

马铃薯密码子用法分析及其 在 t-PA 基因密码子改造上的应用

柏 锡¹, 徐建震², 李 琳³, 郭 政², 李 杰¹, 朱延明¹

(1. 东北农业大学生命科学院, 哈尔滨 150030; 2. 哈尔滨医科大学基础医学院, 哈尔滨 150001;

3. 中国农业大学信息与电气工程学院, 北京 100094)

摘 要:采用 bioperl-1.0 工具在红旗 LINUX 系统下自编了密码子分析软件;通过对马铃薯 98 个蛋白质编码基因序列(codon DNA sequence)的分析,计算出了马铃薯的密码子用法,并确定出了马铃薯的 4 个高表达优越密码子;依据马铃薯密码子用法和高表达优越密码子分析结果,对 t-PA 基因序列进行了密码子的改造,得到了具有马铃薯密码子使用特点的 t-PA 基因序列,从而为以马铃薯为生物反应器高效生产 t-PA 奠定了分子基础。

关键词:马铃薯; t-PA; 密码子用法; 优越密码子; 密码子改造

中图分类号: Q75

文献标识码: A

文章编号: 0253-9772(2004)01-0075-09

Analysis of Codon Usage in Potato and Its Application in the Modification of t-PA Gene

BAI Xi¹, XU Jian-Zhen², LI Lin³, GUO Zheng², LI Jie¹, ZHU Yan-Ming¹

(1. College of Life Sciences, Northeast Agricultural University, Harbin 150030, China;

2. College of Basic Medicine Science, Harbin Medical University, Harbin 150001, China;

3. College of Information and Electrical Engineering, China Agricultural University, Beijing 100094, China)

Abstract: Bioperl-1.0 was used under Hongqi LINUX system to programm the codon analysis software. According to the analysis of 98 codon DNA sequences with this software, the codon usage in potato was calculated and 4 codons have been inferred to the optimal codons. The codons of tissue plasminogen activator (t-PA) gene sequence have been reconstructed according to the results. The t-PA gene sequence containing the optimal codons of potato will be used for t-PA production by potato bioreactor.

Key words: potato; t-PA; codon usage; optimal codons; codon modify

马铃薯作为植物表达系统被广泛地用于外源蛋白质的表达。然而同其他植物表达系统一样,在转基因马铃薯中存在着目的产物表达量低的问题,这极大的限制了其在外源蛋白质生产中的应用。为了提高外源蛋白质的表达量,目前往往采用的方法是提高外源基因转录水平,例如使用高表达启动子、增强子等表达元件^[1]。然而在生物体中,基因特别是高表达基因在密码子的使用上具有偏爱性,不同生物体倾向于使用部分特定的同义密码子,这些同义密码子被称为该种生物的偏爱密码子。由于外源基

因与异源植物表达系统在密码子偏爱上的不同,往往造成外源基因难以有效翻译。因此,通过对一些常用植物表达系统的密码子使用特点进行分析,找出植物所偏爱的密码子,可以为改造外源基因提供依据,实现外源基因在该植物系统中的高水平表达^[2]。

组织型纤溶酶原激活剂(tissue plasminogen activator, t-PA)是重要的溶栓剂之一,其与血栓基质纤维蛋白有较强的特异亲和性,能在血栓局部高效激活纤溶系统,被广泛地应用于血栓栓塞性疾病

收稿日期: 2002-12-09; 修回日期: 2003-04-04

作者简介: 柏 锡(1975-),男(汉族),黑龙江省富锦市人,东北农业大学硕士研究生,专业方向:植物基因工程

通讯作者: 朱延明(1955-),男(汉族),山东省莘县人,博士学位,教授,博士生导师,专业方向:植物基因工程。Tel:0451-5190734,

E-mail: ymzhu2001@hotmail.com

的治疗^[3]。t-PA 是利用重组 DNA 技术开发的首批生物技术药物之一,也是目前市场上销售额最高的基因工程产品之一。目前 t-PA 主要通过微生物发酵和哺乳动物培养细胞途径进行生产,但微生物发酵存在易形成胞含体和产物不能正确折叠等缺点,哺乳动物细胞培养存在培养条件复杂、周期长、成本高等缺点。

本研究通过对马铃薯密码子用法进行分析,找出马铃薯的偏爱密码子和高表达优越密码子。根据密码子分析结果对 t-PA 基因密码子进行改造,得到适于马铃薯中高量表达的 t-PA 基因序列,从而为以马铃薯为生物反应器生产 t-PA 奠定基础。

1 材料和方法

1.1 马铃薯基因的获取

在 GenBank DNA 序列信息库中检索得到 125 个马铃薯基因。

1.2 目的基因的筛选

筛选原则:(1)编码蛋白质的基因,并且编码序列完整;(2)编码基因的序列长度上要求大于 300 个碱基对;(3)在细胞质中翻译,排除细胞器中翻译的基因;(4)排除位于质粒、转座子、病毒上的基因;(5)对多拷贝基因只统计一次。经筛选得到 98 条基因序列作为马铃薯密码子分析的样本(见表 1)。

1.3 密码子有效数 (effective number of codons, N_c) 的计算

N_c 值反映了基因的编码序列中使用密码子种类的多少。 N_c 值是个无偏的估计量,对于短序列的基因(>100 个氨基酸)和偏爱的氨基酸密码子(skewed aa condons)都有很好的反映^[4]。因此该指标已作为讨论密码子用法的重要参考指标之一^[5]。计算方法参见文献。

高表达基因的密码子偏爱程度大,从而 N_c 值较小;低表达基因的密码子偏爱程度小,密码子使用较平均,从而 N_c 值较大。因此可以通过 N_c 值确定内源基因表达量的相对高低。

1.4 相对同义密码子用法 (relative synonymous codon usage, RSCU) 的计算

RSCU 反映了样本或样本组中(多个编码基因)各个同义密码子的相对使用数目,可用于密码子偏爱程度的评价^[6]。计算方法参见文献。

1.5 高表达优越密码子的确定

将检索到的基因按照计算的 N_c 值从高到低排

列,从两端选择约 5% 分别作为低表达组和高表达组样本组。计算出两个样本组中基因序列各个密码子的 RSCU 值,以 t -检验比较($P < 0.05$),找出 RSCU 值差异显著的密码子,从而筛选出高表达基因的优越密码子。

