

souches de *Yersinia pestis* contemporaines. Pour des ancêtres plus éloignés, tel celui des mammifères qui vivait il y a une centaine de millions d'années, une partie de l'information a été perdue par dérive. Cependant, pour les parties du génome contraintes par la sélection, les gènes en particulier, la reconstruction reste tout à fait possible. On peut même connaître le nombre de chromosomes et la façon dont les gènes étaient organisés dans ces génomes. À partir de ces données, on commence à comprendre les modes de vie des espèces disparues (température optimale de croissance, profondeur préférée d'anciens mammifères marins, utilisation de l'oxygène par des bactéries anciennes, alimentation, taille ou longévité de mammifères anciens, histoire des migrations et mélanges humains...), un peu comme si l'on retrouvait le code d'un vieux programme et que l'on reconstituait son fonctionnement à partir de ses versions nouvelles (voir la figure 3).

Cependant, l'application de ces méthodes à l'ensemble du monde vivant

se heurte à une difficulté : nombre d'organismes ont la capacité d'intégrer dans leur génome des gènes de leurs contemporains, proches apparentés ou non. Ces transferts horizontaux violent les hypothèses de reconstruction des génomes ancestraux, qui supposent que l'ancêtre commun des séquences appartient à l'ancêtre commun des espèces.

Le transfert de gène : obstacle ou indice ?

Il est apparu que chaque gène, quelle que soit sa fonction, pouvait, à un moment ou un autre de son histoire, avoir été transféré. À long terme, aucun gène n'a donc la même histoire que les organismes qui le portent. Ce constat semblait remettre en cause la possibilité théorique de reconstruire l'évolution du vivant à partir des gènes.

Ainsi, le gène était trop « individualiste » pour refléter l'histoire des organismes. N'avait-il pour autant rien à en raconter ? De fait, lorsqu'on observe l'ensemble des

gènes et, pour peu que l'on sache interpréter les singularités que révèle chaque histoire de gènes, on s'aperçoit que les transferts horizontaux contiennent une information inédite sur l'histoire du vivant. D'une part, ils peuvent venir à l'appui d'une hypothèse phylogénétique : si un ancêtre a reçu un gène par transfert, ce gène doit être observé chez ses descendants. Ainsi, la présence du gène peut être un caractère partagé permettant de reconnaître les organismes qui sont proches.

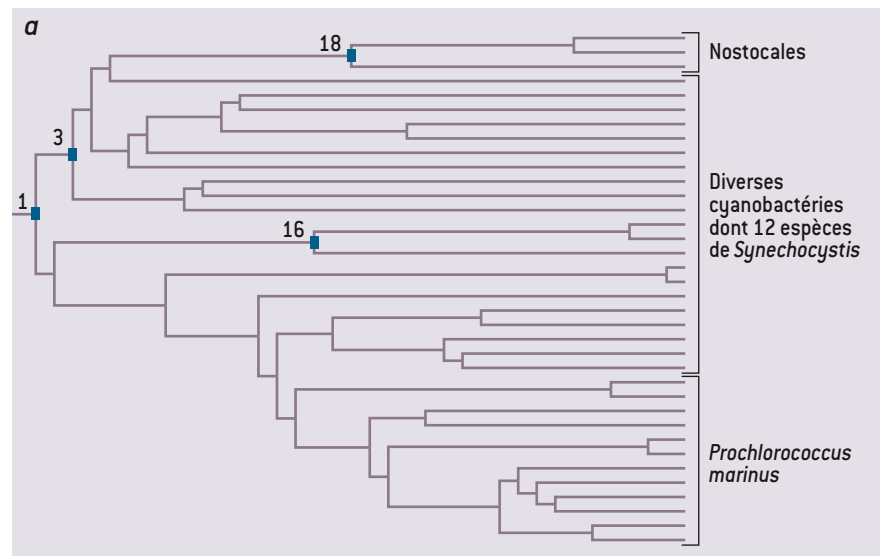
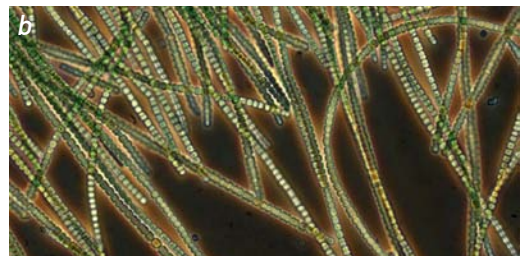
D'autre part, un échange de gènes entre deux organismes ne peut avoir lieu que s'ils sont contemporains. Le transfert horizontal contient donc une information sur les âges relatifs des lignées qui échangent des gènes : on ignore à quelle époque exactement vivaient les organismes anciens, mais, avec des relations de descendance et des échanges entre espèces contemporaines, on peut espérer établir une chronologie de l'histoire de la vie. Or les transferts ont lieu principalement chez les organismes microscopiques (bactéries, virus, champignons unicellulaires, autres

