




Disponible en ligne sur
 ScienceDirect
www.sciencedirect.com

Elsevier Masson France
 EM|consulte
www.em-consulte.com



REVUES GÉNÉRALES ET ANALYSES PROSPECTIVES

Notions de génétique moléculaire pour comprendre l'hérédité

Molecular genetics and heredity: An introduction

J. Lamoril^{c,*}, N. Ameziane^a, J.-C. Deybach^c, P. Bouizegarène^c, M. Bogard^b

^a Laboratoire de biologie polyvalente, centre hospitalier de Sens, 89100 Sens, France

^b Laboratoire de biochimie et biologie moléculaire, centre hospitalier de Meaux, 77100 Meaux, France

^c Laboratoire de biochimie et génétique moléculaire, hôpital Louis-Mourier, 178, rue des Renouillers, 92700 Colombes, France

Reçu le 15 septembre 2008 ; accepté le 3 octobre 2008

Disponible sur Internet le 25 novembre 2008

KEYWORDS

Genetic;
Mendelian heredity;
Epigenetic;
Phenotype;
Polygenic

MOTS CLÉS

Génétique ;
Hérédité
mendélienne ;
Épigénétique ;
Phénotype ;
Polygénique

Summary Understanding on human inheritance has significantly improved these last 20 years. Even if monogenic diseases study brought important informations on the functioning of genes and deciphering of human genome on its organization, knowledge has evolved and has changed our view of heredity. Mendel laws applied to monogenic disorders are not any more sufficient to explain transmission of our characters. Genome regulations and its interactions included with environment demonstrate the complexity of transmission of our traits and the resulting adjustments acquired and hereditary alike for each individual. Moreover, human genetic is inadequately familiar to physicians and biologists and is insufficiently taught in medical universities. In this paper, we remind some important and current notions on human molecular genetic in order to help the non specialist physician or biologist to answer genetic questions to a consultant.

© 2008 Elsevier Masson SAS. Tous droits réservés.

Résumé La compréhension de l'hérédité a considérablement progressé ces 20 dernières années. L'étude des maladies monogéniques a fortement contribué à nous éclairer sur le fonctionnement des gènes et le récent décryptage du génome humain sur son organisation. Les connaissances ont évolué et modifié notre vision de l'hérédité : à l'heure actuelle, les lois de Mendel appliquées aux maladies monogéniques ne suffisent plus à expliquer la transmission de nos caractères. Les régulations du génome et ses interactions, y compris l'impact de l'environnement, démontrent la complexité de la transmission de nos traits et les modulations précises à la fois acquises et héréditaires qui en résultent pour chaque individu. Par ailleurs, la génétique humaine est insuffisamment connue des médecins et biologistes et n'a pas une

* Auteur correspondant.

Adresse e-mail : jerome.lamoril@lmr.ap-hop-paris.fr (J. Lamoril).

place suffisante dans l'enseignement médical. Cet article a pour objectif non pas de décrire les pathologies héréditaires mais de rappeler quelques notions essentielles des mécanismes de transmission. La génétique ayant une place de plus en plus importante en médecine, il nous semble important de connaître ces principes, les consultants étant amenés à poser des questions au médecin non spécialiste et le médecin ou le biologiste confrontés de plus en plus à ces questions se trouvent fréquemment dans la difficulté pour y répondre.

© 2008 Elsevier Masson SAS. Tous droits réservés.

Introduction

Selon la définition générale du dictionnaire, la génétique humaine est une branche de la médecine ayant trait à l'hérédité et aux gènes, l'hérédité se définissant comme l'ensemble des caractères physiques ou non transmis par le père et/ou la mère et les gènes par les éléments issus des chromosomes par lesquels sont transmis les caractères d'un sujet. En pratique, la génétique humaine est bien plus complexe que ne le laisse supposer cette définition comme nous allons l'exposer dans cet article. Par ailleurs, en médecine humaine, la génétique est divisée classiquement en plusieurs branches, chacune résultant de l'utilisation d'outils de travail différents. Schématiquement, on distingue :

- la génétique médicale, branche clinique correspondant aux consultations de génétique assurées par des praticiens communément appelés généticiens et oncogénéticiens, ces derniers s'occupant de la génétique associée aux cancers. Les autres spécialistes en génétique peuvent aussi donner des consultations de génétique. En génétique clinique, existent aussi des conseillers en génétique (ces derniers ne sont pas obligatoirement médecins) ;
- la cytogénétique : son domaine est l'étude des chromosomes et de leurs anomalies ;
- la génétique moléculaire (associée à l'épigénétique) axée sur l'étude des gènes et du génome humain à l'échelle des acides nucléiques.

Cet article se donne pour objectifs d'éclairer le lecteur sur la génétique moléculaire humaine et l'aider à comprendre les dernières évolutions importantes dans ce domaine. L'identification de gènes associés aux maladies a révolutionné la génétique et plus généralement, la médecine. En effet, la génétique prend une place de plus en plus importante dans le domaine du diagnostic, de la prévention, du pronostic, de la thérapeutique, de l'épidémiologie et des connaissances biologiques. La génétique moléculaire et les outils qui l'accompagnent (la biologie moléculaire) progressent quotidiennement [1]. Le modèle classique, celui de l'hérédité mendélienne est associé à la découverte de nombreux gènes et la théorie un gène-une maladie a longtemps été l'unique conception de la génétique. Les nombreuses découvertes sur le mécanisme des maladies héréditaires, sur le fonctionnement des gènes ont fait évoluer cette théorie. Dans les années 1990, on considérait deux modèles, les maladies monofactorielles obéissant dans une certaine mesure aux lois de Mendel (un gène-une maladie) et les maladies multifactorielles (une maladie est la résultante

de plusieurs facteurs environnementaux et de l'action de plusieurs gènes). Cette classification a rapidement paru trop simpliste. Les conceptions des mécanismes génétiques des maladies et de leurs modes de transmission ont évolué. Pour mieux comprendre la génétique moléculaire actuelle, nous rappellerons quelques données de bases (la cytogénétique et la génétique clinique sont exclues de cet exposé). Dans le domaine de la génétique humaine, le concept de maladie monogénique a en effet été la première notion introduite et la mieux connue des médecins. Nous commencerons donc par la rappeler pour développer ensuite les évolutions récentes de la connaissance en génétique et proposer de manière schématique, une approche pratique.

Les maladies monogéniques

Les maladies dites monogéniques (une maladie-un gène) ont une transmission et des mécanismes moléculaires plus complexes que ne les décrit l'hérédité mendélienne. Cependant, afin de mieux comprendre cette complexité, il est indispensable de rappeler certains concepts fondamentaux toujours utilisés en génétique médicale.

Qu'est-ce qu'un gène ?

Lorsqu'on parle de génétique, on parle de gène. Il est donc important de rappeler brièvement sa signification. La notion de gène a en effet évolué ces dernières années suite aux nombreuses découvertes sur ce sujet. Jusque dans les années 1980, un gène était considéré comme une séquence d'ADN dont la lecture aboutissait à la formation d'une protéine. Un gène était reconnu sur la séquence d'ADN à l'identification d'un cadre ouvert de lecture (*open reading frame* [ORF]), ce dernier correspondant à la succession des codons aboutissant après les étapes de transcription (ADN génomique en ARN messenger), de traduction (ARN messenger en polypeptide) et de modifications post-traductionnelles en une protéine fonctionnelle. La connaissance actuelle du génome humain a modifié cette acceptation. Ainsi, en 2002, un gène était-il considéré comme un segment d'ADN contribuant au phénotype d'un individu. Puis, en 2006, le gène est devenu une région identifiée du génome correspondant à une unité transmissible (héritable) associée à des zones de régulation, de transcription et/ou associée à d'autres propriétés fonctionnelles. Pour qu'il n'y ait pas d'ambiguïté sur le terme de gène, il a été précisé que tous les transcrits alternatifs d'une séquence d'ADN (liés par exemple à des sites d'épissage alternatifs) appartiennent au même gène même si le produit de ces transcrits aboutit à la formation de protéines différentes. En 2007, une révision importante de cette définition a été reprise : un gène est constitué de

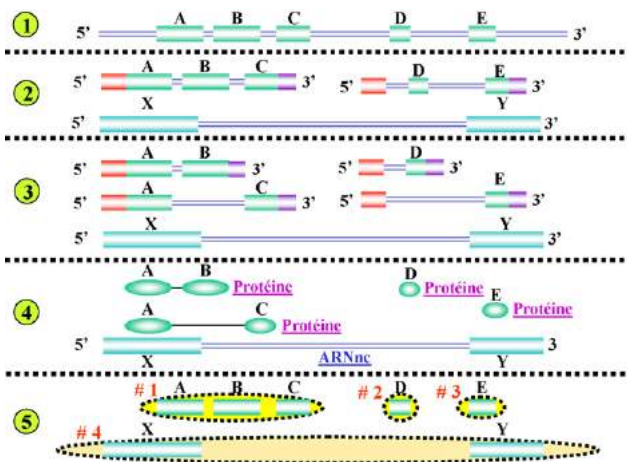


Figure 1 Schéma général définissant le gène.

1. Région génomique située sur un chromosome. Elle est constituée des segments A, B, C, D et E appelés exons.
2. Trois transcrits primaires sont produits à partir de cette région génomique (5'-ABC-3', 5'-DE-3' et 5'-XY-3').
3. Après épissage alternatif des transcrits primaires, on obtient cinq produits différents. Quatre aboutiront à la formation d'une protéine et le cinquième à la formation d'un ARN non codant (ARNnc).
4. Les protéines sont formées à partir de trois groupes d'exons (A, B et C, D, E). Le premier groupe de segments (A, B et C) ont en commun un même segment A. Deux transcrits primaires (D et E) partagent la même partie 5' non codante mais la traduction de leur transcrits ne se chevauche pas. L'ARN non codant (qui n'est donc pas traduit en protéine) comprend des séquences communes (X et Y) avec les segments codants A et E.
5. Il existe donc quatre gènes dans cette région :
 - le gène 1 : il est constitué des segments A, B et C ;
 - le gène 2 : il est constitué du segment D ;
 - le gène 3 : il est constitué du segment E ;
 - le gène 4 : il est constitué du segment X et Y.

séquences génomiques codant pour un ensemble cohérent de produits fonctionnels, ces derniers pouvant se chevaucher (Fig. 1) [24]. Cette définition importante appelle trois remarques fondamentales :

- un gène est une séquence génomique codant directement pour une molécule fonctionnelle (ARN ou protéine) ;
- lorsque plusieurs produits fonctionnels partagent des régions du génome qui se chevauchent, le gène correspond aux séquences chevauchantes dont sont issus ces produits fonctionnels ;
- l'union de ces séquences doit être cohérente. Séparément, elles aboutissent à un produit final déterminé (protéine ou ARN). Cependant, il n'est pas indispensable que tous ces produits partagent nécessairement les mêmes séquences.

Cette définition aux contours flous englobe la notion de gène classiquement connue (un gène code pour une ou quelques protéines) mais aussi la notion selon laquelle un gène peut coder pour une molécule fonctionnelle qui n'est pas une protéine, un ARN non codant par exemple. Les gènes codant pour une (ou quelques) protéines représentent une

partie du génome. Selon l'organisme, elle constitue une part plus ou moins importante de celui-ci. À titre d'exemple, les gènes codant pour une protéine représentent environ 90% du génome des procaryotes, 68% de celui des levures, 25% de celui des nématodes, 17% de celui des insectes, 2% de celui du poulet et 1% de celui des mammifères [59]. Quelle est alors la fonction d'un gène? De manière générale, un génotype aboutit à un phénotype. Au niveau moléculaire, cela signifie qu'une séquence d'ADN détermine la séquence d'une molécule fonctionnelle. L'ADN code donc pour une protéine ou un ARN. Mais, d'une manière plus générale et plus complexe, les gènes sont constitués de modules pouvant se combiner entre eux et générer des produits de multiples façons. Il est important de souligner la propriété de fonctionnalité des gènes, propriété qui permet de relier le génotype au phénotype. Ainsi, il n'y a pas discontinuité avec les définitions antérieures d'un gène mais évolution dans la conception du gène. On peut voir ainsi avec l'évolution des définitions que la notion de gène est de plus en plus difficile à définir. Les apports de la génétique moléculaire nous montrent une complexité du génome plus importante que la vision relativement simple des généticiens jusqu'au début des années 2000.

Modes de transmission obéissant à l'hérédité mendélienne

Nous allons rappeler les principaux modes de transmission des maladies dites monofactorielles ou monogéniques (un gène-une maladie). Pour mémoire, nous rappelons que l'être humain normal possède 23 paires de chromosomes (dans chaque paire, un chromosome est issu du père et l'autre de la mère) dont 22 paires d'autosomes (les chromosomes non sexuels) et une paire de chromosomes sexuels (XX pour la femme et XY pour l'homme).

Transmission autosomique dominante

Une maladie est dite de transmission autosomique dominante lorsqu'un sujet présente sur un des deux autosomes d'une même paire de chromosomes un gène anormal (ou muté). L'autre gène situé sur l'autre chromosome de la même paire est normal. On parle d'allèle muté pour le gène anormal. L'anomalie étant portée par un chromosome non sexuel, elle est donc indépendante du sexe. Le sujet est dit hétérozygote pour l'allèle muté (un gène sur deux est anormal). Sur le plan probabilité, un malade (porteur du gène anormal à l'état hétérozygote) a une chance sur deux (1/2) d'avoir un enfant atteint avec un conjoint sain. La Fig. 2 représente ce type de transmission à travers trois générations. En pratique, ce mode de transmission simplifié présente quelques « variantes » :

- un sujet peut être homozygote (les deux chromosomes atteints). Dans ce cas, la maladie est beaucoup plus grave, voire létale ;
- un sujet peut être porteur du gène anormal et ne pas être malade. On dit qu'il est porteur sain de l'allèle muté. Il peut cependant transmettre le gène anormal à sa descendance. On parle dans ce cas de pénétrance incomplète. La pénétrance représente le pourcentage de sujets malades et porteurs de l'allèle muté sur l'ensemble

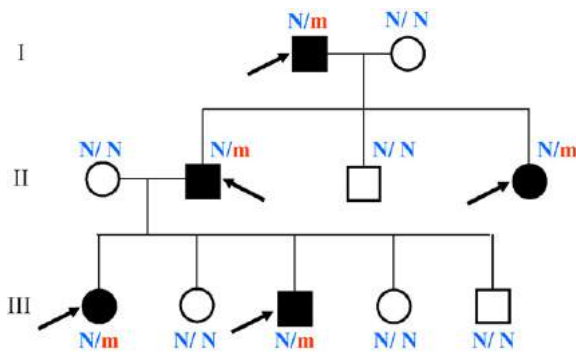


Figure 2 Transmission autosomique dominante. Trois générations sont représentés (I, II et III). Chaque flèche indique un sujet malade, homme (■) ou femme (●). Dans cet exemple, un sujet est malade à chaque génération, indépendamment du sexe. Le sujet est hétérozygote pour le gène anormal (allèle m) et a une chance sur deux de transmettre l'anomalie après union avec un conjoint sain. N : allèle normal ; M : allèle muté.

des sujets malades et porteurs sains. Ainsi, 10% de pénétrance pour une maladie autosomique dominante signifie que 10% des sujets porteurs de l'allèle muté sont malades et 90% sont des porteurs sains (donc, non malades). De nombreuses maladies monogéniques possèdent ce type de transmission telles que la fièvre familiale méditerranéenne (appelée aussi maladie périodique, Omim 134610), la porphyrie aiguë intermittente (Omim 176000), l'hypercholestérolémie familiale (ou hyperlipoprotidémie type IIA, Omim 143890).

