



**HAL**  
open science

# Caractérisation multiparamétrique des cancers de l'oesophage et de l'estomac

Gaël Quéré

► **To cite this version:**

Gaël Quéré. Caractérisation multiparamétrique des cancers de l'oesophage et de l'estomac. Médecine humaine et pathologie. Université de Bretagne occidentale - Brest, 2018. Français. NNT : 2018BRES0100 . tel-03148941

**HAL Id: tel-03148941**

**<https://theses.hal.science/tel-03148941>**

Submitted on 22 Feb 2021

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

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

# THESE DE DOCTORAT DE

L'UNIVERSITE  
DE BRETAGNE OCCIDENTALE

ECOLE DOCTORALE N° 605

*Biologie Santé*

Spécialité : « *Génétique, Génomique, Bioinformatique* »

Par

**Gaël QUERE**

## **Caractérisation multiparamétrique des cancers de l'œsophage et de l'estomac.**

Thèse présentée et soutenue à Brest, le 18/12/2018

Unité de recherche : Inserm UMR 1078 Génétique, Génomique Fonctionnelle et biotechnologies

### **Rapporteurs avant soutenance :**

Alex DUVAL                      Professeur-Praticien hospitalier, INSERM UMR 938  
Béatrice EYMIN                Directrice de recherche, INSERM UMR 1209

### **Composition du Jury :**

Président :  
GENIN Emmanuelle          Directrice de recherche, INSERM 1078

Alex DUVAL                      Professeur-Praticien hospitalier, INSERM UMR 938, rapporteur  
Béatrice EYMIN                Directrice de recherche, INSERM UMR 1209, Rapporteur

Dir. de thèse :  
Laurent CORCOS              Directeur de recherche, INSERM UMR 1078,

Co-dir. de thèse :  
Cédric LE MARECHAL        Professeur-Praticien hospitalier, INSERM UMR 1078

### **Invité(s)**

Astrid LIEVRE                    Professeur-Praticien hospitalier, CHU de Pontchaillou Rennes,

Jean-Philippe METGES        Professeur-Praticien hospitalier, CHRU de Brest,



## Table des matières

|  |    |
|--|----|
| Avant-Propos.....  | 6  |
| Index des figures .....  | 8  |
| Index des tableaux.....  | 11 |
| Listes des annexes.....  | 12 |
| Liste des abréviations.....  | 13 |
| Chapitre 1 Généralités sur le cancer gastrique.....                                | 17 |
| A. Anatomie de l'œsophage et de l'estomac .....                                    | 17 |
| 1. Organisation et fonctions de l'œsophage et de l'estomac .....                   | 17 |
| 2. Aspects microscopiques de l'œsophage et de l'estomac.....                       | 19 |
| Chapitre 2 Mécanismes impliqués dans la pathogenèse du cancer gastrique.....       | 23 |
| A. Description anatomopathologique du cancer gastrique .....                       | 23 |
| 1. Définitions.....  | 23 |
| 2. Classification anatomopathologique.....   | 24 |
| 3. Classification de l'OMS .....   | 24 |
| 4. Classification de Lauren.....   | 25 |
| 5. Classification du stade tumoral .....   | 27 |
| 6. Evolution vers un stade métastatique .....                                      | 30 |
| B. Classification moléculaire du cancer gastrique.....                             | 30 |
| 1. Schéma de pathogenèse séquentielle : modèle de la carcinogenèse gastrique ..... | 30 |
| 2. La classification génomique issue du TCGA.....                                  | 32 |
| Chapitre 3 Epidémiologie du cancer gastrique.....                                  | 36 |
| A. Incidence et prévalence du cancer gastrique .....                               | 36 |
| B. Mortalité du cancer gastrique.....  | 37 |
| C. Facteurs de risques.....  | 37 |
| 1. Facteurs de risques connus .....  | 38 |
| 2. Facteurs de risques possibles .....   | 39 |
| D. Diagnostic et traitement du cancer gastrique .....                              | 39 |
| 1. Diagnostic .....  | 39 |
| 2. Biomarqueurs cliniques .....  | 40 |

|                                |  |     |
|--------------------------------|--|-----|
| E.                             | Prise en charge thérapeutique du cancer gastrique .....                      | 42  |
| 1.                             | Classique .....  | 42  |
| 2.                             | De nouvelle génération (immunothérapie).....                                 | 45  |
| Chapitre 4                     | Oncogenèse MSI.....  | 47  |
| A.                             | Découverte du phénotype tumoral MSI .....                                    | 47  |
| B.                             | Le système MMR et les conséquences de sa déficience.....                     | 47  |
| 1.                             | Fonctionnement du système MMR .....  | 47  |
| 2.                             | Les conséquences de l'inactivation du système MMR.....                       | 49  |
| C.                             | Le syndrome de Lynch.....  | 52  |
| 1.                             | Découverte et définition.....  | 52  |
| 2.                             | Caractéristiques génétiques .....  | 52  |
| 3.                             | Caractéristiques épidémiologiques du Syndrome de Lynch .....                 | 54  |
| 4.                             | Critères d'identification du Syndrome de Lynch .....                         | 55  |
| D.                             | Les tumeurs MSI sporadiques .....  | 56  |
| 1.                             | Cancers gastriques MSI sporadiques.....                                      | 56  |
| 2.                             | Autres cancers MSI sporadiques .....   | 57  |
| E.                             | Détermination du statut MSI .....  | 58  |
| 1.                             | Diagnostic par PCR.....  | 59  |
| 1.                             | Analyse immunohistochimique .....  | 60  |
| 2.                             | Valeur prédictive du statut MSI.....   | 61  |
| 3.                             | Protocoles de traitements spécifiques aux cancers MSI – Immunothérapies..... | 62  |
| Chapitre 5                     | Epissage alternatif et cancer .....  | 63  |
| A.                             | Généralités sur l'épissage alternatif.....                                   | 63  |
| B.                             | Les séquences spécifiques de l'épissage.....                                 | 65  |
| C.                             | Les séquences et facteurs régulateurs de l'épissage .....                    | 67  |
| D.                             | Epissage alternatif et pathologies .....                                     | 70  |
| Objectifs des travaux de thèse | .....  | 75  |
| Résultats I                    | .....  | 77  |
| Résultats II                   | .....  | 119 |
| Bibliographie                  | .....  | 156 |
| Annexes                        | .....  | 173 |



# Avant-Propos

---

L'origine du cancer semble être aussi vieille que celle du vivant. Chez l'Homme, on retrouve déjà des traces d'ostéosarcomes sur des ossements datant du Néolithique. Les premières descriptions des cancers sont retrouvées dans des textes égyptiens notamment dans le papyrus dit d'Edwin Smith (-3000 av. J.C.). Ces textes évoquent déjà l'ablation de tumeurs mammaires, de la tête et du cou et leur prise en charge par cautérisation.

C'est le médecin grec Hippocrate qui définira pour la première fois la maladie, évoquant des grosseurs en forme de crabes, il distinguera à l'époque les tumeurs bénignes appelées carcinos, le cancer curable, squirrhos et le cancer entraînant la mort, carcinoma.

L'autre terminologie adoptée est attribuée au médecin gréco-romain Galien (-130-200 av. J.C.) qui décrit des gonflements (ovκος ou oncos) et dont dérive le terme moderne d'oncologie.

Les travaux qui seront présentés ici dans le cadre de mon doctorat, se porteront sur le cancer de l'estomac, enjeu majeur de santé publique au niveau mondial. En 2016, l'Agence internationale pour la recherche sur le cancer (IARC) estimait à plus de 800 000 le nombre de décès dus au cancer gastrique dans le monde, soit la deuxième cause de mortalité par cancer. L'incidence du cancer gastrique est élevée et il est caractérisé par une grande disparité géographique avec comme régions majoritairement touchées l'Asie, l'Europe de l'ouest et l'Amérique du sud.

L'identification de nouveaux marqueurs permettant une détection précoce de ce cancer, une meilleure caractérisation de celui-ci et une meilleure prise en charge thérapeutique s'avère donc être un enjeu important. Ces nouveaux marqueurs peuvent prendre de nombreuses formes comme des transcrits dont l'expression est dérégulée ou des défauts dans les mécanismes de régulation post-transcriptionnelle (épissage alternatif). La multiplication de l'identification de nouveaux marqueurs participe par conséquent à une meilleure connaissance et une meilleure compréhension du mécanisme de cancérisation ainsi qu'à l'augmentation de la sensibilité et de la spécificité des prédictions de l'évolution de ces cancers.

# Listes des posters

---

- Journée du cancéropôle Grand-Ouest, 16- 17 juin 2016, Sables d'Olonne  
caractérisation multiparamétrique des cancers de l'œsophage et de l'estomac  
Gaël Quéré, Jean-Philippe Metges, Laurent Corcos et Cédric Le Maréchal
  
- Congrès EpiBrest, 7-8 décembre 2016, Océanopolis – Brest  
Molecular characterization of esophageal and stomach adenocarcinoma  
Gaël Quéré, Jean-Philippe Metges, Laurent Corcos et Cédric Le Maréchal
  
- Colloque annuel IBSAM, 7 juin 2017, UFR Médecine – Brest  
Caractérisation multiparamétrique des cancers de l'œsophage et de l'estomac  
Gaël Quéré, Jean-Philippe Metges, Laurent Corcos et Cédric Le Maréchal
  
- Journée des jeunes chercheurs en Santé, 8 juin 2017, UFR Médecine – Brest  
Caractérisation multiparamétrique des cancers de l'œsophage et de l'estomac  
Gaël Quéré, Jean-Philippe Metges, Laurent Corcos et Cédric Le Maréchal
  
- Journée du cancéropôle Grand-Ouest, 29- 30 juin 2017, Vannes  
Caractérisation multiparamétrique des cancers de l'œsophage et de l'estomac  
Gaël Quéré, Jean-Philippe Metges, Laurent Corcos et Cédric Le Maréchal
  
- Journée du cancéropôle Grand-Ouest, 2-3 juillet 2018, Vannes  
Caractérisation multiparamétrique des cancers de l'œsophage et de l'estomac  
Gaël Quéré, Jean-Philippe Metges, Laurent Corcos et Cédric Le Maréchal
  
- Congrès SPLICING 2018, 16-19 juillet 2018, Lisbonne, Portugal  
Functional Genomics of Gastric Cancer  
Gaël Quéré, Jean-Philippe Metges, Laurent Corcos et Cédric Le Maréchal



# Index des figures

---

|  |    |
|--|----|
| Figure 1 : Anatomie de la région gastrique. Adapté de « Les fondamentaux de la pathologie digestives » Chapitre 1 et 2 (CDU-HGE Octobre 2014).....   | 19 |
| Figure 2 : Anatomie de l'œsophage. (INCA : <a href="https://www.e-cancer.fr/Patients-et-proches/Les-cancers/Cancers-de-l-oesophage/L-oesophage">https://www.e-cancer.fr/Patients-et-proches/Les-cancers/Cancers-de-l-oesophage/L-oesophage</a> ) ..... | 20 |
| Figure 3 : Histologie de la paroi stomacale. Adapté de « Physiologie humaine » (Sherwood, 2006).....   | 22 |
| Figure 4 : Classification des cancers de la jonction œsogastrique en fonction de la localisation anatomique décrite par Siewert (Siewert et al, 2005).....   | 23 |
| Figure 5 : Processus de tumorigenèse impliqué dans l'apparition des adénocarcinomes de type intestinal et de type diffus, selon la voie de Correa (Correa, 1992). .....  | 26 |
| Figure 6 : Altérations génomiques retrouvées dans le cancer gastrique (Tan et Yeoh, 2015).   | 31 |
| Figure 7 : Classification moléculaire du cancer gastrique, d'après les résultats obtenus par le TCGA (2014) (The Cancer Genome Atlas Research Network, 2014).....  | 32 |
| Figure 8 : Division des 4 sous-types en fonction des altérations moléculaires majeures retrouvées dans le cancer gastrique (The Cancer Genome Atlas Research Network, 2014)..  | 33 |
| Figure 9 : Répartition géographique des cancers de l'estomac dans le monde selon l'incidence pour les deux sexes ( <a href="http://globo-can.iarc.fr">globo-can.iarc.fr</a> ).....   | 37 |
| Figure 10 : : Voies moléculaires ciblées par les approches de thérapie ciblée utilisant des anticorps monoclonaux «mab» (Riquelme et al., 2015).....   | 46 |
| Figure 11 : Les différents complexes du système MMR chez l'homme.....  | 49 |
| Figure 12 : Mécanisme alternatif de prédisposition au cancer lié à des épimutations constitutionnelles d'après Suter et al, 2004. Certaines mutations sont causées par des   |    |

|  |    |
|--|----|
| altérations génétiques sous-jacentes agissant en cis à proximité du gène affecté et présentent une transmission autosomique dominante. ....  | 53 |
| Figure 13 : Détermination du statut MSI des tumeurs par PCR (pentaplex). ....  | 60 |
| Figure 14 : Détermination du statut MSI des tumeurs par immunohistochimie.....   | 61 |
| Figure 15 : Illustration des mécanismes de transcription et de maturation de l'ARN pré-messager (Samaan et al. 2013). ....   | 63 |
| Figure 16 : Représentation des différents types d'épissage alternatif. (Ghigna et coll., 2008) .....   | 65 |
| Figure 17 : Séquences consensus des sites d'épissage. Il existe des séquences consensus pour les sites d'épissage en 5' et 3', et pour le site de branchement. La taille des lettres pour chaque position est proportionnelle à la fréquence de la présence du nucléotide à cette position. (McManus & Graveley, 2011). .... | 65 |
| Figure 18 : Représentation simplifiée de l'assemblage du splicéosome. Différents complexes sont décrits : les complexes E, A, B, Bact (activé), C et P. (Will & Lührmann, 2011) .....  | 66 |
| Figure 19 : Structure secondaire des snRNA et protéines composant les snRNP chez l'Homme. En gris est noté si les snRNP sont associées avec des protéines Sm (B/B', D3, D2, D1, E, F, et G) ou des protéines LSm (LSm2-8). (Will & Lührmann, 2011).....  | 67 |
| Figure 20 : Représentation des principaux facteurs de régulation agissant en cis ou en trans. Les séquences cis-régulatrices sont situées dans les exons (ESE/ESS) ou dans les introns (ISE/ISS). Les protéines SR et hnRNP peuvent interagir avec les séquences en cis (Lee and Rio 2015).....                              | 68 |
| Figure 21 : Représentation du mécanisme de dégradation des ARNm non-sens. (Baker and Parker 2004) .....  | 70 |
| Figure 22 : Epissage et pathologies .....  | 71 |

|  |     |
|--|-----|
| Figure 23 : Réseau reliant entre eux les 496 gènes surexprimés et les 123 gènes sous-exprimés dans les adénocarcinomes gastriques. Les traits entre les protéines symbolisent l'indice de confiance de la probabilité d'interaction..... | 126 |
| Figure 24 : Cluster de gènes interagissant ensemble issu du réseau d'interaction de la base de données STRING : 9 gènes : 10 liens identifiés. ....  | 127 |
| Figure 25 : visualisation du réseau « papillon » et voies métaboliques communes (seuil de confiance $\geq$ à 0,900). ....  | 134 |
| Figure 26 : Cycle du glutathion.....   | 136 |
| Figure 27 : Cycle de l'élimination des espèces réactives dérivées de l'oxygène. ....   | 137 |
| Figure 28 : Cycle de métabolisation du Benzo[a]pyrène .....  | 138 |
| Figure 29 : Cycle de métabolisation du 1,1-Dichloroéthylène.....   | 138 |
| Figure 30 : Cycle de métabolisation du trichloréthylène .....  | 139 |
| Figure 31 : Cycle de métabolisation de l'irinotecan.....   | 140 |
| Figure 32 : Cycle de métabolisation de l'aflatoxine B1. ....   | 140 |
| Figure 33 : Cycle de biosynthèses des hormones stéroïdes .....   | 142 |
| Figure 34 : Voie de résistance aux traitements de type sels de platine. ....   | 143 |

# Index des tableaux

---

|  |     |
|--|-----|
| Tableau 1 : Comparaison de deux systèmes de classement du cancer gastrique les plus couramment utilisés : la classification de l'OMS et la classification de Lauren (Hu et al., 2012).<br>.....  | 27  |
| Tableau 2 : Classification clinique TNM (8ème édition 2017) des lésions cancéreuses de la jonction œso-gastrique et de l'estomac .....   | 29  |
| Tableau 3 : Correspondance entre la classification TNM et le stade du cancer gastrique.....  | 29  |
| Tableau 4 : Comparaison du niveau d'expression de 9 gènes, sélectionnés à partir des réseaux d'interactions créés dans STRING v10.0, à partir de la base de données transcriptomiques TCGA et de notre cohorte de patients. Méthodes : base de données TCGA : RNAseq; Cohorte des patients: RT-PCR quantitative. N.D. signifie non détectée..... | 129 |
| Tableau 5 : Liste des voies métaboliques selon Kegg Pathway correspondant à la liste de gènes sélectionnés se basant sur réseau d'interactions STRING. ....  | 134 |

# Listes des annexes

---

## *Annexe 1 : Matériels et méthodes*

Données supplémentaires : Création d'un réseau d'interaction à l'aide de la base de données STRING (version 10.5) à partir de la liste des gènes communément dérégulés à la baisse dans le cancer gastrique.

# Liste des abréviations

---

|        |   |
|--------|---|
| 3'SS   | 3' splicing site                            |
| 5-FU   | 5-Fluorouracile                             |
| 5'SS   | 5' splicing site                            |
| ADN    | Acide desoxyrobonucléique                   |
| ADNc   | ADN complémentaire                          |
| ADNg   | ADN génomique                               |
| APC    | <i>Adenomatous polyposis coli</i>           |
| ARN    | Acide ribonucléique                         |
| ARNm   | ARN messenger                               |
| AS     | Alternative pre-mRNA                        |
| CC     | Cancer du côlon                             |
| CDS    | Coding DNA sequence                         |
| CIMP   | CpG island-hyper-méthylation phenotype      |
| CIN    | Chromosomal Instability                     |
| CR     | Cancer du rectum                            |
| CRC    | Cancer colorectal                           |
| DNMT   | DNA Methyltransferase                       |
| EA     | Epissage alternatif                         |
| EBV    | Virus d'Epstein-Barr                        |
| ECL    | entero chromaffin like                      |
| ECM    | Matrice extracellulaire                     |
| EGF    | Epidermal Growth Factor                     |
| EGFR   | EGF Receptor                                |
| EJC    | Exon Junction Complex                       |
| EMT    | Transition Epithélio-Mésenchymateuse        |
| ERO    | Espèces réactives de l'oxygène              |
| ESS    | Exonic splicing silencer                    |
| ESE    | Exonic Splicing Enhancer                    |
| F      | Forward                                     |
| FC     | Fold-change                                 |
| FdUMP  | 5-fluo-rodéoxyuridine 5'-monophosphate      |
| FOLFOX | 5-FU, Leucovorine et Oxaliplatine           |
| GS     | génomiquement stables                       |
| HER1   | Human Epidermal Growth Factor Receptor      |
| HER2   | Human Epidermal Growth Factor Receptor 2    |
| HGFR   | facteur de croissance hépatocytaire         |
| hMLH1  | Human mutL homolog 1                        |
| HNPCC  | Hereditary non-polyposis CRC                |
| hnRNP  | Heterogeneous Nuclear Ribonucleo-Proteins   |
| IARC   | International agency for research on cancer |
| ICGC   | International Cancer Genome Consortium      |
| IL     | interleukine                                |
| InCA   | Institut national du cancer                 |

|             |   |
|-------------|---|
| Inserm      | Institut national de la santé et de la recherche médicale |
| ISE         | Intronic Splicing Enhancer                                |
| ISS         | Intronic Splicing Silencer                                |
| JAK2        | Janus kinase 2  |
| JOG         | jonction œsogastrique                                     |
| <b>KEGG</b> | Kyoto Encyclopedia of Genes and Genomes                   |
| MALT        | mucosa-associated lymphoid tissue                         |
| MAPK        | Mitogen-activated protein kinases                         |
| miARN       | microARN  |
| MIN         | microsatellite instability                                |
| MLH1        | mutL homolog 1  |
| MMP         | Métalloprotéinase   |
| MMR         | Mismatch repair   |
| M-MuLV      | Moloney Murine Leukemia Virus                             |
| mRNA        | messenger RNA   |
| MSH2        | MutS homolog human 2                                      |
| MSH6        | MutS homolog human 6                                      |
| MSI         | Microsatellite Instability                                |
| MSI-H       | MSI- high   |
| MSI-L       | MSI- Low  |
| MSS         | Microsatellite stable                                     |
| NF-κB       | Nuclear Factor-kappa B                                    |
| NMD         | Nonsense Mediated mRNA Decay                              |
| OMS         | Organisation mondiale de la santé                         |
| PABP        | Poly(A) binding protein                                   |
| PCR         | Polymerase chain reaction                                 |
| PIK3        | Phosphoinositide-3- kinase                                |
| PIK3CA      | Gène de codant la phosphoinositide-3- kinase              |
| PMS1        | PMS1 homolog 1, mismatch repair system component          |
| PMS2        | PMS1 homolog 2, mismatch repair system component          |
| qPCR        | PCR quantitative  |
| <b>R</b>    | Reverse   |
| RER         | Replication Errors  |
| RNA         | Ribonucleic acid  |
| RNP         | Ribonucleic protein                                       |
| <b>ROS</b>  | Reactive oxygen species                                   |
| RT-PCR      | Reverse transcription PCR                                 |
| RT-qPCR     | RT-PCR quantitative                                       |
| siARN       | petit ARN interférent                                     |
| Sm          | Smith   |
| snARN       | Small nuclear ARN   |
| snoARN      | petit ARN nucléaire                                       |
| snRNP       | uridine-rich small nuclear ribonucleoprotein              |
| SR          | Sérine-Arginine   |
| SRPK        | Serine-Arginine Protein Kinase                            |
| SRSF        | Serine/arginine splicing factor                           |

|              |                                     |
|--------------|-------------------------------------|
| SVF          | Sérum de veau foetal                |
| TCF          | Facteur de transcription            |
| TCGA         | The Cancer Genome Atlas             |
| TGF- $\beta$ | Transforming Growth Factor- $\beta$ |
| TNM          | Tumor/ Node/ Metastasis             |
| UMR          | Unité mixte de recherche            |
| UTR          | Untranslated Transcribed Region     |
| VEGF         | Vascular Endothelial Growth Factor  |
| vs.          | <i>versus.</i>                      |



# **Introduction générale**

---

# Chapitre 1 Généralités sur le cancer gastrique

## A. Anatomie de l'œsophage et de l'estomac

### 1. Organisation et fonctions de l'œsophage et de l'estomac

L'œsophage est le segment du tube digestif, long d'environ 25 cm, situé entre le pharynx et l'estomac. C'est un conduit musculomembraneux flexible permettant le passage des aliments de la bouche vers l'estomac. Sa paroi est épaisse de 3 mm et sa lumière est large de 2 à 3 cm. Situé dans le segment inférieur de l'œsophage, un renforcement de la musculature lisse, aussi appelé sphincter inférieur de l'œsophage, prévient normalement le reflux du contenu de l'estomac, grâce à son activité tonique.

L'estomac est un organe creux qui permet de stocker les aliments, d'amorcer leur digestion par une dégradation chimique et mécanique, notamment avec la sécrétion de plusieurs substances tels que l'HCl (acide chlorhydrique), la gastrine (hormone stimulant la sécrétion d'HCl), la pepsine (enzyme protéolytique), et le facteur intrinsèque qui permet l'absorption intestinale de la vitamine B12. De l'histamine est également sécrété par les cellules ECL (entero chromaffin like) qui stimule la sécrétion d'HCl par les cellules pariétales. Le bol alimentaire devient le chyme gastrique et est délivré progressivement au duodénum. Il est composé d'une première zone, le cardia, directement relié à l'œsophage dans sa partie supérieure et fait jonction avec le fundus qui est suivi du corps et de l'antrum. L'extrémité inférieure de l'estomac, le pylore, est relié au duodénum, partie supérieure de l'intestin grêle. L'estomac est anatomiquement divisé en deux parties, la première comprenant le cardia et le fundus et la deuxième, l'antrum et le pylore.

- Aspects macroscopiques de l'œsophage et de l'estomac

L'œsophage possède trois portions, reflétant les trois régions anatomiques qu'il traverse : l'œsophage cervical fait suite au pharynx à travers la jonction pharyngo-œsophagienne, l'œsophage thoracique et l'œsophage abdominal qui s'étend jusqu'à la jonction œsogastrique. Il se compose de trois tuniques concentriques et régulièrement superposées qui sont, de l'intérieur vers l'extérieur, une muqueuse, une sous muqueuse et une musculuse, comprenant deux couches : une circulaire interne et une longitudinale externe.

L'estomac est divisé en 5 régions : Le cardia est la région initiale de l'estomac, situé directement sous l'œsophage avec lequel il communique. Il renferme le sphincter cardial, un mince anneau musculaire qui empêche le contenu gastrique de remonter dans l'œsophage ; le fundus est une région arrondie située à côté du cardia et en dessous du diaphragme ; le corps est la partie principale de l'estomac, lieu de brassage et de digestion des aliments ; l'antrum (ou antrum pylorique) est la partie inférieure de l'estomac. L'antrum entrepose les aliments décomposés jusqu'à qu'ils soient prêts à être évacués vers l'intestin grêle et enfin le pylore qui est la partie de l'estomac communiquant avec l'intestin grêle grâce au sphincter pylorique, anneau de muscles épais qui agit comme une valve pour régler le passage des aliments vers le duodénum (première partie de l'intestin grêle).

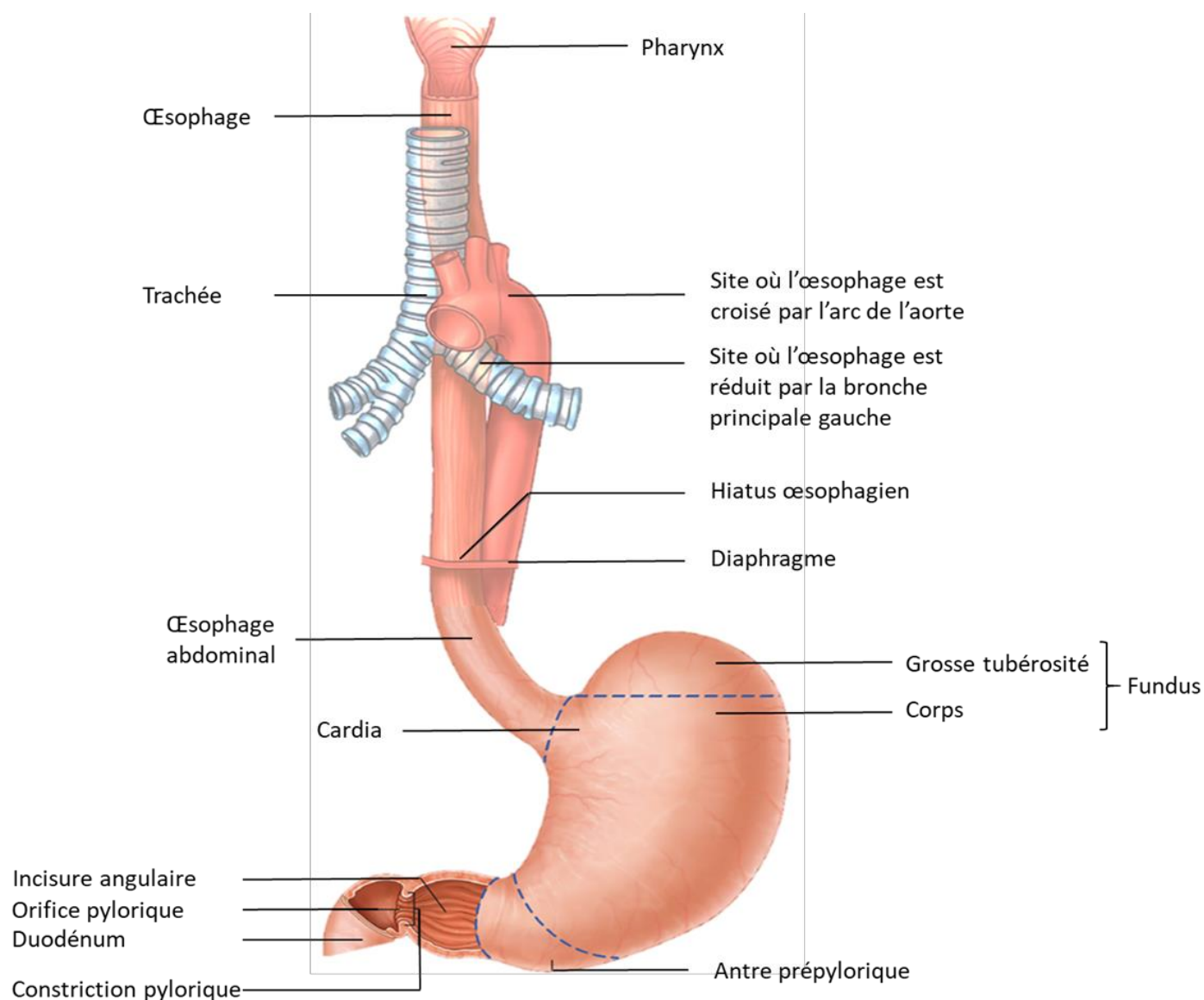


Figure 1 : Anatomie de la région gastrique. Adapté de « Les fondamentaux de la pathologie digestives » Chapitre 1 et 2 (CDU-HGE Octobre 2014).

- Aspects microscopiques de l'œsophage et de l'estomac

La paroi œsophagienne est constituée de la muqueuse œsophagienne formée de trois plans, un épithélium stratifié pavimenteux squameux en continuité avec le recouvrement de l'oropharynx constitue le plan le plus superficiel. Sous l'épithélium, la lamina propria est composée d'un treillis de vaisseaux sanguins et lymphatiques ainsi que des glandes protectrices de mucus. Enfin, le plan le plus profond est la muscularis mucosae composée de fibres musculaires lisses longitudinales.

La sous muqueuse est un plan de tissus de connexion lâche. Il contient des plexus artériels, des fibres élastiques et les corps cellulaires des plexus de Meissner ainsi que deux importants réseaux veineux et lymphatiques.

La musculature est constituée d'un plan circulaire interne et d'un plan longitudinal externe. Entre ces deux plans se situent les ganglions myoentériques qui, avec le plexus de Meissner, coordonnent le mouvement de déglutition. L'œsophage n'est pas recouvert de séreuse mais d'un plan adventiciel qui facilite les mouvements lors de la déglutition.

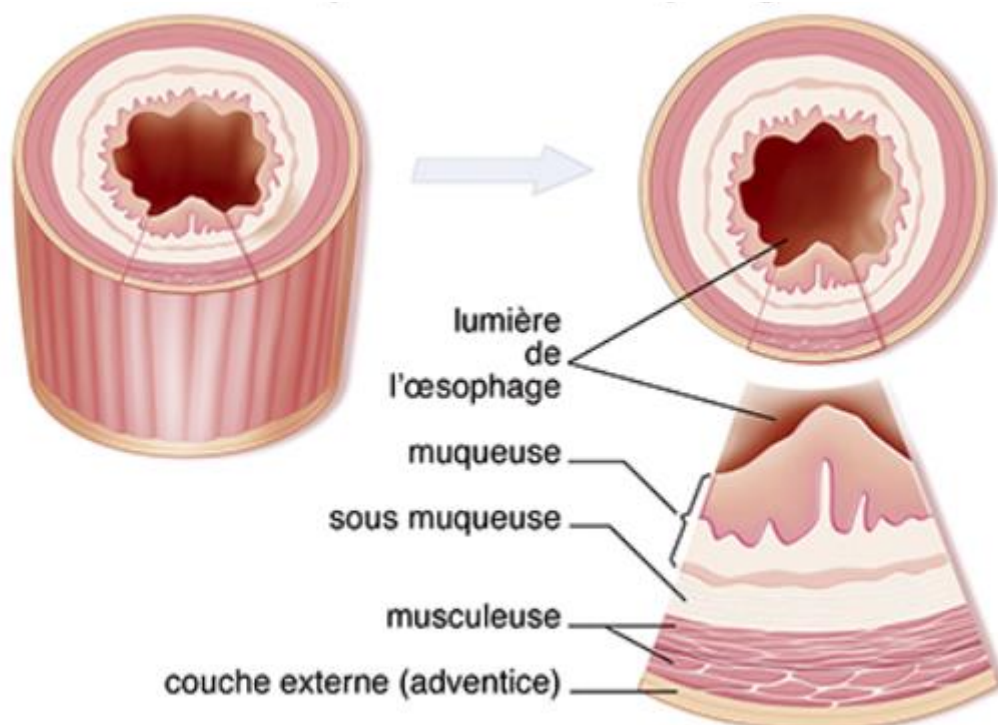


Figure 2 : Anatomie de l'œsophage. (INCA : <https://www.e-cancer.fr/Patients-et-proches/Les-cancers/Cancers-de-l-oesophage/L-oesophage>)

La paroi de l'estomac est constituée, de l'intérieur vers l'extérieur, d'une muqueuse (épithélium, chorion et musculaire muqueuse) qui est le revêtement interne de l'estomac. La couche suivante est la sous-muqueuse, composée de tissus conjonctifs, de gros vaisseaux sanguins et lymphatiques et de cellules nerveuses. La musculature est constituée de trois couches, une oblique interne, une circulaire médiane et une longitudinale externe. La séreuse (péritoine viscéral) est la membrane fibreuse qui recouvre l'extérieur de l'estomac.

On distingue au niveau de la muqueuse gastrique 3 zones : la surface, les cryptes et les glandes. L'épithélium de surface (couche superficielle en regard de la lumière gastrique) est en continuité avec les cryptes gastriques (invagination de l'épithélium de surface). Plus en profondeur, les glandes gastriques communiquent avec le fond des cryptes.

L'épithélium de surface et celui des cryptes est le même dans tout l'estomac : il s'agit d'un épithélium à pôle muqueux fermé. Les glandes gastriques sont différentes en fonction du type de muqueuse. La zone de renouvellement cellulaire est située à la base des cryptes.

Il existe ainsi trois principaux types de muqueuses gastriques :

- La muqueuse de type cardiaal située au niveau du cardia. Les glandes cardiales sont muco-sécrétantes ;
- La muqueuse de type fundique située dans la grosse tubérosité et le corps de l'estomac. Les glandes fundiques sont droites et faites essentiellement de cellules pariétales de cellules principales (sécrétion de pepsinogène), et de cellules neuroendocrines produisant de l'histamine.
- La muqueuse de type pylorique dans l'antra. Les glandes pyloriques sont constituées principalement de cellules mucosécrétantes, de cellules neuroendocrines à gastrine ou cellules G. La gastrine est une hormone stimulant la synthèse d'acide chlorhydrique par les glandes fundiques et la prolifération des cellules ECL du fundus, et de quelques autres cellules neuroendocrines.

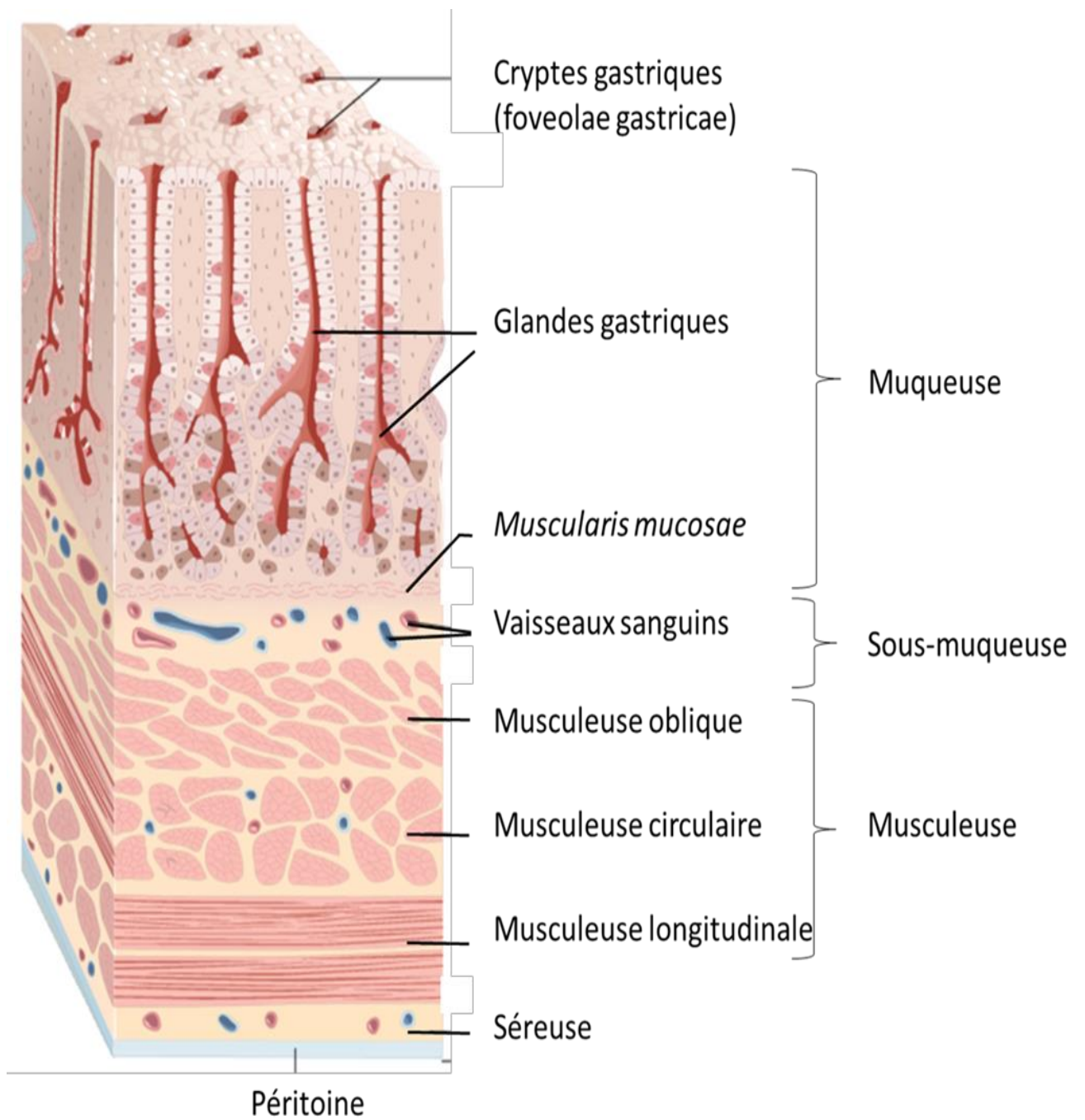


Figure 3 : Histologie de la paroi stomacale. Adapté de « Physiologie humaine » (Sherwood, 2006).

# Chapitre 2 Mécanismes impliqués dans la pathogenèse du cancer gastrique

## A. Description anatomopathologique du cancer gastrique

### 1. Définitions

- Les adénocarcinomes de l'estomac

Les adénocarcinomes primaires de l'estomac se développent à partir des cellules épithéliales de la muqueuse gastrique. On distingue trois types d'adénocarcinomes différents selon la localisation et la forme histologique (Lamarque, 2008). L'adénocarcinome distal de type intestinal et l'adénocarcinome distal de type diffus, qui ont été décrits initialement par Lauren (Lauren, 1965). Puis l'adénocarcinome proximal ou du Cardia ou jonction œsogastrique dont la localisation anatomique précise a été décrite par Siewert (Siewert et al, 2005).

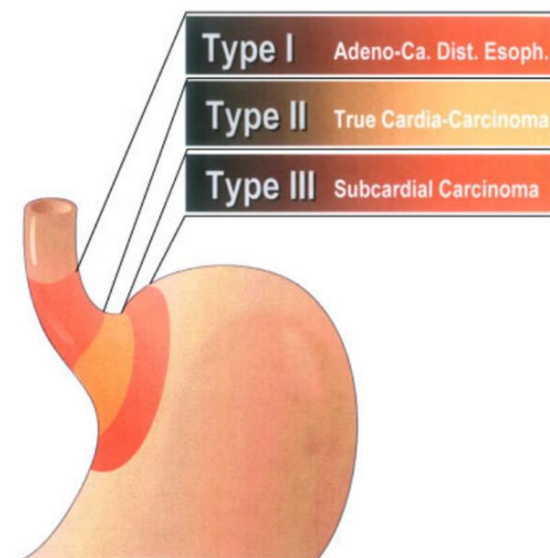


Figure 4 : Classification des cancers de la jonction œsogastrique en fonction de la localisation anatomique décrite par Siewert (Siewert et al, 2005).



Il existe également une autre forme d'adénocarcinome particulière, la linite plastique ou carcinome squirreux. Cette tumeur a la particularité d'être très peu différenciée et de s'infiltrer au niveau de toute l'épaisseur de la paroi stomacale (Ducreux et al, 2000).

- **Les autres formes de cancers**

Les autres formes de cancers ne représentent que 10 % des cancers gastriques. Parmi ces cancers, on retrouve les lymphomes non hodgkiniens qui sont généralement des lymphomes de type MALT (tissu lymphoïde associé aux muqueuses). De façon non exhaustive, ont été décrits dans la littérature des sarcomes (tumeur stromales gastro-intestinales), des tumeurs neuroendocriniennes et des carcinomes lympho-épithéliaux qui ont été associés à l'infection par le virus d'Epstein Barr.

## **2. Classification anatomopathologique**

L'anatomopathologie établit un diagnostic précis du cancer et adapte le niveau de surveillance des lésions précancéreuses comme par exemple l'atrophie gastrique ou la métaplasie intestinale qui évolueront ou non vers un cancer de l'estomac (Correa, 1992).

Il existe actuellement différentes classifications du cancer de l'estomac. De façon non exhaustive, on retrouve la classification de Ming (Ming, 1977), la classification de Goseki (Goseki et al, 1992), mais les classifications les plus utilisées restent celles de Lauren (Lauren, 1965) et de l'OMS (4ème édition, 2010).

## **3. Classification de l'OMS**

La classification de l'OMS datant de 2010 n'est pas limitée aux adénocarcinomes mais classe également les autres formes de cancers de l'estomac (tumeurs non épithéliales : sarcomes, lymphomes...). Cette classification se base sur des critères morphologiques pour définir les différents sous-types de carcinomes de l'estomac : papillaire, tubulaire, mucineux et à cellules rondes dites en bagues à chaton.

## 4. Classification de Lauren

Depuis 1965, selon la classification de Lauren (Lauren, 1965), les adénocarcinomes gastriques peuvent être divisés en 2 types majeurs : intestinal et diffus. Un troisième type, appelé carcinome mixte, caractérise des tumeurs contenant à proportion égale des composants intestinaux et diffus. Cette classification est souvent utilisée par les épidémiologistes.

- L'adénocarcinome de type intestinal

Ce premier type est dit intestinal car d'un point de vue histologique, il est défini par l'apparition de structures glandulaires anormales ressemblant à l'épithélium intestinal. Cette voie cancéreuse fait toujours suite à une gastrite chronique qui évoluera ou non en différents stades pathologiques selon la voie de Correa (Correa, 1992).

- L'adénocarcinome de type diffus

D'un point de vue histologique, ce type ressemble à la classification de l'OMS des carcinomes à cellules rondes dites en bagues à chaton. Les cellules ont la particularité de pouvoir infiltrer toute la paroi de l'estomac. Cette infiltration tumorale anarchique est souvent le résultat de la mutation d'une protéine d'adhésion cellulaire, l'E-cadhérine. Des mutations du gène de l'E-cadherine (gène *CDH1*) sont retrouvées dans 30 % des cas des type diffus (Guilford et al, 1998), dont la voie de la carcinogénèse est encore mal connue et n'est pas aussi bien décrite que pour le type intestinal.

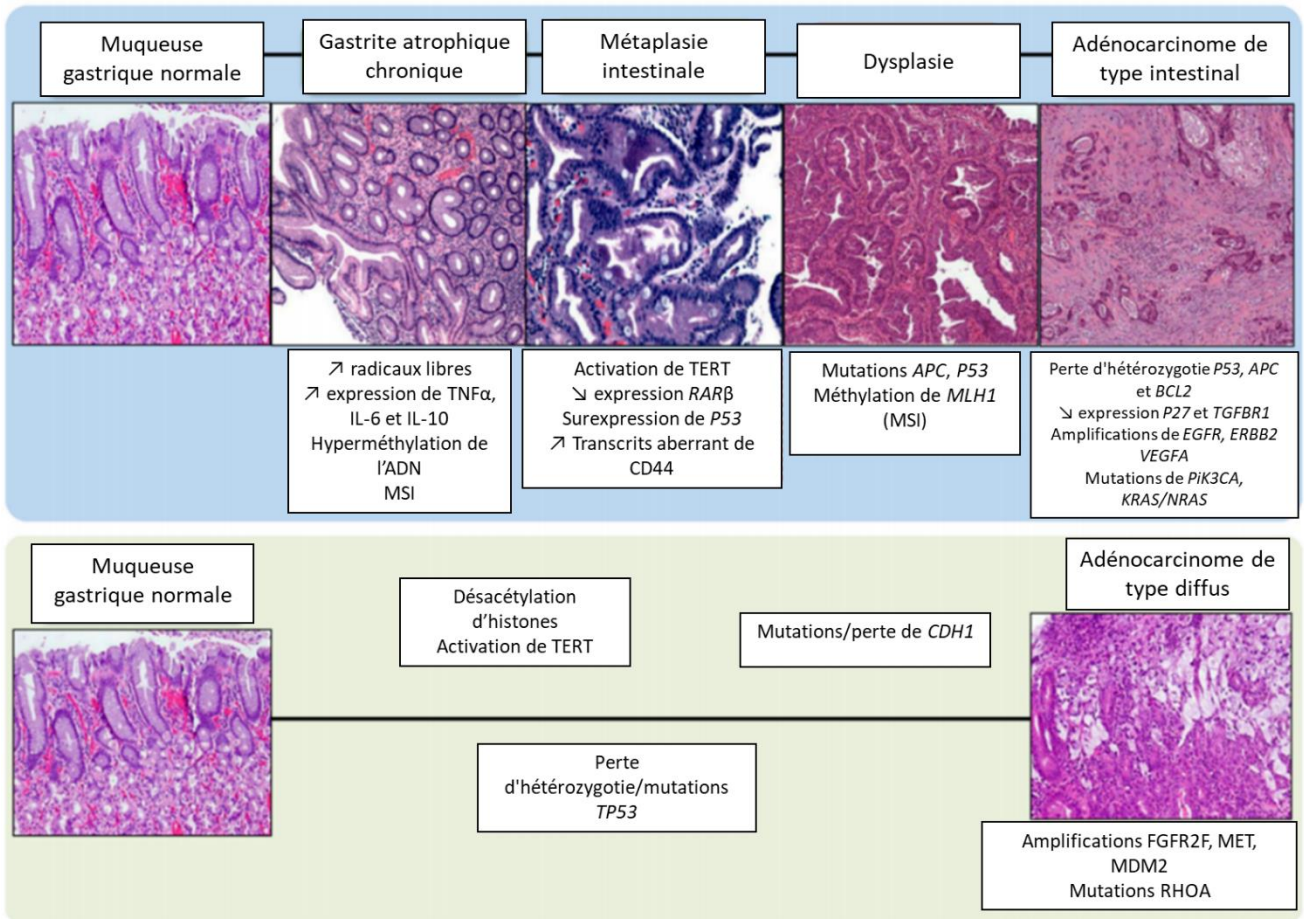


Figure 5 : Processus de tumorigenèse impliqué dans l'apparition des adénocarcinomes de type intestinal et de type diffus, selon la voie de Correa (Correa, 1992).

Pour l'estimation du pronostic, la classification de Lauren est importante au sens où l'adénocarcinome de type diffus est associé à un moins bon pronostic que l'adénocarcinome de type intestinal du fait même de l'infiltration tumorale (Berlth et al, 2014).

Le tableau suivant compare les deux classifications proposées par Lauren et l’OMS (Hu et al., 2012).

| WHO (2010)  | Lauren (1965)      |
|---|--------------------|
| Papillary adenocarcinoma  | Intestinal type    |
| Tubular adenocarcinoma  |                    |
| Mucinous adenocarcinoma   |                    |
| Signet-ring cell carcinoma<br>And other poorly cohesive carcinoma | Diffuse type       |
| Mixed carcinoma   | Indeterminate type |
| Adenosquamous carcinoma   |                    |
| Squamous cell carcinoma   |                    |
| Hepatoid adenocarcinoma   |                    |
| Carcinoma with lymphoid stroma                                    |                    |
| Choriocarcinoma   |                    |
| Carcinosarcoma  |                    |
| Parietal cell carcinoma   |                    |
| Malignant rhabdoid tumor  |                    |
| Mucoepidermoid carcinoma  |                    |
| Paneth cell carcinoma   |                    |
| Undifferentiated carcinoma  |                    |
| Mixed adeno-neuroendocrine carcinoma                              |                    |
| Endodermal sinus tumor  |                    |
| Embryonal carcinoma   |                    |
| Pure gastric yolk sac tumor                                       |                    |
| Oncocytic adenocarcinoma  |                    |

Tableau 1 : Comparaison de deux systèmes de classement du cancer gastrique les plus couramment utilisés : la classification de l’OMS et la classification de Lauren (Hu et al., 2012).

## 5. Classification du stade tumoral

- La classification internationale TNM

Cette classification internationale TNM (tumeur primaire T, l’atteinte des ganglions régionaux N et la présence de métastases à distance M), est applicable à tous les cancers. Elle évalue l’extension de la tumeur et ainsi définit le stade tumoral (Arocena et al, 2006). Cette classification est importante dans la prise en charge du patient, que ce soit pour l’établissement du pronostic ou de l’adaptation des stratégies thérapeutiques.

Deux classifications TNM existent, la classification clinique (cTNM) étant distincte de la classification pathologique (pTNM). La classification clinique repose sur l'examen clinique, l'imagerie, l'endoscopie et l'exploration chirurgicale. La classification pathologique est basée sur l'analyse anatomo-pathologique de la pièce opératoire.

La première classification TNM a été proposée dans les années 1940. Plusieurs révisions ont été publiées, la version actuelle date de 2017. Cette dernière édition semble plus performante. Chaque cancer possède sa propre classification TNM.

La catégorie « T » correspond à la tumeur primitive. Cette catégorie prend par exemple en compte le volume tumoral, le diamètre de la lésion. Pour le stade T4 des subdivisions ont été ajoutées afin d'affiner le diagnostic (Sobin et Compton, 2010). La catégorie « N » correspond aux métastases ganglionnaires. La localisation plus ou moins éloignée de la tumeur, le nombre et la fixation aux tissus voisins des ganglions sont pris en compte dans cette catégorie. La catégorie « M » prend en compte l'absence ou la présence de métastases. En cas de métastases des subdivisions permettent de préciser le nombre d'organes touchés (Tableau 2).

| Tumeur primitive (catégorie T) |   |
|--------------------------------|---|
| Tx                             | Renseignements insuffisants pour classer la tumeur primitive                                      |
| T0                             | Pas de signe de tumeur primitive  |
| Tis                            | Carcinome <i>in situ</i> ou dysplasie de haut grade   |
| T1                             | Tumeur envahissant la lamina propria, la muscularis mucosae ou la sous-muqueuse                   |
| T1a                            | La tumeur envahit la lamina propria ou la muscularis mucosae                                      |
| T1b                            | La tumeur envahit la sous-muqueuse  |
| T2                             | Tumeur envahissant la musculature   |
| T3                             | Tumeur envahissant l'adventice  |
| T4                             | Tumeur envahissant les structures adjacentes  |
| T4a                            | La tumeur envahit la plèvre, le péricarde, la veine azygos, le diaphragme ou le péritoine         |
| T4b                            | La tumeur envahit d'autres structures adjacentes comme l'aorte, un corps vertébral ou la trachée. |

| Adénopathie régionale (catégorie N) |   |
|-------------------------------------|---|
| N <sub>x</sub>                      | Le statut des ganglions régionaux ne peut être évalué   |
| N <sub>0</sub>                      | Pas de signe d'atteinte des ganglions lymphatiques régionaux  |
| N <sub>1</sub>                      | Métastases dans 1 à 2 ganglions lymphatiques régionaux  |
| N <sub>2</sub>                      | Métastases dans 3 à 6 ganglions lymphatiques régionaux  |
| N <sub>3</sub>                      | Métastases dans 7 ganglions régionaux ou plus   |
| pN <sub>0</sub>                     | L'examen histologique d'une pièce opératoire de lymphadénectomie inclura habituellement 6 ganglions lymphatiques ou plus. Si les ganglions ne sont pas atteints, classer pN <sub>0</sub> même si le nombre habituel de ganglions n'est pas atteint. |
| Métastases à distance (catégorie M) |   |
| M <sub>x</sub>                      | Les métastases ne peuvent être évaluées   |
| M <sub>0</sub>                      | Pas de métastase à distance   |
| M <sub>1</sub>                      | Présence de métastase(s) à distance   |

Tableau 2 : Classification clinique TNM (8ème édition 2017) des lésions cancéreuses de la jonction œso-gastrique et de l'estomac

- Autres classifications des lésions cancéreuses gastriques

Entre la classification TNM et les différents stades du cancer colorectal (de I à IV) des correspondances existent. Chaque stade d'avancement du cancer se réfère à une combinaison définie des données T, N et M. Par exemple une lésion au stade 0 correspond à une classification TNM : Tis, N<sub>0</sub>, M<sub>0</sub> (Tableau IV).

| Regroupement en stades |        |                                |                |
|------------------------|--------|--------------------------------|----------------|
| <b>Stade 0</b>         | Tis    | N <sub>0</sub>                 | M <sub>0</sub> |
| <b>Stade IA</b>        | T1     | N <sub>0</sub>                 | M <sub>0</sub> |
| <b>Stade IB</b>        | T2     | N <sub>0</sub>                 | M <sub>0</sub> |
| <b>Stade IIA</b>       | T3     | N <sub>0</sub>                 | M <sub>0</sub> |
| <b>Stade IIB</b>       | T1, T2 | N <sub>1</sub>                 | M <sub>0</sub> |
| <b>Stade IIIA</b>      | T4a    | N <sub>0</sub>                 | M <sub>0</sub> |
|                        | T3     | N <sub>1</sub>                 | M <sub>0</sub> |
|                        | T1, T2 | N <sub>2</sub>                 | M <sub>0</sub> |
| <b>Stade IIIB</b>      | T3     | N <sub>2</sub>                 | M <sub>0</sub> |
| <b>Stade IIIC</b>      | T4a    | N <sub>1</sub> ,N <sub>2</sub> | M <sub>0</sub> |
|                        | T4b    | TousN                          | M <sub>0</sub> |
|                        | Tous T | N <sub>3</sub>                 | M <sub>0</sub> |
| <b>Stade IV</b>        | Tous T | Tous N                         | M <sub>1</sub> |

Tableau 3 : Correspondance entre la classification TNM et le stade du cancer gastrique.

## **6. Evolution vers un stade métastatique**

Du fait de sa position centrale dans la partie abdominale et de la facilité de diffusion des cellules cancéreuses via la perte de fonction d'adhésion (notamment via la mutation du gène CDH1 qui code l'E-Cadhérine), l'évolution vers un stade métastatique est très rapide. Cette extension loco-régionale via le péritoine viscéral, après prolifération dans les quatre tuniques, peut atteindre le foie, le pancréas, la vésicule biliaire ou encore le côlon. En plus de cette diffusion locale, le drainage par les voies lymphatiques permet aux cellules cancéreuses d'atteindre d'autres ganglions situés à distance, et toucher par la suite par exemple les ovaires ou les poumons.

### **B. Classification moléculaire du cancer gastrique**

#### **1. Schéma de pathogenèse séquentielle : modèle de la carcinogenèse gastrique**

Avec l'essor des nouvelles technologies à haut débit comme les puces à ADN et le séquençage de nouvelle génération, plusieurs études ont décrit les différentes altérations génomiques retrouvées dans le cancer gastrique représentées dans la Figure 1 (Tan et Yeoh, 2015). Ces altérations somatiques sont de plusieurs natures : mutations géniques (Lawrence et al, 2013), altérations du nombre de copies (sCNA) (Deng et al, 2012), changements épigénétiques (Niwa et al, 2010) ou encore des changements de niveaux de transcription impliquant les ARN messagers (mRNA) et les ARN non codant (ncRNAs) (Lee et al, 2014 ; Ooi et al, 2009 ; Chang et al, 2014).

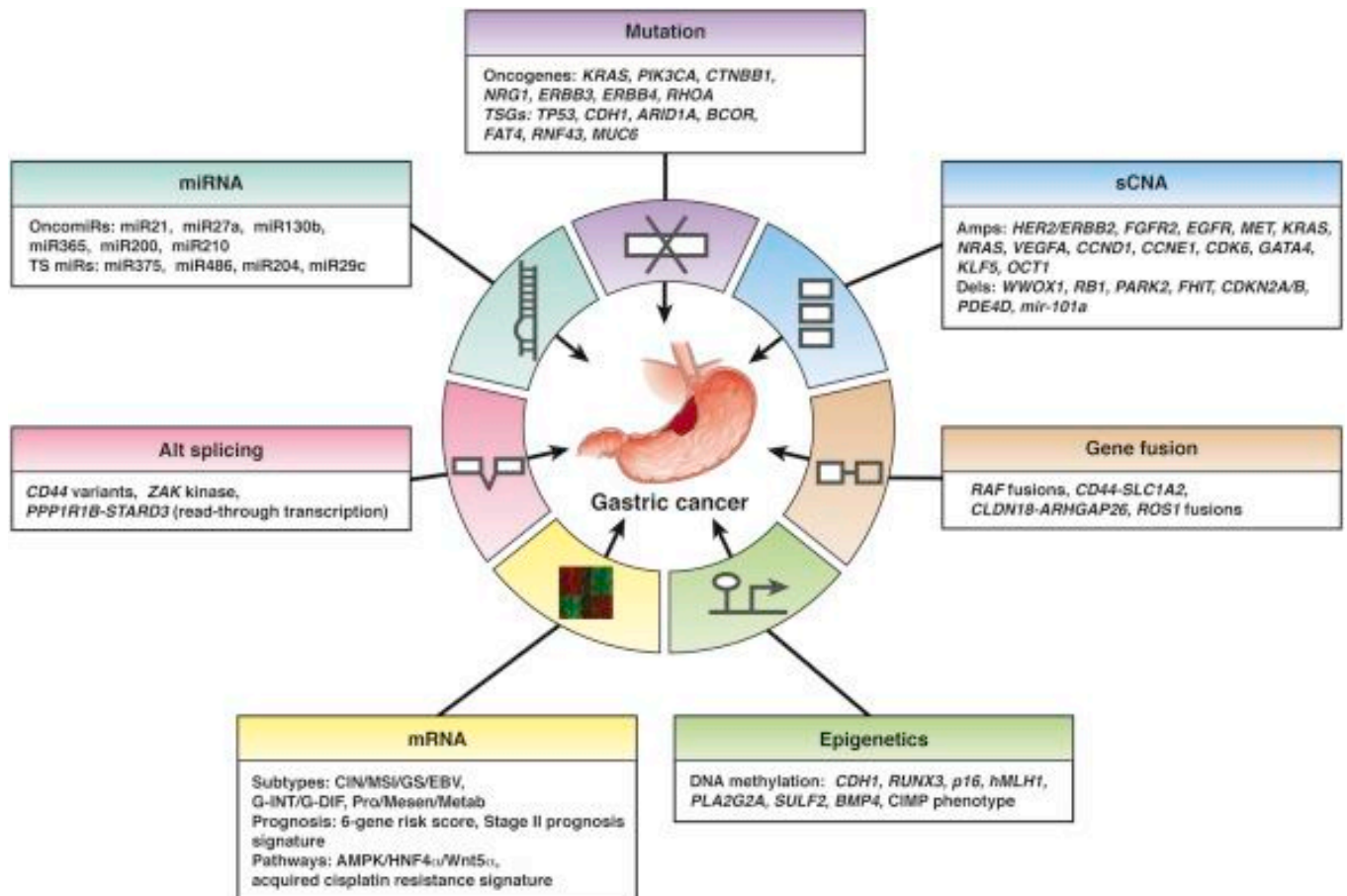


Figure 6 : Altérations génomiques retrouvées dans le cancer gastrique (Tan et Yeoh, 2015).

