

PATRICE  
DAVID

SARAH  
SAMADI

# La théorie de l'évolution

Une logique pour  
la biologie

Champs sciences



# LA THÉORIE DE L'ÉVOLUTION



Patrice DAVID  
Sarah SAMADI

# LA THÉORIE DE L'ÉVOLUTION

Une logique pour la biologie

Édition revue et augmentée 2011

**Champs sciences**

© Flammarion, 2000

© Flammarion, 2011, pour la présente édition  
revue et augmentée  
ISBN 978-2-0812-6597-4

## AVANT-PROPOS

L'objectif de cet ouvrage est de faire un exposé sur la théorie de l'évolution accessible aux étudiants, aux enseignants ainsi qu'à un public plus large, désireux de s'y initier. Il n'est pas nécessaire de posséder une connaissance spécialisée de la biologie pour lire cet ouvrage, dans lequel peu de termes techniques sont utilisés. La définition de ces termes est toujours donnée dans le texte lors de leur première utilisation (signalée en italique).

La théorie de l'évolution explique l'histoire du vivant. Nous avons choisi d'insister sur la cohérence de la théorie et donc sur ses mécanismes plutôt que sur l'histoire elle-même. Du fait de cette cohérence, la théorie de l'évolution offre une logique commune à l'ensemble de la biologie. Cette logique sert de fil conducteur à cet ouvrage. Nous sommes tous concernés par les progrès des biotechnologies dans les domaines agricole et médical (OGM, clonage, séquençage du génome humain...) tout comme par les problèmes d'environnement. Ainsi, la biologie prend une place telle dans notre vie quotidienne qu'il est intéressant pour tous d'en connaître les fondements logiques.

Le texte principal est composé de onze chapitres qui s'achèvent chacun par un résumé retracant sa structure logique. Les références aux exemples traités sont données dans les notes en bas de page. Un glossaire critique

permet de faire brièvement le point sur quelques notions évoquées dans le texte. La bibliographie générale apporte quelques références permettant au lecteur d'approfondir les thèmes abordés dans ce livre. Un index complète l'ouvrage.

Le premier chapitre pose les fondations théoriques de l'évolution. Les trois chapitres suivants détaillent les axiomes de la théorie (mutation, dérive génétique et sélection naturelle) en restant au niveau de l'unité héréditaire fondamentale : le gène. Les chapitres suivants introduisent des niveaux d'intégration supérieurs. Le chapitre 5 explique pourquoi les gènes s'associent entre eux pour former des individus. Les conséquences de cette association sur l'évolution des gènes sont abordées au chapitre 6. Les modalités d'association peuvent elles-mêmes évoluer grâce à une stratégie originale : le sexe, abordé au chapitre 7. Les quatre derniers chapitres introduisent l'évolution des caractères. Le chapitre 8 décrit les relations entre l'évolution des gènes et celle des caractères. Le chapitre 9 intègre l'évolution des caractères des individus dans un contexte spatial et temporel variable. Le chapitre 10 détaille le rôle de la composante biologique de ce contexte : les relations entre organismes. Enfin, le chapitre 11 aborde l'évolution à une échelle supérieure, celle des espèces.

## LES PRINCIPES FONDATEURS DE LA THÉORIE DE L'ÉVOLUTION

### Introduction

**Le domaine de la théorie : les êtres vivants.** Les objets de la théorie de l'évolution, comme de toute la biologie, sont les êtres vivants. Au contraire des objets de la physique ou de la chimie, les êtres vivants sont des objets historiques, dont l'existence est restreinte à un lieu précis de l'Univers et à une époque précise de son histoire. Le lieu est la surface du globe terrestre, autour duquel la matière vivante forme une couche mince appelée *biosphère*. L'époque est une période relativement récente dans la vie de l'Univers, c'est-à-dire, grossièrement, les quatre derniers milliards d'années. C'est en effet dans des roches âgées de trois à quatre milliards d'années qu'on trouve les premières traces d'êtres vivants, sous forme de fossiles d'organismes proches des cyanobactéries actuelles.

Pour construire une théorie de l'évolution, il faut tout d'abord délimiter son champ en donnant une définition permettant de reconnaître les êtres vivants des autres objets. Cette définition est un problème aussi vieux que la biologie. On verra que la théorie elle-même fournit un moyen de définir le vivant. Cependant, sans l'utiliser, les hommes ont des moyens intuitifs de reconnaître les

objets vivants, et aboutissent à un consensus remarquable : la plupart du temps, deux personnes seront d'accord pour dire que tel objet est vivant ou ne l'est pas. Recherchons donc des caractéristiques propres au vivant, sur lesquelles nos intuitions pourraient se fonder :

(i) *La matière constituant les êtres vivants (ou produite par eux) est différente de celle qui constitue les autres objets.* Cette matière, dite *organique*, est constituée principalement de molécules à base d'atomes de carbone et d'hydrogène, auxquels s'ajoutent des proportions variables d'oxygène, d'azote et d'autres éléments minoritaires (phosphore, fer, soufre...). Jusqu'au XIX<sup>e</sup> siècle dominait la conception *vitaliste*, selon laquelle les réactions chimiques de la matière organique sont de nature différente des autres, et réalisables uniquement par des systèmes dotés d'une « énergie vitale » propre aux organismes vivants. La synthèse artificielle de l'urée, petite molécule organique produite par de nombreux animaux, par le chimiste Friedrich Wöhler en 1824, ainsi que le développement ultérieur, spectaculaire, de la chimie organique, ont réfuté cette hypothèse. Les réactions chimiques des molécules organiques suivent les mêmes lois thermodynamiques et chimiques que les autres. Une version atténuée du programme vitaliste peut néanmoins être admise. Parmi les objets non artificiels, les êtres vivants se distinguent par le fait qu'ils sont constitués de molécules organiques souvent très complexes et introuvables dans les objets non vivants. Par ailleurs, ils sont le siège d'une activité chimique intense et originale, le métabolisme, comportant un grand nombre de réactions précisément coordonnées les unes aux autres.

(ii) *Les êtres vivants ont une architecture complexe.* Ils ressemblent au bateau de Thésée qui reste lui-même bien que l'on ait changé toutes ses planches une à une. En

effet, les molécules qui constituent un organisme n'y séjournent que temporairement et sont remplacées par d'autres molécules. Ce qui fait l'organisme, c'est l'agencement très complexe de ses constituants, qui se perpétue malgré le flux permanent de matière.

(iii) *Les êtres vivants se reproduisent.* Cela signifie que des parties d'êtres vivants (par exemple, des cellules sexuelles, spermatozoïdes ou ovules) peuvent s'en séparer pour reformer un être d'apparence et d'organisation semblables. Les descendants héritent de l'essentiel des caractéristiques de leurs parents mais ne leur sont généralement pas strictement identiques. Cependant, la capacité à reproduire des caractéristiques semblables de génération en génération signale qu'une information est transmise des parents aux descendants. Cette information constitue le *patrimoine génétique* et repose sur des molécules présentes au sein de toutes les cellules vivantes. Ces molécules très complexes, appelées *acides nucléiques*, peuvent être considérées, au même titre que la reproduction, comme une caractéristique du monde vivant.

Nous disposons désormais de critères pour distinguer les êtres vivants connus des autres objets. Toutefois, ces critères ne sont pas tous de même nature. Pour le constater, il suffit de se poser la question suivante : lequel, parmi ces critères, me permettrait, en rencontrant un objet extraterrestre, de déterminer s'il est vivant ? Cette question est importante car de telles propriétés doivent nécessairement être incluses dans une théorie de l'évolution. En effet, les propriétés valables uniquement sur Terre peuvent être des accidents historiques. Certes, nous ne connaissons pas d'autres êtres vivants que ceux de notre planète, mais l'histoire de cette planète n'est qu'une réalisation parmi un ensemble d'histoires possibles représentant la classe de phénomènes à laquelle la théorie de

l'évolution choisit de s'intéresser. Il importe donc de classer les critères de définition du vivant en mettant d'un côté ceux qui se conservent dans toutes les histoires possibles compatibles avec notre théorie et, de l'autre côté, ceux qui sont contingents à l'histoire qui s'est effectivement déroulée. Ces derniers ne doivent pas faire partie des bases générales de la théorie, même si les tests expérimentaux doivent les prendre en compte.

Parmi les caractères cités, seule la reproduction (imparfaite) est un critère universel. En effet, la nature chimique ou le degré d'organisation observé chez les organismes terrestres peuvent être des accidents liés à l'histoire de la Terre et aux conditions dans laquelle la vie y est apparue. En conséquence, l'étude de l'évolution de la vie terrestre est double. Certaines parties s'appliqueront uniquement aux êtres vivants produits lors de l'histoire de la Terre, alors que d'autres parties décriront des phénomènes généraux qui concernent des objets ayant comme propriété minimale la reproduction. Cette catégorie comprend les êtres vivants terrestres, mais aussi n'importe quelle entité pour laquelle un processus de type « reproduction pas tout à fait à l'identique » peut être défini. C'est par exemple le cas des virus informatiques (programmes capables de se dupliquer dans les mémoires d'ordinateurs), des rumeurs (capables de se répandre au sein des populations humaines), des langues ou des noms de famille (capables de se transmettre au fil des générations).

**Deux versants de la théorie : historique et mécanistique.** Suite à la discussion précédente, nous distinguerons deux versants dans les sciences de l'évolution : (i) *Un versant descriptif*, dont le but est de raconter l'histoire des êtres vivants sur Terre depuis leur apparition.

Pour cela, à la manière des historiens, il faut compulsier des archives (les fossiles), rechercher des témoins directs ou indirects (traces actuelles de processus évolutifs passés), et établir des chronologies. (ii) *Un versant explicatif*, dont le but est de décrire les lois et les mécanismes qui sous-tendent cette histoire. Il s'agit ici de proposer des liens logiques permettant de prédire les phénomènes observés ou de les comprendre *a posteriori*. Cette partie pourrait, au contraire de la précédente, s'appliquer à n'importe quels objets possédant les propriétés requises (reproduction imparfaite), en particulier à une forme de vie extraterrestre qu'on viendrait à découvrir. Pour l'instant, on ne teste expérimentalement la théorie qu'avec les objets terrestres. Une exigence minimale est donc que ce versant soit cohérent avec la description historique. Remarquons que cette dichotomie se retrouve dans d'autres disciplines scientifiques étudiant des objets non nécessairement vivants (géologie, sciences physiques).

L'aspect historique de l'évolution est le plus connu du grand public, souvent impressionné, à juste titre, par l'étrangeté des faunes passées (les dinosaures de *Jurassic Park*). Il a également fait l'objet de nombreux ouvrages de vulgarisation, surtout paléontologiques. Au cours de ce livre, nous insisterons plutôt sur l'explication des mécanismes, même si les données paléontologiques et biologiques interviennent pour illustrer et tester ces mécanismes tout au long de l'ouvrage. Ces mécanismes ne sont pas connus parfaitement. Ils n'ont pas été découverts en un jour, mais au cours d'un siècle et demi de recherches qui se poursuivent encore activement. Dans ce premier chapitre, nous décrirons successivement les lignes de force des deux versants, historique et mécanistique, de la théorie, telle qu'elle se présente aujourd'hui,

dans un ordre logique plutôt que chronologique (les deux coïncidant parfois).

## L'évolution : un récit de l'histoire des êtres vivants sur Terre

**Les êtres vivants ne sont pas fixes.** La théorie de l'évolution rejette les dogmes *fixistes* selon lesquels les espèces vivantes sont immuables et n'ont pas changé depuis leur apparition. Ces dogmes étaient essentiellement, dans le monde chrétien, fondés sur des textes religieux. Ces derniers décrivent l'origine de la vie comme une création instantanée des êtres vivants par Dieu, suivant un calendrier plutôt serré (les sept jours de la Genèse). Les espèces vivantes seraient restées identiques depuis lors. Sans préjuger de sa portée métaphorique, l'interprétation littérale de ce récit peut être facilement réfutée (au sens scientifique du terme). L'argument le plus simple est l'observation des fossiles, qui représentent les faunes et flores passées. Ces fossiles sont souvent très différents des animaux actuels, et cela d'autant plus qu'ils sont anciens. Personne n'a rencontré d'ammonite, de tyrannosaure ou de trilobite dans la nature actuelle, bien qu'ils fassent partie de grands groupes qui ont des représentants dans nos faunes (mollusques, vertébrés, arthropodes).

**Les espèces vivantes proviennent les unes des autres par descendance avec modification.** Une version remaniée des dogmes fixistes/créationnistes, prenant en compte les changements de faunes observés dans les fossiles, a été proposée par Georges Cuvier<sup>1</sup>. Cette théorie

---

1. G. Cuvier, *Discours sur les révolutions de la surface du globe*, Paris, Bourgois, 1985 (1<sup>er</sup> éd. 1825).

suppose l'existence de plusieurs périodes dans l'histoire de la vie. Chacune commence par la création de nouvelles espèces vivantes, qui restent fixes pendant toute la période puis sont anéanties par une catastrophe planétaire (pluies diluvienues). Une nouvelle période commence alors. Chacune des périodes correspond à une couche de dépôts géologiques qui s'accumule au-dessus des précédentes, séparée par une discontinuité révélant l'épisode catastrophique. Ce « fixisme par morceaux » est rejeté par la théorie de l'évolution, qui postule que (i) tous les organismes peuplant ou ayant peuplé la Terre depuis l'apparition de la vie descendent les uns des autres *via* un processus de reproduction. (ii) Les modifications morphologiques observées se sont donc produites au sein de lignées généalogiques d'organismes et non par création spéciale de nouvelles lignées. Ces deux affirmations sont résumées par Charles Darwin<sup>1</sup> dans l'expression « descendance avec modification ». Quatre arguments principaux permettent d'étayer cette position :

(1) *On observe des variations graduelles au cours du temps.* En effet, on peut souvent regrouper des organismes fossiles proches ayant vécu à des époques successives en une série ordonnée, le long de laquelle des changements directionnels de morphologie sont observés. Par exemple, certaines séries de chevaux et de rongeurs fossiles montrent une augmentation régulière de la taille du corps. La discontinuité intrinsèque des archives géologiques (interruptions de sédimentation) rend difficile l'estimation du degré de continuité ou de discontinuité réelles au sein de ces séries. Cependant, des formes observées à une date *t* peuvent presque toujours être rap-

---

1. Ch. Darwin, *L'Origine des espèces*, Paris, GF-Flammarion, 1992 (1<sup>re</sup> éd. 1859).

prochées de formes observées aux temps immédiatement précédents. Ces observations suggèrent que les modifications morphologiques peuvent avoir lieu au sein d'une même lignée d'animaux descendants les uns des autres, sans faire appel à des créations spéciales de nouvelles lignées.

(2) *On observe des modifications en cours au sein des espèces actuelles.* Ces situations restent exceptionnelles car le rythme des modifications est en général trop lent pour être observé de mémoire d'homme. On a pu montrer néanmoins que la couleur noire a progressivement remplacé la couleur blanche dans les populations d'un papillon, la phalène du bouleau. D'autres exemples concernent des modifications physiologiques plutôt que morphologiques. Ainsi, les micro-organismes pathogènes (comme l'agent unicellulaire du paludisme) développent, en quelques années, des résistances aux nouvelles drogues utilisées contre eux. C'est pourquoi les drogues, efficaces au début de leur mise sur le marché, deviennent rapidement inutiles et doivent être remplacées par de nouvelles drogues. Les modifications les plus spectaculaires restent celles qui accompagnent la domestication par l'homme d'espèces animales ou végétales. On sait que les animaux et plantes domestiques ont été à l'origine prélevés dans la nature, et on connaît en général une espèce sauvage proche qui est la source probable des populations domestiquées. Ainsi, le « chien ancestral » ressemblait à un loup actuel. En l'espace de quelques milliers d'années, ses descendants ont été modifiés jusqu'à atteindre une diversité morphologique étonnante (du berger allemand au pékinois). Des modifications tout aussi spectaculaires ont eu lieu chez le blé et le maïs par rapport à leurs ancêtres sauvages.

(3) *Le principe d'actualisme énoncé par Charles Lyell (géologue contemporain de Darwin) interdit d'expliquer le passé en faisant appel à des causes radicalement différentes de celles qui agissent ou peuvent agir dans le présent.* On s'interdit ainsi d'énoncer des hypothèses gratuites venant opportunément modifier les règles du jeu pour expliquer *a posteriori* des événements historiques. Il repose donc sur le principe d'économie des hypothèses (*parcimonie*) excluant la création divine et les catastrophes à répétition), car aucune de ces causes ne peut être raisonnablement identifiée aujourd'hui.

(4) *Les ressemblances et différences observables entre espèces vivantes et éteintes ont de tout temps été organisées en une classification hiérarchique.* Cette classification a été établie bien avant l'éclosion de la théorie de l'évolution. Elle n'est pas arbitraire mais repose sur des degrés de similitude établis sur des critères morpho-anatomiques dont l'importance est reconnue de tous les biologistes même si des points précis de classification peuvent faire l'objet de débats. Au cours des temps géologiques, les espèces changent mais les changements ne remettent pas en question les grandes lignes de la classification. On s'accorde donc à penser qu'elle reflète une organisation réelle du monde vivant. Pour expliquer la structure hiérarchique de la classification dans le cadre du scénario « créations/catastrophes », il est nécessaire d'ajouter l'hypothèse d'une volonté du Créateur de créer un ordre précis et de le réitérer après chaque catastrophe. L'hypothèse de descendance avec modification économiquement la structure hiérarchique du monde vivant, en lui superposant un arbre généalogique de formes descendant les unes des autres. Les formes ressemblantes sont proches parentes (séparées récemment), alors que les formes très différentes sont des parentes éloignées

(séparées anciennement). Les représentations graphiques de ces généalogies d'espèces vivantes ont une forme d'arbre (arbres *phylogénétiques*). À titre d'exemple, un tel arbre est donné en figure 1.

**Le graphe de l'histoire de la vie est un arbre à une seule racine.** Un prolongement du raisonnement précédent concerne l'unicité de l'origine de la vie. Cette question est ancienne mais elle change de sens suivant que l'on a une vision fixiste ou évolutive du monde vivant. Les fixistes admettaient la génération spontanée de formes complexes, qu'elle soit ou non d'origine divine. L'idée d'évolution des organismes vivants était dès lors inutile comme explication de l'apparition de nouvelles formes.

La première étape du débat fut de rejeter la génération spontanée de formes complexes. Jean-Baptiste de Lamarck<sup>1</sup> considérait par exemple que la génération spontanée ne concernait que des formes de vie simples, qui, par transformation graduelle, auraient été ensuite à l'origine de formes complexes. Il sortit ainsi du cadre fixiste pour rentrer pour la première fois dans celui du *transformisme*. La réfutation définitive en 1862, par Louis Pasteur, de toute forme de génération spontanée permit, dans le cadre de la théorie de l'évolution, de conclure à une origine unique de la vie et donc à une généalogie unique de tous les êtres vivants. La théorie de l'évolution offre donc dans ce cadre une *nouvelle définition des êtres vivants terrestres : un ensemble d'objets reliés les uns aux autres par des liens de parenté*. Cette définition englobe tous les critères précédents : les liens de parenté

---

1. J.-B. Lamarck, *Philosophie zoologique*, Paris, GF-Flammarion, 1994 (1<sup>er</sup> éd. 1809).

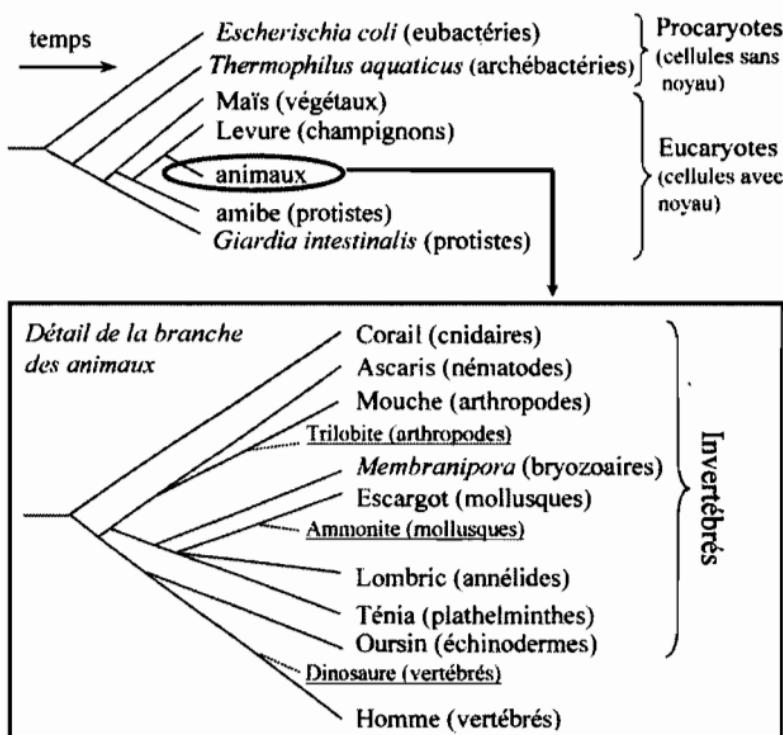


Figure 1 : L'arbre phylogénétique du monde vivant d'après K. M. Halanych<sup>1</sup>. Seules quelques espèces, représentatives des groupes les plus importants (entre parenthèses), sont représentées. Autant que possible, ces espèces ont été nommées par leur nom commun, plus familier au lecteur. L'arbre du haut représente les grandes subdivisions du vivant, la branche des animaux étant détaillée dans l'arbre du bas. Les « branches mortes », soulignées, représentent les groupes éteints (ammonites, trilobites, dinosaures). N.B. : ces arbres sont des reconstructions basées sur les données et techniques actuellement disponibles. Ils sont donc susceptibles de contenir des erreurs, et certains branchements importants sont encore débattus.

1. K. M. Halanych, « The new view of animal phylogeny », *Annual Review of Ecology and Systematics*, 35, 2004, p. 229-256.

impliquent une activité de reproduction et expliquent les ressemblances chimiques (matière organique, ADN).

Une question, déjà soulevée par Darwin, reste cependant en suspens : si la génération spontanée est impossible, comment expliquer l'apparition du premier être vivant ancêtre de tous les autres ? Cette question a été abordée au travers d'approches à la fois expérimentales et théoriques. L'idée généralement admise est que les conditions régnant sur Terre ont changé : elles ont permis, il y a quatre milliards d'années, l'apparition de la vie, chose devenue impossible depuis. Pour étayer cette hypothèse, des expériences ont tenté de reproduire les conditions physico-chimiques de la Terre primitive afin de montrer que ces conditions conduisaient à la synthèse de molécules caractéristiques de la matière vivante et à leur complexification progressive. Sur cette base, de nombreuses théories ont pu être formulées, proposant des scénarios par lesquels ces molécules auraient acquis un pouvoir d'autoréplication<sup>1</sup>. L'histoire géologique et l'action des êtres vivants sur leur environnement auraient ensuite modifié les conditions physico-chimiques, rendant impossible une nouvelle apparition de la vie sur Terre.

**L'arbre possède des branches mortes.** Un élément important dans l'histoire de la vie est la disparition de formes de vie que l'on peut notamment constater dans le registre fossile. On parle *d'extinction* quand un ensemble d'êtres vivants disparaît sans laisser de descendants, c'est-à-dire sans se transformer en de nouvelles formes. Les extinctions sont nombreuses tout au long de l'histoire de la vie. Depuis que l'homme existe, ses actions ont provoqué de nombreuses extinctions : on peut citer le cas

---

1. M.-C. Maurel, *Les Origines de la vie*, Paris, Syros, 1994.

classique du dodo, oiseau terrestre de l'île Maurice entièrement décimé à la suite de l'arrivée de l'homme. Hormis les extinctions d'origine humaine, des *extinctions de masse* spectaculaires (beaucoup d'espèces disparues en peu de temps) jalonnent l'histoire de la vie, telle que la racontent les fossiles. L'exemple le plus célèbre est bien entendu la mort des dinosaures à la fin du Crétacé il y a soixante-cinq millions d'années. À cette époque, l'impact d'une comète géante sur Terre, ou des éruptions volcaniques particulièrement violentes, ont pu être à l'origine de profondes modifications écologiques entraînant la mort de centaines d'espèces. Ces extinctions n'ont pas exclusivement concerné les très médiatiques dinosaures. À la même période, de nombreux organismes ont également disparu comme les ammonites, un groupe alors florissant de mollusques marins (figure 1).

## L'évolution : les grands principes d'une théorie des changements des formes vivantes

**La bibliothèque de Babel et la métaphore typographique de Richard Dawkins.** Richard Dawkins<sup>1</sup> a proposé d'expliquer les grands principes de l'évolution à partir d'une métaphore typographique, où les êtres vivants sont représentés par des textes recopiés sans relâche mais avec des possibilités d'erreur. Si nous reprenons et élargissons cette métaphore pour expliquer la mécanique de la théorie, le monde vivant devient une bibliothèque géante dont il faut expliquer la constitution et la construction.

1. R. Dawkins, *L'Horloger aveugle*, Paris, Robert Laffont, 1989 (1<sup>re</sup> éd. 1986).

Jorge Luis Borges, dans *Fictions*, a imaginé une « bibliothèque de Babel » qui contiendrait tous les livres qu'il est possible de composer à partir d'un nombre fini de caractères. Dans cette bibliothèque se trouvent des livres qui ont un sens, comme ceux de Shakespeare et de Dickens, mais perdus au milieu d'un fatras innombrable de livres dépourvus de sens. Les êtres vivants ressemblent aux livres de Shakespeare et de Dickens : ils ont d'une certaine manière un sens. Bien que n'ayant pas été constitués par un auteur avec une intention précise, ils présentent de nombreuses apparences de *téléconomie* : grossièrement, ils ont l'air « bien faits » pour assurer des fonctions (respiration, reproduction) ou répondre aux contraintes du milieu.

La littérature mondiale ne contient pas tous les livres possibles ayant un sens, et les œuvres qu'elle contient ne peuvent être décrétées comme les « meilleures possibles ». De même, le monde vivant n'est pas optimal. Il ne contient pas tous les êtres vivants « bien faits » possibles et les êtres vivants existants ne sont pas parfaits. Les excès d'une vision trop optimiste de l'adaptation ont été élégamment critiqués par Stephen Jay Gould et Richard Lewontin (voir *Le programme adaptationniste*, p. 337). Pour comprendre le monde vivant, il faut donc expliquer pourquoi il est « bien fait » sans être parfait. Dans notre métaphore, il faut comprendre pourquoi la littérature mondiale, contrairement à la bibliothèque de Babel, ne contient que des livres comme ceux de Shakespeare et de Dickens.

**Le fonctionnement de la bibliothèque.** La première explication possible fait appel à Dieu dans le sens du Grand Horloger, qui devient ici un Grand Bibliothécaire. On se retrouve alors devant un paradoxe :

comment Dieu n'a-t-il pas fait la meilleure des bibliothèques possibles ? Pourquoi a-t-il fait des choix arbitraires ? Une explication plus terrestre du fonctionnement de cette bibliothèque fait appel aux grands principes de l'évolution.

Le premier point à résoudre est celui de la perpétuation des livres, à travers un processus de *réplication*. Imaginons le scénario suivant. Dans notre bibliothèque, ce sont des copistes qui ont en charge cette tâche. Chaque jour, ils se mettent au travail, recopient tous les livres présents en un certain nombre d'exemplaires et jettent les vieux livres abîmés. On obtient ainsi une collection d'entités, les livres, qui se reproduisent (copie) et meurent au bout d'un certain temps. Les générations de livres se succèdent au cours du temps. Au commencement de l'histoire de la bibliothèque il n'y a qu'un seul livre qui ne contient qu'un petit mot, voire une lettre isolée, mais, peu à peu, la bibliothèque se peuple de nombreux livres, copies indirectes de ce lointain ancêtre.

Pourquoi les livres ne sont-ils pas tous identiques à l'ancêtre ? En d'autres termes, quelle est l'origine de la variation entre copies ? Les copistes, débordés de travail, font des erreurs typographiques (changement d'une lettre, oubli d'une phrase ou d'un mot, insertion d'une phrase provenant d'un autre livre). Ce taux d'erreur crée des variations entre les livres. Ces erreurs sont aléatoires et ne se font pas dans un but précis (même si certaines erreurs sont plus fréquentes que d'autres). Par ailleurs, elles seront recopiées par les copistes au jour suivant, comme si de rien n'était, puisque les copistes n'auront pas moyen de vérifier l'original, éliminé en raison de son usure.

Pour résoudre les problèmes de place disponible dans la bibliothèque et de personnel insuffisant, un processus de

tri doit être opéré parmi le très grand nombre de livres à copier. Une première méthode de tri, simple et économique, est la suivante. Parmi un certain nombre de livres qui se ressemblent beaucoup, certains sont jetés sans être recopiés ; ceux qui sont gardés pouvant éventuellement être recopiés en plusieurs exemplaires. Certains textes vont donc être perpétués, d'autres éliminés, *au hasard*. Ce hasard aboutira à une dérive des textes les uns par rapport aux autres. Les textes-fils deviendront de plus en plus différents entre eux et différents du texte-ancêtre original, les étapes intermédiaires étant perdues au cours du tri aléatoire.

Une deuxième méthode de tri fait appel à la population des lecteurs de la bibliothèque. Ceux-ci lisent des échantillons de livres chaque jour et donnent leur avis (j'ai beaucoup aimé, je n'ai pas aimé, je n'ai rien compris). Les livres qui ont un sens sont en général plus appréciés que les livres contenant des suites de caractères inintelligibles. Cependant, ce qui est intéressant pour un lecteur ne l'est pas forcément pour l'autre. Malgré cela, les copistes tiennent compte de l'avis des lecteurs et recopient en priorité les livres qui ont reçu des éloges. Ce processus permet de retirer peu à peu du fatras des œuvres comme celles de Shakespeare ou de Dickens. Bien sûr, beaucoup de chefs-d'œuvre ne voient jamais le jour car ils ont été éliminés par hasard, ou ont été confiés, en cours d'élaboration, à des lecteurs particulièrement difficiles.

**Implémentation biologique.** Le fonctionnement de la bibliothèque a été décrit en trois points : (i) la *copie*, qui assure la pérennité des livres au cours du temps, (ii) la *variation*, qui assure l'apparition de nouveaux livres à la suite d'erreurs de copie, (iii) le *tri*, qui assure une taille finie à la bibliothèque soit par un processus aléatoire,

soit par un processus sélectif. Chacun de ces processus peut être traduit en termes biologiques.

Le processus de copie des livres correspond à la *reproduction* ou *réPLICATION* des êtres vivants. L'original représente, dans ce cas, le parent et la copie le descendant. La notion de réPLICATION comporte implicitement une notion d'hérédité. En effet, le descendant ressemble à son ou à ses parents plus qu'à un organisme pris au hasard. On qualifie d'*héritaire* ou *génétique* cette information qui se transmet de parent à descendant. Traditionnellement, on distingue deux niveaux : celui de l'information elle-même et celui de son expression physique sous forme de caractères visibles. Dans la métaphore de la bibliothèque, l'information (le contenu informationnel du livre) est véhiculée au travers d'un objet-livre qui est sa représentation physique (le papier des pages, l'encre qui dessine les caractères) et interagit ainsi avec les lecteurs. Chez les êtres vivants, l'information est appelée *génotype* (ensemble des informations héréditaires), et sa représentation physique est le *phénotype* (caractères matériels observables sur un individu et interagissant avec l'environnement).

La source de la *variation* (erreurs de copie) est, chez les êtres vivants, la *mutation*. On désigne par là un changement aléatoire en vertu duquel l'information génétique reçue par le descendant est différente de celle des parents. Ces événements sont rares mais constituent la source première de toute nouveauté évolutive ; nous verrons au chapitre 2 leurs mécanismes biologiques et leurs conséquences.

Les deux processus de *tri*, *aléatoire* et *sélectif*, correspondent aux forces évolutives de *dérive génétique* et de *sélection naturelle* (détailées aux chapitres 3 et 4). Dans la nature, comme dans notre bibliothèque fictive, les

différents individus ne laissent pas tous le même nombre de descendants, en particulier parce que certains meurent avant de se reproduire. Cette variabilité a deux composantes. La première (aléatoire) est non prédictible : à la suite de circonstances hasardeuses, tel individu survit et tel autre meurt. Cette composante correspond à la *dérive génétique*. La deuxième composante (déterministe) est fondée sur la confrontation entre les phénotypes et l'environnement (qui correspond dans la métaphore à un lecteur particulier qui donne son avis sur un livre). Le phénotype le plus adapté à l'environnement survit, alors que le phénotype inadapté meurt. Le résultat de cette confrontation peut être partiellement prédit si on connaît bien les deux partenaires (environnement et phénotype). Cette composante correspond à la *sélection naturelle*. On peut associer au processus de sélection naturelle une quantité qui mesure l'aptitude d'un organisme à se perpétuer dans un environnement (dans notre métaphore, la capacité du livre à intéresser les lecteurs) et permet donc de prédire son succès ultérieur. Cette quantité, appelée *valeur sélective* d'un individu (ou *fitness*), est centrale dans la théorie de l'évolution. La *valeur sélective* peut être exprimée comme étant le nombre de descendants laissé par un individu au cours de la totalité de sa vie : elle dépend de son génotype et de l'environnement.

Comment le tri (sélection naturelle et dérive génétique) se combine-t-il à la mutation pour aboutir à une évolution ? Il faut pour le comprendre se placer à l'échelle d'une population d'individus. Dans une population, les mutations créent en permanence de nouveaux génotypes, mais, comme la mutation est un événement rare, ces génotypes sont au départ représentés par très peu d'individus. Pour qu'il y ait évolution, il faut que les

nouveaux génotypes remplacent les anciens, c'est-à-dire que leur *fréquence* augmente dans la population, jusqu'à l'envahir entièrement. Ce résultat peut être obtenu par le hasard (dérive génétique), qui entraîne des changements aléatoires de fréquences. Cependant, il est beaucoup plus facile d'envahir par hasard une population de quatre individus qu'une population d'un million d'individus : la dérive génétique agira donc principalement quand les effectifs sont faibles. Quand les effectifs sont élevés, la sélection naturelle est un scénario beaucoup plus probable : le nouveau génotype envahira la population peu à peu s'il produit plus de descendants que les autres (valeur sélective élevée).

### En quoi l'évolution est-elle une théorie ?

Le mot *théorie* peut être entendu dans deux sens radicalement différents. Une théorie est, dans le sens commun, une hypothèse incertaine expliquant un phénomène particulier. Dans le sens fort, c'est une construction rationnelle qui permet d'expliquer et de prédire des phénomènes naturels. L'exemple type en est la théorie newtonienne du mouvement.

La première question soulevée est donc la pertinence de la théorie de l'évolution dans l'acception commune du mot théorie. En d'autres termes, peut-on douter de la réalité de l'évolution en tant que récit historique ? Dans l'état actuel des connaissances, la réponse est clairement négative. Les preuves sont désormais légion (comme pour la mécanique newtonienne dans son domaine de validité), et il n'y a pas plus de doute sur le récit historique « évolution » que sur l'existence passée de la guerre de Cent Ans. Certaines zones d'ombre persistent

car notre documentation est fragmentaire (comme pour la guerre de Cent Ans) mais la réalité de l'évolution n'est pas en doute, à moins d'adopter un scepticisme extrêmement radical. Les créationnistes cherchent à opposer les récits bibliques et le récit historique de l'évolution. Cependant, cette opposition n'a pas lieu d'être, car le récit biblique n'est pas de nature scientifique, n'étant pas intrinsèquement réfutable. Il est aussi absurde d'opposer la Bible à la théorie de l'évolution que d'opposer une poésie de Guillaume Apollinaire à un annuaire du téléphone. L'annuaire du téléphone peut contenir des erreurs (comme la théorie de l'évolution) car il est censé être une description rationnellement organisée et exacte des objets dont il s'occupe ; la poésie (comme la Bible ou le discours religieux en général) ne peut pas être jugée selon les mêmes critères, ses rapports au réel étant essentiellement métaphoriques. Si l'on nie la nature métaphorique du discours biblique, il est alors nécessaire de le considérer comme une description exacte et rationnelle (scientifique). Or de nombreux éléments bibliques, comme pour ne citer qu'un exemple, l'âge de la Terre (évalué à cinq mille ans d'après la Bible) ne sont tout simplement pas crédibles (comme le montrent les datations utilisant les éléments radioactifs, la Terre a plusieurs milliards d'années). Cette attitude est néanmoins adoptée par une partie des créationnistes, notamment américains, à tel point que le créationnisme est enseigné dans certains États comme une hypothèse alternative à l'évolution. Paradoxalement, c'est également aux États-Unis que se situe aujourd'hui le centre de gravité de la recherche scientifique en biologie évolutive.

Une fois acceptée la pertinence de la théorie de l'évolution comme récit historique, on peut s'interroger sur son statut de théorie scientifique au sens fort. La struc-

ture du discours évolutionniste justifie-t-elle un tel statut ? Nous apporterons une réponse positive à cette question. Les aspects techniques de ce problème dépassent le cadre de cet ouvrage<sup>1</sup> ; nous nous contenterons ici d'illustrer notre propos en montrant que l'évolution est analogue dans sa structure à la mécanique newtonienne, l'exemple type de théorie scientifique.

Tout d'abord il est nécessaire de définir la *classe de phénomènes* que chacune des deux théories cherche à expliquer. La mécanique newtonienne explique *les mouvements d'objets matériels* dans l'espace. La théorie de l'évolution explique les *changements génétiques* au cours du temps, au sein du monde vivant. Ces changements s'expriment comme des modifications temporelles de la fréquence relative des divers génotypes au sein de populations.

Ensuite, on peut pour chaque théorie poser l'*hypothèse nulle*, c'est-à-dire l'ensemble des conditions où aucune modification du système considéré n'entre en jeu. Pour la mécanique newtonienne, l'hypothèse nulle est le *principe d'inertie* : quand aucune force ne s'applique à un système, son mouvement est rectiligne et uniforme (accélération nulle). Pour la théorie de l'évolution, c'est le *principe de Hardy et Weinberg* : en l'absence de sélection et de dérive génétique, la composition génétique d'une population isolée (c'est à dire la fréquence des différents génotypes dans cette population) ne change pas (voir p. 30).

Le point central des deux théories est l'*origine des changements* qu'elles cherchent à expliquer. Dans la

---

1. A. Barberousse, S. Samadi, « Pourquoi et comment formaliser la théorie de l'évolution ? », in *Les Mondes darwiniens* (coordonné par G. Lecointre et M. Silberstein), Éditions Syllèphe, 2009, p 245-264.

### Le principe de Hardy-Weinberg, hypothèse nulle de l'évolution

Ce principe, enseigné maintenant dans tous les cursus universitaires de biologie des populations, est né au début du XX<sup>e</sup> siècle (1906) et a été énoncé par T. Hardy et H. Weinberg, séparément. Notre formulation reprend pour l'instant une partie seulement de leurs énoncés ; la seconde, de nature différente, sera abordée plus tard.

Supposons une population fermée dans laquelle deux variants génétiques coexistent, notés A et B. On suppose la sélection naturelle absente du système, c'est-à-dire qu'aucun des deux n'est meilleur que l'autre pour se perpétuer. Un variant pris au hasard dans la population doit être soit de type A soit de type B. L'état de la population peut être décrit par les fréquences de A et de B (notées  $p$  et  $q$ ), telles que, si l'on tire un variant au hasard dans la population, il a une chance  $p$  d'être A et  $q$  d'être B. Le principe s'énonce ainsi : Si la population a un effectif infini, et en l'absence de mutations, les fréquences  $p$  et  $q$  restent indéfiniment constantes de génération en génération. Ce principe appelle deux remarques :

1) Aucune population réelle ne répond aux conditions (en particulier, l'effectif infini). Le principe n'a donc pas de valeur empirique immédiate, mais représente un état de référence. Tout changement de  $p$  et  $q$  (et l'évolution d'une population n'est rien d'autre qu'une série de tels changements) devra en effet être expliqué par la violation de l'une ou l'autre des hypothèses de départ. Ce jeu d'hypothèses constitue donc une catégorisation des causes possibles de l'évolution. De même, en physique, aucun objet réel n'est en mouvement parfaitement rectiligne et uniforme, mais cet état est un état de référence et certains objets s'y trouvent approximativement.

2) L'énoncé ci-dessus ne se démontre pas. Il constitue un axiome qui contient implicitement la définition de la sélection naturelle. La sélection naturelle se définit comme le phénomène qui engendre des changements déterministes des fréquences des différents variants génétiques ; en l'absence de sélection, ces changements n'existent pas (les variations aléatoires étant, quant à elles, éliminées par l'hypothèse d'effectif infini).

mécanique newtonienne, les changements de position et de vitesse des objets matériels résultent de l'action de forces qui sont directionnelles et déterministes. Dans la théorie de l'évolution, les changements sont expliqués par deux types d'actions. Le premier type de changements résulte de la *sélection naturelle*, qui est, comme les forces newtoniennes, à la fois directionnelle et déterministe. Ces changements peuvent être prédits exactement, connaissant les différences de valeur sélective entre les individus et l'état initial de la population (de même qu'une trajectoire est prédictive connaissant les forces ainsi que la position et la vitesse initiale). Le deuxième type est la *dérive génétique*, c'est-à-dire des effets aléatoires liés à l'effectif fini (voire faible) des populations.

Enfin, il convient de décrire mathématiquement le lien entre les changements et les facteurs à l'origine des changements. Pour la mécanique newtonienne, la force est reliée au mouvement de l'objet matériel par l'équation  $F = mA$ , où  $F$  mesure la force,  $m$  la masse de l'objet et  $A$  son accélération. La définition du concept de force est donnée implicitement par son rapport avec le concept d'accélération. Dans le cas de la théorie de l'évolution, l'action de la sélection naturelle (dans une population asexuée, et en l'absence de dérive pour simplifier) s'exprime par la relation :

$$\frac{dp}{dt} = p (w_A - W)/W$$

où  $p$  est la fréquence du type A dans une population,  $dp/dt$  le changement de cette fréquence au cours du temps,  $w_A$  la valeur sélective associée au type A, et  $W$  une constante égale à la valeur sélective moyenne de la population. Cette formule signifie simplement que la fréquence du type A au sein de la population augmentera

d'autant plus que sa valeur sélective est supérieure à celle de la moyenne de la population. La définition de la valeur sélective est, comme dans le cas de la force, implicite : c'est la mesure qui prédit la propagation ou la régression d'un type d'individus dans une population, comme la force est la mesure qui prédit les mouvements d'un objet.

La comparaison entre la théorie newtonienne et la théorie de l'évolution a bien entendu des limites. La valeur sélective n'agit pas directement (au contraire de la force), ce sont les différences de valeur sélective entre individus qui font la sélection naturelle (qui est donc un concept relatif). Par ailleurs, la dérive génétique n'a pas d'équivalent dans les axiomes de la théorie newtonienne. Dans la théorie de l'évolution, les changements s'expriment en termes probabilistes ou statistiques dès que la dérive intervient. La prédiction de l'état de la population à l'instant  $t$  sera alors accompagnée d'une probabilité. Cependant, cet aspect probabiliste ne diminue pas le pouvoir prédictif de la théorie. On verra (chapitre 3) qu'il existe des analogies profondes entre le phénomène de dérive et des phénomènes physiques (non décrits par la mécanique newtonienne), tels que la diffusion d'un gaz, également étudiés suivant une approche statistique.

Les deux théories reposent sur des postulats (les hypothèses nulles) et sur la description des objets grâce à des variables pertinentes. Certaines sont mesurables directement, comme le mouvement ou le changement de fréquences des types biologiques, alors que d'autres sont implicites, comme la force ou la valeur sélective. Les variables implicites permettent de modéliser et de prédire le comportement des variables observables. Le fait de définir la force par son rapport avec l'accélération, et la

valeur sélective par son rapport avec le changement des fréquences des génotypes, peut paraître circulaire. Cette circularité des postulats de base fait d'ailleurs partie des critiques généralement émises par les créationnistes contre l'évolution. Elle ne pose en réalité pas de problème, puisque l'objectif d'une théorie scientifique est de prédire efficacement les variables mesurables tout en donnant une vision cohérente du fonctionnement du système. La force ou la valeur sélective sont des descripteurs inventés par l'homme, qui n'ont pas d'existence matérielle, mais qui sont utiles pour comprendre et prédir le comportement des objets matériels.

## Conclusion

La théorie de l'évolution est donc une théorie scientifique au même titre que les théories physiques.

**Elle apporte une unité à la biologie en définissant les objets biologiques comme les éléments d'une histoire de filiation (une généalogie).**

**Elle justifie les autres programmes de recherche de la biologie.** Comme l'a remarqué le généticien Theodosius Dobzhansky, « rien en biologie n'a de sens si ce n'est à la lumière de l'évolution ». En effet, les branches de la biologie, autres que l'étude de l'évolution elle-même, se rapportent essentiellement à deux activités :

La première activité (*la systématique*) consiste à décrire et classer les formes vivantes en catégories emboîtées telles que des espèces, des genres, des familles (voir glossaire critique, p. 322). La théorie de l'évolution justifie cette activité dans la mesure où la classification reflète les relations de parenté qui relient entre elles les espèces vivantes (*la phylogénie*).

La deuxième activité consiste à élucider des mécanismes de fonctionnement des êtres vivants. On répond dans ce cadre à deux types de questions : « comment ça marche ? » (par exemple, comment le cœur bat-il ?) et « à quoi ça sert ? » (à quoi servent les battements du cœur ?). La deuxième question introduit le concept de *fonction* (ici, la fonction du cœur) qui sert de fil conducteur pour répondre à la première question. Dans notre exemple, la fonction du cœur est de propulser le sang vers les organes. On recherchera donc les mécanismes par lesquels les battements du cœur s'adaptent aux variations des besoins des organes. Cette démarche est pratiquée non seulement en physiologie, mais aussi en embryologie, en éthologie, etc. La théorie de l'évolution valide cette démarche car elle justifie l'existence de fonctions. Cette fonctionnalité des systèmes vivants était auparavant interprétée comme le dessein du Créateur. On y voit maintenant le produit de la sélection naturelle, qui optimise peu à peu les systèmes vivants en maximisant leur survie et leur reproduction.

Une des forces de la théorie de l'évolution est de prédire, contrairement à la vision créationniste, que les êtres vivants ont des qualités fonctionnelles mais aussi des défauts. En effet, la sélection naturelle est un processus historique. L'adaptation d'un organisme à son milieu est dès lors nécessairement imparfaite, reflet des événements contingents qui ont eu lieu dans le passé comme les phénomènes aléatoires de dérive génétique. Il est donc tout à fait concevable d'observer des organes sans fonction, soit parce qu'ils se sont fixés par hasard (dérive), soit parce qu'ils ont eu autrefois une fonction qu'ils ont perdue. L'existence de tels organes résiduels (comme le coccyx, qui représente un résidu de queue dans le squelette de l'homme) était d'ailleurs un des arguments

utilisés par Darwin pour soutenir l'évolution. En conclusion, la théorie de l'évolution justifie qu'on utilise la notion de fonction comme guide pour comprendre les systèmes vivants (façonnés par le travail de la sélection naturelle) tout en relativisant cette notion (replacée dans le contexte d'une histoire et de ses contingences).

**La théorie de l'évolution clarifie les relations causales utilisées en biologie.** Pour expliciter cette idée, reprenons l'exemple du cœur. Un physiologiste et un généticien évolutionniste ne font pas appel aux mêmes causes pour expliquer les battements du cœur. Le physiologiste fait appel aux mécanismes moléculaires de contraction des cellules du myocarde, et à leur stimulation électrique par d'autres cellules. L'évolutionniste cherche à comprendre en quoi les battements du cœur aident un individu à survivre ou à se reproduire, c'est-à-dire à transmettre son information héréditaire. Dans ce cas, la réponse est triviale car un individu dont le cœur ne bat pas meurt rapidement. La théorie de l'évolution permet de distinguer ces deux types de causes. La *cause ultime* d'un phénomène est la raison pour laquelle il a été favorisé par la sélection naturelle, mais sa *cause proximale* est un ensemble de mécanismes qui déclenchent sa réalisation.

**La théorie de l'évolution, qui apporte une cohérence logique à la biologie, n'est pas une théorie figée.** Au milieu du XIX<sup>e</sup> siècle, Darwin a fourni sa première formulation complète, posant le cadre conceptuel de la théorie actuelle. Il proposait ainsi un vaste programme de recherche, qui s'est développé jusqu'à nos jours. La naissance de la génétique moderne, avec la redécouverte du travail de Mendel au début du XX<sup>e</sup> siècle, loin de remettre en cause ce programme, en fortifie les bases. Les décennies suivantes ont vu éclore la « théorie synthétique de l'évolution », grâce à la mise en commun des

avancées obtenues dans des domaines variés (voir glossaire critique, p. 338). Les principaux artisans de cette synthèse furent Theodosius Dobzhansky (génétique), Ernst Mayr (biogéographie) et George Simpson (paléontologie). Aujourd'hui, la recherche sur les mécanismes de l'évolution est plus active que jamais. Elle s'enrichit en permanence des progrès des autres branches de la biologie (notamment la biologie moléculaire, la biologie du développement et l'écologie). Nous tenterons dans cet ouvrage d'exposer l'état actuel de la théorie de l'évolution.

## LES ORIGINES MOLÉCULAIRES DE LA VARIATION GÉNÉTIQUE

Pour que les mécanismes de tri puissent façonnner la biodiversité au cours de l'histoire de la vie, il faut que les organismes varient et que ces variations soient transmises d'une génération à l'autre. Ces événements de variation héritable sont appelés *mutations*. Les premières mutations ont été identifiées par leurs effets sur la morphologie. Ainsi, parmi une population de drosophiles aux yeux rouges (couleur habituelle pour cette espèce), peuvent apparaître des mutants aux yeux blancs qui, croisés entre eux, ont des descendants aux yeux blancs.

Les mutations n'apparaissent pas en réponse aux besoins de l'organisme, elles sont rares et aléatoires (voir *Caractères acquis ou mutations aléatoires ?*, p. 328). Quels sont donc les mécanismes qui sont à l'origine des mutations ? Le support physique de l'information génétique est aujourd'hui bien connu : c'est une macromolécule nommée ADN (acide désoxyribonucléique) présente dans toutes les cellules. L'ADN est recopié à l'identique et transmis de cellule en cellule lors des divisions cellulaires. Les mutations, qui sont à l'origine de la diversité des organismes vivants, sont donc à rechercher dans les variations de cette molécule. Dans un premier temps, la structure de l'ADN sera examinée, ainsi que la façon

dont il détermine les caractères. Nous nous demanderons ensuite comment les variations de l'ADN, qu'elles soient ponctuelles ou de plus grande ampleur, fournissent un matériau sur lequel la sélection naturelle peut jouer.

## L'ADN et le code génétique

**Comment la molécule d'ADN détermine-t-elle les caractères visibles de l'organisme ?** L'ADN n'agit pas directement sur les caractères : ainsi, la couleur rouge ou bleue d'une fleur n'est pas liée à la couleur rouge ou bleue de la molécule d'ADN de ses cellules, mais aux pigments rouges ou bleus qu'elles synthétisent. Cette synthèse est effectuée par une famille de macromolécules biologiques, les protéines. D'une fleur à l'autre, les protéines chargées de la synthèse des pigments floraux diffèrent : dans l'une, elles fabriquent du pigment rouge, dans l'autre du bleu. De la même façon, tous les caractères observables sont dus à l'action des protéines, qui constituent en quelque sorte les ouvriers chimiques des cellules. Pour déterminer les caractères observables, l'ADN doit donc contrôler la formation des protéines.

L'ADN et les protéines sont des polymères composés de nombreuses molécules élémentaires qui s'enchaînent linéairement comme des perles sur un collier. Dans le cas de l'ADN, ces perles sont appelées *nucléotides* et sont de quatre types différents (adénine, cytosine, thymine, guanine, abrégées en A, C, T, G). Dans le cas d'une protéine, ce sont des acides aminés appartenant à vingt types différents. La fonction de la protéine ne dépend que de l'enchaînement linéaire (*séquence*) des acides aminés. En effet, une séquence donnée détermine une conformation spatiale très précise permettant la réalisation de la

fonction. Par exemple, l'hémoglobine est un complexe protéique dont les repliements ménagent des poches où se fixent l'oxygène et le gaz carbonique. Cela lui permet d'assurer le transport de ces deux gaz dans le sang. L'enchaînement des nucléotides dans la molécule d'ADN forme également une séquence. Sur un fragment d'ADN, la séquence des nucléotides forme des phrases écrites dans un alphabet rudimentaire à quatre lettres (A, C, T, G). Chaque phrase représente la suite d'instructions, appelée *gène*, permettant de construire une protéine.

**Comment passer d'une séquence d'ADN à une séquence protéique ?** L'ADN peut être comparé au code source d'un programme informatique qui peut être compilé puis exécuté. La compilation consiste à rendre lisible par la machine une série de commandes, écrites dans un langage particulier. Le programme pourra alors être exécuté. Dans le cas d'un organisme vivant, la compilation a pour équivalent la *transcription*. Au cours de cette étape un fragment d'ADN est recopié en une molécule d'ARN (acide ribonucléique). L'ARN est un polymère de nucléotides très semblable à une molécule d'ADN, mais plus petit. Contrairement à l'ADN, l'ARN est lisible par la cellule. Au cours d'une étape de *traduction*, la séquence de la molécule d'ARN est décodée et permet l'assemblage d'acides aminés dans un ordre précis, formant une protéine.

Comment coder les vingt acides aminés des protéines avec les quatre nucléotides de l'ADN et de l'ARN ? Le codage se fait à la façon du « Morse », dans lequel deux symboles (point, trait) sont combinés pour exprimer les vingt-six lettres de l'alphabet. Par exemple, la lettre S est codée par trois points successifs. Dans notre cas, chaque acide aminé est codé par une série de trois nucléotides

successifs de l'ADN (un *codon*). Le décryptage du code génétique (correspondance codon-acide aminé) a été effectué dans les années 1960 chez la bactérie *Escherichia coli*. Ce code s'est ensuite révélé identique chez presque tous les êtres vivants. À partir de quatre nucléotides, il y a  $4^3 = 64$  codons possibles alors qu'il n'y a que vingt acides aminés à coder. Cependant, les soixante-quatre codons existent bien dans les molécules d'ADN et sont tous utilisés, car certains acides aminés peuvent être codés par plusieurs codons différents (dits *synonymes*). Le code génétique est donc *redondant*. En outre, trois codons sont utilisés pour la ponctuation. Ces codons, dits *codons-stop*, signalent la fin de la séquence à traduire, qui correspond à l'extrémité de la protéine.

### Mutation ponctuelle

Un premier mode de variation de l'ADN est le changement d'un nucléotide en un autre (par exemple, un C devient A). Ces mutations sont dites *ponctuelles* puisqu'elles n'affectent qu'un endroit de la molécule d'ADN. Les quatre nucléotides qui composent la molécule d'ADN sont chimiquement très proches. Le changement d'un nucléotide en un autre, ou *substitution*, peut avoir lieu spontanément dans des conditions physico-chimiques particulières, par exemple en présence de rayonnements ultraviolets. La cellule possède une machinerie permettant de réparer ces « avaries », mais cette machinerie peut être mise en défaut. Une autre source d'erreurs réside dans le mécanisme de réPLICATION de l'ADN qui précède la division cellulaire. Des erreurs de copie peuvent être commises. Un nucléotide est alors incorporé au lieu d'un autre sur l'ADN copié.

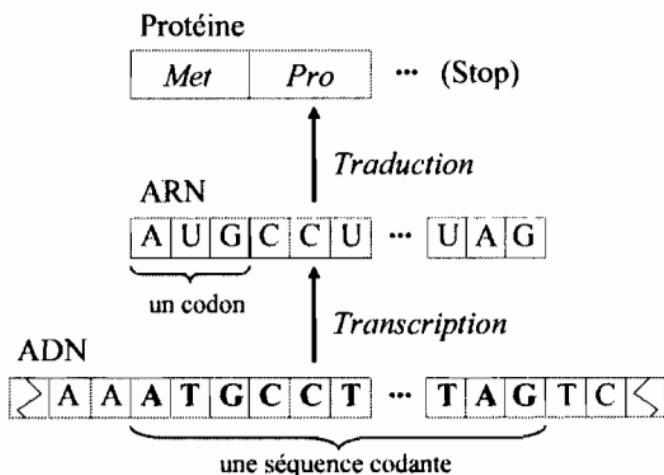


Figure 1 : Schéma des étapes du codage, de l'ADN à la protéine. Les petits carrés représentent les nucléotides et les rectangles, des acides aminés. Un fragment d'ADN est transcrit en un ARN de séquence identique (la thymine T est simplement remplacée par un uracile U). L'ARN est ensuite traduit en protéine, chaque codon correspondant à un acide aminé. Ce schéma résume le principe du codage sans respecter les structures moléculaires (notamment, la structure double-brin de l'ADN) que l'on peut trouver dans tout bon manuel de biologie.

**Quelles sont les conséquences des mutations au niveau de la cellule ?** Le code génétique est redondant : en conséquence, de nombreuses substitutions ne modifient pas la protéine codée puisque la séquence en acides aminés reste la même (*mutations synonymes*). Les trois nucléotides de chaque codon ne sont pas équivalents de ce point de vue. En effet, 70 % des substitutions en troisième position sont synonymes alors qu'aucun changement en seconde position et seulement 4 % des changements en première sont synonymes.

En revanche, si la mutation change la séquence en acides aminés, on parle de mutation *non synonyme*. La

plupart des acides aminés peuvent être substitués sans conséquences importantes sur les propriétés biochimiques de la protéine. Cependant, certains acides aminés ont un rôle clé. Le plus souvent, leur modification provoque une déficience fonctionnelle. Plus rarement, elle améliore son fonctionnement ou lui permet de réaliser une nouvelle fonction.

Beaucoup de maladies génétiques sont des déficiences fonctionnelles dues à des mutations ponctuelles. Ainsi, chez les personnes atteintes d'*anémie falciforme*, le gène codant pour la chaîne protéique  $\beta$  de l'hémoglobine est muté au niveau d'un acide aminé. Cette modification a pour conséquence une mauvaise fixation de l'oxygène par les globules rouges. Cette déficience est telle qu'un sujet ne produisant que des protéines mutées meurt avant d'atteindre l'âge adulte. Toutefois, certains sujets possèdent à la fois des protéines normales et des protéines mutées car ils ont reçu de leur père un gène normal et de leur mère un gène muté (ou inversement). Ces individus ne sont que modérément affectés par la maladie, mais montrent par ailleurs une résistance inattendue à un parasite des cellules sanguines (agent du paludisme). La modification d'un acide aminé dans la protéine a donc détérioré sa fonction première (fixer l'oxygène) mais lui a aussi apporté une nouvelle fonction (la résistance à un parasite). Les mutations, en modifiant la fonction des gènes, créent donc le domaine de variation sur lequel s'opère le tri par la sélection naturelle (voir chapitre 4).

**Les mutations non synonymes sont-elles les seules importantes pour la sélection naturelle ?** Non, car le codage de l'information génétique est plus complexe qu'une traduction automatique de nucléotides en acides aminés. Cette complexité se manifeste notamment par le

biais d'usage du code, c'est-à-dire l'usage préférentiel d'un codon parmi plusieurs synonymes. Dans les langues parlées, il existe souvent plusieurs façons de dire la même chose, mais l'une d'elles est employée beaucoup plus souvent que les autres. En outre, les différents synonymes, comme « bagnole », « voiture » et « automobile » seront utilisés dans des contextes différents ou par des personnes différentes. Il en va de même avec les différents codons synonymes.

Quelle est la cause de ce biais ? La traduction de l'ARN messager en protéine est assurée par une machinerie cellulaire qui n'a pas la même efficacité selon les codons utilisés. Cette machinerie comporte des molécules appelées *ARN de transfert* (ARN<sub>t</sub>), qui jouent le rôle d'adaptateurs permettant l'assemblage des acides aminés. À chaque codon s'adapte un unique ARN de transfert et réciproquement. Il existe donc, pour chaque acide aminé, autant d'ARN<sub>t</sub> différents que de codons synonymes. Bien que tous les ARN<sub>t</sub> existent chez tous les organismes, tous ne sont pas produits dans les mêmes proportions dans la cellule. En général, parmi les ARN<sub>t</sub> correspondant à un acide aminé donné, l'un est très abondant et les autres rares. La présence dans un gène d'un codon dont l'ARN<sub>t</sub> est rare ralentit la fabrication de la protéine codée. Pour les protéines qui doivent être produites en grande quantité, il serait donc intéressant de ne pas utiliser indifféremment tous les codons synonymes. Il vaudrait mieux utiliser en priorité les codons correspondant aux ARN<sub>t</sub> les plus fréquents. On constate en effet que ce biais est particulièrement fort pour les protéines produites en grandes quantités. Bien que le code génétique soit universel, le biais dans son usage varie d'un organisme à l'autre car les ARN<sub>t</sub> les plus fréquents ne sont pas toujours les mêmes. Les mutations

synonymes ont donc des conséquences fonctionnelles, même si elles sont plus subtiles que celles des mutations non synonymes. Elles peuvent alors être soumises à sélection naturelle. Cette interprétation du biais d'usage du code est néanmoins sujette à débat.

**L'accumulation des mutations au cours de l'évolution.** Au cours de l'histoire de la vie, représentée sous la forme d'un arbre phylogénétique, des mutations apparaissent et s'accumulent dans chaque lignée indépendamment. Les différences entre deux organismes proviennent donc des mutations qui se sont accumulées dans les deux lignées depuis leur dernier ancêtre commun (figure 2), c'est-à-dire depuis leur séparation. Si les deux organismes sont proches parents, cette divergence est récente. En moyenne, peu de mutations ont donc eu le temps de s'accumuler. À l'inverse, si les deux organismes sont éloignés dans la phylogénie, beaucoup de mutations ont pu s'accumuler. En comptant les différences entre les séquences d'une même protéine chez plusieurs organismes d'espèces différentes, on peut repérer les mutations intervenues. Cela permet de déterminer l'ordre de succession des divergences dans le temps et de reconstruire la phylogénie (figure 2, légende A). Cependant, la reconstruction phylogénétique est nécessairement imparfaite. Comme la mutation est un phénomène aléatoire, le nombre de mutations accumulées sur un fragment en un temps donné peut varier autour de la valeur moyenne attendue. Cette source d'erreur est d'autant plus importante que le fragment observé est petit et le temps laissé aux mutations pour s'accumuler court. Or, techniquement, on ne peut analyser que des fragments d'ADN relativement petits. Pour utiliser le mieux possible les données disponibles, plusieurs méthodes de reconstruction phylogénétique sont aujourd'hui utilisées

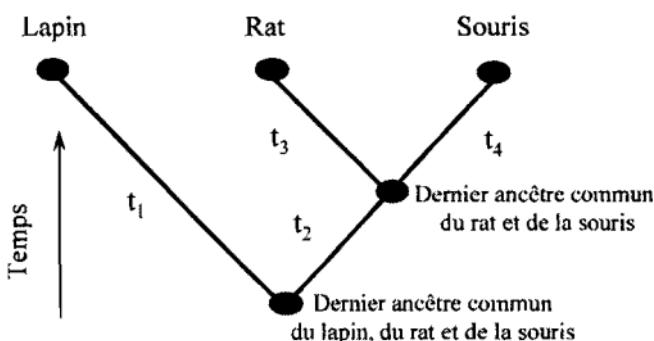
après avoir été âprement débattues (voir *Cladisme et phénéticisme*, p. 329).

En 1965, Emil Zuckerkandl et Linus Pauling<sup>1</sup> établirent la séquence de l'hémoglobine chez plusieurs vertébrés, dont des mammifères comme l'homme et le cheval, un oiseau, un amphibiens, un reptile et des poissons. Pour chaque paire d'espèces, ils calculèrent le nombre d'acides aminés différents entre les deux hémoglobines, ou *taux de substitution*. Par ailleurs, ils utilisèrent les données paléontologiques pour connaître la date approximative de divergence des deux espèces, en millions d'années. Ils firent ainsi une découverte surprenante. Aux imprécisions expérimentales près, les taux de substitution étaient proportionnels aux dates de divergence. Si ce phénomène était général, il fournirait un moyen de dater précisément les divergences en utilisant l'accumulation des mutations de la même façon que l'on date les couches géologiques en mesurant la décomposition des éléments radioactifs. C'est pourquoi cette hypothèse fut qualifiée d'*horloge moléculaire*. Elle implique que l'apparition de nouvelles mutations est un phénomène régulier, qui intervient avec un taux constant au cours du temps, ainsi qu'entre des lignées différentes. L'expérience de Zuckerkandl et Pauling était à l'époque une prouesse technique, car le séquençage direct des protéines était laborieux. Les nombreuses séquences obtenues depuis lors, grâce aux techniques de séquençage direct de l'ADN, ont permis de discuter de la validité de l'*horloge moléculaire*.

Dans ce cadre, un vaste débat à propos de la date de divergence entre l'homme et les grands singes agita la

---

1. E. Zuckerkandl et L. Pauling, « Evolutionary Divergence and Convergence in Proteins », *Evolving Genes and Proteins*, New York, Bryson et Vogel, Academie Press, 1965, p. 97-166.



### Séquences alignées des trois espèces

Lapin G T GGCAA C GTA G CTT A CTT A TGCC T

Rat G A GGCAA G GTA G CTT T CTT T TGCC A

Souris G A GGCAA G GTA C CTT A CTT T TGCC A

Figure 2 : Arbre représentant la phylogénie du lapin, du rat et de la souris.

(A) La construction de l'arbre part du principe que les séquences des espèces proches (ici, le rat et la souris) se ressemblent plus entre elles qu'elles ne ressemblent aux autres (le lapin). Un extrait de séquence est donné à titre d'exemple, montrant en gras les nucléotides différents d'une espèce à l'autre.

(B) Si l'hypothèse de l'horloge moléculaire est vraie, c'est-à-dire si les mutations s'accumulent à un taux constant le long des différentes branches, alors les nombres de mutations qui séparent le rat et la souris actuels de leur ancêtre commun doivent être égaux ( $t_3 - t_4 = 0$ ). On ne peut pas évaluer directement ( $t_3 - t_4$ ). En revanche, le nombre de différences entre les séquences de rat et de souris permet d'estimer ( $t_3 + t_4$ ). Pour tester indirectement l'hypothèse «  $(t_3 - t_4 = 0)$  », on utilise le lapin, dont la divergence est sans ambiguïté antérieure à celle du rat et de la souris. Le nombre de mutations qui sépare le lapin du rat ( $T = t_1 + t_2 + t_3$ ) doit être égal à celui qui sépare le lapin de la souris ( $T' = t_1 + t_2 + t_4$ ). Si  $T = T'$ , alors  $(t_1 + t_2 + t_3) = (t_1 + t_2 + t_4)$  et donc  $(t_3 - t_4) = 0$ , conformément à l'hypothèse d'horloge moléculaire.

communauté des évolutionnistes<sup>1</sup>. En effet, la date calculée selon l'hypothèse de l'horloge moléculaire était de seulement 5 millions d'années, alors que les données disponibles en paléontologie faisaient remonter la divergence à au moins 15 millions d'années. Pourquoi une telle différence dans ces estimations ? Un test de l'horloge moléculaire, indépendant des données paléontologiques, fut proposé. Son principe est illustré dans la figure 2 sur l'exemple du rat et de la souris. Pour une douzaine de protéines différentes, le résultat du test fut positif pour ces deux espèces : les taux de substitution étaient identiques entre les lignées du rat et de la souris. Malheureusement, une expérience similaire sur l'homme et les grands singes montra que le taux de substitution était en moyenne inférieur dans la lignée humaine. De fait, la lignée humaine accumule moins de mutations que les autres lignées de primates. Le temps de divergence homme-singes, calculé sur la base du taux de substitution des primates, avait donc été sous-estimé. Cependant, depuis cette époque, la découverte de nouveaux fossiles par les paléontologues leur a permis de réviser à la baisse leur propre estimation de date de divergence, qu'ils situent maintenant dans un intervalle entre 13 millions d'années (*Pierolapithecus catalaunicus*) et (plus probablement) 7 millions d'années (*Sahelanthropus tchadensis* dit Toumai)<sup>2</sup>. Les écarts entre les deux méthodes (paléontologique et moléculaire) se sont donc réduits des deux côtés.

Suivant les lignées comparées, l'hypothèse de l'horloge moléculaire est donc validée ou infirmée. Quelle est la cause de ces différences de taux de substitution ? Une

1. W.-H. Li, *Molecular Evolution*, Sunderland, Massachusetts, Sinauer Associates, 1997, p. 215-235.

2. P. Picq, *Au commencement était l'homme*, Paris, Odile Jacob, 2009.

explication peut être trouvée dans les différences de nombre de générations qui se succèdent par unité de temps. Une espèce qui se reproduit à l'âge de quatre ans fait vingt-cinq générations en un siècle, alors qu'une espèce se reproduisant à vingt ans ne fait que cinq générations dans le même temps. La première espèce accumulera donc, en un siècle, plus de mutations que la deuxième. Ce raisonnement repose sur l'hypothèse que le nombre de mutations accumulées en une génération est le même pour les deux espèces. Conformément à cette hypothèse, le nombre de mutations ne dépend que du nombre de divisions dans la lignée cellulaire qui donne les ovules ou les spermatozoïdes transmis à la génération suivante. Les données suggèrent que ce nombre de divisions est le même entre l'homme et les singes. Ainsi, l'homme, qui se reproduit plus tard que les autres singes, a un taux d'évolution moléculaire plus lent.

En plus de l'influence du temps de génération, beaucoup d'autres phénomènes peuvent faire varier localement le taux de substitution. Ces problèmes rendent l'hypothèse d'horloge moléculaire inutilisable pour certaines datations précises. Néanmoins, ils ne remettent pas en question le principe général d'une augmentation de divergence au cours du temps.

### Mutation par fragment

**La mutation peut affecter un fragment d'ADN de plusieurs nucléotides** (figure 3). Des nucléotides peuvent être soustraits (*délétion*) ajoutés (*insertion*), copiés (*duplication*) ou encore échangés (*recombinaison*). Comme pour les mutations ponctuelles, la majeure partie de ces muta-

tions a pour origine des erreurs commises lors de la réPLICATION ou de la réPARATION de l'ADN.

La mutation par fragment permet-elle l'innovation au cours de l'évolution ? Suivant la taille du fragment affecté, mais aussi le type d'événement, les conséquences sur l'information génétique seront extrêmement variées. Les mutations concernant des fragments à l'intérieur d'un gène auront des conséquences différentes selon que leur longueur est, ou non, multiple de trois. En effet, le premier acide aminé de la protéine est codé par les nucléotides 1 à 3 du gène, le second par les nucléotides 4 à 6 et ainsi de suite. Si une délétion supprime les deux premiers nucléotides, la protéine obtenue commencera par l'acide aminé correspondant aux nucléotides 3 à 5 puis 6 à 8... La protéine obtenue sera complètement différente de la protéine originale en raison d'un décalage du *cadre de lecture*. Il en va de même pour des insertions non multiples de trois. La modification du cadre de lecture aura pour conséquence une déficience complète dans la fonction codée. En revanche, une insertion ou une délétion ayant une taille multiple de trois ne fera qu'ajouter ou retrancher des acides aminés à la protéine. Elle pourra, suivant les acides aminés affectés, ne modifier que modérément la fonction. La variabilité ainsi produite sera soumise ensuite au tri par le hasard ou la sélection naturelle.

