



HAL
open science

Définition in silico d'épitopes induisant une réponse T cytotoxique en fonction de la variabilité virale du VIH-1 et de l'immunogénétique des patients

Camille Tumiotto

► **To cite this version:**

Camille Tumiotto. Définition in silico d'épitopes induisant une réponse T cytotoxique en fonction de la variabilité virale du VIH-1 et de l'immunogénétique des patients. Médecine humaine et pathologie. Université de Bordeaux, 2018. Français. NNT : 2018BORD0153 . tel-01951600

HAL Id: tel-01951600

<https://theses.hal.science/tel-01951600>

Submitted on 11 Dec 2018

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.

THÈSE PRÉSENTÉE
POUR OBTENIR LE GRADE DE
DOCTEUR DE
L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE SCIENCES DE LA VIE ET DE LA SANTE (SVS)
SPÉCIALITÉ MICROBIOLOGIE-IMMUNOLOGIE

Par Camille TUMIOTTO

**DEFINITION *IN SILICO* D'ÉPITOPES INDUISANT UNE
REPONSE T CYTOTOXIQUE EN FONCTION DE LA
VARIABILITE VIRALE DU VIH-1 ET DE
L'IMMUNOGENETIQUE DES PATIENTS**

Sous la direction de : Hervé FLEURY

Soutenue le 19 Septembre 2018

Membres du jury :

M. MOREAU Jean-Francois	PU-PH	Bordeaux	Président
M. LE GRAND Roger	DVM	Fontenay-aux-Roses	Rapporteur
M. MIRAMBEAU Gilles	MCU	Paris	Rapporteur
M. FLEURY Hervé	PU-PH	Bordeaux	Directeur de Thèse

Remerciements

A mon directeur de thèse

Monsieur le Professeur Hervé Fleury
Professeur des Universités – Praticien Hospitalier
Laboratoire de Virologie
Centre Hospitalier Universitaire de Bordeaux

Vous êtes à l'origine de ce travail et je vous suis reconnaissante de m'avoir permis de travailler sur ce très beau projet. Je vous remercie d'avoir bien voulu diriger ce travail. Merci également pour votre aide et pour votre soutien.

A mon président de jury

Monsieur le Professeur Jean-François Moreau
Professeur des Universités – Praticien Hospitalier
Laboratoire d'Immunologie
Centre Hospitalier Universitaire de Bordeaux

Vous me faites l'honneur de présider cette thèse et de juger mon travail qui unit la Virologie et l'Immunologie. Veuillez recevoir mes plus vifs remerciements et l'expression de ma sincère considération.

A mes rapporteurs

Monsieur le Docteur Gilles Mirambeau
Maitre de Conférence des Universités
Université de la Sorbonne, Paris

Monsieur le Docteur Roger Le Grand
Directeur de Recherche
CEA- Fontenay-aux-Roses

Vous avez accepté avec bienveillance de vous prononcer sur l'intérêt de ce travail. Soyez remerciés également pour avoir accepté de vous plier aux contraintes de la visioconférence. Soyez assurés de ma grande reconnaissance et de mon profond respect.

Lionel, merci pour ta présence, ton soutien, ton calme au quotidien. Je te remercie aussi de ton aide pour le développement du logiciel TutuGenetics et de m'avoir poussée à apprendre à coder. Je suis ravie de cette nouvelle page, nouvelle vie qui s'ouvre devant nous.

Maman, Papa, Charles, merci pour votre soutien inconditionnel et votre confiance en moi. Je n'ose plus dire que cette fois-ci c'est la dernière... ☺ Merci d'être toujours présents malgré les kilomètres que je ne cesse de rajouter entre nous !

Coco, Mamie, Pipof, Baptiste, Papi et Mamie, merci pour votre présence, votre soutien et tous les moments passés ensemble.

Caroline, Thierry, merci encore de m'avoir accueillis il y a déjà 12 ans!

Pantxika, Paty, Isabelle, Pascale, Marie-Edith, merci pour votre soutien, votre compréhension et votre amitié. J'ai adoré et beaucoup appris (et pas seulement en virologie ☺) au cours de ces années partagées !!

Samuel, Emma, Anne-Charlotte, Arnaud, Alexandra, Raphael, Delphine, Clément, Anne-Sophie, Florian, Alice, Ugo, Tiana, Antoine, Cécile, et tous ceux que j'oublie... Merci pour votre soutien, votre présence !!

Merci également à l'équipe du laboratoire de Virologie, l'équipe de Génétique et toutes les personnes avec qui j'ai pris plaisir à partager et travailler.

SOMMAIRE

Sommaire

I. INTRODUCTION	13
A. GENERALITES	13
1. <i>Données épidémiologiques</i>	13
2. <i>Variabilité du VIH</i>	14
a) Classification	14
b) Origine de la variabilité virale.....	17
B. GENERALITES VIROLOGIQUES DU VIH-1	18
C. CYCLE DE REPLICATION DU VIH-1	20
D. HISTOIRE NATURELLE DE L'INFECTION A VIH-1	22
E. REPONSES IMMUNES CONTRE LE VIH	24
1. <i>Réponse CTL et VIH</i>	25
2. <i>Réponse CTL et variabilité génétique</i>	26
3. <i>Réponse CTL et variabilité virale</i>	28
F. ERADICATION VIRALE.....	29
1. <i>Traitements</i>	29
2. <i>Notion de réservoir</i>	30
3. <i>Stratégies d'éradication du VIH-1</i>	32
4. <i>Vaccination thérapeutique</i>	33
II. PROJET PROVIR ET OBJECTIF DE LA THESE.....	37
A. PROJET PROVIR/LATITUDE 45.....	37
B. PROJET DE THESE.....	38
III. 1^{ERE} PARTIE - POPULATION DE L'ETUDE.....	40
A. PATIENTS VIVANTS AVEC LE VIH DANS LA REGION AQUITAINE	40
1. <i>Publication 1</i>	40
2. <i>Publication 2</i>	48
B. POPULATION DE L'ETUDE.....	58

IV. 2^{EME} PARTIE - RECHERCHE IN SILICO D'EPITOPES	
POUVANT INDUIRE UNE REPONSE T CYTOTOXIQUE	65
A. PUBLICATION 3	65
B. LES PATIENTS INCLUS DANS LE PROJET PROVIR/LATITUDE 45 AU CHU DE BORDEAUX.....	83
1. <i>Objectif</i>	83
2. <i>Matériels et méthodes</i>	83
a) Inclusion	83
b) Données immuno-virologiques	83
c) Extractions d'ADN.....	83
d) Séquençage du VIH-1	84
e) PCR ADN VIH.....	88
f) Typage HLA de classe I.....	88
g) Prédiction de la reconnaissance des épitopes en fonction du HLA par bio- informatique	89
h) Logiciel TutuGenetics	89
3. <i>Résultats</i>	90
V. CONCLUSIONS ET PERSPECTIVES	95
VI. BIBLIOGRAPHIE	100
VII. ANNEXES.....	109

Liste des tableaux

<i>Tableau 1 : Fréquences et rangs des 5 HLA A et B majoritaires dans les populations Caucasiennes (Cauc.), Africaines (Afric.), Asiatiques (Asiat.) et Hispaniques (Hisp.) d'après Maiers et al. (Maiers, Gragert, and Klitz, 2007; Martin Maiers, Loren Gragert, William Klitz, n.d.).</i>	26
<i>Tableau 2 : Caractéristiques immuno-virologiques des patients inclus sur Bordeaux dans le Projet Provir/Latitude 45</i>	59
<i>Tableau 3 : Liste des amorces utilisées pour amplifier les gènes Gag, Protease et Reverse Transcriptase et Nef</i>	85
<i>Tableau 4 : Liste « Optimal_Provir » décrivant les épitopes d'intérêt en fonction des HLA étudiés.</i>	91

Liste des figures

<i>Figure 1 : Nombre et répartition des personnes infectées par le VIH dans le monde en 2016 d'après UNAIDS ("Core Epidemiology Slides" 2017)</i>	<i>14</i>
<i>Figure 2 : Diversité génétique des groupes M, N,O et P du VIH-1, VIH-2 et SIVs (d'après (Tebit and Arts 2011)).....</i>	<i>15</i>
<i>Figure 3 : Répartition des sous-types viraux du VIH-1 majoritaires dans le monde, d'après Taylor et al. (B. S. Taylor et al. 2008).</i>	<i>16</i>
<i>Figure 4 : Structure du VIH-1, d'après Campbell et al. (Campbell and Hope 2015). ..</i>	<i>18</i>
<i>Figure 5 : Organisation génomique du VIH-1, d'après Peterlin et al. (Peterlin and Trono 2003).....</i>	<i>18</i>
<i>Figure 6 : Cycle réplcatif du VIH-1, d'après Maartens et al. Les cibles des différents antirétroviraux disponibles à ce jour sont indiquées dans les bulles (Maartens, Celum, and Lewin 2014).....</i>	<i>21</i>
<i>Figure 7: Evolution typique de l'infection à VIH-1, d'après Pantaleo et al. (Pantaleo, Graziosi, and Fauci 1993) et Coffin et al. (John M Coffin, Stephen H Hughes, and Harold E Varmus 1997).</i>	<i>23</i>
<i>Figure 8 : Schéma de la présentation d'un peptide au lymphocyte T par le HLA.</i>	<i>25</i>
<i>Figure 9 : Origine et devenir des réservoirs, d'après Deeks et al. (Deeks et al. 2012).</i>	<i>30</i>
<i>Figure 10 : Schéma représentant le descriptif de l'étude.....</i>	<i>58</i>
<i>Figure 11 : Tris sélectifs ayant permis la constitution de notre liste d'épitopes d'intérêt en fonction des HLA étudiés.</i>	<i>92</i>

ABBREVIATIONS

Abréviations

ADCC	Antibody Dependent Cellular Cytotoxicity
ADN	Acide désoxyribonucléique
ARN	Acide ribonucléique
ARV	Antirétroviral
CA	Capside
CAS9	CRISPR Associated protein 9
CCL5	Chemokine Ligand 5
CCR5	C-C Chemokine Receptor type 5
CD	Cluster de Différenciation
CEA	Commissariat à l'Energie Atomique et aux Energies Alternatives
CHU	Centre Hospitalier Universitaire
CMH	Complexe Majeur Histocompatibilité
CRF	Circulating Recombinant Form
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CTL	Lymphocyte T Cytotoxique
CXCR4	C-X-C Chemokine Receptor type 4
ELISpot	Enzyme-Linked ImmunoSpot
GALT	Gut Associated Lymphoid Tissue
GM-CSF	Granulocyte- Macrophage Colony-Stimulating Factor
GP	Glycoprotéine
GRIC	Groupe de Recherche en Immunologie Clinique
HLA	Human Leukocyte Antigen
IEDB	Immune Epitope DataBase
IFN	Interferons
IL7	Interleukin 7
IN	Intégrase
LRA	Latency Reversing Agent
LTR	Long Terminal Repeats
MA	Matrice
MHC IC50	Constante d'affinité entre une molécule HLA et un épitope
MIP	Macrophage Inflammatory Protein

NAP	Données non disponibles
NC	Nucléo-Capside
ND	Non détectable
NGS	Next Generation Sequencing
ORF	Open Reading Frame
PBMC	Peripheral Blood Mononuclear Cell
PCR	Polymerase Chain Reaction
PD1	Program cell Death protein 1
PR	Protéase
PVVIH	Personnes Vivant avec le VIH
RANTES	Regulated on Activation, Normal T cell Expressed and Secreted
RT	Reverse Transcriptase
SIDA	Syndrome d'Immuno-Depression Acquise
SIV	Simian Immunodeficiency Virus
TALEN	Transcription Activator-Like Effector Nucleases
TAP	Protéines de transport des antigènes
TCR	T Cell Receptor
TNF	Tumor Necrosis Factor
URF	Unique Recombinant Form
VIH	Virus de l'Immunodéficience Humaine
ZFN	Zinc Finger Nuclease

INTRODUCTION

I. Introduction

A. Généralités

Deux types de virus de l'immunodéficience humaine (VIH) existent. Le VIH de type 1 (VIH-1) identifié en 1983 par l'équipe de Barré-Sinoussi et Montagnier, est responsable de la pandémie actuelle (Barre-Sinoussi et al. 1983). Le VIH de type 2 (VIH-2), connu depuis 1986 cause une épidémie localisée en Afrique de l'Ouest (Clavel et al. 1986). Ces virus sont associés au syndrome d'immunodéficience acquise (SIDA), maladie caractérisée par une immunosuppression sévère.

1. Données épidémiologiques

En 2016 (cf *Figure 1*), dans le monde environ 36,7 millions de personnes vivaient avec le VIH-1, et on comptabilisait 1,8 millions de nouvelles infections. L'Afrique subsaharienne est la région la plus touchée avec 69% des personnes vivant avec le VIH (PVVIH) dans le monde. Dans le cadre de la lutte contre l'infection VIH, l'Organisation Mondiale de la Santé a défini les objectifs 90-90-90, à atteindre pour 2020. A savoir : 90% des PVVIH diagnostiquées, 90% des personnes diagnostiquées sous thérapie antirétrovirale (ARV) et 90% des personnes traitées avec une charge virale contrôlée.

Malgré les efforts mis dans les programmes de prévention et de traitement, ces objectifs sont loin d'être atteints dans le monde. En effet, même si les programmes de prévention et de traitements ont permis de réduire la mortalité (diminution de 48% depuis l'an 2000), le nombre de décès dus au SIDA reste encore de 1 million de personnes en 2016 ("UNAIDS DATA 2017," n.d.).

Adults and children estimated to be living with HIV | 2016

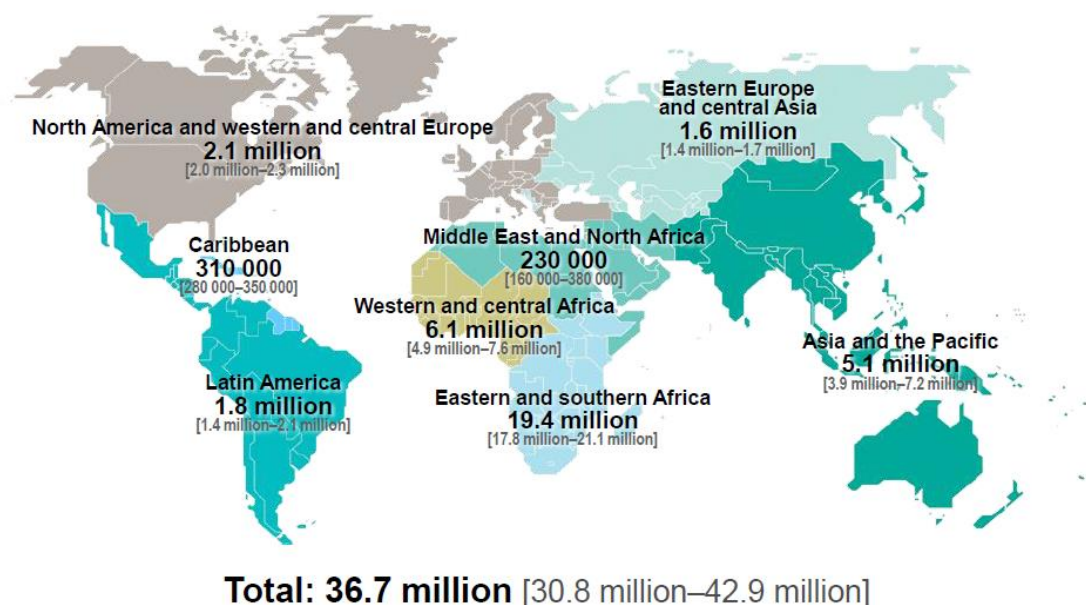


Figure 1 : Nombre et répartition des personnes infectées par le VIH dans le monde en 2016 d'après UNAIDS ("Core Epidemiology Slides" 2017)

En France, en 2016, 180 000 personnes étaient infectées par le VIH-1 et 140 000 étaient sous traitement. Cette même année, 6 000 personnes ont découvert leur séropositivité, chiffre qui reste stable depuis au moins 10 ans ("UNAIDS DATA 2017," n.d.).

2. Variabilité du VIH

a) Classification

Le VIH est apparenté au virus simien SIV (Simian immunodeficiency virus). Selon sa classification phylogénétique, le VIH est divisé en 2 types VIH-1 et VIH-2 (cf Figure 2). Ce dernier est constitué de 8 groupes, les groupes A et B étant majoritaires. Le VIH-1 est quant à lui composé de 4 groupes répartis dans différentes zones géographiques (Hemelaar 2013) :

- Le groupe M (« major » ou majoritaire) avec une distribution mondiale,
- Le groupe O (« outlier » ou externe) retrouvé au Cameroun, au Gabon et en Guinée Equatoriale,
- Le groupe N (« non-M, non-O » ou « new ») isolé uniquement au Cameroun,
- Le groupe P au Cameroun

Le groupe M comprend neuf sous-types (de A à D, de F à H, J et enfin K) et il existe actuellement 98 sous-types recombinants (CRF : Circulating Recombinant Form) ainsi que des Uniques Recombinant Form (URFs), issus de phénomènes de recombinaison entre 2 ou plusieurs virus de sous-type différents co-infectant un même individu.

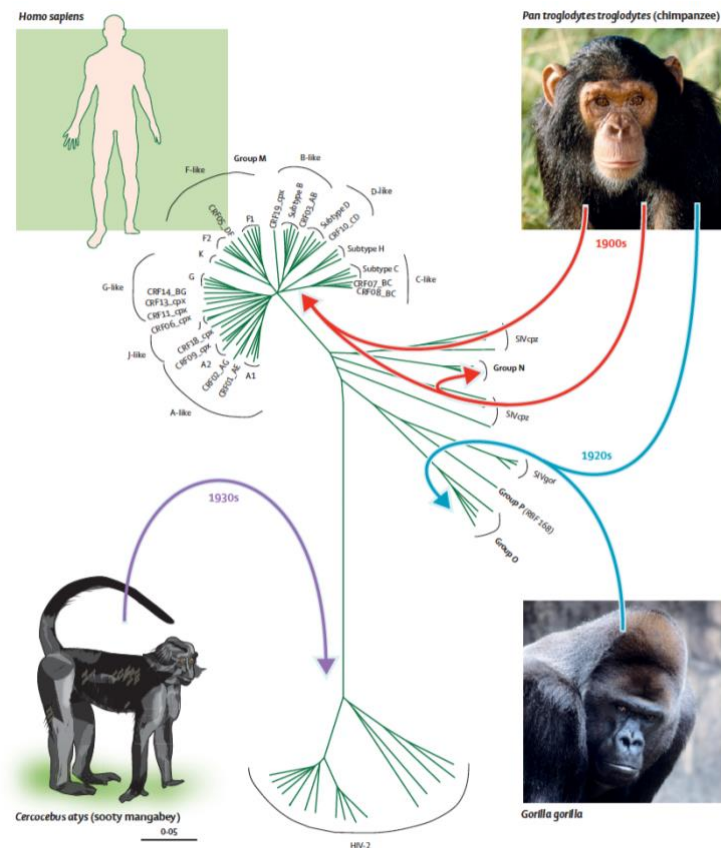


Figure 2 : Diversité génétique des groupes M, N, O et P du VIH-1, VIH-2 et SIVs (d'après (Tebit and Arts 2011))

Les groupes M et N du VIH-1 sont apparentés au SIV du chimpanzé *Pan troglodytes troglodytes* (*Ptt*) alors que les groupes O et P sont proches du SIV du gorille (*Gorilla gorilla*; *SIVgor*). Le VIH-2 trouve son origine dans le SIV infectant les singes verts mangabey (*Cercocebus atys*)

Leur distribution géographique est illustrée dans la Figure 3.

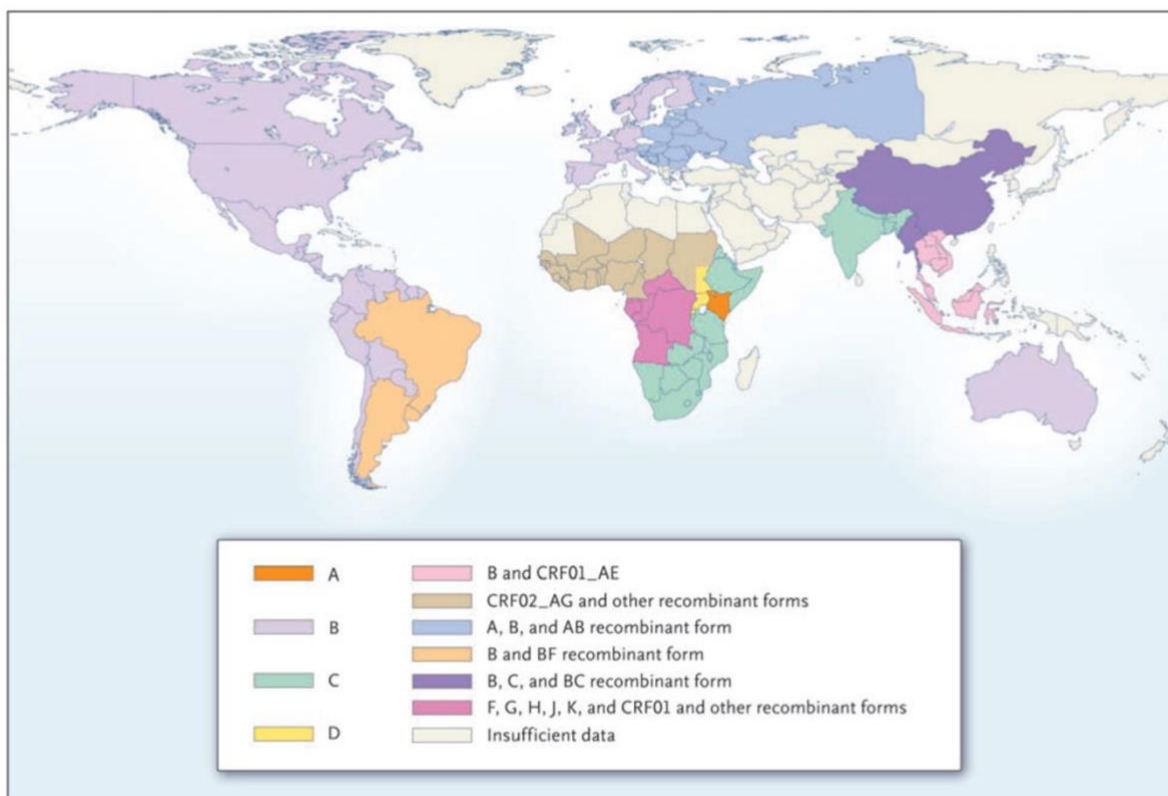


Figure 3 : Répartition des sous-types viraux du VIH-1 majoritaires dans le monde, d'après Taylor et al. (B. S. Taylor et al. 2008).

Le sous-type viral B est majoritaire en France et dans les pays développés, alors que le sous-type viral C est responsable de plus de la moitié des infections à VIH-1 dans le monde.

En France, on note une augmentation de la circulation de virus de sous-type non -B tel que le CRF02_AG (Descamps et al. 2013).

b) Origine de la variabilité virale

Chez un individu infecté, le VIH est présent sous la forme d'une population virale composée de nombreux variants génétiquement distincts mais apparentés définissant la quasi-espèce virale.

Cette variabilité est le résultat de 4 facteurs principaux :

- une réplication intense avec la production quotidienne d'environ 10^{10} - 10^{11} virions (Ho et al. 1995)
- des mutations générées par la transcriptase inverse virale, du fait de sa faible activité de proofreading (Roberts, Bebenek, and Kunkel 1988)
- des hypermutations générées par les enzymes éditrices de type APOBEC (Mangeat et al. 2003)
- des phénomènes de recombinaison lors des cycles réplcatifs (J. E. Taylor and Korber 2005) .

Cette diversité génétique permet au virus de s'adapter aux différentes pressions de sélection imposées par le système immunitaire ou les traitements antirétroviraux mais constitue un obstacle pour l'éradication du virus (Smyth, Davenport, and Mak 2012; Hemelaar 2013; Smyth and Negroni 2016)

B. Généralités virologiques du VIH-1

Le VIH-1 appartient à la famille des rétrovirus et à la sous-famille des *lentivirus*. Ce virus enveloppé a un aspect sphérique avec une taille allant de 90 à 120 nm. Il contient 2 molécules d'ARN.

La Figure 4 ci-dessous schématise la structure du virus.

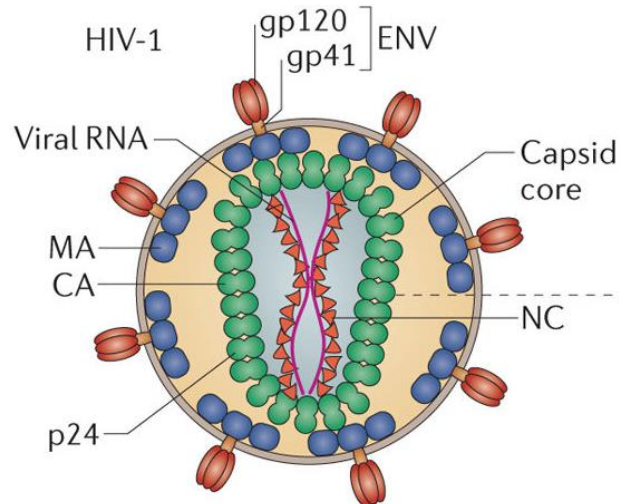


Figure 4 : Structure du VIH-1, d'après Campbell et al. (Campbell and Hope 2015).

La membrane virale est composée de lipides issus des cellules productrices des virions et de glycoprotéines virales nécessaires à l'attachement et l'entrée dans la cellule cible.

Au niveau de la surface interne de l'enveloppe, la matrice p17 forme un réseau stabilisant la particule virale. Au sein de cette enveloppe se trouve le « cœur » ou core viral composé d'environ 1500 copies de capsid p24 contenant les 2 molécules d'ARN viral, la NC ainsi que les enzymes viraux la PR, la RT et l'IN.

Le VIH-1 a un génome ARN simple brin de polarité positive composé de 2 molécules. Une fois intégré, le génome est appelé provirus et fait environ 9200 nucléotides. Son organisation génomique est représentée dans la Figure 5 ci-dessous.

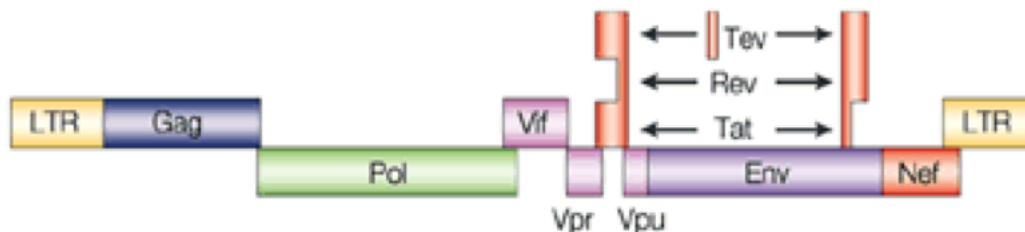


Figure 5 : Organisation génomique du VIH-1, d'après Peterlin et al. (Peterlin and Trono 2003).

Les extrémités du provirus sont flanquées par des séquences répétées nommées LTRs (long terminal repeats) qui se divisent en trois éléments : U3, R et U5. Ces régions contiennent des signaux importants pour l'intégration, la transcription etc...

Le génome viral est constitué de 3 gènes majeurs :

- Le gène Gag code pour une polyprotéine virale GagPol (Pr55gag) qui va être clivée par la protéase virale pour générer 6 protéines : la matrice (MA ou p17), la capside (CA ou p24), la nucléocapside (NC ou p7) le peptide C-terminal P6, le peptide N-terminal P2 ainsi que le peptide C-terminal P1.
- Le gène Pol code pour le précurseur Pr160gag-pol de 160 kDa qui est mûr pour générer les trois enzymes virales : la Protéase (PR), la transcriptase inverse ou Reverse Transcriptase (RT), l'intégrase (IN).
- Le gène Env code pour une polyprotéine d'enveloppe clivée par la protéase cellulaire en une glycoprotéine de surface SU ou gp120 et une glycoprotéine transmembranaire TM ou gp41.

Ce virus a en plus de ces gènes majeurs, des gènes accessoires spécifiques du VIH-1 qui codent pour les protéines accessoires :

- Tat est impliqué dans la régulation de la transcription des gènes viraux.
- Rev stimule le transport nucléo-cytoplasmique des ARNm mono ou non épissés.
- Vif induit la dégradation par le protéasome du facteur cellulaire de restriction APOBEC3G.
- Vpr a un rôle dans l'import nucléaire du complexe de pré-intégration.
- Vpu active le bourgeonnement des nouveaux virions.
- Nef inhibe l'expression de plusieurs récepteurs cellulaires : CD4, CD8, CD28, Complexe Majeur d'Histocompatibilité (CMH) de classe I et II.

C. Cycle de réplication du VIH-1

Le cycle du VIH-1 est représenté dans la *Figure 6* ci-dessous. L'entrée du virus dans la cellule de l'hôte se fait par interaction entre le virus et le récepteur CD4 associé à un corécepteur pouvant être une molécule CCR5 ou CXCR4. La principale cible du VIH est le lymphocyte T CD4+. D'autres cellules possédant le CD4 et le corécepteur sont aussi infectées : les lymphocytes T CD4+ quiescents, les monocytes, les macrophages et les cellules dendritiques. Après l'entrée du virus dans la cellule, la reverse transcriptase virale (RT) rétrotranscrit l'ARN viral en ADN double brin, puis l'intégrase virale assure l'intégration de cet ADN dans le génome humain sous forme d'ADN proviral.

L'ADN proviral est ensuite pris en charge par la machinerie cellulaire pour la transcription qui va générer de nouveaux génomes viraux et des ARNm. S'en suit l'épissage de certains ARNm puis l'export du noyau vers le cytoplasme. A ce niveau, les ARNm sont traduits en polyprotéines virales qui vont être maturées par l'action des protéases virale et cellulaire. Après assemblage, il y a bourgeonnement et relargage des nouveaux virions avec une dernière maturation des protéines par la protéase à l'intérieur du virion (Maartens, Celum, and Lewin 2014).

