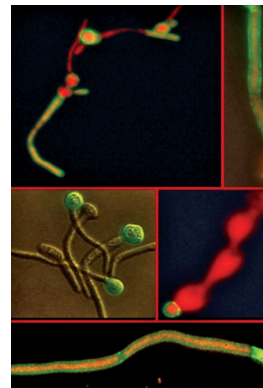


## Microbiotes et métagénomique

Jean Weissenbach<sup>1,2</sup>, Abdelghani Sghir<sup>2</sup>

► Au cours des années 1990, de nouvelles techniques d'analyse ont été appliquées à l'étude des flores et des communautés microbiennes. Ces approches essentiellement analytiques ont consisté en des inventaires moléculaires massifs. Quel bilan vingt ans après ? Sur le plan des compositions microbiennes, de leur suivi et du diagnostic, les résultats sont impressionnants. Mais que font tous ces organismes, quels sont les facteurs qui favorisent ces associations et leur maintien ? Les tentatives d'inventaires de gènes qui approchent l'exhaustivité nous permettent d'approfondir certains processus physiologiques mais ne disent guère plus. Des microbiotes ont même commencé à être manipulés. Mais comme souvent dans la dialectique entre théorie et pratique, « ça marche » mais nous sommes loin d'avoir compris comment. ◀

Le monde microbien représente 50 % de la biomasse sur terre et de l'ordre de  $10^{30}$  cellules [1]. Les microbes ont colonisé l'ensemble de la planète que ce soit sous forme planctonique ou en association, notamment dans des biofilms ou en micro-agrégats [2] sous forme libre ou liée à des organismes hôtes. Antoni van Leeuwenhoek (1632-1723), naturaliste amateur et inventeur du microscope, fut probablement le premier à observer, à l'aide d'un microscope de sa fabrication, une communauté bactérienne provenant du grattage de ses dents. La microbiologie était née, mais il fallut encore attendre deux siècles et notamment des microscopes plus performants pour que cette discipline prenne une réelle importance. Elle fut alors, pour plus d'un siècle, profondément formatée par les concepts de Koch et de Pasteur. On s'intéressa avant tout à des organismes isolés, manipulés sous forme de cultures pures. Ces cultures pures ont permis de nombreuses observations majeures en microbiologie aussi bien dans le domaine de la santé que de l'environnement. Pourtant, il fut assez vite constaté que certaines bactéries observées



<sup>1</sup>CEA ; Institut de Génomique ; Genoscope, 2, rue Gaston Crémieux, 91057 Évry, France ; <sup>2</sup>UMR8030 Génomique Métabolique, CEA/CNRS/ Université d'Évry-Val-d'Essonne, 91057 Évry, France. [jsbach@genoscope.cns.fr](mailto:jsbach@genoscope.cns.fr)

au microscope dans des mélanges ne se retrouvaient pas à la suite des tentatives de culture [3].

Au cours des années 1980, Norman Pace, s'appuyant sur les résultats de séquençage d'ARN ribosomiques (ARNr) 16S de Carl Woese [4], commença à s'affranchir de l'obstacle de la culture en analysant, par PCR (*polymerase chain reaction*) et séquençage, les gènes d'ARN ribosomiques à partir de l'ADN extrait de communautés microbiennes [5]. Les ARN ribosomiques jouent un rôle central dans la synthèse protéique [6] (→), ce qui limite leur dérive évolutive et rend

(→) Voir l'article Nobel de P. Romby et al., m/s n° 11, novembre 2009, page 977

pratiquement impossible leur transfert horizontal. Ils constituent donc un identifiant très stable et fiable des organismes procaryotes et eucaryotes. C'est même autour de cette stabilité que s'est ancrée la notion d'espèce chez les bactéries ainsi que la classification phylogénétique. L'étude des communautés microbiennes, ou micro-

(→) Voir la Synthèse de C. Dauga et al., m/s n° 3, mars 2005, page 290

biotes, fut ainsi (re) lancée. Elle a connu, depuis une vingtaine d'années, un essor remarquable [7] (→). Les microbiotes sont universels et leur fonctionnement a été façonné par l'évolution. Les espèces qui les constituent se sont adaptées à leur voisinage au point parfois de ne plus pouvoir s'en passer. Si nous commençons à percevoir leurs effets sur le reste de la biosphère, sur les biomes<sup>1</sup> dont ils font partie et sur leurs partenaires non-microbiens, nous ne savons guère comment s'exercent ces effets. Quant à ce qui fonde l'existence des microbiotes et leur permet de se maintenir au cours du temps, nous en ignorons l'essentiel.

À défaut de pouvoir appréhender les microbiotes dans leur globalité, les microbiologistes ont adopté la démarche, habituelle, qui considère que les propriétés et les effets de ces microbiotes sont la résultante de l'action des espèces qui les composent. Cette action dépend, à l'évidence, des capacités biologiques encodées dans les gènes de ces espèces. Les travaux de Pace ouvraient la voie à la démarche

Vignette (Photo © Inserm - Daniel Poulain, Chantal Fradin).

<sup>1</sup> Écosystèmes.

moléculaire qui devait donner accès aux espèces et à leurs propriétés. On fit donc appel au ribotypage (c'est-à-dire à l'analyse des ARN ribosomiques) pour aborder des questions comme la composition des microbiotes, leur stabilité, résilience et dynamique, en fonction de divers facteurs comme le temps, le milieu et, le cas échéant, l'âge et l'état de l'hôte, etc.

