

ARTICLE D'OPINION

Utilisation des omics : La génomique et la Métabolomique au service de la santé

Guillaume Grzych¹,
Mariem El Younsi²

1 CHU Lille, Service Hormonologie
Métabolisme Nutrition et
Oncologie, F-59000, Lille France

2 Service des Maladies Congénitales
et Héréditaires, Hôpital Charles
Nicolle, BabSouika, 1006, Tunis,
Tunisie.

INTRODUCTION

Actuellement, une meilleure compréhension de la physiopathologie de nombreuses pathologies complexes est nécessaire afin d'identifier de potentielles cibles thérapeutiques mais également d'identifier des marqueurs non invasifs pour le diagnostic et pour le suivi des évolutions des pathologies chez les patients. Dans ce contexte, les stratégies «omics», telles que la génomique et la métabolomique, peuvent se révéler prometteuses (**Figure 1**).

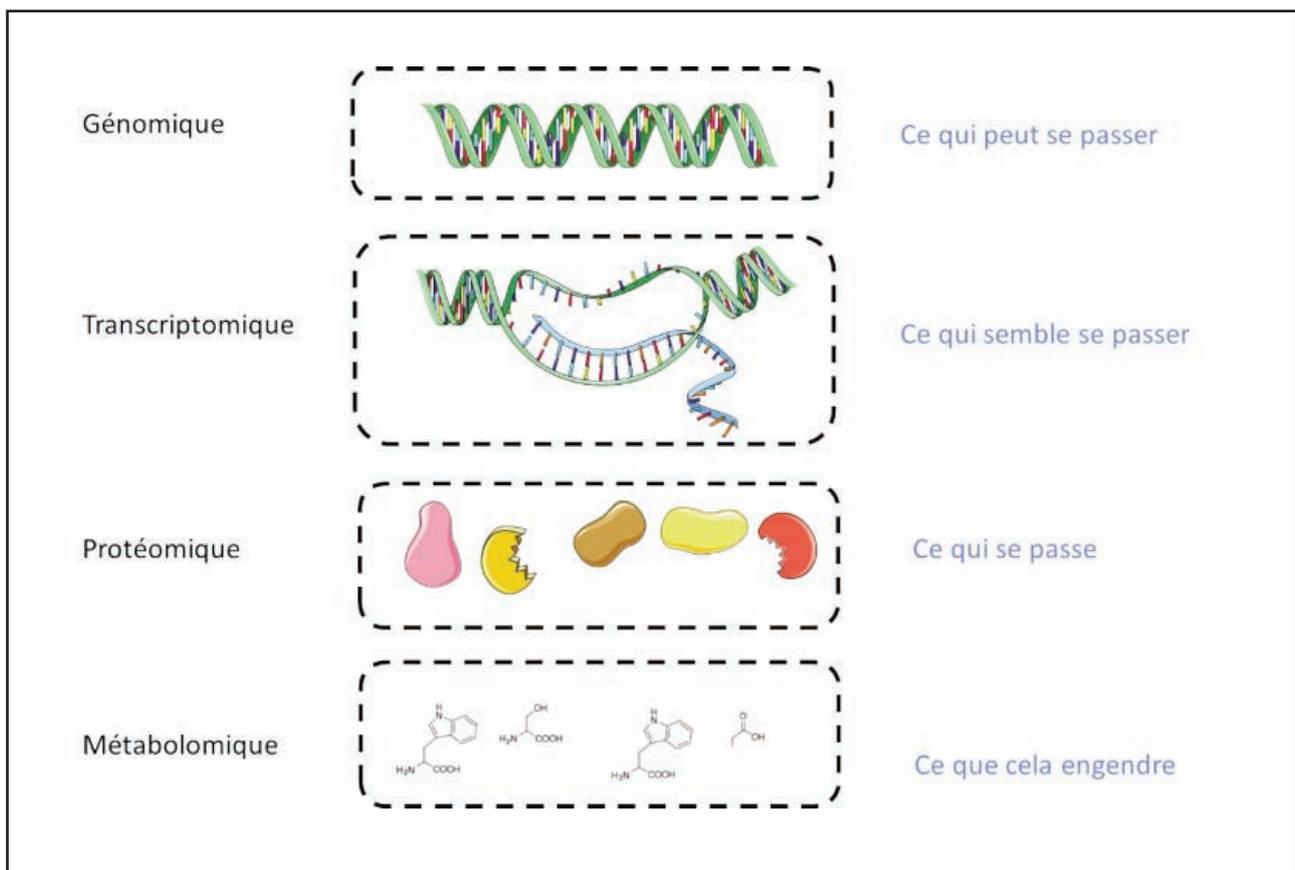


Figure 1: Stratégies omics et place de la génomique et de la métabolomique, extrait de Grzych G, 2017 (1).