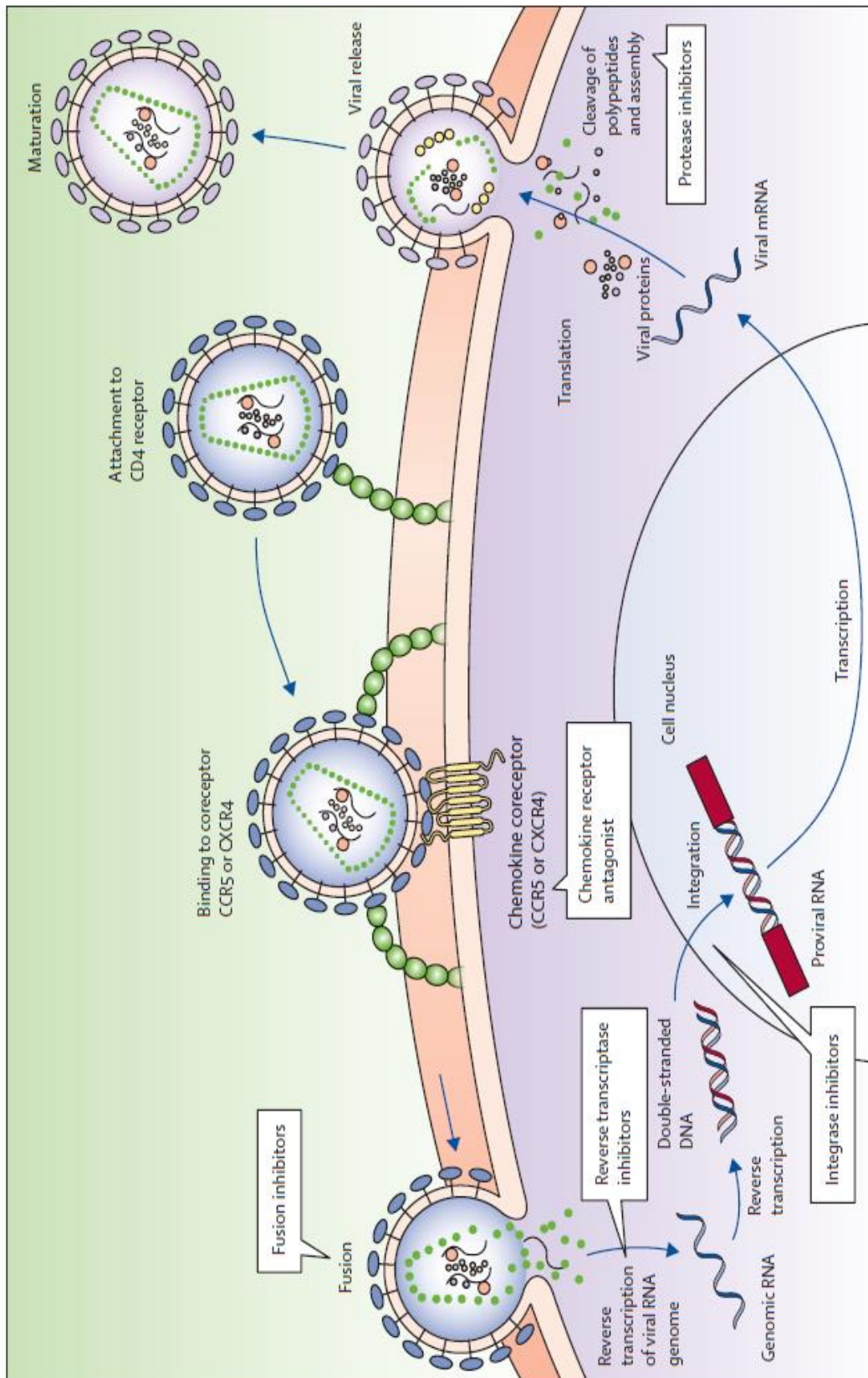


Figure 6 : Cycle répliatif du VIH-1, d'après Maartens et al. Les cibles des différents antirétroviraux disponibles à ce jour sont indiquées dans les bulles (Maartens, Celum, and Lewin 2014)

D. Histoire naturelle de l'infection à VIH-1

La contamination par le virus du VIH-1 se fait selon 3 modes de transmission : sexuelle, materno-fœtale et sanguine (Rugeles, Velilla, and Montoya 2011).

Lors de la primo-infection, les virus ont majoritairement un tropisme macrophagique et utilisent le corécepteur CCR5, peu ont un tropisme lymphocytaire utilisant le corécepteur CXCR4. En France, selon les études et la méthodologie employée, les virus à tropisme CXCR4 au moment de la primo-infection concernent de 6,4 % (Raymond et al. 2010) à 15,9 % (Frange et al. 2009) des cas. Au cours de l'infection, on observe un « switch » vers un tropisme lymphocytaire avec le corécepteur CXCR4, conséquence de la capacité à muter du virus.

Comme précisé plus haut, la cible principale du VIH-1 est le lymphocyte T CD4+. L'évolution de l'infection par le VIH varie d'un individu à l'autre mais elle suit généralement un développement similaire résumé dans la *Figure 7*.

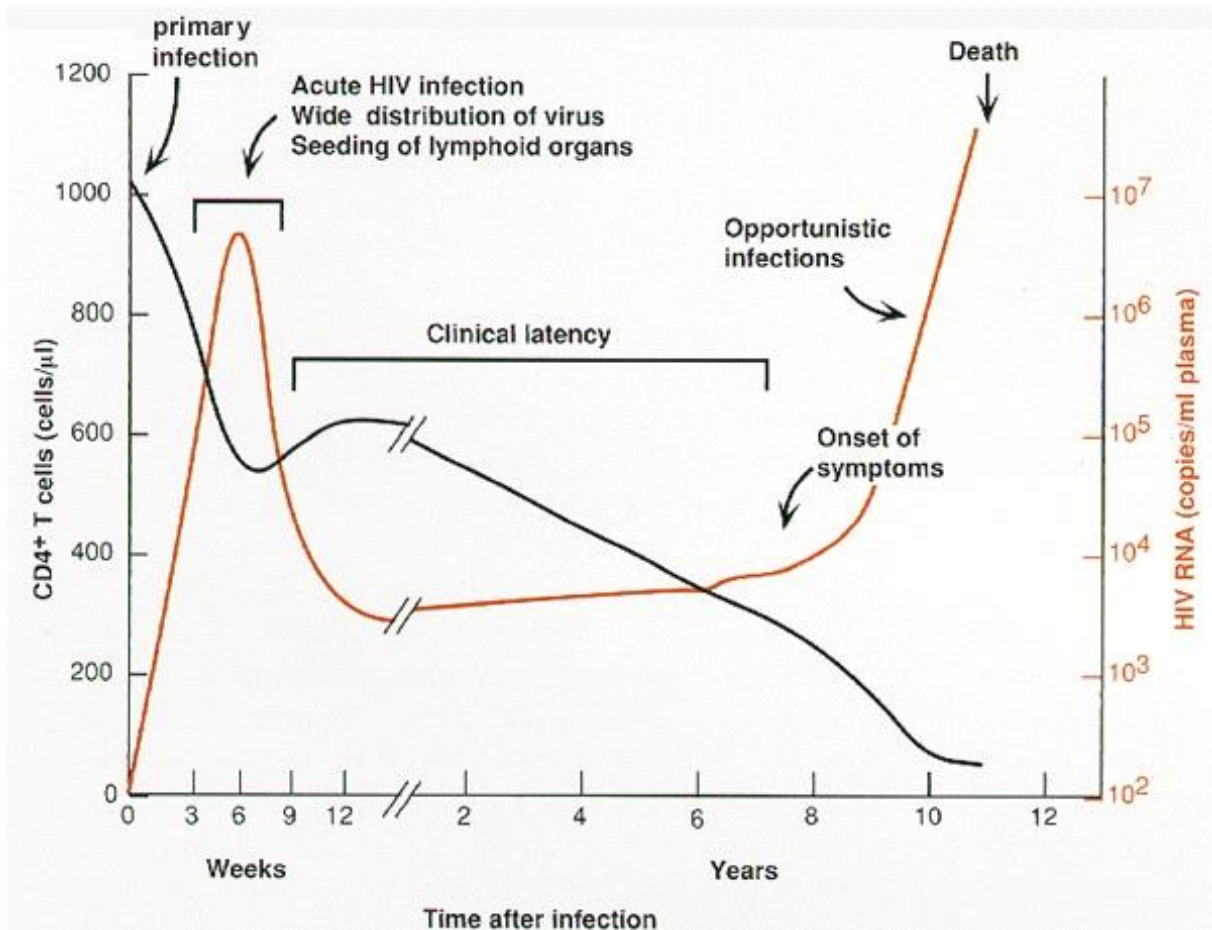


Figure 7: Evolution typique de l'infection à VIH-1, d'après Pantaleo et al. (Pantaleo, Graziosi, and Fauci 1993) et Coffin et al. (John M Coffin, Stephen H Hughes, and Harold E Varmus 1997).

La courbe noire représente l'évolution du nombre de CD4 (cellules/ μ L) et la courbe orangée représente l'évolution de la charge virale (copies/mL). La décroissance des lymphocytes T CD4+, l'augmentation et la valeur de la charge virale VIH-1 varient grandement d'un patient à l'autre.

Au cours de la primo-infection, les signes cliniques miment une mononucléose infectieuse et sont peu spécifiques : fièvre, asthénie intense, myalgies, pharyngite, éruption cutanée, adénopathie, diarrhée, ulcérations buccales et génitales. Rarement les manifestations cliniques sont plus sévères : pneumopathie interstitielle ou symptomatologie neurologique aiguë à type de méningite lymphocytaire aiguë. Les signes cliniques sont présents dans 50 à 75 % des primo-infections et apparaissent entre 3 et 6 semaines après la contamination (Pantaleo, Graziosi, and Fauci 1993; John M Coffin, Stephen H Hughes, and Harold E Varmus 1997; H. Fleury 2009). La réplication virale est intense, les réservoirs se constituent et il y a une diminution du nombre de lymphocytes T CD4+.

Les symptômes disparaissent spontanément au bout de quelques semaines. La réponse liée aux lymphocytes T CD8+ cytotoxiques (CTL) spécifiques du VIH entraîne une diminution de la

charge virale plasmatique jusqu'à une valeur basse plasmatique stable entre 3 et 6 mois après l'infection (Legrand et al. 1997; Pollack et al. 1997).

La phase de latence clinique, qui peut durer de 2 à 15 ans, se caractérise par une charge virale stable associée à une diminution lente mais inexorable du nombre de lymphocytes T CD4+.

En l'absence de traitement, cette phase aboutit au SIDA avec l'apparition d'infections opportunistes (bactériennes, virales, fongiques et parasitaires) et de cancers.

Certaines personnes sont résistantes à l'infection par le virus du VIH-1 : les homozygotes pour la mutation $\Delta 32$ du gène CCR5. Ils représentent 2 à 4 % des Caucasiens. Ces personnes seront naturellement résistantes si elles se contaminent par un virus de tropisme CCR5. Les hétérozygotes vont avoir une évolution plus lente de la maladie (Rugeles, Velilla, and Montoya 2011).

Certains patients infectés par le VIH sont définis comme « Elite Contrôleurs ». Ils sont infectés par le VIH, ont une sérologie positive mais contrôlent spontanément leur charge virale et n'ont pas d'inversion du rapport des CD4/CD8 avec une absence de diminution du nombre de lymphocytes T CD4+. Ces patients représentent environ 1 % du nombre de personnes infectés par le VIH. Même si les mécanismes de contrôle de l'infection chez ces patients ne sont pas élucidés, il a été montré qu'ils développent une réponse T CD8+ cytotoxique très efficace (Stephen A. Migueles and Connors 2012).

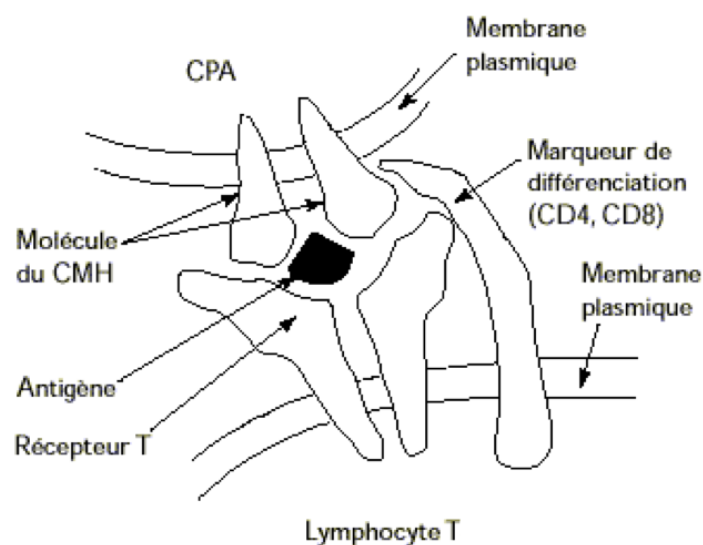
E. Réponses immunes contre le VIH

La réponse immune contre le VIH fait intervenir la réponse immune innée et adaptative. Au sein de la réponse adaptative, le VIH va entraîner une réponse humorale, avec génération d'anticorps spécifiques à l'infection et d'anticorps neutralisants, mais aussi une réponse cellulaire.

1. Réponse CTL et VIH

Les virus sont typiquement éliminés par une réponse lymphocytaire T CD8⁺ spécifique qui reconnaît des antigènes viraux. Ces antigènes proviennent de la segmentation des protéines virales digérées par le protéasome ou d'autres protéases dans le cytosol. Les fragments peptidiques sont ensuite transportés par des protéines de transport et apprêtées comme un complexe par les molécules HLA (Human Leukocyte Antigen) ou CMH (Complexe Majeur d'Histocompatibilité) de classe I à la surface d'une cellule infectée.

Les lymphocytes T CD8⁺ cytotoxiques vont ensuite reconnaître le HLA de classe I associé à l'épitope viral via leurs Récepteurs T Cellulaires T (TCR) (cf *Figure 8*). Cela va initier une cascade d'activation entraînant une libération de granzymes et perforines permettant la lyse de la cellule infectée qui peut survenir avant que des virions soient produits. De plus, l'activation du TCR entraîne la libération de diverses cytokines notamment l'IFN γ , le TNF α , les protéines inflammatoires des macrophages 1 α et β (MIP) et RANTES (CCL5) qui ont des effets antiviraux (Jones and Walker 2016).



CPA cellule présentatrice Ag
CMH = HLA
Antigène = épitope CTL présenté

Figure 8 : Schéma de la présentation d'un peptide au lymphocyte T par le HLA.

Les lymphocytes T CD8⁺ ont des capacités d'inhibition de la réplication du VIH démontrées *in vitro* à l'aide de tests fonctionnels immunologiques (Walker and McMichael 2012). La magnitude et la rapidité de l'activation de la réponse T CD8⁺ spécifique anti-VIH dans l'infection aiguë sont inversement corrélées à la charge virale basale lors de la phase chronique

indiquant que ces cellules assurent la pression antivirale durant le pic de virémie (Jones and Walker 2016).

Bien que la réponse T CD8+ spécifique du VIH soit puissante, le système immunitaire est incapable d'éliminer les cellules infectées. Cela est dû en partie à l'expression réduite ou absente d'antigènes viraux dans les cellules infectées quiescentes latentes, à l'échappement de reconnaissance du virus par les T CD8+, à la dysfonction des T CD8+, à la compartimentation des CD8+ et des réservoirs viraux ce qui limite l'efficacité de la réponse immune induite naturellement pour éliminer l'infection (Jones and Walker 2016). Au cours de la phase chronique, l'épuisement de la réponse CD8+ entraîne une dysfonction progressive des cellules T CD4+ et CD8+ spécifiques anti-VIH. Dans les premières étapes de l'épuisement, les cellules ont une prolifération altérée et produisent moins de cytokines différentes. Dans les étapes tardives, on assiste à une up régulation de multiples molécules activatrices et co-inhibitrices (Jones and Walker 2016).

2. Réponse CTL et variabilité génétique

Les molécules HLA sont codées par un ensemble de gènes présent sur le chromosome 6. Ce sont des glycoprotéines qui assurent la fonction de présentation de l'antigène et d'histocompatibilité. On distingue 3 classes de HLA et la classe I comprend les gènes du HLA A, HLA B et HLA C. Ces molécules sont présentes à la surface de toutes les cellules nucléées (Thomas J. Kindt, Richard A. Goldsby, Barbara A. Osborne 2008; Myriam Labalette, Siamak Bahram, Marie-Christine Béné 2011).

Le système HLA est très polymorphe avec une grande variabilité interindividuelle. Certains HLA sont plus répandus dans certaines populations. Maiers *et al.* ont ainsi étudié la fréquence des allèles HLA de classe I et de classe II dans la population en fonction de l'origine ethnique (Martin Maiers, Loren Gragert, William Klitz, n.d.; Maiers, Gragert, and Klitz 2007). Ainsi pour les allèles HLA A et HLA B, les 5 majoritaires dans chaque population (Caucasienne, Africaine, Asiatique ou Hispanique) sont répertoriés dans le Tableau 1.

Tableau 1 : Fréquences et rangs des 5 HLA A et B majoritaires dans les populations Caucasiennes (Cauc.), Africaines (Afric.), Asiatiques (Asiat.) et Hispaniques (Hisp.) d'après

Maiers *et al.* (Maiers, Gragert, and Klitz 2007; Martin Maiers, Loren Gragert, William Klitz, n.d.).

	Rang Cauc.	% Cauc.	Rang Afric.	% Afric.	Rang Asiat.	% Asiat.	Rang Hisp.	% Hisp.
HLA-A*02:01	1	29,60%	1	12,46%	3	9,46%	1	19,40%
HLA-A*01:01	2	17,18%	8	4,74%	5	5,08%	4	6,70%
HLA-A*03:01	3	14,35%	3	8,13%	11	2,60%	3	7,91%
HLA-A*24:02	4	8,69%	15	2,21%	1	18,24%	2	12,32%
HLA-A*11:01	5	5,64%	18	1,58%	2	17,90%	7	4,62%
HLA-A*23:01	12	1,68%	2	10,77%	27	0,23%	10	3,69%
HLA-A*30:01	13	1,34%	4	6,91%	12	2,06%	15	2,11%
HLA-A*68:02	16	0,85%	5	6,51%	46	0,03%	14	2,46%
HLA-A*33:03	22	0,13%	9	4,45%	4	9,43%	19	1,31%
HLA-A*31:01	10	2,35%	22	1,04%	9	3,25%	5	4,79%
HLA-B*07:02	1	13,99%	2	7,30%	15	2,63%	4	5,45%
HLA-B*08:01	2	12,53%	9	3,84%	21	1,64%	6	4,45%
HLA-B*44:02	3	9,01%	17	2,12%	32	0,76%	9	3,33%
HLA-B*15:01	4	6,65%	23	0,98%	11	3,48%	10	2,88%
HLA-B*35:01	5	5,71%	3	6,49%	5	4,27%	1	6,35%
HLA-B*53:01	32	0,32%	1	11,25%	66	0,09%	21	1,55%
HLA-B*15:03	39	0,09%	4	6,25%	88	0,03%	20	1,60%
HLA-B*42:01	NA	0,00%	5	5,46%	NA	0,00%	39	0,63%
HLA-B*40:01	6	5,64%	21	1,33%	1	7,98%	26	1,35%
HLA-B*51:01	9	4,54%	16	2,18%	2	6,28%	3	5,78%
HLA-B*46:01	NA	0,00%	NA	0,00%	3	6,11%	NA	0,00%
HLA-B*58:01	27	0,47%	11	3,51%	4	5,77%	23	1,45%
HLA-B*44:03	7	4,96%	6	5,37%	6	4,24%	2	6,08%
HLA-B*40:02	20	0,99%	39	0,35%	14	3,06%	5	4,85%

En raison du polymorphisme du CMH, et à partir d'une protéine donnée, chaque allèle HLA présente des peptides conformes à ses contraintes d'ancrage. Une même protéine sera donc reconnue différemment par les CTL de deux individus ayant un typage HLA différent. De même une modification, même minime, de la séquence de la protéine peut affecter la présentation par une même molécule HLA aux CTL.

Au cours de l'infection par le VIH, des associations ont été décrites entre certains allèles HLA de classe I et le degré de progression de la maladie. Certains HLA sont décrits comme protecteurs avec une évolution plus lente de la maladie vers le SIDA : les HLA B27 et B57, alors que d'autres sont impliqués dans une évolution rapide (HLA B35, B58) (Kaslow *et al.* 1996; S. A. Migueles *et al.* 2000; Gao *et al.* 2001; Kawashima *et al.* 2009; Rugeles, Velilla, and Montoya 2011; Goulder and Walker 2012).

3. Réponse CTL et variabilité virale

De nombreuses mutations aléatoires surviennent à chaque cycle de réplication virale du fait de la faible fidélité de la reverse transcriptase. Certaines de ces mutations sont sélectionnées pour leur capacité à échapper à la pression exercée par les CTL. Les mutations d'échappement peuvent modifier les régions qui flanquent les épitopes CTL ou directement la séquence de l'épitope altérant les sites de liaison entre peptide-HLA ou entre peptide-TCR. Celles-ci peuvent également affecter l'action du protéasome ou des transports intracellulaires avant apprêtement par le HLA ce qui va impacter la qualité de la réponse CTL (Lazaro et al. 2009).

En contrepartie, des études de réplication virale indiquent que la pression exercée par la réponse médiée par les T CD8⁺ induit des mutations signatures qui impactent la capacité répllicative (fitness) du virus, particulièrement dans les épitopes reconnus par les HLA protecteurs, suggérant un effet antiviral persistant même dans les situations d'échappement immunitaires (Kløverpris, Leslie, and Goulder 2015).

Durant la primo-infection, la pression de sélection par les CTL est rapide et le virus en échappement au système immunitaire va sélectionner des mutations tout en essayant de ne pas trop altérer son fitness. Durant la phase chronique, la réponse spécifique anti Gag est importante et corrélée à une charge virale basse par opposition à la réponse spécifique anti Env qui est associée à une charge virale haute. Des variations de Gag p24 sont notamment associées à un coût pour le fitness du virus. La réponse CTL contre ces épitopes est plus favorable car soit le virus est contrôlé par le système immunitaire soit il y a une mutation d'échappement mais le virus perd en capacité répllicative et donc en virulence (Kiepiela et al. 2007).

F. Eradication virale

1. Traitements

A l'heure actuelle, le traitement associe une combinaison de molécules antirétrovirales (ARV) ciblant plusieurs étapes du cycle de réplication du virus (cf *Figure 6*). Ces ARV ont amélioré la vie des personnes séropositives en diminuant voire en rendant indétectable leur charge virale plasmatique et en augmentant leur nombre de lymphocytes T CD4, limitant ainsi les infections opportunistes et améliorant leur qualité de vie (Ruelas and Greene 2013). Preuve de cette efficacité, en France, environ 80% des patients sous ARV sont en succès virologique défini par une charge virale < 50 copies/ml (“Rapport_Morlat_2013_Mise_en_ligne.Pdf” 2014).

Cependant ce traitement a ses limites : le virus est contrôlé mais pas éradiqué. Les risques de développer des mutations de résistance et les effets secondaires liés aux ARV persistent. De plus il ne faut pas négliger l'aspect économique d'un traitement au long cours (“Rapport_Morlat_2013_Mise_en_ligne.Pdf” 2014).

La suppression de la réplication VIH par les ARV induit une réduction ou une élimination de l'expression antigénique et altère la magnitude et l'étendue de la réponse T CD8+ effectrice spécifique anti-VIH qui pourrait lyser les cellules infectées par le virus. La réponse T CD8+ dans le sang périphérique diminue sous traitement et pour pouvoir parvenir à l'éradication virale, la réponse T CD8+ doit être activée et augmentée (Jones and Walker 2016).

Si la plupart des patients traités présentent une charge virale indétectable par les techniques de quantification de routine, il est admis qu'une réplication résiduelle subsiste de l'ordre d'une dizaine de copies par millilitre de sang (Palmer et al. 2003, 2008). Son origine n'est pas clairement définie mais selon Chun et ses collaborateurs, elle semble liée aux cellules latentes réservoirs et à une efficacité modérée des ARV dans certains sites anatomiques (Chun et al. 2011a) . Ceci a pour conséquence un rebond de la charge virale VIH dès l'arrêt du traitement.

2. Notion de réservoir

Le réservoir est défini comme une population de cellules quiescentes contenant de l'ADN proviral du VIH latent, potentiellement réactivable pour générer de nouveaux virus infectieux (Eisele and Siliciano 2012). Il est constitué dès le début de l'infection.

L'origine et le devenir de ces réservoirs sont décrits dans la *Figure 9* ci-dessous :

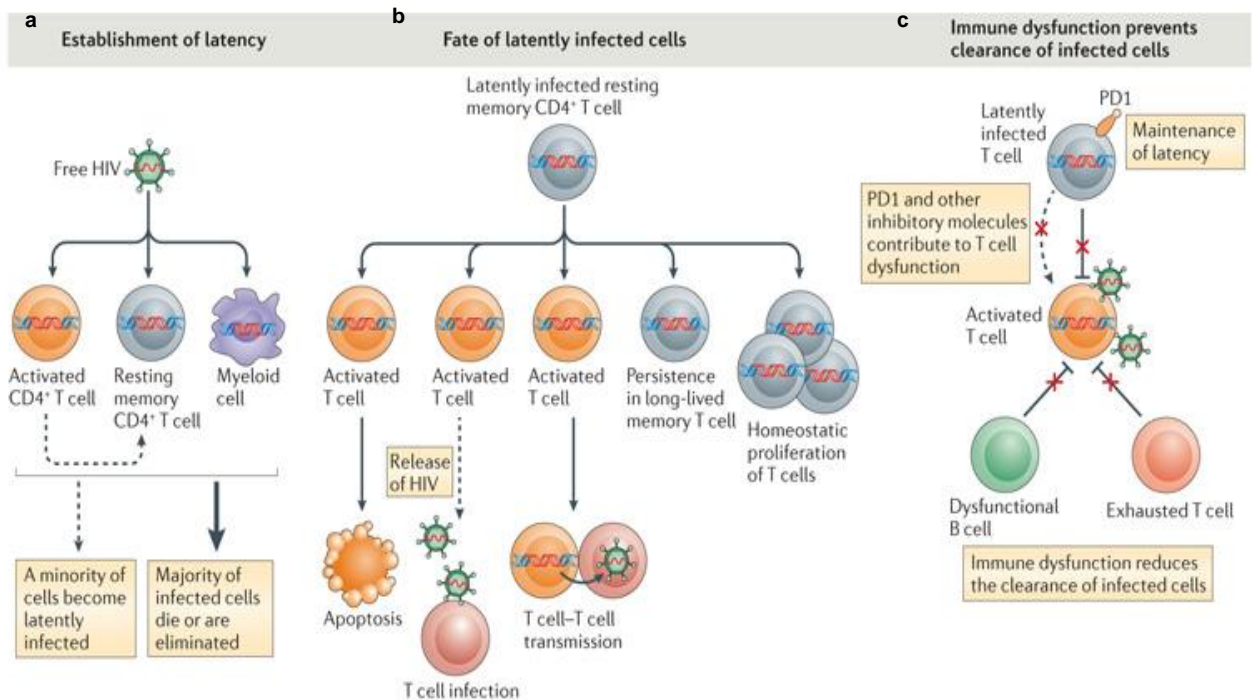


Figure 9 : Origine et devenir des réservoirs, d'après Deeks et al. (Deeks et al. 2012).

- (a) Parmi les cellules CD4⁺ du sang périphérique, la majorité est activée et permet une réplication virale avant d'être détruite. Une faible proportion de ces cellules devient quiescente et constitue le réservoir.
- (b) Ces cellules quiescentes peuvent soit rester dans cet état de latence des années et proliférer, soit être réactivées pour infecter d'autres lymphocytes.
- (c) Un dysfonctionnement du système immunitaire explique l'établissement et le maintien de la latence.

Ces cellules quiescentes peuvent soit rester dans cet état de latence pendant des années et se répliquer soit être réactivées pour infecter d'autres lymphocytes T (Chomont et al. 2009; Deeks et al. 2012). Différentes populations cellulaires constituent le réservoir cellulaire : les cellules T mémoires quiescentes majoritairement mais aussi les macrophages et les cellules folliculaires dendritiques. Si l'on considère que moins de 2% des lymphocytes de l'organisme circulent dans le sang périphérique, on peut supposer que les cellules réservoirs présentes dans d'autres sites anatomiques sont très importantes pour la persistance du VIH. Des réservoirs anatomiques, tels

que le tissu lymphoïde associé au tube digestif, GALT, les autres organes lymphoïdes, le tractus génital, le système nerveux central sont décrits (Eisele and Siliciano, 2012). L'étude de ces réservoirs est limitée soit par leur localisation qui nécessite des biopsies, difficiles à obtenir, soit par leur quantité puisqu'il faut plusieurs centaines de millilitres de sang pour purifier les lymphocytes T CD4+ quiescents à partir du sang périphérique.

Plusieurs études indiquent qu'il y aurait peu de diversité génétique entre les virus présents avant thérapie et les virus intégrés séquencés après plusieurs années de thérapies, que ce soit dans les cellules réservoirs du sang ou du GALT (Evering et al. 2012; Josefsson, Palmer, et al. 2013). Il a toutefois été montré, en étudiant les séquences des ARN viraux plasmatiques isolés chez des patients en succès thérapeutique, que si quelques séquences virales du plasma ressemblaient aux séquences virales obtenues à partir des cellules CD4+ quiescentes du sang, la population virale prédominante était clonale et présentait une origine différente des cellules CD4+ sanguines (Bailey et al. 2006). Il semble donc que la virémie résiduelle ne soit pas générée par une réplication faible persistante dans les cellules CD4 du sang mais par la production de virus soit par un autre type cellulaire du sang comme les macrophages, soit par des cellules CD4+ présentes dans d'autres sites anatomiques. Cette différence entre la séquence virale présente dans les lymphocytes CD4+ quiescents du sang et la séquence virale plasmatique constitue une barrière supplémentaire pour l'éradication du virus.

3. Stratégies d'éradication du VIH-1

En 2007 a été publié le cas du « patient de Berlin », premier et seul patient au monde à avoir guéri du VIH suite à une greffe de cellules souches hématopoïétiques. La perspective d'éradiquer le VIH est ainsi devenue une réalité (Chun and Fauci 2012). Ce « patient de Berlin » infecté par le VIH-1, a présenté une leucémie aigüe et a été allogreffé par des cellules souches hématopoïétiques d'un patient $\Delta 32$ homozygote pour le gène du CCR5. Le virus du VIH-1 qui infectait ce patient étant de tropisme CCR5, ces cellules sont devenues « résistantes » à l'infection et il est maintenant considéré guéri.

La communauté scientifique et médicale s'est tournée vers la recherche thématique sur l'éradication (cure) du HIV-1. Plusieurs voies ont été ou sont explorées (Deeks et al. 2016) : l'intensification thérapeutique, la réversion de l'état de latence (Latency Reversal Agents, LRA, comme l'histone désacétylase) forçant le provirus à s'exprimer avec une reprise de la réplication virale productive et une exposition aux lymphocytes TCD8 cytotoxiques (« Shock and kill »), les anticorps monoclonaux, les immuno-modulateurs (comme l'IL7 et l'inhibiteur PD-1) la thérapie génique (délivrance dans les cellules infectées de nucléases ZFNs, TALENs ou CRISPR/Cas9) et la vaccination thérapeutique .

4. Vaccination thérapeutique

Le but de la vaccination thérapeutique est de stimuler et augmenter la réponse CTL spécifique contre le VIH. Cela s'appuie sur l'hypothèse que le système immunitaire du patient va ainsi cibler et détruire efficacement les cellules productrices du virus. Cette stratégie pourrait permettre l'arrêt du traitement ARV car la réponse CTL serait suffisamment puissante pour bloquer la multiplication virale. Elle pourrait aussi être combinée à l'activation des cellules réservoirs (approche « kick and kill »). En effet, les patients « progresseurs » présentent une réponse CTL déficiente, en partie responsable de la chronicité de l'infection, qui n'est pas restaurée par un traitement ARV. Par opposition, les patients « Elites Contrôleurs » qui présentent un contrôle de la charge virale et un taux normal de CD4 sans traitement ARV, ont une réponse CTL très efficace (Pereyra et al. 2010; Stephen A. Migueles and Connors 2012). La stimulation de la réponse des lymphocytes CD8+ apparaît donc comme un élément clé pour détruire les cellules infectées. Cette hypothèse a été validée *in cellulo* par Shan et ses collaborateurs qui ont montré qu'une pré-stimulation antigénique des lymphocytes T CD8+ avant réactivation de la production virale est nécessaire à la destruction efficace des cellules infectées (Shan et al. 2012). Ceci a été confirmé récemment par l'équipe de Margolis (Sung et al. 2015) qui a montré qu'après pré-stimulation des cellules CD8+ avec un cocktail de 150 peptides, la réponse CTL pouvait cibler et éliminer les cellules réservoirs latentes réactivées par des LRAs.