Toutes ces études ont conduit à une meilleure compréhension de la pathogenèse séquentielle, en mettant en évidence les altérations génétiques et épigénétiques impliquées dans le processus de cancérogenèse gastrique décrit initialement par la voie de Correa (Correa, 1992; Riquelme et al, 2015).

Cependant, cette caractérisation moléculaire du cancer gastrique, via le séquençage de l'ADN (Wang et al, 2011), ne suffit pas à la mise en place d'une classification moléculaire en raison de l'hétérogénéité de ces mutations dans le cancer gastrique (Shah et al, 2011). C'est pourquoi une nouvelle terminologie, basée sur toutes les données de «Omiques» sont intégrées afin d'établir une nouvelle classification moléculaire en cancérologie (The Cancer Genome Atlas Research Network, 2014).



## 2. La classification génomique issue du TCGA

En 2014, le projet The Cancer Genome Atlas (TCGA, [cancergenome.nih.gov](http://cancergenome.nih.gov)) a été lancé avec pour objectifs de développer une classification moléculaire robuste afin d'identifier les différentes voies dérégulées dans plus de 30 types de cancers et notamment dans le cancer gastrique (The Cancer Genome Atlas Research Network, 2014). Ce projet a abouti à proposer une classification moléculaire du cancer gastrique en 4 sous-types : les tumeurs infectées par le virus d'Epstein Barr (EBV) ; les tumeurs avec une instabilité des microsatellites (MSI) ; les tumeurs génomiquement stables (GS) et les tumeurs chromosomiquement instables (CIN).

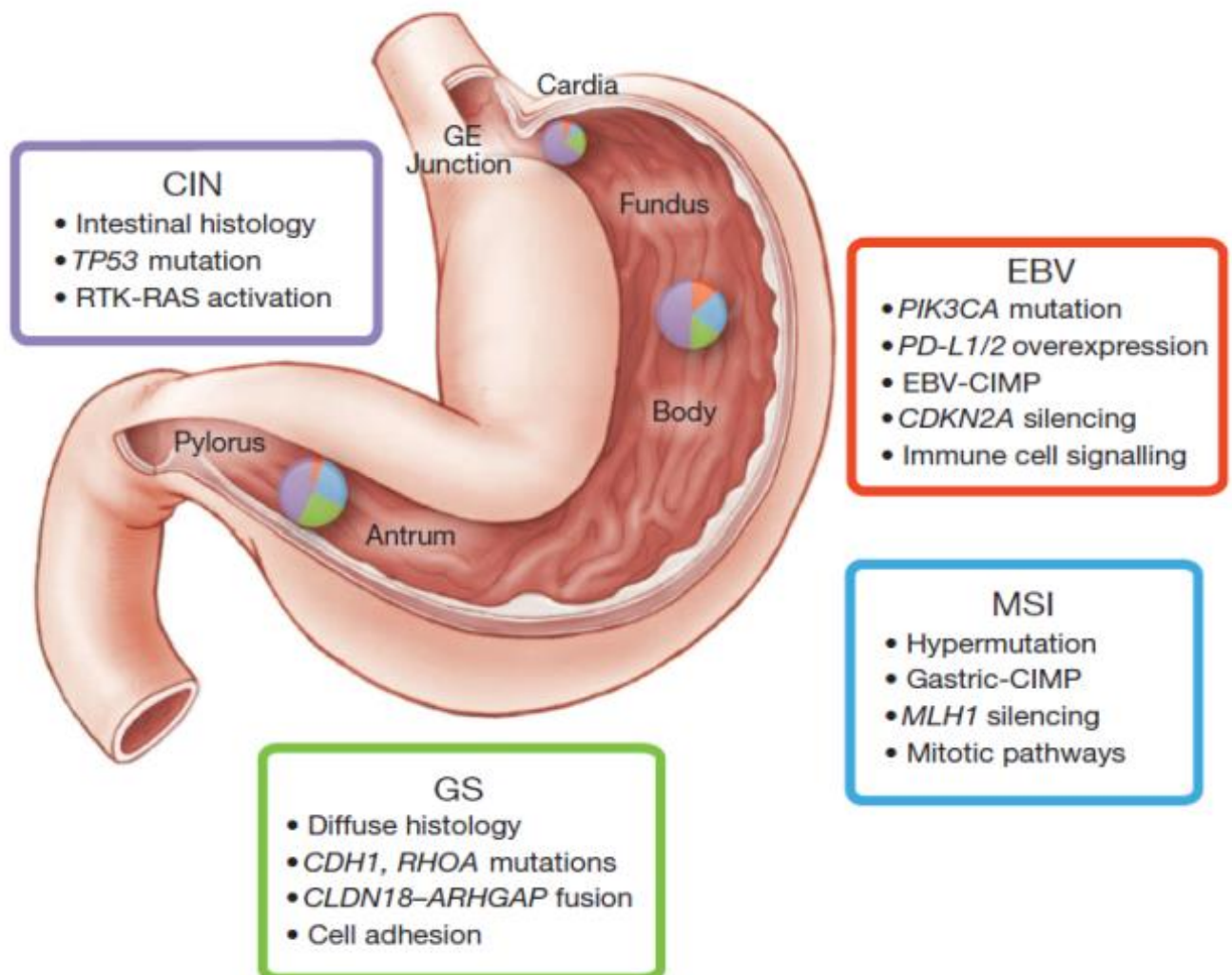


Figure 7 : Classification moléculaire du cancer gastrique, d'après les résultats obtenus par le TCGA (2014) divisée en 4 sous-types : les tumeurs infectées par Epstein Barr Virus (EBV) ; les tumeurs avec une instabilité des microsatellites (MSI) ; les tumeurs génomiquement stables (GS) et les tumeurs chromosomiquement instables (CIN). Les différentes régions de l'estomac (la jonction œsogastrique ou cardiale, le fundus, le corps et l'antra pylorique) sont indiquées avec la répartition des différents sous-types associés (The Cancer Genome Atlas Research Network, 2014).

D'après les résultats obtenus par le TCGA, divisée en 4 sous-types. Les différentes régions de l'estomac (la jonction oeso-gastrique ou cardiale, le fundus, le corps et l'antré pylorique) sont indiquées avec la répartition des différents sous-types associés (The Cancer Genome Atlas Research Network, 2014).

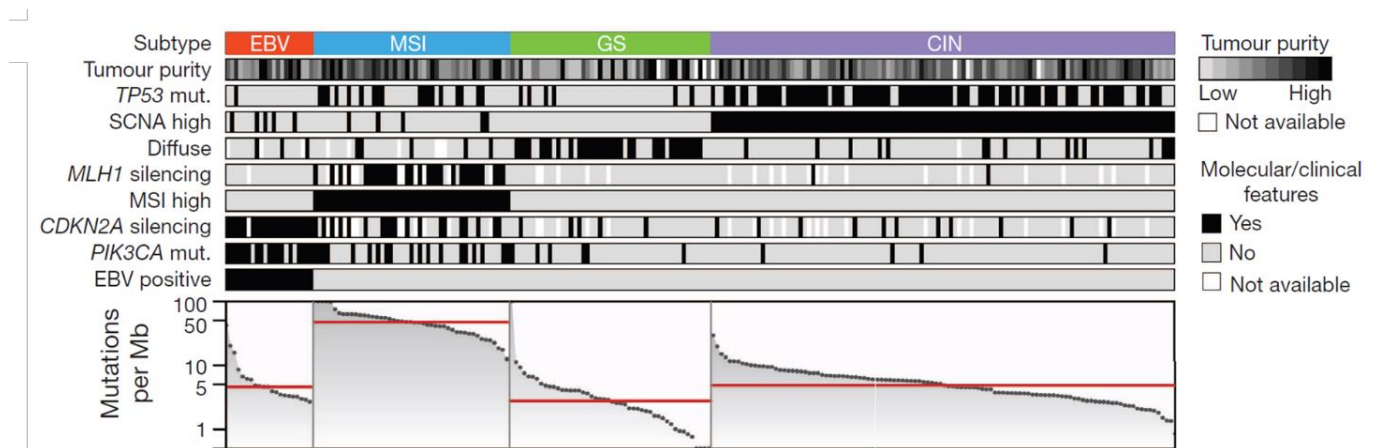


Figure 8 : Division des 4 sous-types en fonction des altérations moléculaires majeures retrouvées dans le cancer gastrique (The Cancer Genome Atlas Research Network, 2014).

- Les tumeurs infectées par le virus d'Epstein Barr (EBV)

Le sous type EBV positif représente environ 10% des cancers gastriques. Il est associé à la pathogénicité du virus et montre la présence d'aberrations chromosomiques et de modifications épigénétiques via la méthylation de régions promotrices de gènes suppresseurs de tumeurs.

Les études du TCGA ont démontré une dérégulation de la voie de signalisation phosphoinositide 3-kinase/ protéine kinase B (PI3K/Akt) et de la voie des Janus kinase 2 (JAK2) dans ces tumeurs (Chen, 2012; Sunakawa and Lenz, 2015).

- Les tumeurs génomiquement stables (GS)

Les tumeurs génomiquement stables (GS) représentent 20% des cancers gastriques mais sont représentatives de près de 75% des formes héréditaires. Ce sous-type est caractérisé par la présence de mutations des gènes *RHOA* et *CDH1*, ainsi que de la fusion des gènes *CLDN18* et *ARHGAP26*.

Le gène *RHOA* code la GTPase RhoA intervenant dans différentes voies de signalisation et notamment dans le cycle cellulaire, l'organisation du cytosquelette et la motilité cellulaire. Les mutations de ce gène vont entraîner une progression tumorale et la formation de métastases due à l'augmentation de la motilité et à la dégradation du cytosquelette.

Le gène *CDH1* code la protéine E-cadhérine impliquée dans l'adhésion cellulaire. La perte de son expression est un marqueur génétique du type diffus (Guilford et al., 1998; Riquelme et al., 2015) entraînant dans un premier temps la formation néoplasique, puis du fait de la perte d'adhésion, la progression métastatique précoce.

- Les tumeurs chromosomiquement instables (CIN)

Les tumeurs chromosomiquement instables (CIN) représentent près de 50 % des cancers gastriques mondiaux. Ce sous-type est caractérisé par un fort taux d'aneuploïdie. En plus de ces aberrations chromosomiques, le sous type CIN-positif est caractérisé par la présence de mutations du gène *TP53* (71%). De plus, de nombreuses voies de signalisation sont dérégulées dans les cancers gastriques CIN-positifs, notamment la voie HER2, qui est actuellement le seul biomarqueur théranostique validé par la communauté scientifique du cancer gastrique (Matsui et al, 2005 ; Fujimoto-Ouchi et al, 2007 ; Bang et al, 2010). Parmi les autres voies dérégulées, on retrouve : le récepteur du facteur de croissance épidermique (EGFR), le récepteur du facteur de croissance endothélial vasculaire (VEGFR), le récepteur au facteur de croissance hépatocytaire (HGFR), et le récepteur 2 du facteur de croissance des fibroblastes (FGFR2).

- Les instabilités des microsatellites (MSI)

Les cancers gastriques avec instabilités des microsatellites (MSI) sont retrouvés dans environ 20% des cas. Dans ce sous-type les altérations génétiques majeures entraînent une inactivation de gènes impliqués dans la réparation des mésappariements de l'ADN, les gènes Mismatch repair (MMR). L'anomalie somatique la plus fréquente est l'hyperméthylation du promoteur du gène *MLH1* (*mutL* *homologue 1*) dont l'inactivation va induire un taux de mutations somatiques élevé dans le génome à l'origine de la tumorigenèse. Cette instabilité est également associée à de nombreuses mutations touchant les gènes *TP53*, *KRAS*, *ARID1A*, *PIK3CA*, *ERBB3*, *PTEN*, *CHRD*, *HLA-B*, *RNF43*, *B2M* et *NF1* (Riquelme et al., 2015).

Bien que plusieurs études aient analysé la relation entre le sous-type MSI et la chimiothérapie adjuvante à base de 5-fluorouracile dans le cancer du côlon, les informations concernant les bénéfices de la chimiothérapie dans le cancer gastrique MSI sont limitées, impliquant des résultats non cohérents et controversés (Ribic et al, 2003). Le statut MSI ne peut donc pas encore être utilisé comme marqueur pronostic, mais il peut jouer un rôle important dans la classification de la tumeur ou la mise en place d'un traitement adjuvant (Sunakawa et Lenz, 2015).

- **Altérations transcriptomiques**

La littérature fait état d'un très grand nombre d'analyses du transcriptome des cancers gastriques. Il existe de nombreuses références dans PubMed avec les mots clés « gastric cancer and transcriptome, les objectifs principaux de ces études sont de découvrir des biomarqueurs pour le diagnostic et/ou le pronostic mais également les voies de signalisation oncogénique impliquées dans l'initiation et/ou la progression du cancer gastrique. Certaines sont disponibles en libre accès et les données brutes d'expression sont référencées dans des bases de données comme GEO (Gene Expression Omnibus, <http://www.ncbi.nlm.nih.gov/gds>) ou Array Express (<https://www.ebi.ac.uk/arrayexpress/>).

Ce nombre important d'études transcriptomiques reflète la grande hétérogénéité du cancer gastrique. On retrouve ainsi de multiples types de comparaisons : cancer gastrique (GC) vs normal, GC non métastatique vs métastatique, ou encore des comparaisons entre les sous-types de GC (diffus, intestinal etc..).

# Chapitre 3 Epidémiologie du cancer gastrique

Du fait de leurs similitudes en termes de symptômes et de facteurs de risque, les cancers de la jonction œsogastrique (JOG) (tumeur correspondant au type II de la classification de Siewert défini par un centre tumoral situé entre 1 cm au-dessus et 2 cm au-dessous du cardia anatomique) et de l'estomac seront regroupés ici sous le terme de cancer gastrique. Ainsi seront détaillés ici à la fois les chiffres des cancers de l'œsophage et de l'estomac.

De plus, en raison des différences multiples et importantes entre l'Asie et l'Occident les résultats des études cliniques asiatiques sont difficilement transposables à la population occidentale. L'analyse de la littérature a été limitée le plus souvent aux études ayant inclus des patients non asiatiques.

## A. Incidence et prévalence du cancer gastrique

A l'échelle mondiale, le Centre International de Recherche sur le Cancer a estimé à 952 000 l'incidence des cas de cancers de l'estomac par an, soit 6.8% des cancers mondiaux, ce qui en fait le 5ème cancer le plus fréquent.

On retrouve une grande disparité géographique avec une majorité de cas de cancers gastriques retrouvés en Asie, à l'Est de l'Europe et en Amérique du Sud avec une incidence annuelle de plus de 40 nouveaux cas pour 100 000 habitants.

En France, le nombre de cancers de l'estomac diagnostiqués est passé de 9000 environ en 1980 à 6550 cas en 2012 et 6600 cas en 2017 et 4 800 nouveaux cas de cancers de l'œsophage. La moitié des cas est diagnostiquée après 75 ans (Binder-Foucard, 2013). Le taux d'incidence annuel standardisé est estimé à 10.5 pour les hommes et de 4 pour les femmes pour 100 000 personnes. Concernant le Finistère, on relevait 168 nouveaux cas de cancer de l'estomac en 2012, se situant alors au premier rang des départements français.

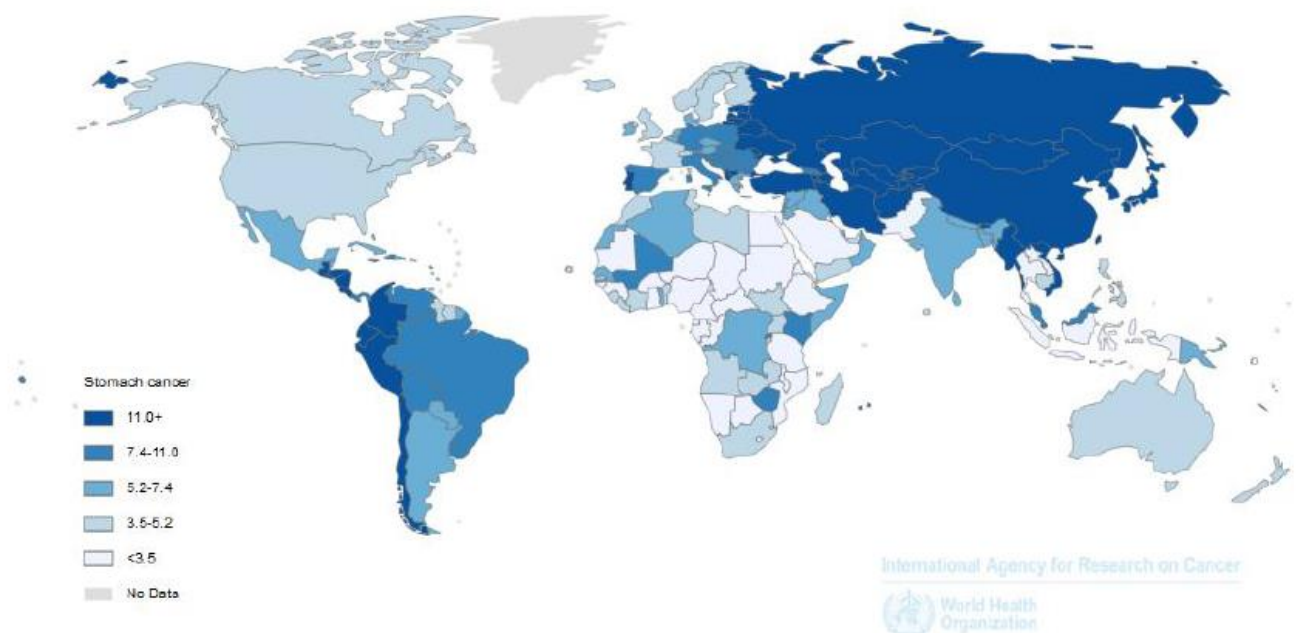


Figure 9 : Répartition géographique des cancers de l'estomac dans le monde selon l'incidence pour les deux sexes ([globo-can.iarc.fr](http://globo-can.iarc.fr)).

## B. Mortalité du cancer gastrique

Dans le monde, le cancer de l'estomac est la deuxième cause de mortalité par cancer en 2016 soit 834 171 décès (environ 8%), après le cancer du poumon (1,71 million de décès, environ 20%), et devant le cancer colorectal (829 557 décès) et le cancer du foie (828 945 décès). Le cancer de l'œsophage est le septième cancer à l'échelle mondiale avec 414 890 décès. Le nombre de décès dû au cancer de l'estomac est estimé à 7421 en France, et 5 159 dû au cancer de l'œsophage en 2017.

La survie estimée à 5 ans est de 25% pour le cancer de l'estomac et de 50% pour le cancer de l'œsophage (INCa, 2010). Le pronostic de ce cancer est particulièrement mauvais avec 30% de stades métastatiques au diagnostic.

## C. Facteurs de risques

Les facteurs de risques de cancers gastriques sont généralement classés du plus au moins important mais dans la plupart des cas il est impossible de les classer avec une certitude absolue. Nous les diviserons donc en risques connus ou certains, risques possibles et risques inconnus.

## 1. Facteurs de risques connus

Des études ont montré que le cancer gastrique faisait partie des cancers dont le risque est augmenté par :

- L'infection à *Helicobacter pylori* : *H. pylori* est une bactérie observée fréquemment dans l'estomac. Elle peut causer de l'inflammation, des ulcères et des changements dans les cellules du revêtement de l'estomac.
- L'infection au virus d'Epstein-Barr pourrait faire augmenter le risque d'adénocarcinome gastrique (type le plus fréquent), ainsi qu'accroître le risque d'apparition de lymphome non hodgkinien dans l'estomac (type rare). Dix % des cancers gastriques seraient positifs pour EBV.
- Le tabagisme : Certaines études ont démontré que l'exposition à la fumée secondaire fait aussi augmenter le risque de cancer de l'estomac.
- Des antécédents familiaux de cancer de l'estomac : Cette hausse du risque peut être attribuable à des troubles génétiques héréditaires. Elle peut aussi être causée par des facteurs de risque communs à d'autres membres de la famille comme l'alimentation, le tabagisme et les infections au *H. pylori* (Hartgrink et al, 2009)
- Certains troubles génétiques héréditaires de prédisposition au cancer comme le cancer gastrique diffus héréditaire (CGDH), la polypose adénomateuse familiale (PAF), le syndrome de Lynch (aussi appelé cancer colorectal héréditaire sans polypose, ou HNPCC), le syndrome de Peutz-Jeghers, le syndrome de polypose juvénile, le syndrome de Li-Fraumeni ou des mutations des gènes *BRCA*.
- Certains troubles d'estomac comme la gastrite atrophique chronique, la métaplasie intestinale, la dysplasie épithéliale gastrique, l'adénome, ou polype adénomateux, la maladie de Ménétrier ou l'anémie pernicieuse.
- Le Reflux gastro-œsophagien : trouble caractérisé par la remontée du contenu de l'estomac, y compris de l'acide gastrique, dans l'œsophage, ce qui cause des « brûlures d'estomac » et des douleurs dans la partie supérieure de l'abdomen ou au thorax. L'acide gastrique peut endommager le tissu de l'œsophage.
- Une chirurgie gastrique antérieure aussi appelée gastrectomie partielle pour le traitement des ulcères de l'estomac

- Une exposition à des rayonnements ionisants : Les personnes qui ont été exposées aux rayonnements ionisants au Japon à la suite de l'explosion d'une bombe atomique pendant la Seconde Guerre mondiale risquent davantage d'être atteintes d'un cancer de l'estomac, en particulier si elles étaient enfants lors de l'exposition à la radiation. Des études ont démontré que les personnes exposées aux rayonnements émis lors d'accidents nucléaires ou à des retombées nucléaires risquent plus d'être atteintes d'un cancer de l'estomac.
- Le travail dans l'industrie du caoutchouc : par une haute exposition à des produits chimiques cancérogènes.

## 2. Facteurs de risques possibles

Des études sont actuellement en cours pour rechercher l'association éventuelle entre le développement du cancer gastrique et :

- La consommation d'alcool
- La consommation de viandes fumées, traitées ou transformées (charcuterie) montre une augmentation du risque avec l'augmentation de la consommation, mais le World Cancer Research Fund International (WCRFI) et l'American Institute for Cancer Research (AICR) considèrent que plusieurs de ces études sont de mauvaise qualité. Ainsi, il est considéré que le lien est possible et non probable.
- L'obésité peut accroître le risque de cancer de l'estomac, en particulier dans le cardia.
- Une exposition professionnelle au plomb / à l'amiante
- Fibrose kystique

## D. Diagnostic et traitement du cancer gastrique

### 1. Diagnostic

Les symptômes du cancer gastrique ne sont pas spécifiques de ce cancer et ne permettent donc pas à eux seuls son diagnostic, ce qui explique un diagnostic souvent tardif : fatigue, perte de l'appétit, douleurs épigastriques, perte de poids, nausées, vomissements, anémie. Les hémorragies digestives et la perforation de l'estomac sont des complications possibles qui peuvent révéler la maladie mais ces événements sont relativement rares.



La localisation de la tumeur peut aussi être responsable de symptômes spécifiques : un cancer du cardia peut entraîner une gêne à la déglutition (dysphagie), tandis que le cancer du pylore peut engendrer un rétrécissement local responsable de douleurs et de vomissements.

Différents examens sont réalisés pour confirmer la présence d'un cancer, le localiser précisément et en définir le type :

- L'examen clinique de l'abdomen permet d'identifier tout gonflement, toute douleur anormale de l'abdomen. Il s'agit notamment de rechercher une masse abdominale ou des adénopathies sus-claviculaires (présence de ganglions lymphatiques de taille augmentée situés au-dessus de la clavicule), ou encore les signes d'une anémie.
- L'endoscopie œsogastrique (ou gastroscopie ou fibroscopie) sert à rechercher des signes éventuels d'inflammation ou un cancer de l'estomac en examinant la muqueuse de l'œsophage, l'estomac et la première partie de l'intestin grêle. En cas de suspicion de cancer, des biopsies sont effectuées.
- La biopsie permet de prélever des fragments de tissus afin de réaliser leur analyse macro et microscopique.
- L'examen anatomopathologique a pour but d'établir le diagnostic, de préciser le type de cancer, notamment grâce à un test moléculaire visant à rechercher la présence de certains biomarqueurs cliniques comme une surexpression de la protéine HER2 et ainsi orienter le choix des traitements

## 2. Biomarqueurs cliniques

### ▪ Le statut HER2

Le gène *HER2* est situé sur le chromosome 17 et code la protéine HER2, un récepteur membranaire de type tyrosine kinase. HER2 fait partie de la famille des récepteurs membranaires ubiquitaires appelés erbB ou HER (HER1, HER2, HER3, HER4) du facteur de croissance EGF dont le développement et le fonctionnement entraîne des signaux prolifératifs et antiapoptotiques (Graus-Porta et al, 1997). Ce récepteur est à l'origine de différents ciblage thérapeutiques notamment par des anticorps monoclonaux ou inhibiteurs de tyrosine kinase neutralisant ainsi son activité. Ce ciblage ne peut être efficace que si la tumeur est dépendante (addiction oncogénique) de la voie her2, c'est à dire présentant une amplification du gène

*her2* (scores dits 3+ ou 2+ amplifiés objectivés par hybridation *in situ* fluorescente (FISH)). Quinze à 25 % des adénocarcinomes gastriques se caractérisent par une surexpression et/ou une amplification d'HER2. Et c'est un essai de phase III ToGa qui a déterminé la nécessité de déterminer le statut HER2 comme facteur prédictif de réponse au traitement en comparant une chimiothérapie (cisplatine + 5-FU) associée au trastuzumab (anticorps anti-HER2) à une chimiothérapie seule, dans les cancers gastriques localement avancés et métastatiques (Bang et al, 2010).

Ainsi la recherche du statut HER2 dans les cancers gastriques et la jonction œsogastrique est devenue indispensable et fait partie des tests de routine en anatomie pathologique. Dans le cancer gastrique, tout comme dans le cancer du sein, immunohistochimie et hybridation *in situ*, respectivement pour la surexpression et l'amplification d'HER2, vont être les témoins de cette dépendance et autoriser un traitement par anticorps monoclonal anti-HER2 (Chapiro, 2012).

Les techniques et le système d'évaluation permettant la détermination du statut HER2 dans les adénocarcinomes gastriques et de la jonction œsogastrique s'appuient sur certaines particularités liées au statut HER2 d'après l'étude ToGA :

- Le statut HER2 est fréquemment hétérogène au sein d'un adénocarcinome gastrique ou de la jonction œsogastrique.
- HER2 est souvent plus surexprimé dans les formes intestinales de cancers gastriques que dans les adénocarcinomes diffus.
- Les adénocarcinomes de la jonction œsogastrique présentent des taux de surexpression de HER2 plus élevés que les adénocarcinomes gastriques (33 versus 21 % selon dans l'étude ToGA).

Selon certains auteurs, cette hétérogénéité en terme d'amplification peut concerner également la tumeur primitive et ses métastases correspondantes. De plus, il est possible que des traitements néoadjuvants par chimiothérapie puissent conduire à une négativité du statut HER2 et entraîner une disparité d'expression (statut positif sur biopsie et négatif sur pièce opératoire post-traitement) (Watson et Louvet, 2014).

- **Formes héréditaires de cancers gastriques**

Les formes héréditaires du cancer de l'estomac sont souvent associées aux adénocarcinomes distaux de type diffus. Dans la majorité des cas, cette forme héréditaire serait le résultat d'une altération constitutionnelle du gène *CDH1* (Jolissaint et al., 2011). Ce gène code une protéine d'adhésion, la E-Cadhérine localisée au niveau des cellules épithéliales. Des études ont montré qu'une mutation constitutionnelle du gène *CDH1* engendrait un risque accru de cancer de l'estomac chez 67% des hommes et chez 83% des femmes (Pharoah et al., 2001). Cette forme familiale est le plus souvent retrouvée chez le sujet jeune (< 40 ans), alors que les formes sporadiques sont plutôt retrouvées chez le sujet âgé.

- **Biomarqueur émergent : C-met**

C-met correspond au récepteur de l'hépatocyte growth factor (HGF). Il possède des propriétés anti-apoptotiques, de prolifération et de dissémination des cellules tumorales. C-met et son ligand HGF font actuellement l'objet d'essais thérapeutiques par ciblage d'anticorps monoclonaux et inhibiteurs de tyrosine-kinase. Les premières données indiquent que les patients porteurs d'adénocarcinomes gastriques présentant une expression élevée de C-met répondraient de façon plus efficace aux anticorps thérapeutiques anti-HGF (Inokuchi et al, 2015).

## **E. Prise en charge thérapeutique du cancer gastrique**

### **1. Classique**

- **Chirurgie**

L'acte chirurgical constitue l'essentiel du traitement dans le cas du cancer de l'estomac. En fonction de la localisation de la tumeur, une gastrectomie totale est privilégiée pour les tumeurs distales et une gastrectomie partielle pour les tumeurs proximales. En plus de la gastrectomie, un curage ganglionnaire est en général réalisé, permettant de retirer tous les ganglions lymphatiques proches de l'aire ganglionnaire de l'estomac.

- Chimiothérapie et radiothérapie (pré et post chirurgicale)

Des études ont montré que la pratique de l'acte chirurgical seul engendre un fort taux de récurrence tumorale (Marrelli et al., 2005). C'est pourquoi différentes approches de chimiothérapie péri-opératoire (2 mois avant et après la chirurgie) ou de chimiothérapie et de radiothérapie post chirurgicale sont envisagées dans le traitement du cancer de l'estomac. Le traitement des métastases à distance de la tumeur primaire ne va généralement pas faire appel à l'acte chirurgical mais plutôt à des chimiothérapies dites palliatives.

- Chimiothérapies

Les stratégies thérapeutiques sont différentes selon la classification TNM (source/description); la haute autorité de santé préconise un traitement endoscopique pour les stades T1aN0M0, une chirurgie seule pour les stades T1-2N0M0, une chimiothérapie adjuvante pour tous les autres stades non métastatiques et enfin une chimiothérapie palliative pour les stades métastatiques.

En règle générale, les chimiothérapies dites combinées sont plus efficaces, en terme de survie globale, que les chimiothérapies uniques (Wagner et al., 2006). Parmi les différentes combinaisons de molécules, on retrouve les platines (cisplatine et oxaliplatine), les fluoropyrimidines (5-fluorouracile, S-1 et capécitabine), l'irinotécan et les taxanes (docétaxel et paclitaxel).

*a. Les platines*

Les sels de platine se comportent comme des agents alkylants. En effet, ils se lient à l'ADN et forment un complexe platine/ADN ou adduit qui empêche la réplication de l'ADN engendrant ainsi la mort par apoptose de la cellule. Il a été montré également que les sels de platines, notamment le cisplatine et l'oxaliplatine, pouvaient interagir avec des protéines cytoplasmiques. Ces sels vont induire par exemple la translocation de la calréticuline à la surface de la cellule. Cette exposition sert alors de signal d'interaction avec les cellules dendritiques et les macrophages (Hato et al., 2014).

De nombreux mécanismes de résistance ont été décrits vis-à-vis des sels de platine, tels qu'une augmentation de glutathion (se lie aux sels de platine), une capacité de réparation augmentée, ainsi qu'une production accrue de métallothionéines.

#### *b. Les fluoropyrimidines*

La famille des fluoropyrimidines fait partie de la classe thérapeutique des anti-métabolites. Le 5-fluorouracile (5-FU), le plus couramment utilisé, interagit avec l'ADN et l'ARN. En effet, l'uracile catabolisée par la thymidilate synthétase est le précurseur de la thymine, via sa méthylation, qui est nécessaire à la synthèse du brin d'ADN. Le 5-FU métabolisé en 5-fluorodéoxyuridine 5'-monophosphate (FdUMP) va alors inhiber la thymidilate synthétase empêchant ainsi la méthylation de l'uracile. Il va pouvoir également entrer en compétition directement avec l'uracile et être incorporé dans l'ARN, ayant pour conséquence la mauvaise traduction des protéines. Tous ces effets anti-métaboliques vont donc être à l'origine de l'arrêt du cycle cellulaire et de l'entrée en apoptose des cellules.

#### *c. L'irinotécan*

L'irinotécan est un inhibiteur de l'ADN topoisomérase de type I. Cette enzyme intervient lors de différents processus tels que la réplication, la transcription, la réparation de l'ADN, la condensation des chromosomes ainsi que leur ségrégation lors de la mitose. Son action est de créer des coupures simples brin de l'ADN avant de les religuer par la suite. L'irinotécan et les autres inhibiteurs de topoisomérase empêchent cette religation, ce qui aboutit à l'apoptose des cellules.

#### *d. Les taxanes*

Les taxanes (paclitaxel et docétaxel) sont issus initialement de l'écorce de l'if (*Taxus baccata*). L'effet majeur des taxanes est de bloquer la division cellulaire en perturbant le fuseau mitotique. Les taxanes favorisent l'assemblage des tubulines et les stabilisent inhibant leur dépolymérisation et ainsi la dynamique des microtubules (Ringel and Horwitz, 1991). Cette perturbation dynamique est à l'origine du blocage du cycle cellulaire en phase G2/M (Fabbri et al., 2006). Le docétaxel (Taxotère®) a montré d'autres effets : au niveau de l'apoptose en phosphorylant la protéine Bcl-2, entraînant la dégradation de cette protéine anti-apoptotique (Berchem et al., 1999), ou encore au niveau de l'angiogenèse induite par les cellules cancéreuses (Sweeney et al., 2001). Néanmoins, de nombreuses études ont montré l'acquisition d'une résistance au docétaxel par des mécanismes d'échappement développés par les cellules cancéreuses (Seruga et al., 2011).

#### *e. Les essais cliniques*

De nombreux essais cliniques de phase III ont été réalisés à ce jour, que ce soit pour des chimiothérapies adjuvantes ou des chimiothérapies palliatives pour les stades avancés, dont la survie globale (OS) est de 9,5 mois.

## **2. De nouvelle génération (immunothérapie)**

Les thérapies ciblées reposent sur le principe d'une bonne connaissance des mécanismes moléculaires impliqués dans le processus de tumorigenèse. Parmi les voies de signalisation les plus ciblées, on retrouve la voie des facteurs de croissance épidermique (EGF et HER) et la voie du facteur de croissance de l'endothélium vasculaire (VEGF) (Riquelme et al., 2015). Ces deux voies sont les principales cibles d'anticorps monoclonaux dirigés contre les récepteurs ou les ligands. Cette fixation d'anticorps entraîne donc l'inhibition de ces voies de signalisation impliquées dans la prolifération cellulaire, l'angiogenèse, le blocage de l'apoptose ainsi que l'invasion cellulaire. De nombreux essais cliniques utilisant ces anticorps monoclonaux sont décrits dans la littérature (Riquelme et al., 2015). Cependant les thérapies ciblées n'ont pas démontré, à ce jour, un taux de survie supérieur à celui de la chimiothérapie. Seul le trastuzumab semble montrer une réelle efficacité, limitée uniquement aux cancers gastriques surexprimant la cible HER2 (ERBB2) qui ne représentent au maximum que 20 % des cas.

D'autres formes de thérapies ciblées émergent également sous l'influence du projet The Cancer Genome Atlas qui a permis d'identifier les différentes voies de signalisation impliquées dans le processus de tumorigenèse, et d'en comprendre leurs rôles. Toutes ces nouvelles thérapeutiques portent le suffixe «mab» (anti-corps monoclonal) et ciblent une voie particulière impliquée dans le processus cancéreux (Figure 8), comme précédemment cité pour C-met et son ligand HGF.

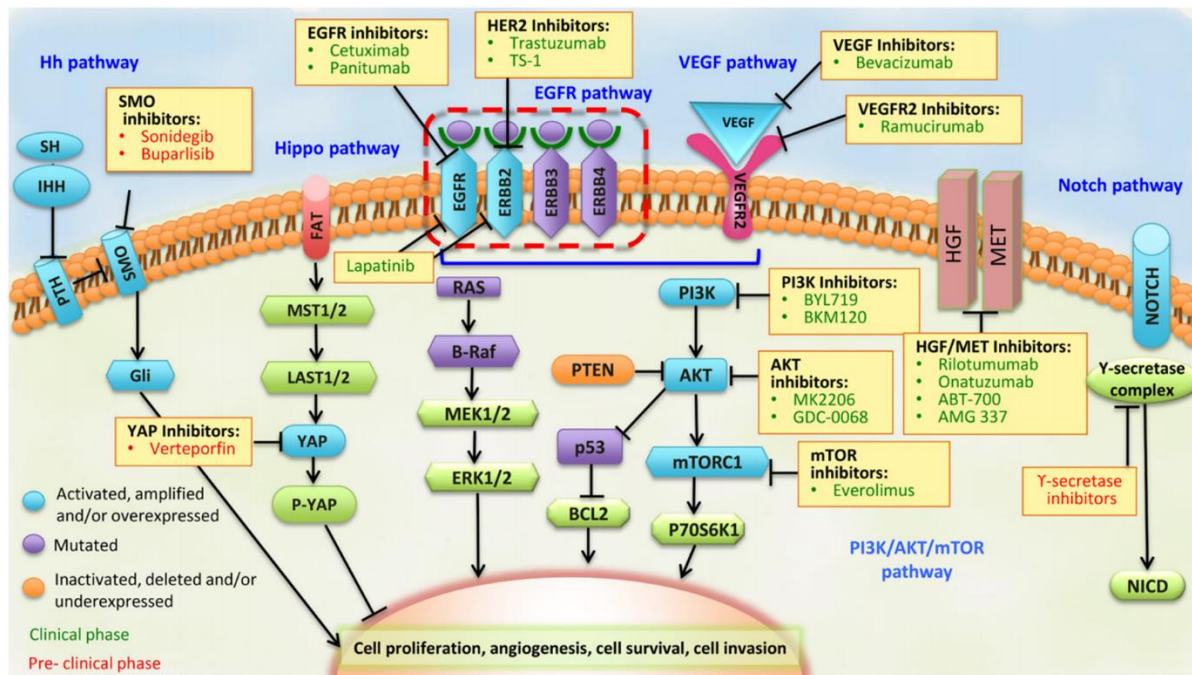


Figure 10 : Voies moléculaires ciblées par les approches de thérapie ciblée utilisant des anticorps monoclonaux «mab» (Riquelme et al., 2015).

# Chapitre 4 Oncogenèse MSI

## A. Découverte du phénotype tumoral MSI

Le mécanisme d'instabilité des microsatellites est décrit pour la première fois en 1993, dans certains cancers colorectaux sous la forme de « mutations somatiques ubiquitaires » dans des séquences homopolymérique (Perucho et al, 1993), (Aaltonen et al, 1993) (Thibodeau et al, 1993).

Le mécanisme observé dans ces tumeurs coliques avec instabilité des régions microsatellites est similaire au défaut du mécanisme MthLS chez *Escherichia Coli*, il est décrit comme impliqué dans le contrôle de la réplication de l'ADN. Il a ensuite été montré que le système équivalent au mécanisme MthLS chez l'Homme était orchestré par les gènes MLH1 et MSH2 situés sur les chromosomes 2 et 3, respectivement (Peltomaki et al, 1993 et Lindblom et al, 1993). Ce système fonctionne par l'intermédiaire de complexes protéiques qui sont impliqués dans le système MMR (mismatch repair). Conséquence de la déficience du système de réparation de l'ADN dans les tumeurs, l'accumulation des mutations se produit sous la forme de délétions et d'insertions au niveau des séquences microsatellites (Fishel et al, 1993 ; Leach et al, 1993 ; Bronner et al, 1994).

Ce phénotype tumoral, nommé RER (Replication Errors) ou MIN (microsatellite instability) dans un premier temps a été renommé phénotype MSI en 1997.

## B. Le système MMR et les conséquences de sa déficience

### 1. Fonctionnement du système MMR

Dès 1953, Watson et Crick proposent, après avoir découvert l'architecture de la molécule d'ADN, un modèle semi-conservatif de la réplication. Confirmé quelques années plus tard, l'interprétation de ce modèle indique que lors de la réplication d'une molécule mère d'ADN, celle-ci en donne deux et chacun de ces nouveaux brins est formé à partir d'un brin matrice issu de la molécule d'ADN mère. La réplication de l'ADN est assurée par l'ADN polymérase.



Cet enzyme réplique l'ADN à environ 50 nucléotides par seconde. Lors de la réplication, des erreurs surviennent à la fréquence moyenne d'1 erreur toutes les  $10^5$  paires de bases. Mais l'ADN polymérase est douée d'une capacité de relecture et de correction de ces erreurs grâce à son activité 3'>5' exonucléase. Le taux d'erreur moyen peut alors être diminué d'un facteurs 100 (Stojic et al, 2004 ; Kunz et al, 2009).

Malgré tout, il persiste des erreurs qui passeraient inaperçues pour l'ADN polymérase. Un second système joue alors le rôle de contrôle la qualité de la réplication : le système MMR. Il est composé de protéines qui, en s'associant, vont identifier et réparer directement sur l'ADN les erreurs de mésappariement suite à l'insertion ou la délétion (création de boucles) ou de substitution de nucléotides. C'est un système très conservé, depuis les bactéries jusqu'aux mammifères. Appelé système MutHLS chez les bactéries, il est composé de trois protéines principales MutS, MutL et MutH. Chez les mammifères, ce système est composé de MSH2, MSH3, MSH4, MSH5, MSH6, homologue de MutS et MLH1, MLH2, MLH3, PMS1 et PMS2, homologues de MutL. Il n'existe aucun homologue de MutH chez les mammifères.

Les mésappariements de l'ADN sont reconnus par les complexe MutS $\alpha$  (MSH2-MSH6) ou MutS $\beta$  (MSH2-MSH3) en fonction de la structure du mésappariement. Les mésappariements dûs à l'insertion ou la délétion de 1 à 2 nucléotides (création de structure en forme de boucles) sont pris en charge par MutS $\alpha$ , les insertions ou délétions supérieures à 2 nucléotides sont prises en charge par MutS $\beta$ , pendant que l'autre brin est protégé vis-à-vis des exonucléases par la protéine RPA (Replication protein A1).

Après l'identification de ces mésappariements, le complexe MutL $\alpha$  formé par MLH1 et PMS2 subit une modification de sa conformation spatiale grâce à une activité ATP dépendante permettant sa mobilisation au niveau du mésappariement. MutL $\alpha$  va alors interagir avec la protéines PCNA (Proliferating Cell Nuclear Antigen) ce qui entrainera l'activation de l'activité endonucléasique de PMS2 afin de rétablir la séquence d'origine. L'exonucléase EXO1 est ensuite recrutée afin d'exciser le mésappariement. L'ADN polymérase est de nouveau mobilisée afin de synthétiser le/les nucléotide(s) manquant correspondant (Li et al., 2016).

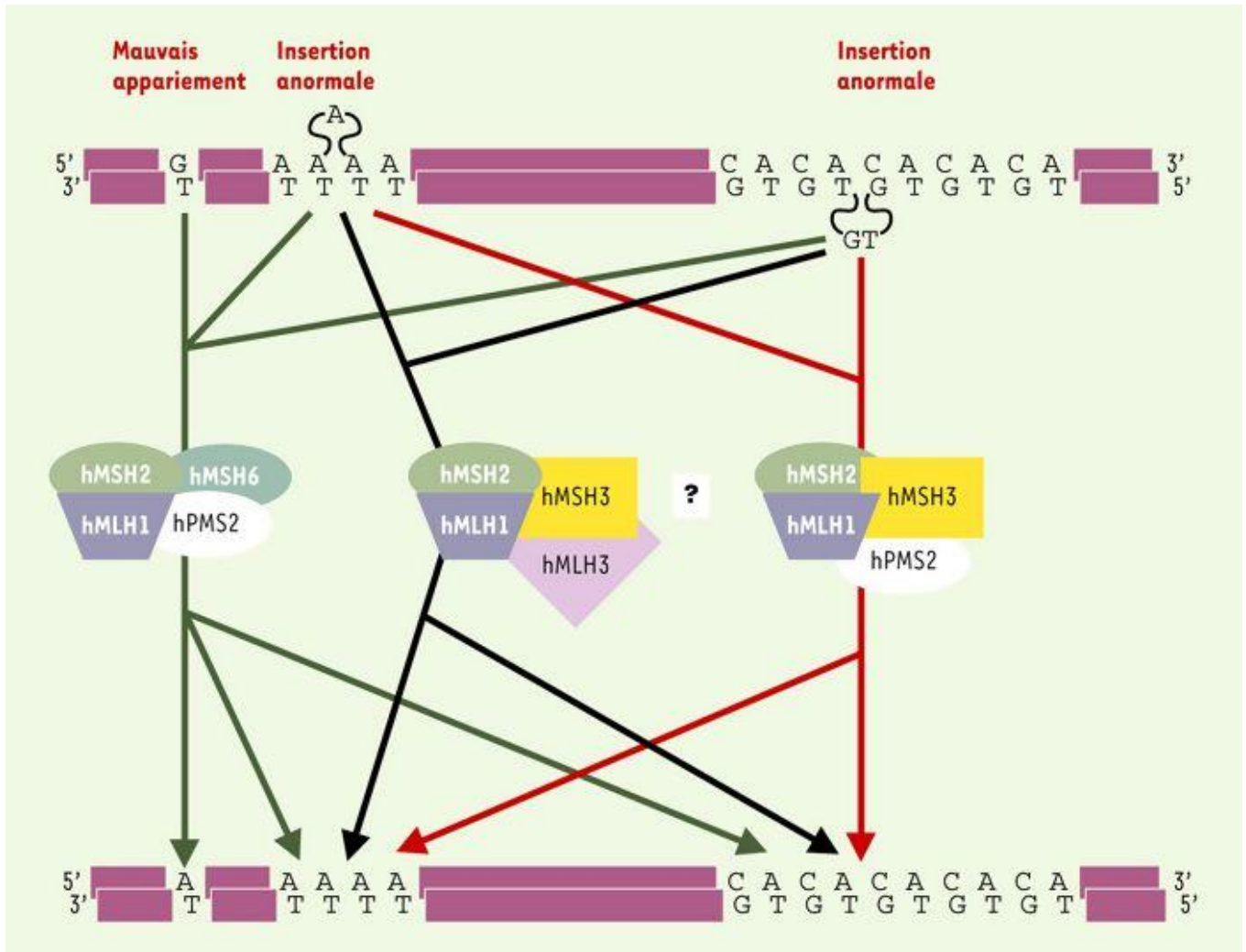


Figure 11 : Les différents complexes du système MMR chez l'homme.

Selon le type de mésappariement différents complexes de réparation vont être mis en place. Les complexes associent MSH2 à MSH6 ou MSH3 et MLH1 à PMS2 ou MLH3. Chaque complexe MMR est impliqué dans la reconnaissance et la réparation des mésappariements de la molécule d'ADN en fonction de leur nature (mauvais appariement, insertion anormale d'un ou de plusieurs nucléotides) (Duval et Hamelin, 2003).

## 2. Les conséquences de l'inactivation du système MMR

Le phénotype MSI est induit par la déficience du système MMR dû à l'inactivation bi-allélique d'un ou de plusieurs gènes de ce système. La majorité des cancers MSI sporadiques trouvent leur origine dans une inactivation épigénétique du promoteur de MLH1 (Kane et al, 1997 et Veigl et al, 1998). Dans les cas héréditaires, ce sont majoritairement MSH2 et MLH1 qui sont inactivés suite à une mutation germinale (Fishel et al, 1993, Nicolaides et al, 1994 et Bronner et al, 1994). Ces inactivations induisent alors une accumulation importante de mutations dans les séquences microsatellites (séquences formées par une répétition continue de motifs composés de 1 à 4 nucléotides).

Ces régions microsatellites sont difficiles à reproduire à l'identique lors de la réplication de l'ADN. Il est établi que dans le phénotype MSI, c'est l'accumulation de mutations secondaires, dans des gènes en lien avec l'oncogenèse qui est le phénomène transformant et non pas la déficience du système MMR en lui-même (Duval et Hamelin, 2003).

- **Les microsatellites non codants**

Une grande partie des mutations affectant les séquences microsatellites est localisée au niveau des régions non codantes inter- ou intra-géniques (introns, régions 5' et 3' UTR). Les régions UTR peuvent ainsi posséder des séquences répétées modifiées dues à la déficience du système MMR. Ainsi, dans le cas de tumeurs MSI, l'expression de certains gènes peut être impactée par la présence de ces mutations dans ces régions régulatrices et ainsi être impliquée dans le développement de ces tumeurs. Des études récentes ont en effet signalé que l'expression de plus de 130 gènes étaient altérée alors que des délétions de 1 ou plusieurs nucléotides étaient identifiées dans leurs régions UTR (Yoon et al, 2013).

Par exemple, une instabilité au niveau d'un microsatellite localisé dans le promoteur du gène MMP-3 (Matrix MetalloProteinase-3) a été associée à une perte d'expression de la protéine correspondante dans des tumeurs colorectales MSI (Morán et al, 2002).

Très peu d'études se sont intéressées aux conséquences de mutations situées dans des régions introniques. Néanmoins, nous pouvons citer quelques situations où ces mutations ont un impact sur la physiologie tumorale, touchant les gènes MRE11, ATM ou encore HSP110. Une description plus détaillée de ces phénomènes sera effectuée dans le chapitre consacré à l'épissage alternatifs et ses altérations dans les cancers MSI.

- **Les microsatellites codants**

Environ 20% des gènes humains contiennent au niveau de leur région codante une répétition microsatellitaire supérieure ou égale à 7 nucléotides (El-Bchiri et al, 2008), ce qui rend très élevé le nombre potentiel de gènes cibles de l'instabilité microsatellitaire, aussi appelés gènes cibles MSI. Les altérations au niveau de ces répétitions entraînent majoritairement un décalage du cadre de lecture et l'apparition d'un codon stop prématuré au niveau du transcrit

messenger. Dans la majorité des cas, une inactivation fonctionnelle de la protéine correspondante est observée, conférant un avantage sélectif à la cellule tumorale MMR-déficiente (Woerner et al, 2005).

En 2002, un modèle a été suggéré afin d'associer la fréquence mutationnelle au rôle du gène cible dans la cellule tumorale. Ainsi, une classification des gènes cibles MSI a été proposée incluant les gènes dits «transformers», dont les mutations sont soumises à une pression de sélection positive et qui donc présentent un taux de mutation élevé et, d'autre part, les gènes dits «survivors», soumis à une pression de sélection négative et qui donc seraient mutés à très basse fréquence dans les cellules tumorales MSI (Duval et al, 2002).

Une grande partie des gènes possédant une région microsatellite peuvent être considérés comme des points chauds de mutations dans les tumeurs MSI. Malgré tout, il semble que la fréquence mutationnelle varie en fonction de certains paramètres tels que l'organe touché ou la contribution du gène à l'oncogenèse (Duval et al, 2002).

Divers travaux ont mis en avant la présence de mutations au niveau de régions microsatellites codantes dans divers cancers (côlon, estomac, endomètre). On notera la présence de mutations affectant des gènes impliqués dans l'apoptose (*APAF1*, *BAX*, *CASP5*, ...), la prolifération cellulaire (*AXIN2*, *IGF2R*, *LMAN1*, *PTEN*, ...) ou encore la signalisation et la réparation des dommages à l'ADN (*ATM*, *BLM*, *MLH3*, *MSH6*, ...) (Imai et Yamamoto, 2008 et Duval et al, 2002).

Phénomène intéressant, l'oncogène MYB possède une séquence de 19 thymines située dans l'intron 1 qui a été identifiée comme une séquence régulatrice de sa propre élongation lors de la transcription de ce gène. Il a également été montré que ce gène était surexprimé dans les tumeurs MSI et potentiellement cela peut être associé à un mauvais facteur pronostic (Biroccio et al, 2001 et Hugo et al, 2006).

## C. Le syndrome de Lynch

### 1. Découverte et définition

Le syndrome de Lynch est décrit pour la première fois en 1913 par Aldred Scott Warthin qui voit là une forme de prédisposition génétique au développement de tumeurs de plusieurs sites (colon, estomac, endomètre). Bien plus tard, c'est le docteur Henry Lynch qui décrira plus précisément ce syndrome et l'appellera « syndrome des cancers familiaux ». Il fut ensuite renommé HNPCC (hereditary non-polyposis colorectal cancer) puis syndrome de Lynch afin d'inclure tout type de cancer, colique ou non.

Le syndrome de Lynch est une forme de cancer colorectal héréditaire sans polypose, de transmission autosomique dominante ou HNPCC (Olschwang et al, 2004). Ce syndrome augmente le risque de développer certains types de cancers, comme les cancers du côlon et du rectum mais aussi de l'estomac, l'intestin grêle, le foie, la vésicule biliaire, les voies urinaires, le cerveau ou encore la peau. Il représente environ 5% des cancers colorectaux soit environ 10% des cas de CCR MSI (Hampel et al, 2005).

### 2. Caractéristiques génétiques

Le syndrome de Lynch est maladie autosomique dominante, qui résulte de mutations germinales dans un allèle portant sur un des gènes du système de réparation des mésappariements de l'ADN (majoritairement MLH1 et MSH2), le deuxième évènement à l'origine du dysfonctionnement de ce système sera somatique. On signalera aussi l'existence du syndrome cMMRD (Congenital mismatch repair deficiency) qui se distingue par des mutations germinales des 2 allèles d'un même gène MMR. D'après les informations recueillies dans les bases de données (*InSIGHT et ICGC*), ces mutations sont majoritairement des substitutions d'un seul nucléotide à l'origine de l'apparition de codon stop. On peut également trouver des mutations de *MSH6* ou *PMS2* mais elles sont plus rares (moins de 10% des cas).

Pour les personnes atteintes du syndrome de Lynch, le risque de développer un cancer colorectal au cours de sa vie est de l'ordre de 10% à 50 ans et 40% à 70 ans (ERISCAM, 2015).

En effet, c'est environ à cette période que la mutation somatique sur le deuxième allèle peut intervenir de manière aléatoire, induisant alors une inactivation de système MMR, une augmentation du taux de mutations non prises en charge par le système de réparation, devenu défaillant, conduisant possiblement à une prolifération cellulaire non contrôlée susceptible d'entraîner un développement tumoral.

Plus récemment, il a été observé, chez deux patients non apparentés atteint de HNPCC, l'absence de mutations sur les gènes MMR. En revanche, il a été montré que le promoteur de MLH1 était hyperméthylé entraînant une diminution de son expression de ce gène. Il s'avère que l'hyperméthylation de ce promoteur est constitutionnelle car retrouvée dans les trois feuilletts embryonnaires indiquant alors que l'épimutation est constitutionnelle et serait apparue *de novo* dans la lignée germinale de l'un des parents, ou très tôt au cours du développement de l'individu (Suter et al, 2004).

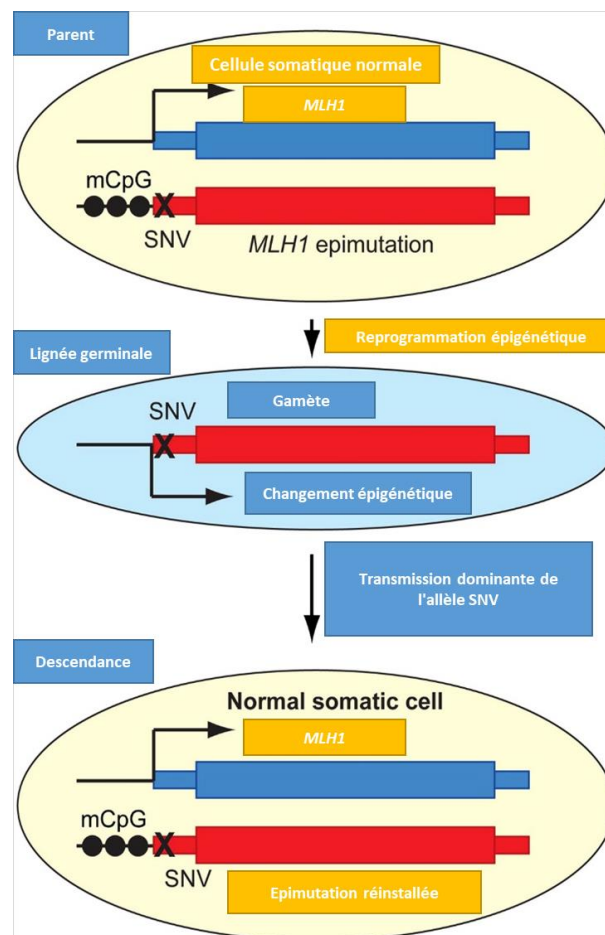


Figure 12 : Mécanisme alternatif de prédisposition au cancer lié à des épimutations constitutionnelles d'après Suter et al, 2004. Certaines mutations sont causées par des altérations génétiques sous-jacentes agissant en cis à proximité du gène affecté, elles présentent une transmission autosomique dominante.

