

# Application de la génétique multifactorielle

Gilles Thomas

► La génétique multifactorielle présente une difficulté par rapport à la génétique classique, telle qu'on l'entendait jusqu'à présent, qui était une génétique de type monogénique, la plupart du temps très pénétrante. C'est-à-dire que la présence d'une altération génétique avait la plupart du temps une traduction phénotypique. Dans le domaine du cancer, la plupart des prédispositions connues se transmettent sur le mode dominant, si bien que l'étude des structures familiales a permis d'envisager des études paramétrées, qui ont elles-mêmes permis d'identifier les gènes responsables. Dans l'hérédité multifactorielle, on cherche des gènes selon deux paramètres : la fréquence de l'allèle « à risque », qui varie de 0 à 1, et le niveau de risque par rapport au risque normal, qui diffère de 1 selon la probabilité de l'allèle à augmenter, par sa présence, l'apparition du phénotype. Par exemple, à un allèle dont la présence augmente le risque de cancer de 20 % par rapport au risque de la population générale sera attribué un risque de 1,2. ◀



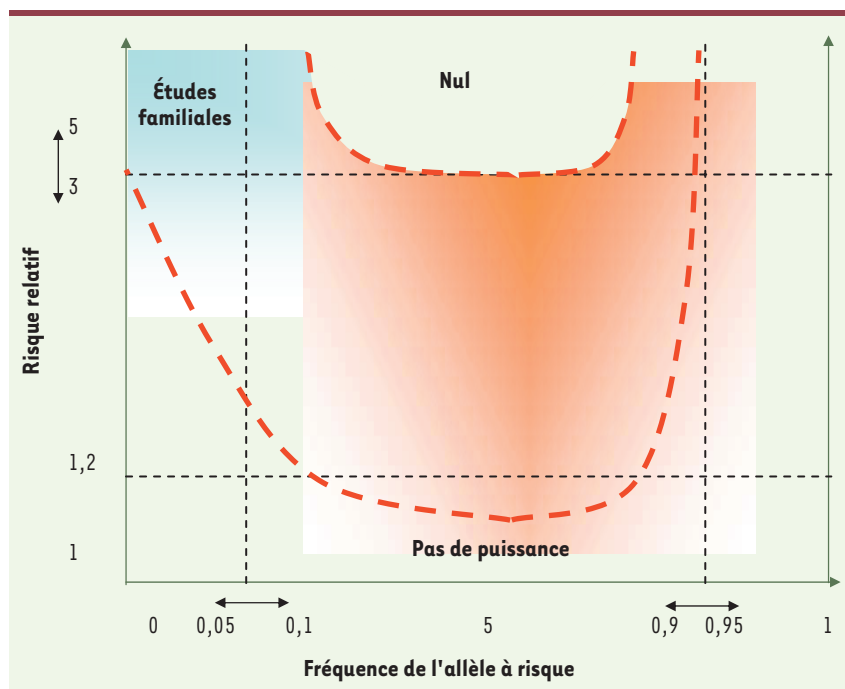
Synergie Lyon Cancer,  
150, cours Albert Thomas,  
69008 Lyon, France.  
[gilles.thomas@inserm.fr](mailto:gilles.thomas@inserm.fr)

## Les études d'association dans les cancers du sein et de la prostate

Les études familiales explorent un espace où les fréquences alléliques sont relativement rares, mais où les risques relatifs sont élevés, qui se traduisent par la présence de structures familiales (*Figure 1*). Les recherches génomiques explorent le reste, et l'on sait qu'en dessous d'un certain risque, même si la fréquence allélique est importante, d'énormes groupes de cas et de témoins seraient nécessaires pour les repérer. Les recherches d'associations se situent donc dans la zone des fréquences alléliques et des risques relatifs intermédiaires.

Les travaux sur le cancer du côlon ont été réalisés grâce à la création d'un consortium européen autour de l'équipe de Richard Houlston, qui a permis le recueil

