

Genome size evolution within *Drosophila*

Contact

Annabelle Haudry, maître de conférence Université Lyon 1

annabelle.haudry@univ-lyon1.fr

Tel. +33 (0) 472 432 918

Lab

Laboratoire de Biométrie et Biologie Evolutive -Université Lyon 1 - UMR CNRS 5558

43 bd du 11 novembre 1918

69622 Villeurbanne cedex FRANCE

Project

Context

Genome size can vary up to 200,000 fold among eukaryotes genomes. Such variations can not be fully explained by changes in the level of "complexity" of species (estimated via the number of genes), as it was originally suggested¹. Sequencing projects of the past two decades have revealed that the differences between genomes were mainly due to differences in their content in repeated DNA. Transposable elements (TEs) are mobile segments of DNA capable of being cut or copied in the genome, that make them potentially highly repeated. TEs represent a very large proportion of Eukaryotic genomes in general (~ 50% of the human genome) and their content in the genome is positively correlated with the genome size². The evolutionary history (leading to changes in demography) and life history traits of species are major factors affecting the intensity of genetic drift acting within a species³, and therefore affecting the efficacy of selection in this species⁴. The model proposed by Lynch and Conery⁵ predicts an accumulation of TEs, and thus larger genomes, in species of small effective population size, in which the genetic drift is strong, due to a loosening of the selection against deleterious TE insertions.

Objectives

To test this hypothesis on closely related species, we want to study both environmental and genomic data for 25 *Drosophila* species whose genome has already been sequenced. These species show significant variations in their ecologies and in their genome sizes. By combining molecular features calculated on the sequences (dn / ds, TE content) with the phylogenetic relationships between species (reconstructed from identified orthologous genes), the objective consists in reconstructing the dynamics of proliferation and elimination of TE over time⁶⁻⁸. These results should provide a better understanding of the ecological factors affecting changes in genome sizes.

Skills

This project implies using bioinformatics tools to analyze whole-genome data. The knowledge of some programming languages will be required for file manipulation (perl, python or C) and statistics analysis (R).

1. Mirsky, A. E. & Ris, H. The desoxyribonucleic acid content of animal cells and its evolutionary significance. *J. Gen. Physiol.* 451–462 (1951).
2. Biémont, C. & Vieira, C. Junk DNA as an evolutionary force. *Nature* **443**, 521–524 (2006).
3. Haudry, A. *et al.* Grinding up Wheat: A Massive Loss of Nucleotide Diversity Since Domestication. *Mol. Biol. Evol.* **24**, 1506–1517 (2007).
4. Haudry, A. *et al.* Mating system and recombination affect molecular evolution in four Triticeae species. *Genet. Res.* **90**, 97–109 (2008).
5. Lynch, M. & Conery, J. S. The origins of genome complexity. *Science* **302**, 1401–4 (2003).
6. Hawkins, J. S., Proulx, S. R., Rapp, R. A. & Wendel, J. F. Rapid DNA loss as a counterbalance to genome expansion through retrotransposon proliferation in plants. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 17811–6 (2009).
7. Sun, C. *et al.* LTR retrotransposons contribute to genomic gigantism in plethodontid salamanders. *Genome Biol. Evol.* **4**, 168–83 (2012).
8. Bennetzen, J. L., Ma, J. & Devos, K. M. Mechanisms of recent genome size variation in flowering plants. *Ann. Bot.* **95**, 127–32 (2005).