

## 油脂降解培养基优化的遗传算法实验研究

李绍新<sup>1</sup> 邢 达<sup>\*1</sup> 秦华明<sup>2</sup> 杨湘波<sup>1</sup> 谭石慈<sup>1</sup>

<sup>1</sup>(华南师范大学激光生命科学研究所, 广州 510631) <sup>2</sup>(华南理工大学生物工程系, 广州 510631)

**摘 要** 我们运用遗传算法,首次对利用热带假丝酵母降解油脂的培养基优化问题进行了研究。通过 5 代实验完成了 50 个实验样本、6 种培养基成分、64 个浓度水平的优化任务。按照优化后的培养基组成,热带假丝酵母降解油脂的降解率为 95.4%,比正交设计法提高了 7.63%。实验结果表明:利用遗传算法可优化培养基成分含量,取得更好的效果;遗传算法是一种优于单因素试验法和正交设计法的新型、高效的培养基优化方法。

**关键词** 油脂降解,培养基优化,遗传算法

### 1 引 言

在微生物降解油脂并应用其进行油脂废水强化处理的研究中,当筛选出高效降解油脂菌株后,微生物的发酵培养成为关键因素。而且微生物发酵培养的培养基成分复杂,用一般数理统计方法也难于优化。传统的优化培养基配方的方法是单因素试验法或正交设计法。单因素试验法是以因素间没有交互作用为前提,对于大多数培养基组成而言,这种假设往往不成立;正交设计如果组分多,则实验数量呈几何级数增长,得到的结果也只是次优解。为了优化培养基配方,采用遗传算法 (genetic algorithms,简称 GA) 是一种行之有效的方法<sup>[1]</sup>。

遗传算法首先由美国的 Holland 提出来<sup>[2]</sup>,它基于达尔文进化论和孟德尔遗传学说,仿效生物的进化与遗传,根据“生存竞争”和“优胜劣汰”的原则,借助复制、交换、突变等操作,使所要解决的问题从初始解一步步地逼近最优解。与其它传统搜索方法相比,GA 具有随机性、鲁棒性、并行性全局搜索的优越性<sup>[3]</sup>。在培养基优化方面 GA 的最大的优越性是不需要建立数学模型确定各个因素之间的相互影响,有目标函数值即可<sup>[4]</sup>。最早报道应用 GA 于培养基优化的是 Freyer 等<sup>[5]</sup>,其后 Zuzek 等也进行了尝试<sup>[6]</sup>。但目前国际上将之应用于发酵方面的研究相对较少。

基于在筛选高效降解油脂菌株研究的基础上<sup>[7]</sup>,本实验采用遗传算法进行油脂降解培养基优化的研究,结果表明,利用遗传算法以较少的实验样本优化了培养基的成分含量,提高了油脂的降解率。

### 2 基本原理

遗传算法是一种基于生物自然选择与遗传机理的随机搜索算法。与传统搜索算法不同,遗传算法从一组随机产生的种群的初始解开始搜索过程。种群中的每个个体是问题的一个解,称为“染色体”。染色体是一串符号,比如一个二进制字符串。这些染色体在后继迭代中不断进化,成为遗传。在每一代中用“适值”来测量染色体的好坏,生成的下一代染色体称为后代。后代是由前一代染色体通过交叉或者变异运算形成的。在新一代形成过程中,根据适应度的大小选择部分后代,淘汰部分后代,从而保持种群大小是常数。适值高的染色体被选中的概率较高。这样经过若干代之后,算法收敛于最好的染色体很可能就是问题的最优解或次优解。

### 3 实验部分

#### 3.1 材料

菌种为热带假丝酵母 (*Candida tropicalis*),由华南理工大学生物工程实验室分离所得。4 条件下用

2003-02-19 收稿;2003-09-08 接受  
本文系广东省自然科学基金资助项目 (No. 015012)