1.6 密码子分析软件编程

采用 bioperl-1.0^[7] 工具在红旗 LINUX 系统下进行软件编程,经测试得到了密码子分析程序 CODON ANALYSE,该程序能够正确地进行单或多基因密码子出现次数统计、 N_c 值计算、RSCU 值计算。

2 结果

2.1 N_c 值的计算和高低表达组样本组的确定

通过 CODON ANALYSE 计算筛选基因的 N_c 值,结果见表 1。按照 N_c 值由小到大的顺序将 98 个基因进行排序,从两端选取约 5% (各取 5 个) 分别作为高表达组和低表达组样本组。高表达样本组所包含的基因是 AB041521 (逆境和成熟时表达)、POTRBCS (与能量代谢相关)、STHIGHMOB (与 DNA 损伤修复相关)、POTMUCSTTN (与抗虫害相关)、POTUBI3A (在胚胎或迅速增殖的细胞中大量表达)。这些基因要求在转录后能够迅速大量合成,在翻译中要避免使用稀有密码子,因此这些基因的 N_c 值低。低表达样本组所包含的基因是 STU02497 (与芳香族化合物的降解和烯烃的利用相关)、STU318070 (铁硫氧还蛋白还原酶催化亚基)、POTU2B (mRNA 拼接)、POTCAM (钙调蛋白)、STADPRF1 (与细胞内囊泡运输调节相关)。这些基因起调节和催化作用,在植物体内表达量比较低,因此它们的 N_c 值较高。

2.2 样本总体 RSCU 值的计算

通过 CODON ANALYSE 统计样本总体中各个密码子的出现次数,计算样本总体的 RSCU 值(见表 2)。由表 2 的结果可看出 14 种氨基酸其密码子第 3 位是 T 时 RSCU 值较高;1 种氨基酸(Gly)其密码子第 3 位是 T 或 A 时 RSCU 值较高;3 种氨基酸(Gln、Glu、Arg)其密码子第 3 位是 A 时 RSCU 值较高;1 种氨基酸(Lys)其密码子第 3 位是 G 时 RSCU 值较高,而 Gln、Glu、Arg、Lys 均不含有第 3 位是 T 的简并密码子;2 种氨基酸(Met、Trp)只对应一个密码子。由此可以看出,在马铃薯中偏爱使用第 3 位为 T 的密码子。终止密码子偏爱使用 TAA。

表 1 马铃薯密码子分析基因样本及其密码子有效数

Table 1 The gene samples for analysis of codon usage in potato and their effective number of codons

基因 Gene	描述 Description	登录号 Accession number	密码子数 Codon quantities	N_c	参考文献 References
AB041521	ppACS1 mRNA for ACC synthase	AB041521	487	37.89	Direct Submission
POTRBCS	ribulose biphosphate carboxylase (rbcS) mRNA	J03613	182	39.24	<i>Proc Natl Acad Sci, USA</i> , 85 (3): 846-850
STHIGHMOB	mRNA for high mobility group protein	AJ002391	142	39.29	Direct Submission
POTMUCSTTN	multicystatin gene	L16450	757	39.41	<i>Plant Mol Biol</i> , 23 (4): 801-812
POTUBI3A	ubiquitin/ribosomal fusion protein (ubi3) gene	L22576	157	41.00	<i>Plant Mol Biol</i> , 24 (1): 119-127
STU307584	mRNA for tropinone reductase I (TRI gene)	AJ307584	265	42.03	Direct Submission
POTADH3	alcohol dehydrogenase 3 (ADH-3) mRNA	M25152	381	44.41	Direct Submission
STTKETMR	mRNA for transketolase	Z50099	742	44.55	Direct Submission
AF082893	methionine synthase (MS) mRNA	AF082893	766	44.82	<i>Plant Mol Biol</i> , 48 (3): 255-265
POTPATA	patatin class I gene	M18880	387	45.45	<i>Gene</i> , 62 (1): 27-44
POTSTHA	pathogenesis-related protein (STH-2) gene	M29041	156	46.31	<i>Plant Mol Biol</i> , 22 (2): 279-291
POTINHWI	wound-inducible proteinase inhibitor I gene	M17108	108	46.77	<i>Plant Mol Biol</i> , 8: 199-207
AF384821	1-aminocyclopropane-1-carboxylate oxidase (ACO2) mRNA	AF384821	320	46.79	Direct Submission
POTPR1A	glutathione S-transferase (gst1) gene	J03679	218	47.26	Direct Submission
STU277244	tuberosum mRNA for alpha-glucosidase (mal2 gene)	AJ277244	929	47.48	<i>Planta</i> , 213 (2): 258-264
POTPSTH21	pSTH-21 protein mRNA	M25156	156	47.66	<i>Mol Plant Microbe Interact</i> , 2 (6): 325-331
STFLAVSYN	mRNA for flavonol synthase	X92178	350	47.71	<i>Plant J</i> , 11 (1): 105-113
POTPATATB	patatin mRNA clone pPATB1	M21879	387	47.74	<i>Plant Mol Biol</i> , 11: 255-269
POTIHK	inhibitor II-chloramphenicol acetyltransferase gene (IHK)	M15186	154	47.85	<i>Proc Natl Acad Sci, USA</i> , 84: 744-748
STACCAS3	(ST ACS2) gene for 1-Aminocyclopropane-1-carboxylate synthase	Z27235	442	48.11	<i>Mol Gen Genet</i> , 246 (4): 496-508
STU318822	pt3 gene for transmembrane protein	AJ318822	536	48.13	<i>Nature</i> , 414 (6862): 462-470
STPCPI	mRNA for DNA/RNA binding protein	X82328	510	48.20	<i>Mol Gen Genet</i> , 247 (6): 759-763
STTSGDC	mRNA for T subunit of glycine decarboxylase multi-enzyme complex	Z25862	407	48.35	<i>Plant Physiol</i> , 104 (3): 1079-1080
S74514	S-adenosylmethionine decarboxylase	S74514	361	48.36	<i>Plant Mol Biol</i> , 26 (1): 327-338
STU345003	mRNA for spermidine synthase (SpdSy gene)	AJ345003	348	48.42	Direct Submission
ST39EFHAN	mRNA for 39 kDa EF-Hand protein	AJ001310	356	48.52	Direct Submission
AF067863	1,3-beta-glucan glucanohydrolase gene	AF067863	364	48.54	Direct Submission
POTPPIE	gene for proteinase inhibitor	D17332	214	48.66	<i>Plant Cell Physiol</i> , 35 (2): 303-312
STBETFRC	invertase gene encoding beta-fructofuranosidase	Z22645	583	48.76	<i>Gene</i> , 145 (2): 211-214
AF193846	branched-chain amino acid aminotransferase (BCAT2) mRNA	AF193846	378	48.77	<i>Plant Physiol Biochem</i> , 39: 855-860
STBETTUBI	(TUBST1) mRNA for beta-tubulin	Z33382	452	48.78	<i>Plant Mol Biol</i> , 26 (3): 1013-1018
POTAROIA	3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (aro1) mRNA	J05191	510	48.87	<i>J Biol Chem</i> , 265 (3): 1608-1614
STU72489	remorin mRNA	U72489	199	49.14	<i>Plant Cell</i> , 8 (12): 2265-2276
STU65648	homeodomain protein POTH1 mRNA	U65648	346	49.24	Direct Submission
AF233342	DNA-binding protein p24 mRNA	AF233342	275	49.40	<i>Plant Cell</i> , 12 (8): 1477-1490
STU007739	mRNA for w-3 desaturase	AJ007739	432	49.52	Direct Submission
STU18548	mRNA for lipoxygenase	Y18548	862	49.55	<i>Biochem J</i> , 353 (Pt 2): 345-355
STU310520	mRNA for fatty acid hydroperoxide lyase (hpl gene).	AJ310520	481	49.68	<i>Proc Natl Acad Sci, USA</i> , 98 (14): 8139-8144