### Recensement des espèces des microbiotes

La pratique du ribotypage fut rapidement généralisée et de nombreuses espèces inconnues furent identifiées. On eut ainsi la démonstration que la majorité des espèces bactériennes n'avait pas été isolée par les techniques classiques de microbiologie, en particulier, de culture microbiologique. On put aussi procéder à des recensements assez étendus de la composition en espèces des communautés microbiennes. Le degré variable d'appareillage des séquences d'ARNr 16S à une référence fut utilisé comme échelle de mesure pour inclure, ou exclure, un *ribotype* d'une espèce, d'un genre ou d'un phylum bactérien. À cette approche, appuyée sur le séquençage, s'ajouta rapidement une technique d'hybridation *in situ* à ces mêmes ARNr au moyen de sondes nucléiques fluorescentes [8], confirmant, d'une part, leur réalité au microscope et donnant, d'autre part, une estimation approximative de leur abondance au sein d'un microbiote. L'observation du signal d'hybridation permettait aussi de montrer que les bactéries concernées synthétisaient bel et bien des ARN ribosomiques et étaient donc biologiquement actives. Mais ces bactéries restaient en général réfractaires à la culture. On se disputa pour qualifier les espèces concernées : sont-elles non cultivables, non cultivées, viables non cultivables, etc. [9] ? De nombreuses hypothèses furent avancées pour expliquer notre incapacité à les mettre en culture.

L'hypothèse la plus simple fut l'absence d'un ou plusieurs nutriments, ce qui reflétait notre ignorance des besoins nutritifs et métaboliques des espèces non cultivées. Dans un microbiote, ces nutriments peuvent éventuellement devenir disponibles à la suite de leur production par des organismes présents dans le milieu. Des transferts entre espèces et la formation de véritables guildes<sup>2</sup> assurant une circulation de nutriments pouvaient rendre compte du phénomène [10]. Cette possible auxotrophie<sup>3</sup> peut aussi être due à l'absence de molécules signal émises par certaines cellules pour stimuler la croissance des autres.

Alternativement, des conditions physico-chimiques inappropriées pourraient provoquer une mise en veille du métabolisme et d'autres activités qui serait réversible. Ces cellules devenues dormantes constitueraient une réserve capable de ressusciter et de se propager avec le retour de conditions de croissance qui seraient plus favorables [11]. Cette dormance pourrait aussi bien s'appliquer à des espèces cultivables connues qu'à des espèces non cultivées [7].

<sup>2</sup> En zoologie et botanique, la guilda désigne l'ensemble des espèces d'un même habitat qui exploitent la même ressource.

<sup>3</sup> Un organisme auxotrophe pour un nutriment est un organisme auquel il faut fournir le nutriment qu'il ne peut synthétiser.

### Les limites du recensement par l'ARN ribosomique

Pendant une bonne dizaine d'années, on s'appuya sur l'inventaire des espèces par ribotypage en extrapolant les caractéristiques métaboliques et physiologiques des représentants des espèces connues d'un microbiote. Mais, d'une part, ces inventaires ne nous disaient rien de plus sur les bactéries non cultivées dont les génomes restaient inconnus et, d'autre part, le séquençage du génome de plusieurs isolats d'une espèce définie par son ARNr 16S montrait souvent des différences sensibles entre les isolats, notamment, dans leur composition en gènes. Un cas assez emblématique est celui d'*Escherichia coli* dont les souches, MG1655 et O157:H7, ont des génomes qui diffèrent d'un tiers de leur contenu [12]. Les génomes bactériens sont fréquemment l'objet de pertes ou de gains de régions génomiques pouvant aller de plusieurs kilobases (kb) à plusieurs dizaines, voire au-delà de la centaine de kb. Sans surprise, cette plasticité se traduit par des modifications sensibles de leur physiologie dont la variabilité, d'un isolat à l'autre, avait souvent été observée par des approches classiques de biochimie et d'immunologie.

La découverte de cette grande plasticité est le résultat direct des programmes de séquençage de génomes complets. Elle provoqua une révision majeure de notre compréhension des mécanismes de l'évolution des procaryotes. Alors qu'on avait longtemps cru que les génomes évoluaient essentiellement par de petits événements mutationnels, mutations ponctuelles et insertions ou délétions d'un ou quelques nucléotides (insertions/délétions ou INDEL), on comprit qu'ils étaient aussi profondément remaniés par des gains et des pertes de segments chromosomiques pouvant porter plusieurs dizaines de gènes, voire plus. Des analyses plus fines montrèrent de manière convaincante que les acquisitions de ces segments, ou îlots génomiques, résultaient d'un transfert horizontal entre bactéries pouvant être phylogénétiquement éloignées. Ces acquisitions d'îlots procurent aux procaryotes une capacité d'adaptation considérable que ce soit pour la colonisation de nouvelles niches, la résistance à des changements brutaux et drastiques du milieu ou la présence de prédateurs.

La variabilité des génomes de procaryotes nécessita aussi d'être intégrée dans notre conception de la taxonomie des espèces procaryotes basée pour l'essentiel sur l'ARNr 16S [13] (→).

(→) Voir la Nouvelle de V. Daubin et S.S. Abby, *m/s* n° 8-9, août-septembre 2012, page 695

Ceci aboutit au concept de pan-génome [14] dans lequel on distingua (1) le « *core genome* », c'est-à-dire

un ensemble de gènes toujours présents et retrouvés dans tous les isolats appartenant à une même espèce et (2) le « *dispensable genome* », un autre ensemble de gènes, dont une part (très variable) pouvait être facultativement présente ou absente d'une souche à l'autre.