Transmission autosomique récessive

Une maladie est dite de transmission autosomique récessive lorsqu'un sujet présente sur les deux autosomes d'une même paire de chromosome un gène anormal (Fig. 3). Le sujet est homozygote pour l'allèle muté si la nature de la mutation est identique sur les deux chromosomes. Si la mutation est différente pour le gène muté sur chaque chromosome homologue, on parle d'hétérozygote composite. Un sujet porteur d'un seul chromosome malade est normal. Les deux parents sont donc normaux puisqu'ils sont chacun porteur d'un seul gène muté (sur un des deux chromosomes homologues). Le risque pour un enfant d'être homozygote (et donc malade) lorsque ses deux parents sont hétérozygotes pour une mutation est de 0,25 (une chance sur quatre). Parmi les pathologies possédant ce type de transmission, on peut citer la mucoviscidose (Omim 219700), la sclérose latérale amyotrophique familiale de type I (Omim 105400).

Transmission récessive liée à l'X

Les hommes sont XY (porteur d'un chromosome X et d'un chromosome Y) et les femmes XX (porteuses de deux chromosomes X). On dit que les hommes sont hémizygotés pour le chromosome X parce qu'ils ne portent qu'un seul chromosome X. Un garçon porteur d'une mutation sur le chromosome X est malade. Il transmet à toutes ses filles le gène muté et à aucun de ses garçons. Une fille hétérozygote pour le gène muté a une probabilité de 0,25 (1 chance sur 4) d'avoir un garçon atteint, 0,25 d'avoir une fille hétérozygote (dite aussi conductrice), 0,25

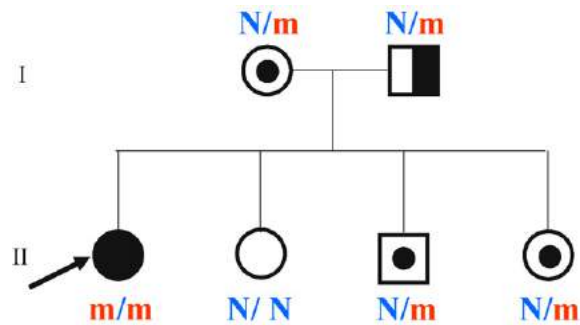


Figure 3 Transmission autosomique récessive. Les parents sont hétérozygotes pour une mutation sur un gène situé dans un autosome. Ils ont chacun transmis le gène muté à leur fille (flèche). Si la mutation issue du père et de la mère est identique, la fille est homozygote pour la mutation. Si la mutation sur le gène anormal est différente pour chaque parent, la fille est hétérozygote composite. Une fille et un garçon (indiqué par le point noir) sont porteurs de la mutation à l'état hétérozygote. Ils ne sont pas malades mais peuvent transmettre le gène muté à leur descendance. L'autre fille (cercle blanc) ne porte aucune mutation. Il a hérité des deux allèles normaux de chacun des parents. Bien évidemment, il n'est pas malade et ne transmettra pas d'allèle anormal pour ce gène. N : allèle normal ; M : allèle mute.

d'avoir un garçon sain et 0,25 d'avoir une fille saine en cas d'union avec un conjoint sain (Fig. 4). Les hémophilias A (Omim 306700) et B (Omim 306900) en sont deux exemples.

Transmission liée à l'Y

Dans de rares cas, on observe une pathologie par transmission d'un gène muté sur le chromosome Y. Le chromosome Y présente en fait deux régions de séquence identique avec le chromosome X. Ces deux petites régions du chromosome Y s'appellent régions pseudo-autosomiques. En cas de mutation dans une de ces deux régions, une pathologie peut être

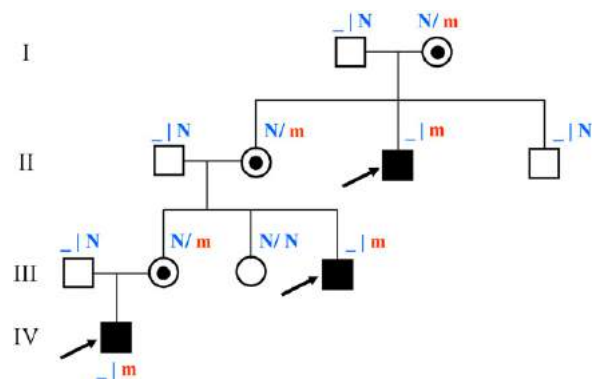


Figure 4 Transmission récessive liée à l'X. Tous les garçons ayant hérité la mutation sont atteints (carrés noirs et flèche). Les filles sont conductrices (cercle avec point noir). Leur probabilité d'avoir un garçon atteint avec un conjoint normal est une chance sur 4. N : allèle normal ; M : allèle muté. L'homme possède un seul chromosome X, l'autre chromosome étant le chromosome Y.

observée. C'est le cas par exemple dans certaines infertilités de l'homme telle que l'azoospermie non obstructive liée à l'Y (Omim 415000) [22].

Les néomutations

Dans certains cas, il apparaît que les parents sont normaux (aucun gène muté) et que leur enfant soit atteint d'une maladie héréditaire. Après s'être assuré de la paternité (le père supposé est le vrai père : un test de paternité permet de le confirmer en cas de doute), une telle mutation chez l'enfant est possible quand une mutation s'est produite au cours de la gamétogenèse d'un des deux parents. On parle de mutation de novo ou néomutation. Ces mutations peuvent survenir dans la lignée germinale, lors de la mitose au cours de la spermatogenèse ou de l'ovogenèse ou pendant la méiose. Ces néomutations sont fréquemment observées dans certaines pathologies. Par exemple, dans la myopathie de Duchenne (Omim 310200), 30 % des mutations sont des néomutations.

Le mosaïcisme

Le mosaïcisme se définit comme l'existence de deux populations cellulaires au minimum, génétiquement distinctes et issues d'un seul œuf. Les mutations provoquant ce phénomène peuvent apparaître au cours de l'embryogenèse (on parle de mosaïque somatique). Le nombre de cellules atteintes est donc variable. Le phénotype résultant de cette double population est déterminé par la période où est apparue la mutation et le type de mutation en cause. On observe alors une hétérogénéité clinique (expressivité variable) pour la maladie causale. L'inactivation de l'X (voir ci-dessous) est un exemple physiologique de mosaïcisme cellulaire. Elles peuvent être aussi plus rarement germinales (un clone de cellules germinales est porteur d'une mutation) et aboutir alors à une néomutation.

La consanguinité

En cas de consanguinité (lien de parenté entre les deux parents ; par exemple, cousins), la probabilité d'avoir un enfant avec un gène muté dont la transmission est autosomique récessive augmente considérablement. On estime que dans le monde environ 20–50 % des mariages sont consanguins, ce qui représente environ un milliard d'individus.

L'hérédité mitochondriale

Le génome de la cellule est constitué en fait de deux types de génomes. Le principal génome est celui contenu dans le noyau de la cellule et que nous appelons communément le génome humain. L'autre génome de la cellule est indépendant de celui du noyau et est contenu dans une organelle cellulaire appelée mitochondrie. Celui-ci, de petite taille (environ 16 kb) présente un mode de transmission spécifique : il est uniquement transmis par la mère.

Le catalogue des maladies de transmission mendélienne

L'ensemble des maladies héréditaires monogéniques peut être consulté sur un site Internet américain, le site *Online Mendelian Inheritance in Man (OMIM)*, encyclopédie exhaustive, mise à jour en permanence et d'accès gra-

tuit (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM>). Dans notre article, chaque pathologie ou gène s'y rapportant sont associés à un numéro, le numéro Omim. Ce numéro permet d'avoir accès directement aux renseignements cliniques, biologiques, cytogénétiques et moléculaires de la pathologie ou du gène impliqué.

Qu'est-ce qu'un gène muté ?

De nombreuses maladies héréditaires et certaines maladies acquises (par exemple, les cancers) présentent des anomalies sur le(s) gène(s) impliqué(s) ou sur les zones régulant l'expression de ces gènes. Ces anomalies sont appelées mutations. Une mutation de manière générale est une altération de l'ADN modifiant le génotype du sujet. La conséquence d'une mutation est l'altération du fonctionnement normal du gène ou de l'expression de celui-ci. Sans rentrer dans les détails des mutations, on distingue schématiquement plusieurs types de mutations dans les maladies monogéniques [1] :

- les mutations ponctuelles : une base (A, C, G ou T) est remplacée par une autre base ;
- une modification (de petite taille) du gène : délétion (perte de quelques bases) ou insertion (ajout de quelques bases) ;
- un réarrangement du gène : il s'agit alors de grandes délétions (un exon, une partie ou la totalité du gène) ou d'une insertion (tout ou partie du gène), d'une translocation (une partie du gène est retrouvée sur un autre chromosome à proximité d'un autre gène), d'une inversion (changement d'orientation d'un gène dans le génome). On parle aussi de modification structurale du gène (voir infra) ;
- une modification du nombre de copies du gène : la majorité des gènes sont présents en un seul exemplaire dans le génome. Il peut arriver qu'un gène soit anormalement présent en plusieurs exemplaires, le plus souvent en double (duplication) ou en triple (triplication) [44]. Il s'agit là aussi d'une modification structurale du gène.

Conséquences des mutations. Un grand nombre de mutations (ponctuelles notamment) modifient la séquence d'acides aminés qui constituent le squelette de la protéine. Les conséquences fonctionnelles peuvent alors être de deux sortes :

- une mutation avec gain de fonction :
 - la protéine mutée présente une activité augmentée ou une nouvelle activité qui peut interagir ou non avec la protéine normale (en cas de mutation hétérozygote, par exemple, un gène étant normal et permettant la synthèse de la protéine normale et l'autre gène muté aboutissant à la synthèse de la protéine anormale, les deux protéines normale et anormale pouvant interagir l'une sur l'autre),
- une mutation avec perte de fonction :
 - la protéine présente une perte ou une abolition de sa fonction (par exemple, abolition de l'activité d'une enzyme).

Des mutations peuvent aussi se trouver dans des régions non codantes d'un gène, par exemple, sur des zones de régulation de l'expression d'un gène modifiant alors son expression ou sur des zones intervenant dans la stabilité de l'ARNm (par exemple, dans la région 3' non codante d'un gène) [1].

Les mutations acquises

Lorsqu'on parle de maladies génétiques, il s'agit traditionnellement des maladies héréditaires c'est-à-dire des maladies qui se transmettent de génération en génération et liées à une modification du génome (qualitative ou quantitative) ou du nombre de chromosomes (par exemple, aneuploïdie, translocation déséquilibrée...). Cependant, outre une anomalie chromosomique acquise, une maladie peut être génétique et acquise soit par néomutation soit par modification secondaire de novo du génome que ce soit une mutation ponctuelle, une insertion, une délétion ou une variation du nombre de copies (par exemple, par influence de l'environnement comme le tabac ou autre toxique). Par opposition à la génétique classique, on parle de génétique somatique (par exemple, dans l'étude des cancers non héréditaires).

Phénotype et mutation

Certaines maladies présentent un phénotype similaire. Néanmoins, les mutations responsables se retrouvent dans de nombreux gènes différents. On parle alors d'hétérogénéité génétique [2]. C'est le cas par exemple de la rétinite pigmentaire (Omim 268000) pour laquelle existent pour un phénotype similaire plusieurs gènes et différentes formes de transmission héréditaire (autosomique dominante, récessive, liée au sexe et mitochondriale). Au moins 42 gènes différents sont répertoriés [17].

Inversement, dans un même gène, différentes mutations induisent des phénotypes différents. C'est le cas par exemple sur le gène codant pour la lamine A/C (LMNA, Omim 150330). Selon la mutation, différentes pathologies peuvent être observées telles que la dystrophie musculaire d'Emery-Dreifuss (Omim 181350), une forme de cardiomyopathie dilatée (Omim 115200), une lipodystrophie partielle (Omim 151660), une dystrophie musculaire des ceintures de type IB (Omim 159001), la maladie de Charcot-Marie-Tooth de type 2B1 (Omim 605588), la progéria de Hutchinson-Gilford (Omim 176670), la dermatopathie restrictive (Omim 275210).

La relation génotype–phénotype

Outre ce qui a déjà été dit dans le paragraphe précédent, pour un même gène, la recherche d'une relation génotype–phénotype est importante à connaître, les conséquences pouvant être importante tant sur le plan pronostic que thérapeutique par exemple. En fait, une relation génotype–phénotype consiste à chercher si une mutation ou un groupe de mutations situées sur un gène sont associées à un phénotype déterminé. Par exemple, dans les tumeurs endocrines multiples de type 1 (*multiple endocrine neoplasia* type 1 (MEN-1), Omim 131100) de transmission autosomique dominante, caractérisée par au moins deux tumeurs endocrines (cf. adénome de la parathyroïde, tumeur neuroendocrine pancréaticoduodé-

nale, adénome pituitaire), de nombreuses mutations (> 400) ont été identifiées sur le gène *MEN1*. Aucune corrélation génotype–phénotype n'a été observée. Par conséquent, la conduite à tenir dans la prise en charge du patient sera identique quel que soit la mutation présente sur le gène. Le gène *APC* (Omim 611731) est impliqué dans la polypose familiale adénomateuse (*familial adenomatous polyposis* (FAP), Omim 175100) de transmission autosomique dominante. Elle se traduit plus particulièrement par la présence de nombreux polypes dans le colôn et le rectum (de 10 à plus de 1000) et prédispose au cancer du colôn. Dans ce gène, selon la position de la mutation, les conséquences phénotypiques sont différentes. Par exemple, outre les manifestations colorectales, les mutations sur les codons 542-1309 sont associées à des lésions pigmentaires de la rétine alors que les mutations sur les codons 1465, 1546, et 2621 du gène sont associées à de nombreuses manifestations extra-intestinales. Il existe donc bien dans ce cas, une corrélation génotype–phénotype dont les conséquences pour la prise en charge du patient doivent être prises en compte [25]. Comme nous l'avons précisé dans la définition d'un gène, la complexité du génome explique la difficulté d'interprétation de la relation génotype–phénotype ainsi que les notions d'expressivité de la maladie et de pénétrance de celle-ci. En effet, son étude et la comparaison de celui-ci aux autres espèces démontrent que celui-ci est bien plus compliqué que sa première description complète ne laissait penser. Bien que le génome humain soit environ 30 fois plus grand que celui des nématodes, il comprend quasiment le même nombre de gènes (22 726 versus 20 060) [28]. La quantité de gènes ne suffit donc pas à expliquer la complexité d'un organisme. Des études ont montré par ailleurs que les génomes produisaient bien plus de transcrits (produits du génome) qu'il n'existait de gènes. Par conséquent, un grand nombre de transcrits ne produisent pas de protéines (pour rappel, ces dernières sont issues de gènes codants). Dans le génome humain notamment, ces transcrits constituent la majorité des produits de transcription. En fait, de nombreuses études ont démontré la complexité de la transcription (c'est-à-dire la transformation de l'ADN génomique en ARN). Parmi les transcrits non codants identifiés (sur lesquels de nombreuses études sont en cours), on peut citer les ARN ribosomaux, de transfert, les petits ARN nucléaires ou *small nuclear* RNA (snRNAs), les petits ARN nucléolaires ou *small nucleolar* RNA (snoRNA), les composants ARN de ribonucléases (RNases) telles que les RNases P et autres RNases, les microARN (miARN) et les petits ARN interférents, les *small interfering* RNA (siRNA), les ARN non codants similaires aux ARNm, les *messenger-like non coding* RNA (mlRNA). Le monde des ARN (en addition au « monde » des ADN) joue un rôle extrêmement important dans la chaîne d'événements génétiques tant à l'échelle moléculaire que cellulaire. Par ailleurs, l'analyse de nombreux transcrits a démontré qu'au cours de la transcription d'un segment d'ADN dans le sens classique (5' vers 3'), il y avait souvent simultanément une transcription sur le même segment mais en antisens (donc de 3' vers 5'), ce dernier pouvant être chevauchant ou à distance du transcrit sens [28,38] sans pour autant qu'il y ait une simultanéité d'expression des transcrits sens et antisens même s'il a été démontré que le transcrit antisens pouvait jouer un rôle dans la régulation du transcrit sens pour un certain

nombre d'entre eux. Par ailleurs, outre l'existence de transcrits antisens, de nombreux transcrits codants sont des isoformes de la protéine de référence (la principale protéine codée par un gène codant). Enfin, des pseudogènes dont la séquence est proche de la séquences du gène du même nom mais qui ne produisent pas le plus souvent de protéine fonctionnelle (et souvent aucun transcrit), peuvent cependant dans certains cas être transcrits, quelques uns d'entre eux étant fonctionnels [68]. Les notions décrites dans ce paragraphe bien que simplifiées montrent la complexité d'organisation du génome. Par conséquent, l'interprétation des relations génotype–phénotype ne se résume pas uniquement à l'étude des mutations sur les gènes : ce concept évolue. Par ailleurs, d'autres événements influent sur cette relation à savoir l'environnement dans son ensemble. Donc, les facteurs génétiques, leurs modulations, leurs régulations complexes ainsi que les facteurs épigénétiques et environnementaux modulent la pénétrance et l'expressivité des maladies génétiques et expliquent en grande partie la variabilité de ces maladies tant d'un individu à l'autre qu'au sein d'une même famille pour une même mutation sur un même gène. Ces inter-relations complexes expliquent les nombreuses difficultés pour comprendre la relation génotype–phénotype de nombreuses pathologies.

