



HAL
open science

Dynamique adaptative des virus hautement variables à un nouvel environnement répliatif

Christophe Rodriguez

► **To cite this version:**

Christophe Rodriguez. Dynamique adaptative des virus hautement variables à un nouvel environnement répliatif. Médecine humaine et pathologie. Université Paris-Est, 2012. Français. NNT : 2012PEST0074 . tel-00916824

HAL Id: tel-00916824

<https://theses.hal.science/tel-00916824>

Submitted on 10 Dec 2013

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

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

Dynamique adaptative des virus hautement variables à un nouvel environnement réplcatif

Adaptive dynamics of highly variable viruses to new replicative environment

Présentée et soutenue publiquement le 23 octobre 2012

Membres du jury :

Rapporteur : Pr Vincent CALVEZ, Université Pierre et Marie Curie, PARIS

Rapporteur : Pr Fabien ZOULIM, Université Lyon 1, LYON

Examineur : Pr Christophe HEZODE, Université Paris Est-Créteil, CRETEIL

Examineur : Pr Esteban DOMINGO, Université autonome de Madrid, MADRID

Examineur : Pr Françoise BRUN-VEZINET, Université Paris-Diderot, PARIS

Directeur de thèse : Pr Jean-Michel PAWLOTSKY, Université Paris Est-Créteil, CRETEIL



REMERCIEMENTS

Je tiens en premier lieu à remercier ma femme, Emelyne, pour son soutien sans faille et sans qui rien de tout cela n'aurait été possible et aussi mes enfants pour leur soutien intéressé car particulièrement sensibles aux infections virales...

Je tiens aussi à remercier mon mentor, Jean-Michel, qui m'a soutenu et poussé dans la direction que j'ai suivi jusqu'ici depuis mon internat et surtout pour la qualité scientifique de son encadrement. Je tiens également à m'excuser auprès de lui pour la longueur de la thèse qui a failli avoir raison de sa santé...

Enfin, je souhaite également remercier Coralie Pallier qui m'a montré à quel point les anciennes techniques de séquençages étaient rébarbatives et grâce à qui je me suis tourné vers d'autres alternatives qui m'ont mené jusqu'ici.

SOUTIENS DE RECHERCHE

Merci à l'Agence nationale de recherche contre le sida et les hépatites virales pour le financement de ma bourse et d'une partie des études réalisées durant cette thèse :



Merci également aux autres sponsors provenant d'agences publiques :



Et enfin les sponsors provenant des industriels :



SOMMAIRE

INTRODUCTION.....	6
PARTIE 1 : du concept d'évolution au concept de résistance virale.....	6
La sélection naturelle, de Charles Darwin à Manfred Eigen	6
Du concept de quasi-espèce au développement de la résistance.....	11
La résistance : un concept multifactoriel complexe	13
PARTIE 2 : Thérapeutique antivirale et résistance, aujourd'hui et demain	18
Thérapeutique et résistance.....	18
Le traitement du VIH	18
Le traitement du VHC	20
Le traitement du VHB	22
Résistance aux traitements antiviraux.....	23
Inhibiteurs de polymérase virale (nucléos(t)idiques et non nucléosidiques)	23
Inhibiteurs de protéase (IPs)	26
Inhibiteurs de cyclophilines.....	27
Autres DAAs.....	28
PARTIE 3 : Outils d'étude de la résistance : de la charge virale à l'UDPS	31
Les outils d'étude de la résistance	31
Tests biochimiques enzymatiques	31
Tests phénotypiques cellulaires	32
Tests phénotypiques du VIH.....	32
Tests phénotypiques du VHC	33
Tests phénotypiques du VHB	34
Tests génotypiques.....	35
Techniques de génotypage conventionnelles.....	36
Techniques de séquençage à haut-débit.....	38
Analyse des données	42
PARTIE 4 : Prise en charge clinique de la résistance	45
Prise en charge initiale et prévention du risque	45
Pré-existence des variants viraux résistants	45
Dynamique des variants viraux résistants au cours du traitement antiviral.....	48

OBJECTIFS	50
PATIENTS	52
Etude HBV POL	52
Etude HBV InVS	53
Etude HBV RNase H	53
Etude HBV PreC	54
Etude HCV Prove2.....	55
Etude HIV génotropisme	56
MATERIELS ET METHODES.....	57
Préparation des échantillons pour les techniques de biologie moléculaires d'étude des variants viraux.....	57
Etude HBV POL et HBV InVS	57
Etude HBV RNaseH	58
Etude HBV PreC	58
Etude HCV Prove2	59
Etude HIV génotropisme	60
Pyroséquençage à haut débit (technologie 454).....	60
Etude HBV POL	60
Etude HBV InVS.....	61
Etude HBV RNase H	62
Etude HBV PreC	62
Etude HCV prove2	62
Etude HIV génotropisme	63
Contrôles	63
Autres techniques de biologie moléculaire	65
Etude HBV POL	65
Etude HBV PreC	66
Analyses des données.....	66
Séquençage Sanger	66
Séquençage à haut-débit.....	66
Analyses statistiques	67
RESULTATS.....	68
Création du package PyroPack® pour l'analyse des données de pyroséquençage à haut débit (ARTICLES 2 ET 5)	68

Etude HBV POL (ARTICLES 1 ET 2)	72
Profils cinétiques de charge virale.....	72
Cinétiques virales sous adefovir chez les patients de l'étude HBV Pol (ARTICLE 1)	74
Etude de la dynamique des variants sauvages et résistants des patients de l'étude HBV Pol par clonage-séquençage (ARTICLE 1)	76
Etude de la dynamique des variants sauvages et résistants des patients de l'étude HBV Pol par pyroséquençage à haut débit (ARTICLE 2)	78
Etude HBV InVS (ANALYSE PRELIMINAIRE-MANUSCRIT EN PREPARATION par Chevaliez S, Rodriguez C, et coll.)	85
Etude HBV RNase H (ARTICLE 3)	86
Etude HBV PreC (ANALYSE PRELIMINAIRE- 2 MANUSCRITS EN PREPARATION par Rodriguez C, et coll.).....	89
Comparaison des techniques INNO-LIPA et de pyroséquençage à haut débit pour l'analyse des mutations de la région PreC/C	90
Prédiction de la réponse à l'adefovir par les mutations PreC/C avant le début du traitement....	91
Dynamique des populations virales portant des mutations PreC/C sous traitement antiviral	92
Etude HCV Prove 2 (ANALYSE PRELIMINAIRE-MANUSCRIT EN PREPARATION par Chevaliez S, Rodriguez C (participation égale), et coll.)	93
Mutations de résistance aux inhibiteurs de protéases préexistantes	94
Dynamique des populations virales résistantes au cours du traitement.....	94
Etude HIV géotropisme	96
Comparaison des méthodes d'étude du géotropisme.....	97
Analyse de la variabilité génétique de la boucle V3.....	98
<i>DISCUSSION</i>	<i>99</i>
Prédiction de la résistance et de l'échec thérapeutique avant le début du traitement	100
Etude dynamique de la résistance et de l'échec thérapeutique au cours du traitement	105
Quasi espèce de départ ou étude dynamique au cours du traitement ?	108
Optimisation des outils d'étude et de lutte contre la résistance	109
<i>CONCLUSION-PERSPECTIVES</i>	<i>112</i>
<i>BIBLIOGRAPHIE</i>	<i>114</i>
<i>ANNEXE 1 : MUTATIONS de RESISTANCE CONNUES pour le VHB, VHC, VIH..</i>	<i>125</i>
<i>ANNEXE 2 : BREVETS</i>	<i>128</i>
<i>ANNEXE 3 : articles relatifs à la thèse.....</i>	<i>129</i>

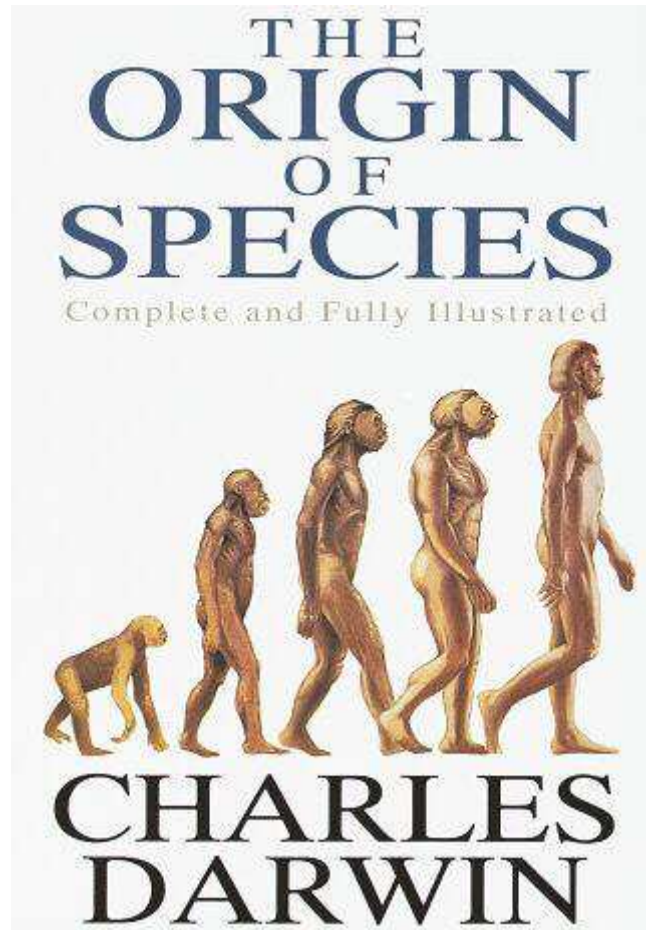


Illustration 1 : Première de couverture d'une réédition de l'œuvre de Darwin sur la théorie de l'évolution des espèces (Darwin 1859).

INTRODUCTION

PARTIE 1 : du concept d'évolution au concept de résistance virale

La sélection naturelle, de Charles Darwin à Manfred Eigen

Il y a un peu plus de 150 ans naissait le principe fondamental de la théorie de l'évolution (Darwin 1859), qui allait bouleverser notre compréhension de la vie et poser les fondements de la biologie moderne (illustration 1). Ce dogme scientifique, d'abord limité à l'explication de la biologie du règne du vivant, est devenu très tôt philosophique car il redéfinissait les règles qui régissent la vie sur Terre, domaine réservé jusqu'alors à la religion. Il suscita des affrontements verbaux restés célèbres, comme le débat qui eut lieu à Oxford en 1860 lors d'une réunion de l'Association britannique pour l'Avancement des Sciences au cours de laquelle l'évêque d'Oxford, Samuel Wilberforce, s'en prit à Darwin à travers Thomas Huxley en lui demandant s'il descendait du singe par son grand-père ou par sa grand-mère, ce à quoi Huxley rétorqua qu'il « préférerait descendre d'un singe plutôt que d'un homme instruit qui utilise sa culture et son éloquence au service du préjugé et du mensonge ». A ce jour, peu de théories ont été l'objet d'autant de tentatives de contradiction. La théorie de Darwin est pourtant toujours présente à l'ère de la génétique moléculaire, de plus en plus d'arguments s'accumulant au fil des ans pour confirmer la justesse de l'intuition initiale, et s'étend même à de plus vastes domaines, comme par exemple l'adaptation des virus à leur environnement.

Charles Darwin, lors de ses études à Cambridge, a été un lecteur assidu et un adepte des théories théologiques de William Paley, qui proposait dans sa « *Théologie naturelle* » (Paley 1802) une explication du phénomène d'adaptation des êtres vivants à leur

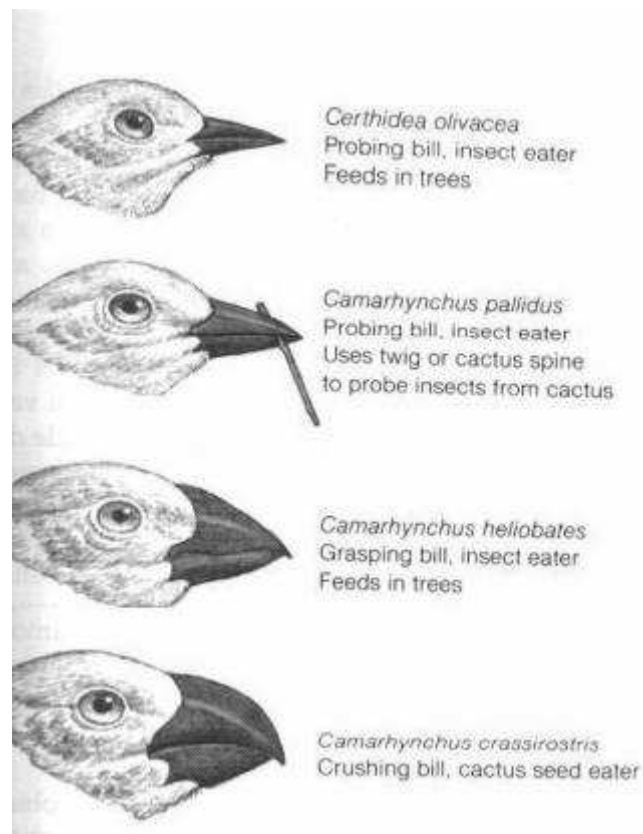


Illustration 2 : Exemple d'adaptation étudiée par Darwin lors de son voyage sur le Beagle. Ici, les oiseaux de la même espèce se sont adaptés en fonction des ressources disponibles dans leur milieu environnemental (Starr and Taggart 2006).

environnement comme un dessein de Dieu : « Si vous trouvez une montre, vous ne doutez pas qu'elle a été fabriquée par un horloger. De même, si vous considérez un organisme un peu complexe, avec l'évidente finalité de tous ses organes, comment ne pas conclure qu'il a été produit par la volonté d'un Créateur? Car il serait simplement absurde de supposer que l'œil d'un mammifère, par exemple, avec la précision de son optique et sa géométrie, aurait pu se former par pur hasard » [traduction de (Jacob 1981)]. Cependant, les observations que fit Darwin lors de son voyage sur le Beagle n'étaient pas compatibles avec cette doctrine, elles tendaient même plutôt à suggérer une autre explication, de laquelle la main de Dieu était absente. Darwin commença à se détourner de la thèse créationniste et chercha à expliquer le phénomène d'adaptation des êtres vivants à leurs conditions d'existence ; ce sera pour lui le mécanisme de la « sélection naturelle ». Pour Darwin, il n'y a donc pas de puissance surnaturelle qui sélectionnerait les individus afin d'améliorer les espèces ; la sélection doit être le produit d'un ressort non intentionnel, d'un mécanisme non dirigé, émanant d'un ensemble de conditions spontanées et nécessaires, qui aboutissent automatiquement à l'adaptation de l'être vivant à son milieu (Illustration 2). Ses arguments scientifiques s'inspirent notamment de la lecture de Thomas Robert Malthus, « *Essai sur le principe de population* » (Malthus 1798), dans lequel est exposée l'idée selon laquelle la croissance des populations est nécessairement limitée par la disponibilité des ressources naturelles. Ainsi, en période de pénurie naît une lutte pour la survie entre les individus de la même espèce et entre les espèces qui se disputent les ressources rares ; seuls alors survivent et parviennent à se reproduire les êtres les plus adaptés à ces circonstances. C'est le philosophe Herbert Spencer, fervent défenseur de Darwin, dans son « *Principles of biology* » (Spencer 1864) qui résumera le mieux le principe de l'évolution par sa phrase « *survival of the fittest* » .

Le Darwinisme ne s'imposa pourtant pas immédiatement, notamment en France où sera soutenue pendant de longues années la thèse du naturaliste français Jean-Baptiste de Lamarck, souvent présentée comme antinomique. Ce dernier, dans « *Philosophie géologique* » (Lamarck 1830), ne remet cependant pas en cause le principe d'évolution mais plutôt le mécanisme de sélection. Il explique que si la tendance à la complexification avait été seule à jouer, la progression de la composition des animaux eût été régulière ; or, le degré de diversification biologique évolue au gré des circonstances rencontrées par l'être vivant ; ce sont donc les circonstances qui sont responsables de l'évolution des espèces. Il plaide ainsi pour une théorie héréditaire acquise. C'est le biologiste Auguste Weismann dans « *The Germ Plasm: A theory of heredity* » (Weismann *et al.* 1893) qui apportera un premier élément déterminant pour rompre avec le Lamarckisme, avec son expérience célèbre dans laquelle il a entrepris de couper la queue de souris sur de nombreuses générations pour montrer qu'aucune adaptation n'était observée. Il y défend également la théorie du plasma germinatif, selon laquelle les organismes pluricellulaires sont constitués de cellules germinales, contenant l'information héréditaire, et de cellules somatiques, effectuant les fonctions vitales. Les cellules germinales ne sont influencées ni par ce que le corps apprend, ni par aucune capacité que celui-ci acquiert au long de sa vie, et ne peuvent donc pas transmettre ces capacités à la génération suivante. Ces avancées sont à l'origine de la redécouverte de l'œuvre de Gregor Mendel : « *Recherche sur des hybrides végétaux* », qui mettent en évidence le caractère héréditaire de la transmission des gènes (Mendel 1865) et signent l'entrée de la biologie dans l'ère de la génétique.

Le début du XXème siècle fut d'abord marqué par la confrontation de deux théories, le saltationnisme et le gradualisme. Pour les mutationnistes saltationnistes, les mutations d'effet important sont le moteur de l'évolution, l'évolution d'une génération à l'autre se

faisant par sauts, et non graduellement. La sélection naturelle jouerait donc un rôle mineur dans l'évolution et les espèces apparaîtraient à la suite de macro-mutations. Thomas Morgan, dans « *Le mécanisme de l'hérédité mendélienne* » (Morgan 1915), montra cependant que les traits à hérédité mendélienne pouvaient parfaitement correspondre à de très petites variations (gradualisme). Autrement dit, la nature discrète de l'hérédité est parfaitement compatible avec la nature continue de la variation et donc le darwinisme. La synthèse néo-darwinienne pouvait alors survenir.

Ronald Aylmer Fisher, plus connu aujourd'hui pour les tests statistiques qui portent son nom, a été un des principaux artisans de cette synthèse en formalisant mathématiquement pour la première fois la théorie de l'évolution dans son théorème fondamental de la sélection naturelle publié dans « *The genetical theory of natural selection* » (Fisher 1930). Il sera d'ailleurs un des auteurs du livre de Julian Huxley « *Evolution, the modern synthesis* » (Huxley 1942) qui expose les principes de base de la génétique des populations :

- La base de la variation continue est discontinue. Le phénotype est en effet la résultante de l'effet des gènes, de celui du milieu, et de leur interaction. Même pour des caractères à déterminisme génétique simple, l'effet du milieu est tel que la variation phénotypique est continue. En outre, la plupart des caractères sont déterminés par plusieurs loci d'effet individuel faible, d'où un nombre de combinaisons potentiellement infini, ce qui, même en l'absence d'effet du milieu sur l'expression phénotypique, conduit à une variation phénotypique de nature continue, gaussienne.

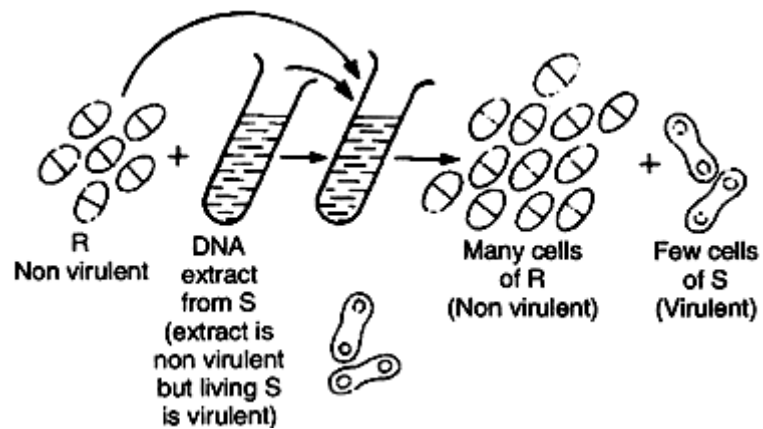


Figure 1: Expérience réalisée par Avery pour démontrer le rôle de l'ADN dans la transmission de l'information héréditaire : par transformation bactérienne de bactéries non virulentes avec de l'ADN de bactéries virulentes, il obtient des bactéries virulentes (Avery *et al.* 1944).

- La reproduction sexuée permet de faire apparaître de nouveaux génotypes sans qu'il soit nécessaire qu'une mutation se produise.
- La sélection naturelle peut agir même sur des petites variations à travers l'augmentation de fréquences des allèles sélectionnés, qui se traduit par une évolution continue.
- Le hasard ne facilitant pas l'action de la sélection naturelle : une mutation unique a peu de chances d'augmenter en fréquence, même si elle est avantageuse.
- Les macro-mutations, sans doute plus rares que les micro-mutations, contribuent peu à l'évolution, en particulier à la diversification.

Si les bases du concept de la sélection naturelle étaient établies, le support de l'information génétique restait inconnu. L'existence de l'Acide désoxyribonucléique (ADN) était connue depuis Friedrich Miescher en 1869 (Dahm 2008), mais c'est Oswald Avery dans son article de 1944 qui, le premier, démontra le rôle de l'ADN dans la transmission de l'information héréditaire (Avery *et al.* 1944) (Figure 1). Peu de temps avant, Wendell Meredith Stanley cristallisait pour la première fois le virus de la mosaïque du tabac (Stanley 1935), permettant d'établir que les virus sont constitués en partie d'acide Ribonucléique (ARN) ou d'ADN et ouvrant la voie à l'extension de la théorie de l'évolution au monde viral. Il fallut attendre quelques années de plus et les travaux du biochimiste Manfred Eigen pour que les principes de la sélection naturelle, essentiellement pensés pour les êtres vivants à reproduction sexuée, soit étendus au monde des virus. C'est ainsi que le concept de « quasi-espèce » apparut pour la première fois dans l'article « *Self-organization of matter and the evolution of biological macromolecules* » (Eigen 1971), principe qui vaudra à son créateur le Prix Nobel en 1967 et sera repris et développé avec Peter Schuster dans « *The Hypercycle, a*

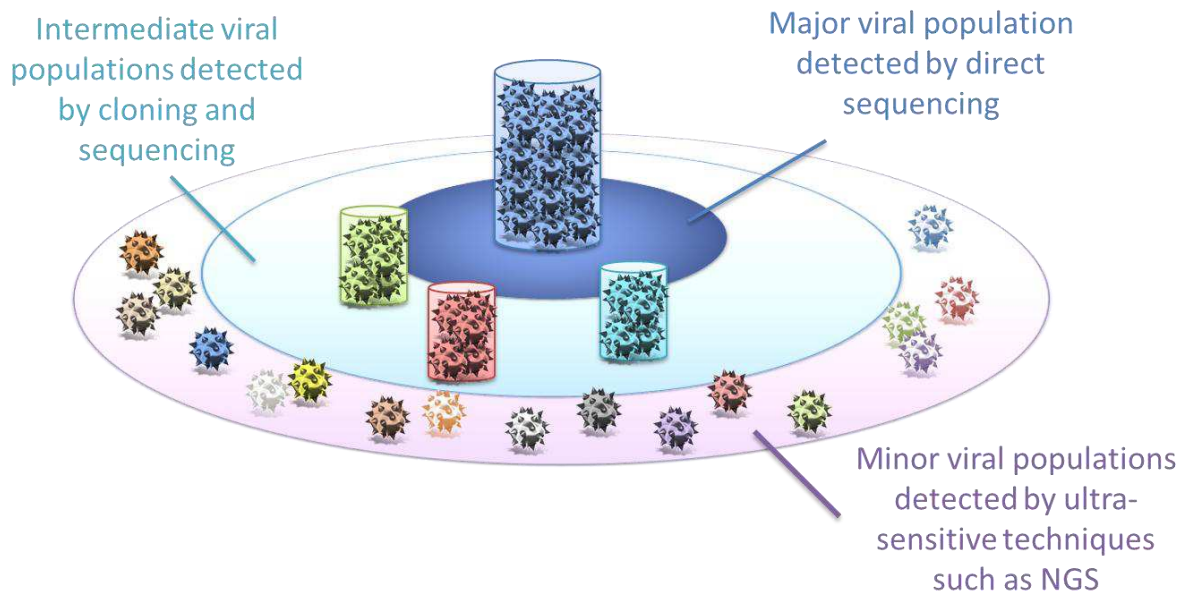


Figure 2 : Organisation en quasi espèce des virus hautement variables. Chaque variant possède un potentiel génétique différent associé à une capacité répliquative différente et une distribution différente au sein de la quasi-espèce (Chevaliez et al. 2012).

principle of natural self-organization » (Eigen and Schuster 1979) pour prendre la forme qu'on lui connaît aujourd'hui.

Du concept de quasi-espèce au développement de la résistance

Eigen a décrit en 1971 les bases fondamentales, encore valides aujourd'hui, permettant de comprendre les mécanismes d'émergence de variants résistants chez les virus hautement variables comme le virus de l'immunodéficience humaine (VIH), le virus de l'hépatite B (VHB) ou le virus de l'hépatite C (VHC). Chez ces virus, l'absence de mécanismes correctifs lors des cycles de réplication entraîne à chaque cycle viral la synthèse d'un nombre de mutations sans commune mesure avec le reste du règne du vivant : $1,4 \times 10^{-4}$ mutations par nucléotide et par cycle pour le VIH, $1,15 \times 10^{-4}$ pour le VHB, $6,3 \times 10^{-4}$ pour le VHC (Ji and Loeb 1992; Park *et al.* 2003; Cuevas *et al.* 2009), contre seulement 10^{-9} par cycle cellulaire pour l'homme (Lynch 2010). Ainsi, ces virus ne sont pas caractérisés par une entité génétique définie, mais au contraire par une multitude de sous-populations virales apparentées et génétiquement proches qui forment collectivement l'entité. Les potentiels génétiques différents des sous-populations virales composant la quasi-espèce s'expriment par des capacités fonctionnelles différentes, aboutissant à des capacités répliquatives différentes. Ainsi, dans un environnement répliquatif donné, une sous-population peut posséder un potentiel génétique plus favorable, qui lui permet d'occuper une place plus importante au sein de la quasi-espèce ; elle sera donc la population majoritaire. D'autres populations, moins adaptées, se partagent des parts de moins en moins importantes de la niche écologique à mesure que leur potentiel répliquatif diminue dans leur environnement répliquatif, donnant naissance à des sous-populations intermédiaires et minoritaires (Figure 2)

(Chevaliez *et al.* 2012). Les sous-populations virales non capables de répliquer dans cet environnement sont éliminées.

La reconnaissance de la distribution en quasi-espèce des virus est la clé de la compréhension de leur évolution dans un milieu environnemental donné (Pawlotsky 2005). Ce phénomène, dynamique, s'inscrit dans l'espace-temps. Il s'agit d'une adaptation constante aux pressions sélectives exercées par le milieu extérieur, de telle sorte que la quasi-espèce étudiée à un instant donné est le résultat des pressions exercées dans le passé et de la lutte du virus pour sa survie sous ces pressions consécutives. Prédire l'évolution dynamique d'une quasi-espèce virale revient donc à « parier » sur le poids de chacune des pressions exercées dans le présent. Modélisée mathématiquement, la quasi-espèce est donc un « objet » en mouvement dans l'espace des possibilités génétiques du virus, que l'on peut caractériser par sa position initiale et un « vecteur vitesse » qui lui permet de se diriger dans une direction précise. L'étude de la position initiale se fait par des études statiques et permet de caractériser l'influence du potentiel de départ de la quasi-espèce sur sa capacité à se diriger. Au contraire, les études dynamiques permettent de définir le vecteur vitesse, qui correspond à la somme des vecteurs représentant chacun une pression, caractérisés eux-mêmes par une vitesse et une direction propres (Rodriguez *et al.*, 2012). Pour aller plus avant dans la description de cet objet mobile, il convient de définir précisément le contexte d'étude, car l'intérêt que l'on porte à la description précise du mouvement est essentiellement lié à la destination que le virus est susceptible d'atteindre et que l'on veut à tout prix éviter, en l'occurrence l'état de résistance.

La résistance : un concept multifactoriel complexe

Si l'on reprend la notion « d'objet mobile » utilisée précédemment pour décrire la quasi espèce virale, défini par des coordonnées de départ et un vecteur vitesse, il est possible de déterminer la position de l'objet à un temps précis, comme par exemple l'issue d'un traitement antiviral, mais ceci nécessite de connaître parfaitement l'ensemble des paramètres de l'équation.

L'espace-temps de la quasi espèce n'étant pas fini et n'ayant pas d'origine, la position de départ ne pourra être décrite qu'en terme de distance avec la position d'arrivée que l'on souhaite étudier, c'est-à-dire la situation de résistance. Il est aisé d'imaginer que l'objectif de résistance sera atteint d'autant plus rapidement que certains événements primaires sont déjà survenus, car le virus se trouvera à une distance originelle plus courte de la destination finale. Ainsi, la présence de variants viraux majoritaires résistants à un traitement antiviral (comme dans le cas de la transmission d'hôte à hôte de virus résistants) est rapidement associée à l'observation d'une résistance clinique chez les patients infectés par le VIH traités par trithérapie (Little *et al.* 2002). De la même façon, la présence de variants minoritaires chez des patients naïfs séropositifs pour le VIH augmente le risque de survenue d'une résistance à la trithérapie antivirale, et ce d'autant que leur prévalence dans la quasi espèce est plus élevée (Johnson *et al.* 2008; Simen *et al.* 2009). Un virus plus proche de l'espace de résistance trouvera donc plus fréquemment et plus rapidement la voie vers celui-ci. Le concept peut très vraisemblablement être étendu au VHB et au VHC, dans la mesure où les conditions de réalisation de ces événements sont également réunies, bien que ce point reste à démontrer pour ces virus.

	VIH	VHC	VHB
Production virale (virion/jour)	10^{10}	10^{12}	$10^{12/13}$
Demi-vie d'un virion (h)	1	2-3	3-34
Taux de mutation/nucléotide/cycle	$1,4 \cdot 10^{-4}$	$6,3 \cdot 10^{-4}$	$1,15 \cdot 10^{-4}$
<i>Nombre de changements nucléotidiques % de tous les mutants possibles créés par jour</i>			
1	100%	100%	100%
2	2%	100%	66%
3	$<10^{-5}$	0.0034%	$<10^{-5}$

Tableau 1 : Caractéristiques de réplication des virus hautement variables VIH, VHC et VHB

d'après (Murray et al. 2006; Soriano et al. 2008; Rong et al. 2010)

Du point de vue de la dynamique virale, comme on l'a dit, l'évolution d'une quasi espèce virale vers l'état de résistance est en partie déterminée par un vecteur vitesse résultant de l'addition des différentes pressions exercées. Cependant, il est important de noter qu'en l'absence de toute pression, le virus possède une capacité intrinsèque d'évolution aléatoire qui peut influencer le vecteur vitesse final. Cette propriété est liée à trois caractéristiques virales importantes (Tableau 1):

- la production quotidienne de virion
- la demi-vie du virus dans l'organe de production et dans la circulation générale
- le taux de mutations spontanées du génome viral

Pour le VHC, il a été montré que l'ensemble des génomes possibles possédant une ou deux mutations par rapport à la séquence consensus était produit chaque jour, contre seulement 2% des génomes doublement mutés possibles pour le VIH (Murray *et al.* 2006; Soriano *et al.* 2008; Rong *et al.* 2010). Il est aisé d'imaginer que la rapidité avec laquelle la quasi espèce virale tend naturellement vers l'état de résistance est influencée par le nombre de mutations que le virus est capable de produire par unité de temps. Néanmoins, les mutations acquises qui font progresser le virus vers son objectif de résistance sont souvent accompagnées d'un déficit de capacité répliquative par rapport au virus sauvage. Ainsi, la charge virale ou nombre de virions produit chaque jour diminue, altérant de ce fait la vitesse de progression du virus vers l'espace de résistance à mesure qu'il se rapproche de son objectif. Poussé à l'extrême, ce processus pourrait prendre un temps infini, incompatible avec la survie du virus, conduisant au « crash » viral et à la guérison du patient. Pour éviter cet écueil, la quasi espèce virale progresse vers l'espace de résistance grâce à la sélection d'une combinaison de mutations de résistance et de mutations qui améliorent la répllication des virus résistants

(mutations dites de « fitness ») afin d'atteindre son objectif dans un délai compatible avec la survie du virus. Il apparaît ainsi que la charge virale et le temps nécessaire pour atteindre l'état de résistance sont intrinsèquement liés et que deux leviers sont possibles pour infléchir la trajectoire du virus vers la résistance : la diminution de la rapidité d'évolution du virus par la réduction de la charge virale et/ou l'éloignement de l'espace de résistance.

Les facteurs qui influencent la charge virale sont de deux ordres, les capacités intrinsèques de lutte contre le virus possédées par le patient et des facteurs extrinsèques sur lesquels il est possible d'intervenir. Les premières sont liées essentiellement au fond génétique du patient. En effet, si tous les êtres humains sont dotés des mêmes outils de défense antivirale, tous ne sont pas capables de les mettre en œuvre avec la même efficacité (Pawlotsky 2009). Il a par exemple été montré que des polymorphismes situés à proximité du gène de l'IL28B (ou interféron lambda 3) étaient associés à un taux de guérison spontanée significativement plus élevé chez les patients en phase aigüe d'infection par le VHC (Thomas *et al.* 2009) et à un plus fort taux de guérison sous traitement par interféron pégylé chez les patients ayant une infection chronique. D'autres exemples existent dans le domaine du VIH, comme les patients HLA B57 dits « contrôleurs » qui sont capables de maintenir un taux de répllication virale très bas en l'absence de traitement antiviral (Kosmrlj *et al.* 2010) ou les patients présentant un effet de barrière à l'entrée virale retardant ou empêchant la progression de l'infection car dépourvus du corécepteur lymphocytaire CCR5 (Goh *et al.* 1999). Le métabolisme peut également jouer un rôle important sur l'émergence de la résistance. En effet, dans le contexte de la thérapeutique antivirale, il conditionne la biodisponibilité des médicaments. Selon que le médicament doit être métabolisé pour être actif et/ou être éliminé, il est possible d'observer un défaut d'activité antivirale ou une toxicité qui pourra avoir des répercussions en termes d'observance.

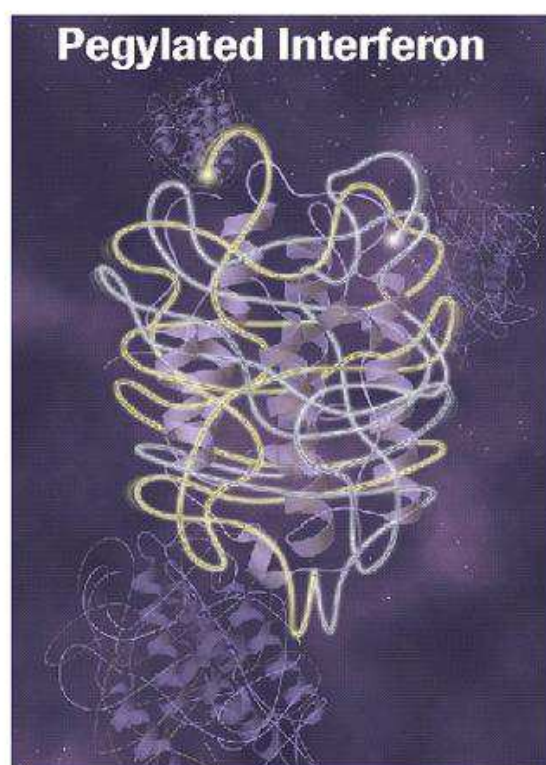


Illustration 3 : La pégylation de la molécule d'interféron alpha permet sa protection contre les processus protéolytiques et augmente sa biodisponibilité.

Si l'on considère les aspects interventionnels, la thérapeutique antivirale tient évidemment une place prépondérante et fera l'objet d'un chapitre spécifique. Pour tous les médicaments, la puissance antivirale induit la diminution de la charge virale et influence donc la vitesse de progression de la quasi espèce vers la résistance. Cependant, l'effet est maximal si les conditions de traitement et la biodisponibilité du médicament sont optimales, en particulier grâce à une stricte observance du traitement. L'adhésion du patient est un facteur fondamental de la réussite ou de l'échec du traitement. Par exemple, la fréquence de la prise des médicaments antiviraux peut jouer un rôle important. Si les efforts de développement thérapeutique ont permis d'optimiser les thérapeutiques antivirales du VHB et du VIH, avec souvent une seule prise de médicament par jour, il subsiste des difficultés pour les thérapeutiques anti-VHC et les thérapeutiques anti-VIH de 2ème ligne. Par exemple, les nouveaux médicaments inhibiteurs de protéase du VHC nécessitent trois prises par jour à heure fixe associées à une prise d'aliments gras pour en augmenter la biodisponibilité. Les contraintes organisationnelles sur le plan professionnel et privé peuvent donc compromettre la poursuite des traitements antiviraux à long terme et favoriser la survenue de la résistance par la moindre pression antivirale exercée. Des recherches visant à l'amélioration de la biodisponibilité des médicaments sont poursuivies selon plusieurs axes :

- « protection » du médicament par greffe de molécules diminuant son accessibilité aux processus d'élimination, par exemple la pégylation de l'interféron qui permet aujourd'hui une injection hebdomadaire et un taux sérique élevé stable alors que l'interféron non pégylé devait être administré trois fois par semaine, voire quotidiennement (illustration 3).
- inhibition directe des cytochromes responsable de la dégradation ; par exemple le « boost » par le ritonavir utilisé dans le traitement par les inhibiteurs de protéase du

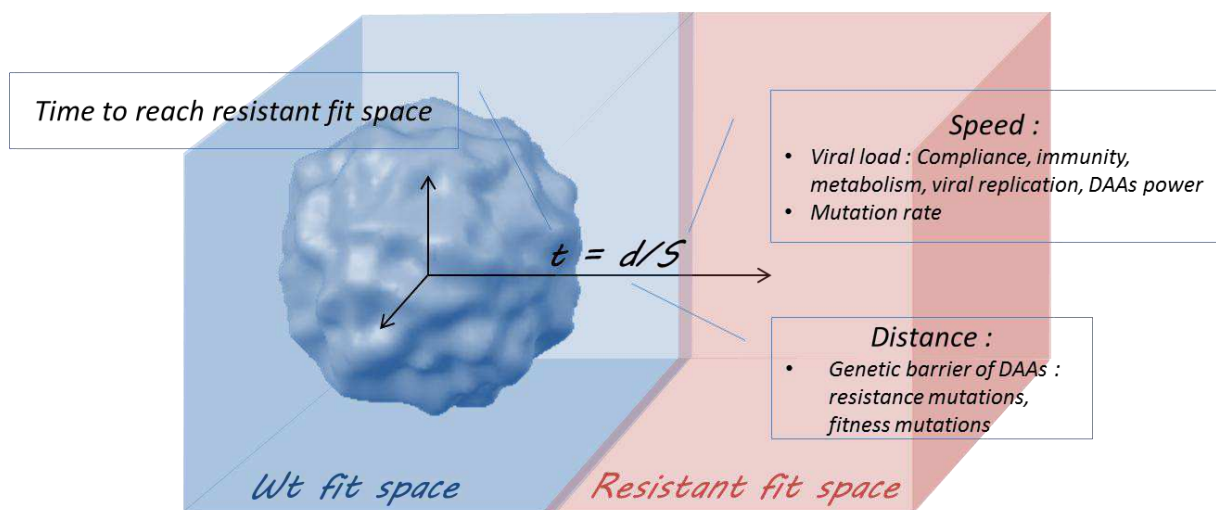


Figure 3 : Paramètres influençant la trajectoire du virus vers la résistance. (Rodriguez *et al.*, 2012). Le temps mis pour atteindre la résistance est proportionnel à la distance entre le virus initial et l'état de résistance et inversement proportionnel à la vitesse, calculée à partir des paramètres de charge virale et des taux de mutation. La charge virale est elle-même fonction de nombreux paramètres comme la compliance du patient au traitement, l'efficacité du système immunitaire, la capacité de réplication du virus et la puissance antivirale du médicament utilisé.

VIH ou du VHC pour augmenter la biodisponibilité des inhibiteurs de protéase de l'un et l'autre virus (Zeldin and Petruschke 2004; Gane *et al.* 2011).

La seconde possibilité d'action pour éviter la résistance est d'éloigner la quasi espèce virale de son objectif. Pour cela, il est possible de choisir des thérapeutiques qui obligent à un nombre d'événements « mutation » plus important, c'est-à-dire des médicaments ou des associations thérapeutiques dotés d'une barrière génétique à la résistance élevée. Cet aspect est souvent développé pour les nouvelles générations de traitement après qu'un chef de file ait démontré son efficacité. C'est le cas par exemple de la lamivudine dans le traitement du VHB ou de la zidovudine dans le traitement du VIH, molécules à faible barrière génétique, qui ne sont plus utilisées aujourd'hui au profit de médicaments possédant une barrière génétique beaucoup plus élevée.

En résumé, la quasi-espèce virale est un objet mobile dans l'univers des possibilités génétiques et le temps nécessaire pour atteindre la résistance est (Figure 3) :

- proportionnel à la distance qui le sépare de cet objectif, autrement dit dépendant de la barrière génétique (qui tient compte des mutations de résistance nécessaires pour conférer la résistance à une classe ou une combinaison données), ou plus généralement de la barrière à la résistance (qui tient compte des mutations de résistance mais aussi de la capacité répliquative des variants viraux résistants).
- inversement proportionnel à la vitesse de sélection des mutations, qui dépend de la charge virale, elle-même dépendante de la compliance, de l'environnement immunitaire, de la réplication virale, de la puissance antivirale du ou des médicaments et du taux de mutation du virus

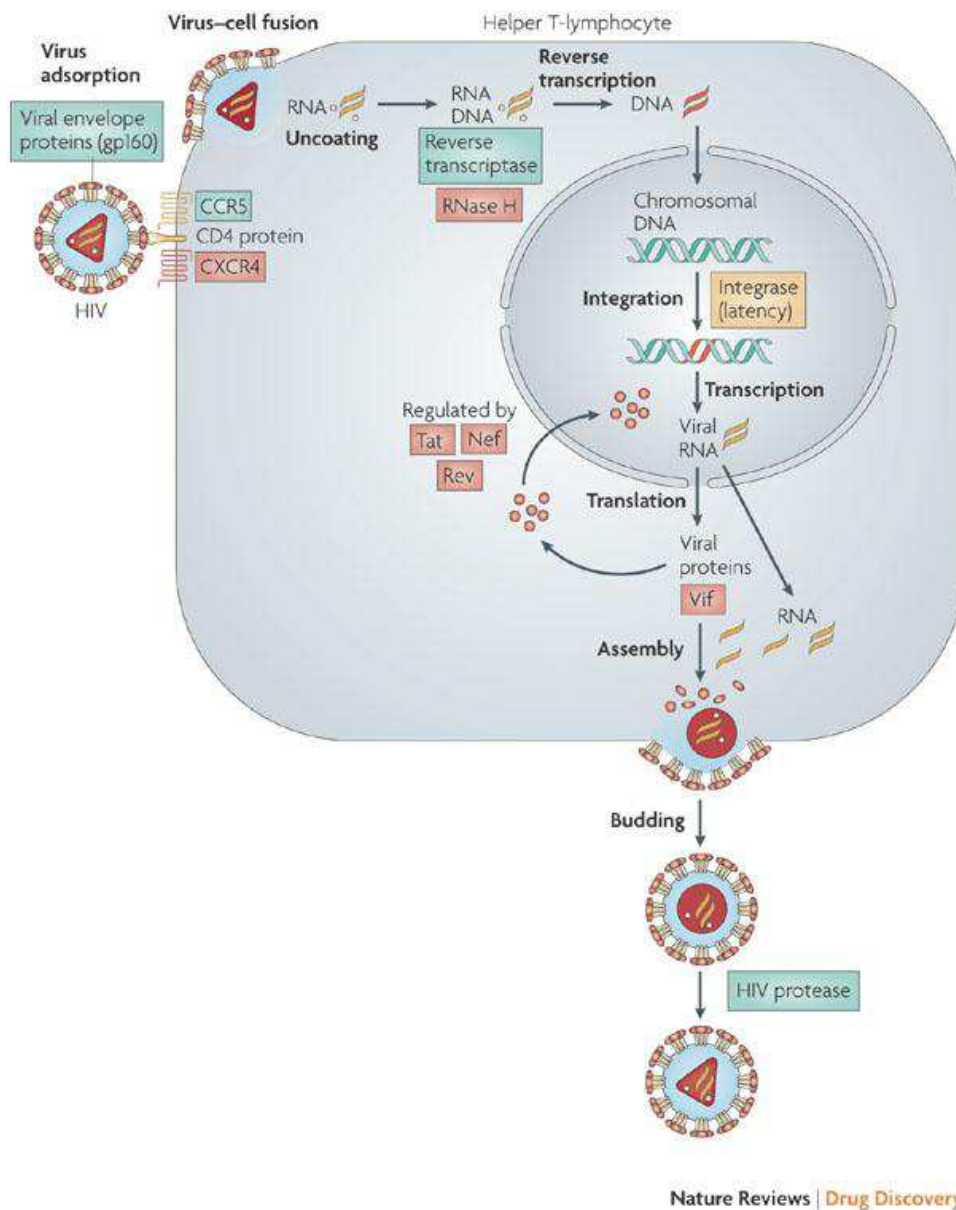


Figure 4 : Cibles antivirales connues dans le cycle cellulaire du VIH. En vert et jaune sont indiquées les cibles thérapeutiques pour lesquelles des médicaments existent sur le marché, en rouge les cibles pour lesquelles des médicaments sont en développement (Flexner 2007).

Thérapeutique et résistance

Les antiviraux jouent un rôle primordial dans la lutte contre l'émergence de la résistance en intervenant à 2 niveaux. Le premier niveau est lié à leur puissance antivirale, qui conditionne la rapidité et l'amplitude avec laquelle la charge virale sera réduite. Le deuxième niveau consiste à éloigner l'espace de résistance en augmentant la barrière du traitement. Cela a été souvent réalisé en modifiant la structure du médicament et en associant des médicaments entre eux afin de créer des combinaisons dont la barrière à la résistance est élevée. La course à la découverte de nouvelles classes thérapeutiques a abouti au développement de nombreuses familles de médicaments pour le VHC, le VHB et le VIH (Sarrazin *et al.* 2012). Dans ce chapitre, nous décrirons les cibles thérapeutiques principales et les mécanismes de résistance associés, en préambule à la description des aspects pratiques de la prise en charge clinique de la résistance (partie 4) (Pawlotsky 2012).

Le traitement du VIH

Le cycle viral du VIH (Figure 4) (Flexner 2007) débute par l'entrée du virus dans le lymphocyte T CD4+. Cette étape, qui constitue une cible thérapeutique importante, nécessite la fixation de la protéine d'enveloppe virale gp120 au récepteur CD4 et aux co-récepteurs CCR5 ou CXCR4 du lymphocyte. Cette fixation entraîne l'ouverture de la gp120 et le déploiement de la protéine gp41. Cette dernière possède 3 domaines importants : HR1, HR2 et le peptide de fusion (PF). Lorsque la gp41 se déploie, le peptide de fusion vient s'insérer dans la membrane cellulaire, puis un repliement de la protéine résultant du rapprochement des domaines HR1 et HR2 permet le rapprochement des membranes

cellulaires et virales et, finalement, la fusion des membranes. Cette étape du cycle viral est la cible de deux médicaments ayant reçu une autorisation de mise sur le marché :

- le maraviroc, inhibiteur du CCR5 (Dorr *et al.* 2005)
- l'enfuvirtide, analogue du domaine HR2 de la gp41 empêchant le repliement de la gp41 et la fusion des membranes (Lalezari *et al.* 2003)

Des inhibiteurs de CXCR4 sont en cours de développement mais semblent associés à une toxicité importante (Choi *et al.* 2012).

Le cycle viral se poursuit par une phase cytoplasmique au cours de laquelle surviennent la décapsidation et la transcription inverse du virus. C'est cette étape de transcription inverse, assurée par la transcriptase inverse, qui a été la cible des premiers médicaments anti-VIH et reste une cible thérapeutique très importante. La transcriptase inverse possède deux activités enzymatiques distinctes, dépendantes l'une de l'autre :

- l'activité de transcription inverse, cible de deux classes médicamenteuses sur le marché : les inhibiteurs nucléosidiques ou nucléotidiques de la transcriptase (INTI) et les inhibiteurs non-nucléosidiques de la TI (INNTI)
- l'activité RNase H, qui dégrade la partie ARN de l'hétéroduplex ADN/ARN à mesure que celui-ci est synthétisé. Cette activité est l'objet d'intenses recherches d'inhibiteurs qui n'ont pour le moment pas abouti en raison d'une toxicité excessive liée à sa proximité de structure avec les RNase H humaines (Tramontano and Di Santo 2010).

L'ADN synthétisé est ensuite transporté vers le noyau puis intégré à l'ADN chromosomique cellulaire grâce à une intégrase ; les inhibiteurs d'intégrase constituent une nouvelle classe thérapeutique prometteuse (Pommier *et al.* 2005). L'ADN intégré est ensuite transcrit en

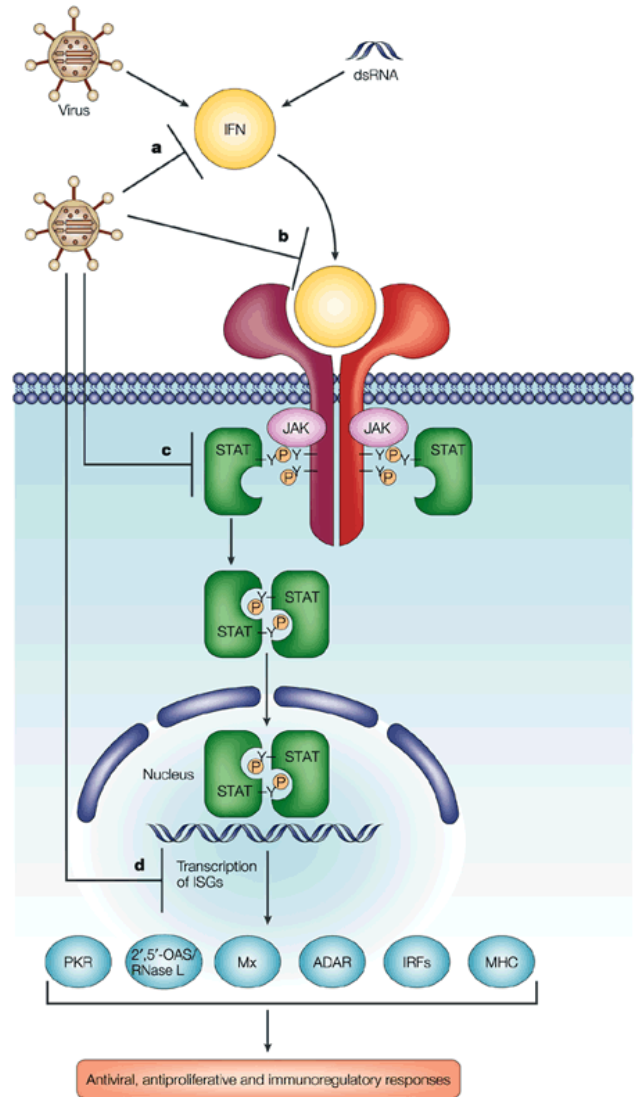
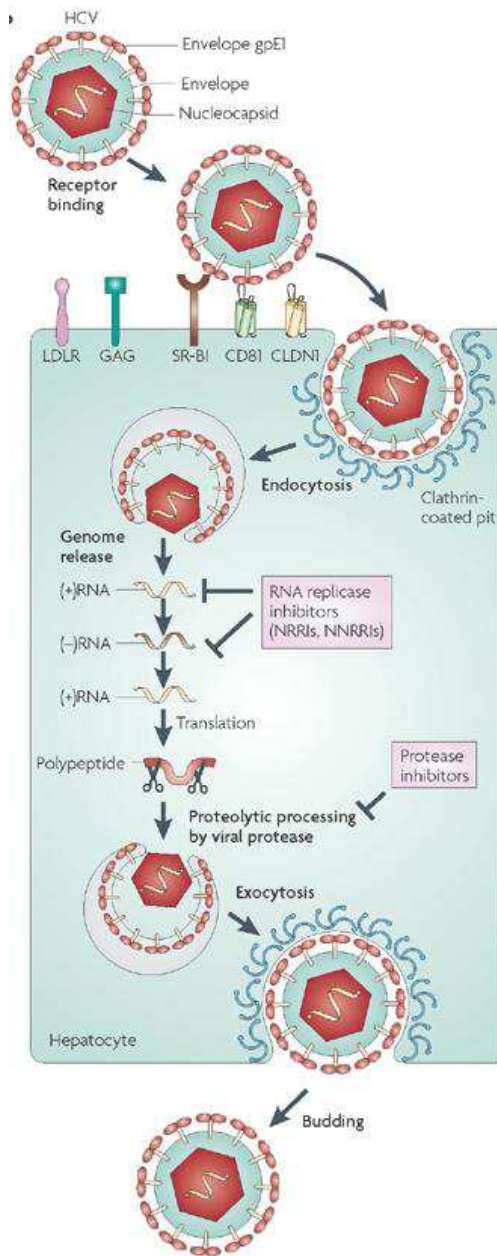


Figure 5 (à gauche) : Cibles antivirales connues sur le cycle cellulaire du VHC. Les médicaments antiviraux ciblant l'hôte ne sont pas représentés (De Clercq 2007).

Figure 6 (à droite): La thérapie par interféron alpha permet de stimuler l'activité des ISGs (interferon-stimulated genes), permettant la synthèse de protéines à activité antivirale non spécifique comme la PKR ou la 2'5'OAS (Katze *et al.* 2002).

ARNs qui sont en partie traduits en protéines structurales et non structurales. Parmi ces dernières, les protéines à activité nucléaire Tat, Nef et Rev qui, respectivement, augmentent la quantité d'ADN viral transcrit, maintiennent un niveau d'activation élevée de la cellule et protègent l'ADN viral du splicing, font l'objet de développements thérapeutiques (Davidson *et al.* 2009; Wick *et al.* 2009; Vercruysse *et al.* 2010). La protéine Vif, qui inhibe la protéine APOBEC3G qui protège le génome viral d'une hypermutation délétère pour le virus, fait également l'objet de recherche d'inhibiteurs (Nathans *et al.* 2008). Les nouveaux génomes synthétisés, sont ensuite encapsidés et enveloppés avant d'être largués dans le milieu extracellulaire. Finalement, les virions subissent une phase de maturation des protéines de capsid assurée par une protéase virale, qui est également la cible de nombreux inhibiteurs utilisés en thérapeutique, appelés inhibiteurs de protéase.

Le traitement du VHC

Le traitement du VHC est en partie fondé sur des approches antivirales directes (DAAs) par ciblage de protéines virales impliquées dans le cycle viral (Figure 5). Le cycle viral du VHC (De Clercq 2007) débute par l'entrée du virus dans l'hépatocyte. Cette étape met en jeu de nombreuses protéines réceptrices, plus particulièrement les récepteurs scavenger receptor-B1, CD81, occludine ou claudine-1. Elle fait l'objet de recherches actives de médicaments inhibiteurs. Cependant, la méconnaissance de la fonction exacte des molécules impliquées dans le complexe récepteur est un frein important au développement de médicaments non toxiques (Lemon *et al.* 2010). Le cycle cellulaire du VHC est exclusivement cytoplasmique. Il comprend une phase de réplication du génome et de synthèse de l'ARN codant catalysée par une ARN polymérase dépendante de l'ARN (protéine NS5B), qui est la cible de nombreux

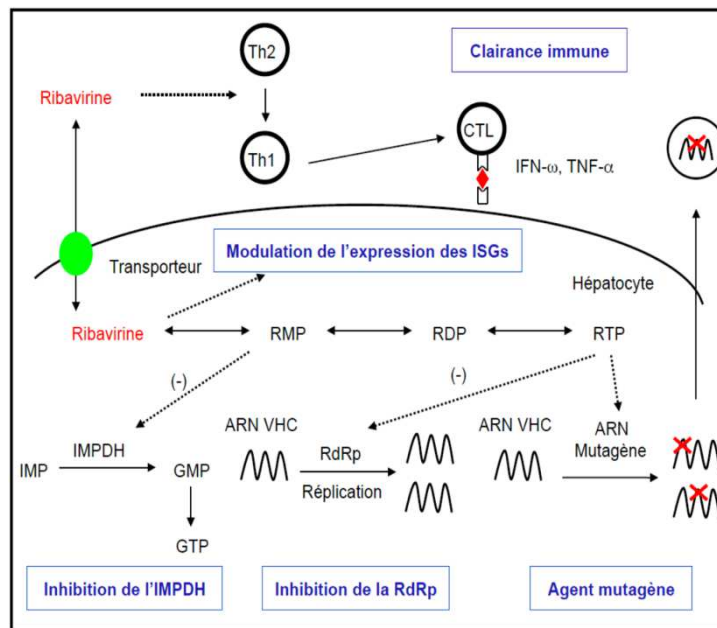


Figure 7 : Mécanismes antiviraux directs et indirects possibles de la ribavirine dans la lutte contre le VHC (Adapté de Lau *et al.* 2002).

médicaments en développement (Soriano *et al.* 2011; Haudecoeur *et al.* 2012). La traduction du cadre de lecture ouvert unique du VHC aboutit à la synthèse d'une polyprotéine qui est ensuite clivée par des protéases cellulaires et virales. La protéase virale NS3/4A, qui assure le clivage en *trans* de l'ensemble des protéines situées en aval, est la cible de nombreux médicaments en développement, dont deux sont récemment arrivés sur le marché (Hofmann and Zeuzem 2011; Pawlotsky 2011; Sarrazin *et al.* 2012). Le cycle se poursuit par l'encapsidation, le bourgeonnement de la particule virale dans la lumière du réticulum endoplasmique, la maturation des protéines virales dans l'appareil de Golgi et l'exocytose du virus.

L'interféron alpha et la ribavirine restent cependant pour encore quelques années à la base du traitement des hépatites virales C, seuls (génotypes non 1) ou associés à un inhibiteur de protéase (patients infectés par un VHC de génotype 1) (Chevaliez and JM 2008). L'efficacité de l'interféron alpha semble essentiellement liée à ses propriétés antivirales indirectes, elles-mêmes liées à l'activation de l'expression de protéines codées par les ISGs (*interferon stimulated genes*), par exemple la protéine kinase R (PKR) ou la 2'-5'-oligoadénylate synthétase (2'5'-OAS), qui ont essentiellement pour fonction de diminuer l'activité traductrice de la cellule en ciblant les facteurs de traduction (Figure 6)(Katze *et al.* 2002). Le mécanisme d'action de la ribavirine sur le VHC reste encore objet de débat. Son activité antivirale directe est faible (Pawlotsky *et al.* 2004) et elle ne semble pas agir par le biais de l'inhibition de l'enzyme cellulaire inosine monophosphate déshydrogénase (IMPDH). Un effet d'accélération de la clairance immune ou la modulation des ISGs ont été évoqués et pourraient expliquer la synergie observée avec l'interféron (Figure 7) (Lau *et al.* 2002).

A côté des DAAs qui ciblent directement les fonctions virales, des inhibiteurs de cyclophiline sont également en cours de développement dans le traitement de l'hépatite chronique C.

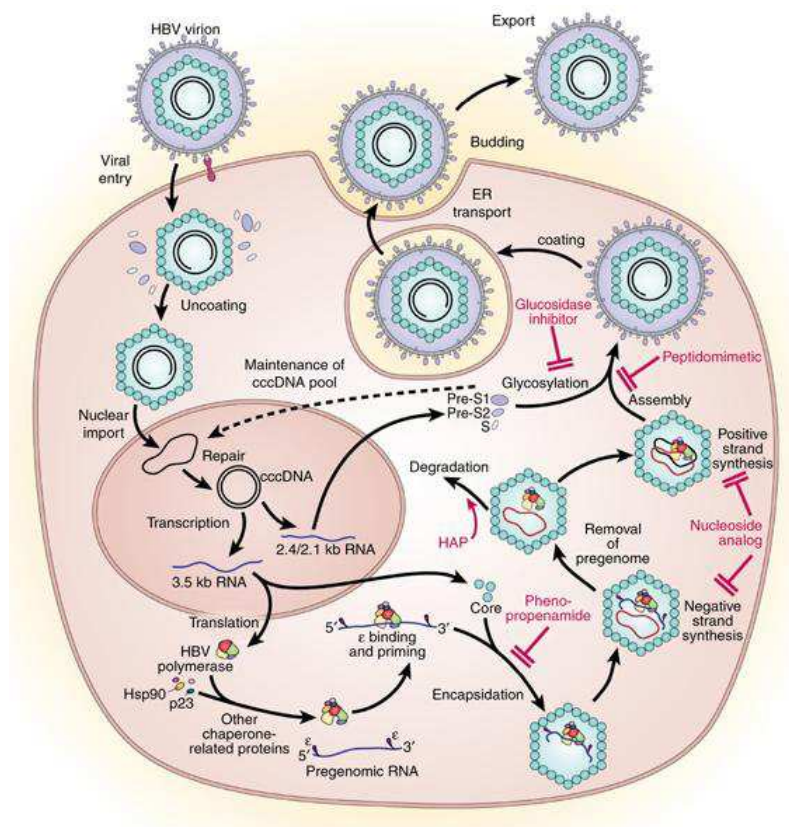


Figure 8 : Cibles antivirales connues sur le cycle cellulaire du VHB (Ghany and Liang 2007).

Leur efficacité antivirale s'explique notamment par le rôle important de l'association de la cyclophylène A au complexe de réplication, qui semble indispensable à sa réalisation (Kaul *et al.* 2009; Liu *et al.* 2009; Coelmont *et al.* 2010; Gallay 2012)

Le traitement du VHB

Le traitement de l'hépatite chronique B est fondé sur l'interféron alpha ou sur des inhibiteurs spécifiques de la transcriptase inverse du virus. Dans le cas du VHB, l'interféron alpha agit à deux niveaux, en bloquant de façon non spécifique la production de virus et en induisant les réponses immunes qui contrôlent l'infection (Ghany and Liang 2007). Le cycle du VHB (Figure 8) commence par une phase d'entrée qui utilise des récepteurs non identifiés à ce jour. Le virus est ensuite internalisé, décapsidé et le génome est transporté jusqu'au noyau sous forme d'ADN partiellement double brin. Une phase de réparation de l'ADN conduit à la formation d'un ADN entièrement double-brin et à son compactage en ADNccc (*covalently closed circular DNA*) ou ADN super-enroulé. L'ADNccc est utilisé par la cellule comme un mini-chromosome (Bock *et al.* 2001), la transcription du génome viral aboutissant à la synthèse d'un ARN pré-génomique et de plusieurs ARN messagers codant les protéines virales, notamment les protéines de capsid et l'ADN polymérase virale. Ces dernières qui s'assemblent avec un génome viral pour former une pré-particule virale non enveloppée. L'assemblage peut-être la cible du phenopropenamide et des heteroaryldihydropyrimidines (HAPs), qui ont fait la preuve d'une activité antivirale *in vitro* modeste (King *et al.* 1998; Weber *et al.* 2002). La polymérase virale assure, au sein de la particule provirale : (i) la synthèse d'un hétéroduplex ADN/ARN grâce à son activité transcriptase inverse ; (ii) la dégradation du brin d'ARN grâce à son activité RNase H ; (iii) la synthèse du brin

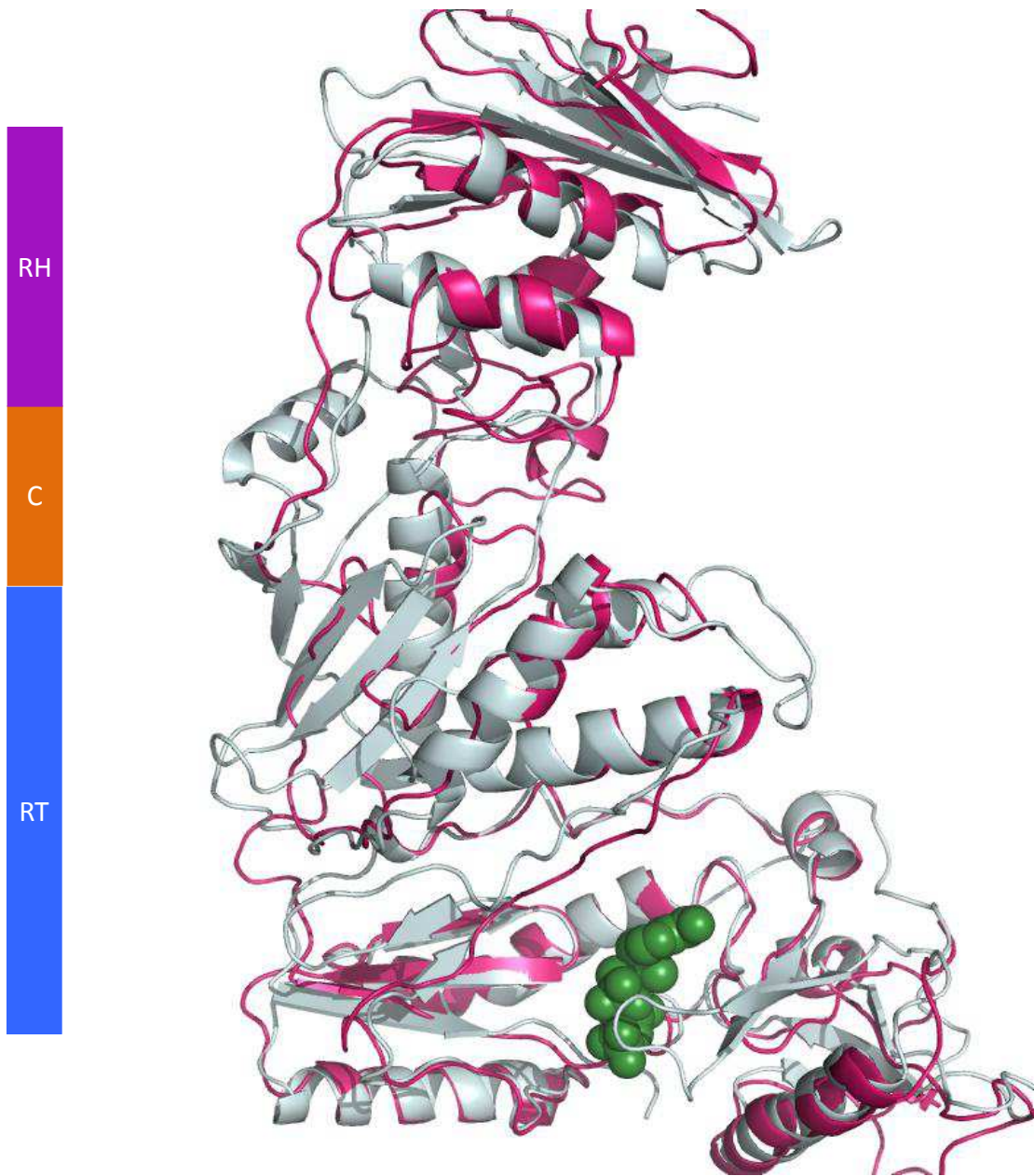


Figure 9 : Superposition des structures transcriptase inverse (RT), connecteurs (C) et RNAse H (RH) du VIH, obtenues par cristallographie (Tuske *et al.* 2004), et du VHB, obtenues par homologie (Hayer *et al.* 2012), modélisées en présence de tenofovir (inhibiteur nucléotidique de la transcriptase inverse du VHB et du VIH). Le pouce et le connecteur sont plus courts pour la protéine du VHB mais les deux structures sont très proches.

complémentaire d'ADN grâce à son activité ADN polymérase. Ces étapes aboutissent à la forme particulière mature du virion qui contient un ADN viral partiellement double-brin. Les deux activités enzymatiques de l'ADN polymérase virale peuvent être la cible de médicaments : les inhibiteurs nucléos(t)idiques de la transcriptase inverse sont la classe thérapeutique la plus utilisée aujourd'hui ; l'activité RNase H est une cible thérapeutique possible qui fait l'objet de recherches (Hayer *et al.* 2012). Les étapes finales d'assemblage peuvent être ciblées, par exemple par des inhibiteurs de glycosylation des protéines de surface virale, mais ces approches restent à ce jour au stade d'études *in vitro* (Block *et al.* 1998). Le virus est finalement exporté à l'extérieur de la cellule via le reticulum endoplasmique.

Résistance aux traitements antiviraux

Inhibiteurs de polymérase virale (nucléos(t)idiques et non nucléosidiques)

Les polymérases du VIH, du VHB, et du VHC sont des cibles thérapeutiques idéales dans la mesure où elles sont essentielles à la réplication et présentent des motifs protéiques spécifiques de chaque virus, permettant le développement d'inhibiteurs très spécifiques et a priori non toxiques pour les polymérases cellulaires. Ces molécules peuvent être classées en deux groupes en fonction de leurs cibles. Le premier groupe cible les ADN polymérases du VHB et du VIH, principalement leur activité enzymatique ADN polymérase dépendante de l'ARN ou transcriptase ; le 2^{ème} groupe cible l'ARN polymérase dépendante de l'ARN du VHC.

Malgré les différences importantes entre le VIH et le VHB, la parenté de fonction et de séquence entre les deux polymérases virales permet de superposer leurs structures (Figure 9) (Hayer *et al.* 2012) et d'identifier des homologies de position amino acidiques. Plusieurs inhibiteurs nucléos(t)idiques utilisés en pratique clinique sont efficace à la fois contre la

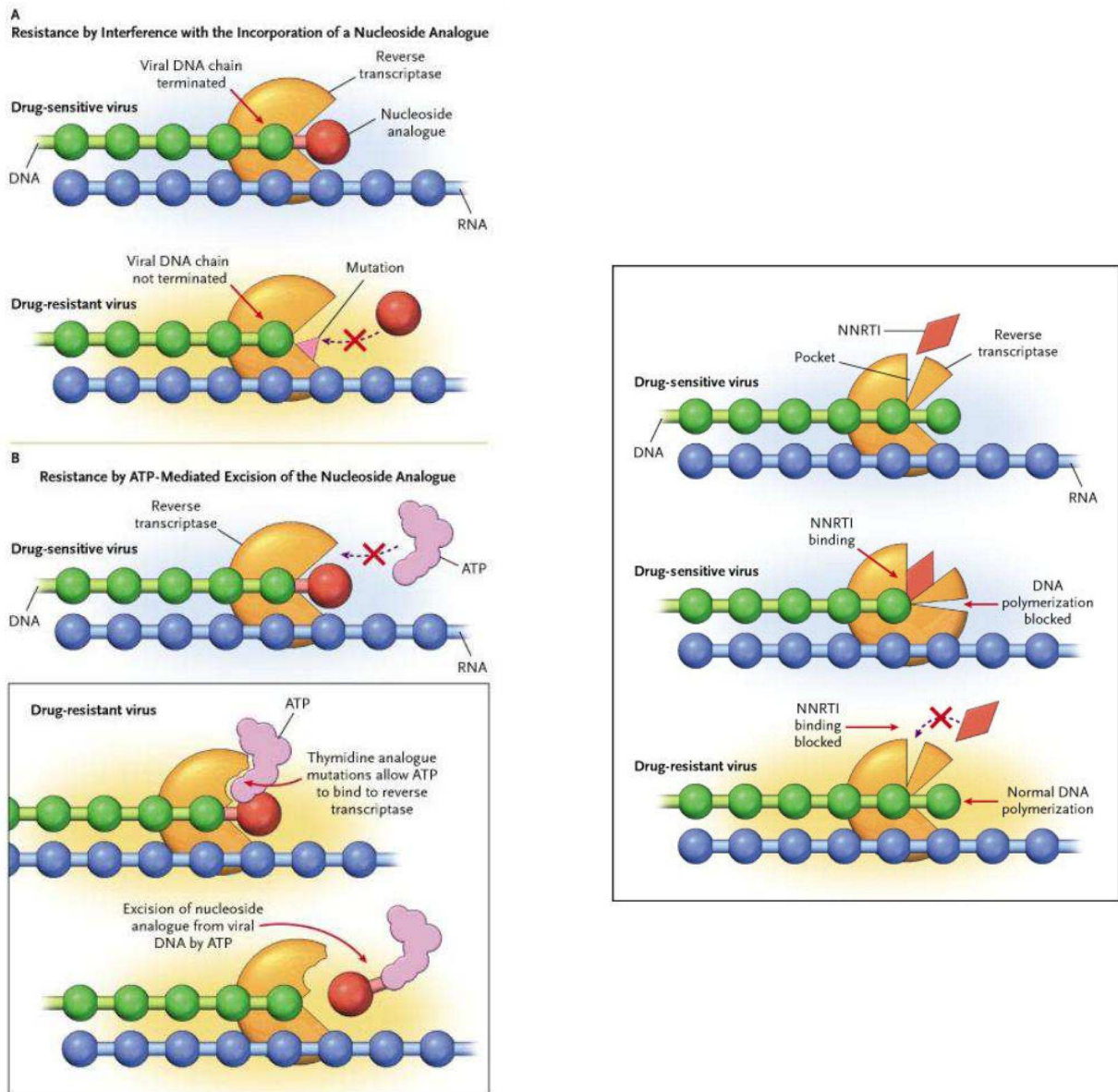


Figure 10 : Mécanismes d'action et principaux mécanismes de résistance observés pour les analogues nucléosidiques inhibiteurs de la transcriptase inverse (à gauche) et pour les inhibiteurs non nucléosidiques de la transcriptase inverse (NNRTI) (à droite) du VIH (Clavel and Hance 2004). Pour la première classe de molécule, la résistance à l'activité terminatrice de chaîne résulte principalement de 2 mécanismes : le défaut d'incorporation de la molécule antivirale ou l'excision de celle-ci, tandis que pour la 2^{ème} classe, il s'agit essentiellement d'un blocage du site de fixation de la molécule antivirale.

transcriptase inverse du VIH et celle du VHB (Bartholomeusz *et al.* 2004; Tuske *et al.* 2004). En revanche, l'ARN polymérase ARN-dépendante du VHC est une enzyme dont la structure et la fonction sont très différentes de celles des polymérases du VHB et du VIH, et aucune efficacité croisée n'est possible entre les inhibiteurs nucléos(t)idiques du VIH et du VHB et ceux du VHC.

L'effet des inhibiteurs nucléos(t)idiques est de bloquer la synthèse de nouveaux génomes par incorporation de didéoxynucléotides triphosphates dans la chaîne virale en cours de polymérisation (Figure 10) (Clavel and Hance 2004). Cette activité terminatrice de chaîne peut être évitée par deux principaux mécanismes de résistance, identifiés initialement pour le VIH :

- un défaut d'incorporation de l'inhibiteur dans la chaîne en cours de polymérisation,
- une excision de la molécule incorporée liée à l'incorporation d'ATP au niveau d'un site protéique adjacent suivie de l'attaque de la liaison phospho-diester établie entre la chaîne d'ADN en cours de polymérisation et le métabolite actif du médicament (Boyer *et al.* 2001).

Plusieurs générations d'inhibiteurs nucléos(t)idiques du VIH et du VHB se sont succédées, avec une amélioration progressive de leur puissance antivirale et de leur barrière génétique à la résistance (Figure 11A et B). Cependant, les analogues de nucléos(t)ides restent globalement des médicaments à faible barrière contre la résistance dans le traitement des infections par le VIH, si bien que leur utilisation n'est pas recommandée sans association à une autre classe médicamenteuse au minimum (Tuske *et al.* 2004). La situation est différente avec le VHB pour lequel la dernière génération de molécules (entecavir et tenofovir) permet de maintenir une suppression de la charge virale à long terme en

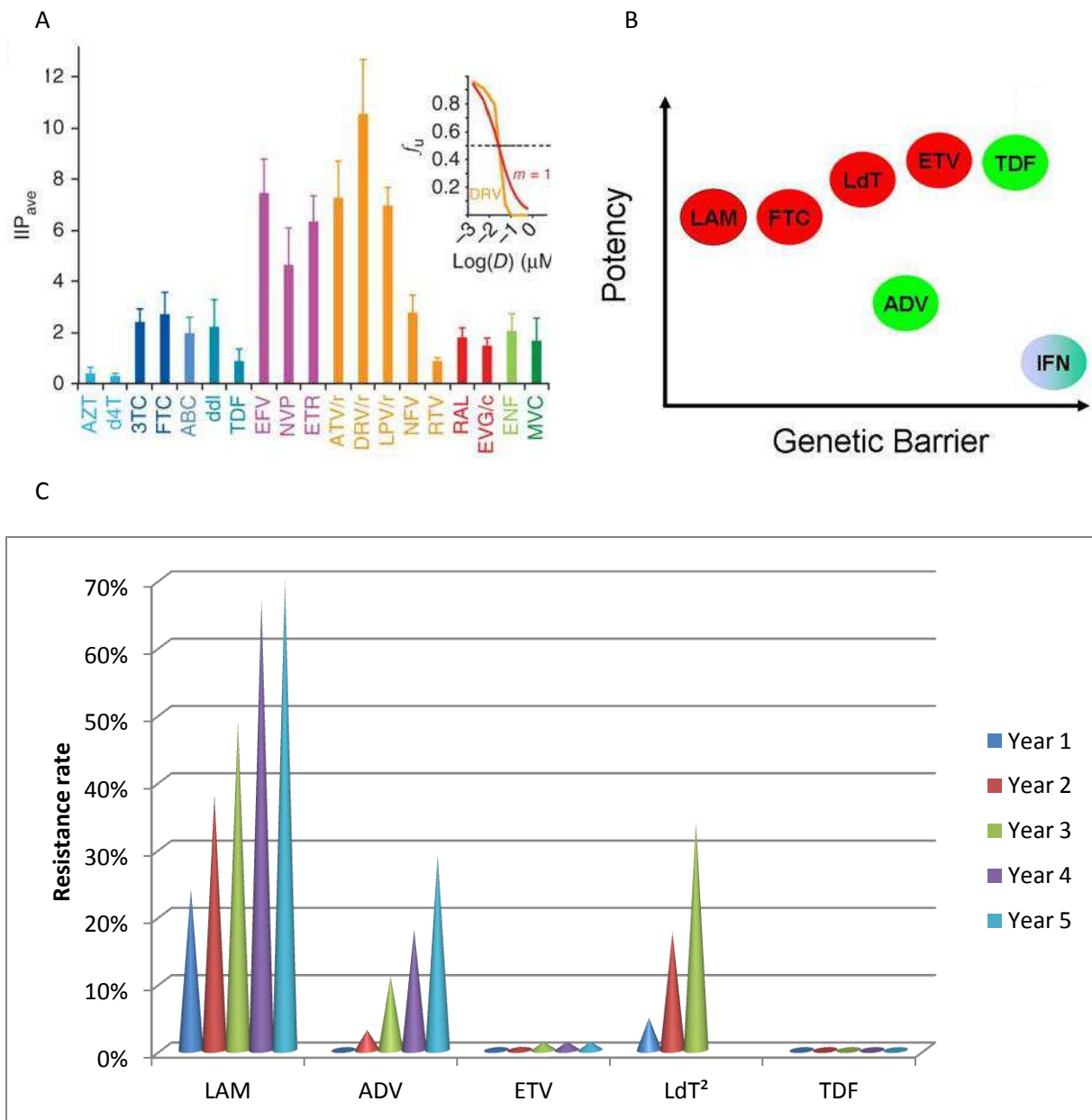


Figure 11 : A : comparaison des potentiels inhibiteurs instantanés des principales molécules utilisées dans le traitement du VIH (Jilek *et al.* 2012). B : Comparaison des potentiels antiviraux des molécules utilisées contre le VHB en fonction de leur barrière génétique à la résistance (Locarnini 2006). C : Incidence cumulative de la résistance aux traitements de l'hépatite chronique B chez les patients naïfs de traitement sur 5 ans (EASL 2009, Zhu *et al.* 2011).

monothérapie, avec un risque minime de survenue de résistance, sur un délai de 5 ans et plus (Figure 11 C) (EASL 2009; Zhu *et al.* 2011).

Les analogues nucléos(t)idiques utilisés dans le traitement de l'hépatite chronique C ont, in vitro comme in vivo, une barrière à la résistance élevée. La raison est que les mutations de résistance sélectionnées par ces traitements entraînent une faible perte de la puissance antivirale du médicament, associée à une importante perte de fitness (McCown *et al.* 2008; Pawlotsky *et al.* 2012). Ces résultats doivent être confirmés par des études cliniques à grande échelle. Les inhibiteurs nucléos(t)idiques représentent cependant une classe thérapeutique prometteuse dans la lutte contre l'infection par le VHC.

Des inhibiteurs non nucléosidiques des polymérases virales sont disponibles, uniquement pour le traitement des infections par le VIH et le VHC. Leur mécanisme d'action, du même type pour les deux virus, consiste à inhiber l'action enzymatique de la polymérase en la bloquant dans une conformation inactive, grâce à leur fixation à un site allostérique de surface de l'enzyme (Esnouf *et al.* 1995; Dorr *et al.* 2005; Haudecoeur *et al.* 2012). Les inhibiteurs non nucléosidiques de la transcriptase inverse du VIH ont un potentiel inhibiteur instantané parmi les plus élevés (Figure 11A), qui reste cependant associé à une barrière génétique faible. En effet, la présence d'une seule mutation de résistance est suffisante pour entraîner une diminution d'efficacité significative des premières générations d'inhibiteurs, deux mutations pour les dernières générations de traitement comme l'etravirine (Das *et al.* 2004; Chen *et al.* 2008). Dans une stratégie de trithérapie anti-VIH de première ligne, le choix se porte surtout aujourd'hui sur l'utilisation d'un inhibiteur non nucléosidique ou d'un inhibiteur de protéase en association à 2 analogues de nucléos(t)ides. L'intérêt des inhibiteurs non nucléosidiques réside surtout dans leur faible toxicité, à l'inverse des inhibiteurs de protéase qui présentent une barrière génétique plus élevée mais une toxicité

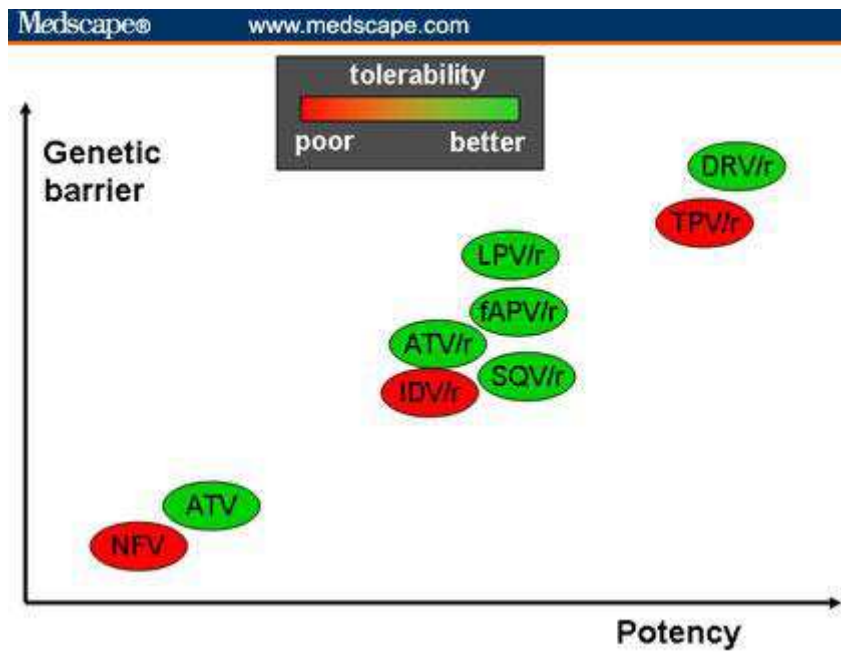


Figure 12 : Comparaison des barrières génétiques des inhibiteurs de protéase utilisés en thérapeutique anti-VIH en fonction de leur potentiel antiviral (Fernández-Montero *et al.* 2009). En rouge sont indiqués les médicaments dont la tolérance est faible, en vert celles présentant une meilleure tolérance.

plus importante pouvant altérer l'observance et, finalement, favoriser la résistance. Dans le cas des inhibiteurs non nucléosidiques de l'ARN polymérase du VHC, les molécules en cours de développement ont un profil du même type que celles utilisées dans le traitement du VIH, avec une barrière à la résistance faible (Le Pogam *et al.* 2006). Les premières générations de molécules souffrent cependant d'un spectre génotypique étroit.

Inhibiteurs de protéase (IPs)

La protéine NS3 du VHC possède un domaine serine protéase, qui forme un hétérodimère avec la protéine NS4A, et un domaine hélicase. La protéase NS3-4A assure quatre clivages protéolytiques de la polyprotéine du VHC, produite par traduction du cadre de lecture ouvert unique du virus (Lin 2006). La protéase du VIH est quant à elle classée parmi les protéases homodimériques aspartiques, dont le rôle consiste à permettre la maturation de la particule virale par clivage de la protéine p55 en p17, p24, p9 et p7 (Henderson *et al.* 1992). Les inhibiteurs de protéase dirigés contre ces deux virus appartiennent à des classes thérapeutiques très différentes, avec des mécanismes d'action très différents.

Dans le cas du VIH, la puissance antivirale des inhibiteurs de protéase est la plus élevée, surtout si elle est boostée par l'adjonction de ritonavir, un inhibiteur du cytochrome P450 (CYP450) dont l'activité antivirale est modeste (Figure 11A). Les inhibiteurs de protéase du VIH bénéficient également d'une barrière génétique très élevée et sont à ce jour les médicaments de référence dans la prévention de la résistance (Tuske *et al.* 2004). Ils possèdent toutefois un inconvénient majeur lié à leur toxicité élevée, qu'il convient de prévenir par un suivi fréquent et éventuellement la mise en place de thérapeutiques préventives. Les inhibiteurs de protéase de dernière génération, comme le

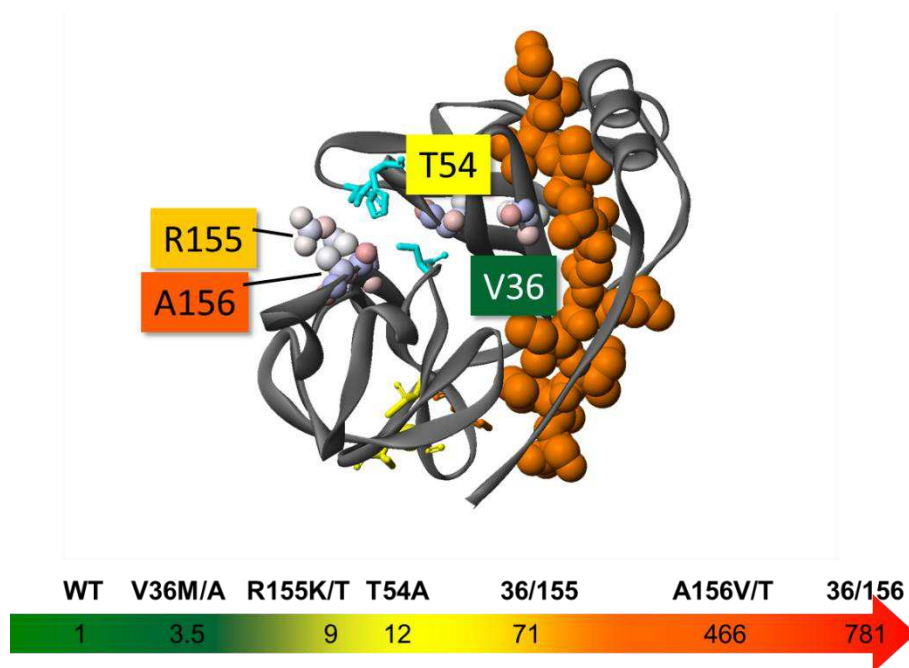


Figure 13 : Représentation en 3D de la protéase NS3 du VHC avec les principales mutations de résistance aux inhibiteurs de protéase connues. Ces mutations sont surlignées avec une couleur allant du vert (faible) au rouge (fort) selon le degré d'altération d'efficacité du traitement antiviral. Elles sont également représentées sur une échelle chiffrée correspondant à l'augmentation de la dose efficace à 50% par rapport à la souche sauvage (WT) in vitro.

darunavir/ritonavir (DRV/r), bénéficient d'un fort potentiel antiviral avec une toxicité plus limitée (Figure 12) (Ortiz *et al.* 2008; Fernandez-Montero *et al.* 2009).

