

► Le cancer colorectal atteint plus de 500 000 personnes chaque année dans le monde. Le bénéfice majeur en termes de dépistage est la détection et le retrait des polypes adénomateux, ce qui sous-tend l'importance de trouver des biomarqueurs utiles pour le diagnostic précoce. La présence d'une histoire familiale de cancer colorectal est retrouvée dans environ 30 % des cas nouvellement diagnostiqués, mais moins de 5 % d'entre eux sont associés à un syndrome identifié de prédisposition héréditaire. Les études d'association ont permis d'identifier de multiples locus auxquels des allèles fréquents peuvent influencer le risque de développer un cancer colorectal. Le risque est généralement faible. Les effets combinés de plusieurs facteurs peuvent cependant être suffisamment importants pour être utilisés dans l'évaluation globale du risque, définissant des groupes de personnes à risque élevé, pour qui une adaptation du dépistage et de la prévention pourrait être proposée, d'autant plus que le nombre de locus ainsi identifiés croît rapidement. ◀

Identification d'allèles de susceptibilité au cancer colorectal à faible pénétrance

Richard Houlston

Adaptation libre de Sylviane Olschwang



R. Houlston : ICR, Royaume-Uni, Section of Cancer, Genetics Institute of Cancer Research, 15 Cotswold Road Belmont, Sutton SM2 5NG, Royaume-Uni.
r.houlston@icr.ac.uk
 S. Olschwang : UMR 891, Département d'Oncogénétique, Centre de Recherche en Cancérologie de Marseille, Institut Paoli-Calmettes, 27, boulevard Leï Roure, 13009 Marseille, France.
sylviane.olschwang@inserm.fr

qui l'on pourrait prévenir la maladie par le biais d'un dépistage et d'une chimio-prévention. Sur le plan mécanistique, cela nous éclaire également sur la biologie de la tumeur, nous aide à formuler des hypothèses étiologiques et peut permettre le développement de nouveaux traitements.

Les prédispositions génétiques majeures

Le cancer colorectal est un problème de santé publique majeur. Dans les pays occidentaux, ce type de cancer est devenu la première cause de mortalité liée au cancer parmi la population de non-fumeurs. C'est donc un enjeu très important. Le cancer colorectal se distingue des autres tumeurs par le fait que le pronostic dépend largement de son stade de détection mais le taux de guérison plafonne à 60 % dans la plupart des séries de patients.

Le mode de vie intervient dans le développement de la maladie, même si son rôle semble limité. Les facteurs connus à ce jour sont les comportements alimentaires et probablement la consommation d'anti-inflammatoires non stéroïdiens. Viennent ensuite les facteurs familiaux, d'ordre génétique ou non. L'intérêt de la recherche de gènes de susceptibilité est double. Tout d'abord, elle permet d'identifier des personnes à risque chez

En ce qui concerne le risque familial, nous savons, depuis les années 1950 grâce aux travaux de Wolf, qu'environ 10 % des patients atteints de cancer colorectal ont un parent de premier degré atteint. Les études épidémiologiques ont toutes montré que le risque est multiplié par deux à trois chez les parents au premier degré. Le risque est multiplié par quatre pour les sujets dont un parent de premier degré a développé le cancer avant 45 ans mais également pour ceux qui ont plus d'un parent au premier degré atteint par la maladie. Ces observations témoignent d'une prédisposition génétique, qui pourrait être le fait de plusieurs allèles avec un effet limité, conjugué à des facteurs environnementaux et non-génétiques ayant des effets plus importants. Cette hypothèse a été pour la première



fois formulée en 1974 par Anderson. Elle méconnaissait l'existence de gènes de prédisposition majeure à forte pénétrance, identifiés depuis : le cancer colorectal héréditaire sans polypose ou syndrome de Lynch résulte ainsi des mutations de *MLH1*, de *MSH2*, mais aussi de *MSH6*, et confère un risque cumulé de cancer de 70 %. La polypose adénomateuse familiale est associée à un risque supérieur à 90 %. Le risque est également très élevé (plus de 50 %) dans le cas de la polypose avec mutations du gène *MYH*, du syndrome de Peutz-Jeghers et de la polypose juvénile.

Il existe deux catégories de cancer colorectal, avec ou sans instabilité de type MSI. Le profil MSI est caractéristique du syndrome de Lynch. Toutes les autres prédispositions majeures se traduisent par des cancers MSS. La contribution de ces syndromes héréditaires aux deux groupes de cancers a été établie grâce aux registres finlandais. Sur environ 1 000 cancers colorectaux, 7 compliquaient une polypose adénomateuse, 2 liées à *APC* et 5 à *MYH* ; il y avait également 1 polypose juvénile et 29 syndromes de Lynch, soit un total de 3 % des cas liés à une prédisposition génétique majeure. Une proportion importante de cancers survenant dans un contexte d'agrégation familiale, essentiellement des cancers de type MSS, reste donc sans explication.