续表 1

基因 Gene	描述 Description	登录号 Accession number	密码子数 Codon quantities	N_c	参考文献 References
POTPFPA	pyrophosphate-fructose 6-phosphate 1-phospho-transferase alpha-subunit	M55190	618	49.89	<i>J Biol Chem</i> , 265 (30): 18366-18371
STRPS4	mRNA for ribosomal protein S4	X76651	265	50.04	<i>Biochim Biophys Acta</i> , 1218 (3): 435-438
STU27082	catalase (CAT1) mRNA	U27082	493	50.08	<i>Plant Physiol</i> , 108 (4): 1748
STU23757	transcription factor (POTM1-1) mRNA	U23757	251	50.12	<i>Gene</i> , 166 (2): 329-330
STU69633	cold-stress inducible protein (C17) gene	U69633	209	50.40	<i>Plant Physiol</i> , 104 (2): 445-452
SBGLURICH	mRNA for glutamic acid-rich protein	X95984	241	50.46	<i>Plant Mol Biol</i> , 33 (2): 291-300
AF008652	MADS transcriptional factor (Stmads11) mRNA	AF008652	222	50.47	<i>Planta</i> , 207 (2): 181-188
STU47738	chalcone synthase 2 mRNA	U47738	390	50.61	Direct Submission
AY063128	G protein alpha subunit 3 mRNA	AY063128	393	50.61	Direct Submission
AF483209	trehalose synthase (TS1) gene	AF483209	858	50.81	Direct Submission
STU308597	kco1b gene for K ⁺ channel protein	AJ308597	350	51.01	Direct Submission
AY098940	CDPK-like protein mRNA	AY098940	512	51.16	Direct Submission
AY039111	gibberellin 3-beta-hydroxylase 1 mRNA	AY039111	374	51.30	Direct Submission
AY027522	starch associated protein R1 mRNA	AY027522	1465	51.33	Direct Submission
STHXK	mRNA for hexokinase	X94302	499	51.45	Direct Submission
SOT223252	mRNA for UDP-glucose:protein transglucosylase (uptg1 gene)	AJ223252	366	51.53	Direct Submission
POTST4C11	4-coumarate-CoA ligase (St4C1-1) gene	M62755	546	51.56	<i>J Biol Chem</i> , 266 (13): 8551-8559
STPPC1	ppc1 gene	X90982	966	51.66	<i>Plant Mol Biol</i> , 23 (4): 881-888
A23741	GBSS gene	A23741	607	51.90	Direct Submission
STU24087	clone gPOSS16 sucrose synthase gene	U24087	806	51.92	<i>Plant Cell</i> , 7 (9): 1369-1385
STU52079	P-glycoprotein (pmdr1) mRNA	U52079	1314	51.96	<i>Plant Mol Biol</i> , 31 (3): 683-687
POTAGPTHI	alpha-glucan phosphorylase type H isozyme mRNA	M69038	839	51.97	<i>J Biol Chem</i> , 266 (28): 18446-18453
STPYRK	mRNA for pyruvate kinase (cytosolic isozyme)	X53688	511	52.06	<i>Plant Mol Biol</i> , 15 (4): 665-669
STSBE	mRNA for starch-branching-enzyme	X69805	862	52.23	<i>Plant Physiol</i> , 102 (3): 1053-1054
AF067859	lactate dehydrogenase-2 (LDH-2) mRNA	AF067859	347	52.25	Direct Submission
STU76701	NADH nitrate reductase (StNR2) mRNA	U76701	912	52.27	<i>J Exp Bot</i> , 51 (347): 1017-1026
POTHMGR1	hydroxymethylglutaryl coenzyme A reductase (hmgr) mRNA	L01400	597	52.37	<i>Plant Cell</i> , 4 (10): 1333-1344
POTSACPD	stearoyl-acyl carrier protein desaturase mRNA	M91238	394	52.40	Direct Submission
AF044172	cysteine synthase mRNA	AF044172	326	52.47	Direct Submission
STFUM1	mRNA for fumarase	X91615	494	52.48	Direct Submission
STU95923	transaldolase (PotTal1) mRNA	U95923	439	52.65	<i>Plant Mol Biol</i> , 32 (3): 447-452
STCHITIN	endochitinase gene	X15494	329	52.67	<i>Nucleic Acids Res</i> , 17 (14): 5855-5856
POTST4C11	4-coumarate-CoA ligase (St4C1-1) gene	M62755	546	52.68	<i>J Biol Chem</i> , 266 (13): 8551-8559
AB062143	StNR6 mRNA for nitrate reductase	AB062143	912	52.74	Direct Submission
STSTARCH	mRNA for starch phosphorylase	X52385	967	52.91	<i>Plant Cell</i> , 1 (5): 559-566
STCDINR	Cdi mRNA for cathepsin D inhibitor	X67843	220	53.09	<i>Plant Cell</i> , 4 (9): 1157-1170
AB041031	HMG2 gene for 3-hydroxy-3-methylglutaryl coenzyme A reductase	AB041031	596	53.17	Direct Submission
STU310910	mRNA for UDP-glucose:protein transglucosylase (uptg2 gene)	AJ310910	367	53.22	Direct Submission
STTATABP	mRNA for TATA-binding protein	X62494	201	53.46	<i>Plant Mol Biol</i> , 19 (3): 455-464
POTCATHD	cathepsin D inhibitor protein mRNA	M96257	222	53.46	<i>Plant Physiol</i> , 101 (2): 703-704
STU314612	mRNA for storekeeper protein (STK gene)	AJ314612	400	53.61	Direct Submission