1. GENOMIQUE

Selon l'Organisation mondiale de la santé (OMS), la « génétique » est l'étude de l'hérédité, et la « génomique » est l'étude des gènes et de leurs fonctions, et des techniques associées. L'achèvement du projet du génome humain (HGP) en 2003 a permis le développement de nouveaux outils en vue de son exploration qui permettent des diagnostics plus précis et plus rapides. Les technologies de séquençage de nouvelle génération (NGS) sont largement utilisées en clinique et en recherche. Ce séquençage peut être partiel et concerner des séquences cibles (panel de gène ou Whole Exome Sequencing WES) ou concerner la totalité du génome (Whole Genome Sequencing WGS).

La médecine génomique est déjà une réalité. Elle transforme la manière dont on prévient diagnostique, soigne et pronostique l'évolution d'une maladie. Elle concerne plusieurs enjeux (2, 3) :

Enjeu de santé publique : La médecine génomique révolutionne le parcours de soin, et donc l'organisation de la santé publique. Un nombre important de patients affectés par des maladies rares, des cancers, de maladies métaboliques, cardiovasculaires, neurologiques bénéficieront grâce au séquençage en routine de leur génome d'une prise en charge diagnostique et thérapeutique plus personnalisée et précise.

Enjeu scientifique et clinique : L'exploration moléculaire des pathologies jusqu'au bénéfice thérapeutique du patient en passant par la constitution et l'appariement de bases de données multiples qu'il s'agisse de données biologiques, cliniques, paracliniques ou environnementales est un enjeu majeur.

Au regard de la variabilité des différents génomes, il y a un intérêt majeur et primordial de disposer des données du génome Tunisien. En effet, les données importantes générées par le séquençage haut débit de l'ADN des patients, devront être comparées à un génome de référence. Or, chaque population possède des variations et des polymorphismes concernant le fonctionnement de ce génome et qui lui sont spécifiques. C'est pourquoi le projet « *Génome Tunisien* » doit être parmi les priorités de l'Etat tunisien qui doit assurer la plateforme technologique nécessaire pour le réaliser, d'autant plus que les compétences tunisiennes pour réussir ce projet existent déjà.

Enjeu technologique: les sciences et technologies de l'information et la communication sont appelées à converger avec les sciences de la vie et de la santé, la capacité à acquérir, stocker, distribuer, apparier, et interpréter des Big-Data. C'est une véritable filière d'excellence en sciences du calcul et des données en biologie de la santé qui va émerger. Aux côtés de serveurs, cœurs et méthodes de calcul intensif, le développement de logiciels pour l'interprétation et la modélisation de données permettra de relever le défi de la médecine génomique et plus largement de la médecine personnalisée.

Enjeu économique à la fois en termes de coût pour le système de santé mais également d'opportunité de développer une nouvelle filière industrielle. L'importance de cet enjeu économique est résumée dans la phrase suivante qu'on doit à Juan Enriquez, Harvard Business School : «Le moteur économique dominant de ce siècle sera la génomique. Ceux qui n'en apprendront pas le langage ne comprendront pas les forces qui s'exercent sur leur vie».

La forte compétition internationale autour de la médecine génomique s'explique par les enjeux qui lui sont rattachés. Dans le monde, plusieurs pays ont commencé à intégrer la médecine génomique dans leur système de santé : Estonie, Pays-Bas, Slovaquie, France, Japon, Qatar, Arabie Saoudite. Si la Tunisie ne s'organise pas pour développer sa médecine génomique, elle risquera de voir émerger un tourisme médical vers d'autres pays offrant ce type de services à forte valeur médicale ajoutée d'où l'émergence d'une inégalité des citoyens dans l'accès aux soins. De plus, une fuite massive des compétences tunisiennes dans les domaines de la génomique pourra avoir lieu.

