

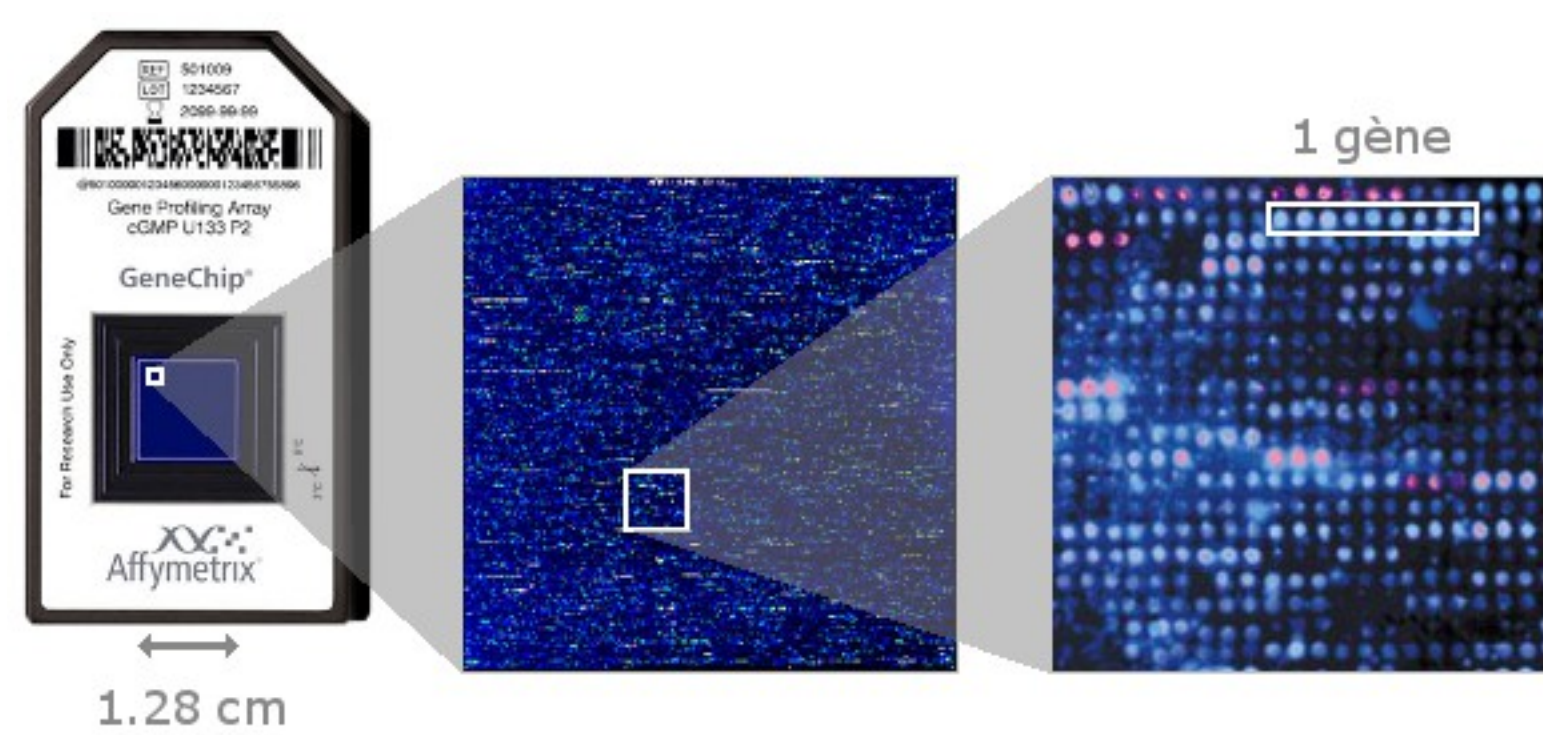
Les Lymphomes Diffus à Grandes Cellules de phénotype B (LDGCB) sont des tumeurs issues de la prolifération de lymphocytes B, l'un des principaux types de cellules immunitaires circulant dans le sang. Ces cancers constituent la majorité des lymphomes dits « Non-Hodgkiniens » (16 000 cas et 11 000 décès en France en 2008), et sont actuellement pris en charge par une chimiothérapie (CHOP ou ACVBP) couplée à un anticorps monoclonal (Rituximab). Grâce à ce traitement généraliste, environ 60 % des patients sont guéris de façon durable à l'heure actuelle (EFS à 5 ans). Afin d'améliorer la prise en charge de ces patients, il est nécessaire de mettre au point des traitements adaptés à chacun des nombreux sous-types de LDGCB, et donc de mieux comprendre les mécanismes impliqués.

C'est dans cette optique que nous allons caractériser une série de 10 tumeurs réfractaires au traitement par des méthodes à « haut débit », et en intégrer les résultats par des méthodes bio-informatiques :

- quantification de l'expression de l'ensemble des gènes connus (puces Affymetrix)
- recherche de réarrangements chromosomiques récurrents (puces CGH)
- recherche de mutations acquises récurrentes (séquençage de nouvelle génération)

Cette étude s'inscrit dans le cadre du projet GHEDI (*deciphering the Genetic HETerogeneity of Diffuse large B-cell lymphoma in the rituximab era*) mené par le GELA (Groupe d'Étude des Lymphomes de l'Adulte), projet à dimension nationale qui vise à mutualiser les moyens et patients pour mettre en évidence de nouvelles cibles thérapeutiques dans ces lymphomes.