续表 1

基因 Gene	描述 Description	登录号 Accession number	密码子数 Codon quantities	N_c	参考文献 References
AF225512	soluble NSF attachment protein mRNA	AF225512	289	53.64	<i>Plant Mol Biol</i> , 43 (4):473-482
STPT2	mRNA for inorganic phosphate transporter	X98891	528	53.65	<i>Plant Cell</i> , 9 (3):381-392
STU309541	mRNA for divinyl ether synthase (des gene)	AJ309541	479	53.74	<i>FEBS Lett</i> , 507 (3):371-376
POTBFRUASE	beta-fructosidase mRNA	L29099	640	54.11	<i>Plant Physiol</i> , 106 (1):397-398
POTAMY23A	alpha-amylase (Amy23) mRNA	M79328	408	54.21	Direct Submission
STU47740	chalcone synthase 1b mRNA	U47740	390	54.35	<i>Plant Physiol</i> , 111:348
AY050221	pathogenesis-related protein 1b precursor (pr1b) mRNA	AY050221	160	54.60	Direct Submission
STU242853	mRNA for Dof zinc finger protein (dof1 gene)	AJ242853	325	54.62	Direct Submission
AB062138	StMPK1 mRNA for mitogen-activated protein kinase	AB062138	397	54.96	Direct Submission
AY089960	aspartic proteinase inhibitor precursor P8G5 mRNA	AY089960	186	55.06	Direct Submission
STTPTMR	mRNA TPT for triose phosphate translocator	X67045	415	55.23	<i>Mol Gen Genet</i> , 238 (3):357-361
STSUCTR	mRNA for sucrose transporter	X69165	317	55.91	<i>Plant Cell</i> , 5 (11):1591-1598
STSTS14	sts14 gene	X82652	215	56.98	<i>Plant Mol Biol</i> , 30 (1):171-176
STCYF1	cy-F1 mRNA for cytosolic fructose-1,6-biphosphate	X76946	341	57.53	<i>Plant J</i> , 9 (5):671-681
STU02497	Lemhi Russet clone EH10.1 epoxide hydrolase mRNA	U02497	322	57.60	<i>Plant J</i> , 6 (2):251-258
STU318070	mRNA for ferredoxin-thioredoxin- reductase catalytic subunit B	AJ318070	149	58.69	Direct Submission
POTU2B	spliceosomal protein (U2B) mRNA	M72892	232	59.75	<i>Gene</i> , 107 (2):197-204
POTCAM	clone PCM1 calmodulin mRNA	J04559	130	61.47	<i>Proc Natl Acad Sci, USA</i> , 86 (10):3644-3648
STADPRF1	mRNA for ADP-ribosylation factor 1	X74461	198	61.56	<i>Plant Cell Rep</i> , 14:180-183

表 2 马铃薯相对同义密码子用法

Table 2 The relative synonymous codon usage of potato

氨基酸 Amino acid	密码子 Codon	N	RSCU	氨基酸 Amino acid	密码子 Codon	N	RSCU	氨基酸 Amino acid	密码子 Codon	N	RSCU	氨基酸 Amino acid	密码子 Codon	N	RSCU
Phe	TTT	1101	1.13	Ser	TCT	859	1.65	Tyr	TAT	802	1.12	Cys	TGT	401	1.20
	TTC	853	0.87		TCC	391	0.75		TAC	627	0.88		TGC	266	0.80
Leu	TTA	528	0.80		TCA	789	1.52	Stop	TAA	43	1.33	Stop	TGA	38	1.18
	TTG	1024	1.55		TCG	185	0.36		TAG	16	0.49		Trp	TGG	588
Leu	CTT	1125	1.71	Pro	CCT	891	1.63	His	CAT	631	1.27	Arg	CGT	367	1.19
	CTC	515	0.78		CCC	287	0.52		CAC	365	0.73		CGC	132	0.43
	CTA	386	0.58		CCA	800	1.46	Gln	CAA	904	1.22	CGA	209	0.68	
	CTG	377	0.57		CCG	211	0.38		CAG	574	0.78	CGG	97	0.32	
Ile	ATT	1318	1.57	Thr	ACT	878	1.55	Asn	AAT	1169	1.21	Ser	AGT	548	1.05
	ATC	665	0.79		ACC	384	0.68		AAC	762	0.79		AGC	348	0.67
	ATA	528	0.63		ACA	823	1.45	Lys	AAA	1412	0.95	Arg	AGA	583	1.90
Met	ATG	983	1.00	ACG	185	0.32	AAG		1560	1.05	AGG		457	1.49	
Val	GTT	1324	1.75	Ala	GCT	1472	1.82	Asp	GAT	1682	1.43	Gly	GGT	1065	1.40
	GTC	482	0.64		GCC	530	0.66		GAC	667	0.57		GGC	465	0.61
GTA	GTG	747	0.99	GCA	993	1.23	Glu	GAA	1582	1.10	GGA	1067	1.40		
	GTG	747	0.99		GCG	239		0.30	GAG	1301		0.90	GGG	453	0.59

N:密码子出现次数。

N:Frequency of codon application.