Les délétions affectant des fragments d'ADN plus grands (un gène complet ou une grande partie d'un gène) ont le plus souvent comme conséquence une déficience complète dans la fonction. Dans la plupart des cas, de telles délétions sont très graves pour la survie de l'organisme (la mutation conférant la myopathie de Duchenne en est un exemple). Dans le meilleur des

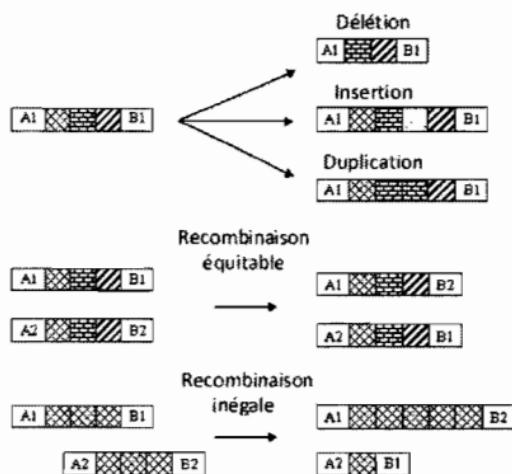


Figure 3 : Mutations par fragments. Les différents rectangles correspondent à des segments d'ADN dont les séquences sont schématisées par des figurés différents. Les mutations de type délétion, insertion et duplication se font à partir d'une seule molécule. Les recombinaisons se font à partir de deux molécules, au niveau de fragments de séquences semblables. Dans le cas d'une recombinaison équitable, il y a échange des extrémités B1 et B2, mais pas de changement de longueur des molécules d'ADN. La présence d'un fragment répété permet un appariement décalé qui conduit à une recombinaison inégale. Ce mécanisme conduit à augmenter la longueur d'une des molécules et à diminuer la longueur de la seconde.

cas, l'absence de la fonction est suppléeée par d'autres gènes et diminue peu les chances de survie de l'individu.

La duplication de fragments d'ADN est un mécanisme de création de variabilité. À la suite de la duplication d'un gène complet, les mutations ponctuelles s'accumulent sur deux fois plus de longueur d'ADN. Cette accélération dure tant que les mutations n'altèrent pas la fonction des deux copies. Par la suite, deux évolutions sont possibles. Dans le premier cas (le plus fréquent), des mutations

délétères apparaissent dans l'une des deux copies du gène. La copie mutée devient inopérante, et se transforme en un *pseudogène*. Par définition, le pseudogène est un fragment d'ADN dont la séquence ressemble à une séquence codante pour une protéine, mais qui n'est plus fonctionnel. Il ne présente ni coût ni bénéfice pour l'organisme puisque la seconde copie du gène est intacte et fonctionne normalement.

Une deuxième évolution possible (moins fréquente) a lieu lorsqu'une mutation modifie la fonction d'une des deux copies du gène. L'organisme ne perd pas pour autant l'ancienne fonction qui est conservée sur la seconde copie. Ce mutant aura ainsi acquis une fonction supplémentaire. Si cette innovation améliore la survie ou la reproduction de l'organisme, la sélection naturelle favorisera cet organisme au détriment des autres. La duplication permet donc d'innover puisqu'elle augmente l'espace de variabilité.

La répétition d'événements de duplication au cours de l'évolution aboutit à la construction de *familles de gènes* (ensembles de gènes issus de duplications à partir d'un même gène ancestral, reconnaissables à des similarités de séquences). C'est par exemple le cas de la famille des globines<sup>1</sup>. Les différentes séquences de globines ont soit gardé la fonction originale, soit acquis une nouvelle fonction, ou encore perdu toute fonction suite à des mutations délétères (pseudogènes). Cette famille comporte deux sous-familles de gènes fonctionnels : les myoglobines et les hémoglobines, issues de la duplication d'une seule globine ancestrale il y a six cents à huit cents millions d'années. Ces deux sous-familles se sont spécialisées au sein de l'organisme. La myoglobine a une très

---

1. W.-H. Li, *op. cit.*, p. 289-292.

forte affinité pour l'oxygène qu'elle stocke au niveau des muscles, grands consommateurs d'oxygène. En revanche, l'hémoglobine, située dans les globules rouges du sang, a un rôle de transporteur. Elle doit prendre en charge l'oxygène à un endroit (poumon) pour le distribuer à un autre (organes consommateurs). Son affinité pour l'oxygène est moins forte que celle de la myoglobine, ce qui permet le passage de l'oxygène du sang aux muscles.

Dans la sous-famille des hémoglobines, des duplications postérieures à la divergence myoglobine/hémoglobine ont donné naissance à plusieurs gènes, dénotés par des lettres grecques (globines  $\alpha$ ,  $\beta$ , etc.). Plusieurs globines doivent s'associer pour faire une hémoglobine fonctionnelle. Dans les globules rouges de l'adulte, on trouve des complexes formés de deux globines  $\alpha$  et deux globines  $\beta$ . Cette association entre deux types de globines permet une régulation fine de la prise en charge de l'oxygène. En effet, grâce à l'interaction entre les deux types de globines, l'affinité pour l'oxygène change en fonction des conditions, de sorte à s'adapter exactement aux besoins de l'organisme. D'autres globines ( $\delta$ ,  $\epsilon$ ) sont utilisées chez le fœtus. Elles confèrent à l'hémoglobine fœtale une affinité pour l'oxygène plus élevée que celle de l'hémoglobine adulte produite par sa mère, assurant ainsi le transfert de l'oxygène du sang de la mère à celui du fœtus au travers du placenta. Cet exemple montre que la duplication de gènes permet une diversification adaptative des fonctions, par spécialisation des protéines en fonction de l'organe et de l'âge où elles interviennent.

**Comment faire du neuf avec du vieux ?** Une des découvertes les plus surprenantes de la biologie moléculaire est que la plupart des gènes (sauf chez les bactéries) ne sont pas d'un seul tenant. Ils ont une structure en mosaïque (figure 4). Par exemple, les gènes de globine sont morcelés

en plusieurs fragments (appelés *exons*) entre lesquels s'intercalent des séquences d'ADN non codant (appelées *introns*). Seuls les exons sont traduits en acides aminés. L'expression d'un gène morcelé inclut donc une étape supplémentaire : la *maturat<sup>i</sup>on* de l'ARN messager (ARN<sub>m</sub>). L'ARN issu de la transcription est une copie conforme de l'ADN, exons et introns compris. Les introns sont ensuite éliminés, ce qui raccourcit la molécule d'ARN.

Pourquoi les gènes sont-ils morcelés ? Une hypothèse explicative considère que chaque exon constituait autrefois un petit gène autonome codant pour une protéine active. Différents exons auraient été ensuite associés de diverses façons, grâce à des duplications et à des recombinaisons pour produire des protéines composites. Dans ce cadre, les introns représenteraient sur l'ADN des zones d'articulation entre deux exons. Cette hypothèse (réassortiment d'exons ou *exon shuffling*) implique que chaque exon devrait correspondre à un domaine structural et fonctionnel plus ou moins indépendant au sein de la protéine. Afin de tester cette hypothèse, la structure spatiale d'une des globines a été examinée. Si l'hypothèse est juste, chaque « raccord » entre deux domaines de la protéine correspond à un intron sur l'ADN. Dans le cas de l'hémoglobine, cette prédition est relativement bien vérifiée puisque deux des trois « raccords » observés correspondent exactement à l'emplacement des introns dans le gène. Cela tend à montrer que l'association de différentes régions du génome permet de créer de nouvelles fonctions protéiques.

### Des segments d'ADN sans fonction

La taille du génome varie énormément entre organismes. *Drosophila melanogaster* (la mouche du vinaigre)

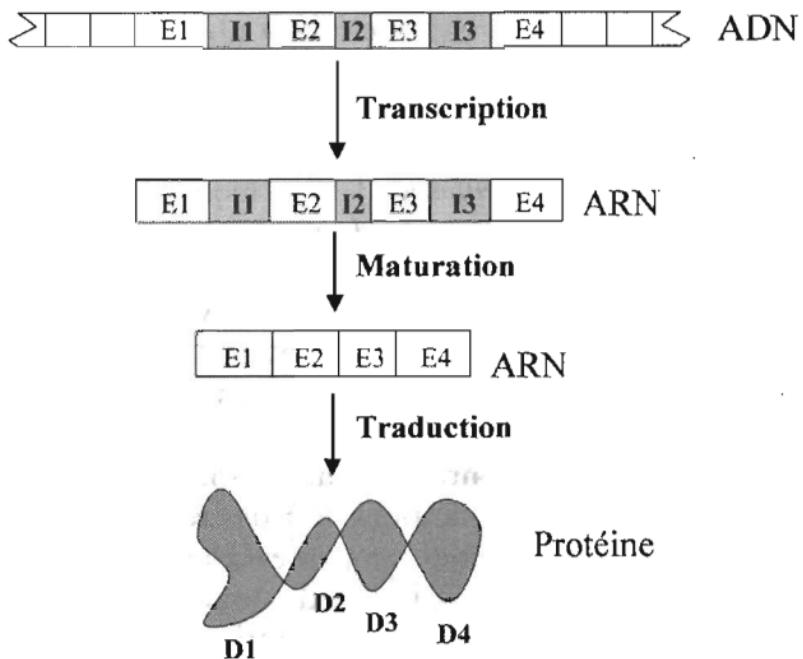


Figure 4 : Expression d'un gène morcelé. E1, E2... correspondent aux exons. I1, I2... aux introns et D1, D2... aux domaines correspondants de la protéine.

possède cent soixante millions ( $1,16 \cdot 10^8$ ) de nucléotides alors que le pin *Pinus resinosa* en possède soixante-huit milliards ( $6,8 \cdot 10^{10}$ ) et l'amibe *Amœba dubia* six cent soixante-dix milliards ( $6,7 \cdot 10^{11}$ ). Même des organismes proches peuvent avoir des tailles de génome très différentes : de  $1,4$  à  $5,7 \cdot 10^9$  chez les mammifères et de  $0,98 \cdot 10^8$  à  $16,9 \cdot 10^9$  nucléotides chez les insectes. Contrairement à l'intuition, la taille du génome ne dépend pas de la taille du corps ni de la complexité structurale des organismes (une amibe est plus petite et a une organisation plus simple qu'une drosophile), mais principalement de la quantité d'ADN non codant. Par exemple, le génome humain contient  $3,4 \cdot 10^9$  nucléotides environ, dont seulement 10 à 20 % sont dans des séquences codantes.

**L'ADN non codant a-t-il une fonction ?** Les ARN de transfert, que nous avons déjà rencontrés, ne sont pas traduits en protéines, mais ont pourtant une fonction. Ils servent d'adaptateurs lors de la fabrication de protéines. Ces ARN sont produits par la transcription d'un segment de l'ADN, non suivie de traduction. Il existe donc dans l'ADN non codant des séquences fonctionnelles servant à produire un ARN directement utilisé par la cellule.

Il existe également dans l'ADN non codant des séquences dites *régulatrices* qui ne sont jamais transcrives, mais qui inhibent ou activent l'expression de gènes voisins. Cette régulation nécessite la fixation spécifique d'une *protéine de contrôle* sur la séquence régulatrice. Les mutations, qu'elles soient ponctuelles ou par fragment, affectent les séquences régulatrices en modifiant leur fonctionnement. Contrairement aux séquences codantes, les effets ne sont pas soumis aux contraintes du code génétique. Les séquences régulatrices ont cependant leurs propres contraintes, liées au maintien de l'interaction chimique avec la protéine de contrôle. Une partie de l'ADN non codant est donc pourvue de fonctions. Elle ne représente néanmoins qu'une faible part de l'ADN d'un organisme. L'essentiel du génome est constitué d'ADN dit *ADN-poubelle*, sans fonction connue.

**Comment se remplissent les poubelles du génome ?** Nous avons déjà rencontré un mécanisme possible : la mutation par duplication, associée à la mutation ponctuelle, produit des pseudogènes non fonctionnels. Mais ce mécanisme est insuffisant car seule une faible fraction de l'ADN non codant est constituée de pseudogènes. Une grande partie du génome est en fait constituée d'*ADN répété*, c'est-à-dire de répétitions de motifs pouvant aller de un à plusieurs centaines de nucléotides.

Certaines de ces séquences ressemblent aux gènes de certains virus, qui intègrent leur ADN en divers endroits de celui des cellules qu'ils infectent, et utilisent la machinerie cellulaire pour se répliquer. Ce faisant, ils finissent tôt ou tard par détruire la cellule. Les séquences répétées, contrairement à ces virus, s'abstiennent de détruire la cellule hôte, mais elles gardent la capacité de se multiplier et de se déplacer dans l'ADN. On appelle ces séquences répétées des *transposons* ou *éléments mobiles*. Elles peuvent s'extraire de l'ADN et se réinsérer à différents points, grâce à des protéines spécialisées. Celles-ci sont parfois codées directement par des gènes inclus dans la séquence répétée, et le transposon contient alors de l'ADN codant et non codant. D'autre fois, les protéines spécialisées sont codées par le reste du génome, et le transposon est entièrement non codant. Contrairement aux virus, les transposons n'ont pas de forme extracellulaire capable d'infester de nouvelles cellules. À titre d'exemple, 80 % du génome du maïs et 15 % du génome de *Drosophila melanogaster* (la mouche du vinaigre) sont composés de séquences d'ADN répétées. Chez la drosophile, la moitié de cet ADN est composé de plusieurs familles de transposons. L'un d'eux, appelé *Copia*, est répété trente à cinquante fois dans chaque individu. L'étude d'individus provenant du monde entier a montré qu'il existait au moins deux cents points d'insertion possibles de *Copia* dans le génome. Les individus varient donc par le nombre et l'emplacement des répétitions de *Copia*. Le mouvement de *Copia* peut provoquer une mutation délétère, s'il s'insère au milieu d'un gène fonctionnel et inactive ce gène. La mutation *white-apricot* qui affecte la couleur des yeux de la drosophile, est due à une telle insertion. Chez l'homme, des cancers peuvent être provoqués par le déplacement d'éléments

mobiles dans des séquences régulatrices. Cela active les gènes voisins de façon inappropriée.

La majeure partie des séquences répétées est constituée non de transposons, mais de motifs beaucoup plus courts répétés un très grand nombre de fois. Par exemple, chez *Dipodomys ordii* (le rat-kangourou), 50 % de l'ADN total est composé de trois séquences, de moins de dix nucléotides chacune, répétées environ deux milliards de fois. Ce type de séquence a été appelé *ADN-satellite*. Le terme est maintenant utilisé de façon générale pour nommer les séquences d'ADN non codant répétées à la queue leu leu (en tandem) le long de la molécule.

Comment les séquences satellites apparaissent-elles et évoluent-elles dans les génomes ? Cette question a suscité plusieurs hypothèses. Certaines supposent que ces séquences dérivent d'éléments mobiles qui auraient perdu leur mobilité. Il est cependant plus probable que les séquences satellites se forment sur place par addition successive de répétitions. En effet, la structure répétée favorise par elle-même les erreurs du système de réPLICATION et de réPARATION de l'ADN. De fait, la répétition favorise les recombinations inégales et les appariements décalés entre brins d'ADN ce qui peut conduire à augmenter le nombre de répétitions (figure 3, *supra*). Dans ce cadre, les séquences répétées se perpétueraient juste en exploitant un défaut des systèmes de réPLICATION, occupant de plus en plus de place dans le génome malgré leur inutilité pour l'organisme. Comme nous le verrons au chapitre 5, ce comportement peut être interprété comme un acte de parasitisme.

## Résumé

L'ADN, support moléculaire de l'information génétique, est à l'origine de la variabilité des organismes. La variabilité entre individus repose sur des changements aléatoires de cette molécule, appelés *mutations*. Ces dernières affectent aussi bien les gènes eux-mêmes (modification entraînant directement un changement dans les caractères de l'individu) que l'organisation de leur expression dans le temps et dans l'espace (la façon dont les informations s'enchaînent pour construire un individu). Les mutations sont souvent destructives mais elles peuvent aussi être créatives. En effet, elles peuvent rajouter au sein du génome de nouveaux fragments d'ADN, qui permettent à l'évolution d'inventer sans cesse de nouvelles fonctions. Toutefois, ces mutations créent également des fragments d'ADN qui s'intègrent au patrimoine génétique bien qu'ils ne soient d'aucune utilité à l'organisme. Le matériel génétique est donc un objet soumis aux lois de l'évolution et en particulier à la sélection naturelle, mais il a également sa propre dynamique liée à son fonctionnement moléculaire.

## 3

## LE TRI ENTRE LES VARIANTS GÉNÉTIQUES PAR LE HASARD

Dans les archives fossiles, les changements semblent souvent brusques : dans une strate, on observera des organismes de type  $A$ , et dans la suivante, des organismes de type  $A'$ . Les variations phénotypiques résultent de différences génétiques. Les organismes de type  $A$  ont-ils tous muté instantanément en  $A'$ ? Non, car les mutations sont des événements rares. Il est plus réaliste de considérer qu'un seul mutant  $A'$  est apparu au sein d'un ensemble de  $A$ , puis que cette mutation a été transmise génétiquement, au fil des générations, à des descendants de plus en plus nombreux, jusqu'à ce que les  $A'$  aient envahi toute la population. On dit que  $A'$  s'est fixé. Entre l'apparition du type  $A'$  et sa fixation existent des états intermédiaires où l'on observe du *polymorphisme*, c'est-à-dire la coexistence dans une population de plusieurs types, ici  $A$  et  $A'$ . Les états *polymorphes* sont transitoires : ils peuvent être enregistrés dans les fossiles mais sont souvent omis, en fonction des caprices de la sédimentation et de la fossilisation.

Quelle variable utiliser pour décrire les changements d'état d'une population ? L'individu n'est pas un niveau de description pertinent car la base de la transmission d'information est le *gène*. Dans ce chapitre, nous négligerons

donc le niveau de l'individu et assimilerons la population à un ensemble de gènes. Le terme de gène est ici pris au sens de *copie physique*, c'est-à-dire d'un fragment concret d'ADN codant. Chaque gène est situé à un emplacement bien précis de la molécule d'ADN, appelé *locus*. En général, tous les gènes qui occupent un locus donné ont la même fonction. Le locus est alors nommé d'après cette fonction (par exemple, locus de la couleur des yeux). — N.B. : on emploie souvent, improprement, le terme de gène pour désigner un locus ; nous chercherons à éviter cette confusion classique.

À quoi correspondent les types *A* et *A'* au niveau des gènes ? Les classes (ou types) de gènes situés à un même locus sont appelées *allèles*. Idéalement ces types sont définis par leur séquence d'ADN : les gènes de type *A* ont une séquence, et les gènes *A'* une autre. Dans la pratique, l'établissement de la séquence d'un gène est lent et coûteux. On utilise donc des catégories moins précises, souvent définies fonctionnellement (par exemple, l'allèle *A* comprend les gènes donnant des yeux sombres, l'allèle *A'* les gènes donnant des yeux clairs). L'évolution d'une population est alors décrite par l'augmentation (ou la diminution) du nombre de gènes représentant un allèle donné dans cette population. Ainsi, pour décrire la progression de l'allèle *A'*, la variable pertinente est la *fréquence* de *A'* dans la population. *A'* étant un allèle, on parle de *fréquence allélique*. Nous recherchons ainsi *quelques mécanismes déterminent les changements de fréquences alléliques au sein d'une population*, et en particulier *ce qui conduit une mutation apparue en très faible fréquence à se fixer*.

Nous avons vu au chapitre 1 que les évolutionnistes font appel, pour répondre à de telles questions, à deux types de mécanismes : aléatoires (*dérive génétique*) et

déterministes (*sélection naturelle*). Nous nous intéresserons ici aux mécanismes aléatoires.

## La mécanique de la dérive génétique

La dérive génétique repose sur un phénomène d'échantillonnage aléatoire de gènes. Illustrons ce phénomène par un exemple. Un enfant a une probabilité  $1/2$  d'être un garçon,  $1/2$  d'être une fille. Parmi quatre enfants, on attend donc en moyenne deux garçons et deux filles. Cependant, on observe souvent des fratries de trois garçons et une fille, ou l'inverse. La raison en est la nature aléatoire du sexe, tiré au hasard pour chaque enfant entre les deux choix possibles. La fratrie est un échantillonnage aléatoire, c'est-à-dire une collection finie (ici, 4) de tirages au hasard. La proportion de garçons dans une fratrie est donc une variable aléatoire. On connaît sa valeur moyenne ( $1/2$ ) mais on ne peut pas prédire avec certitude sa valeur précise.

De même, si, dans une population, un allèle  $A$  est en fréquence  $p$ , la fréquence de cet allèle dans toutes les populations de descendants possibles sera en moyenne  $p$ . Toutefois, les descendants réels ne sont qu'un échantillon parmi tous les possibles. La fréquence réelle de  $A$  parmi les descendants sera alors probablement différente de  $p$ , car c'est une variable aléatoire, au même titre que le nombre de garçons dans une fratrie. La fréquence allélique change donc d'une génération à l'autre, par le simple phénomène d'échantillonnage aléatoire de la population de descendants.

Ce processus se réitère à chaque génération : on passe ainsi d'une fréquence  $p$  à une fréquence  $p'$  puis  $p''$ , etc. La fréquence allélique varie aléatoirement à partir de son

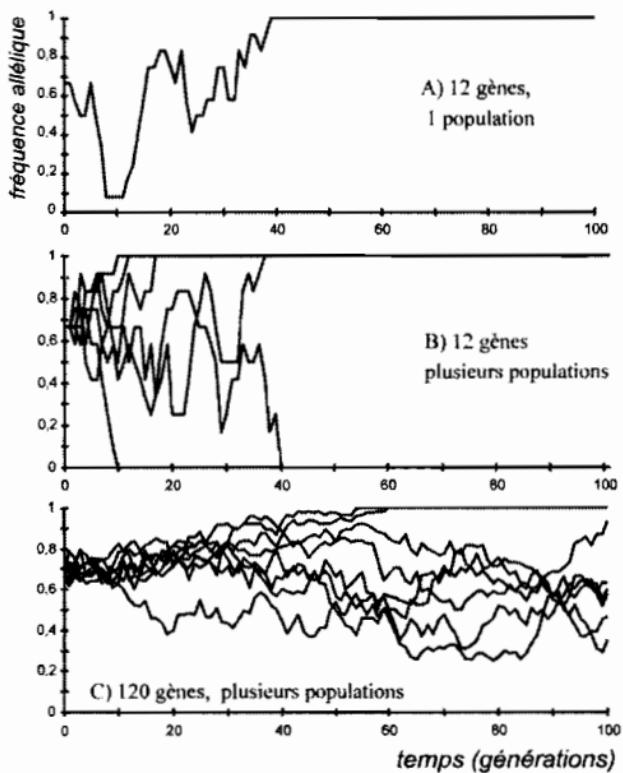


Figure 1 : Simulations du processus de dérive. Le processus d'échantillonnage aléatoire a été simulé dans des populations fictives à l'aide d'un ordinateur. Le panneau A représente une population de douze gènes dans laquelle un allèle *A* est initialement en fréquence 2/3 (huit gènes *A* et quatre gènes non-*A* à la génération zéro). On observe que la fréquence de *A* change à chaque génération sans direction précise. Au bout d'un moment cette fréquence atteint la valeur un, correspondant à une population de douze gènes tous de type *A* : l'allèle *A* s'est fixé. La fréquence ne peut alors plus changer car il n'y a que des *A* dans la population : la fixation met un terme au processus de dérive génétique. Dans le panneau B sont représentées plusieurs populations parties de la même situation de départ (huit *A* sur un total de douze gènes). On observe que la dérive génétique rend les populations, initialement identiques, de plus en plus différentes entre elles. Chacune finit par se fixer soit pour l'allèle *A*

état initial, sans direction précise, d'où le nom de *dérive génétique* (voir les simulations en figure 1). En outre, diverses populations parties du même état initial divergent entre elles.

L'exemple des groupes sanguins A, B, O chez l'homme permet de montrer qu'il existe, dans les populations naturelles, des variations de fréquence allélique dues à la dérive génétique. La figure 2 représente la répartition des groupes sanguins dans les populations humaines. Ces groupes sont déterminés par trois allèles ( $I_A$ ,  $I_B$ ,  $I_O$ ) d'un même locus. Ces trois allèles coexistent au sein de la plupart des populations, qui sont donc *polymorphes* pour ce locus. Cependant, certaines populations, comme celles natives d'Amérique du Sud, sont *monomorphes* : elles contiennent 100 % de O. La fréquence de O varie beaucoup entre les populations, bien que ces dernières dérivent toutes d'une même population ancestrale. La fréquence des groupes sanguins a donc dû, à partir de la population ancestrale, changer dans chaque population, au point qu'elles soient actuellement toutes différentes. La sélection naturelle n'explique pas ces changements : par exemple, aucun argument ne suggère que le groupe sanguin O soit plus favorable en Amérique du Sud qu'en Afrique. Ces variations géographiques dépendent plutôt de l'histoire hasardeuse des migrations et des fondations des différentes populations. À la façon

(fréquence un), soit pour l'autre allèle (fréquence zéro). Dans le panneau C, on part également d'une fréquence initiale de 2/3 pour A mais dans des populations dix fois plus grandes que précédemment (cent vingt gènes). La différenciation entre populations, ainsi que leur fixation, se fait beaucoup plus lentement. La dérive est donc plus forte dans les petites populations que dans les grandes.



Figure 2 : Variation géographique de la fréquence de l'allèle I<sub>O</sub> du groupe sanguin O dans les populations humaines autochtones, sans tenir compte des immigrés récents (simplifié d'après Cavalli-Sforza et Bodmer<sup>1</sup>). Le groupe sanguin est un caractère déterminé par un locus à trois allèles I<sub>A</sub>, I<sub>B</sub>, I<sub>O</sub>. Chaque individu possède deux gènes à ce locus, et leur combinaison détermine le groupe sanguin. Les individus I<sub>A</sub>I<sub>A</sub> et I<sub>A</sub>I<sub>O</sub> sont de groupe A, les individus I<sub>B</sub>I<sub>B</sub> et I<sub>B</sub>I<sub>O</sub> sont de groupe B, les individus I<sub>A</sub>I<sub>B</sub> sont de groupe AB et les individus I<sub>O</sub>I<sub>O</sub> sont de groupe O.

des populations fictives de la figure 1, les populations humaines ont divergé par un processus aléatoire de dérive génétique.

Quels phénomènes naturels recouvrent le terme de hasard ou celui d'échantillonnage aléatoire ? En général, il s'agit juste d'un ensemble de causes situées hors de notre domaine d'étude. Pour le sexe d'un enfant, le hasard est présent dans les circonstances qui font qu'un spermatozoïde de type Y plutôt que X féconde l'ovule. Ces circonstances pourraient être prédictes si les infimes

1. L. L. Cavalli-Sforza et W. F. Bodmer, *Genetics of Human Populations*, San Francisco, W. H. Freeman, 1971.

détails des positions relatives des spermatozoïdes présents au moment de la fécondation étaient connus. Mais il est évident qu'ils sont hors de notre portée. Il en va de même pour la dérive génétique. Le *hasard* est un ensemble de circonstances, souvent triviales, qui font que tel individu survit et pas tel autre, ou se reproduit plus que tel autre, indépendamment des gènes qu'il porte. Les accidents d'automobile, par exemple, sont des causes de mortalité indépendantes du groupe sanguin. Ils trient les individus de façon aléatoire par rapport aux allèles  $I_A$ ,  $I_B$ ,  $I_O$ , et contribuent ainsi à la dérive génétique des fréquences de ces allèles. Leurs causes ne sont pas pour autant mystérieuses (la vitesse, l'alcool sont des causes bien identifiées), mais elles sont hors de notre domaine d'étude.

**Les variations aléatoires dépendent de l'effectif des populations.** Sur quatre enfants, la fréquence de garçons peut s'écarte considérablement de 1/2 (en prenant les valeurs 0, 1/4, 3/4 ou 1), mais sur les quelques millions de naissances annuelles en France, la fréquence globale est très proche de 1/2. De même, la dérive génétique est rapide dans des populations d'effectif restreint, mais très lente dans de grandes populations (figure 1). Dans une population infiniment grande, il n'y aurait pas de dérive : on retrouve ici la *loi de Hardy-Weinberg* exposée au chapitre 1. Le rôle de l'effectif a une conséquence théorique importante : pour comprendre l'évolution (génétique) des populations, nous devons prendre en compte leur démographie. Ainsi, les populations humaines n'auraient pas eu le temps d'accumuler, par dérive génétique, les différences que nous observons pour les allèles  $I_A$ ,  $I_B$ ,  $I_O$ , si l'espèce avait toujours compté plusieurs milliards d'individus comme aujourd'hui. Ces différences se sont établies à une époque où chaque population se comptait

en milliers ou en dizaines de milliers, avant l'expansion démographique des derniers dix mille ans.

La réitération du processus de dérive aboutit nécessairement à la fixation d'un allèle, car la fréquence allélique finit toujours par arriver à un (fixation de l'allèle et disparition de l'allèle alternatif), ou à zéro (réci-proque) (figure 1). Ces deux états sont définitifs, car aucun tirage au hasard ne peut faire réapparaître un allèle perdu : seule une nouvelle mutation ou l'arrivée d'un individu en provenance d'une autre population peut le faire. Ainsi, dans les populations natives d'Amérique du Sud, l'allèle  $I_O$  s'est fixé. Dans les populations sud-américaines actuelles, seul le métissage avec les Européens a fait réapparaître les groupes sanguins A et B.

La dérive détermine des généalogies de gènes. Au lieu de représenter une population par des fréquences alléliques, on peut considérer cette population comme un ensemble de gènes individuels empaquetés dans des organismes. Or, un fragment d'ADN présent dans un organisme est toujours la copie d'un fragment plus ancien, présent chez sa mère ou chez son père. Chaque gène de la population descend donc d'un gène de la génération précédente par un simple processus de copie (réPLICATION d'ADN). Ce lien de descendance permet d'établir une généalogie, en forme d'arbre, de tous les gènes de la population (figure 3). On peut alors exprimer la *dérive* en termes de *généalogies de gènes*.

Si chaque gène laissait toujours un et un seul descendant, la population de gènes resterait identique de génération en génération et la généalogie correspondrait à des lignées parallèles. Il n'y aurait pas de dérive génétique. La dérive vient de la variation aléatoire du nombre de descendants : certains gènes laissent deux descendants ou plus, d'autres aucun. Quand deux gènes se trouvent ainsi

avoir le même gène ancêtre, les deux lignées se rejoignent dans la généalogie (figure 3) formant ce qu'on appelle une *coalescence*. Lorsqu'on remonte dans le temps, le nombre de lignées diminue inexorablement, car chaque coalescence réduit deux lignées à une seule. Au bout d'un moment, il ne restera qu'une lignée : *tous les gènes d'une population descendant nécessairement d'un seul gène ancêtre, plus ou moins éloigné dans le passé.*

Comme la plupart des organismes, l'homme possède des mitochondries, qui sont des compartiments cellulaires pourvus d'un ADN propre. Contrairement aux autres, les gènes mitochondriaux sont transmis uniquement par les femmes. Les hommes les héritent de leur mère mais ne les transmettent pas à leurs descendants (figure 3). Ainsi, mes gènes mitochondriaux proviennent de ma mère, qui les tenait de sa propre mère, etc. Cette situation fournit un exemple du raisonnement précédent. La généalogie des gènes mitochondriaux forme des lignées maternelles qui coalescent entre elles, dès que deux femmes ont la même mère (cas de deux sœurs). Il existe quelque part dans le passé une et une seule mère de toutes les mères. Tous nos ADN mitochondriaux sont donc des copies de celui de cette lointaine ancêtre, surnommée « Ève mitochondriale », bien qu'ils ne soient pas identiques (en raison de mutations postérieures). Cette Ève a probablement vécu quelque part en Afrique vers 200 000 ans avant notre ère<sup>1</sup>. Cependant, malgré la métaphore biblique, elle n'a pas été la seule femme à se reproduire à cette époque. Chacun de nos gènes nucléaires actuels résulte d'une histoire semblable à celle de l'ADN mitochondrial. Toutefois, comme leur transmission est maternelle et paternelle, chacun

1. R.L. Cann *et al.*, « Mitochondrial DNA and Human Evolution », *Nature*, 325, 1987, p. 10-13.

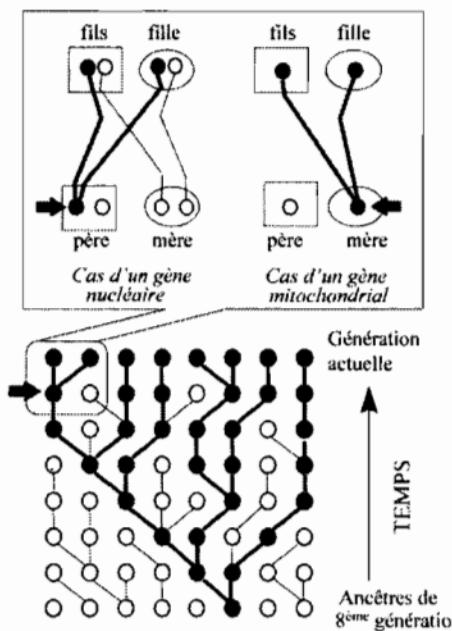


Figure 3 : La dérive s'exprime sous la forme d'une généalogie de gènes en forme d'arbre. Dans le schéma du bas, chaque point représente un gène et les alignements horizontaux représentent des générations successives d'une population de huit gènes. Les traits représentent la relation « être la copie de... », chaque gène étant la copie d'un gène de la génération précédente. On observe ainsi que les différents gènes de la génération actuelle sont reliés par un graphe arborescent, dans lequel les branchements correspondent à des événements de coalescence (flèches). Si l'on remonte suffisamment dans le passé, toutes les lignées des gènes de la population actuelle coalescent sur une lignée ancêtre unique. Le détail du mécanisme de la coalescence est donné dans la partie supérieure de la figure, sur le cas d'une famille de deux enfants (un garçon et une fille : les mâles sont représentés par des carrés, les femelles par des cercles). La plupart des gènes, comme par exemple ceux du groupe sanguin A, B, O, sont codés par l'ADN du noyau des cellules (gènes nucléaires). Dans ce cas, chaque individu contient deux gènes, un venant du père, l'autre de la mère. Une coalescence a lieu lorsque les deux enfants reçoivent des copies du même gène d'un des deux parents (dans

d'eux provient certainement d'un ancêtre différent, mâle ou femelle.

## Comment prédire et tester quantitativement la dérive ?

Par nature, le phénomène de dérive est aléatoire d'où l'impossibilité de prédire les fréquences alléliques futures avec certitude. Ne peut-on pour autant rien dire des phénomènes aléatoires ? Intuitivement, nous avons déjà émis des prédictions quantifiées : la *moyenne* de la proportion de garçons dans une famille sera 1/2. De même, la fréquence moyenne d'un allèle dans une population de descendants sera  $p$ , sa fréquence dans la population de parents. La moyenne est un exemple de prédiction de type *statistique*, c'est-à-dire qui porte sur une collection de situations futures possibles. Elle nous informe sur l'avenir, même si elle ne nous le décrit pas avec certitude. Quelle est l'information la plus complète que nous puissions obtenir sur les valeurs futures d'une fréquence allélique ? Il s'agit de la *distribution statistique* de cette fréquence, soit l'ensemble de toutes les fréquences futures possibles accompagnées chacune d'une *probabilité*. L'interprétation la plus simple des probabilités – celle du sens commun – est une valeur limite de

le cas illustré, le père). Pour les gènes de la mitochondrie, chaque individu n'a qu'un gène, donné par sa mère. Les coalescences ont donc lieu exclusivement par voie maternelle, chaque fois que deux individus ont la même mère. N.B. : Un gène donné est soit nucléaire, soit mitochondrial. Les deux cas ne sont représentés sur la même figure que par commodité, les généalogies de gènes obtenues étant semblables.

proportion : la pièce a une probabilité 1/2 de tomber sur pile car, si je la lance un très grand nombre de fois, j'obtiendrai une proportion de *pile* qui se rapproche de 1/2. Pour quantifier la dérive génétique, on cherchera donc à répondre aux questions suivantes :

(1) *Étant donné une population actuelle, quelles sont les probabilités des différentes fréquences alléliques possibles dans les populations de descendants et comment changeront-elles au cours du temps ?*

(2) *En particulier, avec quelle probabilité un allèle se fixera-t-il ou à l'inverse sera-t-il perdu ?*

(3) *Combien de temps, en moyenne, nécessitera la fixation d'un allèle ?*

Des modèles mathématiques permettent de calculer les distributions statistiques des fréquences alléliques. Le modèle de base est la *loi binomiale*. Pour simplifier, on se donne une population de départ comportant deux allèles *A* et *B*, en fréquences respectives  $p_0$  et  $1 - p_0$  (par exemple, 1/3 et 2/3). Pour constituer la population de descendants, on échantillonne aléatoirement  $n$  gènes, chacun étant la copie d'un des gènes parentaux pris au hasard. Les mathématiciens formalisent la distribution statistique du nombre de *A* dans cet échantillon sous le nom de *distribution binomiale*, symbolisée par  $B(n, p)$ . Leurs formules (données dans la plupart des manuels de mathématiques) donnent les valeurs exactes de toutes les fréquences alléliques possibles dans la première génération de descendants (valeurs de  $p_1$ ) et de leurs probabilités. Prenons par exemple le cas d'une population de  $n = 3$  gènes et  $p_0 = 1/3$  (un gène *A* et deux gènes *B*). La distribution  $B(3, 1/3)$  donne quatre fréquences possibles chez les descendants :  $p_1 = 0$  (trois gènes *B* et aucun *A*) avec une probabilité de 8/27 ;  $p_1 = 1/3$  (un gène *A* et deux *B*) avec une probabilité de 12/27 ;  $p_1 = 2/3$

( $2A + 1B$ ) avec une probabilité de  $6/27$ ; et  $p_1 = 1$  ( $3A$  et aucun  $B$ ) avec une probabilité de  $1/27$ .

Il faut maintenant réitérer ce processus pour obtenir les générations suivantes. La distribution de la fréquence au temps  $t$ , notée  $p_t$ , dépendra évidemment de la fréquence réalisée une génération avant ( $p_{t-1}$ ). Ce type de processus, où la distribution au temps  $t$  ne dépend que de la valeur réalisée à  $t-1$ , est appelé *processus de Markov*. Connaissant les formules de la loi binomiale, on peut écrire les équations de récurrence de ce processus (modèle de Wright-Fisher) et les résoudre : on obtient ainsi les probabilités des différentes valeurs possibles de la fréquence allélique de  $A$  à un temps  $t$  quelconque. Cette résolution ne sera pas détaillée ici. Il importe seulement de savoir que, même à l'aide des meilleurs ordinateurs, les formules obtenues ne peuvent être écrites et calculées exactement que pour des effectifs  $n$  relativement faibles ( $n < 30$ ). Il faut donc recourir à des approximations, car nombre de populations réelles ont des effectifs bien supérieurs (qui se comptent en centaines ou en milliers).

On doit au Japonais Motoo Kimura l'invention d'une technique d'approximation permettant de résoudre ce problème. Cette technique part du principe que lorsque l'effectif  $n$  est suffisamment élevé (quelques dizaines ou plus), le nombre de fréquences alléliques possibles (tous les multiples de  $1/n$  entre 0 et 1) devient grand : le spectre des fréquences ressemble alors à un spectre continu. L'évolution de la distribution des fréquences dans cet espace de fréquences possibles, par dérive génétique à partir de la fréquence initiale  $p_0$ , ressemble à un phénomène de *diffusion* progressive de particules initialement concentrées au point  $p_0$  (voir figure 1, où chaque population serait une particule). La seule exception à

cette analogie est le fait que les « particules » arrivant sur les bords du spectre (fréquences 0 ou 1) s'y retrouvent piégées. En effet, comme nous l'avons vu, ces situations correspondent à la fixation, ou à la perte (définitive), d'un allèle. Il s'agit donc d'une diffusion avec « bords absorbants ». Des équations différentielles (non détaillées ici), analogues à celles employées en physique pour décrire la diffusion de particules, peuvent être écrites et donnent la distribution des fréquences alléliques à un temps  $t$  quelconque, connaissant la valeur de départ  $p_0$ .

Que pouvons-nous déduire de ces modèles ? Le premier résultat marquant est qu'*une population comprenant initialement un mélange de deux types de gènes A et B, finira par être composée uniquement de A ou uniquement de B* (fixation de A ou de B), quelles que soient les fréquences relatives de A et B au départ. La probabilité de fixation d'un allèle est égale à sa fréquence initiale. Ainsi, si la population contient 90 % de A au départ, à long terme, dans 90 % des cas, elle sera composée uniquement de A. Cependant, dans 10 % des cas, A sera éliminé et la population ne contiendra que des B. Ce n'est donc pas toujours le type majoritaire qui finit par dominer. Nous pouvons appliquer cela au cas d'une mutation qui vient d'apparaître dans une population de  $n$  gènes (par exemple, 1 000). En général, la mutation n'affecte qu'un gène parmi les  $n$  présents, et se trouve au départ en fréquence  $1/n = 1/1\,000$ . Une nouvelle mutation a par conséquent une chance sur 1 000 d'être transmise à toute la population, et 999 chances sur 1 000 d'être perdue. Si la proportion de  $1/1\,000$  semble faible, il arrive, au cours du temps évolutif, tant de mutations que cela suffit à faire changer substantiellement les populations.

Les modèles donnent également le temps nécessaire à une nouvelle mutation pour envahir la population, si elle y arrive : ce temps est en moyenne proche de  $2n$ , soit 2 000 générations pour l'exemple choisi. En faisant un raisonnement de type coalescence, c'est-à-dire en remontant le temps, il faut en moyenne remonter 2 000 générations pour trouver le gène ancêtre de la totalité des 1 000 gènes d'une population actuelle. Pour une population de 100 gènes, il suffit de remonter 200 générations. On retrouve ici le rôle de l'effectif : la dérive est beaucoup plus rapide dans les petites populations.

Un second résultat théorique important est que les états intermédiaires entre la population de départ et l'état définitif (population soit fixée pour A, soit fixée pour B) sont hautement imprédictibles : en effet, à part les fréquences 1 (fixation de A) et 0 (fixation de B), toutes les autres valeurs de fréquences sont rapidement aussi probables les unes que les autres.

Pour évaluer la pertinence de ces prédictions, une expérience a été menée par le généticien P. Buri<sup>1</sup> en 1956, utilisant des populations artificielles de mouche du vinaigre (*Drosophila melanogaster*). Chez cette mouche, on s'intéresse à deux allèles d'un locus qui contrôle la couleur des yeux : *bw* (blanc) et *bw<sup>75</sup>* (orange). Chaque mouche possède deux gènes pour ce locus, et peut donc être *bwbw*, *bwbw<sup>75</sup>* ou *bw<sup>75</sup>bw<sup>75</sup>*. Buri a fondé cent huit populations comportant initialement chacune seize individus de type *bwbw<sup>75</sup>*. De cette façon, un gène sur deux est *bw*, et la fréquence allélique initiale de *bw* est donc de 1/2. Il a ensuite laissé les populations se reproduire isolément pendant dix-neuf

1. P. Buri, « Gène Frequencies in Small Populations of Mutant *Drosophila* », *Evolution*, 18, 1956, p. 379-383.

générations, en gardant toujours seize individus, soit trente-deux gènes. Au cours de ces générations, il a régulièrement évalué la fréquence de l'allèle *bw* dans chaque population. Par ailleurs, il a calculé la distribution des fréquences alléliques, donnée par les équations de la dérive (c'est-à-dire le nombre attendu de populations ayant 0, 1, 2, ..., 32 gènes *bw*). La comparaison entre les données et les prédictions théoriques montre une adéquation globalement satisfaisante. La dérive génétique observée se révèle même plus forte que prévu : les données correspondent à ce qu'on attendrait s'il y avait douze et non seize mouches dans chaque population, ce qui suggère que certaines mouches sont inaptes à se reproduire. Ce type d'expérience valide les calculs faits, et ouvre la porte à leur application aux populations naturelles.

## Diversité génétique dans les populations naturelles et dérive génétique

**Comment relier les prédictions théoriques faites sur la dérive à des observations en populations naturelles ?** Ces dernières sont bien plus complexes que les populations artificielles ! Tout d'abord, leur délimitation est difficile : la nature ne contient pas de limites aussi nettes que les parois d'une cage à drosophiles. Ensuite, l'effectif des populations naturelles n'est souvent connu que par un ordre de grandeur, généralement élevé (bien plus que les seize individus de l'expérience de Buri). Enfin, chaque population est unique : on ne dispose pas de réplicats expérimentaux. La plus grosse difficulté rencontrée reste néanmoins qu'on ne connaît pas beaucoup de caractères

qui, à la façon des yeux blancs chez la drosophile, sont des indicateurs de l'existence d'un allèle précis à un locus précis. La plupart des caractères morphologiques qui varient dans les populations naturelles ont un déterminisme génétique encore inconnu, impliquant plusieurs locus.

Cette dernière difficulté a été levée à la fin des années 1960, avec l'invention d'une méthode qui permet d'identifier rapidement, chez n'importe quelle espèce animale ou végétale, des caractères servant de révélateurs d'un locus génétique. On appelle ces caractères des *marqueurs génétiques*. Richard Lewontin et Jack Hubby<sup>1</sup> ont utilisé une technique classique de biochimie, l'électrophorèse, qui permet de séparer des protéines légèrement différentes par leur composition en acides aminés. Le principe est de placer les échantillons de protéines dans un milieu, gélatineux (« gel »), et de les soumettre à un champ électrique. Les protéines sont chargées (en général négativement) et migrent vers le pôle de signe opposé (positif) à travers le gel. Deux protéines qui n'ont pas exactement la même composition en acides aminés ont des charges légèrement différentes, et migreront à des vitesses distinctes. L'électrophorèse est une sorte de course entre protéines : on distingue à l'arrivée des lentes et des rapides. La découverte majeure de Lewontin et Hubby est que, lorsqu'on applique cette technique à des extraits protéiques provenant d'individus différents d'une même population, ils possèdent souvent des formes différentes de la même protéine. Ainsi, on peut observer, pour la protéine du sang nommée *hémoglobine*, des formes

1. R. C. Lewontin et J. L. Hubby, « A Molecular Approach to the Study of Genetic Heterozygosity in Natural Populations. II. Amount of Variation and Degree of Heterozygosity in Natural Populations of *Drosophila pseudoobscura* », *Genetics*, 54, 1966, p. 595-609.

lentes, des formes rapides ou divers degrés d'intermédiaires. L'intérêt de ces différences est que chaque forme correspond en fait à un allèle précis du locus hémoglobine. Cela se répète pour beaucoup d'autres protéines bien connues. On a donc, grâce à l'électrophorèse, une source importante de caractères qui varient dans les populations naturelles, sous la dépendance de quelques allèles précis à des locus précis. L'innovation de Lewontin et Hubby a été suivie d'un grand nombre de travaux utilisant la technique d'électrophorèse des protéines sur des populations d'espèces variées. Par ailleurs, d'autres techniques, liées au développement de la biologie moléculaire à partir des années 1980, ont permis d'identifier des locus et des allèles directement sur l'ADN. Ces caractères (migration des protéines dans un gel ou variations de séquence d'un morceau d'ADN) sont appelés des « marqueurs génétiques » ou « marqueurs moléculaires » car ils donnent à chaque individu une « étiquette » moléculaire transmise génétiquement à ses descendants. Les groupes sanguins A, B, O sont aussi un exemple de tels marqueurs. Une de leurs applications est la recherche de paternité : pour établir qu'un enfant est ou n'est pas de son père présumé, on comparera leurs étiquettes moléculaires (en commençant, en général, par le groupe sanguin).

La conclusion majeure des nombreux travaux qui ont analysé des populations naturelles à l'aide de marqueurs moléculaires est que la variation est quasiment omniprésente, c'est-à-dire qu'on trouve facilement des individus porteurs d'allèles différents dans une même population. Ce polymorphisme fréquent des populations naturelles a beaucoup surpris les chercheurs, lorsqu'il a été découvert. En effet, la conception dominante (plus ou moins implicite) présentait la sélection naturelle comme un phéno-

mène omniprésent : toute nouvelle mutation devait être soit sélectionnée positivement, soit négativement. Dans les deux cas, la théorie prédit que le polymorphisme est perdu très vite. Dans le premier cas, effectivement, la mutation envahit rapidement la population, et dans le second cas elle est rapidement perdue. Il aurait donc dû être improbable d'observer les états transitoires où coexistent le mutant et l'allèle original. L'idée d'une sélection omniprésente avait déjà été critiquée sur des bases théoriques (voir *Neutralistes contre sélectionnistes*, p. 333). Elle n'a cependant été définitivement réfutée que par la découverte du polymorphisme moléculaire. Les données moléculaires ont alors été utilisées pour défendre l'idée que, contrairement à l'hypothèse initiale, beaucoup de mutations ne sont pas soumises à la sélection naturelle (théorie neutraliste de l'évolution moléculaire de Motoo Kimura<sup>1</sup>). Ces mutations, dites *neutres*, vont évoluer au sein des populations par dérive génétique pure.

*Ainsi, la dérive génétique a acquis une importance égale à celle de la sélection naturelle pour comprendre l'évolution moléculaire.* Il s'est produit, au sein du microcosme des évolutionnistes, une sorte de minirévolution conceptuelle, au terme de laquelle le hasard (formalisé à travers la dérive génétique) a été admis parmi les forces évolutives au même titre que la sélection déterministe. L'aspect aléatoire de l'évolution avait été intuitivement évoqué dès Darwin (qui insiste à plusieurs reprises sur la contingence historique de l'évolution), mais il était plus ou moins relégué au rang de « bruit de fond » qui brouillait l'action de la sélection, à la manière du frottement de

---

1. M. Kimura, *Théorie neutraliste de l'évolution*, Paris, Flammarion (Nouvelle Bibliothèque Scientifique), 1990 (1<sup>re</sup> éd. 1983).

l'air qui ralentirait la chute d'un objet. Ayant pris conscience de l'importance de la dérive pour la variation moléculaire, les évolutionnistes l'incluent désormais dans les facteurs déterminants de l'évolution biologique en général.

**La théorie de la dérive permet-elle de comprendre les données moléculaires ?** Pour cela, il faut que cette théorie fasse la prédiction suivante : les phases transitoires, pendant lesquelles plusieurs variants moléculaires coexistent dans une population, sont suffisamment longues pour qu'on ait une chance raisonnable de les observer à un moment donné. Supposons que chaque nouvelle mutation introduit un allèle nouveau (ce qui est en général le cas). Nous pouvons représenter l'histoire de la population comme une succession d'événements de fixation d'allèles apparus par mutation (figure 4).

Entre la mutation et la fixation, la population comporte au moins deux allèles différents, mais elle peut éventuellement en compter trois ou plus, si une nouvelle mutation intervient avant que la précédente ne soit fixée ou éliminée. Si les mutations sont suffisamment fréquentes ou si le temps moyen nécessaire à la fixation est suffisamment long, on observera, *la plupart du temps*, deux ou plusieurs allèles au sein de la population. Cela est exactement conforme aux observations sur les marqueurs moléculaires.

Pour quantifier cette prédiction, commençons par trouver un moyen de quantifier le polymorphisme, c'est-à-dire la diversité des allèles présents à un moment donné dans une population. La manière la plus simple est d'imaginer que nous tirons au hasard deux gènes dans la population et de regarder s'ils sont identiques ou différents. Si, dans 90 % des cas, on obtient deux gènes diffé-

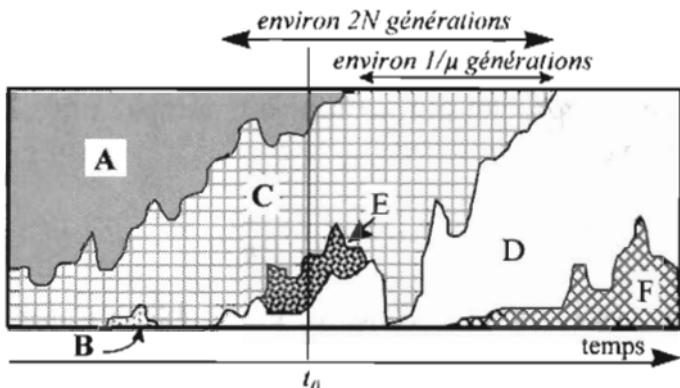


Figure 4 : L'histoire génétique d'une population est une suite de cycles de mutation-fixation. A, B, ..., F représentent des allèles qui apparaissent par mutation d'allèles précédents (par exemple, F apparaît à partir d'une mutation d'un gène D, B à partir d'une mutation d'un gène C). Les fréquences alléliques (représentées par la largeur de la bande correspondant à chaque allèle) varient au cours du temps par dérive génétique. Certains allèles (C, D dans la figure) envahissent complètement la population (fixation), alors que d'autres (B, E) disparaissent rapidement. Le temps mis par une mutation (ici, D) pour se fixer est en moyenne de  $2n$  générations, où  $n$  est le nombre de gènes de la population. De nouvelles mutations peuvent apparaître avant qu'une mutation se fixe, si bien qu'à un temps donné ( $t_0$ ) on peut observer plusieurs allèles dans la population (ici, quatre allèles : A, C, D, E). Le taux de nouvelles fixations est proche de  $\mu$ , le taux de mutation, si bien qu'un intervalle moyen de  $1/\mu$  générations sépare deux fixations successives.

rents, alors la diversité est grande, si cela n'arrive que dans 5 % des cas, elle est faible. La probabilité de tirer deux gènes différents dans la population, notée  $H$ , sera donc une mesure de la diversité génétique dans la population. On peut démontrer, en utilisant les calculs théoriques de la dérive génétique, que dans une population

de  $n$  gènes, qui mutent à chaque génération avec un très faible taux de mutation  $\mu$ , la diversité génétique moyenne approche au cours du temps une limite donnée par :

$$\hat{H} = 1 - \frac{1}{1 + 2n\mu}$$

Que nous apprend cette formule ? Elle montre tout simplement que l'on ne doit pas observer de polymorphisme quand  $2n\mu \ll 1$  (auquel cas  $H$  est proche de zéro). En revanche, dès lors que l'ordre de grandeur de  $2n\mu$  est proche de 1, la diversité n'est pas du tout négligeable. Les taux de mutation  $\mu$  sont en général faibles, et les effectifs des populations ( $n$ ) souvent grands, si bien que l'un compensant l'autre, le produit  $2n\mu$  peut tout à fait être proche de 1. On a pu estimer par exemple que les protéines mutent avec une probabilité proche de  $\mu = 10^{-6}$  (une sur un million) par génération. Dans une espèce animale ou végétale, comme la drosophile, l'ensemble des populations sauvages totalise largement plus d'un million d'individus (soit deux millions de gènes). On obtient donc  $2n\mu > 1$ , et une diversité attendue élevée, conforme à ce qui est observé. Comme on le voit, la connaissance précise de  $n$  et de  $\mu$  n'est pas nécessaire, et c'est heureux car il est impossible de compter toutes les mouches du monde ! L'ordre de grandeur suffit pour comprendre que les données sont qualitativement cohérentes avec les prédictions de la dérive génétique.

Un autre test de la formule précédente est de nature comparative : si nous pouvons trouver des situations qui diffèrent pour  $n$  et/ou pour  $\mu$ , cette différence devrait se répercuter sur  $H$ . Par exemple, si deux populations P1 et P2 diffèrent par leur taille ( $n_2 > n_1$ ) on devrait observer une diversité plus forte dans la population la plus grande

$(H_2 > H_1)$ . De même, si deux marqueurs moléculaires, M1 et M2, diffèrent par leur taux de mutation ( $\mu_2 > \mu_1$ ) alors la diversité devrait être plus grande pour le marqueur qui mute le plus ( $H_2 > H_1$ ). Ces prédictions ont pu être vérifiées statistiquement dans les populations naturelles d'animaux et de végétaux, toujours en s'en tenant à des ordres de grandeurs plutôt qu'à des valeurs précises impossibles à obtenir. On peut ainsi comparer chez une même espèce la variabilité au niveau de deux classes de marqueurs génétiques comme les protéines enzymatiques (analysées par électrophorèse) et les séquences d'ADN microsatellite (petits motifs répétés). L'ordre de grandeur de taux de mutation des premières est estimé à 1 pour un million par génération ( $\mu_1 = 10^{-6}$ ), alors que les seconds mutent beaucoup plus fréquemment ( $\mu_2 = 10^{-3}$ ). Un calcul rapide à partir de la formule ci-dessus montre que le rapport

$$R_{1,2} = \frac{H_1 (1 - H_2)}{H_2 (1 - H_1)}$$

doit être égal au rapport des taux de mutation  $\mu_1/\mu_2$ , dans notre cas  $10^{-6}/10^{-3} = 10^{-3}$ . Chez l'escargot d'eau douce *Bulinus truncatus*, plusieurs populations ont été analysées pour plusieurs marqueurs de type enzymatique et microsatellite<sup>1</sup>. La diversité génétique des marqueurs enzymatiques se révèle faible (en moyenne,  $H_1 = 0,000$  à 0,005), alors que celle des marqueurs microsatellites

---

1. Cf. F. Njiokou *et al.*, « Do Self-Fertilization and Genetic Drift Promote a Very Low Genetic Variability in the Allotetraploid *Bulinus truncatus* (Gastropoda, Planorbidae) ? », *Genetical Research*, 62, Cambridge, 1993, p. 89-100 ; et F. Viard *et al.*, « Microsatellites and the Genetics of Highly Selfing Populations in the Freshwater Snail *Bulinus truncatus* », *Genetics*, 142, 1996, p. 1237-1347.

est élevée (en moyenne  $H_2 = 0,3$  à  $0,9$ ). Le rapport  $R_{1,2}$  observé est donc de l'ordre de grandeur de  $0,001$  ( $10^{-3}$ ) comme prévu.

## Les gènes neutres, marqueurs de l'histoire et de la structure des populations

**Les marqueurs génétiques moléculaires peuvent renseigner sur l'effectif des populations.** En effet, dans une population stable, la formule précédente relie la diversité génétique  $H$  à l'effectif  $n$  de la population. L'utilisation de cette formule nécessite de connaître le taux de mutation des marqueurs utilisés et de faire l'hypothèse que leur variation moléculaire est neutre (non soumise à sélection naturelle). Sous cette même hypothèse, les marqueurs peuvent aussi nous donner nombre d'autres informations : ils jouent le rôle de témoins de l'histoire et de la structure des populations naturelles. Le terme *structure des populations* décrit leur fonctionnement actuel, et en particulier inclut la réponse aux questions suivantes : comment peut-on partitionner l'ensemble des individus d'une espèce en sous-ensembles (populations locales ou dèmes) et où se trouvent leurs limites géographiques ? Quel est l'effectif d'une population locale ? Existe-t-il des échanges d'individus entre sous-ensembles (flux migratoires) et quelle est leur intensité ?

Au contraire, lorsqu'on parle d'*histoire des populations*, on s'intéresse à des événements du passé, et en particulier aux changements d'effectif et de structure des populations au cours du passé « récent ». Comme nous l'avons vu, la totalité des gènes d'une population provient, en moyenne, d'un seul gène situé  $2n$  générations dans le

passé. On ne peut donc guère espérer que des événements vieux de plus de  $2n$  générations laissent des traces dans les gènes actuels.

Pour décrire la structure spatiale des populations, nous utiliserons une propriété de la dérive illustrée en figure 2 (*supra*). Plusieurs populations, au départ identiques, tendront à avoir des fréquences alléliques de plus en plus différentes lorsqu'elles restent complètement isolées les unes des autres, c'est-à-dire qu'elles n'échangent pas de migrants. Pour mesurer la structure des populations, on évaluera donc la *variabilité des fréquences alléliques* (techniquement, la variance de ces fréquences) *entre populations*. Cette mesure a été définie par Sewall Wright<sup>1</sup>, en 1951, sous le nom de  $F_{ST}$  et peut être obtenue en utilisant des marqueurs génétiques. En effet, il suffit, pour un gène marqueur, de déterminer les allèles (variants moléculaires) portés par les individus de plusieurs populations différentes, de calculer leur fréquence dans chaque population, et enfin de calculer la variance de ces fréquences.

Bien sûr, dans la nature, il existe peu de populations complètement isolées : des individus migrent entre populations. Le résultat de ces migrations est d'homogénéiser les fréquences alléliques entre populations. La migration permet effectivement aux gènes de se déplacer d'une population à l'autre. *La migration, à l'encontre de la dérive, tend donc à diminuer  $F_{ST}$ .* Cette loi générale permet de chercher à estimer les migrations entre populations en utilisant le paramètre de différenciation génétique  $F_{ST}$ .

---

1. S. Wright, « The Genetical Structure of Populations », *Annals of Eugenics*, 15, 1951 p. 323-354.

Deux problèmes se posent alors. Premièrement, la dérive est un phénomène aléatoire et une même structure de populations pourra engendrer des  $F_{ST}$  variables simplement du fait du hasard. Pour s'affranchir de ce problème, il faut travailler sur des moyennes, en utilisant un grand nombre de populations, et si possible un grand nombre de marqueurs différents. Deuxièmement, la migration n'est pas simple à décrire. À la limite, si l'on a dix populations, les flux migratoires entre les populations 1 et 2 peuvent être différents de ceux entre 2 et 3 ou entre 1 et 5. Cette variation décourage, *a priori*, le chercheur de décrire la structure en utilisant un seul paramètre. Pour surmonter ce problème, on recourt à des modèles idéalisés, dans lesquels des simplifications sont volontairement introduites.

Le plus célèbre de ces modèles est le « modèle en îles » de Wright<sup>1</sup>, élaboré en 1931. Dans ce modèle, on imagine un grand nombre de populations réparties sur un territoire à la manière d'îles sur l'océan. Toutes les îles contiennent le même nombre d'individus (noté  $N$ ; le nombre de gènes,  $n$ , est égal au double du nombre d'individus soit  $2N$ ) et, à chaque génération, une fraction  $m$  d'immigrants arrive dans chaque île en provenance des autres îles. Chacun des migrants navigue sur une barque individuelle et débarque sur une autre île, choisie au hasard. Dans ce modèle, la mutation est supposée négligeable. Dans ces conditions, Wright montre que l'action de la migration (qui homogénéise les populations) et celle de la dérive (qui les différencie) s'équilibrivent et atteignent un état stationnaire caractérisé par :

---

1. S. Wright, « Evolution in Mendelian Populations », *Genetics*, 16, 1931, p. 97-159.

$$F_{ST} \approx \frac{1}{1 + 4Nm}$$

Cette formule est remarquable car elle relie un paramètre mesurable ( $F_{ST}$ ) au produit  $Nm$ , qui représente exactement le nombre de migrants qui arrivent dans chaque population à chaque génération. Quand ce nombre est élevé, les populations sont peu différenciées génétiquement et  $F_{ST}$  est proche de zéro. Au contraire, des populations très isolées ( $Nm$  très petit) ont un  $F_{ST}$  proche de sa valeur maximale : 1. D'autres modèles, plus complexes que le modèle en îles, peuvent être construits, par exemple en supposant que les migrants ne débarquent pas vers une île choisie au hasard mais se dirigent systématiquement vers l'île la plus proche. On obtient alors des formules légèrement différentes de la formule ci-dessus, mais la philosophie générale reste la même. Dans tous les cas, ces modèles ne peuvent pas être « pris au pied de la lettre », car on sait bien que les populations réelles ne vérifient pas les conditions plutôt exigeantes du modèle. Cependant, ils donnent des ordres de grandeur que l'on suppose « robustes », valables même si les conditions ne sont pas intégralement respectées. De plus, ils permettent de comparer diverses situations. De même qu'une monnaie permet d'exprimer dans la même unité la valeur d'objets très différents, le  $F_{ST}$  quantifie la structure génétique dans une unité commune à toutes les populations et à toutes les espèces.

L'exemple de l'escargot *Bulinus truncatus* illustrera l'utilisation du  $F_{ST}$  pour analyser la structure des populations. Cet escargot vit dans les eaux douces tropicales. Une étude<sup>1</sup> a été menée au Niger, où l'habitat principal

---

1. F. Viard *et al.*, *op. cit.*

de l'espèce est un ensemble de mares dispersées et isolées. Ces conditions ressemblent à celles du modèle en îles. En effet, du point de vue de l'escargot, les mares sont des îlots d'habitat favorable, à l'intérieur d'un océan de territoire défavorable. Des escargots peuvent passer d'une mare à l'autre (par exemple, transportés involontairement par les hommes ou les animaux), mais le phénomène est peu fréquent. Un échantillonnage a été effectué de sorte à pouvoir comparer entre elles des populations de mares différentes ainsi que des populations situées à différents endroits dans la même mare. Ces populations ont été analysées pour quatre marqueurs génétiques de type ADN microsatellite, et les  $F_{ST}$  ont été estimés. Sur l'ensemble des populations (dix mares), le  $F_{ST}$  vaut 0,38 (moyenne sur les quatre locus marqueurs). Il existe donc de fortes différences entre populations. Pour des populations prises à l'intérieur d'une même mare (respectivement *Mari* et *Namaga*), la différenciation est beaucoup plus faible :  $F_{ST} = 0,01$  pour les deux populations de *Mari* et  $F_{ST} = 0,07$  pour les trois populations de *Namaga*. Cela correspond qualitativement aux prédictions *a priori* : des populations situées dans la même mare échangent plus de migrants que des populations de mares différentes. En introduisant les valeurs de  $F_{ST}$  dans une formule du type de celle de Wright (ci-dessus), on estime que un à cent individus passent d'une population à l'autre dans la même mare à chaque génération. À l'inverse, deux populations de mares différentes échangent de l'ordre de un migrant toutes les dix générations. Un résultat surprenant a été obtenu en comparant plusieurs populations de la mare *Boyzé*. On a en effet obtenu  $F_{ST} = 0,30$ , plus élevé que pour *Namaga* ou *Mari*, et proche de l'ordre de grandeur du  $F_{ST}$  « entre mares différentes ». Cela montre que les populations de

cette mare, contrairement à ce qu'on aurait pu penser, sont relativement isolées l'une de l'autre. Ainsi, les résultats des marqueurs génétiques permettent de révéler des propriétés insoupçonnées de certaines populations, comme, dans ce cas, leur isolement.

**La dérive peut également servir de révélateur de l'histoire d'une population.** En effet, elle est sensible à l'effectif des populations : les accidents démographiques passés laissent donc des traces sur la diversité génétique actuelle. Le cas le plus étudié est celui dit des « *goulets d'étranglement* », lorsqu'une population passe par une phase à très faibles effectifs, suivie éventuellement d'un retour à des effectifs élevés. Plusieurs espèces, comme le guépard ou l'éléphant de mer, ont ainsi été dangereusement près de l'extinction. On estime que la population mondiale d'éléphants de mer était réduite au début du siècle à quelques dizaines d'individus, bien qu'elle soit actuellement de plusieurs dizaines de milliers. Le passage par de faibles effectifs a augmenté énormément la dérive, conduisant à la fixation rapide d'allèles, et donc à la diminution de la diversité. Depuis ce goulot d'étranglement, et malgré les effectifs redevenus élevés, la diversité, évaluée par des données moléculaires (électrophorèse des protéines et ADN), est restée très faible ( $H$  proche de zéro), car les nouvelles mutations n'ont pas encore eu le temps de la régénérer. Les événements de goulot d'étranglement sont également impliqués chez des espèces qui, loin de s'éteindre, colonisent de nouveaux milieux. Les populations nouvelles (par exemple, celles qui s'établissent sur une île non encore atteinte par l'espèce) sont effectivement souvent fondées par très peu d'individus, même si elles connaissent ensuite une forte croissance démographique. La perte de la diversité génétique agit encore une fois comme une signature de cet événement.

Un exemple est fourni par le rat commun (*Rattus norvegicus*), chez qui des marqueurs génétiques de type ADN microsatellite ont été analysés. Les populations continentales de rats présentent une variabilité importante pour ces marqueurs, avec en moyenne huit allèles par population. Cependant, dans les populations des îles situées le long des côtes bretonnes, on observe fréquemment deux ou trois, voire un seul allèle. Une explication probable est que seul(s) un (ou quelques) couple(s) fondateur(s) est à l'origine de la totalité de la population d'une île donnée. Si cette colonisation a eu lieu dans un passé récent, la probabilité que de nouvelles mutations introduisent des allèles différents dans la population insulaire est faible. Cette hypothèse est confirmée dans l'île de Trielen et l'île aux Chrétiens en mer d'Iroise, où on a pu dater l'introduction, car elle correspondait au naufrage d'un bateau (dont les cales abritaient des rats), en 1912. Le faible nombre de générations de rats écoulées depuis cette date, compte tenu du taux de mutation des marqueurs génétiques utilisés et de l'effectif des populations, permet de supposer qu'aucune mutation ne s'est fixée depuis et que la variabilité génétique des populations insulaires reflète directement l'événement historique de colonisation.

**La dérive agit aussi à l'échelle de communautés d'espèces.** Comprendre comment différentes espèces coexistent pour former ce qu'on appelle des communautés d'espèces (par exemple une forêt comprenant plusieurs essences d'arbres) a longtemps été considéré comme le travail des écologues, sans rapport direct avec la théorie de l'évolution. Leur concept principal était l'idée de niche écologique : ils postulaient que deux espèces ayant les mêmes niches écologiques, c'est-à-dire exploitant les mêmes ressources et supportant les mêmes

conditions (climats, types de sol, etc.) ne pouvaient pas coexister, car la compétition était si forte que l'une éliminait l'autre. Cette idée a inspiré de nombreuses études, visant à rechercher quelles différences de niche permettaient à tel ou tel assemblage d'espèces de persister. En s'inspirant de la théorie neutraliste de l'évolution moléculaire, Stephen Hubbell<sup>1</sup> a récemment proposé une théorie radicalement différente, dite *théorie neutraliste de la biodiversité*. Hubbell part du principe que les espèces d'une communauté sont des catégories d'individus, comme les allèles d'une population sont des catégories de gènes. Ainsi, même s'il n'existe aucune différence de niche entre les espèces, celles-ci se comporteraient comme des allèles neutres, et les processus d'échantillonnage aléatoire que nous nommons dérive génétique pour les allèles d'une population se dérouleraient de même manière entre espèces d'une communauté, où ils prendraient le nom de dérive écologique. Toute la théorie de la dérive génétique, pourrait ainsi être reprise pour comprendre l'histoire des communautés et leur richesse spécifique. Les tenants de cette théorie citent comme exemples des communautés telles que les forêts tropicales, comportant de très nombreuses espèces d'arbres dont les niches écologiques ne semblent pas très différentes. Cependant la théorie neutraliste de la biodiversité est encore débattue car il existe bien d'autres cas où les différences de niches semblent indiscutablement importantes. En fait, l'étude détaillée de deux espèces montre toujours qu'elles n'utilisent pas strictement les mêmes ressources et n'ont pas exactement les mêmes exigences écologiques ; même s'il n'est en général pas clair que ces

---

1. S. P. Hubbell, *The Unified Neutral Theory of Biodiversity and Biogeography*, Princeton NJ, Princeton University Press, 2001.

différences suffisent à expliquer leur coexistence. Ainsi, un des mérites de la théorie neutraliste de la biodiversité, à l'instar de la théorie neutraliste de l'évolution moléculaire qui l'a inspirée, est d'avoir souligné le rôle des facteurs stochastiques et des processus d'échantillonnage (la dérive écologique), au même titre que les facteurs déterministes (liés aux niches écologiques), dans la régulation de la biodiversité.

