Comparative genomics of gene families in relation with metabolic pathways for gene candidates highlighting

Delphine Larivière [*][†][1], David Couvin [*][‡][1], Gaëtan Droc[1], Jonathan Lorenzo[2], Valentin Guignon[3], Mathieu Rouard[3], Eric Lyons[4], Dominique This[5], Stéphanie Bocs[6], Jean-François Dufayard[6]

Poster 88

[1] Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD) – Avenue Agropolis, Montpellier, France
[2] Institut Français de Bioinformatique (IFB) – Gif-sur-Yvette, Île-de-France, France
[3] Bioversity International – Consultative Group on International Agricultural Research [CGIAR] – Parc Scientifique Agropolis II, F-34 397 Montpellier Cedex 5, France
[4] University of Arizona – University of Arizona Tucson AZ 85721 USA/États-Unis
[5] Montpellier SupAgro – Montpellier SupAgro – 2 place Pierre Viala, F-34 060 Montpellier Cedex 02, France
[6] CIRAD UMR AGAP (AGAP) – Montpellier SupAgro, Institut national de la recherche agronomique (INRA) : UMR1334, CIRAD-BIOS – TA A-108/03, Avenue Agropolis, F-34 398 Montpellier Cedex 5, France

The study of gene families is an important field of comparative genomics, allowing, by the analysis of evolutionary history, to identify homology relationships, gene losses and to help in annotation transfers. The addition of metabolic information improves the identification of candidate genes by addition of functional and gene network data. We propose new web systems to facilitate and improve the analysis of gene family for the search of candidate genes in plants. GenFam (`http://genfam.southgreen.fr/`) is dedicated to the manual and precise analysis of gene families and includes specific workflows running under a Galaxy platform, allowing to gather several data sources, analysis and visualization tools, in order to (i) build custom families (ii) run analysis workflows dedicated to the analysis of gene families (iii) visualize analysis results and functional evidences through a dedicated visualization dashboard.

In complement to the integration of data sources, tools and visualizations, we also suggest a new way to find evidences for the identification of evolutionary events through syntenic analyses. The IDEVEN algorithm is based on the study of syntenic blocks linked to a gene family to identify speciation, Whole Genome Duplication (WGD) events, and other duplications in a family history. The identified events will be reported on the phylogeny and aim to bring complementary evidences to have a clearer view of the evolutionary history of a gene family. To extend this tool to the analysis of multiple gene families and integrate metabolic pathways data, this tool has been integrated in genesPath, which will allow a deep identification and highlighting of candidate genes of interest for a specific project called "Biomass For the Future (BFF)". This online tool will be soon available and could be notably used for searching candidate genes involved in biosynthesis of lignin and cellulose in various plant species (such as maize and sorghum).

Mots clefs : gene family, data integration, synteny, evolutive events, comparative genomics, metabolic pathways, phylogeny

---

*. Intervenant
†. Corresponding author : `lariviere.delphine@gmail.com`
‡. Corresponding author : `david.couvin@googlemail.com`