

用神经网络和遗传算法优化怀槐悬浮细胞合成异黄酮

罗建平* 罗 凯 陈晓燕 姜绍通

(合肥工业大学生物与食品工程学院,合肥 230009)

摘 要 为了获得怀槐悬浮细胞合成异黄酮的最适培养条件,采用 ANNs(人工神经网络)结合 RAGA(实数编码加速遗传算法)对培养基组成进行全局寻优。培养基中影响异黄酮染料木素产率的主要组成是 KNO_3 、 $(\text{NH}_4)_2\text{SO}_4$ 、2,4-D 和 6-BA。在它们的有效作用浓度范围内,用随机 10 组培养基组合及细胞染料木素产率为输入和输出由 ANNs 对数据建模,由 RAGA 优化模型参数。建立的优化模型准确性高,依赖模型由 RAGA 全局寻优获得的最佳培养基组合是 149.68 mg/L $(\text{NH}_4)_2\text{SO}_4$ 、2936.10 mg/L KNO_3 、0.01 mg/L 2,4-D 和 0.19 mg/L 6-BA,染料木素产率达 14.13 mg/L,与模型预测值的误差为 7.38%。结果表明,运用神经网络结合遗传算法优化怀槐细胞合成异黄酮的培养条件是可行的,优化后的培养基使怀槐细胞异黄酮合成能力比优化前有很大的提高。

关键词 怀槐,人工神经网络,实数编码加速遗传算法,染料木素,悬浮培养

中图分类号 Q813 **文献标识码** A **文章编号** 1000-3061(2004)05-0759-05

怀槐(*Maackia amurensis* Rupr. et Maxim)是我国珍贵的野生药用植物^[1],具有祛风湿、消炎、镇痛、健胃、止血等作用^[2]。药理学研究证明,怀槐心材多酚类提取物关键活性成分是异黄酮化合物,可通过阻断肝组织中胶原蛋白和粘多糖的合成有效抑制肝组织纤维化,达到保肝和治疗肝炎的功效^[3-4],并具有肺泡表面活性剂损伤修复的活性^[5]。由于怀槐生长缓慢,尚未人工栽培,野生资源十分有限^[1]。为了保护 and 可持续利用怀槐野生药用资源,我们进行了怀槐细胞工程的研究,已经证明怀槐培养细胞具有异黄酮合成能力,由培养细胞生产的异黄酮具有显著的减轻小鼠 CCl_4 肝炎的作用^[6]。

利用植物细胞工程技术开发野生药用植物资源,关键之一是对培养条件进行优化以最大程度地提高细胞生产名贵药用次生物质的水平。优化培养因素的方法很多^[7],近年来基于遗传算法和神经网络的微生物发酵过程优化和参数寻优都取得较好效果^[8],但在植物细胞培养体系优化方面还没有报道。由于染料木素具有重要的医疗保健功效^[3-5],本文以染料木素作为怀槐细胞生产异黄酮的指标,在单因素及其有效浓度筛选的基础上,采用实数编码加速遗传算法(real coding based accelerating genetic algo-

rithm, RAGA)优化的人工神经网络(artificial neural networks, ANNs)即 ANNs-RAGA 对怀槐培养基优化进行建模,再由 RAGA 进行全局寻优,期望获得怀槐培养细胞生产异黄酮的最优条件,为生物反应器发酵生产怀槐异黄酮奠定基础。

1 材料与方法

1.1 细胞悬浮培养

悬浮培养用怀槐愈伤组织按前文方法进行继代培养^[9]。取生长 21d 的怀槐愈伤组织(平均鲜重 5g),接种于盛有 60 mL 液体培养基的 250 mL 三角瓶中暗处、 $25 \pm 2^\circ\text{C}$ 下悬浮培养。培养基为附加 2.0 mg/L 2,4-D 和 0.5 mg/L 6-BA 的 B5 培养基,水平振荡摇床转速为 110 r/min。每个处理 3~5 个重复。进行培养基单因素试验的因子分别是 $(\text{NH}_4)_2\text{SO}_4$ 、 KNO_3 、2,4-D、NAA、6-BA、KT 和 GA_3 。