La notion de polymorphisme

Le génome humain contient environ 25 000 gènes codant pour des protéines soit environ 5 % du génome. Le projet de séquençage humain, le séquençage individuel de génomes humains, eucaryotes et procaryotes et l'étude des maladies génétiques ont démontré la présence de nombreuses variations du génome d'un individu à l'autre. Ces variations ne sont pas appelées mutations mais polymorphismes car elles ne provoquent pas de maladie a priori. Un polymorphisme peut se localiser n'importe où sur le génome que ce soit dans un gène ou non avec une fréquence variable dans les populations (différences ethniques et géographiques). De manière générale, pour parler de polymorphisme, il est nécessaire que la fréquence de celui-ci soit supérieure ou égale à 1 %. Si la fréquence de cette variation est inférieure à 1 %, on parle de variant. Classiquement, on distingue plusieurs types de polymorphismes :

- le polymorphisme bi-allélique : une base (A, C, G ou T) peut être remplacée par une autre base sur un gène (un allèle se définit dans ce cas comme une des formes du gène pour la base considérée, sur un même locus chromosomique). Un polymorphisme peut être présent à l'état hétérozygote ou homozygote. Par exemple, on peut avoir un C ou un T sur la 150^e position du gène codant. Les anglosaxons définissent le polymorphisme bi-allélique par le terme *single nucleotide polymorphism* (SNP), terme souvent repris en français. Les SNPs sont distribués partout dans le génome humain et sont souvent utilisés dans les études d'association de maladies avec des gènes (notamment pour les maladies multifactorielles, voir infra). On estime que le génome de chaque individu contient environ dix millions de SNPs. L'association de SNPs dans une région du génome située sur un même chromosome s'appelle un haplotype. Certains SNPs proches physiquement sur le génome ont tendance à être transmis en bloc (ensemble).

On dit qu'ils sont en déséquilibre de liaison (qui peut être partiel ou total) ;

- la variation du nombre de séquences répétées : sur le génome, certaines séquences nucléotidiques sont répétées un certain nombre de fois en tandem. Par exemple, –CACACACA–. Les deux nucléotides C et A (dinucléotide CA) sont répétées quatre fois (CA₄). Ce nombre peut dans certains cas varier pour un même locus chromosomique. Par exemple, il peut être répété quatre fois sur le chromosome issu de la mère et huit fois sur celui issu du père. On aura alors un polymorphisme de séquences répétées à l'état hétérozygote [CA]₄/[CA]₈. On parle souvent de microsatellites (séquence répétée <8 pb) ou, pour les séquences répétées de plus grande taille, de minisatellites. Il existe de nombreuses variantes de minisatellites comme par exemple les *variable number tandem repeats* (VNTRs) ;
- la variation du nombre de copies d'un gène (variation structurale). Certains gènes ou séquences extragéniques sont présents en plusieurs copies. Ce nombre de copies peut varier d'un sujet à l'autre [58]. On parle de variation du nombre de copies (*copy number variation* [CNV]). Par exemple, le nombre de copies du gène AMY1A (Omim 104700, codant pour l'amylase salivaire) est variable d'une population à l'autre (1 à 10 copies) et aboutit à une concentration salivaire d'amylase variable et proportionnelle au nombre de copies du gène d'un sujet à l'autre et même d'une population à l'autre. À ce jour, on estime que le génome humaine comprend au moins 1500 CNV et représente environ 12 % du génome. Les CNV peuvent directement affecter l'expression d'un gène ou le dosage de ce gène (la quantité de protéines issues de l'expression du gène) ;
- outre la variation du nombre de copies, on peut observer d'autres variations appelées aussi variations structurales et modifiant la structure chromosomique (*structural variation* [SV]) telle que par exemple, une insertion, une délétion ou une duplication (Fig. 5). On parle de variation structurale lorsque le changement du nombre de copies d'un segment d'ADN ou de sa variation est supérieur ou égale à 1 Kb [14]. Il est important de noter par ailleurs que les éléments transposables (par exemple, LINE, SINE) sont exclus de cette définition (nous ne les décrivons pas : elles sortent du cadre de cet article). Ces variations structurales peuvent être selon les cas, des polymorphismes fonctionnels ou des mutations responsables de pathologies [37].

Conséquences fonctionnelles des polymorphismes

Le plus souvent, les polymorphismes sont neutres. Dans ce cas, un polymorphisme ne présente aucune portée fonctionnelle. Cependant, un polymorphisme peut avoir des conséquences importantes. En effet, il peut arriver qu'un polymorphisme module l'expression d'un gène, voire même accroît ou déclenche l'expression pathologique d'une mutation. On parle alors de polymorphisme fonctionnel. Par exemple, dans le cadre de la protoporphyrie érythroïdique liée à un déficit en ferroxalase, enzyme de la dernière étape de la synthèse de l'hème (Omim 177000), pathologie se traduisant par une photosensibilité invalidante, il a été décrit une hérédité particulière liée à

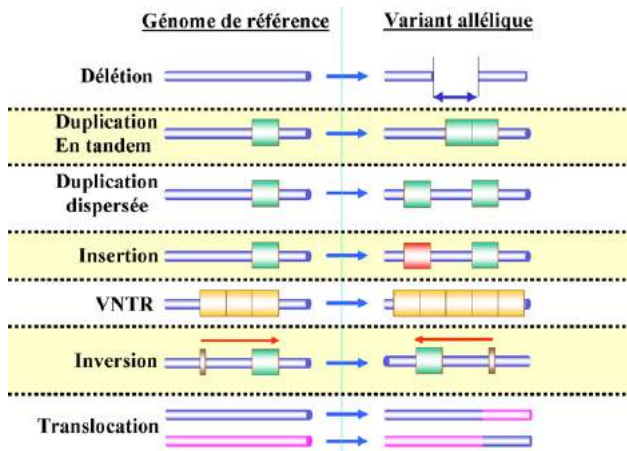


Figure 5 Principales variations structurales [37]. Les variations structurales se définissent comme des variants génétiques modifiant la structure chromosomique et dont la taille est supérieure à 1 kb. Ces variations peuvent induire des pathologies mais aussi être des polymorphismes fonctionnels. Une des variations structurelles actuellement la plus étudiée est la variation du nombre de copies d'un fragment d'ADN (gène ou partie d'un gène, par exemple) (*copy number variation* [CNV]) qui aboutit dans environ 80 % des cas à une augmentation de l'expression du gène (effet de dosage de gène) et dans 20 % des cas à l'inverse (pour plus d'informations, voir le site internet Decipher, decipher.sanger.ac.uk). *Variable number of tandem repeats* (VNTR), séquence répétée de grande taille.

l'existence ou non d'un polymorphisme en *trans* d'une mutation (c'est-à-dire sur le même gène mais situé sur l'autre chromosome par rapport à la mutation). Cette maladie est considérée habituellement comme de transmission autosomique dominante avec pénétrance incomplète (des formes récessives ont aussi été décrites). Les sujets mutés hétérozygotes ont habituellement 50 % de déficit enzymatique et ne sont pas malades. Lorsqu'ils sont malades et hétérozygotes pour une mutation sur le gène, ils présentent un déficit enzymatique de 15–25 %. Il a été démontré que la pathologie était liée à l'expression d'un polymorphisme fonctionnel sur le gène en *trans* du gène muté (Fig. 6). Ce polymorphisme provoque une diminution accrue de l'expression du gène en synergie avec l'absence d'expression liée au gène muté [29].

Un cas particulier : le phénomène de Knudson

En 1971, le docteur A.G. Knudson qui étudiait un cancer héréditaire rare, une tumeur maligne le plus souvent bilatérale de la rétine appelée rétinoblastome (Omim 180200) émet une hypothèse appelée par la suite théorie de Knudson (largement démontrée depuis) [40]. Deux formes de rétinoblastomes existent, la forme familiale de transmission autosomique pseudodominante et la forme sporadique (essentiellement unilatérale). Le phénomène de Knudson décrit dans les formes héréditaires, la présence d'une mutation germinale sur le chromosome 13q14 issu d'un des deux parents. Le sujet est donc hétérozygote et non malade. Pour que la maladie se déclare, il est nécessaire qu'une deuxième mutation ou une épimutation (voir infra) apparaisse, muta-

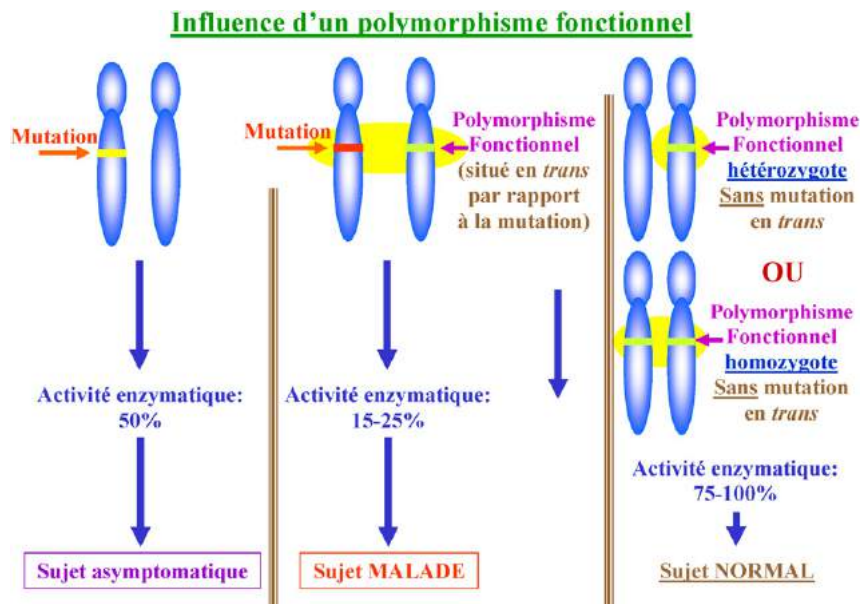


Figure 6 Conséquence d'un polymorphisme fonctionnel. Lorsqu'une mutation est présente à l'état hétérozygote chez un patient (exemple de la protoporphyrie liée à un déficit de l'enzyme ferrochélatase, conséquence d'une mutation sur le gène codant), le sujet présente un déficit enzymatique de 50 % (par rapport à l'activité normale). Il n'est cependant pas malade dans l'immense majorité des cas. Lorsqu'en *trans* du gène muté, se trouve un polymorphisme fonctionnel (IVS3-48C/T), l'activité résiduelle de la ferrochélatase est de 15–25 %. Le polymorphisme fonctionnel est lui-même responsable d'une diminution modérée de l'activité de l'enzyme (qui reste dans les limites de la normale lorsque le polymorphisme est isolé). L'association de la mutation et du polymorphisme fonctionnel entraîne ainsi une diminution de l'activité telle qu'elle se retrouve en dessous d'un seuil dont la conséquence est l'apparition de la maladie.

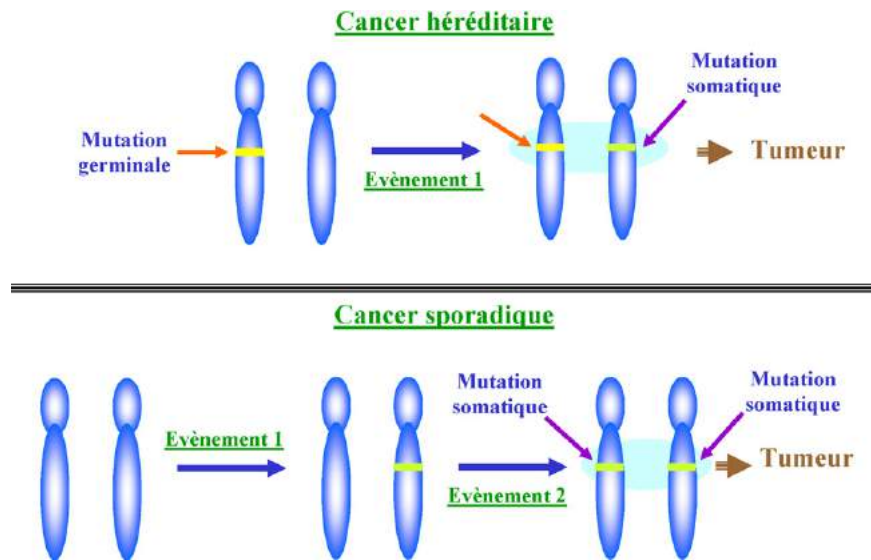


Figure 7 Théorie de Knudson. Dans certains cancers héréditaires, (par exemple, rétinoblastome, le cancer du rein familial), les sujets malades présentent une mutation germinale sur un des deux chromosomes issus du père ou de la mère. Un deuxième évènement survient au cours de la vie, une mutation somatique (acquise) sur le même locus de l'autre chromosome. Il peut s'agir d'une mutation sur la séquence de l'ADN ou d'une méthylation anormale (épimutation). Le gène codant pour une protéine intervenant dans le cycle cellulaire (gène suppresseur de tumeur) est inactivé. Le gène n'étant plus fonctionnel, une tumeur se développe. En revanche, ces mêmes cancers peuvent être acquis (donc, non héréditaires). Dans ce cas-là, deux évènements successifs sur le locus de chaque chromosome sont nécessaires pour que la tumeur se développe.

tion ou épimutation somatique et non germinale. Le gène étant un gène suppresseur de tumeur, la deuxième mutation sur l'autre chromosome inactive le gène provoquant alors le développement d'une tumeur (Fig. 7).