Plusieurs stratégies de vaccination thérapeutiques ont été développées (revue générale par publiée par notre équipe (Fleury et al. 2018)) basées sur différentes constructions permettant entre autre, une réponse CD8+, une réponse ADCC mais aussi la synthèse d'Anticorps neutralisant.

Nous résumons brièvement les données de cette revue générale qui concerne plusieurs essais réalisés chez l'homme visant la stimulation de la réponse CTL.

Plusieurs équipes ont testés diverses stratégies vaccinales variant entre autres par le choix des antigènes ou le vecteur vaccinal utilisé pour délivrer les peptides :

- l'équipe de Pollard : 4 peptides issus de p24 Gag (Vacc-4X) chez des patients en succès thérapeutique (Pollard et al. 2014),

- l'équipe de Mothe : un vaccin MVA recombinant HIV-1 exprimant la gp120 ainsi qu'une polyprotéine Gag-Pol-Nef du HIV-1 de sous type B avec ou sans l'action d'un LRA (Disulfiram) (Mothe et al. 2015),
- l'équipe de Achenbach : un ADN HIV-1 (6 plasmides exprimant Gag, Pol, Nef et Env de sous types A, B C) associé à un Adénovirus recombinant (rAd5 contenant 4 vecteurs rAd5 exprimant les protéines de fusion Gag-Pol sous type B et les glycoprotéines Env des sous types A, B et C) (Achenbach et al. 2015)
- une équipe Danoise : Vacc-4X associé à GM-CSF et la Romidepsine (Leth et al. 2016),
- l'équipe française de Loret : une préparation de protéine Tat (Loret et al. 2016)

Ou à base de cellules dendritiques :

- des cellules dendritiques autologues pulsées par du virus HIV-1 autologue inactivé à des patients sous traitement antirétroviral (Garcia et al. 2013),
- des cellules dendritiques générées *ex vivo* et chargées en lipopeptides génériques (Levy et al. 2014),
- des cellules dendritiques chargées avec des ARNm viraux autologues exprimant Gag, Rev, Vpr et Nef (Jacobson et al. 2016).

Dans la plupart de ces études, il y a une amélioration et augmentation de la réponse cellulaire associée parfois à une diminution de la charge virale sans réussir à contrôler la réplication virale. Le manque d'efficacité de ces vaccins vient probablement du fait qu'ils sont « génériques » et qu'ils ne prennent pas en compte la variabilité du VIH inter- mais aussi intra- patient (McMichael and Koff 2014). En effet, pour que la réponse CTL soit optimale et élimine toutes les cellules infectées, il faut que les antigènes viraux présentés par toutes les cellules soient reconnus. Pour ce faire, ces antigènes ne doivent pas présenter une trop grande variabilité car une différence d'un acide aminé sur un épitope a 30-50 % de chance de diminuer la réponse CTL. Ceci reste à démontrer car les séquences des provirus intégrés peuvent différer entre les différents réservoirs, donnant des épitopes plus ou moins bien reconnus par les lymphocytes CD8+ stimulés et laissant donc la possibilité au virus d'échapper au système immunitaire. De plus, l'efficacité de présentation d'un même épitope peut différer d'un individu à un autre, puisqu'elle est régie par le HLA du patient.

Les précédentes études menées au sein du laboratoire dans le cadre du projet Provir débuté en 2012 ont montré que les épitopes CTL codés par l'ADN proviral chez les patients VIH en succès virologique avaient des séquences variables et différentes de la séquence consensus HxB2, référence utilisée pour concevoir les peptides immunogènes apportés par les vaccins dit

« génériques », expliquant certainement le manque d'efficacité des essais vaccinaux (Papuchon et al. 2013, 2014). De plus, les séquences des épitopes sont variables au cours du temps au sein d'un même individu. Et comme attendu, la présentation des différents épitopes est différente selon les individus, du fait des allèles HLA (Papuchon et al. 2013, 2014).

Un travail très approfondi allant dans le même sens a été publié par l'équipe de R Siliciano (Deng et al. 2015). Elle a d'ailleurs montré que plus le délai entre l'infection et l'initiation du traitement est long, moins la réponse T CD8⁺ autologue est en capacité de reconnaître les épitopes archivés dans l'ADN proviral. Bien que les mutants d'échappements aux épitopes communs CD8⁺ soit relativement rares chez les individus traités au cours de la primo-infection, chez les patients traités durant la phase chronique, la majorité des provirus présente des mutations d'échappement aux épitopes dominants ce qui va les rendre non reconnaissables par les cellules T CD8⁺. Toutefois, une réponse T CD8⁺ dominante ciblant des épitopes non porteurs de mutations d'échappement a été identifiée chez chaque sujet testé et des études *in vitro* ont permis de confirmer l'élimination des cellules infectées avec le VIH autologue grâce à l'expansion des lymphocytes T CD8⁺. Donc chez les sujets traités en phase chroniques, les stratégies pour augmenter spécifiquement les réponses T CD8⁺ contre des épitopes non « échappeurs » pourraient être une des clés pour aboutir à l'éradication virale.

**PROJET PROVIR ET
OBJECTIF DE LA
THESE**

II. Projet Provir et objectif de la thèse

A. Projet Provir/Latitudo 45

Le projet Provir/Latitudo 45, initié par H. Fleury avec autorisation du CPP « Sud-Ouest et Outremer III » en 2012 avec le soutien financier de MSD puis MSDAvenir, est un projet de recherche à visée vaccinale. Il concerne des patients infectés par le VIH-1 en succès virologique depuis l'initiation du traitement antirétroviral. La finalité est d'identifier des épitopes conservés en termes de présentation au système HLA pouvant induire une réponse CTL soit pour l'ensemble de la population soit par groupe de population afin de servir de base à un vaccin thérapeutique pour un grand groupe de patients.

Les patients sont recrutés à Bordeaux, au Canada et au Brésil afin de prendre en compte un maximum de variabilité virale (différents sous-types) et immunogénétique (HLA différents selon les populations Caucasiennes, Africaines, Asiatiques ou Hispaniques). L'extension au Canada a généré de la part des collègues Canadiens l'appellation Provir/Latitudo 45, Montréal étant à la même latitude que Bordeaux.

Pour cela, on étudie la variabilité interindividuelle pour la prédiction de présentation des épitopes CTL issus des protéines Gag, Pol et Nef du VIH-1 par le HLA de classe I (allèles A et B uniquement). Cette prédiction est réalisée au moyen d'outils de bio-informatique (IEDB, Immune Epitope DataBase) en combinant les séquences virales et les allèles HLA.

De plus, avec les patients inclus au Canada, nous allons analyser la variabilité intra-individuelle en étudiant les séquences des épitopes VIH codés par le provirus au niveau des PBMC et du GALT (Gut Associated Lymphoid Tissue) et leur présentation aux molécules HLA.

Pour la plupart des patients inclus, la variabilité virale est analysée par séquençage populationnel (Sanger) mais aussi par un séquençage haut débit (NGS).

Toutes les données obtenues sont ensuite compilées et analysées afin d'identifier des épitopes viraux conservant une bonne affinité pour la poche peptidique d'un HLA donné. Les épitopes sélectionnés seront testés par des tests d'immunologie fonctionnelle afin de valider leur capacité à induire une réponse CTL. Le cocktail vaccinal prévisionnel sera testé sur macaques selon un programme vaccinal calqué sur les essais publiés chez l'homme et la réponse CTL des macaques sera évaluée au bout de 6 mois.

B. Projet de thèse

Ce travail de thèse se concentre sur la cohorte de patients VIH-1 en succès virologique depuis l'initiation de traitement, recrutés au CHU de Bordeaux dans le cadre du projet Provir/Latitude 45. L'objectif principal est d'identifier des épitopes induisant une réponse T cytotoxique en fonction de la variabilité virale du VIH-1 et de l'immunogénétique des patients.

En premier lieu, nous avons étudié la répartition des différents sous-types viraux du VIH-1 chez les patients suivis en Aquitaine et pour lesquels des données génotypiques étaient disponibles au laboratoire. Au sein de cette population, des patients remplissant les critères d'inclusion ont été recrutés dans le projet Provir/Latitude 45.

Dans un deuxième temps, nous avons réalisé le séquençage populationnel de certains gènes de l'ADN proviral du VIH-1 extrait à partir de sang total périphérique. En parallèle le typage HLA a été défini pour chacun des patients. Les données génétiques individuelles et virales ainsi obtenues, ont été combinées au moyen d'IEDB permettant la prédiction de la présentation des épitopes CTL issus des protéines virales par le système HLA. Etant donné la multiplicité des paramètres (séquences virales et typage HLA) pour un même patient mais aussi pour l'ensemble de la population étudiée, un logiciel, TutuGenetics, a été développé pour automatiser cette partie de l'analyse. Ce logiciel instrumentalise des algorithmes d'IEDB en prenant les alignements nucléotidiques des séquences virales et les typages HLA des patients. Ainsi les potentiels épitopes ont été :

- (i) Analysés à partir de la liste des « Best Defined CTL epitope list » disponible sur Los Alamos HIV Database (Llano et al. 2013).
- (ii) Obtenus par découpage séquentiel des séquences virales par pas de 8 à 10 acides aminés.

Toutes les données obtenues sont ensuite compilées et analysées par patients, par HLA afin d'identifier des épitopes viraux conservés induisant une potentielle réponse CTL.

1^{ERE} PARTIE

III. 1^{ère} partie - Population de l'étude

A. Patients Vivants avec le VIH dans la région Aquitaine

1. Publication 1

Diversity of HIV-1 in Aquitaine, South-Western France, 2012-2016

Tumiotto C, Bellecave P, Recordon-Pinson P, Groppi A, Nikolski M, Fleury H.
AIDS Res Hum Retroviruses. 2018 May; 34(5):471-473.

Une cartographie de la répartition des sous-types viraux du VIH-1 présents dans la région Aquitaine a été réalisée afin de vérifier la représentativité de l'échantillonnage des patients inclus dans le projet Provir/Latitude 45.

Dans le cadre de la prise en charge clinique des patients infectés par le VIH-1, des génotypages de résistance aux antirétroviraux sont effectués au laboratoire de Virologie du CHU de Bordeaux pour les différents hôpitaux de la région Aquitaine.

Les génotypages VIH peuvent être réalisés sur l'ARN plasmatique, à l'initiation du traitement antirétroviral ou au cours d'un échec virologique, ou sur l'ADN proviral à partir de sang total périphérique lorsqu'un changement thérapeutique est envisagé et que le patient est en succès virologique.

Les séquences obtenues pour les gènes codant pour la Protease, Reverse Transcriptase et Intégrase entre janvier 2012 et décembre 2016, ont ainsi été collectées et analysées par 3 ou 4 méthodes phylogénétiques afin de définir le sous-type viral.

Entre 2012 et 2016, 3204 génotypages de résistance VIH-1 ont été effectués, 1784 sur ARN plasmatique et 1420 sur ADN proviral.

D'un point de vue global, plus de 71 % des patients sont infectés par un virus de sous type B. Le CFR02_AG qui infecte plus de 11 % des patients est majoritaire parmi les sous-types viraux non B.

Si on considère les génotypages effectués sur ADN proviral donc chez des patients en succès virologique, on retrouve 79 % d'infections par un sous-type B et 8,2 % par CRF02_AG.

DIVERSITY OF HIV-1 IN PATIENTS FROM AQUITAINE, SOUTH WESTERN FRANCE, 2012-2017

Camille Tumiotto^{1,2}, Pantxika Bellecave^{1,2}, Patricia Recordon-Pinson², Alexis Groppi³, Macha Nikolski³ and Hervé Fleury^{1,2}

1 University Hospital of Bordeaux, Bordeaux2 CNRS UMR 5234 MFP - University of Bordeaux

3 Bordeaux Bioinformatics Center (CBiB)-University of Bordeaux, France

Abstract

We have estimated the prevalence of the different viral subtypes between January 2012 and December 2016 in HIV-1-infected patients of the Aquitaine region (south-west part of France) who had a routine HIV-1 genotype resistance testing (GRT) centralized at the Bordeaux University Hospital. GRT was performed on viral RNA (1784 samples) before treatment initiation or at failure whereas proviral DNA was used as template (1420 samples) in the event of a treatment switch in patients with viral load below 50 copies/mL. Pol and integrase sequences were obtained; subtypes, circulating recombinant forms (CRF) and unique recombinant forms (URF) were assigned by combining the results of SCUEAL, REGA, COMET and HIV BLAST. Globally, subtype B was predominant with 71.7% while non-B subtypes accounted for 28.3%. Within the non-B viruses, CRF02_AG was the most prominent (11.6%) followed by non-B non-URF(13.5%), A, CRF01_AE, G, CRF06_cpx, F, C, D, H, J and finally URF (3.2%). The analysis of the two compartments separately showed that RNA exhibits higher percentages of non-B viruses than DNA. This study reveals a high degree of diversity of HIV-1 non-B subtype strains in Aquitaine, with an increasing prevalence of CRF02_AG and URF in the population investigated for viral RNA *id est* including more recently detected HIV-1 infected patients. Future studies should attempt to identify the transmission clusters while paying special attention to URF, since they seem to be increasing in the population and could potentially host CRF.

Key words : HIV-1, subtypes, circulating recombinant forms (CRF), unique recombinant forms (URF), non-B viruses, South Western France

Introduction

HIV infection is a worldwide health threat with more than 36 million people living with HIV, one million having died of AIDS in 2016 and more than 19 million using antiretroviral therapy (UNAIDS report 2017). The pandemic is associated with one of the two HIV viruses: HIV-1. The virus originates from zoonotic transmission in Central Africa. Since it exhibits extensive variability due to the absence of proofreading of the transcriptase reverse plus a high potential for recombination, HIV-1 has been classified into four groups: M, O, N and P. The M group is responsible for the pandemic and is composed of subtypes (A to K) and circulating recombinant forms (CRF) (90 have been recorded so far) (HIV Databases, Los Alamos, USA). While all viruses can be detected in Central Africa, some subtypes or CRF have migrated toward other continents. For example, subtype B is predominant in Europe, the Americas and Australia, subtype C in India and South Africa and CRF01_AE in South-East Asia. CRF02_AG is the major strain circulating in West Africa within a background of high diversity (WHO UNAIDS report 2011). The SPREAD study¹ has shown that in countries of Western Europe (France was not considered in this study), subtype B is the major strain followed by A1, C and CRF02_AG. However, these trends mask important differences, subtype B being less prevalent in the southern part of Europe. In France, subtype B is predominant followed by CRF02_AG². Overall, the percentage of non-B viruses is rising up mainly with the increase of immigration from the south. In Aquitaine, the south-west part of France, an ancient and a more recent study^{3,4} have demonstrated the prevalence of subtype B followed by CRF02_AG, but there is a background of high diversity and the percentage of non-B infections in situ is increasing⁴.

There are several reasons for studying the molecular epidemiology of HIV-1 circulating in different areas, but the main ones are to assess the dynamics of introduction and transmission of non-B viruses and to be able to identify any technical gaps in the commercial assays devoted to antibody-antigen detection and viral load quantitation.

In the present study, we analyzed our routine production of HIV-1 sequences which is carried out for all hospitals in the Aquitaine region (now a part of “Nouvelle Aquitaine”) and including 3204 different non-redundant patients from January 2012 to December 2016.

Methods

Patients

HIV pol and integrase sequences obtained between January 2012 and December 2016 were generated in the context of routine HIV-1 genotype resistance testing (GRT) of patients followed in Aquitaine, France (9 hospital centers: Saint André, Pellegrin, Haut-Leveque, Bayonne, Dax, Orthez, Arcachon, Libourne, Mont de Marsan). GRT was performed on viral RNA before treatment initiation or at failure whereas proviral DNA was used as template in the event of a treatment switch in patients with viral load below 50 copies/mL. The global prevalence of the different viral subtypes and CRF was determined by removing duplicate sequences for a given patient.

GRT

GRT was performed from plasma viral RNA (1784 samples) or proviral DNA (1420 samples) according to the French ANRS procedures. Briefly, viral RNA was extracted by using the High Pure Viral RNA kit (Roche). Five microliters were used for reverse transcription followed by a nested PCR. Proviral DNA was amplified by two successive PCR by using 5 microliters of total DNA extracted on TECAN (Biorad) from whole blood. Sequencing was done on ABI 3500xL and sequences were analyzed and stored in the SmartGene database (SmartGene, Lausanne, Switzerland).

Sequences

Nucleotide sequences of protease, reverse transcriptase and integrase were extracted from the Smartgene database as fasta files. For each gene, sequences were identified as patient ID_sample ID_date of analysis_template nature (RNA or DNA).

Subtyping

Subtypes were assigned using SCUEAL⁵, REGA⁶ and COMET⁷. For a given gene, the subtype was directly assigned if the three analyses were concordant. In the event of discrepancy, sequences were submitted to BLAST in order to conclude. Unique Recombinant Form (URF) was concluded in the event of a difference between the subtyped genes.

Results

We retrieved a total of 3204 GRT carried out between 2012 and 2016 in different contexts: before ART initiation in naïve patients, at virologic failure or before a switch of ART when no historical GRT was available. The global prevalence of the different

HIV-1 subtypes was determined by removing duplicate sequences for a given patient. Therefore, we analyzed 2409 sequences. The results are presented in Figure 1a. Subtype B was predominant with 71.7% while non-B subtypes accounted for 28.3%. Within the non-B viruses, CRF02_AG was the most prominent (11.6%) followed by non-B non-URF (13.5%), A, CRF01_AE, G, CRF06_cpx, F, C, D, H, J and finally URF (3.2%). In our presentation here, the non-B non-URF denomination encompasses the HIV-1 group M subtypes and CRF described in the Los Alamos database, except for B and CRF02_AG.

We then estimated the prevalence of the different subtypes according to the compartment analyzed (viral RNA or proviral DNA). They are presented in Figure 1b (RNA) and 1c (DNA). The percentage of subtype B HIV-1 infected patients who had a GRT from viral RNA is lower than from proviral DNA (64.3% vs 79%) whereas we observed an opposite trend for CRF02_AG infected patients (15% vs 8.2%). URF viruses accounted for 4% of GRT from RNA and 2.4% from DNA while non-B non URF accounted for 16.7% and 10.4% respectively.

The Figure 1d shows the evolution of the respective parts of B and non-B HIV-1 viruses between 2012 and 2016 according to the RNA and DNA compartments. There was a decrease of B subtype viruses genotyped from the RNA compartment while the non-B viruses increased over time. Regarding the viruses identified by genotyped proviral DNA, the proportion of these two groups of HIV-1 subtypes remains constant with a higher prevalence of subtype B.

Discussion

These data are in accordance with data from other parts of Western Europe with a largely higher prevalence of subtype B versus non-B viruses. There is a high diversity within the non-B viruses as previously described, although CRF02_AG is the most common. France and especially the southwestern part of the country have long-standing relationships with the French-speaking countries of Sub-Saharan Africa and most of the non-B viruses have already been described in West and Central Africa. An increase in the prevalence of non-B viruses has been observed in France in the last 15 years as reported by national collaborative studies^{8,9}. Although the prevalence of non-B viruses in Aquitaine seems lower than their global prevalence in France (~30% vs ~45%), the amount of non-B viruses has increased in this area compared to the 16.4% prevalence described in recent seroconverters during the 1996-2006 period¹⁰.

We have also differentiated the data from RNA and DNA compartments. GRT can be performed on proviral DNA when no previous sequence data are available, as mentioned by French recommendations. This gave us supplementary data for this epidemiological study, allowing us to evaluate more precisely the global distribution of HIV-1 subtypes and CRF in our area. With regard to putative circulating viruses, GRT performed with viral RNA is more informative since it includes newly diagnosed patients. The proportion of non-B viruses (mainly CRF02_AG) has been higher in the RNA compartment than in the DNA compartment since 2012. We hypothesize that recently detected HIV-positive patients who have been investigated for transmitted resistance before initiation of ART are at the origin of the increase in the amount of non-B viruses within this compartment, while DNA was investigated in patients who have received long-term treatment and whose therapeutic initiation was established at a moment when the percentage of B viruses was higher. Finally, future studies should attempt to identify the transmission clusters while paying special attention to URF, since they seem to be increasing in the population and could potentially host CRF.

Acknowledgments

We thank Ray Cooke for revising the manuscript.

This work is part of the Provir/Latitude 45 project funded by MSD Avenir (DS-2016-0005), GERMATAN and DRCI (Direction de la Recherche Clinique et de l'Innovation) at the University Hospital of Bordeaux.

The different virologic and immunologic parameters of the patients are stored by the Unite de Surveillance Biologique (USB) at the University Hospital of Bordeaux under the terms of agreement 106897 of the Commission Nationale Informatique et Liberté (CNIL).

Author Disclosure Statement

No competing financial interests exist

References

1. Abecasis AB, Wensing AMJ, Paraskevis D, Vercauteren J, Theys K, Van de Vijver DAMC, et al. HIV-1 subtype distribution and its demographic determinants in

newly diagnosed patients in Europe suggest highly compartmentalized epidemics. *Retrovirology*. 2013 Jan 14;10:7.

2. Brand D, Moreau A, Cazein F, Lot F, Pillonel J, Brunet S, et al. Characteristics of patients recently infected with HIV-1 non-B subtypes in France: a nested study within the mandatory notification system for new HIV diagnoses. *J Clin Microbiol*. 2014 Nov;52(11):4010–6.

3. Fleury H, Recordon-Pinson P, Caumont A, Faure M, Roques P, Plantier J-C, et al. HIV type 1 diversity in France, 1999-2001: molecular characterization of non-B HIV type 1 subtypes and potential impact on susceptibility to antiretroviral drugs. *AIDS Res Hum Retroviruses*. 2003 Jan 1;19(1):41–7.

4. Javaugue F-C, Recordon-Pinson P, Decoin M, Masquelier B, Cazanave C, Neau D, et al. Molecular characterization of non-B HIV type 1 isolates from patients of a department of infectious diseases, University Hospital of Bordeaux, France,. *AIDS Res Hum Retroviruses*. 2012 Sep;28(9):1124–30.

5. Kosakovsky Pond SL, Posada D, Stawiski E, Chappey C, Poon AFY, Hughes G, et al. An evolutionary model-based algorithm for accurate phylogenetic breakpoint mapping and subtype prediction in HIV-1. *PLoS Comput Biol*. 2009 Nov;5(11):e1000581.

6. Alcantara LCJ, Cassol S, Libin P, Deforche K, Pybus OG, Van Ranst M, et al. A standardized framework for accurate, high-throughput genotyping of recombinant and non-recombinant viral sequences. *Nucleic Acids Res*. 2009 Jul;37(Web Server issue):W634-642.

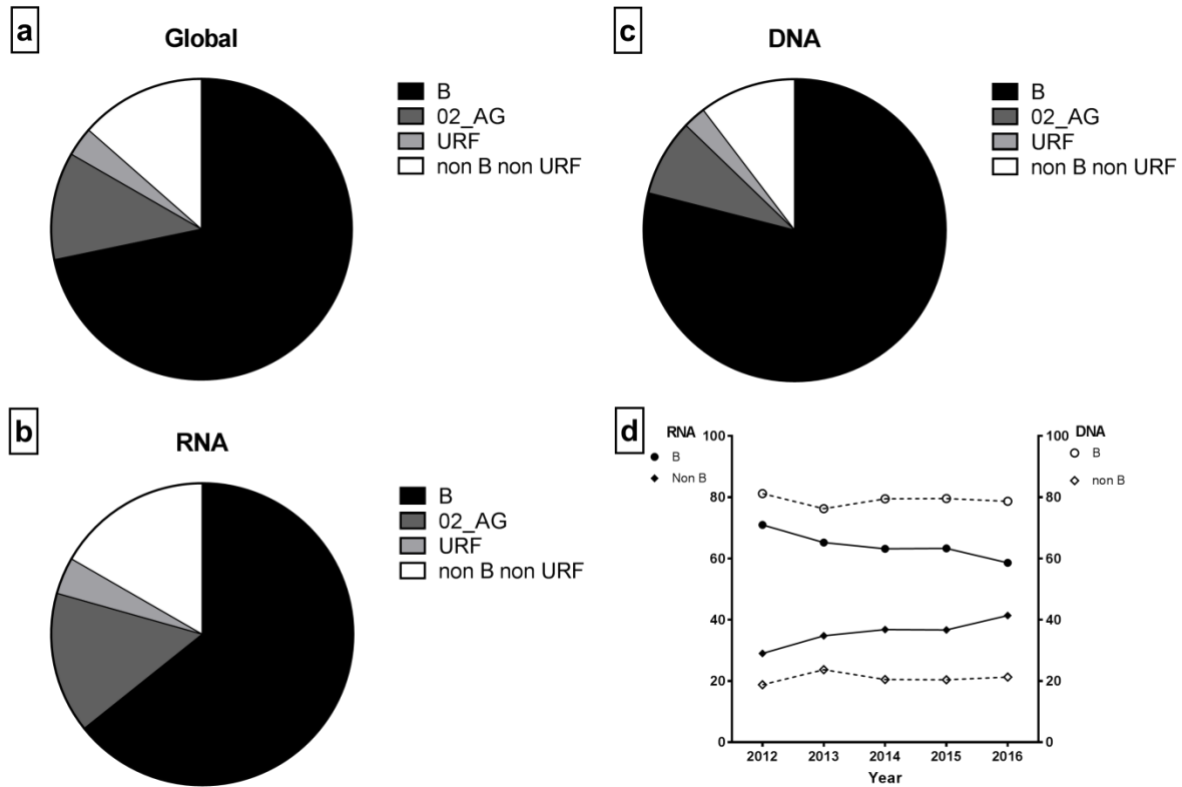
7. de Oliveira T, Deforche K, Cassol S, Salminen M, Paraskevis D, Seebregts C, et al. An automated genotyping system for analysis of HIV-1 and other microbial sequences. *Bioinforma Oxf Engl*. 2005 Oct 1;21(19):3797–800.

8. Frange P, Assoumou L, Descamps D, Cheret A, Goujard C, Tran L, et al. HIV-1 subtype B-infected MSM may have driven the spread of transmitted resistant strains in France in 2007-12: impact on susceptibility to first-line strategies. *J Antimicrob Chemother*. 2015 Jul;70(7):2084–9.

9. Descamps D, Assoumou L, Chaix M-L, Chaillon A, Pakianather S, de Rougemont A, et al. National sentinel surveillance of transmitted drug resistance in antiretroviral-naïve chronically HIV-infected patients in France over a decade: *J Antimicrob Chemother*. 2013 Nov;68(11):2626–31.

10. Recordon-Pinson P, Anies G, Bruyand M, Neau D, Morlat P, Pellegrin J-L, et al. HIV type-1 transmission dynamics in recent seroconverters: relationship with transmission of drug resistance and viral diversity. *Antivir Ther.* 2009;14(4):551–6.

Figure 1



2. Publication 2

A new HIV-1 circulating recombinant form (CRF98_cpx) between CRF06_cpx and subtype B identified in Southwestern France

P. Recordon-Pinson, B. Alves, **C. Tumioto**, P. Bellecave, F. Bonnet, D. Neau,
E. Soares, M. Soares, H. Fleury

Sous presse dans AIDS Research and Human Retroviruses

Lors de la réalisation de la première publication, où nous avons déterminé le sous-type des virus infectant les patients suivis au CHU de Bordeaux ou dans les CH de la région Aquitaine, nous avons identifié un cluster de patients présentant un virus recombinant, non référencé sur la database de Los Alamos.

Nous avons donc approfondi la caractérisation de ce nouveau virus composé de CRF_06cpx et de sous-type B. Pour ce faire un séquençage « near full genome » a été réalisé en collaboration avec l'équipe Brésilienne rattachée au projet Provir/Latitude 45. L'ADN proviral de 4 patients a été amplifié en 4 parties et soumis à un séquençage par Illumina. La reconstitution du consensus combinée à une analyse bio-informatique a permis d'établir la cartographie du génome de ce nouveau virus, référencé comme le CRF98_cpx.

A new HIV-1 circulating recombinant form (CRF98_cpx) between CRF06_cpx and subtype B identified in Southwestern France

Patricia Recordon-Pinson¹, Brunna M. Alves², Camille Tumiotto^{1,3}, Pantxika Bellecave³, Fabrice Bonnet⁴, Didier Neau⁵, Esmeralda A. Soares², Marcelo A. Soares² and Hervé Fleury^{1,3}

¹ CNRS MFP-UMR 5234, Bordeaux, France

² Instituto Nacional de Câncer, Rio de Janeiro 20231-050, Brazil.

³ Department of Biology and Pathology, University Hospital of Bordeaux, Bordeaux, France

⁴ Department of Internal Medicine, University Hospital of Bordeaux, Bordeaux, France

⁵ Department of Infectious and Tropical Diseases, University Hospital of Bordeaux, Bordeaux, France

Running Title : HIV-1, New CRF, France

KEYWORDS: HIV-1, Virus Evolution /Diversity, Epidemiology, Phylogenetics

Corresponding author:

Hervé Fleury

Laboratoire de Virologie, université de Bordeaux, 146 rue Léo Saignat, 33076

Bordeaux,

France

33 5 57571363

herve.fleury@u-bordeaux.fr

Abstract

During a recent study on the sequencing data of our database between 2012 and 2016 in Southwestern France, we observed that 8 patients harbored what seemed to be the same virus. Indeed, routine genotyping at the time of HIV diagnosis showed that protease (Prot) and reverse transcriptase (RT) were related to CRF06_cpx and subtype B, respectively. The integrase (INT) sequences (available for 3 patients) were clustering with CRF06_cpx and envelope (Env) gp120 sequences (available for 2

patients) with subtype B. Since such a recombinant has not been recorded in the Los Alamos database, we decided to characterize the full-length genome of this virus. The data suggest the identification of a new CRF between CRF06_cpx and subtype B, the structure of which is very complex with multiple breakpoints. We will refer this CRF as CRF98_cpx.

Introduction

HIV-1 can be classified in groups (M, N, O, P). The M group is associated with the pandemic and is itself subdivided into subtypes (A to K). Since the virus has a high potential for recombination, the subtypes can recombine and yield recombinant circulating forms (CRFs) or unique recombinant forms (URFs). The molecular identification and distribution of HIV-1 subtypes and CRFs in the world is an important issue with potential implications for disease progression, antiretroviral therapy and vaccine developments¹. To date, 96 CRFs have been recorded in the Los Alamos HIV database (<https://www.hiv.lanl.gov/>). To qualify as a CRF, these recombinant viruses must be characterized molecularly in three or more unrelated persons to attest to their epidemic potential, which is not the case for URFs. While most CRFs have been found in Africa and mainly central Africa, new recombinations may occur on other continents and lead to new CRFs. Such locations are China, South-East Asia, South America (particularly Brazil² and Argentina), the Caribbean isles (Cuba) and other parts of the world including Europe (Russia^{3, 4, 5} Portugal and/or Spain^{6, 7, 8} Italy⁹, UK¹⁰ and Luxembourg¹¹). In France one CRF has been recorded (CRF56_cpx)¹² and a second one is currently being recorded¹³.

