

aux bactéries, qui intervenaient en particulier dans la transcription et la réplication de l'ADN.

Trouver la racine s'est cependant révélé impossible. Chez les eucaryotes, un gène était plus proche des archées, un autre des bactéries. Les discussions n'en finissaient pas.

PLS

A-t-on invoqué les virus comme groupe extérieur possible ?

H. L. G. : Absolument pas à ce niveau-là. On les considérait comme différents. Les virus n'ont pas d'ARN ribosomique, donc on ne pouvait pas les ajouter à l'arbre. On construisait séparément des groupes de virus à l'aide des séquences

virales que l'on connaissait. À cette époque, quasi personne ne pensait que les virus pouvaient être un groupe extérieur ou un vestige d'un groupe extérieur disparu, comme certains le proposent aujourd'hui (voir les articles de Patrick Forterre, pages 42 à 49, et de Chantal Abergel et Jean-Michel Claverie, pages 50 à 53). De toute façon, l'information contenue dans les virus alors connus n'était pas assez importante pour que l'on puisse en déduire un groupe extérieur.

PLS

Qu'a donné la recherche de la racine ?

H. L. G. : Deux choses sont apparues. D'une part, l'idée que si on trouve une racine à cet arbre, il

existe un dernier ancêtre commun à tous ces groupes, LUCA (*Last Universal Common Ancestor*), nom proposé par les biologistes grecs Christos Ouzounis et Nikos Kyrpides en 1994. D'autre part, en 1998, Woese a avancé le concept de progénote : l'idée que les trois groupes proviendraient non pas d'un ancêtre commun unique, mais d'un ensemble de précellules, des cellules qui n'ont pas les caractéristiques des cellules actuelles, procaryotes ou eucaryotes, mais qui ont des génomes très labiles et qui échangent énormément de gènes. Les trois phylums archées, bactéries, eucaryotes auraient pris source dans cette soupe de progénotes. En d'autres termes, LUCA ne serait pas un organisme précis, mais quelque chose de très diffus.

PLS

Carl Woese pensait-il ici déjà à l'idée que les chloroplastes et les mitochondries puissent provenir de bactéries ?

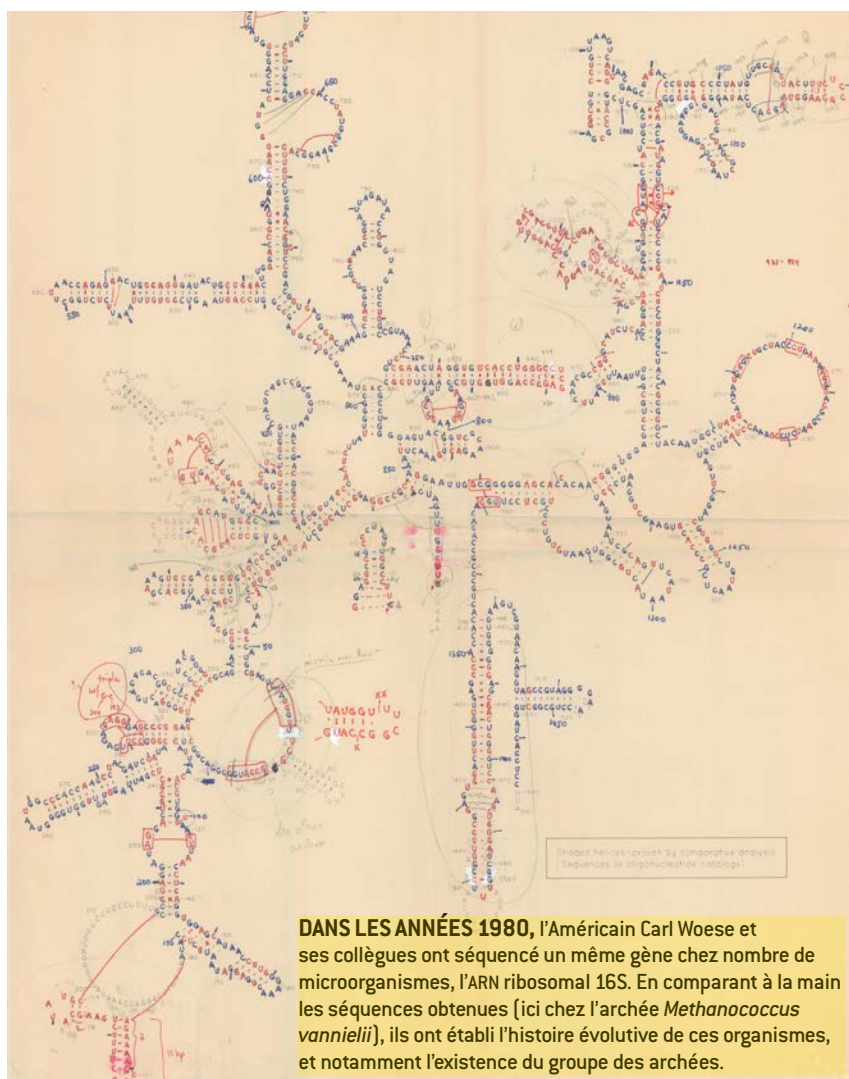
H. L. G. : Déjà au XIX^e siècle, des microbiologistes y pensaient. Mais Woese l'a montré avec brio. En partant d'ARN ribosomiques de maïs issus de la mitochondrie, du chloroplaste et du cytoplasme des cellules (donc purement eucaryote), il a montré que dans sa phylogénie, l'ARN ribosomique mitochondrial se plaçait dans les alpha-protéobactéries et l'ARN ribosomique des chloroplastes dans les cyanobactéries. Cela signifiait que les mitochondries et les chloroplastes étaient des vestiges de bactéries anciennes. Woese a donc fait là une découverte fantastique : la cellule eucaryote est chimérique.

