

الجمهورية الجزائرية الديمقراطية الشعبية
RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
وزارة التعليم العالي والبحث العلمي
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE
جامعة محمد بوضياف/المسيلة
UNIVERSITÉ MOHAMED BOUDIAF DE M'SILA



FACULTÉ DES SCIENCES

DÉPARTEMENT DE MICROBIOLOGIE ET BIOCHIMIE

MÉMOIRE : Master Académique

FILIÈRE : Sciences Biologiques

OPTION : Biochimie Appliquée

Présenté par :

BOUDIAF Karima & HERZI Souhir

Thème :

La phylogénie de VIH en Algérie
Caractérisation phylogénétique à travers l'ENV

DEVANT LE JURY :

Dr. Kheniche Abdelhakim

MCB

Encadreur

Dr. Yahiaoui Merzoug

MCA

Examineur

Dr. Madi Bouaoudia Nadia

MCB

Examinatrice

Mr. Ben azzi Nabil

Medecin

Co-encadreur

Année universitaire: 2019/2020



REMERCIEMENT

Avant tout on remercie dieu tout puissant de nous avoir donné le privilège, la chance d'étudier et de nous avoir donné force, courage, et patience pour accomplir ce travail. Sans oublier nos parents qui ont veillé sur nous durant toute notre vie.

Nous remercions naturellement nos encadreurs, **Dr. KHENICHE Abdelhakim** et nos Co encadreur **Dr. BENAZI Nabil**. Pour leur orientation éclairée, et leur aide dans l'élaboration de notre mémoire.

Nous remercions les membres du jury **Mdme. Madi Bouaoudia Nadia** D'avoir acceptés de juger notre modeste travail et **Mr. Yahiaoui Merzoug** pour l'honneur qui 'elle nous a fait en acceptant de présider le jury de soutenance.

Nous exprimons nos vifs remerciements à tous ceux qui nous ont aidés de près ou de loin, et encouragé pour la réalisation de ce mémoire spécifiquement le docteur et au même temps le directeur de l'institut pasteur l'annexe de M'SILA Monsieur **BENAZI NABIL**, et les travailleurs présents dans l'institut.

Nous remercions vont également à tous nos enseignants qui ont contribué à notre formation durant toute notre scolarité.

KARIMA & SOUHIR



Dédicace

Je dédie ce modeste travail à :

A mes parents **BOUDIAF KADDOUR** et **ZAKKER FATIHA** .Aucun hommage ne pourrait être à la hauteur de l'amour Dont ils ne cessent de me combler. Que dieu leur procure bonne santé et longue vie.

A celui que j'aime beaucoup et qui m'a soutenue tout au long de ce projet : ma sœur **RADIA**, et bien sûr à mes frères **AMINE** et **HICHAM**, sans oublié mes cousines et spécialement à ma cousine **SOUMIA** et son mari **TAREK** et ses fils **ANES** et **ANIS** que j'aime.

A toute ma famille **BOUDIAF** et **ZAKKER** et toutes mes amis,
A mon binôme **SOUHIR** et toute la famille **HERZI** .

Et à tous ceux qui ont contribué de près ou de loin pour que ce projet soit possible, je vous dis merci.

A toute ma famille pour leur soutien tout au long de mon parcours universitaire....

Que ce travail soit l'accomplissement de vos vœux tant allégués, et le fruit de votre soutien infaillible....

Merci d'être toujours là pour moi.

KARIMA



Dédicace

À mes chers parents **ABDERRAHMANE HERZI** et **KHADIDJA BOUDRAF**, pour tous leurs sacrifices, leur amour, leur tendresse, leur soutien et leurs prières tout au long de mes études

À ma deuxième mère, **SAADI SOUAD**, et mon père, **OKACHA BISKER**, pour leur soutien tout au long de cette année scolaire difficile....

A mes chères sœurs **HOUDA, ILHAM, IMANE BISKER, ABLA, HANA** et **HADJER** pour leurs encouragements permanents, et leur soutien moral....

A ma cher frère **RACHID HERZI** pour leur appui et leur encouragement....

À mon cher fiancé **OUSSAMA BISKER** pour son soutien, son intérêt et ses encouragements à rester fort dans les pires circonstances....

À ma cousine e sœur **SAFIA SAEEDANI** pour son soutien et son aide exceptionnelle dans la réalisation mémoire....

A toute ma famille pour leur soutien tout au long de mon parcours universitaire....

Que ce travail soit l'accomplissement de vos vœux tant allégués, et le fruit de votre soutien infallible....

Merci d'être toujours là pour moi.

SOUHIR

Table des matières

Sommaire

Résumé

LISTES DES ABREVIATIONS I

LISTE DES TABLEAUX I

LISTE DES FIGURES I

Introduction générale 1

PARTIE TEORIQUE

Chapitre I: généralité et diversité génétique du VIH

1. Introduction	4
2. Classification taxonomique des VIH	4
3. Caractères virologiques du VIH	5
3.1. Le génome du VIH	6
4. Origines du VIH.....	7
4.1. Les Virus de l'Immunodéficience Simien (SIV).....	8
4.2. Origines du VIH-1_M et du VIH-1_N	8
5. Mécanismes de la variabilité	10
5.1. La recombinaison génomique.....	10
6. Diversité génétique du VIH.....	11
6.1. Diversité génétique du VIH-1	11
6.2. Diversité génétique du VIH-2	12

Chapitre II : la phylogénie moléculaire

1. Introduction	13
2. Les arbres phylogénétiques	14
3. Bases/ banques de données biologiques.....	15
3.1. Les bases généralistes	15
3.2. Les bases spécialistes.....	15
4. Alignements.....	16
5. L'évolution moléculaire et ses modèles.....	17
6. L'analyse phylogénétique.....	18
6.1. Les méthodes de reconstruction phylogénétique.....	18
6.1.1. Les Méthodes basée sur les distances	19
6.1.2. La méthode basée sur les caractères.....	19

Table des matières

6.2. Evaluation de la robustesse des phylogénies	21
--------------------------------------------------------	----

PARTIE PRATIQUE

1. Énoncé et justificatif de l'étude	23
2. Patients et méthodes	23
2.1. Lieu d'étude.....	23
2.2. Patients	24
2.3. Méthodes	28
2.3.1. Choix des taxons	28
2.3.2. Etude de la GenBank	29
2.3.3. Traitement des séquences	29
2.3.4. Alignement multiple des séquences	30
2.3.5. Analyse phylogénétique	31
3. Résultat et discussion	33
3.1. Caractéristiques des patients	33
3.1.1. Caractéristiques anthropométriques	33
3.2. Résultats du génotypage	34
3.2.1. Fichier FASTA env	34
3.2.2. Capture des sorties de résultat fournie par logiciel BEAST-2.....	44
3.2.3. La date du premier cas de sida en Algérie (tMRCA) avec le taux d'évolution :	45
3.2.4. Réparation phylogénétique des cas de sida en Algérie	46
Conclusion Générale	48
Bibliographie	
Glossaries	

Résumé

La variation génétique et l'évolution rapide du VIH sont l'un des facteurs les plus importants de sa propagation mondiale. En effet, La répartition des différents génotypes varie d'une région continentale à l'autre. Notre étude a pour objectif d'établir une étude phylogénétique du VIH-1 en Algérie, pour déterminer les souches de VIH-1 circulantes, d'analyser leurs origines et de déterminer l'ancêtre commun le plus récent de VIH en Algérie.

Cette étude a inclus 134 patients résidents en Algérie, infectés par le VIH originaires de différentes régions géographiques de l'Algérie. L'analyse des gènes de l'enveloppe a été réalisée sur le logiciel d'analyse moléculaire « BEAST2 ». Un arbre phylogénique construit avec le modèle sélectionné, tel que mis en œuvre dans « TRACER 1.7.1 », puis édité et affiché à l'aide « d'Icy.Tree ».

Notre population d'étude est constituée de 60% femme et 40% d'homme. L'analyse phylogénique des souches identifiées a montré que le sous-type B est prédominant, alors que les sous-types non-B représentent une grande diversité qui provient principalement de l'Afrique subsaharienne par le biais de l'immigration et par l'activité d'échange entre les différentes populations.

Mots clés : VIH-1, sous-types, diversité génétique, épidémiologie moléculaire, génotype, étude phylogénétique.

Abstract

One of the most important factors in its global spread is genetic variation and the rapid progression of HIV. Indeed, the distribution of the different genotypes varies from one continental region to another. Our study aims to establish a phylogenetic study of HIV-1 in Algeria, to determine the circulating strains of HIV-1, to analyze their origins and to determine the most recent common ancestor of HIV in Algeria.

This study included 134 patients residing in Algeria HIV-infected patients from different geographic regions of Algeria. Envelope gene analysis was performed on « BEAST2 ».Molecular analysis software. A phylogenetic tree built with the selected model, as implemented in « TRACER 1.7.1 » then edited and displayed using « d'Icy.Tree ».

Our study population consists of 60% women and 40% men. The phylogenetic analysis of the identified strains showed that the B subtype is predominant, while the non-B subtypes represent a great diversity which originates mainly from sub-Saharan Africa through immigration and through the exchange activity between different populations.

Keywords: HIV-1, subtypes, genetic diversity, molecular epidemiology, genotype, phylogenetic study.

ملخص

يعد التباين الجيني والتطور السريع لفيروس نقص المناعة البشرية من أهم العوامل في انتشاره العالمي. في الواقع، يختلف توزيع الأنماط الجينية المختلفة من منطقة قارية إلى أخرى. تهدف دراستنا إلى إنشاء دراسة فيلو جينية لفيروس HIV-1 في الجزائر، لتحديد السلالات المنتشرة لفيروس HIV-1، لتحليل أصولها وتحديد أحدث سلف مشترك لفيروس نقص المناعة البشرية في الجزائر.

شملت هذه الدراسة 134 مريضا مصابا بفيروس نقص المناعة البشرية مقيمين في الجزائر و من مناطق جغرافية مختلفة. تم إجراء التحليل الجيني للمغلف على برنامج التحليل الجزيئي «BEAST2» شجرة سلالة تم إنشاؤها باستخدام النموذج المحدد، كما هو مطبق في «TRACER 1.7.1» ثم تم تحريرها وعرضها باستخدام «d'Icy.Tree» يتكون مجتمع الدراسة لدينا من 60% نساء و 40% رجال. أظهر التحليل الوراثي للسلالات التي تم تحديدها أن النوع الفرعي B هو السائد، في حين أن الأنواع الفرعية غير B تمثل تنوعًا كبيرًا نشأ بشكل رئيسي من أفريقيا جنوب الصحراء الكبرى من خلال الهجرة وعبر نشاط التبادل بين مختلف السكان.

الكلمات المفتاحية HIV-1: الأنواع الفرعية، التنوع الوراثي، الوبائيات الجزيئية، التركيب الجيني، دراسة علم الوراثة.

Listes d'abréviations

LISTES DES ABREVIATIONS

AND	acide désoxyribonucléique	LTR	Long Terminal Repeat
ARN	acide ribonucléique	MA	Matrice
ARNm	acide ribonucléique messager	Mac	Macacus rhesus
ARNsb+	ARN simple brin de polarité positive	MCMC	Markov chain Monte Carlo
ARNt	ARN de transfert	Mnd	Mandrill
ARV	Antirétroviraux	NC	Nucléocapside
ATDB	Arabidopsis thaliana database	NCBI	National center for biotechnology information
BIV	Bovine immunodeficiency virus	Nef	negative factor
Ca	Capside	OMS	Organisation Mondiale de la Santé
CAEV	Caprine arthritis encéphalites virus	ONUSIDA	programme commun des nations unies sur le VIH/SIDA
CCR5	chemokine (C-C motif) Receptor 5	OUT	Unites taxonomiques opérationnelles
CD	cluster of differentiation	Pb	paire de baise
Cpx	Complex	PGD	Preimplantation genetic diagnostic
Cpz	Chimpanzee	PIR	Principale banque protéique
CRF	formes recombinantes circulantes	PMBC	peripheral blood mononuclear cell
CXCR4	Chemokine (X-C motif) Receptor4	Pol	Polymerase
DDBJ	DNA data bank of Japan	PSN	primates non humains
EAV	Equine infectious anemia virus	Rcm	mangabey à tête rousse
EBI	European bioinformatics	RDC	la République démocratique du Congo
ECDC	European centre for disease prevention and control	RT	Rétro transcriptase

Listes d'abréviations

ECF	Effet cytopathique	SGD	Swiss institute of bioinformatics
ELISA	enzyme linked sorbent assay	SIB	Saccharomyces genom database
EMBL	European molecular biology laboratory	SIDA	Syndrome d'immuno déficience acquise
Env	Envelope	SIV	Simian immunodeficiency virus
FIV	Feline imunodeficiency virus	SIVcpzPts	SIVcpz Pan troglodytes schweinfurthii
Gag	group-specific antigen	SIVcpzPtt	SIVcpz Pan troglodytes troglodytes
Gor	Gorille	SIVgor	SIV du gorille
Gp	Glycoprotéine	Smm	sooty mangabey
Gsn	cercopithèque hocheur	tMRCA	time to most recent common ancestor
GTR	general time reversible	UPGMA	Un weighted pair group method with arthmeic mean
HAART	highly active antiretroviral therapy	URF	uniques formes recombinantes
HIV	Human Immunodeficiency Virus	Vif	Viral infectivity factor
HTLV	Human T-Lymphotropic Virus	VIS	Simian Immunodeficiency Virus
IN	Intégrase	Vpr	Viral protein r
KB	Kilo bases	Vpu	Viral protein u
LT	Lymphocyte T	Vpx	Viral protein x

Liste des tableaux

LISTE DES TABLEAUX

Tableaux	Titres	Page
Tableau 1	Classification des lentivirus	5
Tableau 2	Rôle des protéines essentielles du VIH	7
Tableau 3	Origine du VIH	9
Tableau 4	Caractéristique moléculaire et répartition géographique de 134 échantillons algériens du VIH-1	24

Liste des figures

LISTE DES FIGURES

Figures	Titres	Page
Figure 1	Structure 3D du VIH	4
Figure 2	Structure schématique du VIH-1	6
Figure 3	Organisation du génome du VIH1	6
Figure 4	Les différentes hypothèses de l'évolution du SIV au VIH-1	8
Figure 5	relation phylogénétique entre les différents groupes de VIH-1, SIVcpz et SIVgor	9
Figure 6	Cycle de réplication du VIH	10
Figure 7	Arbre phylogénétique du VIH-1 et du VIH-2	11
Figure 8	Schéma illustrant les différentes notions de la représentation d'un arbre Phylogénétique	14
Figure 9	Un exemple d'alignement de deux séquences	16
Figure 10	Position géostratégique d'Algérie	24
Figure 11	Les étapes initiales pour récupérer les séquences de la GenBank	29
Figure 12	l'étape téléchargement des séquences env sous format FASTA	29
Figure 13	Menu principal de logiciel MEGA version 5	30
Figure 14	Exemple des séquences ajoutées dans le logiciel MEGA	30
Figure 15	fenêtre du logiciel BEAST2	32
Figure 16	Sortie de l'écran BEAST2 pour l'analyse	32
Figure 17	Répartition des patients selon le sexe	33
Figure 18	Répartition des patients selon l'origine géographique	34
Figure 19	Traceur montrant un résumé de la série BEAST2	44
Figure 20	L'arbre phylogénétique obtenu à partir de l'alignement des gènes env	46
Figure 21	carte indiquant la répartition géographique des échantillons [gènes env]	47

INTRODUCTION GENERALE

Introduction

L'altération causée par le virus de l'immunodéficience humaine (VIH) constitue l'un des plus grandes pandémies que l'humanité a connues, à cause de son vulnérabilité et le taux de létalité élevées en l'absence de traitement approprié. Le VIH présente une variabilité et une diversité génétique en deux types (VIH-1 et VIH-2), quatre groupes principaux pour le VIH-1 : M, N, O, P et huit sous-groupes de A à H pour le VIH-2. Les virus du groupe M du VIH-1 sont responsables de l'épidémie mondiale et sont sous-divisés en neuf sous-types (A à D, F à H, J et K), plus de 45 formes recombinantes circulantes (CRF), de plus de 200 uniques formes recombinantes (URF) et d'innombrables inter-recombinants. Certains de ces variantes génétiques ont une implication différente dans l'épidémie mondiale, et leur distribution est hétérogène (**Roquebert *et al.*, 2009 ; Vessière *et al.*, 2010 ; Hemelaar, 2012**).

La répartition géographique des souches du VIH-1 varie d'un continent à l'autre, d'une région continentale à l'autre. À la lumière des mouvements de population, y compris la migration, le tourisme et les voyages internationaux. La carte moléculaire du VIH-1 n'est pas fixe, car elle se développe en pénétrant des souches importées à partir des zones où les gens se déplacent de manière dynamique à travers les continents et les régions (**Vidal *et al.*, 2008 ; chaix *et al.*, 2009**).

La diversité génétique de VIH est l'une des difficultés majeures pour l'élaboration d'une médication universelle, pour cela Certains projets scientifiques s'attachent d'ailleurs à surveiller l'apparition de nouvelles variantes génétiques, ou de nouvelles mutations de résistance aux thérapies antirétrovirales, afin d'adapter au mieux les traitements et les tests de dépistage (**Hemelaar *et al.*, 2012**). Pour ce faire ces études scientifiques, il faut premièrement séquencer le génome du virus, ou une portion de celui-ci, et analyser les séquences obtenues à l'aide d'outils informatiques. Les séquences trouvées peuvent servir à estimer la vitesse évolutive du VIH. L'estimation de cette vitesse évolutive est nécessaire pour dater l'origine d'une épidémie locale ou mondiale (**Matthieu, 2012**).

Les méthodes généralement utilisées pour ce genre d'étude sont précises, mais lourdes en temps de calcul ; elles ne permettent donc pas de traiter simultanément un grand nombre de séquences nucléotidiques (au plus quelques centaines) et ne considèrent donc seulement qu'une partie de l'information disponible. Les séquences échantillonnées en des lieux différents peuvent fournir d'autres informations, comme, par exemple, la région géographique à l'origine de la diffusion d'une épidémie au niveau local ou mondial. À l'aide de ces séquences, il est aussi possible d'identifier les flux migratoires du virus ainsi que les chaînes de transmission (**Holmes, 2004, 2008**).

Introduction

Nous y apportons une contribution supplémentaire dans notre mémoire, avec une étude ayant pour objectif de déterminer l'origine temporelle de l'épidémie du VIH-1 en Algérie, à l'aide d'outils phylogénétiques classiquement utilisés.

Le sujet de notre recherche, la phylogénie de VIH en Algérie a été choisi, à cause du manque d'études à ce sujet et par conséquent au manque d'informations sur cette pandémie, nous évoquons par exemple, le premier cas réel de SIDA en Algérie et d'autres informations importantes. Dans ce contexte nous avons apportés une contribution supplémentaire dans notre mémoire, avec une étude afin de savoir : Quelle sont les différentes souches de VIH-1 circulent en Algérie ? quel est l'ancêtre commun le plus récent de VIH en Algérie ? et aussi quel est le taux d'évolution (de mutation) de VIH-1 ?

Pour répondre à ces questions, nous avons développé un plan de recherche, que nous avons divisé en deux parties :

Dans la première partie nous réaliserons une recherche bibliographique sur le VIH qui est divisé en deux chapitres ; le premier englobe des généralités sur le VIH avec la diversité génétique de ce virus, le deuxième est consacré aux méthodes de la phylogénie moléculaire dont nous avons besoin pour notre étude pratique.

La deuxième partie est consacrée à la présentation de l'énoncé et le justificatif de l'étude, quelques informations sur les patients, ainsi que les étapes du travail qui a permis d'obtenir l'arbre phylogénétique qui contient les résultats de l'analyse phylogénétique.

En fin de compte, nous avons terminé le travail par une conclusion générale qui résume nos résultats tout au long des étapes d'étude.

PARTIE

THEORIQUE

Chapitre I: généralité et diversité génétique du VIH

1. Introduction

En 1981, les États-Unis d'Amérique ont annoncé l'émergence d'une maladie rare chez les homosexuels, appelée syndrome homosexuel. La maladie est causée par une déficience grave du système immunitaire accompagnée d'une pneumonie rare à *Pneumocystis carinii* (*Pneumocystisjiroveci*). Fin 1981 et début 1982, Des cas ont été signalés chez des patients hémophiles et des receveurs de transfusion sanguine (Catie, 2019)

Le Virus de l'Immunodéficience Humaine (VIH) appartient à la Famille des Rétrovirale et au genre lentivirus. Cette famille est caractérisée par la présence des enzymes du transcriptase inverse ou rétro transcriptase virale (RT) qui permet la rétro transcription de l'ARN en ADN proviral, l'intégrase virale (IN) qui insère l'ADN proviral dans l'ADN de la cellule hôte, et le Protéase(PR) qui permet de décomposer les poly-protéines qui migrent vers la membrane cytoplasmique en protéines et enzymes internes (Peeterset *al.*, 2008 ; Roquebertet *al.*,2009).

2. Classification taxonomique des VIH

On distingue actuellement trois familles pour les rétrovirus avec leurs différences morphologiques et pathogéniques : les spumavirus, les oncovirus et les lentivirus à laquelle appartient les virus de l'immunodéficience humaine (figure1) et également ceux d'autres espèces comme : féline, bovine, simienne, etc. Les Lentivirus sont responsables de maladies à évolution lente, caractérisées par une longue période de latence aboutissant à la mort. Ils ont la particularité d'être cytopathogènes, c'est-à-dire qu'ils tuent les cellules qu'ils infectent (Brun-Vézinet *et al.*, 2000 ; Huraux *et al.*,2003).



Figure 1 : Structure 3D du VIH (Pandreaet *al.*, 2009).

Nom commun : Virus du SIDA

Nom taxonomique: Human Immunodeficiency Virus (HIV)

Types : HIV-1 et HIV-2

Famille : Retroviridae

Sous-famille : Orthoretrovirinae

Genre : Lentivirus

Tableau 1 : Classification des quelques lentivirus (**Brun-Vézinet et al., 2000**).

<i>Espèce de lentivirus</i>	<i>Espèce d'hôte</i>	<i>Hôte</i>	<i>Pathologie</i>
<i>Simian immunodeficiency virus</i>	Primates	Singe	SIDA
<i>Macacus rhesus (SIVmac)</i>		Macaque	
<i>SIVagm (African green monkey)</i>		Singe vert	Aucune
<i>SIVsmm (sooty mangabey)</i>		Mangabey	Aucune
<i>SIVmnd (mandrill)</i>		Mandrille	Aucune
<i>SIVcpz (chimpanzé)</i>		Chimpanzé	Aucune
<i>VIH-1, HIV-2</i>		Homme	SIDA

Les lentivirus humains VIH-1 et VIH-2 dérivent des Virus de l'Immunodéficience Simiens (VIS) des primates : le VIH-1 du groupe M et N est très proche du SIVcpz du chimpanzé, le VIH-1 du groupe O du SIVgor des gorilles et le VIH-2 quant à lui se rapproche davantage du SIVsmm du Mangabey (**Coffin,1996 ; Peeters et al., 2008 ; chaixet et al., 2011**).

D'après le tableau I, le singe macaque est le seul à avoir le SIDA, alors que le singe vert ne la contracte pas, et cela s'explique par sa capacité à tolérance acquise à l'infection VIH à la suite d'une longue période d'adaptions (**Courgnaud et al.,2004 ; Rigaudet et al., 2011**).

3. Caractères virologiques du VIH

Les rétrovirus se présentent d'un diamètre de 110 à 125 nm et un corps excentré tronculaire et enveloppe (**Gilles Furelaud, Benjamin Pavie, 2002**). Le VIH est constitué comme suit :

- Une enveloppe membranaire ou péplos : est une bicouche phospholipidique provient par bourgeonnement de la membrane cytoplasmique des cellules hôtes. Celles-ci comportent une partie interne (la gp transmembranaire gp41 pour VIH-1 et gp36 pour VIH-2), et une

partie externe (la gp120 pour VIH-1, ou la gp105 pour VIH-2) qu'est fixé sur la gp transmembranaire (**Rigaud et al., 2011 ; Belari, 2019**).

- Une matrice : Tapisse la face interne de l'enveloppe (**Lemahieu et Decoste, 2006**)
Fait de protéines oligomérisées p17 pour VIH-1, et p16 pour VIH-2 (**Alfadhliet al., 2009 ; Ghanamet al., 2012**).

- Une capside : c'est une coque en forme de tronc de cône qui entoure et protégé l'acide nucléique viral, et elle est formée par l'assemblage de la protéine CA p24 pour VIH-1, p26 pour VIH-2 (**Belari, 2019 ; Lemahieu et Decoste, 2006**).

- Nucléocapsides (p8 ou p7) : Au nombre de deux (02), chacune contient un ARNs_{b+} et le protège des enzymes virales (**Rigaud et al., 2011**).

Voir la figure 2 ci-dessous.

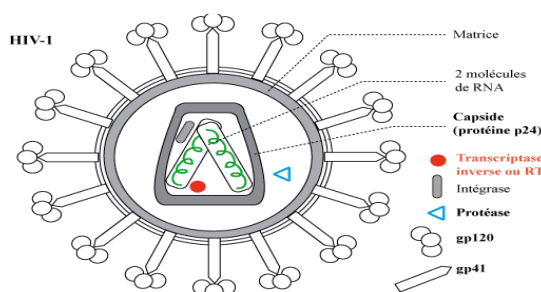


Figure 2. Structure schématique du VIH-1 (**Huraux, 2006**).

3.1. Le génome du VIH

C'est un génome diploïde (deux ARN monocaténares identiques), avec une longueur d'environ 10 kilo bases et une polarité positive. Le génome de VIH code pour les protéines virales, et il est encadré par 2 régions non codantes : 5'LTR et 3'LTR. Les LTR sont utilisés après la transcription inverse pour l'intégration initiale de l'ADN et la transcription (**Lemahieu et Decoste, 2006**).

