

# Filière Technologies du vivant

Orientation Biotechnologie

## Travail de bachelor Diplôme 2018

*Séverine Arcioni*

*Diagnostic du cancer du poumon :  
Développement d'une méthode pour la  
détection de la mutation EGFR T790M*

- *Professeur*  
Sergio Schmid
- *Expert*  
Thomas Von Känel
- *Date de la remise du rapport*  
17.08.2018

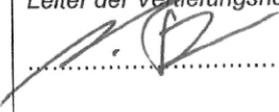
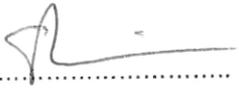
Ce rapport est l'original remis par l'étudiant.  
Il n'a pas été corrigé et peut donc contenir des inexactitudes ou des erreurs.

SYND	ETE	TEVI
X	X	X

Filière / Studiengang TEVI	Année académique / Studienjahr 2017/2018	No TD / Nr. DA bt\2018\10
Mandant / Auftraggeber <input type="checkbox"/> HES—SO Valais <input type="checkbox"/> Industrie <input checked="" type="checkbox"/> Etablissement partenaire Partnerinstitution	Etudiant / Student Arcioni Séverine  Professeur / Dozent Sergio Schmid	Lieu d'exécution / Ausführungsort <input type="checkbox"/> HES—SO Valais <input type="checkbox"/> Industrie <input checked="" type="checkbox"/> Etablissement partenaire Partnerinstitution
Travail confidentiel / vertrauliche Arbeit <input type="checkbox"/> oui / ja <sup>1</sup> <input checked="" type="checkbox"/> non / nein	Expert / Experte (données complètes) Dr. Thomas von Känel, Service de génétique médicale, Institut Central des Hôpitaux ; Av. du Grand-Champsec 86, CP 800, CH-1950 Sion	

<p>Titre / Titel</p> <p style="text-align: center;"><b>Diagnostic du cancer du poumon :</b></p> <p style="text-align: center;"><b>Développement d'une méthode pour la détection de la mutation EGFR T790M</b></p>
<p>Description / Beschreibung</p> <p>Les carcinomes pulmonaires non à petites cellules (NSCLC) représentent près de 85% des cancers pulmonaires. Le gène du récepteur du facteur de croissance épidermique (EGFR) est impliqué dans la régulation de la prolifération cellulaire. Des mutations dans sa séquence, particulièrement dans la région codant pour la poche de fixation de l'ATP, peuvent conduire à sa surexpression et donner des propriétés oncogènes à la cellule. Les médicaments de 2<sup>ème</sup> génération pour les NSCLC ciblaient donc ce gène. Mais au cours du traitement, de nouvelles résistances peuvent apparaître, notamment la mutation EGFR T790M. Le développement de l'Osimertinib qui se fixe de manière unique et irréversible au gène EGFR s'il est muté, a permis depuis début 2016 d'améliorer l'efficacité du traitement pour les patients présentant la mutation T790M. Il est donc important de connaître le statut de la tumeur afin de choisir le traitement adéquat.</p> <p>L'obtention de matériel tumoral lors d'une récurrence ou dans des métastases étant particulièrement difficile et invasif, et la quantité de matériel obtenu étant critique, la détection de l'ADN tumoral circulant dans le sang (ctDNA) a été explorée pour dépister la mutation EGFR T790M. Différentes études ont permis de confirmer l'utilité de cette méthode, d'autant plus qu'elle permet d'avoir une vue plus globale de la présence ou non de la mutation dans les éventuelles métastases également. Actuellement la sensibilité de la détection, que ce soit en digital PCR ou en NGS est parfois trop faible pour qu'un test négatif présuppose une non-sensibilité à l'Osimertinib. C'est pourquoi la mise en place de la détection de cette mutation par séquençage à haut débit (NGS) avec éventuellement un barcodage moléculaire serait profitable pour le patient d'une part, et pour le choix du traitement d'autre part, car le seuil de détection pourrait atteindre 0.1%. C'est pourquoi, en partenariat avec le service d'oncologie du RSV et de pathologie de l'ICH, le service de génétique médicale de l'ICH a décidé de mettre au point d'une part l'extraction d'ADN circulant et d'autre part la détection de la mutation EGFR T790M. Cela nécessite bien sûr d'obtenir des échantillons de patients, positifs et négatifs confirmés, afin de valider la méthode.</p> <p>A cette fin, différents kits d'extraction vont devoir être testés, selon les disponibilités du marché et ce qui se pratique dans d'autres laboratoires. La qualité de cette extraction va ensuite être évaluée afin d'obtenir le maximum d'ADN de qualité et de longueur suffisante.</p> <p>La mise au point de la détection de la mutation EGFR T790M se fera dans un deuxième temps. Des kits commerciaux existent mais ils imposent de tester d'autres mutations en parallèles. Une méthode « home-made » d'amplification de la région d'intérêt peut prendre tout son sens. Cela permettrait d'éviter les « incidental findings », d'améliorer la couverture de la mutation et donc la sensibilité du test, et de pouvoir réaliser l'analyse pour plusieurs patients simultanément tout en maîtrisant le coût.</p> <p>Objectifs :</p> <ul style="list-style-type: none"> <li>• Finaliser la mise au point d'un système PCR pour l'amplification de la région cible (définir les amorces, confirmer la spécificité du produit par séquençage)</li> <li>• Tester différentes méthodes d'extraction d'ADN circulant en considérant les paramètres suivants : qualité d'ADN, applicabilité, rendement, coût, ...</li> <li>• Design des primers pour l'amplification par NGS (avec ou sans molecular barcoding)</li> <li>• Faire des analyses par NGS (s'il reste du temps)</li> </ul>

<sup>1</sup> Par sa signature, l'étudiant-e s'engage à respecter strictement la directive DI.1.2.02.07 liée au travail de diplôme.  
Durch seine Unterschrift verpflichtet sich der/die Student/in, sich an die Richtlinie DI.1.2.02.07 der Diplomarbeit zu halten.

<p>Signature ou visa / <i>Unterschrift oder Visum</i></p> <p>Responsable de l'orientation / filière <i>Leiter der Vertiefungsrichtung / Studiengang:</i></p>  <p>.....</p> <p><sup>1</sup> Etudiant / Student :</p>  <p>.....</p>	<p>Délais / <i>Termine</i></p> <p>Attribution du thème / <i>Ausgabe des Auftrags:</i> <b>30.04.2018</b></p> <p>Fin des travaux de diplôme dans les laboratoires de l'école / <i>Ende der Diplomarbeiten in den Labors der Schule :</i> <b>13.07.2018</b></p> <p>Remise du rapport / <i>Abgabe des Schlussberichts:</i> <b>17.08.2018 / 12h00</b></p> <p>Expositions / <i>Ausstellungen der Diplomarbeiten:</i> <b>31.08.18</b></p> <p>Défense orale / <i>Mündliche Verfechtung:</i> <b>dès la semaine 35 / ab Woche 35</b></p>
---	--

Ce rapport est l'original remis par l'étudiant.  
Il n'a pas été corrigé et peut donc contenir des inexactitudes ou des erreurs.



## Diagnostic du cancer du poumon : Développement d'une méthode pour la détection de la mutation *EGFR* T790M

Diplômante Séverine Arcioni

### Objectif du projet

Le but de ce travail est de comparer le rendement de 4 kits d'extraction de ctDNA dans le plasma dans le cadre du cancer du poumon NSCLC-*EGFR* positif. Une bonne qualité d'ADN permettrait de rechercher la mutation de résistance *EGFR* p.T790M.

### Méthodes | Expériences | Résultats

Un total de 4 kits commerciaux ont été testés parallèlement sur 18 échantillons de sang provenant de volontaires sains et 6 échantillons de patients atteints de NSCLC. Le rendement et la qualité de l'ADN circulant ainsi obtenu ont été déterminés par Qubit, Fragment Analyzer et qPCR.

Au minimum une région de 96 pb de l'exon 20 d'*EGFR* a été amplifiée par PCR dans le ctDNA des patients, puis séquencée sur un MiSeq. Du ctDNA synthétique a également été utilisé pour déterminer le seuil inférieur de détection (LLOD).

Le coût, la méthode et les résultats obtenus avec le kit « QIAamp MinElute ccfDNA » en ont fait le meilleur choix.

La quantification avec le kit « Qubit dsDNA HS Assay » a donné des résultats équivalents à la qPCR pour un coût en temps et en réactifs moins important.

Le bruit de fond observé par séquençage à haut-débit (NGS) dans la région de la mutation p.T790M du gène *EGFR* a indiqué la nécessité d'utiliser des identifiants moléculaires uniques (UMI) afin d'abaisser la limite de détection à 0.1%.

Travail de diplôme  
| édition 2018 |

Filière

*Technologies du vivant*

Domaine d'application

*Biotechnologie*

Professeur responsable

*Sergio Schmid*

*sergio.schmid@hevs.ch*

Partenaire

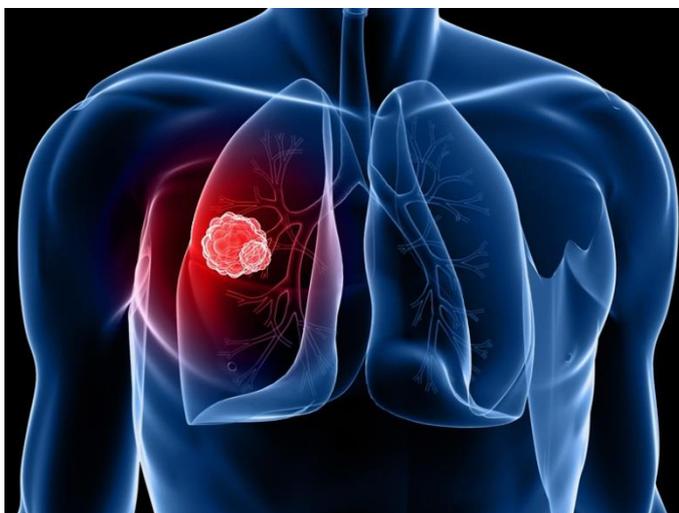
*ICH, Sion*



Hôpital du Valais  
Spital Wallis



Institut Central des Hôpitaux  
Zentralinstitut der Spitäler



Ce rapport est l'original remis par l'étudiant.  
Il n'a pas été corrigé et peut donc contenir des inexactitudes ou des erreurs.

# Table des matières

<b>Abréviations</b> .....	<b>1</b>
<b>1. Introduction</b> .....	<b>2</b>
1.1. Carcinomes pulmonaires non à petites cellules (NSCLC) .....	2
1.2. Le gène EGFR .....	2
1.3. ADN tumoral circulant dans le sang (ctDNA) .....	4
1.4. Techniques de détection des mutations .....	5
1.4.1. NGS .....	6
<b>2. Matériel et méthodes</b> .....	<b>7</b>
2.1. Prélèvement des échantillons .....	7
2.2. Extraction du ctDNA .....	7
2.3. Quantification du ctDNA .....	8
2.4. Détermination de la qualité du ctDNA .....	9
2.5. Concentration du ctDNA extrait avec Promega .....	10
2.6. Séquençage à haut-débit .....	10
2.6.1. Amplification d'une partie de l'exon 20 d'EGFR .....	10
2.6.2. Amplification des exons 18 à 21 avec le kit Agilent .....	11
2.6.3. Amplification d'une partie des exons 20 et 21 d'EGFR .....	11
2.6.4. Utilisation de 4 ctDNA synthétiques .....	11
2.6.5. Analyse des résultats .....	12
<b>3. Résultats</b> .....	<b>12</b>
3.1. Quantification du ctDNA .....	12
3.1.1. Qubit™ dsDNA HS Assay Kit .....	12
3.1.2. Qubit™ ssDNA Assay Kit .....	14
3.1.3. Fragment Analyzer™ Automated CE System .....	14
3.2. Qualité du ctDNA .....	15
3.3. Comparaison inter-kits de la quantité absolue de ctDNA .....	16
3.4. Amplifiabilité du ctDNA .....	17
3.5. Concentration du ctDNA extrait avec Promega .....	18
3.6. Séquençage à haut-débit .....	18
3.6.1. Amplification des exons 18 à 21 avec le kit Agilent .....	18
3.6.2. Détection de la mutation de résistance (p.T790M) .....	19
3.6.3. Détection de la mutation de référence .....	20
3.6.4. Utilisation de 4 ctDNA synthétiques .....	21
<b>4. Discussion</b> .....	<b>21</b>
4.1. Extraction de l'ADN circulant .....	21
4.1.1. Kit « AmoyDx Circulating DNA » .....	21

4.1.2.	Kit « Maxwell RSC ccfDNA » .....	22
4.1.3.	Kits Qiagen.....	22
4.1.4.	Méthode de quantification.....	23
4.2.	<i>Séquençage à haut-débit</i> .....	23
4.2.1.	Amplification des exons 18 à 21 avec le kit Agilent.....	23
4.2.2.	Détection de la mutation de résistance (p.T790M).....	23
4.2.3.	Détection de la mutation de référence .....	23
4.2.4.	Utilisation de 4 ctDNA synthétiques.....	24
<b>5.</b>	<b>Conclusions et perspectives .....</b>	<b>24</b>
<b>6.</b>	<b>Remerciements .....</b>	<b>25</b>
<b>7.</b>	<b>Bibliographie.....</b>	<b>26</b>
<b>8.</b>	<b>Annexes .....</b>	<b>1</b>

## Abréviations

<b>ADN/DNA</b>	Acide désoxyribonucléique/Deoxyribonucleic acid
<b>ARMS-PCR</b>	Amplification Refractory Mutation System PCR (amplification à l'aide d'amorces spécifiques)
<b>ATP</b>	Adénosine triphosphate
<b>Cp</b>	Crossing point (point d'intersection entre le signal et le seuil de référence)
<b>ctDNA</b>	circulating tumour DNA (ADN tumoral circulant)
<b>ddPCR</b>	droplet digital PCR (PCR digitale en gouttelette)
<b>dPCR</b>	digital PCR (PCR digitale)
<b>dsDNA</b>	double strand DNA (ADN double brin)
<b>EGF</b>	Epidermal Growth Factor (facteur de croissance épidermique)
<b>EGFR</b>	Epidermal Growth Factor Receptor (récepteur du facteur de croissance épidermique)
<b>FGFR</b>	Fibroblast Growth Factor Receptor (récepteur du facteur de croissance des fibroblastes)
<b>IASLC</b>	International Association for the Study of Lung Cancer (Association internationale pour l'étude du cancer du poumon)
<b>LLOD</b>	Lower Limit Of Detection (seuil de détection inférieur)
<b>NGS</b>	Next Generation Sequencing (séquençage à haut-débit)
<b>NSCLC</b>	Non-Small Cell Lung Cancer (cancer pulmonaire à non petites cellules)
<b>PCR</b>	Polymerase Chain Reaction (réaction de polymérisation en chaîne)
<b>PDGFR</b>	Platelet-Derived Growth Factor Receptor (récepteur du facteur de croissance plaquette-dérivé)
<b>qPCR</b>	quantitative PCR (PCR quantitative ou PCR en temps réel)
<b>ssDNA</b>	single strand DNA (ADN simple brin)
<b>tag</b>	balise fixée sur une amorce
<b>TE</b>	Tris-EDTA
<b>TK</b>	Tyrosine Kinase
<b>TKI</b>	Tyrosine Kinase Inhibitor (inhibiteur de la tyrosine kinase)
<b>UMI</b>	Unique Molecular Identifier (identifiant moléculaire unique)
<b>VF</b>	Variant Fraction (fréquence de variant)

# 1. Introduction

## 1.1. Carcinomes pulmonaires non à petites cellules (NSCLC)

En Suisse, le cancer du poumon est le 2<sup>ème</sup> cancer le plus fréquent chez l'homme (après celui de la prostate) et le troisième chez la femme (après celui du sein et le colo-rectal) [1]. Il représente plus de 4200 cas par an pour près de 3200 décès, ce qui en fait le plus meurtrier annuellement tous sexes confondus. Il est rarement diagnostiqué avant 30 ans, l'âge médian des patients étant de 70 ans. Environ 80 à 85% des cancers du poumon sont classés comme *Non-small Cell Lung Cancer* (NSCLC) selon leurs caractéristiques histologiques [2], puis subdivisés en adénocarcinomes (60% des cas), carcinomes épidermoïdes (30% des cas) ou carcinomes à grandes cellules (rares).

