

# Sélection et décryptage

## L'innovation en cancérologie à travers la presse scientifique et médicale

### Diversité mutationnelle des sous-clones des cancers colorectaux\*

Les cancers contiennent de multiples mutations dont la signification est différente. Les mutations d'origine, pilotes (*driver*) et clonales, sont ubiquitaires dans la tumeur et assurent l'essentiel du phénotype cancéreux et de la progression maligne. À l'opposé, les mutations sous-clonales sont dispersées dans la tumeur et constituent un réservoir de cellules mutantes qui peuvent être sélectionnées et repeupler la tumeur en fonction des variations des pressions qu'exerce l'environnement tumoral, qui peut être d'origine thérapeutique. Les auteurs ont cherché à quantifier les mutations sous-clonales avec précision dans un échantillonnage de cinq cancers colorectaux, en utilisant le NGS (*next generation sequencing*) avec une très grande profondeur d'analyse.

Ils ont ainsi recherché les mutations présentes dans les gènes des ADN polymérase de réplication et dans les gènes fréquemment mutés dans les cancers colorectaux. Ils observent un taux de mutation très élevé ( $7 \times 10^{-7}$ ). Ces mutations sous-clonales surviennent successivement, de façon ramifiée comme les branches d'un arbre (évolution « branchée »), les mutations les plus rares se situant aux extrémités des branches. L'accumulation des mutations sous-clonales suit un modèle d'évolution neutre, aussi longtemps que la pression de sélection ne change pas.

La conclusion d'un tel travail est extrêmement importante en pratique, malgré l'aspect purement théorique qu'il revêt. En effet, les auteurs peuvent conclure de leur

modélisation mathématique qu'il ne peut exister dans une tumeur, au moment du diagnostic, une seule cellule ne présentant pas au moins une mutation sous-clonale. En d'autres termes, chaque nucléotide de l'ADN est muté dans au moins une cellule de la tumeur... En particulier, toutes les mutations possibles conférant la résistance au traitement sont déjà présentes dans la tumeur au moment du diagnostic.

C'est une mauvaise nouvelle que nous apporte cet article, une de plus, comme chaque fois que l'on creuse de plus en plus profondément dans le génome de la cellule cancéreuse. On s'en doutait un peu, depuis que l'on a vu émerger systématiquement des cellules porteuses de la mutation T790M dans les cancers du poumon traités par un inhibiteur de tyrosine kinase de première génération [1].

\* Loeb LA, Kohn BF, Loubet-Seneor KJ, et al. Extensive subclonal mutational diversity in human colorectal cancer and its significance. *Proc Natl Acad Sci USA* 2019 ; 116 (52) : 26863-72. Doi : 10.1073/pnas.1910301116.

#### RÉFÉRENCE

1. Robert J. Des nouvelles de la course aux armements. *Innov Ther Oncol* 2019 ; 5 : 243-4.

Jacques Robert

### Rien ne ressemble plus au génome des métastases que celui des tumeurs primitives\*

Dès que l'on a su collectionner les profils d'expression génique des cancers, on s'est posé la question de savoir si l'on pourrait identifier un « portrait » caractéristique des métastases, globalement distinct de celui des tumeurs primitives, révélant ainsi des cibles potentielles originales pour combattre le processus métastatique en tant que tel [1]. Il a fallu rapidement déchanter : rien ne ressemble plus au transcriptome d'une métastase que celui de la tumeur primitive dont elle dérive... S'il existe des « gènes de métastase », dont l'expression est activée pour mettre en place la dissémination des cellules cancéreuses, ils sont bien cachés...

Qu'en est-il en ce qui concerne le génome ? Peut-on identifier des gènes dont les mutations, survenues

secondairement au développement du primitif, permettent de franchir les étapes successives nécessaires à la formation des métastases ? Dans le cadre du modèle linéaire de la progression métastatique, proposé par le groupe de Vogelstein dès 1990 [2], cela serait envisageable ; mais il semble que le modèle parallèle explique plus volontiers cette progression : tout ce qui permettra l'éclosion d'une métastase est inscrit dans le génome du primitif.

L'article que publie la revue *Nature* était nécessaire, d'une part, pour confirmer globalement cette hypothèse et, d'autre part, pour établir des règles générales relatives à l'évolution mutationnelle des métastases – car si le primitif évolue sur ce plan, essentiellement en raison de

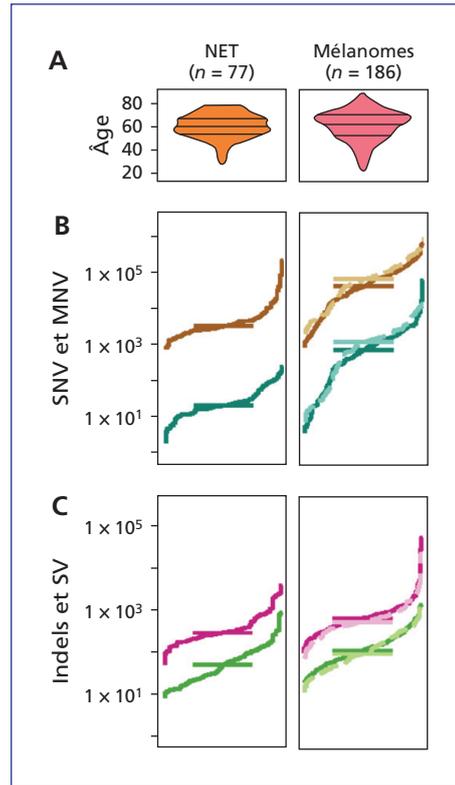
l'instabilité génomique des cellules cancéreuses, il n'y a pas de raison que la métastase n'en fasse pas autant. Soumises à des pressions de sélection distinctes de celles auxquelles sont soumises les tumeurs primitives, les métastases sélectionneraient peut-être des sous-clones originaux pouvant être représentatifs du phénomène de dissémination.

