

## 一种新的 EST聚类方法

张利达, 袁德军, 张建伟, 王石平, 张启发

华中农业大学作物遗传改良国家重点实验室;武汉 430070

收稿日期 修回日期 网络版发布日期 接受日期

**摘要** 该研究发展了一种EST(expressed sequence tag)聚类方法(ESTClustering), 用于分析大规模EST测序中所产生的大量数据, 以获得高质量、非重复表达序列。该方法在聚类过程中采用MEGABLAST工具对一致序列进行序列同源比较, 并用phrap程序对每一EST簇进行拼接检验。这一聚类策略能降低测序错误带来的影响, 有效识别基因家族成员, 并避免选择性剪接的干扰。与NCBI(National Center for Biotechnology Information)的UniGene clustering方法相比, ESTClustering的聚类结果可以更好地反映表达序列的多样性。用ESTClustering对112 256条拟南芥EST聚类测试, 产生23 581个EST簇, 其中13 597个EST簇有对应拟南芥基因组编码序列, 与该基因组中有EST作为依据的预测基因数目接近。应用该方法对收集的147191条水稻EST序列进行聚类, 形成33 896个EST簇。

**关键词** [EST聚类](#) [一致序列](#) [无冗余cDNA文库](#)

分类号

### Abstract

### Key words

DOI:

通讯作者

### 扩展功能

#### 本文信息

- ▶ [Supporting info](#)
- ▶ [PDF\(335KB\)](#)
- ▶ [\[HTML全文\]\(0KB\)](#)
- ▶ [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“[EST聚类](#)”的 [相关文章](#)
- ▶ 本文作者相关文章

- [张利达](#)
- [袁德军](#)
- [张建伟](#)
- [王石平](#)
- [张启发](#)