Exploiter la séquence du génome

17

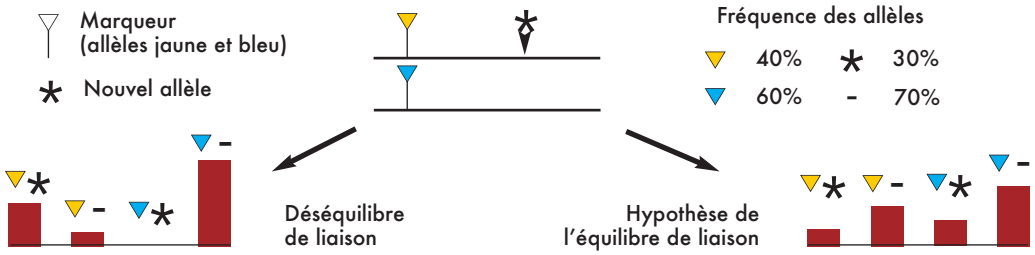
II. La part génétique des maladies communes

Au cours des années 1990, les chercheurs ont réussi à identifier de nombreux gènes de maladies génétiques par l'approche du clonage positionnel *(voir la fiche précédente)*. Enhardis par ces succès, ils ont appliqué cette même stratégie aux maladies "communes", telles que le diabète, les maladies cardiovasculaires ou la maladie d'Alzheimer. Les résultats furent décevants. Ces maladies, dites également multifactorielles ou polygéniques, constituent en effet un défi pour les généticiens : elles possèdent certes une composante génétique, mais celle-ci est difficile à définir. De nombreux gènes semblent contribuer chacun pour une faible part au risque de développer la maladie ; qui plus est, ces gènes interagiraient de façon complexe entre eux et avec les facteurs liés à l'environnement et au style de vie. Démêler cet écheveau est pourtant crucial si l'on veut mieux comprendre les mécanismes intimes des maladies, identifier de nouvelles cibles thérapeutiques et définir la part de l'environnement afin de mettre en place une prévention efficace.

Dans le cas des maladies génétiques, le fait d'hériter d'une ou de deux copies (ou allèles) mutées du gène suffit en général pour que la maladie se déclare. Cette corrélation forte entre génotype et phénotype explique le succès de l'analyse de liaison dans la recherche de ces gènes *(voir la fiche précédente)*. Dans le cas des maladies communes, au contraire, il s'agit de rechercher des allèles qui ne sont ni nécessaires ni suffisants au déclenchement de la maladie ; autrement dit, ces allèles ne sont pas présents chez tous les malades, et peuvent être présents chez certaines personnes bien portantes. Ils ne créent qu'une susceptibilité, ou prédisposition, vis-à-vis de la maladie. C'est ce qui explique qu'ils soient passés jusqu'ici, pour l'essentiel, entre les mailles des filets des généticiens.

L'approche positionnelle a certes permis d'identifier plusieurs gènes responsables de formes "familiales", à forte composante héréditaire, de maladies telles que le diabète, le cancer du sein ou la maladie d'Alzheimer. Ces formes sont rares, mais l'analyse du gène identifié peut éclairer les mécanismes de la maladie dans sa version la plus commune. Au delà de ces cas particuliers, les généticiens ont utilisé l'analyse de liaison pour incriminer des régions chromosomiques dans de nombreuses maladies communes. Toutefois, ces résultats se sont avérés fragiles, difficiles à reproduire. Qui plus est, les régions délimitées par ces travaux mesurent souvent plusieurs dizaines de millions de bases ! Avec de la chance, les chercheurs sont guidés par divers indices vers un gène candidat déjà localisé dans l'intervalle : ils recherchent alors des variations de séquence, dans le gène ou aux alentours, qui soient associées de façon significative à la maladie. En l'absence d'indices sur l'identité possible du gène coupable, les chercheurs sont réduits à explorer les variations de séquence dans un immense intervalle.

(suite au dos)