Les auteurs ont réalisé le séquençage par NGS (*next generation sequencing*) de 2 520 tumeurs appariées au tissu normal correspondant. Travail de rouleau compresseur auquel nous sommes maintenant habitués, conduisant à l'identification de 59 472 629 variations de nucléotide unique (SNV), 839 126 variations de plusieurs nucléotides consécutifs (MNV), 9 598 205 insertions et délétions (indels) et 653 452 variations structurales (SV, amplifications, translocations et délétions).

Qu'ont-ils trouvé dans les métastases ? Comme dans les tumeurs primitives, un nombre très variable de SNV par échantillon, pour tous les types de cancers avec une charge mutationnelle élevée, en particulier dans les mélanomes et les cancers du poumon, et faible dans les tumeurs neuro-endocrines ou les mésothéliomes (*figure 1*). Les MNV sont également plus fréquents dans les métastases des cancers du poumon et des mélanomes que dans les métastases des autres tumeurs ; rien de surprenant quand on sait que le tabac induit des transversions souvent dinucléotidiques de type CC>AA, et les rayons UV des transitions dinucléotidiques CC>TT. Les indels se rencontrent dans les métastases des cancers dont le primitif présente le phénotype d'instabilité microsatellitaire (MSI), tumeurs du cerveau, de l'endomètre et de la prostate, les cancers colorectaux étant sous-représentés en raison du caractère peu métastasiant des cancers MSI.

Quant aux variations structurales, elles présentent la même répartition dans les métastases que dans les tumeurs primitives. Elles concernent surtout, en ce qui concerne les amplifications, les régions contenant les gènes *EGFR*, *CCND1*, *CCNE1*, *MDM2*, ainsi que *TERT* et *MYC*. Pour les délétions, elles correspondent à de nombreuses pertes d'hétérozygotie survenant dans des régions contenant les gènes suppresseurs de tumeurs classiques *RB1*, *PTEN*, *CDKN2A* ou *TP53*. Rien d'original par rapport aux tumeurs primitives des mêmes organes. La *figure 2* présente les altérations structurales les plus fréquemment rencontrées dans les métastases. Une seule différence entre métastases et tumeurs primitives considérées en bloc : la fréquence des duplications du génome entier dans les métastases, beaucoup plus élevée (56 %) que dans les tumeurs primitives (30 %).

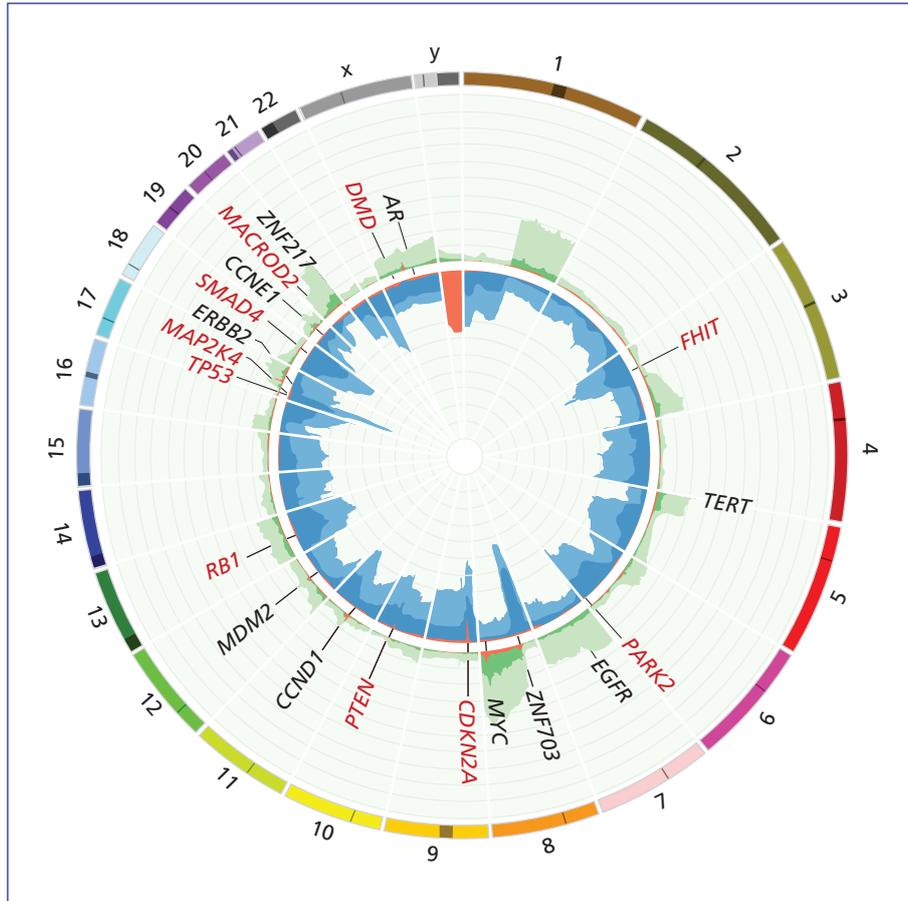
Les auteurs dressent enfin un catalogue des mutations pilotes rencontrées dans les métastases (*figure 3*) ; ce catalogue est à mon avis en tout point analogue à celui établi pour les tumeurs primitives. Bien malin celui qui pourrait discerner des différences autres