During a recent study¹⁴ on the sequencing data of our database between 2012 and 2016, we observed that 8 patients harbored what seemed to be the same virus. Routine genotyping at the time of HIV diagnosis showed that protease (Prot) and reverse transcriptase (RT) were related to CRF06_cpx and subtype B, respectively. The integrase (INT) sequences (available for 3 patients) were clustering with CRF06_cpx and envelope (Env) gp120 sequences (available for 2 patients) with subtype B. Since such a recombinant has not been recorded in the Los Alamos database, we decided to characterize the full-length genome of this virus. We present the data which suggest the identification of a new CRF between CRF06_cpx and subtype B in patients from Southwestern France.

Material and methods

Samples

Samples were available for 4 patients, all followed for their HIV infection at the University Hospital of Bordeaux, receiving a successful cART with a viral load below the threshold (40 copies or 20 copies/mL) and being investigated for therapeutic follow-up. Total blood on EDTA was used for DNA extraction. DNA was analyzed by full-length sequencing of HIV-1.

Phylogenetic analysis of sequences from routine HIV-1 genotype resistance testing

Prot, RT, INT and gp120 V3 sequences obtained from routine HIV-1 genotype resistance testing were used for Neighbor-Joining analysis as previously described.¹⁵

Viral DNA isolation, polymerase chain reaction and sequencing

Viral DNA was extracted with an automated procedure Tecan (Biorad) according to the manufacturer's instructions and sequences were obtained for four samples using an amplification strategy with the Platinum TAQ High Fidelity kit (Thermofisher) based on a nested PCR with four overlapping amplicons as previously described^{16, 17}. These PCR products were pooled for Illumina library preparation including a unique bar code for each sample, and sequenced using Illumina MiSeq 250-bp paired-end technology at the National Institute of Cancer, Rio de Janeiro, Brazil. The consensus sequence for each sample was determined with Geneious software (reference or company).

HIV-1 subtype determination and recombinant identification

The subtype assignments and recombination patterns were determined by phylogenetic analysis using ClustalW2.1 and Simplot version 3.5.1¹⁸. We analyzed the recombination by bootscanning in SimPlot 3.5.1 with a sliding window of 200 bp, an increment step of 20 bp and the Kimura two-parameter model. The resulting breakpoint intervals were then mapped with the LANL recombinant HIV-1 drawing tool (http://www.hiv.lanl.gov/content/sequence/DRAW_CRF/recom_mapper.html). The sequences have been sent to the Los Alamos database under access numbers MH479262 to MH479284.

Results

All patients were MSM with an estimated date of HIV primary infection between 2007 and 2014 (Table 1); the median age was 31 years (range 22-53). HIV diagnosis was made at primary infection time (patients E and G) or less than 18 months after HIV acquisition (determined according to the date of the last negative serological assay). The median viral load at the time of diagnosis was 4.8 log₁₀ copies/mL (range 2.4-5.4) and the median CD4+T cell count was 419/uL (range 351-658). The time between detection of HIV infection and initiation of cART ranged from 2 weeks (patient D) to 5.5 years (patient B). Two patients (F and G) were lost to follow-up after initiation of therapy.

One isolate, patient C, could be analyzed over the whole length while the other three were analyzed partially: sample A with a gap between 5000 and 6000 bp, samples B and D with sequences available from origin down to 5000 bp. Comparison of the bootscanning analysis for the four samples in the regions obtained by NGS sequences exhibited the same recombinant structure (same composition and same breakpoints). The bootscanning analysis of sample C plus the mapping of the recombination points are presented in Figures 1A and B while Table 2 lists the different positions (relative to HxB2) of the genomic segments. The virus is a recombinant between CRF06_cpx and subtype B with multiple breakpoints. Among the gag gene which belongs to CRF06_cpx, 100bp encompassing P1 have similarity with subtype B (2082-2181 of HxB2). The pol gene is composed of both subtypes with several recombination points: the Prot gene is fully of CRF06_cpx origin; the RT is mostly of subtype B and INT mostly of CRF06_cpx origin; both will cluster respectively with subtype B and CRF06_cpx in the phylogenetic trees (see below).

The Prot, RT, INT and Env gp120 parts of the C sample have been introduced into phylogenetic trees together with sequences from other samples when available. The trees are presented in Figure 1 C (C1 to C4). The origins of the different parts of the viruses are as follows: Prot:CRF06_cpx; RT: B; INT: CRF06_cpx; Env gp120: B. These data are in accordance with what we had observed during the routine follow-up of the patients and which had drawn attention to the identity of these viral strains. There was some inaccuracy by SimPlot at a position between 700 and 921 bp. The corresponding sequence was used for a BLAST with references from the Los Alamos database and clearly yielded a CRF06_cpx identity.

Discussion

We describe herein a new HIV-1 circulating recombinant form CRF98_cpx between CRF06_cpx and subtype B which seems to be observed in some patients in Southwestern France. This CRF is characterized by a high number of breakpoints. While subtype B is predominant in France, CRF06_cpx circulates mainly in countries in West Africa.

Interestingly, our group has also described in Algeria an epidemic switch from subtype B to CRF06_cpx¹⁹, so we were concerned by the entrance of CRF06_cpx in France in view of the strong relationships and population mix between Algeria, West Africa and France. The history of this new CRF remains elusive except for the fact that it has been observed so far in a cluster of contamination (MSM) in Southwestern France. Of note, no new cases of infection with this CRF have been identified in our geographical area since 2014. We draw the attention of the commercial companies and diagnostic laboratories in France and elsewhere to the possibility that the reagents for serological and molecular assays may no longer be fully adequate.

Acknowledgments

We thank Dr J Siqueira and I Prellwitz, Instituto Nacional de Câncer, Rio de Janeiro for their scientific contribution to the study. Our thanks also go to Pr Lazaro, Dr Duvignaud, Dr Ribeiro, Dr Paccalin and Dr Hessamfar who are clinicians at the University Hospital of Bordeaux.

We also thank Ray Cooke for revising the article.

This study is part of the Provir/Latitude 45 project, which is funded by Merck Sharp and Dohme (MSD Avenir DS-2016-0005) and partly by GERMATAN.

References

- 1 Hemelaar J, et al.: Global trends in molecular epidemiology of HIV-1 during 2000-2007. *AIDS* 2011;25:679-689
- 2 Santos AF, et al.: Characterization of a new circulating recombinant form comprising HIV-1 subtypes C and B in southern Brazil. *AIDS* 2006; 20:2011-2019.
- 3 Liitsola K, et al.: HIV-1 genetic subtype A/B recombinant strain causing an explosive epidemic in injecting drug users in Kaliningrad. *AIDS* 1998; 12:1907-1919.
- 4 Lukashov VV, et al.: Circulation of subtype A and gagA/envB recombinant HIV type 1 strains among injecting drug users in St. Petersburg, Russia, correlates with geographical origin of infections. *AIDS Res Hum Retroviruses* 1999;15:1577-1583.

- 5 Baryshev PB, Bogachev VV and Gashnikova NM. HIV-1 genetic diversity in Russia: CRF63_02A1, a new HIV type 1 genetic variant spreading in Siberia. *AIDS Res Hum Retroviruses* 2014;30:592-597.
- 6 Fernández-García A, et al. Identification of an HIV-1 BG Intersubtype Recombinant Form (CRF73_BG), Partially Related to CRF14_BG, Which Is Circulating in Portugal and Spain. *PLoS One* 2016;11:e0148549.
- 7 Delgado E, et al. Identification of a newly characterized HIV-1 BG intersubtype circulating recombinant form in Galicia, Spain, which exhibits a pseudotype-like virion structure. *J Acquir Immune Defic Syndr* 2002; 29:536-543.
- 8 Fernández-García A, et al. Identification of a new HIV type 1 circulating BF intersubtype recombinant form (CRF47_BF) in Spain. *AIDS Res Hum Retroviruses* 2010; 26:827-832.
- 9 Simonetti FR, et al. Identification of a new HIV-1 BC circulating recombinant form (CRF60_BC) in Italian young men having sex with men. *Infect Genet Evol* 2014;23:176-181.
- 10 Foster GM, et al. Novel HIV-1 recombinants spreading across multiple risk groups in the United Kingdom: the identification and phylogeography of Circulating Recombinant Form (CRF) 50_A1D. *PLoS One* 2014;9:e83337.
- 11 Struck D, et al. Near Full-Length Characterization and Population Dynamics of the Human Immunodeficiency Virus Type I Circulating Recombinant Form 42 (CRF42_BF) in Luxembourg. *AIDS Res Hum Retroviruses* 2015;31:554-558.
- 12 Leoz M, et al. Characterization of CRF56_cpx, a new circulating B/CRF02/G recombinant form identified in MSM in France. *AIDS* 2013;27:2309-2312.
- 13 Wirden M, et al. A new B/CRF02 circulating recombinant spreading quickly in Paris area Abstract 958, Conference on Retroviruses and Opportunistic Infections (CROI) , Boston, USA, 2018
- 14 Tumiotto C, et al. Diversity of HIV-1 in Aquitaine, Southwestern France, 2012-2016. *AIDS Res Hum Retroviruses* 2018; in press.
- 15 Recordon-Pinson P, et al. HIV type-1 transmission dynamics in recent seroconverters: relationship with transmission of drug resistance and viral diversity. *Antivir Ther* 2009;14:551-556.
- 16 Ode H, et al. Quasispecies Analyses of the HIV-1 Near-full-length Genome With Illumina MiSeq. *Front. Microbiol.* 2015; 6:1258.

17 Alves BM, et al. Characterization of HIV-1 Near Full-Length Proviral Genome Quasispecies from Patients with Undetectable Viral Load Undergoing First-Line HAART Therapy. *Viruses* 2017; 9: 392.

18 Lole KS, 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.

19 Abdellaziz A, et al. Predominance of CRF06_cpx and Transmitted HIV Resistance in Algeria: Update 2013-2014. *AIDS Res Hum Retroviruses* 2016;32:370-372

Patient (sample)	Origin	Risk group	Estimated year of HIV seroconversion	VL (log ₁₀ copies/mL)	CD4 (cells/mm ³)	Delay between first positive serology and treatment initiation (months)
A*	France	MSM	2009	4.5	575	16.8
B*	France	MSM	2009	4.8	628	66.2
C*	France	MSM	2014	5.0	658	1.2
D*	France	MSM	2009	5.0	486	0.5
E	France	MSM	2007	5.4	412	40.1
F	France	MSM	2010	4.7	878	4.3
G	France	MSM	2011	4.8	426	1.5
H	France	MSM	2013	2.4	351	1.1

Table 1: Patients belonging to our sequence database and exhibiting HIV-1 sequences with a CRF06-cpx protease and a subtype B reverse transcriptase.

* Patients for whom HIV-1 proviral DNA was sequenced

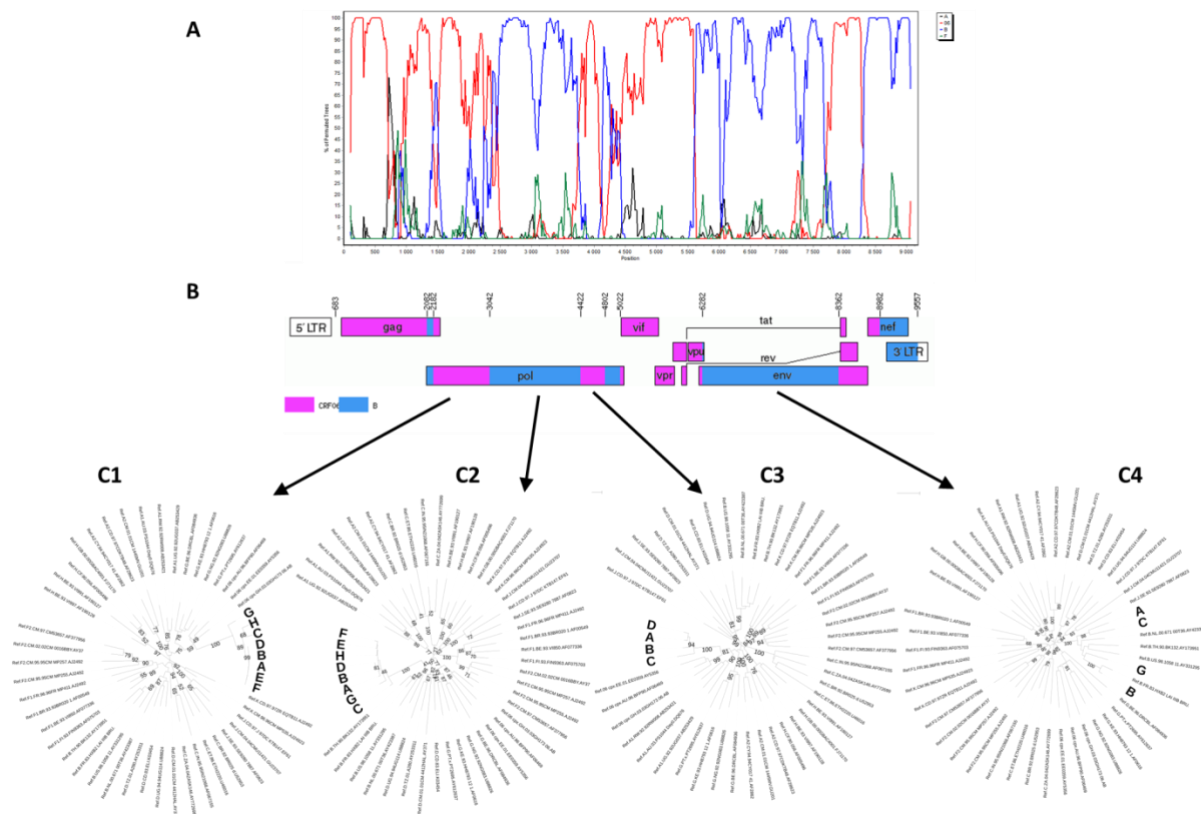
** Estimated VL and CD4 counts at time of HIV diagnosis.

Subtype Position (relative to HxB2)

Subtype	Position (relative to HxB2)

CRF06_cpx	683-2081
B	2082-2181
CRF06_cpx	2182-3041
B	3042-4421
CRF06_cpx	4422-4801
B	4802-5021
CRF06_cpx	5022-6281
B	6282-8361
CRF06_cpx	8361-8981
B	8982-9557

Table 2: Positions (relative to HxB2) of the different fragments of the sequenced viral recombinant



Bootscreening analysis, recombinant pattern and phylogenetic analysis of the new CRF98_cpx

A. Bootscreening analysis of the near full length genome of sample C. Horizontal axes represent nucleotide positions in this region, whereas vertical axes depict bootscreening values (%) that support the grouping of the isolate with each HIV-1 subtype (red for CRF06_cpx and blue for subtype B).

B. Recombination pattern of the new CRF. The subtypes and breakpoints were mapped using LANL recombinant HIV-1 Drawing tool (<http://www.hiv.lanl.gov/>). The regions are highlighted by coloured boxes (pink for CRF06_cpx and blue for subtype B). The positions indicated are according to HxB2 numbering.

C. Neighbor joining trees obtained with Sanger sequences from the eight samples and a set of reference sequences from the Los Alamos database in four regions of the recombinant genome (Prot **C1**, RT **C2**, INT **C3** and Env **C4**). Bootstrap values for the new CRF clustering are 100% in all regions analyzed.

B. Population de l'étude

Des patients infectés par le VIH-1, suivis au CHU de Bordeaux sur l'un des trois sites hospitaliers (Saint André, Haut Lévêque et Pellegrin) et en succès thérapeutique depuis plus de 6 mois (charges virales inférieures à 50 copies/mL ou indétectables) ont été inclus dans le projet Provir/Latitude 45 sur Bordeaux. Un avis favorable a été obtenu auprès du « Comité de protection des personnes du Sud-Ouest et Outremer III » (DC 2012/48) et un consentement écrit a été signé par les patients.

Pour chaque patient, à l'inclusion dans le protocole, sur le sang total, le séquençage de l'ADN proviral du VIH-1 et le typage HLA A et B sont réalisés selon le schéma descriptif ci-dessous (*Figure 10*). Pour certains patients, si des données génotypiques antérieures sont disponibles, les séquences virales ont été récupérées et analysées : celles issues de l'ARN plasmatique avant mise en place du traitement antirétroviral (J0) ou à quelques mois de l'initiation (Mx), ou celles issues de l'ADN proviral sous traitement antirétroviral efficace.

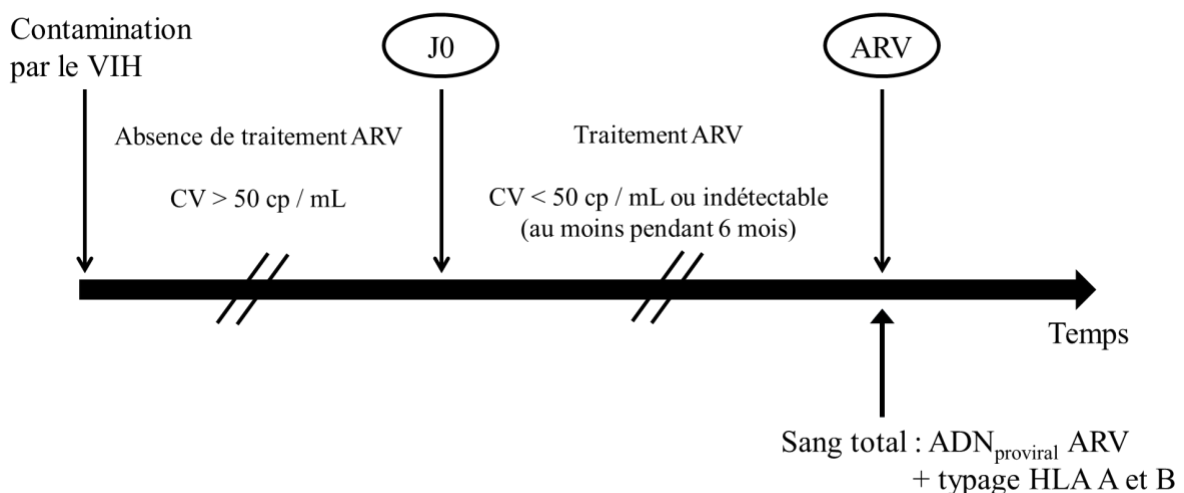


Figure 10 : Schéma représentant le descriptif de l'étude.

Ainsi, 140 patients ont été inclus dans ce projet de 2012 à 2017 au temps « ARV ». Pour 96 patients, des séquences issues de l'ARN plasmatique avant mise en place du traitement ont pu être retrouvées. Pour 4 d'entre eux, ce sont des séquences d'un génotype de résistance réalisé au cours du suivi sur ADN proviral qui ont été récupérées.

Ci-dessous le *Tableau 2* décrit les caractéristiques immuno-virologiques des 140 patients inclus dans le Projet Provir/Latitude 45.

Tableau 2 : Caractéristiques immuno-virologiques des patients inclus sur Bordeaux dans le Projet Provir/Latitude 45

Les HLA majoritaires décrits dans le Tableau 1 sont soulignés.

ND : Non détectable. NAP : données non disponibles.

Patient ID	Sexe	Age (ans)	Allèle HLA A	Allèle HLA A	Allèle HLA B	Allèle HLA B	Sous-type viral	Durée traitement (ans)	CD4 (cellules /mm ³)	Ratio CD4/CD8	PCR ADN VIH (copies /10 ⁶ cellules)
BOR001	H	54	02:06	<u>03:01</u>	<u>44:02</u>	<u>51:01</u>	02_AG	5	265	1,3	817
BOR002	H	27	<u>01:01</u>	<u>02:01</u>	<u>08:01</u>	57:01	B	8	734	0,8	ND
BOR003	H	45	<u>02:01</u>	<u>24:02</u>	27:05	<u>51:01</u>	B	3	521	1,4	480
BOR004	F	48	<u>23:01</u>	<u>68:02</u>	<u>07:02</u>	81:01	11_cpx	0,75	131	0,1	2334
BOR005	H	41	<u>01:01</u>	<u>02:01</u>	<u>07:02</u>	<u>51:01</u>	B	6	382	1,0	789
BOR006	H	55	<u>03:01</u>	<u>23:01</u>	<u>07:02</u>	<u>35:01</u>	B	10	502	0,8	572
BOR007	H	62	<u>02:01</u>		<u>40:01</u>	<u>44:02</u>	B	9	321	0,4	458
BOR008	F	42	<u>23:01</u>	<u>24:02</u>	41:02	<u>53:01</u>	C	6	372	1,1	139
BOR009	H	46	<u>02:01</u>	26:01	39:01	<u>40:01</u>	B	3	240	0,5	3854
BOR010	H	46	<u>24:02</u>	33:01	14:02	55:01	B	7	612	1,1	412
BOR011	F	40	26:01	30:02	13:03	18:01	URF	9	674	0,7	1490
BOR012	H	27	<u>02:01</u>	32:01	13:02	<u>40:02</u>	B	1	581	1,1	ND
BOR013	H	23	<u>03:01</u>	68:01	<u>07:02</u>	<u>44:02</u>	B	1	694	0,8	NAP
BOR014	H	27	<u>01:01</u>	<u>02:01</u>	<u>44:02</u>	57:01	B	2	846	1,4	< 70
BOR015	F	22	29:02	36:01	<u>53:01</u>	<u>58:01</u>	11_cpx	1	747	1,5	246
BOR016	H	42	<u>02:01</u>	29:02	39:01	44:03	B	4	578	1,0	791
BOR017	H	30	<u>24:02</u>	29:02	<u>07:02</u>	44:03	01_AE	1	609	1,8	NAP
BOR018	H	46	<u>02:01</u>	26:01	<u>08:01</u>	38:01	B	3	508	1,3	943
BOR019	H	33	<u>02:01</u>	29:02	27:05	44:03	B	1	695	1,2	71
BOR020	H	59	<u>02:01</u>	68:02	<u>35:01</u>	<u>53:01</u>	01_AE	6	1257	1,9	ND
BOR021	H	65	<u>02:01</u>	<u>24:02</u>	27:05	<u>44:02</u>	B	NAP	175	0,2	NAP
BOR022	H	43	<u>02:01</u>	<u>03:01</u>	14:01	<u>44:02</u>	B	3	995	0,9	1221
BOR023	H	25	<u>11:01</u>	68:01	<u>15:01</u>	35:03	B	2	514	1,2	ND
BOR024	H	60	<u>02:01</u>		<u>15:01</u>	<u>44:02</u>	B	4	378	0,6	1228
BOR025	H	52	<u>01:01</u>	<u>02:01</u>	35:03	<u>44:02</u>	01_AE	3	632	1,0	ND
BOR026	H	56	<u>02:01</u>	<u>30:01</u>	13:02	40:01	B	4	552	0,9	475
BOR027	F	28	<u>68:02</u>	<u>33:03</u>	07:05	<u>53:01</u>	02_AG	1	442	0,4	ND
BOR028	H	30	<u>02:01</u>	<u>31:01</u>	37:01	44:03	B	6	1014	1,9	ND
BOR029	H	40	<u>02:01</u>		<u>07:02</u>	27:05	02_AG	1	416	0,8	188
BOR030	H	51	<u>03:01</u>	30:02	<u>08:01</u>	18:01	B	21	761	0,8	< 70
BOR031	H	23	<u>02:01</u>	02:22	14:01	<u>15:01</u>	URF	2	346	2,1	< 70
BOR032	H	40	<u>11:01</u>	66:01	15:25	58:02	B	12	546	0,8	< 70
BOR033	F	46	26:01	<u>33:03</u>	<u>40:02</u>	<u>58:01</u>	B	18	584	1,5	ND
BOR034	H	74	<u>02:01</u>	02:05	<u>15:03</u>	18:01	URF	2	179	0,5	< 70
BOR035	H	51	<u>02:01</u>	68:01	15:02	<u>51:01</u>	02_AG	12	496	0,9	230
BOR036	H	63	<u>02:01</u>	<u>24:02</u>	<u>07:02</u>	37:01	B	9	1447	1,3	758
BOR037	H	77	<u>11:01</u>	26:01	<u>40:01</u>	<u>51:01</u>	NAP	16	438	2,8	ND

Patient ID	Sexe	Age (ans)	Allèle HLA A	Allèle HLA A	Allèle HLA B	Allèle HLA B	Sous-type viral	Durée traitement (ans)	CD4 (cellules /mm ³)	Ratio CD4/CD8	PCR ADN VIH (copies /10 ⁶ cellules)
BOR038	H	48	<u>01:01</u>	31:01	35:02	39:01	B	12	709	1	< 70
BOR039	F	44	<u>01:01</u>	<u>03:01</u>	18:01	38:01	B	10	766	0,8	597
BOR040	H	45	<u>02:01</u>	<u>30:01</u>	<u>07:02</u>	44:05	B	5	527	0,6	417
BOR041	H	64	<u>24:02</u>	30:02	<u>07:02</u>	18:01	B	10	768	0,5	547
BOR042	H	41	<u>23:01</u>	<u>31:01</u>	49:01	57:01	B	4	1044	1,5	225
BOR043	H	56	02:05	26:01	49:01	55:01	B	8	892	1,2	340
BOR044	H	53	02:05	<u>11:01</u>	35:03	50:01	B	7	629	1,1	ND
BOR045	F	29	<u>03:01</u>	<u>23:01</u>	<u>07:02</u>	<u>58:01</u>	URF	5	549	1	ND
BOR046	H	40	<u>03:01</u>	<u>24:02</u>	<u>07:02</u>	56:01	B	2	344	0,3	292
BOR047	F	26	<u>24:02</u>	25:01	18:01	35:08	01_AE	1	1403	3	96
BOR048	H	53	<u>02:01</u>	<u>03:01</u>	<u>07:02</u>	39:06	B	2	629	1,3	176
BOR049	H	27	<u>11:01</u>		<u>35:01</u>	55:01	B	4	533	1	157
BOR050	F	51	<u>68:02</u>	74:01	<u>44:03</u>	<u>53:01</u>	02_AG	12	1067	1,3	NAP
BOR051	H	51	30:02	<u>68:02</u>	<u>07:02</u>	14:02	02_AG	6	1223	1,1	306
BOR052	F	37	<u>02:01</u>	74:01	15:10	<u>53:01</u>	NAP	8	621	1	ND
BOR053	H	60	<u>01:01</u>	26:08	<u>15:01</u>	<u>58:01</u>	B	15	791	0,6	< 70
BOR054	F	63	<u>02:01</u>		14:01	<u>15:01</u>	B	7	454	0,8	685
BOR055	F	42	<u>03:01</u>	68:01	58:02		22_01A1	6	315	0,3	ND
BOR056	F	45	<u>02:01</u>	<u>33:03</u>	<u>08:01</u>	15:16	NAP	NAP	679	1,9	NAP
BOR057	H	34	<u>11:01</u>	33:01	14:02	<u>44:03</u>	B	3	1010	1,6	587
BOR058	F	36	NAP		18:01	<u>44:03</u>	C	8	727	1,1	318
BOR059	H	50	<u>11:01</u>	<u>24:02</u>	<u>35:01</u>	41:01	B	10	460	1,5	205
BOR060	F	58	<u>24:02</u>	29:02	<u>07:02</u>	13:02	B	17	1741	4,9	866
BOR061	H	54	<u>01:01</u>	26:01	52:01		B	18	1113	1,4	781
BOR062	F	60	34:02	66:01	<u>44:03</u>	58:02	A	3	728	2	381
BOR063	F	43	30:02	<u>68:02</u>	15:17	18:01	B	8	582	1,5	250
BOR064	F	35	<u>24:02</u>	<u>68:02</u>	35:02	50:01	C	11	224	1	364
BOR065	H	68	<u>24:02</u>		<u>07:02</u>	<u>15:03</u>	B	8	1025	2,1	776
BOR066	H	42	<u>24:02</u>	29:02	18:01	<u>51:01</u>	B	15	594	1,7	171
BOR067	H	35	<u>02:01</u>		39:01		B	4	1622	1,5	ND
BOR068	H	52	<u>02:01</u>		<u>40:01</u>	56:01	B	12	317	1	ND
BOR069	H	52	<u>03:01</u>		<u>07:02</u>	<u>35:01</u>	B	21	859	0,8	84
BOR070	H	51	<u>02:01</u>	02:66	14:01	<u>51:01</u>	B	9	1036	1,5	329
BOR071	H	52	<u>23:01</u>	34:02	<u>15:03</u>	<u>58:01</u>	URF	7	400	0,6	369
BOR072	H	55	25:01	<u>68:02</u>	14:02	<u>44:02</u>	B	3	397	1	195
BOR073	H	59	<u>11:01</u>	<u>24:02</u>	<u>08:01</u>	27:05	B	10	1069	0,8	392
BOR074	H	35	<u>02:01</u>	<u>30:01</u>	<u>35:01</u>	<u>51:01</u>	B	2	737	2,2	< 70
BOR075	H	62	<u>23:01</u>	68:01	14:03	55:01	11_cpx	11	924	1,9	1300
BOR076	H	41	<u>02:01</u>	<u>68:02</u>	<u>53:01</u>	<u>58:01</u>	D	11	497	0,7	1196
BOR077	F	49	<u>02:01</u>	<u>33:03</u>	13:02	<u>44:03</u>	02_AG	3	529	0,8	208
BOR078	F	28	<u>02:01</u>	30:02	18:01	39:01	C	6	1168	1,1	216