### 3. Caractéristiques épidémiologiques du Syndrome de Lynch

Comme dit précédemment, le syndrome de Lynch expose à un risque majeur de développer un CCR autour de 50 ans. Il est également connu que le gène impliqué a son importance dans l'expression de ce syndrome. En effet, les individus porteurs d'une altération germinale du gène *MLH1* ou *MSH2* ont généralement un risque plus élevé que ceux porteurs de mutations sur *MSH6* ou *PMS2* et le plus large spectre clinique de cancers liés au syndrome de Lynch.

De plus, certaines études ont montré que les mutations touchant le gène *MSH2* participeraient davantage au développement de tumeurs hors colon/rectum (Ramsoekh et al, 2009 ; Lin-Hurtubise et al, 2013). Il a aussi été montré que les mutations touchant le gène *EPCAM* entraînerait un risque de développer un cancer colorectal comparable à celui de patients portant une mutation dans le gène *MSH2*. De la même manière, il semblerait que le développement de cancers gynécologiques soit favorisé dans le cas de mutations de *MSH6* (Baglietto et al, 2010 et Bonadona et al 2011). Enfin, il a été observé que seulement 25% des hommes et un tiers des femmes présentant une mutation du gène *PMS2* seraient diagnostiqués avec un cancer associé au syndrome de Lynch, soit un risque plus faible que les autres cas de développer un cancer colique ou gynécologique (Senter et al, 2008). De la même façon, un syndrome de Lynch augmente également le risque de développer un cancer gastrique (Kohlmann et Gruber, 1993).

Il a aussi été montré que les patients atteints du syndrome de Lynch survivraient plus longtemps que les patients atteints de tumeurs MSI sporadique. Néanmoins, cela pourrait s'expliquer simplement par la prise en charge plus précoce de ces cancers à forme familiale (Stigliano et al 2008).

De la même manière que les cancers colorectaux, la possibilité de développer un cancer gastrique associé au syndrome de Lynch semble être corrélée à une forme mutée du gène *MMR* ainsi qu'au sexe. Il semble donc plus probable de développer un cancer gastrique lorsque l'on est un homme et que les mutations touchent les gènes *MLH1* ou *MSH2*. Les femmes semblent en effet présenter un risque plus faible entre 10 et 50 % contre 20 à 70% pour les hommes (Kohlmann et Gruber, 1993).

Il est à noter que les syndromes de Lynch ne sont pas cantonnés aux cancers colorectaux, aux cancers gastriques ou aux cancers de l'endomètre. On observe également, mais de manière moins fréquente, des tumeurs des voies urinaires, des ovaires, du pancréas, de l'intestin grêle et du cerveau (Vasen et al, 1999).

#### **4. Critères d'identification du Syndrome de Lynch**

- Critères anatomopathologiques

La majorité des cas de syndrome de Lynch apparaissent au niveau de la partie du colon droit la plus proche de l'intestin grêle, le colon proximal. D'un point de vue histologique, ces tumeurs présentent un fort infiltrat lymphocytaire ainsi que des follicules lymphoïdes aux extrémités de ces tumeurs (type Crohn) (Umar et al, 2004). Ces tumeurs évoluent en grande majorité à partir d'adénomes préexistants avec une augmentation de leur taille et une progression rapide vers le carcinome en moyenne en l'espace de 3 ans contre 10 ans pour des tumeurs MSI sporadiques. Ces tumeurs présentent une architecture en bague de chatons et sont peu différenciées.

Les tumeurs gastriques sont divisées en deux types histologiques majoritaires (classification de Lauren) (Lauren, 1965) : le type intestinal est caractérisé par la présence de cellules cancéreuses bien différenciées et se développant lentement, alors que le type diffus présente des cellules cancéreuses peu différentes et beaucoup plus agressives (Capelle et al, 2010).

- Critères génétiques

Après 1993 et la découverte de mutations spécifiques au syndrome de Lynch, la recherche de ce phénotype MSI est réalisée en routine. Les critères de dépistage ont été élargis après la conférence de Bethesda en 1997, avec une analyse plus spécifique et précise des gènes touchés. Avant cela, l'identification des syndromes de Lynch était basée uniquement sur des données cliniques et familiales (critères d'Amsterdam). De nos jours, un criblage moléculaire systématique est mis en place pour identifier les cas qui échapperaient aux critères cliniques (Hampel et al, 2005).



## D. Les tumeurs MSI sporadiques

Contrairement au syndrome de Lynch, il est très peu probable de retrouver des mutations touchant les zones codantes des gènes MMR dans les cas sporadiques. Quarante-vingt-dix pourcent des cas de tumeurs sporadiques MSI sont dûs à une méthylation épigénétique du promoteur du gène MLH1, avec pour conséquence la forte diminution ou l'absence d'expression de la protéine correspondante (Kim et al, 2013). Il a été montré que l'inactivation de ce gène par l'hyperméthylation de son promoteur ne fonctionne que si cette méthylation se situe dans la région proximale du promoteur, mais pas dans la partie distale (Morán et al, 2002).

En outre, il est peu fréquent, dans ces cancers, d'observer des mutations sur les séquences codantes des gènes MMR tel que MLH1 et MSH2, mutations beaucoup plus fréquentes dans les cancers de formes familiales.

### 1. Cancers gastriques MSI sporadiques

Les cancers gastriques sporadiques MSI peuvent être caractérisés par d'autres altérations responsables de l'inhibition du système MMR, en plus de l'hyperméthylation du promoteur du gène MLH1. Il a été montré que la déficience du système MMR peut être la conséquence d'une infection par la bactérie *Helicobacter pylori*. Plusieurs travaux ont en effet montré que des lignées cellulaires cancéreuses gastriques infectées par cette bactérie présentent une diminution de l'expression des gènes MLH1, MSH2, MSH6 et PMS2 (Machado et al, 2009 ; Yao et al, 2006).

Les mêmes observations ont été faites chez des patients infectés par cette bactérie, en comparaison de patients non infectés (Mirzaee et al, 2008). Il a aussi été montré que l'élimination de cette bactérie entraînait un rétablissement de niveau d'expression de ces gènes dans les lignées cellulaires traitées (Di, 2005).

D'un point de vue clinique, plusieurs études ont évalué l'effet pronostic du phénotype MSI dans les CG, toutes montrant que le statut MSI semble être associé à une survie améliorée à tous les stades par rapport aux cancers gastriques de type MSS (Lee et al, 2002 ; Beghelli et al, 2006).

Ces résultats peuvent être expliqués en partie par les critères clinico-pathologiques des CG MSI. En effet, outre l'association avec le type histologique intestinal, caractérisé par une progression lente (Kim et al, 2011 ; Buhard et al, 2006), les cancers gastriques MSI sporadiques sont associés à une faible capacité de métastases ganglionnaires et généralement de bas grade (Shokal et Sharma, 2012).

## **2. Autres cancers MSI sporadiques**

Le développement des tumeurs dans les CCR MSI de type sporadique est majoritairement dû à une altération épigénétique de MLH1, mais il existe également des mutations touchant les gènes MMR qui peuvent être à l'origine de ces tumeurs. Dans certains cancers colorectaux et de l'endomètre, il a en effet été observé des mutations somatiques dans les gènes MMR ainsi que dans le gène codant l'ADN polymérase POLE, à l'origine d'un nombre important d'erreurs lors de la réplication de l'ADN (Mensenkamp et al, 2014 ; Haraldsdottir et al 2012). Il est à signaler que dans le cas de ces cancers, la seule présence de mutations de l'ADN polymérase n'est pas synonyme de développement tumoral ; il faut que celle-ci soit associée aux mutations des gènes MMR. Une étude a également montré que dans un peu plus de 75% des cas de cancers MSI sporadiques on retrouvait une hyperméthylation du promoteur et dans les 25% restant c'était des mutations somatiques touchant MLH1, MSH2, MSH6 ou encore PMS2 (Cancer Genom Atlas, 2012). De manière intéressante, on peut également trouver d'autres mécanismes de dérégulation du système de réparation des mésappariements de l'ADN : Une dérégulation de l'expression de ces gènes par le miR-155 dans des tumeurs MSI sans causes antérieures connues d'inactivation somatique du système MMR ou encore triméthylation de l'histone H3 sur la Lysine 36 (H3K36me3) nécessaire au recrutement du complexe MutS $\alpha$ . Ainsi, les cellules dépourvues de cet enzyme sont caractérisées par une instabilité de leurs régions microsatellites.

D'un point de vue clinique, on peut signaler que les tumeurs colorectales MSI sporadiques sont retrouvées plus fréquemment chez les femmes que chez les hommes (2 pour 1) de plus de 70 ans.

Enfin, d'un point de vue histologique, les tumeurs colorectales sporadiques MSI sont assez similaires aux cellules tumorales des cancers colorectaux associées au syndrome de Lynch : on observe des cellules mucineuses peu différenciées ainsi qu'une infiltration lymphocytaire. On peut également noter que de nombreuses études montrent que les cellules tumorales des cancers colorectaux MSI sporadiques se développent à partir d'adénomes sessiles, résultant d'une hyperplasie des cryptes (Jass et al, 2002 ; Mäkinen et al, 2007). Des études suggèrent également que les CCR MSI sporadiques se développent d'une manière différente des CCR LS (adénomes préexistants) (Chan et al, 2003 ; Xang et al, 2003).

Concernant la valeur pronostique du statut MSI, il a été montré que les patients atteints d'un des cancers colorectaux MSI ont un meilleur taux de survie que les patients MSS de même stade (Samowitz et al, 2001 ; Sinicrope et al, 2006). Ces résultats sont retrouvés dans une méta-analyse, de 32 études incluant 1277 tumeurs colorectales MSI parmi un total de 7642 patients tous stades confondus (Popat et Houlston, 2005). Il a aussi été montré que le meilleur pronostic dans les tumeurs MSI est plus lié à un stade précoce qu'à un stade avancé (Roth et al, 2012). En effet les tumeurs MSI sporadiques de stade 4 font exceptions à cette règle (pronostic plus sévère comparé aux tumeurs MSS de même stade), malgré le potentiel métastatique réduit des tumeurs MSI (Koopman et al, 2009 ; Malesci et al 2007).

Il est à noter également qu'une mutation récurrente (p.Val600Glu) située au niveau du gène BRAF est régulièrement retrouvée dans les cancers MSI sporadiques et que celle-ci est absente dans les tumeurs colorectales associées au syndrome de Lynch et cette mutation pourrait être à l'origine d'effets délétères dans ce contexte (Venderbosch et al, 2014 ; Goldstein et al 2014).

## **E. Détermination du statut MSI**

De par son aspect pronostique, la détermination de phénotype RER (Replicative Error) est aujourd'hui réalisée en routine. En effet, comme précédemment souligné, les tumeurs MSI sont associées à un meilleur pronostic comparé au phénotype MSS du fait de leurs caractéristiques clinico-pathologiques (immunogène, faible potentiel métastatique).

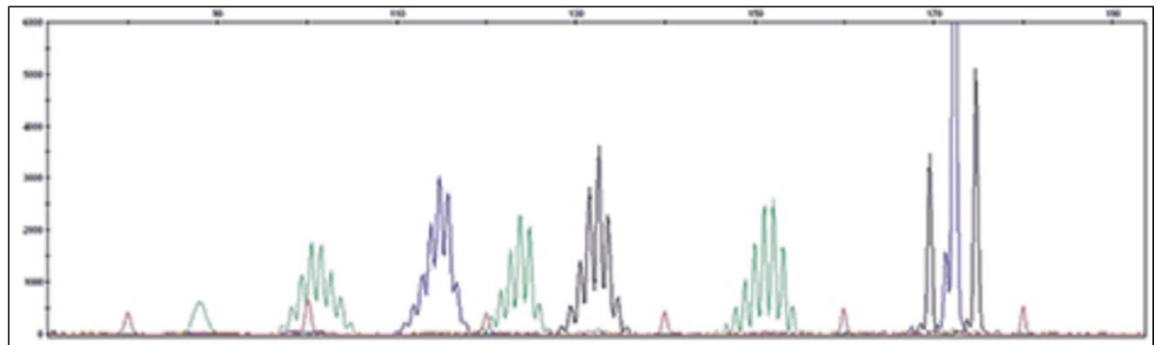
Deux techniques complémentaires sont utilisées en routine pour déterminer le phénotype RER : la structure des séquences microsatellites est analysée par PCR (réaction en chaîne par polymérase) et l'expression des protéines MMR par IHC (immunohistochimie).

## **1. Diagnostic par PCR**

Après la conférence de Bethesda en 1997, la caractérisation du statut MSI reposait sur l'analyse comparative des produits d'amplification par PCR, obtenus à partir de l'ADN tumoral et de l'ADN normal du patient. La PCR est réalisée afin d'amplifier 5 régions microsatellites: 2 marqueurs mononucléotidiques (BAT25 et BAT26) et 3 marqueurs dinucléotidiques (D5S346, D2S123 et D17S250). Les instabilités touchant au moins 2 des 5 marqueurs sont appelées MSI-H (high), et celles ayant une instabilité sur un seul marqueur, MSI-L (low). Les tumeurs à microsatellites stables (MSS) sont celles ne présentant d'instabilité sur aucun des 5 microsatellites. Les tumeurs MSI-L et MSS ont ensuite été regroupées sous le même nom MSS du fait qu'aucune des caractéristiques cliniques ne parvenaient à les différencier. Les critères pour reconnaître les tumeurs MSI ont été révisés en 2002, recommandant alors d'utiliser en priorité des marqueurs mono nucléotidiques (BAT25, BAT26, NR21, NR22 et NR24) qui, en plus d'être très sensibles et spécifiques, ont l'avantage d'être quasi monomorphes (c'est-à-dire présentant très peu de variantes de taille dans la population générale), si bien que l'analyse comparative de l'ADN constitutionnel des patients n'est plus indispensable [3]. Une tumeur est considérée comme MSI si elle montre une instabilité d'au moins 3 de ces 5 microsatellites mono nucléotidiques. Cette méthode reconnue au niveau international est connue en France sous le nom de pentaplex.

Il est également à noter que cette méthode de diagnostic présente quelques limites ; on peut en effet signaler un possible manque de sensibilité et de spécificité du fait d'un certain polymorphisme dans la population mondiale pouvant mimer le phénotype MSI dans certaines situations (faux positifs), ainsi que des allèles rares, en dehors des zones de polymorphismes fréquents qui peuvent échapper à l'analyse (faux négatifs).

Tissu normal



Tissu  
Tumorale MSI

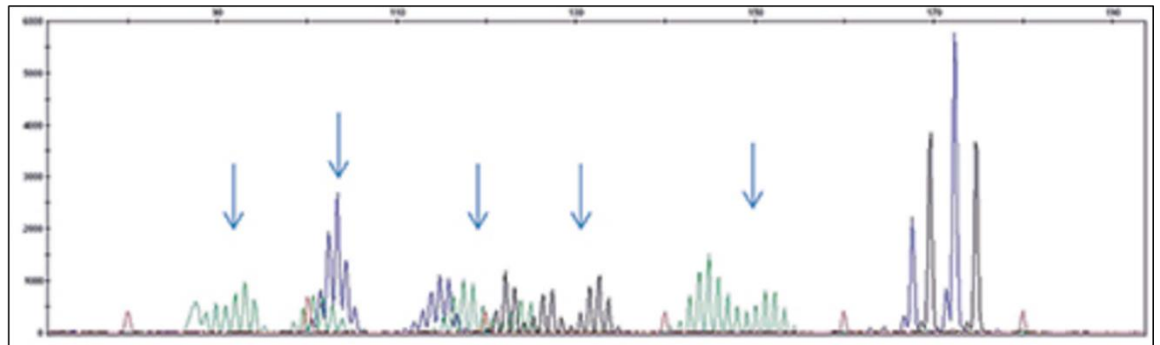


Figure 13 : Détermination du statut MSI des tumeurs par PCR (pentaplex). Mise en évidence de l'instabilité microsatellitaire par génotypage en fluorescence suite à la coamplification de cinq marqueurs (BAT25, BAT26, NR21, NR24 et NR27). Les produits amplifiés sont séparés par électrophorèse sur gel capillaire. Dans un contexte tumoral MSI, les délétions au niveau de ces microsatellites entraînent l'apparition de pics (indiqués par une flèche). Adapté de Bustamante et Chabot-Richards, 2015.

## 2. Analyse immunohistochimique

L'immunohistochimie représente une alternative ou un complément de l'allélotypage des microsatellites, qui permet la détermination indirecte du phénotype MSI en mettant en évidence la perte d'expression d'une ou plusieurs (combinaison de MLH1-PMS2 ou MSH2-MSH6) protéines issues d'un des gènes du système MMR dans la cellule tumorale.

Pour être complète, cette analyse nécessite l'emploi de 4 anticorps dirigés contre les protéines MLH1, MSH2, MSH6 et PMS2. Cette peut être intégrée au tests de routine d'IHC réalisés en pathologie et est moins coûteuse que l'allelotypage des microsatellites. De plus, elle permet de « cibler » le gène MMR en cause et donc d'orienter la recherche de la mutation constitutionnelle dans les familles prédisposées au syndrome de Lynch.

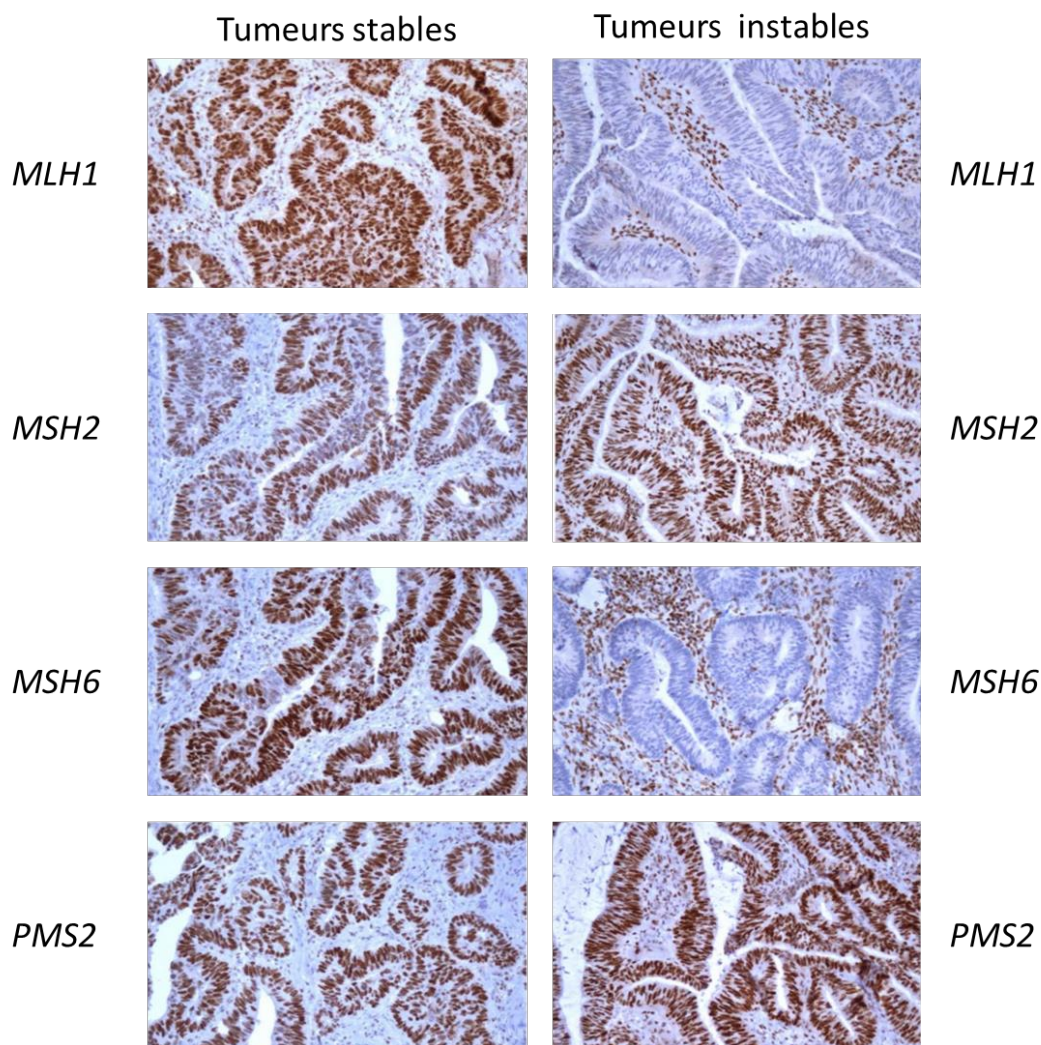


Figure 14 : Détermination du statut MSI des tumeurs par immunohistochimie. Immunomarquages des protéines MLH1, MSH2, MSH6 et PMS2. Un marquage positif dans l'épithélium colique normal adjacent peut être observé du fait des cellules normales contaminantes. Adapté de « immunohistochimie des protéines mmr » de Frédéric Bibeau.

### 3. Valeur prédictive du statut MSI

Plusieurs facteurs moléculaires ont été mis en évidence comme pouvant influencer le pronostic des cancers colorectaux, dont la déficience du système MMR, le phénotype CIMP ainsi que les mutations des oncogènes BRAF et KRAS. Les tumeurs MSI sporadiques, et non héréditaires, portent fréquemment une mutation somatique activatrice p.Val600Glu (p.V600E) dans l'exon 15 de l'oncogène BRAF. Cette mutation est présente dans environ 35% des tumeurs MSI et 5% des tumeurs MSS colorectales.

#### **4. Protocoles de traitements spécifiques aux cancers MSI – Immunothérapies**

Il s'avère que les protocoles de chimiothérapie des patients atteints de CCR MSI ne diffèrent pas des protocoles spécifiques aux tumeurs MSS. Rapidement, en Europe, les patients atteints d'une CCR de stade 2 sont traités par chirurgie, certains patients atteints d'un CCR de stade 2 peuvent recevoir une thérapie adjuvante afin de diminuer les risques de rechutes. Une chimiothérapie adjuvante est prescrite aux patients de stade 2 et 3. Aux Etats-Unis, seuls les CCR de stade 3 sont traités par chirurgie et chimiothérapie. Les stades 4 peuvent avoir une chimiothérapie non adjuvante en plus. Les CCR MSI sont rarement de stade 4 au moment du diagnostic.

Les trois agents chimiothérapeutiques utilisés en routine chez les patients atteints de CCR sont le 5FU, l'oxaliplatine et l'irinotécan. Il a été observé qu'en comparaison des tumeurs coliques MSS, les tumeurs MSI avaient une résistance supérieure à ces drogues, et ont même été clairement remises en question chez ces malades en raison d'un gain de survie nul en comparaison d'une absence de chimiothérapie.

Le traitement FOLFOX est une combinaison de 5FU, d'acide folique et d'oxaliplatine. Il est considéré aujourd'hui comme un standard dans le traitement des tumeurs colorectales quel que soit le phénotype moléculaire de la tumeur (MSI ou non). Une étude a permis de montrer une survie sans récurrence significativement supérieure chez les patients traités par FOLFOX versus 5FU seul lorsque la tumeur était MSI. On peut alors supposer qu'au minimum, l'ajout de l'oxaliplatine et de l'acide folique compenserait les effets négatifs du 5FU seul et pourrait donc être un médicament d'intérêt dans ces tumeurs.

Récemment, la détermination du statut MSI peut orienter le patient vers un traitement par immunothérapie (anti-PDL1) (révision avril 2018, Thésaurus National de Cancérologie Digestive).

# Chapitre 5 Epissage alternatif et cancer

## A. Généralités sur l'épissage alternatif

Suite à la transcription, le transcrit primaire ou ARN pré-messager subit une maturation en ARNm avant d'être exporté du noyau vers le cytoplasme, qui est le lieu de sa traduction en protéines.

Ce phénomène de maturation comprend trois étapes simultanées à la transcription par l'ARN polymérase II. Le transcrit acquiert une coiffe 7-méthyl-guanosine à son extrémité 5' ; une polyadénylation a lieu à son extrémité 3' ; et les séquences non codantes, ou introns, sont éliminées. Cette dernière étape, nommée épissage constitutif, est cruciale pour la grande majorité des gènes codant des protéines. Les processus de transcription et d'épissage apparaissent comme deux étapes indissociables spatialement et temporellement.

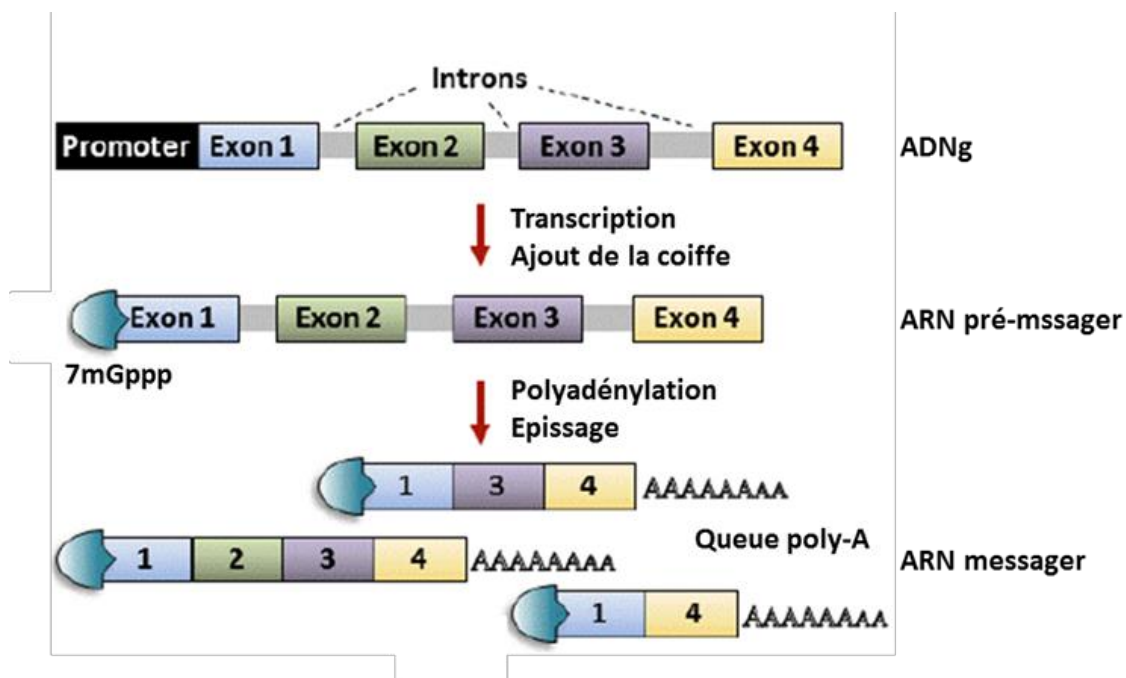


Figure 15 : Illustration des mécanismes de transcription et de maturation de l'ARN pré-messager (Samaan et al. 2013).

En outre, l'épissage est sujet à de nombreuses variantes, ou épissages alternatifs (EA).



Cinq types principaux d'EA sont observés :

- Un exon peut être inclus ou exclu du transcrit.
- Des exons peuvent être mutuellement exclusifs, c'est-à-dire que le transcrit ne contiendra qu'un seul des différents exons possibles.
- Un intron peut également être retenu.
- Il peut exister des extrémités 5' ou 3' alternatives, ce qui a pour conséquence, l'obtention d'exons de taille différente ou en nombres différents.
- Des promoteurs et des sites de polyadénylation alternatifs contribuent aussi à la diversité des transcrits codés par un seul gène.

L'EA concerne entre 80 et 90% des gènes humains (Nilsen et Graveley, 2010; Wang et al, 2008). Il permet de produire plusieurs transcrits distincts (six à huit en moyenne) à partir d'un seul gène, et participe ainsi à la diversité protéique, en façonnant à partir du même pré-ARNm, des transcrits différents codant des protéines aux fonctions variées.

Plusieurs études ont montré qu'environ 75% des événements d'EA affectent des régions codantes, c'est-à-dire qu'à terme, il est possible d'observer des changements d'acides aminés ou des délétions de segments protéiques plus ou moins grands, conduisant possiblement à des pertes ou des gains d'activités (Black et Grabowski, 2003; Blencowe, 2006; Matlin et al, 2005). Ces isoformes protéiques peuvent avoir des fonctions identiques ou différentes (voire antagonistes), être requises de manière transitoire ou traduire une adaptation pathologique (Kim et al, 2008; Nilsen et Graveley, 2010). L'EA est très finement régulé en fonction du type cellulaire ou du stade de développement. Les mécanismes de l'épissage différentiel impliquent les mêmes éléments régulateurs que l'épissage constitutif.

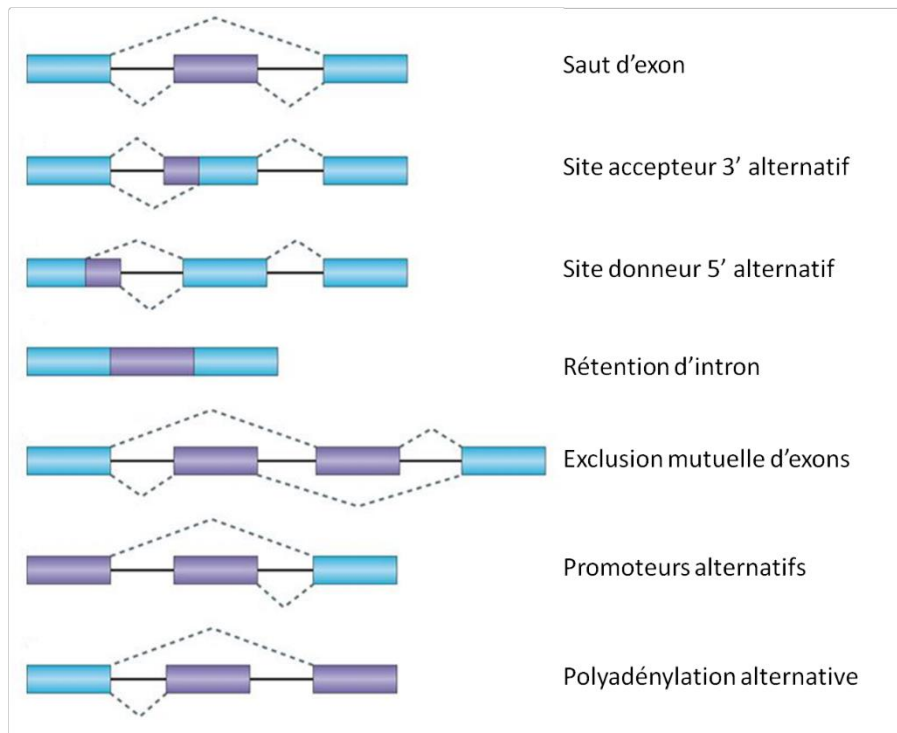


Figure 16 : Représentation des différents types d'épissage alternatif. (Ghigna et coll., 2008)

## B. Les séquences spécifiques de l'épissage

L'excision d'un intron d'un ARN pré-messager se déroule en deux étapes successives, chacune impliquant une réaction de *trans*-estérification. Trois déterminants de séquence sont requis : le site d'épissage en 5' ou site donneur (5'SS pour 5' splice site), le site d'épissage en 3' ou site accepteur (3'SS pour 3' splice site), qui sont situés à la jonction exon-intron, respectivement en 5' et 3' de l'intron, et le site de branchement BPS (branch point site), en amont du site 3' accepteur. Des éléments de séquence conservés caractérisent ces sites.

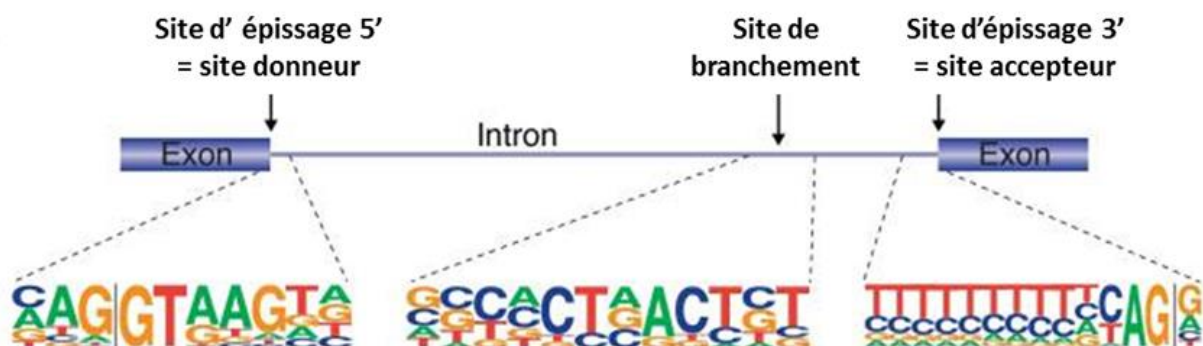


Figure 17 : Séquences consensus des sites d'épissage. Il existe des séquences consensus pour les sites d'épissage en 5' et 3', et pour le site de branchement. La taille des lettres pour chaque position est proportionnelle à la fréquence de la présence du nucléotide à cette position. (McManus & Graveley, 2011).

La réaction de *trans*-estérification a lieu au sein d'un complexe d'épissage, appelé spliceosome, dont l'assemblage et le fonctionnement requièrent des ribonucléo-protéines (snRNP pour uridine-rich small nuclear ribonucleo-proteins), ainsi que plus de 300 polypeptides.

Les snRNP sont constituées d'une molécule d'ARN (snRNA pour small nuclear RNA) complexée avec sept protéines Sm (Smith) ou apparentées et quelques autres protéines spécifiques (Nilsen, 2003). Il existe 5 snRNA nommées U1, U2, U4, U5, U6 (Will et Luhrmann, 2011). Les fonctions du spliceosome sont la reconnaissance des sites d'épissage et les étapes couplées, d'élimination des introns et de re-ligation des exons.

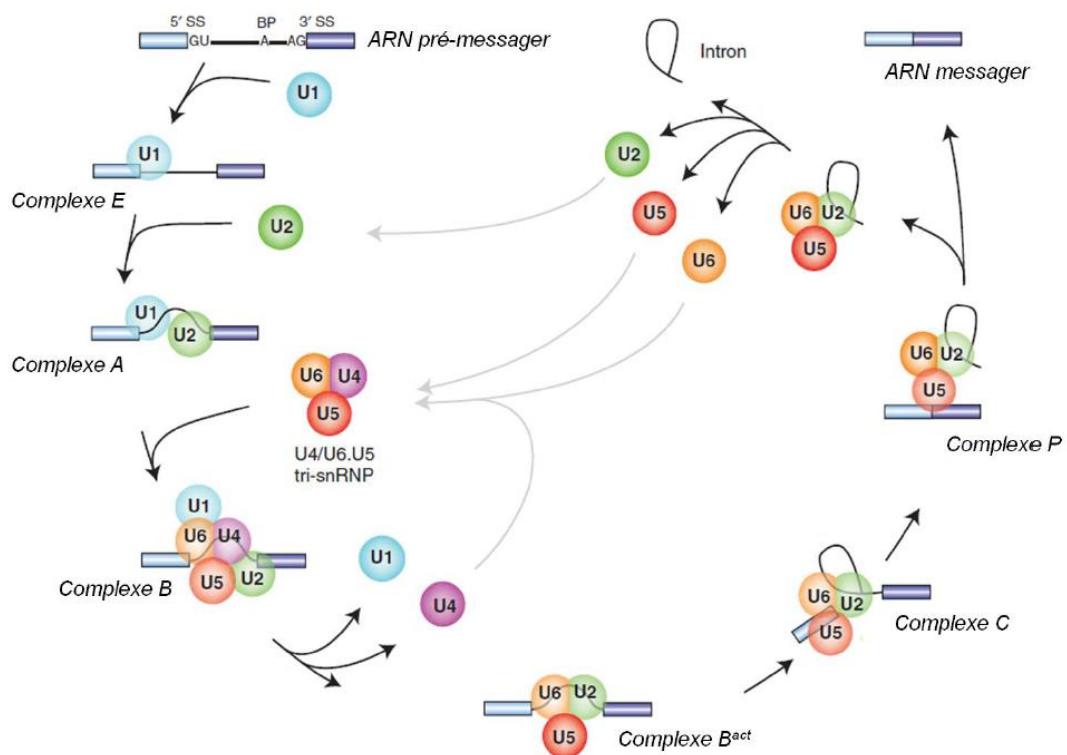


Figure 18 : Représentation simplifiée de l'assemblage du spliceosome. Différents complexes sont décrits : les complexes E, A, B, Bact (activé), C et P. (Will & Luhrmann, 2011)

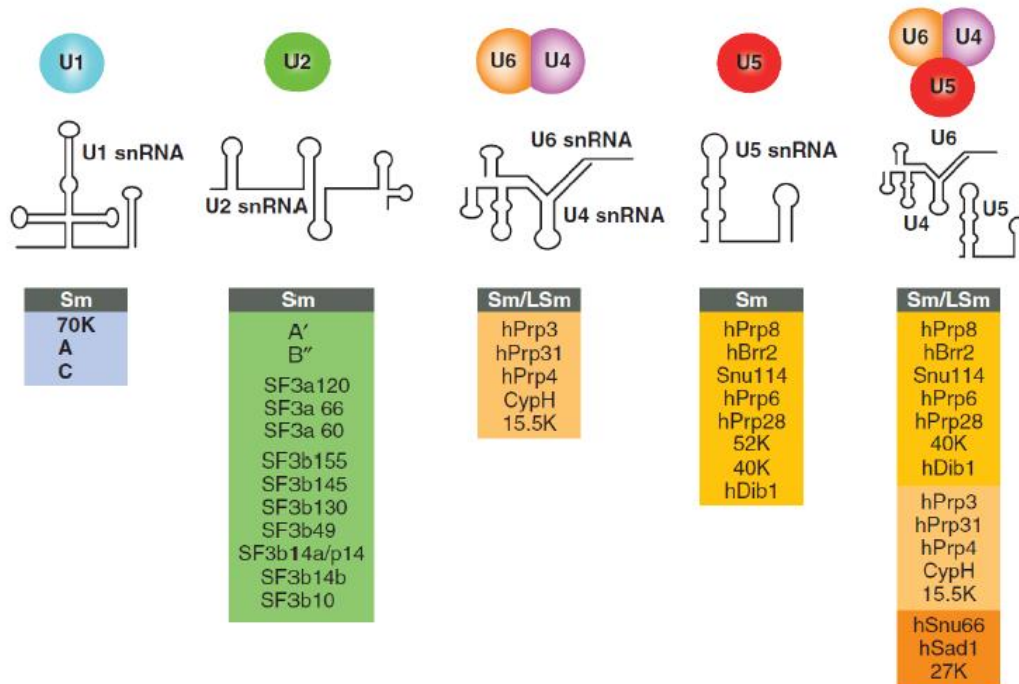


Figure 19 : Structure secondaire des snRNA et protéines composant les snRNP chez l'Homme. En gris est noté si les snRNP sont associées avec des protéines Sm (B/B', D3, D2, D1, E, F, et G) ou des protéines LSm (LSm2-8). (Will & Lüthmann, 2011)

### C. Les séquences et facteurs régulateurs de l'épissage

Les sites d'épissage sont des séquences courtes et dégénérées (Burge et al, 1998). Les introns sont, dans la plupart des cas, de longues séquences. Cependant, l'épissage est un mécanisme très fidèle, ceci grâce à l'existence de signaux additionnels, des facteurs *trans* et des séquences en *cis*. Quatre niveaux de régulation de l'épissage alternatif peuvent être considérés : les protéines SR, qui sont des activateurs d'épissage, les protéines hnRNP (heterogenous nuclear ribonucleo-proteins), qui sont des répresseurs d'épissage, les snRNP et les éléments *cis*-régulateurs.

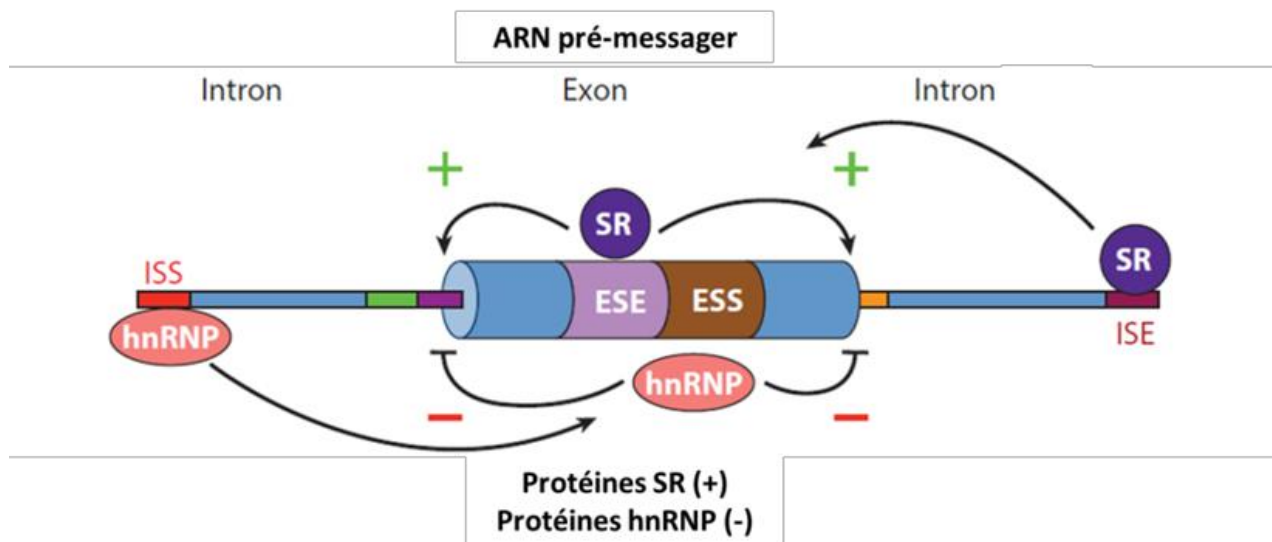


Figure 20 : Représentation des principaux facteurs de régulation agissant en cis ou en trans. Les séquences cis-régulatrices sont situées dans les exons (ESE/ESS) ou dans les introns (ISE/ISS). Les protéines SR et hnRNP peuvent interagir avec les séquences en cis (Lee and Rio 2015).

Les séquences *cis*-régulatrices sont classées selon leur localisation et selon leur effet sur l'épissage. On discerne ainsi :

- Les séquences ESE (exonic splicing enhancer) et ESS (exonic splicing silencer), situées au niveau des exons.
- Les séquences ISE (intronic splicing enhancer) et ISS (intronic splicing silencer), situées au niveau des introns.
- Les ESE et ISE sont des séquences régulatrices qui permettent le recrutement de certaines protéines à effet positif, favorisant ainsi la reconnaissance du site d'épissage.
- Les ESS et les ISS sont quant à elles, des séquences régulatrices négatives induisant généralement une répression de l'épissage par la liaison de protéines qui diminuent l'efficacité de reconnaissance du site d'épissage par le spliceosome (Adams et al, 1996).

Les protéines SR et apparentées sont caractérisées par leur structure modulaire. Il existe douze membres dans la famille des protéines SR (Manley et Krainer, 2010). Elles comportent en position C-terminale, un domaine RS riche en sérines et en arginines, de longueurs variables, phosphorylé de façon réversible, et en position N-terminale, un à deux domaines de liaison à l'ARN, nommés RRM pour RNA recognition motif (Graveley, 2000).

Les protéines SR participent à la reconnaissance du site d'épissage en 5', facilitent ou stabilisent l'assemblage du complexe au site d'épissage en 3', et sont impliquées dans le rapprochement des sites d'épissage en 5' et 3'. Les protéines SR comportent un domaine RS riche en sérines et en arginines et en N-terminal un à deux domaines de liaison à l'ARN (= RRM pour RNA recognition motif) (Aenkoe, 2014). Elles sont nommées SRSFs (pour SR Splicing Factor) numérotées de 1 à 12. Elles assurent les interactions entre des facteurs liés au site d'épissage en 5' et d'autres liés au site en 3' en amont, et permettent ainsi de délimiter les exons.

Pour remplir cette fonction, les protéines SR interagissent de façon spécifique avec les séquences activatrices ESE (Shen et Green, 2004). Certaines protéines SR font la navette du noyau jusqu'au cytoplasme (SRSF1) ; d'autres ne le font pas (SRSF2).

Environ 35% des événements d'épissage alternatif génèrent des transcrits avec des codons stop prématurés, qui pourront être dégradés par le NMD (Nonsense Mediated mRNA Decay) (Green *et al.* 2003). Plus généralement le NMD permet le contrôle de la qualité des ARNm. Ce mécanisme est couplé à la traduction (Maquat and Carmichael 2001).

Les protéines principales du NMD sont Upf1, Upf2 et Upf3. Le NMD va reconnaître et dégrader les transcrits présentant un cadre ouvert de lecture tronqué dû à l'apparition d'un codon stop prématuré (=PTC : Premature Termination Codons). Ces PTCs résultent d'une erreur dans la transcription ou dans l'épissage, d'une mutation ou d'un décalage du cadre de lecture. Upf1 a un rôle de reconnaissance des PTCs sur l'ARNm. Upf2 et Upf3 modulent l'activité d'Upf1.

Plusieurs mécanismes de reconnaissance d'un PTC existent. Un codon stop « normal » se distingue d'un PTC par la présence en aval du complexe de reconnaissance de la dernière jonction exon-exon (EJC pour Exon Junction Complex). L'EJC est une marque résiduelle de l'épissage situé 20 à 24 nucléotides en amont de la jonction exon-exon (Le Hir *et al.* 2000). Dans les transcrits présentant un PTC, le complexe EJC reste associé à l'ARNm et est détecté par Upf1. Un PTC augmente la distance entre le ribosome et la terminaison de la traduction et la queue poly(A). La conformation des particules ribonucléo-protéiques (mRNP pour messenger ribonucleoprotein) contenant le transcrit est alors reconnue comme anormale. Un signal positif de la particule PABP (Poly(A) binding protein), une mRNP, fixée à la queue poly(A) entraîne la traduction du transcrit.

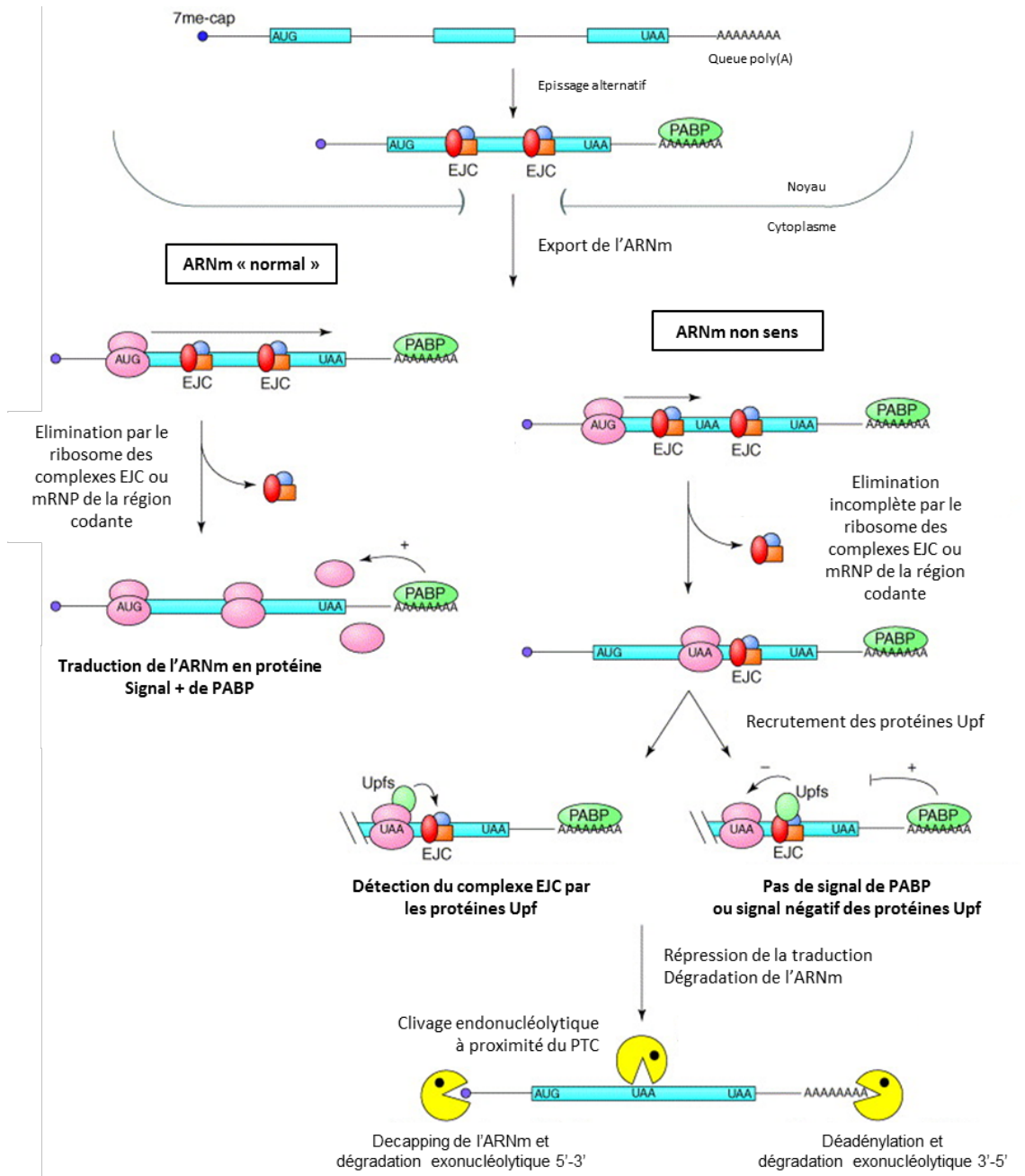


Figure 21 : Représentation du mécanisme de dégradation des ARNm non-sens. (Baker and Parker 2004)

## D. Epissage alternatif et pathologies

L'épissage fait intervenir de nombreux éléments (splicéosome, facteurs cis, facteurs trans) pouvant chacun être à l'origine de dérégulation. Des défauts d'EA sont retrouvés dans différentes maladies. Environ 15% des mutations induisant une maladie génétique affectent l'épissage du pré-ARNm (Krawczak et al. 1992).

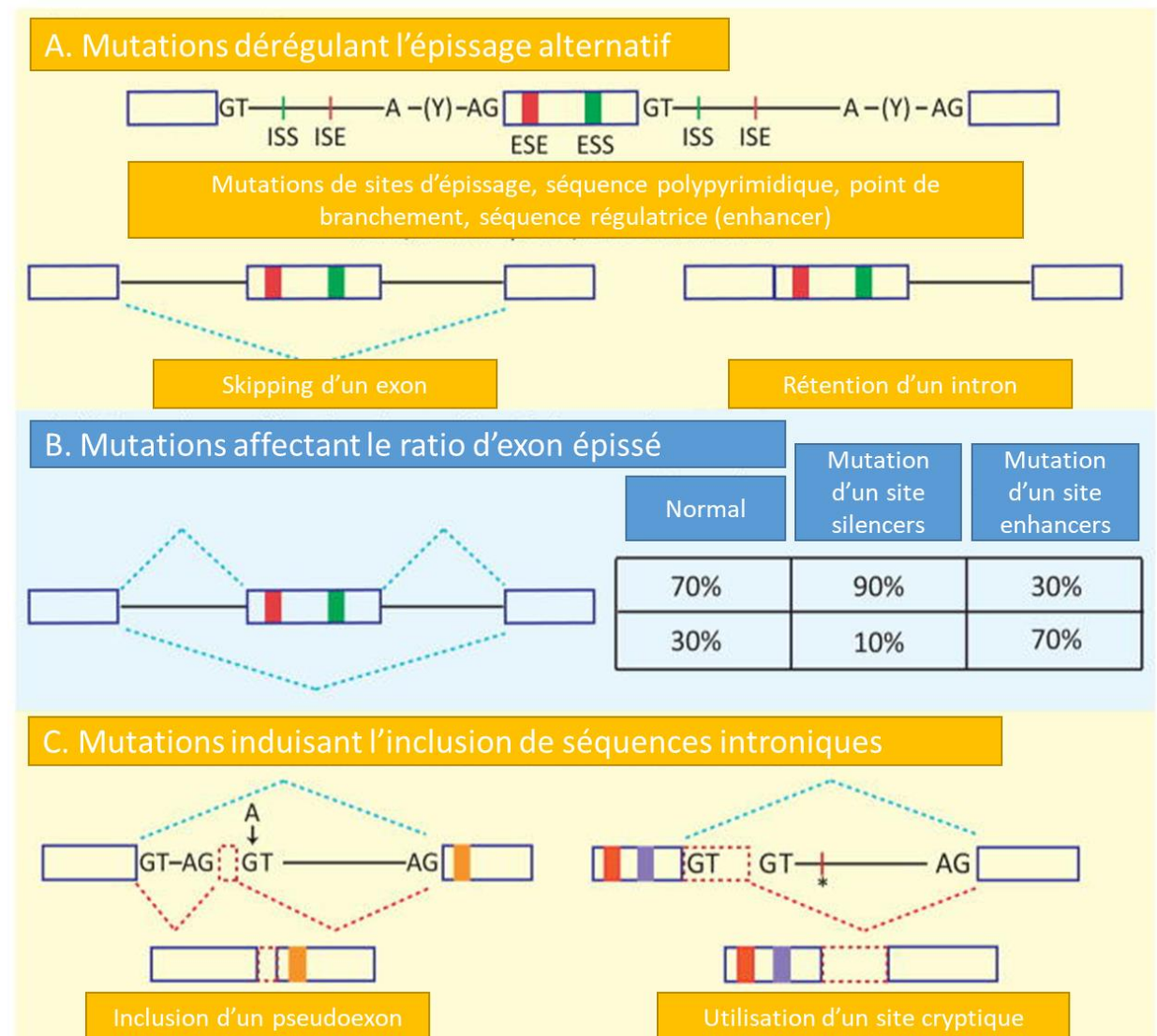


Figure 22 : Epissage et pathologies. D'après Singh et Cooper, 2012. A. Les éléments cis basiques sont indiqués : site d'épissage 5' et 3' (représentés par GT et AG), la séquence polymiridique (Y), point de branchement (A) et séquences régulatrices (ESE, OSE, ESS et ISS), B. Mutations affectant le ratio d'isoforme d'un gène, C. Mutations au sein d'introns pouvant conduire à l'inclusion de séquences introniques (indiquées par des rectangles en pointillés rouges), créer un nouveau site d'épissage ou un nouveau site régulateur (indiqué par un astérisque). Les lignes pointillées bleues indiquent le schéma d'épissage normal, tandis que les lignes pointillées rouges indiquent le schéma d'épissage provoqué par la mutation.

- Dans les cancers

Des défauts d'EA sont impliqués dans différentes maladies, et concernent soit des gènes spécifiques dans le cas de maladies héréditaires, soit des gènes variés. Un certain nombre d'articles a montré un lien entre une modification de l'épissage alternatif et la survenue d'un cancer (Brinkman, 2004).



Un cancer peut émerger du fait d'anomalies génétiques, correspondant à des mutations au niveau des sites d'épissage ou au niveau des séquences *cis*-régulatrices, avec pour conséquence, la modification de l'épissage d'un gène unique ; ou bien peut être lié à des altérations de l'expression des protéines d'épissage, contrôlant la sélection des sites d'épissage. Les mutations au niveau des sites d'épissage représentent au moins 10 à 15% de toutes les mutations qui surviennent dans le cancer (Bérout et al, 2005).

D'autres mutations affectent les séquences de régulation de l'épissage. Un certain nombre d'altérations décrites dans le cancer touchent des oncogènes (Brinkman, 2004 ; Domínguez et al, 2006). De plus, on retrouve certains de ces oncogènes impliqués dans des modifications d'EA (Pospisil et al, 2006). Des gènes suppresseurs de tumeur peuvent être inactivés, certains affectant l'EA d'autres gènes. Les processus de transcription et d'épissage des pré-ARNm sont régulés par plusieurs protéines qui peuvent être mutées dans les cancers (Auboeuf et al, 2007 ; Sanchez et al, 2008). Le catalogue « Cancer Gene Census » décrit un certain nombre de gènes codant des protéines d'épissage, qui sont porteurs de mutations et impliqués dans le cancer (<http://www.sanger.ac.uk/genetics/CGP/Census>). Un certain nombre de protéines d'épissage sont surexprimées dans les cancers (Grosso et al, 2008 ; Karni et al, 2007), parmi lesquelles se trouvent des protéines SR et apparentées (SRSF1, SRSF2, SRSF3, SRSF5, SRSF6, SRSF10) et des protéines hnRNP (HNRNPA1, HNRNPA2B1, HNRNPF, HNRNPK, HNRNPL). De plus, l'altération de l'expression de certains gènes impliqués dans l'assemblage du spliceosome a été observée dans des lésions précancéreuses du sein (André et al, 2009).

En outre, l'expression anormalement élevée d'un gène pourrait compromettre la disponibilité d'un facteur de régulation, et ainsi changer les profils d'épissage des pré-ARNm. Les altérations dans les profils d'EA observées dans les cancers mènent parfois à l'apparition de nouvelles isoformes. Dans d'autres cas, les isoformes correspondent à celles exprimées dans d'autres types cellulaires ou à des étapes précises du développement. Le plus souvent, les altérations d'épissage provoquent des différences dans l'abondance relative d'isoformes déjà exprimées dans ce type de tissu. Parmi ces changements, nombreux sont ceux qui n'auront aucun impact fonctionnel sur la cellule, d'autres, par contre, provoqueront un avantage sélectif en neutralisant les signaux visant à freiner sa prolifération et à provoquer sa mort, ou en favorisant une meilleure dissémination dans l'organisme (Revil et al, 2006).

On peut noter également l'existence de résistances, spécifiques à certains traitements, associées à des épissages alternatifs (Dehm, 2013). Plusieurs études ont noté l'existence de différences dans les profils d'EA entre tissus normaux et cancéreux, notamment dans le cas du cancer colorectal (Gardina et al, 2006; Thorsen et al, 2011). L'expression des protéines SR et hnRNP est altérée dans des adénocarcinomes du côlon (Ghigna et al, 1998). Les variants d'épissage pourraient constituer des biomarqueurs du cancer (Omenn et al, 2013), mais aussi des cibles thérapeutiques (Miura et al, 2012).

