



# 格微<sup>®</sup>甄选

## 16S 基因 FISH 探针设计报告

### Lactobacillus 菌

客户： \_\_\_\_\_

单位： \_\_\_\_\_

日期： \_\_\_\_\_

## 目录

一、引言 .....	3
二、 格微 <sup>®</sup> 甄选 FISH 荧光探针的设计流程 .....	4
1. 数据库的使用 .....	4
2. 格微 <sup>®</sup> 甄选设计过程 .....	4
三、 格微 <sup>®</sup> 甄选 FISH 荧光探针的设计结果 .....	5
1. 设计的第一条探针在库中匹配情况统计 .....	5
2. 设计的第二条探针在序列库中匹配情况统计 .....	6
3. 两条探针，共同击中非目标序列库的序列条数 .....	6
4. 两条探针，共同击中目标序列库的序列条数 .....	6
四、 联系方式 .....	6

## 一、引言

荧光原位杂交 (Fluorescence In Situ Hybridization FISH) 是核酸分子杂交技术的一种, 根据碱基互补的原理, 用已知的荧光素标记的单链核酸为探针, 经变性、退火、复性, 与待检材料中未知的单链核酸进行特异性结合, 形成可被检测的杂交双链核酸。通过荧光显微镜观测荧光信号位置、大小及数量来判断待测序列的缺失、扩增及易位等情况。

微生物 FISH 检测是以微生物中较稳定的 rRNA (主要是 16S 和 23S) 作为分析的目标, 以 rRNA 序列设计探针进行杂交, 对待检测的微生物分布情况和特征性的微生物进行鉴定与定量分析, 提供微生物形态学、空间分布和生物体的细胞数量等信息。

FISH 技术的关键优势在于靶标的多样性及其探针设计的多样性。探针可以和细菌的 rRNA 特异性结合, 也就是一种探针只结合一种细菌的 rRNA。微基生物针对细菌特异性 FISH 探针的科研需求独立设计研发一套自编程序, 可满足科研人员对探针特异性的严格要求。提高实验的准确性、高效性。

### **格微<sup>®</sup>甄选系列产品：16S 基因 FISH 探针设计服务**

选择高特异性的探针序列是 FISH 实验中重要的一步, 目前 rRNA 序列已累计近千万条, 从中挑选出符合要求的“种”、“属”特异性序列, 既对目标微生物群体实现高覆盖率, 又对其他物种具有高特异性, 是件不容易的任务。目前很多文献及公司设计的特异性探针、引物的设计没有经过系统、全面的评测, 使得我们很难评测特定引物、探针对目标微生物群的覆盖率及特异性。

微基生物历时 2 年的自主研发及验证, 成功搭建了格微<sup>®</sup>甄选设计平台, 使得我们能检索出覆盖率最高、特异性最佳的探针、引物, 并提供详尽的统计报告, 让客户清晰的了解到得到所选用探针序列的覆盖率及特异性。

截止目前格微<sup>®</sup>甄选已为客户完成 16S FISH 菌种特异性探针设计 100 余种。

格微®甄选还提供了双 16S FISH 探针筛选、评测功能，当单条 FISH 探针特异性较差或覆盖率较低时，会帮助客户筛选、评测双探针检测方案（两条标记不同荧光的 FISH 探针同时杂交），并提供评测报告，由客户自主决定选择何种方案。

## 二、格微®甄选 FISH 荧光探针的设计流程

silva138.1 数据库：<https://www.arb-silva.de/documentation/release-1381/>

### 1. 数据库的使用

本次引物设计采用 silva138.1 的数据库（2020 年 8 月 27 日发布，目前为最新版本），整个数据库包含 9,469,124 条 rRNA SSU（核糖体小亚基）序列，该数据库进行优化聚类（99%相似性聚类）后一共含有 510508 条高质量代表序列及其物种信息，上述信息组成 rRNA 小亚基代表序列数据库：SSU Ref NR 99 138.1，其中古菌 16S 代表序列有 20389 条，细菌 16S 代表序列有 431329 条，真核生物 18S 代表序列有 58790 条，该数据库是目前细菌 16S 序列最为完善的数据库之一。

由于 silva138.1 核糖体大亚基序列数据量（131 万条）远小于核糖体小亚基序列数据量（947 万条），故我们在设计细菌 rRNA 特异性 FISH 探针时，优先采用核糖体小亚基数据库。

本次目标物种名称为：Lactobacillus，属于目标物种的 16S 代表序列有 1408 条（下称目标 16S 序列库），Lactobacillaceae 的物种有 6372 条，属于非目标物种的代表序列有 509100 条（下称非目标序列库）。

### 2. 格微®甄选设计过程

采用专门开发的生物信息分析流程，根据目标 16S 序列库中全部共有序列，排除掉其中的非特异性区域，列出全部可能的 FISH 探针序列，将上述候选探针序列与全部非目标序列库进行匹配，挑选特异性最高的前 150 条探针，分别统计错配碱基（mismatch）为 0、1、2、3 时，击中非目标序列库的总条数。