马铃薯-琼脂保存。斜面培养基采用马铃薯-琼脂培养基用来激活菌种。种子培养基的成分有 6 种,包括橄榄油 0.5%,酵母浸膏 0.1%,蛋白胨 1%, $\text{KH}_2\text{PO}_4$  0.2%, $(\text{NH}_4)_2\text{SO}_4$  0.2%, $\text{Na}_2\text{HPO}_4$  0.2%, $\text{MgSO}_4$  0.02%, $\text{CaCl}_2$  0.01%,在 pH 5 的酸性条件下培养。所考察的油脂降解培养基由 6 种成分组成,除橄榄油固定为 20 g/L 外,其它 6 种成分及其取值范围如表 1 所示。种子的培养方法为:选取保存在斜面上的菌 1 环,接到盛有 50 mL 种子培养基的 250 mL 三角瓶中,于 30 ℃ 下培养 24 h,摇床转速 200 r/min。发酵采用摇床培养方式,将 5 mL 种子悬浮液接入发酵培养基中,30 ℃ 下培养 24 h,摇床转速 200 r/min。

### 3.2 分析方法

细胞含量用干重法。油脂含量的测定<sup>[3,4]</sup>:采用紫外分光光度法。将发酵液离心后用石油醚萃取,经适当处理后,在 230 nm 波长下测定光密度值。

表 1 用于优化的各培养基成分取值范围及等分值

Table 1 Different culture medium components, the lower and upper limit concentrations and the step of each component

变量 Variable	下限 Lower limit	上限 Upper limit	步长 Step	变量 Variable	下限 Lower limit	上限 Upper limit	步长 Step
$(\text{NH}_4)_2\text{SO}_4$	0.2	25.6	0.4	酵母浸膏 Yeast extract	0.5	12.8	0.2
$\text{Na}_2\text{HPO}_4$	0.2	12.8	0.2	$\text{KH}_2\text{PO}_4$	0.2	12.8	0.2
$\text{MgSO}_4$	0.002	2.56	0.04	表面活性剂 Surfactant	0.00	1.28	0.02

### 3.3 GA 设计与实现

**3.3.1 编码** 采用多参数二进制编码方式。为了提高寻优概率及保证一定的精度,把每个组别划分为 64 等分,即成为 6 因素 64 水平的培养基优化问题。此时各参数的等分值如表 1 所示。因每个参数分为 64 个水平,故每个参数的子串长度  $l$  为 6 即可 ( $2^6 = 64$ )。6 个参数组合在一起形成长度  $L$  为 36 的染色体(即 1 个个体)。

**3.3.2 初始培养方案的产生** 通过算法随机产生由 10 个个体组成的群体,解码后得到初始配方及实验结果。实验的目的是提高油脂降解率(已经降解的油脂量/降解前的油脂含量),因此,以油脂降解率做为适应度来评价组分好坏的标准。

**3.3.3 选择配对与杂交** 选择的过程是个体竞争生存的过程,优秀的个体被选中生存的概率较大。在这里我们使用了锦标赛选择法<sup>[2]</sup>,随机的从种群中选择 2 个个体比较,适应度大的个体就生存下来。这样连续进行 2 次选择,产生了两个亲本。杂交是两个优秀亲本基因的遗传过程。一般杂交是在一定的概率下进行。对于要杂交的两个个体,随机地选择一个位,以这个位为分界点互换两边的基因信息,这样得到了两个新的个体。实验取杂交概率 0.7。

**3.3.4 变异** 变异是为了不使种群向单一的方向发展,保证种群的多样性。变异的概率不能太大,否则种群的发展就不能收敛了。本次实验取变异概率 0.05,也就是说有 5% 的位发生了逆变。变异的也是随机选取的。

**3.3.5 新代的产生** 通过选择、配对、杂交、变异后产生了新的种群。新的种群中必须包括上一代中的最优个体,否则算法的收敛会不稳定。由于新的种群是在旧的种群基础上择优进化的,因此,新的种群的适应度应该得到提高。

## 4 结果和讨论

上述的实验过程经过 5 代循环(共 100 瓶摇瓶实验)。对每个参数的取值平均划分为 10 个浓度水平(即 A、B、C、D、E、F、G、H、I 和 J),第 5 代个体各参数取值分布情况如表 2 所示。由表 2 可见,随着进化的进行,各个参数逐渐收敛到一个或几个区域里,每个区域里个体数目的多少与该区域上一代的适应度函数值大小有关。适应度越大,该段的基因越健康,遗传的可能性越大。