Le diagnostic est souvent posé à un stade avancé (métastatique) rendant la chirurgie impossible. Les progrès de ces dernières années dans le domaine de la médecine personnalisée ont permis d'établir un lien entre des mutations dans certains gènes et des traitements particuliers, plus efficaces et moins pénibles pour le patient, mais plus onéreux. De nombreux gènes ont été décrits dans ce contexte, pour différents cancers, notamment le gène du récepteur du facteur de croissance épidermique (*EGFR*) dans le cadre des NSCLC.

Cependant, la prévalence de mutations semble différer selon de nombreux critères, comme l'origine ethnique des patients : 10% des caucasiens et jusqu'à 35% des asiatiques avec un adénocarcinome du poumon présentent une mutation du gène *EGFR* [3], [4]. De plus, une mutation dans ce gène est également retrouvée chez 40% des malades non-fumeurs et seulement 7% des fumeurs. Des cas de mutations du gène *EGFR* ont également été décrits dans les carcinomes épidermoïdes, particulièrement lorsqu'il s'agit de patients non-fumeurs [5].

## 1.2. Le gène *EGFR*

Le gène *EGFR* code pour une glycoprotéine transmembranaire de la famille des tyrosine kinases. Cette famille comprend de nombreux autres membres, dont notamment le récepteur à l'insuline R, le récepteur du facteur de croissance plaquette-dérivé (PDGFR) et le récepteur du facteur de croissance des fibroblastes (FGFR) pour les plus connus. Le récepteur du facteur de croissance épidermique est constitué d'un domaine de liaison au ligand extracellulaire, relié par un segment hélicoïdal transmembranaire à un domaine de tyrosine kinase intracellulaire (figure 1) [6].

La liaison du ligand induit une activation du récepteur puis sa dimérisation, ce qui active plusieurs voies de signalisation en aval, impliquées entre autres dans la prolifération et la survie cellulaire [7]. Des mutations situées dans le domaine TK (tyrosine-kinase), codé par les exons 18 à 24, entraînent l'activation constitutive de la signalisation, en l'absence de ligand, et par conséquent un effet oncogénique.

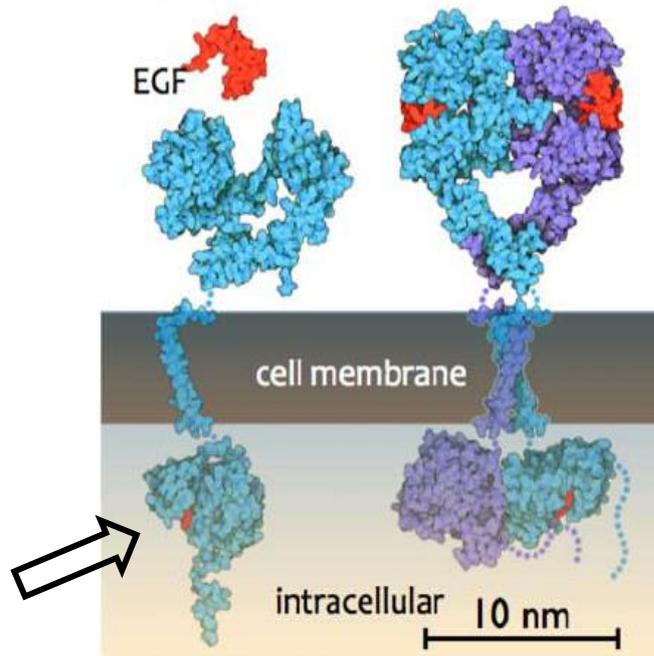


Figure 1: structure EGFR : à gauche : EGFR non-activé, à droite : EGFR dimérisé, en rouge ligand activateur, flèche : position de l'acide aminé 790

Les principales mutations somatiques décrites sont illustrées dans la figure 2 [8].

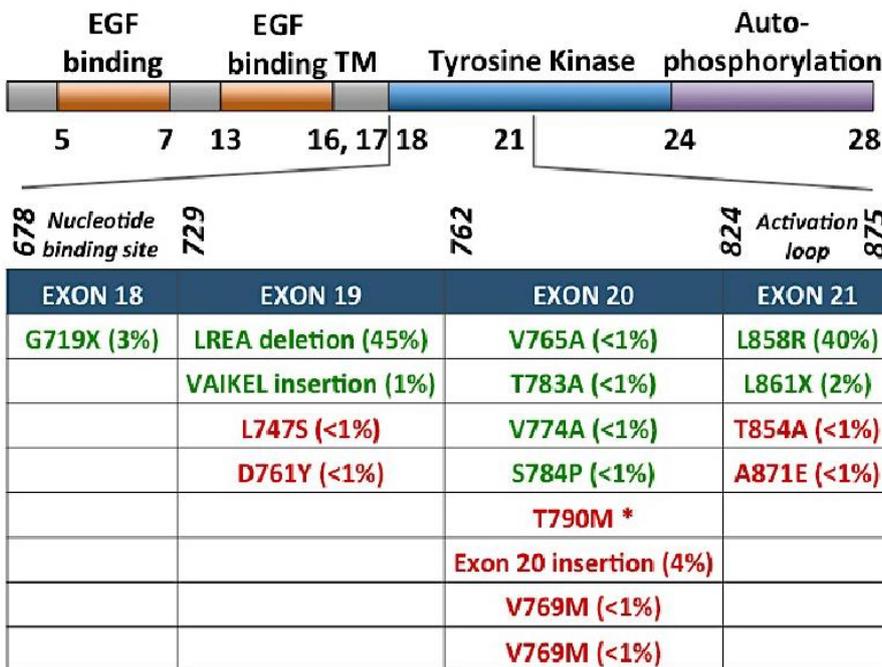


Figure 2: principales mutations somatiques du gène EGFR. En vert les mutations entraînant une sensibilité aux inhibiteurs EGFR TKI, en rouge une résistance. La fréquence est notée entre parenthèses. \* : Pour p.T790M, la fréquence initiale est de 5%, puis 60% après traitement EGFR TKI.

Les mutations sont situées dans la poche de fixation de l'adénosine triphosphate (ATP), soit les exons 18 à 21. La présence d'une mutation dans cette région d'EGFR permet le traitement à base d'inhibiteur de la tyrosine kinase (EGFR TKI) de première ou deuxième génération. Ces traitements sont très efficaces et peuvent même aboutir à une régression tumorale complète dans 60 à 70% des cas [9]. De plus, les traitements par EGFR TKI présentent une survie sans progression plus longue et des effets secondaires invalidants moindres que la chimiothérapie

standard [10]. Cependant, dans certains cas, suite au traitement, une résistance peut apparaître. La principale mutation (50-60% des cas) liée à cette résistance, nommée p.T790M, est un changement de C en T au nucléotide c.2369 de l'exon 20, qui résulte en une substitution d'une méthionine, non-polaire et hydrophobe, par une thréonine, hydrophile. Cette substitution implique un changement de la conformation de la protéine qui empêche l'adhésion des molécules EGFR TKI de première génération. Il s'agit d'une résistance acquise, apparaissant généralement dans les 9 à 14 mois [11] suivant le début du traitement et concernant la majorité des patients. En cas d'apparition de cette résistance, le cancer progresse à nouveau. Un traitement à base d'EGFR TKI de troisième génération peut alors être mis en place [9], pour autant que la présence de la mutation T790M ait été confirmée chez le patient

Le médicament de choix est alors l'Osimertinib, un EGFR TKI de 3<sup>ème</sup> génération. Il s'agit d'un composé mono-anilino-pyrimidique dont la structure chimique est illustrée à la figure 3 [12].

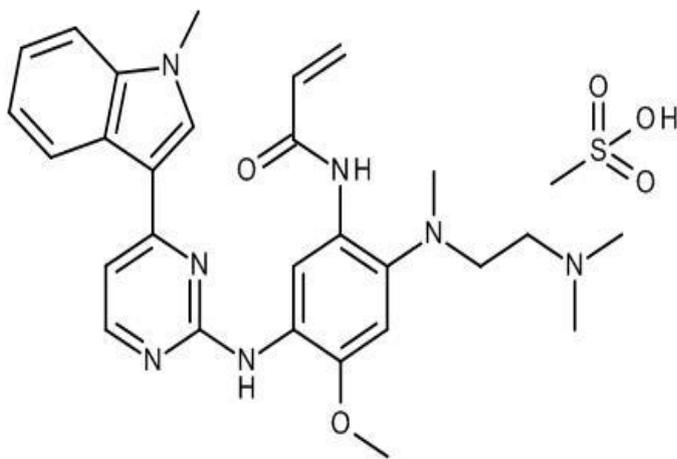


Figure 3: formule développée de l'Osimertinib

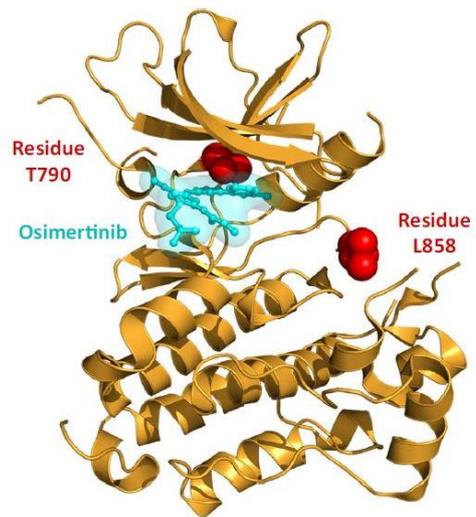


Figure 4: complexe EGFR – Osimertinib

Il se fixe de manière unique et irréversible à la protéine si elle est mutée (figure 4) [13], que ce soit la mutation primaire ou la mutation de résistance. De plus, il ne se fixe quasiment pas à la version sauvage de la protéine. Mis sur le marché début 2015, ce médicament a permis une nette amélioration dans le traitement des patients avec un NSCLC métastatique ayant acquis une résistance aux EGFR TKI traditionnels [14]. Par contre, pour le remboursement par les caisses maladies, la démonstration de la présence d'une mutation est exigée.

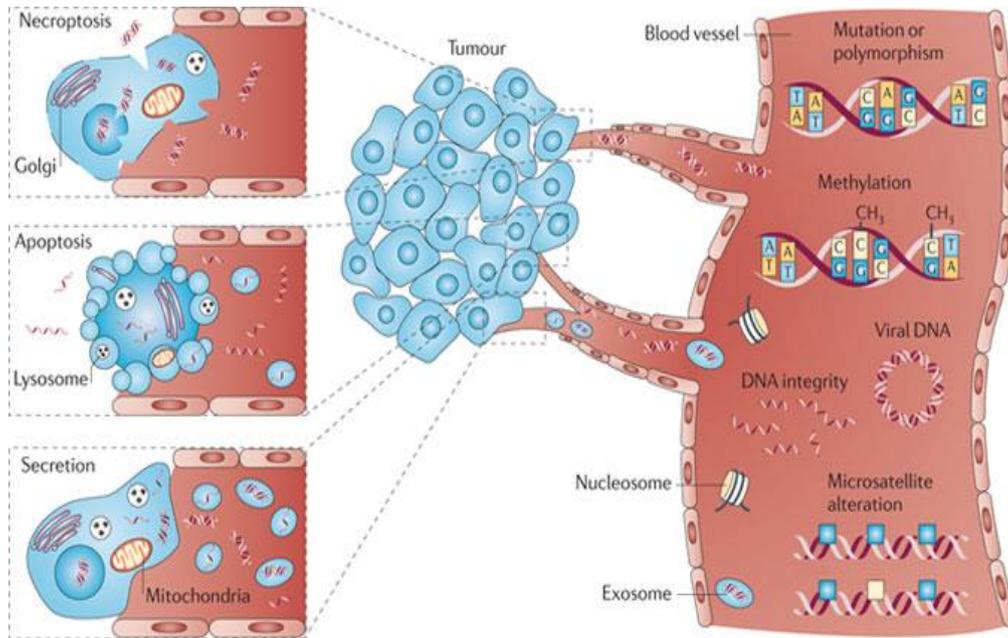
Afin de connaître le statut mutationnel d'un cancer, la méthode de choix était, jusqu'à présent, le prélèvement d'un fragment de tumeur par biopsie. Cette technique montre cependant des faiblesses. En effet, toute intervention chirurgicale implique un risque pour le patient. De plus les résultats en découlant ne sont représentatifs que d'une partie de la tumeur, et non des mutations métastatiques éventuelles. Sans oublier que selon la taille et la localisation tumorale, ce type de prélèvement peut tout simplement être impossible. Une autre méthode est donc nécessaire afin de détecter l'éventuelle apparition d'une résistance au traitement EGFR TKI.

### 1.3. ADN tumoral circulant dans le sang (ctDNA)

En 1948 déjà, Mandel [15] signalait la présence d'acide désoxyribonucléique (ADN) circulant dans le sang d'individus en bonne santé. Puis en 1977, Leon [16] démontrait que la concentration était

plus élevée dans le sang de patients atteints d'un cancer. Malgré ces découvertes précoces ce n'est que depuis 2012 [17] que l'analyse de l'ADN circulant dans une perspective de suivi de l'évolution d'un cancer sort gentiment du domaine de la recherche pour investir celui du diagnostic.

L'ADN circulant dans le système sanguin provient majoritairement de l'apoptose ou de la nécrose cellulaire. Comme illustré à la figure 5 [18], l'analyse de cet ADN peut concerner une recherche de mutation, de méthylation voire d'altération des microsatellites. Cependant, dans le cadre du NSCLC, seules les mutations ponctuelles des exons 18 à 21 du gène *EGFR* sont pertinentes.