Les inhibiteurs de protéase du VHC de première génération ont une faible barrière génétique, avec seulement une mutation nécessaire à l'augmentation importante de la concentration inhibitrice à 50% *in vitro* (Figure 13) (Sarrazin *et al.* 2007; Barbotte *et al.* 2010). Cette particularité se traduit cliniquement par la sélection en seulement quelques jours de variants viraux résistants qui deviennent majoritaires, en monothérapie (Chevaliez *et al.* 2011) comme en trithérapie associant interféron alpha pégylé, ribavirine et inhibiteur de protéase chez les patients mauvais répondeurs à l'interféron (McHutchison *et al.* 2010). Le traitement de l'hépatite chronique C par les inhibiteurs de protéase nécessite donc un suivi très régulier afin d'évaluer la réponse au traitement et de prévenir l'échec thérapeutique associé à la sélection de variants viraux résistants (EASL 2011; Pawlotsky 2011). Les effets secondaires sont également fréquents avec les inhibiteurs de protéase de première génération du VHC.

Inhibiteurs de cyclophilines

Les inhibiteurs de cyclophiline (Cyp) sont une classe prometteuse de médicaments antiviraux. Leur chef de file est l'alisporivir qui présente une activité antivirale modeste contre le VIH (Ptak *et al.* 2008), plus importante contre le VHC (Paeshuyse *et al.* 2006). Une activité anti-VHB a également été récemment rapportée, mais elle semble marginale (moins de 1 log de décroissance) (McMillan *et al.* 1995; Chokshi *et al.* 2012). Dans le cas du VIH, l'alisporivir inhibe l'interaction Cyp-A/capside ; cette effet est associé à une diminution de la réplication du virus dont le mécanisme n'est pas entièrement élucidé (Daelemans *et al.* 2010). Il a également été rapporté une diminution d'activité du facteur de restriction

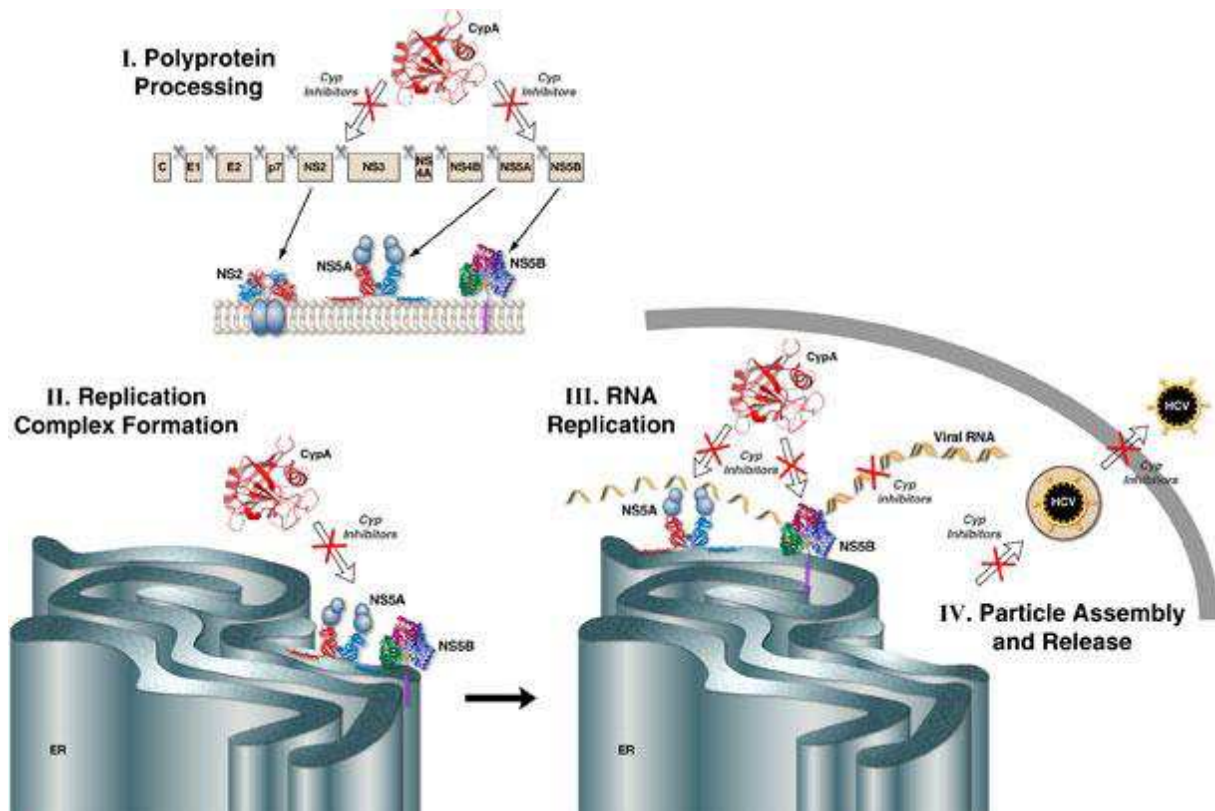


Figure 14 : Mécanismes d'action des inhibiteurs de cyclophiline A anti-VHC. Les 4 mécanismes suivants pourraient nécessiter l'intervention de la cyclophiline A et seraient donc affectées par les inhibiteurs de cyclophiline : 1) clivage des protéines NS5A/NS5B et NS2/NS3 et folding de la protéine NS2, 2) formation du complexe de réplication, 3) réplication, 4) assemblage et transport dans les gouttelettes lipidiques (Gallay 2012, Kaul *et al.* 2009, Liu *et al.* 2009, Coelmont *et al.* 2010).

TRIM5 α , qui nécessite une liaison à la Cyp-A pour exercer son activité (Hatzioannou *et al.* 2005). Dans le cas du VHC, l'action antivirale semble se faire à plusieurs niveaux (Figure 14) (Gallay 2012). Un premier niveau d'inhibition pourrait survenir à l'étape de processing de la polyprotéine du VHC, notamment au niveau des clivages NS5A/NS5B et NS2/NS3 et du folding de la protéine NS2 (Kaul *et al.* 2009). Le mécanisme le plus important semble cependant lié au rôle important de la cyclophiline dans la formation du complexe de réplication (Liu *et al.* 2009) et dans la réplication elle-même (Coelmont *et al.* 2010), ainsi que dans l'assemblage et la libération des virions en lien avec le trafficking des gouttelettes lipidiques (Kaul *et al.* 2009). Pour le VHC, des mutations de résistances ont été rapportées essentiellement au niveau des protéines NS3, NS5A et NS5B, avec une prédominance de l'effet des mutations portées par NS5A qui modifieraient les interactions avec la Cyp-A (Puyang *et al.* 2010). Le niveau de résistance conféré est cependant faible et son rôle dans l'échec thérapeutique n'est pas démontré. Dans le cas du VIH, les mutations de résistance sont sélectionnées essentiellement au niveau du domaine de fixation de la Cyp-A à la capsid virale et entraînent une résistance au mécanisme de défense TRIM5 α (Chatterji *et al.* 2005).

Autres DAAs

Les autres inhibiteurs directs utilisés ou en cours de développement sont spécifiques de chaque virus et ne peuvent donc être comparés. Ils comprennent les inhibiteurs d'intégrase du VIH, les inhibiteurs d'entrée virale du VIH et les inhibiteurs de NS5A du VHC.

L'intégrase du VIH a pour fonction d'intégrer le génome viral au sein du génome humain. Pour exercer cette fonction, l'intégrase utilise trois domaines reliés entre eux par un connecteur (Chen *et al.* 2008). Les études de docking ont montré que les principales

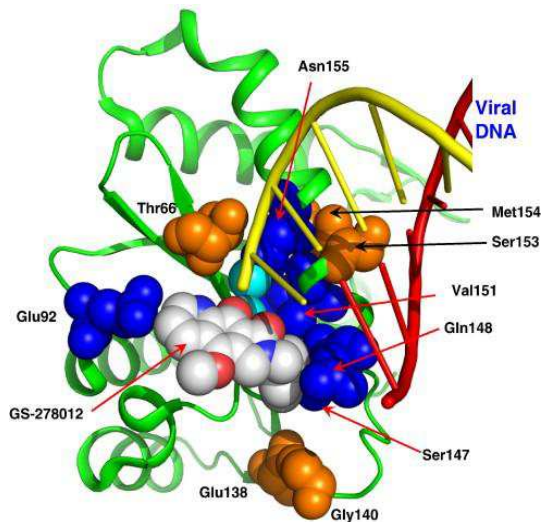


Figure 15 : Localisation des mutations de résistance aux INSTI sur la structure de l'intégrase du VIH. L'ADN viral est représenté en jaune et rouge, les mutations principales en bleu et les mutations secondaires en marron. Le GS-278012 (INSTI) est représenté en blanc (Chen *et al.* 2008).

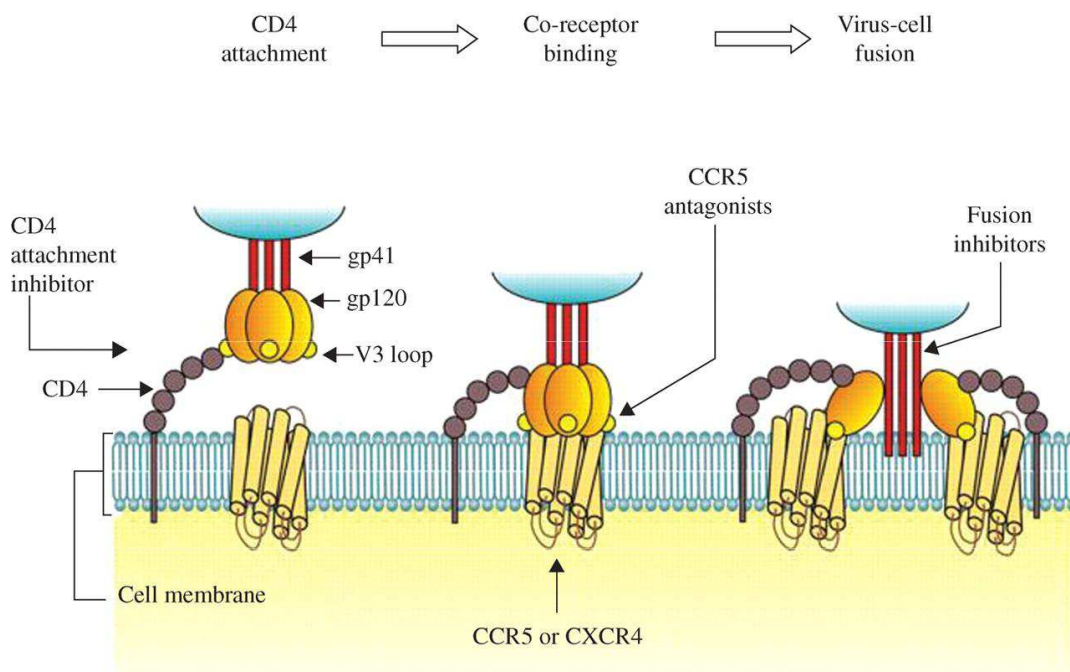


Figure 16 : Représentation schématique de l'entrée virale du VIH dans le lymphocyte T CD4 positif (Vandekerckhove *et al.* 2009). La première étape de l'entrée consiste en la fixation du récepteur viral gp120 sur le récepteur principal CD4. La deuxième étape est la fixation du virus au corécepteur CCR5 ou CXCR4 selon son tropisme, qui entraîne l'ouverture de la gp120 et le déploiement de la molécule gp41 au sein de la membrane plasmique cellulaire. Enfin le repliement des domaines HR1 et HR2 de la gp41 permet le rapprochement des membranes virales et cellulaire et leur fusion.

mutations de résistance aux inhibiteurs d'intégrase étaient toutes situées à proximité immédiate de leur zone de fixation (Figure 15). Le mécanisme de résistance supposé est un encombrement stérique défavorable à la fixation de l'inhibiteur dans la poche catalytique. La barrière génétique et la toxicité des inhibiteurs d'intégrase sont comparables à celles des analogues nucléos(t)idiques et font que ces deux classes médicamenteuses peuvent jouer un rôle voisin dans la stratégie thérapeutique. Les inhibiteurs d'intégrase sont cependant plus chers.

Les inhibiteurs d'entrée du VIH comprennent deux classes médicamenteuses distinctes dont les chefs de file sont : (i) le maraviroc, qui inhibe la fixation de la gp120 au corécepteur CCR5 (Figure 16)(Dorr *et al.* 2005; Vandekerckhove *et al.* 2009), et (ii) l'enfuvirtide, analogue du domaine HR2 de la gp41 qui empêche le rapprochement des domaines HR1 et HR2 et par conséquent la fusion entre les membranes cellulaires et virales (Figure 16)(Kilby and Eron 2003). La résistance clinique au maraviroc peut-être liée à un « glissement » du phénotype R5 vers le phénotype X4. Il s'agit dans ce cas de la sélection de souches X4 pré-existantes associée à une décroissance relative de la prévalence des souches R5 (Westby *et al.* 2007). Un mécanisme de résistance par sélection de mutants de souches R5 a également été observé. Ces mutants conservent un tropisme R5 et leur capacité de se fixer au récepteur CCR5 malgré la présence de l'inhibiteur (Westby *et al.* 2007). Les mécanismes de résistance à l'enfuvirtide sont liés à la sélection de mutations dans les domaines HR1 et HR2 de la gp120 (Greenberg and Cammack 2004). Les mutations HR1 causeraient un défaut de fixation de la molécule, au prix d'un défaut d'attachement de la chaîne HR2 et donc d'une perte importante de fitness (Rimsky *et al.* 1998). Les mutations pourraient également modifier la cinétique d'entrée virale, en diminuant de façon importante le temps d'exposition de la gp41 à l'enfuvirtide (Reeves *et al.* 2005).

Des inhibiteurs de la protéine NS5A du VHC sont en cours de développement pré-clinique et clinique (Gao *et al.* 2010). Le daclatasvir est le chef de file de cette classe, actuellement en Phase III. Un essai de phase II a été réalisé où le daclatasvir était associé à l'inhibiteur de protéase asunaprevir (BMS-650032), avec ou sans interféron alpha et ribavirine, chez des patients répondeurs nuls à un premier traitement par interféron et ribavirine (Lok *et al.* 2010). Dans cette étude, des échappements ont été rapportés chez les patients recevant les deux DAAs en bithérapie. Ils étaient liés à une résistance aux deux molécules, associée à la sélection de mutations dans les régions NS5A et NS3 présentes sur le même brin viral. *In vitro* dans le système du réplicon, il a été montré que certaines mutations de la protéine NS5A pouvaient affecter l'efficacité de l'inhibiteur. Le mécanisme de la résistance reste inconnu, la fonction de cette protéine restant à ce jour peu claire. NS5A semble participer à la cohésion du complexe de réplication et réguler la fonction de l'ARN polymérase virale ; elle est également dotée d'une fonction transactivatrice en cours d'investigation dans notre laboratoire (Polyak *et al.* 2001; Maqbool MA *et al.* 2012).

Les outils d'étude de la résistance

La résistance virologique se traduit par l'augmentation de la charge virale de 1 log au-dessus du nadir chez un patient sous traitement ayant initialement répondu par une décroissance significative de celle-ci. Naturellement, c'est donc la charge virale qui représente l'outil de référence du dépistage de la résistance. Une remontée de la charge virale sous traitement indique que le virus a atteint un niveau de résistance et acquis un fitness élevés, à tel point que le processus a atteint un point de non-retour avec la thérapeutique utilisée. Cependant, aucune information d'ordre mécanistique, pronostique ou utile à l'adaptation du traitement n'est fournie par cette approche. Des outils capables d'anticiper l'émergence de la résistance clinique le plus tôt possible au cours du traitement sont donc nécessaires. Ces outils doivent être très sensibles, fournir une information clinique utile donc être très spécifiques, et permettre l'adaptation rapide du traitement pour éviter son échec (Lok *et al.* 2007; Pawlotsky *et al.* 2008; Kwong *et al.* 2011). Trois types d'approches peuvent être utilisés.

Tests biochimiques enzymatiques

Ils permettent de quantifier le degré de résistance conféré par une ou plusieurs mutations d'une protéine virale à un de ses inhibiteurs spécifiques en dehors du contexte de répllication du virus. Ils supposent que la cible antivirale soit connue et que la protéine puisse être synthétisée en quantité importante, sans et avec la mutation suspectée. Ces méthodes ont l'avantage de pouvoir être automatisées. Elles représentent donc un test de choix dans le contexte du screening de nouvelles drogues antivirales. Leur faiblesse réside dans leur

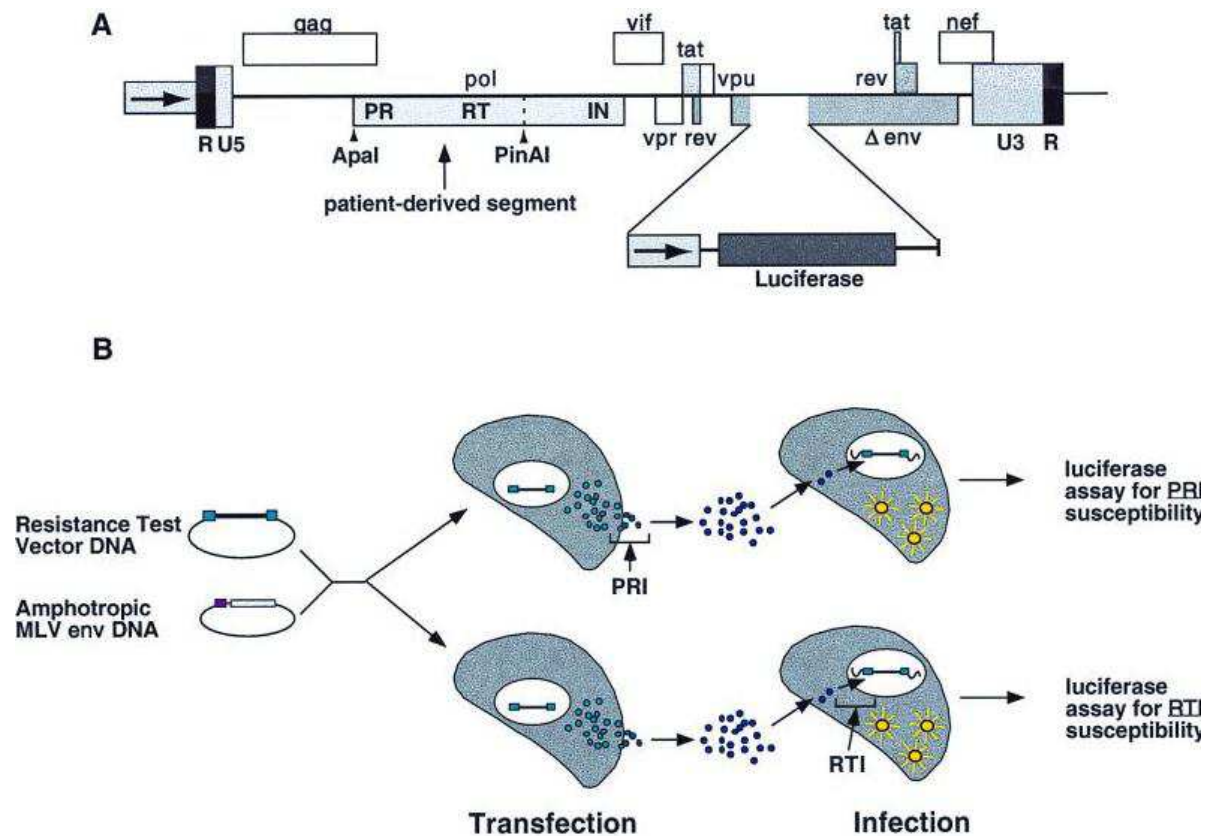


Figure 17 : Schéma du fonctionnement du test Phenosense® (Petropoulos *et al.* 2000). A. Le test débute par le clonage du segment génique viral d'intérêt (PR : protéase, RT : transcriptase inverse, IN : intégrase) à partir de la souche de VIH infectant le patient sur un fond génétique VIH intégrant le gène luciférase à la place du gène d'enveloppe. B. Une première transfection cellulaire est ensuite réalisée avec un vecteur MLV contenant le gène d'enveloppe du VIH. Les cellules contenant les 2 vecteurs produiront un virus infectieux par transcomplémentation. Ces particules virales sont ensuite utilisées pour infecter des cellules cultivées en présence d'un inhibiteur (PRI : protease inhibitor, RTI : reverse transcriptase inhibitor). Les cellules infectées produisent une quantité de luciférase qui dépend de la sensibilité du virus au médicament. Une mesure de luciférase permet de définir la sensibilité du virus par rapport à une référence.

relevance clinique parfois faible, du fait de l'isolement du processus de résistance de tout contexte cellulaire.

Tests phénotypiques cellulaires

Ces tests permettent d'établir avec certitude, dans un contexte répliatif cellulaire, la diminution de sensibilité d'une souche virale à une drogue antivirale par rapport à une souche de référence. La limite principale de cette approche est la nécessité de disposer d'un modèle cellulaire répliatif du virus étudié. En pratique, cette condition n'est satisfaite que pour le VIH sans modification du fond génétique de la souche testée. Malgré cela, les tests VIH commerciaux actuels utilisent un fond viral génétiquement modifié afin de faciliter la quantification, standardisé pour permettre les comparaisons. A l'inverse, les souches virales de VHB ou de VHC obtenues directement à partir de prélèvements de patients ne sont pas suffisamment répliatives pour être utilisées dans un test phénotypique. Les tests phénotypiques du VIH, du VHB ou du VHC ont donc tous nécessité l'utilisation d'artifices géniques pour pouvoir être réalisés de manière performante et standardisée, au prix d'un éloignement avec la répliation naturelle du virus prélevé chez le patient. Cette limite doit être prise en considération lors de l'interprétation des résultats et nécessite souvent de renouveler les expériences sur différents modèles.

Tests phénotypiques du VIH

Actuellement, deux tests commerciaux (Phenosens® et Antivirogram®) sont utilisés pour la réalisation de tests phénotypiques de la résistance aux antiviraux de première ligne du VIH. Ils sont fondés sur une stratégie de clonage des gènes de la protéase, de la transcriptase inverse, de l'intégrase et de la région 3' de gag de la souche étudiée dans une souche de

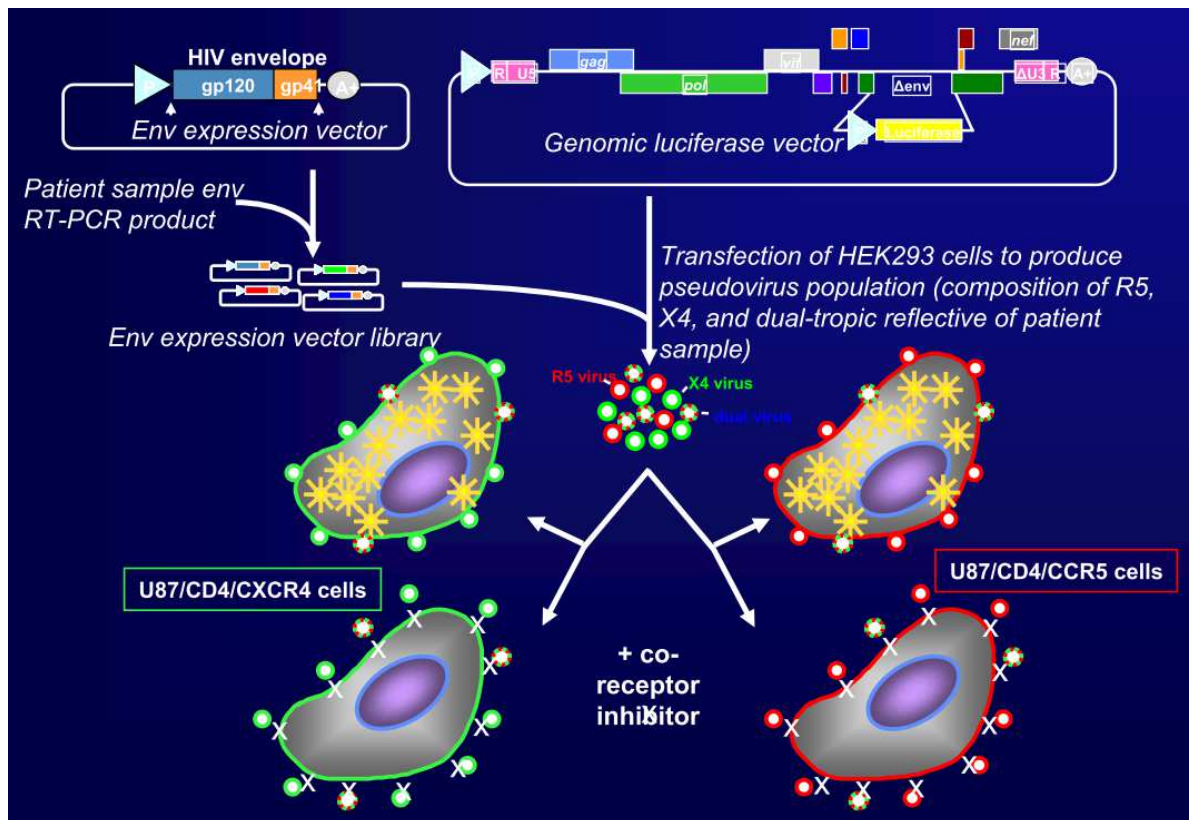


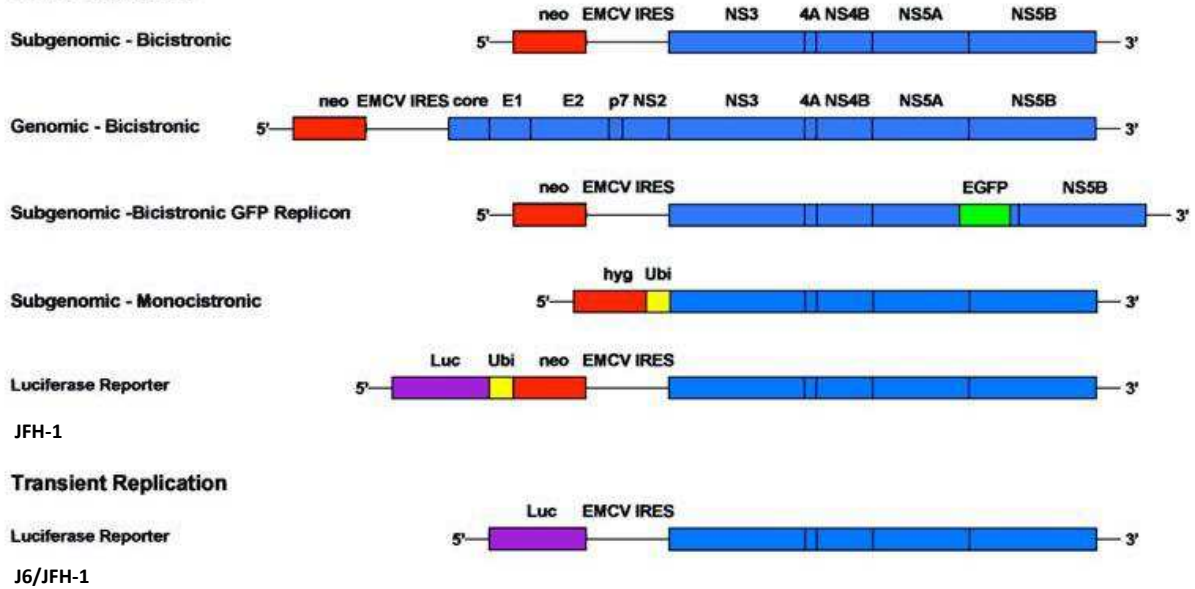
Figure 18 : Schéma du fonctionnement du test Phenosense® Entry (Coakley *et al.* 2005). Le test débute par le clonage des gènes d'enveloppe gp120 et gp41 et se poursuit par un clonage dans un vecteur luciférase puis une transfection sur cellules HEK afin de produire des virions dont le tropisme est en relation avec celui observé chez le patient. Les virions sont ensuite introduits dans le surnageant de culture de cellules CCR5 ou CXCR4 afin de vérifier leur capacité d'infection. La quantité de luciférase produite dans les différents types cellulaires permet de déterminer le tropisme R5, X4 ou R5X4 des souches virales isolées du patient.

référence fournie dans les kits (Figure 17) (Petropoulos *et al.* 2000). Il existe également des tests évaluant le géotropisme CCR5 ou CXCR4 du VIH qui peuvent être utilisés avant la mise en place d'un traitement par maraviroc. Le principe est identique, fondé sur le clonage de la gp120 et de la gp41 dans un vecteur VIH luciférase puis la transfection de cellules HEK293 qui produisent des virions dont le tropisme est identique à celui observé chez le patient. Ces virus sont ensuite utilisés pour infecter des cellules R5 ou X4. Les cellules R5 et/ou X4 expriment la luciférase, dont l'activité peut être mesurée, en fonction du tropisme R5, X4 ou mixte de la souche du patient (Figure 18) (Coakley *et al.* 2005).

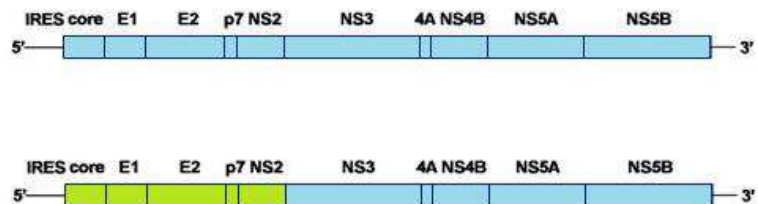
Tests phénotypiques du VHC

Le VHC ne se réplique pas naturellement en culture cellulaire. Il a fallu adapter des isolats cliniques pour les rendre utilisables dans le cadre de tests phénotypiques, fondés sur l'utilisation de réplicons subgénomiques ou génomiques (pour revue : (Blight and Norgard 2006)). A ce jour, des réplicons sont disponibles pour 4 des 6 géotypes (Figure 19). Les premières générations de réplicons ont été adaptées à partir de souches virales de géotype 1b, les souches con1 (Lohmann *et al.* 1999) et N (Guo *et al.* 2001), qui incluaient des gènes de sélection tels que le gène de résistance à la néomycine ou à l'hygromycine afin de permettre leur stabilisation dans la lignée cellulaire d'hépatocarcinome humain Huh7. Ces souches subgénomiques, bicistroniques et stables ont ensuite été améliorées pour faciliter leur étude à l'aide de gènes rapporteurs tels que la GFP (Moradpour *et al.* 2004), la luciférase (Krieger *et al.* 2001) ou les phosphatases alcalines (Yi *et al.* 2002), ou pour étudier le virus doté de ses protéines structurales dans des réplicons génomiques (Ikeda *et al.* 2002). Les réplicons de géotype 1a ont été plus difficiles à développer. Ils dérivent de la souche H77, infectieuse chez le chimpanzé (Kolykhalov *et al.* 1997), qui a dû subir de nombreuses mutations adaptatives (Gu *et al.* 2003; Yi and Lemon 2004). Une innovation majeure est

Stable Replication



Genotype 1a/1b



Genotype 2a

Figure 19 : Schéma des principales constructions de réplicons du VHC utilisées (adapté de Brass *et al.* 2006). Les réplicons peuvent être utilisés sous forme stable dans des cellules adaptées, ou sous forme transitoire ; des réplicons de génotypes 1, 2, 3 et 4 sont aujourd'hui disponibles et les rapporteurs peuvent être de différents types, notamment la luciférase ou la GFP.

intervenue avec le développement du réplicon de génotype 2a, nommé JFH-1 (Kato *et al.* 2003). Cette souche, obtenue à partir d'un isolat clinique de patient ayant présenté une hépatite fulminante, est le seul exemple de VHC capable d'exécuter un cycle cellulaire complet générateur de particules virales infectieuses en culture cellulaire. Cette souche a été adaptée pour constituer une chimère avec les gènes structuraux d'une autre souche de génotype 2a, la souche J6. La souche chimérique J6-JFH-1 a l'avantage d'être hautement répliquative et constitue la souche de génotype 2a la plus utilisée en laboratoire pour les tests phénotypiques (Lindenbach *et al.* 2005). Des réplicons de génotypes 3a et 4a ont été récemment développés. Cependant, il est aujourd'hui impossible de tester des souches complètes provenant de patients infectés dans les systèmes réplicons. L'approche consiste à introduire des portions de génome à tester sous forme de cassette ou à travailler par mutagenèse dirigée pour tester l'effet d'une mutation ponctuelle.

Tests phénotypiques du VHB

Comme le VHC, le VHB doit s'adapter pour atteindre un niveau répliquatif suffisant pour être utilisé dans un test phénotypique. Deux types d'approches ont été testées avec succès, fondées sur l'utilisation ou non d'un vecteur (Figure 20) (Durantel *et al.* 2005). L'approche sans vecteur consiste à réaliser une amplification du génome complet du VHB avec des amorces contenant des enzymes de restriction qui permettent, à posteriori, une circularisation simulant un génome viral (Gunther *et al.* 1995). Après transfection, la transformation de l'amplicon circulaire en cccDNA permet la production virale en cellules Huh7. Cependant, cette production étant faible et l'amplification par PCR du génome complet étant difficile à mettre en œuvre, ce sont les approches fondées sur la vectorisation qui se sont imposées (Durantel *et al.* 2004). Les approches vectorielles permettent notamment d'introduire un promoteur fort tel que le CMV/chicken β -actin en amont du

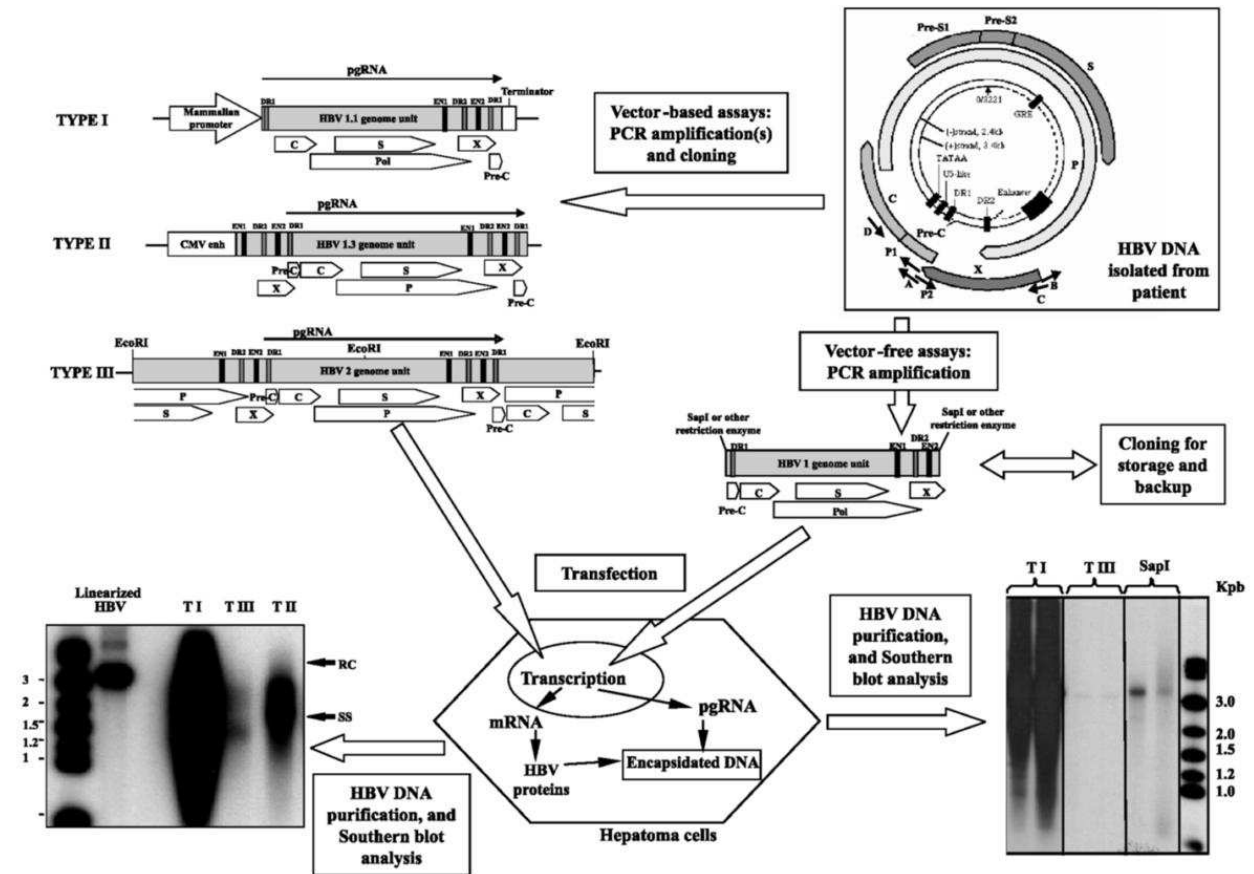


Figure 20 : Schéma des principales constructions de réplicons du VHB utilisées (Durantel *et al.* 2005). Les réplicons se classent en 2 groupes, les systèmes non vectorisés (à droite) et les systèmes vectorisés (à gauche). Parmi les systèmes vectorisés, il existe différents fonds génétiques de 1.1, 1.3 ou 2.0 génomes, éventuellement sous la dépendance de promoteurs forts. Les systèmes les plus répliatifs, qui sont les plus utilisés, sont notés Type II sur le schéma.

génomique viral afin d'accroître considérablement la production d'ARN. Cependant, les constructions de ce type se comportent comme si elles étaient linéaires ; or, l'ARN pré-génomique, qui sert de matrice à la synthèse de nouveaux génomes, a une longueur de 1,1 génome. Il est donc nécessaire d'introduire au minimum 1,1 génome dans les constructions vectorisées. Des systèmes à 1,1 (Durantel *et al.* 2004), 1,3 (Delaney and Isom 1998) ou 2,0 génomes ont été testés avec succès. Ces approches restent cependant difficiles à mettre en œuvre au quotidien pour tester des souches virales issues d'isolats cliniques, notamment en raison de la faible capacité répliquative de nombreuses souches de patients dans ce système et de la difficulté à mettre en œuvre une PCR du génome complet. L'utilisation de systèmes répliquatifs à partir de souches éprouvées modifiées par mutagenèse dirigée permettant d'introduire les mutations d'intérêt est une approche possible et fonctionnelle (Angus *et al.* 2003).

Tests génotypiques

Les tests phénotypiques sont importants pour caractériser l'effet d'une mutation dans un système global ou semi-global de répllication. Cependant, les mécanismes de la résistance ne peuvent être expliqués, car ces tests ne permettent pas de déterminer les causes de l'échappement du virus au traitement. Les tests génotypiques, rapides et moins coûteux, permettent de préciser la nature et la position génomique des mutations sélectionnées au cours des échappements thérapeutiques et de caractériser les dynamiques des populations virales au cours du temps. Ces techniques seront abordées en 2 parties : (i) les techniques conventionnelles, telles que le séquençage Sanger, les techniques d'hybridation inverse de

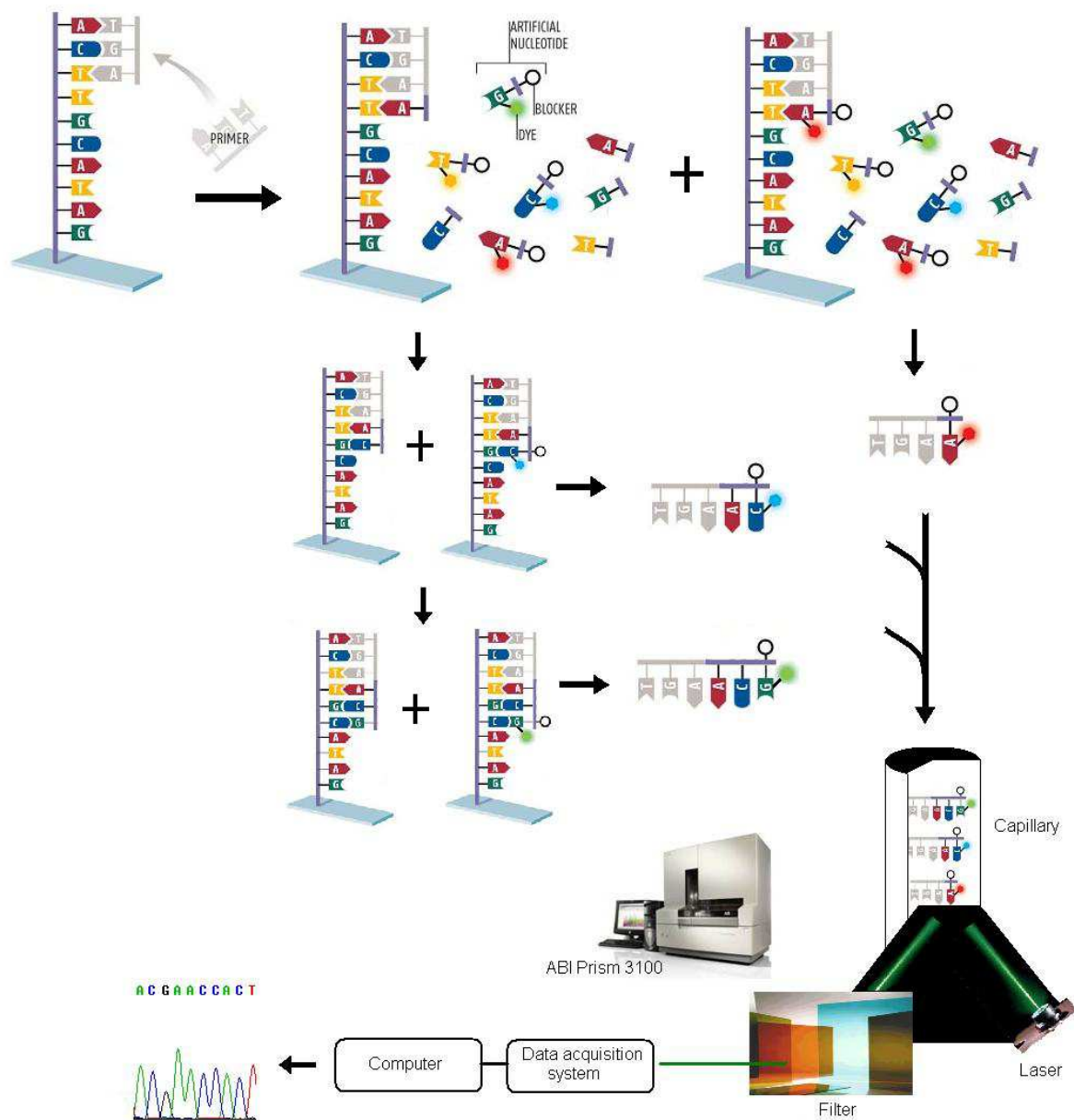


Figure 21 : Principe du séquençage capillaire : la PCR de séquençage permet, après fixation de l'amorce (primer), une élongation à l'aide de bases marquées « terminateurs de chaîne ». Leur incorporation aléatoire génère des fragments d'ADN marqués de tailles différentes. Après migration dans un capillaire, les fragments sont séparés en fonction de leur taille puis détectés par l'intermédiaire du fluorochrome excité par un laser. Le signal recueilli est analysé et traduit en électrophorégramme, permettant d'accéder à la séquence.

type LIPA et le clonage-séquençage (CS) ; (ii) les techniques de séquençage à haut-débit, largement utilisées dans les travaux présentés dans cette Thèse.

Techniques de génotypage conventionnelles

La méthode la plus ancienne, de loin la plus utilisée, est la méthode de Sanger, qui permet la réalisation de séquençages directs ou « population sequencing » (PS). La méthode de Sanger a été décrite pour la première fois en 1977 (Sanger *et al.* 1977) et améliorée par l'introduction de l'électrophorèse capillaire, qui a permis d'accroître rendement et sensibilité, et de la fluorescence qui a permis de s'affranchir de la radioactivité (Prober *et al.* 1987; Smith 1991). Le séquençage direct se déroule comme une PCR, avec des cycles répétés de trois étapes : dénaturation, hybridation des amorces de séquençage et élongation. Au cours de l'étape d'élongation (Figure 21), la polymérase introduit de façon aléatoire des déoxynucléotides (dNTP) ou des didéoxynucléotides (ddNTP) marqués par un fluorochrome différent. Ces derniers, dits terminateurs de chaîne, arrêtent la synthèse lorsqu'ils sont incorporés. A la fin du processus, le mélange réactionnel contient un mélange de produits d'amplification de différentes longueurs qui se terminent par un ddNTP fluorescent, dont la couleur correspond à l'une des bases A, T, C ou G. Les produits d'amplification sont migrés dans un capillaire afin de les séparer en fonction de leur taille, puis détectés à l'aide d'un détecteur de fluorescence après excitation par un laser. La couleur, l'intensité et le temps d'apparition du pic de fluorescence émis par chaque amplicon migré conduit à la création d'un électrophorégramme qui permet de déterminer la composition en bases du génome étudié. Lorsque plusieurs variétés de génomes coexistent dans le mélange réactionnel, comme c'est le cas pour les virus hautement variables, la séquence obtenue par séquençage direct correspond à la moyenne des fluorescences détectées à chaque position, c'est-à-dire à la séquence consensus. On peut parfois noter la présence de doubles pics quand les

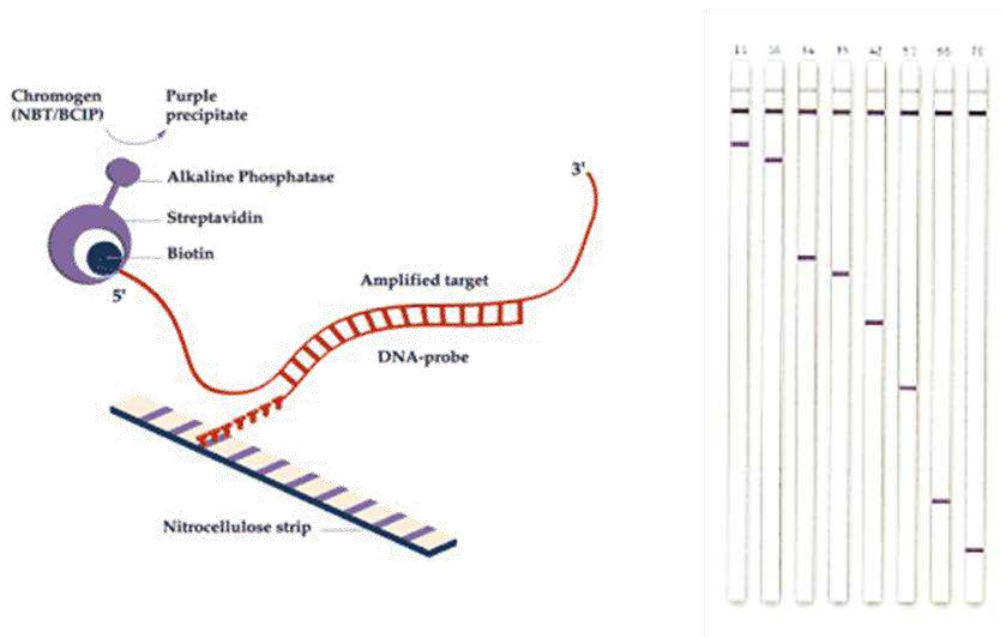


Figure 22 : Schéma du principe de la technique INNO-LiPA. Le gène étudié est amplifié par PCR à l'aide d'amorces biotinylées, puis les amplicons sont hybridés à une bandelette de nitrocellulose sur laquelle des amorces complémentaires de la séquence recherchée ont été fixées (à gauche). Une révélation sous forme de bande colorée est ensuite réalisée par utilisation d'une phosphatase alcaline streptavidinée se fixant sur la biotine du produit d'hybridation (à droite).

populations virales majoritaires sont présentes en proportions équivalentes. Il est donc possible d'identifier la présence de plusieurs populations virales, mais pas possible d'en déterminer les séquences car l'information sur le lien entre les bases est perdue. La sensibilité de la technique ne permet pas de détecter la présence de variants dont la fréquence est inférieure à 15 à 20%, car leur signal se confond avec le bruit de fond de la technique.

Une technique fondée sur l'hybridation inverse, appelée INNO-LiPA (pour Line Probe Assay), a été décrite pour la première fois en 1993 (Stuyver *et al.* 1993). Son principe consiste à réaliser une amplification par PCR du gène étudié, puis une hybridation du produit de PCR à des amorces complémentaires spécifiques du génotype recherché fixées au préalable sur une bandelette de nitrocellulose (Figure 22) (Wang *et al.* 2008). L'utilisation d'amorces de PCR biotinylées permet la révélation de l'hybridation avec une phosphatase alcaline streptavidinée. Cette technique performante, simple à mettre en œuvre, ne permet cependant pas de détecter des mutations différentes de celles qui sont connues et insérées dans les amorces d'hybridation fixées aux bandelettes du test. Par ailleurs, s'il est possible de déterminer la présence d'un variant rare, il reste impossible de le quantifier puisque la technique est uniquement qualitative.

Le clonage-séquençage (CS) a longtemps été considéré comme la technique de référence, car il permet d'obtenir à la fois une bonne sensibilité, la quantification des populations virales et d'établir le lien entre les mutations. En pratique, la méthode consiste à réaliser une séparation physique des variants viraux par clonage (Figure 23), suivie d'un séquençage des clones par la méthode de Sanger. La sensibilité de cette méthode dépend du nombre de clones séquencés ; par exemple le séquençage de trente clones permet de détecter 95% des variants qui représentent au moins 10% de la quasi espèce (Keele *et al.* 2008). On peut donc

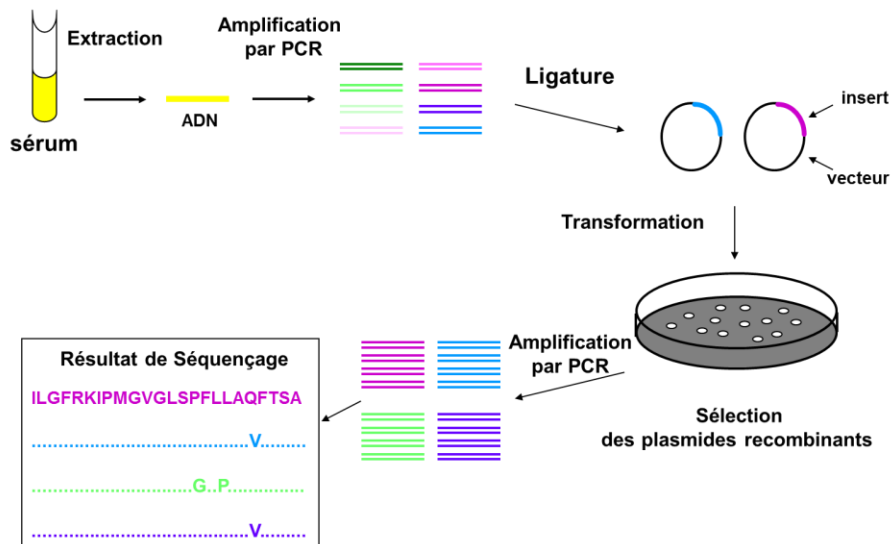


Figure 23 : Schéma du principe de clonage-séquençage (CS). D'abord, le gène étudié est amplifié par PCR, puis chaque amplicon est introduit de façon clonale dans un vecteur d'amplification. Après transformation bactérienne, les bactéries sont cultivées de manière à ce que chaque colonie contienne un seul clone d'amplicon. Les colonies sont ensuite lysées, le plasmide extrait puis séquencé indépendamment pour chaque colonie. Le nombre de colonies choisi dépend de la sensibilité requise pour l'expérimentation.

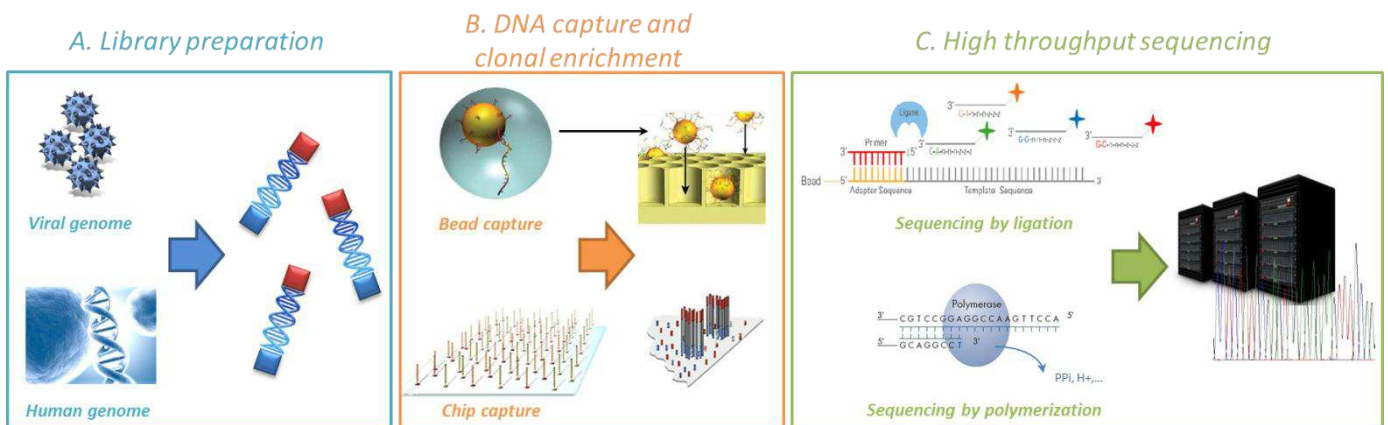


Figure 24 : Le séquençage de nouvelle génération (next-generation sequencing, NGS) fait appel à 3 étapes : A) la préparation des bibliothèques, B) la capture clonale de l'ADN, C) le séquençage (Metzker 2010 ; Chevaliez *et al.* 2012).

adapter le protocole à la sensibilité requise, en multipliant le nombre de clones étudiés pour l'améliorer. En pratique, cette technique est très consommatrice de temps et il est rare que des études ayant généré plus d'une trentaine de clones par échantillon soient réalisées.

Techniques de séquençage à haut-débit

Les techniques de séquençage à haut-débit, appelées également « next-generation sequencing » (NGS), ont été développées pour augmenter de façon très importante la quantité de séquences clonales obtenues en une seule expérimentation. Le processus se déroule en trois étapes, quelle que soit la technique employée (Figure 24) (Metzker 2010; Chevaliez *et al.* 2012). La première étape consiste à préparer des librairies d'ADN afin de formater l'ADN sous une forme lisible par les séquenceurs ; la deuxième étape consiste à séparer de manière clonale chaque fragment d'ADN puis à l'enrichir (sauf dans la technologie PacBio RS) ; la troisième étape est celle du séquençage proprement dit, qui utilise des techniques de polymérisation ou de ligation et une détection du signal fondée sur la détection de pyrophosphate (pyroséquençage), d'ion H⁺ (conductimétrie) ou de fluorescence. Le signal est ensuite converti en séquences qui doivent être analysées par des logiciels spécifiques afin d'être transformées en données biologiquement interprétables.

Les systèmes de séquençage à haut débit sont de deux types : (i) les systèmes à haute capacité, qui comprennent le Genome Analyzer et les séquenceurs HiSeq (Illumina Inc., San Diego, Californie) et les séquenceurs de série 5500 utilisant la technologie SOLiD (Applied Biosystems, Carlsbad, California) ; (ii) les systèmes générant de longues séquences, qui comprennent le Genome Sequencer (GS) FLX ou Junior (454 Life Sciences, Roche Diagnostics Corp., Branford, Connecticut), le système Ion Torrent (Applied Biosystems, Carlsbad, Californie) et le système PacBio RS (Pacific Biosciences, Menlo Park, Californie).

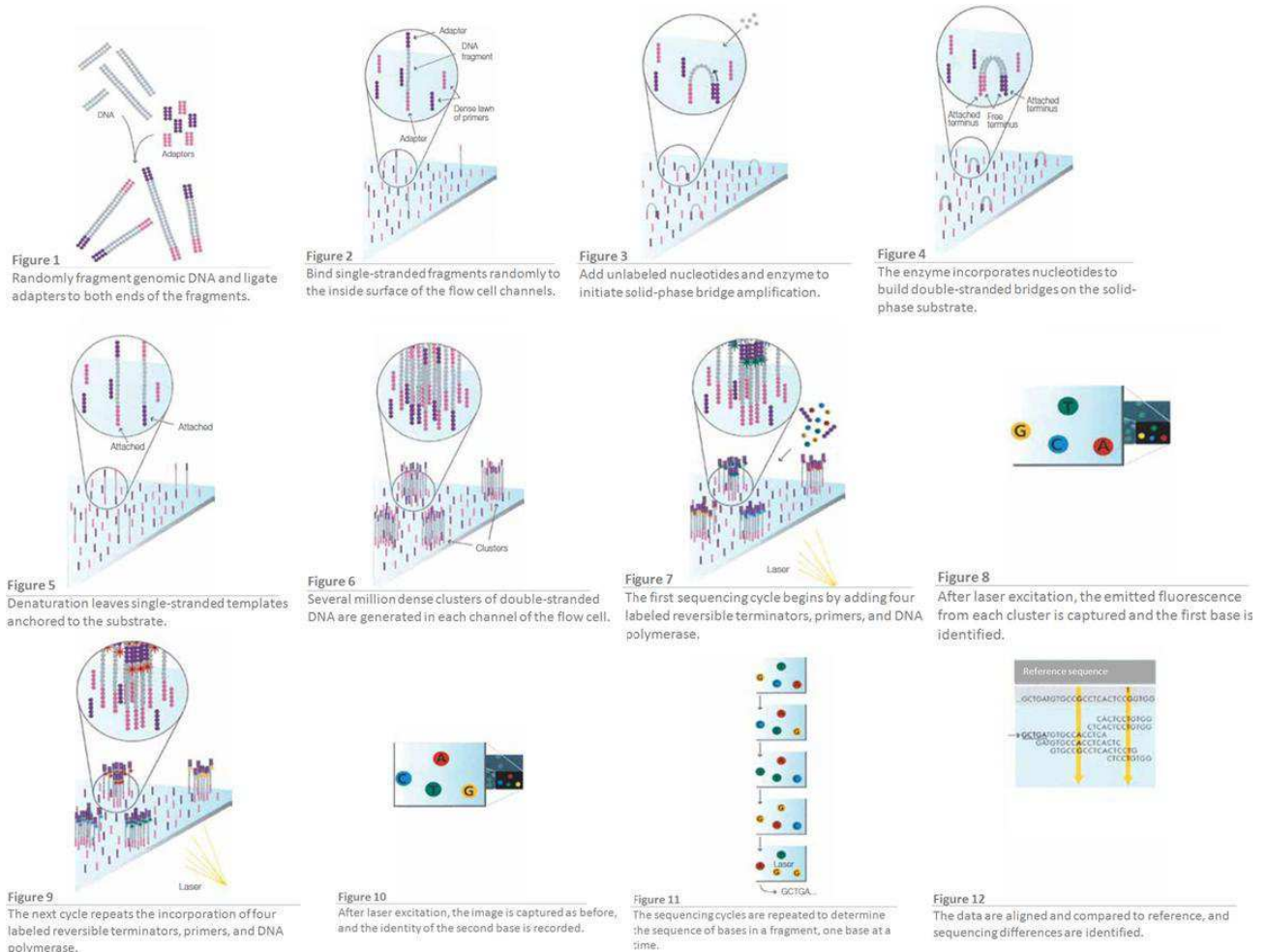


Figure 25 : Processus de séquençage par la technologie Illumina. Après fragmentation et marquage de l'ADN d'intérêt (figure 1), celui-ci est fixé sur la « flow cell » (figure 2) puis amplifié en clusters clonaux (figures 3 à 6). Le séquençage est ensuite réalisé par synthèse du brin complémentaire à l'aide de dNTPs fluorescents (figures 7 à 11) qui permet l'obtention d'un signal convertible en séquence (figure 12) (Source Illumina, Bennett 2004).

Les séquenceurs à très haut débit utilisent essentiellement deux technologies concurrentes : Illumina et SOLiD. La technologie Illumina est fondée sur la fragmentation de l'ADN d'intérêt puis son marquage par des amorces particulières (Figure 25). Des sondes complémentaires des amorces, fixées sur une « flow cell », sont ensuite utilisées pour fixer de manière clonale les fragments d'ADN devant être séquencés. Ensuite, une « bridge PCR » permet l'amplification locale qui génère des clusters d'ADN clonaux fixés sur la cellule. Les amplicons sont alors séquencés à l'aide d'une polymérase qui ajoute des dNTPs fluorescents, lus à chaque cycle (Bennett 2004). La technologie SOLiD est fondée sur une approche de PCR en émulsion qui permet l'enrichissement clonal de l'ADN préalablement fixé sur des billes (cf technologie Roche, Figure 26, proche). Les billes sont ensuite transférées sur une « flow chip » afin de réaliser le séquençage clonal par synthèse du brin complémentaire par ligature d'amorces fluorescentes (Pandey *et al.* 2008). Ces deux technologies permettent de générer des millions de séquences courtes (Tableau 2), facilement utilisables pour le reséquençage (séquence de référence disponible), comme par exemple celui du génome humain dans le cadre de projets de type GWAS (genome-wide association) ou RNA-Seq. Dans ce cadre, 100 à 150 pb sont suffisantes pour identifier un gène ou un ARNm connu. Ces technologies sont également bien adaptées à la détection de mutations de résistance aux traitements antiviraux, ce d'autant que la profondeur obtenue (nombre de séquence couvrant la portion de génome étudié) est souvent très élevée, assurant ainsi une très bonne sensibilité. Cependant, les courtes séquences ne permettent pas de faire le lien entre des mutations éloignées de plus de 100 à 150pb et ne permettent donc pas toujours de répondre aux questions cliniques liées à la résistance.

Le second groupe de méthodes est doté d'un débit plus faible, mais peut générer des séquences de beaucoup plus grande taille. La technologie en tête sur ce segment est celle

Manufacturer	Sequencing device	Technology (template preparation/ NGS chemistry)	Type	Number of single reads per run* (x 10 ⁶)	Number of nucleotides per run* (Gb)	Maximum sequence length* (bp)	Accuracy
Applied Biosystems	5500	emPCR/ligation	High throughput	800	9	75	99.6-99.8%
	5500xl		High throughput	1600	15	75	
	Ion Torrent (ChiP 316)	emPCR/real-time sequencing	Long reads	6.2	>1	>400	99.97%*
Illumina	MiSeq	Solid capture/reversible terminator	High throughput	3.4	>1	150	96.7-100%
	Genome Analyzer IIx		High throughput	320	95	150	
	HiSeq 1000		High throughput	1500	300	100	
	HiSeq 2000		High throughput	3000	600	100	
454 Life Sciences/Roche Diagnostics	GS Junior	emPCR/pyrosequencing	Long reads	0.1	0.035	400	99%*
	GS FLX+		Long reads	1	0.7	1000	97.4-99.9%
Pacific Biosciences/Gen-Probe	PacBio RS	Single molecule/real-time sequencing	Long reads	0.035	0.045	1200	99.99%*

Tableau 2 : Résumé des caractéristiques des différents séquenceurs commercialisés en 2012.

Les chiffres indiqués* ont été fournis par les fabricants (Chevaliez *et al.* 2012).

développée par Roche sur les séquenceurs GS. Les étapes nécessaires à l'obtention des séquences sont: 1) la préparation de la librairie par fragmentation ou, plus souvent, par amplification de la cible à séquencer, 2) la fixation clonale de l'amplicon marqué sur une bille qui sert de support pour la PCR en émulsion, 3) la PCR en émulsion qui consiste à enrichir de façon clonale la cible ADN à séquencer à l'intérieur de chaque gouttelette de l'émulsion, 4) le transfert des billes contenant l'ADN enrichi sur la « picotiter plate » (PTP) qui contient jusqu'à un million de puits (chaque puits ne pouvant contenir qu'une seule bille et donc un seul variant d'ADN), 5) la réaction de pyroséquençage (Figure 26) (Mardis 2008). Cette réaction débute par l'élongation du brin complémentaire en cours de séquençage par un nouveau nucléotide, qui largue en parallèle un pyrophosphate (PPi). Ce dernier permet la synthèse d'ATP en présence d'APS (adenosine 5' phosphosulfate) et d'ATP sulfurylase, induisant la génération de lumière en présence de luciférine et de luciférase (Ahmadian *et al.* 2000; Margulies *et al.* 2005). A la fin de l'année 2012, la technologie Ion Torrent (Applied Biosystems, Carlsbad, Californie) devrait être disponible. Techniquement, cette méthode est proche de la technologie Roche. Elle s'en différencie essentiellement par le support de type puce, qui rend l'évolution de la technique plus simple, et par la méthode de détection fondée sur la génération d'un proton à chaque ajout de base lors de la synthèse du brin complémentaire (Rothberg *et al.* 2011). Il s'agit d'une technologie séduisante, car la détection des variations du pH est moins coûteuse que les techniques décrites précédemment. La technique Pacbio RS est fondée sur des principes très différents. Après préparation de la librairie sans phase de pré-amplification, l'ADN est transféré sur un système optique appelé « zero-mode waveguide » (ZMW), qui a la particularité d'amplifier fortement les signaux de fluorescence émis. Ce système est constitué d'un puits contenant des milliers de polymérases ne pouvant accueillir chacune qu'une seule molécule d'ADN. La

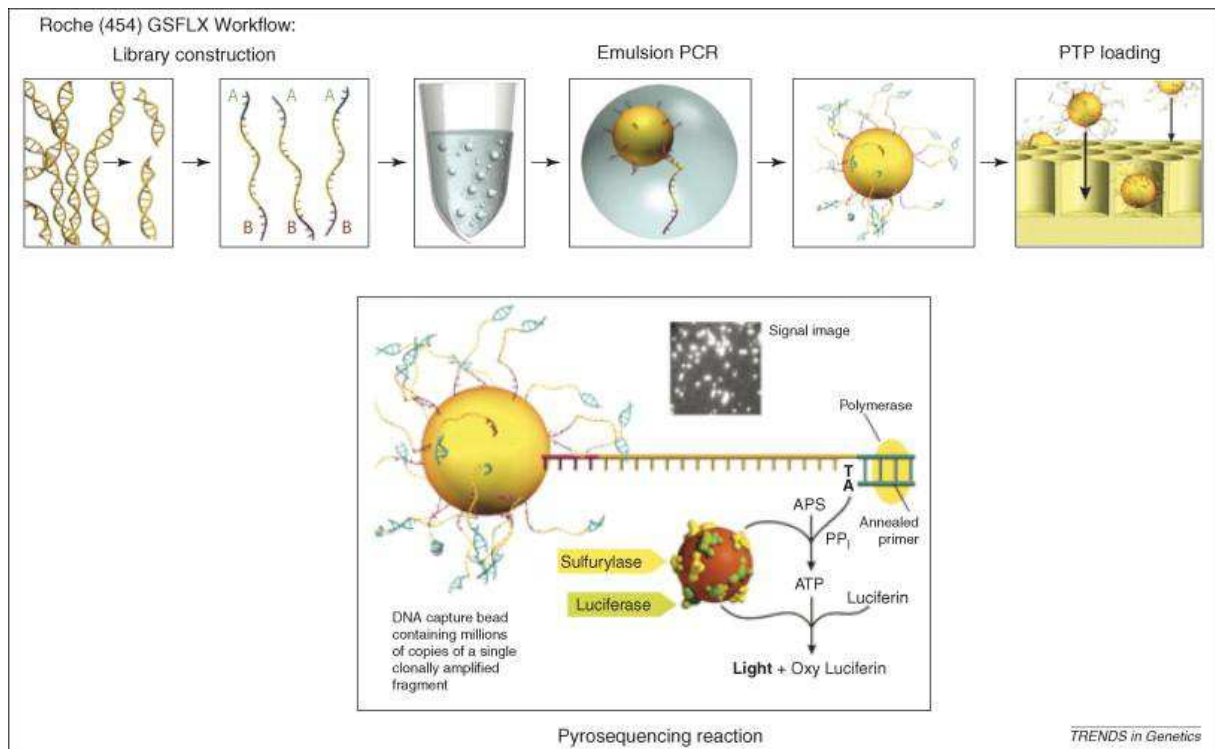
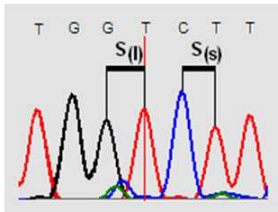


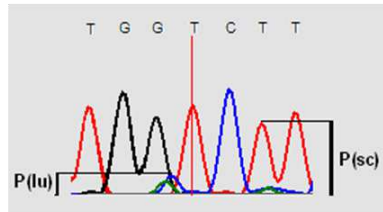
Figure 26 : Processus de séquençage par la technique 454/Roche. Une librairie est d'abord constituée par fragmentation ou par PCR. Les fragments sont ensuite clonés et les cibles devant être séquençées sont enrichies par PCR à émulsion à l'aide de billes. Les billes sont transférées sur la PTP (picotiter plate) afin de permettre la réaction de pyroséquençage. Cette réaction se déroule en 3 étapes : 1) ajout d'un nouveau nucléotide au brin complémentaire du brin à séquencer et relargage d'un pyrophosphate (PPi), 2) synthèse d'ATP à partir du PPi en présence d'APS (adenosine 5' phosphosulfate) et de sulfurylase, 3) génération de lumière en présence d'ATP et de luciférase (Mardis 2008).

réaction de séquençage est fondée sur la fluorescence, dont le signal est fortement amplifié par la ZMV, ce qui permet de travailler sur une seule molécule d'ADN par puits, donc sur une quantité de matériel génétique très faible (Eid *et al.* 2009). Cette méthode a la particularité de générer des séquences longues, un atout pour le séquençage *de novo* et en virologie du fait de la possibilité d'établir des liens entre les différentes mutations présentes sur les mêmes brins génomiques.

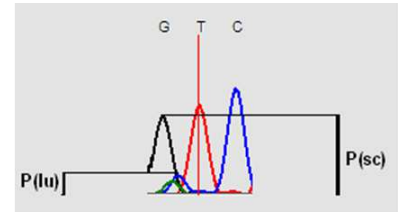
Toutes les méthodes de séquençage à haut débit sont susceptibles de générer des erreurs lors du séquençage. Les évaluations du Genome Analyzer, du 5500 et du GS FLX ont montré une précision de 96.7% à 100%, c'est-à-dire voisine de celle des méthodes classiques de séquençage (Tableau 2) (Harismendy *et al.* 2009). Cependant, une plus grande transparence des fournisseurs, une meilleure standardisation des études de performance et des études indépendantes sont aujourd'hui indispensables. Si l'on considère par exemple la qualité brute des séquences, évaluée à l'aide du score de PHRED qui répertorie le risque d'attribuer par erreur une base à une position donnée (Figure 27) (Ewing and Green 1998; Ewing *et al.* 1998; Richterich 1998), le plus grande production de séquences ininterprétables, appelées « junk data », est observée avec le system SOLiD (Shen *et al.* 2008; Harismendy *et al.* 2009; Walter *et al.* 2009; Suzuki *et al.* 2011). Eliminer les données ininterprétables est facile, mais la technologie SOLiD induit un surcoût liée à la production de nombreuses données inutilisables, sans altération majeure de la qualité des données finales. Si l'on considère les taux de faux positifs et de faux négatifs, c'est la technologie Illumina qui pose le plus de problèmes, par rapport aux technologies GS FLX et SOLiD. Des algorithmes correctifs, fondés sur des calculs de probabilité d'erreurs, peuvent corriger ces défauts et font converger les trois technologies vers des performances similaires si la profondeur d'analyse est suffisante (Harismendy *et al.* 2009; Suzuki *et al.* 2011).



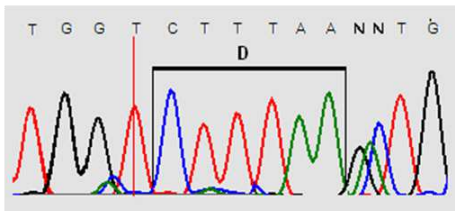
A. Pour chaque lot de 7 bases, le ratio entre le plus grand et le plus petit espace séparant 2 pics ($S(l)/S(s)$) est établi.



B. Pour chaque lot de 7 bases, le ratio entre la hauteur du plus grand pic $P_{(sc)}$ et le plus petit $P_{(lu)}$ est établi.



B.bis Pour chaque lot de 3 bases, le ratio entre la hauteur du plus grand pic $P_{(sc)}$ et le plus petit $P_{(lu)}$ est établi.



C. Enfin, le dernier paramètre, D, représentant la distance entre le pic étudié et le premier pic non résolu (N), est calculé.

$$Q = -10 \log \log_{10} P$$

D. L'ensemble des paramètres calculés en A/B ou Bbis/C sont utilisés pour établir la probabilité d'erreur sur la base considérée. Le score de PHRED pour chaque base est ensuite établi selon l'équation ci-dessus.

PHRED QUALITY SCORE	Probabilité que la base soit incorrect
10	10%
20	1%
30	0,1%
40	0,01%

E. Exemple de correspondance entre le score de PHRED et la probabilité d'attribuer par erreur une base à une position donnée

Figure 27 : Méthode de calcul du score de PHRED et exemples de correspondance entre le score de PHRED et la probabilité d'erreur sur une base (Ewing and Green 1998; Ewing et al. 1998).

Si l'on considère certains types de données, comme par exemple le séquençage d'organismes très riche en GC ou en AT, il a été montré que les performances étaient sensiblement moins bonnes avec les technologies Illumina et SOLiD (Dohm *et al.* 2008) ; alors que dans le cas de génomes possédant un fort taux de mutation de type insertion/délétion (« indel »), le système GS-FLX a montré une meilleure performance que ses concurrents (Harismendy *et al.* 2009; Hoppman-Chaney *et al.* 2010). Au contraire, cette méthode se révèle moins performante lorsque la séquence contient des homopolymères (Dohm *et al.* 2008; Gilles *et al.* 2011). Au total, les études tendent à montrer qu'il n'existe pas de technique idéale, dépourvue de défauts, mais plutôt des solutions adaptées à chaque type de projet, le choix de l'une ou de l'autre devant être mûrement réfléchi.

Analyse des données

Les outils disponibles pour l'analyse de données de séquençage direct sont adaptés à la pratique courante de non-informaticiens. Plusieurs étapes d'analyse sont nécessaires pour lesquelles de nombreux logiciels ont été développés, comme Chromas pour le basecalling et la correction de séquences, (McCarthy 1996), ClustalX pour l'alignement (Chenna *et al.* 2003) ou pour la phylogénie (Page 1996). Ces logiciels ont été créés pour une utilisation manuelle et l'analyse de quelques dizaines de séquences au maximum. Dans le cas du séquençage à haut débit, des millions de séquences sont générées, obligeant à une analyse des données la plus automatisée possible.

Le processus d'analyse de données de séquençage à haut débit se déroule en trois étapes. La première étape, dite de « basecalling », ne sera pas abordée ici car elle est gérée par des logiciels internes aux séquenceurs et peut être considérée comme fiable avec les technologies actuellement commercialisées. Cette étape consiste à convertir les signaux

déTECTÉS (pH, fluorescence, luminescence, etc) en séquences. Les données brutes obtenues à l'issue de cette première phase sont constituées de séquences au format FASTA, associées à des scores de qualité au format PHRED. Ce score, fourni pour chaque base séquencée, est la clé qui permet l'automatisation du processus d'analyse des données (Altshuler *et al.* 2000). La deuxième étape consiste à convertir automatiquement les données brutes en données analysables annotées, c'est-à-dire débarrassées des « junk data » et alignées. Cette étape inclut plusieurs processus, dont l'élimination des séquences ou portions de séquences dont le score de PHRED est inférieur à 20 (risque d'erreur de 1%), l'élimination des séquences trop courtes, et l'alignement. Cette dernière opération est critique, car les algorithmes d'alignement et les paramètres permettant sa mise en œuvre sont nombreux et elle est consommatrice d'une énorme puissance de calcul. Des progrès importants ont été réalisés afin d'optimiser le temps de calcul. Cependant, les paramètres de réglage varient selon les algorithmes d'alignement, d'une étude à l'autre (fréquence des indels, richesse en AT ou CG...), ou d'une technologie à l'autre (prise en charge des problèmes d'homopolymères par exemple) et sont donc difficile à standardiser, si bien que les résultats peuvent varier en fonction du logiciel utilisé. Il est cependant possible, pour une application donnée, par exemple l'étude de virus hautement variables, et pour une technologie donnée, de créer des algorithmes spécifiques utilisant des paramètres préétablis. De nombreuses équipes ont mis au point des processus spécialement dédiés à la gestion des séquences virales (Wang *et al.* 2007; Tsibris *et al.* 2009; Archer *et al.* 2010; Poon *et al.* 2010; Gianella *et al.* 2011; Rodriguez *et al.* 2011). Une fois les données « nettoyées » des junk data et alignées, le processus est complété par une nouvelle phase de gestion des erreurs. Il est en effet possible d'observer la présence d'erreurs aléatoires, associées à un bon score de qualité PHRED. Deux types d'approches existent aujourd'hui pour la correction de ces erreurs : (i) un calcul statistique

qui compare la fréquence observée à celle obtenue selon l'hypothèse d'une présence aléatoire fondée sur la loi binomiale puis détermine si la différence est significative ou non (Macalalad *et al.* 2012; Rodriguez *et al.* 2012), (ii) l'exploration de l'environnement génétique de la base fondé sur l'hypothèse selon laquelle une vraie mutation proviendra plus vraisemblablement de variants apparentés (notion d'haplotype), contrairement à une erreur qui survient de façon aléatoire sur n'importe quel fond génétique (Zagordi *et al.* 2010; Zagordi *et al.* 2010). La troisième étape consiste à rendre ces millions de données de séquences interprétables d'un point de vue biologique. Cette nécessité implique d'adapter le logiciel à la question scientifique posée. Il peut donc exister autant de logiciels que de projets. Pour le domaine qui concerne cette Thèse, on peut citer les logiciels d'interprétation des mutations de résistance comme geno2pheno (Beerenwinkel *et al.* 2003; Altmann *et al.* 2009; Seclen *et al.* 2011; Nozza *et al.* 2012; Santoro *et al.* 2012; Thielen and Lengauer 2012) et PyroPack[®] que nous avons développé (Chevaliez *et al.* 2011; Rodriguez *et al.* 2011; Rodriguez *et al.* 2012).

PARTIE 4 : Prise en charge clinique de la résistance

Prise en charge initiale et prévention du risque

Pré-existence des variants viraux résistants

Comme il a été décrit précédemment, la distance devant être parcourue pour atteindre l'état de résistance peut-être raccourcie lorsque des variants viraux résistants sont présents avant la mise en place de la thérapeutique antivirale. Les variants présents en proportion majoritaire ou intermédiaire sont le plus souvent transmis par un patient source traité en échec thérapeutique. Les variants peuvent également être présents à l'état minoritaire, dans le cadre de la génération spontanée de mutations sur le génome viral.

La résistance aux traitements de première génération du VHB, tels que la lamivudine ou l'adefovir, est un événement fréquent. En revanche, la résistance est exceptionnelle avec l'entecavir lorsqu'il est administré en première intention et n'a pas été rapportée avec le tenofovir. Dans ce contexte, même si l'impact des mutations de résistance à la lamivudine ou à l'adefovir sur les nouvelles générations d'antiviraux ne semble pas négligeable en terme de retard à la négativation de la charge virale par exemple, l'absence de connaissance précise de la prévalence des variants viraux résistants aux premières générations de traitement chez les patients naïfs, conjuguée au choix limité d'alternatives thérapeutiques, rend l'intérêt d'un dépistage systématique hasardeux. En conséquence, il n'est pas recommandé pour le moment de rechercher la présence de variants viraux résistants chez un nouveau patient avant la mise en place d'un traitement anti-VHB (EASL 2009).

Les patients infectés par le VHC bénéficient de l'autorisation de mise sur le marché d'inhibiteurs de protéase depuis un peu plus d'un an (AMM 19/09/2011 pour le telaprevir et 18/07/2011 pour le boceprevir, en combinaison à l'interféron alpha pégylé et à la ribavirine pour les malades infectés par un VHC de génotype 1, source ANSM). La prévalence des mutations de résistance transmises chez les patients naïfs de traitement ne constitue pas pour le moment une préoccupation de premier plan. La grande majorité, si ce n'est tous les malades, abritent naturellement des variants viraux potentiellement résistants aux inhibiteurs de protéase sans avoir jamais été exposés à ces molécules. L'impact de ces mutations n'est pas aujourd'hui démontré, l'échec thérapeutique étant essentiellement lié à une mauvaise réponse à l'action antivirale de l'interféron dans le cas des triples combinaisons, si bien qu'il n'est pas recommandé d'en établir le diagnostic avant la mise en place d'un traitement incluant un inhibiteur de protéase (EASL 2011).

Dans le cas des patients séropositif pour le VIH, l'épidémiologie est bien connue et la proportion de patients infectés par une souche résistante aux antirétroviraux détectable par séquençage direct est comprise entre 10 et 20% dans les pays occidentaux (Wensing *et al.* 2005; Shet *et al.* 2006). La présence de ces populations virales a pour effet de retarder la négativation de la charge virale et de multiplier par cinq le risque d'échec virologique (Yerly *et al.* 1999; Grant *et al.* 2002; Little *et al.* 2002; Kuritzkes *et al.* 2008). Par conséquent, les recommandations de pratique clinique soulignent toutes l'importance de tester la résistance chez tous les patients nouvellement pris en charge (Hirsch *et al.* 2008; Paredes and Clotet 2010; Thompson *et al.* 2010).

L'implication des variants minoritaires (<1%) dans l'échec virologique est incertaine dans le cas de l'infection par le VHC, lorsqu'un DAA est associé à l'interféron et à la ribavirine. Elle joue un rôle déterminant dans la résistance du VHB et VIH aux traitements antiviraux ciblant

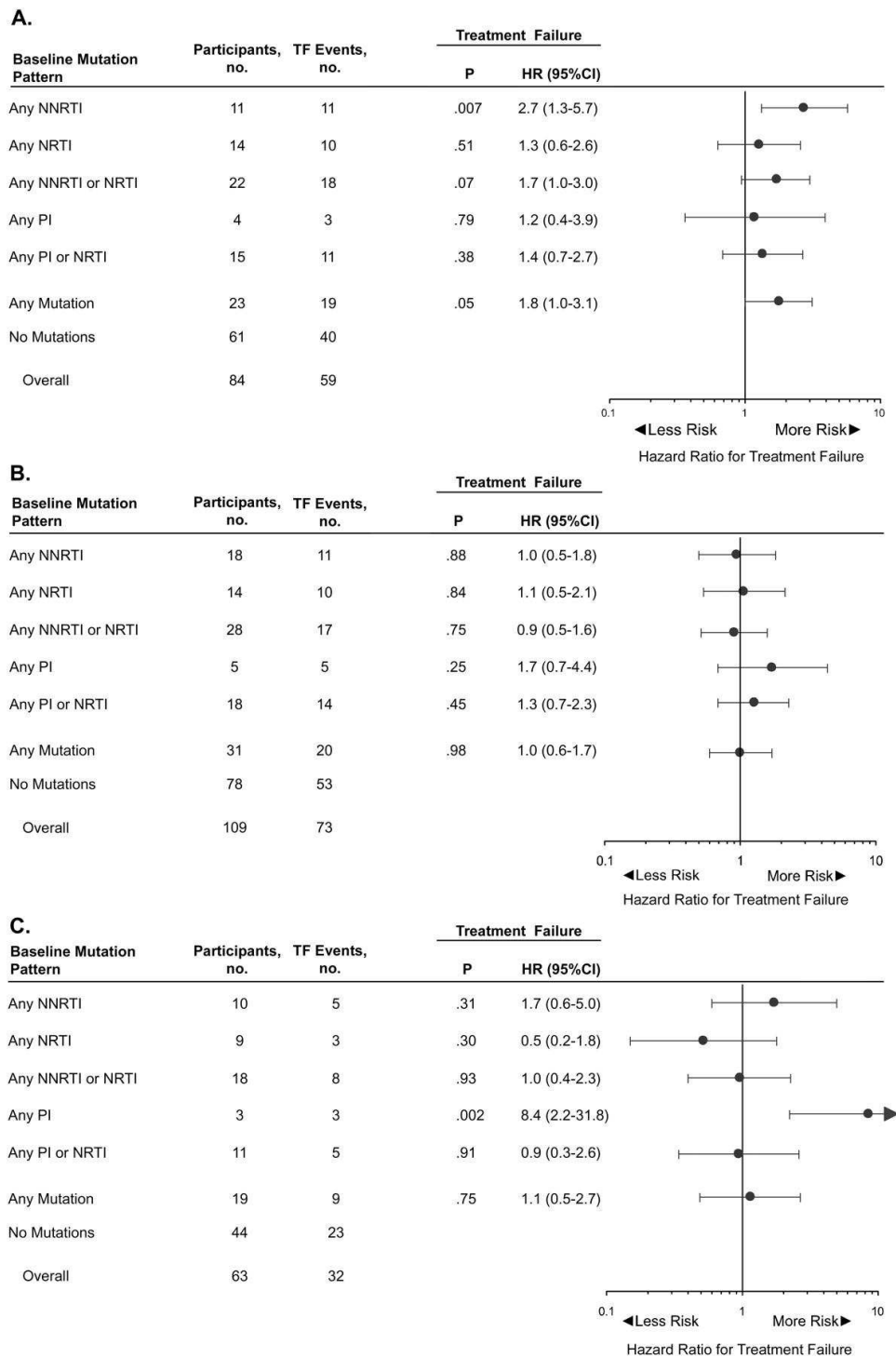


Figure 28 : Risque d'échec virologique par trithérapie antvirale (A : INTI+INNTI, B : INTI+IP, C : INTI+INNTI+IP) associé à la présence de variants résistants minoritaires avant le début du traitement (Simen *et al.* 2009).

directement les fonctions virales. Dans le cas du VIH, il a été clairement montré que la présence de variants viraux résistants était associée à un risque plus élevé d'échec virologique (Figure 28) (Palmer *et al.* 2006; Svarovskaia *et al.* 2007; Johnson *et al.* 2008; Simen *et al.* 2009). Ainsi, le risque d'échec virologique est multiplié par 2,7 chez les patients traités par deux analogues nucléos(t)idiques et un inhibiteur non nucléosidique chez qui des mutations de résistance à ces molécules sont détectables au début du traitement. Il a également été montré que la présence de variants CXCR4 minoritaires à 2% était associée à des échecs thérapeutiques plus fréquents sous maraviroc (Archer *et al.* 2010). Cependant, le manque de consensus concernant les approches techniques, l'analyse des données et la démarche à entreprendre en fonction des résultats ne permet pas actuellement de dégager de stratégie univoque pour prévenir la résistance lorsque les variants sont présents à un faible niveau.

Récemment, l'introduction d'une phase de « lead-in » précédant l'introduction de l'inhibiteur spécifique a été proposée pour réduire l'incidence de la résistance au boceprevir (inhibiteur de protéase du VHC). La lead-in consiste à débiter le traitement par interféron pégylé et ribavirine d'abord, afin d'abaisser la charge virale à un niveau qui permet de réduire le risque de sélection de variants résistants à l'inhibiteur, dans l'hypothèse où la réplication des variants résistants pré-existants serait contrôlée (Susser *et al.* 2009; Bacon *et al.* 2011; Poordad *et al.* 2011). Cette approche ne s'est finalement pas révélée efficace pour réduire le risque d'échec thérapeutique associé à l'émergence de la résistance (Fried *et al.* 2011). Elle permet tout au plus de prédire la réponse au traitement, qui dépend de l'efficacité antivirale de la combinaison interféron-ribavirine, et d'éventuellement décider de ne pas introduire l'inhibiteur de protéase si la réponse à l'interféron est nulle.