Avant l'avènement des techniques de NGS, la stratégie adoptée était de cibler les différents exons du gène le plus fréquemment impliqués dans l'apparition de la maladie par la méthode de séquençage Sanger. Cependant, cette méthode est très coûteuse et lourde, surtout quand il s'agit de pathologie dont le gène muté est de très grande taille, sans hotspot de mutations ou encore avec une hétérogénéité de locus (plusieurs gènes pour un phénotype). Le séquençage classique présente l'inconvénient de ne pouvoir être réalisé que sur des fragments d'une taille moyenne < à 1000 pb. Actuellement, avec l'avènement des techniques de NGS, il est possible de faire l'analyse de tous les gènes impliqués dans le phénotype observé en une seule réaction soit par l'étude d'un panel de gènes soit par la réalisation d'un WES ou WGS, dans le cas où le résultat a été négatif avec le panel (4).

Dans la population tunisienne, la particularité est que plusieurs maladies sont dues à des mutations géniques avec un effet fondateur. Il est alors recommandé d'utiliser la méthode de screening par Sanger d'une mutation bien précise. Si cette mutation est absente, on passera alors au séquençage du reste des exons des gènes si leur nombre est limité. Si les résultats sont négatifs, on indiquera alors la technique du NGS. C'est l'exemple de la maladie de Crigler-Najjar, où il est primordial de cibler la mutation c.1070A>G dans le gène *UGT1A1* (5).

Faut-il faire un séquençage ciblé ou exome ?

Quand les différents gènes d'une pathologie sont parfaitement définis, et que leur nombre n'est pas très important, il est préférable de faire un panel de gènes. Ses avantages sont une meilleure couverture et profondeur de lecture, une analyse bioinformatique plus simplifiée

et un espace de stockage des données minimisé. Son principal inconvénient est la nécessité d'une mise à jour de la librairie. Si la maladie n'est pas bien définie, un séquençage de WES/WGS est préférable, afin d'avoir une meilleure idée sur la totalité des gènes. Les WES/WGS permettent aussi la possibilité d'identifier de nouveaux gènes et d'une ré-analyse bioinformatique des données si nécessaire (des nouvelles découvertes des associations gène-maladie). Comparativement aux panels, les WES/WGS ont une couverture et une qualité de séquençage moindre ; de plus l'analyse bioinformatique est plus lourde et par conséquent un délai de rendu de résultat plus long. Les WES/WGS soulèvent aussi des préoccupations d'ordre éthique telles que les «*incidental findings*» (6, 7). Il est important de souligner l'importance d'une bonne description clinique et une concertation collégiale en prélude à toute indication d'une analyse en NGS.

La génomique a fait émerger un nombre de plus en plus important de variants moléculaires par rapport à la séquence de référence, la totalité de ces variants ne peuvent pas être expliqués et certains restent de signification indéterminée (appelés variants de signification indéterminée ou VSI). Même si des outils de prédiction informatique existent, les prédictions de ces derniers restent limitées et le choix de l'outil reste encore compliqué (8). De plus, les variations du génome ne peuvent expliquer seules, la physiopathologie complète des maladies. Il est donc nécessaire, de compléter cette approche génomique à d'autres approches afin de mieux comprendre l'ensemble des phénomènes biologiques en jeu.

2. LA MÉTABOLOMIQUE

La métabolomique est l'étude de l'ensemble des métabolites (petites molécules avec un poids moléculaire inférieur à 1 kDa) présents dans un organite, une cellule, un tissu, un organe ou un organisme (9). La métabolomique complète les techniques «omics» (génomique, transcriptomique, protéomique) permettant l'étude de propriétés biologiques dynamiques (Figure 1). Ce domaine émergent fournit des données complémentaires pour une approche intégrée afin d'appréhender des voies biologiques dans toutes leurs dimensions (génétique, expression génique, protéines, lipides, ...) (10). La métabolomique a déjà montré un grand potentiel dans les études de toxicité des médicaments, la modélisation de la maladie et le diagnostic, et peut être intégrée aux données génomiques et protéomiques à l'avenir pour fournir une compréhension approfondie des systèmes, des voies et de leurs interactions fonctionnelles dynamiques (11). Par exemple, l'approche intégrée a pu être utilisée pour reconstruire des réseaux biologiques dans le cadre de l'exploration du diabète type 2 en combinant et en confrontant des données omics chez des patients sains contre des patients diabétiques

de type 2.

