**Introduction**

1. **Contexte**

Le diagnostic et le pronostic d’une lésion cancéreuse reposent sur des critères cliniques et anatomopathologiques. Pour le diagnostic, dans les cas difficiles, le laboratoire dispose de technologies de screening moléculaire, permettant l’identification d’altérations moléculaires spécifiques ou caractéristiques du diagnostic. Ces données moléculaires peuvent permettre également d’orienter les choix thérapeutiques, en particulier vers des thérapies ciblées.

L’unité de Pathologie moléculaire est dotée d’un panel de technologies de screening de biologie moléculaire et de cytogénétique, permettant l'identification d'altérations génétiques spécifiques contribuant au diagnostic, au pronostic ou l'identification de cibles thérapeutiques. Grâce à son expertise et l'accès à des technologies innovantes (figure 1), l'unité participe activement aux activités de routine et de recherche.

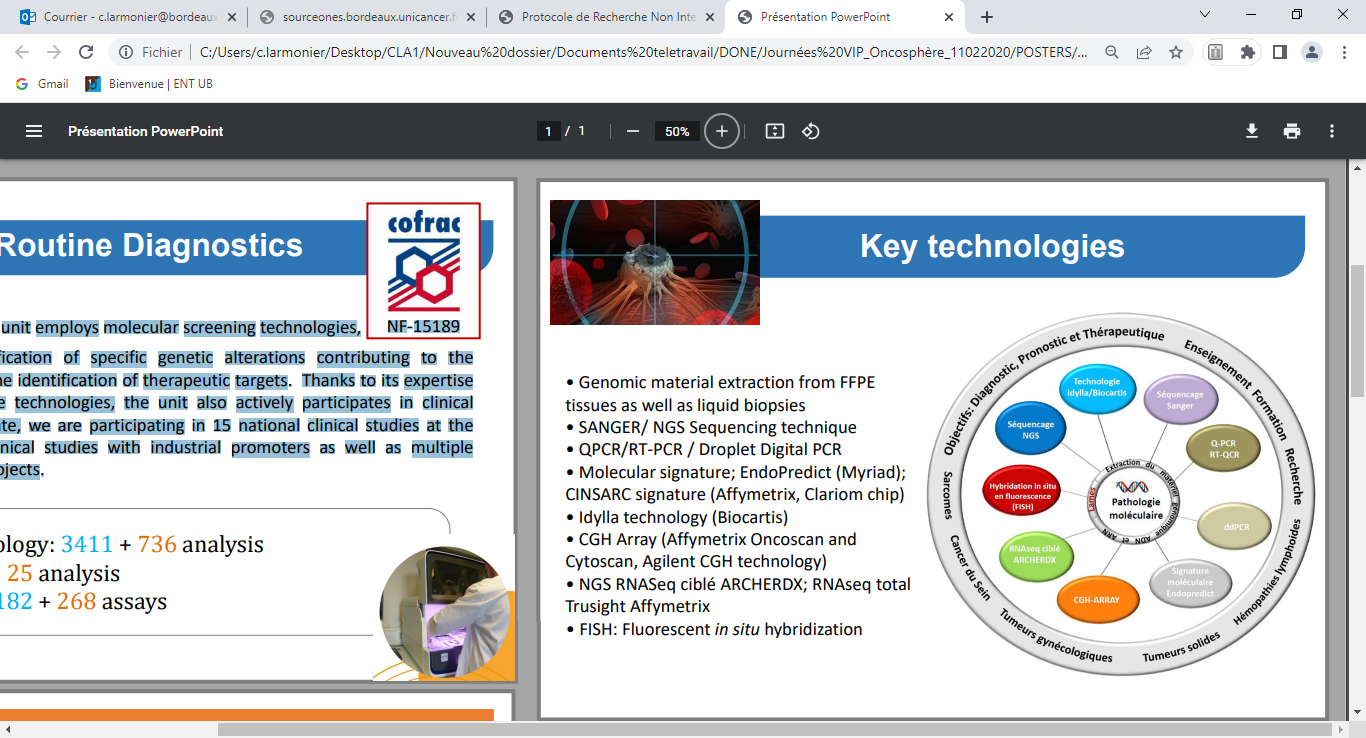


Figure 1 : Représentation schématique des technologies présente dans l’unité de pathologie moléculaire , au sein de département de biopathologie de l’institut bergonié (source : support de communication / institut bergonié)

1. **Stratégie :**

* Les génomes des cellules cancéreuses contiennent souvent des altérations chromosomiques complexes telles que des variations du nombre de copies = CNV (pertes, gains, points de cassures) qui peuvent conduire à la formation et à la progression des tumeurs. L’étude de ces anomalies chromosomiques et de l'instabilité génomique fournit des informations précieuses sur la biologie de ces tumeurs, leur évolution et leur résistance aux traitements.