**Figure 1.** Nombre des altérations moléculaires somatiques présentes dans les métastases de deux types de cancers, aux extrémités du spectre : tumeurs neuro-endocrines (NET) et mélanomes. Sont présentés en (A) l'âge des patients, en (B) les variations de nucléotide unique (SNV) et les variations de plusieurs nucléotides consécutifs (MNV), et en (C) les insertions et délétions (indels) et les variations structurales (amplifications, translocations et délétions, SV). On peut voir, par exemple, que les NET contiennent une valeur médiane de 4 100 SNV dans leur génome et les cancers du poumon une valeur médiane de 36 000.

**Figure 1.** Number of somatic molecular alterations found in metastases of two cancer types at either end of the spectrum: neuro-endocrine tumors (NET) and melanoma. These are represented according to age of the patients (A), single nucleotide variations (SNV) and multiple nucleotide variations (MNV) (B), and insertions and deletions (indels) and structural variations (amplifications, deletions, translocations, SV) (C). For example, the median number of SNV in the genome associated with NETs and lung cancers is 4,100 and 36,00, respectively.

qu'aléatoires (dépendantes de l'échantillonnage des cancers étudié) entre primitif et métastase... Signalons tout de même que les mutations pilotes sont globalement plus fréquentes que dans les tumeurs primitives dans chaque type de cancer, ce qui pourrait simplement témoigner d'un état plus avancé du cancer lors de la biopsie. Les gènes *ESR1* et *AR* présentent en particulier des mutations plus fréquentes dans les métastases des cancers du sein et de la prostate (16 et



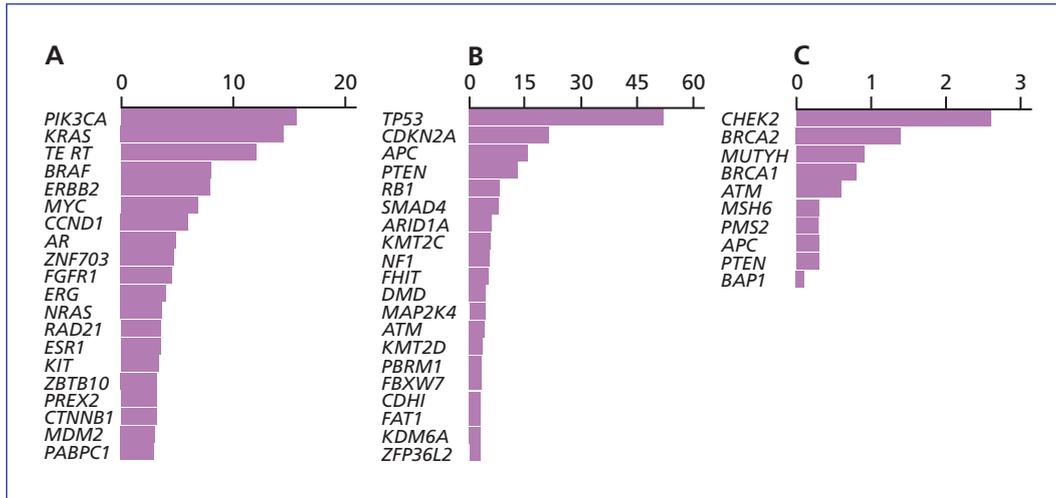
**Figure 2.** Circos plot indiquant la proportion des échantillons de métastases présentant des événements d'amplification et de délétion géniques selon leur position dans le génome. Le cercle intérieur montre le pourcentage de tumeurs avec délétion homozygote (orange), perte d'hétérozygotie (bleu foncé ou clair selon son importance). Le cercle extérieur montre le pourcentage de tumeurs avec amplification (orange, vert foncé ou vert clair selon son importance). Les amplifications les plus fréquentes indiquées en noir concernent les oncogènes, et les délétions homozygotes les plus fréquentes indiquées en rouge concernent les gènes suppresseurs de tumeurs.

**Figure 2.** Circos plot showing the proportion of metastasis samples bearing gene amplification and deletion events according to their localization within the genome. The inner ring shows the percentage of tumours with homozygous deletion (orange) and loss of heterozygosity (dark or light blue according to magnitude). The outer ring shows the percentage of tumours with amplification (orange, dark green or light green according to magnitude). The most frequent gene amplifications in black refer to oncogenes, and the most frequent homozygous deletions in red refer to tumour suppressor genes.

44 % respectivement) que dans les primitifs (chiffres non fournis...). Enfin, les métastases présentent globalement un nombre plus élevé de mutations pilotes que les primitifs ; mais la comparaison est toujours faite entre types tumoraux, et pas entre la métastase et le primitif correspondant.

