

médicale, Martinus Beijerinck en microbiologie agronomique (étude des maladies des plantes, des bactéries du sol, etc.) – se fichaient de l'évolution. Ils avaient besoin d'une classification utilitaire pour leurs disciplines. Mais l'idée d'une classification naturelle restait dans certaines têtes et l'on cherchait comment y parvenir.

Des éléments clés sont arrivés après la Seconde Guerre mondiale, vers les années 1950, en particulier quand on a précisé le concept de procaryote. Les procaryotes étaient bien séparés des eucaryotes, mais à un moment donné, on a considéré que les virus en faisaient partie.

PLS

En quoi cette idée fut-elle un élément clé ?

H. L. G. : Ce n'est pas tant cette idée que la réflexion qu'elle a suscitée. On pensait alors qu'il existait une sorte de continuum entre virus et bactéries. La différence entre eux portait surtout sur la taille, les virus traversant les filtres à bactéries tel le filtre Chamberland, mis au point dans le laboratoire de Pasteur par Charles Chamberland. Mais en 1957, dans un article passionnant, le microbiologiste français André Lwoff a défendu l'idée que « les virus devaient être traités comme des virus parce que les virus sont des virus ». En d'autres termes, le cycle du virus est tellement à part, avec notamment une phase parasitaire obligatoire, qu'il ne peut y avoir de forme de transition entre bactéries et virus. Pour Lwoff, l'élément déterminant n'était pas la taille, mais l'autonomie.

Les microbiologistes se sont ainsi retrouvés avec trois concepts biologiques définis de manière tranchée : les eucaryotes (qui apparaissent bien plus complexes que les autres au microscope électronique tout juste inventé), les procaryotes et les virus. La classification utilitaire a continué sur cette base, puis s'est produit le vrai déclic : l'avènement de la biologie moléculaire.

PLS

Les débuts du séquençage du matériel génétique ?

H. L. G. : Notamment, mais aussi celui des acides aminés qui composent les protéines. Très vite, certains, comme le microbiologiste allemand Ernst Pringsheim, ont pensé que la biologie moléculaire naissante allait peut-être fournir une classification

phylogénétique. De premiers essais, à partir des séquences de certaines protéines communes à différentes branches du vivant, ont cependant été peu concluants. La solution est venue en particulier du séquençage des ARN ribosomiques et d'un homme, Carl Woese.

Au début des années 1960, ce microbiologiste américain s'est penché sur le problème de l'origine du code génétique. Il s'intéressait alors à la machinerie moléculaire de la traduction des gènes en protéines et notamment aux ARN ribosomiques (des éléments constitutifs des ribosomes – les organites qui, dans les cellules, déchiffrent l'information portée par l'ARN messager et synthétisent les protéines correspondantes). Woese a séquencé bon nombre d'ARN ribosomiques de différents organismes et construit une phylogénie en les classant par ressemblances. Il l'a publiée en 1977 avec George Fox.

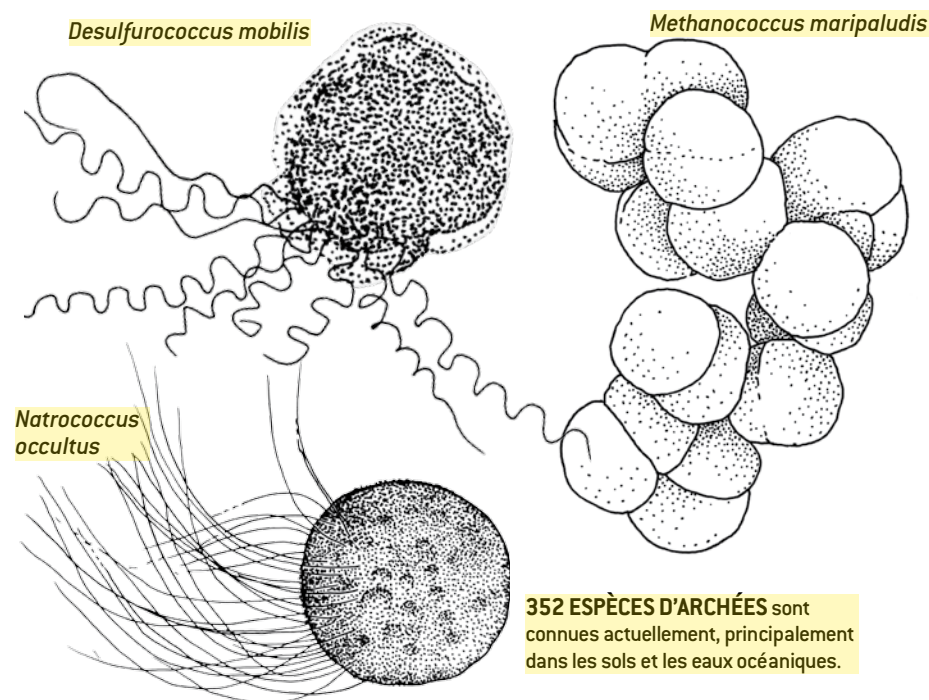
Surprise : premièrement, construire une phylogénie à partir des ARN ribosomiques est possible. Deuxièmement, les procaryotes se retrouvent non pas en un groupe, mais deux : les archéobactéries (aujourd'hui les archées) et les eubactéries (aujourd'hui les bactéries). La découverte est considérable. En effet, jusqu'alors on définissait les procaryotes sur des critères négatifs : pas de noyau, pas de cytosquelette, pas de vraie sexualité, pas de vraie mitose, etc. Or on savait très bien que construire une classification