PLS

Que s'est-il passé ensuite ?

H. L. G. : Des biologistes moléculaires tels que Ford Doolittle, de l'université Dalhousie, au Canada, se sont aussi aperçus, en comparant les séquences des génomes des chloroplastes et des mitochondries avec celles de leurs bactéries sources et des cellules eucaryotes qui les hébergent, qu'il y a aussi eu exportation, vers le noyau des cellules, de nombre de gènes des chloroplastes et des mitochondries. Tout d'un coup, la vision s'est élargie : le génome eucaryote était encore plus chimérique qu'on ne le pensait !

À ce moment-là, les biologistes ont eu l'impression d'avoir résolu un problème, mais pas pour



DANS LES ANNÉES 1980, l'Américain Carl Woese et ses collègues ont séquencé un même gène chez nombre de microorganismes, l'ARN ribosomal 16S. En comparant à la main les séquences obtenues (ici chez l'archée *Methanococcus vannielii*), ils ont établi l'histoire évolutive de ces organismes, et notamment l'existence du groupe des archées.

www.ma.cctb.uleuax.edu/ J. Camone et al. BioMed Central Bioinformatics 3:2 (2002)

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...

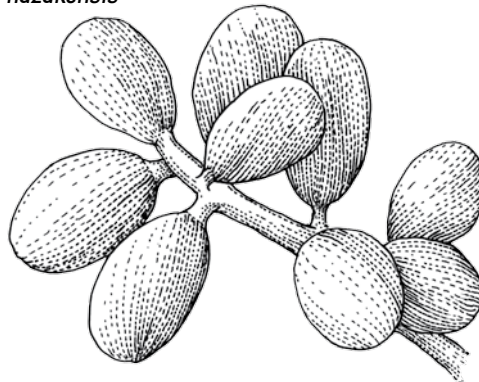
La notion de transfert horizontal a amené à penser le vivant en termes non plus de phylogénie, mais de réseau. On a commencé à comprendre pourquoi trouver une racine à l'arbre du vivant de Woese était impossible : on imaginait que les gènes étaient obtenus par descendance, par transfert vertical, alors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

PLS

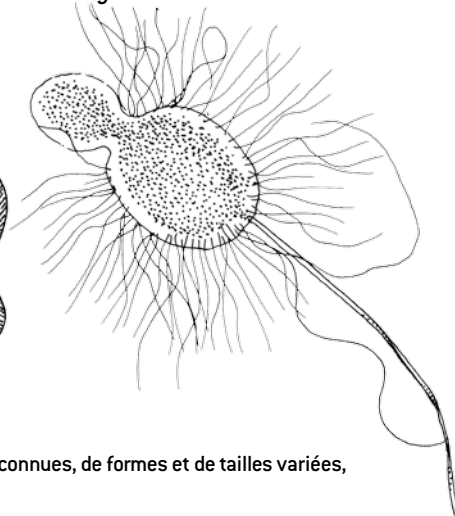
Si l'arbre du vivant est un réseau, peut-on encore l'étudier ?

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.