



<http://portaildoc.univ-lyon1.fr>

Creative commons : Paternité - Pas d'Utilisation Commerciale -
Pas de Modification 2.0 France (CC BY-NC-ND 2.0)



<http://creativecommons.org/licenses/by-nc-nd/2.0/fr>



UNIVERSITE CLAUDE BERNARD - LYON 1 FACULTE DE PHARMACIE
INSTITUT DES SCIENCES PHARMACEUTIQUES ET BIOLOGIQUES

2020

THESE n° 27

THESE

pour le DIPLOME D'ETAT DE DOCTEUR EN PHARMACIE

présentée et soutenue publiquement le 28 mai 2020 par

Mme CHANUT Marion

Née le 09/02/1995

A Tours

EVALUATION DE LA PLATEFORME DE SEQUENÇAGE A HAUT DEBIT VELA
DIAGNOSTICS POUR LE GENOTYPAGE ET LA DETECTION DE VARIANTES DE
RESISTANCE DU VIRUS DE L'HEPATITE C

JURY

Mme MORFIN SHERPA Florence, Professeur des Universités, Praticien Hospitalier

Mme SCHOLTES Caroline, Maître de Conférences des Universités, Praticien
Hospitalier

Mme FROBERT Emilie, Maître de Conférences des Universités, Praticien Hospitalier

Mme LARRAT Sylvie, Professeur des Universités, Praticien Hospitalier

UNIVERSITE CLAUDE BERNARD LYON 1

- Président de l'Université
- Présidence du Conseil Académique
- Vice-Président du Conseil d'Administration
- Vice-Président de la Commission Recherche
- Vice-Président de la Formation et de la Vie Universitaire

Frédéric FLEURY
Hamda BEN HADID
Didier REVEL
Jean François MORNEX
Philippe CHEVALIER

Composantes de l'Université Claude Bernard Lyon 1

SANTE

UFR de Médecine Lyon Est	Directeur : Gilles RODE
UFR de Médecine Lyon Sud Charles Mérieux	Directrice : Carole BURILLON
Institut des Sciences Pharmaceutiques et Biologiques	Directrice : Christine VINCIGUERRA
UFR d'Odontologie	Directeur : Dominique SEUX
Institut des Sciences et Techniques de Réadaptation (ISTR)	Directeur : Xavier PERROT
Département de formation et centre de recherche en Biologie Humaine	Directrice : Anne-Marie SCHOTT

SCIENCES ET TECHNOLOGIES

UFR Fédération Sciences (Chimie, Mathématiques, Physique)	Directeur : M. Bruno ANDRIOLETTI
UFR Biosciences	Directeur : Mme Kathrin GIESELER
Département composante Informatique	Directeur : M. Behzad SHARIAT
Département composante Génie Electrique et des Procédés (GEP)	Directeur : Mme Rosaria FERRIGNO
Département composante Mécanique	Directeur : M. Marc BUFFAT
UFR de Sciences et Techniques des Activités Physiques et Sportives (STAPS)	Directeur : M. Yannick VANPOULLE
Polytech Lyon	Directeur : M. Emmanuel PERRIN
I.U.T. LYON 1	Directeur : M. Christophe VITON
Institut des Sciences Financières et d'Assurance (ISFA)	Directeur : M. Nicolas LEBOISNE

LISTE DES DEPARTEMENTS PEDAGOGIQUES

**DEPARTEMENT PEDAGOGIQUE DE SCIENCES PHYSICO-CHIMIQUES ET PHARMACIE
GALENIQUE**

• **CHIMIE ANALYTIQUE, GENERALE, PHYSIQUE ET MINERALE**

Monsieur Raphaël TERREUX (Pr)
Madame Julie-Anne CHEMELLE (MCU)
Madame Anne DENUZIERE (MCU)
Monsieur Lars-Petter JORDHEIM (MCU-HDR)
Madame Christelle MACHON (MCU-PH)
Monsieur Waël ZEINYEH (MCU)

• **PHARMACIE GALENIQUE - COSMETOLOGIE**

Madame Marie-Alexandrine BOLZINGER (Pr)
Madame Stéphanie BRIANCON (Pr)
Monsieur Fabrice PIROT (PU-PH)
Monsieur Eyad AL MOUAZEN (MCU)
Madame Sandrine BOURGEOIS (MCU)
Madame Danielle CAMPIOL ARRUDA (MCU)
Madame Ghania HAMDI-DEGOBERT (MCU-HDR)
Monsieur Plamen KIRILOV (MCU)
Madame Giovanna LOLLO (MCU)
Madame Jacqueline RESENDE DE AZEVEDO (MCU)
Monsieur Damien SALMON (MCU-PH)
Madame Eloïse THOMAS (MCU)

• **BIOPHYSIQUE**

Monsieur Cyril PAILLER-MATTEI (Pr)
Madame Laurence HEINRICH (MCU)
Monsieur David KRYZA (MCU-PH-HDR)
Madame Sophie LANCELOT (MCU-PH)
Madame Elise LEVIGOUREUX (AHU)

DEPARTEMENT PEDAGOGIQUE PHARMACEUTIQUE DE SANTE PUBLIQUE

• **DROIT DE LA SANTE**

Madame Valérie SIRANYAN (Pr)
Madame Maud CINTRAT (MCU)

• **ECONOMIE DE LA SANTE**

Madame Nora FERDJAOUI MOUMJID (MCU-HDR)
Monsieur Hans-Martin SPÄTH (MCU-HDR)

• **INFORMATION ET DOCUMENTATION**

Monsieur Pascal BADOR (MCU-HDR)

• **INGENIERIE APPLIQUEE A LA SANTE ET DISPOSITIFS MEDICAUX**

Monsieur Xavier ARMOIRY (PU-PH)
Madame Claire GAILLARD (MCU)

- **QUALITOLOGIE – MANAGEMENT DE LA QUALITE**
Madame Alexandra CLAYER-MONTEMBault (MCU)
Monsieur Vincent GROS (MCU - enseignant contractuel temps partiel)
Madame Audrey JANOLY-DUMENIL (MCU-PH)
Madame Pascale PREYNAT (MCU - enseignant contractuel temps partiel)
- **MATHEMATIQUES – STATISTIQUES**
Madame Claire BARDEL-DANJEAN (MCU-PH-HDR)
Madame Marie-Aimée DRONNE (MCU)
Madame Marie-Paule GUSTIN (MCU-HDR)
Madame Delphine HOEGY (AHU)

DEPARTEMENT PEDAGOGIQUE SCIENCES DU MEDICAMENT

- **CHIMIE ORGANIQUE**
Monsieur Pascal NEBOIS (Pr)
Madame Nadia WALCHSHOFER (Pr)
Monsieur Zouhair BOUAZIZ (MCU-HDR)
Madame Christelle MARMINON (MCU)
Madame Sylvie RADIX (MCU-HDR)
Monsieur Luc ROCHEBLAVE (MCU-HDR)
- **CHIMIE THERAPEUTIQUE**
Monsieur Marc LEBORGNE (Pr)
Monsieur Thierry LOMBERGET (Pr)
Monsieur Laurent ETTOUATI (MCU-HDR)
Monsieur François HALLE (MCU)
Madame Marie-Emmanuelle MILLION (MCU)
- **BOTANIQUE ET PHARMACOGNOSIE**
Madame Marie-Geneviève DIJOUX-FRANCA (Pr)
Madame Anne-Emmanuelle HAY DE BETTIGNIES (MCU)
Madame Isabelle KERZAON (MCU)
Monsieur Serge MICHALET (MCU)
- **PHARMACIE CLINIQUE, PHARMACOCINETIQUE ET EVALUATION DU MEDICAMENT**
Madame Roselyne BOULIEU (PU-PH)
Madame Christelle CHAUDRAY-MOUCHOUX (PU-PH)
Madame Catherine RIOUFOL (PU-PH)
Madame Magali BOLON-LARGER (MCU-PH)
Madame Céline PRUNET-SPANO (MCU)
Madame Florence RANCHON (MCU-PH)
Monsieur Teddy NOVAIS (AHU)

DEPARTEMENT PEDAGOGIQUE DE PHARMACOLOGIE, PHYSIOLOGIE ET TOXICOLOGIE

- **TOXICOLOGIE**
Monsieur Jérôme GUITTON (PU-PH)
Madame Léa PAYEN (PU-PH)
Monsieur Bruno FOUILLET (MCU)
- **PHYSIOLOGIE**
Monsieur Christian BARRES (Pr)
Madame Kiao Ling LIU (MCU)
Monsieur Ming LO (MCU-HDR)

- **PHARMACOLOGIE**
 - Monsieur Sylvain GOUTELLZ (PU-PH)
 - Monsieur Michel TOD (PU-PH)
 - Monsieur Luc ZIMMER (PU-PH)
 - Monsieur Roger BESANCON (MCU)
 - Monsieur Laurent BOURGUIGNON (MCU-PH)
 - Madame Evelyne CHANUT (MCU)
 - Monsieur Nicola KUCZEWSKI (MCU)
 - Madame Dominique MARCEL CHATELAIN (MCU-HDR)
- **COMMUNICATION**
 - Monsieur Ronald GUILLOUX (MCU)
- **ENSEIGNANTS CONTRACTUELS TEMPS PARTIEL**
 - Madame Anne INIGO PILLET (MCU – enseignant contractuel temps partiel)
 - Madame Pauline LOUBERT (MCU – enseignant contractuel temps partiel)

DEPARTEMENT PEDAGOGIQUE DES SCIENCES BIOMEDICALES A

- **IMMUNOLOGIE**
 - Monsieur Guillaume MONNERET (PU-PH)
 - Monsieur Sébastien VIEL (MCU-PH)
 - Madame Morgane GOSSEZ (AHU)
- **HEMATOLOGIE ET CYTOLOGIE**
 - Madame Christine VINCIGUERRA (PU-PH)
 - Madame Sarah HUET (MCU-PH)
 - Monsieur Yohann JOURDY (MCU-PH)
- **MICROBIOLOGIE ET MYCOLOGIE FONDAMENTALE ET APPLIQUEE AUX BIOTECHNOLOGIES INDUSTRIELLES**
 - Monsieur Frédéric LAURENT (PU-PH)
 - Madame Florence MORFIN (PU-PH)
 - Madame Veronica RODRIGUEZ-NAVA (Pr)
 - Monsieur Didier BLAHA (MCU-HDR)
 - Madame Ghislaine DESCOURS (MCU-PH)
 - Madame Anne DOLEANS JORDHEIM (MCU-PH-HDR)
 - Madame Emilie FROBERT (MCU-PH)
 - Monsieur Jérôme JOSSE (MCU)
- **PARASITOLOGIE, MYCOLOGIE MEDICALE**
 - Monsieur Philippe LAWTON (Pr)
 - Madame Nathalie ALLIOLI (MCU)
 - Madame Samira AZZOUZ-MAACHE (MCU-HDR)
 - Madame Camille LOURS (AHU)
 - Madame Amy DERICQUEBOURG (AHU)

DEPARTEMENT PEDAGOGIQUE DES SCIENCES BIOMEDICALES B

- **BIOCHIMIE – BIOLOGIE MOLECULAIRE – BIOTECHNOLOGIE**

Madame Pascale COHEN (Pr)
Madame Caroline MOYRET-LALLE (Pr)
Madame Emilie BLOND (MCU-PH)
Monsieur Karim CHIKH (MCU-PH)
Madame Carole FERRARO-PEYRET (MCU-PH-HDR)
Monsieur Anthony FOURIER (MCU-PH)
Monsieur Boyan GRIGOROV (MCU)
Monsieur Hubert LINCET (MCU-HDR)
Monsieur Olivier MEURETTE (MCU-HDR)
Madame Angélique MULARONI (MCU)
Madame Stéphanie SENTIS (MCU)
Monsieur David GONCALVES (AHU)
Monsieur Alexandre JANIN (AHU)
Madame Nadjat LEBSIR (ATER)

- **BIOLOGIE CELLULAIRE**

Madame Bénédicte COUPAT-GOUTALAND (MCU)
Monsieur Michel PELANDAKIS (MCU-HDR)

- **INSTITUT DE PHARMACIE INDUSTRIELLE DE LYON**

Madame Marie-Alexandrine BOLZINGER (Pr)
Monsieur Philippe LAWTON (Pr)
Madame Sandrine BOURGEOIS (MCU)
Madame Marie-Emmanuelle MILLION (MCU)
Madame Alexandra MONTEMBault (MCU)
Madame Angélique MULARONI (MCU)
Madame Marie-Françoise KLUCKER (MCU – enseignant contractuel temps partiel)
Madame Valérie VOIRON (MCU – enseignant contractuel temps partiel)

Pr : Professeur

PU-PH : Professeur des Universités, Praticien Hospitalier

MCU : Maître de Conférences des Universités

MCU-PH : Maître de Conférences des Universités, Praticien Hospitalier

HDR : Habilitation à Diriger des Recherches

AHU : Assistant Hospitalier Universitaire

Remerciements

Aux membres du jury,

Au Pr Florence Morfin, pour avoir accepté de présider ce jury, mais également pour les cours de virologie dispensés en pharmacie qui ont fait naître mon intérêt pour ce domaine.

Au Dr Caroline Scholtès, pour ton encadrement au cours de ma 5AHU, ta confiance en moi qui a permis de me responsabiliser, ton extrême gentillesse, tes conseils, nos discussions et pour m'avoir accompagnée au cours de la réalisation de cette thèse.

Au Dr Emilie Frobert, pour m'avoir accueillie au cours de mon stage de découverte de recherche en virologie qui m'a permis de confirmer ma passion pour la virologie et pour ton aide au fil de mon cursus.

Au Dr Sylvie Larrat, pour avoir accepté de faire partie de mon jury et pour l'apport de votre expertise dans le domaine.

Aux biologistes, techniciens, internes, thésards et autres personnels de l'Hôpital de la Croix-Rousse qui m'ont aidée au cours de mon stage de 5AHU et qui ont contribué à la réalisation de mon projet de thèse.

A mes proches,

A ma team A1 Laënnec, avec et grâce à qui tout a commencé, Océane, Julian, Johanna, Johana, et notre complicité qui perdure malgré nos chemins plus ou moins différents depuis toutes ces années ;

A Ali, pour notre amitié née grâce à notre 5AHU partagée avec notre super tutrice ;

A l'équipe des soulards, Alex, Matthieu, Loïs, Tom, Pierre-Antoine, Sid, avec qui j'ai partagé tant de soirées, de week-end, de bons moments ;

A mes pharma girls d'amour, Camille, Laurine, Angélique, Steffi, Caro, pour votre présence pendant ces belles années de fac et ces moments de rires et de pleurs partagés ensemble ;

A ma binôme de choc à qui je voudrais faire une mention spéciale, pour notre complémentarité en TP mais pas que... ;

A Audrey, pour nos commérages mais surtout ton soutien inconditionnel et tes conseils pendant ces mois partagés au labo.

A tous mes amis, plus ou moins éloignés de mon monde de pharmacienne : Loïc, Julian, Laurine, Julien, Marine, Joana ainsi que toute la petite équipe Thomas, Mélanie, Romain, Vlada, Aymeric, Raphael, Angélique, Benoit, Marie-Emilie. Merci pour votre amitié.

A ma famille, Papa, Valérie pour m'avoir soutenue depuis le début de mes études et m'avoir aidée à donner le meilleur même dans les moments les plus difficiles ;

A mon petit frère, pour ta bonne humeur motivante ; à Matthieu et Johanna, pour votre soutien depuis Gex ; à ma sœur chérie et Mathis, pour tous nos moments inoubliables partagés ensemble et votre soutien sans faille ;

Et mille fois, à mon amoureux ; grâce à toi j'ai persévéré sans cesse et c'est la raison pour laquelle j'en suis là aujourd'hui. Merci infiniment de m'avoir soutenue, aidée, motivée ; mais surtout de m'avoir supportée.

Merci à vous tous, vous avez joué un rôle important dans mon cursus et vous êtes surtout indispensables à mon bonheur.

Table des matières

Table des matières	8
Liste des tableaux	10
Liste des figures.....	10
Table des abréviations	12
Introduction	14
Rappels Bibliographiques.....	16
1. Généralités sur le Virus de l'Hépatite C.....	16
1.1 Epidémiologie de l'infection par le Virus de l'Hépatite C.....	16
1.1.1 A l'échelle mondiale	16
1.1.2 A l'échelle de l'Europe.....	18
1.1.3 A l'échelle de la France	19
1.2 Le génome du Virus de l'Hépatite C.....	20
1.2.1 Le génome viral	20
1.2.2 La variabilité	22
1.2.3 Les fonctions des protéines virales.....	25
1.3 La lipidation du VHC	28
1.4 Le cycle viral (voir Figure 8)	29
1.5 Les caractéristiques de l'infection par le VHC	32
1.5.1 Les modes de transmission	32
1.5.2 Pathogenèse du VHC.....	34
1.6 Les traitements du VHC : vers une révolution thérapeutique	37
1.6.1 Traitements basés sur l'utilisation d'interféron et de ribavirine	39
1.6.2 Traitements basés sur l'utilisation de nouvelles molécules	42
1.6.3 Recommandations actuelles	46
1.6.4 Mutations de résistance.....	49
2. Génotypage et séquençage du Virus de l'Hépatite C	52
2.1 Le séquençage de Sanger	52
2.1.1 Principe général	52
2.1.2 Apports de la technique vis-à-vis du VHC	57
2.2 Le séquençage de nouvelle génération ou Next Generation Sequencing (NGS)	60
2.2.1 Principe général	60
2.2.2 Technologie 454	63

2.2.3 Technologie Illumina.....	65
2.2.4 Technologie SOLiD.....	67
2.2.5 Technologie Ion torrent.....	71
2.2.6 Caractéristiques du séquençage.....	73
2.2.7 Apports du séquençage de nouvelle génération.....	74
2.3 Analyse comparative entre les techniques de séquençage.....	75
Partie expérimentale.....	79
1. Introduction.....	79
2. Méthodes.....	81
2.1 Echantillons cliniques.....	81
2.2 Séquençage.....	82
2.2.1 Extraction et préparation de la librairie.....	82
2.2.2 PCR en émulsion, enrichissement et séquençage.....	84
2.2.3 Contrôles qualité.....	85
2.2.4 Analyse bio-informatique.....	86
3. Résultats.....	90
3.1 Caractéristiques des patients.....	90
3.2 Comparaison à Sanger.....	92
3.3 Caractéristiques de séquençage.....	96
3.4 Capacité d'amplification des 3 gènes.....	102
3.5 Détermination du génotype.....	105
3.6 Détermination des mutations de résistance.....	111
3.6.1 Mutations dans NS3.....	111
3.6.2 Mutations dans NS5A.....	112
3.6.3 Mutations dans NS5B.....	113
3.7 Cas d'échec de séquençage.....	115
3.8 Cas clinique d'échec de traitement.....	117
4 Discussion.....	120
Bibliographie.....	128

Liste des tableaux

Tableau 1. Efficacité de l'association thérapeutique IFN pégylé – ribavirine selon le génotype.	41
Tableau 2. Caractéristiques des différents antiviraux à action directe selon leur classe.	42
Tableau 3. Spécialités utilisées dans le traitement de l'infection par le VHC.	45
Tableau 4. Substitutions de résistance répertoriées in vitro et/ou chez les patients en échec de traitement, considérant l'AAD et le génotype.	51
Tableau 5. Caractéristiques des différentes techniques de séquençage.	77
Tableau 6. Présentation des avantages et inconvénients des différentes technologies de séquençage par NGS.	78
Tableau 7. Cycles suivis par le thermocycleur pour la synthèse puis l'amplification de l'ADN.	83
Tableau 8. Caractéristiques des patients.	91
Tableau 9. Comparaison des données obtenues par séquençage de Sanger et par NGS, pour chaque région d'intérêt.	94
Tableau 10. Données de séquençage obtenues pour les 29 réactions analysées.	97
Tableau 11. Données de séquençage obtenues pour les 424 échantillons analysés.	99
Tableau 12. Données obtenues pour chaque région d'intérêt.	101
Tableau 13. Bilan d'amplification des 3 gènes d'intérêt.	104
Tableau 14. Données de charges virales selon les génotypes.	110
Tableau 15. Bilan des mutations retrouvées dans chaque région d'intérêt.	114
Tableau 16. Comparaison de la capacité d'amplification des régions d'intérêt des deux trousse Vela Diagnostics.	123

Liste des figures

Figure 1. Incidence de l'infection par le VHC dans la population générale en 2015.	17
Figure 2. Structure du VHC.	21
Figure 3. Arbre phylogénétique des génotypes du VHC.....	22
Figure 4. Distribution géographique des génotypes du VHC dans le monde.	23
Figure 5. Répartition des génotypes du VHC en France.	24
Figure 6. Informations structurales de la protéine NS3 et de molécules inhibitrices anti-NS3. .	26
Figure 7. Modèles de LVP.	28
Figure 8. Modèle d'entrée du VHC.	30
Figure 9. Cycle du VHC.....	31
Figure 10. Histoire naturelle du VHC.	35
Figure 11. Cinétique des marqueurs de l'infection par le VHC, pour une infection aiguë (a) ou chronique (b).	36
Figure 12. Evolution de l'efficacité des traitements anti-VHC disponibles.....	38
Figure 13. Comparaison des cinétiques de concentration d'interféron obtenues selon la forme. 40	
Figure 14. Structures des anti-VHC, selon la protéine cible.	44
Figure 15. Algorithme de prise en charge des patients infectés par le VHC.....	48
Figure 16. Cinétique de persistance des variants associés à une résistance au cours des différents traitements.	50
Figure 17. Représentations structurales du dATP et du ddATP.....	53
Figure 18. Déroulement du séquençage par Sanger, avec marquage radioactif.....	55
Figure 19. Déroulement du séquençage par Sanger, avec marquage par fluorochromes.....	56
Figure 20. Exemples de chromatogrammes.	58
Figure 21. Etapes du séquençage par NGS.	62
Figure 22. Principe du pyroséquençage de la technologie 454.	64
Figure 23. Amplification par "bridge PCR".....	65
Figure 24. Approche de séquençage par synthèse d'Illumina.	66
Figure 25. Principe du séquençage par ligation.	69
Figure 26. Principe de la détermination de séquence par séquençage SOLiD.....	70
Figure 27. Déroulement de la PCR en émulsion.	71
Figure 28. Séquençage sur puce semi-conductrice.	72
Figure 29. Histogramme de distribution de l'âge des patients.	91
Figure 30. Bilan de la distribution des génotypes retrouvés.	106
Figure 31. Distribution des génotypes et sous-types retrouvés.	109

Table des abréviations

AAD	: agents antiviraux à action directe
ADN	: acide désoxyribonucléique
AES	: accident d'exposition au sang
AFEF	: Association Française pour l'Etude du Foie
ALT	: transaminases
AMM	: autorisation de mise sur le marché
ApoB/C1/E	: apolipoprotéine B/C1/E
ARN	: acide ribonucléique
ATP	: adénosine triphosphate
CF-1	: control fragment (fragment contrôle)
CHC	: carcinome hépatocellulaire
CLDN1	: claudine 1
CPA	: cellule présentatrice d'antigène
ddNTP	: didésoxyribonucléotides
dNTP	: désoxyribonucléotides
EC	: extraction control (contrôle d'extraction)
ELISA	: enzyme linked immunosorbent assay (méthode immune-enzymatique)
GAG	: glycosaminoglycane
HSH	: hommes ayant des rapports sexuels avec des hommes
HVR1	: région hypervariable 1
IAI	: Institut des Agents Infectieux
IFN	: interféron
IRES	: internal ribosome entry site
ISG	: interferon-stimulated gene (gène inductible par l'interféron)
kb	: kilobase
LANL	: Los Alamos National Lab
LDL-R	: low-density lipoprotein receptor (récepteur aux lipoprotéines de basse densité)
NCBI	: National Center for Biotechnology Information
NFS	: numération formule sanguine
NGS	: Next Generation Sequencing (séquençage de nouvelle génération)

OCLN : occludine

OMS : Organisation Mondiale de la Santé

PAMPs : pathogen associated molecular patterns (motifs moléculaires associés aux pathogènes)

pb : paire de bases

PCR : polymerase chain reaction (réaction de polymérisation en chaîne)

pegIFN : interféron pégylé

PRR : pattern recognition receptor (récepteur de reconnaissance de motifs moléculaires)

RIG-I : retinoic acid-inducible gene I (gène I inductible par l'acide rétinoïque)

RT-PCR : rétrotranscription – polymerase chain reaction

SC : system control (contrôle système)

SR-BI : scavenger receptor class B type 1

TLR : Toll-like receptor (récepteur de type Toll)

VHA : Virus de l'Hépatite A

VHB : Virus de l'Hépatite B

VHC : Virus de l'Hépatite C

VIH : Virus de l'Immunodéficience Humaine

VLDL-R : very-low density lipoprotein receptor (récepteur aux lipoprotéines de très basse densité)

VMT : Virus de la Mosaïque du Tabac

5'/3'UTR : untranslated region (régions non traduites)

Introduction

Le Virus de l'Hépatite C (VHC), découvert en 1989 (1), fait partie des virus responsables d'hépatites ayant un tropisme primaire pour le foie, ciblant donc les hépatocytes, cellules du foie. L'infection chez le patient peut se manifester par une hépatite aigüe mais devient chronique dans la majorité des cas. Selon les dernières données disponibles de Santé Publique France, anciennement Institut National de Veille Sanitaire (InVs), le taux de mortalité imputable au VHC en France s'élevait à 4,5 pour 100 000 habitants en 2001, représentant ainsi 2646 décès, reflétant l'impact du virus au niveau national. Parmi ces décès, 95% des cas présentaient une cirrhose et 33% un carcinome hépatocellulaire (CHC) compliquant la cirrhose (2).

Avec 71 millions de personnes infectées chroniquement dans le monde aboutissant à 400 000 décès annuels, l'hépatite C est un problème de santé publique majeur. Malgré les mesures de prévention mises en place, 1,75 million de nouveaux cas sont recensés chaque année, en raison de l'absence de vaccin. L'Organisation Mondiale de la Santé (OMS) a défini un objectif d'élimination à l'horizon 2030, c'est-à-dire éviter 90% des nouvelles infections par le VHC et réduire la mortalité due aux infections par le VHC de 65% (3,4). Des programmes de prévention, de dépistage et d'accès aux soins nécessaires pour contrôler la pandémie ont également été mis en place (5).

En termes diagnostiques, la recherche des anticorps anti-VHC dans le sérum du patient ainsi que du génome à ARN du VHC est utilisée pour mettre en évidence une infection par le VHC. La variabilité génétique du VHC est très importante, avec, à ce jour, 7 génotypes différents présentant jusqu'à 30% de divergence entre eux, de nombreux sous-types décrits, ainsi que des quasi-espèces au sein d'un même individu (6,7). Cela explique le fait que la majorité des agents antiviraux à action directe (AAD) développés pour traiter l'infection contre le VHC jusqu'à très récemment n'étaient efficaces que contre certains génotypes, rendant ainsi nécessaire la détermination du génotype et du sous-type avant l'instauration d'un traitement afin d'éviter les échecs sous AAD. En effet, il a été démontré, en 2016, que jusqu'à 12% des échecs sous AAD pouvaient être expliqués par une erreur de génotypage (8). Ces molécules vont cibler des protéines virales particulières impliquées dans le cycle viral, mais le choix du traitement dépend bien souvent du génotype, bien que de nouvelles molécules pan-génotypiques, actives quel que soit le

génotype, soient disponibles depuis peu. Elles doivent impérativement être utilisées sous forme de combinaisons (bi ou multithérapies) afin d'éviter l'apparition rapide, sous quelques jours en monothérapie, de mutations de résistance (9). Ces traitements ont une efficacité de plus de 90% chez les patients infectés (10,11).

Jusqu'à récemment, au laboratoire de virologie de l'Hôpital de la Croix-Rousse à Lyon, le génotypage viral des patients infectés par le VHC était réalisé par séquençage de Sanger après amplification par réaction de polymérisation en chaîne (PCR). Ce principe de séquençage fut développé par Frederick Sanger en 1977 et repose sur l'utilisation de didésoxyribonucléotides marqués par des fluorochromes qui vont agir en tant que terminateurs de chaîne lors de l'amplification (12). Cependant, les techniques de séquençage de nouvelle génération (Next Generation Sequencing NGS), développées au cours des 20 dernières années (13,14), sont attrayantes pour le génotypage et la caractérisation des mutations de résistance notamment pour leur sensibilité, leur capacité à quantifier les variants présents et à détecter des co-infections. Elles pourraient ainsi présenter une méthode de choix pour le séquençage du VHC, notamment en routine.

Au cours de cette thèse, nous avons évalué un automate de séquençage Vela Diagnostics, utilisé par ailleurs en routine pour le séquençage du Virus de l'Immunodéficience Humaine (VIH) au laboratoire de virologie de l'Hôpital de la Croix-Rousse à Lyon, pour son utilisation en routine dans le séquençage, le génotypage et la détection de mutations de résistance chez les patients infectés par le VHC ; nous nous sommes attachés à mettre en lumière l'apport de cette nouvelle technologie vis-à-vis des techniques plus anciennes de séquençage.

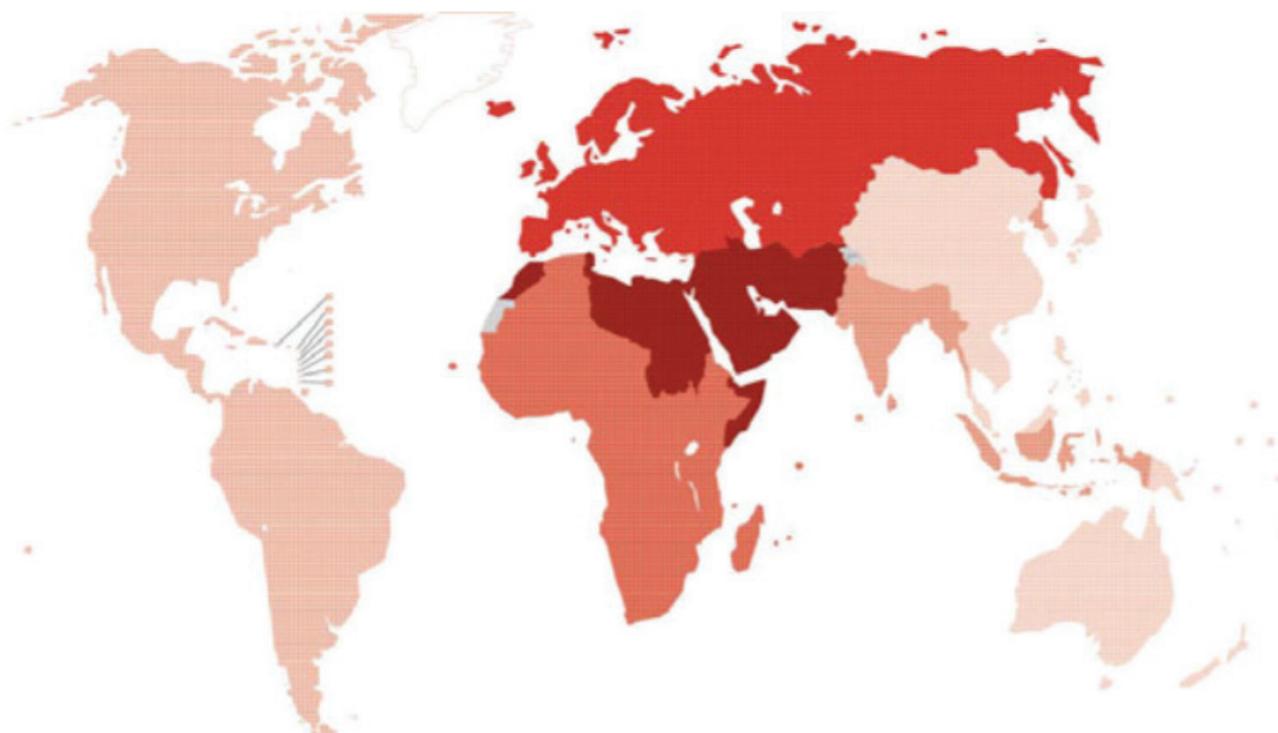
Rappels Bibliographiques

1. Généralités sur le Virus de l'Hépatite C

1.1 Epidémiologie de l'infection par le Virus de l'Hépatite C

1.1.1 A l'échelle mondiale

Selon les estimations de l'OMS de 2015, environ 71 millions de personnes sont porteuses d'une infection par le VHC, représentant 1% de la population et l'incidence annuelle des nouvelles infections est estimée à 1,75 millions (15). Les taux de prévalence les plus élevés sont retrouvés en région méditerranéenne orientale (2,3%) ainsi qu'en Europe (1,5%) (15). Ces zones présentent également un niveau d'infection élevé avec des taux d'incidence de 62,5 pour 100 000 habitants et 61,8 pour 100 000 habitants, respectivement, tandis que les taux d'incidence les plus faibles sont retrouvés dans la région du Pacifique occidental et en Amérique (15) (voir Figure 1).



WHO region	Map key	Incidence of HCV infection			
		Incidence rate (per 100 000)		Total number (000)	
		Best estimate	Uncertainty interval	Best estimate	Uncertainty interval
African Region		31.0	22.5–54.4	309	222–544
Region of the Americas		6.4	5.9–7.0	63	59–69
Eastern Mediterranean Region		62.5	55.6–65.2	409	363–426
European Region		61.8	50.3–66.0	565	460–603
South-East Asia Region		14.8	12.5–26.9	287	243–524
Western Pacific Region		6.0	5.6–6.6	111	104–124
Global		23.7	21.3–28.7	1 751	1 572–2 120

Figure 1. Incidence de l'infection par le VHC dans la population générale en 2015.

Données recueillies par l'OMS concernant l'incidence de l'infection par le VHC selon les régions du monde.

D'après l'Organisation Mondiale de la Santé, 2017 (15).

Les hépatites virales ont mené, en 2015, à 1,34 million de décès, dont 96% étaient le résultat de complications d'hépatites chroniques dues, dans 30% des cas, au VHC (15). Les infections par le VHC ont ainsi entraîné près de 399 000 décès en 2015, pour cause de cirrhose majoritairement, mais également de CHC (15). Contrairement aux infections causées par le Virus de l'Hépatite B (VHB) et le VIH, le VHC est une infection curable, soit spontanément, soit à l'aide d'antiviraux. Les anticorps anti-VHC persistent généralement à vie mais ne sont pas protecteurs. Parmi les patients séropositifs – présentant les anticorps – pour le VHC, on distingue ainsi les patients dits virémiques chez qui l'ARN, génome du VHC, est détectable ; des patients guéris chez qui l'ARN du VHC est indétectable. Parmi les 71 millions de personnes estimées comme infectées par le VHC en 2015, seulement 14 millions d'entre elles, soient 20%, seraient au courant de leur situation et seulement 1,1 million, soient 7%, a débuté un traitement antiviral (15). La couverture mondiale du traitement fut cependant améliorée à 13% en 2016, avec le début de 1,76 million de traitements supplémentaires (16).

Globalement, étant donné que l'estimation annuelle du nombre de nouvelles infections (1,75 million) est supérieure au nombre de personnes décédées pour cause d'infection par le VHC (399 000 personnes) ou guéries de l'infection (843 000 personnes), l'épidémie mondiale pourrait continuer à prendre de l'ampleur en l'absence d'intervention à grande échelle.

1.1.2 A l'échelle de l'Europe

En Europe du Nord, la prévalence de l'infection par le VHC est estimée entre 0,1% et 1% et touche en majorité des adultes de 30 à 50 ans. En Europe centrale, le niveau de prévalence est intermédiaire, allant de 0,2% à 1,2%, tandis qu'en Europe du Sud, il est évalué entre 2,5% à 3,5% (17).

Les travaux de Perz *et al.* (18) ont suggéré que le VHC contribuerait au développement de 38% des cirrhoses en Europe occidentale et 34% en Europe orientale, tandis qu'il serait responsable, respectivement, de 44% et 15% des CHC dans ces régions.

Finalement, l'infection par le VHC serait à l'origine d'un taux de mortalité variable, de 0,1 à 31,5 décès pour 100 000 habitants (19) selon les pays, avec par exemple un taux de 4,5 pour 100 000 habitants en France (20). La mortalité du VHC par CHC suivrait un gradient Est-Ouest, avec des taux de mortalité supérieurs à 3 pour 100 000 habitants en

Europe occidentale et des taux plus faibles, inférieurs à 1 pour 100 000 habitants, en Europe orientale (19). Cela peut être expliqué par un risque plus élevé, à l'Est, de mortalité causée par cirrhose, en raison d'une consommation concomitante d'alcool supérieure. De ce fait, l'infection n'atteint pas le stade du CHC. Également, il est possible qu'à l'Est, les CHC soient sous-diagnostiqués ou classés à tort comme cirrhose, en raison de l'accès limité aux techniques d'imagerie nécessaires.

1.1.3 A l'échelle de la France

En France en 2011, 344 500 personnes ont été estimées comme ayant été en contact avec le VHC soit 0,75% de la population pour laquelle on retrouve les anticorps anti-VHC, avec seulement 57% au courant de cette séropositivité (21,22). Le taux de prévalence est estimé à 0,42% pour les infections virémiques en 2011 (21), et à 0,3% en 2016 (23). Ces valeurs sont abaissées par rapport à celles estimées en 2004, qui étaient respectivement de 0,84% (prévalence des anticorps anti-VHC) et 0,53% (prévalence de l'ARN du VHC) (22).

En outre, l'initiation de traitement est en expansion avec l'arrivée des nouvelles molécules, et pourrait encore tendre à diminuer la prévalence des infections actives et des nouvelles infections. En effet, selon des données du Système national d'information interrégimes de l'Assurance Maladie (Sniiram), le nombre annuel de traitements initiés par ces AAD était seulement de 40 en 2013, et s'est élevé à 8700 en 2014, à 14650 en 2015, pour atteindre environ 60000 traitements cumulés en 2017 (10,24,25).

Concernant la mortalité causée par le VHC en France, une enquête réalisée en 2001 a permis d'estimer que 3618 décès étaient associés à une infection par le VHC, soit un taux de mortalité de 6,1 pour 100 000 habitants, incluant 2646 décès directement imputables au virus, soit un taux de mortalité de 4,5 pour 100 000 habitants dû à une cirrhose ou un CHC (10). D'autre part, il a été montré qu'en 2012, 9% des transplantations hépatiques étaient liées au VHC, alors qu'il était la cause de seulement 3,1% d'entre elles en 2016, reflétant l'impact positif de l'utilisation des AAD sur la maladie hépatique (26).

1.2 Le génome du Virus de l'Hépatite C

1.2.1 Le génome viral

Le VHC est un virus enveloppé, composé d'une molécule d'ARN simple brin de polarité positive. Le génome est protégé par une capsidie icosaédrique elle-même contenue dans une enveloppe lipidique d'origine cellulaire. Le VHC a la caractéristique de présenter une hétérogénéité de morphologie ainsi que de taille, variant entre 40 et 140 nanomètres de diamètre (27,28). D'autre part, il présente une densité particulièrement faible, due à l'association des particules virales avec des lipoprotéines (29). Il appartient à la famille des *Flaviviridae* et au genre des *Hepacivirus* (30). Le génome compte environ 9600 nucléotides, avec deux régions hautement conservées que sont la 5'UTR (untranslated region) et la 3'UTR. Le cadre de lecture ouvert est retrouvé entre ces deux régions non codantes, et contient entre 9030 et 9099 nucléotides selon le génotype viral. Suite à la transcription, la traduction aboutira, grâce à l'initiation par la région IRES (internal ribosome entry site) qui recrute les ribosomes (31), à la production d'un unique précurseur poly-protéique. Le clivage de cette poly-protéine par des protéases virales et cellulaires sera à l'origine de la formation des 10 protéines structurales et non structurales (voir Figure 2). Parmi elles, les protéines core, E1, E2 et p7 sont les protéines structurales impliquées dans la constitution du virion, tandis que NS2, NS3, NS4A, NS4B, NS5A et NS5B sont non structurales et ont un rôle dans la réplication, la morphogenèse virale et la régulation cellulaire, mais ne sont pas constitutives des virions (32).

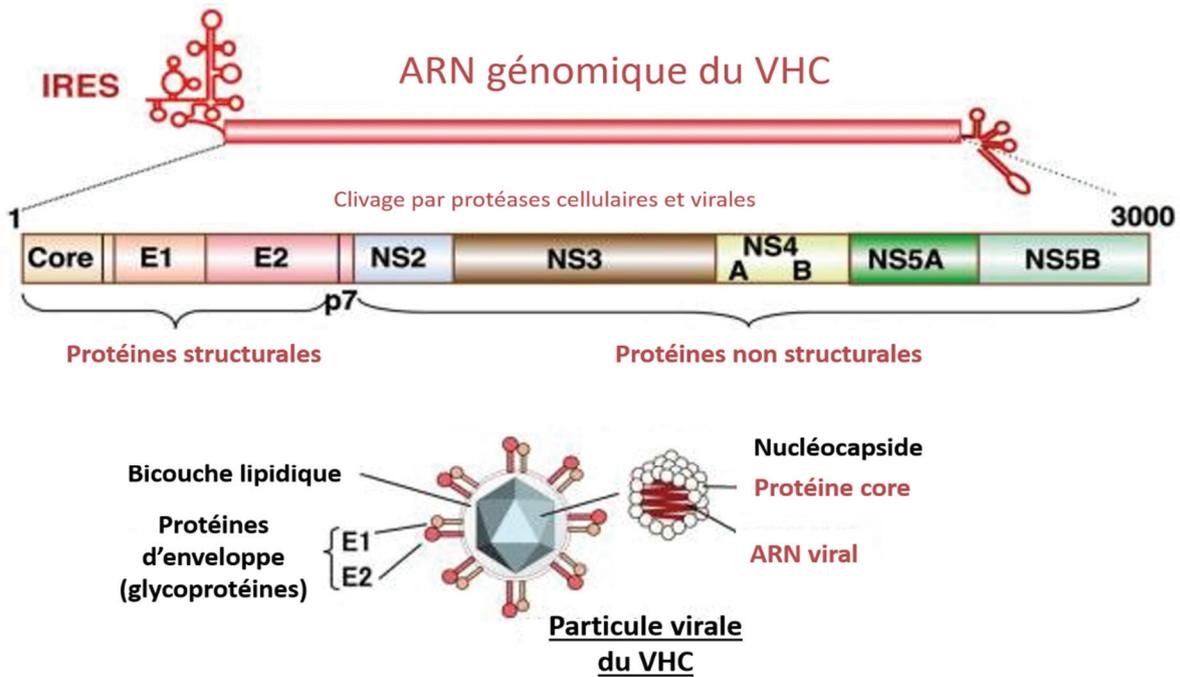


Figure 2. Structure du VHC.

Le VHC code pour une poly-protéine unique constituée d'environ 3000 acides aminés. Les protéines structurales comprenant la protéine core, les glycoprotéines d'enveloppe (E1 et E2) et p7 prenant part à la formation de la particule virale sont distinguées des protéines non structurales (NS2, NS3, NS4A, NS4B, NS5A et NS5B). Adapté de Moriishi et al., 2012 (33).

1.2.2 La variabilité

1.2.2.1 Génotypes et sous-types

Le VHC est un virus présentant une variabilité importante. En effet, la classification du VHC est faite selon 7 génotypes, classés de 1 à 7 (voir Figure 3) ; et de nombreux sous-types sont décrits (6). Il existe plus de 30% de divergence entre les différents génotypes (6). D'autre part, la distinction entre les séquences virales concernant le sous-type peut être faite de la manière suivante : les isolats pour lesquels les séquences diffèrent de moins de 13% sur la région codante appartiennent au même sous-type, tandis que si les séquences diffèrent de plus de 15%, il est considéré que les isolats appartiennent à des sous-types différents.