Le génome du VIH est composé de neuf gènes. Les trois principaux sont gag, pol et env, qui définissent la structure du virus et sont communs à tous les rétrovirus. Les six autres gènes sont tat, rev, nef, vif, vpr et vpu (ou vpx pour le VIH-2), qui codent des protéines régulatrices.

Voir la figure 3 ci-dessous (**Huraux, 2006**).

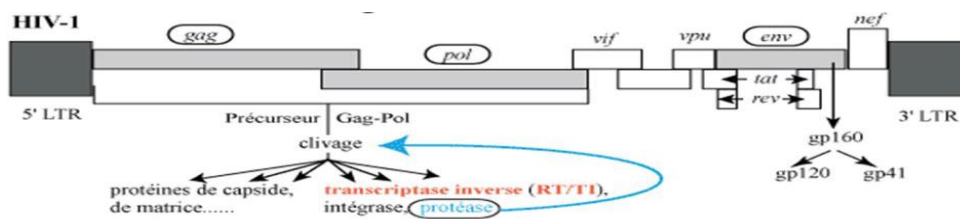


Figure 3 : Organisation du génome du VIH-1 (Huraux, 2006).

Tableau 2 : Rôles des protéines essentielles du VIH.

<i>Gène et produit</i>	<i>Nomenclature</i>	<i>Fonctions</i>
<i>Env</i>	Protéines d'enveloppe	
<i>gp120</i>	Glycoprotéine de surface	Permettent la fixation aux récepteurs
<i>gp41</i>	Glycoprotéine transmembranaires	CD4et co-récepteurs CCR5 ou CXCR4 des cellules cibles (Panceraet al., 2010).
<i>Gag</i>	Protéines de structure	
<i>Gag Pr55</i>	Poly protéine précurseur Gag (Groupe antigène)	Gère le conditionnement et le bourgeonnement viral (Cen et al., 1999).
<i>Pol</i>	Protéines enzymatiques	
<i>PR</i>	Protéase	Décomposer les poly protéines qui migrent vers la membrane cytoplasmique en protéines et enzymes internes. (Lemahieu et Decoste 2006).
<i>RT</i>	Reverse transcriptase	Copier l'ARN en ADN simple brin, hydrolyser le brin d'ARN, copier l'ADN simple brin pour former un ADN bi caténaire. (Lemahieu et Decoste, 2006).

4. Origines du VIH

Grâce aux progrès techniques dans de nombreux domaines scientifiques : sérologie, génomique, le séquençage et l'analyse phylogénétique, l'origine du VIH est devenue connue et a été acceptée comme étant le résultat du métrissage de plusieurs transmissions du virus d'immunodéficience simienne (SIV) à l'homme (**Peeters et al., 2008 ; Vessière al.,2010**). (Figure 4).

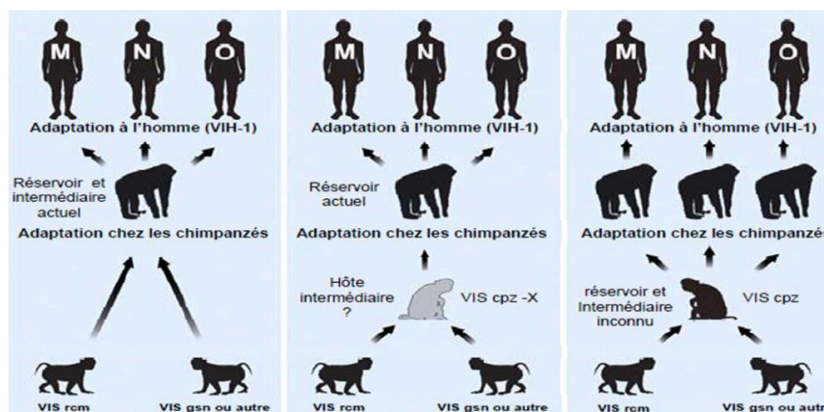


Figure 4 : Les différentes hypothèses de l'évolution du SIV au VIH-1.

Les singes constituent des réservoirs et des intermédiaires qui ont permis au virus d'évoluer jusqu'à devenir pathogène pour l'homme. rcm: mangabey à tête rouge ; gsn : cercopithèque hocheur ; cpz : chimpanzé (en captivité) (Heeney *et al.*, 2006).

4.1. Les Virus de l'Immunodéficience Simien (SIV)

Il se trouvent dans un grand nombre de primates africains et les êtres humains continuent d'être exposés à ces virus par la chasse et la manipulation de la viande de brousse des primates (Liégeois *et al.*, 2009). En 2008, seulement 19 SIV ont été séquencés laissant suggérer qu'il existerait des SIV à découvrir. Chaque espèce de singe est infectée par un SIV qui lui est spécifique (Vande Woude *et al.*, 2006). Il y a 6 grandes lignées de VIS : la lignée VIScpz (chimpanzés), VISmm (mangabeys), VISagm (singes verts), VISsyk (singes sykes), VIShoest (cercopithèque de l'Hoest) et VIScol (colobe guéréza). En revanche, des découvertes récentes de plusieurs souches recombinantes nous montrent la présence de flux VIS entre différentes espèces de singes. En 2010, Plus de 40 SIV, la contrepartie animale du VIH a été décrits (Liégeois *et al.*, 2009 ; Vessière *et al.*, 2010).

Les SIVcpz se trouvent dans deux des quatre sous-espèces de chimpanzés, dont chacune porte une sous-espèce de SIVcpz : SIVcpzPtt dans la sous-espèce de troglodytes d'Afrique de l'Ouest centrale et SIVcpzPts dans toutes les troglodytes Schweinfurth du centre de l'Afrique de l'Est. (Peeters *et al.*, 2008 ; Vessière *et al.*, 2010).

4.2. Origines du VIH-1_M et du VIH-1_N

En 1998, une publication a rapporté la découverte d'un nouveau type de VIH-1 isolé en 1995 chez un patient camerounais atteint du SIDA, c'est le VIH Groupe N pour « non M, non O ». Cette découverte a été la première preuve d'une transmission inter-espèce des VIS des

chimpanzés à l’homme et de l’importance des mécanismes de recombinaison entre les diverses souches de lentivirus des primates (Vessière *et al.*, 2010).

C’est le SIVcpzPtt (Pan troglodytes troglodytes) qui a donné naissance au VIH-1 groupe M et N. le SIVcpz et VIH-1 partagent 80 à 90% d’homologie de séquence nucléotidique (figure 5)(Brun-Vézinet *et al.*,2000 ;Peeters *et al.*, 2008 ; van Heuverswyn et Peeters, 2008). De plus, l’épicentre du VIH-1_M et N est superposé à l’aire de répartition du chimpanzé de l’Ouest (Brun-Vézinet *et al.*, 2000 ; Peeters *et al.*, 2008).

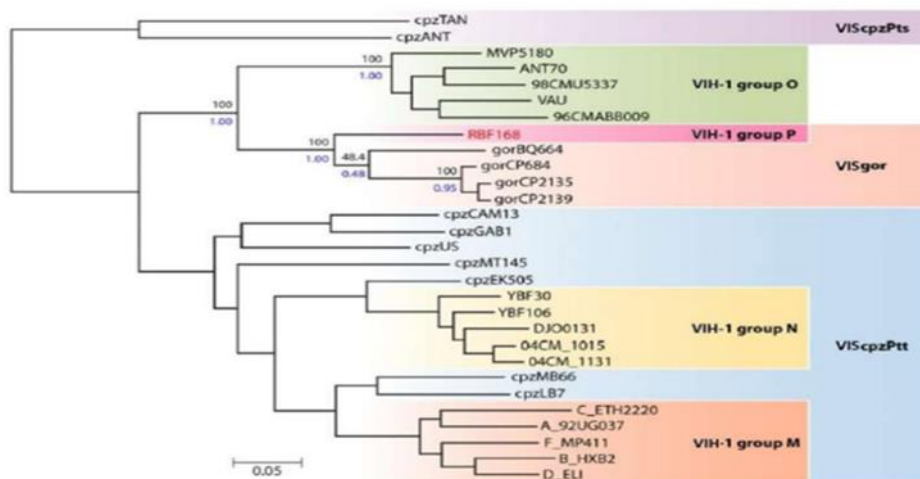


Figure 5 : Relations phylogénétiques entre les différents groupes de VIH-1, SIVcpz et SIVgor.

L’arbre phylogénétique illustre la proximité phylogénétique des séquences de VIH-1 groupe M et N avec les séquences de SIVcpzPtt ainsi que la proximité des séquences de SIVgor avec les séquences de VIH-1 groupe O (Mourez *et al.*, 2013).

- Les origines des autres groupes de VIH-1 a été montré dans le tableau 3.

Tableau 3 : Origines du VIH

Type de virus	Origine	Hôte	Année	Origine géographique	Référence
VIH-1_M	SIVcpzPtt	Chimpanzés pan troglodytes troglodytes	1983	Sud de Cameroun	(Peeters <i>et al.</i> , 2008 ; Vessière <i>et al.</i> ,2010).
VIH-1_N	SIVcpzPtt	Chimpanzés pan	1995	Sud de Cameroun	(Peeters <i>et al.</i> , 2008 ; Vessière <i>et al.</i>

		troglydotes troglydotes			<i>al.,2010).</i>
VIH-1_O	SIVgor	Gorille	1997	Sud du Cameroun	(D’arc et al., 2015 ;Van Heuverswyn et al., 2006).
VIH-1_P	SIVgor	chimpanzés troglydotes	2009	Cameron	Plantier et al., 2009).
VIH-2	SIV smm	Mangabey enfumé	1986	Afrique de l’ouest	(Peeters et al., 2008).

5. Mécanismes de la variabilité

Le changement génétique du VIH est dû à l'émergence d'une mutation, à l'infidélité de la transcription inverse, à la dynamique de reproduction virale et à la recombinaison génétique entre les souches du VIH qui infectent le même organisme, La pléthore de protéines virales et cellulaires de régulation de la réplication virale et aussi les agents pharmacologiques (**Bigaillon et al., 2010; Jameset al.,2011**) (figure 6).

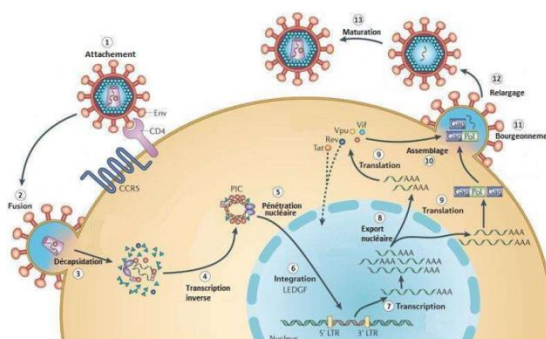


Figure 6 : Cycle de réplication du VIH (Bigaillon et al., 2010).

5.1. La recombinaison génomique

La recombinaison a lieu au cours de la rétro transcription, lorsque la RT utilise les deux brins alternativement comme matrice, par transfert ou déplacement de brin (**Negroni and Buc, 2001**). La recombinaison ne semble pas être limitée par la similitude des séquences, car la recombinaison a été démontrée entre les souches de différents groupes du VIH-1 (groupes M et O) ainsi qu'entre et au sein des sous-types du groupe M. La recombinaison intra-sous-type s'est révélée très étendue au sein du sous-type C du groupe M, bien que de telles données manquent pour d'autres sous-types. Les recombinants entre différents sous-types du groupe M

recombinantes (URF). Par ailleurs, Le sous-type A est subdivisé en sous-sous-type A1, A2 et plus récemment A3 et A4, et le sous-type F est lui-même subdivisé en sous-sous-type du F1 et du F2 (**Roquebert *et al.*, 2009 ; Vessière *et al.*, 2010**).

- Le VIH-1_O qui a une similarité de 73 % pour la protéine Pol et de 50 % pour la protéine env avec le VIH-1_M, elle se caractérise aussi par une très grande diversité génétique mais n'existe pas de classification consensus pour les souches VIH-1-O. L'organisation phylogénétique des souches apparaît moins structurée que celle des VIH-1 du groupe Mavec de nombreuses souches divergentes non classées, on connaît surtout 03 clades A, B et C (**Roquebert *et al.*, 2009**).

- Le VIH-1_N a été découvert en 1995, donc l'histoire naturelle de l'infection dans ce groupe n'est pas bien documentée et les causes de sa faible prévalence ne sont pas expliquées. La particularité de ce VIH est de porter une enveloppe de VIScpzCam, les gènes gag et pol étant proches des VIH mais ne se rapportant à aucun VIH connu à ce jour. Il est unique à ce jour (**Roquebert *et al.*, 2009**).

- Le VIH-1_P a été récemment découvert en août 2009, jusqu'à présent, il n'a été identifié que chez 2 individus (originaire de Yaoundé, Cameroun).il est donc génétiquement unique à ce jour (**Vessière *et al.*, 2010 ; Hemelaar, 2012**).

6.2. Diversité génétique du VIH-2

Le VIH-2 a été génétiquement caractérisé permettant une classification en 8 groupes (A-H), dont les sous-types A et B jouant un rôle dans l'épidémie du VIH-2 en Afrique de l'Ouest (**Peeters *et al.*, 2008**).

Chapitre II : la phylogénie moléculaire

1. Introduction

Au cours des 30 dernières années, la collecte de données en biologie a connu un bond en avant, notamment grâce au développement de nouveaux moyens techniques permettant de comprendre l'ADN et d'autres composants des organismes vivants. Pour analyser ces données extraordinairement volumineuses et complexes, les scientifiques sont tournés vers des nouvelles approches basées sur la modélisation et l'informatique. Ces technologies se rencontrent entre la biologie et l'informatique, et c'est ce qu'on appelle la bioinformatique (Kobbi et Saadi,2017).

La bioinformatique cède des bases de données centrales accessibles à l'échelle mondiale qui permet aux scientifiques de présenter, rechercher et analyser des informations. Elle désigne donc tout à la fois le développement d'infrastructures et d'outils, tels des systèmes de stockage de données, des logiciels de base de données pour les études de données et les comparaisons, et le développement et la mise en œuvre de méthodes mathématiques et informatiques, s'appuyant sur ces outils pour la modélisation, la visualisation, l'exploration et l'interprétation des données (Matthieu,2012 ; KOBBI et SAADI,2017).

La phylogénie est une discipline scientifique qui étudie la formation et l'évolution des organismes vivants afin de prouver la parenté entre eux. Les premières phylogénies se basaient sur des caractères morphologiques, anatomiques et/ou physiologiques afin de comparer les organismes vivants et d'étudier leur parenté. Le but de la phylogénie est de construire un modèle des événements des espèces liées par le passé afin d'obtenir un arbre qui représente fidèlement l'histoire de la biologie, et pour ce faire il est nécessaire de suivre les règles de la Cladistique. Les premières études phylogénétiques principalement basées sur des séquences protéiques remontent au début des années 1960, puis donnent naissance à une nouvelle branche de la phylogénie : la phylogénie moléculaire (Matthieu, 2012).

La phylogénie moléculaire vise à faire des comparaisons pour les séquences de nucléotides ou d'acides aminés dans le but de déterminer les liens de parenté qui les unissent ainsi que pour appréhender leur histoire évolutive (phylogénèse). Elle permet d'étudier l'histoire évolutive des espèces à la base d'une portion de leur séquence moléculaire. Cette discipline est très utilisée en génomique fonctionnelle, science qui étudie le rôle des gènes et aussi très utilisée

Chapitre II : la phylogénie moléculaire

par les épidémiologistes car elle permet de mettre en évidence des liens entre différentes souches virales, liens qui reflètent des chaînes de transmission. La nature quantitative des relations, ainsi établies, a permis le développement de méthodes et de règles rigoureuses pour la construction des arbres phylogénétiques (Matthieu, 2012 ; Kobbi et Saadi,2017).

Nous avons choisi le virus de VIH comme un modèle pour une étude phylogénétique en raison du manque d'informations sur ce virus en Algérie, concernant le premier cas du SIDA, la répartition phylogénétique et le taux d'évolution.

2. Les arbres phylogénétiques

Un arbre phylogénétique ou dendrogramme est une représentation graphique (figure 8) d'origine biologique, mathématique et informatique d'un groupe de taxon pour classer les espèces selon leurs liens de parenté. Chaque espèce est représentée comme un nœud ou une feuille, la relation entre les espèces est appelée bord ou branche, et la racine représente l'ancêtre commun à tous les objets de l'arbre. Les longueurs des branches représentent l'estimation du temps d'évolution ou du nombre de mutations entre les espèces (distance évolutive). Il existe deux principaux types d'arbres qui peuvent être trouvés : arbres enracinés ceux qui ont un seul nœud à partir duquel tous les nœuds sont dérivés, et les arbres non enracinés ceux qui ne le font pas proviennent d'un nœud clair (Jeffrey and Eric,2007).

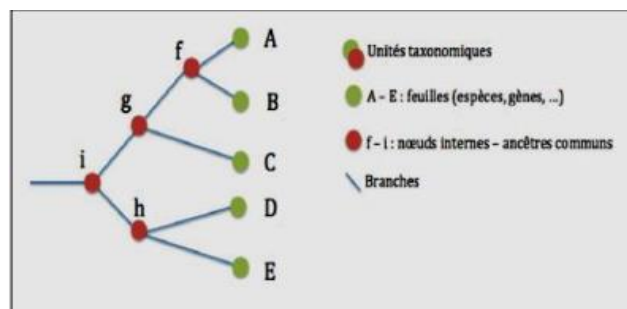


Figure 8 : Schéma illustrant les différentes notions de la représentation d'un arbre phylogénétique (Taguan, 2015).

Actuellement, Il existe plusieurs programmes qui permettent la visualisation des arbres phylogénétiques, par exemple, les programmes : PhyloDraw, NJplot, GeneTree, Phylip (DRAWTREE et DRAWGRAM), Genedoc, Dambe, Treecon, TreeView, et Spectrum (Kobbi et Saadi, 2017).

3. Bases/ banques de données biologiques

Les bases de données fournissent aussi une pléthore d'outils pour manipuler ou analyser les séquences, mais aussi des informations supplémentaires sur chacune d'elles. Les premières banques de données qui ont été créées concernent les informations structurales sur les protéines puis, très rapidement après, les séquences protéiques et nucléotidiques (**Perrière, 2000**).

3.1. Les bases généralistes

Il existe trois banques généralistes de séquences nucléotidiques accessibles au public : l'EMBL en Europe, GenBank aux Etats Unis et la DDBJ au Japon (**Perrière, 2000 ; Golding et al., 2003**). Concernant les banques dédiées aux protéines, elles ont une organisation centrée autour des types d'informations : annotations biologiques et biochimiques d'une part et séquences d'autre part (**Perrière, 2000**).

3.2. Les bases spécialistes

En raison de l'augmentation spectaculaire du nombre de séquences de nucléotides et de protéines publiées, le besoin de développer des banques spécialisées est rapidement apparu. Ces banques contiennent des données homogènes qui sont une collecte établie autour d'une thématique particulière (**Blanchet, 1999**).

D'après (**Perrière, 2000**), on distingue deux types de banques spécialisées : les banques thématiques qui se consacrent à une thématique biologique bien précise (par exemple : banques sur les structures de l'ARN). On peut par exemple citer :

La base de données PDB est une base de données de structures moléculaires, elle combine les structures tridimensionnelles des macromolécules avec des données liées à la structure primaire et secondaire des protéines (**Golding et al., 2003**). La RDP est une banque américaine qui contient les séquences des ARN ribosomiques, et aussi des alignements et des arbres phylogénétiques (**Perrière, 2000 ; Guindon, 2003**).

Le deuxième type de bases de données spécialisées est les bases de données génomiques qui tournent autour des génomes d'organismes étudiés en biologie (séquences d'ADN des chromosomes d'organismes spécifiques). Les modèles d'organisme peuvent être des virus, bactéries, champignons, plantes et animaux. Ces bases aussi rassemblent d'autres données (description de phénotypes, mutations, expression de gènes..., etc.) (**Pillet, 2000**).

Chapitre II : la phylogénie moléculaire

Dans cette catégorie on peut citer : ECDC (est une banque qui regroupe des séquences d'E. Coli contenues dans EMBL et dans GenBank), La base de données génomique de Saccharomyces "SGD" (recueille les séquences nucléiques des gènes codant pour les protéines de Saccharomyces cerisaie), PGD (une base de données qui recueille des données sur les génomes de nombreuses plantes céréalières). En outre, l'ATDB (collecte toutes les informations relatives aux données génétiques, au clonage et aux séquences moléculaires d'Arabidopsis thaliana) (**Kobbi et Saadi,2017**).

4. Alignements

L'alignement des séquences est le processus de comparaison et de détection des similitudes entre les séquences biologiques. Les similitudes détectées dépendront des objectifs du processus particulier d'alignement. L'alignement des séquences semble être extrêmement utile dans un certain nombre d'applications bio- informatiques.

Par exemple, le moyen le plus simple de comparer deux séquences de même longueur est de calculer le nombre de symboles correspondants. Cette étape est une étape cruciale qui permet de choisir les sites qui seront utilisés dans les analyses phylogénétiques avec un but de s'assurer que chacun des sites choisis est homologue. La valeur qui mesure le degré de similitude de séquence est appelée le score d'alignement de deux séquences. La valeur opposée, correspondant au niveau de dissimilarité entre les séquences, est généralement appelée distance entre les séquences. Le nombre de caractères non correspondants est appelé la distance de Hamming. La figure 9 montre un exemple de deux séquences avec une distance de Hamming (**Bookstein et al., 2002**) égale à 3.

```
-----D-PGDF--DRNVPRICGVCGDRATGFHFNAMTCEGCKGFFRRSMKRKA--LFTCP-FNGDCRITKDNRRHCQACRLKRCVDIGMMKEFILTD
IRPQKRK-KGPAP-KMLGNELCSVCGDKASGFHYNVLSCEGCKGFFRRSVIKGA--HYICH-SGGHCPMDTYMRRKCOECLRLKRCRQAGMREECVLSE
SVPGKPS-VNADE-EVGGPQICRVCGDKATGYHFNVMTCCEGCKGFFRRAMKRNA--RLRCPFRKGACEITRKRTRRQCACRLKRCLESGMKKEMIMSD
EPERKRK-KGPAP-KMLGHELCRVCGDKASGFHYNVLSCEGCKGFFRRSVVRRGGARRYACR-GGGTCQMDAFMRRKCOECLRLKRCKEAGMREQCVLSE
PVTKKPRMGASAG-RIKGDDELVVCGDRASGYHYNALTCCEGCKGFFRRSITKNA--VYCK-NGNCVMDMYMRRKCOECLRLKRCCKEMGLAECMYTG
QTEKKC-KGYIPSYLDKDELVVCGDKATGYHYRCITCEGCKGFFRRTIQNLHPSYSCK-YEGKCVIDKVTRNQCQECRPFKCIYVGMATDLVLDLDD
----SPS-PPPPP---RVYKPCFVQDKSSGYHYGVSSCEGCKGFFRRSIQKNM--VYTCH-RDKNCIINKVTRNRCCYCRLOKCFEVGMSKEAVRND
----PPS-PLPPP---RVYKPCFVQDKSSGYHYGVSSCEGCKGFFRRSIQKNM--IYTCH-RDKNCVINKVTRNRCCYCRLOKCFEVGMSKESVRND
----PPS-PPPLP---RIYKPCFVQDKSSGYHYGVSSCEGCKGFFRRSIQKNM--VYTCH-RDKNCIINKVTRNRCCYCRLOKCFEVGMSKESVRND
```

Figure 9 : un exemple d'alignement de deux séquences (**Bookstein et al., 2002**).

En pratique, la reconstruction phylogénétique nécessite l'alignement de plusieurs séquences nucléotidiques. L'alignement est une opération qui permet de comparer des séquences de nucléotides pour repérer les éléments correspondants. Cette étape est une étape

Chapitre II : la phylogénie moléculaire

cruciale qui permet de choisir les sites qui seront utilisés dans les analyses phylogénétiques avec un but de s'assurer que chacun des sites choisis est homologue, c'est-à-dire des séquences nucléotidiques partageant le même ancêtre commun, puisqu'elle consiste à identifier, pour chaque séquence, les nucléotides dérivant du même nucléotide ancestral et à les positionner en regard. Le résultat de cette étape est d'obtenir une matrice, appelée alignement, où chaque ligne correspond à une séquence et où chaque colonne, appelée site, contient les nucléotides issus du même nucléotide ancestral (**Faurnel *et al.*, 1834 ; Guindon, 2003 ; Matthieu, 2012**). Pour faire un alignement il faut que les résidus (nucléotides, acides-aminés) sont superposés de façon à maximiser la similarité entre les séquences. Mais l'alignement de certaines régions peut être difficile et des fois impossibles à cause d'un trop grand nombre d'événement mutationnels qui se sont produits au cours de l'évolution, parmi lesquels des mutations de substitution (remplacement d'un nucléotide par un autre) et/ou d'insertion (ajout d'un ou de plusieurs nucléotides), et aussi de délétion (perte d'un ou de plusieurs nucléotides). Dans certaines séquences de l'alignement des gaps (ou indels) ont pu être introduits, ils correspondent aux phénomènes biologiques d'insertions ou délétions. Pour le biologiste, généralement, le bon alignement est celui qui contient le moins d'évènements de mutation possibles, avec des pondérations différentes pour les différents évènements mutationnels (**Faurnel *et al.*, 1834 ; Guindon, 2003 ; Matthieu, 2012**).