**Figure 5: origine de l'ADN circulant : l'ADN tumoral est relâché par la tumeur et se retrouve dans la circulation sanguine sous forme de simple ou de double brin. Les phénomènes à l'origine du relargage de l'ADN peuvent être l'apoptose, la nécrose ou la sécrétion. La physiologie et le taux de libération ne sont pas encore totalement compris, la charge tumorale et le taux de prolifération des cellules tumorales peuvent jouer un rôle important dans ces événements. Le type de tumeur influence également le taux de ctDNA retrouvé.**

Un prélèvement de ctDNA se nomme une biopsie liquide, et consiste en une simple prise de sang. Elle peut donc être renouvelée facilement, et est moins dangereuse qu'une intervention chirurgicale. Afin de garantir la présence majoritaire de ctDNA, les tubes utilisés contiennent un stabilisateur qui préserve les cellules nucléées sanguines, évitant ainsi le relargage d'ADN cellulaire. Le volume prélevé est de 10 mL, afin d'obtenir environ 4 mL de plasma suite à 2 cycles de centrifugation. Puis l'étape clé consiste à isoler de l'ADN circulant possédant la meilleure qualité et la plus grande taille possible, sachant que la taille moyenne se situe généralement entre 120 et 200 pb [11]. A titre de comparaison, l'ADN génomique contient plusieurs milliers de paires de bases après extraction. Pour y parvenir, les méthodes commerciales actuellement disponibles sont basées soit sur des membranes de silice (par exemple : kits « AmoyDx Circulating DNA kit » (Amoy Diagnostics) et « QIAamp Circulating Nucleic Acid » (Qiagen)), soit sur une sélection par billes magnétiques (par exemple : kits « QIAamp MinElute ccfDNA » (Qiagen) et « Maxwell RSC ccfDNA » (Promega)).

#### 1.4. Techniques de détection des mutations

Dès lors que l'ADN obtenu est de bonne qualité, il est intéressant de rechercher d'une part la mutation de la tumeur primaire, afin de déterminer la fraction tumorale de l'échantillon, et d'autre

part la présence ou non de la mutation de résistance suite au traitement. Pour cela, plusieurs techniques sont utilisées. Au début des années 2000, la PCR digitale ou dPCR est apparue, qui a évolué ensuite en ddPCR (pour *droplet digital PCR*). Cette méthode permet le partitionnement de l'échantillon en 20'000 gouttelettes par émulsion eau-huile. L'amplification de la séquence cible, avec utilisation d'une sonde fluorescente, a lieu ensuite dans chaque gouttelette en parallèle. Le seuil de détection atteignable est de l'ordre de 0.2% [19]. Une autre approche utilisée est le séquençage à haut-débit (*Next Generation Sequencing*), qui permet un seuil de détection relativement bas, de l'ordre de 0.5% de mutation. D'autres méthodes ont été décrites, comme par exemple l'ARMS-PCR (*amplification refractory mutation system PCR*), mais leur sensibilité, spécificité et/ou coûts ne sont pas adéquats pour l'ADN circulant. Etant donné le matériel disponible au laboratoire et selon les recommandations de l'Association internationale pour l'étude du cancer du poumon (IASLC) [20], le NGS est choisi.

#### 1.4.1. NGS

Cette technique consiste en une première amplification en parallèle de toutes les régions d'intérêt par PCR multiplexe, suivi d'une deuxième amplification avec ajout de codes-barres (afin d'identifier le patient) et d'adaptateurs (nécessaires à l'analyse sur le séquenceur). Selon le fournisseur, les séquences des adaptateurs sont différentes. Une amélioration du seuil de détection semble être possible grâce à l'ajout lors de la première PCR de ce que l'on appelle du *molecular barcoding* ou également UMI et qui consiste à utiliser une amorce spécifique de la région mais avec une séquence variable à l'extrémité. Ce système permet ensuite de différencier les *PCR duplicates* (dus à des erreurs générées par la polymérase) des mutations réelles (figure 6) [21].

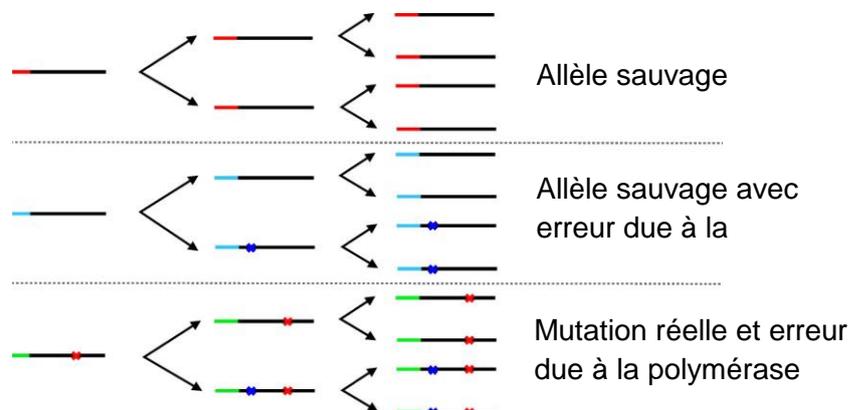


Figure 6: illustration du molecular barcoding : chaque molécule d'ADN est barcodée avec une séquence unique et aléatoire (symboles : rouge, turquoise et vert). A chaque cycle de PCR une copie de la séquence originale est générée (barcode identique). Lorsqu'une erreur due à la polymérase se produit (losanges bleus), seule une partie des molécules d'ADN avec le même barcode possèdera cette erreur spécifique. A l'inverse, une mutation réelle (losanges rouges) sera présente dans tous les amplicons, et pourra donc être interprétée comme « vraie mutation ».

Il semblerait que cette dernière méthode permette de descendre le seuil de détection à 0.1% de mutation, ce qui est particulièrement intéressant pour la mutation p.T790M qui indique une récurrence du cancer et implique donc un changement de traitement. Par contre, pour la recherche de la mutation de la tumeur primaire, une amplification avec un seuil de détection entre 1 et 2% est suffisante [22].

Certains kits commerciaux proposent l'amplification en parallèle de multiples régions du génome. Les inconvénients de cette approche sont le risque d'*incidental findings* (découvertes fortuites) ainsi que la diminution de la sensibilité. En effet un nombre maximum de *reads* (lecture) peut être

atteint par *run* sur le séquenceur. Selon que ces *reads* sont répartis en 50 cibles ou en 5, la couverture de chaque région changera d'un facteur 10 en moyenne.

La limite de détection dépend principalement du prélèvement. En effet, sachant que le génome humain diploïde correspond à 6.6 pg d'ADN, qu'avec 4 mL de plasma il est possible d'obtenir au moins 200 ng de ctDNA à une concentration de 30 ng/μL et que la quantité maximale de ctDNA dans la réaction de PCR est de 5 μL, cela correspond à 150 ng et donc 22'700 génomes diploïdes. En conséquence, 45'000 copies du gène *EGFR* sont potentiellement présentes, ce qui permet d'atteindre une sensibilité théorique de 0.002% (1 copie mutée parmi les 45'000). La concentration minimale requise pour un seuil de 0.1% serait de 1 ng/μL (500 génomes diploïdes). Il s'agit néanmoins d'une limite théorique, car des biais de sélection peuvent intervenir à chaque étape de la procédure. Il faut de plus tenir compte du bruit de fond de la position génomique au séquençage, principalement dû à un problème de désamination spontanée de cytosine (C), ce qui produit une thymine (T) [23], tout comme la mutation p.T790M.

La qualité de l'ADN circulant obtenu, sa concentration ainsi que le rendement global d'extraction sont donc les critères majeurs dans la mise en place de la biopsie liquide dans un but de détection, avec une LLOD (seuil de détection inférieur) de 0.1% de la mutation de récurrence p.T790M de l'exon 20 du gène *EGFR* dans le contexte des NSCLC.

## 2. Matériel et méthodes

### 2.1. Prélèvement des échantillons

De 1 à 4 tubes Streck « Cell-Free DNA BCT® CE » ont été utilisés pour les prélèvements de sang, tant sur les volontaires sains que pour les patients atteints de NSCLC, grâce à une collaboration avec le service d'oncologie de l'Hôpital du Valais. Selon les recommandations de la commission cantonale d'éthique de la recherche sur l'être humain (CER-VD), tous les tubes ont été anonymisés, et les patients, ainsi que les volontaires, informés du but du prélèvement (mise au point d'une analyse). Les tubes ont été centrifugés à 4000g pendant 10 minutes à 4°C, puis le plasma a été transféré dans des tubes Falcon 15 mL, et centrifugé à nouveau aux mêmes conditions.

L'ADN des volontaires a été utilisé afin de comparer le rendement et la qualité des différents kits d'extraction. Seuls les ctDNA de patients ont par la suite été séquencés, ceci afin d'éviter des découvertes fortuites.

Tous les échantillons ont été traités dans un délai maximal de 48h après le prélèvement. Le nombre et le choix des kits testés pour les échantillons s'est fait majoritairement en fonction du volume de plasma obtenu.

### 2.2. Extraction du ctDNA

Un volume de 4 mL de plasma a été utilisé avec les kits « AmoyDx Circulating DNA kit » (Réf : ADx-BL03) de la firme Amoy Diagnostics et « Maxwell RSC ccfDNA Large Volume » (Réf : AX1115) de la firme Promega (sur le robot « Maxwell® RSC Instrument » (Réf : AS4500)). Pour les 2 kits de la firme Qiagen : « QIAamp MinElute ccfDNA Midi Kit » (Réf : 55284) et « QIAamp Circulating Nucleic Acid Kit » (Réf : 55114), les volumes ont été de 3 à 5 mL selon la quantité de plasma à disposition. Les protocoles complets des kits sont en annexe.

Le nombre d'ADN de volontaires, respectivement de patients, extraits avec chacun des kits est le suivant : 14 et 6 avec chacun des kits de Qiagen, 10 et 4 avec le kit d'Amoy Diagnostics, 9 et 6 sur le robot de Promega (tableau 1).

**Tableau 1: liste des prélèvements de sang extraits avec les différents kits, en noir : volontaires sains, en rouge : patients atteints de NSCLC en processus de suivi, en gris : échantillons extraits selon les 4 méthodes**

N° ADN	QIAamp MinElute ccfDNA	QIAamp Circulating Nucleic Acid	AmoyDx Circulating DNA	Maxwell RSC ccfDNA
1	X			
2	X	X	X	
3	X	X	X	
4	X	X	X	
5	X	X	X	
6	X	X	X	
7	X	X	X	
8	X	X		X
9	X	X	X	
10	X	X		X
11	X	X	X	
12				X
13	X	X		X
14	X	X	X	X
15	X	X	X	X
16	X	X	X	X
17	X	X	X	X
18	X			X
19		X		X
20		X		X
21		X		X
22	X	X	X	X
23	X	X	X	X
24	X			X
<i>total</i>	<i>20</i>	<i>20</i>	<i>14</i>	<i>15</i>

L'ADN circulant a été remis en suspension selon les prescriptions du fabricant de chaque kit, soit dans un tampon fourni (« QIAamp Circulating Nucleic Acid Kit », « AmoyDx Circulating DNA kit » et « Maxwell RSC ccfDNA Large Volume ») soit dans de l'eau ultrapure (« QIAamp MinElute ccfDNA Midi Kit »).

### 2.3. Quantification du ctDNA

Plusieurs méthodes ont été appliquées afin de déterminer la quantité d'ADN présente après l'extraction. Dans un premier temps, une quantification de l'ADN avec les kits « Qubit™ dsDNA HS Assay Kit » (Réf : Q32851, limites de détection : 0.010 à 100 ng/μL), pour l'ADN double brin, et « Qubit™ ssDNA Assay Kit » (Réf : Q10212, limites de détection : 0.050 à 200 ng/μL), pour l'ADN simple brin, de la firme Invitrogen a été effectuée en ajoutant 2 à 5 μL d'échantillon au mélange réactionnel, afin de lire la concentration sur un Qubit® 2.0 avec le programme adéquat, selon le protocole du fournisseur.

Puis, afin de quantifier ainsi que de déterminer la taille de l'ADN obtenu, 5 µL de chacun des échantillons, mélangés à 15 ou 20 µL de tampon Tris-EDTA, ont été passés sur le Fragment Analyzer™ Automated CE System de la firme Advanced Analytical, avec le kit « dsDNA 905 Reagent Kit » (Réf : DNF-905), et le « Long Array » (Réf : A2300-1250-5580) pour une résolution de 3-5 pb et une LLOD de 0.5 ng/µL.

La quantité d'ADN circulant étant très variable, notamment entre les patients en rémission et ceux en récédive, la médiane a été utilisée préférentiellement à la moyenne lors de l'analyse des résultats.

## 2.4. Détermination de la qualité du ctDNA

Afin de connaître également la qualité de l'ADN extrait, deux tests par qPCR ont été réalisés, l'un amplifiant un fragment de 495 pb du gène *ZFY*, (ou *ZFX*) présent tant sur le chromosome X que Y (NM\_003411.3) (amorces *ZFY-F* : 5'-ACCRCTGACTGACTGTGATTACAC-3', *ZFY-R* : 5'-GCACYCTTTGGTATCYGAGAAAGT-3') et l'autre amplifiant 96 pb dans l'exon 20 du gène *EGFR* (LRG\_304t1) (amorces *EGFR\_E20F* : 5'-CGCCTGCTGGGCATCTG-3', *EGFR\_E20R* : 5'-GTCTTTGTGTTCCCGGACATAGT-3', selon Seki et al [24]). Toutes les amorces utilisées lors de ce travail ont été synthétisées par la firme Microsynth.

