

Reconstruction de réseaux phylogénétiques à structure arborée depuis un ensemble de clusters

Daniel H. Huson* Regula Rupp* Vincent Berry† Philippe Gambette†
Christophe Paul†

20 décembre 2008

Résumé

Les réseaux phylogénétiques généralisent les arbres phylogénétiques en représentant des événements d'hybridation, de recombinaison ou de transferts horizontaux par des parties réticulées. Des méthodes et logiciels ont été développés pour les reconstruire à partir de séquences ou d'arbres, voire triplets, quadruplets, clusters et bipartitions [1].

Nous présentons ici une méthode permettant d'obtenir à partir d'un ensemble de clusters des réseaux phylogénétiques enracinés avec une structure arborée ("galled networks") [3]. Ces réseaux permettent un bon compromis entre la représentativité biologique et la complexité théorique. Nous montrons que tout ensemble de clusters peut être représenté par un réseau à structure arborée (contrairement à d'autres sous-classes de réseaux phylogénétiques, comme les "galled trees" ou réseaux de niveau k). De plus, il est possible de décider en temps polynomial si un cluster est représenté dans un réseau à structure arborée, alors que le problème général est NP-complet [4].

Nous proposons une approche de reconstruction en deux étapes, qui consiste à éliminer un nombre minimal de feuilles pour que les clusters deviennent compatible avec un arbre, puis attacher ces feuilles élaguées en utilisant un nombre minimal d'arêtes. Bien que ces deux étapes correspondent à des problèmes NP-complets, leur implémentation à l'aide d'un algorithme FPT et d'une approche par séparation et évaluation aboutit en quelques secondes pour des clusters faisant intervenir des centaines de taxons. Cette méthode est disponible dans le logiciel Dendroscope [2].

Références

- [1] Daniel H. Huson et David Bryant. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23(2):254–267, 2006. Programme SplitsTree disponible sur <http://www.splitstree.org>.
- [2] Daniel H. Huson, Tobias DeZulian, Markus Franz, Christian Rausch, Daniel C. Richter, et Regula Rupp. Dendroscope - an interactive tree drawer. *BMC Bioinformatics*, 8(460), 2007. Programme Dendroscope disponible sur <http://www.dendroscope.org>.
- [3] Daniel H. Huson et Tobias Klopper. Beyond galled trees - decomposition and computation of galled networks. In *Proceedings of the eleventh Annual International Conference on Research in Computational Molecular Biology (RECOMB'07)*, volume 4453 of *Lecture Notes in Computer Science*, pages 211–227. Springer Verlag, 2007.
- [4] Iyad A. Kanj, Luay Nakhleh, Cuong Than, et Ge Xia. Seeing the trees and their branches in the network is hard. *Theoretical Computer Science*, 401:153–164, 2008.

* Center for Bioinformatics ZBIT, Tübingen University, Sand 14, 72076 Tübingen, Germany

† Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier (LIRMM, UMR 5506, CNRS), Université Montpellier II, 161, rue Ada, 34392 Montpellier Cedex 5, France