TRANSFERT DE GÈNES ET DATATION

Par définition, le transfert de gènes entre deux espèces n'est possible que si les deux espèces sont contemporaines. Ce processus, qui transparaît dans le génome des espèces actuelles, donne aux phylogénéticiens de précieuses informations sur les âges relatifs d'espèces anciennes dont les traces fossiles sont rares. En 2012, à partir de 36 génomes d'espèces différentes, les auteurs ont ainsi reconstitué la phylogénie chronologique des cyanobactéries, les premiers organismes à avoir fait de la photosynthèse et produit de l'oxygène en grande quantité (a). Les plantes ne récupèrent cette faculté en intégrant des cyanobactéries au fonctionnement de leurs cellules.

Les nœuds de l'arbre sont positionnés de gauche à droite selon leur chronologie. Le nœud 1, le plus ancien, est parfois daté à 3,8 milliards d'années. Les plantes ne figurent pas sur cet arbre, mais on voit la naissance au nœud 18 du clade des cyanobactéries multicellulaires, les nostocales, qui forment des filaments verts de cellules différenciées (b). Comme ce clade a reçu, par transfert, des gènes d'autres mor-

ceaux de la phylogénie, on estime que cette forme de multicellularité est relativement récente. De même, la présence de nombreux transferts du clade du nœud 3 vers le clade du nœud 16 lui donne une origine relativement récente, alors que les hypothèses habituelles plaçaient cette diversification à la racine des cyanobactéries.



Adapté de G. J. Szallasi et al., PNAS, vol. 109, pp. 17513-17518, 2012

organismes unicellulaires tels que l'amibe ou la paramécie). Ces organismes ne laissent presque aucune trace fossile, ce qui rend la datation des trois quarts de l'histoire de la vie inaccessible aux méthodes classiques. Le transfert fournit ainsi l'outil de datation dont la petite taille de ces organismes nous privait.

Le transfert recèle encore d'autres traces inexploitées de l'histoire de la vie. Parmi les êtres vivants anciens, certains ont laissé des descendants, d'autres non. Il est raisonnable de penser que ceux qui n'ont laissé aucun descendant sont bien plus nombreux. Pourtant, on reconstitue l'histoire à partir des génomes actuels – les descendants de ceux qui, par définition, ont laissé des descendants ! On n'a ainsi accès qu'à une faible partie de la biodiversité ancienne.

Et si l'on détecte un échange de gènes entre deux organismes anciens, c'est à partir des descendants de celui qui a reçu le gène. En revanche, il n'y a aucune raison pour que celui qui l'a donné, lui, fasse partie de la minorité qui a eu des descendants. Statistiquement, il est probable qu'on ait ainsi accès à un gène d'une espèce ancienne sans descendant, qui n'aura donc pas transmis ses gènes en dehors de ce transfert. On est donc face à une biodiversité beaucoup plus importante que celle que nous offrirait la seule transmission d'ascendant à descendant.

Les gènes ne sont pas les marqueurs dont on avait rêvé pour reconstruire l'histoire de la vie. Ils racontent parfois une autre histoire, la leur. Mais si on sait modéliser la complexité de ces histoires, ils permettent de trouver, dans les génomes,

des informations dont on pensait avoir perdu la trace. Il reste sans doute encore beaucoup à découvrir, dans nos cellules, sur notre passé.