Le mélange réactionnel était composé, pour les deux réactions, de 1x « LightCycler® FastStart DNA Master SYBR Green I » (Réf : 03003230001) de la firme Roche Diagnostics, 1.25 mM de MgCl<sub>2</sub>, de 0.5 µM de chacune des amorces sens et anti-sens, de 2 µL d'échantillon, et du volume d'eau ultrapure nécessaire pour un volume de 10 µL final. Les programmes utilisés sur le LC2.0 de Roche sont décrits dans les tableaux 2 et 3 ci-dessous :

Tableau 2: Programme qPCR ZFY

ZFY	Denaturation	Amplification			Melting			Cooling
<b>Parameter</b>								
Analysis Mode	None	Quantification			Melting Curves			None
Cycles	1	40			1			1
Segment	1	1	2	3	1	2	3	1
Target [°C]	95	95	55	72	95	65	98	40
Hold [hh:mm:ss]	00:08:00	00:00:00	00:00:15	00:00:20	00:00:10	00:00:20	00:00:00	00:00:30
Ramp Rate [°C/s]	20	20	20	20	20	20	0.1	20
Acquisition Mode	None	None	None	Single	None	None	Continu.	None

Tableau 3: Programme qPCR EGFR

EGFR	Denaturation	Amplification			Melting			Cooling
<b>Parameter</b>								
Analysis Mode	None	Quantification			Melting Curves			None
Cycles	1	40			1			1
Segment	1	1	2	3	1	2	3	1
Target [°C]	95	95	60	72	95	65	98	40
Hold [hh:mm:ss]	00:08:00	00:00:00	00:00:10	00:00:12	00:00:10	00:00:20	00:00:00	00:00:30
Ramp Rate [°C/s]	20	20	20	20	20	20	0.1	20
Acquisition Mode	None	None	None	Single	None	None	Continu.	None

La PCR quantitative avec utilisation d'un standard interne peut permettre également de réaliser une quantification relative de l'ADN présent. Dans ce but, il est recommandé de réaliser les expériences en triplicata et de faire une série de dilutions du standard. Etant donné la faible quantité de ctDNA à disposition, le nombre d'échantillons à tester et sachant que le but premier de la qPCR lors de ce travail était de tester la qualité de l'ADN à disposition, les expériences ont été réalisées à simple avec un seul standard interne. Néanmoins les résultats obtenus ont été quantifiés selon l'équation (1) ci-dessous [25] et comparés aux résultats obtenus par les méthodes décrites plus haut (3.3).

$$\mathit{conc. ech.} = \frac{\mathit{conc. std}}{2^{(Cp\ ech. - Cp\ std)}} \quad \text{Équation (1)}$$

*conc.* : concentration en ADN

*ech.* : échantillon

*std* : standard

*Cp* : nombre de cycles nécessaires pour que la fluorescence émise dépasse significativement le bruit de fond

De plus entre 2 et 3 échantillons obtenu avec chacun des kits d'extraction ont également été amplifiés avec le kit « QC Plex » (Réf : QC-0521.100) de la firme Agilent selon le protocole du fournisseur puis passés sur le Fragment Analyzer, avec le kit DNF-910 (35-1500 pb). Le QC Plex est une méthode PCR multiplexe qui permet de générer 6 fragments (100, 200, 300, 400, 600 et 700 pb) afin de déterminer la qualité de l'ADN et sa capacité à être amplifié. Il contient également un standard interne de 150 pb afin de vérifier l'absence d'inhibiteurs de la PCR.

## 2.5. Concentration du ctDNA extrait avec Promega

Le volume minimal d'élution du kit Promega étant élevé (75 µL), le fabricant propose d'utiliser ensuite leur kit « ReliaPrep™ DNA Clean-Up and Concentration System » (A2891) afin d'obtenir une concentration en ADN plus élevée. Dans ce but, 50 µL de 4 échantillons extraits avec le kit « Maxwell RSC ccfDNA Large Volume » ont donc été passés sur ces colonnes, élués dans 30 µL d'eau ultrapure, puis quantifiés à nouveau avec le kit « Qubit™ dsDNA HS Assay Kit ».

## 2.6. Séquençage à haut-débit

Seuls les échantillons provenant de patients en processus de suivi pour un NSCLC-EGFR positif, ont ensuite été séquencés sur le MiSeq (Illumina) afin de confirmer la faisabilité de la détection de mutations du gène EGFR sur de l'ADN circulant provenant de biopsies liquides.

Dans ce but, 3 méthodes de préparation de bibliothèques ont été utilisées. Les deux premières ont été développées lors de ce travail, afin de détecter une ou plusieurs mutations précises mais avec un seuil de détection le plus bas possible. La troisième est une méthode commerciale, permettant d'amplifier entièrement les exons 18 à 21 du gène EGFR, mais elle est par conséquent plus coûteuse.

### 2.6.1. Amplification d'une partie de l'exon 20 d'EGFR

Pour 9 extraits obtenus, seules les amorces couvrant une partie de l'exon 20 du gène EGFR, (EGFR\_E20F et EGFR\_E20R, voir 2.4), auxquelles des tags ont été ajoutés (tag amorce F : 5'-AAGACTCGGCAGCATCTCCA-3', tag amorce R : 5'-GCGATCGTCACTGTTCTCCA-3') afin de permettre l'ajout des barcodes (MID) dans un deuxième temps, ont été utilisées. La préparation des bibliothèques consiste en une première amplification avec 2x « Taq PCR Master Mix Kit » (Réf :

201443) de la firme Qiagen, 0.5  $\mu$ M de chacune des amorces, 2  $\mu$ L d'ADN et de l'eau ultrapure pour un volume final de 25  $\mu$ L. La PCR, sur le thermocycleur « Trio » (Réf : 846-x-070-723) de la firme Biometra, consiste en une dénaturation initiale de 3' à 94°C, puis 35 cycles 30''/94°C – 30''/58°C – 30''/72°C, et finalement 10' d'élongation à 72°C. A la fin du programme, 20  $\mu$ L d'échantillon sont purifiés avec un ratio 1:2 de billes « Agencourt AMPure XP » (Réf : A63881) de la firme BeckmanCoulter, puis repris dans 40  $\mu$ L d'eau ultrapure.

Après une dilution de 300x, le produit de PCR est amplifié à nouveau afin d'y ajouter les barcodes ainsi que les adaptateurs nécessaires au passage sur le séquenceur. Pour cela, un mélange réactionnel contenant 0.5x « Universal PCR mix », 0.5x tampon « AR1 » et 0.625 U de Taq Polymérase (kit « drMID for Illumina NGS Systems » (Réf : ML-0208.240) de la firme Agilent) par échantillon est préparé. Puis 20  $\mu$ L de ce mix ainsi que 2  $\mu$ L de chacun des MID choisis et 2  $\mu$ L du produit PCR dilué sont mélangés pour une amplification selon le programme suivant : 10' à 98°C, puis 20 cycles 45''/95°C – 45''/64°C – 2'/68°C, et finalement 10' à 72°C.

Suite à cette deuxième amplification, le produit obtenu est à nouveau purifié avec un ratio 1:1.8 à l'aide des billes « Agencourt AMPure XP », puis repris dans 20  $\mu$ L de low TE (Réf : 12090015) de la firme ThermoFisher. Ensuite, 2  $\mu$ L de la librairie ainsi obtenue sont quantifiés avec le kit « Qubit™ dsDNA HS Assay Kit ».

Un mélange équimolaire contenant tous les échantillons à séquencer est réalisé et une dilution à 4'000 pM de ce mélange est utilisée pour le chargement de la cartouche (v2 ou v3 selon le nombre d'échantillons) qui est installée sur le séquenceur (selon protocole Illumina MiSeq). A la fin de l'analyse, les fichiers BAM, BAI, FASTQ et VCF sont récupérés.

### **2.6.2. Amplification des exons 18 à 21 avec le kit Agilent**

Pour 6 extraits, la librairie a été réalisée avec le kit « EGFR 18-21 MASTR » (Réf : MR-0130.024) de la firme Agilent, afin de couvrir les exons 18 à 21 du gène EGFR, selon le protocole fourni [26]. Suite à la première purification avec les billes magnétiques, le déroulement est identique à la procédure décrite précédemment (voir point 2.6.1).

### **2.6.3. Amplification d'une partie des exons 20 et 21 d'EGFR**

Les librairies des 8 derniers extraits ont été réalisées avec une amplification en parallèle d'une partie des exons 20 (voir point 2.6.1) et 21 du gène *EGFR* (amorce EGFR\_E21F : 5'-AGCCAGGGAACGTACTGGTGA-3' et amorce EGFR\_E21R : 5'-TGCCTCCTTCTGCATGGTAT-3', dessinées à l'aide de Primer3 [27] selon la séquence de référence LRG\_304t1, pour une taille de 107 pb et une Tm de 60°C, amorces auxquelles des tags identiques à ceux de l'exon 20 ont été ajoutés). Les deux réactions utilisant le même programme d'amplification, leurs produits ont été mélangés à un ratio de 1:0.75 lors de la première purification. La suite de la préparation a été identique (voir point 2.6.1).

### **2.6.4. Utilisation de 4 ctDNA synthétiques**

Le set de référence « EGFR Multiplex Cell Free DNA » (Réf : HD825) d'Horizon Discovery Ltd a été utilisé afin de déterminer la LLOD du NGS pour différentes mutations du gène *EGFR*. Il est composé de 4 tubes, contenant chacun 350 ng d'ADN fragmenté (environ 160 pb), avec des fréquences alléliques de respectivement 5, 1, 0.1 et 0% des 10 "variants" suivants du gène *EGFR* (tableau 4) :

Tableau 4: variants présents dans le set de référence « EGFR cfDNA » d'Horizon

position génomique	exon	nucléotide	protéine
chr7:55259515	21	c.2573T>G	p.L858R
chr7:55242466-55242480	19	c.2236_2250del15	p.E746_A750delELREA
chr7:55249071	20	c.2369C>T	p.T790M
chr7:55249009-55249010	20	c.2307_2308insGCCAGCGTG	p.V769_D770insASV
chr7:55259524	21	c.2582T>A	p.L861Q
chr7:55241707	18	c.2155G>A	p.G719S
chr7:55249091	20	c.2389T>A	p.C797S
chr7:55227924	12	c.1391C>T	p.S464L
chr7:55227926	12	c.1393G>A	p.G465R
chr7:55249005	20	c.2303G>T	p.S768I

Les amorces des exons 20 et 21 ont été utilisées comme décrit ci-dessus (point 2.6.3) pour un premier test du seuil de détection.

### 2.6.5. Analyse des résultats

Les résultats du séquençage à haut-débit (NGS) ont été visualisés à l'aide du logiciel IGV [28] et la couverture ainsi que la fréquence de chacune des bases aux positions d'intérêt ont été relevées. Les calculs statistiques ont été réalisés sur Excel, grâce à l'utilitaire d'analyse « Test d'égalité des espérances: observations paires ». Le test de Student (t-Test), la médiane et les quartiles ont été déterminés avec les formules idoines sur Excel (seuil de confiance utilisé : 95%).

L'analyse du gène *EGFR* pour les 6 patients a été réalisée en parallèle, dans un but diagnostique, à l'institut de médecine génétique et pathologie de l'hôpital universitaire de Bâle, avec le kit « Oncomine™ Lung cfDNA Assay » de ThermoFisher, sur un séquenceur S5 ou PGM. Une copie des résultats a été obtenue afin de comparer les données obtenues lors de ce travail avec celles de Bâle.

## 3. Résultats

Après centrifugation et décantation du sang prélevé, l'ADN circulant présent dans les plasmas a été extrait avec l'un des 4 kits à disposition (« QIAamp MinElute ccfDNA », « QIAamp Circulating Nucleic Acid », « AmoyDx Circulating DNA » ou « Maxwell RSC ccfDNA »).

### 3.1. Quantification du ctDNA

Le dosage de l'ADN circulant, simple et double brin, a été réalisé selon plusieurs méthodes, afin d'en déterminer la concentration.

#### 3.1.1. Qubit™ dsDNA HS Assay Kit

Tout d'abord, la quantification avec le kit dsDNA HS a donné, pour les volontaires, une médiane entre 0.196 ng/μL (kit « Maxwell RSC ccfDNA ») et 0.919 ng/μL (kit « QIAamp Circulating Nucleic Acid »). Les valeurs extrêmes se sont situées entre 0.434 et 1.29 ng/μL pour le kit « QIAamp MinElute ccfDNA », entre 0.459 et 2.10 ng/μL pour le kit « QIAamp Circulating Nucleic Acid », entre 0.149 et 0.439 ng/μL avec le kit « AmoyDx Circulating DNA » et entre 0.104 et 0.552 ng/μL

pour le kit « Maxwell RSC ccfDNA ».

En ce qui concerne l'ADN circulant des patients, les médianes ont été déterminées avec les mêmes kits et se sont établies entre 0.369 ng/μl (kit « AmoyDx Circulating DNA ») et 1.01 ng/μL (kit « QIAamp Circulating Nucleic Acid »). Les valeurs minimales ont été semblables aux valeurs obtenues dans le plasma des volontaires. Par contre les valeurs maximales se sont révélées bien plus élevées, s'établissant à 10.7, 17.7, 1.93 et 3.83 ng/μL (kits « QIAamp MinElute ccfDNA », « QIAamp Circulating Nucleic Acid », « AmoyDx Circulating DNA » et « Maxwell RSC ccfDNA », respectivement). Les valeurs obtenues sont récapitulées dans le tableau 5 et illustrées dans les figures 7 et 8 (volontaires sains, respectivement patients).

Tableau 5: valeurs obtenues (en ng/μL) avec le Qubit dsDNA HS Assay kit selon le kit d'extraction utilisé

	Volontaires sains			Patients		
	médiane	minima	maxima	médiane	minima	maxima
QIAamp MinElute ccfDNA	0.646	0.434	1.29	0.897	0.528	10.7
QIAamp Circulating Nucleic Acid	0.919	0.459	2.10	1.01	0.647	17.7
AmoyDx Circulating DNA	0.237	0.149	0.439	0.369	0.209	1.93
Maxwell RSC ccfDNA	0.196	0.104	0.552	0.381	0.199	3.83

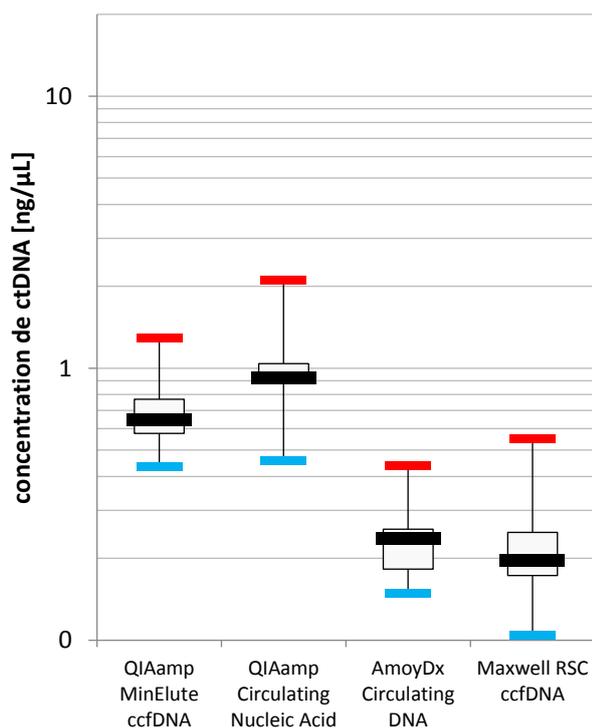


Figure 7: concentration en ctDNA des volontaires sains par Qubit dsDNA HS kit ; en noir : la médiane, en bleu : minima, en rouge : maxima, boîte : 1er et 3ème quartile

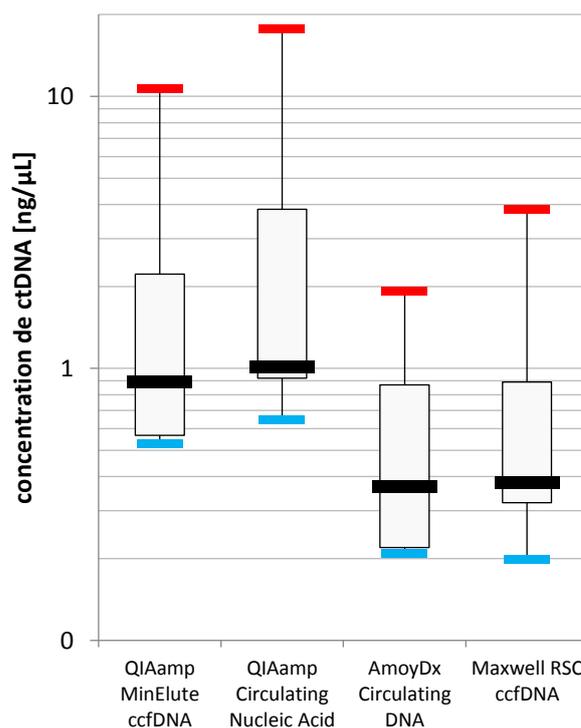


Figure 8: concentration en ctDNA des patients par Qubit dsDNA HS kit en noir : la médiane, en bleu : minima, en rouge : maxima, boîte : 1er et 3ème quartile

Un test de répétabilité, avec deux échantillons de chacun des kits d'extraction, a permis de constater la fiabilité de cette méthode de quantification, comme illustré dans la figure 9.