### La dérive génétique et l'horloge moléculaire

**Comment la dérive va-t-elle agir à l'échelle des temps géologiques ?** À cette échelle de temps, la division pertinente n'est pas celle des populations (entités trop temporaires) mais des espèces. Les populations peuvent fusionner entre elles, mais on considère que les espèces, au contraire, sont isolées les unes aux autres et reliées dans le passé par un graphe en forme d'arbre (voir chapitre 1, figure 1, p. 19). Deux espèces différentes E1 et E2 sont donc assimilables, du point de vue génétique, à deux grandes populations complètement isolées depuis la date de leur divergence, notée  $t$  (âge de l'espèce considérée comme le plus récent ancêtre commun de E1 et E2). Au cours de ce temps, la mutation et la dérive ont agi indépendamment à l'intérieur de chacune de ces deux « super-populations » de sorte qu'elles ont accumulé des différences génétiques. Nous nous intéresserons ici aux différences génétiques fixées entre E1 et E2, c'est-à-dire aux différences telles que tous les individus de E1 sont d'un type, et tous les individus de E2 de l'autre. Ces différences fixées nous renseignent beaucoup mieux sur la divergence à long terme des deux espèces que la variabilité génétique à l'intérieur de E1 ou de E2, qui est

comme on l'a vu de nature *transitoire et sans cesse renouvelée*.

À quel rythme s'accumulent les divergences ? Supposons un locus particulier dans une espèce d'effectif  $n$  quelconque. À chaque génération, de nouvelles mutations apparaissent avec un taux (très faible)  $\mu$ . Le nombre de nouveaux mutants à chaque génération sera donc en moyenne  $\mu n$ . La théorie de la dérive nous dit qu'un allèle apparu par mutation a une chance faible de se fixer, égale à sa fréquence de départ. Chaque mutant peut donc envahir la population avec une probabilité  $1/n$  puisqu'il apparaît dans un seul des  $n$  gènes de la population. *Il en résulte que le taux de fixation au long terme de nouvelles mutations est égal à  $1/n \times \mu n = \mu$ .* Ce résultat est remarquable : *les fixations s'accumulent à un taux constant* qui ne dépend pas de l'effectif de la population. Si toutes les mutations sont nouvelles, c'est-à-dire qu'elles sont suffisamment rares pour ne jamais arriver deux fois sur le même locus, le nombre de différences fixées entre l'espèce E1 et son espèce-ancêtre est donc  $\mu t$ , où  $t$  est l'âge de cet ancêtre. En sachant que l'espèce E2 a accumulé le même nombre de mutations fixées pendant ce temps, et qu'il n'y a jamais deux fois la même mutation, *le taux de différences fixées entre E1 et E2 est  $2\mu t$ .* On peut donc *estimer la date de divergence t entre deux espèces grâce au nombre de différences fixées.* Ce nombre peut, par exemple, être déterminé en utilisant des séquences d'ADN ou de protéines. Ce phénomène, dit *d'horloge moléculaire* a déjà été évoqué à propos des mutations (chapitre 2). Nous avions alors raisonné comme si les espèces étaient des individus, en se préoccupant uniquement du nombre de mutations accumulées dans un génome, et non du fait que ces mutations envahissent la totalité de l'espèce (fixation). Pourtant, un taux

constant de mutation ( $\mu = \text{constante}$ ) ne garantit pas un taux de fixation constant. En effet, même si  $\mu$  ne varie pas, les espèces d'effectif faible produiront nécessairement moins de nouveaux mutants que les espèces très abondantes. Pour que les espèces d'effectif faible changent aussi vite que les espèces abondantes, il faut que la proportion de nouveaux mutants parvenant à se fixer soit plus importante dans les premières que dans les secondes. Comme le montre le calcul ci-dessus, la dérive génétique prédit une telle compensation. Un des grands succès de la théorie de la dérive est donc son aptitude à expliquer simplement l'horloge moléculaire. La dérive explique les changements de fréquences alléliques à court terme mais permet également de retracer l'histoire des divergences anciennes entre espèces.

## Résumé

En l'absence de sélection naturelle, la dérive génétique est un phénomène de variation aléatoire de la composition génétique des populations d'autant plus prononcé que leurs effectifs sont restreints. La dérive élimine à moyen terme la diversité créée par les mutations à l'intérieur des populations, mais elle tend à augmenter les différences entre populations, et, à long terme, entre espèces. La dérive a été admise comme facteur d'évolution au même titre que la sélection naturelle quand les populations ont été analysées au niveau moléculaire. L'abondance de la variation moléculaire à l'intérieur des espèces suggérait, en effet, que beaucoup de mutations étaient neutres du point de vue de la sélection naturelle, c'est-à-dire ni avantageuses ni délétères, et que leur devenir était donc déterminé entièrement par la dérive génér-

tique. La variation moléculaire a permis de vérifier partiellement la validité des équations théoriques de la dérive génétique dans les populations naturelles, et de les utiliser pour inférer des propriétés de la structure et de l'histoire récente de ces populations. À travers l'hypothèse de l'horloge moléculaire, la dérive a aussi permis d'expliquer l'accumulation de différences génétiques entre les espèces à l'échelle de plusieurs millions à centaines de millions d'années, et de dater la divergence des lignées évolutives.



## LA SÉLECTION NATURELLE

Le concept d'évolution, c'est-à-dire de changement graduel des espèces au cours du temps, a été proposé avant Darwin, notamment par Lamarck. L'apport majeur de Darwin<sup>1</sup> est l'idée que la *sélection naturelle*, ou survie du plus apte, est un moteur de cette évolution. Le raisonnement est simple : les espèces sont constituées d'individus en nombre fini car les ressources disponibles sont elles-mêmes limitées. Un individu plus apte que les autres à utiliser ces ressources laissera plus de descendants à la génération suivante. Si cette aptitude est héréditaire, elle sera transmise aux descendants. Il s'ensuit une évolution vers des individus de plus en plus adaptés à leur environnement. Cela permet d'expliquer pourquoi les changements des organismes vivants au cours du temps conduisent à des *adaptations*. L'argumentation de Darwin en faveur de la théorie de la sélection naturelle s'appuie sur deux idées centrales : le raisonnement malthusien et la sélection artificielle.

À la charnière du XVIII<sup>e</sup> et du XIX<sup>e</sup> siècle, les travaux de Thomas R. Malthus<sup>2</sup> concernent la démographie des

1. Ch. Darwin, *L'Origine des espèces*, Paris, GF-Flammarion, 1992 (1<sup>re</sup> éd. 1859).

2. T. R. Malthus, *An Essay on the Principle of Population*, 1798, réédité dans *On Population*, textes rassemblés par G. Himmelfarb, New York, Modern Library, 1960.

populations humaines. À partir d'un calcul simple, il montre que ces populations ont un potentiel de croissance infini. Admettons, par exemple, qu'en l'absence de contrôle, chaque couple a en moyenne trois descendants. La taille de la population est alors multipliée par  $3/2$  à chaque génération. Pour une taille de population initialement égale à  $n_0$ , nous obtenons à la première génération  $n_1 = (3/2)n_0$ , et au bout de  $t$  générations,  $n_t = (3/2)^t n_0$ . La croissance de la population est dite *géométrique* et tend très rapidement vers une taille infinie. Si la taille réelle des populations reste finie, raisonne Malthus, c'est que les ressources en espace et nourriture sont limitées, et qu'un certain nombre de descendants, privés de ressources, ne survivent pas jusqu'à l'âge adulte. Le niveau de ressources détermine donc la population maximale viable, malgré le potentiel d'augmentation indéfini de cette population.

Aux concepts malthusiens, Darwin ajoute celui de variation populationnelle. La population est constituée d'individus uniques qui diffèrent par des variations hérétiques. Étant donné que la croissance des populations est limitée, l'augmentation en nombre des variants favorables s'accompagne forcément de la disparition d'autres variants. Par le biais de cette *sélection naturelle*, la fréquence des variants favorables augmente, et les espèces peuvent évoluer. Cette conception a permis à Darwin d'abandonner la conception *essentialiste*, selon laquelle chaque espèce est définie par un type, les variations observées entre individus étant réduites à des « accidents » sans conséquence.

Pour illustrer sa théorie, Darwin s'appuie sur les effets de la *sélection artificielle* et de la *domestication*. Il prend notamment l'exemple des pigeons, et montre que la diversité des formes domestiques, toutes issues d'une même espèce ancestrale (le pigeon biset), a été obtenue

par les éleveurs en ne faisant reproduire, de génération en génération, que les quelques individus qui satisfaisaient leurs critères esthétiques. Des formes exubérantes, comme le pigeon paon, pourvu d'une queue élargie et capable de faire la roue, ont ainsi été produites. Le même raisonnement s'applique aux autres animaux de compagnie. Les chiens fournissent un exemple très spectaculaire, car les différentes races ont été sélectionnées en grande partie sur des critères esthétiques fantaisistes, variables d'une société à l'autre et d'une époque à l'autre. L'évolution des animaux et plantes domestiqués pour l'agriculture a subi une sélection artificielle plus « rationnelle », sur des critères de rendement (voir *Amélioration génétique des espèces*, p. 315). Cette action de l'homme était essentiellement inconsciente, guidée par une intuition pragmatique : à la manière de Monsieur Jourdain, les hommes du Néolithique, inventeurs de l'agriculture, faisaient de l'amélioration génétique sans le savoir. De nos jours, la théorie de la sélection naturelle s'est développée et ce chapitre en expose les grandes lignes.

### Une vision démographique de la sélection naturelle

La sélection naturelle dépend de « performances démographiques individuelles ». Afin de comprendre l'action de la sélection naturelle, prenons un exemple simple, celui d'une bactérie, organisme unicellulaire se reproduisant de façon asexuée, par simple division de la cellule. On peut dans un milieu de culture artificiel élever une bactérie (A) unique qui, par divisions cellulaires successives, forme un *clone*, c'est-à-dire une population de cellules génétiquement identiques. Le milieu de culture est une gelée contenant des éléments nutritifs,

préparée dans une boîte circulaire (boîte de Pétri). Au début, les bactéries se multiplient rapidement et recouvrent peu à peu toute la surface de la boîte. Puis, la population cesse de s'accroître. C'est à ce stade que le raisonnement malthusien évoqué plus haut s'applique. La limitation des ressources est telle que le nombre de bactéries qui meurent est exactement compensé par le nombre de bactéries qui naissent : chaque mort bactérienne libère un « trou » dans les ressources, immédiatement occupé par une jeune bactérie issue de division cellulaire. Les jeunes bactéries qui ne trouvent pas de place pour s'installer dépérissent.

Supposons maintenant que la population contienne deux souches de bactéries, A et B, identiques en tous points, sauf que les A se divisent deux fois plus que les B. Chaque fois qu'une bactérie meurt et laisse un trou, chacune des bactéries vivantes a une petite chance d'être en phase de division à ce moment là, et de produire une bactérie fille qui ira opportunément s'installer dans le trou. Mais une bactérie de type A a deux fois plus de chances qu'une bactérie de type B d'être en division au bon moment. La bactérie A a donc une probabilité deux fois supérieure d'occuper la place, au fur et à mesure qu'elle se libère. La population de bactéries se renouvelera peu à peu et la proportion de A augmentera progressivement : la sélection naturelle fait son œuvre.

Est-il nécessaire pour cela qu'il y ait limitation par les ressources ? Imaginons que ce n'est pas le cas et prenons nos bactéries au moment où il reste encore largement assez de place dans la boîte de Pétri pour que toutes les jeunes bactéries produites par division puissent s'installer. Les deux types de bactéries ne se reproduisent pas à la même vitesse : pendant que l'une (B) se divise une fois, l'autre (A) fait deux cycles de division. Ainsi, pendant une unité de temps, la population de

bactéries B double alors que la population de A quadruple : la proportion de A augmente donc au cours du temps, et très rapidement, les A auront une majorité écrasante. Contrairement à l'intuition initiale de Darwin, la limitation malthusienne des ressources n'est donc pas une condition nécessaire à l'action de la sélection naturelle, si l'on définit celle-ci comme une modification déterministe des fréquences relatives des types au sein d'une population. La base de cette sélection réside simplement dans une variation héritable d'une « performance démographique individuelle », ici, la vitesse de division des bactéries. Nous allons maintenant modéliser et généraliser ce concept.

Le modèle général de la valeur sélective permet de quantifier l'action de la sélection naturelle. Ce modèle joue un rôle central dans la formalisation de la théorie de l'évolution (voir chapitre 1). On suppose pour simplifier que les générations sont bien distinctes (par exemple, toutes les bactéries se reproduisent en même temps, et une génération représente l'intervalle de temps entre deux épisodes de reproduction). La *valeur sélective* est une variable qui décrit l'aptitude d'un variant génétique à se perpétuer. Elle dépend de deux composantes majeures : survie et fécondité. Le nombre de descendants moyen d'un variant dépend en effet de sa probabilité de survivre jusqu'à l'âge de reproduction, et du nombre de descendants qu'il produit une fois cet âge atteint.

Dans le cas des bactéries, nous définirons la valeur sélective du variant A,  $w_A$ , comme le nombre moyen de descendants laissés à chaque génération par une bactérie A, et  $w_B$  comme celui d'une bactérie B. En supposant qu'il y a au départ  $N_A$  bactéries A et  $N_B$  bactéries B, la fréquence de A est  $p = N_A/(N_A + N_B)$ . Après une génération, nous aurons  $w_A N_A$  bactéries A et  $w_B N_B$  bactéries B ; la fréquence de A sera alors :

$$p' = w_A N_A / (w_A N_A + w_B N_B) = w_A p / W$$

où  $W = w_A p + w_B (1 - p)$  représente la valeur sélective moyenne de la population. Cette équation est équivalente à l'équation générale  $dp/dt = p (w_A - W)/W$  donnée dans le chapitre 1. En réitérant ce processus, on peut prévoir l'évolution de la fréquence sur de nombreuses générations. La connaissance de  $w_A$  et  $w_B$  est donc suffisante pour prévoir l'action de la sélection naturelle. Il n'est même pas nécessaire de connaître  $w_A$  et  $w_B$  exactement, mais simplement le rapport entre les deux. Ainsi, dans l'exemple précédent, les taux de division des bactéries ne sont pas connus numériquement, mais on sait que  $w_A = 2w_B$ , ce qui donne  $p' = 2p/(1 + p)$ , qui est parfaitement déterminé. Nous n'avons donc besoin que de valeurs sélectives relatives, c'est-à-dire de quantités proportionnelles aux nombres moyens de descendants laissés par les différents variants en une génération. Dans le cas des bactéries, on peut donc poser de manière équivalente  $w_A = 2$  et  $w_B = 1$  ou  $w_A = 1$  et  $w_B = 1/2$ .

Le modèle de la valeur sélective est très général et s'applique à toutes sortes d'organismes. Il peut également s'adapter au cas de générations non séparées (reproduction asynchrone comme dans les populations humaines) au prix de calculs supplémentaires, le principe restant le même. Nous allons maintenant montrer que ce modèle peut être testé expérimentalement et donc utilisé pour comprendre des observations faites *in natura*.

**Peut-on reproduire artificiellement l'action de la sélection naturelle ?**

**L'évolution par mutation et sélection naturelle *in vitro*.** La théorie de l'évolution stipule que les populations

s'adaptent par le jeu de la mutation et de la sélection : des mutations qui permettent d'exploiter plus efficacement le milieu se produisent aléatoirement chez des individus isolés, et se fixent ensuite dans l'ensemble de la population grâce à la sélection naturelle. Dans ce cadre, le modèle de la valeur sélective prédit que, si on laisse une population dans un milieu constant pendant longtemps, sa capacité à exploiter le milieu augmente : la valeur sélective des descendants doit être supérieure à celle de leurs ancêtres. Pour tester cette hypothèse, une expérience ambitieuse a été menée par l'équipe de Richard E. Lenski<sup>1</sup> à l'aide de la bactérie *Escherichia coli*. L'avantage de cet organisme est double. Premièrement, il se reproduit rapidement en tube à essai et forme des populations très importantes (plusieurs milliards d'individus), suffisantes pour que des mutations, même rares, finissent par arriver. Deuxièmement, on peut pendant un temps indéfini congeler les bactéries, sans les tuer, et donc stopper leur reproduction. On peut ainsi « ressusciter » un ancêtre congelé pour comparer ses performances avec celles de ses descendants. Douze populations de cette bactérie furent cultivées dans un milieu strictement identique pendant mille cinq cents jours, soit environ dix mille générations. Les douze populations furent fondées par douze bactéries isolées provenant toutes du même clone ancestral. Il n'y avait donc pas de différence génétique entre les douze populations à leur fondation.

1. R. E. Lenski et M. Travisano, « Dynamics of Adaptation and Diversification : a 10,000-Generation Experiment with Bacterial Populations », *Proceedings of the National Academy of Sciences USA*, 91, 1994, p. 6808-6814. Pour une vision globale de cette vaste expérimentation, voir R. E. Lenski, « Evolution in Action : a 50,000-Generation Salute to Charles Darwin », *Microbe*, 6, 2011, p. 30-33.

Le clone fondateur fut par ailleurs conservé par congélation. La question centrale de cette expérience était : *la mutation et la compétition entre individus conduit-elle à une augmentation de la valeur sélective des individus ?*

La valeur sélective relative des différentes populations fut calculée en les mettant en compétition avec la population ancestrale. On estime ainsi la différence de taux de croissance entre bactéries « évoluées » et bactéries « ancestrales ». Au départ (première génération), la valeur sélective de toutes les populations était la même. Au fil des générations, la valeur sélective des douze populations augmenta par rapport à celle de la population ancestrale. La réponse à la question posée est donc positive. Ce résultat est une preuve expérimentale de la sélection naturelle car il ne dépend pas d'un choix délibéré d'individus particuliers comme dans les cas de sélection artificielle. La seule contrainte imposée aux populations était un environnement fixé (milieu de culture, température, conditions de laboratoire). Pour reprendre la terminologie darwinienne, l'augmentation de la valeur sélective observée pour toutes les populations de l'expérience prouve que la compétition entre des individus ayant des traits héritables différents préserve à long terme le type le plus apte dans ce milieu.

Cependant, les douze populations n'avaient pas toutes la même valeur sélective à la fin de l'expérience. Ces différences prouvent qu'une même sélection peut aboutir à des résultats différents. En effet, les mutations favorables apparaissent aléatoirement et chaque population a donc eu sa propre histoire de mutation-sélection. Le fait que toutes les lignées améliorent leur valeur sélective montre que la sélection est, dans une certaine mesure, un phénomène déterministe. Cependant, le hasard des

mutations est suffisant pour que l'on ne puisse pas prédire avec exactitude l'avenir d'une lignée particulière.

**Peut-on généraliser ces résultats à des organismes plus complexes ?** On ne peut reproduire les expériences menées sur les bactéries sur des organismes complexes en raison de leur temps de génération. En effet, même la drosophile, choisie par les généticiens pour son temps de génération court, ne réalise qu'une génération tous les dix jours (contre 6, 7 générations par jour pour la bactérie *E. coli* ci-dessus). Or, pour obtenir des mutations favorables spontanées, il faut manipuler des millions d'individus pendant de nombreuses générations. En conséquence, pour observer l'action possible de la sélection, nous ne pouvons attendre, chez un organisme complexe, que des mutations nouvelles engendrent de la variation. La solution la plus simple consiste à former artificiellement une population ne contenant que deux variants génétiques puis à observer les changements de leur fréquence relative. Les premières expériences de ce type furent initiées par Philippe L'Héritier et Georges Teissier<sup>1</sup> dans les années 1930 sur des drosophiles. Ils inventèrent des « cages à populations », systèmes alimentés en nourriture régulièrement, de sorte à maintenir le niveau de ressource (et donc l'effectif des mouches) constant, comme dans une population malthusienne idéale. En partant d'une population comprenant un mélange de drosophiles standard et de drosophiles caractérisées par une mutation défavorable diminuant la valeur sélective (*Bar*, œil déformé), ils observent une

1. Ph. L'Héritier et G. Teissier, « Une expérience de sélection naturelle. Courbe d'élimination du gène "Bar" dans une population de Drosophiles en équilibre », *Comptes rendus de la Société de biologie de Paris*, 117, 1934, p. 1051.

décroissance régulière de la fréquence de ces mutants. Cette décroissance adopte bien la forme de courbe théorique prédite par le calcul fondé sur les valeurs sélectives. Elle reflète la sélection naturelle contre les mouches *Bar*. Des expériences inverses (introduction d'un variant favorable) ont également donné le résultat attendu. Ainsi, Philip Hedrick<sup>1</sup> introduit, dans des populations de quarante drosophiles aux yeux blancs, porteuses d'allèles *white*, des individus aux yeux rouges, porteurs d'allèles standard. L'allèle standard confère aux individus qui le portent une valeur sélective supérieure à celle des individus *white*. L'allèle standard est au départ en fréquence 0,1. Sur un ensemble de vingt populations, la fréquence moyenne de l'allèle standard augmente jusqu'à 0,35 en moins de quatre générations, conformément aux prédictions.

### Et dans la nature ?

**Un cas d'école : la résistance des moustiques aux insecticides.** L'homme a développé au cours du XX<sup>e</sup> siècle des opérations de démoustication fondées sur une lutte chimique (insecticides). Dans ce cadre, la zone littorale du Languedoc, gîte favori du moustique *Culex pipiens*, a été abondamment traitée par des insecticides organophosphorés à partir de 1968. Ce type d'insecticide tue les larves de moustique en altérant le fonctionnement de l'*acétylcholinestérase*, une enzyme clé dans le fonctionnement du système nerveux. À partir de 1972, les moustiques languedociens sont devenus résistants à ces

1. Ph. Hedrick, « Selection in Finite Populations, III. An Experimental Examination », *Genetics*, 87, 1980, p. 297-313.

insecticides, suite à l'apparition et à la sélection d'une série de nouveaux variants génétiques<sup>1</sup>. L'analyse moléculaire a permis de mieux comprendre les mécanismes à l'origine de ces résistances. La première résistance apparue était due à une mutation affectant la production d'une enzyme appelée *estérase*. Cette enzyme a la propriété de détoxifier la molécule d'insecticide. Elle est au départ présente chez tous les moustiques, mais pas en quantité suffisante pour résister aux doses d'insecticide reçues. La mutation conduisait à une *surexpression* de ce gène : la quantité d'estérase était alors suffisante pour augmenter le taux de survie des moustiques soumis au traitement. L'activité estérase fut affectée en 1984 par une nouvelle mutation. Cette dernière permettait une augmentation encore plus forte de la quantité d'enzyme, grâce à une *amplification* (c'est-à-dire une série de duplications) directe de deux gènes voisins codant chacun pour une estérase. Un grand nombre de copies du gène étaient ainsi disposées côte à côte sur l'ADN.

Entre-temps (1977), était apparue une troisième mutation sur un gène différent, le gène de l'acétylcholinestérase, c'est-à-dire la cible même des insecticides. Cette mutation ponctuelle suffisait à rendre l'acétylcholinestérase moins sensible à l'action des organophosphorés. Cette mutation présentait (comme les précédentes) un coût : l'acétylcholinestérase résistante fonctionne moins bien que la forme standard. En l'absence d'insecticide, les moustiques standards sont donc favorisés. Vers 1994, est apparue une dernière mutation qui permet au moustique de contourner ce problème. Cette dernière

1. T. Guillemaud *et al.*, « Evolution of Resistance in *Culex pipiens* : Allele Replacement and Changing Environment », *Evolution*, 52, 1998, p. 443-453.

mutation correspond à une duplication grâce à laquelle le gène standard et le gène résistant sont associés (côte à côte sur l'ADN). De cette manière, un moustique possédant cette combinaison dispose à la fois du gène de résistance lui permettant de se développer dans un environnement traité, et du gène normal qui restaure sa valeur sélective dans un environnement non traité.

Les quatre variants génétiques décrits sont apparus dans la région de Montpellier successivement, entre 1970 et 1994. On ignore si ces mutations sont nées sur place ou arrivées par migration. En effet, certains allèles de résistance se sont répandus dans le monde entier en quelques années seulement<sup>1</sup>. La migration des moustiques est donc très efficace.

Les mutations (nées sur place ou introduites) sont initialement rares puis augmentent en fréquence rapidement, processus qui a été bien suivi dans la région de Montpellier. Certaines d'entre elles ont ensuite été remplacées par d'autres. Ainsi, l'estérase surexprimée est peu à peu remplacée par les estérases amplifiées et l'acétylcholinestérase résistante par une combinaison résistante-standard. L'exemple des moustiques est riche d'enseignements sur les processus d'adaptation. Premièrement, il montre que l'adaptation à une même pression de sélection peut impliquer différents mécanismes biologiques. Deuxièmement, il illustre la dynamique progressive de cette adaptation. Lors d'un changement brutal de l'environnement (apparition d'insecticides), des mutations à fort effet sur la valeur sélective (invention de la résistance) sont sélectionnées, permettant de se rapprocher à grands pas d'un

---

1. M. Raymond *et al.*, « Worldwide Migration of Amplified Insecticide Resistance Genes in Mosquitoes », *Nature*, 350, 1991, p. 151-153.

phénotype optimum dans cet environnement. Ensuite, si l'environnement reste constant, l'adaptation progresse à petits pas, les premières mutations étant remplacées par d'autres, légèrement plus efficaces. Ce ralentissement a été prédit théoriquement par Ronald Fisher<sup>1</sup> en 1930. Il explique aussi pourquoi l'action de la sélection naturelle est surtout visible dans les cas où l'homme introduit une modification radicale de l'environnement. En effet, cette situation maximise la probabilité d'observer des mutations qui augmentent fortement la valeur sélective. À l'opposé, les populations dont l'environnement est resté stable longtemps ont eu le temps de s'y adapter. Elles sont donc proches de l'optimum sélectif dans cet environnement. Les améliorations par sélection naturelle sont alors rares et de faible amplitude.

**Les bactéries font de la résistance.** Les bactéries, qui ont permis d'étudier l'évolution *in vitro*, évoluent également dans la nature. Un élément important dans leur évolution récente est l'utilisation par l'homme d'antibiotiques. Ce cas est similaire à celui des moustiques, mais les populations de bactéries étant encore plus grandes que celles des moustiques, la probabilité d'apparition d'une mutation favorable est plus grande. Depuis l'utilisation des antibiotiques à des fins médicales, de nombreuses bactéries résistantes sont apparues, particulièrement abondantes dans les milieux hospitaliers, où l'usage des antibiotiques est intense. Ces souches posent des problèmes de santé publique avec la résurgence de maladies, telles que la tuberculose, que l'on ne peut plus traiter par les antibiotiques classiques. Des spécialistes de l'OMS

---

1. R. A. Fisher, *The Genetical Theory of Natural Selection*, Oxford, Oxford University Press, 1930.

ont suggéré que les résistances disparaîtraient d'elles-mêmes si on diminuait fortement la fréquence d'utilisation des antibiotiques, ce qui diminuerait de manière générale leur avantage sélectif. Pour traiter les patients, il faudrait en outre utiliser conjointement plusieurs antibiotiques de types différents : la probabilité d'apparition simultanée de plusieurs mutations de résistance est en effet quasi nulle. Ce raisonnement implique qu'en l'absence de l'antibiotique, la bactérie résistante disparaît, donc qu'elle a une valeur sélective diminuée par rapport à une bactérie non résistante. Les expériences *in vitro* montrent effectivement que le taux de croissance, le pouvoir de multiplication dans un animal ou la survie dans un environnement artificiel dépourvu d'antibiotique, sont souvent plus faibles chez les souches résistantes. Ce phénomène, semblable à celui observé pour la résistance des moustiques aux insecticides, est appelé « coût de la résistance ». Cependant, il a été montré que ces résultats ne sont pas toujours transposables aux conditions naturelles. Deux souches de salmonelles (*Salmonella typhimurium*), résistantes à la rifampicine, ont toutes deux un taux de croissance réduit par rapport à la souche sensible *in vitro*<sup>1</sup>. Toutefois, *in vivo*, en l'occurrence inoculée à un lapin, l'une d'elles s'est révélée aussi prolifique que la souche sensible. Comme dans le cas de la résistance aux insecticides, différentes voies conduisent à résister à un même agent de sélection, mais le coût de ces différentes résistances peut varier. Des mutations ultérieures, sur d'autres gènes, peuvent de plus amoindrir le coût de la résistance initiale. La sélection naturelle agit sur la diver-

1. J. Björkman *et al.*, « Virulence of Antibiotic-Resistant *Salmonella Typhimurium* », *Proceedings of the National Academy of Sciences USA*, 95, 1998, p. 3949-3953.

sité présente en favorisant les génotypes résistants ayant le coût le moins élevé.

Les exemples précédents ont permis aux scientifiques de suivre en direct la sélection naturelle. Dans chaque situation c'est l'action directe de l'homme qui a déclenché ces évolutions en modifiant les milieux. Dans le cas du moustique ou des bactéries, c'est la pression introduite volontairement par l'homme contre eux qui les a conduits à s'adapter pour survivre face à la pression exercée. L'action humaine modifie rapidement et parfois de façon spectaculaire les milieux et crée donc de nouvelles pressions de sélection qui fournissent les exemples les mieux étudiés.

**Peut-on étudier l'action de la sélection naturelle quand ce n'est pas l'action de l'homme qui a modifié les milieux ?** L'exemple des littorines (*Littorina obtusata*), petits escargots des côtes rocheuses d'Amérique du Nord, illustre une évolution rapide en l'absence d'intervention humaine directe. Robin H. Seeley<sup>1</sup> observe que les littorines récoltées dans les années 1980 en Nouvelle-Angleterre (États-Unis) sont de forme plus arrondie et de coquille plus épaisse que celles récoltées à la fin du XIX<sup>e</sup> siècle dans les mêmes sites. Dans le site de Nahant (Maine), où l'on dispose d'un échantillon intermédiaire, on observe que ce changement s'est produit rapidement : entre 1898 et 1915 la hauteur de spire (rapport hauteur/largeur) avait déjà sensiblement diminué. Ce changement coïncide avec l'apparition dans cette région du crabe *Carcinus mænas*, envahisseur en provenance de sites situés plus au sud. Des expériences en nature et

1. R. H. Seeley, « Intense Natural Selection Caused a Rapid Morphological Transition in a Living Marine Snail », *Proceedings of the National Academy of Sciences USA*, 83, 1986, p. 6897-6901.

au laboratoire montrent que ce crabe est un prédateur important des littorines, dont il casse la coquille avec ses pinces. Ces expériences montrent également que les littorines à coquille épaisse et arrondie survivent mieux aux attaques du prédateur, car elles sont beaucoup plus difficiles à casser. Par ailleurs, encore actuellement, les sites où *Carcinus* est le plus abondant sont ceux où l'on trouve les coquilles les plus arrondies. Il semble donc que, à la date de l'apparition de *Carcinus*, les littorines aient subi une sélection naturelle en faveur des formes arrondies, résistantes à la prédation. Ce raisonnement repose néanmoins sur le postulat que la variation de forme est de nature génétique. Il pourrait en être autrement, si cette variation n'est due qu'à une réaction immédiate des littorines, modifiant leur développement pour prendre une forme arrondie lorsqu'elles perçoivent la présence de crabes dans l'environnement. Cependant, cette dernière hypothèse peut être testée. Si elle était vraie, on devrait retrouver des littorines pointues dans les milieux où les crabes sont actuellement rares ou absents. Or, on ne trouve plus de littorines aussi pointues qu'au début du siècle : les formes les plus pointues ont donc été définitivement éliminées par la sélection naturelle au profit des formes arrondies.

### Sélection purifiante

Grâce à l'apparition d'une mutation bénéfique, le processus de la sélection naturelle conduit à une augmentation de l'adaptation des individus d'une espèce à un environnement donné. Toutefois, nous avons vu dans le chapitre 2 que la plupart des mutations ne sont pas bénéfiques mais

au mieux neutres, voire délétères. Dans le cas des mutations sans effet sur la valeur sélective, seule la dérive conduit à l'augmentation ou à la diminution des fréquences dans les populations naturelles (voir chapitre 3). Dans le cas des mutations délétères, la sélection naturelle agit en éliminant les individus mutés.

**L'élimination des mutations délétères est un phénomène symétrique à la fixation des mutations favorables.** Si nous reprenons l'exemple des bactéries A et B, dont les valeurs sélectives sont respectivement 2 et 1, nous pouvons envisager de manière équivalente A comme un mutant favorable dans une population de B, ou B comme un mutant défavorable au sein d'une population de A. L'équation donnant l'évolution de la fréquence de A ( $p$ ) d'une génération sur l'autre détermine également l'évolution de la fréquence de B, qui n'est autre que  $1 - p$ . L'itération de cette équation aboutit à la fixation de A et à la disparition de B. Cette disparition prend un certain temps lorsque B est initialement bien représenté dans la population. En revanche, si B est une nouvelle mutation, il apparaît seulement en un ou quelques exemplaires. Sa fréquence initiale est donc très faible. La sélection aura tôt fait d'éliminer ces quelques bactéries B ou de réduire leur fréquence à une valeur indécelable. La durée de vie des mutations délétères est donc courte, et il est difficile de les observer dans les populations naturelles.

**Le génome est soumis à un flux constant de mutations délétères nombreuses mais rapidement éliminées par une « sélection purifiante ».** Ces mutations délétères sont si fréquentes qu'elles ont servi de matériau de base aux généticiens pour établir les mécanismes génétiques fondamentaux. Historiquement, une étape importante fut l'invention de la *génétique biochimique*, dont le but

était de comprendre la relation entre des mutations génétiques et l'activité biochimique des cellules. Dans les années 1940, George Beadle et Edward Tatum<sup>1</sup> eurent l'idée d'utiliser le champignon *Neurospora* et de tester les performances de plusieurs souches sur des milieux de compositions chimiques différentes. Ils identifièrent tout d'abord un « milieu minimum » ( $M_0$ ), contenant juste assez de composants chimiques pour que les souches standard de *Neurospora* puissent pousser. Ils récupérèrent ensuite des descendants de ces souches et, après les avoir soumis à des radiations de sorte à augmenter les taux de mutations, les cultivèrent pendant plusieurs générations sur des milieux enrichis, par exemple  $M_0$  additionné d'arginine. Le champignon est normalement capable de synthétiser l'arginine, acide aminé nécessaire à sa croissance, à partir des éléments présents dans  $M_0$ . Au bout de plusieurs générations de culture sur  $M_0 +$  arginine, les souches furent repiquées sur  $M_0$ . Certaines d'entre elles ne parvenaient plus à pousser, suggérant qu'elles avaient perdu la capacité à synthétiser de l'arginine. L'analyse génétique montra que chaque souche déficiente avait en fait subi une mutation entraînant l'arrêt du fonctionnement d'une enzyme, participant à la chaîne de réactions chimiques qui produit l'arginine (mutants perte de fonction). De telles mutations sont relativement fréquentes en conditions normales, même si cette fréquence est considérablement augmentée par l'action de mutagènes (radiations par exemple). Ces mutations sont dites *délétères* car elles diminuent la valeur sélective du champignon en l'empêchant de se

1. G. W. Beadle et E. L. Tatum, « Genetic Control of Biochemical Reactions in *Neurospora* », *Proceedings of the National Academy of Sciences USA*, 27, 1941, p. 499-506.

développer sur le milieu minimum. Elles le rendent dépendant de la présence d'arginine dans le milieu ou, en termes scientifiques, *auxotrophe* pour l'arginine. Dans un milieu  $M_0$ , ces mutations se produisent également car la probabilité d'apparition d'une mutation ne dépend pas de son caractère adaptatif ou non (nous avons déjà signalé ce caractère aléatoire de l'apparition des mutations). Toutefois, les mutants auxotropes apparus sur  $M_0$  sont immédiatement éliminés. Cette action *purifiante* de la sélection, permanente dans le milieu minimum, est supprimée dans le milieu enrichi en arginine, ce qui neutralise l'effet délétère des mutations et permet de les détecter.

Le cas de *Neurospora* est une situation expérimentale extrême : la mutation est induite par des rayonnements et ne permet aucune survie dans un milieu normal. Cependant, les populations naturelles de tous les organismes subissent des mutations délétères *spontanées*. Celles qui sont *létales* (en entraînant la mort de l'individu) sont immédiatement éliminées ; d'autres, plus tempérées, diminuent plus ou moins la valeur sélective. Plusieurs expériences menées chez des organismes animaux et végétaux (la drosophile, le ver *Cœnorhabditis elegans*, la plante à fleurs *Arabidopsis thaliana*, le blé) suggèrent qu'en moyenne le nombre de nouvelles mutations qui apparaissent dans un individu en une génération se situe entre 0,01 et 1<sup>1</sup>. Chacune de ces nombreuses mutations ne diminue en général la valeur sélective que d'un faible pourcentage. Cependant, mis bout à bout l'ensemble de ces mutations produit une pression de mutation considé-

1. M. Lynch *et al.*, « Perspective : Spontaneous Deleterious Mutations », *Evolution*, 53, 1999, p. 645-663.

rable qui doit donc en permanence être contrebalancée par une sélection purifiante.

## Sélection naturelle et dérive génétique : l'interaction entre facteurs déterministes et stochastiques

**Les populations naturelles sont soumises à un équilibre entre dérive génétique et sélection naturelle.** Nous avons envisagé jusqu'ici la sélection naturelle comme un phénomène purement déterministe. Pourtant, nous avons vu dans le chapitre 3 que la sélection naturelle n'est pas la seule force à agir sur les fréquences alléliques : les populations naturelles ont toutes un effectif fini, et sont donc soumises à des variations stochastiques (dérive génétique). Nous avons abordé la dérive génétique en considérant des caractères neutres (non soumis à sélection). Que se passe-t-il pour des allèles sélectionnés au sein d'une population finie, quand sélection naturelle et dérive génétique agissent ensemble ?

Les modèles théoriques qui permettent de répondre à cette question combinent les formalismes mathématiques de la sélection et de la dérive. Ils utilisent la *valeur sélective* comme variable indicatrice de l'aptitude d'un génotype à se perpétuer. Ils envisagent également une variation stochastique dans le nombre de descendants. La valeur sélective représente le nombre de descendants *moyen* d'un individu du génotype considéré, mais le nombre *réel* de représentants de ce génotype à la génération suivante est envisagé comme une variable aléatoire suivant une loi binomiale (voir chapitre 3). Le résultat de ces modèles est intuitivement simple : *en moyenne*, la

fréquence allélique obtenue est proche de celle obtenue quand la sélection agit seule. Toutefois, plusieurs populations partant du même état initial n'arrivent pas au même état final : les fréquences alléliques finales varient autour de la moyenne à cause de la dérive. Cette variation autour de la moyenne est d'autant plus importante que la dérive est forte par rapport à la sélection. Elle augmente donc quand l'effectif est faible (dérive forte) ou quand les différences de valeur sélective entre génotypes sont faibles (sélection faible).

Les tests expérimentaux de l'interaction sélection-dérive sont conformes aux prédictions théoriques. Dans l'expérience de Hedrick, déjà mentionnée, un allèle standard (noté +) était introduit dans vingt populations de drosophiles *white*. Sa fréquence augmentait de 0,1 (valeur initiale fixée par l'expérimentateur) à une moyenne de 0,4 au bout de quatre générations, suite à l'action de la sélection. Cependant, l'effectif des populations étant fini (quarante mouches), les modèles d'interaction sélection-dérive prédisent une variation aléatoire entre les populations. Effectivement, certaines populations ont atteint une fréquence proche de 0,7 alors que d'autres ont complètement perdu l'allèle +. La dérive étant d'autant plus forte que les populations sont à faible effectif, l'expérience a été refaite avec des populations de dix individus (au lieu de quarante). Comme attendu, une augmentation des variations autour de la moyenne fut observée.

En général, la dérive conduit à long terme à une mal-adaptation des populations naturelles. Dans l'expérience précédente, on considère un variant favorable qui envahit les populations. Dans un modèle de sélection pure (supposant un effectif infini), la fréquence finale de l'allèle + devrait approcher 100 %. Or, nous avons vu

que la variation aléatoire due à l'effectif fini est telle que certaines populations perdent complètement l'allèle + (fréquence 0). Les mutations étant très improbables à l'échelle de cette expérience, cette perte peut être considérée comme définitive. À long terme, toutes les populations seront soit 100 % (+), soit 100 % (*white*). La valeur sélective de *white* étant inférieure à celle des +, les populations seront donc, en moyenne, moins bien adaptées que si la sélection agissait seule.

Ce principe général peut s'appliquer également aux populations naturelles. Lorsqu'une mutation favorable apparaît, elle est au départ représentée par un faible nombre de copies, quelle que soit la taille de la population. Elle est ainsi vulnérable à la dérive et a une forte probabilité d'être perdue. Dans ce cas, la population a manqué une occasion de s'adapter. Ce phénomène, illustré en situation expérimentale dans l'exemple des drosophiles *white* et +, est très difficile à observer en nature. En effet, il faudrait pour cela observer une mutation favorable qui disparaîtrait par malchance. Or, par définition, de telles mutations n'atteignent jamais des fréquences assez importantes pour être décelées.

La situation inverse est plus facilement observable. Effectivement, de la même façon qu'il est possible de perdre par malchance une mutation favorable, il est également possible de fixer par malchance une mutation faiblement délétère. Une expérience simple consiste à croiser des individus provenant de populations restées longtemps isolées. On prédit que les deux populations ont fixé, par hasard, des mutations délétères à des locus différents. Pour un locus donné, lors du croisement, les mutations apportées par une population sont corrigées par les allèles non mutés de l'autre population. On prédit donc que les hybrides auront, en moyenne, une valeur

sélective supérieure à leurs parents. Ce phénomène de *vigueur hybride* ou *hétérosis*, initialement identifié dans des lignées artificielles de maïs, a été depuis observé sur des populations naturelles de nombreux autres organismes.

## Résumé

La sélection naturelle se manifeste par une augmentation de la fréquence des variants génétiques favorables et une diminution de celle des variants défavorables. Son action peut être modélisée à l'aide de quantités appelées *valeurs sélectives* qui décrivent les performances des différents variants. Les modèles de sélection ont été validés expérimentalement chez différents organismes. Dans la nature, l'action de la sélection naturelle se manifeste de deux façons. (i) L'invasion des populations par des mutations favorables aboutit à une adaptation des organismes à leur environnement. Cette adaptation est facilement mise en évidence lors de changements drastiques de l'environnement consécutifs à l'intervention de l'homme. (ii) Les mutations délétères, qui remettent en cause en permanence l'adaptation des organismes, sont éliminées au fur et à mesure de leur apparition. L'efficacité de la sélection naturelle, dans ces deux mécanismes, est réduite par les variations stochastiques liées à l'effectif fini des populations naturelles.



## L'INDIVIDU, UNE SOCIÉTÉ DE GÈNES EN HARMONIE ?

Peu de notions biologiques semblent aussi accessibles à l'intuition que la notion d'*individu biologique*, qu'on appelle aussi un *organisme*. Cette intuition est anthropomorphe ; elle repose sur la conscience de sa propre individualité par chacun de nous, étendue par analogie à d'autres personnes, puis aux animaux et aux plantes. Sans entrer dans la notion philosophique d'individu, le biologiste cherchera, modestement, à définir l'individu par des propriétés strictement matérielles (gages, à ses yeux, de simplicité et de non-ambiguïté). Dans ce chapitre, nous verrons ce que la théorie de l'évolution peut apporter au concept d'individu biologique. Dans un premier temps, nous chercherons à préciser les contours de cette notion, et nous verrons que, contrairement à ce que suggère l'étymologie, un individu biologique peut être considéré comme un assemblage de parties. La logique évolutive nous conduit alors à envisager la possibilité d'un mécanisme de sélection naturelle à ce niveau, c'est-à-dire une « lutte pour la vie » entre les diverses parties de l'individu. Malgré cela, ces dernières restent associées et semblent coopérer entre elles au lieu d'entrer en compétition. Au cours de ce chapitre, nous mettrons donc la théorie de l'évolution à l'épreuve de ce paradoxe fondamental :

le regroupement de la matière vivante en individus biologiques.

## Qu'est-ce qu'un individu biologique ?

**L'individu est une société de cellules.** L'invention du microscope, à la fin du XVII<sup>e</sup> siècle, a permis à la biologie de faire un extraordinaire bond en avant. Auparavant, la seule connaissance que l'homme avait de son corps (et de celui d'autres animaux) était de nature morpho-anatomique. Les dissections avaient permis de décrire le corps comme un assemblage d'organes auxquels s'attachent des fonctions (par exemple, la respiration pour les poumons). Avec le microscope s'est développée, au XIX<sup>e</sup> siècle, la *théorie cellulaire*, selon laquelle la totalité d'un organisme (animal ou végétal) est découpée en compartiments élémentaires (en général, de 10 à 100 µm de taille) appelés *cellules*. Les éléments permettant de reconnaître une cellule étaient principalement la fine membrane qui l'entoure, et le noyau, corps dense situé à l'intérieur. La théorie cellulaire n'a cessé d'être confirmée depuis, et a été complètement admise au début du XX<sup>e</sup> siècle. Notre connaissance des cellules s'est progressivement améliorée. Nous savons maintenant que la membrane limitante est de même nature chimique dans tous les organismes (composée de deux couches symétriques de molécules appelées *phospholipides*), et que toute cellule contient de l'ADN. En revanche, l'inclusion de cet ADN dans un noyau, ainsi que l'existence d'autres compartiments intracellulaires délimités par une membrane, est une propriété réservée à une lignée évolutive apparue relativement récemment : les *eucaryotes*, dont nous faisons partie avec les animaux, les plantes et les champignons.

Les autres lignées sont dépourvues de noyau (*procaryotes*, comprenant toutes les bactéries au sens large).

Beaucoup d'organismes (eucaryotes et procaryotes) sont constitués d'une seule cellule. C'est le cas de nombreuses bactéries ainsi que des protistes comme les amibes. Cependant, dans les organismes complexes comme l'homme, on trouve plusieurs milliards de cellules. Dans les différentes cellules de l'organisme, deux caractéristiques principales sont reconnues. Premièrement, *elles dérivent les unes des autres par un processus de division cellulaire à partir d'une cellule fondatrice unique*. Deuxièmement, bien que toutes descendent d'une même fondatrice, *les cellules d'un organisme ont des apparences diversifiées*. Ce phénomène, nommé *différenciation cellulaire*, s'acquiert peu à peu, chaque lignée cellulaire adoptant progressivement, au cours des divisions, des caractéristiques propres. La différenciation ne se fait pas au hasard. Les différents types cellulaires correspondent à une *spécialisation fonctionnelle* ; par exemple, les cellules musculaires peuvent effectuer des mouvements de contraction. Les cellules se disposent selon un agencement spatial précis qui permet une *coopération* entre types cellulaires. Ainsi, les cellules musculaires sont accrochées au tissu osseux par des tendons, ce qui permet à la contraction musculaire d'engendrer un mouvement. Cette coopération est telle que certaines cellules se牺牲 pour le bien de l'ensemble, en déclenchant un processus de mort cellulaire programmée. De fait, chez l'embryon, les cinq doigts se forment à partir d'une structure arrondie au départ, dans laquelle des morts cellulaires creusent ensuite quatre interstices.

Les organismes pluricellulaires n'ont pas toujours existé. Les premiers organismes étaient tous des procaryotes unicellulaires de type bactérien. La pluricellularité

est apparue plusieurs fois indépendamment chez plusieurs lignées évolutives eucaryotes (les animaux, les champignons, les plantes) et, à un moindre degré, chez des procaryotes (cyanobactéries). On connaît également des cas de pluricellularité facultative. Les cellules peuvent alors exister séparément ou se regrouper en un organisme cohérent. Par exemple, les amibes du genre *Dictyostelium* peuvent se fusionner en une petite masse compacte capable de mouvements autonomes (surnommée « limace »). Dans ce chapitre, nous nous intéresserons surtout aux organismes pluricellulaires.

**L'individu est une société de gènes.** Quelle est la constitution génétique des cellules d'un individu ? À quelques exceptions près, les divisions cellulaires, nommées *mitoses*, qui engendrent un individu à partir d'une cellule fondatrice sont de type *conforme*. Cela signifie que l'ensemble des gènes d'une cellule (son *génome*) est recopié intégralement dans chaque cellule fille au cours de la division (figure 1). Des divisions non conformes (*méioses*) peuvent néanmoins se produire. Dans ce cas, chaque cellule fille ne reçoit que la moitié du matériel génétique de la cellule mère. Cette dernière est *diploïde* : elle possède, pour chaque locus, deux gènes. Chaque cellule fille reçoit un seul des deux gènes, choisi au hasard, et possède donc un seul gène pour chaque locus (cellule *haploïde*, voir figure 1).

Les cellules filles de méiose quittent en général l'individu qui les produit. Chez les animaux, ces cellules sont des cellules sexuelles haploïdes ou *gamètes* (gamète mâle ou spermatozoïde et gamète femelle ou ovule). Elles fusionnent lors de la *fécondation* pour donner un œuf ou *zygote*, diploïde, qui, par une succession de mitoses, redonne un individu diploïde complet. Chez les végétaux, les cellules filles de méiose sont appelées *méiospores*

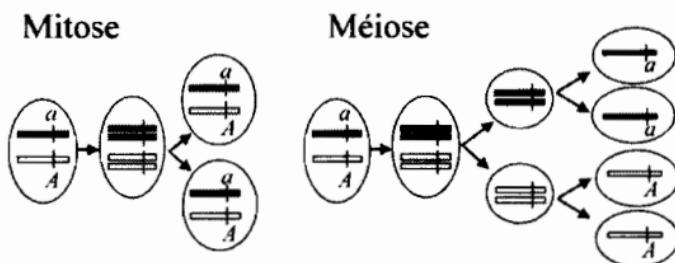
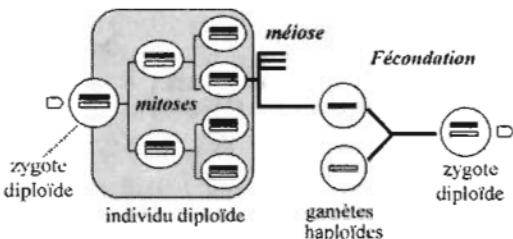


Figure 1 : Les deux grands types de division cellulaire sont illustrés ici. Les ovales représentent des cellules. Les gènes du noyau (seuls représentés ici) sont portés par des chromosomes, représentés par des barres horizontales noires ou blanches. On a illustré le devenir de deux allèles,  $A$  et  $a$ . L'endroit du chromosome (locus) qui contient le fragment d'ADN correspondant à  $A$  ou  $a$  est représenté par une petite barre verticale. Dans les deux types de divisions, la cellule commence par dupliquer tous ses chromosomes. Dans le cas de la mitose, il s'ensuit une seule division, qui donne deux cellules filles chromosomiquement et génétiquement identiques à la cellule mère (reproduction conforme). Dans le cas de la méiose, la duplication est suivie de deux divisions, donnant quatre cellules filles qui ne possèdent que la moitié de l'équipement chromosomique et génétique de la cellule mère. Dans le cas représenté ici, la cellule mère était hétérozygote, c'est-à-dire possédait deux allèles différents ( $A$  et  $a$ ) ; chaque cellule fille de méiose ne possède qu'un de ces deux allèles,  $a$  ou  $A$ , pris au hasard (probabilité 1/2 pour  $a$ , et 1/2 pour  $A$ ).

et se divisent pour former directement un individu haploïde. Cet individu va alors produire des gamètes par mitose. La fécondation entre deux gamètes formera un nouvel individu diploïde (figure 2).

À ce stade, nous pouvons donc donner une définition à la fois génétique et spatiale de l'individu biologique. Il s'agit d'un assemblage d'une ou plusieurs cellules *spatialement solidaires*, et issues par division conforme d'une seule

## ANIMAUX



## VEGETAUX

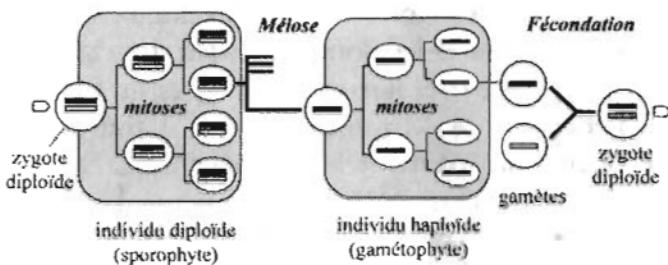


Figure 2 : Les cycles de vie des animaux et des végétaux. Les zones grisesées représentent des individus (cellules restant groupées ensemble). Chez les animaux, on observe un seul type d'individus pluricellulaires (individus diploïdes) car les cellules filles de méiose ne se divisent pas avant la fécondation. Chez les végétaux, il y a deux types d'individus car les cellules filles de méiose se divisent et forment un individu haploïde. Pour simplifier, seules deux divisions, donnant quatre cellules, ont été représentées pour chaque type d'individu. Dans la réalité, les divisions peuvent être beaucoup plus nombreuses, aboutissant souvent à plusieurs milliards de cellules. Les deux gamètes participant à une fécondation proviennent de deux individus différents (dont un seul est représenté).

*cellule*. Chacun des deux critères (génétique et spatial) n'est pas suffisant en lui-même : les deux sont nécessaires *à la fois* pour qu'on puisse parler d'individu biologique. Une cellule peut donc fonder un nouvel individu si elle remplit au moins une des deux conditions suivantes :

1) elle se sépare physiquement de l'ensemble de ses cellules sœurs, 2) elle est le produit d'une division cellulaire non conforme (méiose).

Pénétrons maintenant à l'intérieur d'une cellule pour examiner les gènes, c'est-à-dire les séquences d'ADN qu'elle contient. L'essentiel du génome est contenu dans le *noyau* de la cellule. L'ADN du noyau, emballé sous forme de chromosomes, peut être haploïde ou diploïde selon la cellule considérée (figures 1 et 2). Dans une cellule diploïde, les chromosomes vont par deux : l'un originaire de la mère, l'autre du père. À la façon d'une HLM à deux étages, on peut donc organiser le génome du noyau selon deux axes. L'axe vertical contient les deux étages, correspondant aux chromosomes d'origines paternelle et maternelle. L'axe horizontal représente la succession des séquences d'ADN le long d'un chromosome, à la façon d'une série d'appartements le long d'un couloir. Deux appartements situés l'un au-dessus de l'autre correspondent à deux gènes d'un même locus, ayant des séquences similaires (mais pas toujours identiques). En particulier, si ce sont des séquences codantes, elles permettent à la cellule de produire deux stocks de protéines ayant la même fonction (par exemple deux stocks de globine). En revanche, lorsqu'on marche le long du couloir, on trouve une grande diversité d'occupants (voir chapitre 2) : séquences codantes, séquences régulatrices, pseudogènes et séquences répétées dépourvues de fonction (ADN-poubelle).

Il existe également des parties du génome en dehors du noyau. On les trouve dans deux types d'inclusions intracellulaires : les *chloroplastes*, responsables de la photosynthèse, présents uniquement dans des cellules végétales, et les *mitochondries*, responsables de la respiration cellulaire, dans presque toutes les cellules eucaryotes.

L'ADN des mitochondries et des chloroplastes est court et, contrairement à l'ADN nucléaire, son héritage est généralement uniparental : il provient soit de la mère, soit du père, mais pas des deux.

En conclusion, l'individu est, spatialement, un assemblage de cellules ayant le même génome ; ce dernier est lui-même un assemblage de gènes. La logique de la sélection naturelle (survie du plus apte) peut-elle s'appliquer entre les cellules d'un organisme, ou entre les gènes d'un génome, comme elle s'applique entre les individus d'une population ?

### Compétition ou coopération entre cellules ?

Chez beaucoup d'organismes, y compris l'homme, les cellules destinées à la production de gamètes forment une lignée cellulaire à part, dite *lignée germinale*, définie très tôt dans le développement embryonnaire. August Weismann, au début du XX<sup>e</sup> siècle, a compris le premier l'importance de cette ségrégation précoce de la lignée germinale, toutes les autres cellules de l'embryon formant la *lignée somatique*. Ces dernières se multiplient mais leur descendance reste confinée à l'intérieur de l'individu, et meurt avec lui. Ces cellules somatiques ne participent jamais à la formation de gamètes. Elles ne laissent donc pas de descendantes à la génération suivante. En revanche, elles participent à la survie de l'individu, en constituant des organes comme les poumons qui lui permettent de respirer, ou les muscles qui lui permettent de bouger. Comme le suggérait Weismann, la lignée somatique forme ainsi une sorte d'étui très élaboré, qui protège les cellules germinales et assure leur pérennité.

Ce comportement des cellules somatiques semble être contraire au principe de la sélection naturelle. Des cellules altruistes, sacrifiant leur descendance pour le bien d'autres cellules, devraient être éliminées par la sélection, qui privilégie par définition les variants qui laissent le plus de descendants. La société de cellules qui constitue un individu devrait donc être envahie par des lignées cellulaires égoïstes qui maximisent leur propre transmission, fût-ce au détriment de l'ensemble de la société.

Un exemple de lignée égoïste est donné par les cellules cancéreuses. Celles-ci apparaissent au sein de cellules somatiques, lorsqu'une mutation (ou plutôt, un ensemble de mutations) vient par hasard perturber la régulation de la division cellulaire. La cellule mutée se met à se diviser anarchiquement, produisant des descendantes qui font de même. La lignée cellulaire ainsi formée prolifère sur place, formant une tumeur, et peut finir par coloniser l'ensemble du corps, sous forme de métastases. L'exemple du cancer est un cas pathologique : il met en évidence l'existence, chez les individus sains, de mécanismes forts qui restreignent la prolifération des cellules. L'évolution d'une telle restriction est facilement expliquée par la sélection naturelle, dans ce cas. En effet, les cellules qui ne se restreignent pas (cellules cancéreuses) ont un avantage à court terme sur les autres cellules au sein de l'individu. Elles l'envahissent certes rapidement, mais elles ne peuvent coloniser la lignée germinale. La survie des cellules cancéreuses est donc de courte durée, puisqu'elles entraînent la mort rapide de l'individu sans pouvoir se transmettre à ses descendants. On pourrait cependant concevoir d'autres types de variants égoïstes qui n'auraient pas cette limite : une lignée cellulaire somatique pourrait par exemple entrer en compétition pour l'accès à la lignée germinale. Elle

pourrait par ce biais se perpétuer jusqu'à la génération suivante sans nécessairement entraîner la mort de l'individu.

Pourquoi toutes les lignées cellulaires du corps laissent-elles à une seule d'entre elles le privilège de se perpétuer ? Pourquoi les cellules musculaires ou sanguines n'envoient-elles pas leurs descendantes coloniser la lignée germinale ? C'est en fait la génétique qui détermine le comportement « altruiste » des cellules somatiques. Celles-ci, selon la définition même de l'individu, sont effectivement issues par divisions conformes d'une même cellule fondatrice, et sont donc génétiquement identiques, aux mutations près. À court terme, ces cellules ne sont donc pas en compétition. Nous pouvons exprimer cela de deux façons. *Du point de vue d'une cellule*, cela ne change rien que les cellules de la lignée germinale soient ses propres descendantes ou celles d'autres cellules, puisqu'elles seront de toute façon des copies conformes d'elle-même. *Du point de vue d'un gène*, les cellules ne sont que des véhicules permettant de se transmettre. Toutes les cellules de l'organisme contiennent les mêmes gènes et représentent des véhicules interchangeables. Il peut être alors intéressant de spécialiser certaines lignées cellulaires pour la fabrication des gamètes (lignée germinale) et d'autres pour assurer la survie et la protection des précédentes (lignée somatique).

Ainsi, la sélection naturelle, mécanisme de tri des gènes les plus adaptés, n'agit pas entre les cellules d'un même individu, car toutes ces cellules ont les mêmes gènes. La similarité génétique entre cellules est donc la raison profonde de leur cohésion au sein d'un organisme. Dès qu'une cellule acquiert une composition génétique différente des autres, à la suite d'une division non conforme (méiose) ou d'un événement de fécondation,

cette cohésion est perdue : la cellule se sépare et fonde un nouvel organisme.

Ce raisonnement a-t-il une valeur universelle ? Le plus souvent, des cellules d'un même organisme ont exactement les mêmes gènes. Cependant, il est intéressant de rechercher des exceptions à cette règle pour confirmer le raisonnement utilisé. On s'attend, si notre raisonnement est juste, à observer de la sélection naturelle entre lignées cellulaires dès lors que des différences génétiques existent au sein d'un même individu. Ces dernières ne peuvent apparaître qu'à la suite de mutations intervenues au cours des divisions conformes qui forment l'individu à partir du zygote. Nous avons déjà vu un exemple de telles mutations, avec le cas des cellules cancéreuses. Comme prévu, on observe alors une sélection naturelle entre lignées cellulaires au sein de l'individu, puisque la lignée obtenue adopte un comportement de compétition (et non de coopération) par rapport aux autres lignées.

Un autre exemple est donné par les *lymphocytes*. Ces cellules sanguines ont pour fonction de reconnaître une molécule étrangère (*antigène*) et de produire une molécule dite *anticorps* capable de se fixer spécifiquement à l'antigène de sorte à le détruire. Elles permettent ainsi de lutter contre des pathogènes comme les virus. Il existe, chez un mammifère adulte, des milliers de lignées lymphocytaires, chacune spécialisée dans la lutte contre un antigène donné. Cette spécialisation repose sur des différences génétiques entre les lignées : au cours de la formation des lymphocytes, à partir de cellules souches identiques, chaque lignée subit des mutations particulières dites *réarrangements génétiques*, fondées sur un découpage-collage aléatoire de l'ADN codant pour les

anticorps. En fonction de la manière dont le découpage-collage a été fait, l'anticorps produit se fixera sur tel ou tel antigène.

Chaque lignée est au départ représentée par quelques lymphocytes seulement (un nombre insuffisant pour lutter contre une attaque virale). Lorsqu'un virus attaque, on observe une *activation clonale* de la lignée lymphocytaire produisant les anticorps spécifiquement dirigés contre ce virus : les lymphocytes de cette lignée se mettent à proliférer de sorte à former une population cellulaire suffisante pour lutter contre l'infection. Du point de vue des lymphocytes, cette activation clonale représente un *événement de sélection naturelle entre lignées*, par lequel une lignée devient prépondérante sur les autres du fait d'une reproduction cellulaire accrue, l'agent de cette sélection étant l'antigène (ici, les molécules du virus). Nous voyons que dans ce cas, la compétition entre lignées cellulaires se fait pour le bien des autres cellules, car l'organisme entier est ainsi préservé de l'infection. Contrairement aux mutations cancéreuses (qui sont pathologiques), les mutations qui arrivent au cours de la formation des lignées lymphocytaires font partie du programme de développement normal.

On pourrait cependant s'attendre à ce que ces mutations, source d'une diversité génétique considérable, entraînent par hasard la formation de lignées égoïstes, et non seulement de lignées de lymphocytes destinées à protéger le reste de l'organisme de l'infection. C'est effectivement le cas : on observe au cours du développement l'apparition de lignées produisant des anticorps, non dirigés contre des molécules étrangères mais contre des molécules normalement produites par l'individu lui-même (le *soi*). Si ces lignées « égoïstes » se développaient, elles mineraient l'organisme de l'intérieur car elles détruirait

d'autres cellules. C'est ce qui arrive dans le cas de maladies dites *auto-immunes* (par exemple, la spondylarthrite rhumatoïde ou certains diabètes). Comme dans le cas du cancer, la nature pathologique de ces événements met en relief l'existence chez les organismes sains de mécanismes de contrôle forts. Pour les lymphocytes de mammifères, ce contrôle a lieu dans l'organe où sont produits les lymphocytes, le thymus. On observe dans cet organe une phase dite d'*éducation* : chaque lymphocyte est testé et impitoyablement éliminé s'il produit un anticorps dirigé contre le *soi*. On voit ici que l'organisme, s'il joue sur la compétition entre cellules pour lutter contre les pathogènes, se prémunit efficacement contre les dangers de cette compétition.

Nous avons vu que la cohésion de l'individu vient de l'homogénéité génétique de ses cellules, qui rend la sélection naturelle inopérante à ce niveau. Cela a été confirmé par le fait qu'on observe de la sélection entre cellules dans les cas (exceptionnels) où celles-ci sont génétiquement différentes. Un autre test peut être fait par un raisonnement symétrique : si des individus sont génétiquement identiques (*clones*), ils devraient tendre à s'associer sous forme de « super-individus ». Les clones naturels sont issus de la reproduction asexuée, dans laquelle, contrairement aux cycles de reproduction sexuée décrits en figure 2, les cellules fondatrices des nouveaux individus sont produites par reproduction conforme (et non par méiose ou par fécondation). Chez les animaux, conformément à notre prédition, il est fréquent que les individus issus de reproduction asexuée restent solidaires et forment une colonie. C'est le cas de nombreux invertébrés marins à vie fixée comme les bryozoaires, les ascidiens, les coraux. Le degré d'interdépendance des individus au sein de la colonie est variable,

mais dans certains cas les individus ne sont plus autonomes et/ou se spécialisent dans des fonctions différentes (par exemple, nutrition et reproduction). Il est alors difficile de dire s'il s'agit d'un « super-organisme » ou de plusieurs individus. Les plantes, quant à elles, pratiquent fréquemment la reproduction asexuée, par exemple par marcottage. Un pied de fraisier ou de ronce peut ainsi engendrer d'autres pieds par l'intermédiaire d'une tige allongée (stolon) qui prend racine à une certaine distance du pied mère. Le stolon constitue une connexion physique par laquelle le pied mère envoie de la sève qui nourrit la plantule fille. Cette connexion peut être perdue dès que la plantule a la capacité de se développer en autonomie. On atteint ici les limites de la notion d'individu : ce qui est un individu devient instantanément deux individus dès que la connexion est rompue.

### Coopération ou conflit entre gènes ?

Nous avons vu que la similarité génétique entre celles-là est le ciment de leur coopération au sein de l'individu. Mais à l'intérieur de ces cellules, la molécule d'ADN est un ensemble de séquences et l'on peut concevoir qu'il y ait compétition entre les divers types de séquences pour occuper un maximum de place au sein de la molécule, ou pour maximiser leur transmission d'une cellule à l'autre.

**Le génome est un ensemble hétéroclite de séquences.** Le nombre de séquences codantes augmente principalement par duplication de gènes (chapitre 2). Cependant, la plupart des séquences codantes ne se dupliquent pas à qui mieux mieux pour occuper la molécule d'ADN (les *transposons* constituent une exception). Les séquences

codantes n'entretiennent pas entre elles des rapports de compétition mais de *coopération*. Chacune d'elles est en effet responsable d'une fonction qui permet une meilleure survie de l'individu, donc de l'ensemble des autres séquences. Comme nous l'avons vu au chapitre 2 dans le cas des hémoglobines, les différents gènes ont des actions complémentaires. Tout produit de duplication qui ne confère pas d'avantages en termes de survie est voué à l'élimination, soit par mutation (il devient alors un pseudogène non fonctionnel), soit par l'action de la sélection naturelle qui élimine les individus qui le portent (s'il représente un handicap, par exemple une dépense inutile d'énergie). L'ensemble des séquences codantes forme donc une *association à bénéfice mutuel*.

**Les gènes mitochondriaux et chloroplastiques fournissent un exemple frappant d'un tel mutualisme.** Les mitochondries et les chloroplastes ressemblent à des bactéries par leur structure, leur séquence, et l'existence d'un équipement de transcription-traduction autonome. Sur cette base, la *théorie endosymbiotique* stipule que la mitochondrie et le chloroplaste dérivent de bactéries qui seraient entrées à l'intérieur de cellules eucaryotes pour y établir une symbiose (c'est-à-dire, une association à bénéfices réciproques). En échange de la protection et de l'humidité offertes par le milieu intracellulaire, la mitochondrie assure la respiration de la cellule, et le chloroplaste sa photosynthèse. Au niveau génétique, il s'agit d'une *association mutualiste* entre deux populations de gènes, des gènes bactériens d'une part, des gènes de la cellule eucaryote d'autre part. Ce mutualisme est particulièrement frappant car ces gènes proviennent d'organismes différents et non de duplications des gènes d'un même organisme.

**Certaines séquences ne participent pas à une association mutualiste.** C'est le cas des séquences sans fonction, qui représentent un volume inutile, souvent considérable, dans le génome (voir chapitre 2). Elles sont un fardeau pour les cellules car elles font perdre du temps et de l'énergie aux cellules lors de la réPLICATION de l'ADN. On peut donc considérer que l'extension de ces séquences représente une invasion du génome par un parasite égoïste, qui profite de la machinerie cellulaire pour se perpétuer, mais sans participer à la survie et l'entretien de cette machinerie (voire, en compliquant son fonctionnement). On envisage ainsi l'existence d'une forme de sélection naturelle qui privilégie les séquences les plus capables de s'étendre au sein du génome. Les séquences privilégiées dans ce cas sont les séquences formées de répétitions d'un même motif (séquences satellites). En effet, comme nous l'avons vu, ces séquences sont des sites particulièrement favorables à des mutations par fragment (de type « recombinaison inégale » ou « duplication ») dont le résultat est d'augmenter le nombre de motifs. Les séquences répétées peuvent ainsi s'accroître progressivement. On suppose, sans l'avoir démontré, que cet accroissement finit par avoir des conséquences néfastes sur la survie des individus, si bien que la sélection naturelle empêche ces séquences de grandir indéfiniment. Un exemple de conséquence néfaste est la maladie humaine dite du « X-fragile », qui est due à un microsatellite (répétition de très courts motifs, dans ce cas un trinucléotide CGG) situé sur le chromosome X. Chez les individus normaux, on trouve six à soixante répétitions du motif. Les individus malades en ont jusqu'à deux cents. Le microsatellite entraîne alors l'inactivation de gènes voisins, ce qui, par un mécanisme

encore inconnu, provoque une série de symptômes (notamment un retard mental).

Les transposons ou éléments mobiles (voir chapitre 2) sont d'autres séquences égoïstes, encore plus spécialisées pour envahir le génome. En effet, elles codent pour des protéines, dont la fonction est précisément de permettre la formation de multiples copies de la séquence et leur insertion n'importe où dans le génome. Ainsi, le transposon nommé élément *P* a envahi en quelques années les génomes de drosophiles sauvages. Les transposons sont en général des gènes parasites qui ne servent à rien à la cellule ou à l'organisme et peuvent même lui occasionner quelques dégâts. Ils peuvent par exemple s'insérer au milieu d'un gène et l'empêcher de fonctionner. Beaucoup de virus agissent à la manière des transposons, en insérant leur ADN au sein du génome de l'hôte. La seule différence est qu'ils finissent par détruire complètement la cellule hôte pour reprendre une vie libre et aller infecter d'autres cellules. Beaucoup de transposons pourraient être des virus qui ont perdu la capacité à sortir des cellules hôtes. Doit-on interpréter cette perte comme une adaptation ultime du parasite, s'installant à demeure pour ménager son hôte ? Quelle que soit l'origine des transposons, il est clair que le moteur de leur évolution n'est pas l'avantage qu'ils confèrent à l'organisme qui les véhicule, mais bien leur aptitude à utiliser celui-ci pour assurer leur réPLICATION et leur transmission.