de la ressource nécessaire. Cette approche a également été appliquée pour les cancers du sein et de la prostate et un ensemble de 14 000 cas de cancers de la prostate et presque 13 000 cas de témoins a été réuni par un consortium international autour du NCI. L'exigence statistique pour déclarer par cette approche un locus associé à l'augmentation du risque de cancer est actuellement de  $10^{-7}$ . En effet, les observations ne sont pas reproductibles pour des seuils moindres [1-4]. Les trois grandes études sur le cancer de la prostate ont permis de repérer 15 associations sur le génome. Huit d'entre elles étaient trouvées dans deux études, dont deux dans les trois. Trois études ont été réalisées pour confirmer les observations sur 3 locus. En revanche, il y a une discordance entre deux études pour un locus situé sur le chromosome 19 dans le gène *KLK3*. La réplification de la première observation a de plus diminué sa valeur statistique de manière importante. La raison de cette discordance tient à un biais dans la sélection des cas, basée précisément sur le dosage sanguin du PSA, codé par le gène *KLK3*. En effet, les témoins de l'étude révélant l'association avaient été choisis pour un taux faible de PSA et non pas pour l'absence clinique de cancer de la prostate ; l'absence de ce critère de sélection dans la réplification explique la baisse de signification de l'observation initiale. Cette étude amène donc des conclusions de nature différente et souligne l'importance d'intégrer les critères de recrutement dans les calculs de risque relatif. Les observations faites dans les études sur le cancer du sein sont comparables, avec actuellement six locus trouvés associés dans au moins



**Figure 1.** Paramètres influençant le type et l'efficacité des recherches de gènes de prédisposition.

et restreindre la région où chercher un polymorphisme fonctionnel. Mais il est important de coupler cette démarche avec des approches fonctionnelles. Une hypothèse comparable a été formulée pour le locus à proximité de *FGFR2* et des différences d'expression ont pu être mises en relation avec le génotype. Pour argument, deux des SNP liés au cancer du sein sont localisés dans des facteurs de transcription, *CEBPB* et *RUNX*. L'étude du statut des récepteurs hormonaux révèle également que l'allèle à risque de *FGFR2* pourrait favoriser le développement de cancers RE\*. Chez les femmes porteuses d'une mutation constitutionnelle du gène *BRCA2*, le génotype du locus *FGFR2*

influe sur le risque, alors qu'il n'a aucun effet en cas de mutation de *BRCA1*. La même observation est faite pour *MAP3*-kinase, mais pas pour *TOX3*. Il y a donc des indications pour penser que ces locus interagissent de manière sensible.

### Les applications prévisibles

La première question qui se pose naturellement concerne les capacités d'identifier des personnes dont la prise en charge préventive justifierait d'être modifiée. Plusieurs sociétés privées proposent aujourd'hui la caractérisation de cinq locus favorisant le cancer de la prostate. Les conditions de réalisation de ces tests ne comportent aucun des trois paramètres nécessaires à une application médicale quelconque, à savoir l'adéquation entre les indications du test, sa nature et l'impact du résultat sur recommandations antérieures.

En d'autres termes, la réalisation d'un test génétique est-elle susceptible de modifier l'ordre des différentes options de prise en charge existantes avant le test ? Lorsqu'il s'agit de chercher une mutation sur un gène *MMR*, *BRCA* ou *APC*, le test est proposé à une personne qui a un risque d'au moins 10 % de l'avoir d'après les caractéristiques familiales. La présence de cette mutation fait passer son risque cumulé, pour le cancer du sein par exemple, de 20 % à 70 %, et les options de prise en charge sont modifiées dès lors qu'une mutation est effectivement identifiée.

La présence de l'allèle « à risque » de *FGFR2* dans la population générale fera passer le risque cumulé de

deux études. Les études encore en cours devraient permettre d'identifier quelques locus supplémentaires. Les risques relatifs se situent tous aux alentours de 1,2. Les fréquences des allèles à risque sont en revanche très variables, certaines étant très élevées, ce qui indique qu'ils ne sont pas soumis à une pression de sélection. En effet, ces cancers sont de survenue tardive, et même dans les formes de mauvais pronostic, ils n'entravent pas les capacités normales de reproduction.

### Recherche des variations fonctionnelles

L'enjeu ensuite est d'identifier les variations fonctionnelles justifiant les associations observées. Dans le cancer du sein, quelques locus sont d'emblée liés à de bons gènes candidats, comme *FGF10*, *FGFR2*, *MYC* par exemple. D'autres locus sont moins évocateurs *a priori*, c'est le cas de *MAP3*-kinase et *TOX3*. Dans le cancer de la prostate, deux locus orientent vers des gènes impliqués dans le diabète de type II. Enfin, il y a des locus situés dans des régions vierges de gènes, qui pourraient être des régions de contrôle de l'expression génique.

L'avenir immédiat est de rendre les observations individuelles robustes, en regroupant les données comparables des différentes études, et en complétant les études partielles. Il est également important de prévoir des études sur les tumeurs, qui cherchent à corréliser les génotypes et les données d'expression, pour lesquelles des panels de 500 cas environ seraient suffisants, des modèles cellulaires, animaux...