Principe des études d'association

Le génome humain présente de nombreuses variations sans conséquence phénotypique, les polymorphismes : petites insertions, délétions, variations de l'ADN mini- et micro-satellite. Il existe également des substitutions nucléotidiques, appelées SNP, qui présentent un intérêt particulier du fait d'une stabilité, de leur très grand nombre et de leur répartition homogène dans le génome. À l'inverse des études de liaison génétique, où il s'agit d'associer un phénotype avec une région chromosomique dans des familles atteintes, par une méthode de *Lod Score*, les études d'association consistent à prendre des échantillons de malades et des échantillons témoins et de comparer les fréquences alléliques dans les deux groupes. Une différence de fréquence peut traduire un lien entre un allèle et un phénotype. Les études d'association peuvent être centrées sur des SNP fonctionnels ou dans des gènes candidats. Au contraire, elles ne peuvent se baser sur aucun *a priori*. L'hypothèse sera alors qu'une variation allélique puisse présenter un

déséquilibre de liaison avec un phénotype, transmis au cours de l'évolution de l'espèce. Plusieurs SNP, par leur proximité physique, peuvent ainsi présenter un déséquilibre de liaison et constituer des haplotypes, dont la carte génomique a été établie grâce à un programme d'analyse international, le projet Hap Map. L'intérêt de cette carte pour les recherches d'association est de limiter le nombre de SNP analysés, ceux d'un même haplotype étant connus pour donner des informations identiques ou très proches. Actuellement, une analyse complète du génome sans *a priori*, à la recherche d'une association, peut être réalisée par génotypage de 300 000 à 500 000 SNP indépendants, dont la fréquence de l'allèle mineur est d'au moins 0,05. En effet, le génotypage de SNP moins polymorphes ne permettrait pas d'identifier des associations avec des allèles plus rares. Cette approche analytique s'appelle GWAS (*genome-wide association study*). Plusieurs sociétés de biotechnologie ont développé des outils expérimentaux efficaces, comme Illumina, Agilent, Affymetrix par exemple. Tous sont basés sur l'amplification *in vitro* par la technique PCR.

En résumé, les capacités à identifier des associations sont donc fonction du taux d'hétérozygotie des SNP et de leur degré de liaison génétique, mais également du nombre de SNP analysés, qui conditionne le nombre de tests statistiques. Enfin, un élément indépendant et essentiel sera la contribution des allèles repérés au phénotype étudié. Une recherche d'association nécessite donc une simulation préalable, qui tiendra compte de ces éléments et de la valeur minimale du risque à identifier. Cette simulation permettra de définir la taille des échantillons de cas, de témoins, pour l'étude initiale et pour les études de confirmation, focalisées sur les locus repérés lors de l'étude initiale.

Application au cancer colorectal

Une étude systématique, prévue en Grande-Bretagne, a consisté à identifier des allèles « à risque » dans deux populations : une première constituée de 900 cas familiaux et 1 000 témoins, la seconde de 2 000 patients de moins de 55 ans et 3 000 témoins. Chaque SNP présentant un risque de 1,1 était ensuite testé sur deux nouvelles populations, la première de 2 000 cas et 2 000 témoins, la seconde de 3 000 cas et 3 000 témoins.

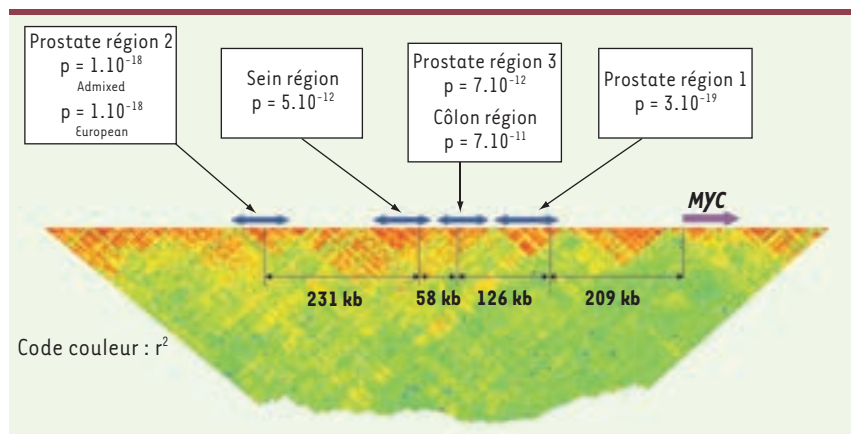


Figure 1. Identification du premier locus de susceptibilité mineure au cancer colorectal.



Pour identifier un allèle « à risque » de manière certaine, il était nécessaire que la fréquence de l'allèle mineur soit d'au moins 0,20 et que le risque soit d'au moins 1,2, pour une valeur statistique de 10^{-4} . Parmi les 550 000 SNP proposés par la technique Illumina, 2 561 n'étaient pas exploitables, soit environ 1 %. En général, les études de simulation font l'hypothèse que 10 % des données ne sont pas exploitables, et ce taux d'échec était donc très satisfaisant.