Pour étudier un microbiote, la meilleure manière de surmonter cette difficulté fut de prendre cette plasticité en compte et d'inventorier non seulement les espèces par l'ARNr 16S mais aussi les autres gènes (constants ou facultatifs) de ces espèces. Comme le séquençage massif se répandit et que les capacités des grands centres de séquençage devenaient plus disponibles, on commença à inclure la dimension génique dans l'étude des microbiotes. La métagénomique était née [15] et devait permettre, bien plus que le ribotypage, de renouveler en profondeur l'étude des communautés microbiennes de toute nature et de toute origine. Fut alors proposé le néologisme de microbiome (voir Encadré).

### Analyses métagénomiques tous azimuts

La métagénomique fut présentée comme la panacée pouvant apporter une réponse satisfaisante aux questions concernant les microbiotes. L'arrivée du séquençage de deuxième génération en particulier devait apporter la masse de données qui permettrait des progrès décisifs. On s'intéressa à de très nombreux types de microbiotes colonisant une grande variété d'environnements et occupant un nombre considérable de niches, soit en interaction avec d'autres organismes uni- ou multicellulaires (surfaces animales, rhizosphère, feuilles, coraux, etc.), soit libres, dans des milieux naturels, comme les sols et les eaux. Les communautés bactériennes sont ainsi devenues des sujets d'études majeurs illustrant leur importance dans les domaines de la santé et de l'environnement. Le volume de données produites est considérable. Il est impossible de résumer l'ensemble des résultats de ces analyses. Une partie notable des résultats concernant plus spécifiquement les microbiotes humains a été rapportée dans médecine/sciences et sera très peu discutée ici [16-20] (→).

On conçoit aisément que l'analyse de données fragmentaires issues d'un mélange de nombreuses espèces de génomes est une gageure. Une méthodologie s'est ainsi progressivement dégagée au fur et à mesure du développement de nouveaux outils informatiques [21] (Figure 1). Les microbiotes environnementaux furent les premiers à être étudiés [22], mais l'anthropocentrisme aidant, les résultats au plus fort impact et les plus nombreux concernent les métagénomiques humaines [23].

Chez l'homme, les approches s'appuyant sur des cultures laissent entrevoir l'existence d'un socle commun à l'ensemble des microbiotes intestinaux. Cette notion est aujourd'hui remise en question. On observe une variation considérable d'une population à l'autre, entre individus d'une même population, mais aussi pour un même individu au cours du temps et selon le régime alimentaire [45] (→).

(→) Voir : la Synthèse de R. Burcelin *et al.*, *m/s* n° 8-9, août-septembre 2013, page 800 ; la Nouvelle de P. Cani et A. Everard, *m/s* n° 2, février 2014, page 125 ; la Synthèse de C. Andréjak et L. Delhaes, *m/s* n° 11, novembre 2015, page 971 ; la Synthèse de C. Audebert *et al.*, *m/s* n° 12, décembre 2014, page 1144 ; la Synthèse de A. El Kaoutari *et al.*, *m/s* n° 3, mars 2014, page 259