Plus spécifiquement, dans les cancers MSI, le gène *MRE11* a été décrit comme cible de cette instabilité en 2004. En effet on observe la présence d'une délétion au niveau du microsatellite T11 situé en amont de l'exon 5. Localisée au niveau d'un site accepteur d'épissage, cette mutation entraîne le saut de l'exon suivant (exon 5) et la production d'un transcrit aberrant. Ce transcrit est responsable d'une perturbation de la réparation des cassures double brin de l'ADN dans les cancers colorectaux (Giannini et al, 2004). Il a également été décrit un phénomène similaire au niveau de sites accepteurs d'épissage du gène *ATM* aboutissant à la formation d'un transcrit aberrant. Le gène *ATM* code une serine-thréonine kinase impliquée dans la signalisation des cassures double brin de l'ADN et dans la surveillance de l'intégrité du génome. Les épissages alternatifs aberrants sont un saut de l'exon 6 dans les cancers gastriques et des sauts des exons 8 et 12 dans les cancers colorectaux MSI. De plus, il est également décrit une diminution de l'expression de la protéine sauvage dans les tumeurs portant ces mutations ainsi que la production d'une protéine tronquée (Kim et al, 2013 ; Ejima et al, 2000). D'autres mutations affectent une région microsatellite intronique du gène codant la protéine HSP110. Sous le contrôle de délétions somatiques au niveau du microsatellite T17 en amont de l'exon 9, l'épissage du transcrit sauvage est perturbée au profit du transcrit mutant dépourvu d'exon 9 et possédant un codon stop prématuré dans l'exon 10 en raison du décalage du cadre de lecture induit par le saut d'exon. Ce transcrit alternatif code une forme tronquée de la chaperonne HSP110, ne possédant que son domaine de liaison à l'ATP. La mutation entraîne alors une perte d'expression du transcrit sauvage HSP110 corrélée à la taille de la délétion au niveau du microsatellite T17. Ces délétions favorisent un épissage aberrant qui atteint un plateau lorsqu'une perte allélique supérieure ou égale à 5 paires de bases est atteinte, entraînant une perte totale de l'expression du transcrit sauvage.

A l'inverse, l'expression du mutant délété de l'exon 9 ne semble pas augmenter en fonction de la taille de la délétion suggérant une contre régulation par le NMD (Dorard et al, 2011).

- Dans d'autres pathologies

Plus spécifiquement, dans les maladies par expansion de triplets qui représentent plus de 40 troubles neurologiques différents à l'origine d'altérations de l'épissage en lien avec une élongation de ces séquences (Groh et al, 2014). Plusieurs maladies, dont la dystrophie myotonique de type 1 et au moins deux types d'Ataxie spino-cérébelleuse, impliquent une mauvaise régulation de l'épissage due à la séquestration de protéines de liaison à l'ARN par des répétitions étendues dans les régions microsatellites correspondantes (Echeverria et Cooper, 2012 ; Galka-Marciniak et al, 2012). Ces effets ont également été observés dans la maladie de Huntington. En effet, le fragment N-terminal tronqué, toxique, de la protéine HTT mutante est généré par un épissage alternatif défectueux dépendant de la longueur de la séquence de microsatellite « CAG » (Sathasivam et al, 2013).

# Objectifs des travaux de thèse

---

Ces travaux de recherches ont été réalisés sous la direction du Dr. Laurent Corcos et du Pr. Cédric Le Maréchal au sein de l'équipe ECLA (Epissage-Cancer-Lipides-Apoptose), dans le laboratoire de Génétique, Génomique Fonctionnelle et Biotechnologies (Unité Inserm UMR 1078) du Dr. Emmanuelle Génin, à Brest.

Le premier axe de ce projet a consisté à rechercher l'existence de variations génomiques fréquentes dans le cancer gastrique chez l'Homme, en particulier des mutations pouvant affecter l'épissage des ARN pré-messagers pour produire des transcrits alternatifs. Parmi ces ARN, nous émettons l'hypothèse que certains pourraient être associés à la cancérogenèse du tissu gastrique.

La seconde partie a consisté à identifier des gènes dont l'expression modifiée pourrait être associée à une grande susceptibilité accrue à développer un cancer gastrique et/ou à rendre compte d'une réduction de l'efficacité des traitements anticancéreux

# Résultats I

---

# Article

Microsatellite instability linked mutations drive altered pre-mRNA splicing in cancer.

Gaël Quéré, Catherine Le Jossic-Corcos, Brigitte Simon, Danielle Arzur, Jean-Philippe Metges, Cédric Le Maréchal\* and Laurent Corcos\*

## Objectifs de l'étude

Ces travaux de recherches ont eu pour but d'identifier des mutations susceptibles d'altérer l'épissage dans les cancers gastriques avec comme paramètres d'être à la fois récurrentes ( $\geq 5\%$ ) et positionnées dans des régions supposées importantes pour le contrôle de l'épissage.

Le programme International Cancer Genome Consortium (**ICGC**), a été initié en avril 2008 et avait pour objectif de mettre à la disposition de la communauté scientifique une base de données recensant les mutations somatiques caractéristiques de 50 cancers ou sous-types de cancers, considérés comme les plus préoccupants aux plans clinique et sociétal.

Afin d'identifier les mutations somatiques retrouvées dans les cancers gastriques, deux bases de données représentées par deux projets distincts ont été compilées : le projet STAD-US regroupant les données d'adénocarcinomes de l'estomac et le projet ESAD-UK pour les adénocarcinomes de l'œsophage. Le projet STAD-US regroupe les données de 443 Tumeurs dont 289 ayant bénéficié d'un séquençage entier de leur génome.

Nous avons ainsi pu identifier deux mutations affectant respectivement le gène *DNAJC18* codant une protéine chaperonne (HSP40), ainsi que le gène *LMAN1* codant une protéine transmembranaire dénommée ERGIC-53, une protéine principalement exprimée au niveau des membranes du compartiment intermédiaire, ou ERGIC, un compartiment tubulo-vésiculaire situé entre le réticulum endoplasmique et l'appareil de Golgi. Dans ces deux cas, les mutations se caractérisent par la délétion somatique d'un nucléotide dans une région microsatellite au niveau du site accepteur d'épissage de l'intron.

Nous avons ensuite analysé les transcrits codant ces protéines dans une série de cancers gastriques et coliques (et de quelques autres) de phénotype MSI (« MicroSatellite instability) et MSS (MicroSatellite Stability). Nos résultats montrent qu'il existe un lien entre le phénotype MSI et l'apparition des transcrits alternatifs caractérisés par le saut de l'exon suivant, dans la séquence, la région microsatellitaire intronique. Le transcrit aberrant contient alors, au niveau de l'exon 3 pour *DNAJC18*, et de l'exon 9 pour *LMAN1*, un codon stop prématuré et code potentiellement une forme mutante tronquée de la protéine correspondante. On peut signaler que le transcrit mutant de *DNAJC18* (détecté dans la lignée cellulaire HCT116) ne semble pas être sensible au nonsense-mediated mRNA decay (NMD).

Afin d'étudier les mécanismes mis en jeu par ces mutations MSI, plusieurs minigènes ont été construits à partir des séquences d'intérêt, puis transfectés dans des cellules cancéreuses MSI et NON-MSI.



## Microsatellite instability linked mutations drive altered pre-mRNA splicing in cancer

Gaël Quéré, Catherine Le Jossic-Corcoc, Brigitte Simon, Danielle Arzur, Jean-Philippe Metges, Cédric Le Maréchal\* and Laurent Corcos\*

### Abstract

Microsatellite (MS) sequences are short repeated nucleotide stretches of 6 or more nucleotides preferentially located within introns, often in the vicinity of intron-exon boundaries. In this study, we addressed the possibility that such sequences could influence pre-mRNA splicing decisions, such as altering the rates of exon inclusion.

Microsatellites instability (MSI) accounts for about 20% of gastric cancers. Upon analyzing gastric cancer data from the TCGA databases (esophagus, n=203 and stomach, n=289), we sorted out those 200 gene mutations that were highly frequent, at least 5%. Among these, we selected 8 mutations localizing in close proximity to splice acceptor sites (genes *ATL3*, *DDX6*, *DNAJC18*, *LMAN1*, *PREPL*, *PSME4*, *SLCO6A1* and *ZNF43*). Five of these mutations were found in samples from 3 patients among our local set of 34 patient samples (genes *DNAJC18*, *LMAN1*, *PREPL*, *PSME4* and *ZNF43*), as revealed by fluorescent multiplex-PCR.

We next analyzed the splicing of these five genes by RT-PCR, in gastric and colon cancer samples carrying or not the identified mutations, as compared to matched mutation-free non-tumor tissue. Strikingly, MS mutations in intron 2 of *DNAJC18* and in intron 7 of *LMAN1* were associated with *DNAJC18* exon 3 and *LMAN1* exon 8 partial skipping, respectively, specifically in MSI gastric cancers, but not in MS stable gastric or colon cancers, or healthy tissue. Several of the same mutations were also found in a series of 96 MSI cancers from diverse localization.

To further confirm the links between presence of the mutations and the skipping of these exons, we engineered minigenes carrying either the normal or the mutated sequences of *DNAJC18* and *LMAN1*, and transfected these into MSI and MSS cell lines. Our results showed that all MSI cell lines, but none of the MSS cell lines, showed a commonly altered splicing pattern - namely the skipping of the exon next to the MS alteration in the preceding intron.

Taken together, our results demonstrated that MSI cancers displayed pre-mRNA splicing defects linked to the instability of microsatellite sequences located in the vicinity of 3' splice acceptor intron sites of genes frequently mutated in such cancers.

# Introduction

Gastric cancer is a leading cause of mortality by cancer, as patients often present at late stages in absence of high specificity warning signs. About 1 million subjects are diagnosed with gastric cancer worldwide each year (Rugge et al, 2015). Roughly 90% of gastric cancers are adenocarcinomas that develop from epithelial cells belonging to the gastric mucosa (ref). At the molecular level, several genes are frequently mutated, including *APC*, *ARID1A*, *BRAF*, *CDH1*, *CDKN2A*, *CTNNB1*, *ERBB2*, *HRAS*, *KRAS*, *NRAS*, *PDGFRA*, and *P53*, or silenced by epigenetic mechanisms (Cancer Genome Atlas Research Network, 2014).

Reports from the TCGA consortium permitted classifying gastric cancers into 4 types characterized by specific phenotypic and genotypic alterations (CIN-chromosomally unstable, GS-genomically stable, EBV-associated and MSI-microsatellite instability). MSI cancers account for 20% of gastric cancers, and are characterized by defects in expression of genes involved in DNA repair such as *MLH1* and *MSH6*, resulting in a high level of deleterious mutations. In addition, microsatellites may be located in any genomic region, either coding or noncoding (Pérez-Jiménez et al 2013; Vieira et al, 2016; Phumichai, 2015). Repair defects of MSI mutations lead to the synthesis of aberrant proteins, or even to recruiting the Nonsense Mediated mRNA Decay (NMD) machinery that destabilizes transcripts with a premature STOP codon, and thus strongly reducing protein levels. Because MSI-associated gastric cancers have a relatively good long-term prognosis (Velho, 2014), it is important to identify MSI regions genome-wide. Pre-mRNA splicing is often altered in cancers of various origins, including both solid cancers and hematological cancers. Such alterations can result from *cis*-acting mutations affecting the activity of regulatory regions, or from reduced level or activity of RNA binding proteins onto these regions (Oltean and Bates, 2014). Indeed, a large number of such proteins participate in splicing control, as members of the spliceosome complex which catalyzes the splicing reaction, or as satellite proteins that help recruiting or preventing spliceosome subunits to interact with nascent pre-mRNA molecules (Wang, 2015).

In the present study, surveying the TCGA database, we aimed at sorting out putative splicing-altering mutations in GC cancers that were both frequent (at least 5%) and occurring in regions putatively important for splicing control. In doing so, we identified roughly 200 genomic alterations, including in repeated mononucleotide stretches, some of which were present in introns, and eventually accounted for specific exon skipping events. We further analyzed the functional effects of these mutations in MSI and in non-MSI cancer tissues and cells, at the level of the endogenous gene, or with engineered minigenes encompassing the contigs of the affected genes.

# Material and methods

## Database search

The Cancer Genome Atlas Data: Raw Sequencing Data and Simple Nucleotide Variation data were downloaded from the Cancer Genomics Hub (CGHub, <https://cghub.ucsc.edu>) and the TCGA Data Portal (<https://tcga-data.nci.nih.gov/tcga/>). Utilization of data was conducted in accordance with TCGA Human Subjects Protection and Data Access Policies.

## Patient sample set and MSI status

Gastric adenocarcinoma primary tumor tissues and normal gastric tissues (fresh frozen) were obtained from 34 patients undergoing surgery and included in the clinical and biological database of Brest University Hospital. None of the patients had been treated with prior chemotherapy or radio-therapy. All patients provided informed consent.

MSI determination was carried out using a modified version of the multiplex PCR described by (Suraweera et al, 2002). The markers were co-amplified in a standard multiplex PCR with universal fluorescent labeling (denaturation at 95°C for 5 minutes, 35 cycles of denaturation at 95°C for 30 seconds, annealing at 55°C for 30 seconds, and extension at 72°C for 30 seconds, followed by an extension step at 72°C for 5 minutes) and were analyzed with ABI 3130xl genetic analyzer (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions.

## DNA and RNA preparation, mutation analyses and sequencing, RT-PCR

DNA and RNA were extracted with The AllPrep DNA/RNA Mini Kit (Qiagen) according to the manufacturer's instructions. RNA integrity was evaluated with an Agilent 2100 bioanalyzer using the RNA Nano Chips kit (Agilent) for all primary tumor tissues and normal gastric tissues. The presence of mutations was determined by amplicon length analysis by capillary electrophoresis with fluorescence detection in our MSI/MSS gastric cohort of 34 patients as well as in Co115, HCT116, LoVo, and HCT8 (MSI colorectal cancer cell lines), Snu1 (MSI gastric cancer), and SW480 (MSS colon cancer), AGS, and HGT-1 (MSS gastric cancer) cell lines.

All the mutated sequences were verified by sequencing. Primer pairs were used to target mutated sequences:

|                      |                                    |                                    |
|----------------------|------------------------------------|------------------------------------|
| ADAM28:              | F: 5'-GCTTGCTAATGACTACACAGATAG-3', | R: 5'-GGAGGTTCTCTGGAGAAAGAA-3',    |
| ATL3:                | F: 5'-AGTGCTGTAGTCAAAGCAAAG-3',    | R: 5'-GTGCTTAGAGCAAAGATGGTAG-3'    |
| ATM:                 | F: 5'- TTTCTGTATGGGATTATGGA -3',   | R: 5'- CCTCGCACACTGAATAG -3',      |
| CDC14A:              | F: 5'-TCAAAGTGGAGGAAGTAAAATG-3',   | R: 5'-CCAAGTCTCACCAGGAGTTAT-3',    |
| DDX6:                | F: 5'-GAGAGTTGCGTCCCTTACA-3',      | R: 5'-GGGTTACTCCCTTCAGAGTTAG-3',   |
| DNAJC18:             | F: 5'-GATTGGAGGTAGGATCGATGT-3',    | R: 5'-CATCTCGAGAAACTCCCAGAA-3',    |
| FHOD3:               | F: 5'-GTCTGTGAAAGGATGCCAAG-3',     | R: 5'-CACTTGGTTGGGCTTGAA-3',       |
| FRMD4A:              | F: 5'-AGCTCTTTCACCTGTGTCT-3',      | R: 5'-CATAGAAGGAGAACTCACCCA-3',    |
| GSE1:                | F: 5'-TCCACAGGGTCTCACT-3',         | R: 5'GGTGTCTGGAATGTAAACTCAAC-3',   |
| HSP110:              | F: 5'- AAACCTGTCCATCCATTG -3',     | R: 5'- CATGGTTCCAGATCAGAGATATT -3' |
| LARP4B:              | F: 5'-TAGGAGGAAATGAGTCTCAACC-3',   | R: 5'-AAACCACCACCAGCCTTA-3',       |
| LMAN1 (rs759519259): | F: 5'- TTTAATGATCTATTCCAAAATC-3',  | R: 5'- TCTTGTGAAAGTGCTCAAA-3',     |
| LMAN1 (rs746845401): | F: 5'- AGCCACACCAGATAAAGAA -3',    | R: 5'- CAAGCGTCCATGATCAACAG -3',   |
| MRE11:               | F: 5'- GAGTATCATTTGAACCCGAGAG -3', | R: 5'- ATCCAAGAAGGTAGGTAGATA -3',  |
| PAMR1:               | F: 5'-GGGCTCTGACACCTTTCT-3',       | R: 5'-GAGATGGAGCTAGGTCAAGG-3',     |
| PREPL:               | F: 5'-ATTACAGGCCCTGGTTAAGG-3',     | R: 5'-TCTAATGTCAGAGGAAGTGTAGTG-3', |
| PSME4:               | F: 5'-GAGGGCAAAGGAAAGAGTAG-3',     | R: 5'-TGTTGTGCTCAGTGACTAAC-3',     |
| RNF43:               | F: 5'-CAGTACCAGCAGTCTGTTC-3',      | R: 5'-CTGGAGGTCCACAGATCAAG-3',     |
| SLCO6A1:             | F: 5'-ACTGCCCTGAAGTCAGTAG-3',      | R: 5'-ACAATGACACCTCCCAGAA-3',      |
| XYLT2:               | F: 5'-CTACAACAGCAGCAGGAAA-3',      | R: 5'-ACGTCACTCTCTGGTTGG-3',       |
| ZBTB20:              | F: 5'-CCAGTGTAGTATCTGCAACAAG-3',   | R: 5'-CAGACGGAGCAGACGTAA-3',       |
| ZNF43:               | F: 5'-GTCAGTACCTGTTCTTTACTC-3',    | R: 5'-GGAATCTGAAACTCATGAATGC-3',   |

Forward primers (F) of each pair of primers were labeled with -FAM fluorochromes by universal labeling during the PCR reaction; primers pairs were synthesized at Eurogentec (Seraing, Belgium). In addition, primers were divided in 2 groups based on the size of the fragments. The PCRs were carried out in a final 25µl volume with 12.5µl of 2x Multiple PCR Master Mix (Qiagen), primers and distilled water to a final volume of 23µl. Twenty nanograms of DNA extract were used as the template for each PCR. Thermal cycling was performed under the following conditions: 95°C for 15 min followed by 25 amplification cycles of 94°C for 30 s,

57°C for 30 s, and 72°C for 60 s and a final extension step at 72°C for 10 min. Each amplification run contained a No template Control. Fluorescent fragment size analysis was performed on an ABI 3130xl genetic analyzer (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions. Data were collected with ABI 3130xl data collection software (version 3.0, Applied Biosystems) and interpreted using GeneMarker software (version 2.4, Softgenetics).

After RNA extraction, one microgram of RNA was reverse transcribed with the M-MLV reverse transcriptase (Invitrogen) and oligo-dT primer, and the cDNA was amplified with primers:

|                      |                                      |                                   |
|----------------------|--------------------------------------|-----------------------------------|
| ADAM28:              | F: 5'-GGGTGAAGAAGTATGAAGTGGT-3',     | R: 5'-GGCTTGTGGTGATCTCCTTT-3',    |
| ATM:                 | F: 5'-CTCAGCAACAGTGGTTAG-3',         | R: 5'-GATATGATTTAGACCTGAAGAG-3',  |
| DDX6:                | F: 5'-TGATAGTATTGGATGAGGCAGATAAG-3', | R: 5'-ACTCGCTGAGAGGAGTTACA-3',    |
| DNAJC18:             | F: 5'-TGACGCAGTTAGAAGAAACAAATAC-3',  | R: 5'-GTCCTCCAAGAAGACGTTGA-3',    |
| FRMD4A:              | F: 5'-CGTCAAGGCTCAGTTCAAGA-3',       | R: 5'-GTACTTTGGTGAGGTGGAGAC-3',   |
| HSP110:              | F: 5'-GCTACACGAATTCAGCTGTGA-3',      | R: 5'-GAGCAGCATGGTTTCGACTAAA-3'   |
| LARP4B:              | F: 5'-GAGCTGAACCCTAATGCAGAA-3',      | R: 5'-GCACATACTGGTCACTATCCATC-3', |
| LMAN1 (rs759519259): | F: 5'-AGATGACCATGATGTCTTTCTT-3',     | R: 5'-CGGTTGAGCTGCTTGATTTTC-3',   |
| MRE11:               | F: 5'-CCAGGGTTCTTGGAGAAG-3',         | R: 5'-CCAGCACAACTTAAAATGTC-3',    |
| PREPL:               | F: 5'-TCTAACAAACCCTGACTGCTTTC-3',    | R: 5'-CGATGGCTTCCTTGAGTTTCT-3',   |
| PSME4:               | F: 5'-CAGCATGATGCAGGGATTTG-3',       | R: 5'-TCGCCATTCTTCTAGCATCTC-3',   |
| RNF43:               | F: 5'-GGTTACATCAGCATCGGACTT-3',      | R: 5'-TCACACAGCCTGTTTACAC-3',     |
| ZNF43:               | F: 5'-CTCTGTGGCTCTGTGACCT-3',        | R: 5'-ATAGGCTCCCAAGGCTCTTT-3'.    |

### Minigene constructions

The genomic sequences surrounding DNAJC18 exon 3 and LMAN1 exon 8 were inserted into the pMG1.1\_polR2G plasmid (A gift from G. Le Gac and C. Ka, Inserm U1078-Brest) to obtain the DNAJC18 minigène and the LMAN1 minigène. Briefly, the genomic segment encompassing the sequences of interest, containing 150 bp of the upstream intron, the next exon and 150 bp of the downstream intron was amplified by PCR from tumor genomic DNA of the patient using, as primers: for the DNAJC18 exon 3 cassette: forward primer 5'-GGATCTTAGCCTGAACCTTG-3', reverse primer 5'-AGTGGGAGCTATGCTGAT-3' and for the LMAN1 exon 8 cassette: forward primer 5'-CGCGACTCTGTCTCTAAA-3', reverse primer 5'-GGTCCCATTCCTCCAATAA-3'.

To permit subsequent cloning, an EcoRI target site was introduced at the end of the sequences. PolR2G vectors, harboring wt or mutant sequences were selected by sequencing (Supplementary Figure 2 and 3).

### **Cell culture and transfection**

All cell lines were grown in 90 mm plates at 37 °C under a humidified atmosphere of 5% CO<sub>2</sub>: LoVo, HCT-8, Co-115 and SNU-1 in Dulbecco's modified Eagle's medium: nutrient mixture F-12 (DMEM/F-12) containing 4.5 g L<sup>-1</sup> glucose and supplemented with 10% fetal bovine serum (FBS) (Gibco-Invitrogen, Cergy-Pontoise, France). HCT116 cells were maintained in DMEM/F-12, supplemented with 5% FBS. HT29 cells were cultured in Dulbecco's modified Eagle's medium (DMEM; 4.5 g/l glucose) (Lonza, Belgium) supplemented with 10% fetal bovine serum. HGT-1 cells were cultured in Dulbecco's modified Eagle's medium (DMEM; 4.5 g/l glucose) (Lonza, Belgium) supplemented with 5% fetal bovine serum. AGS cells were maintained in DMEM/F-12, supplemented with 10% FBS. All cell lines were grown without antibiotics. The medium was changed every two days, and cells were passaged using 0.05% trypsin/EDTA (Gibco Invitrogen).

Cells were plated at a density of 75.10<sup>4</sup> cells per well in 6-well plates 24hr before transfection. Minigenes (1.5 µg per well) were transfected with the Lipofectamine 2000 reagent (Invitrogen). Cells were harvested after 48 h, washed in PBS and used to prepare total RNA.

### **RNA isolation and RT-PCR analyses**

Total RNA was extracted with Tri-Reagent (Ambion) according to the manufacturer's instructions. RNA was quantified by spectrophotometry and the integrity was ascertained by electrophoresis. One microgram of RNA was reverse transcribed with the M-MLV reverse transcriptase (Invitrogen) and oligo-dT primer, and the cDNA was amplified with primers:

PolR2G minigene: F: 5' - TAATACGACTCACTATAGGG -3' , R: 5' - TAGAAGGCACAGTCGAGG -3' ,



# Results

## *1-Database analyses*

The TCGA and ICGC databases were interrogated for gastric cancers (203 esophagus adenocarcinomas and 289 stomach adenocarcinomas) for the presence of somatic mutations. We identified 5 million somatic mutations, from which 200 such mutations, corresponding to a frequency of at least 5%, were selected. Among these, we filtered the data to focus onto those mutations that occurred within microsatellite repeats in the 3' sites of introns, *i.e* near the splice acceptor sites (Figure 1). Among these mutations (Table 1), we focused on one mutation in *DNAJC18* intron 2 (deletion of one T residue) and one mutation in *LMAN1* intron 7 (deletion of an A residue, opposite strand).

## *2-Characterization of the MSI status of our gastric cancer sample set and in gastric and colorectal cell lines*

Among our 34 cancers sample set, 3 belonged to the MSI class, as shown by PCR amplification of microsatellite loci from healthy tissue DNA and gastric tumors DNA. We could observe a new allele for at least 3 of the 5 studied markers (Figure 2a), and confirmed by the specific decreases in the levels of MLH1, MSH6, MSH2 and PMS2 mRNAs, most marked for MLH1 (15-fold) and MSH6 (5-fold) (Figure 2b). The electrophoretic profiles of gastric cancers and normal gastric tissues, at the gene loci for the identified mutations, together with those of MSI and non-MSI cell lines are shown in figure 3.

## *3-Transcript analyses*

In order to determine the splicing status of the *DNAJC18* and *LMAN1* transcripts according to the mutation status (figure 4a and 4b), we performed RT-PCR analyses with samples from MSI and control tissue. Two cDNA amplification products were observed in MSI cancer samples but not in control tissue, one, more abundant, corresponding to the normal sequence, with exon 3 (*DNAJC18*) or exon 8 (*LMAN1*) inclusions, and a faster migrating band that corresponded to incomplete mRNA sequences, without exon 3 (*DNAJC18*) or exon 8 (*LMAN1*),

as determined by sequence analysis of each amplification product (Figure 4c and 4d). Similarly, MSI cell lines showed partial exon skipping for both genes, whereas MSS cell lines expressed only the exon-containing transcript (*DNAJC18*, Figure 5a, and *LMAN1*, Figure 5b). As controls, RT-PCR analyses of already reported data in MSI cancers (Giannini et al, 2004 ; Kim et al, 2013 ; Ejima et al, 2000; Dorard et al, 2011), we analyzed *HSP110* and *MRE11* splicing products. Similarly to the effects of the *DNAJC18* and *LMAN1* mutations, gastric and colorectal cancers bearing *HSP110* or *ATM* mutations within their intronic microsatellite region also showed altered splicing (next exon exclusion).

#### *6-Functional links between mutation occurrence and mRNA splicing alterations*

We next set out to determine the direct links between the mutations in either *DNAJC18* or *LMAN1* introns 2 or 7, respectively, with exon skipping events. To this end, we engineered plasmid minigenes in the PolR2G backbone, encompassing intron 2- exon 3- intron 3 for *DNAJC18* and intron 7- exon 8- intron 9 for *LMAN1*, carrying or not the mutations. These plasmid constructs were transfected into both MSI and non-MSI gastric and colorectal cancer cell lines. Post-transfection RNA was then analyzed by RT-PCR.

Our results showed that the RT-PCR product patterns were similar among all cell lines, regardless of whether these were MSI or non-MSI (HGT-1, AGS, SW480). In addition, aberrant splicing occurred in presence of the mutated sequences, as compared to the wt minigene. Non-mutated or mutated *LMAN1* minigenes, transfected into either MSI (HCT116, HCT8, CO115, LoVo and SNU1) and non-MSI cell lines (HGT-1, AGS, SW480) showed the same profiles, which were different between non-mutated or mutated sequences. For non-mutated *DNAJC18*, we observed the same profile between MSI (HCT116, HCT8, LoVo and SNU1) and non-MSI cell lines (HGT-1, AGS) and for mutated *DNAJC18*, MSI (HCT116, HCT8, Co115) and non-MSI cell lines (AGS) showed the same profile. Until now, the exact meaning of this splice effect was unknown (Figure 7).

### *5- MSI mutations in other cancer types*

We surveyed 96 additional tumors from tissues archived in Brest tumor bank, for the presence of somatic mutations. Several of the same mutations, as those found in gastric cancer samples, were shared with other cancers, mostly with colon cancers, albeit with somewhat different frequencies (supplementary Table 1). In addition, 2 mutations were common to nearly all cancer types (rs746845401 (LMAN1) and rs781215815 (ZNF43)).

# Discussion

In the present study, we identified somatic genomic alterations that occurred at high frequency in MSI gastric cancers. Previous characterization of cancer genomes had already reported that the *LMAN1* gene was frequently mutated in MSI gastric cancers (Roeckel 2009, Nagarajan 2012) and in MSI colorectal cancers (Roeckel, 2009). The mutation characterized here, from the repertoire of TCGA gastric cancers, was distinct from that identified by Roeckel *et al.* Although the mutant mRNA was predicted to undergo a Nonsense Mediated mRNA Decrease, we did not obtain any evidence that this was the case (data not shown), as also observed in the Roeckel study. Although the defect in mRNA assembly ensuing from skipping of exon 8 in a fraction of the *LMAN1* transcript pool could not be readily associated with a presumptive loss of function of the *LMAN1* protein, it is possible that it could lack part of its activity, as demonstrated by the Roeckel study, which reported a reduced secretion of Alpha 1 anti-trypsin in mutant cell lines.

Several of the 49 DNAJ members of the HSP40 protein family have specific roles in cancer, including in cell proliferation, motility or invasion (for review Mitra 2009 and Wu 2017). HSP40 chaperones belonging to the C subclass (DNAJC) need to interact - through their J domain - with either HSP or HSC70 proteins to select their substrates, are located in different sub-cellular compartments and expressed in selected tissues or cancer types. However, the function of the DNAJC18 protein is unknown. Because we could not find an antibody that would faithfully reveal expression of the normal DNAJC18 protein, it was not possible to address the potential effect of the mutation in intron 7 on the level of the normal protein, or of the putative, shorter protein, resulting from exon 3 exclusion. Additional experiments will be required to study this further. Nevertheless, as a frequently mutated protein in MSI gastric cancers, it is most probable that the increase in the shorter DNAJC18 protein and/or the decrease in the full-length protein will play a role in this cancer.

The mutations characterized in the *DNAJC18* and *LMAN1* genes led to a significant fraction of specific exon skipping in those genes in tissues and cells with a MSI phenotype.

We failed to report the effect of these mutations, we observed in vivo, in our minigenes transfection experiments. The only conclusion we can make right now is that the effect of these mutations is not dependent on the MSI phenotype. To show the existence of a direct link between MS mutations in these genes and altered pre-mRNA splicing, more experiments are necessary. Interestingly, the skipping of the exons just next to the intronic mutations was exact, extending from the first to the last nucleotide of the exons. This observation suggested that these mono-nucleotide deletions prevented spliceosome assembly at the 3' boundary site of the intron, and thus partially alleviated the recognition of intron-exon junctions.

Mutations, and specifically deletions of one or more nucleotides within the poly-pyrimidine intronic tracts have already been observed in digestive cancers with MS instability (Dorard et al., Nat. Med. 2011). Hence, splicing alterations in mononucleotide repeats have been found in MSI colorectal cancers and leukemia in the *MRE11* gene (Ham et al., Cancer Sci 2006), as also observed here in our gastric and colorectal cancer samples.

It was previously demonstrated that the deletion of one microsatellite base within a T(17) repetition a few nucleotides upstream from the acceptor site of HSP110 intron 8 led to exon 9 skipping, resulting in the synthesis of a truncated HSP110 isoform that opposed the full-length protein. Importantly, the chaperone anti-apoptotic function of HSP110 was lost (Dorard et al., Nat. Med. 2011). It was concluded that chemotherapy resistance could be suppressed in presence of the truncated HSP110 isoform. These data agreed to the fact that MSI gastric cancer patients have better prognosis when expressing a mutant HSP110 (Collura et al., Gastroenterology 2014).

These results, as well as ours strongly suggested that a single nucleotide deletion in an intronic poly-pyrimidine mononucleotide stretch was sufficient to trigger abnormal splicing, at exactly the proper intron-exon junction. Although the precise mechanism involved was not elucidated, it is quite plausible that either the binding of a protein onto the repetition was hampered or its ability to interact the U2 spliceosome subunit - responsible for the proper recognition of the 3' intron-exon boundary - was decreased, or that the RNA region was subjected to a defective folding, or even a combination of both.

The fact that the molecular phenotype of the MS mutations manifested itself regardless of the context, MSI or non-MSI, suggested that, although the somatic MS mutations occurred in MSI cancers, the splicing control elements may not depend on specifically altered *trans*-acting factors required for splicing the minigenes-encoded RNA precursors. This is, however, distinct for splicing the endogenous RNAs, as the splicing status depends on the MSI context of the cell lines.

Taken together, our results further document the functional links between cancer-associated alterations of microsatellite sequences and aberrant pre-mRNA splicing. Importantly, the mononucleotide sequences where these mutations occur belong to the mandatory *cis*-acting signals required for intron splicing. Much like it happens in genetic diseases, variations in the poly-pyrimidine tract are indeed well known to control the inclusion level of the downstream proximal exon (Dujardin et al, 2014). How gene-specific mutations are selected in the course of MSI cancer occurrence certainly remains a wide-opened question.

# Bibliographie

---

- Collura, Ada, Anaïs Lagrange, Magali Svrcek, Laetitia Marisa, Olivier Buhard, Agathe Guilloux, Kristell Wanherdrick, *et al.* "Patients with Colorectal Tumors with Microsatellite Instability and Large Deletions in Hsp110 T17 Have Improved Response to 5-Fluorouracil-Based Chemotherapy." *Gastroenterology* 146, no. 2 (2014): 401-11. e1.
- Dorard, Coralie, Aurélie De Thonel, Ada Collura, Laetitia Marisa, Magali Svrcek, Anaïs Lagrange, Gaetan Jego, *et al.* "Expression of a Mutant Hsp110 Sensitizes Colorectal Cancer Cells to Chemotherapy and Improves Disease Prognosis." *Nature medicine* 17, no. 10 (2011): 1283.
- Dujardin, Gwendal, Celina Lafaille, Manuel de la Mata, Luciano E Marasco, Manuel J Muñoz, Catherine Le Jossic-Corcós, Laurent Corcos, and Alberto R Kornblihtt. "How Slow Rna Polymerase Ii Elongation Favors Alternative Exon Skipping." *Molecular cell* 54, no. 4 (2014): 683-90.
- Ejima, Yosuke, Lichun Yang, and Masao S Sasaki. "Aberrant Splicing of the Atm Gene Associated with Shortening of the Intronic Mononucleotide Tract in Human Colon Tumor Cell Lines: A Novel Mutation Target of Microsatellite Instability." *International journal of cancer* 86, no. 2 (2000): 262-68.
- Giannini, Giuseppe, Christian Rinaldi, Elisabetta Ristori, Maria Irene Ambrosini, Fabio Cerignoli, Alessandra Viel, Ettore Bidoli, *et al.* "Mutations of an Intronic Repeat Induce Impaired Mre11 Expression in Primary Human Cancer with Microsatellite Instability." *Oncogene* 23, no. 15 (2004): 2640.
- Ham, Maria Francisca, Tetsuya Takakuwa, Wen-Juan Luo, Angen Liu, Akira Horii, and Katsuyuki Aozasa. "Impairment of Double-Strand Breaks Repair and Aberrant Splicing of Atm and Mre11 in Leukemia-Lymphoma Cell Lines with Microsatellite Instability." *Cancer science* 97, no. 3 (2006): 226-34.
- Kim, Tae-Min, Peter W Laird, and Peter J Park. "The Landscape of Microsatellite Instability in Colorectal and Endometrial Cancer Genomes." *Cell* 155, no. 4 (2013): 858-68.
- Mitra, Aparna, Lalita A Shevde, and Rajeev S Samant. "Multi-Faceted Role of Hsp40 in Cancer." *Clinical & experimental metastasis* 26, no. 6 (2009): 559-67.
- Nagarajan, Niranjan, Denis Bertrand, Axel M Hillmer, Zhi Jiang Zang, Fei Yao, Pierre-Étienne Jacques, Audrey SM Teo, *et al.* "Whole-Genome Reconstruction and Mutational Signatures in Gastric Cancer." *Genome biology* 13, no. 12 (2012): R115.
- Network, Cancer Genome Atlas Research. "Comprehensive Molecular Characterization of Gastric Adenocarcinoma." *Nature* 513, no. 7517 (2014): 202.
- Oltean, Sebastian, and DO Bates. "Hallmarks of Alternative Splicing in Cancer." *Oncogene* 33, no. 46 (2014): 5311.
- Pérez-Jiménez, Marga, Guillaume Besnard, Gabriel Dorado, and Pilar Hernandez. "Varietal Tracing of Virgin Olive Oils Based on Plastid DNA Variation Profiling." *PLoS One* 8, no. 8 (2013): e70507.
- Phumichai, Chalernpol, Thitaporn Phumichai, and Arunee Wongkaew. "Novel Chloroplast Microsatellite (Cpssr) Markers for Genetic Diversity Assessment of Cultivated and Wild Hevea Rubber." *Plant molecular biology reporter* 33, no. 5 (2015): 1486-98.
- Roeckel, Nina, Stefan M Woerner, Matthias Kloor, Yan-Ping Yuan, Georgios Patsos, Roland Gromes, Juergen Kopitz, and Johannes Gebert. "High Frequency of Lman1 Abnormalities in Colorectal Tumors with Microsatellite Instability." *Cancer research* 69, no. 1 (2009): 292-99.

- Rugge, Massimo, Matteo Fassan, and David Y Graham. "Epidemiology of Gastric Cancer." In *Gastric Cancer*, 23-34: Springer, 2015.
- Suraweera, Nirosha, Alex Duval, Maryline Reperant, Christelle Vaury, Daniela Furlan, Karen Leroy, Raquel Seruca, Barry Iacopetta, and Richard Hamelin. "Evaluation of Tumor Microsatellite Instability Using Five Quasimonomorphic Mononucleotide Repeats and Pentaplex Pcr." *Gastroenterology* 123, no. 6 (2002): 1804-11.
- Velho, Sérgio, Maria Sofia Fernandes, Marina Leite, Ceu Figueiredo, and Raquel Seruca. "Causes and Consequences of Microsatellite Instability in Gastric Carcinogenesis." *World Journal of Gastroenterology: WJG* 20, no. 44 (2014): 16433.
- Vieira, Maria Lucia Carneiro, Luciane Santini, Augusto Lima Diniz, and Carla de Freitas Munhoz. "Microsatellite Markers: What They Mean and Why They Are So Useful." *Genetics and molecular biology* 39, no. 3 (2016): 312-28.
- Wang, Yang, and Zefeng Wang. "Efficient Backsplicing Produces Translatable Circular Mrnas." *Rna* 21, no. 2 (2015): 172-79.
- Wu, Jianming, Tuo Liu, Zechary Rios, Qibing Mei, Xiukun Lin, and Shousong Cao. "Heat Shock Proteins and Cancer." *Trends in pharmacological sciences* 38, no. 3 (2017): 226-56.



# Figures

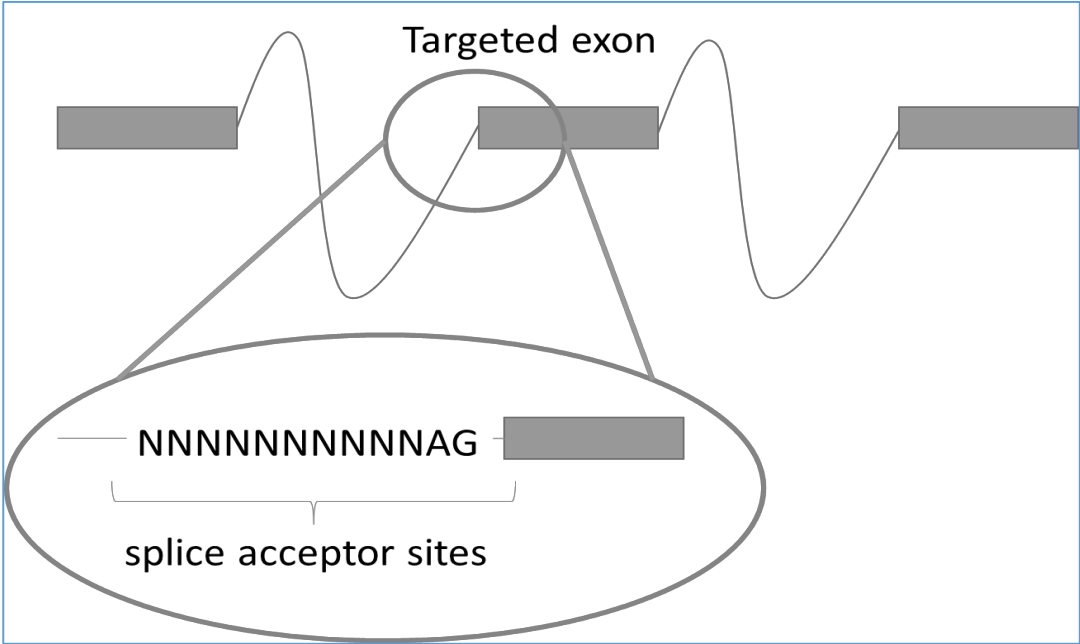


Figure 1: Identification of new target genes with frequent mutations in the TCGA gastric adenocarcinoma database

| GENE    | NAME  | DNA change (GRCh37/hg19/Human) | UCSC code   |
|---------|---|--------------------------------|-------------|
| ATL3    | atlastin GTPase 3   | chr11:g.63420050 A>-           | rs776250621 |
| DDX6    | DEAD (Asp-Glu-Ala-Asp) box helicase 6                       | chr11:g.118629614 ->G          | rs782154217 |
| DNAJC18 | DnaJ heat shock protein family (Hsp40) memberC18            | chr5:g.138764375 A>-           | rs781620531 |
| LMAN1   | lectin, mannose binding 1                                   | chr18:g.57013285 T>-           | rs759519259 |
| PREPL   | prolyl endopeptidase-like                                   | chr2:g.44550553 A>-            | rs773390043 |
| PSME4   | proteasome activator subunit 4                              | chr2:g.54167140 G>-            | rs768165363 |
| SLCO6A1 | solute carrier organic anion transporter family, member 6A1 | chr5:g.101755728 A>-           | rs764367397 |
| ZNF43   | zinc finger protein 43                                      | chr19:g.22002026 A>-           | rs766585124 |

Table 1 :List of mutations localized in close proximity to splice acceptor sites according to the TCGA gastric adenocarcinoma database

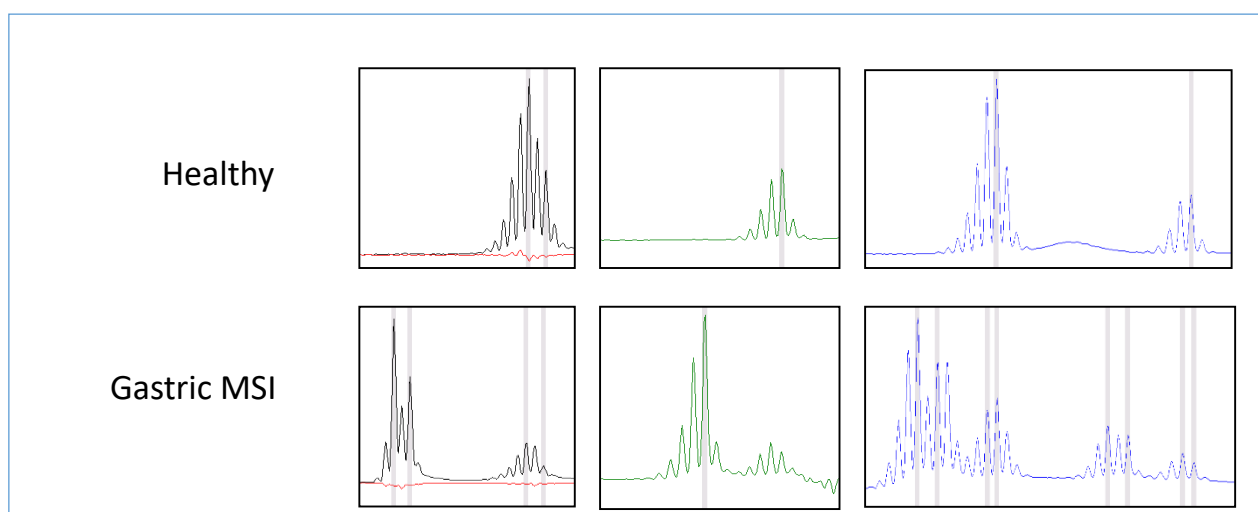


Figure 2a : Fragment analysis of microsatellite loci from healthy tissue DNA and gastric tumors' DNA.

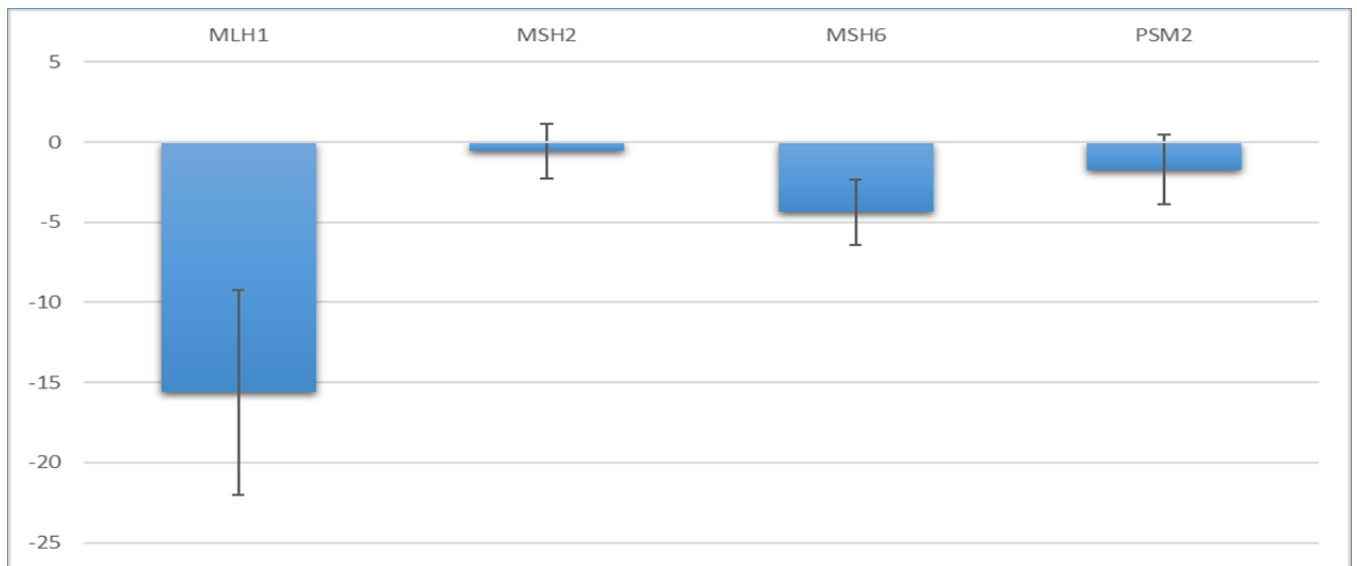


Figure 2b : Analysis of the relative expression level of 4 genes (*MLH1*, *MSH2*, *MSH6*, *PSM2*) involved in the repair of DNA replication errors (MMR system) between gastric MSI tumors (N = 3) and healthy unpaired tissues (N = 5) by quantitative RT-PCR.

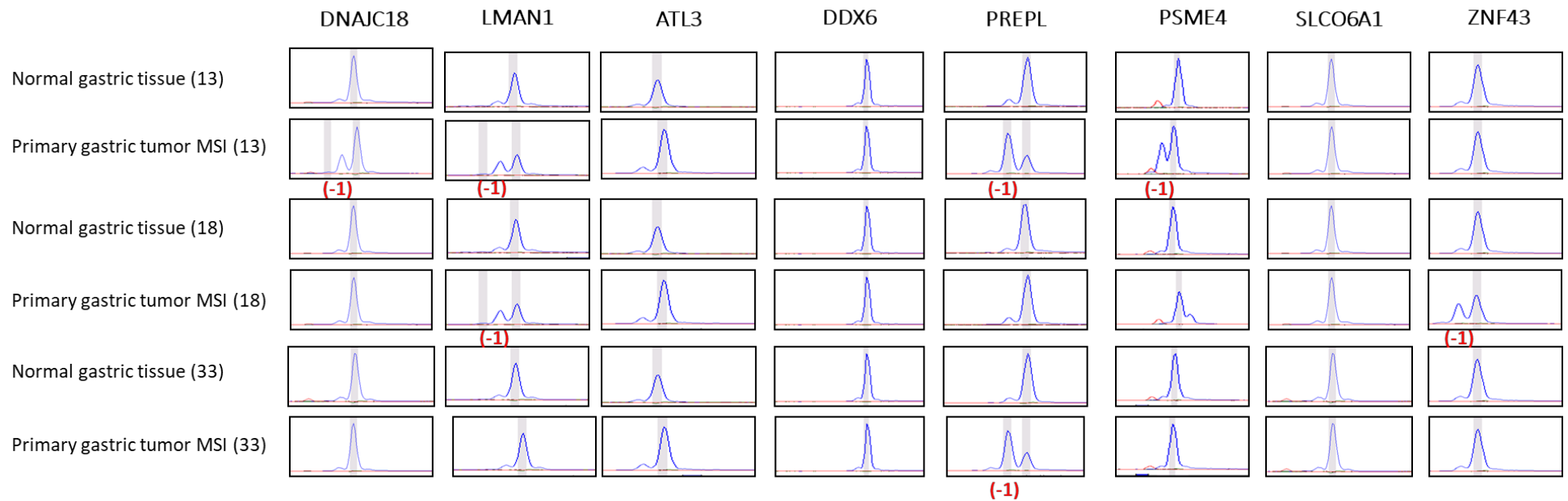


Figure 3a : Allelic profiles for several MSI primary gastric tumors and controls at the sequence of interest for intronic *DNAJC18*, *LMAN1*, *ATL3*, *DDX6*, *PREPL*, *PSME4*, *SLCO6A1* and *ZNF43*. Numbers in parentheses indicate the size of the deletion in MSI tumor samples (in base pairs).

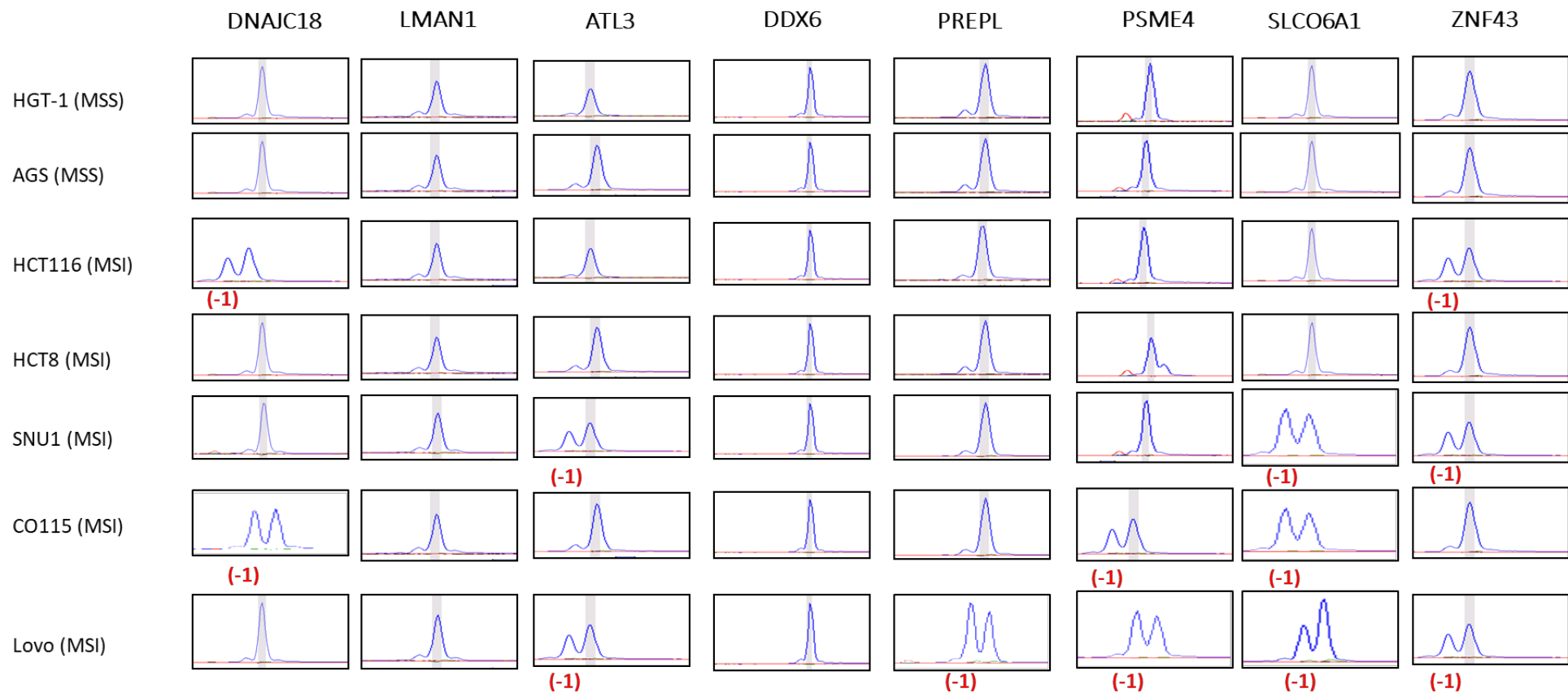


Figure 3b : Allelic profiles for several CRC and gastric cell lines at the sequence of interest for intronic *DNAJC18*, *LMAN1*, *ATL3*, *DDX6*, *PREPL*, *PSME4*, *SLCO6A1* and *ZNF43*. Numbers in parentheses indicate the size of the deletion in MSI tumor samples (in base pairs).

### DnajC18 : DnaJ heat shock protein family (Hsp40) member C18

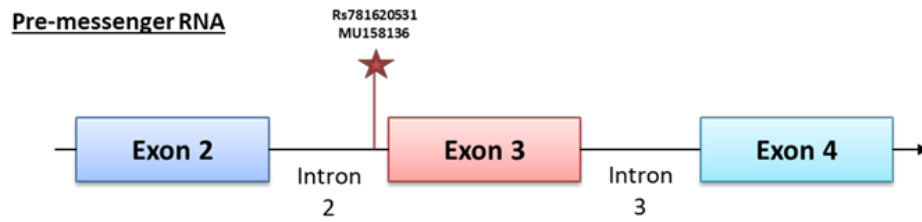


Figure 4a: Localization of the rs781620531 mutation identified in the TCGA database and found in our sample of MSI patients.

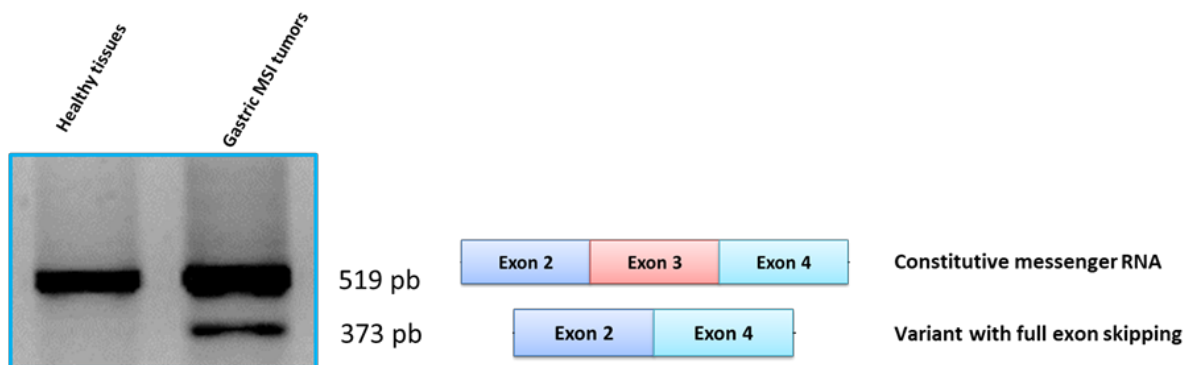


Figure 4c: DnajC18 expression in MSI gastric tissue and matched healthy tissue

### LMAN1 : Lectin, mannose binding 1

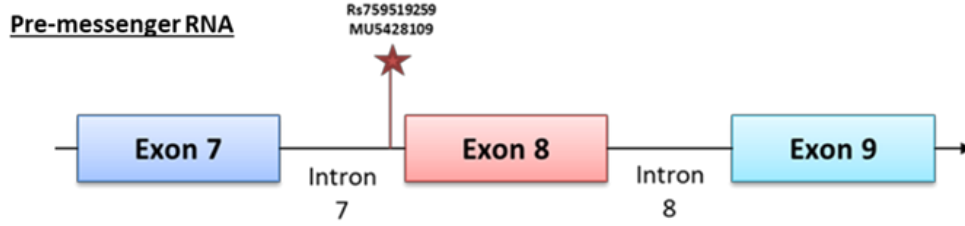


Figure 4b: Localization of the rs759519259 mutation identified in the TCGA database and found in our sample of MSI patients.