如经过上述筛选，有单条候选 FISH 探针序列，其特异性能达到较好效果：错配碱基分别为 0、1、2 时，与非目标序列库对比时，没有或只有少量序列击中，我们将选择最佳的一条序列作为 FISH 探针序列，设计任务完成。

如果任何单条候选 FISH 探针序列，错配碱基分别为 0、1、2 时，都会击中多条非目标序列库中序列，说明使用单条 FISH 探针序列，其特异性难达到较好的效果。我们将评测特异性最高的前 150 条探针，两两组合时，击中非目标序列库中序列的交集和并集，并选择 2 条不同的候选 FISH 探针序列组合使用时，特异性最佳的组合，报告出来，设计任务完成。

客户如需更高的特异性，可以选择使用不同颜色的荧光探针，分别标记上述的两条 FISH 探针，对样本同时进行杂交检测，双色荧光标记的微生物为目标微生物的机率可大幅提高。

### 三、格微®甄选 FISH 荧光探针的设计结果

#### 1. 设计的第一条探针在库中匹配情况统计

p1_非目标 mismatch0_count	p1_非目标 mismatch1_count	p1_非目标 mismatch2_count	p1_非目标 mismatch3_count	p1_非目标 mismatch0_percent (%)	p1_非目标 mismatch1_percent (%)	p1_非目标 mismatch2_percent (%)	p1_非目标 mismatch3_percent (%)
1444	1576	2537	4119	0.284	0.310	0.498	0.809

p1_目标 mismatch0_count	p1_目标 mismatch1_count	p1_目标 mismatch2_count	p1_目标 mismatch3_count	p1_目标 mismatch0_percent (%)	p1_目标 mismatch1_percent (%)	p1_目标 mismatch2_percent (%)	p1_目标 mismatch3_percent (%)
1225	1353	1362	1371	87.003	96.094	96.733	97.372

说明：

1. p\*\_n\_非目标 mismatch\*\_count: 探针 prob1 或 2 与非目标序列库进行匹配，分别统计 mismatch 为 0.1.2.3 时的序列条数。
2. p\*\_start: 探针 prob1 或 2 在 16S 序列上的起始位置。
3. p\*\_目标 mismatch\*\_count: 探针 prob1 或 2 与目标 16S 序列库匹配的条数，分别统计 mismatch 为 0.1.2.3 时的序列条数。
4. p\*\_目标 mismatch\*\_percent: 探针 prob1 或 2 与目标 16S 序列库匹配序列条数对应的百分比，分别统计 mismatch 为 0.1.2.3 时的序列条数的百分比。

**建议：单条 FISH 探针特异性较差，如需提高特异性，建议合成第二条探针（标**

记不同荧光），进行双 FISH 探针杂交检测。请老师决定是否选择双探针法。

## 2.设计的第二条探针在序列库中匹配情况统计

p2_非目标 mismatch0_count	p2_非目标 mismatch1_count	p2_非目标 mismatch2_count	p2_非目标 mismatch3_count	p2_非目标 mismatch0_percent (%)	p2_非目标 mismatch1_percent (%)	p2_非目标 mismatch2_percent (%)	p2_非目标 mismatch3_percent (%)
4784	31655	69241	130343	0.940	6.218	13.601	25.603

p2_目标 mismatch0_count	p2_目标 mismatch1_count	p2_目标 mismatch2_count	p2_目标 mismatch3_count	p2_目标 mismatch0_percent (%)	p2_目标 mismatch1_percent (%)	p2_目标 mismatch2_percent (%)	p2_目标 mismatch3_percent (%)
1277	1351	1364	1373	90.696	95.952	96.875	97.514

两条探针分别与非目标序列库匹配后，击中相同序列的个数。（具体的序列交集的信息见附件 Excel 表格）

## 3.两条探针，共同击中非目标序列库的序列条数

p12_非目标 mismatch0_count	p12_非目标 mismatch1_count	p12_非目标 mismatch2_count	p12_非目标 mismatch3_count	p12_非目标 mismatch0_percent (%)	p12_非目标 mismatch1_percent (%)	p12_非目标 mismatch2_percent (%)	p12_非目标 mismatch3_percent (%)
134	839	1528	2159	0.026	0.165	0.300	0.424

## 4.两条探针，共同击中目标序列库的序列条数

p12_目标 mismatch0_count	p12_目标 mismatch1_count	p12_目标 mismatch2_count	p12_目标 mismatch3_count	p12_目标 mismatch0_percent (%)	p12_目标 mismatch1_percent (%)	p12_目标 mismatch2_percent (%)	p12_目标 mismatch3_percent (%)
1376	1395	1396	1402	97.727	99.077	99.148	99.574

客户可根据以上报告给出的设计的探针特异性情况，选择合适的探针序列，以用于后继的实验研究。

附件：非目标序列交集和目标序列的并集的信息见 Excel 表格。

## 四、联系方式

技术咨询热线：021-50763698

E-mail: support@tinygene.com

公司名称：微基生物科技（上海）有限公司

公司地址：上海浦东川沙路 159 弄 88 号 1 栋 3 楼