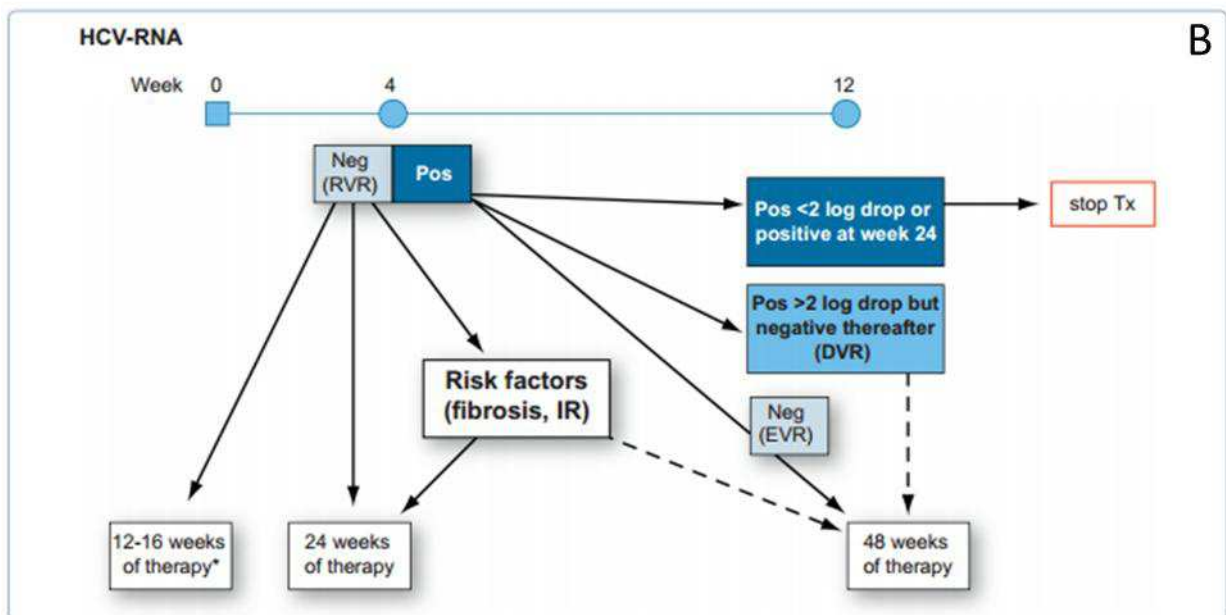
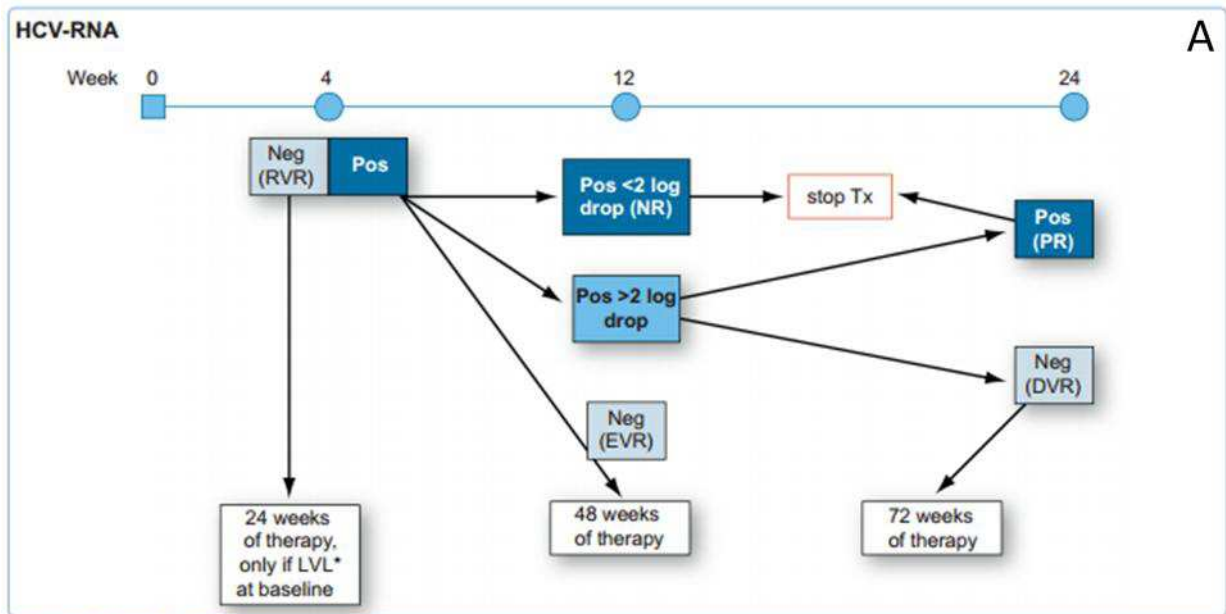


Figure 29 : Algorithme décisionnel permettant de prévenir la sélection de variants viraux résistants aux inhibiteurs de protéase du VHC dans un contexte d'échec thérapeutique chez les patients infectés par un VHC de génotype 1 ou 4 (A), ou 2 ou 3 (B), traités par trithérapie associant un inhibiteur de protéase, l'interféron alpha péglylé et la ribavirine (EASL 2011).

Dynamique des variants viraux résistants au cours du traitement antiviral

L'étude de la résistance au cours du traitement peut permettre d'éviter l'échec virologique lié à celle-ci. Pour tous les virus, le dépistage repose sur la quantification de la charge virale à une fréquence qui est établie par pathologie et par type de traitement. Par exemple, chez les patients infectés par le VHC traités par un inhibiteur de protéase en combinaison à l'interféron alpha pégylé et à la ribavirine, les contrôles de charge virale doivent être effectués avant le traitement, et aux semaines 4, 12 et 24 du traitement par l'inhibiteur. Si la présence de l'ARN viral est établie après la mise en place du traitement par inhibiteur de protéase, la poursuite du traitement doit être discutée selon un algorithme consensuel (Figure 29), afin d'éviter la sélection de variants résistants, conséquence de l'échec thérapeutique. En effet, la présence de tels variants en proportions importantes pourrait entamer les chances de guérison lors de futurs traitements fondés sur des médicaments appartenant à la même classe thérapeutique (Figure 29) (EASL 2011). Chez les patients séropositifs pour le VIH traités par trithérapie, les mesures de charge virale sont effectuées tous les quatre à six mois si le taux de CD4 est $>500/\text{mm}^3$, tous les trois mois dans les autres cas et ce, quel que soit le traitement antiviral instauré (Hirsch *et al.* 2008; Thompson *et al.* 2010). Chez les patients infectés par le VHB, la charge virale doit être mesurée toutes les douze semaines (EASL 2009).

Lorsque la charge virale augmente de plus de 1 log au-dessus de son nadir sous traitement, le rebond virologique doit être exploré et documenté. Pour cela, il convient de s'assurer que le patient respecte la prescription, de vérifier les éventuelles interactions médicamenteuses et l'influence sur la compliance des effets indésirables. Dans le cas du VIH, un dosage de la concentration plasmatique des antiviraux est nécessaire avant d'attribuer l'échec du traitement à la résistance. Par conséquent, le dosage plasmatique des médicaments et

l'analyse génotypique de la résistance font partie des recommandations de pratique clinique en cas de remontée de la charge virale observée sous traitement (Yeni 2010). Dans le cas du VHC, l'arrêt du traitement est recommandé en cas d'échec thérapeutique pour éviter effets secondaires et coûts inutiles (règle de futilité). Dans le cas du VHB, l'échec thérapeutique associé à la résistance aux analogues nucléos(t)idiques doit conduire à l'ajout d'un deuxième antiviral sans résistance croisée avec le premier, sans nécessité de documenter les causes pharmacologiques ou génotypique de l'échec. Pour le VHB comme pour le VHC, aucune modification de la prise en charge thérapeutique ne dépend des résultats de l'analyse génotypique de la résistance. Cependant, ces examens sont réalisés à des fins de recherche dans les études cliniques thérapeutiques ou de cohorte incluant des malades infectés par le VHB ou le VHC. Les tests génotypiques de résistance, en général fondés sur le séquençage direct, permettent d'établir une cartographie des mutations de résistance qui est comparée à des abaques (Annexe 1A pour le VHC, Annexe 1B pour le VHB, Annexe 1C pour le VIH) (Mohanty *et al.* 2006; Paredes and Clotet 2010; Zeuzem and Foster 2012) ou directement interprétée à l'aide d'algorithmes en ligne tel que celui de l'ANRS (<http://www.hivfrenchresistance.org/>), celui de l'université de Stanford (<http://hivdb.stanford.edu/>) ou geno2pheno pour le VIH, ou les banques de séquences du VHB et du VHC (<http://hcv.bioinf.mpi-inf.mpg.de/index.php> ; <http://hbv.bioinf.mpi-inf.mpg.de/>). Les résultats sont ensuite transmis au clinicien, en général accompagnés de recommandations pour le choix d'un nouveau traitement.

OBJECTIFS

Nous avons vu que la prévention de la résistance aux antiviraux pouvait reposer sur la prédiction de son apparition à l'aide de différents marqueurs. L'identification de nouveaux marqueurs et l'évaluation des marqueurs existants sont indispensables à la meilleure définition des algorithmes prédictifs et à leur implémentation dans le processus décisionnel en pratique clinique. Dans ce contexte, il est particulièrement important de parfaitement connaître le contexte épidémiologique des marqueurs étudiés afin d'évaluer précisément l'intérêt d'un dépistage ou d'un suivi systématiques.

Les objectifs de ce travail ont été :

- 1/ De développer de nouveaux outils informatiques d'analyse de la variabilité génétique virale statique et dynamique, fondés sur le séquençage à haut débit de nouvelle génération (pyroséquençage ultra-profond) ;
- 2/ D'utiliser les approches classiques et ces outils pour caractériser la variabilité génétique du VHB, du VHC et du VIH et comprendre l'influence de celle-ci sur les caractéristiques cliniques et la résistance aux antiviraux dans différentes circonstances cliniques ;
- 3/ D'utiliser ces nouveaux outils pour analyser la dynamique temporelle des populations virales sensibles et résistantes au cours des traitements antiviraux ;
- 4/ D'utiliser les données récoltées pour améliorer la prise en charge thérapeutique avec les molécules existantes (par exemple les inhibiteurs d'entrée du VIH) ou développer de nouvelles approches thérapeutiques (par exemple par la découverte d'inhibiteurs de la RNase H du VHB).

Les résultats présentés ici s'appuient sur 6 études, dont l'analyse a atteint des niveaux différents d'achèvement (4 articles originaux sont présentés, la rédaction des deux articles correspondant aux dernières études n'ayant pas débuté). Quatre études sont consacrées au VHB. L'étude HBV Pol a porté sur des malades infectés par le VHB traités par adefovir et ayant présenté un échec à ce traitement lié au développement d'une résistance virale ; l'étude HBV InVS a inclus un grand nombre de patients infectés par le VHB vus pour la première fois dans les centres de référence français au cours de la période 2008-2011 ; l'étude HBV RNase H a porté sur un groupe homogène de patients grecs infectés par un VHB de génotype D ; enfin, l'étude HBV PreC a porté sur des enfants antigène HBe-positifs infectés par le VHB traités par adefovir. L'étude HCV Prove2 a porté sur les malades de l'hôpital Henri Mondor inclus dans l'étude de Phase II PROVE2 ayant testé l'efficacité et la tolérance de la triple combinaison d'interféron alpha pégylé, de ribavirine et de telaprevir chez des sujets naïfs de tout traitement infectés par un VHC de génotype 1. L'étude HIV génotropisme a porté sur des malades infectés par le VIH en échec de leur trithérapie et traités par le maraviroc.

Patient	Sexe/ Age	Origine ethnique	Mode de transmission	Histopathologie (score d'Ishak)					Ag HBe / Ac Anti-HBe	ALAT	Charge virale de base (UI/mL)	Durée de suivi/nb d'échantillons (mois/nb)
				grade			stade					
				a	b	c	d	e				
A	H/46	caucasien	Soin dentaire	4	0	3	2	2	-/+	170	57,351,946	56/19
B	H/59	caucasien	Inconnu	2	0	1	2	3	-/+	166	6,123,596	46/24
C	F/51	caucasien	Inconnu	4	0	4	2	4	-/+	52	41,531,750	48/15
D	H/55	caucasien	Transmission mère- enfant	4	0	4	3	5	-/+	239	3,604,304	24/18
E	H/43	caucasien	Transmission mère- enfant	2	0	2	2	3	-/+	43	1,337,363	46/16
F	F/56	caucasien	Transfusion sanguine	3	0	3	3	3	-/+	88	1,692,250	57/17
G	H/26	caucasien	Inconnu	4	0	2	4	6	-/+	99	283,804	50/19

Tableau 3 : Caractéristiques du groupe de patients étudiés dans l'étude HBV POL

PATIENTS

Etude HBV POL

L'étude a porté sur un groupe de 7 patients (2 femmes et 5 hommes) d'origine caucasienne, âgés de 26 à 59 ans au début de l'étude. Les 7 patients étaient séronégatifs pour le VHC, le virus de l'hépatite delta et le VIH. Tous étaient infectés par un VHB mutant precore (Ag HBe négatif) et atteints d'hépatite chronique, traités par adefovir en monothérapie et avaient présenté un échec thérapeutique lié au développement d'une résistance génotypique à ce médicament. La charge virale avant traitement était comprise entre 5,5 et 7,8 log UI/mL. La sévérité des lésions hépatiques a été caractérisée par le score d'Ishak (Tableau 3). Les patients, naïfs de tous traitements, ont reçu de l'adefovir en monothérapie à la dose de 10mg/j. Les patients A, B, E et G ont également été traités par lamivudine associée à l'adefovir après l'échec du traitement par l'adefovir seul. Des échantillons de sérum ont été prélevés régulièrement avant et au cours du traitement pour l'analyse des quasi-espèces virales. La durée de suivi était de 24 à 57 mois selon les patients.

Le dosage de l'alanine amino-transférase (ALAT) a été réalisé avant et au cours du traitement par adefovir à l'aide de réactifs et de l'analyseur Roche Modular Chemistry Analyser (Roche, Bâle, Suisse). Les charges virales ont été mesurées avant et pendant le traitement par les techniques de PCR Cobas Amplicor® HBV Monitor test ou PCR Cobas TaqMan 48 HBV test® (Roche, Bâle, Suisse), avec des limites inférieures de détection respectives de 36 UI/mL et 17 UI/mL. Tous les résultats ont été exprimés en UI/mL.

AgHBe-négatif [n(%)]	550(88,0%)
<hr/>	
ADN du VHB (Log ₁₀ UI/mL) (moyenne±DS)	3,66±1,42
ADN du VHB ≥3,30 Log ₁₀ UI/mL [n(%)]	226 (36,2%)
ADN du VHB <3,30 Log ₁₀ UI/mL [n(%)]	399 (63,8%)
ADN VHB indétectable (<1,20 Log ₁₀) [n(%)]	105 (16,8%)
<hr/>	
Titre de l'AgHBs (Log ₁₀ UI/mL) (moyenne±DS)	3,38±0,78
<hr/>	
Génotype du VHB [n(%)]	
Déterminé	459 (73,4%)
Non amplifié	166 (26,6%)

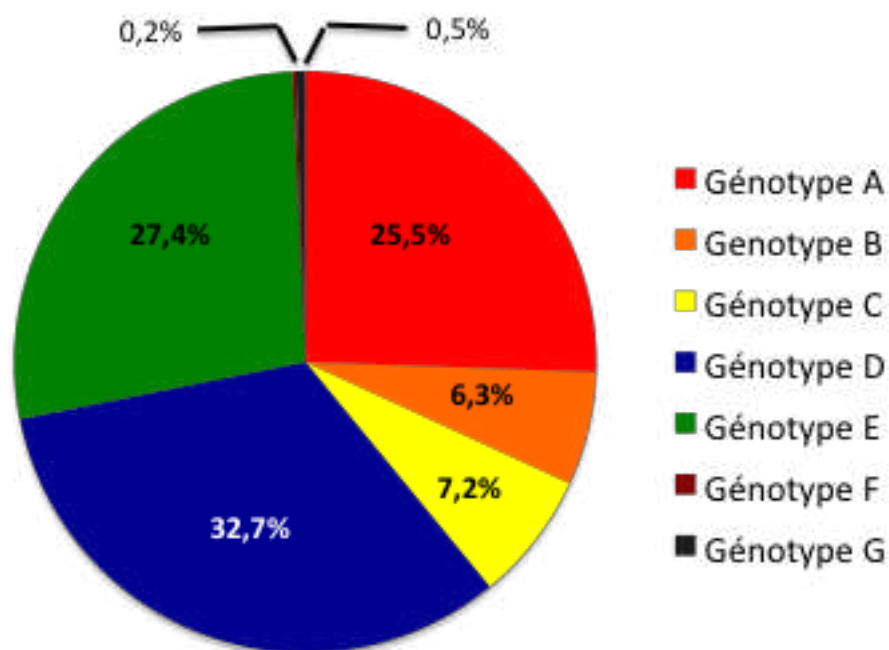


Tableau 4 : Caractéristiques des patients de la cohorte HBV InVS

Etude HBV InVS

Les patients de cette étude étaient issus de la surveillance systématique des patients atteints d'hépatite chronique B nouvellement pris en charge au sein des Pôles de Référence et des Réseaux Hépatites français. Au cours de la période 2008-2011, 2845 patients (59% d'hommes, d'âge moyen 36,0 ans, dont 19% nés dans un pays de faible endémicité et 81% dans un pays de forte ou moyenne endémicité) ont été inclus dans l'enquête de surveillance. Le volet virologique, réalisé par le Centre National de Référence des Hépatites Virales B, C et Delta, a porté sur 625 patients dont les prélèvements ont été collectés par les différents laboratoires de virologie participant à cet observatoire. Les caractéristiques virologiques des patients sont présentées dans le Tableau 4. La grande majorité d'entre eux (88,0%) étaient infectés par un virus mutant de la région pré-C/C (AgHBe-négatif). Environ 64% des patients avaient un ADN du VHB inférieur à 3,3 Log UI/mL (<2000 UI/mL). Les génotypes D (32,7%), E (27,4%) et A (25,5%) étaient les plus fréquents. Le titre de l'AgHBs, mesuré par une méthode de chimioluminescence automatisée (Architect, Abbott), était en moyenne de $3,4 \pm 0,8$ Log UI/mL. Les titres d'AgHBs étaient comparables quel que soit le génotype viral.

Etude HBV RNase H

579 patients infectés par un VHB de génotype D, naïfs de tout traitement et exempts de coinfection par le VHC ou le VIH, avec une hépatite chronique B Ag HBe négative, ont été recrutés dans cette étude épidémiologique menée dans le Nord de la Grèce. Cette cohorte comportait 374 hommes (64,4%) et 206 femmes (35,6%) ayant pour moyenne d'âge 52 ± 13 ans. La sévérité de leur maladie hépatique a été évaluée à l'aide du score Metavir, obtenu à partir de biopsies d'un échantillon représentatif de 147 patients : F0-F1 : 32,6%, F2-F3 :

Baseline Characteristics	2–6 Years ^a		7–11 Years ^a		12–17 Years ^a		Total	
	ADV (n = 23)	PLB (n = 12)	ADV (n = 36)	PLB (n = 19)	ADV (n = 56)	PLB (n = 27)	ADV (n = 115)	PLB (n = 58)
Sex (n, %)								
Female	14 (61)	8 (67)	13 (36)	4 (21)	14 (25)	7 (26)	41 (36)	19 (33)
Male	9 (39)	4 (33)	23 (64)	15 (79)	42 (75)	20 (74)	74 (64)	39 (67)
Age (years)								
Mean (SD)	4.1 (1.5)	4.6 (1.2)	9.5 (1.5)	9.8 (1.1)	14.5 (1.7)	14.1 (1.5)	10.8 (4.3)	10.7 (3.9)
Median	4.0	5.0	10.0	10.0	14.0	14.0	11.0	11.0
Race (n, %)								
White†	6 (26)	3 (25)	23 (64)	17 (90)	41 (73)	21 (78)	70 (61)	41 (71)
Asian	8 (35)	6 (50)	8 (22)	1 (5)	13 (23)	5 (19)	29 (25)	12 (21)
Black	7 (30)	1 (8)	3 (8)	1 (5)	1 (2)	1 (4)	11 (10)	3 (5)
American Indian or Alaska Native	0	0	0	0	1 (2)	0	1 (<1)	0
Other‡	2 (9)	2 (17)	2 (6)	0	0	0	4 (4)	2 (3)
Ethnicity (n, %)								
Hispanic/Latino	0	0	0	0	1 (2)	0	1 (<1)	0
BMI z-score								
Mean (SD)	-0.6 (0.34)	-0.6 (0.61)	-0.4 (0.67)	-0.3 (0.82)	0.3 (1.00)	0.2 (1.21)	0.0 (0.95)	0.0 (1.10)
Prior CHB treatment (n, %)								
Yes	2 (9)	2 (17)	24 (67)	13 (68)	38 (68)	18 (67)	64 (56)	33 (57)
No	21 (91)	10 (83)	12 (33)	6 (32)	18 (32)	9 (33)	51 (44)	25 (43)
HBV DNA (log ₁₀ copies/mL)								
Mean (SD)	9.23 (0.589)	9.01 (1.245)	8.63 (0.735)	8.52 (0.653)	8.60 (1.024)	8.63 (1.114)	8.74 (0.894)	8.67 (1.016)
Median	9.17	9.13	8.72	8.61	8.85	8.69	8.84	8.78
Min, max	8.27, 10.30	5.44, 10.25	6.72, 10.17	7.22, 9.40	5.20, 9.91	5.40, 10.26	5.20, 10.30	5.40, 10.26
HBV Genotype (n, %)								
A	5 (22)	1 (8)	17 (47)	13 (68)	29 (52)	18 (67)	51 (44)	32 (55)
B	4 (17)	3 (25)	3 (8)	1 (5)	6 (11)	1 (4)	13 (11)	5 (9)
C	1 (4)	1 (8)	4 (11)	0	5 (9)	3 (11)	10 (9)	4 (7)
D	11 (48)	5 (42)	9 (25)	4 (21)	15 (27)	5 (19)	35 (30)	14 (24)
E	2 (9)	1 (8)	1 (3)	1 (5)	0	0	3 (3)	2 (3)
F	0	0	1 (3)	0	1 (2%)	0	2 (2)	0
Not done	0	1 (8)	1 (3)	0	0	0	1 (<1)	1 (2)
ALT (U/L)								
Mean (SD)	90 (54.4)	105 (64.8)	105 (60.9)	81 (25.3)	123 (99.5)	109 (59.4)	111 (81.6)	99 (52.8)
Median	72	89	84	78	94	99	84	85
Min, max	33, 227	49, 265	23, 249	43, 130	24, 448	31, 250	23, 448	31, 265
ALT as multiple of ULN								
Mean (SD)	2.6 (1.60)	3.1 (1.91)	2.8 (1.67)	2.1 (0.77)	3.0 (2.39)	2.7 (1.44)	2.9 (2.03)	2.6 (1.40)
Median	2.1	2.6	2.3	2.0	2.3	2.3	2.3	2.2
Min, max	1.0, 6.7	1.4, 7.8	0.7, 6.0	1.1, 3.8	0.7, 10.4	0.7, 5.8	0.7, 10.4	0.7, 7.8
HBeAg † (n, %)								
Negative	0	1 (8)	0	0	2 (4)	0	2 (2)	1 (2)
Positive	23 (100)	11 (92)	36 (100)	19 (100)	54 (96)	27 (100)	113 (98)	57 (98)
Anti-HBe † (n, %)								
Positive	0	1 (8)	0	0	2 (4)	0	2 (2)	1 (2)

Tableau 5 : Caractéristiques générales des patients de l'étude HBV PreC (Jonas *et al.* 2008).

42,2%, F4 : 25,2% pour la fibrose, A0-A1 : 51%, A2-A3 : 49% pour l'activité. La charge virale a été mesurée à l'aide du kit Versant HBV DNA 3.0 (Bayer Healthcare, Leverkusen, Allemagne) selon les recommandations du fournisseur et était de 4.2 ± 2.5 Log UI/mL en moyenne. Parmi les patients de cette cohorte, un échantillon représentatif de 80 individus a été sélectionné de façon aléatoire pour la détermination des caractéristiques génotypiques de la RNaseH. Cet échantillon présentait des caractéristiques similaires à celles de la cohorte principale.

Etude HBV PreC

L'étude HBV PreC a porté sur des patients mineurs ayant une infection chronique par un VHB Ag HBe positif inclus dans un essai clinique visant à étudier l'efficacité d'un traitement par adefovir dans cette population (Jonas *et al.* 2008). Il s'agissait de 173 patients répartis en 3 groupes d'âge : 2-6 ans, 7-11 ans et 12-17 ans (Tableau 5), majoritairement infectés par un VHB de génotype A ou D et exempts de coinfection par le VHC ou le VIH. Les patients de chaque classe d'âge ont été randomisés en deux groupes, l'un recevant l'adefovir et l'autre un placebo (2:1) pendant 48 semaines. A l'issue de 44 semaines de traitement, tous les patients n'ayant pas seroconverti dans le système HBe se sont vus proposer la poursuite ou la mise en route (sujets sous placebo) du traitement par adefovir. Au cours du traitement et pendant le suivi (jusqu'à 180 semaines au total pour certains patients), la charge virale a été quantifiée régulièrement à l'aide du kit COBAS TaqMan HBV (Roche, Bâle, Suisse), l'Ag HBe et les anticorps anti-HBe ont été recherchés respectivement par le kit ETI-EBK-PLUS EIA et ETI-AB-EBK-PLUS EIA (DiaSorin, Saluggia, Italie), et l'Ag HBs et les Ac anti-HBs ont été recherchés respectivement par les kits AxSYM HBsAg (V2) et AxSYM AUSAB (Abbott, Chicago, USA). Un sous-groupe de 158 patients, représentatifs de la cohorte complète, a été

Characteristic	T12PR24 (N=81)	T12PR12 (N=82)	T12P12 (N=78)	PR48 (N=82)
Years of age — median (range)	46 (19–65)	44 (22–65)	45 (20–64)	45 (18–64)
BMI — median (range)†	24 (17–35)	23 (17–32)	24 (18–41)	24 (17–35)
Male sex — no. (%)	54 (67)	49 (60)	43 (55)	46 (56)
Race or ethnic group — no. (%)‡				
White	75 (93)	76 (93)	77 (99)	76 (93)
Black	1 (1)	2 (2)	1 (1)	2 (2)
Asian	3 (4)	2 (2)	0	4 (5)
Hispanic	1 (1)	1 (1)	0	0
Other	1 (1)	1 (1)	0	0
HCV RNA — log ₁₀ IU/ml	6.5±0.6	6.4±0.6	6.3±0.6	6.4±0.6
HCV RNA ≥800,000 IU/ml — no. (%)	73 (90)	67 (82)	63 (81)	68 (83)
Fibrosis on previous liver biopsy — no. (%)				
None or minimal	35 (43)	30 (37)	31 (40)	28 (34)
Portal	37 (46)	46 (56)	43 (55)	46 (56)
Bridging	9 (11)	6 (7)	3 (4)	8 (10)
Cirrhosis§	—	—	1 (1)	—
ALT, IU/ml — median (range)	56 (18–277)	50 (15–259)	58 (18–303)	55 (20–315)
HCV genotype 1 subtype — no. (%)¶				
1a	31 (38)	37 (45)	40 (51)	35 (43)
1b	50 (62)	45 (55)	38 (49)	45 (55)
Indeterminate	0	0	0	2 (2)

* Plus-minus values are means ±SD. Patients were randomly assigned to receive one of four treatments involving various combinations of telaprevir (1250 mg on day 1, then 750 mg every 8 hours), peginterferon alfa-2a (180 µg weekly), and ribavirin (dose according to body weight). The T12PR24 group received telaprevir, peginterferon alfa-2a, and ribavirin for 12 weeks, followed by peginterferon alfa-2a and ribavirin for 12 more weeks. The T12PR12 group received telaprevir, peginterferon alfa-2a, and ribavirin for 12 weeks. The T12P12 group received telaprevir and peginterferon alfa-2a without ribavirin for 12 weeks. The PR48 (control) group received peginterferon alfa-2a and ribavirin for 48 weeks. There were no significant differences between the control group and any of the three telaprevir-based groups. ALT denotes alanine aminotransferase, and HCV hepatitis C virus.

† The body-mass index (BMI) is the weight in kilograms divided by the square of the height in meters.

‡ Race or ethnic group was self-reported. Percentages do not total 100 because of rounding.

§ Patients with cirrhosis on previous biopsy were excluded from the study; however, one such patient was enrolled in error by a site investigator.

¶ HCV genotype and subtype were determined through sequence analysis of the coding region of the nonstructural 3 protease, performed as part of resistance studies.

Tableau 6 : Caractéristiques générales des patients de l'étude HCV Prove2 (Hézode *et al.* 2009).

étudié sur une période couvrant jusqu'à trois années de traitement. Cette étude faisait suite à des résultats préliminaires obtenus par Gilead chez 30 patients de mêmes caractéristiques suggérant que la prévalence des mutations PreC avant la mise en place du traitement était prédictive du succès thérapeutique.

Etude HCV Prove2

L'étude HCV PROVE2 est un essai de Phase II visant à évaluer l'efficacité et la tolérance d'un traitement par telaprevir et interféron alpha pégylé avec ou sans ribavirine chez des patients ayant une infection chronique par un VHC de génotype 1 (Tableau 6) (Hezode *et al.* 2009). Dans cette étude, les patients ont été randomisés en quatre groupes de 78 à 82 individus, d'âge moyen de 45 ans, majoritairement masculins. La sévérité de leur pathologie hépatique a été évaluée par biopsie hépatique et a révélé une maladie peu évoluée dans plus d'un tiers des cas, avec une fibrose uniquement portale chez plus de la moitié d'entre eux. La charge virale initiale, quantifiée par COBAS TaqMan HCV assay, version 1.0 (Roche, Bâle, Suisse) a été établie à 6,4 log UI/mL en moyenne. Les patients des 4 groupes ont reçu : soit 12 semaines de trithérapie par ribavirine + interféron (PR) + telaprevir suivies de 12 semaines de PR (T12PR24) ; soit 12 semaines de trithérapie (T12PR12) ; soit 12 semaines de telaprevir et d'interféron sans ribavirine (T12P12) ; soit 48 semaines de traitement par interféron et ribavirine (groupe contrôle : PR48). Les patients ont été suivis régulièrement pendant le traitement et jusqu'à 254 semaines après son arrêt. Parmi les patients des 4 groupes, les 18 sujets inclus dans l'essai à l'hôpital Henri Mondor ont été choisis pour réaliser l'étude longitudinale : 6 appartenaient au groupe T12P12, 7 aux groupes T12PR12 et T12PR24 et 5 au groupe PR48.

Characteristic ^a	Result at:	
	Screening (n = 189)	Baseline MVC (n = 112)
General		
Male (%)	78.1	77.5
Median age, yr (IQR)	45.3 (41.8-50.7)	45.7 (42.1-51.2)
Median CD ₄ cell count/ μ l (IQR)	247 (110-396)	257 (123-394)
Median nadir CD ₄ cell count/ μ l (IQR)	75 (14-160)	99 (27-177)
Median plasma HIV-1 RNA log ₁₀ copies/ml (IQR)	4.2 (3.5-4.8)	4.2 (3.4-4.9)
Previous antiretroviral treatment		
Median no. of NRTIs (IQR)	6 (5-7)	6 (5-7)
Median no. of NNRTIs (IQR)	1 (1-2)	1 (1-2)
Median no. of PIs (IQR)	4 (3-6)	4 (3-6)
Enfuvirtide (%)	45.0	45.0
Raltegravir (%)	22.2	22.2
HIV-1 subtype B (%)	88.7	90.4
Profile result (%)		
CCR5	71.4	100
CXCR4	2.7	0
D/M	25.9	0
Coprescribed antiretroviral drugs (%)		
Raltegravir	NA ^b	67.9
Darunavir	NA	53.6
Etravirine	NA	28.6
Enfuvirtide	NA	17.0
Median wGSS (IQR)	NA	1 (0.5-1)

^aMVC, maraviroc; IQR, interquartile range; NRTIs, nucleoside RT inhibitors; NNRTIs, nonnucleoside RT inhibitors; PIs, protease inhibitors; D/M, dual/mixed; wGSS, weighed genotypic sensitivity score.

^bNA, not applicable.

Tableau 7 : Caractéristiques générales des patients de l'étude HIV génotropisme (Recordon-

Pinson *et al.* 2010).

Etude HIV géotropisme

189 patients infectés par un VIH-1 traités par HAART ont été inclus dans cette étude visant à évaluer les marqueurs de réponse au traitement par maraviroc (Recordon-Pinson *et al.* 2010). Les patients, majoritairement masculins, avaient une moyenne d'âge de 45 ans, un taux de CD4 médian de 247/ μ L et une charge virale médiane de 4.2 Log copies/mL. Les données du test phénotypique de géotropisme Trofile (Monogram Biosciences, San Francisco, USA) ont été recueillies et les patients présentant un tropisme CCR5 ont été sélectionnés pour recevoir du maraviroc ; 113 patients ont ainsi été traités par une trithérapie optimisée, additionnée de maraviroc (Tableau 7). Des prélèvements ont été réalisés 1, 3 et 6 mois après le début du traitement par maraviroc. A chacun de ces prélèvements, la charge virale a été mesurée à l'aide du kit Cobas TaqMan (Roche, Bâle, Suisse) en log UI/mL.

Amorces	Virus/gène	Position	Séquences
POL1 (Sens)	HBV/POLrt	186-206	5'-CCC TGC TCG TGT TAC AGG CGG-3'
POL2 (Antisens)	HBV/POLrt	1196-1174	5'-GTT GCG TCA GCA AAC ACT TGG CA-3'
POL3 (Sens)	HBV/POLrt	251-272	5'-GAC TCG TGG TGG ACT TCT CTC A-3'
POL4 (Antisens)	HBV/POLrt	1058-1033	5'-GGC ATT AAA GCA GGA TAA CCA CAT TG-3'
HBV5_XL_F	HBV/POLrt	324-343	5'-CCTCCAATCACTCACCAACC-3'
HBV5_XL_R	HBV/POLrt	901-919	5'-GGAACCTTGCCACAGGATC-3'
HBV3_XL_F	HBV/POLrt	901-919	5'-GGAACCTTGCCACAGGATC-3'
HBV3_XL_R	HBV/POLrt	324-343	5'-CCTCCAATCACTCACCAACC-3'
RNaseH-F	HBV/POLrh	1031-1058	5'-CACAATGTGGATATCCTGCCTTAATGC-3'
RNaseH-R	HBV/POLrh	1702-1728	5'-CCTACTTCAAAGACTGTGTGTTAAG-3'
3719_F	HBV/POLrh	1134-1153	5'-CATGAACCTTTACCCCGTTG-3'
3719_R	HBV/POLrh	1661-1680	5'-GACTCTTGGACTCCCAGCAA-3'
Outer Primer mix	HBV/PreCore/Core	Inconnues	Amorces du kit InnoLipa PreC
HBV_PreC_F	HBV/PreCore/Core	1704-1726	5'-TACTTCAAAGACTGTGTGTTTAC-3'
HBV_PreC_R	HBV/PreCore/Core	1924-1942	5'-ATTTGGAGCTACTGTGGAG-3'
3ES	HCV/NS3	3317-3336	5'-TGT GGG GAC ATC ATC TTG G- 3'
ZekriSE	HCV/NS3	3364-3379	5'-GCG TGC GAC ATC ATC- 3'
NS3SE	HCV/NS3	3221-3244	5'-CA(CT) (AG) CG GGC CT(AG) CGA GAC CTT GC-3'
3EAS	HCV/NS3	4039-4059	5'-AGC GGT AAG AGC ACC AAG GT-3'
3IS	HCV/NS3	3364-3385	5'-GAG ATA CTT TTG GGA CCG GCT-3'
ZekriSi	HCV/NS3	3364-3383	5'-GAG ATA CTG CTC GGG CCG G-3'
NS3Si	HCV/NS3	3296-3316	5'-TGG GGG GC (AG) GAT AC(CT) GCG GC- 3'
3IAS	HCV/NS3	3967-3988	5'-GGA GAT GAG TTG TCT GTG AA-3'
HIV_V3-F-outer	HIV/Env	6855-6879	5'-GAG CCA ATT CCC ATA CAT TAT TGT-3'
HIV_V3-R-outer	HIV/Env	7786-7817	5'-TAA GTC TCT CAA GCG GTG GTA GCT GAA-3'
HIV_V3-F-inner	HIV/Env	6877-6898	5'-TGT GCC CCA GCT GGT TTT GCG AT-3'
HIV_V3-F-inner	HIV/Env	6877-6898	5'-GGA TCT GTC TCT GTC TCT CTC TCC A-3'
M13S (Sens)	/	/	5'-GTA AAA CGA CGG CCA GT-3'
M13R (Antisens)	/	/	5'-CAG GAA ACA GCT ATG AC-3'

Tableau 8 : Amorces utilisées pour les différentes approches de biologie moléculaire.

Préparation des échantillons pour les techniques de biologie moléculaires d'étude des variants viraux

Etude HBV POL et HBV InVS

L'ADN viral contenu dans 200µL de sérum a été extrait à l'aide du coffret QIAamp[®] DNA blood kit[™] (QIAGEN GmbH, Hilden, Allemagne) selon les instructions du fabricant, après élution dans 50 µL d'eau stérile. L'amplification par PCR (PCR1) du domaine A à E de la région transcriptase inverse du gène P a été réalisée à l'aide des deux amorces POL1 et POL2 (Tableau 8). Quand la charge virale était inférieure à 1000 UI/mL, elle a été suivie d'une seconde PCR (PCR2) utilisant les amorces POL3 et POL4 (Pallier *et al.* 2009). La réaction de PCR1 a été réalisée avec 5 µL d'ADN viral extrait et la PCR2 avec 2 µL d'amplicons de PCR1. Le milieu réactionnel des deux PCR était composé de 10 pmol de chaque amorce, de tampon Ampli Taq Gold 10X, 0,75 µmol de MgCl₂, 10 nmol de dNTP, 2,5 UI d'Ampli Taq Gold (Applied Biosystems, Carlsbad, Californie) et d'eau stérile qsp 45 µL. L'amplification s'est déroulée dans un thermocycleur ABI 9700 (Applied Biosystems, Carlsbad, Californie) selon le programme suivant : PCR1 : [10 min, 95°C ; 40 cycles de : 30 s, 94°C ; 30 s, 48°C ; 1 min, 72°C ; 10 min, 72°C], PCR2 : [10 min, 95°C ; 35 cycles de : 30 s, 94°C ; 30 s, 55°C ; 1 min, 72°C ; 10 min, 72°C]. Afin d'évaluer l'efficacité de l'amplification, un dixième du produit de PCR additionné de bleu de bromophénol a été analysé par migration électrophorétique en gel d'agarose à 2% dans du tampon TBE 0.5X (Tris base 50 mM, acide borique 45 mM, EDTA 0.5 mM, pH 8.4) additionné de 1% de SybrSafe DNA gel Stain (Applied Biosystems, Carlsbad, Californie). Les bandes obtenues ont été visualisées sous lumière UV. Les produits de PCR

ont ensuite été purifiés à l'aide du kit Amicon-Ultra column (Milipore, Billerica, Massachusetts).

Etude HBV RNaseH

L'ADN du VHB a été extrait à partir de 200 à 500 µL de sérum à l'aide du kit QIAamp DNA blood kit sur automate QiaCube (Qiagen GmbH, Hilden, Allemagne) selon les instructions du fabricant. Un fragment de 697 pb couvrant la totalité de la RNase H de la polymérase du VHB a été amplifié à l'aide des amorces RNaseH-F et RNaseH-R (Tableau 8) et de la Taq polymérase AmpliTaqGold (Life Technologies, Carlsbad, Californie) selon les recommandations du fabricant, sur un thermocycleur ABI 9700 (Life Technologies, Carlsbad, Californie) suivant le programme : [95°C, 10min ; 40 cycles de : 95°C, 30s ; 58°C, 30s ; 72°C, 40s ; 72°C, 7min]. Les amplicons ont ensuite été purifiés à l'aide du kit Amicon-Ultra column (Milipore, Billerica, Massachusetts) et leur taille vérifiée par migration électrophorétique sur gel d'agarose à 2% comme décrit précédemment.

Etude HBV PreC

L'ADN viral contenu dans 200µL de sérum a été extrait à l'aide du kit QIASymphony DSP Virus/Pathogen Kit sur automate QIASymphony (Qiagen, Hilden, Allemagne). L'amplification par PCR du domaine PreC/C a été réalisée de manière à pouvoir comparer les performances des kits INNO-LiPA HBV PreCore (Innogenetics, Gand, Belgique) et du séquençage de nouvelle génération. La préparation des échantillons a été réalisée selon le protocole du kit. Brièvement, un mélange contenant le tampon Taq Stratagene 10x, les dNTP (25 mM), le « Outer Primer Mix » et la Taq DNA Polymerase 5 UI/µl (Stratagene, La Jolla, Californie) a été utilisé pour réaliser une amplification par PCR de 20 µl d'échantillon extrait sur thermocycleur ABI 9700 (Applied Biosystems, Carlsbad, Californie) selon le programme [4

min, 94°C ; 40 cycles de : 30s, 94°C ; 30s, 50°C ; 30s, 72°C ; 10 min, 72°C]. Afin d'évaluer l'efficacité de l'amplification, un dixième du produit de PCR additionné de bleu de bromophénol a été analysé par migration électrophorétique en gel d'agarose à 2% dans du tampon TBE 0.5X (Tris base 50 mM, acide borique 45 mM, EDTA 0.5 mM, pH 8.4) additionné de 1% de SybrSafe DNA gel Stain (Applied Biosystems, Carlsbad, Californie). Les bandes obtenues ont été visualisées sous lumière UV.

Etude HCV Prove2

200 µL de sérum de patient ont été extraits à l'aide du kit QIAamp Viral RNA sur l'automate QIAcube (Qiagen, Hilden, Allemagne) selon les instructions du fournisseur et élués dans 60µL d'eau. 15 µL d'extrait ont ensuite été utilisés pour la RT-PCR one-step multiplexée à l'aide du kit Qiagen OneStep RT PCR (Qiagen, Hilden, Allemagne) selon les instructions du fournisseur à l'aide des amorces sens NS3SE (10µM), ZekriSE (10µM) et 3ES (10µM), et antisens 3EAS (10µM) (Tableau 8) sur thermocycleur ABI 9700 (Applied Biosystems, Carlsbad, Californie) selon le programme : [60min, 50°C ; 15 min, 95°C ; 40 cycles de : 30s, 94°C ; 30s, 55°C ; 1min, 72°C ; 10 min, 72°C]. Les produits d'amplification ont ensuite été purifiés à l'aide du kit Amicon-Ultra column (Miliipore, Billerica, Massachusetts). La vérification qualitative et quantitative des produits d'amplification a été réalisée par PCR en temps réel avec 5 µL de RT-PCR one-step dilués au 1/50^{ème}, le kit SYBR® Green PCR Master Mix (Applied Biosystems, Carlsbad, Californie) et les amorces ZekriSi (10µM) et 3IAS (10µM) (Tableau 8) sur thermocycleur ABI 7300 (Applied Biosystems, Carlsbad, Californie) selon le programme d'amplification [2min, 50°C ; 10 min, 95°C ; 45 cycles de : 15s, 95°C ; 1min, 60°C] additionné d'une étape de dissociation [95°C, 15s ; 60°C, 15s ; rampe de +0.5°C/s jusqu'à 95°C]. Les échantillons présentant un mauvais profil de dissociation ou un signal d'amplification inférieur à 25 cycles ont été écartés de l'étude.

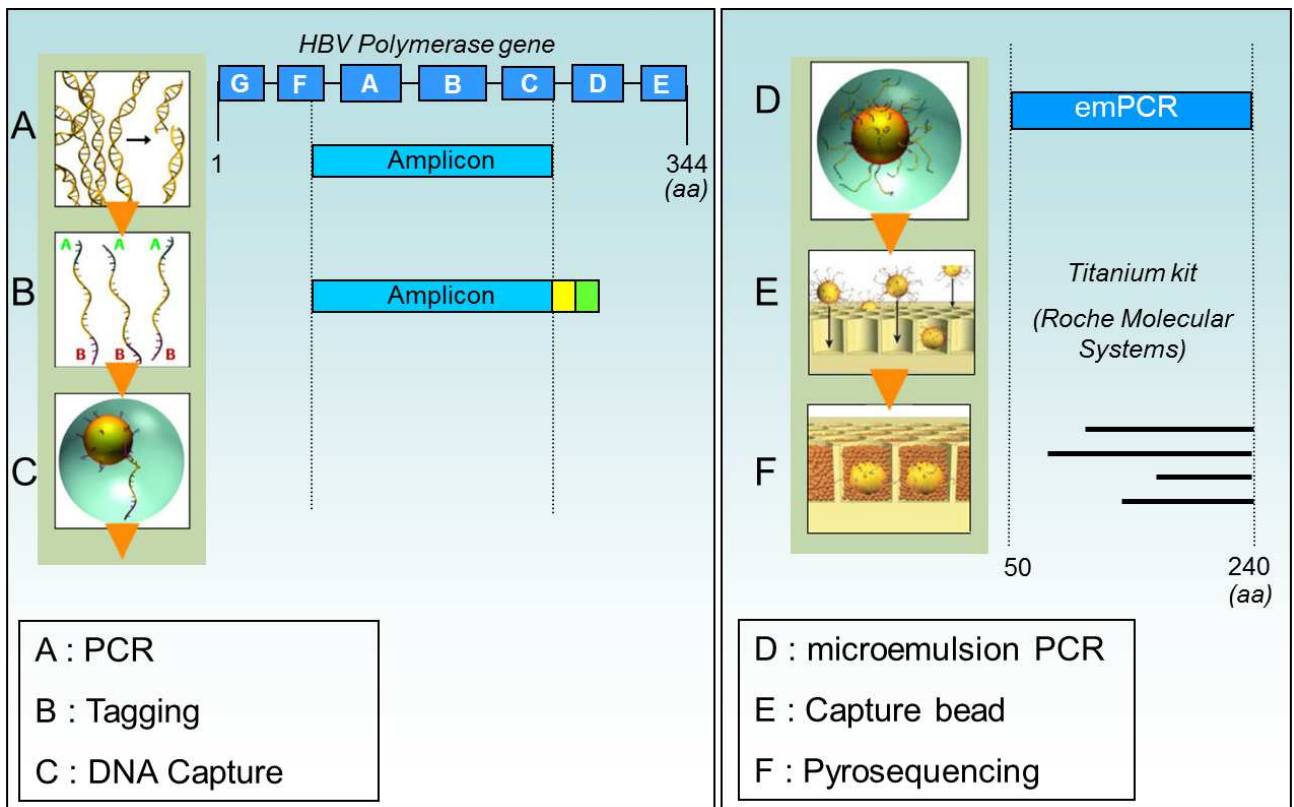


Figure 30 : Représentation schématique des étapes techniques de réalisation du pyroséquençage à haut-débit sur le séquenceur GS FLX de Roche.

Etude HIV géotropisme

Les échantillons de cette étude ont été traités selon une méthode décrite précédemment (Swenson *et al.* 2010; Swenson *et al.* 2011; Swenson *et al.* 2011). L'ARN du VIH a été extrait à partir de 500 µL de plasma à l'aide du kit NucliSENS easyMAG (BioMérieux, Marcy-L'étoile, France). Trois RT-PCR one-step ont été réalisées à partir de 4 µL d'extrait pour chaque amplification à l'aide du kit Qiagen OneStep RT PCR (Qiagen, Hilden, Allemagne) selon les instructions du fournisseur à l'aide des amorces HIV_V3-F-outer et HIV_V3-F-inner (Tableau 8) selon le programme [30min, 52°C ; 2 min, 94°C ; 40 cycles de : 15s, 94°C ; 30s, 55°C ; 90s, 68°C ; 5min, 68°C]. L'amplification a été vérifiée par migration sur gel d'agarose selon le protocole décrit précédemment. Puis les amplicons ont été purifiés sur colonne Amicon-Ultra column (Milibore, Billerica, Massachusetts).

Pyroséquençage à haut débit (technologie 454)

Etude HBV POL

Les différentes étapes du pyroséquençage sont résumées Figure 30. 5 µL d'amplicon purifiés ont été utilisés pour la réalisation de la seconde PCR à l'aide des amorces POL3 et POL 4 (Tableau 8), selon les conditions décrites ci-dessus. POL4 a été modifiée pour permettre l'introduction des adaptateurs de fixation sur les billes utilisées pour le séquençage et des tags permettant d'identifier chaque échantillon (MIDs: Multiplex Identifiers). Ainsi, 8 amorces possédant 8 MIDs différents ont été utilisées pour identifier les échantillons. Une purification post-PCR a été réalisée à l'aide du kit Nucleofast® 96 PCR plates (Clontech, Mountain View, Californie) selon les instructions du fabricant. Les produits de PCR ont été quantifiés à l'aide du kit Quant-iT™ PicoGreen® dsDNA kit (Invitrogen, Carlsbad, Californie)

sur le lecteur DTX 880 Multimode Detector (Beckman-Coulter, Brea, Californie) puis normalisés à 1.10^9 copies/mL. Ensuite, la phase de PCR en émulsion, au cours de laquelle chaque amplicon se fixe de manière clonale à une bille, a été réalisée à l'aide du kit GS FLX Titanium emPCR Kit (454 Life Sciences, Roche Diagnostics Corp.), suivie par une purification et un enrichissement selon les instructions du fabricant. A la fin de ces étapes, les billes fixant les amplicons clonaux ont été comptées et normalisées à l'aide du compteur Beckman Coulter Z1 particle counter (Beckman Coulter, Brea, Californie) avant d'être répartis en 16 lots de 8 échantillons possédant chacun un MID différent. Puis chaque lot a été déposé sur une région de la lame de séquençage 16 régions GS FLX Titanium PicoTiterPlate (454 Life Sciences, Roche Diagnostics Corp.). Le séquençage a ensuite été réalisé sur le séquenceur GS FLX à l'aide du kit GS FLX Titanium Sequencing Kit (454 Life Sciences, Roche Diagnostics Corp.).

Etude HBV InVS

L'étude HBV InVS a été réalisée de la même manière que l'étude HBV POL, mais le séquençage a cette fois été réalisé en sens et antisens à l'aide des amorces de PCR 2 : HBV5_XL_F, HBV5_XL_R et HBV_3_XL, HBV3_XL_R (Tableau 8). Les amorces HBV5_XL_F et HBV3_XL_R ont été modifiées pour recevoir les MIDs (16 différents). Après PCR et purification, la qualité des amplicons a été contrôlée à l'aide du profil obtenu par dépôt sur puce Agilent DNA1000 (Agilent, Santa Clara, Californie) puis, lorsque les profils étaient corrects, une quantification a été réalisée comme précédemment décrit pour permettre la normalisation à 10^9 copies/ μ L. La PCR en émulsion, la purification, l'enrichissement et le comptage des billes ont été réalisés de la même manière que précédemment, à l'exception de l'organisation des lots qui s'est faite en 8 lots de 16 échantillons possédant chacun un identifiant différent. Les billes ont été transférées sur une lame de séquençage 8 régions GS

FLX Titanium PicoTiterPlate (454 Life Sciences, Roche Diagnostics Corp.), puis le séquençage a été réalisé sur séquenceur GS FLX+ à l'aide du kit GS FLX+ Titanium Sequencing Kit (454 Life Sciences, Roche Diagnostics Corp.).

Etude HBV RNase H

Le protocole de pyroséquençage HBV RNase H a suivi grossièrement celui de l'étude HBV InVS avec une amplification secondaire à l'aide des amorces 3719_F et 3719_R, toutes deux modifiées pour recevoir l'un des 40 MIDs utilisés dans ce projet. Ainsi, 4 lots de 40 échantillons ont été préparés selon le protocole décrit précédemment. Le séquençage a été réalisé en sens et antisens sur une lame de séquençage 4 régions de type GS FLX Titanium PicoTiterPlate (454 Life Sciences, Roche Diagnostics Corp.) sur séquenceur GS FLX à l'aide du kit GS FLX Titanium Sequencing Kit (454 Life Sciences, Roche Diagnostics Corp.).

Etude HBV PreC

Le pyroséquençage à haut-débit du projet HBV PreC a été réalisé selon les mêmes modalités que celui du projet HBV POL, avec les amorces HBV_PreC_F et HBV_PreC_R. Dans ce cas particulier, c'est l'amorce HBV_PreC_F qui a reçu les MIDs (25 différents). Le séquençage a été réalisé en sens uniquement, par lots de 25 échantillons, chaque lot étant déposé sur une région de Picotiter plate 8 régions. Deux lames ont été nécessaires pour couvrir le nombre d'échantillons étudiés.

Etude HCV prove2

Le projet HCV Prove 2 a été réalisé exactement dans les mêmes conditions techniques que le projet HBV InVS, c'est-à-dire par séquençage double sens sur les kits et séquenceurs GS FLX+. Les amorces utilisées pour le deuxième round de PCR étaient celles utilisées pour la

quantification en PCR temps réelle décrites dans la partie « préparation des échantillons », modifiées pour accueillir un MID (8 différents utilisés). Le séquençage a été réalisé sur 2 lames 4 régions, chaque région ayant reçu un lot de 8 échantillons possédant chacun un MID différent.

Etude HIV géotropisme

Dans cette étude, le second round de PCR a été réalisé à l'aide des amorces HIV_V3-F-inner et HIV_V3-R-inner, modifiées chacune pour accueillir un des 12 MIDs utilisés pour ce projet. Après purification à l'aide du kit Agencourt Ampure PCR Purification beads (Beckman-Coulter, Brea, Californie), les amplicons ont été quantifiés par le kit Quant-iT Picogreen dsDNA Assay Kit (Invitrogen) sur le lecteur DTX 880 Multimode Detector (Beckman-Coulter, Brea, Californie), et normalisés à 2.10^5 copies/mL. Les amplicons ont ensuite été amplifiés de façon clonale sur bille lors de l'étape de PCR en émulsion selon les mêmes conditions que celles décrites précédemment. Après traitement, les échantillons répartis par lot de 12, chacun possédant un MID différent, ont été transférés sur une lame de séquençage 4 régions.

Contrôles

La procédure a été identique pour les projets concernant le VHB et le VHC. En pratique, deux amplicons provenant d'un patient, l'un avant traitement et l'autre au moment de l'échappement thérapeutique pour lequel la présence majoritaire d'une mutation a été décrite et pour lequel une amplification selon le protocole décrit dans la partie « Préparation des échantillons pour les techniques de biologie moléculaires d'exploration des variants viraux » a pu être obtenue, ont été sélectionnés. Les produits de PCR correspondants ont été purifiés sur colonne Amicon-Ultra column (Milipore, Billerica, Massachusetts) selon les

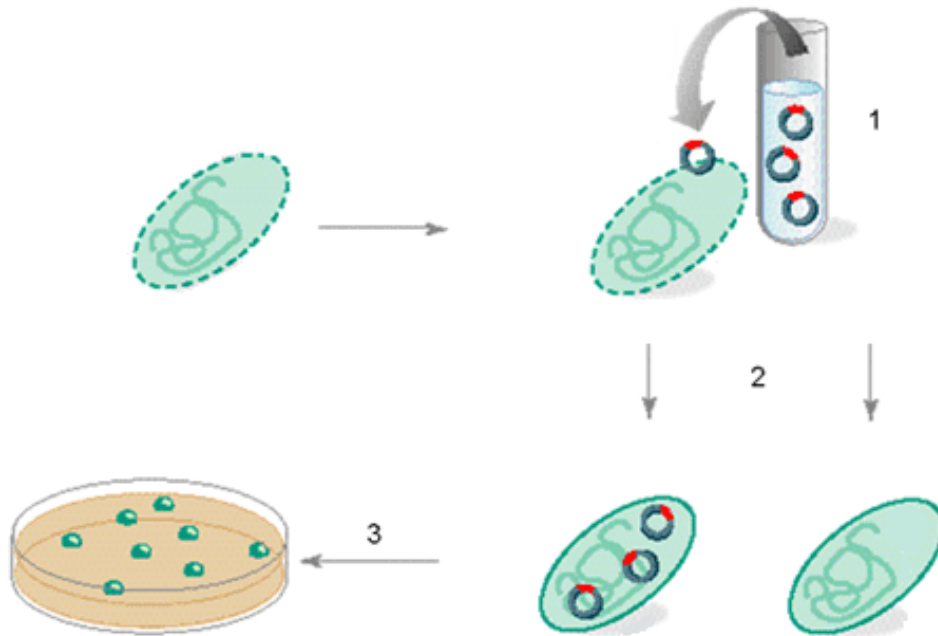


Figure 31 : Représentation schématique des étapes techniques du clonage-séquençage : les bactéries compétentes TOPO10 sont transformées avec le plasmide TOPO TA clonage ayant intégré préalablement l'ADN d'intérêt et possédant un gène de résistance à l'ampicilline (1). Les bactéries sont ensuite incubées dans un milieu de culture liquide enrichi (2), puis ensemencées sur des boîtes de Pétri agar/ampicilline (3). Les bactéries ayant intégré le vecteur résistant à l'ampicilline forment des colonies (3), tandis que les autres sont éliminées.

instructions du fabricant puis ligaturés dans un plasmide pCR[®]4-TOPO[®] (Life Technologies, Carlsbad, Californie) selon le protocole recommandé. 50 µL de bactéries *Escherichia coli* résistantes à l'ampicilline et compétentes one-shot[®] TOP 10[®] (Life Technologies, Carlsbad, Californie) ont été transformées avec 3 µL d'ADN ligaturé par choc thermique (30 min, 0°C ; 30 s, 42°C ; 2min, 0°C) puis transférées à 37°C sous agitation pendant 1 heure dans 250 µL de SOC (2% de tryptone, 0,5% d'extrait de levure, NaCl 10 mM, KCl 2,5 mM, MgCl₂ 10 mM, MgSO₄ 10 mM, glucose 20 mM) (Figure 31). Les bactéries transformées ont étéensemencées sur gélose cœur-cerveille (BD, Franklin Lakes, New Jersey) en présence de 100 µg/mL d'ampicilline et incubées à 37°C une nuit. Pour chaque prélèvement, dix colonies ont été repiquées et mises en culture dans 150 µL de bouillon cœur-cerveille en présence de 100 µg/mL d'ampicilline (BHI-Ampi) et incubées à 37°C 4 heures sous agitation, puis congelées à -80°C dans du glycérol à 50% volume à volume. En parallèle, une amplification par les amorces M13R et M13F du kit pCR[®]4-TOPO[®] (Life Technologies, Carlsbad, Californie) a été réalisée à partir de 6 µL de bouillon bactérien selon le programme d'amplification (5 min, 94°C ; 35 cycles de 30 s, 94°C ; 30 s, 55°C ; 2 min, 72°C ; 10 min, 72°C). L'efficacité d'amplification a été vérifiée par migration électrophorétique des produits de PCR en gel d'agarose à 2%. Les produits de PCR positifs ont été purifiés sur colonne d'affinité Wizard SV 96 Plasmid DNA Purification System (Promega) selon les instructions du fabricant, puis la séquence nucléotidique a été déterminée par séquençage Sanger à l'aide des amorces M13R et M13F et du kit BigDye Terminator v3.1 Cycle Sequencing kit (Life Technologies, Carlsbad, Californie) selon le programme suivant : 1 min, 96°C ; 25 cycles de : 10 s, 96°C ; 5 s, 55°C ; 2 min, 60°C, sur séquenceur capillaire ABI Prism 3100 (Life Technologies, Carlsbad, Californie). Après correction des séquences (voir plus loin), un variant sauvage (wt) du VHB, un variant sauvage du VHC, 2 variants de VHB résistants portant respectivement les mutations rtN236T

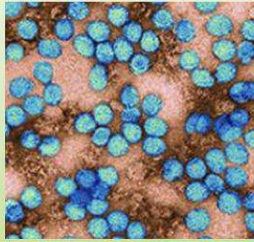
et rtA181T, 2 variants de VHB précore preCA1762T et preCG1896A, et 2 variants du VHC portant respectivement les mutations V36M et R155K du domaine NS3 conférant la résistance aux inhibiteurs de protéase ont été sélectionnés. Les colonies congelées dans le glycérol correspondant aux variants d'intérêt ont été cultivées une nuit dans 200 mL de BHI-ampi, puis le plasmide d'intérêt a été extrait à l'aide du kit PureLink® HiPure Plasmid Filter Maxiprep Kit (Life Technologies, Carlsbad, Californie). Les plasmides ont ensuite été dosés par spectrophotométrie à 260 et 280 nm pour vérifier la qualité de la préparation (rapport 260/280nm >1,8), afin de calculer le nombre de copies d'ADN et de normaliser tous les plasmides à 10^{10} copies/mL. Les solutions mères ont été à nouveau dosées avec le kit Quant-iT Picogreen dsDNA Assay Kit (Life Technologies, Carlsbad, Californie) sur le spectrophotomètre Mithrax LB 940 (Berthold, Bad Wilbad, Germany) selon les instructions du fabricant. Des solutions standards filles établies à $5 \cdot 10^3$ copies/mL, 10^5 copies/mL et 10^8 copies/mL ont été générées en triplicat pour chaque variant. Ces solutions ont ensuite été traitées de la même manière que les échantillons de patients à tester.

Autres techniques de biologie moléculaire

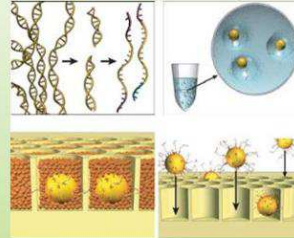
Etude HBV POL

Chaque échantillon prélevé a été étudié en parallèle du pyroséquençage à haut débit par clonage-séquençage. Le clonage-séquençage a été réalisé selon un protocole identique à celui décrit pour la réalisation des contrôles, à l'exception du nombre de clones réalisés qui a été fixé à 40 par point étudié.

Data collection

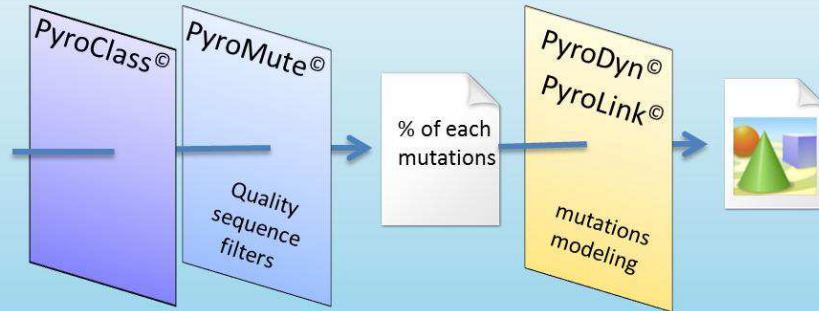


Viral extraction
from patient serum



RT. Amplification.
emPCR.
pyrosequencing

Analysis



Pyroclass®. Pyromute® and Pyrodyn® are protected under IDDN

Figure 32 : Représentation schématique du processus technique et analytique du pyroséquençage à haut-débit utilisant le package de logiciels PyroPack®.

Etude HBV PreC

Chaque échantillon de l'étude HBV PreC a été étudié en parallèle par la technique INNO-LiPA HBV PreCore (Innogenetics, Gent, Belgique) selon les préconisations du fournisseur. En pratique, après amplification du domaine pré-C, les bandelettes preCore ont été hybridées, lavées puis colorées à l'aide de l'automate Auto-LiPA 48 (Innogenetics, Gent, Belgique). Les bandelettes ont ensuite été lues à l'œil et les résultats (+ ou -) pour chaque mutation testée ont été reportés dans un fichier Excel (Microsoft, Redmond, WA).

Analyses des données

Séquençage Sanger

L'analyse des séquences a été réalisée à l'aide du logiciel SeqScape v2.5 (Life Technologies, Carlsbad, Californie) qui permet de corriger les séquences nucléotidiques et de les traduire en acides aminés. Elles ont ensuite été alignées à l'aide du logiciel ClustalX v1.83 (Chenna et al. 2003), puis des arbres phylogéniques ont été construits à l'aide de TreeView v1.6.6 (Page 1996). La séquence et le pourcentage de présence de chaque variant ont été déterminés et répertoriés à l'aide du logiciel GeneDoc v2.6.003 (Nicholas 1997).

Séquençage à haut-débit

Les données obtenues par séquençage à haut-débit ont été analysées au moyen de 5 logiciels (Figure 32) inclus dans le package PyroPack®, créé au laboratoire, qui constitue la première partie des résultats de ce mémoire (voir ci-dessous). Dans le cas du génotropisme du VIH, les résultats ont été comparés à ceux générés par un autre logiciel nommé geno2pheno[454] (Max-Planck-Institut, Saarbrücken, Allemagne) (Swenson *et al.* 2010;

Thielen and Lengauer 2012), suivant les recommandations allemandes et utilisant un taux de faux positifs fixé à 2,5%.

Analyses statistiques

R est un langage informatique spécialement dédié à l'analyse statistique des données. L'ensemble des tests statistiques simples (Test de Mann-Whitney, test de Student) ont été réalisés à l'aide du package de base fourni par RGui(64-bit) v2.14.1 (R_Development_Core_Team 2008). Un package additionnel, pROC v1.5 (Robin *et al.* 2011), qui permet de tracer les courbes de ROC (Receiver Operating Characteristic), d'établir l'aire sous courbe (AUROC) et de réaliser des comparaisons d'aires sous courbe à 0,5 (seuil de prédictivité) à l'aide de tests statistiques (bootstrap methods) (Venkatraman ES 1996; Venkatraman 2000; Bandos *et al.* 2005) a été utilisé pour analyser les données des études HBV PreC et HIV génotropisme. Enfin, pour certaines études, le Q de Yule, qui mesure l'intensité de liaison entre deux variables (Bakeman and Gottman 1989) a été calculé manuellement. L'interprétation des résultats a été réalisée de la manière suivante :

- Liaison nulle : $0 < Q < 0,09$
- Liaison faible : $0,10 < Q < 0,29$
- Liaison modérée : $0,30 < Q < 0,49$
- Liaison forte : $0,49 < Q < 0,69$
- Liaison très forte : $0,70 < Q < 1$

RESULTATS

Création du package PyroPack® pour l'analyse des données de pyroséquençage à haut débit (ARTICLES 2 ET 5)

Les données obtenues par séquençage a haut-débit ont été analysées au moyen de 5 logiciels (Figure 31), codés en ActivePerl64 v5.14.2.1402 (ActiveState, Vancouver, Canada) pour Windows 7 Professionel 64 bits (Microsoft, Redmond, Washington) sur un PC Z800 64bits (Hewlett Packard, Palo Alto, Californie) à double processeur hexacore Intel Xeon E5630 cadencés à 2.53GHz et associés à 24Go de Ram. Les logiciels sont inclus dans un package nommé PyroPack®, créé dans le laboratoire dans le contexte de cette Thèse, et dont les brevets ont été déposés à l'APP (agence de protection des programmes) sous les IDDN (Inter Deposit Digital Number) suivants (Annexe 2):

- PyroClass® : IDDN.FR.001.360005.001.S.P.2010.000.31230
- PyroMute® : IDDN.FR.001.120008.001.S.C.2010.000.31230
- PyroDyn® : IDDN.FR.001.370024.000.S.P.2010.000.31230
- PyroLink® : IDDN.FR.001.180032.000.S.P.2011.000.31230
- PyroTrop® : IDDN.FR.001.180031.000.S.P.2011.000.31230

Pyroclass® permet le démultiplexage des séquences générées. En effet, les données sont générées sous la forme de paquets de séquences correspondant chacun à une région de la lame de séquençage, chaque région contenant un lot d'échantillons. A l'intérieur de ce lot, chaque échantillon possède un identifiant différent (MID) qui correspond à une séquence de 8 bases inséré au début de chaque séquence. Pyroclass® prend en entrée le fichier de séquence, le fichier qualité correspondant et un fichier indiquant pour chaque échantillon la

séquence identité (fichier identité). Il lit ensuite les 8 premières bases de chaque séquence afin de relier la séquence à son identité. Le logiciel est conçu de manière à tolérer une erreur dans l'identité (une mutation) ; au-delà, la séquence est jugée difficilement identifiable et est définitivement écartée du processus d'analyse. A l'issue de l'attribution d'identité, le MID est excisé de la séquence ainsi que les données qualités correspondantes, puis la séquence et les données de qualité sont écrites dans un nouveau fichier échantillon. A l'issue du processus, il existe autant de fichiers échantillons que de MID et chacun contient uniquement la séquence virale correspondante. En parallèle, il est possible de préciser la séquence des amorces sens et antisens dans le fichier identité. Dans ce cas, le logiciel procède à une reverse complémentation des séquences 3'5' et des données de qualité, de manière à ce que le fichier final contienne uniquement des séquences positionnées en sens.

PyroMute® est un logiciel qualité qui prend en entrée les données formatées par PyroClass® et élimine les séquences, portions de séquences ou mutations pouvant être considérées comme des erreurs de séquençage. Afin d'effectuer ce tri, différents filtres ont été programmés. Le premier filtre retire, comme dans le contexte du séquençage Sanger, les séquences ou portions de séquences dont la qualité est faible, c'est à dire celles possédant à la fois un score de qualité de PHRED inférieur à 20 et un score d'environnement de la base inférieur à 20 (Altshuler *et al.* 2000). Les séquences trop courtes (moins de 50 pb) sont également définitivement éliminées. Les séquences sont ensuite alignées à l'aide de 3 processus successifs. Le premier applique l'algorithme de Smith-Waterman (Smith and Waterman 1981) auquel des modifications mathématiques ont été apportées pour améliorer la rapidité de résolution des matrices d'alignement ; un second algorithme, spécialement dédié à l'exploration du positionnement des insertions/délétions, prend le relais pour corriger les défauts d'alignement des indels ; enfin, un algorithme de détection

des homopolymères permet leur correction afin de pallier les erreurs fréquemment introduites à ce niveau par le séquenceur GS-FLX (Gilles *et al.* 2011). Le deuxième filtre de correction d'erreur est ensuite mis en œuvre. Les approches antérieures de correction d'erreurs étaient essentiellement fondées sur le calcul de probabilité d'évènements rares, appliquant la loi binomiale (Macalalad *et al.* 2012). Cette loi suppose que les n évènements (lecture d'une base) sont réalisés dans des conditions identiques, impliquant que n'importe quel évènement rare (par exemple une mutation) peut être analysé de la même manière. Il est évident que ces conditions ne sont pas remplies dans le cas du séquençage à haut débit. En effet, il est possible d'observer des variations de qualité importantes (score de PHRED) au niveau de chaque base de chaque séquence obtenue par ces méthodes, de telle sorte qu'une mutation rare détectée avec une très mauvaise qualité doit être plus facilement considérée comme une erreur qu'une mutation rare de très bonne qualité. Nous avons donc intégré au schéma de Bernoulli la possibilité de traiter la notion de qualité de l'information en utilisant le score qualité de PHRED fourni pour chaque base. Ainsi, pour chaque position i , il est possible de calculer, en fonction de la profondeur P_i (nombre de séquence en i) et de la qualité globale de l'information Q_i , un score t_i correspondant au seuil de significativité permettant de différencier une erreur d'une mutation au risque α . Chaque mutation de chaque position est alors testée de manière itérative contre le score t_i . Le risque α de première espèce est fixé à 5% pour un test, corrigé par la correction de Bonferroni en fonction du nombre de tests devant être réalisés pour maîtriser ce risque. A l'issue du test, un tableau regroupant les mutations significativement différentes d'une erreur aléatoire est créé afin de permettre les calculs de pourcentage de présence de chacune d'entre elle. Les données de pourcentage de présence et les alignements sont stockés en format texte en nucléotides et en acides aminés. Enfin, un rapport indiquant le taux de rejet de chaque filtre

ainsi que des données « métriques » de séquences (qualité globale, longueur moyenne, nombre de séquences, Q30) est fourni afin de permettre le contrôle du bon fonctionnement du processus et la qualité globale des données de l'échantillon.

PyroDyn[®] utilise les données corrigées de PyroMute[®] générées à partir de plusieurs échantillons prélevés à différentes étapes du traitement chez un même patient. Il établit pour chaque mutation détectée dans au moins un échantillon testé un tableau de pourcentage de présence à chaque temps d'étude. A partir de ce tableau, les calculs suivants sont réalisés : (i) régression non-linéaire de la série sur un modèle exponentiel et binomial ; (ii) calcul des coefficients de corrélation correspondants. Le modèle exponentiel correspond aux mutations en cours d'expansion ou de déclin, tandis que le modèle binomial correspond à des mutations qui ont eu une expansion suivie d'un déclin. En parallèle, un calcul de régression linéaire est établi à partir de tous les points situés entre le pourcentage le plus élevé et le pourcentage le plus bas. Les pentes obtenues se répartissent de manière gaussienne et permettent de calculer une moyenne, une variance et un seuil de significativité au-delà duquel la pente est considérée comme significativement supérieure à la moyenne. Ce calcul est fondé sur l'hypothèse que la majeure partie des mutations observées sont liés à des polymorphismes dont l'avantage sélectif est transitoire et donc non maintenu par la sélection naturelle. Finalement, ne sont retenues que les mutations associées à un coefficient de corrélation en modèle binomial ou exponentiel supérieur à 0,8 et les mutations qui présentent un taux de croissance global significativement supérieur à la moyenne. PyroDyn[®] analyse ainsi automatiquement l'ensemble des mutations et génère, pour chaque série de données, un tableau Excel (Microsoft, Redmond, Washington) associé à un graphique en cône 3D montrant exclusivement les mutations ayant subi une croissance ou une décroissance exponentielles, donc potentiellement liées à la résistance.

Pour chaque mutation retenue comme ayant subi des changements exponentiels par PyroDyn®, l'ensemble des combinaisons de liens possibles (linkage) est vérifié par PyroLink® de la manière suivante : à partir des données corrigées de PyroMute®, un inventaire des séquences couvrant l'ensemble des mutations étudiées pour le linkage est établi. Si cet inventaire répertorie au minimum 100 séquences, il a été établi un ratio entre les séquences possédant le lien étudié et celles ne le possédant pas. Ainsi, le pourcentage de présence de chaque famille de variants potentiellement résistants est déterminé. Le pourcentage de présence des variants sauvages (ensemble des variants ne possédant aucune mutation d'intérêt) est également calculé. L'ensemble est à nouveau stocké dans un fichier Excel répertoriant tous les résultats et un graphique en cône 3D est automatiquement créé dans une feuille de calcul séparée. Cette étape permet la description des dynamiques de populations virales complexes, portant une ou plusieurs substitutions, par rapport à la population wt. C'est cette description qui est la plus pertinente sur le plan clinique.

Pyrotrop® est un logiciel dérivé de PyroMute® dont la fonction est de fournir, en plus du pourcentage de mutations, les proportions respectives de variants CXCR4 et CCR5. Pour cela, à l'ensemble des fonctionnalités de PyroMute® a été ajouté un module de calcul du géotropisme selon les règles 11/25-charge (De Jong *et al.* 1992; de Jong *et al.* 1992; Fouchier *et al.* 1992; Fouchier *et al.* 1995; Resch *et al.* 2001; Delobel *et al.* 2007). Les résultats sont stockés dans le rapport d'analyse des séquences identique à celui décrit précédemment pour PyroMute®.

Etude HBV POL (ARTICLES 1 ET 2)

Profils cinétiques de charge virale

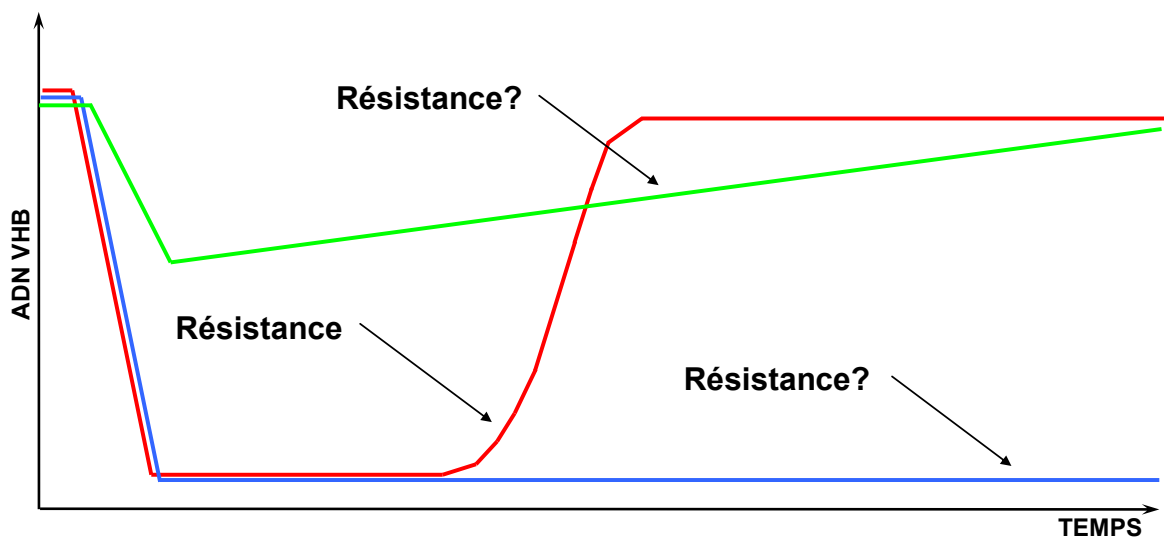


Figure 33 : Représentation des différents profils cinétiques de la charge virale du VHB dans le contexte de la résistance à l'adefovir. En rouge est présenté le profil classique, caractérisé par une réponse virologique suivi d'un échappement ; en vert le profil suboptimal, caractérisé par une réponse suboptimale suivie d'une lente réascension de la charge virale ; en bleu le troisième profil, caractérisé par une réponse sans échappement virologique malgré la sélection de variants viraux résistants.

L'analyse des profils cinétiques de charge virale de l'ensemble des patients de l'étude de Phase III ayant subi un échec thérapeutique associé au développement d'une résistance à l'adefovir (dont étaient issus les patients de l'étude HBV POL) a permis de mettre en évidence trois profils d'échec thérapeutique.

Le profil « classique » était caractérisé par une réponse virologique initiale à l'adefovir, définie par une diminution de 4 Log ou plus de la charge virale après 6 mois de traitement environ (Figure 33). Une phase de plateau plus ou moins longue a ensuite été observée avec une charge virale stable qui s'est maintenue à un faible niveau. Cette phase était suivie d'une remontée brutale de la charge virale, associée à la sélection des variants résistants. Le rebond virologique était caractérisé par une augmentation de plus de 1 Log de la charge virale au-dessus du nadir atteint sous traitement. L'échappement virologique était habituellement suivi d'un rebond biochimique. Ce profil a été décrit chez les patients traités par la lamivudine développant une résistance à ce médicament (Pallier *et al.* 2006).

Le profil « suboptimal » était défini par une réponse virologique initiale modérée à l'adefovir, caractérisée par une diminution de la charge virale d'environ 2 à 3 Log après environ 6 mois de traitement (Figure 33). Une lente réascension de la charge virale jusqu'à son niveau de base avant traitement a ensuite été observée. Les mutants viraux résistants sont apparus en séquençage direct après quelques mois de traitement, mais il était difficile d'affirmer que leur émergence était la cause plutôt que la conséquence de l'échappement virologique.

Le profil « échappement sans rebond » était caractérisé par une bonne réponse virologique à l'adefovir, suivie d'une phase de plateau. Dans ce cas, la sélection de variants viraux

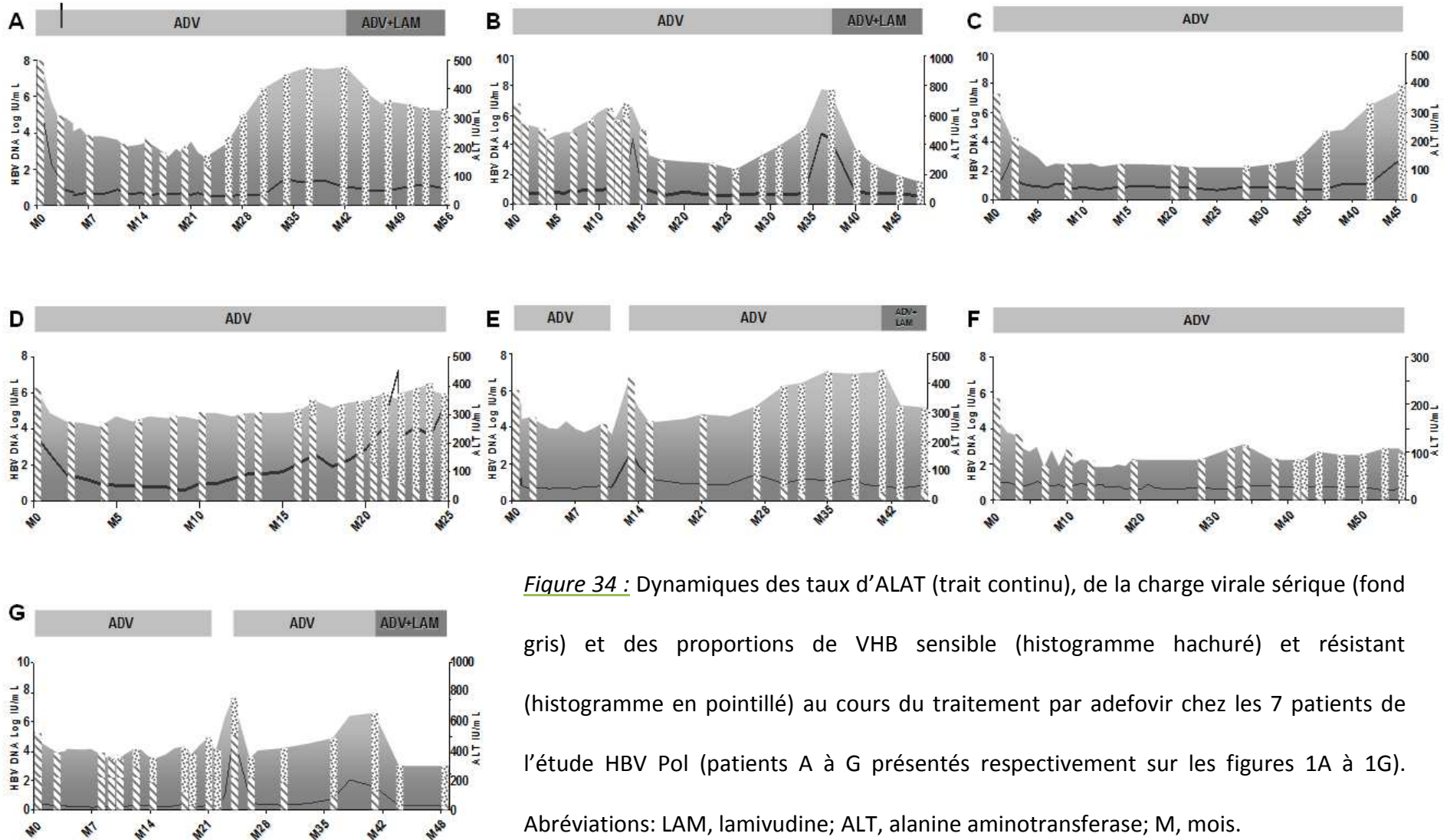


Figure 34 : Dynamiques des taux d'ALAT (trait continu), de la charge virale sérique (fond gris) et des proportions de VHB sensible (histogramme hachuré) et résistant (histogramme en pointillé) au cours du traitement par adefovir chez les 7 patients de l'étude HBV Pol (patients A à G présentés respectivement sur les figures 1A à 1G).
 Abréviations: LAM, lamivudine; ALT, alanine aminotransferase; M, mois.

portant des mutations de résistance à l'adefovir n'était pas suivie d'un rebond virologique (Figure 33).

Cinétiques virales sous adefovir chez les patients de l'étude HBV Pol (ARTICLE 1)

Les patients A, B, et C présentaient un profil de réponse classique, selon la classification présentée ci-dessus. Les patients A et B ont été traités par adefovir seul, puis par une combinaison d'adefovir et de lamivudine. Le patient C a reçu l'adefovir seul pendant 46 mois. L'administration d'adefovir a entraîné une décroissance importante de la charge virale, de 4,6 Log UI/mL en 12 mois pour le patient A, 2,4 Log UI/mL en 4 mois pour le patient B et 5,1 Log UI/mL en 8 mois pour le patient C (Figure 34). Un rebond virologique a été observé respectivement aux mois 28, 31 et 37 pour les patients A, B et C. La charge virale est ensuite revenue à son niveau de base en quelques mois. L'échappement virologique a été suivi d'un rebond biochimique (élévation des ALAT) lorsque la charge virale a atteint son niveau d'avant traitement dans les trois cas. Chez les trois patients, la population sauvage sensible a été progressivement remplacée par une population virale résistante, qui est devenue majoritaire au moment de l'échappement virologique en séquençage direct. L'ajout de lamivudine a permis de restaurer une réponse virologique partielle chez le patient A, avec une diminution de 2 Log de la charge virale, et une réponse importante chez le patient B, avec une chute de la charge virale d'environ 6 Log. Cependant, en dépit de la réponse virologique sous lamivudine chez ces deux patients, les variants viraux résistants à l'adefovir sont restés majoritaires sous la combinaison d'adefovir et de lamivudine.

Les patients D et E ont présenté un profil suboptimal de réponse au traitement. Le patient D (Figure 34) a été traité 25 mois par l'adefovir. Après une réponse virologique suboptimale

(décroissance de 2,5 Log UI/mL en 4 mois), la charge virale est remontée progressivement. Le rebond biochimique est survenu lorsque cette dernière a atteint son niveau initial. Le patient E (Figure 34) a été traité en monothérapie par adefovir 20 mois puis, après un arrêt de 2 mois, le traitement a été repris tout d'abord avec l'adefovir seul pendant 20 mois, puis avec une combinaison d'adefovir et de lamivudine. Une réponse suboptimale (-1,9 Log UI/mL en 10 mois) a été observée lors de la première séquence de traitement. Après la fenêtre thérapeutique, une nouvelle réponse suboptimale (-2,4 log UI/mL en 2 mois) a été suivie d'un retour progressif de la charge virale à son niveau de base. Chez les deux patients, les populations virales sauvages, sensibles à l'adefovir, ont été progressivement remplacées par des populations virales résistantes en séquençage direct. Chez le patient D, les populations résistantes sont devenues majoritaires au moment de l'échappement biochimique.

Le patient F a présenté un profil d'échappement sans rebond. Il a été traité 57 mois par adefovir et, après une réponse virologique initiale importante (-4,3 Log UI/mL en 7 mois), n'a pas présenté de rebond virologique malgré l'émergence de populations virales résistantes (Figure 34). La population sauvage sensible a été progressivement remplacée par les populations résistantes qui sont devenues majoritaires après 42 mois de traitement.

Finalement, le patient G n'a pas pu être classé. Il a d'abord été traité par adefovir seul, puis a observé une fenêtre thérapeutique de 2 mois avant de reprendre un traitement par adefovir puis par une combinaison d'adefovir et de lamivudine (Figure 34). Une réponse suboptimale (-1,6 Log UI/mL en 9 mois) a été observée lors de la première séquence de traitement. Après la fenêtre thérapeutique, une nouvelle réponse suboptimale (-2 Log UI/mL en 2 mois) a été suivie d'une phase en plateau qui a précédé le rebond virologique. L'ajout de lamivudine a permis la restauration de la réponse virologique. Dans ce cas, la population sauvage a été

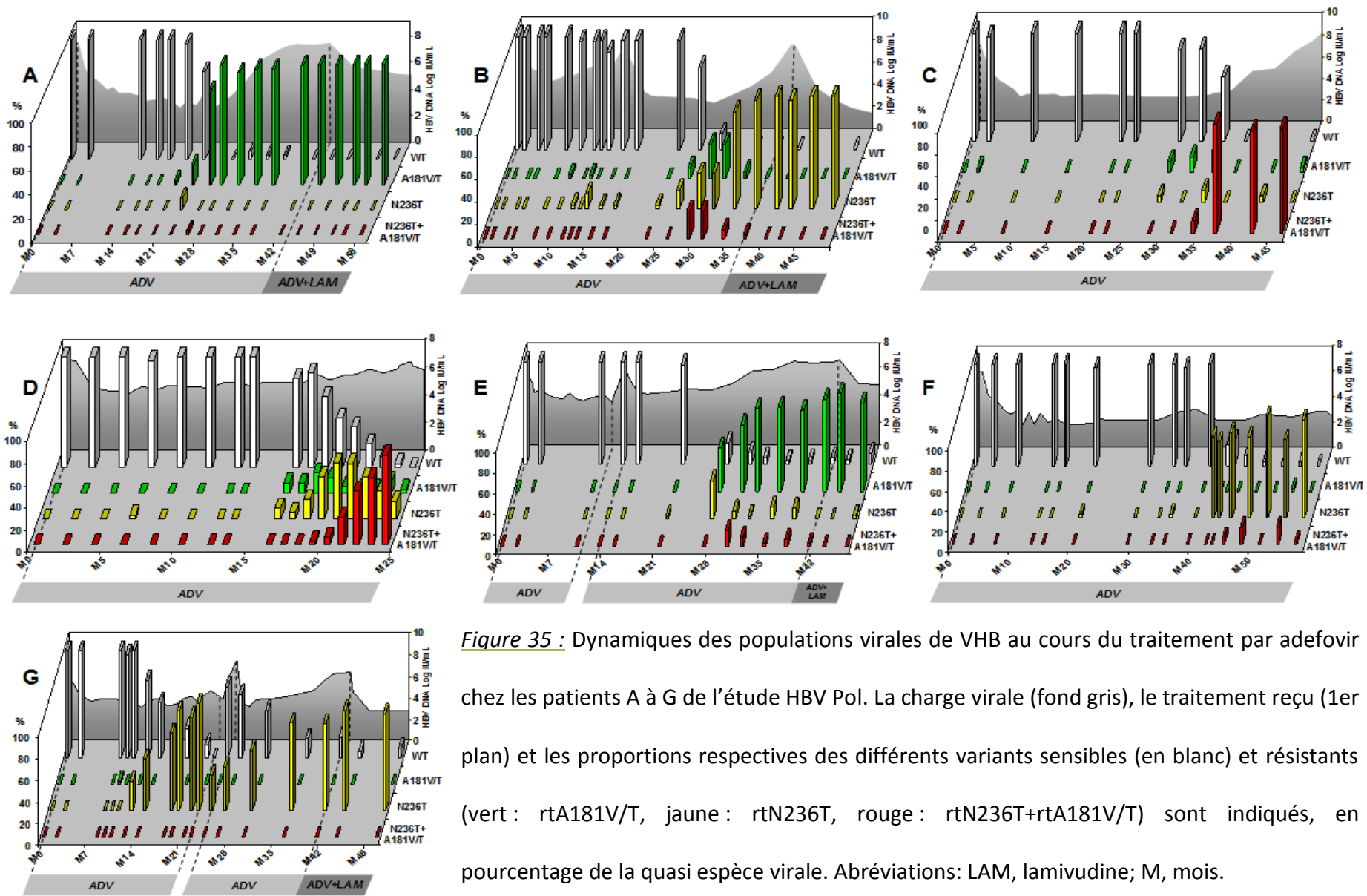


Figure 35 : Dynamiques des populations virales de VHB au cours du traitement par adefovir chez les patients A à G de l'étude HBV Pol. La charge virale (fond gris), le traitement reçu (1er plan) et les proportions respectives des différents variants sensibles (en blanc) et résistants (vert : rtA181V/T, jaune : rtN236T, rouge : rtN236T+rtA181V/T) sont indiqués, en pourcentage de la quasi espèce virale. Abréviations: LAM, lamivudine; M, mois.

remplacée par les populations résistantes lors de la première séquence de traitement, sans échappement virologique et avec une réponse suboptimale à l'adefovir ; un rebond virologique est survenu lors de la deuxième séquence de traitement à la suite d'une réponse suboptimale. Comme chez les autres patients, la population sauvage a été remplacée par des populations résistantes qui sont devenues majoritaires sous traitement par adefovir et le sont restées lorsque la lamivudine a été ajoutée.