Parmi les technologies citées en figure 1, l’analyse globale du génome à la recherche d’anomalies de nombre de copies utilise la technique de CGH-array (array-based Comparative Genomic Hybridization). Plus précisément, deux méthodologies sont utilisées : SurePrint G3 d’Agilent et OncoScan CNV d’Affymetrix.

Le principe repose sur la capacité de 2 séquences d’acides nucléiques, complémentaires et antiparallèles, à s’apparier entre elles. L’ADN extrait du tissu à analyser est fragmenté à l’aide d’enzymes de restriction. Les fragments sont marqués à l’aide d’un traceur (biotine ou fluorochrome) puis hybridés sur un support. Les signaux quantifiés par un scanner, sont comparés à ceux d’un ADN normal, utilisé comme référence. Les signaux sont analysés après calcul des Log2 ratios relatifs (LRR). Pour mettre en évidence ces relations et identifier des anomalies de continuité, les LRR sont ordonnés selon leur position génomique, puis analysés à l‘aide d’algorithmes de segmentation. Leur principe est d’identifier des points de cassure témoignant de changements de niveau dans les signaux. Ces points de cassure délimitent alors des segments à l’intérieur desquels le signal peut être résumé par la moyenne des LRR des sondes qu’ils contiennent.

Les régions anormales sont qualifiées de gagnées, ou amplifiées, lorsque leur nombre de copies est supérieur à 2 (nombre de copies attendu dans un ADN normal), et « délétées » lorsqu’il est inférieur à 2. Après calcul des ratios sonde-à sonde, Log transformation et segmentation, les régions d’intérêt seront donc celles présentant un LRR > 0 pour les régions en gain, ou < 0 pour celles en perte (Frederic Commo. Analyse génomique en médecine de précision : Optimisations et outils de visualisation. Bio-Informatique, Biologie Systémique [q-bio.QM]. Université Paris-Saclay, 2015.).

A titre d’exemple, la plateforme Affymetrix propose une hybridation simple : seul l’ADN analysé est couplé à un fluorochrome, puis hybridé. L’analyse se fera par comparaison à un ADN de contrôle hybridé relativement à une base d’ADN virtuelle. Basée sur la technologie de la sonde d'inversion moléculaire (MIP) (7), **OncoScan CNV Assay** fournit une couverture du génome entier avec une résolution accrue dans tous types de cancers.

(Chekaluk Y, Wu CL, Rosenberg J, Riester M, Dai Q, Lin S, *et al*. Identification of nine genomic regions of amplification in urothelial carcinoma, correlation with stage, and potential prognostic and therapeutic value. PLoS One. 2013;8(4):e60927. / Xie T, D’ Ario G, Lamb JR, Martin E, Wang K, Tejpar S, *et al*. A comprehensive characterization of genome-wide copy number aberrations in colorectal cancer reveals novel oncogenes and patterns of alterations. PLoS One. 2012;7(7):e42001. / Skirnisdottir I, Mayrhofer M, Rydåker M, Akerud H, Isaksson A. Loss-of-heterozygosity on chromosome 19q in early-stage serous ovarian cancer is associated with recurrent disease. BMC Cancer. 2012;12:407./Bellido F, Pineda M, Sanz-Pamplona R, Navarro M, Nadal M, Lázaro C, *et al*. Comprehensive molecular characterization of hereditary non-polyposis colorectal tumours with mismatch repair proficiency. Eur J Cancer.2014;50(11):1964–72.) Elle constitue la plateforme de choix pour la détection des variantes structurelles (VS) de l’ADN telles que les insertions, duplications et délétions chromosomiques. Outre ces variations du nombre de copies (CNV), des puces de l’ensemble du génome qui couvrent non seulement les régions polymorphes (SNP), mais également des régions non polymorphes peuvent détecter des déséquilibres chromosomiques et un déséquilibre allélique indiquant l’absence d’hétérozygotie (AOH), la perte d’hétérozygotie (LOH), ou de longues extensions contiguës d’homozygotie (LCSH).