Il existe deux types d'alignements qui diffèrent suivant leur complexité : une l'alignement par paires (Soit global, c'est-à-dire entre les deux séquences sur toute leur longueur, ou local qui se fait entre une séquence et une partie de l'autre séquence), et une l'alignement multiple, qui est un alignement global, consiste à aligner plus de deux séquences. Dans ce cas en utilisant un alignement multiple qui permet de produire un alignement optimal qui a pour but de faire correspondre les positions homologues de diverses séquences et aussi l'ordre des séquences qui est un facteur crucial qui joue sur la qualité de l'alignement final (**Perrière et Brochier-Armanet, 2010**). Ces alignements sont réalisés par des logiciels spécialisés tels que : MEGA, MUSCLE, TreeAlign, clustal W, DAMBE, T.coffee, BIRCH, EMBOSS.

5. L'évolution moléculaire et ses modèles

L'évolution moléculaire se concentre sur l'étude des modifications moléculaires apparaissant notamment au niveau nucléotidique de l'ADN et de l'ARN, à travers des mutations ponctuelles ou d'autres modifications plus importantes du génome. L'évolution moléculaire

Chapitre II : la phylogénie moléculaire

de manière plus macroscopique, s'intéresse à l'impact de ces événements moléculaires sur les processus de spéciation (apparition de nouvelles espèces) et la relation évolutive des espèces.

Les modèles évolutifs occupent une place centrale en phylogénie moléculaire. La distance évolutive entre deux séquences nucléotidiques est définie comme « le nombre moyen de substitutions par site s'étant produites depuis que ces séquences ont divergé de leur ancêtre commun » (**Perrière et Brochier-Armanet, 2010**). Pour calculer la distance évolutive qui sépare deux séquences dans l'alignement, une approche simpliste consisterait à compter le nombre de dissemblances (c'est-à-dire le nombre de sites différents) et de le diviser par la longueur de l'alignement. La construction d'une matrice de distances nécessite le calcul des distances évolutives pour chaque paire de séquences, des modèles d'évolution sont donc proposés pour estimer au mieux la distance évolutive. Chacun fait des hypothèses différentes en ce qui concerne les fréquences d'apparition des nucléotides et les taux de substitution. Parmi ces modèles, on trouve le modèle général time réversible (GTR) (**Lanave et al., 1984**) qui suppose une fréquence d'apparition différente pour chacun des quatre nucléotides et un taux de substitution relatif différent pour chacune des deux transitions ($A \leftrightarrow T$ et $C \leftrightarrow G$) et des quatre transversions possibles ($A \leftrightarrow C$, $A \leftrightarrow G$, $T \leftrightarrow C$ et $T \leftrightarrow G$), le modèle de Kimura à deux paramètres où la fréquence des transitions est souvent plus élevée que celle des transversions, Le modèle HKY de Hasegawa, Kishino et Yano prend en compte la distinction entre transitions et transversions. Les taux de substitution sont proportionnels à la fréquence stationnaire du nucléotide d'arrivée.

Pour tous ces modèles de substitution, vous effectuerez une analyse de Monte Carlo par chaîne de Markov (MCMC) pour estimer la phylogénie et d'autres paramètres du modèle. Les arbres estimés seront des arbres non racinés avec des paramètres de longueur de branche indépendants (**Matthieu, 2012**).

6. L'analyse phylogénétique

6.1. Les méthodes de reconstruction phylogénétique

Dans le domaine de la reconstruction phylogénétique, il existe plusieurs programmes traitant de différents aspects de la phylogénie, allant des problèmes de comparaison de séquences au calcul des distances évolutives, à la reconstruction et à la visualisation des arbres phylogénétiques et au nettoyage de l'alignement.

Chapitre II : la phylogénie moléculaire

6.1.1. Les Méthodes basée sur les distances

La distance entre chaque paire de séquences est calculée, et la matrice de distance est utilisée pour la construction des arbres. Les méthodes basées sur la distance utilisent des modèles de substitution de bases. Il existe deux méthodes basées sur la distance :

6.1.1.1. La méthode de Neighbor-joining (NJ)

C'est une méthode de construction d'un arbre phylogénétique sans racine, à partir d'un indice de déviation (distance ou dis similarité entre séquences). Il est basé sur la recherche d'un couple d'UTO (unités taxonomiques opérationnelles) qui minimisent la longueur totale des branches de l'arbre et ce à chaque étape de regroupement (**Saitou et Nei, 1987**).

6.1.1.2. La méthode de l'UPGMA

C'est une approche directe de la construction d'un arbre phylogénétique à partir d'une matrice de distances. Cette méthode est la seule méthode où les arbres résultants sont enracinés. Elle suppose un taux de substitution constant au fil du temps et des lignées phylogénétiques (connu sous le nom d'hypothèse de l'horloge moléculaire) (**Kurtzman et al., 2010**).

6.1.2. La méthode basée sur les caractères

L'approche basée sur les caractères correspond à des méthodes plus robustes statistiquement que les méthodes de distances mais elles sont plus lentes. Cette approche regroupe les méthodes de parcimonie, de maximum de vraisemblance et nouvellement les méthodes bayésiennes.

6.1.2.1. La méthode de parcimonie

Les méthodes de maximum de parcimonie ont été initialement développées pour être utilisées sur des données morphologiques (**Hennig, 1966**). la parcimonie, à la différence des méthodes de distances, considère chaque site individuellement pour l'ensemble des séquences. Le principe fondamental des méthodes de parcimonie consiste à diminuer au maximum le nombre d'évènements génétiques (mutations, substitutions) pour comparer deux séquences ou deux individus. L'arbre qui en découle compte le moins de pas évolutifs, c'est-à-dire le moins de mutations possibles ; c'est l'arbre optimal (**Perrière et Brochier-Armanet, 2010**). Dans cette approche, les sites peuvent être traités de manière indépendante. Le nombre de mutations correspondant à un jeu de séquences particulier est égal à la somme des mutations requises par chacun des sites (**Ranwez et al., 2003**). Ces parcimonies sont réalisées par des logiciels spécialisés tels que : PHYLIP (Phylogeny Inference Package), PAUP, PAML, EDIBLE, PTP, PHYML, PARDOOT, ALIFRITZ, CRUX, PRAP.

Chapitre II : la phylogénie moléculaire

6.1.2.2. *La méthode du maximum de vraisemblance*

Les méthodes de maximum de vraisemblance utilisent un modèle mathématique du processus d'évolution des séquences pour définir la probabilité qu'une phylogénie puisse produire les séquences observées. Le maximum de vraisemblance est la méthode d'estimation la moins affectée par l'erreur d'échantillonnage et donne de meilleurs résultats que les méthodes de distances et de parcimonie, permettant ainsi de comparer facilement les hypothèses. L'analyse du maximum de vraisemblance se compose de trois parties : premièrement choisir un modèle d'évolution spécifique, à partir de ces modèles différentes hypothèses sur l'histoire de l'évolution sont évaluées en termes de probabilité. Enfin, on choisit l'hypothèse qui présente la probabilité la plus élevée. L'inconvénient majeur de cette méthode est le temps de calcul qui est plus élevé que les autres méthodes (**Swofford *et al.*, 1996 ; Strimmer, 1997 ; Ranwez, 2013**). Cette méthode est réalisée par des logiciels spécialisés tels que : PHYLIP, DAMBE, EMBOSS, PRAP, PAUP, ARB, FastTree, Creux, Vanilla, McRate, DART.

6.1.2.3. *La méthode d'Inférence bayésienne*

La méthode de Bayes, plus exactement appelée méthode Bayesian Markov Chain Monte Carlo (MCMC), a récemment apporté un nouveau moyen de construire l'arbre phylogénétique le plus probable. Cette méthode est très proche des méthodes de vraisemblance donner des résultats prometteurs et son emploi s'est dernièrement généralisé.

Le principe de cette méthode est relativement simple : une série d'arbre est construite sur le principe que chaque arbre, très voisin du précédent, doit apporter une amélioration à la construction générale. Ainsi, un premier arbre phylogénétique est construit au hasard. Puis un élément de cet arbre est changé, également au hasard. Ce nouvel élément peut être la longueur d'une branche, ou le changement de position d'un taxon. Puis un consensus des arbres obtenus est calculé après la suppression des premiers arbres calculés avant l'état d'équilibre (burn-in). Plus le nombre d'arbres parcourus est grand, meilleure est l'approximation. Généralement, cet algorithme nécessite au minimum un million de générations (nombre d'arbres parcourus) et reste réservé pour des phylogénies ne dépassant pas quelques centaines d'OTU. La différence fondamentale entre les méthodes bayésiennes par rapport aux méthodes de vraisemblance est que les paramètres suivent une distribution donnée a priori par l'utilisateur (**Matthieu, 2012 ; Kobbi et Saadi, 2017**). Cette méthode est réalisée par des logiciels spécialisés tels que : PAML, BAMBE, PAL, Vanilla, MrBayes, Mesquite, PHASE, BEAST, MrBayes tree scanners, P4, SIMMAP, IMa2, BALI-Phy.

6.2. Evaluation de la robustesse des phylogénies

Pour tester la fiabilité des arbres phylogénétiques, en particulier celle des branches internes, il faut utiliser une méthode statistique spécifique selon la méthode utilisée dans la reconstruction phylogénétique. Différentes méthodes statistiques permettent de tester les arbres phylogénétiques. Ces techniques sont couramment utilisées pour déduire des arbres phylogénétiques. Parmi ces techniques ou tests on cite : méthodes de ré-échantillonnage (bootstrap) pour déterminer la pertinence des nœuds internes, méthodes de jackknife pour tester des phylogénies dans un cadre moins paramétrique que celui du maximum de vraisemblance, les tests de Kishino-Hasegawa (KH) et Shimodaira-Hasegawa (SH) qui sont utilisés pour comparer des arbres dont les topologies diffèrent (ex. : tester la monophylie d'un groupe) (Matthieu, 2012).

PARTIE

PRACTIQUE

1. Énoncé et justificatif de l'étude

En Algérie, le VIH/SIDA reste un problème de santé publique. La prévention, la clinique, les techniques de diagnostic, la pharmacothérapie anti-HIV et le développement de vaccin tiennent de plus en plus compte de l'épidémiologie moléculaire du VIH-1.

Notre objectif principal est alors, d'établir une étude phylogénétique du VIH-1 qui circule en Algérie, avec les différentes méthodes d'analyses phylogénétique.

De plus d'autres objectifs ont été associés à savoir :

- Analyser les origines des différentes souches de VIH-1 que nous aurions identifiées.
- déterminer le tMRCA (l'ancêtre commun le plus récent de VIH en Algérie).
- Estimer le taux d'évolution (de mutation) de VIH en Algérie.

2. Patients et méthodes

2.1. Lieu d'étude

QUELQUES CHIFFRES SUR L'ALGERIE

- Nom officiel : république algérienne démocratique et populaire
- Superficie : 2381741 Km²
- Capitale : Alger (4,4 millions d'habitants pour le grand Alger)
- Population totale : 43583000(2019)
- Densité : 15,88 hab /Km²
- Croissance démographique : 2,17%
- Espérance de vie : 77,2 ans
- Alphabétisation : 77,9%
- Religion principale : islam

Source: PNUD, ONS (WWW.diplomatie.gouv.fr)



Figure 10: Position géostratégique d'Algérie

(https://www.voyagesphotosmanu.com/carte_maghreb.html)

L'Algérie est bordée par la rive sud de la mer Méditerranée, face à l'Europe au nord et aux pays africains du Sahel au sud. Cette situation géographique stratégique a fait de l'Algérie un point d'influence, de transit et de migration, ce qui a permis aux autochtones de se mêler aux habitants de l'Afrique noire et de la rive nord de la mer (Europe) (Figure 10).

2.2. Patients

Pour réaliser notre analyse phylogénétique, nous avons utilisés une population d'étude comptait 134 patients résidents en Algérie, et infectés par le VIH originaires de différentes régions géographiques de l'Algérie (tableau 3). Les patients n'étaient pas traités ou recevaient des médicaments antirétroviraux. Les séquences analysées sont pour des gènes de l'enveloppe (env).

Tableau 4 : caractéristique moléculaire et répartition géographique de 134 échantillons algériens du VIH-1 (**Bouzeghoub et al., 2006**).

<i>Patients</i>	<i>RT</i>	<i>Prot</i>	<i>Env</i>	<i>Subtype</i>	<i>City</i>	<i>Loc.</i>
58 01	B	B	B		Alger	
11 03	B	B	ND		Alger	
33 03	B	B	B		Alger	
38 01	B	B	B		Tiemcen	
08 01	B	B	B		Alger	

50 01	B	B	B		Alger
50 03	B	B	B		Alger
30 03	B	B	B		Alger
43 03	B	B	ND		Alger
09 01	B	B	B		Tizi ousou
69 02	B	B	B		Alger
47 01	B	B	ND		Alger
67 02	B	B	B		Oran
45 01	B	B	B		Tizi ousou
74 01	B	B	B		Alger
84 01	B	B	B		Oran
90 01	B	B	B		Djelfa
91 01	B	B	B		Alger
89 01	B	B	B		Alger
22 01	B	B	B		Djelfa N
33 02	B	B	B		Alger
33 01	B	B	B		Alger
29 02	B	ND	B		Tziz ousou
112 02	B	B	B		Oran
46 01	B	B	B		Alger
24 01	B	B	B	B	Alger
11 02	B	B	B		Alger
20 02	B	B	B		Bejaia
26 03	B	B	ND		Alger
44 03	B	B	B		Oran
17 03	B	B	B		Alger
02 03	ND	B	B		Alger
28 02	B	B	ND		Alger
32 02	B	B	B		Alger
36 02	B	B	B		Alger
42 02	B	B	ND		Alger
44 02	B	B	ND		Alger
18 01	B	B	B		Bejaia
48 01	B	B	B		Alger
83 01	B	B	ND		Djelfa
25 01	B	B	B		Alger

35 02	B	B	B	Alger	
67 01	B	B	B	Alger	
24 03	B	B	B	Alger	
42 03	B	B	B	Djelfa	
70 02	B	B	B	Alger	
29 03	ND	B	B	Alger	N
37 02	ND	B	B	Alger	
56 02	B	B	B	Alger	
87 01	B	ND	B	Alger	
82 02	B	B	B	Alger	
19 03	B	B	B	Alger	
23 03	B	B	B	Alger	
51 03	ND	B	B	Alger	
13 01	B	B	ND	Alger	
69 01	B	B	B	Ain bessam	
58 02	B	ND	B	Lakhdaria	
49 03	B	B	ND	Mostaganem	
75 01	B	B	B	Annaba	
54 01	B	B	B	Sétif	
58 03	B	B	B	Ain defla	
113 02	B	B	B	Mostaganem	
67 03	B	B	B	Mazouna	
19 02	ND	B	B	Kmis	
80 03	B	B	ND	B	Relizane
61 02	B	B	ND	Ain defla	
111 02	B	B	ND	Boumerdes	
47 03	B	B	B	Bordj m'nail	
79 03	B	B	B	Biskra	
55 02	B	B	B	Biskra	
23 01	B	B	B	Laghouat	
09 03	B	B	B	Eloued	
39 03	B	B	ND	Timimoun	S
06 02	B	B	ND	Sapin	Europe
20 03	B	B	B	ND	
17 01	ND	B	B	ND	ND
71 03	B	B	B	ND	

104 03	B	B	B		Biskra	S
52 01	02	02	02		Alger	
36 01	ND	02	02		Alger	
03 03	ND	02	02		Tiziouzou	N
12 02	02	02	02		Alger	
13 02	02	02	02	02	Alger	
48 02	02	02	02		Alger	
14 03	02	02	ND		Alger	
83 03	02	02	02		Constantine	
116 02	02	02	02		Tiaret	
93 01	ND	02	02		Chlef	N
66 02	02	02	02		Chlef	
84 02	02	02	02		Bourd m'nail	
36 03	ND	02	02	02	Tamanrasset	S
106 02	ND	02	02		Tamanrasset	
78 03	02	02	ND		Ghana	Africa
08 03	ND	02	02		Mali	
57 03	ND	02	02		ND	ND
101 03	06	06	ND		Saida	
35 03	06	06	ND		Tairret	N
88 03	06	06	ND	06	Sidi Belabes	
103 03	06	06	ND		M'sila	
01 02	06	06	ND		Barika	
10 01	06	02	02		Algr	
12 01	06	02	02		Bejaia	
21 03	06	2	02		Oran	N
114 02	06	ND	02		Alger	
70 01	06	02	02		Oran	
82 03	06	06	02		Saida	
03 02	06	06	02	06 / 02	Mascara	
72 03	06	06	02		Souk ahras	
04 03	06	06	02		M'sila	
68 01	06	02	02		Siki Belabes	
15 01	06	02	ND		Eloued	
39 02	06	06	02		Tougourt	

80 01	02	06	ND		Tamanrasset	S
21 02	ND	G	ND	G	Alger	
74 02	G	G	G		Alger	N
59 03	ND	G	G		Oum bouaghi	
76 01	A1	A1	ND	A	Bejaia	N
34 01	D	D	D	D	Bejaia	N
31 03	06	06	B		Tamanrasset	S
12 03	ND	06	B	B-06	El oued	
105 03	06	06	B		Saida	N
57 01	02	02	B	B-02	Ain salah	S
26 02	02	B	ND		Tamanrasset	S
96 02	ND	02	A	A-02	Ain salah	S
115 02	02	02	A		Sidi Belabes	N
72 01	G	06	G	G-06	Tamanrasset	S
31 01	D	05	D	D-05	ND	ND
14 01	09	02	ND	09-02	Ghana	africa
71 01	02	02	G	G-02	Tamanrasset	S
103 02	F2	B	ND	B-F2	Tizi ghenif	N
107 02	ND	A2	H	A-H	Benin	Africa
61 03	B	06		B-06-02	M'sila	
			02			
106 03	B	06	02		Maghnia	N
77 03	G	02	B	G-02-B	M'sila	N

02, 03, 06, 05, 09, CRF ; ND, données non fournies ou cible nucléique non amplifiée ; Loc, localisation ; N, nord ; S, sus (dans l'Algérie)

Les séquences du VIH-1 ont été téléchargées à partir de la base de données du Laboratoire national de Los Alamos.

2.3. Méthodes

2.3.1. Choix des taxons

On choisit de travailler avec les gènes d'enveloppe du VIH-1 parce que la vitesse évolutive du VIH est plus élevée sur le gène env que sur les gènes gag et pol. En effet, le gène env code pour un précurseur des glycoprotéines gp120 et gp41 qui sont exposées à la surface du virion et mutent beaucoup afin de chercher à échapper au système immunitaire.

2.3.2. Etude de la GenBank

La GenBank est une base de données américaine « Nucléotide », en accès libre , a été créée au Centre national pour l'information biotechnologique (NCBI) dans le cadre de la collaboration internationale sur le séquençage des nucléotides (INSDC selon le sigle anglais). Cette banque est mise à jour régulièrement grâce à des échanges quotidiens de séquences avec la banque européenne EMBL et la banque japonaise DDBJ. Afin d'une reconstruction phylogénétique du VIH, nous récupérons les séquences qui se trouvent au niveau de la GenBank. Les numéros d'accès GenBank pour les séquences rapportées dans cette étude sont AY828839 à AY828948 pour les séquences env. pour nous récupérons ces séquences nous passons par plusieurs étapes distinctes :

1. Nous construisons l'URL suivante : <http://www.ncbi.nlm.nih.gov/Genbank/>
2. nous avons écrit les numéros de séquences de AY828839 à AY828948 (figure 11).
3. nous avons téléchargés les séquences sous forme fichier FASTA (figure 12).

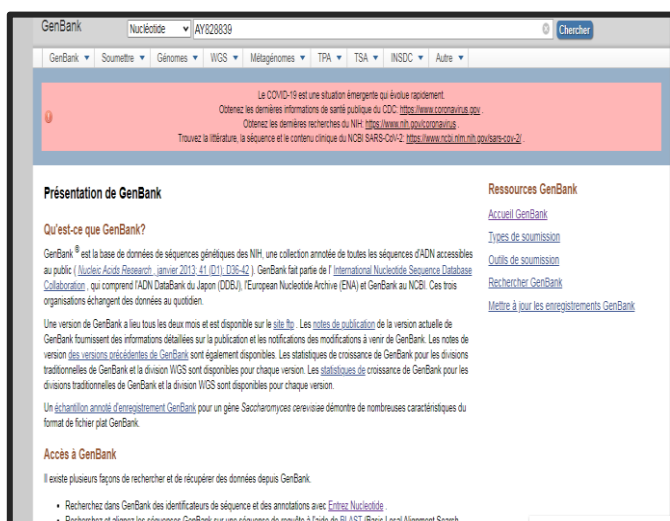


Figure 11 : les étapes initiales pour récupérer séquences de la GenBan



Figure 12 : l'étape téléchargement des séquences env sous format FASTA.

2.3.3. Traitement des séquences

Les séquences sans information géographique ont été écartées de l'analyse. Les séquences du VIH-1 en provenance d'Algérie ont été séparées et sélectionnées manuellement afin de maximiser la longueur du segment analysé ainsi que le nombre de séquences. Selon ce

critère, 129 séquences virales algériennes du VIH-1, couvrant environ 347 pb de l'enveloppe gp120, ont été sélectionnées. Les séquences de référence algériennes ont été séparées, et importées dans le logiciel d'analyse moléculaire de la génétique évolutive version 5.0 (MEGA 5.0) individuellement.

2.3.4. Alignement multiple des séquences

Avant l'étape de l'alignement les séquences nucléotidiques ont été vérifiées à l'aide du logiciel (MEGA 5.0).MEGA est un logiciel simple et efficace pour l'extraction des bases de données en ligne, construction des alignements de séquence et des arbres phylogénétiques. Le plus récent ajout au MEGA5 est une collection de maximum de vraisemblance (ML) analyse pour inférer des arbres évolutifs, sélection de meilleur ajustement des modèles de substitution (de nucléotides ou d'acides aminés), la déduction et les séquences ancestrales membres (avec de probabilités), et l'estimation des taux d'évolution des sites. Cette version de MEGA est destinée pour la plate-forme Windows, et il a été configuré pour une utilisation efficace sur Mac OS X et Linux (Tamura, 2011) (figure 13, 14).

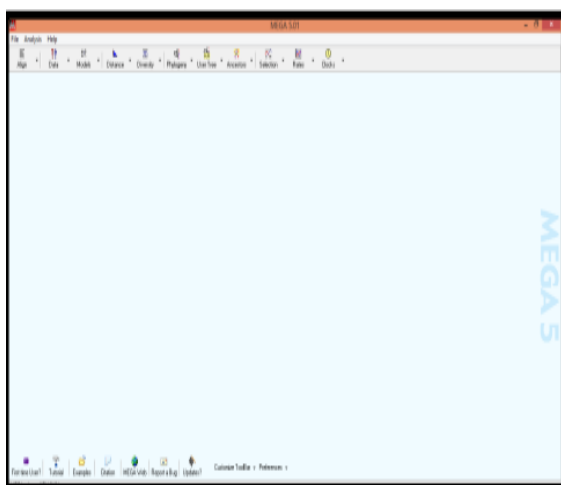


Figure 13 : Menu principal de logiciel MEGA version 5.

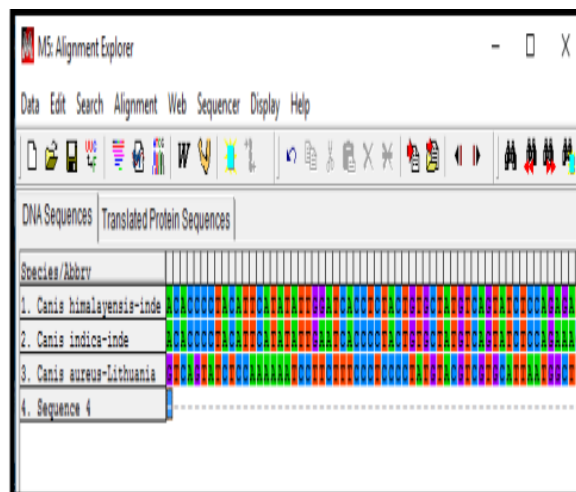


Figure 14 : Exemple des séquences ajoutées dans le logiciel MEGA.

Après l'intégration de ces séquences dans le programme MEGA5, Toutes les séquences sélectionnées ont été alignées dans Clustal W, éditées manuellement si nécessaire. Nous avons choisi ce type d'alignement par ce qu'il est très utile et permet d'étudier les différences et les similitudes entre les séquences d'un même marqueur moléculaire.

Le logiciel ClustalW, comme les autres outils Clustal, est utilisé de manière efficace pour l'alignement multiple des séquences homologues. Il utilise des méthodes d'alignement progressif, qui alignent d'abord les séquences les plus similaires et descendent jusqu'aux séquences les moins similaires jusqu'à ce qu'un alignement global soit créé. ClustalW est un algorithme basé sur une matrice. L'alignement permettra de choisir les sites qui seront utilisés dans les analyses phylogénétiques.

Le programme JMODEL TEST a été utilisé pour sélectionner le modèle le mieux adapté à la substitution de nucléotide, ce qui a abouti au modèle HKY+4 de l'ensemble de données. Ce modèle est avec 4 paramètres : le temps, le taux d'évolution, tMRCA, horloge moléculaire qui sont ajoutées manuellement.

Le JMODEL TEST est un outil pour effectuer la sélection statistique des meilleurs modèles de substitution nucléotidique. Il met en œuvre cinq stratégies de sélection de modèles différentes : des tests de rapport de vraisemblance hiérarchique et dynamique, des critères d'information Akaike et bayésiens et une méthode de théorie de la décision. Il fournit également des estimations de l'incertitude de sélection du modèle, des importances des paramètres et des estimations des paramètres moyennés par le modèle, y compris les topologies d'arbres moyennés par le modèle.