Il existe deux approches en métabolomique : l'analyse globale, dite «non ciblée» sans a priori et l'analyse «ciblée» (12). L'analyse globale non ciblée permet de détecter un grand nombre de métabolites mais ne permet pas de les quantifier: les concentrations sont alors relatives (rapport numérique entre les groupes étudiés). Cette approche a l'avantage de pouvoir détecter un grand nombre de métabolites connus mais également des métabolites non identifiés (pics inconnus), qu'il est possible d'identifier a posteriori par des expérimentations complémentaires (structures moléculaires ...). Cependant, cette approche ne permet pas de quantification exacte et les analyses manquent parfois de reproductibilité. L'approche ciblée permet la détection et la quantification d'un nombre restreint de métabolites qui est fonction des standards internes utilisés dans le protocole expérimental. L'avantage de cette méthode est la forte reproductibilité et l'obtention de la concentration absolue des métabolites mais se limite à un nombre fini de métabolites. L'utilisation de l'une ou l'autre des approches sera donc dépendante de la question scientifique posée. Si l'objectif est la découverte sans a priori de métabolites impliqués dans un processus biologique donné, l'approche non ciblée apparaît celle de choix. Si la question biologique est précise, la métabolomique ciblée est plus adaptée.

Les deux principaux outils permettant les analyses métabolomiques sont la résonance magnétique nucléaire (RMN) (13) et la chromatographie couplée à la spectrométrie de masse (GC-MS ou LC-MS) (14). L'utilisation de l'approche métabolomique nécessite la combinaison d'une technique séparative d'une part (chromatographie en phase gazeuse GC ou liquide LC), et d'une technique analytique d'autre part (spectrométrie de masse). Il est également possible de travailler sans étape de séparation (en injection directe: DI) ou directement par RMN. L'analyse par SM présente l'avantage d'être plus sensible que la RMN ; elle permet donc l'analyse d'un plus grand nombre de métabolites présents en concentrations moins élevées que l'analyse par RMN, mais aura une moins bonne spécificité que cette dernière (15). Le choix de la technologie repose donc en partie sur les concentrations attendues de métabolites dans le milieu biologique d'intérêt. Par exemple, dans le cadre du dosage des acides biliaires plasmatiques, les concentrations attendues sont de l'ordre de la nanomole par litre et seule la spectrométrie de masse a une sensibilité suffisante pour les détecter dans un volume classique d'échantillon (autour du mL).

Les avancées technologiques en spectrométrie de masse et en RMN permettent aujourd'hui l'analyse simultanée d'un grand nombre de métabolites. Cependant, ces technologies génèrent de grandes quantités de données qui nécessitent l'utilisation d'outils bio-statistiques et bio-

informatiques. Une expérience d'analyse de métabolome se compose donc de plusieurs parties : le design expérimental qui permet la sélection des échantillons (sujets, milieux biologiques), le traitement des échantillons, l'analyse des échantillons, l'identification des métabolites, le traitement des données par des outils bioinformatiques et biostatistiques (16).

La métabolomique permet alors de compléter la génomique, lorsque des profils métaboliques faisant intervenir des métabolites issus d'une voie commune sont modifiés, un séquençage des gènes codants pour les enzymes de cette voie peut alors être envisagé. Ceci permet alors de pouvoir identifier les processus physiopathologiques en cause. Cependant, les variations du génome ne sont pas forcément présentes pour expliquer des variations dynamiques observées en métabolomique. Ces variations peuvent alors être dues à des perturbations intervenant entre la transcription et la traduction, empêchant les ARN messagers de pouvoir être traduits (défaut de transcription, demi-vie raccourci des

ARNm, perturbation de l'épissage ...), ce qui peut être exploré par les techniques de transcriptomique. Par ailleurs, des perturbations peuvent également avoir lieu après la traduction, par des modifications dites «post traductionnelles», qui pourront alors être explorées par des techniques dites de «protéomique».