L’avantage de cette technique est la faible quantité d’ADN nécessaire (≥ 80ng d’ADN), contrairement à la technologie Agilent (> 1µg) ainsi que l’utilisation de sondes de petites tailles permettant l’application de cette technique à des ADN partiellement dégradés (échantillons FPPE ; formalin fixed paraffin embedded tissus).

**L’analyse d’un profil moléculaire par CGH permet globalement :**

* Le ratio en log2
* Le ratio en log2 lissé
* La fréquence allélique de l’allèle B (BAF : B allele frequency)
* Le nombre de copies pour chaque région de la puce (copy number)
* La perte d’hétérozygotie par région (LOH)
* La différence allélique ; proportion de l’allèle B par rapport à l’allèle A
* Les gènes présents dans chaque région

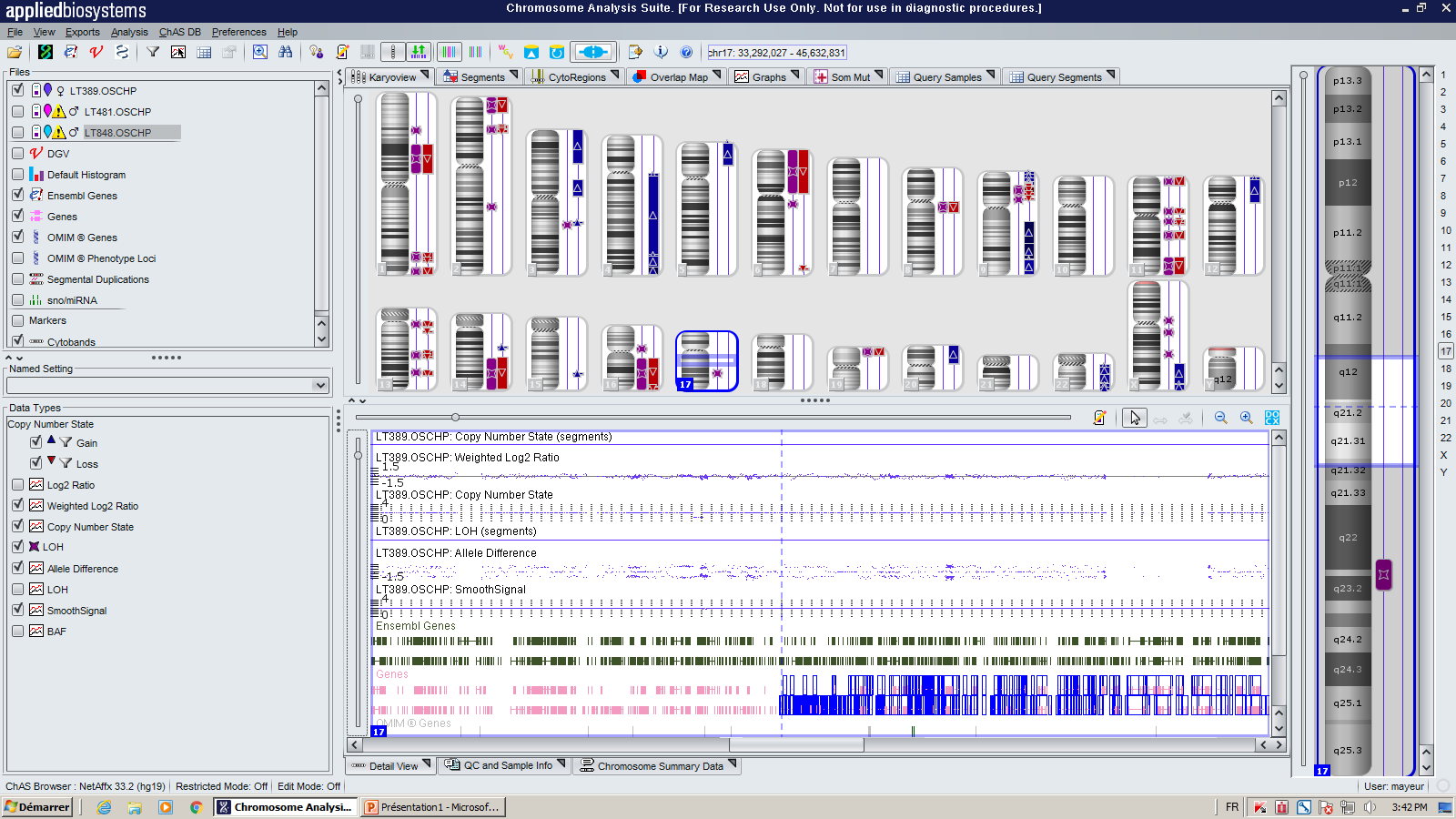


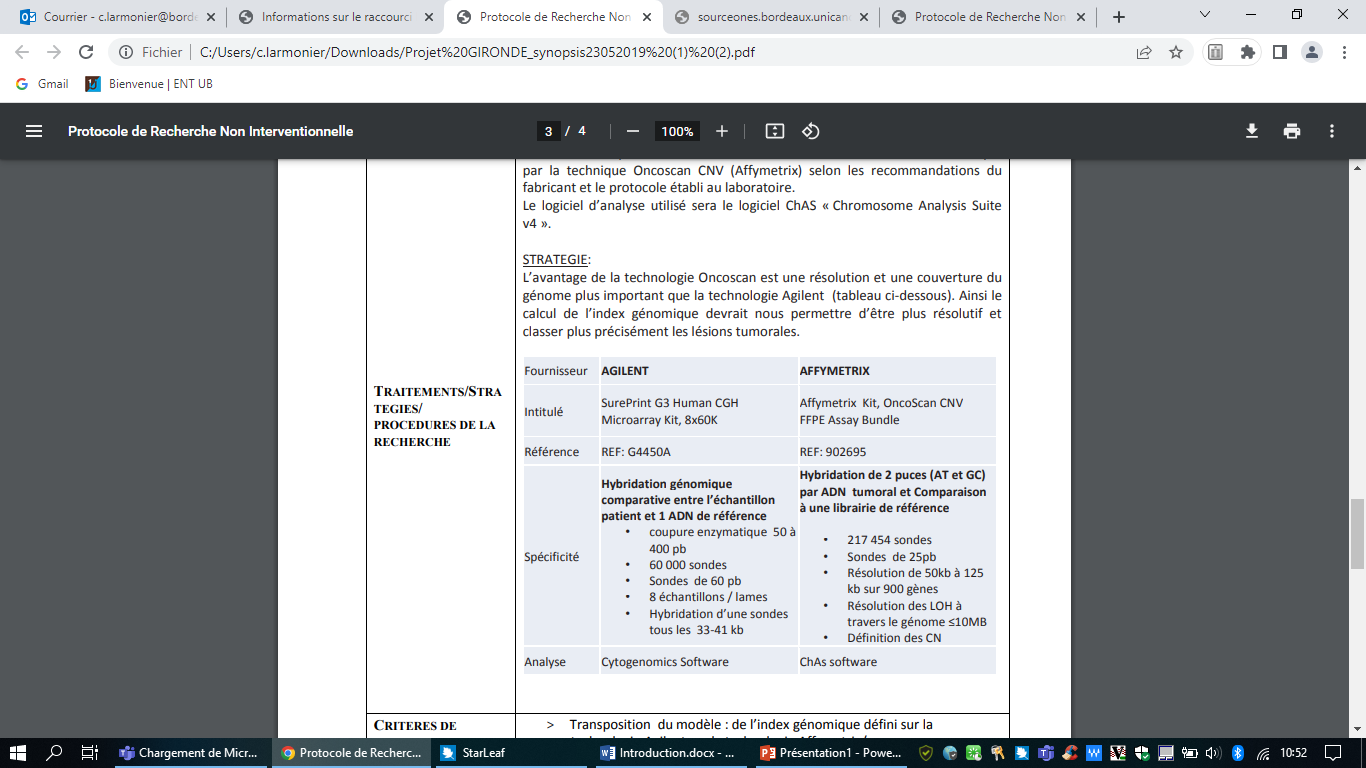
Figure 2 – A/ : Représentation d’un profil sur le logiciel CHAS (Affymetrix, actuellement utilisé pour l’analyse des profils Oncoscan), B/ Pistes log Ratio (en haut) et BAF (en bas) de l’échantillon 5-LD. Les numéros des chromosomes sont indiqués sur les deux pistes. Chaque point correspond à une valeur de log Ratio ou BAF pour une position

