

CET AMAS DE COLIBACILLES (*Escherichia coli*), vu en microscopie électronique, révèle la forme en cylindres oblongs des individus de cette famille de bactérie. Très commune dans la flore intestinale aérobie des mammifères, cette famille de bactéries est constituée surtout de souches inoffensives et de quelques souches pathogènes. Ces bactéries vivaient avec nous depuis l'aube de l'évolution.

L'ESSENTIEL

- **Le séquençage d'ADN à haut débit a révélé la diversité insoupçonnée des bactéries et des archées.**
- **Les eucaryotes, ou organismes à cellules dotées d'un noyau, sont issus de l'endosymbiose d'une bactérie avec une cellule procaryote proche des archées.**
- **Beaucoup de gènes sont capables de sauter d'une cellule à l'autre, de même espèce ou pas.**
- **Ces gènes sauteurs sont au fondement de la capacité des bactéries et des archées à s'adapter et à évoluer.**

Qui a vécu sur Terre en premier ? Les procaryotes ! Ces organismes constitués d'une seule cellule dépourvue de noyau – les bactéries et les archées – ont été les seuls habitants de notre planète pendant presque un milliard d'années. Les eucaryotes unicellulaires, qui ont un noyau dans leur cellule, sont apparus bien après, peut-être il y a quelque deux milliards d'années.

Selon toute vraisemblance, la cellule eucaryote est issue de l'intégration de cellules procaryotes dans une autre, un événement appelé endosymbiose. Une fois intégrés au sein d'une cellule, ces procaryotes internes ont continué à évoluer pour devenir des organites, c'est-à-dire des composants de la cellule qui accomplissent certaines fonctions, notamment la production d'énergie. C'est cette machinerie cellulaire qui a rendu les eucaryotes capables d'investir de nombreux milieux inaccessibles aux procaryotes et, ensuite, de participer à la formation des organismes multicellulaires complexes.

Pour autant, les eucaryotes multicellulaires s'entourent de nombreux microorganismes et en contiennent. Cette interaction intime joue-t-elle un rôle dans l'évolution ? Nous venons de faire des progrès majeurs dans l'élucidation de cette question. D'une part, une nouvelle description de l'incroyable diversité des procaryotes fournit de nouveaux indices sur les liens de parenté entre procaryotes et eucaryotes. D'autre part, nous avons compris que les populations d'unicellulaires procaryotes ont évolué en partageant des gènes et qu'elles utilisent cet échange d'information génétique pour continuer à évoluer avec notre corps.

Très tôt, les humains ont su que les animaux et les végétaux changent de forme pour s'adapter, ce qu'ils ont mis à profit en domestiquant certains. Pour élaborer la théorie de l'évolution, Charles Darwin est justement parti des connaissances ancestrales acquises sur la sélection : le concept de « sélection naturelle » traduit l'idée que l'organisme mieux adapté à son environnement survit et se reproduit plus facilement. La théorie de l'évolution a d'abord été naturaliste, procédant de la comparaison minutieuse entre les formes vivantes actuelles ou fossiles. Elle le reste aujourd'hui, mais à une nouvelle échelle, car les progrès immenses de la microscopie et de la biologie moléculaire permettent

désormais d'envisager une classification des organismes à partir d'observations faites au niveau cellulaire et moléculaire, notamment grâce au séquençage de l'ARN et de l'ADN.

Des indices précieux sur les liens de parenté entre organismes sont en effet transmis par l'ADN depuis l'aube de la vie. Leur exploitation a commencé dès les années 1970, lors des débuts du séquençage de l'ADN. Le microbiologiste américain Carl Woese et des collègues ont alors proposé de subdiviser les procaryotes en « archées » et en « bactéries », alors que, jusque-là, on les confondait sous le nom unique de « bactéries ».

Depuis ces années pionnières, les chercheurs ont séquençé toujours plus de gènes, même si le processus restait coûteux et lent. Cela a changé vers 2006-2007 : les coûts du séquençage se sont mis à plonger si vite que l'on a pu commencer à séquençer de façon routinière des génomes entiers. Environ 30 000 génomes sont aujourd'hui enregistrés dans des bases de données partagées.

À la recherche du métagénome

Or l'une des méthodes de séquençage – le séquençage à haut débit – permet de lire en une passe et à faible coût les millions de fragments d'ADN se trouvant dans un échantillon d'eau de mer, de terre de jardin, d'extrait intestinal, de sédiment de source hydrothermale... Les bio-informaticiens ont mis au point des algorithmes exécutés sur des calculateurs massivement parallèles (de nombreux processeurs travaillent simultanément) afin de superposer et raccorder toutes les séquences (on superpose les parties communes) et ainsi réassembler des génomes entiers. Ces techniques de séquençage rapide sont très sensibles et fonctionnent avec des échantillons pauvres en matériel génétique, voire ne contenant que celui d'une unique cellule... L'assemblage de tous les rubans d'ADN contenu dans un échantillon produit son « métagénome », c'est-à-dire le génome global de tous les organismes présents.

On comprend aisément que le séquençage à haut débit rende possible l'identification par l'ADN de souches de microorganismes impossibles à isoler et à cultiver en laboratoire, ce qui ne peut qu'élargir nos vues sur la diversité de la vie. C'est