

PRÉFACE

L'étude de la diversité du vivant a d'abord été l'apanage de la systématique. Cette science qui voit le jour au XVIIIème siècle se donne pour but d'organiser l'apparent fouillis de la biodiversité. Le « Système de la Nature » publié par Carl Linné en 1735 lance un mouvement d'ensemble des naturalistes en postulant que la vraie diversité est celle des espèces, qui ont été créées dès l'origine. Les variations entre individus sont des "caprices de la nature".

Dans la suite, les naturalistes vont progressivement se convaincre qu'il existe un ordre interne à la nature et tenter d'établir une classification naturelle. Lorsque Charles Darwin publie « L'Origine des Espèces » en 1859, il affirme que toutes les formes vivantes étant issues d'une ou quelques formes originelles, les lignées qui divergent partagent d'autant plus de caractères qu'elles se sont séparées depuis moins longtemps. Il en déduit qu'inconsciemment, les naturalistes ont en réalité cherché à retrouver ces liens de parenté et que la classification du vivant doit se faire en fonction de la phylogénie. Son vœu ne sera réalisé pleinement qu'à la fin du XXème siècle.

Darwin insiste aussi sur l'arbitraire qui préside à l'établissement des niveaux hiérarchiques de la systématique. Il affirme que les différences entre espèces, entre sous-espèces, entre variétés, entre individus "se fondent l'une dans l'autre constituant une véritable série". Les lignées existent bien et sont réellement différenciées mais le statut que le naturaliste décide de leur accorder est nécessairement arbitraire et peut toujours être discuté. Il prédit donc que lorsque sa théorie sera comprise, les systématiciens cesseront leurs interminables débats sur ce qu'est une espèce et sur le statut à accorder à tel ou tel taxon. Il se rend cependant vite compte que la toute puissance de l'idée d'espèce constitue un obstacle majeur à la compréhension de ses idées et il écrit à un ami dès l'année suivante « je désespère souvent de voir la majorité des naturalistes simplement le comprendre, des gens intelligents qui ne sont pas des naturalistes et qui n'ont pas une idée bigote du terme espèce montrent plus de clarté d'esprit ». Aujourd'hui, l'espèce reste, dans l'imaginaire et dans l'expression des biologistes, un niveau prépondérant à tel point qu'on continue, contre toute logique, à séparer la biodiversité "entre espèces" de la diversité "dans l'espèce". Il est clair que, dans ce domaine, la génomique environnementale fera nécessairement bouger les lignes.

Darwin prédisait encore l'émergence d'un nouveau domaine scientifique, étudiant l'hérédité des variations entre individus. La génétique qui naît au début du XXème siècle exauce assez rapidement ses vœux. La génétique des populations, issue au cours des années 1910-1930 se constitue sur des bases théoriques réunissant darwinisme et mendélisme. Des modèles théoriques divers fleurissent ainsi que des expérimentations en cages à populations de mouches mais, dans la nature, les études restent rares. En effet, trouver un polymorphisme génétique dont on puisse repérer les différents génotypes dans

une population naturelle demande un coup de chance. La plupart des variations sont multigéniques et influencées par l'environnement. On reste donc confiné à des variations de couleur (comme la célèbre couleur noire de certains papillons) ou de composés secondaires. C'est dans ce contexte qu'en 1966, Richard Lewontin a l'idée, avec Jack Hubby, de chercher la variation génétique au plus près du génome. Il développe donc des méthodes biochimiques permettant de repérer les variations des protéines et rapidement on trouvera des méthodes pour observer les variations de l'ADN lui-même. Tout un chacun peut enfin trouver des variations génétiques identifiables dans son espèce préférée.

Bien sûr, on dispose maintenant de gènes aux allèles identifiables mais on n'a aucune idée de ce que ces variations génétiques induisent en termes de différences de fonctionnement de l'organisme. La théorie neutraliste de Motoo Kimura apportera une réponse : elles sont neutres et seule la dérive aléatoire préside à leur devenir. Si ces gènes ne présentent pas d'intérêt en eux-mêmes, ils constituent en revanche des marqueurs de l'histoire des populations, des lignées. Ils permettent de connaître les échanges génétiques, les migrations, l'isolement. Alors que la génétique s'était préoccupée de comprendre les processus (*process*), les marqueurs permettent de connaître le pattern. Une fraction importante de la communauté scientifique concernée change alors de sujet de recherche sous l'influence de cette technique nouvelle. La phylogénie moléculaire se développe et renforce la systématique phylogénétique.

Dans un premier temps, l'étude de la diversité des génomes a donc détourné les chercheurs de l'étude du *process* et les a orientés vers l'étude du *pattern*. La génomique environnementale actuelle, en fournissant chaque jour des milliards de bits d'information, permet d'interroger tous les champs de l'écologie et des sciences de l'évolution: de continuer l'étude des patterns bien sûr mais aussi de reprendre l'étude des processus.

Les textes qui suivent montrent l'extraordinaire pouvoir de ces techniques. La finesse des données permet de détecter des phénomènes de sélection agissant sur une partie du génome, et influençant de ce fait le devenir de séquences voisines neutres. Pour trouver les gènes responsables d'un phénotype particulier, ou d'une interaction écologique particulière, la transcriptomique peut être un outil d'une puissance remarquable... Bien sûr, il est probable que, parfois dépassés par la puissance de l'outil dont ils disposent, les chercheurs n'utilisent un peu facilement un marteau pilon pour écraser un moustique. On peut encore une fois craindre que la facilité l'emporte et que plutôt que de rechercher l'originalité, on se contente d'accumuler des données. Mais globalement, la génomique devient un outil indispensable pour le développement des études d'écologie et d'évolution.

En particulier, la métagénomique, qui nous fournit non plus un génome particulier mais l'ensemble des génomes présents dans un sol, un fromage ou un intestin par exemple permet de découvrir et d'étudier d'innombrables organismes vivants

inconnus jusqu'alors et dont nous ignorons tout, y compris la forme. Gageons que ces nouvelles méthodes modifieront notre vision de la diversité du vivant et que les différents niveaux de la systématique devront être révisés pour décrire ces formes vivantes dont nous ne connaissons que le génome. Un concept nouveau dont on verra la puissance dans cet ouvrage a émergé : celui d'OTU (*operational taxonomic unit*). Le vieux rêve de Darwin de voir émerger une vision du vivant qui soit dynamique et qui prenne en compte la diversité dans son ensemble sans s'embarrasser de savoir s'il s'agit de variétés, d'espèces ou de genres pourrait bien advenir sous la pression de cette nouvelle démarche.

Gageons aussi que, devant l'afflux de ces *big data*, la nécessité absolue de développer de nouveaux outils théoriques redonnera aux chercheurs l'envie de reprendre à la fois une recherche plus conceptuelle et, sur une base renouvelée, une approche naturaliste. Depuis une cinquantaine d'années, la biologie a été fascinée par les outils qu'elle produisait. Les démarches théoriques, moins rentables à court terme, ont perdu de leur importance. Il devient clair que pour interpréter toutes les informations accumulées par la génomique, de nouveaux concepts devront émerger, issues de l'interaction entre recherches naturalistes, génomiques et théoriques. Une démarche réductionniste tentant de regarder chaque élément de l'ensemble devra se compléter d'une démarche globale où l'ensemble est étudié dans toute sa complexité. C'est la suite qu'on peut attendre de cette fascinante aventure et cet ouvrage montre de nombreuses pistes qui s'ouvrent à nous et qui à coup sûr, inspireront les biologistes futurs.