Patient ID	Sexe	Age (ans)	Allèle HLA A	Allèle HLA A	Allèle HLA B	Allèle HLA B	Sous-type viral	Durée traitement (ans)	CD4 (cellules /mm ³)	Ratio CD4/CD8	PCR ADN VIH (copies /10 ⁶ cellules)
BOR079	H	39	<u>03:01</u>	03:02	18:01	49:01	B	5	885	1,9	202
BOR080	H	46	<u>01:01</u>	68:01	<u>08:01</u>	57:01	C	2	468	0,7	235
BOR081	H	52	<u>03:01</u>	29:02	37:01	<u>44:03</u>	B	13	1141	2	342
BOR082	H	37	<u>01:01</u>	<u>03:01</u>	<u>07:02</u>	<u>08:01</u>	F1	4	777	0,7	2254
BOR083	F	46	<u>02:01</u>	30:02	41:01	<u>44:02</u>	B	17	1447	1,1	94
BOR084	H	42	<u>02:01</u>	<u>11:01</u>	13:02	<u>51:01</u>	B	1	861	2,1	< 70
BOR085	F	41	<u>03:01</u>	<u>30:01</u>	<u>07:02</u>	<u>53:01</u>	02_AG	11	517	0,6	785
BOR086	H	48	29:01	<u>68:02</u>	49:01	<u>53:01</u>	B	3	1279	4,4	ND
BOR087	H	62	<u>03:01</u>	26:01	<u>44:02</u>		NAP	9	519	0,9	240
BOR088	F	58	<u>24:02</u>	32:01	13:02	39:01	B	18	1180	2,5	ND
BOR089	H	48	<u>03:01</u>	30:02	<u>07:02</u>	15:10	B	13	670	1,3	430
BOR090	F	51	<u>01:01</u>		<u>08:01</u>	37:01	B	8	725	0,8	364
BOR091	H	28	<u>24:02</u>	29:02	<u>15:01</u>	<u>51:01</u>	B	3	482	0,9	316
BOR092	H	37	02:20	32:01	<u>15:01</u>	<u>35:01</u>	A	3	366	0,5	378
BOR093	F	34	68:01		<u>08:01</u>	35:03	A	3	757	2	1005
BOR094	H	53	<u>02:01</u>	30:02			B	6	1058	1,4	120
BOR095	H	49	<u>02:01</u>	<u>24:02</u>	18:01	<u>44:02</u>	BF1	3	1124	2,7	253
BOR096	H	54	<u>03:01</u>	66:01	18:01	<u>40:02</u>	B	12	778	1,2	247
BOR097	H	38	<u>03:01</u>	<u>33:03</u>	15:16	<u>44:03</u>	02_AG	10	493	1,8	277
BOR098	H	61	68:01		35:03	<u>44:02</u>	B	14	919	1,5	599
BOR099	F	51	<u>03:01</u>	<u>30:01</u>	<u>35:01</u>	58:02	URF	11	542	1,2	513
BOR100	H	44	<u>11:01</u>	<u>31:01</u>	27:05	<u>51:01</u>	B	9	1124	2	< 70
BOR101	F	35	33:01	34:02	<u>08:01</u>	82:01	02_AG	3	496	1,8	228
BOR102	H	27	<u>11:01</u>		<u>08:01</u>	<u>35:01</u>	BF1	2	768	1,7	621
BOR103	F	32	<u>02:01</u>	<u>03:01</u>	<u>07:02</u>	<u>44:02</u>	B	7	554	1,8	352
BOR104	H	49	<u>01:01</u>	<u>02:01</u>	<u>07:02</u>	13:02	NAP	17	293	2,3	532
BOR105	F	43	29:02		14:01	14:02	02_AG	11	387	1,7	688
BOR106	H	38	<u>24:02</u>	29:02	<u>07:02</u>	18:01	98_cpx	7	863	2,1	903
BOR107	F	59	<u>03:01</u>	<u>30:01</u>	13:02	27:05	B	2	639	0,8	439
BOR108	H	44	29:02	<u>68:02</u>	<u>15:01</u>	49:01	B	5	655	1,6	248
BOR109	H	49	<u>02:01</u>	26:01	27:05	39:01	B	8	348	1,6	391
BOR110	H	44	<u>02:01</u>	<u>24:02</u>	<u>07:02</u>		B	17	1115	1,6	ND
BOR111	H	53	<u>03:01</u>	26:01	38:01	<u>40:01</u>	B	4	1224	2,4	315
BOR112	H	44	<u>11:01</u>	32:01	40:19	50:01	B	7	419	0,5	728
BOR113	H	33	<u>02:01</u>		41:01	<u>51:01</u>	01_AE	5	702	0,8	164
BOR114	F	32	<u>01:01</u>		52:01	57:01	01_AE	2	604	1,4	ND
BOR115	H	54	<u>01:01</u>		<u>08:01</u>		B	9	620	1	651
BOR116	F	41	<u>01:01</u>	02:02	<u>07:02</u>	18:01	02_AG	6	428	0,4	1227
BOR117	H	43	<u>02:01</u>	<u>68:02</u>	14:01	14:02	B	2	642	0,5	314
BOR118	H	60	<u>01:01</u>	29:02	<u>08:01</u>	<u>53:01</u>	A1	3	968	2	< 70
BOR119	H	38	<u>03:01</u>	25:01	<u>08:01</u>	<u>35:01</u>	C	2	678	2,1	ND

Patient ID	Sexe	Age (ans)	Allèle HLA A	Allèle HLA A	Allèle HLA B	Allèle HLA B	Sous-type viral	Durée traitement (ans)	CD4 (cellules /mm ³)	Ratio CD4/CD8	PCR ADN VIH (copies /10 ⁶ cellules)
BOR120	F	51	68:01	74:03	<u>51:01</u>		B	9	487	1,2	205
BOR121	F	39	<u>24:02</u>	31:01	13:02	<u>35:01</u>	A6	1	893	1,5	645
BOR122	H	34	<u>24:02</u>		35:03	<u>51:01</u>	B	6	723	1	ND
BOR123	H	36	<u>02:01</u>	68:01	15:17	<u>51:01</u>	B	1	187	0,3	1300
BOR124	F	38	30:02	<u>33:03</u>	50:01	<u>53:01</u>	22_01A1	5	447	0,8	ND
BOR125	H	62	<u>02:01</u>	74:01	<u>15:03</u>	<u>58:01</u>	URF	16	426	0,5	1561
BOR126	H	30	<u>02:01</u>	68:01	<u>44:02</u>		B	6	982	1,1	350
BOR127	H	31	<u>02:01</u>	<u>23:01</u>	<u>07:02</u>	15:24	B	9	580	1,9	ND
BOR128	F	39	26:12	74:01			URF	4	665	1,6	ND
BOR129	H	44	<u>01:01</u>	26:01	<u>08:01</u>	14:01	B	2	402	0,5	555
BOR130	H	63	<u>02:01</u>	<u>24:02</u>			B	1	74	0,4	300
BOR131	H	35	<u>01:01</u>	<u>31:01</u>	<u>08:01</u>	<u>44:02</u>	B	8	557	1,2	< 70
BOR132	H	61	<u>11:01</u>	29:02	<u>40:01</u>	<u>44:02</u>	B	6	473	0,5	1072
BOR133	H	38	<u>01:01</u>	29:02	<u>44:03</u>	<u>57:01</u>	B	12	351	0,8	< 70
BOR134	F	69	<u>03:01</u>	<u>24:02</u>	27:05	<u>35:01</u>	B	2	495	0,5	717
BOR135	H	59	<u>11:01</u>	30:02	<u>51:01</u>	57:03	B	10	808	0,6	213
BOR136	H	41	26:01	33:01	14:02	<u>40:02</u>	B	9	888	0,7	307
BOR137	F	61	<u>03:01</u>	<u>24:02</u>	<u>08:01</u>	<u>40:01</u>	B	12	578	1,9	ND
BOR138	F	27	<u>02:01</u>	<u>30:01</u>	<u>15:03</u>	<u>35:01</u>	13_cpx	3	355	1	385
BOR139	H	62	<u>01:01</u>	30:02	<u>08:01</u>	18:01	B	11	530	0,7	159
BOR140	F	31	29:02	68:01	58:02	81:01	02_AG	3	629	0,9	123

Dans ce projet Provir/Latitude 45, 71 % des patients recrutés sont des hommes avec un âge médian de 44,5 ans [22-77]. Au sein de notre population d'étude, les 5 allèles HLA A et B les plus représentés chez nos patients sont :

- Pour l'allèle A : A*02:01 (35,7%), A*03:01 (18,6%), A*24:02 (17,9%), A*01:01(14,3%) et A*11:01 (10%)
- Pour l'allèle B : B*07:02 (17,9%), B*08:01 (12,9%), B*44:02 (12,1%), B*51:01 (11,4%) et B*35:01 (9,3%).

Au moment de l'inclusion dans le projet Provir/Latitude 45, les patients sont traités depuis 6 ans en médiane [9 mois – 21 ans], ils ont un nombre médian de lymphocytes T CD4+ de 620,5 cellules/mm³ [74-1741] avec un ratio lymphocytes T CD4+/lymphocytes T CD8+ médian de 1,1 [0,1-4,9]. La quantité d'ADN proviral VIH présent dans les cellules des patients inclus dans l'étude, évaluée par la PCR LTR, varie de non détectable (ND, pas de signal lors de la PCR) ou < 70 copies/10⁶ cellules (notre seuil de quantification) à 3854 copies/10⁶ cellules. La médiane de cette quantité d'ADN proviral VIH est de 253 copies/10⁶ cellules.

Le sous-type viral majoritaire est le B (63,2 %) avec 86 patients infectés par ce sous-type. Le CFR02_AG qui infecte 10,3 % des patients est majoritaire parmi les sous-types viraux non B.

2^{EME} PARTIE

IV. 2^{ème} partie - Recherche in silico d'épitopes pouvant induire une réponse T cytotoxique

A. Publication 3

Sanger and NGS data for CTL epitopes characterization in archived HIV-1 proviral DNA

**Tumiotto C, Riviere L, Bellecave P, Recordon-Pinson P, Vilain-Parce A, Guidicelli GL,
Fleury H; Provir/Latitude 45 collaborating group.
PLoS One. 2017 Sep 21;12**

L'objectif du projet Provir/Latitude 45 est d'identifier des épitopes candidats pouvant servir de base à un vaccin thérapeutique pour l'ensemble de la population ou pour un groupe de la population en fonction de leurs caractéristiques immunogénétiques. Des données de séquençage NGS par la technologie Roche 454 ont été générées par des études précédentes au laboratoire (Papuchon et al. 2013, 2014).

Pour automatiser l'analyse de ces données issues de séquençage NGS mais aussi de séquençage populationnel un logiciel TutuGenetics a été développé. Ce logiciel est capable à partir (i) d'un alignement nucléotidique par rapport à la référence HxB2, (ii) d'un gène spécifié et (iii) d'un HLA donné, d'instrumentaliser les algorithmes d'IEDB pour déterminer la prédiction de l'affinité entre un HLA et un peptide.

Ainsi, l'alignement nucléotidique est traduit en acide aminé selon l'ORF (Open Reading Frame) du gène spécifié, puis, en fonction des allèles HLA du patient considéré, les épitopes décrits pour ses allèles HLA selon la liste de « Best Defined CTL Epitopes List » de la database de Los Alamos sont caractérisés. Puis grâce aux algorithmes d'IEDB, la présentation de ces épitopes par les différents HLA du patient est étudiée grâce à un score MHC IC₅₀ déterminé pour chaque couple HLA-séquence virale.

Dans cet article, nous avons validé le logiciel TutuGenetics, en comparant l'analyse des données issues du séquençage NGS avec le séquençage Sanger.

Des données comparatives Sanger-NGS qui ne sont pas présentées dans cette publication nous ont permis de considérer que la variabilité constatée par NGS aux points d'intérêt était faible et justifiait le reste de l'analyse par Sanger.

Sanger and Next-Generation Sequencing data for characterization of CTL epitopes in archived HIV-1 proviral DNA

Camille Tumiotto¹, Lionel Riviere¹, Pantxika Bellecave¹, Patricia Pinson¹, Alice Vilain-Parce¹, Gwenda-Line Guidicelli², Hervé Fleury¹ and Provir/Latitude 45 collaborating group : Michel Dupon³, Didier Neau³, Mark Wainberg⁴ (+), Jean-Pierre Routy⁴, Jorge Sanchez⁵, Christian Brander^{6,7,8}

¹Laboratoire de Virologie, CHU de Bordeaux et CNRS UMR 5234, Bordeaux, France

²Laboratoire d'Immunologie, CHU de Bordeaux, Bordeaux, France

³Service des maladies Infectieuses et Tropicales, Hopital Pellegrin, CHU de Bordeaux, Bordeaux, France

⁴McGill University, Montreal, Canada

⁵Centro de Investigaciones Tecnológicas, Biomedicas y Medioambientales, Lima, Peru and Department of Global Health, University of Seattle, WA, USA

⁶IrsiCaixa AIDS Research Institute, Hospital Universitari Germans Trias i Pujol, Badalona (Barcelona) Spain.

⁷University of Vic – Central University of Catalonia, Spain

⁸Institució Catalana de Recerca i Estudis Avançats (ICREA), Pg. Lluís Companys 23, 08005 Barcelona, Spain

Abstract:

One of the strategies for curing viral HIV-1 is a therapeutic vaccine involving the stimulation of cytotoxic CD8-positive T cells (CTL) that are Human Leucocyte Antigen (HLA)-restricted. The lack of efficiency of previous vaccination strategies may have been due to the immunogenic peptides used, which could be different from a patient's virus epitopes and lead to a poor CTL response. To counteract this lack of specificity, conserved epitopes must be targeted. One alternative is to gather as many data as possible from a large number of patients on their HIV-1 proviral archived epitope variants, taking into account their genetic background to select the best presented CTL epitopes. In order to process big data generated by Next Generation Sequencing (NGS) of the DNA of HIV-infected patients, we have developed a software package called TutuGenetics. This tool combines an alignment derived either from sanger or NGS files, HLA typing, target gene and a CTL epitope list as input files. It allows automatic translation after correction of the alignment obtained between the HxB2

reference and the reads, followed by automatic calculation of the MHC IC₅₀ value for each epitope variant and the HLA allele of the patient by using NetMHCpan 3.0, resulting in a csv file as output result. We validated this new tool by comparing Sanger and NGS (454, Roche) sequences obtained from the proviral DNA of patients at success of ART included in the Provir Latitude 45 study and showed a 90% of correlation between the quantitative results of NGS and Sanger. This automated analysis combined with complementary samples should yield more data regarding the archived CTL epitopes according to the patients' HLA alleles and will be useful for screening epitopes that in theory are presented efficiently to the HLA groove, thus constituting promising immunogenic peptides for a therapeutic vaccine.

Introduction:

Human Immunodeficiency Virus type 1 (HIV-1) induces a chronic infection leading to immunosuppression. While viral replication can be controlled by anti-retroviral treatment (ART), the latter cannot be interrupted since this would lead in most cases to a rapid viral rebound (1). The next step is viral cure for which there are various strategies (2), therapeutic vaccination of patients under treatment being particularly promising. This vaccination involves the stimulation of a cytotoxic CD8-positive T cell (CTL)-mediated immune response which should be able to target the virus within the latent viral reservoir (3). CTL epitopes are Human Leucocyte Antigen (HLA)-restricted and must be presented by molecules encoded by HLA alleles, thereby introducing a genetic and individual parameter into the physiopathology of the immune response and the evolution of HIV disease (4-5). Some studies have shown that the viruses recovered from archived DNA are different from pre-ART circulating viruses and (6). Therefore the characterization of CTL epitopes encoded by archived DNA in patients according to their HLA alleles should generate information on epitopes that could be used for therapeutic vaccination in future trials.

Provir Latitude 45 is a project aiming at identifying CTL epitopes in archived HIV-1 DNA in patients during ART, taking into account their HLA A and B genotype. It was initiated in Bordeaux University Hospital and then extended to other groups (Montreal, Canada and Lima, Peru). It aims at gathering as broad a dataset as possible on archived epitope sequences from a large number of patients and at identifying epitopes that are conserved and potentially used for vaccination, taking into account the patients' genetic background.

The present study concerned the sequence analysis of archived HIV-1 DNA using Sanger technology and Next-Generation Sequencing (NGS). As previously detailed (7), the lack of frame awareness in the alignments in the Amplicon Variant Analyzer (AVA) software provided by Roche and the frequent shifts due to homopolymers complicates the analysis of amino acid sequences generated by 454 Roche sequencing. Therefore, we have developed software able to perform an automatic analysis of nucleic acid sequences of Gag and Pol among other genes, leading to the identification of CTL epitopes and MHC IC₅₀ values calculated from the Immune Epitope DataBase (IEDB) according to patients' specific HLA A and B alleles. This software can be used for the population analysis of archived epitopes using either Sanger or NGS technology. It was validated in this study by comparing both sequencing strategies with Sanger as the reference.

Methods:

Study patients:

Thirty-five patients were recruited at Bordeaux University Hospital, France, 6 patients from the primary infection cohort in Montreal, Canada and 6 patients at IMPACTA, Lima, Peru. All had been treated successfully with antiretroviral drugs for at least 6 months and all had viral loads below the threshold of available local molecular assays (< 40 copies then < 20 copies/mL in Bordeaux, < 40 copies/mL in Montreal and < 400 copies/mL in Lima). The study received authorization from the "Comité de protection des personnes du Sud-Ouest" (DC 2012/48).

DNA extraction, Gag and Pol PCR amplification and Sanger sequencing:

Total whole blood was collected from the patients and DNA was extracted from the PBMC followed by amplification and Sanger sequencing as described previously (4-5). Sanger sequences are available in GenBank under accession numbers MF580925 to MF580945.

Gag and Pol Ultra Deep Pyro Sequencing (UDPS):

Gag and Pol ultra-deep-sequencing were performed by using the Roche 454 technology. Amplicons for Gag (obtained with the same primers as for Sanger sequencing) and Pol (4-5) were pooled at equimolar concentrations. Clonal amplification on beads (EmPCR) was performed by using the 454 Life Science

Reagents that enable bi-directional sequencing, composed of 30 cycles of PCR amplification. DNA-containing beads were recovered and UDPS was performed on the GS Junior Sequencer (454 Life Science Roche). Analysis of the obtained data is detailed below. Raw data are available in GenBank under accession numbers SRP114367.

HLA class I typing:

The method used for the molecular characterization of HLA alleles A and B in the Bordeaux and Montreal cohorts has been described previously (4-5). For the patients from Lima, high-resolution HLA typing was performed by Sequence-Based Typing (SBT) at the University of Oklahoma Health Science Center, a CLIA/ASHI-accredited HLA typing laboratory, using in-house methods. Briefly, genomic DNA was extracted from PBMCs using a QIAamp DNA blood kit (QIAGEN). After confirmation, the PCR product was purified using an ExoSAP-IT kit (USB) and was sequenced using BigDye® Terminator v3.1 (APPLIED BIOSYSTEMS) chemistry. Dye was removed by ethanol precipitation. Sequencing reactions were performed on a 3730 Capillary Electrophoresis DNA Sequencer (APPLIED BIOSYSTEMS). Four-digit HLA types were determined using the HLA typing program Assign SBT (Conexio Genomics).

Immune Recognition tools:

Viral epitopes were selected from the Los Alamos HIV Immunology Database using the most recent update of the optimal CTL epitope list (8). Peptide binding to its restricting HLA class I molecule was predicted by using the NetMHCpan 3.0 software from the Immune Epitope Database (IEDB) (9). A value of 500 nM for the MHC IC₅₀ (IC₅₀) was used as the threshold between binder and non-binder epitopes and their variants according to the recommendation given by IEDB and other groups (10-14).

NGS sequence data analysis:

Analysis of reads was performed using a part of the AVA Software pipeline. After de-multiplexing, reads obtained for a sample were grouped and aligned with the HxB2 reference (2). The alignment of the individual reads was then recovered from AVA. UDPS is known to generate errors of around 1% (15-16). These errors are mostly insertions and deletions and occur especially in homopolymeric areas. Consequently, read alignments generate gaps and cannot be translated directly. As stated in (7),

analysis of raw UDPS data is not possible using only the AVA pipeline. The nucleotide alignment cannot be translated directly without correction because of the lack of frame awareness of AVA. Since our goal was to analyze the CTL epitopes archived in the proviral HIV-1 DNA, we needed to translate the reads into amino-acid sequences and then calculate the MHC IC₅₀ score for each epitope/HLA combination. This involves: (i) having as accurate a translation as possible to produce an in-frame correct amino acid sequence; (ii) retrieving the CTL epitopes described in the Best-Defined CTL Epitopes List (BDCEL) (8) (https://www.hiv.lanl.gov/content/immunology/tables/optimal_ctl_summary.html); (iii) calculating the MHC IC₅₀ of the epitopes using the MHC-I Processing Prediction from IEDB (<http://tools.immuneepitope.org/processing/>), considering all combinations of HLA alleles, peptide size, epitope and variant sequences; and then (iv) merging the results from each step into a single spread-sheet for further analysis. Performing all these steps manually is time-consuming and subject to error.

TutuGenetics software:

TutuGenetics is a command line software tool written in C taking as inputs numerous files (Figure1):

- The BDCEL which list HIV-1 epitopes described for given HLA alleles.
- The patient's HLA alleles A and B.
- HxB2 and the patients' sequences alignment.
- The specified target gene required to determine the correct open reading frame for the amino-acid translation.

Figure 1: Description of TutuGenetics software. Four input files are used to generate an output final file merging the quantification (% variants) and qualification (amino-acid sequence) of the variability of HIV epitopes as well as their theoretical presentation to their HLA molecules (IC₅₀ value in nM). It allows the automated analysis of nucleic acid alignments from Sanger or UDPS sequences, corrects them (for the UDPS) and translates each read into its corresponding amino acids. Then, according to the HLA genotype of the patient, it retrieves the CTL epitopes and their variants according to the Los Alamos HIV Immunology Database CTL epitope list. Finally, given the patient's HLA, the IC₅₀ scores of each epitope and their variants are calculated using the stand-alone of the NetMHCpan 3.0 software

As stated earlier, the UDPS alignment cannot be translated directly. Prior to translation, TutuGenetics is able to rearrange the UDPS alignment according to some filtering criterion such as a minimum reads length and alignment correction heuristic.

The alignment issued from a Sanger sequencing or the corrected UDPS are then translated into amino acid sequences given the open reading frame (ORF) of the gene specified by the user.

Based on translated reads/sequences, a given gene and a given HLA allele, TutuGenetics collects epitopes and their variants defined for the corresponding HLA it results in a file that qualifies and quantifies the genetic variability of the selected epitopes. It also generates a shell script instrumenting the stand-alone NetMHCpan 3.0 program from the IEDB for the MHC-I Processing Prediction. This NetMHCpan 3.0 program calculates the MHC IC₅₀ score for each epitope according to the patient HLA allele. Then, TutuGenetics merges all the results in a single spread-sheet.

The output file is a csv file summarizing the HLA alleles, the position of the epitope relative to the HxB2 reference, the epitope sequence, the number of reads encoding this epitope, the percentage of reads containing this epitope and the MHC IC₅₀ value. The first line of an epitope described for an HLA allele is the reference HxB2 and its corresponding IC₅₀.

Results:

Patients' characteristics:

To date, the analysis of archived proviral HIV-1 DNA in Sanger and UDPS has been obtained for 22 patients of the 47 included in this study. The comparison has been done in 19 patients for Pol and in 12 for Gag. Table 1 summarizes the characteristics of the viral clades and HLA alleles for the 22 patients.

Table 1: Patients' characteristics. The viral clade and HLA allele A and B typing results for each individual from the French, Canadian or Peruvian cohort are presented.

Sample ID	Geographic origin	Viral Clade	HLA allele
B01_Bg	Bordeaux	B	A*02:01 B*15:01 B*44:02

B02_BI	Bordeaux	B	A*02:01 A*29:02 B*27:05 B*44:03
B03_Ba	Bordeaux	B	A*02:01 A*24:02 B*27:05 B*51:01
B04_Ce	Bordeaux	B	A*02:01 A*29:02 B*39:01 B*44:03
B05_Ca	Bordeaux	B	A*02:01 A*02:22 B*14:01 B*15:01
B06_Cc	Bordeaux	B	A*02:01 A*26:01 B*08:01 B*38:01
B07_Dj	Bordeaux	CRF02_AG	A*02:01 A*68:01 B*15:02 B*51:01
B08_Jc	Bordeaux	B	A*02:01 A*30:01 B*13:02 B*40:01
B09_Ma	Bordeaux	CRF01_AE	A*01:01 A*02:01 B*35:03 B*44:02
B10_Mj	Bordeaux	B	A*03:01 A*68:01 B*07:02 B*44:02
B11_Sy	Bordeaux	B	A*02:01 A*26:01 B*39:01 B*40:01
B12_Sb	Bordeaux	A1	A*02:01 A*02:05 B*15:03 B*18:01
B13_TEp	Bordeaux	B	A*02:01 B*40:01 B*44:02
B14_TIp	Bordeaux	B	A*01:01 A*02:01 B*07:02 B*51:01
B15_Vp	Bordeaux	B	A*01:01 A*02:01 B*44:02 B*57:01
M01_A4	Montreal	B	A*02:01 A*33:03 B*50:01 B*58:01
M02_A7	Montreal	B	A*03:01 A*30:02 B*07:02 B*18:01
M03_G0	Montreal	B	A*02:01 A*24:02 B*08:01 B*18:01
M04_G1	Montreal	B	A*02:01 A*29:02 B*08:01 B*44:03
M05_H9	Montreal	B	A*01:01 B*35:01 B*51:01
L01_H6	Lima	B	A*24:02 A*29:02 B*35:10 B*44:03
L02_H0	Lima	B	A*24:02 B*35:43 B*58:01

Comparison between theoretical presentations of CTL epitopes archived in proviral DNA for the sequences obtained with Sanger and NGS:

To validate our tool TutuGenetics, we focused on epitopes described on the BDCEL. As Sanger sequencing is the gold standard method for describing HIV sequences, NGS results were compared to Sanger sequencing results after TutuGenetics processing.

TutuGenetics provided the following: sample name, HLA allele, position of examined peptide in Gag, RT or Protease proteins, amino-acid sequence of epitope referenced in HxB2 (first lane) and potential variants (the following lanes), theoretical MHC IC₅₀ value (nM) according to IEDB, number of reads encoding the epitope and fulfilling TutuGenetics criteria as described above and their percentages. The last column

indicates the presence or absence of the different epitopes in the sequences obtained by the Sanger method. The threshold of 1% for the detection of variants was determined according to previous data from our group (17).

Globally we analyzed 133 Gag or Pol epitopes. The results obtained with Sanger sequencing and AVA combined with TutuGenetics data were concordant for 120/133 epitopes (90.2%). Results obtained for 8 samples are presented in Table 2.

Table 2: Output of TutuGenetics and comparison with Sanger sequencing results: eight HLA/peptide associations are detailed. For each sample and HLA, the first line shows the putative epitope described in the HxB2 reference sequence and the calculated IC₅₀ (nM). For each epitope, the number of NGS reads as well as the percentage of variants is also shown. The mutated amino-acids are in bold. In the last column, “yes” means that this epitope obtained with NGS was retrieved in the Sanger sequence; “no” indicates that the epitope can be quantified by NGS but was not observed by Sanger.

<u>Sample ID</u>	<u>HLA</u>	<u>Subprotein</u>	<u>Epitope</u>	IC ₅₀ <u>(nM)</u>	<u>nb</u> <u>reads</u>	<u>%</u> <u>variant</u> <u>s</u>	<u>Concor-</u> <u>dance</u> <u>with</u> <u>Sanger</u>
<i>HxB2</i>	<i>HLA-B*27:05</i>	<i>p17 (19-27)</i>	<i>IRLRPGGKK</i>	<i>192.5</i>			
B03_Ba	HLA-B*27:05	p17 (19-27)	IRLRPGGKK	192.5	4949	93%	yes
B03_Ba	HLA-B*27:05	p17 (19-27)	L RLRPGGKK	144.5	71	1.3%	no
<i>HxB2</i>	<i>HLA-B*15:01</i>	<i>p24 (137-145)</i>	<i>GLNKIVRMY</i>	<i>392.9</i>			
B01_Bg	HLA-B*15:01	p24 (137-145)	GLNKIVRMY	392.9	8336	96.8%	yes
B01_Bg	HLA-B*15:01	p24 (137-145)	E LNKIVRMY	1467.1	108	1.3%	no

HxB2	HLA-A*24:02	p17 (28-36)	KYKCLKHIVW	256.9			
M03_G0	HLA-A*24:02	p17 (28-36)	QYKCLKHLVW	405	896	89.8%	yes
M03_G0	HLA-A*24:02	p17 (28-36)	QYKCLKHIVW	517.9	53	5.3%	no
HxB2	HLA-B*40:01	p24 (44-52)	SEGATPQDL	492.6			
B13_TEp	HLA-B*40:01	p24 (44-52)	SEGATPQDL	492.6	1160	76.5%	yes
B13_TEp	HLA-B*40:01	p24 (44-52)	SERATPQDL	600.9	312	20.6%	no
HxB2	HLA-B*40:01	p2p7p1p6 (118-126)	KELYPLTSL	16.8			
B13_TEp	HLA-B*40:01	p2p7p1p6 (118-126)	KELYPLASL	14.7	2603	94.9%	yes
B13_TEp	HLA-B*40:01	p2p7p1p6 (118-126)	KELYPLAPL	11.9	38	1.4%	no
B13_TEp	HLA-B*40:01	p2p7p1p6 (118-126)	EELYPLASL	51.7	31	1.1%	no
HxB2	HLA-A*29:02	p17 (78-86)	LYNTVATLY	25			
B04_Ce	HLA-A*29:02	p17 (78-86)	LFNTVATLY	12.1	5315	94.4%	yes
HxB2	HLA-A*24:02	p17 (28-36)	KYKCLKHIVW	256.9			
L01_H6	HLA-A*24:02	p17 (28-36)	QYRLKHIVW	528.9	1360	88.3%	yes

L01_H6	HLA-A*24:02	p17 (28-36)	EYRLKHIVW	1437.3	26	1.7%	no
L01_H6	HLA-A*24:02	p17 (28-36)	QYSLKHIVW	244.2	49	3.2%	no
L01_H6	HLA-A*24:02	p17 (28-36)	EYSLKHIVW	850.8	16	1%	no
HxB2	HLA-B*40:01	p17 (11-19)	GELDRWEKI	358.3			
B11_Sy	HLA-B*40:01	p17 (11-19)	GKLDKWERI	20805.1	2114	95.4%	yes
HxB2	HLA-B*35:01	RT (118-127)	VPLDEDFRKY	424.1			
M05_H9	HLA-B*35:01	RT (118-127)	VPLDEDFRKY	424.1	532	76.9%	yes
M05_H9	HLA-B*35:01	RT (118-127)	VPLDEDFRKC	29962.8	140	20.2%	no

We then analyzed for each sample the variability of the epitope sequences and their predicted binding to the patient HLA allele.

B03_Ba expresses the HLA-B*27:05 allele which allows presentation to CTL of the Gag p17 (19-27) epitope; the reference antigen sequence according to the Los Alamos database is noted in the first line. The same sequence was found in the patient's provirus and observed by Sanger while the 454 NGS yields a percentage of 92.99% of the peptide. Interestingly, an epitope variant with a change from I to L at the first aa position was not observed by Sanger but was quantified at 1.33% by NGS. This change did not affect the predicted epitope binding to HLA-B*27:05.

B01_Bg expresses the HLA-B*15:01 allele which has been shown to present the Gag p24 (137-145) epitope, as indicated in the first line ($IC_{50}=395.9$ nM). This epitope was observed by Sanger and quantified at 96.83% by NGS. Remarkably, an epitope variant

with G to E mutation was detected only by NGS (1.26%) and does not theoretically bind to HLA-B*15:01.

M03_G0 expresses the HLA-A*24:02 allele associated with recognition of the Gag p17 (28-36) epitope. The sequence analysis yielded two epitopes which are different from the HxB2 peptide. One epitope was identified by both technologies and should be presented to CD8 T cells and a second one was quantified only at 5.31% by NGS and is at the limit of theoretical binding affinity for effective CTL recognition.

B13_TEp expresses the HLA B*40:01 allele with potential presentation of the Gag p24 (44-52) epitope; the reference peptide was identified by Sanger and quantified by NGS at 76.47%. An epitope variant with a predicted MHC IC₅₀ above the threshold for theoretical recognition was not observed by Sanger but was quantified at 20.57% by NGS. With this same HLA allele, potential recognition for Gag p2p7p1p6 (118-126) was possible. In this case, three different epitope sequences were identified and quantified by NGS at 94.86, 1.39 and 1.13%. The sequence was also identified by Sanger and all three variants are predicted to bind HLA-B*40:01.

With HLA allele A*29:02, patient B04_Ce could present the Gag p17 (78-86) epitope. Although exhibiting a Y to F mutation compared to the HxB2 reference, the peptide encoded by provirus should still be presented. The sequence was detected by Sanger while quantified at 94.42 % by NGS.

Patient L01_H6 expresses the HLA A*24:02 allele which can present the Gag p17 (28-36) epitope. The reference peptide was not observed in the archived virus but four mutated variants were quantified by NGS. Only the variant detected at 88.25% was found by Sanger. Three of the four variants had MHC IC₅₀ values above the threshold for recognition, while the fourth was detected at a very low level and is predicted to bind HLA-A*24:02.