Le paragraphe précédent décrivait une compétition entre divers types de séquences pour envahir « horizontalement » le génome, c'est-à-dire, pour occuper une longueur maximale sur la molécule. Mais on peut également concevoir une invasion « verticale ». Dans ce cadre, la sélection naturelle favoriserait les gènes qui maximisent

leur transmission lors des divisions cellulaires. Cela ne peut avoir lieu lors des divisions conformes, puisque celles-ci transmettent l'ensemble du génome en bloc aux cellules filles. En revanche, cela peut arriver lors d'une méiose. Le fonctionnement normal d'une méiose (illustré en figure 1) est de répartir aléatoirement les deux copies (maternelle et paternelle) d'un même gène entre les cellules filles. Chacune peut avoir soit la copie d'origine paternelle, soit la copie d'origine maternelle, avec une probabilité  $1/2$ . On a cependant observé des cas de *distorsion méiotique*, où l'une des deux copies a une probabilité de transmission supérieure à l'autre. Chez la souris domestique, on a ainsi mis en évidence un ensemble de gènes, collectivement appelés *haplotype t*<sup>1</sup>. Chez un animal ayant un chromosome porteur de ces gènes et un chromosome non porteur, les produits de méiose contiennent, dans presque 100 % des cas, le chromosome porteur. De fait, l'haplotype *t* a un grand avantage de transmission et tend à envahir les populations de souris. Malheureusement pour elles, cet haplotype *t* est néfaste à l'individu : les hétérozygotes (individus ayant un chromosome porteur et un non porteur) sont normaux, mais les mâles homozygotes (ayant deux chromosomes porteurs) sont stériles. L'haplotype *t* se maintient donc en fréquence moyenne par suite d'un équilibre entre l'avantage de transmission à la méiose et le désavantage de fécondité lié à son effet stérilisant.

**De tels gènes responsables de distorsions méiotiques se retrouvent notamment chez les insectes.** Ils sont particulièrement évidents lorsqu'ils touchent les chromosomes sexuels (chromosomes X et Y), car ils affectent

---

1. L. M. Silver, « Mouse *t* Haplotypes », *Annual Review of Genetics*, 19, 1985, p. 179-208.

alors la proportion relative de mâles et de femelles dans la descendance. Chez la mouche *Cyrtodiopsis dalmanni*, par exemple<sup>1</sup>, il existe des chromosomes distorteurs  $X^d$  par opposition aux chromosomes X normaux. Les mâles de type XY donnent classiquement 50 % de spermatozoïdes Y (donnant des descendants mâles) et 50 % de spermatozoïdes X (donnant des descendants femelles). En revanche, des mâles  $X^dY$  donnent des proportions plus proches de 90 % de spermatozoïdes  $X^d$  et 10 % de Y, soit neuf fois plus de descendants femelles que de descendants mâles. Il existe cependant chez cette espèce des chromosomes  $Y^r$  résistants à la distorsion, qui restaurent une ségrégation équitable chez les mâles  $X^dY^r$ . Les mécanismes par lesquels un chromosome maximise sa transmission aux dépens de l'autre sont mal connus, mais on suspecte la production de molécules toxiques. Ainsi, les spermatozoïdes porteurs de chromosomes  $X^d$  pourraient sécréter un poison qui tue les autres spermatozoïdes, en même temps qu'un contrepoison qui les protégerait eux-mêmes. Quel que soit le mécanisme, la distorsion méiotique est certainement un phénomène fréquent mais difficile à observer. Dans la plupart des cas, le chromosome avantagé envahit totalement la population : l'avantage de transmission est alors perdu puisque tous les chromosomes ont la même aptitude à se transmettre, et on en revient à une méiose équitable classique. Pour pouvoir observer des méioses inéquitables, il faut que le gène distorteur n'arrive pas à envahir toute la population. C'est le cas lorsqu'il est désavantage par la sélection naturelle, soit parce qu'il a des effets délétères (cas de l'haplotype *t*), soit parce qu'il entraîne un dés-

---

I. Presgraves *et al.*, « Sex Chromosome Meiotic Drive in Stalk-Eyed Flies », *Genetics*, 147, 1997, p. 1169-1180.

équilibre des proportions relatives des deux sexes (cas des chromosomes sexuels X<sup>d</sup>).

On peut, dans le même ordre d'idées, s'attendre à l'existence d'un autre conflit, cette fois entre des gènes ayant des modes de transmission différents. Ainsi, les gènes nucléaires sont équitablement représentés dans les produits de méiose mâles et femelles. Dans une plante à fleurs par exemple, chaque gène nucléaire de la plante mère aura une chance sur deux d'être présent, que ce soit dans un grain de pollen (produit mâle) ou dans un ovule (produit femelle). Il n'en va pas de même des gènes mitochondriaux. L'hérédité des mitochondries est exclusivement maternelle, c'est-à-dire que les gènes mitochondriaux sont transmis intégralement par les ovules, mais pas du tout par le pollen. Il y a donc un conflit d'intérêt entre ces deux catégories de gènes. Les gènes mitochondriaux ont intérêt à ce que la plante investisse toute son énergie dans la production d'ovules et ne produise pas du tout de pollen. Les gènes nucléaires ont intérêt à ce que la plante produise du pollen, qui les transmet avec la même efficacité que les ovules, mais qui est relativement plus économique à produire. Le thym (*Thymus vulgaris*) offre un exemple de ce conflit<sup>1</sup>. C'est une espèce dite *gynodioïque*, c'est-à-dire que les individus peuvent être soit hermaphrodites, soit femelles. Les hermaphrodites possèdent des étamines productrices du pollen, et un pistil producteur d'ovules, alors que les femelles n'ont pas d'étamines fonctionnelles. Ces femelles sont aussi appelées « mâle-stériles » vu qu'elles montrent une stérilisation des parties mâles par rapport

1. P.-H. Gouyon et D. Couvet, « A Conflict Between Two Sexes, Females and Hermaphrodites », *The Evolution of Sex and its Consequences*, S.C. Stearns, Birkhäuser Verlag, Bâle, 1987.

au type hermaphrodite classique. On a montré que la stérilité-mâle est due à des gènes à hérédité maternelle, ce qui est une indication de leur origine mitochondriale. Chez le thym, on a en outre mis en évidence l'existence de gènes qui contrecurrent l'action des précédents en restaurant la fertilité mâle. Ces gènes, conformément à l'hypothèse du conflit de transmission, sont à hérédité biparentale, signe de leur localisation nucléaire.

## Résumé

L'individu, et le génome qui met en mémoire ses caractéristiques génétiques, sont des assemblages d'éléments, respectivement des cellules et des gènes. La cohésion de ces éléments n'est pas une donnée intangible, mais le produit historique d'une évolution guidée par la sélection naturelle. Cette sélection privilégie les gènes qui se transmettent le mieux. Le regroupement de cellules génétiquement identiques en individus est un des moyens d'améliorer cette transmission. De même, le regroupement de plusieurs gènes au sein d'un même génome est une association à bénéfices mutuels qui augmente leurs chances de se perpétuer ensemble. Cependant, ces associations ne sont pas exemptes d'éléments égoïstes. Ceux-ci peuvent profiter des bonnes conditions de transmission créées par les autres sans apporter eux-mêmes de contribution à la survie de l'ensemble. Ils peuvent même diminuer cette survie, ou imposer leur transmission au détriment d'autres gènes. Derrière une apparence d'unité et d'harmonie, l'individu et son génome sont donc des lieux de conflit, et on peut parler d'une sélection naturelle entre parties de l'individu.

La mise en évidence de ces conflits est une des grandes avancées de la théorie de l'évolution, liée au développement de la biologie moléculaire qui a permis l'exploration du génome, à partir des années 1960. Les fixistes utilisaient, et les anti-évolutionnistes actuels utilisent encore, la perfection des organismes vivants pour justifier l'existence d'un créateur intelligent. L'imperfection même de ces organismes, objets de manipulations contradictoires par des gènes en conflit, est une illustration éclatante de la puissance des arguments darwiniens et un fait décisif pour rejeter les thèses créationnistes. Par ailleurs, comme l'a signalé Richard Dawkins<sup>1</sup>, elle nous conduit à repenser notre nature en tant qu'individus. Nous ne sommes après tout que des véhicules très perfectionnés, élaborés par les gènes pour assurer leur transmission.

---

1. R. Dawkins, *Le Gène égoïste*, Paris, Armand Colin, 1990 (1<sup>re</sup> éd. 1976).

## UN GÈNE N'ÉVOLUE PAS SEUL

Les individus sont des véhicules que les gènes empruntent pour se propager. Cependant, l'individu n'est pas la voiture particulière d'un gène, c'est plutôt un autobus qui transporte, le temps d'un voyage, de nombreux gènes ensemble. Les gènes forment donc des associations temporaires, appelées *génomes* des individus. À chaque épisode de reproduction sexuée, il y a dissociation partielle des génomes individuels (méiose) et réassocation entre parties de génomes différents (fécondation). Ces associations, bien qu'éphémères, modifient le comportement d'un gène particulier face à la sélection naturelle. Dans les situations examinées au chapitre 4, nous avons raisonné comme si les individus ne différaient entre eux que par un gène : le gène cible de la sélection. Par exemple, on considérait des bactéries pourvues d'un gène de résistance à un antibiotique, et des bactéries qui en étaient dépourvues, sans se préoccuper des autres gènes. Toutefois, ce raccourci ne peut pas représenter fidèlement la réalité biologique. La diversité génétique au sein des populations naturelles est si grande qu'il est très difficile de trouver deux individus qui diffèrent par un seul gène. Loin d'une telle homogénéité, le génome est constitué de nombreux gènes en interaction, voire

en conflit. En quoi ce contexte génomique affecte-t-il l'évolution d'un gène particulier ?

### La diploïdie implique des relations verticales entre allèles du même locus

Les animaux et les plantes possèdent une phase diploïde, souvent prépondérante. Lors de cette phase, chaque locus est représenté par deux gènes, provenant respectivement du père et de la mère. Malgré cela, l'individu n'exprime qu'un seul état du caractère correspondant : ce dernier dépend non du type (ou *allèle*) caractérisant un gène, mais du *génotype* qui caractérise l'assemblage de gènes porté par l'individu. Dans une population diploïde où coexistent deux allèles A et B, un gène de type A peut se trouver dans deux états : en compagnie d'un autre gène similaire (dans un individu de génotype AA dit *homozygote*), ou en compagnie d'un gène différent (dans un individu AB dit *hétérozygote*). Le gène A peut donc être associé à deux phénotypes différents. Nous qualifierons de *verticales* les relations entre A et B, allèles du même locus, par opposition aux « relations horizontales » qui lient entre eux des allèles de locus différents (figure 1).

Recherchons tout d'abord une situation où l'évolution d'un allèle n'est pas affectée par les relations verticales. Dans une telle situation, la valeur sélective est affectée par un gène (par exemple, de type A) de la même manière, quel que soit l'autre gène (A ou B). En d'autres termes, en remplaçant un A par un B, on change la valeur sélective du même montant, que ce soit dans un individu AA ou dans un individu AB. Si nous notons  $w_{AA}$ ,  $w_{AB}$  et  $w_{BB}$  les

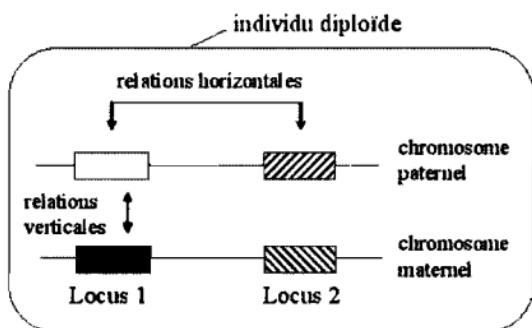


Figure 1 : Relations horizontales et verticales entre gènes dans un individu diploïde. Les rectangles représentent des gènes appartenant à différents allèles, représentés par les figurés (noir, blanc, hachures). L'individu possède deux gènes pour chaque locus, l'un sur un chromosome hérité de son père (chromosome paternel), l'autre sur un chromosome hérité de sa mère (chromosome maternel).

valeurs sélectives des individus AA, AB et BB, cette condition s'écrit :  $w_{AB} - w_{BB} = w_{AA} - w_{AB}$ . En conséquence, l'hétérozygote AB a une valeur sélective exactement intermédiaire entre les deux homozygotes AA et BB. Cette situation est appelée *additivité*. A et B évoluent alors tout simplement comme s'ils étaient chez des haploïdes.

Malheureusement, ce cas reste théorique. Pour trouver des exemples empiriques, il faut rechercher les cas où l'hétérozygote a un caractère intermédiaire entre les deux homozygotes. Ainsi, pour les gènes de pigmentation de certaines fleurs, l'hétérozygote est rose alors que les homozygotes sont respectivement blanc et rouge. Cependant, ce n'est pas parce que l'hétérozygote est intermédiaire pour la couleur qu'il l'est aussi pour la valeur sélective. La valeur sélective dépend de l'attractivité de la fleur pour les insectes. Mais on ignore si l'attractivité des fleurs roses est égale à la moyenne des attractivités des

blanches et des rouges. Globalement, les exemples de gènes à effet purement additif sont rares. En fait, l'additivité est souvent posée comme hypothèse par défaut dans les modèles évolutifs sans autre justification que la facilité des calculs.

Dans la plupart des gènes connus, on observe des relations de hiérarchie entre allèles. Le fondateur de la génétique moderne, Gregor Mendel, a établi ce fait en croisant des petits pois (*Pisum sativum*). Les descendants du croisement entre une lignée à grains jaunes (de génotype  $JJ$ ) et une lignée à grains verts ( $vv$ ) étaient de génotype ( $Jv$ ), mais leurs grains étaient tous jaunes. On qualifie dans ce cas l'allèle  $J$  de *dominant* puisqu'un seul gène  $J$  suffit à donner la couleur jaune. L'allèle  $v$  est dit *récessif*, car la couleur verte n'est pas exprimée en présence d'un  $J$ .

Un exemple intéressant est donné par les *mutants perte de fonction*. Beaucoup de maladies génétiques humaines sont dues à de tels allèles. Ainsi, la  $\beta$ -thalassémie est due à une mutation perte de fonction (souvent une délétion) dans le gène de la sous-unité  $\beta$  de l'hémoglobine  $\alpha_2\beta_2$ , normalement produite à partir de l'âge de quelques mois. Les conséquences de la mutation ne se manifestent que chez les homozygotes possédant deux gènes mutés. Dans ce cas, aucune hémoglobine  $\alpha_2\beta_2$  fonctionnelle n'est produite, ce qui aboutit à une anémie mortelle vers l'âge d'un an. Les hétérozygotes, qui possèdent un gène muté et un gène fonctionnel, sont des porteurs sains : ils peuvent transmettre un gène muté à leurs descendants, mais ils ne souffrent guère d'anémie, car la dose de protéine  $\beta$  produite par un seul gène fonctionnel est suffisante. Ce cas est fréquent : la plupart des mutations délétères sont presque complètement récessives.

**Quel est l'effet de la dominance et de la récessivité des allèles sur la sélection naturelle ?** Prenons le cas d'une maladie récessive comme la  $\beta$ -thalassémie. Le destin de l'allèle muté est de disparaître, vu la force de la sélection naturelle contre lui. Le fait qu'il soit récessif peut toutefois retarder sa disparition. En effet, lorsque la fréquence de l'allèle muté est faible, la probabilité que deux gènes mutés se trouvent ensemble dans le même individu est infime : un gène muté se trouve le plus souvent en compagnie d'un gène sain, chez un individu hétérozygote. Du fait de la récessivité, cet hétérozygote est un porteur sain qui a une valeur sélective égale à celle d'un individu non porteur. Les gènes mutés se protègent donc temporairement de la sélection naturelle en se « cachant » dans un hétérozygote. La récessivité ralentit ainsi la disparition des allèles délétères. Les maladies génétiques observées chez l'homme sont le produit d'un équilibre entre la mutation, qui engendre en permanence des allèles délétères, et la sélection qui les élimine. Les fréquences des allèles mutés sont typiquement de l'ordre de 1/100 ou 1/1 000. Cette fréquence est faible mais bien supérieure à la fréquence des nouvelles mutations à chaque génération. Donc ces gènes mutés, bien que délétères, persistent plusieurs générations avant de disparaître. Cette persistance est permise par leur récessivité.

La récessivité ralentit l'action de la sélection sur un allèle. Ce principe, valable pour les allèles délétères, est également vrai pour un allèle favorable. Lorsqu'une mutation favorable apparaît, elle est nécessairement en faible fréquence (touchant un ou quelques individus parmi toute la population). Elle se trouve essentiellement dans des individus hétérozygotes. Si l'allèle favorable est récessif, les performances de ces hétérozygotes ne seront pas supérieures à la moyenne. Pour envahir la popula-

tion, cet allèle muté devra compter initialement sur le hasard (dérive génétique), jusqu'à ce que sa fréquence soit suffisante pour que des homozygotes mutés apparaissent avec une probabilité non négligeable. La sélection naturelle pourra alors agir, puisque ces homozygotes ont une valeur sélective supérieure aux autres individus. L'invasion d'une population par un allèle favorable sera donc aléatoire et lente s'il est récessif, et, inversement, sûre et rapide s'il est dominant ou semi-dominant. Pour cette raison, on s'attend à ce que les allèles favorables, qui envahissent avec succès les populations naturelles diploïdes, soient plutôt dominants. Dans la plupart des cas où on a identifié de tels allèles, cette prédiction est effectivement vérifiée. Ainsi, les allèles conférant une résistance aux insecticides chez le moustique *Culex pipiens* sont partiellement dominants. En effet, en présence d'insecticide, les hétérozygotes ont une valeur sélective supérieure à celles des homozygotes non résistants. Cependant, les homozygotes résistants supportent des doses d'insecticides plus élevées que les hétérozygotes. Ils survivent donc mieux dans les zones abondamment traitées.

**La diploïdie autorise également des formes originales de sélection qui favorisent la coexistence de plusieurs variants génétiques.** Jusqu'ici nous avons évoqué des cas où l'hétérozygote AB a une valeur sélective inférieure ou égale à celle de l'homozygote le plus performant. Que peut-on attendre lorsque l'hétérozygote a une valeur sélective supérieure à celles des deux homozygotes ? Cette situation, appelée *superdominance*, a lieu lorsque les gènes A et B coopèrent entre eux, et donnent de meilleurs résultats lorsqu'ils sont réunis que pris séparément. Plaçons-nous dans ce cas et imaginons une population contenant beaucoup de A et très peu de B. B sera

alors essentiellement représenté dans les hétérozygotes AB. Ces individus laissent plus de descendants que la moyenne, B va donc augmenter en fréquence. Par un raisonnement symétrique, dans une population où B est en forte majorité, l'allèle A augmente en fréquence. En conclusion, ce mode de sélection favorise un allèle dès qu'il devient rare ; des calculs plus précis (non détaillés ici) montrent qu'il aboutit à stabiliser la population dans un état intermédiaire, où A et B coexistent dans des fréquences qu'il est possible de prédire exactement lorsqu'on connaît les valeurs sélectives des génotypes AA, AB et BB.

L'exemple classique d'une telle sélection stabilisante est l'anémie falciforme<sup>1</sup>. L'anémie falciforme est une maladie génétique due à une mutation ponctuelle dans la chaîne  $\beta$  de l'hémoglobine. L'allèle muté est noté S. La solubilité de l'hémoglobine S est diminuée, et elle tend à cristalliser dans les globules rouges du sang. Les globules déformés prennent alors l'aspect d'une lame de faux (d'où le nom de la maladie). Ils engorgent les capillaires sanguins, provoquant chez les homozygotes SS une anémie sévère qui aboutit le plus souvent à la mort. Cette maladie génétique existe en très faible fréquence dans la plupart des populations humaines. Elle est récessive : les hétérozygotes AS (où A désigne l'allèle normal) ont une espérance de vie semblable aux homozygotes normaux AA. La prévalence de la maladie est cependant beaucoup plus forte en Afrique tropicale. Cette prévalence est géographiquement corrélée à celle du paludisme, grave maladie provoquée par le parasite unicellulaire *Plasmodium falciparum*. De fait, un lien existe entre l'anémie falciforme et le

1. L. L. Cavalli-Sforza et W. F. Bodmer, *Genetics of Human Populations*, San Francisco, W. H. Freeman, 1971.

paludisme : *Plasmodium* parasite les globules rouges, et pour des raisons encore peu claires, les individus AS résistent mieux au parasite que les AA. L'une des causes possibles de cette résistance est que chez un individu AS, le parasite favorise la cristallisation des hémoglobines S, conduisant à l'élimination du globule infecté. Quoique qu'il en soit, nous nous trouvons devant un cas de superdominance : l'hétérozygote AS, à la fois non anémique et résistant au paludisme, a une valeur sélective plus élevée que les homozygotes AA (non anémiques mais sensibles au paludisme) ou SS (insensibles au paludisme mais gravement anémiques). Cette situation, comme prévu, maintient l'allèle S en fréquence non négligeable, alors que celui-ci est pratiquement inexistant dans les zones non impaludées. Les valeurs sélectives estimées à partir des taux de survie des trois groupes d'individus ( $w_{AA} = 0,9$  ;  $w_{AS} = 1$  et  $w_{SS} = 0,2$ ) conduisent à une prédition théorique de 0,11 pour la fréquence d'équilibre du gène S. Cela est très proche de la valeur observée (0,09 en Afrique de l'Ouest).

La superdominance est donc un moyen efficace de stabiliser le polymorphisme. Cette propriété en fait un candidat de choix pour expliquer les forts taux de polymorphisme observés sur les protéines, dans les populations naturelles d'animaux et de plantes. Elle fut donc mise à profit par les sélectionnistes, dans le cadre de la controverse qui les opposait aux neutralistes (voir glossaire). La conception sélectionniste impliquait un polymorphisme stabilisé, contrairement au polymorphisme transitoire postulé par la théorie neutraliste. Malheureusement, le cas de l'anémie falciforme reste un exemple isolé. Il semble maintenant acquis que la superdominance n'est pas responsable du polymorphisme protéique

en général. Pour expliquer ce dernier, les sélectionnistes devront rechercher des formes de sélection plus subtiles.

## Des relations horizontales entre locus différents

**Les gènes s'associent en combinaisons dont la permanence dépend d'une liaison physique sur la molécule d'ADN.** En effet, dans le génome, chaque locus n'est pas représenté par un petit fragment d'ADN autonome : on trouve un nombre restreint de molécules d'ADN, très grandes, qui contiennent chacune beaucoup de locus différents. Au moment des divisions cellulaires, ces molécules se condensent et s'entourent d'un habillage protéique qui les rend visibles, au microscope, sous forme de corpuscules colorés : les *chromosomes*. Si toutes les divisions cellulaires étaient de type conforme, tous les gènes portés sur un jeu de chromosomes resteraient solidaires de génération en génération. Cependant, chez la plupart des organismes, il existe une étape de division cellulaire non conforme : la méiose (figure 2). Au cours de cette étape les *recombinations* créent de nouvelles associations de gènes.

L'efficacité des recombinations entre deux locus dépend de leur position moléculaire : elle est maximale quand les locus sont situés sur des chromosomes différents (recombinaison interchromosomique). Dans ce cas, toutes les combinaisons possibles sont produites avec des probabilités égales. Les associations entre gènes portés par des chromosomes différents sont donc éphémères, remises en cause à chaque méiose. La recombinaison intrachromosomique est, quant à elle, moins efficace. En effet, quand les locus sont sur des sites différents du même chromosome, ils sont physiquement liés. Pour les

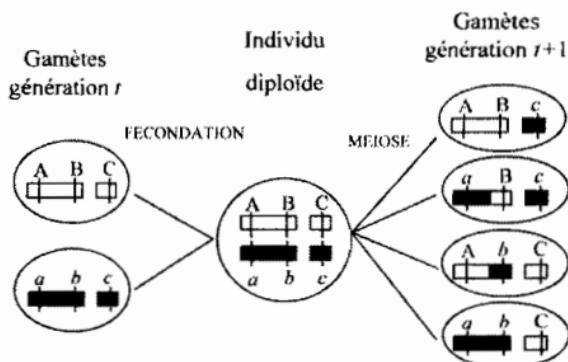


Figure 2 : Les brassages génétiques lors de la reproduction sexuée. Deux paires de chromosomes sont représentées par des rectangles de taille différente. La fécondation réunit chez un même individu diploïde des chromosomes issus de gamètes différents distingués, dans ce schéma, par leur couleur. Trois locus sont indiqués. Pour chaque locus, les gènes contenus dans les deux gamètes appartiennent à des allèles différents (A et  $a$  pour le premier locus, par exemple). La figure montre les produits (gamètes) d'une des méioses possibles chez cet individu. La *recombinaison génétique* désigne les événements amenant à l'association au sein d'un même gamète de gènes auparavant situés dans des gamètes différents. Elle a lieu à deux niveaux. La *recombinaison interchromosomique* se manifeste par la présence dans un même gamète d'un chromosome blanc et d'un chromosome noir (gamètes  $ABc$  et  $abC$ ). La *recombinaison intrachromosomique* produit des chromosomes associant des fragments blancs et noirs (gamètes  $AbC$  et  $aBc$ ).

dissocier, il faut couper deux chromosomes au même point, et les recoller en échangeant les morceaux (*crossing-over*). Les crossing-over sont assez rares. Pour deux locus du même chromosome, la méiose tend donc à reproduire préférentiellement les associations de gènes déjà présentes chez les parents. Les associations recombinées sont donc toujours minoritaires. Leur pourcentage dépend de

la distance qui sépare les deux locus le long de la molécule. Il est maximal (50 %) pour des sites très éloignés, et d'autant plus faible que les sites sont proches. Ainsi, deux gènes resteront associés d'autant plus longtemps qu'ils sont voisins, physiquement liés sur la molécule d'ADN présente dans le chromosome.

**La sélection sur un gène retentit sur les gènes voisins par un phénomène d'auto-stop génétique.** Une mutation favorable n'apparaît en général que chez un seul individu. L'allèle favorable est donc initialement associé avec tous les allèles portés par cet individu sur les autres locus. En l'absence de recombinaison, l'invasion de la population par l'allèle mutant entraîne avec elle tous les autres allèles de l'individu. Ces allèles sont pris en *auto-stop* par le mutant favorable et envahissent avec lui la population<sup>1</sup>. Cependant, le plus souvent, l'absence de recombinaison n'est vérifiée que dans le voisinage chromosomique immédiat du locus sélectionné. Quand on considère des locus de plus en plus éloignés, la probabilité de recombinaison augmente. Plus ils sont loin du locus sélectionné, plus les gènes auto-stoppeurs ont donc de chances d'être désolidarisés de la mutation favorable avant qu'elle n'ait fini d'envahir la population. Une fois cette invasion terminée, on attend donc des caractéristiques différentes d'une région du chromosome à l'autre. Loin du locus sélectionné, les gènes n'ont pas été affectés par la sélection et proviennent de l'ensemble de la population ancestrale. À l'inverse, à proximité de ce locus, on ne trouvera que les gènes initialement présents dans le premier chromosome mutant. La prédiction théorique est donc la suivante : la diversité génétique est très faible

---

1. J. Maynard-Smith et J. Haigh, « The Hitch-Hiking Effect of a Favorable Gene », *Genetical Research Cambridge*, 23, 1974, p. 23-35.

près du locus sélectionné et augmente progressivement quand on s'en éloigne jusqu'à retrouver une valeur « standard » pour la population considérée.

Cette situation a été observée chez le maïs (*Zea mays*), en comparant les formes cultivées et leur ancêtre sauvage, la téosinte. Lors de la domestication (il y a environ dix mille ans), l'homme a sélectionné chez la téosinte des variants morphologiques particuliers. Il a ainsi favorisé un allèle particulier du gène *tb1* (*Teosinte Branched 1*), qui produit une protéine réprimant la croissance des branches latérales au profit des branches porteuses de grains. La mutation touche la région régulatrice du gène, en amont de la séquence codante, augmentant son expression. Cette augmentation permet de passer d'une morphologie de type téosinte à une morphologie de type maïs. Le polymorphisme moléculaire de la région du gène *tb1* a été étudié chez le maïs et la téosinte<sup>1</sup>. On observe dans la région amont du gène une diversité génétique quasiment nulle chez le maïs, alors que cette diversité est élevée chez la téosinte. Lorsqu'on s'éloigne de cette région, le maïs retrouve des valeurs de diversité comparable à celles de la téosinte. Ces observations sont conformes aux prédictions d'un effet auto-stop dû à l'intense sélection exercée par l'homme sur la région amont.

**La sélection purifiante a également pour effet de diminuer la variabilité génétique quand la recombinaison est faible.** Nous avons déjà mentionné que les mutations les plus fréquentes sont délétères et rapidement éliminées par la sélection naturelle (sélection purifiante). De même qu'un allèle favorable prend ses voisins en

1. R. L. Wang *et al.*, « The Limits of Selection During Maize Domestication », *Nature*, 398, 1999, p. 236-239.

auto-stop lors de son invasion, les mutations délétères entraînent dans leur chute les petites régions de chromosomes qui ne sont pas assez éloignées de l'allèle délétère pour s'en dissocier par la recombinaison. La grande différence avec l'auto-stop est que cette élimination ne concerne qu'une petite partie des gènes de la population. En effet, les mutations délétères atteignent rarement des fréquences importantes dans la population. Au total, le mutant délétère ne fait que favoriser la disparition des quelques gènes qui ont le malheur de se trouver à proximité. Ainsi, il érode faiblement la variabilité génétique de la population dans cette région du chromosome. L'effet d'une seule mutation délétère pourrait donc être considéré comme négligeable. Cependant, celles-ci sont très fréquentes et touchent de nombreux locus répartis dans l'ensemble du génome. Il y a donc une « sélection de fond » (*background selection*), à l'œuvre en permanence pour éliminer ces mutants délétères<sup>1</sup>. Seules les portions de chromosome indemnes de mutations se maintiennent et participent à long terme à la constitution des générations futures. Cette sélection a donc un effet semblable à une réduction de l'effectif de la population. Comme nous l'avons vu au chapitre 3, l'effet d'une réduction d'effectif est d'augmenter la dérive (variations aléatoires), ce qui diminue la variabilité génétique populationnelle. Cet effet est plus prononcé quand le taux de recombinaison intrachromosomique est faible. Dans ce cas, chaque mutation délétère entraîne avec elle une plus grande portion de chromosome. Cette prédiction théorique a été remarquablement vérifiée chez la drosophile.

1. B. Charlesworth, M. T. Morgan et D. Charlesworth, « The Effect of Deleterious Mutations on Neutral Molecular Variation », *Genetics*, 134, 1993, p. 1289-1303.

Chez cette mouche, les taux de recombinaison intrachromosomique sont connus et varient d'une région à l'autre du génome. Des données moléculaires montrent sans ambiguïté que les régions les plus recombinantes sont aussi celles où la diversité génétique est la plus élevée<sup>1</sup>.

**Les locus sont également interdépendants en raison d'interactions fonctionnelles.** De fait, indépendamment de sa position sur le chromosome, l'effet physiologique d'un gène (et, *a fortiori*, son effet sur la valeur sélective de l'individu) peut dépendre de la présence ou de l'absence d'un autre gène. Ce type d'interaction est appelé *épistasie*. Les mutants auxotropes pour l'arginine chez le champignon *Neurospora*, déjà évoqués au chapitre 4, en fournissent un exemple simple. George Beadle et Edward Tatum ont obtenu plusieurs mutants différents, tous caractérisés par l'incapacité à synthétiser de l'arginine. Ces mutations étaient donc toutes de type perte de fonction. Toutefois, elles inactivaient chacune une enzyme différente de la chaîne de biosynthèse de l'arginine. Pour produire de l'arginine, il faut effectivement que *toutes* les enzymes de la chaîne soient fonctionnelles. En d'autres termes, l'effet d'une mutation sur l'enzyme 1 dépend de l'état de l'enzyme 2. Si l'enzyme 2 est fonctionnelle, l'inactivation de l'enzyme 1 est lourde de conséquences, puisqu'elle fait passer de l'état normal à l'état auxotrophe pour l'arginine. Si l'enzyme 2 est inactive, l'inactivation de l'enzyme 1 ne change rien.

L'étude des conséquences de telles relations épistatiques sur l'évolution des gènes est une tâche complexe,

---

1. D. J. Begun et C. F. Aquadro, « Levels of Naturally Occurring DNA Polymorphism Correlate with Recombination Rates in *Drosophila melanogaster* », *Nature*, 356, 1992, p. 519-520.

vu la diversité des situations envisageables. Nous envisagerons trois types d'épiastie.

Tout d'abord, dans les cas de *redondance fonctionnelle*, deux gènes situés à des locus différents peuvent se remplacer réciproquement pour effectuer une même fonction. Dans ce cas, on prédit qu'il y a un relâchement général de la pression de sélection purifiante sur un des deux gènes. En réalité, les mutations délétères sur ce gène sont peu contre-sélectionnées, vu qu'il existe toujours l'autre gène pour suppléer à la déficience qu'elles provoquent. Ce cas est notamment observé à la suite d'une duplication. Les deux locus sont au départ fonctionnellement identiques et substituables l'un à l'autre. Ensuite, le plus souvent, l'un des deux dégénère en pseudogène sous l'effet des mutations délétères (cas des globines, voir chapitre 2). Dans de rares cas, une mutation favorable intervient avant cette dégénérescence et donne à l'un des deux gènes une fonction nouvelle. Dès lors, la redondance fonctionnelle est perdue et on obtient deux gènes fonctionnels différents.

Dans les cas de *hiérarchie fonctionnelle*, la présence d'un gène (ou d'un groupe de gènes) est nécessaire pour que d'autres aient une utilité. De telles hiérarchies sont notamment observées au cours du développement embryonnaire. Les gènes exprimés très tôt dans le développement définissent un plan d'organisation général sur lequel d'autres gènes élaborent par la suite des structures de plus en plus complexes (voir chapitre 8). Dans un tel cas, la sélection purifiante est plus forte sur les gènes « maîtres » que sur les gènes subordonnés. En effet, une mutation délétère sur un gène « maître » a des conséquences en cascade sur le fonctionnement de tous les gènes qui lui sont subordonnés. L'évolution tend donc à être plus lente pour les gènes « maîtres » que pour les subordonnés (voir chapitre 8, gènes du développement).

Enfin, dans les cas de *coadaptation fonctionnelle*, deux gènes, situés à deux locus différents, ont un fonctionnement optimal lorsqu'ils sont ensemble, chacun étant « adapté » à l'autre. Par exemple, les communications entre cellules ou entre compartiments cellulaires font souvent appel à des interactions moléculaires entre protéines différentes. C'est le cas entre une hormone et son récepteur. La précision de ces interactions repose sur une complémentarité de forme des molécules, à la façon d'une clé et d'une serrure. Le résultat d'une telle interaction est de ralentir l'évolution de chacun des deux gènes, puisque la modification de l'un d'entre eux suffit à remettre en question leur complémentarité. Quand une mutation parvient néanmoins à se fixer sur l'un des deux gènes, on prédit qu'il existe une forte sélection en faveur de mutations *compensatrices* arrivant sur l'autre gène. La complémentarité peut ainsi être restaurée : les deux gènes évoluent de manière concertée.

La métaphore des paysages adaptatifs, imaginée par Sewall Wright<sup>1</sup> en 1932, montre que la prise en compte de l'épistasie conduit à une conception nouvelle de l'adaptation. Dans l'esprit de Darwin et de ses successeurs, notamment les premiers généticiens des populations comme R. A. Fisher, l'adaptation était conçue comme un phénomène essentiellement régulier. Une population devait s'améliorer continûment jusqu'à atteindre la composition génétique optimale pour exploiter son environnement. Cette conception peut être illustrée sur un graphique (figure 3) dans lequel l'axe vertical

1. S. Wright, « The Roles of Mutation, Inbreeding, Crossbreeding, and Selection in Evolution », *Proceedings of the VI International Congress of Genetics*, 1, 1932, p. 356-366. Réédition dans M. Ridley, *Evolution*, Oxford, Oxford University Press, 1997, p. 32-40.

représente la valeur sélective et l'axe horizontal la variation génétique. Bien sûr, il s'agit d'un raccourci graphique : la variation génétique a de nombreuses composantes, et ne peut être quantifiée par une seule variable. L'optimum génétique est représenté par un pic de valeur sélective. La sélection naturelle, qui retient à chaque génération les génotypes ayant la plus grande valeur sélective, permet à la population de gravir peu à peu les pentes de ce pic jusqu'à l'optimum.

Un autre pionnier de la génétique des populations, Sewall Wright, propose une image très différente. En généralisant le concept d'éistasie, il imagine plusieurs combinaisons de gènes coadaptés qui forment des optimums locaux (plusieurs « pics adaptatifs »). Entre ces pics existent des « vallées » correspondant à des combinaisons de gènes qui ne fonctionnent pas bien ensemble. Dans ce cadre, la sélection naturelle guide toujours la population vers les valeurs sélectives les plus élevées localement, ce qui peut la « piéger » dans un optimum local. En effet, une fois en haut du « pic », tout changement entraîne une diminution de valeur sélective. Pour passer d'un pic à l'autre, il faut un événement capable de faire varier brusquement les fréquences des différents allèles en dépit de la sélection naturelle. Sewall Wright suggère que des accidents démographiques peuvent jouer ce rôle (*shifting balance theory*). Suite à une réduction brutale d'effectif, les variations aléatoires (dérive génétique) domineraient temporairement les variations déterministes (sélection naturelle). Ainsi, la population se déplacerait dans le paysage adaptatif et prendrait un nouveau départ. Ernst Mayr<sup>1</sup>, en 1943, a exprimé une idée similaire sous le nom de « révolution génétique ».

1. E. Mayr, *Population, espèces et évolution*, Paris, Hermann, 1974 (1<sup>re</sup> éd. 1943).

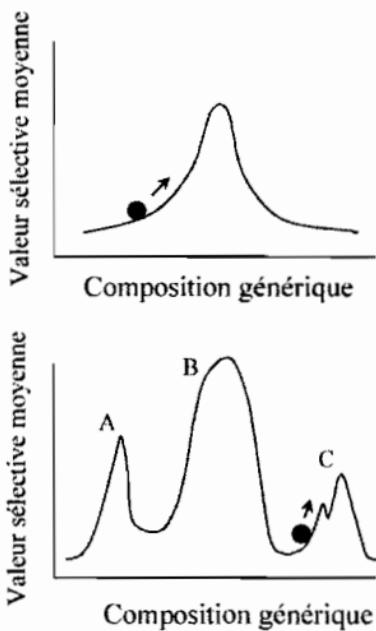


Figure 3 : Les paysages adaptatifs. La conception classique (haut) montre un seul maximum de valeur sélective ou pic adaptatif. La population (bille noire) progresse vers la valeur optimale sous l'effet de la sélection. Dans la conception de Wright (bas), il existe des maximums locaux (pics secondaires : A, C) en plus du maximum général (B). Dans ce paysage accidenté, la population gravite le pic le plus proche (C). Une fois au sommet de ce pic, la sélection naturelle tend à stopper son évolution. Elle peut donc maintenir la population loin de l'optimum général (B).

Cependant, des événements non démographiques sont également capables de stimuler la dynamique évolutive. Par exemple, un changement écologique important modifie le paysage adaptatif lui-même. Les combinaisons auparavant optimales ne se trouvent plus adaptées aux nouvelles conditions du milieu et l'évolution reprend son cours.

La conception de Wright présente une évolution irrégulièrue qui fait la part belle aux interactions entre gènes et aux événements aléatoires ou imprévisibles. Cette conception rencontre de plus en plus de succès, en liaison avec les progrès d'autres domaines. En effet, la complexité du développement embryonnaire est peu à peu décryptée, mettant au jour la richesse des interactions épistatiques entre gènes. Les paléontologues, quant à eux, établissent l'irrégularité des changements morphologiques observés dans les fossiles.

### Modèles multigènes : manifestations de la sélection naturelle sur des groupes de gènes corrélés

Nous avons montré jusqu'ici comment un gène pouvait influer sur l'évolution d'un autre. Ce modèle à deux partenaires n'est qu'un premier pas vers la prise en compte de la totalité des effets, liés au regroupement des gènes en individus. Mathématiquement, l'entreprise est démesurée. De même qu'il est vain d'essayer de décrire le comportement d'un gaz en considérant une par une chacune de ses molécules, il semble peu probable de pouvoir décrire l'évolution des génomes en détaillant toutes les interactions entre paires de gènes. Les modèles d'évolution doivent donc s'orienter vers une prise en compte plus générale de la complexité. Ce développement est encore à venir. En attendant, nous vous proposons deux exemples où la prise en compte du comportement conjoint de plusieurs gènes en corrélation permet de répondre à des questions évolutives.

**Pourquoi le chromosome Y est-il dégénéré ?** Chez beaucoup d'organismes, le déterminisme du sexe est

chromosomique. Chez l'homme, par exemple, le sexe mâle est déterminé par la présence d'un chromosome Y. Ce chromosome se distingue de son homologue (X) par plusieurs aspects : il est plus court, et n'échange d'ADN avec le chromosome X (recombinaison) que sur une petite région (*la région pseudo-autosomal*). Mais sa caractéristique la plus frappante est une dégénérescence fonctionnelle : en dehors de la région pseudo-autosomal, le chromosome Y possède très peu de gènes fonctionnels. Ces gènes codent pour des protéines impliquées dans l'orientation du développement embryonnaire vers la voie mâle. Le chromosome X est très différent, il contient sur toute sa longueur de nombreux gènes fonctionnels. Pourtant, de nombreuses données suggèrent que X et Y dérivent de chromosomes non sexuels, non dégénérés et au départ identiques.

Un modèle d'évolution (ici simplifié) permet d'expliquer la dégénérescence du chromosome Y. À l'origine, le sexe mâle est déterminé par un allèle dominant à un locus situé sur un chromosome normal. D'autres allèles apparaissent ensuite sur des locus voisins, coopérant avec le premier pour la fonction mâle. Ces gènes forment dès lors un groupe fonctionnel très interdépendant, puisque, en l'absence de l'un d'entre eux, la différenciation sexuelle est incomplète. En conséquence, il existe une forte sélection pour empêcher ces gènes de se séparer. Cette sélection favorise la suppression de la recombinaison chez les mâles dans cette région du chromosome. Ainsi, ces chromosomes « mâles » ne recombinent plus que sur une partie de leur longueur. À ce stade, deux populations de chromosomes se sont donc différenciées. Les chromosomes Y portent une région non recombinante, contenant les allèles « mâles ». Les chromosomes X ressemblent à des chromosomes normaux. Les

individus XX sont femelles et les XY sont mâles. Aucun individu YY ne peut apparaître puisque cela impliquerait que ses parents soient... deux mâles.

Imaginons maintenant un gène fonctionnel situé dans la région du chromosome Y qui ne recombine pas. Comme tous les gènes, il est soumis à des mutations délétères qui tendent à le rendre inactif. Ces mutations sont, comme on l'a vu, essentiellement récessives et ne s'expriment que chez un individu homozygote, porteur de deux mutations. Cependant, cette situation ne se produit pas. Les individus YY n'existant pas, la mutation ne pourrait être à l'état homozygote que chez des individus XY. Cela impliquerait la présence de la mutation sur un chromosome X. Or, en raison de l'absence de recombinations, un gène muté apparu sur un Y ne peut jamais passer sur un X. Dans un individu XY, le chromosome X fournit donc toujours une copie fonctionnelle. À l'inverse, une mutation apparue sur un chromosome X finit par apparaître à l'état homozygote dans un individu femelle (XX), et sera ainsi contre-sélectionnée.

Pour cette raison, la sélection purifiante peut agir sur le chromosome X, mais pas sur le chromosome Y. Cela explique que la région non recombinante du chromosome Y dégénère complètement, perdant peu à peu tous ses gènes fonctionnels sous l'effet de mutations délétères. Seuls les gènes déterminant la fonction mâle, donc dépourvus d'homologue sur le X, échappent à ce phénomène. La partie recombinante du chromosome se comporte, elle, comme un chromosome normal non dégénéré.

L'exemple du chromosome Y est intéressant à plusieurs titres. Il montre tout d'abord que la recombinaison, qui conditionne l'évolution conjointe de plusieurs gènes, peut elle-même évoluer, si cela procure un avantage sélectif à l'organisme. Il montre également que cette évolution,

dans un organisme diploïde, modifie fortement les pressions de sélection sur les gènes concernés. Dans le cas du chromosome Y, la suppression de la recombinaison entraîne un relâchement de la sélection purifiante, et, à long terme, la dégénérescence de l'essentiel du chromosome.

**Pourquoi les populations très différenciées ont-elles du mal à se mélanger ?** Notre deuxième exemple se place à l'échelle du génome entier, et fait appel à l'épiostasie entre de nombreux gènes. Dans une population qui reste isolée, la composition des génomes est relativement stable, car elle n'est pas modifiée à chaque génération par l'arrivée de nouveaux gènes provenant d'autres populations. Si l'isolement est assez long, la sélection aura le temps de privilégier les combinaisons de gènes qui fonctionnent bien entre eux (coadaptation fonctionnelle). Ces gènes augmenteront en fréquence de manière concertée et finiront par envahir la population. Pendant le même temps, une autre population isolée évoluera de la même façon, mais la combinaison de gènes coadaptés qui s'y fixera ne sera probablement pas la même. Nous pouvons noter par exemple ABCDEF un ensemble de six allèles (à six locus différents) coadaptés entre eux, fixés dans la première population, et *abcdef* les six allèles correspondants fixés dans la seconde. Si les deux populations viennent à se mélanger, des génomes recombinés (comportant par exemple les allèles *a*, B, *c*, D, E et *f*) finiront par apparaître, par le jeu des croisements entre individus. Or les allèles *a*, *c* et *f* fonctionnent bien avec *b*, *d* et *e* mais pas avec B, D et E : un génome recombiné aura donc une valeur sélective inférieure aux génomes « purs » trouvés dans les populations d'origine.

Ce phénomène est illustré par l'exemple des moules (*Mytilus sp.*). En croisant au laboratoire des moules pro-

venant de sites éloignés, l'un en Méditerranée (M) et l'autre sur la côte Atlantique française (A), Nicolas Bierne<sup>1</sup> et ses collaborateurs ont obtenu des hybrides de première génération puis, croisant ces hybrides entre eux, des descendants de deuxième génération. Cette procédure permet d'obtenir une grande diversité de génotypes recombinés. Au début de leur vie, les moules ne sont que de très petites larves nageuses (et ne sont pas encore devenus les coquillages qui nous sont familiers) mais grâce à l'usage de marqueurs génétiques moléculaires (cf. chapitre 3), il fut possible de déterminer pour chacune de ces petites larves l'origine de son ADN dans deux régions de chromosome (*mac-1* et *EF-bis*). Pour chaque région, une larve peut en effet avoir hérité de ses parents soit deux fragments de chromosome d'origine méditerranéenne (MM), soit deux fragments d'origine atlantique (AA), soit un de chaque (AM). On put ainsi estimer les taux de survie relatifs des différentes catégories de larves, qui sont exprimés ci-dessous.

	Origine de la 1 <sup>re</sup> région de chromosome ( <i>mac-1</i> )			
Origine de la 2 <sup>ème</sup> région de chromosome ( <i>EF-bis</i> )		AA	AM	MM
	MM	0,08	0,65	1,00
	AM	0,38	0,54	0,31
	AA	0,15	0,23	0,00

1. N. Bierne *et al.* « Fitness Landscapes Support the Dominance Theory of Postzygotic Isolation in the Mussels *Mytilus edulis* and *M. galloprovincialis* », *Proceedings of the Royal Society Ser. B*, 273, 2006, p. 1253-1260.

On observe que les génotypes de larves ayant les moins bonnes survies (en gras) sont ceux pour lesquels une des régions de chromosome ne comporte que des gènes d'origine A, et l'autre que des gènes d'origine M. En d'autres termes, les gènes A du premier chromosome ne s'entendent pas bien avec les gènes M du deuxième et inversement.

Sur le terrain, ce type de croisement peut avoir lieu lorsque deux populations d'une même espèce, séparées depuis longtemps, viennent à rentrer en contact (phénomène appelé *Hybridation*). On prédit alors qu'elles auront du mal à se mélanger, car les génomes recombinés, créés dans la zone de contact, seront défavorisés par la sélection naturelle. On aboutit à un équilibre entre la migration, qui tend à mélanger les génomes des deux populations, et la sélection qui tend à éliminer les produits de ce mélange. Dans cette situation, les génomes des deux populations ne se mélangent que sur une ligne de contact et restent « purs » de part et d'autre de cette zone, si la sélection est suffisamment forte. De nombreux exemples (dont celui des moules, déjà cité) soutiennent cette idée. Ainsi la zone de contact entre deux sous-espèces de souris domestiques en Europe constitue-t-elle une barrière que les gènes de l'une ou l'autre population ne peuvent franchir : de part et d'autre, les deux populations restent « pures » (exemple détaillé au chapitre 11).

## Résumé

La sélection n'agit pas sur des gènes isolés mais sur des ensembles de gènes plus ou moins solidaires.

Chez les diploïdes, les gènes d'un même locus sont temporairement associés par paires au sein des individus.

L'interaction entre les deux allèles associés peut moduler l'efficacité de la sélection naturelle sans en changer la direction : les allèles dominants sont ainsi plus exposés à la sélection que les récessifs. Elle peut aussi stabiliser la coexistence des deux allèles, si leur combinaison procure un avantage spécifique (superdominance).

Chez tous les organismes, les gènes situés sur des locus différents évoluent également de manière corrélée, pour deux raisons. Premièrement, ils peuvent être physiquement liés sur une même molécule d'ADN. Dans ce cas, la sélection sur un gène agira indirectement sur les gènes liés avec lui. Deuxièmement, deux gènes peuvent coopérer pour la réalisation de fonctions utiles à l'organisme. Dans ce cas, ils évolueront de manière concertée, chacun s'adaptant à l'autre.

La prise en compte des interactions entre gènes est complexe, mais elle permet de comprendre des phénomènes d'évolution à des niveaux d'intégration élevés, comme le niveau du chromosome (dégénérescence du chromosome Y) ou celui du génome entier (maladaptation des hybrides entre populations différenciées).



## LE SEXE, STRATÉGIE DE GESTION DE LA VARIABILITÉ GÉNÉTIQUE

Comme nous l'avons vu au chapitre 1, les êtres vivants appartiennent à un réseau généalogique commun. La *reproduction* est l'ensemble des mécanismes à l'origine de cette généalogie, par lesquels un individu donne naissance à des individus qui lui ressemblent, créant ainsi une lignée de descendance. Les organismes véhiculent des gènes (voir chapitre 5). Ces derniers se perpétuent donc à travers le réseau généalogique et sont soumis à un tri par sélection naturelle et dérive génétique. Le fait que plusieurs gènes soient associés dans un seul individu modifie les modalités du tri (voir chapitre 6). Mais la façon dont ces gènes s'associent dans les individus résulte elle-même des modalités de la reproduction des individus. Quelles seront alors les conséquences d'un changement de ces modalités, et donc de la structure du réseau généalogique sur l'évolution des organismes et des gènes ?

Deux éléments déterminent la structure du réseau généalogique : (i) La reproduction est-elle sexuée ou non ? Autrement dit, les descendants héritent-ils d'un mélange de gènes issus de deux gamètes différents ou sont-ils des copies génétiques d'un parent unique ? (ii) Quand la reproduction est sexuée, impliquant deux

gamètes, de quel(s) individu(s) viennent ces gamètes et comment leur rencontre a-t-elle lieu ? Les partenaires sexuels sont-ils pris au hasard dans la population, sont-ils apparentés entre eux, voire sont-ils un seul et même individu ?

Ces éléments déterminent les différences génétiques entre les individus au sein du réseau généalogique et dans ce chapitre, nous les aborderons successivement. Le premier est celui du sexe. Nous avons tendance à associer systématiquement la reproduction avec le sexe, mais nous verrons qu'il s'agit d'une intuition anthropomorphique. Quand et par qui le sexe est-il réellement pratiqué dans la nature ? Quels sont ses enjeux évolutifs ? Le deuxième élément est le choix du partenaire sexuel : peut-on pratiquer le sexe avec soi-même (ou avec des individus apparentés) et quelles en sont les conséquences ? En d'autres termes que faire du sexe une fois qu'il a été acquis au cours de l'évolution ?

### Qui fait du sexe et quelles en sont les conséquences génétiques ?

**Les bactéries pratiquent majoritairement, mais pas totalement, une reproduction conforme de type asexué.** Elles se reproduisent par simple division cellulaire. Cette division intervient après réPLICATION de l'ADN, ce qui assure la transmission intégrale de l'ADN maternel à chacune des cellules filles. Seules les mutations peuvent introduire des différences entre cellules filles (voir chapitre 2). Les bactéries ont donc une reproduction qualifiée de *conforme ou clonale*.

La reproduction strictement conforme entraîne que les gènes évoluent en blocs indissociables (voir chapitre 6), car elle exclut la réassocation dans une même bactéries de matériel génétique provenant de deux bactéries différentes. Cependant, en 1946, Joshua Lederberg et Edward Tatum<sup>1</sup> ont réfuté cette hypothèse. Ils disposaient de deux souches de bactéries *Escherichia coli* : la souche 1 était capable de synthétiser la thréonine, la leucine, la thiamine, mais incapable de synthétiser la méthionine et la biotine ; la souche 2 présentait les caractéristiques inverses. Aucune de ces deux souches ne se développait sur milieu minimum, chacune ne poussant que si on lui fournissait les composés qu'elle n'arrivait pas à synthétiser. En revanche, des bactéries capables de pousser sur milieu minimum apparurent lorsque les deux souches furent mises en contact : ces bactéries combinaient les gènes de synthèse de thréonine, leucine et thiamine, provenant de la souche 1 et ceux de méthionine et biotine, provenant de la souche 2. Cela montre que la recombinaison génétique existe chez les bactéries, et peut engendrer des combinaisons favorables de gènes. En effet, la survie sur milieu minimum eût été impossible si la reproduction avait été purement clonale, les deux clones initiaux étant incapables de pousser sur ce milieu.

Les bactéries ne pratiquent pas la méiose mais disposent d'autres mécanismes d'échanges génétiques. Deux bactéries peuvent en effet entrer en *conjugaison*. À cette occasion, elles s'accroissent et échangent des fragments d'ADN. L'ADN échangé peut s'intégrer par une recombinaison dans celui de la bactérie receveuse. La conjugaison permet donc de briser des combinaisons génétiques

1. J. Lederberg et E. L. Tatum, « Gene Recombination in *Escherichia coli* », *Nature*, 158, 1946, p. 558.

et d'en créer de nouvelles. Ce mode d'échanges génétiques n'est pas une reproduction car le nombre de cellules n'augmente pas. Cependant, les deux individus voient leur information génétique modifiée et donc, leurs descendants, produits ultérieurement par division conforme, posséderont des combinaisons de gènes différentes de celles de leurs parents. L'association conjugaison/division cellulaire est donc une forme rudimentaire de reproduction sexuée. La reproduction, chez les bactéries, peut être considérée comme facultativement sexuée.

**Chez les eucaryotes le sexe est caractérisé par deux étapes, la méiose et la fécondation,** qui alternent au cours d'un cycle de vie plus ou moins complexe (voir chapitre 5). La reproduction sexuée engendre ainsi une diversité considérable de combinaisons génétiques. Prenons le cas de l'espèce humaine. Chez cette espèce, la reproduction sexuée est strictement bi-parentale : deux individus doivent s'unir pour donner naissance à de nouveaux individus qui héritent, lors de la fécondation, de la moitié du stock chromosomique de chaque parent. Un premier niveau de diversité génétique est lié au grand nombre de couples possibles au sein de la population. En outre, la formation des gamètes par méiose engendre elle-même une diversité considérable entre les descendants possibles d'un même couple. Le gamète hérite d'un chromosome pris au hasard dans chacune des paires de chromosomes du parent (recombinaison interchromosomique, voir chapitre 6). Pour  $n$  paires de chromosomes (chez l'homme,  $n = 23$ ), il y a  $2^n$  jeux de chromosomes possibles ( $2^{23}$  représentent environ 8 millions). Les deux gamètes étant soumis au même tirage, les descendants de chaque couple peuvent hériter de  $2^n \times 2^n$  combinaisons chromosomiques différentes (soit  $7.10^{13}$  pour un couple humain). Enfin, la recombinaison intrachromosomique

augmente encore cette diversité, chaque chromosome pouvant échanger des fragments avec son homologue selon une infinité de configurations, avant d'être transmis au gamète. Ce mode de reproduction conduit donc à une énorme variabilité des combinaisons de gènes individuelles.

Chez certains eucaryotes, la reproduction sexuée est associée à une différenciation des gamètes. Cette différenciation est manifeste chez l'homme ainsi que chez un très grand nombre d'animaux : il existe deux catégories de gamètes différents (spermatozoïdes et ovules) et une fécondation implique toujours un gamète de chaque catégorie. Ce phénomène, dit d'*anisogamie*, reflète une asymétrie d'investissement dans les gamètes : l'ovule contient des réserves pour le développement futur de l'embryon, et représente un investissement important de l'individu qui le produit. Le spermatozoïde est petit et n'apporte que ses gènes : ses caractéristiques ne reflètent pas une mise en réserve mais plutôt une optimisation de l'accès aux ovules (mobilité). Chez l'homme, certains individus (mâles) sont spécialisés dans la production de spermatozoïdes et d'autres (femelles) dans la production d'ovules. Chez beaucoup d'animaux et de plantes, les deux types de gamètes sont néanmoins produits par un même individu (*hermaphrodite*), qui peut donc avoir la capacité de s'autoféconder. L'anisogamie n'accompagne pas nécessairement le sexe chez les eucaryotes : chez de nombreux unicellulaires, la fécondation a lieu entre deux cellules semblables (*isogamie*).

La plupart des eucaryotes, à l'instar des bactéries, font du sexe facultatif. Dans le cycle de vie de l'homme, et de beaucoup d'animaux et de plantes, la reproduction ne se fait que par voie sexuée. Cependant, la majeure partie de la diversité évolutive des eucaryotes ne se trouve

pas chez les animaux et les plantes, mais chez les organismes unicellulaires ou protistes. Dans les cycles de vie des protistes, les phases de reproduction ne sont pas toutes associées au sexe. Comme les bactéries, les protistes se reproduisent généralement par division cellulaire conforme. Ils sont pourtant capables d'échanges génétiques, lors d'épisodes sporadiques de reproduction sexuée. La reproduction des eucaryotes unicellulaires est donc majoritairement asexuée, et facultativement sexuée.

La reproduction asexuée ne se limite pas aux unicellulaires. En effet, chez certains pluricellulaires, ce n'est pas une cellule unique mais un fragment pluricellulaire de l'individu qui donne naissance à un descendant. De nombreuses plantes sont ainsi capables de *multiplication végétative* (comme les fraisiers, voir chapitre 5). Ce type de reproduction asexuée est également fréquent chez plusieurs grands groupes d'animaux, notamment chez les coraux, les éponges, ou les ascidies. Chez les végétaux comme chez les animaux, la reproduction asexuée est associée à une reproduction sexuée plus ou moins régulière.

Dans la plupart des lignées d'eucaryotes pluricellulaires, il existe des espèces qui ont un mode de reproduction majoritairement clonal dérivé de la reproduction sexuée : la *parthénogenèse*. Chez ces espèces, les femelles produisent des ovules qui se développent sans intégrer de matériel génétique mâle. Ces organismes présentent des structures et des mécanismes (glandes sexuelles, divisions méiotiques modifiées) qui témoignent de l'existence passée d'une reproduction sexuée normale. La parthénogenèse est donc une modification de la reproduction sexuée dont le résultat est de supprimer ou d'atténuer les recombinaisons génétiques traditionnellement liées au sexe. Elle existe chez des

organismes divers selon des mécanismes tous plus originaux les uns que les autres<sup>1</sup>.

L'escargot aquatique *Melanoides tuberculata* illustre le cas le plus simple. Chez cet escargot, l'ovule est produit non par une méiose mais par une division cellulaire conforme (mitose). Il dispose donc du jeu complet de chromosomes maternels, et se développe en un nouvel individu génétiquement identique à sa mère. On parle alors de *parthénogenèse apomictique* ou *apomixie*. La conséquence génétique de ce mode de reproduction, qui existe chez des organismes très variés, est la suppression du réassortiment des gènes par la méiose.

De nombreux organismes combinent au cours de leur cycle de vie plusieurs modalités de reproduction. Ainsi, pendant la belle saison, les femelles de certains pucerons produisent d'autres femelles par parthénogenèse. À l'approche de l'automne, la dernière parthénogenèse de la saison donne naissance à des mâles et à des femelles qui se reproduisent par reproduction sexuée biparentale. Les femelles fécondées pondent alors des œufs, formes de résistance qui se maintiennent tout l'hiver. Ces œufs éclosent au printemps et produisent des femelles à reproduction parthénogénétique.

Le poisson *Poeciliopsis* a, quant à lui, un mode de reproduction parthénogénétique particulièrement original<sup>2</sup>. En effet, les femelles produisent des œufs diploïdes qui nécessitent d'être activés par un spermatozoïde d'une

1. G. Bell, *The Masterpiece of Nature : the Evolution and Genetics of Sexuality*, Berkeley, University of California Press, 1982. O. Judson, *Manuel universel d'éducation sexuelle. À l'usage de toutes les espèces, selon Madame le Dr Tatiana*, Paris, Coll. Science ouverte, Seuil, 2004.

2. R. C. Vrijenhoek, « Unisexual Fish : Model Systems for Studying Ecology and Evolution », *Annual Review in Ecology and Systematics*, 25, 1994, p. 71-96.

autre espèce pour se développer. Le spermatozoïde pénètre dans l'ovule comme si ces deux gamètes étaient de la même espèce. Le développement de l'œuf se déclenche et un nouvel individu femelle, triploïde, se construit. Cependant, dans la lignée germinale, le matériel génétique apporté par le spermatozoïde est rejeté : les ovules produits ne contiennent donc que le matériel génétique maternel. Cette parthénogénèse originale est appelée *gynogenèse*.

**Origine et histoire évolutive de la reproduction.** Comme nous venons de le voir, la reproduction clonale existe chez une grande variété d'organismes. À part les mammifères et les oiseaux, tous les grands groupes vivants connaissent ce mode de reproduction. Les organismes unicellulaires, représentant les lignées eucaryotes les plus anciennes, pratiquent majoritairement la reproduction asexuée, et sporadiquement la reproduction sexuée. Méiose et fécondation ont donc été acquises très tôt au cours de l'évolution des eucaryotes, mais elles n'étaient initialement pas obligatoires pour la reproduction. L'anisogamie est ensuite apparue dans plusieurs lignées d'eucaryotes. La reproduction sexuée obligatoire est également apparue plusieurs fois indépendamment au sein de lignées pluricellulaires relativement récentes : animaux, algues rouges et végétaux vasculaires<sup>1</sup>. Dans ces lignées, certaines espèces ont toutefois détourné les mécanismes de la reproduction sexuée pour adopter une reproduction parthénogénétique, qui annule ou affaiblit fortement les conséquences génétiques du sexe. Cette situation semble paradoxale : pourquoi acquérir une

1. J. Dacks et A. J. Roger : « The First Sexual Lineage and the Relevance of Facultative Sex », *Journal of Molecular Evolution*, 48, 1999, p. 779-783.

reproduction sexuée obligatoire et en annuler par la suite les effets ? Pour aborder cette question, il faut comparer les performances de ces modes de reproduction vis-à-vis de la sélection naturelle.

### Le sexe apporte-t-il un avantage sélectif ?

**Pourquoi le sexe ?** Cette question taraude les biologistes depuis la naissance de la théorie de l'évolution. Les mécanismes impliqués dans la reproduction sexuée, depuis les processus cellulaires et moléculaires (méiose, recombinaison) jusqu'aux processus comportementaux et psychologiques de la séduction, sont tellement complexes et nombreux que le sexe ne peut pas être fortuit. Il faut des raisons impérieuses pour justifier l'apparition et le maintien de ce que Graham Bell appelle « chef-d'œuvre de la nature ». Dans nul autre domaine de la théorie de l'évolution on n'a échafaudé plus de théories, et il y a lieu de croire que le débat sur la valeur sélective du sexe réserve encore bien des rebondissements. Nous présenterons ici ses arguments principaux, assortis d'hypothèses qui, pour être provisoires, sont cependant souvent citées.

**Une vision démographique de la reproduction : le coût des mâles chez les espèces anisogames.** La sélection naturelle favorise les gènes qui se transmettent le plus efficacement. De ce point de vue, les reproductions, sexuée et clonale, diffèrent. Imaginons deux femelles, l'une à reproduction sexuée et l'autre à reproduction clonale, ayant une fécondité égale (par exemple, deux descendants par femelle). La femelle clonale laisse à chaque génération deux descendants femelles. La femelle sexuée produit également deux descendants, mais, en supposant

un sex-ratio équilibré, seul un des deux en moyenne sera femelle. À la génération suivante, les deux femelles cloniales donnent chacune deux descendants, soit en tout quatre individus. En revanche, l'unique femelle sexuée donne à nouveau deux descendants dont, en moyenne, un seul est femelle. Clairement, la démographie de la lignée clonale est beaucoup plus avantageuse que celle de la lignée sexuée. Cela est dû au fait que les mâles, par l'intermédiaire du spermatozoïde, donnent la moitié de leurs gènes à chaque descendant mais ne lui donnent pas de ressources. Le nombre de descendants produits par une femelle est donc déterminé uniquement par l'investissement énergétique de cette dernière et n'est pas plus grand chez la femelle sexuée que chez la femelle clonale. Ce phénomène est donc lié à l'anisogamie (asymétrie d'investissement entre les gamètes mâle et femelle) et on l'appelle couramment le « coût des mâles ». Une vision complémentaire de ce coût part d'un point de vue génétique. La femelle sexuée ne transmet que la moitié de ses gènes dans chaque œuf, laissant la place à autant de gènes en provenance du mâle mais supportant tout l'investissement énergétique. La femelle clonale place, elle, deux fois plus de gènes dans chaque œuf, pour le même prix.

Dans une espèce isogame, le coût des mâles n'existe pas : il n'a donc pas pu s'opposer à l'apparition initiale du sexe, puisque l'anisogamie n'a pu apparaître qu'après le sexe lui-même (avant cela elle n'est pas définie). *A contrario*, la majeure partie des espèces eucaryotes sexuées sont actuellement anisogames : s'il apparaît, dans une de ces espèces, une mutation conférant une reproduction clonale (par exemple, une forme de parthénogenèse), la sélection naturelle devrait lui permettre d'envahir la population. Si l'on en reste là, la reproduction parthénogénétique devrait être le mode de reproduction le plus

courant. Le sexe a-t-il des avantages génétiques pour compenser ce handicap ?

Le sexe permet-il de nettoyer les génomes de leurs allèles délétères ? Quand une mutation délétère apparaît, l'organisme à reproduction clonale a peu de chances de s'en débarrasser. Le gène fonctionnel ne peut être restauré que par une mutation « *réverse* », qui a une probabilité très faible de survenir. Bien sûr, au niveau de la population, la sélection agit en permanence pour favoriser les génomes dépourvus de mutations délétères, mais, si la population est petite, ces génomes optimaux peuvent très bien disparaître en raison de fluctuations aléatoires (dérive génétique). Il n'est alors pas possible de revenir en arrière, et tous les individus de la population se retrouvent avec au moins un gène délétère. Ce phénomène, propre aux populations à reproduction clonale, est appelé « *cliquet de Muller* », par référence aux dispositifs de cliquets qui empêchent le mouvement vers l'arrière d'une pièce mécanique, et à l'auteur qui décrivit le premier ce phénomène. Le sexe apparaît dans ce cadre comme un moyen économique de débarrasser le génome des gènes délétères : en recombinant deux génomes présentant chacun une mutation, mais sur des locus différents, il permet en effet de recréer un génome exempt de mutations. Ce modèle simple a néanmoins été remis en question : des simulations ont montré que lorsque les effectifs sont suffisamment faibles pour qu'une population asexuée accumule réellement des mutations délétères, la dérive est si forte que même dans une population sexuée beaucoup d'entre elles se fixent. Une fois une telle fixation arrivée, aucun chromosome de la population ne possède l'allèle sain correspondant : la recombinaison ou le sexe ne peuvent donc plus régénérer de génomes sans mutations.

Une autre objection peut être apportée : le bénéfice du sexe dans la lutte contre l'accumulation de mutations délétères ne s'exprime qu'à long terme. En effet, la plupart des mutations délétères sont récessives. Imaginons une lignée parthénogénétique dérivée d'une espèce sexuée : les mutations délétères s'y accumulent, certes, mais elles n'auront un effet que lorsque les deux gènes homologues auront été touchés, ce qui peut prendre du temps ; d'ici là, la lignée aurait eu le temps d'envahir l'espèce vu son avantage démographique. Cet effet est amplifié par la polyploidie car chaque gène est alors présent en plus de deux exemplaires. Cela retarde d'autant plus l'effet des mutations délétères. Il se trouve justement que beaucoup d'organismes parthénogénétiques sont polyploïdes.

**Au niveau de la population, la reproduction sexuée permet de combiner des mutations favorables et accélère les « progrès » génétiques.** Imaginons trois mutations favorables A, B, C portant sur des locus différents. Dans une population où ces trois gènes sont absents au départ, chacun d'eux peut apparaître indépendamment des deux autres par mutation. Dans une population clonale (figure 1, bas), les trois gènes ne peuvent se fixer que si une même lignée est touchée successivement par trois événements de mutation. Si par malchance, la mutation B touche des individus non porteurs de la mutation A, A et B se trouvent sur des clones différents qui entrent en compétition. Au lieu de se combiner, les mutations s'élimineront donc l'une l'autre. Ce désavantage de la clonalité est encore plus fort si les gènes A, B et C ne sont favorisés que lorsqu'ils se trouvent ensemble dans un même individu (épistasie coopérative). Le temps nécessaire à combiner les trois mutations sera alors très long, car la fixation d'un gène, en l'absence des deux

autres, n'est pas stimulée par la sélection et repose sur la dérive génétique.

En revanche, dans une population sexuée, les trois gènes mutants, portés initialement par des individus différents, pourront par le jeu des croisements être rapidement combinés dans un même individu, lui conférant l'avantage sélectif inhérent à l'association ABC. Les bactéries de Lederberg et Tatum (voir p. 169) fournissent un exemple d'une combinaison de mutations favorables produite par reproduction « sexuée ». Le sexe permet donc une incorporation rapide des nouvelles mutations, et l'obtention des combinaisons de gènes les plus favorables. Il accélère ainsi la vitesse d'évolution d'une espèce qui le pratique. Cette accélération est particulièrement appréciable lorsque l'environnement change rapidement, mettant l'espèce au défi de trouver sans cesse des combinaisons de gènes adaptées.