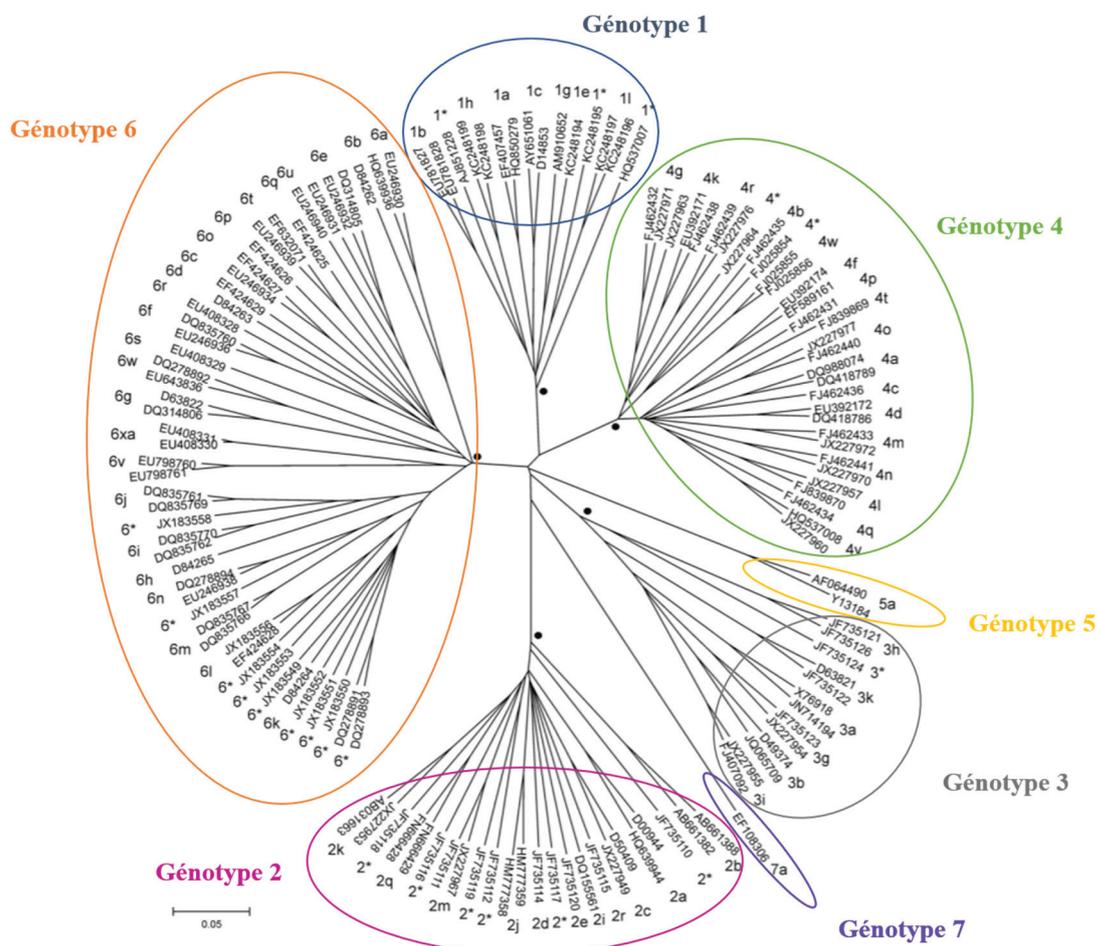


Figure 3. Arbre phylogénétique des génotypes du VHC.

L'arbre représente 129 séquences codantes complètes du VHC, construit selon la méthode de neighbor-joining, mettant en évidence les 7 génotypes du VHC.

Adapté de Smith et al., 2014 (6).

Ces génotypes et sous-types se distribuent de manière préférentielle selon la localisation géographique. Le génotype 1 est ubiquitaire ; il est le génotype le plus fréquemment rencontré, représentant plus de 46% des génotypes infectants, suivi du génotype 3 à 30%. Parmi les cas recensés comme étant infectés par le génotype 1 dans l'étude de Messina *et al.* (34), il s'est révélé que 31% des cas étaient attribuables au sous-type 1a, et 68% au sous-type 1b.

22% des personnes infectées par le VHC le sont par l'un des trois génotypes 2, 4 et 6 tandis que le génotype 5 n'est que rarement rencontré, dans moins de 1% des infections par le VHC. Les génotypes 1 et 3 sont distribués de manière aléatoire dans la plupart des pays, indépendamment du statut économique, alors que les génotypes 4 et 5 touchent davantage les pays à faibles revenus (34). Plus précisément, le génotype 2 domine en Afrique de l'Ouest, le génotype 3 en Asie du Sud, le génotype 4 en Afrique Centrale et en Afrique du Nord, le génotype 5 en Afrique du Sud et le génotype 6 en Asie du Sud Est (voir Figure 4).

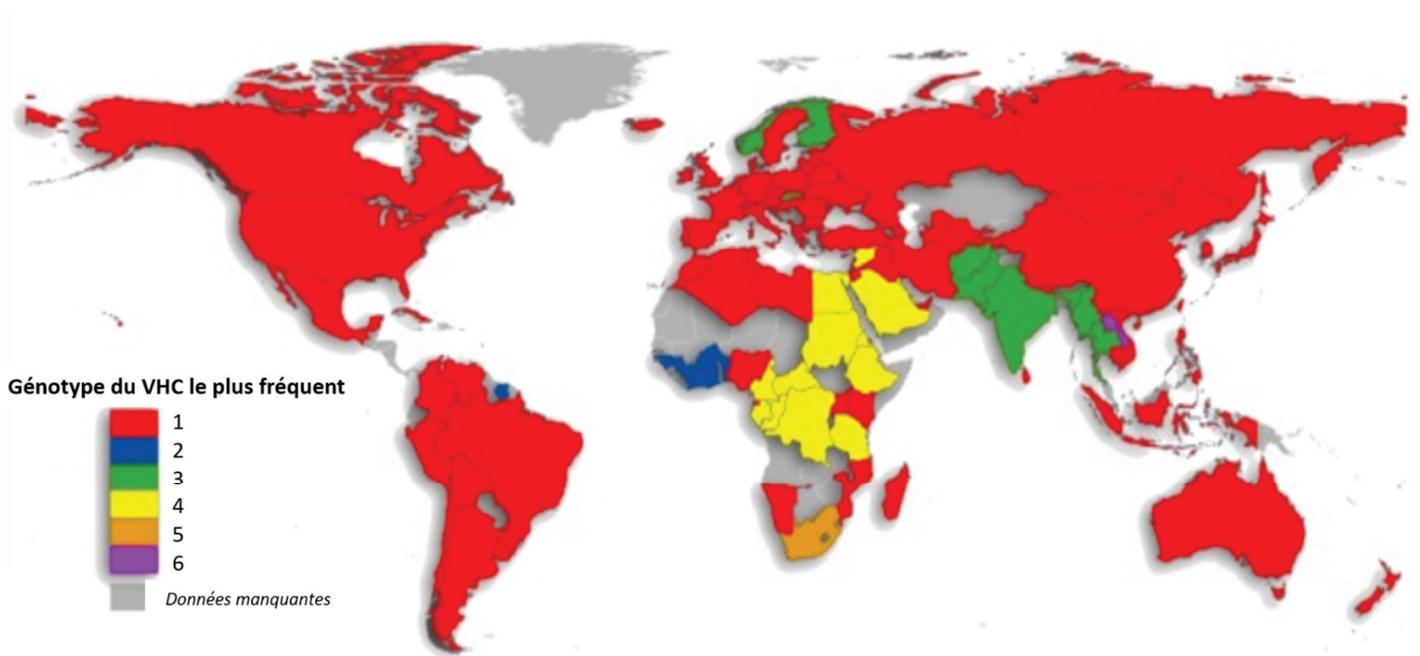


Figure 4. Distribution géographique des génotypes du VHC dans le monde.

La carte représente, pour chaque pays, le génotype du VHC majoritaire. Les pays pour lesquels les données ne sont pas connues apparaissent en gris.

Adapté de Messina et al., 2015 (34).

De manière plus localisée en France, les génotypes prédominants sont le génotype 1b à 30%, le génotype 3a à 22% et le génotype 1a à 16% (voir Figure 5). Certains génotypes sont retrouvés plus volontiers dans une population particulière. Par exemple, le génotype 1b est plus fréquemment retrouvé chez les personnes plus âgées ainsi que les personnes qui présentent un antécédent de transfusion sanguine ou lorsque le mode de transmission virale est inconnu. Le génotype 3a est, lui, plus souvent retrouvé parmi la population toxicomane, plus jeune (35–37).

Hormis ces génotypes bien décrits, des virus recombinants ont également été identifiés. Il s'agit de virus présentant une partie du génome viral correspondant à un génotype voire un sous-type précis et une autre partie à un autre génotype, ou à un autre sous-type du même génotype. Le premier virus recombinant inter-génotypique 2k/1b a été décrit en Russie en 2002 (38) et semble toucher beaucoup de patients originaires de Géorgie (39). Aujourd'hui, au moins 7 autres recombinants inter-génotypiques (2i/6p, 2b/1b, 2/5, 2b/6w, 2b/1b, 2b/1a, 2b/1b) ainsi que 9 recombinants intra-génotypiques (1b/1a, 1b/1a, 4d/4a, 6a/6o, 6e/6h, 6e/6o, 6n/6o et deux recombinants 1a/1c) bien que moins fréquents, sont identifiés (40).

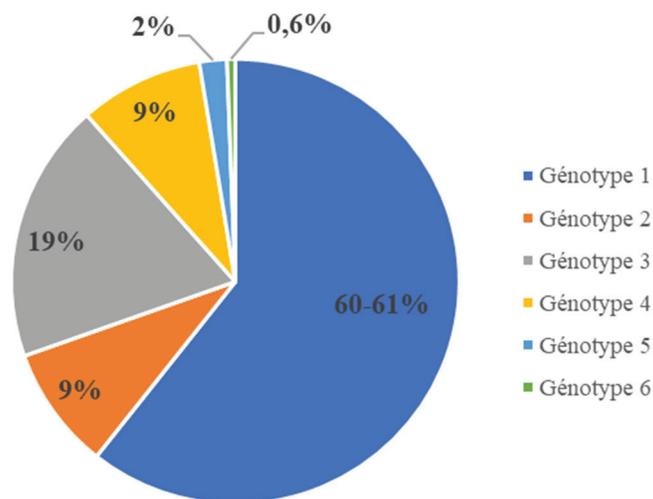


Figure 5. Répartition des génotypes du VHC en France.

Sont représentés chacun des génotypes du VHC ainsi que leur fréquence en France. D'après les données de Brouard et al., 2009 (41).

1.2.2.2 Quasi-espèces

Hormis la variabilité du VHC reflétée par cette classification en géotypes et sous-types, on décrit également des quasi-espèces pour le VHC (42). En effet, le virus code pour une ARN polymérase ARN dépendante (NS5B), utilisée lors de sa réplication. Cette enzyme, ne présentant pas d'activité de correction exonucléasique 3'-5', est en effet responsable d'erreurs aléatoires à un taux élevé (10^{-4} à 10^{-5} par nucléotide répliqué). En lien avec la rapidité de réplication virale qui permet la production de 10^{12} particules virales par jour, cela est à l'origine de modifications de nucléotides dans le génome du VHC au cours des cycles viraux (43,44). Cela aboutit à la formation de nouvelles populations virales formées par ces nouveaux variants viraux infidèlement répliqués, retrouvées en faible proportion au sein d'un même hôte, que l'on appelle ainsi quasi-espèces. Ces variants ne se distinguent que par quelques mutations, conservant une homologie de séquence supérieure à 90% (45).

1.2.3 Les fonctions des protéines virales

La première partie de la poly-protéine obtenue par traduction du génome du VHC est clivée par la signal peptidase cellulaire pour aboutir à la formation de la protéine core, des glycoprotéines E1 et E2 ainsi que de p7. La seconde partie de la poly-protéine est elle aussi clivée du fait des activités protéasiques virales pour donner les différentes protéines non structurales (46). Trois de ces protéines sont plus particulièrement ciblées par les traitements actuels contre le VHC : NS3, NS5A et NS5B. Afin d'échapper à ces traitements, elles sont ainsi susceptibles de développer des mutations, d'où l'importance de leur séquençage.

1.2.3.1 La protéine NS3

NS3, aussi appelée p70, est une protéine multifonctionnelle de 70 kDa composée de 631 acides aminés. La partie C-terminale de la protéine a deux activités ATPase et ARN hélicase, permettant de catalyser la liaison et de dérouler l'ARN viral au cours de la réplication (47). La partie N-terminale présente à la fois des activités ATPase et sérine protéase, responsables de clivages au cours du cycle, notamment en association avec NS4A (47,48). C'est cette activité protéasique qui est ciblée par les molécules « anti-protéases » également appelées anti-NS3 (voir Figure 6).

L'importance de NS3 dans l'assemblage du virus a également été suggérée à travers diverses études. En effet, des mutations dans le domaine hélicase, essentiel au moment de la réplication, perturbent l'assemblage du VHC (49). Il est possible que cette même activité soit impliquée dans l'assemblage de l'ARN viral dans la nucléocapside (50).

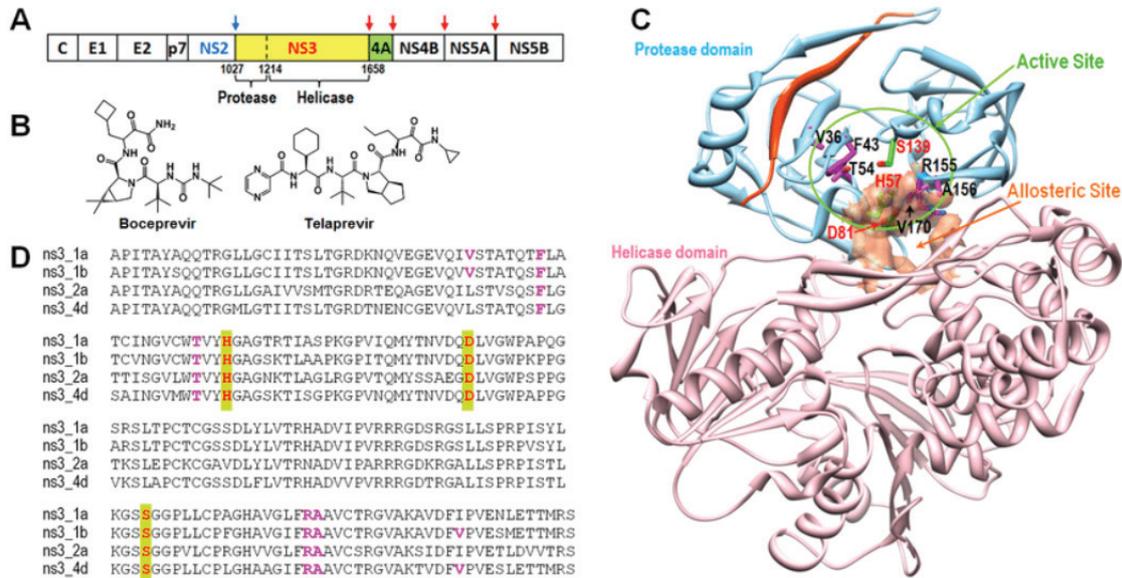


Figure 6. Informations structurales de la protéine NS3 et de molécules inhibitrices anti-NS3.

La protéine NS3 présente une activité protéase ainsi qu'une activité hélicase (A). En (B) sont représentées les structures des deux premières molécules anti-NS3, et en (C) la structure de la protéine NS3 en bleu, et du cofacteur NS4A en rouge. Les résidus actifs de la protéine NS3 (H57, D81 et S139) sont représentés en rouge (C et D) et les positions susceptibles de mutation (V36, F43, T54, R155 et A156) entraînant résistance aux anti-NS3 sont représentées en rose (C et D).

D'après Lee et al., 2013 (51).

1.2.3.2 La protéine NS5A

La protéine NS5A est une phosphoprotéine constituée de 485 acides aminés de 56 kDa ou 58 kDa selon qu'elle se trouve sous sa forme phosphorylée, alors appelée pp56, ou hyper-phosphorylée, pp58. Elle présente, tout comme NS3, plusieurs fonctions : elle est impliquée dans la pathogenèse virale, la réplication virale, l'assemblage viral et la modulation des voies de signalisation cellulaire (52,53). Plus précisément, elle comporte 3 domaines fonctionnels. Tous sont importants au cours de la réplication, avec notamment certains résidus des domaines II et III étant critiques pour cette étape (54). Le domaine III pourrait également avoir un rôle au moment de l'assemblage par interaction avec la protéine core (55,56). Ce sont ces activités de réplication et d'assemblage qui sont visées par les anti-NS5A.

1.2.3.3 La protéine NS5B

La protéine NS5B, ou p68, présente 591 acides aminés, et possède l'activité ARN polymérase ARN dépendante. Elle est ainsi indispensable à la réplication virale, permettant l'initiation de la synthèse virale à partir du brin d'ARN de polarité positive en tant que matrice. Ne possédant pas d'activité exonucléasique 3'-5', de nombreuses mutations sont générées au cours des cycles de réplication. Ces erreurs, associées au fort taux de réplication virale, sont à l'origine de la diversité virale du VHC et notamment de l'existence de quasi-espèces (57).

1.3 La lipodation du VHC

Une caractéristique unique du VHC est la densité particulièrement faible des virions, appelés lipo-viro-particules (LVP) (58). En effet, dans le sérum de patients infectés mais aussi de chimpanzés infectés expérimentalement, les fractions de faible densité contenues entre 1 et 1,10 contiennent la plupart des particules virales (58–60), en raison de leur composition lipidique particulière et spécifique. Cette caractéristique permet de distinguer le VHC des autres virus enveloppés qui ne présentent pas cette caractéristique liée aux lipides, d'où leur densité plus élevée que celle du VHC (50).

En effet, les virions du VHC dérivés d'un patient sont capables de s'associer aux composants des lipoprotéines (voir Figure 7) ; ils peuvent notamment contenir des lipides neutres tels que des triglycérides ou des esters de cholestérol. En outre, plusieurs apolipoprotéines présentes dans la circulation sanguine, comme par exemple l'apolipoprotéine E (apoE), l'apolipoprotéine B (apoB) et l'apolipoprotéine C1 (apoC1), peuvent également s'associer aux particules du VHC (29,58,60,61).

Cette lipodation pourrait prendre place dans le milieu extracellulaire, après sécrétion des particules néo-formées par la cellule infectée (62). Ce processus de lipodation nécessiterait à la fois des composants viraux et notamment la région hypervariable HVR1 de la glycoprotéine E2, les lipoprotéines comme source de lipides mais également des facteurs sériques non lipidiques, comprenant notamment l'albumine sérique (62).

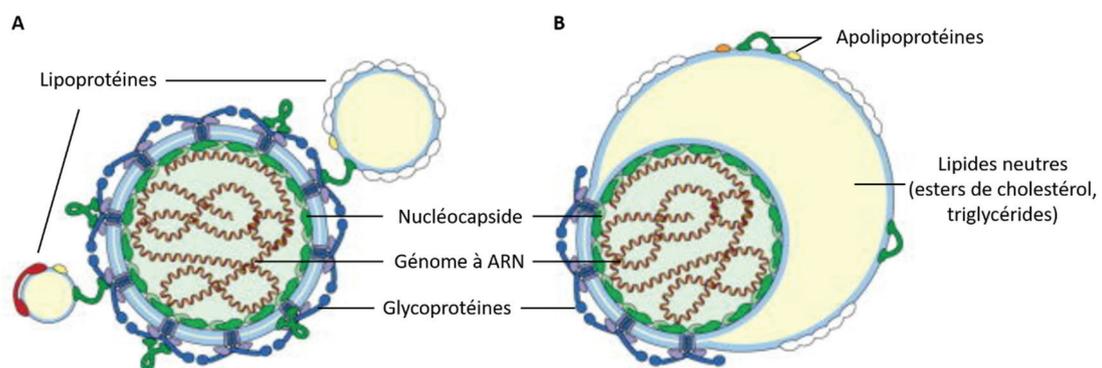


Figure 7. Modèles de LVP.

Deux modèles différents de LVP sont classiquement décrits. Le premier est un modèle où la particule de VHC interagit de manière transitoire avec les différentes lipoprotéines (A). Dans le deuxième modèle, il est proposé que la particule du VHC et la lipoprotéine partagent leur membrane (B).

Adapté de Lindenbach et al., 2013 (50).

1.4 Le cycle viral (voir Figure 8)

Le VHC présente un processus d'entrée particulièrement complexe dans les hépatocytes, cellules cibles. Les virions circulent dans la circulation sanguine et accèdent à ces hépatocytes par leur surface basolatérale. La fixation primaire des particules de VHC aux cellules se fait par une liaison peu spécifique aux protéoglycanes à héparane sulfate (63,64) ou par l'intermédiaire de SR-BI (scavenger receptor class B type 1) (65,66). D'autre part, les particules du VHC peuvent aussi interagir fonctionnellement avec le récepteur des lipoprotéines de faible densité (LDL-R) (67), par interaction avec l'apoE présent sur le VHC (68,69). Il a également été démontré qu'il peut y avoir une interaction avec le récepteur des lipoprotéines de très basse densité (VLDL-R) (70). Globalement, le VLDL-R, le LDL-R et SR-BI pourraient présenter des fonctions redondantes, en raison de leur activité de transfert lipidique (67). Cependant, même si ces récepteurs ont été montrés comme impliqués au cours des étapes d'entrée, leur rôle précis et l'interaction de chacun d'entre eux avec les particules virales et/ou leurs composants lipidiques en surface demeurent mal compris.

Après ces étapes, le récepteur CD81 entre en jeu au cours des étapes suivant l'attachement viral, ainsi que le récepteur de jonction cellulaire, claudine 1 (CLDN1). Leur interaction aboutira à l'internalisation des particules virales avec les complexes CD81-CLDN1 (71) et à la pénétration cellulaire. Une autre protéine à jonction serrée, l'occludine (OCLN), est également essentielle à l'entrée virale (72,73). L'entrée efficace des virions dans les cellules est finalement induite par endocytose clathrine-dépendante (74), suivie de leur transport vers les endosomes où l'acidification pourrait déclencher des changements de conformation des glycoprotéines E1 et E2 induisant la fusion entre les membranes virales et endosomales (75,76).

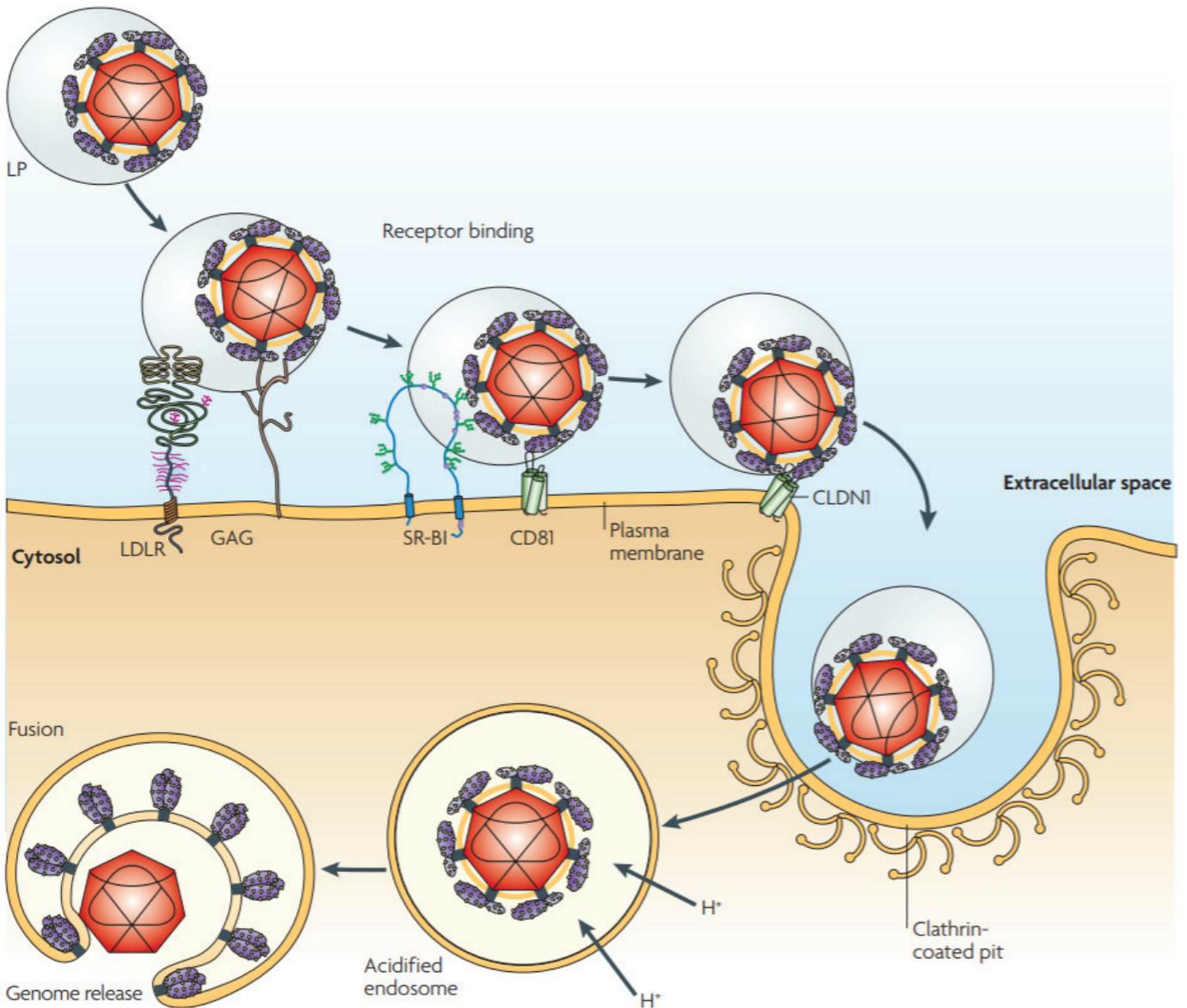


Figure 8. Modèle d'entrée du VHC.

Les particules de VHC sont capables de s'associer aux composants des lipoprotéines (LP) puis s'attachent à la surface des hépatocytes via les glycosaminoglycane (GAG), le LDL-R, le VLDL-R et/ou SR-BI avant d'interagir avec les protéines CD81 et CLDN1. L'internalisation dépend ensuite de l'endocytose médiée par la clathrine. Dans l'endosome, l'acidification induit la fusion membranaire des virions qui aboutit à la décapsulation et à la libération du génome dans le cytosol.

D'après Moradpour et al., 2007 (77).

La capside virale est alors détruite, laissant l'ARN libre dans le cytoplasme. A partir de la molécule d'ARN simple brin se feront, d'une part, la réplication, et d'autre part, la traduction. La réplication se fait grâce à la production du brin de polarité négative par l'ARN polymérase ARN dépendante (NS5B), avec l'ARN simple brin positif comme matrice. La traduction sera ensuite effectuée et donnera le précurseur de la polyprotéine, contenant les dix protéines. Après clivage protéolytique pour donner les différentes protéines, l'assemblage viral a lieu, permettant ainsi, après encapsidation, la formation de nouvelles particules virales, en détournant la voie de synthèse des lipoprotéines. Les virions sont alors libérés sous forme de lipo-viro-particules et peuvent aller infecter de nouveaux hépatocytes (78) (voir Figure 9).

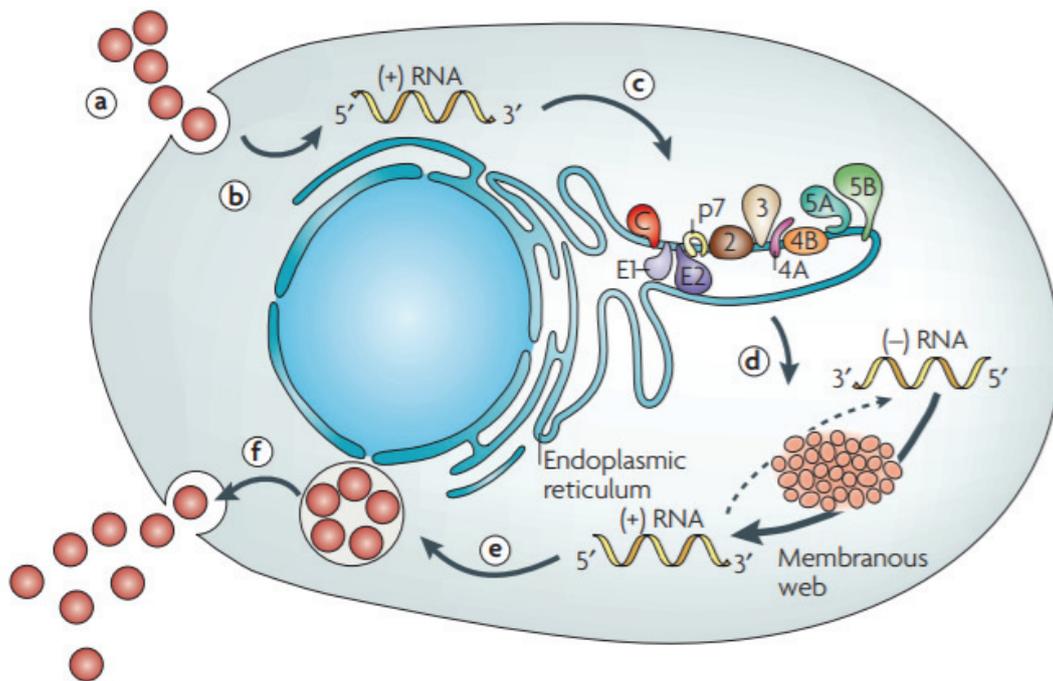


Figure 9. Cycle du VHC.

Le virus se lie aux hépatocytes puis est internalisé (a) ; s'ensuivent la décapsidation et la libération du génome viral dans le cytoplasme (b). La traduction est médiée par l'IRES et aboutit à la polyprotéine qui permettra la formation des protéines virales après clivage (c). L'ARN se réplique (d) ; puis l'assemblage viral et l'encapsidation du génome ont lieu (e) ; avant de relarguer les particules néoformées après maturation (f).

D'après Moradpour et al., 2007 (77).

1.5 Les caractéristiques de l'infection par le VHC

1.5.1 Les modes de transmission

Le VHC est transmis par exposition à du sang contaminé, incluant majoritairement la transfusion de sang et dérivés ou l'utilisation de drogue par injection intraveineuse (voie parentérale). Hormis ces modes de transmission largement décrits, le virus pourrait également être diffusé par des voies moins classiques, par contact sexuel, mais aussi, rarement, de la mère à l'enfant. Ces deux derniers cas sont retrouvés en cas de lésions, impliquant ainsi des contacts sanguins.

1.5.1.1 Les drogues

Au niveau mondial, la principale source de contamination par le VHC est représentée par la toxicomanie intraveineuse, soit l'injection de drogues. En effet, au cours de ces pratiques, la probabilité de transmission du VHC par aiguille souillée est très importante. Une étude de 2017 a rapporté l'usage de drogues par injection dans 179 pays et a estimé à 15,6 millions le nombre de personnes suivant ces pratiques (79), avec une grande majorité d'hommes. Parmi les drogues injectées dans ces pays ont été recensés notamment des opioïdes à 82,9% et des stimulants à 33%.

Globalement, cela porte à 52,3% la prévalence des anticorps anti-VHC dans cette population, avec des variations géographiques (79). En France, elle fut estimée à 59,8% par l'enquête Coquelicot réalisée en 2004 (80), puis revue à la baisse, à 44%, en 2011 (81) suggérant l'efficacité de la politique de réduction des risques, incluant l'accès à des seringues propres afin d'éviter le partage de matériel. Cette source de contamination reste le mode principal de transmission du VHC et la prévalence de l'infection chronique par ce virus dans ces populations est sensiblement plus élevée que celle du VHB ou du VIH (82).

1.5.1.2 La transfusion

La transfusion de sang ou de produits stables dérivés du sang (fractions coagulantes, immunoglobulines) a été l'une des premières causes décrites de transmission du VHC et a joué un rôle important dans la dissémination du virus jusqu'à la fin du 20ème siècle. De cette manière, de nombreux patients hémophiles, hémodialysés, transplantés qui ont été transfusés à cette période ont été contaminés par le VHC. Cependant, la mise en place de

mesures de dépistage des donneurs dès 1990 a permis la réduction de la dissémination du VHC par cette voie (83). A l'heure actuelle, le risque résiduel de propagation par transfusion est considéré comme infime, de 1 pour plus de 500 000 dons, engendrant 5 hépatites post-transfusionnelles par an en France (36). Cependant, l'infection par le VHC pouvant rester silencieuse de nombreuses années, des personnes transfusées au 20^{ème} siècle ont été contaminées mais n'ont toujours pas été dépistées.

1.5.1.3 Autres contacts avec du sang

Le VHC peut également être transmis de façon nosocomiale par le personnel soignant, ou par du matériel, des objets souillés par le virus. Aujourd'hui, ce mode de dissémination est réduit du fait des mesures hygiéniques et de l'utilisation de matériel à usage unique mais il a été important à l'époque où les méthodes n'étaient pas efficaces pour désinfecter le matériel.

Une effraction cutanée ou des actes invasifs tels qu'une chirurgie peuvent aussi être à l'origine de la transmission du VHC lorsque des instruments contaminés sont utilisés. Il est possible que l'acupuncture, le piercing ou les tatouages soient aussi une source de transmission. Des études ont en effet montré une association significative entre la présence d'anticorps anti-VHC et des antécédents de tatouage (84) ou d'acupuncture (85). Le risque de transmission est également présent chez les soignants, par blessure accidentelle notamment après un accident d'exposition au sang (AES). Le risque pour le soignant d'être infecté par du matériel souillé est estimé entre 3 et 5% lorsque le sang présent sur le matériel contient en effet du virus, et peut s'élever encore davantage si la virémie est élevée (36).

1.5.1.4 La transmission sexuelle

En Europe, mais également en Australie ainsi qu'aux Etats-Unis, la transmission sexuelle du VHC a été rapportée pour des hommes séropositifs pour le VIH ayant des rapports sexuels avec des hommes (86).

Des rapports hétérosexuels avec un partenaire infecté et avec plusieurs partenaires ont été identifiés comme étant des facteurs de risque de transmission du VHC (87), bien que la transmission sexuelle du VHC soit beaucoup moins efficace que celle d'autres virus transmis par cette voie.

1.5.1.5 La transmission mère-enfant

La transmission verticale du VHC de la mère à l'enfant au moment de l'accouchement est également décrite bien que rare (88). Cette infection par transmission périnatale se produirait chez 2,7 à 8,4 % des nourrissons nés de mères infectées par le VHC (89), et pour une proportion encore plus élevée pour les nourrissons nés de mères co-infectées par le VIH et le VHC (90).

1.5.2 Pathogenèse du VHC

Après infection par le VHC, une clairance spontanée médiée par le système immunitaire est possible. Dans ce cas de figure, l'infection sera simplement répertoriée comme hépatite C aiguë. Lorsque le génome viral est détectable depuis plus de 6 mois chez le patient, on parle d'hépatite C chronique. Ce passage à la chronicité est rapporté dans 70% des cas (91). L'infection pourra alors persister si le patient n'est pas traité.

1.5.2.1 Evolution de l'infection

Classiquement, la primo-infection par le VHC est asymptomatique, d'où une difficulté de diagnostic. En cas d'infection symptomatique, le patient peut présenter de la fatigue, de la fièvre, des douleurs abdominales, des nausées et possiblement un ictère, plus spécifique d'une atteinte hépatique (16). Dans le cas où l'hépatite se développe de manière chronique, des dommages hépatiques peuvent être observés, incluant une stéatose, accumulation anormale de graisses au niveau des hépatocytes ou une fibrose, niveau anormalement élevé de tissu conjonctif fibreux, dans 15 à 25% des cas (91), dont les symptômes peuvent n'être ressentis qu'après plusieurs dizaines d'années d'évolution silencieuse. Un CHC peut également apparaître, chez les patients atteints ayant développé de la cirrhose. Parmi cette dernière population, le risque est estimé à 3% par an. L'apparition du cancer est la dernière étape possible de l'évolution du virus (voir Figure 10).

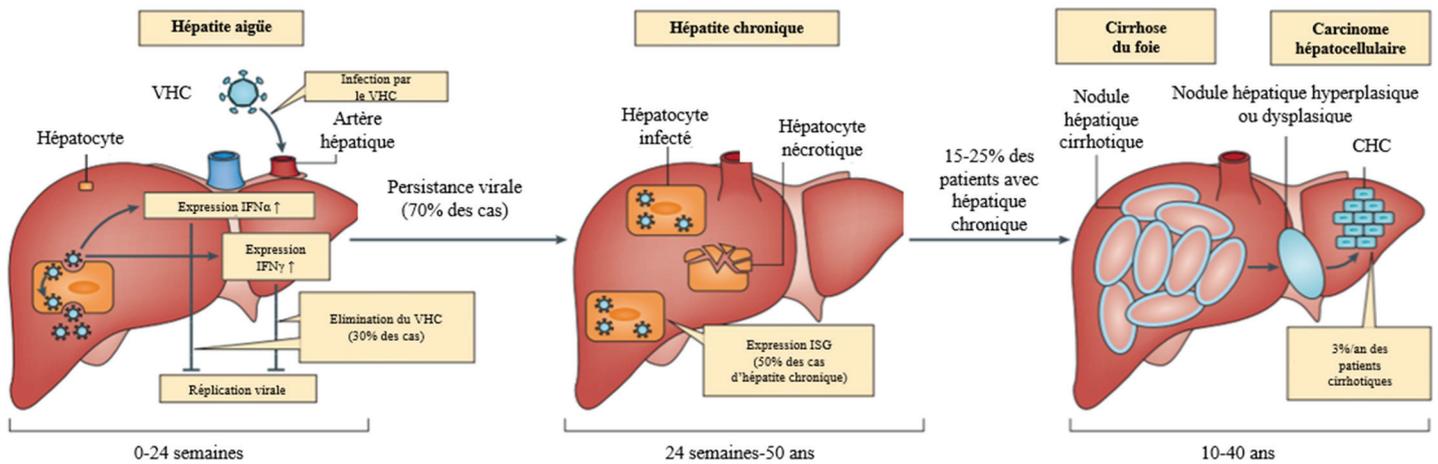


Figure 10. Histoire naturelle du VHC.

Le VHC peut provoquer une infection aigüe d'une durée inférieure à 6 mois au cours desquels l'interféron (IFN) sera exprimé pour son effet antiviral, permettant la guérison dans 30% des cas. 70% des patients infectés vont développer une infection chronique et selon la capacité des hépatocytes à contrer l'infection, notamment grâce aux gènes inductibles par l'IFN (ISG), ils pourront être à risque de développer une fibrose, une cirrhose voire un CHC.

Adapté de Heim et al., 2013 (91).

1.5.2.2 Signes biologiques et moléculaires de l'infection par le VHC

Le premier marqueur de l'infection qui apparaît chez le patient est l'ARN du virus dans le sérum, détectable par RT-PCR (transcription inverse de l'ARN en ADN puis amplification par PCR) dès une à trois semaines après contamination par le VHC (92). A partir de la troisième semaine et jusqu'à 12 semaines après exposition au VHC, les transaminases vont s'élever dans le sang (93), signe de la lyse des hépatocytes. Plus tardivement, environ 3 mois en moyenne après l'infection, seront détectables les anticorps anti-VHC, notamment par technique ELISA (enzyme-linked immuno sorbent assay) (voir Figure 11a). Cependant, en cas de co-infection avec le VIH notamment, les anticorps peuvent apparaître de façon retardée, plus d'un an après le contagé (94). Après 6 mois de portage de l'ARN viral sans éradication, il est possible que le patient présente un ou plusieurs symptômes précédemment énoncés. Le taux de transaminases est alors variable, reflétant les lésions hépatiques, et peut fluctuer (voir Figure 11b). Sans traitement, l'infection peut persister à vie et avoir des conséquences hépatiques.

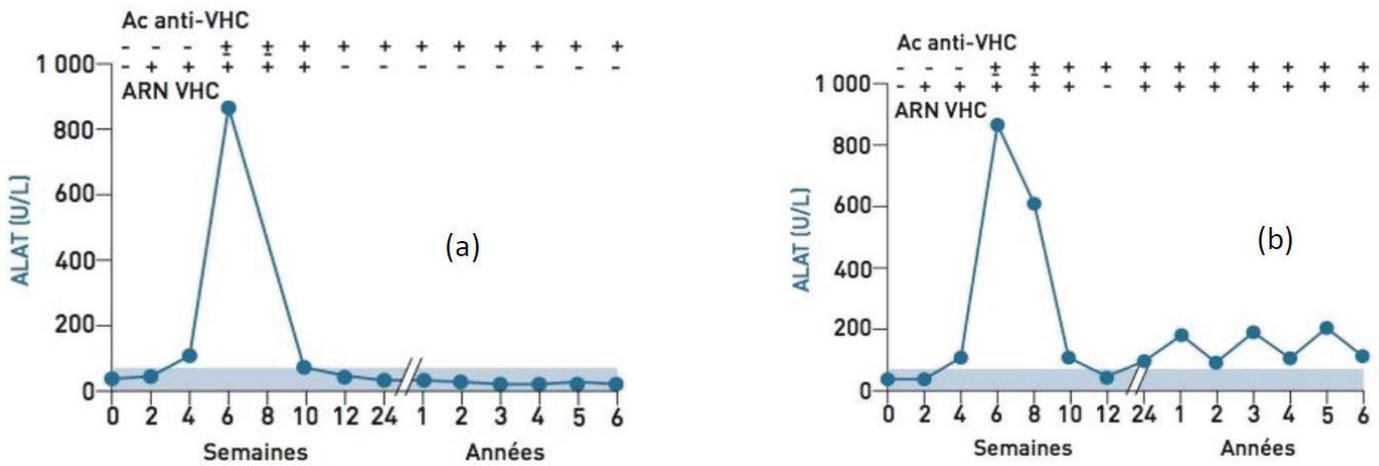


Figure 11. Cinétique des marqueurs de l'infection par le VHC, pour une infection aigüe (a) ou chronique (b).

Sont représentés les niveaux de transaminases (ALT), ainsi que la détectabilité de l'ARN viral (ARN VHC) et des anticorps anti-VHC (Ac anti-VHC) dans le cas d'une infection aigüe (a) ou chronique (b) par le VHC.

Selon Santé Publique France, 2007 (95).

1.5.2.3 Réponse immunitaire innée

Chez les patients infectés par le VHC, l'immunité innée représente la première barrière antivirale et va permettre la production de cytokines et de différents types d'IFN.

Les particules virales et leurs composants, comprenant des motifs moléculaires associés aux pathogènes (pathogen associated molecular patterns PAMPs) vont en effet être reconnus par des récepteurs particuliers spécifiques (pattern recognition receptor PRR). Plus précisément, l'ARN viral du VHC, certaines parties spécifiques du génome ou encore les protéines virales vont être reconnus par des récepteurs de type Toll (Toll-like receptors) et notamment les TLR2, 3, 7 et 8 (96). Dans la cellule, le gène RIG-I (retinoic acid-inducible gene I) joue aussi un rôle de reconnaissance du génome viral (97). L'interaction des composants viraux avec ces récepteurs variés est à l'origine de l'activation de voies intracellulaires spécifiques par des cascades d'évènements, qui vont aboutir à la synthèse et à la sécrétion d'IFN, de cytokines mais aussi de gènes inducibles par l'IFN (ISGs), qui vont avoir une action antivirale. Cette réponse médiée par l'IFN est effective rapidement, pendant 4 à 10 semaines suivant l'infection par le VHC (91).

1.5.2.4 Réponse immunitaire adaptative

La réponse immunitaire innée est ensuite suivie de la réponse immunitaire adaptative, qui se décompose en réponse cellulaire et en réponse humorale.

Il a en effet été montré que l'élimination du VHC est associée à la présence de lymphocytes T à la fois cytotoxiques (lymphocytes T CD8+) ou non (lymphocytes T CD4+) qui vont reconnaître des épitopes sur les protéines virales (98). Ces lymphocytes entraîneront l'apoptose des cellules infectées ainsi que la sécrétion de cytokines, d'action antivirale. Ces effecteurs correspondent à la réponse cellulaire. Si elle est efficace, elle peut résulter en une clairance virale ; à l'inverse, si elle n'est pas assez importante, la virémie sera maintenue et l'infection pourra progresser vers la chronicité.

La réponse humorale va également jouer un rôle prépondérant. Les antigènes viraux vont être reconnus puis présentés par les cellules présentatrices d'antigène (CPA) pour permettre la différenciation des lymphocytes B en plasmocytes capables de sécréter des anticorps spécifiques aux antigènes du VHC. Notamment, ces anticorps rendront possible la neutralisation des particules virales circulantes. De manière importante, la région HVR1 de la glycoprotéine E2 a été révélée comme étant une cible majeure d'anticorps neutralisants (99). Cependant, le VHC étant un virus à ARN, il présente une capacité importante à muter afin d'échapper à la réponse immunitaire et notamment aux anticorps.

1.6 Les traitements du VHC : vers une révolution thérapeutique

Depuis la découverte du VHC, les traitements ont considérablement évolué. En effet, les infections ont d'abord été traitées par IFN α , en associant ou non la ribavirine avec un taux de guérison autour de 40% (100). Hormis cette efficacité moyenne, les traitements étaient longs et présentaient de nombreux effets indésirables chez le patient traité.

Depuis 2011, une nouvelle génération de traitements plus efficaces et mieux tolérés est disponible, avec le développement d'inhibiteurs de la protéase ou anti-NS3, d'inhibiteurs de NS5A et d'inhibiteurs de la polymérase ou anti-NS5B. Du fait de la grande variabilité du VHC, la majorité des molécules disponibles ne sont efficaces que contre quelques génotypes ; afin d'obtenir la meilleure efficacité possible, il est donc important de déterminer le génotype *a priori* pour choisir un traitement adapté.

Cependant, depuis 2018, de nouvelles molécules pan-génotypiques, c'est-à-dire actives sur l'ensemble des géotypes, sont disponibles afin de traiter les infections par le VHC quel que soit le géotype infectant.