(→) Voir la Synthèse de R. Burcelin *et al.*, page 952 de ce numéro

### Sémantique du néologisme microbiome

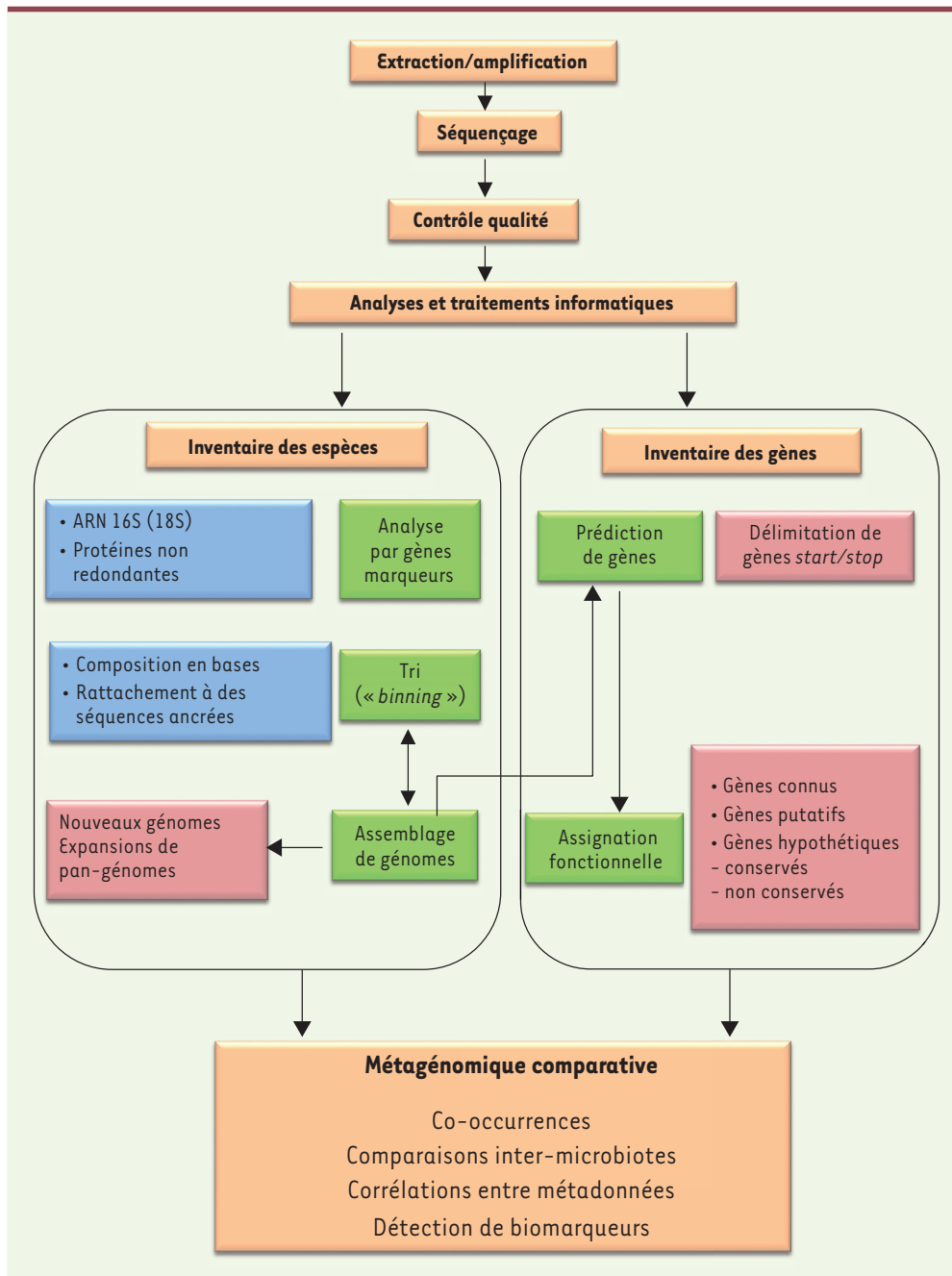
Si la définition du terme de microbiote semble assez simple, son usage assez ancien et généralement admis, le terme de microbiome pose plusieurs problèmes. Le néologisme microbiome résulte d'une association, mais celle-ci semble différente selon les sources. Si l'on s'en tient à la définition du terme tel qu'il apparaît dans l'édition de langue française de Wikipedia (06/09/2016), le microbiome serait le biome du microbiote : c'est micro-biome soit le biome « micro ». Mais le biome étant par définition l'ensemble des écosystèmes d'une unité de lieu, un concept éminemment globalisant tendant à tout inclure, n'y a-t-il pas une contradiction dans le terme de microbiome ? Un biome qui ne serait que « micro » est-il encore un biome ? ! Du côté de la langue anglaise autre problème : l'association est microbi\_ome, c'est-à-dire l' « ome » microbien. « Microbome » aurait eu l'avantage d'éviter cette confusion. Mais le terme n'en resterait pas moins ambigu. S'agit-il d'un microbiote ou si l'on s'en réfère à ce qui est souvent considéré comme la validation étymologique [44] du métagénome d'un microbiote ? Constatons aussi que l'entrée « microbiome » de l'édition en langue anglaise de Wikipedia (06/09/2016) pointe sur l'article « microbiota » !

Au fur et à mesure de l'accumulation de nouvelles études, l'existence du socle commun devient de plus en plus improbable [24].

Des observations concernant la composition et les dynamiques des microbiotes sont souvent bien antérieures à l'usage de la métagénomique qui n'a fait que les documenter extensivement. Et, s'il est devenu possible de procéder à des inventaires de fonctions (voir ci-après), les analyses de métagénomique se heurtent au côté très fragmentaire des données de séquence produites par séquençage qui ne permettent que rarement d'associer des gènes (traduits par des phénotypes physiologiques) à un genre ou une espèce (Figure 1). En outre, l'observation d'un fragment d'ADNr 16S ne prouve pas la présence stable de l'espèce bactérienne concernée, qui peut être fortuite (espèce passagère) ou accidentelle.

### L'inventaire des fonctions

L'inventaire des gènes est encore plus impressionnant que celui des espèces rencontrées [25-27]. Mais que peut-on en dire ? En dépit d'une différence de composition des espèces entre les microbiotes intestinaux, on retrouve une forte redondance fonctionnelle sous forme d'un socle de



**Figure 1. Déroulé d'une analyse métagénomique de microbiote.** Dans un premier temps le microbiote d'intérêt est soumis à une extraction de l'ADN. L'ADN métagénomique obtenu peut être soumis à une étape d'amplification, pour augmenter les quantités disponibles. Il est ensuite séquencé par une méthode de nouvelle génération (en général). Les lectures obtenues sont souvent de taille très courte (100 pb, paires de bases). Les données brutes obtenues sont alors examinées par des procédures de contrôle qualité puis traitées à l'aide de logiciels capables de produire différents types d'analyses visant à déterminer le contenu en espèce (analyse taxonomique) ou en gènes (analyse fonctionnelle). Les principaux logiciels d'analyse taxonomique vont identifier les fragments d'intérêt et les comparer à des collections de séquences de référence. Les gènes de référence utilisés sont principalement et très majoritairement les gènes qui encodent les ARN ribosomiques qui permettent une assignation à une espèce, un genre ou un phylum. Si on souhaite une analyse plus fine, on inclut des comparaisons à des gènes marqueurs présents en simple copie par cellule (par exemple