L'inactivation de l'X

L'homme est 46XY (46 chromosomes dont un chromosome X et un chromosome Y) et la femme 46XX (46 chromosomes dont 2 chromosomes X). Le chromosome Y est un petit chromosome ne contenant que 80 gènes et deux petites régions homologues au chromosome X, les régions pseudo-autosomiques du chromosome Y (constituées de 29 gènes). Le chromosome X est beaucoup plus important (5% du génome et 1090 gènes identifiés). Pour pallier à cette différence du nombre de chromosomes et du risque de surdosage génique (liée à l'expression potentielle des gènes sur chacun des 2 chromosomes X chez la femme contre un pour l'homme), un mécanisme a été démontré : la femme possède deux chromosomes X dont l'un est inactivé de manière aléatoire dans chaque cellule : ce dernier peut être identifié physiquement (masse compacte intranucléaire appelée corpuscule de Barr) : la femme possède donc un chromosome X actif (issu du père ou de la mère) et un chromosome X inactif (issu du père ou de la mère) [43]. On parle d'inactivation de l'X. Chez la femme, il s'agit donc d'un mosaïcisme somatique naturel. Normalement aléatoire, ce mécanisme peut ne pas l'être chez 1 à 20% des femmes selon les études et augmente d'ailleurs avec l'âge de la femme. On la retrouve chez moins de 7% des femmes de moins de 25 ans et chez 16% des femmes à 60 ans. Sans rentrer dans les détails du mécanisme, l'inactivation de l'X est la conséquence d'une méthylation de la presque totalité du chromosome X ainsi que des modi-

fications épigénétiques des histones [56]. Elle a lieu sur 85% de l'X inactivé (expression mono-allélique). Par conséquent, environ 15% des gènes sur l'X inactif sont en fait actifs et 10% supplémentaires partiellement inactivés (variable d'une femme à l'autre) [20]. Complexe et partiellement élucidé, le mécanisme est lié en partie à l'action d'un ARN non codant sur l'X inactivé, ARN exprimé à partir du centre d'inactivation de l'X (*Xic*, *X inactivation center*) et appelé *Xist* (*X inactivation specific transcript*). Le mécanisme de maintenance de l'inactivation de l'X est complexe et n'est pas complètement résolu [56,65].

L'expression mono-allélique sur les autosomes

De la même façon qu'il existe une inactivation au hasard d'un des deux allèles sur le chromosome X, on estime que 5–10% des gènes sur les chromosomes non sexuels (les autosomes) sont exprimés par un seul des deux allèles [27]. Des mécanismes épigénétiques ont été démontrés en ce qui concerne l'expression mono-allélique de certains de ces gènes (voir infra). Les mécanismes en cause, à différencier de l'empreinte parentale, se retrouvent par exemple dans le réarrangement des gènes d'immunoglobulines, dans l'expression des récepteurs d'antigènes du système immunitaire et dans l'expression des récepteurs du système olfactif.

Phénotype et gènes modificateurs

L'étude des maladies monogéniques a permis de démontrer l'existence d'une hétérogénéité génétique et d'une variabilité phénotypique d'un grand nombre de maladies héréditaires. On observe ainsi que dans certaines pathologies héréditaires la pénétrance, l'expressivité et la sévérité d'une maladie varient souvent chez les sujets porteurs d'une

même mutation y compris à l'intérieur d'une même famille (variabilité inter- et intrafamiliales). L'existence de gènes ou de loci modificateurs a donc été suggérée puis démontrée, même si l'influence précise et leur nature sont encore mal connues. Un gène modificateur est un gène dont le produit affecte le phénotype d'une maladie liée à un autre gène en interagissant soit directement soit sur une voie biologique alternative de la voie pathologique liée au gène causal de manière fonctionnelle et/ou quantitative. La conséquence peut en être un phénotype plus sévère ou moins sévère. Par ailleurs, le plus souvent, pour une maladie déterminée, plusieurs gènes modificateurs peuvent agir en combinaison sur le phénotype (combinatoire d'action). Ils jouent un rôle dans la pénétrance et l'expressivité de la maladie (par exemple, âge d'apparition, sévérité des symptômes, signes cliniques et/ou biologiques). Ils jouent un rôle important dans la relation génotype–phénotype. À titre d'exemple, dans la maladie polykystique rénales de transmission autosomique dominante (Omim 173900), la sévérité de la maladie est la conséquence de facteurs environnementaux mais aussi génétiques. Selon les études, on estime que les gènes modificateurs sont responsables de 43–78 % de l'âge d'apparition du stade terminal de l'insuffisance rénale observée dans cette pathologie [60]. Dans la mucoviscidose, l'existence de gènes modificateurs a ainsi été rapidement suspectée. En effet, bien que les mutations décrites dans le gène codant pour *cystic fibrosis transmembrane conductance regulator* (CFTR), Omim 602421 responsables de la mucoviscidose présentent une certaine corrélation avec l'atteinte pancréatique, il n'en était pas de même avec l'atteinte pulmonaire qui fait la gravité de la maladie. En effet, l'atteinte pulmonaire est plus difficile à prédire. L'influence d'autres gènes a alors été soulevée. Des gènes modificateurs jouant un rôle dans les manifestations digestives et influençant la gravité des signes pulmonaires ont ainsi été démontrés [3,8]. Il est important de noter qu'il existe une subtile différence entre gène modificateur et gène de susceptibilité. Un gène modificateur est un variant génétique qui module l'expression génétique d'une maladie. Un gène de susceptibilité est un variant génétique prédisposant à une maladie (en général, le plus souvent dans les maladies multifactorielles, voir infra).

L'ensemble des mécanismes brièvement exposés ci-dessus montrent la complexité des maladies dites monogéniques. Les nombreuses études démontrent en fait l'existence d'un continuum entre les maladies dites monogéniques et les autres maladies héréditaires plus complexes telles que les maladies à transmission oligogénique, polygéniques et multifactorielles.

Les maladies oligogéniques

La difficulté à établir une relation génotype–phénotype dans un certain nombre de maladies comme la mucoviscidose, par exemple, démontre que les maladies monogéniques à transmission mendélienne sont certes des modèles utiles pour décrire un grand nombre de maladies héréditaires mais que ces modèles sont incomplets. L'évolution de la connaissance des gènes et de leurs pathologies a amené à découvrir certaines maladies que l'on croyait monogéniques à être en fait la conséquence de l'action/modulation d'un petit nombre d'autres loci. On

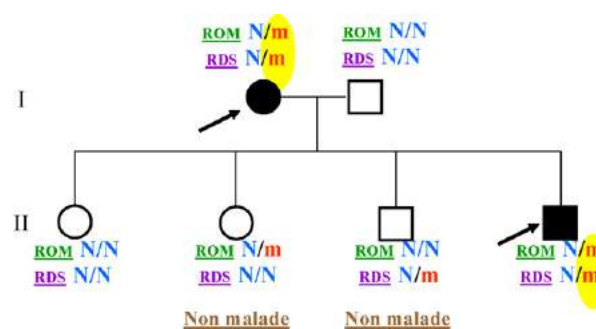


Figure 8 Transmission oligogénique. Dans cet arbre généalogique, un homme et une femme (flèche) sont atteints d'une forme de rétinite pigmentaire (Omim 268000). Parmi les nombreuses formes de cette maladie, des cas de transmission digénique ont été décrits. Pour qu'un sujet soit malade, il est nécessaire que les deux gènes soient mutés. Un sujet muté hétérozygote pour le gène *retinal outer segment membrane protein 1* (*ROM1*), Omim 180721 ou hétérozygote muté pour le gène *retinal degeneration, slow, mouse, homolog of (RDS)* appelé aussi *peripherin 2, mouse, homolog of (PRPH2)*, Omim 179605) ne sont pas malades. Les deux patients possèdent une mutation à l'état hétérozygote sur les deux gènes *ROM1* et *RDS*. N : allèle normal. m : allèle mute.

parle de maladies oligogéniques. Il s'agit donc de maladies qui ne sont ni monogéniques ni de transmission plus complexe (voir infra). Les maladies oligogéniques sont donc héréditaires mais nécessitent une synergie d'action d'allèles mutés sur un petit nombre de loci (Fig. 8). C'est le cas par exemple dans une forme de surdité de transmission récessive autosomique (*deafness, neurosensory, autosomal recessive 1*; DFNB1 Omim 202290) où certaines formes digéniques ont été décrites à savoir des mutations sur les gènes *GJB2* [*gap junction protein connexin 26*] et *GJB6* [*gap junction protein, beta-6 connexin 30*] [51]. De manière générale, dans les formes oligogéniques, les mutations peuvent avoir l'une sur l'autre une synergie d'action ou agir comme modificateur de l'une par rapport à l'autre. Un autre exemple démontrant la complexité des maladies héréditaires est celui de l'hérédité tri-allélique. Prenons l'exemple d'une maladie rare, le syndrome de Bardet-Biedl (Omim 209900), maladie héréditaire hétérogène caractérisé schématiquement par une obésité, un hypogonadisme, une rétinopathie, une polydactylie, des malformations génito-urinaires et un retard mental [4]. Dans certaines formes de cette maladie rare, l'hypothèse tri-allélique a été formulée [39]. Initialement décrite comme une maladie de transmission autosomique récessive, l'analyse haplotypique et l'étude des mutations démontrent que certaines formes de la maladie n'obéissent pas complètement au modèle récessif attendu. Une étude portant sur un des nombreux gènes impliqués dans cette maladie, le gène *BBS6* (*bardet-biedl syndrome 6*, appelé aussi gène *MKK*, *McKusick-Kaufman*, Omim 604896) illustre ce propos (Fig. 9). Schématiquement, dans l'analyse d'une famille consanguine, le père avait transmis à l'état hétérozygote une mutation à ses deux enfants dont un seul était malade. L'analyse haplotypique a démontré que les deux enfants avaient hérité du même chromosome maternel. Ainsi, même si la mère possédait une mutation non décelée, elle devrait la transmettre

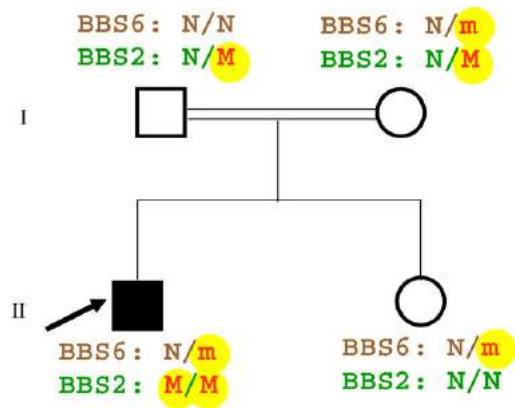


Figure 9 Hérité tri-allélique – Exemple du syndrome de Bardet-Biedl (BBS). Autre exemple de modèle oligogénique : dans cette famille consanguine, le patient (flèche) et sa sœur présentent la même mutation sur le gène *BBS6* à l'état hétérozygote. Le patient présente aussi une mutation à l'état homozygote sur le gène *BBS2* alors que sa sœur ne présente aucune mutation. Le sujet malade a donc trois mutations sur deux gènes (3 allèles mutés). N : allèle normal. M : allèle muté sur le gène *BBS2*. m : allèle muté sur le gène *BBS6*.

aux deux enfants dont un était non malade et avait deux mutations sur le gène *BBS6* à l'état hétérozygote. Cette observation indique que dans certains cas, plus de deux mutations sont nécessaires pour que la maladie se manifeste. Des études ont alors démontré que le sujet malade était aussi homozygote muté sur un autre gène impliqué dans cette maladie, le gène *BBS2* (*bardet-biedl syndrome 2*, Omim 606151). Le sujet malade avait donc deux mutations sur le gène *BBS2* et une mutation sur le gène *BBS6* alors que son frère n'avait aucune mutation sur le gène *BBS2* [3].

L'hérédité non mendélienne

Les lois de Mendel ont participé fondamentalement à la description des maladies monogéniques. Néanmoins, elles ne permettent pas d'expliquer un grand nombre de maladies génétiques pour lesquelles aucune hérédité mendélienne claire ne peut être établie. Les progrès de la génétique moléculaire ont permis la compréhension de nombreuses maladies monogéniques et ont permis la description d'autres modes d'hérédité plus complexes. L'hérédité oligogénique, l'existence de gènes modificateurs, l'hérédité tri-allélique sont des modes de transmission intermédiaires entre l'hérédité mendélienne (plus complexe que les premières descriptions ne laissaient penser) et les hérédités non mendéliennes, terme impropre mais habituellement utilisé, l'existence d'un continuum entre hérédité mendélienne et non mendélienne ayant été démontré (Tableau 1) [32].

L'anticipation

L'anticipation se caractérise par une sévérité accrue d'une maladie héréditaire de génération en génération. Classiquement, elle est la conséquence de l'expansion de trinucleotides au-delà d'un seuil, le nombre de séquences

répétées croissant de génération en génération. La pénétrance de la maladie augmente aussi de génération en génération et apparaît à un âge plus précoce. Ce phénomène est observé par exemple dans la maladie de Steinert (Omim 160900), dystrophie musculaire de transmission autosomique dominante se traduisant sur le plan moléculaire par l'expansion d'un trinucleotide [CTG] situé dans la région 3' non codante du gène codant pour *dystrophia myotonica protein kinase* (DMPK) et dont le nombre de séquences répétées dépasse le seuil de 50 unités pour s'accroître de génération en génération.

L'hérédité épigénétique

La génétique telle que décrite ci-dessus ne peut expliquer toute la variété phénotypique physiologique et/ou pathologique dans une population. Elle ne peut pas expliquer non plus la diversité phénotypique et la susceptibilité différentielle aux maladies observées chez les jumeaux monozygotes « clones naturels ». L'épigénétique explique une partie de ces variabilités non liée à la génétique classique. Elle a été décrite pour la première fois en 1939 et définie à cette époque comme les interactions observées entre les gènes et leur produits dont la résultante est le phénotype de l'individu [61]. Cette définition a évolué. Il s'agit d'un terme général pour désigner les modifications transmissibles de l'expression des gènes qui ne résultent pas de modifications de la séquence d'ADN [35]. Ce terme, définissant un certain nombre de concepts tels que l'empreinte génétique ou les paramutations, a également été employé pour décrire la propagation stable de programmes d'expression de gènes au cours de la division cellulaire (également appelée hérédité épigénétique mitotique). Pour aborder l'épigénétique, nous rappellerons quelques notions de base. L'ADN et les histones qui les entourent constituent les nucléosomes, éléments de base de la chromatine (147 pb d'ADN entouré d'un octamère de quatre protéines histones H2A, H2B, H3 et H4). Des modifications covalentes de l'ADN et des histones peuvent se produire à l'aide d'enzymes spécifiques (par exemple, des ADN méthyl-transférases) : ces modifications sont dites épigénétiques. Certaines de ces modifications aboutissent à une modulation de l'expression des gènes [18]. Les modifications covalentes observées au niveau de l'ADN sont restreintes à la méthylation des cytosines qui précèdent des guanines au niveau de régions riches en dinucleotides CG et appelés îlots CpG. Les îlots CpG sont caractérisés par une taille minimum de 200 pb, par un contenu en [G+C] d'au moins 50% et sont situés essentiellement en 5' des gènes dans les zones de régulation de l'expression de ces derniers. Ils représentent environ 0,7% du génome humain et contiennent environ 7% des dinucleotides CpG. 60% des promoteurs de gènes possèdent des îlots CpG (dont la plupart ne sont pas méthylés). Ces modifications aboutissent en général à l'abolition de l'expression du gène méthylé (mais, ce n'est pas systématique). Les histones présentent plusieurs types de modifications possibles tels que l'acétylation, la méthylation, l'ubiquitination ou la phosphorylation. Selon la nature des modifications, on peut observer une abolition ou une activation de l'expression d'un gène. La plupart des modifications observées aboutissent à une interaction complexe entre la chromatine et les facteurs qui régulent l'expression des gènes. Les modifications épigénétiques de l'ADN sont

Tableau 1 Exemples de traits influençant la variabilité de l'hérédité [32].