**Au niveau de l'individu, le sexe, au contraire de la reproduction clonale, tend néanmoins à détruire les combinaisons avantageuses acquises.** En effet, lors de la reproduction sexuée, les combinaisons de gènes sont défaites et refaites à chaque génération. Un individu ayant la bonne combinaison ABC a donc tout intérêt à se reproduire clinalement : il serait sûr de produire des descendants à forte valeur sélective. Le même raisonnement peut être tenu du point de vue des gènes. Imaginons un gène S qui confère la reproduction sexuée ou augmente la recombinaison. Grâce à cette propriété, il favorisera l'apparition d'individus performants de type SABC. Associé au génotype favorable ABC, S devrait tendre à envahir la population. Toutefois, la même propriété entraîne la dislocation de la combinaison SABC, par recombinaison avec des individus de génotype différent. Le gène S ne profitera donc pas longtemps de son

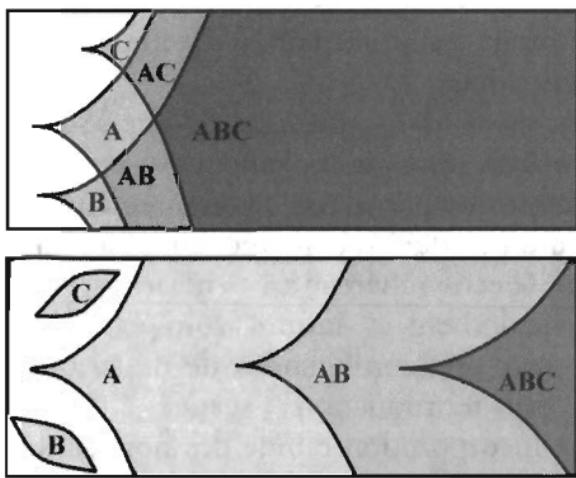


Figure 1 : D'après Hermann J. Muller<sup>1</sup>. L'évolution d'une population sexuée et celle d'une population clonale au cours du temps sont représentées respectivement dans les schémas du haut et du bas. Le temps est décrit par l'axe horizontal alors que l'axe vertical représente la fréquence de différents génotypes dans les populations. Au début, la population est homogène (blanc). Ensuite, des gènes mutants (A, B, C) apparaissent. Dans la population sexuée (haut), par le jeu des croisements, les trois gènes mutants sont rapidement associés dans un même individu ce qui permet à ces trois gènes de se fixer rapidement dans la population. Dans la population clonale (bas), il faudra que les mutations apparaissent successivement dans la même lignée clonale pour que la combinaison puisse envahir la population.

avantage. Au contraire, un gène de clonalité parvenant à s'associer avec la combinaison ABC envahira rapidement la population.

**Le sexe est donc intéressant à long terme pour la population ou l'espèce, mais coûteux à court terme**

1. H. J. Muller : « Some Genetic Aspect of Sex », *American Naturalist* 66, 1932, p. 118-138.

**pour les individus.** Grâce à la diversité génétique extraordinaire produite par le sexe, les espèces sexuées ont un potentiel évolutif supérieur. Ce potentiel se révèle particulièrement intéressant si l'espèce ou la population est mise au défi de s'adapter à des changements de l'environnement. Cependant, cette diversité n'est, quasiment jamais, un avantage sélectif en soi pour l'individu qui pratique le sexe. Les seuls cas où un individu peut avoir intérêt à produire une descendance génétiquement diversifiée et différente de lui-même se produisent dans des situations extrêmes : par exemple, quand l'individu est manifestement mal adapté à son milieu, où quand ce milieu change de manière si rapide que les descendants connaîtront des conditions différentes de celles des parents. À court terme, l'avantage démographique de la reproduction clonale devrait donc l'emporter : les modèles théoriques prédisent qu'en dehors des situations extrêmes, un organisme strictement clonal a toujours un avantage sélectif par rapport aux individus pratiquant le sexe obligatoire.

**Le sexe facultatif est un compromis possible.** La pratique occasionnelle de la reproduction sexuée dans une espèce globalement clonale semble souvent la solution optimale<sup>1</sup>. La reproduction clonale majoritaire évite le coût démographique du sexe ; la reproduction sexuée sporadique permet aux individus de diversifier leur descendance, favorisant son adaptation à un changement des conditions environnementales. De fait, les épisodes de reproduction sexuée facultative sont souvent liés à un stress environnemental (par exemple, diminution des

---

1. L. D. Hurst et J. R. Peck, « Recent Advances in Understanding of the Evolution and Maintenance of Sex », *Trends in Ecology and Evolution*, 11, 1996, p. 46-52.

ressources alimentaires ou augmentation de la densité, chez les unicellulaires). Il semble donc relativement aisé de comprendre le maintien d'une reproduction sexuée facultative chez les unicellulaires, mais pas celui d'une reproduction sexuée obligatoire chez beaucoup d'organismes pluricellulaires.

**Pourquoi le sexe obligatoire se maintient-t-il au cours de l'évolution dans les groupes qui l'ont acquis ?** Dans ces groupes, la parthénogénèse et le sexe facultatif n'existent que sporadiquement, chez quelques espèces. Paradoxalement, la parthénogénèse apparaît toujours comme un trait récemment fixé dans une lignée particulière, ce qui suggère que les lignées parthénogénétiques tendent à disparaître plus vite que les autres ou à revenir rapidement à l'état ancestral sexué obligatoire. Il existe néanmoins des cas de clonalité ancienne : chez les rotifères bdelloïdes, de très petits invertébrés aquatiques, les données moléculaires suggèrent que la parthénogénèse apomictique a persisté plusieurs millions d'années<sup>1</sup>. Ce cas reste néanmoins exceptionnel.

Qu'est-ce qui empêche les espèces de devenir entièrement parthénogénétiques, étant donné l'avantage à court terme de ce type de reproduction ? Les réponses à cette question semblent décourager toute tentative de généralisation, car dans chaque espèce, le sexe semble maintenu en raison d'une contrainte particulière. Ainsi, les œufs parthénogénétiques des pucerons ne sont pas résistants au froid, au contraire des œufs résistants issus de reproduction sexuée. Un mutant chez lequel la phase sexuée du cycle de vie serait supprimée ne pourrait donc pas

1. D. Mark Welch et M. Meselson, « Evidence for the Evolution of Bdelloid Rotifers without Sexual Reproduction or Genetic Exchange », *Science*, 288, 2000, p. 1211-1215.

survivre à l'hiver (N.B. : dans les régions à hiver doux, on trouve effectivement des pucerons à cycle de vie entièrement parthénogénétique). De telles explications particulières à une espèce, d'apparence anecdotique, peuvent être trouvées pour la plupart des espèces qui maintiennent de la sexualité alors, qu'elles sont capables de reproduction clonale. Il en va de même pour les espèces à reproduction sexuée obligatoire.

**Avantages à court terme et avantages à long terme : la sélection d'espèces.** La sélection naturelle au niveau de l'individu est défavorable au sexe. En effet, pour un individu, le désavantage démographique du sexe est immédiat, mais l'avantage génétique ne s'exprimera qu'au bout de nombreuses générations : pour être plus précis, à la génération future de descendants qui subira un changement de l'environnement. Or, la sélection naturelle ne préserve pas les individus en fonction des performances de leur descendance future.

On peut alors imaginer qu'une sélection à un niveau supérieur, celui de la population ou de l'espèce, serait responsable du maintien d'une reproduction sexuée obligatoire<sup>1</sup>. Imaginons un ensemble d'espèces pratiquant, le sexe facultatif (SF) (figure 2). Dans l'une d'elles, un mutant pratiquant le sexe obligatoire (SO) peut apparaître. Si l'espèce a un effectif faible, ce mutant a une probabilité non nulle de se fixer par dérive génétique : l'espèce entière devient alors SO. Si l'environnement restait stable, une espèce ne pourrait rester éternellement SO : tôt ou tard elle serait envahie par des mutants SF, qui ont une démographie plus efficace que les SO. Imaginons maintenant que les conditions environnementales

1. Nunney, « The Maintenance of Sex by Group Selection », *Evolution*, 43, 1989, p. 245-255.

changent brutalement. Les espèces SO auront de bonnes chances de survivre à ce changement, car la variété des combinaisons génétiques produites leur permet de trouver rapidement une combinaison adaptée aux nouvelles conditions. Au contraire, une espèce SF majoritairement clonale aura de grandes chances de s'éteindre lors de ces épisodes de changement (figure 2). Cette survie différentielle sélectionne les espèces SO au détriment des SF. Ce scénario est soutenu par la relative jeunesse de la plupart des lignées parthénogénétiques chez les animaux actuels, qui suggère qu'elles ont un taux d'extinction élevé et une faible durée de vie dans l'évolution. Par ailleurs, le scénario justifie également l'apparition dans chaque espèce de mécanismes particuliers qui empêchent à court terme le retour à une reproduction clonale (comme les œufs d'hiver des pucerons). Une espèce sexuée obligatoire privée de tels mécanismes est en effet défavorisée, car elle court le risque d'être envahie par un mutant parthénogénétique, et de s'éteindre au prochain changement d'environnement.

Ce scénario pose néanmoins deux problèmes. Premièrement, la sélection d'espèce n'explique pas l'apparition de la première espèce SO : en effet, à l'intérieur d'une espèce SF, les individus SO ne sont pas favorisés, voire sont défavorisés si l'espèce est anisogame. Nous avons donc dû faire intervenir au départ un événement hypothétique de dérive génétique. Dans la suite du scénario, ce problème disparaît. En effet, le caractère SO est transmis par héritage phylogénétique : une espèce est SO parce que ses ancêtres étaient SO. Le deuxième problème posé par le scénario est la nature des changements environnementaux récurrents qui maintiennent l'avantage des espèces SO. De ce point de vue, l'hypothèse la plus

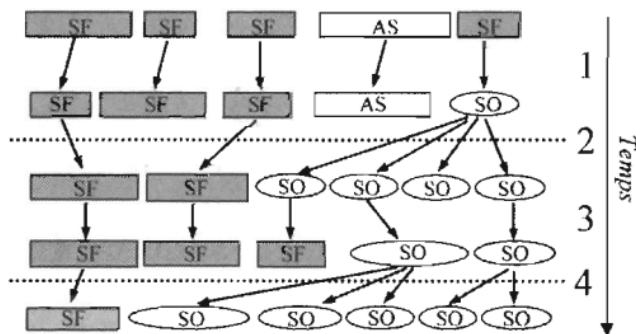


Figure 2 : Scénario idéalisé expliquant comment les potentialités évolutives du sexe ont pu permettre la mise en place de la reproduction sexuée obligatoire dans certaines lignées. Différentes espèces sont représentées par des rectangles blancs quand elles sont totalement asexuées (AS), des rectangles gris quand elles pratiquent le sexe facultatif (SF) et par des ellipses blanches quand elles pratiquent le sexe obligatoire (SO). Au début aucune espèce n'est de type SO. Le sexe obligatoire apparaît dans une espèce à faible effectif, lors de la phase 1, et s'y fixe fortuitement par dérive génétique. Au cours de la phase 2 a lieu un changement important de l'environnement (trait pointillé) : ce changement entraîne l'extinction de la plupart des espèces AS et SF, incapables de trouver rapidement des combinaisons génétiques adaptées. L'espèce SO survit et fonde de nouvelles espèces filles dans les emplacements laissés vacants par les extinctions. Au cours de la phase 3, l'environnement est redevenu stable et les espèces SF survivantes prospèrent. Certaines populations SO deviennent SF, envahies par des mutants SF, démographiquement plus efficaces que les individus SO. En phase 4, un nouveau changement de l'environnement fait encore augmenter la proportion des espèces SO. Au long terme, si les changements environnementaux sont fréquents, toutes les espèces pratiqueront le sexe obligatoire.

citée est l'hypothèse parasitaire. En effet, une composante importante de l'environnement d'une espèce est le cortège d'espèces parasites qui l'exploitent : virus, bactéries pathogènes, vers parasites, etc. La mise au point de mécanismes de défense contre ces parasites est un défi adaptatif permanent pour l'espèce hôte : en effet, la sélection naturelle pousse les parasites à mettre au point sans cesse de nouvelles stratégies de contournement des défenses de l'hôte. L'hôte est donc en permanence mis au défi d'évoluer plus vite que le parasite et inversement. Les deux espèces auront donc intérêt à maximiser leur potentiel d'évolution en maintenant une reproduction sexuée très fréquente.

**Pourquoi la reproduction sexuée obligatoire est-elle alors plus répandue chez les pluricellulaires que chez les unicellulaires ?** On peut émettre l'hypothèse d'une différence liée au temps de génération. Un même changement de l'environnement peut être perçu comme très lent par un protiste (qui fait plusieurs générations par jour) et très rapide par un mammifère comme l'homme (qui fait une génération tous les vingt-cinq ans). Il existe donc des conditions dans lesquelles l'homme a intérêt à se reproduire sexuellement à chaque génération (sexualité obligatoire), alors que le protiste a intérêt à se reproduire clinalement pendant un certain nombre de générations avant chaque épisode de sexe. Imaginons un protozoaire parasite de l'homme : le protozoaire, une fois rentré dans un hôte, produit des descendants qui restent dans le même environnement tant qu'ils ne sortent pas de cet hôte. Ils ont donc intérêt à se reproduire clinalement au moins jusqu'au moment de leur sortie en direction d'un nouvel hôte. Au contraire, le descendant d'un homme sera confronté à des parasites différents de ceux de ses

parents : l'homme aurait donc intérêt à pratiquer le sexe à chaque génération de reproduction.

## Que faire du sexe quand il est devenu obligatoire ?

**La reproduction sexuée implique le choix d'un partenaire.** De ce point de vue, les espèces animales et végétales montrent une grande diversité. Beaucoup d'espèces, comme le blé ou l'escargot *Bulinus truncatus*, pratiquent de manière très majoritaire l'*autofécondation* : les individus sont hermaphrodites et sont à la fois le père et la mère de leurs descendants. Cependant, les hermaphrodites peuvent également pratiquer l'*allofécondation*, c'est-à-dire le croisement entre individus différents. Certaines plantes s'assurent ainsi une alloffécondation quasi obligatoire par divers mécanismes d'*auto-incompatibilité* empêchant le pollen d'accéder aux ovules de la même plante. Ces mécanismes peuvent être physiologiques (maturation des organes mâles et femelles décalée dans le temps) ou chimiques (rejet de l'autopollen sur la base de molécules reconnues par les organes femelles). En général, une plante est soit très majoritairement autofécondante, soit très majoritairement alloffécondante, bien qu'il existe quelques espèces à taux d'autofécondation proche de 50 %<sup>1</sup>. Une situation similaire est retrouvée chez les escargots hermaphrodites.

Le moyen le plus simple d'empêcher l'autofécondation reste l'existence de sexes séparés ou *gonochorie*, répandue chez beaucoup d'animaux dont les vertébrés et, bien sûr, l'homme. Cependant, même quand les sexes sont sépa-

1. S. C. H. Barrett et L. D. Harder, « Ecology and Evolution of Plant Mating », *Trends in Ecology and Evolution*, 11, 1996, p. 73-79.

rés, les partenaires peuvent être plus ou moins proches dans le réseau généalogique (apparentés). La reproduction entre apparentés, dite *consanguine*, peut être considérée comme une forme atténuée de l'autofécondation. Dans le cas d'un croisement frère-sœur par exemple, les descendants n'ont que deux grands-parents au lieu de quatre : cette situation est analogue aux descendants d'autofécondation, qui n'ont qu'un parent au lieu de deux. Beaucoup d'espèces gonochoriques ont mis en place des comportements d'évitement des accouplements consanguins. Les souris femelles rejettent les mâles apparentés sur la base de l'odeur de leur urine, les chimpanzés mâles vont se reproduire dans une troupe différente de celle où ils sont nés (*exogamie*). Quant à l'homme, il a inscrit cet évitement dans ses coutumes ou dans ses lois (*tabou de l'inceste*). Pour comprendre la diversité des systèmes de reproduction, il faut rechercher quels sont les avantages et inconvénients, du point de vue de la sélection naturelle, de la consanguinité. Pour simplifier, nous considérerons surtout l'alternative la plus tranchée : autofécondation *versus* allofécondation.

**L'autofécondation transmet mieux les gènes que l'allofécondation.** Son premier avantage est d'éviter à l'individu la recherche d'un partenaire sexuel. Non seulement il économise ainsi du temps et de l'énergie, mais cela lui permet en outre de fonder une nouvelle population, même s'il se trouve seul dans son milieu. Cette *assurance de se reproduire* peut être particulièrement intéressante chez des espèces vivant dans des milieux instables qui ne doivent leur survie qu'à la capacité de coloniser sans cesse de nouveaux sites. Cependant, l'avantage majeur de l'autofécondation est la *fidélité de transmission* des gènes. Les ovules produits par un individu autofécondant ne contiennent que ses propres

gènes : la dépense énergétique de fabrication de ces ovules n'est donc pas « gaspillée » à véhiculer des gènes provenant d'un spermatozoïde étranger. Dans une population d'hermaphrodites, un individu qui s'autoféconde transmet des gènes par trois voies : ses ovules, ses spermatozoïdes fécondant les ovules d'autres individus, et ses spermatozoïdes fécondant ses propres ovules. Un hermaphrodite qui ne s'autoféconde pas perd la troisième voie et transmet donc moins de gènes. Les gènes qui déterminent l'autofécondation devraient donc se transmettre mieux que les autres et envahir les populations, si toutes choses étaient égales par ailleurs. Mais l'autofécondation a-t-elle des conséquences sur la qualité génétique des descendants ?

**L'autofécondation entraîne une uniformité génétique des deux chromosomes d'un individu.** Chez un diploïde, les deux copies de chaque gène, héritées l'une du père, l'autre de la mère, peuvent être de même type (*individu homozygote*) ou différentes (*hétérozygote*). En cas d'autofécondation, le père et la mère sont un seul et même individu ; les deux gènes ont donc la possibilité d'être des copies d'un même gène. En conséquence, on trouvera dans une population autofécondante plus d'homozygotes que dans une population allofécondante (voir p. 193). Si une population pratique 100 % d'autofécondation pendant de nombreuses générations, tous les individus deviendront homozygotes pour tous leurs locus, et leurs descendants seront des copies conformes d'eux-mêmes malgré l'existence de méioses. Cette situation extrême est donc équivalente à une reproduction clonale. Toutefois, elle est très difficile à trouver dans la nature : quelques rares épisodes d'allofécondation suffisent en effet à restaurer une hétérozygotie élevée et annuler l'effet précédent. Pour les mêmes raisons qu'on

observe rarement une clonalité stricte (voir p. 182), on observera rarement une autofécondation stricte. Beaucoup d'espèces ont néanmoins des taux d'autofécondation proches de (mais pas égaux à) 100 %.

**L'homozygotie entraîne une baisse de valeur sélective.** Nous avons déjà vu, grâce à l'exemple des maladies génétiques humaines, que des gènes délétères existent en faible fréquence dans les populations, et que ces mutations sont le plus souvent récessives, c'est-à-dire qu'elles ne s'expriment qu'à l'état homozygote. Les descendants d'autofécondation ont une probabilité élevée d'être homozygotes et ce, en particulier, pour des allèles délétères : ils auront donc en moyenne une valeur sélective plus faible que les descendants d'allofécondation. Ce phénomène, appelé *dépression de consanguinité*, a été observé chez de très nombreuses espèces, y compris l'espèce humaine. Cette dernière étant incapable de s'autoféconder, il faut, pour, observer la dépression de consanguinité, comparer les descendants de croisements consanguins (forme atténuée de l'autofécondation) à ceux de croisements entre individus non apparentés. Une étude menée à Strasbourg sur 238 942 naissances<sup>1</sup> montre que les malformations congénitales sont environ dix fois plus fréquentes chez les descendants de couples de cousins germains que chez ceux des couples non consanguins. Cette étude montre également que plus le père et la mère sont proches parents, plus les risques de malformation augmentent. D'autres exemples d'intérêt direct pour l'homme sont les plantes cultivées (voir *Hétérosis et dépression de consanguinité*, p. 319).

1. C. Stoll *et al.*, « Parental Consanguinity as a Cause for Increased Incidence of Births Defects in a Study of 238 942 Consecutive Births », *Annales de génétique*, 42, 1999, p. 133-139.

Pourquoi l'autofécondation évolue-t-elle dans certaines lignées et pas dans d'autres ? L'autofécondation a un avantage intrinsèque de fidélité de transmission des gènes, mais cet avantage est contrecarré par la dépression de consanguinité. Nous avons vu que l'avantage intrinsèque est dû au fait qu'il y a deux fois plus de gènes parentaux dans les descendants d'autofécondation que dans les descendants d'allofécondation : un parent fournit par autofécondation la totalité des gènes de son descendant, mais seulement la moitié par allofécondation. Cet avantage sera exactement compensé si les descendants d'autofécondation ont une valeur sélective deux fois moins bonne que ceux d'allofécondation (dépression de consanguinité égale à 1/2). Selon ce raisonnement, les espèces devraient évoluer vers 100 % d'autofécondation si la dépression de consanguinité est inférieure à 1/2, et 100 % d'allofécondation si elle est supérieure à cette valeur. Bien que ce raisonnement soit très simplifié, de nombreux modèles mathématiques ont montré la généralité et la robustesse de cette prédiction. Dans la nature, la plupart des espèces ont effectivement des taux d'autofécondation proches de zéro ou de 100 %, bien qu'ils n'atteignent presque jamais 100 % exactement (un taux de 100 % d'autofécondation correspond à une reproduction sexuée totalement inefficace, où jamais les gènes ne recombinent entre eux, et de telles espèces seraient condamnées au même sort que les espèces totalement asexuées).

La corrélation attendue entre l'intensité de la dépression de consanguinité et le système de reproduction est globalement observée : les espèces majoritairement autofécondantes ont une dépression de consanguinité faible ou nulle, alors que les espèces allofécondantes ont en général

une dépression élevée<sup>1</sup>. Considérons par exemple deux escargots d'eau douce de la famille des *Planorbidae* : le bulin (*Bulinus truncatus*) et la physe (*Physa acuta*). Ces deux espèces sont toutes deux hermaphrodites, et leurs caractéristiques écologiques sont proches. La mesure du degré d'hétérozygotie des individus pour des marqueurs moléculaires permet d'inférer le type de reproduction. Chez le bulin, presque tous les individus prélevés dans une population du Niger sont homozygotes pour quatre locus microsatellites, malgré l'existence de plusieurs allèles dans la population à chaque locus, ce qui indique un taux d'autofécondation très élevé, estimé à 80-90 %<sup>2</sup>. Chez la physe, des analyses similaires (sur des marqueurs enzymatiques) ont permis d'estimer ce taux à moins de 10 % dans la région de Montpellier<sup>3</sup>. Corrélativement, chez le bulin, aucune dépression de consanguinité significative n'a été mise en évidence ; chez les physes de la région de Montpellier, la dépression de consanguinité (perte de valeur sélective des descendants d'autofécondation) est estimée à plus de 90 %.

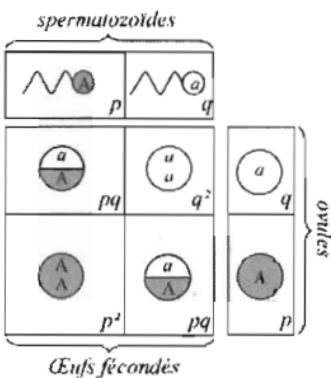
**La dépression de consanguinité dépend elle-même, à long terme, du taux d'autofécondation.** Nous avons jusqu'ici raisonné comme si la dépression était une donnée

1. J.S. Escobar *et al.*, « Patterns of Mating System Evolution in Hermaphroditic Animals : Correlations among Selfing Rate, Inbreeding Depression and the Timing of Reproduction », *Evolution*, 65, 2011, p. 1233-1253

2. F. Viard *et al.*, « Selfing, Sexual Polymorphism and Microsatellites in the Hermaphroditic Freshwater Snail *Bulinus truncatus* », *Proceedings of the Royal Society of London B*, 264, 1997, p. 39-44.

3. P. Jarne *et al.*, « The Influence of Self-Fertilization and Grouping on Fitness Attributes in the Freshwater Snail *Physa acuta* : Population and Individual Inbreeding Depression », *Journal of Evolutionary Biology*, 13, 2000, p. 645-655.

## A) Panmixie



## B) Autofécondation

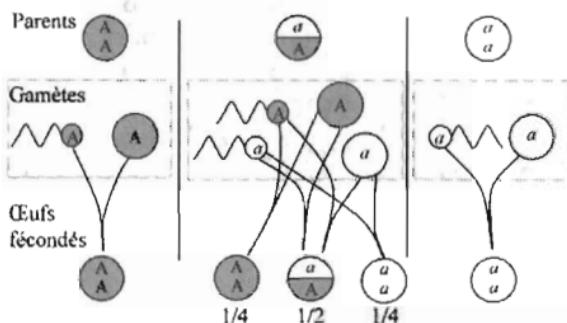


Figure 3 : voir encadré page suivante.

fixe, caractéristique d'une espèce. Or, elle repose sur la fréquence des allèles délétères récessifs dans les populations des espèces considérées. Ces allèles étant soumis à sélection, la dépression peut donc évoluer. Toutefois, cette évolution dépend elle-même du taux d'autofécondation. Dans une espèce non autofécondante, une mutation délétère récessive persistera longtemps dans une population : tant que sa fréquence n'est pas trop grande, elle est en effet surtout à l'état hétérozygote, ce qui lui permet d'échapper à la sélection. Au contraire, dans une population très autofécondante, les mutations délétères sont vite démasquées : mises à l'état homozygote par l'autofécondation, elles seront éliminées beaucoup plus efficacement par la sélec-

### Hétérozygotie et consanguinité

La figure 3A représente une population dans laquelle la consanguinité est totalement absente : les individus se croisent au hasard. Cette situation est appelée *panmixie* (littéralement : mélange total). Dans ce cas, tout se passe comme si on mettait en contact l'ensemble des spermatozoïdes produits par les mâles de la population avec l'ensemble des ovules produits par les femelles. On suppose qu'il existe un locus à deux allèles, A et *a*, présents dans la population en fréquence  $p$  et  $q$  ( $= 1 - p$ ) : on a donc une proportion  $p$  de spermatozoïdes de type A, et une proportion  $q$  de spermatozoïdes de type *a*, et il en va de même pour les ovules. La probabilité qu'un spermatozoïde A rencontre un ovule A est  $p \times p = p^2$ . Le schéma de croisement donne également les probabilités des autres configurations possibles. On en déduit un théorème dit de *Hardy et Weinberg* :

*Une génération de panmixie suffit à obtenir les fréquences génotypiques suivantes dans les œufs fécondés :  $p^2$ ,  $2pq$ ,  $q^2$  pour AA, Aa et aa respectivement, si les fréquences alléliques sont p (pour A) et q (pour a).*

Cet énoncé a été formulé au début du XX<sup>e</sup> siècle (1906) en même temps que le principe, de nature différente, vu au chapitre 1.

Que se passe-t-il dans une population consanguine ? Le cas le plus extrême de consanguinité est l'autofécondation, illustrée dans la figure 3B. Dans ce cas, seuls les spermatozoïdes et ovules produits par un même individu peuvent se rencontrer. La figure montre que les individus homozygotes (AA et aa) ne donnent que des descendants homozygotes. En revanche, le descendant d'un hétérozygote (Aa) n'a qu'une probabilité 1/2 d'être également hétérozygote. Si tous les descendants survivent de la même façon, la proportion d'hétérozygotes devrait donc être divisée par deux à chaque génération d'autofécondation stricte dans une population. À ce rythme, les hétérozygotes disparaissent rapidement de la population. Des formes de consanguinité moins sévères (autofécondation partielle, reproduction entre apparentés) aboutissent également à diminuer l'hétérozygotie de la population, bien qu'elles ne l'annulent pas complètement.

tion naturelle. L'autofécondation permet donc à la population de se purger rapidement de ses allèles délétères, ce qui diminue la dépression de consanguinité. En conclusion, l'évolution de l'autofécondation est un processus auto-entretenu : l'autofécondation fait diminuer la dépression de consanguinité ce qui favorise l'augmentation du taux d'autofécondation. L'allofécondation permet à la dépression de consanguinité d'augmenter, ce qui favorise l'augmentation du taux d'allofécondation. Dans cette optique, les taux d'autofécondation proches de 50 % ne pourraient être que des situations provisoires. Rappelons cependant qu'il s'agit là d'une vision simplifiée du problème : des modèles théoriques, que nous ne détaillerons pas, prédisent que sous certaines conditions de tels taux peuvent être stables.

## Résumé

La capacité à pratiquer le sexe, mais aussi la façon de le pratiquer, conditionnent les modalités de transmission de l'information génétique, et par là même les conditions de l'évolution de tous les caractères. Le sexe lui-même est également un caractère présentant une diversité de modalités et est donc susceptible d'évoluer. La grande diversité des pratiques observées chez les espèces actuelles en est une illustration.

La reproduction clonale et l'autofécondation ont un avantage immédiat sur la reproduction sexuée et l'allofécondation (respectivement), car elles permettent une transmission fidèle des gènes parentaux aux descendants : l'énergie investie par la mère pour produire des œufs est entièrement consacrée à propager ses propres gènes, et non à propager aussi ceux d'un individu mâle. La repro-

duction clonale stricte ou l'autofécondation stricte sont néanmoins très rares car elles entraînent à long terme une incapacité de l'espèce à trouver rapidement de nouvelles combinaisons génétiques. Elle ne peut donc s'adapter aux changements de l'environnement. FAVORABLE au niveau individuel, la reproduction clonale est donc défavorable au niveau de l'espèce. En fonction de la fréquence relative des changements d'environnement et des phases de reproduction, le sexe à chaque événement de reproduction (reproduction sexuée obligatoire), ou seulement de temps en temps (reproduction sexuée facultative) peut être avantage à l'échelle d'une espèce.

L'évolution de l'autofécondation est, quant à elle, déterminée principalement par une sélection au niveau de l'individu. L'autofécondation a pour conséquence une diminution de valeur sélective des descendants (dépression de consanguinité). Si cette diminution compense ou dépasse l'avantage dû à la fidélité de transmission, l'autofécondation sera défavorisée ; dans le cas contraire, elle sera favorisée. L'autofécondation peut ainsi être très fréquente, mais les espèces la pratiquant de manière exclusive sont très rares car, comme les espèces totalement asexuées, elles ont à long terme une faible capacité d'adaptation et une probabilité élevée de s'éteindre.

## DES GÈNES AUX CARACTÈRES

Les chapitres précédents s'inscrivent principalement dans le cadre de la génétique des populations. Née au début du XX<sup>e</sup> siècle, cette discipline modélise l'évolution à l'aide des concepts formels de la génétique mendélienne : gènes, locus, allèles. Son efficacité est maximale lorsque les allèles considérés correspondent directement à des caractères simples, comme la résistance aux insecticides. Malheureusement, il est frustrant de constater que le cadre théorique très élaboré de la génétique des populations reste impuissant à expliquer l'évolution de la majorité des caractères observables. En effet, l'évolution la plus facilement observable, sur des restes fossiles ou sur des organismes actuels, est celle de traits individuels tels que la taille ou la forme d'organes (cf. l'exemple des variations de la forme de la coquille de l'escargot *Littorina obtusata*, vu au chapitre 4). Bien que ces traits soient transmis de parent à descendant, ils ne dépendent pas d'un gène précis, mais de nombreux gènes souvent inconnus. Leur niveau d'intégration est celui de l'individu, non du gène. Comment décrire l'évolution de ces caractères individuels ? Deux démarches complémentaires peuvent être envisagées :

(1) La démarche « holiste » consiste à élaborer des modèles décrivant le comportement « macroscopique » des caractères héréditaires sans se préoccuper du niveau

« microscopique », c'est-à-dire des gènes particuliers qui les déterminent. Il faut, dans ce but, inventer une génétique différente de la génétique mendélienne : il s'agit de la *génétique quantitative*.

(2) La démarche « réductionniste » consiste, à l'inverse, à disséquer le lien entre les caractères individuels complexes et les gènes. Ce lien n'est pas direct : les gènes interagissent les uns avec les autres tout au long de la construction de l'organisme. L'étude de ces interactions est l'objet de la *génétique du développement*.

### L'évolution des caractères quantitatifs, un modèle alternatif à la génétique des populations

L'hérédité des caractères continus n'est pas bien décrite par la génétique mendélienne. Au début du XX<sup>e</sup> siècle, la génétique mendélienne a permis de décrire la transmission de caractères discontinus, comme la couleur jaune ou verte des petits pois. Les modèles de Mendel furent validés par de nombreux résultats expérimentaux portant sur ces caractères discontinus. Cependant, les caractères qui varient sur une échelle continue, comme la taille du corps, s'accordent mal des modèles mendéliens. Pour obtenir une distribution continue à partir d'un locus « *taille* », il faudrait effectivement imaginer une infinité d'allèles correspondant chacun à une taille particulière. Même si de tels caractères ne répondent pas au modèle mendélien, ils sont indéniablement héritables : la taille des descendants est en général proche de celle de leurs parents. Ces caractères à spectre de variation continu nécessitent donc une modélisation génétique différente, qui ne fasse pas appel

à des locus ou à des allèles. Ces modèles sont l'objet de la génétique quantitative. Les pionniers de cette approche (appelés *biométriciens*) l'envisageaient comme un modèle alternatif au modèle mendélien, ce qui donna lieu à une longue controverse au début du XX<sup>e</sup> siècle (voir *Biométriciens et Mendéliens*, p. 325). Cette controverse s'acheva par la victoire des mendéliens, soutenus par des preuves expérimentales de plus en plus nombreuses. Néanmoins, la génétique mendélienne restait impuissante à expliquer des variations continues comme celle de la taille. Les modèles de la génétique quantitative sont donc encore utilisés. Actuellement, on ne les considère plus comme une alternative mais comme un complément aux modèles mendéliens. En effet, il est admis que tous les caractères reposent *in fine* sur des locus et allèles à comportement mendélien. Cependant, pour la plupart des caractères continus, en particulier la taille du corps, les locus et les allèles impliqués sont nombreux ou mal connus, et seuls les modèles quantitatifs permettent de prédire la transmission de ces caractères.

Le modèle de base de la génétique quantitative décrit la variation d'un caractère comme la somme d'une composante génétique et d'une composante environnementale. Dire que les descendants ressemblent à leurs parents (ce qui est l'essence de l'hérédité) n'a de sens qu'en référence à une comparaison implicite : les descendants ressemblent à leurs parents *plus qu'à un individu pris au hasard dans la population*. La génétique quantitative part donc du principe qu'un caractère ne peut être considéré comme héritable que par rapport à une population. Par exemple, on dira que la taille est un caractère héréditaire si les individus plus grands que la moyenne de la population donnent des descendants également plus grands que la moyenne. La variable utile en génétique

quantitative n'est donc pas la taille d'un individu elle-même mais la différence de taille entre cet individu et la moyenne de la population, ou *déviation phénotypique* (notée  $P$ ).  $P$  peut être positive ou négative selon les individus, et, par construction, sa moyenne est nulle dans la population.

Deux raisons, non exclusives, peuvent expliquer pourquoi un individu est plus grand ou plus petit que la moyenne. Premièrement, cela peut tenir à ses gènes. Deuxièmement, cela peut dépendre de ses conditions de vie, par exemple de la nourriture qu'il a reçue pendant sa croissance. Notre modèle de base intègre ces deux composantes dans l'équation  $P = G + E$ , où  $G$  représente la valeur codée par les gènes et  $E$  la contribution de l'environnement au sens large (conditions de vie). De même que  $P$ , les variables  $G$  et  $E$  prennent des valeurs positives ou négatives, et leur moyenne est nulle. La valeur de  $G$  est fixée par le génotype. Ainsi, deux vrais jumeaux, qui ont exactement le même génotype, ont la même valeur de  $G$ . Cependant ils peuvent avoir des valeurs de  $E$  différentes, par exemple s'ils ont été séparés à la naissance et élevés dans des conditions différentes. La plupart du temps on n'est capable ni de connaître parfaitement les conditions de vie de tous les individus, ni de savoir exactement comment elles affectent le caractère observé. On considère donc que  $E$  varie aléatoirement, de manière indépendante de  $G$ .

La variabilité du phénotype peut être mesurée par la dispersion des valeurs de  $P$  autour de leur moyenne (zéro). Cette dispersion est quantifiée par une mesure statistique, la variance de  $P$  (notée  $V_P$ ). Si l'on représente la distribution des valeurs de  $P$  par une courbe de fréquences,  $V_P$  correspond à « l'étalement » de la courbe : plus  $V_P$  est élevée, plus les individus ont des phénotypes

différents les uns des autres.  $V_P$  peut être évaluée facilement à partir d'un jeu de mesures de  $P$  obtenues sur un échantillon d'individus d'une population, en faisant la moyenne de ces valeurs mises au carré. Puisque  $P$  est la somme d'une composante génétique et d'une composante environnementale ( $P = G + E$ ), la variance de  $P$  est la somme de la variance de la composante génétique  $G$ , et de celle de la composante environnementale  $E$ . On obtient donc  $V_P = V_G + V_E$ . En d'autres termes, la variation de taille résulte d'une variation génétique et d'une variation environnementale.

Un caractère sera déterminé d'autant plus fortement par les gènes que  $V_G$  sera grand devant  $V_E$ . On appelle *héritabilité au sens large* ( $H^2$ ) le rapport  $V_G/V_P$ . Quand  $V_P$  est grand par rapport à  $V_G$ , ce rapport est proche de zéro, l'effet des gènes est négligeable devant les effets environnementaux. Quand  $V_P$  est proche de  $V_G$ , le rapport est proche de un, le caractère est déterminé entièrement par les gènes et très peu par l'environnement ; il est donc génétiquement transmis de manière très fidèle (héritable).

La connaissance de l'héritabilité permet de prédire  $P$  chez un individu dont on connaît le parent, avec une certaine marge d'erreur (figure 1). Dans le cas d'une reproduction asexuée, tous le génome du parent est transmis en bloc à ses descendants. La valeur de  $G$  est donc la même entre parent et descendant. En revanche, dans le cas d'une reproduction sexuée, seule une moitié des gènes du parent est transmise au descendant. En conséquence la valeur de  $G$  ne sera pas transmise intégralement. Il ne s'agit cependant pas d'une simple division par deux de  $G$  : la valeur de  $G$  du parent dépend en effet non seulement des gènes présents, mais de leurs interactions verticales (dominance) ou horizontales (épistasie) (voir

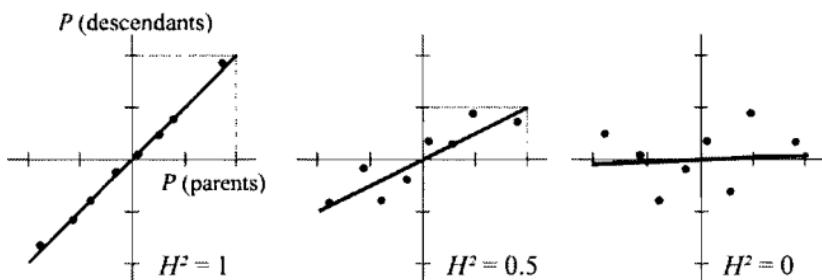


Figure 1 : Diverses valeurs d'héritabilité correspondent à divers degrés de ressemblance parent-descendant. Pour les trois cas présentés, la reproduction est asexuée et  $H^2$  désigne l'héritabilité au sens large. Dans chaque graphique, un point représente le phénotype d'un parent (en abscisse) et de son enfant (en ordonnée). Quand  $H^2 = 1$ , le caractère est entièrement déterminé par les gènes. Comme ceux-ci sont entièrement transmis *via* la reproduction asexuée, les descendants ont toujours le même phénotype que leur parent (pente de la droite de régression = 1). Quand  $H^2 = 0$ , le caractère est entièrement déterminé par l'environnement. Parents et descendants ne doivent leur variation de taille qu'à l'environnement dans lequel ils se développent. En supposant qu'un descendant ne connaît pas nécessairement les mêmes conditions de vie que ses parents, les descendants de parents grands n'ont pas une taille plus grande en moyenne que les descendants de parents petits (pente de la droite = 0). En revanche, on observe une forte dispersion des points, reflétant les aléas environnementaux.  $H^2 = 0,5$  correspond à une situation intermédiaire. Dans tous les cas, la pente de la droite égale la valeur de l'héritabilité. Si l'on connaît  $H^2$ , on peut donc construire la droite et l'utiliser pour prédire le phénotype d'un descendant connaissant son parent. Cette prédiction est d'autant plus précise que  $H^2$  est élevée. Si la reproduction est sexuée, les résultats restent les mêmes, mais il faut substituer  $H^2$  par  $h^2$  l'héritabilité au sens étroit, et utiliser la moyenne de  $P$  des deux parents.

chapitre 6). Lors de la reproduction sexuée, la méiose défait complètement les associations verticales et la recombinaison modifie les associations horizontales. La valeur de  $G$  peut être décomposée en une partie liée aux interactions, non transmise par la reproduction sexuée, et une partie indépendante des interactions, notée  $A$ , qui se transmet par la reproduction sexuée.  $A$  est traditionnellement appelée *valeur reproductive (breeding value)*. Le rapport  $V_A/V_P$  désigne l'*héritabilité au sens étroit*, c'est-à-dire la part de variation phénotypique transmise de parent à descendant lors de la reproduction sexuée.

La génétique quantitative fut d'abord appliquée à la sélection artificielle en vue d'améliorer les espèces domestiques. Lors de la domestication d'espèces à usage agricole, l'homme a opéré une sélection très forte mais inconsciente, portant essentiellement sur des caractères continus, tels que le poids de grains produits par une plante (voir *Amélioration génétique des espèces*, p. 315). Avec la naissance de la génétique quantitative, au début du XX<sup>e</sup> siècle, l'homme a cherché à rationaliser cette sélection. Dans ce cadre, la question la plus importante concerne la vitesse de réponse à la sélection. Par exemple, pour augmenter la production de lait chez une race bovine, il faut à chaque génération retenir les meilleures vaches laitières et les faire se reproduire. Suivant l'héritabilité du caractère « production de lait » dans le troupeau, le nombre de générations nécessaires pour augmenter de manière intéressante son rendement ne sera pas le même. L'éleveur a donc intérêt à pouvoir prédire par avance la *réponse à la sélection*, c'est-à-dire le changement de la moyenne du phénotype (ici, la production de lait) obtenu en une génération de sélection artificielle. En effet, si les modèles prédisent, par exemple, que deux cents générations de sélection (donc,

plusieurs centaines d'années) sont nécessaires, le sélectionneur fera mieux de chercher une autre stratégie de sélection ou un autre troupeau de départ.

Les calculs de génétique quantitative donnent un résultat remarquablement simple : la réponse à la sélection est  $R = h^2 S$ , où  $h^2$  est l'héritabilité au sens étroit, et  $S$  le *differentiel de sélection*, c'est-à-dire la différence entre la moyenne des individus reproducteurs gardés par le sélectionneur (la production de lait des vaches sélectionnées) et la moyenne du caractère dans la population de départ. Avant de se lancer dans un programme de sélection, l'améliorateur mesurera l'héritabilité du caractère choisi dans son troupeau (par exemple, en utilisant une comparaison parent-enfant comme dans la figure 1). Cela lui permettra de prédire de combien la production moyenne de lait augmentera au cours du temps, en fonction de la sélection qu'il est capable d'imposer.

Cette prédiction est cependant à court terme, la sélection artificielle ne faisant que trier, parmi tous les gènes présents dans une population de départ, ceux qui agissent sur un caractère dans une direction précise (par exemple, augmentation de la production de lait). À chaque locus impliqué dans la production du lait, les allèles stimulant cette production augmentent progressivement en fréquence avec la sélection, et finissent par se fixer. Lorsque tous ces allèles sont fixés, la population ne peut plus progresser génétiquement, puisque tous les individus possèdent les mêmes allèles. Il y a donc un épuisement de la variation génétique qui ne peut être contrecarré que par des mutations introduisant de nouveaux allèles (si l'on exclut l'introduction d'individus extérieurs). De telles mutations sont rares et leur arrivée est aléatoire. En général, le sélectionneur ne peut se permettre d'attendre leur apparition. On prédit donc que

L'héritabilité du caractère diminue au fil des générations de sélection, pour atteindre une valeur nulle quand la population obtient le phénotype maximum possible. En général, si l'on part d'une population de départ à forte diversité génétique, l'héritabilité décroît assez lentement pour qu'une amélioration génétique importante soit obtenue.

Une expérience très démonstrative a été menée chez le maïs<sup>1</sup> (*Zea mays*). À partir d'une même population de départ, deux lignées, « grasse » et « maigre », furent construites en retenant les épis ayant respectivement le contenu en huile le plus élevé ou le plus faible. L'expérience, commencée en 1896, s'est poursuivie pendant soixante-seize générations. Le contenu en huile moyen, initialement de 5 %, augmenta régulièrement dans la lignée grasse (jusqu'à environ 20 %) et diminua dans la lignée maigre (jusqu'à moins de 0,5 %). Au bout de soixante-seize générations, on observait encore une réponse à la sélection bien qu'elle ait progressivement faibli depuis 1896. Ainsi, lors des dix premières générations, les héritabilités calculées étaient de 0,3 pour la lignée grasse et de 0,5 pour la lignée maigre. Après soixante-seize générations elles n'étaient plus que de 0,12 et 0,15 respectivement.

Cette expérience illustre également une des caractéristiques communes à beaucoup d'études de ce type : les phénotypes obtenus après de nombreuses générations de sélection étaient des phénotypes nouveaux, très en dehors de la gamme initiale de variation. Autrement dit, personne n'avait jamais, avant cette expérience, observé

---

1. J. W. Dudley, « Seventy-Six Generations of Selection for Oil and Protein Percentage in Maize », in E. Pollack et al., *Proceedings of the International Conference on Quantitative Genetics*, Ames, Iowa, Iowa State University Press, 1977, p. 459-473.

de grains de maïs aussi gras ou aussi maigres que ceux obtenus dans les lignées sélectionnées. Il n'y a pourtant pas besoin d'imaginer que de nouvelles mutations soient apparues au cours de la sélection pour comprendre comment ces variants ont pu émerger. Supposons que de nombreux allèles, augmentant chacun un peu la teneur en graisses, existaient dans la population de départ, mais qu'ils étaient rares et dispersés dans des individus différents. À l'issue de la sélection, ces allèles sont devenus fréquents et peuvent être présents simultanément dans un même individu, où leurs effets se cumulent. Cette accumulation confère à l'individu une très haute teneur en graisses, inobservée jusque là. Cet exemple montre qu'il existe en permanence des réserves de variation génétique qui ne s'expriment pas sur la variation phénotypique, mais qui permettent à un moment donné à la sélection naturelle de faire émerger des phénotypes nouveaux en quelques générations. Parce qu'elle semble seulement trier dans la variation existante, beaucoup doutent de la capacité de la sélection naturelle à faire émerger rapidement des nouveautés importantes. C'est oublier que la variation phénotypique visible à un moment donné ne représente qu'une petite fraction de la variation génétique sous-jacente.

**La génétique quantitative permet également d'analyser l'évolution des caractères phénotypiques dans la nature.** Les effets à court terme de la sélection naturelle sur des traits quantitatifs peuvent être étudiés dans le même cadre que ceux de la sélection artificielle. On se fonde alors sur des valeurs d'héritabilité que l'on détermine expérimentalement sur un échantillon d'individus (comparaison parent-enfant), et que l'on suppose constantes pendant au moins quelques générations.

L'héritabilité d'un trait dans une population naturelle est une mesure du potentiel d'évolution à court terme de ce trait.

Une application de ce raisonnement a été proposée par les époux Grant<sup>1</sup> (1993) sur le cas des pinsons de Darwin (*Geospiza fortis*) des îles Galápagos. Chaque espèce de pinson est caractérisée par une taille particulière du bec, qui lui permet de consommer de manière optimale une catégorie précise de graines. Lors d'une année particulièrement sèche (événement « *El Niño* » de 1982), la production de graines, dans l'île où vit *Geospiza fortis*, fut très affectée, la taille moyenne des graines diminuant considérablement. Des pinsons furent marqués et suivis individuellement à l'aide de bagues, et il s'avéra que les oiseaux à bec plus petit que la moyenne, mieux adaptés à l'exploitation de petites graines, survivaient mieux que les autres. Pour plusieurs caractères (dont la longueur et la largeur du bec), il fut possible d'estimer un différentiel de sélection en calculant la différence entre la moyenne des survivants et celle de l'ensemble de la population. Par ailleurs, la mesure simultanée de ces caractères chez les parents et leurs descendants donna des valeurs d'héritabilité (méthode illustrée figure 1, p. 202). Les deux caractères (longueur et largeur du bec) avaient des héritabilités élevées ( $b^2 = 0,65$  et  $0,90$  respectivement) mais les différentiels de sélection ( $S$ ) n'étaient pas les mêmes. Les oiseaux à bec moins long que la moyenne étaient très légèrement favorisés ( $S = -0,03$  mm), alors qu'un bec moins large que la moyenne augmentait fortement la survie ( $S = -0,17$  mm). On prédit, grâce à

1. B. R. Grant et P. R. Grant, « Evolution of Darwin's Finches Caused by a Rare Climatic Event », *Proceedings of the Royal Society of London B*, 251, 1993, p. 111-117.

l'équation  $R = h^2 S$ , une réponse à la sélection très faible pour la longueur du bec (- 0,02 mm, indétectable en pratique), mais mesurable pour la largeur (- 0,15 mm). Les mesures effectuées l'année suivante confirmèrent cette prédiction : la longueur moyenne du bec ne changea pas de manière significative, alors que la largeur moyenne diminua de 0,12 mm (correspondant aux prédictions, à l'erreur expérimentale près).

**À long terme, la variation génétique dans les populations naturelles est érodée par la sélection naturelle et la dérive génétique, mais régénérée par la mutation.** L'approche de la génétique des populations envisage la *mutation* comme la source de la variabilité (chapitre 2), alors que la dérive (chapitre 3) et la sélection (chapitre 4) sont des facteurs d'érosion de cette variabilité. La prise en compte de ces trois facteurs dans des modèles de génétique quantitative est complexe. Ils aboutissent néanmoins à des résultats qualitativement identiques à ceux de la génétique des populations. La diversité intrapopulation pour les caractères quantitatifs est engendrée par la mutation, et réduite sous l'effet de la sélection naturelle et de la dérive. Cette dernière engendre cependant une diversité à une échelle supérieure, entre des populations différentes (figure 2). La diversité génétique, qu'elle soit mesurée par  $H$  (l'hétérozygotie) sur des locus à comportement mendélien (génétique des populations) ou par  $V_P$  (variance phénotypique) sur des caractères phénotypiques (génétique quantitative), répond donc de la même façon aux trois pressions évolutives majeures (*mutation, sélection, dérive*).

**Peut-on réconcilier la génétique quantitative et la génétique mendélienne ?** La génétique quantitative et la génétique des populations reposent sur des formalismes

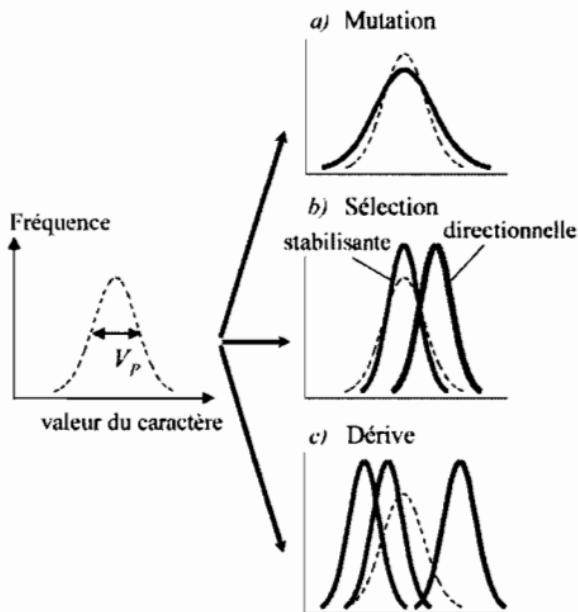


Figure 2 : Impact de la mutation, de la sélection naturelle et de la dérive génétique sur la distribution d'un caractère quantitatif. Ces effets interviennent tous en même temps dans la nature, mais ils ont été représentés séparément pour plus de clarté. La quantité de variation présente dans la population est mesurée par la variance phénotypique  $V_p$  qui correspond visuellement à l'étalement de la distribution. (a) La mutation a pour effet d'augmenter cette variance car les gènes mutants provoquent des variations génétiques aléatoires dans l'un ou l'autre sens. (b) La sélection diminue la variance en érodant la variabilité génétique. Dans le cas d'une sélection directionnelle (phénotypes extrêmes favorisés), la moyenne de la distribution se déplace vers les valeurs favorisées par la sélection. Dans le cas d'une sélection stabilisante (phénotypes moyens favorisés), la distribution se resserre autour de sa moyenne. (c) La dérive diminue la variance phénotypique dans une population mais augmente les différences entre populations (ici, trois populations possibles sont représentées).

différents mais donnent des visions similaires des processus évolutifs. Ces deux formalismes sont-ils cohérents entre eux ? De même que la pression d'un gaz s'explique par des mouvements moléculaires, on aimerait montrer que les variables macroscopiques des modèles quantitatifs (la déviation phénotypique  $P$ , la valeur génétique  $G$ , la valeur environnementale  $E$ , la valeur reproductive  $A$ ) peuvent être interprétées, et leur comportement prédit, en fonction des variables microscopiques mendéliennes (locus, allèles). Ronald Fisher<sup>1</sup> aborda ce problème par un modèle, en supposant que la valeur de  $P$  est contrôlée par un locus à deux allèles chez une espèce diploïde sexuée. Il montra que la transmission mendélienne des gènes avait pour conséquence une transmission des phénotypes selon le mode illustré par la régression parent-enfant (figure 1), et établit des formules de correspondance entre les variables quantitatives ( $A$ ,  $G$ ) et les caractéristiques des allèles (effet sur le phénotype, relations de dominance/récessivité, fréquence dans la population). À ce stade, deux problèmes se posent. Premièrement, la variable  $E$  n'a pas d'équivalent mendélien : cela est une insuffisance de la génétique mendélienne. Les caractères discontinus employés en génétique mendélienne sont en effet le fruit d'une simplification mettant de côté une partie de la variation : les petits pois, attribués volontairement par l'expérimentateur à deux classes, jaunes et verts, peuvent en fait être plus ou moins jaunes ou plus ou moins verts. La variation environnementale ( $V_E$ ) existe donc pour tous les caractères. Les caractères choisis

---

1. R. Fisher, « The Correlation Between Relatives and the Supposition of Mendelian Inheritance », *Transactions of the Royal Society of Edinburgh*, 52, 1918, p. 399-433.

pour valider le modèle mendélien sont tels qu'il est possible de l'occulter, ce qui est impossible avec des caractères comme la taille du corps. Deuxièmement, les modèles quantitatifs supposent une distribution de  $P$ ,  $A$  et  $G$  de type continu et normal (en forme de cloche, voir figure 1). Cependant, avec un locus à deux allèles, on n'obtient que trois génotypes. En ajoutant une variable environnementale, on étale la distribution du caractère, mais il reste toujours trois pics. Ce problème peut être résolu en supposant que le phénotype est influencé non par un seul locus mais par la somme des effets d'un très grand nombre de locus, chacun ayant un faible impact. Cette supposition est irréaliste car les locus influençant un trait donné sont forcément en nombre limité. De plus, leurs effets peuvent ne pas être faibles ou ne pas s'additionner entre eux (épistasie, voir chapitre 6). Des modèles plus subtils ont tenté de prendre en compte ces aspects. On peut tirer comme conclusion générale de ces modèles que la distribution normale utilisée dans les modèles quantitatifs est une approximation robuste, proche de la réalité dans la plupart des cas envisagés. Cela explique l'efficacité des modèles quantitatifs pour prédire l'évolution naturelle et artificielle à long terme. Au prix de quelques hypothèses, on peut donc réduire la génétique quantitative à la génétique mendélienne.

**Les bases mendéliennes de la génétique quantitative peuvent aussi être recherchées empiriquement.** La technique des QTL (*Quantitative Trait Loci*) a été mise au point dans ce but. Il s'agit, pour un caractère dont la variation génétique a été établie de manière quantitative, de localiser les locus particuliers qui déterminent cette variation (QTL). Le principe de cette technique est illustré en figure 3.

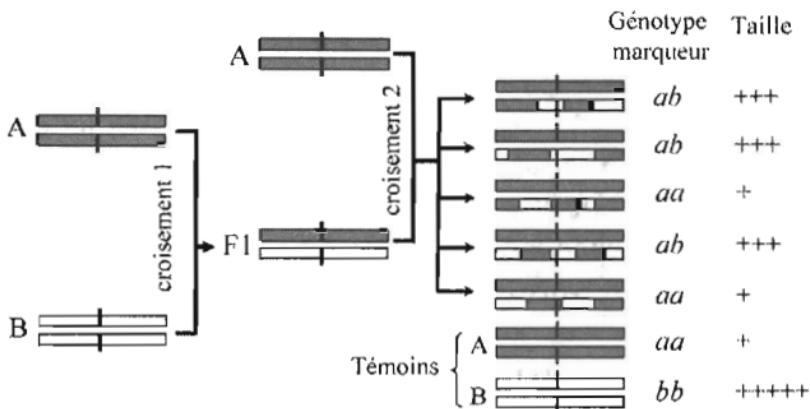


Figure 3 : Principe de détermination d'un QTL pour le caractère taille. On part de deux souches qui diffèrent par la taille : A (petits) et B (grands). Cette différence repose sur la présence de gènes différents : quelque part sur les chromosomes B (blancs) se trouvent des gènes qui confèrent une grande taille (gènes « *grands* »). Le but de l'expérience est de les localiser précisément sur le chromosome. Pour cela on fait deux croisements : le croisement 1 (lignée A x lignée B) fournit des descendants tous identiques (F1). Lors du croisement 2, les F1 sont recroisées avec le parent A (*backcross*) et donnent une collection d'individus différents. Suite aux recombinaisons, chacun d'eux possède un chromosome d'origine A (gris) et un chromosome mosaïque avec des régions A (grises) et des régions B (blanches). Parmi ces individus, les plus grands sont ceux qui ont hérité, dans le chromosome mosaïque, de fragments de chromosomes B (blancs) sur lesquels se trouvaient des gènes « *grands* ». Si on observe que les individus ayant hérité d'un fragment B (blanc) dans une région du chromosome sont en moyenne plus grands que ceux qui ont hérité d'un fragment A (gris) au même endroit, on en conclut qu'il y a dans cette région un locus agissant sur la taille. Pour repérer ces régions, il suffit donc de mesurer les individus et de savoir, pour chacun, quelles parties du chromosome mosaïque sont d'origine B. On utilise à cet effet une batterie de marqueurs

L'approche des QTL fut suivie pour identifier les gènes responsables des différences morphologiques entre le maïs (*Zea mays mays*) et la téosinte (*Zea mays parviglumis*), son ancêtre sauvage<sup>1</sup>. La téosinte possède une tige ramifiée, chaque branche portant un petit épi femelle, alors que le maïs n'est pas ramifié, et concentre tout son effort reproducteur femelle sur un seul gros épi situé en haut de la tige. Grâce au protocole schématisé en figure 2, furent isolées deux petites régions de chromosomes responsables de presque toutes ces différences d'architecture. Dans l'une de ces régions se trouvait le locus *Tenosinte Branched 1* (*tb1*). L'implication de ce gène dans le contrôle de l'architecture végétale était connue grâce à des mutants obtenus sur le maïs. Il fut ensuite confirmé que maïs et téosinte n'avaient effectivement pas les mêmes allèles au locus *tb1*. La domestication du maïs a donc impliqué une étape de sélection (peut-être inconsciente) favorisant l'architecture de type « maïs », à l'issue de laquelle un allèle particulier du locus *tb1*, impliqué

génétiques répartis tous le long du chromosome : la présence d'un allèle typique de la souche B pour un marqueur montre que la région qui l'entoure est d'origine B. Nous avons ici représenté un seul des nombreux marqueurs utilisés (repéré sur chaque chromosome par une barre noire). Pour ce marqueur, l'allèle typique de la souche A est noté *a*, et l'allèle typique de B est *b*. Seuls les individus *ab* ont hérité d'un fragment blanc au voisinage du locus marqueur. Ces individus ayant une taille supérieure aux *aa*, il existe dans cette zone un locus (QTL) agissant sur la taille.

1. J. Doebley, « Mapping the Genes that Made Maize », *Trends in Genetics*, 8, 1992, p. 302-307.

dans cette architecture, s'est fixé dans les populations domestiquées<sup>1</sup>.

### Le développement et son évolution : du génotype au phénotype

Les QTL donnent un moyen de relier l'évolution des caractères à celle des gènes, en partant de la variation du caractère pour aller vers les gènes. Une autre démarche, celle de la génétique du développement, consiste à modifier aléatoirement les gènes par des mutations, puis à en étudier les effets sur la mise en place des caractères au cours du développement. Le *développement* ou *ontogenèse* est défini comme l'ensemble des étapes permettant à une cellule unique (l'œuf fécondé) de devenir un organisme adulte complet. C'est au cours de ces étapes qu'il faut chercher les événements et les gènes qui déterminent les caractères de l'adulte. La génétique du développement s'est développée à partir des années 1960 grâce à l'étude de mutants obtenus initialement chez la drosophile, mais avant cela l'étude du développement par des méthodes non génétiques (embryologie) avait déjà permis d'en appréhender les mécanismes.

L'essor de la biologie du développement au XIX<sup>e</sup> siècle, avant celui de la génétique, est associé à la théorie de l'évolution. L'embryologiste Ernst Haeckel en est la figure de proue. Il échafaude la *théorie de la recapitulation*, selon laquelle les stades embryonnaires successifs représentent la morphologie de la lignée d'ancêtres d'un organisme actuel. On résume classiquement cette

1. J. Doebley *et al.*, « The Evolution of Apical Dominance in Maize », *Nature*, 386, 1997, p. 485-488.

théorie par la phrase : « L'ontogenèse récapitule la phylogénèse. » Ainsi, l'embryon de mammifère est, à un stade précoce, pourvu de fentes branchiales, typiques d'ancêtres de type poisson. La première implication de cette théorie est que l'embryologie fournit un moyen direct de reconstituer la phylogénie et l'histoire évolutive des organismes. Une deuxième implication concerne l'évolution du développement embryonnaire lui-même. Celle-ci doit nécessairement se faire sur le mode de l'*addition terminale* : de nouvelles étapes sont ajoutées par l'évolution en fin de développement, si bien que l'embryon d'un organisme actuel représente le stade adulte d'une espèce ancêtre.

Les recherches ultérieures ont montré que cette théorie était trop caricaturale. Ainsi, le placenta, une innovation évolutive des mammifères, se met en place très tôt dans l'embryogenèse et ne correspond à aucun caractère adulte ancestral. La logique générale de l'addition terminale ne peut cependant pas être complètement rejetée. Les événements de développement s'organisent en effet selon une *hiérarchie fonctionnelle*. Par exemple, la différenciation du système nerveux dépend chez les vertébrés d'une *induction*, c'est-à-dire de l'envoi de messagers chimiques par un autre tissu (le mésoderme) qui s'est différencié avant lui. Les nouveaux événements de développement qui apparaissent au cours de l'évolution tendent donc à se placer sous la dépendance d'événements préexistants, ce qui traduit une forme d'addition terminale en termes fonctionnels (et non temporels, comme dans la théorie de la récapitulation).

Le développement fait intervenir les gènes de manière hiérarchisée. Grâce à l'étude de mutations, les caractéristiques du développement ont pu être traduites en termes génétiques. Les événements de développement

(comme la différenciation du système nerveux) reposent sur l'activation ou la répression de gènes particuliers dans les tissus. La notion de hiérarchie fonctionnelle entre événements se traduit donc par une interaction de type *épistatique* entre gènes (voir chapitre 6) : un gène contrôle l'expression de l'autre, et ainsi de suite, en cascade. Dans le cas de l'induction, le gène qui déclenche la production du messager chimique par le mésoderme contrôle l'expression des gènes de différenciation du système nerveux. Depuis les années 1970, les généticiens ont élucidé une partie des interactions génétiques qui sont à la base du développement. Bien que ces connaissances soient encore fragmentaires, elles ont permis d'avancer des hypothèses pour expliquer l'évolution à long terme de caractères phénotypiques fondamentaux. Nous illustrerons ces avancées par un modèle (encore spéculatif) sur l'émergence des grands plans d'organisation dans le règne animal.

Les animaux pluricellulaires se rangent parmi un petit nombre de plans d'organisation apparus pour la plupart rapidement lors de la radiation cambrienne. Les zoologistes regroupent la totalité des animaux en seulement une trentaine d'embranchements, qui correspondent chacun à un plan d'organisation bien défini. Par exemple, les arthropodes (insectes, araignées, crustacés, etc.) se caractérisent par un corps recouvert d'un squelette externe articulé, divisé d'avant en arrière en segments portant des paires d'appendices (pattes, ailes, mandibules). Ils présentent une symétrie bilatérale et leur système circulatoire est ouvert. De nombreux autres caractères, caractérisant l'organisation de tous les arthropodes pourraient ainsi être listés. Il en va de même pour les autres embranchements (vertébrés, mollusques, etc.). De plus, la plupart des fossiles connus, même parmi les

plus anciens, se rangent dans ces embranchements. En effet, de manière surprenante, les plans d'organisation n'ont guère changé au cours de l'évolution : un grand nombre d'entre eux (arthropodes, vertébrés, mollusques, etc.) sont apparus lors d'une période relativement rapide (une ou quelques dizaines de millions d'années), au Cambrien, il y a 530 millions d'années. Depuis, ces plans d'organisation fondamentaux sont restés fixes. Peut-on trouver dans les mécanismes génétiques du développement les raisons de ce rythme particulier de l'évolution morphologique ?

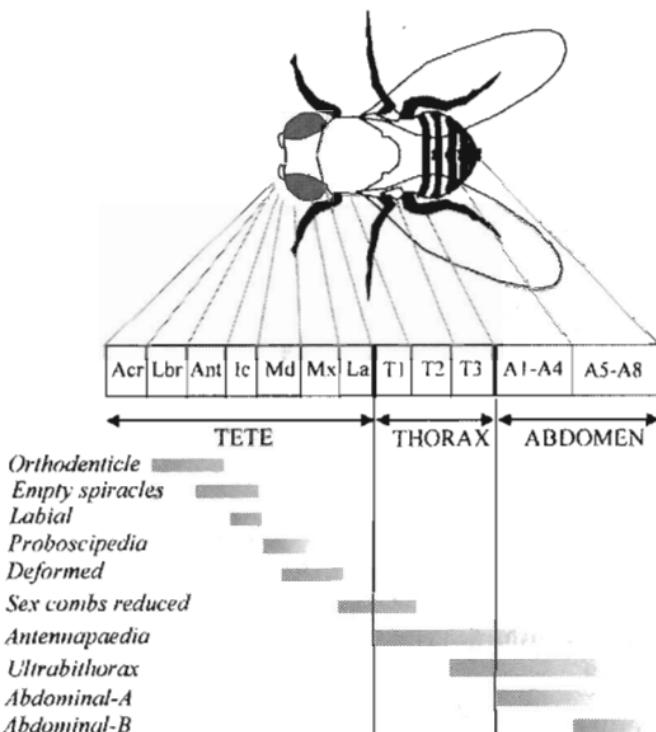
L'examen d'un des organismes descendants de la radiation cambrienne, la *drosophile*, éclaire certains aspects de l'ontogenèse des plans d'organisation. Chez cette espèce, on a obtenu des mutations très spectaculaires, qui provoquaient une modification de l'identité d'un segment. Ces mutations ont été appelées *homéotiques*, car elles rendaient un segment semblable à un autre. Ainsi, le mutant *Antennapedia* montrait des pattes à la place des antennes, signe de la modification d'un segment céphalique (normalement porteur d'antennes) vers le type thoracique (normalement porteur de pattes). De même, les mutants *Bithorax* avaient le troisième segment thoracique (T3) modifié à la façon du deuxième segment (T2) : le balancier (organe d'équilibration normalement présent sur T3) était partiellement ou totalement remplacé par une aile (portée normalement par le T2). L'étude de telles mutations permit d'identifier, dans les années 1970-1980, une famille de gènes (gènes *Hox*) impliqués dans une véritable cartographie du corps de la *drosophile*. À un stade précoce du développement, ces gènes s'expriment dans des positions précises d'avant en arrière, et l'identité d'un segment est déterminée par la combinaison de gènes homéotiques qui s'y expriment

(figure 4). En termes de hiérarchie fonctionnelle, les gènes homéotiques établissent donc un plan d'ensemble dont dépendent tous les gènes impliqués ultérieurement dans les différenciations segmentales. En fait, chaque gène homéotique produit une protéine de type *facteur de transcription*, c'est-à-dire qui modifie l'expression de gènes cibles (plusieurs centaines dans le cas des gènes *Hox*). Par ailleurs, des familles de gènes homologues aux *Hox* de drosophile ont été retrouvées chez d'autres organismes issus de la radiation cambrienne, où ils jouent un rôle similaire. Le nombre de gènes et la position de leurs zones d'expression varient entre les grands embranchements ou classes. Ainsi, les vertébrés (dont l'homme) ont quatre jeux de gènes *Hox*, dont certains s'expriment de manière proximo-distale (de la base à l'extrémité des membres) plutôt que le long de l'axe antéro-postérieur. Cependant, les gènes *Hox* semblent varier très peu à l'intérieur de ces groupes : leurs positions semblent s'être stabilisées au Cambrien, en même temps qu'étaient établis les grands plans d'organisation.