Etude de la dynamique des variants sauvages et résistants des patients de l'étude HBV Pol par clonage-séquençage (ARTICLE 1)

3843 séquences de la transcriptase inverse du VHB ont été générées et analysées afin de caractériser les quasi-espèces virales avant et pendant le traitement par adefovir. A chaque date de prélèvement, les populations virales ont été quantifiées et exprimées en pourcentage de la population totale. Les résultats ne montraient pas de sélection d'autres mutations que celles déjà décrites dans la littérature (rtA181V/T, rtN236T et/ou rtA181V/T+rtN236T).

Les populations virales résistantes du patient A sont devenues majoritaires au moment du rebond virologique (Figure 35). Malgré l'apparition simultanée des trois populations résistantes, une seule a rapidement pris le dessus sur toutes les autres ; il s'agissait du variant rtA181V. Celui-ci est resté majoritaire tout au long du traitement par adefovir seul ou associé à la lamivudine.

Chez le patient B, l'échappement virologique est survenu au mois 25 de traitement. Il était également associé à la sélection de variants viraux résistants (Figure 35). La disparition progressive du virus sauvage a été contemporaine de l'apparition simultanée des trois

populations virales résistantes en proportions équivalentes. Lors de la poursuite du traitement, c'est le mutant rtN236T qui est devenu majoritaire aux dépens des autres populations virales résistantes (rtA181V/T ou rtN236T+rtA181V/T) et sauvages. Le variant rtN236T sélectionné est resté majoritaire après l'ajout de lamivudine, malgré une bonne réponse virologique à ce nouveau traitement.

Un rebond virologique a été observé au mois 35 chez le patient C (Figure 35). Comme chez le patient B, l'analyse des quasi-espèces virales par clonage-séquençage a montré l'apparition simultanée des trois populations virales résistantes au moment du rebond virologique. Chez ce patient, c'est la population rtA181V/T+rtN236T qui a été sélectionnée et a remplacé la population virale sauvage sensible. Chez les patients A, B, et C, aucune mutation de résistance autre que celles correspondant aux mutations rtA181V/T ou rtN236T n'a été décelée dans le cadre de lecture de l'Ag HBs.

Chez le patient D, l'analyse des quasi-espèces a mis en évidence trois vagues successives de sélection de variants viraux résistants, associées à la disparition progressive de la population virale sauvage (Figure 35). La première vague concernait le variant rtA181V/T, la seconde le variant rtN236T, la troisième le double mutant. C'est ce dernier qui est devenu majoritaire au stade le plus tardif du suivi.

Chez le patient E, c'est la population rtA181V/T qui est devenue majoritaire lorsque la charge virale a atteint son niveau initial (Figure 35). Le mutant sélectionné est resté majoritaire après l'ajout de lamivudine. Aucune mutation de résistance autre que celles correspondant aux mutations rtA181V/T ou rtN236T n'a été décelée dans le cadre de lecture de l'Ag HBs.

L'analyse des quasi-espèces virales du patient F a montré la sélection d'un variant rtN236T associée à la disparition progressive du variant sauvage (Figure 35). Contrairement au variant rtN236T du patient B, celui du patient F est devenu majoritaire sans entraîner de rebond virologique. Aucune mutation de résistance autre que celles correspondant aux mutations rtA181V/T ou rtN236T n'a été décelée dans le cadre de lecture de l'Ag HBs.

Lors de la première séquence de traitement du patient G, le mutant rtN236T a progressivement remplacé le virus sensible, sans entraîner de rebond virologique (Figure 35). La fenêtre thérapeutique ayant permis un retour du virus sauvage, le virus résistant rtN236T a de nouveau remplacé ce dernier lors de la réintroduction de l'adefovir, entraînant cette fois un échappement virologique. Malgré une bonne réponse virologique lors de l'introduction de la lamivudine en combinaison à l'adefovir, le virus sauvage est resté détectable sous forme minoritaire. Aucune mutation de résistance autre que celles correspondant aux mutations rtA181V/T ou rtN236T n'a été décelée dans le cadre de lecture de l'Ag HBs.

Etude de la dynamique des variants sauvages et résistants des patients de l'étude HBV Pol par pyroséquençage à haut débit (ARTICLE 2)

L'analyse des contrôles de pyroséquençage a montré que l'erreur maximum générée pour cette étude était globalement de $0,120 \pm 0,005\%$. Cependant, ce taux d'erreur variait selon la position considérée. Seules les substitutions dont la fréquence était supérieure au taux maximum additionné de 2 déviations standards par position ont été considérées dans l'analyse. Les taux d'erreurs étaient reproductibles à toutes les concentrations testées, couvrant des charges virales comprises entre $5 \cdot 10^3$ et 10^8 copies/mL.

	ADN du VHB (Log UI/mL)	rtV173L	rtL180M	rtA181V/T	rtT184S/A/I/L	rtR202G	rtM204V/I	rtN236T
Patient 1	6,6	<seuil	<seuil	0,47%	<seuil	<seuil	0,31%	0,25%
Patient 2	7,8	<seuil	<seuil	0,18%	<seuil	<seuil	<seuil	<seuil
Patient 3	6,8	<seuil	<seuil	0,22%	<seuil	<seuil	0,32%	<seuil
Patient 4	5,5	<seuil	<seuil	<seuil	0,17%	<seuil	<seuil	<seuil
Patient 5	7,6	<seuil	<seuil	0,33%	<seuil	<seuil	<seuil	<seuil
Patient 6*	6,2	ND	ND	ND	ND	ND	ND	ND
Patient 7	6,1	<seuil	<seuil	0,33%	<seuil	<seuil	<seuil	<seuil
seuil de détection de la substitution[¶]	-	0,05%	0,03%	0,08%	0,15%	0,18%	0,21%	0,22%

*The number of sequences obtained with the baseline sample was too small in this patient to yield reliable results.

The detection cutoff of each substitution was calculated as the mean maximal error rate for the given substitution plus two standard deviations

Tableau 9 : Prévalence au début du traitement des substitutions amino acidiques connues pour conférer la résistance du VHB à l'adefovir chez

les 7 patients de l'étude HBV Pol, détectés par pyroséquençage et analyse par PyroClass[®] et PyroMute[®]. ND : non déterminé.

Le pyroséquençage à haut débit a été utilisé sur l'ensemble des sérums prélevés avant traitement et à différents temps sous thérapie antivirale. Au total, 119 échantillons provenant des 7 patients ayant développé une résistance à l'adefovir en monothérapie ont été étudiés (15 à 24 échantillons par patient). A partir de ces échantillons, 480 000 séquences (soit 111Mbp) ont été générées, soit 4010 ± 843 séquences par échantillon, d'une longueur moyenne de 382 ± 31 nt après élimination des séquences impropres à l'analyse (soit 10,2% des séquences écartées par les filtres de PyroMute®).

■ Préexistence des mutations de résistance avant le traitement :

Le tableau 9 montre la prévalence des substitutions amino- acides associées à la résistance aux analogues nucléos(t)idiques détectées par pyroséquençage à haut débit chez les 7 patients étudiés. Chez un patient, le nombre de séquences obtenues était trop faible pour atteindre une sensibilité exploitable. Chez 5 des 6 patients restants, la mutation rtA181V/T a été trouvée en faible proportion. L'un d'entre eux présentait également une mutation rtN236T et 2 autres la mutation rtM204V/I en faible proportion. Aucune mutation connue pour augmenter le fitness des mutations rtM204V/I en présence de lamivudine, d'entecavir ou de telbivudine (rtV173L et rtL180M) n'a été trouvée à un taux supérieur au seuil avant le traitement. Un patient présentait une substitution rtT184S/A/I/L, associée à la résistance à l'entecavir lorsqu'elle est présente avec les mutations rtM204V/I et rt L180M.

■ Caractérisation moléculaire de l'échec thérapeutique sous adefovir chez le patient 1 par pyroséquençage à haut débit fondée sur l'analyse des données de séquence par PyroPack®

Le patient 1 a répondu de manière suboptimale à l'adefovir et la charge virale du VHB a augmenté graduellement, de son nadir à M6 jusqu'à la fin du suivi à M24 (Figure 1A à F-

Article 2-Annexe 3). La Figure 1 de l'article 2 montre les dynamiques de la charge virale et des substitutions amino-acidiques de la transcriptase inverse par rapport au variant sauvage sous adefovir, obtenues par pyroséquençage à haut débit et grâce à l'analyse des données par PyroPack®. Les séquences ont été classées par PyroClass®, puis filtrées par PyroMute®. Les dynamiques des variants viraux résistants à l'adefovir portant des mutations de résistance connues (rt181V, rt181T et rtN236T) ont été établies (Figure 1A-Article 2-Annexe 3). En parallèle, les mêmes données ont été analysées par PyroDyn®, qui caractérise les dynamiques de croissance et/ou de décroissance exponentielle des populations virales, sans a priori concernant les mutations recherchées (Figure 1B-Article 2-Annexe 3). La liaison (linkage) a été établie à partir des données de PyroDyn® à l'aide du logiciel PyroLink® (Figure 1C et 1D-Article 2-Annexe 3).

L'analyse des données de pyroséquençage par PyroMute® a révélé la croissance des substitutions amino-acidiques rtA181V, rtA181T et rtN236T à partir du 13^{ème} mois de traitement, c'est-à-dire 5 mois avant le clonage-séquençage. L'analyse par PyroDyn® a détecté la croissance exponentielle de ces mutations, confirmant la capacité du logiciel à détecter les mutations de résistance connues. Cette analyse a apporté de nombreuses informations additionnelles en montrant d'une part la sélection d'autres substitutions amino-acidiques en faible proportions, incluant les mutations rtY245H et rtD238N, d'autre part une vague précoce de sélection de mutations inconnues, incluant les substitutions rtN138K, rtR139K et rtR212T. Cet enrichissement est apparu quelques mois seulement après le début du traitement, de façon transitoire, et ces mutations ont disparu au moment où les mutations de résistances connues ont fait leur apparition.

L'analyse combinée des données de pyroséquençage par PyroDyn® et PyroLink® (Figure 1C et 1D-Article 2-Annexe 3) a permis de caractériser la dynamique des populations virales,

fournissant l'information clinique attendue. La Figure 1C de l'article 2 (Annexe 3) montre la dynamique des variants viraux, la proportion de chaque variant au sein de la quasi-espèce étant représentée à chaque instant, tandis que la Figure 1D de l'article 2 (Annexe 3) montre la même dynamique avec cette fois la quantité de chaque population virale exprimée en valeur absolue, c'est-à-dire en Log UI/mL. Cette dernière représentation est apparue la mieux adaptée pour décrire la dynamique des variants viraux sous traitement par adefovir ; elle a été choisie pour décrire les dynamiques des populations virales des autres patients. La représentation en valeur absolue de charge virale des variants résultant de l'analyse par PyroDyn® et PyroLink® a apporté les informations suivantes chez le patient 1 :

- immédiatement après l'initiation du traitement, une première vague de sélection est survenue, en fait liée au maintien, en valeur absolue, de variants minoritaires présents avant le début du traitement en parallèle à la décroissance des variants sauvages, sensibles à l'adefovir.
- immédiatement après que le nadir eut été atteint (M2), la quantité de virus sauvage a commencé à augmenter de nouveau, tandis que les variants minoritaires initialement sélectionnés diminuaient pour disparaître.
- la première vague de sélection de variants indubitablement résistants à l'adefovir a été détectée à M17 et a atteint un pic à M21/22, lorsque le variant sauvage est devenu indétectable. Cette vague était composée de variants portant des substitutions amino-acidiques connues pour être associées à la résistance à l'adefovir, en particulier les mutations rtN236T (majoritaire) et rtA181T (minoritaire)
- une seconde vague de sélection de variants viraux résistants est ensuite apparue. Ces variants ont progressivement remplacé les variants de la première vague de résistance, probablement grâce à l'acquisition de mutations de fitness et/ou de

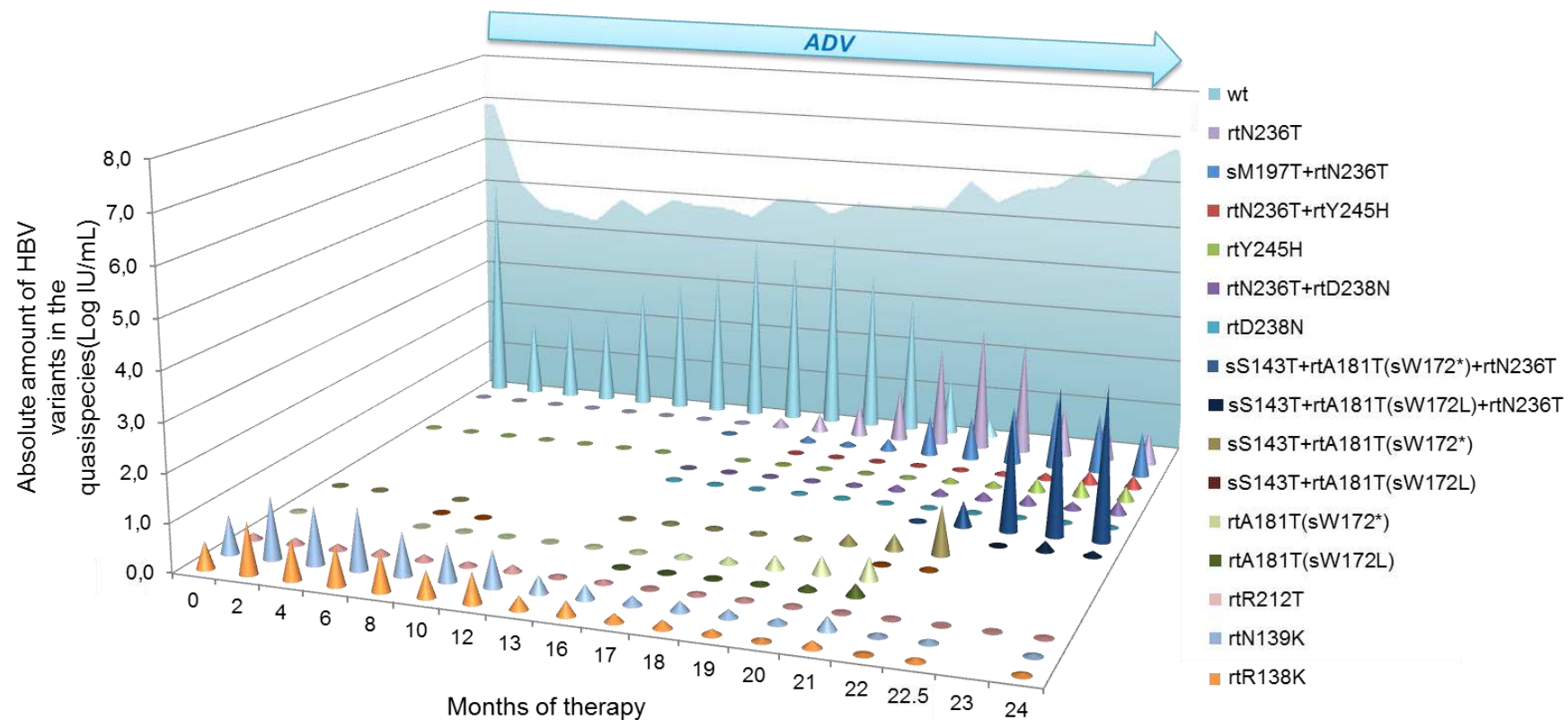


Figure 36 : (Figure 1F de l'article 2) Dynamique des populations virales de VHB au cours du traitement par adefovir chez le patient 1 de l'étude HBV Pol obtenu par analyse par Pyropack®. La charge virale (fond bleu), le traitement reçu et les proportions respectives des différents variants sensibles et résistants sont indiqués en valeur absolue. Les variants sont représentés à la fois avec leurs mutations dans l'Ag HBs et dans le domaine transcriptase inverse de la polymérase.

résistance additionnelle. Les variants sélectionnés incluait, par ordre de fréquence, les substitutions rtN236T+rtA181T, rtY245H, rtN236T+rtY245H, et rtN236T+rtD238N.

En parallèle, l'analyse des changements dans le cadre de lecture de l'Ag HBs a été réalisée à l'aide de PyroPack® (Figure 36 ou Figure 1E et 1F-Article 2-Annexe 3). La figure 1E de l'article 2 (Annexe 3), qui montre la dynamique des substitutions amino-acidiques sur l'Ag de surface, a permis de mettre en évidence la croissance des substitutions associées aux mutations de résistance dans le cadre de lecture chevauchant de la polymérase. D'autres substitutions, sS143T et sM197T, ont été sélectionnées au cours du traitement, bien qu'elles soient silencieuses dans le cadre de lecture de la polymérase. La figure 36 (Figure 1F-Article 2-Annexe 3) montre les quantités, en valeur absolue, des différents variants aux différents temps de suivi. Cette figure montre la liaison et la croissance simultanée des mutations sS143T et rtA181T d'une part, sM197T et rtN236T d'autre part. Ces mutations semblent donc contribuer au gain de fitness de la seconde vague de variants résistants, avec une prépondérance pour le variant sM197T+rtN236T et le variant sS143T+rtA181T+rtN236T.

- Caractérisation moléculaire de l'échec thérapeutique sous adefovir chez les patients 2 à 7 par pyroséquençage à haut débit fondée sur l'analyse des données de séquence par PyroPack®

La comparaison des différentes approches analytiques chez les 6 patients restants a confirmé la supériorité de la combinaison de PyroDyn® et PyroLink® pour l'analyse, suivie de l'expression des résultats en valeur absolue de charge virale. Les résultats sont présentés sur les figures 2A à 2F de l'article 2 (Annexe 3). Aucune représentation de l'Ag HBs n'est présentée car, au contraire du patient 1, aucune substitution n'a été trouvée par PyroPack® dans cette région.

Chez le patient 2 (Figure 2A-Article 2-Annexe 3), les variants sauvages ont progressivement diminué en proportion après la mise en place du traitement par adefovir, puis ont de nouveau rempli l'espace de réplication pendant l'arrêt du traitement d'une durée d'environ un an. Lorsque l'adefovir a été réintroduit quelques semaines plus tard, les variants sauvages ont à nouveau diminué en proportion, et se sont maintenus en plateau à environ 10^4 UI/mL. L'émergence de la résistance a été caractérisée par la sélection simultanée des mutations rtN236T et rtA181V à au 27^{ème} mois. La compétition entre les variants a abouti à la dominance du variant rtA181V, qui était associé au rebond virologique. La multiplication de ce variant a ensuite été partiellement inhibé, mais est resté majoritaire, lorsque la lamivudine a été ajoutée à l'adefovir.

Chez le patient 3 (Figure 2B- Article 2-Annexe 3), l'émergence de variants résistants a été observée au 29^{ème} mois. Elle était caractérisée par l'apparition de variants simples ou de doubles mutants aux positions rt181 et rt236. Chez ce patient, le variant rtN236T est devenu majoritaire et était associé au rebond virologique. Lors de l'ajout de la lamivudine, les variants viraux se sont comportés comme chez le patient 2, avec une inhibition partielle de leur réplication, mais le maintien du variant en tant que population majoritaire.

Le patient 4 a présenté une réponse suboptimale à l'ADV. Son profil de résistance était plus complexe (Figure 2C-Article 2-Annexe 3). Au cours d'une phase de plateau d'environ 20 mois, caractérisée par de faibles fluctuations de la charge virale, le variant sauvage a été progressivement remplacé par un mélange de variants portant une seule mutation (rtY124H ou rtN236T), de doubles mutants (rtY124H+rtN236T) et de triples mutants (rtY124H+rtN236T+rtN238T) répliquant à un faible niveau. Le variant sauvage est ensuite redevenu majoritaire lors d'un arrêt de traitement de 2 mois, puis la résistance s'est développée à nouveau avec un rebond virologique typique associé à la sélection d'un variant

Patient	Augmentation du taux d'ADN du VHB >1 Log ₁₀ au-dessus du nadir	Séquençage direct*	Clonage-séquençage*	Pyroséquençage + PyroMute [®] †	Pyroséquençage + Pyromute [®] + PyroDyn [®] /PyroLink [®] †
Patient 1	17	18	6	0	16
Patient 2	30	27	27	15	27
Patient 3	31	29	9	0	11
Patient 4	21	14	9	2	12
Patient 5	37	37	28	2	28
Patient 6	33	42	20	15	15
Patient 7	28	26	20	7	17
Mediane (IQR)	30 (24-32)	27 (22-33)	20 (9-23)	2 (1-11)	16 (13-22)

Tableau 10 : Date du diagnostic de la résistance du VHB à l'adefovir (en mois après le début du traitement) chez les 7 patients de l'étude HBV

Pol, en fonction de l'analyse réalisée. IQR: interquartile range.

simple rtN236T. La réplication de ce variant a ensuite été partiellement inhibée par la lamivudine. Sous bithérapie, c'est le mélange de variants sélectionné lors de la première phase de traitement qui a émergé et répliqué à un faible niveau.

Le patient 5 (Figure 2D-Article 2-Annexe 3) était caractérisé par un profil de résistance très simple, avec une réponse suboptimale au traitement par adefovir et un rebond virologique au 37^{ème} mois lié à la sélection d'un double mutant (rtA181V+rtN236T).

Le patient 6 (Figure 2E-Article 2-Annexe 3) se distinguait du patient 5 par l'absence de rebond virologique clairement identifiable, malgré la persistance d'une faible virémie et la sélection d'un double mutant (rtN236T+rtN238E) comme population majoritaire.

Finalement, chez le patient 7 (Figure 2F-Article 2-Annexe 3), le rebond virologique était lié à la sélection d'une population majoritaire portant la mutation rtA181V et d'une population double mutante rtA181V+rtF221Y. Les deux variants ont été partiellement inhibés lors de l'ajout de lamivudine, mais sont restés majoritaires.

■ Comparaison des techniques de diagnostic de la résistance à l'adefovir

Le tableau 10 montre, pour chaque patient traité par adefovir, la date du diagnostic de résistance par les méthodes suivantes: (i) augmentation de 1 log au-dessus du nadir de la charge virale du VHB (rebond virologique); (ii) détection d'une mutation de résistance connue par séquençage direct; (iii) détection d'une mutation de résistance connue par pyroséquençage suivi de l'analyse des données par PyroMute®; (iv) détection de l'émergence de mutations de résistance non connues a priori par pyroséquençage suivi d'une analyse par PyroMute® et PyroDyn®/PyroLink®. Chez la plupart des patients, le pyroséquençage à haut-débit a détecté la présence de variants viraux résistants à l'adefovir en très faibles proportions dès le début du traitement. La combinaison du pyroséquençage

et de l'analyse des données par PyroDyn®/PyroLink® a mis en évidence la présence de variants viraux résistants au même moment que le clonage-séquençage, c'est-à-dire plusieurs mois avant le séquençage Sanger et le rebond virologique de la charge virale. Ces deux dernières méthodes étaient les plus tardives à diagnostiquer la résistance.

Etude HBV InVS (ANALYSE PRELIMINAIRE-MANUSCRIT EN PREPARATION par Chevaliez S, Rodriguez C, et coll.)

Cette étude vise à caractériser sur le plan virologique les malades infectés par le VHB, jamais traités, vus pour la première fois dans un centre de référence français dans le cadre de l'étude de surveillance coordonnée par l'InVS, et à mesurer la prévalence des mutations préexistantes conférant la résistance aux analogues de nucléosides et de nucléotides. L'analyse préliminaire a porté sur un échantillon de 111 patients. Les 111 prélèvements ont permis la génération de 901 605 séquences, soit 8125 ± 2812 séquences par échantillon. Après rejet de 13% des séquences par PyroMute®, considérées comme de qualité insuffisante, les séquences analysées avaient pour longueur moyenne 366 ± 76 pb et présentaient une moyenne de qualité PHRED de $31 \pm 2,3$ et un Q30 atteint par en moyenne $70 \pm 6\%$ des séquences. L'analyse des séquences a porté sur les positions connues de résistance aux différents traitements antiviraux du VHB (cf Annexe 1B), dont la présence n'a été considérée que si leur proportion au sein de la quasi espèce virale était supérieure au seuil de détection établi avec les contrôles. Les mêmes contrôles que dans l'étude HBV POL ont été séquencés dans les 2 sens de lecture, ce qui a permis de fixer à 0,2% le seuil sur l'ensemble de la portion génomique étudiée.

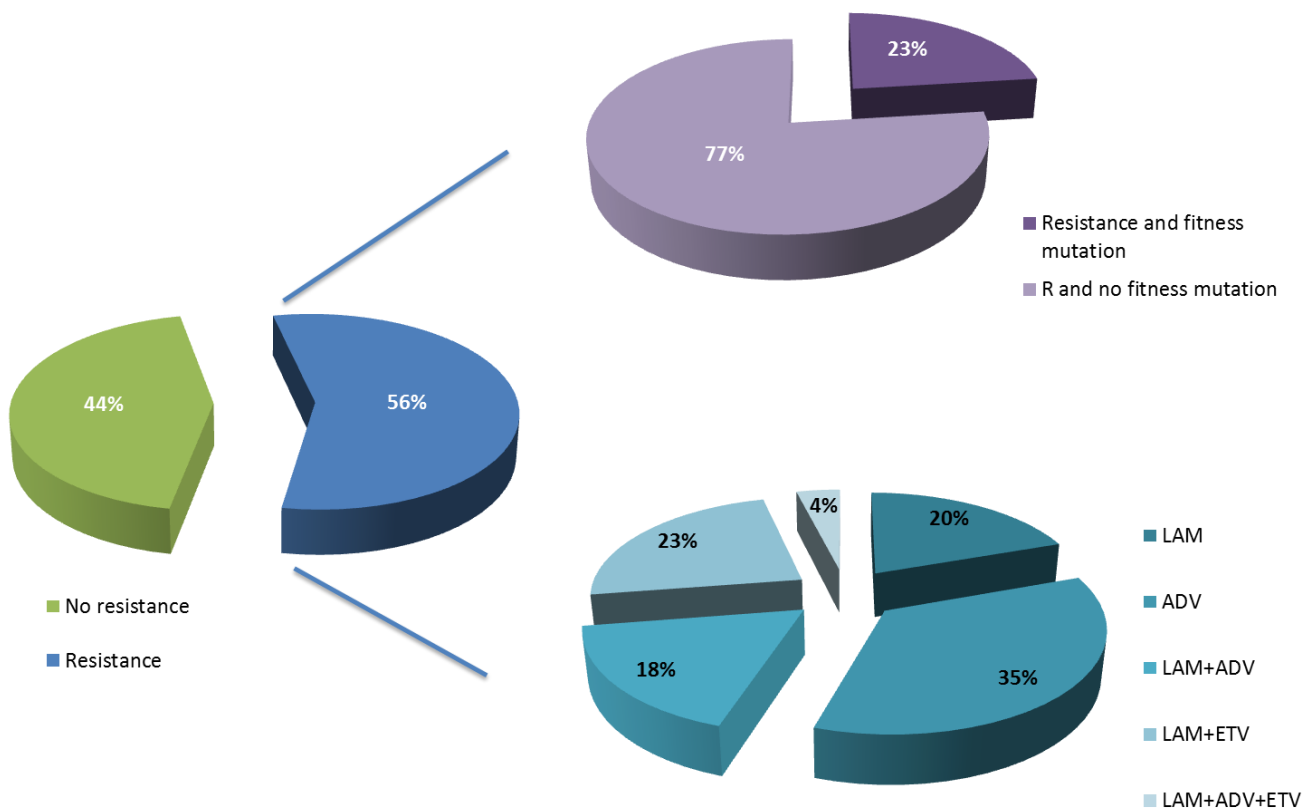


Figure 37 : Distribution des variants du VHB résistants aux différents traitements antiviraux dirigés contre le VHB présents sous la forme de populations virales minoritaires détectables en pyroséquençage. Ces résultats ont été obtenus à partir d'un échantillon de 111 patients de la cohorte HBV InVS. Le camembert de gauche montre la prévalence de patients présentant des variants minoritaires en proportion supérieure à 0.2%. Le camembert en haut à droite présente la proportion de ces patients possédant des mutations de fitness secondaires en plus des variants résistants. Le camembert du bas à droite montre la répartition des mutations de résistance par traitement dans la population de l'étude abritant ces variants minoritaires (56% de la cohorte).

Au total, 56% des patients abritaient des variants viraux résistants en faible proportion et ce, en l'absence de tout antécédent de traitement (Figure 37). Vingt-trois pour cent d'entre eux abritaient également des variants présentant des mutations de fitness (mutations en positions 184, 202 ou 250 associées à la résistance à l'entecavir), mais aucune d'entre elles n'était associée à une mutation de résistance sur le même brin d'ADN viral. Chez 9 patients sur 111 (8%), les variants résistants représentaient plus de 5% de la quasi espèce virale ; 2 d'entre eux étaient porteurs de mutations de résistance à l'adefovir (A181T), 7 de mutations de résistance à la lamivudine (6 M204V/I et 1 M204V/I+L180M). Enfin, parmi les patients abritant des variants viraux résistants, 20% ne présentaient que des mutations de résistance à la lamivudine, 35% que des mutations de résistance à l'adefovir, 18% des mutations de résistance à la lamivudine et à l'adefovir, 23% des mutations de résistance à la lamivudine et à l'entecavir et 4% des mutations de résistance aux 3 médicaments.

Les séquences des autres patients de l'étude sont en cours de génération. Un article décrivant l'ensemble des paramètres virologiques et leurs relations ainsi que la prévalence de la résistance primaire analysée par pyroséquençage à haut débit et PyroPack® est en cours de rédaction.

Etude HBV RNase H (ARTICLE 3)

L'article 3, en cours de finalisation, combine les données (présentées ici) d'analyse de la séquence de la RNase H de 80 patients infectés par un VHB de génotype D, issus d'une étude de cohorte grecque incluant des patients ayant une hépatite chronique B antigène HBe négative, et des études de modélisation de cette portion de la polymérase virale et de mapping des mutations observées réalisées par l'équipe de Christophe Combet à l'Institut de

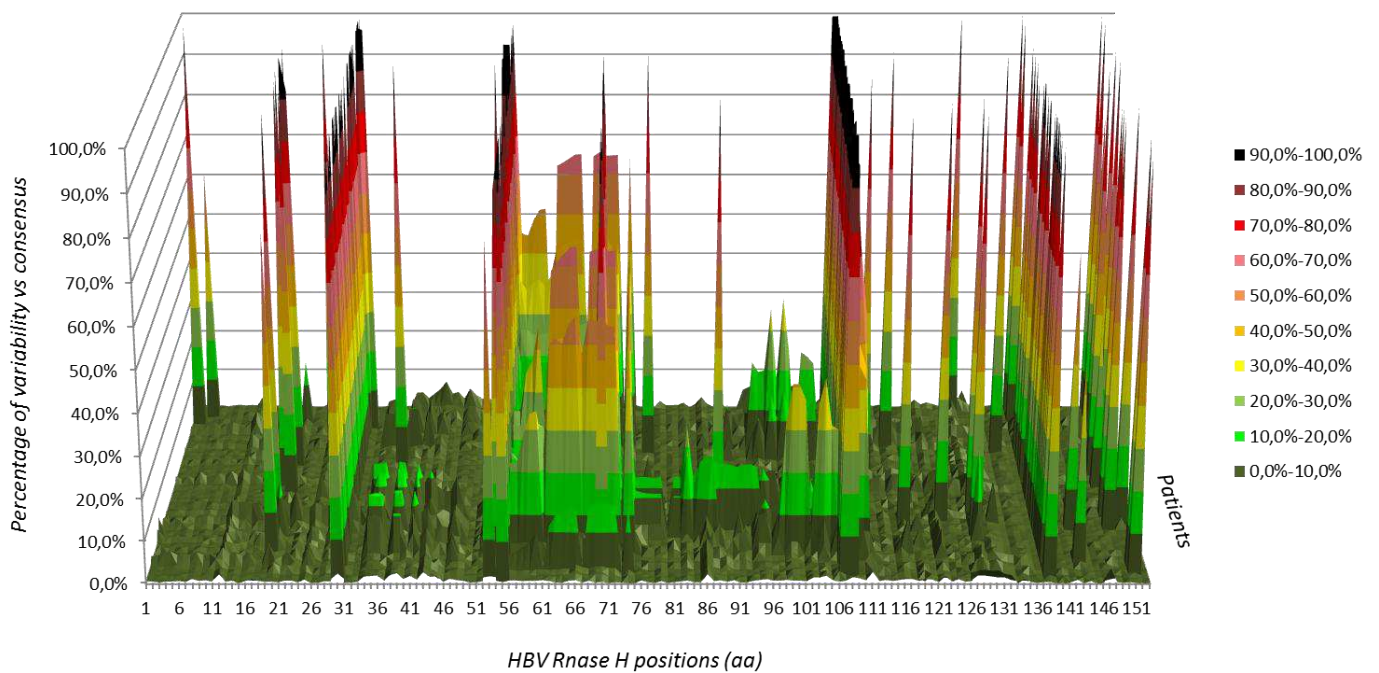


Figure 38 : Cartographie de la variabilité génétique de la RNase H du VHB infectant les patients de l'étude HBV RNaseH. Cette carte génétique fait apparaître des domaines de variabilité et de non variabilité, laissant supposer des contraintes génétiques imposées par la fonctionnalité de la structure.

Biologie et Chimie des Protéines de Lyon. L'ADN du VHB a été amplifié et séquencé avec une profondeur suffisante chez 73 des 80 patients sélectionnés. Au total, 958 000 séquences ont été générées, soit 12949 séquences par patient en moyenne (extrêmes : 4350-14165). La longueur moyenne des séquences étudiées était de 302 pb (221-369), avec un score de qualité PHRED moyen de 31,0 (extrêmes : 30,2 à 31,7). Le Q30 (séquences jugées de très bonne qualité) a été atteint par $63,8 \pm 2,6\%$ des séquences. Après analyse par PyroPack® d'un peu plus de 900 000 séquences, la variabilité de la RNase H des souches virales de chaque patient a été étudiée par rapport à la séquence consensus de tous les patients. Celle-ci a été générée en identifiant l'acide aminé le plus fréquent à chaque position à partir du lot complet de données. Une carte génétique de la variabilité a ainsi pu être établie pour l'intégralité de la RNase H (Figure 38).

La carte génétique de la variabilité de la RNase H faisait apparaître des domaines de variabilité et de non variabilité, laissant supposer l'existence de contraintes génétiques imposées par la fonctionnalité de la structure. Ces données ont été utilisées dans un premier temps pour améliorer le positionnement des différents domaines et acides aminés indispensables à l'activité catalytique de la RNase H par modélisation par homologie avec les autres RNase H existantes, en particulier celles d'HIV-1 (PDB : 1T05) (Tuske *et al.* 2004), d'*Escherichia coli* (PDB : 1RDB) (Katayanagi *et al.* 1993; Katayanagi *et al.* 1993) et d'*Homo sapiens* (PDB : 2QKK) (Nowotny *et al.* 2007). Le positionnement des 4 acides aminés catalytiques D-E-D-D des RNase H de type I n'a jamais pu être affirmé avec certitude dans les études précédentes (Chen *et al.* 1994; Chen and Marion 1996). La structure de la RNase H du VHB, prédite par modélisation et présentée en superposition avec les structures de RNase H du VIH, d'*E. coli* et d'*H. sapiens*, identifiait les acides aminés nécessaires à l'activité catalytique aux positions 689, 718, 737, 817, avec cependant une incertitude entre les

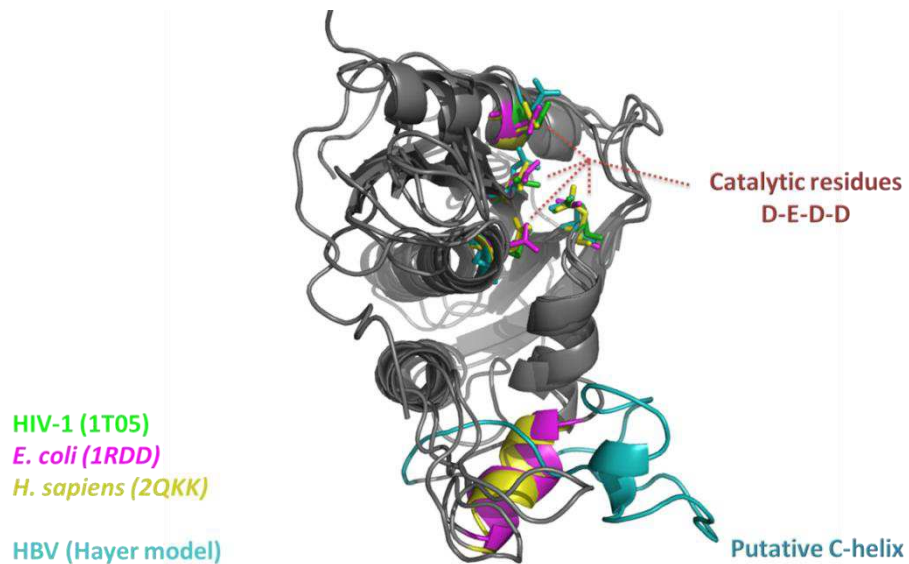


Figure 39 : Superposition des structures des RNase H du VHB (modèle par homologie, Hayer *et al.* 2012), d'E.coli, d'H. sapiens et du VIH. Les résidus catalytiques D-E-D-D sont positionnés sur la structure, ainsi qu'un domaine spécifique du VHB : le domaine putatif C en hélice alpha.

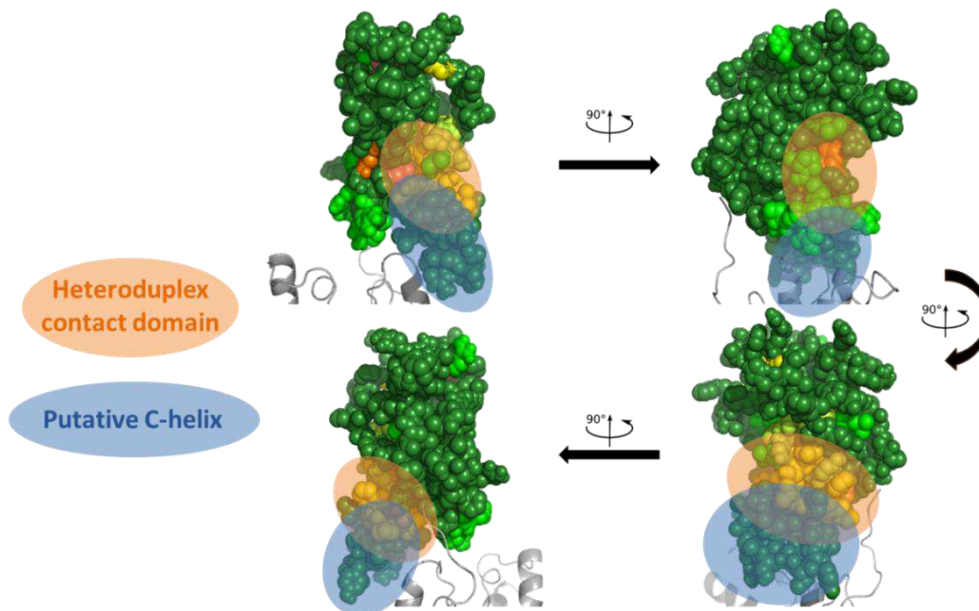


Figure 40 : Représentation 3D sous 4 angles de la variabilité de la structure de la RNase H du VHB (rouge : fortement variable au vert : faiblement variable). Le domaine putatif C du VHB est faiblement variable, à l'inverse du domaine de contact avec l'hétéroduplex ADN/ARN.

positions 807 et 817 pour le 4^{ème} acide aminé catalytique (Figure 39). Les données de variabilité montraient des taux de conservation moyens des acides aminés D689 et E718 de respectivement 99,9% et 96,2%. Le 3^{ème} site catalytique présentait une variabilité plus importante, avec 12 patients portant une mutation D737A représentant 15 à 50% de la quasi-espèce. 70 des 73 patients avaient un D en position 807, contre seulement 51 patients en position 817, favorisant la position 807 pour le 4^{ème} acide aminé catalytique. Trois patients sur 73 présentaient une mutation D807V. L'un d'entre eux présentait également une mutation D817V, excluant la possibilité que cet acide aminé se substitue à la mutation D807V pour maintenir une fonctionnalité de la RNase H. Plus largement, aucune substitution par un acide aminé acide n'a été trouvée chez ce patient sur l'ensemble de la RNase H, malgré une charge virale du même ordre que celle observée dans le reste de la cohorte (respectivement 3,5 Log UI/mL et 4,26±1,83 Log UI/mL).

Les données de variabilité naturelle ont été exploitées pour identifier les zones de variabilité et de conservation structurales (Figure 40). Quatre domaines conservés ont été identifiés aux positions 683-698, 713-730, 755-785, et 790-816, ainsi qu'un domaine hautement variable à la position 731-754. Les RNase H de type 1 sont typiquement organisées en 4 domaines incluant les 4 acides aminés catalytiques et un domaine fonctionnel appelé domaine de protrusion basique dont la fonction est d'attirer l'hétéroduplex vers l'enzyme (Cerritelli and Crouch 2009). Les 2 premiers domaines conservés aux positions 683-698 et 713-730 incluaient effectivement les 2 premiers acides aminés du site catalytique, tandis que le dernier incluait le dernier acide aminé catalytique (si on le considère comme positionné en 807). Cependant, le 3^{ème} acide aminé catalytique et la position 817 qui peut également correspondre au 4^{ème} acide aminé catalytique étaient situés au sein de domaines variables. Le domaine de protrusion basique a été localisé sur la structure prédite par homologie sur le

3^{ème} domaine conservé, selon une organisation typique de RNase H. Enfin le domaine variable identifié à l'aide de la carte génétique correspondait au domaine de contact avec l'hétéroduplex ADN/ARN.

Un premier domaine d'intérêt, le « domaine putatif C » ou domaine de protrusion basique, avait la particularité structurale d'être spécifique du VHB et peu variable. Cette spécificité d'espèce en fait une cible intéressante pour un usage thérapeutique, car des médicaments dirigés contre ce domaine pourraient être spécifiques, sans effet sur les autres RNase H, évitant une toxicité potentielle. La conservation de ce domaine pourrait également permettre aux médicaments le ciblant d'atteindre une barrière génétique élevée. Le second domaine d'intérêt, le domaine de contact avec l'hétéroduplex ADN/ARN, est très variable et ne représente donc pas une cible prometteuse du fait du risque élevé de sélection de variants viraux résistants.

Etude HBV PreC (ANALYSE PRELIMINAIRE- 2 MANUSCRITS EN PREPARATION par Rodriguez C, et coll.)

La région PreC/C du génome du VHB a été étudiée en pyroséquençage à haut débit dans 337 échantillons provenant de 157 patients pédiatriques avec une hépatite chronique B Ag HBe-positive inclus dans un essai thérapeutique évaluant l'efficacité de l'adefovir contre placebo. Au total, 1 280 000 séquences ont été générées, soit en moyenne 3774 ± 781 séquences par échantillon, de longueur moyenne 213 ± 22 pb, avec un score de qualité PHRED moyen de $33,6 \pm 3,5$, et un Q30 (séquences jugées de très bonne qualité) atteint pour $73,8 \pm 7,1\%$ d'entre elles. Après analyse par PyroPack® d'un peu plus d'1 050 000 séquences (17,7% d'échec), la relation entre la présence de mutations du domaine PreC/C au début du traitement par

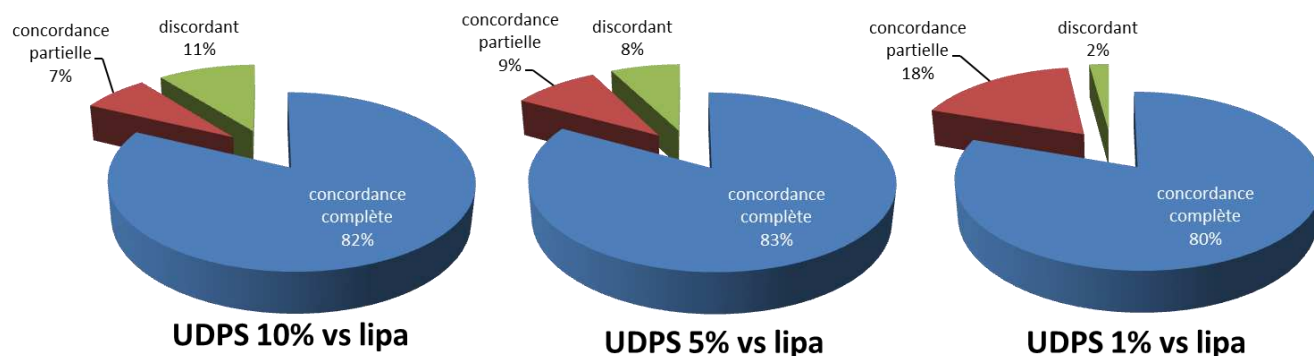


Figure 41 : Concordance entre la méthode INNO-LIPA et le pyroséquençage (UDPS) pour l'identification des mutations de la région PreC/C, utilisant des seuils de positivité de 1%, 5% ou 10% pour le pyroséquençage. La concordance partielle (rouge) indique le pourcentage d'échantillons (N=240) pour lesquels une partie des mutations trouvées ne l'ont été qu'avec le pyroséquençage, tandis que les discordants correspondent à ceux pour lesquels une partie des mutations n'a été trouvée qu'en INNO-LIPA.

Patient	LIPA	UDPS 10%	UDPS 5%	UDPS 1%
Sensibilité	36.9%	23,1%	29.2%	49.2%
Spécificité	67.7%	80.6%	77.4%	64.5%
VPP	54.5%	55.6%	57.6%	59,3%
VPN	50.6%	50%	51.1%	54.8%
Q de Yule	0.10	0.11	0.17	0.30

■ : négligeable (0,01-0,09) ■ : léger (0,10-0,29) ■ : modéré (0,30-0,49) ■ : fort (0,50-0,69) ■ : très fort (0,70-1)

Tableau 11 : Capacité des tests INNO-LIPA PreC/C et pyroséquençage (UDPS) aux seuils de 10, 5 et 1% de prédire la séroconversion HBe. Le Q de Yule, qui mesure l'intensité de la liaison entre la présence de mutations PreC/C et la séroconversion HBe, indique que la présence de ces mutations a une faible capacité de prédiction de la séroconversion.

adefovir et l'efficacité de celui-ci a été étudiée. Les résultats de pyroséquençage ont également été comparés aux résultats obtenus sur 240 des 337 échantillons séquencés par une technique de recherche de mutations de la région PreC/C fondée sur l'hybridation inverse (INNO-LIPA).

Comparaison des techniques INNO-LIPA et de pyroséquençage à haut débit pour l'analyse des mutations de la région PreC/C

L'étude a permis d'évaluer une technique qualitative (INNO-LiPA) par rapport à une technique quantitative (pyroséquençage) pour la détection des mutations de la région PreC/C. La capacité de déceler ou non une mutation connue au-dessus d'un seuil donné par pyroséquençage a été comparée à la capacité de détection de la technique INNO-LiPA. La figure 41 montre les pourcentages de concordance complète (toutes les mutations étaient trouvées de façon identique par les 2 techniques), de concordance partielle (certaines mutations n'étaient pas détectées par la technique INNO-LIPA alors qu'elles étaient détectées par le pyroséquençage) ou de discordance (certaines mutations étaient présentes en INNO-LIPA, mais pas en pyroséquençage) pour différents seuils de positivité fixés pour le pyroséquençage. Le minimum de discordance (2%) a été obtenu pour un seuil de 1% en pyroséquençage. La discordance a été observée pour 7 échantillons, dont 3 pour lesquels les mutations ont été détectées entre 0,5 et 1%, 3 entre 0,1 et 0,5%, et 1 échantillon qui restait discordant quel que soit le seuil de positivité. Pour une valeur seuil de 1%, le pourcentage de concordance partielle augmentait fortement, suggérant que nombre de mutations peuvent être détectées en pyroséquençage si cette valeur seuil est utilisée, sans être détectées par la technique INNO-LIPA. La comparaison des 2 techniques sur un résultat « muté en PreC »

versus « non muté en PreC » par un test de χ^2 de McNemar, a montré une perte de différence entre les deux techniques parallèle à l'augmentation de la valeur du seuil de pyroséquençage, qui n'était plus significative au-delà de 5%. Ces résultats montrent la meilleure sensibilité de la technique de pyroséquençage à haut débit pour la détection des mutations de la région PreC/C.

Prédiction de la réponse à l'adefovir par les mutations PreC/C avant le début du traitement

En comparant les différentes méthodes (pyroséquençage à 10% de seuil, pyroséquençage à 5%, pyroséquençage à 1% et INNO-LIPA) (Tableau 11), nous avons observé que la sensibilité, la spécificité, la valeur prédictive positive et la valeur prédictive négative sur la séroconversion HBe étaient supérieures avec la technique de pyroséquençage à 1% (Tableau 11). Le Q de Yule, qui mesurait l'intensité de la liaison entre la séroconversion et la présence de variants mutés au départ, montrait une intensité de liaison modérée, avec un Q à 0,3 pour le pyroséquençage à 1% et une liaison faible pour les autres méthodes. Le risque relatif associant la présence de variants minoritaires détectés par pyroséquençage au seuil de 1% avant le début du traitement et la séroconversion ultérieure sous adefovir était de 1,76 (intervalle de confiance à 95% : 0,86-3,59 ; $p = 0,11$). Ces résultats ne suggéraient donc pas que les chances de séroconversion soient plus élevées en présence de variants PreC/C minoritaires au début du traitement, même si ces variants étaient détectés par une technique sensible fondée sur le pyroséquençage.

La relation entre la décroissance de la charge virale maximale et la présence de variants PreC/C avant le début du traitement a également été étudiée. Un nuage de points

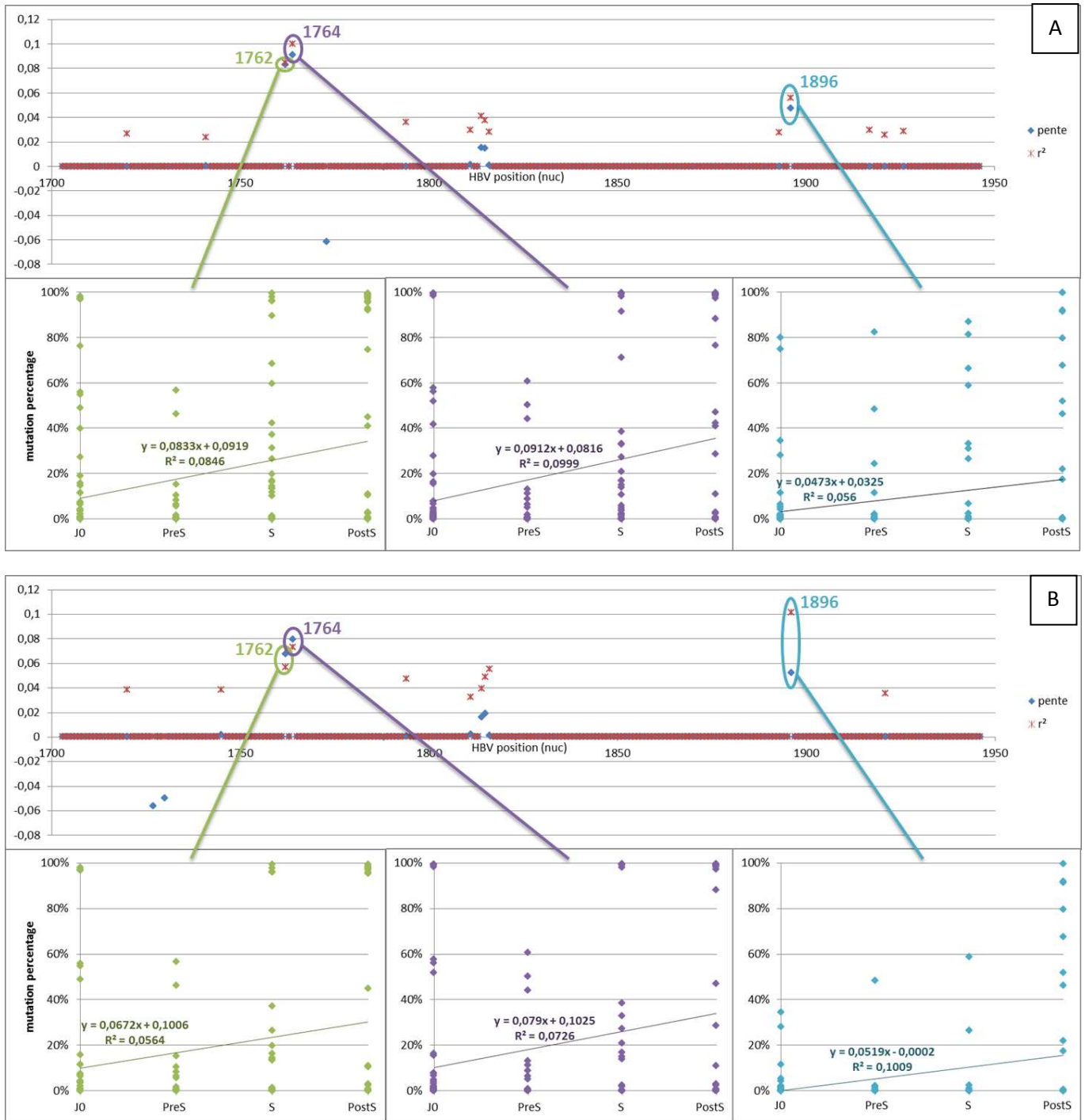


Figure 43 : Relation entre dynamique évolutive des mutations sur l'ensemble du domaine preC/C (graphique du haut en A et B) ou des mutations d'intérêt (1762, 1764 et 1896) chez les patients séroconvertisseurs (A) ou séroconvertisseurs avec charge virale négative (B) sous traitement par adefovir au temps J0, pré-séroconversion (PreS), séroconversion (S) et post séroconversion (PostS).

caractérisant la relation entre la décroissance de la charge virale et le pourcentage de mutations trouvées à une position donnée a été généré. Pour chacun des graphiques, une droite de régression a été tracée, et un test de pente et un calcul du coefficient de détermination r^2 ont été réalisés. Cette approche a permis de montrer qu'aucune mutation testée n'était prédictive de l'amplitude de la décroissance de la charge virale sous traitement.

Finalement, la relation entre la présence de mutations PreC/C avant le traitement et le succès thérapeutique, défini comme l'association d'une séroconversion HBe et d'une charge virale indétectable, a également été étudiée. Dans ce cas, le Q de Yule montrait une liaison négligeable ($<0,09$), quelle que soit la technique employée. Au total, ces résultats montraient que la présence ou l'absence de mutations dans la région PreC/C n'avait pas d'impact sur l'issue du traitement.

Dynamique des populations virales portant des mutations PreC/C sous traitement antiviral

La dynamique des populations virales portant des mutations PreC/C a été étudiée chez 2 groupes de patients de l'étude: les séroconvertisseurs HBe quel que soit le statut de la charge virale à l'issue du traitement (Figure 43A) et les patients traités avec succès qui ont obtenu une négativation de la charge virale et une séroconversion HBe (Figure 43B). Pour l'ensemble des mutations observées aux différentes positions de la portion de génome du VHB étudiée, une droite de régression a été tracée en fonction du temps (pré-traitement, 6 mois pré-séroconversion, séroconversion et 6 mois post séroconversion) et du pourcentage de mutations observés chez les patients sous traitement. Le coefficient de détermination (r^2) et la pente ont été calculés pour chacune des positions et pour chacun des 2 groupes de patients. Chez les patients séroconvertisseurs (Figure 43A), les mutations qui présentaient la

plus forte pente associée aux meilleurs coefficients de détermination étaient celles présentes en position 1762, 1764 et 1896, qui sont également les positions connues de mutations PreC/C. Cependant, les pentes associées à ces mutations (respectivement 8,3%, 9,1% et 4,7% par unité de temps en moyenne) étaient associées à des coefficients de détermination faibles, suggérant que cette variable est faiblement liée à l'évènement « séroconversion ». Si l'on considère les patients ayant obtenu un succès thérapeutique, les mêmes mutations avaient la plus forte croissance et le meilleur coefficient de détermination (Figure 43B). Cependant, ces coefficients (respectivement 5,6%, 7,3% et 5,2% pour les mutations 1762, 1764 et 1896) étaient moins élevés que chez les patients séroconvertisseurs, ce qui suggérait que l'effet observé était essentiellement lié à la séroconversion et non à la négativation de la charge virale.

Etude HCV Prove 2 (ANALYSE PRELIMINAIRE-MANUSCRIT EN PREPARATION par Chevaliez S, Rodriguez C (participation égale), et coll.)

18 des patients inclus dans l'essai de Phase II PROVE2 ont été étudiés. 74 échantillons ont été prélevés avant et pendant le traitement par bithérapie par telaprevir et interféron, bithérapie par interféron et ribavirine, ou trithérapie par telaprevir, interféron et ribavirine et l'ARN viral a été séquencé par pyroséquençage à haut débit. Plus de 1 666 000 séquences ont été obtenues, soit en moyenne 22 525 séquences par échantillon, de longueur moyenne 323 ± 61 pb, avec un score de qualité PHRED moyen de $26,6 \pm 1,6$, le Q30 (séquences jugées de très bonne qualité) étant atteint pour $41,1 \pm 5,6\%$ des séquences (plus bas niveau obtenu pour l'ensemble des études réalisées). Après analyse par PyroPack® d'un peu plus de 910 000 séquences (45% d'échec), la présence de mutations de résistance connues aux

Patient	IFN	RBV	TVR	statut	geno	nb seq	V36A/M	T54A/S	V55A	Q80R/K	R155K/T/ Q	A156S/T/ V	D168A/V/ T/H	I170A/T
1				NR	1a	8405		90.0%				0.4%		0.5%
2				NR	1a	4101						1.1%		
3				Rel	1b	609					0.5%	0.5%		
4				Rel	1b	877		29.4%				1.3%		
5				Rel	1a	3762						2.9%		
6				Rel	1b	3825	4.2%							
7				SVR	1a	4643		11.1%		0.7%				
8				SVR	1a	4190						0.5%		
9				SVR	1a	1338					0.6%	1.8%		
10				SVR	1a	1103					0.6%			
11				Rel	1a	8324			100.0%		6.0%	3.2%		
12				SVR	1b	2685								
13				SVR	1b	1572								0.8%
14				NR	1b	2259								
15				SVR	1b	3252					0.4%			
16				SVR	1a	17981			1.3%	0.5%	7.8%			
17				SVR	1a	6338		47.4%				0.4%		
18				SVR	1b	11009		20.0%				0.4%		

NR : Non responder. Rel : relapser. SVR : sustained virological response

Tableau 12 : Prévalence des principales mutations de résistance aux inhibiteurs de protéase du VHC connues chez les 18 patients de l'étude Prove 2 répartis en 3 groupes, (i) telaprevir (TVR)+interféron (IFN), (ii) IFN+TVR+ribavirine (RBV), (iii) IFN+RBV, et selon l'issue du traitement : (i) réponse virologique soutenue (SVR), (ii) répondeur-rechuteur (Rel) ou (iii) non répondeur (NR).

inhibiteurs de protéases (cf ANNEXE 1A) ont été étudiées avant le traitement et leurs dynamiques sous traitement ont été caractérisées.

Mutations de résistance aux inhibiteurs de protéases préexistantes

Le tableau 12 présente les réponses aux traitements et le taux de présence des mutations de résistance connues aux inhibiteurs de protéase du VHC pour chacun des patients, répartis en 3 groupes selon le traitement reçu. Chez pratiquement tous les patients, des mutations de résistance aux inhibiteurs de protéases du VHC étaient présentes en l'absence de toute exposition aux médicaments, sous la forme de populations virales minoritaires. Il n'existait pas de relation entre la présence de ces mutations au départ du traitement et l'issue du traitement, confirmant le rôle prépondérant de la réponse à l'interféron sur celui-ci.

Dynamique des populations virales résistantes au cours du traitement

L'analyse par PyroDyn® a permis de caractériser les dynamiques des mutations potentiellement impliquées dans l'émergence de la résistance des patients en échec thérapeutique pour lesquels des prélèvements séquentiels étaient disponibles. Deux exemples de cinétiques sont décrits, l'un correspondant à un patient non répondeur : patient 1 (Figure 44A et B), l'autre à un patient répondeur-rechuteur : patient 11 (Figure 44C).

Le patient 1 a reçu 85 jours de traitement par telaprevir et interféron pégylé sans ribavirine ; des prélèvements ont été réalisés régulièrement jusqu'à plus de 2 ans après l'arrêt du traitement. Durant la phase thérapeutique (Figure 44A), deux groupes de mutations ont été sélectionnés. Le premier, composé des mutations H28Q, Y52C, S54T et R155K, a émergé très

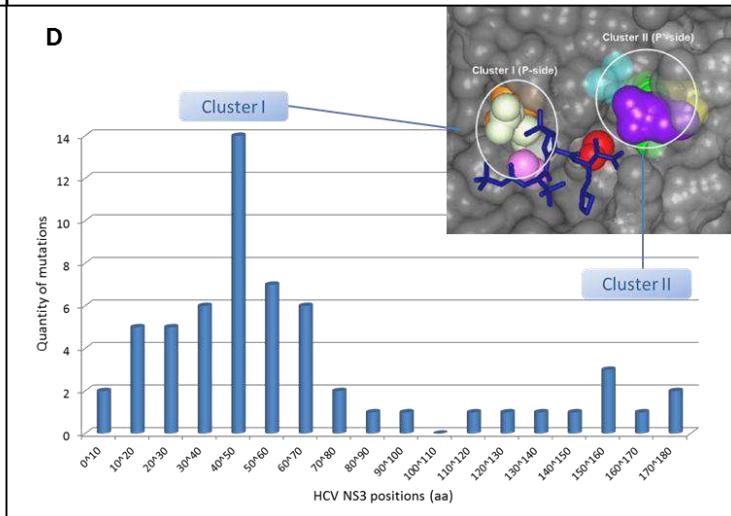
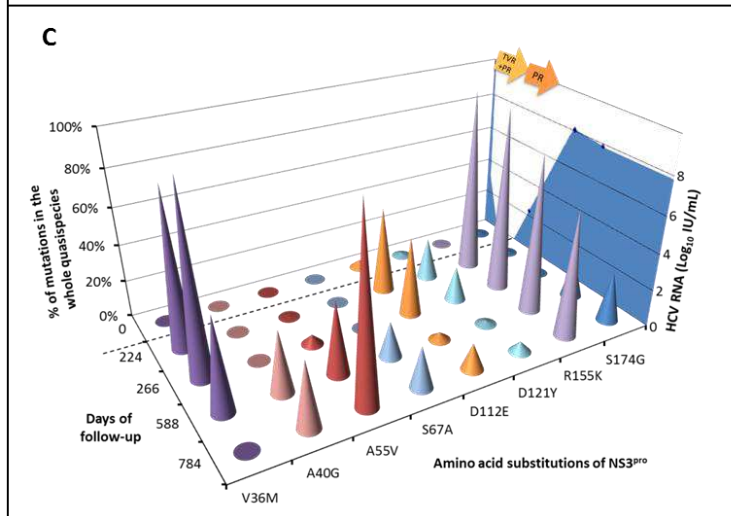
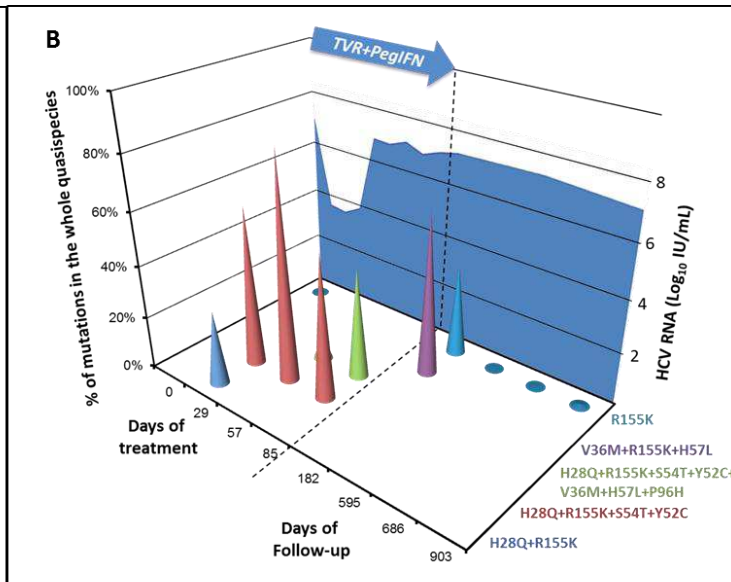
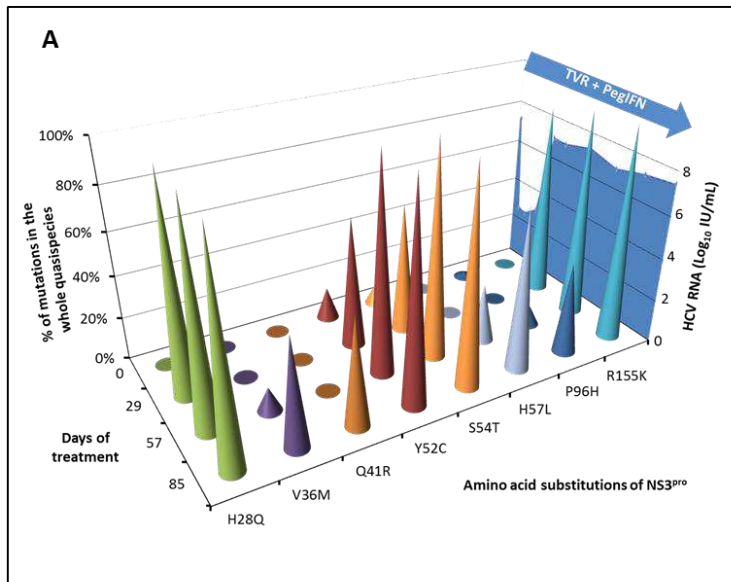


Figure 44 : Dynamiques des mutations sélectionnées comme potentiellement résistantes au telaprevir étudiées par pyroséquençage et analysées par PyroMute® au cours du traitement par telaprevir et interféron pégylé chez le patient 1 (A) et après arrêt du traitement par telaprevir, interféron et ribavirine chez le patient 11 (C). La figure B, décrit la dynamique des variants viraux chez le patient 1 à l'aide de PyroLink®, pendant et après traitement. L'étude des cinétiques des variants chez les patients de l'étude HCV Prove 2 a permis de créer un répertoire des mutations positionnées sur le domaine NS3 du VHC (D). Il est possible d'observer 2 clusters de mutations correspondant à ceux décrit auparavant comme les zones de fixations des inhibiteurs de protéase de type kétoamide (molécule en bleue) c'est-à-dire de la même famille que le telaprevir (Tong *et al.* 2008)

rapidement, avec des proportions dépassant 80% de la quasi-espèce seulement 29 jours après le début du traitement et atteignant une prévalence proche de 100% après 57 jours de traitement. Ce groupe était composé de mutations connues pour conférer une baisse de sensibilité aux inhibiteurs de protéase, comme la S54T (diminution de sensibilité modérée) ou la R155K (forte diminution de sensibilité), et de mutations inconnues, qui seront caractérisées ultérieurement par un test phénotypique. Le 2^{ème} groupe de mutations, composé des mutations V36M, H57L et P96H, est apparu plus tardivement, en proportions majoritaires uniquement en fin de traitement. Comme le 1^{er} groupe, ce dernier comprenait des mutations de résistance connues (V36M) et inconnues, qui seront caractérisées sur le plan phénotypique. L'utilisation de PyroLink® (Figure 44B) a montré que le profil dynamique de la quasi espèce virale était caractérisé par la sélection de 3 vagues successives de variants dérivant les uns des autres en se complexifiant par l'acquisition de nouvelles mutations de fitness ou de résistance. Lors de l'arrêt du traitement, le suivi trop espacé du patient n'a malheureusement pas permis de décrire avec précision la décroissance des variants sélectionnés au cours du traitement. Cependant, il a été possible d'observer une diminution progressive de la complexité de la quasi espèce virale, aboutissant à un retour à un génotype majoritairement sauvage, avec un variant minoritaire résistant (R155K), situation identique à celle observée avant le début du traitement.

Le patient 11 a reçu une trithérapie par telaprevir, interféron pégylé et ribavirine et a rechuté au terme de ce traitement au cours duquel sa charge virale était restée nulle (Figure 44C). Au moment de la rechute, la composition du variant majoritaire avait changé par rapport à l'étude pré-thérapeutique, avec un accroissement de la présence de populations virales résistantes. L'arrêt du traitement s'est accompagné d'une décroissance progressive de la proportion de populations virales mutées. Les mutations sélectionnées comprenaient

	Consensus/ /g2p % (n)	Consensus/ Pyrotrop[®] % (n)	UDPS/ g2p % (n)	UDPS/ Pyrotrop[®] % (n)
Consensus/ g2p	100% (102)	76.5% (78)	77.5% (79)	71.6% (73)
Consensus/ Pyrotrop[®]		100% (102)	83.3% (85)	94.1% (96)
UDPS/g2p			100% (102)	86.3% (88)
UDPS/ Pyrotrop[®]				100% (102)

Tableau 13 : Concordance entre les 4 méthodes utilisées pour analyser les données de génotropisme HIV chez les 102 patients pour lesquels des données de pyroséquençage étaient disponibles. Deux sources de données ont été utilisées : (i) la séquence consensus, qui correspond à la séquence moyenne obtenue à l'aide de l'ensemble des séquences disponibles pour un échantillon ; cette séquence consensus est comparable à celle qui pourrait être obtenue par séquençage direct ; (ii) l'intégralité des données de pyroséquençage (UDPS). L'analyse des données a été réalisée à l'aide de 2 logiciels : geno2pheno[454] (Max Planck Institut, Saarbrücken, Allemagne) et le logiciel maison exploitant les données qualité obtenues lors du séquençage PyroTrop[®].

des mutations connues pour conférer un niveau moyen à élevé de résistance (V36M, R155K), et des mutations non majoritaires dont le rôle reste à préciser dans un test phénotypique (D112E, D121Y). Un autre groupe de mutations, ayant crû en prévalence lors de l'arrêt du traitement, pourrait correspondre à des mutations de fitness permettant une meilleure adaptation des souches virales aux conditions de réplication sans traitement antiviral.

L'analyse de l'ensemble des profils dynamique des patients a permis de construire un répertoire des mutations sélectionnées chez les différents patients à l'aide de PyroDyn® (figure 44D). Lorsque les mutations étaient répertoriées par position sur le domaine NS3 du VHC, il était possible d'observer 2 clusters préférentiels de mutations sélectionnés, qui correspondent aux domaines de fixation des inhibiteurs de type kétoamide (famille du telaprevir) (Tong *et al.* 2008).

Du fait de la qualité insuffisante des séquences générées dans cette étude, certains séquençages sont en cours de répétition afin de finaliser l'article décrivant la préexistence des mutations de résistance avant le début du traitement et les dynamiques des populations virales au cours de l'étude Prove2.

Etude HIV génotropisme

Les données de pyroséquençage avant introduction du maraviroc ont été obtenues chez 84 patients de l'étude HIV génotropisme, infectés par le VIH en échec de trithérapie avec une indication de traitement par le maraviroc. 18 patients supplémentaires ont été étudiés en pyroséquençage, mais aucune donnée n'était disponible concernant la réponse au traitement à M1 de ces patients. Enfin, 11 patients n'ont pu être étudiés en raison d'une

	Sensibilité			Spécificité			VPP			VPN		
	0,5 log	1,0 log	2,0 log	0,5 log	1,0 log	2,0 log	0,5 log	1,0 log	2,0 log	0,5 log	1,0 log	2,0 log
Réponse virologique												
Consensus/g2p	78,3%	77,6%	76,5%	20,0%	16,7%	17,9%	87,1%	83,9%	62,9%	11,8%	11,8%	29,4%
Consensus/Pyrotrop[®]	95,6%	95,8%	94,6%	10,0%	8,3%	3,6%	88,8%	86,3%	66,3%	25,0%	25,0%	25,0%
UDPS/g2p	82,4%	81,9%	80,4%	40,0%	33,3%	21,4%	91,0%	88,1%	67,2%	23,5%	23,5%	35,3%
UDPS/Pyrotrop[®]	91,9%	91,7%	91,1%	30,0%	25,0%	14,3%	90,7%	88,0%	68,0%	33,3%	33,3%	44,4%

Tableau 14 : Sensibilité, spécificité, valeur prédictive positive (VPP) et valeur prédictive négative (VPN) des 4 approches utilisées pour prédire la réponse au maraviroc avec 3 degrés différents de magnitude de réponse virologique à M1 (0,5, 1, et 2 log de décroissance de la charge virale) à partir des données de variabilité des 84 patients de l'étude obtenues à J0.

impossibilité d'obtenir une amplification de bonne qualité. L'étude de comparaison des techniques a donc porté sur 102 patients et l'étude de prédictivité de la réponse au traitement par maraviroc sur 84 patients.

Le pyroséquençage a permis d'obtenir 626 874 séquences, parmi lesquelles 539 096 valides pour l'analyse par geno2pheno[454] (taux de rejet 14,3%), soit en moyenne 5285 séquences par échantillon. En parallèle, 461 044 séquences ont été validées par PyroTrop® (taux de rejet : 26,6%), soit en moyenne 4520 séquences par échantillon.

Comparaison des méthodes d'étude du génotropisme

Le tableau de concordance des techniques utilisées montre que la meilleure concordance a été obtenue pour les techniques utilisant PyroTrop® (Tableau 13). Le Tableau 14 et la Figure 45 décrivent la performance de chaque technique pour la prédiction du génotropisme HIV. Celle-ci était fondée sur l'étude de la séquence de la boucle V3 du VIH à J0, dont la valeur prédictive sur la réponse au traitement par maraviroc a été étudiée. Cette réponse a été évaluée à 3 niveaux, correspondant respectivement à une décroissance de la charge virale de 0,5, 1 ou 2 Log UI/mL à M1. Quel que soit le niveau de réponse considéré, les 2 méthodes utilisant Pyrotrop® (celle utilisant la séquence consensus ou celle utilisant l'ensemble des données générées par le pyroséquençage) ont été plus sensibles que les méthodes fondées sur l'algorithme geno2pheno, tandis que la spécificité était toujours meilleure lorsque l'ensemble des données de pyroséquençage étaient utilisées. Les valeurs prédictives positives sur la réponse au traitement étaient du même ordre pour les 4 méthodes, mais les valeurs prédictives négatives étaient meilleures pour les techniques analysées par Pyrotrop®. Finalement, le score Q de Yule, qui mesure l'intensité de la liaison entre prédiction et réponse antivirale au maraviroc, a été calculé pour l'ensemble des 4 méthodes

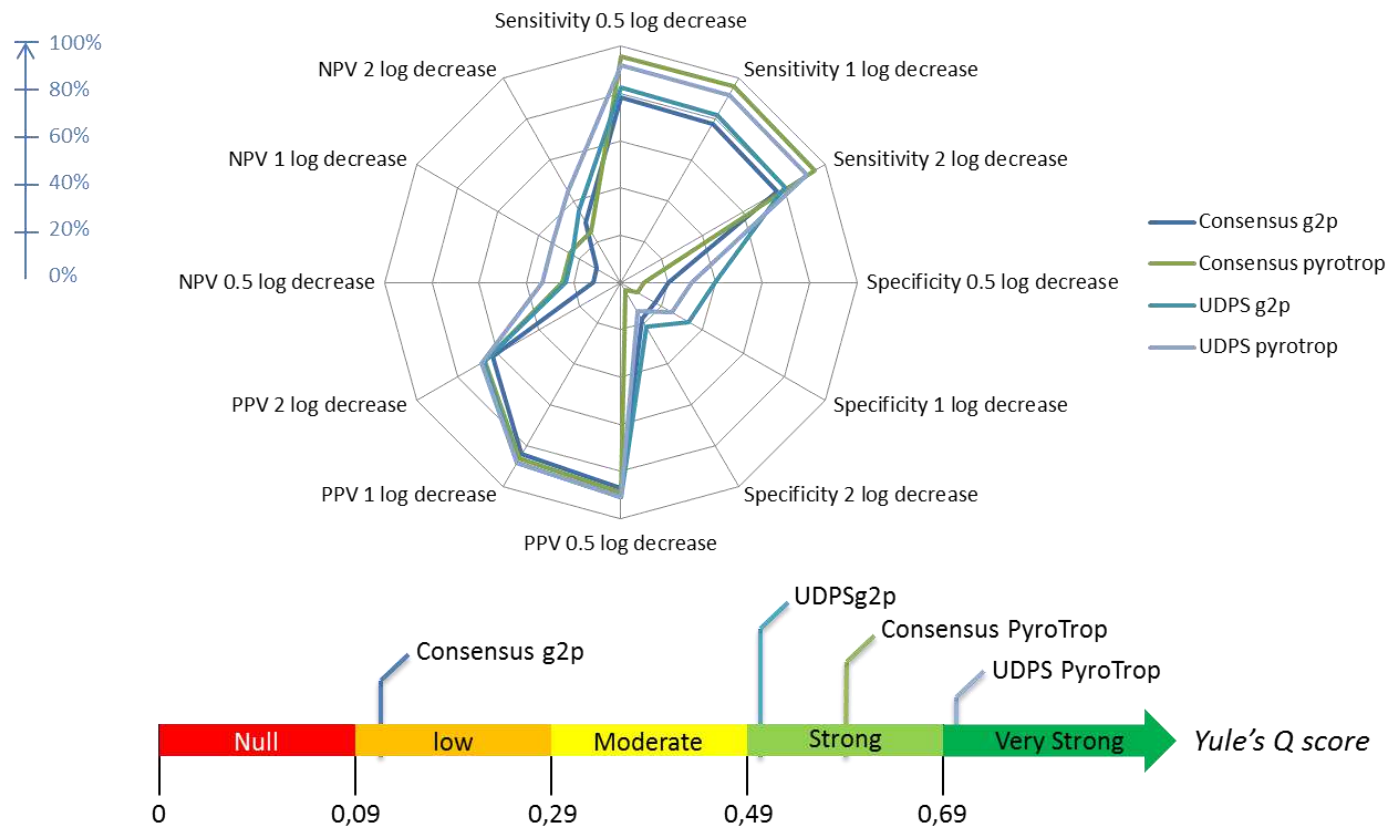


Figure 45 : Sensibilité, spécificité, valeur prédictive positive (VPP) et valeur prédictive négative (VPN) des 4 approches utilisées pour prédire la réponse au maraviroc avec 3 degrés différents de magnitude de réponse virologique à M1 (0,5, 1, et 2 log de décroissance de la charge virale) à partir des données de variabilité des 84 patients de l'étude obtenues à J0. La partie basse de la Figure indique la liaison entre prédiction et réponse antivirale selon le score Q de Yule.

(Figure 45). La liaison était faible pour la méthode utilisant consensus/g2p ($Q = 0,11$), forte pour les techniques utilisant pyroséquençage/g2p et consensus/PyroTrop® (respectivement $Q=0,51$ et $Q=0,59$) et très forte pour la technique pyroséquençage/Pyrotrop® ($Q = 0.71$).

Analyse de la variabilité génétique de la boucle V3

La prédictivité testée dans la partie précédente était fondée sur l'utilisation de la règle 11/25 + charge. Afin de vérifier le potentiel prédictif sur la réponse au maraviroc d'autres mutations, la prédictivité de chaque mutation de la boucle V3 a été testée à l'aide d'un module additionnel de PyroTrop® permettant le recueil des pourcentages de mutations à chaque position pour chaque patient. Les données recueillies et analysées par le package pROC en R ont permis d'établir la prédictivité de chacune des mutations à l'aide d'une courbe de ROC testée statistiquement (Figure 46). Il a ainsi été montré que la substitution H34Y/S possédait le potentiel prédictif le plus élevé ($p=0,011$) alors que les autres mutations en position 34 ou les mutations à d'autres positions étaient dénuées de potentiel prédictif significatif dans la cohorte étudiée. H34Y/S était présente chez 23% des patients répondeurs au maraviroc, à un niveau représentant 79,6% de la quasi espèce virale en moyenne. Cette mutation n'a en revanche jamais été trouvée chez les patients non répondeurs. A l'inverse, les mutations en position 11 et 25, utilisées dans la règle 11/25, sont apparues à la limite de la significativité d'un potentiel prédictif à titre individuel, avec $p=0,13$ pour les substitutions S11R/K et $p=0,08$ pour les substitutions D25R/K.

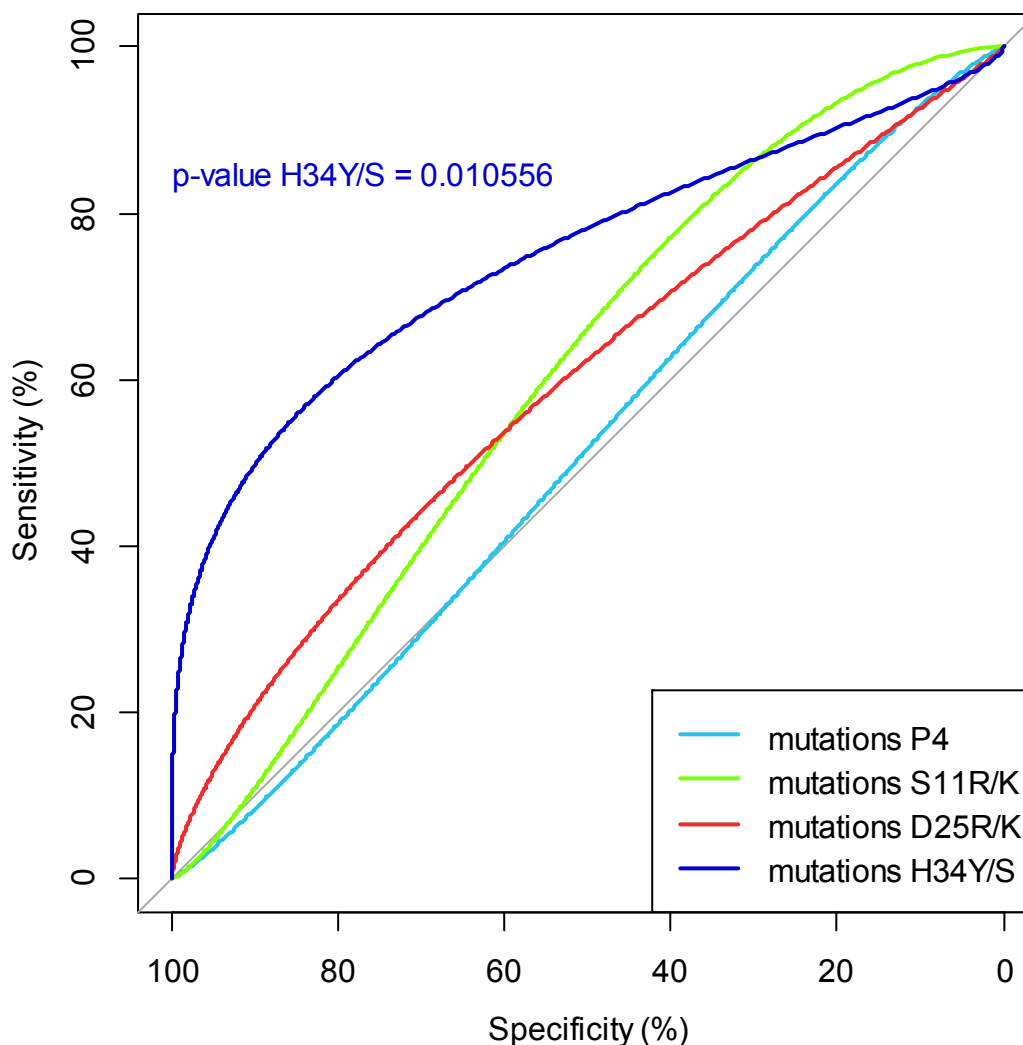


Figure 46 : Courbes ROC générées par le package R pROC évaluant la capacité à prédire la réponse au traitement par maraviroc des différentes mutations observées sur la boucle V3 du VIH chez les 84 patients de l'étude HIV génotropisme. Le test statistique utilisé a permis de mettre en évidence le potentiel prédictif significatif de la mutation H34Y/S, tandis que les mutations en position 11 et 25 avaient un potentiel prédictif à la limite de la significativité (respectivement $p = 0,11$ et $p = 0,08$). La courbe ROC des mutations en position 4, sans potentiel prédictif comme le reste des mutations testées, est incluse à titre indicatif.

DISCUSSION

Depuis quelques années, les cliniciens et virologues s'interrogent sur le bienfondé de débiter un traitement antiviral, parfois toxique et souvent coûteux, pour un bénéfice qui n'est pas toujours à la mesure des efforts entrepris par le patient, les soignants et le système de protection sociale. Dans certains cas, le choix d'une combinaison thérapeutique est difficile, car le rapport bénéfices/risques des différentes options semble équivalent si l'on se fonde sur les données des essais cliniques. Dans ces situations, délivrer au patient le traitement approprié, à la bonne dose et au bon moment, pour un bénéfice évaluable et un risque maîtrisé et accepté par tous les acteurs œuvrant pour la guérison apparaît comme une nécessité. Cette approche ne peut être seulement fondée sur des données statistiques issues des études de cohortes et des essais thérapeutiques. Elle doit être taillée au cas particulier du patient, avec sa maladie, son virus, ses capacités génétiques : on parle aujourd'hui de médecine personnalisée. Lorsque la décision de traiter a été prise, l'échec, toujours possible, ne doit pas être considéré comme une fatalité, car les mêmes approches peuvent être utilisées pour ajuster la probabilité de l'échec ou du succès et faire pencher la balance du côté de la guérison ou du contrôle de l'infection virale. Il y a peu, les techniques permettant d'étudier les paramètres nécessaires à ces ajustements n'étaient pas adaptées en terme de sensibilité, de débit ou de coût. L'avènement des techniques de séquençage à haut débit pourrait, dans ce contexte, représenter un tournant.