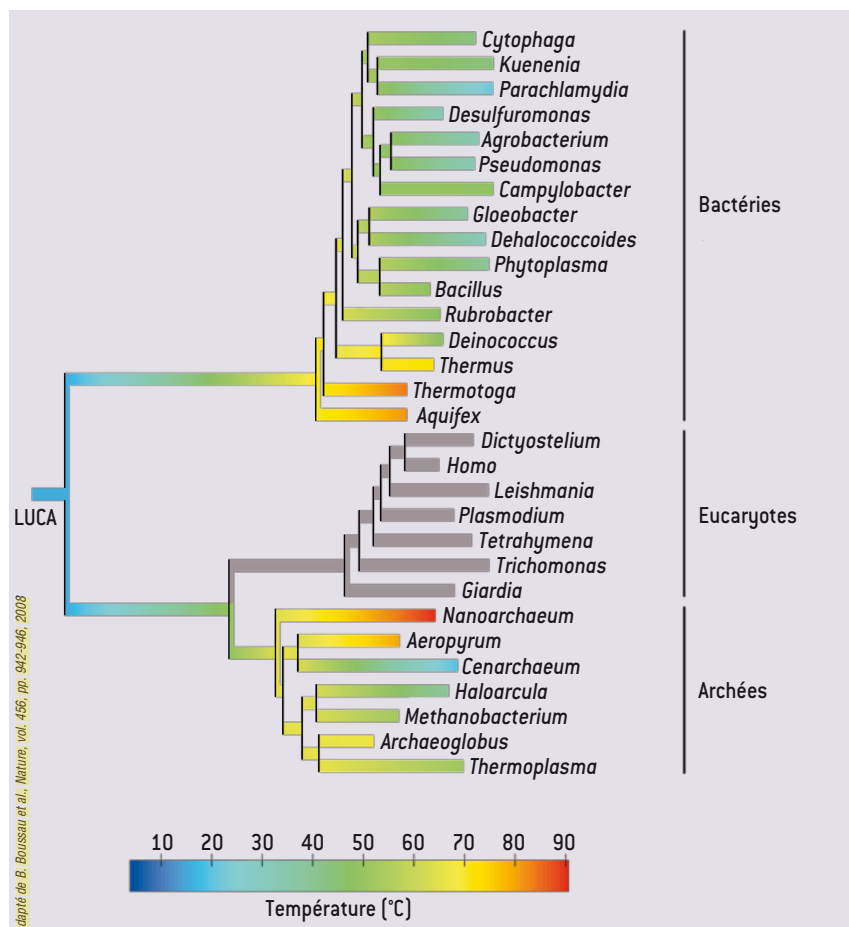
D'autres supports que l'ADN ?

En informatique, les supports de l'information numérique ont beaucoup évolué depuis les cartes perforées. Ils sont encore très divers aujourd'hui. Dans le monde vivant, il existe également plusieurs types de supports, mais c'est la double hélice d'ADN qui s'est établie comme le support universel chez les organismes cellulaires. Pourtant, ce support n'aurait pas été le premier.

L'ARN, toujours utilisé aujourd'hui par certains virus et conservé comme intermédiaire pour la lecture de l'information par les cellules, pourrait bien l'avoir précédé. Il avait l'avantage d'être capable de catalyser des réactions enzymatiques, offrant la possibilité d'organismes primitifs assez simples, dont le support d'information génétique est en même temps actif biochimiquement. Mais cette activité est également un facteur d'instabilité, et les organismes de ce monde à ARN auraient finalement adopté un média dédié pour préserver leur précieuse information génétique.

Cette transition est comparable à l'invention de l'imprimerie qui, plus rapide pour copier l'information écrite et moins sujette à l'erreur, a remplacé le scribe. Ou, en informatique, à l'abandon des cartes perforées qui stockaient l'information numérique lisible par les ordinateurs, pour des supports électromagnétiques.

Aujourd'hui, l'ensemble du vivant a adopté l'ADN comme média de stockage et de transmission de l'information numérique. Même l'informatique envisage d'effectuer cette transition : l'ADN pourrait remplacer les supports magnétiques pour stocker l'information numérique et être utilisé par les ordinateurs à la place des disques durs (voir l'article de Ch. Desimoz dans ce numéro). Cette innovation pourrait avoir des conséquences sur le monde vivant lui-même : pour des organismes habitués à récupérer l'ADN dans l'environnement par transfert horizontal, les grandes quantités d'ADN présentes dans des machines constitueraient un substrat pour inventer de nouveaux gènes. L'informatique elle-même changerait alors le cours de l'évolution !



3. RECONSTITUTION DE LA TEMPÉRATURE OPTIMALE DE CROISSANCE de l'ancêtre commun à tout le vivant (LUCA) à partir de l'analyse phylogénétique d'organismes actuels (65 eucaryotes, 60 archées et 331 bactéries). En se fondant sur l'observation selon laquelle chez les procaryotes (archées et bactéries), la température optimale de croissance est corrélée à la proportion de nucléotides C et G dans le génome, des chercheurs du laboratoire des auteurs ont décrit une évolution de la température en deux temps : LUCA vivait à environ 20 °C, puis la tolérance thermique de ses descendants a augmenté, avant de diminuer pour les espèces les plus récentes.