

longtemps. Une multitude de génomes entiers de bactéries et d'archées ont été séquencés. Quand on a comparé le génome d'une espèce bactérienne à celui de nombreux clones provenant de celle-ci, on s'est aperçu qu'il existe beaucoup plus de gènes au niveau de l'espèce qu'au niveau de l'organisme. Tout se passe comme si une bactérie comportait d'une part des gènes essentiels – des gènes informationnels clés pour la réplication, la transcription et la traduction – et, d'autre part, des gènes qui changent d'un individu à l'autre, agissant plus sur le métabolisme et la pathogénicité.

Ainsi est apparue une notion ahurissante, qui n'existe pas chez les eucaryotes *a priori* : le pangénome – l'ensemble de ces gènes changeants (voir l'article de Bianca Sclavi et John Herrick, pages 34 à 41). Cette notion contient celle de transfert horizontal, c'est-à-dire d'échanges de gènes entre organismes. L'idée existait déjà pour les chloroplastes et les mitochondries, mais tout à coup, on se rendait compte que ce phénomène se produit peut-être en permanence. Comment se font ces transferts horizontaux ? Via les virus, probablement, ou au moins en grande partie. Les virus sont ainsi de retour dans l'histoire du vivant...

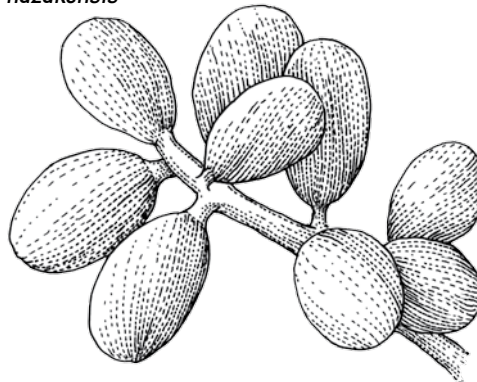
La notion de transfert horizontal a amené à penser le vivant en termes non plus de phylogénie, mais de réseau. On a commencé à comprendre pourquoi trouver une racine à l'arbre du vivant de Woese était impossible : on imaginait que les gènes étaient obtenus par descendance, par transfert vertical, alors que sans doute beaucoup provenaient de transferts horizontaux. L'arbre du vivant cachait en fait une forêt.

PLS

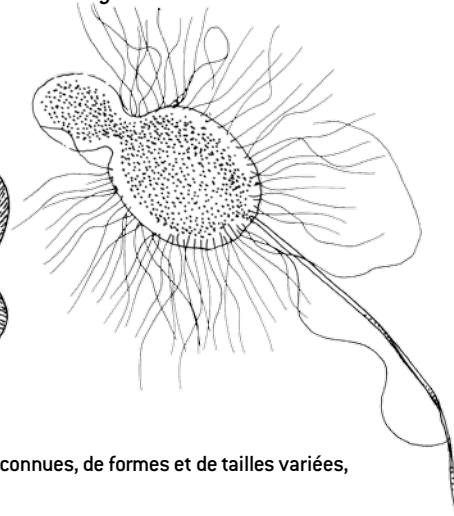
Si l'arbre du vivant est un réseau, peut-on encore l'étudier ?

H. L. G. : Il y a deux écoles. La première considère qu'il est toujours possible de faire une phylogénie des archées et des bactéries : l'idée est que si l'on est capable de repérer les gènes transférés horizontalement, il suffit de les éliminer et de ne garder que les gènes informationnels, qui ont beaucoup plus de chance d'être transmis verticalement. C'est alors ceux-là que l'on utilise pour faire une phylogénie, sans être brouillé par les gènes transmis horizontalement. C'est ce que font par exemple Céline Brochier-Armanet à l'université Lyon-1 ou Simonetta Gribaldo à l'institut Pasteur.

Thermosporothrix hazakensis



Planctomyces maris



LE GROUPE DES BACTÉRIES compte 16 760 espèces connues, de formes et de tailles variées, et de voies métaboliques très diversifiées.