On est déçu, en fin de compte, de ne pas trouver dans cet article une comparaison « face à face » des mutations rencontrées dans les métastases et dans les tumeurs primitives correspondantes, mais seulement une comparaison globale... J'aurais aimé savoir, par exemple,

le nombre médian de mutations pilotes dans les métastases de tel ou tel type de cancer par rapport à celui trouvé dans les primitifs correspondants. *Nature* nous a habitués à présenter des catalogues purement descriptifs : en voici un de plus ! Peut-on appeler cela de la recherche ? J'aimerais que les auteurs poursuivent leur travail sur cet impressionnant corpus de données et que l'on puisse savoir si, comme je le pense, « rien ne ressemble plus au génome d'une métastase que celui de la tumeur primitive dont elle dérive ». Car c'est la vraie question...



**Figure 3.** Altérations moléculaires les plus fréquentes rencontrées dans les métastases au niveau somatique : (A) oncogènes ; (B) gènes suppresseurs de tumeurs. (C) Altérations moléculaires de gènes de prédisposition héréditaire au cancer rencontrées au niveau germlinal.

**Figure 3.** The most prevalent somatically mutated oncogenes found in metastases: (A) oncogenes and (B) tumour-suppressor genes. (C) Germline gene variants associated with a predisposition to cancer.

\* Priestley P, Baber J, Lolkema MP, et al. Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature* 2019 ; 575 : 210-6. Doi : 10.1038/s41586-019-1689-y.

2. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* 1990 ; 61 : 759-67.

## RÉFÉRENCES

1. Perou CM, Sørlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature* 2000 ; 406 : 747-52.

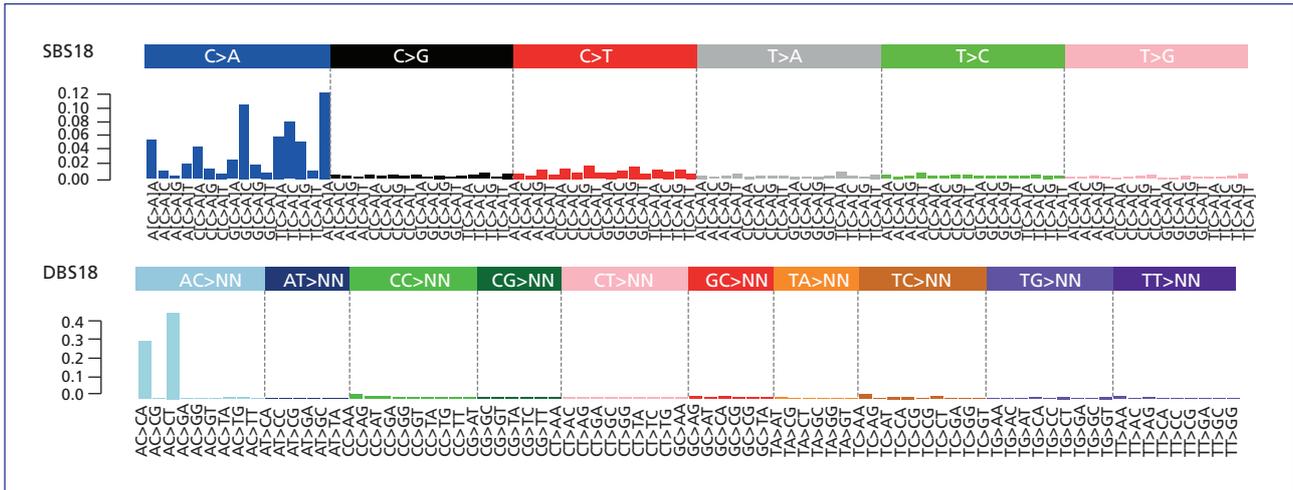
Jacques Robert

## Les tissus épithéliaux normaux portent aussi des mutations de gènes de cancer\*

La présence de mutations de « gènes de cancer » dans un tissu épithélial normal est d'observation courante : on sait très bien qu'une mutation oncogénique ne fait pas le cancer et que, de la cellule cancéreuse porteuse de mutations au cancer constitué, il existe une série d'étapes à franchir – et qui souvent ne sont jamais franchies : acquisition du phénotype d'immortalité, angiogenèse, capacité migratoire, échappement à la surveillance immunologique, etc. Cet article de *Nature*, qui présente le séquençage d'environ 2 000 villosités non cancéreuses prélevées chez 42 individus, vient nous le rappeler.

Notre muqueuse colique est plissée en plus de 15 millions de villosités contenant chacune environ 2 000 cellules. Au fond de ces villosités sont installées les cellules souches qui donnent naissance aux cellules différenciées qui migrent vers la surface de la muqueuse. Les mutations acquises par ces cellules souches seront donc présentes dans les 2 000 cellules qui tapissent la villosité. Les auteurs se sont donc attaqués à l'évaluation de la charge mutationnelle, des signatures mutationnelles et de la fréquence des mutations pilotes rencontrées dans ces cryptes.