B11_Sy expresses HLAB*40:01 with theoretical presentation of Gag p17 (11-19). Compared to the HxB2 reference sequence, the epitope was present in archived provirus containing two aa mutations found both by Sanger and NGS (95.4%). These

changes affect binding to the HLA molecule since its MHC IC₅₀ corresponds to 20805 nM (compared to IC₅₀ 358.3 nM for HxB2).

Patient M05_H9 expresses HLA-B*35:01 and should be able to present the RT (118-127) epitope. An epitope sequence similar to our reference was noted by Sanger with a NGS quantitation at 76.88%. An epitope variant which is theoretically unable to bind to HLA-B*35:01 was not detected by Sanger but was quantitated by NGS (20.23%)

Discussion

TutuGenetics allows the rapid and reliable analysis of NGS data with direct and HLA allele-dependent identification of the CTL epitope according to a reference HIV-1 epitope list and predication of epitope binding affinity. The correlation between the quantitative results of NGS and Sanger was high and in accordance with a threshold of variant detection around 20 to 30% for NGS quantitation/Sanger detection, as generally admitted in the field (18). This tool is thus able to facilitate the identification of archived CTL epitopes according to the patients' HLA alleles for their potential inclusion in therapeutic vaccine immunogens.

As already pointed out by Deng et al (6), a large number of archived HIV-1 DNA genomes carry CTL epitope escape mutations even in patients close to primary infection. It renders infected cells insensitive to CTL directed at conserved/wild type epitopes when ART is interrupted. It is therefore crucial to identify CTL epitopes that are non-mutated or, if they do show sequence mutations, that have retained their ability to bind to the presenting HLA class I molecule. The pattern is very complicated since data indicate that a process of viral adaptation to high-frequency HLA alleles can occur in a given population (19), therefore conferring an advantage to some rare HLA supertypes in HIV disease progression (20).

TutuGenetics may be able to overcome this issue by generating big data from a large number of patients with the hypothesis that some archived CTL epitopes cannot vary for viral functional reasons, even though they are under pressure from the most frequent HLA alleles in a population. One example of a study similar to Provir but devoted to the analysis of a preventive vaccine trial (RV144) has shown that there is an association between vaccine efficacy and HLA A*02:01 (21). TutuGenetics has the advantage to be modular. Therefore it accepts as input data files generated by other NGS platforms (PGM Ion Torrent and Illumina). Also, the search of epitopes may not

be restricted to the BDCEL list and can be used to identify novel epitopes for other viruses.

Some important issues remain to be solved. One is the variability of these epitopes within the quasi-species. This variability, if present, is mainly measurable by NGS; if there is a high variability of the CTL epitopes within the quasi-species with some variants outside the range of theoretical presentation, it will be more complicated to design a peptide cocktail reacting with the most frequent HLA alleles. As noted above, most of the DNA sequences that are studied by NGS carry defective proviruses and do not allow identification of the viruses that will resume replication at ART interruption (22). The only way to study a virus resuming from the archived DNA is to induce replication in cell culture using memory TCD4 cell-enriched samples. However, these techniques require large volumes of blood and clearly do not reflect the full range of the viral reservoir (6). Moreover, they do not allow the investigation of the numerous patients required by such a strategy, in which the quest would be to define highly frequent conserved archived epitopes able to be presented by the most common HLA alleles given the above mentioned theoretical limitation. Sequencing analysis also has its limitations. First, sequenced full-length genomes are not available and even if they were, this would not necessarily mean that they indicate that the genomes come from single viruses: they may arise from computerized alignments of parts of different genomes covering the full length of the virus. To overcome these limitations, an even more labor-intensive single-genome sequencing approach would need to be employed. However, since T cells are reactive with short, linear epitopes, the issue of wrongful reconstruction of full-length genomes may not have a dramatic effect.

The second issue is the variability of the CTL epitopes observed in different compartments, particularly PBMC and gut-associated lymphoid tissues (GALT) (23). While some studies seem to indicate that the viral reservoir in patients on long-term suppressive ART is stable with few genetic changes (24), it would still be necessary to check by NGS whether such a conclusion would be applicable to CTL epitopes of archived virus in various compartments. If the corresponding CTL epitopes were to be different from one compartment to the other, this would add an additional hurdle to cross. Further studies are thus needed to investigate these issues. They will be facilitated by our TutuGenetics software, which can use raw data from different NGS sequencers, namely the PGM Ion Torrent and the Illumina.

Acknowledgments

Provir/Latitude 45 is funded by Merck, GERMATAN (Groupe d'Etude et de Recherche sur les Maladies Animales Transmissibles et les ANthropozoonoses) and received a personal donation from André Fleury (+).

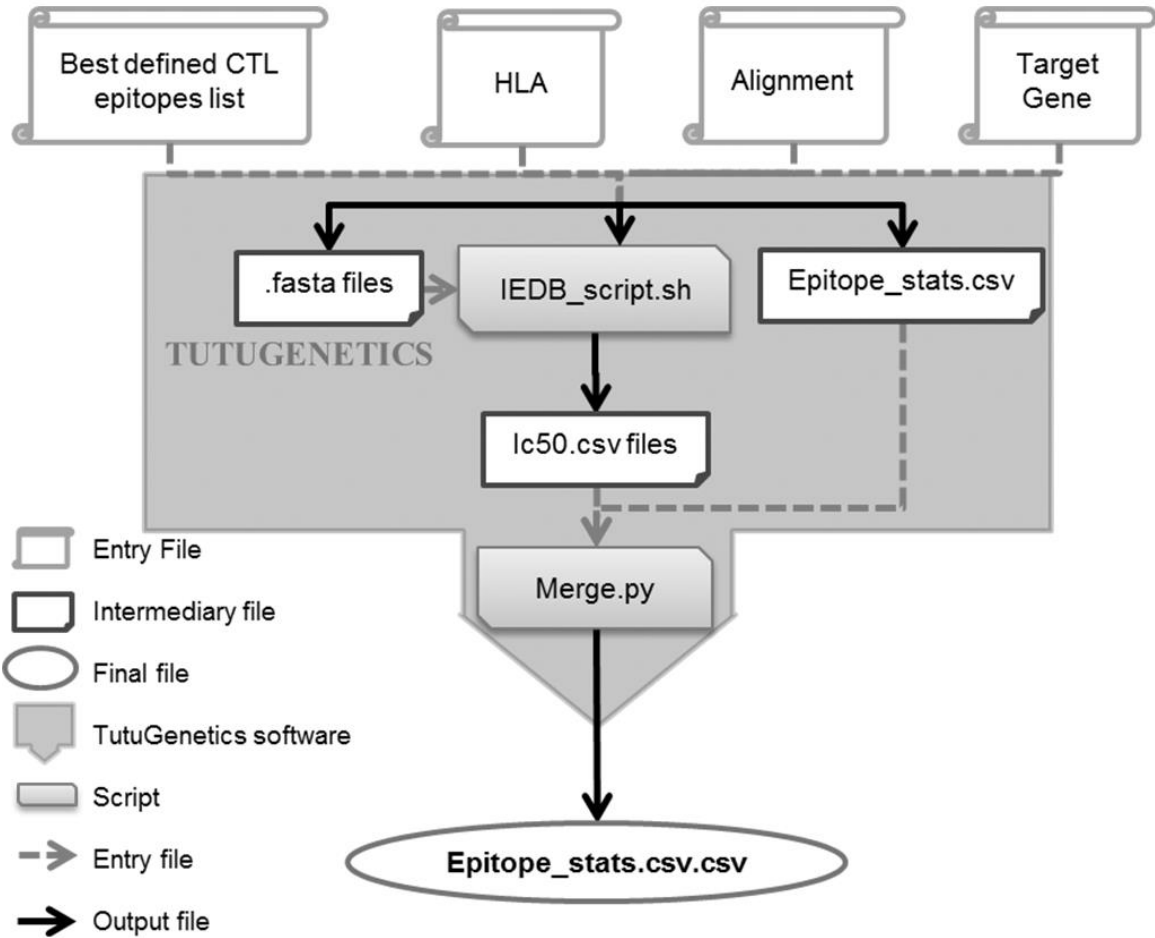
References

1. Chomont N, El-Far M, Ancuta P, Trautmann L, Procopio FA, Yassine-Diab B, et al. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat Med*. 2009 Aug;15(8):893–900.
2. Deeks SG, Autran B, Berkhout B, Benkirane M, Cairns S, Chomont N, et al. Towards an HIV cure: a global scientific strategy. *Nat Rev Immunol*. 2012 Aug;12(8):607–14.
3. Migueles SA, Osborne CM, Royce C, Compton AA, Joshi RP, Weeks KA, et al. Lytic granule loading of CD8+ T cells is required for HIV-infected cell elimination associated with immune control. *Immunity*. 2008 Dec 19;29(6):1009–21.
4. Papuchon J, Pinson P, Lazaro E, Reigadas S, Guidicelli G, Taupin J-L, et al. Resistance mutations and CTL epitopes in archived HIV-1 DNA of patients on antiviral treatment: toward a new concept of vaccine. *PloS One*. 2013;8(7).
5. Papuchon J, Pinson P, Guidicelli G-L, Bellecave P, Thomas R, LeBlanc R, et al. Kinetics of HIV-1 CTL epitopes recognized by HLA I alleles in HIV-infected individuals at times near primary infection: the Provir/Latitude45 study. *PloS One*. 2014;9(6).
6. Deng K, Perteau M, Rongvaux A, Wang L, Durand CM, Ghiaur G, et al. Broad CTL response is required to clear latent HIV-1 due to dominance of escape mutations. *Nature*. 2015 Jan 7;
7. Bellecave P, Recordon-Pinson P, Fleury H. Evaluation of Automatic Analysis of Ultradeep Pyrosequencing Raw Data to Determine Percentages of HIV Resistance Mutations in Patients Followed-Up in Hospital. *AIDS Res Hum Retroviruses*. 2016 Jan;32(1):85–92.
8. Llano A, Williams A, Olvera A, Silva-Arrieta S, Brander C. Best-Characterized HIV-1 CTL Epitopes: The 2013 Update. *HIV Molecular Immunology* 2013; 2013.

9. Kim Y, Ponomarenko J, Zhu Z, Tamang D, Wang P, Greenbaum J, et al. Immune epitope database analysis resource. *Nucleic Acids Res.* 2012 Jul;40(Web Server issue):W525-530.
10. Nielsen M, Lundegaard C, Worning P, Lauemoller SL, Lamberth K, Buus S, et al. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci Publ Protein Soc.* 2003 May;12(5):1007–17.
11. Nielsen M, Lundegaard C, Blicher T, Lamberth K, Harndahl M, Justesen S, et al. NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. *PloS One.* 2007 Aug 29;2(8):e796.
12. Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, Nielsen M. NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8-11. *Nucleic Acids Res.* 2008 Jul 1;36(Web Server issue):W509-512.
13. Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinforma Oxf Engl.* 2016 Feb 15;32(4):511–7.
14. Paul S, Weiskopf D, Angelo MA, Sidney J, Peters B, Sette A. HLA class I alleles are associated with peptide-binding repertoires of different size, affinity, and immunogenicity. *J Immunol Baltim Md 1950.* 2013 Dec 15;191(12):5831–9.
15. Gilles A, Meglec E, Pech N, Ferreira S, Malausa T, Martin J-F. Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. *BMC Genomics.* 2011 May 19;12:245.
16. Rask TS, Petersen B, Chen DS, Day KP, Pedersen AG. Using expected sequence features to improve basecalling accuracy of amplicon pyrosequencing data. *BMC Bioinformatics.* 2016 Apr 22;17:176.
17. Vandenhende M-A, Bellecave P, Recordon-Pinson P, Reigadas S, Bidet Y, Bruyand M, et al. Prevalence and evolution of low frequency HIV drug resistance mutations detected by ultra deep sequencing in patients experiencing first line antiretroviral therapy failure. *PloS One.* 2014;9(1).
18. Palmer S, Kearney M, Maldarelli F, Halvas EK, Bixby CJ, Bazmi H, et al. Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. *J Clin Microbiol.* 2005 Jan;43(1):406–13.

19. Kawashima Y, Pfafferoth K, Frater J, Matthews P, Payne R, Addo M, et al. Adaptation of HIV-1 to human leukocyte antigen class I. *Nature*. 2009 Apr 2;458(7238):641–5.
20. Trachtenberg E, Korber B, Sollars C, Kepler TB, Hraber PT, Hayes E, et al. Advantage of rare HLA supertype in HIV disease progression. *Nat Med*. 2003 Jul;9(7):928–35.
21. Gartland AJ, Li S, McNevin J, Tomaras GD, Gottardo R, Janes H, et al. Analysis of HLA A*02 association with vaccine efficacy in the RV144 HIV-1 vaccine trial. *J Virol*. 2014 Aug;88(15):8242–55.
22. Ho Y-C, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DIS, et al. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell*. 2013 Oct 24;155(3):540–51.
23. Barton K, Winckelmann A, Palmer S. HIV-1 Reservoirs During Suppressive Therapy. *Trends Microbiol*. 2016 May;24(5):345–55.
24. Josefsson L, von Stockenstrom S, Faria NR, Sinclair E, Bacchetti P, Killian M, et al. The HIV-1 reservoir in eight patients on long-term suppressive antiretroviral therapy is stable with few genetic changes over time. *Proc Natl Acad Sci U S A*. 2013 Dec 17;110(51):E4987-4996.

Figure 1



B. Les patients inclus dans le projet Provir/Latitude 45 au CHU de Bordeaux

1. Objectif

L'objectif principal est d'identifier des épitopes induisant une réponse T cytotoxique en fonction de la variabilité virale du VIH-1 et de l'immunogénétique des patients en croisant, *in silico*, les données immunogénétiques (HLA) et génomiques (provirus) de PVVIH en succès thérapeutique.

2. Matériels et méthodes

a) Inclusion

Comme abordé dans le chapitre *III- 1^{ère} partie – Population de l'étude – B - Population de l'étude* 140 patients infectés par le VIH-1 suivis au CHU de Bordeaux et en succès thérapeutique depuis plus de 6 mois (charges virales inférieures à 50 copies/mL ou indétectables) ont été inclus dans notre étude. Un avis favorable a été obtenu auprès du « Comité de protection des personnes du Sud-Ouest et Outremer III » (DC 2012/48) et un consentement écrit a été signé par les patients. Les 11 premiers patients inclus ont fait l'objet d'une publication par notre équipe (Papuchon et al. 2013).

b) Données immuno-virologiques

Les données immuno-virologiques des patients sont recueillies le long du suivi. Le dosage de la charge virale VIH (en copies/mL) est réalisé au laboratoire de virologie et le dosage du nombre de lymphocytes T CD4+ (en nombre de cellules/mm³) est réalisé au laboratoire d'immunologie. Ces dosages sont réalisés en routine dans le cadre de la prise en charge de l'infection par le VIH-1.

c) Extractions d'ADN

A partir d'aliquotes de sang total (1,5 à 3 mL) des patients, l'ADN total, contenant l'ADN proviral intégré, est extrait pour la réalisation du séquençage VIH et du typage HLA. Cette extraction est réalisée sur le Tecan EVO selon le protocole du fournisseur.

d) Séquençage du VIH-1

Afin d'amplifier les gènes Gag, Nef, Protease et Reverse Transcriptase en vue de les séquencer, une première PCR est réalisée à partir de l'éluat d'ADN total extrait. Pour essayer de limiter les biais liés à la PCR et pour augmenter la variabilité analysée, les PCR ont été réalisées en double. Ces produits sont ensuite soumis à une PCR nichée. Ces PCR sont effectuées sur un Primus Thermocycler.

La première PCR est réalisée à l'aide du kit ThermoStar® 2 Hot Start Taq contenant l'enzyme ThermoStar (Eurobio) dans un volume réactionnel de 25 µL au final dont 5 µL d'ADN.

Pour réaliser la PCR nichée, 2 microlitres des produits de première PCR sont prélevés et amplifiés avec le kit Eurobio dans un volume réactionnel de 50 µL au final.

Les séquences des amorces utilisées sont colligées dans le *Tableau 3* ci-dessous.

Tableau 3 : Liste des amorces utilisées pour amplifier les gènes Gag, Protease et Reverse Transcriptase et Nef

Nom de l'amorce et du gène cible	Position sur HxB2	Séquence (5' → 3')
Gag 1-5, 1 ^{ère} intention, (1 ^{ère} PCR + PCR nichée)	764-785	GACTAGCGGAGGCTAGAA
Gag 1a3, 1 ^{ère} intention, (PCR nichée)	1219-1241	GTTCTAGGTGATATGGCCCTGATG
Gag 1b5, 1 ^{ère} intention, (PCR nichée)	1232-1256	CACCTAGAACTTTAAATGCATGGGT
Gag 1-3, 1 ^{ère} intention, (1 ^{ère} PCR + PCR nichée)	1631-1658	TTTGGTCCTTGTCTTATGTCCAGA
Gag 2-5, 1 ^{ère} intention, (1 ^{ère} PCR + PCR nichée)	1544-1573	ATAATCCACCTATCCCAGTAGGAGAAATT
Gag 2a3, 1 ^{ère} intention, (PCR nichée)	2264-2281	AGGGGTCGTTGCCAAAGA
Gag 2b5, 1 ^{ère} intention, (PCR nichée)	2135-2163	TCAGAGCAGACCAGAGCCAACAGCCCCA
Gag 2-3, 1 ^{ère} intention, (1 ^{ère} PCR + PCR nichée)	2621-2650	AATGCTTTTATTTTTTCTTCTGTCAATGGC
Gag 5'O, 2 ^{ème} intention (1 ^{ère} PCR)	762-783	TTGACTAGCGGAGGCTAGAAGG
Gag 3'O, 2 ^{ème} intention (1 ^{ère} PCR)	2390-2415	GATAAAACCTCCAATTCCCCCTATC
Gag 5'I, 2 ^{ème} intention (PCR nichée)	785-807	GAGAGATGGGTGCGAGAGCGTC
Gag 3'I, 2 ^{ème} intention (PCR nichée)	2316-2338	CTGCTCCTGTATCTAATAGAGC
RT MJ3, 1 ^{ère} intention, (1 ^{ère} PCR)	2480-2499	AGTAGGACCTACACCTGTCA
Prot 5P1 (1 ^{ère} PCR)	2082-2109	TAATTTTTTAGGGAAGATCTGGCCTTCC
Prot 3P1 (1 ^{ère} PCR)	2703-2734	GCAAATACTGGAGTATTGTATGGATTTTCAGG

Prot 5P2 (PCR nichée)	2136-2163	TCAGAGCAGACCAGAGCCAACAGCCCCA
Prot 3P2 (PCR nichée)	2621-2650	AATGCTTTTATTTTTTCTTCTGTCAATGGC
RT MJ4, 1 ^{ère} intention, (1 ^{ère} PCR)	3399-3420	CTGTTAGTGCTTTGGTTCCTCT
RT A35, 1 ^{ère} intention (PCR nichée)	2530-2558	TTGGTTGCACTTTAAATTTTCCATTAGTCCTATT
RT NE35, 1 ^{ère} intention (PCR nichée)	3300-3334	CCTACTAACTTCTGTATGTCATTGACAGTCCAGCT
RT RT18, 2 ^{ème} intention, (1 ^{ère} PCR)	2377-2407	GGAAACCAAAAATGATAGGGGGAATTGGAGG
RT RT21, 2 ^{ème} intention, (1 ^{ère} PCR)	3509-3539	CTGTATTTCTGCTATTAAGTCTTTTGATGGG
RT RT1, 2 ^{ème} intention (PCR nichée)	2604-2632	CCAAAAGTTAAACAATGGCCATTGACAGA
RT RT4, 2 ^{ème} intention (PCR nichée)	3231-3250	AGTTCATAACCCATCCAAAG
Nef 1-5 (1 ^{ère} PCR)	8673-8699	GCCACAGCCATAGCAGTAGCTGAGGGG
Nef 1-3 (1 ^{ère} PCR)	9512-9540	CCAGTACAGGCAAAAAGCAGCTGCTTATA
Nef 2-5 (PCR nichée)	8754-8782	CCTAGAAGAATAAGACAGGGCTTGGAAG
Nef 2-3 (PCR nichée)	9443-9468	ACGCCTCCCTGGAAAGTCCCCAGCGG

Les températures d'hybridation sont toutes de 55°C sauf pour les amorces Gag de 2^{ème} intention à 57°C.

La qualité et la taille des produits de la PCR sont ensuite évaluées par électrophorèse sur gel d'agarose 1,5%.

Les produits PCR sont purifiés soit sur des micro-colonnes Séphadex de type MicroSpin™ S-400 HR (Amersham Biosciences), soit selon le protocole de Purification du kit Agencourt

AMPure XP (Beckman Coulter) afin d'éliminer entre autres les amorces non utilisées et les dNTPs libres.

Pour le séquençage populationnel, la méthode Sanger est utilisée. Les réactions de séquences bidirectionnelles sont réalisées dans un volume final de 20 µL. Le système Applied Biosystems 3500 xls Dx Genetic Analyser est utilisé pour le séquençage.

e) PCR ADN VIH

La PCR ADN VIH a été réalisé à partir d'ADN total extrait selon le protocole décrit dans la publication sur le projet Provir/Latitude (Papuchon et al. 2013).

f) Typage HLA de classe I

Le typage des gènes A et B du HLA des patients est réalisé au laboratoire d'Immunologie/Immunogénétique selon un protocole déjà décrit dans les publications sur le projet Provir/Latitude 45 (Papuchon et al. 2013, 2014).

g) Prédiction de la reconnaissance des épitopes en fonction du HLA par bio-informatique

L'affinité du HLA pour les épitopes est évaluée par la valeur du MHC IC₅₀ (nM). Elle combine pour un peptide donné les prédictions de dégradation par le protéasome, leur transport par TAP et leur fixation au HLA. Pour ce faire, les séquences des épitopes sont croisées avec le HLA du patient avec les algorithmes du site Immuno Epitope Data Base (<http://tools.immuneepitope.org/processing>) ce qui permet de prédire l'affinité du HLA pour le peptide et d'identifier la capacité intrinsèque de chaque peptide à être un épitope potentiel pour les lymphocytes T. Plus la valeur du MHC IC₅₀ est basse, plus l'affinité est bonne. Un score < 50 nM indique une forte affinité entre le HLA et le peptide. Entre 50 et 500 nM une affinité moyenne. Au-delà de 500 nM, on considère que la reconnaissance de l'épitope par ce HLA n'est plus possible en théorie (Nielsen et al. 2003; Peters and Sette 2005; Nielsen et al. 2007; Lundegaard et al. 2008; Sidney et al. 2008; Hoof et al. 2009; Kim et al. 2012; Paul et al. 2013; Andreatta and Nielsen 2016; Nielsen and Andreatta 2016)..

h) Logiciel TutuGenetics

Comme abordé dans le chapitre *IV – 2^{ème} partie – Recherche in silico d'épitopes pouvant induire une réponse T cytotoxique – A – Publication 3*, le logiciel TutuGenetics a été développé. Il permet d'automatiser les algorithmes d'IEDB pour étudier la prédiction de reconnaissance entre un potentiel épitope et un HLA. Ce logiciel peut soit étudier des épitopes en fonction de la database « Best Defined CTL Epitope List » de Los Alamos selon le HLA du patient, soit par découpage séquentiel de la séquence virale par pas de 8 à 10 acides aminés afin de trouver des potentiels nouveaux épitopes induisant une réponse CTL.

3. Résultats

Comme abordé dans le chapitre *III – 1^{ère} partie – Population de l'étude – B – Population de l'étude*, 140 patients ont été inclus dans le projet Provir/Latitude 45 sur le CHU de Bordeaux. En ce qui concerne les amplifications et séquençage obtenus à partir de l'ADN proviral sur les prélèvements récupérés à l'inclusion dans ce projet au temps « ARV », nous avons obtenu :

- 25/140 patients avec une absence totale d'amplifications des gènes Pol, Gag et Nef
- 105/140 patients avec un séquençage de la Reverse Transcriptase virale
- 90/140 patients avec un séquençage complet ou partiel du gène Gag
- 52/140 patients avec un séquençage de la Protease virale
- 44/140 patients avec un séquençage du gène Nef.

La réussite de l'amplification des gènes RT, Gag, Prot ou Nef ne semble pas dépendante du nombre de copies d'ADN total du VIH (cf *Tableau 2*).

Pour ces 140 patients, toutes les séquences qui ont pu être amplifiées ou récupérées (pour les génotypes de résistance antérieurs) ont été alignés par rapport à une séquence B de référence HxB2 pour chacun des gènes disponibles.

Ensuite, ces fichiers ont été analysés par le logiciel TutuGenetics en fonction du HLA des patients en demandant au logiciel un découpage séquentiel de la séquence virale par pas de 8 à 10 acides aminés (méthode Slice).

Pour la Protease, qui est le gène étudié le plus petit, faisant 99 acides aminés, en méthode Slice : 92 épitopes potentiels sont générés si la séquence virale est découpée par fraction de 8 acides aminés, 91 par pas de 9 et 90 par pas de 10 soit 273 épitopes potentiels pour un HLA donné. Chaque patient ayant 2 allèles HLA A et 2 allèles HLA B éventuellement différents, il y a 1092 épitopes possibles à analyser rien que pour la Protease virale. Si l'on considère le gène Gag qui est le plus long étudié dans ce projet, 500 acides aminés, 4428 épitopes potentiels sont ainsi analysés.

Les données générées sont ensuite compilées par gène et par HLA pour identifier les épitopes conservés entre les individus et induisant la meilleure réponse T cytotoxique théorique possible.

Pour chaque gène, chaque HLA majoritaire des populations Caucasiennes, Africaines, Asiatiques et Hispaniques, un premier filtre est appliqué : les épitopes retenus sont ceux ayant eu au moins une fois un MHC IC₅₀ < 50 nM. Ensuite un second filtre est utilisé : seuls les épitopes ayant un MHC IC₅₀ < 50 nM pour plus de 80% des échantillons analysés sont gardés. En dernier lieu, seuls les couples HLA-épitopes avec plus de 10 échantillons patients séquencés sont conservés. Une nouvelle liste d'épitopes d'intérêt en fonction du HLA est ainsi établie : « Optimal_Provir » décrit dans le *Tableau 4* ci-dessous.

Tableau 4 : Liste « Optimal_Provir » décrivant les épitopes d'intérêt en fonction des HLA étudiés.

HLA	Nom de l'épitope	Epitope comprenant une position avec mutation de résistance aux ARV
HLA-A*02:01	Pol01	OUI
HLA-A*02:01	Pol02	OUI
HLA-A*02:01	Pol03	NON
HLA-A*03:01, HLA-A*11:01	Pol04	NON
HLA-A*03:01, HLA-A*11:01	Pol05	OUI
HLA-A*11:01	Pol06	NON
HLA-A*68:02	Pol07	OUI
HLA-B*07:02, HLA-B*35:01	Pol08	NON
HLA-B*08:01	Pol09	NON
HLA-B*15:01	Pol10	NON
HLA-B*35:01	Pol11	OUI
HLA-A*02:01	Gag01	NON
HLA-A*02:01	Gag02	NON
HLA-B*07:02	Gag03	NON
HLA-A*24:02	Nef01	NON

Les tris successifs ayant permis de définir cette liste « Optimal_Provir » sont décrits dans la *Figure 11* ci-dessous.

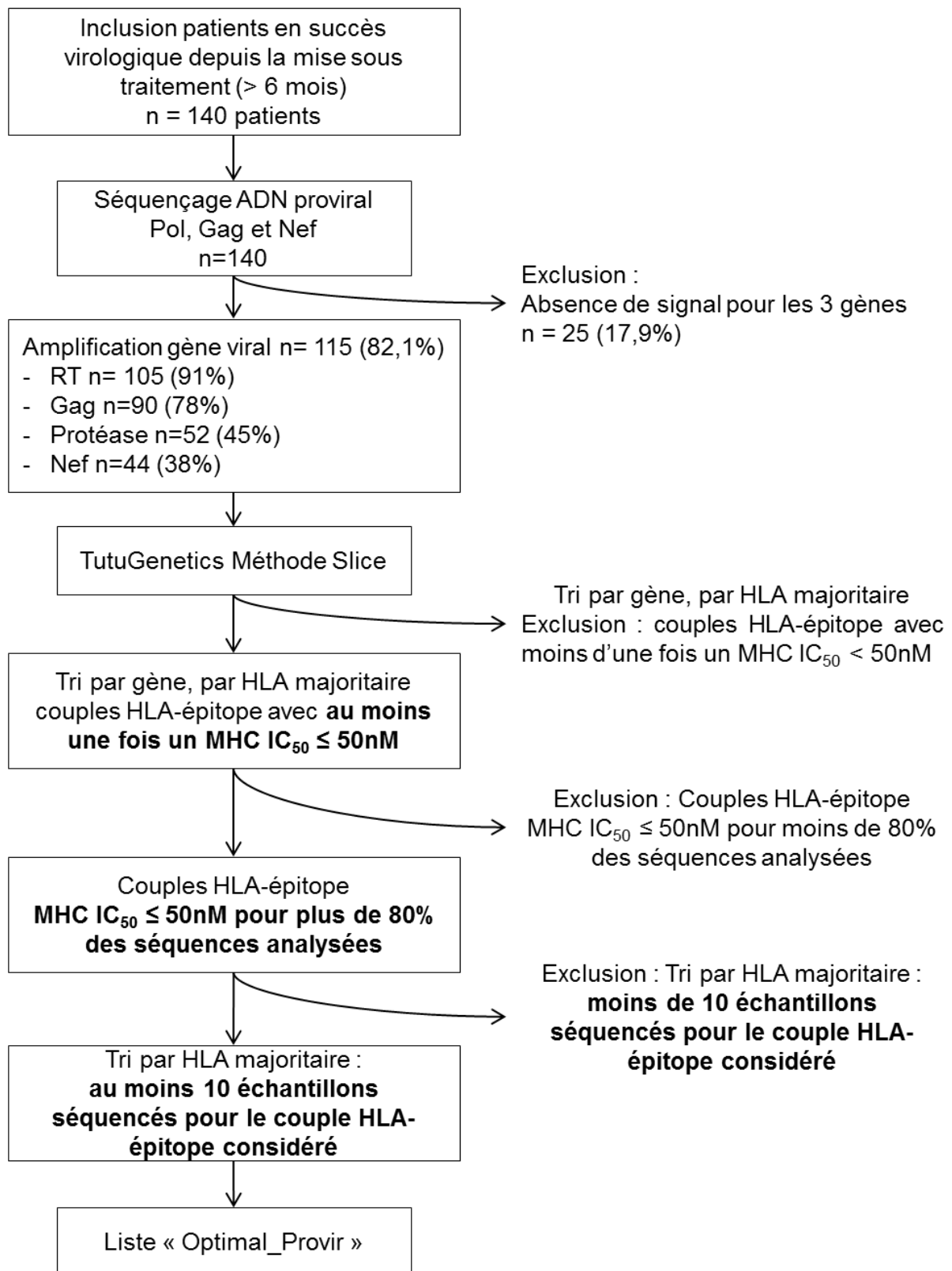


Figure 11 : Tris sélectifs ayant permis la constitution de notre liste d'épitopes d'intérêt en fonction des HLA étudiés.

