

华中农业大学学士学位论文答辩

枯草芽孢杆菌群比较基因组分析

生信 2001 张子栋
指导教师：郑金水 教授

华中农业大学
信息学院

2024 年 6 月 13 日



华中农业大学
HUAZHONG AGRICULTURAL UNIVERSITY

目录

- 1 引言
- 2 材料与amp;方法
- 3 结果
- 4 讨论



1 引言

- 比较基因组分析
- 生物被膜
- 生物被膜形成相关基因

2 材料与方法

3 结果

4 讨论

引言

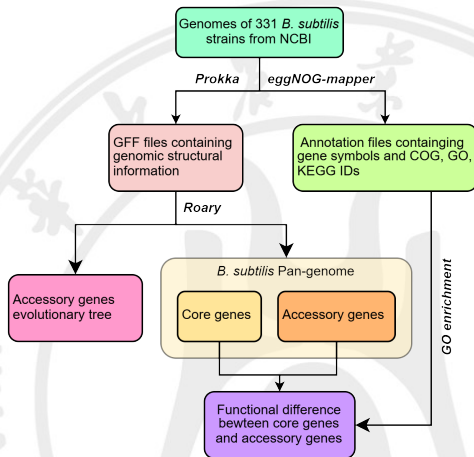
为更好利用枯草芽孢杆菌 (*Bacillus subtilis*) 这一微生物资源，本研究对从 NCBI 获取的 331 个枯草芽孢杆菌基因组进行了全面的比较基因组分析，旨在筛选出具有生物被膜 (Biofilm) 形成潜力较高的菌株。

引言

比较基因组分析

本研究中比较基因组分析以右图流程展开，包括以下几个分析步骤：

- 1 基因组结构、功能注释
- 2 泛基因组分析
- 3 基因功能富集分析
- 4 生物被膜形成相关基因的比对与识别

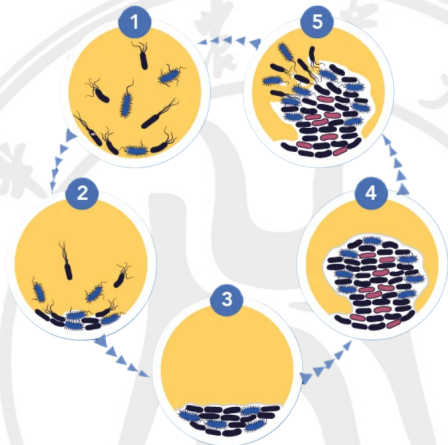


图：比较基因组分析流程

引言

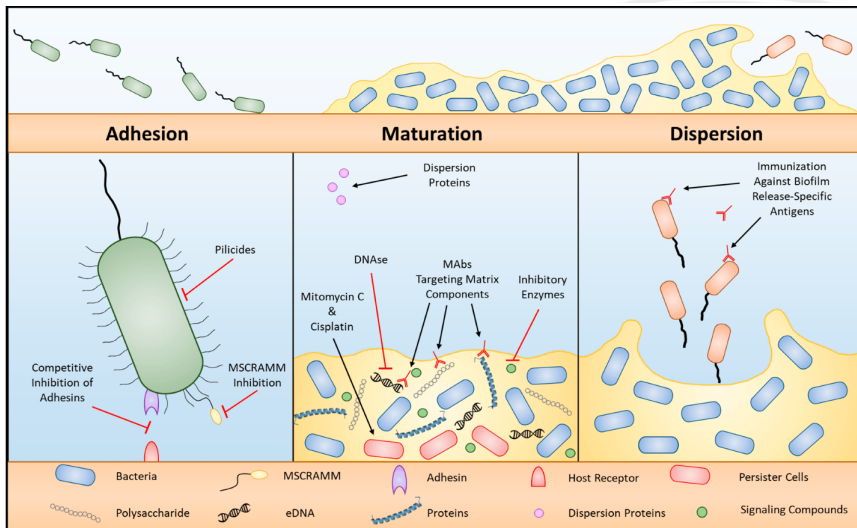
生物被膜的作用

- 保护作用
 - 抗生素耐受性
 - 免疫逃避
 - 环境抗性
- 基因交换
 - 水平基因转移
- 定殖与生存
 - 稳定附着
 - 生长优势
- 资源获取
 - 营养浓缩
 - 代谢合作



