

## Introduction

Les applications de la génétique envahissent notre quotidien. Qu'il s'agisse de diagnostic prénatal, d'agronomie ou de police scientifique, nous apprécions la précision des méthodes génétiques en même temps que nous redoutons leur puissance. Ce que l'on est capable d'accomplir aujourd'hui était inimaginable il y a peu de temps. Nous entrons dans l'ère de la médecine personnalisée et de la thérapeutique génétique en même temps que nous quittons l'empirisme qui a prévalu jusqu'à présent pour la description et l'exploitation de la biosphère. Et les progrès s'accroissent. Mais de quoi s'agit-il ? Comment comprendre ce qui se passe si l'on n'a pas des notions claires sur les principes fondamentaux du monde vivant ? Principes qui ne se sont révélés que lentement aux scientifiques, dont les investigations ont souvent suivi des trajectoires complexes, peu explicites aux non-spécialistes.

Lorsque l'on parle d'hérédité, le langage courant donne l'apparence de la simplicité. On entend dire, par exemple, que cet enfant aux yeux clairs a hérité des *gènes\** de sa grand-mère. Qui a aussi les yeux clairs, c'est sous-entendu, tout le monde le sait dans la famille. Cette simple phrase cache toute la complexité du gène. D'un côté, le gène y est bien perçu comme cet élément – un peu mystérieux – transmis de génération en génération. Et pouvant même sauter les générations, car c'est la grand-mère dont il est question. De l'autre, le gène est ici confondu avec le caractère observable qu'on désigne par *phénotype\**. Ici, les yeux sont clairs, mais, dans la population, il existe une variété de teintes, pas seulement des yeux clairs opposés aux yeux foncés. Y aurait-il un gène par teinte ? Cela semble bien difficile à imaginer. Et puis la génétique nous apprend aussi que cet enfant aux yeux clairs a hérité d'exactly autant de gènes de sa grand-mère aux yeux clairs que de chacun de ses autres grands-parents qui, eux, avaient les yeux foncés. Alors, qu'est-ce qu'un gène ? Comment ça marche ?

On a longtemps cru que c'était le sang qui véhiculait les caractères héréditaires, d'où l'adjectif « consanguin » utilisé habituellement pour désigner les filiations. On sait maintenant qu'il s'agit de l'*ADN*\* (pour acide désoxyribonucléique), une macromolécule présente dans toutes nos cellules, c'est-à-dire qu'il en existe plusieurs dizaines de milliers de milliards de copies chez un être humain adulte, dérivant toutes des molécules uniques présentes dans l'œuf qui lui a donné naissance. Avec les progrès de la génétique, le sigle ADN s'est banalisé. C'est devenu un terme courant pour désigner l'essence d'une personnalité ou même d'une chose. On entend parler de l'ADN d'une entreprise ou de l'ADN d'un club sportif. Ce qui n'a littéralement aucun sens. L'ADN n'existe que dans les organismes vivants. Et même pour des personnalités, invoquer l'ADN d'un artiste pour signifier son talent ou l'ADN d'un enfant passionné d'équitation pour justifier son goût des chevaux n'apporte rien de plus à la compréhension des causes que de parler de sang. Sauf l'apparence d'être instruit. Mal instruit. Car le passage du sang à l'ADN s'est accompagné d'une révolution conceptuelle considérable totalement ignorée ici : le caractère particulière des déterminants des *traits*\* héréditaires. La génétique est née à partir du moment où la vision ancienne de mélange de « fluides parentaux » qui était invoquée pour expliquer l'hérédité des caractères de la descendance a été remplacée par la notion de « grains » unitaires d'hérédité. Alors que ces fluides imprécis donnaient l'illusion de continuité, chaque grain, qui recevra plus tard le nom de gène, devenait une entité propre indépendante des autres grains avec lesquels il se trouvait mélangé en proportions discrètes, quantifiables par l'expérimentateur. C'est ce qu'avait compris Gregor Mendel dès 1865. Mais alors quelle est la relation entre gène et ADN ?

### Les multiples facettes du gène

Il ne faut pas s'attendre à une réponse simple. Le gène est un concept. Mais un concept manipulable dans un tube à essai ! On trouvera dans les chapitres suivants l'évolution des idées à mesure que la nature, l'organisation et le fonctionnement du matériel génétique se clarifiaient. Qu'en est-il aujourd'hui ? Le gène reste contre-intuitif, car il oppose trois difficultés principales à sa compréhension : l'une opérationnelle, l'autre temporelle et la troisième essentielle.