Globalement, le traitement vise à inhiber la réplication virale afin d'obtenir une diminution puis une suppression totale et maintenue de la charge virale dans le sérum du patient, caractérisant ainsi la guérison du patient à 3 mois après la fin du traitement. A l'heure actuelle, il est possible de guérir plus de 95% des patients avec des traitements d'une durée de 8 à 12 semaines seulement (16), soulignant l'efficacité de ces traitements par rapport aux traitements antérieurs (voir Figure 12).

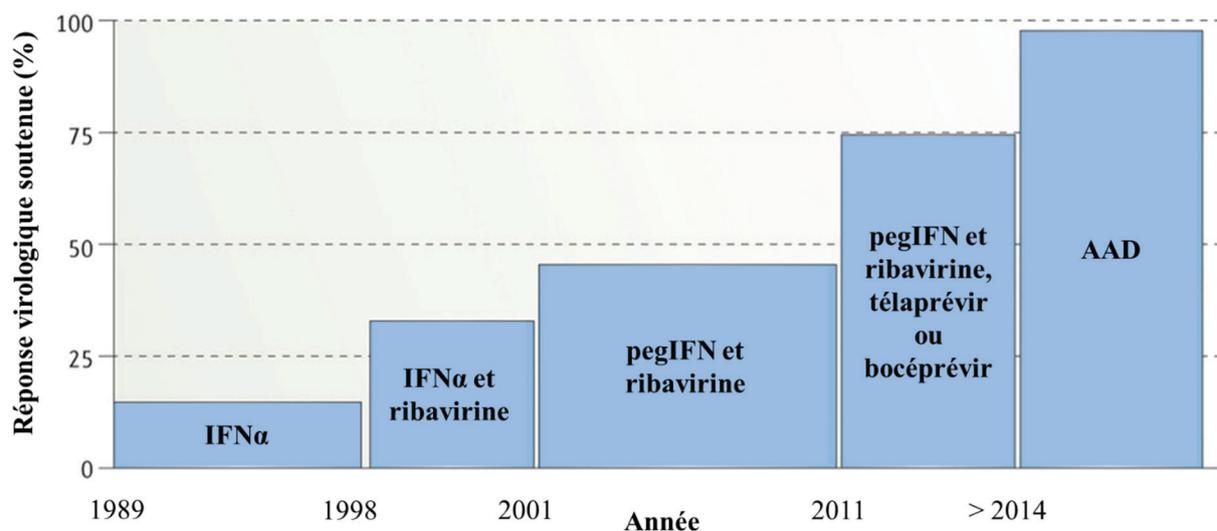


Figure 12. Evolution de l'efficacité des traitements anti-VHC disponibles.

Le graphique représente, par type de traitement, le niveau de réponse virologique obtenu estimé par la charge virale. IFN : interféron. pegIFN : interféron pégylé. AAD : agents antiviraux à action directe.

Adapté de Moradpour et Müllhaupt, 2015 (101).

1.6.1 Traitements basés sur l'utilisation d'interféron et de ribavirine

1.6.1.1 L'interféron

Les IFN sont de petites molécules protéiques produites par les cellules pour répondre à des infections, notamment virales. Dans le cadre d'une infection par le VHC sont surtout utilisés des IFN de classe α produits par génie génétique. Bien que leur mode d'action ne soit pas clairement déterminé, ils permettraient d'inhiber la réplication virale par stimulation d'un certain nombre de voies de signalisation et ainsi de production protéique. L'IFN α , utilisé en monothérapie, ne permettait d'atteindre que des taux faibles de réponse virologique à long terme, autour de 15%, avec 3 injections par semaine (100). D'autre part, le traitement présentait des effets indésirables, et notamment un syndrome pseudo-grippal fréquent, se manifestant par une fièvre, des frissons, des maux de tête, des courbatures. Deux laboratoires produisaient des spécialités différentes, l'Introna® et le Viraféron® (Merck & Co) ainsi que le Roféron® (Roche) (102).

1.6.1.2 L'interféron pégylé

Une forme pégylée de cet IFN α fut ensuite utilisée, composée de l'IFN classique conjugué à un polyéthylène glycol. Cette nouvelle molécule a permis une augmentation de 10 fois de la demi-vie, atteignant 40 heures, par diminution de la clairance de l'IFN. De ce fait, les concentrations retrouvées dans le plasma des patients traités étaient davantage stabilisées et équilibrées dans le temps (voir Figure 13), réduisant la posologie à une unique injection hebdomadaire et limitant certains effets secondaires provoqués, pour la forme non pégylée, par des pics de concentration importante d'IFN. En outre, cette forme permet une diminution de l'immunogénicité, limitant la réponse anticorps contre le traitement, d'où une meilleure tolérance (100).

Cependant, en monothérapie, la réponse prolongée était obtenue chez moins de 20% des patients. Merck & Co a élargi la production au Viraféronpeg® et Roche a ainsi produit le Pegasys®.

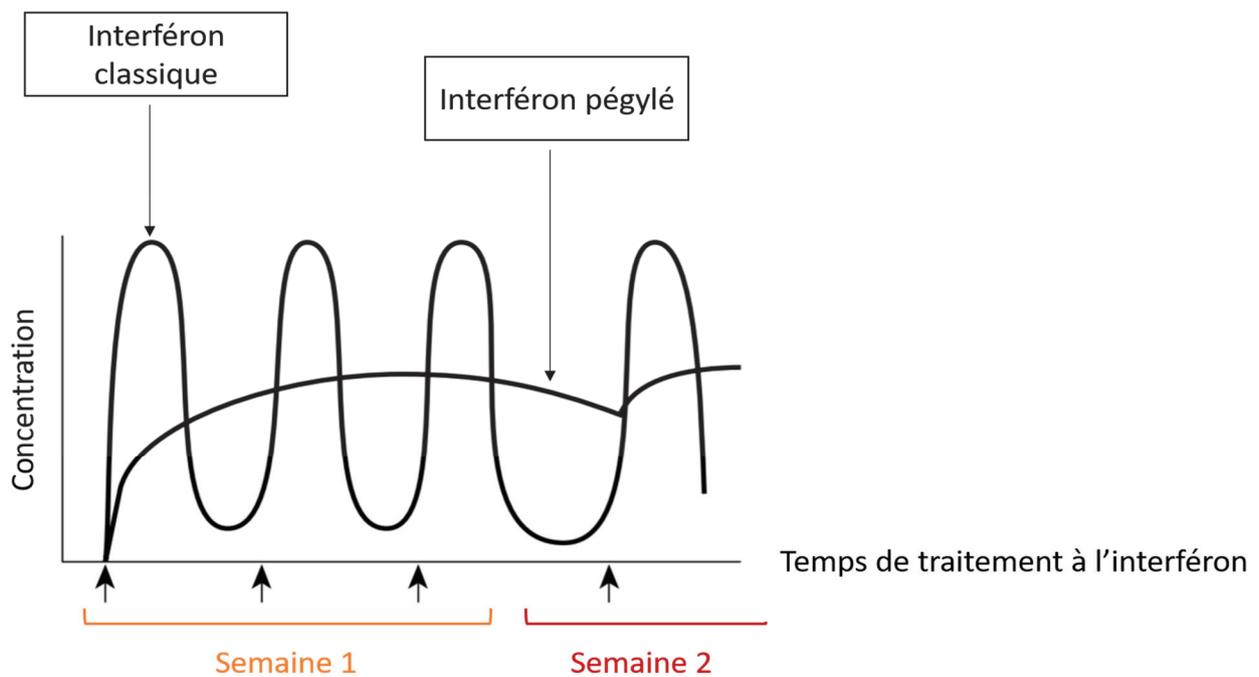


Figure 13. Comparaison des cinétiques de concentration d'interféron obtenues selon la forme.

La concentration d'IFN est représentée dans le temps, après les différentes injections indiquées par les flèches de l'axe des abscisses. Les cinétiques de l'IFN classique sont observées pour les 3 injections par semaine, tandis que les cinétiques de l'IFN pégylé sont obtenues avec la première injection de chaque semaine de traitement.

Adapté de Boyer et al., 2002 (100).

1.6.1.3 La ribavirine

L'IFN α fut ensuite utilisée en bithérapie avec la ribavirine. Il s'agit d'un analogue nucléosidique, agissant probablement comme inhibiteur de la polymérase virale, affectant alors la production des protéines virales du VHC mais ayant également montré son action contre d'autres virus. Par exemple, son insertion dans les génomes en cours de synthèse peut provoquer des mutations ou des erreurs d'appariement entre les bases nucléotidiques (103).

Aux effets indésirables provoqués par l'IFN s'ajoutent ceux possiblement induits par la ribavirine au cours d'une bithérapie. Le principal effet décrit lors de la prise de ribavirine est une anémie, soit une diminution du taux des globules rouges dans le sang, exprimée par une pâleur, une fatigue, un essoufflement et/ou une augmentation du rythme cardiaque.

L'efficacité directe de la ribavirine en monothérapie reste modeste ; l'intérêt de cette molécule dans le cadre de l'infection par le VHC réside dans la mise en place d'une bithérapie alliant la ribavirine à l'IFN pégylé dont elle potentialise l'effet antiviral.

Cette association a permis d'augmenter la réponse virologique autour de 40% pour les génotypes 1, et cette bithérapie fut longtemps le traitement de référence pour les patients infectés par le VHC (100). Cependant, l'efficacité de cette association thérapeutique présentait une grande variabilité selon le génotype, rendant le génotypage d'autant plus important (voir Tableau 1).

Tableau 1. Efficacité de l'association thérapeutique IFN pégylé – ribavirine selon le génotype.

D'après Manns et al., 2001 (102).

Association pegIFN + ribavirine pendant 48 semaines	Génotype 1	Génotypes 2/3	Génotypes 4/5/6
Niveau de réponse virologique soutenue obtenu à 24 semaines	34-42%	80-82%	33-50%

1.6.2 Traitements basés sur l'utilisation de nouvelles molécules

Malgré l'avancée qu'a permis cette bithérapie, une partie importante des patients ne répondait pas au traitement ou présentait une rechute. Au vu de ces constats et grâce aux nouveaux moyens d'étude *in vitro* du virus, de nouvelles molécules sont apparues sur le marché afin de traiter le VHC de manière plus efficace. En effet, l'IFN est resté pendant de nombreuses années la seule option thérapeutique du fait de la difficulté à cultiver le VHC *in vitro* pour la mise en place de nouveaux traitements.

Ces molécules visent des étapes spécifiques du cycle viral (104) et plus précisément des protéines virales impliquées au cours du cycle du VHC, par interaction directe, en lien avec leurs diverses structures (voir Figure 14), et sont ainsi dénommées AAD. Les différentes classes se distinguent quant à leur activité vis-à-vis des génotypes du VHC, leur niveau de barrière aux résistances, leurs effets indésirables ainsi que le potentiel d'interactions (voir Tableau 2). Ces molécules peuvent être utilisées seules ou en association, selon les recommandations (voir Tableau 3 et Rappels bibliographiques, 1.6.3 *Recommandations actuelles*).

Tableau 2. Caractéristiques des différents antiviraux à action directe selon leur classe.

Adapté de Liang et Ghany, 2016 (105).

	Inhibiteurs de la protéase NS3	Inhibiteurs de NS5A	Inhibiteurs nucléotidiques de la polymérase NS5B	Inhibiteurs non nucléotidiques de la polymérase NS5B
Efficacité	Elevée	Elevée	Elevée	Faible/intermédiaire
Couverture génotypique	Etroite	Intermédiaire	Large	Etroite
Probabilité de développement de résistance	Elevée	Intermédiaire	Faible	Elevée
Risque d'effets indésirables	Substantiel	Faible/intermédiaire	Faible/intermédiaire	Faible/intermédiaire
Potentiel d'interactions médicamenteuses	Substantiel	Intermédiaire	Faible	Intermédiaire

1.6.2.1 Inhibiteurs de la protéase ou anti-NS3

La première classe des AAD à avoir obtenu une autorisation de mise sur le marché (AMM) est celle des anti-NS3 ou anti-protéase. Les molécules ont le suffixe « -prévir ». Les premières molécules utilisées en 2011, le télaprévir et le bocéprévir, étaient utilisées en combinaison avec l'IFN et la ribavirine. Elles agissent par compétition avec le substrat enzymatique par liaison covalente à la protéase. Elles ont rapidement été supplantées par des molécules de 2^{ème} génération avec moins d'effets secondaires et une plus forte barrière de résistance. Notamment, les molécules présentant une AMM avec indication pour le traitement du VHC sont le paritaprévir, le grazoprévir, le voxilaprévir et le glécaprévir, dont les deux dernières à activité pan-génotypique (106). Elles vont cibler le domaine NS3 et peuvent être, hormis pour le voxilaprévir et le glécaprévir, actives sur certains génotypes spécifiquement, d'où une couverture génotypique considérée étroite.

La protéine virale NS3 est nécessaire pour la réplication virale et la fixation des anti-NS3 sur le site catalytique bloque la libération des protéines virales non structurales fonctionnelles.

1.6.2.2 Anti-NS5A

Les anti-NS5A ont le suffixe « -asvir » et on retrouve sur le marché le daclatasvir, le lédipasvir, l'ombitasvir, l'elbasvir, le velpatasvir et le pibrentasvir, dont les deux dernières d'action pan-génotypique (106). Ces molécules vont agir possiblement sur un ou plusieurs génotypes.

Ces molécules agissent, d'une part, en limitant la réplication, et, d'autre part, en bloquant les étapes d'assemblage des composants viraux pour former la particule virale finale. La protéine NS5A ne présentant pas d'activité enzymatique, les mutations qui peuvent apparaître dans cette région ont peu d'impact sur la viabilité virale et persistent ainsi longtemps. De ce fait, ces inhibiteurs sont à l'origine d'une diminution rapide de la réplication, effective à partir de quelques jours de traitement.

1.6.2.3 Inhibiteurs de la polymérase ou anti-NS5B

La dernière classe de molécules présente le suffixe « -buvir », avec le sofosbuvir et le dasabuvir (106).

Elles vont agir sur l'ARN polymérase virale ARN dépendante ou NS5B, empêchant alors sa fonction lors de la réplication virale. Le dasabuvir est un inhibiteur non nucléosidique non compétitif qui va induire un changement de conformation de l'enzyme par fixation à un site allostérique de NS5B, responsable de la perte de son activité.

Le sofosbuvir, inhibiteur nucléotidique, nécessite une activation intracellulaire par phosphorylation puis va s'incorporer dans l'ARN à la place du substrat naturel et agir en tant que terminateur de la synthèse d'ARN du VHC en cours. Il agit donc en tant qu'inhibiteur compétitif. Grâce à ce mode d'action, il est capable d'agir sur tous les génotypes et le risque de résistance est faible (107). D'éventuelles mutations peuvent apparaître au niveau du site catalytique de la polymérase mais entraînent une perte trop importante de la capacité répliquative du VHC pour être sélectionnées.

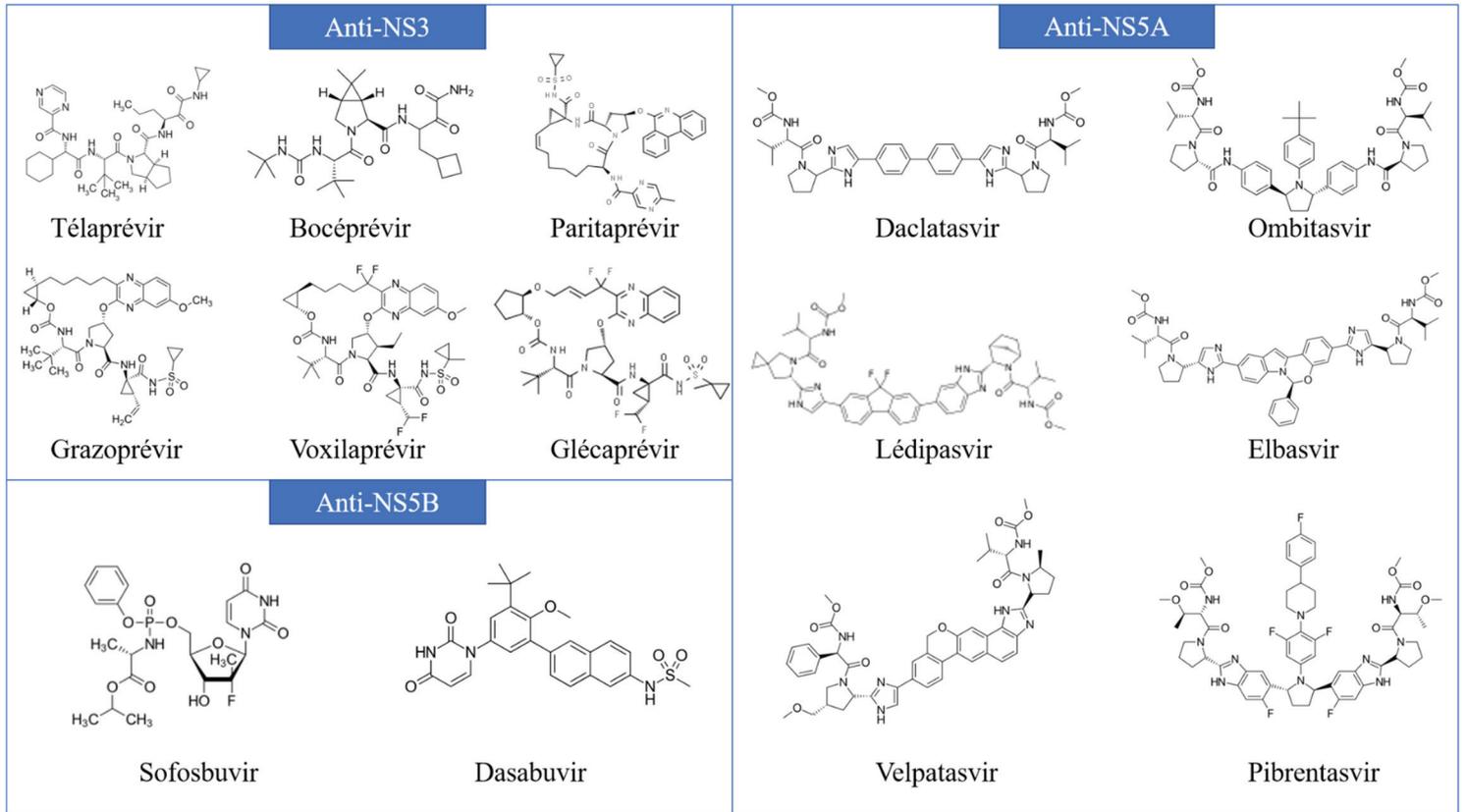


Figure 14. Structures des anti-VHC, selon la protéine cible.

Selon les informations de PubChem (108).

Spécialité	Dénomination commune internationale	Génotype(s) ciblé(s)	Date AMM	Arrêt commercialisation	Retrait AMM
Anti-NS3					
Incivo®	Télaprévir	1	09/2011	04/2015	-
Victrelis®	Bocéprévir	1	07/2011	11/2015	06/2018
Olysiso®	Siméprévir	1 et 4	05/2014	04/2018	05/2018
Anti-NS5A					
Daklinza®	Daclatasvir	1, 3, 4	08/2014	-	-
Anti-NS5B					
Sovaldi®	Sofosbuvir	1 à 6	01/2014	-	-
Exviera®	Dasabuvir	4	01/2015	10/2018	-
Associations thérapeutiques					
Viekierax®	Paritaprévir (anti-NS3) Ombitasvir (anti-NS5A) Ritonavir (inhibiteur du métabolisme)	1 et 4	01/2015	10/2018	-
Zepatier®	Grazoprévir (anti-NS3) Elbasvir (anti-NS5A)	1 et 4	07/2016	-	-
Vosevi®	Voxilaprévir (anti-NS3) Velpatasvir (anti-NS5A) Sofosbuvir (anti-NS5B)	1 à 6	07/2017	-	-
Harvoni®	Lédipasvir (anti-NS5A) Sofosbuvir (anti-NS5B)	1, 3, 4, 5 et 6	11/2014	-	-
Epclusa®	Velpatasvir (anti-NS5A) Sofosbuvir (anti-NS5B)	1 à 6	07/2016	-	-
Maviret®	Glécaprévir (anti-NS3) Pibrentasvir (anti-NS5A)	1 à 6	07/2017	-	-

Tableau 3. Spécialités utilisées dans le traitement de l'infection par le VHC.

D'après les données de l'Agence Nationale de Sécurité du Médicament et des Produits de Santé (106,109).

1.6.3 Recommandations actuelles

L'arrivée de nouvelles combinaisons thérapeutiques à un rythme très rapide ces dernières années, avec des coûts initialement exorbitants, a fait évoluer les recommandations de traitement tous les 6 mois entre 2012 et 2018. En 2018, l'arrivée de combinaisons pan-génotypiques très efficaces et très bien tolérées, ainsi qu'une diminution des prix, ont permis une simplification de la prise en charge, rendant possible depuis mai 2019, la prescription d'AAD par l'ensemble des médecins en France (110). Il n'y a, à l'heure actuelle, plus d'autres molécules en développement.

Le principe du traitement du VHC repose sur l'association de plusieurs molécules provenant de différentes classes, pour une durée de 8 ou 12 semaines selon l'association utilisée et/ou la sévérité de la maladie.

Actuellement, les recommandations de 2018 de l'Association Française pour l'Etude du Foie (AFEF) proposent différents schémas selon la présence d'antécédent de traitement et/ou l'avancée des lésions hépatiques, et notamment le niveau de fibrose (111).

1.6.3.1 Bilan initial du patient

Quand un patient est diagnostiqué comme ayant une infection active par le VHC, c'est-à-dire présentant une charge virale positive dans le sang, il est nécessaire d'évaluer les comorbidités et de rassembler toutes les informations nécessaires à la mise en place du traitement le plus approprié (112) : la consommation d'alcool, un surpoids, une obésité, un diabète ou encore un syndrome métabolique ; une insuffisance rénale sévère ; une co-infection avec le VIH et/ou le VHB ; la notion d'échec à un traitement anti-VHC antérieur. Dans ces cas-là, la prise en charge du patient sera dite spécialisée.

Pour estimer l'atteinte hépatique, différents tests non invasifs peuvent être réalisés : le Fibroscan® qui mesure l'élasticité hépatique, le Fibrotest® ou le Fibromètre® qui sont des tests sanguins évaluant les niveaux de marqueurs révélateurs de fibrose. Si les valeurs obtenues à ces tests sont respectivement plus élevées que 10kPa, 0,58 et 0,786, le patient nécessite en effet un parcours spécialisé.

Un bilan sanguin initial est également réalisé comprenant les bilans hépatique et rénal ainsi que la numération de la formule sanguine (NFS). La mesure de la charge virale avant initiation du traitement est également indispensable afin de fournir une référence au cours du traitement.

1.6.3.2 Recommandations pour une prise en charge simplifiée

Lorsque la prise en charge est dite simplifiée, dans le cas où il n'y a pas d'antécédent de traitement pour le VHC et pour une fibrose modérée ou minime, le traitement repose sur l'utilisation d'une association pan-génotypique récente Epclusa® (sofosbuvir + velpatasvir) pendant 12 semaines ou Maviret® (glécaprevir + pibrentasvir) pendant 8 semaines (voir Figure 15).

Le contrôle de la charge virale est effectué 12 semaines après la fin du traitement : si la charge virale est indétectable, on considère que le traitement a permis une réponse virologique soutenue, et que le patient est guéri.

Cependant, il est important de noter la persistance des anticorps anti-VHC même après guérison, signe du contact du système immunitaire avec le virus. Malgré tout, ces anticorps ne protègent pas d'une infection ultérieure, d'où le risque de réinfection si le patient présente un comportement à risque (voir Rappels bibliographiques, **1.5.1 Les modes de transmission**). Un suivi adapté peut donc être proposé chez ces patients, avec une recherche régulière de la charge virale du VHC.

De même, les patients chez qui des comorbidités hépatiques ont été détectées doivent se voir suivre régulièrement. En effet, la guérison du VHC ne permet pas systématiquement de s'affranchir de l'altération hépatique viro-induite, ni du risque de CHC si les lésions étaient déjà trop avancées (111).

1.6.3.3 Recommandations pour une prise en charge spécialisée

Si le patient présente un antécédent de traitement pour le VHC et/ou des lésions hépatiques importantes, la prise en charge est dite spécialisée, avec utilisation possible de molécules génotype-dépendant.

Ici encore, Epclusa® peut être utilisé pour 12 semaines, ainsi que Maviret® pendant 8 à 16 semaines. Cependant, en raison notamment de leur efficacité et de leur facilité d'emploi, deux autres associations peuvent être utilisées selon le génotype. Harvoni® (sofosbuvir + lédirasvir) peut être donné pour 8 semaines pour des génotypes 1 sans maladie hépatique sévère. La deuxième association Zepatier® (grazoprévir + elbasvir) pendant 12 semaines peut être indiquée chez les patients infectés par le génotype 1b, le génotype 1a avec une charge virale initiale inférieure à 800000 UI/mL mais également chez les patients infectés par un génotype 4 et sans antécédent de bithérapie par IFN et

ribavirine. Ces deux associations de molécules permettent ainsi d'obtenir une réponse virologique soutenue chez plus de 95% des patients traités (111).

Le même suivi est indiqué pour les patients présentant une réponse virologique soutenue après le traitement (voir Rappels bibliographiques, 1.6.3.2 Recommandations pour une prise en charge simplifiée).

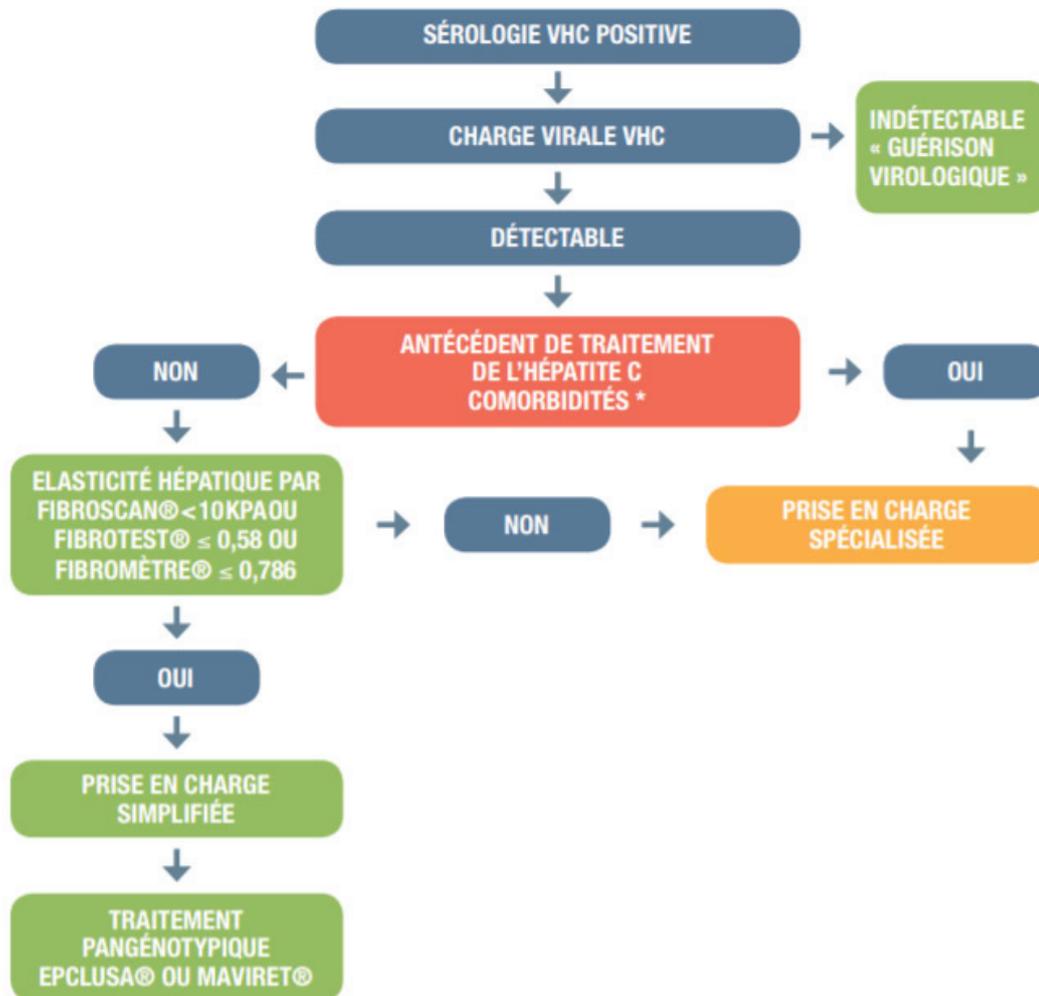


Figure 15. Algorithme de prise en charge des patients infectés par le VHC.

Dès lors qu'un patient est diagnostiqué pour le VHC, le bilan initial permettra d'initier un traitement, selon une prise en charge simplifiée ou spécialisée.

D'après l'Association Française pour l'Etude du Foie, 2018 (111).

1.6.4 Mutations de résistance

Une mauvaise observance, l'arrêt des antiviraux avant la durée recommandée, une interaction du traitement avec d'autres médicaments, un foie fibrosé peuvent entraîner une sous-exposition des cellules infectées aux concentrations efficaces et donc être à l'origine d'une perte d'efficacité du traitement. D'autre part, certaines mutations préexistantes dans des régions d'intérêt du virus, et notamment NS3, NS5A, NS5B, peuvent également avoir un impact sur la sensibilité aux molécules au cours du traitement. Ces protéines virales étant les cibles du traitement, des mutations sont susceptibles d'être sélectionnées afin de permettre au VHC de continuer à se disséminer tout en échappant à l'effet des molécules antivirales. Pour cette raison, le séquençage est utile pour la détection de ces mutations dans le génome viral pour permettre d'adapter le traitement, selon les recommandations (113).

La variabilité du VHC est telle que l'impact d'une mutation sur la sensibilité vis-à-vis d'une molécule dépend du génotype voire du sous-type. Le tableau 4 recense les mutations associées à une résistance, selon la classe antivirale ou la molécule utilisée et en fonction du génotype infectant pour des patients en échec d'AAD.

L'échec aux AAD peut être expliqué par l'apparition de ces mutations dans les régions virales cibles du traitement. Tandis que les mutations présentes dans NS3 disparaissent globalement quelques mois après leur apparition, celles trouvées dans NS5A peuvent persister pendant plusieurs années. Les mutations pouvant être retrouvées au niveau de la polymérase virale NS5B, et plus particulièrement sous sofosbuvir, sont généralement rarement sélectionnées du fait de leur impact sur le « fitness » viral (voir Figure 16), contrairement à celles développées sous inhibiteur de NS5B non nucléosidique.

Variants associés à
une résistance (%)

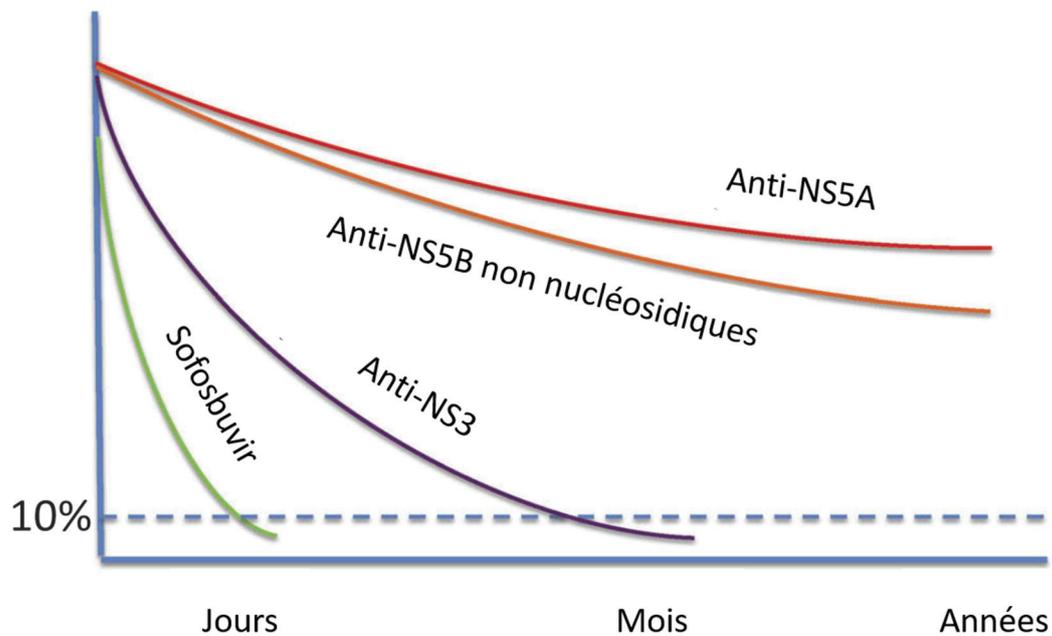


Figure 16. Cinétique de persistance des variants associés à une résistance au cours des différents traitements.

La persistance des variants associés à une résistance est représentée en fonction du temps, pour différents types de traitements.

Adapté de Benítez-Gutiérrez et al., 2016 (114).

Récemment, des mutations préexistantes dans NS5A et NS5B du VHC ont été recensées avant l'initiation de tout traitement pour 16% et 2,5% des patients infectés, respectivement (115). Le niveau de prévalence de ces substitutions virales est variable selon la région géographique mais dépend également du génotype infectant (116,117).

Notamment, les mutations présentes dans NS5A ont une importance capitale dans l'issue du traitement antiviral : une étude de 2017 a montré que pour 1765 patients infectés par un génotype 1, la bithérapie sofosbuvir/lédipasvir présente une réduction d'efficacité importante pour les patients pré-traités par antiviral et présentant, avant l'initiation de cette bithérapie, des mutations de résistance dans NS5A (76% contre 97% chez les patients sans mutation [génotype 1a], 89% contre 98% [génotype 1b]). Ces substitutions ne montrent qu'un impact réduit pour les patients de génotype 1a naïfs de traitement (91% contre 99% chez les patients sans mutation), et aucun impact dans le cas du génotype 1b (99% contre 99%) (118).

	Position acide aminé	Génotype (sous-type)						
		1a	1b	2	3	4	5	6
Sofosbuvir	159	L159F	L159F	L159F	L159F			
	282	S282T/R	S282T	S282T	S282T	S282T	S282T	
	320	L320I/F/V						
	321	V321A			V321A			
Inhibiteurs NSSA	24	K24G/N/R			T24A			Q24H
	26	K26E						
	28	M28A/G/T/S/V	L28M/T	L/F28M/V/S	M28T	L28S/V	L28I	F28L
	29		P29S					
	30	Q30C/D/E/G/H/I/K/L/Q/R/S/T/Y	R30G/H/P/Q/R	L30H/S	A30K/S	L30H		
	31	L31I/F/M/P/V	L31F/I/M/V	L31M/V	L31I/M/V	L31I/M	L31V	L31M/V
	32	P32L/S	P32F/L/S					P32L/S
	38	S38F						
	58	H58D/L/R	P58D/S			T58P/S		T58A/N/S
	62		Q/E62D					
	92	A92K/T	A 92K					
	93	Y93C/F/H/L/N/R/S/T/W	Y93C/H/N/S	Y93H	Y93H	Y93HY93H/R		
	Inhibiteurs de protéase	36	V36A/C/G/L/M	V36A/C/G/L/M				
41		Q41R	Q41R					
43		F43L	F43I/S/V					
54		T54A/S	T54A/C/G/S					
55		V55A/I	V55A					
56		Y56H	Y56H/L	Y56H		Y56H		Y56H
80		Q80H/K/L/R	Q80H/K/L/R					L80K/Q
122		S122G/R	S122D/G/I/N/R/T					S122T
155		R155G/I/K/M/S/T/W	R155C/G/I/K/Q/M/S/T/W					
156		A156S/T/V	A156G/F/S/T/V					
158		V158I	V158I					
168		D168A/C/E/F/G/H/I/K/L/N/T/V/Y	D168A/C/E/F/G/H/I/K/L/N/T/V/Y		Q168R	D168V		D168E/Y
170		I/V170F/T/V	I/V170A/L/T					I170V
175		M175L						
Dasabuvir	314	L314H						
	316	C316Y	C316H/N/Y/W					
	368		S368T					
	411		N411S					
	414	M414I/T/V	M414I/T/V					
	445		C445F/Y					
	446	E446K/Q						
	448	Y448C/H	Y448C/H					
	451	C451R						
	553	A553T	A553V					
	554	G554S	G554S					
	555	Y555H						
	556	S556G/R	S556G/R					
	557	G557R						
	558	G558R	G558R					
559	D559G/N	D559G/N						
561	Y561H/N							

Tableau 4. Substitutions de résistance répertoriées in vitro et/ou chez les patients en échec de traitement, considérant l’AAD et le génotype.

Selon Pawlotsky et al., 2016 (119).

2. Génométypage et séquençage du Virus de l'Hépatite C

La détermination du génotype du VHC infectant chez le patient présente plusieurs intérêts. Comme vu précédemment, le traitement mis en place peut dépendre du génotype, d'où la nécessité de faire le génotypage. D'autre part, cela présente un intérêt épidémiologique, afin de connaître la répartition, la fréquence et la distribution géographique des génotypes. Il permet aussi de distinguer un échec de traitement d'une réinfection, qui ne seront pas traités de la même façon. La mise en évidence de réinfection nécessite par ailleurs un travail pédagogique pour limiter les comportements à risque.

Plusieurs méthodes peuvent être utilisées pour la détermination du génotype, notamment des techniques d'hybridation à des sondes spécifiques ou encore du séquençage direct du génome. En confrontant la séquence obtenue à des séquences répertoriées dans des bases de données (par exemple, à l'aide des sites <https://hcv.geno2pheno.org/>, ou <https://www.ncbi.nlm.nih.gov/refseq/>), le génotype peut être déterminé en fonction des homologies de séquences retrouvées. En plus de permettre le génotypage, le séquençage permet d'obtenir la séquence de régions d'intérêt voire du génome complet afin de pouvoir observer d'éventuelles mutations, importantes à détecter pour la mise en place du traitement et/ou en cas d'échec du traitement.

2.1 Le séquençage de Sanger

2.1.1 Principe général

Le premier séquençage par la technique de Sanger fut réalisé en 1977 par Frederick Sanger, qui obtient ainsi la séquence d'un virus bactérien (12). Actuellement, cette technique est largement utilisée et repose sur l'amplification du génome par PCR.

2.1.1.1 Les acteurs du séquençage

Au cours de la réaction de séquençage, différents composants sont requis :

- Le génome sous forme d'ADN, ayant pour origine l'organisme, ici le VHC, dont on souhaite connaître la séquence. Le VHC étant un virus à ARN, une étape préalable de rétro-transcription est nécessaire pour obtenir l'ADN correspondant. Si l'ADN se trouve sous forme double-brin, il est nécessaire qu'il soit dénaturé afin d'obtenir des brins séparés. Cette étape de dénaturation est classiquement effectuée par l'obtention d'une température élevée au cours du séquençage. Un

- des brins, le brin appelé « matrice », servira ainsi de modèle au cours du séquençage ;
- Les désoxyribonucléotides (dNTP) ou bases (A, T, G et C), qui vont être incorporés dans la chaîne néoformée par appariement aux bases présentes sur le brin matrice. Les bases seront liées entre elles au cours de l'élongation à l'aide d'un groupement OH particulier en 3' du ribose ;
 - Les didésoxyribonucléotides (ddNTP), nécessaires en quantité moins importante que les dNTP, qui vont agir en tant que terminateurs de chaîne à cause de l'absence des groupements 2'OH et 3'OH du ribose (voir Figure 17), empêchant l'élongation de la chaîne de manière irréversible ;
 - Les amorces, qui sont des séquences d'ADN courtes comptant une vingtaine de nucléotides, qui vont permettre de s'hybrider au brin matrice par complémentarité afin de commencer l'élongation du second brin ;
 - L'ADN polymérase, qui va permettre l'incorporation des dNTP et ddNTP par son activité enzymatique par complémentarité entre les bases du brin matrice et du brin néosynthétisé. Elle débute son activité grâce à l'amorce hybridée sur le brin matrice. La synthèse du brin se fait de manière à ce que les bases soient complémentaires, soit un A faisant face à un T, et un C faisant face à un G.

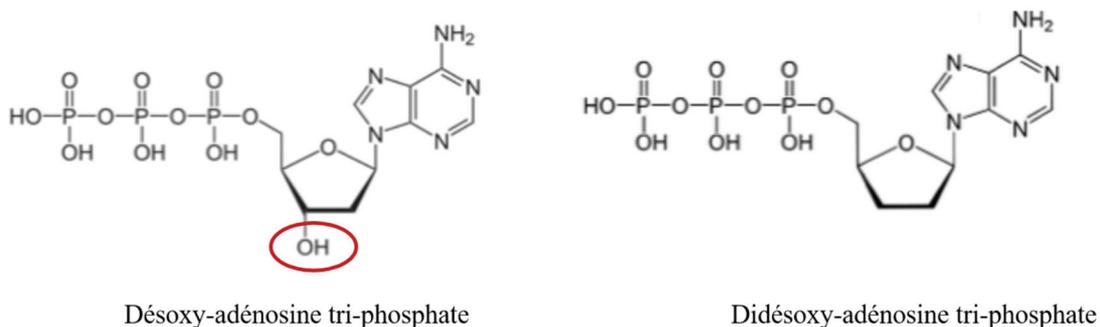


Figure 17. Représentations structurales du dATP et du ddATP.

Sont représentés chimiquement le dATP (désoxy-adénosine tri-phosphate à gauche) avec le groupement OH, et le ddATP (didésoxy-adénosine tri-phosphate à droite) privé du groupement OH et à l'origine de l'interruption de l'élongation d'ADN.

D'après Young, 2017 (120).

2.1.1.2 Les étapes du séquençage

Le séquençage commence en présence du mélange réactionnel précédemment détaillé (voir Rappels bibliographiques, 2.1.1.1 Les acteurs du séquençage), dont chaque composant est présent en quantité spécifique pour permettre la réaction de séquençage de manière efficiente. Tout d'abord, les brins vont être séparés à haute température afin de permettre l'hybridation des amorces sur les brins. L'ADN polymérase pourra alors incorporer, selon la complémentarité avec le brin matrice et de manière aléatoire, les dNTP, auquel cas la synthèse du brin néoformé continue, ou les ddNTP, auquel cas la réaction s'arrête. De nombreuses réactions ont lieu simultanément et entraînent la formation de chaînes d'ADN de tailles différentes, correspondant respectivement à une chaîne interrompue par l'incorporation d'un ddNTP à chaque position. Ces différents fragments vont pouvoir être séparés selon leur taille, par migration sur un gel d'acrylamide à l'aide d'un courant électrique. La migration du fragment se fera d'autant plus loin sur le gel qu'il sera court et ainsi léger. Inversement, les fragments plus longs migrent de manière plus lente et iront moins loin. Après migration, une succession de bandes est obtenue et la lecture permettra de connaître quel nucléotide est présent à chacune des positions, afin de déterminer la séquence de l'ADN d'intérêt (12,121).

Cette lecture est rendue possible grâce à un marquage des ddNTP. Historiquement, le marquage radioactif étant utilisé, il n'était pas possible de discriminer les ddNTP entre eux. Pour cette raison, 4 réactions distinctes étaient effectuées dans 4 tubes différents, avec un type unique connu de ddNTP pour chacun d'entre eux (ddATP, ddCTP, ddGTP ou ddTTP). La migration des fragments de chacun des tubes en parallèle permettait de déterminer la succession des bases de l'ADN d'intérêt (voir Figure 18) (121).

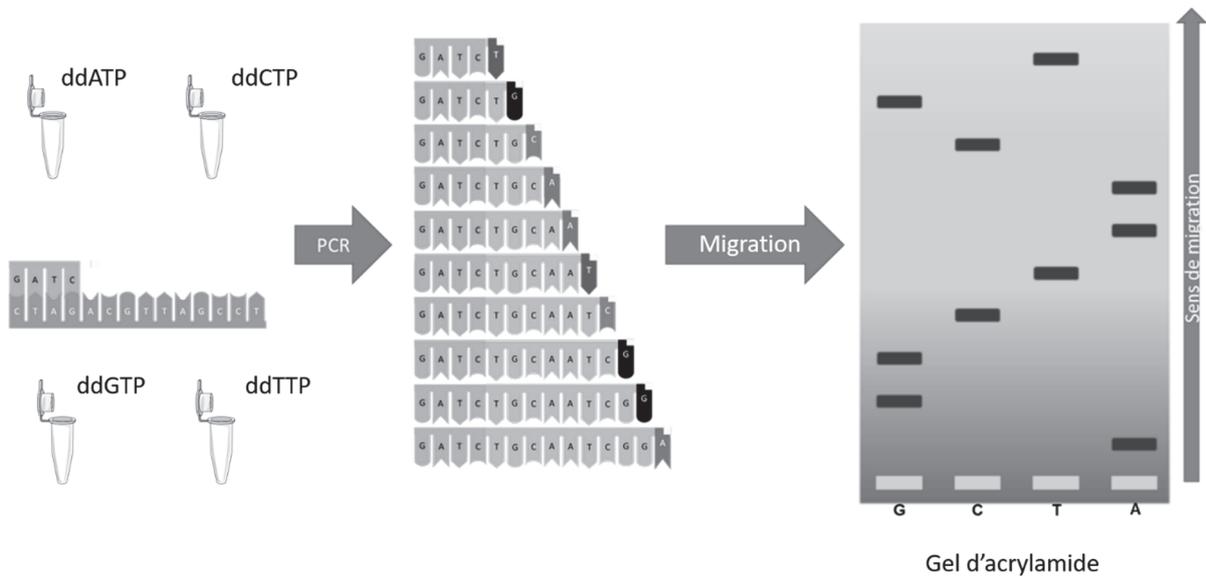


Figure 18. Déroulement du séquençage par Sanger, avec marquage radioactif.

