研究、探讨

从基因芯片数据快速有效地挖掘共调控基因

赵 倩，尚学群

西北工业大学 计算机学院，西安 710129

摘要　针对基因芯片数据高噪音、列（基因）数比行（实验条件）数多几个数量级的特殊性，为了进一步提高从基因芯片数据挖掘共调控基因的时间效率和挖掘结果的有效性，首先根据所有两两基因对之间的Pearson相关系数对原始完整数据集进行分组，然后使用列（基因）枚举方法对各组数据分别进行闭合频繁模式挖掘，并对活化和抑制共调控关系的挖掘分别进行处理。实验结果证明：算法快速有效地挖掘出了两种共调控基因。

关键词　　基因芯片数据　共调控基因　Pearson相关系数　闭合频繁模式

分类号　TP311

Mining co-regulated genes from microarray data quickly and effectively

ZHAO Qian，SHANG Xue-qun

School of Computer，Northwestern Polytechnical University，Xi'an 710129，China

Abstract

Microarray data sets typically contain strong noise and an order of magnitude more genes than experiments.To further reduce the running time and improve the validity of co-regulated genes mined from microarray data，a new method is proposed which firstly groups all genes according to the Pearson correlation coefficient between every two genes，then uses column（gene）enumeration to mine closed frequent patterns as positive or negative co-regulated genes for each group.The experimental results show that the proposed approach can quickly and effectively mine two kinds of co-regulated genes from microarray data.

Key words　　microarray data　co-regulated genes　Pearson correlation coefficient　closed frequent pattern

通讯作者 赵 倩 zhaoqian_qiezi@126.com

扩展功能

本文信息
- Supporting info
- PDF(1165KB)
- [HTML全文](0KB)
- 参考文献

服务与反馈
- 把本文推荐给朋友
- 加入我的书架
- 加入引用管理器
- 复制索引
- Email Alert
- 文章反馈
- 浏览反馈信息

相关信息
- 本刊中 包含"基因芯片数据"的相关文章
- 本文作者相关文章
- · 赵 倩
- · 尚学群