La première difficulté tient au fait que le gène est à la fois un segment d'ADN et une entité fonctionnelle. En d'autres termes, sa nature même change selon la façon dont on l'examine. Un peu comme des particules élémentaires en physique quantique peuvent être à la fois grain de matière et onde. Or, si l'ADN est une molécule parfaitement définie au niveau chimique (voir le chapitre 2), que l'on peut manipuler *in vitro*, la notion de fonction reste une abstraction propre à la biologie qui s'applique à différentes échelles depuis la molécule – une *enzyme*\* par exemple

– jusqu’à l’organisme entier – la couleur des yeux par exemple. Et si l’ADN est sécable en ses composants élémentaires, les fonctions, elles, ne le sont pas. Ce qui fait du gène une unité primordiale, un « grain », mais composée d’éléments séparables les uns des autres.

La deuxième difficulté tient au fait que le gène est à la fois l’élément qui traverse les générations et celui qui agit à chacune d’elles. À chaque génération, c’est l’aspect fonctionnel qui prévaut. Entre les générations, c’est la qualité de la transmission du contenu informationnel qui importe. Or cette dualité repose sur deux types de molécules et deux processus différents. La transmission du contenu informationnel repose sur l’ADN et son dédoublement (on parle de *réplication\**) avant chaque division cellulaire. L’aspect fonctionnel repose sur la copie de l’ADN (on parle de *transcription\**) sous forme de molécules d’*ARN\**, pour acide ribonucléique, l’autre catégorie d’acides nucléiques (voir le chapitre 2).

Ceci nous conduit à la troisième difficulté, car la permanence intergénérationnelle est immatérielle – c’est une information –, mais elle est portée par des molécules et contenue dans l’ordre de leurs éléments. Et il est important de bien comprendre qu’au niveau matériel ce n’est pas le gène lui-même qui est transmis de génération en génération, mais ses copies avec, inévitablement, leur potentiel d’erreur, c’est-à-dire de *mutation\**. Le gène, élément de permanence, devient donc automatiquement élément de variation. C’est ainsi qu’au cours des générations successives d’une lignée, et donc au sein des populations d’une espèce, un même<sup>1</sup> gène acquerra de multiples formes, que l’on nomme *allèles\**. Cette notion très importante fait de la génétique « la science qui utilise la variation pour étudier la permanence », comme le disait fort bien Philippe L’Héritier, l’un des pionniers de cette discipline en France.

Pour résumer, on pourrait dire que le gène est une information nécessaire à la réalisation d’une fonction biologique, inscrite dans un *acide nucléique\** et transmise de génération en génération avec un certain degré d’imperfection. Mais d’où proviennent alors les traits phénotypiques, si l’enfant aux yeux clairs a hérité d’autant de gènes de chacun de ses grands-parents, quelle que soit la couleur de leurs yeux ?

## Génotype et phénotype : la réalité du déterminisme génétique

On entre ici dans des phénomènes qui ne furent compris que lentement à travers un siècle et demi de recherches. Éliminons d’emblée le vocabulaire simpliste utilisé sur

---

1. On n’entrera pas ici dans le problème philosophique insoluble qui consiste à savoir à quel point un ensemble reste le même lorsque ses parties changent.

Internet ou par certains médias qui parlent de gène à propos de n'importe quoi dès que les humains sont concernés. On y trouve pêle-mêle le gène de la violence, de la dépression, de la bougeotte, de la gloutonnerie, de la fesse généreuse, de la lâcheté, de la résilience, de la révolte, de la minceur, de l'obésité, de la criminalité, de la schizophrénie, de l'homosexualité, de l'intelligence mathématique... Curieusement autant de caractéristiques au féminin. Encore plus curieusement, on n'y trouve pas le gène de l'imbécillité qui caractérise pourtant cette liste. Car le déterminisme génétique est tout sauf simple et direct, comme le suggère ce vocabulaire trompeur.