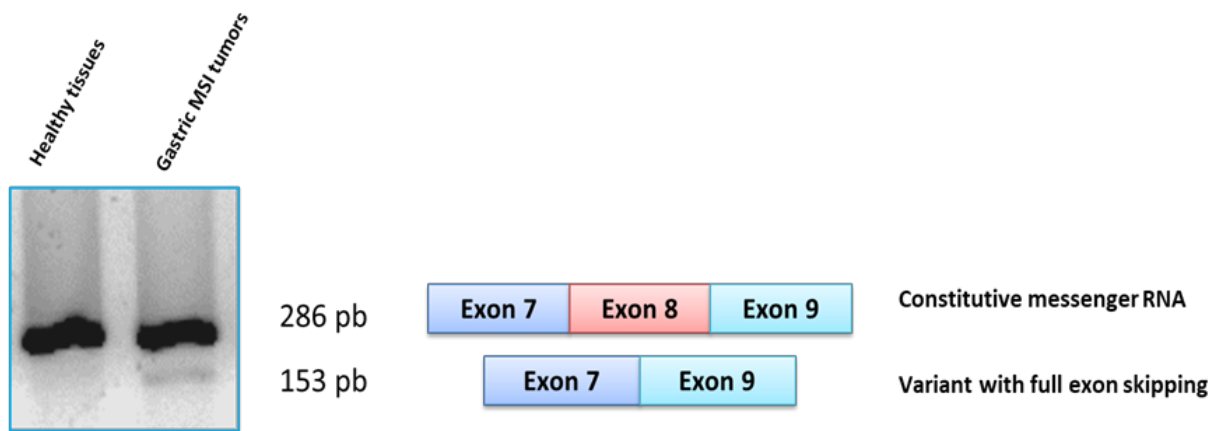


Figure 4d: LMAN1 expression in MSI gastric tissue and matched healthy tissue

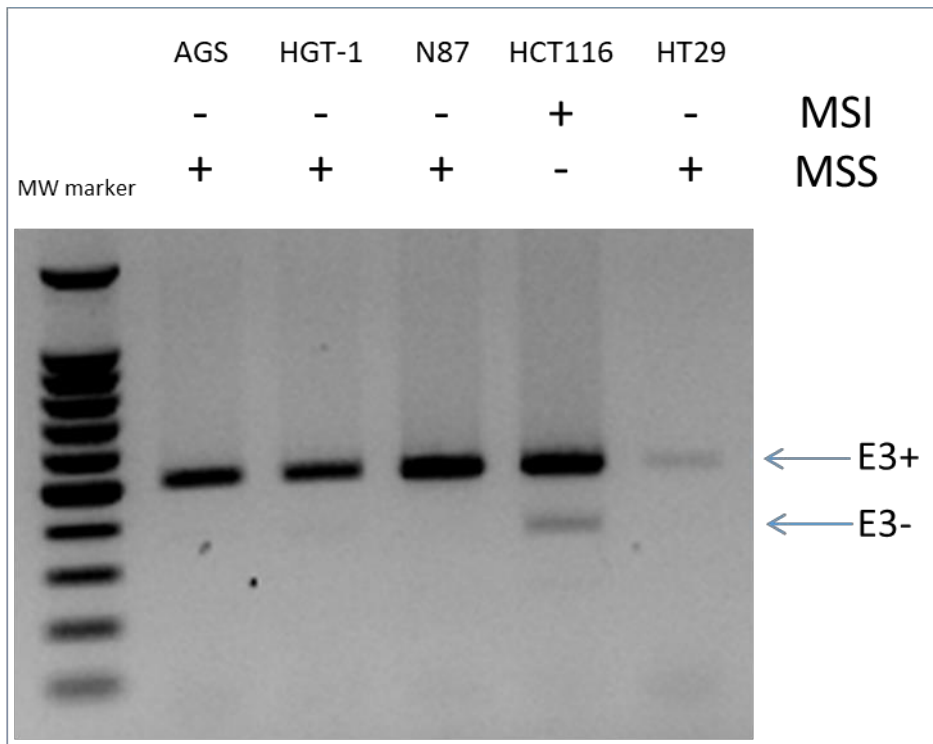


Figure 4e: RT-PCR analysis of DNAJC18 transcript in gastric and colorectal cell lines. Partial exon 3 skipping is associated with the MSI (HCT116) status.

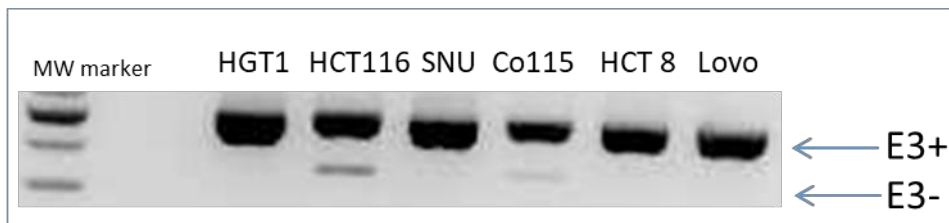


Figure 4f: RT-PCR analysis of DNAJC18 transcript in gastric and colorectal cell lines. Partial exon 3 skipping is associated with the MSI (HCT116 and Co115) status.



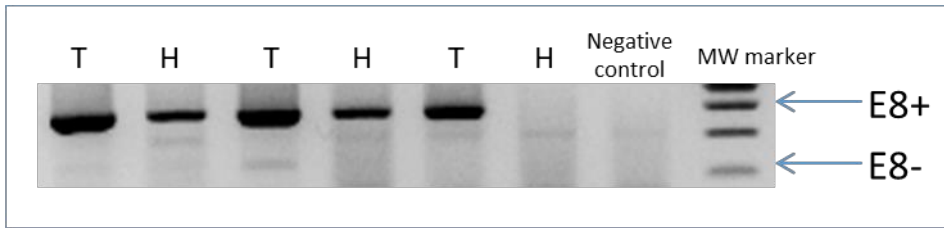


Figure 4g: RT-PCR analysis of LMAN1 transcript in MSI gastric cancer and healthy tissue. Partial exon 8 skipping is associated with the cancer status. T, tumor; H, healthy tissue.

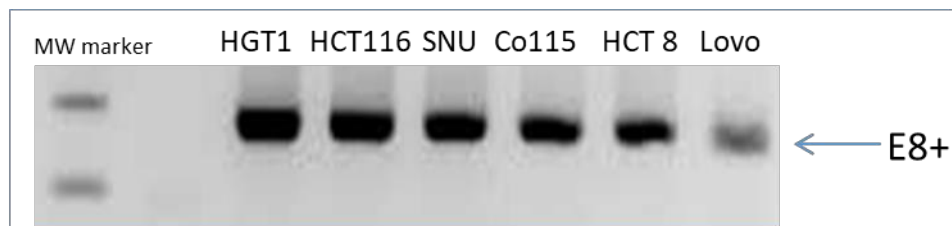


Figure 4h: RT-PCR analysis of LMAN1 transcript in gastric and colorectal cell lines. No exon skipping in this MSI cell lines.

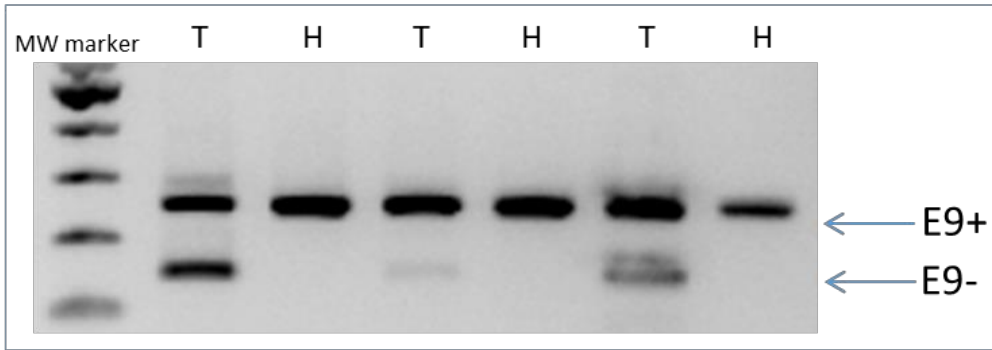


Figure 5a: RT-PCR analysis of HSP110 transcript in MSI gastric cancer and healthy tissue. Partial exon 9 skipping is associated with the cancer status. T, tumor; H, healthy tissue.

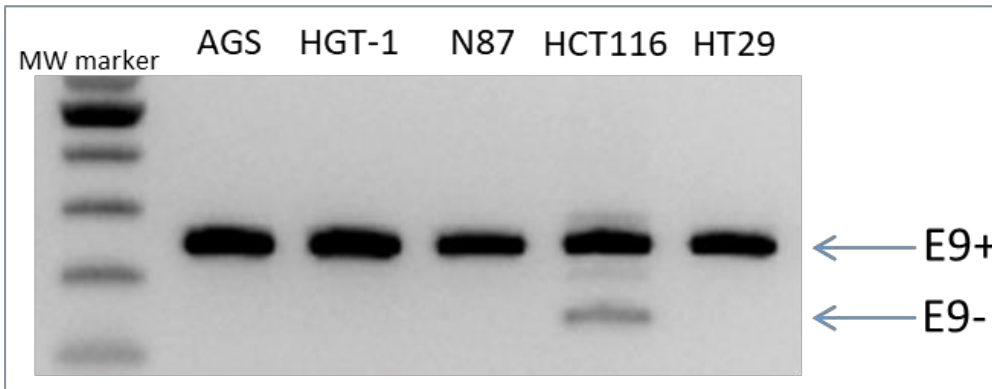


Figure 5b: RT-PCR analysis of HSP110 transcript in gastric and colorectal cell lines. Partial exon 9 skipping is associated with the MSI (HCT116) status.

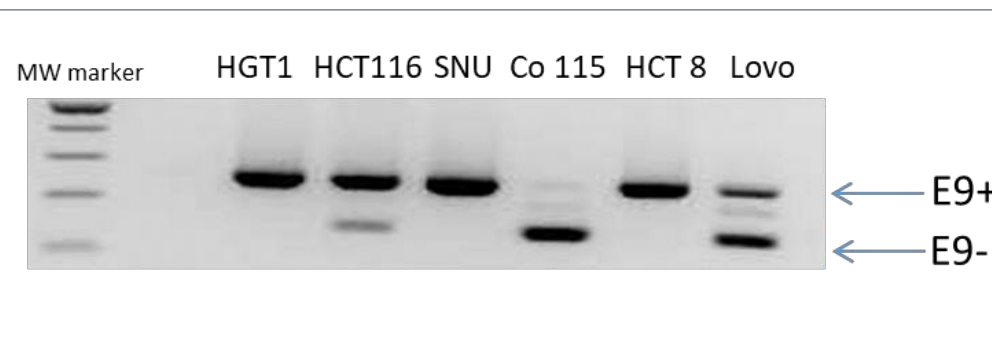


Figure 5c: RT-PCR analysis of HSP110 transcript in gastric and colorectal cell lines. Partial exon 9 skipping is associated with the MSI status.

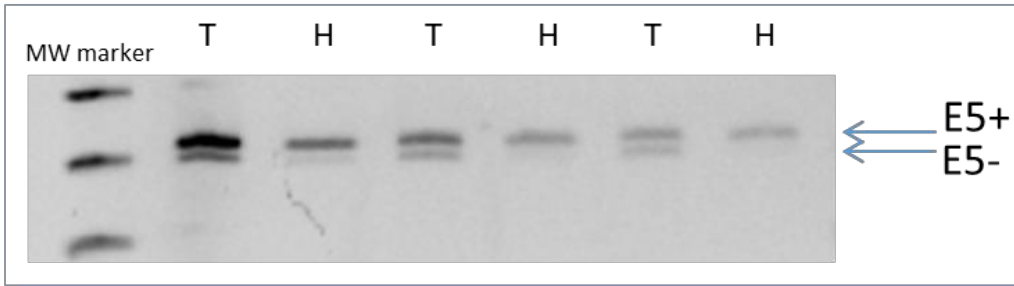


Figure 6a: RT-PCR analysis of ATM expression in MSI gastric cancers and healthy tissues. Partial exon 8 skipping is associated with the cancer status. T, tumor; H, healthy tissue.

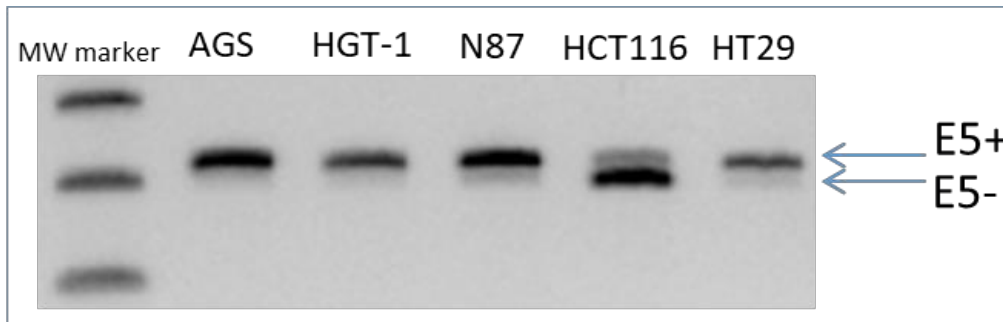


Figure 6b: RT-PCR analysis of ATM transcript in gastric and colorectal cell lines. Strong exon 8 skipping is associated with the MSI (HCT116) status.



Figure 6c: RT-PCR analysis of ATM transcript in gastric and colorectal cell lines. Strong exon 8 skipping is associated with the MSI status.

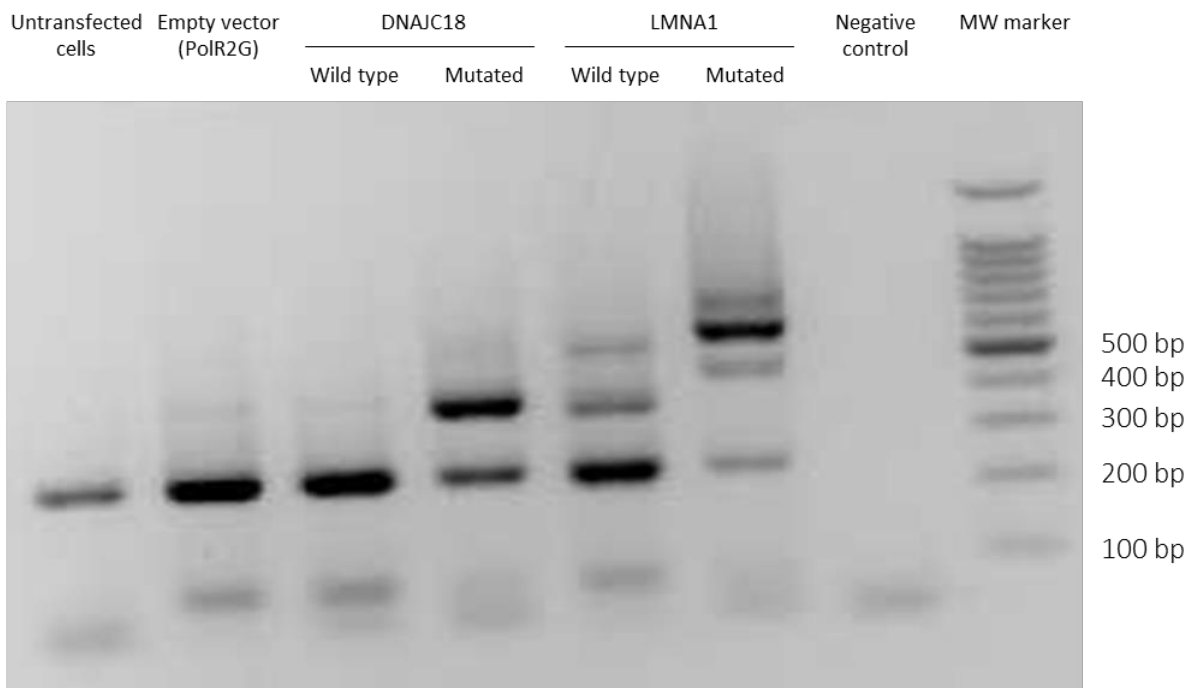


Figure 7a : RT-PCR analysis of products generated by the splicing assay in MSI colorectal cell lines HCT116.

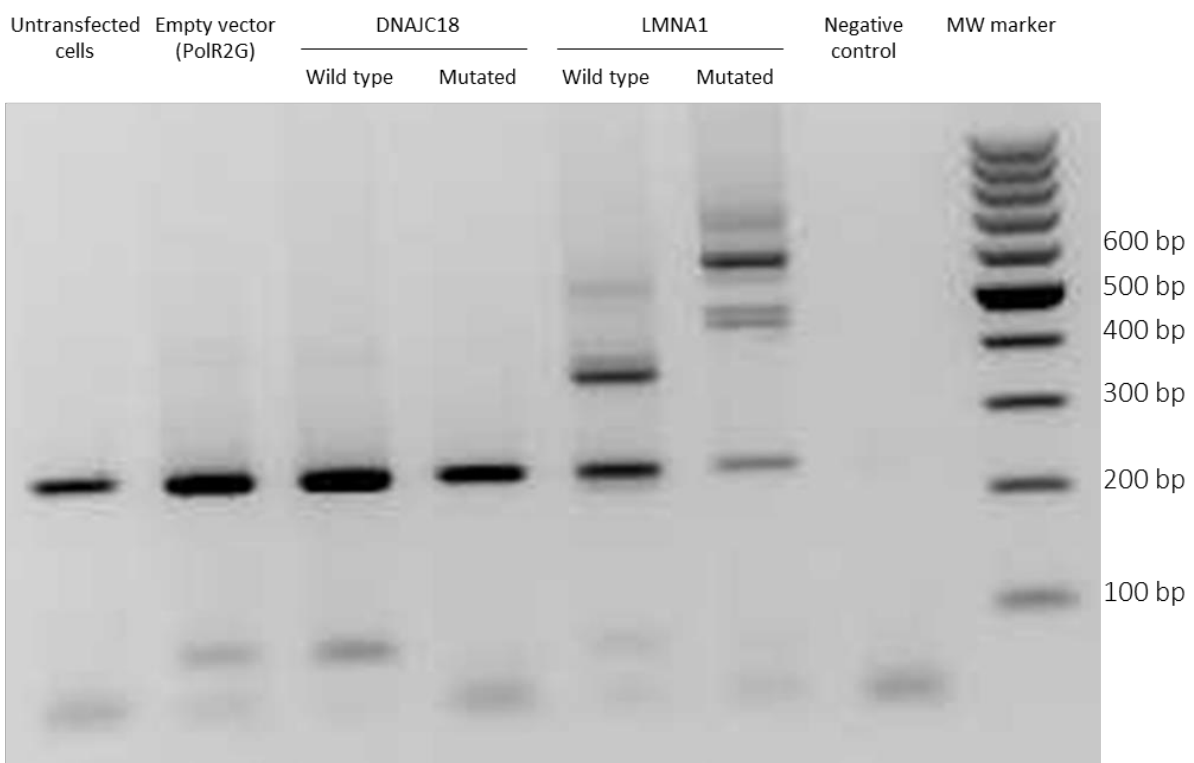


Figure 7b : RT-PCR analysis of products generated by the splicing assay in MSI colorectal cell lines HCT8.

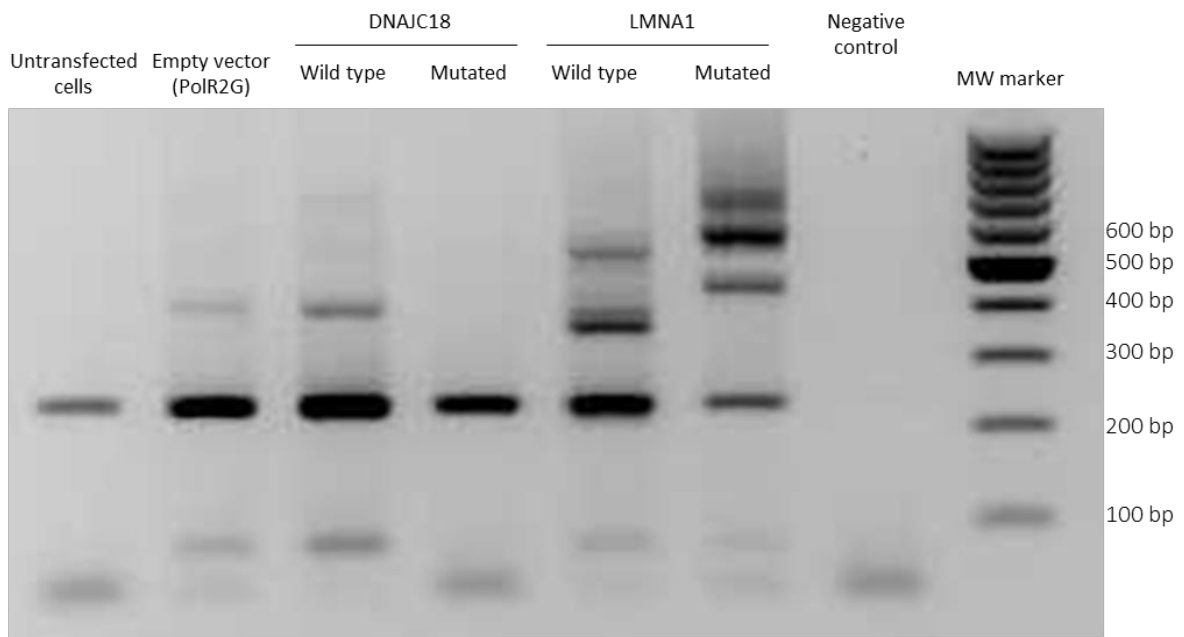


Figure 7c : RT-PCR analysis of products generated by the splicing assay in MSI colorectal cell lines CO115.

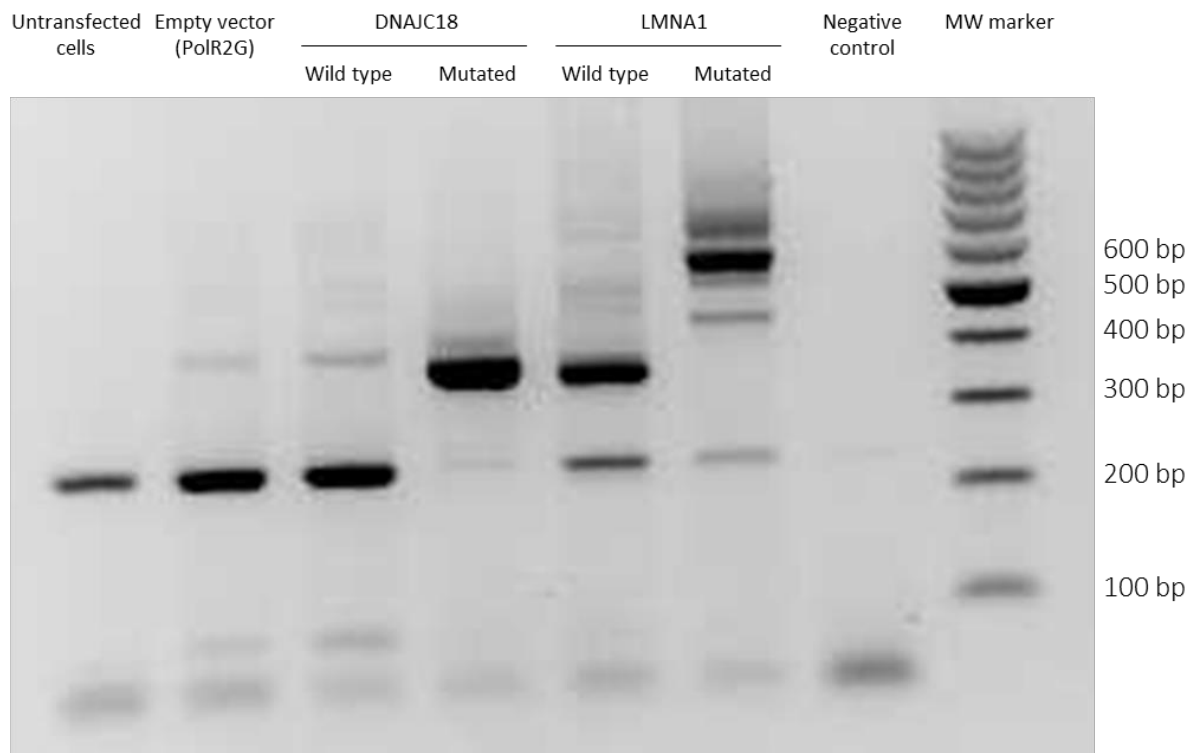


Figure 7d : RT-PCR analysis of products generated by the splicing assay in MSI colorectal cell lines LoVo.

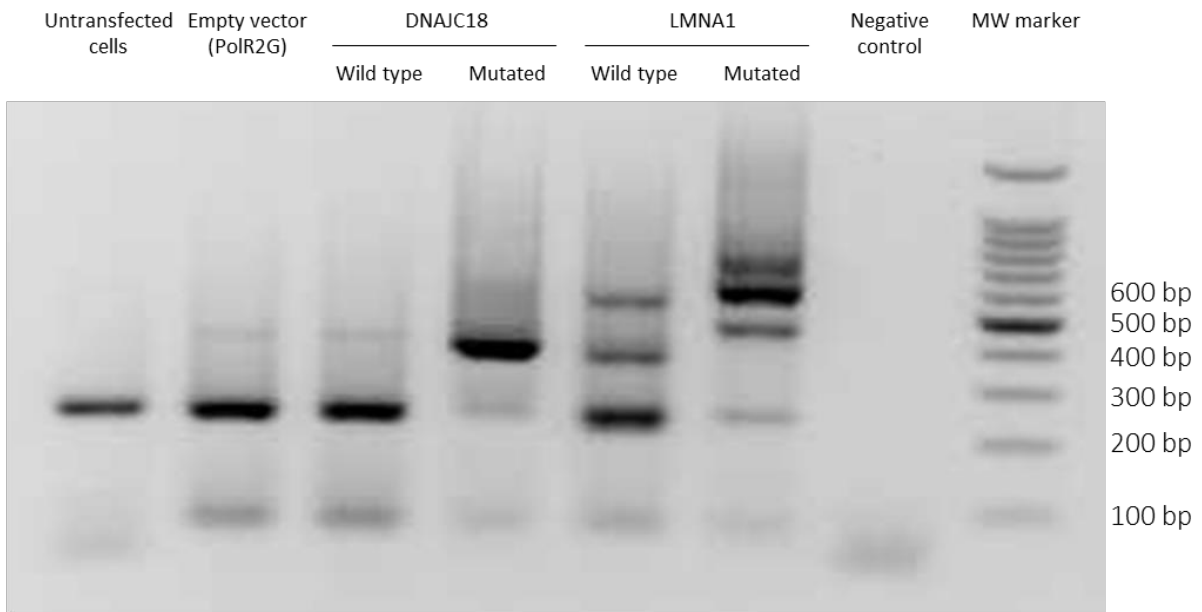


Figure 7e : RT-PCR analysis of products generated by the splicing assay in MSI gastric cell lines SNU1.

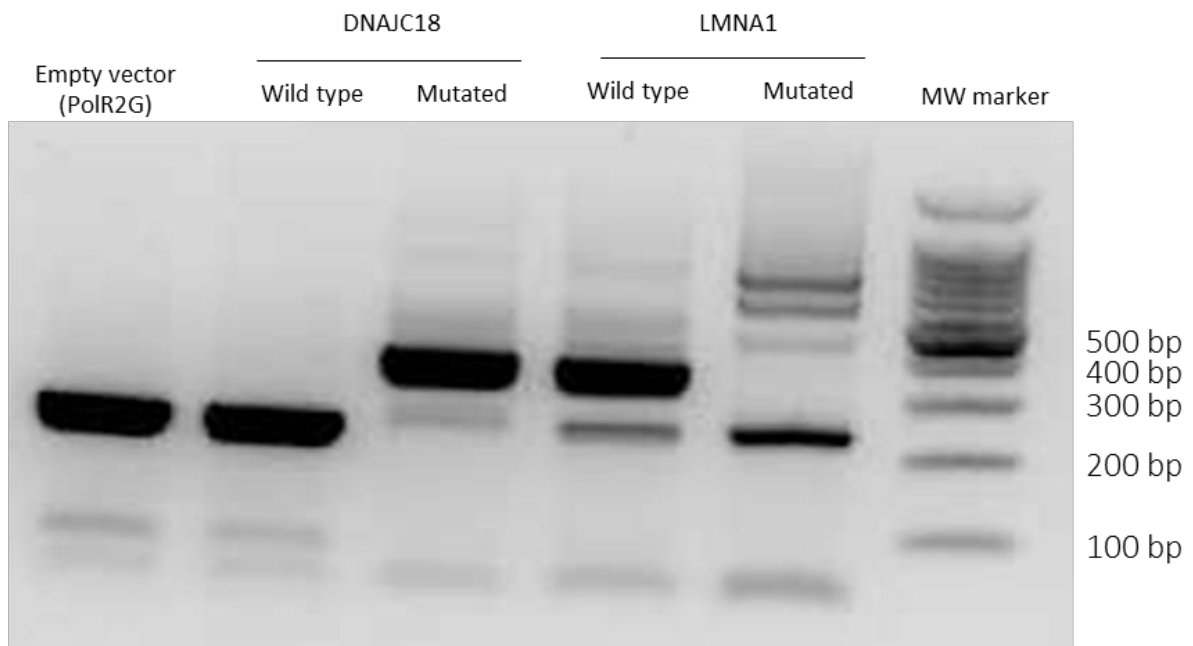


Figure 7f : RT-PCR analysis of products generated by the splicing assay in MSS gastric cell lines HGT-1.

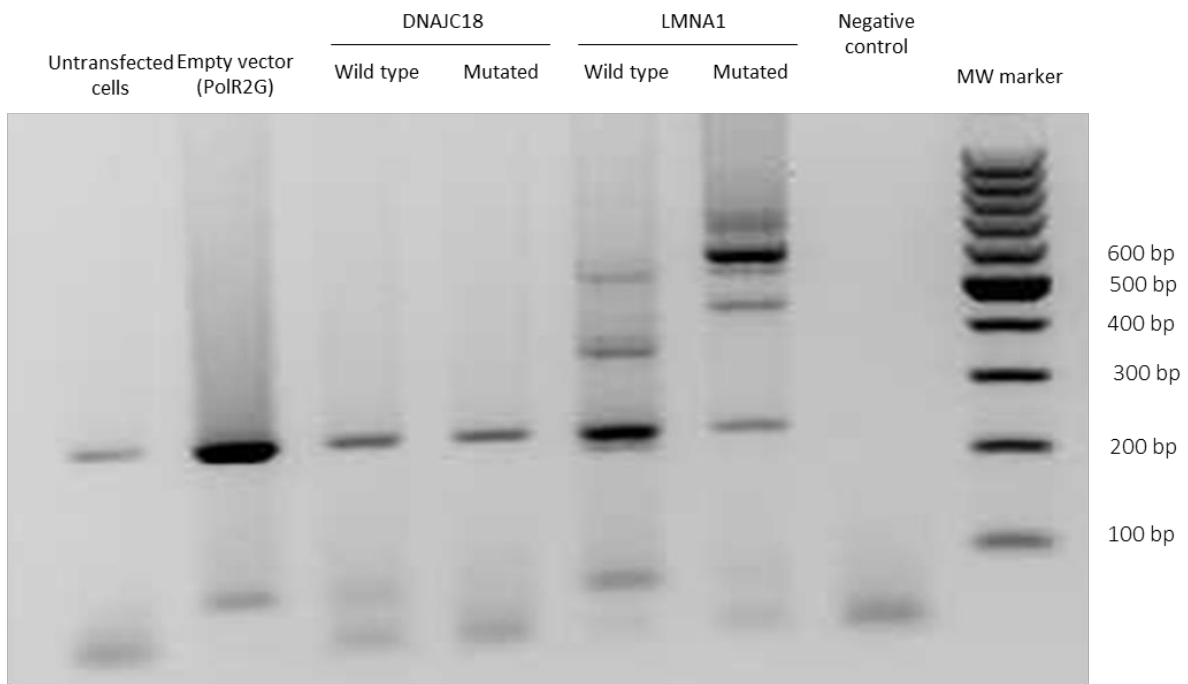


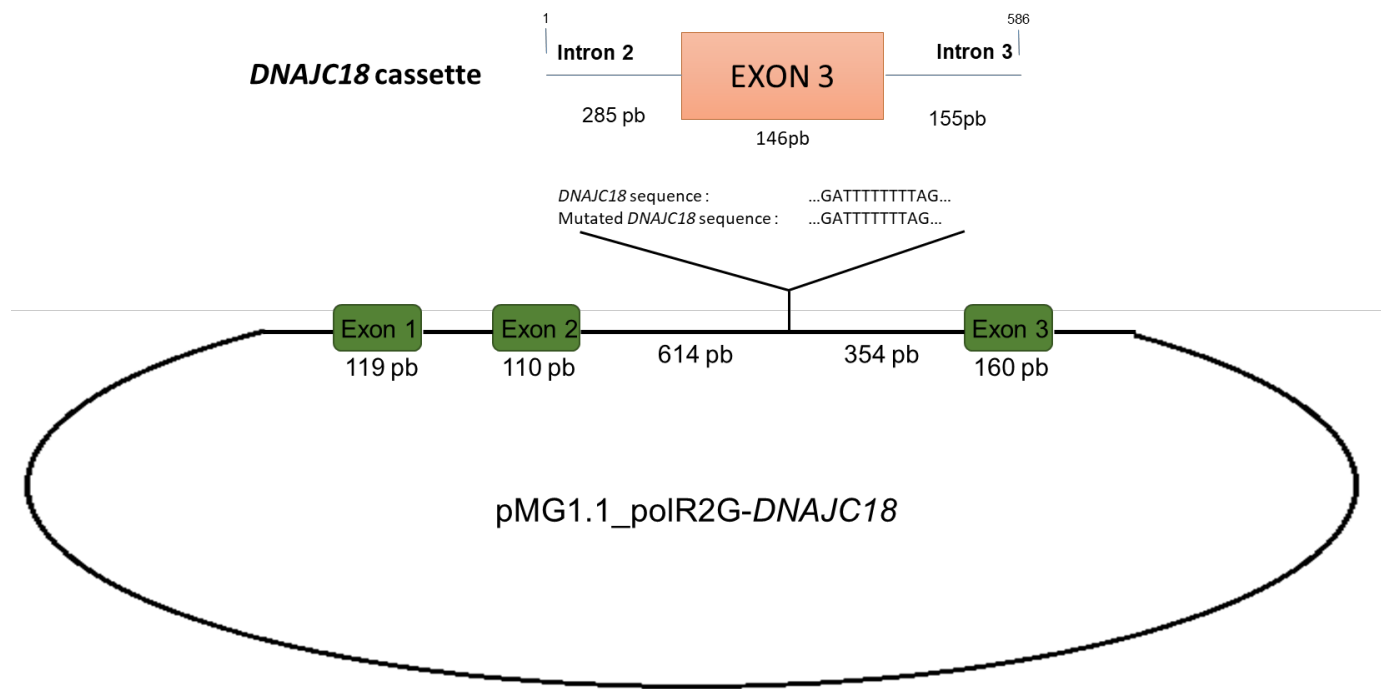
Figure 7g : RT-PCR analysis of products generated by the splicing assay in MSS gastric cell lines AGS.

**Supplementary Table 1: MSI mutations in cancer types other than gastric cancer.** Recurrent somatic alterations through different types of tumors archived at the Brest CHRU with microsatellite region instability (MSI) represented the number of alterations found in the tumors (white = none to blue = high).

| Primary Site                | N  | CDC14       | ADAM28      | PAMR1       | XYLT2        | DNAJC18     | FRMD4       | ATL3        | DDX6        | FHOD3       | ZBTB20      | PSME4       | GSE1        | LARP4B | SLCO6A1     | ZNF43       | PREPL       | RNF43        | LMAN1       |             |
|-----------------------------|----|-------------|-------------|-------------|--------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|--------|-------------|-------------|-------------|--------------|-------------|-------------|
|                             |    | RS773911500 | RS770638998 | RS939607516 | RS1019214909 | RS781620531 | RS990652950 | RS776250621 | RS782154217 | RS777980924 | RS777745612 | RS768165363 | RS754199513 | N.D.   | RS764367397 | RS766585124 | RS773390043 | RS7812158115 | RS759519259 | RS746845401 |
| Colorectal (adenocarcinoma) | 66 | 2           | 4           | 4           | 5            | 7           | 7           | 9           | 9           | 11          | 12          | 13          | 14          | 14     | 16          | 17          | 24          | 38           | 40          | 44          |
| Colorectal (adenoma)        | 5  | 3           |             |             |              |             |             |             | 2           |             |             | 2           |             |        | 1           |             |             |              |             | 1           |
| Gastric (adenocarcinoma)    | 5  |             | 1           |             | 1            | 1           | 1           | 2           | 1           | 1           |             |             | 1           | 1      | 1           | 2           | 2           | 3            | 3           | 4           |
| Ovary / Peritoneum          | 10 |             | 1           |             | 1            | 2           |             |             | 1           | 1           | 1           | 2           |             | 2      |             | 5           | 2           | 3            | 2           | 6           |
| Skin                        | 4  |             |             |             |              |             |             | 2           |             |             | 1           | 1           | 1           |        | 1           | 1           | 1           |              |             |             |
| Utérus                      | 4  |             | 1           |             |              |             |             |             |             |             | 1           | 1           |             | 1      |             | 2           |             | 1            | 1           | 3           |
| Brain                       | 1  |             |             |             |              |             |             |             |             |             |             |             | 1           |        |             |             |             |              |             |             |

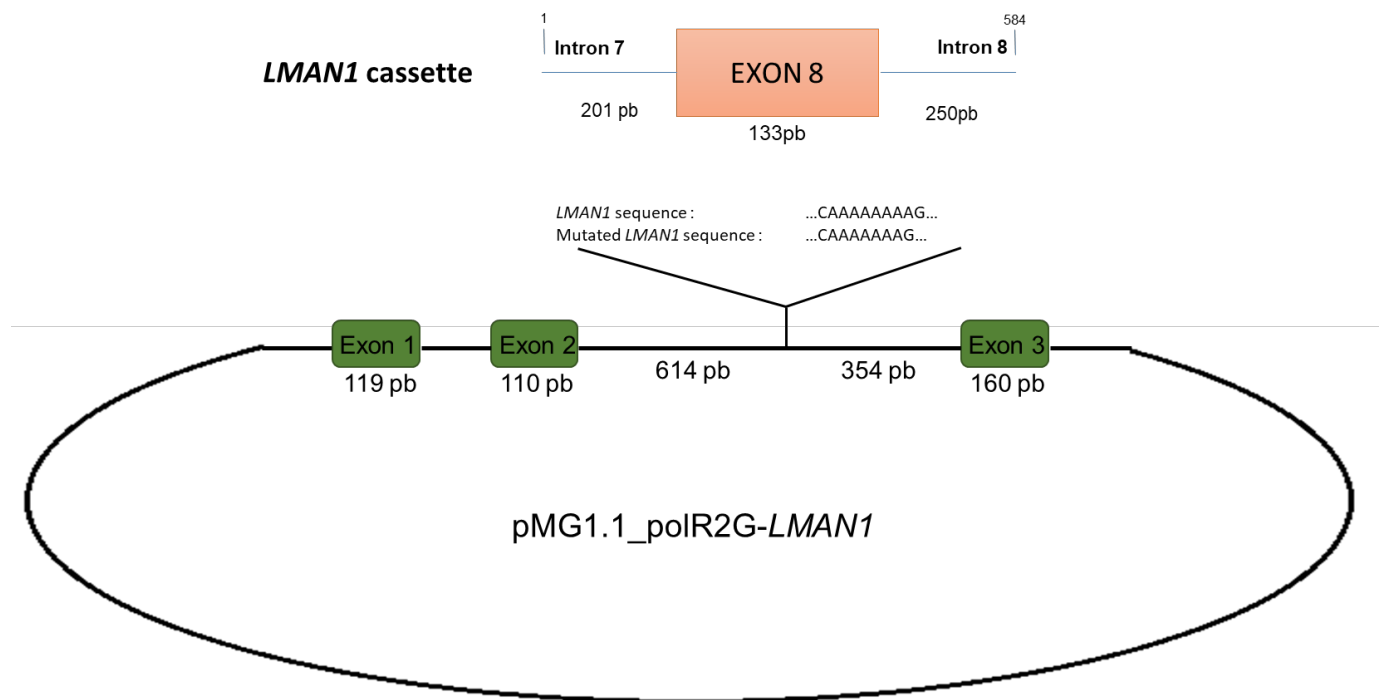


## Supplementary Figure 1



**Supplementary Figure 1: Structure of the pMG1.1\_polR2G-DNAJC18 minigene.** The upper part shows the *DNAJC18* genomic cassette, with the numbering starting at the first base of the cassette and ending at the last base of the cassette (586 bp). The lower part shows the pMG1.1\_polR2G-DNAJC18 minigene plasmid with the *DNAJC18* or the mutated *DNAJC18* exon 3 sequence. The white boxes on each side of the *DNAJC18* sequence are from the pMG1.1\_polR2G plasmid, and derived from the polR2G genomic sequence.

## Supplementary Figure 2



**Supplementary Figure 2: Structure of the *pMG1.1\_polR2G-LMAN1* minigene.** The upper part shows the *LMAN1* genomic cassette, with the numbering starting at the first base of the cassette and ending at the last base of the cassette (584 bp). The lower part shows the *pMG1.1\_polR2G-LMAN1* minigene plasmid with the *LMAN1* or the mutated *LMAN1* exon 8 sequence. The white boxes on each side of the *LMAN1* sequence are from the *pMG1.1\_polR2G* plasmid, and derived from the *polR2G* genomic sequence.

# Discussion

---

Dans ce qui va suivre, nous allons diviser la discussion en deux aspects directement liés à ces deux mutations retrouvées dans les gènes DNAJC18 et LMAN1. Le premier aspect concerne les mutants DNAJC18 et LMAN1 et leurs potentielles propriétés. Le deuxième aspect concerne l'effet de ces mutations et viendra aborder un aspect plus mécanistique de l'épissage.

Tout d'abord les membres de la famille HSP40 / DNAJ constituent l'un des plus grands groupes de chaperons moléculaires. Les HSP40 ont pour rôle de contrôler les HSP70, régulant leur activité et leur spécificité. Les différentes HSP40 sont impliquées dans des processus tels que le repliement des HSP70, le transport intracellulaire et transmembranaire, le remodelage des complexes protéiques et la dégradation de certaines protéines. Plusieurs de ces 49 membres de la famille de protéines HSP40 ont également des rôles spécifiques dans le cancer, y compris dans la prolifération, la motilité ou l'invasion de cellules (Mitra 2009 et Wu TIPS 2017). Les chaperons HSP40 appartenant à la sous-classe C (DNAJC) doivent pouvoir interagir, via leur domaine J, avec les protéines HSP ou HSC70 pour sélectionner leurs substrats. Elles sont situées dans différents compartiments intracellulaires et exprimées dans des tissus ou types de cancers sélectionnés. Cependant, la fonction de la protéine DNAJC18 est inconnue. Un exemple intrigant a récemment été présenté concernant le cycle de vie du virus de la dengue (DENV). DENV exploite la machinerie HSP70 des cellules, via des DNAJ distinctes, pour favoriser l'entrée, la réplication et l'assemblage du virus (Taguwa et al., 2015). Les protéines DNAJC9 et DNAJC16 et DNAJC18 sont ancrées dans la membrane, où elles exposent le domaine J au cytosol. Elles sont impliquées dans la facilitation de l'entrée du virus, tandis que les étapes post-entrée (réplication et production virales) nécessitent DNAJA2, DNAJB6, DNAJB7, DNAJB11 et DNAJC10, à différentes étapes. Guey et al décrivent en 2010, l'existence d'un variant associé à une forme agressive du carcinome de la vessie. Zoppino et al, en 2018 ont étudié les changements d'expression des HSP dans le cadre de différents cancers. Ils ont trouvé de nombreuses HSP dont l'expression étaient dérégulée, parfois de façon spécifique à certains sous-types moléculaires, tandis que d'autres étaient partagées par différents sous-types intrinsèques. DNAJC18 a présenté une diminution de son niveau d'expression dans tous les sous-types moléculaires de cancers du sein, tout comme DNAJB4, HSPA12A, HSPA12B, HSPB2, HSPB6 et HSPB7.

Il est également à noter que nous avons montré que l'ARN de DNAJC18 mutant contient un codon stop prématuré dans l'exon 4, ce qui devrait le rendre sensible au NMD, c'est-à-dire qu'il devrait être reconnu et dégradé par le système de surveillance et de dégradation de l'ARNm. Après 24h de traitement à la cycloheximide, les transcrits de DNAJC18 étaient facilement détectables par analyse RT-PCR dans la lignée cellulaire HCT116, où la mutation est retrouvée et semble donc résister au NMD. Comme nous ne pouvions pas trouver d'anticorps qui révélerait fidèlement l'expression de la protéine DNAJC18 normale, il n'a pas été possible d'aborder l'effet potentiel de la mutation de l'intron 2 sur le niveau de la protéine normale, ni de la protéine plus courte supposée, qui résulterait de l'exclusion de l'exon 3. Des expériences supplémentaires seront nécessaires pour approfondir cette question. Néanmoins, en tant que protéine fréquemment mutée dans les cancers gastriques MSI, il est fort probable que l'augmentation de la protéine DNAJC18 mutante et/ou la diminution de la protéine constitutive joue un rôle dans ce type de cancer.

Concernant LMAN1, des caractérisations antérieures de génomes de cancer MSI avaient déjà signalé une fréquence de mutation importante (Roeckel 2009, Nagarajan 2012) et principalement dans le cancer colorectal MSI (Roeckel 2009). La protéine LMAN1 est située entre le compartiment intermédiaire du réticulum endoplasmique et du Golgi, une structure membranaire mobile qui transporte des protéines sécrétoires le long des microtubules, du RE au Golgi (Stephens et Pepperkok, 2001). La mutation décrite ici dans le répertoire des cancers gastriques TCGA était distincte de celle identifiée par Roeckel et al. Bien que l'on prévoyait que l'ARNm mutant subirait une diminution induite par un non-sens, nous n'avons obtenu aucune preuve que ce soit le cas, comme l'a également observé l'étude Roeckel. Bien que le défaut d'assemblage d'ARNm résultant du saut de l'exon 8 dans une fraction du pool de transcrits de LMAN1 ne puisse pas être facilement associé à une perte présomptive de la fonction de la protéine LMAN1, il est possible que celle-ci manque d'une partie de son activité, comme l'a démontré l'étude de Roeckel, qui a rapporté une sécrétion réduite d'alpha 1 anti-trypsine (A1AT) dans des lignées cellulaires mutantes. L'A1AT exerce non seulement une fonction anti-protéolytique, mais confère également une activité angio-inhibitrice. Des études sur des xénogreffes de cellules tumorales chez des souris nude ont montré que l'administration systémique d'A1AT retardait la progression tumorale et réduisait la densité en vaisseaux sanguins (Huang et al, 2004).

Des analyses comparatives de l'expression de puces à ADN sur des tissus humains normaux et tumoraux ont révélé que des taux plus bas de transcrits d'A1AT étaient corrélés à une taille de tumeur plus grande (Huang et al, 2004) et que des taux plus élevés d'A1AT non circulante dans les tumeurs étaient généralement associés à un meilleur pronostic (Allgayer et al, 1998). Fait intéressant, il a été émis l'hypothèse que A1AT protège contre le développement du cancer MSI-H – colorectal (Grodstein et al, 1999). Les travaux de Huang et ses collègues (Huang et al, 2004) montrent également des variations significatives des taux locaux d'A1AT dans les tissus tumoraux humains. Notre observation de l'hétérogénéité intratumorale du déficit en LMAN1 local pourrait fournir une explication moléculaire à cette découverte (Nyfeler et al, 2008).

Les mutations caractérisées dans les gènes DNAJC18 et LMAN1 ont entraîné une fraction significative du saut d'exon spécifique dans ces gènes dans les tissus et les cellules présentant un phénotype MSI. Nous avons ensuite montré que l'effet de ces mutations pouvait être en partie récapitulé dans des expériences de transfection de minigènes, confortant ainsi l'existence d'un lien entre ces mutations MSI et l'épissage altéré des pré-ARNm. Fait intéressant, ces mutations introniques entraînent dans les deux cas que nous avons observé un saut d'exon complet, du premier au dernier nucléotide des exons concernés.

Cette observation suggère que ces délétions mononucléotidiques peuvent empêcher l'assemblage du spliceosome au niveau du site 3' de l'intron et atténue ainsi partiellement la reconnaissance des jonctions intron-exon. Des mutations, et en particulier des délétions d'un ou de plusieurs nucléotides au sein de régions introniques poly-pyrimidiques, ont déjà été observées dans les cancers de l'appareil digestif présentant une instabilité microsatellitaire (Dorard et al, 2011). De plus, des modifications d'épissage dû à des mutations dans des répétitions mononucléotidiques ont également été trouvées dans les cancers colorectaux MSI et un type de leucémie dans les gènes MRE11 (Ham et al, 2006), comme observé ici dans nos échantillons de cancers gastriques et colorectaux. Il a déjà été démontré que la suppression de nucléotides d'une région microsatellite dans une répétition de T (17), quelques nucléotides en amont du site accepteur de l'intron 8 de HSP110, conduisait au saut d'exon 9, aboutissant à la synthèse d'une isoforme HSP110 tronquée.

Fait important, la fonction anti-apoptotique du chaperon de HSP110 était ainsi perdue (Dorard et al, 2011). Il a été conclu que la résistance à la chimiothérapie pouvait être supprimée en présence de l'isoforme HSP110 tronqué. Ces données concordaient avec le fait que les patients atteints d'un cancer gastrique MSI avaient un meilleur pronostic lorsqu'ils exprimaient un HSP110 mutant (Collura et al, 2014).

Ces résultats associés aux nôtres, suggéraient fortement qu'une délétion d'un seul nucléotide dans un segment intronique de mononucléotide poly-pyrimidique était suffisante pour déclencher un épissage anormal. Bien que le mécanisme précis impliqué n'ait pas été élucidé, il est tout à fait plausible que, soit la liaison d'une protéine à la répétition ait été entravée, soit sa capacité à interagir avec la sous-unité de spliceosome de U2 - responsable de la reconnaissance correcte de la frontière 3' intron-exon – ait été diminuée, ou que la région d'ARN ait été soumise à un repliement défectueux, ou même une combinaison des deux. Le fait que le phénotype moléculaire des mutations MSI ne se manifeste que dans le contexte spécifique de cancers MSI suggère que l'expression des gènes impliqués dans l'activité des spliceosomes pourrait être altérée dans ces cancers . Néanmoins, nos transfections de minigènes DNAJC18 et LMAN1 montrent une absence de lien direct avec le phénotype MSI, semblant indiquer que ce phénotype n'est pas absolument requis pour observer un effet de la mutation et donc que ce saut d'exon serait causé uniquement par la séquence en elle-même. On peut cependant remarquer, d'une part, que les cellules en culture ne sont pas assimilables à un cancer en place et, d'autre part, que la transfection de minigènes est un système artificiel non nécessairement superposable à un cancer.

Pris ensemble, nos résultats documentent davantage les liens fonctionnels entre les altérations des séquences de microsatellites associées au cancer et l'épissage aberrant des pré-ARNm. Il est important de noter que les séquences mononucléotidiques dans lesquelles ces mutations se produisent appartiennent aux signaux obligatoires agissant en cis requis pour l'épissage des introns. Proche de ce qu'il se passe dans d'autres types de maladies génétiques, les variations de séquence dans des régions poly-pyrimidiques sont bien connues pour contrôler le niveau d'inclusion ou d'exclusion d'un exon (Dujardin). Il reste encore à éclaircir de manière exact le mécanisme de survenue de ce saut d'exon et comment ces mutations spécifiques à un gène sont sélectionnées au cours de la survenue d'un cancer MSI.

# Résultats II

---



# Article

Identification de nouveaux biomarqueurs moléculaires dans le cancer gastrique. Gaël Quéré, Catherine Le Jossic-Corcoc, Brigitte Simon, Danielle Arzur, Jean-Philippe Metges, Cédric Le Maréchal\* and Laurent Corcos\*

## Objectifs de l'étude

Ces travaux de recherche ont pour but d'analyser des données RNAseq issues de la base de données du Cancer Genome Atlas (TCGA) afin d'identifier et caractériser l'expression de nouveaux marqueurs moléculaires du cancer gastrique. Cette étude fait partie de l'objectif global d'identification, sur le plan moléculaire, de nouvelles voies de signalisation qui seraient caractéristiques des adénocarcinomes gastriques.

## Matériels et méthodes

- ***Analyses des bases de données***

The Cancer Genome Atlas (TCGA) a catalogué les aberrations moléculaires existant dans 295 adénocarcinomes gastriques. Des analyses d'expression d'ARNs messagers et de micro-ARNs, de méthylation des promoteurs de la transcription et de variabilité du nombre de copies de gènes (CNV) ont été réalisées pour ces tumeurs et pour des tissus sains provenant des mêmes patients. Nous nous sommes intéressés plus précisément aux données de séquençage de l'ARN de tumeurs gastriques (Cancer Genome Atlas Research Network. (2014). Comprehensive molecular characterization of gastric adenocarcinoma. Nature, 513(7517), 202-209.). Le séquençage de l'ARN (RNAseq) est une technologie à haut débit permettant une caractérisation du transcriptome avec une très grande précision. Le RNAseq est capable d'identifier rapidement des transcrits, des isoformes, des nouveaux transcrits, des fusions de gènes et des ARN non codants, rares ou courants, parmi un large éventail d'échantillons (lignées cellulaires, patients...). Pour l'analyse du transcriptome, le consortium TCGA a utilisé une plateforme basée sur le système Illumina. Le séquençage et l'alignement des short-reads ont été réalisés dans 3 grands centres : Broad Institute, Baylor College of Medicine et

Washington University Sequencing Center. La base de donnée « NCBI dbGaP » est le référentiel des données de séquence générées par le TCGA.

Dans cette étude, 295 adénocarcinomes gastriques ont été analysés. Ils ont été subdivisés en quatre sous-types: les tumeurs positives au virus Epstein-Barr (N=26), qui présentent des mutations récurrentes de *PIK3CA*, l'hyperméthylation extrême de l'ADN et l'amplification de *JAK2*, *CD274* (également appelé PD-L1) et *PDCD1LG2* (également connu comme PD-L2); les tumeurs instables microsatellites (N=64), qui présentent des taux de mutation élevés, y compris des mutations de gènes codant des protéines de signalisation oncogènes pouvant être ciblées; les tumeurs génomiquement stables (N=58), qui sont enrichies pour la variante histologique diffuse et les mutations de *RHOA* ou des fusions impliquant des protéines activant la GTPase de la famille RHO; et les tumeurs présentant une instabilité chromosomique (N=147), qui présentent une aneuploïdie et une amplification focale marquées des récepteurs tyrosine kinases.

L'identification des gènes différentiellement exprimés a été effectuée après une analyse non appariée à deux classes en utilisant SAMseq avec un seuil FDR de 0,05 (Normalisation, Estimation de la dispersion (i.e. variabilité) et Tests statistiques). Les résultats ont été filtrés en éliminant les gènes dont l'expression médiane était inférieure à 5 RPKM. Les gènes pour lesquels la p-value était inférieure à 0.05 après un test de Wilcoxon ont été sélectionnés.

L'ensemble de ces données est disponible à l'adresse : [https://tcga-data.nci.nih.gov/docs/publications/stad\\_2014/](https://tcga-data.nci.nih.gov/docs/publications/stad_2014/)

- ***Extraction d'ARN et PCR quantitative.***

L'ARN des tissus tumoraux et des tissus gastriques sains ont été extraits avec le kit AllPrep DNA/RNA Mini (Qiagen) conformément aux instructions du fabricant. L'intégrité des ARNs a été évaluée à l'aide d'un bioanalyseur Agilent 2100 utilisant le kit RNA Nano Chips (Agilent) pour tous les tissus tumoraux primaires et les tissus gastriques sains.

Le niveau d'expression de l'ensemble des gènes a été analysée par PCR en temps réel (ABI 7000, Applied Biosystems, France). La PCR a été réalisée avec le Power SYBR-Green PCR

Master Mix, conformément aux instructions du fabricant (Applied Biosystems) à l'aide des amorces :

|          |                                     |  |
|----------|-------------------------------------|--|
| ADH1B:   | F: 5'-GAGATGCCTTCCCCTGTAGC -3',     | R: 5'- TGGCCCAGCATGTGTATGTT -3'        |
| AKR1C3:  | F: 5'-GGTTCCGCCATATAGATTCTGCTCA-3', | R: 5'- TGGCCTCCAGGTGGTACAGAGATC-3',    |
| ALDH3A1: | F: 5'-ACTCAGCAGGACGAGCTCTA -3',     | R: 5'- CAGGTCTGGCCACTGTTTCAT -3',      |
| CYP2C19: | F: 5'-CCAAAGGACCTTGACACAAC -3',     | R: 5'- GAGATGACGGGTCAGAAGAA -3',       |
| GGT6:    | F: 5'-CAGAGGCGCTGGTTCTAAA -3',      | R: 5'- GATTCTGGAGCTGCCTCACA -3',       |
| GSTA1:   | F: 5'-CCATGAAATGTGTGGGAGTG-3',      | R: 5'-GTCTCCTGGAGGTTTCTCTAAG-3',       |
| GPX3:    | F: 5'-ATTCTGGGCTTCCCTGCAA -3',      | R: 5'- AAAGTTCAGCGGATGTCGT -3',        |
| SOD2:    | F: 5'-AAGGGAGATGTTACAGCCAGATA-3',   | R: 5'- TCCAGAAAATGCTATGATTGATATGAC-3', |
| UGT1A1:  | F: 5'-CTTGCCTCAGAATTCCTTCAG-3',     | R: 5'- GCCTAGGGTAATCCTTCACA-3',        |

Toutes les conditions ont été normalisées par rapport au transcrit de contrôle GAPDH. Les résultats ont été analysés à l'aide de la méthode  $2^{-\Delta\Delta Ct}$  (Livak et Schmittgen, 2001).

#### ▪ ***Création d'un réseau d'interaction***

Afin d'obtenir des informations biologiques sur l'ensemble des gènes dérégulés dans les adénocarcinomes gastriques, nous avons recherché des associations directes (physiques) et indirectes (associées fonctionnellement sans interaction directe) entre ces gènes identifiés dans la base de données STRING (version 10.5).

STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) (Jensen et al., 2009) est une base de données d'interactions protéine-protéine prédites ou démontrées expérimentalement chez un certain nombre d'organismes dont l'Homme. Il peut s'agir d'interactions directes (physiques) ou indirectes (fonctionnelles). Elle est distribuée depuis 2000 par l'EMBL et est développée conjointement avec le SIB et l'université de Zurich.

Les données d'interactions sont entièrement précalculées à partir d'un grand nombre de sources de données :

- Importation des banques de données d'interactions protéiques directes IntAct (Kerrien et al., 2007), BioGRID (Breitkreutz et al., 2008), HPRD (Prasad et al., 2009), MINT (Chatr-aryamontri et al., 2007), DIP (Salwinski et al., 2004), GO (Ashburner et al., 2000), et BIND (Alfarano et al., 2005),

- Importation des banques de voies métaboliques KEGG (Kyoto Encyclopedia of Genes and Genomes) (Kanehisa et al., 2008), Reactome (Matthews et al., 2009), PID (Schaefer et al., 2009) et EcoCyc (Keseler et al., 2009),
- Extraction de données à partir de bases de connaissances PubMed, OMIM (Hamosh et al., 2005), FlyBase (Tweedie et al., 2009), et SGD (Hong et al., 2008) et recherche de cooccurrences de noms de gènes.