L'autre école est celle de Bill Martin, de l'université de Düsseldorf, en Allemagne, que résume bien le titre d'un article qu'il a publié en 2006, *The tree of one percent* (L'arbre de 1%). Si vous enlevez tous les gènes que vous pensez transmis horizontalement, vous ne gardez que 1 % du génome et vous construisez votre phylogénie avec 1 % des gènes. Est-ce vraiment raisonnable ? L'idée est de dire que les phylogénies ne tiennent pas la route. Les arbres que l'on construit sur les archées ou les bactéries servent à faire une classification, mais pas une phylogénie. Ils ne reflètent pas l'histoire puisque celle-ci est plus compliquée. Il faut utiliser les réseaux.

PLS

Concrètement, comment étudie-t-on les réseaux ?

H. L. G. : Ce champ d'étude mêle trois domaines : la biologie moléculaire pour collecter des génomes à comparer, la bio-informatique pour la gestion des données et les mathématiques pour l'interprétation. Il a explosé avec le séquençage à haut débit. La biologie moléculaire a engrangé beaucoup de séquences de gènes et d'ARN à la fin des années 1980 et au début des années 1990, mais à l'époque, tout était encore fait à la main. Puis, dans les années 1990, on a commencé à avoir des génomes entiers, notamment le génome humain à la fin des années 1990 et au début des années 2000.

Aujourd'hui, obtenir le génome d'une bactérie ne coûte que 1 000 euros. D'énormes quantités de séquences sont ainsi engrangées. Dans notre

laboratoire, par exemple, Philippe Lopez et Éric Baptiste traitent ainsi des millions de séquences. Il y a quelques années, le goulet d'étranglement était l'obtention des séquences, maintenant c'est le traitement de l'information, tant son volume est grand.

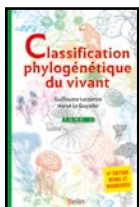
L'interprétation des données, quant à elle, n'est pas plus simple. Elle fait appel à l'étude des réseaux d'un point de vue mathématique, notamment à la théorie des graphes. Il s'agit d'un domaine extrêmement ardu des mathématiques, aussi les biologistes travaillent-ils avec des mathématiciens purs. Le domaine n'avance pas vite, mais il rassemble une communauté de plus en plus importante de chercheurs, notamment canadiens, américains, irlandais, anglais et allemands.

PLS

Qu'a apporté cette notion de réseau ?

H. L. G. : Elle a montré que la vie est beaucoup plus complexe qu'on ne le pensait. Le point clé est que les transferts horizontaux continuent toujours à se produire. Cela change complètement la vision des microorganismes. Si vous imaginez des archées dans les sédiments marins ou ailleurs, il faut les penser avec des échanges de gènes. Les chercheurs essaient ainsi de tirer des enseignements de ces échanges actuels pour reconstituer une histoire des bactéries et des archées. L'idée est la suivante : si les transferts horizontaux sont aussi importants aujourd'hui, ils l'étaient certainement davantage auparavant. Cette idée a permis de clarifier notre compréhension des eucaryotes.

■ POUR EN SAVOIR PLUS



G. Lecointre et H. Le Guyader, **Classification phylogénétique du vivant**, Belin, 2016.

Entièrement revue, cette

quatrième édition apporte la vision que l'on a actuellement de l'ensemble du vivant. Elle fait notamment la part belle aux microorganismes et aux débats qu'ils suscitent depuis l'avènement de l'ère génomique.

