

Statistiques sur les groupes et familles de gènes

David Sankoff

Département de mathématiques et de statistique

Université d'Ottawa

585 King Edward Ave.

Ottawa, Ontario

CANADA K1N 6N5

Résumé

Une approche importante permettant d'étudier les forces guidant l'organisation des génomes et leur évolution consiste à comparer des ordres de gènes entre deux ou plusieurs espèces apparentées. Les groupes de gènes similaires identifiés dans des génomes sont souvent utilisés pour prouver que les génomes sont proches d'un point de vue évolutif et fonctionnel. Cependant, l'ordre des gènes est progressivement modifié au cours de l'évolution suite à des réarrangements de grande envergure, des transferts horizontaux, des duplications et des pertes de gènes. De ce fait, il devient de plus en plus difficile de déterminer si les similarités observées dans les génomes à un niveau local sont réellement les traces de l'ordre des gènes d'un ancêtre commun, ou simplement le fruit du hasard. C'est pourquoi il est nécessaire de développer des méthodes permettant de distinguer, parmi les régularités dans un ou plusieurs génomes, celles dues au hasard de celles dues à une histoire ou une fonctionnalité commune.

Dans cet exposé, nous présentons des tests permettant de distinguer des groupements significatifs de gènes par rapport à des groupements aléatoires. Cette étude prend en compte les groupements incomplets, les génomes multichromosomaux ainsi que les familles de gènes. Nous évaluons aussi bien la pertinence des groupements individuels de gènes pré-spécifiés que le degré global de groupement dans un génome entier.