Les 4 réactions avec chacun des ddNTP se font dans 4 tubes différents. Après amplification par PCR, la migration permet de séparer, pour chaque tube, les fragments de différentes tailles sur gel d'acrylamide. La lecture de la radioactivité permet de déterminer la séquence.

Adapté de Biomnigène (122).

Une évolution importante de cette technique fut le remplacement de ce marquage radioactif par un marquage avec 4 fluorochromes différents, avec une couleur propre à chacun des 4 ddNTP. Ces fluorophores sont des composés chimiques qui vont dégager une lumière de couleur spécifique. Cette avancée permet ainsi d'exécuter la réaction complète de séquençage dans un seul tube. Après électrophorèse capillaire de tous les fragments, la lecture optique grâce à un détecteur de fluorescence après excitation par un laser aboutira à l'établissement de la séquence voulue. La séquence ainsi reconstituée peut être représentée sur un chromatogramme après traitement informatique des données (voir Figure 19) (121).

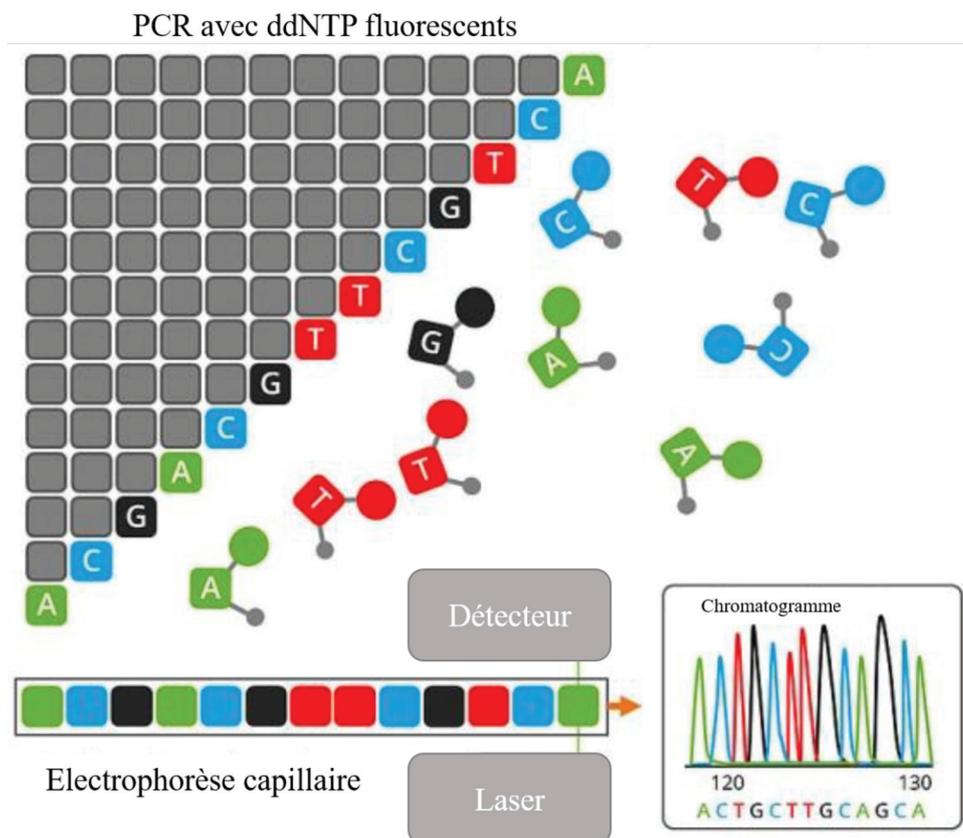


Figure 19. Déroulement du séquençage par Sanger, avec marquage par fluorochromes.

L'avancée majeure du séquençage de Sanger réside en l'utilisation d'un marquage par fluorochromes pour chacun des ddNTP afin de permettre toutes les réactions dans un même tube. L'électrophorèse capillaire puis la détection après excitation par laser permettront l'obtention du chromatogramme avec la succession des bases de l'ADN. Adapté d'ABM (123).

2.1.2 Apports de la technique vis-à-vis du VHC

Le séquençage Sanger est utilisé d'une part, pour déterminer le génotype et le sous-type du VHC (124), mais également pour la détection de polymorphismes (125) ou de mutations dans les régions d'intérêt, possiblement associées à une résistance aux traitements (126,127). En effet, la séquence virale obtenue après traitement informatique peut être alignée pour comparaison à des séquences tirées de bases de données, répertoriées comme appartenant à un génotype et sous-type déterminés. Les homologies entre les séquences permettront d'assimiler la séquence d'intérêt à un génotype et sous-type connu ; tandis que des mutations pourront être révélées, pouvant avoir un impact sur l'efficacité d'un éventuel traitement (mutation de résistance) ou non (polymorphisme).

Dans le cas où différentes séquences du VHC coexistent au sein de l'échantillon d'intérêt, la séquence finale correspond à la moyenne des fluorescences détectées à chaque position, soit la séquence dite « consensus », la plus fréquente. Sur le chromatogramme, il est cependant possible d'observer un double pic pour la position de nucléotide concernée si la proportion de la deuxième séquence est assez importante. Il est donc possible de détecter la présence de plusieurs populations virales, mais uniquement si ces variants représentent plus de 15 à 20% de la population virale totale, permettant la détection de la coexistence de quasi-espèces. Il n'est cependant pas possible de déterminer leur fréquence au sein du mélange. De plus, si plusieurs génotypes ou sous-types coexistent, le chromatogramme devient illisible (voir Figure 20).

L'inconvénient majeur de cette technique pour la détermination des mutations de résistance réside en la nécessité de relecture et de correction des chromatogrammes, qui restent subjectives. Cela est chronophage et nécessite une expertise avancée.

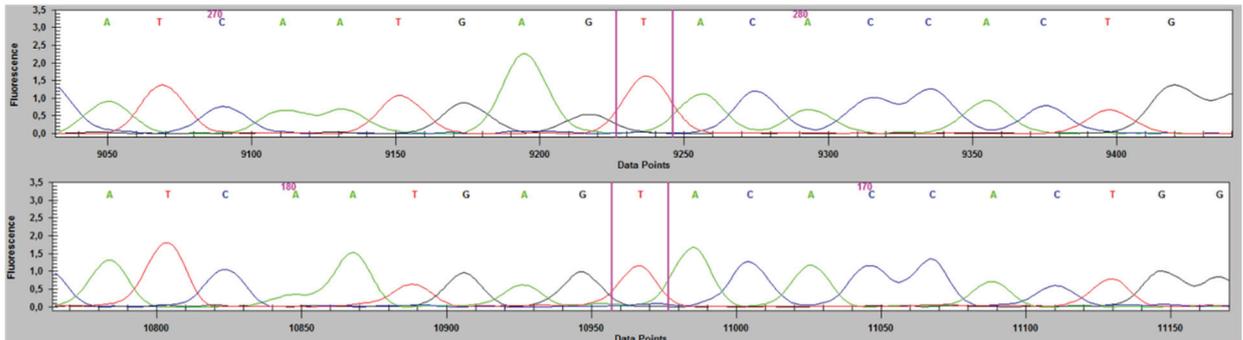
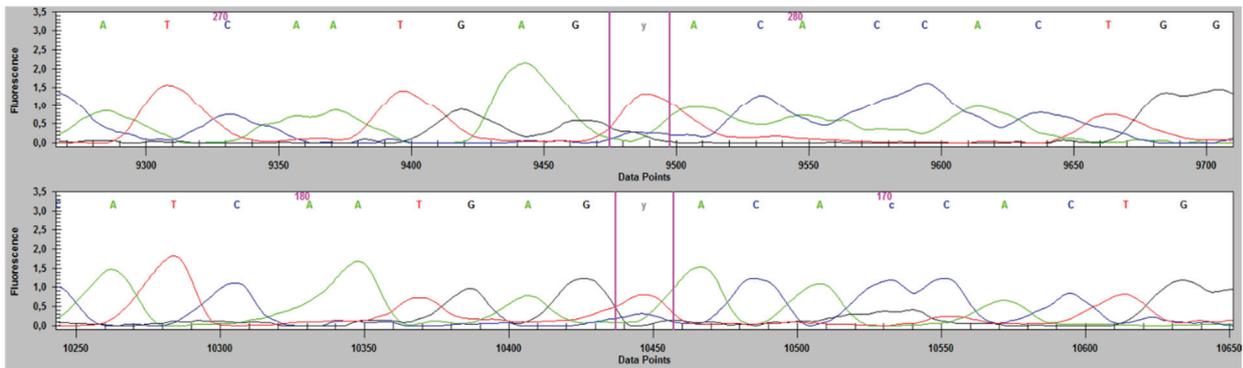
A**B**

Figure 20. Exemples de chromatogrammes.

Les chromatogrammes A montrent des séquences lisibles, avec la détermination des nucléotides à chaque position rendue possible par une qualité de séquençage suffisante, à la fois sur la séquence sens (séquence du haut, séquençage dans le sens 5'-3') et sur la séquence antisens (séquence du bas, séquençage dans le sens 3'-5'). Les chromatogrammes B montrent des séquences de plus mauvaise qualité, montrant notamment à la position d'intérêt différents pics, d'où la difficulté de déterminer le nucléotide exact de manière objective. L'ambiguïté du mélange probable de T et de C à cette position est notée par la lettre Y.

Images obtenues sur le logiciel GenomeLab™ GeXP de Beckman Coulter.

Au laboratoire de virologie de l'Hôpital de la Croix-Rousse, le génotypage du VHC était réalisé systématiquement chez les patients diagnostiqués pour une infection du VHC avant la mise en place d'un traitement antiviral. La détermination du génotype se faisait par amplification de l'ARN du VHC par RT-PCR puis séquençage par cette technique de Sanger. Le génome du VHC faisant autour de 10 kb (kilobases) de longueur, le séquençage par Sanger rend nécessaire le choix de régions à séquencer car cette technique est limitante sur la taille des fragments à analyser, qui ne peut dépasser 800 à 1000 nucléotides. Ainsi, uniquement les régions NS3, NS5A et NS5B, suffisantes pour la détermination du génotype, étaient séquencées.

Le choix des régions séquencées a évolué au cours du temps, en fonction de l'évolution des traitements. Historiquement, le génotypage était effectué par séquençage de la région 5' non codante ; cependant, cette région est insuffisamment discriminante pour la distinction correcte des sous-types. Désormais, les 3 régions virales NS3, NS5A et NS5B sont considérées comme assez variables pour permettre de distinguer les différents génotypes et sous-types, mais suffisamment conservées pour la conception d'amorces nécessaires à leur amplification (128,129). Le choix de la région est également orienté par l'intérêt d'obtenir simultanément une information supplémentaire sur les mutations de résistance. Ainsi, suite à l'arrivée des premiers anti-NS3, le séquençage était fait sur la région NS3, puis l'importance prise par les anti-NS5A a mené à séquencer la région NS5A en première intention. Les régions NS3 et NS5B étaient alors séquencées en seconde ligne, dans le cas où la séquence de NS5A déterminée par la technique de Sanger ne permettait pas la détermination du génotype, en cas de doute ou d'échec de traitement.

2.2 Le séquençage de nouvelle génération ou Next Generation Sequencing (NGS)

La possibilité de traiter l'ensemble des patients depuis l'arrivée des nouveaux traitements entraîne une augmentation exponentielle de la demande de séquençage, nécessitant l'analyse de quantités importantes de données rapidement. Cela est possible grâce à la mise au point des techniques de séquençage à haut débit, dont celles de NGS, dites de deuxième génération, et celles de troisième génération permettant le séquençage de longs fragments, incluant la technologie Nanopore. Cette technique est très attirante ces dernières années pour des raisons de performance, de pratique et de coût. En effet, le séquenceur MinION proposé par Oxford Nanopore Technologies présente la taille d'une clef USB pour un coût inférieur à 1000€. Il présente l'avantage de ne pas nécessiter d'amplification du génome, lorsque celui-ci est sous forme d'ADN, pour le séquençage de très longs fragments à travers des pores de la taille du nanomètre (130).

Parmi les méthodes de séquençage de deuxième génération ou NGS, il existe plusieurs technologies, et notamment la technologie 454, la technologie Illumina (anciennement Solexa), la technologie SOLiD ainsi que la technologie Ion Torrent. Ces techniques présentent généralement trois étapes distinctes et successives comptant la préparation de la librairie par fragmentation de l'ADN à séquençer ; l'amplification de cette librairie par PCR puis le séquençage en lui-même (voir Figure 21), avec différentes approches et caractéristiques selon la technologie employée.

2.2.1 Principe général

2.2.1.1 Déroulement général

La première étape est la construction de la librairie qui consiste à réaliser, après amplification du génome d'intérêt, la fragmentation en petites séquences de taille adaptée à la technologie utilisée. Des adaptateurs d'ADN sont ensuite fixés grâce à une ligase aux extrémités de ces fragments, permettant ensuite leur amplification grâce à des amorces complémentaires à ces séquences. Ces adaptateurs sont de courtes séquences dépendant de la technique de séquençage. D'autre part, des index identiques sont ajoutés pour tous les fragments appartenant à un même échantillon, de manière à pouvoir reconnaître l'appartenance de chacun des fragments car tous les fragments de tous les échantillons

sont ultérieurement mélangés pour le séquençage : on parle de multiplexage. On obtient ainsi une librairie unique qui contient tous les fragments de tous les patients.

L'objectif de la seconde étape est d'amplifier chacun des fragments obtenus lors de la réalisation de la librairie, afin d'avoir un signal suffisant lors du séquençage. Selon la technique employée, les fragments vont être capturés sur différentes surfaces (microbilles ou plaques) à l'aide de séquences complémentaires aux adaptateurs fixés lors de l'étape précédente.

Au cours de la dernière étape s'effectue le séquençage en lui-même, avec des méthodes propres à chaque technologie, dont les processus sont détaillés ci-dessous. Le séquençage Ion Torrent utilise la détection par semi-conducteur, Illumina emploie une technique de séquençage par synthèse avec des bases marquées par des fluorochromes, tandis que le séquençage 454 de Roche exploite le pyroséquençage, et la technologie SOLiD le séquençage par ligation (131).

2.2.1.2 Traitement bio-informatique des données

Après séquençage, une partie bio-informatique est indispensable afin d'obtenir la séquence. L'algorithme de « base calling » convertit les données informatiques sous forme de successions de lettres correspondant aux bases, obtenues dans un fichier Fastq avec un score de qualité (phred quality score) pour chaque séquence, appelée « read » correspondant à un fragment séquencé. Ce score permet dans un premier temps de trier les reads afin de garder uniquement ceux présentant une qualité suffisante pour ne pas identifier de fausses mutations. De même, les reads de petite taille ou ceux polyclonaux, non représentatifs de la réalité, sont éliminés. Il est également nécessaire d'éliminer les séquences correspondant aux adaptateurs préalablement fixés aux fragments d'ADN, car elles ne représentent pas des séquences d'intérêt.

A partir de la multitude de reads obtenus et provenant des différents patients, il est ensuite nécessaire de reconstruire les séquences dans leur totalité : il peut s'agir d'un génome complet, mais également de régions particulières, selon les amorces choisies précédemment au cours de l'amplification. Pour le VHC, il est par exemple possible d'amplifier les régions NS3, NS5A et NS5B à l'aide d'amorces spécifiques pour ces régions. L'amplification du génome complet avec des amorces appropriées est aussi possible mais compliquée et aléatoire selon les patients et la charge virale (132,133).

Pour recomposer les séquences, deux méthodes existent : l'assemblage *de novo* ou l'alignement sur référence :

- L'assemblage *de novo* consiste à recréer la séquence totale sans modèle, à l'aide des fragments chevauchants appartenant à chacun des échantillons. Cela aboutit à la reconstruction de séquences que l'on appelle « contigs ». L'ensemble des contigs représente alors le « scaffold », qui est la séquence totale mais qui peut ainsi présenter des bouts de séquences manquants appelés « gaps » lorsque les reads ne couvrent pas certaines zones. Cette technique est utilisée notamment pour les génomes non connus.
- La deuxième méthode est celle d'alignement sur référence, c'est-à-dire que l'on s'aide cette fois d'une séquence dite modèle, ici du VHC d'un certain génotype, pour recréer la séquence totale à l'aide des reads.

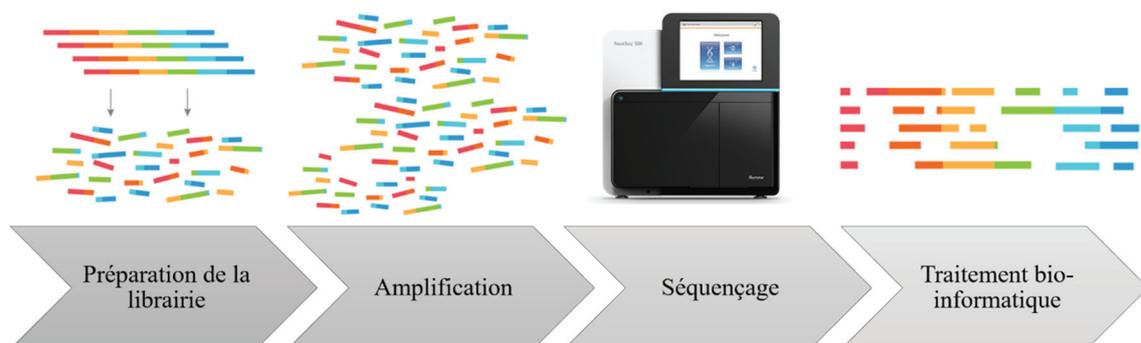


Figure 21. Etapes du séquençage par NGS.

Les technologies de NGS nécessitent la réalisation d'une librairie, qui sera ensuite amplifiée avant le séquençage. L'étape finale de traitement bio-informatique est ensuite indispensable afin d'obtenir les séquences.

2.2.2 Technologie 454

Première technologie appartenant au NGS, le pyroséquençage 454 a été décrit en 2005 (134). L'ADN est fragmenté au cours de la production de la librairie et des adaptateurs ainsi que des index, permettant de connaître l'appartenance des fragments aux différents patients, sont attachés aux fragments. Comme lors du séquençage par Ion Torrent et par SOLiD (voir Rappels bibliographiques, **2.2.4 Technologie SOLiD** et **2.2.5 Technologie Ion Torrent**), l'amplification est réalisée sur billes par une PCR en émulsion grâce aux adaptateurs. Les billes sont ensuite déposées dans les puits de la surface où se déroule le séquençage.

Cette technologie repose ensuite sur la technique de pyroséquençage, basée sur un séquençage par synthèse. Le brin complémentaire est ainsi synthétisé à l'aide d'une polymérase. Cette méthode permet la détection, à chaque incorporation du nucléotide complémentaire dans la chaîne en cours de synthèse, de la libération de pyrophosphate. En présence d'adénosine et d'ATP sulfurylase, cette molécule de pyrophosphate est convertie en ATP utilisée ensuite pour la conversion de luciférine en oxyluciférine, réaction catalysée par la luciférase. A l'issue de la cascade enzymatique, la lumière émise est détectée à l'aide d'une caméra et l'intensité relative est proportionnelle au nombre de bases identiques incorporées à la position d'intérêt (131,135) (voir Figure 22).

Les appareils utilisant cette technologie étaient produits par Roche et on distinguait le GS Junior, le GS Junior Plus, le GS FLX Titanium XL+ ainsi que le GS FLX Titanium XLR70 ; mais la commercialisation de ces appareils a été abandonnée courant 2016 (136).

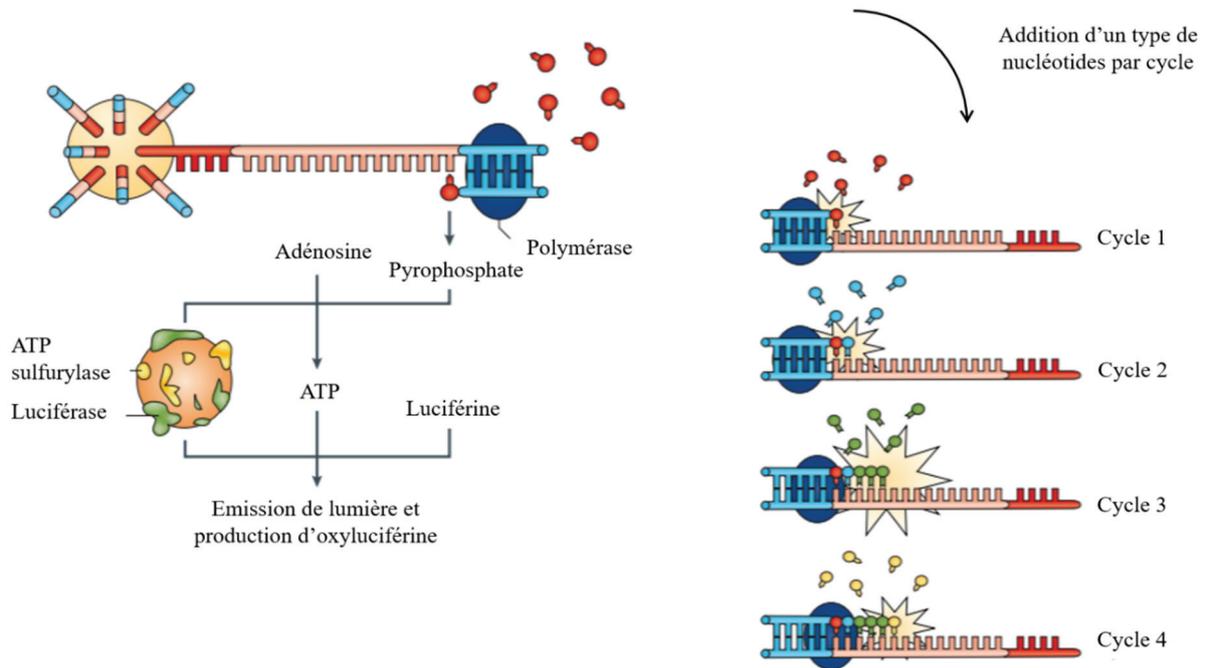


Figure 22. Principe du pyroséquençage de la technologie 454.

L'élongation de la chaîne par la polymérase relâche une molécule de pyrophosphate qui permettra la production d'ATP à partir d'adénosine, permettant lui-même de catalyser la conversion de luciférine en oxyluciférine. La lumière est ainsi émise lorsque le nucléotide effectivement complémentaire est ajouté, auquel cas l'intensité est proportionnelle au nombre de ce même nucléotide incorporé à la suite.

Adapté de Goodwin et al., 2016 (137).

2.2.3 Technologie Illumina

La technologie Illumina représente une seconde méthode parmi les techniques de NGS. Elle diffère notamment de la précédente par la méthode d'amplification appelée « bridge PCR ».

La première étape de librairie suit le même principe de fragmentation des ADN ainsi que l'ajout des index et des adaptateurs. Les fragments, via les adaptateurs, vont ainsi pouvoir s'attacher par complémentarité, aux amorces attachées à une phase solide. Les brins complémentaires aux fragments sont alors synthétisés ; chacun des brins fixés crée ainsi des ponts par hybridation avec l'amorce liée sur la phase solide. Les ADN double-brin nouvellement produits sont ensuite dénaturés. L'objectif est d'obtenir environ 1 million de copies pour chaque fragment d'ADN, formant ainsi des amas ou « clusters » pour chacun d'entre eux (voir Figure 23). Ici encore, il est possible de ne séquencer que des régions particulières du génome cible, ou de choisir la séquence complète.



Figure 23. Amplification par "bridge PCR".

Les fragments de la librairie s'attachent, via leurs adaptateurs et par complémentarité, aux amorces attachées sur la phase solide (1) et la synthèse des brins complémentaires prend place (2), entraînant la formation de ponts (3). La dénaturation des double-brins (4) aboutit à la formation de clusters comportant chacun un nombre important du même fragment (5).

D'après ABM (123).

La technique de séquençage se rapproche de celle utilisée dans la technique de Sanger. En effet, les 4 types de nucléotides utilisés sont marqués par des fluorochromes distincts afin de synthétiser les nouveaux brins. Cependant, ces nucléotides marqués ne vont pas agir en tant que terminateurs irréversibles : l'élongation sera seulement bloquée de manière provisoire. Grâce à l'excitation par un laser, la fluorescence émise par chaque cluster à chaque incorporation d'un nucléotide est détectée, permettant de déterminer quel nucléotide a été intégré pour chaque cluster (131) (voir Figure 24).

L'étape bio-informatique pour l'assemblage des reads est la dernière étape nécessaire pour obtenir les séquences.

Le système fut initialement développé par Solexa, actuellement Illumina, en 2006 (138). Parmi les séquenceurs utilisant cette technique, on compte le MiSeq, le NextSeq500, le HiSeq2500, 3000 et 4500 de capacité de séquençage croissante (123,137,139).

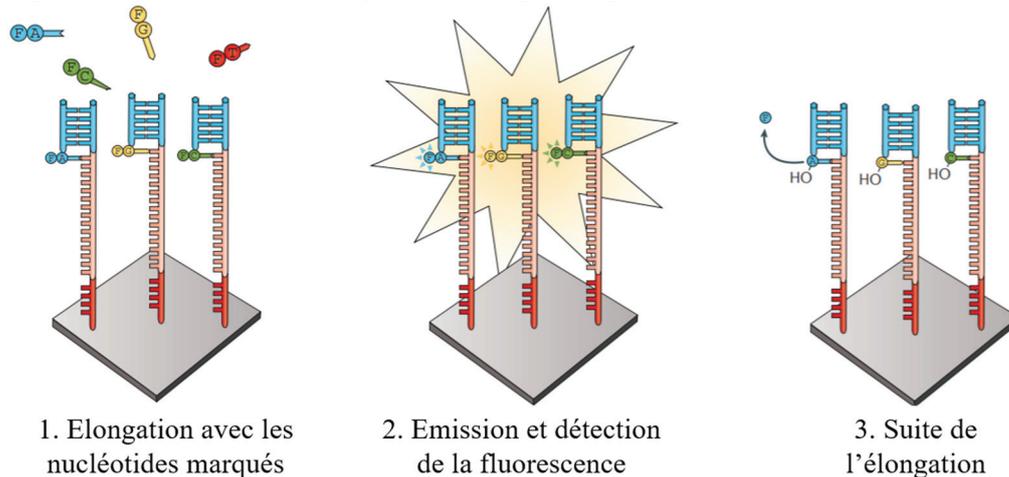


Figure 24. Approche de séquençage par synthèse d'Illumina.

Après la fixation des fragments sur la phase solide, le mélange réactionnel est ajouté, contenant notamment les nucléotides (A, G, C, T) marqués par un fluorophore (F) spécifique à chacun d'entre eux. Etant donnée la structure des nucléotides, un seul peut s'incorporer au cours de l'élongation, bloquant ensuite temporairement l'addition d'un nouveau nucléotide. Une fois le nucléotide complémentaire incorporé, les autres nucléotides sont évacués et la fluorescence est détectée, indiquant la nature du nucléotide incorporé à chaque position. Le fluorophore est ensuite clivé, permettant de recommencer un nouveau cycle.

Adapté de Goodwin et al., 2016 (137).

2.2.4 Technologie SOLiD

Cette technologie débute, comme pour les autres, par la production de la librairie. L'amplification est réalisée suivant le modèle détaillé pour la méthode Ion Torrent (voir Rappels bibliographiques, **2.2.5 Technologie Ion torrent** et Figure 27), par PCR en émulsion à l'aide d'adaptateurs pré-fixés sur les fragments, complémentaires aux amorces retrouvées sur les billes. Cela aboutit à la production de billes où sont fixées idéalement des répétitions du même fragment. Ces billes vont ensuite être attachées à une surface de verre, où se déroulera finalement l'étape de séquençage. Comme pour la technologie Ion Torrent, chaque bille sera ainsi le siège d'une réaction de séquençage.

Cependant, cette technologie SOLiD (Sequencing by Oligonucleotide Ligation and Detection) diverge des autres méthodes quant à l'étape de séquençage. Ici, l'élongation du brin n'utilise pas de ddNTP marqués ni de polymérase, mais un mélange d'oligonucléotides marqués et l'intervention d'une ligase, enzyme présentant la capacité de ligaturer des brins d'ADN entre eux par catalyse d'une liaison covalente.

Une première amorce spécifique à la technologie va permettre l'hybridation aux adaptateurs présents sur les fragments et la synthèse du brin va prendre place grâce à la ligation d'oligonucléotides comprenant 8 bases, préalablement marqués à l'aide de fluorescence. Tout d'abord, ces oligonucléotides comportent deux bases xy qui iront s'hybrider selon leur complémentarité à la séquence cible ; sont retrouvées ensuite trois bases dégénérées désignées n, puis trois bases universelles z (inosines) (voir Figure 25). Chaque oligonucléotide est marqué en 5' à l'aide d'un fluorochrome discriminant ainsi, de manière propre, les deux premières bases. La ligase permet alors la fixation de l'oligonucléotide à l'amorce. Un traitement à la phosphatase permet d'éviter que d'autres brins ne soient ligués de manière aléatoire. Puis la fluorescence est détectée par excitation à chaque cycle, permettant l'acquisition de l'image. Les trois bases universelles comportant le fluorochrome sont ensuite clivées de manière chimique, laissant ainsi la sonde comptant les 5 bases liée à l'amorce. Un autre cycle de ligation peut alors débiter par l'ajout d'un autre oligonucléotide, et permettra l'élongation du brin. Ces cycles de ligation continuent jusqu'à que la longueur entière du fragment soit hybridée avec les oligonucléotides.

De ce fait, il est obtenu une séquence présentant une succession de dinucléotides spécifiques et de bases dégénérées. Lorsque ce brin est dénaturé, le cycle suivant de

séquençage peut commencer à l'aide d'une amorce différente de séquençage, décalée d'un nucléotide vis-à-vis de celle utilisée ultérieurement. La répétition de ces étapes avec des amorces différentes permet d'obtenir des séquences qui se chevauchent, assurant le séquençage du fragment entier de manière fidèle (135,137).

Etant donné que chaque signal fluorescent permet l'identification d'un dinucléotide, le codage des données est réalisé sur ces deux nucléotides. Dès lors que l'on connaît la première base, il est possible de convertir les couleurs en bases, grâce aux différents cycles qui permettront de déduire les nucléotides à chaque position (voir Figure 26).

Life Technologies a ainsi commercialisé 2 appareils qui sont le 5500W System et le 5500xl W System.

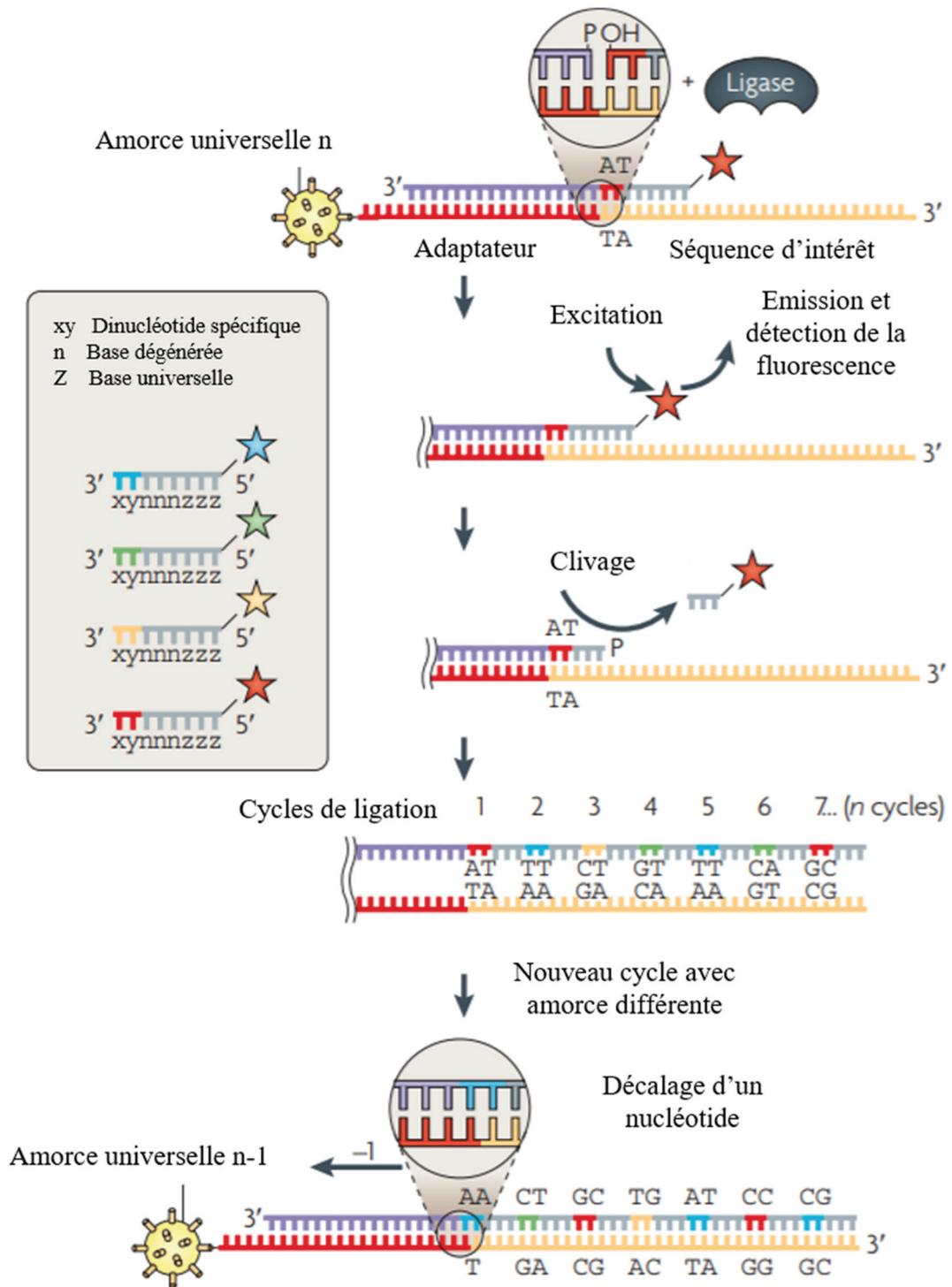
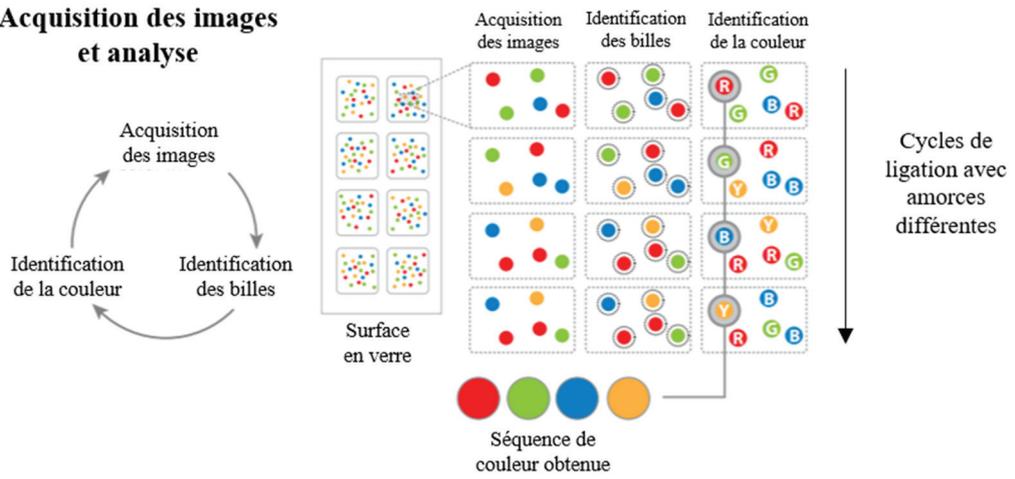


Figure 25. Principe du séquençage par ligation.

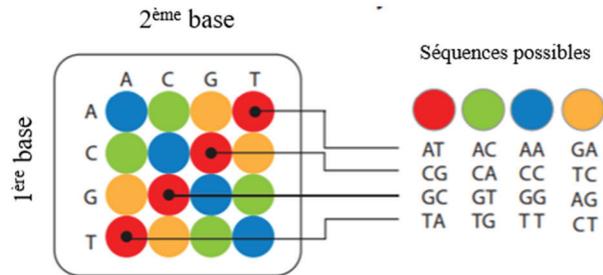
Le séquençage commence à l'aide d'une amorce universelle. L'oligonucléotide est ligé à l'adaptateur et la fluorescence peut être détectée. Les cycles de ligation permettent de couvrir la longueur entière du fragment à séquençer, et un nouveau cycle avec une amorce décalée d'une base permettra de séquencer l'ensemble du fragment de manière fidèle.

Adapté de Metzker, 2010 (140).

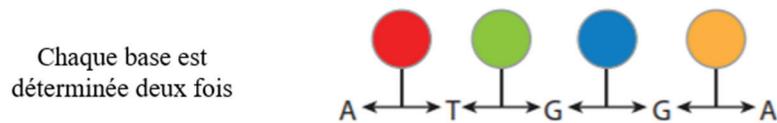
Acquisition des images et analyse



Combinaisons de nucléotides selon la couleur



Double interrogation



Décodage

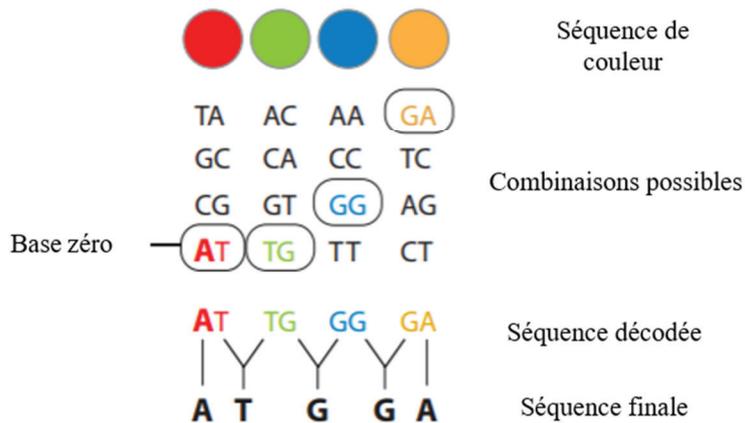


Figure 26. Principe de la détermination de séquence par séquençage SOLiD.

L'acquisition, pour chaque position de billes, des couleurs au fil des cycles permettra, grâce à la double interrogation, de décoder la séquence afin de déterminer la séquence d'intérêt.

Adapté de Mardis, 2008 (141).

2.2.5 Technologie Ion torrent

Après production de la librairie, l'amplification des fragments est obtenue par PCR en émulsion à l'aide d'amorces ciblant les adaptateurs préalablement fixés. Une émulsion huile dans eau est réalisée dans le but de piéger, dans chaque microréacteur formé par les gouttelettes aqueuses, une microbille où sont fixées les amorces pour l'amplification, un fragment d'ADN à séquencer ainsi que les réactifs nécessaires à la PCR (voir Figure 27). La finalité est d'obtenir, après amplification, un maximum de microbilles recouvertes chacune par des milliers de copies d'un seul fragment de l'ADN cible afin d'atteindre le seuil de détection nécessaire pour le séquençage, sachant que seulement 15% des microréacteurs présentent cette configuration idéale.

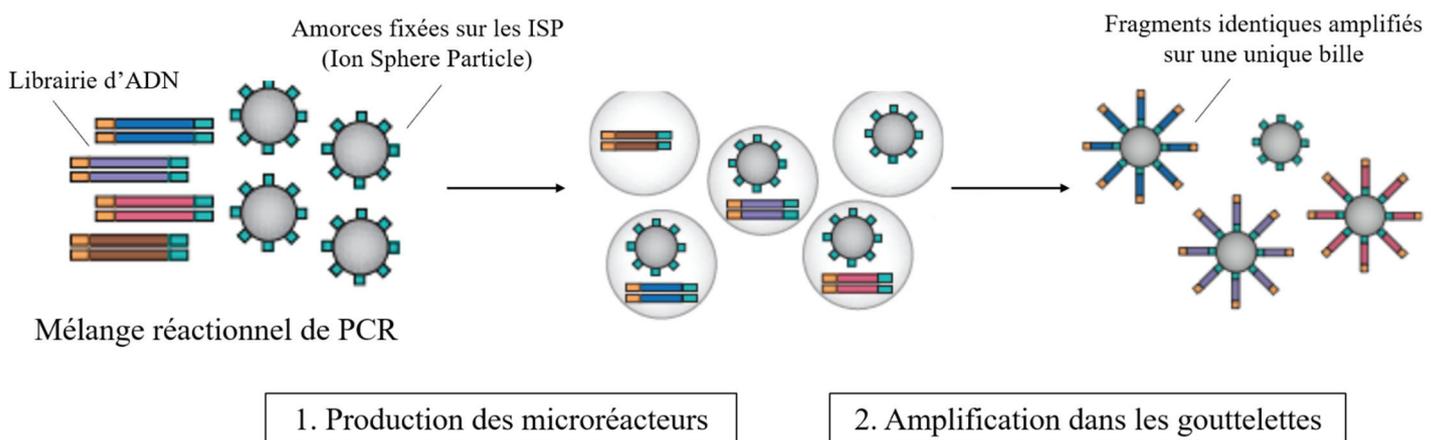


Figure 27. Déroulement de la PCR en émulsion.

Le mélange réactionnel de PCR contenant l'ensemble des fragments générés lors de la production de librairie et les billes (Ion Sphere Particle ISP) auxquelles sont attachées les amorces complémentaires aux adaptateurs des fragments d'ADN pour leur amplification permet la production de microréacteurs où se déroulera l'amplification, idéalement monoclonale, d'un fragment unique sur une seule bille.

La phase suivante consiste à rompre l'émulsion et à réaliser l'enrichissement en ne gardant que les billes en configuration monoclonale, par capture à l'aide de billes recouvertes de streptavidine qui viendront fixer la biotine liée aux amorces présentes sur les fragments amplifiés.

Ce sont finalement ces billes qui seront déposées dans une puce semi-conductrice contenant plus de 10 millions de puits où se déroulera le séquençage en lui-même. Dans ces puits, un nouveau brin d'ADN, complémentaire à celui fixé à la bille, va être synthétisé base par base. A chaque incorporation de nucléotide, la puce est capable de détecter les ions hydrogène produits, permettant la détermination du nucléotide intégré grâce au capteur de pH du séquenceur (voir Figure 28) (137).

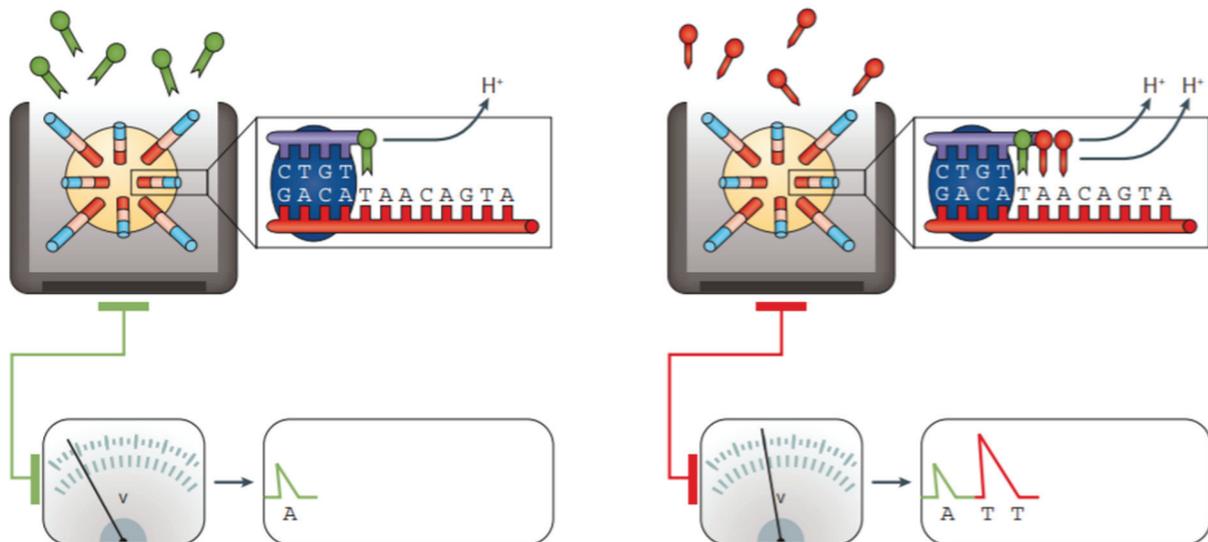


Figure 28. Séquençage sur puce semi-conductrice.