Ainsi, 15 épitopes ont pu être mis en évidence pour les HLA majoritaires : 3 épitopes dans le gène Gag, 1 dans le gène Nef et 11 dans le gène Pol et plus particulièrement sur la Reverse Transcriptase virale. En majorité (9/15), les épitopes font une taille de 9 acides aminés. Sur ces 11 épitopes situés sur la RT, 5 englobent des positions associées à des mutations de résistance. Parmi les séquences virales des patients analysés, très peu sont porteuses de ces mutations de résistance et, lorsque ces dernières sont présentes, elles ne modifient pas la présentation théorique de l'épitope au HLA considéré. Pour étudier l'impact des mutations de résistance possibles sur ces épitopes, nous avons simulé par bio-informatique toutes les combinaisons de mutations possibles. Pour les épitopes Pol01 et Pol02, la majorité des couples HLA-épitope mutés a un MHC IC₅₀ > 50 nM suggérant une perte d'affinité entre l'épitope et le HLA.

Etant donné la prépondérance d'infection par du sous-type B dans notre population d'étude et la disparité des autres sous-types viraux infectant les patients inclus dans l'étude, l'analyse a été réalisée sur l'ensemble des données disponibles.

CONCLUSIONS ET PERSPECTIVES

V. Conclusions et Perspectives

L'objectif de ce travail de thèse est d'identifier des épitopes induisant une réponse T cytotoxique en considérant d'une part la variabilité du VIH-1 et d'autre part la variabilité immunogénétique des patients infectés par le VIH-1 en succès thérapeutique. L'idée, à terme, est d'utiliser ces peptides comme base de vaccin thérapeutique et de permettre au système immunitaire de prendre le relai du traitement antirétroviral pour contrôler l'infection virale.

Cent quarante patients infectés par le VIH-1, suivis au CHU de Bordeaux sur l'un des trois sites hospitaliers (Saint André, Haut Lévêque et Pellegrin) et en succès thérapeutique depuis plus de 6 mois (charges virales inférieures à 50 copies/mL ou indétectables) ont été inclus dans le projet Provir/Latitude 45 à Bordeaux entre 2012 et 2017.

Nous avons initialement réalisé une cartographie de la répartition des sous-types viraux du VIH-1 en Aquitaine pour s'assurer que notre population d'étude était bien représentative des patients vivant avec le VIH dans la région (Tumiotto et al. 2018). L'analyse rétrospective de plus de 3200 génotypes de résistance VIH-1 effectués entre 2012 et 2016 a permis de déterminer que le sous-type viral majoritaire infectant les patients vivants avec le VIH dans la région est le sous-type B. Le CRF02_AG vient en deuxième position et est le majoritaire parmi les sous-types viraux non B, comme cela a été décrit dans des études nationales portant chez des patients en primo-infection, naïfs de traitement ou en échec virologique (Frange et al. 2015; Descamps et al. 2013; Assoumou et al. 2017).

Si l'on se focalise sur les patients inclus dans ce projet Provir/Latitude 45, on retrouve la prédominance du sous-type B (63%) suivi du CRF02_AG (10,3%).

Cette première analyse phylogénétique a par ailleurs mis en évidence un cluster composé de 8 patients infectés par un nouveau virus recombinant composé de CRF06_cpx et de sous-type B. Nous avons donc cartographié le génome de ce nouveau virus en collaboration avec l'équipe Brésilienne associée au projet Provir/Latitude 45. L'article est sous presse dans *AIDS Research and Human Retroviruses* (Recordon-Pinson et al. 2018). Il s'agit du premier recombinant décrit avec entre du CRF06_cpx et du sous-type B et a été référencé en tant que CRF98_cpx.

Un patient inclus au sein du projet Provir/Latitude 45 est d'ailleurs infecté par un virus de sous-type CRF98_cpx.

Comme dit précédemment, l'objectif principal du projet Provir/Latitude 45 est d'identifier des épitopes candidats pouvant servir de base à un vaccin thérapeutique pour l'ensemble de la population ou pour un groupe de la population en fonction de leurs caractéristiques immunogénétiques. Pour cela, nous combinons les données des séquences virales avec le typage HLA des patients afin de prédire *in silico* l'affinité entre un HLA et un peptide via les algorithmes d'IEDB. Etant donné le nombre de patient inclus au sein de l'étude (n = 140), la diversité des allèles HLA (n = 35 pour les allèles HLA-A et n = 53 pour les allèles HLA-B), le nombre de gènes viraux analysés (n = 4), mais aussi les techniques de séquençage (populationnel en Sanger et/ou NGS), nous avons développé un logiciel, TutuGenetics, afin d'automatiser l'analyse de ces données.

Ce logiciel à partir d'une séquence nucléotidique instrumentalise les algorithmes IEDB et permet d'étudier la variabilité de la présentation des épitopes par les différents HLA du patient grâce à un score MHC IC₅₀ déterminé pour chaque couple HLA-séquence virale. Ce logiciel modulable permet d'analyser les épitopes soit selon la liste des « Best Defined CTL Epitopes List » de la database de Los Alamos, soit après découpage de la séquence virale par pas de 8 à 10 acides aminés pour identifier de nouveaux épitopes potentiels. Dans cet article, nous avons validé le logiciel TutuGenetics, en comparant l'analyse des données issues du séquençage NGS avec le séquençage Sanger (Tumiotto et al. 2017).

Dans la dernière partie de ce travail de thèse, nous avons réalisé le séquençage de l'ADN proviral des 140 patients inclus ainsi que leur typage HLA. En utilisant le logiciel TutuGenetics, nous avons collecté les valeurs MHC IC₅₀ pour toutes les combinaisons HLA-épitope. Une analyse plus fine nous a ensuite permis de déterminer une liste de 15 épitopes (cf *Tableau 4*) avec une forte affinité *in silico* pour les HLA majoritaires et induisant donc potentiellement une bonne réponse CTL. Ces épitopes ont été sélectionnés en appliquant différents filtres (détaillés dans la *Figure 11*) après découpage des séquences virales par pas de 8 à 10 acides aminés. Seuls les couples HLA-peptide avec plus de 10 patients séquencés par position et par HLA ont été conservés dans cette liste « Optimal_Provir ».

Sur ces 15 épitopes identifiés, 1 seul (Pol03) n'est pas enregistré dans la liste « CTL epitope variants and escape mutations » ni dans la « CTL/CD8+ T-cell Epitope Database » du site de Los Alamos HIV Database ; 5 de ces épitopes (Pol02, Pol05, Pol06, Pol07 et Pol10) sont caractérisés mais leur pouvoir immunogène n'a pas été évaluée pour les mêmes HLA que ceux présentés dans ce travail. Pour les 9 couples HLA-épitopes restants, des études *in silico* ou immuno-fonctionnelles ont déjà mis en évidence ces associations.

Une protection juridique est à l'étude sur ce cocktail peptidique, limitant ainsi la discussion autour de leurs localisations, compositions et/ou données préexistantes publiées.

Toutefois, si l'on regarde la composition de vaccins génériques tels que le vaccin Lipo-5, 3 de ces 5 lipopeptides ne sont pas communs avec notre liste « Optimal_Provir » (Frey et al. 2014). De manière comparable, le vaccin Vacc-4x évalué par l'équipe de Pollard (Pollard et al. 2014) a montré une absence de bénéfice dans le groupe vacciné par rapport au groupe Placebo. Aucun des peptides contenus dans ce vaccin n'est identique à ceux déterminés dans ce travail.

Nous pouvons donc espérer que notre approche donne un choix supplémentaire, plus efficace à l'échelle populationnelle.

Des données comparatives Sanger-NGS qui ne sont pas présentées dans ce travail nous ont permis de considérer que la variabilité constatée par NGS aux points d'intérêt était faible dans le sang total périphérique et justifiait le reste de l'analyse par Sanger. Néanmoins, le partenariat avec le Brésil, qui ne séquence uniquement qu'avec une technologie NGS, va nous permettre de confirmer ces données préliminaires et cette liste « Optimal_Provir ». Ainsi, nous vérifierons l'absence de variants minoritaires, au niveau des épitopes identifiés, modifiant la qualité de la présentation théorique entre le HLA et le peptide. Ce partenariat, nous donne également l'opportunité de pouvoir recruter au sein du projet des patients infectés par différents sous-types du VIH-1 et ayant une immunogénétique différente.

Dans la littérature, la faible variabilité virale dans le sang périphérique, dès qu'un traitement par ARV est mis en place, semble unanime (Josefsson, von Stockenstrom, et al. 2013; van Zyl, Bale, and Kearney 2018). De même, en ce qui concerne les réservoirs viraux, différentes études montrent une absence de différence dans l'ADN proviral VIH entre le sang périphérique et les réservoirs tels que GALT ou ganglions lymphatiques (Imamichi et al. 2011; Evering et al. 2012; Josefsson, von Stockenstrom, et al. 2013; van Zyl, Bale, and Kearney 2018; William R. McManus et al. 2018). En revanche, ces études s'attachent à regarder les différences d'un point de vue phylogénétique et ne s'intéressent pas aux modifications des peptides viraux et donc des potentiels épitopes. Grâce à notre partenariat avec Montréal, nous disposons de 16 couples d'ADN proviraux issus de sang total et de biopsies intestinales. Nous sommes en cours de séquençage NGS sur le full génome du VIH-1 (Ode et al. 2015; Alves et al. 2017) pour comparer les provirus dans les deux compartiments. Si les séquences virales sont différentes et induisent une modification du peptide, donc du score MHC IC₅₀, une telle stratégie vaccinale pourra être remise en question.

Dans la deuxième phase de ce projet Provir/Latitude 45, nous allons au moyen de tests d'immunologie fonctionnelle valider *in vitro* ce cocktail peptidique issu de la liste « Optimal_Provir ». Pour cela, pour des patients inclus au CHU de Bordeaux, des ELISpots vont être réalisés en partenariat avec le Groupe de Recherche en Immunologie Clinique (GRIC). Des tétramères de CMH vont aussi être effectués afin de confirmer *in vitro*, l'immunogénicité prédite *in silico* entre les HLA et les peptides qui ne sont pas déjà connus : Pol02, Pol03, Pol05, Pol06, Pol07 et Pol10.

Enfin, comme cela est fait pour plusieurs constructions vaccinales (Liu et al. 2009; Dereuddre-Bosquet et al. 2015) ; des essais chez l'animal de ce cocktail vaccinal sont envisagés afin de vérifier l'effet immunologique dans un premier temps puis l'efficacité au niveau de la réponse virologique. Un protocole est en cours de rédaction avec le Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA) pour des essais chez le macaque.

Le travail présenté dans ce mémoire, prenant en compte la variabilité virale et immunogénétique des patients, a permis d'identifier de nouveaux épitopes susceptibles d'induire une réponse CTL. Même si les études cellulaires et chez l'animal sont nécessaires pour valider leur efficacité sur le VIH, ces données *in silico* présentent un grand intérêt dans le cadre d'une nouvelle conception de vaccination thérapeutique.

BIBLIOGRAPHIE

VI. Bibliographie

- Achenbach, Chad J., Lambert Assoumou, Steven G. Deeks, Timothy J. Wilkin, Baiba Berzins, Joseph P. Casazza, Sidonie Lambert-Niclot, et al. 2015. "Effect of Therapeutic Intensification Followed by HIV DNA Prime and RAd5 Boost Vaccination on HIV-Specific Immunity and HIV Reservoir (EraMune 02): A Multicentre Randomised Clinical Trial." *The Lancet. HIV* 2 (3): e82-91. [https://doi.org/10.1016/S2352-3018\(15\)00026-0](https://doi.org/10.1016/S2352-3018(15)00026-0).
- Alves, Brunna M., Juliana D. Siqueira, Marianne M. Garrido, Ornella M. Botelho, Isabel M. Prellwitz, Sayonara R. Ribeiro, Esmeralda A. Soares, and Marcelo A. Soares. 2017. "Characterization of HIV-1 Near Full-Length Proviral Genome Quasispecies from Patients with Undetectable Viral Load Undergoing First-Line HAART Therapy." *Viruses* 9 (12). <https://doi.org/10.3390/v9120392>.
- Andreatta, Massimo, and Morten Nielsen. 2016. "Gapped Sequence Alignment Using Artificial Neural Networks: Application to the MHC Class I System." *Bioinformatics (Oxford, England)* 32 (4): 511–17. <https://doi.org/10.1093/bioinformatics/btv639>.
- Assoumou, L., C. Charpentier, P. Recordon-Pinson, M. Grude, C. Pallier, L. Morand-Joubert, S. Fafi-Kremer, et al. 2017. "Prevalence of HIV-1 Drug Resistance in Treated Patients with Viral Load >50 Copies/ML: A 2014 French Nationwide Study." *The Journal of Antimicrobial Chemotherapy* 72 (6): 1769–73. <https://doi.org/10.1093/jac/dkx042>.
- Bailey, Justin R., Ahmad R. Sedaghat, Tara Kieffer, Timothy Brennan, Patricia K. Lee, Megan Wind-Rotolo, Christine M. Haggerty, et al. 2006. "Residual Human Immunodeficiency Virus Type 1 Viremia in Some Patients on Antiretroviral Therapy Is Dominated by a Small Number of Invariant Clones Rarely Found in Circulating CD4+ T Cells." *Journal of Virology* 80 (13): 6441–57. <https://doi.org/10.1128/JVI.00591-06>.
- Barre-Sinoussi, F., J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Daugey, et al. 1983. "Isolation of a T-Lymphotropic Retrovirus from a Patient at Risk for Acquired Immune Deficiency Syndrome (AIDS)." *Science (New York, N.Y.)* 220 (4599): 868–71.
- Campbell, Edward M., and Thomas J. Hope. 2015. "HIV-1 Capsid: The Multifaceted Key Player in HIV-1 Infection." *Nature Reviews. Microbiology* 13 (8): 471–83. <https://doi.org/10.1038/nrmicro3503>.
- Chomont, Nicolas, Mohamed El-Far, Petronela Ancuta, Lydie Trautmann, Francesco A Procopio, Bader Yassine-Diab, Genevieve Boucher, et al. 2009. "HIV Reservoir Size and Persistence Are Driven by T Cell Survival and Homeostatic Proliferation." *Nat Med* 15 (8): 893–900. <https://doi.org/10.1038/nm.1972>.
- Chun, Tae-Wook, and Anthony S. Fauci. 2012. "HIV Reservoirs: Pathogenesis and Obstacles to Viral Eradication and Cure." *AIDS (London, England)* 26 (10): 1261–68. <https://doi.org/10.1097/QAD.0b013e328353f3f1>.
- Chun, Tae-Wook, Danielle Murray, J. Shawn Justement, Claire W. Hallahan, Susan Moir, Colin Kovacs, and Anthony S. Fauci. 2011a. "Relationship between Residual Plasma Viremia and the Size of HIV Proviral DNA Reservoirs in Infected Individuals Receiving Effective Antiretroviral Therapy." *The Journal of Infectious Diseases* 204 (1): 135–38. <https://doi.org/10.1093/infdis/jir208>.
- Clavel, F., M. Guyader, D. Guetard, M. Salle, L. Montagnier, and M. Alizon. 1986. "Molecular Cloning and Polymorphism of the Human Immune Deficiency Virus Type 2." *Nature* 324 (6098): 691–95. <https://doi.org/10.1038/324691a0>.
- "Core Epidemiology Slides." 2017. UNAIDS.

- Deeks, Steven G., Brigitte Autran, Ben Berkhout, Monsef Benkirane, Scott Cairns, Nicolas Chomont, Tae-Wook Chun, et al. 2012. "Towards an HIV Cure: A Global Scientific Strategy." *Nature Reviews. Immunology* 12 (8): 607–14. <https://doi.org/10.1038/nri3262>.
- Deeks, Steven G., Sharon R. Lewin, Anna Laura Ross, Jintanat Ananworanich, Monsef Benkirane, Paula Cannon, Nicolas Chomont, et al. 2016. "International AIDS Society Global Scientific Strategy: Towards an HIV Cure 2016." *Nature Medicine* 22 (8): 839–50. <https://doi.org/10.1038/nm.4108>.
- Deng, Kai, Mihaela Perteau, Anthony Rongvaux, Leyao Wang, Christine M. Durand, Gabriel Ghiaur, Jun Lai, et al. 2015. "Broad CTL Response Is Required to Clear Latent HIV-1 Due to Dominance of Escape Mutations." *Nature* 517 (7534): 381–85. <https://doi.org/10.1038/nature14053>.
- Dereuddre-Bosquet, Nathalie, Marie-Laurence Baron, Vanessa Contreras, Leslie Gosse, Isabelle Mangeot, Frederic Martinon, Rahima Yousfi, Pascal Clayette, Yves Levy, and Roger Le Grand. 2015. "HIV Specific Responses Induced in Nonhuman Primates with ANRS HIV-Lipo-5 Vaccine Combined with RMVA-HIV Prime or Boost Immunizations." *Vaccine* 33 (20): 2354–59. <https://doi.org/10.1016/j.vaccine.2015.03.032>.
- Descamps, Diane, Lambert Assoumou, Marie-Laure Chaix, Antoine Chaillon, Sophie Pakianather, Alexis de Rougemont, Alexandre Storto, et al. 2013. "National Sentinel Surveillance of Transmitted Drug Resistance in Antiretroviral-Naive Chronically HIV-Infected Patients in France over a Decade." *The Journal of Antimicrobial Chemotherapy* 68 (11): 2626–31. <https://doi.org/10.1093/jac/dkt238>.
- Eisele, Evelyn, and Robert F. Siliciano. 2012. "Redefining the Viral Reservoirs That Prevent HIV-1 Eradication." *Immunity* 37 (3): 377–88. <https://doi.org/10.1016/j.immuni.2012.08.010>.
- Evering, Teresa H., Saurabh Mehandru, Paul Racz, Klara Tenner-Racz, Michael A. Poles, Amir Figueroa, Hiroshi Mohri, and Martin Markowitz. 2012. "Absence of HIV-1 Evolution in the Gut-Associated Lymphoid Tissue from Patients on Combination Antiviral Therapy Initiated during Primary Infection." *PLoS Pathogens* 8 (2). <https://doi.org/10.1371/journal.ppat.1002506>.
- Fleury, Herve, Camille Tumioto, Pantxika Bellecave, and Patricia Recordon-Pinson. 2018. "Therapeutic Vaccine Against HIV, Viral Variability, Cytotoxic T Lymphocyte Epitopes, and Genetics of Patients." *AIDS Research and Human Retroviruses* 34 (1): 27–30. <https://doi.org/10.1089/aid.2017.0175>.
- Frangé, Pierre, Lambert Assoumou, Diane Descamps, Antoine Cheret, Cecile Goujard, Laurent Tran, Marine Gousset, et al. 2015. "HIV-1 Subtype B-Infected MSM May Have Driven the Spread of Transmitted Resistant Strains in France in 2007-12: Impact on Susceptibility to First-Line Strategies." *The Journal of Antimicrobial Chemotherapy* 70 (7): 2084–89. <https://doi.org/10.1093/jac/dkv049>.
- Frangé, Pierre, Julie Galimand, Cecile Goujard, Christiane Deveau, Jade Ghosn, Christine Rouzioux, Laurence Meyer, and Marie-Laure Chaix. 2009. "High Frequency of X4/DM-Tropic Viruses in PBMC Samples from Patients with Primary." *The Journal of Antimicrobial Chemotherapy* 64 (1): 135–41. <https://doi.org/10.1093/jac/dkp151>.
- Frey, Sharon E., Laurence Peiperl, M. Juliana McElrath, Spyros Kalams, Paul A. Goepfert, Michael C. Keefer, Lindsey R. Baden, et al. 2014. "Phase I/II Randomized Trial of Safety and Immunogenicity of LIPO-5 Alone." *Clinical and Vaccine Immunology* : CVI 21 (11): 1589–99. <https://doi.org/10.1128/CVI.00450-14>.
- Gao, X., G. W. Nelson, P. Karacki, M. P. Martin, J. Phair, R. Kaslow, J. J. Goedert, et al. 2001. "Effect of a Single Amino Acid Change in MHC Class I Molecules on the Rate

- of Progression to AIDS.” *The New England Journal of Medicine* 344 (22): 1668–75. <https://doi.org/10.1056/NEJM200105313442203>.
- Garcia, Felipe, Nuria Climent, Alberto C. Guardo, Cristina Gil, Agathe Leon, Brigitte Autran, Jeffrey D. Lifson, et al. 2013. “A Dendritic Cell-Based Vaccine Elicits T Cell Responses Associated with Control of HIV-1 Replication.” *Science Translational Medicine* 5 (166): 166ra2. <https://doi.org/10.1126/scitranslmed.3004682>.
- Goulder, Philip J. R., and Bruce D. Walker. 2012. “HIV and HLA Class I: An Evolving Relationship.” *Immunity* 37 (3): 426–40. <https://doi.org/10.1016/j.immuni.2012.09.005>.
- H. Fleury. 2009. *Virologie Humaine*. 5 Ème Édition-Masson.
- Hemelaar, Joris. 2013. “Implications of HIV Diversity for the HIV-1 Pandemic.” *The Journal of Infection* 66 (5). <https://doi.org/10.1016/j.jinf.2012.10.026>.
- Ho, D. D., A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, and M. Markowitz. 1995. “Rapid Turnover of Plasma Virions and CD4 Lymphocytes in HIV-1 Infection.” *Nature* 373 (6510): 123–26. <https://doi.org/10.1038/373123a0>.
- Hoof, Ilka, Bjoern Peters, John Sidney, Lasse Eggers Pedersen, Alessandro Sette, Ole Lund, Soren Buus, and Morten Nielsen. 2009. “NetMHCpan, a Method for MHC Class I Binding Prediction beyond Humans.” *Immunogenetics* 61 (1): 1–13. <https://doi.org/10.1007/s00251-008-0341-z>.
- Imamichi, Hiromi, Gerald Degray, Robin L. Dewar, Peter Mannon, Michael Yao, Cheryl Chairez, Irini Sereti, and Joseph A. Kovacs. 2011. “Lack of Compartmentalization of HIV-1 Quasispecies between the Gut and Peripheral Blood Compartments.” *The Journal of Infectious Diseases* 204 (2): 309–14. <https://doi.org/10.1093/infdis/jir259>.
- Jacobson, Jeffrey M., Jean-Pierre Routy, Seth Welles, Mark DeBenedette, Irina Tcherepanova, Jonathan B. Angel, David M. Asmuth, et al. 2016. “Dendritic Cell Immunotherapy for HIV-1 Infection Using Autologous HIV-1 RNA: A Randomized, Double-Blind, Placebo-Controlled Clinical Trial.” *Journal of Acquired Immune Deficiency Syndromes (1999)* 72 (1): 31–38. <https://doi.org/10.1097/QAI.0000000000000926>.
- John M Coffin, Stephen H Hughes, and Harold E Varmus. 1997. *Retroviruses*. Cold Spring Harbor Laboratory Press(NY). <http://www.ncbi.nlm.nih.gov/books/NBK19376/?report=classic>.
- Jones, R. Brad, and Bruce D. Walker. 2016. “HIV-Specific CD8(+) T Cells and HIV Eradication.” *The Journal of Clinical Investigation* 126 (2): 455–63. <https://doi.org/10.1172/JCI80566>.
- Josefsson, Lina, Sarah Palmer, Nuno R. Faria, Philippe Lemey, Joseph Casazza, David Ambrozak, Mary Kearney, et al. 2013. “Single Cell Analysis of Lymph Node Tissue from HIV-1 Infected Patients Reveals That the Majority of CD4+ T-Cells Contain One HIV-1 DNA Molecule.” *PLoS Pathogens* 9 (6). <https://doi.org/10.1371/journal.ppat.1003432>.
- Josefsson, Lina, Susanne von Stockenstrom, Nuno R. Faria, Elizabeth Sinclair, Peter Bacchetti, Maudi Killian, Lorrie Epling, et al. 2013. “The HIV-1 Reservoir in Eight Patients on Long-Term Suppressive Antiretroviral Therapy Is Stable with Few Genetic Changes over Time.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (51): E4987-4996. <https://doi.org/10.1073/pnas.1308313110>.
- Kaslow, R. A., M. Carrington, R. Apple, L. Park, A. Munoz, A. J. Saah, J. J. Goedert, et al. 1996. “Influence of Combinations of Human Major Histocompatibility Complex Genes on the Course of HIV-1 Infection.” *Nature Medicine* 2 (4): 405–11.

- Kawashima, Yuka, Katja Pfafferott, John Frater, Philippa Matthews, Rebecca Payne, Marylyn Addo, Hiroyuki Gatanaga, et al. 2009. "Adaptation of HIV-1 to Human Leukocyte Antigen Class I." *Nature* 458 (7238): 641–45. <https://doi.org/10.1038/nature07746>.
- Kiepiela, Photini, Kholiswa Ngumbela, Christina Thobakgale, Dhanwanthie Ramduth, Isobella Honeyborne, Eshia Moodley, Shabashini Reddy, et al. 2007. "CD8+ T-Cell Responses to Different HIV Proteins Have Discordant Associations with Viral Load." *Nature Medicine* 13 (1): 46–53. <https://doi.org/10.1038/nm1520>.
- Kim, Yohan, Julia Ponomarenko, Zhanyang Zhu, Dorjee Tamang, Peng Wang, Jason Greenbaum, Claus Lundegaard, et al. 2012. "Immune Epitope Database Analysis Resource." *Nucleic Acids Research* 40 (Web Server issue): W525-530. <https://doi.org/10.1093/nar/gks438>.
- Kløverpris, Henrik N, Alasdair Leslie, and Philip Goulder. 2015. "Role of HLA Adaptation in HIV Evolution." *Frontiers in Immunology* 6: 665. <https://doi.org/10.3389/fimmu.2015.00665>.
- Lazaro, Estibaliz, Sasha Blue Godfrey, Pamela Stamegna, Tobi Ogbechie, Christopher Kerrigan, Mei Zhang, Bruce D. Walker, and Sylvie Le Gall. 2009. "Differential HIV Epitope Processing in Monocytes and CD4 T Cells Affects Cytotoxic T Lymphocyte Recognition." *The Journal of Infectious Diseases* 200 (2): 236–43. <https://doi.org/10.1086/599837>.
- Legrand, E., I. Pellegrin, D. Neau, J. L. Pellegrin, J. M. Ragnaud, M. Dupon, B. Guillemain, and H. J. Fleury. 1997. "Course of Specific T Lymphocyte Cytotoxicity, Plasma and Cellular Viral Loads, and Neutralizing Antibody Titers in 17 Recently Seroconverted HIV Type 1-Infected Patients." *AIDS Research and Human Retroviruses* 13 (16): 1383–94.
- Leth, Steffen, Mariane H. Schleimann, Sara K. Nissen, Jesper F. Hojen, Rikke Olesen, Mette E. Graversen, Sofie Jorgensen, et al. 2016. "Combined Effect of Vacc-4x, Recombinant Human Granulocyte Macrophage Colony-Stimulating Factor Vaccination, and Romidepsin on the HIV-1 Reservoir (REDUC): A Single-Arm, Phase 1B/2A Trial." *The Lancet. HIV* 3 (10): e463-472. [https://doi.org/10.1016/S2352-3018\(16\)30055-8](https://doi.org/10.1016/S2352-3018(16)30055-8).
- Levy, Yves, Rodolphe Thiebaut, Monica Montes, Christine Lacabaratz, Louis Sloan, Bryan King, Sophie Perusat, et al. 2014. "Dendritic Cell-Based Therapeutic Vaccine Elicits Polyfunctional HIV-Specific." *European Journal of Immunology* 44 (9): 2802–10. <https://doi.org/10.1002/eji.201344433>.
- Liu, Jinyan, Kara L. O'Brien, Diana M. Lynch, Nathaniel L. Simmons, Annalena La Porte, Ambryce M. Riggs, Peter Abbink, et al. 2009. "Immune Control of an SIV Challenge by a T-Cell-Based Vaccine in Rhesus Monkeys." *Nature* 457 (7225): 87–91. <https://doi.org/10.1038/nature07469>.
- Llano, Anuska, Ashley Williams, Alex Olvera, Sandra Silva-Arrieta, and Christian Brander. 2013. "Best-Characterized HIV-1 CTL Epitopes: The 2013 Update." *HIV Molecular Immunology* 2013.
- Loret, Erwann P., Albert Darque, Elisabeth Jouve, Elvonn A. Loret, Corinne Nicolino-Brunet, Sophie Morange, Elisabeth Castanier, et al. 2016. "Intradermal Injection of a Tat Oyi-Based Therapeutic HIV Vaccine Reduces of 1.5 Log Copies/ML the HIV RNA Rebound Median and No HIV DNA Rebound Following CART Interruption in a Phase I/II Randomized Controlled Clinical Trial." *Retrovirology* 13 (April): 21. <https://doi.org/10.1186/s12977-016-0251-3>.
- Lundegaard, Claus, Kasper Lamberth, Mikkel Harndahl, Soren Buus, Ole Lund, and Morten Nielsen. 2008. "NetMHC-3.0: Accurate Web Accessible Predictions of Human, Mouse and Monkey MHC Class I Affinities for Peptides of Length 8-11." *Nucleic*

- Acids Research* 36 (Web Server issue): W509-512.
<https://doi.org/10.1093/nar/gkn202>.
- Maartens, Gary, Connie Celum, and Sharon R. Lewin. 2014. "HIV Infection: Epidemiology, Pathogenesis, Treatment, and Prevention." *Lancet (London, England)* 384 (9939): 258–71. [https://doi.org/10.1016/S0140-6736\(14\)60164-1](https://doi.org/10.1016/S0140-6736(14)60164-1).
- Maiers, Martin, Loren Gragert, and William Klitz. 2007. "High-Resolution HLA Alleles and Haplotypes in the United States Population." *Human Immunology* 68 (9): 779–88. <https://doi.org/10.1016/j.humimm.2007.04.005>.
- Mangeat, Bastien, Priscilla Turelli, Gersende Caron, Marc Friedli, Luc Perrin, and Didier Trono. 2003. "Broad Antiretroviral Defence by Human APOBEC3G through Lethal Editing of Nascent Reverse Transcripts." *Nature* 424 (6944): 99–103. <https://doi.org/10.1038/nature01709>.
- Martin Maiers, Loren Gragert, William Klitz. n.d. "High Resolution HLA Alleles and Haplotypes in the US Populations as of 2006." <https://bioinformatics.bethematchclinical.org/Templates/DisplayPageContent?id=9376>.
- McMichael, Andrew J., and Wayne C. Koff. 2014. "Vaccines That Stimulate T Cell Immunity to HIV-1: The next Step." *Nature Immunology* 15 (4): 319–22. <https://doi.org/10.1038/ni.2844>.
- Migueles, S. A., M. S. Sabbaghian, W. L. Shupert, M. P. Bettinotti, F. M. Marincola, L. Martino, C. W. Hallahan, et al. 2000. "HLA B*5701 Is Highly Associated with Restriction of Virus Replication in a Subgroup of HIV-Infected Long Term Nonprogressors." *Proceedings of the National Academy of Sciences of the United States of America* 97 (6): 2709–14. <https://doi.org/10.1073/pnas.050567397>.
- Migueles, Stephen A., and Mark Connors. 2012. "Small Molecules and Big Killers: The Challenge of Eliminating the Latent HIV Reservoir." *Immunity* 36 (3): 320–21. <https://doi.org/10.1016/j.immuni.2012.03.006>.
- Mothe, Beatriz, Nuria Climent, Montserrat Plana, Miriam Rosas, Jose Luis Jimenez, Maria Angeles Munoz-Fernandez, Maria C. Puertas, et al. 2015. "Safety and Immunogenicity of a Modified Vaccinia Ankara-Based HIV-1 Vaccine (MVA-B) in HIV-1-Infected Patients Alone or in Combination with a Drug to Reactivate Latent HIV-1." *The Journal of Antimicrobial Chemotherapy* 70 (6): 1833–42. <https://doi.org/10.1093/jac/dkv046>.
- Myriam Labalette, Siamak Bahram, Marie-Christine Béné. 2011. *Le Complexe Majeur d'histocompatibilité Humain (HLA). Polymorphisme et Présentation Des Antigènes Aux Kymphocytes T*. Ouvrage L2 Immunologie Générale de l'ASSIM, Chapitre 4. www.assim.refer.org/raisil/raisil/L02.html.
- Nielsen, Morten, and Massimo Andreatta. 2016. "NetMHCpan-3.0; Improved Prediction of Binding to MHC Class I Molecules Integrating Information from Multiple Receptor and Peptide Length Datasets." *Genome Medicine* 8 (1): 33. <https://doi.org/10.1186/s13073-016-0288-x>.
- Nielsen, Morten, Claus Lundegaard, Thomas Blicher, Kasper Lamberth, Mikkel Harndahl, Sune Justesen, Gustav Roder, et al. 2007. "NetMHCpan, a Method for Quantitative Predictions of Peptide Binding to Any HLA-A and -B Locus Protein of Known Sequence." *PloS One* 2 (8): e796. <https://doi.org/10.1371/journal.pone.0000796>.
- Nielsen, Morten, Claus Lundegaard, Peder Worning, Sanne Lise Lauemoller, Kasper Lamberth, Soren Buus, Soren Brunak, and Ole Lund. 2003. "Reliable Prediction of T-Cell Epitopes Using Neural Networks with Novel Sequence Representations." *Protein Science : A Publication of the Protein Society* 12 (5): 1007–17. <https://doi.org/10.1110/ps.0239403>.