2.3 高低表达样本组 RSCU 值的计算和高表达优越密码子的确定

通过 CODON ANALYSE 统计高表达与低表达样本组中各个密码子的出现次数,计算出高低表达样

本组中各个密码子的 RSCU 值(见表 3)。对两组数据进行 *t* 检验,取 $P < 0.05$,筛选出高表达优越密码子。*t* 检验在 MATLAB6.0 上实现。共筛选出 4 个密码子,分别是 TTG、GTT、TAC、GGA(见表 4)。

表 3 马铃薯中高低表达样本组的相对同义密码子用法

Table 3 The relative synonymous codon usage of high/low expression sample group in potato

氨基酸 Amino acid	密码子 Codon	高表达组 High expression group		低表达组 Low expression group		氨基酸 Amino acid	密码子 Codon	高表达组 High expression group		低表达组 Low expression group		氨基酸 Amino acid	密码子 Codon	高表达组 High expression group		低表达组 Low expression group							
		N	RSCU	N	RSCU			N	RSCU	N	RSCU			N	RSCU	N	RSCU	N	RSCU				
Phe	TTT	49	1.02	30	1.18	Ser	TCT	11	0.86	14	1.68	Tyr	TAT	31	0.94	22	1.29	Cys	TGT	10	1.11	6	0.86
	TTC	47	0.98	21	0.82		TCC	8	0.62	3	0.36		TAC	35	1.06	12	0.71		TGC	8	0.89	8	1.14
Leu	TTA	12	0.60	8	0.55	TCA	26	2.03	14	1.68	Stop	TAA	3	1.80	2	1.20	Stop	TGA	2	1.20	3	1.80	
	TTG	43	2.15	15	1.03		TCG	1	0.08	3		0.36	TAG	0	0	0		0	Trp	TGG	26	1.00	14
Leu	CTT	46	2.30	31	2.14	Pro	CCT	20	1.25	12	0.96	His	CAT	21	1.50	11	0.96	Arg	CGT	20	2.31	9	1.15
	CTC	9	0.45	14	0.96		CCC	8	0.50	8	0.64		CAC	7	0.50	12	1.04		CGC	0	0	6	0.77
	CTA	7	0.35	9	0.63		CCA	36	2.25	23	1.84		CAA	44	1.29	20	0.95		CGA	0	0	5	0.64
	CTG	3	0.15	10	0.69		CCG	0	0	7	0.56		CAG	24	0.71	22	1.05		CGG	1	0.12	1	0.13
Ile	ATT	44	1.42	26	1.39	Thr	ACT	21	1.18	24	1.71	Asn	AAT	59	1.14	28	1.30	Ser	AGT	23	1.79	10	1.20
	ATC	20	0.64	19	1.02		ACC	17	0.96	13	0.93		AAC	44	0.85	15	0.70		AGC	8	0.62	6	0.72
	ATA	29	0.94	11	0.59		ACA	31	1.75	18	1.28		AAA	77	0.81	35	0.90		Arg	AGA	17	1.96	10
Met	ATG	33	1.00	33	1.00	ACG	2	0.11	1	0.07	AAG	112	1.18	43	1.10	AGG	14	1.62	16	2.04			
Val	GTT	87	2.43	27	1.64	Ala	GCT	72	2.22	41	1.95	Asp	GAT	85	1.53	38	1.33	Gly	GGT	29	1.13	22	1.35
	GTC	13	0.36	11	0.67		GCC	17	0.52	13	0.62		GAC	26	0.47	19	0.67		GGC	12	0.47	19	1.17
	GTA	20	0.56	9	0.54		GCA	29	0.89	24	1.14		GAA	66	1.06	44	0.97		GGA	47	1.82	16	0.98
	GTG	23	0.64	19	1.15		GCG	12	0.40	6	0.29		GAG	58	0.94	47	1.03		GGG	15	0.58	8	0.49

N:密码子出现次数。

N:Frequency of codon application.

表 4 马铃薯中的高表达优越密码子列表

Table 4 The optimal codon list of high expressional gene in potato

氨基酸 Amino acid	密码子 Codon	优越密码子 Optimal codon <i>P</i> (0.05)	氨基酸 Amino acid	密码子 Codon	优越密码子 Optimal codon <i>P</i> (0.05)	氨基酸 Amino acid	密码子 Codon	优越密码子 Optimal codon <i>P</i> (0.05)	氨基酸 Amino acid	密码子 Codon	优越密码子 Optimal codon <i>P</i> (0.05)
Phe	TTT	—	Ser	TCT	—	Tyr	TAT	—	Cys	TGT	—
	TTC	—		TCC	—		TAC	TAC(0.0462)		TGC	—
Leu	TTA	—	TCA	—	Stop	TAA	—	Stop	TGA	—	
	TTG	TTG(0.0495)		TCG		—	TAG		—	Trp	TGG
Leu	CTT	—		Pro		CCT	—		His	CAT	—
	CTC	—	CCC		—	CAC	—	CGC		—	
	CTA	—	CCA		—	CAA	—	CGA		—	
	CTG	—	CCG		—	CAG	—	CGG		—	
Ile	ATT	—	Thr	ACT	—	Asn	AAT	—	Ser	AGT	—
	ATC	—		ACC	—		AAC	—		AGC	—
	ATA	—		ACA	—		AAA	—		Arg	AGA
Met	ATG	—	ACG	—	AAG	—	AGG	—			
Val	GTT	GTT(0.0224)	Ala	GCT	—	Asp	GAT	—	Gly	GGT	—
	GTC	—		GCC	—		GAC	—		GGC	—
	GTA	—		GCA	—		GAA	—		GGA	GGA(0.0333)
	GTG	—		GCG	—		GAG	—		GGG	—

“—”:非优越密码子。

“—”:The unoptimal codons.