La charge mutationnelle est très variable et va de 1 500 à 15 000 mutations par échantillon pour des individus autour de la soixantaine, avec une valeur médiane de 3 000. Une dizaine de signatures caractéristiques de processus oncogéniques précis ont été identifiées dans ces villosités : par exemple, la signature SBS1, associée à la désamination spontanée de la 5-méthylcytosine, la signature SBS5, rencontrée en particulier dans les mutations germinales conduisant à la prédisposition héréditaire au cancer, ou la signature SBS18, attribuée aux dommages de l'ADN causés par les radicaux libres oxygénés ; en revanche, les signatures SBS4, liée à l'exposition aux composés de la fumée du tabac, et SBS7, liée à l'exposition aux rayons UV, ne sont pas rencontrées. En outre, de nouvelles signatures, non encore répertoriées, ont été identifiées dans les villosités, dont la signification reste à explorer. Toujours à titre d'exemple, on peut noter que la présence de la signature SBS1 augmente avec l'âge des sujets, mais que sa fréquence est identique dans toutes les villosités du même individu. Enfin, des signatures rares, sporadiques, ont été rencontrées dans les villosités de sujets qui avaient



**Figure 1.** Exemple de la distribution des mutations rencontrées dans les villosités coliques normales dans deux signatures : la SBS (*single base substitution*) 18, attribuée aux dommages de l'ADN causés par les radicaux libres oxygénés, et la DBS (*double base substitution*) 8, de cause oncogénique incertaine.

**Figure 1.** Example showing the distribution of mutations found in normal colonic crypts with two signatures: SBS (*single base substitution*) 18, likely due to DNA damage from oxygen radicals, and DBS (*double base substitution*) 8, of unknown significance.

été traités précédemment pour un cancer, montrant l'impact de la chimiothérapie sur la survenue de mutations dans les tissus normaux.

Les altérations structurales majeures (amplifications, délétions, translocations, SV) sont rencontrées avec une faible fréquence dans les villosités normales. Elles ont été identifiées dans 18 % des villosités : des duplications d'un chromosome sont rencontrées (chromosomes 3, 7 et 9 ; des amplifications focales de ces chromosomes sont d'ailleurs rencontrées dans les cancers colorectaux). Aucune perte de chromosome entier n'a été observée dans les villosités normales, mais des délétions focales de divers bras chromosomiques ont été trouvées dans une douzaine de cas. À l'exception d'un cas, seule une villosité est affectée chez un individu donné.

Les gènes dans lesquels on rencontre des mutations de gènes de cancer sont des gènes qui ne sont pas tous connus pour être oncogéniques dans les cancers colorectaux : pour les oncogènes, *PIK3CA*, *ERBB2*, *ERBB3*, *FBXW7* ; pour les gènes suppresseurs de tumeurs, *AXIN2* et *STAG2* à l'état homozygote, et *ARID2*, *ATM*, *ATR*, *BRCA2*, *CDK12*, *CDKN1B*, *RNF43*, *TBL1XR1* et *TP53* à l'état hétérozygote. Environ 1 % des villosités analysées porte une des mutations pilotes (celles des oncogènes et celles homozygotes des gènes suppresseurs de tumeurs) chez les sujets de la tranche d'âge 50-60 ans. Sachant qu'à 70 ans, 40 % des individus portent un adénome dans leur muqueuse colique, et que 5 % des individus développent un cancer colorectal au cours de leur vie, on peut voir que seule une proportion extrêmement faible de ces micro-cancers présents dans les villosités deviendra un adénome détectable (un sur 400 000 environ) ou un carcinome (un sur 3 millions) dans la décennie suivante.

Poursuivons cette comparaison entre micro-cancers (villosités présentant une mutation pilote d'un gène de cancer) et carcinomes constitués. La charge mutationnelle des cancers colorectaux est de 10 000 à 20 000 substitutions (en dehors des cancers de phénotype hypermutateur lié à l'instabilité microsatellitaire), de 1 000 à 2 000 indels et de plusieurs centaines de SV, à comparer avec les 3 000 substitutions, les 300 indels et les 0 à 4 SV rencontrés dans les villosités normales des sujets de la tranche d'âge 50-60 ans. Aucun phénotype hypermutateur n'est rencontré dans les villosités normales, alors qu'il représente 15 à 20 % des carcinomes colorectaux. La différence la plus importante entre villosités normales et carcinomes est peut-être la nature des gènes mutés ; alors que les mutations des gènes *KRAS*, *TP53* et *APC* sont parmi les plus fréquentes dans les carcinomes colorectaux, elles sont rares dans les villosités normales (une mutation observée dans 14 villosités) ; et celles des gènes *ERBB2* et *ERBB3*, rares dans les carcinomes, sont fréquentes dans les villosités normales. On pourrait conclure que les mutations des gènes *APC*, *TP53* et *KRAS* sont cruciales dans la conversion de ces îlots de cellules tumorales en adénomes et en carcinomes.

Cet article est important, en ce qu'il confirme de façon documentée le fait qu'un îlot de cellules cancéreuses présent dans la muqueuse colique n'est pas destiné à devenir un cancer. Il est à noter qu'aucune des mutations de gènes de cancer rencontrés dans les villosités normales n'est directement liée à l'instabilité génomique (microsatellitaire ou non) ; cela suggère que le *primum movens* de la transformation maligne est bien l'instabilité génomique, comme de nombreux auteurs commencent

à le penser, et que les mutations sporadiques rencontrées dans les tissus épithéliaux normaux ne constituent pas un état précancéreux – au moins jusqu’à ce que survienne une mutation d’un gène impliqué dans la reconnaissance et la réparation des dommages de l’ADN.

\* Lee-Six H, Olafsson S, Ellis P, *et al.* The landscape of somatic mutation in normal colorectal epithelial cells. *Nature* 2019 ; 574 : 532-37. Doi : 10.1038/s41586-019-1672-7.