1.2 细胞生长、异黄酮提取、含量测定及产率计算

怀槐细胞悬浮培养 12 d,收获过滤、 60°C 下空气干燥至恒重,细胞生物量为每升培养基收获的干重细胞,以 g(dry wt)/L 表示。干燥的细胞充分研磨粉碎后于 50 mL 乙醇(95%)中冷浸 24 h,经超声波处理 30 min 后过滤。3 次提取滤液合并, 50°C 减压蒸

收稿日期:2004-02-16,修回日期:2004-05-10。

基金项目:国家自然科学基金资助(No.30170059)。

* 通讯作者。 Tel:86-551-2900090; Fax:86-551-2900096; E-mail: jianpingluo@sohu.com

干后,用 100 mL 乙酸乙酯-水(5:1, V:V)混合溶液萃取 3 次,乙酸乙酯部分减压蒸干,再以甲醇溶解定容至 5 mL,经 0.45 μm 孔径滤膜过滤后,于 Waters 高效液相色谱仪(包括 515 型泵、UV-2487 检测器)上测定异黄酮染料木素含量。HPLC 分析时,色谱柱为美国 Merck 公司的 Purospher® STARC₁₈ 柱(4.6 mm i.d. \times 250 mm, 5 μm),流动相为甲醇-水-乙酸(体积比 10:10:1)溶液,柱温 25 $^{\circ}\text{C}$,进样体积 20 μL ,流速 1.0 mL/min,检测波长 260 nm,灵敏度为 3.0 AUFS。染料木素标准品购自 Sigma 公司,其标准曲线是 $Y = 6.37 \times 10^4 X - 1.45 \times 10^4$ ($r = 0.9992$)。测定数据用 WSC 色谱工作站(中科院大连化物所)处理,按外标法计算培养细胞中染料木素含量,染料木素产率为每升培养基收获的细胞中染料木素的积累量,以 mg/L 表示。

1.3 ANNs-RAGA 建模与 RAGA 优化(图 1)

1.3.1 RAGA 操作过程^[10]:

第 1 步:进行实数编码。用线性变换 $x(j, i) = a(j, i) + y(j, i)[b(j) - a(j)]$, $j = 1, 2, \dots, p$, $i = 1, 2, \dots, n$,把初始变化区间 $[a(j), b(j)]$ 的第 j 个优化变量 $x(j, i)$ 对应到 $[0, 1]$ 区间上的 $y(j, i)$,生成 GA 中的基因,优化问题所有变量的对应的基因依次连在一起构成问题解的编码形式,生成染色体。

第 2 步:进行父代种群初始化。选取初始种群为 n ,在 $[0, 1]$ 区间上随机产生 n 组数,每组 p 个,记为 $u(j, i)$,作为初始群体的父代个体。

第 3 步:对父代群体适应度进行评价。运用 RAGA 进行优化,其适应度为 ANNs-RAGA 的输出值(染料木素产率),其值越大,表示个体适应度越高。因此将适应度函数定义为 $F(i) = f(i)/\max$, \max 为样本中结果的最大值。

第 4 步:进行选择操作、杂交操作和变异操作,产生 $3n$ 个子代个体,从中选择适应度最大的 n 个作为新的父代个体,转入编码操作,进入新一轮的演化。

第 5 步:进行加速循环。用第一、第二次演化迭代所产生的优秀个体形成的子群所对应的变量变化区间作为变量新的初始变化区间, RAGA 转入第 1 步。如此加速循环,优秀个体变化区间逐步收缩,直至循环终止。

1.3.2 模型构建及全局寻优:根据单因素试验确定 $(\text{NH}_4)_2\text{SO}_4$ 、 KNO_3 、2,4-D 和 BA 的有效范围浓度作为网络输入,以细胞生产染料木素的产率为目标输

