-1/RBF208_M3	6,54	168	0,0392	5,08	184	0,0321
VIH-1/RBF208_M4	5,21	208	0,0268	3,98	208	0,0290
$\Delta_{\text{max-min}}$	2,15	48	0,0144	1,63	40	0,0078
<b>moyenne pic</b>	5,97	176	0,0346	4,40	188	0,0296
<b>moyenne réplcats</b>	<b>5,83</b>	<b>184</b>	<b>0,0375</b>	<b>4,65</b>	<b>168</b>	<b>0,0327</b>
$\Delta_{\text{max-min pic OM - RBF208}}$	<b>-0,25</b>	<b>-36</b>	<b>0,0056</b>	<b>0,05</b>	<b>-44</b>	<b>0,0086</b>
$\Delta_{\text{max-min réplcats OM - RBF208}}$	<b>-0,03</b>	<b>-16</b>	<b>-0,0028</b>	<b>-0,14</b>	<b>-48</b>	<b>0,0127</b>

$\Delta_{\text{max-min}}$  : écart calculé entre la valeur maximale et la valeur minimale

*moyenne pic* : moyenne calculée entre les valeurs des différents pics obtenus pour chaque cinétique

*moyenne réplcat* : moyenne calculée entre les valeurs des différents prélèvements correspondants entre chaque cinétique

$\Delta_{\text{max-min pic OM - RBF208}}$  : écart calculé entre la valeur moyenne des pics obtenus pour chaque réplcat des souches OM-TB-2016 et RBF208

$\Delta_{\text{max-min réplcat OM - RBF208}}$  : écart calculé entre réplcats par souches OM-TB-2016 et RBF208 (cf Tableaux 8 et 9)

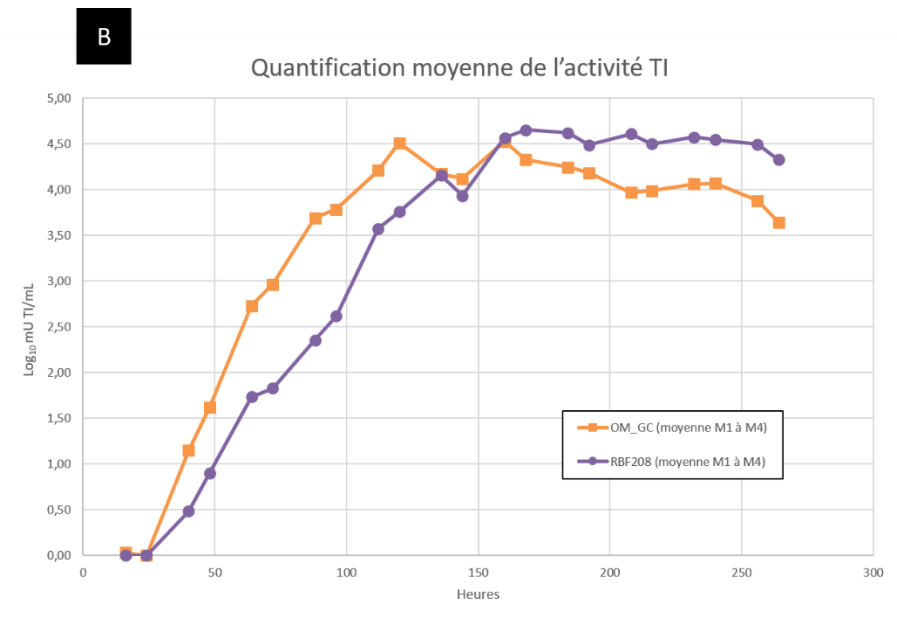
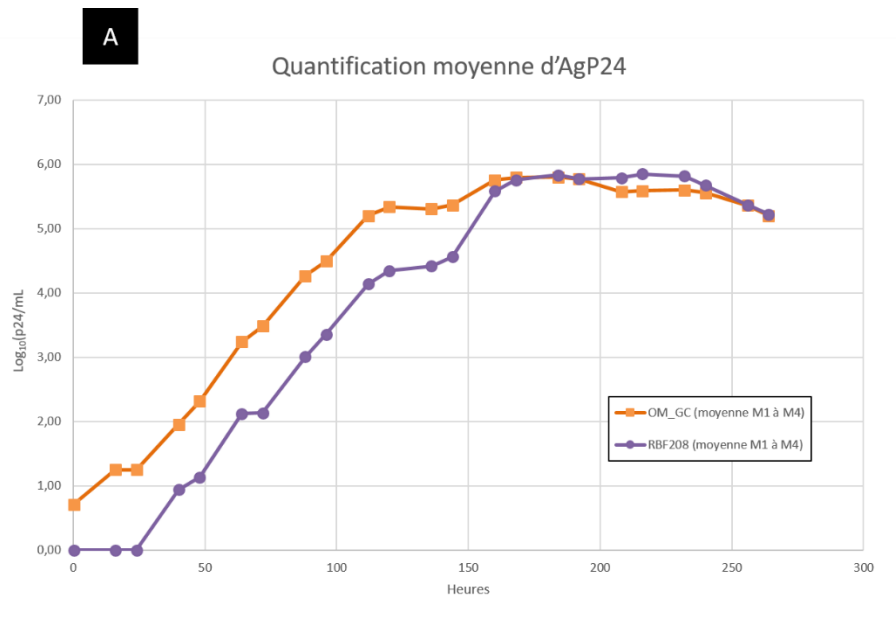


Figure 45 : Comparaison des cinétiques moyenne de répliation des souches OM-TB-2016 vs RBF208

- A) Représentation des courbes de cinétiques moyennes obtenues pour OM-TB-2016 (orange) et RBF208 (violet) après révélation en Innostest de la quantité d'AgP24  
L'axe des abscisses représente le temps écoulé en heures post infection, tandis que l'axe des ordonnées représente la quantité d'AgP24 obtenue en Log<sub>10</sub> pg/mL
- B) Représentation des courbes de cinétiques de trois réplcats pour OM-TB-2016 (orange) et RBF208 (violet) après révélation SG-PERT de l'activité de la TI  
L'axe des abscisses représente le temps écoulé en heures post infection, tandis que l'axe des ordonnées représente la quantité d'activité TI obtenue en Log<sub>10</sub> mU TI/mL

### c. Discussion

Cette étude avait pour objectif de comparer le potentiel répliatif *in vitro* de deux formes recombinantes VIH-1/MO. Les deux souches RBF208 et OM-TB-2016 (110, 137) possédaient des profils de recombinaisons différents, permettant d'étudier un éventuel impact de la diversité génétique et du positionnement du point de cassure sur le potentiel répliatif de ces souches.

Deux profils ont été identifiés entre les répliatés lancés pour les deux souches, avec une concordance entre les résultats des cinétiques M1 et M3 d'une part et ceux des cinétiques M2 et M4 d'autre part. L'utilisation de cellules provenant de donneurs différents, entre les cinétiques M1 à M4, a pu limiter la reproductibilité entre les répliatés. L'obtention de deux profils distincts au cours des cinétiques, pour les deux marqueurs suivis (quantité d'AgP24 et activité TI), a semblé liée au pourcentage de mortalité des CMSPs le jour de l'infection (J0). En effet, des pics d'intensité d'AgP24 et d'activité de la TI plus faibles ont été obtenus lorsque la mortalité à J0 était forte (cas des cinétiques M2 et M4). La moyenne des différents répliatés effectuée pour chaque paramètre (vitesse, délai et intensité d'obtention du pic) de chaque marqueur a permis de comparer nos cinétiques entre elles, en évitant un biais lié à la plus forte mortalité cellulaire observée pour le profil M2/M4. L'analyse du J6 n'a pas été réalisée pour la cinétique M1 des deux souches, car le protocole initialement mis en place prévoyait des prélèvements 6 jours sur 7. Avec l'obtention d'un pic proche de ce point de la culture pour la cinétique M1, nous avons réalisé les prélèvements J6 pour les autres cinétiques, afin de ne pas manquer l'apparition du pic de répliation. Malgré tout, l'absence de J6 pour cette cinétique M1 a probablement conduit à des résultats sous-estimés pour ce point de la culture puisque sur les trois points utilisés pour réaliser la moyenne des quatre cinétiques, deux appartenaient au profil M2/M4 et un seul au profil M1/M3.

Une même intensité de production d'AgP24 a été relevée entre les deux souches, avec 5,80 et 5,83  $\text{Log}_{10}$  pg/mL pour OM-TB-2016 et RBF208, respectivement, une même vitesse de répliation ( $\approx 0,04$   $\text{Log}_{10}$  p24/mL/h) et un pic de répliation atteint à 168h pour les deux souches. Une même intensité d'activité de la TI a été identifiée entre les deux souches, avec 4,51 et 4,65  $\text{Log}_{10}$  mU TI/mL pour OM-TB-2016 et RBF208, respectivement, une différence de vitesse de répliation non significative ( $\approx 0,014$   $\text{Log}_{10}$  mU TI/mL/h) et un pic de répliation atteint à 48h de distance (120h et 168h pour OM-TB-2016 et RBF208 respectivement). Ainsi aucune différence de potentiel répliatif n'a été observée entre ces deux souches au cours de cette étude.

Concernant la souche OM-TB-2016, les données obtenues au cours de cette étude ont été mises en regard des résultats obtenus lors d'une précédente étude (137). Les conditions opératoires retenues dans notre étude n'étaient pas les mêmes, avec l'infection du même nombre de cellules ( $5 \cdot 10^6$ ), mais à une MOI différente ( $10^{-4}$  pour nos cinétiques contre  $10^{-2}$  pour les cinétiques précédentes) et avec des cellules différentes (infection de CMSPs au lieu des cellules Jurkat X4-R5 (137)). Toutefois, les différents prélèvements effectués avaient été révélés par la même technique de quantification d'AgP24 (Innotest), nous permettant de comparer les résultats obtenus au cours de ces deux études. La mesure de l'activité de la TI effectuée au cours de cette précédente étude avec la technique Cavidu a été remplacée au cours de notre étude par la technique de SG-PERT. Cette technique avait été testée au laboratoire sur un panel de souches pour lesquelles des valeurs de Cavidu étaient déjà disponibles, montrant une bonne reproductibilité entre les deux techniques, avec l'absence de différence significative supérieure à  $0,5 \text{ Log}_{10} \text{ mU TI/mL}$  et nous permettant de comparer les résultats entre eux. Le pic d'AgP24 était atteint entre 6,2 et 6,5  $\text{Log}_{10} \text{ pg/mL}$  vers J14 - J19 dans les précédentes conditions, contre 5,2 à 6,3  $\text{Log}_{10} \text{ pg/mL}$  d'AgP24 vers J5 - J7 pour nos expérimentations. Ces résultats montraient une même intensité de répliation maximale, aux alentours de  $6 \text{ Log}_{10} \text{ pg/mL}$  d'AgP24, indifféremment des cellules utilisées (Jurkat X4-R5 versus CMSPs). Concernant le pic d'activité de la TI mesurée en Cavidu, ce dernier était atteint entre 4,4 et 5,2  $\text{Log}_{10} \text{ pg/mL}$  vers J14 - J19 jours, contre 4,2 à 5,0  $\text{Log}_{10} \text{ mU TI/mL}$  mesuré en SG-PERT, vers J5 - J7 pour nos expérimentations. La même intensité d'AgP24 et d'activité de la TI ont donc été obtenues entre ces expérimentations, à partir du même stock de surnageant initial. Ces données ont montré que la modification des paramètres de MOI et de choix de cellules n'a donc pas eu d'impact sur le potentiel répliatif de la souche étudiée en parallèle dans ces deux études. Toutefois, le délai d'obtention du pic pour ces deux marqueurs a été très largement raccourci, avec un gain d'une dizaine de jours de culture en utilisant les CMSPs, au lieu des cellules Jurkat X4-R5. Enfin, les conditions retenues pour le lancement de nos cinétiques impliquaient une MOI plus faible, donc permettaient d'utiliser moins de volume de surnageant.

Les mêmes résultats ont été observés au cours du test préliminaire pour OM-TB-2016, à la MOI de  $10^{-2}$  et  $10^{-4}$ , révélant un pic d'AgP24 d'une même intensité vers  $6 \text{ Log}_{10} \text{ pg/mL}$ , dès J4 donc de façon plus précoce encore que lors du lancement de nos cinétiques. Ces résultats ont confirmé que le potentiel répliatif de la souche OM-TB-2016 n'était pas impacté par le changement de MOI, puisqu'une même intensité du pic de répliation était observé. Néanmoins, pour la même MOI testée de  $10^{-4}$ , le délai de répliation était légèrement impacté par le lot de poches de sang issus des donneurs de CMSPs, avec l'obtention d'un pic à J4 pour le test préliminaire contre un pic entre J5 et J7 pour nos cinétiques, soit une différence de 1 à 3 jours. Ces données montraient une limite de reproductibilité

de notre modèle, mais les données obtenues étaient tout de même comparables d'une cinétique à l'autre.

Contrairement aux travaux précédents, nos résultats ont montré que l'utilisation des cellules Jurkat X4-R5, modifiées et fragilisées, n'était pas recommandée dans notre contexte d'étude (137). La confirmation du problème de tropisme des cellules Jurkat X4-R5, réalisé par cytométrie en flux au cours des expérimentations de clonage (cf **discussion** page 85), a montré que les souches X4 étaient avantagées par rapport aux souches R5, avec une sous-expression de ce co-récepteur au niveau de la membrane des Jurkat X4-R5. Ces cellules avaient été utilisées dans d'autres contextes pour la production de particules virales, sans rencontrer ce problème (191–193). La souche OM-TB-2016 possède un tropisme X4 contrairement à la souche RBF208, qui est R5, confirmant que le potentiel répliatif de ces souches n'aurait pas pu être comparé en utilisant les cellules Jurkat X4-R5 modifiées et justifiant l'utilisation des CMSPs, malgré l'hétérogénéité de ce modèle cellulaire. Les CMSPs n'ont pas été qualifiées par l'EFS ce qui ne permettait pas d'étudier le tropisme des cellules avant infection et constituait également une limite à notre modèle expérimental. En fonction de l'utilisation des CMSPs, issues de donneurs différents, ainsi que de la proportion de cellules ayant un tropisme X4 ou R5, il pourrait y avoir un avantage d'une souche par rapport à une autre en culture. Cette hypothèse pourrait expliquer le léger décalage d'obtention du pic pour nos deux souches, avec une vitesse de réplication plus rapide pour la souche OM-TB-2016 (120h) par rapport à la souche RBF208 (168h), pour le marqueur de la TI notamment. Cependant, l'utilisation de trois poches de sang différentes pour constituer notre mélange de cellules à infecter avait pour objectif de limiter ce risque. Ces résultats pourraient alors traduire une légère différence de potentiel répliatif entre ces souches, du moins pour ce marqueur de la TI.

La souche OM-TB-2016 a induit une mortalité d'environ 80,3% contre 54,2% et 50,5% pour RBF208 et le témoin cellulaire non infecté, respectivement, soit un écart d'environ 27,9% entre les souches OM-TB-2016 et RBF208 et peu d'écart entre le témoin cellulaire et RBF208. La souche RBF208 a donc induit de plus faibles valeurs de mortalité cellulaire, comparativement à la mortalité des CMSPs induite par la souche OM-TB-2016. Ces données, montraient la forte mortalité cellulaire induite par la souche OM-TB-2016 *in vitro*, également visualisées par la présence de nombreux syncytia en culture *in vitro* pour cette souche, qui n'était pas observés pour RB208 et le témoin cellulaire non infecté. Ces résultats ont montré également un écart de 69,5% de mortalité cellulaire entre les réplicats M3 (10,9%) et M2 (80,4%) du témoin non infecté. La moyenne effectuée entre les réplicats pour nos deux souches, couplée à la quantification négative d'AgP24 et la mesure de l'activité de la du témoin cellulaire, permettaient de conclure à cette réelle différence entre nos deux souches OM-TB-2016 et RBF208.



Les deux souches OM-TB-2016 et RBF208 possèdent des fragments parentaux d'origine différente, avec des fragments de VIH-1/M de sous-type B pour OM-TB-2016 et de sous-type D pour RBF208 et des fragments de VIH-1/O de sous-groupe T pour OM-TB-2016 et de sous-groupe H pour RBF208. Cette différence de diversité génétique n'a pas eu d'impact sur le potentiel répliatif, selon les marqueurs d'AgP24 et d'activité de la TI. Les trois paramètres étudiés (vitesse, délais et intensité du pic) pour chaque marqueur, ont montré des différences non significatives. Néanmoins, la différence de tropisme entre ces souches pourrait expliquer la différence de mortalité cellulaire identifiée au cours de cette étude. Les souches X4 sont connues pour être associées à une évolution plus rapide de la maladie, allant jusqu'au stade SIDA, en lien avec une très forte mortalité des lymphocytes CD4+ (194–197). Les souches R5 quant à elles, sont retrouvés durant les phases chroniques ou précoces de l'infection et n'induit pas de syncytia. La souche OM-TB-2016 étant de tropisme X4, contrairement à la souche RBF208 de tropisme R5, a donc pu induire une plus forte mortalité cellulaire en lien avec son tropisme.

Ce travail a permis d'obtenir des données préliminaires sur la comparaison des potentiels répliatifs de deux souches recombinantes différentes, mais, néanmoins, il serait intéressant d'effectuer une comparaison avec d'autres souches de profils plus différents. Le positionnement des points de cassure entre ces deux souches était, certes, différent mais leurs profils de recombinaison étaient similaires. En effet, les deux souches OM-TB-2016 et RBF208 possédaient une TI de groupe O et une enveloppe de groupe M, avec l'hypothèse que ces formes possèdent un fort potentiel répliatif, caractérisé par une production d'Agp24 et une activité enzymatique de la TI importantes et associée à une vitesse de répliation très rapide sur CMSPs *in vitro*. Notre objectif de départ était de comparer deux formes de profil inverse, mais nous nous sommes heurtés à une impossibilité de cloner les souches correspondantes sous forme de CMIC. L'idée de comparer des souches possédant une TI et une enveloppe de groupe différents reste d'intérêt, afin de voir un réel impact du profil de recombinaison sur le potentiel répliatif. L'absence d'autres souches VIH-1/MO isolées en culture au laboratoire ne nous a pas permis d'effectuer cette comparaison. Et la difficulté d'obtenir des outils de clonage adaptés ainsi que le manque de prélèvements de sang total incluant les cellules des patients infectés par ses souches recombinantes ont limité notre étude.

Enfin, une véritable comparaison devra être effectuée avec la mise en compétition des souches, par une étude de *fitness* adaptatif viral, avec le besoin de développer des outils de révélation complexes, non disponibles au moment de notre étude. Une étude récente a montré qu'il était possible de suivre l'expression d'une souche de VIH-1/M de sous-type C, grâce au clonage de cette dernière sous forme de CMI et addition d'un gène rapporteur exprimant la GFP (*Green Fluorescent Protein*). Ce clone présentait un même niveau de répliation que le CMI initial produit au cours de cette

étude (198). Ces résultats permettent d'envisager un marquage spécifique des souches à suivre en culture, avec l'utilisation de différentes protéines fluorescentes, mais cette technologie reste complexe et induirait un niveau de difficulté supplémentaire pour la génération des clones moléculaires.



## II. Transmission verticale d'une forme recombinante VIH-1/MO

Ce travail fait l'objet d'un article présenté en Annexe 8 :

Vautrin M, De Oliveira F, Icard V, Leoz M, Trabaud M-A, Ouziel A, Panetta L, Lemée V, Plantier J-C, Moisan A. 2023. *First evidence of a mother to child transmission of an HIV-1/MO intergroup recombinant form*. En cours de soumission.

La transmission mère-enfant (TME) du VIH, ou transmission verticale, peut avoir lieu à différents moments : lors de la grossesse par voie transplacentaire (15-20%), au moment de l'accouchement et du passage par la filière génitale (45-50%) ou durant l'allaitement (30-40%) (199). Ce risque de TME est le plus élevé à l'accouchement et durant le troisième trimestre de grossesse (199). Ce mode de transmission est la principale cause des nouvelles infections chez les enfants dans le monde (15). Le taux de TME reste variable selon les pays, de l'ordre de 3% en France contre 17% au Cameroun, en 2021 en absence de traitement (15). Néanmoins, une forte réduction du nombre d'enfant nés avec le VIH est observée depuis la mise à disposition plus globale des ARV dans le monde.

Le taux de TME impliquant un VIH-1/M est de l'ordre de 20 à 30% en absence de traitement, contre seulement 0,3% sous traitement en France (21). Au contraire, le VIH-2 est moins transmissible à l'enfant que le VIH-1 (200). De la même façon, une étude de 2022 a montré que le taux de TME du VIH-2 était plus faible que celui du VIH-1 (201) et, en France, de l'ordre de 1 à 3% en absence de traitement et d'environ 0,3% sous traitement ARV (21, 200). Cette différence de taux de transmission est à mettre en lien avec la plus faible répllication du VIH-2 par rapport au VIH-1 (200, 202). En revanche, trop peu de données sont disponibles pour conclure quant à la TME des VIH-1 non M. Quelques cas de TME impliquant le VIH-1/O ont été décrits, dont le premier en 1998 chez une mère d'origine Camerounaise et de son bébé, tous les deux testés positifs au VIH au moment de la naissance (203, 204).

La diversité génétique du VIH aurait un impact sur la transmissibilité du virus. Des études ont montré que le sous-type impliqué et la recombinaison entre sous-types avaient un impact sur la TME (205–208) avec un risque de transmission plus élevé pour le VIH-1/M de sous-type D (207, 209) comparativement aux autres sous-types, ou du sous-type *C in utero* (208). Cependant, une autre étude n'a pas retrouvé de lien entre ces sous-types A et D et la fréquence de TME (210). Aucune étude à ce jour n'a pu démontrer un lien entre TME et groupes, notamment entre les groupes M et O. L'étude de l'influence de la diversité virale et de la recombinaison sur le TME est donc toujours en cours et reste complexe à réaliser, paradoxalement, grâce à la mise à disposition des traitements qui rend le taux de TME parfois très faible dans certains pays (205, 211).

C'est sur ce contexte que se base ce travail. En mai 2022, un bébé (RBF255), né au Cameroun et âgé de 4 mois, a été hospitalisé pour une pneumocystose pulmonaire, en unité de soins intensifs pédiatriques du Centre Hospitalier Universitaire (CHU) de Lyon. Peu d'informations clinico-biologiques étaient disponibles avant son arrivée du Cameroun en France, en janvier 2022. Face à la survenue de cette infection, des sérologies VIH ont été réalisées chez lui puis chez sa mère (RBF254). Pour eux deux, un résultat positif a été obtenu et un diagnostic d'infection à VIH a été posé. Le génotypage de résistance, à Lyon a révélé une discordance moléculaire, suggérant une infection atypique (double infection VIH-1/M+O ou infection par un recombinant VIH-1/MO) et nécessitant une exploration approfondie (Figure 46). Plusieurs prélèvements ont été envoyés au laboratoire associé au Centre National de Référence du VIH (CNR) du CHU de Rouen pour caractérisation séro-moléculaire.

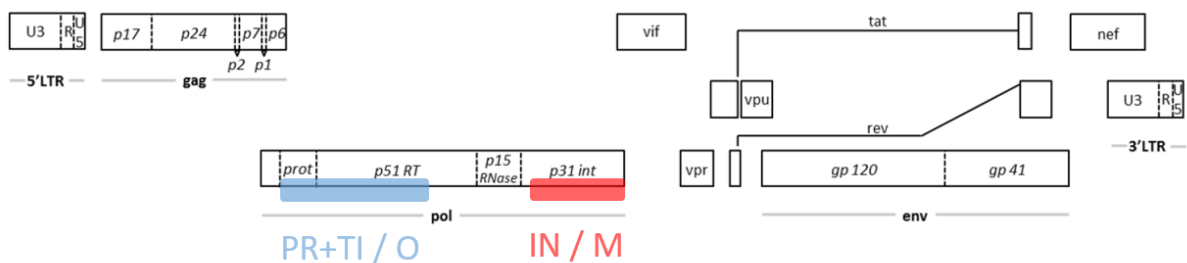


Figure 46 : Discordance moléculaire du génotypage de résistance effectué au CHU de Lyon pour RBF255

Les fragments de VIH-1/O sont indiqués en bleu tandis que les fragments de VIH-1/M sont en rouge. Chaque position a été indiquées par rapport à la souche de référence HxB2 (n° d'accession GenBank K03455), sous les fragments amplifiés correspondants.

## 1. Méthodologie

### *Patients et échantillons*

L'ARN viral a été extrait de façon automatisée à l'aide du kit EZ1® DSP Virus (Qiagen) selon les recommandations du fournisseur, à partir d'un plasma pour RBF255 (16/06/2022) et de deux plasmas pour RBF254 (15/06/2022 et 28/07/2022). Des données viro-cliniques, incluant la CV, le nombre de CD4 et le suivi thérapeutique ont été collectées par le CHU de Lyon, permettant de suivre l'évolution de l'infection tout au long du suivi de ces deux patients jusqu'à ce jour.