D'abord, parce que chaque gène pouvant prendre de multiples formes, ce sont ses allèles qu'il faut considérer. Ensuite, parce que beaucoup d'organismes vivants – dont l'homme – étant *diploïdes\**, c'est la relation entre deux allèles qui compte : lorsque l'effet de l'un domine celui de l'autre, on parlera de *dominance\** du premier et de *récessivité\** du second, mais il est aussi possible que les effets de l'un et de l'autre s'ajoutent en proportions variables. Enfin, parce que les fonctions biologiques étant fortement intriquées, c'est l'interaction entre allèles de différents et parfois nombreux gènes qui entre en jeu. À titre d'exemple, plus de 400 gènes différents interviennent pour déterminer la taille des adultes chez l'homme, chaque gène étant représenté au sein de la population par de nombreux allèles associés par paires dans chaque individu. Rien d'étonnant, donc, à ce que la variation de taille des individus forme un continuum apparent. Est-ce toujours aussi complexe ? Non. La mucoviscidose est causée par le dysfonctionnement d'un transporteur d'ions chlorure de l'épithélium pulmonaire qui est le produit d'un gène précisément défini, localisé sur notre *chromosome\* 7*. On connaît maintenant environ 2 000 allèles de ce gène qui ont pu être reliés à la sévérité du syndrome. Dans ce cas, les allèles de cet unique gène expliquent la quasi-totalité des phénotypes observés. Il en va de même pour un phénotype pourtant complexe, comme la transformation des antennes en pattes chez la mouche *Drosophila*, suite à la mutation d'un seul gène.

En réalité, la grande majorité des traits phénotypiques dépendent simultanément de gènes à effets majeurs – les plus faciles à identifier – et de gènes à effets mineurs dont les listes ne sont généralement pas complètes. La puissance de l'analyse génétique dépend énormément des organismes auxquels elle s'applique, de leur degré de fécondité et du polymorphisme génétique au sein des populations naturelles. Dans beaucoup de cas chez l'homme, les mutations de l'ensemble des gènes déjà identifiés n'expliquent pas l'intégralité des individus porteurs d'un trait phénotypique particulier. On dit qu'il y a un manque de *héritabilité\** de ce trait. Comblé ce manque est un objectif important de la génétique humaine actuelle.

Avant d'aller plus loin, il est nécessaire de revenir sur les traits phénotypiques. Une plante placée sur un sol pauvre et aride poussera généralement moins bien que

sa jumelle génétique (sa bouture par exemple) placée sur un sol riche et convenablement irrigué. Comme un animal mal nourri aura des difficultés de croissance. Ces banales constatations sont importantes au plan pratique, mais sans intérêt pour ce qui nous concerne ici. La comparaison des phénotypes entre individus n'a de sens génétique que lorsque ces individus se trouvent placés dans des conditions extérieures comparables. Mais, même dans ces conditions, le déterminisme génétique du phénotype n'est pas direct. En réalité, des individus génétiquement identiques (les membres d'une *lignée pure\** ou d'un *clone\** ou simplement de vrais jumeaux) placés dans des conditions identiques ne seront jamais strictement identiques au plan phénotypique. Car il existe toujours, pour chaque trait phénotypique considéré, une variation individuelle autour d'une valeur moyenne. Au sein d'une population génétiquement homogène, cette variation génère une distribution statistique particulière à ce trait. Selon le trait considéré, la variance de la distribution sera plus ou moins grande. Il est donc essentiel de bien comprendre que ce qui est transmissible à la descendance, c'est la forme et les valeurs paramétriques (moyenne, variance) de ces distributions, pas le phénotype précis de chaque individu. C'est cette observation fondamentale qui avait amené Wilhelm Johannsen à définir *génotype\** et *phénotype\** dès 1903 (voir le chapitre 1). Lorsque l'on parle de déterminisme génétique, il faut donc conserver à l'esprit que le génotype, élément de permanence intergénérationnelle, ne définit pas directement le phénotype de chaque individu, mais définit la distribution statistique des phénotypes de tous les individus qui portent ce même génotype, placés dans les mêmes conditions extérieures.

Aujourd'hui, on sait définir entièrement le génotype d'un individu par le séquençage intégral de son *génom\**. Pour un diploïde, les deux allèles de chaque gène doivent évidemment être séquencés individuellement, ce qui n'est devenu possible que très récemment. Sauf pour des cas particuliers de diagnostics prénataux par exemple, il reste difficile de prédire le phénotype qui en découlera, car on est loin de connaître toutes les interactions qui peuvent exister entre les allèles de tous les gènes. Le diagnostic génétique reste donc probabiliste.