Une observation intéressante, bien que surprenante *a priori*, a été celle de l'association d'une région avec trois localisations tumorales, le côlon, le sein et la prostate. Un des locus de la région est lié au côlon et à la prostate. Cette région contient le gène *MYC* et il est possible qu'il y ait une relation, sans qu'aucune variation d'expression de *MYC* n'ait été mise en évidence en fonction du statut allélique. Il est donc nécessaire de poursuivre le génotypage pour préciser l'association

cancer du sein de 12,6 à 14 %, donc aucune modification de la prise en charge n'en découlera. L'existence d'un autre allèle à risque, TOX3, pourra augmenter encore ce risque et l'amener à 22,6 % ; cette constatation n'est pas non plus susceptible de modifier de manière importante la prise en charge. Elle concernerait 1,5 % des femmes testées. Il convient donc de caractériser d'autres locus. Mais il y a déjà 9 combinaisons génotypiques avec 2 locus alors qu'il n'y en a que 3 avec un seul. La connaissance précise du risque associé à chacun de ces 9 génotypes nécessite des études épidémiologiques de très grande envergure, non encore réalisées. Le génotypage de 6 locus offre 729 combinaisons, et les 9 connus actuellement presque 20 000. Il n'est donc pas réaliste d'évaluer un risque individuel en dehors d'un modèle. Le modèle de risque le plus simple est de considérer les risques associés à chaque allèle, d'un même locus ou de locus distincts, de manière indépendante, sans interaction. Le risque maximum associé à la présence de 6 allèles à risque est d'environ 33 %. Il concerne 2 femmes sur 10 000. Il sera donc nécessaire de faire 10 000 tests pour observer 2 femmes qui auraient ce risque, et noter le développement ou non d'un cancer du sein. Pour le cancer de la prostate, les constatations sont identiques, avec un proportion de 3 hommes sur 1 000 dont le risque cumulé serait d'au moins 40 % par la présence de 6 allèles à risque. Dans ce dernier cas, peut-être pourrait-on trouver une justification médicale au test. Mais c'est faire l'hypothèse que le modèle est juste...

### En conclusion

Il est essentiel de différer les tests génétiques en population. L'identification des variants fonctionnels donnera probablement des indications sur leur mode d'action, dominant, récessif, additif, des paramètres qu'il faudra intégrer à la construction du modèle, sans parler des interactions entre locus. Ces nouveaux modèles, moins théoriques, devront être ensuite validés dans le cadre d'études cliniques et de protocoles consensuels. Le développement des tests dans des structures privées prive à la fois les personnes qui s'y soumettent, et la

communauté scientifique et médicale, de toute capacité d'interprétation des résultats. Néanmoins, les études d'association sont importantes sur le plan cognitif, car elles apportent de nouveaux moyens d'explorer les risques et les mécanismes d'apparition du cancer et permettront très certainement à court terme une dissection précise des différentes voies biologiques de la tumorigénèse. ♦

### SUMMARY

#### Genome-wide association approach

The genome-wide association approach has been the most powerful and efficient study design thus far in identifying genetic variants that are associated with complex human diseases. This approach became feasible as the result of several key advancements in genetic knowledge, genotyping technologies, statistical analysis algorithms and the availability of large collections of cases and controls. With all these necessary tools in hand, many genome-wide association studies were recently completed, and many more studies which will explore the genetic basis of various complex diseases and quantitative traits are soon to come. The recent large genotyping studies have identified a new repertoire of cancer susceptibility genes and loci which are characterized by common risk alleles. Many of these loci were detected at low power, indicating that many further loci will probably be detected with larger studies. For the most part, the loci were not previously suspected to be related to carcinogenesis, and point to new disease mechanisms. The risks conferred by the susceptibility alleles are low, generally 1.3-fold or less. The combined effects may, however, be sufficiently large to be useful for risk prediction, and targeted screening and prevention, particularly as more loci are identified. ♦

### RÉFÉRENCES

1. Easton DF, Eeles RA. Genome-wide association studies in cancer. *Hum Mol Genet* 2008 ; 17 : 9-15.
2. Hemminki K, Försti A, Lorenzo Bermejo J. New cancer susceptibility loci: population and familial risks. *Int J Cancer* 2008 ; 123 : 1726-9.
3. Torkamani A, Topol EJ, Schork NJ. Pathway analysis of seven common diseases assessed by genome-wide association. *Genomics* 2008 ; 92 : 265-72.
4. Seng KC, Seng CK. The success of the genome-wide association approach: a brief story of a long struggle. *Eur J Hum Genet* 2008 ; 16 : 554-64.

---

### TIRÉS À PART

G. Thomas