La simulation de l'analyse de la première population prédisait que 27 382 SNP montreraient une association au seuil 0,05, et la réalité en a identifié 27 673. Au seuil, 10^{-5} , 5 SNP étaient attendus contre 8 observés. L'association la plus forte, pour un SNP localisé en 8q24 a été confirmée sur les autres populations (Figure 1) [1, 2]. D'autres, moins fortes, étaient également robustes : 15q, 10p, 11q, 18q, avec un risque augmenté de 1,1 à 1,2 [3-7]. Depuis, une méta-analyse a trouvé 4 locus supplémentaires [8].

L'enjeu est ensuite d'identifier le gène en cause, que le SNP localisé en 18q illustre bien. Au voisinage de ce SNP se trouve le gène *SMAD7* [3]. L'hypothèse de son implication tient à cette notion de proximité. Ce gène, accompagné de 3 autres SNP, appartient au même bloc de déséquilibre de liaison. D'un point de vue fonctionnel (membre de la voie de signalisation du TGF β), il était également un excellent candidat, plusieurs autres membres de la famille SMAD ayant un rôle démontré dans la tumorigenèse colorectale. Son étude spécifique a permis d'identifier un variant, appelé NOVEL1, qui présentait une association plus forte que le premier SNP. La présence de ce variant est associée à une diminution importante de l'expression dans une lignée cellulaire de lymphoblastes et des études sont en cours dans des lignées dérivées de cancer colorectal.

L'étude du SNP situé en 8q présente un intérêt majeur, son association avec le cancer colorectal étant la plus forte, et une association forte existant aussi avec le cancer de la prostate. Cependant, le bloc de déséquilibre de liaison auquel il appartient a une taille très importante, d'environ 800 kb, et il n'y a pas de séquence candidate dessus.

Outre la caractérisation fonctionnelle de ces associations, qui permet de comprendre les mécanismes fins de la tumorigenèse, le génotypage de ces SNP comme facteurs de risque présente-t-il un intérêt dans une approche individuelle du dépistage ? Là encore, une simulation de l'impact à l'échelle d'une population est possible : il existe actuellement 7 allèles à risque iden-

tifiés, sur 7 locus distincts. La population présentant le plus fort risque est celle qui aurait les 7 allèles, mais elle ne constitue que 3 % de la population totale, le risque conféré par des associations de nombres d'allèles inférieurs ne justifiant pas pour sa part l'instauration d'un dépistage par coloscopie. Et encore s'adresse-t-on à des locus dont le degré d'hétérozygotie est supérieur à 10 %. Il est actuellement peu réaliste d'envisager l'identification d'allèles à risque plus rares. Reste enfin à analyser la conjugaison de la présence d'un ou plusieurs allèles à risque avec des facteurs de risque environnementaux. C'est un défi en soi, l'autre défi étant de dépasser ce stade d'identification d'associations en identifiant les variations biologiques responsables. ♦

SUMMARY

Identification of low-penetrance alleles associated with colon cancer risk

Colorectal cancer affects over 500,000 individuals yearly. Much of the benefit of colorectal cancer screening has been attributed to detection and removal of adenomatous polyps, highlighting the importance of colorectal polyps as targets for intervention and as biomarkers for colorectal cancer risk. Positive familial history (first or second degree relative) for colorectal carcinoma can be found in approximately 30 % of all newly diagnosed cases, but less than 5 % will be due to a defined genetic category of hereditary CRC. Genome-wide association studies have identified multiple loci at which common variants modestly influence the risk of developing colorectal cancer. The risks conferred by the susceptibility alleles are low. 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. Tomlinson I, Webb E, Carvajal-Carmona L, et al. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. *Nat Genet* 2007 ; 39 : 984-8.
2. Zanke BW, Greenwood CM, Rangrej J, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. *Nat Genet* 2007 ; 39 : 989-94.
3. Broderick P, Carvajal-Carmona L, Pittman AM, et al. A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk. *Nat Genet* 2007 ; 39 : 1315-7.
4. Poynter JN, Figueiredo JC, Conti DV, et al. Variants on 9p24 and 8q24 are associated with risk of colorectal cancer: results from the colon cancer family registry. *Cancer Res* 2007 ; 67 : 11128-32.
5. Jaeger E, Webb E, Howarth K. Common genetic variants at the CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk. *Nat Genet* 2008 ; 40 : 26-8.
6. Tenesa A, Farrington SM, Prendergast JG, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. *Nat Genet* 2008 ; 40 : 631-7.
7. Tomlinson IP, Webb E, Carvajal-Carmona L, et al. A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. *Nat Genet* 2008 ; 40 : 623-30.
8. Houlston RS, Webb E, Broderick P, et al. Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. *Nat Genet* 2008 ; 40 : 1426-35.

TIRÉS À PART

R. Houlston
et S. Olschwang