ceux codant les protéines ribosomiques) ou qui sont spécifiques d'un groupe phylétique (c'est-à-dire spécifiques des espèces d'un même phylum). On peut allonger la longueur des fragments d'intérêt sélectionnés en procédant à l'assemblage des lectures correspondantes. Ces analyses débouchent sur l'identification des séquences de référence phylogénétiquement les plus proches. Elles peuvent être complétées par le positionnement des séquences analysées dans un arbre phylogénétique. Cette première analyse donne une idée sur l'identité et le groupe taxonomique de la séquence considérée. Une idée plus précise de l'appartenance à un groupe taxonomique peut être fournie par le « binning », un tri souvent basé sur la composition nucléotidique des fragments analysés. Il existe aussi des méthodes de tri qui utilisent d'autres stratégies. Pour augmenter la puissance des analyses, il est important de travailler avec les séquences les plus longues possibles et donc de procéder à des assemblages de séquences recouvrantes colinéaires préalablement à d'autres traitements. Mais, inversement, certains de ces traitements, comme le tri préalable, peuvent eux aussi augmenter la qualité des assemblages. Les choix stratégiques sont donc parfois délicats. Il existe de nombreux programmes d'assemblage de métagénomes dont les performances peuvent être différentes en fonction de la nature des données utilisées. Les traitements des inventaires « fonctionnels » des métagénomes sont en général les mêmes que ceux utilisés pour les génomes. Ils se pratiquent surtout sur les séquences assemblées. Traitements généraux (en ocre). Traitements « taxonomiques » et « fonctionnels » (en vert). Méthodes ou ressources utilisées dans les traitements « taxonomiques » (en bleu). Résultats des traitements (en rouge).

fonctions communes. Une grande partie de celles-ci concerne le métabolisme central et intermédiaire. Mais ceci n'est pas une surprise. Un système biologique, même rudimentaire doit disposer d'un degré minimal d'autonomie pour se maintenir et se reproduire. L'obtention de ce type de résultats a mobilisé des ressources qui semblent démesurées par rapport à leur signification. Certes, sans le secours de la bio-informatique, il eut été impossible de procéder à des analyses efficaces des données de métagénomique (Figure 1). Mais, même si elle fait sens, et elle le fait, la grande force de la bio-informatique reste la prédiction du prévisible. C'est plutôt au niveau de fonctions, déjà partiellement connues, qu'on note des progrès. Ainsi a-t-il été possible d'analyser en profondeur certaines fonctions majeures du microbiote du côlon comme la digestion des polysaccharides [20]. De même, des résultats parlants avaient été obtenus pour des microbiotes de processus environnementaux déjà largement étudiés et relativement compris. Ceux-ci concernent, par exemple, le traitement des eaux où se produisent des étapes majeures de la minéralisation de substrats organiques dans le déroulement des cycles du carbone, de l'azote et du phosphore. De nouvelles étapes de ces cycles, comme le processus anammox (*anaerobic ammonium oxidation*) [28], ou l'oxydation anaérobie du méthane [29], avaient été décrites ces dernières décennies et leurs acteurs microbiens en partie identifiés. Les approches moléculaires ont cependant permis d'identifier certains gènes clés jusque-là inconnus.

Après 25 ans de génomique et 15 ans de métagénomique, la fraction des gènes connus au sein d'un génome procaryote n'a guère progressé et se situe toujours autour du tiers. Pour un second tiers il est possible de proposer une fonction vraisemblable. Quant au dernier tiers, il est constitué de gènes hypothétiques, sans fonction connue, se subdivisant en gènes conservés et non conservés. Avec l'accumulation de nouvelles séquences, certains gènes hypothétiques se révèlent être conservés alors que de nouveaux gènes hypothétiques sont prédits. Toutes ces conclusions s'appuient sur l'utilisation des banques de données de séquence qui s'enrichissent sans discontinuer mais sont fortement biaisées par une surreprésentation des séquences provenant de génomes de pathogènes dont l'essentiel se restreint, de surcroît, à 3 phylums bactériens [30]. On note à l'inverse une sous-représentation des espèces environnementales et plus généralement un biais en faveur des cultivables [46] (→).

Le séquençage a montré maintes fois que la fraction des gènes inconnus est plus élevée dans les espèces non cultivées.

Comme dans toute analyse de génomes, le résultat final débouche sur une double frustration : d'une part l'empilement de plus en plus démesuré de représentants additionnels de fonctions connues, notamment les plus étudiées et les plus universellement distribuées dans le vivant et, d'autre part, l'accroissement et l'élargissement inexorables des collections de gènes codant des fonctions inconnues. Les courbes de dénombrement de nouveaux gènes commencent cependant à tendre vers un plateau [25]. Pour le moment ces signes espérés de raréfaction s'observent surtout au niveau de biomes très étudiés (Figure 2). Mais ils montrent qu'on s'achemine vers une saturation progressive de l'inventaire des protéines de la biosphère.