À cela s'ajoute une dimension temporelle, car l'état fonctionnel d'un gène dépend non seulement de sa forme allélique (et éventuellement d'autres gènes du même génome), mais également de son état fonctionnel au cours de la génération précédente. C'est ce que l'on appelle hérédité *épigénétique\**. Les mécanismes moléculaires qui en sont responsables seront abordés dans les chapitres suivants. À ce stade, il suffit de rappeler que les phénomènes épigénétiques sont eux-mêmes génétiquement déterminés et, pour certains d'entre eux, non seulement par le génotype de l'individu concerné, mais aussi par ceux des générations précédentes.

## Les produits des gènes

On imaginerait naïvement que les gènes ont une grande variété de produits, vu le nombre et la diversité des traits phénotypiques observables. C'est exactement l'inverse. Il n'y a qu'un seul type de produit direct des gènes : l'ARN. Les molécules d'ARN naissantes copient l'information de l'ADN dont sont constitués les gènes, mais, selon leurs *séquences*\* nucléotidiques, elles s'engagent ensuite dans des voies fonctionnelles très différentes. Ce sont donc en réalité les ARN et non les gènes eux-mêmes qui sont à l'origine de toutes les fonctions des cellules vivantes. Mais ces molécules d'ARN ont une durée de vie limitée et, sauf exception, ne sont pas transmises à la génération suivante, car elles ne se répliquent pas. Elles n'assurent donc pas la permanence intergénérationnelle. Ce n'est pas le cas des molécules d'ARN de certains *virus*\* qui se répliquent, envahissent les cellules et peuvent se transmettre à la descendance.

Parmi les voies fonctionnelles empruntées par ces molécules d'ARN, l'une aboutit à la synthèse d'une autre catégorie de macromolécules faites d'enchaînements d'*acides aminés*\* dans des ordres précis : les *protéines*\*. La synthèse des protéines est un processus chimique complexe qui implique plusieurs types de molécules d'ARN agissant ensemble de manière coordonnée. Au cours de ce processus, l'information faite de la succession des *nucléotides*\* dans un ARN, la séquence de l'ARN dit « messenger », est traduite en une information faite d'une succession d'acides aminés dans une protéine, c'est-à-dire une autre séquence.

Comme les séquences de nucléotides offrent une combinatoire infinie, la diversité théorique possible des protéines, dont dépendent *in fine* les traits phénotypiques, l'est également. On comprend ainsi comment, par un seul mécanisme, les gènes peuvent être à l'origine de tous les traits phénotypiques. Mais comment se réalise ce passage d'information de l'ARN aux protéines que l'on appelle processus de *traduction*\* ?

C'est ici qu'intervient le *code génétique*\*. Ce terme est souvent utilisé de façon inappropriée dans le public, conduisant à une incompréhension totale de la génétique. Il ne s'agit pas du contenu informationnel du génome d'un organisme vivant qui lui est propre et qui diffère selon les espèces et même entre les membres d'une même espèce. Il s'agit d'un code de déchiffrement, commun à tous les organismes vivants actuellement connus. Le code génétique, comme tous les codes de déchiffrement, est un ensemble de règles simples qui établissent la correspondance entre messages constitués d'éléments de natures différentes. Un peu à la manière dont un message fait d'une suite de points et de traits peut être déchiffré en écriture latine grâce au code Morse. Ici, le code génétique établit la correspondance entre une succession de nucléotides dans un acide nucléique, une molécule d'ARN

messenger, et une succession d'acides aminés dans une protéine qui en est sa traduction. Le code génétique ne décide en rien de la nature de la protéine synthétisée. Celle-ci ne dépend que de l'information portée par l'ARN messenger qui elle-même provient, par un chemin parfois complexe, de l'information portée par le gène. Dire, comme on l'entend trop souvent, qu'une technique de réécriture des génomes modifie le code génétique d'une cellule est donc un contresens complet, qui revient à dire que l'on a changé le code Morse. Heureusement qu'il n'en est rien, car, comme toute communication télégraphique deviendrait impossible, la cellule deviendrait immédiatement incapable de déchiffrer son propre génome et disparaîtrait !