Jacques Robert

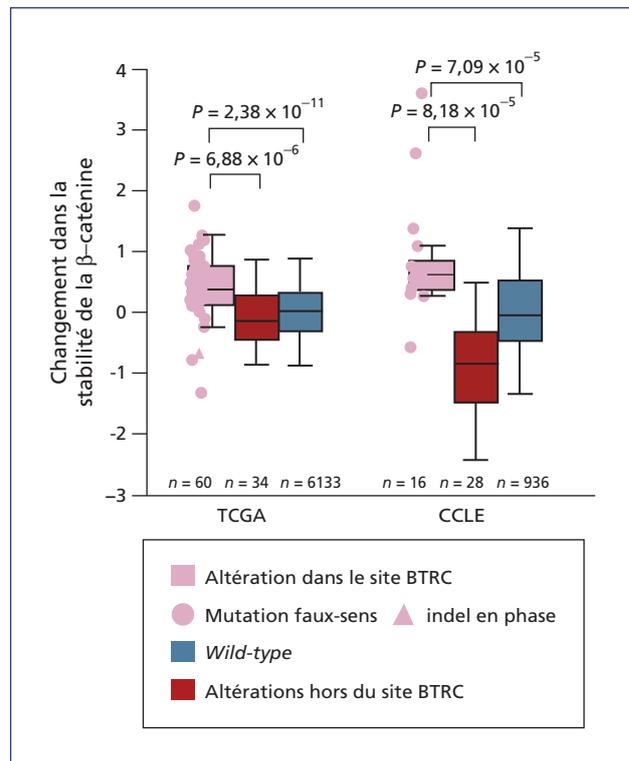
## Un monde à découvrir : les mutations oncogéniques des sites de liaison des protéines à leur E3 ubiquitine ligase\*

La protéolyse constitue un moyen essentiel de régulation de la durée de vie des protéines. Le protéasome assure une dégradation spécifique de nombreuses protéines après qu’elles ont été étiquetées par un polymère de molécules d’ubiquitine. Le système ubiquitine - protéasome (UPS) est, d’une part, un élément de régulation de phénomènes comme l’expression des gènes, la transduction des signaux, la progression du cycle cellulaire, la réparation de l’ADN et l’apoptose. Il permet, d’autre part, l’élimination des protéines mal repliées, endommagées ou mutantes, de conformation anormale, dont l’accumulation pourrait être délétère pour la cellule. Des aberrations dans cette voie conduisent à des phénotypes pathologiques qui font d’elle une cible pour une intervention thérapeutique dans les cancers.

L’ubiquitine (UB) est une protéine de 76 résidus aminoacides qui est conjuguée avec la protéine à détruire à la suite de trois réactions : une première enzyme E1 (*ubiquitin-activating enzyme*) active l’ubiquitine, qui est ensuite transférée sur un résidu cystéine d’une enzyme E2 (*ubiquitin-conjugating enzyme*). Enfin, les enzymes E3 (*ubiquitin protein ligase*) transfèrent l’ubiquitine sur la protéine substrat à détruire grâce à une liaison amide entre le groupe carboxyle du résidu glycine C-terminal de l’ubiquitine et le groupement amine d’un résidu interne de lysine de la protéine substrat. Il existe deux enzymes E1, quelques dizaines d’enzymes E2 et plusieurs centaines d’enzymes E3 qui portent la spécificité de substrat de la protéine à ubiquitinyler.

Les auteurs ont cherché à évaluer l’importance du système ubiquitine - protéasome dans l’oncogenèse. Ils se sont d’abord adressés aux mutations des motifs de reconnaissance, sur les protéines à détruire, mutations qui leur permettent d’échapper à la destruction ; ils se sont ensuite adressés aux mutations des E3 ubiquitine ligases. Pour identifier ces motifs de reconnaissance, que l’on appelle maintenant des *degrons*, sur l’ensemble des protéines, ils ont utilisé un *machine learning classifier*, autrement dit une méthode d’intelligence artificielle, à partir d’un set de 11 caractéristiques connues des

degrons déjà identifiés. La *figure 1* montre la séquence peptidique consensus d’un degron : un total de 28 degrons distincts correspond à cette structure. À partir de 32 022 isoformes de 19 986 protéines humaines, ils identifient ainsi plus de 80 000 degrons potentiels,



**Figure 2.** Augmentation de la stabilité de la  $\beta$ -caténine (gène *CTNNB1*) lors de mutations ou d’indels survenant soit dans la séquence de son degron reconnu par l’E3 ubiquitine ligase  $\beta$ -TRCP (gène *BTRC*), soit en dehors de cette séquence. Travail réalisé sur des échantillons tumoraux humains en provenance du TCGA (*The Cancer Genome Atlas*) ou de lignées cellulaires tumorales en provenance de la CCLE (*Cancer Cell Line Encyclopedia*).

**Figure 2.** Enhanced integrity of *b-catenin* (the *CTNNB1* gene) when mutations or indels occur either within the degron sequence recognised by ubiquitin E3 ligase  $\beta$ -TRCP (the *BTRC* gene), or outside of this sequence. This finding is based on human tumour samples originating from the TCGA (*The Cancer Genome Atlas*) and tumoural cell lines from the CCLE (*Cancer Cell Line Encyclopedia*).



**Figure 1.** Séquence peptidique consensus d’un degron.  
*Figure 1.* Consensus peptide sequence of a degron.

nombre réduit à 20 000 en ne prenant en compte que ceux qui ont plus de 50 % de chances d'être un vrai degron.