## Physiologie et « rôle » des microbiotes

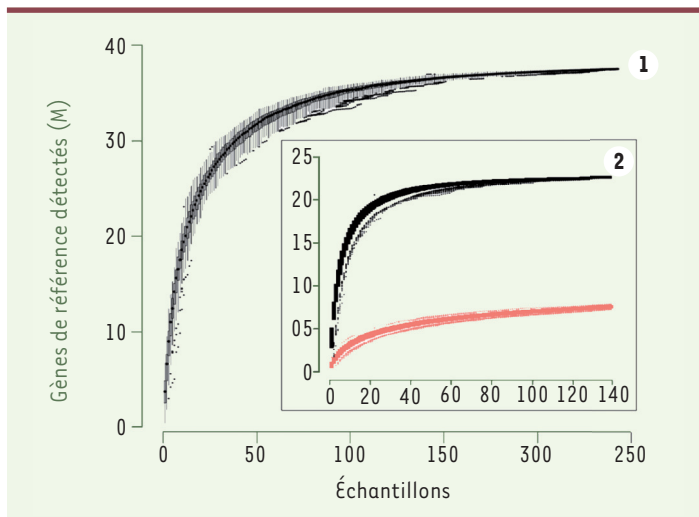
De manière surprenante, on constate des traits communs de la physiologie de microbiotes d'écosystèmes différents en dépit de leurs différences de composition ou de complexité. Ils sont organisés selon une même logique, en groupes fonctionnels ou trophiques, avec des interactions multiples, qu'il s'agisse d'interactions avec leur milieu physique ou d'interactions biotiques (commensalisme, mutualisme, syntrophie, etc.)<sup>4</sup>. Ils sont aussi construits à partir de leur environnement immédiat (dans lequel l'environnement sélectionne et l'espèce façonne) où la redondance fonctionnelle est la règle. Ainsi, le tube digestif humain ou animal et la rhizosphère végétale présentent des similitudes frappantes. Le tube digestif et les racines végétales ont pour rôle majeur l'absorption des nutriments. Les deux habitats présentent des structures d'échanges et de dialogue, sous forme de microvillosités ou des poils racinaires. Chacune de ces structures offre des conditions physicochimiques particulières (pH, concentrations en nutriments, oxygène) qui favorisent l'établissement de micro-niches colonisées par des espèces adaptées prédominantes qui contribuent à la spécificité métabolique de ces microenvironnements [31, 32]. Il en résulte que chaque habitat est colonisé par un microbiote particulier, à la fois complexe et dynamique, composé d'espèces adaptées de bactéries, d'archées, d'eucaryotes et de virus [33].

À côté d'une fraction d'espèces dominantes dans un microbiote, on rencontre une flore d'espèces supplémentaires qui joue un rôle fonctionnel respectable. Plusieurs études moléculaires ont révélé que les microbiotes sont constitués d'un grand nombre de taxons en faible abondance [34-36]. Ces espèces qui paraissent minoritaires en nombre, peuvent contribuer à des processus spécifiques comme la méthanogenèse, la méthylotrophie, la fixation de l'azote dans les tubes digestifs et les rhizosphères [37-39] avec des répercussions majeures sur la nutrition, la croissance de l'hôte ou sur les cycles biogéochimiques [40].

Les microbiotes montrent en général une capacité de résilience remarquable quand ils sont soumis à des perturbations ou des agressions liées à des changements des conditions chez l'hôte ou de paramètres environnementaux [24]. Une forte redondance de certaines

<sup>4</sup> Le commensalisme correspond à un type de relation où une espèce en exploite une autre sans établir de relation de nature parasitaire ; le mutualisme est une interaction entre deux espèces, dans laquelle les organismes impliqués tirent tous les deux profit de cette relation ; la syntrophie est une relation mutualiste établie entre deux espèces au cours de laquelle chaque espèce produit un élément nutritif essentiel au développement de l'autre.





**Figure 2. Comparaison de courbes de raréfaction de catalogues de gènes de référence provenant de quelques grands projets d'analyse de métagénomies. 1.** Les séquences de 243 échantillons d'ADN métagénomiques (68 sites océaniques) ont été assemblées et soumises à une analyse de prédiction de gènes produisant un total de 111,5 millions (M) de gènes. Ces gènes prédits ont été regroupés avec les jeux de séquences de gènes d'autres métagénomies océaniques disponibles publiquement (26 millions de gènes « externes »). L'analyse de regroupement (*clustering*) de séquences identiques à 95 % a fourni un ensemble de gènes de référence de 40 millions de gènes de métagénomies océaniques. La partie droite de la courbe montre un très faible accroissement de diversité pour l'addition des 100 derniers échantillons. **2.** La partie encadrée compare les courbes de raréfaction des gènes procaryotes marins (courbe noire, 24 millions de gènes) et intestinaux (courbe rouge, 7 millions de gènes) à partir d'échantillons normalisés pour la couverture en séquence. La nette différence du niveau des asymptotes fait ressortir une plus grande diversité de gènes de procaryotes marins (d'après [26]).

protéines clés responsables de fonctions spécifiques, capables de s'exprimer sous une large gamme de conditions, contribue vraisemblablement à cette capacité de résilience.

D'une manière générale les microbiotes semblent être un facteur d'équilibre du biome dans lequel ils se développent. Les perturbations de ces équilibres, les dysbioses, peuvent avoir des causes et des conséquences multiples que nous sommes loin d'avoir toutes identifiées et comprises. Le rétablissement de l'équilibre compositionnel par transplantation fécale pour un « bon » fonctionnement de microbiotes a été testé particulièrement dans le traitement contre l'infection par *Clostridium difficile* [41, 47] (→) mais demande à être confirmé sur le long terme [42]. Certes, nous connaissons certaines causes de dysbiose, telles que les traitements antibiotiques. Les analyses par la métagénomique notamment quantitative [23, 43] ont aussi permis de confirmer des hypothèses antérieures. Mais qu'avons-nous compris au bout du compte sur le fonctionnement d'un microbiote ?

(→) Voir la Synthèse de J.C. Lagier et D. Raoult, page 991 de ce numéro

Dans notre vision finaliste et anthropocentrique, nous avons tendance à attribuer des fonctions ou des rôles aux microbiotes, telles que des éco-services ou la participation au maintien en bon état général. Il est notoire que les fonctions assurées par les composants des microbiotes symbiotiques ont une répercussion profonde sur le bien-être et le développement de leur hôte. Ils agissent en particulier sur (1) le développement et la structuration de leur habitat ; (2) la décomposition, la transformation, la détoxification ou la synthèse de nutriments ; (3) la régulation de l'expression des gènes de l'hôte ; et (4) l'immunité et la protection, jouant le rôle de la première ligne de défense et de protection dans le mucus et à la surface des épithéliums contre l'invasion d'agents pathogènes [48, 49] (→).