Les interactions importées sont ensuite complétées par des méthodes de prédiction basées sur le contexte génomique des gènes ou leur profil d'expression :

- Conservation de gènes dans un voisinage proche : une suite de gènes conservée, entre plusieurs génomes, à l'intérieur d'une distance chromosomique compatible avec un mécanisme de régulation commun peut indiquer une relation fonctionnelle.
- Fusion de gènes : la fusion de plusieurs gènes dans certains génomes indique potentiellement une étroite relation fonctionnelle entre ces gènes.
- Cooccurrence de gènes : les familles de gènes partageant un même profil phylogénétique de présence/absence dans plusieurs génomes sont considérées comme ayant une relation fonctionnelle.
- Coexpression : des associations fonctionnelles sont recherchées à l'intérieur des données d'expression de puces à ADN à l'aide du logiciel ArrayProspector (Jensen et al., 2004).

En bref, les interactions dans STRING proviennent de sources multiples (données expérimentales, bases de données sélectionnées, prédiction du contexte génomique, co-expression et extraction de texte automatisée). La version 10.5 actuelle de STRING contient environ 9.5 millions de protéines provenant de 2031 organismes vivants et référençant ainsi plus d'1.3 milliard d'interactions. Le niveau de confiance du bord, c'est-à-dire l'interaction entre 2 nœuds ou protéine / gène, est calculé dans un score combiné allant de 0 (aucune confiance) à 1 (très forte confiance). Dans le cadre de cette recherche, nous nous sommes placés à l'indice de confiance à 0.9.

▪ ***Etude des voies métaboliques : KEGG Pathway***

Nous avons ensuite intégré les données récupérées par l'analyse de réseau grâce à KEGG (Kyoto Encyclopedia of Genes and Genomes) (Kanehisa et al., 2008) qui est une collection de banques de données initiée en 1995 par le GenomeNet Database, Service du Kanehisa Laboratory à l'université de Kyoto dans le cadre du Human Genome Program, programme lancé par le Ministère de l'éducation, la culture, des sports, des sciences et des technologies japonaises. KEGG se définit comme une base de données orientée vers la biologie des systèmes permettant la compréhension des mécanismes et des concepts fonctionnels complexes de la cellule ou d'un organisme à partir de son génome. KEGG est constituée de quatre banques principales connectées entre elles et jouant le rôle de points d'entrée pouvant aboutir aux autres sous-banques qui les composent :

- PATHWAY est une base de connaissances de voies métaboliques dessinées manuellement et de voies non-métaboliques générées automatiquement,
- BRITe est l'ontologie de tous les concepts et connaissances présents dans KEGG,
- GENES est un catalogue de gènes de plusieurs génomes complets,
- LIGAND est un catalogue de substances chimiques et de réactions qui interviennent dans le domaine de la vie.

Chacune des banques de KEGG est conçue pour être représentée sous forme de graphes dont les entrées sont les nœuds et les relations biologiques les arêtes. PATHWAY est sans doute le point central de la banque KEGG.

# Résultats

## *Analyses des données de RNAseq*

Les résultats obtenus à partir de la base de données du Cancer Genome Atlas (TCGA) permettent de recueillir les données de gènes surexprimés et de gènes sous-exprimés en fonction de chaque sous classes d'adénocarcinomes gastriques :

- MSI vs tissus sains : 3618 gènes surexprimés / 832 gènes sous-exprimés
- EBV vs tissu sains : 2383 gènes surexprimés / 849 gènes sous-exprimés
- GS vs tissu sains : 1776 gènes surexprimés / 566 gènes sous-exprimés
- CIN vs tissu sains : 2856 gènes surexprimés / 790 gènes sous-exprimés

Ces données prétraitées, nous les avons compilées afin de sélectionner uniquement les gènes qui étaient dérégulés, à la hausse et à la baisse, au sein de ces 4 catégories. Nous avons ainsi obtenu 496 gènes surexprimés et 123 gènes sous-exprimés au sein des adénocarcinomes gastriques vs les tissus sains.

## Analyses du réseau d'interaction

Cette compilation de gènes surexprimés et sous-exprimés, nous l'avons intégrée afin d'obtenir un réseau in silico :

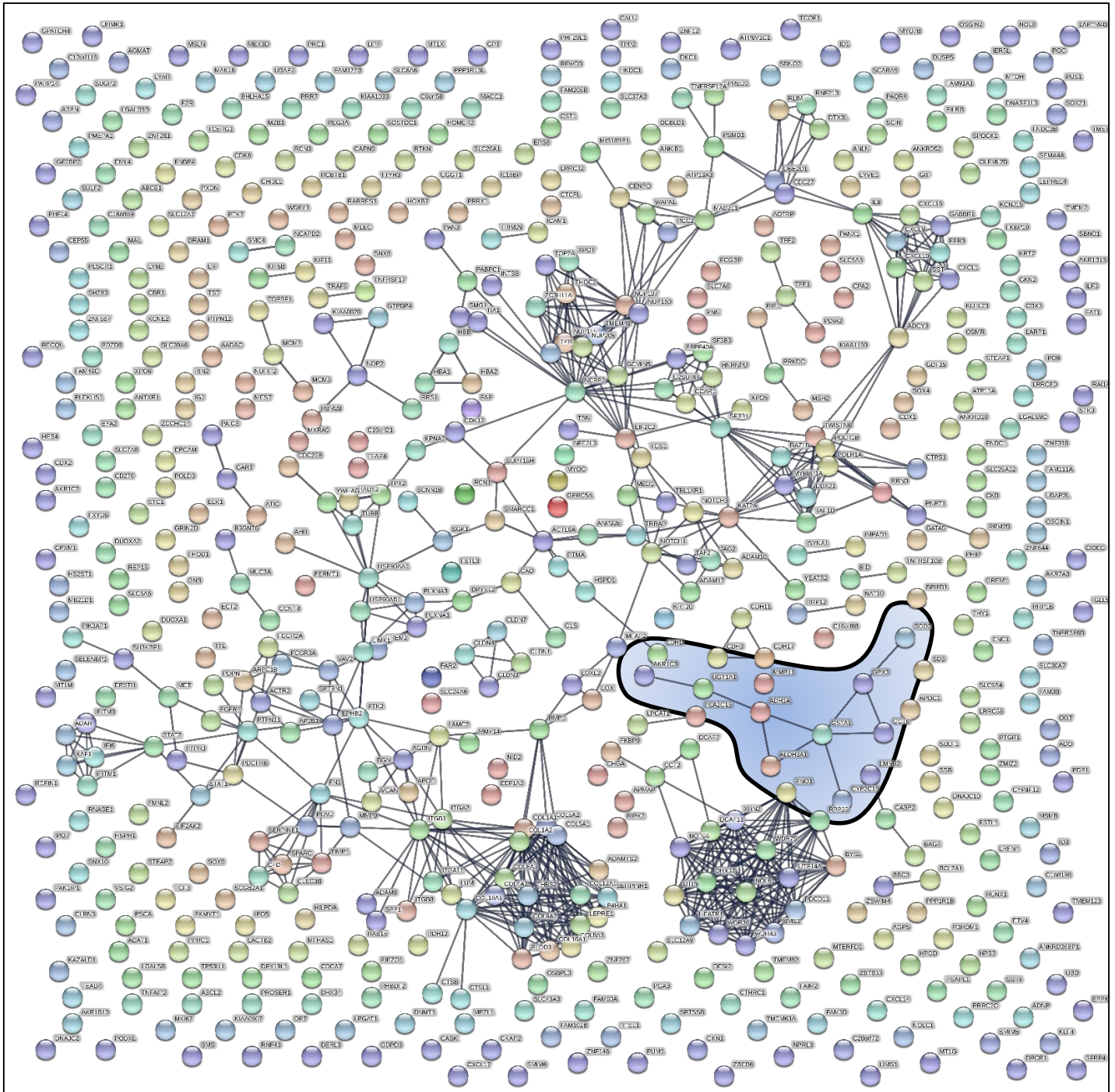


Figure 23 : Réseau reliant entre eux les 496 gènes surexprimés et les 123 gènes sous-exprimés dans les adénocarcinomes gastriques. Les traits entre les protéines symbolisent l'indice de confiance de la probabilité d'interaction. (Seuil de confiance  $\geq 0,900$ ). Chaque noeud du graphe représente une protéine et chaque arête une interaction. Réseau d'interaction basé sur 619 gènes dérégulés dans les tumeurs gastriques, 687 liens identifiés.

Notre attention a été attirée par le comportement d'un petit sous-réseau, baptisé réseau « Papillon », comportant 9 gènes impliqués, collégalement, dans les réactions de détoxification, et majoritairement sous-exprimés dans les échantillons de cancers gastriques (à l'exception de la SOD2) :

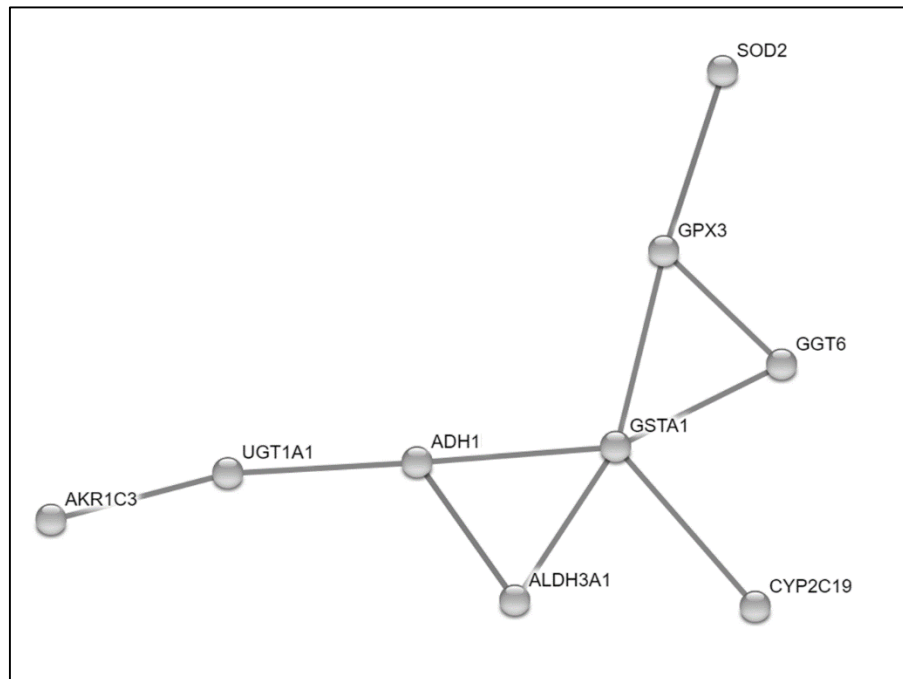


Figure 24 : Cluster de gènes interagissant ensemble issu du réseau d'interaction de la base de données STRING : 9 gènes : 10 liens identifiés.

Au sein de ce réseau d'interaction, 9 gènes ont été identifiés parmi lesquels on trouve :

- *SOD2* : une superoxyde dismutase qui est une métalloprotéine qui catalyse la dismutation de l' $O_2^-$  en  $H_2O_2$ . Cette enzyme a une fonction importante dans le système de défense contre les radicaux libres.
- *GPX3* : une glutathion peroxydase qui métabolise une grande variété d'hydroperoxydes (R-OOH) en plus de l' $H_2O_2$ , bien que celui-ci reste préférentiel.
- *GGT6* : une gamma-glutamyl transférase qui est une enzyme membranaire dimérique, amphiphile, fortement glycosylée (Meister et al, 1981). Il n'existe que peu d'information sur la protéine codée par ce gène. Les GGT sont conservées dans l'évolution et son rôle serait similaire à celui de *GGT1* et *GGT5* (synthèse et homéostasie du glutathion) même si son expression est plus faible.



- *GSTA1* : une glutathione S-transferase qui a pour rôle la détoxification de composés carcinogènes, médicaments, toxines environnementales, ou encore produits du stress oxydatif.
- *CYP2C19* : membre de la famille des cytochromes P450, la protéine CYP2C19 est impliquée dans le métabolisme des xénobiotiques et de nombreux inhibiteurs de la pompe à protons.
- *ADH1B* : l'Alcool déshydrogénase 1B intervient dans le métabolisme de l'éthanol, du rétinol ou encore des produits de la peroxydation lipidique.
- *ALDH3A1* : l'Aldéhyde Déshydrogénase 3 Family Member A1 a pour rôle, la détoxification de l'acétaldéhyde, de corticostéroïdes et de produits de la peroxydation lipidique.
- *UGT1A1* : l'UDP-glycosyltransferase 1 polypeptide A1 intervient dans le métabolisme de stéroïdes ou encore de la bilirubine.
- *AKR1C3* : membre de la famille des aldo-keto reductases, la protéine AKR1C3 possède une activité Kétostéroïde réductase dans le métabolisme des stéroïdes, concernant notamment les œstrogènes et la progestérone.

Nous avons mesuré l'expression des gènes composant ce réseau par PCR quantitative afin de déterminer si nos échantillons étaient représentatifs des données issues de TCGA.

| Gènes          | Niveau d'expression des gènes                          |              |               |               |  |
|----------------|--|--------------|---------------|---------------|--|
|                | Base de données du Cancer Genome Atlas Database (TCGA) |              |               |               | Cohortes de patients finistériens (N=32) |
|                | Normal vs CIN  | Normal vs GS | Normal vs EBV | Normal vs MSI | Normaux vs Tumeurs                       |
| <i>ADH1B</i>   | -5,3   | N.D.         | -7,1          | -6,5          | <b>-3,5</b>                              |
| <i>ALDH3A1</i> | -14,9  | -13,2        | -6,1          | -16,9         | <b>-4,6</b>                              |
| <i>CYP2C19</i> | -5,0   | -5,3         | -2,8          | -6,1          | <b>-2,6</b>                              |
| <i>GGT6</i>    | -4,9   | -3,2         | -1,9          | -3,8          | <b>N.D.</b>                              |
| <i>GPX3</i>    | -5,0   | -3,0         | -6,3          | -7,4          | <b>-3,5</b>                              |
| <i>GSTA1</i>   | -6,0   | -6,2         | -7,1          | -20,8         | <b>N.D.</b>                              |
| <i>UGT1A1</i>  | -2,1   | -2,7         | N.D.          | -4,3          | <b>-2,3</b>                              |
| <i>AKR1C3</i>  | -3   | -3,9         | -7,5          | -5,6          | <b>N.D.</b>                              |
| <i>SOD2</i>    | 1,9  | 1,9          | 2,1           | 2,8           | <b>N.D.</b>                              |

Tableau 4 : Comparaison du niveau d'expression de 9 gènes, sélectionnés à partir des réseaux d'interactions créés dans STRING v10.0, à partir de la base de données transcriptomiques TCGA et de notre cohorte de patients. Méthodes : base de données TCGA : RNAseq; Cohorte des patients: RT-PCR quantitative. N.D. signifie non détectée.

▪ **Etude des voies métaboliques : KEGG Pathway**

Les données de ce réseau ont été intégrées afin d'étudier dans quelles voies métaboliques ces gènes intervenaient de façon commune. Ainsi, nous avons pu classer cette liste de gènes au sein de 4 sous catégories appartenant au métabolisme des dérivés réactifs de l'oxygène, au métabolisme du glutathion, au métabolisme des xénobiotiques et à la voie de biosynthèse des hormones stéroïdes.

De façon plus précise, voici la liste des voies métaboliques concernées par chacun de ces gènes dérégulés :

| <b>Gènes</b> | <b>Voie métabolique d'appartenance - KEGG pathway</b> | <b>code KEGG</b> |
|--------------|---|------------------|
| <i>SOD2</i>  | FoxO signaling pathway                                | hsa04068         |
|              | Peroxisome  | hsa04146         |
|              | Longevity regulating pathway                          | hsa04211         |
|              | Longevity regulating pathway - multiple species       | hsa04213         |
|              | Huntington disease                                    | hsa05016         |
| <i>GPX 3</i> | Glutathione metabolism                                | hsa00480         |
|              | Arachidonic acid metabolism                           | hsa00590         |
|              | Thyroid hormone synthesis                             | hsa04918         |
| <i>GGT6</i>  | Glutathione metabolism                                | hsa00480         |

|                |  |          |
|----------------|--|----------|
|                | Taurine and hypotaurine metabolism           | hsa00430 |
|                | Metabolic pathways                           | hsa01100 |
| <i>GSTA1</i>   | Glutathione metabolism                       | hsa00480 |
|                | Drug metabolism - cytochrome P450            | hsa00982 |
|                | Metabolism of xenobiotics by cytochrome P450 | hsa00980 |
|                | Drug metabolism - other enzymes              | hsa00983 |
|                | Platinum drug resistance                     | hsa01524 |
|                | Pathways in cancer                           | hsa05200 |
|                | Chemical carcinogenesis                      | hsa05204 |
|                | Hepatocellular carcinoma                     | hsa05225 |
|                | Fluid shear stress and atherosclerosis       | hsa05418 |
| <i>ALDH3A1</i> | Metabolism of xenobiotics by cytochrome P450 | hsa00980 |
|                | Drug metabolism - cytochrome P450            | hsa00982 |
|                | Glycolysis / Gluconeogenesis                 | hsa00010 |
|                | Histidine metabolism                         | hsa00340 |

|                |  |          |
|----------------|--|----------|
|                | Tyrosine metabolism                          | hsa00350 |
|                | Phenylalanine metabolism                     | hsa00360 |
|                | beta-Alanine metabolism                      | hsa00410 |
|                | Metabolic pathways                           | hsa01100 |
|                | Chemical carcinogenesis                      | hsa05204 |
| <i>ADH1B</i>   | Drug metabolism - cytochrome P450            | hsa00982 |
|                | Metabolism of xenobiotics by cytochrome P450 | hsa00980 |
|                | Glycolysis / Gluconeogenesis                 | hsa00010 |
|                | Fatty acid degradation                       | hsa00071 |
|                | Tyrosine metabolism                          | hsa00350 |
|                | Retinol metabolism                           | hsa00830 |
|                | Metabolic pathways                           | hsa01100 |
|                | Chemical carcinogenesis                      | hsa05204 |
| <i>CYP2C19</i> | Drug metabolism - cytochrome P450            | hsa00982 |
|                | Arachidonic acid metabolism                  | hsa00590 |

|               |  |          |
|---------------|--|----------|
|               | Linoleic acid metabolism                     | hsa00591 |
|               | Metabolic pathways                           | hsa01100 |
|               | Serotonergic synapse                         | hsa04726 |
|               | Chemical carcinogenesis                      | hsa05204 |
| <i>UGT1A1</i> | Metabolism of xenobiotics by cytochrome P450 | hsa00980 |
|               | Drug metabolism - cytochrome P450            | hsa00982 |
|               | Drug metabolism - other enzymes              | hsa00983 |
|               | Pentose and glucuronate interconversions     | hsa00040 |
|               | Ascorbate and aldarate metabolism            | hsa00053 |
|               | Steroid hormone biosynthesis                 | hsa00140 |
|               | Retinol metabolism                           | hsa00830 |
|               | Porphyrin and chlorophyll metabolism         | hsa00860 |
|               | Metabolic pathways                           | hsa01100 |
|               | Chemical carcinogenesis                      | hsa05204 |
| <i>AKR1C3</i> | Steroid hormone biosynthesis                 | hsa00140 |

|  |                             |          |
|--|-----------------------------|----------|
|  | Arachidonic acid metabolism | hsa00590 |
|  | Folate biosynthesis         | hsa00790 |
|  | Metabolic pathways          | hsa01100 |
|  | Ovarian steroidogenesis     | hsa04913 |

Tableau 5 : Liste des voies métaboliques selon Kegg Pathway correspondant à la liste de gènes sélectionnés se basant sur réseau d'interactions STRING.

L'identification des voies métaboliques dans lesquelles interviennent les gènes issus du réseau « papillon » permet ainsi de diviser notre réseau d'interaction en 4 sous-groupes :

- Métabolisme des dérivés réactifs de l'oxygène (ROS),
- Métabolisme du glutathion,
- Métabolisme des xénobiotiques,
- Voie de biosynthèse des hormones stéroïdes.

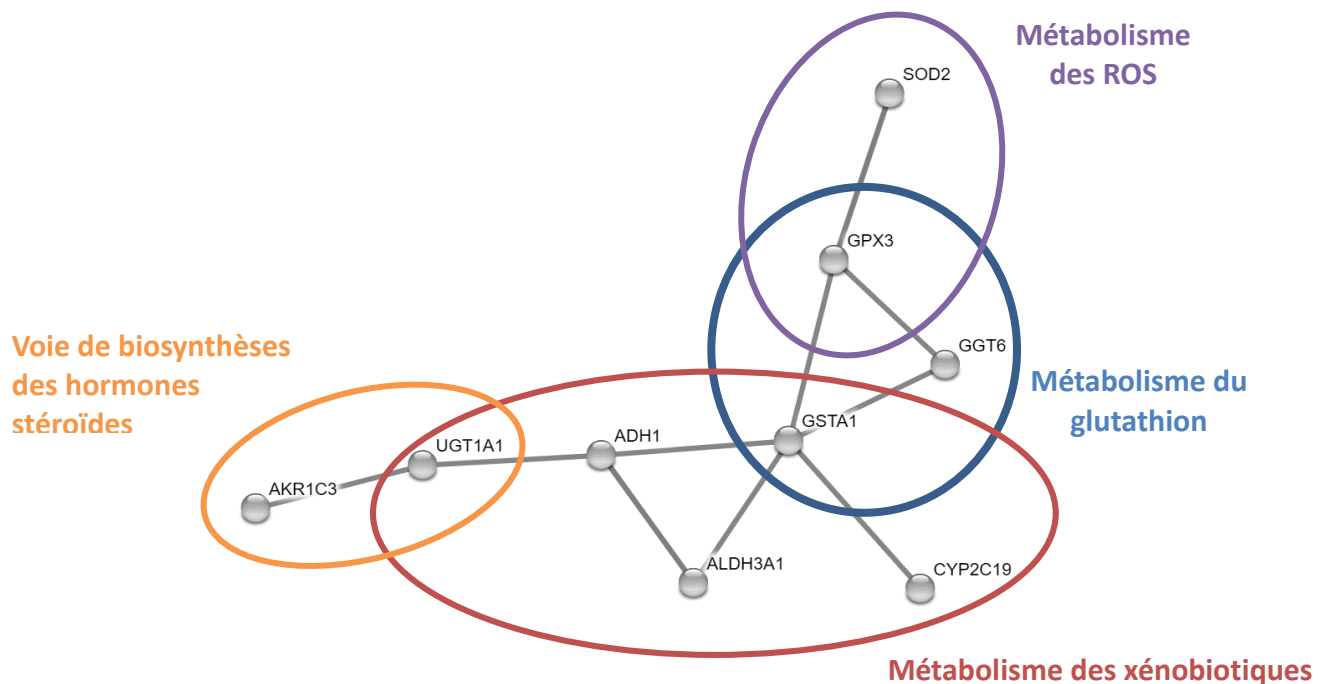


Figure 25 : visualisation du réseau « papillon » et voies métaboliques communes (seuil de confiance  $\geq 0,900$ ). Réseau d'interactions basé sur 9 gènes dérégulés dans les tumeurs gastriques, 10 liens identifiés.

Si, en apparence, une réaction catalysée par une enzyme peut sembler vague, celle-ci prend tout son sens dans le contexte des voies métaboliques qui ont été extraites de la littérature pour être décrites par les banques KEGG PATHWAY (Kanehisa et al., 2008), MetaCyc (Caspi et al., 2008) ou Reactome (Matthews et al., 2009).

En effet, les numéros EC associés à chaque description d'enzyme (cf. Tableau 2) sont souvent utilisés afin de cartographier les gènes d'un organisme sur l'ensemble des voies métaboliques connues afin de mieux comprendre sa physiologie.



C'est ainsi que nous avons sélectionné plusieurs réseaux métaboliques qu'il peut s'avérer intéressant d'étudier :

- **Sous-groupe n°1 : Métabolisme du glutathion**

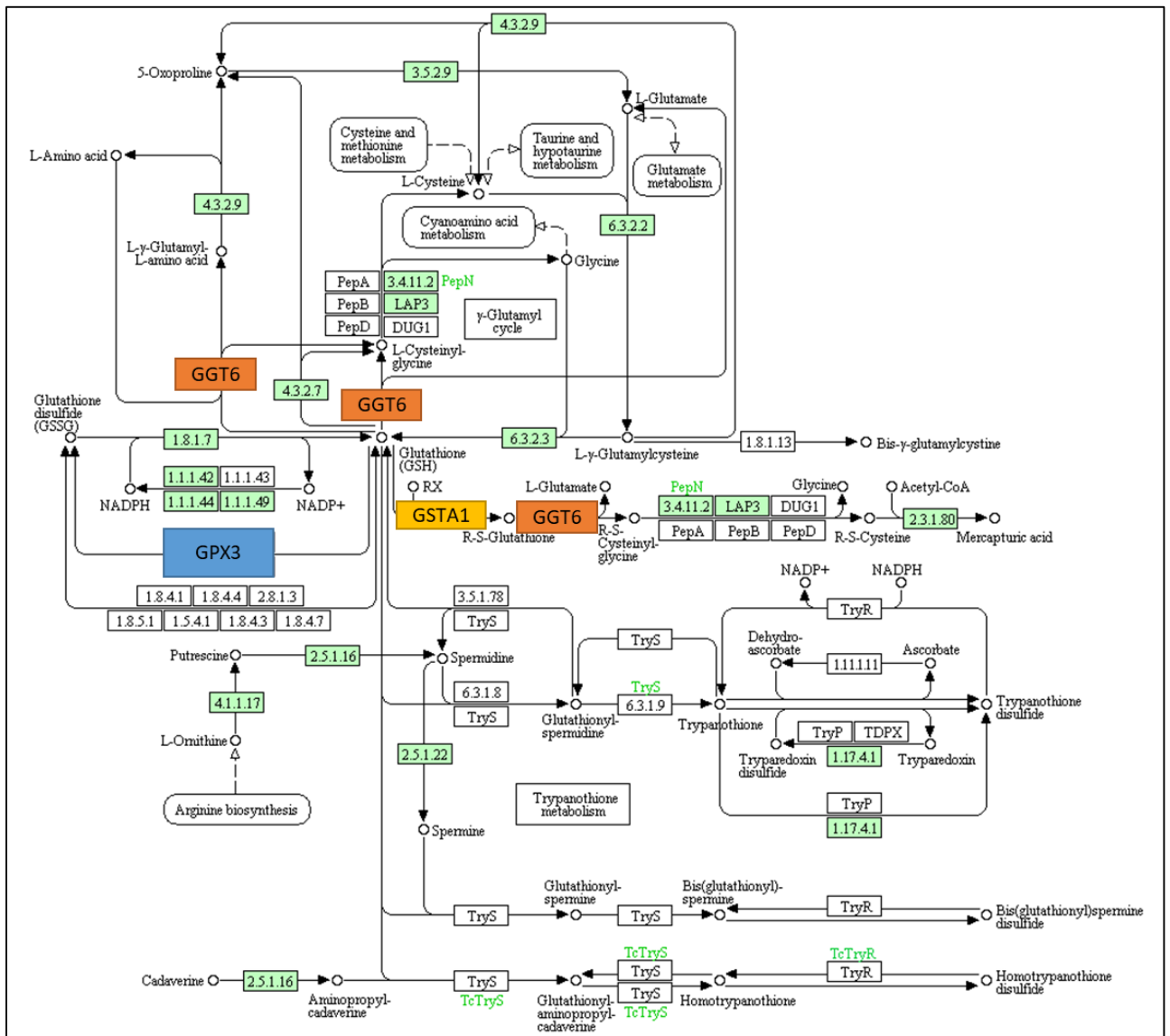


Figure 26 : Cycle du glutathion.

Le glutathion (L- $\gamma$ -glutamyl-L-cysteinylglycine ; GSH) est un tripeptide formé par l'association de L-glutamate, de L-cystéine, et de L-glycine. Il contient un groupement thiol (-SH) au niveau du résidu cystéine qui possède un fort pouvoir réducteur. Les deux tiers du glutathion libre sont sous la forme réduite GSH (Meister, 1996). Le glutathion peut réduire les ponts disulfures non enzymatiques. L' $H_2O_2$  produit lors d'un stress oxydant peut-être réduit par le glutathion pour produire de l'eau et de l'oxygène.

Dans le cadre de tumeurs gastriques, les gènes *GSTA1*, *GGT6* et *GPX3* sont sous exprimés. Cette sous-expression a potentiellement pour conséquence une diminution du taux de glutathion, et donc de sa capacité à réduire l' $\text{H}_2\text{O}_2$  où prendre en charge les molécules excrétables.

- **Sous-groupe n°2 : Métabolisme des ROS**

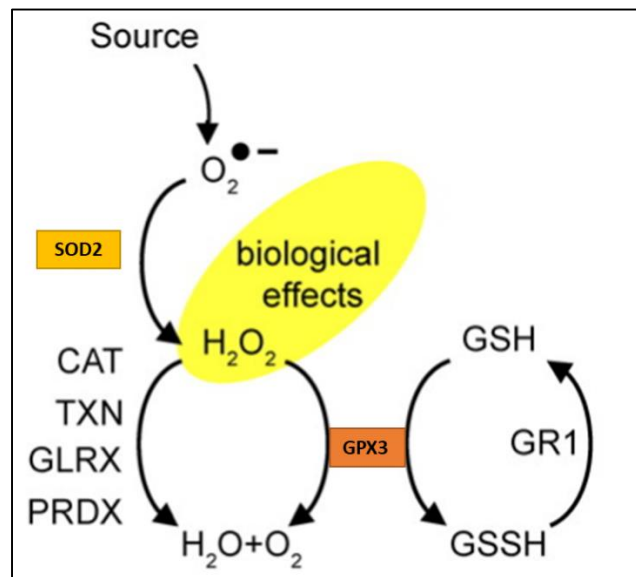


Figure 27 : Cycle de l'élimination des espèces réactives dérivées de l'oxygène.

Dans les conditions physiologiques, 90% des anions superoxydes sont produits au niveau de la chaîne respiratoire mitochondriale. La détoxification de l' $\text{O}_2^-$  est majoritairement effectuée par la SOD2 à manganèse (MnSOD). D'après les données TCGA, on observe dans les cancers gastriques une augmentation de l'expression de la *SOD2* couplée à une diminution de l'expression de *GPX3* qui a pour rôle de le transformer en  $\text{H}_2\text{O} + \text{O}_2$ .

- **Sous-groupe n°3 : Métabolisme des xénobiotiques**

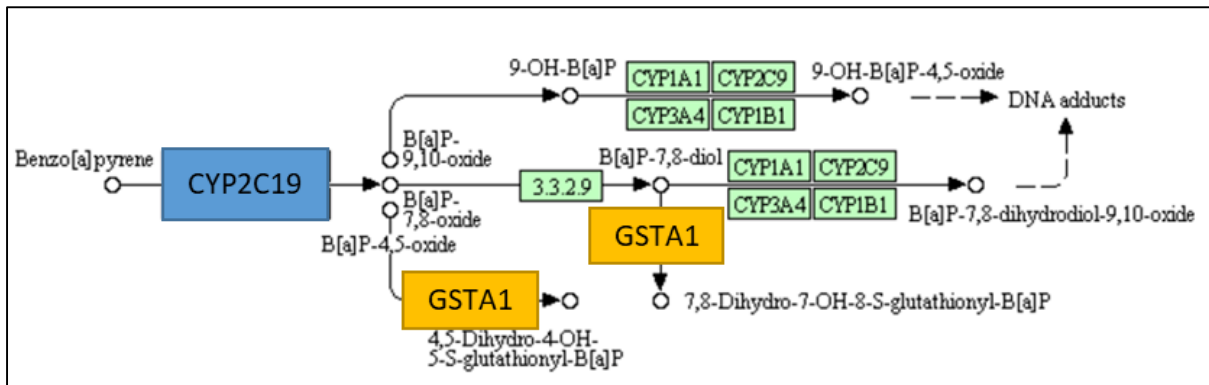


Figure 28 : Cycle de métabolisation du Benzo[a]pyrène

Le Benzo[a]pyrène est un hydrocarbure aromatique polycyclique. Il a été classé cancérigène de catégorie 2 au niveau de l'Union européenne, le CIRC l'a réévalué et introduit dans le groupe 1 des agents cancérigènes pour l'homme. CYP2C19 participe à sa prise en charge au côté de CYP1A1, CYP1B1, CYP2C8, CYP2C9, CYP2C19 dans l'organisme. GSTA1 par la voie du glutathion intervient ensuite dans son excrétion.

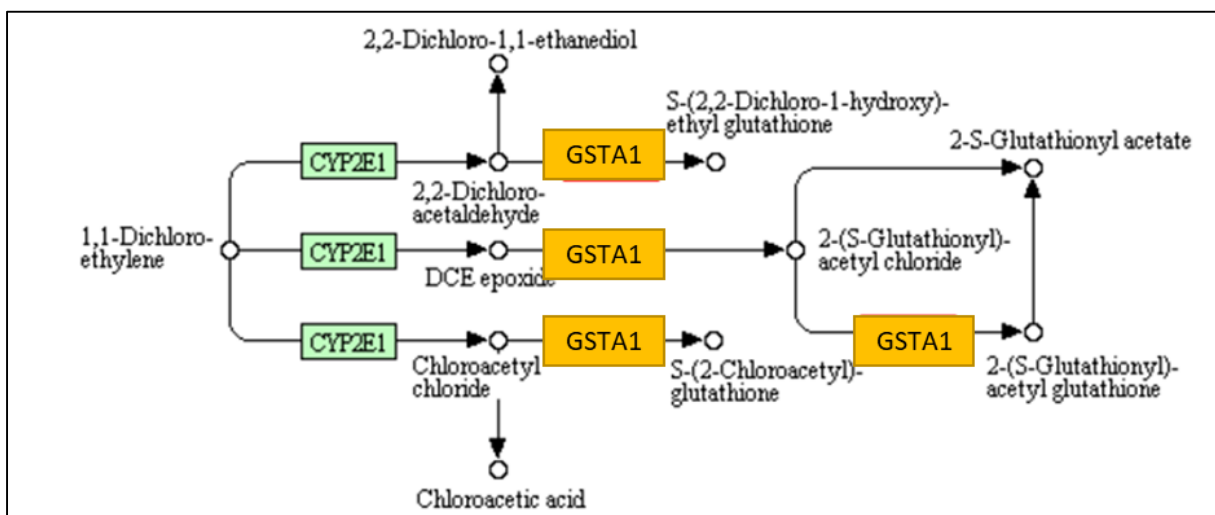


Figure 29 : Cycle de métabolisation du 1,1-Dichloroéthylène

Le dichloroéthylène est retrouvé comme solvant de nombreux produits de synthèse (résines, parfums, colorants, laques, phénols...). Il est également utilisé pour l'extraction à froid de produits tels que la caféine ou le caoutchouc naturel. Selon l'INRS, aucune étude documentée d'intoxication chez l'homme n'est disponible, mais son utilisation ancienne comme anesthésique laisse penser qu'il serait à l'origine de troubles de la conscience. Son excrétion est également effectuée par les GST dont la GSTA1.

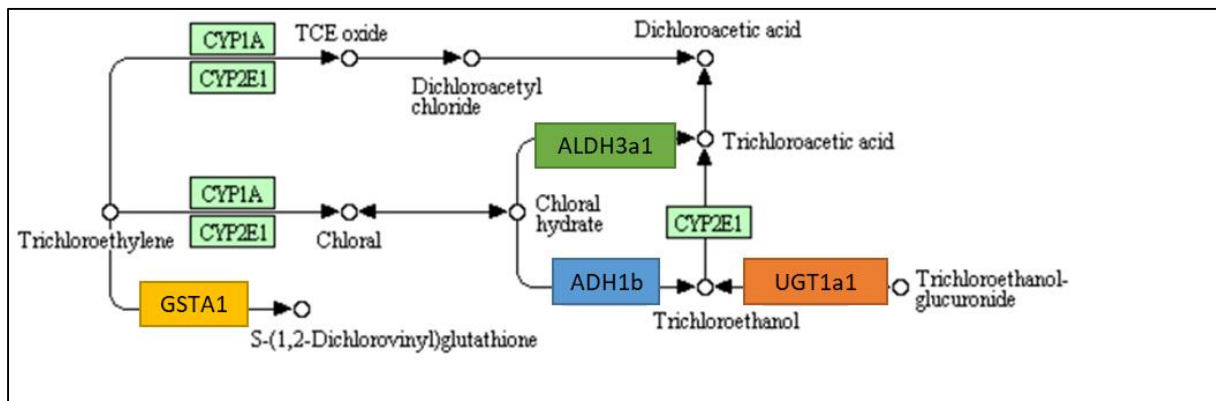


Figure 30 : Cycle de métabolisation du trichloréthylène

Le trichloroéthylène est un solvant chloré utilisé principalement pour le dégraissage et le nettoyage des métaux. Il est aussi utilisé dans l'industrie du caoutchouc, des produits d'entretien, des peintures et vernis. Il est classé cancérigène avéré pour l'homme (groupe 1 du CIRC) pour le cancer du rein. Les gènes *GSTA1*, *ALDH3A1*, *ADH1B* et *UGT1A1* sous exprimés dans les cancers gastriques interviennent dans son élimination.

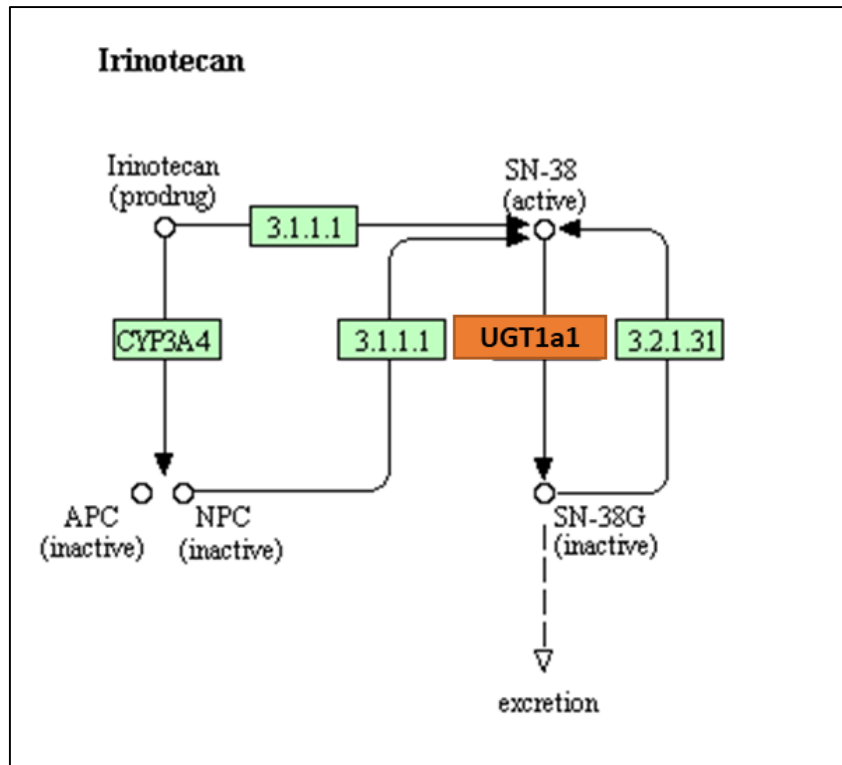


Figure 31 : Cycle de métabolisation de l'irinotecan

L'irinotécan est un dérivé hémisynthétique de la camptothécine. Il s'agit d'un agent antinéoplasique qui agit comme inhibiteur spécifique de l'ADN topo-isomérase I. L'irinotécan est métabolisé par la carboxylestérase en un métabolite actif, le SN-38. L'inhibition de l'ADN topoisomérase I par l'irinotécan ou le SN-38 induit des lésions simple-brin de l'ADN et sont responsables de l'activité cytotoxique. Son élimination est effectuée par voie biliaire et urinaire après sa métabolisation par l'UGT1A1.

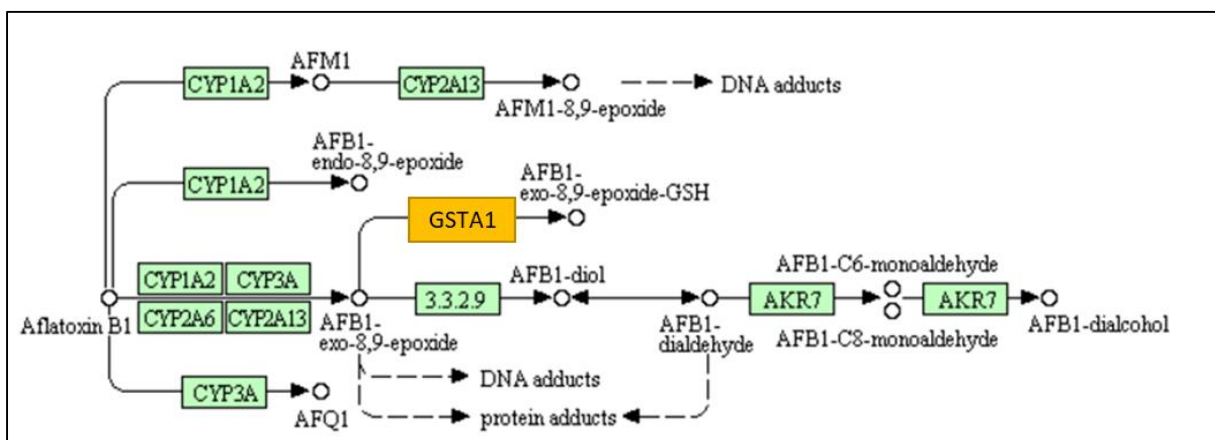


Figure 32 : Cycle de métabolisation de l'aflatoxine B1.

Les aflatoxines sont des toxines (mycotoxines) produites par des moisissures appartenant au genre *Aspergillus* (section *Flavi*), champignons pouvant se développer et produire les toxines plus facilement dans les régions chaudes et humides des pays d'Afrique, d'Asie et d'Inde. Le niveau d'exposition aux toxines peut parfois être élevé pour les régimes alimentaires traditionnels à base de maïs et d'arachide. L'ingestion de denrées alimentaires contaminées par les aflatoxines est la voie d'exposition majeure chez l'homme. Quatre aflatoxines (B1, B2, G1, G2) sont produites dans la nature. L'aflatoxine B1 est la plus fréquente et la plus toxique. Elle possède des propriétés génotoxiques et carcinogènes.



L'ensemble des hormones stéroïdiennes provient du cholestérol. Ce sont les cytochromes P450sc (side chain cleavage) et les cytochromes P450 associés à l'aldo-keto reductase, dans la matrice mitochondriale, qui font entrer le cholestérol dans la chaîne de biosynthèse des stéroïdes. En fonction de la glande endocrine dans laquelle elle se produit, cette biosynthèse peut aboutir à la formation de :

- La testostérone, dans le testicule,
- La progestérone et d'œstradiol, dans l'ovaire,
- Les glucocorticoïdes comme le cortisol dans la glande médullo-surrénale,
- Les minéralocorticoïdes, comme l'aldostérone, dans la glande cortico-surrénale.

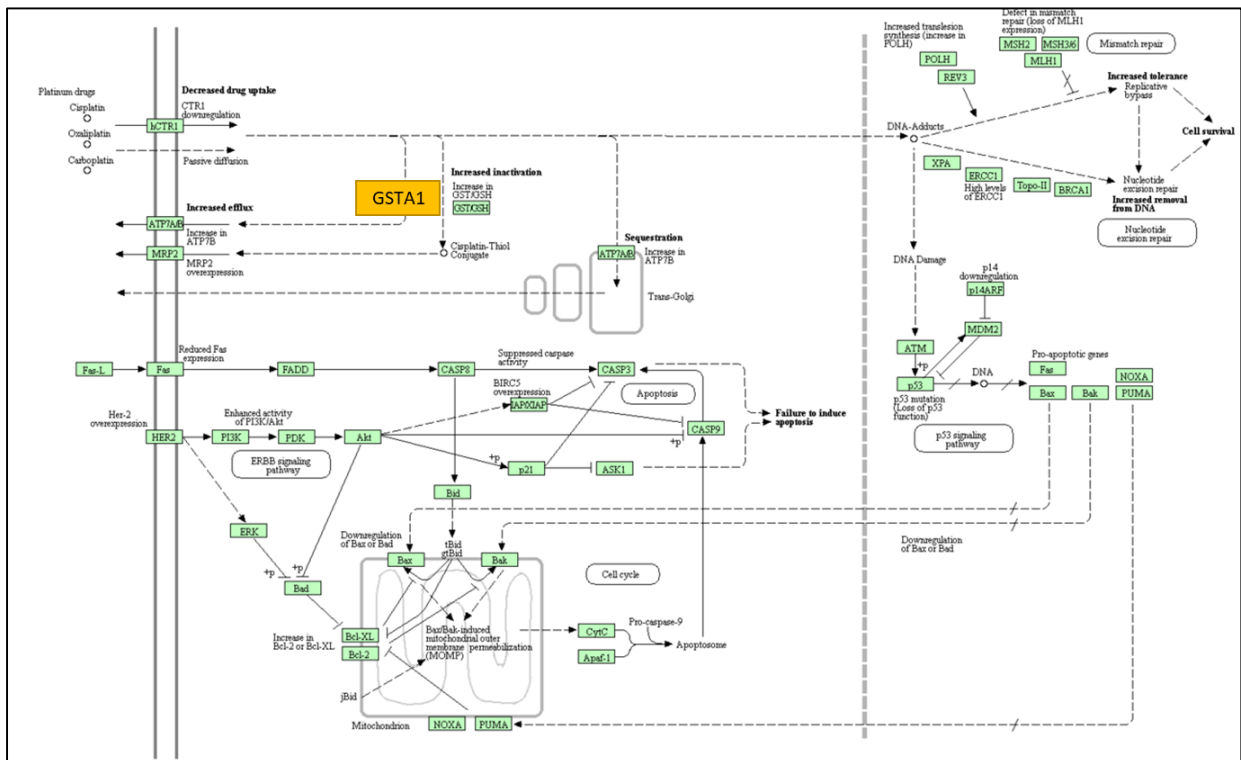


Figure 34 : Voie de résistance aux traitements de type sels de platine.

On observe sur ce réseau le mécanisme d'action des traitements de type sels de platine sur le métabolisme d'une cellule potentiellement tumorale. GSTA1 intervient notamment dans la voie d'élimination de ces sels de platines, qui se lie en particulier à l'ADN (mais aussi à l'ARN et aux protéines), afin de créer des adduits (complexes ADN-Platine) responsable de la cytotoxicité et qui aboutit à l'apoptose de la cellule.



# Discussion

---

Dans cette étude, nous avons découvert un sous ensemble de gènes impliqués collégalement dans diverses réactions de détoxification. La majorité de ces gènes sont sous-exprimés dans les échantillons de cancers gastriques analysés à partir des bases de données TCGA. Ces gènes sont également sous exprimés dans les tumeurs étudiées issues de la base de données clinico-biologiques du CHRU de Brest. Tous les fragments de tissu utilisés pour ces analyses ont été extraits de pièces chirurgicales de patients n'ayant pas subi de traitement anticancéreux au préalable, ce qui s'avère primordial pour étudier les origines de la dérégulation de ce réseau.

La position centrale de GSTA1 dans notre réseau peut être une indication de l'importance de son rôle dans notre réseau. En effet, les GST représentent une famille d'enzymes multigéniques et multifonctionnelles. Ces enzymes utilisent comme substrats une large gamme de composés électrophiles pour les conjuguer au glutathion de façon à diminuer leur réactivité vis-à-vis des macromolécules intracellulaires. Cette conjugaison par le glutathion permet l'excrétion de ces composés, potentiellement toxiques, hors de la cellule où ils seront éliminés. Les GST catalysent l'attaque nucléophile du glutathion sur ces molécules électrophiles. Cette attaque nécessite l'activation de la fonction thiol du glutathion qui se retrouve déprotonée via une interaction avec un résidu tyrosyl du site actif qui joue le rôle de catalyseur base (Rooseboom et al., 2004). Chez l'homme, ces enzymes sont abondantes dans le foie qu'elles protègent contre les effets de divers xénobiotiques mais elles sont également présentes dans d'autres organes comme les poumons, l'estomac ou le colon. Pour évaluer la fonction de GSTA1, des études ont été mené notamment dans des cellules cancéreuses du poumon. Il a été montré une surexpression de GSTA1 dans la lignée cellulaire d'adénocarcinome du poumon A549. Cette surexpression de GSTA1 favoriserait la prolifération cellulaire et la diminution de l'expression de GSTA1 par l'intermédiaire d'un siRNA dans ces cellules induit de manière significative l'apoptose cellulaire *in vitro* (Liu et al, 2018). Au contraire, dans des cellules Caco-2, l'étude d'Adnan et al. en 2012 a montré que la suppression de GSTA1 par l'intermédiaire d'un siRNA a significativement augmenté la prolifération cellulaire et n'a pas modifié le taux d'apoptose. L'hétérogénéité de ces résultats peut être dû à un polymorphisme de GSTA1 dans différents types de cellules cancéreuses. Il a par exemple été montré que les polymorphismes de GSTM1 et GSTT1 étaient associés au risque de cancer du poumon : Liu et al en 2014 ont démontré qu'une GSTM1 non fonctionnelle présentait une association significative avec le carcinome épidermoïde (SC), l'adénocarcinome

(AC) et le carcinome pulmonaire à petites cellules, et qu'une GSTT1 non fonctionnelle présentait une association significative dans SC et AC. Plusieurs études ont également étudié l'association entre les polymorphismes de GSTA1 et les cancers colorectaux (Pettersson et Mannervik, 2001), de la prostate (Komiya et al 2005), du sein et de la vessie (Matic et al, 2013). Deng et al, en 2015, ont rapporté que le génotype GSTA1 BB était associé à un risque de cancer accru, en particulier pour le cancer colorectal chez les Caucasiens.

Les présents résultats ainsi que la position centrale dans notre réseau d'interaction suggèrent que le rôle de GSTA1 est important dans le développement des cancers gastriques.

Au sein de notre réseau, la protéine CYP2C19 possède un lien direct avec GSTA1. CYP2C19 fait partie des monooxygénases P450, autrement appelées cytochromes P450, qui constituent une vaste superfamille d'hémoprotéines utilisant l'oxygène moléculaire pour catalyser différentes réactions telles que les époxydations, les N-désalkylations, les S-oxydations et les hydroxylations (Chang et Kam, 1999). Il existe deux grandes familles de cytochromes P450 en fonction de leur localisation subcellulaire : 1) les enzymes microsomales, qui sont des protéines membranaires appartenant à la membrane du réticulum endoplasmique et dont le rôle majeur est la détoxification des xénobiotiques et le catabolisme de certains composés endogènes et ; 2) les enzymes mitochondriales, qui sont intégrées dans la membrane interne des mitochondries. Ces dernières sont impliquées dans la synthèse des hormones stéroïdiennes. Les cytochromes P450 ont également d'autres rôles. Par exemple, ils interviennent dans le processus de peroxydation des lipides, en générant des ROS. Ces entités produisent elles-mêmes des aldéhydes comme le 4-HNE (Bardag-Gorce et al., 2005). On pourra noter que CYP2C19 est une enzyme majoritairement hépatique et qu'un certain nombre de facteurs cliniques peuvent influencer sur l'activité des enzymes hépatiques, notamment une prise de médicaments associés à ses activités enzymatiques et des maladies diverses du foie. Dans notre étude, nous regardons son expression au niveau de l'estomac. Le cytochrome P450 2C19 (CYP2C19) joue un rôle important dans le métabolisme des xénobiotiques et des médicaments tels que L'oméprazole, le lansoprazole, le rabéprazole, le diazépam, le propranolol et le S-méphénytoïne.

Il contribue aussi au catabolisme des substrats endogènes tels que l'estradiol. La variabilité génétique influe sur l'expression et l'activité du CYP2C19 et peut donc influencer sur le catabolisme des œstrogènes.

Ce qui a principalement été rapporté dans la littérature est son polymorphisme qui pourrait être un facteur associé aux différences interindividuelles de sensibilité à certaines formes de cancers, comme observé avec d'autres CYP dans la population japonaise pour différents cancers dont l'estomac (Suzuki et al, 2004 ; Gao et al, 2002 ; Cai et al, 2001 ; Yokose et al, 1998 ; Tsuneoka et al, 1996). En effet, environ 3% des Caucasiens, mais jusqu'à 30% des Asiatiques expriment le phénotype CYP2C19-PM (poor metabolisers) (Xie et al, 1999). En 2008, Helsby et al ont montré que l'activité de l'enzyme CYP2C19 était diminuée chez les patients caucasiens atteints d'un cancer avancé du sein, du poumon, du rein, de l'ovaire et du mélanome.

On peut ainsi supposer que la diminution de l'expression de *CYP2C19* a un potentiel rôle dans le développement et/ou la progression de plusieurs cancers dont les cancers gastriques sans pouvoir conclure sur son rôle exact, même si nous pouvons supposer que CYP2C19 peut avoir un impact significatif sur l'efficacité et la toxicité des agents chimiothérapeutiques qui sont des substrats pour cette enzyme.

GPx3 possède également un lien direct avec GTS1. Ce gène code un enzyme faisant partie des glutathion peroxydases. Les glutathion peroxydases métabolisent une grande variété d'hydroperoxydes (R-OOH) en plus de l' $H_2O_2$ . Néanmoins pour GPx1, GPx2 et GPx3 le substrat  $H_2O_2$  reste préférentiel. GPx3 est principalement extracellulaire et est retrouvée majoritairement dans le plasma (Takahashi et al., 1987) et au niveau des surfaces muqueuses. Son activité enzymatique permet de réduire le peroxyde d'hydrogène mais également les acides gras peroxydés avant de pouvoir pénétrer dans les cellules (Brigelius-Flohe et Kipp, 2009). Cette caractéristique unique fait de GPx3 un composant extrêmement important non seulement dans la famille GPx, mais également dans les systèmes antioxydants cellulaires, offrant la première ligne de défense contre l'activité des ROS néfastes. Le rôle de GPx3 dans la tumorigenèse humaine a été rapporté dans les carcinomes de l'œsophage (Lee et al, 2005) et de la prostate (Yu et al, 2007).

Une inactivation monoallélique conduisant à une perte partielle d'expression de l'ARNm et de la protéine GPx3 a été observée dans des lésions précurseurs bénignes de l'œsophage, tandis que l'inactivation des deux allèles était détectée dans le carcinome invasif de l'œsophage (Lee et al, 2005). Une inactivation complète du gène GPx3 par perte d'un allèle ou une diminution de l'expression par la méthylation de l'allèle restant a également été rapportée dans le cancer de la prostate (Yu et al, 2007). De plus, l'activité suppressive de la croissance tumorale et des métastases a été démontrée par des études *in vitro* et *in vivo* toujours dans la prostate (Yu et al, 2007). L'analyse de l'expression de GPX3 et la méthylation du promoteur a été étudiée dans des lignées cellulaires humaines dérivées de cancers de la tête et du cou, du poumon, du sein, de la vessie et du mélanome (Chen et al, 2011). Il en résulte une régulation négative significative de l'expression de GPx3, ou une inactivation complète du gène GPx3 dû à une méthylation dense de l'îlot CpG du promoteur GPx3 (38–85%) ; En outre, cette étude suggère que l'inactivation du gène GPx3 par la méthylation du promoteur entraîne une perte partielle ou totale de la fonction antioxydante. Ces découvertes antérieures, combinées au rôle de GPx3 en tant que composant majeur du système antioxydant dans le plasma, suggèrent que l'inactivation du gène GPx3 pourrait contribuer à la tumorigenèse d'un large spectre de types de cancer dont les adénocarcinomes gastriques.

GPx3 est associé à la SOD2 au sein de notre réseau d'interaction. La superoxyde dismutase Mn (MnSOD ou SOD2) assume une fonction de protection contre les composés réactifs de l'oxygène. Elle est de plus en plus reconnue pour son rôle dans la tumorigenèse. On pourra noter que les souris knock-out SOD2 homozygotes sont létales au stade embryonnaire (Li et al, 1995), ce qui rend difficile l'étude du rôle du déficit en SOD2 dans la tumorigenèse chez la souris. Les cas de tumorigenèse sont plus nombreux chez les souris SOD2 +/- Gpx1 - / -, mais les tumeurs sont principalement des lymphomes et, dans une moindre mesure, des adénocarcinomes du poumon (Zhang et al, 2009). Toutefois, l'expression de la SOD2 dans les tumeurs malignes d'origine tissulaire différente varie considérablement, allant d'une régulation à la baisse dans les cancers de l'œsophage (Sun et al, 2011), du pancréas (Cullen et al, 2003) et du sein (Papa et al, 2014) à un niveau élevé dans les cancers de l'ovaire (Hu et al, 2005) et du cancer colorectal (Govatati et al, 2016). Ces observations indiquent que le rôle de la SOD2 est extrêmement complexe et qu'il peut agir en tant que suppresseur de tumeur ou oncogène, en fonction de l'origine du tissu tumoral et/ou des stades de la maladie.

Au niveau pulmonaire, la voie SOD-2 est induite par l'inflammation via la voie NF- $\kappa$ B. Cette voie joue un rôle important dans la transition épithélio mésenchymateuse (TEM) et la migration des cellules tumorales dans l'adénocarcinome pulmonaire induit par des xénobiotiques comme l'aflatoxine B1. Dans le cas des cancers gastriques, une réponse inflammatoire via le TNF- $\alpha$  pourrait être le principal contributeur à la régulation positive de la SOD-2 ainsi qu'à la TEM et à *posteriori*, du développement de cellules cancéreuses.

Ceci couplé aux données de notre réseau d'interaction, montrent un lien entre la diminution de l'expression de GPx3 et l'augmentation de l'expression de la SOD2 suggérant une accumulation d'H<sub>2</sub>O<sub>2</sub> au niveau cellulaire ; une hypothèse pourrait être, que cela puisse servir de mécanisme commun pour le développement du cancer gastrique.