**La radiation cambrienne repose peut-être sur l'acquisition d'un nouveau type de développement.** Davidson et al., (1995) distinguent chez les animaux actuels deux modes de développement<sup>1</sup>. Le premier est représenté par la droso-

Figure 4 : L'expression des gènes homéotiques chez la drosophile. Le plan d'organisation de la drosophile, constituée de segments répartis d'avant en arrière en trois groupes (tête, thorax, abdomen), est déterminé au stade embryon par les zones d'expression des gènes homéotiques (en gris). Les frontières antérieures des zones d'expression correspondent à des

1. E. H. Davidson et al., « Origin of Bilaterian Body Plans : Evolution of Developmental Regulatory Mechanisms », *Science*, 270, 1995, p. 1319-1325.



limites segmentales, et la combinaison de gènes homéotiques exprimée dans un segment lui confère son identité. Ainsi, le segment T2 (porteur d'ailes) est déterminé par l'expression d'*Antennapedia*, et le T3 (porteur de balanciers) par la combinaison *Antennapedia Ultrabithorax*. La suppression, par des mutations, de l'activité *Ultrabithorax* dans le T3, le transforme en T2 (présence d'ailes). Les segments sont en général nommés d'après les appendices qu'ils portent : Acr = Acron, Lbr = Labre, Ant = Antenne, Ic = Intercalaire, Md = Mandibule, Mx = Maxillaire, Lb = Labium, T1 = Prothorax, T2 = Mésothorax, T3 = Métathorax, A1-A8 : segments abdominaux 1 à 8. Remarque : les gènes *orthodenticle* et *empty spiracles* sont les seuls à ne pas faire partie de la famille « *Hox* ».

phile et se caractérise par une définition précoce du plan du corps par l'expression de gènes organisateurs, dont les gènes homéotiques. Avant même de se différencier, chaque tissu est donc caractérisé par une information de position, ou « étiquette génétique », qui détermine son développement futur. Les organismes issus de la radiation cambrienne semblent en majorité suivre ce type de développement. Le deuxième type est celui du ver *Cænorhabditis elegans*, dans lequel les cellules ne se déterminent pas *a posteriori* en fonction de leur position, mais en général dès leur naissance, en fonction d'instructions laissées par la cellule-mère sous forme de messagers chimiques (notons que cette vision est hautement simplifiée, il s'agit d'une tendance générale qui admet des exceptions). Selon Davidson *et al.*, ce mode de développement est limité et ne peut produire de plan d'organisation complexe. En effet, il nécessite de mettre précisément en mémoire la détermination individuelle de chaque cellule, et une cellule aura rarement la liberté de se diviser pour remplacer une autre cellule morte par accident. Ils supposent donc que la diversification cambrienne est liée à l'invention du développement du type *drosophile*, beaucoup plus souple, où les cellules sont déterminées collectivement en fonction de leur position, et peuvent éventuellement se remplacer les unes les autres. Cette invention serait liée à la mise en jeu de gènes organisateurs majeurs, dont les plus représentatifs sont les gènes *Hox*. Chez certains organismes issus de la radiation cambrienne, comme les oursins, il existe un stade larvaire planctonique très petit (0,1 à 1 mm) très différent du stade adulte. Davidson *et al.* interprètent cette situation comme un emboîtement des deux types de développement : le développement de la larve serait de type « *Cænorhabditis* », et celui de l'adulte se ferait suivant le mode « *drosophile* », à partir d'un petit groupe de cellules

indifférenciées « mises de côté » lors du développement larvaire. Ces stades larvaires seraient donc les témoins de l'état ancestral du développement, datant d'avant le Cambrien. Notons que des homologues des gènes *Hox* ont été identifiés chez *Cœnorhabditis*. Cependant, leur rôle dans le développement est méconnu. La radiation cambrienne ne correspondrait donc pas à l'apparition des gènes *Hox*, mais à leur recrutement comme organisateurs du plan du corps.

Selon l'hypothèse de Davidson, au moment de l'invention du développement type « *drosophile* », s'est mise en place une diversité de plans d'organisations en relation avec une évolution des gènes organisateurs et de leurs zones d'expression. Par la suite, des gènes de différenciation se seraient mis sous la dépendance des gènes organisateurs pour créer une variété d'organes dans chaque région du corps. Par exemple, les gènes de différenciation de l'aile se seraient mis sous la dépendance d'*Antennapedia* et *Bithorax* chez les insectes. Or, plus il y a de mécanismes sous la dépendance des gènes organisateurs, plus une mutation affectant l'expression précoce de ces gènes tendra à avoir des effets forts, affectant en cascade toutes les fonctions subordonnées. Ainsi, la plupart des mutants homéotiques de *drosophile* peuvent être considérés comme des individus aberrants, ayant une valeur sélective nulle ou très faible en populations naturelles. À terme, les gènes organisateurs ne peuvent donc plus évoluer, et le plan d'organisation basique est « gelé ». Cela expliquerait la conservation des plans d'organisation apparus au Cambrien, en l'espace de dix ou vingt millions d'années, pendant les quelque cinq cents millions d'années qui ont suivi. Bien sûr, cette hypothèse est encore un modèle spéculatif, qui reste à confirmer par les progrès futurs de la biologie du développement. Cependant, il montre bien que la prise en

compte de la hiérarchie fonctionnelle, caractéristique des gènes du développement, permet d'expliquer les rythmes de l'évolution morphologique observée dans les fossiles.

**Les mécanismes du développement conditionnent la capacité à évoluer des autres caractères.** Les mutations qui affectent des mécanismes basiques du développement peuvent avoir un effet majeur sur l'évolution à long terme. Elles apportent effectivement des innovations fonctionnelles telles qu'elles bouleversent la capacité de diversification évolutive ultérieure des espèces. Ainsi, les mutations qui ont peut-être mis en place au Cambrien un nouveau type de développement (en particulier, le système de gènes organisateurs *Hox*) ont pu permettre une diversification morphologique sans précédent, en formant des plans d'organisations de base sur lequel purent ensuite se greffer des événements de développement apportés par des mutations ultérieures. L'évolution du développement est donc une évolution au second ordre, une évolution de la capacité à évoluer elle-même. Ce type d'évolution pose un autre problème : les gènes *Hox*, par exemple, semblent n'avoir d'intérêt que par rapport aux divers mécanismes qui construisent l'animal sous leur direction, lesquels mécanismes semblent être apparus ultérieurement dans l'évolution. Pourquoi les gènes *Hox* ont-ils été sélectionnés au départ ? Une solution serait de supposer qu'ils avaient initialement une autre fonction que celle pour laquelle ils sont actuellement maintenus. Cette fonction initiale reste encore mystérieuse.

**Peut-on concilier génétique quantitative et génétique du développement ?** Nous avons suggéré au début de ce chapitre que ces deux voies de recherche étaient complémentaires pour l'étude de l'évolution des caractères morphologiques. Mais s'appliquent-elles vraiment aux mêmes caractères ? La variation des traits étudiés en génétique quantitative est régénérée à court terme par

des mutations « courantes » qui maintiennent la variabilité morphologique dans les populations, sur laquelle la sélection naturelle et la dérive génétique peuvent agir. La génétique du développement devrait en théorie permettre de savoir sur quels gènes portent ces mutations et comment ils fonctionnent pour donner des variations morphologiques. Toutefois, les caractères analysés par les biologistes du développement restent généralement très fondamentaux et concernent des échelles de temps et de variation évolutive très grandes : il s'agit par exemple du plan général d'organisation, défini par les gènes *Hox*. C'est logique car, pour commencer à comprendre le développement, il valait mieux s'intéresser d'abord aux mécanismes fondamentaux communs à beaucoup d'organismes, plutôt qu'aux détails subtils propres à une lignée. Les caractères « subtils » sont néanmoins responsables de l'évolution actuelle : ce sont eux qui varient à l'intérieur des populations, et c'est leur évolution que décrit la génétique quantitative. Ce fossé pourra-t-il être comblé ? On peut parier que oui. La technique des QTL donne déjà un moyen d'isoler des gènes particuliers responsables de variations quantitatives à petite échelle. Il reste à disséquer leur architecture fonctionnelle par une approche de génétique du développement, en recherchant des mutants peut-être moins spectaculaires que les mutants homéotiques, mais plus importants pour l'évolution en cours dans les populations.

## Résumé

Les caractères morphologiques reposent rarement sur un seul locus mendélien, et leur évolution ne peut donc pas être analysée par la génétique des populations

classique. La génétique quantitative, inventée au départ pour l'amélioration des espèces domestiquées, permet de modéliser l'évolution à court terme d'un caractère soumis à mutation, sélection et dérive sans connaître son déterminisme mendélien. Son efficacité a été montrée aussi bien en sélection artificielle qu'en situation naturelle. Ses prédictions (sur des caractères) sont qualitativement les mêmes que celles de la génétique des populations (sur des allèles). Les modèles quantitatifs sont en effet théoriquement et empiriquement compatibles avec un comportement mendélien des gènes déterminant le caractère examiné.

La génétique du développement permet de disséquer fonctionnellement les relations gènes-caractères. Les gènes interagissent de manière hiérarchique lors du développement d'un embryon, car les événements de développement qui apparaissent au cours de l'évolution tendent à se mettre sous la dépendance d'événements déjà présents. Ainsi, le plan d'organisation de nombreux métazoaires repose sur un petit nombre de gènes organisateurs. Leur mise en place, au Cambrien, a probablement permis une augmentation considérable du potentiel d'évolution du développement, donnant lieu à une diversification morphologique sans précédent.

Les caractères étudiés jusqu'à présent par les généticiens du développement sont très fondamentaux et anciens dans l'évolution. Cependant, on peut espérer que, dans le futur, cette approche permettra de connaître l'architecture génétique et fonctionnelle des traits qui varient actuellement dans les populations, établissant le lien entre génétique quantitative et génétique du développement.

## QUAND LES RÈGLES DU JEU CHANGENT

Au cours des chapitres précédents, nous avons étudié l'évolution des caractères comme s'ils étaient les seuls déterminants de la valeur sélective d'un individu. Il s'agit d'une approximation simplificatrice. En effet, dès le chapitre 1, nous avions défini la valeur sélective comme une fonction non seulement du génotype de l'individu mais aussi de son environnement. Chaque individu a une existence délimitée dans le temps et dans l'espace. Ce contexte environnemental donne les règles du jeu évolutif, il décide quel individu va perdre (valeur sélective faible) ou gagner (valeur sélective forte). Tout serait simple si le contexte était le même pour tous les individus. Mais il n'en va pas ainsi : l'environnement varie dans l'espace et dans le temps. Ces variations peuvent être prévisibles, comme les saisons au cours d'un cycle annuel. Elles peuvent également être imprévisibles, comme la chute d'un arbre (chablis), qui crée une clairière dans un milieu forestier fermé. Comment s'adapter quand les règles du jeu changent ?

Comment rester adapté dans un environnement qui varie ?

Les individus d'une même espèce peuvent se distribuer sur une surface géographique non uniforme. Ainsi,

les populations humaines se répartissent sur quasiment toute la planète, dans des contrées aux conditions environnementales extrêmement variées. Les pressions de sélection s'exerçant dans un milieu polaire ou dans un milieu tropical ne sont, de façon évidente, pas les mêmes. Comment évoluent les espèces confrontées à de telles disparités spatiales ? Nous illustrerons cette question par plusieurs exemples.

L'adaptation aux variations spatiales du milieu implique une différenciation génétique dans l'espace. Pour illustrer cette idée, notre exemple sera pris dans le milieu marin. Contrairement aux apparences, le milieu marin n'est pas un milieu homogène. Notamment, la salinité est moins forte et plus variable dans les zones marginales (golfes, bras de mer, embouchures) que dans les zones ouvertes (pleine mer). Cette variation fait peser des contraintes fortes sur les organismes vivants. Selon les lois de la diffusion, l'eau tend toujours à se déplacer du compartiment le moins concentré en solutés vers le compartiment le plus concentré. En conséquence, quand l'eau de mer est très salée, les cellules tendent à perdre leur eau (diminution du volume cellulaire), alors qu'elles tendent à se gonfler d'eau quand elles sont plongées dans une eau peu salée.

Une variation génétique impliquée dans la résistance à la salinité a été identifiée chez les moules<sup>1</sup> (*Mytilus edulis*). Elle fut révélée lors d'une étude de la variation protéique par électrophorèse, sur le locus *LAP*. Ce locus code pour une enzyme (leucine amino-peptidase) dont la fonction est de digérer d'autres protéines, libérant ainsi

1. T. J. Hilbish et R. K. Koehn, « Dominance in Physiological Phenotypes and Fitness at an Enzyme Locus », *Science*, 229, 1985, p. 52-54.

des acides aminés dans le milieu intracellulaire. Les allèles de ce locus ont une répartition géographique particulière au niveau du bras de mer de Long Island (près de New York). En effet, la fréquence de l'allèle *LAP94* varie considérablement entre les milieux à faible salinité et les milieux à forte salinité. Atteignant 55 % en zone océanique (très salée), cette fréquence diminue progressivement jusqu'à 10 % vers l'extrémité du bras de mer (eaux moins salées). Ce type de variation géographique progressive, appelé *cline*, suggère que l'allèle *LAP94* est favorable aux fortes salinités, et inversement, défavorable aux faibles salinités. En effet, au fond du bras de mer, la fréquence de *LAP94*, élevée chez les jeunes individus, est de plus en plus faible quand l'âge des moules augmente. Une sélection naturelle forte s'exerce donc sur cet allèle : les moules porteuses de *LAP94* survivent moins bien que les autres en zone peu salée. L'étude physiologique a montré que l'allèle *LAP94* code pour une enzyme plus active que les autres allèles. Donc, la dégradation des protéines et la concentration d'acides aminés libres dans les cellules sont plus élevées chez les moules possédant *LAP94*. De ce fait, cette concentration se rapproche de celle des eaux très salées de la zone océanique. Ainsi, les moules *LAP94* pourraient maintenir leur volume cellulaire plus efficacement dans la zone océanique. En revanche, elles supporteraient moins bien les eaux relativement peu salées du bras de mer. La moule s'adapte donc aux conditions locales, grâce à une différenciation des fréquences alléliques dans l'espace : l'allèle *LAP94* est plus fréquent là où il est le mieux adapté.

Cependant l'adaptation locale n'est pas parfaite. Si tel était le cas, il n'y aurait plus de *LAP94* au fond du bras de mer. Au lieu de cela, à chaque génération, de

jeunes moules porteuses de *LAP94* s'installent dans cette zone, où elles ne sont pas adaptées. Elles sont ensuite impitoyablement éliminées au cours de leur croissance. Quelle est l'origine de cette maladaptation permanente ? Le développement des moules comprend deux phases : l'œuf fécondé se développe d'abord en une larve très petite (< 1 mm) qui vit en pleine eau (*vie planctonique*). Au bout de plusieurs semaines, les larves se métamorphosent en petites moules adultes et adoptent une vie fixée. Les larves sont mobiles : les courants peuvent les entraîner loin de leur lieu de naissance. Malheureusement, elles ne contrôlent guère leur direction : chaque année, des larves parties de la zone océanique (à forte fréquence de *LAP94*) viennent se métamorphoser dans le bras de mer. L'origine de la maladaptation chronique des populations du bras de mer est donc la *migration*, qui dans ce cas est passive. La migration tend à homogénéiser génétiquement les populations, alors que la sélection naturelle tend à les différencier. Les deux forces maintiennent un cline de fréquences. La maladaptation persiste car l'échelle spatiale de la migration est semblable à celle de la variation de salinité.

**L'environnement peut également varier dans le temps.** Pour illustrer cet effet, nous reprendrons l'exemple du moustique *Culex pipiens* et de la résistance aux insecticides dans la région de Montpellier (chapitre 4). Comme nous l'avons vu, les insecticides constituent une pression de sélection qui a conduit à l'installation dans les populations de divers mutants résistants. Cependant, la pression de sélection ne s'exerce que lorsque l'insecticide est présent, en l'occurrence en été, lors de la saison touristique. Pendant l'hiver, la zone littorale n'est pas traitée. Il se passe alors l'inverse de ce

qui a lieu en été : les individus non résistants augmentent en fréquence car, en l'absence d'insecticide, leur valeur sélective est meilleure que celle des individus résistants (coût de la résistance). Les fréquences des moustiques résistants fluctuent donc de manière saisonnière, en suivant le rythme des traitements. Comme pour les moules, la valeur sélective des individus dépend du contexte (présence ou absence d'insecticides). Dans le cas des moustiques, ce contexte varie dans le temps selon un cycle saisonnier prévisible.

**Une dynamique cyclique de l'adaptation locale.** La distribution géographique des moustiques ne se cantonne pas à la zone littorale traitée à l'insecticide durant l'été. L'arrière-pays comporte, en plus faible densité, des gîtes à moustiques. Cette zone, éloignée du littoral, n'est pas traitée : les populations de moustiques et de touristes, toutes deux moins denses, y cohabitent pacifiquement. Les moustiques sont cependant capables de migrer de la zone traitée vers la zone non traitée et réciproquement. Les fréquences de moustiques résistants suivent un cycle qui reflète, à l'instar des moules, les éléments du cycle de vie et les variations, à la fois spatiales et temporelles, de l'environnement.

Les variations de fréquence des allèles de résistance suivent un cycle qui comporte cinq phases (figure 1). Tout d'abord (phase I), pendant l'été, sur une bande de 20 km depuis le littoral, la pression de sélection est forte en raison du traitement à l'insecticide. Il y a donc, dans cette zone, augmentation de la fréquence des résistants. À l'inverse, dans la zone non traitée, la fréquence des individus non résistants baisse car, en l'absence d'insecticide, les individus non résistants survivent mieux (coût de la résistance). Ainsi, le cline de fréquence entre les

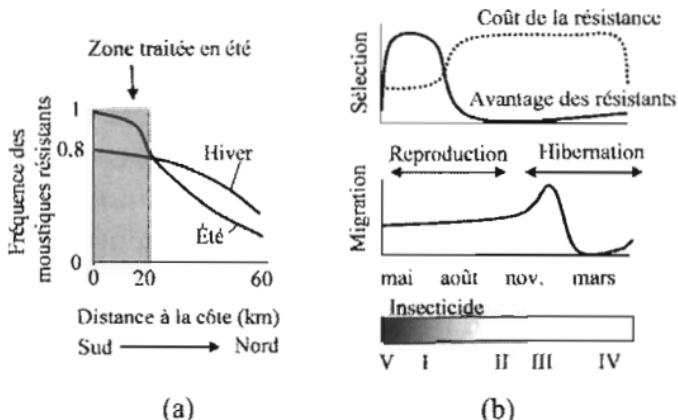


Figure 1 : Les variations cycliques de fréquence des allèles de résistance chez *Culex pipiens*<sup>1</sup>. La variation spatiale des pressions de sélection (moustiques résistants favorisés au sud, moustiques non résistants favorisés au nord) aboutit, comme chez les moules, à la mise en place d'un cline de fréquences des allèles de résistance du littoral (S) vers l'arrière-pays (N). (a) Variations saisonnières du cline. (b) Évolution de la pression de sélection, du coût de la résistance et du taux de migration dans la zone traitée au cours d'une année. Le traitement à l'insecticide couvre, du mois de mai au mois de septembre, une grande partie de la période de reproduction. Durant cette période, l'avantage des résistants est fort et le coût de la résistance faible. La fin de la reproduction a lieu en l'absence de traitement, elle est immédiatement suivie par une période d'intense migration dirigée vers le nord. Enfin, durant l'hibernation, le coût de la résistance est fort et la sélection en faveur des résistants faible. Les cinq phases citées dans le texte sont notées en chiffres romains.

1. T. Lenormand *et al.*, « Tracking the Evolution of Insecticide Resistance in the Mosquito *Culex pipiens* », *Nature*, 400, 1999, p. 861-864.

zones traitée et non traitée s'accentue, malgré les migrations des individus, qui tendent à homogénéiser les fréquences. À la fin de cette phase, le cline se stabilise car les populations sont proches d'un équilibre entre migration et sélection. La seconde phase (II) commence au début de l'automne, alors que le traitement à l'insecticide a pris fin. Elle correspond à la fin de la saison de reproduction : les adultes produisent des descendants qui ne sont pas soumis à la pression de l'insecticide. Le cline commence par conséquent à décroître, car les résistants situés au sud n'ont plus aucun avantage. Le troisième stade (III) correspond à une intense migration vers les sites d'hibernation qui se trouvent au nord de la zone traitée, dans l'arrière-pays (jamais traité). La migration est directionnelle (préférentiellement du sud vers le nord) : les populations du nord (zone non traitée) voient leur fréquence de résistants augmenter par l'arrivée de moustiques résistants venus du sud. Le quatrième stade (IV) correspond à l'hibernation à proprement parler. Durant cette période, les femelles fécondées séjournent dans les grottes et caves de l'arrière-pays, et ne migrent pas. Cependant, le coût de la résistance est fort, ce qui conduit à l'élimination de nombreux individus résistants. La dernière phase (V) correspond au printemps, qui coïncide avec la fin de l'hibernation et le début de la période de traitement à l'insecticide.

Le cycle de variation de la fréquence des variants résistants aux insecticides suit ainsi les variations géographiques et temporelles de l'environnement. Les moustiques n'ont guère le temps de s'installer confortablement dans un optimum adaptatif : ils sont constamment mis en situation de maladaptation, soit parce que l'environnement (ici, le traitement) change, soit parce

qu'ils se déplacent eux-mêmes vers un autre environnement (migration).

## Une autre stratégie face aux variations de l'environnement : la plasticité phénotypique

Les deux exemples qui précèdent (moules et moustiques) ont un point commun : dans une certaine mesure, l'environnement est prédictible. Une fois une larve de moule fixée quelque part, effectivement, la salinité ne varie plus. De même, pour les moustiques, le traitement aux insecticides est saisonnier et géographiquement localisé. Il suffit alors, pour analyser ces situations, de trouver l'échelle géographique et/ou temporelle à laquelle les variations de l'environnement induisent l'action de la sélection naturelle. Que se passe-t-il si les changements ne sont pas prévisibles ? La sélection naturelle ne peut pas favoriser un variant quand les descendants subissent des pressions de sélection différentes de leurs parents. Dans certains cas, l'environnement change à chaque génération. Une stratégie qui peut alors être sélectionnée est celle de la *plasticité phénotypique*, c'est-à-dire la capacité d'un génotype à exprimer des phénotypes différents selon l'environnement où il se développe (figure 2).

Nous reprendrons ici un exemple déjà vu au chapitre 4, celui de l'effet de la présence des crabes *Carcinus mœnas* sur la coquille des littorines (*Littorina obtusata*) dans le golfe du Maine, sur la côte Atlantique des États-Unis. Robin H. Seeley (1986) a montré que la fréquence des formes arrondies est en augmentation depuis le début du siècle. Cette augmentation coïncide avec l'inva-

		Génotype	
		Plastique	Non Plastique
Environnement	A		
	B		

Figure 2 : La plasticité phénotypique. Le génotype plastique acquiert une morphologie différente selon l'environnement. Le génotype non plastique produit le même phénotype quel que soit le milieu.

sion du crabe *Carcinus mœnas*. Grâce à leur forme arrondie les coquilles résistent mieux à l'attaque des crabes. Nous avons vu au chapitre 4 que ce changement de forme ne résulte pas d'une plasticité phénotypique, mais de l'augmentation en fréquence de génotypes qui produisent en toutes circonstances une coquille arrondie.

Le crabe, qui a commencé son invasion par le sud du Cape Cod, s'est ensuite propagé vers le nord. Cependant, en remontant vers le nord, les populations de crabes sont plus petites et moins denses. Une nouvelle étude a été entreprise par Geoffrey C. Trussell et L. David Smith<sup>1</sup>, à la suite de l'observation d'une corrélation entre la

1. G. C. Trussell et L. D. Smith, « Induced Defences in Response to an Invading Crab Predator : an Explanation of Historical and Geographie Phenotypic Change », *Proceedings of the National Academy of Sciences USA*, 97, 2000, p. 2123-2127.

présence des crabes et l'épaisseur des coquilles de littorines. Effectivement, les coquilles sont plus épaisses là où il y a des crabes, contribuant à protéger les littorines de la prédatation. L'épaisseur de la coquille a-t-elle suivi la même évolution que sa forme arrondie ? L'épaississement de la coquille résulterait dans ce cas de l'invasion des populations par des mutants construisant, en toutes circonstances, une coquille épaisse. Cependant, on peut aussi envisager que l'épaississement soit une réaction physiologique induite, chez toutes les littorines, par la présence des crabes (plasticité phénotypique). Dans ce cas, il n'y aurait pas de différences génétiques pour ce caractère entre populations de littorines : en présence du crabe, toutes les littorines alloueraient simplement plus d'énergie à produire une coquille épaisse.

Pour tester ces hypothèses, Trussell et Smith ont étudié la croissance d'individus échantillonnés soit dans le sud soit dans le nord du golfe du Maine. Quatre expériences ont été réalisées (figure 3). Des individus échantillonnés au nord ont été transférés dans le sud et vice versa. Dans chaque cas, les individus ont été marqués et mis dans des filets, ce qui permettait de les suivre tout au long de l'expérience. Chaque filet a été placé dans le milieu naturel, en aval d'un second filet contenant soit uniquement des littorines, soit un crabe et des littorines. Ainsi, les escargots marqués ne pouvaient pas être mangés par le crabe mais pouvaient détecter sa présence par les effluves qu'il dégageait. Le résultat de l'expérience est clair : en présence de crabes, les littorines construisent une coquille plus épaisse quels que soient leur lieu d'origine et leur lieu de transplantation.

Il existe néanmoins une différence entre les littorines placées au nord et au sud. Les littorines placées au nord,

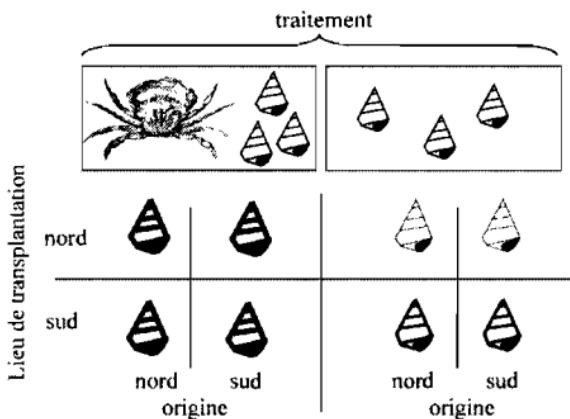


Figure 3 : Protocole expérimental de Trussel et Smith. Des littorines ont été échantillonnées dans un site du nord du golfe du Maine ou dans un site au sud. Les individus échantillonnés dans chaque site ont été réimplantés deux fois dans le site du nord et deux fois dans le site du sud. La première réimplantation a lieu en aval d'un filet contenant uniquement des littorines. Cette série d'expériences permet de tester les effets du lieu de transplantation et du lieu d'origine des littorines, sur leur croissance. La seconde série de réimplantations a lieu en aval d'un filet contenant des crabes et des littorines. Cette série, qui reçoit les effluves du filet contenant des crabes, permet de tester les littorines, de voir si elles sont capables de détecter la présence de crabes et de modifier l'épaisseur de leur coquille en conséquence.

même si elles réagissent à la présence du crabe de la même façon qu'au sud, ont à la base une coquille moins épaisse. Cette différence s'explique en partie par la différence de température de l'eau qui conduit à des différences dans la teneur en carbonate. Il est en effet plus difficile de construire une coquille dans de l'eau froide, pauvre en calcaire, que dans de l'eau chaude, plus riche en calcaire.

La croissance du corps du mollusque a été étudiée parallèlement à l'épaisseur de la coquille. Cette étude a permis de montrer que la croissance du corps suit une variation inverse de celle de l'épaisseur de la coquille. Quand l'individu construit plus de coquille, il ne peut construire en même temps plus de tissus pour son corps. Suivant le contexte (présence ou non de crabes, température de l'eau), l'allocation des ressources varie. Cette différence d'allocation de ressources permet d'optimiser la valeur sélective en fonction du contexte. La valeur sélective a deux composantes : la fécondité et la survie. Quand les crabes sont présents, la probabilité de survie diminue. En épaissant la coquille, la littorine améliore sa probabilité de survie. Quand les crabes sont absents, les littorines produisent une masse corporelle plus importante et améliorent leur fécondité. On sait en effet que chez les mollusques marins la fécondité est directement corrélée avec la masse du corps.

**Quand une espèce peut-elle acquérir une plasticité phénotypique ?** Pour s'adapter à la présence de crabes, la littorine a associé deux stratégies. Tout d'abord, la sélection a favorisé des variants ayant une coquille plus arrondie. D'autre part, la plasticité phénotypique de l'épaisseur de la coquille et de la production de masse corporelle permet, suivant le contexte, d'allouer au mieux les ressources énergétiques en jouant tour à tour sur les deux composantes de la valeur sélective. Pourquoi ces deux adaptations se sont-elles faites sur un mode différent ? L'hypothèse la plus souvent citée pour expliquer la plasticité phénotypique est qu'elle permet à un individu de réagir immédiatement à son environnement. Si l'environnement varie très vite, l'adaptation par la plasticité phénotypique est beaucoup plus efficace qu'un

changement génétique inconditionnel. En effet, les individus, au lieu d'hériter d'une adaptation fixe qui devient vite obsolète, héritent de la possibilité de s'adapter aux conditions (imprévisibles) dans lesquelles ils se trouvent. On devrait donc observer de la plasticité lorsque l'environnement change rapidement de manière imprévisible, et inversement, des variations génétiques non plastiques quand l'environnement change sur le long terme.

Cette théorie, cependant, n'explique pas l'exemple des littorines, puisque la même pression de sélection (les crabes) a conduit à une réponse plastique dans un cas (épaisseur de la coquille) et non plastique dans l'autre (forme de la coquille). Donc, tous les caractères ne sont pas équivalents. La plasticité est une solution possible pour certains, mais pas pour d'autres. Cela suggère qu'il faut prendre en compte les contraintes de développement qui pèsent sur la mise en place de ces caractères. Dans notre cas, des contraintes fortes semblent peser sur l'épaisseur de la coquille. On ne peut effectivement épaisser la coquille qu'en diminuant la masse du corps, donc la fécondité. Cette contrainte pourrait expliquer pourquoi la plasticité a été adoptée plutôt qu'une variation génétique donnant des coquilles inconditionnellement épaisses. En réalité, même dans les régions infestées par le crabe, il existe des sites où il est localement absent, ou des années pendant lesquelles il est moins abondant. Vu le coût des coquilles épaisses en termes de fécondité, les variants capables d'éviter ce coût, en construisant une coquille fine quand il n'y a temporairement pas de crabes, sont favorisés. Les individus qui font systématiquement une coquille épaisse en paient le coût, même quand elle est inutile. On peut imaginer qu'en revanche, le changement de la forme de la coquille n'a pas un coût

très important en termes de fécondité : la plasticité n'apporterait pas vraiment d'avantages pour ce caractère.

### Vivre sous la menace de catastrophes récurrentes

**Les populations vivent souvent sous la menace d'une extinction brusque.** Dans les trois exemples précédents, les changements de l'environnement restent relativement modérés : les individus survivent au changement et les différences de valeurs sélectives entre les individus permettent l'action de la sélection naturelle. Dans certains cas, les changements peuvent prendre l'allure de catastrophes : des phénomènes naturels, mais imprévisibles (*perturbations*), peuvent anéantir instantanément, une population. Les exemples de telles perturbations sont variés. Ainsi, à la suite d'un feu de forêt, les populations aussi bien végétales qu'animales disparaissent, à l'exception de certaines plantes dont les graines ou les parties souterraines sont résistantes. Lors d'une crue, ce sont les populations d'escargots aquatiques qui se trouvent lessivées du lit de la rivière. Parmi un ensemble de populations d'une espèce donnée, ce sont les plus petites qui sont les plus fragiles. Les perturbations d'ampleur faible pourront en effet anéantir une petite population plus facilement qu'une grande. En outre, les populations très petites subissent également des fluctuations stochastiques liées aux aléas de la reproduction des individus qui les composent. Si les populations vivaient isolées, elles s'éteindraient donc les unes après les autres, en commençant par les plus petites.

Le modèle des *métapopulations* suggère qu'un ensemble, constitué de nombreuses petites populations échangeant des migrants, peut se maintenir dans un état

stable malgré les extinctions locales. Ce modèle considère que l'extinction stochastique de petites populations peut être compensée par des capacités de migration et de recolonisation de sites vides. Cette dynamique se place à une échelle supérieure à un site, il s'agit d'une dynamique de *métapopulation* (population de populations). Formellement, quatre critères entrent dans la définition d'une métapopulation. Tout d'abord, l'environnement doit comporter des parcelles discontinues d'habitat favorable, assez grandes pour permettre aux populations locales qui s'y trouvent de se reproduire de façon autonome. Il faut pour cela qu'un nombre suffisant d'individus qui y sont nés restent sur place pour s'y reproduire. Ensuite, aucune des populations de l'ensemble ne doit être assez grande pour assurer sa survie de façon autonome à long terme. D'autre part, les parcelles doivent être suffisamment connectées les unes aux autres pour que les colonisations puissent se faire facilement. Enfin, la dynamique des populations sur les parcelles occupées doit être suffisamment asynchrone pour que les populations ne s'éteignent pas toutes à la fois. Sous certaines conditions, les modèles montrent qu'une métapopulation, dans laquelle toutes les populations locales courrent individuellement le risque de s'éteindre à moyen terme, peut persister à plus long terme.

Le fonctionnement des métapopulations a été remarquablement illustré par l'étude du papillon *Melitaea cinxia*<sup>1</sup>. Cette étude a permis de montrer que les quatre conditions ci-dessus sont effectivement remplies, et qu'elles expliquent la persistance à long terme de ce

1. I. Hanski, M. Kuussaari et M. Nieminen, « Metapopulation and Migration in the Butterfly *Melitaea cinxia* », *Ecology*, 75, 1994, p. 747-762.

papillon en Finlande, alors même qu'il fait partie d'un des groupes d'espèces de papillons les plus menacés d'Europe. *Melitaea cinxia* se trouve en particulier sur l'île de Åland au sud-ouest de la Finlande. Le paysage de cette île est une mosaïque composée de petits champs cultivés, de bois et de prairies. Les parcelles propices au développement du papillon sont des prairies sèches où poussent les deux plantes dont se nourrissent spécifiquement les chenilles : le plantain (*Plantago lanceolata*) et la véronique (*Veronica spicata*). Au cours de l'étude, cinquante parcelles contenant ces plantes ont été explorées, leurs tailles variant entre 12 m<sup>2</sup> et 4,6 ha. Dans un premier temps, la qualité des parcelles (densité de plantes hôtes, exposition au soleil, etc.) ainsi que leur isolement les unes par rapport aux autres ont été évalués. La démographie des populations a été suivie pendant deux années consécutives (1991 et 1992). Malgré ce bref laps de temps, cette étude a permis de mettre en évidence des extinctions dans certaines parcelles et des colonisations dans d'autres. En effet, deux parcelles vides en 1991 étaient occupées en 1992, ce qui atteste deux événements de colonisation. En revanche, cinq sites occupés en 1991 se sont éteints en 1992. Ces données sont en accord avec un modèle de métapopulation où la persistance est due à un équilibre entre extinction et colonisation. Une prédition de ce modèle est que la persistance de la métapopulation n'est possible que si la densité en parcelles favorables est suffisante par rapport aux capacités de dispersion des individus. Dans le cas de *Melitaea cinxia*, il s'avère que sur l'île de Åland, la densité des parcelles, grâce au maintien des pratiques agricoles traditionnelles, est forte. À l'inverse, sur le continent, la densité des parcelles décroît. Cette baisse de la densité des parcelles va

de pair avec un déclin des populations de ce papillon dans la zone continentale.

Les capacités de dispersion des individus ont également été étudiées à l'aide de la technique de capture-marquage-recapture. Pour cela, des individus sont capturés à un endroit, marqués (à l'aide d'un feutre), puis relâchés à l'endroit où ils ont été capturés. Ensuite, on recapture les individus à différents endroits (notamment dans les parcelles adjacentes). Des individus marqués en A et capturés en B se sont déplacés depuis A vers B. Par cette technique, Hanski et son équipe ont pu montrer que 30 % des femelles et 15 % des mâles changeaient de population sur une durée d'une semaine. Les capacités de dispersion des individus sont donc importantes et expliquent le pouvoir de colonisation des sites vides. Dans les parcelles les plus petites, une mauvaise année (sécheresse...) peut conduire facilement à l'extinction des populations locales. La migration permet de compenser ce phénomène : les années suivantes, si la qualité du site redevient bonne, il pourra être réoccupé par de nouveaux immigrants.

Dans une métapopulation, la dispersion peut être un caractère adaptatif. La structure en métapopulation est liée à la fois à un paysage où l'habitat (les ressources) est fragmenté, et à une variation temporelle de la qualité des parcelles. Les modèles théoriques et les données montrent qu'un élément important du maintien d'une telle structure de population est la capacité de dispersion des individus entre les parcelles favorables. Elle a alors un rôle clé dans l'adaptation à ce régime particulier de variation de l'environnement. Nous avons vu, dans les parties précédentes, que la dispersion est une source de maladaptation locale quand l'environnement varie dans l'espace de manière prévisible. Dans le modèle des méta-

populations, où l'environnement varie aléatoirement et brutalement, la dispersion est au contraire un avantage à long terme pour l'espèce. Cependant, en général, les caractères n'évoluent pas pour le bien de l'espèce, mais pour celui des individus. La dispersion peut-elle évoluer par sélection entre individus ?

La première condition d'une telle évolution est l'existence d'une variation génétique de l'aptitude à la dispersion. Les plantes fournissent de nombreux exemples de telles variations. Elles peuvent contrôler le taux de dispersion de leurs graines à travers les structures morphologiques des fruits qui les entourent, notamment des dispositifs de prise au vent comme chez les pissenlits. Certaines plantes, notamment dans la famille des composées, choisissent de ne pas mettre tous leurs œufs dans le même panier. Elles produisent deux types de graines, les unes morphologiquement adaptées à la dispersion (dispositifs de prise au vent, légèreté), les autres destinées à rester près de la plante mère (plus lourdes, sans prise au vent). Le pourcentage des deux types est en général génétiquement variable chez ces espèces. Sachant que la variation génétique existe, dans quelle direction la dispersion devrait-elle évoluer ? Pour le savoir, il faut connaître son effet sur la valeur sélective des individus dans une métapopulation. Cette question est théoriquement et empiriquement complexe, car le taux de dispersion optimal pour un individu résulte d'un compromis entre des pressions antagonistes. Du point de vue de l'individu, la dispersion est une stratégie coûteuse car elle représente une consommation d'énergie et une prise de risque. En effet, on n'est jamais sûr de trouver un milieu aussi favorable que celui d'où l'on vient. D'un autre côté, la dispersion a des avantages : en envoyant ses propagules loin de son site natal, une plante

évitera de les concentrer au même endroit, où elles seraient en compétition les unes avec les autres. En outre, le fait de répartir ses descendants dans plusieurs sites garantit qu'ils ne disparaîtront pas tous d'un coup lors d'une extinction de la population locale. Dans tous les cas, le fonctionnement en métapopulation favorise l'évolution d'un taux de dispersion non nul.

## Résumé

La variabilité spatio-temporelle de l'environnement aboutit à une opposition entre sélection et migration. Si la sélection n'est pas trop forte, elle peut aboutir à un équilibre qui permet de maintenir de la variabilité dans les populations. Cependant, le prix de cet équilibre est une maladaptation chronique. L'adaptation d'une population à son milieu est toujours remise en question par l'arrivée d'immigrants provenant d'un autre milieu.

L'acquisition d'un phénotype optimal dans un milieu donné peut se faire de deux façons : la population peut être envahie soit par un génotype qui produit ce phénotype de manière inconditionnelle, soit par un génotype capable de produire facultativement ce phénotype en réaction à des stimuli environnementaux (plasticité phénotypique). L'évolution favorise la plasticité (i) lorsque l'environnement varie de manière imprévisible de génération en génération, et/ou (ii) lorsque les contraintes du développement rendent cette voie possible, voire plus économique qu'une modification génétique inconditionnelle.

Enfin, quand les ressources sont fragmentées, les populations sont de taille limitée et courent des risques d'extinction suite à des perturbations locales et imprévisibles de l'environnement. La mise en place d'une dyna-

mique de métapopulation, dans laquelle les extinctions dans certains sites sont en permanence compensées par des colonisations provenant d'autres sites, favorise l'évolution de la dispersion. La dispersion est donc adaptative dans un environnement qui varie stochastiquement dans le temps, bien qu'elle soit source de maladaptation dans un environnement constant dans le temps mais variable dans l'espace.

## L'ENVIRONNEMENT, C'EST AVANT TOUT LES AUTRES

Au chapitre précédent, nous avons vu que l'environnement change et remet ainsi souvent en question l'adaptation des populations aux conditions. Les changements évoqués étaient de nature physico-chimique (variations de salinité, présence/absence d'insecticide, perturbations-catastrophes). Dans de tels cas, la situation est relativement simple : ce sont les variations environnementales qui entraînent l'évolution de la population et non l'inverse. Les relations causales sont donc à sens unique. Cependant l'environnement devient beaucoup plus complexe quand on considère sa composante *biologique*. En effet, les conditions de vie d'un organisme dépendent de ses interactions avec d'autres organismes. Ces interactions peuvent être extrêmement complexes, car ces « autres », qui constituent mon environnement, sont eux-mêmes capables d'évoluer. Les relations causales sont alors à double sens. Mon évolution dépend des autres, qui évoluent aussi en fonction de moi. Comment prédire l'issue de ce jeu d'interactions ? Une adaptation réciproque des partenaires est-elle possible, ou doit-on finalement accepter que « l'enfer, c'est les autres » ?

## Comment préserver ses intérêts face aux autres individus de la même espèce ?

Parmi les « autres », il y a avant tout les individus de mon espèce, qui partagent avec moi des liens de parenté et la possibilité d'avoir une descendance commune. De ce fait, nous partageons à la fois des gènes et des ressources. Quelle stratégie puis-je adopter pour tirer mon épingle du jeu ?

**Faire l'original : la sélection fréquence-dépendante.** Les individus d'une même espèce sont généralement en compétition pour les ressources. La sélection naturelle favorise alors les variants génétiques qui n'exploitent pas la même ressource que la majorité de la population. Peu nombreux à se partager cette ressource alternative, ils souffrent moins de la compétition. Les génotypes rares dans la population sont donc avantageés. Dès que leur fréquence devient élevée, ils se retrouvent en situation de compétition les uns avec les autres. On qualifie donc ce type de sélection de *fréquence-dépendante*.

Un exemple simple concerne l'évolution des locus d'auto-incompatibilité chez les plantes. Les gènes de ces locus permettent d'éviter les effets délétères de la reproduction par autofécondation. Le principe général est le suivant : une plante ne peut pas se reproduire avec une plante possédant les mêmes allèles d'auto-incompatibilité qu'elle (notamment elle-même). Ainsi une plante de génotype  $A1A2$  rejette du pollen porteur de l'allèle  $A1$ . Imaginons que l'allèle  $A1$  est très répandu dans une population : les plantes porteuses de cet allèle sont défavorisées car elles ne peuvent se reproduire qu'avec une minorité de la population, celle qui n'a pas l'allèle  $A1$ . Au contraire, une plante possédant un allèle unique (que

ne partage aucune autre plante de la population) peut se reproduire librement avec tout le reste de la population. La reproduction est donc plus facile pour une plante porteuse d'un allèle rare. Chaque allèle étant défavorisé dès lors qu'il est plus fréquent que les autres, on s'attend à ce que tous les allèles soient à l'équilibre en fréquences égales (par exemple, pour quatre allèles : 1/4, 1/4, 1/4, 1/4). Par ailleurs, une nouvelle mutation apparue dans la population est immédiatement favorisée, et augmente en fréquence jusqu'à un nouvel équilibre (dans le cas précédent, cinq allèles avec fréquences 1/5). Le processus peut se réitérer plusieurs fois, augmentant sans cesse le nombre d'allèles. Conformément à ce modèle, on observe en général un grand nombre d'allèles aux locus d'auto-incompatibilité, chacun en fréquence faible.

La sélection fréquence-dépendante est donc un mécanisme fondé sur l'avantage de la rareté, qui conduit à stabiliser la coexistence de plusieurs variants (polymorphisme). Elle s'oppose ainsi à la sélection directionnelle classique et à la dérive génétique, qui tendent à amener l'un des variants à se fixer dans la population.

**Coopérer, mais pas avec n'importe qui.** Les oiseaux et les mammifères poussent souvent un cri d'alarme caractéristique qui prévient leurs conspécifiques de l'arrivée d'un intrus potentiellement dangereux. L'individu qui crie permet aux autres de se protéger mais s'expose à un risque accru, dévoilant sa présence et sa position au prédateur potentiel. Il s'agit donc d'un comportement « *altruiste* », dans le sens où un individu sacrifie un peu de sa probabilité de survie pour augmenter celle des autres. Cela semble contradictoire avec le principe de sélection naturelle : les gènes qui déterminent ces comportements devraient être éliminés de la population.

V. C. Wynne-Edwards<sup>1</sup>, en 1962, tenta d'expliquer ce paradoxe en invoquant un changement du niveau de sélection. Il raisonnait sur un comportement du même type : par exemple la limitation de la taille de ponte, chez les oiseaux. L'intérêt égoïste d'un individu serait de faire le plus de descendants possibles. Cependant, selon Wynne-Edwards, les oiseaux ont des stratégies d'autolimitation de la taille des pontes qui bénéficient au groupe. Ces stratégies permettent d'éviter la surpopulation et donc l'épuisement des ressources qui pourrait conduire à l'extinction. Pour expliquer la mise en place de telles stratégies, il émet l'hypothèse que la sélection naturelle, à l'intérieur d'une population, favorise les individus les plus féconds (« égoïstes »), mais que seules les populations où tous les individus limitent leur propre fécondité (« altruistes ») peuvent survivre. Il s'agit donc d'une *sélection de groupe*. Cependant, les modèles mathématiques ont ultérieurement discrédité cette idée : pour que ce scénario fonctionne, il faut en effet imaginer une extrême sensibilité des groupes à la surpopulation, et une migration entre populations très faible. En effet, dès que les populations égoïstes survivent assez longtemps pour avoir une petite chance d'envoyer des migrants dans des populations altruistes, celles-ci sont envahies par les lignées égoïstes immigrantes.

Ces conditions sont extrêmes et, à ce jour, un seul exemple concret a permis d'illustrer ce modèle : le cas du virus de la myxomatose chez les lapins. Certaines lignées de virus ont des taux de reproduction élevés et sont donc très virulentes. Elles représentent les lignées « égoïstes » du modèle. Ces virus se multiplient très vite,

---

1. V. C. Wynne-Edwards, *Animal Dispersion in Relation to Social Behaviour*, Londres, Oliver et Boyd, 1962.

courant le risque que le lapin meure avant même d'avoir pu transmettre le virus à d'autres lapins par contact sexuel. Un virus se multipliant moins vite (« altruiste ») économise son hôte et a ainsi plus de chances de se transmettre. Si les deux souches se trouvent en compétition dans un même lapin, la population de virus qu'il contient sera vite envahie par la souche virulente. Toutefois, à moyen terme, les populations virulentes s'éteignent avec la mort de leur hôte. L'équilibre est atteint pour une valeur de virulence juste assez basse pour que la transmission ait lieu.

Les lapins introduits en Australie ont permis d'étayer cette prédition. Les populations introduites de lapins causèrent au début du XX<sup>e</sup> siècle des dégâts si considérables que l'on chercha à enrayer cette invasion en introduisant la myxomatose. Cette stratégie réussit au-delà des prévisions : les lapins australiens, n'ayant pas l'expérience du contact avec le virus, y étaient plus sensibles que leurs cousins européens. Ils mouraient vite, souvent avant d'avoir le temps de transmettre le virus. Conformément à l'hypothèse de sélection de groupe, on vit rapidement évoluer les virus de myxomatose vers des taux de multiplication et de virulence plus faibles<sup>1</sup>.

Si cet exemple illustre un cas de sélection de groupe, cette théorie fut cependant rejetée en tant qu'explication générale des comportements coopératifs. Une autre explication possible est le *bénéfice réciproque*. Cette idée se fonde sur le principe « l'union fait la force » : deux individus altruistes qui coopèrent font un bénéfice qui compense et dépasse la dépense d'énergie ou le risque

---

1. F. Fenner, « Biological Control, as Examplified by Smallpox Eradication and Myxomatosis », *Proceedings of the Royal Society of London B*, 218, 1983, p. 259-285.

pris par chacun. Revenons à l'exemple du cri d'alarme. On peut imaginer que le risque pris par l'oiseau, chaque fois qu'il détecte le premier un prédateur, est largement compensé par le fait d'être averti, les autres fois, par d'autres oiseaux. Cependant, cet avantage repose sur l'honnêteté de tous les participants : des oiseaux « tricheurs » pourraient profiter des cris d'alarme poussés par les autres sans jamais en pousser eux-mêmes. Un oiseau « altruiste » fait donc un pari sur le comportement des autres. Sur cette base, John Maynard-Smith<sup>1</sup> a proposé de modéliser ces situations par une approche de *théorie des jeux* (voir p. 324). L'interaction de deux individus est modélisée comme un jeu où le gain obtenu par un individu dépend de ce que joue l'autre. Dans le cas précédent, deux individus gagnent à coopérer réciproquement plutôt qu'à avoir tous les deux un comportement égoïste. Cependant, le meilleur gain est celui d'un individu « tricheur », qui ne prend pas le risque de coopérer, tout en bénéficiant de la coopération de l'autre. Les modèles de théorie des jeux recherchent, sans se préoccuper de la base génétique des comportements, quel est le comportement « évolutivement stable » (ESS pour *evolutionarily stable strategy*) dans la population. Un comportement C est une ESS si aucun autre comportement ne peut envahir une population initialement composée uniquement d'individus C. Pour cela, il faut que le gain du comportement C dans cette population soit meilleur que celui de tout autre comportement. Dans le cas du cri des oiseaux, le comportement coopératif n'est pas une ESS, car s'il apparaît dans une population entièrement coopérative un mutant « égoïste », il aura un gain supérieur aux

1. J. Maynard-Smith, *Evolution and the Theory of Games*, Cambridge, Cambridge University Press, 1982.

autres à chaque rencontre. En revanche, le comportement « égoïste » est une ESS. Dans une population « égoïste », un mutant « altruiste » aidera effectivement tout le monde sans rien recevoir en échange. Son gain sera donc inférieur à celui des autres. Ainsi, l'ESS, qui est l'équilibre atteint par l'évolution, n'est pas la situation la plus intéressante du point de vue de la population. En effet, le gain moyen est supérieur dans une population où tout le monde coopère (non ESS) par rapport à une population d'« égoïstes » (ESS). Malheureusement, nous avons vu que la sélection au niveau de la population ne peut dominer la sélection individuelle que dans des conditions extrêmes.

À la suite de ces résultats, les modélisateurs recherchèrent, dans le cadre de la théorie des jeux, des stratégies plus subtiles qui résisteraient à l'invasion par des mutants « égoïstes ». On conçut ainsi des stratégies mixtes, dans lesquelles un individu coopère ou s'abstient avec une certaine probabilité. En fait, elles se révélèrent n'être que des formes atténuées de coopération : chaque individu prenait moins de risque, mais le résultat général restait le même. L'élément clé n'était donc pas la dose de coopération mais bien l'assurance de réciprocité. En effet, un individu n'a intérêt à coopérer que s'il est raisonnablement sûr qu'on lui rendra la pareille. Finalement, une stratégie de coopération évolutivement stable fut identifiée : *tit for tat* ou, en français, « œil pour œil, dent pour dent ». Cette stratégie impliquait qu'un individu interagisse plusieurs fois avec le même partenaire. Adoptant à la première rencontre un comportement coopératif, l'individu *tit for tat* reproduisait ensuite le comportement de son partenaire lors des rencontres suivantes. Ainsi, si mon partenaire coopère, je coopérerai la prochaine fois, sinon, je m'abstiendrai. Dans une

population de *tit for tat*, un mutant « égoïste » est très défavorisé puisque, dès la deuxième rencontre, personne ne coopère plus avec lui, alors que tous les autres coopèrent entre eux. *Tit for tat* est donc une stratégie évolutivement stable.

Un exemple de stratégie *tit for tat* est donné par les poissons hermaphrodites du genre *Hypolectrus*. Chez ces poissons, les deux partenaires sexuels s'interfécondent en assumant chacun, à tour de rôle, la position de mâle et de femelle et en alternant de nombreuses fois. Cela semble être une complexité inutile par rapport à un seul échange des rôles. La raison est en fait à rechercher dans le coût des gamètes : les spermatozoïdes transmettent le même nombre de gènes tout en étant beaucoup moins coûteux à produire que les ovules. L'individu jouant le rôle de femelle est donc « altruiste » : il paie le prix fort alors que l'autre transmet ses gènes à moindre coût. Pour ne pas être perdant, il doit compter sur la réciprocité. Avec un seul échange des rôles, des « tricheurs », qui joueraient en premier le rôle du mâle puis partiraient sans intervertir les rôles, seraient favorisés. En alternant de nombreuses fois, le problème est évité. De fait, l'accouplement s'arrête dès qu'un des deux individus refuse de recevoir le sperme de l'autre. Comme l'alternance permet de diminuer la quantité de gamètes transférée à chaque tour, le gain du tricheur, qui ne peut pas avoir plus d'un tour d'avance, est faible.

Certains organismes adoptent toutefois un comportement coopératif sans espoir de réciprocité. Un cas extrême est celui des insectes sociaux. Dans une colonie de bourdons, une seule femelle se reproduit (la reine). Pourquoi toutes les autres (les ouvrières) sacrifient-elles leur propre reproduction pour se consacrer à soigner les

descendants de cette reine ? William Hamilton<sup>1</sup>, en 1964, proposa une réponse originale : la *kin-selection* ou *sélection de parentèle*. Le comportement des ouvrières n'est paradoxal qu'en admettant que chaque individu cherche à produire un maximum de descendants. Or l'individu est essentiellement une « machine » destinée à propager des gènes. En général, un gène laisse un maximum de copies de lui-même quand l'individu qui le porte se reproduit beaucoup. C'est pourquoi on raisonne en supposant que l'évolution tend à maximiser la reproduction individuelle. Cependant, pour certains systèmes de reproduction particuliers, comme celui des bourdons, transmission maximale des gènes ne rime pas forcément avec reproduction individuelle. En effet, les bourdons, comme la plupart des hyménoptères sociaux, ont un système de reproduction particulier qui engendre un apparemment très fort entre les descendantes d'une même reine (figure 1). De fait, un gène présent chez une ouvrière a une probabilité  $3/4$  d'être présent chez une de ses sœurs. Un gène de la reine n'a, quant à lui, qu'une probabilité  $1/2$  d'être présent dans chacun de ses descendants. En conséquence, un gène, présent chez un individu femelle, maximisera ses chances de transmission en manipulant le comportement de telle façon que cette femelle participe à la production de ses sœurs plutôt que de produire directement des descendants. Ainsi, du point de vue génétique, les gènes des ouvrières exploitent la reine pour engendrer des copies d'eux-mêmes, et non l'inverse. Ce raisonnement est résumé dans la « règle de Hamilton » : *un comportement coopératif, effectué par un individu acteur au bénéfice d'individus receveurs, se répandra au sein*

---

1. W. H. Hamilton, « The Genetical Evolution of Social Behaviour », *Journal of Theoretical Biology*, 7, 1964, p. 1-52.

*d'une population si et seulement si*  $rb - c > 0$ , où  $r$  représente le nombre de copies d'un gène de l'acteur se trouvant en moyenne dans chaque individu receveur,  $b$  le bénéfice des receveurs, c'est-à-dire l'augmentation moyenne de leur valeur sélective lorsqu'ils bénéficient de la coopération, et  $c$  le coût de la coopération, c'est-à-dire la diminution moyenne de la valeur sélective de l'acteur, suite à son acte coopératif. Cette règle signifie simplement qu'un comportement coopératif évolue si, par ce biais, un gène laisse au total un plus grand nombre de copies. Le terme clé est  $r$ , appelé *coefficient d'apparentement*. En effet,  $r$  est d'autant plus grand que le receveur et l'acteur sont proches parents, d'où le nom de *sélection de parentèle*.

La sélection de parentèle peut expliquer un grand nombre de comportements coopératifs. En effet, ceux-ci sont souvent dirigés vers des individus apparentés. Cependant, la détermination exacte de  $r$ ,  $b$  et  $c$  est complexe. Ainsi, dans le cas du bourdon, les calculs ci-dessus suggèrent que  $r = 3/4$  pour une ouvrière et ses sœurs, contre  $r = 1/2$  pour les descendants directs. Toutefois, ce calcul repose sur l'hypothèse que les ouvrières ont toutes le même père et la même mère. L'analyse des marqueurs moléculaires suggère que c'est effectivement le cas chez *Bombus terrestris* ainsi que chez trois autres espèces de bourdons<sup>1</sup>. En revanche, chez *Bombus hypnorum*, plusieurs mâles fécondent parfois la même reine. Malgré cette difficulté à quantifier précisément les termes de l'équation, la *kin-selection* est actuellement considérée

1. A. Estoup *et al.*, « Monoandry and Polyandry in Bumble Bees (*Hymenoptera Bombinae*) as Evidenced by Highly Variable Microsatellites », *Molecular Ecology*, 4, 1995, p. 89-93.

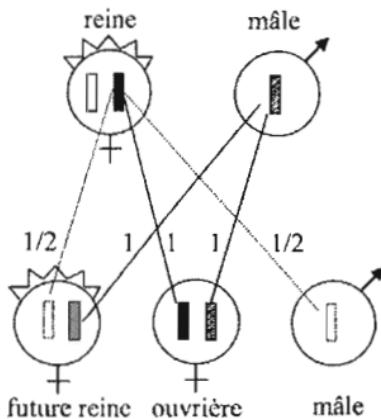


Figure 1 : Les relations de parenté dans une colonie. La reproduction des bourdons est *haplo-diploïde* : la reine produit des œufs fécondés qui donnent des femelles (des ouvrières non sexuées ou des futures reines, sexuées) et des œufs non fécondés qui donnent des mâles haploïdes. Le schéma indique les probabilités de transmission des gènes. Comparons une ouvrière et sa sœur sexuée, c'est-à-dire une future reine sortant de la ruche pour aller fonder une nouvelle colonie. Un gène de l'ouvrière est d'origine paternelle avec une probabilité 1/2 : il a alors obligatoirement été transmis à sa sœur (probabilité de un). Si le gène est d'origine maternelle (probabilité 1/2), il est tiré au hasard parmi les deux gènes homologues de la mère. La mère aura donné le même gène à l'ouvrière et à la future reine avec une probabilité 1/2. Au total, la probabilité qu'un gène de l'ouvrière soit également présent chez sa sœur est donc  $(1/2 \times 1) + (1/2 \times 1/2) = 3/4$ . De même, la probabilité qu'il soit présent chez son frère est de 1/4.

comme un composant majeur de l'évolution des comportements coopératifs.

**Savoir se mettre en valeur.** Beaucoup d'espèces (vertébrés, insectes, etc.) montrent un *dimorphisme sexuel* : les femelles et les mâles ont alors des apparences différentes. Un exemple particulièrement frappant est celui des

oiseaux mâles, qui sont très souvent affublés de caractères très voyants les distinguant des femelles. Par exemple, leur plumage est fortement coloré ou leurs plumes sont hypertrophiées (comme celles de la queue du paon). L'existence de tels caractères est paradoxale : ils rendent les mâles plus vulnérables aux prédateurs et devraient donc diminuer leur valeur sélective. En compilant des données sur de nombreux animaux, Darwin<sup>1</sup> montra que les caractères spectaculaires des mâles avaient pour fonction de leur permettre l'accès aux femelles. Ils se rangeaient parmi deux types. (i) Les *armes de combat*, comme les bois du cerf, utilisés lors de combats entre mâles pour la possession d'un harem de femelles. (ii) Les *ornements*, comme la queue du paon, utilisés lors de parades pour la séduction des femelles. L'évolution d'une arme de combat est relativement simple à comprendre. Les mâles sont en compétition pour une même ressource (les femelles). La sélection priviliege alors les mâles les mieux « armés », plus aptes à évincer leurs concurrents.

En revanche, pour expliquer l'évolution des ornements, il faut supposer que les femelles montrent un comportement de préférence envers les mâles les plus ornementés. L'avantage reproductif de ceux-ci compense alors l'effet négatif de l'ornement sur leur survie. Cette hypothèse a été vérifiée maintes fois. Une expérience célèbre a été menée sur l'oiseau-veuve *Euplectes progne*<sup>2</sup>, un oiseau africain dont les mâles arborent des queues démesurément longues. En découplant et recollant des plumes, la longueur de la queue fut modifiée chez

1. Ch. Darwin, *The Descent of Man, and Selection in Relation to Sex*, Princeton, NJ, Princeton University Press, 1981 (1<sup>re</sup> éd. 1871).

2. M. Andersson, « Female Choice Selects for Extreme Tail Length in a Widowbird », *Nature*, 299, 1982, p. 818-820.

plusieurs mâles. Ceux dont on avait artificiellement allongé la queue obtinrent en moyenne plus de femelles que les témoins, pour lesquels les plumes avaient été coupées et recollées mais sans modification de longueur. Les ornements donnent donc du *sex-appeal*. Darwin avait compris ce phénomène, mais Fisher<sup>1</sup>, cinquante ans plus tard, réalisa que le paradoxe était toujours là. Certes, les mâles s'ornementent pour plaire aux femelles, mais pourquoi les femelles préfèrent-elles les mâles ornem-  
tés ?

Le comportement des femelles semble fantaisiste, car la diversité des préférences est confondante. Les femelles de l'oiseau-veuve, comme celles du paon et des hirondelles de cheminée, préfèrent les mâles pourvus de plumes caudales longues ou colorées. Chez les mouches mâles de l'espèce *Cyrtodiopsis dalmanni*, les yeux sont portés au bout de très longs pédoncules, et les femelles préfèrent s'accoupler avec les mâles aux pédoncules oculaires les plus longs (plus que leur corps !). Les femelles de guppys (petits poissons) préfèrent les mâles pourvus d'une tache latérale rouge. Quant aux oiseaux de paradis de Nouvelle-Guinée, ils font preuve d'une imagination délirante : non contents de se parer de couleurs psychédéliques et de plumes biscornues, les mâles de certaines espèces doivent construire et décorer des édifices hauts de plus d'un mètre (pour un oiseau de dix à vingt centimètres), et se livrer à des danses interminables pour obtenir les faveurs des femelles.

Fisher, considérant que tous ces caractères n'avait aucun point commun, émit un modèle verbal pour expliquer l'évolution de préférences arbitraires chez les

---

1. R. A. Fisher, « The Evolution of Sexual Preference », *Eugenics Review*, 7, 1915, p. 184-192.

femelles. Les prédictions de ce modèle furent ensuite précisées mathématiquement<sup>1</sup>. Il fut baptisé *Fisher's runaway process* (processus de fuite en avant de Fisher). Considérons une population comprenant initialement une faible proportion de mâles ornementés et de femelles préférant ces mâles. La majorité des mâles sont sans ornement et la plupart des femelles indifférentes aux ornements. Dans une telle population, les mâles ornementés ont un avantage. Auprès des femelles indifférentes, majoritaires dans la population, ils ont exactement le même succès que les autres. Leur succès auprès des femelles préférant les mâles ornementés est en revanche supérieur. Ils sont en moyenne plus « sexy » que les autres et fécondent ainsi plus de femelles. Du point de vue des femelles, celles qui préfèrent les ornements s'accouplent plus souvent que les indifférentes avec un mâle ornementé. Leurs fils héritent des gènes d'ornementation de leur père, et auront donc comme eux un succès reproducteur au-dessus de la moyenne. La *sélection sexuelle* favorise donc les femelles ayant une préférence, car leurs fils leur donnent assurément plus de petits-enfants. Le processus s'emballe car la présence de femelles ayant une préférence favorise l'évolution d'une ornementation chez les mâles, et par contrecoup celle de la préférence chez les femelles. Une fois le processus engagé, c'est une fuite en avant (*runaway process*) qui ne s'arrête que lorsque l'ornement devient si handicapant que les mâles perdent plus en survie qu'ils ne gagnent en reproduction. Ce processus peut expliquer l'apparition de traits extravagants, d'apparence arbitraire, car il sélec-

---

1. R. Lande et S. J. Arnold, « Evolution of Mating Preference and Sexual Dimorphism », *Journal of Theoretical Biology*, 117, 1985, p. 651-664.

tionne les phénotypes extrêmes à partir de n'importe quelle petite originalité de départ.

Il existe néanmoins un point commun à tous les ornements des mâles : ils sont tous coûteux et handicapants ! Ainsi, l'efficacité de mâles d'hirondelles pour attraper les insectes dont ils se nourrissent est plus grande lorsqu'on leur raccourcit artificiellement la queue : leur ornement est donc un handicap pour voler. Selon la théorie du *runaway*, ce coût est la conséquence de la course aux phénotypes extrêmes, qui fait évoluer la taille des ornements bien au-delà de sa valeur optimale en termes de survie. Paradoxalement, d'autres théories voient dans ce coût la raison même de la préférence des femelles. Amotz Zahavi<sup>1</sup> a proposé le premier, en 1975, la *théorie du handicap*, selon laquelle l'ornement permet aux femelles d'évaluer la qualité génétique des mâles. Dans cette théorie, l'ornement est si handicapant qu'il met en quelque sorte les mâles à l'épreuve. Ceux ayant de bons gènes sont capables de survivre malgré leur ornement, mais les autres ne peuvent surmonter ce handicap. En choisissant un mâle ornementé, la femelle s'assure donc qu'il donnera de bons gènes à la descendance. Les modèles mathématiques montrèrent toutefois que l'hypothèse, formulée de cette façon, ne fonctionnait pas<sup>2</sup>. Dans ce modèle, effectivement, tout le bénéfice des « bons gènes » est perdu à cause du handicap imposé par l'ornement : les mâles ayant de bons gènes et un ornement ne sont pas meilleurs que les mâles ayant de mauvais gènes et pas

1. A. Zahavi, « Mate Selection – a Selection for a Handicap », *Journal of Theoretical Biology*, 53, 1975, p. 205-214.

2. J. Maynard-Smith, « Sexual Selection and the Handicap Principle », *Journal of Theoretical Biology*, 57, 1976, p. 239-242.

d'ornement. La femelle n'a donc pas vraiment intérêt à choisir un mâle ornementé.

Une version modifiée de l'hypothèse du handicap a alors été proposée, qui, elle, est cohérente avec les modèles mathématiques<sup>1</sup>. Selon cette hypothèse, le développement de l'ornement est physiologiquement contraint : il se développera d'autant mieux que la condition physique du mâle est bonne. Contrairement à l'hypothèse initiale de Zahavi, le coût de l'ornement n'est pas assez fort pour annuler le bénéfice des bons gènes : chaque mâle développe simplement un ornement aussi grand qu'il peut le tolérer, et qui signale honnêtement sa condition physique. Il permet aux femelles d'évaluer cette condition, et par conséquent de se renseigner sur la qualité génétique du mâle ainsi que sur son état de santé. La femelle, en choisissant un partenaire ornementé, assure donc la transmission de bons gènes à ses descendants, et se prévaut contre les maladies contagieuses qu'elle pourrait contracter *via* son partenaire. De nombreuses expériences confirment les prédictions de cette théorie. Ainsi, en manipulant la condition physique des mâles, on peut tester la réponse en termes de développement de l'ornement. Par exemple, chez la mouche *Cyrtodiopsis dalmanni*, les pédoncules oculaires (ornement sexuel des mâles) sont moins longs chez les mâles dont on a diminué la condition physique en réduisant la nourriture<sup>2</sup>. Bien sûr, d'autres organes que les pédoncules oculaires sont également affectés par cette pénurie

1. Y. Iwasa *et al.*, « The Evolution of Costly Mate Preferences. II. the "Handicap" Principle », *Evolution*, 45, 1991, p. 1431-1442.

2. P. David *et al.*, « Male Sexual Ornament Size but not Asymmetry Reflects Condition in Stalk-Eyed Flies », *Proceedings of the Royal Society of London B*, 265, 1998, p. 2211-2216.

de nourriture, mais proportionnellement beaucoup moins.

## Comment gérer ses relations avec d'autres espèces

L'autre est la ressource. Une espèce est toujours une ressource (par exemple un aliment) pour une ou plusieurs autres. Les relations interspécifiques sont classées en fonction des bénéfices ou des coûts pour les deux partenaires. Lorsque chaque partenaire retire un bénéfice, on parle de *mutualisme*. La coopération peut être temporaire. Ainsi, les « oiseaux nettoyeurs » se perchent sur les grands mammifères herbivores (par exemple les gnous) pour manger les insectes et tiques présents sur leur peau. L'oiseau gagne une nourriture et le mammifère se débarrasse de ses parasites. Cependant, le degré de coopération est parfois tel que les deux partenaires restent associés presque toute leur vie. On parle alors de symbiose. Les coraux hébergent ainsi, dans leurs tissus, des organismes unicellulaires (les zooxanthelles) qui sont capables de photosynthèse. Le corail abrite les zooxanthelles qui lui fournissent en échange des métabolites carbonés (c'est le gîte contre le couvert). Dans le cas où deux espèces sont en interaction parce qu'elles convoitent la même ressource, on parle de *compétition*. L'interaction a alors un coût, mais aucun bénéfice pour chacun des deux partenaires. Enfin, l'interaction peut être bénéfique pour l'un et coûteuse pour l'autre. Dans ce cas, une espèce exploite l'autre, soit en consommant des individus (*prédatation*), soit en s'installant durablement sur, ou dans sa victime (*parasitisme*). Les exemples de parasitisme sont innombrables : les espèces pluricellulaires hébergent toujours un cortège de parasites, et l'homme ne fait pas

exception. Ses parasites incluent des virus, des bactéries pathogènes, des protozoaires (comme *Plasmodium*) mais aussi des vers (ténia, douve, trichine, etc.).

Le mutualisme et la compétition sont relativement simples sur le plan évolutif. Les deux espèces ont des intérêts communs : soit d'optimiser les conditions de leur interaction (symbiose), soit d'éviter cette interaction (compétition). Aussi, l'évolution d'un mutualisme est une adaptation réciproque d'une espèce à l'autre, par des mécanismes parfois très élaborés. Dans le cas d'une symbiose, l'adaptation conduit à l'apparition d'une spécificité, c'est-à-dire la spécialisation de différentes lignées de symbiotes vis-à-vis de leurs espèces hôtes. Quand cette spécificité est telle que le symbiose ne peut vivre que sur une espèce d'hôte, la coévolution de l'hôte et du symbiose aboutit à des phylogénies identiques. Ainsi, chaque fois que la lignée des hôtes se sépare en deux, la lignée des symbiotes fait de même. Le cas le plus convaincant est celui des mitochondries et des chloroplastes. Nous avons déjà vu que ces organites dérivent en fait de prokaryotes, entrés il y a très longtemps en symbiose avec des cellules eucaryotes (chapitre 5). Les phylogénies données par le génome original des cellules hôtes (gènes nucléaires) et celles d'origine prokaryote (génome chloroplastique et mitochondrial) sont, à peu de chose près, parfaitement congruentes.

Tout comme la compétition intraspécifique, la compétition entre espèces favorise les variants qui exploitent une ressource originale. La coexistence entre deux espèces facilite leur spécialisation pour des ressources différentes. De fait, la divergence entre les espèces n'est pas remise en cause par les flux de gènes car deux individus d'espèces différentes ne peuvent se croiser. Un exemple classique est celui des escargots marins *Hydrobia ulvae* et

*Hydrobia ventrosa* au Danemark<sup>1</sup>. Chez ces espèces, la taille des particules de nourriture ingérées est directement liée à celle de l'animal. Plus l'animal est gros, plus il ingère de grosses particules et inversement. Lorsque les deux espèces ne sont pas en contact, la taille de leurs individus est semblable, ainsi que celle des particules ingérées. Cependant, dans les sites où coexistent les deux espèces, les individus de *Hydrobia ventrosa* sont plus gros et ingèrent de plus grosses particules que les individus de *Hydrobia ulvae*. Ils évitent ainsi la compétition pour la nourriture. On observe donc un *déplacement de caractère*, c'est-à-dire un changement morphologique lié à la compétition avec une autre espèce.

