



Quel est l'apport de la biologie pour les patientes ?

Les outils de la biologie moléculaire

Molecular oncology in genetic labs

Mots-clés : Génomique – Biopuces – Tests – Prédiction thérapeutique – Biologie moléculaire.

Keywords: *Genomics – Microarrays – Tests – Therapeutic prediction – Molecular biology.*

N. Sévenet*

L'augmentation importante de la quantité d'analyses de biologie moléculaire à réaliser sur une biopsie tumorale ou sur un fragment d'exérèse élargit d'un regard nouveau cette discipline récente dans les laboratoires de biologie, essentiellement hospitaliers. En effet, le corpus des études de génétique moléculaire publiées à propos du cancer du sein ne cesse de croître et démontre l'utilité de cette approche méthodologique dans le dépistage précoce, le diagnostic, le pronostic et le traitement du cancer du sein (1). En introduction au congrès 2009 de l'American Society of Clinical Oncology (ASCO), Richard Schilsky présentait l'ensemble des moyens à mettre en œuvre pour une médecine assurant un confort au patient cancéreux, au premier rang desquels les techniques de génétique moléculaire, qu'elles soient classiques (marqueur) ou innovantes (génomique et ses corollaires) [2].

À mesure que le génome humain est séquencé, que des cartes complètes d'interaction protéique sont publiées, la communauté scientifique et médicale découvre la complexité du réseau cellulaire eucaryote. Ainsi les techniques d'analyse globale sont-elles devenues essentielles, tant d'un point de vue technique qu'économique. L'utilisation de la biologie moléculaire dans le cancer du sein – qui constitue une pathologie remarquablement hétérogène – permet une meilleure compréhension des mécanismes d'oncogenèse mais également une nouvelle description nosologique complétant les classifications anatomopathologiques; enfin, ces techniques visent à établir de nouveaux critères thérapeutiques décisionnels au côté de ceux utilisés dans la pratique clinique quotidienne (âge, taille tumorale, envahissement ganglionnaire, grade, statut des récepteurs hormonaux...). Cette courte revue, basée sur une présentation thématique des objectifs de la biologie moléculaire dans le cancer du sein, ne se veut pas exhaustive mais pointe les applications cliniques des techniques de biologie moléculaire au sens large.

Intérêt de la biologie moléculaire dans le cancer du sein

La recherche de marqueurs génétiques de susceptibilité au cancer du sein

En 1994 et 1995, les deux gènes majeurs de susceptibilité au cancer du sein, BRCA1 et BRCA2 (3, 4), étaient caractérisés après une stratégie classique de liaison génétique et de clonage positionnel. Ces découvertes fondamentales, qui expliquaient une partie

(20-25%) des syndromes de prédisposition héréditaire au cancer du sein, permirent de penser que cette tumeur, bien qu'hétérogène dans sa description anatomopathologique, survenait sur des terrains génétiques d'altérations monogéniques peu fréquentes dans la population et sur des gènes à forte pénétrance. Force est de constater, 15 années plus tard, qu'il n'en est rien. Quelques gènes supplémentaires ont été isolés par les mêmes approches de liaison génétique et de clonage, dont les mutations confèrent un risque tumoral élevé (pénétrance élevée) et sont également à l'origine de syndromes de prédisposition au cancer du sein: PTEN et la maladie de Cowden, p53 et le syndrome de Li-Fraumeni, LKB1/STK11 et le syndrome de Peutz-Jeghers (*pour revue*, 5). Par des approches de génétique des populations ainsi que des analyses de liaison génétique dans les familles, une série d'autres variants a été caractérisée dont le risque tumoral, induit par leur altération, est modéré (pénétrance modérée): CHEK2 1100delC (une mutation spécifique), ATM, BRIP1, PALB2. En raison du faible risque tumoral induit et de leur rareté dans la population, leur implication pathogénique se limite à moins de 3% des syndromes héréditaires de cancer du sein.

Finalement, la grande majorité des cancers du sein obéit à un modèle polygénique résultant probablement de la combinaison de plusieurs variants dont la fréquence dans la population est importante (plus de 5%) mais dont l'implication pathogénique est remarquablement faible (variants de faible pénétrance). Les progrès technologiques majeurs dans les domaines du grand séquençage et du génotypage, dont le séquençage du génome humain et la caractérisation dans plusieurs populations de dizaines de millions de variations nucléotidiques polymorphiques (SNP, *single-nucleotide polymorphism*) issues du programme de génotypage HapMap, ont permis aux scientifiques et aux cliniciens de mener de nouvelles investigations à très grande échelle et sans hypothèse préalable, les études d'association pan-génome (*genome-wide association studies*), en vue de caractériser des variants fréquents dans la population générale et de les lier au cancer du sein sporadique ou familial (*pour revues*, 6, 7).

