Sujet M2 2015-2016

Intitulé du sujet : Exploration de la face cachée de la recombinaison: mutation, conversion et évolution des génomes animaux

Directeur: Laurent DURET (Laurent.Duret@univ-lyon1.fr) LBBE – UMR 5558

Résumé du projet:

La caractérisation des éléments fonctionnels d'un génome repose sur l'identification des traces laissées par l'action de la sélection naturelle. Classiquement, trois processus sont pris en considération pour interpréter l'évolution des séquences: la mutation, la sélection et la dérive génétique aléatoire. Cependant, depuis quelques années, nos travaux ont permis de montrer qu'**un quatrième processus joue un rôle majeur dans l'évolution des génomes: la conversion génique biaisée (gBGC).** Chez les eucaryotes, ce processus est intrinsèquement lié à la recombinaison méiotique: la présence de sites polymorphes au sein des hétéroduplex formés entre les deux chromosomes parentaux conduit à l'apparition de mésappariements, dont la réparation aboutit à la conversion d'un des deux allèles par l'autre. Il a été montré chez la levure et chez l'homme, que les mésappariements AT|GC sont préférentiellement réparés en paire de base G:C plutôt que A:T [1,2]. De ce fait, les allèles G:C ont une probabilité plus forte d'être transmis à la génération suivante. Ce processus augmente donc la probabilité de fixation des allèles G:C dans les populations, et peut même contrer les effets de la sélection et conduire à la fixation de mutations délétères [3].

Nous avons montré que ce processus a un impact majeur sur l'évolution des génomes de mammifères [3] et qu'il affecte de nombreux autres organismes eucaryotes [4]. Il semble cependant que l'intensité du gBGC varie fortement entre espèces, même relativement proches. Par exemple, le gBGC est très fort chez l'abeille [5] mais très faible (voire nul) chez la drosophile [6]; de même le gBGC paraît beaucoup plus intense chez les mammifères et les oiseaux que chez les autres vertébrés [7]. Pourquoi le gBGC est-il intense dans certains taxons mais pas dans d'autres? Et quelle est la *raison d'être* de ce processus? Pourquoi tant d'espèces montrent-elles un biais de conversion en faveur des allèles GC?

Pour tenter de répondre à ces questions, nous proposons de quantifier l'intensité du gBGC ainsi que différents paramètres génomiques clés (taux de recombinaison, taux de mutation, patron de mutation) dans un grand nombre d'espèces animales. Nous pourrons ainsi explorer les relations entre ces différents paramètres, et leur co-variations avec différentes caractéristiques génomiques ainsi que des traits d'histoire de vie (longévité, masse, métabolisme, mode de sexualité, …) pour tenter de comprendre les contraintes qui gouvernent l'évolution du gBGC.

Ce projet se basera sur l'analyse de données de polymorphisme pan-génomique collectées sur 76 espèces, représentatives des principaux phylums animaux. Il impliquera d'analyser de gros volumes de données (RNAseq, polymorphisme). L'étudiant devra développer des scripts pour gérer et analyser les données, mettre en œuvre des outils bioinformatiques d'analyse de séquences, et des outils statistiques d'analyse de données (R). L'étudiant utilisera également les concepts de la génomique des populations pour quantifier le gBGC et les processus mutationnels.

Références:

[1] Mancera et al. (2008). *Nature*, 454:479–85.
[2] Williams et al. (2015). *eLife (in press).*
[3] Duret &Galtier (2009). *Ann. Rev. Genom. Hum. Genet*. 10:285–311.
[4] Pessia et al. (2012). *Genome Biol. Evol*, 4:675–82.
[5] Kent et al. (2012). *Proc. Natl. Acad. Sci. U S A 109*, 18012–18017.
[6] Robinson et al. (2014). *Mol. Biol. Evol. 31*, 425–33.
[7] Escobar et al. (2011). *Mol. Biol. Evol. 28*, 2561–2575.