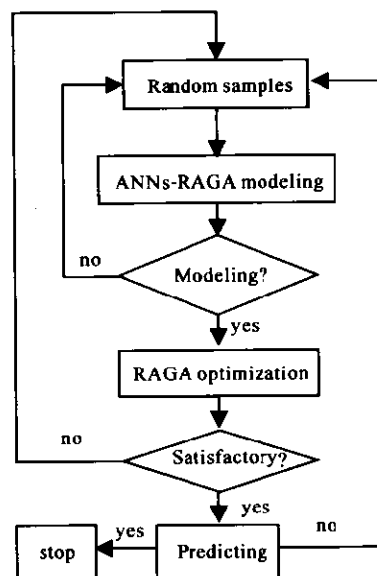


图 1 ANNs 与 RAGA 流程图

Fig.1 The flow chart of ANNs and RAGA

出,用 4 种因子有效浓度范围内随机 10 组试验数据作为学习样本,采用 BP(back-propagation algorithm,误差反向传播算法)进行 ANNs 训练,由 RAGA 调整运行参数,即 ANNs-RAGA 法构建该系统数学模型,寻找输入、输出数据之间的复杂关系,并另取 5 组实验数据对模型预测的准确性进行检验。根据模型寻找出的输入与输出数据间的关系,由 RAGA 随机生成样本来寻找输入样本之间的最佳组合关系,使得输出值达到最优。

1.3.3 实验验证:用获得的最优培养基进行怀槐细胞悬浮培养,同时进行 4 因素 3 水平正交试验(表 1),对模型优化结果进行验证。

表 1 正交试验因素和水平

Table 1 Factors and levels of orthogonal experiments

Levels	Factors			
	A	B	C	D
	KNO_3	$(\text{NH}_4)_2\text{SO}_4$	2,4-D	6-BA
	/(mg/L)	/(mg/L)	/(mg/L)	/(mg/L)
1	1500	200	0	0
2	2250	500	0.5	0.1
3	3000	800	1.0	0.3

2 结果与讨论

2.1 影响细胞合成异黄酮的主要培养基组成及其有效浓度的确定

前文怀槐细胞培养过程中发现培养基中氮源和添加的激素对细胞合成异黄酮影响很大^[11]。为选

择适合的输入参数用于神经网络建模优化,本文首先对不同的氮源、生长素和细胞分裂素影响细胞生产异黄酮染料木素产率进行单因素实验。当改变培养基中氨态氮(NH_4^+) $_2\text{SO}_4$ 或硝态氮 KNO_3 浓度时,细胞生长未表现出显著的差异($P < 0.05$),但细胞中异黄酮合成有较大的变化。 NH_4^+ 浓度在 100 ~ 500 mg/L 范围内染料木素产率维持在较高水平,以 500 mg/L NH_4^+ 时的产率最高,减少或增加 NH 浓度,染料木素产率都迅速下降(图 2A)。 NO_3^- 在实验浓度范围内出现两个适合染料木素生产的浓度,分别是 1500 mg/L 和 3000 mg/L,其它浓度下染料木素产率比最高值(1500 mg/L KNO_3 时)低 23% ~ 60% (图 2B)。

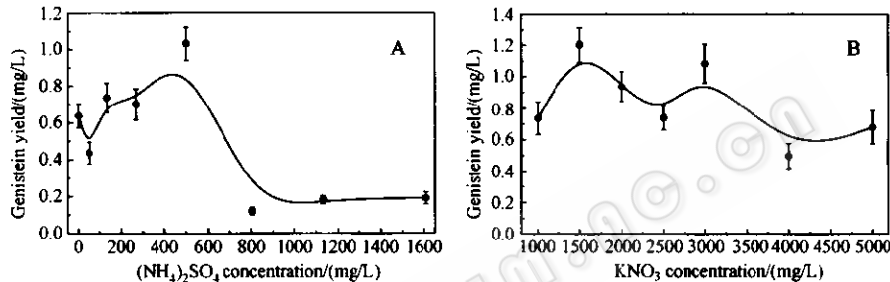


图2 氨态氮(A)和硝态氮(B)对怀槐细胞生产染料木素的影响

Fig.2 Effects of NH_4^+ (A) and NO_3^- (B) on genistein yield of *Maackia amurensis* suspension cells

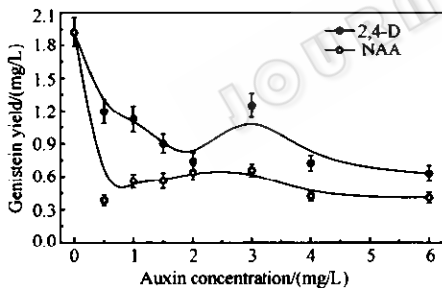


图3 生长激素对怀槐细胞生产染料木素的影响。

Fig.3 Effects of auxins on genistein yield of *Maackia amurensis* suspension cells