2.3.5. Analyse phylogénétique

Il existe plusieurs méthodes pour construire un arbre phylogénétique, parmi ces méthodes nous en avons utilisées la méthode de maximum de vraisemblance. La méthode de maximum de vraisemblance a été réalisée déterminer en termes de probabilités l'ordre des branchements et la longueur des branches d'un arbre sous un modèle évolutif donné. Cette méthode construit un arbre en partant de l'alignement multiple de Clustalw sans passer par les matrices de distances.

Un arbre de maximum de vraisemblance a été construit avec le modèle sélectionné, tel que mis en œuvre dans TRACER 1.7.1. Les valeurs de confiance des branches d'arbres ont été testées à l'aide du test Shimodaira- Hasegawa(SH), qui peut fournir des informations de précision suffisantes par rapport au bootstrap traditionnel et est très rapide pour les données volumineuses. L'arbre a été édité et affiché à l'aide d'Icy.Tree.

Une approche bayésienne de Markov à chaîne de Markov (MCMC) mise en œuvre dans BEAST v1.6.1 a été utilisée pour estimer les caractéristiques d'évolution du VIH-1

algérien. BEAST2 est un logiciel gratuit pour l'analyse bayésienne évolutive de séquences moléculaires à l'aide de MCMC et strictement orienté vers l'inférence utilisant des arbres phylogénétiques enracinés et mesurés dans le temps. Ce tutoriel est écrit pour BEAST v2.5.0. (Figure 15,16).



Figure 15 : fenêtre du logiciel BEAST2.

Les paramètres évolutifs ont été estimés à l'aide de 134 séquences de données chronologiques (de 2000 à 2003) issues de patients algériens infectés par le VIH-1. L'estimation des paramètres du modèle de substitution de nucléotides, le taux d'évolution (μ , substitutions de nucléotides par site par an, sous-site / site par an) et le temps jusqu'au dernier ancêtre commun (tMRCA) ont été estimés avec un arbre coalescent bayésien de Skyline auparavant. Sous le modèle HKY+4 de substitution de nucléotide, et une horloge moléculaire stricte (modèle Lognormal non corrélé). Deux chaînes MCMC distinctes ont été exploitées pendant 2×10^8 générations pour chaque jeu de données, avec un burn-in de 10%.

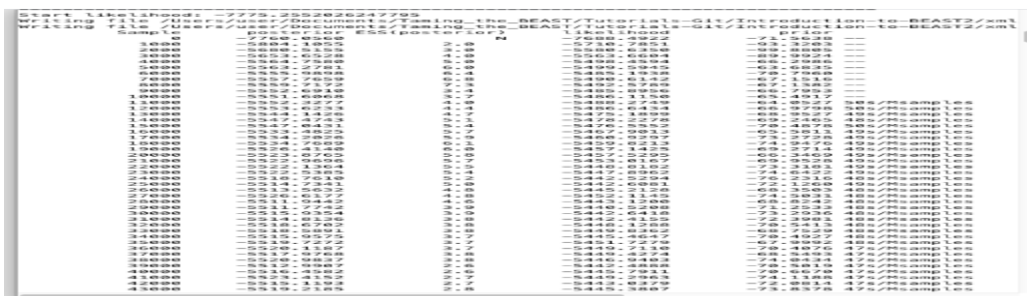


Figure 16 : Sortie de l'écran BEAST2 pour l'analyse.

La sortie de BEAST a été analysée à l'aide de TRACER v1.4, l'incertitude dans les estimations des paramètres étant reflétée par les valeurs de densité de probabilité (HPD) de

95%. Les valeurs de la taille effective de l'échantillon (ESS) pour les estimations étaient supérieures à 200.

Tracer est utilisé pour résumer les estimations postérieures des différents paramètres échantillonnés par la chaîne de Markov. Ce programme peut être utilisé pour l'inspection visuelle et pour évaluer la convergence. Il permet d'afficher rapidement les estimations médianes et les intervalles de densité postérieure les plus élevés à 95% des paramètres, et calcule les tailles d'échantillon effectives (ESS) des paramètres. Il peut également être utilisé pour étudier les corrélations potentielles des paramètres. Nous utiliserons Tracer v1.7.0.

3. Résultat et discussion

3.1. Caractéristiques des patients

3.1.1. Caractéristiques anthropométriques

3.1.1.1. Répartition des patients selon le sexe

Sur un effectif de 143 patients, les femmes représentaient la grande majorité à 60% (80/134) alors que les hommes ne représentent que 40% (53/134) (figure 17).

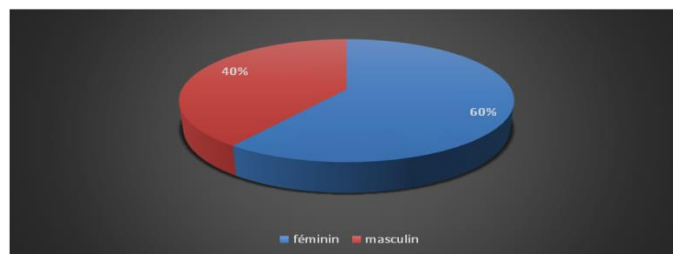


Figure 17 : Répartition des patients selon le sexe.

3.1.1.2. Répartition des patients selon l'origine géographique

Le groupe de patients répartis sur le territoire algérien ont été regroupés selon leurs origines géographiques et sont représentées dans la figure ci-dessous.

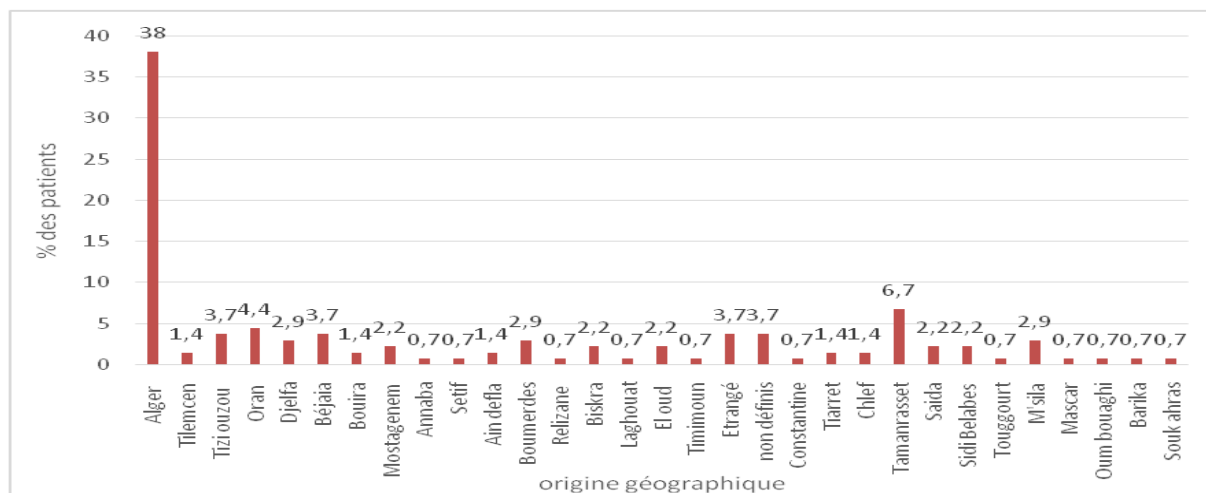


Figure. 18 : Répartition des patients selon l'origine géographique.

3.2. Résultats du génotypage

Les séquences des gènes d'enveloppe (env) de 134 souches de VIH-1 d'individus infectés sont données les résultats suivants :

- Les séquences obtenues d'env sont consignées dans le fichier FASTA env ci-dessous.
- le Traceur montrant un résumé de la série BEAST2 de données des séquences ENV est présenté dans la figure 17.
- L'arbre phylogénétique à partir des séquences d'env est représenté dans la figure 18.
- La répartition des différentes souches est représentée dans la figure 19.

3.2.1. Fichier FASTA env

> AY828841.1 Clone du VIH-1 Clone 0302env Gène de la glycoprotéine (env) d'enveloppe, CDS partiels

```
AATGGCAGCTCTAGCAGAAGAAGAAGTAGTAATTAGATCTGARAATTTACAAACAATGCTAAAATCATAATAGTACAGCTGAATGAATCTGTAA
CAATTAATTGTACAAGACCCAATAACAATACAAGAAAAGTATACATATAGGACCAGGCAGAGCATTATACAACAGGAGAAATAATAGGAG
ATATAAGACAAGCACATTGTAACTTAGTATAGAGTAGAATGGAATAACACTTTACAAAGGATAGTTATGAAATTAAGAGAACAATTTAGAAATAT
AAGCACAAATAGTCTTTAATAAATCTCAGGAGGGGACCCAGAAATTGTAAGACACAGTTTTAACTGTGGAGGGGAATTTTTCTACTGTAA
```

> AY828840.1 Clone du VIH-1 Clone 0302env Gène de la glycoprotéine (env) d'enveloppe, CDS partiels

```
AAATGGCAGTCTAGCAGAAGAAGAAATAGYGATCAGATCTGAARACYTCACAARCAATGTCAAGACCATAATAGTACAGYTAGCTGAGCCTGT
AAAAATTAATTGTAGCAGACCTAACAACAATACAAGAAAAGTATACSTATAGGACCAGGGCAAACATTCTATGCAACAGGTGACATAATAGG
GGATATAAGACAAGCAYATTGTAATATTAGTAGGAAASAATGGGATAGCACATTACAAAAGGTAGCTGMACAATTAAGGAAACACCTTAGGAA
TGAGACAAAATAATCTTTGATAAGCACTCAGGAGGGGATTTAGAAATTACAACACATAGTTTTAATTGTAGAGGGGAATTTTTCTAYTGAA
```

> AY828841.1 Clone du VIH-1 Clone 0303env Gène de la glycoprotéine (env) d'enveloppe, CDS partiels

```
AATGGCAGCTCTAGCAGAAGGAGAAGTAGTAATTAGATYTGAAAATATTACAACAATGCCAAGACCATAATAGTGCAGTTAAGTATCTCTGTAA
GAATTAATTGYACCAGACCAMACAACAATACAAGAACAAGTGTACGTATAGGACCAGGGCAAACATTCTATGCAATAATGACATAATAGGGR
ATATAAGACAAGCACATTGTAATATCAGTGAACAGAATGGCATGCAACTTTAKACAAGGTAGCTACAAAATTAAGRSAGYACTTTAACACAAC
AATAATMTTTRATAAMTCYTCAGGAGGGGATTTAGAAATTACAACACATAGTTTTAATTGTAGAGGGGAATTTTTCTACTGTAA
```

> AY828842.1 Clone du VIH-1 Clone 0403env Gène de la glycoprotéine (env) d'enveloppe, CDS partiels

```
AATGGCAGCTCTAGCAGAAGGACAAATAGCGATCAGRTCTGAARATATCACAACAATGYCAAGACTATAACAGTACAGTTAGTTAATCTCTGTAA
GAATTAATTGTAGCAGACCTRGCAACAATACAAGAAAAGTATACGTATAGGACCAGGGCAAACATTCTATGCAACAAATGACATAATAGGGG
ATATAAGACAAGCACATTGTAATATTAGTAAACAYGAATGGAATAMCATTACAAAAGGTAGCTGAACAATTAYKGAACACTTTAAAAATR
CCACARAAATARTCTTTGCAARCCACTCAGGAGGAGATATAGAAATTACAACACATAGTTTTAATTGTAGAGGGGAATTTTTCTAYTGAA
```

> AY828843.1 Clone du VIH-1 Clone 0602env Gène de la glycoprotéine d'enveloppe (env), CDS partiels
AAAACCATAATAGTACAGCTGAAAGAACCTGTAAAKATTAATTGTTCAAGACCCAACAACMATAGAAGAAARRGGATAACTATGGGACCAGGG
AGAGTATTMTATACAACARRAGAWGTARTAGGAKATATAAGRMGAGCATATTGTAAYATCAGTAGAGCAGAATGGAATAAARCTTTGRGACAK
GTAGTWGAAAWATTAAGAGAACAATTTAATAATAAMACAATAGTCTTTAATAAATCCTCAGGAGGGGACCCAGAAATTGTAATGCATAGCTWT
AATTGTAGAGGGGAATTT

> AY828844.1 gène de la glycoprotéine d'enveloppe (env) du clone du VIH-1 0801env, CDS partiels
TGAAAATTCACGAATAATGCTAAAACCATAATAGTACAGCTGAACCAATCTGTGAAATTAATTGTACAAGACCCAACAACAATACAAGAAAA
GGTGTACATCTAGGACCAGGGGAGCATTCTATACAACAGGACAAATAATAGGAGATATAAGACAAGCATTGTAACATTAGCAGAGCAAAA
TGGAAATRCACCTTAAAACAGATAGTTGGCAAATTAAGAAAAACAATATGAGAATAAAAACAATATCTTTACAACTCCTCAGGAGGGGACCCAG
AAATT

> AY828845.1 Clone du VIH-1 Clone 0803env Gène de la glycoprotéine d'enveloppe (env), CDS partiels
AATGGCAGTCTAGCAGAAGAAGAAATAGTGATTAGATCTGAAAACATCAGTAACAATGCYAAAARCATAATAGTACAGTTGAARRAGGCTGTA
AATATTAMRTGTATGAGACCTGGCAACAATACAAGMAMAAGTGTACGTATAGGACCAGGACAAACATTCTATGCAACAGGAAAAATAAYAGG
GGATAYAAGACAAGCATTGTAATRTCAGTRAAARACWGTGGAACGAACTTTACAGCAGGTRGCTACACAATTAAGGAATCTTACRAATGTC
ACAACAATAATCTTTACTAGTCTCAGGAGGGGATTTAGAAATTACAACACATAGCTTTAGTTGTGGAGGGGAATTTTCTACTGTAA

> AY828846.1 Clone du VIH-1 Clone 0901env Gène de la glycoprotéine d'enveloppe (env), CDS partiels
AATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCCTCCAATTCACAGACAATGCTAAARTCATAATAGTRCAGCTGAATGAATCTGTAG
TYATTAATTGTACAAGACCCAACAACAATACAAGAAAAGGTATACATATAGGACCAGGGAGAACATGGTTYGCRACAGGAGACATAATAGGAA
ATATAAGACAAGCAGACTGTAACTTACTAGAGCAAGATGGAATAAACAATTTAGCAAAGATGTTACAAAATTAAGAGAACAATTTGGGAATA
AAACAATAGTCTTTAAYTCATCCTCAGGAGGGGAYCCAGAAGTGGTAATGCACAGTTTTAATTGTGGAGGGGAATTTTCTACTGT

> AY828847.1 gène de la glycoprotéine d'enveloppe (env) du clone 0903env du VIH-1, CDS partiels
AAATGGCAGTCTAGCAGAAGAAGRGGTAGTAATTAGATCTGACAATTCACAGTCAATACTAAAACCATAACAGTACAATAAAAGACCTGTAA
GTAATTAATTGTACAAGACCTCACACAATACAAGAAAAAGTATACCGATGGGACCAGGGAAAGCATGGTATGCAACAGGAGATATAAYAGGA
GATATAAGACAAGCATTGTAACATTAGTGGAAAYAAAATGGAATRATACTTTAAAACAKATAGTTRAAAAAATTAAARAGAAAAWRITAMTGAA
AYATTTAATGGAACATTTAATGGAACAATAGCCTTTAATCAAYCCTCAGGAGGGGACCCAGAAATTGTAATGCACAGCTTTAATTGTGGAGGGG
AATTTTCTACTGT

> AY828848.1 Clone du VIH-1 1001env gène de la glycoprotéine d'enveloppe (env), CDS partiels
AATGGCAGTCTAGCAGAAGGAGAAATAGCAATCAGATCTGAAAGTCTCACAACAATGCCAAGACCATAACAGTACAATTAATGAGTCTGTAA
AAAATTAATTGTACCAGACCTGGCAACAATACAAGAAAAAGTGTACGTATAGGACCAGGGCAACATTTCTATGCACCAGGTGACATAATAGGA
GATATAAGACAAGCATTGTAATATTAGTAAAAAAGACTGGGAGAACATTACAT AAGGTAGCTGCAACAATTAAGGGAGAGCTTTGGGTAT
GGGAATAAGACAAAAAATCTTTGCTAACCACTCAGGAGGGGATTTAGAAATTACAACACATAGTTTTAATTGTAGAGGGGAATTTTCTACT
GTAA

> AY828849.1 Clone du VIH-1 Clone 1102env Gène de la glycoprotéine d'enveloppe (env), CDS partiels
ATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCCGAGAATTCACGAACAATGCAAAAAYCATAATAGTACAGCTRAACACACYGTAA
ACATTACTTTGTTCAAGACCCAACAACAATACAAGAAAAAGTATACRTATRGGACCAGGCAGCGCAYTTTATACAACAGGAGAAATAATAGGAG
ATATAAGACAAGCATTGTAACCTTAGTAGARCARCATGGAATGACACTTTAAAAGAGGTAGCTAAAAACTAAGTGMMAAATTTGGGAATA
AAACAATAAACTTTACTCAATCCTCAGGAGGGGAYCCAGAAATTGCCATGCATAGTTTCAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828850.1 gène de la glycoprotéine d'enveloppe (env) du clone 1201env du VIH-1, CDS partiels
GAAGGAGAAATAGCGATCAGATCTGAAAATATCAGACAATGCCAAGATCATAACAGTACAATTAATGAGACTGTAAGAATTAATTGTAGC
AGACCTGGCAACAATACAAGAAAAAGTATACGTATAGGACCAGGGCAACATTTCTATGCACATGGTGATATAATAGGGGATATAAGACAAGCA
CATTGTAATATTAGTAAAAAAGACTGGGAKAACACATTAYATAAGGTAGCTGYACAATTAAGGAARCACTTTGGGARTAACACAACAATAAACT
TTACTAGMCACTCAGGAGGGGATTTAGAARTTACAACACATAGTTTTAATTGTAGAGGGAATTTTCTACTGYAATACA

> AY828851.1 gène de la glycoprotéine d'enveloppe (env) du clone 1202env du VIH-1, CDS partiels
TTAAATGGCAGTCTAGCAGAAGAAGATATAGTGATTAGATCTRAAAATATCACAACAATGCCAAAACCATAATAGTACAGTTAGATGMGCCTG
TAAGAATTAATTGTACCAGACCTAACAACAATACAAGACAAAAGTGTACGTATAGGACCAGGGCAACATTTCTATGCAACAGGTGACATAATAG
GGGACATAAGACAAGCATTGTAATGTCAGTAAAACAGAATGGAATAATATGTTACAAAAGGTAGYTACAAAATTAAGGAAGYAYTTTAAGG
ATAATACAACAATAATCTTTGCTAACCAACAGGAGGCGATKTAGAAATTACAACACATAGTTTTAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828852.1 gène de la glycoprotéine d'enveloppe (env) du clone 1203env du VIH-1, CDS partiels
TGGCAGTCTAGCAGAAGAGGAGGTAGTAATTAGTCTGAAAATTTACCGACAATGTTAAAACCATAATAGTACAGCTGAACGAACTGTAAA
AATTAATTGTACAAGACCCAACAACAATACAAGAAAAAGTATACATATAGGACCAGGGAGAGCATTATGCAACAGGAGAYATAATAGGAGA
TATAAGAMAAGCATTGTAACATTAGCAGAGCAGAATGGAATAAAACCTTAGAACAGATAGCTCAAAAATTAAGAGAACAATTTCTGAATAA
AACAATAGCCTTTAATCAATCCTCAGGAGGGGACCCAGAAGTTGCAATGCACAGTTTTAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828853.1 Clone du VIH-1, gène de la glycoprotéine d'enveloppe (env) 1302env, CDS partiels
AATGGCAGTCTAGCAGAAGAAGATATAGTGATTAGATCTCAAAATATCACAACAATGCCAAAACCATAATARTACAGTTAARTAAAGCCTGTAA
GAATTAATTGTACCAGACCTAACAACAATACAAGACAAGTGTACATATGGGACCAGGGCAACATTTCTATGCAACAGGTGACATAATAGGGG
ATATAAGACAAGCATTGTAATGTCAGTGAACAGAATGGAATAGGACRTTACAAGAGGTAGCTACAAAATTAAGGMAGTACTTTAAKGATA
ATACAACAATAATCTTTGCTAACCAACAGGAGGSGATTTAGAAATTACAACWCATAGTTTTAATTGTGGAGGGGAATTTTCTATTGT

> AY828854.1 gène de la glycoprotéine d'enveloppe (env) du clone 1701env du VIH-1, CDS partiels

TAGCAGAAGAAGAGGTAGTAGTGTAGATCTGAAAATTCACGAACAATGCTAAAATAATAATAGTACAGCTGAAGGAAGCTGTAACAATTAATTG
TACAAGACCCAGCAACAATACAAGAAGAAGTATACCTTTAGGACCAGGYAGAGCAATTTATACAACAGACATAATAGGAGATATAAGACAAGC
ACATTGTAACGTTAGTAAAAACAAATGGGAGGGCAGCTTTAAAACAGATAGTTGAAAAATTAAGAGAACAATTTAAAAATAAAACAATAGTCTTT
AATCAATCCGAGGGGGAGACCCAGAAATTGTRRCGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828855.1 gène de la glycoprotéine d'enveloppe (env) du clone 1703env du VIH-1, CDS partiels
ATGGCAGTCTAGCAGAAGGAGAGGTAATAATTAGGTCCGCAATTTCTCGAACAATGSTGACAACATAATAGTACAGCTGAATGAATCTGTAGT
AATTAATTGTACAAGACCMGGCAACAATAMAAGAAAGAGKATAASTATGGGACCAGGYAGAGYATATTATACAACAGGAGAAATARTAGGGG
ATRTAAGACAAGCATTGTAACCTCAGTAGAACAATAATGGGATACCCTTTAAAAAGGATAGCTCTAAAATTAAGRMAACAATTTGAGAATA
AAACAATAGTCTTTGAGAAAACCTTCAGGAGGGGACCCAGAAATTGTAATGCACACTTTAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828856.1 gène de la glycoprotéine d'enveloppe (env) du clone 1801env du VIH-1, CDS partiels
AATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCCGCAATTTACCGACAATGCTAAAGTTATAATAGTGCAGCTGAATGAATCTGTAC
AAATTAATTGTACAAGACCAACAACAACAAGAAAAAGTATATGGTTAGGACCAGGCAGAGCATTATACGACAAAAATAATGGGAGATA
TAAGACAAGCATTGTAACCTTAGTAGAGCACAATGGAATAACACTTTAAGGCAGATAGCTCAAAAATTAAGCAAACAATTTGAGAATAGAA
CAATAGTCTTCAATCAATCCTCAGGAGGGGACCTAGAAATTGTAATGCATAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828857.1 Clone du VIH-1, gène de la glycoprotéine d'enveloppe (env) 1902env, CDS partiels
TTTCACAGACAATGCTAAAACCATAATATGCTACATGTTAAATGAAACTGTACAWATWWWATTGTACAASACCCGGCAATAATACAAGAAGA
WGTATAYATATAGSACCAGGGAGAGYATTTTATACAACGGGASAYATAACAGGAKATATAAGAAAAGCAYATTGCAAYATTAGTGGAGAAGAA
TGGWATAAAACTTTAAMACAKATAGTYGCAAWATYACAAGAACWGTTCACCAACAAMACWATAGTCTWTAMGCWATCCTCAGGAGGGGACC
CAGAAATTGTWATGCACAGTTTTAATTGTRGAGGGGAATTT

> AY828858.1 Clone du VIH-1, gène de la glycoprotéine d'enveloppe (env) 1903env, CDS partiels
ATGGCAGTCTAGCAGAAGAAGGGTAATGATTAGGTCTGAAAATTCACARACAATRCWAAAACCATAATAGTACAGCTGAATGAACTGTAG
AAATTAATTGTACAAGACCAACAACAATACAAGRAAAAGTATACATATGGGACCAGGGARAACATGGTATGCAACAGGAGAYATAATAGGAA
ATATAAGACAAGCATTGTAACCTTAGYAGAGCAAAATGGAATGAACTTTAAGACAGGTAGTTGCCAAATTAAGTGAACAGTTTCCAATAA
AACAATAGTCTTTAATCAATCCTCAGGAGGGGAYCCAGAAATTGTAATGCACAGTTTTAATTGTGSAGGGGAATTTTTCTAYTGAA

> AY828859.1 Clone du VIH-1 2002env gène de la glycoprotéine d'enveloppe (env), CDS partiels
AATGGCAGTCTAGCAGAAGAAGACATAGTAATTAGATCTGAAAATTTTCTGACAATGCTAAGACCATAATAGTACAGCTGAATAAACTGTAC
AAATTCATTGTATAAGACCAACAACAATACAAGAAGAAGTATACCTATAGGACCAGGGAGAGCATTATGCAACAGGGGAAGTAATAGGAA
ATATAAGACAAGCATTGTAACATTAGTAGAGAACAATGGAATGACACTTTAAAAAGATAGCTGACAAGTTAGAAAAACAATTTACGGGTA
GAACAATAGTCTTTAATCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCACAGTTTTAAGTGTGGAGGGGAATTTTTCTACTGTAA