Dans ce travail, limité à l'étude du pathogène sans étude approfondie des facteurs de l'hôte, nous avons caractérisé les paramètres virologiques de la résistance et étudié leur capacité de prédire échec ou succès thérapeutique avant la mise en place et/ou au cours du

traitement, dans une perspective de transfert des connaissances acquises vers le diagnostic et la pratique clinique.

Prédiction de la résistance et de l'échec thérapeutique avant le début du traitement

Cet aspect a été abordé à deux niveaux. Nous avons d'abord caractérisé les marqueurs étudiés, puis évalué leur capacité de prédire l'émergence de la résistance, donc le succès ou l'échec thérapeutique.

La présence de variants viraux minoritaires résistants avant le début du traitement est en théorie la règle puisque, pour les 3 virus étudiés dans ce travail, chaque variant mono-muté est prédit comme étant présent à chaque instant chez tout patient infecté non traité. Dans le cas du VIH, la préexistence des variants viraux résistants a été bien décrite dans la littérature, avec une prévalence de variants résistants dépassant 1% de la quasi-espèce chez environ 30% des patients nouvellement pris en charge (Simen *et al.* 2009). Dans le cas du VHB, même si la préexistence des variants résistants est considérée comme acquise, elle n'a pas été documentée. Notre étude HBV InVS, qui porte sur un grand nombre d'échantillons et dont les données préliminaires sont présentées dans ce travail, répond à cette question. Nous avons en effet observé que 56% des patients étudiés abritaient des variants viraux portant des mutations connues pour conférer la résistance aux analogues nucléos(t)idiques anti-VHB qui représentaient plus de 0,2% de la quasi-espèce virale ; 8% d'entre eux abritaient des variants viraux résistants représentant plus de 5% de la quasi espèce (4,5% des patients de la cohorte totale). Les patients concernés par ce profil, supposés à risque de développer rapidement une résistance aux antiviraux, étaient 27% à posséder des mutations de résistance à l'entecavir, un des médicaments les plus utilisés en première intention à ce jour. Cependant, les profils de résistance à l'entecavir complets n'ont pas été trouvés avant

traitement. Plus de la moitié des patients présentaient des mutations minoritaires de résistance à l'adefovir qui peuvent diminuer la sensibilité du VHB au tenofovir, deuxième molécule la plus utilisée en première intention (Qi *et al.* 2007), ou retarder la négativation de la charge virale lors du traitement par ce médicament. Cependant, en pratique clinique, la résistance à ces deux molécules est un événement exceptionnel, rapporté chez 1,2% des patients seulement après 5 ans de traitement par entecavir (Tenney *et al.* 2009) et jamais rapporté après plus de 5 ans de traitement par le tenofovir. Le décalage apparent entre les observations cliniques et la présence de variants minoritaires potentiellement résistants en l'absence de tout traitement suggère que d'autres paramètres du traitement sont importants dans la sélection de variants viraux résistants, comme la sensibilité intrinsèque des souches aux médicaments et l'exposition, probablement assez élevée pour contrôler la réplication de variants viraux à susceptibilité partiellement réduite. Nos résultats suggèrent que la recherche de variants viraux résistants avant le début d'un traitement antiviral ne pourra pas être utilisée pour guider le choix thérapeutique. La question reste cependant ouverte pour le futur sur plusieurs aspects :

- la possibilité que de plus en plus de variants résistants soient transmis par des patients traités par entecavir en échec thérapeutique (résistance primaire) pourrait rendre nécessaire de reconsidérer le problème dans le cas du traitement par entecavir,
- l'existence de mutations de résistance au tenofovir, identifiées *in vitro* (Delaney *et al.* 2006; Borroto-Esoda *et al.* 2007), soulève la question de la survenue de la résistance après de nombreuses années de traitement, qui pourrait rendre nécessaire de dépister d'éventuels variants résistants minoritaires transmis,

- l'impact de la préexistence de mutations de résistance sur le retard à la négativation de la charge virale et la progression de la maladie hépatique doit être évalué,
- les rechutes fréquentes après l'arrêt du traitement pourraient être liées, après de nombreuses années de traitement et de contrôle efficace de la réplication, à la réplication à bas bruit de variants viraux résistants, source par mutation de la réémergence d'un mutant sauvage.

Dans le cas des mutations PreC/C du VHB, des études préliminaires, réalisées par la société Gilead, ont suggéré que leur présence au début du traitement était corrélée au succès du traitement par adefovir. Nous avons étudié ce marqueur dans une population de 157 patients pédiatriques Ag HBe-positifs traités par adefovir et suivis plusieurs années. La séroconversion HBe est associée à un ralentissement de la progression de la maladie hépatique (Chu *et al.* 2004) ; c'est donc un objectif thérapeutique important, même si la négativation de la charge virale reste l'objectif prioritaire. La séroconversion survient dans un contexte où les anticorps anti-HBe sont produits en excès par rapport à l'Ag HBe. Si l'on pose l'hypothèse que la synthèse d'anticorps anti-HBe est à peu près constante chez les patients, la séroconversion peut donc intervenir dans deux circonstances : soit après arrêt de la production de virions grâce à l'efficacité du traitement, soit par modification génétique de la quasi espèce virale conduisant à un arrêt ou une baisse de production de la protéine HBe (par exemple liée à la présence de mutations dans la région PreC/C). Dans ce dernier cas, la pression qui guide la sélection favorable de ces variants n'est pas connue, même si une origine immunitaire est probable (Hunt *et al.* 2000). Une forme d'immunotolérance permettrait au virus muté PreC de continuer à se répliquer à une phase tardive de sa coexistence avec son hôte. Elle pourrait expliquer la plus grande difficulté à contrôler l'infection chez les patients ayant une hépatite chronique Ag HBe-négative et la rechute plus

fréquente lorsque ces patients arrêtent leur traitement antiviral. La présence de mutations PreC/C avant le début du traitement pourrait donc être le reflet d'un début d'adaptation du virus à la réponse immune et serait par conséquent un prédicteur d'échec thérapeutique. C'est dans ce contexte que l'étude a été menée en parallèle avec 2 techniques capables de détecter la présence des mutations PreC/C. La première partie de l'étude a permis de montrer que la sensibilité de la technique de pyroséquençage était excellente, avec une capacité d'analyse plus étendue que celle de l'INNO-LiPA puisque sans a priori sur la séquence recherchée et couvrant l'intégralité du gène amplifié et séquencé. Cependant, dans notre étude, la présence de mutations PreC/C avant la mise en place du traitement ne présentait aucune relation significative avec la séroconversion ultérieure, ni avec la séroconversion associée à une négativation de la charge virale, qu'elle soit étudiée par la technique INNO-LiPA ou en pyroséquençage à haut débit, même si cette approche améliorerait sensiblement le score Q de Yules. Nos résultats indiquent donc que la présence de mutations PreC/C détectables au début du traitement ne permet pas de prédire son issue, infirmant l'hypothèse de départ.

Dans le cas du VHC, aucune étude fondée sur des techniques très sensibles évaluant le risque d'échec thérapeutique associé à la détection de mutations de résistance aux antiviraux n'a été publiée. Notre étude a montré que la quasi-totalité des patients infectés par le VHC abritent naturellement des variants viraux résistants en proportions non négligeables, et ce en l'absence de toute exposition antérieure à un traitement antiviral. Dans certains cas, ces variants sont majoritaires à la base. Cependant, notre étude montre clairement que la présence de variants résistants au départ et leur type n'influencent pas l'issue du traitement lorsque le telaprevir est associé à l'interféron alpha pégylé avec ou sans ribavirine. Ce résultat n'est pas surprenant dans la mesure où l'échec thérapeutique est

essentiellement lié à une réponse insuffisante à l'action antivirale de l'interféron, qui ne peut contrôler la croissance des variants viraux résistants sélectionnés par l'inhibiteur de protéase. Chez les bons répondeurs à l'interféron, la préexistence de variants résistants n'a pas été délétère en raison de leur sensibilité à l'action antivirale de ce dernier. Dans l'hypothèse probable d'approches thérapeutiques sans interféron, l'impact de la présence au début du traitement de variants viraux minoritaires portant des mutations de résistance aux différentes classes de médicaments utilisées pourrait en revanche jouer un rôle important sur l'issue du traitement. Ce rôle devra être évalué et notre étude montre que la meilleure approche reposera sur l'utilisation du pyroséquençage.

Dans le cas du VIH, les risques associés à la présence de variants viraux minoritaires résistants au début du traitement sont bien documentés dans une littérature abondante. L'utilisation de cette information pour la décision stratégique en pratique clinique reste cependant confuse. Dans notre étude consacrée au génotypisme du VIH, nous avons montré que la prédiction de la réponse au traitement par l'étude du génotypisme avant sa mise en place pouvait être améliorée d'une part par l'utilisation de techniques très sensibles fondées sur le pyroséquençage, d'autre part par l'analyse des données de pyroséquençage à l'aide du logiciel PyroTrop[®]. Nous avons également souligné la difficulté d'identifier des séquences signatures qui différencient de façon définitive les souches CCR5 des souches CXCR4. Une évaluation plus poussée des mutations de la boucle V3 et son extension à l'ensemble de la gp120 seront sans doute nécessaire pour identifier la séquence contiguë ou les motifs de séquence permettant une meilleure caractérisation. Notre étude montre que l'approche fondée sur le pyroséquençage et l'analyse des données par le logiciel PyroTrop[®] développé dans le cadre de cette Thèse est fiable et pourrait trouver son utilité en pratique quotidienne de laboratoire avant la mise en place de traitements par le maraviroc.

Etude dynamique de la résistance et de l'échec thérapeutique au cours du traitement

La prédiction de l'échec thérapeutique et de la résistance à partir de l'analyse réalisée avant traitement présuppose que le patient se comporte selon le modèle prédictif déduit de l'observation d'un grand nombre de sujets. En réalité, des événements imprévus, spécifiques de l'histoire du patient, sont à tout moment susceptibles de modifier l'évolution du pathogène dans le contexte thérapeutique. La prédiction peut alors être ajustée par l'étude des changements dynamiques des populations virales en cours de traitement, qui apportent une information précise sur l'effet réel du traitement sur le virus.

La dynamique d'émergence de variants viraux résistants en cours de traitement antiviral a fait l'objet d'un traitement analytique spécifique, qui s'est traduit par la création du logiciel PyroDyn®. Ce logiciel a pour fonction de détecter la croissance ou la décroissance exponentielle de mutations (ou de populations virales lorsque l'analyse est complétée par PyroLink®) dont l'interprétation permet de caractériser, comprendre et éventuellement prédire le succès ou l'échec thérapeutique. Le principe est que l'ensemble des variants viraux présents chez les patients infectés suivent des cinétiques temporelles fluctuantes, fruit du hasard, à l'origine de courbes de croissance non significatives. Ces variants apparaissent au gré d'une mutation au hasard. Ils n'ont généralement pas la capacité de se multiplier rapidement dans un environnement favorable qui les amène à devenir dominants et disparaissent le plus souvent rapidement. Cependant, certains variants apparus au hasard ont une capacité répliquative suffisante pour survivre. Dans ce cas, les courbes temporelles suivent un modèle particulier (logarithmique, exponentiel, binomial) en fonction du contexte répliquatif, qui peut être modélisé afin de permettre la détection de l'acquisition de fitness

d'une population virale dont la séquence est inconnue a priori. C'est le principe du logiciel PyroDyn® développé dans ce travail. Parmi les multiples représentations possibles des données dynamiques, nous avons choisi de représenter les quantités relatives des différentes populations virales en valeur absolue, par rapport à la charge virale totale. Cette approche est la plus fidèle pour montrer les changements réels qui surviennent au sein d'une quasi espèce virale sous traitement antiviral. Elle permet d'identifier des variants sensibles dont les populations décroissent plus lentement que les variants sauvages majoritaires et de les différencier des variants viraux résistants dont la croissance est rapide au cours du traitement. La croissance exponentielle de ces derniers, lorsqu'elle est avérée, indique clairement l'état de résistance, et notre approche fondée sur PyroDyn® et PyroLink® permet de détecter la résistance sans connaissance a priori de la séquence virale responsable.

La première étude ayant bénéficié de ces outils a été l'étude HBV POL. Dans ce travail, nous avons comparé les résultats du pyroséquençage à haut débit à des résultats obtenus par clonage et séquençage sur les mêmes échantillons. D'un point de vue purement ergonomique, la technique de clonage-séquençage a nécessité plus de 2 ans de travail expérimental et analytique pour environ 4000 clones analysés, alors que la génération des séquences et l'analyse ont nécessité quelques semaines en pyroséquençage pour environ 500 000 séquences générées. C'est donc 100 fois plus de données brutes qui ont été acquises à partir des mêmes échantillons lorsque le pyroséquençage a été utilisé. L'inconvénient majeur lié à la masse de données générées est la difficulté d'analyse. La correction manuelle de 500 000 séquences est simplement impossible. C'est la raison pour laquelle nous avons développé un outil de gestion automatisée des données de pyroséquençage. Une fois corrigée et alignée, la masse d'information de séquence a permis

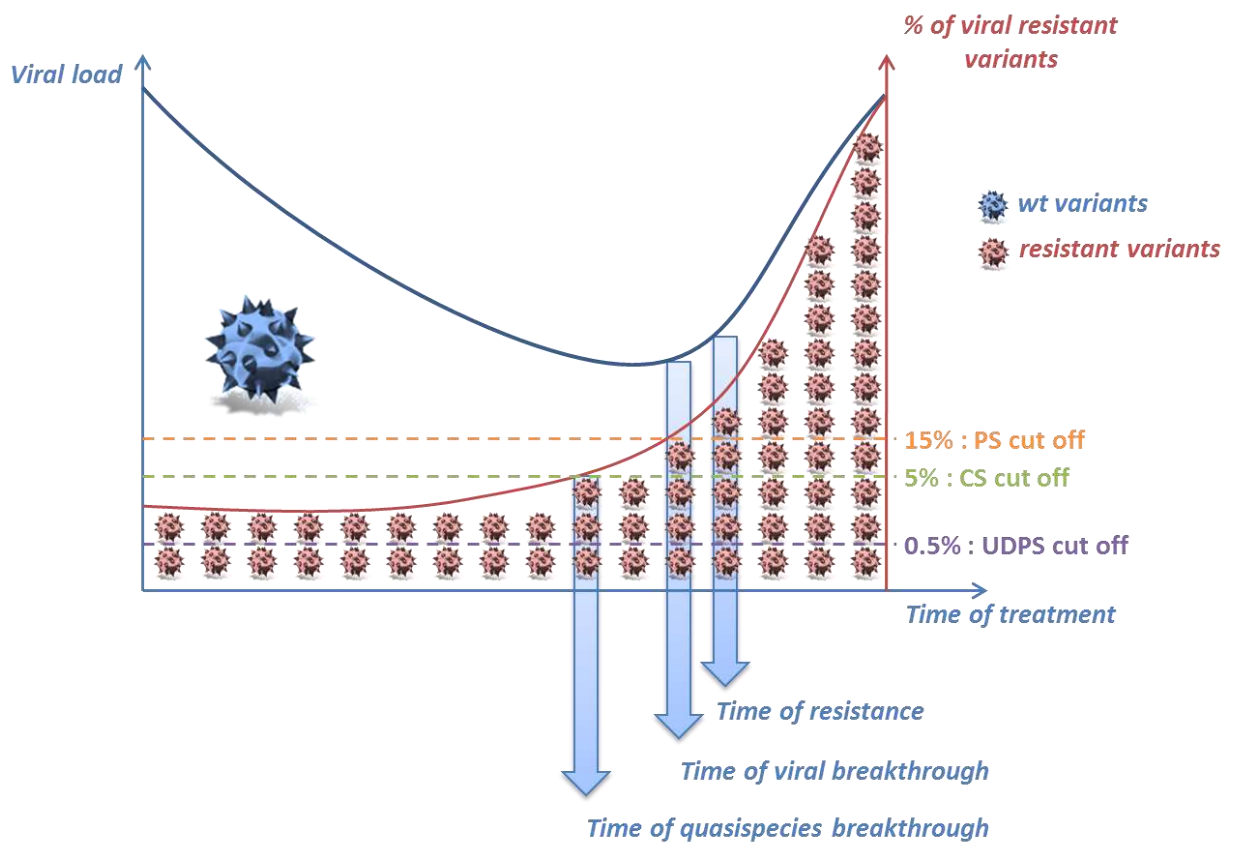


Figure 47 : Représentation schématiques des évènements virologiques lors de l'émergence de la résistance à un traitement antiviral. Les techniques actuelles, comme le séquençage direct (PS), détectent la résistance lorsqu'elle est installée, tandis que les techniques comme le clonage séquençage (CS) ou le pyroséquençage (UDPS) la détectent beaucoup plus tôt. L'idéal est d'anticiper la résistance et d'éviter la sélection de variants très adaptés au traitement.

d'atteindre une sensibilité considérable. Le supplément d'information acquis permet la détection de variants minoritaires non détectés par le clonage-séquençage, car présents en proportions trop faibles. En outre, la description des cinétiques des populations virales en cours de traitement s'est avérée beaucoup plus précise et correspondait mieux aux modèles mathématiques de croissance, permettant la caractérisation de nombreuses populations virales possédant des capacités de résistance au médicament variables. Certaines des mutations nouvellement détectées par la technique de pyroséquençage, comme la mutation rtN238T de l'étude HBV POL ou la substitution S67A du domaine NS3 du VHC de l'étude HCV Prove 2 ont déjà été décrites dans la littérature (Pastor *et al.* 2009; Bae *et al.* 2010; Zhong *et al.* 2012). Les mutations observées dans le cadre de lecture de l'Ag HBs dans l'étude HBV POL sont situées au niveau d'épitopes d'immunodominance, suggérant un mécanisme d'échappement au contrôle immunitaire. Le couplage du pyroséquençage et de l'analyse par PyroPack[®] permet donc une caractérisation sensible, spécifique et sans a priori, de nature à permettre la détection précoce de phénomène débutants à l'origine de mouvements de plus grande ampleur finalement responsables de l'échec virologique.

Le rebond virologique survient au moment où la résistance est durablement installée, le variant majoritaire étant déjà bien adapté à son environnement. L'anticipation de ce phénomène consiste à détecter l'événement en amont de sa survenue (Figure 47). Il a ainsi été possible de détecter l'émergence de variants résistants plusieurs mois avant les techniques conventionnelles dans l'étude HBV POL. Une limite de cette approche est toutefois la nécessité que la charge virale soit détectable pour permettre le diagnostic précoce de la résistance. C'était le cas des patients de notre étude, traités par adefovir, molécule peu puissante associée à des virémies détectables dans un grand nombre de cas au cours du traitement. Il est évident qu'avec des molécules plus puissantes, l'échec

thérapeutique, qui se traduit par un rebond virologique chez des patients à charge virale négative, survient dans un contexte de résistance avérée, la population virale résistante étant majoritaire et parfaitement adaptée à son environnement.

Des limitations pratiques doivent également être soulignées. Dans l'étude HCV Prove 2, l'émergence de variants viraux résistants est souvent survenue quelques jours après le début du traitement, un délai incompatible avec la pratique de la biologie médicale. Si l'approche dynamique n'est pas utilisée en pratique quotidienne, elle reste indispensable pour la génération de données permettant de mieux comprendre les mécanismes de la résistance et son rôle dans l'échec thérapeutique. L'étude de la dynamique des populations virales dans l'étude Prove 2 a permis de mettre en évidence une complexification par étapes des populations majoritaires, avec gain de fitness graduel de populations portant de plus en plus de mutations de résistance et de fitness. L'étude a également montré un retour à un état proche de l'état pré-thérapeutique en plusieurs mois après l'arrêt du traitement, avec un virus dominant sauvage et des variants résistants simples minoritaires.

Quasi espèce de départ ou étude dynamique au cours du traitement ?

Le développement de technologies ultra-sensibles, comme le pyroséquençage, soulève la question de l'intérêt d'une détermination statique de la composition de la quasi espèce virale avant le traitement par rapport au suivi des dynamiques virales au cours du traitement, dans le contexte de la prédiction de l'issue du traitement antiviral. Dans ce cas, un seuil de présence de variants résistants au-delà duquel le patient pourrait être considéré comme à risque d'échec thérapeutique peut-il être défini ? Des travaux antérieurs réalisés dans le domaine du VIH et les travaux présentés ici montrent que la quantité de variants

viraux préexistants, tout au moins sa partie visible, n'influence pas la réponse virologique au traitement. Une des raisons est que même lorsqu'on utilise les techniques de pyroséquençage, les populations détectées sont relativement abondantes et déjà très adaptées à leur environnement réplcatif. Il faudrait une sensibilité beaucoup plus importante, techniquement irréalisable aujourd'hui, pour cartographier avec exactitude la composition de base des quasi espèces virales et étudier son influence sur l'évolution thérapeutique. L'étude de la dynamique des populations virales sous traitement, qui capture à la fois l'information qualitative de la présence de variants résistants et l'information quantitative d'acquisition exponentielle de fitness en cours de traitement, apporte donc beaucoup plus d'information. Elle est cependant limitée, comme discuté plus haut, par le fait que la charge virale doit être détectable et que celle-ci le redevient bien souvent à un stade avancé du processus d'échec thérapeutique lié à la résistance.

Optimisation des outils d'étude et de lutte contre la résistance

La prévention de l'émergence de la résistance aux antiviraux est fondée sur la combinaison de médicaments appartenant à des classes différentes, sans résistance croisée. C'est dans ce contexte que le projet HBV RNase H a été initié, avec pour objectif de caractériser la variabilité naturelle de cette enzyme et d'identifier des cibles potentielles pour le développement d'inhibiteurs spécifiques. En effet, les 2 enzymes du VHB (transcriptase inverse et RNase H) sont portées par l'ADN polymérase virale. Tous les médicaments anti-VHB sont aujourd'hui dirigés contre l'activité transcriptase inverse, qui a bénéficié du transfert d'expertise du monde des antirétroviraux.

La RNase H représente une cible thérapeutique intéressante, car des médicaments dirigés contre son activité enzymatique pourraient trouver une indication en combinaison avec les

inhibiteurs de la transcriptase inverse du VHB, éventuellement en seconde intention en cas d'échec thérapeutique ou de réponse suboptimale. Les tentatives de développement d'inhibiteurs de la RNase H du VIH sont malheureusement restées infructueuses, en raison de la proximité structurale avec les RNase H bactériennes et humaines, faisant courir le risque d'un manque de spécificité associé à une toxicité importante. Notre objectif était d'identifier des domaines spécifiques de la RNase H du VHB pouvant être ciblés par une approche de « drug discovery ». La superposition des structures de RNase H a permis d'identifier un domaine de protrusion C spécifique du VHB (Figure 39), structuré en hélice alpha. Cette structure a été prédite par modélisation mathématique, mais les logiciels de modélisation structurale n'ont pas visualisé ce domaine. D'un point de vue fonctionnel, l'identification de la fonction de ce domaine fait l'objet d'un travail collaboratif avec l'équipe du Pr Mikael Kann (UMR 5234, Bordeaux). Par homologie avec les autres RNase H, ce domaine pourrait être lié à la capacité de l'enzyme de fixer préférentiellement un hétéroduplex ADN/ARN, fonction identifiée comme essentielle pour l'activité enzymatique. Ce domaine pourrait représenter une piste intéressante de développement de nouveaux inhibiteurs ciblant spécifiquement le VHB. De façon intéressante, ce domaine est l'un des moins variables de l'enzyme, ce qui indique qu'il est structuralement et fonctionnellement indispensable et assurerait à des inhibiteurs une barrière génétique élevée. De façon importante, le domaine d'intérêt n'est pas codé par une portion de génome partagée avec le gène X, donc la contrainte conservatoire est exercée sur la RNase H elle-même. Le domaine de protrusion représente donc une cible à fort potentiel thérapeutique qui fera l'objet d'études plus approfondies et d'un screening large pour la découverte d'inhibiteurs de l'activité RNase H.

Les techniques de pyroséquençage présentées dans ce travail peuvent également être utilisées au cours du traitement pour mieux comprendre les mécanismes de la résistance au cours des traitements antiviraux. Il est possible, sans connaissance préalable de type docking, de repérer des domaines de liaison du médicament à la protéine virale cible. C'est ce qui a été rendu possible par l'analyse des données de séquences et leur évolution au cours du temps au niveau du domaine NS3 du VHC sous traitement par inhibiteurs de protéase par le logiciel Pyrodyn®. Il a ainsi été montré que l'échappement était lié à des changements d'acides aminés survenant essentiellement au niveau de 2 sous-domaines de la protéine NS3, confirmant les données de docking rapportées dans la littérature. L'utilisation du pyroséquençage peut donc être utile pour améliorer les connaissances sur les mécanismes de résistance sans passer par les étapes de modélisation ou en complément de celles-ci.

CONCLUSION-PERSPECTIVES

Nous avons créé des outils informatiques d'analyse des données de séquençage de nouvelle génération adaptés à l'étude de la dynamique des populations virales dont nous avons démontré le potentiel et l'intérêt dans le dépistage et la caractérisation de la résistance aux antiviraux. La poursuite de ce travail se fera selon 2 axes, qui ont guidé les travaux de ce mémoire. Nous envisageons d'abord d'adapter ces outils à une pratique courante de laboratoire. Nous confirmerons les résultats obtenus dans plusieurs études de prédiction de la résistance. La mise en place d'une plateforme technologique de séquençage de nouvelle génération sur le site du Groupe Hospitalier Henri Mondor nous permettra d'améliorer les coûts et l'organisation, afin de permettre le transfert de la technologie et des méthodes d'analyse en pratique clinique.

Sur le plan de la recherche proprement dite, nous continuerons d'utiliser l'outil pour mieux comprendre les mécanismes de résistance du VHB, du VHC et du VIH aux nombreux antiviraux utilisés en clinique ou en cours de développement. Dans ce cadre, la plateforme d'étude phénotypique développée au laboratoire par le Dr Abdelhakim Ahmed-Belkacem, qui dispose de tests adaptés à des débits importants pour la caractérisation des mutations de résistance et de fitness, sera mise à disposition pour définir le rôle exact des mutations identifiées comme jouant un rôle important dans les phénomènes observés. Enfin, des transferts sont possibles vers d'autres domaines, comme la découverte de nouveaux médicaments, qui fait déjà l'objet d'études collaboratives, ou l'étude de l'échappement viral dans le contexte de la vaccination contre le VIH ou le VHC (projet intégré à celui du Labex

Vaccine Research Institute pour lequel plusieurs études ont débuté en partenariat avec l'équipe du Pr Yves Levy à Créteil).

BIBLIOGRAPHIE

- Ahmadian, A., B. Gharizadeh, et al. (2000). "Single-nucleotide polymorphism analysis by pyrosequencing." *Analytical biochemistry* **280**(1): 103-110.
- Allen, M. I., M. Deslauriers, et al. (1998). "Identification and characterization of mutations in hepatitis B virus resistant to lamivudine. Lamivudine Clinical Investigation Group." *Hepatology* **27**(6): 1670-1677.
- Altmann, A., M. Daumer, et al. (2009). "Predicting the response to combination antiretroviral therapy: retrospective validation of geno2pheno-THEO on a large clinical database." *The Journal of infectious diseases* **199**(7): 999-1006.
- Altshuler, D., V. J. Pollara, et al. (2000). "An SNP map of the human genome generated by reduced representation shotgun sequencing." *Nature* **407**(6803): 513-516.
- Angus, P., R. Vaughan, et al. (2003). "Resistance to adefovir dipivoxil therapy associated with the selection of a novel mutation in the HBV polymerase." *Gastroenterology* **125**(2): 292-297.
- Archer, J., A. Rambaut, et al. (2010). "The evolutionary analysis of emerging low frequency HIV-1 CXCR4 using variants through time--an ultra-deep approach." *PLoS computational biology* **6**(12): e1001022.
- Avery, O. T., C. M. Macleod, et al. (1944). "Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii." *The Journal of experimental medicine* **79**(2): 137-158.
- Bacon, B. R., S. C. Gordon, et al. (2011). "Boceprevir for previously treated chronic HCV genotype 1 infection." *The New England journal of medicine* **364**(13): 1207-1217.
- Bae, A., S. C. Sun, et al. (2010). "Susceptibility of treatment-naive hepatitis C virus (HCV) clinical isolates to HCV protease inhibitors." *Antimicrobial agents and chemotherapy* **54**(12): 5288-5297.
- Bakeman, R. and J. M. Gottman (1989). *Observing interaction an introduction to sequential analysis*. Cambridge New York Port Chester, Cambridge university press.
- Bandos, A. I., H. E. Rockette, et al. (2005). "A permutation test sensitive to differences in areas for comparing ROC curves from a paired design." *Statistics in medicine* **24**(18): 2873-2893.
- Barbotte, L., A. Ahmed-Belkacem, et al. (2010). "Characterization of V36C, a novel amino acid substitution conferring hepatitis C virus (HCV) resistance to telaprevir, a potent peptidomimetic inhibitor of HCV protease." *Antimicrobial agents and chemotherapy* **54**(6): 2681-2683.
- Bartholomeusz, A., B. G. Tehan, et al. (2004). "Comparisons of the HBV and HIV polymerase, and antiviral resistance mutations." *Antiviral therapy* **9**(2): 149-160.
- Beerenwinkel, N., M. Daumer, et al. (2003). "Geno2pheno: Estimating phenotypic drug resistance from HIV-1 genotypes." *Nucleic acids research* **31**(13): 3850-3855.
- Bennett, S. (2004). "Solexa Ltd." *Pharmacogenomics* **5**(4): 433-438.
- Blight, K. J. and E. A. Norgard (2006). HCV Replicon Systems. *Hepatitis C Viruses: Genomes and Molecular Biology*. S. L. Tan. Norfolk (UK).
- Block, T. M., X. Lu, et al. (1998). "Treatment of chronic hepadnavirus infection in a woodchuck animal model with an inhibitor of protein folding and trafficking." *Nature medicine* **4**(5): 610-614.
- Bock, C. T., S. Schwinn, et al. (2001). "Structural organization of the hepatitis B virus minichromosome." *Journal of molecular biology* **307**(1): 183-196.
- Borroto-Esoda, K., M. D. Miller, et al. (2007). "Pooled analysis of amino acid changes in the HBV polymerase in patients from four major adefovir dipivoxil clinical trials." *Journal of hepatology* **47**(4): 492-498.
- Boyer, P. L., S. G. Sarafianos, et al. (2001). "Selective excision of AZTMP by drug-resistant human immunodeficiency virus reverse transcriptase." *Journal of virology* **75**(10): 4832-4842.
- Cerritelli, S. M. and R. J. Crouch (2009). "Ribonuclease H: the enzymes in eukaryotes." *The FEBS journal* **276**(6): 1494-1505.

- Chatterji, U., M. D. Bobardt, et al. (2005). "Naturally occurring capsid substitutions render HIV-1 cyclophilin A independent in human cells and TRIM-cyclophilin-resistant in Owl monkey cells." *The Journal of biological chemistry* **280**(48): 40293-40300.
- Chen, X., M. Tsiang, et al. (2008). "Modeling, analysis, and validation of a novel HIV integrase structure provide insights into the binding modes of potent integrase inhibitors." *Journal of molecular biology* **380**(3): 504-519.
- Chen, Y. and P. L. Marion (1996). "Amino acids essential for RNase H activity of hepadnaviruses are also required for efficient elongation of minus-strand viral DNA." *Journal of virology* **70**(9): 6151-6156.
- Chen, Y., W. S. Robinson, et al. (1994). "Selected mutations of the duck hepatitis B virus P gene RNase H domain affect both RNA packaging and priming of minus-strand DNA synthesis." *Journal of virology* **68**(8): 5232-5238.
- Chenna, R., H. Sugawara, et al. (2003). "Multiple sequence alignment with the Clustal series of programs." *Nucleic acids research* **31**(13): 3497-3500.
- Chevaliez, S. and P. JM (2008). *Interferons and Their Use in Persistent Viral Infections*, Springer Berlin Heidelberg.
- Chevaliez, S., C. Rodriguez, et al. (2012). "New virologic tools for management of chronic hepatitis B and C." *Gastroenterology* **142**(6): 1303-1313 e1301.
- Chevaliez, S., C. Rodriguez, et al. (2011). "Molecular Characterization of HCV resistance to telaprevir by means of ultra-deep pyrosequencing: preexisting resistant variants and dynamics of resistant populations." *Journal of hepatology* **54**(1): S30.
- Choi, W. T., S. Duggineni, et al. (2012). "Drug discovery research targeting the CXC chemokine receptor 4 (CXCR4)." *Journal of medicinal chemistry* **55**(3): 977-994.
- Chokshi, S., S. Phillips, et al. (2012). "OC-025 Alisporivir inhibition of cellular cyclophilins disrupts hepatitis B virus (HBV) replication and this effect is further enhanced in combination with direct antiviral targeting HBV-DNA polymerase in vitro." *gut* **61:suppl 2**(A11).
- Chu, C. M., S. J. Hung, et al. (2004). "Natural history of hepatitis B e antigen to antibody seroconversion in patients with normal serum aminotransferase levels." *The American journal of medicine* **116**(12): 829-834.
- Clavel, F. and A. J. Hance (2004). "HIV drug resistance." *The New England journal of medicine* **350**(10): 1023-1035.
- Coakley, E., C. J. Petropoulos, et al. (2005). "Assessing chemokine co-receptor usage in HIV." *Current opinion in infectious diseases* **18**(1): 9-15.
- Coelmont, L., X. Hanouille, et al. (2010). "DEB025 (Alisporivir) inhibits hepatitis C virus replication by preventing a cyclophilin A induced cis-trans isomerisation in domain II of NS5A." *PloS one* **5**(10): e13687.
- Cuevas, J. M., F. Gonzalez-Candelas, et al. (2009). "Effect of ribavirin on the mutation rate and spectrum of hepatitis C virus in vivo." *Journal of virology* **83**(11): 5760-5764.
- Daelemans, D., J. M. Dumont, et al. (2010). "Debio-025 inhibits HIV-1 by interfering with an early event in the replication cycle." *Antiviral research* **85**(2): 418-421.
- Dahm, R. (2008). "Discovering DNA: Friedrich Miescher and the early years of nucleic acid research." *Human genetics* **122**(6): 565-581.
- Darwin, C. (1859). *On the origin of species by means of natural selection*. London,, J. Murray.
- Das, K., A. D. Clark, Jr., et al. (2004). "Roles of conformational and positional adaptability in structure-based design of TMC125-R165335 (etravirine) and related non-nucleoside reverse transcriptase inhibitors that are highly potent and effective against wild-type and drug-resistant HIV-1 variants." *Journal of medicinal chemistry* **47**(10): 2550-2560.
- Davidson, A., T. C. Leeper, et al. (2009). "Simultaneous recognition of HIV-1 TAR RNA bulge and loop sequences by cyclic peptide mimics of Tat protein." *Proceedings of the National Academy of Sciences of the United States of America* **106**(29): 11931-11936.
- De Clercq, E. (2007). "The design of drugs for HIV and HCV." *Nature reviews. Drug discovery* **6**(12): 1001-1018.

- De Jong, J. J., A. De Ronde, et al. (1992). "Minimal requirements for the human immunodeficiency virus type 1 V3 domain to support the syncytium-inducing phenotype: analysis by single amino acid substitution." Journal of virology **66**(11): 6777-6780.
- de Jong, J. J., J. Goudsmit, et al. (1992). "Human immunodeficiency virus type 1 clones chimeric for the envelope V3 domain differ in syncytium formation and replication capacity." Journal of virology **66**(2): 757-765.
- Delaney, W. E. t. and H. C. Isom (1998). "Hepatitis B virus replication in human HepG2 cells mediated by hepatitis B virus recombinant baculovirus." Hepatology **28**(4): 1134-1146.
- Delaney, W. E. t., A. S. Ray, et al. (2006). "Intracellular metabolism and in vitro activity of tenofovir against hepatitis B virus." Antimicrobial agents and chemotherapy **50**(7): 2471-2477.
- Delobel, P., M. T. Nugeyre, et al. (2007). "Population-based sequencing of the V3 region of env for predicting the coreceptor usage of human immunodeficiency virus type 1 quasispecies." Journal of clinical microbiology **45**(5): 1572-1580.
- Dohm, J. C., C. Lottaz, et al. (2008). "Substantial biases in ultra-short read data sets from high-throughput DNA sequencing." Nucleic Acids Res **36**(16): e105.
- Dorr, P., M. Westby, et al. (2005). "Maraviroc (UK-427,857), a potent, orally bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5 with broad-spectrum anti-human immunodeficiency virus type 1 activity." Antimicrobial agents and chemotherapy **49**(11): 4721-4732.
- Durantel, D., M. N. Brunelle, et al. (2005). "Resistance of human hepatitis B virus to reverse transcriptase inhibitors: from genotypic to phenotypic testing." Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology **34 Suppl 1**: S34-43.
- Durantel, D., S. Carrouee-Durantel, et al. (2004). "A new strategy for studying in vitro the drug susceptibility of clinical isolates of human hepatitis B virus." Hepatology **40**(4): 855-864.
- EASL (2009). "EASL Clinical Practice Guidelines: management of chronic hepatitis B." Journal of hepatology **50**(2): 227-242.
- EASL (2011). "EASL Clinical Practice Guidelines: management of hepatitis C virus infection." Journal of hepatology **55**(2): 245-264.
- Eid, J., A. Fehr, et al. (2009). "Real-time DNA sequencing from single polymerase molecules." Science **323**(5910): 133-138.
- Eigen, M. (1971). "Selforganization of matter and the evolution of biological macromolecules." Die Naturwissenschaften **58**(10): 465-523.
- Eigen, M. and P. Schuster (1979). The hypercycle, a principle of natural self-organization. Berlin ; New York, Springer-Verlag.
- Esnouf, R., J. Ren, et al. (1995). "Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors." Nature structural biology **2**(4): 303-308.
- Ewing, B. and P. Green (1998). "Base-calling of automated sequencer traces using phred. II. Error probabilities." Genome research **8**(3): 186-194.
- Ewing, B., L. Hillier, et al. (1998). "Base-calling of automated sequencer traces using phred. I. Accuracy assessment." Genome research **8**(3): 175-185.
- Fernandez-Montero, J. V., P. Barreiro, et al. (2009). "HIV protease inhibitors: recent clinical trials and recommendations on use." Expert opinion on pharmacotherapy **10**(10): 1615-1629.
- Fisher, R. A. (1930). The genetical theory of natural selection. Oxford,, The Clarendon press.
- Flexner, C. (2007). "HIV drug development: the next 25 years." Nature reviews. Drug discovery **6**(12): 959-966.
- Fouchier, R. A., M. Brouwer, et al. (1995). "Simple determination of human immunodeficiency virus type 1 syncytium-inducing V3 genotype by PCR." Journal of clinical microbiology **33**(4): 906-911.
- Fouchier, R. A., M. Groenink, et al. (1992). "Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule." Journal of virology **66**(5): 3183-3187.

- Fried, M. W., S. J. Hadziyannis, et al. (2011). "Rapid virological response is the most important predictor of sustained virological response across genotypes in patients with chronic hepatitis C virus infection." Journal of hepatology **55**(1): 69-75.
- Gallay, P. A. (2012). "Cyclophilin inhibitors: a novel class of promising host-targeting anti-HCV agents." Immunologic research **52**(3): 200-210.
- Gane, E. J., R. Rouzier, et al. (2011). "Antiviral activity, safety, and pharmacokinetics of danoprevir/ritonavir plus PEG-IFN alpha-2a/RBV in hepatitis C patients." Journal of hepatology **55**(5): 972-979.
- Gao, M., R. E. Nettles, et al. (2010). "Chemical genetics strategy identifies an HCV NS5A inhibitor with a potent clinical effect." Nature **465**(7294): 96-100.
- Ghany, M. and T. J. Liang (2007). "Drug targets and molecular mechanisms of drug resistance in chronic hepatitis B." Gastroenterology **132**(4): 1574-1585.
- Gianella, S., W. Delpont, et al. (2011). "Detection of Minority Resistance during Early HIV-1 Infection: Natural Variation and Spurious Detection rather than Transmission and Evolution of Multiple Viral Variants." Journal of virology **85**(16): 8359-8367.
- Gilles, A., E. Meglec, et al. (2011). "Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing." BMC genomics **12**: 245.
- Gish, R. G., H. Trinh, et al. (2005). "Safety and antiviral activity of emtricitabine (FTC) for the treatment of chronic hepatitis B infection: a two-year study." Journal of hepatology **43**(1): 60-66.
- Goh, W. C., J. Markee, et al. (1999). "Protection against human immunodeficiency virus type 1 infection in persons with repeated exposure: evidence for T cell immunity in the absence of inherited CCR5 coreceptor defects." The Journal of infectious diseases **179**(3): 548-557.
- Grant, R. M., F. M. Hecht, et al. (2002). "Time trends in primary HIV-1 drug resistance among recently infected persons." JAMA : the journal of the American Medical Association **288**(2): 181-188.
- Greenberg, M. L. and N. Cammack (2004). "Resistance to enfuvirtide, the first HIV fusion inhibitor." The Journal of antimicrobial chemotherapy **54**(2): 333-340.
- Gu, B., A. T. Gates, et al. (2003). "Replication studies using genotype 1a subgenomic hepatitis C virus replicons." Journal of virology **77**(9): 5352-5359.
- Gunther, S., B. C. Li, et al. (1995). "A novel method for efficient amplification of whole hepatitis B virus genomes permits rapid functional analysis and reveals deletion mutants in immunosuppressed patients." Journal of virology **69**(9): 5437-5444.
- Guo, J. T., V. V. Bichko, et al. (2001). "Effect of alpha interferon on the hepatitis C virus replicon." Journal of virology **75**(18): 8516-8523.
- Harismendy, O., P. C. Ng, et al. (2009). "Evaluation of next generation sequencing platforms for population targeted sequencing studies." Genome biology **10**(3): R32.
- Hatzioannou, T., D. Perez-Caballero, et al. (2005). "Cyclophilin interactions with incoming human immunodeficiency virus type 1 capsids with opposing effects on infectivity in human cells." Journal of virology **79**(1): 176-183.
- Haudecoeur, R., M. Peuchmaur, et al. (2012). "Structure-Activity Relationships in the Development of Allosteric Hepatitis C Virus RNA-Dependent RNA Polymerase Inhibitors: Ten Years of Research." Medicinal research reviews.
- Hayer, J., C. Rodriguez, et al. (2012). "HBV RNase H sequence variability investigated by pyrosequencing, sequence analysis and molecular modeling." submitted.
- Henderson, L. E., M. A. Bowers, et al. (1992). "Gag proteins of the highly replicative MN strain of human immunodeficiency virus type 1: posttranslational modifications, proteolytic processings, and complete amino acid sequences." Journal of virology **66**(4): 1856-1865.
- Hezode, C., N. Forestier, et al. (2009). "Telaprevir and peginterferon with or without ribavirin for chronic HCV infection." The New England journal of medicine **360**(18): 1839-1850.
- Hirsch, M. S., H. F. Gunthard, et al. (2008). "Antiretroviral drug resistance testing in adult HIV-1 infection: 2008 recommendations of an International AIDS Society-USA panel." Clinical infectious diseases : an official publication of the Infectious Diseases Society of America **47**(2): 266-285.

- Hofmann, W. P. and S. Zeuzem (2011). "A new standard of care for the treatment of chronic HCV infection." Nature reviews. Gastroenterology & hepatology **8**(5): 257-264.
- Hoppman-Chaney, N., L. M. Peterson, et al. (2010). "Evaluation of oligonucleotide sequence capture arrays and comparison of next-generation sequencing platforms for use in molecular diagnostics." Clinical chemistry **56**(8): 1297-1306.
- Hunt, C. M., J. M. McGill, et al. (2000). "Clinical relevance of hepatitis B viral mutations." Hepatology **31**(5): 1037-1044.
- Huxley, J. (1942). Evolution, the modern synthesis. London,, G. Allen & Unwin Ltd.
- Ikeda, M., M. Yi, et al. (2002). "Selectable subgenomic and genome-length dicistronic RNAs derived from an infectious molecular clone of the HCV-N strain of hepatitis C virus replicate efficiently in cultured Huh7 cells." Journal of virology **76**(6): 2997-3006.
- Jacob, F. (1981). Le jeu des possibles : essai sur la diversité du vivant. Paris, Fayard.
- Ji, J. P. and L. A. Loeb (1992). "Fidelity of HIV-1 reverse transcriptase copying RNA in vitro." Biochemistry **31**(4): 954-958.
- Johnson, J. A., J. F. Li, et al. (2008). "Minority HIV-1 drug resistance mutations are present in antiretroviral treatment-naive populations and associate with reduced treatment efficacy." PLoS medicine **5**(7): e158.
- Jonas, M. M., D. Kelly, et al. (2008). "Safety, efficacy, and pharmacokinetics of adefovir dipivoxil in children and adolescents (age 2 to <18 years) with chronic hepatitis B." Hepatology **47**(6): 1863-1871.
- Katayanagi, K., M. Ishikawa, et al. (1993). "Crystal structures of ribonuclease HI active site mutants from Escherichia coli." The Journal of biological chemistry **268**(29): 22092-22099.
- Katayanagi, K., M. Okumura, et al. (1993). "Crystal structure of Escherichia coli RNase HI in complex with Mg²⁺ at 2.8 Å resolution: proof for a single Mg(2+)-binding site." Proteins **17**(4): 337-346.
- Kato, T., T. Date, et al. (2003). "Efficient replication of the genotype 2a hepatitis C virus subgenomic replicon." Gastroenterology **125**(6): 1808-1817.
- Katze, M. G., Y. He, et al. (2002). "Viruses and interferon: a fight for supremacy." Nature reviews. Immunology **2**(9): 675-687.
- Kaul, A., S. Stauffer, et al. (2009). "Essential role of cyclophilin A for hepatitis C virus replication and virus production and possible link to polyprotein cleavage kinetics." PLoS pathogens **5**(8): e1000546.
- Keele, B. F., E. E. Giorgi, et al. (2008). "Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection." Proceedings of the National Academy of Sciences of the United States of America **105**(21): 7552-7557.
- Kilby, J. M. and J. J. Eron (2003). "Novel therapies based on mechanisms of HIV-1 cell entry." The New England journal of medicine **348**(22): 2228-2238.
- King, R. W., S. K. Ladner, et al. (1998). "Inhibition of human hepatitis B virus replication by AT-61, a phenylpropenamide derivative, alone and in combination with (-)beta-L-2',3'-dideoxy-3'-thiacytidine." Antimicrobial agents and chemotherapy **42**(12): 3179-3186.
- Kolykhalov, A. A., E. V. Agapov, et al. (1997). "Transmission of hepatitis C by intrahepatic inoculation with transcribed RNA." Science **277**(5325): 570-574.
- Kosmrlj, A., E. L. Read, et al. (2010). "Effects of thymic selection of the T-cell repertoire on HLA class I-associated control of HIV infection." Nature **465**(7296): 350-354.
- Krieger, N., V. Lohmann, et al. (2001). "Enhancement of hepatitis C virus RNA replication by cell culture-adaptive mutations." Journal of virology **75**(10): 4614-4624.
- Kuritzkes, D. R., C. M. Lalama, et al. (2008). "Preexisting resistance to nonnucleoside reverse-transcriptase inhibitors predicts virologic failure of an efavirenz-based regimen in treatment-naive HIV-1-infected subjects." The Journal of infectious diseases **197**(6): 867-870.
- Kwong, A. D., I. Najera, et al. (2011). "Sequence and phenotypic analysis for resistance monitoring in hepatitis C virus drug development: recommendations from the HCV DRAG." Gastroenterology **140**(3): 755-760.

- Lai, C. L., N. Leung, et al. (2005). "A 1-year trial of telbivudine, lamivudine, and the combination in patients with hepatitis B e antigen-positive chronic hepatitis B." Gastroenterology **129**(2): 528-536.
- Lalezari, J. P., K. Henry, et al. (2003). "Enfuvirtide, an HIV-1 fusion inhibitor, for drug-resistant HIV infection in North and South America." The New England journal of medicine **348**(22): 2175-2185.
- Lamarck, J. B. P. A. d. M. d. (1830). Philosophie zoologique. Paris,, G. Baillière; etc.
- Lau, J. Y., R. C. Tam, et al. (2002). "Mechanism of action of ribavirin in the combination treatment of chronic HCV infection." Hepatology **35**(5): 1002-1009.
- Le Pogam, S., H. Kang, et al. (2006). "Selection and characterization of replicon variants dually resistant to thumb- and palm-binding nonnucleoside polymerase inhibitors of the hepatitis C virus." Journal of virology **80**(12): 6146-6154.
- Lemon, S. M., J. A. McKeating, et al. (2010). "Development of novel therapies for hepatitis C." Antiviral research **86**(1): 79-92.
- Lin, C. (2006). HCV NS3-4A Serine Protease. Hepatitis C Viruses: Genomes and Molecular Biology. S. L. Tan. Norfolk (UK).
- Lindenbach, B. D., M. J. Evans, et al. (2005). "Complete replication of hepatitis C virus in cell culture." Science **309**(5734): 623-626.
- Little, S. J., S. Holte, et al. (2002). "Antiretroviral-drug resistance among patients recently infected with HIV." The New England journal of medicine **347**(6): 385-394.
- Liu, Z., F. Yang, et al. (2009). "Critical role of cyclophilin A and its prolyl-peptidyl isomerase activity in the structure and function of the hepatitis C virus replication complex." Journal of virology **83**(13): 6554-6565.
- Lohmann, V., F. Korner, et al. (1999). "Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line." Science **285**(5424): 110-113.
- Lok, A., D. Gardiner, et al. (2010). "Quadruple therapy with BMS-790052, BMS-650032 and PEG-IFN/RBV for 24 weeks results in 100% SVR12 in HCV genotype 1 null responders." Hepatology **52**(Suppl): 877A.
- Lok, A. S., F. Zoulim, et al. (2007). "Antiviral drug-resistant HBV: standardization of nomenclature and assays and recommendations for management." Hepatology **46**(1): 254-265.
- Lynch, M. (2010). "Rate, molecular spectrum, and consequences of human mutation." Proceedings of the National Academy of Sciences of the United States of America **107**(3): 961-968.
- Macalalad, A. R., M. C. Zody, et al. (2012). "Highly sensitive and specific detection of rare variants in mixed viral populations from massively parallel sequence data." PLoS computational biology **8**(3): e1002417.
- Malthus, T. R. (1798). An essay on the principle of population. London,, J. Johnson.
- Maqbool MA, Imache MR, et al. (2012). "Regulation of Hepatitis C Virus Replication by Nuclear Translocation of Nonstructural 5A Protein and Transcriptional Activation of Host Genes." J virol in press.
- Mardis, E. R. (2008). "The impact of next-generation sequencing technology on genetics." Trends in genetics : TIG **24**(3): 133-141.
- Margulies, M., M. Egholm, et al. (2005). "Genome sequencing in microfabricated high-density picolitre reactors." Nature **437**(7057): 376-380.
- McCarthy (1996). Chromas, Griffith University, Brisbane.
- McCown, M. F., S. Rajyaguru, et al. (2008). "The hepatitis C virus replicon presents a higher barrier to resistance to nucleoside analogs than to nonnucleoside polymerase or protease inhibitors." Antimicrobial agents and chemotherapy **52**(5): 1604-1612.
- McHutchison, J. G., M. P. Manns, et al. (2010). "Telaprevir for previously treated chronic HCV infection." The New England journal of medicine **362**(14): 1292-1303.
- McMillan, J. S., T. Shaw, et al. (1995). "Effect of immunosuppressive and antiviral agents on hepatitis B virus replication in vitro." Hepatology **22**(1): 36-43.
- Mendel, G. J. (1865). "Versuche uber Pflanzen-Hybriden."

- Metzker, M. L. (2010). "Sequencing technologies - the next generation." Nature reviews. Genetics **11**(1): 31-46.
- Mohanty, S. R., S. S. Kupfer, et al. (2006). "Treatment of chronic hepatitis B." Nature clinical practice. Gastroenterology & hepatology **3**(8): 446-458.
- Moradpour, D., M. J. Evans, et al. (2004). "Insertion of green fluorescent protein into nonstructural protein 5A allows direct visualization of functional hepatitis C virus replication complexes." Journal of virology **78**(14): 7400-7409.
- Morgan, T. H. (1915). The mechanism of Mendelian heredity. New York,, Holt.
- Murray, J. M., R. H. Purcell, et al. (2006). "The half-life of hepatitis B virions." Hepatology **44**(5): 1117-1121.
- Nathans, R., H. Cao, et al. (2008). "Small-molecule inhibition of HIV-1 Vif." Nature biotechnology **26**(10): 1187-1192.
- Nicholas, K. (1997). GeneDoc: analysis and visualization of genetic variation, Embnew.
- Nowotny, M., S. A. Gaidamakov, et al. (2007). "Structure of human RNase H1 complexed with an RNA/DNA hybrid: insight into HIV reverse transcription." Molecular cell **28**(2): 264-276.
- Nozza, S., F. Canducci, et al. (2012). "Viral tropism by geno2pheno as a tool for predicting CD4 decrease in HIV-1-infected naive patients with high CD4 counts." The Journal of antimicrobial chemotherapy **67**(5): 1224-1227.
- Ortiz, R., E. Dejesus, et al. (2008). "Efficacy and safety of once-daily darunavir/ritonavir versus lopinavir/ritonavir in treatment-naive HIV-1-infected patients at week 48." AIDS **22**(12): 1389-1397.
- Paeshuyse, J., A. Kaul, et al. (2006). "The non-immunosuppressive cyclosporin DEBIO-025 is a potent inhibitor of hepatitis C virus replication in vitro." Hepatology **43**(4): 761-770.
- Page, R. D. (1996). "TreeView: an application to display phylogenetic trees on personal computers." Computer applications in the biosciences : CABIOS **12**(4): 357-358.
- Paley, W. (1802). Natural theology. London,, Printed for R. Faulder by Wilks and Taylor.
- Pallier, C., L. Castera, et al. (2006). "Dynamics of hepatitis B virus resistance to lamivudine." Journal of virology **80**(2): 643-653.
- Pallier, C., C. Rodriguez, et al. (2009). "Complex dynamics of hepatitis B virus resistance to adefovir." Hepatology **49**(1): 50-59.
- Palmer, S., V. Boltz, et al. (2006). "Selection and persistence of non-nucleoside reverse transcriptase inhibitor-resistant HIV-1 in patients starting and stopping non-nucleoside therapy." AIDS **20**(5): 701-710.
- Pandey, V., R. Nutter, et al. (2008). Next Generation Genome Sequencing: Towards Personalized Medicine.
- Paredes, R. and B. Clotet (2010). "Clinical management of HIV-1 resistance." Antiviral research **85**(1): 245-265.
- Park, S. G., Y. Kim, et al. (2003). "Fidelity of hepatitis B virus polymerase." European journal of biochemistry / FEBS **270**(14): 2929-2936.
- Pastor, R., F. Habersetzer, et al. (2009). "Hepatitis B virus mutations potentially conferring adefovir/tenofovir resistance in treatment-naive patients." World journal of gastroenterology : WJG **15**(6): 753-755.
- Pawlotsky, J. M. (2005). "The concept of hepatitis B virus mutant escape." Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology **34 Suppl 1**: S125-129.
- Pawlotsky, J. M. (2009). "Hepatitis: HCV variability, the immune system and resistance to antiviral drugs." Nature reviews. Gastroenterology & hepatology **6**(7): 383-385.
- Pawlotsky, J. M. (2011). "Treatment failure and resistance with direct-acting antiviral drugs against hepatitis C virus." Hepatology **53**(5): 1742-1751.
- Pawlotsky, J. M. (2012). "Is hepatitis virus resistance to antiviral drugs a threat?" Gastroenterology **142**(6): 1369-1372.
- Pawlotsky, J. M., H. Dahari, et al. (2004). "Antiviral action of ribavirin in chronic hepatitis C." Gastroenterology **126**(3): 703-714.

- Pawlotsky, J. M., G. Dusheiko, et al. (2008). "Virologic monitoring of hepatitis B virus therapy in clinical trials and practice: recommendations for a standardized approach." Gastroenterology **134**(2): 405-415.
- Pawlotsky, J. M., I. Najera, et al. (2012). "Resistance to mericitabine, a nucleoside analogue inhibitor of HCV RNA-dependent RNA polymerase." Antiviral therapy **17**(3): 411-423.
- Petropoulos, C. J., N. T. Parkin, et al. (2000). "A novel phenotypic drug susceptibility assay for human immunodeficiency virus type 1." Antimicrobial agents and chemotherapy **44**(4): 920-928.
- Polyak, S. J., K. S. Khabar, et al. (2001). "Hepatitis C virus nonstructural 5A protein induces interleukin-8, leading to partial inhibition of the interferon-induced antiviral response." Journal of virology **75**(13): 6095-6106.
- Pommier, Y., A. A. Johnson, et al. (2005). "Integrase inhibitors to treat HIV/AIDS." Nature reviews. Drug discovery **4**(3): 236-248.
- Poon, A. F., L. C. Swenson, et al. (2010). "Phylogenetic analysis of population-based and deep sequencing data to identify coevolving sites in the nef gene of HIV-1." Molecular biology and evolution **27**(4): 819-832.
- Poordad, F., J. McCone, Jr., et al. (2011). "Boceprevir for untreated chronic HCV genotype 1 infection." The New England journal of medicine **364**(13): 1195-1206.
- Prober, J. M., G. L. Trainor, et al. (1987). "A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides." Science **238**(4825): 336-341.
- Ptak, R. G., P. A. Gallay, et al. (2008). "Inhibition of human immunodeficiency virus type 1 replication in human cells by Debio-025, a novel cyclophilin binding agent." Antimicrobial agents and chemotherapy **52**(4): 1302-1317.
- Puyang, X., D. L. Poulin, et al. (2010). "Mechanism of resistance of hepatitis C virus replicons to structurally distinct cyclophilin inhibitors." Antimicrobial agents and chemotherapy **54**(5): 1981-1987.
- Qi, X., S. Xiong, et al. (2007). "In vitro susceptibility of adefovir-associated hepatitis B virus polymerase mutations to other antiviral agents." Antiviral therapy **12**(3): 355-362.
- R_Development_Core_Team (2008). R: A Language and Environment for Statistical Computing.
- Recordon-Pinson, P., C. Soulie, et al. (2010). "Evaluation of the genotypic prediction of HIV-1 coreceptor use versus a phenotypic assay and correlation with the virological response to maraviroc: the ANRS GenoTropism study." Antimicrobial agents and chemotherapy **54**(8): 3335-3340.
- Reeves, J. D., F. H. Lee, et al. (2005). "Enfuvirtide resistance mutations: impact on human immunodeficiency virus envelope function, entry inhibitor sensitivity, and virus neutralization." Journal of virology **79**(8): 4991-4999.
- Resch, W., N. Hoffman, et al. (2001). "Improved success of phenotype prediction of the human immunodeficiency virus type 1 from envelope variable loop 3 sequence using neural networks." Virology **288**(1): 51-62.
- Richterich, P. (1998). "Estimation of errors in "raw" DNA sequences: a validation study." Genome research **8**(3): 251-259.
- Rimsky, L. T., D. C. Shugars, et al. (1998). "Determinants of human immunodeficiency virus type 1 resistance to gp41-derived inhibitory peptides." Journal of virology **72**(2): 986-993.
- Robin, X., N. Turck, et al. (2011). "pROC: an open-source package for R and S+ to analyze and compare ROC curves." BMC bioinformatics **12**: 77.
- Rodriguez, C., C. Chevaliez, et al. (2011). "Ultra-deep pyrosequencing, a powerful new tool to study HBV Resistance to Nucleoside/Nucleotide analogues, identifies pre-existing HBV Variants bearing resistance mutations and characterizes their on-treatment kinetics." Hepatology **52**(S4): 117A.
- Rodriguez, C., S. Chevaliez, et al. (2012). "Dynamics of Hepatitis B Virus Resistance Revisited by Ultra-Deep Pyrosequencing and PyroPack© Analysis." Journal of Virology (in press).
- Rodriguez, C., C. Soulie, et al. (2011). "HIV-1 coreceptor usage prediction by "Ultra-deep pyrosequencing" increases the rate of patients treated with success by CCR5 antagonists." Antiviral therapy **16**(Supplement 1): A100.

- Rong, L., H. Dahari, et al. (2010). "Rapid emergence of protease inhibitor resistance in hepatitis C virus." Science translational medicine **2**(30): 30ra32.
- Rothberg, J. M., W. Hinz, et al. (2011). "An integrated semiconductor device enabling non-optical genome sequencing." Nature **475**(7356): 348-352.
- Sanger, F., G. M. Air, et al. (1977). "Nucleotide sequence of bacteriophage phi X174 DNA." Nature **265**(5596): 687-695.
- Santoro, M. M., D. Armenia, et al. (2012). "The lowest X4 Geno2Pheno false-positive rate is associated with greater CD4 depletion in HIV-1 infected patients." Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases **18**(8): E289-298.
- Sarrazin, C., C. Hezode, et al. (2012). "Antiviral strategies in hepatitis C virus infection." Journal of hepatology **56 Suppl 1**: S88-100.
- Sarrazin, C., T. L. Kieffer, et al. (2007). "Dynamic hepatitis C virus genotypic and phenotypic changes in patients treated with the protease inhibitor telaprevir." Gastroenterology **132**(5): 1767-1777.
- Seclen, E., V. Soriano, et al. (2011). "High concordance between the position-specific scoring matrix and geno2pheno algorithms for genotypic interpretation of HIV-1 tropism: V3 length as the major cause of disagreement." Journal of clinical microbiology **49**(9): 3380-3382.
- Sheldon, J., N. Camino, et al. (2005). "Selection of hepatitis B virus polymerase mutations in HIV-coinfected patients treated with tenofovir." Antiviral therapy **10**(6): 727-734.
- Shen, Y., S. Sarin, et al. (2008). "Comparing platforms for C. elegans mutant identification using high-throughput whole-genome sequencing." PloS one **3**(12): e4012.
- Shet, A., L. Berry, et al. (2006). "Tracking the prevalence of transmitted antiretroviral drug-resistant HIV-1: a decade of experience." Journal of acquired immune deficiency syndromes **41**(4): 439-446.
- Simen, B. B., J. F. Simons, et al. (2009). "Low-abundance drug-resistant viral variants in chronically HIV-infected, antiretroviral treatment-naive patients significantly impact treatment outcomes." The Journal of infectious diseases **199**(5): 693-701.
- Smith, L. M. (1991). "High-speed DNA sequencing by capillary gel electrophoresis." Nature **349**(6312): 812-813.
- Smith, T. F. and M. S. Waterman (1981). "Identification of common molecular subsequences." Journal of molecular biology **147**(1): 195-197.
- Soriano, V., A. S. Perelson, et al. (2008). "Why are there different dynamics in the selection of drug resistance in HIV and hepatitis B and C viruses?" The Journal of antimicrobial chemotherapy **62**(1): 1-4.
- Soriano, V., E. Vispo, et al. (2011). "Directly acting antivirals against hepatitis C virus." The Journal of antimicrobial chemotherapy **66**(8): 1673-1686.
- Spencer, H. (1864). The principles of biology. London etc., William and Norgate.
- Stanley, W. M. (1935). "Isolation of a Crystalline Protein Possessing the Properties of Tobacco-Mosaic Virus." Science **81**(2113): 644-645.
- Stuyver, L., R. Rossau, et al. (1993). "Typing of hepatitis C virus isolates and characterization of new subtypes using a line probe assay." The Journal of general virology **74 (Pt 6)**: 1093-1102.
- Susser, S., C. Welsch, et al. (2009). "Characterization of resistance to the protease inhibitor boceprevir in hepatitis C virus-infected patients." Hepatology **50**(6): 1709-1718.
- Suzuki, S., N. Ono, et al. (2011). "Comparison of sequence reads obtained from three next-generation sequencing platforms." PloS one **6**(5): e19534.
- Svarovskaia, E. S., N. A. Margot, et al. (2007). "Low-level K65R mutation in HIV-1 reverse transcriptase of treatment-experienced patients exposed to abacavir or didanosine." Journal of acquired immune deficiency syndromes **46**(2): 174-180.
- Swenson, L. C., T. Mo, et al. (2011). "Deep sequencing to infer HIV-1 co-receptor usage: application to three clinical trials of maraviroc in treatment-experienced patients." The Journal of infectious diseases **203**(2): 237-245.

- Swenson, L. C., T. Mo, et al. (2011). "Deep V3 sequencing for HIV type 1 tropism in treatment-naive patients: a reanalysis of the MERIT trial of maraviroc." Clinical infectious diseases : an official publication of the Infectious Diseases Society of America **53**(7): 732-742.
- Swenson, L. C., A. Moores, et al. (2010). "Improved detection of CXCR4-using HIV by V3 genotyping: application of population-based and "deep" sequencing to plasma RNA and proviral DNA." Journal of acquired immune deficiency syndromes **54**(5): 506-510.
- Tenney, D. J., S. M. Levine, et al. (2004). "Clinical emergence of entecavir-resistant hepatitis B virus requires additional substitutions in virus already resistant to Lamivudine." Antimicrobial agents and chemotherapy **48**(9): 3498-3507.
- Tenney, D. J., R. E. Rose, et al. (2009). "Long-term monitoring shows hepatitis B virus resistance to entecavir in nucleoside-naive patients is rare through 5 years of therapy." Hepatology **49**(5): 1503-1514.
- Thielen, A. and T. Lengauer (2012). "Geno2pheno[454]: a Web server for the prediction of HIV-1 coreceptor usage from next-generation sequencing data." Intervirology **55**(2): 113-117.
- Thomas, D. L., C. L. Thio, et al. (2009). "Genetic variation in IL28B and spontaneous clearance of hepatitis C virus." Nature **461**(7265): 798-801.
- Thompson, M. A., J. A. Aberg, et al. (2010). "Antiretroviral treatment of adult HIV infection: 2010 recommendations of the International AIDS Society-USA panel." JAMA : the journal of the American Medical Association **304**(3): 321-333.
- Tong, X., S. Bogen, et al. (2008). "Characterization of resistance mutations against HCV ketoamide protease inhibitors." Antiviral research **77**(3): 177-185.
- Tramontano, E. and R. Di Santo (2010). "HIV-1 RT-associated RNase H function inhibitors: Recent advances in drug development." Current medicinal chemistry **17**(26): 2837-2853.
- Tsibris, A. M., B. Korber, et al. (2009). "Quantitative deep sequencing reveals dynamic HIV-1 escape and large population shifts during CCR5 antagonist therapy in vivo." PloS one **4**(5): e5683.
- Tuske, S., S. G. Sarafianos, et al. (2004). "Structures of HIV-1 RT-DNA complexes before and after incorporation of the anti-AIDS drug tenofovir." Nature structural & molecular biology **11**(5): 469-474.
- Vandekerckhove, L., C. Verhofstede, et al. (2009). "Maraviroc: perspectives for use in antiretroviral-naive HIV-1-infected patients." The Journal of antimicrobial chemotherapy **63**(6): 1087-1096.
- Venkatraman, E. S. (2000). "A permutation test to compare receiver operating characteristic curves." Biometrics **56**(4): 1134-1138.
- Venkatraman ES, B. C. (1996). "A distribution-free procedure for comparing receiver operating characteristic curves from a paired experiment." Biometrika **83**: 835-848.
- Vercruysse, T., E. Pardon, et al. (2010). "An intrabody based on a llama single-domain antibody targeting the N-terminal alpha-helical multimerization domain of HIV-1 rev prevents viral production." The Journal of biological chemistry **285**(28): 21768-21780.
- Walter, N. A., D. Bottomly, et al. (2009). "High throughput sequencing in mice: a platform comparison identifies a preponderance of cryptic SNPs." BMC genomics **10**: 379.
- Wang, C., Y. Mitsuya, et al. (2007). "Characterization of mutation spectra with ultra-deep pyrosequencing: application to HIV-1 drug resistance." Genome research **17**(8): 1195-1201.
- Wang, Y., F. Kong, et al. (2008). "Use of a multiplex PCR-based reverse line blot (mPCR/RLB) hybridisation assay for the rapid identification of bacterial pathogens." Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases **14**(2): 155-160.
- Weber, O., K. H. Schlemmer, et al. (2002). "Inhibition of human hepatitis B virus (HBV) by a novel non-nucleosidic compound in a transgenic mouse model." Antiviral research **54**(2): 69-78.
- Weismann, A., W. N. Parker, et al. (1893). The germ-plasm: a theory of heredity. New York,, C. Scribner's sons.
- Wensing, A. M., D. A. van de Vijver, et al. (2005). "Prevalence of drug-resistant HIV-1 variants in untreated individuals in Europe: implications for clinical management." The Journal of infectious diseases **192**(6): 958-966.

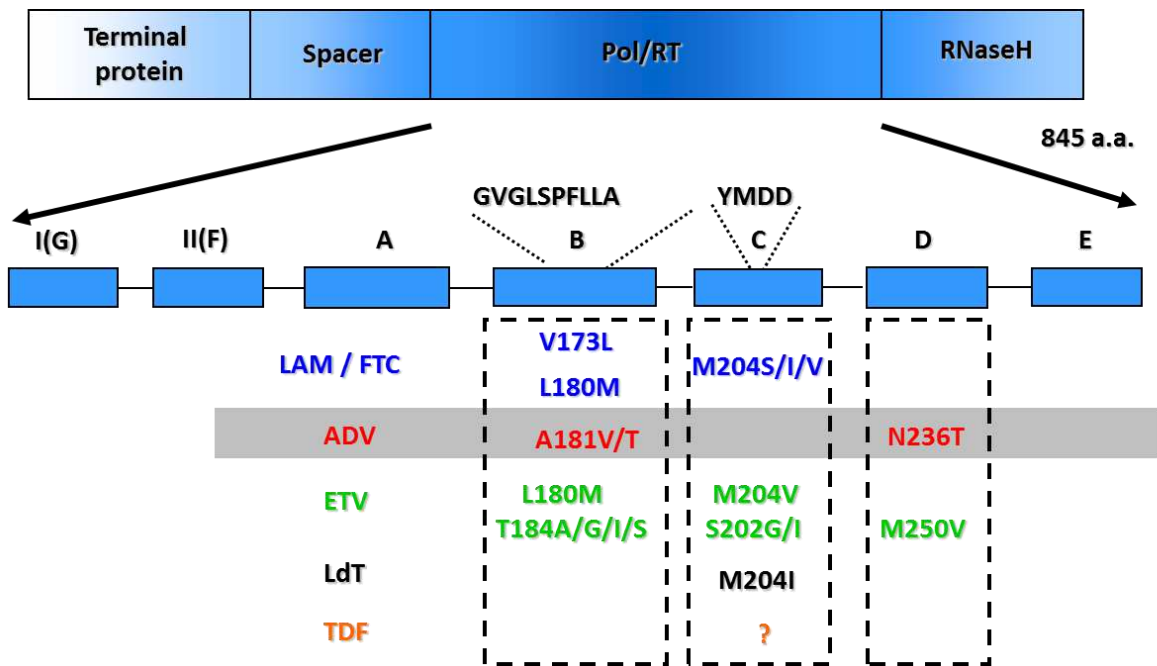
- Westby, M., C. Smith-Burchnell, et al. (2007). "Reduced maximal inhibition in phenotypic susceptibility assays indicates that viral strains resistant to the CCR5 antagonist maraviroc utilize inhibitor-bound receptor for entry." Journal of virology **81**(5): 2359-2371.
- Wick, W. D., P. B. Gilbert, et al. (2009). "Predicting the impact of blocking human immunodeficiency virus type 1 Nef in vivo." Journal of virology **83**(5): 2349-2356.
- Yeni, P. (2010). Prise en charge médicale des personnes infectées par le VIH. santé.
- Yerly, S., L. Kaiser, et al. (1999). "Transmission of antiretroviral-drug-resistant HIV-1 variants." Lancet **354**(9180): 729-733.
- Yi, M., F. Bodola, et al. (2002). "Subgenomic hepatitis C virus replicons inducing expression of a secreted enzymatic reporter protein." Virology **304**(2): 197-210.
- Yi, M. and S. M. Lemon (2004). "Adaptive mutations producing efficient replication of genotype 1a hepatitis C virus RNA in normal Huh7 cells." Journal of virology **78**(15): 7904-7915.
- Zagordi, O., L. Geyrhofer, et al. (2010). "Deep sequencing of a genetically heterogeneous sample: local haplotype reconstruction and read error correction." Journal of computational biology : a journal of computational molecular cell biology **17**(3): 417-428.
- Zagordi, O., R. Klein, et al. (2010). "Error correction of next-generation sequencing data and reliable estimation of HIV quasispecies." Nucleic acids research **38**(21): 7400-7409.
- Zeldin, R. K. and R. A. Petruschke (2004). "Pharmacological and therapeutic properties of ritonavir-boosted protease inhibitor therapy in HIV-infected patients." The Journal of antimicrobial chemotherapy **53**(1): 4-9.
- Zeuzem, S. and G. R. Foster (2012). Clinical Implications of HCV Resistance, Clinical Care Options Hepatitis.
- Zhong, Y., J. Lv, et al. (2012). "Prevalence, virology and antiviral drugs susceptibility of hepatitis B virus rtN238H polymerase mutation from 1865 Chinese patients with chronic hepatitis B." Antiviral research **93**(1): 185-190.
- Zhu, X. F., L. X. Lu, et al. (2011). "Effect and Predictive Elements for 52 Weeks' Telbivudine Treatment on Naive HBeAg positive Chronic Hepatitis B." Hepatitis monthly **11**(12): 980-985.

ANNEXE 1 : MUTATIONS de RESISTANCE CONNUES pour le VHB, VHC, VIH

Target	Variant	NS3 Linear	NS3 Macrocyclic	NS5A inhibitor	NS5B nucleoside	NS5B Palm	NS5B Thumb	NS5B Finger	IFN	RBV
NS3 Protease	V36M	R	S	S	S	S	S	S	S	S
	T54A	R	S	S	S	S	S	S	S	S
	R155K	R	R	S	S	S	S	S	S	S
	A156T	R	R	S	S	S	S	S	S	S
	D168V	S	R	S	S	S	S	S	S	S
NS5A	L28V	S	S	R	S	S	S	S	S	S
	Y93H	S	S	R	S	S	S	S	S	S
NS5B	S282T	S	S	S	R	S	S	S	S	S
	C316Y	S	S	S	S	R	S	S	S	S
	M414T	S	S	S	S	R	S	S	S	S
	R422K	S	S	S	S	S	R	S	S	S
	M423T	S	S	S	S	S	R	S	S	S
	P495S	S	S	S	S	S	S	R	S	S

IFN, interferon; R, resistant (> 4-fold increase in 50% effective concentration); RBV, ribavirin; S, susceptible (< 4-fold shift in HCV replicon 50% effective concentration).

Annexe 1A : Principales mutations de résistance sélectionnées au cours des traitements du VHC par inhibiteurs de protéase (NS3 protease), inhibiteurs de la protéine NS5A et inhibiteurs de polymérase (NS5B) (Zeuzem and Foster 2012)



Annexe 1B : Principales mutations associées à la résistance du VHB aux inhibiteurs nucléos(t)idiques : lamivudine (LAM), emtricitabine (FTC), entecavir (ETV), telbivudine (LdT), adefovir (ADV) ou tenofovir (TDF) (Allen *et al.* 1998; Tenney *et al.* 2004; Gish *et al.* 2005; Lai *et al.* 2005; Sheldon *et al.* 2005).

MUTATIONS IN THE REVERSE TRANSCRIPTASE GENE ASSOCIATED WITH RESISTANCE TO REVERSE TRANSCRIPTASE INHIBITORS

Nucleoside and Nucleotide Reverse Transcriptase Inhibitors (NRTIs)

Multi-nRTI Resistance: 69 Insertion Complex; affects all nRTIs currently approved by the US FDA

L 41 62 69 70 **210 215 219**
M A K **W Y Q**
Insert R **W Y Q**

Multi-nRTI Resistance: 151 Complex; affects all nRTIs currently approved by the US FDA except tenofovir

V 62 75 77 **116 151**
A V F O **W Y Q**

Multi-nRTI Resistance: Thymidine Analogue-associated Mutations (TAMs); affect all nRTIs currently approved by the US FDA

M D K **210 215 219**
L N R **L T K**
W Y Q

Abacavir

K 65 74 **115 184**
R V Y F M

Didanosine

K 65 74 **184**
R V Y F M

Emtricitabine

K 65 **184**
R V Y F M

Lamivudine

K 65 **184**
R V Y F M

Stavudine

M D K **210 215 219**
L N R **L T K**
W Y Q

Tenofovir

K 65 70 **184**
R E Y F M

Zidovudine

M D K **210 215 219**
L N R **L T K**
W Y Q

Nonnucleoside Reverse Transcriptase Inhibitors (NNRTIs)

L K V V Y Y G **225**
I N M I C L A **P H**

Etravirine

V A L K V V Y G
I G I F I T V A S

Nevirapine

L K V V Y Y G **181 188 190**
I N A I C C A **L H**

MUTATIONS IN THE PROTEASE GENE ASSOCIATED WITH RESISTANCE TO PROTEASE INHIBITORS

Atazanavir +/- ritonavir

L G K L V L E M G I F I D I I A G V I I N L I
10 16 40 24 32 33 34 36 46 48 50 53 54 60 62 64 71 73 82 84 86 88 90 93
F E R I T Q L L V L Y V E V I V C T V V S M M
C I I V V M M I T I A F I

Darunavir/ritonavir

V V L I I I J I G L I L
11 32 33 47 50 54 73 76 84 89
I F V V M S V V V

Fosamprenavir/ritonavir

L V V M I I J I G L V I L
10 32 46 47 50 54 73 76 82 84 90
I I V V V S V A V M

Indinavir/ritonavir

L K L V M M I I A G L V V I L
10 20 24 32 36 46 54 71 73 76 77 82 84 90
R R I I V V Y S V I A V M

Lopinavir/ritonavir

L K L V L M I I F I J I A G L V I L
10 20 24 32 33 46 47 50 53 54 63 71 73 76 82 84 90
F M I I F L A V L V P V S V A V M
R R A A A T S

Nelfinavir

L D M M A V V I N L
10 30 36 46 71 77 82 84 88 90
F N I L I A V D M
I L T F S

Saquinavir/ritonavir

L L L G I I A G V V I L
10 24 48 54 62 71 73 77 82 84 90
I I V V V V V I A V M
R V L T S

Tipranavir/ritonavir

L I K L E M K M I J Q H T V N I L
10 13 20 33 35 36 43 46 47 54 58 69 74 82 84 90
V V M R F G I T L V A E K P L D V M
I R V V

MUTATIONS IN THE ENVELOPE GENE ASSOCIATED WITH RESISTANCE TO ENTRY INHIBITORS

G I V Q N N
36 37 38 39 40 42 43
S M E
D V A R H T D

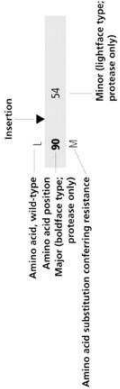
Maraviroc

See User Note (ref 7)

MUTATIONS IN THE INTEGRASE GENE ASSOCIATED WITH RESISTANCE TO INTEGRASE INHIBITORS

O N 148 155
H H K K R

MUTATIONS



Annexe 1C : Principales mutations associées à la résistance du VIH aux traitements par INTI, INNTI, inhibiteurs de protéase, inhibiteurs d'entrée et inhibiteurs d'intégrase (Hirsch et al. 2008)

ANNEXE 2 : BREVETS

- PyroClass® : IDDN.FR.001.360005.001.S.P.2010.000.31230
- PyroMute® : IDDN.FR.001.120008.001.S.C.2010.000.31230
- PyroDyn® : IDDN.FR.001.370024.000.S.P.2010.000.31230
- PyroLink® : IDDN.FR.001.180032.000.S.P.2011.000.31230
- PyroTrop® : IDDN.FR.001.180031.000.S.P.2011.000.31230

Inter Deposit Digital Number

Certificat délivré par

Agence pour la Protection des Programmes

249, rue de Crimée - 75019 PARIS. Tél 33 (1) 40 35 03 03. Fax 33 (1) 40 38 96 43

IDDN.FR.001.180031.000.S.P.2011.000.31230

Pour l'oeuvre : **Pyrotrop**

Identité du(des) titulaire(s) des droits
INSERM-Institut National de la Santé
et de la Rech. Médicale
101 rue de Tolbiac
75654 PARIS Cedex 13
Siret : 180038048

Adhérent sous le numéro : **07.75.1**
Support utilisé : 1 CD-Rom en dou

Le titulaire *

Fait à Paris, le 05/05/2011



Logibox conservé par l'adhérent : 67295
Logibox conservé par l'APP : 67296

(1) Inter Deposit Digital Number
(2) Nationalité de l'oeuvre

(3) Numéro de l'organisme d'enregistrement
(4) Numéro d'ordre de l'enregistrement

(5) Numéro de version
(6) Type d'enregistrement

(7) Type de l'oeuvre
(8) Année d'enregistrement

* Le titulaire s'engage à informer l'APP de toute cession ou aliénation, totale ou partielle, de ses droits de propriété intellectuelle.
Seules les inscriptions de type S et C permettent un éventuel accès au programme source.

Agence pour la Protection des Programmes
249, rue de Crimée - 75019 PARIS
Tél. 33 (1) 40 35 03 03
Fax 33 (1) 40 38 96 43
Site : www.appr.fr
Siret : 38538588900027 - APE 913 E

Inter Deposit Digital Number

Certificat délivré par

Agence pour la Protection des Programmes

249, rue de Crimée - 75019 PARIS. Tél 33 (1) 40 36 03 03. Fax 33 (1) 40 38 96 43

IDDN.FR.001.180032.000.S.P.2011.000.31230

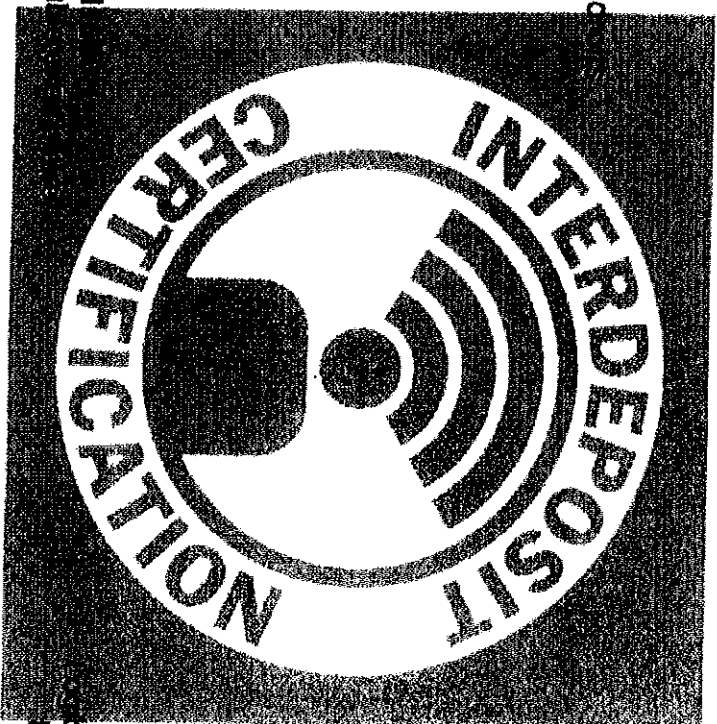
Pour l'oeuvre : **Pyrolink**

Identité du(des) titulaire(s) des droits
INSERM-Institut National de la Santé
et de la Rech. Médicale
101 rue de Tolbiac
75654 PARIS Cedex 13
Siret : 180036048

Adhérent sous le numéro : **07.75.1**
Support utilisé : 1 CD-Rom en dou

Le titulaire *

Fait à Paris, le 05/05/2011



Logibox conservé par l'adhérent : 67297
Logibox conservé par l'APP : 67298

(1) Inter Deposit Digital Number
(2) Nationalité de l'oeuvre

(3) Numéro de l'organisme d'enregistrement
(4) Numéro d'ordre de l'enregistrement

(5) Numéro de version
(6) Type d'enregistrement

(7) Type de l'oeuvre
(8) Année d'enregistrement

249, rue de Crimée - 75019 PARIS
Tél 33 (1) 40 36 03 03
Fax 33 (1) 40 38 96 43

* Le titulaire s'engage à informer l'APP de toute cession ou attribution totale ou partielle, de ses droits de propriété intellectuelle.
Seules les inscriptions de type S et C permettent un éventuel accès au programme source.

Agence pour la Protection des Programmes
249, rue de Crimée - 75019 PARIS
Tél 33 (1) 40 36 03 03
Fax 33 (1) 40 38 96 43

Inter Deposit Digital Number

Certificat délivré par

Agence pour la Protection des Programmes

249, rue de Crimée - 75019 PARIS. Tél 33 (1) 40 35 03 03. Fax 33 (1) 40 38 96 43

IDDN.FR.001.370024.000.S.P.2010.000.31230

(1) (2) (3) (4) (5) (6) (7) (8) (9) (10)

Pour l'oeuvre : Pyrodyn

Identité du(des) titulaire(s) des droits
INSERM-Institut National de la Santé
et de la Rech. Médicale
101 rue de Tolbiac
75654 PARIS Cedex 13
Siret : 180036048

Adhérent sous le numéro : 07.75.11.11.11
Support utilisé : 1 CD-Rom en dou

Le titulaire *

Fait à Paris, le 17/09/2010



Logibox conservé par l'adhérent : 63800
Logibox conservé par l'APP : 63801

(1) Inter Deposit Digital Number
(2) Nationalité de l'oeuvre

(3) Numéro de l'organisme de rattachement
(4) Numéro d'ordre de l'enregistrement

(5) Numéro de version
(6) Type d'enregistrement

(7) Type de l'oeuvre
(8) Année d'enregistrement

* Le titulaire s'engage à informer l'APP de toute cession ou aliénation totale ou partielle, de ses droits de propriété intellectuelle.
Seules les inscriptions de type S et C permettent un éventuel accès au programme source.

Le Président
Agence pour la Protection des Programmes
249 rue de Crimée - 75019 PARIS
Tél 33 (1) 40 35 03 03 - Fax 33 (1) 40 38 96 43
103, Classe de présidents - APE 913 E
Siret - 3853584400027 - APE 913 E

Inter Deposit Digital Number

Certificat délivré par

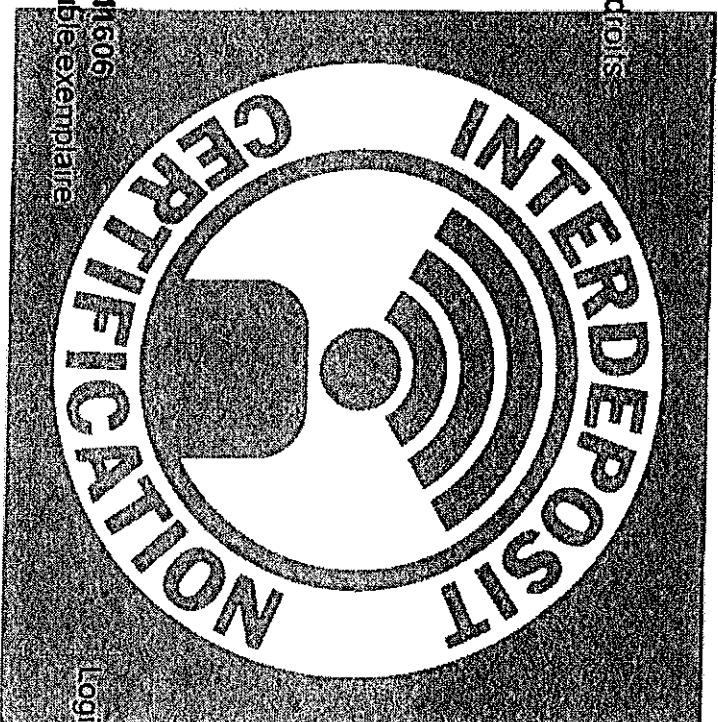
Agence pour la Protection des Programmes

249, rue de Clichy - 75019 PARIS, Tel 33 (1) 40 35 03 03, Fax 33 (1) 40 38 96 43

IDDN.FR.001.360005.000.S.P.2010.000.31230

Pour l'oeuvre : **Pyroclass**

Identité du(des) titulaire(s) des droits
INSERM-Institut National de la Santé
et de la Rech. Médicale
101 rue de Tolbiac
75654 PARIS Cedex 13
Siret : 180036048



Adhérent sous le numéro : **07.75.41606**
Support utilisé : 1 CD-Rom en double exemplaire

Le titulaire *

Fait à Paris, le 06/09/2010

Logibox conservé par l'adhérent : 66590
Logibox conservé par l'APP : 66591

Le Président

(1) Inter Deposit Digital Number

(3) Numéro de l'organisme d'enregistrement

(4) Numéro d'ordre de l'organisme

(5) Numéro de version

(6) Type d'enregistrement

(7) Type de l'oeuvre

(8) Année d'enregistrement

Agence pour la Protection des Programmes
(9) Zone réservée (Clé d'impression) PARIS
249, rue de Clichy - 75019 PARIS
Classe de produit

Siret : 38538584400027 - APE 913 E

* Le titulaire s'engage à informer l'APP de toute cession ou aliénation, totale ou partielle, de ses droits de propriété intellectuelle.
Seules les inscriptions de type S et C permettent un éventuel accès au programme source.