Trois études d'association pan-génome menées dans le cancer du sein ont ainsi caractérisé 7 loci de susceptibilité: FGFR2, TNRC9, MAP3K1, LSP1, 8q24, 2q35, et 5p12. Aucun de ces variants ne se trouve dans ou à proximité d'une région codante (exon d'un gène), mais dans des régions introniques, voire intergéniques. L'association la plus forte est un SNP situé dans l'intron 2 du gène FGFR2, dont l'allèle causal est présent chez environ 40% de la population européenne, et dont le génotype homozygote rare entraîne un risque tumoral mammaire chez les femmes de 1,63 comparati-

* Institut Bergonié, laboratoire de génétique constitutionnelle, Bordeaux.

vement à l'autre génotype (homozygote de l'allèle non causal). Bien que la caractérisation de FGFR2 résulte d'une approche sans hypothèse, l'implication de ce gène – qui code une tyrosine kinase transmembranaire – dans la tumorigenèse mammaire avait déjà été décrite, ainsi que sa surexpression dans plus de 10 % des tumeurs du sein ou sa corrélation aux tumeurs RE+ (*pour revue*, 8). Ces résultats démontrent comment sont caractérisés, par des approches techniques différentes, des mécanismes biologiques similaires.

Les classifications moléculaires

Si la première publication rapportant le dépôt d'acide nucléique sur une lame de verre traitée date de 1991, la première application de cette technologie afin d'étudier en une seule expérimentation les échantillons complexes d'ADN a été réalisée par le laboratoire de Patrick Brown à l'Université de Stanford aux États-Unis, dès 1995 (9). C'est en cancérologie que l'utilisation des biopuces est désormais la plus courante. Ainsi, il est possible de déterminer l'état d'expression de groupes tumoraux définis et homogènes (*profiling*), avec pour finalité une classification des tumeurs basée sur ces profils d'expression (*clustering*) [10]. Il est important de noter que toutes ces signatures moléculaires générées à partir des profils d'expression génique de tumeurs recoupent les classifications anatomopathologiques et/ou immuno-histochimiques, mais font également apparaître de nouveaux sous-groupes tumoraux de définition ou d'évolution différente des précédents. Cela permet de confirmer en partie la réalité phénotypique des résultats des biopuces.

Les biopuces à ADN (*microarrays*) définissent un ensemble de procédés caractérisés par un même schéma expérimental : un acide nucléique (ADN/ARN) à caractériser (la cible) est extrait d'un groupe homogène de cellules ou d'un tissu d'intérêt, puis "marqué" (coloré avec un fluorochrome ou un traceur radioactif) et cohybridé avec un ADN/ARN provenant d'un groupe cellulaire contrôle, sur un support solide (lame de verre, membrane de nylon, etc.) présentant à sa surface entre 100 et 1 000 000 de spots d'acides nucléiques immobilisés (les sondes). Après hybridation longue, puis une série de lavages stringents, le support est placé dans un scanner (ou un détecteur de radioactivité) soumis à un laser excitant les molécules de fluorescence des cibles retenues par hybridation spécifique sur les sondes immobilisées. Le signal émis est capturé et converti informatiquement en signal d'intensité (échelle de gris). Ce signal est retraité informatiquement avec attribution de couleurs standard (cyanine 5 : rouge ; cyanine 3 : vert). Selon l'intensité et la caractéristique du signal émis, l'expérimentateur connaît ainsi précisément le type d'acide nucléique (la sonde est connue) et sa quantité relative par rapport aux autres dans la solution complexe d'acides nucléiques de l'échantillon d'intérêt. Après analyse de l'image, les résultats sont traités statistiquement au moyen de logiciels et d'algorithmes adaptés. Suivant le type de biopuces à ADN utilisé, il sera possible d'étudier le génome ou le transcriptome d'un groupe homogène de cellules (ou d'un tissu). Les biopuces à ADN génomique sont notamment utilisées en hybridation génomique comparative (CGH, *comparative genomic hybridization*) afin de tester l'intégrité du génome de cellules d'intérêt. L'intensité et la couleur du signal émis en fin d'hybridation par chacun des spots d'ADN génomique permettent de déterminer le type d'altération génomique pour ce fragment d'ADN. Par exemple, un spot rouge démontrera une surreprésentation de cette région d'ADN génomique dans les cellules étudiées (amplification, duplication), tandis qu'un spot vert démontrera une perte de matériel génétique