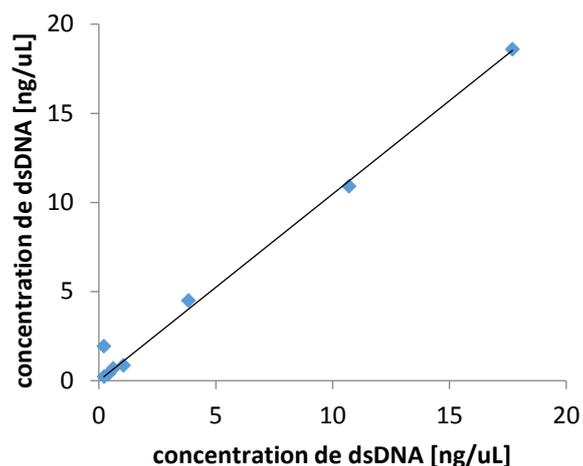


Figure 9: test de répétabilité du kit « Qubit dsDNA HS »

Le coefficient de corrélation de Pearson obtenu entre les échantillons est de 99.5%.

### 3.1.2. Qubit™ ssDNA Assay Kit

La méthode permettant le dosage de l'ADN simple brin détecte également tout l'ADN double brin. Les résultats obtenus ont donc montré des valeurs de concentration médiane plus élevées. La différence a été de l'ordre de 7x plus d'ADN simple brin que la valeur mesurée pour l'ADN double brin. Cela quel que soit le kit d'extraction utilisé.

Le test de répétabilité effectué avec 2 à 3 extraits pour chacun des kits (tableau 6), a donné un coefficient de corrélation de Pearson de 99.1% pour les 9 échantillons quantifiés à double.

Tableau 6: test de reproductibilité de la quantification avec le kit Qubit ssDNA ; ssDNA-1 : concentration obtenue lors du 1<sup>er</sup> dosage, ssDNA-2 : concentration obtenue lors du 2<sup>ème</sup> dosage

Kit	N° ADN	ssDNA - 1	ssDNA - 2
QIAamp MinElute ccfDNA	15	3.2 ng/μL	3.9 ng/μL
	23	70 ng/μL	79 ng/μL
QIAamp Circulating Nucleic Acid	4	6.7 ng/μL	8.2 ng/μL
	8	5.4 ng/μL	6.0 ng/μL
	15	27 ng/μL	21 ng/μL
AmoyDx Circulating DNA	4	2.6 ng/μL	6.6 ng/μL
	9	1.0 ng/μL	1.7 ng/μL
Maxwell RSC ccfDNA	14	0.94 ng/μL	1.3 ng/μL
	15	0.89 ng/μL	1.4 ng/μL

### 3.1.3. Fragment Analyzer™ Automated CE System

La médiane pour la quantification effectuée sur le Fragment Analyzer chez les volontaires s'est établie entre 0.363 ng/μl (kit « AmoyDx Circulating DNA ») et 1.47 ng/μL (kit « QIAamp Circulating Nucleic Acid »). Quant à la médiane pour les échantillons de patients, elle s'est située entre 2.12 ng/μL (kit « AmoyDx Circulating DNA ») et 7.08 ng/μL (kit « QIAamp Circulating Nucleic Acid »). Les résultats des autres kits ainsi que les valeurs maximales et minimales sont illustrées dans le tableau 7.

Tableau 7: concentrations, en ng/ $\mu$ L, obtenues sur Fragment Analyzer ; en rouge : valeurs en dessous du seuil de détection

	Volontaires sains			Patients		
	médiane	minima	maxima	médiane	minima	maxima
QIAamp MinElute ccfDNA	0.973	0.390	15.4	3.92	2.02	116
QIAamp Circulating Nucleic Acid	1.47	0.633	5.02	7.08	3.16	302
AmoyDx Circulating DNA	0.363	0.000	1.01	2.12	0.150	7.66
Maxwell RSC ccfDNA	0.607	0.146	3.85	2.57	0.779	28.5

Il faut cependant relever que la plupart des concentrations mesurées étaient proches, voire inférieures à la LLOD de la méthode, particulièrement avec les kits « AmoyDx Circulating DNA » et « Maxwell RSC ccfDNA ».

Un test de reproductibilité a été effectué avec 2 échantillons, ayant une concentration supérieure à la LLOD, obtenus avec chacune des méthodes d'extraction. Les résultats montrent des différences allant de 2 fois moins à 3 fois plus d'ADN détecté, bien que la taille moyenne obtenue soit, elle, comparable (tableau 8).

Tableau 8: test de reproductibilité pour le Fragment Analyzer ; FA-1 : concentration obtenue au 1er passage, FA-2 : concentration obtenue au 2ème passage

Kit	N° ADN	FA - 1	FA - 2
QIAamp MinElute ccfDNA	22	2.1 ng/ $\mu$ L	2.9 ng/ $\mu$ L
	24	3.1 ng/ $\mu$ L	5.3 ng/ $\mu$ L
QIAamp Circulating Nucleic Acid	15	4.5 ng/ $\mu$ L	2.9 ng/ $\mu$ L
	16	31 ng/ $\mu$ L	15 ng/ $\mu$ L
AmoyDx Circulating DNA	7	1.0 ng/ $\mu$ L	0.46 ng/ $\mu$ L
	23	7.7 ng/ $\mu$ L	9.8 ng/ $\mu$ L
Maxwell RSC ccfDNA	15	3.8 ng/ $\mu$ L	1.8 ng/ $\mu$ L
	23	28 ng/ $\mu$ L	92 ng/ $\mu$ L

La taille moyenne des fragments des échantillons s'est située entre 152 et 169 pb, quel que soit le kit, autant pour le ctDNA des volontaires que pour celui des patients.

### 3.2. Qualité du ctDNA

Concernant le control négatif, le test par qPCR pour un amplicon de 495 pb (gène *ZFY*) effectué sur les ADN extraits n'a pas montré d'amplification pour la plupart des échantillons. Uniquement 1 échantillon extrait avec le kit « QIAamp MinElute ccfDNA », 3 avec le kit « QIAamp Circulating Nucleic Acid » et 5 avec le kit « AmoyDx Circulating DNA » ont permis de détecter un signal avant le 35<sup>ème</sup> cycle. Le kit « Maxwell RSC ccfDNA » n'a, quant à lui, présenté aucun échantillon amplifiant à 495 pb.

Par contre, l'amplification d'une partie de l'exon 20 du gène *EGFR*, à 96 pb, a été détectée, comme attendu, dans tous les échantillons. Il est toutefois intéressant de noter qu'il est sorti en moyenne plus tardivement (32<sup>ème</sup> cycle) avec les extraits du kit « Maxwell RSC ccfDNA » qu'avec les autres, qui sont sortis respectivement au 29<sup>ème</sup> cycle (kit « QIAamp MinElute ccfDNA »), 30<sup>ème</sup> cycle (kit « QIAamp Circulating Nucleic Acid ») et 28<sup>ème</sup> cycle (kit « AmoyDx Circulating DNA »).

La robustesse de la méthode a été testée par la répétition, lors d'une nouvelle analyse, de 2

échantillons au hasard (1 de volontaire et 1 de patient) extrait avec chacun des kits (figure 10).

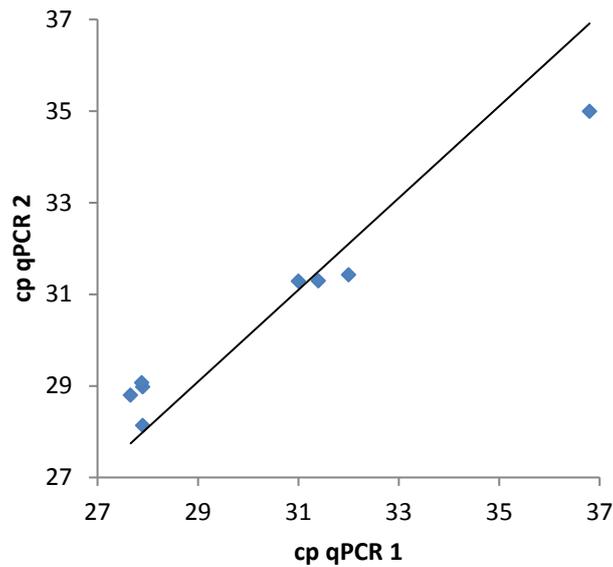


Figure 10: test de répétabilité de la qPCR pour *EGFR*

Le coefficient de corrélation de Pearson est de 98.9% pour *EGFR*. Pour *ZFY* les résultats obtenus (négatifs) ont été semblables.

### 3.3. Comparaison inter-kits de la quantité absolue de ctDNA

Pour 6 des échantillons, les 4 méthodes d'extraction ont été appliquées. Le contrôle positif, de concentration connue, utilisé lors de la qPCR a permis d'établir une corrélation entre le cp observé (point d'intersection signal - seuil de référence) et la quantité d'ADN circulant dans les échantillons selon l'équation (1) (voir point 2.4). Une comparaison des quantités absolues (nanogrammes par millilitre de plasma) a pu ainsi être effectuée. Les quantifications au Qubit, que ce soit en double ou en simple brin, et au Fragment Analyzer étant comparables ( $p > 0.05$ ), seuls ceux du « Qubit™ dsDNA HS Assay Kit » sont illustrés à la figure 11. Les résultats par qPCR sont eux représentés dans la figure 12.

Au Qubit, seul de kit « AmoyDx Circulating DNA » montre des résultats de quantification clairement abaissés par rapport aux 3 autres méthodes. Par contre, par qPCR, ce sont les résultats du kit « Maxwell RSC ccfDNA » qui sont abaissés pour tous les échantillons, et les extraits obtenus avec le kit « AmoyDx Circulating DNA » présentent des résultats intéressants.

L'analyse statistique comparant la quantification par Qubit dsDNA HS et par qPCR pour l'ensemble des échantillons (n entre 14 et 20 selon les kits, voir point 2.2) donne les résultats suivants selon la méthode d'extraction utilisée (tableau 9) :

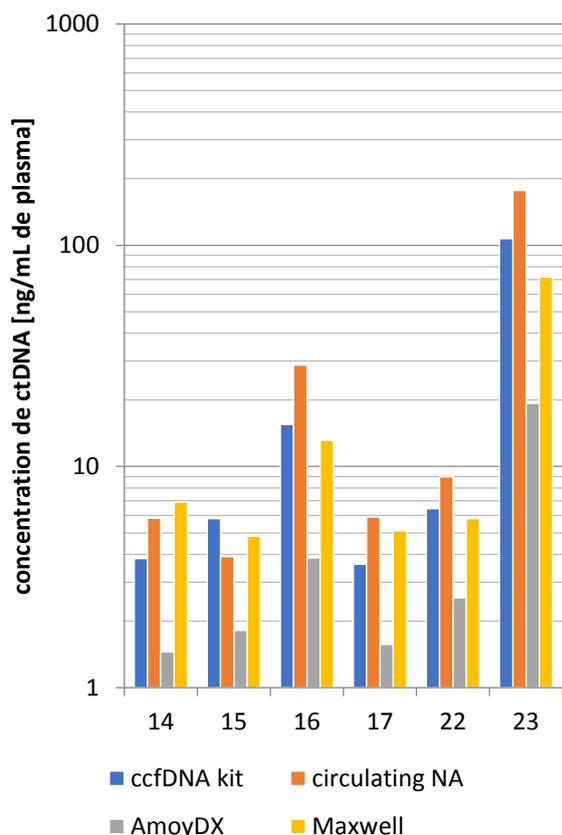


Figure 11 : comparaison des concentrations obtenues par "Qubit dsDNA HS" entre les différents kits pour 6 échantillons

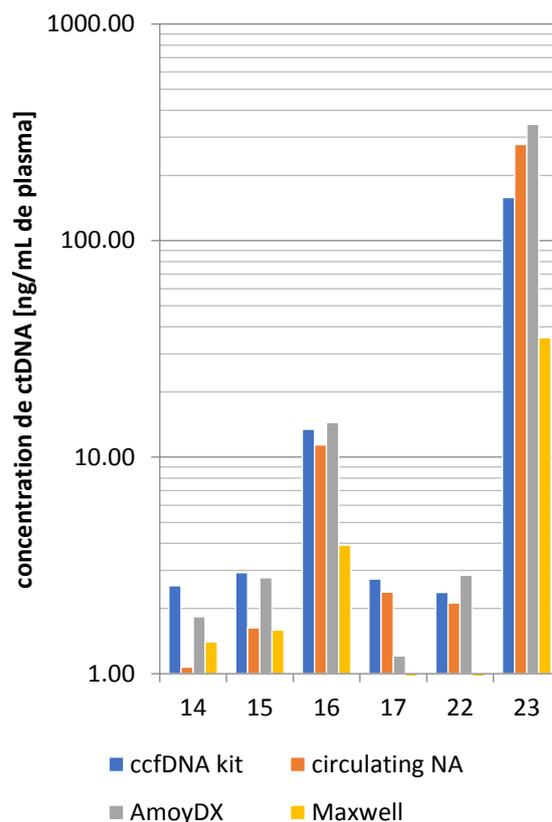


Figure 12: comparaison des concentrations obtenues par qPCR (96 pb du gène EGFR) entre les différents kits pour 6 échantillons

Tableau 9: résultats statistiques de comparaison des méthodes Qubit dsDNA HS et qPCR selon les kits utilisés

	Coefficient de corrélation de Pearson	T-test
QIAamp MinElute ccfDNA	0.996	0.37
QIAamp Circulating Nucleic Acid	0.994	0.91
AmoyDx Circulating DNA	0.989	0.29
Maxwell RSC ccfDNA	0.996	0.02

Un coefficient de corrélation de Pearson proche de 1 signifie une corrélation totale entre les valeurs obtenues par les 2 méthodes. Lorsque le résultat de T-test est supérieur à 0.05, cela signifie qu'il n'y a pas de différence significative entre les 2 types de mesure.