## Caractérisation moléculaire des populations virales

### - RT-PCRs nichées spécifiques de groupes M et O

Quatre RT-PCR (*Reverse Transcriptase-PCR*) nichées spécifiques des groupes M et O et ciblant les régions d'intérêt PR, TI, IN et GP41 ont été réalisées sur les extraits d'ARN de RBF254 et RBF255. Afin d'amplifier les régions PR, TI et GP41, des amorces spécifiques du VIH-1/M ont été utilisées selon les recommandations de l'ANRS (212) et des amorces spécifiques du VIH-1/O ont été utilisées, en se basant sur un protocole conçu par le CNR de Rouen (213). Pour amplifier la région de l'IN, d'autres amorces spécifiques de groupes M et O ont également été sélectionnées (214). Une fois les différents fragments amplifiés, les amplicons ont été révélés par migration sur gel d'agarose, puis purifiés à l'aide du kit Nucleospin Gel and PCR clean-up (Macherey-Nagel) avant séquençage par une méthode dérivée de Sanger (Annexe 6). L'analyse des séquences a été réalisée grâce à l'alignement des séquences obtenues avec des souches de référence utilisées dans le cadre de l'activité CNR. Enfin, l'analyse des profils de résistance selon l'algorithme de l'ANRS MIE v32 (octobre 2021) a été réalisée pour chaque fragment séquençé.

### - Séquençage des génomes complets

L'ensemble des étapes est représenté dans la Figure 47.

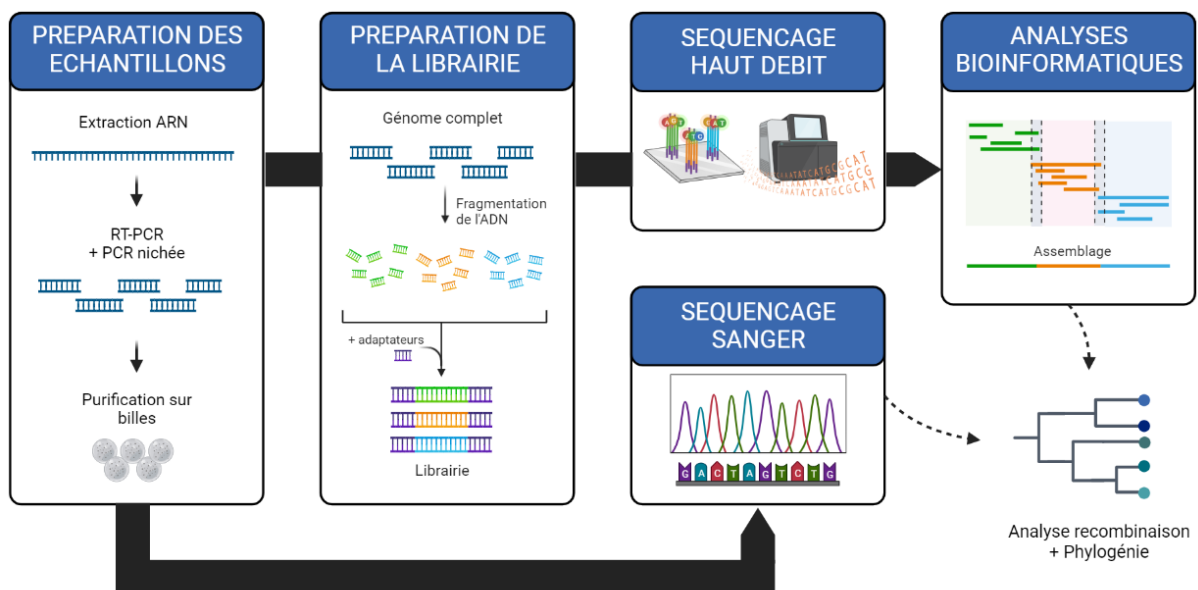


Figure 47 : Etapes de séquençage des génomes complets par séquençage nouvelle génération et Sanger

○ **Amplification des fragments pour le séquençage des génomes complets**

Six fragments chevauchants (A, B, C, D, E1 et E2) ont été amplifiés par un système de RT-PCRs nichées, à l'aide du kit DeepChek® Assay Whole Genome HIV-1 Genotyping V1.x (ABL, Luxembourg), selon les recommandations du fournisseur (215) (Figure 48). Ce protocole développé pour des variants VIH-1/M, a néanmoins dû être adapté, avec l'utilisation d'amorces non spécifiques de groupes permettant également l'amplification du VIH-1/O (Tableau 11). Après amplification, les différents amplicons obtenus ont été quantifiés à l'aide du Qubit™ 4 fluorometer en utilisant le kit dsDNA HS Assay (ThermoFisher Scientific, Etats Unis).

Ces amplicons purifiés ont permis de réaliser en parallèle le séquençage Sanger et le séquençage nouvelle génération (*Next generation sequencing*, NGS) par technologie Illumina®. Lorsque la qualité des séquences Sanger ou le volume d'échantillon étaient insuffisants, une nouvelle PCR nichée était réalisée dans les mêmes conditions que décrites ci-dessus, mais les amplicons étaient purifiés sur colonne à l'aide du kit Nucleospin Gel and PCR Clean up (Macherey-Nagel) après vérification sur gel d'agarose. Une purification sur gel a été effectuée à l'aide de ce même kit, pour les fragments E1 et E2.

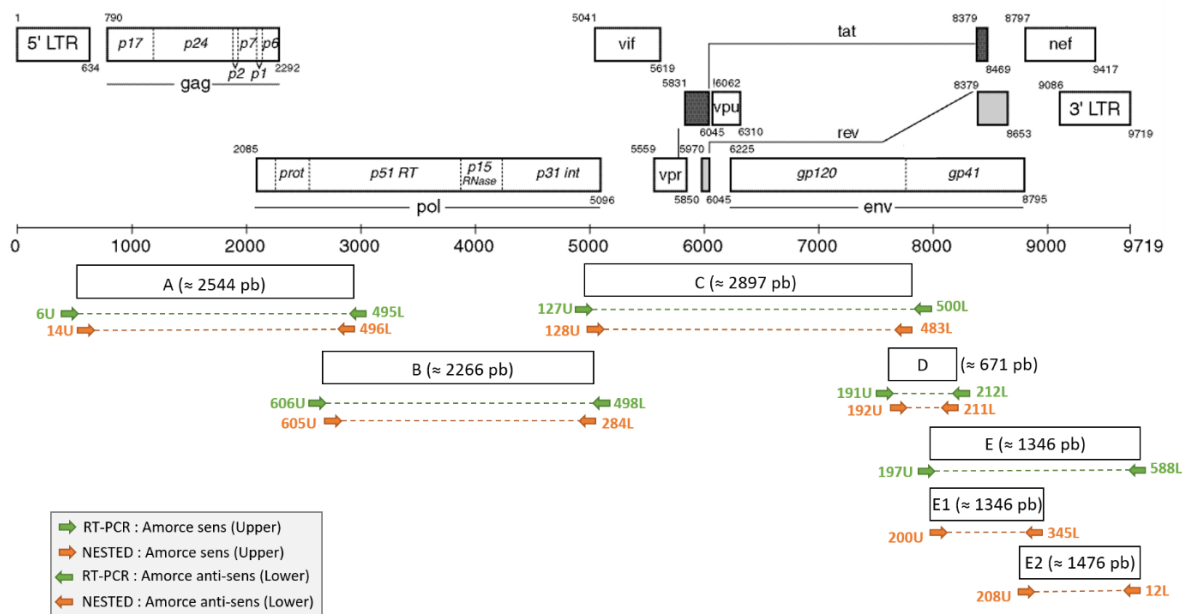


Figure 48 : Position des amorces utilisées pour l'amplification des génomes complets de RBF254 et RBF255

Chaque position est indiquée par rapport à la souche HxB2 et chaque taille de fragment est renseignée dans le rectangle correspondant à l'amplicon. Les amorces ayant servi à lancer la RT-PCR sont représentées en vert, tandis que celles ayant servi pour la Nested PCR sont représentées en orange, sous chaque amplicon attendu.

Tableau 11 : Amorces utilisées pour l'amplification des génomes complets de RBF254 et RBF255

Nom du fragment	Région génomique	Etape PCR	Nom de l'amorce	Séquence de l'amorce (5' - 3')	Taille amplicons (pb)
A	5'LTR/pol	RT-PCR	6U*	ACCAGRITYTGAGCCYGGGAG	2544
			495L**	GCTGGTGAWCCTTTCCATCC	
		NESTED	14U	CTGCTTAASSCTCAATAAAGCTTGC	
			496L	CCYGYGGGAGMACATTGTA CTG	
B	pol	RT-PCR	605U	AGTGGAGAAAATTGGTAGACTTCAG	2266
			498L	GCCACACWATCAKYACCTGCCAT	
		NESTED	606U	AAGAACAAGACTTCTGGGAGG	
			284L	ACCTGCCATCTGTTTCCATA	
C	pol/env	RT-PCR	127U	GGGTYTATTACAGRGACAGCAGAG	2897
			500L	TTGYCTGCTGYTGCATACTACC	
		NESTED	128U	AGAGAYCCWATTGGAAAGGACC	
			483L	CCCATAGTGCTWCCWGCTGC	
D	pol/env	RT-PCR	191U	TGGAGGAGGAGATATGAGG	671
			212L	GTGAGTATCCCTGCCTAACTCTAT	
		NESTED	192U	ATGAGGACAATTGGAGAAGTG	
			211L	CTACCAAGCCTCTACTATC	
E	env/3'LTR	RT-PCR	197U	TCTTAGGAGCAGCAGGAAGCACTAT GGG	1826
			588L	CACTCAAGGCAAGCTTTATTG	
		NESTED (E)	200U	CAGCAGGWAGCACTATGGG	1346
			12L	GCAAGCTTTATTGAGSSTTAAGCAG	
NESTED (E1)	200U	CAGCAGGWAGCACTATGGG	1476		
	345L	GCAAGCTTTATTGAGSSTTAAGCAG			
NESTED (E2)	208U	TGGCAAGTYTGGAAATTGGTT			
	12L	GCAAGCTTTATTGAGSSTTAAGCAG			

\* : Upper (Amorce sens)

\*\* : Lower (Amorce antisens)

Les amorces ayant servi à lancer la RT-PCR sont représentées en vert, tandis que celles ayant servi pour la Nested PCR sont représentés en orange, sous chaque amplicon attendu.

#### ○ Séquençage par la méthode de Sanger

Un séquençage des génomes complet a d'abord été entrepris par la méthode de Sanger, selon les mêmes modalités que décrites en Annexe 6. L'ensemble des amorces ayant servi au séquençage des souches RBF254 et RBF255 a été listé dans le Tableau 12.



Tableau 12 : Amorces utilisées pour le séquençage Sanger des génomes complets de RBF254 et RBF255

Fragment identifiant	Genomic location	Primer name	Primer sequence (5' - 3')
A	5'LTR	14U	CTGCTTAASSCTCAATAAAGCTTGC
	gag	235U	ACATTTAGTATGGGCAAGCA
		236L	TGCTTGAGAGCTGGYTCTA
		40U	GGYGYTGGAATGTGG
	pol	255L	TGGTCTGTCCCAAAGAGG
		51U	CCTCARATCACTCTTTGG
B	pol	606U	AAGAACAAGACTTCTGGGAGG
		266U	GAAARCTAAATTGGGCAAGTC
		268U	GCCTCCACACAAATGAKATAAG
		95U	CATGGGTNCCAGCHCAYAARGG
		274U	AATCCACAAAGTCAAGGAGTAGTAGA
		284L	ACCTGCCATCTGTTTCCATA
	env	117L	CAGTACAATGTACACATGG
	pol	128U	AGAGAYCCWATTTGAAAGGACC
C	acc	294U	ATCTCCYATGGCAGGAAGAAG
		295L	GCTTCTCCTGCCATRGAG
	env	160L	GTACCCATAATAGACYGTGACCC
		447U	CYTTTGARCCAATCCAATACAYTAYTGTGC
		601U	CAGTACAATGYACACATGGAAT
		179L	CCATTTAACAGCAGTTGAGTTG
483L	CCCATAGTGCTWCCWGCTGC		
D1	env	192U	ATGAGGGACAATTGGAGAAGTG
		211L	CTACCAAGCCTCCTACTATC
D2	env	200U	CAGCAGGWAGCACTATGGG
		201U	ACAATTATTGTCTGGTATAGTGCAACAGCA
		208U	TGGGCAAGTYTGTGGAATTGGTT
		335U	AACATTAGGCAGGGATATCAAC
		339L	TAAGTTGCTCAAGAGGTGGTA
	218L	CAGTCCHCCCTTTCTTTTA	
3'LTR	4L	595U	AGATAYCCACTGACMTTGGATG
		4L	TTATATGCAGCWTCTGAGGG
		12L	GCAAGCTTATTGAGSSTTAAGCAG

\* : Upper (Amorce sens)

\*\* : Lower (Amorce antisens)

### ○ Séquençage nouvelle génération par technologie Illumina

Une technique de séquençage de nouvelle génération a également été utilisée au cours de ce projet, en utilisant le kit Illumina DNA prep (Illumina, États-Unis). La librairie a été préparée selon les recommandations du fournisseur, à l'aide des amplicons utilisés générés pour le séquençage Sanger. Pour ce faire, les amplicons de chaque échantillon A, B, C, D, E1 et E2, ont été mélangés. Tous les échantillons ont été marqués et fragmentés au cours d'une même étape de tagmentation, correspondant à l'utilisation de billes couplées à des transposomes et permettant d'obtenir des ADN de taille de 350 pb, compatibles avec la technique de séquençage Illumina utilisée (Figure 49).

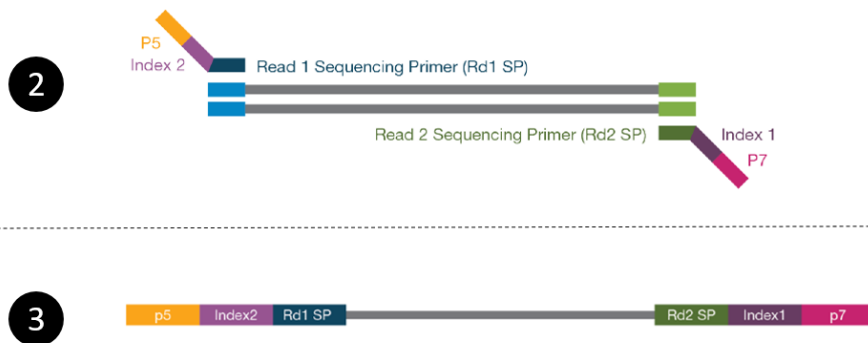
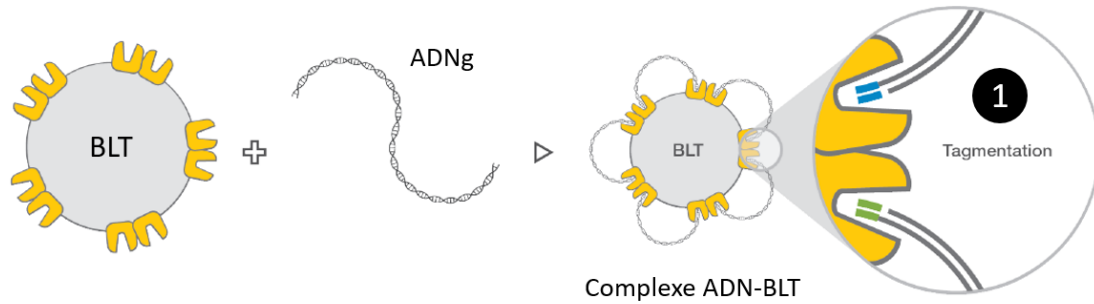


Figure 49 : Principe de la tagmentation

Schéma adapté du protocole Illumina DNA Prep (216).

**1) Tagmentation.** Les billes couplées à des transposomes ont été mise en contact avec l'ADNg de nos échantillons à séquencer afin de permettre à la fois de normaliser la taille des fragments (environ 350 pb) et d'ajouter à leurs extrémités des séquences adaptatrices permettant la fixation des index. **2) Amplification de la librairie par PCR.** Ajout des adaptateurs P5 et P7 et des index nécessaires à la fixation de l'échantillon sur la *flow cell* **3) Séquençage des fragments** après lavage et mélange.

Les fragments sans adaptateurs ont été éliminés à l'aide d'un lavage, puis la librairie a été amplifiée, selon les recommandations du fournisseur, avant d'être à nouveau purifiée par un lavage. L'ensemble des fragments a ensuite été normalisé après quantification sur le Qubit™ 4 fluorometer en utilisant le kit dsDNA HS Assay (Thermo Fisher Scientific, Etats-Unis). Un mélange de tous ces échantillons a été réalisé, puis 13 pM de librairie finale ont été chargés sur la *flow cell*, après de 5% de PhiX, d'un calibrateur et d'un contrôle qualité de séquençage Illumina. Le séquençage a été réalisé à l'aide du système MiSeq (Illumina, États-Unis) en utilisant le kit MiSeq Reagent Nano kit v2.

Le séquençage Illumina repose sur les deux principes suivants : l'amplification par pont (*bridge PCR*) et le séquençage par synthèse (*Sequencing By Synthesis, SBS*). Une fois la librairie d'ADN constituée, celle-ci a été déversée sur le support de séquençage recouvert de séquences complémentaires des adaptateurs fixés à chaque extrémité des fragments. Les fragments se fixent alors au niveau du support et une élongation a lieu à l'aide d'une polymérase afin de générer le brin complémentaire, qui reste bloqué sur la *flow cell*. Chaque fragment restant est amplifié à partir de l'autre extrémité non fixée, qui peut alors s'hybrider sur la séquence complémentaire de la *flow cell*,

formant ainsi un pont. Chaque pont créé permet de synthétiser le fragment complémentaire, qui sera à son tour dénaturé, puis bloqué sur la *flow cell*. L'amplification progresse dans un espace restreint de la *flow cell* et à partir d'un même fragment de départ, entraînant ce que l'on appelle un cluster (Figure 50).

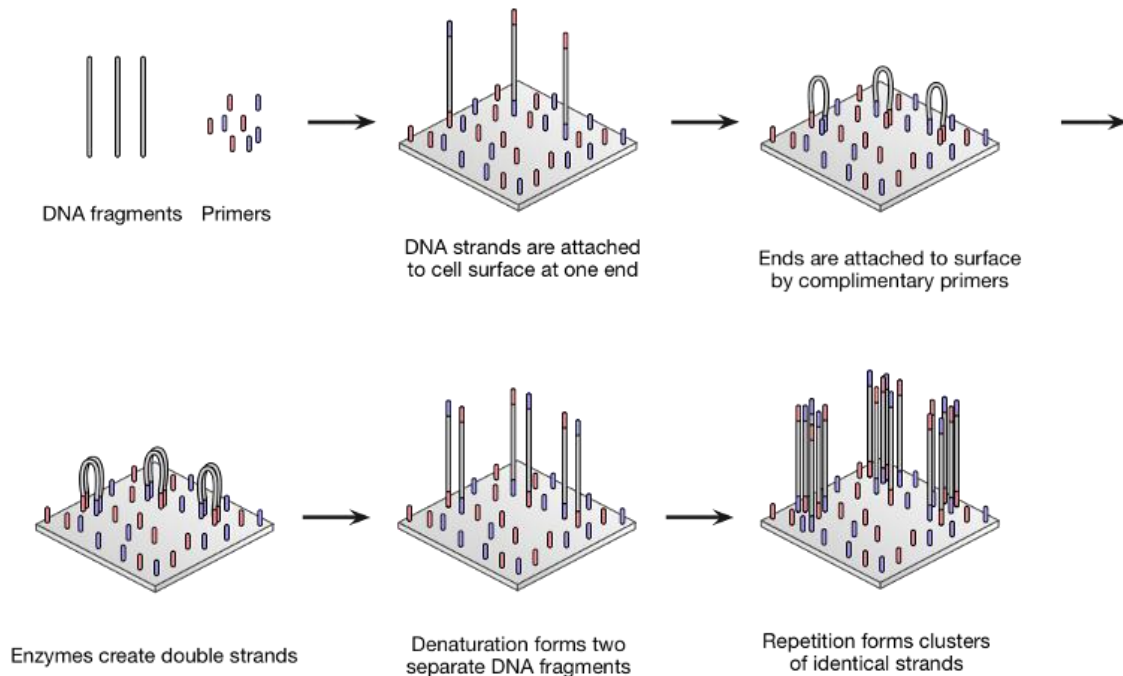


Figure 50 : Principe du séquençage par la technologie Illumina

Issu de ATDBio (217).

Le principe du SBS est basé sur la détection des bases durant la synthèse du brin complémentaire. Chaque base couplée à un fluorophore émet un signal lumineux spécifique, lors de l'incorporation de l'une d'entre elles. Après incorporation, le fluorophore est clivé, permettant à la polymérase de poursuivre l'analyse avec l'incorporation des dNTPs suivants. Chaque signal de fluorescence est converti en base permettant d'obtenir de courtes séquences de nucléotides appelée *reads*, correspondant à la succession des bases lues pour chaque fragment par le séquenceur. Chaque base a été associée à un score de qualité de séquençage, qui sert à déterminer les bases ambiguës et la qualité finale de la séquence obtenue.

Les données statistiques ont été visualisées à l'aide du logiciel Sequencing Analysis Viewer v.2.4.7 (Illumina) afin de contrôler le bon déroulement du séquençage NGS (218). La densité des clusters et le nombre total de *reads* ont été analysés, ainsi que le nombre de *reads* dépassant le seuil du Q30, correspondant à la probabilité que la base attribuée soit bien à bonne à 99,9%. D'autres paramètres ont été contrôlés, tels que le pourcentage de *reads* attribués au PhiX, ceux attribués au témoin négatif et enfin, le pourcentage de *reads* ayant été bien identifiés par les index Illumina.

Les données de séquençage, incluant tous les *reads* séquencés ont été récupérées dans un fichier informatique au format .FASTQ. L'ensemble des analyses bio-informatiques a été réalisé au laboratoire à l'aide du pipeline HAPHPIPE (*H*Aplotype and *P*Hylodynamics *PI*PEline) (219), permettant de supprimer les adaptateurs et les *reads* de mauvaise qualité, de corriger les erreurs de séquençage et de reconstituer le génome complet consensus des différentes souches par deux méthodes d'assemblage :

- Par l'assemblage « *ref-base* », correspondant à l'assemblage des *reads* sur un génome de référence.
- Par l'assemblage « *de novo* », correspondant à l'assemblage des *reads* entre eux jusqu'à obtenir des séquences intermédiaires, jusqu'à finalement reconstituer une séquence consensus sans l'aide d'une référence.

### **Comparaison des deux technologies de séquençage**

Les séquences consensus des deux patients, obtenues par les deux technologies de séquençage, ont été comparées entre elles avec l'outil BLAST, afin de mettre en évidence la concordance des résultats entre les deux technologies de séquençage utilisées. Un alignement de ces séquences obtenues par Sanger et NGS a également été réalisé sur les souches de références HxB2 et ANT70, à l'aide de l'outil MEGA 7, afin d'identifier et de positionner les différences de nucléotides et les décalages de séquences. Une autre analyse a été effectuée afin de visualiser des distances génétiques entre ces séquences à l'aide de MEGA 7. Les distances génétiques, correspondant aux différences identifiées entre chaque paire de séquences, ont été calculées à l'aide de la méthode de Kimura à 2 paramètres. La fiabilité de cette analyse a été estimée par Bootstrap avec 1000 répétitions. Une dernière analyse comparant ces séquences a été effectuée avec l'outil Highlighter de Los Alamos (220), afin de visualiser les différences observées entre nos séquences. Toutes ces analyses ont été effectuées afin de valider notre protocole expérimental de séquençage de nouvelle génération sur les génomes complets de VIH non-M.

Le logiciel HAPHPIPE a pu être utilisé sur les données de NGS, afin d'étudier les quasi-espèces au sein d'un même échantillon, pour nos deux souches, contrairement aux données issues du séquençage Sanger qui ne permettaient pas cette analyse. Cette étape, effectuée avec le module Predict\_haplo du logiciel HAPHPIPE, nous a permis d'étudier les différentes populations présentes dans nos prélèvements séquencés, correspondant aux différentes séquences représentatives de la quasi-espèce.

### *Etudes phylogénétiques et analyse de la recombinaison*

Les séquences obtenues pour les deux patients par méthode Sanger ont été comparées à une banque de données VIH mondiale, à l'aide de l'outil BLAST de Los Alamos (Basic Local Alignment Search Tool) (221). Cet outil a été utilisé afin d'identifier simplement les groupes de VIH-1 (VIH-1/M ou VIH-1/O) et sous-types de chaque fragment impliqué dans la forme recombinante VIH-1/MO.

Une première étape de visualisation des points de cassure a été effectuée sur toutes ces séquences à l'aide de l'outil HIV Genotyping du NIBSC (222). Ces dernières ont ensuite été alignées avec un ensemble de séquences VIH-1/M et VIH-1/O, provenant de la base de données VIH de Los Alamos (24) et de notre laboratoire, à l'aide de MEGA 7 (223), afin de préciser plus finement la localisation des points de cassure.