Influence sur la transmission héréditaire	Principe	Exemple de pathologie	Conséquence sur le phénotype
Gène modificateur	Un gène différent situé sur un autre locus peut modifier le phénotype de la maladie	Mucoviscidose	Sévérité de l'atteinte pulmonaire Réponse à l'infection
Hérédité di-génique	Le phénotype résulte de l'interaction entre deux génotypes	Certaines formes de surdité	Sévérité de la perte auditive
Hérédité tri-allélique	3 mutations situées sur des gènes en 2 loci différents ségrége avec l'expression de la maladie	Syndrome de Bardet-Biedl	La présence d'une troisième mutation modifie le phénotype
L'empreinte	Un gène au même locus s'exprime différemment selon l'origine parentale	Syndrome de Beckwith-Wiedemann	Phénotype dépendant de l'origine parentale de la mutation

transmissibles d'une génération à l'autre tant au cours de la mitose que de la méiose [45]. Elles sont également réversibles. La séquence de l'ADN est par ailleurs normale. Le niveau de méthylation dans un tissu est variable de cellule à cellule. L'épigénétique en général (en incluant le processus de méthylation de l'ADN) joue un rôle important dans un grand nombre de fonctions cellulaires telles que l'expression tissu-spécifique de certains gènes, la différenciation cellulaire, l'empreinte génétique (voir infra), l'inactivation des séquences exogènes d'insertion (par exemple, les transposons), l'inactivation du chromosome X, la régulation de la structure chromatinienne, la carcinogenèse, le vieillissement et le développement embryonnaire [6]. Les anomalies épigénétiques sont responsables d'un certain nombre de pathologies. Ces anomalies peuvent être héréditaires (cf anomalies de l'empreinte parentale) ou acquises (cf certains cancers comme celui du colôn) [55]. En effet, la méthylation de promoteurs de certains gènes notamment ceux impliqués dans la régulation du cycle cellulaire provoque leur inactivation. Cette dernière participe au mécanisme d'oncogenèse aboutissant à la formation d'une tumeur. Un certain nombre de maladies génétiques associent anomalies épigénétiques et mutations sur l'ADN : c'est le cas par exemple du syndrome de l'X fragile (Omim 300624), conséquence de la répétition de la séquence tri-nucléotidique [CGG] supérieure à 200 fois et dans lequel est observé le phénomène d'anticipation. Chez un sujet normal, cette séquence est répétée moins de 50 fois. Cette séquence est située sur la partie 5' non codante du gène. Elle est aussi hyperméthylée aboutissant alors à l'inactivation du gène. Deux anomalies coexistent donc, l'augmentation anormale du nombre de séquences répétées et l'hyperméthylation du site.

Relation génome–épigénome. Les régions de méthylation de l'ADN et d'interactions avec les histones modifiées constituent l'épigénome. Elles sont encore mal connues. On sait par exemple que le développement précoce au cours de l'embryogenèse se traduit par une reprogrammation épigénétique du génome humain, fondamental pour le développement à venir. Certaines régions génomiques interagissant avec l'épigénome commencent à être identifiées [6]. Par ailleurs, de nombreuses études dont celles portant

sur des jumeaux monozygotes discordants ont démontré le rôle important de l'épigénétique en plus des manifestations environnementales dans le développement des maladies [23].

L'empreinte génétique. Chaque cellule eucaryote possède deux copies d'un gène sur chaque chromosome issu du père et de la mère. À chaque locus chromosomique, les allèles normaux sont transcrits et fonctionnent de la même manière. L'empreinte génétique est une exception à cette règle générale [64]. L'empreinte est une forme de régulation d'un gène dépendant de l'origine parentale. L'expression dans ce cas-là est mono-allélique (un seul des deux gènes parentaux est exprimé). Cette empreinte parentale a lieu en général au moment de la gamétogenèse. Lorsque l'embryon est constitué et avant son implantation, la programmation épigénétique réalisée au cours de la gamétogenèse est effacée dans sa quasi-totalité [55]. Après l'implantation de l'œuf, une reprogrammation de novo de la méthylation a lieu (elle peut aboutir dans certains cas à des erreurs de méthylation). Son rôle dans le développement humain est considérable. Cette régulation est transmise à partir de zones spécifiques du génome appelées région de contrôle de l'empreinte, ou *imprinting control regions* (ICRs). Les ICRs sont des régions où existe une méthylation différentielle de l'ADN mais aussi des histones (méthylations et acétylations différentielles) ainsi que des modifications locales de la structure chromatinienne. Le mécanisme de méthylation différentielle de cette région explique l'expression parentale différentielle des gènes en dépendant ainsi que leur régulation. Certaines maladies héréditaires sont la conséquence d'une anomalie de fonctionnement de ce mécanisme. C'est le cas par exemple, du syndrome de Beckwith-Wiedemann (Omim 130650), syndrome complexe caractérisé notamment par une grande taille à la naissance, une macroglossie, un omphalocèle, une hypoglycémie néonatale, un syndrome dysmorphique et une augmentation du risque de certaines tumeurs. La majorité des cas sont sporadiques mais la transmission est autosomale dominante dans environ 15% des cas. Plusieurs mécanismes expliquent cette pathologie hétérogène sur le plan génétique et notamment des anomalies de méthylation [32]. Par ailleurs, il a été démontré que dans l'empreinte

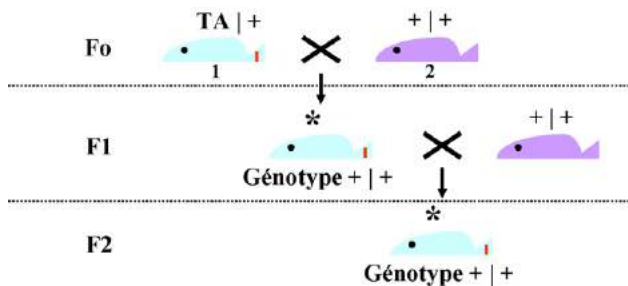


Figure 10 Paramutation chez la souris [11]. À la première génération (F0), la souris 1 (mâle ou femelle) est hétérozygote pour l'allèle TA (TA/+), allèle ne permettant pas la production de la protéine KIT, tyrosine kinase impliquée notamment dans la mélanogenèse. Elle possède des pattes blanches ainsi que l'extrémité de la queue blanche (bande rouge). En croisant celle-ci avec une souris normale (souris 2 : +/+), on obtient à la génération F1 des souris génétiquement normales homozygotes +/+ et possédant des pattes blanches ainsi que l'extrémité de la queue blanche. Le niveau de l'ARNm de KIT est diminué à un niveau similaire à celui de son parent hétérozygote 1. Ce descendant au phénotype paramutant est désigné *. La fréquence observée lors du croisement n'est pas de 100% et varie selon les croisements effectués. Par ailleurs, l'extension des zones blanches observées varie d'une souris à l'autre. À la génération suivante (F2), après un croisement de la souris paramutante et d'une souris normale (+/+), on observe encore des souris au phénotype paramutant. Le taux de souris paramutantes est néanmoins inférieur à la génération F1. On constate qu'avec les générations successives, la fréquence du paramutant diminue progressivement. Ces modifications transmissibles sont liées à l'expression d'ARN non codants modifiant l'expression du gène KIT. Ces ARN seraient transmis par les gamètes et provoqueraient la dégradation des ARNm codés à partir du gène KIT chez les paramutés [53].

génomique, mécanisme complexe, certains ARN non codants jouent un rôle déterminant notamment en participant à l'inactivation d'un des deux allèles. Ils participent donc au contrôle de l'expression monoallélique dans ce mécanisme [66].

Les épipolymorphismes. Les variations épigénétiques ont été démontrées chez les végétaux et ont été observées dans les études de jumeaux monozygotes et sur le vieillissement chez les êtres humains. Il est encore difficile cependant de différencier ce qui pourrait être acquis de ce qui est héréditaire ainsi que de la possibilité de transmettre ces épipolymorphismes d'une génération à l'autre. Cependant, des preuves s'accumulent et des études en cours cherchent à démontrer définitivement leur existence chez les êtres humains [54].

La notion de paramutation. Décrite pour la première fois en 1956, la notion de paramutation a évolué avec les connaissances de génétique moléculaire [9]. Une paramutation est un transfert épigénétique d'une information d'un allèle à un autre (devenant ainsi « un épi-allèle », Fig. 10). Ce transfert d'information induit un état d'expression variable quantitativement d'un gène, transmissible de génération en génération [11]. La séquence de l'ADN n'est pas modifiée (il ne s'agit pas d'une mutation). Plutôt

qu'un effet oui/non (*on/off*) de l'expression d'un gène, on observe une modulation quantitative de son expression [16]. Une paramutation possède trois caractéristiques principales :

- le nouvel état d'expression d'un gène est transmissible d'une génération à l'autre même si avant la paramutation, l'état d'expression n'était pas transmissible ;
- le locus ainsi modifié agit de la même façon sur des séquences homologues ;
- les modifications ne sont pas observées sur la séquence d'ADN. Les nouvelles instructions d'expression du gène ainsi que la mémoire acquise sont donc liées à des modifications épigénétiques.

Observées chez des plantes (comme le maïs) et la souris, ce mécanisme n'a pas encore été objectivé chez les humains. À priori, il n'y a aucune raison de croire que cela n'existerait pas même s'il est nécessaire de le démontrer. Le rôle joué par les ARN non codants dans ce mécanisme bien qu'encore mal connu est fondamental. Le transfert d'ARN par les gamètes a ainsi été démontré. Le sperme humain contient de grandes quantités d'ARN [46]. Alors que dans les plantes, de nombreux exemples existent, les paramutations sont-elles rares chez les mammifères ? Est-ce un phénomène « anecdotique » ou sont-elles fréquentes ? À ce jour, il n'y a pas de réponse à ces questions [16].

Les épimutations acquises.

Le rôle de l'épigénétique dans les cancers. Sans rentrer dans une description complète du rôle de l'épigénétique dans les cancers, nous allons décrire certains aspects notables notamment leur place et leur interaction avec la génétique. Lorsque l'on compare les cellules tumorales et normales, le niveau globale de la méthylation dans les cellules cancéreuses est bas comparé à celui des cellules non tumorales [21]. Cette hypométhylation est en fait la conséquence d'une déméthylation des séquences répétées et de certains introns et régions codantes. Cette hypométhylation augmente progressivement au cours de la transformation de la tumeur bénigne vers son évolution cancéreuse. Il en résulte une instabilité chromosomique, une réactivation des transposons et une perte d'empreinte génomique. La déméthylation peut aussi provoquer l'expression de gènes habituellement réprimés. Par exemple, l'hypométhylation du gène *PAX2* (Omim 167409), gène codant pour un facteur de transcription (impliqué notamment dans la néphrogenèse au cours de l'embryogenèse et dans la prolifération cellulaire) provoque son activation et favorise le développement du cancer de l'endomètre. Néanmoins, il a été démontré que c'est en fait l'hyperméthylation d'ilôts CpG des promoteurs des gènes suppresseurs de tumeur qui constituait un événement majeur à l'origine de nombreux cancers. C'est le cas par exemple dans le cas des cancers du côlon où l'hyperméthylation du gène *hMLH1* (*Mutl*, *E. coli*, *homolog of*, *1*; Omim 120436) provoque son inactivation, ce gène étant impliqué dans la réparation de l'ADN. Il en résulte l'absence de réparation de certaines erreurs notamment lors de la mitose cellulaire ouvrant ainsi la voie vers la transformation néoplasique. De nombreux gènes impliqués dans le cycle cellulaire, la réparation de l'ADN, le métabolisme des carcinogènes, l'apoptose,

les interactions cellulaires, l'angiogenèse peuvent être méthylés et par conséquent inactivés [19]. Par ailleurs, les anomalies de méthylation peuvent aussi toucher les zones régulant l'expression des microARN (ARN non codant impliqués dans la régulation de l'expression de nombreux gènes). Certains de ces microARN jouent un rôle de suppresseur de tumeur. L'hyperméthylation provoque leur inactivation [10]. En général, les profils anormaux de méthylation sont spécifiques d'un type de cancer définissant ainsi un hyperméthylome. Ces profils de méthylation observés dans les cancers sporadiques (non héréditaires) peuvent aussi se voir dans des cancers héréditaires. En parallèle à ces modifications de méthylation, des modifications des histones se produisent. En fait, la combinaison des dysméthylations associées aux modifications complexes des histones aboutissent au développement progressif de la tumeur. L'association de ces modifications épigénétiques aux mutations somatiques d'autres gènes aboutissent à une cancérisation progressive des cellules atteintes.

Épigénétique et environnement. L'environnement joue un rôle dans les modifications épigénétiques du génome. L'étude de jumeaux monozygotes (même patrimoine génétique) a démontré qu'il existait entre eux des modifications épigénétiques de leur génome, modifications liées à l'environnement (par exemple, le tabac) [23]. Par ailleurs, certaines études tendent à démontrer que des caractères acquis liés à l'environnement provoquent des modifications épigénétiques transmissibles. Bien que cela ait été démontré chez un ver de terre, *Caenorhabditis elegans* [41], on peut penser qu'un tel mécanisme acquis puisse exister chez les humains : cela reste à démontrer.

Les ARN non codants. Comme nous l'avons déjà rappelé, une faible part du génome est constituée de gènes codants. Un certain nombre de gènes code pour un ARN, on parle alors d'ARN non codant (ARNnc). Parmi ces derniers outre les ARN ribosomiaux (ARNr) et les ARN de transfert (ARNt), on peut citer :

- les microARNs (miARN) et les *small interfering* ARN (siARNs) : le rôle de ces ARN est complexe [69]. Ils interviennent notamment dans la régulation post-traductionnelle des gènes codants, dans la dégradation de certains ARNm et de manière plus globale dans la régulation épigénétique de nombreux gènes. Par exemple, ils jouent un rôle dans la méthylation de l'ADN, les modifications des histones et la conformation de la chromatine. Ils participent aussi à l'inactivation de l'X et l'inactivation des transposons [15] ;
- les petits ARN nucléaires ou snRNAs ;
- les petits ARN nucléolaires ou snoRNA : ces petits ARN sont souvent issus d'introns d'ARNm (chez les eucaryotes évolués), ce qui démontre l'importance fonctionnelle des introns et l'organisation complexe des gènes. Ces ARN agissent sur les ARNr et modifient leurs structures en association avec des complexes protéiques. Ils participeraient aussi à la régulation de l'épissage alternatif de certains gènes [28] ;
- les composants ARN de ribonucléases (RNases) telles que les RNases P et autres RNases ;
- les ARN non codants similaires aux ARNm, les mlARN : ce sont des ARN dont la structure est proche des ARNm mais

qui ne présentent pas de cadre ouvert de lecture (*open reading frame* [ORF]). Bien qu'encore mal connus, ces ARN semblent avoir une fonction proche des miARN et semble interagir avec ceux-ci dans un réseau de régulation complexe miARN/mlARN [69] ;

- les *piwi-interacting* (ARN piARN), les piwi étant une classe de protéines de la famille Argonaute impliquée dans la spermatogenèse). Ces piARN sont aussi appelés *repeat-associated* ARN (raARN). Ils sont importants pour le contrôle des transposons. Ainsi, chez la souris, on les retrouve associés aux rétrotransposons SINES, LINEs et LTRs.

Cette description simplifiée des ARN non codants démontre une fois de plus que la génétique ne se limite pas aux gènes codants et que de nombreuses interactions complexes modulent le phénotype global d'une cellule, d'un tissu et finalement de l'organisme entier.

Les maladies polygéniques et multifactorielles

De nombreuses maladies appelées improprement « communes » dans de nombreux ouvrages telles que le diabète, l'hypertension artérielle, l'obésité, de nombreux cancers, certaines maladies mentales comme la schizophrénie sont en fait des maladies complexes, conséquences de l'interaction de nombreux acteurs ou facteurs (d'où le terme multifactoriel) parmi lesquels des facteurs génétiques. Les études notamment chez les jumeaux et dans un grand nombre de familles démontrent pour nombre d'entre elles une prédisposition génétique. On retrouve des histoires familiales, sans schéma de transmission héréditaire évident, une pénétrance et une expressivité variables. Les outils bioinformatiques et notamment statistiques ont largement prouvé le rôle important (mais non suffisant) joué par les facteurs génétiques. Les études d'épidémiologie génétique et de biologie moléculaire jouent un rôle majeur dans ces déterminations. Grâce à ces études, on sait désormais qu'un seul gène ne suffit pas à expliquer ces maladies fréquentes. Mais, les prédispositions génétiques observées dans ces maladies sont la conséquence de variations mineures dans un grand nombre de gènes. Ces variations, isolées, ne suffisent pas à provoquer une maladie. Leur association (combinatoire) et le résultat de leur interaction avec d'autres facteurs externes (agents infectieux, environnement, toxiques...) dans le métabolisme cellulaire, tissulaire et à l'échelle de l'organisme, déclenchent la maladie [42]. Ce principe est appelé modèle « allèle commun-maladie commune ». D'autres études ont proposé un schéma alternatif (mais, non exclusif du précédent) pour certaines maladies multifactorielles (par exemple pour la schizophrénie). Certaines mutations prédisposeraient à la schizophrénie et présenteraient une forte pénétrance. Ces mutations aboutiraient à l'abolition de l'expression de plusieurs gènes sur une ou plusieurs voies métaboliques impliquées dans le développement neuronal. Elles seraient rares à l'échelle individuelle et d'origine récente dans l'évolution humaine [62]. Pour illustrer la complexité des facteurs génétiques en cause dans les maladies multifactorielles, prenons l'exemple de la maladie de Crohn (Omim 266600). C'est une maladie chronique inflammatoire du tube

Tableau 2 Exemples de gènes impliqués dans la maladie de Crohn, maladie multifactorielle.