引言

生物被膜形成



引言

生物被膜形成相关基因

- ① *srfA* 操纵子：合成表面素
 - 包括 *srfAA*, *arfAB*, *srfAC*, *srfAD* 4 个基因。
- ② *abrB*：负转录调控因子
 - 抑制生物被膜基质中多糖和蛋白质相关编码基因。
- ③ *codY*：负转录调控因子
 - 营养丰富的条件下活性高，抑制与生物被膜形成相关基因的表达。
- ④ *comA*：正转录调控因子
 - 激活多种与生物被膜形成相关基因。
- ⑤ *degU*：响应转录调控因子
 - 能正向、负向调控生物被膜的形成。
- ⑥ *sinR*：负转录调控因子
 - 抑制与生物被膜形成相关基因的表达。

1 引言

2 材料与方法

- 基因组序列
- 基因组结构与功能注释
- 泛基因组分析
- 生物被膜形成相关基因的比对与识别

3 结果

4 讨论

材料与方法

基因组序列

本研究中分析的 331 个基因组下载自 NCBI，组装水平均为 complete genome。

Bacillus subtilis ★

Bacillus subtilis is a species of firmicute in the family *Bacillaceae*.

[Browse taxonomy](#)

Current scientific name	<i>Bacillus subtilis</i>
Taxonomic rank	species
NCBI Taxonomy ID	1423

材料与方法

基因组结构与功能注释

- 使用 Prokka 完成基因组结构注释
 - 对编码序列、tRNA 和 rRNA 作箱线图
- 使用 eggNOG-mapper 进行功能注释
 - 提取 Gene symbol 与注释到的 GO ID，用于下游功能富集分析

材料与方法

泛基因组分析

- 使用 Roary 构建泛基因组
 - 构建泛基因组矩阵
 - 构建物种进化树
- 使用 cluserProfiler 分别对核心基因、辅助基因完成功能富集分析

材料与方法

生物被膜形成相关基因的比对与识别

- 使用 DIAMOND 将生物被膜形成相关基因比对到基因组
 - 统计各个基因组中生物被膜形成相关基因频率
 - 作生物被膜形成相关基因的频率分布热图
 - 筛选出具有高频 *srfA* 基因和正转录调控因子、低频负转录调控因子的菌株。

1 引言

2 材料与方法

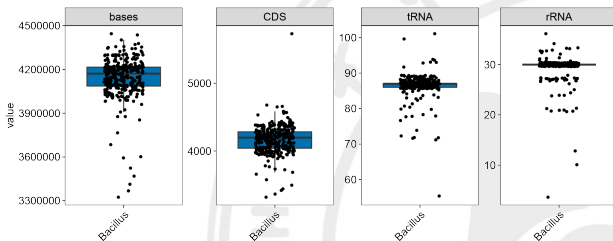
3 结果

- 基因组结构与功能注释
- 泛基因组构建
- 泛基因组矩阵
- 泛基因组组成
- 泛基因组功能富集分析
- 核心基因、软核基因、壳层基因和云层基因功能富集
- 生物被膜形成相关基因频率分布

结果

基因组结构与功能注释

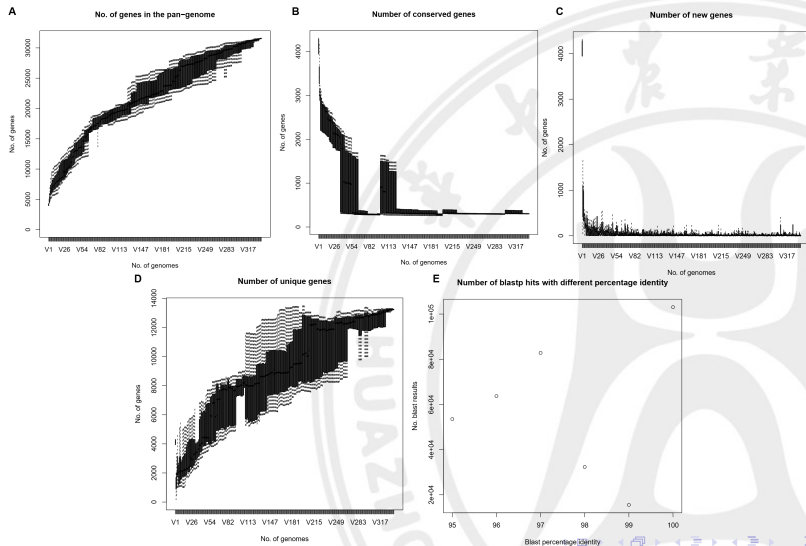
- 各基因组大小、编码序列、tRNA 和 rRNA 数量:



- 通过 eggNOG-mapper 共注释到 2 503 个 GO term, 用于下游功能富集分析。

结果

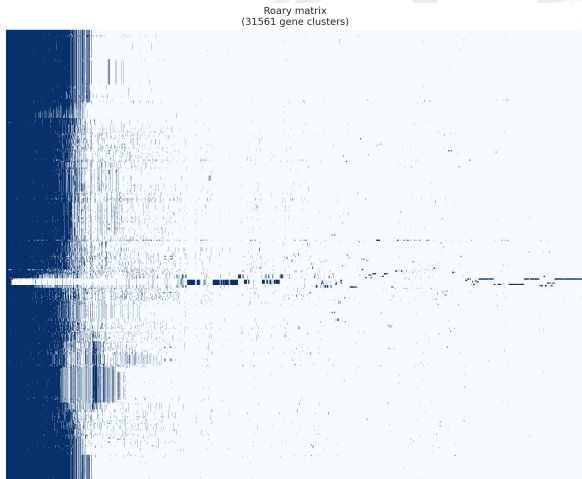
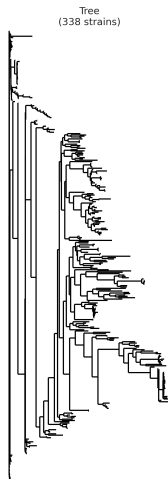
泛基因组构建



结果

泛基因组矩阵

通过 Roary 构建的枯草芽孢杆菌泛基因组共 31 561 个基因



结果

泛基因组组成

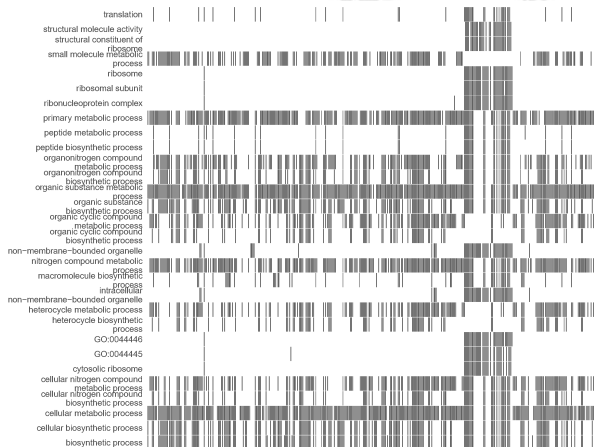
表: 泛基因组组成

Gene Category	Strain Range (% of strains)	Number of Genes
Core genes	$99\% \leq \text{strains} \leq 100\%$	303
Soft core genes	$95\% \leq \text{strains} < 99\%$	2 284
Shell genes	$15\% \leq \text{strains} < 95\%$	2 390
Cloud genes	$0\% \leq \text{strains} < 15\%$	26 584
Total genes	$0\% \leq \text{strains} \leq 100\%$	31 561