Alors que trois des quatre kits d'extraction présentent des résultats comparables entre les deux méthodes de quantification, seuls les résultats obtenus avec le kit « Maxwell RSC ccfDNA » sont significativement différents.

### 3.4. Amplifiabilité du ctDNA

Pour tous les échantillons amplifiés avec le kit « QC Plex », des fragments ont été détectés. Les concentrations maximales obtenues ont été retrouvées dans les amplicons de 100 pb. Avec les deux kits de la société Qiagen et celui de Amoy Diagnostics, des fragments de plus de 400 pb ont pu être amplifiés. Cependant, le kit « QIAamp Circulating Nucleic Acid » a présenté les concentrations des fragments de plus de 300 pb les plus élevées. Les 2 échantillons extraits avec

le kit « Maxwell RSC ccfDNA » n'ont permis de détecter que des fragments jusqu'à 300 pb y compris. La figure 13 illustre les résultats de 7 des 10 échantillons testés sur le Fragment Analyzer.

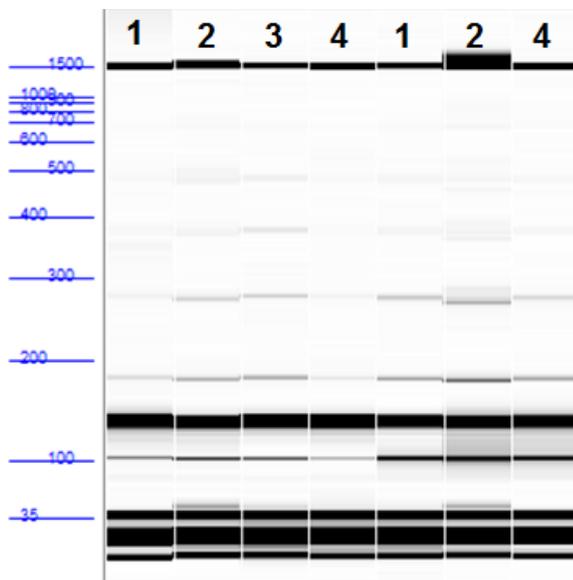


Figure 13: Image du Fragment Analyzer représentant 7 des échantillons testés par « QC Plex » extraits avec les kits suivants : « QIAamp MinElute ccfDNA » (1), « QIAamp Circulating Nucleic Acid » (2), « AmoyDx Circulating DNA » (3) et « Maxwell RSC ccfDNA »(4). La bande intense, à 150 pb, correspond au contrôle interne du kit.

### 3.5. Concentration du ctDNA extrait avec Promega

Le volume minimal d'élution avec le kit « Maxwell RSC ccfDNA » étant important, les extraits ont présenté des concentrations en ADN circulant faibles. Afin d'obtenir une quantité d'ADN par microlitre plus élevée, 4 des extraits obtenus avec ce kit ont été passés sur les colonnes fournies par Promega afin de concentrer l'ADN. Or, après dosage avec le kit « Qubit™ dsDNA HS Assay Kit », tous ont perdu au mieux la moitié, voire tout l'ADN présent, ce dernier n'étant plus quantifiable.

### 3.6. Séquençage à haut-débit

L'ADN obtenu avec les différents kits pour les 6 patients a été amplifié selon les 3 méthodes à disposition afin de rechercher des mutations somatiques du gène *EGFR* dans le but de valider la qualité et l'origine de l'ADN extrait.

#### 3.6.1. Amplification des exons 18 à 21 avec le kit Agilent

L'amplification des 4 exons ciblés par le kit Agilent « EGFR 18-21 MASTR » a permis d'obtenir une couverture moyenne de chacun des exons dans les 6 extraits testés entre 1'600 et 9'000 reads environ. Les résultats sont illustrés dans la figure 14.

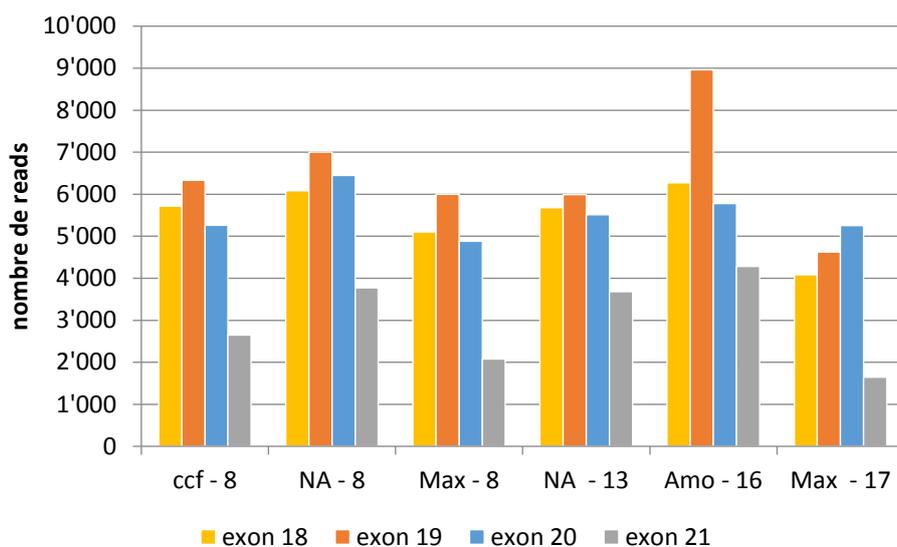


Figure 14: couverture des différents exons pour les 6 extraits avec le kit « EGFR 18-21 MASTR » ; ccf : kit « QIAamp MinElute ccfDNA », NA : kit « QIAamp Circulating Nucleic Acid », Amo : kit « AmoyDx Circulating DNA » et Max : kit « Maxwell RSC ccfDNA », le chiffre correspond au numéro d'ADN

L'écart maximal de couverture, pour un même échantillon, entre ses différents exons a passé du simple au triple.

### 3.6.2. Détection de la mutation de résistance (p.T790M)

Tous les extraits provenant de patients atteints de NSCLC ont été testés pour la mutation c.2369C>T, p.T790M, dans l'exon 20 du gène *EGFR*, selon une des 3 méthodes décrites aux points 2.6.1, 2.6.2 ou 2.6.3. Les résultats obtenus sur le séquenceur MiSeq ont présenté une couverture entre 4'900 et 36'500 *reads* pour cette position. Etant donné le bruit de fond à cette position, et en comparaison avec les résultats des analyses effectuées à l'institut de médecine génétique et pathologie de l'hôpital universitaire de Bâle, le seuil de détection se situe un peu au-dessus de 0.2%. En effet dans les extraits d'échantillons, négatifs à Bâle, la mutation a été détectée à une fréquence moyenne de 0.12% mais avec un maximum à 0.21% à Sion.

Néanmoins, ces résultats divergent légèrement selon les kits. En effet, pour les 5 patients négatifs, la fréquence moyenne, en pourcents, de la mutation s'est établie à  $0.097 \pm 0.061$  (n=5) pour le kit « QIAamp MinElute ccfDNA », à  $0.13 \pm 0.037$  (n=5) pour le kit « QIAamp Circulating Nucleic Acid », à  $0.13 \pm 0.074$  (n=3) pour le kit « AmoyDx Circulating DNA » et à  $0.11 \pm 0.064$  (n=5) pour le kit « Maxwell RSC ccfDNA ».

Etant donné les 30% de fréquence allélique moyenne du nucléotide T à la position c.2369, de l'échantillon n°23, faisant partie des 6 patients testés, il est le seul à présenter une récurrence accompagnée d'une résistance. Les résultats de ce patient, selon le kit d'extraction utilisé, sont illustrés à la figure 15, sachant que le résultat de Bâle donne une présence de 40% de la mutation de résistance.

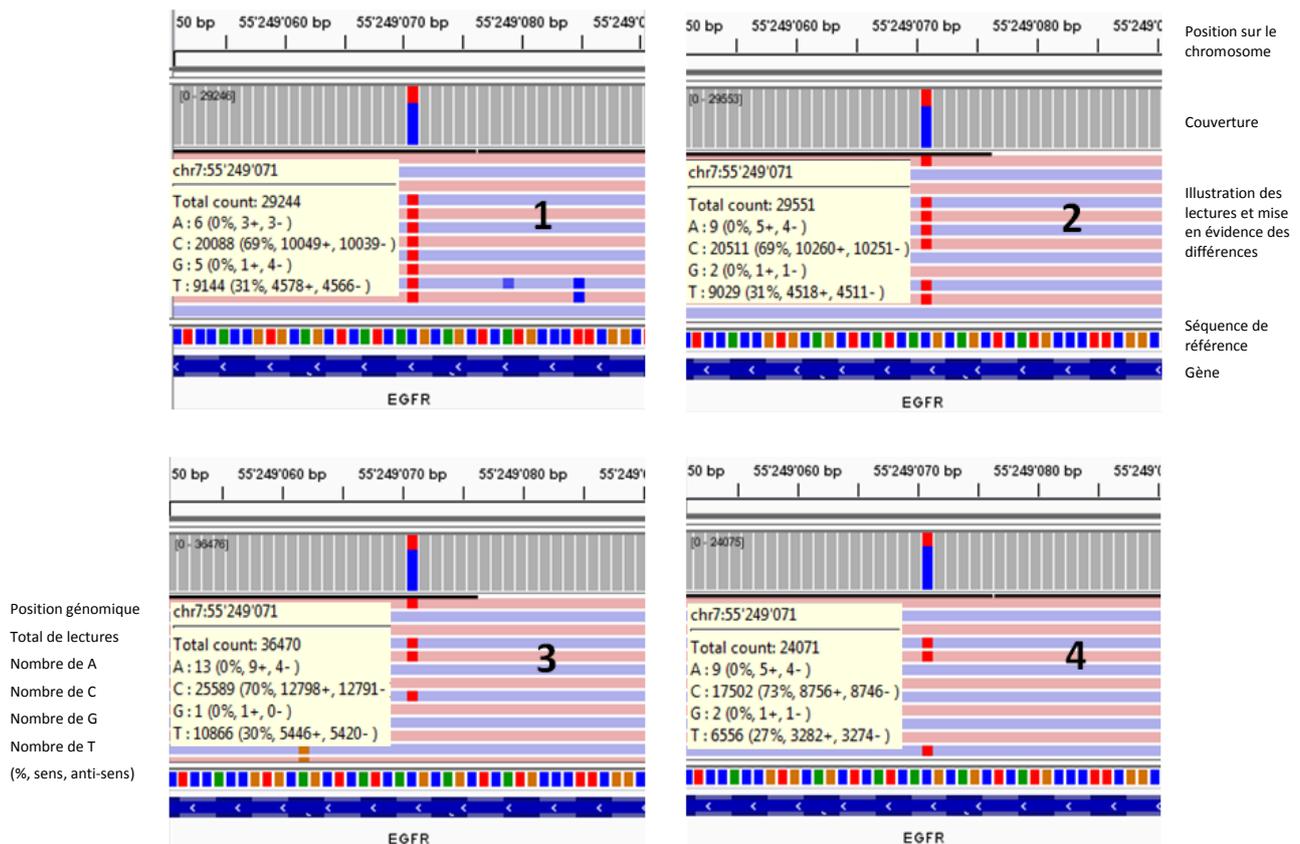


Figure 15: Visualisation de la couverture et de la fréquence de chacun des nucléotides à la position c.2369 du gène *EGFR* pour les kits « QIAamp MinElute ccfDNA » (1), « QIAamp Circulating Nucleic Acid » (2), « AmoyDx Circulating DNA » (3) et « Maxwell RSC ccfDNA »(4) sur IGV ; mutation recherchée : C>T ; barre bleue : nucléotide C, barre rouge : nucléotide T, en vert :nucléotide A, en orange : nucléotide G ; lignes rose et bleu : séquences identiques à la référence en sens et anti-sens.

La mutation est retrouvée dans les 4 extraits du patient, à des fréquences alléliques entre 27 et 31%, quel que soit le sens de lecture, avec une couverture d'au moins 24'000x. Quelques nucléotides A sont également toujours trouvés à cette position.

### 3.6.3. Détection de la mutation de référence

Chez 2 des patients, il a été possible de retrouver la mutation de référence détectée au diagnostic.

La première mutation, p.E746\_S752delinsV, située dans l'exon 19 du gène *EGFR*, n'a pu être confirmée que sur un extrait (kit « AmoyDx Circulating DNA »), amplifié avec le kit « EGFR 18-21 MASTR ». En effet, les autres extraits pour ce patient ont été amplifiés uniquement avec des amorces pour une partie de l'exon 20. Le résultat du séquençage a permis de démontrer la présence de la mutation à une fréquence allélique de 48.7% avec une couverture de 8'854x. Le laboratoire de Bâle a quant à lui rendu une fréquence de 38%.

La deuxième mutation de référence présente était située dans l'exon 21 (p.L858R) du gène *EGFR*, et les 4 ctDNA obtenus pour ce patient ont été amplifiés pour les exons 20 et 21 partiels. Cela a permis de retrouver ce variant selon les fréquences indiquées dans le tableau 10.

Tableau 10: détection de la mutation de référence p.L858R selon les différentes méthodes d'extraction en fonction de la couverture et de la fréquence de la base mutée en pourcents

	couverture	% muté
QIAamp MinElute ccfDNA	15'567	46.1
QIAamp Circulating Nucleic Acid	15'747	45.4
AmoyDx Circulating DNA	20'023	49.5
Maxwell RSC ccfDNA	21'126	46.3
résultat Bâle	-	59.0

La fréquence retrouvée pour la mutation à Sion est plus faible que celle donnée par Bâle, quel que soit le kit d'extraction utilisé.

### 3.6.4. Utilisation de 4 ctDNA synthétiques

L'amplification partielle des exons 20 et 21 pour les 4 ctDNA synthétiques a été utilisée afin de vérifier le seuil de détection et la fiabilité des VF (fréquence de variant) obtenus. Seules 4 des 10 positions de variants ont pu être analysées, les autres se trouvant dans des régions du gène non-amplifiées avec cette méthode. Les résultats de couverture ainsi que de VF pour les 4 échantillons sont donnés dans le tableau 11.

Tableau 11: résultats de couverture et VF obtenus avec les échantillons du set de référence « EGFR cfDNA » d'Horizon ; en rouge les valeurs hors seuil

	attendu : 5% [± 1.5]		attendu : 1% [± 0.4]		attendu : 0.1% [± 0.05]		attendu : 0% [<0.1%]	
	couverture	%	couverture	%	couverture	%	couverture	%
p.T790M	31'842	6.5	32'503	1.6	30'984	0.25	32'404	0.19
p.C797S	31'846	5.9	32'506	1.1	30'990	0.11	32'408	0.012
p.L858R	17'541	2.7	18'140	0.52	17'326	0.15	16'785	0.083
p.L861Q	17'541	6.3	18'142	1.5	17'329	0.15	16'875	0.059

La corrélation entre le résultat attendu et la fréquence retrouvée est bonne pour les mutations p.C797S et p.L861Q. Elle est légèrement sous-estimée pour la mutation p.L858R et globalement surestimée pour la mutation p.T790M.