Le logiciel SimPlot (224) a été utilisé, après l'alignement des séquences préalablement effectué, afin d'identifier et de localiser les points de cassure le long du génome en utilisant la méthode statistique de Kimura à 2 paramètres et un rapport transition-transversion de 2,0. Ce logiciel permet d'analyser l'ensemble de la séquence d'intérêt en calculant la similarité de cette séquence par rapport à d'autres séquences de référence préalablement choisies, afin de déterminer le profil de recombinaison sur l'ensemble du génome des deux virus. Ce logiciel permettait non seulement la visualisation des points de cassure mais également de la diversité génétique des fragments impliqués.

Des arbres phylogénétiques ont été réalisés, à l'aide de la méthode de Neighbor-joining, afin de visualiser les liens phylogénétiques entre échantillons, avec les différents sous-types et sous-groupes de VIH-1, ainsi qu'avec les 12 URF\_MO déjà décrites *in vivo* et complètement séquencées. Les distances évolutives, correspondant aux substitutions par position, ont été calculées à l'aide de la méthode de Kimura à 2 paramètres. La fiabilité de ces arbres a été estimée par une analyse Bootstrap avec 1000 répétitions.

## 2. Résultats

### *Patients et échantillons*

Le suivi des patients au CHU de Lyon a permis d'obtenir les données associées aux différents prélèvements de RBF254 et RBF255 (Tableau 13, Figure 51).

Pour les deux patients, une trithérapie antirétrovirale active a été initiée dès juin 2022 avec l'association bictégavir-emtricitabine-tenofovir pour RBF254 et zidovudine-lamivudine-dolutégravir, pour RBF255, rapidement changé pour la combinaison abacavir-lamivudine-dolutégravir, du fait d'une adaptation pharmacologique. Le suivi séquentiel et rapproché des deux patients a permis de mettre en évidence une bonne réponse au traitement ARV pour les deux patients.

En effet, pour RBF254, le zénith de la CV était de 5,4 Log<sub>10</sub> cp/mL. Une diminution de 3,4 Log<sub>10</sub> cp/mL a été observée après 3 semaines de traitement et l'indéteçtabilité a été finalement obtenue en 3 mois environ (Tableau 13). Le nadir de CD4 avait été observé au moment du diagnostic (308 CD4/mm<sup>3</sup>) et la mise sous traitement a permis une bonne restauration immunitaire.

Concernant RBF255, le zénith de la CV était de 6,3 Log<sub>10</sub> cp/mL au moment du diagnostic. La mise sous traitement a été associée à une diminution rapide de la CV, d'environ 3 Log<sub>10</sub> cp/mL en une semaine seulement (Tableau 13). Cette diminution de la CV était associée à une restauration immunitaire intense (+ 800 CD4/mm<sup>3</sup> environ après un mois de traitement ARV) à partir du nadir de CD4, observé comme pour RBF254 au moment du diagnostic (141 CD4/mm<sup>3</sup>). En revanche, du fait de difficultés d'observance, l'indéteçtabilité n'avait toujours pas été obtenue après 9 mois de traitement ARV.

Tableau 13 : Données viro-cliniques et thérapeutiques des échantillons pour RBF254 et RBF255

	Date de prélèvement	CV plasmatique (Log <sub>10</sub> cp/mL)	Comptage CD4 (/mm <sup>3</sup> )	Données de séquençage									
				Lieux d'analyse	PR	RT	IN	GP41	Génome complet	VIH-1/M	VIH-1/O		
<b>RBF254</b>	14-06-2022	5,4	331	Lyon	VIH-1/O	VIH-1/O	-*						
	16-06-2022	5,4	308	Rouen	VIH-1/O	VIH-1/O	VIH-1/M	VIH-1/M	VIH-1/MO	CRF02_AG	T		
	<i>30-06-2022 initiation bictegavir-emtricitabine-tenofovir</i>												
	22-07-2022	2,0	412										
	22-08-2022	1,7	430										
	20-09-2022	Indétectable											
	19-10-2022	1,8	465										
	15-12-2022	1,7	444										
	23-03-2023	Indétectable	514										
	<i>23-03-2023 initiation dolutegravir-emtricitabine-tenofovir</i>												
<b>RBF255</b>	31-05-2022	6,3	141										
	02-06-2022		141										
	<i>07-06-2022 initiation zidovudine-lamivudine-dolutegravir</i>												
	08-06-2022			Lyon	VIH-1/O	VIH-1/O	VIH-1/M						
	<i>14-06-2022 initiation abacavir-lamivudine-dolutegravir</i>												
	15-06-2022	3,3		Rouen	VIH-1/O	VIH-1/O	-	VIH-1/M					
	23-06-2022	2,7											
	07-07-2022	4,2	927										
	12-07-2022	4,0											
	28-07-2022	4,8		Rouen			VIH-1/M		VIH-1/MO	CRF02_AG	T		
	04-08-2022	2,7											
	23-08-2022	2,5											
	13-09-2022	2,1	685										
	11-10-2022	2,4	850										
	15-11-2022	2,4	1220										
	14-12-2022	2,1	1456										
	19-01-2023	1,9	1260										
	14-02-2023	2,0											
	14-03-2023	2,2	1110										

\*- : Pas d'amplification

PR: Protéase, RT: Reverse transcriptase, IN: Intégrase

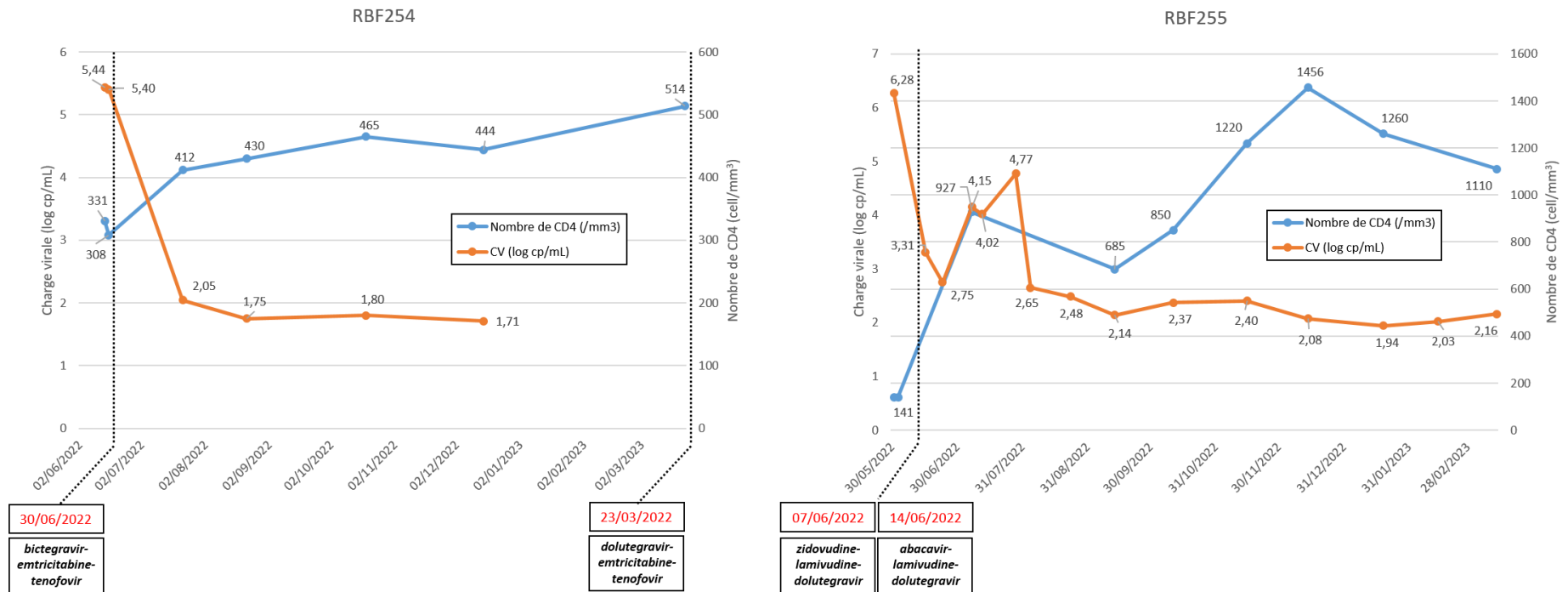


Figure 51 : Suivi de la CV et du nombre de CD4 pour les échantillons de RBF254 et RBF255

A gauche est représenté le suivi de la CV et du nombre de CD4 pour RBF254 et à droite pour RBF255, en fonction du temps. Au-dessus de chaque point est indiquée la valeur de la CV en Log<sub>10</sub> cp/mL ou en nombre de CD4/mm<sup>3</sup>. Le suivi de la CV a été représenté en orange et le nombre de CD4 en bleu. La mise sous traitement a été indiquée par une date en rouge en abscisse, avec le traitement correspondant en dessous.



## Caractérisation moléculaire des populations virales

### - RT-PCRs nichées spécifiques de groupes M et O

Après les RT-PCR nichées spécifiques des groupes M et O, il s'est avéré que la région PR+TI était amplifiée seulement par les amorces de groupe O pour les deux souches, tandis que la GP41 était amplifiée avec les amorces de groupe M uniquement. La région de l'IN de RBF254 a été amplifiée avec les amorces de VIH-1/M, mais des difficultés ont été rencontrées pour amplifier cette région pour RBF255. L'IN de RBF255 a été amplifiée avec des amorces de VIH-1/O, mais les données de séquençage ont finalement confirmé qu'il s'agissait d'une amplification non spécifique, montrant une IN de groupe M après séquençage, comme pour RBF254. Pour les deux patients, les régions PR et TI correspondaient à du VIH-1/O, alors que les régions IN et GP41 correspondaient à du VIH-1/M (Figure 52). Ces données suggéraient la présence d'une forme recombinante VIH-1/MO avec, au moins, un point de cassure de motif OM suspecté au niveau de la région TI-IN, sans forme parentale VIH-1/M ou VIH-1/O associée.

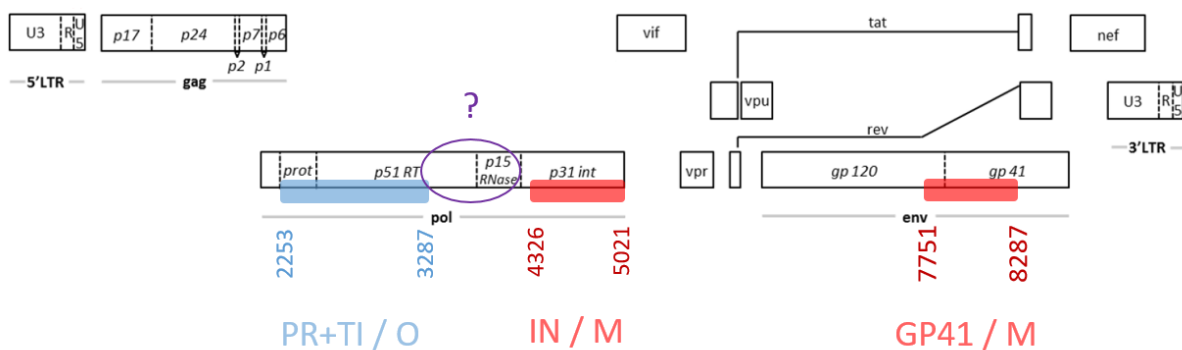


Figure 52 : Discordance moléculaire observée pour RBF254 et RBF255 après investigation par le CNR de Rouen

Chaque position a été indiquée par rapport à la souche HxB2 et les positions de chaque début et fin de fragment amplifiés ont été indiquées en dessous du rectangle coloré correspondant, avec la représentation suivante : en bleu figurent les portions de génome appartenant à du VIH-1/O, tandis que celles appartenant à du VIH-1/M ont été indiquées en rouge. En violet figure la portion non séquencée suggérant un point de cassure de profil OM dans cette région.

L'analyse des séquences a révélé les mêmes profils de résistance chez la mère et son enfant, avec les mêmes mutations détectées : A98G et K103R dans la TI, L10I, Q58E et A71V dans la PR et aucune mutation associée à de la résistance dans les régions IN et GP41. Ces mutations n'entraînaient pas de résistance aux traitements ARV initiés et n'entraînaient qu'une résistance possible à la rilpivirine (INNTI) et l'atazanavir (IP).

## - Séquençage des génomes complets

### ○ Amplification des fragments pour le séquençage des génomes complets

Les six fragments chevauchants A, B, C, D, E1 et E2, conçus pour couvrir l'intégralité des génomes viraux de RBF254 et RBF255, ont été obtenus avec succès après RT-PCR nichées non-spécifiques de groupes et révélation sur gel d'agarose. L'amplification du fragment D, correspondant au début de la GP41, a nécessité l'utilisation d'amorces spécifiques du VIH-1/M. Le fragment E initialement amplifié en un seul fragment, allant de la région GP41 de l'enveloppe à la fin du génome, a montré la présence d'une bande d'intérêt de très faible intensité et de multiples bandes non spécifiques sur le gel d'agarose. Deux fragments (E1 et E2) ont donc été amplifiés à l'aide d'amorces non spécifiques de groupes (à l'exception de l'amorce Upper du fragment E2), à partir des amplicons de la RT-PCR (fragment E complet initialement). Les amplicons E1 et E2 ont été purifiés directement à partir du gel d'agarose, afin d'obtenir une bonne qualité d'échantillon.

### ○ Séquençage par la méthode de Sanger

Le positionnement d'amorces tous les 500 pb le long du génome environ a permis de séquencer les génomes complets des souches de RBF254 et RBF255. Néanmoins, des difficultés ont été rencontrées pour le séquençage du début de l'IN (fragment B) et de l'enveloppe (Fragments D, E1 et E2). Plusieurs essais de séquençage ont été réalisés pour ces fragments, dû à un bruit de fond important. Ce dernier a été relié à la présence des points de cassures, à une quantité d'ADN insuffisante et la présence de bandes non spécifiques sur le gel d'agarose à l'issue de l'amplification. Pour la GP41, une étape de purification sur gel et l'utilisation d'amorces spécifiques de groupe a été nécessaire pour obtenir la séquence finale. Les séquences finales obtenues pour les deux patients ont montré une plus forte diversité intra-échantillon pour RBF254, suggérant la présence d'une quasi-espèce virale plus ancienne que pour RBF255. La présence de plusieurs populations au sein de cet échantillon a été investiguée par séquençage NGS.

### ○ Séquençage nouvelle génération par technologie Illumina

Les données du séquençage de chaque souche ont été analysées après lancement sur l'automate MiSeq, via le logiciel Sequencing Analysis Viewer d'Illumina. Un total de 83 366 *reads* ont été obtenus pour nos échantillons, dont 94,6 % (n=78 897) possédaient un score de qualité supérieur au Q30 (Tableau 14, Figure 53). Sur la totalité des *reads* obtenus, une proportion d'identification

homogène a été observée entre les échantillons allant de 3,3 à 4,7% pour RBF254 et RBF255 respectivement et de 0,0002% pour le témoin négatif (Tableau 14).

Tableau 14 : Résultats de séquençage obtenus pour RBF254 et RBF255 par technologie Illumina

	RBF254	RBF255	Total du run
Nombre de <i>reads</i> totaux	83 366		1 034 505
Nombre de <i>reads</i> PF*	46 456	32 441	78 897
Pourcentage de <i>reads</i> PF identifiés	4,7	3,3	94,6%

PF : *Reads* ayant un score de qualité supérieur à Q30

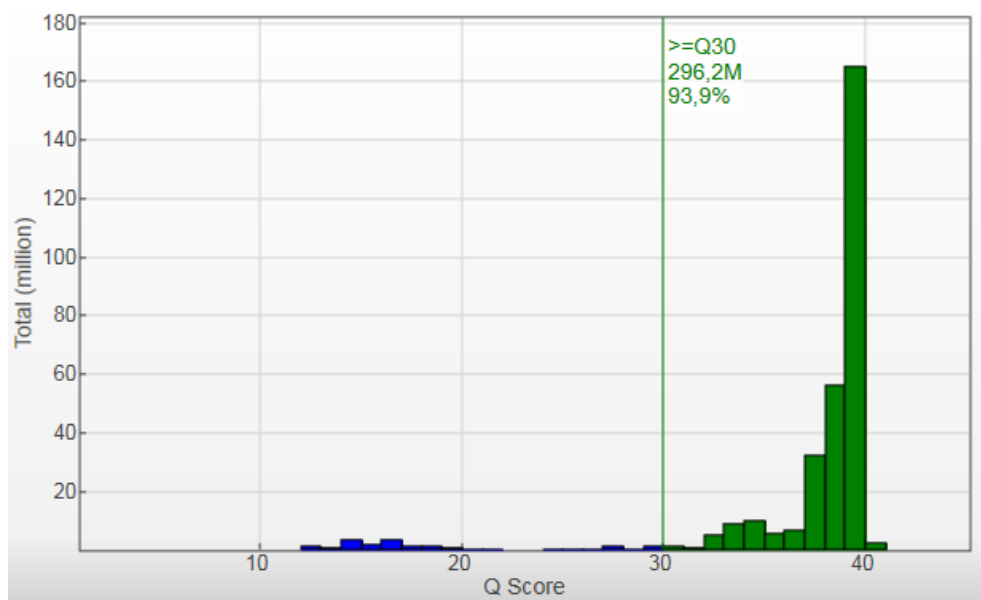


Figure 53 : Histogramme présentant la qualité des *reads* obtenus après séquençage NGS

En bleu figure les *reads* de mauvaise qualité, possédant un score de qualité inférieur au seuil fixé par Illumina Q30, qui est en vert, représentant les *reads* de bonne qualité (Graphique issu du logiciel Sequencing Analysis Viewer).

L'analyse par l'outil HAPHPIPE a permis d'obtenir les génomes complets des deux souches. La séquence consensus du bébé, RBF255, a été obtenue directement à partir de l'alignement des *reads* entre eux (assemblage *de novo*). L'alignement des *reads* entre eux par méthode *de novo*, avec l'outil HAPHPIPE, n'a pas permis d'obtenir d'informations entre les positions 8226 – 8299 pb pour la mère, RBF254 (Figure 54). Une étape supplémentaire d'alignement des *reads* sur la séquence de RBF255, prise comme référence, a été réalisée afin de reconstituer le génome complet de RBF254. Au final, les deux séquences obtenues avaient une taille totale de 9264 et 9255 pb. Enfin, cet outil n'a pas mis en évidence la présence de populations parentales au sein des échantillons.

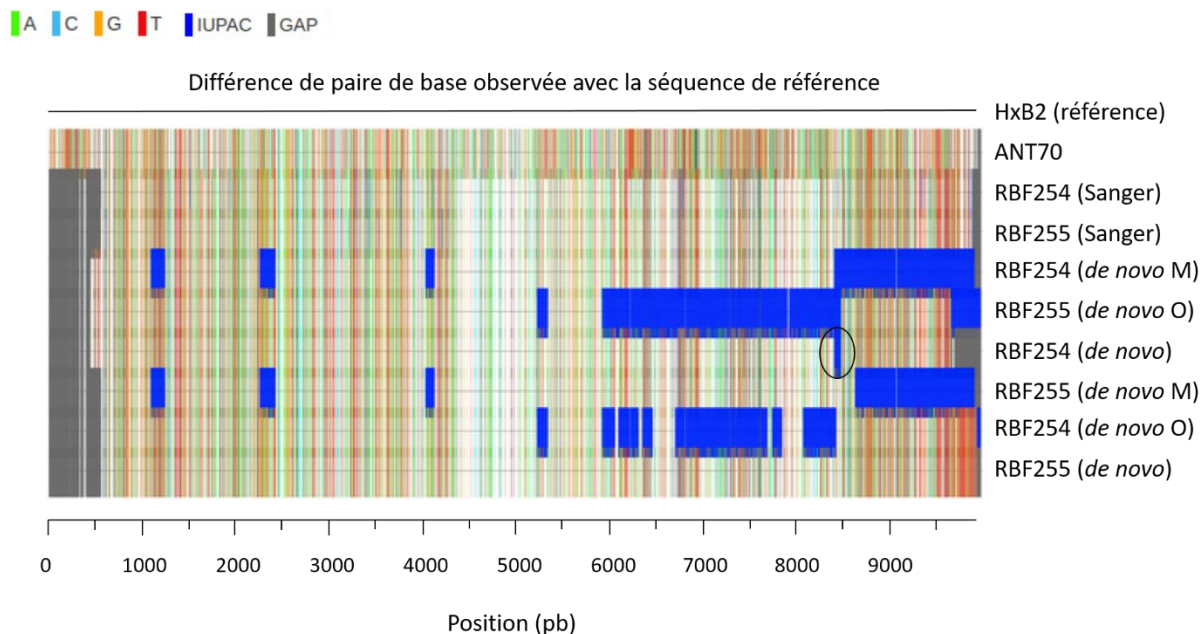


Figure 54 : Différences de paires de base identifiées entre les séquences Sanger et NGS pour RBF254 et RBF255

Cette analyse a permis de visualiser par un code couleur, les nucléotides de chaque séquence qui ne correspondant pas à ceux de la séquence de référence HxB2. Le code IUPAC (International Union of Pure and Applied Chemistry) indique les positions pour lesquelles la paire de bases identifiée n'était pas définie dû à la présence de plusieurs populations, différentes de celles de la référence HxB2, ou dû à la présence d'une base non identifiée. Les GAP, indiquent les décalages de séquences identifiées par rapport à la souche de référence HxB2.

### Comparaison des deux technologies de séquençage

La comparaison des séquences effectuée pour les deux souches par l'outil BLAST, a permis de montrer qu'il n'y avait que 1% de différence entre le Sanger et le NGS (Tableau 15). Cette différence constatée est la conséquence de plusieurs éléments. Tout d'abord, la taille des séquences était similaire quelle que soit la technologie utilisée (environ 9200 pb). Néanmoins, un léger décalage entre le début et la fin de séquence a été observé pour nos deux souches. Enfin, les séquences consensus obtenues présentaient quelques différences de nucléotides, notamment entre les séquences *de novo* et celle obtenue en Sanger pour un même patient, ainsi que quelques différences entre les séquences de RBF254 et RBF255.

**Tableau 15** : Comparaison des séquences consensus de RBF254 et RBF255 obtenues par séquençage NGS et Sanger

	Technologie de séquençage	Taille totale (pb)	Début de la séquence (pb)	Fin de la séquence (pb)	Pourcentage et score d'identité (BLAST)
<b>RBF254</b>	Sanger	9204	527	9619	99% (8854/8981)
	NGS ( <i>de novo</i> )	9024	410	9420	
	NGS ( <i>ref-base</i> )	9264	526	9680	99% (9147/9207)
<b>RBF255</b>	Sanger	9203	522	9619	99% (9102/9207)
	NGS ( <i>de novo</i> )	9255	526	9680	

Chaque position a été indiquée par rapport à la souche de référence HxB2 (n° d'accèsion GenBank K03455).

Le pourcentage d'identité obtenu par l'outil BLAST montre le nombre de paires de bases identiques entre les deux séquences comparées, sur le nombre total de bases analysées entre les deux séquences. Chaque séquence obtenue par séquençage NGS a été comparée à la séquence obtenue par séquençage Sanger.

L'étude des distances génétiques, effectuée à l'aide de MEGA 7, a montré qu'il y avait très peu de différence entre les séquences obtenues selon les deux techniques de séquençage (Tableau 16). Elles indiquaient, en effet, un chiffre très proche de 0, correspondant à une séquence identique, à la fois entre RBF254 et RBF255 et entre les séquences obtenues par séquençage Sanger et NGS. Cette distance était plus importante avec les souches de référence utilisées au cours de l'alignement (HxB2 et ANT70).

**Tableau 16** : Matrice de distances entre les séquences de RBF254 et RBF255 obtenues par séquençage Sanger et NGS

	1. HxB2	2. ANT70	3. RBF254 Sanger	4. RBF255 Sanger	5. RBF254 NGS	6. RBF255 NGS
<b>1. HxB2</b>	-	-	-	-	-	-
<b>2. ANT70</b>	0.589	-	-	-	-	-
<b>3. RBF254 Sanger</b>	0.378	0.328	-	-	-	-
<b>4. RBF255 Sanger</b>	0.382	0.330	0.005	-	-	-
<b>5. RBF254 NGS</b>	0.381	0.329	0.001	0.007	-	-
<b>6. RBF255 NGS</b>	0.382	0.329	0.005	0.001	0.006	-

Cette matrice de distance a été calculée à l'aide d'un test Kimura 2-paramètres. Chaque séquence représentée par un numéro a été comparée aux numéros correspondants, figurant en haut du tableau. Les distances génétiques mesurées entre les séquences ont été renseignées dans les cases correspondantes. Les tirets indiquent la comparaison impossible entre les mêmes séquences ayant un numéro identique, ou une comparaison déjà renseignée dans la matrice.

### **Etudes phylogénétiques et analyse de la recombinaison**

Toutes les séquences obtenues par Sanger ont été comparées avec une banque de données via l'outil BLAST, afin de déterminer selon une première méthode, la diversité des fragments impliqués dans la forme recombinante de RBF254 et RBF255. Ces différentes comparaisons effectuées pour chaque région (PR, TI, IN et GP41) et pour les génomes complets, ont montré que les fragments de de

VIH-1/M appartenait à un sous-type CRF02\_AG. Les deux génomes de VIH-1/M et de VIH-1/O les plus proches obtenus par BLAST, à savoir AB485634 (VIH-1/M) et AY169810 (VIH-1/O), ont été téléchargés pour servir de séquences de référence pour la suite des analyses.

