

DES PANGÉNOMES DES MICROORGANISMES AUX MICROBIOTES

par David **VALLENET**
(Genoscope)

Les micro-organismes présentent la plus grande biodiversité avec l'histoire évolutive la plus longue. Ceci se traduit par une importante dynamique de leurs génomes dont le contenu en gènes est très variable même parmi les organismes d'une même espèce. Cette plasticité génomique, qui est le résultat de nombreux échanges par transferts horizontaux, explique la capacité des microbes à être pathogènes ou à s'adapter à des environnements très diversifiés voire extrêmes.

De par les progrès considérables des technologies de séquençage, nous accédons actuellement à des millions de séquences de génomes complets dans les banques publiques. Comparer toutes ces séquences pour en donner du sens biologique constitue un challenge pour la bioinformatique. Ainsi, le concept de pangénome qui regroupe l'ensemble des séquences partagées par les génomes d'une espèce offre de nouvelles perspectives pour analyser ces données massives. La modélisation de ces données sous forme de graphes permet de capturer l'intégralité du contenu en gènes tout en conservant l'information de voisinage chromosomique qui est un élément clé pour comprendre l'expression des gènes et le flux des transferts horizontaux. A partir de ces graphes de pangénomes, des méthodes sont développées pour comparer des milliers de génomes en quelques minutes pour identifier des régions variables et des groupes de gènes potentiellement liés fonctionnellement.

L'utilisation de ces graphes de pangénomes pour l'étude des microbiotes devrait permettre de mieux comprendre la dynamique des espèces dans leurs écosystèmes.



Directeur de recherche CEA, Genoscope UMR8030 Génomique Métabolique

Direction de l'équipe LABGeM (<https://labgem.genoscope.cns.fr>) qui conduit des développements et analyses bio-informatiques pour l'étude des génomes de micro-organismes et de leur métabolisme. Concernant la pangénomique, l'équipe développe plusieurs méthodes autour de la suite logicielle PPanGGOLiN (<https://github.com/labgem/PPanGGOLiN>) pour représenter les génomes sous la forme de graphes de pangénomes permettant l'étude des régions de plasticité (panRGP) et de modules conservés (panModule).

Responsable scientifique de la plateforme MicroScope (<https://mage.genoscope.cns.fr/microscope>) qui offre à une communauté internationale de microbiologistes (>6000 comptes utilisateurs) un environnement intégré web pour l'annotation des génomes procaryotes.