Chromosome	Gène impliqué
1p31	<i>IL23R</i>
2q37	<i>ATG16L1</i>
3p21	Nombreux gènes dont <i>MST1</i>
5p13	<i>PTGER4</i>
5q31	Nombreux gènes dont <i>SLC22A5</i>
5q33	Nombreux gènes dont <i>IRGM</i>
5q33	<i>IL12B</i> (fraction p40)
10q21	<i>ZNF365</i>
10q24	<i>NKX2-3</i>
16q12	<i>NOD2</i>
17q21	Nombreux gènes dont <i>STAT3</i>
18p11	<i>PTPN2</i>

ATG16L1: autophagy related 16-like protein 1; *IL12B*: interleukin 12 beta; *IL23R*: interleukin-23 receptor; *IRGM*: immunity-related GTPase family M; *NKX2-3*: NK2 transcription factor related locus 3; *NOD2*: nucleotide-binding oligomerization domain protein 2; *PTGER4*: prostaglandin receptor EP4; *PTPN2*: protein tyrosine phosphatase non-receptor type 2; *SLC22A5*: solute carrier family 22, member 5; *STAT3*: signal transducer and activator of transcription 3; *ZNF365*: zinc-finger protein 365.

digestif, conséquence d'un dysfonctionnement de la réponse immunitaire aux bactéries commensales du tube digestif. De nombreux facteurs génétiques et loci chromosomiques ont été observés dans cette pathologie (5–20 % des patients ont une histoire familiale). Le **Tableau 2** résume quelques gènes impliqués [13]. La bibliographie sur ce vaste sujet démontre l'importance des recherches sur les maladies polygéniques et multifactorielles. On peut observer que les technologies des puces d'ADN révolutionnent le monde de la génétique et notamment des maladies dans lesquelles une composante génétique existe (soit l'immense majorité des maladies humaines) [7]. Grâce à ces technologies (étudiant notamment les SNPs répartis sur l'ensemble du génome humain), les études d'association tout génome deviennent plus exhaustives et à un coût de plus en plus faible (les puces commercialisées par les sociétés leaders dans le monde, Illumina et Affymetrix atteignent un prix d'environ 400–800 \$ US en 2008) [30]. En parallèle d'autres études tentent d'établir des réseaux moléculaires (génotypiques) plus particulièrement en étudiant les variations génotypiques de l'ADN et en évaluant les conséquences phénotypiques que ces variants peuvent provoquer à l'échelle d'un système global (cellulaire, tissulaire...). Ces études s'appuyant sur les expériences (moléculaires, cellulaires et métaboliques) et les modèles bio-informatiques sont fondamentales pour comprendre comment les combinatoires moléculaires même mineures en apparence peuvent provoquer une pathologie [5,12,5,50]. On parle alors de biologie des systèmes.

La notion d'interaction épistatique

L'épistasie (appelée aussi interaction gène–gène) est un phénomène complexe dans lequel les effets d'un gène sur un trait biologique dépendent de l'action d'un ou plusieurs gènes [47,48]. Ce phénomène s'observe en particulier

dans les réseaux génétiques. Les interactions épistatiques permettraient de tamponner l'action des mutations sur un phénotype déterminé. L'architecture génétique serait telle que le réseau génétique présenterait des redondances (des gènes aux effets similaires) limitant ainsi l'action de mutations sur le phénotype global résultant de ce même réseau. Un effet visible sur le phénotype ne s'observerait alors qu'en présence de plusieurs mutations dans le réseau concerné. L'épistasie permettrait donc de moduler les conséquences d'une mutation par l'existence d'une dépendance entre les gènes d'un même réseau. On parle d'épistasie génétique. Le même principe s'applique au niveau biologique (du fait de l'interdépendance biologique cellulaire–génétique) [47]. Comme cela a été évoqué dans le paragraphe ci-dessus, les interactions épistatiques jouent un rôle important dans les maladies polygéniques. Ainsi, si on raisonne du point de vue épistatique, les conséquences d'un polymorphisme ne peuvent être analysées isolément mais en association avec d'autres polymorphismes afin de déterminer les conséquences phénotypiques résultant de ce polymorphisme (en association avec les autres). Comme on peut le voir, les interactions épistatiques quand elles existent, ajoutent une couche de complexité supplémentaire à l'analyse des maladies polygéniques [47].

Le rôle de l'environnement

Nous avons déjà évoqué le rôle de l'environnement dans les modifications épigénétiques. Il nous semble important de souligner l'importance de celui-ci en sachant qu'il n'est pas toujours facile expérimentalement de différencier et donc d'attribuer ce qui est observé aux facteurs génétiques et/ou aux facteurs environnementaux [26]. De nombreux éléments influent sur l'expression des gènes parmi lesquels on peut citer par exemple, les infections, les toxiques, les stress (y compris psychologique), le régime alimentaire, les médicaments.

Un exemple de maladie multifactorielle : la sclérose en plaques

Pour illustrer brièvement la complexité des maladies multifactorielles et la difficulté à décrypter celles-ci, nous prendrons l'exemple d'une maladie inflammatoire chronique et neurodégénérative du système nerveux central (SNC), la sclérose en plaques (SEP) [49]. Schématiquement, elle se traduit le plus souvent par des poussées inflammatoires focales rémittentes du SNC pouvant évoluer vers une forme progressive chronique, cette dernière forme pouvant plus rarement se développer d'emblée. Le pronostic à long terme est souvent mauvais. En moyenne, 15 ans après le diagnostic, plus de 80 % des patients ont des handicaps fonctionnels et/ou cognitifs et 50–60 % nécessitent une aide pour se déplacer. Cette pathologie présente deux composantes pathogéniques principales environnementale et multigénétique.

Les facteurs environnementaux. Les études de prévalence et d'incidence de la maladie dans différentes populations et régions du monde ont démontré l'importance de ces facteurs et ont permis d'émettre certaines hypothèses. Au moins deux millions de personnes sont touchées dans le monde avec une fréquence plus importante dans le nord de l'hémisphère, une population malade essentiellement

blanche et féminine (plus particulièrement d'origine nord-européenne). L'incidence de la maladie augmente depuis un siècle, temps trop court pour qu'une influence génétique ait pu être modifiée. L'incidence de la maladie est particulièrement élevée en Europe du nord (1–2/1000) par opposition au sud de l'Europe. Il semblerait qu'au cours de la période de vie comprise entre 5 et 15 ans, un agent déclenchant (à découvrir) participerait à la genèse de la maladie. Elle est rare par ailleurs chez les Asiatiques, les Africains et les indigènes d'Australie, d'Amérique et de Nouvelle-Zélande. Ces observations sont en faveur d'un agent pathogène dont la distribution géographique varie. Elles sont cependant également compatibles avec des facteurs génétiques ethniques différents. Une autre hypothèse pour expliquer l'effet lié à la latitude est l'exposition au soleil (et donc à la quantité de vitamine D synthétisée). Une diminution de synthèse de cette vitamine dans le nord de l'hémisphère provoquerait une diminution de son effet sur la régulation immunitaire. Par ailleurs, parmi les agents pathogènes impliqués, de nombreux virus ont été suspectés (rougeole, rubéole, varicelle, oreillons...). Un d'entre eux, le virus d'Epstein-Barr (EBV) semble jouer un rôle important dans la physiopathologie de la maladie [57]. En parallèle, les gènes de susceptibilité à l'infection par l'EBV font l'objet de recherches pour étayer cette hypothèse intéressante.

Les facteurs génétiques. En Europe du nord, de nombreuses études ont démontré que le risque de développer une SEP pour un apparenté est 15 à 20 fois supérieur à celui de la population générale. Un terrain génétique a donc été largement démontré. Non seulement l'existence des gènes de susceptibilité pour la maladie sont démontrés par les études de transmission mais celles-ci révèlent également que des gènes modulant le phénotype et l'évolution de la maladie existent [36,49]. Par ailleurs, les études chez les jumeaux monozygotes montrent un risque élevé (30% de concordance) par rapport aux jumeaux dizygotes (5% de concordance) confirmant l'origine génétique complexe de la maladie [63]. Les études actuelles déterminent la part des allèles dits communs (polymorphismes) des allèles rares (variants) dans la susceptibilité à la maladie. Dans tous les cas, une transmission mendélienne de la maladie a été écartée et les études actuelles s'orientent vers l'association de polymorphismes « communs » dans l'origine multigénique de la maladie. Les nombreuses études d'analyse de liaison et d'études d'association réalisées notamment par le consortium international d'études génétiques sur la SEP, l'*International Multiple Sclerosis Genetics Consortium* (IMSGC) ; site : imgsc.org ont démontré le rôle majeur joué par le locus HLA de classe II (plus précisément, au niveau du locus HLA DR, le gène *HLA DRB1*) et ont estimé que ce locus comptait pour 20–60% de la susceptibilité à la SEP [34]. L'étude de cette région complexe est rendue difficile par l'existence de nombreux déséquilibres de liaison et des variations génotypiques importantes (région hautement polymorphique). La région associée à la SEP mesure en effet 1 mégabase et même si certains génotypes ressortent (ainsi que la certitude de la causalité de cette région dans la SEP), les débats continuent pour savoir quels génotypes sont réellement en cause et pour savoir si la région HLA proche à savoir le locus HLA de classe III, ne joue pas aussi un rôle. L'étude fine du locus de classe II a néanmoins démontré l'existence d'haplotypes impliqués à des

Tableau 3 Exemples de gènes impliqués dans la sclérose en plaques en dehors du locus HLA DRB1.

Chromosome	Gène impliqué	Risque relatif (<i>odds ratio</i>)
1p13	<i>CD58 (LFA3)</i>	1,24
1p22	<i>RPL5</i>	1,15
1p22	<i>FAM69A</i>	1,12
2p23	<i>ALK</i>	1,37
5p13	<i>IL7R</i>	1,18
9q33	<i>DBC1</i>	1,17
10p15	<i>IL2RA</i>	1,25
16p13	<i>CLEC16A</i>	1,14

LFA3 : *lymphocyte function-associated antigen*; RPL5 : *ribosomal protein 5*; FAM69A : *family with sequence similarity 69A*; ALK : *anaplastic lymphoma receptor tyrosine kinase*; IL7R : *interleukin 7 receptor*; DBC1 : *deleted in bladder cancer 1*; IL2RA : *interleukin 2 receptor alpha*; CLEC16A : *C-type lectin domain family 16, A*.

degrés variables dans la maladie [49]. Les études fonctionnelles du gène *HLA-DRB1* ont été aussi réalisées mais les résultats obtenus ne sont pas faciles à interpréter. Selon les haplotypes observés, il existerait un allèle de forte susceptibilité à la SEP (*HLA DRB1*15*) et d'autres allèles de susceptibilité variable au même locus. La combinatoire de ces allèles en *trans* donnerait un gradient de susceptibilité à la maladie (de très susceptible à résistant). Les études sur des modèles murins sont en cours d'étude pour valider cette intéressante hypothèse [31]. Ces travaux sont difficiles à réaliser tant par le nombre de sujets nécessaires (plus de 10 000) que par les analyses statistiques à employer et par les technologies à appliquer (puces d'ADN par exemple pour l'analyse du transcriptome) ou les modèles animaux utilisés. Ces études (non terminées) commencent néanmoins à porter leur fruit. Les études actuelles suggèrent qu'entre 20 et 100 polymorphismes, chacun ayant un risque relatif compris entre 1,2 et 1,5 suffiraient pour expliquer la prévalence et l'héritabilité dans la SEP. Cependant, entre quelques centaines, voire milliers de variants rares pourraient individuellement expliquer la susceptibilité à la SEP avec des risques pouvant être multipliés par 10–20 [67]. Ainsi, certains gènes de susceptibilité ont été découverts (Tableau 3). Parmi ceux-ci (par exemple le gène codant pour *IL2RA*), certains d'entre eux sont impliqués dans d'autres pathologies inflammatoires : *IL2RA* est aussi impliqué dans le diabète de type I et la maladie de Basedow (hyperthyroïdie auto-immune). Les études se poursuivent donc tant pour identifier de manière exhaustive les SNPs sur le génome humain (plus de 500 000 SNPs peuvent être analysés en même temps à l'aide des technologies actuelles) que pour trouver tous les variants impliqués dans la pathologie ou des variants associés à une réponse thérapeutique individuelle (domaine de la pharmacogénétique) (Fig. 11).

Quelle est la place des tests génétiques en médecine ?

Nous évoquerons quelques pistes de base, ce sujet vaste étant impossible à traiter en quelques lignes. Actuelle-

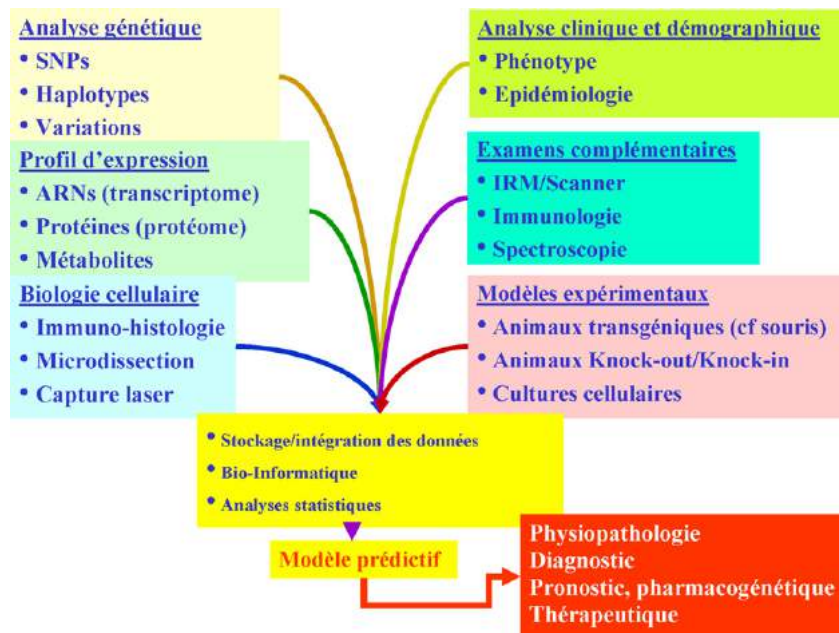


Figure 11 Principes des études entreprises pour l'étude de maladies polygéniques et multifactorielles. La complexité des maladies multifactorielles rend nécessaire l'approche dite intégrative. L'analyse des systèmes en biologie nécessite l'intégration de nombreuses données à traiter par les outils de la bio-informatique. En effet, on observe de nombreuses intégrations non linéaires tant au niveau moléculaire qu'au niveau cellulaire et organique. Dans la biologie des systèmes, l'analyse est effectuée sur un réseau de gènes et de protéines plutôt que sur chaque élément pris isolément. Les informations obtenues sont tant qualitatives que quantitatives. Ces données nécessitent bien évidemment une approche multidisciplinaire (cliniciens, biologistes, physiciens, mathématiciens notamment)

ment, hormis la médecine légale et la police scientifique, l'immense majorité des tests en génétique moléculaire sont prescrits pour des maladies mendéliennes (le diagnostic pré-implantatoire étant un cas à part). Des tests pour quelques maladies multifactorielles existent aussi (par exemple, dans les cancers du sein, les cancers du côlon). Certains d'entre eux sont prescrits dans le cadre d'un conseil génétique (notamment dans le cadre d'une enquête familiale ou chez un sujet asymptomatique, par exemple). Dans d'autres cas, elle peut être utilisée pour confirmer un diagnostic ou une maladie à présentation atypique. Dans le cadre d'une prescription (réglementée en France), le médecin doit pouvoir évaluer les tests dont le résultat permettra une meilleure prise en charge du patient. Ce n'est pas toujours simple. Par exemple, dans la prescription d'un test pour recherche d'une mutation sur le gène *HFE* impliqué dans l'hémochromatose (Omim 235200) et de transmission autosomique récessive, la présence de la mutation *C282Y* à l'état homozygote est responsable d'hémochromatose. Cependant, une minorité de patients seulement développeront la maladie (la pénétrance clinique est inférieure à 10%). Par ailleurs, un test génétique peut aussi être discutable dans les cas où malgré un résultat positif (présence d'une mutation), aucune mesure préventive ou prédictive ne peuvent être prise pour la pathologie étudiée. Dans ce cas, il n'existe pas de bénéfice pour le patient. Il est important aussi de savoir que dans un certain nombre de cas, le test génétique peut être négatif. En aucun cas, un test négatif élimine une maladie génétique. En effet, les tests génétiques explorent seule-

ment une partie de la région potentiellement mutée. Ces tests ne sont donc pas exhaustifs dans leur analyse génomique. Il est également important de savoir qu'un test génétique doit toujours être interprété en conjonction avec la clinique, la biologie et les examens complémentaires.