Ces données ont été confirmées par une deuxième méthode, avec la construction de deux arbres phylogénétiques pour RBF254 et RBF255 des régions PR-RT d'une part et IN d'autre part. L'alignement et la comparaison de ces deux régions du génome, avec une banque de séquences de différents sous-types de VIH-1/M et de sous-groupe VIH-1/O, a bien confirmé la diversité génétique retrouvé au sein de ces fragments pour nos deux souches (Figure 55). Cette analyse a également montré que les fragments de VIH-1/O appartenait à un sous-groupe T. Pour ces deux régions étudiées, les arbres ont montré une proximité génétique de 100% entre les séquences PR-TI et IN des deux patients, indiquant qu'ils étaient tous les deux infectés par la même souche recombinante (Figure 55).

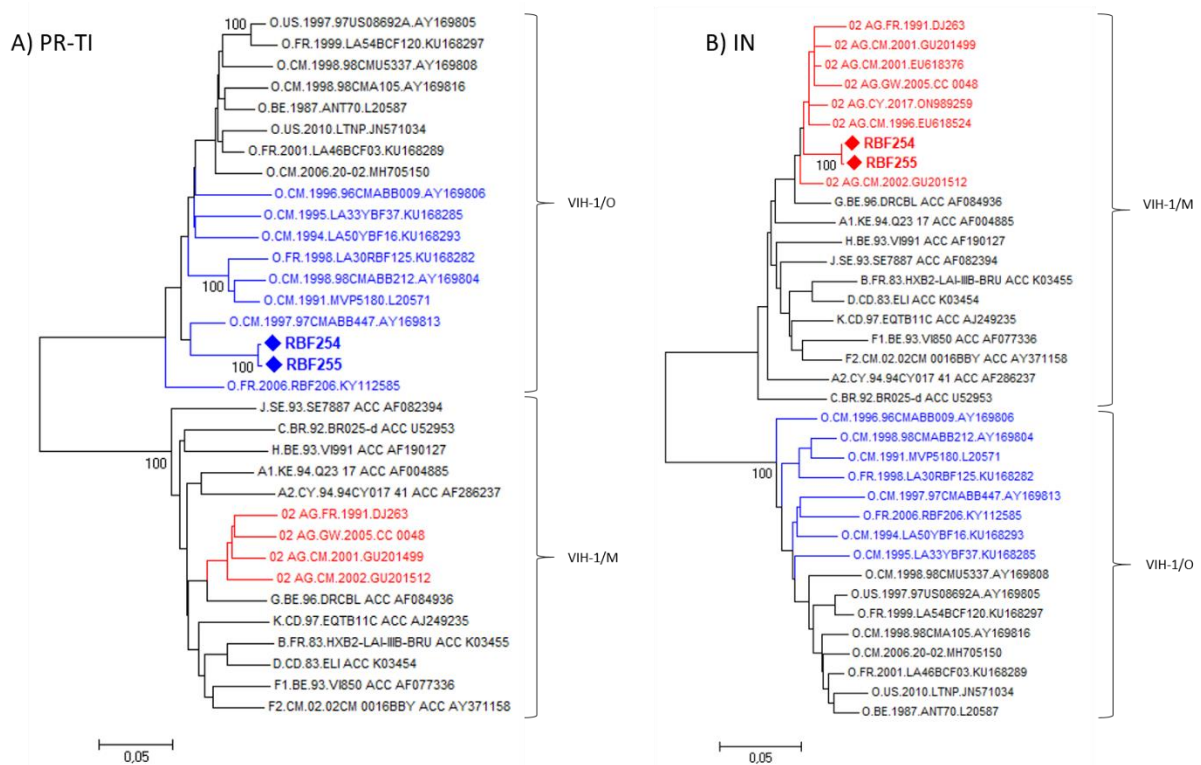


Figure 55 : Arbres phylogénétiques des régions PR-TI (A) et IN (B) de RBF254 et RBF255

Toutes les séquences ont été alignées avec le logiciel MEGA 7 avec un ensemble de séquences VIH-1/M et VIH-1/O de la base de données VIH de Los Alamos et de la base de données de notre laboratoire. Les arbres phylogénétiques ont été obtenus à l'aide de la méthode d'assemblage neighbor-joining. Les distances évolutives, correspondant aux substitutions par site, ont été calculées à l'aide de la méthode à 2 paramètres de Kimura et indiquées par une échelle au bas de chaque arbre. Les souches RBF254 et RBF255 sont mises en évidence en gras, précédées d'un symbole en forme de losange. Les VIH-1/M de sous-type CRF02\_AG ont été indiqués en rouge, tandis que les VIH-1/O de sous-groupe T ont été indiqués en bleu. La fiabilité a été estimée par une analyse bootstrap avec 1000 répétitions et tous ceux inférieurs à un score de 100 ont été supprimés.

L'outil Simplot a permis d'analyser l'ensemble du génome complet des deux souches obtenues, révélant la présence de deux points de cassure, aux mêmes positions pour RBF254 et RBF255 (Figure 56). Cette analyse a également permis de mettre en évidence la grande similarité entre les séquences de la mère et de son enfant sur la totalité du génome, du fait de la quasi superposition de RBF255 sur RBF254 (Figure 56).

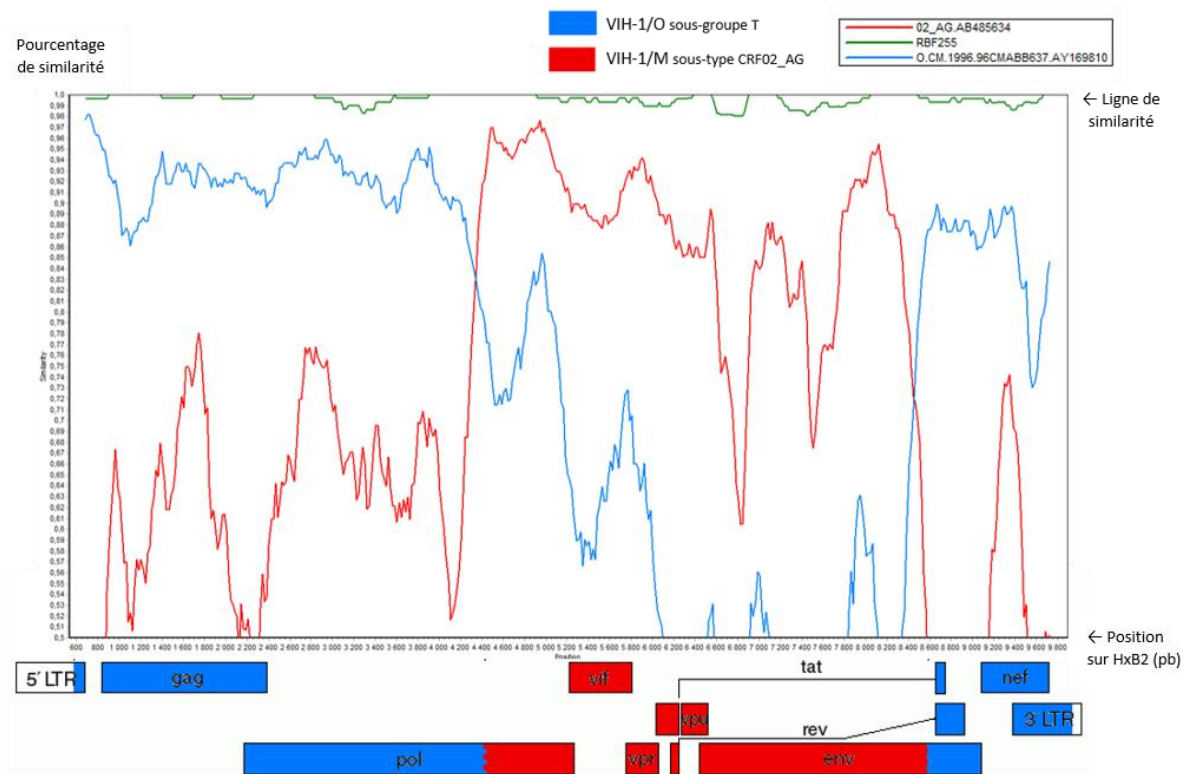


Figure 56 : Simplot de la séquence du génome complet de RBF254

L'analyse Simplot a permis d'identifier et de localiser les points de cassure le long du génome avec une fenêtre de 300 nt, un pas de 20 nt et une analyse statistique par méthode Kimura-2 avec un ratio de transition-transversion de 2,0. La séquence de RBF254 est représentée par la ligne 1.0 en haut du simplot, indiquant la ligne de similarité maximale. La séquence de RBF255 est représentée en vert. La séquence du VIH-1/M le plus proche (AB485634) est représentée en rouge. La séquence du VIH-1/O le plus proche (AY169810) est représentée en bleu. Le génome de la souche recombinante VIH-1/MO correspondante a été représenté en bas du simplot.

Les différents alignements et le Simplot ont permis de localiser précisément les deux points de cassure, le premier au tout début du gène de l'IN et le second, au début du domaine transmembranaire de la GP41, aux positions 4253 - 4293 pb et 8297 - 8318 pb, respectivement. Enfin, ces analyses ont permis de schématiser le profil des génomes complets de RBF254 et RBF255 (Figure 57).

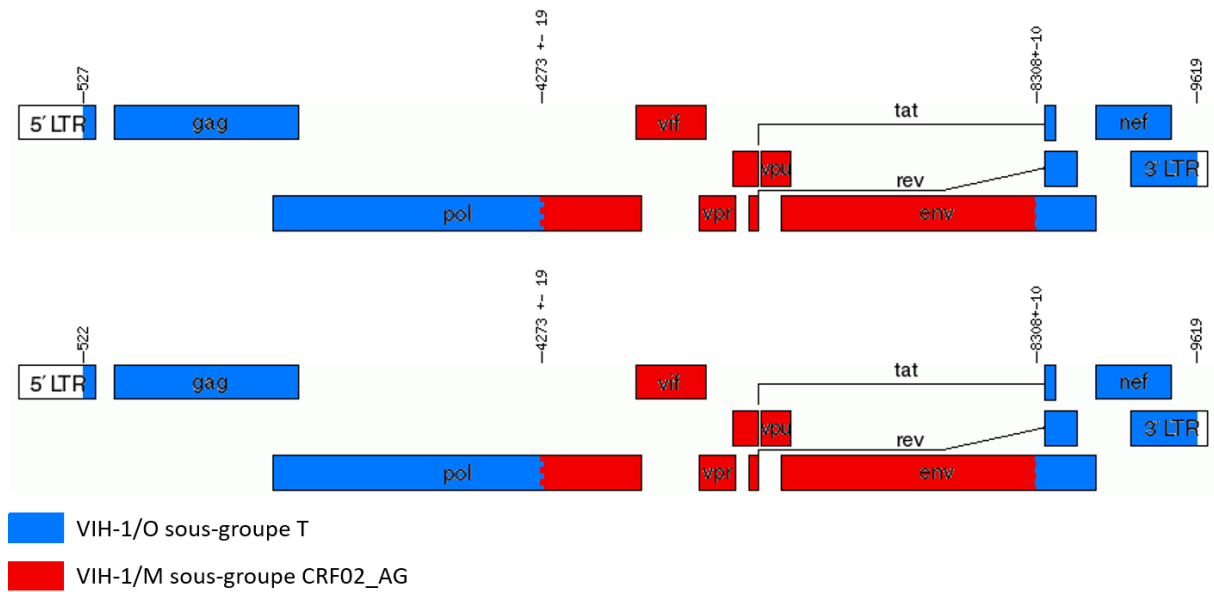


Figure 57 : Génomes complets de RBF254 et RBF255

Le profil moléculaire des deux formes RBF254 et RBF255 sont représentés en haut et bas de la figure respectivement. Les fragments de VIH-1/O ont été indiqués en bleu tandis que les fragments de VIH-1/M sont en rouge. Chaque position a été indiquée par rapport à la souche de référence HxB2. Les deux points de cassure ont été positionnés sur les génomes au début du gène de l'IN, aux positions 4253-4293 pb (4273±19) et dans la GP41 aux positions 8297-8318 pb (8308±10).

Une dernière analyse phylogénétique a été effectuée pour comparer cette forme recombinante VIH-1/MO avec 15 des 23 formes VIH-1/MO déjà complètement séquencées, représentant au final 12 URF\_MO identifiées *in vivo* (Figure 58). Les sept autres URF\_MO, représentés par des génomes partiels ont été inclus dans un second temps, sur la base de la caractérisation des fragments disponibles (données non montrées). Les données de ces deux analyses ont permis de conclure qu'il n'y avait pas de lien entre RBF254/RBF255 et l'une de ces 19 URF\_MO, amenant à 20 le nombre total d'URF\_MO désormais décrites.



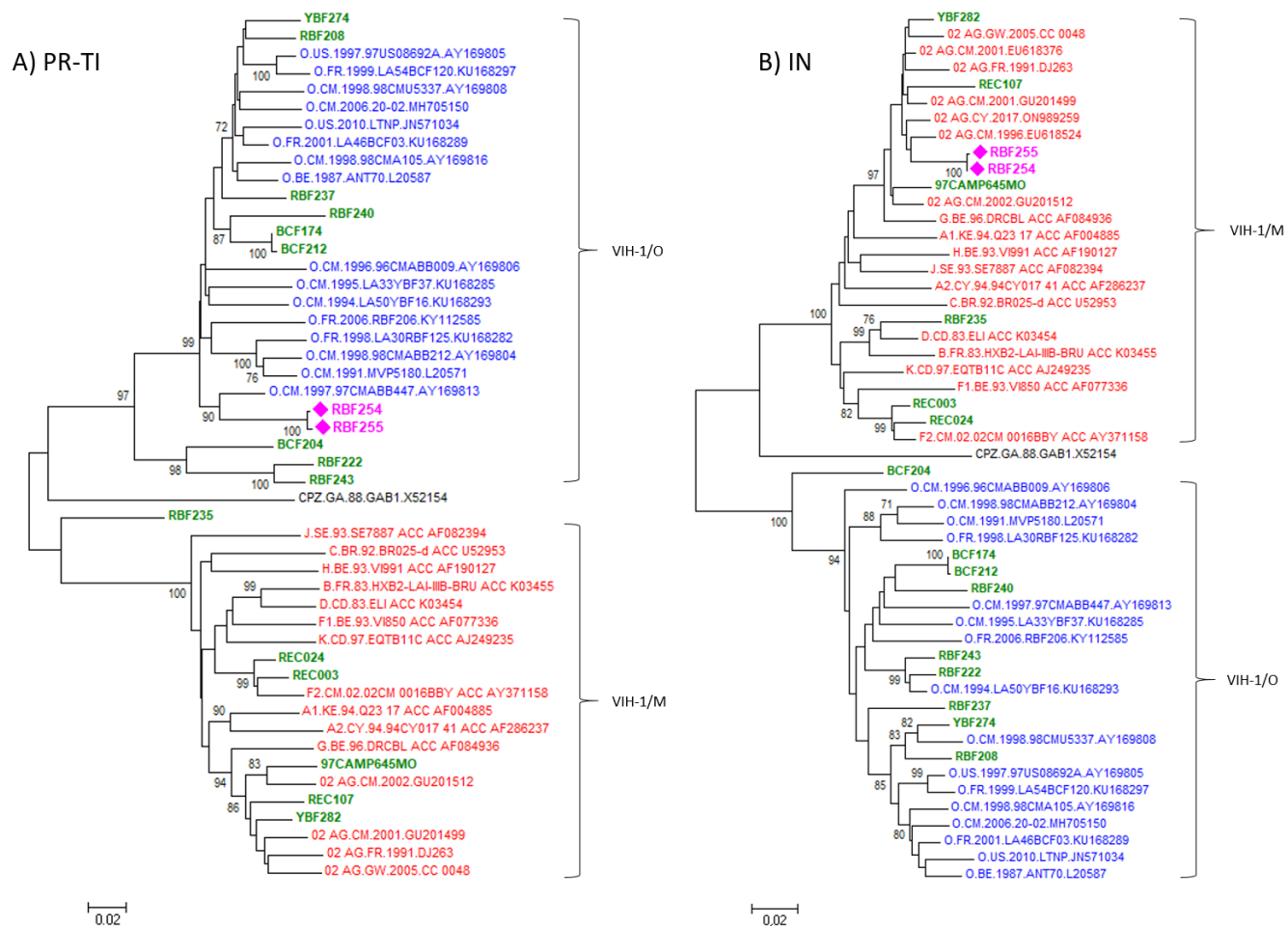


Figure 58 : Arbres phylogénétiques des régions PR-TI (A) et IN (B) de RBF254, RBF255 et des 15 VIH-1/MO complètement séquencées

Toutes les séquences ont été alignées avec le logiciel MEGA 7 avec un ensemble de séquences VIH-1/M et VIH-1/O de la base de données VIH de Los Alamos et de la base de données de notre laboratoire. Les arbres phylogénétiques ont été obtenus à l'aide de la méthode d'assemblage neighbor-joining. Les distances évolutives, correspondant aux substitutions par site, ont été calculées à l'aide de la méthode à 2 paramètres de Kimura et indiquées par une échelle au bas de chaque arbre. Les souches RBF254 et RBF255 sont mises en évidence en rose et en gras, précédées d'un symbole en forme de losange. Les 15 des 23 formes VIH-1/MO déjà complètement séquencées, représentant les 12 URF\_MO ont été indiquées en gras et vert. Les VIH-1/M ont été indiqués en rouge, tandis que les VIH-1/O ont été indiqués en bleu. Cet arbre phylogénétique a été enraciné à l'aide d'une séquence de VIScpz de Gabon (accession n°X52154) (49). La fiabilité a été estimée par une analyse bootstrap avec 1000 répétitions et tous ceux inférieurs à un score de 70 ont été supprimés.

### 3. Discussion

Tous les patients infectés par les formes recombinantes VIH-1/MO mises en évidence depuis 1999, présentaient un lien avec le Cameroun, où une importante diversité de VIH circule (15, 83). Le taux de TME était de 17% dans ce pays en 2021 en absence de traitement. La TME du VIH est un phénomène rare, qui touche majoritairement les pays à ressources limitées. Notre étude avait pour objectif de caractériser la transmission de l'une de ces formes recombinantes, identifiée au cours de l'année 2022, chez un bébé (RBF254) et sa maman (RBF254), tous deux d'origine camerounaise.

Dans notre étude, RBF255 a développé une infection opportuniste très rapidement après sa naissance, la pneumocystose, fréquemment retrouvée chez les personnes infectées par le VIH en stade SIDA. Cette maladie a pu être provoquée à cause d'un taux de CD4 inférieur à  $200/\text{mm}^3$  chez le nouveau-né, indiquant une forte immunodépression. Concernant la mère (RBF254), aucune donnée viro-clinique antérieure n'était disponible avant son arrivée en France, en janvier 2022. Au moment de son diagnostic, sa CV était de 5,4  $\text{Log}_{10}$  cp/mL, montrant une réplication intense du virus et associée à un taux de CD4 inférieur à  $500/\text{mm}^3$ . Comme ce qui est observé pour le VIH-1/M, une primo-infection en cours de grossesse a pu être associée à un risque de TME très important, même si les données dont nous disposons ne permettent pas de trancher entre une primo-infection *pré-* ou *per-partum* (225).

La TME du VIH-1/M a été largement étudiée et est désormais bien prise en charge, avec des moyens de prévention efficaces sous traitement ARV et contrôle de la CV maternelle. En revanche, peu de données sont disponibles sur les variants non-M du VIH. Deux études en France ont rapporté une TME des VIH-1/O, malgré la faible prévalence de ces infections (203, 204). Des cas de TME du VIH-2 ont également été décrits malgré un taux de transmission nettement inférieur à celui du VIH-1/M (201). Dans une étude française, une cohorte périnatale a montré que ce taux de TME du VIH-2 était d'environ 0,3% avec seulement deux cas identifiés chez 585 cas investigués (200). Les données actuelles sur le VIH-2 indiquent que la CV et le nombre de CD4, plus faibles, seraient en lien avec un taux de TME plus faible, même en l'absence de traitement de la mère.

Les données obtenues pour les VIH-1/M et pour le VIH-2, semblent indiquer que la diversité génétique influence le taux de TME, avec une transmission facilitée pour les VIH-1 au contraire des VIH-2 qui provoquent une infection atténuée. Notre étude sur la TME d'un VIH-1/MO apporte donc quelques informations complémentaires sur la transmission et la gravité de ces formes recombinantes. Le stade SIDA retrouvé chez RBF255 seulement 4 mois après sa naissance suggère que cette forme recombinante VIH-1/MO avait provoqué une infection grave et n'était pas atténuée *in vivo*. Les données de CV et du nombre de CD4 indiquaient une forte réplication de la forme recombinante,

comme chez RBF254, au contraire de ce qui est observé chez les VIH-2. Néanmoins, le faible nombre d'études sur la TME des VIH-1/O et la description de ce premier cas, pour les formes recombinantes VIH-1/MO, ne nous permettent pas de conclure quant à la gravité et l'impact de ces formes par rapport aux VIH-1/M et VIH-1/O.

Sur le plan thérapeutique, les différentes données collectées par le CHU de Lyon ont permis de voir une stagnation de la CV à 2 Log<sub>10</sub> cp/mL depuis août 2022 pour les deux patients, malgré la mise sous traitement. De plus, une remontée de la CV de RBF255 a été observée au cours du mois de juillet 2022, probablement lié un défaut d'observance. Ces données posaient le problème de l'émergence de résistance aux ARVs, avec l'acquisition de mutations sous pression thérapeutique inconstante.

La présence de fragments de VIH-1/O au sein du génome de ces formes n'a pas perturbé le diagnostic puisque les outils de détection sont désormais adaptés à la diversité génétique des VIH. Toutefois, l'utilisation d'amorces non adaptées au cours du génotypage de résistance en routine dans les laboratoires pose toujours des difficultés pour le séquençage de résistance. Le cas de RBF254 et RBF255 montre bien l'intérêt du génotypage spécifique de groupes M et O, qui est désormais le seul à pouvoir mettre en évidence des discordances suggérant des infections atypiques. Ces dernières doivent impérativement être explorées avec des outils adaptés, spécifiques de groupes, disponibles uniquement dans les laboratoires d'expertise (110).

L'analyse de la répartition des points de cassure le long du génome a été réalisée entre cette forme recombinante RBF254/RBF255 et les 12 URF-MO décrites *in vivo* et complètement séquencées. Ces données ont montré la présence d'un point de cassure dans la GP41, dont la position avait déjà été décrite en 2008 dans la forme recombinante RBF208 (135). Ce point de cassure était localisé à l'intérieur d'un gène, celui de l'enveloppe, contrairement à la majorité des autres formes recombinantes possédant des points de cassure aux extrémités des gènes (137). Des études avaient montré que la recombinaison avait lieu majoritairement au début ou à la fin des gènes, pour ne pas entraver leurs fonctionnalités et notamment pour ce gène (122–124). Néanmoins, ce point de cassure localisé au début du domaine transmembranaire de la GP41 n'a pas entravé sa fonctionnalité puisque le virus s'est montré très répliatif *in vivo* (135) et *in vitro* lors de ce travail expérimental (2. **Étude *in vitro* du potentiel répliatif de deux formes VIH-1/MO ; page 86**). Ces deux URF\_MO sont néanmoins différentes car elles ne possèdent ni le même second point de cassure, ni la même diversité génétique de fragments et sans lien phylogénétique.

Les deux techniques de séquençage réalisées au laboratoire ont permis d'obtenir les génomes complets de RBF254 et RBF255. Ces séquences ont été identifiées comme étant identiques, validant

notre protocole de séquençage NGS. Cette technologie est de plus en plus utilisée par les laboratoires de virologie (226–229), mais le défi reste de standardiser ces méthodes dans les laboratoires et de s'adapter constamment à la diversification de l'épidémie à VIH. Quelques protocoles ont été adaptés pour le séquençage des génomes complets du VIH-1 par technique NGS, mais se sont révélés peu adaptés à la diversité des souches identifiées en Afrique Sub-Saharienne et complexes à mettre en place (230, 231). Une étude plus récente réalisée par *Berg et al* en 2016 et adaptée par *Yamaguchi et al* en 2018, a proposé l'utilisation de sondes ciblant des régions assez conservées entre différentes souches du VIH pour pouvoir les détecter sans distinction et réaliser une librairie « universelle », permettant ainsi le séquençage des différents groupes de VIH-1 et le séquençage du VIH-2 par technique NGS (232, 233). Cette technique s'est révélée fastidieuse à mettre en place et ne pouvait être utilisée dans notre contexte d'étude. Notre étude a montré qu'il était possible de séquencer, les formes recombinantes VIH-1/MO, à l'aide d'un protocole fournisseur simple, de nouvelle génération, en utilisant des extraits ARN issus de plasmas et une adaptation des conditions d'amplification des fragments avant préparation de la librairie. Concernant ce protocole, l'amplification de nos souches a été plus difficile pour les fragments impliquant un point de recombinaison et notamment la GP41 au sein de l'enveloppe, déjà difficilement amplifiable car soumise à une large diversité génétique. Une adaptation des amorces d'amplification des génomes complets, plus adaptées à la diversité génétique des VIH, devra donc être effectuée dans de futurs travaux. Enfin, les prélèvements utilisés pour cette analyse présentaient de fortes CV, de 5,4 et 4,8  $\text{Log}_{10}$  cp/mL pour RBF254 et RBF255, respectivement. Il sera donc nécessaire par la suite de déterminer l'impact de cette CV sur la possibilité de séquencer les échantillons par cette technique.