Inter Deposit Digital Number

Certificat délivré par

Agence pour la Protection des Programmes

249, rue de Crimée - 75019 PARIS. Tel 33 (1) 40 35 03 03, Fax 33 (1) 40 38 96 43

IDDN.FR.001.120009.001.S.C.2010.000.31230

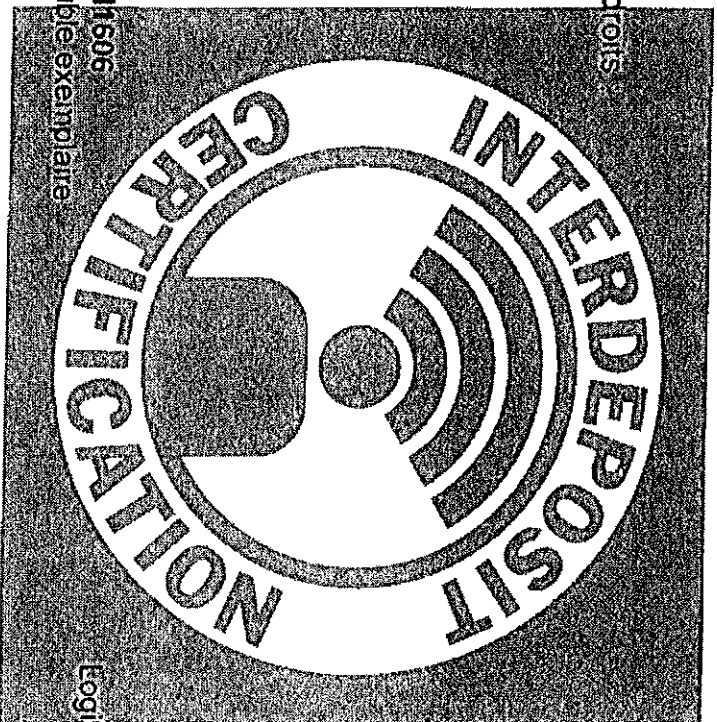
Pour l'oeuvre : PYROTREE V2

Identité du(des) titulaire(s) des droits
INSERM-Institut National de la Santé
et de la Rech. Médicale
101 rue de Tolbiac
75654 PARIS Cedex 13
Siret : 180038048

Adhérent sous le numéro : 07.75.11596
Support utilisé : 1 CD-Rom en double exemplaire

Le titulaire *

Fait à Paris, le 06/09/2010



Logibox conservé par l'adhérent : 63769
Logibox conservé par l'APP : 63770

(1) Inter Deposit Digital Number

(3) Numéro de l'organisme demeuré

(5) Numéro de version

(7) Type de l'oeuvre

(2) Nationalité de l'oeuvre

(4) Numéro d'ordre de l'enregistrement

(6) Type d'enregistrement

(8) Année d'enregistrement

* Le titulaire s'engage à informer l'APP de toute cession ou attribution totale ou partielle, de ses droits de propriété intellectuelle

Seules les inscriptions de type S et C permettent un éventuel accès au programme source.

Le Président

Agence pour la Protection des Programmes

2 Place des Saussaies (5^e arrondissement) 75019 PARIS

Tel 33 (0)1 40 35 03 03 - Fax 33 (0)1 40 38 96 43

Site Internet : www.appprotection.com

APPProtection des Programmes - 38538584400027 - APPE 913 E

Inter Deposit Digital Number

Certificat délivré par

Agence pour la Protection des Programmes

249, rue de Cimée - 75019 PARIS. Tél 33 (1) 40 35 03 03. Fax 33 (1) 40 38 96 43

IDDN.FR.001.120008.001.S.C.2010.000.31230

Pour l'oeuvre : PYROMUTE V2

Identité du(des) titulaire(s) des droits
INSERM-Institut National de la Santé
et de la Rech. Médicale
101 rue de Tolbiac
75654 PARIS Cedex 13
Siret : 180036048



Adhérent sous le numéro : 07.75.11606
Support utilisé : 1 CD-Rom en double exemplaire

Logibox conservé par l'adhérent : 63777
Logibox conservé par l'A.P.P. : 63778

Le titulaire *

Fait à Paris, le 10/09/2010

(1) Inter Deposit Digital Number
(2) Nationalité de l'oeuvre

(3) Numéro de l'organisme d'enregistrement
(4) Numéro d'ordre de l'ordre de l'enregistrement

(5) Numéro de version
(6) Type d'enregistrement

(7) Type de l'oeuvre
(8) Année d'enregistrement

* Le titulaire s'engage à informer l'A.P.P. de toute cession ou aliénation, totale ou partielle, de ses droits de propriété intellectuelle.
Seules les inscriptions de type S et C permettent un éventuel accès au programme source.

Le Président

Agence pour la Protection des Programmes

249, rue de Cimée - 75019 PARIS

Tel: 33 (0)1 40 35 03 03 - Fax 33 (0)1 40 38 96 43

Site: 38538584400027 - A.P.P. 913 E

ARTICLE 1

Pallier C*, **Rodriguez C***, Brillet R, Nordmann P,
Hézode C, Pawlotsky JM

Complex dynamics of hepatitis B virus resistance to adefovir

Hepatology 2009 ; 49 :50-59

*participation égale

Complex Dynamics of Hepatitis B Virus Resistance to Adefovir

Coralie Pallier,^{1,2,3*} Christophe Rodriguez,^{1,2*} Rozenn Brillet,^{1,2} Patrice Nordmann,³ Christophe Hézode,^{1,2,4} and Jean-Michel Pawlotsky^{1,2}

In patients with hepatitis B e antigen-negative chronic hepatitis B, adefovir dipivoxil administration selects variants bearing reverse transcriptase rtN236T and/or rtA181V/T substitutions in 29% of cases after 5 years. The aim of this study was to characterize the dynamics of adefovir-resistant variant populations during adefovir monotherapy in order to better understand the molecular mechanisms underlying hepatitis B virus resistance to this class of nucleotide analogues. Patients included in a 240-week clinical trial of adefovir monotherapy who developed adefovir resistance-associated substitutions were studied. The dynamics of hepatitis B virus populations were analyzed over time, after generating nearly 4,000 full-length reverse transcriptase sequences, and compared with the replication kinetics of the virus during therapy. Whatever the viral kinetics pattern, adefovir resistance was characterized by exclusive detection of a dominant wild-type, adefovir-sensitive variant population at baseline and late and gradual selection by adefovir of several coexisting resistant viral populations, defined by the presence of amino acid substitutions at position rt236, position rt181, or both. The gain in fitness of one or the other of these resistant populations during adefovir administration was never associated with the selection of additional amino acid substitutions in the reverse transcriptase. *Conclusion:* Our results suggest that adefovir administration selects poorly fit preexisting or emerging viral populations with low-level adefovir resistance, which subsequently compete to fill the replication space. Viral kinetics depends on the initial virological response to adefovir. Lamivudine add-on restores some antiviral efficacy, but adefovir-resistant variants remain predominant. Whether these adefovir resistance-associated substitutions may confer cross-resistance to tenofovir *in vivo* will need to be determined. (HEPATOLOGY 2009;49:50-59.)

Abbreviations: ALT, alanine aminotransferase; HBeAg, hepatitis B e antigen; HBV, hepatitis B virus; PCR, polymerase chain reaction.

From the ¹French National Reference Center for Viral Hepatitis B, C, and delta, Department of Virology, Hôpital Henri Mondor, Université Paris 12, Créteil, France; ²INSERM U841, Créteil, France; the ³Department of Bacteriology and Virology, Hôpital de Bicêtre, Université Paris XI, Le Kremlin-Bicêtre, France; and the ⁴Department of Hepatology and Gastroenterology, Hôpital Henri Mondor, Université Paris 12, Créteil, France.

Received April 6, 2008; accepted September 9, 2008.

This work was part of the VIRGIL European Network of Excellence on Antiviral Drug Resistance and was supported by grant LSHM-CT-2004-503359 from the Priority 1 "Life Sciences, Genomics, and Biotechnology for Health" program in the 6th Framework Programme of the European Union. It was also supported by an unrestricted grant from Gilead Sciences.

*These authors contributed equally to the study.

Address reprint requests to: Professor Jean-Michel Pawlotsky, M.D., Ph.D., Department of Virology, Hôpital Henri Mondor, 51 Avenue du Maréchal de Lattre de Tassigny, 94010 Créteil, France. E-mail: jean-michel.pawlotsky@hmn.apbp.fr; fax: (33)-1-4981-4831.

Copyright © 2008 by the American Association for the Study of Liver Diseases.

Published online in Wiley InterScience (www.interscience.wiley.com).

DOI 10.1002/hep.22634

Potential conflict of interest: Dr. Pawlotsky advises and received grants from Gilead. He also advises Bristol-Myers Squibb.

Treatment of chronic hepatitis B virus (HBV) infection is aimed at driving viral replication to the lowest possible level in order to halt the progression of chronic hepatitis B and to prevent complications such as cirrhosis, decompensation of cirrhosis, and hepatocellular carcinoma. Two categories of drugs can be used to treat chronic hepatitis B, namely pegylated interferon- α and nucleoside/nucleotide analogues that directly inhibit the reverse transcriptase function of HBV DNA polymerase. Five HBV inhibitors (lamivudine, adefovir dipivoxil, entecavir, telbivudine, and tenofovir) are currently approved. Lamivudine and adefovir dipivoxil are most widely used worldwide.

Chronic administration of nucleoside/nucleotide analogues in monotherapy frequently leads to the selection of HBV variants with amino acid substitutions in the reverse transcriptase domain of HBV DNA polymerase, conferring reduced susceptibility to the inhibitory action of these drugs.¹⁻⁴ HBV has a quasispecies distribution in each infected individual, meaning that the virus circulates as a complex mixture of genetically distinct but closely

related variants that are in equilibrium at a given time in a given replicative environment.⁵ The quasispecies distribution of HBV implies that any newly generated mutation conferring a selective advantage in a given replicative environment will allow the corresponding viral population to overtake the other variants, following a classical Darwinian evolutionary process. As a result, HBV resistance is generally acquired gradually through the selection of preexisting resistant variants and gradual accumulation of new amino acid substitutions. The latter may confer stepwise increases in the level of drug resistance, but they mostly improve the replication capacity of resistant variants in the presence of the relevant drug. HBV resistance to antiviral therapy is a major issue in clinical practice because it can lead to HBV treatment failure and progression of liver disease.

Long-term administration of adefovir dipivoxil has been shown to select variants bearing reverse transcriptase rtN236T and/or rtA181V/T substitutions, both of which confer low-level resistance to adefovir *in vitro*. In patients with hepatitis B e antigen (HBeAg)-negative chronic hepatitis B, these variants are selected in 29% of cases (cumulative incidence) after 5 years of monotherapy, essentially in patients who do not achieve HBV DNA levels below 1,000 IU/mL.⁶ Their selection is associated with different patterns of HBV DNA kinetics in different patients: (1) typical escape, with a return to baseline replication levels in patients who had initially responded to the drug by a more than 3 log reduction of HBV DNA level; (2) a gradual re-increase in viral replication preceding the emergence of the mutant virus by several months, generally in patients with a suboptimal response to adefovir (less than 3 log reduction of HBV DNA level); and (3) no change in adefovir antiviral efficacy in patients who initially responded well to the drug.⁷⁻¹⁰ These different patterns point to differences in the replicative fitness of selected resistant variants *in vivo*.

To understand the mechanisms underlying viral resistance to a specific inhibitor, it is first necessary to determine the quasispecies dynamics during the emergence and amplification of resistance, together with the impact of the relevant amino acid substitutions on viral replication *in vivo*.¹¹⁻¹³ We have recently shown that, *in vivo*, the fluctuating environment in which lamivudine-resistant variants replicate during lamivudine administration to patients potentially favors one variant population over another, regardless of the intrinsic replication capacities of the viral populations *in vitro*.¹¹ The aim of the present study was to characterize the dynamics of adefovir-resistant HBV variant populations during adefovir monotherapy to better understand the molecular mechanisms underlying HBV resistance to adefovir.

Patients and Methods

Patients

This was a substudy of a recently published 240-week clinical trial of adefovir monotherapy in 185 nucleoside/nucleotide-naïve HBeAg-negative patients, with a double-blinded phase of 96 weeks followed by an open-label, long-term safety and efficacy study of 144 weeks.⁶ Serial HBV DNA assays were systematically performed during therapy, as well as direct sequence analysis of the entire HBV reverse transcriptase gene at baseline and, in patients with detectable serum HBV DNA, at weeks 48, 96, 144, 192, and 240. Twenty-nine patients developed adefovir resistance-associated substitutions during the 240-week study. Their viral dynamics were characterized by means of frequent HBV DNA assays, classifying the patients into the following three groups: (1) responders (HBV DNA level reduction of more than 3 log IU/mL) who experienced secondary treatment failure (reincrease in the HBV DNA level of more than 1 log IU/mL above the nadir); (2) suboptimal responders (HBV DNA level reduction of less than 3 log IU/mL) who developed amino acid substitutions in a context of a slow, gradual re-increase in viral replication; and (3) responders who developed the amino acid substitutions without any virological breakthroughs. We selected the first seven patients displaying these three dynamic patterns in order to study HBV population dynamics during adefovir exposure.

The seven patients were five men and two women aged 26 to 59 years at the start of the study. Their baseline characteristics are shown in Table 1. They were all Caucasian, and none was coinfecting with hepatitis delta virus, hepatitis C virus, or human immunodeficiency virus. All were nucleoside/nucleotide treatment-naïve at inclusion. They were treated with 10 mg/day adefovir for the full study period. In four cases (patients A, B, E, and G), lamivudine was added during adefovir therapy because of virological failure. Quasispecies analysis was performed with serum samples taken at baseline and frequently during therapy (15 to 24 samples). At every time point and in each patient, HBV DNA was extracted and the full-length HBV reverse transcriptase domain was polymerase chain reaction (PCR)-amplified. The PCR products were directly sequenced and cloned. For quasispecies analysis, 25 to 41 clones per time point were sequenced. The dynamics of HBV populations over time were studied.

Methods

Alanine Aminotransferase and HBV DNA Assays.

Alanine aminotransferase (ALT) levels were determined at baseline and during treatment using a Roche Modular Chemistry Analyzer (Roche Diagnostics, Basel, Switzerland). HBV DNA was measured at each time point by

Table 1. Clinical and Virological Characteristics of the Seven Patients Included in the Study

Patient	Sex/Age	Race	Route of Transmission	HBV Genotype	HBeAg/Anti-HBe Ab	HBV DNA (log IU/mL)	Baseline ALT	Ishak Score*				
								a	b	c	d	e
A	M/46	Caucasian	Unknown	D	-/+	7.8	170	4	0	3	2	2
B	M/59	Caucasian	Unknown	D	-/+	6.8	166	2	0	1	2	3
C	F/51	Caucasian	Unknown	D	-/+	7.6	52	4	0	4	2	4
D	M/55	Caucasian	Childhood acquisition	D	-/+	6.6	239	4	0	4	3	5
E	M/43	Caucasian	Childhood acquisition	D	-/+	6.1	43	2	0	2	2	3
F	F/56	Caucasian	Blood transfusion	D	-/+	6.2	88	3	0	3	3	3
G	M/26	Caucasian	Unknown	D	-/+	5.5	99	4	0	2	4	6

Abbreviations: M, male; F, female; Ag, antigen; Ab, antibodies.

*Ishak score: a, periportal injury; b, confluent necrosis; c, parenchymal injury; d, portal inflammation; e, fibrosis.

means of a “classical” reverse transcriptase PCR assay (Cobas Amplicor HBV Monitor v2.0, Roche Molecular Systems, Pleasanton, CA) or a real-time PCR assay (Cobas TaqMan 48 HBV test, Roche Molecular Systems). The results were expressed in IU/mL.

PCR Amplification of the HBV Reverse Transcriptase Region. DNA was extracted from 200–500 mL of serum with the QIAamp DNA blood kit (Qiagen GmbH, Hilden, Germany) according to the manufacturer’s instructions. A 630-bp fragment encompassing domains A to E of the HBV reverse transcriptase, between amino acid positions 66 and 275, was amplified via PCR using primer pair pol1 (5′-CCCTGCTCGTGTTACAG-GCGG-3′, nucleotide position 186–206) and pol2 (5′-GTTGCGTCAGCAAACACTTGGCA-3′, nucleotide position 1196–1174).¹⁴ Forty cycles of amplification were performed, with a denaturation step at 94°C for 30 seconds, annealing at 48°C for 30 seconds and extension at 72°C for 1 minute. When the HBV DNA level was below 1,000 IU/mL (3.0 log IU/mL), a nested PCR technique was used, with internal primers pol3 (5′-GACTCGTG-GTGGACTTCTCTCA-3′, nucleotide position 251–272) and pol4 (5′-GGCATTAAAGCAGGATAACCACATTG-3′, nucleotide position 1058–1033). PCR amplification consisted of 30 cycles at 94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 1 minute. PCR products were analyzed via electrophoresis through 2% agarose gel and ethidium bromide staining. They were then purified with Montage PCR centrifugal filter devices (Millipore Corporation, Bedford, MA) according to the manufacturer’s protocol.

Direct Sequence Analysis of PCR Products and Determination of the HBV Genotype. Purified PCR products were directly sequenced with the primers used for PCR amplification. The HBV genotype was determined by phylogenetic analysis of the 630-bp fragment amplified at baseline, with respect to reference sequences of the principal HBV genotypes retrieved from GenBank, using the Phylogeny Inference Package (PHYLIP) program version 3.65.

Quasispecies Analysis of HBV RT. Purified PCR products were cloned into the pCR4-TOPO vector (TOPO TA cloning kit, Invitrogen Corporation, Carlsbad, CA) according to the manufacturer’s instructions. Recombinant plasmid DNA was transformed into One Shot TOP10 chemically competent *Escherichia coli*, as recommended by the manufacturer, and transformants were grown on ampicillin plates. Cloned DNA was reamplified via PCR and purified for sequencing. Twenty-five to 41 clones per sample were sequenced, and their nucleotide and amino acid sequences were aligned with CLUSTAL X, version 1.83. Because all but one sample had an HBV DNA level above 1,000 IU/mL, resampling resulting in expansion of multiple copies of the same viral species was unlikely to occur.

Nucleotide Sequence Accession Numbers. The sequence data from this study have been deposited in GenBank under accession numbers EU026558 to EU028300.

Results

Twenty-nine patients developed adefovir resistance-associated substitutions during the 240-week study. Based on their viral dynamics, the 29 patients were classified into response pattern 1 (initial virological response with secondary treatment failure; n = 9 [31%]), response pattern 2 (suboptimal virological response followed by a gradual re-increase in viral replication; n = 4 [14%]), response pattern 3 (sustained antiviral response throughout follow-up; n = 15 [52%]), and mixed response pattern (n = 1 [3%]). Given the number of samples studied per patient and the number of clones generated per sample, not all patients could be extensively studied. We selected the first seven patients, including three patients with response pattern 1, two patients with response pattern 2, one patient with response pattern 3, and one patient with a mixed response pattern.

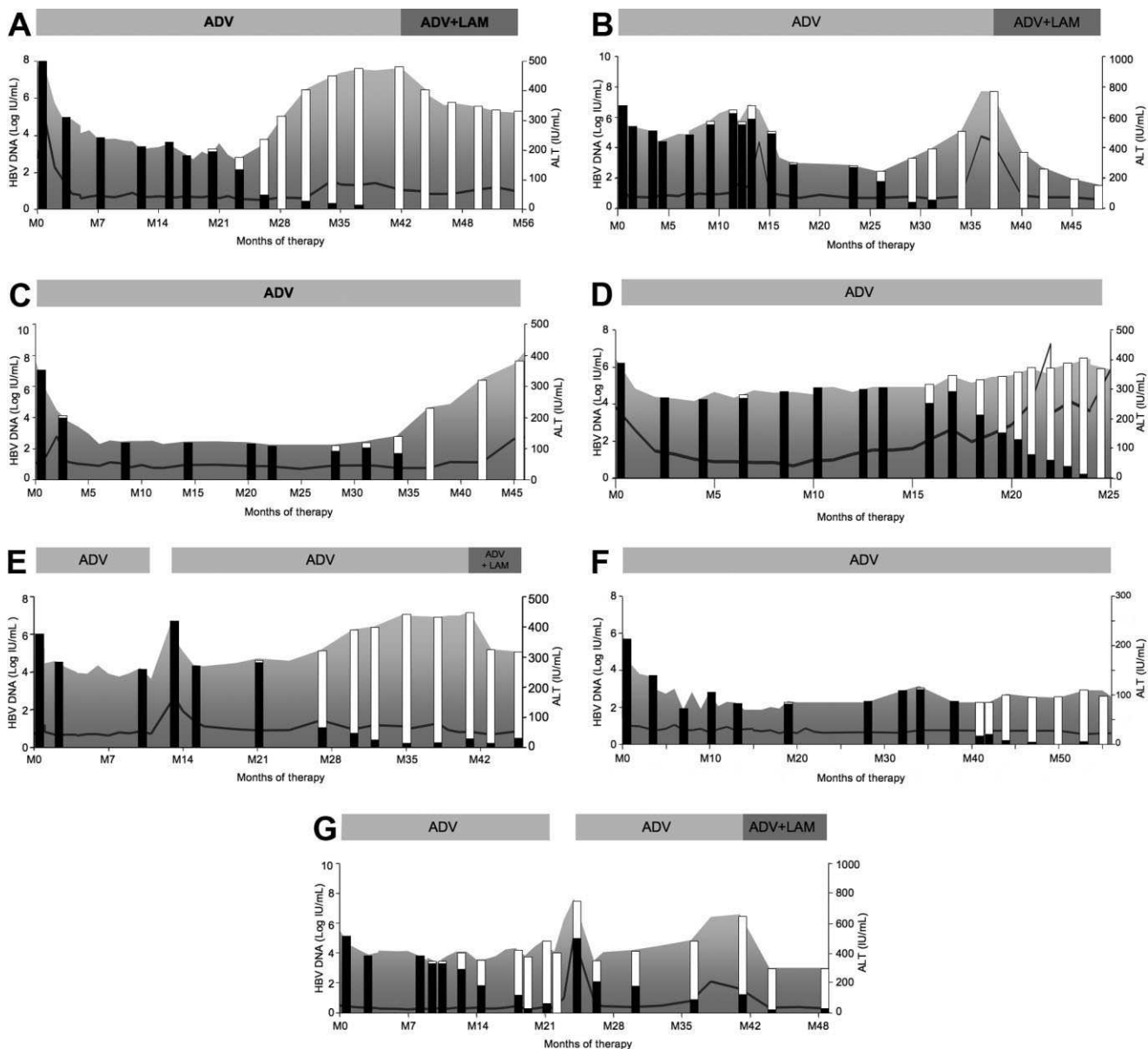
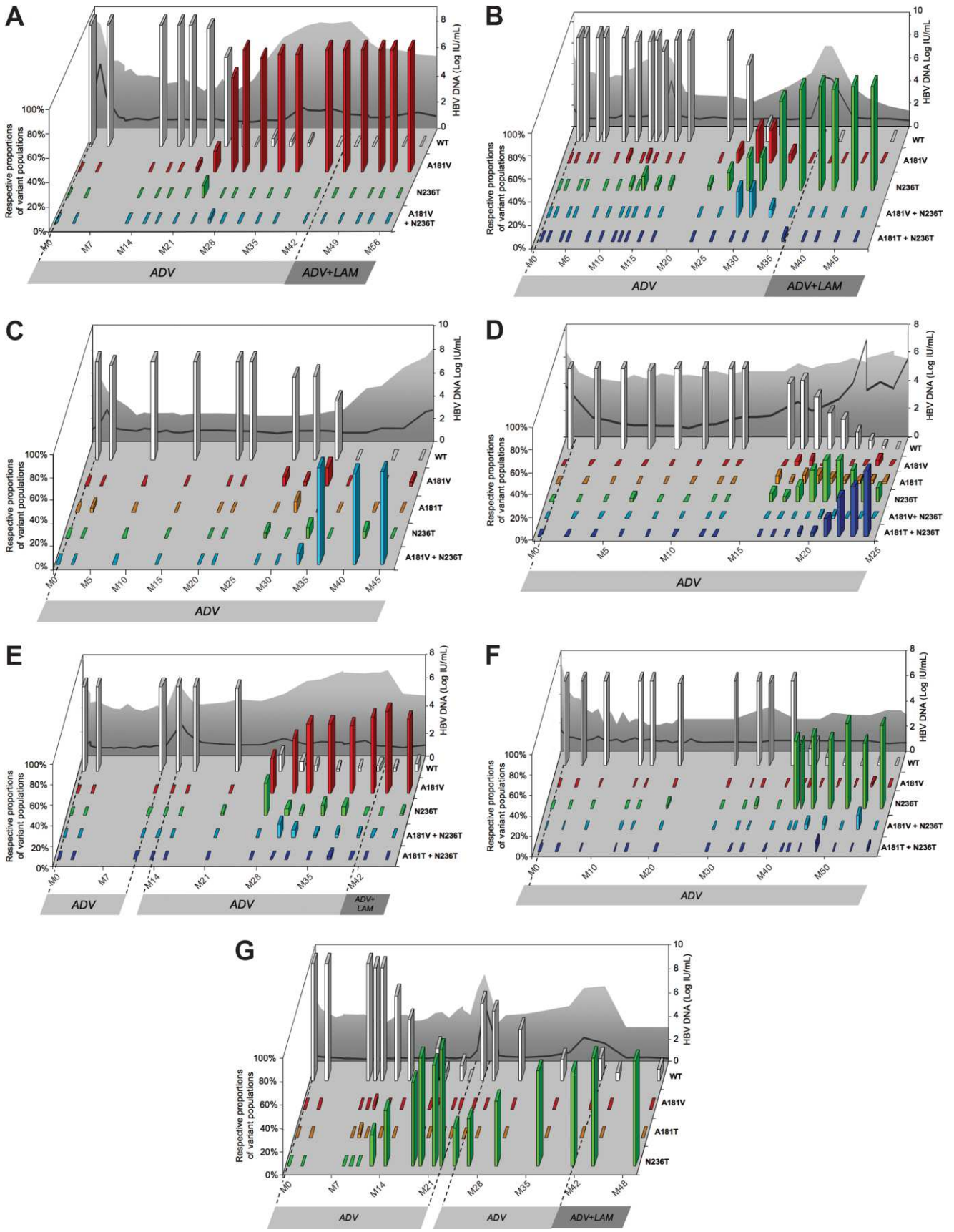


Fig. 1. Dynamics of serum HBV DNA levels (gray area), serum aminotransferase activities (gray line), adefovir-sensitive (black bars), and adefovir-resistant (white bars) HBV viral populations during adefovir (ADV) and eventually adefovir plus lamivudine (ADV + LAM) therapy in the patients with response pattern 1 [initial virological response with secondary treatment failure: (A) patient A, (B) patient B, and (C) patient C], response pattern 2 [suboptimal virological response followed by a gradual re-increase in viral replication: (D) patient D and (E) patient E], response pattern 3 [sustained antiviral response throughout follow-up: (F) patient F], and a mixed response pattern [(G) patient G].

A total of 3843 nearly full-length HBV reverse transcriptase quaspecies sequences were generated and analyzed. At baseline, the HBV reverse transcriptase always had a typical quaspecies distribution, characterized by the coexistence of distinct but closely related viral populations, including one to three major variants, each representing more than 10% of the quaspecies variants, and a variable number of minor variants bearing amino acid substitutions at sporadic positions relative to the most frequent variant. Figure 1 shows the HBV DNA dynamics over time in each patient and the proportions of resis-

tant versus wild-type virus at baseline and at each time point during therapy. In all the patients, only a wild-type, adefovir-sensitive virus population (that is, viruses without amino acid substitutions at positions rt181 and rt236) was detected prior to therapy (Fig. 1). Figure 2 shows the dynamics of the different adefovir-resistant viral populations over time.

Dynamics of Resistance Associated with Response Pattern 1 (Initial Virological Response with Secondary Treatment Failure). Three patients (patients A, B, and C) initially responded to adefovir, with a significant



fall in the HBV DNA level (-4.6 , -3.8 , and -5.1 log IU/mL after 12, 17, and 8 months of adefovir administration, respectively; Figs. 1A-C). The antiviral response only occurred after 14 months in patient B, after an ALT flare; this patient had initially responded suboptimally. The virological breakthroughs occurred at months 28, 31 (17 months after the antiviral response), and 37, respectively, and were always followed by a biochemical breakthrough (ALT elevation) when the HBV DNA level returned to baseline. Lamivudine add-on in patients A and B partially restored an antiviral response (Figs. 1A,B).

No viral populations bearing amino acid substitutions at position rt181 or rt236 were detected before therapy. The wild-type viral population was always gradually replaced by the resistant viruses during therapy. The size of the resistant virus population grew gradually for several months before the virological breakthrough and became predominant at the time of the biochemical breakthrough (Figs. 1A-C). The adefovir-resistant viral populations remained predominant when lamivudine was added to adefovir in patients A and B (Figs. 1A,B). Five resistant viral populations were found during adefovir administration in the three patients with viral dynamic pattern 1 (Figs. 2A-C), including viruses with an rtA181V substitution alone (patients A, B, and C), viruses with an rtA181T substitution alone (patient C), viruses with an rtN236T substitution alone (patients A, B, and C), viruses harboring both rtA181V and rtN236T substitutions (patients A, B, and C), and viruses harboring both rtA181T and rtN236T substitutions (patient B).

In the three patients, several different viral populations were present at several time points. However, the dynamics of HBV viral populations varied from one patient to the next. In patient A (Fig. 2A), the rtA181V variant population predominated over the other adefovir-resistant populations throughout therapy. Despite the partial restoration of an antiviral response when lamivudine was added, the rtA181V population remained predominant

during adefovir-lamivudine combination therapy (Fig. 2A). In patient B (Fig. 2B), the wild-type sensitive virus was initially replaced by a mixture, in similar proportions, of three variants harboring the rtA181V substitution, the rtN236T substitution, and both substitutions together, respectively. However, after a few months the rtN236T-only variants took over and remained predominant until the end of adefovir monotherapy, even when lamivudine was added, despite partial restoration of the antiviral response (Fig. 2B). In patient C (Fig. 2C), quasispecies analysis showed the simultaneous emergence of four resistant variants, but variants bearing both rtA181V and rtN236T substitutions rapidly became predominant. No reverse transcriptase positions other than rt181 or rt236 showed specific amino acid changes during therapy that might have led to improved fitness and thus explained how one specific resistant population overtook the others during adefovir exposure.

Dynamics of Resistance Associated with Response Pattern 2 (Suboptimal Virological Response Followed by a Gradual Re-increase in Viral Replication). Two patients (D and E) initially responded suboptimally to adefovir. In patient D, the HBV DNA level fell by -2.5 log IU/mL after 4 months of treatment and then gradually increased. A biochemical breakthrough occurred at month 22 in this patient, when the HBV DNA level had returned to baseline (Fig. 1D). Patient E also responded suboptimally to adefovir (-1.9 log IU/mL). Adefovir withdrawal at month 10 was followed by HBV DNA and ALT flares. Treatment was restarted at month 13 and the patient again responded suboptimally (-2.4 log IU/mL in 2 months). The HBV DNA level then started to increase gradually, returning to baseline (the value before the second course of treatment) at month 30 (Fig. 1E). No biochemical breakthrough occurred in this patient. Lamivudine was added at month 40 and resulted in a 2-log fall in HBV DNA (Fig. 1E).

In these two patients the wild-type sensitive virus was gradually replaced by resistant viruses, which became predominant when the HBV DNA level had returned to baseline (Figs. 1D,E). In patient D, five resistant populations (rtA181V alone, rtA181T alone, rtN236T alone, rtA181V plus rtN236T, and rtA181T plus rtN236T) were detected during follow-up. Their dynamics were characterized by three successive waves, with rtA181V/T initially dominating, then being replaced by a dominant rtN236T population and, finally, by a dominant population bearing both rtA181T and rtN236T substitutions (Fig. 2D). The biochemical breakthrough coincided with the shift from a dominant sensitive viral population to a resistant one at the time of the second wave (dominant rtN236T population) (Fig. 2D).

Fig. 2. Dynamics of HBV variant populations during adefovir (ADV) and eventually adefovir plus lamivudine (ADV + LAM) therapy in the patients with response pattern 1 [initial virological response with secondary treatment failure: (A) patient A, (B) patient B, and (C) patient C], response pattern 2 [suboptimal virological response followed by a gradual re-increase in viral replication: (D) patient D and (E) patient E], response pattern 3 [sustained antiviral response throughout follow-up: (F) patient F], and a mixed response pattern [(G) patient G]. White bars represent the proportion of wild-type, adefovir-sensitive HBV variants at the different time points. The proportions of HBV variants bearing the rtA181V substitution, the rtA181T substitution, the rtN236T substitution, both rtA181V and rtN236T substitutions, and both rtA181T and rtN236T substitutions are represented by red, orange, green, light blue, and dark blue, respectively. HBV DNA levels are represented by the gray area, and ALT levels are represented by the gray line.

In patient E, four resistant populations were detected during follow-up (rtA181V alone, rtN236T alone, rtA181V plus rtN236T, and rtA181T plus rtN236T), but the viral population bearing the rtA181V substitution was dominant throughout follow-up (Fig. 2E). The same pattern (predominance of rtA181V variants) was observed when lamivudine was added (Fig. 2E). No positions other than rt181 or rt236 showed specific amino acid changes during therapy in patients D and E.

Dynamics of Resistance Associated with Response Pattern 3 (Sustained Antiviral Response Throughout Follow-up). Patient F was treated with adefovir for 57 months. He initially responded with a -4.3 log IU/mL decrease in HBV DNA after 7 months, and this level of response persisted throughout follow-up, despite gradual replacement of the wild-type, sensitive virus by the resistant virus population after month 40 (Fig. 1F). Despite the predominance of resistant viruses, no virological or biochemical breakthrough was observed between month 40 and month 47, when follow-up ended for this patient (Fig. 1F). As in the other patients, regardless of the viral dynamics on therapy, several resistant populations were detected, but the population with only the rtN236T substitution predominated from the onset of viral resistance until the end of follow-up. No positions other than rt181 or rt236 showed specific amino acid changes during therapy.

Mixed Response. Patient G had a mixed virological response (Fig. 1G). This patient initially responded suboptimally to adefovir (-2.0 log IU/mL in 9 months), and the HBV DNA level remained stable thereafter, despite gradual replacement of the wild-type, sensitive viral population by a resistant population that predominated at month 22 when treatment was stopped (Fig. 1G). Patient G had an HBV DNA flare 1 month after treatment withdrawal and, although the resistant viral population was still detectable, most of the variants present at the time of the flare were sensitive. The patient was then retreated at month 24 and responded (-4.0 log IU/mL in 2 months). However, the resistant population took over rapidly, and virological and biochemical breakthroughs occurred a few months later. Lamivudine was added, and this restored an effective antiviral response, even though an adefovir-resistant population predominated (Fig. 2G). As shown in Fig. 2G, the rtN236T population was nearly the only resistant population present during the first course of therapy, at the time of the flare after treatment withdrawal, during the second course of therapy, at the time of the virological and biochemical breakthrough, and when lamivudine was added (Fig. 2G). No positions other than rt181 or rt236 showed specific amino acid changes during therapy.

Discussion

We generated nearly 4,000 full-length HBV reverse transcriptase sequences in order to unravel the dynamics of HBV variants during adefovir dipivoxil administration and thereby to understand how HBV resistance to this drug develops. The dynamics of HBV viral populations were compared with the replication kinetics of the virus during several years of adefovir administration.

Whatever the viral kinetics, adefovir resistance was always characterized by: (1) exclusive detection of a dominant wild-type, adefovir-sensitive variant population at baseline; (2) late (several years) and gradual selection by adefovir of coexisting resistant viral populations, defined by the presence of amino acid substitutions at position rt236, position rt181, or both; and (3) during adefovir administration, an apparent gain in fitness of one or other of these resistant populations, which became predominant after several months of resistance. This gain in fitness was never associated with the selection of additional amino acid substitutions in the reverse transcriptase. In contrast, substitutions that improve viral replication without substantially enhancing resistance have been observed during lamivudine resistance (substitution at position rt180 or rt173)¹⁵⁻¹⁸ and during entecavir resistance (substitution at position rt180 and at positions rt184, rt202, and/or rt250).¹⁹⁻²²

Our results are in keeping with, although they do not demonstrate, the hypothesis that the adefovir-resistant viral populations preexist in a subset of HBV-infected patients, who are those exposed to develop adefovir resistance. Indeed, amino acid substitutions occur randomly as a result of viral replication. They are not induced by the administration of an antiviral drug that can only select viral variants that bear them when they are present. Therefore, although it cannot be excluded that some substitutions occurred during the few months of therapy in patients with reduced HBV replication levels, the probability is much greater that these substitutions randomly occurred during the many years that preceded therapy during which replication levels were higher. HBV variants bearing rtN236T and/or rtA181V/T polymorphisms, however, remain undetectable at baseline, owing to their relatively poor *in vivo* fitness. Although these variants were not observed at baseline in our clonal analysis, the sporadic detection of one or both mutations early in therapy, prior to clear-cut outgrowth of resistant virus populations, was consistent with the hypothesis of their preexistence. In addition, the substitutions at positions rt236 and rt181 confer only low-level HBV resistance to adefovir (up to a three-fold increase in the 50% inhibitory concentration *in vitro*), whereas the rtM204V/I substitu-

tion confers a >1,000 fold increase in the lamivudine 50% inhibitory concentration.²³⁻²⁵ Thus, their low-level resistance and poor *in vivo* fitness together explain why adefovir-resistant variants emerge late and very slowly during adefovir administration.

The HBV replication kinetics provide further information on the relative *in vivo* fitness of adefovir-resistant variants in the presence of adefovir. In the patients with pattern 1 replication kinetics (virological response followed by virological breakthrough), the profound inhibition of wild-type, sensitive virus replication opens the replication space to resistant variants. In this situation, the gain in fitness can allow HBV DNA to return to the baseline level when the resistant variant becomes predominant. As during lamivudine resistance, selection of adefovir-resistant variants preceded virological breakthrough by several months in our study. The interval between the emergence of resistant variants and virological breakthrough was longer, however, than during lamivudine administration,¹¹ owing to the poorer fitness of adefovir-resistant variants *in vivo* relative to lamivudine-resistant variants. The biochemical breakthrough generally occurred when the resistant population accounted for 100% of circulating variants.

In viral kinetic pattern 3, no virological relapse was observed despite an initial virological response (as in pattern 1) and gradual replacement of the wild-type, sensitive population by adefovir-resistant variants. Two mutually compatible hypotheses can explain this observation: (1) the selected variants may have had such poor *in vivo* fitness that they could not be produced at levels similar to those of the wild-type, sensitive virus present before therapy, in which case a viral breakthrough will be observed only if the variants acquire additional mutations susceptible to improve their *in vivo* fitness; and (2) owing to the poor replication capacity of the variants, more time may have been required for HBV DNA to reach high levels, meaning that relapse could occur later in these patients. This observation could be reassuring in that detection of adefovir resistance-associated substitutions during therapy does not mean that a clinically relevant viral breakthrough will occur. However, such detection indicates that a breakthrough may subsequently occur, and further research should be conducted to determine whether treatment should be altered in order to prevent it.

Viral kinetic pattern 2 was more difficult to interpret. Indeed, these patients had a suboptimal response to adefovir, owing to the use of a low dose in clinical practice (10 mg/day).²⁶ A suboptimal response to adefovir has been reported to occur in approximately 20% of patients with HBeAg-negative chronic hepatitis B. In these patients, the only moderate decrease in wild-type, sensitive

virus replication did not open the replication space sufficiently for adefovir-resistant viruses to break through rapidly. This probably explains why the increase in HBV DNA levels was gradual and matched the progressive selection of resistant viruses, which became predominant roughly when HBV DNA levels had nearly returned to baseline.

Interestingly, we found no substitutions at positions other than rt236 and rt181 in the HBV reverse transcriptase that could explain a gain or loss of fitness by any of the adefovir-resistant populations in our patients. However, we cannot exclude the possibility that substitutions in other genomic regions than the reverse transcriptase, which were not analyzed here, could have played a role (for example, by modifying the intrinsic replication capacity of the corresponding variants or by inducing a shift in epitope immunodominance). It was indeed striking that, in two patients, the response to adefovir was directly associated with an ALT flare, with no visible sequence changes in the reverse transcriptase region. Indeed, patient B initially responded suboptimally to adefovir, but then responded fully to the same dose after an ALT flare occurred during therapy. Patient G also initially responded suboptimally, experiencing an ALT flare just after treatment was interrupted and subsequently responding fully to the same dose of adefovir. In both cases, the response to adefovir was followed by virological breakthrough as a result of adefovir-resistant variant selection. Pharmacodynamic differences in adefovir metabolism between patients may also influence viral fitness and, as a result, the kinetics of viral populations and HBV DNA during therapy by influencing the replicative environment. For instance, differences in the uptake of antivirals, including adefovir, by organic anion transporters have been reported.²⁷

In patients on adefovir monotherapy who developed resistance to adefovir, lamivudine add-on partially restored the antiviral response. However, the circulating variants remained adefovir-resistant. This was not surprising, because adefovir-resistant variants theoretically have similar lamivudine sensitivity to wild-type, sensitive variants but remain fitter than the latter in the presence of adefovir. If adefovir had been replaced by lamivudine, an adefovir-sensitive viral population would likely have replaced the adefovir-resistant virus, allowing lamivudine resistance to occur rapidly, as in a treatment-naïve population receiving lamivudine monotherapy. The fact that adefovir-resistant variants bearing substitutions at positions rt236 and/or rt181 would need to accumulate additional substitutions at position rt204 (and possibly positions rt180 and rt173) in order to become resistant to lamivudine and to become fit *in vivo* is reassuring with

respect to the time necessary for resistance to develop in patients receiving both drugs. Conversely, it has been demonstrated that adefovir add-on is a better strategy than a switch from lamivudine to adefovir in patients with lamivudine resistance.²⁸ More potent antiviral drugs with no cross-resistance with adefovir, such as telbivudine and entecavir, should be added in adefovir-resistant patients. It is important to note that substitutions at position rt181 have recently been reported to confer cross-resistance to lamivudine and adefovir *in vivo* and could also confer resistance to entecavir and telbivudine.^{29,30}

Taken together, our results suggest that adefovir administration selects poorly fit preexisting or emerging viral populations with low-level adefovir resistance, which subsequently compete to fill the replication space. In patients who initially respond to adefovir, resistance may be associated with a virological breakthrough followed by a biochemical breakthrough. In patients who initially respond suboptimally to adefovir, resistance emerges slowly and is associated with a gradual re-increase in viral replication, in keeping with the small difference in *in vivo* fitness between wild-type, sensitive, and adefovir-resistant viral variants. Lamivudine add-on restores some antiviral efficacy, but adefovir-resistant variants remain predominant.

Tenofovir disoproxil fumarate has recently been reported to reduce HBV replication more potently than adefovir in both HBeAg-positive and HBeAg-negative patients with chronic hepatitis B.^{31,32} Tenofovir has now received approval for use in the treatment of chronic hepatitis B in Europe and the United States. Conflicting results have been reported as to whether adefovir resistance-associated substitutions confer cross-resistance to tenofovir. However, several recent reports suggest that substitutions at positions rt236 may confer reduced HBV susceptibility to tenofovir *in vitro*.^{33,34} It remains to be seen whether adefovir resistance mutations also confer reduced susceptibility to tenofovir *in vivo*, as suggested by the recent observation that adefovir resistance substitutions persist when adefovir is switched to tenofovir in non- or suboptimal responders to adefovir.³⁵ If so, the dynamics of resistant viral populations will need to be carefully assessed. This will largely determine whether tenofovir eventually becomes the natural partner of nucleoside analogues in fixed-dose combinations for first- and second-line treatment of HBV infection.

Acknowledgment: We are grateful to Katyna Borroto-Esoda for her help.

References

1. Locarnini S, Hatzakis A, Heathcote J, Keeffe EB, Liang TJ, Mutimer D, et al. Management of antiviral resistance in patients with chronic hepatitis B. *Antivir Ther* 2004;9:679-693.
2. Lok AS, Zoulim F, Locarnini S, Bartholomeusz A, Ghany MG, Pawlotsky JM, et al. Antiviral drug-resistant HBV: standardization of nomenclature and assays and recommendations for management. *HEPATOLOGY* 2007;46:254-265.
3. Shaw T, Bartholomeusz A, Locarnini S. HBV drug resistance: mechanisms, detection and interpretation. *J Hepatol* 2006;44:593-606.
4. Pawlotsky JM, Dusheiko G, Hatzakis A, Lau D, Lau G, Liang TJ, et al. Virologic monitoring of hepatitis B virus therapy in clinical trials and practice: recommendations for a standardized approach. *Gastroenterology* 2008;134:405-415.
5. Pawlotsky JM. The concept of hepatitis B virus mutant escape. *J Clin Virol* 2005;34(Suppl 1):S125-S129.
6. Hadziyannis SJ, Tassopoulos NC, Heathcote EJ, Chang TT, Kitis G, Rizzetto M, et al. Long-term therapy with adefovir dipivoxil for HBeAg-negative chronic hepatitis B for up to 5 years. *Gastroenterology* 2006;131:1743-1751.
7. Yang H, Westland CE, Delaney WE 4th, Heathcote EJ, Ho V, Fry J, et al. Resistance surveillance in chronic hepatitis B patients treated with adefovir dipivoxil for up to 60 weeks. *Hepatology* 2002;36:464-473.
8. Angus P, Vaughan R, Xiong S, Yang H, Delaney W, Gibbs C, et al. Resistance to adefovir dipivoxil therapy associated with the selection of a novel mutation in the HBV polymerase. *Gastroenterology* 2003;125:292-297.
9. Westland CE, Yang H, Delaney WE 4th, Gibbs CS, Miller MD, Wulfsohn M, et al. Week 48 resistance surveillance in two phase 3 clinical studies of adefovir dipivoxil for chronic hepatitis B. *HEPATOLOGY* 2003;38:96-103.
10. Villeneuve JP, Durantel D, Durantel S, Westland C, Xiong S, Brosgart CL, et al. Selection of a hepatitis B virus strain resistant to adefovir in a liver transplantation patient. *J Hepatol* 2003;39:1085-1089.
11. Pallier C, Castera L, Soulier A, Hezode C, Nordmann P, Dhumeaux D, et al. Dynamics of hepatitis B virus resistance to lamivudine. *J Virol* 2006;80:643-653.
12. Villet S, Ollivet A, Pichoud C, Barraud L, Villeneuve JP, Trepo C, et al. Stepwise process for the development of entecavir resistance in a chronic hepatitis B virus infected patient. *J Hepatol* 2007;46:531-538.
13. Villet S, Pichoud C, Villeneuve JP, Trepo C, Zoulim F. Selection of a multiple drug-resistant hepatitis B virus strain in a liver-transplanted patient. *Gastroenterology* 2006;131:1253-1261.
14. Pichoud C, Seigner B, Wang Z, Trepo C, Zoulim F. Transient selection of a hepatitis B virus polymerase gene mutant associated with a decreased replication capacity and famciclovir resistance. *HEPATOLOGY* 1999;29:230-237.
15. Allen MI, Deslauriers M, Andrews CW, Tipples GA, Walters KA, Tyrrell DL, et al. Identification and characterization of mutations in hepatitis B virus resistant to lamivudine. *HEPATOLOGY* 1998;27:1670-1677.
16. Delaney WE 4th, Yang H, Westland CE, Das K, Arnold E, Gibbs CS, et al. The hepatitis B virus polymerase mutation rtV173L is selected during lamivudine therapy and enhances viral replication *in vitro*. *J Virol* 2003;77:11833-11841.
17. Fu L, Cheng YC. Role of additional mutations outside the YMDD motif of hepatitis B virus polymerase in L(-)SddC (3TC) resistance. *Biochem Pharmacol* 1998;55:1567-1572.
18. Seigner B, Pichoud C, Martin P, Furman P, Trepo C, Zoulim F. Inhibitory activity of dioxolane purine analogs on wild-type and lamivudine-resistant mutants of hepadnaviruses. *HEPATOLOGY* 2002;36:710-722.
19. Colonno RJ, Rose R, Baldick CJ, Levine S, Pokornowski K, Yu CF, et al. Entecavir resistance is rare in nucleoside naive patients with hepatitis B. *HEPATOLOGY* 2006;44:1656-1665.
20. Colonno RJ, Rose RE, Pokornowski K, Baldick CJ, Eggers B, Yu D, et al. Four year assessment of entecavir resistance in nucleoside-naive and lamivudine refractory patients. *J Hepatol* 2007;46(Suppl 1):S294.

21. Tenney DJ, Levine SM, Rose RE, Walsh AW, Weinheimer SP, Discotto L, et al. Clinical emergence of entecavir-resistant hepatitis B virus requires additional substitutions in virus already resistant to lamivudine. *Antimicrob Agents Chemother* 2004;48:3498-3507.
22. Tenney DJ, Rose RE, Baldick CJ, Levine SM, Pokornowski KA, Walsh AW, et al. Two-year assessment of entecavir resistance in lamivudine-refractory hepatitis B virus patients reveals different clinical outcomes depending on the resistance substitutions present. *Antimicrob Agents Chemother* 2007;51:902-911.
23. Sozzi T, Edwards R, Shaw T, Xiong S, Borroto-Esoda K, Locarnini S. Antiviral cross-resistance between clinically important HBV mutants: phenotypic testing using the recombinant HBV-baculovirus assay system. *Global Antiviral J* 2005;1(Suppl 2):79-80.
24. Brunelle MN, Jacquard AC, Pichoud C, Durantel D, Carrouee-Durantel S, Villeneuve JP, et al. Susceptibility to antivirals of a human HBV strain with mutations conferring resistance to both lamivudine and adefovir. *HEPATOLOGY* 2005;41:1391-1398.
25. Zhu Y, Curtis M, Snow-Lampart A, Yang H, Delaney W, Miller MD, et al. In vitro drug susceptibility analysis of hepatitis B virus clinical quasispecies populations. *J Clin Microbiol* 2007;45:3335-3341.
26. Hezode C, Chevaliez S, Bouvier-Alias M, Roudot-Thoraval F, Brillet R, Zafrani ES, et al. Efficacy and safety of adefovir dipivoxil 20 mg daily in HBeAg-positive patients with lamivudine-resistant hepatitis B virus and a suboptimal virological response to adefovir dipivoxil 10 mg daily. *J Hepatol* 2007;46:791-796.
27. Truong DM, Kaler G, Khandelwal A, Swaan PW, Nigam SK. Multi-level analysis of organic anion transporters 1, 3, and 6 reveals major differences in structural determinants of antiviral discrimination. *J Biol Chem* 2008;283:8654-8663.
28. Lampertico P, Vigano M, Manenti E, Iavarone M, Sablon E, Colombo M. Low resistance to adefovir combined with lamivudine: a 3-year study of 145 lamivudine-resistant hepatitis B patients. *Gastroenterology* 2007;133:1445-1451.
29. Gerolami R, Bourliere M, Colson P, Halfon P, Borentain P, Henry M, et al. Unusual selection of rtA181V HBV mutants cross-resistant to adefovir following prolonged lamivudine monotherapy: report of two cases. *Antivir Ther* 2006;11:1103-1106.
30. Karatayli E, Karayalcin S, Karaaslan H, Kayhan H, Turkyilmaz AR, Sahin F, et al. A novel mutation pattern emerging during lamivudine treatment shows cross-resistance to adefovir dipivoxil treatment. *Antivir Ther* 2007;12:761-768.
31. van Bommel F, Wunsche T, Mauss S, Reinke P, Bergk A, Schurmann D, et al. Comparison of adefovir and tenofovir in the treatment of lamivudine-resistant hepatitis B virus infection. *HEPATOLOGY* 2004;40:1421-1425.
32. van Bommel F, Zollner B, Sarrazin C, Spengler U, Huppe D, Moller B, et al. Tenofovir for patients with lamivudine-resistant hepatitis B virus (HBV) infection and high HBV DNA level during adefovir therapy. *HEPATOLOGY* 2006;44:318-325.
33. Qi X, Xiong S, Yang H, Miller M, Delaney WE 4th. In vitro susceptibility of adefovir-associated hepatitis B virus polymerase mutations to other antiviral agents. *Antivir Ther* 2007;12:355-362.
34. Viller S, Pichoud C, Billioud G, Barraud L, Durantel S, Trepo C, et al. Impact of hepatitis B virus rtA181V/T mutants on hepatitis B treatment failure. *J Hepatol* 2008;48:747-755.
35. Tan J, Degertekin B, Wong SN, Husain M, Oberhelman K, Lok AS. Tenofovir monotherapy is effective in hepatitis B patients with antiviral treatment failure to adefovir in the absence of adefovir-resistant mutations. *J Hepatol* 2008;48:391-398.

ARTICLE 2

Rodriguez C, Chevaliez S, Bensadoun P, Pawlotsky JM

*Dynamics of hepatitis B virus resistance revisited by
ultra-deep pyrosequencing*

J Virol ; en révision

Dynamics of Hepatitis B Virus Resistance Revisited by Ultra-Deep Pyrosequencing

Christophe Rodriguez,^{1,2} Stéphane Chevaliez,^{1,2} Paul Bensadoun,²
and Jean-Michel Pawlotsky^{1,2#}

¹National Reference Center for Viral Hepatitis B, C and Delta, Department of Virology,
Hôpital Henri Mondor, Université Paris-Est, Créteil, France;

²INSERM U955, Créteil, France

RUNNING TITLE : HBV Resistance and Ultra-Deep Pyrosequencing

#Corresponding author: Professor Jean-Michel Pawlotsky, MD, PhD, Department of
Virology, Hôpital Henri Mondor, 51 avenue du Maréchal de Lattre de Tassigny, 94010
Créteil, France

Tel: +33-1-4981-2827 ; Fax: +33-1-4981-4831

E-mail: jean-michel.pawlotsky@hmn.aphp.fr

ABSTRACT

Hepatitis B virus (HBV) resistance to nucleoside/nucleotide analogues is frequent. Ultra-deep pyrosequencing (UDPS) is a powerful new tool that can detect minor viral variants and characterize complex quasispecies mixtures. However, suitable software for clinically meaningful analysis of virological data generated by UDPS is lacking. We used UDPS and an original in-house software package to analyze the dynamics of adefovir-resistant HBV variants in patients with chronic HBV infection in whom adefovir resistance occurred during treatment. Amino acid substitutions known to confer resistance to adefovir were detected at baseline in most of the patients. The dynamics of adefovir-resistant variants were complex and differed among the patients, as a result of evolving differences in variant fitness. UDPS analysis revealed successive waves of selection of HBV populations with single and multiple amino acid substitutions. Adefovir-resistant variants were partially inhibited by lamivudine but remained fit in its presence. Resistance was diagnosed several months earlier with UDPS than with population sequencing. In conclusion, using original software for the analysis of viral sequences generated by UDPS and other next-generation sequencing methods in the context of antiviral drug resistance, we show that substitutions conferring HBV resistance to nucleoside/nucleotide analogues exist before treatment and that the dynamics of adefovir-resistant populations are much more complex and heterogeneous than previously thought and involve so far unknown amino acid substitutions. The approach described here is likely to have important implications for the assessment of antiviral drug resistance and may provide a valuable tool for early diagnosis of viral resistance in clinical practice..

INTRODUCTION

Approximately 240 million individuals worldwide are chronically infected with hepatitis B virus (HBV) (18). Chronic HBV infection is the leading cause of chronic liver disease and accounts for nearly a million deaths every year (9, 14, 34). Chronic hepatitis B can be treated with either pegylated interferon- α or nucleoside/nucleotide analogues. The latter drugs act by directly inhibiting the enzymatic function of HBV reverse transcriptase, the enzyme responsible for viral replication. Five such drugs have been approved for HBV therapy, namely three nucleoside analogues (lamivudine, telbivudine and entecavir) and two nucleotide analogues administered as prodrugs (adefovir dipivoxil and tenofovir disoproxil fumarate). The vast majority of HBV-infected patients have an indication for therapy with nucleoside/nucleotide analogues.

The main issues with nucleoside/nucleotide analogues are the need for long-term (probably life-long) administration and the possible emergence of viral resistance (36). Resistance is characterized by outgrowth of viral populations bearing amino acid substitutions that confer reduced sensitivity to the drug. This is due to the quasispecies distribution of HBV in infected individuals, i.e. the coexistence of a mixture of genetically distinct but closely related viral populations in an unstable equilibrium that depends strongly on their relative fitness (i.e. their ability to propagate efficiently) in a specific replicative environment (21, 22, 36). Resistant variants that emerge during treatment are thought to pre-exist as minor populations prior to treatment, but this remains to be demonstrated in the case of HBV. The fitness cost of drug resistance can gradually be offset by the accumulation of “compensatory” amino acid substitutions during replication (21, 22, 36). HBV resistance generally results in virological and biochemical breakthrough, followed by accelerated liver disease progression (21, 36).

Few techniques are available to study HBV resistance in the clinical setting. Population sequencing (or direct sequencing) is the most widely used, but it can only detect the dominant viral population(s). Reverse hybridization with the line probe assay can only detect variants representing at least 5% of the viral quasispecies, and can only identify substitutions already known to confer HBV resistance to a given drug (10, 17). Sequencing of multiple clones generated after polymerase chain reaction (PCR) amplification is cumbersome and time-consuming (19, 20, 32, 35). In addition, analysis of 20 clones per time point provides only a 95% probability that variants representing 10% or more of the viral quasispecies will be identified, whereas random minor variants with no clinical significance may also be highlighted with this method. Novel technical approaches are therefore needed to study antiviral drug resistance.

Next-generation sequencing techniques are capable of generating vast quantities of data without prior knowledge of a particular gene or sequence of interest. Most of these methods can only read short sequences, however, which precludes linkage analysis when studying viral genomes in a quasispecies (4, 16). The 454 technology (454 Life Sciences, Roche Diagnostics Corp., Branford, CT), based on ultra-deep pyrosequencing (UDPS), provides longer reads than most other techniques and is well suited to viral resistance studies (12, 13, 23, 29).

The vast amount of sequence data generated by next-generation sequencing techniques such as UDPS needs to be stored, filtered and analyzed in order to draw physiological and/or clinical inferences. Informatics tools developed thus far have been successful in pre-treating and classifying sequence data generated by next-generation methods (11), but they have failed to address the specific aspects of viral population dynamics in the clinical context of viral resistance. We therefore created our own set of software programs, the PyroPack[®] package, for the analysis of next-generation

sequencing data in the context of viral resistance and used it together with UDPS to study the dynamics of HBV resistance to adefovir.

MATERIALS AND METHODS

Patients

We studied 7 patients (5 men and 2 women, 26 to 59 years old) with chronic HBV infection who developed adefovir resistance during a pivotal, 240-week clinical trial of adefovir monotherapy (8). The characteristics of the 7 patients have been described elsewhere (20). None of them had previously received nucleoside/nucleotide treatment, and all received 10 mg/day adefovir for the full study period. In four cases (patients 1, 2, 5 and 7), lamivudine was added during adefovir therapy because of virological failure. Serial serum samples taken at baseline and during adefovir therapy were analyzed by UDPS. The data were then analyzed and interpreted with the PyroPack[®] package.

First HBV DNA PCR round

HBV DNA was extracted from 200-500 μ L of every available serum sample with the QIAamp DNA blood kit (Qiagen GmbH, Hilden, Germany), according to the manufacturer's instructions. A 630-bp fragment encompassing domains A to E of HBV reverse transcriptase was PCR-amplified with primers pol1 and pol2, as previously described (20, 24).

Controls

The first-round PCR amplicon from patient 1's baseline sample was cloned into TOPO TA Cloning 2.1 vector (Invitrogen, Carlsbad, CA), transformed by means of One Shot® TOP10 chemically competent *Escherichia Coli* (Invitrogen) and cultured in brain heart infusion (BHI) agar Petri dishes with 1 mg/mL ampicilline. Ten colonies were sequenced with M13 primers, according to the TOPO TA cloning protocol (Invitrogen). The sequences were aligned with ClustalX v2.0.9. One wild-type colony was selected and amplified into BHI medium containing 1mg/mL ampicilline. It was then purified by means of PureLink™ HiPure Plasmid Filter Maxiprep Kit (Invitrogen). Plasmid DNA was quantified by means of Quant-iT™ dsDNA Assay Kit (Invitrogen) and diluted in order to achieve final concentrations of 10^8 , 10^5 , and 5×10^3 copies/mL. The final DNA amounts were confirmed by means of a quantitative real-time PCR technique on ABI 7500 (Applied Biosystems, Carlsbad, CA). Each plasmid dilution was then amplified in triplicate in three independent PCR reactions using primers pol1 and pol2, as described above.

Ultra-deep pyrosequencing

A second “nested” PCR amplification was performed with internal primers pol3 and pol4 (20). that were modified to introduce a GS FLX bead adaptor and a specific identity tag (multiplex identifier, MID). A combination of 8 different MIDs was used to identify each sample. Amplicons containing the bead adaptor and MID were then purified in Nucleofast® 96 PCR plates (Clontech, Mountain View, CA), according to the manufacturer's instructions. The amplicons were then quantified with the Quant-iT™ PicoGreen® dsDNA kit (Invitrogen), fixed to beads, and amplified in a microemulsion

with the GS FLX Titanium emPCR kit (454 Life Sciences). The amplified beads were purified and enriched according to the manufacturer's instructions, counted with a Beckman Coulter Z1 particle counter (Beckman Coulter, Brea, CA), and deposited in a GS FLX Titanium PicoTiterPlate (454 Life Sciences). The pyrosequencing reaction was performed with the GS FLX Titanium sequencing kit on an FLX Genome Sequencer (454 Life Sciences).

Computer analysis of UDPS data

Data generated with the UDPS method were analyzed with our four in-house software programs included in the PyroPack[®] package and protected under IDDN, including PyroClass[®] (IDDN.FR.001.360005.001.S.P.2010.000.31230), PyroMute[®] (IDDN.FR.001.120008.001.S.C.2010.000.31230), PyroDyn[®] (IDDN.FR.001.370024.000.S.P.2010.000.31230), and PyroLink[®] (IDDN.FR.001.180032.000.S.P.2011.000.31230), designed respectively to classify, filter, model and link viral sequences generated with these methods.

Sequence data analysis with PyroPack[®] is based on the following procedure. Sequences generated by the GS-FLX sequencer are merged into a single file. Each sequence is identified by means of its sample tag and assigned to the file corresponding to the sample of origin by PyroClass[®].

PyroMute[®] then uses a number of quality filters to eliminate unreliable sequences. Sequence portions with a low Phred quality score are removed (1). Too short sequences (<50 bp), including those generated by the previous filtering step, are also eliminated. Sequences are then aligned by means of a modified version of the Smith-Waterman algorithm (28), in which the resolution of alignment matrices is accelerated

and the identification of insertions/deletions (so-called indels) and the correction of length errors in homopolymeric sequences are improved. Quality filters subsequently remove sequences with an identity score <80% relative to the consensus sequence of the patient's baseline sample. In the next step, an array of nucleotide substitutions and corresponding Phred quality scores is built and tested by means of a modified statistical test based on the binomial law (3), in order to eliminate sequences that are too rare and/or of poor quality, likely due to sequencing errors. Finally, the remaining nucleotide sequences, considered reliable, are converted into amino acid sequences and their respective frequencies are calculated. Each amino acid change is ascribed to its sequence of origin in order to subsequently analyze linkages between substitutions.

PyroDyn[®] uses data generated by PyroMute[®] to detect and quantify increases or decreases in amino acid substitutions through mathematical modeling of their variations and correlation with an exponential growth model, which provided the best fit with the observed data. Briefly, in every patient, the frequency of each amino acid at each position is established at each time point. Assuming that HBV resistance is governed by exponential outgrowth of selected resistant variants, the best fit curve is established for each substitution in each patient. Combined cut-off values have been established ($r^2 > 0.8$ and growth rate $>$ mean of the growth rates of all substitutions at all positions plus 2 standard deviations) to differentiate significant exponential changes from polymorphism fluctuations.

Finally, PyroLink[®] has been designed to analyze genetic linkages between amino acid substitutions which have been selected by PyroDyn[®] and to characterize the dynamics of viral populations bearing one or several amino acid substitutions over time. Briefly, linkages are automatically sought for every substitution identified with PyroDyn[®] as exponentially growing or decreasing over time. Then, sequences than span

all of the identified substitutions are extracted from PyroMute® data and, when more than 100 such sequences are available, the proportion of sequences bearing no (wild-type), one, 2, 3, etc, substitutions is calculated and used for subsequent analyses.

RESULTS

UDPS sequence data

UDPS analysis of the 9 controls revealed that the mean maximum error rate generated by the technology was $0.120 \pm 0.005\%$. However, this number varied according to the considered amino acid position. Therefore, only substitution frequencies higher than the mean maximal error rate for the corresponding amino acid position plus two standard deviations were taken into account in the analysis. The plasmids were tested in triplicate at different dilutions ranging from 5×10^3 to 10^8 copies/mL. No impact of the DNA amount on UDPS results was observed (data not shown).

UDPS was then applied to serum samples taken at baseline and frequently during therapy, representing a total of 119 serial samples from 7 patients who developed resistance to adefovir monotherapy (15 to 24 samples per patient). Approximately 480,000 sequences (111 Mbp) were generated, with 4010 ± 843 sequences per sample and a mean length of 382 ± 31 nt after eliminating excessively short sequences (≤ 50 nt) with PyroMute®. Overall, 10.2% of the generated sequences were eliminated by the software because of inadequate quality.

Presence of amino acid substitutions known to confer HBV resistance to nucleoside/nucleotide analogues in baseline samples

Table 1 shows the prevalence of amino acid substitutions known to confer HBV resistance to nucleoside/nucleotide analogues detected by UDPS at baseline in the 7 patients. In one patient the number of sequences generated in the reaction was too small for reliable assessment of baseline status. Five of the remaining six baseline samples were found to harbor rtA181V/T substitutions, one harbored the rtN236T substitution, and two harbored rtM204V/I substitutions. No substitutions known to improve the fitness of rtM204V/I variants in the presence of lamivudine, telbivudine or entecavir (rtV173L and rtL180M) were found in the baseline samples. One sample harbored the rtT184S/A/I/L substitution and none the rtR202G substitution; both these substitutions are known to confer full resistance to entecavir when associated with rtM204V/I and rtL180M.

Characterization of HBV resistance to adefovir by UDPS and PyroPack[®] analysis in patient 1

Patient 1 responded suboptimally to adefovir, and the HBV DNA level started to increase gradually after a nadir at month 6, until the end of follow-up at month 24. Fig. 1 shows the time course of the HBV DNA level, together with the dynamics of amino acid substitutions and HBV viral populations during adefovir therapy in this patient, as assessed by UDPS and PyroPack[®] analysis. UDPS-generated sequences were classified with the PyroClass[®] program. PyroMute[®] was used to characterize the dynamics of amino acid substitutions known to confer resistance to adefovir, including rt181V,

rt181T and rtN236T. Exponential outgrowth or an exponential decline of any amino acid substitution, including substitutions not known to be associated with resistance to adefovir or other HBV nucleoside/nucleotide analogues, were characterized by analyzing UDPS data with PyroDyn[®]. Finally, linkages were studied with PyroLink[®] which, when combined with PyroDyn[®], allowed us to characterize the dynamics of HBV viral populations during therapy.

PyroMute[®] analysis. As shown in Fig. 1A, outgrowth of amino acid substitutions at positions rtA181V, rtA181T and rtN236T, all known to confer resistance to adefovir, was detected at month 13 when UDPS sequences were analyzed with PyroMute[®] alone (5 months before they were detected by population sequencing), but no further information could be extracted from this analysis.

PyroDyn[®] analysis. Analysis with PyroDyn[®] allowed us to detect exponential outgrowth or decline of any amino acid substitution during the study period (Fig. 1B). This module confirmed the outgrowth of amino acid substitutions at positions rtA181 and rtN236, and also revealed: (i) simultaneous outgrowth of other amino acid substitutions present at baseline in small proportions, including rtY245H and rtD238N; (ii) an earlier wave of relative enrichment in amino acid substitutions not known to confer resistance to adefovir, including rtN138K, rtR139K and rtR212T. This enrichment occurred a few months after treatment initiation, was transient, and disappeared when outgrowth of amino acid substitutions at positions rt181 and rtN236 began.

Combined PyroDyn[®]/PyroLink[®] analysis. In Figs. 1C and 1D, we used both PyroDyn[®] and PyroLink[®] to take into consideration linkages between different amino acid substitutions (i.e. their presence in the same viral variant) and to characterize the dynamics of HBV populations that constituted the quasispecies. Fig. 1C shows the proportion of each viral variant in the quasispecies at each time point, while Fig. 1D

shows the absolute amount of each viral variant, in Log_{10} international units (IU)/mL, at each time point. The latter representation is more suitable for interpreting the dynamics of viral variants during antiviral therapy, which can be described as follows in patient 1.

(i) Immediately after treatment initiation, the apparent outgrowth of a first wave of variants with the single amino acid substitutions rtN138K, rtR139K and rtR212T (increased proportions in Fig. 1C) in fact corresponded to minor variants that were present at baseline and persisted, quantitatively unchanged, during adefovir administration, while the wild-type virus was profoundly inhibited. (ii) Immediately after the HBV DNA nadir was reached at month 2, the absolute amount of wild-type virus started to increase again, whereas the minor variants gradually lost their relative fitness and became nearly undetectable when outgrowth of adefovir-resistant variants started to be observed. (iii) The first wave of resistant-variant outgrowth was detected at month 17 and peaked at months 21-22, when the wild-type virus became undetectable. This wave was composed of viral variants bearing single amino acid substitutions known to confer adefovir resistance, including a majority of rtN236T and a minority of rtA181T. (iv) A second wave of outgrowth of adefovir-resistant variants then gradually replaced the first wave, probably following fitness acquisition by resistant variants bearing single and double amino acid substitutions, including, by order of frequency, rtN236T+rtA181T, rtY245H, rtN236T+rtY245H, and rtN236T+rtD238N.

Hepatitis B surface antigen substitutions selected together with adefovir resistance substitutions in patient 1

As shown in Fig. 1E, the prevalence of amino acid substitutions in the hepatitis B surface antigen (HBsAg) region increased in parallel to that of reverse transcriptase

substitutions associated with adefovir resistance. The former substitutions included sS143T and sM197T. They were not encoded by the nucleotide changes responsible for substitutions in the reverse transcriptase region selected by adefovir. In Fig. 1F, PyroDyn[®] and PyroLink[®] were used together to link substitutions in the reverse transcriptase and HBsAg regions and to describe the dynamics of viral variants during adefovir therapy, represented as the absolute amount (Log_{10} IU/mL) of each variant at each time point. HBsAg substitutions were linked to reverse transcriptase substitutions selected by adefovir (sS143T with rtA181T and sM197T with rtN236T). These HBsAg substitutions appeared to contribute to the fitness gain of the second wave of adefovir-resistant variants, with preferential outgrowth of sM197T+rtN236T and sS143T+rtA181T+rtN236T variants over single or other double mutants at this late stage of follow-up (Fig. 1F).

Characterization of HBV resistance to adefovir in the other 6 patients by means of UDPS and PyroDyn[®]/PyroLink[®] analysis

We compared the different analytical approaches in the remaining 6 patients (data not shown) and confirmed that the most meaningful description of HBV variant dynamics over time was obtained by combining PyroDyn[®] and PyroLink[®] to analyze UDPS data and by expressing the results as absolute amounts of HBV variants in Log_{10} IU/mL at each time point. The results of these analyses are shown in Fig. 2 and briefly summarized below. Contrary to patient 1, amino acid substitutions were not selected in the HBsAg region in the other patients.

In patient 2 (Fig. 2A), wild-type HBV declined gradually during adefovir administration but re-increased when treatment was stopped after approximately a

year. When adefovir was reintroduced a few weeks later, wild-type HBV declined again slowly and plateaued at approximately 10^4 IU/mL. The emergence of resistance was characterized by simultaneous selection of variants with the single rtN236T and rtA181V substitutions at week 27. Subsequently, the rtA181V variant became predominant and was responsible for the virological breakthrough. This variant was partially inhibited, but remained dominant, when lamivudine was added to adefovir after 43 months of therapy.

In patient 3 (Fig. 2B), resistance occurred at month 29 and was characterized by initial outgrowth of HBV variants with single or double amino acid substitutions at positions rt181 and rt236. In contrast to patient 2, a variant with the single rtN236T substitution took over and was responsible for the virological breakthrough. As in patient 2, this variant was partially inhibited by lamivudine but remained predominant on combination therapy.

Patient 4 exhibited a more complex resistance pattern (Fig. 2C). This patient had a suboptimal response to adefovir. During the plateau phase, which lasted approximately 20 months, with mild fluctuations, wild-type HBV was gradually replaced by a mixture of variants with single (rtY124H and rtN236T), double (rtY124H+rtN236T) and triple (rtY124H+rtN236T+rtN238T) amino acid substitutions that replicated at low levels. Wild-type virus returned when adefovir treatment was interrupted. Adefovir was reintroduced about 2 months later, and resistance then developed, along with a typical virological breakthrough due to outgrowth of a viral population bearing the single rtN236T substitution. This variant was partially inhibited by lamivudine. Subsequently, on adefovir-lamivudine combination therapy, the same mixture of single, double and triple variants as observed earlier during therapy emerged and replicated at a low level.

Patient 5 (Fig. 2D) exhibited a simple resistance pattern, characterized by a suboptimal response to adefovir and virological breakthrough at month 37, due to outgrowth of a resistant variant bearing two amino acid substitutions (rtA181V+rtN236T).

Patient 6 (Fig. 2E) responded to adefovir but low-level viremia persisted, with transient selection of a variant bearing an rtN238E substitution. Subsequently, a variant bearing the double rtN236T+rtN238E substitution took over and persisted at a low level (about 10^2 IU/mL).

Finally, in patient 7 (Fig. 2F), the virological breakthrough was related to the selection of a major viral population bearing the rtA181V substitution and a minor population bearing the double rtA181V+rtF221Y substitution. Both variants were partially inhibited but remained dominant when lamivudine was added to adefovir.

Diagnosis of HBV resistance to adefovir with different methods

Table 2 shows the time points at which HBV resistance was diagnosed in each patient with different methods, in months after the start of adefovir therapy. The methods compared were: (i) a 1-Log₁₀ HBV DNA increase over the nadir (virological breakthrough); (ii) population sequencing and clonal sequence analysis (these results have previously been reported (20)); and (iii) UDPS and PyroMute[®] analysis, and UDPS and PyroDyn[®]/PyroLink[®] combined analysis, as described in this study. In most patients, UDPS detected amino acid substitutions in low amounts at baseline. Diagnosis of resistance by UDPS and PyroDyn[®]/PyroLink[®] combined analysis was made approximately at the same time as with clonal sequence analysis, i.e. many months before population sequencing and virological breakthrough.

DISCUSSION

Drug resistance is the principal cause of antiviral treatment failure, which may result in clinical disease progression. Next-generation sequencing technologies, such as UDPS, have the capacity to generate thousands of sequences from complex genomic mixtures, including sequences from dominant, intermediate and minor viral populations (4, 16). UDPS-based GS FLX technology provides sequence reads of sufficient length to span the region of interest when studying HBV resistance to nucleoside/nucleotide analogues. To interpret the data generated with this method, we developed an original package of four complementary software programs, PyroPack[®], capable of analyzing large numbers of sequences (nearly 500,000 in this study) in the specific context of viral resistance.

Software developed thus far failed to provide clinically meaningful information in terms of the actual dynamics of sensitive and resistant viral populations over time and the discovery of new substitutions involved in these processes (2, 7, 11, 25, 30, 33). PyroDyn[®] is designed to detect exponential changes, based on a minimum of 3 blood samples from the patient in question. With PyroDyn[®], potential errors due to the first PCR reaction are eliminated, based on their inconsistent presence in serial samples. In addition, PyroDyn[®] can detect and track amino acid substitutions not known to play a role in viral resistance. Linkages between amino acid substitutions present in the same viral variant are established with PyroLink[®]. As shown in Table 1 and Fig. 1, PyroClass[®]/PyroMute[®] analysis is sufficient to characterize the viral quasispecies at a given time point. However, PyroDyn[®] and PyroLink[®] must then be used to obtain a sensitive and accurate description of viral population dynamics on antiviral treatment.

With this approach, the results of UDPS are better represented in IU/mL rather than as simple proportions of each viral population, which do not take changes in the total HBV DNA level into account.

Pre-existence of resistant HBV variants in patients never exposed to nucleoside/nucleotide analogues is an accepted concept. However, classical techniques, such as population sequencing, reverse hybridization or even cloning of PCR products followed by sequencing of the clones, are generally not sensitive enough to detect variants present in very small amounts due to their poor relative fitness compared to wild-type virus. UDPS has been used to assess the pre-existence of HBV variants resistant to nucleoside/nucleotide analogues in a few studies. These works however suffer from important methodological flaws, including insufficient sensitivity, lack of consideration of the error rate of the method to establish reliable cutoffs and ensure specificity, too short genomic region analyzed, and/or no linkage studies (15, 27, 29). Using the UDPS/PyroPack[®] approach, we found that variants with amino acid substitutions at positions rtA181 and rtN236 were already present as minor populations at baseline in most of the treatment-naïve patients who subsequently developed adefovir resistance, with a sensitivity $\leq 0.22\%$. These substitutions were also detected during therapy in the remaining patients, suggesting that they may also have been present at baseline but in amounts too small to be detected by UDPS. The frequency of adefovir resistance substitutions at baseline may have been overestimated compared to the general population, as the patients we studied were selected because adefovir resistance occurred during treatment. Nevertheless, substitutions at position rtM204, that confer cross-resistance to lamivudine, telbivudine and entecavir, were also found as minor populations at baseline in two patients, whereas amino acid substitutions that confer resistance to entecavir when associated with rtM204 substitutions were rarely

found. Larger-scale UDPS studies are now needed to assess the prevalence of primary resistance to nucleoside/nucleotide analogues in the general HBV-infected population.

During treatment, UDPS data analysis with the PyroPack[®] package, and particularly the PyroDyn[®] and PyroLink[®] modules, allowed us to characterize the complex dynamics of HBV populations more thoroughly than with classical methods. In this respect, adefovir resistance is a particularly interesting model, as adefovir lacks potency at the dose used in HBV therapy (10 mg daily): the patients we studied all had detectable HBV DNA on therapy, allowing us to apply UDPS to a large number of serial samples obtained over several years. UDPS sequence data analysis revealed successive waves of selection of HBV viral populations (Figs. 1 and 2). Our findings can be summarized as follows.

- (i) The dynamics of viral variants differed among the patients, despite the fact that they were receiving the same treatment, emphasizing the importance of the HBV quasispecies composition at the start of therapy and of individual viral variant fitness, both of which are patient-specific.
- (ii) In patients 1 and 4 the wild-type virus rapidly declined at the beginning of treatment, whereas a few pre-existing variants did not appear to be affected, suggesting primary resistance. These variants were soon replaced by variants with single amino acid substitutions (including at position rtA181 or rtN236), which were in turn replaced by more complex variants with multiple amino acid substitutions.
- (iii) Other patients exhibited a different pattern of resistance, with initial selection of variants with multiple substitutions, that were subsequently replaced by simpler variants.

- (iv) Other substitutions, such as those at positions rt238 in patients 1, 4 and 6 (including reversion to N in patient 1) and at rtY245 in patient 1, appeared to be associated with a fitness gain of rtN236T variants. Interestingly, rtN238T has already been reported in two patients who developed resistance to adefovir (31, 35).
- (v) Variants with substitutions at both rtA181 and rtN236 were present during follow-up in several patients and finally took over in two cases (patients 1 and 5).
- (vi) The addition of lamivudine always reduced HBV DNA levels but did not alter the relative fitness of adefovir-resistant variants, which remained dominant during combination therapy. Interestingly, variants with rtA181V/T substitutions partially responded to lamivudine.
- (vii) As shown in patient 1, substitutions in other genomic regions (such as HBsAg) can be associated with HBV resistance substitutions in the reverse transcriptase region and play a role in viral fitness *in vivo*. However, this was seen in only one patient in this study.

A substantial amount of novel information was generated with the UDPS/PyroPack[®] approach when compared to methods previously used to characterize HBV resistance to nucleoside/nucleotide analogues (6, 8, 10, 17, 19, 20, 32, 35). Indeed, population sequencing is not sensitive enough, detecting only dominant variant populations representing >25% of the viral quasispecies, and it does not allow for linkage studies, making it impossible to characterize the actual dynamics of viral populations on therapy (6, 8). Reverse hybridization methods are slightly more sensitive, being able to detect variants that represent 5%-10% of the viral quasispecies, but they require prior knowledge of the selected amino acid substitutions and can

hardly be used to detect numerous variants with single and multiple substitutions (10, 17). Finally, cloning of PCR products followed by sequencing is cumbersome and has limitations. In two prior studies in which we used this approach (19, 20), almost two years were needed to generate the full, clean set of sequence data, versus one week in the present study with UDPS plus a few days for analysis with PyroPack®. A large amount of information extracted from UDPS/PyroPack® analysis could not be generated with cloning/sequencing. In particular, this approach did not allow us to detect the outgrowth and/or decrease of a number of minority variant populations. In addition, the maximum number of clones analyzed (20-30 per time point) did not allow us to identify all of the different viral variants with single and multiple substitutions at the different time points or to characterize their actual dynamics over time, especially when these variants never became dominant.

As shown in Table 2, next-generation sequencing methods such as UDPS could be of considerable interest for early diagnosis of viral resistance during antiviral therapy. With the PyroDyn® module, resistance could be diagnosed at approximately the same time as with cloning and sequencing, and thus considerably earlier than with population sequencing (20). One limitation of this approach is that it requires detectable HBV DNA levels and may not therefore be feasible when more potent drugs than adefovir are used. Nevertheless, in our experience, UDPS can be successfully applied to samples containing HBV DNA levels as low as 100 IU/mL.

In conclusion, using PyroPack®, an original software package for analyzing viral sequences generated by UDPS and other next-generation sequencing methods in the context of antiviral resistance, (i) we showed that substitutions conferring resistance to nucleoside/nucleotide analogues pre-exist in patients who have never been exposed to these drugs; (ii) we characterized the complex and heterogeneous dynamics of adefovir-

resistant viral populations in a group of patients in whom resistance emerged during long-term adefovir therapy; (iii) we identified so far unknown amino acid substitutions that appeared to play an important role in HBV resistance to adefovir. These findings will also be helpful for understanding resistance to tenofovir, which shows cross-resistance with adefovir *in vitro*. Our findings imply that next-generation sequencing data analysis and PyroPack[®] software will have a number of applications in viral resistance assessment, as we recently reported with HCV and HIV (5, 26). Our aim is now to expand our experiments, further improve the software package and resolve practical issues in order to make the software publicly available for next-generation sequencing data analysis through an open web-based platform.

ACKNOWLEDGEMENTS

This work was funded by the French National Agency for Research on AIDS and Viral Hepatitis (ANRS), and we are grateful to Prof. Jean-François Delfraissy, the Director of ANRS, for his continuous support to our UDPS projects. Christophe Rodriguez is the recipient of a doctoral fellowship from Fondation pour la Recherche Médicale (FRM). The hosting laboratory is funded by FRM as an “FRM Team”.

The authors thank Thierry Ravard for his help with mathematical modeling, Françoise Roudot-Thoraval for her help with statistical tests, and Katyna Borroto-Esoda and Gilead Sciences for providing the patient samples used in this study.

REFERENCES

1. **Altshuler, D., V. J. Pollara, C. R. Cowles, W. J. van Etten, J. Baldwin, L. Linton, and E. S. Lander.** 2000. An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature* **407**:513-6.
2. **Archer, J., A. Rambaut, B. E. Taillon, P. R. Harrigan, M. Lewis, and D. L. Robertson.** 2010. The evolutionary analysis of emerging low frequency HIV-1 CXCR4 using variants through time-an ultra-deep approach. *PLoS Comput Biol* **6**:e1001022.
3. **Campbell, P. J., E. D. Pleasance, P. J. Stephens, E. Dicks, R. Rance, I. Goodhead, G. A. Follows, A. R. Green, P. A. Futreal, and M. R. Stratton.** 2008. Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. *Proc Natl Acad Sci USA* **105**:13081-6.
4. **Chevaliez, S., C. Rodriguez, and J. M. Pawlotsky.** 2012. New virologic tools for management of chronic hepatitis B and C. *Gastroenterology* **142**:1303-13.
5. **Chevaliez, S., C. Rodriguez, A. Soulier, A. Ahmed-Belkacem, C. Hezode, and J. M. Pawlotsky.** 2011. Molecular characterization of HCV resistance to telaprevir by means of ultra-deep pyrosequencing: preexisting resistant variants and dynamics of resistant populations. *J Hepatol* **54 (Suppl. 1)**:S30.
6. **Colonna, R. J., R. Rose, C. J. Baldick, S. Levine, K. Pokornowski, C. F. Yu, A. Walsh, J. Fang, M. Hsu, C. Mazzucco, B. Eggers, S. Zhang, M. Plym, K. Kluszczewski, and D. J. Tenney.** 2006. Entecavir resistance is rare in nucleoside naive patients with hepatitis B. *Hepatology* **44**:1656-65.
7. **Gianella, S., W. Delpont, M. E. Pacold, J. A. Young, J. Y. Choi, S. J. Little, D. D. Richman, S. L. Kosakovsky Pond, and D. M. Smith.** 2011. Detection of minority

- resistance during early HIV-1 infection: natural variation and spurious detection rather than transmission and evolution of multiple viral variants. *J Virol* **85**:8359-67.
8. **Hadziyannis, S. J., N. C. Tassopoulos, E. J. Heathcote, T. T. Chang, G. Kitis, M. Rizzetto, P. Marcellin, S. G. Lim, Z. Goodman, J. Ma, C. L. Brosgart, K. Borroto-Esoda, S. Arterburn, and S. L. Chuck.** 2006. Long-term therapy with adefovir dipivoxil for HBeAg-negative chronic hepatitis B for up to 5 years. *Gastroenterology* **131**:1743-51.
 9. **Kim, W. R.** 2009. Epidemiology of hepatitis B in the United States. *Hepatology* **49**:S28-34.
 10. **Lok, A. S., F. Zoulim, S. Locarnini, A. Mangia, G. Niro, H. Decraemer, G. Maertens, F. Hulstaert, K. De Vreese, and E. Sablon.** 2002. Monitoring drug resistance in chronic hepatitis B virus (HBV)-infected patients during lamivudine therapy: evaluation of performance of INNO-LiPA HBV DR assay. *J Clin Microbiol* **40**:3729-34.
 11. **Macalalad, A. R., M. C. Zody, P. Charlebois, N. J. Lennon, R. M. Newman, C. M. Malboeuf, E. M. Ryan, C. L. Boutwell, K. A. Power, D. E. Brackney, K. N. Pesko, J. Z. Levin, G. D. Ebel, T. M. Allen, B. W. Birren, and M. R. Henn.** 2012. Highly sensitive and specific detection of rare variants in mixed viral populations from massively parallel sequence data. *PLoS Comput Biol* **8**:e1002417.
 12. **Margeridon-Thermet, S., N. S. Shulman, A. Ahmed, R. Shahriar, T. Liu, C. Wang, S. P. Holmes, F. Babrzadeh, B. Gharizadeh, B. Hanczaruk, B. B. Simen, M. Egholm, and R. W. Shafer.** 2009. Ultra-deep pyrosequencing of hepatitis B virus quasispecies from nucleoside and nucleotide reverse-transcriptase

- inhibitor (NRTI)-treated patients and NRTI-naive patients. *J Infect Dis* **199**:1275-85.
13. **Margulies, M., M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bembgen, J. Berka, M. S. Braverman, Y. J. Chen, Z. Chen, S. B. Dewell, L. Du, J. M. Fierro, X. V. Gomes, B. C. Godwin, W. He, S. Helgesen, C. H. Ho, G. P. Irzyk, S. C. Jando, M. L. Alenquer, T. P. Jarvie, K. B. Jirage, J. B. Kim, J. R. Knight, J. R. Lanza, J. H. Leamon, S. M. Lefkowitz, M. Lei, J. Li, K. L. Lohman, H. Lu, V. B. Makhijani, K. E. McDade, M. P. McKenna, E. W. Myers, E. Nickerson, J. R. Nobile, R. Plant, B. P. Puc, M. T. Ronan, G. T. Roth, G. J. Sarkis, J. F. Simons, J. W. Simpson, M. Srinivasan, K. R. Tartaro, A. Tomasz, K. A. Vogt, G. A. Volkmer, S. H. Wang, Y. Wang, M. P. Weiner, P. Yu, R. F. Begley, and J. M. Rothberg.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**:376-80.
 14. **McMahon, B. J.** 2009. The natural history of chronic hepatitis B virus infection. *Hepatology* **49**:S45-55.
 15. **Mello, F. C., B. V. Lago, L. L. Lewis-Ximenez, C. A. Fernandes, and S. A. Gomes.** 2012. Detection of mixed populations of wild-type and YMDD hepatitis B variants by pyrosequencing in acutely and chronically infected patients. *BMC Microbiol* **12**:96.
 16. **Metzker, M. L.** 2010. Sequencing technologies-the next generation. *Nat Rev Genet* **11**:31-46.
 17. **Niesters, H. G., F. Zoulim, C. Pichoud, M. Buti, F. Shapiro, N. D'Heuvaert, L. Celis, J. Doutreloigne, and E. Sablon.** 2010. Validation of the INNO-LiPA HBV DR assay (version 2) in monitoring hepatitis B virus-infected patients receiving nucleoside analog treatment. *Antimicrob Agents Chemother* **54**:1283-9.