pour cette région d'ADN génomique (déletion, perte d'allèle). Les biopuces d'expression permettent l'étude du transcriptome. Étant donné que la plupart des variations phénotypiques cellulaires sont liées à des modifications du transcriptome, il paraît pertinent de documenter l'expression génique d'un groupe cellulaire tumoral. Les biopuces d'expression sont probablement les lames les plus utilisées compte tenu de leur rapport qualité/prix avantageux et de la grande quantité d'informations pertinentes qu'elles génèrent. L'ARN total ou messenger (la cible), est extrait de deux groupes de cellules (ou tissu) : un groupe contrôle et le groupe de cellules à étudier. Après hybridation, lavages, puis lecture, l'intensité et la coloration des spots d'acides nucléiques du support reflètent la quantité relative des ARN les uns par rapport aux autres.

Les principaux résultats en CGH tendent à montrer que les tumeurs du sein sporadiques se répartissent essentiellement en quatre groupes : les tumeurs de génome stable, les tumeurs avec des réarrangements chromosomiques récurrents (1p, 16q), les tumeurs avec des pertes/gains de bras entiers chromosomiques, et les tumeurs de génome très remanié, présentant une grande instabilité génétique et dans lesquelles on retrouve la majorité des tumeurs héréditaires mutées pour BRCA1 ou BRCA2 ainsi que les tumeurs de pronostic défavorable (*pour revue*, 11).

Les études d'expression génique ont démontré également quatre classes tumorales : les tumeurs du sein *basal-like*, les tumeurs amplifiées pour le gène ERBB2 (HER2), les tumeurs lumineales A, et les tumeurs lumineales B. Ces groupes correspondent assez bien aux descriptions cliniques basées sur le statut des récepteurs hormonaux et de la protéine HER2, ainsi que sur le grade histologique ou les marqueurs de prolifération cellulaire. Ces résultats feront l'objet de la présentation d'A. Vincent-Salomon "La nouvelle classification des cancers du sein : apport de la pathologie moléculaire".

Le pronostic clinique

Utilisant la technologie des biopuces d'expression, de nombreuses équipes ont déterminé des signatures moléculaires prédisant l'évolution clinique avec une précision plus importante que les critères habituellement utilisés (*pour revue*, 12). Selon le type d'étude rétrospective menée et, surtout, selon le schéma expérimental utilisé, les signatures moléculaires sont différentes et ne présentent que très peu de gènes en commun. Enfin, la plupart de ces signatures moléculaires pronostiques sont désormais mises en œuvre par des *start-up* commercialisant des "tests génomiques". Dans ces cas-là, l'analyse bio-informatique et statistique l'emporte sur la partie biologie moléculaire expérimentale.

Ainsi, une approche *top-down* supervisée (l'analyse statistique était contrainte par les données d'évolution tumorale) a permis la caractérisation d'un set de 78 gènes (MammaPrint®, Agendia) facilitant la classification en bon ou mauvais pronostic des femmes présentant un cancer du sein sans atteinte ganglionnaire, de moins de 5 cm, de stade I ou II et âgées de moins de 61 ans ("signature d'Amsterdam"). Utilisant une approche de type *bottom-up*, une autre équipe a caractérisé une signature de 97 gènes appelée "grade génomique", différenciant un bas grade génomique et un haut grade, correspondant peu ou prou au "grading" histologique (*pour revue*, 13).

Deux autres signatures génomiques sont issues de la RT-PCR semi-quantitative, technologie basée sur le principe de la PCR, permettant l'estimation quantitative du nombre de transcrits (donc du niveau d'expression d'un gène) au sein d'un groupe

cellulaire donné. Utilisant 21 gènes (*Recurrence score*, Oncotype DX®, Genomic Health) ou le rapport d'expression de 2 gènes (HOXB13/IL17R) [*2-gene ratio*, Theros®, Biotheranostics], ces signatures définissent des scores de rechute tumorale. Deux essais cliniques (TAILORx et MINDACT) sont en cours afin de valider les signatures pronostiques utilisant respectivement Oncotype DX® et MammaPrint®.

