



Université Claude Bernard



DIPLÔME NATIONAL DE DOCTORAT

(Arrêté du 25 mai 2016)

Date de la soutenance : **13 novembre 2018**

Nom de famille et prénom de l'auteur : **AOUAD Monique**

Titre de la thèse : « *Etude phylogénomique de l'histoire évolutive des archées et leur lien avec l'eucaryogenèse* »



Résumé

L'explosion des données de séquençage a permis d'identifier et d'analyser des gènes conservés, permettant de résoudre la plupart des relations phylogénétiques chez les archées. Ceci constitue un cadre essentiel pour comprendre leur évolution et leur adaptation à une grande variété d'environnements, ainsi que leur lien évolutif avec les deux autres domaines du vivant, les bactéries et les eucaryotes. Néanmoins, de nombreuses questions restent à résoudre à l'échelle du domaine des archées et à l'échelle des trois domaines du vivant. Parmi elles, les relations phylogénétiques au sein du cluster II, notamment la position des lignées halophiles extrêmes qui ont été placées à différentes positions dans l'arbre en fonction des marqueurs et des modèles de reconstruction utilisés, ainsi que la position de la racine des archées et la position des eucaryotes à la lumière des lignées d'archées nouvellement séquencées.

Au cours de ma thèse, j'ai contribué à (i) affiner la phylogénie du domaine des Archaea en se concentrant sur les relations phylogénétiques au sein du cluster II, en particulier les positions des lignées halophiles extrêmes à travers des analyses dédiées à cette partie spécifique de l'arbre, et (ii) établir une phylogénie globale des archées afin de comprendre leur histoire évolutive ancienne et leur lien avec les eucaryotes à travers une analyse phylogénomique en deux étapes à l'échelle des trois domaines du vivant.

D'abord, en utilisant des approches de génomique comparée sur 155 génomes complets appartenant aux Halobacteria, Nanohaloarchaea, méthanogènes de classe II, Archaeoglobales et Diaforarchaea, j'ai identifié 258 protéines portant un signal phylogénétique fiable pour étudier les relations de parenté au sein du cluster II. En combinant différentes approches limitant l'impact du signal non phylogénétique sur l'inférence phylogénétique (comme la méthode Slow-Fast et le recodage des acides aminés), j'ai montré que les Nanohaloarchaea branchent avec les Methanocellales et les Halobacteria branchent avec les Methanomicrobiales. Ce jeu de données a ensuite été utilisé pour étudier la position d'une troisième lignée halophile extrême, les Methanonatronarchaeia, qui se positionnent entre les Archaeoglobales et les Diaforarchaea. Ces résultats impliquent que l'adaptation à la salinité extrême est apparue au moins trois fois de manière indépendante chez les archées et que les similitudes phénotypiques observées chez les Nanohaloarchaea, Halobacteria et Methanonatronarchaeia résultent d'une convergence évolutive, éventuellement accompagnée de transferts de gènes horizontaux. Enfin, ces résultats suggèrent que le groupement des Nanohaloarchaea avec d'autres lignées des DPANN serait la conséquence d'un artéfact de reconstruction.

Pour la deuxième partie de ma thèse, j'ai appliqué une stratégie consistant à analyser séparément les trois domaines du vivant considérés deux à deux, en mettant à jour 72 familles protéiques précédemment identifiées par Raymann et al. (2015) pour inclure toutes les nouvelles lignées d'archées séquencées depuis la publication de cette étude comme les Asgard,

les DPANN, les Stygia, les Acherontia, etc. Au total, mon échantillonnage taxonomique comprend 435 archées, 18 eucaryotes et 67 bactéries. Les résultats des analyses par la méthode Slow-Fast soutiennent une racine des Archaea située entre le superphylum basal des DPANN et le reste des archées séparées en cluster I et cluster II, comme décrit par Raymann et al. (2015), et montrent que la monophylie des Euryarchaeota est liée aux positions évoluant vite. Mes résultats placent les eucaryotes en tant que groupe frère du superphylum des TACK et montrent que leur regroupement avec les Asgard est liée aux positions évoluant vite. Ces résultats ont des implications majeures sur les inférences de la nature du dernier ancêtre commun des archées et sur l'histoire évolutive de ce domaine qui a conduit à l'apparition de la première cellule eucaryote.