第 1 代随机产生的实验方案中,各参数浓度几乎覆盖整个取值范围;而到第 5 代(表 2),90% 的  $(\text{NH}_4)_2\text{SO}_4$  浓度集中在浓度为 C 的范围内,即收敛到同一个浓度区域,这表明该参数与其它参数没有交

互作用。此外,在第 5 代,有 90 % 的酵母膏浓度集中在 E、F 的范围内;100 %MgSO<sub>4</sub> 浓度分别分布在和 G、H 的范围内(即收敛到相邻的浓度区域),这可认为它们与其它参数没有明显的交互作用。而 KH<sub>2</sub>PO<sub>4</sub>、Na<sub>2</sub>HPO<sub>4</sub> 在最后一代则收敛到两个有间隔的浓度区域。40 % 的 KH<sub>2</sub>PO<sub>4</sub> 浓度值在 C 范围内,50 % 在 G 范围内;50 % 的 Na<sub>2</sub>HPO<sub>4</sub> 浓度值分布在 B 范围内,40 % 在 F 范围内。说明这些参数与其它参数有交互作用。

表 2 (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>、Na<sub>2</sub>HPO<sub>4</sub>、KH<sub>2</sub>PO<sub>4</sub>、酵母浸膏、表面活性剂和 MgSO<sub>4</sub> 优化后浓度区域分布 (%)  
Table 2 Distributions for end concentration of (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>、yeast extract、surfactant、KH<sub>2</sub>PO<sub>4</sub> and MgSO<sub>4</sub> (%)

	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	酵母浸膏 Yeast extract	KH <sub>2</sub> PO <sub>4</sub>	Na <sub>2</sub> HPO <sub>4</sub>	MgSO <sub>4</sub>	表面活性剂 Surfactant
A						
B				50		
C	90		40			
D						
E		40				
F		60		40		
G			50		60	
H					40	
I						90
J						

优化过程中各代及不同培养基配方所得到的油脂降解率的最高值和平均值如图 1 所示。由图可以看出,第二代后,最高值的变化不大,但是还有稍微的上升。说明该算法很快地搜索到了最优配方的邻域附近。平均值也在上升,接近最优值,这是由于当算法搜索到最优值邻域时,群体的适应度得到了提高,群体的配方浓度收敛到了如表 2 所示的一个较小的范围内。图 1 中第二代的平均值和最高值比第三代还高,这是由于实验中种子活力的差别引起的。在每一代实验中,个体菌种的活力是相同的;在不同代的实验中,由于菌种培养条件或者操作上的细微的差异而导致不同代培养出的菌种的活力有所差别,因此降解油脂的能力也有所不同。但是并不影响最优配方的搜索。

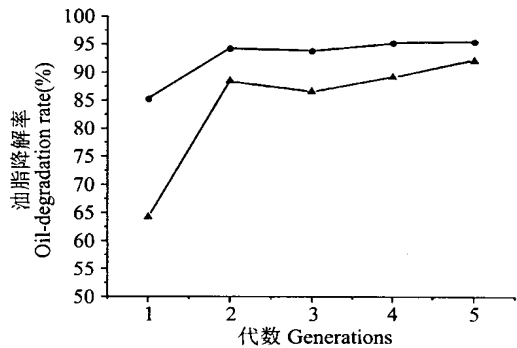


图 1 不同代不同培养基组成油脂降解率的最高值和平均值(和 分别表示最大的油脂降解率和平均降解率)

优化前后油脂降解培养基及油脂降解率的变化情况如表 3 所示。由表看出,利用正交法实验得到的油脂降解率比遗传算法得到的油脂降解率要低 6.95 个百分点。然而正交法做的实验次数远比遗传算法做的实验次数多,而且所费的时间长。

Fig. 1 Maximum and average oil-degradation rate under different culture medium of each generation. ( and represent maximum and average oil-degradation rate respectively)

