

## Modèles mutationnels en évolution moléculaire intraspécifique

approche analytique, statistique et bioinformatique

<http://www.biologie.ens.fr/modmut/>

### CONTEXTE

Notre projet s'inscrit dans la recherche fondamentale en **génomique comparative** et concerne l'étude de la variabilité génétique entre les individus d'une même population ou espèce (**polymorphisme**) voire entre espèces (**divergence**). Il entre dans le cadre des tentatives de rapprochement entre ces deux niveaux d'étude, qui sont encore aujourd'hui abordés largement indépendamment, avec des approches sensiblement différentes.

La génomique comparative est très développée essentiellement au niveau interspécifique avec les analyses phylogénomiques. Le choix du modèle mutationnel est déterminant dans ce type d'analyse. Au contraire, sur de petites échelles taxonomiques (intraspécifiques), l'approche est très différente. A cette échelle, les modèles mutationnels font des approximations très simples avec des mutations indépendantes, à un taux constant au cours du temps et le long d'une séquence. Les écarts aux modèles mutationnels sont rarement considérés lors de tests de scénarii évolutifs, alors que de résultats récents indiquent qu'ils biaisent sensiblement les analyses [0].

[0] Baudry, E., and F. Depaulis. Effect of misoriented sites on neutrality tests with outgroup. *Genetics* **165**:1619-22, 2003.

### OBJECTIFS

Sur la base de la problématique des modèles mutationnels, nous visons à réassocier les domaines de la variabilité de séquence inter- et intra-spécifique, encore largement indépendants car concernant des échelles de temps très différentes. Il s'agit ici d'aborder cette question du modèle mutationnel sur une échelle taxonomique réduite autour du niveau spécifique.

### PROJET

- Tester les hypothèses mutationnelles classiques à la base des traitements des données en phylogénie moléculaire et analyse du polymorphisme nucléotidique.
- Comprendre et modéliser les processus de mutation et fixation à une échelle populationnelle. Etablir en particulier un modèle mutationnel adapté à la D-loop de l'ADNmt. En tirer également des méthodes de détection de régions fonctionnelles dans les régions non codantes.
- Corriger les méthodes d'analyse de données par prise en compte d'un modèle mutationnel adéquat.

### RESULTATS

- Probabilités de fixation en modèles écologiques impliquant de la densité dépendance et différentes stratégies en terme de nombre moyen,  $v$  variance du nombre de descendants et capacités compétitives [1, 2]
- ADN ancien et génétique des populations: effets de données étalonnées dans le temps sur les analyses, effets mutationnels et apparente signature de la recombinaison [3, Fig2].
- Détection des effets d'auto-stop par les tests haplotypiques [4].
- Estimations des taux de mutation de microsatellites en population expérimentales de blé [5].
- Microsatellites et modèles mutationnels [6].
- Effet de la démographie sur la fréquence des mutations délétères [7].

### PUBLICATIONS PRINCIPALES

- [1] Lambert A. The branching process with logistic growth. *Ann. Appl. Prob.* 2005 vol.15 (2) 1506-1535.
- [2] Lambert, A. Probability of fixation under weak selection: a branching process unifying approach. (Soumis).
- [3] Depaulis, F., L. Orlando, V. Ung, and C. Hänni. Population genetics of temporally spaced data: an application on cave bear ancient DNA (Encouragé à resoumission pour *Mol Biol Evol*).
- [4] Depaulis, F., S. Mousset, and M. Veuille 2005. Detecting selective sweeps with haplotype tests. In: *Nurminsky D (ed) Selective Sweep. Landes Bioscience, Georgetown, USA*
- [5] Raquin A. L., F. Depaulis, A. Lambert, N. Galic, P. Brabant, and I. Goldringer. Microsatellites evolution in an experimental wheat population: detection of mutation and derivation of coalescent-related methods to assess mutation rates (Soumis à *Genetics*).
- [6] Balaresque P, Toupance B, Quintana-Murcilluis, Crouau-Roy B, Heyer E. Sex-specific selection on the human X chromosome? *Genet Res.* 83:169-76, 2004
- [7] Heyer E, Sibert A, Austerlitz F. Cultural transmission of fitness: genes take the fast lane. *Trends Genet.* 21(4):234-9, 2005