**Engager la course aux armements.** Quand, dans une relation entre espèces, les deux partenaires n'ont pas les mêmes intérêts (parasitisme et prédation), l'évolution des relations devient complexe. Au sein des espèces parasitées, la sélection naturelle favorise les individus qui mettent en place des défenses efficaces. Chez les vertébrés, il existe ainsi un ensemble de cellules (le système immunitaire) spécialisées dans la détection et l'élimination des pathogènes. Les parasites, de leur côté, seront sélectionnés pour arriver à contourner ces défenses. On observe donc une « course aux armements » entre hôtes et parasites. Dans certains cas, le déterminisme génétique des mécanismes d'attaque et de défense des espèces sont connus. Le cas le plus simple a été observé pour la première fois chez une plante, le lin (*Linum usitatissimum*), et un champignon parasite de cette plante, *Melampsora lini*<sup>2</sup>. Dans ce cas, c'est l'interaction d'un gène du parasite avec

1. T. Fenchel, « Character Displacement and Coexistence in Mud Snails (*Hydrobiidae*) », *Ecologia*, 20, 1975, p. 19-32.

2. H. H. Flor, *Phytopathology*, 45, 1955, p. 680-685.

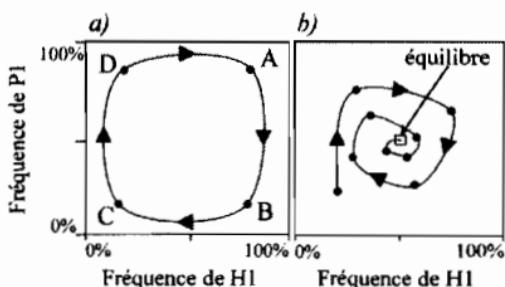


Figure 2 : Évolution des fréquences d'allèles de l'hôte et du parasite dans le modèle de Clarke (1976). Chez l'hôte existe un locus H à deux allèles : H1 et H2 (la fréquence de H2 est 100 % quand celle de H1 est 0 %). Chez le parasite existe également un locus à deux allèles, P1 et P2 (idem). Les hôtes porteurs de l'allèle H1 sont plus résistants aux parasites P1 qu'aux parasites P2, et inversement pour les hôtes H2. *a)* Évolution dans un cas de spécificité forte (H1 résiste beaucoup mieux à P1 qu'à P2). En partant d'une population d'hôtes essentiellement H1 (A), les parasites P1 sont défavorisés et diminuent en fréquence au profit des P2. Lorsque les parasites P2 sont nombreux (B), ce sont les hôtes H2 qui sont favorisés car ils résistent mieux aux P2. Les H2 envahissent alors la population d'hôtes (C), ce qui favorise les parasites P1. Une fois les P1 devenus majoritaires (D), les hôtes H1 sont à nouveau favorisés et le cycle recommence. *b)* Évolution dans un cas de spécificité faible (H1 résiste un peu mieux à P1 qu'à P2). Une dynamique cyclique s'engage mais les oscillations convergent peu à peu vers un équilibre polymorphe où H1, H2 coexistent chez les hôtes et P1, P2 chez les parasites.

un gène de l'hôte qui détermine la réaction de défense de la plante (modèle « gène pour gène »).

Des modèles inspirés de cette situation ont été construits pour simuler l'évolution conjointe d'un hôte et d'un parasite (figure 2). Ainsi, Clarke suppose, pour simplifier, deux espèces haploïdes (hôte et parasite),

possédant chacune un locus à deux allèles. Ce modèle prédit une diversité de dynamiques possibles, mais le départ est toujours une série d'oscillations. Selon les cas, cette dynamique cyclique peut être entretenue sans fin ou s'atténuer peu à peu et converger vers un équilibre où les différentes souches de l'hôte et du parasite coexistent dans leurs populations respectives.

Expérimentalement, ce modèle a été testé en laissant coévoluer dans des tubes à essai des populations de bactéries *Escherichia coli* (hôtes) et des virus capables d'attaquer ces bactéries (parasites). Après des oscillations initiales, caractérisées par le développement de nouvelles souches virales suivies de l'apparition de bactéries résistantes, le système se stabilise finalement en un équilibre comprenant des bactéries particulièrement résistantes et quelques virus survivants. Il est difficile d'observer ce type d'évolution *in natura*, mais l'idée d'une course aux armements peut être testée. Une première prédition est que, si l'hôte et le pathogène sont restés longtemps séparés, leurs relations seront très déséquilibrées (succès foudroyant du pathogène ou échec immédiat). C'est, en effet, quand les deux espèces sont en interaction qu'elles peuvent coévoluer et donc rester en phase dans la course aux armements. Cette prédition est confirmée par de nombreuses données : nous avons déjà vu l'exemple de la myxomatose, à laquelle les lapins australiens, isolés depuis de nombreuses générations, étaient devenus très sensibles. De même, l'introduction de la variole et de la syphilis en Amérique, lors de la conquête espagnole, fit des ravages dans une population amérindienne n'ayant jamais développé de résistance naturelle à ces maladies.

Une deuxième conséquence de la course aux armements est d'accélérer l'évolution des gènes impliqués dans la virulence et la résistance. Cette prédition peut

être testée en évaluant les taux d'évolution moléculaire de ces gènes : leur *horloge moléculaire* (voir chapitre 2) doit être plus rapide que celles des autres gènes. Le virus du SIDA en fournit un exemple spectaculaire. Dans ce virus, le gène *env* code pour une protéine de l'enveloppe externe, partie susceptible d'être reconnue par le système immunitaire de l'hôte, et d'être donc le siège des interactions de virulence/résistance. Dans une région du gène *env*, le taux de substitution a été évalué à  $17,2 \cdot 10^{-3}$  substitutions par site nucléotidique et par an, pour les substitutions synonymes (ne changeant pas la protéine produite) contre  $14,0 \cdot 10^{-3}$  pour les substitutions non synonymes<sup>1</sup>. Ces taux sont exceptionnels à deux titres. Premièrement, d'habitude, les taux de substitutions non synonymes sont cinq à dix fois moins forts que les taux synonymes. Cette règle est vérifiée pour d'autres gènes du même virus. Le comportement original du gène *env* suggère que la sélection favorise les changements d'acides aminés dans la protéine d'enveloppe, plus que dans les autres protéines, conformément aux prédictions. Deuxièmement, ces taux de substitution sont extrêmement élevés : pour comparaison, des valeurs typiques chez les mammifères sont de l'ordre de  $5 \cdot 10^{-9}$  substitutions par site et par an (synonymes) et  $10^{-9}$  substitutions par site et par an (non synonymes), soit environ un million de fois moins que pour le gène *env*. La vitesse d'évolution très rapide de *env* est conforme à l'hypothèse de la course aux armements.

On peut, symétriquement, s'intéresser aux gènes de résistance des hôtes. Les gènes du CMH (complexe

---

I. W. H. Li *et al.*, « Rates and Dates of Divergence Between AIDS Virus Nucleotide Sequences », *Molecular Biology and Evolution*, 5, 1988, p. 313-330.

majeur d'histocompatibilité), chez l'homme, produisent des protéines utilisées par les cellules immunitaires pour détruire les pathogènes. Une cellule infectée présentera à sa surface une molécule typique du pathogène (*antigène*) associée à une protéine du CMH. Cette *présentation de l'antigène* déclenche une réaction spécifique des cellules immunitaires : elles se divisent et produisent des anticorps, ou détruisent la cellule infectée. Dans ces conditions, un pathogène qui produirait des molécules incapables de se complexer avec les protéines CMH déjouerait le système immunitaire. Les protéines du CMH humain montrent une diversité génétique considérable<sup>1</sup> : entre les séquences de deux allèles d'un même locus CMH, existent plus de nucléotides différents (environ 5 %) qu'entre des allèles d'un locus quelconque (typiquement, 0,5 %). Deux arguments supplémentaires suggèrent que c'est la sélection naturelle qui maintient cette diversité élevée. Premièrement, les substitutions non synonymes sont plus nombreuses que les synonymes (14,1 % contre 4,7 %), contrairement à ce que l'on constate pour tous les gènes classiques (voir chapitre 2). Deuxièmement, cette diversité n'est trouvée que dans une petite région des gènes CMH. Cette région correspond précisément au domaine de la protéine qui se lie avec l'antigène (domaine ABS ou *antigen binding site*). Pourquoi, la sélection naturelle maintient-elle une diversité élevée dans cette partie du gène ? On peut supposer que, s'il n'y avait qu'un seul allèle CMH, ou des allèles CMH peu différents les uns des autres, il pourrait

---

1. Ph. Hedrick *et al.*, « Polymorphism and Evolution of the Major Histocompatibility Complex Loci in Mammals », *Evolution at the Molecular Level*, Sunderland, Massachusetts (Selander *et al.* éditeurs), Sinauer Associates, 1991, p. 222-248.

apparaître un pathogène produisant des molécules qui ne se lieraient pas à ces protéines CMH : il échapperait ainsi aux défenses immunitaires et envahirait les populations humaines. La diversification est une stratégie efficace pour se prémunir contre cet échappement : si je possède un CMH très différent de celui de mon voisin, aucun pathogène ne sera capable d'échapper aux deux, et ainsi mon voisin ne pourra pas me transmettre ses maladies.

**Même les mutualismes portent en eux les germes d'une course aux armements.** Contrairement aux parasitismes, les mutualismes sont des associations bénéfiques aux deux espèces impliquées. Pourtant des divergences d'intérêt peuvent à tout moment apparaître entre les partenaires et empoisonner leurs relations, entraînant éventuellement une course aux armements. Nous avons déjà rencontré un tel cas : celui du conflit entre mitochondrie et gènes nucléaires chez le thym (*Thymus vulgaris*, chapitre 5, p. 138-139). Il s'agit bien d'une symbiose obligatoire et ancienne entre un procaryote (devenu la mitochondrie) et un eucaryote (la plante). Pourtant ils n'ont pas les mêmes intérêts, le premier perpétuant ses gènes par la voie maternelle (graines) et le second par les voies maternelle et paternelle (graines et pollen). On observe ainsi une forme de course aux armements entre les gènes de mitochondrie, qui tendent à supprimer la production de pollen, et ceux du noyau, qui tendent à la restaurer.

**Courir pour rester sur place ?** Nous avons vu au chapitre 9 que des événements de changements physico-chimiques du milieu comme les événements géologiques ou climatiques, indépendants de l'action des êtres vivants, pouvaient périodiquement relancer le processus

éolutif en mettant les espèces en situation de maladaptation provisoire. Qu'arriverait-il si de tels événements n'existaient pas : l'évolution s'arrêterait-elle ? Cette question peut sembler purement spéculative. Elle recouvre cependant un problème intéressant : les êtres vivants créent-ils une dynamique qui entretient les changements évolutifs ou ceux-ci sont-ils initiés par des changements externes ? Leigh Van Valen<sup>1</sup> formalisa le premier l'idée que les interactions entre espèces suffisent à entretenir un changement éolutif continu dans les espèces vivantes. Si nous prenons l'exemple d'un hôte et de son parasite, la course aux armements est en effet un processus sans fin : dès que le parasite s'adapte à son hôte, c'est ce dernier qui se trouve en situation de maladaptation, et doit s'adapter à son tour. En raison de leurs intérêts contradictoires, il n'est pas possible que les deux espèces atteignent un optimum adaptatif. En effet, du fait de la course aux armements, les deux espèces ne « s'améliorent » pas, elles cherchent seulement chacune à devancer l'autre. Ce raisonnement a été transcrit par Van Valen, en 1973, dans la célèbre métaphore de la « reine rouge », empruntée à Lewis Carroll (*De l'autre côté du miroir*). La reine rouge explique à Alice que, dans son pays, le paysage est en perpétuel mouvement : « Ici, voyez-vous, il faut courir de toutes ses forces pour rester sur place. » Cette idée a été acceptée avec enthousiasme par la communauté des évolutionnistes, et a été appliquée à une diversité de modèles qui impliquent une variation des pressions de sélection importante et continue. En particulier, elle a été citée comme une des raisons possibles du maintien de la reproduction sexuée (voir chapitre 7).

1. L. Van Valen, « A New Evolutionary Law », *Evolutionary Theory*, 1, 1973, p. 1-30.

Il reste hautement spéculatif de supposer que la « Reine rouge » serait un *pacemaker* nécessaire de l'évolution en l'absence de changements physico-chimiques. La majorité des chercheurs est cependant convaincue de son importance dans l'histoire réelle des êtres vivants.

## Résumé

Les autres organismes, qu'ils soient ou non de la même espèce, représentent une composante importante de l'environnement d'un individu et exercent donc sur lui des pressions de sélection. Au sein d'une espèce, comme entre espèces, les relations peuvent être de nature conflictuelle (compétition, prédation, parasitisme) ou coopérative (comportements « altruistes », symbioses, mutualisme).

Les relations de compétition favorisent une diversification génétique : au niveau intraspécifique, elles favorisent les génotypes rares (sélection fréquence-dépendante) et stabilisent le polymorphisme. Au niveau interspécifique, elles accélèrent la divergence entre espèces.

Les relations coopératives évoluent dans l'intérêt des deux partenaires. En particulier, les comportements altruistes ne sont en fait qu'une apparence : si un individu semble se sacrifier pour en aider un autre ou pour le bien du groupe, c'est parce qu'il transmet mieux ses gènes de cette manière, soit parce qu'il obtient une aide réciproque, soit parce qu'il aide un individu transportant ses propres gènes.

Enfin, les relations entre individus entretiennent la dynamique de l'évolution, en renouvelant en permanence les pressions de sélection. Ainsi, les préférences sexuelles des femelles et les ornements sexuels des mâles

favorisent mutuellement leur évolution, malgré le handicap que représentent ces ornements pour la survie. De même, les parasites et leurs hôtes, voire les partenaires d'un mutualisme, cherchent toujours à s'adapter les uns aux autres dans une interminable course aux armements.



## LA MACRO-ÉVOLUTION, DES INDIVIDUS AUX ESPÈCES

Nous avons abordé l'évolution au niveau des gènes, des individus, des populations, puis nous avons introduit le niveau de l'espèce à partir du chapitre 7. La définition de l'espèce fait l'objet d'une discussion qui reste aujourd'hui encore très animée<sup>1</sup>. Cette discussion a pris de l'importance avec la préoccupation croissante de préserver la biodiversité mondiale (voir p. 316). En effet, l'inventaire de la diversité biologique se fait principalement par le dénombrement des espèces. De même, l'histoire de la vie se raconte comme une chronologie, où se succèdent naissance et extinction d'espèces. Comment naissent et meurent les espèces ? Enfin, certains voient dans les espèces un niveau supplémentaire de l'évolution. Dans ce cadre, les principes de la théorie de l'évolution s'appliquent-ils à l'évolution des espèces au même titre qu'à celle des gènes ou des individus ?

---

1. S. Samadi et A. Barberousse « L'espèce selon Darwin, l'espèce aujourd'hui », *Les Mondes darwiniens* (coordonné par G. Lecointre et M. Silberstein, M), Éditions Syllèphe, 2009, p 155-174.

## Qu'est-ce qu'une espèce ?

Dans la dénomination des objets biologiques, l'espèce a un statut particulier. C'est pour beaucoup l'entité ultime de la classification qui, contrairement aux autres niveaux (genres, familles, classes, embranchements), n'est pas un simple rang supplémentaire dans une série hiérarchique, mais une classe d'objets, définie par une ou des propriétés de nature différente des autres. Cette différence, qui peut paraître subtile, est en fait fondamentale. Les autres niveaux de classification font référence à une hiérarchie des objets classés. Depuis Darwin, on s'attache à ce que cette hiérarchie reflète l'histoire évolutive des organismes vivants, c'est-à-dire la succession des branchements dans l'arbre phylogénétique. Le nombre de niveaux utilisés dépend alors à la fois des connaissances sur les organismes étudiés, de leur nombre ainsi que de l'épaisseur de la tranche d'histoire considérée. Dans le cas de la catégorie « espèce », le débat est ouvert. Certains pensent que l'espèce, à l'instar des autres catégories taxinomiques, n'est qu'un rang qui reflète un niveau de hiérarchie. D'autres pensent que la proposition « appartenir à l'espèce A » indique que les individus de l'ensemble A partagent des propriétés biologiques particulières.

**L'espèce intuitive.** Dans toutes les cultures humaines, les êtres vivants sont nommés. La concordance des noms d'une culture à l'autre, mais aussi avec les dénominations scientifiques, est étonnante. Ainsi, deux cultures qui connaissent les mêmes animaux vont utiliser des noms, propres à chaque langue, pour désigner un même ensemble d'être vivants. Souvent, cet ensemble est également reconnu par les scientifiques et possède un nom d'espèce en latin. Ainsi, « chat », « cat » ou « *Felis catus* »

désigne un même ensemble d'animaux. Les désaccords entre le Français, l'Anglais ou le systématicien qui établiront le classement seront mineurs, tant qu'il s'agit d'organismes qui présentent une valeur d'usage pour l'homme. Les organismes plus éloignés de nos préoccupations quotidiennes, comme les petits insectes du sol, sont soit ignorés, soit désignés dans le langage courant par des noms regroupant beaucoup d'espèces définies par les scientifiques. Toutefois, dans les organismes qui l'intéressent, l'homme reconnaît intuitivement l'existence d'ensembles disjoints. Cette intuition se fonde à la fois sur leur ressemblance physique (on peut reconnaître visuellement les chats des chiens), et sur leur continuité généalogique (un chat fait des chatons, un chien fait des chiots).

**La vision essentialiste.** Selon cette vision chaque espèce correspond à un type qui se définit par un ensemble de caractères correspondant à un organisme idéal. Ainsi, un individu qui aurait tous les caractères du type serait le représentant « idéal » de l'espèce. Cependant, les individus biologiques réels ne sont que des représentants matériels du type. Leur histoire propre dans le monde matériel comporte des « accidents » conduisant à des déviations par rapport au type. Les variations observées entre les individus d'une même espèce résultent dans ce cadre des contingences historiques. Ce type de définitions est hérité de Platon. Il implique la distinction d'un monde d'idées (ici de types) et d'un monde matériel (la réalisation matérielle, nécessairement imparfaite, des idées).

Cette vision de l'espèce est très ancienne. Elle a été mise en pratique à grande échelle par Carl von Linné, au XVIII<sup>e</sup> siècle, dans le cadre d'une vision théologique du monde vivant. Dieu a créé des types (les espèces) dont

les individus biologiques sont l'expression matérielle. Les systématiciens associerent les diverses catégories de la classification, dont les espèces, à un *archétype* : il s'agissait en général d'un dessin représentant un individu idéal, comportant tous les éléments caractéristiques de la catégorie considérée. Cette approche se fonde au départ sur une conception fixiste : les espèces sont des types intemporels qui restent tels qu'ils ont été créés. Les variations observées en leur sein sont purement contingentes, sans rôle pour l'avenir de l'espèce. Paradoxalement, les schémas d'archétypes ont été utilisés dans les manuels de zoologie longtemps après l'abandon des théories fixistes.

**A la recherche d'une définition cohérente avec la théorie de l'évolution.** Darwin a pu élaborer la théorie de l'évolution parce qu'il a compris le rôle essentiel de la variation entre les individus au sein d'une espèce. Ainsi, c'est le tri des individus sur la base de leurs différences qui permet les changements évolutifs. La vision fixiste, qui ne considère les variations intraspécifiques que comme des accidents, n'est pas cohérente avec la théorie de l'évolution et au cours du XX<sup>e</sup> siècle, le débat autour d'une définition cohérente avec le cadre évolutionniste a fait rage. Ce débat a fourni une foison de définitions (recensées par Philippe Lherminier et Michel Solignac<sup>1</sup>). Parmi ces définitions, nous distinguerons trois grandes familles.

**Les définitions typologiques** considèrent que l'espèce est une entité « pratique » qui permet de manipuler des objets biologiques. Ces groupes sont des inventions humaines et n'ont pas une place particulière dans la théorie de l'évolution. Les espèces existent de par la

---

1. Ph. Lherminier et M. Solignac, « L'espèce : définitions d'auteurs », *Comptes rendus de l'Académie des sciences*, 323, 2000, p. 153-165.

nécessité de nommer des ensembles pour les manipuler. Ce travail typologique est l'activité de base de la taxinomie (voir *La systématique*, p. 322).

À ces définitions sont principalement associées des pratiques dites *phénétiques*. Cette méthodologie a été développée par Peter H. A. Sneath et Robert R. Sokal<sup>1</sup>, en 1973. Elle se fonde sur des méthodes statistiques complexes, permettant de regrouper un grand nombre de caractères différents pour établir, le plus objectivement possible, des groupes au sein d'un ensemble d'individus. Plus on utilisera de caractères, meilleur sera le classement. Les caractères utilisés sont divers. Bien qu'ils soient le plus souvent morphologiques, ils peuvent également être moléculaires (séquences d'ADN, de protéines...). La critique majeure que l'on peut apporter à cette méthode est que, suivant le caractère choisi, un même individu peut être classé dans deux espèces différentes. De cette façon, en utilisant le critère de la couleur, les fleurs de pensée et de lilas, de couleur violette, seront classées dans un groupe et celles de couleurs blanches dans un autre. Toutefois, en utilisant le critère de la forme de la fleur, les pensées violettes et blanches seront classées dans le même groupe, de même que les lilas blancs et violets. Les tenants de cette méthode considèrent néanmoins qu'en multipliant le nombre de caractères, on se rapproche toujours d'une classification finale unique.

On considère généralement que la méthode phénétique va de pair avec une conception typologique. Pourtant, certains considèrent que, bien que les espèces ne

1. P. H. A. Sneath et R. R. Sokal, *Numerical Taxonomy – the Principles and Practice of Numerical Classification*, San Francisco, Freeman & C<sup>e</sup>, 1973.

soient pas uniquement des groupes que l'on nomme par commodité, le meilleur moyen de les délimiter reste la méthode phénétique. Les pratiques phénétiques pourraient être des moyens fiables de reconnaissance des espèces définies sur une autre propriété, de même que le nom de famille peut être un moyen assez efficace pour discerner des liens de parenté entre des personnes, même s'il ne définit pas cette parenté. Les questions de la nature de l'espèce et du moyen de la reconnaître sont donc de types différents. Elles sont souvent imbriquées dans les débats sur la définition de l'espèce, les rendant parfois confus.

Les définitions *biologiques* se focalisent sur une propriété biologique. La définition la plus connue est probablement celle d'Ernst Mayr<sup>1</sup>, qui définit l'espèce comme « un groupe de populations naturelles réellement ou potentiellement interfécondes, et reproductivement isolées d'autres groupes semblables ». De la sorte, les individus appartenant à la même espèce sont capables de se reproduire entre eux et uniquement entre eux. Du fait de cette communauté génétique, due aux échanges de la reproduction sexuée, les espèces ont tendance à se différencier génétiquement les unes des autres et à constituer des entités disjointes. D'un point de vue pratique, bien que cette définition propose un critère absolu (*l'interfécondité*), il est difficile de le mettre en œuvre. Effectivement, il est impossible de vérifier de façon exhaustive avec qui chaque individu est, ou non, capable de se reproduire. Plutôt qu'une simple définition, l'espèce biologique propose un programme de recherche. Elle suggère, que les différences observées entre les espèces

---

1. E. Mayr, *Systematics and the Origin of Species*, New York, Columbia University Press, 1942.

« intuitives » sont liées à l'isolement reproducteur entre les espèces.

Un problème majeur, maintes fois soulevé, est l'instantanéité de ce type de définition. Ainsi, il est difficile de trouver un sens, dans le cadre des définitions biologiques, à la question « appartiens-je à la même espèce que Jules César ? ». Or, la naissance de nouvelles espèces est nécessairement un phénomène historique, qui se déroule au cours du temps. Cette idée est difficile à concilier logiquement avec une définition qui n'établit des relations qu'entre individus contemporains. Mayr et d'autres auteurs ont tenté de résoudre cette incohérence, sans parvenir à convaincre toute la communauté scientifique.

Un second problème classiquement soulevé par cette définition, est celui de la reproduction uniparentale : les individus qui ne se reproduisent pas par reproduction strictement biparentale ne peuvent entrer dans le cadre théorique de cette définition. En particulier, si la reproduction biparentale n'est que sporadique, à un instant  $t$  donné, tous les individus qui se reproduisent uniparentalement appartiennent chacun à une espèce différente. À l'instant  $t + 1$ , certains individus qui s'étaient jusqu'alors reproduit uniparentalement se croisent : ils font maintenant partie de la même espèce. Le problème posé par la reproduction uniparentale facultative rejoint ici celui de l'instantanéité de la définition biologique.

Que faire des espèces à reproduction uniparentale ou biparentale facultative ? La solution proposée par beaucoup d'auteurs est *disjonctive*, c'est-à-dire que l'on donne une définition différente de l'espèce pour ces organismes à reproduction uniparentale. Ainsi, la définition biologique fondée sur l'interfécondité s'applique aux organismes à reproduction biparentale obligatoire, mais pas

aux autres. Cette solution n'est toutefois pas satisfaisante car l'espèce désignerait alors des catégories différentes selon les organismes.

Les définitions *phylogénétiques* définissent l'espèce sur la base de la descendance commune. William Hennig<sup>1</sup>, père de la *systématique phylogénétique* actuelle, propose une vision intéressante de l'espèce. La diversité biologique est, en général, décrite sous la forme d'un arbre phylogénétique (p. 284, figures 1a et 1b). Les relations entre les éléments sont alors uniquement divergentes : plusieurs branches de l'arbre peuvent partir d'un même point, mais elles restent ensuite séparées jusqu'au bout. Or, au sein des populations, les individus ne se comportent pas ainsi : leurs lignées de descendants se rejoignent par le biais de la reproduction biparentale. Dès lors, si l'on retrace les relations de descendance entre les individus dans une population, c'est un réseau et non un arbre qui est obtenu. Si l'on regarde ce réseau de loin, on verra cependant qu'il existe des portions de réseau qui forment des groupes fermés. Ces groupes ne sont pas reliés aux autres, ils représentent des entrelacs séparés (figure 1c). À grande échelle, le réseau se comporte donc comme un arbre. Hennig définit l'espèce sur cette base. Les espèces sont les portions du réseau généalogique qui, entre elles, ont des relations divergentes de type « arbre ». Ainsi, les relations entre espèces se représentent sous la forme d'un arbre, alors que les relations entre les individus d'une espèce forment un réseau généalogique (figure 1c).

La définition de Hennig présente une autre difficulté, celle de se fonder uniquement sur le passé. Imaginons

1. W. Hennig, *Phylogenetic Systematics*, Urbana, University of Illinois Press, 1966.

qu'à un instant  $t$ , A et B représentent des parties terminales du réseau, restées disjointes depuis suffisamment longtemps pour que deux espèces au sens phylogénétique soient reconnues (figure 1d). S'il advient qu'au temps  $t + 1$ , des individus issus de A et de B se reproduisent entre eux (hybridation), cette séparation est annulée. A et B, espèces considérées jusque-là différentes, ne le sont plus. Pour pallier ce problème, Kornet<sup>1</sup> ajoute à la définition phylogénétique le concept de *séparation permanente*. A et B sont de bonnes espèces si (i) elles vérifient les critères de Hennig sur la structure du réseau passé, et si (ii) tous leurs descendants futurs possibles restent dans des portions différentes du réseau. Le problème posé par cette définition est simple : comment prédire ce que vont faire les descendants futurs ?

Peut-on réconcilier dans une même définition théorique les deux aspects de la théorie de l'évolution pris en compte par ces deux familles de définitions ? Suivant l'échelle de temps, le réseau global de l'histoire de la vie passée a un aspect d'arbre ou de réseau. Les définitions biologiques s'intéressent à l'échelle « réseau » et se fondent sur une partie de la théorie de l'évolution : la *transmission des gènes par la reproduction* entre les individus. Les définitions phylogénétiques s'intéressent quant à elles à l'échelle « arbre » et se fondent sur un autre versant de la théorie : la *descendance commune*. D'un point de vue pratique, les méthodes proposées sont radicalement différentes : les critères sont comportementaux et physiologiques pour les uns, phylogénétiques pour les autres.

---

1. D. J. Kornet, « Permanent Splits as Speciation Events : a Formal Reconstruction of the Internodal Species Concept », *Journal of Theoretical Biology*, 164, 1993, p. 407-435.

Nous pensons pourtant que les critères biologiques sont complémentaires aux critères phylogénétiques. Comme nous l'avons vu, ce qui manque à la définition phylogénétique c'est une garantie de la permanence des événements de séparation entre des portions du réseau généalogique. Deux types de garanties peuvent être apportés. Le premier, trivial, est l'extinction. Une espèce éteinte n'aura plus de descendants : ceux-ci ne pourront alors pas se reproduire avec des descendants d'autres branches. Les espèces éteintes sont donc reconnaissables sans ambiguïté. Un deuxième critère, plus intéressant, est précisément le critère de Mayr, c'est-à-dire un obstacle biologique à la reproduction entre espèces. Pour que la barrière reproductive détectée entre A et B soit une garantie de séparation permanente, il faut qu'elle se maintienne entre les descendants de A et ceux de B. De fait, quand une telle barrière se met en place, elle ne peut plus disparaître. En effet, une barrière biologique à la reproduction repose sur une divergence génétique portant sur les caractères reproductifs. Dès lors, une fois A et B reproductivement isolés, leurs descendants hériteront de ces caractères. En l'absence d'échanges entre A et B, les caractères reproductifs continueront à diverger au même titre que tous les autres caractères génétiques. Par exemple, il est certain qu'aucun descendant futur de l'espèce humaine ne se reproduira avec un descendant futur de l'espèce *Drosophila melanogaster*. La situation n'est pas toujours aussi évidente car, dans le cas d'organismes proches, l'étanchéité de la barrière est plus difficile à établir. Toutefois, ces cas plus épineux ne remettent pas en question la définition de l'espèce : ils posent juste un problème pratique de reconnaissance. L'évaluation des critères biologiques (isolement reproducteur actuel) donne, sinon des certitudes, du moins des

arguments motivés pour reconnaître les divergences permanentes entre portions du réseau généalogique.

En conséquence, les espèces peuvent être définies comme des portions disjointes de manière permanente du reste du réseau généalogique. Nous pensons que le seul mécanisme pertinent qui assure la permanence de cette structure, et donc sa forme d'arbre au réseau généalogique du vivant, est l'acquisition de caractères génétiques qui empêchent la reproduction entre individus appartenant à différentes branches de l'arbre. Les limites temporelles de l'espèce sont données par la date à laquelle elle acquiert ces caractères génétiques, et par la date à laquelle elle s'éteint ou donne naissance à de nouvelles espèces.

Dans la pratique l'identification de telles limites est difficile ; on peut considérer les noms scientifiques d'espèces comme des hypothèses provisoires sur ces limites, susceptibles d'être rejetées ou affinées en fonction de l'arrivée de nouvelles données. La méthodologie de délimitation repose sur des critères reliés aux trois familles de définitions. Ces réflexions récentes, initiées par Kevin de Queiroz en 1998<sup>1</sup>, ont clarifié le débat et permis d'identifier de nouvelles méthodes de délimitations d'espèces qui s'appuient sur la multiplication des données génétiques, les progrès méthodologiques de la génétique des populations et de la reconstruction phylogénétique, ainsi que sur la puissance des programmes de calculs<sup>2</sup>.

Cette intégration des deux approches permet de résoudre les problèmes évoqués plus haut. Tout d'abord,

1. K. de Queiroz, « The General Lineage Concept of Species, Species Criteria, and the Process of Speciation : a Conceptual unification and Terminological Recommendationst », in D. J. Howard & S. H. Berlocher (Ed.), *Endless forms. Species and speciation*, Oxford University Press, Oxford, 1998, p. 57-75.

2. Cf. note 1 p. 266

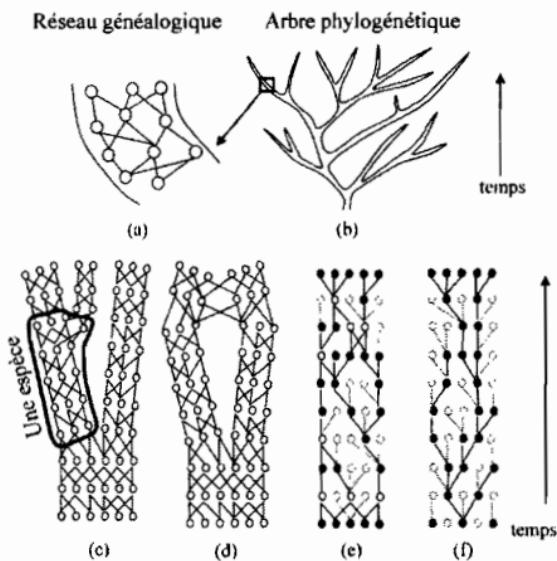


Figure 1 : L'arbre et le réseau. Haut : schémas montrant l'arbre phylogénétique du monde vivant et les relations généalogiques entre les individus. (a) Les individus sont représentés par des ronds, les liens entre parents et descendants sont représentés par les traits. (b) Chaque branche de l'arbre phylogénétique est en fait une portion du réseau généalogique. Bas : différentes structures de réseau généalogique. (c) Spéciation. Le réseau se sépare définitivement en deux portions divergentes. (d) Hybridation. La séparation que l'on croyait définitive se révèle temporaire. (e) Reproduction biparentale facultative. Les descendants de reproduction biparentale sont en blanc, ceux de reproduction uniparentale en noir. Les lignées éteintes sont en pointillé. Les événements occasionnels de reproduction biparentale suffisent à obtenir des relations réticulées entre individus. (f) Reproduction uniparentale obligatoire. Les relations entre individus sont uniquement divergentes. Chaque individu fonde une lignée qui reste séparée des autres.

le problème de l'instantanéité de l'espèce biologique disparaît car l'épaisseur temporelle, permise par l'approche phylogénétique, a été introduite. La question de la reproduction clonale est également résolue, notamment pour les espèces à sexe facultatif. Du moment que des reproductions biparentales existent de temps en temps dans le réseau, on reste dans la même espèce car cela signifie que les individus n'ont pas acquis de caractères assurant leur isolement reproducteur (figure 1e). En revanche, dans un cas de clonalité stricte (figure 1f), chaque lignée clonale est une espèce. Les lignées purement clonales ont effectivement entre elles des relations divergentes et, à défaut d'exhiber un mécanisme de cohésion entre elles, autre que la reproduction sexuée, nous n'avons pas de raison de les regrouper dans une même espèce. En outre, la clonalité stricte est extrêmement rare dans le monde vivant, contrairement à la sexualité facultative.

Une objection supplémentaire est régulièrement faite aux approches fondées uniquement sur les critères d'isolement reproducteur. Souvent deux entités, traditionnellement reconnues comme des espèces différentes, s'hybrident. Leur différenciation n'était pas suffisante pour assurer une séparation permanente. Le statut d'espèce donné à ces entités était donc abusif. Ces entités étaient en fait des *sous-espèces*, c'est-à-dire des populations différencierées de la même espèce. Cette confusion s'explique par le caractère progressif de l'acquisition de mécanismes génétiques d'isolement reproducteur. Elle illustre également le statut d'hypothèse scientifique des espèces nommées : l'acquisition de nouvelles données, ici l'observation d'un croisement entre deux groupes d'individus que l'on croyait isolés sur la base d'autres critères, permet de rejeter l'hypothèse initiale de deux espèces et d'en proposer une nouvelle. L'hybridation n'est donc pas

un problème qui remet en cause la définition de l'espèce d'un point de vue théorique mais qui illustre bien la difficulté pratique d'identifier les limites entre espèces.

**Est-il possible de décrire les millions d'espèces estimées de la biodiversité actuelle ?** Il y a une trentaine d'années, la diversité biologique était considérée comme majoritairement connue. Cette vision a radicalement changé et actuellement cette diversité est évaluée avec un à deux ordres de grandeur supplémentaires. Le problème du recensement exhaustif de la biodiversité ne se pose donc pas dans les mêmes termes. Le développement des méthodes de séquençage de l'ADN a accru nos capacités à acquérir des données génétiques, mais aussi à gérer ces données ainsi que les autres données biologiques (morphologiques, écologiques, etc.) à l'aide de bases de données informatisées. De plus, le développement des moyens de communication permet de partager ces données. L'ensemble mis bout à bout permet d'envisager l'accélération de la connaissance de la diversité des espèces. C'est dans ce contexte que se développe actuellement le projet « Barcode of Life » lancé en 2003 par Paul Hebert et ses collaborateurs<sup>1</sup>. Ce projet vise à lier les données taxonomiques classiques aux données génétiques et à rendre l'ensemble accessible, via Internet, par des bases de données publiques.

### La naissance de nouvelles espèces

Nous avons vu qu'une propriété fondamentale du monde vivant est d'être divisé en espèces très nombreuses (voir *La biodiversité et sa préservation*, p. 316), ayant une

1. P. D. Hebert *et al.*, « Biological Identifications through DNA Bar-codes », *Proceedings of the Royal Society of London B, Biological Sciences*, 270, p. 313-321, 2003.

origine unique. La théorie de l'évolution comprend trois principes (voir chapitre 1) : la reproduction imparfaite (avec mutation), un tri aléatoire (dérive génétique) et un tri déterministe (sélection naturelle). Nous avons vu jusqu'ici que ces principes expliquent beaucoup de caractéristiques du monde vivant. Peuvent-ils également expliquer l'existence des espèces ou doit-on ajouter celle-ci comme un axiome indépendant ? Pour le savoir, il faut rechercher si la naissance des espèces ou *spéciation* découle ou non de ces principes.

Les espèces sont caractérisées par un isolement reproducteur définitif entre les individus d'espèces différentes. Fondamentalement, le mécanisme pertinent qui assure la pérennité des séparations est une divergence génétique. Un scénario de spéciation doit donc avoir comme propriété minimale de prédire l'apparition de différences génétiques suffisantes entre deux groupes d'individus issus de la même espèce pour qu'ils ne puissent plus se reproduire entre eux (ou, ce qui est équivalent, donner des descendants viables). Ces caractères génétiques sont appelés *mécanismes d'isolement reproducteur*.

Le scénario canonique de la spéciation implique un isolement géographique qui entraîne un isolement génétique. Ce scénario de spéciation, dit *allopatrique*, se déroule en plusieurs phases. À l'origine, l'espèce ancêtre occupe une surface continue dans laquelle les individus circulent, et sont tous capables de se reproduire entre eux. La deuxième phase suppose un événement de fragmentation de cette population en deux ou plusieurs populations isolées (*allopatriques*). Pendant cette phase, les migrations entre populations sont absentes ou très rares. À ce stade, les individus de deux populations A et B ne se reproduisent pas entre eux car ils ne se rencontrent jamais. Cependant, si on les transplantait d'une

population à l'autre durant cette phase, ils seraient physiologiquement capables de le faire. Lors de la troisième phase, des variants génétiques différents se fixent progressivement à l'intérieur de chacune des populations. L'isolement géographique est crucial pour cette divergence, car les migrations d'individus homogénéiseraient ces deux populations en transportant les gènes de l'une vers l'autre. Les différences génétiques qui s'accumulent entre les deux espèces portent sur une variété de caractères différents, dont ceux qui déterminent la compatibilité reproductive entre individus. Au cours de la quatrième phase, la différence entre les populations est telle que les individus d'origine A et B ne peuvent plus se reproduire entre eux, ou ne peuvent plus avoir ensemble de descendants viables. A et B ont donc acquis un isolement reproducteur. Dès lors, A et B sont deux espèces : même si elles sont remises en contact, les individus originaires de A et de B formeront deux branches séparées dans le réseau généalogique.

Ce scénario peut être testé dans la nature de deux façons : premièrement, la répartition des espèces nouvellement formées doit refléter l'isolement géographique de leurs populations d'origine. Deuxièmement, le scénario prédit l'existence de formes inachevées de spéciation, c'est-à-dire de populations allopatриques qui ont commencé à diverger génétiquement mais qui, si elles sont remises en contact, ne sont pas complètement isolées reproductivement.

**La répartition des espèces récemment formées apporte de nombreuses confirmations au modèle allopatrique.** Les cas les plus typiques sont trouvés dans les archipels. En effet, la mer isole de nombreuses îles, formant un obstacle difficile à franchir pour des organismes terrestres. Ainsi, les populations se trouvent fragmentées

et isolées. Elles ont par conséquent l'occasion de diverger génétiquement. On s'attend *a priori* à observer dans les archipels des espèces différentes de celle des continents et, dans chaque île, une espèce différente.

Dans le cas des martins-pêcheurs du genre *Tanysiptera* en Nouvelle-Guinée, Ernst Mayr rapporte la présence sur l'île principale d'une seule espèce (*Tanysiptera galatea*), divisée en trois populations différencierées (ou sous-espèces) occupant des parties différentes de l'île. Ces trois sous-espèces sont très difficiles à distinguer morphologiquement. À l'inverse, quatre petites îles périphériques abritent chacune une espèce différente de *Tanysiptera* morphologiquement très différenciée, conformément au scénario de spéciation allopatrique. Pourtant, une cinquième espèce semble contredire ce scénario. Elle se trouve en partie sur une île périphérique, et en partie au sud de l'île principale (Nouvelle-Guinée). Les données géologiques suggèrent que ces deux territoires faisaient autrefois partie d'une même île qui s'est ensuite fragmentée. Une partie de cette île s'est raccordée au sud de la Nouvelle-Guinée et la seconde a formé l'île périphérique. La cinquième espèce se serait donc différenciée quand ces deux territoires étaient réunis. Finalement, la relation une île/une espèce est vérifiée pour les cinq espèces.

Cette relation est cependant loin d'être générale. Dans les îles Hawaii, par exemple, les criquets et les drosophiles ont eu des taux de spéciation exceptionnels, donnant naissance à des centaines d'espèces. Dans chaque île, on peut trouver plusieurs espèces de criquets et plusieurs espèces de drosophiles. Si les espèces se forment à la suite de la séparation des populations dans différentes îles, on s'attend à trouver les espèces nouvellement formées dans des îles différentes, et à retrouver les espèces les plus proches phylogénétiquement (espèces sœurs) sur

des îles distinctes. Cette prédition semble vérifiée pour les drosophiles, mais pas pour les criquets. Par exemple, parmi les criquets du genre *Laupala*, seules les divisions phylogénétiques les plus anciennes semblent être associées à des événements de colonisation d'îles. À l'intérieur de chaque île, ces branches anciennes se sont récemment divisées en trois voire six espèces<sup>1</sup>. À première vue, cet exemple semble aller à l'encontre du modèle allopatrique. Néanmoins, la mer n'est pas le seul élément qui fragmente les populations. À l'échelle d'un criquet, par exemple, chaque île est une mosaïque de milieux isolés entre eux, notamment par le relief. Il semble donc que les taux de spéciation élevés observés dans les archipels ne sont pas uniquement dus à la séparation des îles par la mer.

L'**histoire évolutive des souris (*Mus musculus*) illustre les premiers stades de la spéciation allopatrique**. L'étude de la variation morphologique et des marqueurs moléculaires chez *Mus musculus* a permis de déterminer quatre groupes différenciés qui occupent des zones séparées dans l'ancien monde<sup>2</sup> (figure 2). Ces groupes ont le statut de sous-espèce car l'isolement reproducteur entre les groupes n'est pas complet. Les sous-espèces de *Mus musculus* permettent d'illustrer la première phase de la spéciation allopatrique : l'acquisition de différences génétiques en allopatrie.

Les données moléculaires révèlent trois caractéristiques. (i) Les quatre groupes ont une origine commune

1. D. Otte, « Speciation in Hawaiian Crickets », *Speciation and its Consequences*, édité par D. Otte et J. A. Endler, Sunderland, Massachusetts, Sinauer Associates, 1989, p. 482-526.

2. P. Boursot *et al.*, « The Evolution of House Mice », *Annual Review of Ecology and Systematics*, 24, 1993, p. 119-152.

récente : les comparaisons de séquences d'ADN mitochondrial, grâce à l'horloge moléculaire, permettent d'estimer l'âge du dernier ancêtre commun à moins d'un million d'années. (ii) La phylogénie des quatre sous-espèces forme une étoile, c'est-à-dire que les branches correspondantes semblent toutes partir d'un même point. Cette forme de phylogénie suggère que les quatre groupes se sont séparés les uns des autres à la même époque sur un laps de temps court. (iii) La sous-espèce indo-afghane *M. m. bactrianus* n'est pas génétiquement homogène. Elle contient des génotypes très variés, si bien que, génétiquement, les trois autres sous-espèces semblent être des branches particulières émergeant d'un ensemble de souris artificiellement regroupées sous le nom *bactrianus*. En particulier, les souris iraniennes semblent, conformément à leur position géographique, être génétiquement intermédiaires entre les souris indiennes et la sous-espèce *domesticus*.

À partir de ces données, et des restes fossiles (dents) retrouvés dans l'ancien monde, l'histoire des souris peut être reconstituée avec un bon degré de confiance. L'origine des souris se trouve dans la région indienne. Il y a moins d'un million d'années, plusieurs vagues de colonisation sont parties en même temps de cette région en prenant des directions différentes. Une fois séparées, les populations ont génétiquement divergé jusqu'à former les sous-espèces actuelles : *musculus*, *domesticus* et *castaneus*. Les populations restées sur place, correspondant au groupe *bactrianus*, ont également divergé. Cependant, elles gardent une plus grande partie des caractères ancestraux que les autres sous-espèces. Les souris illustrent donc bien le premier stade de la spéciation allopatrique : l'acquisition de différences génétiques en allopatrie.



Figure 2 : Répartition géographique des quatre sous-espèces de souris domestique *Mus musculus*.

Peut-on observer également, comme le prédit le modèle, l'acquisition d'un isolement reproducteur ? Il faut pour cela s'intéresser aux zones dans lesquelles des souris de sous-espèces différentes se rencontrent. En effet, les sous-espèces initialement isolées ont ensuite étendu leur aire de répartition et sont entrées en contact en plusieurs endroits : Europe (*musculus-domesticus*), Chine (*domesticus-castaneus*). La zone la mieux étudiée se situe en Europe où *domesticus* et *musculus* se sont rencontrées très récemment (3 000 à 6 000 ans). *M. m. domesticus* arrivait d'Afrique du Nord et *musculus* d'Europe de l'Est. Plusieurs arguments indiquent l'existence d'un isolement reproducteur partiel. Premièrement, des croisements en laboratoire montrent une baisse de valeur sélective des descendants de croisements entre les deux groupes, par rapport aux descendants de croisements à l'intérieur des groupes. Deuxièmement, la zone de

contact, qui traverse toute l'Europe du nord au sud, est très étroite (au maximum, quarante kilomètres). Les gènes typiques de *musculus* n'ont guère pénétré dans le territoire *domesticus* et réciproquement. Les modèles expliquent cette absence de mélange par un désavantage sélectif des génomes mixtes combinant des gènes *musculus* et *domesticus*. Cependant, ce désavantage est difficile à mettre en évidence dans la nature. Enfin, les femelles *musculus* montrent souvent un comportement de préférence pour les mâles de leur propre sous-espèce. Toutefois, dans la zone de contact, les individus ont des génomes mixtes. Pour produire ces génomes, il a fallu non seulement que des hybrides naissent, mais encore qu'ils aient plusieurs générations de descendants. La divergence génétique de populations qui ont été séparées géographiquement est donc bien accompagnée d'un isolement reproducteur comme le prédit le modèle de la spéciation allopatrique. Dans le cas de la souris, la spéciation n'est pas achevée car la séparation des deux groupes n'a pas été assez longue avant la remise en contact pour que l'isolement reproducteur soit complet.

**Comment se met en place l'isolement reproducteur ?** L'exemple précédent montre que l'isolement reproducteur entre deux populations A et B a deux composantes. La première, dite *isolement postzygotique*, est caractérisée par une valeur sélective réduite (voire nulle) des descendants hybrides entre A et B. A et B peuvent donc s'accoupler et se féconder, mais la descendance est peu viable. La deuxième, dite *isolement prézygotique*, est un ensemble de mécanismes qui conduisent à éviter les croisements entre A et B. Les deux reposent sur des différences génétiques que A et B ont accumulées. Or, la théorie de l'évolution prévoit deux mécanismes permettant une telle différenciation : la sélection et la dérive

génétique. Quel est le rôle de ces deux mécanismes dans l'acquisition de l'isolement reproducteur ?

De nombreuses données suggèrent que l'isolement postzygotique est souvent dû à des interactions épistatiques entre gènes (voir chapitre 6). Pour illustrer cette idée, imaginons deux locus (1 et 2), pour lesquels chaque population est fixée pour un allèle différent ( $A_1$  et  $A_2$  chez A,  $B_1$  et  $B_2$  chez B). Les gènes de type  $A_1$  et  $A_2$  fonctionnent bien ensemble, de même que les gènes  $B_1$  et  $B_2$ . Un individu hybride porte des gènes  $A_1$  et  $B_2$  (ou  $B_1$  et  $A_2$ ), qui ne fonctionnent pas bien ensemble. La généralisation de ce raisonnement aboutit au concept de *coadaptation génomique* : à l'intérieur d'une population, les allèles de tous les locus du génome ont été sélectionnés pour fonctionner efficacement ensemble. Chez un hybride sont réunis des allèles qui n'ont pas suivi cette histoire de coadaptation réciproque. La faible valeur sélective des hybrides s'explique donc par une mauvaise coadaptation des gènes. Les différences écologiques entre espèces fournissent un exemple simple de ce raisonnement. De fait, de nombreux cas de spéciation allopatrique vont de pair avec des divergences écologiques. Ainsi, parmi deux espèces sœurs de drosophile (*Drosophila simulans* et *Drosophila sechellia*), la première exploite une large gamme de fruits, alors que la seconde n'exploite que les fruits d'un arbre particulier, le morinda. Or, pour exploiter un fruit, il faut à la fois des gènes qui déterminent un comportement de choix de ce fruit, et des gènes qui permettent d'en digérer les composants. Les gènes du comportement et ceux de la digestion sont coadaptés au sein d'une espèce, car ils permettent d'exploiter le même fruit. Par conséquent, un hybride risque de choisir un fruit A alors qu'il ne saura que digérer les fruits B, ce qui est clairement désavantageux.

La sélection naturelle est donc impliquée dans l'isolement postzygotique : elle maintient la cohésion interne des ensembles de gènes dans des populations isolées. S'il apparaît une différence entre les deux populations pour un gène, la sélection tendra à la généraliser à d'autres gènes. Cependant, l'apparition d'une première différence peut se faire aussi bien par dérive génétique, que par sélection. La proportion relative de ces deux phénomènes est discutée. Elle dépend des différences de pressions de sélection rencontrées par les groupes qui sont isolés. Ainsi, les forts taux de spéciation observés dans les archipels peuvent être expliqués de deux manières. Selon les uns, ils sont dus à la fréquence des événements de colonisation d'îles par un ou quelques migrants. Chacun de ces événements serait alors à l'origine d'une nouvelle espèce. Le faible nombre de migrants favoriserait la dérive, mettant en place une divergence initiale rapide. Selon les autres, ils sont dus à la variété des niches écologiques disponibles sur les îles. Il y a peu d'espèces sur chaque île et les ressources ne sont souvent pas toutes exploitées. Dans ce cas, une espèce arrivant sur l'île se séparerait en autant de groupes qu'il y a de types de ressources disponibles (*niches écologiques*). Une divergence entre ces groupes conduirait à une *radiation adaptative*, c'est-à-dire à la formation rapide de nombreuses espèces présentant des adaptations variées. Bien que les chercheurs ne soient pas d'accord sur l'importance relative de chacun des facteurs, ils interviennent probablement tous deux. L'évolution de l'isolement postzygotique dépend donc à la fois de la dérive et de la sélection. Toutefois, en aucun cas il n'est sélectionné pour lui-même : il serait absurde que la sélection naturelle favorise des individus qui ont des descendants inviables ! L'isolement postzygotique apparaît comme sous-produit d'une

divergence génétique, qu'elle soit due à la sélection pour d'autres caractères (écologiques...), ou à la dérive.

L'isolement prézygotique peut apparaître suivant le même processus. Ainsi, les signaux de reconnaissance entre partenaires sexuels peuvent diverger initialement par dérive ou par adaptation à des conditions différentes. La sélection à l'intérieur de chaque population maintient la coadaptation entre les systèmes d'émission de signaux et les systèmes de réception et interprétation de ces signaux, ce qui accentue les divergences entre populations. Les préférences écologiques jouent également un rôle dans l'isolement prézygotique, car elles déterminent les probabilités de rencontre entre partenaires. De fait, les mâles et les femelles de *Drosophila sechellia* se rencontrent sur le fruit qui constitue leur ressource principale (le morinda). Cependant, contrairement à l'isolement postzygotique, l'isolement prézygotique peut être sélectionné directement. Cela peut avoir lieu lorsqu'il existe déjà un isolement postzygotique. En effet, lorsque les hybrides entre deux sous-espèces sont peu viables, un individu qui choisit de se reproduire en priorité avec des partenaires de la même sous-espèce est favorisé. Ce phénomène, appelé *renforcement*, a une conséquence testable : considérons deux entités (sous-espèces ou espèces) dont les hybrides sont défavorisés (isolement postzygotique). Quand les deux entités sont restées géographiquement séparées, le renforcement ne peut avoir eu lieu : l'isolement prézygotique devrait donc être moins fort que quand elles sont en contact. Cette prédiction a été superbement confirmée par une étude sur un grand nombre de paires d'espèces du genre *Drosophila*<sup>1</sup>. Ces auteurs ont montré que pour les paires

1. J. A. Coyne et H. A. Orr, « Patterns of Speciation in *Drosophila* », *Evolution*, 43, 1989, p. 362-381.

d'espèces en situation d'allopatrie, l'isolement prézygotique apparaît progressivement : il est élevé pour des espèces séparées depuis longtemps mais faible pour des espèces séparées récemment. Au contraire, les espèces au contact l'une de l'autre acquièrent très rapidement un isolement prézygotique fort. Cette différence n'est pas retrouvée pour l'isolement postzygotique, qui augmente progressivement avec le temps de divergence, qu'on soit ou non en allopatrie. Cela suggère qu'en allopatrie, les deux types d'isolement évoluent progressivement, en fonction d'une divergence due à la dérive génétique ou à la sélection sur d'autres caractères. Inversement, quand les populations sont en contact, l'isolement prézygotique est renforcé par une sélection directe.

**Existe-t-il d'autres modes de spéciation que la spéciation allopatrique ?** La plupart des chercheurs estiment que la spéciation allopatrique explique une grande partie, voire l'essentiel, des naissances d'espèces. D'innombrables modèles ont néanmoins suggéré que d'autres modes de spéciation étaient possibles. Ces modèles ne supposent pas de phase d'isolement géographique : l'isolement reproducteur apparaît soit entre populations adjacentes échangeant des migrants à travers leurs frontières (*spéciation parapatrique*), soit entre deux groupes d'individus occupant le même territoire (*spéciation sympatrique*). Dans les deux cas, le problème est de trouver des mécanismes suffisamment forts pour que l'isolement reproducteur se développe sans être remis en cause par les flux de gènes. Si la spéciation parapatrique peut être considérée comme une forme atténuée de spéciation allopatrique, la spéciation sympatrique est, elle, beaucoup plus difficile à expliquer. Les modèles montrent

cependant qu'elle est possible<sup>1</sup>. En effet, s'il existe une pression de sélection forte favorisant la spécialisation de deux catégories d'individus dans des ressources différentes A et B, les individus intermédiaires entre ces deux catégories seront désavantagés, n'étant compétitifs ni pour exploiter A ni pour exploiter B. Dans ce cas, il est intéressant pour un spécialiste de A de se croiser uniquement avec d'autres individus spécialistes de A : ils éviteront ainsi de produire des descendants mal adaptés. Cela conduit à l'évolution d'un isolement prézygotique. Une fois acquis, cet isolement permet une divergence des deux catégories pour beaucoup d'autres caractères. La pression de sélection favorisant les individus spécialistes doit cependant être très forte pour contrecarrer les flux géniques entre les deux catégories d'individus, et il est difficile de dire si de telles conditions existent dans la nature. Certains exemples le suggèrent. Chez les plantes, on observe souvent la persistance de deux types morphologiques en sympatrie malgré un isolement reproducteur imparfait. Les exemples de spéciation sympatrique restent pour le moment moins nombreux que les exemples de spéciation allopatrique.

Les modes de spéciation envisagés jusqu'ici proposent des solutions différentes à un même problème théorique : celui de l'apparition progressive d'un isolement reproducteur entre deux groupes d'individus. Nous avons vu que la solution la plus simple était l'allopatrie, mais que la sélection en faveur des spécialistes, en sympatrie, était une autre solution envisageable. Cependant, le problème disparaît si un isolement total ou presque total

---

1. U. Dieckmann et M. Doebeli, « On the Origin of Species by Sympatric Speciation », *Nature*, 400, 1999, p. 354-357.

est obtenu instantanément, en une seule mutation. Imaginons qu'une telle mutation apparaisse dans la lignée germinale d'un individu. Ses descendants héritent de la mutation et ne peuvent plus se croiser avec le reste de la population, ou alors les descendants de ces croisements sont inviables. Si ces mutants parviennent à se reproduire entre eux et si leurs descendants survivent, ils fonderont une nouvelle espèce. Les mutations chromosomiques donnent un exemple de telles spéciations rapides. Par exemple<sup>1</sup>, dans les populations de la sous-espèce de souris *Mus musculus domesticus*, on trouve plusieurs races chromosomiques. En effet, deux chromosomes de paires différentes peuvent fusionner entre eux (*fusions robertsonniennes*). L'équipement chromosomique (ou *caryotype*) peut donc varier selon les types de fusions intervenues. Les hybrides entre individus de caryotypes différents sont souvent partiellement ou totalement stériles, car lors de la méiose, les lots chromosomiques issus des deux parents, configurés de manière différente, ne ségrègent pas correctement. Les souris *Mus musculus domesticus* ont colonisé la petite île de Madère il y a seulement cinq cents ans, avec l'arrivée des Portugais. Il n'en a pas fallu plus pour que six races chromosomiques distinctes s'y établissent. Leurs différences chromosomiques semblent suffisantes pour induire une stérilité des hybrides. Dans cet exemple, l'allopatrie entre des populations situées dans des vallées isolées a, sans doute, favorisé l'accumulation progressive de fusions différentes dans chacune. En effet, une différence d'une seule fusion ne suffit pas à établir une stérilité totale des hybrides. Cependant, la rapidité du phénomène est clairement due au fait qu'un

1. Britton-Davidian *et al.*, « Rapid Chromosomal Evolution in Island Mice », *Nature*, 403, 2000, p. 158.

petit nombre de mutations chromosomiques suffit à acquérir cet isolement reproducteur.

En conclusion, l'apparition de limites spécifiques entre des groupes d'individus, par l'évolution d'un isolement reproducteur, est une conséquence des principes de la théorie de l'évolution : mutation, sélection et dérive. Elle nécessite cependant des événements contingents qui déterminent le lieu et le moment des spéciations, par la présence de barrières géographiques (spéciation allopatrique) ou l'arrivée aléatoire de mutations chromosomiques. À la contingence près, l'existence des espèces ne nécessite pas de principe supplémentaire à la théorie.

### La macro-évolution, une évolution à l'échelle de plusieurs espèces

L'apparition de nouvelles espèces et leur extinction doivent être intégrées dans notre description de l'évolution. Nous avons exploré, dans les chapitres précédents, les mécanismes par lesquels les changements évolutifs ont lieu à l'intérieur des espèces. Que se passe-t-il quand on change d'échelle et qu'on examine l'évolution sur un pas de temps à l'intérieur duquel des espèces peuvent naître et mourir ? À cette échelle, notre principale source d'information est l'évolution morphologique observée chez les fossiles.

**La morphologie évolue-t-elle de manière continue ?** En 1972, deux paléontologues américains, Niles Eldredge et Stephen J. Gould<sup>1</sup>, proposèrent la *théorie des*

1. N. Eldredge et S. J. Gould, « Punctuated Equilibria : an Alternative to Phyletic Gradualism », *Models in Paleobiology*, édité par T. J. M. Schopf, San Francisco, Freeman & Cie, 1972, p. 82-115.

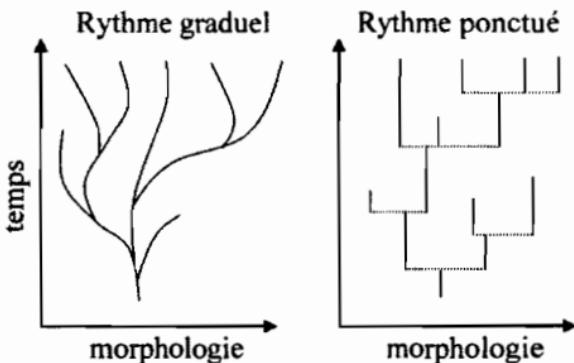


Figure 3 : Images gradualistes et ponctualistes de l'évolution morphologique. Dans le cas d'un rythme ponctué, on observe une alternance de stases (traits pleins) et de variations brusques correspondant à des spéciations (traits pointillés).

*équilibres ponctués*, dont l'élément central est une hypothèse sur le *tempo* de l'évolution morphologique. Selon eux, les changements morphologiques observés dans les fossiles sont en général discontinus. Ainsi, une lignée reste longtemps stable morphologiquement (*stase*), et tous les changements ont lieu lors d'épisodes très brefs de diversification morphologique (*ponctuations*), au cours desquels apparaissent de nouveaux types morphologiques. Très rapidement émergent donc de nouvelles lignées statiques. Cette image s'oppose au concept, selon eux classique, de *gradualisme phylétique*, selon lequel la morphologie des lignées change progressivement, sans faire de sauts (figure 3).

À partir de la constatation d'une image ponctuée de la variation morphologique dans les fossiles, ils proposent une hypothèse sur le *mode d'évolution* de la morphologie. Selon eux, la stase s'explique par l'existence de fortes contraintes de développement empêchant tout changement général de la morphologie à l'intérieur d'une espèce. Ces contraintes sont dues à l'interaction entre les

allèles majoritaires au sein d'une population. Ces allèles déterminent collectivement une homéostasie de développement, c'est-à-dire une résistance aux changements induits par de nouvelles mutations. Un changement ne peut se produire que si (i) cette homéostasie est rompue, soit par l'arrivée d'une mutation à effets très forts, soit par l'apparition par hasard d'une combinaison d'allèles minoritaires. (ii) Les génotypes déviants ainsi produits sont vite isolés génétiquement du reste de la population, et fondent une nouvelle espèce. Protégée, cette espèce acquiert dès lors une homéostasie propre qui stabilise sa nouvelle morphologie. L'originalité de ce scénario est que la variation morphologique est exclusivement associée à des événements de spéciation. En conséquence, la sélection naturelle ne peut agir sur la morphologie qu'au niveau interspécifique : les espèces ayant la morphologie la plus favorable seront favorisées (taux d'extinction faibles et taux de spéciation forts) par rapport aux autres. Cela s'oppose aux concepts traditionnels, selon lesquels la morphologie peut évoluer par sélection des individus adaptés à l'intérieur d'une espèce.

**Les paléontologues ont livré une bataille d'exemples mais ceux-ci soulèvent des difficultés d'interprétation.** Le registre fossile montre le plus souvent une variation discontinue d'une couche sédimentaire à l'autre. Cependant, la plupart des exemples ne peuvent être cités à l'appui de la théorie des équilibres ponctués. En effet, le dépôt des couches géologiques est un phénomène en général très discontinu : les variations brusques de morphologie peuvent être dues à des interruptions longues de la sédimentation, conduisant à supprimer des formes intermédiaires dans une série à variation continue. Les paléontologues cherchent donc des séries exceptionnelles dans lesquelles

l'enregistrement sédimentaire était le plus complet possible pour échapper à ce biais. Néanmoins, les exemples cités à l'appui d'une hypothèse étaient toujours critiqués sur des bases techniques par les partisans de l'autre hypothèse. Certains d'entre eux semblent cependant être dignes de confiance. Ainsi, dans plusieurs lignées de trilobites (arthropodes fossiles de l'ère primaire), on observe un changement graduel du nombre moyen de côtes sur le pygidium qui est la partie postérieure de l'animal<sup>1</sup>. À chaque niveau stratigraphique, les trilobites d'une même lignée ne sont pas tous identiques : leur nombre de côtes se répartit selon une distribution centrée sur une moyenne. Cette moyenne augmente progressivement d'un niveau à l'autre sur une durée de trois millions d'années. Ce même schéma se répète pour huit lignées distinctes. Cet exemple illustre un rythme graduel, d'autres soutiennent néanmoins l'idée d'un rythme ponctué. Un autre exemple provient de la série sédimentaire du lac Turkana en Afrique (4,5 millions d'années) qui présente une continuité et une résolution temporelle exceptionnelle. Dans cette série, Peter G. Williamson<sup>2</sup>, en 1981, a suivi l'évolution de plusieurs lignées de mollusques. Leur morphologie était généralement stable, mais les périodes de stabilité étaient entrecoupées de périodes de variations rapides, globalement synchrones pour toutes les lignées. Ces périodes voyaient l'émergence instantanée de nouveaux types morphologiques, suivies en général d'un retour rapide au type initial. Parfois, le nouveau type avait une persistance

1. P. R. Sheldon, « Parallel Gradualistic Evolution of Ordovician Trilobites », *Nature*, 330, 1987, p. 561-563.

2. P. G. Williamson, « Paleontological Documentation of Speciation in Caenozoic Molluscs from the Turkana Basin », *Nature*, 293, 1981, p. 214-215.

plus longue. D'après ces exemples, l'évolution morphologique semble tantôt graduelle, tantôt ponctuée. La généralité de la description ponctualiste de l'évolution est donc encore débattue. Le cas de la lignée de rongeurs du genre *Stephanomys* au tertiaire permet de relativiser encore plus cette description : dans ce groupe, dont les restes fossiles sont constitués principalement de dents, la taille des dents évolue de manière continue alors que leur forme semble évoluer par sauts<sup>1</sup>. L'image obtenue peut donc dépendre du caractère étudié.

Les données biologiques peuvent servir à évaluer les interprétations des séries paléontologiques. La série du lac Turkana et d'autres séries fossiles semblent soutenir l'hypothèse ponctualiste sur le rythme de l'évolution. Cependant, cela n'est pas une garantie de la validité de l'hypothèse ponctualiste sur les mécanismes de la variation. Peter G. Williamson interprète l'apparition brusque de variants morphologiques, dans la série du Turkana, par une série de spéciations rapides, chaque morphotype étant considéré comme une espèce distincte. Malheureusement, on ne peut pas vérifier directement si les critères de définition de l'espèce s'appliquent sur les fossiles. Cependant, lorsque l'on s'intéresse à des organismes récents, on peut évaluer la vraisemblance des hypothèses à l'aide de données sur les animaux actuels. C'est le cas pour la série du lac Turkana. Parmi les lignées étudiées dans cette série, l'escargot *Melanoides tuberculata* possède des représentants dans les eaux douces actuelles. Ces escargots ont un mode de reproduction principalement parthénogénétique et sont actuellement présents dans

1. S. Renaud *et al.*, « Fourier Analysis Applied to Stephanomys (Rodentia, Muridae) Molars : Non-Progressive Evolutionary Pattern in a Gradual Lineage », *Paleobiology*, 22, 1996, p. 255-265.

l'ensemble de la zone intertropicale. L'étude de la variation morphologique et génétique chez ces escargots a révélé plusieurs caractères surprenants<sup>1</sup>. (i) Le groupe est divisé en types morphologiques (ou *morphes*) bien distincts, qui correspondent chacun à une lignée parthé-nogénétique séparée. (ii) Ces morphes peuvent exister seuls dans un site, ou coexister temporairement, voire se remplacer les uns les autres sur un même site par des invasions successives. (iii) La forme de la coquille peut varier entre escargots génétiquement identiques placés dans des milieux différents (plasticité phénotypique). (iv) La parthénogenèse n'est pas le seul mode de reproduction, il apparaît de temps en temps des mâles ayant une coquille de forme distincte de celle des femelles. (v) Des événements de reproduction sexuée biparentale ont épisodiquement lieu entre morphes différents : les descendants fondent alors une nouvelle lignée parthéno-génétique caractérisée par une morphologie nouvelle. Cela indique que tous les morphes de *Melanoides tuberculata* doivent être considérés comme des variétés d'une seule et même espèce, reliés entre eux par des liens de parenté et de reproduction. Les observations (i) à (v) suggèrent au moins trois explications autres que la spéciation pour expliquer les variations brusques observées dans la série du Turkana. Premièrement, il peut s'agir d'un effet de la plasticité suite à un changement rapide du milieu. Les variations coïncident effectivement avec des changements du niveau d'eau du lac. Deuxièmement, ils peuvent être dus à l'apparition sporadique de

1. S. Samadi *et al.*, « Variation of Shell Shape in the Clonal Freshwater Snail *Melanoides tuberculata*, and its Consequences for the Interpretation Fossil Series », *Evolution*, 54, 2000, p. 492-502.

mâles, morphologiquement différents des femelles. Troisièmement, ils peuvent être dus au remplacement d'un morphé par un autre dans la localité étudiée ou à un épisode de reproduction sexuée engendrant un nouveau morphé. L'étude des animaux actuels permet donc d'émettre de sérieux doutes sur l'interprétation d'un des cas d'école cités par les ponctualistes.

**La sélection entre espèces est-elle incompatible avec la sélection entre individus ?** Nous avons vu que l'observation de rythmes ponctués ne prouve pas le modèle de sélection des ponctualistes. Ces derniers ont opposé la sélection entre espèces à la sélection entre individus, comme s'il s'agissait de deux mécanismes incompatibles. Selon eux, la sélection entre espèces peut se définir comme la préservation des caractères conférant aux espèces la capacité d'engendrer beaucoup d'espèces filles (spéciation), et/ou réduisant leurs probabilités d'extinction. Aucun exemple convaincant de caractère sélectionné pour favoriser la spéciation n'a pu être identifié. En revanche, des caractères de résistance à l'extinction ont pu être identifiés, comme la reproduction sexuée évoquée au chapitre 7. De fait, la plupart des auteurs admettent que les espèces parthénogénétiques ont, en général, des durées de vie moins longues que les espèces à reproduction biparentale. Dans ce cas, un même caractère semblait avantageux à l'espèce, mais désavantageux aux individus. Cependant, la majorité des caractères qui favorisent les individus diminuent également la probabilité d'extinction de l'espèce. La sélection entre individus et la sélection entre espèces sont alors deux manifestations du même phénomène, et on peut s'attendre à ce que les caractères changent dans le même sens à l'échelle intraspécifique et à l'échelle interspécifique. On observe ainsi une tendance évolutive à long terme. L'exemple des escargots marins illustre une telle

tendance : le registre fossile montre une augmentation progressive de la proportion d'espèces ayant une coquille résistante à la prédateur (ouverture étroite et épaisse), à partir de l'ère secondaire<sup>1</sup>. Cette évolution est probablement à mettre en relation avec une augmentation parallèle du nombre d'espèces possédant des spécialisations pour la prédateur des escargots. En conclusion, la sélection entre espèces et la sélection entre individus ne sont pas incompatibles. Au contraire, elles vont le plus souvent dans le même sens. Il est plausible de supposer une sélection entre espèces, mais il paraît très hardi de nier, comme le font les ponctualistes, l'efficacité de la sélection entre individus pour l'évolution morphologique, celle-ci ayant été observée en détail chez des organismes actuels (voir chapitre 8).