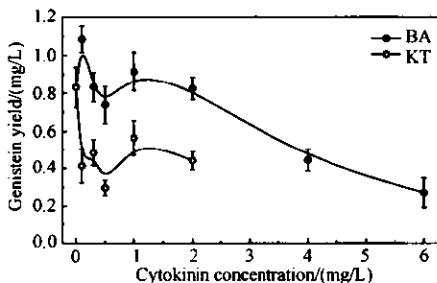


图4 细胞分裂素对怀槐细胞生产染料木素的影响

Fig.4 Effects of cytokinins on genistein yield of *Maackia amurensis* suspension cells

生长素。尽管培养基中添加 2,4-D 和 NAA 能不同程度促进细胞生长,但显著抑制细胞生产染料木素, NAA 的抑制作用大于 2,4-D,染料木素产率以不含生长素最好(图 3)。调节培养基中细胞分裂素浓度可以明显改变细胞生产染料木素能力,BA 在低浓度时有利于染料木素生产,高浓度时抑制染料木素生产,以 0.1 mg/L 时产率最高,但 KT 对细胞合成染料木素表现出显著的抑制作用(图 4)。通过单因素实验分析,确定出 $(\text{NH}_4)_2\text{SO}_4$ 、 KNO_3 、2,4-D 和 BA 是培养基中影响细胞生产染料木素的主要组成成分,它们的适合作用浓度范围分别在 100 ~ 800 mg/L、1500 ~ 3000 mg/L、0 ~ 3 mg/L 和 0 ~ 1 mg/L 之间。

2.2 ANNs-RAGA 的模型建立

人工神经网络是把简单的处理单元广泛连接成网络,用以模拟大脑神经系统的结构和功能^[10]。根据单因素实验确定的 4 种因素,所建立的 ANNs 拓扑结构为 4-4-1。以 4 种因素有效作用浓度范围内随机生成 10 种组合与实验结果进行 ANNs-RAGA 建模。结果表明样本训练能较快地达到收敛(图 5),模型模拟结果与试验值的吻合度较好(图 6),两者之间的相对误差都在 10% 以内。由另外随机 5 组实验样本数据进行模型运算,对模型预测进行检验,运行结果的误差分析表明,随机样本实验值与模型预测值之间的相对误差均在 6% 以内(表 2),说明模型预测准确性较高。因此,只要知道培养基中这 4 种成分的组成便可由模型估测出细胞生产异黄酮的产率。

2.3 培养基优化预测及实验验证

交替使用 RAGA 和 ANNs-RAGA 模型进行培养基组合的全局寻优,得到最优培养基组合(表 3)。用该优化的培养基和优化前的培养基同批进行细胞悬浮培养,结果表明优化的培养基可使细胞生产染料木素产率比优化前有很大的提高。分析表明优化

前后的培养基中细胞合成染料木素的实际产率和模型预测结果的误差都小于 10%,说明建立的模型是稳定的,所寻优得出的培养基组合是可信的。建模同时,在单因素有效作用浓度范围内进行 4 因素 3

水平的正交试验。结果表明由模型优化的培养基中实际细胞染料木素产率比正交试验优化结果提高近 1.5 倍,证明了神经网络和遗传算法用于培养基优化明显优于正交试验。

表 2 实验测定值与模型预测值比较

Table 2 Comparison between experimental values and predictive values

Medium combination	Factor concentration/(mg/L)				Genistein yield/(mg/L)		Relative error /%
	(NH ₄) ₂ SO ₄	KNO ₃	2,4-D	BA	Experimental value	Predictive value	
1	134	2500	0	0.5	5.78	6.92	-5.38
2	134	3000	2	0.5	3.02	2.91	3.60
3	268	2500	2	0.5	2.93	2.85	2.69
4	134	2500	2	0.1	4.20	4.23	-1.91
5	134	2500	2	0.5	0.89	0.85	4.59