Le NGS permet un gain de temps considérable, tout en étant plus informatif car il permet d'étudier la diversité des populations composant une même quasi-espèce (228, 234–237). De plus, la technique Sanger mise en place a été très lourde, avec la réalisation de plusieurs essais de séquençage pour les fragments impliquant un point de recombinaison, la purification sur gel de la GP41 et l'utilisation d'amorces spécifiques de groupe pour obtenir la séquence finale. Le protocole NGS mis en place au cours de ce projet permettait de reconstituer les génomes complets de nos deux souches de façon très rapide et également d'étudier la diversité des populations présentes dans ces échantillons, incluant la recherche de formes parentales VIH-1/M et VIH-1/O. La reconstitution des génomes complets par la méthode *de novo* nécessitait d'assembler des fragments de VIH-1/M et VIH-1/O de part et d'autre d'un point de cassure, ne facilitant pas la reconstruction finale. L'obtention du génome complet pour RBF254 a nécessité un deuxième alignement des *reads* sur les séquences obtenues pour RBF255, du fait d'une plus forte diversité intra-échantillon. Aucune forme parentale n'a été mise en évidence au cours de l'analyse étayant l'hypothèse d'une transmission directe de la forme

recombinante à RBF254. Ces données montraient une infection plus ancienne chez la mère, avec la présence d'une quasi-espèce plus diversifiée que pour son bébé. Ces données préliminaires montrent le potentiel du NGS et la nécessité d'adapter cette technologie à la diversité des VIH existante actuellement. Des investigations sont néanmoins encore nécessaires, afin de caractériser au mieux les populations constituant les échantillons, car le NGS nécessite des analyses bio-informatiques complexes, du fait de la présence de plusieurs fragments de groupes M et O.

Peu de données sont disponibles sur ces formes recombinantes VIH-1/MO. La description d'une transmission directe de l'une d'entre elle au sein d'un couple a démontré la transmissibilité de ces formes (136); ici, nous l'avons démontré dans le contexte d'une TME. Le séquençage des génomes complets et l'analyse phylogénétique ont permis de prouver la transmission directe de cette nouvelle URF\_MO de RBF254 à RBF255. Cette nouvelle URF\_MO possède une forte réplication et entraîne une infection grave puisqu'en l'absence de traitement, le taux de CD4 a très largement diminué chez ces deux patients, provoquant un stade SIDA après 4 mois de vie seulement pour RBF255. L'ensemble de ces données montre que ces formes recombinantes MO ont un potentiel de transmissibilité et répliatif très important et qu'il est donc essentiel de les détecter et de bien les caractériser.

## Conclusion / perspectives

Malgré la divergence génétique importante entre les VIH-1/M et VIH-1/O, 20 URF\_MO sont désormais recensées à ce jour. Cette étude visait l'étude du potentiel répliatif de ces formes recombinantes entre elle et avec leurs formes parentales et a permis de mettre en évidence la transmission de ces formes de la mère à l'enfant.

Malgré de nombreuses tentatives de génération de CMI et CMIC effectuées au cours de ce projet, à la fois par notre laboratoire et par la société Genecust, nous n'avons pas obtenu de clone stable et/ou infectieux après clonage. Notre modèle expérimental basé sur la génération de ces clones selon la méthode de Gibson Assembly, avait été construit à partir des données de la littérature pour le positionnement des points de cassure et pour la modélisation de nos candidats. Ces éléments devaient faciliter le clonage des formes recombinantes VIH-1/MO sous forme de CMIC ; néanmoins nos travaux ont confirmé la complexité d'obtention de clone stables et infectieux. Au final, un nouveau CMI VIH-1/M (pM-A-062021) a été produit au cours de cette étude, s'ajoutant à ceux déjà décrits dans la littérature. Ces outils restent nécessaires pour l'étude de la diversité génétique des VIH et d'autres méthodologies devront être développées à l'avenir.

La comparaison de deux souches recombinantes de VIH\_1/MO, OM-TB-2016 et RBF208, a montré qu'elles avaient le même potentiel répliatif, malgré une diversité génétique et un profil de recombinaison différents, dans nos conditions de lancement des cinétiques *in vitro*. Les données obtenues ont permis de mesurer le taux de répliation de chaque souche au travers de deux paramètres : la quantité d'antigène p24 et l'activité de la TI. Aucune autre étude n'a jusqu'à présent permis de comparer le potentiel répliatif des souches recombinantes VIH-1/MO entre elles et notre étude apporte donc des résultats préliminaires sur cette question, même si les souches testées incluaient toutes les deux une TI de VIH-1/O et une GP120 de VIH-1/M. De plus, ce projet a montré que l'utilisation de CMSPs était préférable aux cellules de lignées cellulaires, telles que les Jurkat X4-R5, pour la comparaison des potentiels répliatifs de telles formes, puisqu'elle permettait de s'affranchir des problématiques d'expression des différents corécepteurs du VIH. La technique mise en place au cours de cette étude a consisté à mettre les souches en culture parallèlement, afin de réaliser des comparaisons de cinétiques de répliation. D'autres techniques permettent d'estimer le *fitness* répliatif d'un virus, par la mise en compétition de plusieurs souches au sein d'une même culture (238, 239). Cependant, une véritable étude du *fitness* n'a pas pu être mise en place car elle nécessitait des moyens de révélation discriminants entre nos souches, non disponibles à l'heure actuelle.

Notre travail a également permis de décrire pour la première fois un cas de TME d'une forme recombinante VIH-1/MO. L'analyse phylogénétique a permis de confirmer que ces formes sont bien transmissibles et les données cliniques ont montré qu'elles peuvent provoquer des infections graves. L'observation de très fortes CV pour nos deux patients, couplée à de faibles taux de CD4, ont montré une réplication intense, suggérant une bonne adaptation de cette nouvelle URF\_MO *in vivo*. En ajoutant les données obtenues au cours de cette étude pour cette nouvelle URF\_MO, à celles obtenues en 2019 dans une précédente étude, 67% des VIH-1/MO identifiées *in vivo* possèdent ce profil de recombinaison [O<sub>gag/pol</sub>-M<sub>env</sub>], contre seulement 33% ayant le profil inverse [M<sub>gag/pol</sub>-O<sub>env</sub>]. La mise en évidence d'un avantage répliatif de telles formes n'a pas pu être effectuée au cours du précédent volet, mais la description de cette nouvelle forme de profil [O<sub>gag/pol</sub>-M<sub>env</sub>], ayant une enveloppe de VIH-1/M couplée à une TI de VIH-1/O, est un argument supplémentaire en faveur de notre hypothèse selon laquelle les formes recombinantes ayant cette combinaison de fragments génétiques pourraient être avantagées. Enfin, cette étude a permis de montrer qu'il était possible de séquencer ces VIH-1/MO par une technique de séquençage NGS, mise en place au cours de ce projet ; celle-ci devra néanmoins être optimisée, à la fois pour inclure une plus grande diversité de VIH et pour permettre une étude plus approfondie des populations de la quasi-espèce intra-échantillon.

En conclusion, la diversité génétique liée à la recombinaison des VIH-1 reste un défi important pour les chercheurs, puisque de nouvelles souches recombinantes émergent chaque année. Ce phénomène est largement étudié mais reste un sujet complexe malgré l'ensemble des connaissances déjà disponibles sur le sujet. La recombinaison inter-groupes MO en particulier, implique la génération de formes, dont les propriétés restent à étudier afin de déterminer si elles représentent un risque accru ou non par rapport aux formes parentales actuelles. La recombinaison inter-groupes MO implique aussi de rester vigilant lors de l'observation de discordances au moment du génotypage de résistance des VIH en routine.

## REFERENCES

1. Centers for Disease Control (CDC). 1981. Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California. *MMWR Morb Mortal Wkly Rep* 30:305–308.
2. Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vézinet-Brun F, Rouzioux C, Rozenbaum W, Montagnier L. 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 220:868–871.
3. Broder S, Gallo RC. 1984. A pathogenic retrovirus (HTLV-III) linked to AIDS. *N Engl J Med* 311:1292–1297.
4. Popovic M, Sarngadharan MG, Read E, Gallo RC. 1984. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science* 224:497–500.
5. Poiesz BJ, Ruscetti FW, Gazdar AF, Bunn PA, Minna JD, Gallo RC. 1980. Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc Natl Acad Sci U S A* 77:7415–7419.
6. Wain-Hobson S, Sonigo P, Danos O, Cole S, Alizon M. 1985. Nucleotide sequence of the AIDS virus, LAV. *Cell* 40:9–17.
7. Mboup S, Essex M, Kank P, Ndoye I, Barin F, Samba null, Boye CS, Ricard D, Denis F, Romet JL. 1986. [The HTL IV virus: a new human retrovirus without pathogen protection against the HTLV III/LAV virus?]. *Afr J Sex Transmi Dis* 2:81–84.
8. Clavel F, Guétard D, Brun-Vézinet F, Chamaret S, Rey MA, Santos-Ferreira MO, Laurent AG, Dauguet C, Katlama C, Rouzioux C. 1986. Isolation of a new human retrovirus from West African patients with AIDS. *Science* 233:343–346.
9. De Leys R, Vanderborght B, Vanden Haesevelde M, Heyndrickx L, van Geel A, Wauters C, Bernaerts R, Saman E, Nijs P, Willems B. 1990. Isolation and partial characterization of an unusual human immunodeficiency retrovirus from two persons of west-central African origin. *J Virol* 64:1207–1216.
10. Simon F, Mauclère P, Roques P, Loussert-Ajaka I, Müller-Trutwin MC, Saragosti S, Georges-Courbot MC, Barré-Sinoussi F, Brun-Vézinet F. 1998. Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nat Med* 4:1032–1037.
11. Plantier J-C, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Lemée V, Damond F, Robertson DL, Simon F. 2009. A new human immunodeficiency virus derived from gorillas. *Nat Med* 15:871–872.
12. Ayouba A, Akoua-Koffi C, Calvignac-Spencer S, Esteban A, Locatelli S, Li H, Li Y, Hahn BH, Delaporte E, Leendertz FH, Peeters M. 2013. Evidence for continuing cross-species transmission of SIVsmm to humans: characterization of a new HIV-2 lineage in rural Côte d'Ivoire. *AIDS* 27:2488–2491.
13. Données épidémiologiques VIH/sida monde - 2021. Sidaction. <https://www.sidaction.org/donnees-epidemiologiques-vihsida-monde-2021>. Retrieved 24 April 2023.
14. 2021. Global HIV & AIDS statistics — Fact sheet. <https://www.unaids.org/en/resources/fact-sheet>. Retrieved 2 January 2023.



15. AIDSinfo | UNAIDS. <https://aidsinfo.unaids.org/>. Retrieved 2 January 2023.
16. SPF. Bulletin de santé publique VIH-IST. Décembre 2022. <https://www.santepubliquefrance.fr/maladies-et-traumatismes/infections-sexuellement-transmissibles/vih-sida/documents/bulletin-national/bulletin-de-sante-publique-vih-ist.-decembre-2022>. Retrieved 24 April 2023.
17. Données épidémiologiques VIH/sida France 2021. Sidaction. <https://www.sidaction.org/donnees-epidemiologiques-vihsida-france-2021>. Retrieved 5 June 2023.
18. Girard P-M, Katlama C, Pialoux G. 2011. VIH. Doin.
19. Transmission mère-enfant du VIH : un succès de la prévention. <https://www.larevuedupraticien.fr/article/transmission-mere-enfant-du-vih-un-succes-de-la-prevention>. Retrieved 29 June 2023.
20. SPF. Transmission mère-enfant du VIH en France : l'impact majeur des stratégies de prévention - Résultats de l'Enquête périnatale française ANRS-EPF. Numéro thématique - Infections congénitales et transmises de la mère à l'enfant en France : des progrès not... <https://www.santepubliquefrance.fr/maladies-et-traumatismes/infections-sexuellement-transmissibles/vih-sida/transmission-mere-enfant-du-vih-en-france-l-impact-majeur-des-strategies-de-prevention-resultats-de-l-enquete-perinatale-francaise-anrs-epf.-nu>. Retrieved 29 June 2023.
21. CNS W. 2019. Prise en charge du VIH - Recommandations du groupe d'experts. Conseil national du sida et des hépatites virales. <https://cns.sante.fr/actualites/prise-en-charge-du-vih-recommandations-du-groupe-dexperts/>. Retrieved 12 May 2023.
22. Mandelbrot L, Tubiana R, Le Chenadec J, Dollfus C, Faye A, Pannier E, Matheron S, Khuong M-A, Garrait V, Reliquet V, Devidas A, Berrebi A, Allisy C, Elleau C, Arvieux C, Rouzioux C, Warszawski J, Blanche S, ANRS-EPF Study Group. 2015. No perinatal HIV-1 transmission from women with effective antiretroviral therapy starting before conception. *Clin Infect Dis* 61:1715–1725.
23. Mourez T, Burrel S, Boutolleau D, Pillet S. 2019. *Traité de virologie médicale*. Société française de microbiologie.
24. HIV Databases. <https://www.hiv.lanl.gov/content/index>. Retrieved 1 February 2023.
25. Cassan E, Arigon-Chifolleau A-M, Mesnard J-M, Gross A, Gascuel O. 2016. Concomitant emergence of the antisense protein gene of HIV-1 and of the pandemic. *Proc Natl Acad Sci U S A* 113:11537–11542.
26. Gholizadeh Z, Iqbal MohdS, Li R, Romerio F. 2021. The HIV-1 Antisense Gene ASP: The New Kid on the Block. *Vaccines (Basel)* 9:513.
27. Mancarella A, Procopio FA, Achsel T, De Crignis E, Foley BT, Corradin G, Bagni C, Pantaleo G, Graziosi C. 2019. Detection of antisense protein (ASP) RNA transcripts in individuals infected with human immunodeficiency virus type 1 (HIV-1). *J Gen Virol* 100:863–876.
28. Affram Y, Zapata JC, Gholizadeh Z, Tolbert WD, Zhou W, Iglesias-Ussel MD, Pazgier M, Ray K, Latinovic OS, Romerio F. 2019. The HIV-1 Antisense Protein ASP Is a Transmembrane Protein of the Cell Surface and an Integral Protein of the Viral Envelope. *J Virol* 93:e00574-19.
29. Sathiyamani B, Daniel EA, Ansar S, Esakialraj BH, Hassan S, Revanasiddappa PD, Keshavamurthy A, Roy S, Vetrivel U, Hanna LE. 2023. Structural analysis and molecular dynamics simulation studies of HIV-1 antisense protein predict its potential role in HIV replication and pathogenesis. *Front Microbiol* 14:1152206.

30. Klasens BIF, Das AT, Berkhout B. 1998. Inhibition of polyadenylation by stable RNA secondary structure. *Nucleic Acids Research* 26:1870–1876.
31. Klasens BI, Thiesen M, Virtanen A, Berkhout B. 1999. The ability of the HIV-1 AAUAAA signal to bind polyadenylation factors is controlled by local RNA structure. *Nucleic Acids Res* 27:446–454.
32. Brand D, Truong C, Barin F. 1994. Domaines fonctionnels de l'enveloppe du VIH-1 et anticorps neutralisants. *Med Sci (Paris)* 10:417.
33. Chen B. 2019. Molecular Mechanism of HIV-1 Entry. *Trends in Microbiology* 27:878–891.
34. Zila V, Müller TG, Müller B, Kräusslich H-G. 2021. HIV-1 capsid is the key orchestrator of early viral replication. *PLoS Pathog* 17:e1010109.
35. Toccafondi E, Lener D, Negroni M. 2021. HIV-1 Capsid Core: A Bullet to the Heart of the Target Cell. *Front Microbiol* 12:652486.
36. René B, Mauffret O, Fossé P. 2018. Retroviral nucleocapsid proteins and DNA strand transfers. *Biochimie Open* 7:10–25.
37. Simon-Lorière E, Galetto R, Negroni M. 2006. Caractéristiques mécanistiques de la recombinaison chez le virus de l'immunodéficience humaine 10:12.
38. Bandera A, Gori A, Clerici M, Sironi M. 2019. Phylogenies in ART: HIV reservoirs, HIV latency and drug resistance. *Curr Opin Pharmacol* 48:24–32.
39. Van Lint C, Bouchat S, Marcello A. 2013. HIV-1 transcription and latency: an update. *Retrovirology* 10:67.
40. Marini B, Kertesz-Farkas A, Ali H, Lucic B, Lisek K, Manganaro L, Pongor S, Luzzati R, Recchia A, Mavilio F, Giacca M, Lusich M. 2015. Nuclear architecture dictates HIV-1 integration site selection. *Nature* 521:227–231.
41. Engelman A, Cherepanov P. 2012. The structural biology of HIV-1: mechanistic and therapeutic insights. 4. *Nat Rev Microbiol* 10:279–290.
42. Daugherty MD, Liu B, Frankel AD. 2010. Structural basis for cooperative RNA binding and export complex assembly by HIV Rev. 11. *Nat Struct Mol Biol* 17:1337–1342.
43. Beemon KL. 2022. Retroviral RNA Processing. *Viruses* 14:1113.
44. Emery A, Swanstrom R. 2021. HIV-1: To Splice or Not to Splice, That Is the Question. *Viruses* 13:181.
45. Hanson HM, Willkomm NA, Yang H, Mansky LM. 2022. Human Retrovirus Genomic RNA Packaging. *Viruses* 14:1094.
46. Ye L, Ambi UB, Olguin-Nava M, Gribling-Burrer A-S, Ahmad S, Bohn P, Weber MM, Smyth RP. 2021. RNA Structures and Their Role in Selective Genome Packaging. *Viruses* 13:1788.
47. Freed EO. 2015. HIV-1 assembly, release and maturation. 8. *Nat Rev Microbiol* 13:484–496.
48. Peeters M, Honoré C, Huet T, Bedjabaga L, Ossari S, Bussi P, Cooper RW, Delaporte E. 1989. Isolation and partial characterization of an HIV-related virus occurring naturally in chimpanzees in Gabon. *AIDS* 3:625–630.
49. Huet T, Cheynier R, Meyerhans A, Roelants G, Wain-Hobson S. 1990. Genetic organization of a chimpanzee lentivirus related to HIV-1. 6273. *Nature* 345:356–359.
50. Sauter D, Kirchhoff F. 2019. Key Viral Adaptations Preceding the AIDS Pandemic. *Cell Host & Microbe* 25:27–38.

51. Keele BF. 2006. Chimpanzee Reservoirs of Pandemic and Nonpandemic HIV-1. 5786. *Science* 313:523–526.
52. Van Heuverswyn F, Li Y, Bailes E, Neel C, Lafay B, Keele BF, Shaw KS, Takehisa J, Kraus MH, Loul S, Butel C, Liegeois F, Yangda B, Sharp PM, Mpoudi-Ngole E, Delaporte E, Hahn BH, Peeters M. 2007. Genetic diversity and phylogeographic clustering of SIVcpzPtt in wild chimpanzees in Cameroon. *Virology* 368:155–171.
53. D'arc M, Ayouba A, Esteban A, Learn GH, Boué V, Liegeois F, Etienne L, Tagg N, Leendertz FH, Boesch C, Madinda NF, Robbins MM, Gray M, Cournil A, Ooms M, Letko M, Simon VA, Sharp PM, Hahn BH, Delaporte E, Mpoudi Ngole E, Peeters M. 2015. Origin of the HIV-1 group O epidemic in western lowland gorillas. *Proc Natl Acad Sci U S A* 112:E1343-1352.
54. Sharp PM, Hahn BH. 2011. Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect Med* 1:a006841.
55. Hirsch VM, Olmsted RA, Murphey-Corb M, Purcell RH, Johnson PR. 1989. An African primate lentivirus (SIVsmclosely related to HIV-2. 6223. *Nature* 339:389–392.
56. Tebit DM, Arts EJ. 2011. Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. *The Lancet Infectious Diseases* 11:45–56.
57. Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, Tatem AJ, Sousa JD, Arinaminpathy N, Pépin J, Posada D, Peeters M, Pybus OG, Lemey P. 2014. The early spread and epidemic ignition of HIV-1 in human populations. *Science* 346:56–61.
58. Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, Bunce M, Muyembe J-J, Kabongo J-MM, Kalengayi RM, Van Marck E, Gilbert MTP, Wolinsky SM. 2008. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature* 455:661–664.
59. Leoz M, Feyertag F, Kfutwah A, Mauclère P, Lachenal G, Damond F, De Oliveira F, Lemée V, Simon F, Robertson DL, Plantier J-C. 2015. The Two-Phase Emergence of Non Pandemic HIV-1 Group O in Cameroon. *PLoS Pathog* 11:e1005029.
60. Wertheim JO, Worobey M. 2009. Dating the Age of the SIV Lineages That Gave Rise to HIV-1 and HIV-2. *PLoS Comput Biol* 5:e1000377.
61. Sauter D, Hué S, Petit SJ, Plantier J-C, Towers GJ, Kirchhoff F, Gupta RK. 2011. HIV-1 Group P is unable to antagonize human tetherin by Vpu, Env or Nef. *Retrovirology* 8:103.
62. Visseaux B, Bertine M, Le Hingrat Q, Ferré V, Charpentier C, Collin F, Damond F, Matheron S, Hué S, Descamps D. 2021. HIV-2 diversity displays two clades within group A with distinct geographical distribution and evolution. *Virus Evol* 7:veab024.
63. Peeters M, Courgnaud V, Abela B, Auzel P, Pourrut X, Bibollet-Ruche F, Loul S, Liegeois F, Butel C, Koulagna D, Mpoudi-Ngole E, Shaw GM, Hahn BH, Delaporte E. 2002. Risk to Human Health from a Plethora of Simian Immunodeficiency Viruses in Primate Bushmeat. *Emerg Infect Dis* 8:451–457.
64. Wain LV, Bailes E, Bibollet-Ruche F, Decker JM, Keele BF, Van Heuverswyn F, Li Y, Takehisa J, Ngole EM, Shaw GM, Peeters M, Hahn BH, Sharp PM. 2007. Adaptation of HIV-1 to Its Human Host. *Molecular Biology and Evolution* 24:1853–1860.
65. Alessandri-Gradt E, De Oliveira F, Leoz M, Lemee V, Robertson DL, Feyertag F, Ngoupo P-A, Mauclere P, Simon F, Plantier J-C. 2018. HIV-1 group P infection: towards a dead-end infection? *AIDS* 32:1317–1322.
66. Vallari A, Holzmayer V, Harris B, Yamaguchi J, Ngansop C, Makamche F, Mbanya D, Kaptué L, Ndemi N, Gürtler L, Devare S, Brennan CA. 2011. Confirmation of Putative HIV-1 Group P in Cameroon. *J Virol* 85:1403–1407.

67. Toccafondi E, Kanja M, Winter F, Lener D, Negroni M. 2023. A snapshot on HIV-1 evolution through the identification of phylogenetic-specific properties of HIV-1 integrases M/O. *PLoS Pathog* 19:e1011207.
68. Neil SJD, Zang T, Bieniasz PD. 2008. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* 451:425–430.
69. Sauter D, Schindler M, Specht A, Landford WN, Münch J, Kim K-A, Votteler J, Schubert U, Bibollet-Ruche F, Keele BF, Takehisa J, Ogando Y, Ochsenbauer C, Kappes JC, Ayouba A, Peeters M, Learn GH, Shaw G, Sharp PM, Bieniasz P, Hahn BH, Hatzioannou T, Kirchhoff F. 2009. Tetherin-driven adaptation of Vpu and Nef function and the evolution of pandemic and nonpandemic HIV-1 strains. *Cell Host Microbe* 6:409–421.
70. Hu W-S, Rhodes T, Dang Q, Pathak V. 2003. Retroviral recombination: review of genetic analyses. *Front Biosci* 8:d143-155.
71. Mansky LM, Temin HM. 1995. Lower *in vivo* mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J Virol* 69:5087–5094.
72. Preston BD, Poiesz BJ, Loeb LA. 1988. Fidelity of HIV-1 reverse transcriptase. *Science* 242:1168–1171.
73. Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* 373:123–126.
74. Domingo E, García-Crespo C, Perales C. 2021. Historical Perspective on the Discovery of the Quasispecies Concept. *Annu Rev Virol* 8:51–72.
75. Singh K, Mehta D, Dumka S, Chauhan AS, Kumar S. 2023. Quasispecies Nature of RNA Viruses: Lessons from the Past. *Vaccines* 11.
76. Chen J, Zhou T, Zhang Y, Luo S, Chen H, Chen D, Li C, Li W. 2022. The reservoir of latent HIV. *Front Cell Infect Microbiol* 12:945956.
77. Rambaut A, Posada D, Crandall KA, Holmes EC. 2004. The causes and consequences of HIV evolution. 1. *Nat Rev Genet* 5:52–61.
78. Yamaguchi J, Vallari A, McArthur C, Sthreshley L, Cloherty GA, Berg MG, Rodgers MA. 2020. Brief Report: Complete Genome Sequence of CG-0018a-01 Establishes HIV-1 Subtype L. *J Acquir Immune Defic Syndr* 83:319–322.
79. Hemelaar J, Gouws E, Ghys PD, Osmanov S. 2006. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* 20:W13-23.
80. Robertson DL. 2000. HIV-1 Nomenclature Proposal. 5463. *Science* 288:55d–555.
81. 2023. HIV Circulating Recombinant Forms (CRFs). <https://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html>. Retrieved 6 September 2022.
82. Montavon C, Toure-Kane C, Nkengasong JN, Vergne L, Hertogs K, Mboup S, Delaporte E, Peeters M. 2002. CRF06-cpx: a new circulating recombinant form of HIV-1 in West Africa involving subtypes A, G, K, and J. *J Acquir Immune Defic Syndr* 29:522–530.
83. Hemelaar J, Elangovan R, Yun J, Dickson-Tetteh L, Fleminger I, Kirtley S, Williams B, Gouws-Williams E, Ghys PD, WHO–UNAIDS Network for HIV Isolation Characterisation. 2019. Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis. *Lancet Infect Dis* 19:143–155.