Quand penser à une maladie génétique ?

La présence de la génétique dans tous les domaines de la médecine est démontrée dans de nombreuses études en sciences fondamentales et appliquées. Qu'en est-il en pratique? **Quand penser à une maladie héréditaire en consultation?** En dehors de cas évidents de maladie héréditaire connue ou au tableau typique, il est important d'y penser dans certaines situations. Sans être exhaustif, voici quelques exemples :

- **une notion d'histoire familiale** : l'absence d'histoire familiale n'élimine cependant pas une maladie génétique et à l'inverse, la présence d'une histoire familiale ne signifie pas maladie héréditaire ;
- **l'apparition de la pathologie à un âge jeune** (enfant, adolescent ou jeune adulte), plus particulièrement pour une maladie connue pour apparaître habituellement à un âge plus tardif (par exemple, cancer du côlon) ;
- **l'absence d'étiologie évidente** pour une maladie (avec souvent des explorations multiples et négatives).

En pratique, que retenir et comment aborder la génétique avec les patients ?

Nous avons évoqué de manière simplifiée un certain nombre de connaissances actuelles en génétique humaine. Ces connaissances sont fondamentales pour comprendre les maladies (héréditaires ou non) en 2008. Comme cela a été brièvement exposé, la génétique humaine n'est pas simple même si des avancées majeures ont été et continuent de se réaliser. L'information du sujet au cours d'une consultation non spécialisée est difficile. Un médecin traitant doit cependant avoir des notions de bases importantes dans ce domaine (même s'il ne peut pas tout savoir) tout simplement parce que la génétique est présente dans tous les domaines de la médecine et joue un rôle majeur tant dans les origines de nombreuses maladies fréquentes (obésité, diabète, hypertension...) que dans l'adaptation de traitements de plus en plus spécialisés (par exemple, en cancérologie) et le pronostic de certaines maladies (par exemple, l'infection par le VIH). Il peut arriver par ailleurs qu'en consultation, un patient ou un sujet ayant dans sa famille un malade atteint d'une pathologie héréditaire aborde le thème des maladies génétiques. De nombreuses questions sont possibles, on peut citer par exemple : pouvez-vous m'expliquer cette maladie héréditaire ? Est-ce que ma maladie a un rapport avec une cause génétique ? Comment savoir si j'ai la même maladie héréditaire que mon oncle ? Est-ce que mon enfant risque d'être atteint ? Quand dois-je faire un test génétique ? Ma femme est enceinte et mon frère a la maladie. Que dois-je faire pour savoir si mon bébé risque d'être malade ? etc. Un rapport récent du service de santé britannique (publié en septembre 2007) a rapporté le ressenti de consultants (ayant un proche ou étant eux-mêmes atteints de maladie génétique) sur les informations en génétique transmises par leurs médecins non généticiens (rapport disponible en anglais sur le site <http://www.geneticseducation.nhs.uk>). Bien qu'imparfaite, cette étude nous semble intéressante pour illustrer certains points importants et inciter à la réflexion le lecteur sur ce sujet :

- la majorité des consultants préfèrent recevoir une information génétique de leur médecin habituel plutôt que d'un spécialiste. Bien que des spécialistes soient formés en génétique médicale, ils estiment que le médecin traitant doit participer à l'information, au suivi et au traitement des patients en coordination avec le spécialiste ;
- en parallèle de l'information orale, les consultants souhaitent un document écrit complétant celle-ci afin de le lire plus tard et faciliter ainsi l'assimilation des informations reçues ;
- le meilleur moment pour donner une information génétique (notamment un diagnostic) n'est pas évident. Les consultants reconnaissent que le médecin doit s'adapter à son interlocuteur et transmettre l'information sans émettre de jugement ou sans avis biaisé sur le sujet. Par ailleurs, la quantité d'information à transmettre doit être adaptée au consultant. Le stress psychologique important reçu par le sujet (et/ou par son entourage) doit être pris en compte. Le médecin doit savoir que le sujet (et/ou son entourage) peut ensuite ressentir un sentiment de culpa-

bilité ou subir des reproches de son entourage. Il est donc important d'en parler avec lui ;

- l'information doit être simple sans être simpliste. Expliquer des termes de génétique est souvent une tâche difficile. Il est nécessaire que le consultant comprenne les termes employés (par exemple, des termes comme « mutant » ou « risque » sont souvent mal compris voire pas du tout du sujet). Comme cela a été précisé ci-dessus, la remise d'une brochure écrite pour relire les informations orales est importante. Proposer au consultant de revenir pour poser d'autres questions s'il le souhaite est également essentiel ;
- les maladies génétiques sont mal connues des médecins non généticiens. Les consultants ont conscience que les médecins ne peuvent pas tout savoir et ne peuvent être des experts en maladies rares. Ils souhaitent cependant que les médecins évoquent une maladie génétique plus souvent et prennent en compte l'inquiétude des patients au sujet d'une possible transmission héréditaire d'une maladie. Ils souhaitent aussi que les médecins reconnaissent leur manque de compétence dans ce domaine et les adressent alors vers les spécialistes adéquats. Les médecins devraient savoir où les adresser si besoin. Le site Orphanet permet d'obtenir des informations précieuses (<http://www.orpha.net>) ;
- les patients ressentent fréquemment le manque de communication entre les différents spécialistes impliqués dans leur pathologie. Ils ont souvent peur des répercussions possibles sur la prise en charge de la maladie pour eux ou leurs proches ;
- les forums sur Internet, les associations de patients par exemple sont des sources d'aide importantes pour ces consultants. Le médecin généraliste devraient pouvoir les diriger vers ces groupes de support.

Cette étude montre le rôle important que doit jouer le médecin généraliste dans ce domaine même si des spécialistes en génétique sont formés spécifiquement au conseil génétique [52]. Le conseil génétique est une consultation à part réalisée par des spécialistes spécifiquement formés. Sans rentrer dans les détails du conseil génétique qui sort du cadre de cet article et qui fait l'objet de nombreuses discussions, nous donnerons quelques repères de base qui peuvent être utiles au non-spécialiste (pour plus de renseignements, voir le site <http://www.eurogentest.org/guidelines.xhtml>) :

Au cours du conseil génétique, plusieurs sujets sont abordés :

- options thérapeutiques, risques d'avoir la maladie (en terme de probabilité), but d'un test génétique, nature (par exemple, le diagnostic prénatal) et conséquences d'un test génétique, limites liées au test génétique (l'interprétation), conséquences d'un test positif, les risques pour la famille (possibilité d'une enquête familiale), l'existence de groupes de soutien (associations, forums sur Internet...);
- la nécessité de s'assurer de la compréhension du consultant sur le sujet évoqué ;
- le soutien psychologique est fondamental. La réponse du spécialiste aux réactions émotionnelles du consultant est très importante. Il est essentiel de se rappeler que le

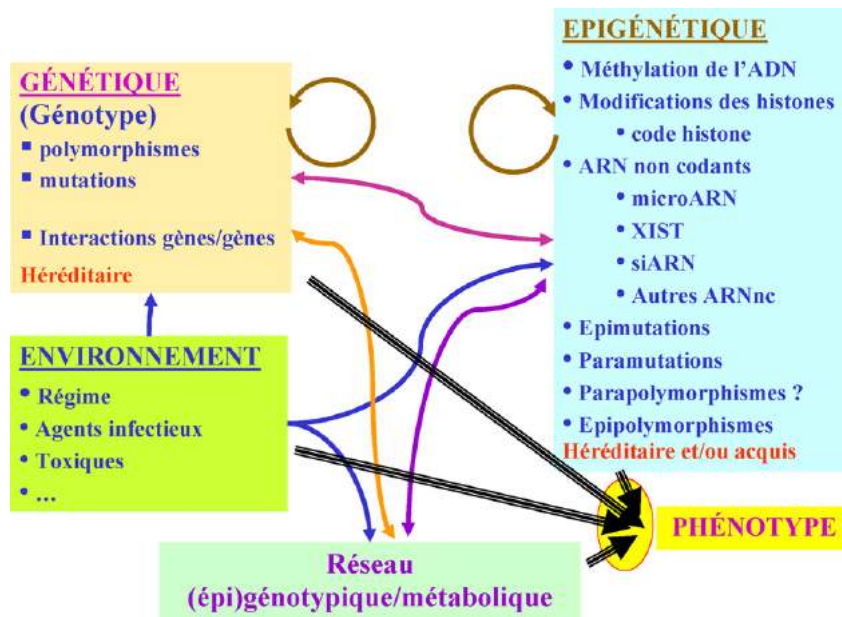


Figure 12 La relation genotype–phénotype. La relation genotype–phénotype est difficile à établir. La pénétrance et l'expressivité d'une maladie résultent d'une combinatoire associant de nombreuses interactions. Que la maladie soit monogénique, oligogénique ou multifactorielle, son phénotype est modulé par des facteurs tant génétiques, qu'épigénétiques ou environnementaux. L'importance de chacun de ces facteurs est variable d'une pathologie à l'autre et dépend souvent du réseau métabolique dont il fait partie. Ce schéma montre que l'interprétation d'un test génétique isolé ne suffit pas le plus souvent pour déterminer une maladie héréditaire ou à composante héréditaire.

stress émotionnel altère le plus souvent les perceptions du consultant ;

- les conséquences sociales et familiales sont importantes (retentissement sur la vie familiale, professionnelle, privée, possibilité d'une enquête familiale notamment). Les facteurs culturels (voire ethniques) sont à prendre en compte.

Bien que cela soit difficile, la neutralité est un objectif recommandé pour le conseiller vis-à-vis du consultant. Cependant, ce concept est critiqué par certains [52]. Le patient doit être aidé dans sa réflexion afin de l'encourager à prendre une décision en toute indépendance (l'éducation du public en général prend ici toute son importance). Bien évidemment, il existe des cas où le consultant est incapable ou peine à prendre une décision autonome. La non-directivité d'un conseiller étant une règle souvent acceptée, ce dernier cas rend le conseil génétique plus problématique. Certains spécialistes pensent que la décision doit être partagée entre le consultant et le spécialiste, ce dernier devant s'adapter au consultant qu'il trouve en face de lui [52].

Trois règles de base aident donc à diriger un entretien de conseil génétique, l'information, le support psychologique et les aspects éthiques.

En conclusion

La génétique avance à pas de géant. Les connaissances évoluent mais la place de la génétique dans la médecine quotidienne n'est pas encore claire. Bien que cette branche

de la médecine soit un domaine spécialisé, elle ne peut plus être mise de côté par les non-spécialistes, la génétique étant maintenant présente dans tous les domaines de la médecine [33]. Il serait donc important de modifier le programme de génétique dans les universités médicales pour diffuser les connaissances actuelles et leurs conséquences possibles dans un avenir proche. Il en est de même dans les revues médicales. L'enseignement scolaire devrait intégrer quelques notions de génétique médicale de base afin d'éduquer et sensibiliser la population générale sur ce sujet. De nombreux généticiens appellent à un développement majeur de l'enseignement en génétique et pas uniquement comme science fondamentale [33]. Pourquoi est-ce si important ? Tout n'est pas génétique et de nombreux facteurs notamment environnementaux jouent un rôle majeur dans les multiples pathologies humaines. Les études montrent cependant que les facteurs génétiques jouent un rôle certain (Fig. 12). Les connaître permettront donc d'améliorer la santé publique notamment la prévention, le diagnostic, le pronostic ou le traitement. Bien sûr, aujourd'hui, beaucoup de médecins non spécialistes ne voient pas l'intérêt de la génétique dans leur pratique quotidienne (hormis les cas des maladies héréditaires qu'en pratique, ces derniers voient rarement). Qu'ils ne se trompent pas, la génétique s'insinue progressivement dans la pratique. Et dans les années à venir, il leur sera nécessaire de pouvoir répondre aux questions des consultants sur ce vaste sujet.

Conflits d'intérêts

Aucun.