**Les extinctions d'espèces soulignent le rôle de la contingence dans l'histoire de la vie.** L'exemple de l'évolution des coquilles « antiprédateurs » chez les escargots marins reflète le caractère déterministe et prévisible des tendances macro-évolutives dues à la sélection. Quelle est la part dans la macro-évolution des facteurs déterministes et de la contingence ? L'étude des extinctions d'espèces fossiles offre l'occasion d'aborder cette question. C'est une tâche ardue car, lorsque une espèce A cesse d'apparaître dans les couches géologiques, cela peut signifier deux choses : soit elle s'est réellement éteinte, soit elle a changé morphologiquement et a été appelée B dans les couches géologiques suivantes. En prenant garde d'éviter ces pseudo-extinctions, les paléontologues ont établi deux faits remarquables. (i) La probabilité d'extinction n'augmente pas avec l'âge d'une espèce : les espèces ne vieillissent pas. (ii) Au cours de l'histoire de la vie, au

1. G. J. Vermeij, *Evolution and Escalation*, Princeton, Princeton University Press, 1987.

moins cinq crises d'*extinctions de masse* ont eu lieu. Pendant ces crises, un très grand nombre d'espèces (voire de genres et de familles) disparaissent rapidement. Ces crises modifient si profondément les faunes et les flores fossiles qu'elles délimitent les grandes périodes géologiques. En effet, ces périodes sont définies sur la base des changements paléontologiques. La crise la plus nootore correspond à la limite Crétacé-Tertiaire (soixante-six millions d'années) : elle voit très rapidement disparaître les dinosaures, les ammonites, ainsi que de nombreux brachiopodes et organismes planctoniques. Cette limite correspond aussi à l'emplacement d'une très fine couche sédimentaire, retrouvée en de nombreux endroits du globe, exceptionnellement riche en iridium, élément normalement rare sur terre. Cette observation a suscité la théorie de la météorite géante<sup>1</sup>. Selon cette théorie, une météorite de plusieurs kilomètres de diamètre aurait percuté la Terre, libérant une poussière riche en iridium. Ce choc très violent aurait perturbé durablement le climat (obscurcissement de l'atmosphère par les poussières). De nombreuses espèces, incapables de s'adapter à un changement si brusque, se seraient immédiatement éteintes. Ce scénario catastrophe a reçu d'autres arguments, notamment la présence de roches intensément déformées, et celle, plus controversée, des restes du cratère d'impact. Cependant, même en acceptant la présence d'une météorite, il est difficile d'être sûr que c'est elle qui a provoqué les extinctions. Dans certaines séries, celles-ci semblent commencer un million d'années avant la couche d'iridium. Ces données sont cependant débattues : il pourrait s'agir d'artefacts dus à la perturbation postérieure des

1. L. W. Alvarez *et al.*, « Extraterrestrial Cause for the Cretaceous-Tertiary Extinction », *Science*, 208, 1980, p. 1095-1108.

couches sédimentaires par des animaux fouisseurs. Quoi qu'il en soit, les autres extinctions de masse, dont la plus importante se situe à la limite Permien-Trias (- 225 millions d'années), ne semblent pas associées à des météorites. D'autres éléments comme un volcanisme particulièrement actif peuvent de la même façon modifier profondément le climat. Quoi qu'il en soit, les extinctions de masse sont dues à des catastrophes exceptionnelles qui n'ont peut-être pas toutes des causes semblables.

À la suite de ces crises, on observe une diversification de groupes autrefois discrets, allant de pair avec un anéantissemement ou une persistance très faible de groupes autrefois florissants. Ainsi, de toutes les branches de dinosaures, seule une (celle qui a donné les oiseaux actuels) a passé la crise Crétacé-Tertiaire. Elle a ensuite connu une diversification importante, de même que les mammifères, autres rescapés de la crise. Les hypothèses pour expliquer ces extinctions inégales entre groupes ne manquent pas. Certains ont par exemple proposé que l'homéothermie aurait permis aux mammifères, contrairement aux dinosaures, de résister au refroidissement climatique. Même si l'on pouvait vérifier ces hypothèses, on ne peut interpréter l'homéothermie comme une adaptation à l'arrivée d'une météorite-surprise. C'est par chance que certains groupes ont les caractères leur permettant de survivre lors des grandes crises. La contingence joue donc un rôle déterminant dans la grandeur et la décadence des grands groupes fossiles.

## Résumé

Le monde vivant est divisé en unités fondamentales qui ne sont pas des catégories arbitraires : les espèces.

Ces unités sont utilisées pour mesurer la diversité du monde vivant. Les individus vivants sont reliés entre eux par une généalogie en réseau, mais ce réseau, vu de loin, forme un graphe arborescent (l'arbre phylogénétique) dont les branches, une fois séparées, ne se rejoignent plus. Les espèces sont les « entre-nœuds » de cet arbre. Leur existence repose sur des mécanismes qui conduisent à la divergence permanente entre des groupes d'êtres vivants. Dans la pratique, deux types de critères essentiels sont retenus pour les délimiter : les relations de parenté entre les individus qui les composent (critère phylogénétique) et la capacité de ces individus à se reproduire entre eux (critère biologique). Les espèces nommées avec des noms scientifiques en latin sont des hypothèses sur la structure du réseau généalogique que l'accumulation des connaissances peut remettre en cause.

Les espèces naissent principalement par spéciation allopatrique. Dans ce modèle, les populations se différencient génétiquement grâce à une séparation géographique prolongée. Cette différenciation conduit à un isolement reproducteur entre les populations qui deviennent ainsi des espèces. D'autres modèles, moins bien documentés, proposent une spéciation ne nécessitant pas de séparation géographique. Dans tous les cas, la spéciation résulte de l'action de la sélection naturelle et de la dérive génétique. L'existence des espèces est donc une conséquence de la théorie de l'évolution.

L'évolution des caractères peut s'étudier à un niveau supérieur à l'espèce à l'aide des fossiles. Elle présente alors une apparence discontinue. La théorie des équilibres ponctués propose une interprétation de cette image. Elle propose que seule la sélection entre espèces, à l'exclusion de la sélection entre individus d'une même espèce, est responsable des changements morphologiques

à long terme. Les supports empiriques de cette interprétation sont faibles. Sélection d'individus et sélection d'espèce ne sont pas contradictoires. Elles introduisent toutes deux des tendances déterministes dans la macroévolution. Toutefois, la contingence se manifeste de manière spectaculaire lors de crises d'extinctions de masse qui réorganisent en profondeur la diversité biologique.



## ÉPILOGUE

Plusieurs éléments attestent le succès de la théorie de l'évolution. La population de chercheurs en biologie évolutive ne cesse de s'accroître, notamment en France. Cette branche a intégré fructueusement des outils et des personnes venant d'autres disciplines (mathématiques, épistémologie, chimie...). Les raisonnements évolutifs ont infiltré les autres domaines de la biologie comme l'écologie ou la biologie du développement.

La théorie de l'évolution unifie les biologistes en regroupant leurs objets d'étude, les organismes, dans une histoire commune régie par des lois communes. Chaque biologiste en trouvera les conséquences dans son travail quotidien. La théorie de l'évolution a su absorber les révolutions majeures de la biologie du XX<sup>e</sup> siècle, comme la génétique (1900-1930) ou plus récemment la biologie moléculaire (1960-1970). Elle s'apprête aujourd'hui à absorber le renouveau de la biologie du développement.

À partir des modèles simples de la génétique des populations des années 1930, la théorie a mûri, gagnant en complexité et en réalisme. Elle prend désormais en compte les interactions entre les unités sélectionnées, et les propriétés « émergentes » qui apparaissent à des niveaux d'intégration de plus en plus élevés : gènes, génomes, individus, groupes, espèces. Cette « évolution du deuxième ordre » s'intéresse à l'histoire des relations

entre objets biologiques au moins autant qu'à celle des objets eux-mêmes. Enfin, l'évolution du sexe et celle du développement embryonnaire ouvrent la perspective d'un autre type d'évolution du « deuxième ordre » : l'évolution de la capacité à évoluer.

Au terme de ce livre, nous espérons vous avoir convaincu(e) que la théorie de l'évolution est le cadre logique de toute la biologie. C'est une théorie touche-à-tout. Elle prétend expliquer presque toutes les propriétés des systèmes vivants, des plus immédiates, comme le blindage d'une coquille protégeant des prédateurs, aux plus surprenantes, comme les préférences des oiseaux femelles pour des mâles décoratifs mais handicapés. Cette prétention pourrait lui être reprochée : les évolutionnistes n'ont-ils pas toujours une anecdote toute prête à sortir de leur sac ? Pas vraiment, car leur imagination débordante se retourne contre eux. Ils s'emparent d'un phénomène qui semblait tout naturel, comme le sexe, et le transforment en un problème déroutant. Auparavant, nous pensions simplement que le sexe servait à la reproduction. Maintenant, nous nous demandons pourquoi nous utilisons un mode de reproduction aussi bizarre au lieu de nous cloner tout simplement. Quelles que soient les réponses, on ne peut plus voir la biologie de la même façon après avoir écouté les arguments évolutionnistes.

Les progrès de la biologie dépassent le simple cadre des spécialistes. Ils touchent notre quotidien à travers leurs applications à l'agriculture, la médecine, l'environnement. La théorie de l'évolution apporte un éclairage essentiel sur ces pratiques. Elle peut par exemple nous aider à repenser nos relations avec le milieu naturel, qu'il s'agisse de lutter contre des pathogènes inventant sans cesse de nouvelles résistances ou de gérer une biodiversité issue de millions d'années d'évolution.

## GLOSSAIRE CRITIQUE

### Notions

#### *Amélioration génétique des espèces*

Dès le Néolithique, la domestication de plantes et d'animaux est une forme intuitive de sélection artificielle. Durant cette période, l'homme a isolé, dans la variabilité génétique des espèces sauvages, des variants répondant à des critères de rendement. Par exemple, le maïs est une forme à gros grains d'une espèce sauvage, la téosinte. Dans différentes populations humaines, des lignées différentes ont été sélectionnées. Ainsi, il existe, malgré une intense sélection, une diversité génétique considérable entre les différentes lignées de maïs.

L'amélioration génétique des espèces a été rationalisée avec le développement de la génétique quantitative au XX<sup>e</sup> siècle, qui établit des protocoles efficaces pour trier les individus en fonction de leur valeur génétique. De ce fait, les qualités agronomiques des espèces domestiquées, comme le blé, ont connu une progression rapide. Des voies plus complexes qu'une simple sélection directionnelle ont été inventées pour obtenir rapidement des variétés performantes. Chez le maïs, la technique des hybrides a ainsi été développée (voir *Hétérosis et dépression de consanguinité*, p. 319).

Une nouvelle technique d'amélioration génétique des plantes fait aujourd'hui grand bruit : ce sont les « organismes génétiquement modifiés » ou OGM. On introduit directement dans l'organisme un gène provenant d'un autre organisme, qui apportera un caractère précis tel que la résistance à un herbicide. Cependant, même si le gène introduit est bien connu, on ne sait pas forcément prédire quelles seront ses interactions avec les autres gènes du génome. En outre, l'espèce modifiée, une fois plantée dans un champ, risque de transmettre le nouveau gène aux espèces sauvages apparentées si elle est capable de les polliniser. De même que l'invasion d'une espèce introduite dans un écosystème, l'invasion de plantes sauvages par un gène échappé des OGM pourrait avoir des conséquences difficiles à prévoir.

### *La biodiversité et sa préservation*

La *biodiversité* désigne l'ensemble de la diversité biologique, qu'il s'agisse de diversité génétique ou du nombre d'espèces. Elle se dénombre principalement à l'aide de la catégorie « espèce ». Il est difficile d'estimer avec précision le nombre d'espèces du globe. Le 1,7 million d'espèces décrites ne représente sans doute que 1 à 10 % de la diversité totale. Les différents groupes ne sont pas également connus. Par exemple, il est probable que le nombre d'espèces d'oiseaux connues (9 672) reflète bien la diversité de ces organismes qui sont faciles à observer. En revanche, bien que l'on pense que le million d'espèces d'arthropodes décrites est loin de couvrir la diversité complète de ce groupe, on est incapable d'estimer le nombre restant à découvrir.

Au cours de l'évolution se succèdent naissances et extinctions d'espèces. L'histoire de la vie a connu plusieurs grandes crises au cours desquelles se sont éteintes de nombreuses espèces. La plus connue est l'extinction de masse du Crétacé. Aujourd'hui, les extinctions d'espèces s'accélèrent, résultat de l'action de l'homme sur le milieu naturel. Certains pensent que la crise actuelle a une ampleur comparable ou supérieure aux grandes crises passées (on parle de la sixième extinction), après lesquelles la biodiversité n'a regagné son niveau antérieur qu'en plusieurs millions d'années.

Deux comportements humains conduisent principalement à cette accélération. Tout d'abord, l'homme modifie profondément les milieux, que ce soit avec l'agriculture ou, de façon plus récente, avec l'urbanisation. Ainsi, la réduction de la forêt amazonienne entraîne l'extinction d'espèces qui y sont inféodées. Ensuite, l'homme transporte aux quatre coins du monde des espèces qui deviennent envahissantes. Dans certains milieux, notamment les îles, les invasions ont des conséquences dramatiques. Les espèces envahissantes prennent la place d'espèces locales qui n'existaient qu'à cet endroit.

Pour préserver la biodiversité, différentes stratégies sont possibles. On peut choisir de protéger des espèces menacées d'extinction ou encore de protéger des milieux menacés. Pour faire des choix de conservation, comme celui de l'espèce à préserver en priorité, il faut des connaissances sur le fonctionnement des populations naturelles. De fait, certaines espèces ont un rôle clé dans le maintien global d'un écosystème. Par exemple, la disparition de la loutre de mer sur les côtes pacifiques américaines entraîna une prolifération des oursins et la fragmentation d'un système de « forêt » d'algues marines. On parle d'espèce *clé de voûte* (*key-stone species*). Il est

difficile de prédire par avance quelles espèces jouent ce rôle. L'étude du fonctionnement des écosystèmes à des fins de préservation de la biodiversité fait l'objet de la *biologie de la conservation*.

### *Génétique et eugénisme*

Né en Angleterre au XIX<sup>e</sup> siècle, *l'eugénisme* préconise une sélection en vue d'améliorer l'espèce humaine. Défendu par des scientifiques reconnus (Galton), il suscita de nombreuses publications scientifiques.

Divers gouvernements occidentaux (États-Unis, Suède, Allemagne...) mirent en pratique l'eugénisme dès les années 1930 par des campagnes massives de stérilisation. Elles visaient à empêcher la prolifération d'individus jugés indésirables eu égard à des critères variés : débilité, délinquance, pauvreté, couleur de peau. Les pratiques eugénistes culminèrent en Allemagne nazie, où la stérilisation laissa la place à des plans d'extermination.

L'eugénisme, comme tout programme de sélection artificielle, suppose l'existence d'une variation génétique des caractères sélectionnés. Si certains caractères comme la couleur de la peau sont clairement génétiques, d'autres, comme la pauvreté, ne le sont pas. Les débats les plus animés portèrent sur l'*« intelligence »*. C. Burt inventa les statistiques multivariées pour obtenir des scores d'intelligence synthétisant les résultats de nombreux tests. D'après une étude des ressemblances entre jumeaux séparés à la naissance, il affirma que ces scores étaient génétiques. Mais son jeu de données semble avoir été inventé. Les Américains généralisèrent un autre score : le quotient intellectuel (QI). Des études récentes sur des jumeaux suggèrent que le QI dépend faiblement

des gènes, et très fortement de l'environnement. Globalement, elles suggèrent qu'il est plus efficace d'agir sur l'environnement (santé, éducation) que sur les gènes (eugénisme) si l'on veut améliorer le QI moyen d'une population. Cependant, elles ne permettent pas de savoir si le QI a une quelconque signification.

Les objections quant à l'efficacité de l'eugénisme ne sont cependant pas les plus pertinentes. Le problème éthique fondamental est de savoir si nous avons le droit de stériliser ou de tuer pour modifier l'évolution de l'espèce dans un sens précis. Cette question n'est pas du ressort des scientifiques. Bien que l'eugénisme ait été massivement rejeté après la période nazie, le problème pourrait à nouveau se poser. En effet, les techniques scientifiques de diagnostic prénatal rendent aujourd'hui possibles des choix individuels apparentés à l'eugénisme.

### *Hétérosis et dépression de consanguinité*

La *dépression de consanguinité* est la diminution de valeur sélective observée chez les descendants d'autofécondation ou d'autres croisements entre individus apparentés. Au XX<sup>e</sup> siècle, des généticiens américains l'étudièrent chez le maïs afin d'améliorer son rendement. En autofécondant de manière répétée des plantes, ils créèrent des *lignées consanguines*. Les performances agronomiques diminuaient au fur et à mesure des générations d'autofécondation. Cet effet disparaissait en recroisant ces lignées entre elles : les descendants obtenus (*hybrides*) avaient des performances égales, voire supérieures, aux performances de départ. Cet effet, appelé *hétérosis* ou *vigueur hybride*, est à la base des protocoles d'améliora-

tion génétique du maïs : on cherche à isoler des lignées donnant les meilleurs hybrides possibles.

Pour la majorité des chercheurs, la dépression de consanguinité et l'hétérosis résultent de la présence d'allèles récessifs délétères. La consanguinité met ces allèles à l'état homozygote, ce qui diminue la valeur sélective (dépression). Le croisement entre deux lignées homozygotes pour des allèles différents entraîne une hétérozygotie généralisée dans le génome, ce qui empêche les effets délétères de s'exprimer (hétérosis). Cependant, la qualité des hybrides est éphémère. Dès qu'ils se croisent, ils produisent des descendants ayant deux fois moins de gènes hétérozygotes, ce qui diminue les performances agronomiques. Il est donc peu rentable de semer les graines issues d'un champ d'hybrides. Il faut conserver les lignées consanguines pour produire chaque année des semences hybrides de première génération. Selon certains auteurs, cet effet maintient une dépendance des agriculteurs vis-à-vis des entreprises qui leur vendent les semences, les obligeant à racheter des hybrides à chaque cycle de culture.

### *Sociobiologie et darwinisme social*

On désigne ainsi deux tentatives faites à des époques différentes pour appliquer des raisonnements évolutionnistes aux sociétés humaines.

Spencer, précurseur du darwinisme social, expliquait le fonctionnement des sociétés par la « survie du plus apte ». Sur cette base, certains suggérèrent que le fonctionnement inégalitaire de la société était un mal nécessaire pour progresser vers un bien-être général. Avec l'essor du capitalisme au XIX<sup>e</sup> siècle, le darwinisme social

fut utilisé comme argument pour justifier l'abandon de l'assistance aux démunis sous prétexte qu'elle ne ferait que ralentir le progrès social. D'autres tendances issues du darwinisme social préconisaient, à l'inverse, une coopération entre les classes sociales, considérée comme plus efficace pour progresser. Le darwinisme social est une transposition sociale de l'eugénisme.

La sociobiologie, quant à elle, a été fondée par l'entomologiste Edward O. Wilson. La compréhension théorique de l'évolution des comportements coopératifs a beaucoup progressé dans les années 1950-1970, avec les travaux de Wynne-Edwards (sélection de groupe), Maynard-Smith (théorie des jeux) et Hamilton (sélection de parentèle). Wilson (*Sociobiology, The New Synthesis*) associe ces théories avec les connaissances naturalistes afin de comprendre l'évolution des sociétés animales. Cette tentative serait restée très consensuelle si le dernier chapitre de cet ouvrage n'avait abordé l'espèce humaine, déclenchant une polémique virulente. En effet, les écrits de Wilson, à l'instar de ceux de Spencer un siècle plus tôt, suggéraient un déterminisme social. Ils suggéraient que les comportements humains étaient dictés principalement par des gènes, la structure des sociétés résultant de la sélection naturelle. Ainsi, l'investissement parental des femmes (fabrication des ovules, gestation) les pousserait à s'occuper du foyer et des enfants, alors que l'incertitude des hommes sur leur paternité les pousserait à chercher des aventures. Ces positions furent très critiquées. Premièrement, les spécialistes jugeaient très approximatives les généralisations faites par Wilson sur les sociétés humaines. Deuxièmement, le déterminisme des comportements humains est en partie génétique, mais la transmission culturelle est plus grande que pour toute autre espèce. Les théories sociobiologiques font

l'impasse sur la transmission non génétique des comportements, ou la traitent de manière jugée superficielle. Troisièmement, le fait que l'homme est un produit de l'évolution n'implique pas qu'il ne puisse pas lui-même modifier les conditions de cette évolution. Il peut créer une morale pour contrôler ses conduites.

L'intensité des réactions suscitées par la sociobiologie a été variable. En France, elle a été diabolisée, comme si *a priori* toute approche biologique des conduites humaines sentait le soufre. Selon Gouyon *et al.*, cette réaction extrême est liée au souvenir traumatisant des pratiques eugénistes du régime de Vichy. Aux États-Unis, la réaction générale a été plus modérée, fondée sur des débats argumentés. Certes, les sociétés humaines ne sont pas qu'une affaire de gènes ; on peut néanmoins penser que les concepts évolutionnistes peuvent utilement être intégrés aux côtés des concepts traditionnellement utilisés par les sociologues et anthropologues pour comprendre l'histoire des sociétés humaines.

### *La systématique*

La systématique est une discipline scientifique qui vise à ordonner le vivant. Au sein de cette discipline, on peut distinguer deux grandes tâches : la définition des taxons et la classification de ces taxons selon un système prédéfini. Généralement, l'activité de définition des taxons est appelée « taxonomie » et concerne principalement la délimitation des espèces.

La description d'une espèce est associée aux noms des auteurs de cette description car elle représente une hypothèse émise par eux au vu des données disponibles à la date où ils ont travaillé et des théories auxquelles ils

adhéraient. Ces hypothèses sont révisées quand des données nouvelles sont disponibles : nouveaux spécimens collectés, nouvelles méthodes pour examiner les organismes (microscopie, ADN, etc.), nouvelles conceptions théoriques de la biologie. Les noms scientifiques en latin permettent de suivre, grâce aux règles des codes de nomenclature, ce processus de révision des hypothèses. Souvent, les descriptions de nouvelles espèces sont fondées sur des critères de ressemblance morphologique. Cette démarche a souvent été critiquée et qualifiée de typologique ou de « phénétique ». Cependant, le choix des caractères morphologiques utilisés dans ces descriptions n'est pas arbitraire et repose sur l'expérience acquise par les taxonomistes. Par exemple, chez les insectes, avant l'invention du microscope, les mâles et les femelles de la même espèce, d'apparences souvent différentes, étaient parfois décrits comme des espèces séparées. Le microscope ayant permis d'observer plus finement l'anatomie, ces hypothèses furent révisées et permirent d'attribuer à la même espèce les mâles et femelles correspondants. Depuis, les entomologistes prennent en compte ces connaissances quand ils analysent les caractères morphologiques dans le but de proposer – avec un critère de ressemblance – la délimitation de nouvelles espèces.

Les espèces ainsi décrites sont ensuite classées, depuis Linné, selon un système hiérarchique reposant sur l'observation. Par exemple, un chien ressemble plus à un chat qu'à une fourmi. Nous classerons donc chiens et chats dans un même groupe (les mammifères, plus précisément les carnivores) qui ne comprend pas la fourmi. Pour Linné, cette hiérarchie est justifiée car elle illustre un ordre établi par une puissance divine. Darwin proposa quant à lui une explication scientifique à ces observations : le rapprochement du chat et du chien résulte

du partage d'un ancêtre commun plus récent qu'avec la fourmi. Darwin propose donc le premier que la classification des êtres vivants repose sur l'histoire évolutive. En conséquence, après des débats houleux au cours du XX<sup>e</sup> siècle, les classifications s'appuient aujourd'hui principalement sur les différentes méthodes de reconstructions phylogénétiques.

### *La théorie des jeux*

La théorie des jeux a été proposée par des mathématiciens (von Neumann et Morgenstern) au début du XX<sup>e</sup> siècle afin de comprendre les systèmes économiques et militaires. L'objectif est de savoir, dans un contexte où les décisions des uns ont des conséquences sur les autres, quelle stratégie permet de perdre le moins possible. Un modèle classique est le « dilemme du prisonnier ». Deux prisonniers sont jugés. Si les deux avouent le crime, ils sont tous deux légèrement condamnés. Si l'un avoue et l'autre nie, celui qui avoue est récompensé et l'autre lourdement condamné. Si les deux nient, ils sont tous deux libérés. Si les deux prisonniers pouvaient s'accorder, ils nieraient tous deux. Cependant, sans information, la stratégie qui minimise la perte est d'avouer pour ne pas risquer d'être lourdement condamné à la suite de l'aveu de l'autre.

Ces modèles ont été appliqués en biologie évolutive, notamment par Maynard-Smith, afin de comprendre le comportement d'individus ayant des conflits d'intérêt. Le dilemme du prisonnier a ainsi été utilisé pour comprendre pourquoi les oiseaux poussaient des cris d'alarme prévenant les autres membres du groupe de l'arrivée d'un prédateur. L'oiseau qui crie prend le risque de se faire

repérer. Par analogie avec le modèle du prisonnier, la meilleure stratégie est de ne pas prendre ce risque. La théorie des jeux pose donc le problème de l'existence de comportements coopératifs normalement contre-sélectionnés.

La théorie des jeux a permis de modéliser l'évolution de nombreux caractères. Elle a néanmoins ses limites car elle prédit un optimum sans tenir compte des mécanismes génétiques qui permettraient de l'atteindre.

## Acteurs et controverses

### *Biométriciens et mendéliens*

Le débat entre les biométriciens et les mendéliens (fin XIX<sup>e</sup>- début XX<sup>e</sup>) portait à la fois sur la nature de l'hérité dit et sur celle de l'évolution. La démarche des biométriciens, initiée par Galton, était d'analyser la variation de caractères métriques (comme la taille) et leur transmission d'une génération à l'autre. Le but était de comprendre l'action de la sélection naturelle sur ces caractères. Galton et ses successeurs, notamment K. Pearson, mirent au point des techniques mathématiques (analyse de distributions, corrélations, régressions) qui ont fondé les statistiques modernes et la génétique quantitative. Ils décrivirent l'hérité dit comme un ensemble de lois mathématiques liant les valeurs des parents (ou d'ancêtres plus éloignés) à celles des enfants. Il s'agissait d'une hérité dit-mélange (*blending inheritance*) : la valeur du descendant était une sorte de moyenne des valeurs de ses ascendants. Pearson défendait l'idée d'une évolution graduelle fondée sur une sélection naturelle agissant sur des caractères à variation continue.

Cette idée se heurtait néanmoins à un paradoxe : l'hérédité-mélange supprimait rapidement la variation car les descendants étaient toujours plus proches de la moyenne que leurs parents. Cette idée était renforcée par des données sur les tailles humaines, analysées par Galton, montrant que les enfants de couples très grands étaient en moyenne moins grands que leurs parents, et les enfants de couples petits moins petits. Ce phénomène de « *régression vers la moyenne* » (*regression towards mediocrity*) indiquait selon lui une diminution de la variabilité génétique de génération en génération, toute la population devenant « moyenne ». Nous savons maintenant que ce n'est qu'une impression due à l'existence d'une variabilité non transmise (variance environnementale). Certains parents grands doivent leur taille à leurs gènes, d'autres à un environnement exceptionnellement favorable lors de la croissance. Les descendants héritent des gènes mais pas de l'environnement : ils sont donc en moyenne moins grands. Quoi qu'il en soit, la régression vers la moyenne soulevait à l'époque un sérieux doute sur l'efficacité de l'action de la sélection naturelle sur la variation continue.

Dans ce contexte, les mendéliens (G. Bateson, H. De Vries, W. Johannsen, T. Morgan) défendaient l'idée d'une variation discontinue en ressuscitant les modèles établis par Mendel (locus, gènes, allèles). Ces modèles avaient un avantage sur l'hérédité-mélange : les gènes sont des particules d'hérédité qui se transmettent en bloc, sans se mélanger ou se moyenner. La variation génétique se transmet donc intacte au fil des générations. Cette absence de changement à court terme semblait contraire à l'idée d'évolution. Les mendéliens réussirent néanmoins à rendre les gènes compatibles avec un changement évolutif : ils montrèrent l'existence d'événements

de *mutations*, changements instantanés dans les gènes créant de nouveaux allèles à comportement mendélien. Ils allèrent jusqu'à supposer que des mutations de grande ampleur (*macromutations*) étaient à l'origine de l'évolution : celle-ci procédait par sauts. L'évolution graduelle par sélection naturelle était pour eux un mythe. Les travaux expérimentaux des mendéliens établirent définitivement la validité et la généralité des lois de Mendel. Cependant, la théorie des macromutations est maintenant abandonnée : si l'on considère encore possible (mais peu probable) l'apparition de mutations à effet majeur, la génétique des populations nous a appris que l'avenir évolutif d'une mutation dépend de processus de tri (sélection naturelle et dérive) postérieurs à l'événement de mutation lui-même.

On peut rétrospectivement se demander pourquoi la controverse biométriciens-mendéliens fut si violente. Les querelles personnelles et académiques ont sans doute envenimé le débat, en particulier l'opposition entre le University College de Londres, berceau des biométriciens (Galton, Pearson, Weldon) et l'université de Cambridge, personnifiée par un farouche mendélien (Bateson). On admet aujourd'hui qu'il n'y a pas d'incompatibilité théorique mais une complémentarité entre l'hérédité mendélienne et la génétique quantitative (héritière de l'école des biométriciens). La variation d'un caractère continu peut reposer sur des gènes à la façon dont la forme d'un tas de sable repose sur le nombre et l'arrangement de ses grains. On admet aussi que la mutation et la sélection naturelle sont également importantes pour l'évolution, l'une fournissant la variabilité nécessaire à l'action de l'autre. Peut-être la polémique naquit-elle du fait d'avoir mélangé deux problèmes logiquement indépendants : les mécanismes de la transmission à court

terme des caractères entre individus (héritéité) et ceux de leur changement à long terme dans les populations (évolution).

### *Caractères acquis ou mutations aléatoires ?*

Selon l'hypothèse de l'*héritéité des caractères acquis*, popularisée par Lamarck, des changements acquis au cours de la vie d'un organisme, en réponse à une contrainte du milieu, peuvent être transmis à ses descendants. Ainsi, la girafe étire son cou pour accéder aux feuilles des arbres. Cet allongement est transmis à ses descendants, entraînant l'évolution d'un cou de plus en plus long. Selon le concept de *mutation aléatoire*, les changements héritables (mutations) ne sont pas orientés par les contraintes du milieu. L'adaptation est alors due au tri effectué postérieurement par sélection naturelle. En fait, dans les deux hypothèses, les changements sont acquis. Le terme de *caractères acquis* est un héritage historique mais il pourrait avantageusement être remplacé par *mutation dirigée*.

Les premiers arguments contre les *caractères acquis* viennent de Weismann. Celui-ci découvre que les gamètes, qui transmettent l'information héréditaire à la génération suivante chez les organismes pluricellulaires, viennent d'une lignée cellulaire particulière non affectée par les changements du milieu. L'héritéité ne peut donc pas se modifier à la suite de ces changements. Le développement de la génétique apporta des preuves expérimentales des mutations aléatoires, même chez des organismes unicellulaires. Luria et Delbrück montrèrent que chez les bactéries, les mutations conférant la résistance à un virus n'arrivaient pas en réponse à sa présence,

mais avant même leur mise en contact. Les mutations sont donc aléatoires.

L'hypothèse des *caractères acquis* resurgit périodiquement au cours du XX<sup>e</sup> siècle. Le biologiste soviétique Lyssenko imposa cette hypothèse en URSS, à partir d'expériences probablement truquées sur la réaction des plantes au froid. Soutenu par Staline, Lyssenko prit en main la biologie soviétique et fit arrêter les partisans de la génétique classique, considérée comme une théorie occidentale réactionnaire. Il ne fut destitué qu'en 1965, par Khrouchtchev. Loin de ces excès, la problématique a été récemment renouvelée par l'observation de formes de mutation dirigée chez des bactéries. Ces données suggèrent un mode d'adaptation original mais ne remettent pas en cause que les mutations sont le plus souvent aléatoires.

### *Cladisme et phénéticisme*

La phylogénie retrace l'histoire évolutive des espèces, sous forme d'un arbre, en établissant leur ordre d'apparition au cours du temps. Deux écoles ont proposé des méthodes différentes pour reconstruire ces arbres. Le principe général est que plus deux groupes se sont différenciés récemment, plus ils se ressemblent. La *phénétique* utilise l'ensemble des différences entre deux taxons pour estimer le temps qui les sépare. L'approche *cladiste* différencie les caractères en fonction de leur ordre d'apparition dans l'évolution. Elle distingue ainsi des caractères *ancestraux* (hérités des espèces ancestrales) et des caractères *dérivés* (caractères qui se sont modifiés dans une lignée). L'ordre de branchement est établi par le partage de caractères dérivés. Ainsi, si parmi trois taxons, deux

partagent un caractère dérivé, absent chez le troisième, ce sont les deux taxons les plus proches dans l'arbre. En revanche, s'ils partagent un caractère ancestral, absent chez le troisième, cela implique simplement que le troisième a perdu ce caractère. On ne peut donc rien en déduire sur les proximités phylogénétiques de ces taxons. Cette démarche se heurte à une difficulté : le même caractère peut changer plusieurs fois indépendamment le long d'un arbre phylogénétique. Un même état de caractère peut alors être à la fois ancestral et dérivé. Pour surmonter ce problème, les cladistes utilisent la *parcimonie* ou économie d'hypothèses. Ils choisissent parmi tous les arbres possibles celui qui nécessite le moins d'hypothèses de changements indépendants. Théoriquement, la méthode cladiste est la plus satisfaisante. Dans les faits, il est parfois difficile d'établir quel arbre est le plus parcimonieux car les analyses statistiques sous-jacentes deviennent vite complexes. Les méthodes cladiste et phénotypique peuvent toutes deux s'appliquer aux caractères morphologiques et moléculaires. Dans le second cas, la méthode phénotypique donne une estimation fiable de la phylogénie quand l'hypothèse de l'horloge moléculaire est vérifiée (voir chapitre 2).

Les développements techniques de l'informatique ont permis d'ouvrir une troisième voie avec l'approche probabiliste. Cette approche regroupe plusieurs méthodes apparentées (méthode de maximum de vraisemblance, méthode bayésienne) qui utilisent des modèles d'évolution des caractères pour calculer la vraisemblance des arbres phylogénétiques. Ces approches permettent notamment de prendre en compte des différences de taux d'évolution entre les branches de l'arbre ou entre les différents caractères utilisés. Par exemple, dans les séquences d'ADN qui codent pour des protéines, en

raison de la dégénérescence du code génétique (voir chapitre 2), les différentes positions n'évoluent pas à la même vitesse. Les modèles permettent de prendre en compte ces données et donc d'améliorer la reconstruction des arbres. Les débats sur les méthodes de reconstruction phylogénétique, très vifs dans la seconde moitié du XX<sup>e</sup> siècle, ont aujourd'hui été dépassés et les discussions portent plus sur la justification de la méthode de reconstruction en fonction de la nature du jeu de données que sur la validité intrinsèque de telle approche par rapport à telle autre.

### *Équilibres ponctués*

Cette théorie sur le rythme et les modalités de l'évolution morphologique a été énoncée en 1972 par N. Eldredge et S. J. Gould et a suscité une vaste polémique pendant une quinzaine d'années. Elle part d'un constat : dans les données paléontologiques, c'est l'absence de changement morphologique au cours du temps, ou *stase*, qui est l'observation la plus commune, d'où l'énoncé « *stasis is data* ». Pourtant les paléontologues ne s'intéressent qu'aux lignées de fossiles présentant des changements graduels de formes au cours du temps, ils ont donc une vision biaisée du tempo et du mode de l'évolution morphologique. À partir de ce constat, Eldredge et Gould proposent une théorie où les changements de morphologie sont rares et brusques (et donc instantanés à l'échelle fossile). Ils soutiennent que la morphologie est sujette à de fortes contraintes (adaptatives, développementales, etc.). Elle ne varie donc qu'exceptionnellement au sein des espèces. La variation est toujours brusque et associée à la naissance de nouvelles

espèces. Ces auteurs estiment donc que les mécanismes décrits à l'échelle de la biologie des populations (micro-évolution) ne permettent pas de décrire les changements majeurs de la morphologie observables sur les restes fossilisés des êtres vivants (macro-évolution).

Dans les derniers développements de cette théorie, ils suggèrent qu'il existe un niveau de sélection au-dessus de l'individu : celui de l'espèce (*sélection d'espèces*). Cependant, ce modèle n'a jamais été rigoureusement formalisé. La majeure partie de la littérature sur le sujet a pour origine la polémique sur le tempo des changements, graduels ou ponctués. Les arguments empiriques des ponctualistes n'ont pas convaincu la majorité des évolutionnistes. Néanmoins, quelle que soit l'issue de cette polémique, elle aura stimulé un renouveau de la paléontologie évolutive.

### *La génération spontanée*

La *génération spontanée* est l'apparition directe d'un être vivant à partir de matière inanimée. Elle s'oppose à l'idée qu'un organisme est toujours engendré par un autre organisme. Au Moyen Âge, la génération spontanée est admise : la viande pourrie produit naturellement des mouches. Aux XVII<sup>e</sup> et XVIII<sup>e</sup> siècles, le doute s'installe, suscitant des tests expérimentaux. Des organismes apparaissent-ils dans une matière inanimée placée dans un bocal hermétique ? F. Redi montre que les mouches n'apparaissent que dans la viande laissée dans des bocaux ouverts. Mais l'invention du microscope et la découverte d'organismes invisibles à l'œil nu posent de nouveau le problème : s'il n'y a pas de mouches, qu'en est-il des microbes ? Les expériences confirment la génération

spontanée : quelles que soient les précautions prises, les microbes apparaissent toujours, même dans un récipient hermétiquement fermé. Finalement, c'est Pasteur qui infirme définitivement la génération spontanée, en 1868. Il invente un système permettant de garder un bouillon de culture perpétuellement stérile (ballon « à col de cygne »). Il montre aussi que notre environnement grouille de germes microbiens, expliquant l'échec des tentatives précédentes.

Ayant abandonné l'idée de génération spontanée actuelle, les chercheurs spéculent sur le passé. En effet, la théorie de l'évolution implique un événement de génération spontanée : l'apparition de la vie. Ce problème diffère des précédents car l'« organisme » originel est inconnu. Son organisation devait être beaucoup plus simple que celle des organismes actuels. Au lieu de supposer qu'un assemblage aussi complexe qu'une cellule vivante naît de rien, on recherche une chaîne de transformations progressives. Des molécules inorganiques se seraient transformées en molécules organiques de plus en plus complexes, jusqu'à des assemblages supramoléculaires ressemblant à de petites cellules. Les théories sur l'origine de la vie intègrent donc la génération spontanée dans le monde « familier » de la chimie et de la biologie.

### *Neutralistes contre sélectionnistes*

Dès les années 1960-1970, les progrès techniques permirent de mesurer la variation moléculaire entre organismes à trois niveaux : la variation protéique observée par électrophorèse, la variation des séquences des protéines en acides aminés et celle des séquences d'ADN. Ces mesures donnèrent deux résultats majeurs.

1) Les populations actuelles présentent un fort polymorphisme.

2) Les taux de substitution au sein d'une espèce sont élevés : l'apparition et la fixation de nouveaux variants sont fréquentes.

Ces deux observations étaient surprenantes car on pensait jusque-là que la sélection naturelle éliminait efficacement la variation. On supposait que les mutations étaient le plus souvent éliminées et, dans de rares cas, favorables, donc rapidement fixées : le polymorphisme et les taux de substitution auraient dû être faibles.

Kimura émit, pour expliquer la variation moléculaire, la *théorie neutraliste de l'évolution moléculaire*. Elle stipule que la majorité des variations moléculaires observées ne sont pas soumises à la sélection naturelle : le polymorphisme observé actuellement dans les populations est transitoire, composé d'allèles neutres soumis à la seule dérive génétique. De même, les substitutions observées au cours du temps sont dues surtout à des fixations aléatoires de variants neutres. À cette école s'opposèrent des conceptions « *sélectionnistes* ». Ces dernières donnaient à la sélection naturelle le rôle prépondérant, mais supposaient un type particulier de sélection. Les forts taux de substitution s'expliquaient par des mutations favorables relativement fréquentes. Le fort polymorphisme résultait d'une *sélection stabilisante*, maintenant la variation : superdominance (voir chapitre 6), sélection en environnement variable (voir chapitre 9), sélection fréquence-dépendante (voir chapitre 10). Les deux théories étaient en accord sur le fait que la plupart des mutations spontanées sont délétères et vite éliminées (« *sélection purifiante* »). Mais les autres mutations, celles qui se fixent à long terme, étaient neutres pour les uns, et positivement sélectionnées pour les autres. Les deux « camps » se sont

affrontés sur la base d'arguments empiriques et théoriques.

**Arguments sur les taux de substitution.** Pour Kimura, les substitutions observées ne peuvent être dues à la sélection. Un mutant favorable n'envahit une population que s'il élimine les autres allèles. Les individus portant ces allèles doivent donc mourir ou se reproduire peu (*fardeau génétique*). Les individus favorisés doivent compenser ce fardeau en se reproduisant en excès. Les taux de substitutions mesurés sont si élevés que le fardeau calculé par Kimura implique un excès reproducteur irréalisable.

*L'hypothèse du fardeau génétique a cependant été infirmée par la suite.* L'erreur était que les fardeaux correspondant aux différents gènes ne s'additionnent pas. Ainsi, la mort d'un seul individu peut éliminer à la fois plusieurs allèles délétères situés sur des locus différents.

*La constance des taux de substitution (horloge moléculaire) est prédite par les neutralistes.* La théorie neutraliste stipule en effet que le taux de substitution égale le taux d'apparition de mutations neutres, supposé constant (voir chapitre 3). Les sélectionnistes ont deux réponses à cet argument : premièrement, l'horloge moléculaire peut être obtenue par sélection si le taux d'apparition de mutations bénéfiques est constant. Deuxièmement, dans les données empiriques, l'horloge moléculaire est vérifiée en général mais fausse sur de nombreux cas particuliers.

*Les séquences fonctionnellement peu importantes évoluent vite, comme le prédit la théorie neutraliste.* Par exemple, les sites où la substitution n'a pas d'effet (synonyme) évoluent plus vite que les sites où la substitution change un acide aminé (non synonyme). Cela est prédit par la théorie neutraliste, car il y a plus de mutations neutres (bénignes) dans des séquences peu importantes. Comme précédemment, les sélectionnistes ont une objection

double : d'une part, ils proposent des hypothèses sélectionnistes donnant le même résultat (par exemple, les mutations favorables sont plus rares dans les séquences importantes) ; d'autre part, ils exhibent des contre-exemples (les gènes du CMH des mammifères, voir chapitre 10).

**Arguments sur le polymorphisme actuel.** *Les sélectionnistes ont trouvé dans la nature des exemples de sélection stabilisante.* C'est le cas de la superdominance observée pour l'anémie falciforme chez l'homme (voir chapitre 5). Cependant, cet exemple est demeuré isolé. Les autres formes de sélection stabilisante sont mieux documentées, mais les démonstrations expérimentales restent rares.

*Les données d'électrophorèse des protéines montrent que le degré de polymorphisme et les taux de substitutions sont corrélés positivement, comme la théorie neutraliste le prévoyait.* Cette théorie prédit, en effet, qu'ils sont tous deux proportionnels au taux de mutations neutres. Cette prédition n'est pas partagée par les hypothèses sélectionnistes.

La théorie neutraliste a suscité d'après controverses car l'enjeu était de taille. En effet, la neutralité de la variation moléculaire est une hypothèse fondamentale utilisée pour de nombreuses applications : analyse de la structure des populations et de leur histoire récente, datation des divergences par l'horloge moléculaire. Ces activités se sont généralisées, suggérant que la théorie neutraliste est entrée dans les mœurs. Pour quelles raisons ?

L'argument du fardeau génétique s'est révélé invalide et ne peut servir de support à la théorie neutraliste. Cette théorie engendre toutefois un faisceau de prédictions très générales qui sont confirmées par les données. Les sélectionnistes n'arrivent aux mêmes prédictions qu'en invoquant une explication particulière pour chacune. Ils sont

aussi pris dans une incohérence : les uns cherchent à montrer que des hypothèses sélectionnistes font les mêmes prédictions que les neutralistes, les autres exhibent des contre-exemples sensés discrépantes ces prédictions. Même s'ils sont convaincants, ces contre-exemples peuvent être simplement les exceptions qui confirment la règle.

Beaucoup de biologistes actuels se situent entre les deux extrêmes : la théorie neutraliste serait valide dans ses grandes lignes, mais avec des exceptions. Elle s'appliquerait littéralement aux mutations silencieuses ou à l'ADN non codant, et en grande partie aux séquences codantes et aux protéines. Dans ces dernières, les phénomènes de sélection positive semblent plus importants que Kimura ne l'avait imaginé. Un effort de modération a été fait par une élève de Kimura, T. Ohta, qui a imaginé une théorie « quasi neutraliste », incorporant une dose de sélection faible. Sans forcément adhérer à son modèle, on peut facilement supposer que la majorité des molécules se comportent de manière neutraliste, mais que des événements de sélection sont susceptibles d'être détectés occasionnellement.

### *Le programme adaptationniste*

En 1979, S. J. Gould et R. Lewontin publièrent un article consacré à l'évolution qui commençait bizarrement par une description architecturale. Dans le dôme de la cathédrale San Marco à Venise, la disposition des mosaïques est si harmonieuse que l'architecture semble avoir été pensée pour elles. Il n'en est rien, cette architecture répond d'abord à des contraintes mécaniques, dont le décorateur a su ensuite tirer parti. Cette métaphore

illustre une erreur typique en évolution, qualifiée de *programme adaptationniste* : présupposer une sélection naturelle toute-puissante qui aurait optimisé la quasi-totalité des caractères. Nous serions alors dans le *mieux des mondes possibles*, selon l'expression du dogmatique Pangloss (personnage du conte de Voltaire, *Candide*). Cette présupposition stérilise la recherche en biologie. En effet, des hypothèses pertinentes autres que la sélection peuvent expliquer les caractères observés : la dérive génétique, l'héritage phylogénétique de caractères ancestraux ou les contraintes de développement.

Les raisonnements des sociobiologistes fournirent des cibles privilégiées à Gould et Lewontin. Un de ces raisonnements proposait que les sacrifices humains et le cannibalisme rituel chez les Aztèques étaient une adaptation au manque de protéines alimentaires. Il s'agit selon Gould et Lewontin d'une attitude panglossienne, invoquant la sélection naturelle là où des facteurs culturels expliquent les faits de manière plus pertinente. Même si les protéines humaines constituaient, pour les classes sociales restreintes qui les consommaient, un apport nutritif, elles ne sont pas la raison primordiale de l'apparition des sacrifices.

L'article de Gould et Lewontin fait maintenant partie des classiques : les évolutionnistes bannissent unanimement les raisonnements panglossiens, même si, dans la pratique, ils cèdent parfois à la tentation !

### *La théorie synthétique*

La génétique mendélienne fut acceptée vers 1930. Elle apporta à la théorie darwinienne de l'évolution la théorie fiable de l'hérédité qui lui faisait défaut. Cette synthèse

fut réalisée de manière indépendante par les inventeurs de la génétique des populations, Fisher, Haldane et Wright, et fut appelée, pour la première fois, *théorie synthétique* par Huxley en 1942. On la nomme également *néodarwinisme*. Elle impulsa immédiatement de nombreux travaux. En particulier, Dobzhansky entama des recherches, aujourd'hui classiques, sur l'évolution génétique des populations de drosophiles. Ford et Kettlewell entreprirent des travaux dits de « génétique écologique ». Kettlewell, par l'étude de la phalène du bouleau et du mélanisme industriel, fournit le premier exemple de l'action de la sélection *in natura*, sur un locus mendélien.

L'intégration de la théorie mendélienne a également modifié les conceptions portant sur la spéciation. Jusqu'alors, la spéciation était expliquée par des événements exceptionnels tels que des macromutations. Mayr montre qu'elle résulte d'une différenciation génétique normale entre les populations. L'espèce devient un « pool génétique », c'est-à-dire un ensemble d'individus capables d'échanger des gènes. Elle entre ainsi dans le cadre de la génétique des populations. Enfin, Simpson a fourni un apport similaire pour la paléontologie. Auparavant, les paléontologistes considéraient que l'évolution morphologique des fossiles traduisait une tendance intrinsèque à évoluer dans un certain sens (par exemple, augmenter de taille). Simpson soutient que les changements de morphologie observés sont dus à des changements génétiques expliqués par la sélection et la dérive selon les modèles de génétique des populations. Ainsi la synthèse touche-t-elle au milieu des années 1940 tous les domaines de la biologie évolutive.



## REMERCIEMENTS

Nous remercions pour leurs précieux commentaires sur tout ou partie du manuscrit : Anouk Barberousse, Thomas Bataillon, Claude Blanchet, Claire Calmet, Philippe Jarne, Thomas Lenormand, Marie-Laure Parmentier, Nicolle Samadi, Louis Thaler, Michel Veuille. Une mention spéciale à Marie-Laure Parmentier pour sa traque, sans complaisance, des moindres incohérences ainsi qu'à Anouk Barberousse et à Pascal Ludwig pour nous avoir patiemment aidés à structurer nos idées. Cet ouvrage doit beaucoup aux discussions passionnées au sein du SII (Séminaire interdisciplinaire Izotges).



## BIBLIOGRAPHIE GÉNÉRALE

### Textes historiques

- G. Cuvier, *Discours sur les révolutions de la surface du globe*, Paris, Bourgois, 1985 (1<sup>re</sup> éd. 1825).
- Ch. Darwin, *L'Origine des espèces*, Paris, GF-Flammarion, 1992 (1<sup>re</sup> éd. 1859).
- J. B. Lamarck, *Philosophie zoologique*, Paris, GF-Flammarion, 1994 (1<sup>re</sup> éd. 1809).
- E. Mayr, *Populations espèces, et évolution*, Paris, Hermann, 1974 (1<sup>re</sup> éd. 1970).
- M. Kimura, *Théorie neutraliste de l'évolution*, Paris, Flammarion (Nouvelle Bibliothèque Scientifique), 1983.

### Recueils de textes historiques

- J.-M. Drouin et C. Lenay (textes présentés par), *Théories de l'évolution : une anthologie*, Paris, Pocket, 1990.
- E. Mayr (textes présentés par), *Après Darwin : la biologie une science pas comme les autres*, Paris, Éd. Dunod, 2006.
- M. Ridley (textes présentés par), *Evolution*, Oxford, Oxford University Press, 1997.

## Ouvrages de vulgarisation

- C. Bousquet, *Darwin, l'homme qui osa*, Paris, Belin, 2009.
- R. Dawkins, *Qu'est-ce que l'évolution ? Le fleuve de la vie*, Paris, Hachette, 2005.
- J. Deutsch, *Le ver qui prenait l'escargot comme taxi et autres histoires naturelles*, Paris, Éditions du Seuil, 2007.
- S. J. Gould, *Le Pouce du panda*, Paris, Grasset, 1982.
- S. J. Gould, *La vie est belle*, Paris, Seuil (Science ouverte), 1991 (1<sup>re</sup> éd. 1989).
- J. Harder, *Alpha... Directions*, Arles, Actes Sud – L'An 2, 2009.
- R. Lewis, *Pourquoi j'ai mangé mon père*, Arles, Actes Sud, 1990.
- P. Picq, *L'Évolution expliquée à mes petits-enfants*, Paris, Éditions du Seuil, 2009.

## Essais

- R. Dawkins, *Le Gène égoïste*, Paris, Odile Jacob, 1996 (1<sup>re</sup> éd. 1976).
- R. Dawkins, *L'Horloger aveugle*, Paris, Robert Laffont (La fontaine des sciences), 1989 (1<sup>re</sup> éd. 1936).
- S. J. Gould, *La Structure de la théorie de l'évolution*, Paris, Gallimard, 2006 (1<sup>re</sup> éd. 1972).
- F. Jacob, *La Logique du vivant*, Paris, Gallimard, 1970.
- E. Mayr, *Histoire de la biologie : diversité, évolution et hérédité*, 1989 (1<sup>re</sup> éd. 1982).
- J. Monod, *Le Hasard et la Nécessité*, Paris, Seuil, 1970.

- M. Morange, *La Vie, l'évolution et l'histoire*, Odile Jacob, 2011.
- P. Tassy, *L'Arbre à remonter le temps*, Paris, Bourgois, 1991.
- O. Wilson Edward, *Sociobiology. The New Synthesis*, Londres, Press of Harvard University Press, Cambridge (Mass.), Belknap, 1975.

## Manuels

- P.-H Gouyon, J.-P Henry et J. Arnould, *Les Avatars du gène*, Paris, Belin (Regards sur la science), 1997.
- T. Heams, P. Huneman, G. Lecointre, Marc Silberstein (dir.), *Les Mondes darwiniens. L'évolution de l'évolution*. Paris, Syllepse, 2009.
- G. Lecointre, H. Le Guyader, *Classification phylogénétique du vivant*, Paris, Belin, 3<sup>e</sup> édition revue et augmentée, 2006.
- J. Maynard-Smith, *Evolutionary Genetics*, Oxford, Oxford University Press, 1989.
- M. Ridley, *Evolution*, Oxford, Blackwell, 1993.
- M. Solignac, G. Périquet, D. Anxolabéhère et C. Petit, *Génétique et évolution*, t. I et II, Paris, Hermann, 1995.
- F. Thomas, T. Lefevre, M. Raymond (dir.), *biologie évolutive* De Boeck Université, 2010.



## INDEX

Les pages correspondant à des entrées du glossaire critique apparaissent en gras.

- β-thalassémie, 144-145  
Acétylcholinestérase, 104-106  
Acides nucléiques (ADN, ARN), 11, 20, 37-45, 48-50, 53-58, 60, 66-68, 76, 81, 86-88, 91, 105-106, 120, 123, 125-126, 129, 132, 134-135, 149, 151, 160, 163, 165, 168, 169, 277, 286, 291, 323, 330, 333, 337  
Actualisme, 17  
Adaptation locale, 227, 229  
Additivité, 143-144  
ADN répété, satellite, égoïste, 50, 55-58, 81, 86, 88, 125, 127, 130, 134, 135, 139  
Allèle, 60-66, 69-70, 72-79, 83, 87-89, 91, 104, 106, 114-116, 123, 142-148, 150-153, 157, 160, 162, 165, 177, 190, 192-195, 197, 198-199, 204, 206, 210-211, 213, 224, 227, 229-230, 236, 246-247, 264-265, 267, 294, 302, 307, 320, 326-327, 334-335  
Allofécondation, 187-191, 195  
Altruisme, 127-128, 247-252, 270  
*Amœba dubia*, 54  
Amplification, 105  
Anémie falciforme, 42, 147-148, 336  
*Arabidopsis thaliana*, 113  
Autofécondation, 187-196, 246, 319  
Auto-incompatibilité, 187, 246-247, 327  
Auto-stop, 151-153  
Auxotrophe, 113, 154  
*Background selection*, 153  
BEADLE (George Wells), 112, 154  
Biométriciens, 199, 325-328  
Biosphère, 9  
*Bulinus truncatus*, voir Escargot  
*Cænorhabditis elegans*, 113, 220-221  
*Carcinus mœnas*, 109-110, 232-233  
Chloroplaste, 125-126, 133, 262  
Chromosome, 123, 125, 134, 136-138, 143, 149-154, 159-165, 170-173, 177, 189, 212, 213, 299

- Cline, 227-231  
 Clone, 97, 101-102, 131, 169, 178  
 Coalescence, 67-69, 73  
 Complexe majeur d'histocompatibilité (CMH), 266-268, 336  
 Consanguinité (dépression), 188, 190-192, 194-196, 319-320  
*Copia*, 56  
 Coût de la résistance, 108, 229-231  
 Coût des mâles, 175-177  
 Créationnisme, 14, 28, 33-34, 140  
*Crossing-over*, 150  
*Culex pipiens*, voir Moustique  
 CUVIER (Georges), 14  
 Cyanobactéries, 9, 122  
*Cyrtodiopsis dalmani*, 137, 257, 260  
 DARWIN (Charles), 15, 17, 20, 35, 77, 95-96, 99, 156, 207, 256-257, 274, 276, 323-324  
 Darwinisme social, 320-322  
 DAWKINS (R.), 21-22, 140  
 Dérive génétique, 8, 25-27, 29, 31-34, 60-93, 111, 114-117, 146, 153, 157, 160, 167, 177, 179, 183-185, 208-209, 223-224, 247, 287, 293-297, 300, 310, 327, 334, 338-339  
 Descendance avec modification, 14-18  
*Dictyostelium*, 122  
 Différenciation cellulaire, 121  
 Dinosaures, 19, 21, 308-309  
 Diploïde, 122-125, 142, 143, 146, 150, 162, 164, 173, 189, 210, 255  
*Dipodomys ordii*, 57  
 Dispersion, 240-244  
 Distorsion méiotique, 136-137  
 DOBZHANSKY (Theodosius), 33, 36, 339  
 Dodo, 21  
 Domestication, 96, 152, 203, 213, 315  
 Dominance, 145-146, 148, 165, 201, 210, 334, 336  
*Drosophila melanogaster*, 53, 56, 73, 282  
 Duplication, 48, 50-55, 105-106, 123, 132-134, 155  
 Épistasie, 154-157, 162, 178, 201, 211  
 Escargot (*Bulinus truncatus*), 81, 85-86, 109, 173, 187, 192, 197, 234, 238, 262, 304-307  
*Escherichia coli*, 40, 101, 169, 265  
 Essentialiste, 96, 275-276  
 Estérase, 105-106  
 Eucaryote, 120-122, 125-126, 133, 170-172, 174, 176, 262, 268  
*Euplectes progne*, 256  
 Extinction, 20, 87, 184-185, 238-244, 248, 273, 282, 300, 302, 306-309, 311, 317  
 Extinction de masse, 21, 308-309, 311, 317  
 Fécondation (voir également auto-fécondation, allofécondation), 65, 122-124, 128, 131, 141, 150, 170-171, 174, 187-196, 246, 319  
 FISHER (Ronald), 107, 156, 210, 257-258, 339

- Fitness* (voir Valeur sélective)  
Fixisme, 15  
Fonction, fonctionnel, 22, 34-35, 38-39, 42, 44, 49-56, 60, 112, 120-121, 125, 129, 132-135, 138, 144, 154-157, 160-162, 165, 177, 215-216, 218, 221-224, 226, 335
- Gamète, 122-124, 126, 128, 150, 167-168, 170-171, 174, 176, 252, 328
- Génétique biochimique, 111
- Génétique quantitative, patrimoine génétique, 11, 58, 198-199, 203-204, 206, 208, 211, 222-224, 315, 325, 327
- Génotype, 25-27, 29, 33, 109, 114-115, 142, 144, 147, 157, 163-164, 179-180, 194, 200, 211, 214, 225, 232-233, 243, 246, 270, 291, 302
- Geospiza fortis*, voir Pinsons de Darwin
- Gonochorie, 187
- Guppys, 257
- Gynodioïque, 138
- HAECKEL (Ernst), 214
- Haploïde, 122-125, 143, 255, 264
- HARDY (T.), 29-30, 65, 194
- Hémoglobine, 39, 42, 45, 51-53, 75-76, 133, 144, 147-148
- HENNIG (William), 280-281
- Héréditaire, hérédité, 8, 25, 35, 95, 126, 138-139, 197-199, 325-328, 338
- Héritabilité, 201-207
- Hermaphrodite, 138-139, 171, 187, 189, 192, 252
- Hétérosis, 117, 190, 315, 319-320
- Hétérozygote, 123, 136, 142-148, 189, 193-194, 320
- Homéotiques, 217-221, 223
- Homozygote, 136, 142-144, 146-148, 161, 189-190, 192-194, 320
- Horloge moléculaire, 45-48, 90-93, 266, 291, 330, 335, 336
- Hybridation, 164, 281, 284-285
- Hydrobia ulvae*, 262-263
- Hydrobia ventrosa*, 263
- Individu, 8, 25-27, 31-32, 35, 42, 50, 56-60, 64-69, 73-76, 80-91, 95-104, 110-116, 119-154, 159-207, 213, 220-221, 225, 227, 229, 231, 234-242, 246-254, 261-263, 270, 273-288, 293-302, 306-311, 313, 315, 318-319, 324, 328, 332, 335, 339
- Isolement prézygotique, postzygotique, 293-298
- Kin-selection* (sélection de parentèle), 253-254
- LAMARCK (Jean-Baptiste de), 18, 95, 328
- Létale (mutation), 113
- Liaison physique, 149
- Lin (*Linum utatissimum*), 263
- Littorina obtusata*, 109-110, 197, 232-237
- Loi binomiale, 70-71, 114
- LYELL (Charles), 17
- Lymphocytes, 129-131
- Maïs, 16, 56, 117, 152, 205-206, 213, 315, 319-320

- Maladaptation, 115, 165, 228, 231, 241, 244, 269
- Maladie génétique, 147
- MAYNARD-SMITH (John), 250, 321, 324
- MAYR (Ernst), 36, 157, 278-279, 282, 289, 339
- Méiose, 122-125, 128, 131, 136-138, 141, 149-150, 169-170, 173-175, 189, 203, 299
- Melampsora lini*, 263
- Mélanisme industriel, 339
- Melanoides tuberculata*, 173, 304-305
- MENDEL (Gregor), 35, 144, Mendélien, 197-199, 208, 210-211, 223-224, 325-328, 338-339
- Métabolisme, 10
- Métapopulation, 238-244
- Mitochondrie, 67, 69, 125-126, 133, 138, 262, 268
- Mitoses, 122-123, 173
- Moule (*Mytilus edulis*), 162-164, 226-230, 232
- Moustique (*Culex pipiens*), 104-109, 146, 228-232
- Multiplication végétative, 172
- Mus musculus*, voir Souris
- Mutation (délétère, favorable, létale, neutre, spontanée, ponctuelle, par fragment, synonyme), 8, 25, 26, 30, 37, 40-53, 55, 56, 58-60, 66-67, 72-73, 77-84, 87-92, 100-117, 127-130, 133-134, 144-145, 147, 151-156, 161, 168, 176-180, 190, 193, 204-209, 214-223, 247, 287, 299-302, 327-329, 334-339
- Mutualisme, 133, 261-262, 268, 270-271
- Mytilus edulis*, voir Moule
- Myxomatose, 248-249, 265
- Neurospora*, 112-113, 154
- Neutralistes, 77, 148, 333-337
- Oiseaux de paradis, 257
- Ontogenèse, 214-215, 217
- Organique (matière), 10, 20, 333
- Paludisme, 16, 42, 147-148
- Parasites, parasitisme, 47, 57, 134-135, 147-148, 186, 261-265, 268-271
- Parthénogénèse (apomictique, cyclique, gynogenèse), 172-176, 182, 305
- PASTEUR (Louis), 18, 333
- Perturbation, 238, 243, 245, 308
- Phalène du bouleau (*Biston betularia*), 16, 339
- Phénotype, 25-26, 107, 142, 200-206, 209-211, 214, 232, 233, 243, 259
- Phylogénétique (arbre), 18-19, 44, 274, 280, 284, 310, 330
- Phylogénie, 33, 44, 46, 215, 262, 291, 329-330
- Physe (*Physa acuta*), 192
- Pinsons de Darwin (*Geospiza fortis*), 207
- Pinus resinosa*, 54
- Pisum sativum*, 144
- Plantain (*Plasmodium falciparum*), 240
- Plasticité phénotypique, 232-234, 236-238, 243, 305
- Poeciliopsis*, 173-174

- Polymorphisme, 59, 76-78, 80, 148, 152, 247, 270, 334, 336-337  
Polyplioïdie, 178  
Prédation, 110, 234, 247, 250, 256, 261, 263, 270, 307, 314, 324  
Prokaryotes, 121-122, 262  
Processus de Markov, 71  
Pseudogène, 51, 55, 125, 133, 155  
*Quantitative Trait Loci* (QTL), 211-214, 223  
Radiation cambrienne, 216-221  
Récessivité, 145-146, 210  
Recombinaison, 48, 50, 53, 57, 134, 149-154, 160-162, 169-170, 172, 175, 177, 179, 203, 212  
« Reine rouge », 269-270  
RéPLICATION, 20, 23, 25, 40, 49, 57, 66, 134-136, 168  
Reproduction, 11-13, 15, 20, 22, 25, 34, 51, 99-101, 123, 130-132, 141, 150, 167-177, 179, 181-189, 191-192, 194-196, 201-203, 230-231, 238, 246-248, 252-253, 255, 258, 269, 278-285, 287, 304-306, 314  
Résistance (aux insecticides, aux antibiotiques), 16, 42, 104-108, 141, 146, 148, 173, 197, 226, 228-231, 265-266, 302, 306, 314, 316, 328  
Révolution génétique, 157  
Rotifères bdelloïdes, 182  
*Runaway process*, 258-259  
Salinité, 226-228, 232, 245  
Sélection artificielle, 95-97, 102, 203-206, 224, 315, 318  
Sélection directionnelle, 209, 247, 315  
Sélection d'espèces, 183-184, 311, 332  
Sélection de fond (voir *Background selection*)  
Sélection de groupe, 248-249, 321  
Sélection fréquence-dépendante, 246-247, 270, 334  
Sélection naturelle, 8, 25-27, 30-32, 34-34, 38, 42, 44, 49, 51, 58, 61, 63, 76-77, 82, 92, 95-117, 119, 126-131, 133-135, 137, 139, 141, 145-146, 152, 157-159, 164-165, 167, 175-176, 183, 186, 188, 206, 208-209, 223, 227-228, 232, 238, 246-248, 263, 267, 287, 295, 302, 310, 321, 325-328, 334, 338  
Sélection purifiante, 110-114, 152-155, 161-162, 334  
Sélection sexuelle, 258  
Sélection stabilisante, 147, 209, 334, 336  
Sélectionnistes, 77, 148-149, 333-337  
*Sex-appeal*, 257  
Sex-ratio, 176  
Sexe, 167-196, 285  
*Shifting balance theory*, 157  
Soma, lignée somatique, 126-128  
Souris (*Mus musculus*), 290-293  
Sous-espèces, 285, 289-293, 296, 299

- Spéciation allopatrique, 287-294, 297-298, 300, 310  
 Structure des populations, 82-90, 336  
 Substitution, taux de substitution, 40-41, 45, 47-48, 266-267, 334-336  
 Superdominance, 146, 148, 165, 334, 336  
 Surexpression, 105  
 Symbiose, 133, 261-262, 268, 270  
 Systématique, 33, 277, 280, 322-324  
*Tanysiptera*, 289  
 TATUM (Edward), 112, 154, 169, 179  
 Téléconomie, 22  
 Téosinte, 152, 213, 314  
 Théorie de la récapitulation, 214, 215  
 Théorie des jeux, 250-251, 321, 324-325  
 Théorie du handicap, 259  
 Théorie endosymbiotique, 133  
*Thymus vulgaris*, 138, 268  
 Transposons (éléments mobiles), 56-57, 132, 135  
 Trilobites, 14, 19, 303  
 Valeur reproductive (*breeding value*), 203, 210  
 Valeur sélective (*fitness*), 26-27, 31-33, 99-104, 106-108, 111-116, 142-143, 145-146, 148, 154, 157-158, 162, 175, 179, 190-192, 196, 221, 225, 229, 236, 242, 254, 256, 292-294, 319-320  
*Veronica spicata*, 240  
 Vigueur hybride, 117, 319  
 Vitaliste, 10  
 WEINBERG (H.), 29-30, 65, 194  
 WRIGHT (Sewall), 71, 83, 84, 86, 156-159, 339  
*Zea mays*, 152, 205, 213  
 Zooxanthelles, 261

## TABLE

<i>Avant-propos .....</i>	7
1. Les principes fondateurs de la théorie de l'évolution .....	9
2. Les origines moléculaires de la variation génétique.....	37
3. Le tri entre les variants génétiques par le hasard .....	59
4. La sélection naturelle .....	95
5. L'individu, une société de gènes en harmonie ? .....	119
6. Un gène n'évolue pas seul .....	141
7. Le sexe, stratégie de gestion de la variabilité génétique.....	167
8. Des gènes aux caractères .....	197
9. Quand les règles du jeu changent .....	225
10. L'environnement, c'est avant tout les autres.....	245
11. La macro-évolution, des individus aux espèces .	273
<i>Épilogue .....</i>	313

## Glossaire critique

Notions .....	315
<i>Amélioration génétique des espèces, 315 ; La biodiversité et sa préservation, 316 ; Génétique et eugénisme, 318 ; Hétérosis et dépression de consanguinité, 318</i>	

<i>nité</i> , 319 ; <i>Sociobiologie et darwinisme social</i> , 320 ; <i>La systématique</i> , 322 ; <i>La théorie des jeux</i> , 324.	
Acteurs et controverses.....	325
<i>Biométriciens et mendéliens</i> , 325 ; <i>Caractères acquis ou mutations aléatoires ?</i> , 328 ; <i>Cladisme et phénéticisme</i> , 329 ; <i>Équilibres ponctués</i> , 331 ; <i>La génération spontanée</i> , 332 ; <i>Neutralistes contre sélectionnistes</i> , 333 ; <i>Le programme adaptationniste</i> , 337 ; <i>La théorie synthétique</i> , 338.	
<i>Remerciements</i> .....	341
<i>Bibliographie générale</i> .....	343
<i>Index</i> .....	347

Mise en page par Meta-systems - 59100 Roubaix

N° d'édition : L.01EHQN000588.N001

Dépôt légal : octobre 2011

Imprimé en Espagne par Novoprint S. A.

# PATRICE DAVID SARAH SAMADI

## La théorie de l'évolution

À la fois récit historique et explication du monde vivant, la théorie de l'évolution répond à des questions fondamentales et transforme en permanence notre vision du monde. Qu'est-ce que les êtres vivants ont en commun ? Pourquoi ont-ils l'air à la fois bien conçus et imparfaits ? Pourquoi le sexe est-il le mode de reproduction le plus agréable mais le moins efficace ? Comment comprendre l'apparition incessante de nouveaux virus, et la perte d'efficacité de nos médicaments ? Quelles guerres silencieuses ont lieu dans notre ADN ?

Dans ce livre, Patrice David et Sarah Samadi initient le lecteur aux fondements logiques de la théorie de l'évolution. Ils montrent comment les arguments nés avec Darwin ont accompagné et englobé les progrès ultérieurs de la génétique, de la biologie moléculaire, de l'embryologie et de l'écologie, au point qu'ils constituent aujourd'hui le socle conceptuel de toute la biologie.

Ancien élève de l'École normale supérieure, docteur en biologie, **Patrice David** est directeur de recherche au CNRS et enseigne la biologie des populations et la génétique évolutive à l'université de Montpellier-II. Ancienne élève de l'École normale supérieure, docteur en biologie, **Sarah Samadi** est chargée de recherche à l'Institut de recherche pour le développement dans un laboratoire associé au Muséum national d'histoire naturelle.

**Édition revue et augmentée 2011.**

Prix France : 10 €  
ISBN : 978-2-0812-6597-4

En couverture: Image conceptuelle  
sur la recherche en génétique.  
© SPL/PHANIE.



9 782081 265974

**Flammarion**