Pour des raisons que l'on expliquera dans cet ouvrage, le code génétique établit une correspondance entre des séquences de trois nucléotides (il y en a  $4^3 = 64$ ) et les vingt acides aminés principaux qui entrent dans la composition des protéines. On notera que l'universalité<sup>2</sup> du code génétique dans le monde vivant connu démontre l'origine ancestrale commune du mécanisme de synthèse des protéines utilisé par les êtres vivants actuels. Il s'agit d'un mécanisme pratiquement figé en termes d'évolution, ce qui est extrêmement rare dans le monde vivant. Cette universalité permet le transfert de gènes entre organismes contemporains différents. Un phénomène dénommé *transgénése\**, qui, bien avant son utilisation artificielle, a joué un rôle critique dans l'évolution biologique naturelle : en son absence nous ne serions pas là pour en parler.

Mais alors, si tous les organismes vivants utilisent le même code génétique et peuvent même parfois échanger des gènes, peut-on parler de gènes humains, de gènes de chats, de sauterelles, de blé, de perroquet ou de poisson rouge ? Et qu'est-ce qu'une espèce ?

## Unité et diversité du monde vivant

Historiquement, la notion d'espèces vivantes a précédé celle d'évolution biologique. À l'origine, les espèces étaient définies comme des ensembles d'individus présentant des caractères communs suffisamment distinctifs de ceux d'autres ensembles qui constituaient donc d'autres espèces. Si ces caractères se modifiaient avec le temps, alors de nouvelles espèces pouvaient apparaître, d'où l'idée de transformisme et de filiation entre espèces. Pour les organismes à reproduction sexuée, c'est la viabilité et la fertilité effective des hybrides qui indiquaient que les deux parents étaient membres d'une même espèce ou, dans le cas contraire, de deux espèces différentes, comme s'il existait des barrières héréditaires entre espèces. Comment réconcilier ces

---

2. Il s'agit en réalité d'une quasi-universalité, car des changements mineurs sont observés dans certaines espèces.

observations ? La génétique est née dans ce contexte historique et a très rapidement cherché à rationaliser les bases de la spéciation par l'idée d'incompatibilité entre allèles de différents gènes, donc en faisant l'hypothèse que les mêmes gènes devaient être présents sous différentes formes alléliques dans les différentes espèces. Dès lors, il n'existe plus de gènes humains, de gènes de chats, de sauterelles, de blé, etc., mais seulement des formes différentes des mêmes gènes après divergence de leurs séquences suite aux mutations qui se sont accumulées au cours des générations successives depuis leurs ancêtres communs. Évidemment, plus le nombre de générations augmente, plus les divergences rendent difficile la reconnaissance de leur ancestralité commune. Mais en deçà de cette limite, on pourra reconstituer leurs liens probables de parenté en comparant leurs séquences. L'idée de tenter de reconstituer les liens généalogiques entre espèces par des arbres n'est pas nouvelle. On doit en effet une première illustration concrète de cette idée à Antoine Nicolas Duchesne dans son *Histoire naturelle des fraisiers* en 1766, à propos de ces espèces qu'il avait étudiées et croisées. Cette idée de généalogie des espèces, reprise principalement par Jean-Baptiste de Lamarck et Charles Darwin, fut développée par leurs successeurs immédiats, dont Ernst Haeckel, qui proposa le premier arbre du vivant dans son ensemble tel qu'on le connaissait alors. Avec le séquençage des génomes, la recherche des filiations entre espèces au cours de l'évolution a pris récemment une ampleur considérable, à tel point que nos idées sur les différentes branches de l'arbre évolutif du monde vivant se sont beaucoup clarifiées, modifiant parfois les arbres plus anciens (voir le passage « Arbre simplifié du vivant »).

Dès 1937, Édouard Chatton faisait la différence entre les organismes dont les cellules possèdent un *noyau*\* contenant les chromosomes qui, de ce fait, sont isolés du cytoplasme, et d'autres dont les cellules ne sont pas compartimentées, laissant les chromosomes directement dans le cytoplasme. Les premiers, appelés *eucaryotes*\*, incluent animaux, plantes, champignons, algues, mais également une très grande variété de micro-organismes unicellulaires globalement peu connus du public. Les cellules eucaryotes contiennent aussi d'autres *organites*\* qui portent des gènes qui leur sont propres : les *mitochondries*\* et les *chloroplastes*\*. Les seconds, appelés historiquement *procaryotes*\*<sup>3</sup> parce leur plus grande simplicité pouvait suggérer qu'ils étaient plus primitifs, sont presque exclusivement des micro-organismes unicellulaires initialement confondus sous le nom de *bactéries*\*. Les premières comparaisons de séquences devaient confirmer l'unicité des eucaryotes, mais distinguer deux groupes parmi les procaryotes, dont l'un, appelé aujourd'hui *archées*\*<sup>4</sup>,