84. Hemelaar J, Elangovan R, Yun J, Dickson-Tetteh L, Kirtley S, Gouws-Williams E, Ghys PD, WHO-UNAIDS Network for HIV Isolation and Characterisation. 2020. Global and regional epidemiology of HIV-1 recombinants in 1990-2015: a systematic review and global survey. *Lancet HIV* 7:e772–e781.
85. Bbosa N, Kaleebu P, Ssemwanga D. 2019. HIV subtype diversity worldwide. *Current Opinion in HIV and AIDS* 14:153.
86. Leoz M. 2016. HIV-1 Group O Genetic Diversity : Characterization, Evolution and Associated Viral Properties. These de doctorat. Rouen.
87. Roques P, Robertson DL, Souquière S, Damond F, Ayouba A, Farfara I, Depienne C, Nerrienet E, Dormont D, Brun-Vézinet F, Simon F, Maucière P. 2002. Phylogenetic analysis of 49 newly derived HIV-1 group O strains: high viral diversity but no group M-like subtype structure. *Virology* 302:259–273.
88. Quiñones-Mateu ME, Albright JL, Mas A, Soriano V, Arts EJ. 1998. 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* 72:9002–9015.
89. Leoz M, Depatureaux A, Vessière A, Roquebert B, Damond F, Rousset D, Roques P, Simon F, Plantier J-C, RES-O group. 2008. Integrase polymorphism and HIV-1 group O diversity. *AIDS* 22:1239–1243.
90. Depatureaux A, Charpentier C, Leoz M, Unal G, Damond F, Kfutwah A, Vessière A, Simon F, Plantier J-C. 2011. 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* 56:139–145.
91. Kouanfack C, Unal G, Schaeffer L, Kfutwah A, Aghokeng A, Mougnotou R, Tchémgué-Noumsi N, Alessandri-Gradt E, Delaporte E, Simon F, Vray M, Plantier J-C, ANRS 12168 DynaMO Study, Alima M, Essengué L, Mounpou G, Ngang P, Ngoma P, Omgba V, Omam D, Tonfack L, Zé F, Akongnwi E, Dumortier J, Ngoupo P-A, Njouom R, Rousset D, Le Foulher L, Madec Y, Bodelet M, Dupre J-M, Pavie J, Rekacewicz C, Garcia P, Colin G. 2019. Comparative Immunovirological and Clinical Responses to Antiretroviral Therapy Between HIV-1 Group O and HIV-1 Group M Infected Patients. *Clinical Infectious Diseases* ciz371.
92. Unal G, et al. 2018. Human Immunodeficiency Virus Type 1 Group O Infection in France: Clinical Features and Immunovirological Response to Antiretrovirals. 11. *Clinical Infectious Diseases* 66:1785–1793.
93. Santoro MM, Perno CF. 2013. HIV-1 Genetic Variability and Clinical Implications. *ISRN Microbiol* 2013:481314.
94. Yamaguchi J, Vallari AS, Swanson P, Bodelle P, Kaptué L, Ngansop C, Zekeng L, Gürtler LG, Devare SG, Brennan CA. 2002. Evaluation of HIV type 1 group O isolates: identification of five phylogenetic clusters. *AIDS Res Hum Retroviruses* 18:269–282.
95. Tebit DM, Patel H, Ratcliff A, Alessandri E, Liu J, Carpenter C, Plantier J-C, Arts EJ. 2016. HIV-1 Group O Genotypes and Phenotypes: Relationship to Fitness and Susceptibility to Antiretroviral Drugs 13.
96. Zhang L, Carruthers CD, He T, Huang Y, Cao Y, Wang G, Hahn B, Ho DD. 1997. HIV type 1 subtypes, coreceptor usage, and CCR5 polymorphism. *AIDS Res Hum Retroviruses* 13:1357–1366.
97. Tagnoukam Ngoupo PA, Sadeuh-Mba SA, De Oliveira F, Ngo-Malabo ET, Ngonu L, Plantier J-C, Kfutwah A, Njouom R. 2018. Short Communication: Characterization of a New HIV-1 Group N

- Isolate Originating from a Cameroonian Patient. *AIDS Research and Human Retroviruses* 34:621–625.
98. Delaugerre C, De Oliveira F, Lascoux-Combe C, Plantier J-C, Simon F. 2011. HIV-1 group N: travelling beyond Cameroon. *Lancet* 378:1894.
  99. Visseaux B, Damond F, Matheron S, Descamps D, Charpentier C. 2016. Hiv-2 molecular epidemiology. *Infection, Genetics and Evolution* 46:233–240.
  100. Ibe S, Yokomaku Y, Shiino T, Tanaka R, Hattori J, Fujisaki S, Iwatani Y, Mamiya N, Utsumi M, Kato S, Hamaguchi M, Sugiura W. 2010. HIV-2 CRF01\_AB: First Circulating Recombinant Form of HIV-2. *JAIDS Journal of Acquired Immune Deficiency Syndromes* 54:241–247.
  101. Yamaguchi J, Vallari A, Ndembi N, Coffey R, Ngansop C, Mbanya D, Kaptué L, Gürtler LG, Devare SG, Brennan CA. 2008. HIV Type 2 Intergroup Recombinant Identified in Cameroon. *AIDS Research and Human Retroviruses* 24:86–91.
  102. Heyndrickx L, Alary M, Janssens W, Davo N, van der Groen G. 1996. HIV-1 group O and group M dual infection in Bénin. *Lancet* 347:902–903.
  103. Takehisa J, Zekeng L, Ido E, Mboudjeka I, Moriyama H, Miura T, Yamashita M, Gürtler LG, Hayami M, Kaptué L. 1998. Various types of HIV mixed infections in Cameroon. *Virology* 245:1–10.
  104. Vergne L, Bourgeois A, Mpoudi-Ngole E, Mougnotou R, Mbuagbaw J, Liegeois F, Laurent C, Butel C, Zekeng L, Delaporte E, Peeters M. 2003. 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* 310:254–266.
  105. De Oliveira F, Mourez T, Vessiere A, Ngoupo P-A, Alessandri-Gradt E, Simon F, Rousset D, Plantier J-C. 2017. Multiple HIV-1/M + HIV-1/O dual infections and new HIV-1/MO inter-group recombinant forms detected in Cameroon. *Retrovirology* 14.
  106. Ngoupo PA. 2016. Fréquence et profil génétique des doubles infections VIH-1/M+O et formes recombinantes VIH-1/MO au Cameroun. These de doctorat. Rouen.
  107. Yamaguchi J, Bodelle P, Vallari AS, Coffey R, McArthur CP, Schochetman G, Devare SG, Brennan CA. 2004. 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* 20:944–957.
  108. Brand D, Beby-Defaux A, Macé M, Brunet S, Moreau A, Godet C, Jais X, Cazein F, Semaille C, Barin F. 2004. First identification of HIV-1 groups M and O dual infections in Europe. *AIDS* 18:2425–2428.
  109. Plantier J-C, Lemée V, Dorval I, Gueudin M, Braun J, Hutin P, Ruffault A, Simon F. 2004. HIV-1 group M superinfection in an HIV-1 group O-infected patient. *AIDS* 18:2444–2446.
  110. De Oliveira F, Cappy P, Lemée V, Moisan A, Pronier C, Bocket L, Bouvier-Alias M, Chaix M-L, Gault E, Morvan O, Poveda J-D, Schneider V, Wirden M, Alessandri-Gradt E, Mourez T, Plantier J-C. 2018. Detection of numerous HIV-1/MO recombinants in France. *AIDS* 32:1289–1299.
  111. Moisan A, De Oliveira F, Pronier C, Cappy P, Maillard A, Plantier J-C. 2020. *In vivo* emergence of an HIV-1/MO recombinant revealed undiagnosed HIV-1/M+O co-infection. *Clinical Microbiology and Infection* 26:262–264.
  112. Coffin JM. 1979. Structure, Replication, and Recombination of Retrovirus Genomes: Some Unifying Hypotheses. *Journal of General Virology* 42:1–26.

113. Zhuang J, Jetzt AE, Sun G, Yu H, Klarmann G, Ron Y, Preston BD, Dougherty JP. 2002. Human immunodeficiency virus type 1 recombination: rate, fidelity, and putative hot spots. *J Virol* 76:11273–11282.
114. DeStefano JJ, Bambara RA, Fay PJ. 1994. The mechanism of human immunodeficiency virus reverse transcriptase-catalyzed strand transfer from internal regions of heteropolymeric RNA templates. *J Biol Chem* 269:161–168.
115. DeStefano JJ, Mallaber LM, Rodriguez-Rodriguez L, Fay PJ, Bambara RA. 1992. Requirements for strand transfer between internal regions of heteropolymer templates by human immunodeficiency virus reverse transcriptase. *J Virol* 66:6370–6378.
116. Hwang CK, Svarovskaia ES, Pathak VK. 2001. Dynamic copy choice: Steady state between murine leukemia virus polymerase and polymerase-dependent RNase H activity determines frequency of *in vivo* template switching. *Proc Natl Acad Sci U S A* 98:12209–12214.
117. Pathak VK, Hu W-S. 1997. “Might as Well Jump!” Template Switching by Retroviral Reverse Transcriptase, Defective Genome Formation, and Recombination. *Seminars in Virology* 8:141–150.
118. Jetzt AE, Yu H, Klarmann GJ, Ron Y, Preston BD, Dougherty JP. 2000. High rate of recombination throughout the human immunodeficiency virus type 1 genome. *J Virol* 74:1234–1240.
119. Batorsky R, Kearney MF, Palmer SE, Maldarelli F, Rouzine IM, Coffin JM. 2011. Estimate of effective recombination rate and average selection coefficient for HIV in chronic infection. *Proceedings of the National Academy of Sciences* 108:5661–5666.
120. Clavel F, Hoggan MD, Willey RL, Strebel K, Martin MA, Repaske R. 1989. Genetic recombination of human immunodeficiency virus. *Journal of Virology* 63:1455–1459.
121. Hu WS, Temin HM. 1990. Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. *Proceedings of the National Academy of Sciences* 87:1556–1560.
122. Archer J, Pinney JW, Fan J, Simon-Loriere E, Arts EJ, Negroni M, Robertson DL. 2008. Identifying the important HIV-1 recombination breakpoints. *PLoS Comput Biol* 4:e1000178.
123. Smyth RP, Schlub TE, Grimm AJ, Waugh C, Ellenberg P, Chopra A, Mallal S, Cromer D, Mak J, Davenport MP. 2014. Identifying Recombination Hot Spots in the HIV-1 Genome. *J Virol* 88:2891–2902.
124. Simon-Loriere E, Galetto R, Hamoudi M, Archer J, Lefeuvre P, Martin DP, Robertson DL, Negroni M. 2009. Molecular mechanisms of recombination restriction in the envelope gene of the human immunodeficiency virus. *PLoS Pathog* 5:e1000418.
125. Jia L, Li L, Gui T, Liu S, Li H, Han J, Guo W, Liu Y, Li J. 2016. 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* 13:156.
126. Baird HA, Gao Y, Galetto R, Lalonde M, Anthony RM, Giacomoni V, Abreha M, Destefano JJ, Negroni M, Arts EJ. 2006. Influence of sequence identity and unique breakpoints on the frequency of intersubtype HIV-1 recombination. *Retrovirology* 3:91.
127. Motomura K, Chen J, Hu W-S. 2008. Genetic recombination between human immunodeficiency virus type 1 (HIV-1) and HIV-2, two distinct human lentiviruses. *J Virol* 82:1923–1933.
128. Zhang J, Temin HM. 1994. Retrovirus recombination depends on the length of sequence identity and is not error prone. *J Virol* 68:2409–2414.

129. Westesson O, Holmes I. 2009. Accurate Detection of Recombinant Breakpoints in Whole-Genome Alignments. *PLoS Comput Biol* 5:e1000318.
130. Simon-Loriere E, Holmes EC. 2011. Why do RNA viruses recombine? *Nat Rev Microbiol* 9:617–626.
131. Galetto R, Moumen A, Giacomoni V, Véron M, Charneau P, Negroni M. 2004. The structure of HIV-1 genomic RNA in the gp120 gene determines a recombination hot spot <in vivo>. *J Biol Chem* 279:36625–36632.
132. Peeters M, Liegeois F, Torimiro N, Bourgeois A, Mpoudi E, Vergne L, Saman E, Delaporte E, Saragosti S. 1999. Characterization of a Highly Replicative Intergroup M/O Human Immunodeficiency Virus Type 1 Recombinant Isolated from a Cameroonian Patient. 9. *J Virol* 73:7368–7375.
133. Takehisa J, Zekeng L, Ido E, Yamaguchi-Kabata Y, Mboudjeka I, Harada Y, Miura T, Kaptué L, Hayami M. 1999. Human Immunodeficiency Virus Type 1 Intergroup (M/O) Recombination in Cameroon. 8. *J Virol* 73:6810–6820.
134. Vessièrè A. 2009. Caractérisation virologique des doubles infections et des formes recombinantes par les VIH-1 du groupe M et du groupe O au Cameroun : conséquences épidémiologiques, diagnostics et thérapeutiques. These de doctorat. Paris 5.
135. Vessièrè A, Leoz M, Brodard V, Strady C, Lemée V, Depatureaux A, Simon F, Plantier J-C. 2010. First evidence of a HIV-1 M/O recombinant form circulating outside Cameroon. *AIDS* 24:1079.
136. Ngoupo PA, Sadeuh-Mba SA, De Oliveira F, Ngono V, Ngono L, Tchendjou P, Penlap V, Mourez T, Njouom R, Kfutwah A, Plantier J-C. 2016. First evidence of transmission of an HIV-1 M/O intergroup recombinant virus. *AIDS* 30:1–8.
137. Moisan A. 2019. Formes recombinantes inter-groupes M et O du VIH-1 : Etude de leur potentiel réplcatif et de leur émergence *in vitro/in vivo*. phdthesis. Normandie Université.
138. Moisan A. 2016. 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. These de doctorat.
139. Brennan CA, Bodelle P, Coffey R, Devare SG, Golden A, Hackett J, Harris B, Holzmayer V, Luk K-C, Schochetman G, Swanson P, Yamaguchi J, Vallari A, Ndembi N, Ngansop C, Makamche F, Mbanya D, Gürtler LG, Zekeng L, Kaptué L. 2008. The prevalence of diverse HIV-1 strains was stable in Cameroonian blood donors from 1996 to 2004. *J Acquir Immune Defic Syndr* 49:432–439.
140. Tongo M, Martin DP, Zembe L, Mpoudi-Ngole E, Williamson C, Burgers WA. 2013. Characterization of HIV-1 gag and nef in Cameroon: further evidence of extreme diversity at the origin of the HIV-1 group M epidemic. 1. *Virol J* 10:29.
141. Cappy P, Moisan A, De Oliveira F, Plantier J-C, Negroni M. 2017. HIV-1 sequences in the epidemic suggest an alternative pathway for the generation of the Long Terminal Repeats. *Sci Rep* 7:13715.
142. Fan J, Negroni M, Robertson DL. 2007. The distribution of HIV-1 recombination breakpoints. *Infect Genet Evol* 7:717–723.
143. Suzuki M, Christians FC, Kim B, Skandalis A, Black ME, Loeb LA. 1996. Tolerance of different proteins for amino acid diversity. *Mol Divers* 2:111–118.
144. Wargo AR, Kurath G. 2012. Viral fitness: definitions, measurement, and current insights. *Curr Opin Virol* 2:538–545.



145. Song H, Giorgi EE, Ganusov VV, Cai F, Athreya G, Yoon H, Carja O, Hora B, Hraber P, Romero-Severson E, Jiang C, Li X, Wang S, Li H, Salazar-Gonzalez JF, Salazar MG, Goonetilleke N, Keele BF, Montefiori DC, Cohen MS, Shaw GM, Hahn BH, McMichael AJ, Haynes BF, Korber B, Bhattacharya T, Gao F. 2018. Tracking HIV-1 recombination to resolve its contribution to HIV-1 evolution in natural infection. *Nat Commun* 9:1928.
146. Drummond DA, Silberg JJ, Meyer MM, Wilke CO, Arnold FH. 2005. On the conservative nature of intragenic recombination. *Proceedings of the National Academy of Sciences* 102:5380–5385.
147. Voigt CA, Martinez C, Wang Z-G, Mayo SL, Arnold FH. 2002. Protein building blocks preserved by recombination. *Nat Struct Mol Biol* 9:553–558.
148. Ritchie AJ, Cai F, Smith NMG, Chen S, Song H, Brackenridge S, Abdool Karim SS, Korber BT, McMichael AJ, Gao F, Goonetilleke N. 2014. Recombination-mediated escape from primary CD8+ T cells in acute HIV-1 infection. *Retrovirology* 11:69.
149. Streeck H, Li B, Poon AFY, Schneidewind A, Gladden AD, Power KA, Daskalakis D, Bazner S, Zuniga R, Brander C, Rosenberg ES, Frost SDW, Altfeld M, Allen TM. 2008. Immune-driven recombination and loss of control after HIV superinfection. *J Exp Med* 205:1789–1796.
150. Smyth RP, Davenport MP, Mak J. 2012. The origin of genetic diversity in HIV-1. *Virus Res* 169:415–429.
151. Iwabu Y, Mizuta H, Kawase M, Kameoka M, Goto T, Ikuta K. 2008. 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* 10:504–513.
152. Quan Y, Xu H, Wainberg MA. 2014. 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* 69:21–27.
153. Nora T, Charpentier C, Tenailon O, Hoede C, Clavel F, Hance AJ. 2007. Contribution of Recombination to the Evolution of Human Immunodeficiency Viruses Expressing Resistance to Antiretroviral Treatment. *J Virol* 81:7620–7628.
154. Moutouh L, Corbeil J, Richman DD. 1996. Recombination leads to the rapid emergence of HIV-1 dually resistant mutants under selective drug pressure. *Proc Natl Acad Sci U S A* 93:6106–6111.
155. Alessandri-Gradt E, Moisan A, Plantier J-C. 2023. HIV-1 Non-Group M Strains and ART. *Viruses* 15:780.
156. Domingo E, de Ávila AI, Gallego I, Sheldon J, Perales C. 2019. Viral fitness: history and relevance for viral pathogenesis and antiviral interventions. *Pathogens and Disease* 77:ftz021.
157. Domingo E, Holland JJ. 1997. RNA virus mutations and fitness for survival. *Annu Rev Microbiol* 51:151–178.
158. Daar ES. 2005. HIV-1 virulence, fitness and replication capacity. *Therapy* 2:131–140.
159. Novella IS, Duarte EA, Elena SF, Moya A, Domingo E, Holland JJ. 1995. Exponential increases of RNA virus fitness during large population transmissions. *Proc Natl Acad Sci U S A* 92:5841–5844.
160. Yuste E, Sánchez-Palomino S, Casado C, Domingo E, López-Galíndez C. 1999. Drastic fitness loss in human immunodeficiency virus type 1 upon serial bottleneck events. *J Virol* 73:2745–2751.
161. Bordería AV, Lorenzo-Redondo R, Pernas M, Casado C, Alvaro T, Domingo E, Lopez-Galindez C. 2010. Initial fitness recovery of HIV-1 is associated with quasispecies heterogeneity and can occur without modifications in the consensus sequence. *PLoS One* 5:e10319.

162. Kong X, West JT, Zhang H, Shea DM, M'soka TJ, Wood C. 2008. The human immunodeficiency virus type 1 envelope confers higher rates of replicative fitness to perinatally transmitted viruses than to nontransmitted viruses. *J Virol* 82:11609–11618.
163. Woo J, Robertson DL, Lovell SC. 2014. Constraints from protein structure and intra-molecular coevolution influence the fitness of HIV-1 recombinants. *Virology* 454–455:34–39.
164. Bagaya BS, Vega JF, Tian M, Nickel GC, Li Y, Krebs KC, Arts EJ, Gao Y. 2015. Functional bottlenecks for generation of HIV-1 intersubtype Env recombinants. *Retrovirology* 12:44.
165. Maldonado JO, Mansky LM. 2018. The HIV-1 Reverse Transcriptase A62V Mutation Influences Replication Fidelity and Viral Fitness in the Context of Multi-Drug-Resistant Mutations. *Viruses* 10:376.
166. Martinez-Picado J, Martínez MA. 2008. HIV-1 reverse transcriptase inhibitor resistance mutations and fitness: A view from the clinic and ex vivo. *Virus Research* 134:104–123.
167. Cappy P. 2016. Profil de recombinaison, genèse et étude fonctionnelle des recombinants VIH-1/MO : à la frontière de nouvelles espèces de VIH. These de doctorat. Rouen.
168. Ariën KK, Abraha A, Quiñones-Mateu ME, Kestens L, Vanham G, Arts EJ. 2005. The Replicative Fitness of Primary Human Immunodeficiency Virus Type 1 (HIV-1) Group M, HIV-1 Group O, and HIV-2 Isolates. 14. *J Virol* 79:8979–8990.
169. Gilbert PB, McKeague IW, Eisen G, Mullins C, Guéye-NDiaye A, Mboup S, Kanki PJ. 2003. Comparison of HIV-1 and HIV-2 infectivity from a prospective cohort study in Senegal. *Stat Med* 22:573–593.
170. Tebit DM, Lobritz M, Lalonde M, Immonen T, Singh K, Sarafianos S, Herchenröder O, Kräusslich H-G, Arts EJ. 2010. Divergent evolution in reverse transcriptase (RT) of HIV-1 group O and M lineages: impact on structure, fitness, and sensitivity to nonnucleoside RT inhibitors. *J Virol* 84:9817–9830.
171. Mack K, Starz K, Sauter D, Langer S, Bibollet-Ruche F, Learn GH, Stürzel CM, Leoz M, Plantier J-C, Geyer M, Hahn BH, Kirchhoff F. 2017. Efficient Vpu-Mediated Tetherin Antagonism by an HIV-1 Group O Strain. *J Virol* 91:e02177-16.
172. Bolivar F, Rodriguez RL, Greene PJ, Betlach MC, Heyneker HL, Boyer HW, Crosa JH, Falkow S. 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* 2:95–113.
173. Watson N. 1988. A new revision of the sequence of plasmid pBR322. *Gene* 70:399–403.
174. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6:343–345.
175. Gibson DG. 2011. Enzymatic Assembly of Overlapping DNA Fragments, p. 349–361. *In* *Methods in Enzymology*. Elsevier.
176. 2023. Gibson assembly. Wikipedia.
177. NEB® Stable Competent E. coli (High Efficiency) | NEB. <https://international.neb.com/products/c3040-neb-stable-competent-e-coli-high-efficiency#Product%20Information>. Retrieved 30 May 2023.
178. Montefiori DC. 2004. Evaluating Neutralizing Antibodies Against HIV, SIV, and SHIV in Luciferase Reporter Gene Assays. *Current Protocols in Immunology* 64.
179. Platt EJ, Wehrly K, Kuhmann SE, Chesebro B, Kabat D. 1998. Effects of CCR5 and CD4 cell surface concentrations on infections by macrophagetropic isolates of human immunodeficiency virus type 1. *J Virol* 72:2855–2864.

180. Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74:5463–5467.
181. Wu S, Yan Y, Yan P, Huang H, Wang H. 2010. Construction and characterization of a full-length infectious clone from a fast-replicating, X4-tropic HIV-1 subtype B' isolate. *Arch Virol* 155:1923–1931.
182. Rodriguez MA, Chen Y, Craig JK, Chatterjee R, Ratner D, Tatsumi M, Roy P, Neogi D, Gupta P. 2006. Construction and characterization of an infectious molecular clone of HIV-1 subtype A of Indian origin. *Virology* 345:328–336.
183. Baalwa J, Wang S, Parrish N, Decker JM, Keele BF, Learn GH, Yue L, Ruzagira E, Ssemwanga D, Kamali A, Amornkul PN, Price MA, Kappes JC, Karita E, Kaleebu P, Sanders E, Gilmour J, Allen S, Hunter E, Montefiori DC, Haynes BF, Cormier E, Hahn BH, Shaw GM. 2013. Molecular identification, cloning and characterization of transmitted/founder HIV-1 subtype A, D and A/D infectious molecular clones. *Virology* 436:33–48.
184. NIBSC - AIDS Reagents catalogue. [https://www.nibsc.org/science\\_and\\_research/idd/cfar/reagents\\_catalogue.aspx](https://www.nibsc.org/science_and_research/idd/cfar/reagents_catalogue.aspx). Retrieved 17 June 2023.
185. Tebit DM, Zekeng L, Kaptué L, Gürtler L, Fackler OT, Keppler OT, Herchenröder O, Kräusslich H-G. 2004. Construction and characterization of an HIV-1 group O infectious molecular clone and analysis of vpr- and nef-negative derivatives. *Virology* 326:329–339.
186. Adachi A, Gendelman HE, Koenig S, Folks T, Willey R, Rabson A, Martin MA. 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J Virol* 59:284–291.
187. Grisson RD, Chenine A-L, Yeh L-Y, He J, Wood C, Bhat GJ, Xu W, Kankasa C, Ruprecht RM. 2004. Infectious molecular clone of a recently transmitted pediatric human immunodeficiency virus clade C isolate from Africa: evidence of intraclade recombination. *J Virol* 78:14066–14069.
188. Deymier MJ, Claiborne DT, Ende Z, Ratner HK, Kilembe W, Allen S, Hunter E. 2014. Particle infectivity of HIV-1 full-length genome infectious molecular clones in a subtype C heterosexual transmission pair following high fidelity amplification and unbiased cloning. *Virology* 468–470:454–461.
189. Zyryanova DP, Totmenin AV, Bogacheva NV, Gashnikova NM. 2020. Construction and Characterization of Infectious Molecular Clones of HIV-1 CRF63\_02A6. *AIDS Res Hum Retroviruses* 36:227–233.
190. Pizzato M, Erlwein O, Bonsall D, Kaye S, Muir D, McClure MO. 2009. A one-step SYBR Green I-based product-enhanced reverse transcriptase assay for the quantitation of retroviruses in cell culture supernatants. *J Virol Methods* 156:1–7.
191. Vermeire J, Naessens E, Vanderstraeten H, Landi A, Iannucci V, Van Nuffel A, Taghon T, Pizzato M, Verhasselt B. 2012. Quantification of reverse transcriptase activity by real-time PCR as a fast and accurate method for titration of HIV, lenti- and retroviral vectors. *PLoS One* 7:e50859.
192. Meuwissen PJ, Stolp B, Iannucci V, Vermeire J, Naessens E, Saksela K, Geyer M, Vanham G, Arien KK, Fackler OT, Verhasselt B. 2012. Identification of a highly conserved valine-glycine-phenylalanine amino acid triplet required for HIV-1 Nef function. *Retrovirology* 9:34.
193. Vermeire J, Roesch F, Sauter D, Rua R, Hotter D, Van Nuffel A, Vanderstraeten H, Naessens E, Iannucci V, Landi A, Witkowski W, Baeyens A, Kirchhoff F, Verhasselt B. 2016. HIV Triggers a cGAS-Dependent, Vpu- and Vpr-Regulated Type I Interferon Response in CD4+ T Cells. *Cell Reports* 17:413–424.

194. Gorry PR, Ancuta P. 2011. Coreceptors and HIV-1 pathogenesis. *Curr HIV/AIDS Rep* 8:45–53.
195. Richman DD, Bozzette SA. 1994. The impact of the syncytium-inducing phenotype of human immunodeficiency virus on disease progression. *J Infect Dis* 169:968–974.
196. Locher CP, Witt SA, Kassel R, Dowell NL, Fujimura S, Levy JA. 2005. Differential effects of R5 and X4 human immunodeficiency virus type 1 infection on CD4+ cell proliferation and activation. *Journal of General Virology* 86:1171–1179.
197. Weinberger AD, Perelson AS. 2011. Persistence and Emergence of X4 Virus in HIV Infection. *Math Biosci Eng* 8:605–626.
198. Rai M, Timilsina U, Gaur R. 2022. Construction and characterization of a full-length, replication-competent and infectious enhanced green fluorescence protein-tagged HIV-1 subtype C molecular clone. *Virology* 571:34–38.
199. Rouzioux C, Costagliola D, Burgard M, Blanche S, Mayaux MJ, Griscelli C, Valleron AJ. 1995. Estimated timing of mother-to-child human immunodeficiency virus type 1 (HIV-1) transmission by use of a Markov model. The HIV Infection in Newborns French Collaborative Study Group. *Am J Epidemiol* 142:1330–1337.
200. Burgard M, Jasseron C, Matheron S, Damond F, Hamrene K, Blanche S, Faye A, Rouzioux C, Warszawski J, Mandelbro L, ANRS French Perinatal Cohort EPF-CO1. 2010. Mother-to-child transmission of HIV-2 infection from 1986 to 2007 in the ANRS French Perinatal Cohort EPF-CO1. *Clin Infect Dis* 51:833–843.
201. Ter Schiphorst E, Hansen KC, Holm M, Hønge BL. 2022. Mother-to-child HIV-2 transmission: comparison with HIV-1 and evaluation of factors influencing the rate of transmission. A systematic review. *Trans R Soc Trop Med Hyg* 116:399–408.
202. O'Donovan D, Ariyoshi K, Milligan P, Ota M, Yamuah L, Sarge-Njie R, Whittle H. 2000. Maternal plasma viral RNA levels determine marked differences in mother-to-child transmission rates of HIV-1 and HIV-2 in The Gambia. MRC/Gambia Government/University College London Medical School working group on mother-child transmission of HIV. *AIDS* 14:441–448.
203. Gueudin M, Lemée V, Ferre V, Beby-Defaux A, Pathé J-P, Guist'hau O, Braun J, Simon F, Plantier J-C. 2004. Virologic Diagnosis and Follow-up of Children Born to Mothers Infected by HIV-1 Group O. *JAIDS Journal of Acquired Immune Deficiency Syndromes* 36:639.
204. Chaix-Baudier M-L, Chappey C, Burgard M, Letourneur F, Igual J, Saragosti S, Rouzioux C. 1998. First Case of Mother-to-Infant HIV Type 1 Group O Transmission and Evolution of C2V3 Sequences in the Infected Child. *AIDS Research and Human Retroviruses* 14:15–23.
205. Mouafo LCM, Dambaya B, Ngoufack NN, Nkenfou CN. 2017. Host Molecular Factors and Viral Genotypes in the Mother-to-Child HIV-1 Transmission in Sub-Saharan Africa. *J Public Health Afr* 8:594.
206. Koulinska IN, Villamor E, Msamanga G, Fawzi W, Blackard J, Renjifo B, Essex M. 2006. Risk of HIV-1 transmission by breastfeeding among mothers infected with recombinant and non-recombinant HIV-1 genotypes. *Virus Res* 120:191–198.
207. Yang C, Li M, Newman RD, Shi Y-P, Ayisi J, van Eijk AM, Otieno J, Misore AO, Steketee RW, Nahlen BL, Lal RB. 2003. Genetic diversity of HIV-1 in western Kenya: subtype-specific differences in mother-to-child transmission. *AIDS* 17:1667.
208. Renjifo B, Gilbert P, Chaplin B, Msamanga G, Mwakagile D, Fawzi W, Essex M, Group and TV and HS. 2004. Preferential in-utero transmission of HIV-1 subtype C as compared to HIV-1 subtype A or D. *AIDS* 18:1629.

209. Krivine A, Launay O, Firtion G, Finkielsztejn L, Jullien V, Rozenberg F, Pannier E. 2009. Does HIV-1 Subtype D Have a Higher Risk of Vertical Transmission than Other HIV Subtypes? *The Journal of Infectious Diseases* 199:1553–1554.
210. Murray MC, Embree JE, Ramdahin SG, Anzala AO, Njenga S, Plummer FA. 2000. Effect of human immunodeficiency virus (HIV) type 1 viral genotype on mother-to-child transmission of HIV-1. *J Infect Dis* 181:746–749.
211. Bulterys PL, Dalai SC, Katzenstein DA. 2010. Viral Sequence Analysis from HIV-Infected Mothers and Infants: Molecular Evolution, Diversity, and Risk Factors for Mother-To-Child Transmission. *Clin Perinatol* 37:739–viii.
212. HIV-1 genotypic drug resistance interpretation's algorithms. <https://hivfrenchresistance.org/>. Retrieved 12 April 2023.
213. Vessièrè A, Rousset D, Kfutwah A, Leoz M, Depatureaux A, Simon F, Plantier J-C. 2010. Diagnosis and Monitoring of HIV-1 Group O-Infected Patients in Cameroun. *JAIDS Journal of Acquired Immune Deficiency Syndromes* 53:107.
214. Heyndrickx L, Janssens W, Gürtler L, Zekeng L, Loussert-Ajaka I, Vereecken K, Willems B, Coppens S, Ndumbe P, Fransen K, Saman E, Alary M, van der Groen G. 1998. 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* 14:973–977.
215. DeepChek® Assay Whole Genome HIV-1 Genotyping - ABL Diagnostics S.A. <https://www.abldiagnostics.com/products/deepchek-assay-whole-genome-hiv-1-genotyping/>. Retrieved 27 March 2023.
216. Illumina DNA Prep | Flexibility for many whole-genome sequencing applications. <https://www.illumina.com/products/by-type/sequencing-kits/library-prep-kits/illumina-dna-prep.html>. Retrieved 12 June 2023.
217. ATDBio - Next generation sequencing. <https://atdbio.com/nucleic-acids-book/Next-generation-sequencing#Bridge-PCR>. Retrieved 14 May 2023.
218. Sequencing Analysis Viewer v2.4.7 Software. <https://support.illumina.com/downloads/sequencing-analysis-viewer-software-v2-4-7.html>. Retrieved 21 June 2023.
219. Bendall ML, Gibson KM, Steiner MC, Rentia U, Pérez-Losada M, Crandall KA. 2020. HAPHPIPE: Haplotype Reconstruction and Phylodynamics for Deep Sequencing of Intra-host Viral Populations. *Mol Biol Evol* 38:1677–1690.
220. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, Sun C, Grayson T, Wang S, Li H, Wei X, Jiang C, Kirchherr JL, Gao F, Anderson JA, Ping L-H, Swanstrom R, Tomaras GD, Blattner WA, Goepfert PA, Kilby JM, Saag MS, Delwart EL, Busch MP, Cohen MS, Montefiori DC, Haynes BF, Gaschen B, Athreya GS, Lee HY, Wood N, Seoighe C, Perelson AS, Bhattacharya T, Korber BT, Hahn BH, Shaw GM. 2008. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proceedings of the National Academy of Sciences* 105:7552–7557.
221. BLAST search. [https://www.hiv.lanl.gov/content/sequence/BASIC\\_BLAST/basic\\_blast.html](https://www.hiv.lanl.gov/content/sequence/BASIC_BLAST/basic_blast.html). Retrieved 31 January 2023.
222. Genotyping. <https://www.ncbi.nlm.nih.gov/projects/genotyping/formpage.cgi>. Retrieved 15 March 2023.
223. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution* 33:1870–1874.

224. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, Ray SC. 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* 73:152–160.
225. Amin O, Powers J, Bricker KM, Chahroudi A. 2021. Understanding Viral and Immune Interplay During Vertical Transmission of HIV: Implications for Cure. *Front Immunol* 12:757400.
226. Parikh UM, McCormick K, van Zyl G, Mellors JW. 2017. Future technologies for monitoring HIV drug resistance and cure. *Curr Opin HIV AIDS* 12:182–189.
227. Barzon L, Lavezzo E, Costanzi G, Franchin E, Toppo S, Palù G. 2013. Next-generation sequencing technologies in diagnostic virology. *J Clin Virol* 58:346–350.
228. Van Laethem K, Theys K, Vandamme A-M. 2015. HIV-1 genotypic drug resistance testing: digging deep, reaching wide? *Curr Opin Virol* 14:16–23.
229. Casadellà M, Paredes R. 2017. Deep sequencing for HIV-1 clinical management. *Virus Res* 239:69–81.
230. Gall A, Ferns B, Morris C, Watson S, Cotten M, Robinson M, Berry N, Pillay D, Kellam P. 2012. Universal amplification, next-generation sequencing, and assembly of HIV-1 genomes. *J Clin Microbiol* 50:3838–3844.
231. Grossmann S, Nowak P, Neogi U. 2015. Subtype-independent near full-length HIV-1 genome sequencing and assembly to be used in large molecular epidemiological studies and clinical management. *J Int AIDS Soc* 18:20035.
232. Berg MG, Yamaguchi J, Alessandri-Gradt E, Tell RW, Plantier J-C, Brennan CA. 2016. A Pan-HIV Strategy for Complete Genome Sequencing. *J Clin Microbiol* 54:868–882.
233. Yamaguchi J, Olivo A, Laeyendecker O, Forberg K, Ndembi N, Mbanya D, Kaptue L, Quinn TC, Cloherty GA, Rodgers MA, Berg MG. 2018. Universal Target Capture of HIV Sequences From NGS Libraries. *Front Microbiol* 9:2150.
234. Tzou PL, Ariyaratne P, Varghese V, Lee C, Rakhmanaliev E, Villy C, Yee M, Tan K, Michel G, Pinsky BA, Shafer RW. 2018. Comparison of an In Vitro Diagnostic Next-Generation Sequencing Assay with Sanger Sequencing for HIV-1 Genotypic Resistance Testing. *Journal of Clinical Microbiology* 56:e00105-18.
235. Parkin NT, Avila-Rios S, Bibby DF, Brumme CJ, Eshleman SH, Harrigan PR, Howison M, Hunt G, Ji H, Kantor R, Ledwaba J, Lee ER, Matías-Florentino M, Mbisa JL, Noguera-Julian M, Paredes R, Rivera-Amill V, Swanstrom R, Zaccaro DJ, Zhang Y, Zhou S, Jennings C. 2020. Multi-Laboratory Comparison of Next-Generation to Sanger-Based Sequencing for HIV-1 Drug Resistance Genotyping. *Viruses* 12:694.
236. Ávila-Ríos S, Parkin N, Swanstrom R, Paredes R, Shafer R, Ji H, Kantor R. 2020. Next-Generation Sequencing for HIV Drug Resistance Testing: Laboratory, Clinical, and Implementation Considerations. *Viruses* 12:617.
237. Teo CHY, Norhisham NHB, Lee OF, Png S, Chai CN, Yan G, Tang JW-T, Lee CK. 2022. Towards Next-Generation Sequencing for HIV-1 Drug Resistance Testing in a Clinical Setting. *Viruses* 14:2208.
238. De Luca A. 2006. The impact of resistance on viral fitness and its clinical implications, p. *In* Geretti, AM (ed.), *Antiretroviral Resistance in Clinical Practice*. Mediscript, London.
239. Dykes C, Demeter LM. 2007. Clinical Significance of Human Immunodeficiency Virus Type 1 Replication Fitness. *Clinical Microbiology Reviews* 20:550–578.



## Annexe 1 : Amorces ayant servi à la génération des CMI

Souche VIH-1 pour CMI	Diversité	Fragment	Amorce Forward (Upper)	Amorce (5' -> 3')	Amorce Reverse (Lower)	Amorce (5' -> 3')	Taille
YBF206	O-H	A	Rec020U <b>523U</b>	GCGACCACACCCGTCCTGTGCTGGAAGGGTTAATTTACTCC TTGGGTGTATCACACTCAGG	Rec021L <b>285L</b>	TGTTTTCCATAAATGTCTGATTATTTTGCCTTTCTTC CTGTCCACTTGCCATACTATCAG	5 126 pb 5093 pb
		B	<b>Rec022U</b>	ATCAGACATTATGGAAAACAGATGGCAGG	<b>Rec023L</b>	ACGATGCGTCCGGCGTAGAGGATCTCTAGATACCAGGGTTGAAC	4 786 pb
RBF130	O-H	A	<b>Rec024U</b>	CACACCCGTCCTGTGGGATGGGTTAATTTACTCCCTGAGA	Rec025L	TGTTTTCCATAAATTTCTGATTATTTTGCCTTTCTTC	5 119 pb
					284L	ACCTGCCATCTGTTTTCCATA	5 129 pb
				280L	CCYTCACCTTTCCACAGGAG	4996 pb	
				<b>282L</b>	CCYTCACCTTTCCACAGGAG	5038 pb	
		B	Rec026U 297U <b>281U</b>	ATCAGAAAATTATGGAAAACAGATGGCAG AGAGAYCCTATYTGAAAGGAC	<b>Rec023L</b>	ACGATGCGTCCGGCGTAGAGGATCTCTAGATACCAGGGTTGAAC	4 788 pb 4811 pb 4899 pb
BCF108	O-H	A	<b>Rec027U</b>	CACACCCGTCCTGTGTTGGAAGGGCTAATTTACTCC	Rec028L <b>285L</b>	TGTTTTCCATAAATGTCTAATTATTTTGCCTTCC CTGTCCACTTGCCATACTATCAG	5 154 pb 5120 pb
		B	<b>Rec029U</b>	ATTAGACATTATGGAAAACAGATGGCAG	<b>Rec030L</b>	ACGATGCGTCCGGCGTAGAGACTCAAGCAAGCTTTATTG	4 656 pb
AAS638	M-CRF02_AG	A	<b>Rec042U</b>	CACACCCGTCCTGTGTTGGAAGGGTTAATTTGGTCC	<b>Rec013L</b>	ATCTGTTTTCCATAATGCCTAATGATTTTAC	5078 pb
		B	<b>Rec014U</b>	AAATCATTARRCATTATGGAAAACAGATGGC	<b>Rec043L</b>	ATGCGTCCGGCGTAGAGGCTAGAGATTTTACACAGTCTC	4723 pb
M-A-062021	M-A	A	<b>Rec031U</b>	CACACCCGTCCTGTGTTGGAAGGGTTAATTTACTCC	Rec032L <b>284L</b>	TGTTTTCCATAAATCTAATGATTTTAGCCTTTC ACCTGCCATCTGTTTTCCATA	5 077 pb 5067 pb
		B	<b>Rec033U</b>	ATTAGAGATTATGGAAAACAGATGGCAG	<b>Rec034L</b>	ACGATGCGTCCGGCGTAGAGTGCTAGAGATTTTACAAAGTCTAGAGTG	4 774 pb
M-F-062021	M-F	A	<b>Rec035U</b>	ACACCCGTCCTGTGTTGGAAGGGTTAATTTACTCC	<b>Rec036L</b>	TGTTTTCCATAAATCCCTAATGATCTTTGCTTTTC	5 103 pb
		B	<b>Rec037U</b>	ATTAGGGATTATGGAAAACAATGGCAG	<b>Rec038L</b>	GCGTCCGGCGTAGAGTATTGAAGTACTCAAGGCAAG	4 680 pb

U : Upper (amorces sens)

L : Lower (amorces antisens)

Plusieurs tentatives de génération des fragments ont été réalisées pour certains CMI (YBF026, RBF130, BCF108 et M-A-062021). Les amorces représentées en gras ont permis l'obtention d'une bande d'amplification et ont servi à la génération des derniers assemblages GA testés au laboratoire.



## **Annexe 2 : Protocole de transformation chimique des bactéries NEB Stable®**

Après décongélation d'un aliquot de 20 µL bactéries dans de la glace, 1 à 2µL de CMI, soit environ 100 pg à 100 ng selon les recommandations du fournisseur, ont été ajoutés et le mélange a été incubé dans de la glace pendant 30min. Pendant ce temps, des témoins positif (3µL de bactéries sur géloses LB sans antibiotiques) et négatif (3µL de bactéries non transformées sur géloses LB + Amp 100) ont été réalisés pour chaque test de clonage et ensemencés sur géloses LB (*Lysogeny Broth*), préchauffées à 30°C. Après incubation, un choc thermique a été réalisé à 42°C pendant 30 s exactement, puis le mélange a été incubé 5 min dans la glace, avant d'y ajouter 500µL du milieu NEB® Outgrowth medium préchauffé à 30°C. Les tubes ont été par la suite placés à 30°C pour 1h d'agitation en position inclinée, à 125 rpm. À la fin de l'incubation, 50µL et 100µL du mélange ont été ensemencés sur des géloses LB + Amp 100. Le reste du mélange a été centrifugé pendant 7min à 5000 rpm afin d'éliminer le surnageant et remettre en suspension les bactéries concentrées dans 50 à 100µL de NEB® Outgrowth medium. Les trois boîtes de pétri ensemencées par clone ont ensuite été mises à incuber à 30 °C pendant minimum 24h. Le lendemain, si un tapis était observé, un ré-isolement était effectué. Si non, quelques colonies étaient sélectionnées après vérification des boîtes témoins. Pour chaque souche testée, 12 colonies ont été ensemencées dans des bouillons LB + Amp 100 et mises à incuber à nouveau durant 24h à 30°C sous agitation à 125-150 rpm. Après 24h, 850µL de suspension bactérienne ont été congelés avec 150µL glycérol dans des cryotubes pour conservation du clone à -20°C. Par ailleurs, 5mL ont été centrifugés 3mn à 10 000g afin de procéder à la purification de l'ADN plasmidique. Le protocole de purification par miniprep a été réalisé avec le kit NucleoSpin Plasmid - *low copy* protocol (Macherey-Nagel) selon les recommandations fournisseurs.

### Annexe 3 : Amorces ayant servi à la génération des CMIC

CMIC	Fragment	Amorce (Upper)	Amorce (5' -> 3')	Amorce (Lower)	Amorce (5' -> 3')	Taille
MOMO	MA	Rec031U	CACACCCGTCCTGTGTGGATGGGTTAATTCACCTCC	Rec044L	GTTAAGCAGCGGGTTCCTAGCTAGCCAG	538 pb
	OB	Rec045U	TAGGGAACCCGCTGCTTAACGCTCAATAAAG	Rec046L	TTCTTCTAGTAGCTCTAGTGTCCATTCATTG	5152 pb
	MC	Rec047U	CACTAGAGCTACTAGAAGAAGCTTAAACATGAAGCTGTTAG	Rec048L	GTTAAGCAGCGGGTTCCTAGCTAGCCAG	4060 pb
	OD	Rec049U	TAGGGAACCCGCTGCTTAACGCTCAATAAAG	Rec050L	ACGATGCGTCCGGCGTAGAGTGCTAGAGATTTTTCTGCTTC	148 pb
OMOM	OA	Rec051U	CACACCCGTCCTGTGTGGATGGGTTAATTTACTCC	Rec052L	CTCAAGCAGTGGGTTAGCTAGAGGC	574 pb
	MB	Rec053U	AGCTGAACCCACTGCTTGAGCCTCAATAAAG	Rec054L	CTCTTCTAGGAGTTCTAACATCCATTCCTG	5132 pb
	OC	Rec055U	TGTTAGAAGCTCTAGAAGAGCTAAAAGCAGAAGC	Rec056L	CTCAAGCAGTGGGTTAGCTAGAGGC	4115 pb
	MD	Rec057U	AGCTGAACCCACTGCTTGAGCCTCAATAAAG	Rec034L	ACGATGCGTCCGGCGTAGAGTGCTAGAGATTTTTACAAAGCTAGAGTG	151 pb

U : Upper (amorces sens)

L : Lower (amorces antisens)

En bleu sont représentées les portions de l'amorces ciblant du VIH-1/O , en rouge du VIH-1/M et en vert le plasmide pBR322.

#### **Annexe 4 : Protocole de génération des fragments de CMIC**

Les fragments de petite taille (< 300 pb) ont été amplifiés à l'aide du kit HotStart Taq® Master Mix (Qiagen) selon le protocole suivant : 25µL de mix ont été complétés avec 2,5µL de chaque amorce sens et antisens à 10µM, ainsi que 1µL d'ADN et complétés avec de l'eau RNase free pour un volume réactionnel final de 50µL. Les échantillons ont été placés dans un thermocycleur 15min à 95°C, puis 35 cycles de 30s à 94°C, 30s à 58°C, 1min à 72°C et enfin une élongation finale de 10min à 72°C. Les grands fragments (> 3000pb) ont été obtenus après utilisation du kit Q5® Hot Start High-Fidelity 2X Master Mix (NEB) en complétant 25µL de mix par 2,5µL de chaque amorce sens et antisens à 10µM, 10ng d'ADN et par de l'eau RNase free pour un volume réactionnel final de 50µL. Une première étape de dénaturation a été entreprise à 98°C pendant 30s, suivie de 35 cycles de 10s à 98°C, 58°C pendant 30s, 72°C pendant 10min, puis une élongation finale à 72°C durant 2min.

#### **Annexe 5 : Protocole de génétique inverse (transfection / co-culture)**

La veille de la transfection (J-3), 750 000 HEK293T ont été réparties dans une plaque 6 puits dans un volume final de 2mL de milieu DMEM + 10% de SVFd + Gentamicine (50µg/mL), afin d'obtenir 60 à 80% de confluence le jour de la transfection (J-2). Cette transfection, a été réalisée en diluant 2µg de plasmide dans 200µL de tampon JetPRIME® (Polyplus Transfection, France) avant agitation et centrifugation. Ensuite, 4µL de réactif JetPRIME® (Polyplus Transfection) ont été ajoutés à ce mélange avant agitation et centrifugation également. L'ensemble a été incubé à température ambiante pendant 10min. En parallèle, le milieu des cellules HEK293T a été remplacé par du DMEM + 10% SVFd + Gentamicine (50µg/mL) frais. Après incubation, 250µL du mélange ont été déposés goutte à goutte sur la surface du puits correspondant. La plaque a ensuite été mise à incuber 72h à 37°C avec 5% de CO<sub>2</sub>.