Après l'enrichissement des billes en configuration souhaitée, les billes sont disposées dans une plaque de telle sorte qu'une unique bille occupe un seul puits de réaction. Les nucléotides sont ajoutés aux puits un par un et la réaction d'élongation prend place lorsque le nucléotide est en effet celui complémentaire pour cette position. Lorsque l'incorporation est effective, un ion H^+ est généré, provoquant une variation de pH détectée par la puce. Si le nucléotide n'est pas complémentaire, un lavage permet son évacuation avant d'ajouter le suivant.

D'après Goodwin et al., 2016 (137).

A l'issue du séquençage, un fichier de données brutes avec les ionogrammes est obtenu, transféré alors vers le serveur Torrent.

Le premier outil utilisant cette technique fut commercialisé en 2010 par Ion Torrent avant d'être racheté par Life Technologies. Deux tailles de séquenceurs sont disponibles, avec différents modèles d'Ion PGM (Personal Genome Machine 314, 316, 318) ainsi que le Ion Proton d'une capacité supérieure (139).

2.2.6 Caractéristiques du séquençage

Lors d'une réaction de séquençage, diverses caractéristiques sont importantes à prendre en compte, afin de connaître la qualité du séquençage et d'estimer la fiabilité des données obtenues. Pour cela, différentes données peuvent être analysées.

2.2.6.1 Nombre et qualité des reads

Le nombre de reads est le nombre de fragments générés lors de la production de la librairie. Lors d'une réaction de séquençage, selon la qualité des reads en termes de taille, d'incorporation des adaptateurs et index, une partie plus ou moins importante des reads produits lors de la première étape sera fiable pour le séquençage. On les appelle les reads valides, et ce sont eux qui permettent alors une réaction de séquençage correcte après leur amplification.

Pour cette raison, le contrôle de la qualité des reads est importante ; en effet, si les reads ne sont pas suffisamment qualitatifs et ne répondent pas aux critères de sélection, leur alignement ultérieur à une séquence sera difficile. Pour évaluer l'aspect qualitatif des reads, le score de qualité phred peut être utilisé (voir Rappels bibliographiques, 2.2.1.2 Traitement bio-informatique des données) et permet ainsi d'estimer le niveau de confiance que l'on peut avoir pour les résultats obtenus avec chacun des fragments (142).

2.2.6.2 Couverture et profondeur

On appelle couverture de séquençage la région couverte par un minimum de 1 read. Elle peut être exprimée en % : si la couverture est de 90%, cela signifie que 90% de la séquence d'intérêt a été séquencé au moins une fois (un read) et que 10% ne présente pas de read et ainsi pas de données de séquençage. La couverture peut être discontinue, c'est-à-dire présentant des gaps (voir Rappels bibliographiques, 2.2.1.2 Traitement bio-

informatique des données). De ce fait, plus le nombre de reads valides est important, plus la couverture peut être élevée, évitant les régions peu voire pas séquencées car permettant statistiquement l'étalement des reads tout au long de la séquence d'intérêt.

La profondeur, exprimée en X, représente le nombre de reads présents pour chaque base séquencée. Par exemple, une profondeur de 10X signifie que, pour chaque position de la séquence d'intérêt, au moins 10 reads indépendants ont été utilisés pour le séquençage. Pour garantir une fiabilité optimale, la profondeur doit être suffisante et l'on doit ainsi s'assurer que le nombre de reads de qualité est effectif (143).

2.2.6.3 Longueur du séquençage

Il est important également de prendre en considération la longueur des contigs obtenus. En effet, lorsque le séquençage s'intéresse à une/plusieurs régions d'intérêt dans le génome, il est nécessaire de vérifier que la séquence entière a en effet été séquencée ou si la longueur ne correspond pas à celle souhaitée, ce qui rejoint la notion de la couverture.

2.2.7 Apports du séquençage de nouvelle génération

Ces techniques de NGS présentent un certain nombre d'intérêts pour la virologie par rapport au séquençage de Sanger, bien que ce dernier soit encore largement utilisé. Tout d'abord, il est possible, par ce type de séquençage, de faire des études de métagénomique virale, c'est-à-dire de caractériser les génomes viraux directement sur l'échantillon. La métagénomique virale ou virome se rapporte à toutes les populations virales retrouvées dans un échantillon, qu'il soit humain, animal ou qu'il provienne de l'environnement. Cela permet donc la détection de tous les virus présents dans l'échantillon d'intérêt sans *a priori*, pouvant aboutir notamment à la découverte de nouveaux virus, ou à la compréhension du rôle de certains virus dans des processus physiologiques ou pathologiques (144,145).

Le NGS représente également une avancée importante en termes de sensibilité et de résolution moléculaire, avec une capacité à distinguer les différentes séquences de différents VHC. Cela a ainsi permis l'identification d'un nouveau sous-type du génotype 4, mais également de recombinants du VHC (146). Grâce à sa sensibilité, le NGS apparaît également comme l'outil de choix pour l'étude des quasi-espèces. Dans le cas du VHC, cette technologie a en effet permis d'étudier les variants viraux minoritaires ainsi que la

diversité des quasi-espèces (147,148). Des ARN non codants en lien avec le CHC causé par le VHC ont également pu être mis en évidence (149).

Le séquençage par NGS montre aussi un intérêt pour ce qui est de la résistance du VHC aux traitements, et notamment aux AAD. En effet, une étude du gène NS3 par NGS a permis par exemple, chez des patients naïfs de traitement par AAD, d'observer dans plus de 80% des cas la présence de substitutions dans cette région, ayant un impact sur la sensibilité aux AAD (150). D'autres études ont montré que le NGS permettait aussi de détecter et suivre des mutations dans d'autres régions génomiques, et notamment NS5B, intéressantes à exploiter pour la mise en place du traitement, le suivi et la prédiction de la résistance au traitement (151,152). L'évolution virale du VHC au cours de l'infection chez l'humain, et par exemple la distribution des variants minoritaires peut également être suivie au cours du traitement, sous la pression de sélection des traitements (153–155). Ces méthodes de séquençage par NGS apportent ainsi de nombreuses applications pour le VHC, que ce soit en termes de génotypage, de détection des variants mais également pour la détermination des mutations potentiellement en lien avec la sensibilité aux traitements.

2.3 Analyse comparative entre les techniques de séquençage

La technique de Sanger a pendant longtemps, et encore actuellement, montré un succès important pour le séquençage, et notamment du VHC. Cependant, ce séquençage dit de première génération présente une évolutivité limitée, d'où l'avantage du séquençage par NGS, plus particulièrement en cas de nécessité de séquençage de nombreux échantillons, ou de génomes importants en termes de longueur. A noter cependant un des atouts de la technique de Sanger, qui est la longueur plus importante des fragments dans cette méthode de séquençage vis-à-vis de celles de NGS de deuxième génération. Ce séquençage NGS est ainsi appelé séquençage de deuxième génération, développé afin de pallier certaines limites retrouvées par le séquençage de Sanger, et notamment la sensibilité.

En plus de présenter un nombre d'applications important, la technologie NGS génère 3 à 4 fois plus d'informations en comparaison à la technique de Sanger, pour un coût moins important en termes de prix de séquençage par nucléotide (156). Notamment, la technologie Illumina présente un coût faible lorsque le multiplexage est suffisamment

important, c'est-à-dire lorsqu'un nombre assez élevé d'échantillons est séquençé au cours de la même réaction, soient 96 échantillons. Ce type de multiplexage n'est cependant pas adaptable à un séquençage utilisé en routine à l'hôpital, plus particulièrement pour le VHC. En effet, si la réaction de séquençage ne pouvait être réalisée que lorsqu'elle compte effectivement un minimum de 96 échantillons, les délais de rendu des résultats seraient de plusieurs mois.

Les performances sont intrinsèques à chaque technique et diffèrent notamment sur la longueur des reads, la précision ou encore le temps de séquençage (140,157,158). Entre la technique de Sanger mais aussi parmi les techniques de NGS sont également recensées des caractéristiques différentes (voir Tableau 5). Selon l'application souhaitée, chaque technique de nouvelle génération présente des avantages et des inconvénients (voir Tableau 6).

Tableau 5. Caractéristiques des différentes techniques de séquençage.

D'après Van Dijk et al., 2014 (159), ABM (123), Goodwin et al., 2016 (137), Université d'Angers (131), Pareek et al., 2011 (160), Biorigami, 2013 (161). NR : non renseigné.

	Sanger	Life Technologies		Illumina	Roche
		Ion Torrent PGM	SOLiD		
Date de lancement	1977	2006	2006	2006	2005
Méthode d'amplification	PCR	PCR émulsion	PCR émulsion	Bridge PCR	PCR émulsion
Méthode de séquençage	Synthèse	Synthèse	Ligation	Synthèse	Synthèse (pyroséquençage)
Longueur maximale des reads	1100 pb	400 pb	120 pb	150 à 300 pb	400 à 700 pb
Capacité par séquençage	2 à 300 kpb	2 Gpb	70 à 150 Gpb	1,5 à 200 Gpb	35 à 900 Mpb
Précision	99,9999%	99%	99,95 à 99,99%	99,90%	99%
Durée de séquençage	20min à 3h	2h	6 à 12j	26h à 14j	10 à 23h
Coût de séquençage	NR	445 à 850 €	5500 à 9500 €	650 à 18000 €	1000 à 5500 €
Coût de l'appareil	80 000 €	45 000 €	310000 à 540000 €	100000 à 620000 €	100000 à 450000 €
Coûts annexes	NR	18 000 €	50 000 €	NR	22500 à 27000 €

Tableau 6. Présentation des avantages et inconvénients des différentes technologies de séquençage par NGS.

Selon les données de Van Dijk et al., 2014 (159).

	Points forts	Limites
Ion Torrent	<ul style="list-style-type: none"> ▸ <u>Technologie semi-conductrice</u> : → pas besoin de balayage optique → pas besoin de nucléotides fluorescents 	<ul style="list-style-type: none"> ▸ <u>Taux d'erreurs</u> : élevé pour régions homopolymères
Illumina	<ul style="list-style-type: none"> ▸ <u>Leader dans l'industrie NGS</u> : → compatibilité d'une majorité des modes de réalisation de librairie → débit le plus élevé → coût de séquençage par base le plus bas en cas de multiplexage important ▸ <u>Longueur des reads</u> : > 300 pb → compatible avec la majorité des applications 	<ul style="list-style-type: none"> ▸ <u>Chargement de l'échantillon</u> : techniquement difficile ▸ <u>Diversité et complexité des séquences</u> : nécessaire → éventuel mélange avec d'autres séquences pour générer la diversité si besoin
SOLiD	<ul style="list-style-type: none"> ▸ <u>Performance</u> : deuxième système le plus efficace ▸ <u>Taux d'erreurs</u> : faible 	<ul style="list-style-type: none"> ▸ <u>Longueur des reads</u> : courte ▸ <u>Durée de séquençage</u> : longue ▸ <u>Assemblage de novo</u> : moins adapté ▸ <u>Utilisation</u> : moins importante qu'Illumina → kits de préparation des échantillons moins développés
454	<ul style="list-style-type: none"> ▸ <u>Longueur des reads</u> : 1kb maximum → facile à aligner à un génome de référence → avantage pour assemblage de novo → avantage pour métagénomique ▸ <u>Durée de séquençage</u> : assez rapide 	<ul style="list-style-type: none"> ▸ <u>Débit</u> : relativement faible ▸ <u>Réactifs</u> : coût élevé ▸ <u>Taux d'erreurs</u> : élevé pour régions homopolymères

Partie expérimentale

1. Introduction

La détermination du génotype infectant du VHC chez le patient ainsi que sa séquence, présentant d'éventuelles mutations pouvant impacter la susceptibilité au traitement, s'avère essentielle à la fois pour la mise au point du traitement le plus approprié possible, pour le suivi de l'apparition de résistances, mais également pour l'épidémiologie et la répartition des génotypes et/ou sous-types selon les populations étudiées ou la géographie.

Au laboratoire de virologie de l'hôpital de la Croix-Rousse, le séquençage du VHC était réalisé, jusqu'en 2017, par la méthode de Sanger (voir Rappels bibliographiques, **2.1 Le séquençage de Sanger**). En plus de renseigner le génotype, le séquençage a pour objectif d'apporter le maximum d'informations pertinentes pour le choix du traitement. De ce fait, le séquençage par cette technique de Sanger a évolué dans le temps avec l'évolution des AAD, en fonction des molécules utilisées en première intention et de la problématique des mutations de résistance associées. Ainsi, les premiers AAD utilisés en 2011 étant des anti-protéases, le génotypage a été initialement réalisé par séquençage de la région NS3 spécifique au génotype 1, puis avec des amorces pan-génotypiques lorsque des anti-protéases ont été disponibles pour d'autres génotypes. Lorsque les anti-NS5A sont arrivés en première ligne, le génotypage a alors été réalisé par séquençage de cette seule région NS5A. Les régions NS3 et NS5B étaient séquencées uniquement en seconde ligne, en cas d'échec d'amplification de la région NS5A ou après un échec au traitement, selon les molécules utilisées, pour caractériser la présence de mutations de résistance.

Par ailleurs, la détection de variants viraux minoritaires, de co-infections ou de quasi-espèces est rendue difficile par séquençage de Sanger. Pourtant, la co-infection est fréquente, notamment chez les patients appartenant à des groupes considérés à risque élevé comme les toxicomanes ou les hommes ayant des rapports sexuels avec des hommes (HSH) (162). En outre, cette méthode de séquençage présente des étapes non automatisées et nécessite une expertise technique assez longue à acquérir. Cela justifie la nécessité de procéder à un séquençage précis du VHC par des outils plus appropriés, afin d'optimiser la gestion clinique de l'infection par ce virus.

De ce fait, le séquençage du VHC par des automates récents présents à l'hôpital, d'ores et déjà utilisés pour le VIH, a été proposé pour son efficacité et ses capacités importantes. Ce circuit de séquençage développé par Vela Diagnostics (Allemagne) utilise une technique de NGS et plus spécifiquement la méthode de détection par semi-conducteur Ion Torrent. Une première version de la trousse HCV Vela Diagnostics a été tout d'abord évaluée mais ne donnait pas des résultats suffisamment satisfaisants pour le séquençage du VHC en routine, du fait de sa limitation aux génotypes 1a et 1b pour l'amplification de NS3 et NS5A, et à l'absence de détection de certaines mutations clés.

Une deuxième version a été développée et a fait l'objet d'essais que nous avons réalisés au dernier trimestre de l'année 2017. Ces essais s'étant avérés concluants, ils ont conduit à la mise en place de cette méthode en routine pour le séquençage du VHC des patients suivis à l'hôpital de la Croix-Rousse.

Cet automate et le kit utilisé permettent de cibler de manière systématique les 3 régions d'intérêt NS3, NS5A et NS5B et de séquencer simultanément 15 échantillons par réaction. Dans cette étude, la performance de la plateforme de NGS Ion Torrent de Vela Diagnostics est évaluée, basée sur l'analyse de 424 échantillons de plasmas ou sérums incluant 8 contrôles qualité. Plus précisément, les performances entre les réactions de séquençage, la sensibilité et la fréquence des génotypes et des mutations ont été analysées.

2. Méthodes

2.1 Echantillons cliniques

L'étude a été réalisée à l'Institut des Agents Infectieux (IAI) de l'hôpital de la Croix-Rousse, sur 424 échantillons cliniques, *sera* ou plasmas provenant de patients infectés par le VHC. 45 de ces échantillons, préalablement caractérisés par la technique de Sanger, ont été utilisés pour l'évaluation de la nouvelle trousse HCV de Vela Diagnostics et la comparaison de ce séquençage de nouvelle génération à la technique de Sanger. 8 contrôles externes de qualité ont également été intégrés à l'une des réactions de séquençage. Les autres échantillons ont été analysés entre mai 2018 et juin 2019, dans le cadre du suivi de routine des patients. Le séquençage de ces échantillons biologiques résulte de la demande d'hépatologues et d'infectiologues, majoritairement des Hospices Civils de Lyon, après consultation ; soit pour suspicion d'infection par le VHC lorsque les anticorps anti-VHC se sont révélés positifs ; soit pour des patients suivis au cours d'un traitement notamment. L'anonymisation a été effectuée avant chaque réaction de séquençage par l'utilisation du code de demande associé dans le système de gestion du laboratoire. Tous les échantillons ont été systématiquement inclus indépendamment du sexe du patient, de l'âge ou de la sévérité de l'infection. Les charges virales effectuées au sein de l'hôpital de la Croix-Rousse ont été évaluées par l'automate m2000 RealTime System de chez Abbott (charges virales effectuées à une date antérieure à avril 2019) ou par le Cobas® 6800 de Roche (charges virales effectuées à une date ultérieure à avril 2019).

En pratique, le séquençage des échantillons est réalisé pour les patients infectés, lors du premier diagnostic de l'infection, en cas d'échec de traitement, de prises de risques multiples et/ou de recontamination. Il est aussi effectué si le génotypage antérieur avait été réalisé par une technique ayant une moins bonne précision pour l'identification des sous-types et des recombinants, incluant le séquençage de la région 5' non codante ou le génotypage par hybridation moléculaire (Inno-Lipa).

2.2 Séquençage

L'analyse des échantillons, depuis l'extraction jusqu'au séquençage puis au traitement informatique des données, a été effectuée selon les instructions du fabricant (Vela Diagnostics, Allemagne). Le circuit de séquençage comprend l'extraction de l'ARN du VHC pour chaque échantillon et la préparation de la librairie sur le Sentosa™ SX101, la PCR en émulsion sur le Sentosa™ ST401e, l'enrichissement sur le Sentosa™ SQ401i, le séquençage sur le Sentosa™ SQ 301 et l'analyse des données grâce au Sentosa™ Reporter.

2.2.1 *Extraction et préparation de la librairie*

La première étape consiste en l'extraction de l'ARN du VHC des échantillons biologiques à partir d'un volume de 730 µL de sérum ou de plasma pour chaque patient. Chaque réaction de séquençage permet l'analyse de 15 échantillons ainsi qu'un échantillon contrôle. L'extraction est réalisée sur le Sentosa™ SX101 à l'aide du kit d'extraction Sentosa™ SX Virus Total Nucleic Acid Plus II Kit stocké à température ambiante. La purification de l'ARN est effectuée à l'aide de billes magnétiques. L'automate procède ensuite à la préparation de la plaque 96 puits qui sera utilisée ultérieurement dans le thermocycleur pour la RT-PCR. Pour cela, le kit de préparation de la librairie Sentosa™ SQ HCV Genotyping Reagents v2.0, conservé à -20°C, est nécessaire.

Cette étape précédant la PCR nécessite ainsi une préparation d'une vingtaine de minutes (décongélation et reconstitution des réactifs, préparation de la table de travail de l'automate) avant l'extraction et la préparation de la plaque PCR par l'appareil, étapes d'une durée de 2h10. La plaque PCR rendue par l'automate nécessite d'être scellée puis centrifugée à 100g pendant 30 secondes avant d'être placée dans un thermocycleur (Veriti 96 well thermal cyler, Applied Biosystem pour cette étude) pour la RT-PCR, d'une durée de 2h40, selon le cycle indiqué dans le tableau 7. Le volume de réaction est ici de 25 µL.

Tableau 7. Cycles suivis par le thermocycleur pour la synthèse puis l'amplification de l'ADN.

Etape	Température (°C)	Durée	Nombre de cycle(s)
Reverse-transcription ARN en ADN	25	2 min	1
	50	15 min	
Dénaturation	94	2 min	1
Amplification PCR 1 (dénaturation – hybridation – élongation)	94	15 sec	20
	57	30 sec	
	68	2 min	
Amplification PCR 2 (dénaturation – hybridation – élongation)	94	15 sec	20
	57	30 sec	
	68	1 min 30	
Extension finale	68	1 min	1
Maintien	10	∞	1

Au cours de l'amplification, 3 régions d'intérêt sont ciblées : 798 pb (paires de bases) de la région NS3 correspondant aux codons 1 à 266, 564 pb de la région NS5A (codons 13 à 200) et 1659 pb retrouvées au sein de NS5B (codons 7 à 559).

Après l'obtention de l'ADN s'effectue la préparation de la librairie. Pour cela, l'automate est chargé avec les différents consommables et réactifs, à l'aide des kits Sentosa™ SQ Genotyping Solutions et Sentosa™ SQ HCV Genotyping Reagent Plate stockés respectivement à 4°C et à -20°C ainsi que la plaque contenant l'ADN suite à la RT-PCR. Au cours de cette étape, l'automate permet la production de la librairie dans un unique tube. L'ADN de chaque échantillon est fragmenté aléatoirement de manière enzymatique, et à chacun des fragments d'un même échantillon sera ligué un index identique, permettant ultérieurement la reconnaissance de l'appartenance de chaque fragment aux différents échantillons. D'autre part, à ces fragments seront également fixés des adaptateurs, utiles au cours de l'étape de PCR en émulsion car nécessaires à l'attachement des fragments sur les billes par complémentarité (voir Rappels bibliographiques, **2.2.5 Technologie Ion torrent**). Cette étape demande 10 minutes de préparation de l'automate suivies de 3h10 de réactions réalisées par l'appareil, à l'issue desquelles un seul tube contenant tous les fragments mélangés de tous les patients, comportant adaptateurs et index, est obtenu.

2.2.2 PCR en émulsion, enrichissement et séquençage

A partir de la librairie, chacun des fragments va être amplifié à l'aide de billes par une PCR en émulsion à l'aide des kits Sentosa™ ST Template Solutions et Sentosa® ST Template Reagents. Pour cela, la librairie est chargée dans la cartouche de l'automate Sentosa™ ST401e. Au cours de cette étape, des microréacteurs sont formés par l'émulsion huile dans eau où seront idéalement retrouvés à la fois une bille et un fragment d'ADN (conformation monoclonale). La fixation de l'ADN sur la bille est rendue possible par complémentarité grâce aux adaptateurs préalablement ligués au cours de la préparation de la librairie. L'amplification monoclonale de chaque fragment se réalise ainsi dans les gouttelettes, aboutissant à des billes recouvertes par des copies du même fragment (voir Rappels bibliographiques, **2.2.5 Technologie Ion torrent**). La PCR en émulsion nécessite 5h45, incluant 15min de temps de préparation de l'automate et 5h30 de réaction.

Après rupture de l'émulsion et afin d'optimiser le séquençage, un enrichissement en billes présentant la conformation monoclonale est réalisé. Cela est effectué sur l'automate Sentosa™ ST401e par capture avec des billes présentant de la streptavidine en surface

(kit Sentosa™ ST Template ES beads stocké à 4°C). Cette protéine permet d'aller se lier à la biotine, ajoutée avec les amorces lors de la PCR en émulsion. Finalement, les fragments sont dénaturés par de la soude afin d'être monobrins. L'enrichissement est ainsi réalisé en 35min par l'automate, après une préparation de 15min.

L'étape ultime est réalisée par l'automate Sentosa™ SQ 301 sur une puce de séquençage où sont chargées les billes enrichies mélangées à l'enzyme et aux amorces nécessaires au séquençage (kit Sentosa™ SQ Sequencing Reagents). L'automate nécessitera également les 4 types de dNTP (kit Sentosa™ SQ Sequencing Reagents) et des solutions utiles au cours du séquençage pour l'ajustement du pH ainsi que le rinçage (kit Sentosa™ SQ Sequencing Solutions). Un par un, les nucléotides vont être présentés dans chaque puits de la puce où se trouve une bille recouverte d'un unique fragment à séquencer. La détection de la modification du pH à l'incorporation du nucléotide par complémentarité dans l'ADN en cours de synthèse permet de déterminer la séquence d'intérêt (voir Rappels bibliographiques, **2.2.5 Technologie Ion torrent**). Le temps de séquençage est de 5h.

2.2.3 Contrôles qualité

En termes de contrôles, on distingue le contrôle « système » (system control SC) qui est un plasmide fortement concentré du génome du Virus de la Mosaïque du Tabac (VMT) mais ne contenant pas de VHC ; le contrôle d'extraction (extraction control EC), petit plasmide de VMT faiblement concentré, qui sera distribué dans tous les échantillons incluant le SC ; et le CF-1 (control fragment). Ces contrôles permettent de surveiller et de mettre en évidence une éventuelle contamination, un échec d'extraction ou d'une autre étape, ainsi que la présence d'inhibiteurs dans les échantillons.

En effet, le SC est un contrôle qui doit être amplifié seulement par des amorces spécifiques au VMT et non par les amorces spécifiques au VHC. Cet échantillon étant traité comme les autres, il est considéré comme un contrôle positif pour le processus complet de séquençage, de l'extraction jusqu'à l'analyse des données de séquençage. S'il n'est pas amplifié par ses amorces, on peut conclure à un incident au cours du processus, auquel cas aucun échantillon, *a priori*, ne pourra être séquencé. Il sert également de

contrôle négatif pour le VHC (appelé alors non-template control), permettant de s'assurer de l'absence de contamination au cours du séquençage, puisque ne contenant pas de VHC. L'EC est ajouté automatiquement dans tous les échantillons au début de l'extraction. C'est un témoin positif pour cette étape. En cas de problème au cours de celle-ci, cela sera visible dans le rapport final. Cela permet de valider la négativité éventuelle des échantillons si l'EC est effectivement extrait dans les échantillons.

Le dernier contrôle, CF-1, est ajouté au cours de l'amplification par PCR en émulsion et sert à la vérification des étapes de préparation puis de séquençage (étapes ultérieures à la préparation de la librairie).

2.2.4 Analyse bio-informatique

(voir Rappels bibliographiques, 2.2.1.2 Traitement bio-informatique des données).

2.2.4.1 Analyse des données par les serveurs

A l'issue du séquençage, les données brutes avec les ionogrammes sont générées et transférées vers le serveur Sentosa™ SQ Suite pour l'analyse primaire puis vers Sentosa™ SQ Reporter pour l'analyse finale. De nombreuses données sont ainsi disponibles dans les rapports finaux générés (voir Annexe 1).

Le premier serveur va analyser les données générées, et attribue notamment le score de qualité à chacun des reads. Cela permet de filtrer les reads de faible qualité et de conserver uniquement ceux qui sont fiables. Ces données sont disponibles dans des fichiers Fastq qui pourront être consultés ultérieurement. D'autre part, les reads sont débarrassés de leurs adaptateurs dans une étape appelée « trimming ».

Le second serveur vérifie les contrôles qualité (voir Partie Expérimentale, **2.2.3 Contrôles qualité**) à l'aide des données de couverture, de rendement, de chargement de la puce. Puis l'alignement des reads est réalisé avec des séquences de référence. Ici, l'alignement des reads est réalisé sur 1306 génomes du VHC provenant de NCBI (National Center for Biotechnology Information) et LANL (Los Alamos National Lab) (163,164). D'autre part, les contrôles sont alignés avec des séquences correspondant au VMT. Les données de génotypage et les mutations d'intérêt seront cependant déduites seulement des régions d'intérêt précédemment ciblées au cours de l'amplification.

Premièrement, le serveur détermine le génotype selon la correspondance des séquences obtenues vis-à-vis des séquences des différents génotypes de référence, en ne considérant

que la région NS5B, suffisamment variable pour le génotypage. De ce fait, si cette région n'est pas amplifiée chez un patient, aucun génotype ne sera déterminé par le serveur, même si des séquences pour les autres régions NS3 et NS5A sont disponibles. Après récupération des données, ces séquences pourront cependant être « manuellement » confrontées à des séquences connues afin de déterminer le génotype, par exemple sur www.hcv.genopheno.org/ (165). Cette analyse des différents contigs obtenus est en réalité systématiquement réalisée après l'obtention des séquences par le serveur de Vela Diagnostics, afin de contrôler les génotypes et sous-types rendus, ainsi que de les déterminer dans des cas ambigus.

Puis, le serveur, via le « Torrent Variant Caller » est capable de définir les positions différentes de celles de(s) séquence(s) de référence, afin de déterminer les mutations présentes dans les séquences d'intérêt. Cela est réalisé pour les génotypes 1a, 1b et 3 uniquement, de par leur fréquence élevée, leur distribution et la disponibilité d'algorithmes d'interprétation. Pour cette raison, les mutations éventuelles d'un virus d'un autre génotype sont moins bien caractérisées et ne sont pas répertoriées dans l'algorithme d'interprétation de Vela Diagnostics qui ne référence ainsi aucune mutation dans le rapport de séquençage pour ces autres génotypes. De la même manière que pour le génotype, il sera alors possible de les déterminer manuellement à partir des séquences consensus que rendra le séquençage.

2.2.4.2 Données obtenues et rapports de séquençage (voir Annexe 1)

Sur le serveur Sentosa™ SQ Reporter sont retrouvés des fichiers relatifs à chaque échantillon. A l'issue du séquençage, différentes données y sont disponibles.

Les données de contrôle qualité sont répertoriées :

- Le rendement du séquençage : nombre de bases des reads présentant un adaptateur valide pour l'exécution du séquençage. Il peut présenter le statut « passed » si la valeur est supérieure à la valeur seuil recommandée, « failed » si la valeur est sous la valeur seuil minimum, ou « warning » si la valeur se trouve entre ces deux seuils ;
- Le chargement de la puce : quantité relative de puits contenant une bille. Il peut présenter le statut « passed » ou « failed » selon la valeur ;

- Les bases et reads valides/invalides : pourcentage de bases ou reads valides ou non, par rapport à la totalité de celles/ceux de la librairie utilisée pour le séquençage, il peut avoir le statut « passed » ou « failed » selon les valeurs seuils ;
- Les contrôles système comprenant le contrôle positif (VMT) dont la couverture médiane doit être supérieure au seuil, et le contrôle négatif (non-template control) dont la couverture médiane pour le VHC ne doit pas excéder le seuil. Ils peuvent avoir le statut « passed » ou « failed » ;
- Les contrôles qualité des échantillons incluant le contrôle d'extraction qui vérifie qu'un nombre suffisant d'amplicons des échantillons présente une couverture suffisante ; et que l'amplicon contrôle ait une couverture plus haute que le seuil.

On retrouve également les données globales de la réaction de séquençage regroupant les informations suivantes :

- La librairie finale : nombre de reads qui ont passé les filtres ;
- Le nombre de bases valides : nombre de bases obtenues après filtration et « trimming » ;
- Le nombre et le pourcentage de bases alignées ou non : quantité de bases provenant des reads qui ont été alignés correctement ou non à la séquence de référence ;
- Les reads invalides (voir ci-dessus) ;
- Les reads valides alignés ou non (voir ci-dessus) ;
- Le chargement de la puce ;
- Le niveau d'enrichissement des billes ;
- La quantité relative de billes mono ou polyclonales.

Sont également rapportées des données similaires pour chaque échantillon :

- Le nombre total de bases valides ;
- Le nombre de bases alignées et non alignées ;
- Le nombre de reads valides ;
- Le nombre de reads alignés, comprenant les reads spécifiques représentés par le pourcentage de reads alignés de manière correcte dans les régions cibles de la séquence de référence, les reads non spécifiques qui sont la quantité de reads

alignés en dehors de ces régions, et les reads rejetés. Les reads sont susceptibles d'être rejetés lorsque leur longueur est trop courte, s'ils sont de faible qualité (évaluation par le score de qualité), s'ils ne sont pas utilisés dans l'analyse ou s'ils sont peu alignés aux régions cibles ;

- Les reads non alignés ;
- La longueur moyenne de tous les reads générés pour l'échantillon ;
- Le pourcentage de bases G et C retrouvées dans la séquence (GC%).

Sont ensuite listés le nombre de contigs, en lien avec la région ciblée (contig pour NS3, NS5A ou NS5B) et la profondeur de chacun d'entre eux. La profondeur pour chaque position des 3 régions d'intérêt est également disponible, avec l'attribution d'un statut « passed », « warning » ou « failed ».

Le génotype voire sous-type retrouvé grâce au séquençage de NS5B est indiqué. Dans le cas où des mutations sont trouvées, elles sont répertoriées par région. La position est précisée, et la profondeur ainsi que la fréquence de cette mutation sont également définies.

3. Résultats

3.1 Caractéristiques des patients

Dans cette étude, 424 échantillons ont été collectés afin d'être séquencés par l'appareil de séquençage Sentosa SQ HCV Genotyping Assay de Vela Diagnostics (Allemagne) à l'hôpital de la Croix-Rousse, au cours de 29 réactions de séquençage. Parmi eux, 45 ont également été séquencés et génotypés par la méthode de Sanger afin de permettre une comparaison des deux techniques et de valider la fiabilité de la technique de NGS, et 8 étaient des contrôles qualité. Ces échantillons représentent ainsi les patients infectés par le VHC et suivis pour cette infection, *a minima*, à l'hôpital de la Croix-Rousse. Cette étude n'a pas nécessité d'accord particulier provenant d'un comité éthique. La détermination de la charge virale des échantillons prélevés au sein de l'hôpital de la Croix-Rousse fut réalisée préalablement au séquençage. Cependant, la charge virale de certains échantillons envoyés par des laboratoires extérieurs n'était pas connue.

Parmi les échantillons, 69,32% d'entre eux provenaient d'hommes pour 30,68% de femmes, représentant 287 hommes et 127 femmes (voir Tableau 8). En ce qui concerne l'âge des patients infectés, tous présentaient une année de naissance ultérieure à 1930. 34 patients étaient âgés de moins de 30 ans soient 8,21%, et 229 patients étaient âgés de plus de 50 ans soient 55,31% (voir Tableau 8 et Figure 29).

Les échantillons présentaient une charge virale quantifiable entre 1,20 log UI/mL (valeur minimum) et 7,70 log UI/mL (valeur maximum), avec une moyenne de 5,55 log UI/mL et une médiane de 5,70 log UI/mL. 23 charges virales parmi les échantillons n'étaient pas connues, et 1 était non détectable, avec un nombre de copies inférieur au seuil de détection.

Tableau 8. Caractéristiques des patients.

Caractéristique	
Sexe	Femmes : n = 127 (30,68%)
	Hommes : n = 287 (69,32%)
Âge	< 30 ans : n = 34 (8,21%)
	30-50 ans : n = 151 (36,47%)
	> 50 ans : n = 229 (55,31%)
Charge virale (log UI/mL)	Moyenne = 5,55
	Médiane = 5,70
	Min = 1,20
	Max = 7,70

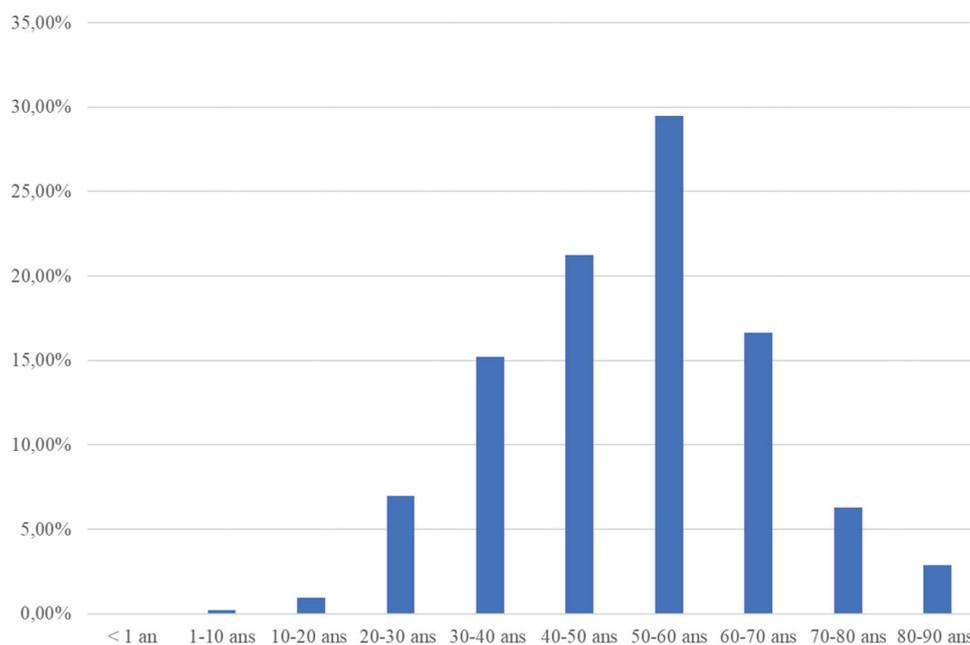


Figure 29. Histogramme de distribution de l'âge des patients.

L'âge de chacun des patients a été calculé selon leur date de naissance, à la date de séquençage. Le pourcentage de chaque classe d'âge est indiqué, calculé par le nombre de personnes dans chaque tranche d'âge par rapport au nombre total de patients.

3.2 Comparaison à Sanger

Initialement, l'automate de séquençage Sentosa SQ HCV Genotyping Assay de Vela Diagnostics (Allemagne) a été testé sur 3 réactions de séquençage, soient 45 échantillons. Ce test a permis la comparaison entre les données obtenues préalablement par technique de Sanger (génotype et mutations de résistance éventuelles sur NS5A *a minima*, voire sur NS3 et NS5B selon les cas) et les données obtenues par le nouveau circuit de séquençage de Vela Diagnostics afin d'estimer les performances, la fiabilité et la sensibilité de la technique de NGS.

Tout d'abord, la cohérence des résultats obtenus entre les deux types de séquençage a été évaluée. Le séquençage par ce circuit Vela Diagnostics a révélé une concordance de génotypage de 98% (42 génotypes identiques sur 43). La concordance n'est pas totale du fait d'un échantillon présentant une charge virale basse (2 log) pour lequel l'amplification, et donc le séquençage, furent difficiles. De ce fait, aucun génotype ni sous-type n'ont pu être établis par NGS pour cet échantillon, alors qu'il avait été décrit que le virus était de génotype 1b selon la technique de Sanger. Concernant le sous-typage, la concordance est de 92% (36 sous-types identiques sur 39) par rapport au séquençage de Sanger. Pour deux échantillons, le séquençage de Vela Diagnostics a rendu un génotype 1 pour un sous-type g, alors qu'il avait été établi que le virus infectant appartenait au génotype 1e selon Sanger ; et un génotype 4a pour un génotype initialement considéré 4v par Sanger. A partir de ces résultats différents, il est cependant difficile de savoir quelle technique permet de donner le génotypage correct.

Dans ces cas-là, la soumission des séquences consensus des différents gènes obtenues sur geno2pheno permet d'orienter les résultats vers un génotype et un sous-type. Pour le premier patient de génotype 1 présentant deux sous-types différents par les deux techniques, l'alignement des contigs obtenus par NGS correspondant aux trois régions d'intérêt à des séquences de référence sur cette base de données a rendu un génotype 1e, avec une similarité de séquence supérieure à 87%. Ce résultat est concordant avec le génotype déterminé par séquençage de Sanger. Pour chacun des 3 contigs, la similarité de séquence est également importante avec les séquences de référence de génotype 1g, bien que moins importante par rapport aux séquences de génotype 1e.

Pour le deuxième patient, l'alignement de l'unique contig de NS5B obtenu par NGS rend un génotype 4a sur geno2pheno avec 88% de similarité, mais une similarité de 87% est

également retrouvée pour les séquences de génotype 4v. Le génotypage effectué par technologie de Sanger avait été réalisée grâce à la séquence de NS5A. L'alignement de cette séquence, précédemment obtenue par Sanger, aux séquences de références sur geno2pheno détermine une similarité à 87% pour les génotypes 4v et 83% pour les génotypes 4a.

Les résultats sont ainsi discordants entre les techniques de séquençage et notamment pour ce qui est de l'analyse des séquences. Cela peut être expliqué par des divergences dans les jeux de séquences de référence utilisées pour l'alignement, notamment pour les sous-types rares qui ne sont pas forcément référencés pour toutes les régions virales.

Globalement, le séquençage par NGS était ainsi capable de déterminer correctement le génotype et le sous-type, de la même manière que par séquençage de Sanger. D'autre part, tandis que la technique de Sanger n'avait pas permis le sous-typage de 4 des échantillons (1 de génotype 1 et 3 de génotype 2) le séquençage par NGS a rendu des sous-types 1a, 2q, 2m et 2b.

Parmi les 43 échantillons de patients séquencés par Sanger, 7 d'entre eux avaient nécessité un séquençage de NS3 et/ou NS5B en plus de NS5A, automatiquement réalisé. Concernant les mutations retrouvées, des mutations supplémentaires, non retrouvées par séquençage de Sanger, ont été retrouvées sur NS5A grâce au séquençage par NGS pour 33% des patients (14 patients sur 43), et sur NS3 et/ou NS5B pour 30,23% des patients (13 patients sur 43). Cependant, pour au moins 2 patients pour lesquels le séquençage NGS a permis la détection de mutations supplémentaires sur NS5A (échantillons 9 et 10), une contamination par un autre échantillon a été suspectée. Du fait du génotype de ces échantillons (6e), le séquençage par NGS ne permet habituellement pas de rendre les mutations. Les mutations effectivement trouvées par NGS correspondent ainsi aux séquences contaminantes. Cela réduit ainsi à 28% le taux de patients pour lesquels le séquençage NGS a permis la détection de mutations supplémentaires sur NS5A.

D'autre part, certaines mutations retrouvées en Sanger n'étaient pas toutes retrouvées en NGS. Ce fut le cas pour 2 patients pour le séquençage de NS3, soient 4,65% ; 2 patients pour NS5A, soient 4,65% ; et 1 patient pour le séquençage de NS5B, soient 2,33% (voir Tableau 9).

Finalement, la capacité de la technique de NGS à détecter les mélanges de virus et ainsi les co-infections par au moins deux virus différents a été testée. Pour cela, différents

échantillons présentant des charges virales connues et des génotypes préalablement déterminés par séquençage de Sanger ont été mélangés. Le séquençage du premier mélange de deux échantillons, présentant des charges virales respectivement élevées à 4,8 et 4,1 log UI/mL, a permis la détection des 2 génotypes 1a et 2c dans les proportions relatives attendues, soient 4748 reads séquencés pour le premier échantillon et 4701 pour le second. Le deuxième mélange, avec un génotype 3a à 3,1 log UI/mL et 1b à 6,2 log UI/mL, n'a cependant pas été détecté. Seul le second génotype, présentant la charge virale la plus élevée, a pu être déterminé par séquençage de NGS.

Tableau 9. Comparaison des données obtenues par séquençage de Sanger et par NGS, pour chaque région d'intérêt.

Patient	NS3		NS5A		NS5B		Génotype*
	Sanger	NGS	Sanger	NGS	Sanger	NGS	
1	-	-	93H	62A 93H	-	-	3a
2	-	NA	-	NA	-	NA	4r
3	-	NA	93H	NA	-	NA	2a
4	-	56F	-	30R 54R 58P	-	-	1b
5	-	-	-	54N 58P	-	-	1/1a
6	-	-	93H	30R 54H 93H	-	414L 446Q 556G	1b
7	-	80K	58P	58P	-	-	1a
8	-	NA	-	NA	-	NA	6e
9	-	-	-	62H 93H	-	-	6e + contamination 3a suspectée
10	-	-	-	62A 93H	-	-	1e + contamination 3a suspectée /1g
11		56F 170V		30R		499A	1b
12	-	NA	-	NA	-	NA	1b/-
13	-	-	-	62T 93H	-	445F	3a
14	-	NA	-	NA	-	NA	2/2q
15	-	NA	-	NA	-	NA	4d
16	-	56F 122T	31M	-	-	-	1b
17	-	-	93H	93H	282T	282T	3a
18	-	-	-	93H	-	-	3a
19	-	-	28V 30R 58P	28V/A 30R 58P 62D	-	556G	1a
20	-	NA	-	NA	-	NA	2c

Patient	NS3		NS5A		NS5B		Génotype*
	Sanger	NGS	Sanger	NGS	Sanger	NGS	
21	-	-	-	-	-	-	3a
22	-	80K	-	-	-	316Y 556G	1a
23	-	-	30H 93H	30H 93H	-	-	1a
24	-	NA	-	NA	-	NA	4f
25	-	56F	-	-	-	159F 316N 556G	1b
26	36L 168Q	-	-	93H	-	-	3a
27	-	NA	-	NA	-	NA	2/2m
28	122N	122N/T	-	30Q	-	-	1b
29	-	-	-	-	-	-	3a
30	-	-	-	-	-	-	1a
31	-	NA	-	NA	-	NA	2k
32	-	NA	-	NA	-	NA	2/2b
33	-	-	31M 93H	31M 93H	159F 316N	159F 316N	1b
34	-	-	-	-	-	-	3a
35	-	-	-	93H	-	-	3a
36	36L 168Q	-	92E 93H	93H	282T	-	3a
37	-	-	-	-	-	-	3a
38	-	-	-	30K	-	-	3a
39	-	36L 55A 80K	-	-	-	-	1a
40	-	NA	-	NA	-	NA	4v/4a
41	-	-	31M 92T	31M 92T	-	159F 316N 556G	1b
42	-	56F 122T	-	-	-	-	1b
43	-	56F 122T	-	-	-	-	1b

* Un unique génotype est indiqué lorsque les deux méthodes de séquençage ont en effet déterminé le même génotype et sous-type. Dans le cas où les deux génotypes et/ou sous-types varient, le premier génotype indiqué est celui déterminé par séquençage de Sanger, tandis que le deuxième est celui déterminé par NGS. NA : non applicable, - : aucune mutation déterminée.