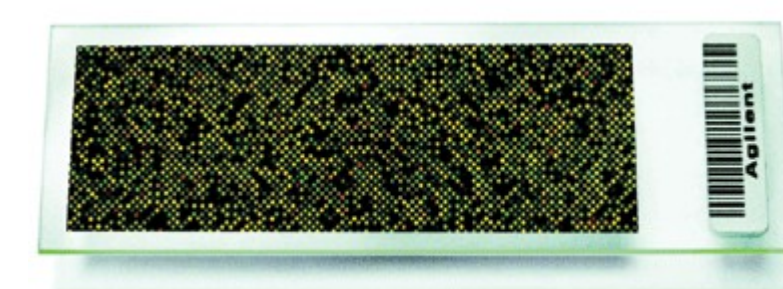
## Transcriptome



Les 10 lymphomes séquencés font partie d'une série de 225 LDGCB au sein desquels les niveaux d'expression de plus de 21 000 gènes ont été mesurés par des puces Affymetrix U133+2. L'intégration de ces données avec les résultats de CGH et de séquençage permettra de se concentrer sur les anomalies génétiques ayant un impact direct sur l'expression des gènes impliqués.

Ces données sont également précieuses pour déterminer les classes GCB (Germinal Center B-cell like) et ABC (Activated B-Cell like), deux sous-types de LDGCB connus pour leurs profils d'expression et pronostiques très distincts.

## Réarrangements chromosomiques



Conjointement aux puces transcriptomiques, le projet GHEDI incluait pour ces 10 patients l'analyse des réarrangements chromosomiques par un autre type de puces à ADN : des puces CGH (*Comparative Genomic Hybridization*). Pour chacune des 180 000 sondes de ces puces réparties le long du génome humain, une hybridation compétitive entre l'ADN tumoral et un ADN de référence permet de mettre en évidence les régions dupliquées ou perdues dans l'ADN tumoral.

L'étude de ces puces, qui a fait l'objet de mes deux années d'alternance en master 2, a d'ores et déjà permis de mettre en évidence de nouveaux gènes potentiellement impliqués dans la pathogénicité des lymphomes.

## Séquençage exomique

La recherche de mutations récurrentes dans la série de 10 LDGCB constituera le cœur de l'étude. Elle sera réalisée par une technique de séquençage dite de « nouvelle génération », qui permet de séquencer en quelques jours les 30 millions de bases (A, C, G ou T) d'un exome humain (partie de l'ADN codant effectivement des gènes, soit un peu moins de 1% du génome entier). A titre de comparaison, le séquençage du génome humain débuté par l'*Human Genome Project* en 1990 a pris 10 ans pour faire ce que cette technique est capable de réaliser en quelques jours.

Cette expérience se décompose en plusieurs étapes :

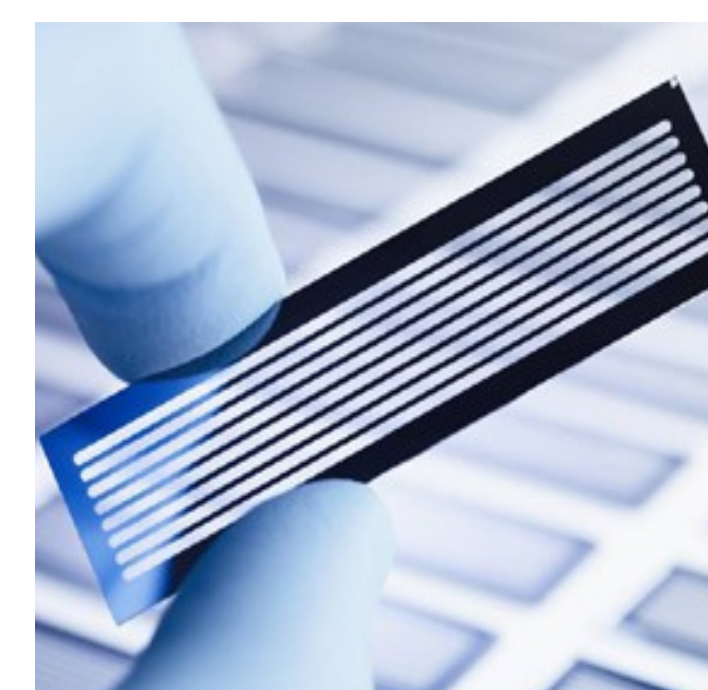
- l'ADN tumoral est extrait, fragmenté et modifié pour les besoins de la technique.
- chaque fragment est fixé et dupliqué localement en un point précis de la *flow cell* appelé *cluster*.
- les 4 bases de l'ADN sont testées successivement sur l'ensemble de la *flow cell* : les fragments dont la séquence débute par le complémentaire de la base testée émettent un signal fluorescent qui est localisé avec précision et enregistré.
- l'opération est répétée pour chaque position des différents fragments : 2<sup>ème</sup> base, 3<sup>ème</sup> base ... jusqu'à une centaine de bases.
- les séquences des différents fragments sont reconstruites en recoupant les signaux enregistrés.

La partie la plus délicate consiste ensuite à localiser les milliards de fragments séquencés dans les gènes d'où ils proviennent, dans la mesure où il existe plusieurs milliers de gènes et que les mutations que l'on recherche font que les séquences obtenues diffèrent des séquences connues. Cette étape de *mapping* repose sur des outils bio-informatiques encore expérimentaux, dont l'utilisation et la mise au point constituera la partie centrale de mon travail. Les mutations récurrentes détectées seront ensuite recoupées avec l'expression des gènes mutés afin de quantifier leur impact sur le fonctionnement de la tumeur.



### Genome Analyzer IIx

Appareil réalisant la phase de séquençage proprement dite.



### Flow cell

Jusqu'à 8 exomes peuvent être analysés sur la même *flow cell*.

Les 3 années de thèse dédiées à ce projet (2012-2015) sont financées par la Région Haute-Normandie. Ce travail est supervisé par le Pr Fabrice Jardin (UMR INSERM 918) et le Pr Thierry Lecroq (EA 4108).