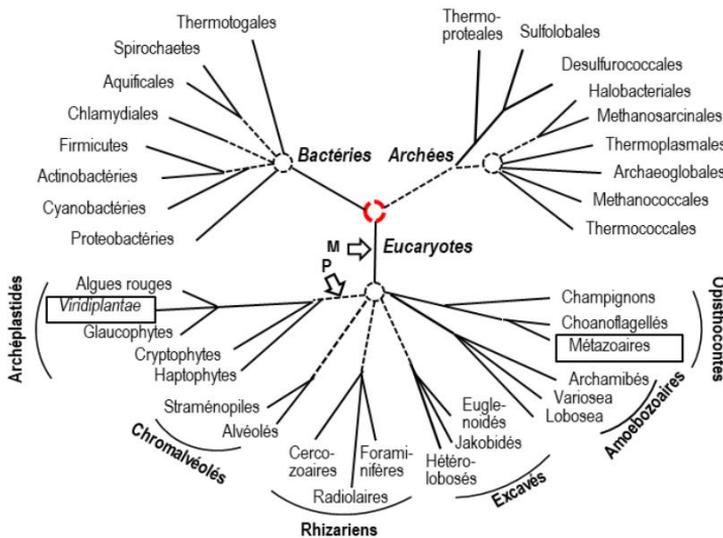
---

3. Le terme exact aurait été « acaryotes ».

4. Pour des raisons historiques, les archées ont été initialement désignées « archéobactéries », tandis que l'autre groupe procaryote était désigné « eubactéries ». Contrairement à cette désignation, rien n'indique que les archées soient plus primitives que les bactéries.

avait été beaucoup moins étudié que l'autre, qui a conservé l'appellation de *bactéries\**. Le monde vivant serait donc fait de trois grands domaines dont l'ordre de séparation initiale reste sujet de discussion. Archées, bactéries et eucaryotes sont eux-mêmes subdivisés chacun en nombreux sous-groupes. Les eucaryotes en comptent au moins cinq, peut-être huit, parmi lesquels animaux et plantes ne représentent que deux lignées particulières parmi un nombre beaucoup plus élevé de lignées de micro-organismes unicellulaires. Mais cette classification ignore le monde des *virus\**, car il ne s'agit plus directement de cellules vivantes, mais de particules moléculaires qui doivent infecter des cellules vivantes pour se multiplier. On connaît des virus qui infectent les archées, d'autres les bactéries, d'autres encore les eucaryotes, chacun avec des spécificités d'hôtes plus ou moins larges. On est loin de connaître tous les virus existants et il est probable qu'il s'en forme de nouveaux constamment. Les derniers résultats d'exploration des océans suggèrent que les virus représentent à eux seuls la plus grande masse du monde vivant.

#### ARBRE SIMPLIFIÉ DU VIVANT.



L'arbre est construit selon les principes phylogénétiques communément admis en se basant principalement sur les séquences de l'ARN ribosomique, molécule commune à tous les organismes vivants cellulaires. Les virus en sont donc exclus. Seules les principales branches sont représentées avec les noms des groupes correspondants. Beaucoup de branches comptent de très grands nombres d'espèces connues, parfois des millions, et on ne sait pas combien d'espèces restent à découvrir, probablement beaucoup. Les traits pleins indiquent les relations

établies. Les pointillés indiquent les incertitudes. Les petites flèches placent les événements d'endosymbiose primaire de lignées bactériennes supposés être à l'origine de la formation des mitochondries (M) et des chloroplastes (P) des cellules eucaryotes. Les mitochondries proviendraient de la lignée des Protéobactéries tandis que les chloroplastes proviendraient de la lignée des Cyanobactéries. Il est évidemment important de bien comprendre qu'il ne s'agit pas des bactéries actuelles, mais de leurs très lointains ancêtres, aujourd'hui disparus.