## 4. Discussion

### 4.1. Extraction de l'ADN circulant

Selon les figures 7 et 8 et comme attendu la quantité d'ADN circulant est plus élevée chez les patients que chez les volontaires sains. Ceci s'est avéré particulièrement vrai pour deux des patients, qui sont probablement en récurrence. En effet, pendant la phase de rémission, le taux d'ADN circulant revient à la normale [29].

#### 4.1.1. Kit « AmoyDx Circulating DNA »

Les résultats de quantification pour le kit « AmoyDx Circulating DNA » montrent des valeurs bien plus faibles que pour les autres méthodes. Les bons résultats obtenus par quantification en qPCR montrent vraisemblablement une contamination de l'extrait avec de l'ADN leucocytaire. En effet, selon le test du gène ZFY, près de la moitié des échantillons extraits par cette méthode présentent

un résultat avant le 35<sup>ème</sup> cycle. Ce kit n'est donc pas adéquat pour l'utilisation de biopsies liquides dans le cadre des NSCLC.

#### 4.1.2. Kit « Maxwell RSC ccfDNA »

Concernant le kit « Maxwell RSC ccfDNA », le robot n'ayant été disponible que pendant 3 semaines sur les 8 durant lesquelles des prélèvements ont été effectués, cela a limité le nombre d'extractions réalisées. De plus, la quantification par qPCR suggère la présence d'un inhibiteur dans le tampon d'éluion (fournit avec le kit) des extraits obtenus avec ce kit. En effet, la quantité d'ADN utilisée, bien que semblable à celle du kit « AmoyDx Circulating DNA », n'a permis une amplification de l'exon 20 partiel du gène EGFR que 2 cycles plus tard. La taille des fragments (en moyenne à 160 pb) n'explique pas non plus ces résultats de cp tardifs.

En outre, toujours pour le kit « Maxwell RSC ccfDNA », la concentration d'ADN circulant dans les extraits était très faible, tant dans les dosages en ssDNA qu'en dsDNA, ce qui est dû au volume d'éluion minimal élevé du kit. La méthode devant permettre une concentration des extraits n'ayant pas fonctionné, ce kit n'est pas idéal pour des méthodes de séquençage à haut-débit. En effet, le volume d'ADN requis par échantillon étant généralement faible, une concentration de ctDNA faible induirait une limite de détection insuffisante pour un diagnostic de récurrence de cancer. Sachant que plus tôt la récurrence est détectée, plus efficace sera le traitement, un seuil de détection le plus bas possible est recommandé.

#### 4.1.3. Kits Qiagen

Les quantités de ctDNA obtenues se sont révélées les meilleures avec les deux kits de la société Qiagen.

Par dosage au Qubit de l'ADN double brin, le kit « QIAamp Circulating Nucleic Acid » (extraction sur membrane de silice) obtient les résultats les plus élevés (échantillons de 14 volontaires et 6 patients). Par contre, si l'on considère la quantification par qPCR, le kit « QIAamp MinElute ccfDNA » (extraction suite à une sélection par billes magnétiques) permet une détection des 96 pb de l'exon 20 du gène *EGFR* en moyenne plus précoce que les autres kits.

Il n'y a pas de différence significative entre les valeurs de quantification obtenues par « Qubit™ dsDNA HS Assay Kit » et celles calculées selon les cycles de qPCR en ce qui concerne les 2 kits de la société Qiagen (T-test supérieur à 0.05). Le coefficient de corrélation entre les 2 méthodes est supérieur à 99,4%, et cela indépendamment de la méthode d'extraction utilisée (extraction sur membrane de silice ou suite à une sélection par billes magnétiques).

De plus, pour le kit « QIAamp MinElute ccfDNA » l'absence d'amplification d'une partie du gène *ZFY* dans 95% des cas indique que la contamination de l'extraction avec de l'ADN cellulaire représente un phénomène marginal, bien que le test « QC Plex » ait permis d'amplifier, très faiblement, des fragments jusqu'à 600 pb.

Le kit « QIAamp MinElute ccfDNA » avec pourtant un rendement légèrement inférieur au kit « QIAamp Circulating Nucleic Acid » montre globalement de meilleurs résultats pour une utilisation de l'ADN circulant en NGS. Ceci peut provenir de la différence dans la méthode d'extraction, la sélection par des billes étant moins agressive pour l'ADN que le passage sur une colonne. Des tests ultérieurs avec ces 2 kits pourraient permettre de confirmer les observations réalisées lors de ce travail.

#### 4.1.4. Méthode de quantification

Concernant la méthode de dosage, le kit dsDNA HS utilisé avec le Qubit étant plus simple et moins coûteux, il est à privilégier. La différence observée entre les valeurs d'ADN simple et double brin est attendue de par les différents mécanismes de relargage d'ADN tumoral [30]. L'ADN simple brin étant moins stable, il est préférable d'utiliser la valeur obtenue avec l'ADN double brin pour les expériences ultérieures d'amplification.

Comme décrit également par Hussing et al [31] le Qubit obtient des valeurs correctes, et surestime peu l'ADN présent, contrairement au Fragment Analyzer. Le kit utilisé sur le Fragment Analyzer ayant un coefficient de corrélation de seulement 70.7% lors du test de reproductibilité, il ne permet pas de quantifier avec fiabilité l'ADN circulant. L'utilisation d'un autre kit de détection sur le Fragment Analyzer (éventuellement le DNF-474 « High Sensitivity NGS Fragment Analysis kit », avec gamme de concentration allant de 0.05 à 5 ng/μL) pourrait améliorer la détection et plus particulièrement permettre d'obtenir des valeurs au-dessus de la LLOD. Pour des questions de temps et de coûts, cette option n'a pas été envisagée lors de ce travail.

## 4.2. Séquençage à haut-débit

Bien que n'étant pas le but premier de ce travail, le séquençage à haut-débit a été nécessaire afin d'assurer que l'ADN extrait soit d'une part de qualité suffisante pour une utilisation ultérieure par cette méthode et soit d'autre part effectivement de l'ADN tumoral et non leucocytaire. Ce dernier point a pu être validé par la détection de mutations somatiques.

### 4.2.1. Amplification des exons 18 à 21 avec le kit Agilent

Bien que l'amplification des 4 exons se fasse conjointement, le kit « EGFR 18-21 MASTR » ne donne pas une couverture identique pour chaque exon, un facteur 3 ayant pu être retrouvé. La couverture de l'exon 20 est généralement moins bonne que celle notamment de l'exon 19, alors qu'il nécessite un seuil de détection élevé pour la mutation p.T790M. Néanmoins, les couvertures obtenues seraient suffisantes pour détecter les mutations présentes dans ces 4 exons à un taux de 0.5%, par exemple pour la recherche de la mutation de référence.

### 4.2.2. Détection de la mutation de résistance (p.T790M)

Les résultats obtenus à la position c.2369C>T dans l'exon 20 du gène *EGFR* pour les différents extraits chez les patients sont concordants avec ceux de Bâle. Cependant, le taux de nucléotides T à cette position dans les extraits sauvages indique une région bruitée, et par conséquent une limite de détection plus élevée que désirée. Le bruit observé est légèrement plus faible pour les échantillons de patients négatifs extraits avec le kit « QIAamp MinElute ccfDNA », néanmoins, étant donné le faible nombre d'échantillons, cette observation nécessite une confirmation ultérieure.

### 4.2.3. Détection de la mutation de référence

Les valeurs obtenues à Bâle et à Sion sont là aussi comparables. Cependant, pour 2 mutations détectées la fréquence allélique est plus élevée à Bâle et pour la 3<sup>ème</sup> c'est l'inverse. La quantité de plasma extrait, la méthode d'extraction ainsi que technique de détection, peuvent influencer ces valeurs. Le laboratoire de Bâle extrait le ctDNA de 10 mL de plasma, la méthode d'extraction utilisée n'est pas communiquée et l'amplification est réalisée avec le kit « Oncomine™ Lung cfDNA Assay » de ThermoFisher sur S5 ou PGM. Afin d'éviter le prélèvement d'un volume de sang trop

important, il a été décidé de ne pas ponctionner plus de 3 tubes chez les patients en plus des 2 destinés aux analyses effectuées à Bâle. Ceci explique le volume restreint de plasma pour les différents tests d'extraction et éventuellement les légères différences observées.

#### 4.2.4. Utilisation de 4 ctDNA synthétiques

Les résultats obtenus avec l'ADN synthétique pour 2 des mutations sont adéquates. Par contre, il est intéressant de noter, que comme chez les patients, la fréquence observée de la mutation de référence p.L858R a tendance à être plus faible qu'attendu. De l'autre côté, pour la mutation de résistance p.T790M dans l'exon 20 du gène *EGFR*, elle est plus élevée qu'attendu, particulièrement aux faibles concentrations. Ceci provient certainement, au moins en partie, du bruit de fond inhérent à la séquence, également observée lors de l'amplification de cette même région avec les extraits des patients. La VF de 0.5% du NGS retrouvée dans la littérature est due à ce phénomène. Avec déjà 30'000 reads à cette position, il est manifeste qu'une augmentation de la couverture ne serait pas suffisante pour améliorer la LLOD, car elle ne prendrait pas en compte les erreurs dues à la polymérase, tant lors de la préparation des librairies que de la détection sur le séquenceur. Seul un ajout d'identifiants moléculaires uniques pourrait permettre de descendre le seuil de détection à 0.1% ou de s'en approcher.

## 5. Conclusions et perspectives

Le kit « QIAamp MinElute ccfDNA », avec un coût par échantillon le moins élevé [32], une durée d'extraction courte (moins d'une heure), nécessitant peu de matériel additionnel et surtout présentant des résultats très bons tant au niveau qualité que rendement d'ADN circulant, est donc recommandé. De plus, concernant la quantification de l'ADN extrait, la méthode de dosage préférentielle est le kit « Qubit™ dsDNA HS Assay Kit ».

Le séquençage à haut-débit nécessite l'utilisation du *molecular barcoding* afin d'abaisser le seuil de détection au niveau désiré de 0.1%.

Dans la perspective de mettre en place la détection de la mutation de résistance p.T790M de l'exon 20 du gène *EGFR*, une comparaison des 4 sociétés sur le marché, proposant des kits personnalisés avec UMI, a été établie. Les critères retenus ont été : la possibilité de séquencer parallèlement des librairies obtenues avec d'autres fournisseurs, la préparation des librairies, le coût par échantillon, la validation ou non du kit par le fournisseur et la méthode d'analyse des résultats.

Le kit de la société Swift Biosciences, avec une durée de préparation donnée de 2h, un volume de 10 µL d'ADN circulant utilisable, une pré-validation effectuée par le fournisseur, un coût par échantillon de moins de 100 CHF, une compatibilité avec d'autres kits de préparation et permettant l'amplification des 4 exons d'intérêt du gène *EGFR*, semble répondre le mieux aux critères évalués. Il va donc être commandé et testé principalement pour la LLOD, avec l'utilisation notamment des ADN synthétiques d'Horizon.

## 6. Remerciements

Je tiens à remercier toutes les personnes m'ayant permis de mener à bien ce travail, plus particulièrement :

- le Professeur Sergio Schmid pour sa disponibilité
- le Dr Thomas von Känel ainsi que C. Rey Berthod et C. Forré, du laboratoire de génétique de l'ICH, pour leur soutien sans faille
- le Dr Sandro Anchisi ainsi que tous les membres du service d'oncologie du RSV pour leur agréable collaboration
- Yann Antille et nos enfants (Coralie et Delhia) pour leur patience
- S. Guenat et F. Secretan pour leurs conseils avisés lors de la relecture

## 7. Bibliographie

- [1] « Le cancer en Suisse: les chiffres », *Ligue Suisse contre le cancer*, p. 13, 2017.
- [2] L. A. Pikor, V. R. Ramnarine, S. Lam, et W. L. Lam, « Genetic alterations defining NSCLC subtypes and their therapeutic implications », *Lung Cancer*, vol. 82, n° 2, p. 179-189, nov. 2013.
- [3] G. Bethune, D. Bethune, N. Ridgway, et Z. Xu, « Epidermal growth factor receptor (EGFR) in lung cancer: an overview and update », p. 4.
- [4] P. Perez-Moreno, E. Brambilla, R. Thomas, et J.-C. Soria, « Squamous cell carcinoma of the lung: molecular subtypes and therapeutic opportunities », *Clinical Cancer Research*, vol. 18, n° 9, p. 2443-2451, mai 2012.
- [5] X. Song et Z. Wang, « Clinical efficacy evaluation of tyrosine kinase inhibitors for non-adenocarcinoma lung cancer patients harboring EGFR-sensitizing mutations », *Onco Targets Ther*, vol. 10, p. 3119-3122, juin 2017.
- [6] D. B. Peckys et N. de Jonge, « Liquid scanning transmission electron microscopy: imaging protein complexes in their native environment in whole eukaryotic cells », *Microscopy and Microanalysis*, vol. 20, n° 02, p. 346-365, avr. 2014.
- [7] R. S. Herbst et P. A. Bunn, « Targeting the epidermal growth factor receptor in non-small cell lung cancer », *Clin Cancer Res*, vol. 9, n° 16, p. 5813-5824, déc. 2003.
- [8] E. L. Stewart, S. Z. Tan, G. Liu, et M.-S. Tsao, « Known and putative mechanisms of resistance to EGFR targeted therapies in NSCLC patients with EGFR mutations—A review », p. 16, 2014.
- [9] M. Santarpia *et al.*, « Osimertinib in the treatment of non-small-cell lung cancer: design, development and place in therapy », *Lung Cancer: Targets and Therapy*, vol. Volume 8, p. 109-125, août 2017.
- [10] T. S. Mok *et al.*, « Osimertinib or Platinum–Pemetrexed in EGFR T790M–positive lung cancer », *New England Journal of Medicine*, vol. 376, n° 7, p. 629-640, févr. 2017.
- [11] I. B. Hench, J. Hench, et M. Tolnay, « Liquid biopsy in clinical management of breast, lung, and colorectal cancer », *Frontiers in Medicine*, vol. 5, janv. 2018.
- [12] « tagrisso-product-monograph-fre.pdf ». .
- [13] N. C. Dracopoli et M. S. Boguski, « The evolution of oncology companion diagnostics from signal transduction to immuno-oncology », *Trends in Pharmacological Sciences*, vol. 38, n° 1, p. 41-54, janv. 2017.
- [14] D. A. E. Cross *et al.*, « AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer », *Cancer Discovery*, vol. 4, n° 9, p. 1046-1061, sept. 2014.
- [15] P. Mandel et P. Métais, « Les acides nucléiques du plasma sanguin chez l'homme ». [En ligne]. Disponible sur: <https://www.ncbi.nlm.nih.gov/pubmed/?term=18875018>. [Consulté le: 16-avr-2018].
- [16] S. A. Leon, B. Shapiro, D. M. Sklaroff, et M. J. Yaros, « Free DNA in the serum of cancer patients and the effect of therapy », p. 6, 1977.
- [17] L. Fernandez-Cuesta *et al.*, « Identification of circulating tumor DNA for the early detection of small-cell lung cancer », *EBioMedicine*, vol. 10, p. 117-123, août 2016.
- [18] H. Schwarzenbach, D. S. B. Hoon, et K. Pantel, « Cell-free nucleic acids as biomarkers in cancer patients », *Nature Reviews Cancer*, vol. 11, n° 6, p. 426-437, juin 2011.
- [19] X. Li, C. Zhou, X. Li, et C. Zhou, « Comparison of cross-platform technologies for EGFR T790M testing in patients with non-small cell lung cancer », *Oncotarget*, vol. 8, n° 59, p. 100801-100818, juill. 2017.