ADH1B participe, entre autre, au métabolisme médicamenteux. Ce gène est directement lié à GSTA1 et ALDH3a1 dans notre réseau. ADH1B voit son expression diminuer dans les adénocarcinomes de l'estomac. En 2017, Durand et al sont les premiers à avoir signalé un cas de dérégulation d'un cancer pour l'ARNm d'ADH1B dans les cancers colorectaux. Une publication récente a également identifié ADH1B et XYLT1 (une enzyme impliquée dans la biosynthèse des chaînes de glycosaminoglycanes) comme partenaires d'interaction avec ADH1C (Huttlin et al, 2015) qui est également dérégulé à la baisse dans 3 types cancers gastriques sur 4 (MSI, GS et CIN). Toujours dans le cancer colorectal, il s'avère que XYLT1 est un partenaire d'interaction de MYC, le deuxième gène le plus surexprimé dans les CRCs. MYC a abaissé la régulation des ADH1B et ADH1C, soulignant une contribution de ces enzymes de détoxification dans la pathogenèse des CCR. Dans nos données de RNAseq, MYC n'est pas dérégulé, la diminution de l'expression de ces enzymes pourrait avoir une autre cause. Toutefois une contribution de ces enzymes de détoxification dans la pathogenèse du cancer gastrique semble probable.

Les aldéhydes déshydrogénases (ALDH) oxydent les aldéhydes en acides carboxyliques en utilisant NAD ou NADP comme coenzymes. Les aldéhydes sont des molécules aliphatiques ou aromatiques hautement réactives qui jouent un rôle important dans de nombreux processus physiologiques, pathologiques et pharmacologiques. Le génome humain contient 19 gènes ALDH connus, ainsi que de nombreux pseudogènes. Il convient de noter le fait que l'activité élevée de divers ALDH, à savoir ALDH1A2, ALDH1A3, ALDH1A7, ALDH3A1 ou encore ALDH4A1 ont été observés dans des cellules souches normales et cancéreuses. La famille des ALDH3 comprend des enzymes capables d'oxyder les aldéhydes aliphatiques et aromatiques à chaîne moyenne, tels que les aldéhydes peroxydes et gras. En outre, ces enzymes ont également des fonctions non catalytiques, notamment des fonctions antioxydantes. Ces gènes sont modulés par la présence de régions régulatrices positives ou négatives. Leur activation se fait par l'intermédiaire du récepteur Ah, qui est un facteur de transcription se liant à la séquence consensus, l'élément sensible à Ah (AhRE) (Whitlock et al, 1996 ; Xiao et al, 1995). Dans la littérature, l'activité de ALDH3A1 a principalement été décrite dans le cancer du sein et le cancer du foie. Au niveau du sein, les taux d'ALDH3 varient considérablement entre les échantillons de tissus mammaires normaux, malins primaires et métastatiques. Une activité accrue de l'ALDH3A1 dans les cellules épithéliales mammaires humaines est associée à une croissance et à une survie cellulaires accrues, qui peuvent jouer un rôle important dans la promotion de la tumeur (Burchiel et al, 2007). Une étude récente a aussi démontré que ALDH3A1 est présent dans des quantités plus élevées dans les fibroblastes du cancer du sein que dans les fibroblastes mammaires normaux (Zhang et al, 2011). Dans les cellules hépatiques, l'ALDH3A1 cytosolique est très peu exprimé, contrairement à l'ALDH1 cytosolique qui, dans le foie humain normal, contribue à la détoxification de divers aldéhydes aliphatiques et aromatiques exogènes et endogènes à différentes longueurs de chaîne. De plus, les deux protéines de liaison à l'ADN découvertes dans les cancers du foie et responsables de l'augmentation de l'expression de ALDH3A1 ne sont pas retrouvées dans les tissus gastriques normaux ni dans les cellules KATO III du carcinome de l'estomac (Chang et al, 1998). Les niveaux d'ALDH3A1 sont généralement induits non seulement par des contaminants environnementaux ou par des polyphénols, comme par exemple le catéchol (Sreerama et al, 2007). Étant donné que le tube digestif et les poumons sont les principaux "points d'entrée" pour les substances alimentaires et environnementales, il est possible de supposer que ces substances peuvent induire des niveaux élevés d'ALDH3A1 sur ces sites.

On pourra aussi noter que des taux élevés d'ALDH3A1 ont été observés dans d'autres tissus comme par exemple la muqueuse gastrique normale, les tumeurs de la glande parotide, les lignées cellulaires du cancer du côlon (Vasiliou et al, 1999 ; Sreerama et al ;1997).

Au sein de notre réseau, nous observons une diminution de l'expression du gène ALDH3A1 comparé au tissu sain, ce qui laisse supposer une moins bonne prise en charge de contaminants environnementaux, de polyphénols... pouvant être à l'origine d'une promotion facilitée des cellules tumorales.

L'UDP-glucuronosyltransférase 1A1 (UGT1A1) est une enzyme essentielle impliquée dans le métabolisme du SN-38, le métabolite actif de l'irinotécan utilisé couramment dans le traitement du cancer colorectal métastatique. Il a été montré une méthylation aberrante de d'îlots CpG dans des cellules UGT1A1 négatives révélant ainsi que l'état de méthylation de son promoteur est négativement corrélé à sa transcription génique (Bélanger et al, 2010). Cette même étude suggère que le statut de méthylation du CpG-4 pourrait être un indicateur pertinent de la méthylation du promoteur proximal d'UGT1A1 et par conséquent un marqueur épigénétique potentiel de l'expression du gène UGT1A1. En outre, la régulation épigénétique du gène HNF1A, un facteur transcriptionnel, pourrait également jouer un rôle important dans la régulation du métabolisme cellulaire de l'irinotécan. Dans l'ensemble, la régulation épigénétique des gènes HNF1A et UGT1A1 pourrait être déterminante dans la prise en charge de l'agent anticancéreux SN-38.

Les AKR constituent une superfamille d'oxydoréductases à cofacteur NAD(P)(H) impliqués dans la biotransformation de nombreux composés tels que les aldéhydes aliphatiques et aromatiques, les oses aromatiques ou polycycliques, les stéroïdes, les prostaglandines, les phytoalexines isoflavinoïdes et certains xénobiotiques (Jez et al., 1997). Il a été montré que les cellules neuroendocrines du pancréas, du tractus gastro-intestinal et du poumon n'exprimaient pas AKR1C3. Les cellules neuroendocrines ont la capacité de produire, de stocker et de sécréter des peptides et des amines biogènes (Kaltsas et al, 2004).



AKR1C3 est connu pour réguler la conversion des hormones stéroïdes, des xénobiotiques et des cancérogènes chimiques. La signification biologique de l'absence d'AKR1C3 dans les cellules neuroendocrines est incertaine. Il a été montré une forte expression de AKR1C3 dans les carcinomes épidermoïdes et les adénocarcinomes du poumon, mais pas d'expression dans les carcinomes à petites cellules, une tumeur neuroendocrine du poumon extrêmement agressive (Miller et al, 2012). Il a été démontré que AKR1C3 était largement exprimé dans d'autres carcinomes non-neuroendocriniens, notamment ceux provenant de la prostate (Fung et al, 2006), de l'endomètre (Zakharov et al, 2010), de l'urothélium et du rein (Azzarello et al, 2010), et du sein (Lewis et al 2004 ; Han et al, 2008). Compte tenu des fonctions diversifiées de AKR1C3 prises en compte, ce gène peut jouer un rôle important dans les effets thérapeutiques des chimiothérapies et des radiothérapies sur ces cancers. On peut alors penser que le faible taux d'expression de AKR1C3 dans les tumeurs neuroendocrines et les adénocarcinomes gastriques, et le taux élevé d'expression dans les carcinomes épidermoïdes, comme le démontrent plusieurs études (Fung et al, 2006 ; Azzarello et al, 2009 et Park et al, 2010), pourraient fournir un marqueur de diagnostic supplémentaire entre les tumeurs NE et non NE.

Enfin, dernier gène présent dans notre réseau, GGT6 est une gamma-glutamyl transferase (GGT ou, S-glutamyl-peptide-amino-acide-S-glutamyltransférase). C'est une enzyme membranaire dimérique, amphiphile, fortement glycosylée (Meister et al, 1981). Il existe en fait plusieurs isoformes de la GGT différentes par leurs mobilités électrophorétiques. Elles sont essentielles au maintien de l'homéostasie de la cystéine, à la synthèse des leucotriènes, au métabolisme des conjugués de glutathion et au catabolisme du glutathion extracellulaire. Peu d'études concernant les gamma-glutamyl transferases ont analysé l'expression et l'activité de GGT6. Malgré tout, Lasseigne et al (2014) ont noté que le premier intron du gène codant la GGT6 est hyperméthylé dans les cancers du rein par rapport aux tissus normaux et, corrélativement, les niveaux de transcription de GGT6 étaient significativement réduits. Ce niveau d'expression pourrait être lié à l'infiltration immunitaire et à l'inflammation, et potentiellement à des modifications de la méthylation de l'ADN par le biais de l'évolution des niveaux de glutathion et de la disponibilité des groupes méthyles (Heisterkamp et al, 2008 ; Lertratanangkoon et al, 1997).

Il semble ainsi, qu'en grande majorité, la diminution de l'expression des gènes de notre réseau retrouvé dans d'autre type de cancers soit due à l'hyperméthylation des promoteurs de ces gènes. Il est donc important de caractériser la méthylation de ces gènes au sein de nos tumeurs gastriques.

Cette sous-expression des gènes du réseau Papillon dans le cancer gastrique semble ainsi pouvoir servir de nouveaux marqueur de diagnostic. Il reste aussi à déterminer de façon plus précise si cette susceptibilité accrue de l'estomac à un ensemble d'agressions génotoxiques peut être associée l'induction de ce cancer. Pour répondre à cette question, il faudrait pouvoir disposer d'un ensemble de lésions précoces, comme par exemple des prélèvements d'œsophage de Barrett. Bien que ce type de prélèvements étagés le long de la muqueuse soit difficile à obtenir, il est possible d'analyser les conséquences d'une modulation de l'expression de l'un ou de plusieurs des gènes du réseau Papillon, puis de déterminer si les traitements de chimiothérapie seraient plus efficaces dans des cellules cancéreuses de cancer gastrique.

# Conclusion Générale

---

L'objectif majeur de ce projet est d'aboutir à l'identification de nouvelles cibles pronostiques et/ou thérapeutiques. Le but, à terme, étant la mise en œuvre d'une médecine de plus en plus personnalisée et qui permettrait de prendre en compte un ensemble de données caractérisant de la manière la plus précise la tumeur de chaque patient, et ainsi lui fournir le traitement le plus adapté. Ce travail, de par son lien avec le domaine clinique s'inscrit dans cet objectif.

Ces travaux de recherche se sont intéressés à étudier l'effet de mutations qui touchent des répétitions non codantes du génome, caractéristiques de l'instabilité microsatellitaire. Ces observations ont permis d'identifier de nouvelles altérations d'épissages spécifiques des tumeurs MSI aussi bien dans l'estomac que dans le colon ou l'endomètre. Ainsi, ces travaux viennent naturellement compléter des observations récemment effectuées dans ces cancers, sur les gènes *HSP110*, *ATM* ou *MRE11*. Et même s'il reste encore à éclaircir de manière exacte le mécanisme de survenue de ces anomalies d'épissage, la multiplication de ces observations suggère qu'il existe un lien étroit entre anomalies d'épissage et cancers MSI.

Mes travaux ont également permis d'identifier un ensemble de gènes, dont les fonctions communes de biotransformation et de régulation du stress oxydant semblent jouer un rôle clé dans la progression tumorale et la réponse au traitement de type chimiothérapie. La continuité de ce travail, notamment par l'étude des profils de méthylation de ces gènes, participera certainement à faire la lumière sur la question de leur rôle, à la fois au niveau moléculaire (compréhension des mécanismes d'initiation et de progression tumorale et intégration des différents processus génétiques et épigénétiques) et cliniques (diagnostic, pronostic et thérapeutique).

# Bibliographie

---

- Aaltonen, Lauri A, Paivi Peltomaki, Fredrick S Leach, Pertti Sistonen, Lea Pylkkanen, Jukka-Pekka Mecklin, Heikki Jarvinen, Steven M Powell, and Jin Jen. "Clues to the Pathogenesis of Familial Colorectal Cancer." *Science* 260, no. 5109 (1993): 812-16.
- Adams, Melissa D, David Z Rudner, and Donald C Rio. "Biochemistry and Regulation of Pre-Mrna Splicing." *Current opinion in cell biology* 8, no. 3 (1996): 331-39.
- Aenkoe, Minna-Liisa. "Regulation of Gene Expression Programmes by Serine–Arginine Rich Splicing Factors." Paper presented at the Seminars in cell & developmental biology, 2014.
- Alfarano, C, CE Andrade, K Anthony, N Bahroos, M Bajec, K Bantoft, Doron Betel, *et al.* "The Biomolecular Interaction Network Database and Related Tools 2005 Update." *Nucleic acids research* 33, no. suppl\_1 (2005): D418-D24.
- Allgayer, H., R. Babic, K. U. Grutzner, B. C. Beyer, A. Tarabichi, F. W. Schildberg, and M. M. Heiss. "Tumor-Associated Proteases and Inhibitors in Gastric Cancer: Analysis of Prognostic Impact and Individual Risk Protease Patterns." [In eng]. *Clin Exp Metastasis* 16, no. 1 (Jan 1998): 62-73.
- André, Fabrice, Stefan Michiels, Philippe Dessen, Veronique Scott, Voichita Suci, Catherine Uzan, Vladimir Lazar, *et al.* "Exonic Expression Profiling of Breast Cancer and Benign Lesions: A Retrospective Analysis." *The lancet oncology* 10, no. 4 (2009): 381-90.
- Arocena, M Gloria, A Barturen, L Bujanda, O Casado, MM Ramírez, JM Oleagoitia, M Galdiz Iturri, *et al.* "Mri and Endoscopic Ultrasonography in the Staging of Gastric Cancer." *Revista Espanola de Enfermedades Digestivas* 98, no. 8 (2006): 582-90.
- Auboeuf, Didier, Eric Batsché, Martin Dutertre, Christian Muchardt, and Bert W O'Malley. "Coregulators: Transducing Signal from Transcription to Alternative Splicing." *Trends in Endocrinology & Metabolism* 18, no. 3 (2007): 122-29.
- Azzarello, Joseph T, Hsueh-Kung Lin, Awet Gherezghiher, Vladislav Zakharov, Zhongxin Yu, Bradley P Kropp, Daniel J Culkun, Trevor M Penning, and Kar-Ming Fung. "Expression of Akr1c3 in Renal Cell Carcinoma, Papillary Urothelial Carcinoma, and Wilms' Tumor." *International journal of clinical and experimental pathology* 3, no. 2 (2010): 147.
- Baglietto, Laura, Noralane M Lindor, James G Dowty, Darren M White, Anja Wagner, Encarna B Gomez Garcia, Annette HJT Vriends, *et al.* "Risks of Lynch Syndrome Cancers for Msh6 Mutation Carriers." *Journal of the National Cancer Institute* 102, no. 3 (2010): 193-201.
- Bang, Yung-Jue, Eric Van Cutsem, Andrea Feyereislova, Hyun C Chung, Lin Shen, Akira Sawaki, Florian Lordick, *et al.* "Trastuzumab in Combination with Chemotherapy Versus Chemotherapy Alone for Treatment of Her2-Positive Advanced Gastric or Gastro-Oesophageal Junction Cancer (Toga): A Phase 3, Open-Label, Randomised Controlled Trial." *The Lancet* 376, no. 9742 (2010): 687-97.
- Bardag-Gorce, Fawzia, Jun Li, Barbara A French, and Samuel W French. "The Effect of Ethanol-Induced Cyp2e1 on Proteasome Activity: The Role of 4-Hydroxynonenal." *Experimental and molecular pathology* 78, no. 2 (2005): 109-15.
- Beghelli, Stefania, Giovanni de Manzoni, Stefano Barbi, Anna Tomezzoli, Franco Roviello, Carmela Di Gregorio, Carla Vindigni, *et al.* "Microsatellite Instability in Gastric Cancer Is Associated with Better Prognosis in Only Stage II Cancers." *Surgery* 139, no. 3 (2006): 347-56.
- Bélanger, Anne-Sophie, Jelena Tojcic, Mario Harvey, and Chantal Guillemette. "Regulation of Ugt1a1 and Hnf1 Transcription Factor Gene Expression by DNA Methylation in Colon Cancer Cells." *BMC molecular biology* 11, no. 1 (2010): 9.
- Berchem, GJ, M Bosseler, Natacha Mine, and Bernard Avalosse. "Nanomolar Range Docetaxel Treatment Sensitizes Mcf-7 Cells to Chemotherapy Induced Apoptosis, Induces G2m Arrest and Phosphorylates Bcl-2."

- Anticancer research* 19, no. 1A (1999): 535-40.
- Berlth, Felix, Elfriede Bollschweiler, Uta Drebber, Arnulf H Hoelscher, and Stefan Moenig. "Pathohistological Classification Systems in Gastric Cancer: Diagnostic Relevance and Prognostic Value." *World journal of gastroenterology: WJG* 20, no. 19 (2014): 5679.
- Béroud, Christophe, Dalil Hamroun, Gwenaëlle Collod-Béroud, Catherine Boileau, Thierry Soussi, and Mireille Claustres. "Umd (Universal Mutation Database): 2005 Update." *Human mutation* 26, no. 3 (2005): 184-91.
- Binder-Foucard, F, Aurélien Belot, Patricia Delafosse, L Remontet, AS Woronoff, and N Bossard. "Estimation Nationale De L'incidence Et De La Mortalité Par Cancer En France Entre 1980 Et 2012." *Partie*, 2013.
- Biroccio, Annamaria, Barbara Benassi, Igea D'Agnano, Carmen D'Angelo, Simonetta Buglioni, Marcella Mottolese, Andrea Ricciotti, *et al.* "C-Myb and Bcl-X Overexpression Predicts Poor Prognosis in Colorectal Cancer: Clinical and Experimental Findings." *The American journal of pathology* 158, no. 4 (2001): 1289-99.
- Black, DL, and PJ Grabowski. "Alternative Pre-Mrna Splicing and Neuronal Function." In *Regulation of Alternative Splicing*, 187-216: Springer, 2003.
- Blencowe, Benjamin J. "Alternative Splicing: New Insights from Global Analyses." *Cell* 126, no. 1 (2006): 37-47.
- Bonadona, Valérie, Bernard Bonaïti, Sylviane Olschwang, Sophie Grandjouan, Laetitia Huiart, Michel Longy, Rosine Guimbaud, *et al.* "Cancer Risks Associated with Germline Mutations in Mlh1, Msh2, and Msh6 Genes in Lynch Syndrome." *Jama* 305, no. 22 (2011): 2304-10.
- Breitkreutz, Bobby-Joe, Chris Stark, Teresa Reguly, Lorrie Boucher, Ashton Breitkreutz, Michael Livstone, Rose Oughtred, *et al.* "The Biogrid Interaction Database: 2008 Update." *Nucleic acids research* 36, no. suppl\_1 (2007): D637-D40.
- Brigelius-Flohe, Regina, and Anna Kipp. "Glutathione Peroxidases in Different Stages of Carcinogenesis." *Biochimica et Biophysica Acta (BBA)-General Subjects* 1790, no. 11 (2009): 1555-68.
- Brinkman, Brigitta MN. "Splice Variants as Cancer Biomarkers." *Clinical biochemistry* 37, no. 7 (2004): 584-94.
- Bronner, C Eric, Sean M Baker, Paul T Morrison, Gwynedd Warren, Leslie G Smith, Mary Kay Lescoe, Michael Kane, *et al.* "Mutation in the DNA Mismatch Repair Gene Homologue Hmlh 1 Is Associated with Hereditary Non-Polyposis Colon Cancer." *Nature* 368, no. 6468 (1994): 258.
- Buhard, Olivier, Francesca Cattaneo, Yick Fu Wong, So Fan Yim, Eitan Friedman, Jean-Francois Flejou, Alex Duval, and Richard Hamelin. "Multipopulation Analysis of Polymorphisms in Five Mononucleotide Repeats Used to Determine the Microsatellite Instability Status of Human Tumors." *J Clin Oncol* 24, no. 2 (2006): 241-51.
- Burchiel, Scott W, Todd A Thompson, Fredine T Lauer, and Tudor I Oprea. "Activation of Dioxin Response Element (Dre)-Associated Genes by Benzo (a) Pyrene 3, 6-Quinone and Benzo (a) Pyrene 1, 6-Quinone in Mcf-10a Human Mammary Epithelial Cells." *Toxicology and applied pharmacology* 221, no. 2 (2007): 203-14.
- Burge, Christopher B, Richard A Padgett, and Phillip A Sharp. "Evolutionary Fates and Origins of U12-Type Introns." *Molecular cell* 2, no. 6 (1998): 773-85.
- Cai, Lin, Shun-Zhang Yu, and Zuo-Feng Zhang. "Cytochrome P450 2e1 Genetic Polymorphism and Gastric Cancer in Change, Fujian Province." *World journal of gastroenterology* 7, no. 6 (2001): 792.
- Capelle, Lisette G, Annemarie C de Vries, Jelle Haringsma, Frank Ter Borg, Richard A de Vries, Marco J Bruno, Herman van Dekken, *et al.* "The Staging of Gastritis with the Olga System by Using Intestinal Metaplasia as an Accurate Alternative for Atrophic Gastritis." *Gastrointestinal endoscopy* 71, no. 7 (2010): 1150-58.

- Chang, CHENG, LILY C Hsu, VIBHA Davé, and AKIRA Yoshida. "Expression of Human Aldehyde Dehydrogenase-3 Associated with Hepatocellular Carcinoma: Promoter Regions and Nuclear Protein Factors Related to the Expression." *International journal of molecular medicine* 2, no. 3 (1998): 333-41.
- Chang, GWM, and PCA Kam. "The Physiological and Pharmacological Roles of Cytochrome P450 Isoenzymes." *Anaesthesia* 54, no. 1 (1999): 42-50.
- Chang, Wen-Jun, Yan Du, Xin Zhao, Li-Ye Ma, and Guang-Wen Cao. "Inflammation-Related Factors Predicting Prognosis of Gastric Cancer." *World journal of gastroenterology: WJG* 20, no. 16 (2014): 4586.
- Chapiro, J. "Différentes Thérapeutiques Anti-Her2 Dans Le Cancer Du Sein: Anticorps Monoclonaux (Trastuzumab, Pertuzumab, T-Dm1), Inhibiteurs De Tyrosine Kinase (Lapatinib, Nératinib). Mécanismes D'action Et De Résistance." *Cancer du sein* (2012): 489-97.
- Chatr-Aryamontri, Andrew, Arnaud Ceol, Luisa Montecchi Palazzi, Giuliano Nardelli, Maria Victoria Schneider, Luisa Castagnoli, and Gianni Cesareni. "Mint: The Molecular Interaction Database." *Nucleic acids research* 35, no. suppl\_1 (2006): D572-D74.
- Chen, Baishen, Xi Rao, Michael G House, Kenneth P Nephew, Kevin J Cullen, and Zhongmin Guo. "Gpx3 Promoter Hypermethylation Is a Frequent Event in Human Cancer and Is Associated with Tumorigenesis and Chemotherapy Response." *Cancer letters* 309, no. 1 (2011): 37-45.
- Chen, Tie, Kun Yang, Jianhua Yu, Wentong Meng, Dandan Yuan, Feng Bi, Fang Liu, *et al.* "Identification and Expansion of Cancer Stem Cells in Tumor Tissues and Peripheral Blood Derived from Gastric Adenocarcinoma Patients." *Cell research* 22, no. 1 (2012): 248.
- Collura, Ada, Anaïs Lagrange, Magali Svrcek, Laetitia Marisa, Olivier Buhard, Agathe Guilloux, Kristell Wanherdrick, *et al.* "Patients with Colorectal Tumors with Microsatellite Instability and Large Deletions in Hsp110 T17 Have Improved Response to 5-Fluorouracil-Based Chemotherapy." *Gastroenterology* 146, no. 2 (2014): 401-11. e1.
- Correa, Pelayo. "Human Gastric Carcinogenesis: A Multistep and Multifactorial Process—First American Cancer Society Award Lecture on Cancer Epidemiology and Prevention." *Cancer research* 52, no. 24 (1992): 6735-40.
- Corcos, Laurent, and Stéphanie Solier. "Épissage Alternatif, Pathologie Et Thérapeutique Moléculaire." *médecine/sciences* 21, no. 3 (2005): 253-60.
- Corso, Giovanni, Sérgia Velho, Joana Paredes, Corrado Pedrazzani, Diana Martins, Fernanda Milanezi, Valeria Pascale, *et al.* "Oncogenic Mutations in Gastric Cancer with Microsatellite Instability." *European Journal of Cancer* 47, no. 3 (2011): 443-51.
- Cullen, Joseph J, Frank A Mitros, and Larry W Oberley. "Expression of Antioxidant Enzymes in Diseases of the Human Pancreas: Another Link between Chronic Pancreatitis and Pancreatic Cancer." *Pancreas* 26, no. 1 (2003): 23-27.
- Dehm, Scott M. "Mrna Splicing Variants: Exploiting Modularity to Outwit Cancer Therapy." *Cancer research* 73, no. 17 (2013): 5309-14.
- Deng, Niantao, Liang Kee Goh, Hannah Wang, Kakoli Das, Jiong Tao, Iain Beehuat Tan, Shenli Zhang, *et al.* "A Comprehensive Survey of Genomic Alterations in Gastric Cancer Reveals Systematic Patterns of Molecular Exclusivity and Co-Occurrence among Distinct Therapeutic Targets." *Gut* (2012): gutjnl-2011-301839.
- Deng, Qiwen, Bangshun He, Yuqin Pan, Huiling Sun, Xian Liu, Jie Chen, Houqun Ying, *et al.* "Polymorphisms of Gsta1 Contribute to Elevated Cancer Risk: Evidence from 15 Studies." *J BUON* 20, no. 1 (2015): 287-95.



- Domínguez, Gemma, José M García, Cristina Peña, Javier Silva, Vanesa García, Lara Martínez, Constanza Maximiano, *et al.* "Δtap73 Upregulation Correlates with Poor Prognosis in Human Tumors: Putative in Vivo Network Involving P73 Isoforms, P53, and E2f-1." *Journal of clinical oncology* 24, no. 5 (2006): 805-15.
- Dorard, Coralie, Aurélie De Thonel, Ada Collura, Laetitia Marisa, Magali Svrcek, Anaïs Lagrange, Gaetan Jego, *et al.* "Expression of a Mutant Hsp110 Sensitizes Colorectal Cancer Cells to Chemotherapy and Improves Disease Prognosis." *Nature medicine* 17, no. 10 (2011): 1283.
- Ducreux, Michel, Valérie Boige, Julien Taieb, Emmanuel Mitry, and Philippe Rougier. "La Linéite: Une Forme Particulière De Cancer Gastrique." *Hépatogastro & Oncologie Digestive* 7, no. 4 (2000): 263-7.
- Dujardin, Gwendal, Celina Lafaille, Manuel de la Mata, Luciano E Marasco, Manuel J Muñoz, Catherine Le Jossic-Corcós, Laurent Corcos, and Alberto R Kornblihtt. "How Slow Rna Polymerase Ii Elongation Favors Alternative Exon Skipping." *Molecular cell* 54, no. 4 (2014): 683-90.
- Durand, Stéphanie, Killian Trillet, Arnaud Uguen, Aude Saint-Pierre, Catherine Le Jossic-Corcós, and Laurent Corcos. "A Transcriptome-Based Protein Network That Identifies New Therapeutic Targets in Colorectal Cancer." *BMC genomics* 18, no. 1 (2017): 758.
- Duval, Alex, and Hamelin, Richard. "Mutations at Coding Repeat Sequences in Mismatch Repair-Deficient Human Cancers: Toward a New Concept of Target Genes for Instability." *Cancer research* 62, no. 9 (2002): 2447-54.
- Duval, Alex and Hamelin, Richard. "Réparation Des Erreurs De Réplication, Microsatellites Et Cancer." *médecine/sciences* 19, no. 1 (2003): 55-62.
- Echeverria, Gloria V, and Thomas A Cooper. "Rna-Binding Proteins in Microsatellite Expansion Disorders: Mediators of Rna Toxicity." *Brain research* 1462 (2012): 100-11.
- Edge, Stephen B, and Carolyn C Compton. "The American Joint Committee on Cancer: The 7th Edition of the Ajcc Cancer Staging Manual and the Future of Tnm." *Annals of surgical oncology* 17, no. 6 (2010): 1471-74.
- Ejima, Yosuke, Lichun Yang, and Masao S Sasaki. "Aberrant Splicing of the Atm Gene Associated with Shortening of the Intronic Mononucleotide Tract in Human Colon Tumor Cell Lines: A Novel Mutation Target of Microsatellite Instability." *International journal of cancer* 86, no. 2 (2000): 262-68.
- El-Bchiri, Jamila, Agathe Guilloux, Peggy Dartigues, Etienne Loire, Dominique Mercier, Olivier Buhard, Iradj Sobhani, *et al.* "Nonsense-Mediated Mrna Decay Impacts Msi-Driven Carcinogenesis and Anti-Tumor Immunity in Colorectal Cancers." *PLoS One* 3, no. 7 (2008): e2583.
- Fabbri, Francesco, Silvia Carloni, Giovanni Brigliadori, Wainer Zoli, Rosa Lapalombella, and Marina Marini. "Sequential Events of Apoptosis Involving Docetaxel, a Microtubule-Interfering Agent: A Cytometric Study." *BMC Cell Biology* 7, no. 1 (2006): 6.
- Fishel, Richard, Mary Kay Lescoe, MRS Rao, Neal G Copeland, Nancy A Jenkins, Judy Garber, Michael Kane, and Richard Kolodner. "The Human Mutator Gene Homolog Msh2 and Its Association with Hereditary Nonpolyposis Colon Cancer." *Cell* 75, no. 5 (1993): 1027-38.
- Fujimoto-Ouchi, Kaori, Fumiko Sekiguchi, Hideyuki Yasuno, Yoichiro Moriya, Kazushige Mori, and Yutaka Tanaka. "Antitumor Activity of Trastuzumab in Combination with Chemotherapy in Human Gastric Cancer Xenograft Models." *Cancer chemotherapy and pharmacology* 59, no. 6 (2007): 795-805.
- Fung, KM, ENS Samara, C Wong, A Metwalli, R Krlin, B Bane, CZ Liu, *et al.* "Increased Expression of Type 2 3α-Hydroxysteroid Dehydrogenase/Type 5 17β-Hydroxysteroid Dehydrogenase (Akr1c3) and Its Relationship with Androgen Receptor in Prostate Carcinoma." *Endocrine-related cancer* 13, no. 1 (2006): 169-80.
- Funkhouser Jr, William K, Ira M Lubin, Federico A Monzon, Barbara A Zehnbauser, James P Evans, Shuji Ogino, and Jan A Nowak. "Relevance, Pathogenesis, and Testing Algorithm for Mismatch Repair-Defective Colorectal Carcinomas: A Report of the Association for Molecular Pathology." *The Journal of Molecular Diagnostics* 14,

no. 2 (2012): 91-103.

- Galka-Marciniak, Paulina, Martyna O Urbanek, and Włodzimierz J Krzyżosiak. "Triplet Repeats in Transcripts: Structural Insights into Rna Toxicity." (2012).
- Gao, Feng, Diane L Johnson, Sean Ekins, John Janiszewski, Kevin G Kelly, R Daniel Meyer, and Michael West. "Optimizing Higher Throughput Methods to Assess Drug-Drug Interactions for Cyp1a2, Cyp2c9, Cyp2c19, Cyp2d6, Rcy2d6, and Cyp3a4 in Vitro Using a Single Point Ic50." *Journal of biomolecular screening* 7, no. 4 (2002): 373-82.
- Gardina, Paul J, Tyson A Clark, Brian Shimada, Michelle K Staples, Qing Yang, James Veitch, Anthony Schweitzer, *et al.* "Alternative Splicing and Differential Gene Expression in Colon Cancer Detected by a Whole Genome Exon Array." *BMC genomics* 7, no. 1 (2006): 325.
- Ghigna, Claudia, Mauro Moroni, Camillo Porta, Silvano Riva, and Giuseppe Biamonti. "Altered Expression of Heterogeneous Nuclear Ribonucleoproteins and Sr Factors in Human Colon Adenocarcinomas." *Cancer research* 58, no. 24 (1998): 5818-24.
- Giannini, Giuseppe, Christian Rinaldi, Elisabetta Ristori, Maria Irene Ambrosini, Fabio Cerignoli, Alessandra Viel, Ettore Bidoli, *et al.* "Mutations of an Intronic Repeat Induce Impaired Mre11 Expression in Primary Human Cancer with Microsatellite Instability." *Oncogene* 23, no. 15 (2004): 2640.
- Goldstein, J, B Tran, J Ensor, P Gibbs, HL Wong, SF Wong, E Vilar, *et al.* "Multicenter Retrospective Analysis of Metastatic Colorectal Cancer (Crc) with High-Level Microsatellite Instability (Msi-H)." *Annals of Oncology* 25, no. 5 (2014): 1032-38.
- Goseki, N, T Takizawa, and M Koike. "Differences in the Mode of the Extension of Gastric Cancer Classified by Histological Type: New Histological Classification of Gastric Carcinoma." *Gut* 33, no. 5 (1992): 606-12.
- Govatati, Suresh, Sravanthi Malempati, Bulle Saradamma, Dasi Divyamaanasa, B Prathap Naidu, Pallaval Veera Bramhachari, Nagesh Narayana, *et al.* "Manganese-Superoxide Dismutase (Mn-Sod) Overexpression Is a Common Event in Colorectal Cancers with Mitochondrial Microsatellite Instability." *Tumor Biology* 37, no. 8 (2016): 10357-64.
- Graus-Porta, Diana, Roger R Beerli, John M Daly, and Nancy E Hynes. "ErbB-2, the Preferred Heterodimerization Partner of All ErbB Receptors, Is a Mediator of Lateral Signaling." *The EMBO journal* 16, no. 7 (1997): 1647-55.
- Graveley, Brenton R. "Sorting out the Complexity of Sr Protein Functions." *Rna* 6, no. 9 (2000): 1197-211.
- Grodstein, F., P. A. Newcomb, and M. J. Stampfer. "Postmenopausal Hormone Therapy and the Risk of Colorectal Cancer: A Review and Meta-Analysis." [In eng]. *Am J Med* 106, no. 5 (May 1999): 574-82.
- Groh, Matthias, Michele MP Lufino, Richard Wade-Martins, and Natalia Gromak. "R-Loops Associated with Triplet Repeat Expansions Promote Gene Silencing in Friedreich Ataxia and Fragile X Syndrome." *PLoS genetics* 10, no. 5 (2014): e1004318.
- Grosso, Ana Rita, Sandra Martins, and Maria Carmo-Fonseca. "The Emerging Role of Splicing Factors in Cancer." *EMBO reports* 9, no. 11 (2008): 1087-93.
- Guilford, Parry, Justin Hopkins, James Harraway, Maybelle McLeod, Ngahiraka McLeod, Pauline Harawira, Huriana Taite, *et al.* "E-Cadherin Germline Mutations in Familial Gastric Cancer." *Nature* 392, no. 6674 (1998): 402.
- Ham, Maria Francisca, Tetsuya Takakuwa, Wen-Juan Luo, Angen Liu, Akira Horii, and Katsuyuki Aozasa. "Impairment of Double-Strand Breaks Repair and Aberrant Splicing of Atm and Mre11 in Leukemia-Lymphoma Cell Lines with Microsatellite Instability." *Cancer science* 97, no. 3 (2006): 226-34.

- Hamosh, Ada, Alan F Scott, Joanna S Amberger, Carol A Bocchini, and Victor A McKusick. "Online Mendelian Inheritance in Man (Omim), a Knowledgebase of Human Genes and Genetic Disorders." *Nucleic acids research* 33, no. suppl\_1 (2005): D514-D17.
- Hampel, H, WL Frankel, E Martin, M Arnold, K Khanduja, P Kuebler, H Nakagawa, *et al.* "Screening for the Lynch Syndrome (Hereditary Nonpolyposis Colorectal Cancer) N Engl J Med. 2005; 352: 1851–1860. Doi: 10.1056." *NEJMoa043146.[PubMed][Cross Ref]* (2005).
- Han, B, S Li, D Song, D Poisson-Pare, G Liu, V Luu-The, J Ouellet, F Labrie, and G Pelletier. "Expression of 17 $\beta$ -Hydroxysteroid Dehydrogenase Type 2 and Type 5 in Breast Cancer and Adjacent Non-Malignant Tissue: A Correlation to Clinicopathological Parameters." *The Journal of steroid biochemistry and molecular biology* 112, no. 4-5 (2008): 194-200.
- Haraldsdottir, Sigurdís, Heather Hampel, Jerneja Tomsic, Wendy L Frankel, Rachel Pearlman, Albert De La Chapelle, and Colin C Pritchard. "Colon and Endometrial Cancers with Mismatch Repair Deficiency Can Arise from Somatic, Rather Than Germline, Mutations." *Gastroenterology* 147, no. 6 (2014): 1308-16. e1.
- Hartgrink, Henk H, Edwin PM Jansen, Nicole CT van Grieken, and Cornelis JH van de Velde. "Gastric Cancer." *The Lancet* 374, no. 9688 (2009): 477-90.
- Hato, Stanleyson V, Andrea Khong, I Jolanda M de Vries, and W Joost Lesterhuis. "Molecular Pathways: The Immunogenic Effects of Platinum-Based Chemotherapeutics." *Clinical Cancer Research* 20, no. 11 (2014): 2831-37.
- Heisterkamp, Nora, John Groffen, David Warburton, and Tam P Sneddon. "The Human Gamma-Glutamyltransferase Gene Family." *Human genetics* 123, no. 4 (2008): 321-32.
- Helsby, NA, Wing Yee Lo, K Sharples, G Riley, M Murray, K Spells, M Dzhelai, A Simpson, and Michael Findlay. "Cyp2c19 Pharmacogenetics in Advanced Cancer: Compromised Function Independent of Genotype." *British journal of cancer* 99, no. 8 (2008): 1251.
- Hong, Eurie L, Rama Balakrishnan, Qing Dong, Karen R Christie, Julie Park, Gail Binkley, Maria C Costanzo, *et al.* "Gene Ontology Annotations at Sgd: New Data Sources and Annotation Methods." *Nucleic acids research* 36, no. suppl\_1 (2007): D577-D81.
- Hu, Bing, Nassim El Hajj, Scott Sittler, Nancy Lammert, Robert Barnes, and Aurelia Meloni-Ehrig. "Gastric Cancer: Classification, Histology and Application of Molecular Pathology." *Journal of gastrointestinal oncology* 3, no. 3 (2012): 251.
- Hu, Yumin, Daniel G Rosen, Yan Zhou, Li Feng, Gong Yang, Jinsong Liu, and Peng Huang. "Mitochondrial Manganese-Superoxide Dismutase Expression in Ovarian Cancer Role in Cell Proliferation and Response to Oxidative Stress." *Journal of Biological Chemistry* 280, no. 47 (2005): 39485-92.
- Huang, Hanhua, Steven C Campbell, Thomas Neliús, Dhugal F Bedford, Dorina Veliceasa, Noel P Bouck, and Olga V Volpert. "A1-Antitrypsin Inhibits Angiogenesis and Tumor Growth." *International journal of cancer* 112, no. 6 (2004): 1042-48.
- Hugo, Honor, Alina Cures, Nirosha Suraweera, Yvette Drabsch, Damian Purcell, Theo Mantamadiotis, Wayne Phillips, *et al.* "Mutations in the Myb Intron I Regulatory Sequence Increase Transcription in Colon Cancers." *Genes, Chromosomes and Cancer* 45, no. 12 (2006): 1143-54.
- Huttlin, Edward L, Lily Ting, Raphael J Bruckner, Fana Gebreab, Melanie P Gygi, John Szpyt, Stanley Tam, *et al.* "The Bioplex Network: A Systematic Exploration of the Human Interactome." *Cell* 162, no. 2 (2015): 425-40.
- Inokuchi, Mikito, Sho Otsuki, Yoshitaka Fujimori, Yuya Sato, Masatoshi Nakagawa, and Kazuyuki Kojima. "Clinical Significance of Met in Gastric Cancer." *World journal of gastrointestinal oncology* 7, no. 11 (2015): 317.

- Ionov, Yuriy, Miguel A Peinado, Sergei Malkhosyan, Darryl Shibata, and Manuel Perucho. "Ubiquitous Somatic Mutations in Simple Repeated Sequences Reveal a New Mechanism for Colonic Carcinogenesis." *Nature* 363, no. 6429 (1993): 558.
- Jass, Jeremy R, Vicki LJ Whitehall, Joanne Young, and Barbara A Leggett. "Emerging Concepts in Colorectal Neoplasia." *Gastroenterology* 123, no. 3 (2002): 862-76.
- Jensen, Lars J, Michael Kuhn, Manuel Stark, Samuel Chaffron, Chris Creevey, Jean Muller, Tobias Doerks, *et al.* "String 8—a Global View on Proteins and Their Functional Interactions in 630 Organisms." *Nucleic acids research* 37, no. suppl\_1 (2008): D412-D16.
- Jez, Joseph M, T Geoffrey Flynn, and Trevor M Penning. "A New Nomenclature for the Aldo-Keto Reductase Superfamily." *Biochemical pharmacology* 54, no. 6 (1997): 639-47.
- Jolissaint, Laurianne, Antoine de Pauw, and Bruno Buecher. "Les Formes Héritaires Et Familiales Des Cancers De L'estomac." (2011).
- Kaltsas, Gregory A, G Michael Besser, and Ashley B Grossman. "The Diagnosis and Medical Management of Advanced Neuroendocrine Tumors." *Endocrine reviews* 25, no. 3 (2004): 458-511.
- Kane, Michael F, Massimo Loda, Gretchen M Gaida, Jennifer Lipman, Rajesh Mishra, Harvey Goldman, J Milburn Jessup, and Richard Kolodner. "Methylation of the Hmlh1 Promoter Correlates with Lack of Expression of Hmlh1 in Sporadic Colon Tumors and Mismatch Repair-Defective Human Tumor Cell Lines." *Cancer research* 57, no. 5 (1997): 808-11.
- Kanehisa, Minoru, Michihiro Araki, Susumu Goto, Masahiro Hattori, Mika Hirakawa, Masumi Itoh, Toshiaki Katayama, *et al.* "Kegg for Linking Genomes to Life and the Environment." *Nucleic acids research* 36, no. suppl\_1 (2007): D480-D84.
- Karni, Rotem, Elisa de Stanchina, Scott W Lowe, Rahul Sinha, David Mu, and Adrian R Krainer. "The Gene Encoding the Splicing Factor Sf2/Asf Is a Proto-Oncogene." *Nature Structural and Molecular Biology* 14, no. 3 (2007): 185.
- Kerrien, Samuel, Yasmin Alam-Faruque, Bruno Aranda, I Bancarz, Alan Bridge, C Derow, Emily Dimmer, *et al.* "Intact—Open Source Resource for Molecular Interaction Data." *Nucleic acids research* 35, no. suppl\_1 (2006): D561-D65.
- Keseler, Ingrid M, César Bonavides-Martínez, Julio Collado-Vides, Socorro Gama-Castro, Robert P Gunsalus, D Aaron Johnson, Markus Krummenacker, *et al.* "Ecocyc: A Comprehensive View of Escherichia Coli Biology." *Nucleic acids research* 37, no. suppl\_1 (2008): D464-D70.
- Keshava Prasad, TS, Renu Goel, Kumaran Kandasamy, Shivakumar Keerthikumar, Sameer Kumar, Suresh Mathivanan, Deepthi Telikicherla, *et al.* "Human Protein Reference Database—2009 Update." *Nucleic acids research* 37, no. suppl\_1 (2008): D767-D72.
- Kim, Eddo, Amir Goren, and Gil Ast. "Insights into the Connection between Cancer and Alternative Splicing." *Trends in Genetics* 24, no. 1 (2008): 7-10.
- Kim, Tae-Min, Peter W Laird, and Peter J Park. "The Landscape of Microsatellite Instability in Colorectal and Endometrial Cancer Genomes." *Cell* 155, no. 4 (2013): 858-68.
- Kohlmann, Wendy, and Stephen B Gruber. "Lynch Syndrome." In *Genereviews*® [Internet]: University of Washington, Seattle, 2018.
- Komiya, Yasuhiro, Hiromasa Tsukino, Hiroyuki Nakao, Yoshiki Kuroda, Hirohisa Imai, and Takahiko Katoh. "Human Glutathione S-Transferase A1, T1, M1, and P1 Polymorphisms and Susceptibility to Prostate Cancer in the Japanese Population." *Journal of cancer research and clinical oncology* 131, no. 4 (2005): 238-42.

- Koopman, M, GAM Kortman, L Mekenkamp, MJL Ligtenberg, N Hoogerbrugge, NF Antonini, CJA Punt, and JHJM Van Krieken. "Deficient Mismatch Repair System in Patients with Sporadic Advanced Colorectal Cancer." *British journal of cancer* 100, no. 2 (2009): 266.
- Kunz, C, Y Saito, and P Schär. "DNA Repair in Mammalian Cells." *Cellular and molecular life sciences* 66, no. 6 (2009): 1021-38.
- Lamarque, Dominique. "Épidémiologie De L'adénocarcinome De L'estomac." *Hépto-Gastro & Oncologie Digestive* 15, no. 2 (2008): 101-10.
- Lasseigne, Brittany N, Todd C Burwell, Mohini A Patil, Devin M Absher, James D Brooks, and Richard M Myers. "DNA Methylation Profiling Reveals Novel Diagnostic Biomarkers in Renal Cell Carcinoma." *BMC medicine* 12, no. 1 (2014): 235.
- Lauren, Pekka. "The Two Histological Main Types of Gastric Carcinoma: Diffuse and So-Called Intestinal-Type Carcinoma: An Attempt at a Histo-Clinical Classification." *Acta Pathologica Microbiologica Scandinavica* 64, no. 1 (1965): 31-49.
- Lawrence, Michael S, Petar Stojanov, Paz Polak, Gregory V Kryukov, Kristian Cibulskis, Andrey Sivachenko, Scott L Carter, *et al.* "Mutational Heterogeneity in Cancer and the Search for New Cancer-Associated Genes." *Nature* 499, no. 7457 (2013): 214.
- Leach, Fredrick S, Nicholas C Nicolaides, Nickolas Papadopoulos, Bo Liu, Jin Jen, Ramon Parsons, Päivi Peltomäki, *et al.* "Mutations of a Muts Homolog in Hereditary Nonpolyposis Colorectal Cancer." *Cell* 75, no. 6 (1993): 1215-25.
- Lee, Hye Seung, Seung Im Choi, Hyeon Kook Lee, Hee Sung Kim, Han-Kwang Yang, Gyeong Hoon Kang, Yong Il Kim, Byung Lan Lee, and Woo Ho Kim. "Distinct Clinical Features and Outcomes of Gastric Cancers with Microsatellite Instability." *Modern pathology* 15, no. 6 (2002): 632.
- Lee, Na Keum, Jung Hwa Lee, Chan Hyuk Park, Dayeon Yu, Yong Chan Lee, Jae-Ho Cheong, Sung Hoon Noh, and Sang Kil Lee. "Long Non-Coding Rna Hotair Promotes Carcinogenesis and Invasion of Gastric Adenocarcinoma." *Biochemical and biophysical research communications* 451, no. 2 (2014): 171-78.
- Lee, Ok-Jae, Regine Schneider-Stock, Patricia A McChesney, Doerthe Kuester, Albert Roessner, Michael Vieth, and Christopher A Moskaluk. "Hypermethylation, Loss of Expression of Glutathione Peroxidase-3 in Barrett's Tumorigenesis." *Neoplasia* 7, no. 9 (2005): 854-61.
- Lertratanangkoon, K., C. J. Wu, N. Savaraj, and M. L. Thomas. "Alterations of DNA Methylation by Glutathione Depletion." [In eng]. *Cancer Lett* 120, no. 2 (Dec 9 1997): 149-56.
- Lewis, Michael J, John P Wiebe, and J Godfrey Heathcote. "Expression of Progesterone Metabolizing Enzyme Genes (Akr1c1, Akr1c2, Akr1c3, Srd5a1, Srd5a2) Is Altered in Human Breast Carcinoma." *BMC cancer* 4, no. 1 (2004): 27.
- Li, Yibing, Ting-Ting Huang, Elaine J Carlson, Simon Melov, Philip C Ursell, Jean L Olson, Linda J Noble, *et al.* "Dilated Cardiomyopathy and Neonatal Lethality in Mutant Mice Lacking Manganese Superoxide Dismutase." *Nature genetics* 11, no. 4 (1995): 376.
- Li, Zhongdao, Alexander H Pearlman, and Peggy Hsieh. "DNA Mismatch Repair and the DNA Damage Response." *DNA repair* 38 (2016): 94-101.
- Lindblom, Annika, Pia Tannergård, Barbro Werelius, and Magnus Nordenskjöld. "Genetic Mapping of a Second Locus Predisposing to Hereditary Non-Polyposis Colon Cancer." *Nature genetics* 5, no. 3 (1993): 279-82.
- Lin-Hurtubise, Kevin M, Christopher G Yheulon, Ronald A Gagliano Jr, and Henry T Lynch. "Excess of Extracolonic Non-Endometrial Multiple Primary Cancers in Msh2 Germline Mutation Carriers over Mlh1." *Journal of surgical oncology* 108, no. 7 (2013): 433-37.

- Liu, Huan, Zhouping Yang, Linquan Zang, Guixiang Wang, Sigui Zhou, Guifang Jin, Zhicheng Yang, and Xuediao Pan. "Downregulation of Glutathione S-Transferase A1 Suppressed Tumor Growth and Induced Cell Apoptosis in A549 Cell Line." *Oncology letters* 16, no. 1 (2018): 467-74.
- Livak, Kenneth J, and Thomas D Schmittgen. "Analysis of Relative Gene Expression Data Using Real-Time Quantitative Pcr and the 2-  $\Delta\delta$ ct Method." *methods* 25, no. 4 (2001): 402-08.
- Machado, Ana Manuel Dantas, Ceu Figueiredo, Eliette Touati, Valdemar Máximo, Sonia Sousa, Valérie Michel, Fátima Carneiro, *et al.* "Helicobacter Pylori Infection Induces Genetic Instability of Nuclear and Mitochondrial DNA in Gastric Cells." *Clinical Cancer Research* 15, no. 9 (2009): 2995-3002.
- Mäkinen, MJ. "Colorectal Serrated Adenocarcinoma." *Histopathology* 50, no. 1 (2007): 131-50.
- Malesci, Alberto, Luigi Laghi, Paolo Bianchi, Gabriele Delconte, Ann Randolph, Valter Torri, Carlo Carnaghi, *et al.* "Reduced Likelihood of Metastases in Patients with Microsatellite-Unstable Colorectal Cancer." *Clinical Cancer Research* 13, no. 13 (2007): 3831-39.
- Manley, James L, and Adrian R Krainer. "A Rational Nomenclature for Serine/Arginine-Rich Protein Splicing Factors (Sr Proteins)." *Genes & development* 24, no. 11 (2010): 1073-74.
- Marrelli, Daniele, Alfonso De Stefano, Giovanni de Manzoni, Paolo Morgagni, Alberto Di Leo, and Franco Roviello. "Prediction of Recurrence after Radical Surgery for Gastric Cancer: A Scoring System Obtained from a Prospective Multicenter Study." *Annals of surgery* 241, no. 2 (2005): 247.
- Matic, Marija, Tatjana Pekmezovic, Tatjana Djukic, Jasmina Mimic-Oka, Dejan Dragicevic, Biljana Krivic, Sonja Suvakov, *et al.* "Gsta1, Gstm1, Gstp1, and Gstt1 Polymorphisms and Susceptibility to Smoking-Related Bladder Cancer: A Case-Control Study." Paper presented at the Urologic Oncology: Seminars and Original Investigations, 2013.
- Matlin, Arianne J, Francis Clark, and Christopher WJ Smith. "Understanding Alternative Splicing: Towards a Cellular Code." *Nature reviews Molecular cell biology* 6, no. 5 (2005): 386.
- Matsui, Yoko, Masafumi Inomata, Manabu Tojigamori, Kazuya Sonoda, Norio Shiraishi, and Seigo Kitano. "Suppression of Tumor Growth in Human Gastric Cancer with Her2 Overexpression by an Anti-Her2 Antibody in a Murine Model." *International journal of oncology* 27, no. 3 (2005): 681-85.
- Matthews, Lisa, Gopal Gopinath, Marc Gillespie, Michael Caudy, David Croft, Bernard de Bono, Phani Garapati, *et al.* "Reactome Knowledgebase of Human Biological Pathways and Processes." *Nucleic acids research* 37, no. suppl\_1 (2008): D619-D22.
- Meister, A., S. S. Tate, and O. W. Griffith. "Gamma-Glutamyl Transpeptidase." [In eng]. *Methods Enzymol* 77 (1981): 237-53.
- Mensenkamp, Arjen R, Ingrid P Vogelaar, Wendy AG van Zelst-Stams, Monique Goossens, Hicham Ouchene, Sandra JB Hendriks-Cornelissen, Michael P Kwint, *et al.* "Somatic Mutations in Mlh1 and Msh2 Are a Frequent Cause of Mismatch-Repair Deficiency in Lynch Syndrome-Like Tumors." *Gastroenterology* 146, no. 3 (2014): 643-46. e8.
- Miller, Valerie L, Hsueh-Kung Lin, Paari Murugan, Michael Fan, Trevor M Penning, Lacy S Brame, Qing Yang, and Kar-Ming Fung. "Aldo-Keto Reductase Family 1 Member C3 (Akr1c3) Is Expressed in Adenocarcinoma and Squamous Cell Carcinoma but Not Small Cell Carcinoma." *International journal of clinical and experimental pathology* 5, no. 4 (2012): 278.
- Ming, Si-Chun. "Gastric Carcinoma: A Pathobiological Classification." *Cancer* 39, no. 6 (1977): 2475-85.
- Mitra, Aparna, Lalita A Shevde, and Rajeev S Samant. "Multi-Faceted Role of Hsp40 in Cancer." *Clinical & experimental metastasis* 26, no. 6 (2009): 559-67.

- Morán, Alberto, Pilar Iniesta, Carmen de Juan, Rosa González-Quevedo, Andrés Sánchez-Pernaute, Eduardo Díaz-Rubio, Santiago Ramón y Cajal, *et al.* "Stromelysin-1 Promoter Mutations Impair Gelatinase B Activation in High Microsatellite Instability Sporadic Colorectal Tumors." *Cancer research* 62, no. 13 (2002): 3855-60.
- Nagarajan, Niranjana, Denis Bertrand, Axel M Hillmer, Zhi Jiang Zang, Fei Yao, Pierre-Étienne Jacques, Audrey SM Teo, *et al.* "Whole-Genome Reconstruction and Mutational Signatures in Gastric Cancer." *Genome biology* 13, no. 12 (2012): R115.
- Network, Cancer Genome Atlas Research. "Comprehensive Molecular Characterization of Gastric Adenocarcinoma." *Nature* 513, no. 7517 (2014): 202.
- Nicolaides, Nicholas C, Nickolas Papadopoulos, Bo Liu, Ying-Fei Weit, Kenneth C Carter, Steven M Ruben, Craig A Rosen, *et al.* "Mutations of Two P/Ws Homologues in Hereditary Nonpolyposis Colon Cancer." *Nature* 371, no. 6492 (1994): 75.
- Nilsen, Timothy W. "The Spliceosome: The Most Complex Macromolecular Machine in the Cell?". *Bioessays* 25, no. 12 (2003): 1147-49.
- Nilsen, Timothy W, and Brenton R Graveley. "Expansion of the Eukaryotic Proteome by Alternative Splicing." *Nature* 463, no. 7280 (2010): 457.
- Niwa, Tohru, Tetsuya Tsukamoto, Takeshi Toyoda, Akiko Mori, Harunari Tanaka, Takao Maekita, Masao Ichinose, Masae Tatematsu, and Toshikazu Ushijima. "Inflammatory Processes Triggered by Helicobacter Pylori Infection Cause Aberrant DNA Methylation in Gastric Epithelial Cells." *Cancer research* (2010): 0008-5472. CAN-09-2755.
- Nyfeler, B., V. Reiterer, M. W. Wendeler, E. Stefan, B. Zhang, S. W. Michnick, and H. P. Hauri. "Identification of Ergic-53 as an Intracellular Transport Receptor of Alpha1-Antitrypsin." [In eng]. *J Cell Biol* 180, no. 4 (Feb 25 2008): 705-12.
- Olschwang, Sylviane, Catherine Bonaïti, Josué Feingold, Thierry Frébourg, Sophie Grandjouan, Christine Lasset, Pierre Laurent-Puig, *et al.* "Identification Et Prise En Charge Du Syndrome Hnpcc (Hereditary Non Polyposis Colon Cancer), Prédisposition Héritaire Aux Cancers Du Côlon, Du Rectum Et De L'utérus." *Bulletin du cancer* 91, no. 4 (2004): 303-15.
- Oltean, Sebastian, and DO Bates. "Hallmarks of Alternative Splicing in Cancer." *Oncogene* 33, no. 46 (2014): 5311.
- Omenn, Gilbert S, Rajasree Menon, and Yang Zhang. "Innovations in Proteomic Profiling of Cancers: Alternative Splice Variants as a New Class of Cancer Biomarker Candidates and Bridging of Proteomics with Structural Biology." *Journal of proteomics* 90 (2013): 28-37.
- Ooi, Chia Huey, Tatiana Ivanova, Jeanie Wu, Minghui Lee, Iain Beehuat Tan, Jiong Tao, Lindsay Ward, *et al.* "Oncogenic Pathway Combinations Predict Clinical Prognosis in Gastric Cancer." *PLoS genetics* 5, no. 10 (2009): e1000676.
- Papa, Luena, Mary Hahn, Ellen L Marsh, Bradley S Evans, and Doris Germain. "Sod2 to Sod1 Switch in Breast Cancer." *Journal of biological chemistry* 289, no. 9 (2014): 5412-16.
- Park, Aubrey L, Hsueh-Kung Lin, Qing Yang, Chor Wing Sing, Michael Fan, Timothy B Mapstone, Naina L Gross, *et al.* "Differential Expression of Type 2 3 $\alpha$ /Type 5 17 $\beta$ -Hydroxysteroid Dehydrogenase (Akr1c3) in Tumors of the Central Nervous System." *International journal of clinical and experimental pathology* 3, no. 8 (2010): 743.
- Park, Dong Il, Seung Ha Park, Sang Hoon Kim, Jeong Wook Kim, Yong Kyun Cho, Hong Joo Kim, Chong Il Sohn, *et al.* "Effect of Helicobacter Pylori Infection on the Expression of DNA Mismatch Repair Protein." *Helicobacter* 10, no. 3 (2005): 179-84.