À partir de l'origine hypothétique (placée au centre, rouge), on distingue trois domaines majeurs, les bactéries, les archées et les eucaryotes. Les noms des principaux sous-groupes d'eucaryotes sont indiqués en gras. On constate qu'à cette échelle globale, l'ensemble des animaux d'une part et l'ensemble des plantes et algues vertes d'autre part ne représentent, chacun, qu'une seule branche terminale (Métazoaires et *Viridiplantae*, respectivement). La grande majorité de la diversité du monde vivant est représentée par des micro-organismes. Cet arbre est inspiré des travaux initiaux de Carl Woese et collaborateurs (Woese *et al.* 1990) avec de nombreuses modifications pour accommoder les résultats ultérieurs.

Les apports à la génétique des trois groupes d'organismes cellulaires et des virus ont été très différents. Comme on le verra, la génétique est née de l'étude de l'hérédité chez les plantes et s'est initialement développée avec l'étude des insectes, des champignons et de quelques autres organismes, tous eucaryotes. Elle est ensuite devenue moléculaire, avec l'étude de bactéries et de leurs virus, les *bactériophages\**, qui ouvrait la voie de l'analyse fine des gènes et fournissait les premiers outils qui serviront plus tard à l'analyse des génomes. Les outils génomiques permettent aujourd'hui d'aborder tous les organismes et leurs virus, mais il reste de cette histoire un biais de connaissance entre les différentes branches du monde vivant qui laisse entrevoir de nouvelles découvertes importantes par l'étude de lignées restées pratiquement inexplorées. On en verra des exemples. Ce biais n'a jamais remis en cause l'universalité des principes établis, mais il interroge sur leur complétude.

## Des changements permanents : les leçons des génomes

Les spectaculaires progrès de la génomique illustrent cette interrogation. Elle nous apprend que les espèces ne diffèrent pas seulement les unes des autres par les formes alléliques des gènes qu'elles partagent en commun, mais aussi par la présence, l'absence ou encore la multiplication de certains gènes. Il en va de même pour les membres d'une même espèce, à la dimension quantitative près. Deux

individus pris au hasard dans l'espèce humaine ne diffèrent pas seulement par les formes alléliques de leurs gènes, mais aussi par le nombre exact de leurs gènes. Ici, la différence reste faible au plan numérique – quelques unités tout au plus – mais pas forcément au plan fonctionnel. Pour d'autres organismes, en particulier des micro-organismes, le nombre de gènes commun à tous les membres de l'espèce peut être significativement inférieur au nombre de gènes de chaque individu, lui-même très inférieur au nombre total estimé de gènes de tous les membres de l'espèce.

Les gains ou pertes de gènes qui expliquent ces variations reposent sur des mécanismes que l'on commence seulement à mieux comprendre. La perte des gènes de génération en génération semble plus ou moins constante, limitée seulement par les effets délétères éventuels qui peuvent en découler. Ces pertes sont compensées numériquement par des gains qui peuvent avoir deux sources : l'acquisition par transfert horizontal à partir d'autres organismes contemporains ou la formation *de novo* de gènes par différents processus de mutation de séquences existantes. Ces événements restent rares au niveau individuel, mais ils n'en jouent pas moins un rôle considérable à l'échelle de l'évolution des populations. Il s'ensuit que tout génome est par essence imparfait, image instantanée de processus dynamiques enchevêtrés dont les effets se trouvent atténués par la résilience de leurs produits. On est loin d'un déterminisme génétique simple et naïf où toute modification d'un génome serait à proscrire.

## Avenir de la génétique : espoirs et craintes

Il est rare que l'on garde présents à l'esprit les progrès des siècles passés. Leurs bienfaits sont considérés comme des acquis. Leurs méfaits sont connus ou oubliés. Les progrès remarquables du XX<sup>e</sup> siècle en matière de production agricole, d'élevage, de santé ou de domestication des micro-organismes avaient une large base génétique. Qui s'en souvient ? Et puisqu'ils ont été réalisés, que reste-t-il à faire ? Les progrès actuels de la génétique semblent soulever autant de questionnements dans le public qu'ils suscitent d'espoir parmi les spécialistes. Lorsqu'une anomalie génétique est responsable d'une maladie, quoi de plus naturel que de chercher les moyens de la réparer, d'autant plus que cette dernière est invalidante ou réduit fortement l'espérance de vie ? Cette idée n'est pas nouvelle, elle a émergé il y a environ soixante ans, avec les découvertes de la structure de l'ADN et du code génétique. La connaissance des génomes nous permet-elle d'anticiper l'apparition des maladies ? La transgénése permet-elle de les traiter ? On vient de montrer la possibilité de transformer des lymphocytes en médicament contre des cellules malignes. Peut-on créer des organismes vivants à partir de génomes obtenus par synthèse chimique de l'ADN ? Aujourd'hui avec des micro-organismes, demain

avec des animaux ou des plantes ? La vie s'est bâtie sur la chimie des acides nucléiques et des protéines. Y avait-il d'autres choix possibles ? Une *xénobiologie\**, c'est-à-dire une nouvelle biologie séparée de notre monde vivant, car utilisant d'autres atomes et d'autres structures moléculaires, est-elle envisageable ?