结果

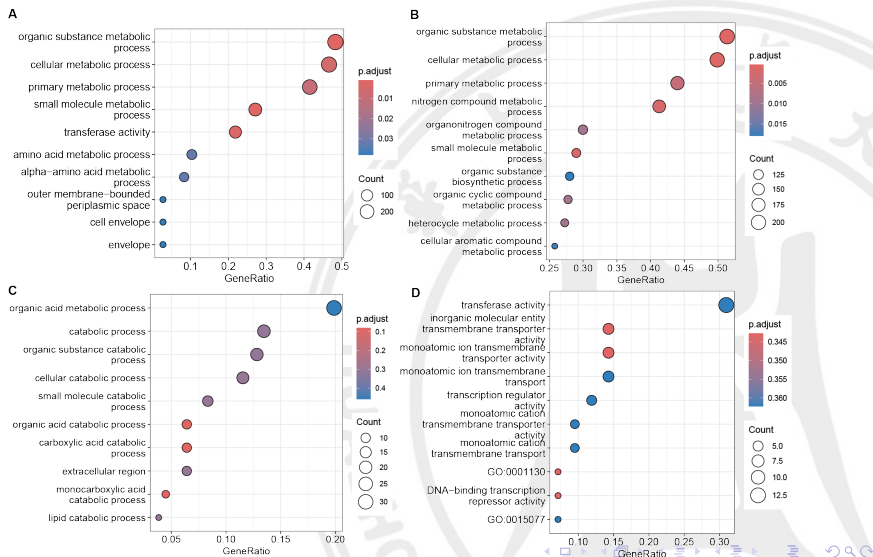
泛基因组功能富集

图: GO 富集 Top 30 功能及对应基因



结果

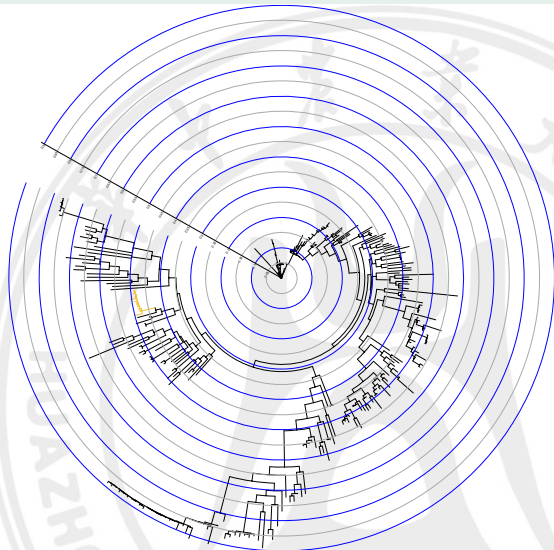
核心基因、软核基因、壳层基因和云层基因功能富集



结果

辅助基因较多的分支

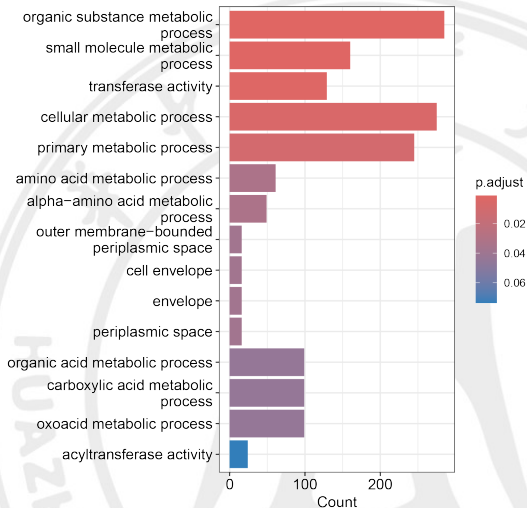
- 从泛基因组矩阵中观察到一个辅助基因较多的分支
 - 该分支存在较多与其他菌株差异大的基因
 - 进一步分析这些基因的功能



结果

辅助基因较多的分支功能富集

- 该分支的功能与核心基因功能基本重合
 - 功能冗余与替代
 - 环境适应
 - 进化压力与选择
 - 基因组重组
 - 保护机制



结果

生物被膜形成相关基因频率分布

- 筛选出 10 个生物膜合成相关基因与其他菌株有差异的菌株
 - 可能具有更高的生物膜形成潜力
 - 可以分离这些菌株做进一步的转录组分析



① 引言

② 材料与方法

③ 结果

④ 讨论



讨论

- 遗传多样性和进化机制
 - 开放的泛基因组
 - 核心基因的保守性
 - 辅助基因的多样性
- 生物被膜相关基因的分布频率
 - 相关基因在所有基因组中的拷贝数
 - 具有较高生物膜形成潜力的菌株
- 生物技术的应用前景
 - 微生物资源的保存与运输
 - 微生物肥料
 - 病原性微生物的防治
- 研究的局限性和未来展望
 - 缺乏湿实验的进一步验证
 - 未能揭示生物被膜形成相关基因的调控网络
 - 比较基因组联合多组学分析

致谢

感谢郑金水导师对本研究的悉心指导，感谢同门师兄翟昱翔对本研究开展期间提出的建议！

感谢新助研生物科技有限公司提供的计算资源，感谢程超、陈文博士对本研究的指导！

感谢同课题组的颜旭、姚代洪和张敦彪同学！在平时与他们的交流学习中，我受益匪浅。

谢谢！
请老师批评指正！