* En parallèle de signature moléculaire plus complexe, le laboratoire a précédemment démontré l’intérêt diagnostique et pronostique de l’index génomique (GI) par CGH-array dans différents types de sarcomes comme les GIST (ref 3,5) , les synovialosarcomes (ref 4) et les tumeurs musculaires lisses de l’utérus (ref 6) . Le GI est défini par le nombre d’altérations de nombre de copies 2 sur le nombre de chromosomes qui les portent. Ce « genomic index » est le reflet direct du degré de complexité moléculaire et d’instabilité génomique de la tumeur et s’avère être un puissant prédicteur de l’agressivité tumorale et de la rechute métastatique des tumeurs. Ces études ont été réalisées avec les microarrays 8x60 K whole genome d’Agilent, à partir de matériel fixé en formol et inclus en paraffine ce qui rend pertinent et accessible leur utilisation en clinique

Figure 3 – Les profils CGH des échantillons 9-LA (en haut) et 16-DD (en bas). En noir : valeurs brutes de log Ratio. En orange : régions altérées déterminés à partir de ces valeurs. Les numéros des chromosomes sont indiqués en bas de chaque chromosome.

1. **Objectif :**

Dans le but d’étendre l’utilisation de l’évaluation de l’index génomique (GI) par CGH, nous souhaiterions transposer l’approche validée précédemment au laboratoire sur puces AGILENT/SurePrint G3 8x60 K à la technologie Affymetrix/Oncoscan plus récente, plus résolutive et demandant moins de matériel moléculaire, également déjà utilisée au laboratoire (figure 4). Les 2 technologies n’ayant pas du tout la même couverture du génome, il est nécessaire de les comparer sur une série de tumeurs parmi les cas précédemment publiés et pour lesquels nous disposons de toutes les données cliniques et biologiques.



Pour répondre à cette problématique, nous avons 1/ Evalué des outils bioinformatiques permettant de calculer le GI à partir des données OncoScan, 2/ Comparer leurs spécificités et leurs performances. Cela inclut le temps de calcul, mais aussi la sensibilité et la spécificité des outils dans l’identification d’altérations de nombre de copies. 3/ Corréler les valeurs de GI définies par un outil et les valeurs obtenues par la méthodologie validée.

⮚La définition de l’index génomique constitue un outil critique pour la classification des tumeurs, préciser un diagnostic et assigner un traitement pertinent. Utilisé couramment au laboratoire sur la technologie Oncoscan, Il permettra de communiquer des données moléculaires complexes, en termes quantitatifs et soutiendra l'application réussie de la génomique dans les soins aux patients.

1. **Spécificité de mon stage :** Interaction Bioinformatique-Biologie

Ce travail s’inscrit dans le projet GIRONDE « : GIRONDE/ Génomique Index Resolution by ONcoscan Definition and Expertise «, pour lequel le laboratoire a obtenu un financement industriel. Ce projet a un caractère concret de par sa problématique et également par mon intégration dans l’unité de pathologie moléculaire. J’ai ainsi pu suivre le parcours des échantillons, de leur arrivée au laboratoire, au suivi d’une technique de CGH et à l’analyse des résultats. J’ai à la fois interagi avec l’équipe de biologistes pour mieux cerner leur préoccupation quotidienne et avec l’équipe bioinformatique du departement de Biopathologie pour comprendre les enjeux de la mise en place de nouvel outils bioinformatiques destinée l’analyse de routine. Lors de réunion de présentation hebdomadaire, un travail de vulgarisation des outils bioinformatiques testés a ainsi été fait pour souligner les points d’intérêt de chacun d’eux. J’ai également participé avec l’équipe à la formation à l’utilisation du logiciel CHAS, dispensée par l’ingénieur d’application Affymetrix en mai.

A ajouter dans état de l’art concernant les échantillons utilisés.

Sur la base des travaux précédemment publiés (REF), le laboratoire souhaite valider l’index génomique défini par la technique Oncoscan, dans un 1er temps dans les tumeurs stromales gastro-intestinales (GIST). En effet, notre laboratoire est impliqué dans un protocole de recherche clinique ayant pour objectif l’évaluation de l’efficacité d’un traitement adjuvant par Imatinib pour les tumeurs GIST de risque intermédiaire présentant un GI de mauvais prognostic ( ref 3 Gironde) . Pour ces tumeurs et sur la base du calcul de l’index génomique, les patients présentant un GI≥10 peuvent bénéficier d’un traitement adjuvant. Un protocole de surveillance est mis en place pour les patients avec un GI