3.3 Caractéristiques de séquençage

Pour chaque réaction de séquençage effectuée par l'automate de Vela Diagnostics, les différentes caractéristiques de la réaction sont obtenues sur les rapports qui sont automatiquement générés. Des informations globales sur la réaction de séquençage en elle-même sont disponibles (nombre de reads total dans la librairie, nombres de bases et reads valides, reads alignés ou non parmi les reads valides), ainsi que des informations plus précises pour chacun des échantillons (voir Partie expérimentale, 2.2.4.2 Données obtenues et rapports de séquençage).

Au cours de ces 29 réactions de séquençage, la moyenne du nombre total de reads par librairie était de 3 772 914 reads (minimum : 1 625 066 reads, maximum : 5 201 564 reads) avec une moyenne de 3 729 756 reads valides, soient 563 536 881 bases valides en moyenne (voir Tableau 10). Parmi ces reads, 16 726 642 en moyenne étaient en effet alignés à la séquence de référence, soient 98,55% des reads valides, et 50 847 reads en moyenne ne l'étaient pas, soient 1,36% des reads valides (voir Tableau 10).

Tableau 10. Données de séquençage obtenues pour les 29 réactions analysées.

Réaction séquençage	Nombre total reads	Reads valides (n)	Bases valides (n)	Reads non mappés (n)	Reads non mappés (%)	Reads mappés (n)	Reads mappés (%)
1	3064565	3022544	435264557	69883	2.31	2952661	97.69
2	4511532	4464799	683231279	75253	1.69	4389546	98.31
3	3983465	3946703	602380631	54483	1.38	3892220	98.6
4	1625066	1611738	237772041	19900	1.23	1591838	98.77
5	4810829	4766976	766482424	59841	1.26	4707135	98.74
6	3425970	3387110	565855853	38421	1.13	3348689	98.87
7	3112467	3077780	507181623	32419	1.05	3045361	98.95
8	4744374	4699771	731957133	70558	1.5	4629213	98.5
9	4317727	4280162	702774910	44296	1.03	4235866	98.97
10	4150477	4109620	662499196	67380	1.64	4042240	98.36
11	4543253	4505508	687766084	50077	1.11	4455431	98.89
12	3031932	3004442	464548251	42704	1.42	2961738	98.58
13	3470946	3431734	526650761	59286	1.73	3372448	98.27
14	3136560	3103973	500149509	34128	1.1	3069845	98.9
15	4319746	4281650	682060594	47448	1.11	4234202	98.89
16	2352320	2319266	344544112	26614	1.15	2292652	98.85
17	4144137	4085397	625968169	53802	1.32	4031595	98.68
18	2745941	2683814	395625369	28087	1.05	381039987	96.31
19	2933429	2854408	403072883	38183	1.34	2816225	98.66
20	5201564	5152279	789466706	57237	1.11	5095042	98.89
21	2644099	2602425	382830983	37017	1.42	2565408	98.58
22	3818184	3762403	547827240	38374	1.02	3724029	98.98
23	4021983	3966162	544397652	56158	1.42	3910004	98.58
24	3980581	3929369	568242555	48674	1.24	3880695	98.76
25	3405523	3368895	481510478	36236	1.08	3322659	98.92
26	3696955	3656593	500300010	57568	1.57	3599025	98.43
27	4324594	4280805	552017589	98584	2.3	4182221	97.7
28	4929056	4885353	726586525	73136	1.5	4812217	98.5
29	4967242	4921266	723604435	58835	1.2	4862431	98.8
Moyenne	3772914.38	3729756.72	563536881.10	50847.66	1.36	16726642.17	98.55
Maximum	5201564	5152279	789466706	98584	2.31	381039987	98.98
Minimum	1625066	1611738	237772041	19900	1.02	1591838	96.31
Médiane	3980581	3929369	552017589	50077	1.26	3892220	98.68

Ces données sont également disponibles individuellement, par échantillon, ainsi que d'autres informations complémentaires. Notamment, la quantité de reads alignés spécifiques, non spécifiques et les reads rejetés parmi les reads alignés est renseignée. D'autre part, la longueur moyenne des reads ainsi que la fréquence des bases G et C (GC%) sont indiquées. Le GC% d'une séquence est la proportion relative de bases qui sont déterminées comme étant soit une guanine (G), soit une cytosine (C) parmi toutes les bases de la séquence. Cette dernière information peut expliquer, lorsque la valeur est élevée, une difficulté à l'amplification et donc au séquençage. En effet, les liaisons établies par ces deux bases sont plus stables que celles établies par l'adénine (A) ou la thymine (T), ce qui entraîne, lorsque le GC% est élevé, une augmentation de la température de fusion d'où une dénaturation des brins possiblement insuffisante. De ce fait, le brin matrice est possiblement moins accessible, rendant le séquençage moins qualitatif.

Ce bilan a ainsi permis d'obtenir les moyennes de chacune de ces données individuelles. Il en ressort que le nombre moyen de reads valides était de 251 457, soit 6,74% du nombre total de reads ; ce qui représente une moyenne de 37 796 672 bases valides, soit 6,74% du nombre total de bases. Parmi ces bases, le rapport permet de distinguer les reads qui ont été alignés ou non. Au cours des différentes réactions, le nombre moyen de reads alignés était de 247 647 (97,74% des reads valides), tandis que celui du nombre de reads non alignés était de 3 406, soit 1,83% en moyenne. Au sein de ces reads alignés, 92,38% des reads en moyenne étaient spécifiques, alors que 4,69% en moyenne ne l'étaient pas, et 2,94% des reads furent rejetés (voir Tableau 11). En ce qui concerne la longueur moyenne des reads, elle était de 150,27, et le GC% estimé à 56,05 en moyenne (voir Tableau 11).

Tableau 11. Données de séquençage obtenues pour les 424 échantillons analysés.

	Moyenne	Maximum	Minimum	Médiane
Reads valides (n)	251457,31	1131915	24	220455
Reads valides (%)	6,74	30,38	0	6,135
Bases valides (n)	37796672,57	167810730	1751	33401920
Bases valides (%)	6,74	30,98	0	6,125
Reads mappés (n)	247647,62	1119710	18	217653
Reads mappés (%)	97,74	99,53	8,72	98,68
◦ Reads spécifiques (%)	92,38	99,17	0,01	96,68
◦ Reads non spécifiques (%)	4,69	95,17	0,04	0,97
◦ Reads rejetés (%)	2,94	53,85	0,65	2,04
Reads non mappés (n)	3406,21	24492	5	2688
Reads non mappés (%)	1,83	25	0,47	1,32
Longueur reads	150,27	184	69	152
GC%	56,05	67,92	5,73	56,515

Comme indiqué précédemment, la moyenne de charge virale des échantillons était de 5,55 log UI/mL. Alors que le seuil limite de détection prévu par Vela Diagnostics est de 1000 UI/mL pour les génotypes 1a, 1b, 2, 3 et 4 ; 15 échantillons avec une charge virale inférieure à ce seuil ont néanmoins pu être évalués. L'automate a montré son efficacité pour séquencer les 3 régions de 9 échantillons présentant une charge virale inférieure à ce seuil, avec un minimum de 1,2 log UI/mL (génotype 1a), permettant de rendre un génotype pour 8 de ces 9 échantillons. 6 autres échantillons à charge virale inférieure au seuil ont été inclus dans l'étude : 1 échantillon pour lequel la région NS5A seule a pu être séquencée sans rendu de génotype, 3 échantillons pour lesquels les régions NS5A et NS5B ont été séquencées avec rendu de génotype et 2 échantillons en échec de séquençage pour les 3 régions, sans possibilité de rendre un génotype. D'autre part, le seuil de détection de 2000 UI/mL pour les génotypes 5 et 6 n'a pas pu être testé, étant donné que tous les échantillons séquencés de génotype 6 recensés au cours de cette étude présentaient une charge virale supérieure au seuil de détection, et qu'aucun génotype 5 n'a été répertorié.

Pour chaque échantillon, il est possible de consulter le nombre moyen de reads pour chaque région d'intérêt (NS3, NS5A et NS5B), la profondeur moyenne ainsi que la longueur moyenne.

Pour NS3, le nombre de reads moyen était de 8 300 avec une profondeur moyenne de 1 422. La longueur moyenne du contig correspondant à cette région était de 934 bases (voir Tableau 12), pour une région à amplifier initiale de 798 pb.

Pour NS5A, le nombre de reads moyen était de 5 530 et la profondeur moyenne était de 1 463. Pour cette région, la longueur moyenne du contig obtenu était de 618 bases (voir Tableau 12) alors que les amorces sont destinées à amplifier 564 pb.

Pour la dernière région NS5B, le nombre de reads moyen était de 13 528 avec une profondeur moyenne de 1 287. La longueur moyenne du contig était ici de 1 579 bases (voir Tableau 12) pour les 1659 pb cibles.

Tableau 12. Données obtenues pour chaque région d'intérêt.

	NS3			NS5A			NS5B		
	Nombre de reads	Profondeur	Longueur moyenne	Nombre de reads	Profondeur	Longueur moyenne	Nombre de reads	Profondeur	Longueur moyenne
Moyenne	8300,46	1422,15	933,57	5530,00	1462,90	618,16	13528,50	1287,14	1579,07
Maximum	15119,00	2048,00	1121,00	9876,00	2151,00	1698,00	26200,00	2063,00	1970,00
Minimum	169,00	26,00	283,00	91,00	25,00	386,00	206,00	10,00	561,00
Médiane	10723,00	1969,00	926,00	7106,00	1980,00	602,00	18723,00	1964,50	1697,00

3.4 Capacité d'amplification des 3 gènes

Contrairement au séquençage par Sanger automatiquement réalisé uniquement pour NS5A, et complété si besoin par celui de NS3 et NS5B, l'automate de Vela Diagnostics procède au séquençage systématique de ces 3 régions d'intérêt, et la séquence de NS5B est utilisée pour la détermination du génotype. Les résultats présentés ici ont été obtenus sur 423 échantillons, de par l'exclusion volontaire d'un échantillon à charge virale indétectable pour lequel, de manière cohérente, aucune des 3 régions n'a été amplifiée ; ceci afin de ne pas biaiser l'analyse des capacités réelles.

Pour NS3, l'amplification a été effectuée dans 85,82% des cas, comptant un contig unique dans 84,87% des cas (359 patients sur 423) et 2 contigs dans 0,95% des cas (4 patients sur 423 dont l'un des deux mélanges volontaires). 14,18% des échantillons n'ont pu être séquencés, soient 60 patients, rendant impossible la détermination d'éventuelles mutations dans cette région (voir Tableau 13). Ces 60 échantillons présentaient une médiane de charge virale à 5,32 log UI/mL, avec la charge virale la plus basse à 1,2 log UI/mL. Bien que le génotype de 35,48% de ces échantillons n'ait pu être déterminé, les autres échantillons en échec d'amplification de NS3 ont été répertoriés en génotypes 4 à 19,35%, génotypes 3 à 17,74%, génotypes 1 à 16,13% et génotypes 2 à 11,29%.

Pour NS5A, l'amplification a pu être réalisée dans 90,54% des cas, comptant un seul contig dans 88,18% des cas (373 patients sur 423), 2 contigs dans 1,89% des cas (8 patients sur 423) et 3 contigs dans 0,47% des cas (2 patients sur 423). 9,46% des échantillons n'ont pu être séquencés, représentant 40 patients pour lesquels les mutations sur cette région n'ont pu être définies (voir Tableau 13). Ces 40 échantillons montraient une médiane de charge virale à 5,30 log UI/mL, avec un minimum à 1,2 log UI/mL. Ici, le génotype correspondant n'a pu être déterminé dans 47,50% des cas, et les autres échantillons en échec d'amplification de NS5A ont été classés en génotypes 4 à 27,50%, génotypes 2 à 20% et génotypes 3 à 5%.

Pour NS5B, région d'intérêt pour l'établissement du génotype, l'amplification a été possible dans 93,14% des cas. Un contig fut retrouvé dans 85,58% des cas (362 sur 423 patients), 2 dans 7,09% des cas (30 patients sur 423 dont l'un des deux mélanges volontaires) et 3 dans 0,47% des cas avec 2 patients sur 423. Ainsi, 6,86% des patients n'ont pu être séquencés sur NS5B, soient 29 patients pour lesquels ni le génotype, ni le sous-type, ni les mutations n'ont pu être rendus par l'analyse automatisée de Vela

Diagnostics (voir Tableau 13), pour une charge virale médiane à 5,37 log UI/mL. Ici, l'analyse des séquences disponibles pour les autres régions d'intérêt par comparaison à des séquences de référence devient ainsi indispensable pour pouvoir rendre un génotypage, bien que le rapport systématique de Vela Diagnostics ne permette pas de l'effectuer.

Parmi les 423 échantillons, les 3 gènes de 338 échantillons ont pu être amplifiés avec succès (79,91%) pour une charge virale médiane à 5,78 log UI/mL (minimum 1,2 log UI/mL, maximum 7,7 log UI/mL). Ces échantillons étaient de génotype 1 à 55,03%, 3 à 27,81%, 4 à 6,51%, 2 à 5,03%, 1b/2k à 2,66%, 6 à 1,18% et de génotype indéterminé pour 1,78%.

57 échantillons (13,48%) présentaient 1 échec d'amplification et 12 échantillons (2,84%) présentaient 2 échecs d'amplification parmi les 3 gènes d'intérêt. La médiane des charges virales pour les échantillons avec 1 échec d'amplification parmi les 3 régions d'intérêt a été estimée à 5,32 log UI/mL (minimum 2,08 log UI/mL, maximum 6,88 log UI/mL), et le séquençage a montré que 28,07% d'entre eux étaient des génotypes 4, 19,30% des génotypes 2, 19,30% des génotypes 3, 17,54% des génotypes 1 et 15,79% présentaient un génotype indéterminable par Vela Diagnostics.

La médiane des charges virales pour les échantillons avec 2 échecs d'amplification parmi les 3 régions d'intérêt a été estimée à 5,13 log UI/mL (minimum 2 log UI/mL, maximum 7,09 log UI/mL), et le séquençage a montré que 25% d'entre eux étaient des génotypes 4, 16,67% des génotypes 2, 8,33% des génotypes 3 et 50% présentaient un génotype indéterminable par Vela Diagnostics. Finalement, on compte 16 échantillons (3,78%) pour lesquels aucun des gènes n'a pu être amplifié. La médiane des charges virales parmi ces 16 échantillons était de 5,34 log UI/mL (minimum 1,2 log UI/mL, maximum 6,70 log UI/mL).

Tableau 13. Bilan d'amplification des 3 gènes d'intérêt.

	NS3	NS5A	NS5B
	Nombre (%)	Nombre (%)	Nombre (%)
Amplification :	363 (85,82)	383 (90,54)	394 (93,14)
1 contig	359 (84,87)	373 (88,18)	362 (85,58)
2 contigs	4 (0,95)	8 (1,89)	30 (7,09)
3 contigs	0 (0)	2 (0,47)	2 (0,47)
Non amplification	60 (14,18)	40 (9,46)	29 (6,86)

3.5 Détermination du génotype

L'objectif majeur du séquençage des virus infectants chez un patient est la détermination du génotype. Elle est réalisée à partir de la séquence obtenue de NS5B uniquement pour le séquenceur Sentosa™ SQ 301 de Vela Diagnostics. Ainsi, pour tout patient dont cette région ne peut être séquencée, l'analyse automatisée ne permet pas de rendre le génotype ni le sous-type. Cependant ce génotype peut être déterminé en alignant la séquence des autres gènes amplifiés à des séquences de référence dans des bases de données ad-hoc. Le bilan de séquençage effectué sur les 424 échantillons provenant de patients a permis d'établir la distribution des génotypes et sous-types retrouvés. Les contrôles qualité ont ici été exclus des analyses de répartition des génotypes ; ainsi, la distribution présentée représente celle de 416 échantillons. Parmi les échantillons pour lesquels le génotypage a été possible, un unique génotype fut déterminé pour la majorité des patients, tandis que deux génotypes furent établis pour 11 patients, démontrant la coexistence de virus dans l'échantillon. Parmi ces 11 échantillons où ont été retrouvés 2 génotypes, l'un d'entre eux était un mélange volontaire de 2 échantillons, effectué pour tester la capacité de l'automate à détecter les co-infections (voir Partie Expérimentale, **3.2 Comparaison à Sanger**). Le second mélange effectué n'a pas été détecté, rendant ainsi un unique génotype à l'issue de l'analyse. L'existence de 2 génotypes différents dans un échantillon peut provenir d'une co-infection, chez le patient, par 2 virus de génotypes différents, mais peut éventuellement résulter d'une contamination, au moment du séquençage, d'un échantillon par un autre, comme cela est suspecté pour plusieurs des 10 autres échantillons. Dans ce cas, la contamination éventuelle est évaluée en fonction de l'homologie de séquence avec celle d'un autre patient de la série et le repassage de l'échantillon dans une nouvelle série.

Concernant les génotypes, le génotype 1 fut le plus fréquent, retrouvé pour 203 échantillons (47,54%). Le second génotype le plus fréquent était le génotype 3 avec 107 échantillons (25,07%). Puis, le génotype 4 a été retrouvé chez 43 patients (10,07%), le génotype 2 chez 25 patients (5,85%), le génotype recombinant particulier 1b/2k chez 9 patients (2,11%) et le génotype 6 chez 4 patients (0,94%) (voir Figure 30). Pour 36 échantillons, aucun génotype n'a pu être déterminé par le circuit Vela Diagnostics, soient 8,43%. Parmi eux, l'impossibilité de déterminer le génotype provient d'une absence

d'amplification de NS5B pour 27 échantillons dont celui déjà évoqué à charge virale indétectable (75%), et le génotype ne put être défini malgré l'amplification de NS5B pour les 9 autres (25%). D'autre part, un unique cas parmi tous les échantillons, où le génotype (1a) a pu être identifié sans amplification de NS5B, a été recensé.

Parmi les patients pour lesquels le séquençage par le circuit Vela Diagnostics n'a pas permis la détermination du génotype, l'analyse des séquences obtenues via geno2pheno ainsi que le repassage des échantillons dans des réactions de séquençage ultérieures ont cependant rendu possible le génotypage pour la majorité des échantillons, avec finalement 6 patients seulement pour lesquels aucun génotype n'a pu être rendu (échantillon à charge virale indétectable exclus).

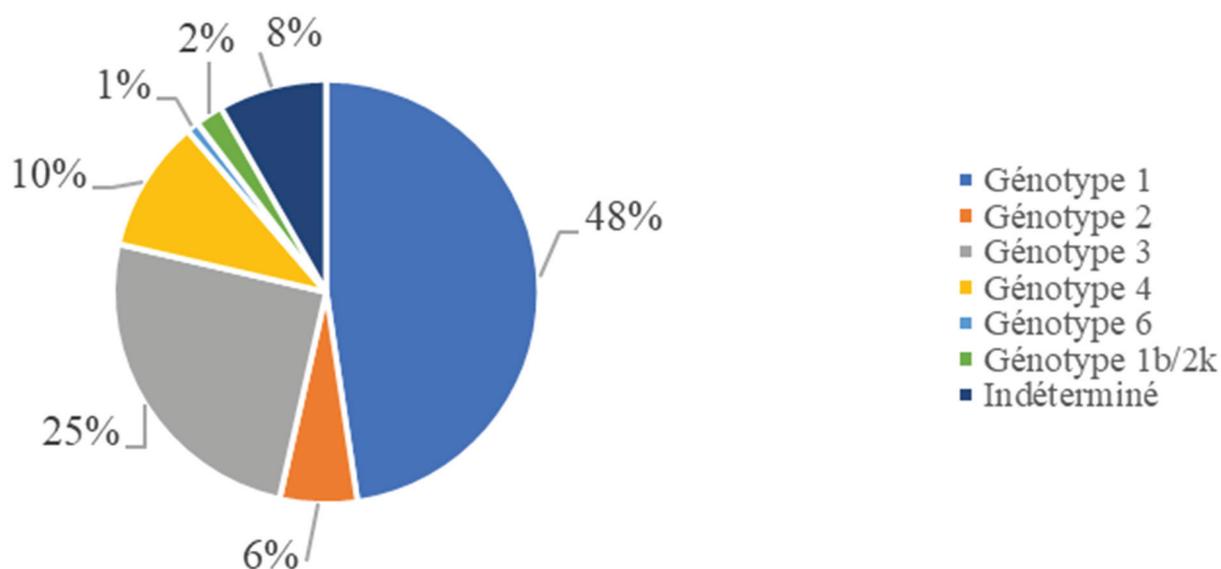


Figure 30. Bilan de la distribution des génotypes retrouvés.

La distribution des génotypes parmi les échantillons est représentée, calculée selon le nombre d'échantillons présentant ce génotype par rapport au nombre total d'échantillons. Dans cette étude, le génotype 1 fut le plus représenté, suivi du génotype 3, en accord avec les données épidémiologiques connues en France.

Parmi les génotypes 1, 2 sous-types furent les plus fréquents : le sous-type 1a chez 113 échantillons (55,67% des échantillons déterminés génotypes 1) et le sous-type 1b chez 80 patients (39,41% des échantillons déterminés génotypes 1). 5 autres sous-types ont été retrouvés en très faibles proportions pour les 10 autres échantillons (voir Figure 31).

Tous les génotypes 3 étaient des sous-types 3a, soient 107 patients (voir Figure 31).

Les génotypes 4 ont été partagés entre 8 sous-types. Les deux plus fréquents étaient les sous-types 4d et 4a, retrouvés respectivement chez 19 et 11 patients, représentant 44,19% et 25,58% des échantillons déterminés génotypes 4. Les 13 autres échantillons, soient 30,23% des génotypes 4, ont été classés en sous-types 4k, 4r, 4b, 4c, 4f et 4g (voir Figure 31).

Pour les génotypes 2, 7 sous-types furent retrouvés : le sous-type 2c (10 patients, soient 40% des échantillons déterminés génotypes 2), le sous-type 2a (5 patients, soient 20% des échantillons déterminés génotypes 2), le sous-type 2k (4 patients, soient 16% des échantillons déterminés génotypes 2), les sous-types 2b et 2q (2 patients de chaque sous-type, soient 8% des échantillons déterminés génotypes 2 pour chacun de ces 2 sous-types) et les sous-types 2i et 2m (1 patient de chaque sous-type, soient 4% des échantillons déterminés génotypes 2 pour chacun de ces 2 sous-types) (voir Figure 31).

Les échantillons déterminés comme génotype 6, le moins fréquent, ont été exclusivement classés en sous-types 6e (voir Figure 31).

Pour 36 échantillons, aucun sous-type n'a pu être déterminé.

Les données de charge virale selon les génotypes sont présentées dans le Tableau 14. Il en ressort une moyenne de 5,67 log UI/mL pour les génotypes 1, avec un minimum à 1,21 log UI/mL et un maximum à 7,23 log UI/mL ; une moyenne de 5,73 log UI/mL pour les génotypes 2, avec un minimum à 3,5 log UI/mL et un maximum à 7,10 log UI/mL ; une moyenne de 5,31 log UI/mL pour les génotypes 3, avec un minimum à 2,36 log UI/mL et un maximum à 7,7 log UI/mL ; une moyenne de 5,71 log UI/mL pour les génotypes 4, avec un minimum à 2,90 log UI/mL et un maximum à 7,30 UI/mL ; une moyenne de 6,45 log UI/mL pour les génotypes 6, avec un minimum à 6,08 log UI/mL et un maximum à 7,10 UI/mL ; une moyenne de 5,86 log UI/mL pour les génotypes recombinants 1b/2k, avec un minimum à 4,98 log UI/mL et un maximum à 6,99 log UI/mL. Les résultats obtenus lors de cette étude montrent ainsi des moyennes de charge virale globalement

similaires entre les géotypes, hormis pour le géotype 6 où les valeurs sont plus élevées ($p = 0,014$ selon test non paramétrique de Kruskal Wallis). Une telle corrélation entre une charge virale élevée et un virus infectant de géotype 6 avait en effet déjà été décrite (166). Cela pourrait révéler une capacité virale plus importante à la réplication et/ou à résister à la réponse immunitaire, mais cette observation peut être biaisée par le nombre faible d'échantillons présentant ce géotype ($n = 4$ dans cette étude).

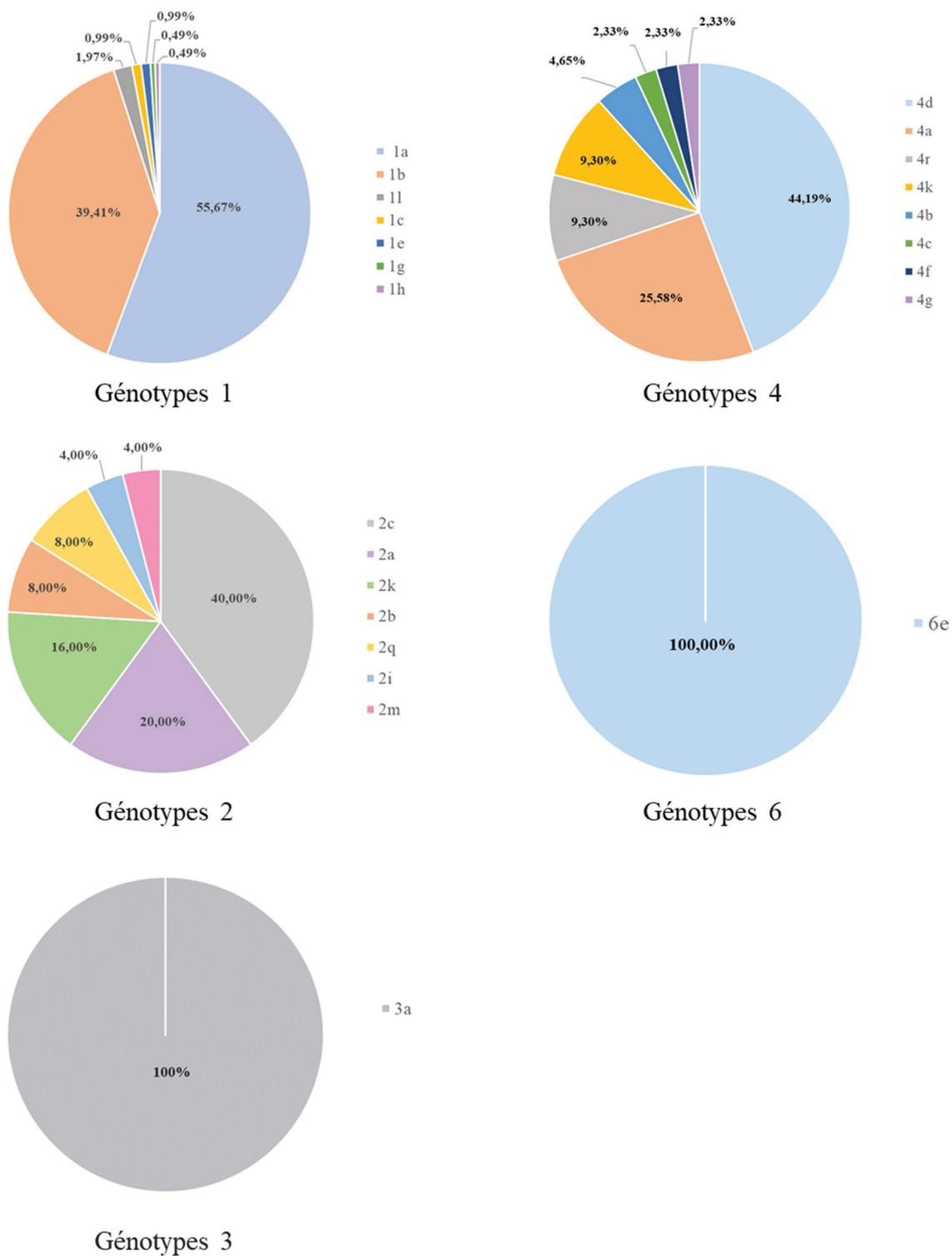


Figure 31. Distribution des génotypes et sous-types retrouvés.

Pour chaque génotype, les différents sous-types recensés sont représentés avec la valeur en pourcentage, calculée selon le nombre d'échantillons présentant ce sous-type par rapport au nombre total d'échantillons du génotype correspondant.

Tableau 14. Données de charges virales selon les génotypes.

Génotype	Charge virale moyenne (log UI/mL)	Médiane (log UI/mL)	Charge virale minimum (log UI/mL)	Charge virale maximum (log UI/mL)
1	5,67	5,80	1,21	7,23
2	5,73	6,16	3,50	7,10
3	5,31	5,52	2,36	7,70
4	5,71	5,85	2,90	7,30
6	6,45	6,31	6,08	7,10
1b/2k	5,86	5,70	4,98	6,99

3.6 Détermination des mutations de résistance

L'un des avantages *a priori* de cette méthode de séquençage est la détermination systématique des mutations éventuelles dans les trois régions d'intérêt NS3, NS5A et NS5B, tandis que le séquençage de Sanger ne déterminait, en première ligne, que celles présentes dans la région NS5A. Cependant, cela est réalisé uniquement pour les séquences déterminées comme appartenant aux génotypes 1a, 1b et 3 par le séquençage Vela Diagnostics. Sur le rapport de séquençage par NGS de ces génotypes sont ainsi rendus le nombre de mutations de résistance par région, leur nature (position, acide aminé de référence à cette position, acide aminé retrouvé) ainsi que leur fréquence. Ces mutations sont considérées comme impliquées dans la sensibilité aux traitements antiviraux. De ce fait, les polymorphismes ne sont pas indiqués dans le rapport. Leur détermination peut être réalisée *a posteriori*, en récupérant la séquence obtenue et par alignement à une séquence de référence de choix. Cela permet aussi, si besoin, la vérification du génotype et/ou du sous-type obtenu par NGS, notamment grâce à des bases de données telles que geno2pheno (165). Les données présentées ici excluent les contrôles qualités, rendant ainsi les résultats sur un ensemble de 416 échantillons.

3.6.1 Mutations dans NS3

Sur NS3, une grande partie des échantillons ne présentait pas de mutation (183 échantillons, soient 43,99%). Sur cette région, tous échantillons confondus, ont été retrouvées 135 mutations. Une seule mutation fut retrouvée dans cette région pour 75 échantillons (18,03%), tandis que d'autres échantillons présentaient deux (27 échantillons, soient 6,49%) ou trois mutations (2 échantillons, soient 0,48%). La détermination des mutations ne fut pas réalisée automatiquement dans 31,01% des cas (129 échantillons), pour cause de génotype différent des génotypes 1a, 1b ou 3 ; ou pour cause de non amplification de cette région.

Les mutations déterminées dans NS3 étaient retrouvées sur 10 positions d'acides aminés (positions 36, 54, 55, 56, 80, 122, 155, 168, 170 et 175). Parmi ces positions, les mutations les plus fréquentes étaient retrouvées en position 80 (25,19% des mutations sur NS3, soient 34 mutations) avec les mutations 80K, 80L et 80R. En position 122 furent retrouvées 22,22% des mutations sur NS3 (30 mutations), avec les mutations 122G, 122T et 122N. La mutation 56F fut fréquente, recensée dans 18,52% des mutations de NS3,

pour 25 échantillons. 17 mutations ont été retrouvées en position 170 (12,59% des mutations sur NS3) avec les substitutions 170V et 170T. La position 36 fut concernée par 7,41% des mutations de NS3 avec 10 mutations, 36M ou 36L. De même, deux mutations différentes furent retrouvées en position 55, 55A et 55I, dans 5,93% des mutations sur NS3, pour 8 échantillons. Les autres positions (54, 155, 168 et 175) ont été impliquées dans 8,15% des mutations (voir Tableau 15).

Parmi les échantillons de génotype 1 pour lesquels la détermination des mutations a été possible, 55,61% présentaient au moins une mutation ; tandis qu'aucun des échantillons de génotype 3 ne présentait de mutation dans cette région NS3.

3.6.2 Mutations dans NS5A

Sur la deuxième région d'intérêt NS5A, près de la moitié des échantillons ne présentait aucune mutation (206 échantillons, soient 49,52% des 416 échantillons). Ici, 139 mutations ont été détectées, avec une mutation dans 17,55% des cas (73 échantillons), deux mutations dans 4,81% des cas (20 échantillons), trois ou quatre mutations dans 1,44% des cas (3 échantillons présentant trois mutations et 3 échantillons en présentant quatre) et cinq mutations pour un unique cas (0,24%). Les mutations n'ont pas été établies dans 26,44% des cas, soient 110 échantillons, pour les mêmes raisons qu'évoquées précédemment (génotype différent des génotypes 1a, 1b ou 3 ; non amplification de cette région).

8 positions d'acides aminés de NS5A ont été touchées par des mutations, les positions 28, 30, 31, 54, 58, 62, 92 et 93. La position la plus fréquemment impactée par une mutation fut la position 93, pour 31,65% des mutations (44 mutations), avec les substitutions 93H, 93F, 93N et 93C. En position 30 ont été recensées 32 mutations, soit 23,02% des mutations de NS5A (30K, 30R, 30H, 30Q, 30L). 2 substitutions différentes ont été retrouvées en position 58 (58P et 58S), représentant 16 mutations, soient 11,51% de toutes les mutations de NS5A. Puis, les positions 28 et 31 ont été touchées par quatre et trois substitutions, respectivement (28V, 28M, 28T, 28A ; et 31M, 31F, 31V) représentant 10,79% des mutations pour chaque position. 7,91% des mutations ont été déterminées en position 62 avec quatre substitutions (62D, 62A, 62H, 62T). D'autres mutations ont été définies en positions 54 (54R, 54N, 54H) et 92 (92T), pour 4,32% des mutations de cette région (voir Tableau 15).

Dans cette région NS5A, 37,11% des échantillons de génotype 1 pour lesquels les mutations ont pu être détectées présentaient une mutation ou plus ; pour 27,62% des échantillons de génotype 3.

3.6.3 Mutations dans NS5B

En ce qui concerne les résultats obtenus sur NS5B, la majorité des échantillons, soient 268 sur 416 (64,42% des cas), ne présentait aucune mutation. Pour les autres échantillons, un total de 66 mutations a été dénombré. 18 échantillons présentaient une seule mutation (4,33% des cas), 9 en présentaient deux (2,16% des cas) et 10 autres en présentaient trois (2,40%). Pour 26,68% des échantillons, soient 111 échantillons, aucune mutation n'a pu être recherchée, pour les mêmes raisons qu'évoquées précédemment (génotype différent des génotypes 1a, 1b ou 3 ; non amplification de cette région).

Différentes mutations ont été détectées sur 8 positions différentes de cette région (159, 282, 316, 414, 445, 446, 499 et 556). Avec les plus grandes fréquences ont été retrouvées 26 mutations en position 556 (556G ou 556R) représentant 39,39% des mutations de cette région, 22 mutations en position 316 (316N ou 316Y) représentant 33,33% des mutations et 13 mutations en position 159 (159F) soient 19,70% des mutations. Cinq autres mutations ont été recensées, touchant respectivement les 5 autres positions (282, 414, 445, 446 et 499) avec une seule substitution (282T, 414L, 445F, 446Q, 499A) représentant chacune 1,52% des mutations de NS5B (voir Tableau 15).

Parmi tous les échantillons pour lesquels la détermination des mutations fut favorable dans cette dernière région d'intérêt, 18,65% des échantillons des génotypes 1 et 1,87% seulement de ceux des génotypes 3 ont montré au moins une mutation.

Tableau 15. Bilan des mutations retrouvées dans chaque région d'intérêt.

NS3		NS5A		NS5B	
Mutation	Nombre (%)	Mutation	Nombre (%)	Mutation	Nombre (%)
36L	8 (5,93)	28V	10 (7,19)	159F	13 (19,70)
36M	2 (1,48)	28M	2 (1,44)	282T	1 (1,52)
54S	4 (2,96)	28T	2 (1,44)	316N	21 (31,82)
55A	6 (4,44)	28A	1 (0,72)	316Y	1 (1,52)
55I	2 (1,48)	30K	11 (7,91)	414L	1 (1,52)
56F	25 (18,52)	30R	9 (6,47)	445F	1 (1,52)
80K	28 (20,74)	30Q	5 (3,60)	446Q	1 (1,52)
80L	4 (2,96)	30H	6 (4,32)	499A	1 (1,52)
80R	2 (1,48)	30L	1 (0,72)	556G	25 (37,88)
122G	12 (8,89)	31M	13 (9,35)	556R	1 (1,52)
122T	11 (8,15)	31F	1 (0,72)		
122N	7 (5,19)	31V	1 (0,72)		
155K	1 (0,74)	54R	1 (0,72)		
155G	1 (0,74)	54N	1 (0,72)		
168E	3 (2,22)	54H	1 (0,72)		
170V	16 (11,85)	58P	13 (9,35)		
170T	1 (0,74)	58S	3 (2,16)		
175L	2 (1,48)	62D	7 (5,04)		
		62A	2 (1,44)		
		62T	1 (0,72)		
		62H	1 (0,72)		
		92T	3 (2,16)		
		93H	39 (28,06)		
		93F	3 (2,16)		
		93N	1 (0,72)		
		93C	1 (0,72)		
Total	135	Total	139	Total	66

3.7 Cas d'échec de séquençage

Parmi la totalité des 424 échantillons, 103 d'entre eux présentaient des paramètres à la limite des seuils d'acceptabilité, voire classant le séquençage comme échec selon les critères de Vela Diagnostics. Différents paramètres ont été répertoriés comme à l'origine de ces échecs :

- Des paramètres de la réaction de séquençage dans son ensemble, et notamment le contrôle système (SC). Celui-ci est invalidé lorsque sa couverture d'amplification est trop faible, que le taux d'erreur est supérieur à 1%, ou que le nombre de reads correspondant au VHC est anormalement haut. En effet, le SC correspond au VMT, différent du VHC, raison pour laquelle il n'est pas attendu d'amplification importante du VHC pour cet échantillon (non-template control) ;
- Des paramètres propres au contrôle qualité de l'échantillon lui-même, pouvant être un débit de paires de bases sous le seuil au cours du séquençage ou une anomalie du contrôle d'extraction, qui sont possiblement des causes d'échec. Concernant le contrôle d'extraction, les échecs ont été dus à une couverture moyenne de la région NS5B ou de l'amplicon contrôle, inférieure au seuil.

Les divers échecs ont touché 16 des 29 réactions de séquençage, soit plus de 55% de toutes les réactions. Pour 6 d'entre elles, les échecs touchaient l'ensemble des échantillons, alors que 10 d'entre elles ne comptaient qu'une mineure partie des échantillons (un échantillon pour 5 réactions de séquençage, deux échantillons pour les 5 autres réactions de séquençage).

Parmi ces réactions où un ou deux échantillons présentaient des paramètres ne satisfaisant pas les seuils, six d'entre eux montraient un contrôle d'extraction dont les valeurs étaient aux limites des seuils, et deux pour lesquels les valeurs ont entraîné la classification en échec. Pour les sept autres échantillons, l'échec entraîné par les valeurs de contrôle d'extraction était associé à des valeurs du débit de paires de bases inférieures au seuil.

Pour les réactions où tous les échantillons étaient qualifiés comme échecs à cause de l'invalidation du contrôle système (soient 88 échantillons au cours de 6 réactions), des échantillons présentaient en outre des valeurs limites pour le contrôle d'extraction (4 échantillons, soient 4,55%) ou hors des seuils (10 échantillons, soient 11,36%). Pour 4 échantillons, le débit de paires de bases séquencées fut considéré comme hors limite,

classifié ainsi en tant qu'échec pour ce paramètre, en plus de l'échec attribué par les valeurs du contrôle de la réaction globale de séquençage (contrôle système). Cependant, l'attribution d'un échec, pour un échantillon et/ou une réaction entière de séquençage, n'empêche pas l'analyse des échantillons si l'amplification des régions d'intérêt est effective. Mais dans ce cas, les résultats devront être contrôlés voire confirmés par l'opérateur afin de s'assurer de leur fiabilité.

Pour ces échantillons voire réactions entières considérées en échec, le rapport indique possiblement le génotype et/ou le sous-type ainsi que le(s) mutation(s) dans les régions d'intérêt. Du fait des valeurs considérées en dehors ou aux limites des normes, les informations obtenues sont à prendre en compte avec précaution. Pour ces données, la séquence brute peut être utilisée pour déterminer et confirmer « manuellement » les informations afin de consolider leur fiabilité.

Hormis ces critères définis par Vela Diagnostics, il est possible de classer le séquençage en échec selon des données cliniquement plus informatives. Notamment, on peut considérer un séquençage en échec partiel si les régions virales NS3 et/ou NS5A ne sont pas amplifiées ; et en échec total lorsque qu'aucune séquence de la région de NS5B ne peut être rendue ou que la séquence obtenue n'a pas permis la détermination du génotype par l'analyse automatisée proposée par Vela Diagnostics, qui fournit l'information de génotype à partir de l'analyse de la séquence NS5B. A noter cependant que l'analyse manuelle des autres gènes peut fournir cette donnée. Dans cette étude, 57 échantillons ont été en échec partiel pour 36 échantillons en échec total, excluant l'échantillon à charge virale indétectable (voir Partie Expérimentale, **3.4 Capacité d'amplification des 3 gènes**).

Dans ce bilan, on compte ainsi 13,44% des échantillons, soient 57 échantillons, en échec partiel. 87,72% des échantillons montraient l'une des deux régions NS3 ou NS5A non amplifiée, et 12,28% des échantillons étaient en échec d'amplification pour ces deux régions. Parmi ces échantillons en échec partiel, 7,02% n'ont pu voir leur génotype déterminé, tandis que 33,33% étaient des génotypes 4, 22,81% des génotypes 2, 21,05% des génotypes 3 et 15,79% des génotypes 1.

Comme déjà indiqué auparavant (voir Partie Expérimentale, **3.4 Capacité d'amplification des 3 gènes**), la région NS5B n'a pu être séquencée pour 30 échantillons

(échantillon à charge indétectable inclus). Cependant, l'analyse Vela Diagnostics a pu rendre un génotype pour l'un de ces 30 échantillons. En outre, le génotype de 9 patients n'a pu être déterminé malgré l'amplification de NS5B. Ainsi, 38 échantillons incluant 2 contrôles qualité (dont l'un valide *in fine* puisque contrôle négatif) de cette étude peuvent être considérés en échec total, sans rendu de génotype par l'analyse de Vela Diagnostics. Cependant, le génotypage complémentaire via géno2phéno ainsi que le re-séquençage a permis de réduire le nombre d'échantillons non génotypés à 6, d'où une moyenne d'1 échantillon en échec pour 4 à 5 réactions de séquençage.

Plus globalement, le séquençage a été réalisé deux fois ou plus pour 34 échantillons, soit 8% de tous les échantillons, incluant 8 échantillons en échec total au premier séquençage, sans rendu de génotype possible même par analyse manuelle. Pour ces derniers échantillons, un séquençage supplémentaire est nécessaire pour le génotypage lorsque les régions d'intérêt n'ont pu être amplifiées ou lorsque l'analyse via géno2phéno n'a pas permis de rendre le génotype de façon fiable. Pour d'autres échantillons, cela est nécessaire pour pouvoir amplifier toutes les régions et ainsi rechercher des mutations pouvant modifier la sensibilité du virus au traitement (cas d'échecs au traitement), pour rechercher les cas de surinfection ; mais également tester la reproductibilité pour la validation de méthode, en profitant d'une place libre pour compléter une série de séquençage.

3.8 Cas clinique d'échec de traitement

Le séquençage du VHC infectant présente un enjeu important au cours du traitement d'un patient. En effet, il permet le génotypage mais aussi la détection de mutations indispensables à déceler pour le choix et/ou la modification du traitement, de façon à obtenir une efficacité optimale. Au cours de cette étude ont été inclus des échantillons d'un patient infecté par le VHC pour lequel le séquençage a présenté un rôle majeur pour sa prise en charge.

Le patient est un homme de 53 ans, présentant un ictère avec cholestase 2 mois après avoir subi une greffe de rein. Les sérologies Virus de l'Hépatite A (VHA), VHB et VHC ont été réalisées, et bien que négatives, ont été considérées insuffisantes pour exclure une hépatite virale du fait de l'immunodépression post-greffe. De ce fait, la charge virale a été réalisée et fut estimée à 7,7 log UI/mL pour le VHC, indiquant ainsi une hépatite C

aigüe chez ce patient pour lequel la sérologie au moment de la greffe était négative. Devant la gravité du tableau, avec une hépatite cholestatique fibrosante rapidement évolutive et une insuffisance rénale aigüe, un traitement pan-génotypique par Epclusa® (sofosbuvir + velpatasvir), qui était alors l'une des options thérapeutiques recommandées (113), fut initié immédiatement sans attendre le génotypage.