2.4 t-PA 基因的密码子改造

根据马铃薯密码子偏爱性及高表达优越密码子分析结果,对组织型纤溶酶原激活剂(tissue plas-

minogen activator, t-PA) 基因^[8] (GenBank 登录号 M15518) 序列进行改造,使其在密码子使用上具有马铃薯基因的特点,结果见图 1。

氨基酸	M	D	A	M	K	R	G	D	E	E	V	L	L	L	C	G	A	V	P	V	S	P	S	U	R	I	H	A	R	F
原序列	ATG	GAT	GCT	GCA	ATG	AAG	AGA	GGG	GTG	TGC	TGT	GTG	GTG	GTG	GTG	TGT	GGA	GTA	GTC	GTG	TGG	GGC	AGC	DAG	GAA	ATC	GAT	GCC	GGA	TTC
改造序列	ATG	GAT	GCT	ATG	AAG	AGA	GGA	TTG	TGT	TGT	GTT	TTG	TTG	TTG	TGT	GGA	GCT	GST	TTT	GTT	TCT	CCA	DCT	CAA	GAA	ATC	GAT	GCC	GGA	TTC
氨基酸	K	R	G	A	I	L	S	V	Q	V	I	G	H	H	K	Y	Q	M	I	V	Q	Q	H	S	W	L	R	P	V	
原序列	AGA	AGA	GGA	DCC	AGA	TCT	TAC	CAA	GTG	ATC	TGC	AGA	GAT	GAA	AAA	ACG	DAG	ATG	ATA	TAC	CAG	CAA	CAT	CAG	TCA	TGG	CTG	GGT	GGT	
改造序列	AGA	AGA	GGA	GCT	AGA	TCT	TAC	CAA	GTT	ATT	TGT	AGA	GAT	GAA	AAA	ACT	CAA	ATG	ATT	TAC	CAA	CAA	DAT	CAA	TCA	TGG	TTG	TTG	AGG	GTT
氨基酸	L	R	S	N	R	V	E	Y	C	W	C	N	S	G	R	A	Q	C	H	S	V	P	V	K	S	L	P	F		
原序列	CTC	AGA	AGC	AAC	CGG	GCG	GAA	TAT	TGC	TGG	TGC	AAG	AGT	GAC	AGG	GCA	CAG	TGC	CAG	TGA	GTC	GGT	GGT	AAA	ACT	TGG	ATC	AGC		
改造序列	TTG	AGA	TTT	AAT	AGA	GTT	GAA	TAT	TGC	TGG	TGT	AAI	TGT	AGA	AGA	GCT	CAA	TGT	CAT	TCA	ATT	GCT	GGT	AAA	TCT	GCT	TGT	GAA		
氨基酸	C	F	N	G	G	T	C	Q	Q	Q	A	L	V	F	S	D	F	V	E	Q	C	F	E	G	Y	A	U	K		
原序列	TCT	TTT	AAI	GGG	DAC	ADL	TGC	CAI	GAG	GCC	ATG	GCC	ATG	TGC	TGC	TCA	GAT	TTG	GTG	TGC	TAT	TTK	TTK	TTK	GAA	TTA	TTT			
改造序列	TCT	TTT	AAT	GGA	DGA	ACT	TGT	CAA	CAA	GCT	TGG	TAC	TTT	TCA	GAT	TTT	GTT	TGT	CAA	TGT	CCA	GAA	GSA	TTT	GCT	GGA	ANG			
氨基酸	I	D	I	R	A	T	G	Y	H	D	U	G	I	S	Y	B	G	I	W	S	Y	A	E	S	D	A	E			
原序列	ATA	TTT	ATC	ATC	CCG	CCG	TGC	TAT	GAT	GAC	CAI	GGC	ATC	AGC	TAC	AGG	GGC	AGG	TGG	ATC	ACA	DCC	DAG	AGC	GGC	GCC	GAG			
改造序列	ATT	GAT	ACT	AGA	GCT	ACT	TGT	TAC	GAA	GAT	CAA	GGA	ATT	TCT	TAC	AGA	GGC	ACT	TGG	TCT	ACA	GCT	GAA	TCT	GGA	GGT	GAT			
氨基酸	W	N	S	A	L	A	L	A	Q	K	P	Y	S	D	R	P	D	A	I	L	L	G	L	G	N	H	S			
原序列	TGG	AAC	AGC	AGC	GGG	TTG	GCC	CAG	AAG	CCG	TAC	AGC	GGG	GGG	AGG	CCA	BAC	GCC	ATC	AGG	CTG	GGC	CTG	GGG	AGC	CAC				
改造序列	TGG	AAT	TCT	TCT	GGT	TTG	GCT	CAA	AAG	CCA	TAC	TCT	GAA	AGA	AGA	CCA	GAT	GCT	ATP	AGA	TTG	GGT	GGT	GAA	AAI	TCA				
氨基酸	N	D	D	R	D	S	K	P	C	Y	V	F	K	A	G	K	A	G	S	B	E	F	C	S	T	P				
原序列	AAC	CCA	GAT	GCA	CAC	TCA	AAG	CCG	TGG	TGC	TAC	GTC	TTT	AAG	GGG	GGG	AGG	TAC	AGC	TCA	GAG	TTT	TCC	AGC	AGC					
改造序列	AA	CCG	CAT	ATA	GAT	TCA	AAG	CCA	TGG	TGT	TAC	GTT	TTT	AAG	GGG	GGG	AAI	TAC	AGC	TCA	GAG	TTT	TCC	AGC	AGC					
氨基酸	G	H	S	D	C	Y	F	G	N	G	S	A	V	R	G	T	H	S	L	T	E	S	G	A	S	C				
原序列	CGA	AAC	AGT	GAC	TGT	TAC	TTT	GGG	AAT	GGG	TCA	GCT	TAC	DGT	GGC	AGC	CAA	AGC	CTC	ATC	GAG	TTT	GGT	GTC	TTT					
改造序列	CGA	AAT	TCT	GAT	TGT	TAC	TTT	GGA	AAT	GGA	TCA	GCT	TAC	AGA	GGC	ACT	CAT	TCT	TTG	ACT	GAA	TCT	GGA	DCT	DCT					
氨基酸	S	M	L	I	Q	K	Y	V	T	A	Q	N	S	A	Q	A	L	L	L	L	L	G	K	H	N	V				
原序列	TCC	ATG	ATC	ATC	ATA	GCC	AAG	GTT	TAC	ACA	CCA	CAG	AAC	CCS	AGT	BCC	CAG	CCA	CTG	GGC	CTG	GGC	AAA	CAT	AAT					