> AY828860.1 Clone du VIH-1 2003env gène de la glycoprotéine d'enveloppe (env), CDS partiels
ATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCCGCAATTTCTCGGACAATGCTAAAGTCATAATAGTACAGTTAAATAAATCTATAGA
AATTAATTGTACAAGACCAACAACAATACAAGAAGAAGTATACATATAGGACCAGGCAGATCATTTTATGCAACAGGAAATATAATAGGAGA
CATAAGAMAAGCAYATTGTACCATTAAATGGAACAGAATGGAATAACACTTTAAAACAGATAGCTACAAAATTAAGAGAACAATTTAAAAATA
AACAATAGCTTTAACAATCCTCAGGAGGAGACCCAGAAGTTACAACACATAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828861.1 gène de la glycoprotéine d'enveloppe (env) du clone 2103env du VIH-1, CDS partiels
ATGGCAGTCTAGCAGAAGAAGGAATAGTGATCAGATCTGAAAATTTCAACAACAATGCCAAGACCATAATAGTACAGTTAGATGAGCCTGTAA
GAATTAATTGTAGCAGACCTAACAACAATACAAGAAARAGTATACGTATAGGACCAGGGCAAACATTCTATGCAACAGGTGACATAATAGGGG
ATATAAGACAAGCAGACTGTAATATTAGTAGAAAAGACTGGAATAACACATTACAAAAGGTAGCTAAACAATTAAGGATACACCTTAAGAAGA
ATGACACAAAAATAATCTTTGCTAAACACTCAGGAGGGGATTTAGAAATTAACAACATAGTTTTAATTGTAGAGGGGAATTTTTCTAYTG
> AY828862.1 gène de la glycoprotéine d'enveloppe (env) du clone 2201env du VIH-1, CDS partiels
TGAAAATATCTCGGACAATGCTAAAACCATAATAGTACAGTTGAACAAAATATAGCAATTAATTGTACTAGACCAACAACAATACAAGAAAA
AGTATATCTATAGGACCAGGGAGAGCATTTTATGCAACAGGAGAAATAATAGGAGATATAAGACAAGCATTGTAACATTAGTGAACAGAG
TGGAATAACTTTAAAACAGGTAGCTAAACAATTAAGAGAACAATAAAGAAATGGGAGAATAAAACAATAGTCTTTAATCAATCCTCAGGA
GGGACCCAGAAATTGTAATGCACA

> AY828863.1 Clone du VIH-1, gène de la glycoprotéine d'enveloppe 2301env (env), CDS partiels
AATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGGTCTGAAAATTTACGAACAATGCTAAGACCATAATAGTACAGCTGAAAGAACCTGTA
CAAATTAAGTGTACAAGACCAACAACAATACAAGAAAAGGTATACATATAGGACCAGGGAGGGCATTATATACAACAGGAGACATAATAGGA
RAWATAAGACAAGCATATTGTAACATTAGTAGAGAAAAATGGAATAMWRCCTTAAAACAGATAGCTTWCAAAATTGAGAGCACAATTAATAA
TAAAACAATAGTCTTTAACAATCCACAGGAGGGGACCCAGAAGTTGTAATGCACAGTTTTAATTGTAGAGGGGAATTTTTCTACTGTAA

> AY828864.1 Clone du VIH-1, gène de la glycoprotéine d'enveloppe 2303env (env), CDS partiels
ATGGCAGTCTAGCAGAAGAAGAAARTARTAAATTAGATCTGAAAATCTCACAACAATGTTAAACCCATAATAGTACATCTGAATGAATCTATAGA
AATTAATTGTACAAGACCAACAACAATACAAGAAGAAGTATTACATTTGGACCAGGGAAAGCATTATATACATCAGACATAATAGGAGATATA
AGACGAGCATATTGTACCCTTAATGGAACAAAATGGAATAACACTCTAAGACTCRTAGCTGCAAAAATTAAGAGAACAATTAATAAAACAATA
WYCTTTSAGCCATCCTCAGGAGGGGACCCAGAAATTGTAACACACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828865.1 gène de la glycoprotéine d'enveloppe (env) du clone 2401env du VIH-1, CDS partiels
AATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCCGCAATATCACRGACAATACTAAAACCATAATAGTACAGCTGAATGAACTGTA
GAAGTTAATTGTAYAAGACCTGGCAACAATACAAGAAAAGGTATACATATAGGACCAGGCAGAGCATTATGCAACAGGAGACATAATAGGA

GATATAAGACAAGCACATTGTAACATTARTGRGAMAMAATGGAATGACACYTTAAAAMAGATAGTTATAAAAATTAAGAGAACAATTTCCGRAT
AAGACAATCAGYTTTAAAGCATTCTCAGGAGGGGACCCAGAAAT

> AY828866.1 Clone du VIH-1 2403env gène de la glycoprotéine d'enveloppe (env), CDS partiels
GGCAGTCTAGCAGAAGAAGAAGTAATAATTAGATCTGACAATTTCCAAAACAATGCTAAAACYATAATAGTACAATTGAAGGAACCCAGTAGAA
ATTAATTGTACAAGACCCAACAACAATACAAGAAAAAGTATGCATMTAGGACCAGGGAGAGCATTATGTCARTAGGAGACATAATAGGAGAT
ATAAGACAAGCACATTGCAACATTAGTAGAARWGMWTGGAATAACACTTTAAGACAGATAGTTAAAAAATTAATAGGACAATWTGGDAATAA
AACAATAGTCTTTAAACAATCTCAGGAGGGGACCCAGAAATTTGTAATGCACAGCTTTAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828867.1 gène de la glycoprotéine d'enveloppe (env) du clone 2501env du VIH-1, CDS partiels
TGGCAGTCTAGCAGAAGAGGTAATAATTAGATCCGCTAATTTCCAGACAATGCTAAAAGCATAATAGTACAGCTGAATGAATCTGTARTAATT
AATTGTACAAGACCCAACAACAATACAAGAAAAAGTATCYCTATAGTCCAGGCAGAGCATTATGCAACAGGAGACATAATAGGAGATATA
AGACAAGCATATTGTAACATTAGTAGAGCAAAATGGAATGACACTTTAAGACAGATAGTTRWAAAATTAAGAGAACAATWTGGGAAAATAA
AACAATAAYCTTTAATCAATCTCAGGAGGGGACCCAGAAATTTGARTGCACAGTTTAATTGTAGAGGGGAATTTTCTACTGTAA

> AY828868.1 Clone du VIH-1 2802env gène de la glycoprotéine d'enveloppe (env), CDS partiels
AAATCTCACAAAATGTTAAACCCATAATAGTACATCTGAATGAATCTATAGAAATTAATTGTACAADACCCAACAACAATACAAGAAGAAGT
ATTACATTTGGACCAGGGAAAGCATTATACATCAGACATAATAGGAGATATAAGACGAGCATATTGTACCCTAATGGAAACAAAATGGAATA
ACACTCTAAGACTRTAGCTGCAAAATTAAGAGAACAACCTCAATAAAAACAATATTCTKTGAGCCATCCTCAGGAGGGGACCCAGAAATTTGTAAC
ACACAGTTTTAATTGTGGAGGGGAATTTTCTACTGTAAAAAA

> AY828869.1 Clone du VIH-1 2902env gène de la glycoprotéine d'enveloppe (env), CDS partiels
ATGGCAGTCTAGCAGAAGAAGAGGTARTAATTAGATCCGCCAATTTACGGAAAATCTAAAGTCATAATAGTACAGCTAAAAGAATCTGTAGA
AATTAATTGTACRAGACCCAACAACAATACAAGAAGAAGTATACATATAGGRCCAGGCAGAGCATTTTATACAACAGGAGATATAATAGGAGA
TATAAGACAAGCACATTGTAAACATTAGTAAAGCAAAATGGAATRACACTTTAGAACAGATAGTTACAAAATTAAGAGAACAATTTGGGAATAAT
AAAACCATAGTCTTTAATCCCCCTCAGGAGGGGACCTAGAAATTTGAATGCACAGTTTTAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828870.1 Clone du VIH-1 2903env gène de la glycoprotéine d'enveloppe (env), CDS partiels
AATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATYTGAAAATTTCCAAAACAATGCTAAAGTCATAATAGTACAGCTGAATAAACTGTGA
GAAATTAATTGTACAAGACCCAACAACAATACAAGGAAAAGTATAAATATAGGACCAGGCAGAGCATGGTATGCAACAGGAGATATAGTAGGA
AATAAAGACAAGCACATTGTAAACATTAATAGAACAAGTGAATGACACTTTAAAACATATAGTTGACAAAATTAAGAGAACAATTTGGGAATA
ATAAAAACAATAATTTTAAATCCATCCTCAGGAGGGGACCCAGAAATTTGTAACACACAGTTTTAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828871.1 gène de la glycoprotéine d'enveloppe (env) du clone 3003env du VIH-1, CDS partiels
AATGGCAGTCTAGCAGAAGATGATGTAGTAATTAGTCTGAAAATTTACGGACAATGTTAAAACCATAAATAGTACAGCTGAAAGAARCTGTAA
AAATTAATTGTACAAGACCCAGCAACAATACAAGAAGAGGTATGCATATAGGACCAGGGAGAGCATTTTTTGAACAGGAAGCATAATAGGAG
ATATAAGACAAGCACATTGTAAACATTAGCARAGGAGAATGGAACAACACCTTAAGWCAGATAGTTGRCAAATTTAGAGAACAATTTCCGAATA
AAACAATAATCTTTAATCAATCCTCAGGAGGGGACCCAGAAATTTCAATGCACAGTTTTAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828872.1 gène de la glycoprotéine d'enveloppe (env) du clone 3101env du VIH-1, CDS partiels
AGTCTAGCAGAAGAAGAGATCATAATTAGATCTGCAAAATATCTCGTCCAATAGTAAAACATAATAGTACAGCTTAATAACTCTGTAGAAATTA
ATTGTACAAGGCCCAACAATAATACAAGACAAGGTGTACATATAGGACCAGGGCAAGCCTCTATACAACAAATATAATAGGAGACATAAGAA
AGGCACATTGTAACATTAGTAGAAAAGCCTGGAATACAACCTTACAGCAAGTAGCAGAAAAATTAAGAAACCTTACTAACATAACAACAATAAT
TTTTCAAATCTCAGGAGGGGACCCAGAAATTTACAACACAC

> AY828873.1 gène de la glycoprotéine d'enveloppe (env) du clone 3103env du VIH-1, CDS partiels
TGGCAGTCTAGCAGAAGAAGGGTGTAGTAATTAGTCTGAAAATTTACGAAACAATGYTAAGACCATAATAGTACAGCTGMAAGAACCTGTAGA
AATTAATTGTACAAGACCCAACAACAATACAAGGAAAAGTATACATATRGGAYCAGGGARRGCATTTATACAACAGGAGACATAATAGGAGA
TATAAGACAAGCACATTGTACCATTAGTGAAGGAAAWTGGAAATGATACCTTAAAMCAGATAGCTGACAAAATTTAGAGWACAATTTCCGAATAA
AACAATAGTCTTTAACCAATCCWCAGGAGGRGACCCAGAAATTTGTAATGCACAGTTTTAATTGTAGAGGGGAATTTTCTACTGTAA

> AY828874.1 Clone du VIH-1 Clone 3202env Gène de la glycoprotéine d'enveloppe (env), CDS partiels
ATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGCCAATTTCCAAAACAATGCTAAGGTCATAATAGTGCAGCTGAATGAATCTGTAGA
AATTAATTGTACAAGACCCAAYAACAATACAAGAAAAAGTATACATTTAGGACCAGGCAGAGCATTCATGCAACAGGAGAAATTAATAGGAGA
TATAAGAAAAGCATATTGTAMKGTARTAGAGCACAATGGAATAACACTTTAAAACAGGTAGTTGCAAAAATTTAGGAAAACAATTTTRWTAAYAM
RACAATAGTCTTYMAGMAMSCCKCAGGAGGGGACCCAGAAATTTGTAATGCACAGTTTTAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828875.1 Clone du VIH-1 3301env gène de la glycoprotéine d'enveloppe (env), CDS partiels
AATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGAAAATATCTCAGACAATGCTAAAACATAATAGTACAGTTGAATGTATCTGTAG
AAATTAATTGTACAAGACCCAACAACAATACAATAAACCTGTAACTATGGGACCAGGGAGAGTACTATACAACAGGACAAATAGTAGGAG
ATATAAGGCGAGCACATTGTAACTGAGTGCAGCAAAATGGAGTGACACTTTAAAACAGGTAGCTGACAAAATTAAGAGAACAATTTACAAGCA
ACACAACAGTAGTCTTTAATCGATCTCAGGAGGGGACCCAGAACTTTGTAATGCACAATTTAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828876.1 Clone du VIH-1 3302env gène de la glycoprotéine d'enveloppe (env), CDS partiels
TGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTCAAAAATTTCCAAAACAATGCTAARGTTATAATAGTACAGCTGAATAAAACCTGTAGAA
ATTAATTGTACAAGACCCAACAACAATACAAGGAAAAGTATACCTATAGGACCAGGCAGAGCATGGTATGCAACRGGAGATATAATAGGAAAT
ATAAGACAAGCACATTGTAACCTTAGTAAAGCAGAATGGRAKAACACTTTAAAACAATAAGTCTGCAAAATTAAGAGAACAATATGAGAATAA
ACAATAATTTTAAATTCATCCTCAGGAGGGGACCCAGAAATTTGTAACACACAGTTTTAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828877.1 Clone du VIH-1 3303env gène de la glycoprotéine d'enveloppe (env), CDS partiels

AATGGCAGTCTAGCAGAAGAAGAAATAGTAATTAGATCTGAAAATCTCACAAACAATGTTAAACCCATAATAGTACATCTGAATGAATCTATAG
AAATTAATTGTACAAGACCCAACAACAATACAAGAAGAAGTATTACATTTGGACCAGGGAAAGCATTATACATCAGACATAATAGGAGATAT
AAGACGAGCATATTGTACCTTAATGGAACAAAATGGAATAACACTCTAAGACTCRTAGCTGCAAAAATTAAGAGAACACTTCAATAAAAAAATA
WCTTTTSAGCCATCCTCAGGAGGGGACCCAGAAAATGTAACACACAGTTTAAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828878.1 Clone du VIH-1 3401env gène de la glycoprotéine d'enveloppe (env), CDS partiels
AATGGCAGTCTAGCAGAAGGAGAGATCATTATTAGATYTGAAAATCTCTCAGACAATACTAAAAACATAATAGTACAGCTTAAATGAATCTGTAG
TAATTACTTGTACAAGGCCAACAAATAACAAGAACAGGTATACACATAGGACCAGGGCAAGCATACTTTACAACAATAATAGGAGACAT
AAGACAAGCACATTGTAATATTACTAGAGGAACTGGACTAAAACCTTAAAGCAAGTAGCTAGAAAATTAGAAAAGCCTTATAACACATCAAA
AATAATTTTGTAGACCATCCTCAGGAGGGGACCCAGAAAATACAACACACACTTTTAAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828879.1 gène de la glycoprotéine d'enveloppe (env) du clone du VIH-1 3502env, CDS partiels
ATGGCAGTCTAGCAGAAGAAGGGGTAGTAGTTAGATCTGAAAATTCACAGACAATGCTAAAACCATAAATAGTACAGCTGAAAGATCCTGTAGT
AATTAATTGTACAAGACCCAATAACAATAACAAGAAAAAGTATACCTATGGGACCAGGGAGAGCATTTTATGCAACAGGAGACATAATAGGAGA
TATAAGACAAGCACATTGTAACCTTAGTGAACAGAATGGAATAACACCTTAAACCAGATAGTTACCAAATTAAGAGAACAATTAATAATAAA
ACAATAGTCTTTAATCATTCTCAGGAGGGGATCCAGAAAATTGTAATGCACAGTTTAAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828880.1 Clone du VIH-1 Clone 3601env Gène de la glycoprotéine (env) d'enveloppe, CDS partiels
ATGGCAGTCTAGCAGAAGAAGAMATAGTGATTAGATCTGAAAATATCACAAACAATGCCAAAACCATAAATAGTACAGTTAGWTRAACTGTAA
GAATTAATTGTACCAGACCWAACAACAATAACAAGAAAAAGTGTACGAATAGGACCAGGGCAACATTCTATGCAACAGGTGATATAATAGGGR
ATATAAGACAAGCACATTGTAATGTTAGTAGAAAAGACTGGRAAGRCACATTACAAAAGGTAGYTRAACAATAAAGGACACACTTTGGTAATGC
AACAATAATCTTTRATAAATCCTCAGGAGGGGATTTAGAAATCACAACACATAGTTTAAATTGTAGAGGGGAATTTTCTACTGTAA

> AY828881.1 Clone du VIH-1 Clone 3602env Gène de la glycoprotéine d'enveloppe (env), CDS partiels
GGCAGTSTAGCAGAAGAGGAGRTAGTAATTAGGTCTGAAAATTCACGAAYAATGCTAAAACCATAAATAGTACAGCTGAAAGAAGCTGTAAAGA
ATTTTTGTACAAGACCCAACAACAATAACAAGAAGAAGTATACCTATAGGACCAGGGAGAGCATTTTATGCAACAGGAGACATAATAGGAAAT
ATAAGACAAGCACATTGTAACATTAGCAGAGGAGCATGGAATARCACCTTAAATCAAACAGTTAGYAAATTAAGAGAACAATTTCCGAATAAA
ACAATAATCTTTAATCATTCTCAGGAGGGGACCCAGAAAATTACAATGCACAGTTTAAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828882.1 Clone du VIH-1 Clone 3603env Gène de la glycoprotéine (env) d'enveloppe, CDS partiels
AATGGCAGTCTAGCAGAAGAAGACATAGTGATTAGATYTGAAAATATCACAAACAATGCCAAAACCATAAATAGTACAGTTAGWTRAACTGTAA
AGAATTAATTGTACCAGACCWAACAACAATAACAAGAAAAAGTGTACGAATAGGACCAGGGCAACATTCTATGCAACAGGTGATATAATAGGG
RATATAAGACAAGCACATTGTAATGTTAGTAGAAAAGACTGGRAAGRCACATTACAAAMAGGTAGYTRWACAATAAAGGACACACTTTGGTAAT
GCAACAATAATCTTTRATAAATCCTCAGGAGGGGATTTAGAAATCACAACACATAGTTTAAATTGTAGAGGGGAATTTTCTACTGTAA

> AY828883.1 Clone du VIH-1 3702env gène de la glycoprotéine d'enveloppe (env), CDS partiels
TGGCAGTCTAGCAGAAGAAGAGGTAATAATCAGATCTAAAAATKCTCGGACAATACTAAAACCATAAATGTACAGCTGAATAYATCTGTAGCA
ATTAATTGTACAAGACCCAGTAACAATAACAAGAAAAAGTATACAGATAGGACCAGGCAGAACATGGTATGCAACAGGAGACATAATAGGAGAT
ATAAGACARGCACAYTGTAACTTAGTAGAGCACAWTGGAAATRATACCTTAAAGRGATAGTWGGCAAATTAAGAGAACAATTTAGGGARAAT
AAAACAATAGTCTTTAATCAACCTCAGGAGGGGACCCAGAAAATTGTAATGCACAGTTTAAATTGTGGAGGGGAATTTTCTACTGT

> AY828884.1 Clone du VIH-1 3801env gène de la glycoprotéine d'enveloppe (env), CDS partiels
GCCAATTTCTCAGACAATGCTAAAATCATAATAGTACAGCTGAATGAATCTGTAGAAATTAATTGTACAAGGCCAACAAACAATAACAAGAAAA
GGATAACTATGGGACCAGGCAGAGYATATTACAACAGGAGATATAATAGGAGACATAAGAAGAGCATATTGTAACATTAGTAAAGCAAAAT
GGAATAACACTTTAAGACARATAGCTATAAAATTAAGAGAACAATTTGGRAATAAYACA

> AY828885.1 Clone du VIH-1 3902env gène de la glycoprotéine d'enveloppe (env), CDS partiels
AATGGCAGTCTAGCAGAAGGAGACATAGCGATCAGATCTGAAAATATCACAAACAATGCCAAGACTATAACAGTACAGTTARTTGGCCTGTAA
GAATTAATTGTAGTAGACCTAACAAACAATAACAAGAAAAAGTATACGTATAGGACCAGGGCAACATTCTATGCAACAGGTGACATAATAGGGA
ATATAAGACAAGCAYATTGTAATATTAGTRAACAGCCTGGAATAAAACATTACAAAAGGTAGCTGRACAATAAAGGAAACACTTTAGAAATG
ACACAAAAAATCTTTGATAGACTCAGGAGGAGATGTAGAAAATTACAACACATAGTTTAAATTGTAGAGGGGAATTTTCTACTGTAAA

> AY828886.1 gène de la glycoprotéine d'enveloppe (env) du clone du VIH-1 4203env, CDS partiels
AATGGCAGTCTAGCAGAAGACGGGTAAATAATTAGATCCAAAATTTACAGGACAATACTAAAACATAATAGTACAGTTGAATACAGCCGTA
AACATTACTTGTCTAAGACCCAACAACAATAACAAGAAGAAGTATAAATATAGGACCAGGCAGAGCATTTTATGCAGCAGGAGACATAATAGGA
GATATAAGACAGGCACATTGTAACATTAGTAAAACAGCATGGAATAACACTTTAAAACAGATAGCTATAARATTAAGCAAAGAATTTGAAAATA
RAACAKTAATCTTTAATCATTCTCAGGAGGGGACCCAGAAAATTGTAATGCACAGTTTAAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828887.1 Clone du VIH-1 4403env gène de la glycoprotéine d'enveloppe (env), CDS partiels
AATGGCAGTCTAGCAGAAGAAGACATAGTACTAAGATCTGAGAATTTACAAACAATGCTAAAACCATAAATAGTACACTGAACGAATCTGTAC
AAATTAATTGTACAAGACCCARTRACAASAGASAAAGKATACATATAGGACCAGGGARAGCWTTTTATACAMATAATATAATAGGAGATA
TAAGAMGRGCATATTGTAACATTAGTAGTACAAAATGGAATAACACATTAAAGCAGATAGTTAAAAAATTAAGAGAACAATTTAGGAATAAAA
CAATAGCCTTTRAKGGATCCTCAGGAGGAGACCCAGAAAATTGTAATGCACAGTTTAAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828888.1 gène de la glycoprotéine d'enveloppe (env) du clone 4501env du VIH-1, CDS partiels
AATGGCAGTCTAGCAGAAGAAGAGATAGTAATTAGGTYTGAAAATCTCACGAACAATGCTAAAACCATAAATAGTACATCTGAATGAATCTATAC
AAATTAATTGTACAAGACCCAACAACAATAACAGCAACCTATAACTATGGGACCAGGCAGAGTATTTTATACAACAGGAGAAGTAACAGGAG
ATATAAGAAAAGCACATTGTAATATTAGTGGAGTAAAATGGAATGACACTTTAAAAGGTAGCTAAAAAATTAAGAGAACAATTAATACTAC
AACAATAATCTTTAAGCAACCTCAGGAGGGGACCCAGAAGTTGTAATGCACATGTTTAAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828889.1 Clone du VIH-1, gène de la glycoprotéine d'enveloppe (env) 4601env, CDS partiels
AATGGCAGTCTAGCAGAAGAAGAGATAGTAATTAGATCTGAAAATTTACRAATAATGCTAAAACCATAACAGTACAGCTGAAGRAACCGTA
ACAATTCRCCTGTGYAAGACCCAACAACAATACAAGAAAAGGTATACATATGGGACCAGGGAGTGTATTKTATACAACAGAAAGTAATAGGGGAT
ATAAGACAAGCATATTGTAACATTAMTAGGRCAGAMTGGAAATAACAYTTTARGACRGATAGYTGTTAAATTACAAGAAAAATTTAACAAAAACA
ATAATCTTTAAKCAATCCTCAGGAGGGGACCCAGAAATTTGTAACCTCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828890.1 gène de la glycoprotéine d'enveloppe (env) du clone 4703env du VIH-1, CDS partiels
AATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCAGAAAATTTACGGACAATGCTAAAAACATAATAGTACAGCTGAATGTATCTGTAG
AAATTAATTGTACAAGGCCAATAACAATACAAGAAAAGTATACATATAGCACCAGGCAGAGCATTATGCAACAGGAGAAAATAATAGGAG
ATATAAGACAAGCACATTGTAACATTAGTAAAACAAAATGGAATAACACTTTAAAACGGATAGTTATAAAATTAAGAGAACAATTTGAGAATAT
AACACAATAATCTTTAATCAATCYTCAGGAGGGGATGCAGAAGTTGCAATGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828891.1 gène de la glycoprotéine d'enveloppe (env) du clone 4801env du VIH-1, CDS partiels
AGAGTTTGTACCTAGACTTGAAAATTTCTCAGACAGTCTAAAACCATAATGGTACAGCTGAACGAAACTGTACAAATTAATTGTACAAGACCC
AAYACAATACAARMAAAGGTATACACATAGGACCAGGAAGAGCATTATGCAACAGGAGAWGTAGTAGGAGATATAAGGAAAGCATATTG
CACCTTAATAAAAACAAAATGGAAAAACACTTTAAAACAGGTAGTTATAAAATTAGCAGAACAATTCAGTAATAAAAACAGAAATAAGCTTTAA
GCAACTCAGGGGGGACCCAGAAATTTAATGCACACT

> AY828892.1 gène de la glycoprotéine d'enveloppe (env) du clone 4802env du VIH-1, CDS partiels
TAAATGGCAGTCTAGCAGAAGAAGGGATAATGATTAGATCTGAAAATATCACAACAATGCCAAAAACATAATAGTACAGTTGAAGKAGSCTG
TAGGAATTARTTGTATCAGACCTGGTAATAATACAAGAAAAGTATACACATAGGACCAGGCAAGCATTCTATGCAACAGGTGACATAACAGG
GAATAAAGACAAGCACATTGTAATRTCAGTAAAGCAGYCTGGAATAAKACTTTAACACAGGTGGCTACACAATTAAGGAGGCCTTTAAAAAT
ACAAAAATAAAATTTGCTAAMTCTCAGGAGGGGATTTGAAATTAACAACACATAGTTTTAATTGTAGAGGGGAATTTTTCTACTRTAAA