- [20] C. Rolfo *et al.*, « Liquid biopsy for advanced non-small cell lung cancer (NSCLC): a statement paper from the IASLC », *Journal of Thoracic Oncology*, juin 2018.
- [21] A. Ståhlberg, P. M. Krzyzanowski, M. Egyud, S. Filges, L. Stein, et T. E. Godfrey, « Simple multiplexed PCR-based barcoding of DNA for ultrasensitive mutation detection by next-generation sequencing », *Nature Protocols*, vol. 12, n° 4, p. 664-682, mars 2017.
- [22] E. Crowley, F. Di Nicolantonio, F. Loupakis, et A. Bardelli, « Liquid biopsy: monitoring cancer-genetics in the blood », *Nature Reviews Clinical Oncology*, vol. 10, n° 8, p. 472-484, août 2013.
- [23] G. Chen, S. Mosier, C. D. Gocke, M.-T. Lin, et J. R. Eshleman, « Cytosine deamination is a major cause of baseline noise in next generation sequencing », *Mol Diagn Ther*, vol. 18, n° 5, p. 587-593, oct. 2014.
- [24] Y. Seki *et al.*, « Picoliter-droplet digital polymerase chain reaction-based analysis of cell-free plasma DNA to assess *EGFR* mutations in lung adenocarcinoma that confer resistance to tyrosine-kinase inhibitors », *The Oncologist*, vol. 21, n° 2, p. 156-164, févr. 2016.
- [25] M. W. Pfaffl, « A new mathematical model for relative quantification in real-time RT-PCR », *Nucleic Acids Res*, vol. 29, n° 9, p. e45, mai 2001.
- [26] N. V. Multiplicom, « Instructions for use part I - Multiplex PCR with amplicon specific primers: EGFR 18-21 MASTRTM assay ». 2013.
- [27] A. Untergasser *et al.*, « Primer3—new capabilities and interfaces », *Nucleic Acids Res*, vol. 40, n° 15, p. e115, août 2012.
- [28] J. T. Robinson *et al.*, « Integrative genomics viewer », *Nature Biotechnology*, vol. 29, p. 24, janv. 2011.
- [29] Z. Wei, W. Wang, Z. Shu, X. Zhou, et Y. Zhang, « Correlation between circulating tumor DNA levels and response to tyrosine kinase inhibitors (TKI) treatment in non-small cell lung cancer », *Med Sci Monit*, vol. 23, p. 3627-3634, juill. 2017.
- [30] A. R. Thierry, S. El Messaoudi, P. B. Gahan, P. Anker, et M. Stroun, « Origins, structures, and functions of circulating DNA in oncology », *Cancer Metastasis Rev*, vol. 35, n° 3, p. 347-376, 2016.
- [31] C. Hussing, M. L. Kampmann, H. S. Mogensen, C. Børsting, et N. Morling, « Comparison of techniques for quantification of next-generation sequencing libraries », *Forensic Science International: Genetics Supplement Series*, vol. 5, p. e276-e278, déc. 2015.
- [32] S. Arcioni, « Mise au point d'une technique d'extraction de l'ADN circulant chez les patients atteints de NSCLC en vue de la détection de la mutation *EGFR* T790M. », HES-SO Valais, Sion, Projet 3, avr. 2018.

## 8. Annexes

### QIAamp MinElute ccfDNA protocole

Avant de commencer, ajouter 12 mL d'isopropanol 100% au tampon ACB et 30 mL d'éthanol 96-100% au tampon ACW2. Bien mélanger après l'ajout.

Préchauffer un agitateur pour microtubes à température ambiante pour l'étape 3 et à 56°C pour l'étape 9.

1. Mélanger dans un tube 15 mL les différents composants selon le tableau 1 et incuber 10 minutes à température ambiante sur un mélangeur à vitesse lente.

Tableau 12: composants à mélanger

Plasma (mL)	Billes magnétiques en suspension (µL)	Protéinase K (µL)	<i>Bead Binding Buffer</i> (µL)
4	120	220	600
5	150	275	750
6	180	330	900

2. Mettre le tube sur un support magnétique, attendre au moins 1 minute que la solution soit claire. Eliminer le surnageant.
3. Enlever le tube du support et ajouter 200 µL de ***Bead Elution Buffer*** sur le culot de billes. Vortexer pour remettre en suspension puis éliminer les billes résiduelles fixées à la paroi par pipetage. Transférer le tout dans un *Bead Elution Tube*, et incuber 5 minutes sur un agitateur pour microtubes à température ambiante et 300 rpm.
4. Mettre le microtube sur un support magnétique, attendre au moins 1 minute que la solution soit claire.
5. Transférer le surnageant dans un nouveau *Bead Elution Tube*, et y ajouter 300 µL de **tampon ACB**. Mélanger sur le vortex, et centrifuger rapidement pour collecter les gouttes dans le bouchon. Eliminer le culot de billes.
6. Pipeter le mélange précédant dans une colonne QIAamp UCP MinElute, et centrifuger 1 minute à 6'000 g. Placer ensuite la colonne sur un tube de collecte, éliminer le *flow-through*.
7. Ajouter 500 µL de **tampon ACW2** dans la colonne, centrifuger à nouveau à 6'000 g pendant 1 minute. Transférer la colonne dans un tube de collecte propre, éliminer le *flow-through*.
8. Centrifuger à vitesse maximale (20'000 g) pendant 3 minutes.
9. Placer la colonne sur un tube d'éluion 1.5 mL fourni. Ouvrir le bouchon et incuber 3 minutes à 56°C dans un agitateur pour microtubes, afin de sécher complètement la membrane.
10. Ajouter 20-80 µL d'**eau ultrapure** (fournie) bien au centre de la membrane. Fermer le bouchon et incuber 1 minute à température ambiante. Centrifuger à vitesse maximale (20'000 g) durant 1 minute afin d'éluier l'ADN. Eventuellement remettre l'éluant sur la colonne et répéter la centrifugation afin d'augmenter le rendement.

## QIAamp DSP Circulating NA Kit protocole

Avant de commencer, préchauffer un bain-marie à 60°C pour des tubes 50 mL. Préchauffer également un bloc chauffant pour microtubes à 56°C.

Préparer les tampons ACB, ACW1 et ACW2. Préparer également le volume nécessaire de mélange *carrier RNA* – tampon ACL selon le tableau 1.

**Tableau 13: Volumes de tampon ACL et de *carrier RNA* pour des échantillons de 4 ou 5 mL de plasma**

Nbre d'échantillons	Tampon ACL (mL)		<i>Carrier RNA</i> dans du tampon AVE (µL)
	4 mL plasma	5 mL plasma	
1	3.5	4.4	5.6
2	7.0	8.8	11.3
3	10.6	13.2	16.9
4	14.1	17.6	22.5
5	17.6	22.0	28.1

1. Mettre 400 µL (4 mL de plasma) ou 500 µL (5 mL de plasma) de **Protéinase K** dans un tube 50 mL.
2. Ajouter le volume de **plasma** prévu.
3. Ajouter également 3.2 mL, respectivement 4.0 mL du **mélange tampon ACL – *carrier RNA*** préparé précédemment. Fermer le bouchon et mélanger par des pulsations au vortex pendant 30 secondes. S'assurer de la présence d'un tourbillon dans la solution. Un mélange homogène est indispensable à une bonne lyse.
4. Incuber immédiatement à 60°C pendant 30 minutes.
5. Remettre les tubes sur un statif et ôter les bouchons.
6. Ajouter 7.2 mL ou 9 mL (selon volume de plasma) de **tampon ACB**. Fermer le bouchon et mélanger à nouveau par des pulsations au vortex durant 15 à 30 secondes.
7. Incuber le mélange 5 minutes sur glace.
8. Insérer la colonne QIAamp Mini dans le connecteur sur le QIAvac 24 Plus. Insérer fermement l'extenseur de colonne 20 mL dans la colonne ouverte. Garder le tube de lavage pour l'étape 13.
9. Transférer délicatement le lysat de l'étape 7 dans l'extenseur de colonne. Allumer la pompe à vide. Quand tout le lysat est passé, éteindre la pompe et relâcher la pression à 0 mbar. Ôter et jeter l'extenseur. Le passage à travers la colonne peut prendre jusqu'à 15 minutes pour le plus grand volume.
10. Mettre 600 µL de **tampon ACW1** dans la colonne. Laisser le capuchon ouvert et enclencher la pompe. Quand tout le tampon est passé, relâcher la pression.
11. Ajouter 750 µL de **tampon ACW2** dans la colonne. Laisser le capuchon ouvert et enclencher la pompe. Quand tout le tampon est passé, relâcher la pression.
12. Ajouter 750 µL d'**éthanol 96-100%** dans la colonne. Laisser le capuchon ouvert et enclencher la pompe. Quand tout le tampon est passé, relâcher la pression.
13. Fermer le capuchon de la colonne et l'ôter de la pompe à vide. Jeter le connecteur. Mettre la colonne dans un tube de lavage propre et centrifuger à vitesse maximale (20'000 g) durant 3 minutes.
14. Mettre la colonne dans un tube de lavage propre, ouvrir le capuchon et incuber l'ensemble à 56°C pendant 10 minutes pour sécher complètement la membrane.
15. Transférer la colonne dans un tube 1.5 mL pour élution fourni. Ajouter 20-150 µL de **tampon AVE** bien au centre de la membrane. Fermer le capuchon et incuber 3 minutes à température ambiante.
16. Centrifuger à vitesse maximale (20'000 g) durant 1 minute afin d'éluer l'ADN. Eventuellement remettre l'éluant sur la colonne et répéter la centrifugation afin d'augmenter le rendement.

## AmoyDx Circulating DNA Kit protocole

Avant de commencer, préchauffer un bain-marie à 60°C pour des tubes 10 mL. Préchauffer également un bloc chauffant pour microtubes à 56°C.

Ajouter 19 mL d'éthanol 96-100% dans le tampon CW1 ainsi que dans le tampon CW2. Bien mélanger après l'ajout. Mettre de l'isopropanol 100% au congélateur (3.5 mL par échantillon).

1. Transférer 4 mL de **plasma** dans un tube 10 mL fourni.
2. Ajouter 2.4 mL de **tampon CDL** et 210 µL de **Digest Solution** dans le tube. Fermer le bouchon et mélanger au vortex plus de 10 secondes.
3. Incuber 15 minutes à dans le bain-marie à 60°C.
4. Mettre le tube sur de la glace 5 minutes pour ramener la solution à température ambiante.
5. Ajouter 400 µL de **DNA Tracer**, mélanger en pipetant gentiment.
6. Ajouter ensuite 3.3 mL d'**isopropanol** froid, fermer le bouchon et mélanger en le retournant plus de 10 fois.
7. Centrifuger à 10'000 g pendant 5 minutes.
8. Bien éliminer tout le surnageant à l'aide d'une pipette.
9. Ajouter 470 µL de **tampon CDB** au précipité. Puis ajouter directement sous la surface du liquide 40 µL de **tampon CDD**.
10. Vortexer le tube 5 secondes et incuber 10 minutes à 60°C dans le bain-marie. Ramener ensuite à température ambiante et ajouter 250 µL d'**éthanol 96-100%**, puis dissoudre le précipité par 15-20 pipetages.
11. Transférer le lysat dans la colonne Micro NA Spin, fermer le capuchon et centrifuger à 10'000 g pendant 30 secondes.
12. Eliminer le *flow-through* dans le tube collecteur.
13. Ajouter 700 µL de **tampon CW1** dans la colonne et centrifuger 30 secondes à 10'000 g.
14. Eliminer le *flow-through* dans le tube collecteur.
15. Ajouter 700 µL de **tampon CW2** dans la colonne et centrifuger 30 secondes à 10'000 g.
16. Eliminer le tube collecteur avec le *flow-through*.
17. Mettre la colonne dans un tube collecteur propre. Centrifuger 1 minute à 13'000 g.
18. Eliminer le tube collecteur avec le *flow-through*.
19. Placer la colonne ouverte dans un tube 1.5 mL sans bouchon et incuber 10 minutes à 56°C, puis jeter le tube 1.5 mL.
20. Mettre la colonne dans un tube à centrifuger 1.5 mL.
21. Mettre 30-50 µL de **tampon CDE** au centre de la membrane. Incuber 5 minutes à température ambiante, puis centrifuger 1 minute à 13'000 g afin d'éluer l'ADN. Eventuellement remettre l'éluant sur la colonne et répéter la centrifugation afin d'augmenter le rendement.

## Promega Maxwell RSC Circulating DNA Purification Kit protocole

### Préparation de l'échantillon :

1. Mettre 4 mL de **plasma** dans un tube 15 mL. Ajouter le même volume de **Binding Buffer**.
2. Agiter la bouteille contenant la résine magnétique jusqu'à complète resuspension.
3. Ajouter 140 µL de **résine** dans le tube 15 mL.
4. Incuber 45 minutes sur un agitateur-rotateur. La résine doit rester en suspension durant toute l'incubation.
5. Centrifuger le tube 2 minutes à 1'000 g.
6. Placer le tube sur un support magnétique et éliminer totalement le surnageant.

### Préparation du robot :

Allumer le robot, puis la tablette, et lancer le programme adéquat. Placer le nombre de cartouche nécessaire sur le support, avec le puits 1 (le plus grand) le plus loin du tube d'élution. Bien fixer la cartouche. Ôter le film protecteur de la cartouche. Préparer le tube d'élution avec 75 µL de tampon d'élution et le placer sur le rack à la suite de la cartouche.

7. Récupérer le **contenu** du puits 1 et le transférer dans le tube 15 mL contenant la résine.
8. Remettre en suspension la résine à l'aide de la pipette afin d'éviter toute adhésion sur le haut du tube.
9. Transférer la résine et le liquide à nouveau dans le puits 1 de la cartouche.
10. Placer un *plunger* dans le puits 8 de la cartouche.
11. Remettre le support dans le robot et lancer le programme. Celui-ci dure environ 30 minutes.
12. A la fin du *run*, vérifier que le *plunger* est à nouveau dans le puits 8. Enlever le support, récupérer le tube contenant l'ADN extrait, et jeter la cartouche.
13. Lancer le programme de nettoyage de l'appareil.
14. Pour éteindre : quitter le programme, éteindre l'appareil à l'arrière, puis la tablette.