18. **Ott, J. J., G. A. Stevens, J. Groeger, and S. T. Wiersma.** 2012. Global epidemiology of hepatitis B virus infection: new estimates of age-specific HBsAg seroprevalence and endemicity. *Vaccine* **30**:2212-9.
19. **Pallier, C., L. Castera, A. Soulier, C. Hezode, P. Nordmann, D. Dhumeaux, and J. M. Pawlotsky.** 2006. Dynamics of hepatitis B virus resistance to lamivudine. *J Virol* **80**:643-53.
20. **Pallier, C., C. Rodriguez, R. Brillet, P. Nordmann, C. Hezode, and J. M. Pawlotsky.** 2009. Complex dynamics of hepatitis B virus resistance to adefovir. *Hepatology* **49**:50-9.
21. **Pawlotsky, J. M.** 2005. The concept of hepatitis B virus mutant escape. *J Clin Virol* **34 (Suppl. 1)**:S125-9.
22. **Pawlotsky, J. M.** 2011. Treatment failure and resistance with direct-acting antiviral drugs against hepatitis C virus. *Hepatology* **53**:1742-51.
23. **Petrosino, J. F., S. Highlander, R. A. Luna, R. A. Gibbs, and J. Versalovic.** 2009. Metagenomic pyrosequencing and microbial identification. *Clin Chem* **55**:856-66.
24. **Pichoud, C., B. Seigneres, Z. Wang, C. Trepo, and F. Zoulim.** 1999. Transient selection of a hepatitis B virus polymerase gene mutant associated with a decreased replication capacity and famciclovir resistance. *Hepatology* **29**:230-7.
25. **Poon, A. F., L. C. Swenson, W. W. Dong, W. Deng, S. L. Kosakovsky Pond, Z. L. Brumme, J. I. Mullins, D. D. Richman, P. R. Harrigan, and S. D. Frost.** 2010. Phylogenetic analysis of population-based and deep sequencing data to identify coevolving sites in the nef gene of HIV-1. *Mol Biol Evol* **27**:819-32.
26. **Rodriguez, C., C. Soulie, A. G. Marcelin, V. Calvez, D. Descamps, C. Charpentier, P. Flandre, P. Bellecave, G. Lee, P. R. Harrigan, J. M. Pawlotsky, and B. Masquelier.** 2011. HIV-1 coreceptor usage prediction by ultra-deep

- pyrosequencing increases the rate of patients treated with success by CCR5 antagonists. *Antiviral Ther* **16 (Suppl. 1)**:A100.
27. **Rodriguez-Frias, F., D. Tabernero, J. Quer, J. I. Esteban, I. Ortega, E. Domingo, M. Cubero, S. Camos, C. Ferrer-Costa, A. Sanchez, R. Jardí, M. Schaper, M. Homs, D. Garcia-Cehic, J. Guardia, R. Esteban, and M. Buti.** 2012. Ultra-deep pyrosequencing detects conserved genomic sites and quantifies linkage of drug-resistant amino acid changes in the hepatitis B virus genome. *PLoS One* **7**:e37874.
 28. **Smith, T. F., and M. S. Waterman.** 1981. Identification of common molecular subsequences. *J Mol Biol* **147**:195-7.
 29. **Solmone, M., D. Vincenti, M. C. Prosperi, A. Bruselles, G. Ippolito, and M. R. Capobianchi.** 2009. Use of massively parallel ultradeep pyrosequencing to characterize the genetic diversity of hepatitis B virus in drug-resistant and drug-naive patients and to detect minor variants in reverse transcriptase and hepatitis B s antigen. *J Virol* **83**:1718-26.
 30. **Tsibris, A. M., B. Korber, R. Arnaout, C. Russ, C. C. Lo, T. Leitner, B. Gaschen, J. Theiler, R. Paredes, Z. Su, M. D. Hughes, R. M. Gulick, W. Greaves, E. Coakley, C. Flexner, C. Nusbaum, and D. R. Kuritzkes.** 2009. Quantitative deep sequencing reveals dynamic HIV-1 escape and large population shifts during CCR5 antagonist therapy in vivo. *PloS One* **4**:e5683.
 31. **Villet, S., C. Pichoud, G. Billioud, L. Barraud, S. Durantel, C. Trepo, and F. Zoulim.** 2008. Impact of hepatitis B virus rtA181V/T mutants on hepatitis B treatment failure. *J Hepatol* **48**:747-55.

32. **Villet, S., C. Pichoud, J. P. Villeneuve, C. Trepo, and F. Zoulim.** 2006. Selection of a multiple drug-resistant hepatitis B virus strain in a liver-transplanted patient. *Gastroenterology* **131**:1253-61.
33. **Wang, C., Y. Mitsuya, B. Gharizadeh, M. Ronaghi, and R. W. Shafer.** 2007. Characterization of mutation spectra with ultra-deep pyrosequencing: application to HIV-1 drug resistance. *Genome Res* **17**:1195-201.
34. **Yang, J. D., and L. R. Roberts.** 2010. Hepatocellular carcinoma: a global view. *Nat Rev Gastroenterol Hepatol* **7**:448-58.
35. **Yim, H. J., M. Hussain, Y. Liu, S. N. Wong, S. K. Fung, and A. S. Lok.** 2006. Evolution of multi-drug resistant hepatitis B virus during sequential therapy. *Hepatology* **44**:703-12.
36. **Zoulim, F., and S. Locarnini.** 2009. Hepatitis B virus resistance to nucleos(t)ide analogues. *Gastroenterology* **137**:1593-608.

FIGURE LEGENDS

Fig. 1. Dynamics of amino acid substitutions and viral variants on antiviral therapy in patient 1. Dynamics of HBV DNA levels are shown at the back in blue. Dynamics of amino acid substitutions and HBV viral variants have been assessed by UDPS and analyzed with different methods, in patient 1. (A) Analysis of UDPS data on the reverse transcriptase domain with PyroMute[®], which detects only amino acid substitutions known to be associated with HBV resistance; (B) Analysis of UDPS data on the reverse transcriptase domain with PyroDyn[®], which detects any exponential increase or decrease in the prevalence of an amino acid substitution (results are expressed as percentages of the entire viral quasispecies); (C) Combined analysis of UDPS data with PyroDyn[®] and PyroLink[®], which detects any exponential increase or decrease in the prevalence of an HBV variant population (results are expressed as percentages of the entire viral quasispecies); (D) Combined analysis of UDPS data on the reverse transcriptase domain with PyroDyn[®] and PyroLink[®] (results are expressed as the absolute amount of each HBV variant in Log₁₀ IU/mL); (E) Analysis of UDPS data on the HBsAg domain with PyroDyn[®] (results are expressed as percentages of the entire viral quasispecies); (F) Combined analysis of UDPS data on both the reverse transcriptase and HBsAg domains with PyroDyn[®] and PyroLink[®] (results are expressed as the absolute amount of each HBV variant in Log₁₀ IU/mL).

Fig. 2. Dynamics of viral variants on antiviral therapy in patients 2 to 7. Dynamics of HBV DNA levels are shown at the back in blue. Dynamics of HBV variants have been assessed by UDPS and analyzed with PyroDyn[®] and PyroLink[®] combined, which detect any exponential increase or decrease in the prevalence of an HBV variant population

(results are expressed as the absolute amount of each HBV variant in Log_{10} IU/mL) in patient 2 (A), patient 3 (B), patient 4 (C), patient 5 (D), patient 6 (E) and patient 7 (F).

Table 1. Prevalence of amino acid substitutions known to confer HBV resistance to nucleoside/nucleotide analogues in the 7 patients at baseline, as detected by UDPS and analyzed with PyroClass[®] and PyroMute[®]. The cutoff value of the method was calculated at 0.13%. ND: not determined. rtM204V/I confers cross-resistance to lamivudine, telbivudine, and entecavir; rtV173L and rtL180M confer improved fitness to rtM204V/I variants; rtN236T confers resistance to adefovir; rtA181V/T confers cross-resistance to lamivudine, telbivudine and adefovir; rtT184S/A/I/L and rtR202G confer additional resistance to entecavir and improved fitness to variants bearing rtM204V/I, in the presence of this drug.

	HBV DNA level (Log IU/mL)	rtV173L	rtL180M	rtA181V/T	rtT184S/A/I/L	rtR202G	rtM204V/I	rtN236T
Patient 1	6.6	<cutoff	<cutoff	0.47%	<cutoff	<cutoff	0.31%	0.25%
Patient 2	7.8	<cutoff	<cutoff	0.18%	<cutoff	<cutoff	<cutoff	<cutoff
Patient 3	6.8	<cutoff	<cutoff	0.22%	<cutoff	<cutoff	0.32%	<cutoff
Patient 4	5.5	<cutoff	<cutoff	<cutoff	0.17%	<cutoff	<cutoff	<cutoff
Patient 5	7.6	<cutoff	<cutoff	0.33%	<cutoff	<cutoff	<cutoff	<cutoff
Patient 6*	6.2	ND	ND	ND	ND	ND	ND	ND
Patient 7	6.1	<cutoff	<cutoff	0.33%	<cutoff	<cutoff	<cutoff	<cutoff
Detection cutoff of the substitution†	-	0.05%	0.03%	0.08%	0.15%	0.18%	0.21%	0.22%

*The number of sequences obtained with the baseline sample was too small in this patient to yield reliable results.

†The detection cutoff of each substitution was calculated as the mean maximal error rate for the given substitution plus two standard deviations

Table 2. Time to diagnosis of HBV resistance to adefovir (in months after treatment initiation) in the 7 patients, based on various analytical methods. IQR: interquartile range.

Patient	>1 Log ₁₀ HBV DNA level increase over nadir	Population sequencing*	Clonal sequence analysis*	UDPS + PyroMute®†	UDPS + Pyromute® + PyroDyn®/PyroLink®‡
Patient 1	17	18	6	0	16
Patient 2	30	27	27	15	27
Patient 3	31	29	9	0	11
Patient 4	21	14	9	2	12
Patient 5	37	37	28	2	28
Patient 6	33	42	20	15	15
Patient 7	28	26	20	7	17
Median (IQR)	30 (24-32)	27 (22-33)	20 (9-23)	2 (1-11)	16 (13-22)

*The results of population sequencing and clonal sequence analysis have previously been published.

†Time to detection of the amino acid substitutions subsequently associated with virological breakthrough

‡Time to detection of an exponential increase in the amount of resistant HBV variants

Figure 1A

Patient 1

ADV

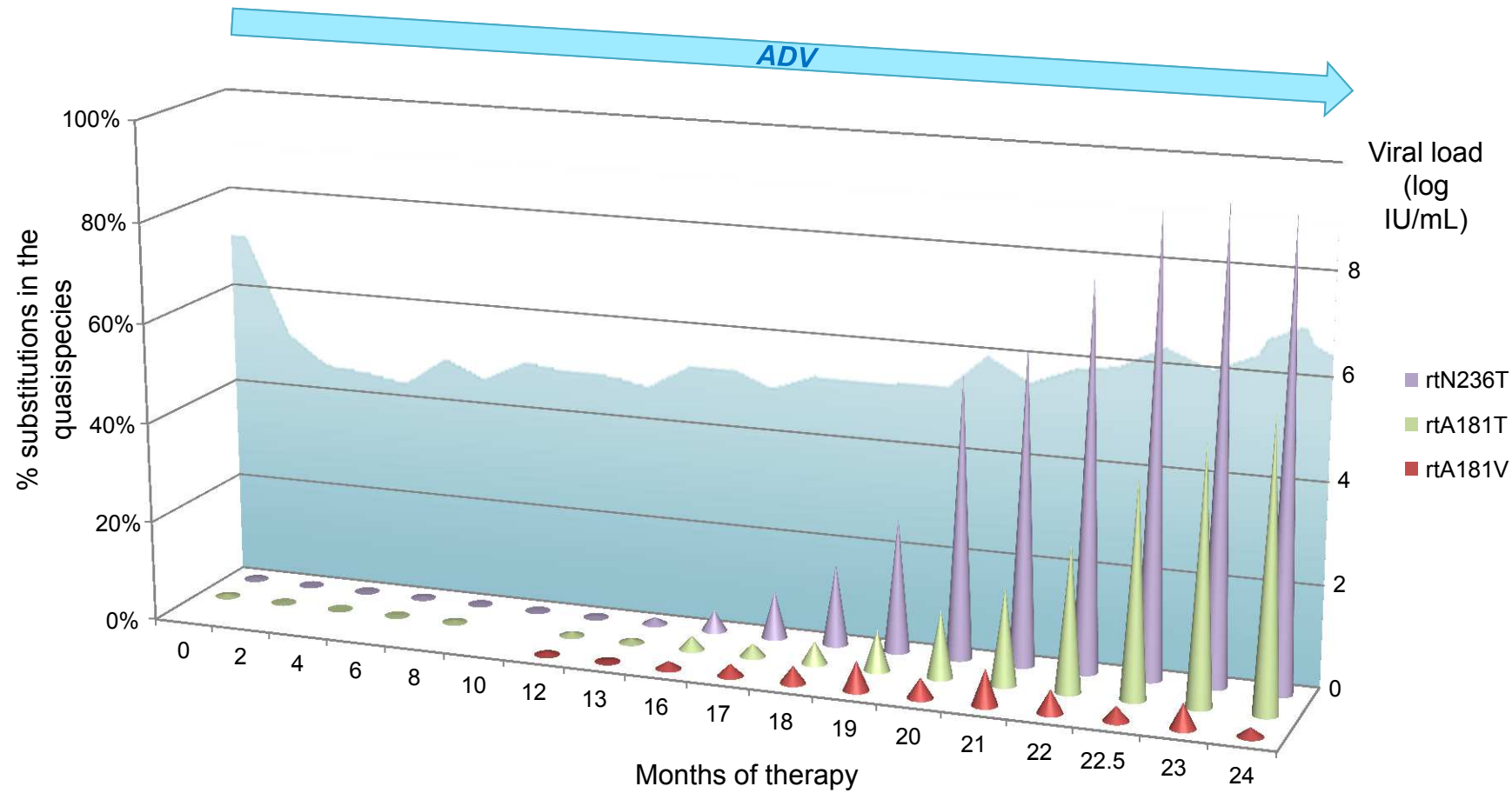


Figure 1B

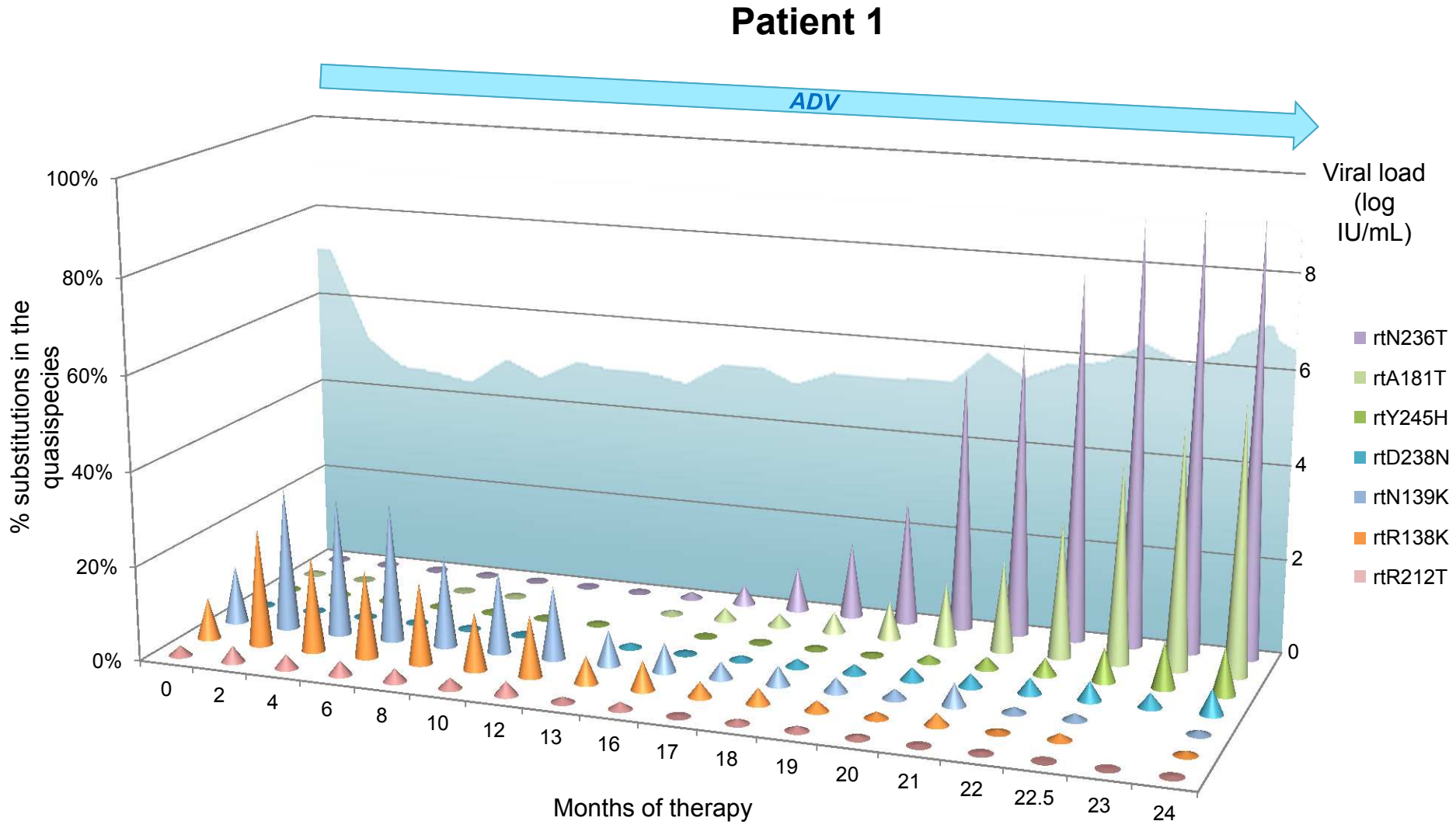


Figure 1C

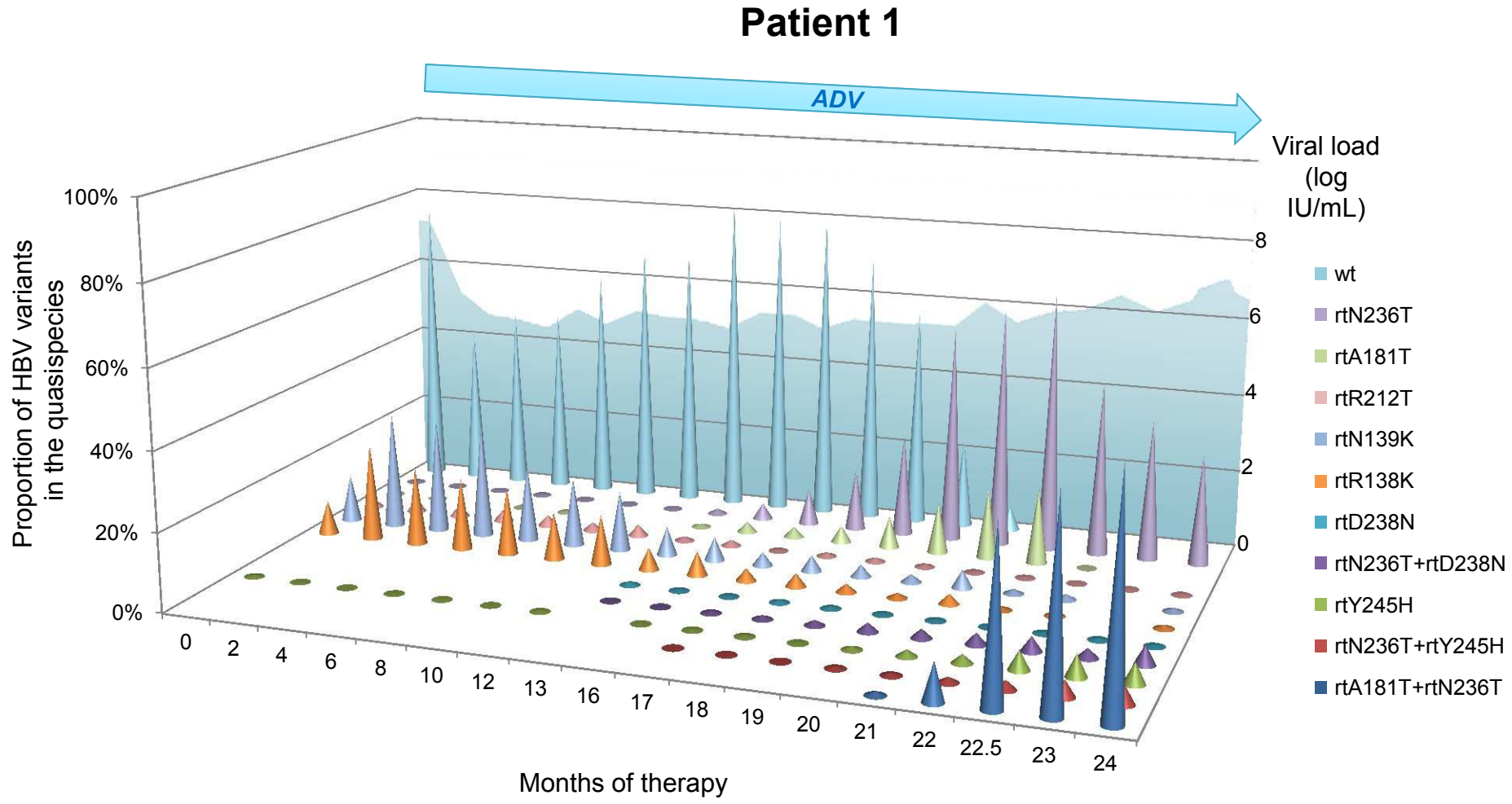


Figure 1D

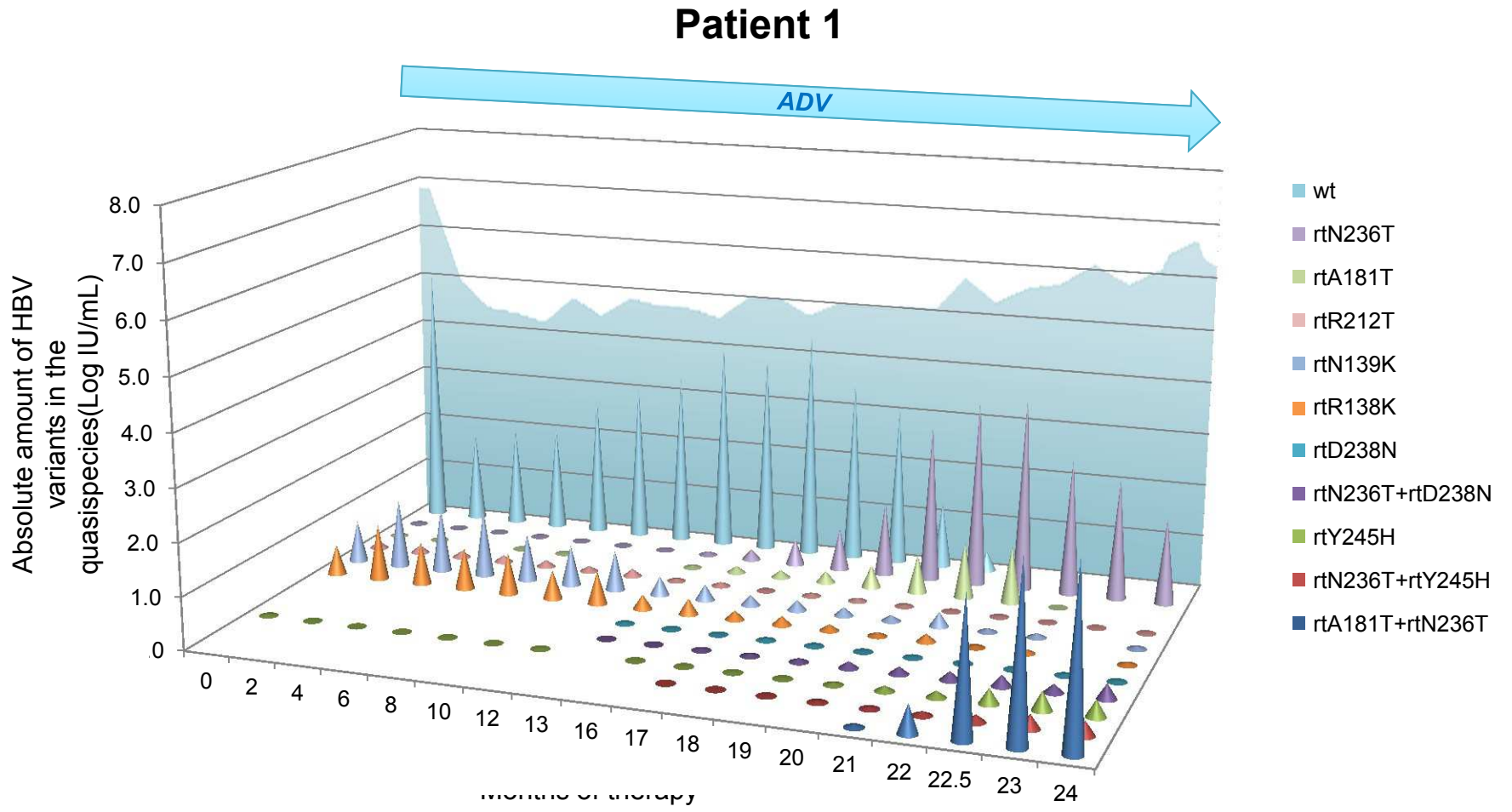


Figure 1E

Patient 1

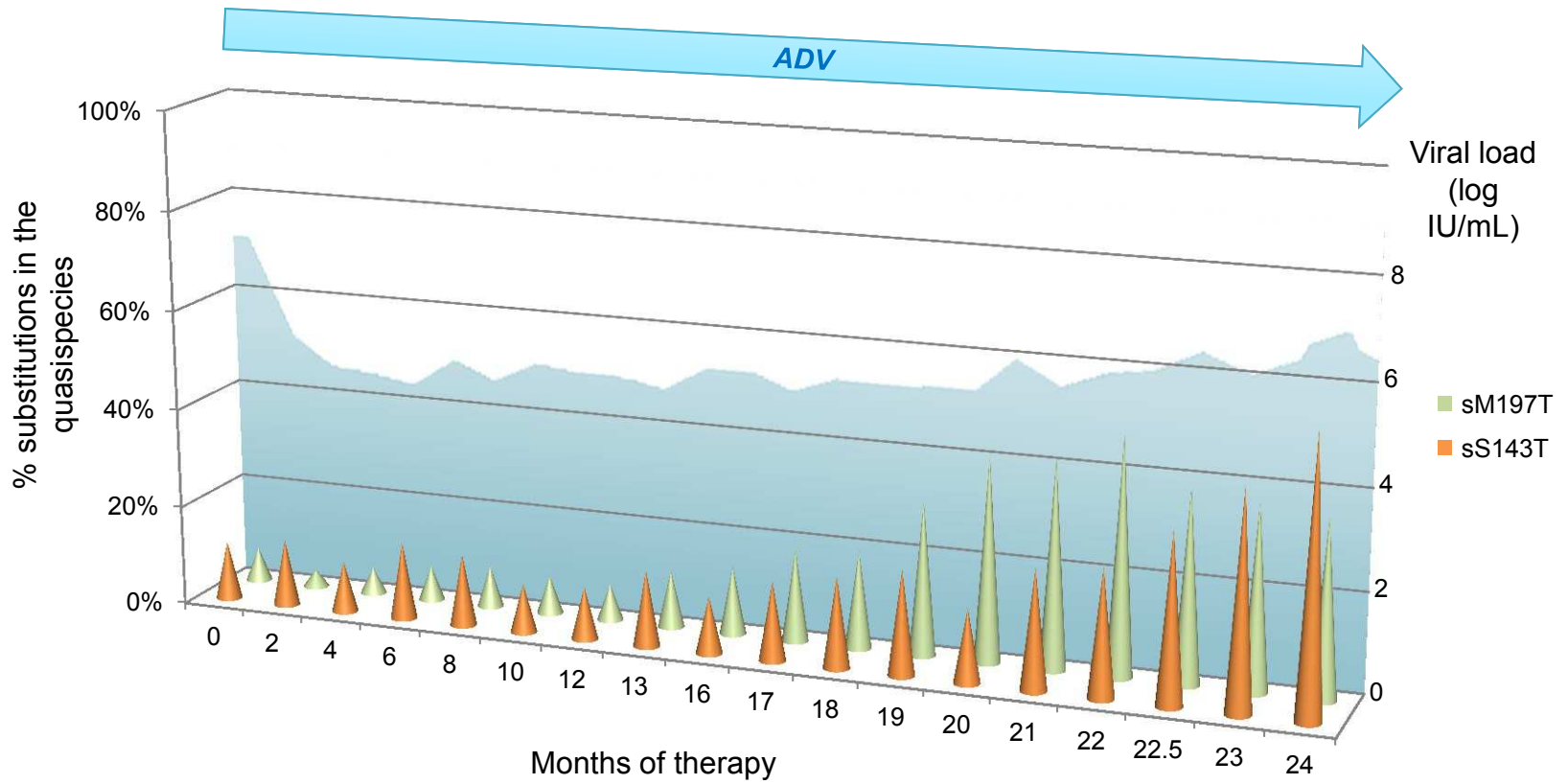


Figure 1F

Patient 1

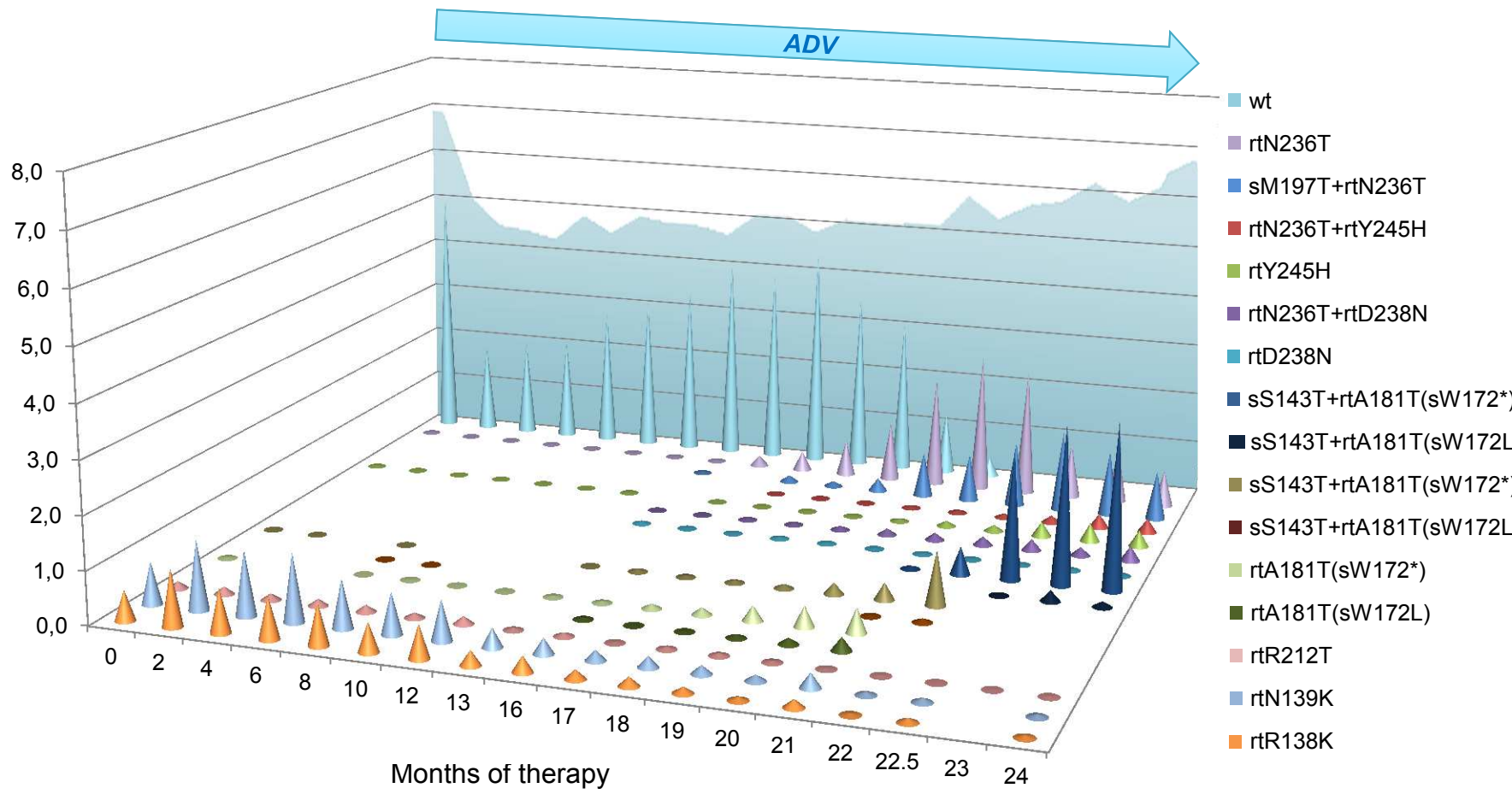


Figure 2A

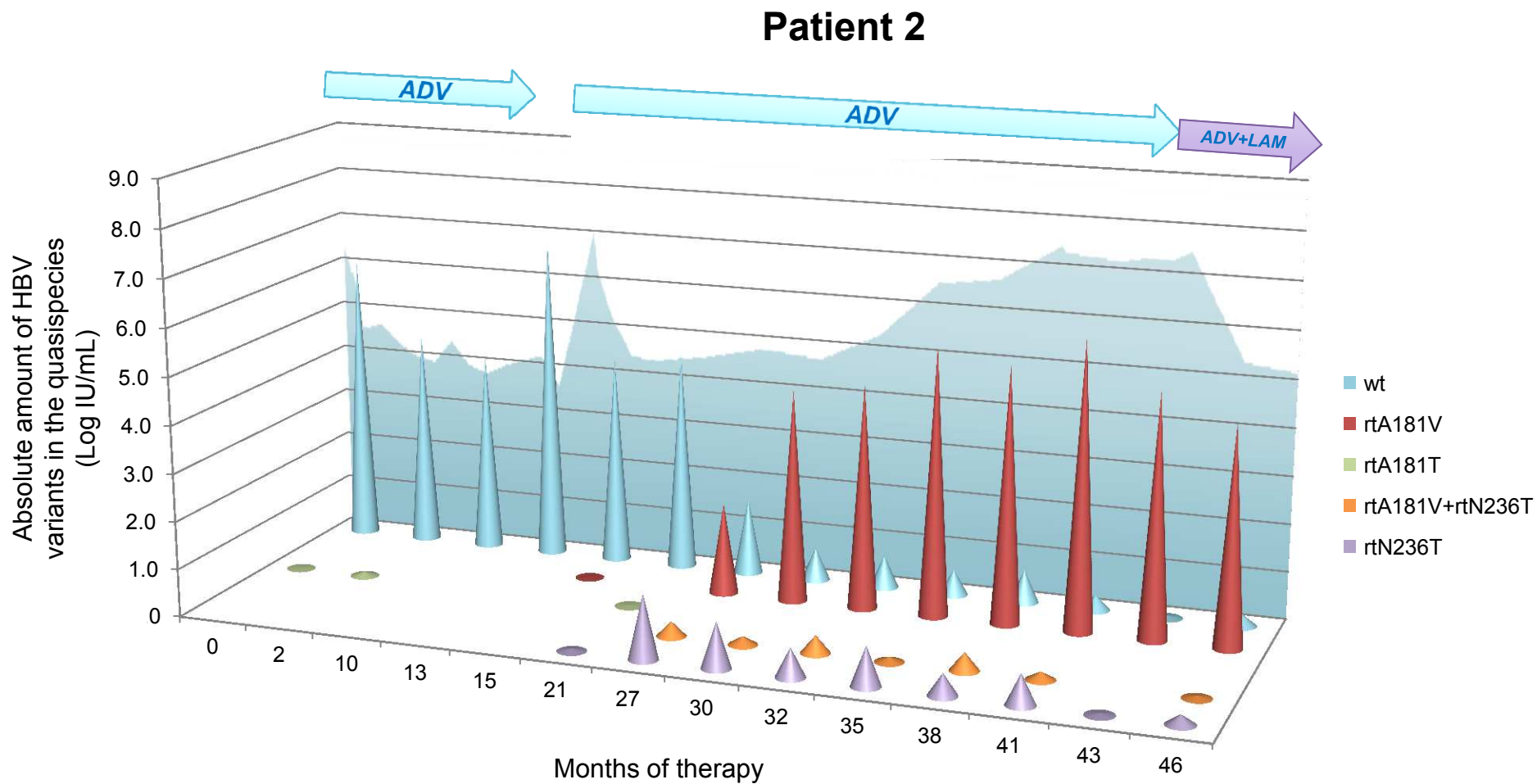


Figure 2B

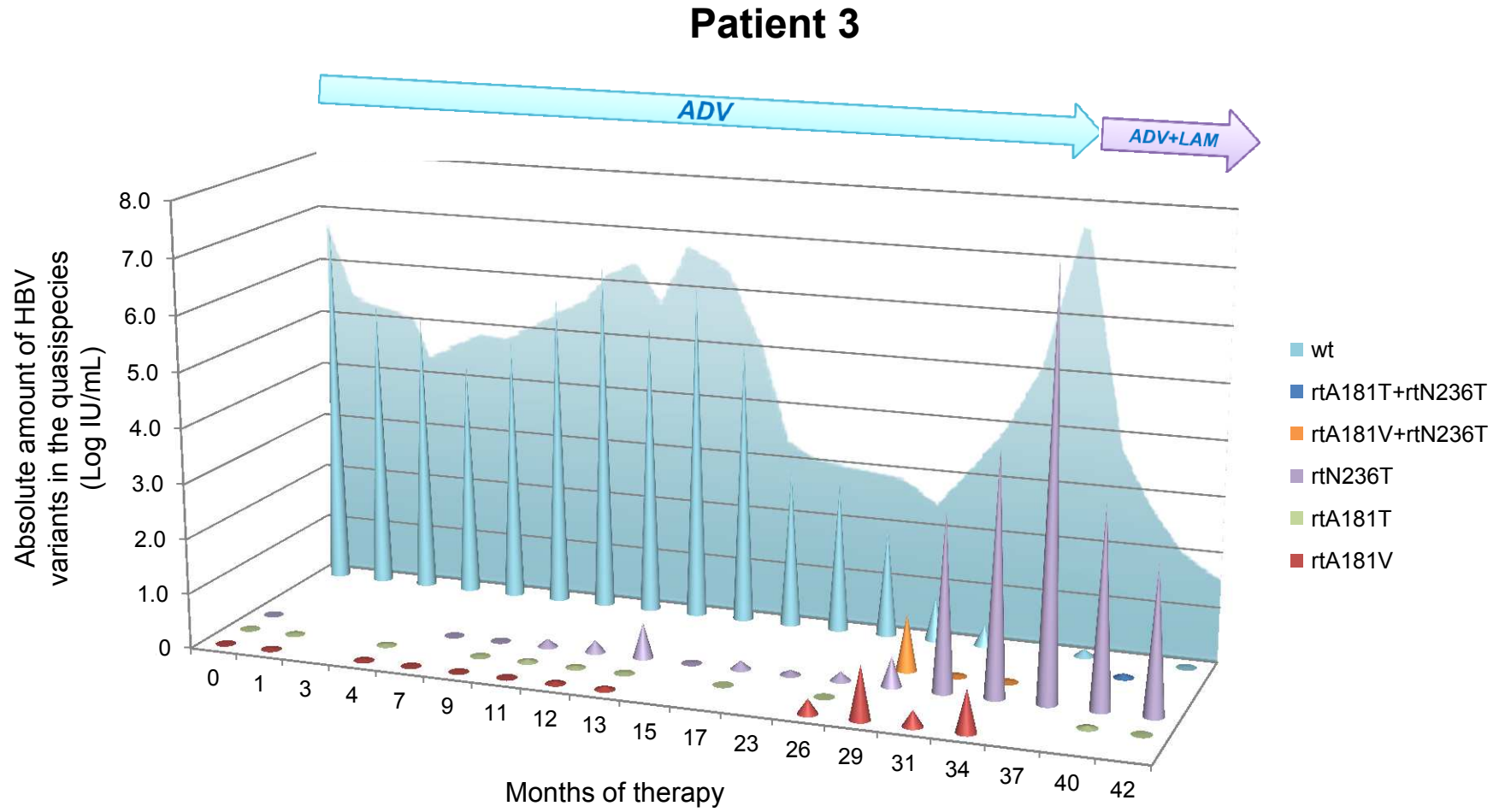


Figure 2C

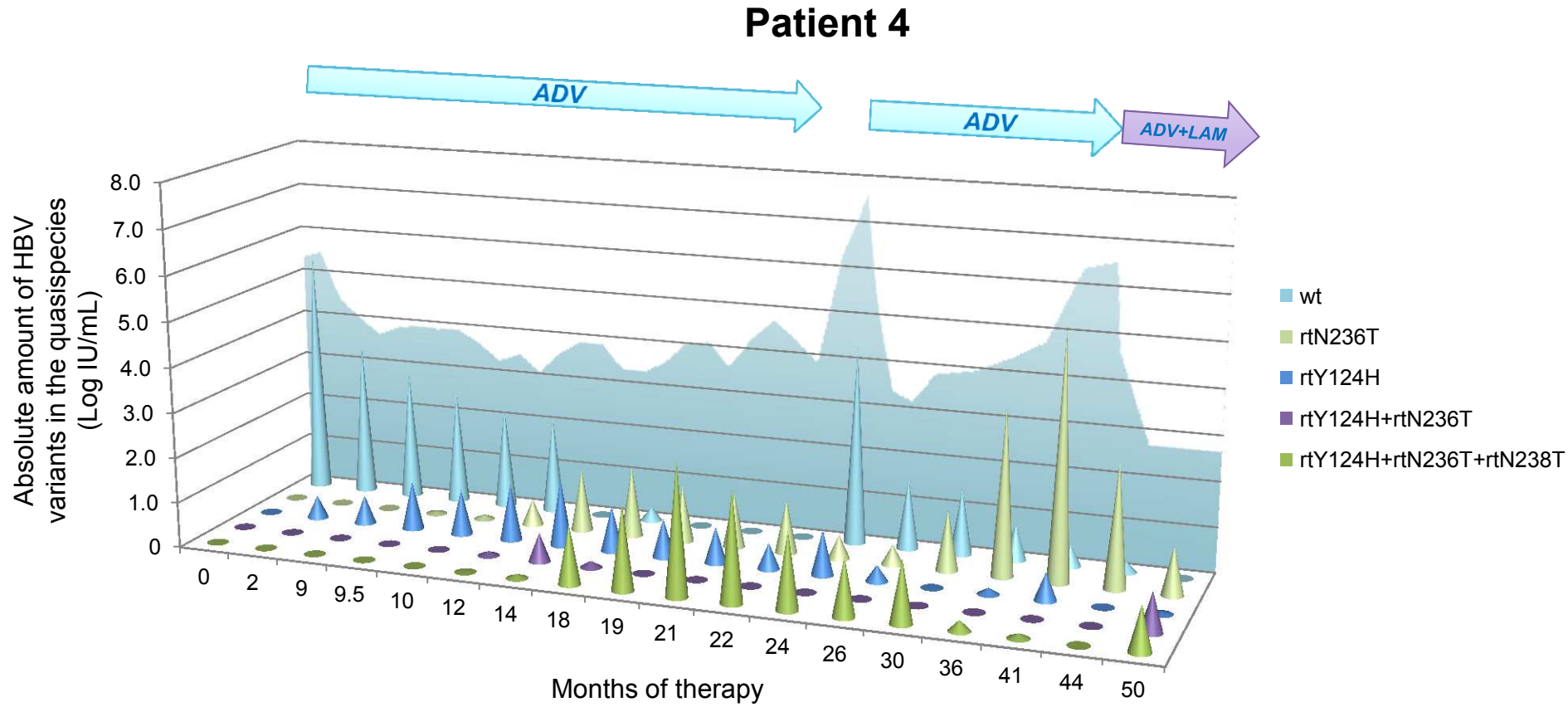


Figure 2D

Patient 5

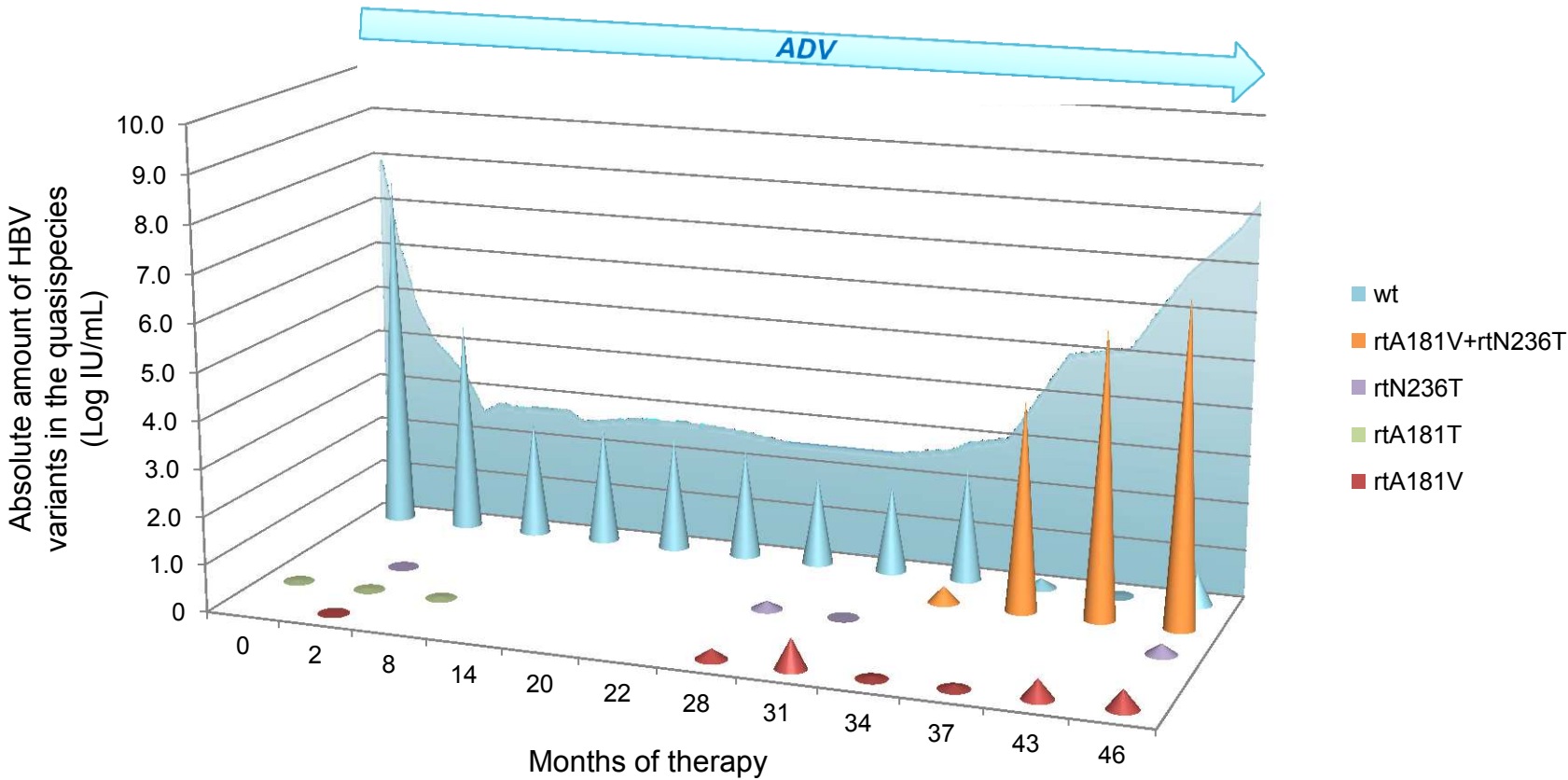


Figure 2E

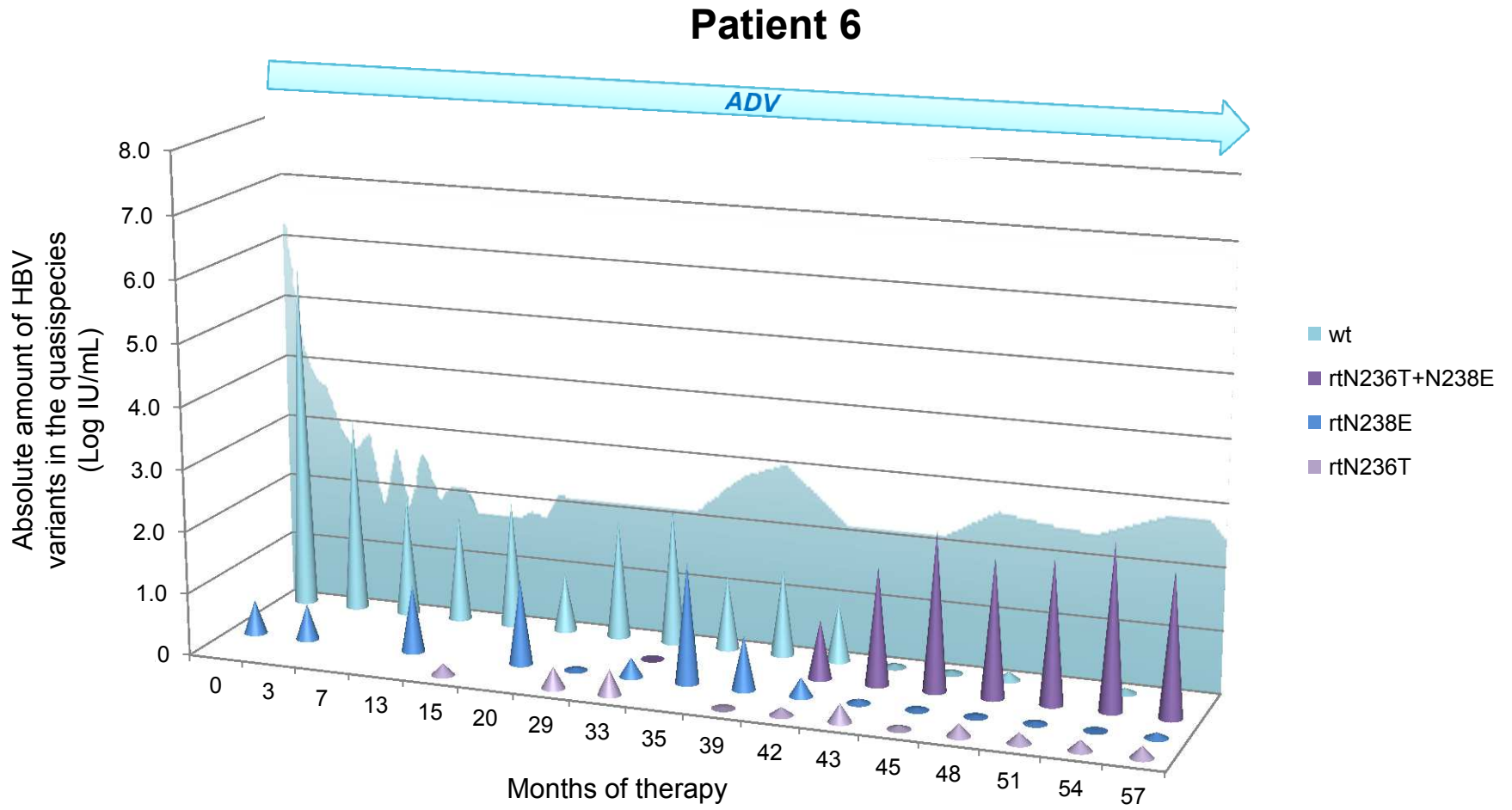
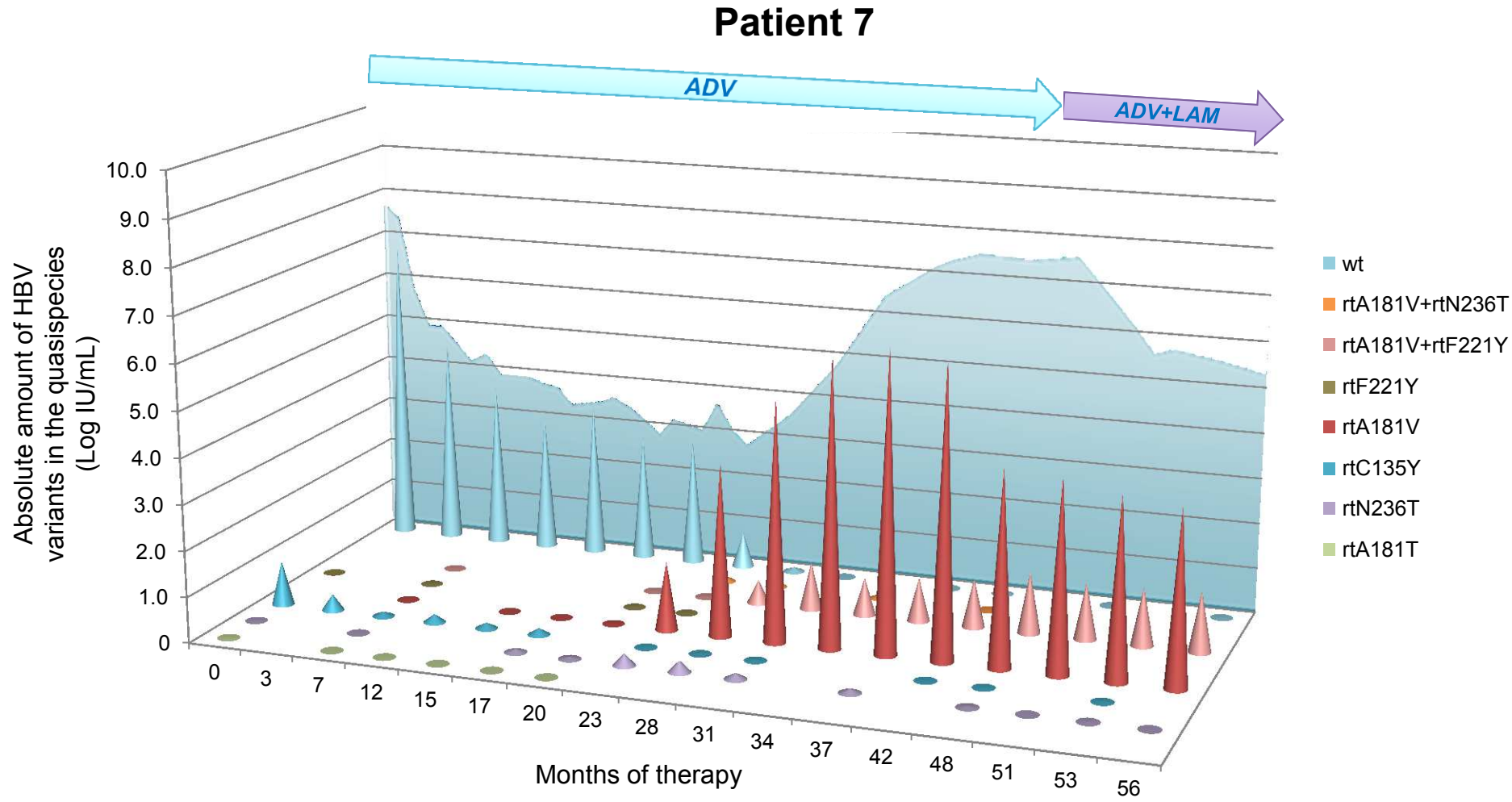


Figure 2F



ARTICLE 3

Hayer J*, Rodriguez C*, Germanidis G, Deléage G, Zoulim F,
Pawlotsky JM, Combet C

***Hepatitis B virus RNase H sequence variability investigated by
pyrosequencing, sequence analysis and molecular modeling***

Manuscrit en préparation

*participation égale

1 **TITLE**

2 Hepatitis B virus RNase H structure revisited through ultra-deep pyrosequencing analysis
3 and molecular modeling of its genetic variability

4
5 **RUNNING TITLE**

6 HBV RNase H variability

7 **AUTHORS AND AFFILIATIONS**

8 Juliette Hayer^{1*}, Christophe Rodriguez^{2,3*}, Georgios Germanidis⁴, Gilbert Deléage¹,
9 Fabien Zoulim⁵, Jean-Michel Pawlotsky^{2,3}, Christophe Combet^{1,#}

10
11 ¹Bases Moléculaires et Structurales des Systèmes Infectieux, UMR 5086, CNRS,
12 University of Lyon 1, IBCP 7 passage du vercors, 69367, Lyon cedex 07, France.

13 ²National Reference Center for Viral Hepatitis B, C and Delta, Department of Virology,
14 Hôpital Henri Mondor, Université Paris-Est, Créteil, France.

15 ³INSERM U955, Créteil, France.

16 ⁴University of Thessaloniki, Greece.

17 ⁵Viral Hepatitis Research Laboratory, U1052, INSERM; Lyon University Hospital, 151,
18 cours Albert Thomas, 69003 Lyon; Hepatology Department, Hospices Civils de Lyon,
19 France.

20 # :
21

21 Dr Christophe Combet

22 BMSSI - IBCP

23 7 passage du Vercors, 69367, Lyon Cedex 07, France.

24 Phone: 04 37 65 29 47 – Fax: 04 72 72 26 01

25 E-mail: c.combet@ibcp.fr

26 **ABSTRACT WORD COUNT**

27 XXX words

28 **TEXT WORD COUNT**

29 X,XXX words

30

34 INTRODUCTION

35

36 Hepatitis B virus (HBV) infection is among the top ten most frequent viral
37 infections worldwide, with approximately 240 million chronic HBV carriers facing a
38 dramatically high risk of life-threatening complications, such as liver cirrhosis or
39 hepatocellular carcinoma (HCC) [Wei 2010]. HBV infection is nowadays the first cause
40 of primary liver cancer worldwide, and mortality due to chronic hepatitis B exceeds one
41 million per year. Prevention of these complications can be achieved if HBV replication is
42 efficiently controlled in the long-term. Two types of antiviral treatments can be used to
43 achieve this goal: a finite treatment with injectable pegylated interferon alpha, or lifelong
44 oral administration of nucleoside/nucleotide analogues that target the HBV DNA
45 polymerase [EASL Clinical Practice Guidelines, J Hepatol 2012].

46 The HBV DNA polymerase is a multifunctional protein that consists of four
47 domains, including a terminal protein (TP) primase, a reverse transcriptase (RT), a
48 ribonuclease H (RNase H), and a variable spacer domain between the TP and RT
49 domains [Nguyen 2008]. Nucleoside/nucleotide analogues are highly effective in
50 controlling HBV replication by specifically suppressing the activity of the HBV RT
51 [Locarnini 2006]. Treatment of chronic hepatitis B with RT inhibitors induces a rapid
52 decrease of viremia. However, long-term treatments with first-generation drugs such as
53 lamivudine or adefovir with a low genetic barrier to resistance have been associated with
54 viral breakthroughs due to the selection of drug-resistant HBV variants [Nguyen 2008].
55 The current first-line HBV drugs, entecavir and tenofovir, are potent and have a higher
56 barrier to resistance. Thus, selection of resistant HBV strains is rare over the first 5 years

57 of therapy. However, resistance may occur later in the course of therapy in patients likely
58 to use these drugs lifelong. In this context, it is safe to develop new generations of highly
59 effective anti-HBV drugs with targets other than the RT.

60 The HBV RNase H represents a potential therapeutic target, because it has
61 enzymatic activity which is essential to the HBV lifecycle. However, targeted drug
62 discovery requires extensive knowledge of RNase H structure and variability. Attempts
63 to experimentally determine the 3D structure of the HBV DNA polymerase or its RNase
64 H domain have so far been unsuccessful. However, several RNase H structures from
65 various organisms have been reported [Katayanagi 1993;Lim 2006]. An important
66 structural feature of RNases H is the existence of an α -helix, called C-helix, involved in
67 substrate recognition and catalytic activity [Champoux 2009], which is present in certain
68 organisms (*e.g.* *E. coli*, Murine Moloney Leukemia Virus), but absent in others (*e.g.*
69 HIV-1). The availability of these models and the notion that RNase H is conserved across
70 species [Tadokoro 2009] open the way to molecular modeling of the enzyme domains
71 based on sequence-template sequence alignments.

72 A molecular model of the HBV RNase H has been published using *E. coli* RNase
73 H structure as the template [Potenza 2007]. In this work, we generated a new molecular
74 model of HBV RNase H from the *E. coli* RNase H structure by using: (i) sequences from
75 patients infected with different HBV genotypes generated by population sequencing
76 available in public databases; and (ii) quasispecies sequences from a homogeneous
77 population of 73 never treated patients infected with HBV genotype D generated by
78 means of a highly sensitive method based on ultra-deep pyrosequencing (UDPS).

79

80 MATERIALS AND METHODS

81

82 RNase H sequences from public databases

83

84 The HBV polymerase dataset were retrieved from the HBVdb database
85 (<http://hbvdb.ibcp.fr>), which contains a collection of computer-annotated HBV entries
86 extracted from the European Nucleotide Archive (ENA) [Hayer 2012, in prep] [Leinonen
87 2011]. The residue repertoire and residue frequencies were retrieved for each HBV
88 genotype and combined to generate an all-genotype dataset. These data were used to
89 identify conserved and variable RNase H positions. Only residues present in at least 1%
90 of the sequences for each column of the alignment were considered.

91

92 RNase H quasispecies sequences generated by ultra-deep pyrosequencing

93

94 A representative subset of 80 patients (52 males, 28 females, mean age 52 ± 13)
95 was randomly selected from a cohort of 579 treatment-naïve HBe-Ag negative patients
96 chronically infected with HBV genotype D for UDPS sequence analysis of the RNase H
97 domain. HBV DNA was extracted from 200-500 μ L of serum with the QIAamp DNA
98 blood kit (Qiagen GmbH, Hilden, Germany), according to the manufacturer's
99 instructions. A 697-bp fragment encompassing the full-length RNase H domain of HBV
100 DNA polymerase was PCR-amplified with primers RNaseH-F (nucleotide positions
101 1031-1058, 5'-CACAATGTGGATATCCTGCCTTAATGC-3') and RNase-H-R
102 (nucleotide positions 1702-1728, 5'-CCTACTTCAAAGACTGTGTGTTTAAG-3') in

103 the presence of AmpliTaqGold (Applied Biosystems, Carlsbad, California), according to
104 the manufacturer's instructions. The following amplification program was used: initial
105 denaturation at 95°C for 10 min; 40 PCR cycles at 95°C for 30 s, 58°C for 30 s, and 72°C
106 for 40 s; final elongation at 72°C for 7 min. PCR products were then purified using
107 Amicon-Ultra columns (Millipore, Billerica, Massachusetts). A second "nested" PCR
108 round was performed with internal primers UDP3719-F (nucleotide positions 1134-1154,
109 5'-CATGAACCTTTACCCCGTTG-3') and UDPS3719-R (nucleotide positions 1661-
110 1681, 5'-GACTCTTGGACTCCCAGCAA-3'), modified to introduce a GS FLX bead
111 adaptor and a specific identity tag (multiplex identifier, MID). Amplicons containing the
112 bead adaptor and MID were then pyrosequenced as previously described (Rodriguez et al,
113 2012). Briefly, the amplicons were purified in Nucleofast[®] 96 PCR plates (Clontech,
114 Mountain View, California), quantified with the Quant-iT[™] PicoGreen[®] dsDNA kit
115 (Invitrogen, Carlsbad, California), and emulsion-PCR amplified by means of emPCR kit
116 (454 Life Sciences, Roche Diagnostics Corp., Branford, Connecticut). The beads were
117 counted with a Beckman Coulter Z1 particle counter (Beckman Coulter, Brea, California)
118 and deposited in a GS FLX Titanium PicoTiterPlate (454 Life Sciences). The
119 pyrosequencing reaction was performed in both directions with the GS FLX Titanium
120 sequencing kit on an FLX Genome Sequencer (454 Life Sciences).

121 Our recently described in-house software package, Pyropack[®], was used to
122 analyze UDPS data (Rodriguez et al. 2012). Briefly, the sequences were ascribed to each
123 patient by means of their identity tags and PyroClass[®]. Then, the sequences were quality
124 filtered, aligned, and substitutions were analyzed statistically by means of PyroMute[®], in

125 order to eliminate low quality data and errors due to the procedure and keep only relevant
126 sequences for final analysis. Finally, a map of RNase H genetic variability was generated.

127

128 **Molecular Modeling**

129

130 Alignments of HBV RNase H (UniProtKB:Swiss-Prot: C7DSI5 residues 680-832)
131 and *E.coli* RNase H sequences (PDB:1RDD) were computed by means of Clustal W
132 (version 1.8). The alignments were manually optimized by combining data from two
133 alignments, including: (i) a structural alignment of MMLV (PDB:2HB5), *E. coli*
134 (PDB:2RN2), HIV-1 (PDB:1HRH) and *B. halodurans* (PDB:1ZBF) RNases H published
135 by Lim *et al.* [Lim 2006], and (ii) a pairwise alignment of HBV and *E. coli* RNases H
136 used to model the HBV RNase H published by Potenza *et al.* [Potenza 2007]. Manual
137 optimizations were guided by the observed secondary structures, deduced from the three-
138 dimensional structure using the DSSP algorithm [Kabsch 1983], the predicted secondary
139 structures obtained by combining PHD [Rost 1993], DSC [King 1996] and SOPMA
140 [Geourjon 1995] algorithms with the NPS@ server [Combet 2000], and visual inspection
141 of the template structures.

142 The molecular model of RNase H domain was computed by means of the Geno3D
143 software [Combet 2002], that uses query-template alignment and PDB:1RDD structure as
144 input files. Geno3D computes the models according to a protocol similar to that used for
145 structure resolution by nuclear magnetic resonance. It uses distance geometry, simulated
146 annealing, molecular dynamics and energy minimization algorithms as implemented in
147 the CNS software (version 1.2). Geno3D measures spatial restraints (dihedral angles and

148 inter-atomic distances) used in the calculation from the template structure for similar
149 residues, according to the query-template alignment. Ten models were computed for the
150 Geno3D run. The stereochemistry of each model was checked with PROCHECK
151 software [Laskowski 1996]. The most accurate model of each run was selected according
152 to energy value, stereochemistry, restraint violations and the root mean square deviation
153 (RMSD) to the template, for further analysis.

154

155 **Mapping of variable positions on the RNase H model**

156

157 Variable positions identified from the HBVdb and UDPS analyses were mapped
158 on the RNase H model using coloration by residues, as described in Pymol [REF].

159

160 **RESULTS**

161

162 **RNase H genetic variability in population sequencing-based sequences** 163 **retrieved from public databases**

164

165 One hundred and ninety-three full-length RNase H sequences were retrieved from
166 the databases, including sequences of the 8 HBV genotypes (676 from genotype A, 121
167 for B, 963 for C, 193 for D, 223 for E, 97 for F, 29 for G, and 24 for H). The RNase H
168 domain contains 153 residues. When combining available data from the eight genotypes,
169 63 positions were found to display sequence differences. Among them, 32 were found in
170 one genotype only, including 12 that were found only in genotype D. When each

171 genotype was assessed individually, the number of amino acid positions with different
172 residues observed was 9 for genotype A, 22 for genotype B, 17 for genotype C, 33 for
173 genotype D, 23 for genotype E, 20 for genotype F, 14 for genotype G and 12 for
174 genotype H.

175 Type 1 RNases H typically display four catalytic residues (D-E-D-D) [Tadokoro
176 2009]. In this analysis, the first three catalytic residues were conserved among all HBV
177 genotypes at residues UniProtKB:Swiss-Prot:C7DSI5:D689, E718 and D737 with
178 frequencies of 0.999. Three possibilities exist for the fourth catalytic residue, including
179 residue UniProtKB:Swiss-Prot:C7DSI5:D777, D807 or D817. However, none of them
180 fulfilled the criteria for being the fourth catalytic residue in our analysis. Indeed, D777 is
181 highly conserved across genotypes (frequency of 0.999) and mutating it dramatically
182 affects the enzyme activity of RNase H [Radziwill 1990; Chen 1992;Chen 1994;Chen
183 1996;Chang 1990]. Nevertheless, our analysis showed that this residue is too far from the
184 three other catalytic residues (XX Å) in the molecular models, a result in keeping with
185 Potenza *et al.* [Potenza 2007]. These authors aligned the fourth catalytic residue of *E. coli*
186 RNase H with residue UniProtKB:Swiss-Prot:C7DSI5:D807 and suggested it as the
187 fourth catalytic residue. Nevertheless, residue UniProtKB:Swiss-Prot:C7DSI5:D807 is
188 not conserved across genotypes (frequency of 0.416). Indeed, as shown in Figure 1A, the
189 most frequent residue at position 807 is a valine for HBV genotypes A, C, E and G. Our
190 alignment suggested that residue UniProtKB:Swiss-Prot:C7DSI5:D817 is well conserved
191 across genotypes (frequency of 0.734) and could act as the fourth residue of the catalytic
192 unit. However, as shown in Figure 1B, this residue is not conserved in HBV genotypes A
193 and H.

194

195 **RNase H genetic variability in patients chronically infected with HBV**

196 **genotype D analyzed by UDPS**

197

198 HBV DNA could be amplified from the serum of 73 of the 80 patients with HBe

199 antigen-negative chronic hepatitis B related to HBV genotype D selected for this study.

200 Nearly 960,000 sequences were generated from these samples, i.e. on average

201 $13,130 \pm 3220$ sequences per patient. Their average length was 302 bp (range 221-369 bp)

202 and 63.8% of the full set of sequences was considered of excellent quality. After

203 eliminating unreliable or too short sequences by means of PyroMute[®], more than 900,000

204 sequences were available for subsequent analysis. The consensus amino acid sequence

205 was built and used as a reference to map genetic variability at each position in the

206 analyzed population (Figure 2A). In this set of data, the first two putative catalytic

207 residues, D689 and E718, were conserved in 99.9% and 96.2% of sequences,

208 respectively. The third putative catalytic residue, D737, was less conserved, with 12

209 patients harboring a D737A substitution that represented 15% to 50% of their viral

210 quasispecies. Seventy of the 73 patients had a D as most frequent amino acid at position

211 807 (Figure 2B), while the remaining three patients had an aliphatic substitution (D807A

212 or D807V) at this position in their major viral population. One of them harbored an

213 additional D817V substitution in his major viral population, ruling out a possible rescue

214 of the fourth catalytic amino acid. This patient's viral level was not substantially lower

215 than in the other patients (3.5 vs 4.3 ± 1.8 Log IU/mL, respectively). Finally, only 51 of the

216 77 patients had a dominant D817 population, while 20 patients had a D817V substitution
217 and 2 patients had a D817A substitution (Figure 2C).

218 The global analysis shown in Figure 2A allowed us to identify conserved and
219 variable domains in the RNase H sequence. Four conserved domains (at positions 683-
220 698, 713-730, 755-785, and 790-816, respectively) and one highly variable domain (at
221 positions 731-754) were identified. Typically, type 1 RNases H contain 4 conserved
222 domains spanning the four catalytic residue positions, and one additional functional
223 domain called “basic protrusion domain” which attracts the heteroduplex to the enzyme.
224 In our mapping of RNase H variability, the first two conserved domains contained the first
225 two catalytic residues D689 and E718, whereas the last one contained D807, a possible
226 fourth catalytic residue. However, if the fourth catalytic residue were D817, this residue
227 would be located in a variable region. The protrusion domain, at positions 755-785, was
228 well conserved. The highly variable domain spanned the region of RNase H interacting
229 with RNA/DNA heteroduplexes.

230

231 **Molecular Modeling**

232

233 When comparing HBV and *E. coli* RNase H sequences, we found only 14.6%
234 identity, but 39.3% similarity and 27.0% gap, confirming the homology. The 24-residue
235 region UniProtKB:Swiss-Prot:C7DSI5:F748-D771 was suggested to correspond to the C-
236 helix present in MMLV and *E. coli* RNases H. In the latter case, a large number of basic
237 residues cluster within the C-helix and interact with the nucleic acid. In the case of HBV

238 RNase H, there was only one arginine at position R762 that could interact with the
239 nucleic acid.

240 The set of 10 models described in the Methods section was computed with XXX
241 distance and XXX dihedral restraints involving XXX residues. The agreement between
242 the models was good, with an RMSD of 1.54 Å over 81 Ca atoms. Model number 1 was
243 selected according to energy, stereochemistry, restraint violations and RMSD to the
244 template for further analyses. The RMSD of model 1 with the PDB:1RDD template
245 structure was 1.49 Å over 81 Ca atoms. While this value indicated good agreement
246 between the model and the template structure, there were local differences in close
247 vicinity to the putative C-helix. Superimposition of the template structure and selected
248 model is presented in Figure 3A.

249

250 **Mapping of variable positions on RNase H model**

251

252 Figures 3B to 3D show a 3D representation of HBV RNase H facing the region of
253 the enzyme in contact with the DNA-RNA heteroduplex. Figure 3B summarizes the
254 positions that were found variable across genotypes in the analysis of sequences retrieved
255 from HBVdb. Variable positions that differed from the majority in more than one
256 genotype were located in the region that does not interact with the DNA/RNA
257 heteroduplex. Figures 3C and 3D summarize variable positions within genotype D, in the
258 193 genotype D sequences generated by population sequencing and retrieved from public
259 databases and in the quasispecies sequences from 73 patients generated by means of
260 UDPS, respectively. As shown in Figure 1C, most variable positions across patients were

261 located in the region of RNase H that does not interact with the DNA/RNA heteroduplex,
262 in sequences that represented the dominant viral population in the quasispecies. In
263 contrast, when UDPS data were analyzed, numerous amino acid changes were found in
264 the region interacting with the heteroduplex, reflecting variability of minority variants
265 present in the viral quasispecies. Variations were found in minor variants that involved
266 the second and third putative catalytic residues, which could correspond to defective
267 variants detected by UDPS. A Venn diagram (Supplementary Figure) summarizes the
268 number of positions displaying amino acid changes in the two datasets of genotype D
269 sequences, as well as in genotypes other than D. Six amino acids positions displayed
270 changes for HBV genotype D only in both HBVdb and sequences analyzed by UDPS; 15
271 positions showed changes in the three datasets whereas 18 displayed changes in
272 genotypes other than D only. The depth of UDPS analysis allowed us to identify 32
273 additional positions with mutations.

274

275 **DISCUSSION**

276

277 **ACKNOWLEDGMENTS**

278

279 JH received a doctoral fellowship from the Ministère de l'Enseignement Supérieur
280 et de la Recherche (contract 0012). CR received a doctoral fellowship from the Fondation
281 pour la Recherche Médicale (FRM). This work was supported by the Finovi foundation
282 (contract 051274 2010-2011), FRM and the National Agency for Research on AIDS and
283 Viral Hepatitis (ANRS). Sequence data analyses and molecular modeling were performed
284 on the PRABI computing platform (Contract GIS-IBiSA 2009-2001).

285 FIGURE LEGENDS

286

287 **Figure 1.** (A) Histogram of amino acids frequencies at position UniProtKB:Swiss-
288 Prot:Q05486:D807 for the eight genotypes. (B) Histogram of amino acids frequencies at
289 position UniProtKB:Swiss-Prot:Q05486:D817 for the eight genotypes.

290

291 **Figure 2.** (A) Mapping of HBV RNase H genetic variability established from UDPS
292 sequence analysis of full-length RNase H from 73 patients with chronic hepatitis B due to
293 HBV genotype D. The figure shows the percent variability observed within patients
294 quasispecies at each amino acid position of the RNase H, showing both highly variable
295 and conserved domains. (B) Distribution of amino acids at position 807 based on UDPS
296 sequence analysis in the 73 patients with HBV genotype D infection. (C) Distribution of
297 amino acids at position 817 based on UDPS sequence analysis in the 73 patients with
298 HBV genotype D infection.

299

300 **Figure 3.** (A) Superimposition of *E. coli* RNase H structure (PDB:1RDD), in gray, and
301 the selected model of HBV RNase H, in purple (cartoon representation). The C-helix of
302 *E. coli* is colored in black and the four catalytic residues are shown in sticks and colored
303 in yellow. The putative HBV C-helix is colored in orange and the three conserved
304 catalytic residues are pink in sticks. (B) Mapping of positions showing mutations in all
305 genotypes, with color gradients going from green to red: positions that do not mutate are
306 colored in dark green, a position with changes in 4 genotypes is in orange and a position
307 with changes in 8 genotypes is in dark red. (C) Mapping of mutated positions (in red)

308 established with HBV genotype D sequences retrieved from HBVdb public sequences.
309 (D) Mapping of mutated positions (in red) established with UDPS data from 73 patients
310 infected with HBV genotype D. (B)(C) and (D) are in sphere representation.

Figure 1

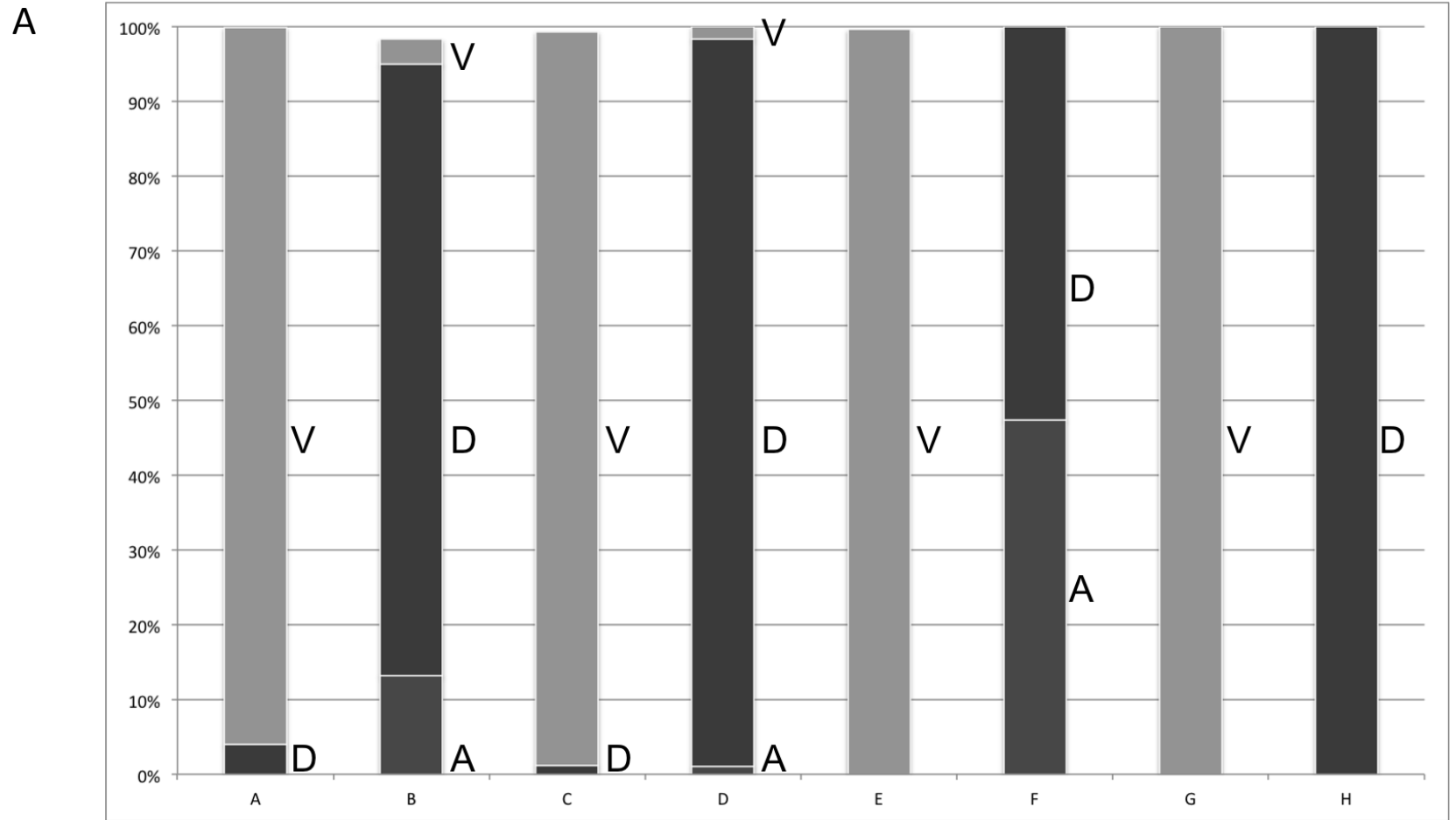


Figure 1

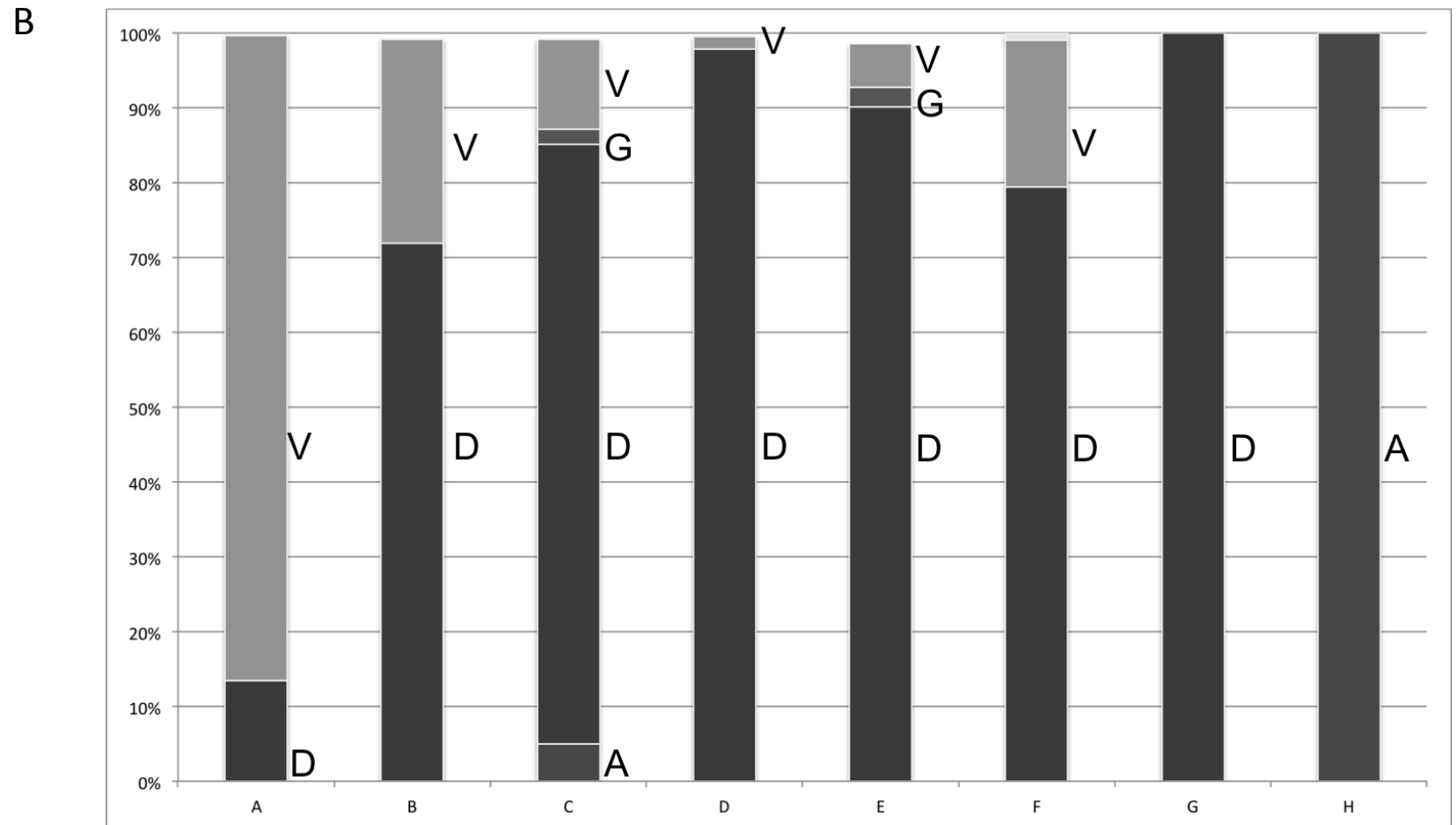


Figure 2

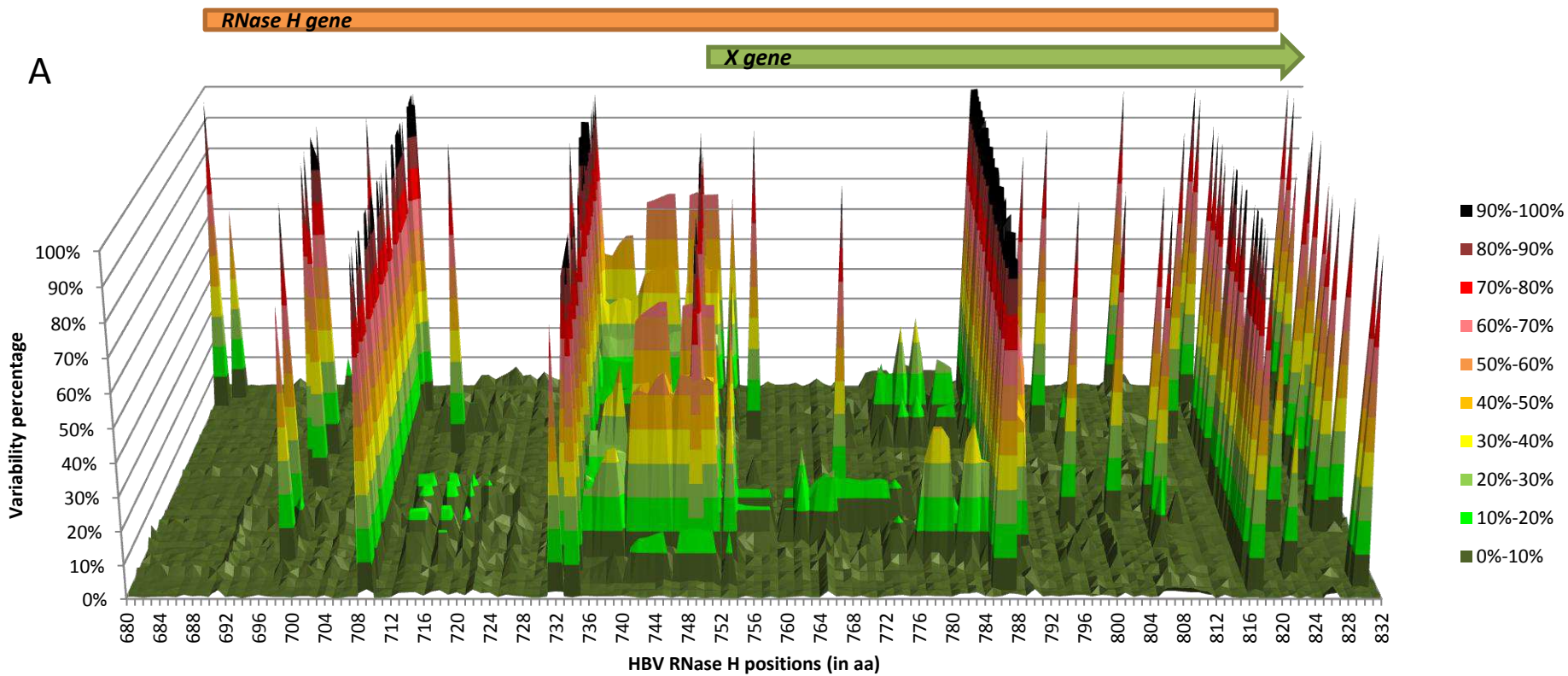


Figure 2

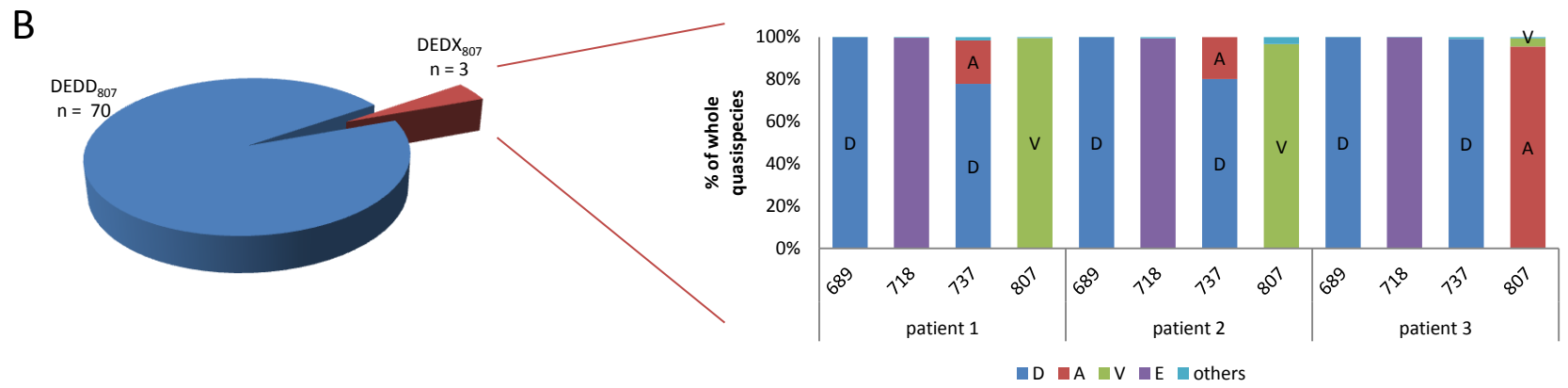
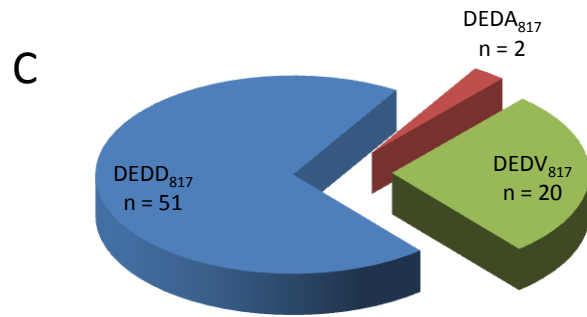
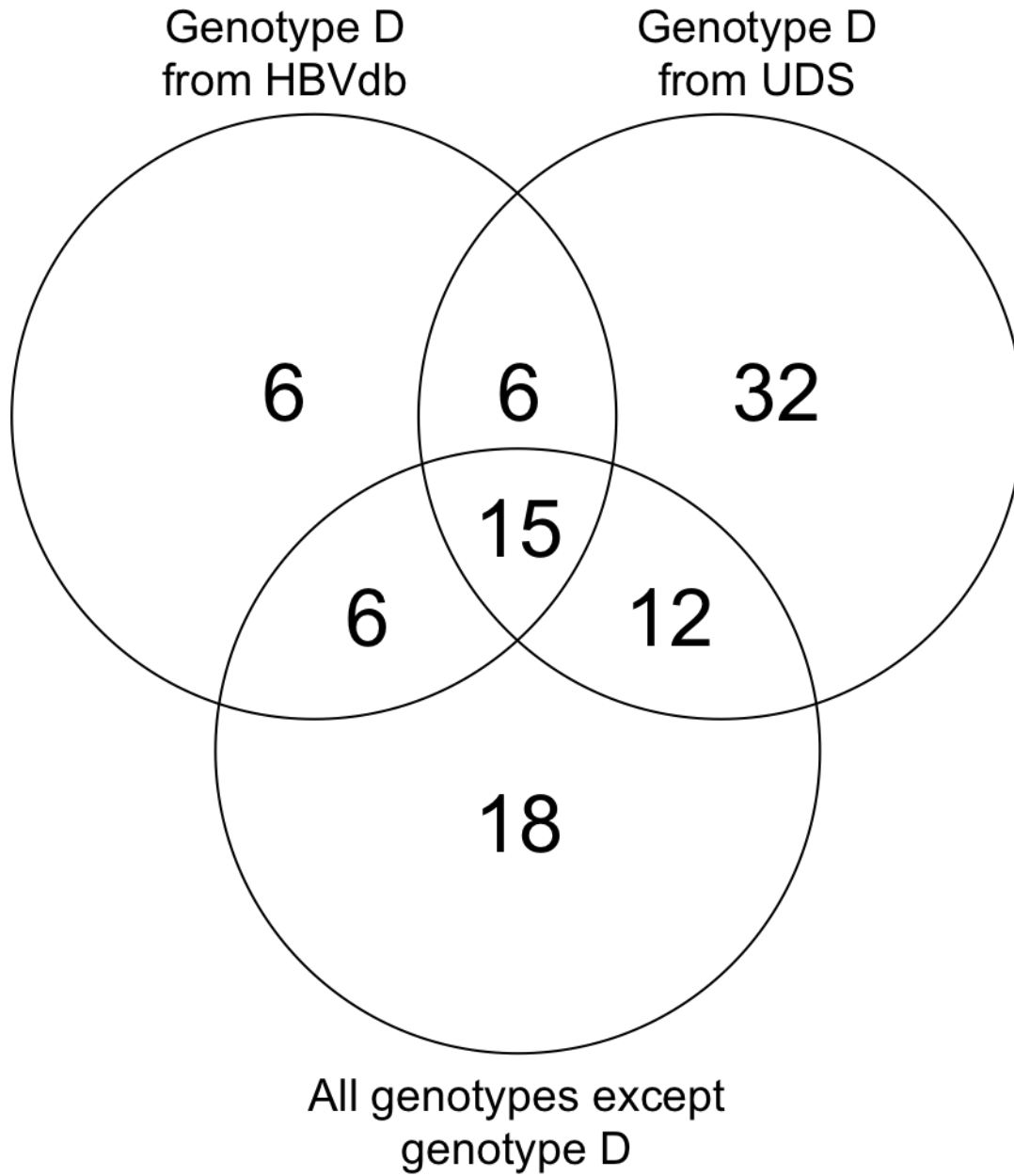


Figure 2



Supplemental
Figure



ARTICLE 4

Rodriguez C, Soulié C, Marcelin AG, Calvez V, Descamps D, Charpentier C, Flandre P, Recordon-Pinson P, Bellecave P, Harrigan R, Pawlotsky JM, Masquelier B, ANRS AC11 Study Group

HIV coreceptor usage prediction by ultra-deep pyrosequencing to predict maraviroc response

Soumis

HIV-1 coreceptor usage prediction by ultra-deep pyrosequencing to predict maraviroc response

Christophe RODRIGUEZ,^{1,2} Cathia SOULIE,³ Anne-Geneviève MARCELIN,³ Vincent CALVEZ,³ Diane DESCAMPS,⁴ Charlotte CHARPENTIER,⁴ Philippe FLANDRE,⁵ Patricia RECORDON-PINSON,⁶ Pantxika BELLECAVE,⁶ Richard HARRIGAN,⁷ Jean-Michel PAWLOTSKY,^{1,2} Bernard MASQUELIER⁶ and the ANRS AC11 Study Group

¹National Reference Center for Viral Hepatitis B, C and delta, Department of Virology; Henri Mondor Hospital, University of Paris-Est, Créteil, France; ²INSERM U955, Créteil, France; ³Department of Virology, Pitié-Salpêtrière Hospital, Paris, France; ⁴Department of Virology, Bichat-Claude Bernard Hospital, HUPNVS, Paris, France; ⁵INSERM U943 UPMC, UMR-S 943 Paris, France; ⁶Laboratoire de Virologie, CHU de Bordeaux and UMR5234, Université de Bordeaux, Bordeaux, France; ⁷British Columbia Centre for Excellence in HIV/AIDS, Vancouver, Canada

Short title: HIV-1 coreceptor usage prediction by UDPS

Number of words: 1775

Corresponding author : Dr Christophe RODRIGUEZ, Department of Virology, Hôpital Henri Mondor, 51 avenue du Maréchal de Lattre de Tassigny, 94010 Créteil, France. Tel: +33-1-4981-2833; Fax: +33-1-4981-2839. E-mail: christophe.rodriguez@hmn.aphp.fr

ABSTRACT

Objectives. Maraviroc (MVC) is an HIV entry inhibitor that alters the conformation of CCR5 and is poorly efficient in patients infected by viruses that use CXCR4 as an entry coreceptor. The goal of this study was to assess the capacity of ultra-deep pyrosequencing (UDPS) and different data analysis approaches to assess HIV tropism and predict the therapeutic outcome on MVC treatment.