La génétique ne nous donne-t-elle pas des moyens d'action éthiquement discutables ? La question n'est pas nouvelle. Ainsi s'exprimait William Bateson, l'un des pères fondateurs de la génétique, dans son discours au banquet de clôture de la 4<sup>e</sup> conférence de génétique qui fut organisée à Paris :

« Que va devenir la génétique ? Nous sommes au début, et quand on envisage les profondeurs et les hauteurs qu'elle peut atteindre, on est véritablement étourdi. La génétique donne à l'espèce humaine un pouvoir qu'on ne pouvait jamais prévoir et qui est extrêmement dangereux... [...] peut-être – et je ne crois pas être ridicule en disant cela – posséderons-nous dans un siècle le pouvoir de régler le destin de la race humaine, et les types dont nous ne voudrions pas ne naîtront pas. Je ne suis pas certain qu'un gouvernement possédant ce pouvoir n'en abusera pas. »

Nous étions en 1911. Le siècle est passé. Les dérives eugénistes n'ont abouti à rien d'autre que des drames, tant leurs bases étaient erronées. Elles étaient d'ailleurs inspirées davantage de C. Darwin que de G. Mendel. Mais la prophétie s'est réalisée. Nous avons le pouvoir d'agir sur le destin de la race humaine, de modifier délibérément les écosystèmes, de fabriquer artificiellement des micro-organismes – y compris pathogènes – et, pourquoi pas, bientôt de faire revivre des espèces éteintes. Comment l'appliquer utilement ? C'est la première question qui ait un sens, l'autre étant de savoir comment conserver la curiosité de l'inconnu.

## **Origine de l'ouvrage et contenu des chapitres**

Cet ouvrage illustrera certaines des trajectoires suivies actuellement par la génétique, en insistant sur les notions conceptuelles sur lesquelles elles se fondent. Il est basé sur un colloque (Comptes Rendus Biologies 2016) qui s'est tenu à l'Académie des sciences à Paris en septembre 2016 à l'occasion du 150<sup>e</sup> anniversaire de la publication du mémoire de G. Mendel, *Versuche über Pflanzen-Hybriden*. Ce colloque a réuni des chercheurs français – pays qui s'est longtemps montré réticent à reconnaître cette discipline à sa juste valeur – et étrangers. L'ouvrage abordera les notions fondamentales et les enjeux sans lesquels il est impossible de comprendre l'évolution des recherches actuelles. Il essaiera de montrer qu'au-delà des retombées immédiates, c'est la curiosité intellectuelle qui

donne à la génétique tout son élan vers le futur. La nouveauté et la puissance des dogmes établis par les découvertes successives de la génétique au cours des décennies 1950 et 1960 pouvaient laisser croire qu'elle avait achevé son œuvre, alors qu'on n'avait pas encore isolé chimiquement les gènes. Les développements ultérieurs de l'ADN recombinant et de la génomique, que d'aucuns considéraient soit dangereux soit superflus, allaient au contraire ouvrir des domaines de recherche insoupçonnés qui éclairent désormais toutes les branches de la biologie, des plus fondamentales aux plus appliquées. L'ouvrage illustrera la transition entre ces deux périodes, les trois premiers chapitres traitant des bases fondamentales, les trois suivants de la génomique et les trois derniers des applications présentes et des développements attendus.

Nous remercions Jean-Yves Chapron et Éric Postaire pour leurs conseils avisés et Jean Weissenbach pour sa relecture attentive du manuscrit.

## Bibliographie

- Comptes Rendus Biologies (2016). Trajectoires de la génétique, 150 ans après Mendel. *Comptes Rendus Biologies*, 339(7/8).
- Woese, C.R., Kandler, O., Wheelis, M.L. (1990). Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences*, 87, 4576–4579.