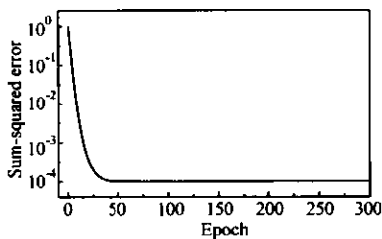


图 5 ANNs-RAGA 模型过程训练误差下降曲线

Fig.5 Sum-squared error decrease curve of ANNs-RAGA training

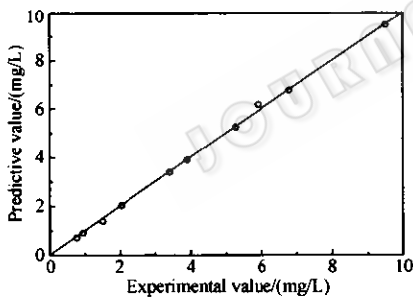


图 6 ANNs-RAGA 建模实验值与模型计算值比较

Fig.6 Comparison between experimental values for modeling and calculating values by ANNs-RAGA

表 3 培养基优化前后细胞生产染料木素的比较

Table 3 Comparison of cells producing genistein in optimized medium with the original medium

Item		Before optimization	After optimization
Optimized factors in medium	(NH ₄) ₂ SO ₄ /(mg/L)	134.00	149.68
	KNO ₃ /(mg/L)	2500.00	2936.10
	2,4-D/(mg/L)	2.00	0.01
	BA/(mg/L)	0.50	0.19
Yield of genistein/(mg/L)	Predictive value	0.74	15.17
	Experimental value	0.73 ± 0.08	14.13 ± 0.33
	Relative error/%	0.41	7.38

3 结论

由于在培养基优化过程中要考虑的因素众多,而且它们之间的关系又十分复杂,各组分之间相互影响,使得优化的过程变得比较困难。仅考虑各种因素单独对优化结果的影响都需要做大量的实验。如果在单因素的基础上做正交实验,虽然有较好的代表性,但考虑到工作量大的因素,使得选择因素浓度的范围很有限,不能在大范围内进行全局寻优,而运用神经网络和遗传算法正好可以弥补以上不足。本文结果也验证了由神经网络和遗传算法进行全局寻优能获得比正交试验结果更加优化的组合。神经网络和遗传算法在工程领域国内外已有很多的运用,随着神经网络和遗传算法本身的逐渐成熟,其运用范围也在不断扩展^[10],已在微生物发酵过程优化中显示出很好应用前景^[12,13]。由于培养基优化的最终目的是寻找到最优结果,因此本文采用神经网络和遗传算法相结合来对怀槐细胞悬浮培养基优化进行建模和寻优,该建模方法和优化方法具有很强的可操作性与简单性,非常适合用于高度非线性的、并含有大量噪声的生物系统^[13]。本文运用此方法取得了较理想的结果,为下一步的怀槐细胞生物反应器大规模发酵培养奠定了基础,并说明了运用此方法进行怀槐细胞生物反应器大规模发酵生产怀槐异黄酮的模型建立和优化应该是可行的。

致 谢 本文研究工作得到了合肥工业大学金菊良教授的帮助,在此表示感谢。

REFERENCES(参考文献)