Le jour de la co-culture (J0), 1mL de surnageant issu de la culture des cellules HEK293T a été récupéré dans chaque puits avant de déposer 1,5.10<sup>6</sup> Jurkat X4-R5 goutte à goutte sur le volume restant, toujours sans décoller le tapis de cellules adhérentes. Après ajout de RPMI pour un volume final de 2mL, la plaque a été mise à incuber 72h à 37°C avec 5% de CO<sub>2</sub> dans une atmosphère humide. Après 3 jours d'incubation (J3), la totalité du surnageant de chaque puits, sans les HEK293T, a été transférée dans une T25 (flasque ayant une surface de culture de 25 cm<sup>2</sup>) identifiée pour chaque candidat avec un volume final de 15mL. La culture a été entretenue sur une vingtaine de jours, deux fois par semaine, avec ajout de milieu RPMI+SVFd+Gentamicine+Généticine frais. Des prélèvements de surnageants ont été effectués également deux fois par semaine.

## **Annexe 6 : Protocole de séquençage Sanger**

Au fond d'une plaque 96 puits, 2µL de chaque amplicon préalablement purifié ont été mélangés à 8µL d'eau RNase free, 7µL de mix DCTS QuickStart (Beckman Coulter, Etats-Unis) et 3µL d'amorces spécifiques des régions à séquencer. Cette plaque a ensuite été placée dans un thermocycleur pour 30 cycles de 96°C pendant 25 secondes, 50°C pendant 25s puis 60°C pendant 4 min. L'ajout d'une solution contenant du NaOAc à une concentration de 3M, de l'eau RNase free, de l'EDTA 0,5M et du glycogène, ainsi que de l'éthanol à 95%, ont permis de purifier et précipiter les échantillons à froid. Les culots d'ADN obtenus ont été remis en suspension à l'aide d'un tampon de chargement, le SLS (Sample Loading Solution, Beckman Coulter). Les différentes données de séquençage ont été obtenues à l'aide du séquenceur Beckman Coulter CEQ8000™.

## **Annexe 7 : Protocole de SG-PERT du laboratoire de virologie du CHU de Rouen**

Un tampon de lyse 2X concentré a été réalisé au préalable avec 0,25% de Triton X-100 (MP Biomedicals, France), 50 mM de KCL (Merck, Allemagne), 1M Tris-HCL (Sigma-Aldrich) à un pH de 7,5 ainsi que du glycérol à 40% (Merk). Dans une première plaque 96 puits à fond en U (Beckton Dickinson, Belgique), 5µL de ce tampon ont été ajoutés et complétés par 0,1µL d'inhibiteur de RNase à la concentration de 40 U/µL (Ribolock™ Fermentas, Allemagne). A ce mélange, 5µL de surnageant à tester ont été ajoutés en répliqués. La solution a ensuite été mise à incubée 10 min à température ambiante après centrifugation, puis diluée par ajout de 90 µL d'eau RNase free avant nouvelle centrifugation puis conservation à +4°C. Dans une autre plaque 96 puits, spécifique au lecteur LC480 (LightCycler 480 plate 96 blanc, Roche Diagnostics, France), 10,4µL de mélange réactionnel ont été déposés pour chaque échantillon. Ce mélange contenait 10 µL de SYBR Green I Master Mix 2x (LightCycler 480 SYBR I Master mix, Roche Diagnostics), 0,1µL Inhibiteur RNase Ribolock™ à 40 U/µL (Fermentas, Allemagne) dilué dans de l'eau RNase free pour une concentration finale de 4 U/µL, 0,1µL de MS2 ARN, 0,1µL des amorces MS2 sens (5'-TCCTGCTCAACTTCCTGTCGAG-3') et antisens (5'-CAGGTCAAACCTCCTAGGAATG-3') à 100µM. Enfin, 9,6µL de lysat contenu dans la première plaque à fond en U, ont été déposés dans la plaque 96 spécifique du LC480 contenant les 10,4µL de mélange réactionnel selon le plan de plaque réalisé. Le volume réactionnel total de 20µL contenu dans chaque puits de la plaque à tester a ensuite été lancé dans un lecteur spécifique LC 480 afin de suivre l'activité de la TI en temps réel par qRT-PCR. La plaque a été mise à incuber 20 min à 42°C pour la réaction de transcription inverse, suivie de 5 min à 95°C pour l'activation de l'ADN polymérase (FastStart Taq, Sigma-Aldrich). La 3ème étape de cette réaction a consisté en 45 cycles d'amplification de 5 secondes à 95°C, 5 secondes à 60°C et 15 secondes à 72°C. Enfin, la plaque a été incubée à 95°C puis 37°C pour acquérir la mesure de fluorescence. L'interprétation des résultats se fait alors sur le logiciel de l'automate et permet de quantifier pour chaque surnageants, l'activité de la TI en mU/mL.

1 **Annexe 8 :**

2 **First evidence of a mother to child transmission of an HIV-1/MO intergroup**  
3 **recombinant form**

4 Manon Vautrin<sup>1\*</sup>, Fabienne De Oliveira<sup>2\*</sup>, Vinca Icard<sup>3</sup>, Marie Leoz<sup>1</sup>, Mary-Anne Trabaud<sup>3</sup> Antoine  
5 Ouziel<sup>4</sup>, Luc Panetta<sup>4</sup>, Véronique Lemée<sup>2</sup>, Jean Christophe Plantier<sup>2</sup>, Alice Moisan<sup>2\*\*</sup>

6  
7 <sup>1</sup>Univ Rouen Normandie, Univ de Caen, INSERM, DYNAMICURE UMR 1311, F-76000 Rouen, France

8 <sup>2</sup>Univ Rouen Normandie, Univ de Caen, INSERM, DYNAMICURE UMR 1311, CHU Rouen, Department of Virology,  
9 National Reference Center of HIV, F-76000 Rouen, France

10 <sup>3</sup>Virology department, Institut des Agents Infectieux, Hôpital de la Croix Rousse, Hospices Civils de Lyon, 69004  
11 Lyon, France

12 <sup>4</sup>Department of Pediatric Emergency & Pediatric Intensive Care, Hôpital Femme-Mère Enfant, Hospices Civils de  
13 Lyon, 69677 Bron cedex, France

14  
15 \*These authors contributed equally to this study

16 \*\*corresponding author: [alice.moisan@chu-rouen.fr](mailto:alice.moisan@chu-rouen.fr)

17

18 **Keywords:** Vertical transmission, intergroup MO recombination, URF\_MO

19

20 **Abstract:**

21 HIV-1 intergroup M and O are rare forms, for which little is known concerning virological properties  
22 and clinical impact, due to the atypical genetic profile. Here, we report the first case of mother to child  
23 transmission of such a form, which present high replicative profile and severe infections.

24

25 **Case:**

26 Since 1999, several studies, in Cameroun and in France, have demonstrated that HIV-1 inter-  
27 group M and O recombinant forms (HIV-1/MO) are existing, despite the strong divergence between  
28 these two groups [1–7]. Although their conditions of emergence remain unclear, multiple URF\_MO  
29 have been described in several contexts, including more recently, the description of the emergence *in*  
30 *vivo* of an HIV-1/MO in an HIV-1/M+O dually infected patient [1–8]. Among these contexts, a single  
31 study reported the direct transmission of an HIV-1/MO recombinant form, within a Cameroonian  
32 couple [5]. The circulation of HIV-1/MO recombinant forms is very low with only 19 URF\_MO reported  
33 in 23 patients by now. However, we report here a new URF\_MO in a very rare context of mother-to-  
34 child-transmission (MTCT).

35 In May 2022, a Cameroonian 4-month-old baby (RBF255) was admitted for pneumocystosis in  
36 pediatric intensive care unit of Lyon University Hospital, in France. Him and his mother (RBF254) of 29

37 years old, were diagnosed HIV positive on May, and June 2022 respectively. At diagnosis, plasma viral  
38 loads (pVL) and CD4 cell counts were 5.4 and 6.3 Log<sub>10</sub> cp/mL (Cobas® HIV-1, Roche Diagnostics, Swiss)  
39 and 331 and 141 CD4/mm<sup>3</sup>, for RBF254 and RBF255 respectively (Table 1). Resistance genotyping at  
40 baseline, revealed protease (PR) and reverse transcriptase (RT) regions belonging to HIV-1 group O,  
41 and integrase (IN) belonging to HIV-1 group M, with no major natural resistance (Table 1). cART were  
42 then started: bicitgravir-emtricitabine-tenofovir for RBF254 and zidovudine-lamivudine-dolutegravir  
43 quickly switched to abacavir-lamivudine-dolutegravir for RBF255 for pharmaceutical reasons. The pVL  
44 kinetics are presented in Table 1. These first molecular results suggested an atypical infection (HIV-  
45 1/M+O dual infection or HIV-1/MO recombinant form) and were then investigated in the National  
46 Reference Centre (NRC) for HIV. Nested group-specific RT-PCRs, targeting three regions in *pol* (PR, RT  
47 and IN) and GP41 in *env*, were simultaneously performed with group M- and group O-specific primers  
48 and followed by sequencing, as previously described [5]. Only the presence of HIV-1/O fragments was  
49 detected in the PR and RT regions, whereas only HIV-1/M fragments were detected in the IN and GP41  
50 regions, for both patients (Table 1). These discordant results suggested the presence of an HIV-1/MO  
51 recombinant form with, at least, a breakpoint between the RT and IN regions and without HIV-1/M  
52 and HIV-1/O parental strains.

53 Full-length genome amplifications were therefore performed, by amplifying five overlapping  
54 fragments using the DeepChek® Assay Whole Genome HIV-1 Genotyping kit (ABL, Advanced Biological  
55 Laboratories Diagnostics, France) [9], adapted with nested group-non-specific primers (Supp. Table 1).  
56 The different amplicons were used to perform NGS sequencing, using the Illumina DNA prep (Illumina,  
57 USA) protocol according to manufacturer's instructions on a MiSeq system (Illumina, USA), using the  
58 MiSeq Reagent Nano kit v2 (Illumina, USA). Full-length sequences were obtained using HAPHIPE bio-  
59 informatic pipeline [10].

60 Comparison of all sequences obtained were performed using the Los Alamos HIV BLAST [11]  
61 and NIBSC Genotyping tools [12], in order to identify the HIV-1/M and HIV-1/O groups of the parental  
62 strains involved in the HIV-1/MO recombinant form. Simplot analysis showed the presence of two  
63 breakpoints on full-length genomes, at the same locations for RBF254 and RBF255: the first at the  
64 beginning of IN region and the second in the GP41 region (Figure 1.A). Phylogenetic analysis also  
65 demonstrated that the fragments of HIV-1/O belonged to subgroup T, whereas the fragments of HIV-  
66 1/M belonged to subtype CRF02\_AG (Figure 1.B, C and D). As expected, the recombination and  
67 phylogenetic analyses showed that the two viruses were very similar and that this HIV-1/MO  
68 recombinant form was not linked to any URF\_MO already described, bringing to 20 the total number  
69 of URF\_MO currently described (Figure 1).

70           These data confirm a very rare MTCT of a HIV-1/MO recombinant form. This report is of  
71 importance for improving our knowledge on these atypical forms. Due to the small number of cases,  
72 it is not known whether such forms are more or less transmissible, or more or less virulent, owing to  
73 the presence of genomic regions from group O. Indeed, recent data showed a 1 Log<sub>10</sub> cp/mL lower viral  
74 replication of and a slower evolution of natural clinical history of group O viruses compared to group  
75 M [13,14]. Thus, these properties could impact the HIV-1/MO recombinant form, depending on the  
76 mosaic profiles, i.e., which regions are from HIV-1/O or HIV-1/M parental strain. In our case, the HIV-  
77 1/MO recombinant form caused a severe infection in both patients, as demonstrated by high pVL  
78 values and low CD4 cell count at diagnosis (Table 1). This form seems well adapted with an effective  
79 and intense replication leading to a rapid development of AIDS for the baby at childbirth. These  
80 properties could be linked to accessory genes (coding viral infectivity factors) or integrase belonging  
81 to group M. Indeed, it has been demonstrated that group O integrase has a lower enzymatic efficiency  
82 than that of group M [15].

83           The mosaic regions could also impact efficiency of antiretroviral therapy; thus, a group O RT  
84 limits the use of non-nucleoside reverse transcriptase inhibitors (NNRTI) [16]. Resistance genotyping  
85 at diagnosis allowed to confirm a natural polymorphism usually found in group O protease and RT, and  
86 group M integrase, and therefore to choose the better cART according to these groups. An effective  
87 immuno-virological response was observed in both patients; but difficulties were encountered for the  
88 baby, who still had a viral load around 2 Log<sub>10</sub> cp/mL at the last points. (Table 1). However, this was  
89 probably due to the very high pVL at cART initiation and difficulty to treat babies rather than to mosaic  
90 nature of the virus. These data showed that these infections could be efficiently treated like group M  
91 or group O mono-infections, with available drugs despite the mosaic structure of the genome.  
92 However, for this point, it is needed to determine the precise profile, confirming that resistance  
93 genotyping at diagnosis is essential. This report also confirms that any result combining two distinct  
94 variants of HIV, especially for resistance genotyping, has to be explored. This strategy allowed to  
95 discover a new recombinant form, distinct from the previous MO described by now, reflecting the  
96 incredible plasticity of the HIV.

97           In conclusion, we confirmed that HIV-1/MO recombinant forms are transmissible, and that  
98 they can lead to a severe infection. However, adapted cART can control these viruses. Our report of  
99 such a rare case of MTCT in a context of sporadic description of MO recombinant forms leads to the  
100 hypothesis of a higher circulation than initially estimated, and to the necessity of vigilance in presence  
101 of non-expected results.

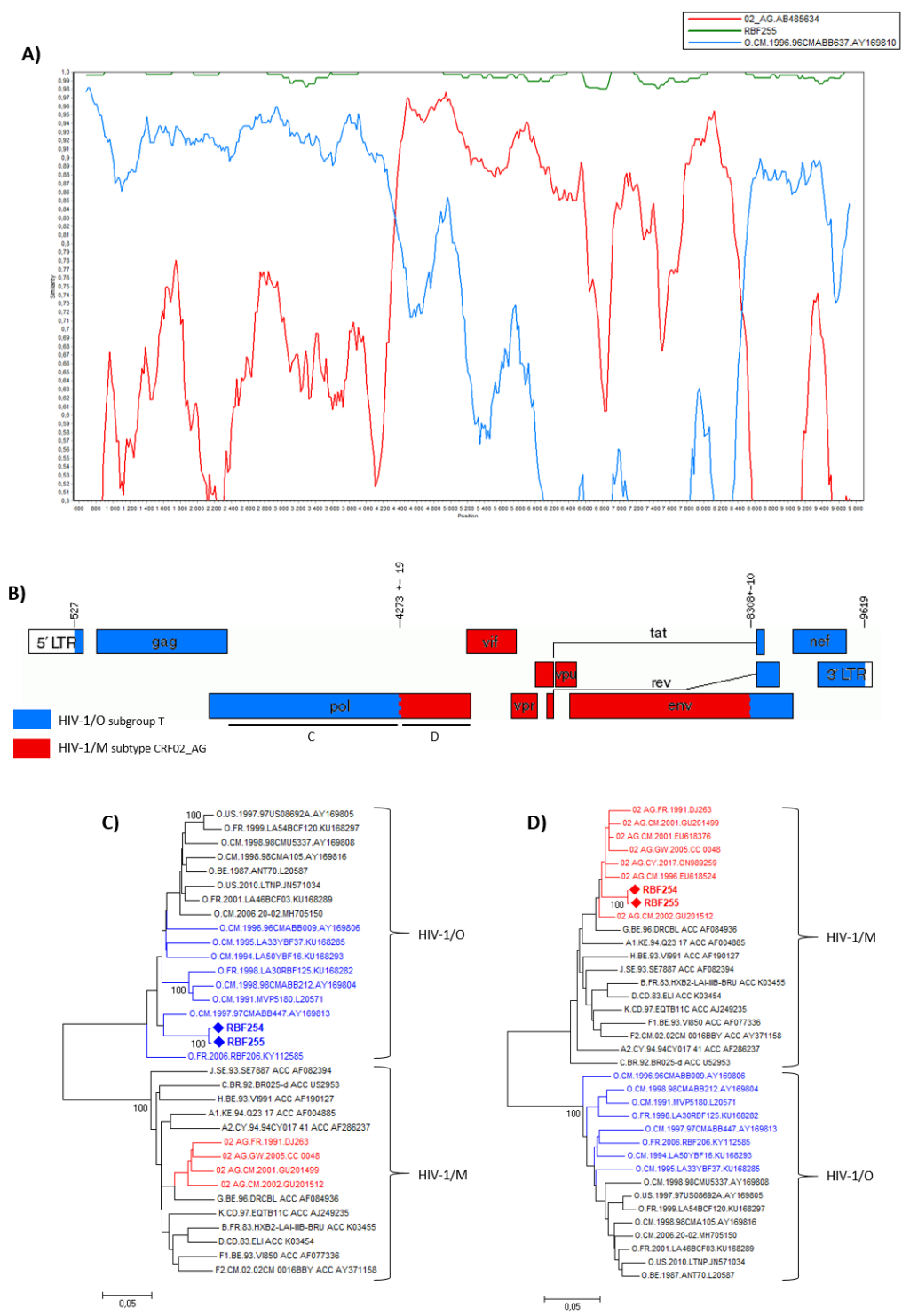


Table 1. Samples data associated with RBF254 and RBF255

	Sample collection date	Plasma viral load (Log <sub>10</sub> cp/mL)	CD4 count (/mm <sup>3</sup> )	Sequencing data									
				Location	PR	RT	IN	GP41	Full-length genome	HIV-1/M	HIV-1/O		
<b>RBF254</b>	2022-06-14	5.4	331	Lyon	HIV-1/O	HIV-1/O	-*						
	2022-06-16	5.4	308	Rouen	HIV-1/O	HIV-1/O	HIV-1/M	HIV-1/M	HIV-1/MO	CRF02_AG	T		
	<i>2022-06-30 initiation of cART (bictegravir-emtricitabine-tenofovir)</i>												
	2022-07-22	2.0	412										
	2022-08-22	1.7	430										
	2022-09-20	Undetectable											
	2022-10-19	1.8	465										
	2022-12-15	1.7	444										
	2023-03-23	Undetectable	514										
	<i>2023-03-23 initiation of cART (dolutegravir-emtricitabine-tenofovir)</i>												
<b>RBF255</b>	2022-05-31	6.3	141										
	2022-06-02		141										
	<i>2022-06-07 initiation of cART (zidovudine-lamivudine-dolutegravir)</i>												
	2022-06-08			Lyon	HIV-1/O	HIV-1/O	HIV-1/M						
	<i>2022-06-14 initiation of cART (abacavir-lamivudine-dolutegravir)</i>												
	2022-06-15	3.3		Rouen	HIV-1/O	HIV-1/O	-	HIV-1/M					
	2022-06-23	2.7											
	2022-07-07	4.2	927										
	2022-07-12	4.0											
	2022-07-28	4.8		Rouen			HIV-1/M		HIV-1/MO	CRF02_AG	T		
	2022-08-04	2.7											
	2022-08-23	2.5											
	2022-09-13	2.1	685										
	2022-10-11	2.4	850										
	2022-11-15	2.4	1220										
	2022-12-14	2.1	1456										
	2023-01-19	1.9	1260										
	2023-02-14	2.0											
	2023-03-14	2.2	1110										

\*-: No amplification

PR: Protease, RT: Reverse transcriptase, IN: Integrase



**Figure 1 : Analysis of recombination profiles and phylogenetic relationships of RBF254 and RBF255**

HIV-1/M subtype CRF02 and HIV-1/O subgroup T are represented in red and blue respectively **A) Simplot Analysis.** Simplot was used to identify and locate recombination breakpoints with 300 nucleotides windows, 20 nucleotides increments, and the Kimura 2-parameters method with a transition-transversion ratio of 2.0 [17]. RBF255 and full-length genome of HIV-1/M (AB485634) and HIV-1/O (AY169810) sequences were mapped to identify sequences similarities on RBF254. **B) Genomic profile of full-length genome of RBF254 and RBF255 HIV-1/MO recombinants form.** Genome map was obtained using the Recombinant HIV Drawing Tool [18]. Two breakpoints were identified at position 4253-4293 at the beginning of IN region, and 8297-8318 bp for the second breakpoint in GP41. All positions in bp, including breakpoints, were indicated on HxB2 HIV-1 reference strain (GenBank accession number K03455). **C) Phylogenetic trees from the PR-RT and D) IN region.** All sequences were aligned with MEGA 7 software [19] with a set of HIV-1/M and HIV-1/O sequences from the Los Alamos HIV database [20] and our laboratory database, in order to perform phylogenetic analysis. Phylogenetic trees were obtained using the neighbor-joining method. The evolutionary distances, corresponding to substitutions per site, were computed using the Kimura 2-parameter method and showed by scale at the bottom of each trees. RBF254 and RBF255 strains are highlighted in bold, preceded by a diamond-shaped symbol. Reliability was estimated by bootstrap analysis with 1000 replicates and all bootstraps less than 100 were removed. The two full-length genome of RBF254 and RBF255 have been respectively submitted to GenBank database, under accession numbers OR162377 and OR162378.

## REFERENCES

1. Peeters M, Liegeois F, Torimiro N, et al. Characterization of a Highly Replicative Intergroup M/O Human Immunodeficiency Virus Type 1 Recombinant Isolated from a Cameroonian Patient. *J Virol* **1999**; 73:7368–7375.
2. Takehisa J, Zekeng L, Ido E, et al. Human Immunodeficiency Virus Type 1 Intergroup (M/O) Recombination in Cameroon. *J Virol* **1999**; 73:6810–6820.
3. Yamaguchi J, Bodelle P, Vallari AS, 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**; 20:944–957.
4. Vessière A, Leoz M, Brodard V, et al. First evidence of a HIV-1 M/O recombinant form circulating outside Cameroon. *AIDS* **2010**; 24:1079.
5. Ngoupo PA, Sadeuh-Mba SA, De Oliveira F, et al. First evidence of transmission of an HIV-1 M/O intergroup recombinant virus. *AIDS* **2016**; 30:1–8.
6. De Oliveira F, Mourez T, Vessiere A, 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. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5237259/>. Accessed 30 March 2021.
7. De Oliveira F, Cappy P, Lemée V, et al. Detection of numerous HIV-1/MO recombinants in France. *AIDS* **2018**; 32:1289–1299.
8. 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. *Clinical Microbiology and Infection* **2020**; 26:262–264.
9. DeepChek® Assay Whole Genome HIV-1 Genotyping - ABL Diagnostics S.A. Available at: <https://www.abldiagnostics.com/products/deepchek-assay-whole-genome-hiv-1-genotyping/>. Accessed 27 March 2023.
10. Bendall ML, Gibson KM, Steiner MC, Rentia U, Pérez-Losada M, Crandall KA. HAPPIPE: Haplotype Reconstruction and Phylodynamics for Deep Sequencing of Intra-host Viral Populations. *Mol Biol Evol* **2020**; 38:1677–1690.
11. BLAST search. Available at: [https://www.hiv.lanl.gov/content/sequence/BASIC\\_BLAST/basic\\_blast.html](https://www.hiv.lanl.gov/content/sequence/BASIC_BLAST/basic_blast.html). Accessed 31 January 2023.
12. Genotyping. Available at: <https://www.ncbi.nlm.nih.gov/projects/genotyping/formpage.cgi>. Accessed 15 March 2023.
13. Kouanfack C, Unal G, Schaeffer L, et al. Comparative Immunovirological and Clinical Responses to Antiretroviral Therapy Between HIV-1 Group O and HIV-1 Group M Infected Patients. *Clinical Infectious Diseases* **2020**; 70:1471–1477.
14. Unal G, Alessandri-Gradt E, Leoz M, et al. Human Immunodeficiency Virus Type 1 Group O Infection in France: Clinical Features and Immunovirological Response to Antiretrovirals. *Clinical Infectious Diseases* **2018**; 66:1785–1793.
15. Depatureaux A, Mespède T, Quashie P, et al. HIV-1 group O integrase displays lower susceptibility to raltegravir and has a different mutational pathway for resistance than HIV-1 group M. *J Int AIDS Soc* **2014**; 17:19738.
16. Alessandri-Gradt E, Moisan A, Plantier J-C. HIV-1 Non-Group M Strains and ART. *Viruses* **2023**; 15:780.
17. Lole KS, Bollinger RC, Paranjape RS, 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**; 73:152–160.
18. Recombinant Genome Drawing Tool. Available at: [https://www.hiv.lanl.gov/content/sequence/DRAW\\_CRF/recom\\_mapper.html](https://www.hiv.lanl.gov/content/sequence/DRAW_CRF/recom_mapper.html). Accessed 31 January 2023.
19. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution* **2016**; 33:1870–1874.
20. HIV Databases. Available at: <https://www.hiv.lanl.gov/content/index>. Accessed 1 February 2023.