改造序列	TCT	ATC	ATP	TTG	ATC	GUA	AAG	GTT	TAC	ACA	GCT	CAA	AAT	CCA	TCT	GCT	CAA	GCT	TTG	GGA	TTG	GGA	AAA	CAT	AAT					
氨基酸	D	G	D	A	K	P	W	E	H	V	L	K	N	R	R	L	T	W	E	Y	C	D	V	P	S					
原序列	CAT	GGT	GAT	GGG	AAG	CCC	TGG	TGC	CAC	CTG	CTG	ATG	AAC	CGC	AGG	CTG	ACG	TGG	GAG	TAC	TGC	GAT	GTG	CCC						
改造序列	GAT	GGA	GAT	GCT	AAG	DCA	TGG	TGT	CAT	GT	TTG	AAQ	AAT	AGA	AGA	TTG	ACT	TGG	GAA	TAL	TGC	GAT	GTT							
氨基酸	L	F	Q	V	S	Q	P	Q	J	R	I	K	G	G	L	P	A	D	L	S	E	P	W	Q	A					
原序列	CTG	AGA	CAN	TAC	AGC	GAG	CCG	CCG	CAI	GGT	ATC	AAA	GGA	GGG	CTG	TTK	GAC	ATC	AGC	TCC	CAC	CCG	CCG	GGT						
改造序列	TTG	AGA	CAA	TAC	TCT	CAA	CCA	CAA	TTT	AGA	ATT	AAA	DGA	DGA	TTG	TTT	GCT	DAT	ATP	GCT	TTT	DAT	CCG							
氨基酸	K	I	R	R	S	P	G	L	R	J	L	C	G	G	L	L	I	S	S	C	W	J	S	A						
原序列	AAG	DAC	AGC	AGG	TTG	TTG	GGG	GGG	TTG	CTG	TGC	GGG	GGT	ATA	CTG	ATK	AUX	TCX	TCC	TGG	ATC	CTT	TTI							
改造序列	AAG	CAT	AGA	AGA	TCT	CCA	GGA	GAA	AGA	TTT	TTG	TGT	GGA	DGA	ATT	TTG	ATT	TCT	TCT	TGT	TGT	TGT								
氨基酸	E	K	F	P	P	H	H	C	T	V	L	D	R	I	Y	H	V	P	P	D	C	F	Q	K						
原序列	GAG	AGC	TTT	CCG	CCG	CAC	CTG	AGC	ATT	ATC	TTC	GGC	ANA	ACA	TAC	GGT	GGT	GGT	GGT	GGT	GGT	GGT								
改造序列	GAA	AGC	TTT	CCA	CCA	CAT	CAT	TTG	ACA	GTT	ATT	TTG	GGA	AGA	ACA	TAC	AGA	GTT	GTT	CCA	GSA	GAA								
氨基酸	K	Y	I	V	H	K	E	F	D	D	D	T	Y	D	N	D	I	A	L	T	D	I	K							
原序列	AAG	TAC	ATC	GGT	CAT	AAG	GAA	TTG	GAT	GAT	GAC	ACT	TAC	GAC	AAT	GAC	ATT	GGG	CTG	CTG	CAG	CTG								
改造序列	AAI	TAC	ATT	GCT	CAT	AAG	GAA	TTT	GAT	GAT	GAT	ACT	TAC	GAT	AAT	GAT	ATT	GCT	TTG	TTG	CAA	TTG								
氨基酸	Q	L	S	S	V	V	T	Y	C	L	P	A	D	L	L	L	L	L	L	L	L	L	L							
原序列	AAT	AGC	AGC	AGT	GGT	GGT	GGT	GGT	GGT	GGT	GGT	GGT	GGT	GGT	GGT	GGT	GGT	GGT	GGT	GGT	GGT	GGT								
改造序列	CAA	GAA	TCT	TCT	GTT	GTT	ATA	ACT	GTT	TGT	TTT	DCA	DCA	CCG	GAT	TTG	CAA	TTG	CCA	GAT	TTG									
氨基酸	K	I	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L								
原序列	AAG	DAT	DAG	GGC	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT									
改造序列	AAG	CAT	GAA	GCT	TTG	TCT	CCA	TTT	TAC	TCT	DAA	AGA	TTR	AGG	GAA	GGT	CAI	GGT	ADA	TTT	TTG									
氨基酸	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L									
原序列	TTA	CGT	AAC	AGA	ACA	GTG	ACC	GAC	AAC	ATG	CTG	TGT	GCT	GGA	BAC	ACT	GGG	AGC	GGC	GGG										
改造序列	TTG	TTG	AAT	AGA	ACA	GCT	ACT	GGT	AAT	ATG	TTG	ZGT	GCT	GGA	DAT	ACT	AGA	TCT	GSA	GGA										
氨基酸	G	D	S	G	G	P	L	V	C	L	N	D	D	B	M	T	L	V	G	I	L	S								
原序列	GAC	GAU	TGG	GGA	GGC	CCC	CTG	GGT	GGT	GGT	CTG	ATC	GAT	GGC	GGC	ATG	ACT	TTG	GTG	GGC										
改造序列	GGC	GAT	TCT	DGA	DGA	CCA	TTG	GGT	GGT	GGT	ATG	AAT	GAT	GGA	AGA	ATG	ACT	TTG	GGT	GGT										
氨基酸	D	V	E	G	G	V	Y	T	K	V	T	N	Y	L	D	W	L	R	D	S	X									
原序列	DAT	GTC	CCC	GGT	GGT	GGT	TAC	ACC	GGT	GGT	GGT	GGT	GGT	GGT	GGT	GGT	GGT	GGT	GGT											
改造序列	DAT	GTT	CCA	GGA	DTT	TAC	ACT	AAG	GTT	ACT	AAT	TAC	TTR	DAT	TGG	ATT	AGA	GAT	AAT											