CONCLUSION

L'avancée des technologies de biologie moléculaire, tel que le séquençage à haut débit, a fait naître l'espoir d'une compréhension totale des mécanismes moléculaires et des explorations cliniques qui en découlent. Cependant, l'expérience nous montre que tout ne peut être expliqué par la génomique, il est donc nécessaire de pouvoir compléter cette méthode par d'autres approches et technologies, par exemple la métabolomique. La combinaison de ces technologies permettra d'améliorer notre compréhension du vivant et la prise en charge des pathologies mais nécessite des investissements en termes de compétences mais aussi de technologies.

RÉFÉRENCES BIBLIOGRAPHIQUES

1. Grzych G. *Génétique*. De Boeck. 2017. <http://www.deboecksuperieur.com/ouvrage/9782807306868-genetique>.
2. Nestor JG, Groopman EE, Gharavi AG. Towards precision nephrology: the opportunities and challenges of genomic medicine. *J Nephrol*. 2018; 31(1):47-60. doi: 10.1007/s40620-017-0448-0.
3. McGuire AL, Gabriel S, Tishkoff SA, Wonkam A, Chakravarti A, Furlong EEM *et al*. The road ahead in genetics and genomics. *Nat Rev Genet*. 2020; 21(10):581-596. doi: 10.1038/s41576-020-0272-6.
4. Ormondroyd E, Border P, Hayward J, Papanikitas A. Genomic health data generation in the UK: a 360 view. *Eur J Hum Genet*. 2021: 1–8. doi: 10.1038/s41431-021-00976-w.
5. Petit FM, Bézieau S, Gajdos V, Parisot F, Scoul C, Capel L *et al*. The Tunisian population history through the Crigler-Najjar type I syndrome. *Eur J Hum Genet*. 2008; 16(7):848-853. doi: 10.1038/sj.ejhg.5201989.
6. Yang Y, Muzny DM, Reid JG, Bainbridge MN, Willis A, Ward PA *et al*. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med*; 369(16):1502-1511. doi: 10.1056/NEJMoal306555.
7. Saudi Mendeliome Group. Comprehensive gene panels provide advantages over clinical exome sequencing for Mendelian diseases. *Genome Biol*. 2015; 16(1):134. doi: 10.1186/s13059-015-0693-2.
8. Grzych G. *Evaluation des outils de prédiction in silico et intérêt des tests fonctionnels dans l'interprétation des variants identifiés par séquençage de nouvelle génération en génétique humaine*. 2018; [Thèse]. Université de Lille.
9. Bujak R, Struck-Lewicka W, Markuszewski MJ, Kaliszan R. Metabolomics for laboratory diagnostics. *J Pharm Biomed Anal*. 2015; 113:108-120. doi: 10.1016/j.jpba.2014.12.017.
10. Mamas M, Dunn WB, Neyses L, Goodacre R. The role of metabolites and metabolomics in clinically applicable biomarkers of disease. *Arch Toxicol*. 2011; 85(1):5-17. doi: 10.1007/s00204-010-0609-6.
11. Yugi K, Kubota H, Hatano A, Kuroda S. Trans-Omics: How To Reconstruct Biochemical Networks Across Multiple 'Omic' Layers. *Trends Biotechnol*. 2016; 34(4):276-290. doi: 10.1016/j.tibtech.2015.12.013.
12. Lu W, Bennett BD, Rabinowitz JD. Analytical strategies for LC-MS-based targeted metabolomics. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2008; 871(2):236-242. doi: 10.1016/j.jchromb.2008.04.031.
13. Li H, Wang L, Yan X, Liu Q, Yu C, Wei H *et al*. A proton nuclear magnetic resonance metabolomics approach for biomarker discovery in nonalcoholic fatty liver disease. *J Proteome Res*. 2011; 10(6):2797-2806. doi: 10.1021/pr200047c.
14. Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. *Mass Spectrom Rev*. 2007 Jan-Feb;26(1):51-78. doi: 10.1002/mas.20108.
15. Ritchie MD, Holzinger ER, Li R, Pendergrass SA, Kim D. Methods of integrating data to uncover genotype-phenotype interactions. *Nat Rev Genet*. 2015; 16(2):85-97. doi: 10.1038/nrg3868.
16. Barupal DK, Fan S, Fiehn O. Integrating bioinformatics approaches for a comprehensive interpretation of metabolomics datasets. *Curr Opin Biotechnol*. 2018; 54:1-9. doi: 10.1016/j.copbio.2018.01.010.