sur des caractères négatifs ne reposait sur rien et risquait de donner des groupes non naturels. Or avec les ARN ribosomiques, Carl Woese a obtenu trois domaines *a priori* monophylétiques, c'est-à-dire constitués d'un ancêtre commun et de l'ensemble de ses descendants. L'étude des ARN ribosomiques fournissait ainsi une vraie phylogénie définie par des relations de parenté. Enfin, pas tout à fait complète tout de même.

PLS

Que manquait-il ?

H. L. G. : La phylogénie des trois domaines de Carl Woese n'avait pas de racine, pas d'orientation. On ne savait pas où mettre l'ancêtre commun de ces trois domaines. Alors on a commencé à y réfléchir. Habituellement, pour trouver la racine d'un arbre, on recherche un groupe extérieur à l'arbre qui présente avec le groupe étudié un ancêtre commun plus ancien. Or l'arbre de Carl Woese rassemblait tous les organismes vivants. Il était donc impossible de trouver un groupe extérieur, et donc *a priori* une racine. Sauf en utilisant des gènes qui se sont dupliqués et dont la duplication est plus ancienne que la racine, c'est-à-dire issue d'un ancêtre plus ancien. C'est l'idée qui est apparue lorsqu'on a compris, avec les génomes que l'on connaissait alors, qu'il existait des gènes communs aux archées et



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 ?

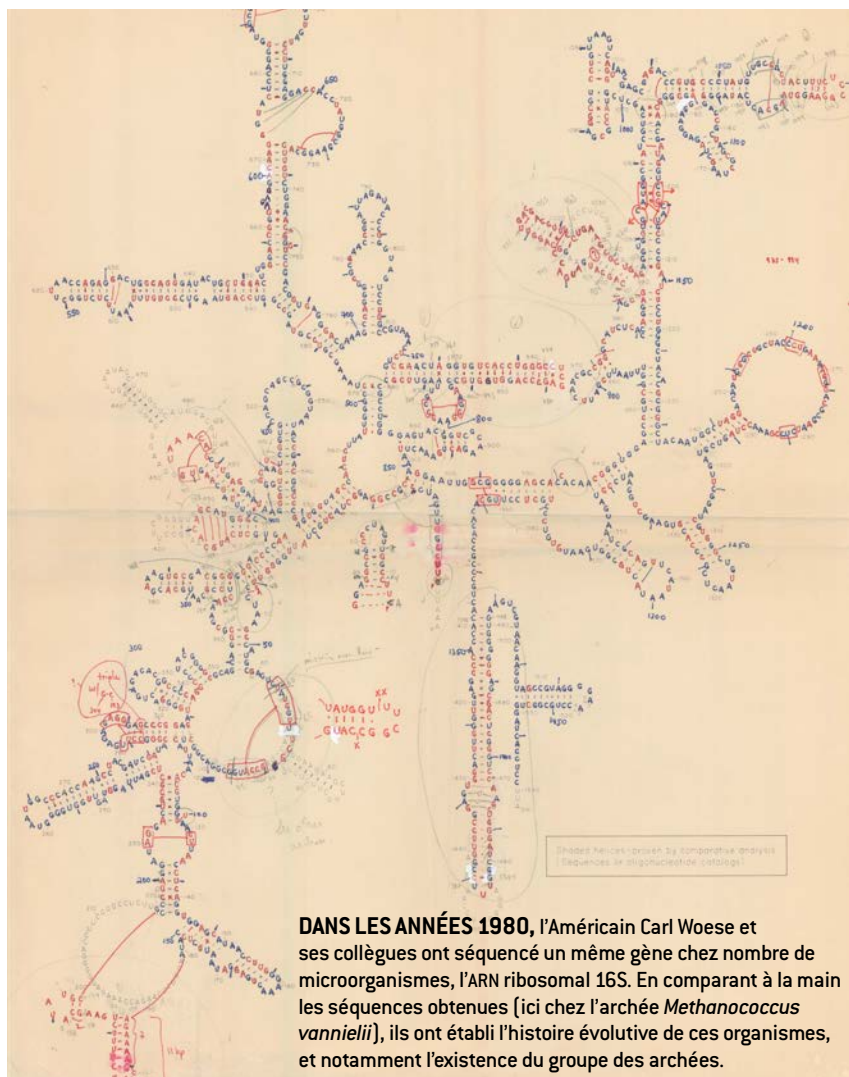
H. L. G. : Déjà au XIX^e siècle, des microbiologistes y pensaient. Mais Woese l'a montré avec brio. En partant d'ARN ribosomiques de maïs issus de la mitochondrie, du chloroplaste et du cytoplasme des cellules (donc purement eucaryote), il a montré que dans sa phylogénie, l'ARN ribosomique mitochondrial 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



www.mnccbb.uhexas.edu/J. Camone et al., BioMed Central Bioinformatics, 3:2 (2002)