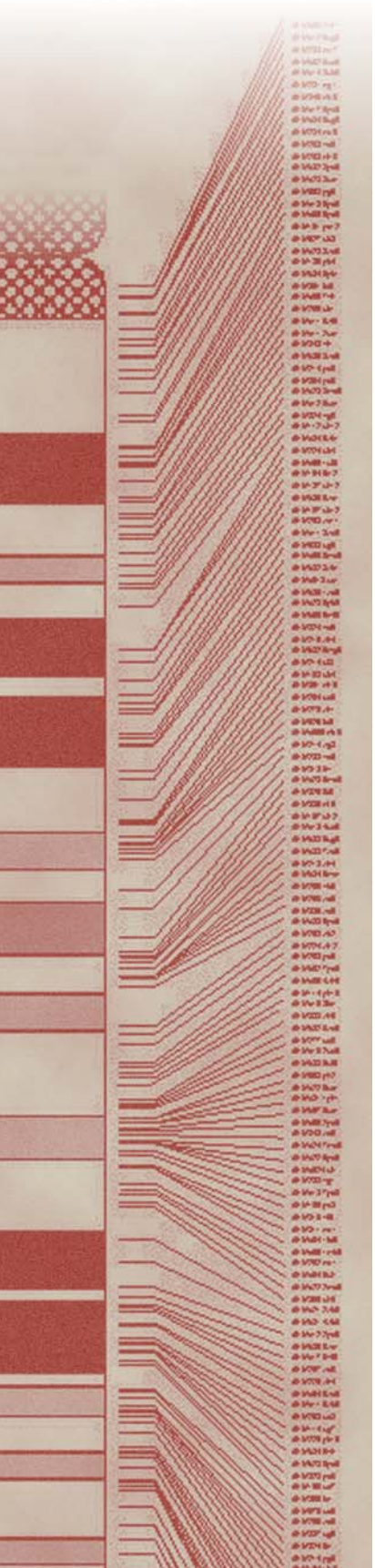


Exemple d'un allèle (astérisque) en déséquilibre de liaison absolu avec un allèle d'un marqueur près duquel il est apparu (triangle orange). A gauche, les fréquences observées des différents haplotypes : l'allèle * est toujours trouvé avec l'allèle jaune du marqueur. A droite, les fréquences attendues si les combinaisons d'allèles étaient dues au seul hasard. Ce déséquilibre de liaison diminuera progressivement à mesure que des événements de recombinaison interviendront entre les deux loci.

Exploiter la séquence du génome

17

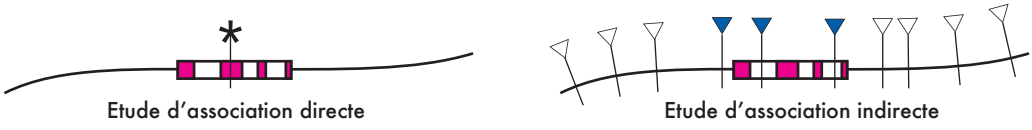
II. La part génétique des maladies communes (suite)



Les variations de séquence dans le génome humain sont pour l'essentiel des changements d'une seule base, nommés SNP ou snips (pour *Single Nucleotide Polymorphism*). Deux chromosomes homologues de deux individus choisis au hasard diffèrent en moyenne en une position toutes les 1 000 bases : chacun d'entre nous est donc génétiquement identique à 99,9% à n'importe quel autre être humain. Si l'on compare plus de deux individus, on trouvera davantage de positions variables dans la séquence. Si l'on considère l'ensemble de la population humaine, on estime qu'il existe au moins 10 millions de snips dans notre génome, soit un snip toutes les 300 bases. Seule une petite fraction de ces snips se traduit par des variations des caractéristiques des individus, telles que la stature ou la susceptibilité à diverses maladies. Ce sont bien sûr les plus intéressants, mais la connaissance des autres snips est aussi très utile aux généticiens. Ils procurent des points de repère nombreux et régulièrement espacés le long du génome, à l'aide desquels on espère localiser les snips fonctionnellement significatifs.

Le projet génome humain a rendu de telles études envisageables. Plus de 4 millions de snips ont déjà été déterminés, nombre d'entre eux dans le cadre du projet, et positionnés sur la séquence du génome humain. Ils procurent un maillage beaucoup plus fin que les marqueurs des cartes génétiques utilisées pour l'analyse de liaison. Les généticiens comptent s'en servir de deux façons : dans les études d'association "directes", on tente de découvrir sans détour, dans et autour de gènes candidats, le ou les polymorphismes influençant le caractère étudié ; dans les études d'association "indirecte", on analyse des snips plus largement espacés, à l'échelle de grandes régions génomiques, dans l'espoir que certains soient fortement associés au mystérieux polymorphisme fonctionnel et délimitent la zone où le rechercher. Une telle association est fondée sur l'existence d'un "déséquilibre de liaison" entre l'allèle recherché et les allèles près desquels il est apparu *(voir la figure)* : jusqu'à ce que de nombreuses générations se soient écoulées et que de nombreux événements de recombinaison soient intervenus, ces allèles sont trouvés ensemble plus souvent que ne le voudrait le hasard.

De telles études sont ambitieuses, car elles nécessitent de déterminer les allèles présents en chaque position chez de nombreux individus souffrant ou non de la maladie. Selon le nombre de snips que l'on estime nécessaire pour aboutir, le coût peut devenir prohibitif. Certains généticiens placent leur espoir dans l'organisation de ces allèles, du fait d'un déséquilibre de liaison, en "blocs haplotypiques" qui réduiraient la diversité réelle du génome et pourraient être typés à l'aide d'un petit nombre de snips. Quoi qu'il en soit, des progrès technologiques seront nécessaires avant que l'on puisse entreprendre de telles études d'association à l'échelle du génome entier. Les succès déjà obtenus, par exemple, dans le cas du diabète ou de la maladie de Crohn, indiquent en tous cas qu'une part des facteurs génétiques de susceptibilité aux maladies est à notre portée.



Dans les études d'association directe (à gauche), on teste l'association avec la maladie d'un allèle (*) situé dans la région codante d'un gène candidat, dans l'espoir qu'il s'agisse d'un variant fonctionnel. Dans les études d'association indirecte (à droite), on détermine, pour un ensemble de SNP régulièrement répartis, les allèles présents chez des personnes malades et bien portantes. L'association de certains allèles (en bleu) avec la maladie indique la région où pourrait se trouver le variant fonctionnel, avec lequel ces allèles seraient en déséquilibre de liaison.