La prédiction thérapeutique

Établir une signature d'expression génique de prédiction de réponse à la chimiothérapie uniquement sur les cellules de la tumeur est probablement trop restrictif, puisque l'environnement tumoral, qu'il soit cellulaire (stroma entourant la tumeur) ou organique (caractéristiques de l'hôte en termes de pharmacogénétique et métabolisme médicamenteux), n'est pas pris en compte (14). La plupart des études se concentrent ainsi sur l'étude de la sensibilité des tumeurs du sein à la chimiothérapie préopératoire (situation néoadjuvante) et analysent la réponse pathologique complète au niveau anatomopathologique. Une étude génomique menée sur 133 patientes présentant une tumeur du sein de stade I, II ou III et ayant reçu, avant la chirurgie, du paclitaxel avec une combinaison de 5-fluorouracile, doxorubicine et cyclophosphamide, a permis de caractériser une signature de 30 gènes prédicteurs d'une réponse pathologique complète. Cette signature ainsi que quelques autres issues d'études pilotes ne sont pas commercialisées et nécessitent d'être validées dans des essais cliniques multicentriques.

Et dans le futur ?

L'extraordinaire avancée de la biologie moléculaire dans le cancer du sein ces dernières années a consisté dans le développement de ces techniques ouvertes ou semi-ouvertes, et dans la recherche de déterminants génétiques sans hypothèse préconçue. Le développement des signatures moléculaires apporte aux cliniciens une aide décisionnelle de thérapie pour les tumeurs du sein où les critères anatomopathologiques sont intermédiaires.

Les progrès technologiques majeurs attendus reposent sur le séquençage à très haut débit ainsi que sur le génotypage de millions de SNP par tumeur (15). Parallèlement, apparaissent sur le marché des automates de dissection des fragments tumoraux de plus en plus efficaces, permettant d'isoler quelques cellules tumorales. Enfin, l'étude des espèces d'ARN différentes des classiques ARN messagers (microARN, ARN issus d'ADN antisens, ARN de pseudogènes, ARN issus de séquences non codantes) devrait contribuer à une nouvelle approche diagnostique.

À la lumière de ces découvertes, une nouvelle spécialité de médecine moléculaire émerge, avec comme finalité, notamment, la prise en compte de l'ensemble des critères moléculaires d'une tumeur dans le diagnostic, le pronostic et la prédiction thérapeutique. Cela nécessite une coopération renforcée entre biologistes, généticiens, anatomopathologistes et cliniciens, et des temps de réponse diagnostique courts et intégrés en routine hospitalière. ■

Références bibliographiques

- [1] Olopade OI, Grushko TA, Nanda R, Huo D. *Advances in breast cancer: pathways to personalized medicine. Clin Cancer Res* 2008;14:7988-99.
- [2] Schilsky R. *Personalizing cancer care: American Society of Clinical Oncology presidential address 2009. J Clin Oncol* 2009;27:3725-30.
- [3] Miki Y, Swensen J, Shattuck-Eidens D et al. *A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science* 1994;266:66-71.
- [4] Wooster R, Neuhausen SL, Mangion J et al. *Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. Science* 1994;265:2088-90.
- [5] Garcia-Closas M, Chanock S. *Genetic susceptibility loci for breast cancer by estrogen receptor status. Clin Cancer Res* 2008;14:8000-8.
- [6] Pharoah PDP, Antoniou AC, Easton DF, Ponder BAJ. *Polygenes, risk prediction and targeted prevention of breast cancer. N Engl J Med* 2008;358:2796-803.
- [7] Hardy J, Singleton A. *Genome wide association studies and human disease. N Engl J Med* 2009;360:1759-68.
- [8] Katoh M. *Cancer genomics and genetics of FGFR2 (review). Int J Oncol* 2008;33:233-7.
- [9] Fodor SP, Read JL, Pirrung MC et al. *Light-directed, spatially addressable parallel chemical synthesis. Science* 1991;251:767-73.
- [10] Perou CM, Sorlie T, Eisen MB et al. *Molecular portraits of human breast tumours. Nature* 2000;406:747-52.
- [11] Sotiriou CJ, Pusztai L. *Molecular origins of cancer: gene-expression signatures in breast cancer. New Engl J Med* ;360:790-800.
- [12] Sotiriou C, Piccart MJ. *Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care? Nat Rev Cancer* 2007;7:545-53.
- [13] Oakman C, Bessi S, Zafarana E, Galardi F, Biganzoli L, Di Leo A. *Recent advances in systemic therapy: new diagnostics and biological predictors of outcome in early breast cancer. Breast Cancer Res* 2009;11:205-15.
- [14] Farmer P, Bonnefoi H, Anderle P et al. *A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer. Nat Medicine* 2009;15:68-74.
- [15] Tucker T, Marra M, Friedman JM. *Massively parallel sequencing: the next big thing in genetic medicine. Am J Hum Genet* 2009;85:142-54.