Methods. One hundred and thirteen patients treated by HAART with detectable HIV-1 RNA were treated by MVC. The response to MVC was assessed at month 1. The sequence of the HIV V3 loop was determined at baseline by UDPS and UDPS data were analyzed by different methods including our in-house software PyroTrop[®] and the geno2pheno algorithm

Results. Using consensus sequences or the full set of UDPS sequences, Pyrotrop[®] software was more sensitive than methods based on the geno2pheno algorithm; specificity was better with methods using the full set of UDPS data. Overall, Pyrotrop[®] analysis of the full set of UDPS sequences offered the best prediction of MVC response. The H34Y/S substitution had the strongest individual predictive value on MVC response, i.e. stronger than substitutions at positions 11 or 25 used in the 11/25 rules to predict HIV coreceptor usage.

Conclusions. UDPS with Pyrotrop[®] software is a powerful tool that can be used with confidence to predict MVC response in HIV-1-infected patients. Improvement of the predictive value of the algorithm used is possible and our results suggest that adding the H34S/Y substitution would substantially improve the performance of the 11/25/charge rule.

KEY-WORDS

Human immunodeficiency virus; coreceptor usage, maraviroc, ultra-deep pyrosequencing

Human immunodeficiency virus (HIV) entry starts with the attachment of the viral envelope glycoprotein gp120 to the CD4-positive T-cell receptor and to either of two coreceptors, coreceptor 5 (CCR5) or coreceptor 4 (CXCR4) [1-4]. Maraviroc (MVC) is an HIV entry inhibitor designed to prevent infection of CD4-positive T-cells by altering CCR5 conformation [5, 6]. As a result, this therapy is poorly effective in patients infected by viruses that use CXCR4 as an entry coreceptor. Thus, characterization of HIV tropism is important prior to deciding to use MVC in HIV therapy [7, 8]. Classically, assessment of HIV tropism has been based on two different techniques. The first approach uses phenotypic assays, such as for instance the Trofile assay (Monogram Biosciences) [9]. Because of the need for cell coculture and *in vitro* infection, this method is challenging in the clinical setting [10]. The second, genotypic approach is based on sequence analysis of the HIV V3 loop, a region involved in the interaction with the coreceptor that drives viral tropism.

HIV has a quasispecies distribution in infected patients, characterized by the coexistence of closely related but distinct viral populations, including major and minor viral populations. If CXCR4 viruses are present as a minor viral population in the presence of a majority of CCR5 viruses, they can be selected by MVC, expand and become predominant, ultimately leading to treatment failure. In these circumstances, MVC treatment is not indicated. Previous works have established that the presence of more than 2% of CXCR4 viruses in patients' quasispecies is highly predictive of MVC failure [11]. However, such sensitivity for the detection of minority viral populations cannot be obtained with population sequencing, also called direct sequencing, the most widely used sequencing approach in the clinical setting. Cloning and sequencing is not sensitive enough, except if a very large number of clones are generated, which is hardly feasible in clinical practice. Therefore, more sensitive techniques are required to genotypically

assess HIV tropism prior to initiating MVC therapy and obtain high predictive values on the outcome of therapy.

Next-generation sequencing methods, such as ultra-deep pyrosequencing (UDPS), have recently been developed to increase sequencing capacity while generating clonal sequences. The main challenge with this technology is the very large amount of sequences generated that require complex analyses to be transformed into clinically meaningful information. Three steps are required: (1) the UDPS data must be formatted into usable files before starting the analysis, including ascribing each generated sequence to the corresponding patient's sample; (2) quality analysis of UDPS sequences must be performed in order to eliminate unreliable information; this step is based on the use of successive filters including, for example, the PHRED sequence quality score [12], the length of the sequences (non-full HIV V3 loop sequences eliminated), or the quality of alignment; and (3) bioinformatics algorithms must be used to determine whether the analyzed sample should be considered CCR5 or CXCR4; this analysis is generally based on rules using substitutions at positions 11 and 25 and the global charge of the V3 loop [13-18] or on comparisons with phenotypic test databases.

In this work, we used UDPS and different analytical approaches to assess HIV tropism and the capacity of these pre-treatment determinations to predict the therapeutic outcome on MVC treatment.

PATIENTS AND METHODS

One hundred and thirteen patients treated by highly active antiretroviral therapy (HAART) with detectable HIV-1 RNA were enrolled in this study and treated by MVC combined with an optimized background therapy [19]. Patient's sera were collected at day 0 (D0) and month 1 (M1) of MVC therapy. The patients were considered as

responders to MVC if their HIV RNA level (assessed by means of COBAS AmpliPrep/COBAS TaqMan HIV-1 Test, Roche Molecular Systems, Pleasanton, California) had decreased by at least 0.5, 1.0 or 2.0 log at M1 compared with D0.

The sequence of the HIV V3 loop at D0 was determined by means of UDPS in the 113 patients, as previously described [11]. Briefly, three independent one-step nested PCR amplifications were performed using tagged primers. Next, amplicons were quantified, fixed onto microbeads, subjected to emulsion PCR and the beads were loaded onto picotiter plates for forward and reverse pyrosequencing by means of the GS-FLX Titanium Kit in the Genome Sequencer-FLX (454 Life Sciences, Roche Diagnostics Corp., Branford, Connecticut).

The generated sequences were first classified by means of Pyroclass[®], an in-house software that identifies each sequence by means of its sample tag, removes the tag, reverse complements reverse sequences, and ascribes the sequence to its sample file. Next, four methods were used to analyze and report the UDPS data. Their intrinsic performance and their ability to predict the response to MVC were compared. *Method 1*: consensus sequences were generated from UDPS data at D0, used as a surrogate of population sequencing, and interpreted with the geno2pheno[coreceptor] algorithm (Max-Planck-Institute), based on the German recommendations; *Method 2*: the same consensus sequences were interpreted by means of the 11/25/charge rules using a module included into PyroTrop[®], an in-house software developed for this application; *Method 3*: the full set of UDPS sequences was interpreted with geno2pheno[454] False Positive Rate 2.5% (Max-Planck-Institute), an adaptation of the geno2pheno algorithm to multiple sequence files, such as those obtained with UDPS; *Method 4*: the full set of UDPS sequences was interpreted with our in-house software PyroTrop[®].

PyroTrop[®] uses three quality filters to eliminate unreliable sequences, including: sequences with a PHRED quality score <20 [12]; sequences not spanning the full-length

HIV V3 loop; and sequences too divergent from HIV strain HXB2 reference sequence (GenBank accession number K03455). Then, PyroTrop[®] ascribes each valid V3 sequence to either CCR5 or CXCR4 by means of the 11/25/charge algorithm and provides the respective percentages of these two populations within the patient's quasispecies (2% of HIV CXCR4 variants was set as the detection threshold) [11]. Additional modules are included in PyroTrop[®] to assess the frequency of each amino acid at each position; additional quality filters based on the Poisson's law eliminate substitutions with a frequency compatible with a technical error.

The Yule's Q score was used to establish the strength of the prediction of each method on the patient's response to MVC, as follows: 0.01-0.09 = no association; 0.10-0.29 = weak association; 0.30-0.49 = moderate association; 0.50-0.69 = strong association, 0.70-1.00 = very strong association. In addition, a Mann-Whitney test of genetic variability was used to differentiate MVC responders and nonresponders based on the amino acid sequence of their V3 loop.

In order to assess the predictive value of V3 loop quasispecies sequences on the response to MVC, Receiver Operating Characteristic (ROC) curves were plotted for each mutation and the corresponding area under the curves (AUROC) was compared with an AUROC of 0.5, considered to be non-predictive, using bootstrap methods [20-22], by means of package pROC (v1.5) [23] using R language software by means of RGui(64-bit) (v 2.14.1) [24].

RESULTS:

HIV RNA could not be amplified in 11 patients at D0. Pyrosequencing data of good enough quality at D0 were obtained in the remaining 102 patients. Eighty-four of them also had HIV RNA levels available at M1 to assess prediction of MVC response. UDPS generated a total of 626,874 sequences, among which 539,096 sequences were valid for

geno2pheno[454] analysis (mean 5285 per sample; rejection rate: 14.3%) and 461,044 were valid for PyroTrop[®] analysis (mean 4520 per sample; rejection rate: 26.6%).

Concordance between the four methods is shown in Table 1 for the 102 patients with available UDPS sequences at D0. Best concordance was obtained with the two methods using Pyrotrop[®] software.

Table 2 shows the performance of the four methods used to assess HIV tropism in predicting different magnitudes of response to MVC therapy (0.5, 1.0 and 2.0 log decrease, respectively) in the 84 patients with available HIV RNA at D0 and M1. Both methods using Pyrotrop[®] software were more sensitive than methods based on the geno2pheno algorithm. Specificity was better with methods using the full set of UDPS data than with methods using the UDPS consensus sequence, used as a surrogate of population sequencing. The positive predictive values were of the same order with the four methods, whereas the best negative predictive value was obtained with Pyrotrop[®] analysis of the full set of UDPS sequences (Table 2).

Using the Yule's Q method, prediction of an HIV RNA decline of 1.0 log at M1 upon MVC was weak with consensus/g2p ($Q=0.11$), moderate with consensus/Pyrotrop[®] and UDPS/g2p ($Q = 0.59$ and 0.51 , respectively), and very strong with UDPS/ Pyrotrop[®] ($Q = 0.71$). Overall, Pyrotrop[®] analysis of the full set of UDPS sequences offered the best prediction of subsequent MVC response (Table 2).

The ability of each observed amino acid at each position in the V3 loop to predict MVC response was assessed by plotting the corresponding individual ROC curves (data not shown). The H34Y/S substitution had the strongest individual predictive value ($p=0.011$). Indeed, H34Y/S was present in 23% of MVC responders, in whom it represented on average 79.6% of the viral population. Conversely, it was not detectable at significant levels in any of the non-responders to MVC ($p=0.01$). The association between MVC response and substitutions at positions 11 and 25, used in the 11/25 rules

to predict HIV coreceptor usage, was not significant when these substitutions were considered individually: $p=0.13$ for S11R/K, and $p=0.08$ for D25R/K.

DISCUSSION

In this study, we assessed the performance of methods based on consensus or UDPS sequence analysis, followed by geno2pheno or Pyrotrop[®] data analysis, to predict the response to MVC in HIV-1-infected patients under HAART therapy. The best predictive value was achieved when UDPS was combined with Pyrotrop[®] analysis, and prediction was better with Pyrotrop[®] whatever sequencing approach used. This improvement in the predictive value is likely due to the sequence quality filters introduced in Pyrotrop[®] for UDPS sequence analysis.

In addition, our analysis suggests that predictions based on the 11/25/charge rule can be substantially improved by adding other parameters to the algorithm. Indeed, we found that the H34S/Y substitutions were significantly associated with CCR5 genotropism. This association was even stronger than the individual association of substitutions at positions 11 or 25 and is not fully reflected by charge changes. Thus, H34S/Y deserves to be included into a new algorithm, and this approach will require prospective confirmation in other cohorts of patients.

In conclusion, UDPS is a powerful tool that can be used with confidence to predict MVC response in HIV-1-infected patients, provided that appropriate software is used for data analysis. Among these, Pyrotrop[®] appears to fulfill the criteria for selection. Improvement of the predictive value of the algorithm used is possible and our results suggest that adding the H34S/Y substitution would substantially improve the performance of the 11/25/charge rule with these new approaches.

REFERENCES

1. Berger EA, Murphy PM, Farber JM. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Ann Rev Immunol* 1999,**17**:657-700.
2. Soriano V, Geretti AM, Perno CF, Fatkenheuer G, Pillay D, Reynes J, *et al.* Optimal use of maraviroc in clinical practice. *AIDS* 2008,**22**:2231-2240.
3. Briz V, Poveda E, Soriano V. HIV entry inhibitors: mechanisms of action and resistance pathways. *J Antimicrob Chemother* 2006,**57**:619-627.
4. Este JA, Telenti A. HIV entry inhibitors. *Lancet* 2007,**370**:81-88.
5. Dorr P, Westby M, Dobbs S, Griffin P, Irvine B, Macartney M, *et al.* Maraviroc (UK-427,857), a potent, orally bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5 with broad-spectrum anti-human immunodeficiency virus type 1 activity. *Antimicrob Agents Chemother* 2005,**49**:4721-4732.
6. Carter NJ, Keating GM. Maraviroc. *Drugs* 2007,**67**:2277-2288.
7. Westby M, Lewis M, Whitcomb J, Youle M, Pozniak AL, James IT, *et al.* Emergence of CXCR4-using human immunodeficiency virus type 1 (HIV-1) variants in a minority of HIV-1-infected patients following treatment with the CCR5 antagonist maraviroc is from a pretreatment CXCR4-using virus reservoir. *J Virol* 2006,**80**:4909-4920.
8. Saag M, Goodrich J, Fatkenheuer G, Clotet B, Clumeck N, Sullivan J, *et al.* A double-blind, placebo-controlled trial of maraviroc in treatment-experienced patients infected with non-R5 HIV-1. *J Infect Dis* 2009,**199**:1638-1647.
9. Whitcomb JM, Huang W, Fransen S, Limoli K, Toma J, Wrin T, *et al.* Development and characterization of a novel single-cycle recombinant-virus assay to determine human immunodeficiency virus type 1 coreceptor tropism. *Antimicrob Agents Chemother* 2007,**51**:566-575.

10. Rose JD, Rhea AM, Weber J, Quinones-Mateu ME. Current tests to evaluate HIV-1 coreceptor tropism. *Curr Opin HIV AIDS* 2009,**4**:136-142.
11. Swenson LC, Mo T, Dong WW, Zhong X, Woods CK, Jensen MA, *et al*. Deep sequencing to infer HIV-1 co-receptor usage: application to three clinical trials of maraviroc in treatment-experienced patients. *J Infect Dis* 2011,**203**:237-245.
12. Altshuler D, Pollara VJ, Cowles CR, Van Etten WJ, Baldwin J, Linton L, *et al*. An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature* 2000,**407**:513-516.
13. Delobel P, Nugeyre MT, Cazabat M, Pasquier C, Marchou B, Massip P, *et al*. Population-based sequencing of the V3 region of env for predicting the coreceptor usage of human immunodeficiency virus type 1 quasispecies. *J Clin Microbiol* 2007,**45**:1572-1580.
14. De Jong JJ, De Ronde A, Keulen W, Tersmette M, Goudsmit J. Minimal requirements for the human immunodeficiency virus type 1 V3 domain to support the syncytium-inducing phenotype: analysis by single amino acid substitution. *J Virol* 1992,**66**:6777-6780.
15. de Jong JJ, Goudsmit J, Keulen W, Klaver B, Krone W, Tersmette M, *et al*. Human immunodeficiency virus type 1 clones chimeric for the envelope V3 domain differ in syncytium formation and replication capacity. *J Virol* 1992,**66**:757-765.
16. Fouchier RA, Brouwer M, Broersen SM, Schuitemaker H. Simple determination of human immunodeficiency virus type 1 syncytium-inducing V3 genotype by PCR. *J Clin Microbiol* 1995,**33**:906-911.
17. Fouchier RA, Groenink M, Kootstra NA, Tersmette M, Huisman HG, Miedema F, *et al*. Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. *J Virol* 1992,**66**:3183-3187.

18. Resch W, Hoffman N, Swanstrom R. Improved success of phenotype prediction of the human immunodeficiency virus type 1 from envelope variable loop 3 sequence using neural networks. *Virology* 2001,**288**:51-62.
19. Recordon-Pinson P, Soulie C, Flandre P, Descamps D, Lazrek M, Charpentier C, *et al.* Evaluation of the genotypic prediction of HIV-1 coreceptor use versus a phenotypic assay and correlation with the virological response to maraviroc: the ANRS GenoTropism study. *Antimicrob Agents Chemother* 2010,**54**:3335-3340.
20. Venkatraman ES BC. A distribution-free procedure for comparing receiver operating characteristic curves from a paired experiment. *Biometrika* 1996,**83**:835-848.
21. Venkatraman ES. A permutation test to compare receiver operating characteristic curves. *Biometrics* 2000,**56**:1134-1138.
22. Bandos AI, Rockette HE, Gur D. A permutation test sensitive to differences in areas for comparing ROC curves from a paired design. *Stat Med* 2005,**24**:2873-2893.
23. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011,**12**:77.
24. R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna (Austria), 2008.

Table 1. Concordance between the 4 methods used to analyze HIV tropism in the 102 patients with UDPS sequences of good quality at D0. Consensus means that the consensus of UDPS sequences has been used as a surrogate of population sequencing; UDPS means that the full set of pyrosequencing sequences has been used; g2p means that the geno2pheno algorithm has been used for analysis; Pyrotrop[©] means that our in-house software has been used for analysis.

	Consensus/g2p % (n)	Consensus/Pyrotrop [©] % (n)	UDPS/g2p % (n)	UDPS/Pyrotrop [©] % (n)
Consensus/g2p	100% (102)	76.5% (78)	77.5% (79)	71.6% (73)
Consensus/Pyrotrop [©]		100% (102)	83.3% (85)	94.1% (96)
UDPS/g2p			100% (102)	86.3% (88)
UDPS/Pyrotrop [©]				100% (102)

Table 2. Sensitivity, specificity, positive and negative predictive values of the 4 methods used to predict three magnitudes of HIV RNA level decrease on MVC therapy (0.5, 1.0 or 2.0 log, respectively) in the 84 patients with available HIV RNA levels at D0 and M1. Consensus means that the consensus of UDPS sequences has been used as a surrogate of population sequencing; UDPS means that the full set of pyrosequencing sequences has been used; g2p means that the geno2pheno algorithm has been used for analysis; Pyrotrop[®] means that our in-house software has been used for analysis.

	Sensitivity			Specificity			Positive predictive value			Negative predictive value		
	0.5 log	1.0 log	2.0 log	0.5 log	1.0 log	2.0 log	0.5 log	1.0 log	2.0 log	0.5 log	1.0 log	2.0 log
Virologic response												
Consensus/g2p	78.3%	77.6%	76.5%	20.0%	16.7%	17.9%	87.1%	83.9%	62.9%	11.8%	11.8%	29.4%
Consensus/Pyrotrop[®]	95.6%	95.8%	94.6%	10.0%	8.3%	3.6%	88.8%	86.3%	66.3%	25.0%	25.0%	25.0%
UDPSg2p	82.4%	81.9%	80.4%	40.0%	33.3%	21.4%	91.0%	88.1%	67.2%	23.5%	23.5%	35.3%
UDPS/Pyrotrop[®]	91.9%	91.7%	91.1%	30.0%	25.0%	14.3%	90.7%	88.0%	68.0%	33.3%	33.3%	44.4%

ARTICLE 5

Chevaliez S, **Rodriguez C**, Pawlotsky JM

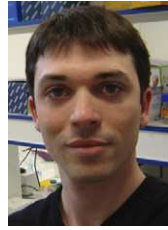
New virologic tools for management of chronic hepatitis B and C

Gastroenterology 2012;142:1303-1313

New Virologic Tools for Management of Chronic Hepatitis B and C



Stéphane Chevaliez*‡



Christophe Rodriguez*‡



Jean-Michel Pawlotsky*‡

*National Reference Center for Viral Hepatitis B, C, and D, Department of Virology, Hôpital Henri Mondor, Université Paris-Est, Créteil, France; and ‡INSERM U955, Créteil, France

Molecular biology techniques are routinely used to diagnose and monitor treatment of patients with chronic hepatitis B virus (HBV) and hepatitis C virus (HCV) infections. These tools can detect and quantify viral genomes and analyze their sequence to determine their genotype or subtype and to identify nucleotide or amino acid substitutions associated with resistance to antiviral drugs. They include real-time target amplification methods, which have been standardized and are widely used in clinical practice to diagnose and monitor HBV and HCV infections, and next-generation sequencing techniques, which are still restricted to research laboratories. In addition, new enzyme immunoassays can quantify hepatitis B surface and hepatitis C core antigens, and point-of-care tests and alternatives to biologic tests that require whole-blood samples obtained by venipuncture have been developed. We review these new virologic methods and their clinical and research applications to HBV and HCV infections.

Keywords: Diagnostic Assay; Real-Time Polymerase Chain Reaction; Real-Time Transcription-Mediated Amplification; Next-Generation Sequencing; Antigen Quantification; Point-of-Care Tests.

Virology techniques developed over the past 20 years are routinely used to diagnose and monitor chronic viral infections, such as those caused by hepatitis B virus (HBV) and hepatitis C virus (HCV). Molecular biology tools can be used to detect and quantify viral genomes, sequence them, assign them to a phylogenetic clade or subclade (genotype or subtype), and identify clinically relevant nucleotide or amino acid substitutions, such as those associated with resistance to antiviral drugs. Recent advances include real-time target amplification methods for detecting and quantifying viral genomes and next-generation sequencing (NGS) techniques. Other new assays detect and quantify viral antigens, whereas point-of-care tests and alternatives to biologic tests that require

whole-blood samples have been developed. We review these methods and how they might be used to diagnose, treat, and manage patients with HBV or HCV infection.

New Methods for Quantifying HBV DNA and HCV RNA

HCV and HBV genomes must be detected and quantified before treatment decisions can be made and to monitor antiviral therapy. The viral genomes can be detected and quantified using target amplification methods, such as polymerase chain reaction (PCR) and transcription-mediated amplification (TMA), and with signal amplification methods, such as hybrid-capture and the branched DNA assay. Recently developed real-time target amplification methods have improved viral genome detection and quantification for clinical and research purposes.

Target Amplification Methods

First-generation target amplification techniques have been widely used to diagnose HBV and HCV infections and to monitor responses to antiviral therapies. In classic PCR or TMA assays, amplicons are detected at the end of the amplification process by their specific hybridization to immobilized oligonucleotide probes; the amplicon-probe hybrids are detected by an enzymatic reaction. They are quantified based on competitive amplification of the viral template with a known amount of synthetic

Abbreviations used in this paper: ALT, alanine aminotransferase; bp, base pair; DAA, direct-acting antiviral; DBS, dried blood spot; emPCR, emulsion polymerase chain reaction; GS, genome sequencer; GWAS, genome-wide association studies; HBeAg, hepatitis B e antigen; HBsAg, hepatitis B surface antigen; HBV, hepatitis B virus; HCV, hepatitis C virus; IFN, interferon; LLOD, lower limit of detection; LLOQ, lower limit of quantification; NGS, next-generation sequencing; PCR, polymerase chain reaction; SVR, sustained virologic response; TMA, transcription-mediated amplification.

© 2012 by the AGA Institute
0016-5085/\$36.00
doi:10.1053/j.gastro.2012.02.027

standard added to each reaction tube; the relative amounts of viral template and standard amplicons are measured, and the results are interpreted with a standard curve established in parallel.

Because amplification reactions are saturable, these methods have a narrow dynamic range of quantification. As a result, high levels of virus are not always accurately quantified and require reanalysis after samples are diluted, whereas low levels of virus (such as in patients receiving antiviral therapy) are often not detectable. This problem was solved by development of real-time target amplification techniques in which quantification takes place during the exponential phase of the amplification reaction. In addition, the reaction is run in a closed system, which prevents carryover contamination and improves specificity.

Real-Time Target Amplification

Real-Time PCR. PCR uses several temperatures and a thermostable DNA polymerase to generate double-stranded DNA amplicons. In assays for HBV, nucleic acids are isolated from a sample, and the HBV DNA is directly amplified by PCR. Assays for HCV require reverse transcription of HCV RNA, to synthesize complementary DNA, which is used as template in the PCR reaction. Each complete PCR cycle doubles the number of DNA copies. The principle of real-time PCR is to detect amplicon synthesis during the PCR reaction and thereby deduce the starting amount of viral genome in a clinical sample. A fluorescent probe is linked to a quencher and annealed to the target sequence between the sense and antisense PCR primers. During each PCR reaction, as the DNA polymerase extends the primer, its intrinsic nuclease activity degrades the probe, releasing the reporter fluorochrome. The amount of fluorescence released during the amplification cycle and detected by the system is proportional to the amount of amplicons generated in each PCR cycle. Software is used to calculate the threshold cycle in each reaction, which has a linear relationship with the initial amount of nucleic acid. In each PCR run, parallel processing of a panel of quantified standards is used to establish a standard curve for quantification.

Real-Time TMA. The TMA reaction is isothermal and uses 2 enzymes: a reverse transcriptase and a T7 RNA polymerase. The amplicons consist of single-stranded RNA. After lysis of the viral envelope, the viral genome (HBV DNA or HCV RNA) is captured by oligonucleotide probes and bound to magnetic microparticles. Amplification involves autocatalytic isothermal production of RNA transcripts with the 2 enzymes. Each newly synthesized RNA re-enters the TMA process and serves as a template for the next round of replication, resulting in exponential amplification of the target RNA. The amplicons created during amplification are detected in real time, using probes that contain a reporter dye (fluorophore) and a quencher. In the absence of amplicons, these probes exist in a closed configuration, and the fluorescence output of the fluorophore is quenched because it is in close prox-

imity to the quencher. During amplification, when these probes bind to amplicons, the fluorophore and quencher are separated, and the fluorescent signal is generated. Software is used to calculate the amount of viral genome in the initial sample by comparison with a panel of quantified standards.

Available methods. Table 1 shows the performance of current and upcoming commercial real-time PCR and TMA assays. These assays accurately quantify HBV DNA and HCV RNA in clinical practice.¹⁻⁵ The first-generation Cobas Ampliprep/Cobas Taqman assay (Roche Molecular Systems, Pleasanton, CA) substantially underestimated levels of HCV RNA in approximately 30% of patients infected with HCV genotype 4, and occasionally failed even to detect this genotype, because of nucleotide polymorphisms at the hybridization site of the PCR primers or TaqMan probe (in the 5' noncoding region of the HCV genome).^{1,3,6,7} This issue has been resolved in the second-generation assay, scheduled for release in 2012.⁸

These assays avoid false-positives that result from carryover contamination. They are fully or partly automated (in the latter case, the extract has to be transferred for automated PCR amplification and quantification, after automated extraction). These techniques are recommended to quantify HBV DNA and HCV RNA in international liver society guidelines.⁹⁻¹⁴

Future Needs

The advent of real-time PCR and TMA methods has resolved many of the problems of quantifying HBV DNA and HCV RNA in clinical practice. HBV DNA assays need further improvement to increase their upper limit of quantification—clinicians frequently encounter patients (particularly those who are immunotolerant) with levels of HBV DNA much higher than 10⁸ IU/mL.

There were confusing results from recent phase 2 and 3 clinical trials of direct-acting antiviral (DAA) agents against HCV, administered with or without pegylated interferon (IFN)- α and ribavirin, according to the use of the lower limit of detection (LLOD) or the lower limit of quantification (LLOQ) defined by the assay manufacturers to assess the on-treatment virologic response and to predict sustained virologic response (SVR). In trials of pegylated IFN- α , ribavirin, and telaprevir or boceprevir, HCV RNA levels below the LLOD (at weeks 4, 8, or 12, depending on the protease inhibitor) were reported to be more predictive of SVR than HCV RNA levels below the LLOQ.^{15,16} However, the LLOQ is, by definition, accurate and reproducible because it is within the dynamic range of quantification, whereas the LLOD is defined statistically, and the actual amount of HCV RNA it indicates varies among patients and samples. Therefore, the LLOD is not suitable for making individual therapeutic decisions. Future assays should have identical LLOD and LLOQ values to ensure accurate definition of undetectable levels of HCV RNA. In the meantime, new time points should be defined for accurate assessment using the LLOQ.

Table 1. Commercially Available Real-Time Target Amplification Assays for HBV DNA and HCV RNA Detection and Quantification

Assay	Manufacturer	Method	Automated extraction device	Amplification device	Volume required (μL)	LLOD (IU/mL)	Dynamic range of quantification (IU/mL)
HBV DNA							
COBAS TaqMan HBV Test, v2.0	Roche Molecular Systems, Pleasanton, CA	Real-time PCR	Manual with high pure system viral nucleic acid kit	COBAS TaqMan	650	6–10 IU/mL	29 ($1.5 \log_{10}$)– 1×10^8 ($8.0 \log_{10}$) IU/mL
COBAS Ampliprep-COBAS TaqMan (CAP/CTM) HBV Test, v2.0	Roche Molecular Systems, Pleasanton, CA	Real-time PCR	COBAS Ampliprep	COBAS TaqMan	650	20 IU/mL	20 ($1.3 \log_{10}$)– 1.7×10^8 ($8.2 \log_{10}$) IU/mL
RealTime HBV	Abbott Molecular, Des Plaines, IL	Real-time PCR	m2000SP	m2000RT	200 or 500	10 IU/mL (for 500 μL) 15 IU/mL (for 200 μL)	10 ($1.0 \log_{10}$)– 1.0×10^9 ($9.0 \log_{10}$) IU/mL
Artus HBV QS-RGQ Assay ^a	Qiagen, Hilden, Germany	Real-time PCR	QIASymphony RGQ	Rotor-Gene Q	1000	10.2 IU/mL	31.6 ($1.5 \log_{10}$)– 2×10^7 ($7.3 \log_{10}$) IU/mL
APTIMA HBV Quantitative Assay	Gen-Probe, San Diego, CA	Real-time TMA		Panther		Available in 2012	
HCV RNA							
COBAS TaqMan HCV Test, v2.0	Roche Molecular Systems, Pleasanton, CA	Real-time PCR	Manual with high pure system viral nucleic acid kit	COBAS TaqMan	650	20 IU/mL	25 ($1.4 \log_{10}$)– 3.0×10^8 to 3.9×10^8 (8.5 – $8.6 \log_{10}$) IU/mL
COBAS Ampliprep-COBAS TaqMan (CAP/CTM) HCV Test, v2.0	Roche Molecular Systems, Pleasanton, CA	Real-time PCR	COBAS Ampliprep	COBAS TaqMan	650	15 IU/mL	15 ($1.2 \log_{10}$)– 1.0×10^8 ($8.0 \log_{10}$) IU/mL
RealTime HCV	Abbott Molecular, Des Plaines, IL	Real-time PCR	m2000SP	m2000RT	200 or 500	12 IU/mL (for 500 μL) 30 IU/mL (for 200 μL)	12 ($1.1 \log_{10}$)– 1.0×10^8 ($8.0 \log_{10}$) IU/mL
Artus HCV QS-RGQ Assay ^a	Qiagen, Hilden, Germany	Real-time PCR	QIASymphony RGQ	Rotor-Gene Q	1000	36.2 IU/mL	67.6 ($1.8 \log_{10}$)– 17.7×10^6 ($7.2 \log_{10}$) IU/mL
VERSANT HCV RNA 1.0 Assay (kPCR)	Siemens Medical Solutions Diagnostics, Tarrytown, NY	Real-time PCR	Sample preparation (SP) module	Amplification and Detection (AD) Module	500	15 IU/mL	15 ($1.2 \log_{10}$)– 1.0×10^8 ($8.0 \log_{10}$) IU/mL
APTIMA HCV Quantitative Assays	Gen-Probe, San Diego, CA	Real-time TMA		Panther		Available in 2012	

LLOD, lower limit of detection.

^aAvailable only for EDTA-plasma specimens.

New Serologic Methods

Hepatitis B Surface Antigen Quantification

The level of hepatitis B surface antigen (HBsAg) has been proposed as a marker of infected liver mass or the amount of covalently closed circular DNA, the episomal form of HBV which persists in hepatocytes. Combined analysis of HBsAg and HBV DNA levels can be used to identify inactive carriers and monitor response to therapy.¹⁷⁻²⁰ Two commercial enzyme immunoassays are available; they have narrow dynamic ranges of quantification, so samples that contain high levels of HBsAg must often be retested after being diluted. The HBsAg assay on Architect system (Abbott Diagnostics, Abbott Park, IL) has a dynamic range of quantification of 0.05–250.00 IU/mL (maximum authorized dilution, 1:999), whereas the HBsAg II Quant assay on Elecsys or Cobas e devices (Roche Diagnostics GmbH, Mannheim, Germany) has a dynamic range of quantification of 0.05–130.00 IU/mL (no recommendation for maximum dilution). Both assays are European Conformity marked, but they have not been approved by the US Food and Drug Administration and are available for research use only in the United States. The assays are inexpensive, easy to perform, and accurately quantify HBsAg in all HBV genotypes.^{21,22}

HCV Core Antigen Detection and Quantification

Levels of HCV core antigen correlate with levels of HCV RNA in patients with chronic HCV infection.²³⁻²⁶ Assays that quantify HCV core antigen can therefore be used as alternatives to those that measure levels of HCV RNA, as shown in various populations of patients.^{23,27-29} The Architect HCV Ag assay (Abbott Diagnostics) is commercially available. Its dynamic range of quantification is 3.0–20,000 fmol/L. This assay is European Conformity marked but has not been approved by the US Food and Drug Administration and is available for research use only in the United States. It is less expensive and easier to use than current HCV RNA tests for diagnosis of chronic hepatitis C and monitoring of antiviral therapy. However, the HCV Ag assay's LLOD corresponds to HCV RNA levels of 500–3000 IU/mL, depending on the HCV genotype.³⁰ Assays for HCV core antigen are therefore not suitable for response-guided therapy, according to current guidelines.

Point-of-Care Tests and Alternatives to Biologic Tests That Require Whole-Blood Samples

Point-of-care tests are used directly at the site of patient care, outside of the biology laboratory. They include immunoassays based on agglutination, immunochromatography, or immunofiltration, as well as assays for nucleic acid detection and quantification. Point-of-care tests can be particularly useful for large-scale screening and improving access to care in regions without molecular biology laboratories.

Alternatives to biologic tests that require whole-blood samples, obtained by venipuncture, are being developed. They can use original specimen matrices, such as oral fluids or blood collected from a finger stick into a capillary tube. Blood can be collected on filter paper, known as the dried blood spot (DBS) method, which makes it possible to store desiccated samples for transport as nonhazardous material via regular mail or courier services. DBS collection and storage is easy, and, although standard procedures must be carefully followed, little training is required for preparation. HBV DNA and HCV RNA can be detected and quantified from DBS by classic or real-time PCR methods, although at a cost of reduced analytical sensitivity, compared with the whole blood-based methods described above. The reduction in sensitivity results from small input volumes, less-efficient nucleic acid extraction, and nucleic acid degradation in extreme storage conditions. One study reported that the limit of detection of HBV DNA from DBS was approximately 1-log₁₀ higher than from serum samples. There was a high level of concordance in HBV genotype analysis of DBS and serum samples and also in detection of precore mutants and drug-resistant variants.³¹ In another report, the detection limit for HCV RNA in DBS was about 1000 IU/mL. Concordant results were also obtained from HCV quantification and genotype analyses of serum and DBS.³² In theory, the analytical sensitivity of viral genome detection from DBS could reach levels similar to those obtained from serum and plasma samples, as recently shown for HIV.³³ Any serum marker, including HBsAg and HCV core antigen, could also be detected and quantified from DBS. However, this has not been tested yet.

Next-Generation Sequencing

Viral genome sequence analysis can be used to determine HBV or HCV genotype, assess mutations in the HBV core or core promoter, and identify amino acid substitutions responsible for resistance to antiviral agents. DNA sequence analysis has been based for many years on the chain termination method, described by Sanger et al.³⁴ NGS methods offer new opportunities but also create new challenges because they generate large amounts of data, and their results can often be difficult to interpret.

Classic Sequencing Techniques

The Sanger method has been improved by the inclusion of fluorescent labels and capillary electrophoresis.^{35,36} This technique involves DNA denaturation, annealing, and elongation. During elongation, DNA polymerase randomly incorporates deoxynucleotides or dideoxynucleotides that have been labeled with different fluorochromes, resulting in chain elongation or termination, respectively. Chains of different lengths are therefore generated, each ending with a labeled nucleotide. The amplicons are then migrated in a capillary system that separates them according to their lengths. Bioinformatic software measures the fluorescence peak at each position

of the analyzed sequence and identifies the corresponding dideoxynucleotides to determine a consensus sequence. The presence of multiple peaks at the same position indicates a mixed sequence.

The Sanger sequencing method is well suited for detecting small numbers of substitutions in known genes, but it cannot be used to sequence large portions of plant or animal genomes for genome-wide association studies (GWAS)³⁷⁻³⁹ or to sequence large amounts of targets for gene expression profiling, so-called RNA-seq.⁴⁰

Viruses that establish chronic infections, such as HIV, HBV, and HCV, behave as quasispecies—complex mixtures of genetically related but distinct viral populations in equilibrium in a given replicative environment (Figure 1). For this reason, at any given time, patients carry a large number of different viral genome sequences. Classic sequencing techniques can only identify the consensus sequence of a viral quasispecies at a given time point, whereas most intermediate and minor viral populations go undetected. In addition, this approach cannot link different substitutions present in the same viral variant or distinguish substitutions present in different variants. PCR products can be cloned and sequenced to identify intermediate virus populations and linked substitutions, but this approach is cumbersome and time-consuming and has therefore been restricted to the research setting.

Principles of NGS

NGS methods have been recently developed to increase sequencing capacity while generating clonal sequences. Current NGS methods use a 3-step sequencing process: library preparation, DNA capture and enrichment, and sequencing (Figure 2).⁴¹ In library preparation, DNAs of appropriate lengths are prepared and labeled

with the primers required for subsequent isolation and sequencing. DNA is then captured, and clonal DNA is enriched (this step is not required with Pacific Biosciences technology). The DNA is then sequenced, via polymerization combined with detection based on pyrophosphate (pyrosequencing), H⁺ ions (pH variation) or fluorescence, or ligation combined with fluorescence detection. The detected signals are transformed into readable sequences by an integrated browser and specific software to produce biologically relevant information.

NGS Methods

NGS technologies developed by various manufacturers (Table 2) include: high-capacity systems, such as the Genome Analyzer, HiSeq sequencers (Illumina Inc, San Diego, CA), and 5500 series sequencers, which use SOLiD technology (Applied Biosystems, Carlsbad, CA); and long-read sequencers, such as the Genome Sequencer (GS) FLX or Junior (454 Life Sciences, Roche Diagnostics Corp, Branford, CT), Ion Torrent (Applied Biosystems), and PacBio RS (Pacific Biosciences, Gen-Probe, Menlo Park, CA).

The Illumina technology fragments DNA and then tags the fragments with primers that are complementary to probes to allow fixation on a flow cell. Clonal DNA is then enriched by PCR to generate clusters of amplicons and sequenced through polymerization with fluorescent deoxynucleotides.⁴² The SOLiD system (Applied Biosciences) uses emulsion PCR (emPCR) to enrich clonal DNAs fixed to microbeads. The beads are then transferred to a flow chip, and DNA is sequenced by ligation with fluorescent primers.⁴³ These 2 technologies generate millions of short reads (Table 2). They are well suited to resequencing when a reference sequence is available (hu-

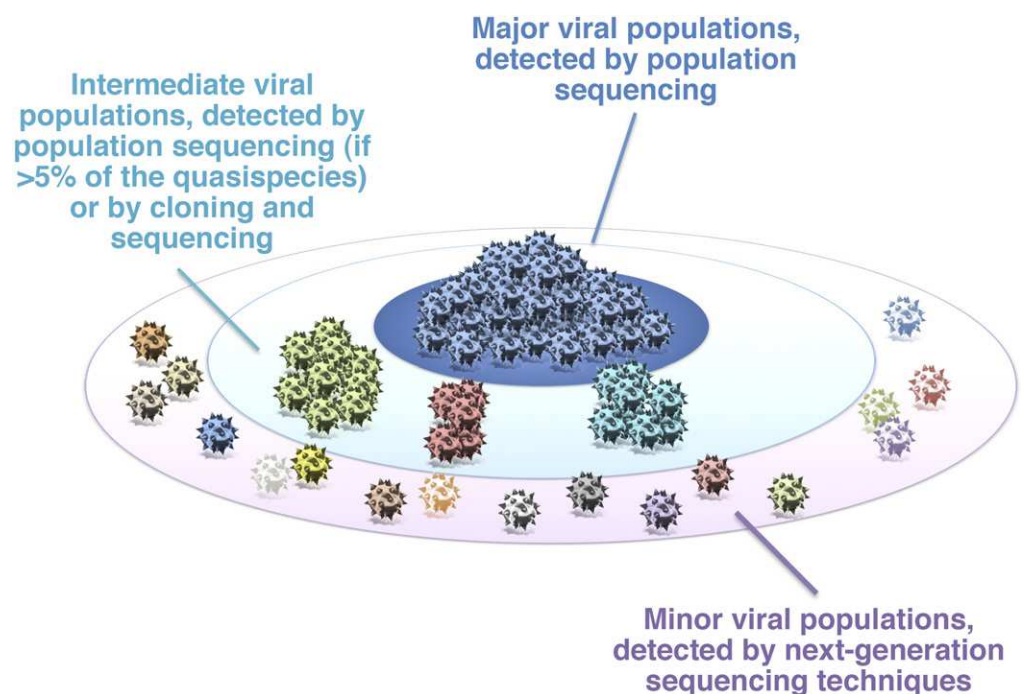


Figure 1. Quasispecies distribution of HBV and HCV genomes and methods to detect major (>10% of the viral quasispecies), intermediate, and minor (<1% of the viral quasispecies) viral populations.

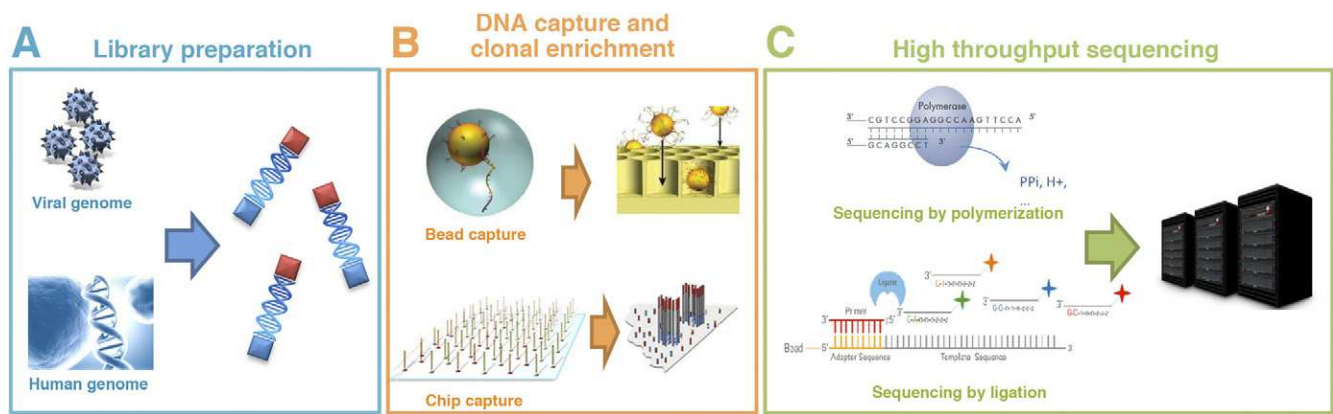


Figure 2. Current workflow for NGS methods, including (A) library preparation, (B) DNA capture and clonal DNA enrichment, and (C) high-throughput sequencing that can be achieved by polymerization, combined with detection based on pyrophosphate (pyrosequencing), H^+ ions (pH variation) or fluorescence, or by ligation combined with fluorescence detection. The detected signals are transformed into readable sequences by integrated browser and specific software to yield biologically relevant information.

man genome or transcriptome), and the read length of 100–150 base pairs (bp) is sufficient to identify a known gene, for example in GWAS or RNA-seq studies. These technologies can also be used to identify individual substitutions in viral genes, such as those associated with drug resistance. Analysis of a large number of sequences (corresponding to the depth of the reaction) ensures high levels of sensitivity for minor viral variants. However, short reads cannot be used to establish linkages between substitutions located on the same viral variants, which is required for viral resistance studies.

The long-read sequencers have more limited sequencing capacities but generate longer sequences. In the GS-based technology, DNA is labeled and clonally enriched by emPCR at the surface of microbeads, which are then loaded in picotiter plates containing over a million wells. During the pyrosequencing reaction, light is emitted every time a new base is inserted by polymerization.⁴⁴ The Ion Torrent technology is technically close to the GS technique, but emPCR amplicons are loaded on a semiconductor sequencing chip; the addition of nucleotides during the polymerization reaction generates H^+ ions, which induce local pH changes that are detected and mea-

sured.⁴⁵ The PacBio RS technology is substantially different. After preparation of the DNA library, unamplified DNAs are transferred to an optical system, called a *zero-mode waveguide*, which amplifies fluorescence emitted during polymerization. In this technique, fluorescent signals from each clonal DNA molecule, processed by thousands of polymerase molecules fixed to the zero-mode waveguide, are detected without the need for prior amplification.⁴⁶ This group of sequencers generates smaller numbers of sequences, but their length is suitable for linkage studies and de novo characterization of genomic sequences for which no reference sequence is available. They are therefore well suited to virology studies.

Limitations of NGS Methods

All existing methods are susceptible to generating sequencing errors. Evaluations of the Genome Analyzer, 5500 sequencer, and GS FLX have calculated accuracy values of 96.7%–100%—similar to those of the classic Sanger method (Table 2).⁴⁷ However, there is controversy about the actual error rates of these methods, so prospective assessments are required by independent investigators.

Table 2. Characteristics of NGS Techniques

Manufacturer	Sequencing device	Technology (template preparation/NGS chemistry)	Type	Number of single reads per run ^a ($\times 10^6$)	Number of nucleotides per run ^a (Gb)	Maximum sequence length (bp) ^a	Accuracy
Applied Biosystems	5500	emPCR/ligation	High throughput	800	9	75	99.6%–99.8%
	5500xl		High throughput	1600	15	75	
	Ion Torrent (ChIP 316)		Long reads	6.2	>1	>400	
Illumina	MiSeq	Solid capture/reversible terminator	High throughput	3.4	>1	150	96.7%–100%
	Genome Analyzer IIx		High throughput	320	95	150	
	HiSeq 1000		High throughput	1500	300	100	
	HiSeq 2000		High throughput	3000	600	100	
454 Life Sciences/ Roche Diagnostics	GS Junior	emPCR/pyrosequencing	Long reads	0.1	0.035	400	99% ^a
	GS FLX+		Long reads	1	0.7	1000	97.4%–99.9%
Pacific Biosciences/ Gen-Probe	PacBio RS	Single molecule/real-time sequencing	Long reads	0.035	0.045	1200	99.99% ^a

Gb, gigabase; bp, base pair; emPCR, emulsion PCR.

^aInformation provided by the manufacturers.

NGS methods generate unreadable sequences (called *junk data*) at variable rates. For instance, these rates appear to be high with SOLiD technology on the 5500 sequencer.⁴⁷⁻⁵⁰ This is associated with an over cost but not with lower quality final results because these sequences can be easily identified and removed. False-positive or false-negative results for substitutions have been reported at higher rates with Illumina technology than the SOLiD and GS FLX technologies.^{47,50} They are easily corrected when the sequencing depth is high, as in GWAS or gene expression profiling studies, but are problematic for studies of rare variants.

Performance can be affected by AT- or GC-rich regions, especially with the Illumina and SOLiD technologies, leading to low overall resolution of genome sequences.⁵¹ For the detection, of insertions or deletions, so-called indels, the GS-FLX system has been shown to provide the best performance,^{47,52} but it performs less well in homopolymeric regions, where it has high rates of error.^{51,53}

Analyzing NGS Data

In the first step of NGS data analysis, signals generated by the sequencing reaction are converted into readable sequences by proprietary base-caller software. At the end of this step, data are composed of sequences in FASTA format, along with quality scores for each nucleotide of each sequence. The most frequently used format is the PHRED quality score,⁵⁴⁻⁵⁶ which is based on base-calling error probabilities.

In the second part of the analysis, the large number of generated sequences must be processed into comprehensible biologic data by software through 3 steps that depend on the type of study and sample analyzed.⁵⁷ The goal of the first step is to classify and format sequences into readable data. In the second step, unreliable sequences are eliminated, including sequences with a Phred quality score below 20 (1% risk of base-caller error)⁵⁸ and excessively short sequences (the precise cut-off depends on the technology used). The third step, sequence alignment, is critical and particularly difficult for viral sequences because of the level of natural genetic variation (numerous substitutions or indels, relative to reference sequences) and the coexistence of numerous viral populations with different sequences. A number of alignment algorithms have been developed but were found to generate different alignments and generally fail to take into consideration the specificities of viral infections.⁵⁹⁻⁶⁴ Additional alignment quality filters are therefore needed. This part of the analysis must ultimately generate a set of formatted sequences that can be used with confidence for specific analyses and to generate clinically meaningful information.

We recently developed in-house software to model the dynamics of viral populations during and after antiviral therapy and characterize drug resistance.⁶⁴⁻⁶⁶ This software package has been used to demonstrate the pre-existence of amino acid substitutions that confer HBV and HCV resistance to specific antiviral drugs. It has also

been used to characterize the complex dynamics of sensitive and resistant viral populations in patients with chronic hepatitis B treated with adefovir and in patients with chronic hepatitis C treated with telaprevir combined with pegylated IFN- α , with or without ribavirin.^{64,65}

Future Needs

The main advantage of NGS techniques for studying chronic viral infections is their improved depth, relative to the Sanger method: this ensures that minor viral populations that are of possible clinical significance are sequenced. No further improvements in depth appear to be needed because existing technologies can detect minor viral populations in complex quasispecies mixtures. However, the clinical significance of small viral populations is not clear. In addition, it is still difficult to distinguish between true substitutions and errors generated by the sequencing reaction, so specificity needs to be improved.

Sequence length is an important factor in identifying linkages between different substitutions located on the same viral strand. Full-length viral genome analysis will improve our understanding of the dynamics of viral populations in different clinical circumstances. Reading lengths will need to be extended considerably to meet this need (the HBV genome contains approximately 3000 bp, the HCV genome 10,000 bp).

Currently, the main need is for bioinformatic tools to analyze NGS sequences in the context of viral infections. Complex tools might be suitable for research purposes, but clinicians will require bioinformatic tools that can translate millions of generated sequences into simple, clinically relevant information on which therapeutic decisions can be made.

Clinical Use of Virologic Tools

HBV Infection

Tests that are useful for patients with HBV infection include enzyme immunoassays that detect HBV antigen and antibodies, assays that detect and quantify HBV DNA, and methods that identify amino acid substitutions associated with HBV resistance to nucleoside/nucleotide analogues. There is debate over the value of assays that quantify HBsAg. Tools that determine HBV genotype or identify core and core promoter mutants are valuable for research but have no clinical use.

Diagnosis and prognosis. Chronic HBV carriers are defined by the persistence of HBsAg in their blood for more than 6 months. Chronic hepatitis B is characterized by a level of HBV DNA >2000–20,000 IU/mL (ie, >3.3–4.3 log₁₀ IU/mL), a persistent or intermittent increase in level of aminotransferase, and specific features of liver biopsies.^{9,13,14} Patients with viral loads below 2000 IU/mL are generally considered to be inactive carriers, although there is controversy over this value.

The level of HBV DNA has prognostic value. A chronically high level of HBV DNA is associated with a significant risk of progression to cirrhosis and hepatocellular

carcinoma.^{67,68} However, no HBV DNA threshold has been established that has a strong predictive value for the occurrence of clinical outcomes at the individual level.

Decision to treat. The decision to treat chronic hepatitis B is based on multiple parameters, including the level of HBV DNA. The European Association for the Study of the Liver Clinical Practice Guidelines states that HBV therapy is indicated for noncirrhotic patients with an HBV DNA level >2000 IU/mL ($>3.3 \log_{10}$ IU/mL), serum levels of alanine aminotransferase (ALT) activity above the upper limit of normal, and a liver biopsy (or a noninvasive test) that indicates moderate to severe activity (METAVIR grade $\geq A2$) and/or fibrosis (METAVIR stage $\geq F2$).⁹ Patients with compensated cirrhosis and detectable HBV DNA should be considered for treatment and patients with decompensated cirrhosis require urgent antiviral treatment, whatever the level of HBV replication.⁹ The Clinical Practice Guidelines of the American Association for the Study of Liver Diseases and the Asian-Pacific Association for the Study of the Liver state that antiviral treatment should be considered for noncirrhotic patients with HBV DNA titers $>20,000$ IU/mL ($>4.3 \log_{10}$ IU/mL) if they are HBeAg positive or >2000 IU/mL ($>3.3 \log_{10}$ IU/mL) if they are HBeAg negative and if they have increased serum levels of ALT activity ($>2 \times$ upper limit of normal) and/or a liver biopsy showing inflammation and/or fibrosis.^{13,14} Patients should not be tested for drug resistance before therapy.

In patients with normal ALT activity, serum levels of HBV DNA and ALT should be measured at 3- to 6-month intervals. More frequent monitoring should be considered during the first year for hepatitis B e antigen (HBeAg)-negative patients with normal levels of ALT and a level of HBV DNA <2000 IU/mL ($<3.3 \log_{10}$ IU/mL) to confirm their status as inactive carriers.

Choice of therapy. Pegylated IFN- α is considered to be the first-line treatment option for HBeAg-positive patients without cirrhosis who have increased levels of ALT and a low level of HBV DNA ($<2 \times 10^6$ IU/mL, such as $<6.3 \log_{10}$ IU/mL). This treatment is more controversial for HBeAg-negative patients with the same characteristics.⁹ The HBV genotype is a significant predictor of the response to pegylated IFN- α . However, the individual predictive value of the HBV genotype for the outcome of pegylated IFN- α therapy is weak and could be biased by the relationship between genotype and ethnicity.^{69,70}

Most patients are treated with nucleoside or nucleotide analogues. The International Liver Society guidelines recommend either tenofovir or entecavir monotherapy as the first-line treatment for chronic hepatitis B.^{9,13,14,71} The HBV genotype has no influence on the effectiveness of nucleoside/nucleotide analogues.⁷⁰

Monitoring treatment. The ideal end point of therapy is a sustained loss of HBsAg, with or without seroconversion to anti-HBs.⁹ However, this is rarely achieved with current drugs. The practical aim is therefore to reduce levels of HBV DNA as far as possible (ideally below 10–20 IU/mL) to achieve biochemical remission and his-

toxic improvement and prevent HBV-related liver disease.^{9,13,14} Treatment of HBeAg-positive patients is therefore monitored by assays for HBV DNA and ALT every 3 to 6 months and HBeAg and antibody to HBeAg every 6 months.

Treatment of HBeAg-positive patients can be stopped after a consolidation period of 6 to 12 months following hepatitis B e seroconversion.^{9,13,14} However, durable hepatitis B e seroconversion, in association with low levels of HBV replication, is rare following treatment with nucleoside/nucleotide analogues. In HBeAg-positive patients who do not develop antibody to HBeAg, and in HBeAg-negative patients, nucleoside/nucleotide analogues must be administered for life, with the aim of achieving profound and durable HBV DNA suppression. If HBV DNA is still detectable (>10 – 20 IU/mL) after 48 weeks of therapy, current guidelines, based on studies of modestly potent nucleoside/nucleotide analogues with low barriers to resistance, recommend adding a second drug with no viral cross-resistance to the first to prevent emergence of resistant variants. However, this is not necessary with entecavir or tenofovir if the HBV DNA level continues to decrease after week 48 because it is likely to become undetectable thereafter.⁷² Serial measurements of HBsAg levels could help identify patients who will subsequently clear HBsAg.

In patients who adhere and respond to treatment, selection of resistant HBV variants should be suspected if virologic breakthrough occurs, defined as a re-increase in the HBV DNA level by 1 \log_{10} or more above the nadir, or as HBV DNA that becomes detectable after being undetectable.^{14,73} A blip (single detection of a low-level of HBV DNA, preceded and followed by undetectable HBV DNA) should not be considered a virologic breakthrough. If resistance occurs, the only effective strategy is to add a second drug against which the selected HBV variant is not cross-resistant.^{9,14} Such drug combinations include tenofovir and a nucleoside analogue. Identification of amino acid substitutions associated with resistance has little clinical value because this information rarely influences the choice of therapy. Thus, resistance testing by means of classic sequencing or reverse hybridization is generally not recommended. The only exception is HBV with the rtA181V/T substitution: it has full sensitivity to only entecavir, which must be combined with tenofovir.

Clinical research. Cohort studies, surveillance studies, and clinical trials of new therapeutic agents or strategies require serial measurements of HBV DNA. In research studies, tests for resistant variants should be systematically performed when antiviral therapy fails, to identify the resistant variant and characterize its nature and dynamics. Sensitive sequencing methods such as NGS are particularly useful for characterizing the dynamics of sensitive and resistant virus populations over time. Information about HBV genotype and core and core promoter mutations might be used in epidemiology and pathophysiology studies.

HCV Infection

Assays that detect antibodies against HCV, detect and quantify HCV RNA, or determine HCV genotype are useful for diagnosis and treatment of HCV infection. Assays that detect and quantify HCV core antigen are alternatives to tests for HCV RNA but are not sensitive enough to use in treatment monitoring or response-guided therapy. Resistance testing is not useful for patient management but is widely used in clinical trials of DAA or host-targeted agents in case of treatment failure.

Diagnosis and decision to treat. Patients are diagnosed with chronic HCV infection based on the detection of antibodies to HCV by enzyme immunoassays and HCV RNA by a sensitive molecular biology-based technique that has a detection limit of about 10 to 15 IU/mL.

All treatment-naïve patients with compensated liver disease and detectable HCV RNA should be considered for therapy.^{10,12} Assessment of liver disease severity is important for decision making.^{10,12} Patients infected with HCV genotype 1, in whom pegylated IFN- α and ribavirin combination therapy failed to eradicate the virus, should be considered for retreatment because they might benefit from the triple combination of pegylated IFN- α , ribavirin, and a protease inhibitor.

Selecting therapy. The HCV genotype should be determined before treatment is started. Patients infected with HCV genotypes other than type 1 should be treated with pegylated IFN- α and ribavirin alone. For patients infected with genotype 2 or 3, the dose of ribavirin is 0.8 g/day, and the treatment duration is 24 weeks; patients with genotypes 2 and 3 with baseline factors that indicate low responsiveness should receive weight-based ribavirin.¹⁰ For patients infected with genotypes 5 and 6, the dose of ribavirin is based on body weight, and treatment should last up to 72 weeks, depending on the virologic response during treatment.

Patients infected with HCV genotype 1 who qualify for therapy should receive the triple combination of pegylated IFN- α , ribavirin, and either telaprevir or boceprevir. The subtype (1a or 1b) does not need to be determined for patients who receive the combination of pegylated IFN- α and ribavirin, or triple drug combinations that include a protease inhibitor, because genotype 1 subtypes do not affect treatment decisions with these therapeutic strategies.⁷⁴ Further trials are needed to improve the selection of patients infected with HCV genotype 1 who might not need protease inhibitors, such as young patients with mild fibrosis, a low baseline level of HCV RNA, or the CC IL28B genotype.⁷⁵

Monitoring treatment. The end point for HCV therapy is an SVR, characterized by undetectable HCV RNA (<10–15 IU/mL) 24 weeks after the end of treatment: this corresponds with viral eradication in more than 99% of cases.

Monitoring of HCV RNA levels during treatment is key in determining virologic response, in guiding duration of treatment, and in deciding futility of treatment. This is

particularly important when DAAs are used, as reviewed by Barritt and Fried in this issue.

Failure of triple combination therapy to clear HCV is associated with outgrowth of variant populations that are resistant to the protease inhibitor used. However, resistance testing based on HCV sequence analysis of the protease region has no utility in clinical practice.⁷⁶ Resistant variants are present in almost all infected patients before therapy begins, but technologies available in clinical virology laboratories are not sensitive enough to detect them⁶⁵; negative results are therefore uninterpretable. In addition, the profile of HCV variants at the start of therapy does not appear to affect the outcome of treatment when the protease inhibitor is combined with pegylated IFN- α and ribavirin. At the time of treatment failure (virologic breakthrough or relapse), the resistant population is always enriched, and phase 2 and 3 trials have reported that it is dominant in 50% to 70% of patients.^{77,78} Resistance testing is still not indicated at this time because the results will have no impact on subsequent treatment decisions.⁷⁶

Clinical research. In cohort studies, surveillance studies, and clinical trials of HCV DAA or host-targeting agents, given alone or in combination, serial HCV RNA testing is needed to characterize the virologic response and detect virologic breakthrough and relapse. Resistance variants should always be analyzed in patients who fail to respond to therapy to characterize the nature and dynamics of the selected viral populations and identify amino acid substitutions selected by the DAAs. When several antiviral drugs with different target sites and mechanisms of action are combined, all the target sites should be studied. Sensitive sequencing methods such as NGS are particularly useful for characterizing the dynamics of sensitive and resistant virus populations over time.

Conclusion

New technologies are available for detecting and quantifying viral antigens and genomes and for analyzing viral genome nucleotide and amino acid sequences. Real-time target amplification (PCR or TMA) methods are well standardized and widely used in clinical practice to diagnose and monitor HBV and HCV infection. In contrast, NGS techniques are largely restricted to research laboratories. These generate large quantities of data; software is needed (and is being developed) that can provide clinicians with meaningful information.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at doi: 10.1053/j.gastro.2012.02.027.

References

1. Chevaliez S, Bouvier-Alias M, Brillet R, et al. Overestimation and underestimation of hepatitis C virus RNA levels in a widely used

- real-time polymerase chain reaction-based method. *Hepatology* 2007;46:22–31.
2. Chevaliez S, Bouvier-Alias M, Laperche S, et al. Performance of version 2.0 of the Cobas AmpliPrep/Cobas TaqMan real-time PCR assay for hepatitis B virus DNA quantification. *J Clin Microbiol* 2010;48:3641–3647.
 3. Chevaliez S, Bouvier-Alias M, Pawlotsky JM. Performance of the Abbott real-time PCR assay using m2000sp and m2000rt for hepatitis C virus RNA quantification. *J Clin Microbiol* 2009;47:1726–1732.
 4. Paba P, Fabeni L, Perno CF, et al. Performance evaluation of the Artus hepatitis C virus QS-RGQ assay. *J Virol Methods* 2012;179:77–80.
 5. Thibault V, Pichoud C, Mullen C, et al. Characterization of a new sensitive PCR assay for quantification of viral DNA isolated from patients with hepatitis B virus infections. *J Clin Microbiol* 2007;45:3948–3953.
 6. Chevaliez S, Bouvier-Alias M, Castera L, et al. The Cobas AmpliPrep-Cobas TaqMan real-time polymerase chain reaction assay fails to detect hepatitis C virus RNA in highly viremic genotype 4 clinical samples. *Hepatology* 2009;49:1397–1398.
 7. Chevaliez S, Fix J, Soulier A, et al. Underestimation of hepatitis C virus genotype 4 RNA levels by the CobasAmpliPrep/CobasTaqMan assay. *Hepatology* 2009;50:1681.
 8. Vermehren J, Colucci G, Gohl P, et al. Development of a second version of the Cobas AmpliPrep/Cobas TaqMan hepatitis C virus quantitative test with improved genotype inclusivity. *J Clin Microbiol* 2011;49:3309–3315.
 9. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of chronic hepatitis B. *J Hepatol* 2009;50:227–242.
 10. European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of hepatitis C virus infection. *J Hepatol* 2011;55:245–264.
 11. Ghany MG, Nelson DR, Strader DB, et al. An update on treatment of genotype 1 chronic hepatitis C virus infection: 2011 practice guideline by the American Association for the Study of Liver Diseases. *Hepatology* 2011;54:1433–1444.
 12. Ghany MG, Strader DB, Thomas DL, et al. Diagnosis, management, and treatment of hepatitis C: an update. *Hepatology* 2009;49:1335–1374.
 13. Liaw YF, Leung N, Kao JH, et al. Asian-Pacific consensus statement on the management of chronic hepatitis B: a 2008 update. *Hepatol Int* 2008;2:263–283.
 14. Lok AS, McMahon BJ. Chronic hepatitis B: update 2009. *Hepatology* 2009;50:661–662.
 15. Lawitz E, Poordad F, Bronowicki JP, et al. The effect of using lower limit of quantitation (LLQ) vs lower limit of detection (LLD) for the definition of undetectable HCV RNA: data from the RESPOND-2 and SPRINT-2 trials. *Hepatology* 2011;54:442A.
 16. Harrington PR, Zeng W, Naeger LK. Clinical relevance of detectable but not quantifiable hepatitis C virus RNA during boceprevir or telaprevir treatment. *Hepatology* 2012;55:1048–1057.
 17. Brunetto MR, Moriconi F, Bonino F, et al. Hepatitis B virus surface antigen levels: a guide to sustained response to peginterferon alfa-2a in HBeAg-negative chronic hepatitis B. *Hepatology* 2009;49:1141–1150.
 18. Moucari R, Mackiewicz V, Lada O, et al. Early serum HBsAg drop: a strong predictor of sustained virological response to pegylated interferon alfa-2a in HBeAg-negative patients. *Hepatology* 2009;49:1151–1157.
 19. Wursthorn K, Jung M, Riva A, et al. Kinetics of hepatitis B surface antigen decline during 3 years of telbivudine treatment in hepatitis B e antigen-positive patients. *Hepatology* 2010;52:1611–1620.
 20. Chan HL, Thompson A, Martinot-Peignoux M, et al. Hepatitis B surface antigen quantification: why and how to use it in 2011. A core group report. *J Hepatol* 2011;55:1121–1131.
 21. Sonneveld MJ, Rijckborst V, Boucher CA, et al. A comparison of two assays for quantification of hepatitis B surface antigen in patients with chronic hepatitis B. *J Clin Virol* 2011;51:175–178.
 22. Wursthorn K, Jaroszewicz J, Zacher BJ, et al. Correlation between the Elecsys HBsAg II assay and the Architect assay for the quantification of hepatitis B surface antigen (HBsAg) in the serum. *J Clin Virol* 2011;50:292–296.
 23. Bouvier-Alias M, Patel K, Dahari H, et al. Clinical utility of total HCV core antigen quantification: a new indirect marker of HCV replication. *Hepatology* 2002;36:211–218.
 24. Park Y, Lee JH, Kim BS, et al. New automated hepatitis C virus (HCV) core antigen assay as an alternative to real-time PCR for HCV RNA quantification. *J Clin Microbiol* 2010;48:2253–2256.
 25. Hosseini-Moghaddam S, Iran-Pour E, Rotstein C, et al. Hepatitis C core Ag and its clinical applicability: potential advantages and disadvantages for diagnosis and follow-up? *Rev Med Virol* 2012 In press.
 26. Descamps V, Op de Beeck A, Plassart C, et al. Strong correlation between liver and serum levels of hepatitis C virus core antigen and RNA in chronically infected patients. *J Clin Microbiol* 2012;50:465–468.
 27. Moscato GA, Giannelli G, Grandi B, et al. Quantitative determination of hepatitis C core antigen in therapy monitoring for chronic hepatitis C. *Intervirology* 2010;54:61–65.
 28. Mederacke I, Potthoff A, Meyer-Olson D, et al. HCV core antigen testing in HIV- and HBV-coinfected patients, and in HCV-infected patients on hemodialysis. *J Clin Virol* 2012;53:110–115.
 29. Miedouge M, Saune K, Kamar N, et al. Analytical evaluation of HCV core antigen and interest for HCV screening in haemodialysis patients. *J Clin Virol* 2010;48:18–21.
 30. Ross RS, Viazov S, Salloum S, et al. Analytical performance characteristics and clinical utility of a novel assay for total hepatitis C virus core antigen quantification. *J Clin Microbiol* 2010;48:1161–1168.
 31. Jardi R, Rodriguez-Frias F, Buti M, et al. Usefulness of dried blood samples for quantification and molecular characterization of HBV DNA. *Hepatology* 2004;40:133–139.
 32. Tuailon E, Mondain AM, Meroueh F, et al. Dried blood spot for hepatitis C virus serology and molecular testing. *Hepatology* 2010;51:752–758.
 33. Marconi A, Balestrieri M, Comastri G, et al. Evaluation of the Abbott Real-Time HIV-1 quantitative assay with dried blood spot specimens. *Clin Microbiol Infect* 2009;15:93–97.
 34. Sanger F, Air GM, Barrell BG, et al. Nucleotide sequence of bacteriophage phi X174 DNA. *Nature* 1977;265:687–695.
 35. Prober JM, Trainor GL, Dam RJ, et al. A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* 1987;238:336–341.
 36. Smith LM. High-speed DNA sequencing by capillary gel electrophoresis. *Nature* 1991;349:812–813.
 37. Wang WY, Barratt BJ, Clayton DG, et al. Genome-wide association studies: theoretical and practical concerns. *Nature Rev Genet* 2005;6:109–118.
 38. Hardy J, Singleton A. Genomewide association studies and human disease. *N Engl J Med* 2009;360:1759–1768.
 39. Ku CS, Loy EY, Pawitan Y, et al. The pursuit of genome-wide association studies: where are we now? *J Hum Genet* 2010;55:195–206.
 40. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 2009;10:57–63.
 41. Metzker ML. Sequencing technologies: the next generation. *Nat Rev Genet* 2010;11:31–46.
 42. Bennett S. Solexa Ltd. *Pharmacogenomics* 2004;5:433–438.
 43. Pandey V, Nutter RC, Prediger E. Applied biosystems SOLiD system: ligation-based sequencing. In: Janits M, ed. Next-generation genome sequencing. Towards personalized medicine. Hoboken, NJ: Wiley 2008:29–42.
 44. Margulies M, Egholm M, Altman WE, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005;437:376–380.

45. Rothberg JM, Hinz W, Rearick TM, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 2011;475:348–352.
46. Eid J, Fehr A, Gray J, et al. Real-time DNA sequencing from single polymerase molecules. *Science* 2009;323:133–138.
47. Harismendy O, Ng PC, Strausberg RL, et al. Evaluation of next-generation sequencing platforms for population targeted sequencing studies. *Genome Biol* 2009;10:R32.
48. Shen Y, Sarin S, Liu Y, et al. Comparing platforms for *C elegans* mutant identification using high-throughput whole-genome sequencing. *PLoS One* 2008;3:e4012.
49. Walter NA, Bottomly D, Laderas T, et al. High throughput sequencing in mice: a platform comparison identifies a preponderance of cryptic SNPs. *BMC Genomics* 2009;10:379.
50. Suzuki S, Ono N, Furusawa C, et al. Comparison of sequence reads obtained from three next-generation sequencing platforms. *PLoS One* 2011;6:e19534.

Received December 4, 2011. Accepted February 6, 2012.

Reprint requests

Address requests for reprints to: Jean-Michel Pawlotsky, MD, PhD, Professor, Department of Virology, Hôpital Henri Mondor, 51 avenue du Maréchal de Lattre de Tassigny, 94010 Créteil, France. e-mail: jean-michel.pawlotsky@hmn.aphp.fr; fax: (33) 1-4981-4831.

Conflicts of interest

The authors have collectively received research grants from Gilead and Roche. Stéphane Chevaliez has served as an advisor for Gilead and Janssen. Jean-Michel Pawlotsky has served as an advisor for Abbott, Abbott Molecular, Achillion, Anadys, Biotica, Boehringer-Ingelheim, Bristol-Myers Squibb, DebioPharm, Gen-Probe, Gilead, Glaxo-SmithKline, Idenix, Inhibitex, Janssen, Madaus-Rottapharm, Sanofi-Aventis, Schering-Plough/Merck, Novartis, Pfizer, Roche, Vertex, and Virco.

References (Online-Only)

51. Dohm JC, Lottaz C, Borodina T, et al. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Res* 2008;36:e105.
52. Hoppman-Chaney N, Peterson LM, Klee EW, et al. Evaluation of oligonucleotide sequence capture arrays and comparison of next-generation sequencing platforms for use in molecular diagnostics. *Clin Chem* 2010;56:1297–1306.
53. Gilles A, Meglec E, Pech N, et al. Accuracy and quality assessment of 454 GS-FLX titanium pyrosequencing. *BMC Genomics* 2011;12:245.
54. Ewing B, Green P. Base-calling of automated sequencer traces using Phred. II. Error probabilities. *Genome Res* 1998;8:186–194.
55. Ewing B, Hillier L, Wendl MC, et al. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Res* 1998;8:175–185.
56. Richterich P. Estimation of errors in “raw” DNA sequences: a validation study. *Genome Res* 1998;8:251–259.
57. Richter BG, Sexton DP. Managing and analyzing next-generation sequence data. *PLoS Comput Biol* 2009;5:e1000369.
58. Altshuler D, Pollara VJ, Cowles CR, et al. An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature* 2000;407:513–516.
59. Wang C, Mitsuya Y, Gharizadeh B, et al. Characterization of mutation spectra with ultra-deep pyrosequencing: application to HIV-1 drug resistance. *Genome Res* 2007;17:1195–1201.
60. Tsibris AM, Korber B, Arnaout R, et al. Quantitative deep sequencing reveals dynamic HIV-1 escape and large population shifts during CCR5 antagonist therapy in vivo. *PLoS One* 2009;4:e5683.
61. Archer J, Rambaut A, Taillon BE, et al. The evolutionary analysis of emerging low frequency HIV-1 CXCR4 using variants through time—an ultra-deep approach. *PLoS Comput Biol* 2010;6:e1001022.
62. Poon AF, Swenson LC, Dong WW, et al. Phylogenetic analysis of population-based and deep sequencing data to identify coevolving sites in the nef gene of HIV-1. *Molecular Biol Evol* 2010;27:819–832.
63. Gianella S, Delpont W, Pacold ME, et al. Detection of minority resistance during early HIV-1 infection: natural variation and spurious detection rather than transmission and evolution of multiple viral variants. *J Virol* 2011;85:8359–8367.
64. Rodriguez C, Chevaliez C, Bensadoun P, et al. Ultra-deep pyrosequencing, a powerful new tool to study HBV resistance to nucleoside/nucleotide analogues, identifies pre-existing HBV variants bearing resistance mutations and characterizes their on-treatment kinetics. *Hepatology* 2011;52:117A.
65. Chevaliez S, Rodriguez C, Soulier A, et al. Molecular characterization of HCV resistance to telaprevir by means of ultra-deep pyrosequencing: preexisting resistant variants and dynamics of resistant populations. *J Hepatol* 2011;54:S30.
66. Rodriguez C, Soulie C, Marcelin A, et al. HIV-1 coreceptor usage prediction by “ultra-deep pyrosequencing” increases the rate of patients treated with success by CCR5 antagonists. *Antiviral Ther* 2011;16:A100.
67. Chen CF, Lee WC, Yang HI, et al. Changes in serum levels of HBV DNA and alanine aminotransferase determine risk for hepatocellular carcinoma. *Gastroenterology* 2011;141:1240–1248.
68. Chen CJ, Yang HI, Iloeje UH. Hepatitis B virus DNA levels and outcomes in chronic hepatitis B. *Hepatology* 2009;49:S72–S84.
69. Buster EH, Janssen HL. The importance of HBV genotype in HBeAg-positive chronic hepatitis B patients in whom sustained response is pursued. *J Hepatol* 2011;54:395–396.
70. Raimondi S, Maisonneuve P, Bruno S, et al. Is response to antiviral treatment influenced by hepatitis B virus genotype? *J Hepatol* 2010;52:441–449.
71. Woo G, Tomlinson G, Nishikawa Y, et al. Tenofovir and entecavir are the most effective antiviral agents for chronic hepatitis B: a systematic review and Bayesian meta-analyses. *Gastroenterology* 2010;139:1218–1229.
72. Zoutendijk R, Reijnders JG, Brown A, et al. Entecavir treatment for chronic hepatitis B: adaptation is not needed for the majority of naive patients with a partial virological response. *Hepatology* 2011;54:443–451.
73. Pawlotsky JM, Dusheiko G, Hatzakis A, et al. Virologic monitoring of hepatitis B virus therapy in clinical trials and practice: recommendations for a standardized approach. *Gastroenterology* 2008;134:405–415.
74. Chevaliez S, Bouvier-Alias M, Brillet R, et al. Hepatitis C virus (HCV) genotype 1 subtype identification in new HCV drug development and future clinical practice. *PLoS One* 2009;4:e8209.
75. Clark PJ, Thompson AJ. Host genomics and HCV treatment response. *J Gastroenterol Hepatol* 2012;27:212–222.
76. Pawlotsky JM. Treatment failure and resistance with direct-acting antiviral drugs against hepatitis C virus. *Hepatology* 2011;53:1742–1751.
77. Sullivan JC, De Meyer S, Bartels DJ, et al. Evolution of treatment-emergent resistant variants in telaprevir phase 3 clinical trials. *J Hepatol* 2011;54:S4.
78. Zeuzem S, Barnard RJ, Howe JA, et al. Boceprevir resistance-associated variants (RAVS) are observed more frequently in HCV (GT1)-infected patients with poor response to peginterferon α -2b/ribavirin. *J Hepatol* 2011;54:S4–S5.

ARTICLE 6

Rodriguez C, Pawlotsky JM

HCV and HIV drug resistance: exploring the similarities and differences

Clin Infect Dis 2012; in preparation.

RESUME

La lutte pour les ressources est un phénomène qui a débuté dès l'apparition d'organismes reproductifs et dont la description a été initiée par Malthus puis remarquablement synthétisée et étendue à la biologie sous le terme d'évolution par Darwin en 1859 dans « *De l'origine des espèces* ». Si le concept est ancien à l'échelle des sciences biologiques, il continue à caractériser des domaines à l'époque insoupçonnés par son auteur tels que la virologie. En effet, les virus hautement variables tels que le virus de l'immunodéficience humaine (VIH), de l'hépatite B (VHB) et de l'hépatite C (VHC) sont présents sous forme de quasi espèce au sein de leur environnement répliatif, c'est à dire qu'une multitude de virus génétiquement proche mais distincts coexistent au sein de cet espace qu'ils doivent partager selon les mêmes règles générales que les êtres vivants. Ainsi, lorsque des pressions de sélection s'exercent (immunitaires, antivirales...), une redistribution des variants majoritaires est observé au bénéfice de variants minoritaires mieux adaptés à cet environnement changeant. La modélisation mathématique et informatique de la capacité mutationnelle et la dynamique d'adaptation des variants minoritaires au travers de 6 études de cohortes de patients infectés, par la technique ultra-sensible de pyroséquençage haut débit associée à des logiciels originaux ont permis de mettre en évidence, caractériser et évaluer l'impact de marqueurs diagnostics permettant de prédire la résistance aux antiviraux mais aussi de caractériser de nouvelles cibles antivirales.

ABSTRACT

Struggle for resources is a worldwide rule which has been first described by Malthus and extended to whole world of living organisms by Darwin in 1859 in "Origin of species". Today, this concept has been enlarged to virological field, and is particularly adapted to describe highly variable viruses like Human Immunodeficiency Virus (HIV), Hepatitis B Virus (HBV) and Hepatitis C Virus (HCV) which have a quasispecies distribution in infected patients characterized by the co-existence of a number of distinct but related viral populations. Selection pressure on viral replicative environment (immune, antiviral drug treatment...), generally lead to a redistribution of the viral quasispecies with an increasing of the best adapted minor viral variants at the expense of major viral populations. Mathematical and bioinformatic modelization of this phenomenon through 6 infected patients cohorts by means of ultra-deep sequencing and an original bioinformatic package allowed discovery, characterization and evaluation of new diagnostic markers that could be used to prevent resistance emergence to antiviral drugs and to characterized new therapeutics antiviral targets.