(→) Voir la Synthèse de V. Gaboriau-Routhiau et N. Cerf-Bensussan, page 961 de ce numéro, et la Nouvelle de B. Chassaing, m/s n° 4, avril 2015, page 355

## Conclusion

À ce jour, les microbiotes ont surtout été l'objet d'études moléculaires massives qui ont établi des inventaires impressionnants. Mais que savons-nous faire de ces catalogues ? Peu importe diront certains, la manipulation de microbiotes est à portée de main. En matière de santé peut-être, et encore s'agit-il de transplantation de flores « quasi naturelles ». Une manipulation peu contrôlée, où l'on ne sait guère ce qui se passe [47] (→).

(→) Voir la Synthèse de J.C. Lagier et D. Raoult, page 991 de ce numéro

Dans les systèmes très ouverts comme le sol ou les océans nous sommes loin de vrais succès. Le bilan de l'utilisation de micro-organismes en dépollution est très mitigé. Gardons à l'esprit que l'adéquation d'un micro-organisme à un microbiote est la résultante des propriétés de cet organisme et des propriétés de tous les autres organismes susceptibles d'interagir avec lui, directement ou indirectement. De réelles chances de progrès se dessineront lorsque l'écologie microbienne s'appuiera sur des approches expérimentales destinées à tester des prédictions et des hypothèses clairement formulées et dont les résultats se prêteront à des interprétations sans ambiguïtés. ♦

## SUMMARY

### Microbiotes and metagenomics

Major technical advances were introduced in the study of microflora and microbial communities in the nineties. These are essentially analytical approaches conducted as frequently by brute force. What conclusions can be drawn from these analyses twenty years later? In terms of microbial compositions, monitoring and

diagnosis, the results are impressive. But what do all these microorganisms do? What are the factors behind their associations and their maintenance? Gene inventories are approaching completion; they allow us to deepen some physiological processes but do not say much more. We have even begun to manipulate microbiota. But as often in the dialectic between theory and practice, it works but we are far from understanding how! ♦

## LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

## RÉFÉRENCES

- Whitman WB, Coleman DC, Wiebe WJ. Prokaryotes: the unseen majority. *Proc Natl Acad Sci USA* 1998 ; 95 : 6578-83.
- Iriberry J, Herndl GJ. Formation and microbial utilization of amorphous aggregates in the sea: Ecological significance. *Microbiologia* 1995 ; 11 : 309-22.
- Staley JT, Konopka A. Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annu Rev Microbiol* 1985 ; 39 : 321-46.
- Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: Proposal for the domains archaea, bacteria, and eucarya. *Proc Natl Acad Sci USA* 1990 ; 87 : 4576-9.
- Olsen GJ, Lane DJ, Giovannoni SJ, et al. Microbial ecology and evolution: a ribosomal RNA approach. *Annu Rev Microbiol* 1986 ; 40 : 337-65.
- Romby P, Marzi S, Westhof E. La structure atomique du ribosome en pleine lumière. Prix Nobel de Chimie 2009 à Venkatraman Ramakrishnan, Thomas A. Steitz et Ada E. Yonath. *Med Sci (Paris)* 2009 ; 25 : 977-81.
- Dauga C, Doré J, Sghir A. La diversité insoupçonnée du monde microbien. *Med Sci (Paris)* 2005 ; 21 : 290-6.
- Amann RL, Ludwig W, Schleifer KH. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev* 1995 ; 59 : 143-69.
- Colwell RR. Viable but nonculturable bacteria: a survival strategy. *J Infect Chemother* 2000 ; 6 : 121-5.
- Tanaka Y, Hanada S, Manome A, et al. *Catellibacterium nectariphilum* gen. Nov., sp. Nov., which requires a diffusible compound from a strain related to the genus *Sphingomonas* for vigorous growth. *Int J Syst Evol Microbiol* 2004 ; 54 : 955-9.
- Lennon JT, Jones SE. Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat Rev Microbiol* 2011 ; 9 : 119-30.
- Perna NT, Plunkett G, 3rd, Burland V, et al. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* 2001 ; 409 : 529-33.
- Daubin V, Abby SS. Les transferts horizontaux de gènes et l'arbre de la vie. *Med Sci (Paris)* 2012 ; 28 : 695-8.
- Vernikos G, Medini D, Riley DR, Tettelin H. Ten years of pan-genome analyses. *Curr Opin Microbiol* 2015 ; 23 : 148-54.
- Handelsman J. Metagenomics: application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* 2004 ; 68 : 669-85.
- Burcelin R, Chabo C, Blasco-Baque V, et al. Le microbiote intestinal à l'origine de nouvelles perspectives thérapeutiques pour les maladies métaboliques ? *Med Sci (Paris)* 2013 ; 29 : 800-6.
- Cani PD, Everard A. *Akkermansia muciniphila* : une nouvelle cible pour contrôler l'obésité, le diabète de type 2 et l'inflammation ? *Med Sci (Paris)* 2014 ; 30 : 125-7.
- Andrejak C, Delhaes L. Le microbiome pulmonaire en 2015 : une fenêtre ouverte sur les pathologies pulmonaires chroniques. *Med Sci (Paris)* 2015 ; 31 : 971-8.
- Audebert C, Hot D, Lemoine Y, Caboche S. Le séquençage haut-débit : vers un diagnostic basé sur la séquence complète du génome de l'agent infectieux. *Med Sci (Paris)* 2014 ; 30 : 1144-51.
- El Kaoutari A, Armougom F, Raoult D, Henricsson B. Le microbiote intestinal et la digestion des polysaccharides. *Med Sci (Paris)* 2014 ; 30 : 259-65.
- Sharpton TJ. An introduction to the analysis of shotgun metagenomic data. *Front Plant Sci* 2014 ; 5 : 209.
- Handelsman J, Rondon MR, Brady SF, et al. Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products. *Chem Biol* 1998 ; 5 : R245-9.
- Qin J, Li R, Raes J, et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 2010 ; 464 : 59-65.
- Lozupone CA, Stombaugh JJ, Gordon JL, et al. Diversity, stability and resilience of the human gut microbiota. *Nature* 2012 ; 489 : 220-30.
- Li J, Jia H, Cai X, et al. An integrated catalog of reference genes in the human gut microbiome. *Nat Biotechnol* 2014 ; 32 : 834-41.
- Sunagawa S, Coelho LP, Chaffron S, et al. Ocean plankton. Structure and function of the global ocean microbiome. *Science* 2015 ; 348 : 1261359.
- Yooseph S, Sutton G, Rusch DB, et al. The Sorcerer II global ocean sampling expedition: expanding the universe of protein families. *PLoS Biol* 2007 ; 5 : e16.
- Jetten MS, Strous M, van de Pas-Schoonen KT, et al. The anaerobic oxidation of ammonium. *FEMS Microbiol Rev* 1998 ; 22 : 421-37.
- Ettwig KF, Butler MK, Le Paslier D, et al. Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. *Nature* 2010 ; 464 : 543-8.
- Kyrpides NC. Fifteen years of microbial genomics: meeting the challenges and fulfilling the dream. *Nat Biotechnol* 2009 ; 27 : 627-32.
- Turroni F, Marchesi JR, Foroni E, et al. Microbiomic analysis of the bifidobacterial population in the human distal gut. *Isme J* 2009 ; 3 : 745-51.
- Aslam Z, Yasir M, Yoon HS, et al. Diversity of the bacterial community in the rice rhizosphere managed under conventional and no-tillage practices. *J Microbiol* 2013 ; 51 : 747-56.
- Nam YD, Chang HW, Kim KH, et al. Bacterial, archaeal, and eukaryal diversity in the intestines of Korean people. *J Microbiol* 2008 ; 46 : 491-501.
- Bartram AK, Lynch MD, Stearns JC, et al. Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end illumina reads. *Appl Environ Microbiol* 2011 ; 77 : 3846-52.
- Gloor GB, Hummelen R, Macklaim JM, et al. Microbiome profiling by illumina sequencing of combinatorial sequence-tagged PCR products. *PLoS One* 2010 ; 5 : e15406.
- Sogin ML, Morrison HG, Huber JA, et al. Microbial diversity in the deep sea and the underexplored rare biosphere. *Proc Natl Acad Sci USA* 2006 ; 103 : 12115-20.
- Gibson GR, Cummings JH, Macfarlane GT. Use of a three-stage continuous culture system to study the effect of mucin on dissimilatory sulfate reduction and methanogenesis by mixed populations of human gut bacteria. *Appl Environ Microbiol* 1988 ; 54 : 2750-5.
- Jarrell KF, Walters AD, Bochiwal C, et al. Major players on the microbial stage: why archaea are important. *Microbiology* 2011 ; 157 : 919-36.
- St-Pierre B, Wright AD. Diversity of gut methanogens in herbivorous animals. *Animal* 2013 ; 1 : 49-56.
- Ohkuma M. Symbioses of flagellates and prokaryotes in the gut of lower termites. *Trends Microbiol* 2008 ; 16 : 345-52.
- Smits LP, Bouter KE, de Vos WM, et al. Therapeutic potential of fecal microbiota transplantation. *Gastroenterology* 2013 ; 145 : 946-53.
- Agrawal M, Aroniadis OC, Brandt LJ, et al. The long-term efficacy and safety of fecal microbiota transplant for recurrent, severe, and complicated *Clostridium difficile* infection in 146 elderly individuals. *J Clin Gastroenterol* 2016 ; 50 : 403-7.
- Le Chatelier E, Nielsen T, Qin J, et al. Richness of human gut microbiome correlates with metabolic markers. *Nature* 2013 ; 500 : 541-6.
- Lederberg J, McCray AT. Ome sweet omics: a genealogical treasury of words. *Scientist* 2001 ; 15 : 8.
- Burcelin R, Nicolas S, Blasco-Baque V. Microbiotes et maladies métaboliques : de nouveaux concepts pour de nouvelles stratégies thérapeutiques. *Med Sci (Paris)* 2016 ; 32 : 952-60.
- Lagier JC, Raoult D. *Culturomics* : une méthode d'étude du microbiote humain. *Med Sci (Paris)* 2016 ; 32 : 923-5.
- Lagier JC, Raoult D. Greffe de microbiote fécal et infections : mise au point, perspectives. *Med Sci (Paris)* 2016 ; 32 : 991-7.
- Gaboriau-Routhiau V, Cerf-Bensussan N. Microbiote intestinal et développement du système immunitaire. *Med Sci (Paris)* 2016 ; 32 : 961-7.
- Chassaing B. Le microbiote intestinal, un acteur de la réponse immunitaire adaptative antivirale ? *Med Sci (Paris)* 2015 ; 31 : 355-7.

TIRÉS À PART

J. Weissenbach