Une fois ce travail achevé, ils ont pris en considération les séquences de 236 protéines dans près de 7 000 tumeurs et 212 protéines dans 861 lignées cellulaires et ont cherché dans ces protéines des mutations présentes au niveau d'un degron et leurs conséquences fonctionnelles. Si la fonction d'un degron est altérée, on peut penser que la protéine aura une stabilité supérieure ou inférieure à celle de la protéine sauvage : stabilité augmentée si la mutation du degron masque le site de reconnaissance de la protéine par l'E3 ligase ; stabilité diminuée si la mutation du degron favorise la reconnaissance de la protéine par l'E3 ligase. La stabilité de la protéine peut être évaluée par une distorsion entre l'expression du gène mesurée par l'abondance de son messager et l'expression de la protéine évaluée par *reverse-phase protein assay* (RPPA). On peut penser *a priori* que les mutations des degrons conféreront une plus grande stabilité aux protéines produits d'oncogènes et une moins grande stabilité aux protéines produits de gènes suppresseurs de tumeurs.

Quelques exemples de résultats obtenus : la protéine NFE2L2 (facteur de transcription) est le produit d'un oncogène muté dans divers cancers – vessie, col utérin, estomac, VADS, poumon (épidermoïde) ; 38 de ses 40 mutations non-synonymes identifiées dans le corpus tumoral étudié sont localisées au niveau d'un degron reconnu par son E3 ligase, KEAP1 ; la stabilité des 38 protéines porteuses de ces mutations est supérieure à celle des 473 protéines non porteuses de ces mutations. Il en est de même pour la  $\beta$ -caténine (gène *CTNNB1*) : 59 tumeurs et 16 lignées cellulaires portent une des 22 mutations et 11 indels (en phase) affectant le degron

$\beta$ -TRCP : la stabilité des protéines porteuses de ces altérations est supérieure à celle des protéines sauvages et des protéines affectées par une autre mutation (*figure 2*).

Au-delà de la confirmation du rôle des mutations affectant les degrons connus et répertoriés [1], l'approche d'intelligence artificielle utilisée permet également de découvrir de nouveaux degrons localisés dans des protéines oncogéniques où il n'en avait été pas encore été identifié, comme HER3, ERK, PKC $\alpha$ , BRAF, etc. Un total de 35 nouveaux degrons ont été ainsi identifiés. Ainsi, de nouveaux mécanismes contribuant à l'oncogenèse et à la progression des cancers ont-ils été identifiés. On savait déjà que de nombreux gènes codant des E3 ligases étaient porteurs de mutations invalidantes dans les cancers (*VHL* dans les cancers du rein, *FBXW7* dans les cancers colorectaux, *BTRC* dans divers cancers, etc.), permettant ainsi aux protéines, qu'elles ne pouvaient plus envoyer vers le protéasome, de jouer un rôle d'oncogènes (HIF1 $\alpha$ , MYC,  $\beta$ -caténine, respectivement). L'ampleur des altérations mutationnelles de l'UPS est ainsi multipliée par deux (au moins).

\* Martínez-Jiménez F, Muiños F, López-Arribillaga E, Lopez-Bigas N, Gonzalez-Perez A. Systematic analysis of alterations in the ubiquitin proteolysis system reveals its contribution to driver mutations in cancer. *Nature Cancer* 2019, sous presse. Doi : 10.1038/s43018-019-0001-2.

## RÉFÉRENCE

1. Mészáros B, Kumar M, Gibson TJ, Uyar B, Dosztányi Z. Degrons in cancer. *Sci Signal* 2017 ; 10 : pii: eaak9982.

Jacques Robert

## Les SNP extra-géniques modulent le risque de cancer en affectant oncogènes, gènes suppresseurs de tumeurs et fonctions immunitaires\*

Les GWAS (*genome-wide association studies*) ont révélé la présence de nombreux SNP (*single nucleotide polymorphism*) associés au risque de cancer (que j'appellerai ci-dessous SNP-AC), localisés en dehors des séquences codantes. Le mécanisme de leur mise en jeu pour moduler le risque de cancer est souvent inconnu ; nous en avons étudié deux dans des brèves précédentes [1, 2]. Les auteurs de cet article ont cherché à comprendre, de façon générale, la façon dont ils pouvaient ainsi être associés au risque de cancer. Pour ce faire, ils ont utilisé une approche dite « *Expression quantitative trait locus* (eQTL) », que je ne détaillerai pas (je la pratique depuis une dizaine d'années sans savoir que ça s'appelait ainsi), pour associer un génotype donné (c'est-à-dire la présence d'un SNP) au niveau d'expression d'un gène n'ayant *a priori* rien à voir avec le SNP. En analysant ainsi les

altérations du profil d'expression des gènes associées au SNP, on peut identifier une liste de gènes candidats qui pourraient être régulés par le SNP en question.

C'est un travail difficile en raison de la multiplicité des faux positifs attendus, des déséquilibres de liaison entre SNP, et de l'ignorance des chaînes d'information entre le SNP-AC et le gène dont l'expression est altérée (la présence du SNP régule « ceci », qui module l'expression de « cela », ce qui aboutit à la diminution de l'expression du gène considéré, « ceci » et « cela » pouvant être des micro-ARN ou des longs ARN non codants encore non identifiés). Plusieurs études ont toutefois déjà identifié des « gènes de cancer » (oncogènes et gènes suppresseurs de tumeurs) dont l'expression était altérée par la présence d'un SNP d'une région extra-génique (non codante) découvert lors d'une GWAS. Les auteurs ont sagement

choisi de s'adresser à des tissus sains en espérant avoir une situation plus nette, non polluée par l'activation de multiples oncogènes ou l'inactivation de multiples gènes suppresseurs de tumeurs.