> AY828893.1 gène de la glycoprotéine d'enveloppe (env) du clone 4902env du VIH-1, CDS partiels
AATGGCAGTCTAGCAGAAGAAGAGRTAGTAATTAGATCAGACAATTTCTCGGAAAATGCTAAAACATAATAGTACAGCTAAATGMWCTGTA
AAMATTAMTTGTAYAAGACCCARCAACAATACAAGAAGAAGTAACTATAGGACCAGGCAGAGCATTCTATRCAACAGGACACATAAYAGG
AGAYATAAGACAAGCACATTGTAACCTTAGTAGAGCAGACTGGGATAACTTTAAGAAAGATAGCYATMAAATTAAGTGAACAATTTAACAA
TACAACAATAGYATTCATCCCTCAGGRGGAGACCCAGAARTTGAAGGCACAKTTTTAWTTGTGGAGGGGAATTTTTCTACTGTAA

> AY828894.1 gène de la glycoprotéine d'enveloppe (env) du clone du VIH-1 5001env, CDS partiels
ATGGCAGTCTAGCAGAAGAAGAGATAGTAGTTAGATCTGCCAATTTCTCRGACAATGCTAAAACCATAATAGTACAGCTGAACAGGTCTATAAA
AATTAATTGTACAAGACCCAACAACAATACAAGAAAAGTATACCTATAGGACCAGGAGAGCATTITAYACAAGTGGAGACATAATAGGRGA
TATAAGACAAGCACATTGTAATRTTAGYAGAAAAGYTTGGAATGACACTTTAAAACAGATAGTTGAAAATTAAGKGAACAATTCAGAAYAA
AACAATAGYCTTTAATCAATCCTCAGGAGGGGAYCCAGAAATTTGAACACACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828895.1 gène de la glycoprotéine d'enveloppe (env) du clone du VIH-1 5003env, CDS partiels
AAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCARACAACCTCTCGGACAATGCTAAAACCATAATAGTRCAGCTGAAAAAGCTAT
AAAAATTAATTGTACAAGACCCAACAACAATACAAGAAAAGTATACATATGGGACCAGGAAAGCATTITATGCGACAGGAGRMATAAYAGG
AGATAAAGACAAGCACATTGTAACCTTAGTAAAGCAGAATGGAATGACACTTTAAACAGGTAGTTAACAAATTAAGAGAACAATTCCTCAAT
AAAACAATAATTTAADCCTCCTCAGGAGGGGACCCAGAAATTTGAATGCACAGTTTTAATTGTAGAGGGGAATTTTTCTACTGTAA

> AY828896.1 gène de la glycoprotéine d'enveloppe (env) du clone du VIH-1 5103env, CDS partiels
ATGGCAGTCTAGCAGAAGAAGAGATAGTAATTAGGTCTGAAAATATCAGTAACAATGCTAAAACCATAATAGTACACCTGAATGAATCTGTAAA
AATTCATTGTGAAAGACCCAACAACAATACAAGAAAAGTATACATATAGCACCAGGAGAACAATTTATGCAACAGGAGACATAATAGGAGA
TATAAGACAAGCATATTGTAACATTAGTAGAGCAGACTGGAATAAGACTTTAGGACAGGTAAGTCCAAATTAAGAGAACAATTTAATAGTACA
ACAATAATCTTTAAGCCATCCTCAGGAGGGGACCCAGAAATTTGATGCGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828897.1 gène de la glycoprotéine d'enveloppe (env) du clone 5201env du VIH-1, CDS partiels
AATGGCAGTCTAGCAGAAGAAGAGATAGTGATYAGATCTAAAATATCAGATAATACAAAATCATATAGTACASTTGAATGAAYCTGTAG
GAATTAATTGTACCAGACCTAAYACAATACCAGGAAAAGTGTACATATAGGACCAGGCAAGCATTCTATGCAACAGGTGCCATAATAGGAA
ATATAAGASAAGCACACTGTAATGTCAGTAGTACAGAATGGWATARCAGKTTAGAAAAGGTAGYTGAAAATTAAGGRAGCACTTAATAAAA
CAATAATCTTTGCTAATTCCTCAGGAGGGGATTTAGAARTTACAACACATAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828898.1 Clone du VIH-1 5401env gène de la glycoprotéine d'enveloppe (env), CDS partiels
AATGGCAGTCTAGCAKAAGAAGACATAGTAATTAGATCTGMMMAATTTACGAACAATGCTAAGAYCATAATAGTACAACCTGAATGAATCTGTA
GAAATTAATTGTACAAGACCCAACAACAATACAAGAAAAGGATAACTATGGGACCAGGAGAGYATATYATRCAACAGGACAAATAATAGG
AGATAAAGACGAGCATATTGTARCRRTARTARARCAMAATGGGAKAACACYTTARRACAGRTAGYTRMAAATTAAGAGAACAATATGCAAA
TAATACAATAGTCTTTAAGAAATCTCAGGAGGGGACCCAGAAATTTGTAATGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828899.1 gène de la glycoprotéine d'enveloppe (env) du clone 5502env du VIH-1, CDS partiels
AAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCCAGCAATYTCACGGACAATGCTAAAGTCATAATAGTACAGCTGAATACAACCTGT
AAAAATTAATTGTACAAGACCCAACAACAATACAAGAAAAGTATACATATAGSRCCAGGCAGAKCATTITATGCAACAGGAAAAATAATAGG
AGATAAAGACAAGCACATTGTAACCTTAGTAGARMAGAAATGGAATAACACTTTAAAACAGATAGYACAAAATTAARGAGAACAATTAAGAA
TAAAACAATAGTCTTTAAGCAATCCTCAGGAGGGGACCCAGAAATTTGTAATGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAA

GGG> AY828900.1 gène de la glycoprotéine d'enveloppe (env) du clone 5602env du VIH-1, CDS partiels
AATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGAAAATTTCAIRAACAATGCTAAAATTAATAGTACAGCTRAATGAWTCTGTAG
AAATTAATTGTACAAGACCCAAYACAATACAAGAAAAGTGTACCTATGGGACCAGGCGAGCAATTTATGCAACAGGAGACATAATAGGAG

ATATAAGAAAAGCATATTGTACAATTAATGAAGCAAAATGGAATAACACTTTAAGACACATAGTTGAAAAATTAAGGAGAAACAATTTAAGAATA
AAACAATAADCTTTAATCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAAAAATTTT
CTACTGTAA

> AY828901.1 Clone du VIH-1 5701env gène de la glycoprotéine d'enveloppe (env), CDS partiels
ATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGAAAATTTACAAAACAATGCTAAAAGTCATAATAGTACAGCTGAATAAAACCTGTAGA
AATTAATTGTACAAGACCCAACAACAATACAAGGAAAAGTATAAATATAGGACCAGGCAGAGCATGGTATGCAACAGGAGATATAGTAGGAAA
TATAAGACAAGCACATTGTAACATTAATAGAACAGAAATGGAATGACACTTTAAAACATATAGTTGACAAATTAAGGAGAAACAATTTGGGAATAAT
AAAACAATAATTTTTAATCCATCCTCAGGAGGGGACCCAGAAATTGTAACACACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828902.1 gène de la glycoprotéine d'enveloppe (env) du clone du VIH-1 5703env, CDS partiels
GAAAATATCACAAAACATGCCAAAACCATAATAATACAGTTGAATGMGSGTGTARGAATTAATTGTACCAGACCTGGCAACAATACAAGAAAA
AGTGTGCGTATAGGACCAGGGCAAGCATTCTATGCAACAGGTGACGTAATAGGGRATATAAGACAAGCACCATTGAATGTCAGTAAAACAAAA
TGGAATRAAGCTTACRТАAGGTAGYTGAWCAATTAAGAAAAACCTTAAGGATAAACCAATAATCTTTAAGAACTCCTCAGGAGGGGATTTAG
AAATTACAACACATAGTTTTACTTGTAGAGGGGAATTTTTCTACTGTAA

> AY828903.1 Clone VIH-1 5801env gène de la glycoprotéine d'enveloppe (env), CDS partiels
AATGGCAGTCTAGCAGAAGAAGGGGTAGTAATTAGATCTGCCAATTTCTCAGACAATACTAAAAMCATAATAGTGCAGCTGGAAGTACCTGTAA
AAATTAATTGTACAAGACCCGGCAACAATACAATAAAAGGKATACATATAGGACCAGGGAGAGCATTTTATACAAGGAGAACAAATAATAGGAG
ATATAAGACARGCACATTGTAAYCTTAGTGAACAGAAATGGAATAATACTTTAMAAAAGATAGTTAAACAATTAAGAAAAACAATTAAGGAATG
AMACAACAATAGTCTTTGATCAATCCTCAGGAGGGGATCCAGAAATTGCCATGCAYAGTTTTAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828904.1 gène de la glycoprotéine d'enveloppe (env) du clone 5802env du VIH-1, CDS partiels
TGGCAGTCTAGCAGAAGARGATATARTACTCAGATCTGCCAATTTCTCAGACAATGCTAAAGTCATAATAGTACATCTGAATGAATCTATACAAA
TTAATTGTACAAGACCCAACAACAATACAACAGGAAAGTATACATATAGGACCAGGCAGAGCATTTTATGCAACAGGACAAATAACAGGAGATA
TAAGAAAGGCATATTGTAATGTTAGTAGAGAACAATGGAATAAAACCTTTAGAAAGGMTAGTTATAAAAATAAGGAGAAACAATTAACAATAMCA
TAACAATCTTTAACCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCACAAATTTAACTGTGGAGGGGAATTTTTCTACTGTAA

> AY828905.1 Clone du VIH-1 5803env gène de la glycoprotéine d'enveloppe (env), CDS partiels
ATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGAAAATTTACRGAACAATGTTAAAACCATAATAGTACAGCTRAAAGAAGCTGTAA
AATTAATTGTACAAGACCCAACAACAATACAAGAAAAAGTATATCTATAGGACCAGGGAGAGCATTTTATGCAACAGGAGACATAATAGGAAA
TATAAGACAAGCACATTGTAACATTAGCAAAAACAACTGGACTAACACCTTAAGACAGATAGTTGRAAAAATTGGGAGAACAAATTCAGAAATA
AACATAAYCTTTARTAAATCTTCAGGAGGGGACCCAGAAATTACAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828906.1 Clone du VIH-1 5903env gène de la glycoprotéine d'enveloppe (env), CDS partiels
TAATCAGAAGTGAACACACAGACAATGCCAAAATCATAATAGTGCAGCTTAATGAAACTRTASAAATTVATTGTACCAGRCCCAACAACAA
TACAAGRARAGGTATAGGAATTGGACCTGGACAAAYGTTCTATGCAACAGATAAAAAT AATAGGAAATATAAGACAAGCACATTGTAATGTAAG
TGGAGAAAAATGGAGGGAGATGTTAAAGAAGGTAGGGGACAATTAAAAAAGGCTTTAAGAAAAATGTAACCTTTAACTACCCACAGGAGG
UNE

> AY828906.1 Clone du VIH-1 5903env gène de la glycoprotéine d'enveloppe (env), CDS partiels
TAATCAGAAGTGAACACACAGACAATGCCAAAATCATAATAGTGCAGCTTAATGAAACTRTASAAATTVATTGTACCAGRCCCAACAACAA
TACAAGRARAGGTATAGGAATTGGACCTGGACAAAYGTTCTATGCAACAGATAAAAAT AATAGGAAATATAAGACAAGCACATTGTAATGTAAG
TGGAGAAAAATGGAGGGAGATGTTAAAGAAGGTAGGGGACAATTAAAAAAGGCTTTAAGAAAAATGTAACCTTTAACTACCCACAGGAGG
UNE

> AY828907.1 gène de la glycoprotéine d'enveloppe (env) du clone 6103env du VIH-1, CDS partiels
TGAAAATATCACAAAACAATGTCAAGACTATAACAGTACAGTTAGWTAATCCTGTAAAGATTAATTGTAGCAGACCTAACAAACATACAAGAAA
AAGTRTWCGTATAGGACCAGGGCAAACATTCTATGCAACAAATGACATAATAGGGRATATAAGACAAGCACATTGTAATATTAGTAAACAAGA
ATGGAATACCACATTACAAAAGGTAGCTGWACAATTAAGGAAACCTTTAACAAATRCACAAWAATCTTTGCAAACTCAGGAGGAGATAT
AGAAATTACAACACATAGTTTTAAYTGTAGAGGGAATTTTCTACTGTAA

> AY828908.1 Clone VIH-1 6602env Gène de la glycoprotéine d'enveloppe (env), CDS partiels
TAAATGGCAGTCTAGCAGAAGAAGAGRTAATGRRTAGATCTGAGAATATCTCAAACAATGCCAAAACCATAATAGTACAATTGAATAAGACTGT
AGCAATTAATTGTACCAGACCTAACAAACAATACMAGAAAAAGTRTACGTATAGGACCAGGGCAAACATTMTATGCAACAGGTGAAATARTAGG
GGATATAAGAMAAGCATATTGTAATGTCAGTAGAGCAGAAATGGAGTAACACTTTACACAAGGTAGCTRCACGATTAAGGGAGCACYTCAACAA
TAGCAAACCAATAGTCTTTGCTAACCCCTCAGGGGGGATATGGAATTAACAACACATAGTTTTAATTGTGGAGGGGAATTTTTCTAYTGAAA

> AY828909.1 Clone VIH-1 6701env gène de la glycoprotéine d'enveloppe (env), CDS partiels
TGGCAGTCTAGCAGAAGAAGAGATAGTRATTAGATCAGAAAACCTCACTAACAAATGCTAAAAGTCATAATAGTACAGCTGAATGAAACTGTACAA
ATTAATTGTACAAGACCCACAACAATACAAGAAAAAGTATAAATATAGGACCAGGCAGAGCATGGTATACAACAGGACAAATAATAGGAGAT
ATAAGAAAAGCATATTGTAACCTTAGTAGHACAAATGGRATAATGCTTTAARAAGGATAAGTATAAAAATAAWGAACAATTAAGAATGCA
ACAACAATAGCTTTTAAKCCATCCTCAGGAGGGGACCCAGAAATTGTAATGCACACTTTAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828910.1 Clone VIH-1 6702env Gène de la glycoprotéine d'enveloppe (env), CDS partiels
TAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGACAATTTCTCAGACAATGTTAAAACCATAATAGTACAGCTGAACCAATCTGT
ARTAATTAATTGTACAAGACCCAACAACAACAAGAAAAAGTATAYCTATAGGACCAGGGAGAGCATTTTATGCAACAGGAGATATAATAGG
AAACATAAGACAAGCACATTGCAACATTARTAGAGCAGCTTGAATTAACACTTTGAAAMAGATAGCTGCAAAAATTAGGARACAATTTAACAA
AACAAAGKATAARATTTGARCAATCCTCAGGAGGGGACCCAGAAATTGTAWTGCACAGTTTTAATTGTCRAGGGGAATTTTTCTACTGTAA

> AY828911.1 Clone VIH-1 6703env gène de la glycoprotéine d'enveloppe (env), CDS partiels
AAATGGCAGTCTAGCAGAAGARGAGGTAGTAATTAGATCTGACAATTTCTCAGACAATGCTAAAACCATAATAGTACAGCTGAAAYGAATCTGTA
GTAATTAATTGTACAAGACCCAACAACAACAAGAAAGGTATACATATAGGACCAGGAAGAGATTTTATGCAACAGATATAATAGGAGAC
ATAAGACAAGCACATTGCAACATTAGTAGACAAAATTGGACTAAYACTTTAAGACAGATAGCTRCAAAATTAGGAAAAGAAATTTARGAATAGC
AATATAAGCTTTAAKCAATCCTCAGGAGGGGACCCAGAAATTGTAWTGCACAGTTTTAAAYGTAGAGGGGAATTTTCTACTGTAATA

> AY828912.1 Clone VIH-1 6801env Gène de la glycoprotéine d'enveloppe (env), CDS partiels
ATGGCAGTCTAGCAGAAGGAGAAATAGCAATCAGATCTGAAAATATCACAAACAAATGCCAAGACCATAACAGTACAGTTAGATCAGCCTGTAA
AAATTAATTGTARCAAGMHTGGCAACAATACAAGAAAAAGTATACGTATAGGACCAGGGCAAACATTCTATGCAACAGGTGACATAATAGGGR
ACAYMAGAMAAGCACATTGTAATATTAGCAGAACAGAATGGAATAACACATTACATAAGGTAGCTRTACAATTAAGGAAACTTTGGGAATG
ACACAACAATAATCTTTAATAAACRCTCAGGAGGGGATTTAGAAATTACATCACATAGTTTTAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828913.1 Clone VIH-1 6901env gène de la glycoprotéine d'enveloppe (env), CDS partiels
ATGGCAGTCTAGCAGAAGAAGAGGTAATAATTAGATYTGAAAATTCACAGACAATGTTAAAACCATAATAGTACAACACTGAAAGACCCAGTAG
AAATTAATTGTACAAGACCCAACAATAATACAAAGAAAAGTATAAATATAGGACCAGGGAGAGCATTGATGTAACAGGAAGCATAATAGGAG
ATATAAGACGAGCACATTGTAACGTTAGTAGAGCAAAATGGAATAACACTTTACATCAGGTAGTTGAAAACTAAAAGAACAATTTAGGAATA
AAACAATAATCTTTAATCATTCTCAGGAGGGGACCCAGAAATTGTAATGCACAGTTTTAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828914.1 Clone VIH-1 6902env gène de la glycoprotéine d'enveloppe (env), CDS partiels
AATGGCAGTCTAGCAGAAGAAGRGGTAGTAATTAGATCTGACAATTCACAGTCAATACTAAAACCATAACAGTACAACACTAAAAGACCTGTAG
TAATTAATTGTACAAGACCTCACAAACAATACAAGAAAAAGTATACCGATGGGACCAGGGAAAGCATGGTATGCAACAGGAGATATAAYAGGAG
ATATAAGACAAGCACATTGTAACATTAGTGGAAYAAAATGGAATRATACTTTAAAACAKATAGTTRAAAAATTAARAGAAAAWRTTAATGAAA
YATTTAATGGAACATTTAATGGAACAATAGCCTTTAATCAAYCCTCAGGAGGGGACCCAGAAATTGTAATGCACAGCTTTAATTGTGGAGGGGA
ATTTTCTACTGTAA

> AY828915.1 gène de la glycoprotéine d'enveloppe (env) du clone 7001env du VIH-1, CDS partiels
GAGATGCTGATCAGATCTGAAAATCTCACAAACAATGCCAAGACCATAAYAGTACAGTTAGCTGAGCCTGTAAAATTAATTGTRYCAGACCYA
GCAACAATACAAGACAAAGTATACGTATAGGACCAGGACAARYATTCTATGCAACAGGTGACATAATAGGGGATATAAGACAAGCACATTGTA
ATATTAGTAGAAAAGATTGGAATARCACATTACAAAAGGTAGCTGTACAATTAAGGAAACATTTTGRGAATGCAACAATAAWCTTTAMTAAAC
CCGYAGGAGGGGATTTAGAAATTAAYAACACAYAGTTTTAATTRTAGAAGRGAATTTTCTA

> AY828916.1 gène de la glycoprotéine d'enveloppe (env) du clone 7002env du VIH-1, CDS partiels
AATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGCCAATTCACAGACAATGCTAAAAGCATAATAGTACAACACTGAATGCATCTGTAG
AAATTAATTGTACAAGACCCAGCAACAATACAAGAAGAAGTATACACATAGGACCAAACAGAGCATTATGCAACAGGAGACATAATAGGAG
ATATAAGACGAGCCTATTGTAACATTAGTAGAGCAAAATGGAATAACACTTTAAAACAGRTAGTTACAAAATTAAGAGAACAATATGGAATA
AAACAATAGTCTTTAACCATTCTCAGGAGGGGACCCAGAAATTGTAATGCACAGTTTTAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828917.1 Clone du VIH-1 7101env gène de la glycoprotéine d'enveloppe (env), CDS partiels
AATGGCAGTCTAGCAGAAGGAAATATAAAAATTAGATCTGACAATTCACAGACAATACCAAAGTCATAATAGTGCAGTTAATGAGACTGTGG
AAATTAATTGTACCAGGCCTAACACAATACAAGAAGAAGTATAACATTTGGACCAGGACAAGCGTTCTATGCAACAGGTGCAATAATAGGAG
ACATAAGACAAGCATATTGTAATGTTAGTGAAGAAAATGGAATCAGATGTTACAGAAGGTCGAAGACATCTACAAAAATCTATAACATCA
GTAACAACAATAACMTTAACTCAKCTCAGGAGGGGACYTAGAAATTACAACACATAGTTTCAATTGTAGAGGGGAATTTTCTACTGTAA

> AY828918.1 gène de la glycoprotéine d'enveloppe (env) du clone 7103env du VIH-1, CDS partiels
ATGGCAGTCTAGCAGAAGAGGAGGTAGTAATTAGGTCTGAAAATTCACGAACAATGCTAAAACCATAATAGTACAGCTGAAAGAAGCTGTAA
TAATTAAGTGTACAAGACCCAACAACAATACAAGAAAAAGTATACCTATRGGACCAGGGAGAGCATTATGCAACAGGARSCATAATAGGAG
ATATAAGACAAGCACATTGTAATATTAGCAGAGAAAAATGGAATARCCTTAAGTCAGATAGTTGGCAAATTAAGAGAACAATTCAGAATAA
AACAATAATCTTTAATCAATCCTCAGGAGGGGACCCAGAAATWATAATGCACASTTTAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828919.1 gène de la glycoprotéine d'enveloppe (env) du clone 7201env du VIH-1, CDS partiels
AATGGCAGTCTAGCAGAAGGAAATATAAAAATTAGRTCTGACAATTCACAGACAATACCAAAGTCATRATAGTGCAGTTAATGAGACTGTGG
AAATTAATTGTACCAGGCCTAACACAATACAAGAAGAAGTATAACATTTGGACCAGGACAAGCGTTCTATGCAACAGGTGCAATAATAGGAG
ACATAAGACAAGCATATTGTAATGTTAGTGRAGAAAATGGAATCAGATGTTACAGMAGGTCGAATACATCTACAAAAATCTATAACATCA
GTAACAACAATAACCTTTAACTCAKCCYAGGAGGGGACCTAGAAATTACAACACATAGTWTCAATTGTAGAGGGGAATTTTCTACTGTAA

> AY828920.1 gène de la glycoprotéine d'enveloppe (env) du clone du VIH-1 7203env, CDS partiels
GGCAGTCTAGCAGAAGGAAAGATAGCGATCAGATCTGAAAATTCACAAACAATGCCAAGATCATAACAGTACAGTTAGTTRAMCCTGTAARA
ATTAATTGTAGCAGACCTGRCAACAATACAAGACAAGGTATACRTATAGGACCAGGGCAAACATTCTATGCAACAGGTGAAATAATAGGGGAY
ATAAGACAAGCAYACTGTAATATTAGTAAAAAGACTGGAATAAAACATTASARWGGTAGCTRAACRATTAAGGGAACGCCTTGACAATAGA
TCTGCAACARTAATCTTTGCTAACCACACAGGAGGGGATTTAGAAATYACAACACATAGTTTTAATTGYAGAGGGGAATTTTCTACTGTAA

> AY828921.1 Clone du VIH-1 7301env gène de la glycoprotéine d'enveloppe (env), CDS partiels
AATGGCAGTCTAGCAGAAGGAGAGGTAATGGTTAGATCAAAAATTTCTCGGACAATACTAAAGTCATAATAGTACAACACTGAATGAAGCTATAG
AAATTAATTGTACAAGACCCAACAACAATACAAGAAAAAGTATACATATGGGACAAGGCAGRCATGGTATACAACAGGAGATATAATAGGAA
ATATAAGACAAGCACATTGTAACCTTAGTAGAGCAGACTGGAATGACACTTTAAGAAGGATAGCTATAAAAATTAAGAGAACAATATGAGAATG
CAACAATAGTCTTTAATAAATCTCAGGAGGGGACCCAGAAATTGACTGCACAGCTTTAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828922.1 Clone du VIH-1 7401env gène de la glycoprotéine d'enveloppe (env), CDS partiels

TTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGGTCTGAAAATTTTACGGMCAATACTAAAACCATRATAGTACAGYTGAATGAATCTG
TAAACAATTAATTGTACCAGACCCAACAACAATACAAGAAAAAGTATAAGTATAGGACCAGGGAGAGCATTATACAACAGGAGACATAATAG
GAGATATAAGACAGGCACATTGTAACCTTAGTAGAACACAATGGAATAACACCTTAAAAACAGGTAGTTGCCAAATTAAGAGAACAATTAATAA
AACAATAGTCTTTAATCACTCCTCGGGAGGGGACCCAGAAATTGTAATGCACAGTTTTAATTGTGGAGGGGAATTT

> AY828923.1 gène de la glycoprotéine d'enveloppe (env) du clone 7402env du VIH-1, CDS partiels

GGCAGTCTAGCAGAAGAAGACATAAGAATTAGATCTGAAAATCTCACAGACAATGCCAAAGTCATAATAGTACAGTTTAAATGAGACTGTAGAG
ATTAATTGTACTAGACCCAACAACAATACAAGAAAAAGTGTARCAKTCGGACCAGGACAAGCGATCTATGCAACAGGTGACATAATAGGAGAC
ATAAGACAAGCACATTGTAATGTTAGTAAAAGCAAATGGAAGGAGATGTTACAGAGAGTTGAAACACAGCTAAAGAATACCTATAAGAATATC
TATAATAGTACTGGGAACATGCATTTGMCTCATCCGAGGAGGGGACCTGGAAATTACAACACACAGTTTCAATTGTAGAGGGGAATTTTACT
AC

> AY828924.1 gène de la glycoprotéine d'enveloppe (env) du clone 7501env du VIH-1, CDS partiels