表 3 优化前后培养基浓度及油脂降解率变化

Table 3 Change of medium and product concentrations before and after optimization

	正交法 (g/L) Orthogonal array	遗传算法 (g/L) Genetic algorithms	变化率 Change (%)
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	5	4.64	- 0.72
酵母浸膏 Yeast extract	4	3.55	- 13.75
Na <sub>2</sub> HPO <sub>4</sub>	2	2.24	12
KH <sub>2</sub> PO <sub>4</sub>	4	6.35	58.75
MgSO <sub>4</sub>	0.5	0.135	- 73
表面活性剂 Surfactant	0.9	0.75	- 16.67
油脂降解率 Oil-degradation (%)	89.5	95.4	6.59

尽管将遗传算法应用到培养基优化方面取得了很大的成功,但是该算法还是有一定的局限性。要想

算法收敛到最优配方,实验的次数还要增加;该算法采用竞争选择,当种群中出现个别适应度相当高的个体时,这种个体在种群中会迅速繁殖,算法可能会迅速收敛到局部最优解。因此,算法本身还要发展。

实验结果表明,运用 GA 能够以较少的实验次数 ( $2 \times 50$  次,全部实验次数为  $64^6$  次)在较大搜索区域内完成培养基优化任务。它不仅提高了油脂降解率,而且精确了培养基成分含量,有利于提高经济效益。遗传算法是一种优于单因素试验法和正交法的快速、高效、经济的培养基优化方法。

## References

- 1 He Xianfeng(何险峰), Zhou Jiaju(周家驹). *Progress in Chemistry*(化学进展), **1998**, 10(3): 312 ~ 318
- 2 Wang Xiaoping(王小平), Cao Liming(曹立明). *Genetic Algorithms Theories, Applications and Software Practices* (遗传算法——理论、应用与软件实现). Xi an (西安): Xi an Traffic University Press (西安交通大学出版社), **2002**: 4 ~ 14
- 3 Zhou Ming(周 明), Sun Shudong(孙树栋). *Genetic Algorithms Principles and Applications*(遗传算法原理及应用). Beijing (北京): National Defence Industry Press (国防工业出版社), **1999**: 11 ~ 15
- 4 Zoe S D, Richard J G, Roger J M, Douglas B K. *Applied and Environmental Microbiology*, **2000**, 66(4): 1435 ~ 1443
- 5 Freyer S, Weuster B D, Wandrey C. *Bio Engineering*, **1992**, 8(5-6): 16 ~ 25
- 6 Zuzek M, Friedrich J, Cestnik B, Karali A, Cimerman A. *Biotechnology Techniques*, **1996**, 10 (12): 991 ~ 996
- 7 Qin Huaming(秦华明), Zhang Ying(张 英), Liang Shizhong(梁世中). *The Third Fermentation Engineering Symposium*(第三次全国发酵工程学术讨论会论文集). Beijing(北京): Chinese Light Industry Press(中国轻工业出版社), **2002**: 86 ~ 90

## Experimental Study on Genetic Algorithms of Medium Optimization for Oil-degradation

Li Shaoxin<sup>1</sup>, Xing Da<sup>\*1</sup>, Qin Huaming<sup>2</sup>, Yang Xiangbo<sup>1</sup>, Tan Shici<sup>1</sup>

<sup>1</sup>( Institute of Laser Life Science, South China Normal University, Guangzhou 510631)

<sup>2</sup>( Department Bioengineering, South China University of Technology, Guangzhou 510631)

**Abstract** Medium optimization of the culture for oil-degradation by *Candida tropicalis* using genetic algorithms (GA) was studied. The 64 concentration levels of 6 medium components were optimized within 50 experiments. The results show that GA can be applied in the medium optimization and better results are obtained. Employing optimized medium composition, a final oil-degradation rate of 95.4% was obtained and it was improved by 6.94% compared to the results with the medium before optimization. GA is an efficient new medium optimum method which is superior to the other optimum means such as orthogonal array method.

**Key words** Oil-degradation, medium optimization, genetic algorithms

(Received 19 February 2003; accepted 8 September 2003)