- Peltomaki, Paivi, Lauri A Aaltonen, Pertti Sistonen, Lea Pylkkanen, Jukka-Pekka Mecklin, Heikki Jarvinen, Jane S Green, JL Weber, and FS Leach. "Genetic Mapping of a Locus Predisposing to Human Colorectal Cancer." *Science* 260, no. 5109 (1993): 810-12.
- Pérez-Jiménez, Marga, Guillaume Besnard, Gabriel Dorado, and Pilar Hernandez. "Varietal Tracing of Virgin Olive Oils Based on Plastid DNA Variation Profiling." *PLoS One* 8, no. 8 (2013): e70507.
- Pettersson, Pär L, and Bengt Mannervik. "The Role of Glutathione in the Isomerization of  $\Delta^5$ -Androstene-3, 17-Dione Catalyzed by Human Glutathione Transferase A1-1." *Journal of Biological Chemistry* 276, no. 15 (2001): 11698-704.
- Pharoah, Paul DP, Parry Guilford, Carlos Caldas, and International Gastric Cancer Linkage Consortium. "Incidence of Gastric Cancer and Breast Cancer in Cdh1 (E-Cadherin) Mutation Carriers from Hereditary Diffuse Gastric Cancer Families." *Gastroenterology* 121, no. 6 (2001): 1348-53.
- Phumichai, Chalernpol, Thitaporn Phumichai, and Arunee Wongkaew. "Novel Chloroplast Microsatellite (Cpssr) Markers for Genetic Diversity Assessment of Cultivated and Wild Hevea Rubber." *Plant molecular biology reporter* 33, no. 5 (2015): 1486-98.
- Popat, S, R Hubner, and RS Houlston. "Systematic Review of Microsatellite Instability and Colorectal Cancer Prognosis." *Journal of clinical oncology* 23, no. 3 (2005): 609-18.
- Pospisil, Heike, Alexander Herrmann, Kristine Butherus, Stefan Pirson, Jens G Reich, and Wolfgang Kemmner. "Verification of Predicted Alternatively Spliced Wnt Genes Reveals Two New Splice Variants (Ctnnb1 and Lrp5) and Altered Axin-1 Expression During Tumour Progression." *BMC genomics* 7, no. 1 (2006): 148.
- Ramsioek, Dewkoemar, Anja Wagner, Monique E van Leerdam, Dennis Dooijes, Carli MJ Tops, Ewout W Steyerberg, and Ernst J Kuipers. "Cancer Risk in Mlh1, Msh2 and Msh6 Mutation Carriers; Different Risk Profiles May Influence Clinical Management." *Hereditary cancer in clinical practice* 7, no. 1 (2009): 17.
- Revil, Timothée, Lulzim Shkreta, and Benoît Chabot. "Épissage Alternatif Des Arn Prémessagers Dans Le Cancer: Impact Fonctionnel, Mécanismes Moléculaires Et Espoirs Thérapeutiques." *Bulletin du cancer* 93, no. 9 (2006): 909-19.
- Ribic, Christine M, Daniel J Sargent, Malcolm J Moore, Stephen N Thibodeau, Amy J French, Richard M Goldberg, Stanley R Hamilton, et al. "Tumor Microsatellite-Instability Status as a Predictor of Benefit from Fluorouracil-Based Adjuvant Chemotherapy for Colon Cancer." *New England Journal of Medicine* 349, no. 3 (2003): 247-57.
- Ringel, Israel, and Susan Band Horwitz. "Studies with Rp 56976 (Taxotere): A Semisynthetic Analogue of Taxol." *JNCl: Journal of the National Cancer Institute* 83, no. 4 (1991): 288-91.
- Riquelme, Ismael, Kathleen Saavedra, Jaime A Espinoza, Helga Weber, Patricia García, Bruno Nervi, Marcelo Garrido, et al. "Molecular Classification of Gastric Cancer: Towards a Pathway-Driven Targeted Therapy." *Oncotarget* 6, no. 28 (2015): 24750.
- Roeckel, Nina, Stefan M Woerner, Matthias Kloor, Yan-Ping Yuan, Georgios Patsos, Roland Gromes, Juergen Kopitz, and Johannes Gebert. "High Frequency of Lman1 Abnormalities in Colorectal Tumors with Microsatellite Instability." *Cancer research* 69, no. 1 (2009): 292-99.
- Rooseboom, Martijn, Jan NM Commandeur, and Nico PE Vermeulen. "Enzyme-Catalyzed Activation of Anticancer Prodrugs." *Pharmacological reviews* 56, no. 1 (2004): 53-102.
- Roth, Arnaud D, Mauro Delorenzi, Sabine Tejpar, Pu Yan, Dirk Klingbiel, Roberto Fiocca, Giovanni d'Ario, et al. "Integrated Analysis of Molecular and Clinical Prognostic Factors in Stage Ii/Iii Colon Cancer." *Journal of the National Cancer Institute* 104, no. 21 (2012): 1635-46.



- Rugge, Massimo, Matteo Fassan, and David Y Graham. "Epidemiology of Gastric Cancer." In *Gastric Cancer*, 23-34: Springer, 2015.
- Samowitz, Wade S, Karen Curtin, Khe-Ni Ma, Donna Schaffer, Landon W Coleman, Mark Leppert, and Martha L Slattery. "Microsatellite Instability in Sporadic Colon Cancer Is Associated with an Improved Prognosis at the Population Level." *Cancer Epidemiology and Prevention Biomarkers* 10, no. 9 (2001): 917-23.
- Sanchez, Gabriel, Danielle Bittencourt, Karine Laud, Jérôme Barbier, Olivier Delattre, Didier Auboeuf, and Martin Dutertre. "Alteration of Cyclin D1 Transcript Elongation by a Mutated Transcription Factor up-Regulates the Oncogenic D1b Splice Isoform in Cancer." *Proceedings of the National Academy of Sciences* 105, no. 16 (2008): 6004-09.
- Salwinski, Lukasz, Christopher S Miller, Adam J Smith, Frank K Pettit, James U Bowie, and David Eisenberg. "The Database of Interacting Proteins: 2004 Update." *Nucleic acids research* 32, no. suppl\_1 (2004): D449-D51.
- Sathasivam, Kirupa, Andreas Neueder, Theresa A Gipson, Christian Landles, Agnesska C Benjamin, Marie K Bondulich, Donna L Smith, *et al.* "Aberrant Splicing of Htt Generates the Pathogenic Exon 1 Protein in Huntington Disease." *Proceedings of the National Academy of Sciences* 110, no. 6 (2013): 2366-70.
- Schaefer, Carl F, Kira Anthony, Shiva Krupa, Jeffrey Buchoff, Matthew Day, Timo Hannay, and Kenneth H Buetow. "Pid: The Pathway Interaction Database." *Nucleic acids research* 37, no. suppl\_1 (2008): D674-D79.
- Schwerk, Christian, and Klaus Schulze-Osthoff. "Regulation of Apoptosis by Alternative Pre-Mrna Splicing." *Molecular cell* 19, no. 1 (2005): 1-13.
- Senter, Leigha, Mark Clendenning, Kaisa Sotamaa, Heather Hampel, Jane Green, John D Potter, Annika Lindblom, *et al.* "The Clinical Phenotype of Lynch Syndrome Due to Germ-Line Pms2 Mutations." *Gastroenterology* 135, no. 2 (2008): 419-28. e1.
- Seruga, Bostjan, Alberto Ocana, and Ian F Tannock. "Drug Resistance in Metastatic Castration-Resistant Prostate Cancer." *Nature reviews Clinical oncology* 8, no. 1 (2011): 12.
- Shah, Manish A, Raya Khanin, Laura H Tang, Yelena Y Janjigian, David S Klimstra, Hans Gerdes, and David P Kelsen. "Molecular Classification of Gastric Cancer: A New Paradigm." *Clinical cancer research* (2011): clincanres.2203.010.
- Shen, Haihong, and Michael R Green. "A Pathway of Sequential Arginine-Serine-Rich Domain-Splicing Signal Interactions During Mammalian Spliceosome Assembly." *Molecular cell* 16, no. 3 (2004): 363-73.
- Sherwood. Appareil digestif. Physiologie humaine. Ed De Boeck 2006; 15:465-509.
- Shokal, Upasana, and Prakash C Sharma. "Implication of Microsatellite Instability in Human Gastric Cancers." *The Indian journal of medical research* 135, no. 5 (2012): 599.
- Siewert, J Rüdiger, Marcus Feith, and Hubert J Stein. "Biologic and Clinical Variations of Adenocarcinoma at the Esophago-Gastric Junction: Relevance of a Topographic-Anatomic Subclassification." *Journal of surgical oncology* 90, no. 3 (2005): 139-46.
- Sinicrope, Frank A, Rafaela L Rego, Kevin C Halling, Nathan Foster, Daniel J Sargent, Betsy La Plant, Amy J French, *et al.* "Prognostic Impact of Microsatellite Instability and DNA Ploidy in Human Colon Carcinoma Patients." *Gastroenterology* 131, no. 3 (2006): 729-37.
- Smith, Barry, Michael Ashburner, Cornelius Rosse, Jonathan Bard, William Bug, Werner Ceusters, Louis J Goldberg, *et al.* "The Obo Foundry: Coordinated Evolution of Ontologies to Support Biomedical Data Integration." *Nature biotechnology* 25, no. 11 (2007): 1251.

- Sobin, Leslie H, and Carolyn C Compton. "Tnm Seventh Edition: What's New, What's Changed: Communication from the International Union against Cancer and the American Joint Committee on Cancer." *Cancer* 116, no. 22 (2010): 5336-39.
- Sreerama, Lakshmaiah, Rainer Grant, Bari Revee, Aaron Gross, and Daniel D Gregory. "Identification of Polymorphic Human Class-3 Aldehyde Dehydrogenases." *Enzymology and Molecular Biology of Carbonyl Metabolism* 13 (2007): 53.
- Stephens, David J, and Rainer Pepperkok. "The Many Ways to Cross the Plasma Membrane." *Proceedings of the National Academy of Sciences* 98, no. 8 (2001): 4295-98.
- Stigliano, Vittoria, Daniela Assisi, Maurizio Cosimelli, Raffaele Palmirotta, Diana Giannarelli, Marcella Mottotese, Lupe Sanchez Mete, Raffaello Mancini, and Vincenzo Casale. "Survival of Hereditary Non-Polyposis Colorectal Cancer Patients Compared with Sporadic Colorectal Cancer Patients." *Journal of Experimental & Clinical Cancer Research* 27, no. 1 (2008): 39.
- Stojic, Lovorka, Richard Brun, and Josef Jiricny. "Mismatch Repair and DNA Damage Signalling." *DNA repair* 3, no. 8-9 (2004): 1091-101.
- Sun, G-G, Y-D Wang, L-Q Chen, S-J Wang, G-L Liu, X-R Yu, Y-J Cheng, and Q Liu. "Novel Cancer Suppressor Gene for Esophageal Cancer: Manganese Superoxide Dismutase." *Diseases of the Esophagus* 24, no. 5 (2011): 346-53.
- Sunakawa, Yu, and Heinz-Josef Lenz. "Molecular Classification of Gastric Adenocarcinoma: Translating New Insights from the Cancer Genome Atlas Research Network." *Current treatment options in oncology* 16, no. 4 (2015): 17.
- Suraweera, Nirosha, Alex Duval, Maryline Reperant, Christelle Vaury, Daniela Furlan, Karen Leroy, Raquel Seruca, Barry Iacopetta, and Richard Hamelin. "Evaluation of Tumor Microsatellite Instability Using Five Quasimonomorphic Mononucleotide Repeats and Pentaplex Pcr." *Gastroenterology* 123, no. 6 (2002): 1804-11.
- Suter, Catherine M, David IK Martin, and Robyn L Ward. "Germline Epimutation of Mlh1 in Individuals with Multiple Cancers." *Nature genetics* 36, no. 5 (2004): 497.
- Suzuki, Shioto, Youko Muroishi, Isao Nakanishi, and Yoshio Oda. "Relationship between Genetic Polymorphisms of Drug-Metabolizing Enzymes (Cyp1a1, Cyp2e1, Gstm1, and Nat2), Drinking Habits, Histological Subtypes, and P53 Gene Point Mutations in Japanese Patients with Gastric Cancer." *Journal of gastroenterology* 39, no. 3 (2004): 220-30.
- Sweeney, Christopher J, Kathy D Miller, Sean E Sissons, Shinichi Nozaki, Douglas K Heilman, Jianzhao Shen, and George W Sledge. "The Antiangiogenic Property of Docetaxel Is Synergistic with a Recombinant Humanized Monoclonal Antibody against Vascular Endothelial Growth Factor or 2-Methoxyestradiol but Antagonized by Endothelial Growth Factors." *Cancer research* 61, no. 8 (2001): 3369-72.
- Taguwa, Shuhei, Kevin Maringer, Xiaokai Li, Dabeiba Bernal-Rubio, Jennifer N Rauch, Jason E Gestwicki, Raul Andino, Ana Fernandez-Sesma, and Judith Frydman. "Defining Hsp70 Subnetworks in Dengue Virus Replication Reveals Key Vulnerability in Flavivirus Infection." *Cell* 163, no. 5 (2015): 1108-23.
- Takahashi, Kazuhiko, Nelly Avissar, John Whitin, and Harvey Cohen. "Purification and Characterization of Human Plasma Glutathione Peroxidase: A Selenoglycoprotein Distinct from the Known Cellular Enzyme." *Archives of biochemistry and biophysics* 256, no. 2 (1987): 677-86.
- Tan, Patrick, and Khay-Guan Yeoh. "Genetics and Molecular Pathogenesis of Gastric Adenocarcinoma." *Gastroenterology* 149, no. 5 (2015): 1153-62. e3.
- Thibodeau, Stephen N, G Bren, and D Schaid. "Microsatellite Instability in Cancer of the Proximal Colon." *Science* 260, no. 5109 (1993): 816-19.

- Thorsen, Kasper, Francisco Mansilla, Troels Schepeler, Bodil Øster, Mads H Rasmussen, Lars Dyrskjøt, Rotem Karni, *et al.* "Alternative Splicing of Slc39a14 in Colorectal Cancer Is Regulated by the Wnt Pathway." *Molecular & Cellular Proteomics* 10, no. 1 (2011): M110. 002998.
- Todd, Peter K, Seok Yoon Oh, Amy Krans, Fang He, Chantal Sellier, Michelle Frazer, Abigail J Renoux, *et al.* "Cgg Repeat-Associated Translation Mediates Neurodegeneration in Fragile X Tremor Ataxia Syndrome." *Neuron* 78, no. 3 (2013): 440-55.
- Tsuneoka, Yutaka, Kouji Fukushima, Yoshinori Matsuo, Yoshiyuki Ichikawa, and Yasuhiro Watanabe. "Genotype Analysis of the Cyp2c19 Gene in the Japanese Population." *Life sciences* 59, no. 20 (1996): 1711-15.
- Tweedie, Susan, Michael Ashburner, Kathleen Falls, Paul Leyland, Peter McQuilton, Steven Marygold, Gillian Millburn, *et al.* "Flybase: Enhancing Drosophila Gene Ontology Annotations." *Nucleic acids research* 37, no. suppl\_1 (2008): D555-D59.
- Umar, Asad, C Richard Boland, Jonathan P Terdiman, Sapna Syngal, Albert de la Chapelle, Josef Rüschoff, Richard Fishel, *et al.* "Revised Bethesda Guidelines for Hereditary Nonpolyposis Colorectal Cancer (Lynch Syndrome) and Microsatellite Instability." *Journal of the National Cancer Institute* 96, no. 4 (2004): 261-68.
- Van Grieken, NCT, T Aoyama, PA Chambers, D Bottomley, LC Ward, I Inam, TE Buffart, *et al.* "Kras and Braf Mutations Are Rare and Related to DNA Mismatch Repair Deficiency in Gastric Cancer from the East and the West: Results from a Large International Multicentre Study." *British journal of cancer* 108, no. 7 (2013): 1495.
- Vasen, Hans FA, Patrice Watson, Jukka-Pekka Mecklin, and Henry T Lynch. "New Clinical Criteria for Hereditary Nonpolyposis Colorectal Cancer (Hnpcc, Lynch Syndrome) Proposed by the International Collaborative Group on Hnpcc." *Gastroenterology* 116, no. 6 (1999): 1453-56.
- Vasiliou, Vasilis, Amos Bairoch, Keith F Tipton, and Daniel W Nebert. "Eukaryotic Aldehyde Dehydrogenase (Aldh) Genes: Human Polymorphisms, and Recommended Nomenclature Based on Divergent Evolution and Chromosomal Mapping." *Pharmacogenetics* 9, no. 4 (1999): 421-34.
- Veigl, Martina L, Lakshmi Kasturi, Joseph Olechnowicz, AiHong Ma, James D Lutterbaugh, Sumudra Periyasamy, Guo-Min Li, *et al.* "Biallelic Inactivation of Hmlh1 by Epigenetic Gene Silencing, a Novel Mechanism Causing Human Msi Cancers." *Proceedings of the National Academy of Sciences* 95, no. 15 (1998): 8698-702.
- Velho, Sérgio, Maria Sofia Fernandes, Marina Leite, Ceu Figueiredo, and Raquel Seruca. "Causes and Consequences of Microsatellite Instability in Gastric Carcinogenesis." *World Journal of Gastroenterology: WJG* 20, no. 44 (2014): 16433.
- Venderbosch, Sabine, Iris D Nagtegaal, Timothy S Maughan, Christopher G Smith, Jeremy P Cheadle, David Fisher, Richard Kaplan, *et al.* "Mismatch Repair Status and Braf Mutation Status in Metastatic Colorectal Cancer Patients: A Pooled Analysis of the Cairo, Cairo2, Coin and Focus Studies." *Clinical Cancer Research* (2014): clincanres. 0332.2014.
- Vieira, Maria Lucia Carneiro, Luciane Santini, Augusto Lima Diniz, and Carla de Freitas Munhoz. "Microsatellite Markers: What They Mean and Why They Are So Useful." *Genetics and molecular biology* 39, no. 3 (2016): 312-28.
- Wagner, Anna D, Wilfried Grothe, Johannes Haerting, Gerhard Kleber, Axel Grothey, and Wolfgang E Fleig. "Chemotherapy in Advanced Gastric Cancer: A Systematic Review and Meta-Analysis Based on Aggregate Data." *J Clin Oncol* 24, no. 18 (2006): 2903-09.
- Wang, Eric T, Rickard Sandberg, Shujun Luo, Irina Khrebtukova, Lu Zhang, Christine Mayr, Stephen F Kingsmore, Gary P Schroth, and Christopher B Burge. "Alternative Isoform Regulation in Human Tissue Transcriptomes." *Nature* 456, no. 7221 (2008): 470.
- Wang, Kai, Junsuo Kan, Siu Tsan Yuen, Stephanie T Shi, Kent Man Chu, Simon Law, Tsun Leung Chan, *et al.* "Exome Sequencing Identifies Frequent Mutation of Arid1a in Molecular Subtypes of Gastric Cancer." *Nature*

- genetics* 43, no. 12 (2011): 1219.
- Wang, Yang, and Zefeng Wang. "Efficient Backsplicing Produces Translatable Circular Mrnas." *Rna* 21, no. 2 (2015): 172-79.
- WARTHIN, ALDRED SCOTT. "Heredity with Reference to Carcinoma: As Shown by the Study of the Cases Examined in the Pathological Laboratory of the University of Michigan, 1895-1913." *Archives of Internal Medicine* 12, no. 5 (1913): 546-55.
- Watson, James D, and Francis HC Crick. "Molecular Structure of Nucleic Acids." *Nature* 171, no. 4356 (1953): 737-38.
- Watson, Sarah, and Christophe Louvet. "Her2 in Esogastric Cancer: Clinical Implications." *Hépto-Gastro & Oncologie Digestive* 21, no. 7 (2014): 502-08.
- Whitlock Jr, JAMES P, Steven T Okino, Liqun Dong, Hyunsung P Ko, Regina Clarke-Katzenberg, Qiang Ma, and Hui Li. "Cytochromes P450 5: Induction of Cytochrome P4501a1: A Model for Analyzing Mammalian Gene Transcription." *The FASEB journal* 10, no. 8 (1996): 809-18.
- Will, Cindy L, and Reinhard Lührmann. "Spliceosome Structure and Function." *Cold Spring Harbor perspectives in biology* 3, no. 7 (2011): a003707.
- Woerner, Stefan M, Matthias Kloor, Annegret Mueller, Josef Rueschoff, Nicolaus Friedrichs, Reinhard Buettner, Moriz Buzello, et al. "Microsatellite Instability of Selective Target Genes in Hnpcc-Associated Colon Adenomas." *Oncogene* 24, no. 15 (2005): 2525.
- Wu, Jianming, Tuo Liu, Zechary Rios, Qibing Mei, Xiukun Lin, and Shousong Cao. "Heat Shock Proteins and Cancer." *Trends in pharmacological sciences* 38, no. 3 (2017): 226-56.
- Xiao, GH, JA Pinaire, AD Rodrigues, and RA Prough. "Regulation of the Ah Gene Battery Via Ah Receptor-Dependent and Independent Processes in Cultured Adult Rat Hepatocytes." *Drug metabolism and disposition* 23, no. 6 (1995): 642-50.
- Xie, HG, RB Kim, CM Stein, GR Wilkinson, and AJJ Wood. "Genetic Polymorphism of (S)-Mephenytoin 4'-Hydroxylation in Populations of African Descent." *British journal of clinical pharmacology* 48, no. 3 (1999): 402.
- Xie, Hong-Guang, Harish C Prasad, Richard B Kim, and C Michael Stein. "Cyp2c9 Allelic Variants: Ethnic Distribution and Functional Significance." *Advanced drug delivery reviews* 54, no. 10 (2002): 1257-70.
- Yan, P Yu, Guoying Yu, George Tseng, Kathleen Cieply, Joel Nelson, Marie Defrances, Reza Zarnegar, George Michalopoulos, and Jian-Hua Luo. "Glutathione Peroxidase 3, Deleted or Methylated in Prostate Cancer, Suppresses Prostate Cancer Growth and Metastasis." *Cancer research* 67, no. 17 (2007): 8043-50.
- Yao, Yuan, Hong Tao, Dong Il Park, Jorge L Sepulveda, and Antonia R Sepulveda. "Demonstration and Characterization of Mutations Induced by Helicobacter Pylori Organisms in Gastric Epithelial Cells." *Helicobacter* 11, no. 4 (2006): 272-86.
- Yokose, Tomoyuki, Mikio Doy, Motoharu Kakiki, Tohru Horie, Yasushi Matsuzaki, and Kiyoshi Mukai. "Expression of Cytochrome P450 3a4 in Foveolar Epithelium with Intestinal Metaplasia of the Human Stomach." *Japanese journal of cancer research* 89, no. 10 (1998): 1028-32.
- Zakharov, Vladislav, Hsueh-Kung Lin, Joseph Azzarello, Scott McMeekin, Kathleen N Moore, Trevor M Penning, and Kar-Ming Fung. "Suppressed Expression of Type 2 3 $\alpha$ /Type 5 17 $\beta$ -Hydroxysteroid Dehydrogenase (Akr1c3) in Endometrial Hyperplasia and Carcinoma." *International journal of clinical and experimental pathology* 3, no. 6 (2010): 608.

- Zhang, Fengchun, Caili Song, Yue Ma, Lei Tang, Yingchun Xu, and Hongxia Wang. "Effect of Fibroblasts on Breast Cancer Cell Mammosphere Formation and Regulation of Stem Cell-Related Gene Expression." *International journal of molecular medicine* 28, no. 3 (2011): 365-71.
- Zhang, Yiqiang, Yuji Ikeno, Wenbo Qi, Asish Chaudhuri, Yan Li, Alex Bokov, Suzanne R Thorpe, *et al.* "Mice Deficient in Both Mn Superoxide Dismutase and Glutathione Peroxidase-1 Have Increased Oxidative Damage and a Greater Incidence of Pathology but No Reduction in Longevity." *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences* 64, no. 12 (2009): 1212-20.
- Zoppino, Felipe CM, Martin E Guerrero-Gimenez, Gisela N Castro, and Daniel R Ciocca. "Comprehensive Transcriptomic Analysis of Heat Shock Proteins in the Molecular Subtypes of Human Breast Cancer." *BMC cancer* 18, no. 1 (2018): 700

# Annexes

---

# Annexe 1

## Données supplémentaires :

### Matériels et méthodes

Création d'un réseau d'interaction à l'aide de la base de données STRING (version 10.5) à partir de la liste des gènes communément dérégulés à la baisse dans le cancer gastrique.

Utilisation des données de « S5.7a\_mRNA\_differential\_expression\_20140128 » publiées dans Comprehensive Molecular Characterization of Gastric Adenocarcinoma, Nature, 2014 ([https://tcga-data.nci.nih.gov/docs/publications/stad\\_2014/](https://tcga-data.nci.nih.gov/docs/publications/stad_2014/)).

Pour chacune des 4 catégories, les gènes différentiellement exprimés ont été identifiés après une analyse non appariée à deux classes en utilisant SAMseq avec un seuil FDR de 0,05. Les résultats ont été filtrés en éliminant les gènes dont l'expression médiane était inférieure à 5 RPKM. Les gènes pour lesquels la p-value était inférieure à 0.05 après un test de Wilcoxon ont été sélectionnés.

Nous avons ensuite sélectionné les gènes communément dérégulés au sein d'au moins 3 des 4 catégories : tumeurs infectées par le virus d'Épstein Barr (EBV) vs normal, tumeurs avec une instabilité des microsatellites (MSI) vs normal, génomiquement stables (GS) vs normal et tumeurs chromosomiquement instables (CIN) vs normal.  $W.pval.unpaired$  : p-value après un test de Wilcoxon unpaired test et  $fold.change : (median.group2 + 1)/(median.group1 + 1)$ .

| Normal vs CIN DOWN |                     |             | Normal vs GS DOWN |                     |             | Normal vs EBV DOWN |                     |             | Normal vs MSI DOWN |                     |             |
|--------------------|---------------------|-------------|-------------------|---------------------|-------------|--------------------|---------------------|-------------|--------------------|---------------------|-------------|
| GENE               | W.pval.unp<br>aired | fold.change | GENE              | W.pval.unp<br>aired | fold.change | GENE               | W.pval.unp<br>aired | fold.change | GENE               | W.pval.unp<br>aired | fold.change |
| GKN1               | 4,91669E-07         | 0,000504116 | GKN1              | 4,72157E-05         | 0,000477722 | GKN1               | 6,46314E-05         | 0,000422071 | GKN1               | 8,23758E-06         | 0,000397974 |
| GKN2               | 4,27411E-06         | 0,002344486 | GKN2              | 0,000107            | 0,002272898 | GKN2               | 0,000260844         | 0,002094913 | GKN2               | 3,66675E-05         | 0,002061887 |
| LIPF               | 0,000157154         | 0,003787593 | LIPF              | 0,000857539         | 0,003847375 | REG3A              | 0,002511799         | 0,003205027 | PGC                | 0,001432544         | 0,002931813 |
| GIF                | 0,001343386         | 0,010047842 | PGC               | 0,012170092         | 0,004370425 | TFF1               | 0,003137997         | 0,004277786 | LIPF               | 8,05742E-05         | 0,003780346 |
| PGA3               | 6,48935E-07         | 0,010066379 | TFF2              | 0,0044907           | 0,006054496 | LIPF               | 0,00318903          | 0,004428473 | REG3A              | 0,001596409         | 0,005700338 |
| REG3A              | 0,000910699         | 0,01223972  | GIF               | 0,001336014         | 0,007721355 | TFF2               | 0,000955498         | 0,004464487 | GIF                | 0,000286947         | 0,007843191 |
| CHGA               | 3,68371E-06         | 0,015938308 | REG3A             | 0,002418327         | 0,009654356 | PSCA               | 2,63489E-06         | 0,006409462 | PGA3               | 2,68701E-07         | 0,009855006 |
| IGJ                | 7,67881E-05         | 0,018454489 | PGA3              | 1,59417E-05         | 0,010115828 | PGC                | 0,013844005         | 0,006442805 | CHGA               | 1,05826E-07         | 0,010758835 |
| KRT20              | 0,005889945         | 0,030392244 | PSCA              | 6,6887E-06          | 0,01625443  | GIF                | 0,011694076         | 0,008927749 | CXCL17             | 1,42814E-07         | 0,011865944 |
| PGC                | 0,018970839         | 0,032415731 | CXCL17            | 2,18125E-05         | 0,016413965 | PGA3               | 2,31005E-05         | 0,009798769 | KRT20              | 0,000827688         | 0,015907681 |
| CXCL17             | 1,03482E-05         | 0,033691859 | TFF1              | 0,002583519         | 0,02397414  | CHGA               | 1,11841E-06         | 0,010238693 | TFF2               | 0,00307994          | 0,023293048 |
| TFF2               | 0,001340801         | 0,033919764 | IGJ               | 0,00149448          | 0,027491925 | KRT20              | 0,001661284         | 0,013227192 | IGJ                | 0,000467721         | 0,026791229 |
| PSCA               | 3,85003E-06         | 0,041328043 | AX747544          | 2,63957E-05         | 0,030511234 | AX747544           | 1,02603E-05         | 0,022365571 | SST                | 1,22334E-06         | 0,033660672 |
| TFF1               | 0,000868147         | 0,043755505 | CHGA              | 3,18795E-05         | 0,033330152 | MT1G               | 4,4418E-09          | 0,024698627 | TFF1               | 0,004126871         | 0,038502707 |
| SST                | 3,52032E-06         | 0,044844421 | AKR1B10           | 2,32348E-06         | 0,047381297 | DPCR1              | 1,02603E-05         | 0,026663155 | PSCA               | 1,06799E-05         | 0,041249747 |
|                    |                     |             |                   |                     |             |                    |                     |             |                    |                     |             |



| Normal vs CIN DOWN   |                  |             | Normal vs GS DOWN    |                  |             | Normal vs EBV DOWN |                  |             | Normal vs MSI DOWN   |                  |             |
|----------------------|------------------|-------------|----------------------|------------------|-------------|--------------------|------------------|-------------|----------------------|------------------|-------------|
| GENE                 | W.pval.unp aired | fold.change |                      | W.pval.unp aired | fold.change |                    | W.pval.unp aired | fold.change |                      | W.pval.unp aired | fold.change |
| AKR1B10              | 2,33415E-07      | 0,046873147 | KCNE2                | 0,000934731      | 0,057107885 | SST                | 5,36719E-06      | 0,028742179 | AKR1B10              | 1,07203E-06      | 0,046078667 |
| KCNE2                | 2,61355E-05      | 0,05913674  | KRT20                | 0,016300757      | 0,059129731 | FAM3B              | 4,64677E-07      | 0,048253645 | GSTA1                | 1,54846E-06      | 0,04848937  |
| DPCR1                | 0,000126338      | 0,060689068 | SLC5A5               | 0,000498683      | 0,062189739 | KCNE2              | 0,00034236       | 0,049119153 | KCNE2                | 7,28659E-06      | 0,04928869  |
| C6orf58              | 0,001366253      | 0,06186147  | C6orf58              | 0,011518757      | 0,064070212 | MT1M               | 3,56751E-13      | 0,051585472 | ALDH3A1              | 9,3267E-07       | 0,05905667  |
| Ig_alpha_1-[alpha]2m | 2,92879E-05      | 0,063475157 | SST                  | 0,000533583      | 0,075007716 | GGT6               | 2,03138E-05      | 0,05382285  | AX747544             | 2,54131E-05      | 0,059566727 |
| AX747544             | 2,00812E-05      | 0,063741089 | ALDH3A1              | 2,6961E-05       | 0,076474007 | SLC5A5             | 0,000829864      | 0,060072666 | SLC5A5               | 7,07786E-05      | 0,062299243 |
| MAL                  | 3,7545E-15       | 0,065767188 | SOX21                | 2,65229E-06      | 0,078904092 | C6orf58            | 0,002751761      | 0,060204218 | Ig_alpha_1-[alpha]2m | 0,000138584      | 0,06313774  |
| ALDH3A1              | 4,16613E-07      | 0,066869799 | DPCR1                | 0,001224718      | 0,079758492 | FCGBP              | 0,004261405      | 0,062040535 | C6orf58              | 0,010124702      | 0,065214969 |
| MYOC                 | 4,15289E-09      | 0,069906506 | CPA2                 | 0,000502076      | 0,095419777 | DUOXA2             | 8,14682E-07      | 0,062317592 | MYOC                 | 4,42293E-10      | 0,06701554  |
| SLC5A5               | 0,000618222      | 0,071228015 | FCGBP                | 0,005526871      | 0,100338634 | CKB                | 5,3564E-12       | 0,06523001  | MAL                  | 1,37098E-13      | 0,067910881 |
| DPT                  | 1,75416E-13      | 0,081072546 | PSAPL1               | 0,00193907       | 0,116084956 | FAM3D              | 1,84617E-05      | 0,066140147 | SOX21                | 7,90225E-08      | 0,072533435 |
| FCGBP                | 0,00044421       | 0,090781244 | Ig_alpha_1-[alpha]2m | 0,004096364      | 0,118829933 | IGJ                | 0,012463503      | 0,067701038 | ADH1C                | 0,001839856      | 0,073854034 |
| SOX21                | 1,7171E-06       | 0,099632449 | SCGB2A1              | 0,000966508      | 0,122408323 | MYOC               | 9,5347E-06       | 0,070630843 | VSIG2                | 2,54131E-05      | 0,074965683 |
| C16orf89             | 5,3079E-12       | 0,101924054 | ADTRP                | 0,0044907        | 0,132116742 | AKR1B10            | 0,00034236       | 0,07327521  | ADTRP                | 0,000467721      | 0,075025214 |
| CPA2                 | 0,000628826      | 0,105097788 | SOSTDC1              | 7,47275E-10      | 0,134041334 | AKR1C2             | 9,4676E-10       | 0,075726836 | FAM3B                | 2,16036E-05      | 0,079989219 |
| MT1M                 | 5,3079E-12       | 0,106843328 | B3GNT6               | 1,95918E-05      | 0,138887934 | HPGD               | 1,81471E-08      | 0,07613152  | CXCL14               | 3,62464E-08      | 0,080701613 |
|                      |                  |             |                      |                  |             |                    |                  |             |                      |                  |             |

| Normal vs CIN DOWN |                 |             | Normal vs GS DOWN |                 |             | Normal vs EBV DOWN   |                 |             | Normal vs MSI DOWN |                 |             |
|--------------------|-----------------|-------------|-------------------|-----------------|-------------|----------------------|-----------------|-------------|--------------------|-----------------|-------------|
| GENE               | W.pval.unpaired | fold.change |                   | W.pval.unpaired | fold.change |                      | W.pval.unpaired | fold.change |                    | W.pval.unpaired | fold.change |
| LGALS9C            | 2,12882E-08     | 0,114400939 | VSIG2             | 0,002583519     | 0,141285717 | HBB                  | 3,76849E-09     | 0,077970528 | MT1M               | 1,24313E-09     | 0,081787122 |
| SCGB2A1            | 0,000100122     | 0,114929551 | GPT               | 0,000616878     | 0,147382153 | C16orf89             | 6,56789E-10     | 0,078731266 | C16orf89           | 3,75269E-11     | 0,087118812 |
| VSIG2              | 0,000174198     | 0,117451884 | SLC9A4            | 0,000191173     | 0,148739518 | Ig_alpha_1-[alpha]2m | 0,007138771     | 0,083545405 | DPT                | 1,12515E-10     | 0,087746767 |
| SOSTDC1            | 1,2129E-11      | 0,120988061 | ADH1B             | 0,017171783     | 0,150129499 | EEF1A2               | 1,13566E-08     | 0,085829064 | SOSTDC1            | 1,60941E-10     | 0,088186121 |
| CFD                | 2,38348E-13     | 0,12277818  | REP15             | 0,00227634      | 0,153994835 | DPT                  | 2,69843E-09     | 0,086017839 | CLEC3B             | 3,5624E-13      | 0,088666818 |
| PSAPL1             | 0,000369759     | 0,12728429  | FAM3B             | 0,000258881     | 0,154996492 | SOSTDC1              | 7,89454E-10     | 0,087538559 | AKR1C2             | 1,11156E-12     | 0,092017444 |
| ADH1C              | 0,000526759     | 0,128941788 | SCNN1B            | 8,95207E-10     | 0,157636739 | MAL                  | 3,07986E-14     | 0,087987547 | CPA2               | 5,0594E-05      | 0,094169377 |
| HPGD               | 1,56006E-11     | 0,131860007 | LGALS9C           | 4,76177E-06     | 0,159554057 | AKR7A3               | 0,017887967     | 0,091882079 | EEF1A2             | 3,8175E-08      | 0,097708139 |
| AKR1C2             | 4,74166E-10     | 0,138376943 | CAPN9             | 0,011845132     | 0,16009115  | CPA2                 | 0,000629085     | 0,094123574 | CKB                | 1,2984E-11      | 0,098357509 |
| CLEC3B             | 2,30385E-13     | 0,141703535 | AKR1C2            | 3,66598E-08     | 0,161446088 | SCGB2A1              | 2,3258E-05      | 0,09797518  | PSAPL1             | 0,000267243     | 0,107298506 |
| B3GNT6             | 5,39855E-07     | 0,14254288  | MYOC              | 0,00063875      | 0,162123871 | GPT                  | 7,91455E-05     | 0,100025488 | SCGB2A1            | 3,98162E-05     | 0,109685546 |
| MT1G               | 3,68449E-06     | 0,145411066 | GSTA1             | 0,00250336      | 0,162479855 | CAPN9                | 0,000317238     | 0,103758734 | HBB                | 2,03841E-08     | 0,119943039 |
| HBB                | 3,35745E-09     | 0,146083048 | DUOXA2            | 0,000387803     | 0,164003071 | CLEC3B               | 9,0838E-11      | 0,107600836 | B3GNT6             | 5,74677E-07     | 0,123523413 |
| SCNN1B             | 1,10396E-13     | 0,151843513 | AKR1B15           | 8,40058E-06     | 0,164945034 | SCIN                 | 6,61784E-08     | 0,112174161 | HPGD               | 1,84796E-09     | 0,12727219  |
| ADTRP              | 0,002846191     | 0,152365944 | HBB               | 1,92405E-06     | 0,166713527 | PSAPL1               | 0,011168433     | 0,11415839  | LGALS9C            | 1,73485E-06     | 0,133475523 |
| GPT                | 6,9794E-05      | 0,153395219 | RDH12             | 4,05733E-06     | 0,168394996 | HBA2                 | 2,44138E-08     | 0,118098279 | REP15              | 0,00050081      | 0,133761617 |
|                    |                 |             |                   |                 |             |                      |                 |             |                    |                 |             |

| Normal vs CIN DOWN |                 |             | Normal vs GS DOWN |                 |             | Normal vs EBV DOWN |                 |             | Normal vs MSI DOWN |                 |             |
|--------------------|-----------------|-------------|-------------------|-----------------|-------------|--------------------|-----------------|-------------|--------------------|-----------------|-------------|
| GENE               | W.pval.unpaired | fold.change | GENE              | W.pval.unpaired | fold.change | GENE               | W.pval.unpaired | fold.change | GENE               | W.pval.unpaired | fold.change |
| AKR1B15            | 3,52186E-07     | 0,165466262 | EEF1A2            | 2,43817E-07     | 0,172093589 | VSIG2              | 0,003775052     | 0,121765691 | GPX3               | 5,3375E-13      | 0,134638945 |
| EEF1A2             | 8,73994E-06     | 0,166856282 | HPGD              | 1,36158E-08     | 0,175487402 | PDIA2              | 1,7231E-07      | 0,125225925 | SCNN1B             | 9,10676E-13     | 0,137606717 |
| GSTA1              | 0,002203574     | 0,167115903 | MAL               | 5,19775E-10     | 0,176838298 | HBA1               | 1,34139E-08     | 0,126229128 | AADAC              | 0,000193749     | 0,138704814 |
| SLC9A4             | 5,37547E-05     | 0,169038791 | MT1M              | 1,41841E-05     | 0,183761305 | HOMER2             | 2,56655E-15     | 0,127699169 | SLC9A4             | 3,70627E-06     | 0,142261027 |
| FAM3D              | 6,34541E-06     | 0,174653327 | CYP2C19           | 0,000278836     | 0,18897352  | LDHD               | 3,01031E-11     | 0,129108145 | CAPN9              | 0,001624489     | 0,142857143 |
| REP15              | 0,000748679     | 0,183091295 | ID1               | 1,44266E-09     | 0,196904293 | LGALS9C            | 3,38947E-05     | 0,129340213 | FCGBP              | 0,01549078      | 0,145719102 |
| RDH12              | 6,31273E-08     | 0,184143975 | BHLHA15           | 0,001446018     | 0,198814582 | B3GNT6             | 0,000306766     | 0,130426563 | RDH12              | 2,92039E-10     | 0,148765639 |
| RNASE1             | 4,89572E-06     | 0,184814643 | AKR7A3            | 0,012843797     | 0,199817988 | AKR1C3             | 1,02603E-05     | 0,132882593 | ADH1B              | 2,04723E-10     | 0,153549697 |
| ADH1B              | 3,72444E-11     | 0,186905196 | LGALS9B           | 7,12702E-06     | 0,20400827  | REP15              | 0,002598779     | 0,139305128 | SCARA5             | 1,76373E-12     | 0,159638593 |
| LGALS9B            | 1,68349E-07     | 0,187731424 | LDHD              | 8,49615E-07     | 0,220092172 | ADH1B              | 3,63142E-07     | 0,140135996 | CYP2C19            | 3,50313E-05     | 0,163267344 |
| GPX3               | 2,88902E-14     | 0,199990532 | PDIA2             | 0,003213655     | 0,234266382 | GSTA1              | 0,012463503     | 0,141433779 | DPCR1              | 0,000573634     | 0,164205323 |
| CYP2C19            | 1,74387E-06     | 0,200944128 | DUOXA1            | 2,47764E-05     | 0,235733999 | SOX21              | 0,010886578     | 0,143408319 | HOMER2             | 1,71219E-10     | 0,164783183 |
| CAPN9              | 0,002288894     | 0,201026169 | RNASE1            | 0,000516907     | 0,238402322 | SCNN1B             | 9,0838E-11      | 0,151720377 | LGALS9B            | 1,01214E-05     | 0,17361454  |
| GGT6               | 0,000554178     | 0,204901637 | HOMER2            | 1,53467E-07     | 0,245532141 | RDH12              | 7,52365E-07     | 0,153709775 | RNASE1             | 5,38801E-06     | 0,17661013  |
| HBA2               | 3,96232E-09     | 0,205983733 | AKR1C3            | 0,000347692     | 0,251055053 | SCARA5             | 6,96562E-12     | 0,156318756 | AKR1C3             | 3,63384E-06     | 0,176842874 |
| HBA1               | 2,52656E-08     | 0,206692801 | MT1G              | 0,003524284     | 0,253181461 | GPX3               | 5,00478E-13     | 0,159266743 | AKR1B15            | 0,000146992     | 0,178386404 |
|                    |                 |             |                   |                 |             |                    |                 |             |                    |                 |             |

| Normal vs CIN DOWN |                 |             | Normal vs GS DOWN |                 |             | Normal vs EBV DOWN |                 |             | Normal vs MSI DOWN |                 |             |
|--------------------|-----------------|-------------|-------------------|-----------------|-------------|--------------------|-----------------|-------------|--------------------|-----------------|-------------|
| GENE               | W.pval.unpaired | fold.change | GENE              | W.pval.unpaired | fold.change | GENE               | W.pval.unpaired | fold.change | GENE               | W.pval.unpaired | fold.change |
| SCARA5             | 4,8208E-12      | 0,206923488 | SCIN              | 9,93849E-05     | 0,260145813 | CFD                | 7,33598E-11     | 0,159629214 | SCIN               | 2,43488E-06     | 0,181313131 |
| AKR7A3             | 0,006183663     | 0,212661676 | EYA2              | 0,001145306     | 0,269440493 | ALDH3A1            | 0,004011674     | 0,164543987 | FAM177B            | 0,00189765      | 0,19670765  |
| AADAC              | 0,00029791      | 0,215532861 | KCNJ15            | 0,001399009     | 0,269446428 | AKR1B15            | 0,000522505     | 0,170804903 | PDIA2              | 0,00067853      | 0,197423302 |
| PDIA2              | 0,000602859     | 0,223201872 | DPT               | 0,000125676     | 0,290745257 | PTGR1              | 3,7301E-10      | 0,174229627 | EYA2               | 3,98264E-07     | 0,198820531 |
| CKB                | 2,33575E-08     | 0,239377605 | HBA2              | 2,81208E-05     | 0,292505248 | FAM177B            | 0,004261405     | 0,175874622 | CFD                | 3,41787E-09     | 0,205570189 |
| LDHD               | 3,1126E-08      | 0,241874706 | MT1X              | 2,58467E-05     | 0,300109815 | DUOXA1             | 1,15077E-08     | 0,179463238 | DUOXA2             | 0,003172498     | 0,209465874 |
| GCNT4              | 2,34134E-11     | 0,243690948 | HBA1              | 4,8237E-05      | 0,300346474 | CXCL14             | 1,84617E-05     | 0,184575857 | HBA2               | 4,93922E-06     | 0,229923695 |
| MT1X               | 1,53497E-09     | 0,245956004 | FAM3D             | 0,004627054     | 0,301273768 | ADH1C              | 0,014582602     | 0,188311042 | GCNT4              | 9,89354E-10     | 0,23052423  |
| DNASE1L3           | 5,81541E-09     | 0,249410528 | GGT6              | 0,028380512     | 0,309758122 | SLC9A4             | 0,00545913      | 0,188341373 | GPT                | 0,008071487     | 0,23134038  |
| LYVE1              | 4,42318E-09     | 0,250187936 | AADAC             | 0,011685117     | 0,312532468 | LGALS9B            | 2,66969E-05     | 0,189945384 | DNASE1L3           | 2,73613E-09     | 0,231630767 |
| FAM177B            | 0,003473424     | 0,253534305 | FAM177B           | 0,01903707      | 0,317254134 | RNASE1             | 0,000829864     | 0,191585912 | UGT1A1             | 0,001007249     | 0,231683438 |
| KCNJ15             | 0,00016797      | 0,271268368 | C16orf89          | 0,00010751      | 0,319385457 | LYVE1              | 3,19177E-09     | 0,214038542 | BHLHA15            | 0,001007249     | 0,234330387 |
| BHLHA15            | 0,015563494     | 0,313766145 | GPX3              | 4,65486E-06     | 0,330653537 | MT1X               | 7,20055E-09     | 0,242942072 | MT1G               | 0,00067853      | 0,24786884  |
| CXCL14             | 0,000582933     | 0,317009573 | GCNT4             | 6,02327E-07     | 0,336764706 | GDPD3              | 2,63489E-06     | 0,248325328 | DUSP5              | 1,85772E-06     | 0,249438679 |
| FAM3B              | 1,92776E-05     | 0,317736498 | TST               | 0,001266303     | 0,341366678 | BHLHA15            | 0,015355038     | 0,278619986 | LYVE1              | 1,34992E-06     | 0,257442873 |
| DUOXA2             | 0,004046056     | 0,327086044 | TMEM92            | 0,014674433     | 0,353806917 | CIDEA              | 5,5633E-06      | 0,281270337 | HBA1               | 1,43142E-05     | 0,264950835 |
|                    |                 |             |                   |                 |             |                    |                 |             |                    |                 |             |

| Normal vs CIN DOWN |                 |             | Normal vs GS DOWN |                 |             | Normal vs EBV DOWN |                 |             | Normal vs MSI DOWN |                 |             |
|--------------------|-----------------|-------------|-------------------|-----------------|-------------|--------------------|-----------------|-------------|--------------------|-----------------|-------------|
| GENE               | W.pval.unpaired | fold.change | GENE              | W.pval.unpaired | fold.change | GENE               | W.pval.unpaired | fold.change | GENE               | W.pval.unpaired | fold.change |
| SCIN               | 5,53858E-05     | 0,32842131  | FAM46C            | 0,004767146     | 0,356158306 | CBR1               | 9,4676E-10      | 0,285361374 | GGT6               | 0,013304779     | 0,266278912 |
| EYA2               | 0,009924749     | 0,328917734 | KLF4              | 1,90026E-08     | 0,359486955 | SELENBP1           | 6,84183E-06     | 0,289833158 | KCNJ15             | 0,000223883     | 0,281664584 |
| AKR1C3             | 0,000421905     | 0,332317465 | CBR1              | 9,09871E-08     | 0,360250652 | DUSP5              | 0,000111249     | 0,293720422 | FAM3D              | 0,000634596     | 0,299971051 |
| SLC7A8             | 5,21372E-09     | 0,345715919 | GDPD3             | 0,00019909      | 0,3628023   | SLC7A8             | 1,48499E-07     | 0,303806257 | DUOXA1             | 0,000161533     | 0,315050156 |
| FAM46C             | 0,002563293     | 0,350057505 | UGT1A1            | 0,000289346     | 0,363876802 | DNASE1L3           | 0,001442721     | 0,334175057 | LDHD               | 0,000173782     | 0,316158524 |
| ID1                | 8,37686E-06     | 0,351875795 | DNASE1L3          | 0,00013065      | 0,365512682 | KLF4               | 4,34891E-08     | 0,343640339 | AKR7A3             | 0,021858878     | 0,322990232 |
| KLF4               | 6,90708E-10     | 0,357587002 | CXCL14            | 0,003980942     | 0,38118529  | ID3                | 0,00018238      | 0,350311909 | SLC7A8             | 1,33295E-08     | 0,323192845 |
| CBR1               | 5,81541E-09     | 0,363403528 | DUSP5             | 0,001309189     | 0,393784088 | CYP2C19            | 0,005404107     | 0,357668257 | TMEM92             | 0,015105832     | 0,323282886 |
| DUSP5              | 6,84694E-05     | 0,363808069 | SELENBP1          | 1,61537E-05     | 0,394863922 | GCNT4              | 3,24527E-05     | 0,364523011 | MT1X               | 4,72848E-06     | 0,327683505 |
| GDPD3              | 5,64695E-05     | 0,371902669 | CLEC3B            | 8,48424E-05     | 0,395629862 | OSGIN1             | 9,5217E-07      | 0,376032805 | SELENBP1           | 5,87568E-06     | 0,343139253 |
| SNRPN              | 1,62409E-09     | 0,37312038  | CKB               | 0,000616878     | 0,397967373 | TST                | 0,027746621     | 0,411411558 | CIDEA              | 2,44049E-05     | 0,356058831 |
| TMEM92             | 0,008690596     | 0,38516402  | OSGIN1            | 5,73242E-07     | 0,407595324 | CYP4F12            | 0,029083096     | 0,448317467 | FAM46C             | 0,012317706     | 0,357571122 |
| DUOXA1             | 0,004804132     | 0,38860703  | CFD               | 0,005691332     | 0,414474757 | ID1                | 0,015355038     | 0,576842773 | ID3                | 0,000161533     | 0,368951665 |
| PTGR1              | 3,82954E-08     | 0,392976374 | PTGR1             | 8,09088E-07     | 0,417050285 | HADH               | 6,09463E-05     | 0,635107622 | ID1                | 6,13509E-06     | 0,381906737 |
| HOMER2             | 0,001298949     | 0,399305519 | SLC7A8            | 2,67411E-06     | 0,43423589  |                    |                 |             | KLF4               | 4,59831E-07     | 0,385874712 |
| TST                | 0,001238429     | 0,416397305 | LYVE1             | 0,005526871     | 0,436925004 |                    |                 |             | CYP4F12            | 0,000134371     | 0,388830627 |
|                    |                 |             |                   |                 |             |                    |                 |             |                    |                 |             |

| Normal vs CIN DOWN |                 |             | Normal vs GS DOWN |                 |             | Normal vs EBV DOWN |                 |             | Normal vs MSI DOWN |                 |             |
|--------------------|-----------------|-------------|-------------------|-----------------|-------------|--------------------|-----------------|-------------|--------------------|-----------------|-------------|
| GENE               | W.pval.unpaired | fold.change | GENE              | W.pval.unpaired | fold.change | GENE               | W.pval.unpaired | fold.change | GENE               | W.pval.unpaired | fold.change |
| ID3                | 0,000105854     | 0,425416047 | CYP4F12           | 0,003861999     | 0,439270987 |                    |                 |             | OSGIN1             | 2,22627E-07     | 0,414541894 |
| CYP4F12            | 0,000147783     | 0,435982284 | CIDEDEC           | 0,007595456     | 0,455258363 |                    |                 |             | GDPD3              | 2,34348E-05     | 0,4147334   |
| SELENBP1           | 5,6974E-06      | 0,445412155 | SCARA5            | 0,001224645     | 0,473825364 |                    |                 |             | PTGR1              | 1,65713E-07     | 0,422270169 |
| CIDEDEC            | 0,0016966       | 0,447676689 | ID3               | 0,00331431      | 0,497196928 |                    |                 |             | TST                | 0,002733818     | 0,434546247 |
| OSGIN1             | 2,00812E-05     | 0,455780648 | HADH              | 1,19118E-05     | 0,599756026 |                    |                 |             | CBR1               | 2,66295E-06     | 0,44857795  |
| UGT1A1             | 0,006999246     | 0,468275768 |                   |                 |             |                    |                 |             |                    |                 |             |
| HADH               | 5,97609E-09     | 0,579313673 |                   |                 |             |                    |                 |             |                    |                 |             |



## Titre : Caractérisation multiparamétrique des cancers de l'œsophage et de l'estomac

**Mots clés : Cancer gastrique, MSI, épissage alternatif, détoxification, antioxydants, résistance aux médicaments.**

**Résumé :** Le cancer gastrique est la deuxième cause de mortalité par cancer dans le monde. En France, la région Grand Ouest est particulièrement touchée, et le département du Finistère se situe au premier rang. L'incidence des cancers de l'estomac et de la jonction œsogastrique augmente et les thérapeutiques disponibles montrent très souvent leurs limites, ce qui en fait un problème de santé publique important.

Afin d'étudier les bases moléculaires de ces cancers, un premier axe de ces travaux a consisté à mener un ensemble d'analyses des données génomiques issues des projets Cancer Genome Atlas (TCGA) et de l'International Cancer Genome Consortium (ICGC). Parmi une cinquantaine de mutations ponctuelles de type insertion, délétion ou substitution, situées dans des régions exoniques et introniques et d'une fréquence supérieure à 5%, huit délétions ont été retrouvées dans les cancers gastriques archivés au CHRU de Brest, deux d'entre elles pouvant conduire à un défaut d'épissage. L'analyse de l'épissage des gènes mutés a montré l'apparition de transcrits alternatifs uniquement dans les cancers (y compris au-delà de l'estomac) et dans les lignées cancéreuses avec instabilité des séquences microsatellites (MSI), suggérant un lien entre la machinerie de réparation de l'ADN et le contrôle de l'épissage. Afin d'analyser le mécanisme d'épissage alternatif résultant de ces mutations, plusieurs minigènes ont été construits à partir des séquences d'intérêt, mutées et non mutées, puis transfectés dans des cellules cancéreuses MSI et NON-MSI.

Les résultats seront présentés, et mis en perspective afin d'en déterminer les répercussions sur la pathogénèse du cancer gastrique.

Le deuxième axe de ces travaux de recherche a consisté en l'analyse de données transcriptomiques issues également du projet Cancer Genome Atlas. Nous avons ainsi mis en évidence des ensembles de miner les répercussions sur la pathogénèse du cancer gastrique.

Plus de 140 gènes dérégulés à la baisse et plus de 400 gènes dérégulés à la hausse dans les adénocarcinomes gastriques. A la suite de quoi, nous avons pu établir un réseau in silico, comportant 687 liens associant 540 gènes dérégulés dans les tumeurs gastriques. Après avoir intégré ces ensembles de données au sein d'un réseau génique d'interactions, nous avons pu mettre en évidence plusieurs groupements de gènes interagissant respectivement ensemble. Notre attention a été attirée par le comportement d'un petit sous-réseau, baptisé « réseau Papillon », comportant 9 gènes impliqués, collégialement, dans les réactions de détoxification, et tous sous-exprimés dans les échantillons de cancers gastriques. Nous espérons maintenant comprendre comment ces voies sont dérégulées et leurs rôles dans le mécanisme de pathogénicité du cancer gastrique, qui pourraient à terme constituer de nouvelles cibles thérapeutiques.

## Titre : Characterization of esophageal and gastric adenocarcinoma

**Keywords : gastric cancer, MSI, splicing, detoxification, antioxidants, drug resistance.**

**Abstract :** Gastric cancer is the world's second leading cause of cancer mortality. In France, 6550 cases were diagnosed in 2012. Brittany is particularly affected; Finistère is the worst department in France. The incidence of stomach and esophageal cancer increases and the few treatment options are ineffective because of late detection, which make it a major public health problem.

To dismember the genetic causes of these cancers, we conducted a series of analyzes on the Cancer Genome Atlas project database, which compiles genomics, transcriptomics and epigenomics data from more than 290 patients with stomach adenocarcinoma and 200 patients with esophagus adenocarcinoma. These analyzes allowed us to identify about fifty point mutations as insertions, deletions or substitutions in exonic, intronic and intergenic regions, and with a frequency higher than 5%. Of eight simple nucleotide deletions found in gastric cancers collected in Brest University Hospital, two could lead to a splice defect. Analysis of these mutated sequences has shown the appearance of alternative transcripts only in cancers (including beyond the stomach?) and in microsatellite instability (MSI) tumor cell

lines, suggesting a link between the DNA repair machinery and the control of splicing. In addition, bioinformatics predictions indicated that hnRNPA1 splice protein could bind to the immediate vicinity of an intronic homopolymeric sequence, which would lead to skipping the next exon.

In order to analyze the alternative splicing mechanism resulting from these mutations, several minigenes encompassing the alternatively spliced region were engineered from the mutated and non-mutated sequences of interest. We have also identified more than 140 genes under-expressed and 400 genes over-expressed in gastric adenocarcinoma. After integrating this data in a gene interaction network, we were able to identify several groups ("clusters") of genes interacting together. Our attention was drawn to the behavior of a small sub-network, comprising 9 genes involved, collegially, in detoxification reactions, and all under-expressed in gastric cancer samples. We hope to understand how these pathways are deregulated and their roles in the pathogenicity mechanism of cancer, which eventually constitute new therapeutic targets.