- Ode, Hirotaka, Masakazu Matsuda, Kazuhiro Matsuoka, Atsuko Hachiya, Junko Hattori, Yumiko Kito, Yoshiyuki Yokomaku, Yasumasa Iwatani, and Wataru Sugiura. 2015. "Quasispecies Analyses of the HIV-1 Near-Full-Length Genome With Illumina MiSeq." *Frontiers in Microbiology* 6: 1258. <https://doi.org/10.3389/fmicb.2015.01258>.
- Palmer, Sarah, Frank Maldarelli, Ann Wiegand, Barry Bernstein, George J. Hanna, Scott C. Brun, Dale J. Kempf, John W. Mellors, John M. Coffin, and Martin S. King. 2008. "Low-Level Viremia Persists for at Least 7 Years in Patients on Suppressive Antiretroviral Therapy." *Proceedings of the National Academy of Sciences of the United States of America* 105 (10): 3879–84. <https://doi.org/10.1073/pnas.0800050105>.
- Palmer, Sarah, Ann P. Wiegand, Frank Maldarelli, Holly Bazmi, JoAnn M. Mican, Michael Polis, Robin L. Dewar, et al. 2003. "New Real-Time Reverse Transcriptase-Initiated PCR Assay with Single-Copy Sensitivity for Human Immunodeficiency Virus Type 1 RNA in Plasma." *Journal of Clinical Microbiology* 41 (10): 4531–36.
- Pantaleo, G., C. Graziosi, and A. S. Fauci. 1993. "New Concepts in the Immunopathogenesis of Human Immunodeficiency Virus Infection." *The New England Journal of Medicine* 328 (5): 327–35. <https://doi.org/10.1056/NEJM199302043280508>.
- Papuchon, Jennifer, Patricia Pinson, Gwenda-Line Guidicelli, Pantxika Bellecave, Rejean Thomas, Roger LeBlanc, Sandrine Reigadas, et al. 2014. "Kinetics of HIV-1 CTL Epitopes Recognized by HLA I Alleles in HIV-Infected Individuals at Times near Primary Infection: The Provir/Latitude45 Study." *PloS One* 9 (6). <https://doi.org/10.1371/journal.pone.0100452>.
- Papuchon, Jennifer, Patricia Pinson, Estibaliz Lazaro, Sandrine Reigadas, Gwendaline Guidicelli, Jean-Luc Taupin, Didier Neau, and Herve Fleury. 2013. "Resistance Mutations and CTL Epitopes in Archived HIV-1 DNA of Patients on Antiviral Treatment: Toward a New Concept of Vaccine." *PloS One* 8 (7). <https://doi.org/10.1371/journal.pone.0069029>.
- Paul, Sinu, Daniela Weiskopf, Michael A. Angelo, John Sidney, Bjoern Peters, and Alessandro Sette. 2013. "HLA Class I Alleles Are Associated with Peptide-Binding Repertoires of Different Size, Affinity, and Immunogenicity." *Journal of Immunology (Baltimore, Md. : 1950)* 191 (12): 5831–39. <https://doi.org/10.4049/jimmunol.1302101>.
- Pereyra, Florencia, Xiaoming Jia, Paul J. McLaren, Amalio Telenti, Paul I. W. de Bakker, Bruce D. Walker, Stephan Ripke, et al. 2010. "The Major Genetic Determinants of HIV-1 Control Affect HLA Class I Peptide Presentation." *Science (New York, N.Y.)* 330 (6010): 1551–57. <https://doi.org/10.1126/science.1195271>.
- Peterlin, B. Matija, and Didier Trono. 2003. "Hide, Shield and Strike Back: How HIV-Infected Cells Avoid Immune Eradication." *Nature Reviews. Immunology* 3 (2). <https://doi.org/10.1038/nri998>.
- Peters, Bjoern, and Alessandro Sette. 2005. "Generating Quantitative Models Describing the Sequence Specificity of Biological Processes with the Stabilized Matrix Method." *BMC Bioinformatics* 6 (May): 132. <https://doi.org/10.1186/1471-2105-6-132>.
- Pollack, H., M. X. Zhan, J. T. Safrit, S. H. Chen, G. Rochford, P. Z. Tao, R. Koup, K. Krasinski, and W. Borkowsky. 1997. "CD8+ T-Cell-Mediated Suppression of HIV Replication in the First Year of Life: Association with Lower Viral Load and Favorable Early Survival." *AIDS (London, England)* 11 (1).
- Pollard, Richard B., Jurgen K. Rockstroh, Giuseppe Pantaleo, David M. Asmuth, Barry Peters, Adriano Lazzarin, Felipe Garcia, et al. 2014. "Safety and Efficacy of the Peptide-Based Therapeutic Vaccine for HIV-1, Vacc-4x: A Phase 2 Randomised,

- Double-Blind, Placebo-Controlled Trial.” *The Lancet. Infectious Diseases* 14 (4): 291–300. [https://doi.org/10.1016/S1473-3099\(13\)70343-8](https://doi.org/10.1016/S1473-3099(13)70343-8).
- “Rapport_Morlat_2013_Mise_en_ligne.Pdf.” 2014. December 5, 2014. http://www.corevih-pdl.fr/images/pieces_jointes/Rapport_Morlat_2013_Mise_en_ligne.pdf.
- Raymond, Stephanie, Pierre Delobel, Maud Mavigner, Michelle Cazabat, Stephanie Encinas, Corinne Souyris, Patrick Bruel, et al. 2010. “CXCR4-Using Viruses in Plasma and Peripheral Blood Mononuclear Cells during Primary HIV-1 Infection and Impact on Disease Progression.” *AIDS (London, England)* 24 (15): 2305–12. <https://doi.org/10.1097/QAD.0b013e32833e50bb>.
- Recordon-Pinson, Patricia, Brunna Alves, Camille Tumiotto, Pantxika Bellecave, Fabrice Bonnet, Didier Neau, Esmeralda A. Soares, Marcelo A. Soares, and Herve J. Fleury. 2018. “A New HIV-1 Circulating Recombinant Form (CRF98_cpx) between CRF06_cpx and Subtype B Identified in Southwestern France.” *AIDS Research and Human Retroviruses*, June. <https://doi.org/10.1089/AID.2018.0122>.
- Roberts, J. D., K. Bebenek, and T. A. Kunkel. 1988. “The Accuracy of Reverse Transcriptase from HIV-1.” *Science (New York, N.Y.)* 242 (4882): 1171–73.
- Ruelas, Debbie S., and Warner C. Greene. 2013. “An Integrated Overview of HIV-1 Latency.” *Cell* 155 (3): 519–29. <https://doi.org/10.1016/j.cell.2013.09.044>.
- Rugeles, Maria T., Paula A. Velilla, and Carlos J. Montoya. 2011. “Mechanisms of Human Natural Resistance to HIV: A Summary of Ten Years of Research in the Colombian Population.” *Biomedica : Revista Del Instituto Nacional de Salud* 31 (2): 269–80. <https://doi.org/10.1590/S0120-41572011000200016>.
- Shan, Liang, Kai Deng, Neeta S. Shroff, Christine M. Durand, S. Alireza Rabi, Hung-Chih Yang, Hao Zhang, Joseph B. Margolick, Joel N. Blankson, and Robert F. Siliciano. 2012. “Stimulation of HIV-1-Specific Cytolytic T Lymphocytes Facilitates Elimination of Latent Viral Reservoir after Virus Reactivation.” *Immunity* 36 (3). <https://doi.org/10.1016/j.immuni.2012.01.014>.
- Sidney, John, Erika Assarsson, Carrie Moore, Sandy Ngo, Clemencia Pinilla, Alessandro Sette, and Bjoern Peters. 2008. “Quantitative Peptide Binding Motifs for 19 Human and Mouse MHC Class I Molecules Derived Using Positional Scanning Combinatorial Peptide Libraries.” *Immunome Research* 4 (January): 2. <https://doi.org/10.1186/1745-7580-4-2>.
- Smyth, Redmond P., Miles P. Davenport, and Johnson Mak. 2012. “The Origin of Genetic Diversity in HIV-1.” *Virus Research* 169 (2): 415–29. <https://doi.org/10.1016/j.virusres.2012.06.015>.
- Smyth, Redmond P., and Matteo Negroni. 2016. “A Step Forward Understanding HIV-1 Diversity.” *Retrovirology* 13 (April): 27. <https://doi.org/10.1186/s12977-016-0259-8>.
- Sung, Julia A., Sharon Lam, Carolina Garrido, Nancie Archin, Cliona M. Rooney, Catherine M. Bollard, and David M. Margolis. 2015. “Expanded Cytotoxic T-Cell Lymphocytes Target the Latent HIV Reservoir.” *The Journal of Infectious Diseases*, January. <https://doi.org/10.1093/infdis/jiv022>.
- Taylor, Barbara S., Magdalena E. Sobieszczyk, Francine E. McCutchan, and Scott M. Hammer. 2008. “The Challenge of HIV-1 Subtype Diversity.” *The New England Journal of Medicine* 358 (15): 1590–1602. <https://doi.org/10.1056/NEJMra0706737>.
- Taylor, Jesse E., and Bette T. Korber. 2005. “HIV-1 Intra-Subtype Superinfection Rates: Estimates Using a Structured Coalescent with Recombination.” *Infection, Genetics and Evolution : Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* 5 (1): 85–95. <https://doi.org/10.1016/j.meegid.2004.07.001>.

- Tebit, Denis M., and Eric J. Arts. 2011. "Tracking a Century of Global Expansion and Evolution of HIV to Drive Understanding and to Combat Disease." *The Lancet. Infectious Diseases* 11 (1): 45–56. [https://doi.org/10.1016/S1473-3099\(10\)70186-9](https://doi.org/10.1016/S1473-3099(10)70186-9).
- Thomas J. Kindt, Richard A. Goldsby, Barbara A. Osborne. 2008. *Immunologie. Le Cours de Janis Kuby Avec Questions de Révision*. 6ème.
- Tumiotto, Camille, Pantxika Bellecave, Patricia Recordon-Pinson, Alexi Groppi, Macha Nikolski, and Herve Fleury. 2018. "Diversity of HIV-1 in Aquitaine, Southwestern France, 2012-2016." *AIDS Research and Human Retroviruses* 34 (5): 471–73. <https://doi.org/10.1089/AID.2017.0298>.
- Tumiotto, Camille, Lionel Riviere, Pantxika Bellecave, Patricia Recordon-Pinson, Alice Vilain-Parce, Gwenda-Line Guidicelli, and Herve Fleury. 2017. "Sanger and Next-Generation Sequencing Data for Characterization of CTL Epitopes in Archived HIV-1 Proviral DNA." *PloS One* 12 (9): e0185211. <https://doi.org/10.1371/journal.pone.0185211>.
- "UNAIDS DATA 2017." n.d. http://www.unaids.org/en/resources/documents/2017/2017_data_book.
- Walker, Bruce, and Andrew McMichael. 2012. "The T-Cell Response to HIV." *Cold Spring Harbor Perspectives in Medicine* 2 (11). <https://doi.org/10.1101/cshperspect.a007054>.
- William R. McManus, Michael J. Bale, Jonathan Spindler, Ann Wiegand, Andrew Musick, Xiaolin Wu, David Wells, et al. 2018. "No Evidence For Ongoing HIV Replication In Lymph Nodes During Suppressive ART." In . Vol. Poster 70. Boston.
- Zyl, Gert van, Michael J. Bale, and Mary F. Kearney. 2018. "HIV Evolution and Diversity in ART-Treated Patients." *Retrovirology* 15 (1): 14. <https://doi.org/10.1186/s12977-018-0395-4>.

ANNEXES

VII. Annexes

Articles publiés

A new HIV-1 circulating recombinant form (CRF98_cpx) between CRF06_cpx and subtype B identified in Southwestern France, *Recordon-Pinson P, Alves B, Tumiotto C, Bellecave P, Bonnet F, Neau D, Soares EA, Soares MA, Fleury HJ*, *AIDS Res Hum Retroviruses*. 2018 Jun 27. doi: 10.1089/AID.2018.0122.

Diversity of HIV-1 in Aquitaine, South-Western France, 2012-2016, *C. Tumiotto, P. Bellecave, P. Pinson, A. Groppi, M. Nikolski and H. Fleury*, *AIDS Res Hum Retroviruses*. 2018 Feb 14.

Therapeutic Vaccine Against HIV, Viral Variability, Cytotoxic T Lymphocyte Epitopes, and Genetics of Patients, *Fleury H, Tumiotto C, Bellecave P, Recordon-Pinson P.*, *AIDS Res Hum Retroviruses*. 2017 Oct 12.

Sanger and NGS data for CTL epitopes characterization in archived HIV-1 proviral DNA, *C. Tumiotto, L. Riviere, P. Bellecave, P. Pinson, A. Vilain-Parcé, G. Guidicelli, H. Fleury and the Provir/Latitude 45 study group*, *PLoS One*. 2017 Sep 21.

Therapeutic Vaccine Against HIV, Viral Variability, Cytotoxic T Lymphocyte Epitopes, and Genetics of Patients

Herve Fleury, Camille Tumiotta, Pantxika Bellecave, and Patricia Recordon-Pinson

Abstract

The scientific and medical community is seeking to cure HIV. Several pathways have been or are being explored including therapeutic vaccination. Viroimmunological studies on primary infection as well as on elite controllers have demonstrated the importance of the cytotoxic CD8 response and have mainly oriented research on vaccine constructs toward this type of response. The results of these trials are clearly not commensurate with the hope placed in them. Might there be one or more uncontrolled variables? The genetics of patients need to be taken into consideration, especially their human lymphocyte antigen (HLA) alleles. There is a need to find a balance between the conservation of cytotoxic T lymphocyte (CTL) epitopes and presentation by HLA alleles. The pathway is a narrow one between adaptation of the virus to HLA I restriction and the definition of conserved proviral CTL epitopes presentable by HLA I alleles. It is likely that the genetics of patients will need to be considered for HIV-1 vaccine studies and that multidisciplinary collaboration will be essential in this field of infectious diseases.

Keywords: HIV-1, CTL epitopes, HLA alleles, vaccine trials, immunogenetics, viral cure

Introduction

HIV-1 INFECTION IS a chronic infection that induces a steady decrease in CD4⁺ T lymphocytes and leads to AIDS. The emergence of tritherapies and diversification of antiretroviral agents have made it possible to control viral replication in treated patients, resulting in the restitution of their immune defenses and allowing them to lead a normal life. However, these treatments are costly and have side effects. Trials in which treatment was discontinued were followed by a resumption of viral replication.¹ Indeed, after the primary infection phase, HIV-1 establishes a cell reservoir, the CD4⁺ memory T cells, where it becomes latent,² and anatomical reservoirs such as the central nervous and genital systems, the lymph nodes that are not fully permeable to some antiretroviral molecules, and the gastrointestinal associated lymphoid tissues. Therefore, the virus replication resumes significantly while the patient is experiencing a therapeutic interruption; moreover, it seems that residual replication persists in the reservoir cells during the control phase of viral replication at loads less than the quantification thresholds.³

The scientific and medical community is now seeking to cure HIV. Several pathways have been or are being explored⁴: therapeutic intensification, latency reversal agents such as histone deacetylase that force the provirus to express itself with a resumption of productive viral replication, and exposure to cytotoxic CD8⁺ T lymphocytes (“shock and

kill”), monoclonal antibodies, immunomodulators (such as IL7 and PD-1 inhibitor), gene therapy (delivery in infected cells of zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), or clustered regularly interspaced short palindromic repeats (CRISPR) nucleases/Cas9), and therapeutic vaccination. At present, only one patient in the world has been successfully cured of HIV. This is the “Berlin patient” who, during his exceptional clinical and therapeutic history, received a graft of delta 32 CCR5 lymphocytes that could no longer serve as a target for viral replication.⁵

Therapeutic Vaccination

It was Salk who first and very early in the history of the discovery of AIDS and HIV-1 had the idea to vaccinate HIV-1-positive patients against the virus using inactivated and viral particles depleted in Gp120 (HIV-1 Immunogen Remune).⁶ This attempt was not successful. Since then, and more recently, other vaccine constructs have been tested, of which this article gives a summary. First, viroimmunological studies on primary infection as well as on long-term nonprogressing patients, called elite controllers, have demonstrated the importance of the CD8⁺ T cell cytotoxic response⁷ and have mainly oriented research on vaccine constructs toward this type of response, without ignoring antibody-dependent cellular toxicity neutralizing antibodies. A special comment must be

made here on the CD4⁺ T cell response; this response is restricted by HLA II alleles and is of help to the synthesis of antibodies by B lymphocytes and to the CD8⁺ T cell response; some vaccine constructions have integrated antigens for both CD4⁺ and CD8⁺ T responses, but due to the general knowledge that CD8⁺ T cells are highly efficient for viral clearance, most of the therapeutic vaccines have focused on stimulation of CD8⁺ rather than CD4⁺ T cell response^{8,9}

Table 1 summarizes the different vaccine constructions and the main trials in HIV-infected patients on successful combination ART.

Viral/Proviral Variability and Patient Genetics

The results of these trials are clearly not commensurate with the hope placed in them. Might there be one or more

uncontrolled variables? It should first be emphasized that most vaccine constructs are based on generic HIV-1 viruses and that circulating strains may exhibit variations, especially regarding the CTL epitopes that must be presentable to antigens linked to HLA I alleles. In infected subjects and owing to a process of drift induced by reverse transcriptase errors and the development of subspecies, the number of mutated CTL epitopes increases with time since the primary infection phase. In contrast, the virus, which multiplies after discontinuation of treatment, originates mainly from the cellular reservoir represented by the CD4⁺ memory T cells, which are reactivated. The virus, therefore, leaves an archived viral DNA that is not strictly identical to the virus that was circulating before the antiretroviral treatment. This means that what has been archived is probably the image of what will be reactivated. In addition, no vaccine construct based on the

TABLE 1. SUMMARY OF THE DIFFERENT VACCINE CONSTRUCTIONS AND OF THE MAIN TRIALS IN HIV-INFECTED PATIENTS ON SUCCESSFUL COMBINATION ART

Recombinant live viruses: canarypox (ALVAC), fowlpox, adenovirus, modified Ankara vaccine (MVA), lentivirus, VSV	MVA expressing gp120+Gag-Pol-Nef clade B ± disulfiram. Increase of Gag-specific T cell response. No impact on rebound and latent reservoir when given alone or in combination with disulfiram. ¹²
Adenovirus+MVA DNA HIV-1 expressing proteins Rev, Nef, Tat, Gag, and Pol DNA+generic lipopeptides DNA+VSV DNA+MVA DNA+adenovirus	DNA prime (Env A, Env B, Env C, Gag B, Pol B, Nef B) followed by rAd5 (Env A, Env B, Env C, Gag B, Pol B, Nef B). Evaluation only of neutralizing antibodies and ADCC. No impact on these parameters and on CD4 ⁺ cell-associated HIV DNA. ¹⁰ Intensification of ART alone or associated with same vaccine as already cited (DNA prime+rAd5). No significant reduction of HIV DNA in the vaccine arm. ¹⁷ Vacc-4X (p24 Gag) (four primary immunizations plus two booster immunizations). Significant decrease of viral load (VL) but no impact on CD4 ⁺ T cell counts and on proportion of patients resuming ART before end of the study. ¹⁴ Vacc-4X (six injections)+GM-CSF+romidepsin ("shock and kill strategy"). No significant decrease of integrated DNA. ¹⁵
Generic peptides containing epitopes restricted by HLA I and II	Three injections of synthetic Tat Oyi protein at four doses, one dose (33 µg/injection) lowering the extent of RNA and DNA rebound. More data expected. ¹⁸
Env C5/gp41 Tat protein	LIPO-5 peptides of French ANRS. Increase of the breadth of immune response+frequency of functional CD4 ⁺ and CD8 ⁺ T cells; 1 log decrease of VL in 50% of patients during ART interruption. ¹³ Three injections of monocyte-derived dendritic cells pulsed with autologous inactivated HIV-1/nonpulsed VL decrease >1 log in 55% of patients at weeks 12 and 24 after ART interruption. ¹¹
Electroporated dendritic cells with generic HIV-1 RNAs Dendritic cells generated <i>in vitro</i> and loaded with generic HIV-1 lipopeptides	mRNA expressing Gag, Vpr, Nef, and Rev. Induction of HIV-specific effector/memory CD8 ⁺ T cell responses but no impact on VL between vaccine and placebo at the end of analytical treatment interruption. ¹⁶
Pulsed dendritic cells with autologous inactivated HIV-1	
Dendritic cells loaded with autologous HIV-1 mRNA	

The right column includes references of the studies cited in the References section of the article and a brief comment on the conclusion of the trials. ADCC, antibody-dependent cellular toxicity; ART, antiretroviral therapy; ANRS, Agence Nationale de Recherche sur le Sida (France); MVA, modified vaccinia ankara.

genetic characteristics of patients has been developed to date. However, whether they are provided by recombinant viruses, plasmid DNA, or directly in the form of lipopeptides, CTL epitopes must be presented by HLA I allele antigens, a process that depends on the genetics of patients. Our team has drawn attention to this discrepancy between the archived epitopes and the theoretical recognition by the HLA alleles,^{19,20} and a very thorough study of a similar nature was published by Siliciano's team.²¹ They raised the issue of the adequacy between archived CTL epitopes and vaccine CTL epitopes, which are preserved epitopes. They have confirmed that, at interruption of ART, the virus resumes from proviral DNA of CD4+ memory T cells; therefore, they have characterized and compared by high-throughput sequencing the CTL epitopes from the archived virus with those from the virus produced by these cells in culture. The aim of such work is to identify nonvariable epitopes in the provirus that can be used to induce a vaccine response. In contrast, proviral sequence studies are promising, even though the majority of proviral sequences do not generate infectious viruses.²²

Of course, the genetics of patients need to be taken into consideration, especially their HLA alleles. Indeed, it can be hypothesized that generic CTL epitopes introduced into vaccine constructs will generate an immune response, but only if the presentation allele is present. There is, therefore, a need to find a balance between the conservation of CTL epitopes and presentation by HLA alleles.

This presentation by HLA alleles is being increasingly understood.²³ Some alleles are associated with a slower evolution of HIV-1 infection such as HLA-B*27 and HLA-B*57. A study of elite controllers confirmed the role of CTL in controlling infection.²⁴ Recently, it was found that the vaccine used in the RV144 preventive vaccination trial was more effective in HLA-A*02 patients.²⁵ To add further complexity to the subject, adaptation of the virus to HLA class I alleles has been demonstrated²⁶: in that study, the authors showed in >2,800 patients that the frequency of their epitopic variants was correlated with their restriction alleles, raising the hypothesis of an adaptation of the virus to class I HLA alleles in a given population. Similarly, a Japanese study showed how an Nef-Y 135F mutant escapes presentation by the HLA-A*24:02 allele, which is expressed by >60% of the Japanese population.²⁷

The pathway is, therefore, a narrow one between adaptation of the virus to HLA I restriction and the definition of conserved proviral CTL epitopes presentable by HLA I alleles. A rational strategy would be to sequence the archived provirus of patients in therapeutic success and then to synthesize peptides corresponding to their presentable CTL epitopes, on the basis of their HLA alleles characterized by molecular biology or even by high-throughput sequencing. This would be tantamount to creating a personalized vaccine. However, the industrial issues are complex, if not insurmountable. Another strategy is to identify CTL epitopes preserved on the basis of the frequency of alleles in specific populations, to synthesize peptide assemblies presentable by the most frequent alleles and haplotypes.²⁸ It is likely that the genetics of patients will need to be considered for HIV-1 vaccine studies and that multidisciplinary collaboration will be essential in this field of infectious diseases.

Author Disclosure Statement

No competing financial interests exist.

References

1. Chun TW, *et al.*: Rebound of plasma viremia following cessation of antiretroviral therapy despite profoundly low levels of HIV reservoir: Implications for eradication. *AIDS* 2010;27:2803–2808.
2. Finzi D, *et al.*: Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science* 1997; 278:1295–1300.
3. Anderson JA, *et al.*: Clonal sequences recovered from plasma from patients with residual HIV-1 viremia and on intensified antiretroviral therapy are identical to replicating viral RNAs recovered from circulating resting CD4+ T cells. *J Virol* 2011;85:5220–5223.
4. Deeks SG, *et al.*: International AIDS Society global scientific strategy: Towards an HIV cure. *Nat Med* 2016;22: 839–850.
5. Siliciano JD, Siliciano RF: Recent developments in the effort to cure HIV infection: Going beyond $N=1$. *J Clin Invest* 2016;126:409–414.
6. Levine AM, *et al.*: Initial studies on active immunization of HIV-infected subjects using a gp120-depleted HIV-1 Immunogen: Long-term follow-up. *J Acquir Immune Defic Syndr Hum Retrovirol* 1996;11:351–364.
7. Mothe B, *et al.*: Definition of the viral targets of protective HIV-1-specific T cell responses. *J Transl Med* 2011;9:208.
8. Walker B, McMichael A: The T-cell response to HIV. *Cold Spring Harb Perspect Med* 2012;2:a007054.
9. Koup R, Douek D: Vaccine design for CD8 T lymphocyte responses. *Cold Spring Harb Perspect Med* 2011;1:a007252.
10. Gach JS, *et al.*: HIV-1-specific antibody response and function after DNA prime and recombinant adenovirus 5 boost HIV vaccine in HIV-infected subjects. *PLoS One* 2016; 11:e0160341.
11. García F, *et al.*: A dendritic cell-based vaccine elicits T cell responses associated with control of HIV-1 replication. *Sci Transl Med* 2013;5:166ra2.
12. Mothe B, *et al.*: Safety and immunogenicity of a modified vaccinia Ankara-based HIV-1 vaccine (MVA-B) in HIV-1-infected patients alone or in combination with a drug to reactivate latent HIV-1. *J Antimicrob Chemother* 2015;70: 1833–1842.
13. Lévy Y, *et al.*: Dendritic cell-based therapeutic vaccine elicits polyfunctional HIV-specific T-cell immunity associated with control of viral load. *Eur J Immunol* 2014;44: 2802–2810.
14. Pollard RB, *et al.*: Safety and efficacy of the peptide-based therapeutic vaccine for HIV-1, Vacc-4x: A phase 2 randomised, double-blind, placebo-controlled trial. *Lancet Infect Dis* 2014;14:291–300.
15. Leth S, *et al.*: Combined effect of Vacc-4x, recombinant human granulocyte macrophage colony-stimulating factor vaccination, and romidepsin on the HIV-1 reservoir (REDUC): A single-arm, phase 1B/2A trial. *Lancet HIV* 2016;3: e463–e472.
16. Jacobson JM, *et al.*: Dendritic cell immunotherapy for HIV-1 infection using autologous HIV-1 RNA: A randomized, double-blind, placebo-controlled clinical trial. *J Acquir Immune Defic Syndr* 2016;1:31–38.
17. Achenbach CJ, *et al.*: Effect of therapeutic intensification followed by HIV DNA prime and rAd5 boost vaccination

- on HIV-specific immunity and HIV reservoir (EraMune 02): A multicentre randomised clinical trial. *Lancet HIV* 2015; 2:e82–e91.
18. Loret EP, *et al.*: Intradermal injection of a Tat Oyi-based therapeutic HIV vaccine reduces of 1.5 log copies/mL the HIV RNA rebound median and no HIV DNA rebound following cART interruption in a phase I/II randomized controlled clinical trial. *Retrovirology* 2016;13:21.
 19. Papuchon J, *et al.*: Resistance mutations and CTL epitopes in archived HIV-1 DNA of patients on antiviral treatment: Toward a new concept of vaccine. *PLoS One* 2013;8:e69029.
 20. Papuchon J, *et al.*: Kinetics of HIV-1 CTL epitopes recognized by HLA I alleles in HIV-infected individuals at times near primary infection: The Provir/Latitude45 study. *PLoS One* 2014;9:e100452.
 21. Deng K, *et al.*: Broad CTL response is required to clear latent HIV-1 due to dominance of escape mutations. *Nature* 2015;517:381–385.
 22. Lorenzi JC, *et al.*: Paired quantitative and qualitative assessment of the replication-competent HIV-1 reservoir and comparison with integrated proviral DNA. *Proc Natl Acad Sci U S A* 2016;113:E7908–E7916.
 23. Goulder PJ, Walker BD: HIV and HLA class I: An evolving relationship. *Immunity* 2012;37:426–440.
 24. Pereyra F, *et al.*: The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. International HIV Controllers Study. *Science* 2010;330:1551–1557.
 25. Gartland AJ, *et al.*: Analysis of HLA A*02 association with vaccine efficacy in the RV144 HIV-1 vaccine trial. *J Virol* 2014;88:8242–8255.
 26. Kawashima Y, *et al.*: Adaptation of HIV-1 to human leukocyte antigen class I. *Nature* 2009;458:641–645.
 27. Katoh J, *et al.*: Rapid HIV-1 disease progression in individuals infected with a virus adapted to its host population. *PLoS One* 2016;11:e0150397.
 28. Maiers M, Gragert L, Klitz W: High-resolution HLA alleles and haplotypes in the United States population. *Hum Immunol* 2007;68:779–788.

Address correspondence to:

Herve Fleury

Laboratoire de Virologie

CHU de Bordeaux et CNRS UMR 5234

Université de Bordeaux

146 rue Leo Saignat

Bordeaux 33076

France

E-mail: herve.fleury@u-bordeaux.fr