La réévaluation rétrospective des *sera* du patient et du donneur de rein ont permis d'exclure une éventuelle erreur lors de l'appel à la greffe. Le receveur du foie du même donneur n'a pas été contaminé. L'interrogatoire du patient n'a pas permis de retrouver de facteurs de risques.

Le génotypage par méthode de Sanger a rendu un génotype 3a. L'analyse phylogénétique de ce virus, avec les autres séquences de génotype 3a effectuées la même année, a montré une homologie proche de 100% avec un autre virus. Or, le patient porteur du virus avec cette forte homologie a lui aussi subi une transplantation rénale la veille du cas, faisant supposer une transmission nosocomiale.

Le suivi de la charge virale au cours du traitement a montré une réduction à 3,12 log UI/mL à 4 semaines après l'initiation du traitement, ce qui correspond à une cinétique de décroissance très faible par rapport à ce qui est habituellement attendu. Bien qu'aucune mutation n'ait été détectée par le séquençage de Sanger, il a été décidé une prolongation du traitement à une durée totale de 24 semaines, en raison de l'origine du virus, provenant d'un patient en double échec d'AAD. A 6 semaines après initiation du traitement, un rebond de charge virale fut observé, avec une ré-augmentation de celle-ci, estimée à 4,01 log UI/mL puis à 5,6 log UI/mL à 8 semaines. Cela indique ainsi que le patient ne répond pas au traitement, malgré une bonne observance, ce qui est rarissime.

Une recherche de mutation a alors été effectuée, montrant la mutation Y93H sur NS5A et S282T sur NS5B par séquençage de Sanger, expliquant l'échec au traitement choisi. Suite à cela, le traitement a été modifié par quadrithérapie associant du Solvadi® (sofosbuvir), du Maviret® (glécaprévir + pibrentasvir) et de la ribavirine (111) (soit une combinaison d'anti-NS5B, anti-NS3 et anti-NS5A) pendant 24 semaines. Suite à l'instauration de ce traitement, la charge virale a finalement été détectable mais non quantifiable à la 8^{ème} semaine, et maintenue non détectée à 3 mois après le traitement et dans le suivi au-delà, ce qui permet de conclure à la guérison du patient.

Les échantillons du patient ont été séquencés rétrospectivement par NGS lors de l'évaluation effectuée au cours de cette thèse. Il en ressort que sur l'échantillon pré-thérapeutique, la mutation Y93H était déjà détectable par cette technique à faible fréquence (9,36%), mais indétectable par séquençage de Sanger. Le séquençage de l'échantillon à l'échec montre à nouveau cette mutation, sélectionnée par le traitement à une fréquence de 99,76% et la mutation S282T à 99,6%.

Chez le patient source, la mutation Y93H était présente à 11,7% au moment de la transplantation, mutation seulement détectable *a posteriori* par NGS. A noter que ce patient avait sélectionné antérieurement la mutation S282T, mais que celle-ci était indétectable même par NGS à ce moment-là. Ce patient en échec de sofosbuvir-siméprévir puis sofosbuvir-daclatasvir a été retraité, lui aussi, avec succès, par une autre combinaison d'anti-NS5B, anti-NS5A et anti-NS3 (sofosbuvir + velpatasvir + voxilaprévir Vosevi®) en association à la ribavirine pendant 12 semaines.

Dans ce cas de figure, la sensibilité de la technique pour la détection de populations mineures avec des mutations de faibles fréquences montre tout son intérêt dans le choix initial du traitement et son adaptation.

4 Discussion

La détermination du génotype, du sous-type et la détection des mutations présentes dans les régions d'intérêt, puisque régions cibles des traitements antiviraux, sont des étapes indispensables chez les patients infectés par le VHC. En effet, la connaissance de ces données permet de choisir au mieux les molécules composant le traitement, sa durée pour éviter les échecs thérapeutiques, et d'évaluer le pronostic de l'infection. L'introduction récente des nouveaux antiviraux (voir Rappels bibliographiques, *1.6.2 Traitements basés sur l'utilisation de nouvelles molécules*) a donné lieu à une efficacité plus importante, mesurée par une réponse virologique plus importante qu'avec les traitements utilisés précédemment (voir Rappels bibliographiques, *1.6.1 Traitements basés sur l'utilisation d'interféron et de ribavirine*), mais aussi à une tolérance bien meilleure, menant à une durée de traitement réduite et à de meilleurs résultats ; à condition d'avoir effectivement choisi le traitement adéquat, conditionné par la connaissance du génotype/sous-type.

Au laboratoire de virologie de l'Hôpital de la Croix-Rousse, ces données fondamentales étaient obtenues par séquençage par la méthode de Sanger. Cependant, les récentes technologies de séquençage appartenant au NGS sont attirantes de par leurs capacités, leurs performances et leur automatisme (140). La présente étude a permis l'évaluation de la plateforme de NGS Ion Torrent de Vela Diagnostics dont les automates étaient déjà disponibles au laboratoire pour le séquençage du VIH, pour le séquençage du VHC, ayant pour but le génotypage et également la surveillance des mutations, possiblement à l'origine de pharmaco-résistance. En première ligne, les résultats d'échantillons obtenus par ces automates ont été comparés à ceux préalablement obtenus par séquençage de Sanger, méthode alors utilisée en routine au laboratoire. Puis, la nouvelle technologie a été implantée en routine pour le suivi de l'infection par le VHC. Cette étude a ainsi permis de dresser un bilan d'utilisation de cette technologie de NGS pour le VHC, et plus particulièrement du circuit de séquençage proposé par Vela Diagnostics, pendant l'année qui a suivi sa mise en place en routine.

Afin d'étudier la fiabilité de la technique de NGS, des données obtenues précédemment par séquençage de Sanger ont été confrontées au circuit de Vela Diagnostics au cours de 3 réactions de séquençage, représentant 43 échantillons de génotypes voire sous-types et mutations éventuelles déterminés par Sanger, ainsi que 2 mélanges prédéfinis.

A l'issue des réactions, il en ressort que les données entre les deux techniques sont concordantes à 98% pour le génotype obtenu et à 92% pour le sous-type (voir Partie Expérimentale, **3.2 Comparaison à Sanger**). Quant aux mutations déterminées par les deux techniques, le NGS a permis d'obtenir des données complémentaires à celles préalablement détectées par technique de Sanger, sur au moins l'une des régions, dans 47% des cas (données calculées sans prise en compte des contaminations ; voir Partie Expérimentale, **3.2 Comparaison à Sanger**). A l'inverse, des mutations définies par séquençage de Sanger n'ont pas été retrouvées par NGS, dans au moins l'une des régions d'intérêt, dans 9,3% des cas (voir Tableau 9). Cela peut être expliqué par la subjectivité de la lecture des séquences obtenues par Sanger, lorsque la qualité des séquences est mauvaise. Les chromatogrammes obtenus peuvent prêter à confusion et être à l'origine de difficultés à déterminer la séquence réelle. D'autre part, pour l'un des patients, l'absence de détection de certaines mutations par NGS est expliquée par la non-amplification de NS3, rendant impossible la détermination des mutations de la région alors qu'elle avait été obtenue par séquençage de Sanger.

Volontairement, deux mélanges d'échantillons de génotypes prédéterminés par Sanger à des charges virales connues ont été effectués afin d'évaluer la faculté de la technique de NGS à détecter d'éventuelles co-infections chez un patient. Le séquençage par NGS a effectivement permis la détection de l'un des deux mélanges, tandis que seul le virus présentant la charge virale la plus haute du deuxième mélange a été décelé.

Ces résultats mettent en évidence les performances de la technologie de NGS à détecter de manière fiable le génotype voire le sous-type d'échantillons infectés par 1 ou plusieurs virus présentant une charge virale suffisamment haute. En outre, les résultats révèlent, d'une part, la capacité de cette technologie de séquençage de nouvelle génération à déterminer des mutations non révélées par technique de Sanger ; et d'autre part, son habilité à distinguer la présence de plusieurs acides aminés différents à une même position, révélant les quasi-espèces et leur fréquence. Cela met en lumière la meilleure sensibilité de cette technique vis-à-vis de la technique de Sanger.

Après ces tests ayant permis la validation de la fiabilité de ce nouveau circuit de séquençage par Vela Diagnostics, le séquençage de routine du VHC fut ainsi réalisé par cette nouvelle plateforme, remplaçant entièrement les fonctions du séquençage de Sanger.

Au cours des 13 mois sur laquelle a été effectué le bilan, 26 réactions de séquençage ont été réalisées, soient 2 réactions par mois, comptant de 13 à 15 patients par réaction.

Au fil des réactions, le séquençage par Vela Diagnostics a permis le génotypage de 91,51% des échantillons, représentant en moyenne 1,24 échantillon en échec par réaction de séquençage, pour lequel aucun génotype n'est rendu. L'échec de détermination du génotype était, en majorité, causée par une absence d'amplification de la région d'intérêt NS5B. La distribution des génotypes et sous-types des échantillons effectivement génotypés a révélé que les génotypes 1, et notamment les sous-types a (fréquence de 28,97%) et b (fréquence de 20,40%) ; et 3a (fréquence de 27,20%) étaient les plus fréquents (voir Partie Expérimentale, **3.5 Détermination du génotype**, Figures 30 et 31). De ce fait, les données obtenues sont globalement cohérentes avec celles connues de l'épidémiologie globale du VHC en France. A noter que, dans les cas où l'automate de Vela Diagnostics n'a pas permis de déterminer le génotype, celui-ci a pu être défini manuellement pour les échantillons où une ou plusieurs des régions d'intérêt avai(en)t été amplifié(es), ou par re-séquençage. Finalement, moins de 2% des échantillons n'ont pu être génotypés (6 échantillons). Parmi ces échantillons, les charges virales étaient inférieures à 3 log UI/mL, seuil officiel de la technique, pour 2 d'entre eux (2,1 et 1,2 log UI/mL), entre 3 et 5 log UI/mL pour 2 autres (3,17 et 4,84 log UI/mL) et supérieures à 5 log UI/mL pour 2 autres (5,1 et 6,7 log UI/mL).

Plus globalement, la capacité d'amplification des différentes régions d'intérêt a été évaluée entre les deux trousse de Vela Diagnostics. Il en ressort que cette deuxième trousse ici testée présente une aptitude plus élevée à amplifier les génotypes 1 dits rares (non a non b), 3 et 4 par rapport à la première version (voir Tableau 16).

Tableau 16. Comparaison de la capacité d'amplification des régions d'intérêt des deux trousse Vela Diagnostics.

V1 : trousse version 1, V2 : trousse version 2. NA : non applicable.

Génotype	Nombre échantillons		Succès d'amplification (%)							
	V1	V2	NS3		NS5A		NS5B		V1	V2
			V1	V2	V1	V2	V1	V2		
1a	159	115	100	98,26	100	100	100	100	100	99,13
1b	118	81	85,11	92,59	100	100	100	100	100	100
1 (non a, non b)	17	10	37,5	80	87,25	100	100	100	100	100
2	43	25	19,61	72	77,85	68	100	100	100	100
3	183	108	36,54	89,81	70,56	98,15	100	100	100	100
4	65	44	61,6	72,73	53,2	75	100	100	100	100
5	32	0	74,44	NA	6,67	NA	100	100	100	NA
6	98	5	0	100	97,57	100	100	100	100	100
Inconnu	2	36	26,39	40,54	0	48,65	0	0	0	24,32

Parmi tous les échantillons, la moyenne de charge virale était de 5,55 log UI/mL. Les échantillons présentant un échec d'amplification sur au moins l'une des régions présentaient une moyenne de 5,22 log UI/mL, pour une moyenne de 4,81 log UI/mL pour 2 échecs d'amplification et une moyenne de 4,88 log UI/mL pour 3 échecs. Cependant, la charge virale la plus basse des échantillons pour lesquels l'amplification s'est réalisée correctement était de 1,2 log UI/mL (génotype 1a), tandis que celle annoncée par Vela Diagnostics était de 1000 UI/mL pour les génotypes 1a, 1b, 2, 3 ou 4 ; et de 2000 UI/mL pour les génotypes 5 et 6. Les échantillons déterminés génotypes 6 pendant cette étude présentaient tous une charge virale importante, supérieure à 6 log UI/mL et ainsi supérieure au seuil de détection.

Concernant la détermination des mutations par l'automate de Vela Diagnostics, effectuée seulement pour les génotypes 1a, 1b et 3a, le séquençage a, parmi ces génotypes, révéla entre 1 et 5 mutations dans NS3 pour 35,74% des échantillons, dans NS5A pour 32,68% des échantillons et dans NS5B pour 12,13% des échantillons. A noter ainsi que, du fait de génotypes différents de 1a, 1b ou 3a majoritairement ; ou de non amplification des régions, 31,01% des échantillons n'ont pas été considérés par l'analyse de Vela Diagnostics pour les mutations de NS3, 26,44% pour NS5A et 26,68% pour NS5B. En effet, 52,71%, 62,73% et 72,97% des échantillons écartés de l'analyse des mutations l'ont été pour cause de leur génotype, pour NS3, NS5A et NS5B respectivement.

La nature des mutations retrouvées est cohérente avec les positions répertoriées, liées aux résistances aux traitements (voir Tableau 4). Notamment, dans la région NS3, les positions 36, 54, 55, 56, 80, 122, 155, 168, 170 et 175 ont en effet été impactées par des mutations, à l'origine de résistances éventuelles aux inhibiteurs de protéase (voir Rappels bibliographiques, 1.6.2.1 Inhibiteurs de la protéase ou anti-NS3 et Tableau 4). Alors que la plupart de ces mutations sont importantes pour la sensibilité au traitement pour les génotypes 1a et 1b, très fréquents en France, d'autres mutations retrouvées dans notre étude telles que 80K, 122T, 168E ou 170V impactent aussi l'efficacité des traitements pour les virus de génotypes 6.

Sur NS5A, des positions là encore importantes pour l'efficacité des traitements antiviraux, et plus particulièrement anti-NS5A, ont été affectées par des mutations, dont les positions 28, 30, 31, 58, et 93 (voir Rappels bibliographiques, 1.6.2.2 Anti-NS5A et

Tableau 4). De la même manière que pour les mutations de NS3, la majorité des mutations de NS5A à ces positions peut entraver la sensibilité des traitements pour les génotypes 1a et 1b. Cependant, les mutations 28V/M/T, 30H/K, 31M/V, 58P/S et 93H, révélées pour certains échantillons, sont possiblement à l'origine de résistance aux anti-NS5A pour tous les autres génotypes, et notamment la mutation 31V, retrouvée pour 1 échantillon, qui peut causer une résistance à ces traitements quel que soit le génotype infectant.

Finalement, sur NS5B, des mutations en positions 159, 282, 316, 414, 445, 446 et 556, telles que déjà décrites comme liées à des résistances aux anti-NS5B, ont été retrouvées (voir Rappels bibliographiques, 1.6.2.3 Inhibiteurs de la polymérase ou anti-NS5B et Tableau 4). Ces mutations impactant là encore les génotypes 1a et 1b en majorité, les mutations 159F et 282T peuvent altérer, en outre, les réponses aux traitements pour les génotypes 2 et 3 ; ainsi que 2, 3, 4 et 5, respectivement.

Ces mutations ont été retrouvées à des fréquences allant de 3,23% (sur NS3) à 100% (sur NS3, NS5A), avec une moyenne globale de 77,92% (80,78% pour les mutations de NS3, 67,55% pour celles de NS5A et 85,43% pour celles de NS5B). Il en ressort que 11,54%, 29% et 13,51% des mutations présentaient une fréquence inférieure à 20%, ainsi indétectables en séquençage de Sanger, pour NS3, NS5A et NS5B, respectivement. Cela met en évidence la capacité des technologies de NGS, et plus particulièrement de l'automate Ion Torrent de Vela Diagnostics, à révéler ces mutations de faibles fréquences. Cela confirme sa capacité à détecter des variants minoritaires. Cette sensibilité est ainsi un avantage important de cette nouvelle génération de séquençage pour le suivi des patients traités, pour la surveillance de l'apparition des mutations et donc l'adaptation du traitement en lien avec le génotype.

Par conséquent, nous pouvons conclure qu'actuellement, le NGS constitue un outil satisfaisant pour l'étude de la diversité génétique des différentes populations éventuellement présentes dans un unique prélèvement sanguin.

CONCLUSIONS

THESE SOUTENUE PAR: CHANUT MARION

Le Virus de l'Hépatite C, responsable d'hépatites, est à l'origine de cirrhoses et carcinomes hépatocellulaires. Le virus est à l'origine de 1,75 millions de nouveaux cas chaque année, représentant au total 71 millions de personnes infectées chroniquement au niveau mondial, et entraînant 400 000 décès annuellement. Le Virus de l'Hépatite C est ainsi un problème de santé publique majeur, d'où l'importance que porte notamment l'OMS à ces infections, avec l'objectif d'élimination du virus d'ici 2030. Récemment, de nouveaux traitements ont été mis en place, et nécessitent, d'une part, la détermination du génotype viral du Virus de l'Hépatite C et d'autre part, la détection d'éventuelles mutations de résistance dans les régions cibles des traitements, facteurs cruciaux pour le choix des molécules composant le traitement optimal de l'infection.

Au laboratoire de virologie de l'Hôpital de la Croix-Rousse à Lyon, ces informations virales étaient obtenues par séquençage de Sanger en routine jusqu'en mai 2018. Cette thèse a eu pour objectif l'évaluation d'un automate de séquençage de Vela Diagnostics, afin de permettre son utilisation en routine pour le génotype et la détermination des mutations de résistance chez les patients infectés par le Virus de l'Hépatite C puis d'effectuer un bilan à 1 an de cette utilisation en routine.

Ce travail nous a permis de valider la fiabilité du séquençage de nouvelle génération (NGS) par la solution automatisée de Vela Diagnostics comparativement au séquençage de Sanger pour l'obtention du génotype et du sous-type ; et a ainsi permis de mettre en place le séquençage Vela Diagnostics pour le Virus de l'Hépatite C en routine au laboratoire de virologie de la Croix-Rousse. En effet, les résultats confirment que cette nouvelle génération de séquençage est tout autant efficace et présente davantage de performances que le séquençage de Sanger pour le génotypage, le sous-typage, et la détection de mutations de résistance.

Selon l'ensemble des résultats, les données obtenues par l'analyse automatisée fournie par la solution Vela Diagnostics permettent d'obtenir des résultats fiables pour les patients infectés présentant les génotypes les plus fréquents, à savoir 1a, 1b et 3a en France, d'où son utilité en pratique clinique. Pour les autres génotypes, les données brutes post-séquençage sont facilement récupérables pour analyser de façon plus manuelle la présence de mutations éventuelles.

D'autre part, le bilan a mis en lumière la capacité de l'automate de séquençage de chez Vela Diagnostics, utilisant la technologie Ion Torrent, pour ces informations, mais également pour l'identification de co-infections, pour la révélation de mutations à faible fréquence liées à l'existence des quasi-espèces virales et pour la recherche et l'identification de variants associés à une résistance chez des patients en échec, en rechute ou au cours d'une ré-infection.

Bien qu'aujourd'hui, il semblerait que les applications du NGS à la prise en charge des patients infectés par le VHC soient limitées du fait de l'arrivée de molécules pan-génomiques à forte barrière de résistance, cet outil reste néanmoins un recours intéressant pour un laboratoire spécialisé devant des cas compliqués. Cette thèse a permis de valider son utilisation en routine au laboratoire de virologie de la Croix-Rousse en lien avec ses performances et son automatisation.

Le Président de la thèse,
Nom : MORFIN SHERPA Florence

Signature :



Vu et permis d'imprimer, Lyon, le 4-5-2020
Vu, la Directrice de l'Institut des Sciences
Pharmaceutiques et Biologiques, Faculté de Pharmacie

Pour le Président de l'Université Claude Bernard Lyon 1,



Professeure C. VINCIGUERRA

Bibliographie

1. Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science*. 21 avr 1989;244(4902):359-62.
2. Autres données épidémiologiques / Hépatite C / Hépatites virales / Maladies infectieuses / Dossiers thématiques / Accueil [Internet]. [cité 1 juill 2019]. Disponible sur: <http://invs.santepubliquefrance.fr/Dossiers-thematiques/Maladies-infectieuses/Hepatites-virales/Hepatite-C/Autres-donnees-epidemiologiques>
3. Organisation Mondiale de la Santé OMS. Global health sector strategy on viral hepatitis 2016–2021 [Internet]. [cité 1 juill 2019]. Disponible sur: <https://apps.who.int/iris/bitstream/handle/10665/246177/WHO-HIV-2016.06-eng.pdf;jsessionid=E089E61B025ACE768954DBC8DD66D0C2?sequence=1>
4. Organisation Mondiale de la Santé OMS. Éliminer l'hépatite: la réponse de l'OMS [Internet]. [cité 1 juill 2019]. Disponible sur: <https://www.who.int/fr/news-room/detail/27-07-2017-eliminate-hepatitis-who>
5. Organisation Mondiale de la Santé OMS. Viral hepatitis [Internet]. 2009 [cité 1 juill 2019]. Disponible sur: http://apps.who.int/gb/ebwha/pdf_files/EB126/B126_15-en.pdf
6. Smith DB, Bukh J, Kuiken C, Muerhoff AS, Rice CM, Stapleton JT, et al. Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: updated criteria and genotype assignment web resource. *Hepatology* Baltim Md. janv 2014;59(1):318-27.
7. Tsukiyama-Kohara K, Kohara M. Hepatitis C Virus: Viral Quasispecies and Genotypes. *Int J Mol Sci*. 22 déc 2017;19(1).
8. Ramière C, Maynard M, Hartig-Lavie K, Bailly F, Virlogeux V, Zoulim F, et al. Real Life Experience in Treating Chronic Hepatitis C with Direct Acting Antivirals: Genotyping and Subtype Matter. *J Hepatol*. 2016;64(2):S751-2.
9. Pawlotsky J-M. Treatment failure and resistance with direct-acting antiviral drugs against hepatitis C virus. *Hepatology*. mai 2011;53(5):1742-51.
10. Institut National de Veille Sanitaire (InVS). VHC, autres données épidémiologiques [Internet]. 2014 [cité 1 juill 2019]. Disponible sur: <http://invs.santepubliquefrance.fr/Dossiers-thematiques/Maladies-infectieuses/Hepatites-virales/Hepatite-C/Autres-donnees-epidemiologiques>
11. Haute Autorité de santé. Prise en charge de l'hépatite C par les médicaments anti-viraux à action directe (AAD) [Internet]. 2014 [cité 1 juill 2019]. Disponible sur: https://www.has-sante.fr/portail/upload/docs/application/pdf/2014-07/hepatite_c_prise_en_charge_anti_viraux_aad.pdf

12. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*. déc 1977;74(12):5463-7.
13. Metzker ML. Emerging technologies in DNA sequencing. *Genome Res*. 1 déc 2005;15(12):1767-76.
14. Hutchison CA. DNA sequencing: bench to bedside and beyond. *Nucleic Acids Res*. 28 août 2007;35(18):6227-37.
15. Organisation Mondiale de la Santé OMS. Global hepatitis report, 2017 [Internet]. 2017 [cité 1 juill 2019]. Disponible sur: <http://apps.who.int/iris/bitstream/10665/255016/1/9789241565455-eng.pdf?ua=1>
16. Organisation Mondiale de la Santé OMS. Principaux repères sur l'hépatite C [Internet]. 2018 [cité 1 juill 2019]. Disponible sur: <https://www.who.int/fr/news-room/fact-sheets/detail/hepatitis-c>
17. Esteban JI, Sauleda S, Quer J. The changing epidemiology of hepatitis C virus infection in Europe. *J Hepatol*. janv 2008;48(1):148-62.
18. Perz JF, Armstrong GL, Farrington LA, Hutin YJF, Bell BP. The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. *J Hepatol*. oct 2006;45(4):529-38.
19. Mühlberger N, Schwarzer R, Lettmeier B, Sroczynski G, Zeuzem S, Siebert U. HCV-related burden of disease in Europe: a systematic assessment of incidence, prevalence, morbidity, and mortality. *BMC Public Health*. 22 janv 2009;9:34.
20. Marcellin P, Pequignot F, Delarocque-Astagneau E, Zarski J-P, Ganne N, Hillon P, et al. Mortality related to chronic hepatitis B and chronic hepatitis C in France: evidence for the role of HIV coinfection and alcohol consumption. *J Hepatol*. févr 2008;48(2):200-7.
21. Santé Publique France. Estimation de la prévalence de l'hépatite C en population générale, France métropolitaine [Internet]. 2016 [cité 1 juill 2019]. Disponible sur: http://invs.santepubliquefrance.fr/beh/2016/13-14/2016_13-14_1.html
22. Meffre C, Le Strat Y, Delarocque-Astagneau E, Dubois F, Antona D, Lemasson J-M, et al. Prevalence of hepatitis B and hepatitis C virus infections in France in 2004: social factors are important predictors after adjusting for known risk factors. *J Med Virol*. avr 2010;82(4):546-55.
23. Santé Publique France. Prévalence des hépatites chroniques C et B et antécédents de dépistage en population générale en 2016 : contribution à une nouvelle stratégie de dépistage , baromètre de santé publique France - BaroTest. [Internet]. 2016 [cité 2 déc 2019]. Disponible sur: http://beh.santepubliquefrance.fr/beh/2019/24-25/2019_24-25_1.html

24. Brouard C, Boussac-Zarebska M, Silvain C, Durand J, de Lédighen V, Pillonel J, et al. Rapid and large-scale implementation of HCV treatment advances in France, 2007-2015. *BMC Infect Dis.* 20 2017;17(1):784.
25. Caisse Nationale d'Assurance Maladie. Les antiviraux à action directe dans le traitement de l'Hépatite C [Internet]. [cité 2 juill 2019]. Disponible sur: https://www.ameli.fr/fileadmin/user_upload/documents/Points_de_repere_n_52_-_Les_antiviraux_a_action_directe_dans_le_traitement_de_l_hepatite_C.pdf
26. Agence de la biomédecine. Le rapport annuel médical et scientifique 2017 - La greffe hépatique [Internet]. 2017 [cité 18 sept 2019]. Disponible sur: <https://www.agence-biomedecine.fr/annexes/bilan2017/donnees/organes/05-foie/synthese.htm>
27. Catanese MT, Uryu K, Kopp M, Edwards TJ, Andrus L, Rice WJ, et al. Ultrastructural analysis of hepatitis C virus particles. *Proc Natl Acad Sci U S A.* 4 juin 2013;110(23):9505-10.
28. Piver E, Boyer A, Gaillard J, Bull A, Beaumont E, Roingeard P, et al. Intriguing structure of the HCV particle. *Gut.* août 2017;66(8):1351-2.
29. Thomssen R, Bonk S, Propfe C, Heermann KH, Köchel HG, Uy A. Association of hepatitis C virus in human sera with beta-lipoprotein. *Med Microbiol Immunol (Berl).* 1992;181(5):293-300.
30. Kapoor A, Simmonds P, Gerold G, Qaisar N, Jain K, Henriquez JA, et al. Characterization of a canine homolog of hepatitis C virus. *Proc Natl Acad Sci U S A.* 12 juill 2011;108(28):11608-13.
31. Fraser CS, Doudna JA. Structural and mechanistic insights into hepatitis C viral translation initiation. *Nat Rev Microbiol.* janv 2007;5(1):29-38.
32. Kazakov T, Yang F, Ramanathan HN, Kohlway A, Diamond MS, Lindenbach BD. Hepatitis C Virus RNA Replication Depends on Specific Cis- and Trans-Acting Activities of Viral Nonstructural Proteins. Tellinghuisen TL, éditeur. *PLOS Pathog.* 13 avr 2015;11(4):e1004817.
33. Moriishi K, Matsuura Y. Exploitation of lipid components by viral and host proteins for hepatitis C virus infection. *Front Microbiol.* 2012;3:54.
34. Messina JP, Humphreys I, Flaxman A, Brown A, Cooke GS, Pybus OG, et al. Global distribution and prevalence of hepatitis C virus genotypes. *Hepatology Baltim Md.* janv 2015;61(1):77-87.
35. Pawlotsky JM, Tsakiris L, Roudot-Thoraval F, Pellet C, Stuyver L, Duval J, et al. Relationship between hepatitis C virus genotypes and sources of infection in patients with chronic hepatitis C. *J Infect Dis.* juin 1995;171(6):1607-10.
36. Menecier D. Hépatite C [Internet]. [cité 3 juill 2019]. Disponible sur: <http://hepatoweb.com/hepatite-virale-C.php>

37. Gower E, Estes C, Blach S, Razavi-Shearer K, Razavi H. Global epidemiology and genotype distribution of the hepatitis C virus infection. *J Hepatol.* nov 2014;61(1 Suppl):S45-57.
38. Kalinina O, Norder H, Mukomolov S, Magnius LO. A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg. *J Virol.* avr 2002;76(8):4034-43.
39. Karchava M, Waldenström J, Parker M, Hallack R, Sharvadze L, Gatsrelia L, et al. High incidence of the hepatitis C virus recombinant 2k/1b in Georgia: Recommendations for testing and treatment. *Hepatol Res Off J Jpn Soc Hepatol.* déc 2015;45(13):1292-8.
40. Galli A, Bukh J. Comparative analysis of the molecular mechanisms of recombination in hepatitis C virus. *Trends Microbiol.* juin 2014;22(6):354-64.
41. Brouard C, Pioche C, Bronowicki JP, Couzigou P, Gorla O, Guyader D, et al. 390 trends in epidemiological and clinical characteristics of HCV-infected patients over six years: a french experience. *J Hepatol.* avr 2009;50:S148.
42. Farci P, Bukh J, Purcell RH. The quasispecies of hepatitis C virus and the host immune response. *Springer Semin Immunopathol.* 1997;19(1):5-26.
43. Cuevas JM, González-Candelas F, Moya A, Sanjuán R. Effect of ribavirin on the mutation rate and spectrum of hepatitis C virus in vivo. *J Virol.* juin 2009;83(11):5760-4.
44. Marascio N, Torti C, Liberto M, Focà A. Update on different aspects of HCV variability: focus on NS5B polymerase. *BMC Infect Dis.* 2014;14 Suppl 5:S1.
45. Asselah T, Martinot M, Boyer N, Marcellin P. Variabilité génétique du virus de l'hépatite C : implications cliniques -. *Gastroenterol Clin Biol.* 2000;24:175-84.
46. Arzumanyan A, Reis HMGPV, Feitelson MA. Pathogenic mechanisms in HBV- and HCV-associated hepatocellular carcinoma. *Nat Rev Cancer.* févr 2013;13(2):123-35.
47. Jin L, Peterson DL. Expression, isolation, and characterization of the hepatitis C virus ATPase/RNA helicase. *Arch Biochem Biophys.* 20 oct 1995;323(1):47-53.
48. Bartenschlager R, Ahlborn-Laake L, Mous J, Jacobsen H. Nonstructural protein 3 of the hepatitis C virus encodes a serine-type proteinase required for cleavage at the NS3/4 and NS4/5 junctions. *J Virol.* juill 1993;67(7):3835-44.
49. Ma Y, Yates J, Liang Y, Lemon SM, Yi M. NS3 helicase domains involved in infectious intracellular hepatitis C virus particle assembly. *J Virol.* août 2008;82(15):7624-39.
50. Lindenbach BD, Rice CM. The ins and outs of hepatitis C virus entry and assembly. *Nat Rev Microbiol.* oct 2013;11(10):688-700.

51. Lee H, Zhu T, Patel K, Zhang Y-Y, Truong L, Hevener KE, et al. High-Throughput Screening (HTS) and Hit Validation to Identify Small Molecule Inhibitors with Activity against NS3/4A proteases from Multiple Hepatitis C Virus Genotypes. Bartosch B, éditeur. PLoS ONE. 9 oct 2013;8(10):e75144.
52. Macdonald A, Harris M. Hepatitis C virus NS5A: tales of a promiscuous protein. J Gen Virol. sept 2004;85(Pt 9):2485-502.
53. Reyes GR. The nonstructural NS5A protein of hepatitis C virus: an expanding, multifunctional role in enhancing hepatitis C virus pathogenesis. J Biomed Sci. juin 2002;9(3):187-97.
54. Tellinghuisen TL, Foss KL, Treadaway JC, Rice CM. Identification of residues required for RNA replication in domains II and III of the hepatitis C virus NS5A protein. J Virol. févr 2008;82(3):1073-83.
55. Appel N, Zayas M, Miller S, Krijnse-Locker J, Schaller T, Friebe P, et al. Essential role of domain III of nonstructural protein 5A for hepatitis C virus infectious particle assembly. PLoS Pathog. 28 mars 2008;4(3):e1000035.
56. Masaki T, Suzuki R, Murakami K, Aizaki H, Ishii K, Murayama A, et al. Interaction of hepatitis C virus nonstructural protein 5A with core protein is critical for the production of infectious virus particles. J Virol. août 2008;82(16):7964-76.
57. Simmonds P. Genetic diversity and evolution of hepatitis C virus – 15 years on. J Gen Virol. 1 nov 2004;85(11):3173-88.
58. André P, Komurian-Pradel F, Deforges S, Perret M, Berland JL, Sodoyer M, et al. Characterization of low- and very-low-density hepatitis C virus RNA-containing particles. J Virol. juill 2002;76(14):6919-28.
59. Lindenbach BD, Meuleman P, Ploss A, Vanwolleghem T, Syder AJ, McKeating JA, et al. Cell culture-grown hepatitis C virus is infectious in vivo and can be recultured in vitro. Proc Natl Acad Sci U S A. 7 mars 2006;103(10):3805-9.
60. Nielsen SU, Bassendine MF, Burt AD, Martin C, Pumeechockchai W, Toms GL. Association between hepatitis C virus and very-low-density lipoprotein (VLDL)/LDL analyzed in iodixanol density gradients. J Virol. mars 2006;80(5):2418-28.
61. Diaz O, Delers F, Maynard M, Demignot S, Zoulim F, Chambaz J, et al. Preferential association of Hepatitis C virus with apolipoprotein B48-containing lipoproteins. J Gen Virol. oct 2006;87(Pt 10):2983-91.
62. Denolly S, Granier C, Fontaine N, Pozzetto B, Bourlet T, Guérin M, et al. A serum protein factor mediates maturation and apoB-association of HCV particles in the extracellular milieu. J Hepatol. avr 2019;70(4):626-38.
63. Lefèvre M, Felmlee DJ, Parnot M, Baumert TF, Schuster C. Syndecan 4 is involved in mediating HCV entry through interaction with lipoviral particle-associated apolipoprotein E. PloS One. 2014;9(4):e95550.

64. Shi Q, Jiang J, Luo G. Syndecan-1 serves as the major receptor for attachment of hepatitis C virus to the surfaces of hepatocytes. *J Virol.* juin 2013;87(12):6866-75.
65. Scarselli E, Ansuini H, Cerino R, Roccasecca RM, Acali S, Filocamo G, et al. The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *EMBO J.* 1 oct 2002;21(19):5017-25.
66. Dao Thi VL, Granier C, Zeisel MB, Guérin M, Mancip J, Granio O, et al. Characterization of hepatitis C virus particle subpopulations reveals multiple usage of the scavenger receptor BI for entry steps. *J Biol Chem.* 7 sept 2012;287(37):31242-57.
67. Yamamoto S, Fukuhara T, Ono C, Uemura K, Kawachi Y, Shiokawa M, et al. Lipoprotein Receptors Redundantly Participate in Entry of Hepatitis C Virus. *PLoS Pathog.* 2016;12(5):e1005610.
68. Hishiki T, Shimizu Y, Tobita R, Sugiyama K, Ogawa K, Funami K, et al. Infectivity of hepatitis C virus is influenced by association with apolipoprotein E isoforms. *J Virol.* nov 2010;84(22):12048-57.
69. Owen DM, Huang H, Ye J, Gale M. Apolipoprotein E on hepatitis C virion facilitates infection through interaction with low-density lipoprotein receptor. *Virology.* 10 nov 2009;394(1):99-108.
70. Ujino S, Nishitsuji H, Hishiki T, Sugiyama K, Takaku H, Shimotohno K. Hepatitis C virus utilizes VLDLR as a novel entry pathway. *Proc Natl Acad Sci U S A.* 5 janv 2016;113(1):188-93.
71. Farquhar MJ, Hu K, Harris HJ, Davis C, Brimacombe CL, Fletcher SJ, et al. Hepatitis C virus induces CD81 and claudin-1 endocytosis. *J Virol.* avr 2012;86(8):4305-16.
72. Benedicto I, Molina-Jiménez F, Bartosch B, Cosset F-L, Lavillette D, Prieto J, et al. The tight junction-associated protein occludin is required for a postbinding step in hepatitis C virus entry and infection. *J Virol.* août 2009;83(16):8012-20.
73. Sourisseau M, Michta ML, Zony C, Israelow B, Hopcraft SE, Narbus CM, et al. Temporal analysis of hepatitis C virus cell entry with occludin directed blocking antibodies. *PLoS Pathog.* mars 2013;9(3):e1003244.
74. Blanchard E, Belouzard S, Goueslain L, Wakita T, Dubuisson J, Wychowski C, et al. Hepatitis C virus entry depends on clathrin-mediated endocytosis. *J Virol.* juill 2006;80(14):6964-72.
75. Op De Beeck A, Voisset C, Bartosch B, Ciczora Y, Cocquerel L, Keck Z, et al. Characterization of functional hepatitis C virus envelope glycoproteins. *J Virol.* mars 2004;78(6):2994-3002.
76. Sharma NR, Mateu G, Dreux M, Grakoui A, Cosset F-L, Melikyan GB. Hepatitis C virus is primed by CD81 protein for low pH-dependent fusion. *J Biol Chem.* 2 sept 2011;286(35):30361-76.

77. Moradpour D, Penin F, Rice CM. Replication of hepatitis C virus. *Nat Rev Microbiol.* juin 2007;5(6):453-63.
78. Lindenbach BD, Rice CM. Unravelling hepatitis C virus replication from genome to function. *Nature.* 18 août 2005;436(7053):933-8.
79. Degenhardt L, Peacock A, Colledge S, Leung J, Grebely J, Vickerman P, et al. Global prevalence of injecting drug use and sociodemographic characteristics and prevalence of HIV, HBV, and HCV in people who inject drugs: a multistage systematic review. *Lancet Glob Health.* 2017;5(12):e1192-207.
80. Jauffret-Roustide M, Razafindratsima N, Vivier G, Oudaya L, Lefevre C, Desenclos J-C. Estimation de la séroprévalence du VIH et du VHC et profils des usagers de drogues en France, étude InVS-ANRS Coquelicot, 2004. :4.
81. Institut National de Veille Sanitaire (InVS). Estimation de la séroprévalence du VIH et de l'hépatite C chez les usagers de drogues en France. [Internet]. 2013 [cité 7 juill 2019]. Disponible sur: http://invs.santepubliquefrance.fr/beh/2013/39-40/2013_39-40_2.html
82. Garfein RS, Vlahov D, Galai N, Doherty MC, Nelson KE. Viral infections in short-term injection drug users: the prevalence of the hepatitis C, hepatitis B, human immunodeficiency, and human T-lymphotropic viruses. *Am J Public Health.* mai 1996;86(5):655-61.
83. Donahue JG, Muñoz A, Ness PM, Brown DE, Yawn DH, McAllister HA, et al. The declining risk of post-transfusion hepatitis C virus infection. *N Engl J Med.* 6 août 1992;327(6):369-73.
84. Neal KR, Jones DA, Killey D, James V. Risk factors for hepatitis C virus infection. A case-control study of blood donors in the Trent Region (UK). *Epidemiol Infect.* juin 1994;112(3):595-601.
85. Sun CA, Chen HC, Lu CF, You SL, Mau YC, Ho MS, et al. Transmission of hepatitis C virus in Taiwan: prevalence and risk factors based on a nationwide survey. *J Med Virol.* nov 1999;59(3):290-6.
86. Chan DPC, Sun H-Y, Wong HTH, Lee S-S, Hung C-C. Sexually acquired hepatitis C virus infection: a review. *Int J Infect Dis IJID Off Publ Int Soc Infect Dis.* août 2016;49:47-58.
87. Alter MJ, Coleman PJ, Alexander WJ, Kramer E, Miller JK, Mandel E, et al. Importance of heterosexual activity in the transmission of hepatitis B and non-A, non-B hepatitis. *JAMA.* 1 sept 1989;262(9):1201-5.
88. Yeung LT, King SM, Roberts EA. Mother-to-infant transmission of hepatitis C virus. *Hepatol Baltim Md.* août 2001;34(2):223-9.
89. Shepard CW, Finelli L, Alter MJ. Global epidemiology of hepatitis C virus infection. *Lancet Infect Dis.* sept 2005;5(9):558-67.

90. Polis CB, Shah SN, Johnson KE, Gupta A. Impact of maternal HIV coinfection on the vertical transmission of hepatitis C virus: a meta-analysis. *Clin Infect Dis Off Publ Infect Dis Soc Am.* 15 avr 2007;44(8):1123-31.
91. Heim MH. 25 years of interferon-based treatment of chronic hepatitis C: an epoch coming to an end. *Nat Rev Immunol.* 2013;13(7):535-42.
92. Farci P, Alter HJ, Wong D, Miller RH, Shih JW, Jett B, et al. A Long-Term Study of Hepatitis C Virus Replication in Non-A, Non-B Hepatitis. *N Engl J Med.* 11 juill 1991;325(2):98-104.
93. Santantonio T, Wiegand J, Gerlach JT. Acute hepatitis C: current status and remaining challenges. *J Hepatol.* oct 2008;49(4):625-33.
94. Thomson EC, Nastouli E, Main J, Karayiannis P, Eliahoo J, Muir D, et al. Delayed anti-HCV antibody response in HIV-positive men acutely infected with HCV. *AIDS Lond Engl.* 2 janv 2009;23(1):89-93.
95. Santé Publique France. Hépatite C ; dépistage, clinique, prise en charge et conseils aux patients [Internet]. 2007 [cité 8 juill 2019]. Disponible sur: <http://inpes.santepubliquefrance.fr/CFESBases/catalogue/pdf/1054.pdf>
96. Chan ST, Ou J-HJ. Hepatitis C Virus-Induced Autophagy and Host Innate Immune Response. *Viruses.* 12 2017;9(8).
97. Saito T, Owen DM, Jiang F, Marcotrigiano J, Gale M. Innate immunity induced by composition-dependent RIG-I recognition of hepatitis C virus RNA. *Nature.* 24 juill 2008;454(7203):523-7.
98. Thimme R, Bukh J, Spangenberg HC, Wieland S, Pemberton J, Steiger C, et al. Viral and immunological determinants of hepatitis C virus clearance, persistence, and disease. *Proc Natl Acad Sci U S A.* 26 nov 2002;99(24):15661-8.
99. Farci P, Shimoda A, Wong D, Cabezon T, De Gioannis D, Strazzer A, et al. Prevention of hepatitis C virus infection in chimpanzees by hyperimmune serum against the hypervariable region 1 of the envelope 2 protein. *Proc Natl Acad Sci U S A.* 24 déc 1996;93(26):15394-9.
100. Boyer N, Marcellin P. L'interféron pegylé : nouveau progrès dans le traitement de l'hépatite chronique virale C. Randomised trial of interferon alfa 2b plus ribavirin for 48 weeks or for 24 weeks versus interferon alfa 2b plus placebo for 48 weeks for treatment of chronic infection with hepatitis C virus. *sept 2000;24(8-9):767.*
101. Moradpour D, Müllhaupt B. Hépatite C: traitement actuel. *Rev Médicale Suisse.* 2015;11:902-6.
102. VIDAL : Base de données médicamenteuse pour les prescripteurs libéraux [Internet]. VIDAL. [cité 9 juill 2019]. Disponible sur: <https://www.vidal.fr/>

103. Crotty S, Cameron C, Andino R. Ribavirin's antiviral mechanism of action: lethal mutagenesis? *J Mol Med Berl Ger.* févr 2002;80(2):86-95.
104. Bartenschlager R, Lohmann V, Penin F. The molecular and structural basis of advanced antiviral therapy for hepatitis C virus infection. *Nat Rev Microbiol.* juill 2013;11(7):482-96.
105. Liang TJ, Ghany MG. Current and future therapies for hepatitis C virus infection. *N Engl J Med.* 16 mai 2013;368(20):1907-17.
106. Haute Autorité de santé. Rapport d'évaluation des antiviraux d'action directe dans le traitement de l'hépatite C [Internet]. 2017 [cité 9 juill 2019]. Disponible sur: https://www.has-sante.fr/upload/docs/application/pdf/2018-03/rapport_actualisation_strat_aad_2017_avis2_06122017_cteval318.pdf
107. Chtioui H, Buclin T, Moradpour D. Nouveaux traitements de l'hépatite C : aspects pharmacologiques et potentiel d'interaction [Internet]. *Revue Médicale Suisse.* [cité 11 juill 2019]. Disponible sur: <https://www.revmed.ch/RMS/2014/RMS-N-440/Nouveaux-traitements-de-l-hepatite-C-aspects-pharmacologiques-et-potentiel-d-interaction>
108. PubChem [Internet]. [cité 20 sept 2019]. Disponible sur: <https://pubchem.ncbi.nlm.nih.gov/>
109. Agence Nationale de Sécurité du Médicament et des Produits de Santé. Répertoire des Spécialités Pharmaceutiques [Internet]. [cité 3 déc 2019]. Disponible sur: <http://agence-prd.ansm.sante.fr/php/ecodex/index.php#result>
110. Ministère des Solidarités et de la Santé. Simplification de l'accès au traitement contre l'hépatite C chronique [Internet]. Ministère des Solidarités et de la Santé. 2019 [cité 20 sept 2019]. Disponible sur: <https://solidarites-sante.gouv.fr/actualites/presse/communiqués-de-presse/article/simplification-de-l-acces-au-traitement-contre-l-hepatite-c-chronique>
111. l'Association Française pour l'Etude du Foie. Recommandations AFEF 2018 [Internet]. 2018 [cité 9 juill 2019]. Disponible sur: <https://afef.asso.fr/wp-content/uploads/2018/06/VF-INTERACTIF-RECO-VHC-AFEF-v2103.pdf>
112. Haute Autorité de santé. Hépatite C : prise en charge simplifiée chez l'adulte [Internet]. 2019 [cité 20 juill 2019]. Disponible sur: https://www.has-sante.fr/upload/docs/application/pdf/2019-03/rapport_elaboration_hepatite_c.pdf
113. l'Association Française pour l'Etude du Foie. Recommandations AFEF 2017 [Internet]. 2017 [cité 9 juill 2019]. Disponible sur: <https://afef.asso.fr/wp-content/uploads/2018/06/RecommandationsAFEFMars2017.pdf>
114. Benítez-Gutiérrez L, Barreiro P, Labarga P, de Mendoza C, Fernandez-Montero JV, Arias A, et al. Prevention and management of treatment failure to new oral hepatitis C drugs. *Expert Opin Pharmacother.* juin 2016;17(9):1215-23.