À partir d'un catalogue recensant les SNP-AC, les auteurs ont extrait 872 SNP associés au risque de cancer avec une probabilité de  $p < 5 \times 10^{-8}$ . Un SNP donné est généralement associé avec le risque de survenue d'un seul type de cancer, mais chaque type de cancer peut être associé à la présence de plusieurs SNP dans le génome. Le risque de cancer est toujours faible : 99 % des SNP-AC sont associés au risque de cancer avec un risque relatif inférieur à 3. Environ 10 % des SNP-AC sont localisés dans un exon ou une zone d'épissage, 40 % dans un intron, et 50 % dans des régions intergéniques, ce qui veut dire que 90 % des SNP-AC jouent un rôle régulateur et n'ont pas de conséquence structurale pour la protéine. Plus intéressant est le fait que 3,3 % des SNP-AC sont localisés dans la région promotrice d'un gène quelconque, alors que seulement 0,9 % des SNP-nonAC sont localisés dans une telle région. Mieux : 84 % des SNP-AC sont localisés à moins de 100 kb d'une origine de transcription, contre 25 % pour les SNP non associés au risque de cancer.

L'analyse à haut débit des associations entre les SNP-AC et l'altération de l'expression de gènes a été effectuée dans 13 tissus normaux dont les profils d'expression sont disponibles dans des bases de données. On ne peut donner de leurs résultats que des données globales ou des exemples. Un exemple, tout d'abord : le SNP rs72699833 est associé au risque de cancer du sein ; sa présence est associée en *cis*<sup>1</sup> avec l'expression du gène *PHGDH*, un gène impliqué dans le métabolisme de la sérine et connu pour être surexprimé dans les cancers du sein, du col utérin, du côlon et du poumon non à petites cellules (NPC). En outre, il est associé en *trans* avec l'expression des gènes *LAD1*, *COL17A1*, *KRT10*, *LGALS7B* et *FERMT1*. Tous ces gènes sont impliqués dans l'interaction de la matrice extracellulaire avec les cellules épithéliales ; *FERMT1* joue en outre un rôle dans la transition épithélio-mésenchymateuse. Ils sont tous surexprimés ou sous-exprimés dans les cancers du sein.

Et, de façon plus globale, les auteurs ont cherché à savoir si les SNP-AC étaient plus associés à des modifications d'expression des « gènes de cancer » que les SNP-nonAC. La réponse est positive : les SNP-AC ont plus de chances d'être associés à la modification de l'expression d'un gène

de cancer que les SNP-nonAC ( $p < 10^{-6}$ ). Par ailleurs les SNP-AC sont plus souvent rencontrés dans les promoteurs des gènes de cancer que dans ceux des autres gènes. Le lien entre SNP-AC et cancer passe donc bien par un effet, direct ou indirect, sur un oncogène ou un gène suppresseur de tumeurs.

Un autre résultat important est le fait que de nombreux SNP-AC sont associés à une altération de l'expression des gènes des complexes majeurs d'histocompatibilité (MHC) des classes I et II. Cette association est clairement en *cis* : ce sont les SNP portés par le chromosome 6 qui sont associés de façon prépondérante à ces altérations d'expression des gènes MHC, tous regroupés sur ce chromosome. Et cette association est, comme celle observée avec les gènes de cancer, présente dans tous les types de tissus analysés. Ce résultat souligne une fois encore l'importance, dans la genèse des cancers, du système immunitaire.

Enfin, certains SNP-AC sont associés à l'expression de certains gènes de façon tissu-spécifique, non pas des « gènes de cancer », mais des gènes importants pour le développement du tissu en question, épithélial ou mésenchymateux. Par exemple, les gènes dont l'expression est altérée concomitamment à la présence d'un SNP-AC de façon spécifique au tissu cutané sont ceux impliqués dans le développement épithélial, le développement de l'épiderme, la mise en place de la barrière cutanée, la régulation des mouvements de l'eau, etc.

Sans résoudre les mécanismes précis selon lequel chaque SNP-AC détermine le risque de cancer, cet article pose clairement les bases qui permettront aux chercheurs futurs de décortiquer chacun de ces mécanismes de façon rationnelle.

\* Fagny M, Platig J, Kuijjer ML, Lin X, Quackenbush J. Nongenetic cancer-risk SNPs affect oncogenes, tumour-suppressor genes, and immune function. *Br J Cancer* 2019, sous presse. Doi : 10.1038/s41416-019-0614-3.

## RÉFÉRENCES

1. Robert J. Un SNP associé au risque du cancer de la prostate est localisé dans le site de liaison de plusieurs facteurs de transcription. *Innov Ther Oncol* 2018 ; 4 : 224-8.
2. Robert J. Comment un SNP peut être en cause dans de multiples phénomènes. *Innov Ther Oncol* 2019 ; 5 : 244-7.

Jacques Robert

<sup>1</sup> On parle d'« association en *cis* » lorsque l'élément régulateur (ici le SNP) et l'élément régulé (ici le gène) sont situés sur le même chromosome ; et d'« association en *trans* » dans le cas contraire.