### LABORATOIRES IMPLIQUES (Paris)

UMR 7625 (CNRS-ENS- P6) Equipe Eco-Evolution Mathématique  
DYOGEN (CNRS UMR8541-ENS) Dynamique et Organisation des Génomes  
UMR 5145 (CNRS/MNHN/P7) Equipe Génétique des Populations Humaines

### MOTS CLES

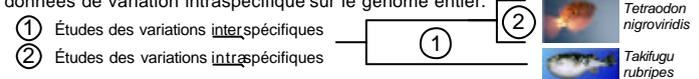
Modèles mutationnels, duplications, D-loop, recombinaison, taux d'évolution, Coalescence, processus de branchements.

### CALENDRIER

- Octobre / Novembre (Paris) Démarrage
- Réunion ACTI les 10/11/2004, 3/5/2005 et
- 8/6/2005: modèle mutationnels adapté à la D-loop.
- 9/2005: Voyage d'étude à l'université Edimbourg ICAPB: projets flux géniques vs homoplasies (UMR7625)

### POLYMORPHISME ET DIVERGENCE

Les génomes des poissons tétraodontiformes présentent un contexte exceptionnel pour étudier conjointement la variation génétique due au polymorphisme intraspécifique et la divergence interspécifique. Les tétraodontiformes *Takifugu rubripes* et *Tetraodon nigroviridis* ont divergé relativement récemment il y a environ 20-30 millions d'années. Leurs génomes ont été entièrement séquencés, au Genoscope (*Tetraodon* [8]) et au Joint Genome Institute (*Takifugu*). De plus, le génome de *Tetraodon* a été séquencé en quantité équivalente à partir de deux individus sauvages, fournissant ainsi des données de variation intraspécifique sur le génome entier.



Pour réaliser ce projet, nous analysons d'une part les lectures originales des assemblages des deux poissons pour y relever les bases polymorphiques, et d'autre part l'alignement des deux génomes pour y documenter les positions ayant divergé au cours de l'évolution. Ces données primaires nous permettront de développer de nouveaux modèles pour prendre en compte le taux de polymorphisme et le taux de mutation neutre des gènes d'une région. Une application de ce projet sera d'identifier de manière fiable les régions non-codantes des régions intergéniques soumises à une sélection négative et donc potentiellement fonctionnelles.

[8] Jaillon et al. (2004) 1. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* 431:946-957

### D-LOOP: RECOMBINAISON OU MUTATIONS

Le déséquilibre de liaison (association statistique entre les mutations/SNPs, DL) décroît avec la distance séparant les mutations sur l'ADNmt humain [9] (Fig. 1). Cette observation a été interprétée comme le résultat de la recombinaison.

Or des modèles mutationnels impliquant une agrégation spatio-temporelle des mutations peuvent rendre compte de cette corrélation (Fig. 2; [3]) et représentent donc une alternative vraisemblable à la présence de recombinaison.

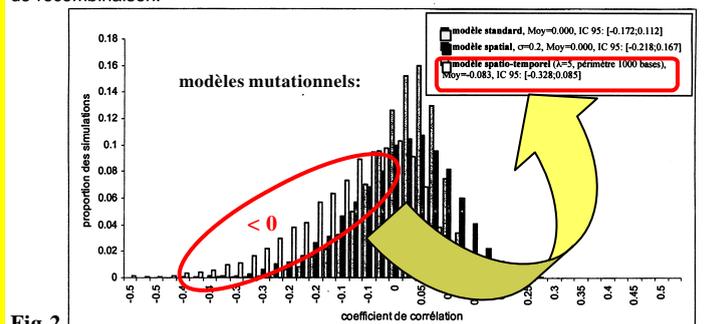
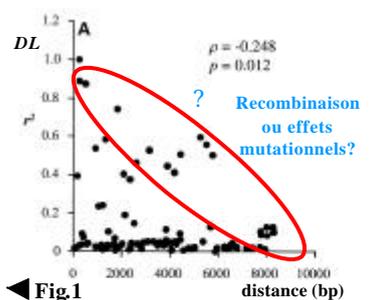


Fig. 2

[9] Awadalla, P. A. Eyre-Walker and J.M. Smith. Linkage disequilibrium and recombination in hominid mitochondrial DNA. *Science* **287**:2524-5 (1999).