

Un outil pour la visualisation de relations entre gènes

Marie Agier*, Jean-Marc Petit**

*LIMOS, UMR 6158 CNRS, Univ. Clermont-Ferrand II

**LIRIS, UMR 5205 CNRS, INSA Lyon

agier@isima.fr, jean-marc.petit@insa-lyon.fr

Résumé. La reconstruction de réseaux de gènes est un des défis majeurs de la post-génomique. A partir de données d'expression issues de puces à ADN, différentes techniques existent pour inférer des réseaux de gènes. Nous proposons dans ce papier une approche pour la visualisation de réseaux d'interactions entre gènes à partir de données d'expression. L'originalité de notre approche est de superposer des règles avec des sémantiques différentes au sein d'un même support visuel et de ne générer que les règles qui impliquent des gènes dits centraux. Ceux-ci sont spécifiés en amont par les experts et permettent de limiter la génération des règles aux seuls gènes qui intéressent les spécialistes. Une implémentation a été réalisée dans le logiciel libre MeV de l'institut TIGR.

1 Introduction

Suite au succès rencontré par les techniques de puces à ADN pour mesurer l'expression des gènes à grande échelle, la **reconstruction de réseaux de gènes** à partir de ces données d'expression a suscité depuis quelques années un intérêt croissant. Dans des travaux antérieurs [1, 2], nous avons proposé une approche ayant pour but de découvrir **différents types de règles entre gènes**. Pour faciliter l'interprétation des règles par les experts, nous proposons dans ce papier une visualisation conviviale des règles générées. Nous montrons comment les règles peuvent être visualisées sous forme de graphe orienté présentant les diverses relations découvertes dans les données. L'originalité de notre proposition est de superposer différents types de règles dans un même support visuel. Nous proposons également aux utilisateurs de spécifier plusieurs gènes dits centraux, à partir desquels seront présentées uniquement les règles impliquant ces gènes centraux et limitant ainsi le coût de la génération des règles.

2 Approche proposée

Nous souhaitons avant tout réaliser un outil convivial et proposer ainsi une méthode de visualisation intuitive pour les experts. D'autre part, nous proposons d'appliquer un filtre sur les règles générées en fonction de cinq indices de qualité (support, confiance, lift, leverage et conviction). Ne seront donc visualisées que les règles les plus intéressantes pour les experts, il est donc suffisant de pouvoir visualiser les indices pour la règle ou l'attribut sélectionnés par un simple clic. L'interprétation des règles est une étape particulièrement délicate et très difficile, puisqu'une règle entre deux gènes implique également divers produits associés (protéines, facteurs de transcription...). C'est pourquoi les biologistes sont rarement intéressés par