Maintenant que l'on regarde les génomes entiers d'eucaryotes, d'archées et de bactéries, on comprend que les génomes eucaryotes comportent des gènes provenant d'archées, d'autres de bactéries. Les gènes de type archée sont surtout importants pour la réplication, la transcription et la traduction, et proviennent d'un groupe particulier d'archées, le superphylum TACK. On découvre aussi que des gènes impliqués dans la dynamique du squelette des cellules eucaryotes sont également chez les archées. L'exploration de ces gènes provenant d'archées et de bactéries a conduit à l'idée que l'eucaryote serait une fusion d'au moins une archée et une bactérie, lesquelles ont ensuite formé un organisme qui a été capable d'accepter une mitochondrie et un chloroplaste. Une chimère avant la chimère, en somme.

De plus, on s'aperçoit que les gènes des archées et des bactéries sont très différents. L'idée actuelle est donc que les phylums de ces deux domaines sont sortis d'une soupe de progénètes de façon indépendante et qu'ils n'ont jamais eu d'ancêtre commun. Ils se sont individualisés dans ce monde où il y avait déjà des transferts horizontaux, puis les eucaryotes ont émergé d'une chimère d'archées et de bactéries, sans doute plus proche des archées que des bactéries puisqu'ils comportent plus de gènes d'archées. C'est l'idée de Bill Martin, Ford Doolittle, Philippe Lopez et Éric Bapteste, par exemple.

PLS

La notion de réseau aide-t-elle à dater ces événements ?

H. L. G. : Le réseau des premiers microorganismes, les échanges de gènes, apportent des possibilités statistiques fantastiques qui pourraient expliquer une découverte récente. Il y a quelques mois, une équipe australienne a mis en évidence des traces de vie remontant à 3,7 milliards d'années, alors que les plus anciennes connues jusqu'alors dataient de 3,2 milliards d'années. Il y a 3,7 milliards d'années, cela ne faisait *grosso modo* que quelques centaines de millions d'années que la Terre était devenue habitable. À l'échelle géologique, tout se passe comme si la vie est apparue dès que les conditions physicochimiques d'apparition de la vie ont été réunies sur Terre. Les échanges de gènes pourraient expliquer une telle rapidité.

PLS

La notion de réseau a-t-elle aussi modifié notre vision des organismes macroscopiques et de leur évolution ?

H. L. G. : En effet, diverses études visent à pousser la notion de réseau microbien jusqu'au bout, à appliquer l'outil réseau à d'autres échelles. Cet outil est déjà utilisé par les écologues pour étudier les écosystèmes. Éric Bapteste, par exemple, cherche à l'utiliser pour relier des niveaux d'organisation différents – des groupes de gènes et des organismes, des groupes de gènes et des communautés, etc. – et pour essayer de voir comment ces interactions et ces échanges se produisent dans le vivant.

PLS

Les études récentes sur le microbiote, l'ensemble des microorganismes d'un environnement, par exemple l'intestin, descendent-elles de cette idée de réseau ?

H. L. G. : Au départ, étudier le microbiote consistait à déterminer quels organismes le composaient : on séquençait les ARN de l'échantillon et on déterminait les organismes associés. C'est le principe de la métagénomique : vous séquencez le génome total d'un échantillon – une poignée de terre, une goutte d'eau de mer, un extrait de salive – et vous l'examinez de façon exhaustive. Cela a permis de mettre en évidence des organismes que l'on ne connaissait pas, car on ne pouvait pas les cultiver. On s'est ainsi aperçu que l'on n'arrivait à cultiver qu'un très faible pourcentage des bactéries qui existent dans la nature.

Mais on se rend compte à présent qu'il existe des échanges de gènes au sein des microbiotes, et qu'il faut les voir en réseau. Il ne s'agit pas juste d'une juxtaposition de bactéries, mais d'une communauté dont les membres interagissent constamment. D'où l'idée que si cette communauté n'est pas « stable », des dérèglements peuvent survenir, qu'il s'agisse d'obésité (voir l'article de Joël Doré et ses collègues, pages 54 à 59) dans le cas du microbiote intestinal ou de maladies des plantes (voir l'article de Stéphane Uroz et ses collègues, pages 74 à 79). ■

Propos recueillis par Marie-Neige CORDONNIER