Références

- [1] Ameziane N, Bogard M, Lamoril J. Principes de biologie moléculaire en biologie clinique. Collection Campus Référence Elsevier, 2005.
- [2] Antonarakis SE, Beckmann JS. Mendelian disorders deserve more attention. *Nat Rev Genet* 2006;7:277–82.
- [3] Badanol JL, Katsanis N. Beyond Mendel: an evolving view of human genetic disease transmission. *Nat Genet Rev* 2002;3:779–89.
- [4] Beales PL. Lifting the lid on Pandora's box: the Bardet-Biedl syndrome. *Curr Opin Genet Dev* 2005;15:315–23.
- [5] Benfey PN, Mitchell-Olds T. From genotype to phenotype: systems biology meets natural variation. *Science* 2008;320:495–7.
- [6] Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. *Cell* 2007;128:669–81.
- [7] Bogard M, Ameziane N, Lamoril J. Microarray d'ADN et profils d'expression des gènes Première partie: concept, fabrication et mise en œuvre. *Immunoanal Biol Spéc* 2008;23:71–88.
- [8] Boyle MP. Strategies for identifying modifier genes in cystic fibrosis. *Proc Am Thorac Soc* 2007;4:52–7.
- [9] Brink RA. A genetic change associated with the R locus in maize which is directed and potentially reversible. *Genetics* 1956;41:872–9.
- [10] Calin GA, Croce CM. MicroRNAs signatures in human cancers. *Nat Rev Cancer* 2006;6:857–66.
- [11] Chandler VL. Paramutation: from maize to mice. *Cell* 2007;128:641–5.
- [12] Chen Y, Zhu J, Lum PY, Yang X, Pinto S, MacNeil DJ, et al. Variations in DNA elucidate molecular networks that cause disease. *Nature* 2008;452:429–35.
- [13] Cho JH. The genetics and immunopathogenesis of inflammatory bowel disease. *Nature* 2008;8:458–66.
- [14] Conrad DF, Hurler ME. The population genetics of structural variation. *Nat Genet* 2007;39:530–6.
- [15] Costa FF. Non-coding RNAs, epigenetics and complexity. *Gene* 2008;410:9–17.
- [16] Cuzin F, Grandjean V, Rassoulzadegan M. Inherited variation at the epigenetic level: paramutation from the plant to the mouse. *Curr Opin Genet Dev* 2008;18:193–6.
- [17] Daiger SP, Bowne SJ, Sullivan LS. Perspective on genes and mutations causing retinitis pigmentosa. *Arch Ophthalmol* 2007;125:151–8.
- [18] Downs JA, Nussenzweig MC, Nussenzweig A. Chromatin dynamics and the preservation of genetic information. *Nature* 2007;447:951–8.
- [19] Esteller M. Epigenetics in cancer. *N Engl J Med* 2008;358:1148–59.
- [20] Federman DD. The biology of human sex differences. *N Engl J Med* 2006;354:1507–14.
- [21] Feinberg AP, Volgestein B. Hypomethylation distinguishes genes of human cancers from their normal counterparts. *Nature* 1983;301:89–92.
- [22] Flaquer A, Rappold GA, Wienker TF, Fischer C. The human pseudoautosomal regions: a review for genetic epidemiologists. *Eur J Hum Genet* 2008;16:771–9.
- [23] Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, et al. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci USA* 2005;102:10604–9.
- [24] Gerstein MB, Bruce C, Rozowsky JS, Zheng D, Du J, Korbelt JO, et al. What is a gene, post-ENCODE? History and updated definition. *Genome Res* 2007;17:669–81.
- [25] Giardiello FM, Petersen GM, Piantadosi S, Gruber SB, Traboulsi EI, Offerhaus GJA, et al. APC gene mutations and extraintestinal phenotype of familial adenomatous polyposis. *Gut* 1997;40:521–5.
- [26] Gibson G. The environmental contribution to gene expression profiles. *Nat Rev Genet* 2008;9:575–81.
- [27] Gimelbrant A, Hutchinson JN, Thompson BR. Chess A Widespread monoallelic expression on human autosomes. *Science* 2007;318:1136–40.
- [28] Gingeras TR. Origin of phenotypes: genes and transcripts. *Genome Res* 2007;17:682–90.
- [29] Gouya L, Martin-Schmitt C, Robreau AM, Austerlitz F, Da Silva V, Brun P, et al. Contribution of a common single-nucleotide polymorphism to the genetic predisposition for erythropoietic protoporphyria. *Am J Hum Genet* 2006;78:2–14.
- [30] Grant SF, Hakonarson H. Microarray technology and applications in the arena of genome-wide association. *Clin Chem* 2008;54:1116–24.
- [31] Gregersen JW, Kranc KR, Ke X, Svendsen P, Madsen LS, Thomsen AR, et al. Functional epistasis on a common MHC haplotype associated with multiple sclerosis. *Nature* 2006;443:574–7.
- [32] Gropman AL, Adams DR. Atypical patterns of inheritance. *Semin Pediatr Neurol* 2007;14:34–45.
- [33] Guttmacher AE, Porteous ME, McInerney JD. Educating health-care professionals about genetics and genomics. *Nat Rev Genet* 2007;8:151–7.
- [34] Haines JL, Terwedow HA, Burgess K, Pericak-Vance MA, Rimmeler JB, Martin ER, et al. Linkage of the MHC to familial multiple sclerosis suggests genetic heterogeneity. The Multiple Sclerosis Genetics Group. *Hum Mol Genet* 1998;7:1229–34.
- [35] Holliday R. The inheritance of epigenetic defects. *Science* 1987;238:163–70.
- [36] Hupperts R, Broadley S, Mander A, Clayton D, Compston DA, Robertson NP. Patterns of disease in concordant parent-child pairs with multiple sclerosis. *Neurology* 2001;57:290–5.
- [37] Hurler ME, Dermitzakis ET, Tyler-Smith C. The functional impact of structural variation in humans. *Trends Genet* 2008;24:238–45.
- [38] Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, et al. RIKEN Genome Exploration Research Group; Genome Science Group (Genome Network Project Core Group); FANTOM Consortium. Antisense transcription in the mammalian transcriptome. *Science* 2005;309:1564–6.
- [39] Katsanis N, Ansley SJ, Badano JL, Eichers ER, Lewis RA, Hoskins BE, et al. Triallelic inheritance in Bardet-Biedl syndrome, a Mendelian recessive disorder. *Science* 2001;293:2256–9.
- [40] Knudson Jr AG. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci USA* 1971;68:820–3.
- [41] Li Y, Alvarez OA, Gutteling EW, Tijsterman M, Fu J, Riksen JA, et al. Mapping determinants of gene expression plasticity by genetical genomics in *C. elegans*. *PLoS Genet* 2006;2:e222.
- [42] Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* 2003;33:177–82.
- [43] Lyon MF. Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Naturwissenschaften* 1961;190:372–3.
- [44] MacCarroll SA, Altshuler DM. Copy-number variation and association studies of human disease. *Nat Genet* 2007;39:537–42.
- [45] Martin C, Zhang Y. Mechanisms of epigenetic inheritance. *Curr Opin Cell Biol* 2007;19:266–72.
- [46] Miller D, Ostermeier GC, Krawetz SA. The controversy, potential and roles of spermatozoal RNA. *Trends Mol Med* 2005;11:156–63.
- [47] Moore JH. A global view of epistasis. *Nat Genet* 2005;37:13–4.
- [48] Moore JH. The ubiquitous nature of epistasis in determining susceptibility to common human diseases. *Hum Hered* 2003;56:73–82.
- [49] Oksenberg JR, Baranzini SE, Sawcer S, Hauser SL. The genetics of multiple sclerosis: SNPs to pathways to pathogenesis. *Nat Rev Genet* 2008;9:516–26.

- [50] Pache RA, Zanzoni A, Naval J, Mas JM, Aloy P. Towards a molecular characterisation of pathological pathways. *FEBS Lett* 2008;582:1259–65.
- [51] Pertersen MB. Non-syndromic, autosomal-recessive deafness. *Clin Genet* 2006;69:371–91.
- [52] Rantanen E, Hietala M, Kristoffersson U, Nippert I, Schmidtke J, Sequeiros J, et al. What is ideal genetic counselling? A survey of current international guidelines. *Eur J Hum Genet* 2008;16:445–52.
- [53] Rassoulzadegan M, Grandjean V, Gounon P, Vincent S, Gillot I, Cuzin F. RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature* 2006;441:469–74.
- [54] Richards EJ. Population epigenetics. *Curr Opin Genet Dev* 2008;18:221–6.
- [55] Santos-Rebouças CB, Pimentel MMG. Implication of abnormal epigenetic patterns for human diseases. *Eur J Hum Genet* 2007;15:10–7.
- [56] Sasltrom JL. X-inactivation and the dynamic maintenance of gene silencing. *Mol Genet Metab* 2007;92:56–62.
- [57] Serafini B, Rosicarelli B, Franciotta D, Magliozzi R, Reynolds R, Cinque P, et al. Dysregulated Epstein-Barr virus infection in the multiple sclerosis brain. *J Exp Med* 2007;204:2899–912.
- [58] Shelling AN, Ferguson LR. Genetic variation in human disease and a new role for copy number variants. *Mut Res* 2007;622:33–41.
- [59] Taft RJ, Pheasant M, Mattick JS. The relationship between non-protein coding DNA and eukaryotic complexity. *Bioessays* 2007;29:288–99.
- [60] Torres VE, Harris PC, Pirson Y. Autosomal dominant polycystic kidney disease. *Lancet* 2007;369:1287–301.
- [61] Waddington CH. Preliminary notes on the development of the wings in normal strains and mutant strains of *Drosophila*. *Proc Natl Acad Sci USA* 1939;25:299–307.
- [62] Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 2008;320:539–43.
- [63] Willer CJ, Dyment DA, Risch NJ, Sadovnick AD, Ebers GC, Canadian Collaborative Study Group. Twin concordance and sibling recurrence rates in multiple sclerosis. *Proc Natl Acad Sci USA* 2003;100:12877–82.
- [64] Wood AJ, Oakey RJ. Genomic imprinting in mammals: emerging themes and established theories. *Plos Genet* 2006;2:e147.
- [65] Wutz A, Gribnau J. X inactivation Xplained. *Curr Opin Genet Dev* 2007;17:387–93.
- [66] Yang PK, Kuroda MI. Noncoding RNAs and intranuclear positioning in monoallelic gene expression. *Cell* 2007;128:777–86.
- [67] Yang Q, Khoury MJ, Friedman J, Little J, Flanders WD. How many genes underlie the occurrence of common complex diseases in the population? *Int J Epidemiol* 2005;34:1129–37.
- [68] Yano Y, Saito R, Yoshida N, Yoshiki A, Wynshaw-Boris A, Tomita M, et al. A new role for expressed pseudogenes as ncRNA: regulation of mRNA stability of its homologous coding gene. *J Mol Med* 2004;82:414–22.
- [69] Zhao Y, He S, Liu C, Ru S, Zhao H, Yang Z, et al. MicroRNA regulation of messenger-like noncoding RNAs: a network of mutual microRNA control. *Trends Genet* 2008;24:323–7.

Pour en savoir plus

Site d'information sur les maladies rares : <http://www.orpha.net>

Site d'information sur les maladies mendéliennes OMIM (en anglais) : <http://www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM>

Site tout public (en français) sur la génétique, Planète Gène : <http://www.planetegene.com>

Site tout public (en anglais) sur la génétique, Genetic home reference : <http://ghr.nlm.nih.gov/>

Glossaire

Analyse haplotypique: Étude de la transmission d'une combinaison d'allèles situés sur un chromosome (issu d'un des deux parents) et en différents sites de celui-ci. L'association sur ce chromosome (du même parent) des différents allèles définit un haplotype

Anticipation: Propriété observée lorsqu'une maladie devient de plus en plus grave de génération en génération

Chromatine: Complexe formé d'ADN et de protéines qui composent les chromosomes. La chromatine englobe l'ADN dans un volume lui permettant d'être contenu dans noyau de la cellule, de réaliser la mitose et la meiose et de contrôler l'expression des gènes. La conformation de la chromatine est modifiée par la méthylation de l'ADN et les modifications sur les histones

Consanguin: Descendant issu d'ancêtres communs (de la même famille)

Conseil génétique: Consultation au cours de laquelle un spécialiste informe un consultant sur une maladie héréditaire notamment sur le diagnostic, l'évolution possible, le pronostic, le traitement quand il existe, la prise en charge, la transmission notamment dans la famille et pour les enfants à venir, le risque de développer celle-ci. Le spécialiste présente et interprète les résultats au consultant et l'aide à comprendre leur signification et les conséquences possibles notamment la conduite à tenir la plus appropriée pour le consultant et éventuellement pour son entourage et/ou sa famille en relation avec la maladie génétique en cause

Déséquilibre de liaison: Association non aléatoire entre allèles situés sur des loci différents

Disomie uniparentale: État dans lequel des allèles homologues à un locus donné sont issus du chromosome du même parent

Empreinte parentale (ou génomique): Expression d'un gène différentielle selon qu'il est d'origine paternelle ou maternelle

Épigénome: État épigénétique global d'une cellule

Épimutation: Méthylation anormale acquise ou héréditaire sur l'AND

Épistasie: De manière générale, interaction entre les gènes. Pour certains, il s'agit de l'action d'un gène masquant l'effet d'un autre gène

Expressivité: Degré d'expression phénotypique d'un gène

Gène: Segment d'ADN contribuant au phénotype

Haplotype: Combinaison d'allèles sur un même chromosome transmis ensemble

Héritabilité: Proportion d'un état attribué à un ou plusieurs facteurs génétiques. Statistiquement, elle se traduit par la proportion de la variation phénotypique globale d'une population qui peut être attribuée à la variance génétique des individus

Hétérogénéité allélique: Situation dans laquelle plus d'un allèle d'un gène sont associés à une pathologie. À différencier de l'hétérogénéité génétique ou d'un locus dans laquelle la variation dans plusieurs gènes peut être responsable de formes identiques ou similaires d'une maladie chez différents sujets

Hétérozygote composite: Présence au même locus de deux allèles sur chaque chromosome et dont l'effet combiné est similaire à un allèle muté homozygote (se voit dans les maladies récessives)

Hétérogénéité génétique: Condition dans laquelle la variation génétique dans plusieurs gènes peut être responsable de formes identiques ou similaires d'une maladie chez différents sujets. Voir aussi Hétérogénéité allélique

Histone: Principal composant protéique de la chromatine. Les histones H2A, H2B, H3 et H4 s'assemblent pour former le nucléo-

some. Chaque nucléosome s'enroule autour de 146 pb de l'ADN. L'histone H1 constitue un histone de liaison qui maintient l'ADN en place et permet une structure compacte plus élaborée

Ilôts CpG: Régions de l'ADN contenant de nombreux cytosines et guanines adjacentes. Le p dans CpG correspond à la liaison phosphodiester entre la cytosine et la guanine. On estime qu'environ 40% des promoteurs de gènes humains possèdent de tels îlots

Interaction épistatique: Interaction de plusieurs loci modulant la variation phénotypique

Jumeaux monozygotes discordants: Jumeaux partageant le même patrimoine génétique mais qui ne sont pas affectés par la même pathologie (un étant malade et l'autre non)

Maladie mendélienne: Maladie se transmettant de génération en génération par l'intermédiaire d'un ou plusieurs gènes

Maladie monogénique: maladie conséquence d'une anomalie sur un seul gène

Méthylation de l'ADN: Addition d'un groupement méthyl sur le carbone 5 du noyau pyrimidine dans le nucléotide cytosine qui précède une guanine

Multifactoriel: Combinaison de traits génétiques et non génétiques, chacun contribuant au phénotype global

Mutation dynamique: Mutation sur des séquences répétées qui se caractérise par le changement du nombre de répétitions soit sous forme d'expansion soit sous forme de contraction à la génération suivante

Open reading frame (ORF): En français, cadre ouvert de lecture, c'est-à-dire séquence nucléotidique dont la lecture (3 bases par 3 bases ou triplets appelés aussi codons) permet de déterminer la séquence des acides aminés qui constituent le peptide et aboutira après maturation à la protéine codée par le gène correspondant. L'ORF débute par un codon initiateur ATG (codon

codant aussi pour l'acide aminé méthionine) et finit un codon stop non codant (TGA, TAA ou TAG)

Paramutation: Modification épigénétique d'expression d'un gène « paramutable » lorsqu'il se transmet (co-ségrège) avec un parent hétérozygote possédant un allèle paramutagène. Le phénotype résulte de la modification quantitative de l'expression du gène paramuté

Pénétrance: Probabilité qu'un génotype se manifeste par un phénotype. Quand la pénétrance est inférieure à 100%, on parle de pénétrance incomplète

Polymorphisme du nombre de copies (CNV, copy-number variant): Variant au niveau d'un locus chromosomique se manifestant par une variation du nombre de copies de la région génomique en question

Pseudo-autosomique: Région chromosomique homologue entre le chromosome X et le chromosome Y

Relation génotype-phénotype: Relation entre la sévérité de l'expression bioclinique d'une maladie (phénotype) et des variants génétiques spécifiques (génotype)

Single nucleotide polymorphisms (SNP) ou polymorphisme bi-allélique: Polymorphisme bi-allélique par substitution d'une base par une autre. C'est le plus fréquent des polymorphismes

Transcriptome: Ensemble des ARNm exprimés dans une cellule ou un tissu à un moment donné. Il constitue un instantané « une photo » de l'expression des gènes codants à un moment donné dans un état physiologique ou pathologique particulier. Parmi les technologies employées pour l'étudier, on retrouve les puces d'expression

Transposon: Séquence mobile d'ADN qui peut se transporter dans un autre endroit du génome d'une cellule. On parle alors de transposition. Un transposon peut provoquer une mutation ou une modification de l'organisation du génome