

Analyse des groupes de gènes co-exprimés (AGGC) : un outil automatique pour l'interprétation de données de biopuces

Ricardo Martinez*, Nicolas Pasquier*,
Claude Pasquier**, Martine Collard*, Lucero Lopez***

*Projet Execo
Laboratoire I3S, UNSA
2000, route des lucioles,
06903 Sophia-Antipolis cedex, France
{rmartine, pasquier, mcollard}@i3s.unice.fr

**Laboratoire Biologie Virtuelle, UNSA
Centre de Biochimie, Parc Valrose,
06108 Nice cedex 2, France
claude.pasquier@unice.fr

***Projet Odyssee
INRIA Sophia Antipolis,
2004 route des Lucioles
06905 Sophia Antipolis, France
lucero.lopez@gmail.com

Résumé. La technologie des biopuces permet de mesurer les niveaux d'expression de milliers de gènes dans différentes conditions biologiques générant ainsi des masses de données à analyser. De nos jours, l'interprétation de ces volumineux jeux de données à la lumière des différentes sources d'information est l'un des principaux défis dans la bio-informatique. Nous avons développé une nouvelle méthode appelée AGGC (Analyse des Groupes de Gènes Co-exprimés) qui permet de constituer de manière automatique des groupes de gènes à la fois fonctionnellement riches, i.e. qui partagent les mêmes annotations fonctionnelles, et co-exprimés. AGGC intègre l'information issue des biopuces, i.e. les profils d'expression des gènes, avec les annotations fonctionnelles des gènes obtenues à partir des sources d'information génomiques comme *Gene Ontology*. Les expérimentations menées avec cette méthode ont permis de mettre en évidence les principaux groupes de gènes fonctionnellement riches et co-exprimés dans des expériences de biopuces¹.