AATGGCAGTCTAGCAGAAGAAGACATAGTAATTAGATCTGAAAATCTCACAGACAATGCTAAAATCATAATAGTACATCTGAATGAATCTATAC
AAATTAATTGTACAGACCAGGCAACAATACAAGAAAGTATACATATAGGGCCAGGGAAGGCATTCTATGCAACAGGAGACATAATAGGAG
ATATAAGACAAGCACATTGTAACATTAGTAAAGCAAAGATGGAATAGCACTTAAATAGGGTACTAACAAATTAAGAGAACAATTTGGAATA
AGACAATAGTCTTTAATCAATCCTCAGGAGGGGATCCAGAAATTGTGACACACAGTTTTAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828925.1 Clone du VIH-1 7703env gène de la glycoprotéine d'enveloppe (env), CDS partiels

AATGGCAGTCTAGCAGAAGAAGACATAGTAATTAGATCTGAAAATCTCACAGACAATGCTAAAATCATAATAGTACATCTGAATGAATCTATAC
AAATTAATTGTACAGACCAGGCAACAATACAAGAAAGTATACATATAGGGCCAGGGAAGGCATTCTATGCAACAGGAGACATAATAGGAG
ATATAAGACAAGCACATTGTAACATTAGTAAAGCAAAGATGGAATAGCACTTAAATAGGGTACTAACAAATTAAGAGAACAATTTGGAATA
AGACAATAGTCTTTAATCAATCCTCAGGAGGGGATCCAGAAATTGTGACACACAGTTTTAATTGTGGAGGGGAATTTTCTACTGTAA

> AY828926.1 Clone du VIH-1 7903env gène de la glycoprotéine d'enveloppe (env), CDS partiels

AATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGAAAATTTACGAAACAATGCTAAAATTATAATAGTACAGCTGAATGAATCTGTAC
AAATTAATTGTACAGACCACAACAATACAAGAAAAAGTGTACATATGGGACCAGGAGAGCATTATATRCAACAGGAGATATAATAGGAG
ATATAAGAAAAGCATATTGTACAATTAATGAAACAAAATGGAATAACACTTAAAGACACATAGTTAAAAAATTAAGAGAACAATTCAGAATA
AAACAATAATCTTTAATCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCAYAGTTTTAATTGTGGAGGGGAATTTTCTACTGT

> AY828927.1 Clone VIH-1 8202env gène de la glycoprotéine d'enveloppe (env), CDS partiels

GTACAGATGATACGATTCTGTGAAAATTTACGGACAATGCTAAAACCATAATAGTACAGCTGAACGATTCTGTAAAAATGAGTGTDTAAGAC
CCAACAACAATACAAGAAGAAGTATACATATVGGACCAGGAGAGCATTATGCAACAGGAGMCATAACAGGAGATATAAGAAAAGCATATT
GTAATCTTAGTAGAGTAGATTGGAATAAAACCTTACACCAGATAGTCGAAAATTAAGAGAACAATTCGAACAATCAAGAATAAAACAATAGT
CTTTAATCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCACAGTTTTAATTGTGGAGGGGAA

> AY828928.1 Clone VIH-1 8203env gène de la glycoprotéine d'enveloppe (env), CDS partiels

AATGGCAGTCTAGCAGAAGGAGAAATAGCAATCAGATCTGAAAATCTCACAAACAATGCCAAGACCATAACAGTACAGTTAGTTGAGCCTGTA
AGAATTAATTGTAGTAGACCTAACACAATACAAGAAAAAGTATACGCATAGGACCAGGACAACATTCTATGCACACGGTGAGATAATAGGG
AATATAAGACAAGCACATTGTAATATTAGTAGGAAAGACTGGAATAACACATTACAAAAGGTAGCTGCACAATTAAGGAAACACTTTAATAATG
ACACAAGAATAATCTTTACTAAACATTCAGGAGGGGATTTAGAAATTACAACACATAGTTTTAATTGTAGAGGGGAATTTTCTACTGTAA

> AY828929.1 Clone VIH-1 8303env gène de la glycoprotéine d'enveloppe (env), CDS partiels

TAAATGGCAGTCTAGCAGAAGAGGAKGTAGTGATTAGATCTGAAAATCTCACAAACAATGCCAAAACCATAATAGTGCAGTTGGMTAAGCCTG
TAAACATTACTTGTATTAGACCTGGCAACAATACAAGAAAAAGTGTACGTATAGGGCCAGGCAACATTCTATGCAACAGGTGACATAACAGG
GGATATAAGACGAGCACATTGTGAGGTCAATAGCACAAAGTGAATCAARCTTTAACACAGGTGGTGAYACAATTAARGRAGCGTTTTAATAAT
AATACAATAATCTTTAAKRACCAYKAGGAGGGGATATAGAAATTACAACACATAGTTTTAATTGTAGAGGGGAATTTTCTACTGTAA

> AY828930.1 Clone VIH-1 8401env gène de la glycoprotéine d'enveloppe (env), CDS partiels

AATTTACGAAACAATGCTAGAAACATAATAGTACAGCTGAATGAACCTGTACAAATTCAGTGTATAAGACCCGATAACAAGATACAAAGAAGG
AGGATACATATAGGACCAGGACAGCATTGTTACAACAGGAGGAGACATAAGAAAAGCATATTGTAATGTTAGTAGAACAGATTGGAAT
AACGCTTTAAACCAGACAGCTATAAAATTAAGAGAACAATTAAGAATAAAACAATAGTCTTTAATCACTCCTCAGGAGGGGACACAGAAAT

> AY828931.1 Clone VIH-1 8402env gène de la glycoprotéine d'enveloppe (env), CDS partiels

AAATGGCAGTCTAGCAGAAGGAGAGATAGTGATAAGATCTGAAAATATCTCAAACAATGCCAAAACCATAATARTACAYTTAAATRAGRCTGTA
AGAATTAATTGTACCAGACTGGCAACAATACAAGAAAAAGTGTACGTATAGGACCAGGCAAGTATTCTATGCAACAGGAGAAATAACAGGG
GATATAAGACAAGCAYATTGTAATGTCARTGGAACAGAATGGWATRAMATGTTGAAAARGGTGGCTRGRAATTAAGGAGATACTTTGRSAAT
ACCACARAMATAATCTTTAMTRMGWCCTCAGGAGGGGATTTAGAAATTACAACACATAGTTTTAATTGTGGAGGGGAATTTTCTACTGTAAA

> AY828932.1 Clone du VIH-1 8701env gène de la glycoprotéine d'enveloppe (env), CDS partiels

AATGGCAGTCTAGCAGAAGAAGAGGTAGTGATTAGATCTAGAAATTTCTCRGAAAATGRGGATAGCATAATAGTACAGCTGAATGCATCTGTAG
AAATTAATTGTACAAGACCCAACAACAATACAAGAAAAAGTATACATATAGGACCAGGTAGAGCATTATACAACAGGAGACATAATAGGAG
ATATAAGACAAGCACATTGTAACATTAGTGGAGAAAAATGGGATAWCACCTTAGAACAGATAGCTATACAATTAAGARGACAATATGGAGAKA
ATAAAAACAATAGTCTTTACCAACATTCAGGGGGGACCCAGAAATTGTAATGCACAGTTTTAATTGTGGAGGGGAATTT

> AY828933.1 Clone du VIH-1 8901env gène de la glycoprotéine d'enveloppe (env), CDS partiels

GCCAAATTCAGTGACAATACTAAAATCATAAATAGTACAGCTGAATGAATCTGTAATAATTAATTGTATAAGACCCAACAACAATACAAGAAAA
GTATACATATAGGACCAGGTAAGCATTATGCAACAGGAGACATAATAGGAGATATAAGACAAGCACATTGTAACATTAGTGCARCAGARTG

GAATAACACCTTAAACAGGTAGCTATAAAATTAAGAACAATTTGGAAATAATAAAACAATAGTCTTTAAACAATCTTCAGGAGGGGACCCA
GAAATTGTAATGCACAGTTTAAATTGTGGAGGGGAATTT

> AY828934.1 gène de la glycoprotéine d'enveloppe (env) du clone du VIH-1 9001env, CDS partiels

CTAGCAGAAGAAGAGGTAGTAGTTAGATCTGAAAATTTTCAGAACAAATGCTAAAATAATAATAGTACAGCTGAAGGAAGCTGTAAACAATTAATT
GTACAAGACCAGCAACAATACAAGAAGAAGTATACCTTTAGGACCAGGCAGAGCAATTTATACAACAGACATAATAGGAGATATAAGACAAG
CACATTGTAACGTTAGTAAAACAAAATGGGAGGGCAGCTTTAAAACAGATAGTTGAAAAATTAAGAGAACAATTTAAAAATAAAACAATAGTCTT
TAATCAATCCGACAGGGGAGACCCAGAAATTTAGCGCACAGTTTAAATTGTGGAGGGGAATTT

> AY828935.1 gène de la glycoprotéine d'enveloppe (env) du clone 9101env du VIH-1, CDS partiels

CTAGCAGAAGAAGAAGTAGTAATTAGATCTCCAATTTTCACAGACAATGCTGCAATCATAATAGTACAGCTAAATGAATCTGTAGTAATTAATT
GTACAAGACCAACAACAATACAAGAAAAAGTATACATATRGACCAGGCAGAGCATGGTATGCAACAGGAGRCATAATAGGAGATATAAGAC
AAGCACATTGTAACATTAATAGAACAATGGAATRACACTTTAAAACAAATAGTTAGAAAATTAAGAGAACAATTTCCAAATAAGACAATCAT
CTTTAATCAATCTCAGGAGGGGACCCAGAAATTTAGCGCACAGTTTAAATTGTGGAGGGGAATTT

> AY828936.1 Clone du VIH-1 9301env gène de la glycoprotéine d'enveloppe (env), CDS partiels

AATGGCAGTCTAGCAGAAGACGAGGTAGTGATTAGATCTGAAAATATCACAACAATGCCAAAATCATAATAGTACAGTTAAATAAGCCTGTAA
AAATCAATTGTACCAGACCTAACACAATACAAGAGAAAGTTGGCGTATAGGACCAGGGCAAACATTCTATGCAACAGGTGAAATAGTAGGGA
ATATAAGACAAGCACATTGTRATGTCARTAAAACAGAATGGGATGCAACTTTAAAACAGGTAGCTACACAATTAAGGAAGCACTTTAAAAATGC
CACACAATAATCTTTGCTAACTCCTCAGGAGGGGATTTAGAAATWACAACATAGTTTTATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828937.1 Clone du VIH-1 9502env gène de la glycoprotéine d'enveloppe (env), CDS partiels

GGCAGTSTAGCAGAAGGAGAAATAAGGGTCAGATCTGAAGATCTCACAACAATGCCAAGACCATAACAGTACAGTTAGYTGAGCCTGTAA
ATTAATTGTASCAGACCTRRCAACAATACAAGAACARGTRTACGTATAGGACCAGGACAAACATTCTATGCAACAGGTGGMATAATAGGGGATA
TAAGACAAGCACATTGTAATATYAGTAGAAAAGACTGGAATAACACATTACAAAAGGTAGCTGMACAATTAAGGAAACACTTTAAKARYAMCA
CCATAATCTTTGCTRAACACTCAGGAGGGGATGTAGAAATTAACAACATAGTWTAAATTGTAGAGGRGAATWTTTCTACTGT

> AY828938.1 gène de la glycoprotéine d'enveloppe (env) du clone du VIH-1 9602env, CDS partiels

AATGGCAGTCTAGCAGAAGARGAGGTAAATGATTAGATCTGAAAATCTCACAACAATGYCAAAAYCATAATAGTGCAGCTTGTAAASCTGTAA
RYATTAMTTGTAYCAGACCCMACACAATACAAGAACAAGTRTACGTATAGGACCAGGACAAGCATTCTATGCAACAGGTGACATAAYAGGGG
ATATAAGAAAAGCACATTGTAATGTCAGTAAAACARAATGGSATRAAGCTTTARAAGRGGTAGYTRAACAGCTAAGACAACACTTTAAAAATAA
ACAATAAAAATTTACRAATCTCAGGAGGGGATATAGAAATTAACAACACAGCTTTAATTGTGGAGGGGAATTTTTCTACTGT

> AY828939.1 Clone du VIH-1 10403env gène de la glycoprotéine d'enveloppe (env), CDS partiels

ATGGCAGTCTAGCAGAAGAAGAGATAGTAATTAGRTCTCAAATTTTCAGCGACTCTGCTAAAACATAATAGTACACCTGAATGCATCTGTAGA
AATTAATTGTACAAGGCCAATAACAATACAAGAAAAAGTATAAATATAGGACCAGGCAGAGCATTTTATACAACAGGACAAATAATAGGAGA
TATAAGAAAAGCATATTGTAACATTAGTAGAGCAAAATGGAATGACTTTARAAWAGATAGTTGTAATAATTGAGAAAACAGTTTGAAAYAA
AACAATAATCTTTAAKACTCCTCAGGAGGGGACCCAGAAATTTGTAATGCACAGTTTAAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828940.1 Clone du VIH-1 10503env gène de la glycoprotéine d'enveloppe (env), CDS partiels

GCCAAATCAGTGACAATACTAAAATCATAATAGTACAGTGAATGAATCTGTAATAATTAAATTGTATAAGACCAACAACAATACAAGAAAA
GTATACATATAGGACCAGGTAAGCATTATGCAACAGGAGACATAATAGGAGATATAAGACAAGCACATTGTAACATTAGTGAACAGAGT
GGAATAACACCTTAAAMCAGGTAGCTATAAAATTAAGAACAATTTGGAAATAATAAAACAATAGTCTTTAAACAMTCTTCAGGAGGGGAYC
CAGAAATTGTAATGCACAGTTTAAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828941.1 Clone du VIH-1 10602env gène de la glycoprotéine d'enveloppe (env), CDS partiels

AAATGGCAGTCTAGCAGAAGGAAACATAACAATTAGATCTGAAAACATCTCAGACAATGCCAAAACCATAATAGTACAGTTAGTTAATCCTGTAA
AATATCACCTGTATCAGACCTAACACAATACAAGAAGAGGTGTACATATAGGACCAGGAAAAACATTCTATACAGGTGAAATATAGGGGAT
ATAAGACAAGCACATTGTAATGTCAGTAGGAAAAATGGAACAATCTTTAACACAGGTGGTACACAATTAAGGAGTTACTTTAAGAATGATA
CAACAATAATCTTTGATAAACCTCAGGAGGGGATTTAGAAGTTACAACATTTAGTTTAAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828942.1 Clone du VIH-1 10603env gène de la glycoprotéine d'enveloppe (env), CDS partiels

AATGGCAGTCTAGCAGAAGGAGAAATAGCGATCAGATCTGAAAATTTACAACAATGCCAAGATCATAACAGTACAGTTAGTTGAKCCTGTAA
GAATYAATTGTAGCAGACCTAAYAACAATACAAGAAAAAGTATACGTATAGGACCAGGCAAACATTCTATGCAACAGGTGACATAATAGGGA
ATATAAGAGAAGCACACTGTAATATTAGTAGAAAAGATTGGARTAAAACATTACAAAAGGTAGCTGAACAATTAAGGACACACCTTAAGAATA
ACACAACAATAACCTTTGCTAAACTCAGGAGGGGATTTAGAAATTAACAACATAGTTTAAATTGTAGAGGGGAATTTTTCTACTGTAA

> AY828943.1 Clone du VIH-1 10702env gène de la glycoprotéine d'enveloppe (env), CDS partiels

AAATGGCAGTCTAGCAGAAGRAGAGGTAAATCATTAGATCTGAAAATATCACAACAATGCCAAAACATAATAGTACAGTTGACCAAGTMTGT
AAACATTAGTTGTAAGACCTAACATAATACAAGGAAAAGTGTACSTATAGGRCCAGGACRAGCATTCTATGCAACAGGTGATRTAATAGGA
GAYATAAGACAAGCATATTGTAARRTCAATGGAACAGAATGGTCCGAAGCTTTAAACCAGACAAAARCAACAATTARGAAAATACTTTGGTAACA
AAACAATARTCTTTCAAACCTCAGGAGGGGACTTAGAAATTAACAACATAGTTTAAATTGTGTRAGGGGAATTTTTCTACTGTAA

> AY828944.1 Clone du VIH-1 11202env gène de la glycoprotéine d'enveloppe (env), CDS partiels

ATGGCAGTCTAGCAGAAGCAGGAGAGGTGRTAATTAGATCAGCCAATTTCTCGACAATRCTAAAATCATAATAGTACATTTGAAYACATCTGT
AGCAATTACTTGTAYAAGACCAACAACAATACAAGRAGARGTATACRTATAGGACCAGGYAGMKCAWTGTATGCAACAGGAGAMATAATAG
GARATATAAGACAAGCACATTGTAACATYAATAGAACAATGGAATRACACTYTAGAASAGRTAGTGAARAATTAAGAGAACAATCCSGA
ATAAAACRATRGYATTTARGCCAYCCGYAGGAGGRGACCCAGAAATAGTAATGCAYAGTTTCAATTGTGGAGGGGAATTTTTCTACTGTAA

> AY828945.1 Clone du VIH-1 11302env gène de la glycoprotéine d'enveloppe (env), CDS partiels

AATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGAGAATTTCACTAACAAATGCTAAAACCATCATAGTACAGCTGAATGAATCTGTAG
 WAATTAATTGTACAAGACCAACAACAATACAAGCAAAAGTATACATMTAGGACCAGGGGTGCATTTTATGCAACAGGAGGAATAATAGGAG
 ATATAAGACAAGCACATTGTAACCTTARTAGTGCARRATGGAATRAGACTTTAAACAGGTAGTTGGAAAATTAAGAGAACAATTTGGAAATAA
 TAAGACAATAGTCTTTRAACAACCTCAGGAGGGGACCCAGAAAATTGACTGCATAGTTTAAATTGTGGAGGGGAATTTTCTACTGTAA
 > AY828946.1 Clone du VIH-1 11402env gène de la glycoprotéine d'enveloppe (env), CDS partiels
 AATGGCAGTCTAGCAGAAGAAGAAATAGCGATCAGATYTGAAAATCTCACAACAATGCCAAGACCATAACAGTACAGTTRGTTACTCTGTAA
 GAATTAATTGTACCAGACCTGGCAACAATACAARAAGGGGTATGCTTATAGGACCAGGGCGAAAAGTCTATGTAACAGATAAAAATAATAGGGG
 ATATAAGACGAGCATATTGTAATATTAGTAGAAAAGAATGGAATAACACATTACACAAGGTAGCTGAACAATTAAGGAAACACTTTAAGAATG
 ACACAAGAATAATCTTTGCTAACCACTCAGGAGGGGATTTAGAAATTACAACACATAGTTTAAAYTGTAGAGGGGAATTTTCTACTGTAA
 > AY828947.1 Clone du VIH-1 11505env gène de la glycoprotéine d'enveloppe (env), CDS partiels
 AATGGCAGTCTAGCACAAGGAGAART AATGATTAGATCTGAAAATCTCACAACAATGCCAAAACCATAATAGTACAGCTTGMTGAGCCTGTA
 GAAATTAATTGTACCAGACCCAAY AAY AATACAATAAAAAGTGTACGTATAGGACCAGGGCAAACATTCTATGCAACAGGTAGTATAATRGGG
 GATATAAGACAAGCAY AHTGT AATGTCARTAGARCMATGGRKRGRCTTTACAAGRGGT AGYTRSAMARTTAMGRCAACACTTYAASAAC
 ACAAY AATAAWATTTGCTA ACTCTCAGGAGGGGATTTAGAAATTACAACACATAGYTTAATTGTGGAGGGGAATTTTCTACTGTAA
 > AY828948.1 Clone du VIH-1 11602env gène de la glycoprotéine d'enveloppe (env), CDS partiels
 AATGGCAGTCTAGCAGAAGACRAGGTAGTGATTAGATCTGAAAATATCACAACAATGCCAAAATCATAATAGTACAGTTAAATAAGCCTGTAA
 AAATCAATTGTACCAGACCTAACACAACAATACAAGARAAGTTGGMGTATAGGACCAGGGCAAACATTCTATGCAACAGGTGAHATARTAGGGR
 ATATAAGACAAGCACATTGTAATGTCAGTAAAACAGAATGGAATGCAACTTACAACAGGTAGCTRCACAATTAAGGACGYACTTTAGTAATRV
 CACAACAATAATCTTTRCTA ACTCTCAGGAGGGGATTWGAAAATTACAACWCATAGTTTATTTGTGGAGGGGAATTTTCTACTGTAA

3.2.2. Capture des sorties de résultat fournie par logiciel BEAST-2

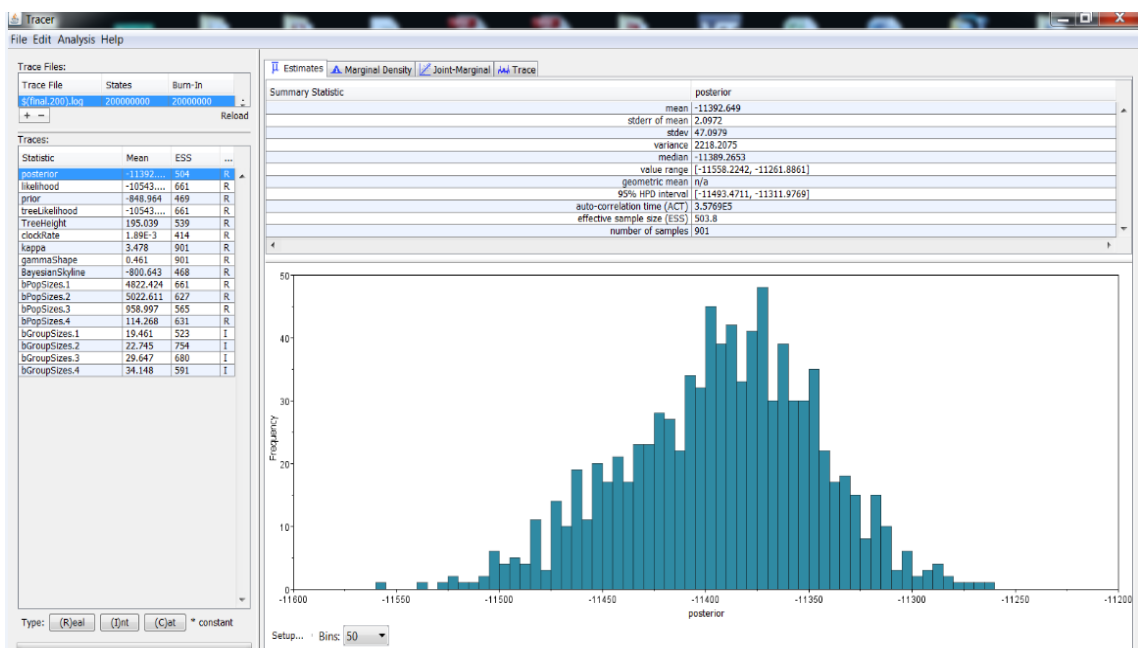


Figure 19 : Traceur montrant un résumé de la série BEAST2 de données des séquences ENV du VIH-1 en Algérie avec une longueur de chaîne MCMC de 200×10^8 .

Tracer fournit quelques statistiques récapitulatives utiles sur les résultats de l'analyse. Sur le côté gauche de la fenêtre supérieure, il fournit une liste des fichiers journaux chargés dans le programme pour le moment. La fenêtre ci-dessous montre la liste des statistiques enregistrées dans chaque fichier. Pour chaque statistique, il donne une liste de valeurs récapitulatives telles que la moyenne, l'erreur standard, la médiane et d'autres qu'il peut calculer à partir des

données. Les valeurs récapitulatives sont affichées dans la fenêtre en haut à droite et un histogramme montrant la distribution de la statistique se trouve dans la fenêtre en bas à droite.

Le fichier journal contient des traces pour le postérieur (c'est le logarithme naturel du produit de la vraisemblance de l'arbre et de la densité a priori), a priori, la vraisemblance, les probabilités de l'arbre et d'autres paramètres continus. La sélection d'une trace sur la gauche fait apparaître les statistiques récapitulatives de cette trace sur le côté droit. Lors de la première ouverture, la trace postérieure est sélectionnée et diverses statistiques de cette trace sont affichées sous l'onglet Estimations.

Pour chaque fichier journal chargé, nous pouvons spécifier un Burn-In, qui est affiché dans le tableau de la liste des fichiers (en haut à gauche) dans Tracer. Le rodage est destiné à donner à la chaîne de Markov le temps d'atteindre sa distribution d'équilibre, en particulier si elle est partie d'un mauvais point de départ. Un mauvais point de départ peut conduire à un sur échantillonnage des régions du postérieur qui ont en fait une très faible probabilité sous la distribution d'équilibre, avant que la chaîne ne s'installe dans la distribution d'équilibre. Burn-in nous permet de supprimer simplement les premiers échantillons d'une chaîne et ne pas les utiliser pour calculer les statistiques récapitulatives. La détermination du nombre d'échantillons à rejeter n'est pas un problème trivial et dépend de la taille de l'ensemble de données, de la complexité du modèle et de la longueur de la chaîne. Une bonne règle de base est de toujours jeter au moins les 10 premiers% de la longueur totale de la chaîne lors du rodage (cependant, dans certains cas, il peut être nécessaire de jeter jusqu'à 50% de la chaîne MCMC).

3.2.3. La date du premier cas de sida en Algérie (tMRCA) avec le taux d'évolution :

Le tMRCA pour le VIH-1 en Algérie est estimé 1985, c'est-à-dire le premier cas de SIDA découvert en Algérie. Ce qui vas en concordance avec la déclaration de L'ONUSIDA ALGERIE

([https://www.unaids.org/sites/default/files/country/documents/DZA_narrative_report_2015.p](https://www.unaids.org/sites/default/files/country/documents/DZA_narrative_report_2015.pdf)

[df](https://www.unaids.org/sites/default/files/country/documents/DZA_narrative_report_2015.pdf)). Nous avons aussi estimé le taux développement du SIDA en Algérie qui est $1.89 \cdot 10^{-3}$ /site/ans.