- [1] Liu J(刘娟), Wang LX(王良信). A new resource of the medicinal

- plant—*Maackia amurensis*. *Chinese Wild Plant Resource* (中国野生植物资源), 1997, 16: 22–25
- [2] Qu XZ(曲香芝), Li JD(李景道), Li GZ(李光洙). Studies on chemical constituents of *Maackia amurensis* Rupr. et Maxim. *Chinese Traditional and Herbal Drugs* (中草药), 1997, 28: 72–73
- [3] Qi QH(齐茄红), Kang LP(康鲁平), Zhang JP(张俊平). Antifibrous effects of genistein and quercetin *in vitro*. *Acta Pharmaceutical Sinica* (药学报), 2001, 36: 648–651
- [4] Barbara L, Vera R, Raffaella V *et al*. Phytoestrogens and liver disease. *Molecular and Cellular Endocrinology*, 2002, 193: 81–84
- [5] Palagina MV, Bezdetko CN, Moisejenko L. Biological activity of *Maackia amurensis* Rupr. et Maxim wood in correction of irradiation damage of lung surfactant in rats. *Russk Resursy*, 2000, 36: 78–82
- [6] Luo JP(罗建平), Shen GD(沈国栋), Jiang ST(姜绍通). Analysis and hepatoprotective function of isoflavones from callus cultures of *Maackia amurensis*. *Chinese Food Science* (食品科学), 2003, 24: 139–142
- [7] Fang BS(方柏山), Chen HW(陈宏文), Xie XL(谢晓兰) *et al*. The medium optimization of xylitol fermentation based on neural networks and genetic algorithms. *Chinese Journal of Biotechnology* (生物工程学报), 2000, 16: 648–650
- [8] Sunitha K, Lee JK, Oh TK. Optimization of medium components for phythase production by *E. coli* using response surface methodology. *Bioprocess Engineering*, 1999, 21: 477–481
- [9] Shen GD(沈国栋), Luo JP(罗建平), Chen XY(陈晓燕). Influences of different signal molecules on suspension cell growth and isoflavone production of *Maackia amurensis*. *Journal of Hefei University of Technology* (合肥工业大学学报) 2003, 26: 1295–1298
- [10] Jin JL(金菊良), Ding J(丁晶). *Water resource systems engineering* (水资源系统工程). Chengdu: Sichuan Science and Technology Press (四川科学技术出版社), 2002
- [11] Luo JP(罗建平), Wu LF(吴礼福), Jiang ST(姜绍通). Callus formation and its isoflavonoid accumulation in *Maackia amurensis*. *Journal of Chinese Materia Medica* (中国中药杂志), 2003, 28: 1138–1141
- [12] Lednicky P, Meszaros A. Neural network modeling in optimization of continuous fermentation process. *Bioprocess Engineering*, 1998, 18: 427–432
- [13] Cai YJ(蔡宇杰), Zhuge B(诸葛斌), Zhang XH(张锡红) *et al*. Optimization of medium of raw starch enzyme fermentation with neural networks and genetic algorithms. *Journal of Wuxi University of Light Industry* (无锡轻工大学学报), 2001, 20: 421–423

Optimization for Isoflavone Production in *Maackia amurensis* Suspension Cells Based on Neural Networks and Accelerating Genetic Algorithm

LUO Jian-Ping* LUO Kai CHEN Xiao-Yan JIANG Shao-Tong

(School of Biotechnology and Food Engineering, Hefei University of Technology, Hefei 230009, China)

Abstract The medium for isoflavone production in *Maackia amurensis* suspension cells has been optimized through the artificial neural networks (ANNs) and the real coding based accelerating genetic algorithm (RAGA). Among the ingredients of the medium, nitrogen sources and plant growth regulators were found to be the main factors affecting the production of isoflavone genistein. $(\text{NH}_4)_2\text{SO}_4$, KNO_3 , 2,4-D and 6-BA, 100–800 mg/L, 1500–3000 mg/L, 0–3 mg/L and 0–1 mg/L respectively, significantly increased genistein yield, in the ranges of effective concentrations. The random ten combinations of these four components generated by RAGA as input data and the genistein yields of ten combinations as output data were used for ANNs-RAGA (the artificial neural networks associated with the accelerating genetic algorithm) modeling. The resultant model showed a high fit between the experimental data and calculating values by ANNs-RAGA. Based on the prediction of the model, the optimum combination of four factors for genistein production was determined on 149.68 mg/L for $(\text{NH}_4)_2\text{SO}_4$, 2936.10 mg/L KNO_3 , 0.01 mg/L 2,4-D and 0.19 mg/L 6-BA. When cells were cultured in the optimized medium, their capability of genistein production was remarkably enhanced to 14.13 mg/L, which was about 19 times higher than that in the original medium. The relative discrepancy between the experimental value and the predictive value of genistein yield from the optimized medium was 7.38%.

Key words *Maackia amurensis* Rupr. et Maxim, artificial neural networks, accelerating genetic algorithm, genistein, cultured cells

Received: 02-16-2004

This work was supported by Grant from the National Natural Science Foundation of China (No. 30170059).

* Corresponding author. Tel: 86-551-2900090; Fax: 86-551-2900096; E-mail: jianpingluo@sohu.com