115. Sarrazin C, Dvory-Sobol H, Svarovskaia ES, Doehle BP, Pang PS, Chuang S-M, et al. Prevalence of Resistance-Associated Substitutions in HCV NS5A, NS5B, or NS3 and Outcomes of Treatment With Ledipasvir and Sofosbuvir. *Gastroenterology*. 2016;151(3):501-512.e1.
116. Krishnan P, Tripathi R, Schnell G, Reisch T, Beyer J, Irvin M, et al. Resistance analysis of baseline and treatment-emergent variants in hepatitis C virus genotype 1 in the AVIATOR study with paritaprevir-ritonavir, ombitasvir, and dasabuvir. *Antimicrob Agents Chemother*. sept 2015;59(9):5445-54.
117. Welzel TM, Bhardwaj N, Hedskog C, Chodavarapu K, Camus G, McNally J, et al. Global epidemiology of HCV subtypes and resistance-associated substitutions evaluated by sequencing-based subtype analyses. *J Hepatol*. 2017;67(2):224-36.
118. Zeuzem S, Mizokami M, Pianko S, Mangia A, Han K-H, Martin R, et al. NS5A resistance-associated substitutions in patients with genotype 1 hepatitis C virus: Prevalence and effect on treatment outcome. *J Hepatol*. 2017;66(5):910-8.
119. Pawlotsky J-M. Hepatitis C Virus Resistance to Direct-Acting Antiviral Drugs in Interferon-Free Regimens. *Gastroenterology*. 2016;151(1):70-86.
120. Young J. Sanger sequencing - a hands-on simulation [Internet]. 2017 [cité 16 juill 2019]. Disponible sur: <http://www.genetics-gsa.org/education/GSAPREP.2018.003.shtml>
121. Kircher M, Kelso J. High-throughput DNA sequencing--concepts and limitations. *BioEssays News Rev Mol Cell Dev Biol*. juin 2010;32(6):524-36.
122. Sanger [Internet]. Biomnigene. [cité 15 juill 2019]. Disponible sur: <https://www.biomnigene.fr/fr/nos-solutions/sanger.html>
123. Applied Biological Materials ABM. An Intro to NGS [Internet]. [cité 15 juill 2019]. Disponible sur: https://old.abmgood.com/marketing/knowledge_base/next_generation_sequencing_introduction.php
124. Virtanen E, Mannonen L, Lappalainen M, Auvinen E. Genotyping of hepatitis C virus by nucleotide sequencing: A robust method for a diagnostic laboratory. *MethodsX*. 2018;5:414-8.
125. Beloukas A, King S, Childs K, Papadimitropoulos A, Hopkins M, Atkins M, et al. Detection of the NS3 Q80K polymorphism by Sanger and deep sequencing in hepatitis C virus genotype 1a strains in the UK. *Clin Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis*. nov 2015;21(11):1033-9.
126. Lisboa-Neto G, Noble CF, Pinho JRR, Malta FM, Gomes-Gouvêa MS, Alvarado-Mora MV, et al. Resistance mutations are rare among protease inhibitor treatment-naïve hepatitis C genotype-1 patients with or without HIV coinfection. *Antivir Ther*. 2015;20(3):281-7.

127. Paolucci S, Premoli M, Novati S, Gulminetti R, Maserati R, Barbarini G, et al. Baseline and Breakthrough Resistance Mutations in HCV Patients Failing DAAs. *Sci Rep.* 22 2017;7(1):16017.
128. Besse B, Coste-Burel M, Bourgeois N, Feray C, Imbert-Marcille B-M, André-Garnier E. Genotyping and resistance profile of hepatitis C (HCV) genotypes 1-6 by sequencing the NS3 protease region using a single optimized sensitive method. *J Virol Methods.* oct 2012;185(1):94-100.
129. Sandres-Sauné K, Deny P, Pasquier C, Thibaut V, Duverlie G, Izopet J. Determining hepatitis C genotype by analyzing the sequence of the NS5b region. *J Virol Methods.* mai 2003;109(2):187-93.
130. Oxford Nanopore Technologies. MinION [Internet]. [cité 24 sept 2019]. Disponible sur: <http://nanoporetech.com/products/minion>
131. Université d'Angers. Génomique : les méthodes de séquençage d'acides nucléiques et l'acquisition des données [Internet]. 2017 [cité 17 juill 2019]. Disponible sur: <http://biochimej.univ-angers.fr/Page2/COURS/9ModulGenFoncVeg/5MethEtudGenFonc/1MethodeSEQUENGAGE/1SEQUENGAGE.htm>
132. Trémeaux P, Caporossi A, Thélu M-A, Blum M, Leroy V, Morand P, et al. Hepatitis C virus whole genome sequencing: Current methods/issues and future challenges. *Crit Rev Clin Lab Sci.* 2 sept 2016;53(5):341-51.
133. Trémeaux P, Caporossi A, Ramière C, Santoni E, Tarbouriech N, Thélu M-A, et al. Amplification and pyrosequencing of near-full-length hepatitis C virus for typing and monitoring antiviral resistant strains. *Clin Microbiol Infect.* mai 2016;22(5):460.e1-460.e10.
134. Shendure J, Balasubramanian S, Church GM, Gilbert W, Rogers J, Schloss JA, et al. DNA sequencing at 40: past, present and future. *Nature.* oct 2017;550(7676):345-53.
135. ATD Bio. Next generation sequencing [Internet]. [cité 22 juill 2019]. Disponible sur: <https://www.atdbio.com/content/58/Next-generation-sequencing#figure-sequencing-by-ligation>
136. Fierce Biotech. Roche ferme les sciences de la vie 454 car il réduit le séquençage des gènes | FierceBiotech [Internet]. 2013 [cité 23 sept 2019]. Disponible sur: <https://www.fiercebiotech.com/medical-devices/roche-to-close-454-life-sciences-as-it-reduces-gene-sequencing-focus>
137. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet.* 17 2016;17(6):333-51.
138. History of Illumina Sequencing and Solexa Technology [Internet]. [cité 23 sept 2019]. Disponible sur: <https://www.illumina.com/science/technology/next-generation-sequencing/illumina-sequencing-history.html>

139. Mardis ER. Next-generation sequencing platforms. *Annu Rev Anal Chem Palo Alto Calif.* 2013;6:287-303.
140. Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet.* janv 2010;11(1):31-46.
141. Mardis ER. Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet.* 2008;9:387-402.
142. Delafontaine J. L'analyse de données RNA-seq: mode d'emploi [Internet]. *bioinfo-fr.net.* 2013 [cité 25 juill 2019]. Disponible sur: <https://bioinfo-fr.net/lanalyse-de-donnees-rna-seq-mode-demploi>
143. Lacoste C, Fabre A, Pécheux C, Lévy N, Krahn M, Malzac P, et al. [Next-generation DNA sequencing in clinical diagnostics]. *Arch Pediatr Organe Off Soc Francaise Pediatr.* avr 2017;24(4):373-83.
144. Lecuit M, Eloit M. The human virome: new tools and concepts. *Trends Microbiol.* oct 2013;21(10):510-5.
145. Delwart E. A roadmap to the human virome. *PLoS Pathog.* févr 2013;9(2):e1003146.
146. Newman RM, Kuntzen T, Weiner B, Berical A, Charlebois P, Kuiken C, et al. Whole genome pyrosequencing of rare hepatitis C virus genotypes enhances subtype classification and identification of naturally occurring drug resistance variants. *J Infect Dis.* juill 2013;208(1):17-31.
147. Ninomiya M, Ueno Y, Funayama R, Nagashima T, Nishida Y, Kondo Y, et al. Use of illumina deep sequencing technology to differentiate hepatitis C virus variants. *J Clin Microbiol.* mars 2012;50(3):857-66.
148. Gregori J, Salicrú M, Domingo E, Sanchez A, Esteban JI, Rodríguez-Frías F, et al. Inference with viral quasispecies diversity indices: clonal and NGS approaches. *Bioinforma Oxf Engl.* 15 2014;30(8):1104-11.
149. Hou W, Bonkovsky HL. Non-coding RNAs in hepatitis C-induced hepatocellular carcinoma: dysregulation and implications for early detection, diagnosis and therapy. *World J Gastroenterol.* 28 nov 2013;19(44):7836-45.
150. Bartolini B, Giombini E, Zaccaro P, Selleri M, Rozera G, Abbate I, et al. Extent of HCV NS3 protease variability and resistance-associated mutations assessed by next generation sequencing in HCV monoinfected and HIV/HCV coinfecting patients. *Virus Res.* 6 nov 2013;177(2):205-8.
151. Nasu A, Marusawa H, Ueda Y, Nishijima N, Takahashi K, Osaki Y, et al. Genetic heterogeneity of hepatitis C virus in association with antiviral therapy determined by ultra-deep sequencing. *PloS One.* 2011;6(9):e24907.

152. Franco S, Casadellà M, Noguera-Julian M, Clotet B, Tural C, Paredes R, et al. No detection of the NS5B S282T mutation in treatment-naïve genotype 1 HCV/HIV-1 coinfecting patients using deep sequencing. *J Clin Virol Off Publ Pan Am Soc Clin Virol.* déc 2013;58(4):726-9.
153. Verbinnen T, Van Marck H, Vandebroucke I, Vijgen L, Claes M, Lin T-I, et al. Tracking the evolution of multiple in vitro hepatitis C virus replicon variants under protease inhibitor selection pressure by 454 deep sequencing. *J Virol.* nov 2010;84(21):11124-33.
154. Bull RA, Luciani F, McElroy K, Gaudieri S, Pham ST, Chopra A, et al. Sequential bottlenecks drive viral evolution in early acute hepatitis C virus infection. *PLoS Pathog.* sept 2011;7(9):e1002243.
155. Quiñones-Mateu ME, Avila S, Reyes-Teran G, Martinez MA. Deep sequencing: becoming a critical tool in clinical virology. *J Clin Virol Off Publ Pan Am Soc Clin Virol.* sept 2014;61(1):9-19.
156. Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, et al. Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol.* mai 2012;30(5):434-9.
157. Liu L, Li Y, Li S, Hu N, He Y, Pong R, et al. Comparison of Next-Generation Sequencing Systems. *J Biomed Biotechnol.* 2012;2012:1-11.
158. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, et al. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics.* 24 juill 2012;13:341.
159. van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. Ten years of next-generation sequencing technology. *Trends Genet TIG.* sept 2014;30(9):418-26.
160. Pareek CS, Smoczynski R, Tretyn A. Sequencing technologies and genome sequencing. *J Appl Genet.* nov 2011;52(4):413-35.
161. Biorigami. Plateformes de séquençage [Internet]. 2013 [cité 30 juill 2019]. Disponible sur: <http://www.biorigami.com/?tag=plateformes-sequencage>
162. Grebely J, Pham ST, Matthews GV, Petoumenos K, Bull RA, Yeung B, et al. Hepatitis C virus reinfection and superinfection among treated and untreated participants with recent infection. *Hepatology.* avr 2012;55(4):1058-69.
163. National Center for Biotechnology [Internet]. [cité 22 août 2019]. Disponible sur: <https://www.ncbi.nlm.nih.gov/>
164. Los Alamos National Lab: National Security Science [Internet]. [cité 22 août 2019]. Disponible sur: <https://www.lanl.gov/>
165. Geno2pheno HCV. [cité 22 août 2019]. Disponible sur: <https://hcv.geno2pheno.org/>

166. Rong X, Lu L, Wang J, Xiong H, Huang J, Chen J, et al. Correlation of viral loads with HCV genotypes: higher levels of virus were revealed among blood donors infected with 6a strains. *PloS One*. 2012;7(12):e52467.

Annexe 1

Quality Control Report

Patient Information:

Name of Patient:		Date of Collection:	
Date of Birth:		Date of Receipt at Lab:	
Age:		Date of Results:	
Gender:		Specimen ID:	019057445001-190409
Patient ID:		Specimen Types:	
Name/ID of Hospital:		Tumor Cellularity:	
Physician(s):		DNA Quality/Yield:	
Assay Name:	16_Sentosa_SQ_HCV_Genotyping_Assay_v2		
Patient History:			

Assay Information:

The *Sentosa*[®] SQ HCV Genotyping Assay is a next generation sequencing-based test intended for viral genotyping from human plasma or serum. This test is intended to be used on the *Sentosa*[®] SX101 with the *Sentosa*[®] SX Virus Total Nucleic Acid Plus II Kit in conjunction with the *Sentosa*[®] ST401 and *Sentosa*[®] SQ301 instruments.

Run Quality Controls:

QC Metric	QC Status	Value	QC Range
Run Throughput (bp)	PASSED	500,300,010	≥ 50,000,000 (≥ 210,000*)
Loading (%)	PASSED	75.83	≥ 15.00
Key Signal	PASSED	72	≥ 30
Invalid Reads / Bases	PASSED		
- Invalid Reads (%)		1.09	≤ 10.00
- Invalid Bases (%)		0.56	≤ 10.00
System Control	PASSED		
- Assembly Quality		99.83%	Completeness ≥ 95.00%
		0.00%	Error rate ≤ 1.00%
		24,705	Median Cov. ≥ 200
- No Template Control		28	HCV Reads ≤ MAX [50, 0.001 * Control Amplicon Reads (82047)]

Sample Quality Controls:

QC Metric	QC Status	Value	QC Range
Sample Throughput (bp)	PASSED	29,984,723	≥ 210,000
Extraction Control	PASSED		
- NS5B 3' Median Coverage		8347	Median Cov. ≥ 200 (50*)
- Control Amplicon Median Coverage		760	≥ 200

Databases and Tools Version:

Reference	Veladx Reference - HCV_v2_ref
MIRA	Mira 4.0.2
TMAP	TMAP 4.0.5
SAM Tools	SAM Tools 0.1.19-44428cd
<i>Sentosa</i>[®] SQ Reporter	SQ Reporter 1.3

Sequencing Information:

Run Info

Run ID 190409_HCV_ABS_ZGQZN
 Run Date April 12, 2019, 01:51 AM
 Key Signal 72
 Loading% 75.83
 Chip Type 318DX
 Machine SQSuite/sn274680013

Bases

Total Valid Bases 500,300,010
 - Mapped Bases 483,072,471 (96.56%)
 - Unmapped Bases 17,227,539 (3.44%)

Reads

Final Library 3,696,955
 Invalid Reads 40,362 (1.09%)
 Invalid Sample Reads 0 (0.0%)
 Total Valid Reads 3,656,593
 - Mapped Reads 3,599,025 (98.43%)
 - Unmapped Reads 57,568 (1.57%)

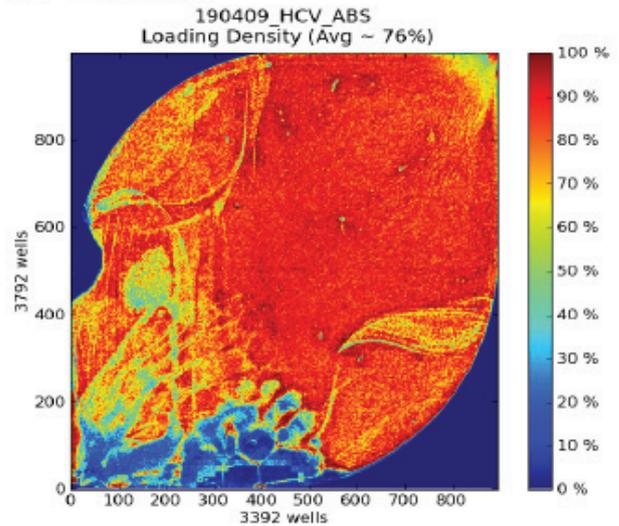
Barcode Info

Name SentosaBC-8

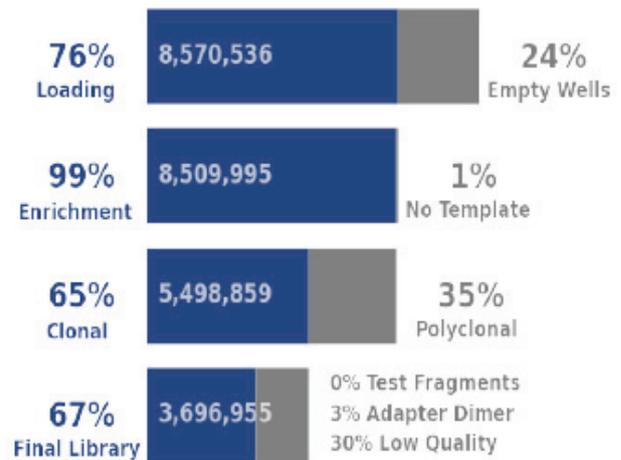
Total Valid Bases 29,984,723
 Mapped Bases 28,885,308 (96.33%)
 Unmapped Bases 1,099,415 (3.67%)

Total Valid Reads 210,772
 Mapped Reads 207,586 (98.49%)
 - Specific Reads 201,935 (95.81%)
 - Non-Specific Reads 1,188 (0.56%)
 - Discarded Reads 4,463 (2.12%)
 Unmapped Reads 3,186 (1.51%)
 Mean Read Length 142
 GC% 58.13%

ISP Quality:



ISP Density



ISP Summary

Run traceability record:

Sample Details	Sample ID Sample Name	019057445001-190409 NA
Assay Details	Name Application Panel Reference Library	HCV-reanalyse-120419 AmpliSeq 16_Sentosa_SQ_HCV_Genotyping_Assay_v2 hcv_20140514
Library Preparation	Kit Name Kit Barcode Part Number Expiration Date Lot Number	16 Sentosa SQ HCV Genotyping Assay v2 91000145100000_0010395171910182100000301 000145 2019-10-18 0000_0010395
Template Preparation	Start Date Completion Date Operator Kit Name Kit Barcode Part Number Expiration Date Lot Number	2019-04-10 2019-04-10 NGSLYON ngslyon Sentosa ST Template Kit 9100007310000000008994172005302100148301 000073 2020-05-30 000000008994
ST401 Prep	Templating Kit Control	Sentosa ST Template Kit Ion Dx Templating Control
SQ301 Sequencing	Sequencing Kit Chip Flow	Sentosa SQ Sequencing Kit Sentosa SQ 318 Chip Kit 500
Sequencing	Start Date Completion Date Operator Kit Name Kit Barcode Part Number Expiration Date Lot Number Chip Chip Manufacture Code (Bottom) Chip Barcode (Top) Chip Manufacturing Date Chip Serial Number Chip Manufacture Number	2019-04-12 2019-04-12 ngslyon Sentosa SQ Sequencing Kit 9100007410000000008988172005302100057301 000074 2020-05-30 000000008988 Sentosa SQ 318 Chip Kit 10p3234122201701 21ax005776791318d 2017-01 ax0057767 318d
Analysis	Date Operator	2019-04-12 Auto
Instrument Summary	SX101 Instrument serial number ST401 Instrument serial Number SQ301 Instrument serial number	5075DH904081 00190 sn274680013

Run QC Evaluation Metrics

Metric Name	Value	Reference Range	QC Status
Key Signal	72	≥ 30	Passed
Percent Loading	75.8	≥ 15	Passed

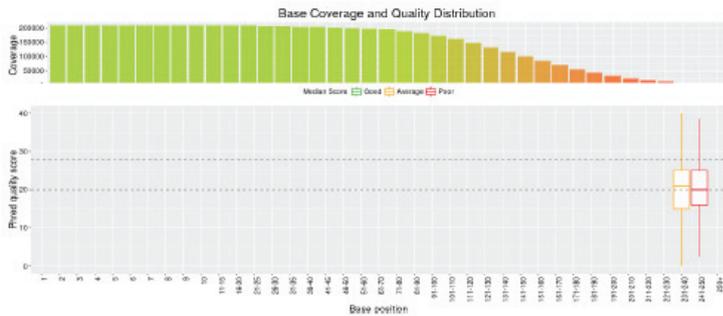
Sample QC Evaluation Metrics

Metric Name	Value	Reference Range	QC Status
Mean AQ20 Read Length (bp)	48	≥ 0	Passed
Mean Read Length (bp)	142	≥ 65	Passed
Percent Reads	1.9	≥ 0	Passed

Control QC Sequence Metrics

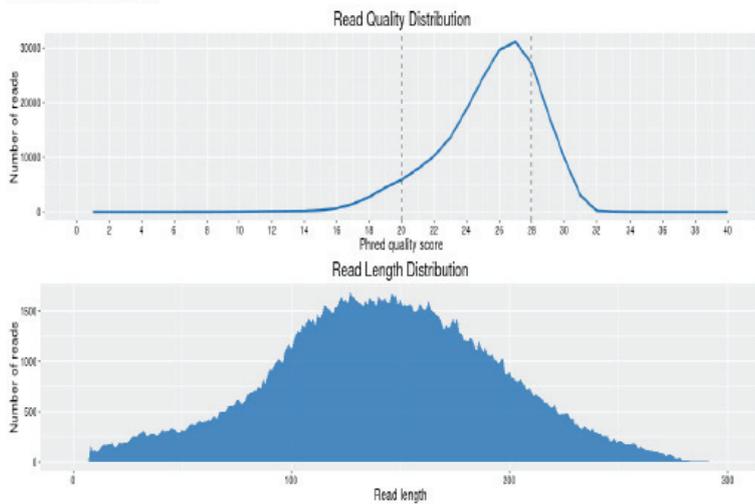
Metric Name	Value	Reference Range	QC Status
Ion Dx CF-1			
Base Call Accuracy	99.4	≥ 97.5	Passed
Key Signal	68	≥ 65	Passed
Mean AQ20 Read Length (bp)	135	≥ 122	Passed
Percent Reads	0.04	≥ 0.01	Passed

Base QC:



The **base quality distribution boxplot** for this barcode (SentosaBC-8). The upper section shows the base coverage across all reads in this sample, while the lower section provides information on the median, lower and upper quantiles of the base quality. In the plot, there are two horizontal lines at position 20 and 28. They demarcate the regions where the score can be considered of **poor quality** ($Q \leq 20$), **average quality** ($20 < Q \leq 28$), and **good quality** ($Q > 28$). $Q > 28$ means that there is a less than 0.15% chance the base is wrong.

Read QC:

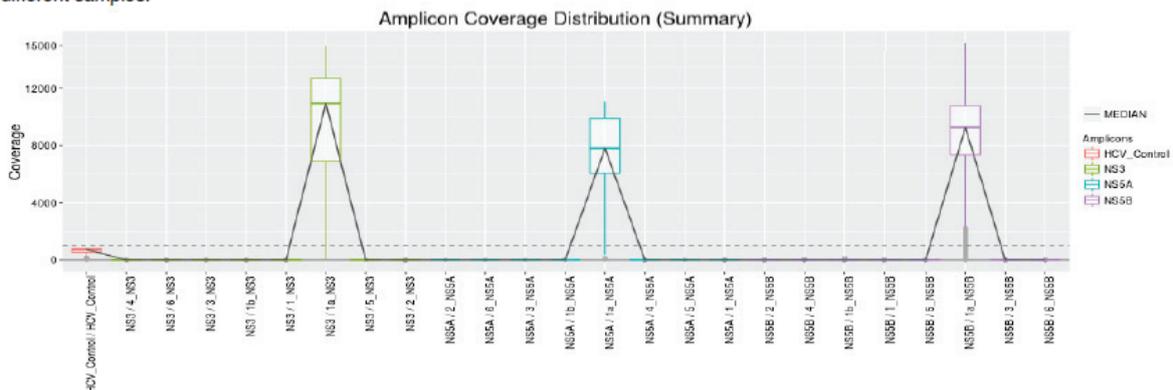


The **read quality plot** shows the number of reads for a given quality score. The read scores are determined by calculating the mean Phred quality scores of their constituent bases.

The **read length plot** shows the number of reads with a certain length, i.e. number of bases. By looking at the peaks on the plot, it is possible to tell what's the length of the majority of the reads.

Amplicon QC:

The **amplicon coverage distribution plot** shows the view of the distribution, showing median, lower and upper quantiles, and outliers (grey dots), if present. The black line connects the all medians together to provide visual help in identifying patterns when comparing plots from different samples.



Phred Quality Score (definition):

The quality score of a base, also known as a Phred or Q score, is an integer value representing the estimated probability of an error, i.e. that the base is incorrect. If P is the error probability, then: $P = 10^{-(Q/10)}$. Put another way: $Q = -10 \log_{10}(P)$. Note that a Q score of 20 means $P = 0.01$, meaning that there is a 1% chance the base is wrong. Higher values represent lower probabilities of error.

Genotype:

No.	Genotype / Subtype	No. of Contigs	Total Reads
1	1a	1	5395

Contigs:

No	Contig Name	Gene target	Median Coverage	Length	Completeness	Total Reads
1	20190412_019057445001-190409_1a_NS3_c1	NS3	1969	926	98.20%	12382
2	20190412_019057445001-190409_1a_NS5A_c1	NS5A	2007	602	99.83%	8154
3	20190412_019057445001-190409_1a_NS5B_c1	NS5B	2024	1698	99.82%	22955

AA Substitution Analysis QC:

Summary

No	Gene target	Total AA substitutions analyzed	No. of substitutions passed optimally	No. of substitutions passed	No. of substitutions failed
1	NS3	81	81 (100.0%)	0 (0.0%)	0
2	NS5A	75	75 (100.0%)	0 (0.0%)	0
3	NS5B	65	56 (86.15%)	9 (13.85%)	0

NS3

No.	AA substitution	Nucleotide positions	Min. Coverage	Status
1	36T	3525-3527	13812	PASSED
2	36M	3525-3527	13812	PASSED
3	36G	3525-3527	13812	PASSED
4	36C	3525-3527	13812	PASSED
5	36L	3525-3527	13812	PASSED
6	36I	3525-3527	13812	PASSED
7	36A	3525-3527	13812	PASSED
8	41R	3540-3542	14978	PASSED
9	43V	3546-3548	14833	PASSED
10	43S	3546-3548	14833	PASSED
11	43I	3546-3548	14833	PASSED
12	43L	3546-3548	14833	PASSED
13	54G	3579-3581	14171	PASSED
14	54S	3579-3581	14171	PASSED
15	54A	3579-3581	14171	PASSED
16	54C	3579-3581	14171	PASSED
17	55A	3582-3584	13151	PASSED
18	55I	3582-3584	13151	PASSED
19	56F	3585-3587	13394	PASSED
20	56H	3585-3587	13394	PASSED
21	56L	3585-3587	13394	PASSED
22	80R	3657-3659	12642	PASSED
23	80N	3657-3659	12642	PASSED
24	80L	3657-3659	12642	PASSED
25	80H	3657-3659	12642	PASSED
26	80K	3657-3659	12642	PASSED
27	107I	3738-3740	13763	PASSED
28	109K	3744-3746	13815	PASSED
29	122T	3783-3785	12506	PASSED
30	122D	3783-3785	12506	PASSED
31	122G	3783-3785	12506	PASSED
32	122I	3783-3785	12506	PASSED
33	122N	3783-3785	12506	PASSED
34	122R	3783-3785	12506	PASSED
35	122A	3783-3785	12506	PASSED
36	132V	3813-3815	12433	PASSED
37	138T	3831-3833	12801	PASSED
38	155W	3882-3884	12545	PASSED
39	155S	3882-3884	12545	PASSED
40	155T	3882-3884	12545	PASSED
41	155M	3882-3884	12545	PASSED
42	155Q	3882-3884	12545	PASSED
43	155I	3882-3884	12545	PASSED
44	155K	3882-3884	12545	PASSED
45	155C	3882-3884	12545	PASSED
46	155G	3882-3884	12545	PASSED
47	156L	3885-3887	12200	PASSED
48	156S	3885-3887	12200	PASSED

49	156T	3885-3887	12200	PASSED
50	156V	3885-3887	12200	PASSED
51	156N	3885-3887	12200	PASSED
52	156G	3885-3887	12200	PASSED
53	156F	3885-3887	12200	PASSED
54	158A	3891-3893	12069	PASSED
55	158I	3891-3893	12069	PASSED
56	168F	3921-3923	10804	PASSED
57	168E	3921-3923	10804	PASSED
58	168G	3921-3923	10804	PASSED
59	168A	3921-3923	10804	PASSED
60	168C	3921-3923	10804	PASSED
61	168R	3921-3923	10804	PASSED
62	168V	3921-3923	10804	PASSED
63	168Y	3921-3923	10804	PASSED
64	168L	3921-3923	10804	PASSED
65	168K	3921-3923	10804	PASSED
66	168I	3921-3923	10804	PASSED
67	168H	3921-3923	10804	PASSED
68	168T	3921-3923	10804	PASSED
69	168S	3921-3923	10804	PASSED
70	168P	3921-3923	10804	PASSED
71	168N	3921-3923	10804	PASSED
72	170T	3927-3929	10799	PASSED
73	170F	3927-3929	10799	PASSED
74	170A	3927-3929	10799	PASSED
75	170T	3927-3929	10799	PASSED
76	170L	3927-3929	10799	PASSED
77	170A	3927-3929	10799	PASSED
78	170V	3927-3929	10799	PASSED
79	170L	3927-3929	10799	PASSED
80	170F	3927-3929	10799	PASSED
81	175L	3942-3944	10708	PASSED

NS5A

No.	AA substitution	Nucleotide positions	Min. Coverage	Status
1	24T	6327-6329	4718	PASSED
2	24R	6327-6329	4718	PASSED
3	24N	6327-6329	4718	PASSED
4	24M	6327-6329	4718	PASSED
5	24G	6327-6329	4718	PASSED
6	26E	6333-6335	4933	PASSED
7	28V	6339-6341	4980	PASSED
8	28S	6339-6341	4980	PASSED
9	28G	6339-6341	4980	PASSED
10	28T	6339-6341	4980	PASSED
11	28L	6339-6341	4980	PASSED
12	28A	6339-6341	4980	PASSED
13	28T	6339-6341	4980	PASSED
14	28M	6339-6341	4980	PASSED
15	29S	6342-6344	5063	PASSED
16	30I	6345-6347	5120	PASSED
17	30K	6345-6347	5120	PASSED
18	30G	6345-6347	5120	PASSED
19	30H	6345-6347	5120	PASSED
20	30E	6345-6347	5120	PASSED
21	30Y	6345-6347	5120	PASSED
22	30S	6345-6347	5120	PASSED
23	30T	6345-6347	5120	PASSED
24	30L	6345-6347	5120	PASSED
25	30R	6345-6347	5120	PASSED
26	30C	6345-6347	5120	PASSED
27	30S	6345-6347	5120	PASSED
28	30D	6345-6347	5120	PASSED
29	30K	6345-6347	5120	PASSED
30	30Q	6345-6347	5120	PASSED
31	30P	6345-6347	5120	PASSED
32	30H	6345-6347	5120	PASSED
33	30G	6345-6347	5120	PASSED
34	31I	6348-6350	5185	PASSED
35	31F	6348-6350	5185	PASSED
36	31V	6348-6350	5185	PASSED
37	31P	6348-6350	5185	PASSED
38	31M	6348-6350	5185	PASSED
39	32F	6351-6353	5275	PASSED
40	32L	6351-6353	5275	PASSED
41	32S	6351-6353	5275	PASSED

42	38P	6369-6371	5842	PASSED
43	38F	6369-6371	5842	PASSED
44	54R	6417-6419	7819	PASSED
45	54H	6417-6419	7819	PASSED
46	54Y	6417-6419	7819	PASSED
47	54T	6417-6419	7819	PASSED
48	54N	6417-6419	7819	PASSED
49	54Y	6417-6419	7819	PASSED
50	58R	6429-6431	8081	PASSED
51	58P	6429-6431	8081	PASSED
52	58L	6429-6431	8081	PASSED
53	58D	6429-6431	8081	PASSED
54	58S	6429-6431	8081	PASSED
55	58D	6429-6431	8081	PASSED
56	62A	6441-6443	7987	PASSED
57	62D	6441-6443	7987	PASSED
58	62P	6441-6443	7987	PASSED
59	62L	6441-6443	7987	PASSED
60	62R	6441-6443	7987	PASSED
61	62D	6441-6443	7987	PASSED
62	62T	6441-6443	7987	PASSED
63	62R	6441-6443	7987	PASSED
64	62E	6441-6443	7987	PASSED
65	92T	6531-6533	7469	PASSED
66	92K	6531-6533	7469	PASSED
67	93T	6534-6536	7585	PASSED
68	93W	6534-6536	7585	PASSED
69	93R	6534-6536	7585	PASSED
70	93S	6534-6536	7585	PASSED
71	93F	6534-6536	7585	PASSED
72	93C	6534-6536	7585	PASSED
73	93H	6534-6536	7585	PASSED
74	93L	6534-6536	7585	PASSED
75	93N	6534-6536	7585	PASSED

NS5B

No.	AA substitution	Nucleotide positions	Min. Coverage	Status
1	61G	7782-7784	7782	PASSED
2	112T	7935-7937	9425	PASSED
3	142T	8025-8027	10781	PASSED
4	159F	8076-8078	10779	PASSED
5	237G	8310-8312	10114	PASSED
6	282T	8445-8447	11107	PASSED
7	282R	8445-8447	11107	PASSED
8	307R	8520-8522	12632	PASSED
9	314F	8541-8543	13091	PASSED
10	314I	8541-8543	13091	PASSED
11	314P	8541-8543	13091	PASSED
12	316H	8547-8549	12834	PASSED
13	316N	8547-8549	12834	PASSED
14	316W	8547-8549	12834	PASSED
15	316Y	8547-8549	12834	PASSED
16	320F	8559-8561	12932	PASSED
17	320V	8559-8561	12932	PASSED
18	320I	8559-8561	12932	PASSED
19	321A	8562-8564	13092	PASSED
20	321I	8562-8564	13092	PASSED
21	335H	8604-8606	11685	PASSED
22	341H	8622-8624	15273	PASSED
23	368T	8703-8705	12876	PASSED
24	406A	8817-8819	8989	PASSED
25	406I	8817-8819	8989	PASSED
26	411S	8832-8834	8654	PASSED
27	414T	8841-8843	8952	PASSED
28	414L	8841-8843	8952	PASSED
29	414I	8841-8843	8952	PASSED
30	414V	8841-8843	8952	PASSED
31	419S	8856-8858	9052	PASSED
32	422K	8865-8867	9263	PASSED
33	423T	8868-8870	9121	PASSED
34	423V	8868-8870	9121	PASSED
35	445Y	8934-8936	8686	PASSED
36	445F	8934-8936	8686	PASSED
37	446Q	8937-8939	8582	PASSED
38	446K	8937-8939	8582	PASSED
39	448H	8943-8945	8465	PASSED
40	448C	8943-8945	8465	PASSED

41	449I	8946-8948	8414	PASSED
42	450V	8949-8951	8333	PASSED
43	451R	8952-8954	8288	PASSED
44	470S	9009-9011	7148	PASSED
45	473T	9018-9020	7359	PASSED
46	495S	9084-9086	6471	PASSED
47	495T	9084-9086	6471	PASSED
48	495L	9084-9086	6471	PASSED
49	495Q	9084-9086	6471	PASSED
50	495A	9084-9086	6471	PASSED
51	499A	9096-9098	6157	PASSED
52	553I	9258-9260	1028	PASSED
53	553T	9258-9260	1028	PASSED
54	553V	9258-9260	1028	PASSED
55	554S	9261-9263	1003	PASSED
56	554D	9261-9263	1003	PASSED
57	555H	9264-9266	981	WARNING
58	556G	9267-9269	968	WARNING
59	556R	9267-9269	968	WARNING
60	557R	9270-9272	954	WARNING
61	558R	9273-9275	837	WARNING
62	559V	9276-9278	743	WARNING
63	559N	9276-9278	743	WARNING
64	559G	9276-9278	743	WARNING
65	559I	9276-9278	743	WARNING

Review Result:

Please tick the appropriate box

Technician 1

APPROVED:

REJECTED:

Comments:

Technician 1's signature

Name printed

Date

Technician 2

APPROVED:

REJECTED:

Comments:

Technician 2's signature

Name printed

Date

Laboratory Director

APPROVED:

REJECTED:

Comments:

Lab. Director's signature

Name printed

Date

Vela is a trademark of Vela Holding Pte Ltd. *Sentosa*[®] is a registered trademark of Vela Holding Pte Ltd in several markets including the US and the European Union.

Laboratory Test Report



Run name : 190409_HCV_ABS / ZGQZN
Barcode : SentosaBC-8

Specimen ID : 019057445001-190409
Date reported : 12-Apr-2019

Laboratory Director Comments:

_____ Laboratory Director Signature	_____ Name Printed	_____ Date
--	-----------------------	---------------

Result:

No.	HCV genotype / subtype	No. of contigs	Total reads
1	1a	1	5395

Interpretation:

HCV genotype(s) / subtype(s) detected: 1a

Assay Method:

The Sentosa(R) SQ HCV Genotyping Assay is a next generation sequencing-based test intended for viral genotyping from human plasma or serum. This test is intended to be used on the Sentosa(R) SX101 with the Sentosa(R) SX Virus Total Nucleic Acid Plus II Kit in conjunction with the Sentosa(R) ST401 and Sentosa(R) SQ301 instruments.

Targeted sequencing was performed on this sample. HCV viral RNA was extracted from sample and amplified by one-step RT-PCR. The amplicons were enzymatically fragmented, ligated with barcoded adaptor and subjected to next generation sequencing on the Sentosa(R) SQ301 Sequencer. Subsequently, the Sentosa(R) SQ Suite software performed primary analysis (signal processing and base-calling) on the raw sequencing data generated by Sentosa(R) SQ301. After primary analysis, the data was transferred to Sentosa(R) SQ Reporter Server for secondary analysis and report generation.

Genotyping analysis was based on the sequencing result of NS5B region for HCV Genotypes 1-6. Resistance-Associated Substitution (RAS) analysis of selected targets of NS3 region, NS5A region and NS5B region are provided for HCV subtypes 1a and 1b and genotype 3.

Gene List:

Resistance-associated Substitution (RAS) analysis of NS5A, NS5B and NS3 genes was performed.

HCV Reference Sequences Used:

Hepatitis C virus subtype 1a: NC_004102, subtype 1b: AJ238799 and genotype 3 GU814263 - for RAS determination only

Disclaimer:

According to Distribution of In Vitro Diagnostic Products Labeled for Research Use Only or Investigational Use Only - Guidance for Industry and Food and Drug Administration Staff, Document issued on November 25, 2013 by U.S. Food and Drug Administration (<http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm253307.htm>), the criteria in section 812.2(c)(3) also include compliance with labeling requirements section CFR 809.10(c), which exempts shipments and other deliveries of IVDs from certain labeling requirements if either (1) the device complies with part 812, or (2) the investigation is not subject to part 812 and one of the following condition is met: For a product in the laboratory research phase of development, and not represented as an effective in vitro diagnostic product, all labeling bears the statement, prominently placed: For Research Use Only. Not for use in diagnostic procedures.

Vela is a trademark of Vela Holding Pte Ltd. Sentosa® is a registered trademark of Vela Holding Pte Ltd in several markets including the US and the European Union.

Contigs:

No.	HCV genotype / subtype	Contig name	Gene target	Completeness	Median coverage
1	1a	20190412_019057445001-190409_1a_NS5B_c1	NS5B	99.82%	2024
2	1a	20190412_019057445001-190409_1a_NS5A_c1	NS5A	99.83%	2007
3	1a	20190412_019057445001-190409_1a_NS3_c1	NS3	98.20%	1969

HCV Resistance-associated Substitutions (RAS):

No.	Gene target	Reference sequence	Nucleotide position	RAS detected	Coverage	Substitution frequency (%)	Coverage QC
1	NS3	NC_004102	3583	V55A	2023	94.22	Passed
2	NS5A	NC_004102	N/A	(see comments)	N/A	N/A	N/A
3	NS5B	NC_004102	N/A	(see comments)	N/A	N/A	N/A

N/A = Not applicable

Comments:

No drug resistance mutations were detected in NS5A region.

No drug resistance mutations were detected in NS5B region.

« L'ISPB-Faculté de Pharmacie de Lyon et l'Université Claude Bernard Lyon 1 n'entendent donner aucune approbation ni improbation aux opinions émises dans les thèses ; ces opinions sont considérées comme propres à leurs auteurs »

« L'ISPB-Faculté de Pharmacie de Lyon est engagé dans une démarche de lutte contre le plagiat. De ce fait une sensibilisation des étudiants et encadrants des thèses a été réalisée avec notamment l'incitation à l'utilisation de méthodes de recherche de similitudes »

CHANUT Marion

Evaluation de la plateforme de séquençage à haut débit Vela Diagnostics pour le génotypage et la détection de variants de résistance du Virus de l'Hépatite C

Th. D. Pharm., Lyon 1, 2020, 154p.

RESUME

Le Virus de l'Hépatite C (VHC), responsable d'hépatites, est à l'origine de cirrhoses et carcinomes hépatocellulaires et infecte chroniquement 71 millions de personnes au niveau mondial, avec 400 000 décès annuellement. Le VHC est ainsi un problème de santé publique majeur. Récemment, de nouveaux traitements ont été mis en place, et nécessitent, d'une part, la détermination du génotype viral et d'autre part, la détection d'éventuelles mutations de résistance dans les régions cibles des traitements, facteurs cruciaux pour le choix des molécules composant le traitement optimal de l'infection.

Au laboratoire de virologie de l'Hôpital de la Croix-Rousse à Lyon, ces informations virales étaient obtenues par séquençage de Sanger en routine jusqu'en mai 2018. Cette thèse a eu pour objectif l'évaluation d'un automate de séquençage de nouvelle génération (NGS) de Vela Diagnostics, afin de permettre son utilisation en routine pour le génotypage et la détermination des mutations de résistance chez les patients infectés par le VHC puis d'effectuer un bilan à 1 an de cette utilisation en routine.

Ce travail nous a permis de valider la fiabilité du séquençage NGS par la solution automatisée de Vela Diagnostics comparativement au séquençage Sanger pour l'obtention du génotype et du sous-type. Il fournit en outre automatiquement les données de résistance pour les génotypes les plus fréquents, à savoir 1a, 1b et 3a en France, d'où son utilité en pratique clinique. Pour les autres génotypes, l'analyse manuelle des séquences brutes, rendues disponibles par ce circuit de séquençage, permet d'obtenir les informations sur les mutations éventuelles. La capacité de cet automate de séquençage à identifier les co-infections et à révéler les mutations de faible fréquence liées à l'existence des quasi-espèces virales a également été mise en lumière.

Le travail effectué au cours de cette thèse a ainsi permis de valider l'utilisation du circuit de séquençage Vela Diagnostics pour le VHC en routine au laboratoire de virologie de la Croix-Rousse en lien avec ses performances et son automatisation.

MOTS CLES

Virus de l'Hépatite C
Séquençage à haut débit (NGS)
Génotypage
Mutations de résistance

JURY

Mme MORFIN SHERPA Florence, Professeur des Universités, Praticien Hospitalier
Mme SCHOLTES Caroline, Maître de Conférences des Universités, Praticien Hospitalier
Mme FROBERT Emilie, Maître de Conférences des Universités, Praticien Hospitalier
Mme LARRAT Sylvie, Professeur des Universités, Praticien Hospitalier

DATE DE SOUTENANCE

Jeudi 28 mai 2020

ADRESSE DE L'AUTEUR

80 rue Pierre Delore – 69008 LYON