3.2.4. Réparation phylogénétique des cas de sida en Algérie

L'analyse phylogénétique de 129 échantillons de VIH-1 d'Algérie et de référence des sous-types de VIH-1 disponibles dans GenBank donne l'arbre phylogénétique représentée dans la figure 20, et la carte géographique représentée dans la figure 21 .

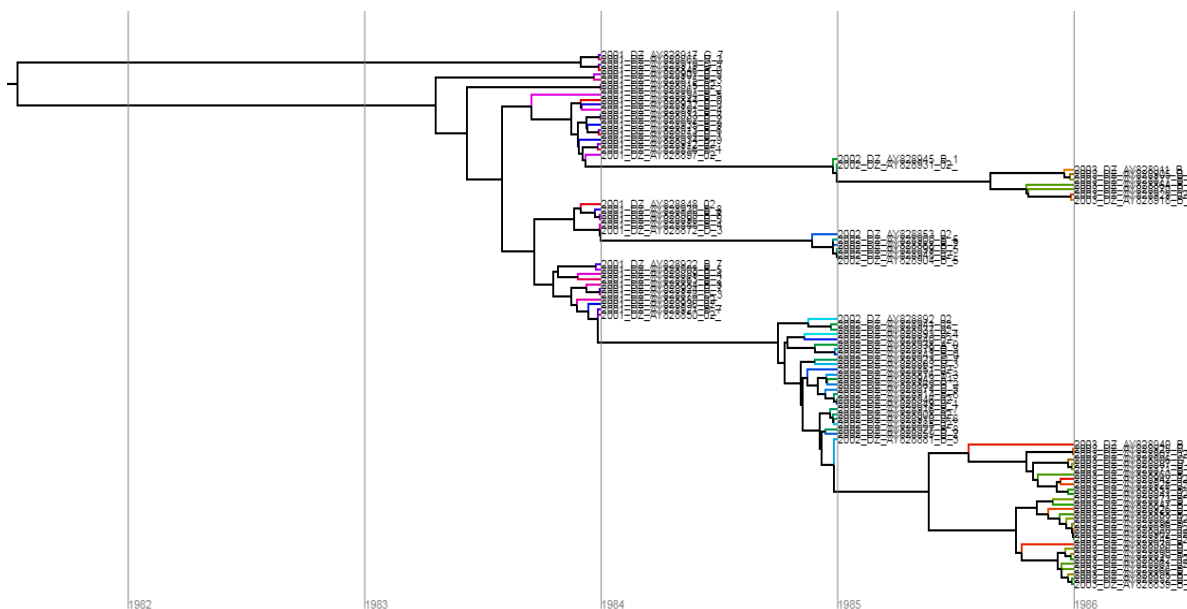


Figure 20 : L'arbre phylogénétique obtenu à partir de l'alignement des gènes env.

Il ressort clairement de l'arbre phylogénétique que le sous-type B est prédominant. À la suite du sous-type B, on note une grande diversité de VIH-1 : CRF02_AG et CRF06_cpx, qui ont été décrits dans d'autres pays d'Afrique subsaharienne et en particulier dans les environs de l'Algérie. Recombinants entre CRF02_AG et CRF06_cpx sont fréquemment détectés comme observé au Niger. Les autres sous-types sont G, A et D. En raison de la grande diversité du VIH-1 dans cette partie du monde et de la recombinaison lors des co-infection, les recombinants inter sous-types et / ou CRF peuvent être enregistrés : B / CRF06, B / CRF03, B / CRF02, A / CRF02, G / CRF06, G / CRF02, D / CRF05, CRF09 / CRF02, B / F2, A / H, B / CRF06 / CRF02 et G / CRF02 / B.

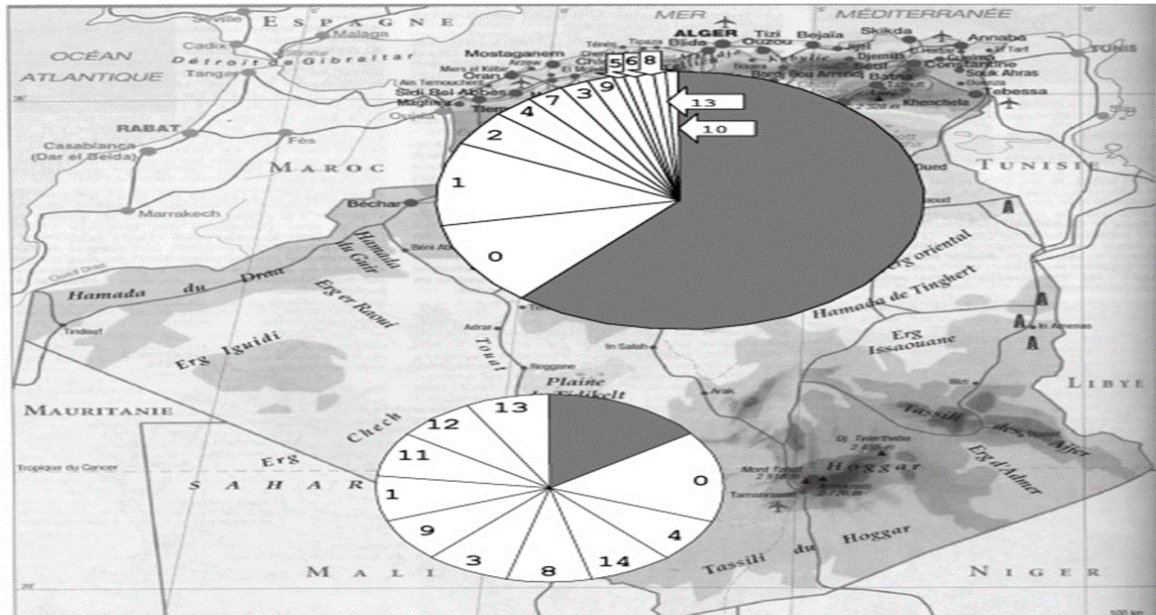


Figure 21 : carte indiquant la répartition géographique des échantillons [gènes env] (bouzeghoub *et al.*, 2006).

La figure 21 montre que dans le sud, le sous-type B est moins prédominant tandis que CRF02_AG, CRF06_cpx et les recombinants entre ces CRF sont plus fréquent. Il est à noter que les pays subsahariens limitrophes de la partie sud de l'Algérie offrent une grande diversité de VIH- 1, y compris les différents CRF mentionnés ci-dessus. L'introduction de ces virus en Algérie peut résulter d'une activité d'échange entre les différentes populations de cette région d'Algérie. En outre, la partie sud de l'Algérie est un point de rencontre important pour de nombreux groupes qui migrent des pays subsahariens vers l'Europe.

Conclusion Générale

L'étude moléculaire visant à déterminer l'origine géographique et temporelle de l'épidémie du VIH en Algérie utilise uniquement des outils phylogénétiques classiquement utilisés. Dans cette étude nous avons récupéré toutes les séquences env du VIH-1 disponibles dans la base de données sur le VIH du laboratoire national de Los Alamos, et pour lesquelles les origines de collecte sont connues.

Un arbre de maximum de vraisemblance (PhyML) est calculé afin de déterminer l'origine géographique de l'épidémie du VIH-1 et l'origine temporelle est étudiée avec le logiciel BEAST. Différents modèles d'horloges moléculaires (stricte, et en log-normal) sont utilisés pour estimer la date de l'ancêtre commun de SIDA en Algérie.

Les différentes analyses révèlent que l'ancêtre commun aux souches du VIH-1 collectées en Algérie est daté au 1985, et aussi permet d'identifier les différents géotypes de VIH-1 circulant en Algérie. De plus, nous avons estimé le taux d'évolution du SIDA en Algérie qui est $1.89 \cdot 10^3$ /site/ans.

En conclusion, nous présentons ici les premières données sur l'origine temporelle du VIH-1 en Algérie, avec une autre donnée qui montre qu'il existe une grande diversité de virus dans ce pays.

Bibliographie

Bibliographie

Alfadhli, A.; Still, A.; Barklis, E., Analysis of human immunodeficiency virus type 1 matrix binding to membranes and nucleic acids. *Journal of virology* 2009, 83 (23), 12196-203.

Anglaret X, Salamon R. Épidémie de sida en Afrique subsaharienne. *Med Sci (Paris)* 2004 ; 20 : 593–8.

Belarbi Amar N. les virus, cours de cytologie de la 1ère année de médecine, université Oran 1 Ahmed Benbella, 2019-2020.

Bigaillon Christine, Mérens Audrey, Rapp Christophe. Intérêt des tests génotypiques de résistance du VIH aux antirétroviraux en pratique clinique quotidienne. *Revue Francophone Des Laboratoires*, 2010(422), 69–82. Doi :10.1016/s1773-035x (10)70512-8.

Bouzeghoub Salima, Jauvin VALÉRIE, Recordon-Pinson PATRICIA, GARRIGUE ISABELLE, AMRANE ACHOUR, BELABBES EL-HADJ, FLEURY HERVÉ J. High diversity of HIV type 1 in Algeria. *AIDS Res Hum Retroviruses*. 2006 ;22 :367-72.

Brun-Vézinet F, Damond F, Descamps D et Simon F. Virus de l'immunodéficience humaine. *Encycl Méd Chir (Editions Scientifiques et Médicales Elsevier SAS, Paris), Maladies infectieuses*, 8-050-B-15, 2000.

Catie.com

Cen, S., Huang, Y., Khorchid, A., Darlix, J.-L., Wainberg, M. A., & Kleiman, L. (1999). The Role of Pr55gag in the Annealing of tRNA³Lys to Human Immunodeficiency Virus Type 1 Genomic RNA. *Journal of Virology*, 73(5), 4485–4488. Doi:10.1128/jvi.73.5.4485-4488.1999.

Chaix ML, Descamps D, Wirden M, et al., Stable frequency of HIV-1 transmitted drug resistance in patients at the time of primary infection over 1996-2006 in France. *AIDS* 2009 ;23 :717-24.

Chollet, Emmanuelle ; Frégaville-Arcas, Olivier ; Rigaud, Nicolas. VIH/sida : nouvelles recherches, nouveaux espoirs. *Science&Santé(Paris)*, 2011, N° 1. URI : <http://hdl.handle.net/10608/6861>, p. 22-35.

Bibliographie

Coffin JM. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. *Science* 1996 ;267 : 483-489.

Cournaud V, Müller-Trutwin M, Sonigo P. Évolution et virulence des lentivirus de primates. *Med Sci (Paris)* 2004 ; 20 : 448-52.

D'arc, M., Ayouba, A., Esteban, A., Learn, G.H., Boué, V., Liegeois, F., Etienne, L., Tagg, N., Leendertz, F.H., Boesch, C., et al. (2015). Origin of the HIV-1 group O epidemic in western lowland gorillas. *Proc. Natl. Acad. Sci. U. S. A.* 112, E1343-1352.

Depatureaux A, Leoz M, De Oliveira F, et al. Specific diagnosis and follow-up of HIV-1 group O infection: RES-O data. *Med Mal Infect* 2010 : doi :10.1016/j.medmal.2010.04.011.

Desenclos J.C, Dabis F, Semaille C. Épidémiologie du VIH dans le monde : particularités de l'épidémie au Nord et au Sud. *Virologie*. Volume 17, Numéro 3, 132-44, Mai-Juin 2013

El Anbassi Sarra, Bianchi Vincent, Duployer Claire (2019) [en ligne]. Bactériologie virologie. 2 ème édition imprimé français. Louvain-la-Neuve : De Boeck supérieur. 192 p. (prépa pharma) disponible sur : https://books.google.dz/books?id=9J6WDwAAQBAJ&dq=livre+bact%C3%A9riologie+virologie&hl=fr&source=gbs_navlinks_s.

Fleury Hervé.J. HIV-1, sous-types et CRF : différences dans la réponse aux traitements et la résistance aux antirétroviraux. Journées Québécoises sur l'infection VIH, Montréal, Canada, mars 2010.

Geretti AM. HIV-1 subtypes: epidemiology and significance for HIV management. *Curr Opin Infect Dis* 2006 ;19 : 1-7.

Ghanam, R. H.; Samal, A. B.; Fernandez, T. F.; Saad, J. S., Role of the HIV-1 MatrixProtein in Gag Intracellular Trafficking and Targeting to the Plasma Membrane for Virus Assembly. *Front Microbiol* 2012, 3, 55.

Gilles Furelaud, Benjamin Pavie. Le virus du SIDA, 2002.

Heeney JL, Dalglish AG, Weiss RA. Origins of HIV and the evolution of resistance to AIDS. *Science* 2006 ; 313 : 462-66.

Bibliographie

Helaers, Raphaël. Développements en phylogénomique : comparaisons de génomes et estimation de grandes phylogénies. 115p. Thèse, FUNDP - SBIO_URBM (unité de recherche en biologie moléculaire) FUNDP - Ecole doctorale en sciences, faculté des sciences département de biologie. Prom. : Depiereux, Éric ; Milinkovitch, Michel.

Hemelaar, J. (2012). L'origine et la diversité de la pandémie de VIH-1. *Trends in Molecular Medicine*, 18 (3), 182–192. Doi : 10.1016 / j. molmed.2011.12.001.

Holmes EC (2004) The phylogeography of human viruses. *Mol. Ecol.* 13 : 745-756.

Holmes EC (2008) Evolutionary history and phylogeography of human viruses. *Annu. Rev. Microbiol.* 62 : 307-328.

<http://www.hiv.lanl.gov/components/sequence/HIV/geo/geo.comp>, Los Alamos HIV-1 database.2011.

<http://www.ncbi.nlm.nih.gov/Genbank/>

https://www.unaids.org/sites/default/files/country/documents/DZA_narrative_report_2015.pdf

https://www.voyagesphotosmanu.com/carte_maghreb.html

Huraux Jean Marie. *Virologie*, 2006-2007.

Huraux Jean-Marie, NICOLAS Jean-Claude, AGUT Henri, PEIGUE-LAFEUILLE Hélène. *Virologie Médicale*. Édition Estem 2003, P699.

James MM, Wang L, Musoke P, Donnell D, Fogel J, et al. Association of HIV Diversity and Survival in HIV-Infected Ugandan Infants. *PLoS ONE* 6(4) : e18642. 2011.

L'établissement public de télévision algérienne « Sida : 673 nouveaux cas au 30 septembre 2019 » [en ligne] disponibles sur le site : [<https://www.entv.dz/sida-673-nouveaux-cas-au-30-septembre-2019/>].

Lemey P, Pybus OG, Wang B, et al. Tracing the origin and history of the HIV-2 epidemic. *Proc Natl Acad Sci USA* 2003 ; 100 : 6588-92.

Bibliographie

Leoz M, Chaix ML, Delaugerre C, Rivoisy C and al. Circulation of multiple patterns of unique recombinant forms B/CRF02_AG in France: precursor signs of the emergence of an upcoming CRF B/02. *AIDS*. 2011 Jul 17 ;25(11) :1371-7.

Liégeois F, Lafay B, et al. Full-length genome characterization of a novel simian immunodeficiency virus lineage (SIV_{olc}) from Olive Colobus (*Procolobus verus*) and new SIV_{wrcPbb} strains from western Red Colobus (*Piliocolobus badius badius*) from the Tai forest in Ivory Coast. *J Virol* 2009 ; 83(1) :428-439.

Matthieu Jung. Évolution du VIH : méthodes, modèles et algorithmes. Bio-informatique [q-bio.QM]. Université Montpellier II - Sciences et Techniques du Languedoc, 2012. Français. (tel-00842785).

Mourez T, Simon F, Plantier JC. Non-M variants of human immunodeficiency virus type 1. *Clin Microbiol Rev* 2013 ; 26(3) :448-461.

Negroni, M., and H. Buc. 2001. Mechanisms of retroviral recombination. *Annu Rev Genet* 35:275-302.

Pancera, M., Majeed, S., Ban, Y.-E. A., Chen, L., Huang, C. -c., Kong, L., ... Kwong, P. D. (2009). Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility. *Proceedings of the National Academy of Sciences*, 107(3), 1166–1171. Doi :10.1073/pnas.0911004107.

Pandrea I, Silvestri G, Apetrei C. AIDS in african nonhuman primate hosts of SIVs: a new paradigm of SIV infection. *Curr HIV Res*. 2009 Jan ; 7(1) : 57-72.

Peeters, M., Chaix, M.-L., et Delaporte, E. (2008). Phylogénie des SIV et des VIH. *Médecine / sciences*, 24 (6-7), 621–628. Doi : 10.1051 / medsci / 20082467621.

Pepin J. *The Origins of AIDS*: Cambridge University Press, Cambridge, UK, First edition, 2011.

Plantier JC, Leoz M, Dickerson JE, et al. A new human immunodeficiency virus derived from gorillas. *Nat Med* 2009 ;15 :871—2.

Plantier, J.-C., & Simon, F. (2011). Les variants rares du VIH-1. *Journal Des Anti-Infectieux*, 13(3), 170–183. Doi : 10.1016/j.antinf.2011.03.009.

Bibliographie

Roberts, J.D., Bebenek, K., and Kunkel, T.A. (1988). The accuracy of reverse transcriptase from HIV-1. *Science* 242, 1171–1173.

Roquebert, B., Damond, F., Brun-Vézinet, F., & Descamps, D. (2009). Diversité génétique des VIH et ses conséquences. *Pathologie Biologie*, 57(2), 142–148. Doi : 10.1016/j.patbio.2008.04.004.

Ross, Lisa; Johnson, Michael; DeMasi, Ralph; Liao, Qiming; Graham, Neil; Shaefer, Mark; Clair, Marty St. Viral genetic heterogeneity in HIV-1-

infected individuals is associated with increasing use of HAART and higher viremia. & 2000 Lippincott Williams & Wilkins *AIDS* 2000, Vol 14 No 7.

Saimot AG, Coulaud JP, Mechali D, et al. HIV-2/LAV-2 in Portuguese man with AIDS (Paris, 1978) who had served in Angola in 1968–74. *Lancet* 1987 ;1 :688.

Shan Cen, Yue Huang, Ahmad Khorchid, Jean-Luc Darlix, Mark A. Wainberg, Lawrence Kleiman, The Role of Pr55gag in the Annealing of tRNA³Lys to Human Immunodeficiency Virus Type 1 Genomic RNA 1998.

<http://www.hiv.lanl.gov/components/sequence/HIV/geo/geo.comp>, Los Alamos HIV-1 database.2011.

Vallari, A., Holzmayer, V., Harris, B., Yamaguchi, J., Ngansop, C., Makamche, F., ... Brennan, C. A. (2010). Confirmation of Putative HIV-1 Group P in Cameroon. *Journal of Virology*, 85(3), 1403–1407. Doi :10.1128/jvi.02005-10.

Van Heuverswyn F and Peeters M. (2008). Human Immunodeficient Viruses: Origin. In: Mahy B.W.J. and van Regenmortel M.H.V. (Editors) *Encyclopedia of Virology*, vol 2, Academic Press, New York, p. 525-534.

Van Heuverswyn F, Li Y, Neel C, et al. Human immunodeficiency viruses: SIV infection in wild gorillas. *Nature* 2006 ;444 :164.

VandeWoude S, Apetrei C. Going wild: lessons from naturally occurring T-lymphotropic lentiviruses. *Clin Microbiol Rev* 2006 ;19 : 728-62.

Bibliographie

Vessièrre, A., Delaugerre, C., Plantier, J.-C., & Simon, F. (2010). Passages inter-espèces des virus de l'immunodéficience humaine et simienne chez les primates. *Revue Francophone Des Laboratoires*, 2010(423), 43–53. Doi :10.1016/s1773-035x (10)70559 1.

Vidal N, Frange P, Chaix ML, Mulanga C, Lepira F, Bazepeo SE, et al. Characterization of an old complex circulating recombinant form, CRF27_cpx, originating from the Democratic Republic of Congo (DRC) and circulating in France. *AIDS Res Hum Retroviruses* 2008 ;24 :315–321.

Visseaux B, Le Hingrat Q, Damond F, Charpentier C, Descamps D. Physiopathologie de l'infection par le VIH-2. *Virologie* 2019 ; 23(5) : 277-91 doi :10.1684/vir.2019.0789.

Zheng NN, Kiviat NB, Sow PS, et al. Comparison of human immunodeficiency virus (HIV)-specific T-cell responses in HIV-1- and HIV-2-infected individuals in Senegal. *J Virol* 2004 ; 78 :13934-42.

Simon, F. et al. (1998) Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nat. Med.* 4, 1032–1037.

Kobbi, N et Saadi, L. (2016). Les méthodes de reconstruction phylogénétique : exploitation des logiciels et application au genre *Canis*. 79p. Mémoire de Master, Faculté de Science de la nature et de la vie, Département de Biologie et Physiologie Cellulaire, Université de Blida -1-, Algérie.

Jeffrey Rizzo and Eric C. Rouchka, “Review of Phylogenetic Tree Construction” University of Louisville, Bioinformatics review(report) 2007.

Taguan (2015). Cours : Les arbres phylogénétiques. Zeste de savoir. P29. <https://zestedesavoir.com/articles/pdf/223/les-arbres-phylogenetiques.pdf>.

Perière G., 2000. Bases de données et outils d'analyse pour la génomique bactérienne. Mémoire de l'habilitation à diriger des recherches. Univ. Claude Bernard, LYON 1. 100 p.

Golding B., Morton D., 2003. Elementary Sequence Analysis. Department of Biology. McMaster. Univ. Hamilton, Ontario. 190 p.

Blanchet C., 1999. Logiciel MPSA et ressources bioinformatiques client-serveur Web dédiés à l'analyse de séquences de protéine. Thèse de doctorat en Bioinformatique. Dir. Thèse : Deleage G. Univ. Claude Bernard-Lyon I : 23-36. 158p.

Bibliographie

Guindon S., Gascuel O., 2003. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood (PHYML). *Syst. Biol.* 52(5): P 696–704.

Pillet V., 2000. Méthodologie d'extraction automatique d'information à partir de la littérature scientifique en vue d'alimenter un nouveau système d'information. Thèse de doctorat en Sciences de l'Information et de la communication. Dir. Thèse : L. Quoniam et B. Jac. Univ. Aix-Marseille III. 20-43. 134p.

Fournel M., Tardif C., 1834. *Traité du voisinage : considérations dans l'ordre judiciaire et dans ses rapports avec le code civil*, Volume 1. Ed. Videcoq, Paris .52p.

Perrière G & Brochier-Armanet C (2010) *Concepts et Méthodes En Phylogénie Moléculaire* 1er éd. Springer Verlag France.

Lanave C, Preparata G, Saccone C & Serio G (1984) A new method for calculating evolutionary substitution rates. *J. Mol. Evol.* 20 : 86-93.

Saitou N & Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4 : 406-425.

Kurtzman C.P., Boekhout T., Fell, J.W. (2010). *The Yeasts, A Taxonomic Study*. 5th Edition. Copyright Elsevier B.V. All rights reserved. Elsevier science: 2354p.

Hennig W. (1966). *Phylogenetic systematic*. University of Illinois Press, Urbana.

Ranwez V., Gascuel O., 2003. "Phylogenetic Reconstruction Algorithms Based on Weighted 4-Trees." *Lecture Notes in Computer Science* 2066 : 337-348.

Swofford D, Olsen G, Waddell P & Hillis D (1996) *Phylogenetic Inference*. In *Molecular Systematics* p 407-509. David Hillis, Graig Moitz, Barbara Mable.

Strimmer K. S. (1997). *Maximum likelihood methods in molecular phylogenetics*. Herbert Utz Verlag.

WWW.diplomatie.gouv.fr

Tamura K., Peterson D., Peterson N., Stecher G., Nei M. & Kumar S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution*, 28(10), 2731-2739.

Glossaires

Glossaries

Monophylétique: un groupe incluant un ancêtre commun et la totalité de ses descendants.

Taxon : Groupe d'êtres vivants ou fossiles qui ont des traits communs.

Robustesse : capacité d'un estimateur statistique à ne pas être modifié par une petite modification dans les données ou dans les paramètres du modèle choisi pour l'estimation.

Méthodes Bayésiennes: qualifie des méthodes d'inférences statistiques fondées sur une évaluation des probabilités des hypothèses, préalablement à l'observation d'un événement aléatoire embranchement (*ou* phylum), classe, ordre, famille, genre et espèce.

Site: une position relative « colonne » dans un alignement de séquences.

Horloge moléculaire : hypothèse selon laquelle les protéines homologues évoluent à la même vitesse quelles que soient les contraintes exercées par l'environnement.

Processus de Markov : Processus probabiliste selon lequel un système passe d'un état à un autre (parmi un nombre fini ou infini) à intervalles de temps réguliers. L'évolution du processus ne dépend que de l'état dans lequel il se trouve présentement. Tout le passé de l'évolution du processus se trouve donc résumé dans son état au dernier instant où on le connaît.

Transition versus transversions : On dit qu'une mutation nucléotidique est une transition lorsqu'elle transforme une purine (A ou G) en purine ou une pyrimidine (C ou T) en pyrimidine. Dans le cas contraire (purine vers pyrimidine ou pyrimidine vers purine), c'est une transversion.

Clade : l'ensemble des taxons qui descendent d'un même ancêtre commun.

FASTA : algorithme proposé par Pearson basé sur l'identification rapide de fragments strictement identiques entre la séquence cible et les séquences de la banque.

UPGMA : une méthode utilisée pour construire des arbres phylogénétiques si les séquences ne sont pas trop divergentes .

Bootstrap : est une méthode souvent utilisée pour tester la fiabilité des branches internes des arbres phylogénétiques.

Glossaires