图 1 t-PA 基因密码子改造结果

Fig. 1 The results of codons modification of t-PA gene

3 讨论

3.1 生物反应系统

基因工程类蛋白质药物的生产一直是基因工程技术的重要研究领域之一,目前许多生物都被尝试着用于这类药物的生产,例如微生物、动物细胞、转基因动物及转基因植物等。作为生物反应系统,这些生物均存在着不同的优缺点。微生物表达系统的产量较高,但是不能对真核生物蛋白质进行准确的翻译后加工和糖基化修饰,并且在发酵过程中常常产生一些不溶性聚合物,而将这些聚合物重新溶解并折叠成天然蛋白质需要很高的成本。此外,发酵常需要庞大的设备投资。动物细胞培养系统在一定程度上能满足真核蛋白质的翻译后加工和折叠,但动物细胞培养系统所需要的培养基相当昂贵。而以转基因动物作为生产系统技术难度高,并可能会成为更多公众或伦理关注的焦点。相对来说植物表达系统——植物生物反应器既能满足真核蛋白质的翻译后加工和折叠,又以相对低廉的生产成本而成为一个极具吸引力的生产系统,已经被广泛用于次生代谢物、糖类、脂类、蛋白质类药物、疫苗等的生产,并有可能替代外源蛋白质的发酵生产系统。然而,目前在植物表达系统的应用中却存在着目标产物产量低的问题,从而限制了植物生物反应器的应用。

为了提高外源蛋白在转基因植物中的表达量,目前主要通过如下 3 条途径来解决:(1)在转录水平上,使用高表达启动子、增强子等表达元件。Ichiro Mitsuhara^[9]通过对 CaMV 35S 启动子核心序列及其 5' 上游 3 个不同序列、菜豆蛋白基因内含子 1 序列和 Ω 序列进行组合研究,构建了一系列嵌合启动子,相对于 35S 启动子使 GUS 表达活性提高了 20~70 倍。(2)在翻译水平上,应用翻译增强元件和植物偏爱密码子。Joshi CP^[10]研究发现在植物起始密码子周围存在着共有序列 AACAAATGG,该序列同 Kozak 序一样是核糖体在与起始密码子结合之前结合的序列,具有增强翻译的作用。Mason^[11]等将不耐热肠毒素 B 亚基(LT-B)基因密码子改造成植物偏爱密码子,使 LT-B 在马铃薯叶和块茎中表达量提高了 1~2 倍。(3)在翻译后水平上,将转基因产物定向表达到一个合适的细胞分隔,例如细胞质、质外体和内质网腔^[12]。Wandelt^[13]等将四聚肽 KDEL 序列与豌豆球蛋白基因融合并转化烟草和苜

蓿,转基因烟草和苜蓿叶中豌豆球蛋白的积累水平提高 20 到 100 倍,而豌豆球蛋白—KDEL 与内源豌豆球蛋白在 mRNA 水平或翻译稳定性上没有差异。其积累水平的提高主要是因为豌豆球蛋白 C 端的 KDEL 可引导蛋白质定位到宿主植物的内质网中,使其在植物中的降解水平降低。

3.2 生物密码子分析

生物的遗传密码具有简并性,每一种氨基酸可由 1~6 个密码子编码,对于不同生物体及同一生物体的不同种类基因,由于 tRNA 丰度、(G+C)% 和 wobble 摆动效应等因素的差异,致使在密码子的使用上存在着偏爱性。日本 Kazusa DNA 研究所根据 NCBI-GenBank 中的基因资源得到了密码子用法数据库^[14],该数据库以密码子频率为参数列出了 16591 种生物的密码子使用表,并且随着新基因的发现而不断更新统计结果。由于该数据库分析的样本量大并能得到样本的及时更新,而使密码子使用频率能够有效反映生物体密码子用法。但对于分析生物体中含量较少的高表达基因和已知基因较少的生物的密码子用法时,由于可用于分析的样本数较少,密码子使用频率就不能很好地反映真实情况了。RSCU 能有效地反映样本中各个同义密码子的相对使用数目,引入这一参数可以实现在不同大小和氨基酸组分的样本间进行密码子频率比较^[5,6]。赵翔^[5,15]等采用参数 RSCU 对酵母 *Yarrowia lipolytica* 和 *Pichia pastoris* 中 43 和 28 个蛋白质编码基因进行分析,得出了两种酵母各自的密码子用法表和高表达优越密码子(17 个和 19 个)。

目前在密码子分析上常使用的软件是 GCG 软件包中的 codon frequency 命令,例如 Elizabeth E. Murray^[16]等采用该软件对来自单、双子叶植物的 207 个基因进行分析,得出了包括玉米、水稻和大豆等在内的高等植物密码子用法表。

本研究采用 Perl 语言编写了可以进行密码子频率统计、 N_c 值和 RSCU 值计算的编程软件,进行了马铃薯密码子用法分析,其结果与 <http://www.kazusa.or.jp/codon/> 提供的以密码子频率为参数的马铃薯密码子用法表相符。并且通过对高低表达样本组的 RSCU 值进行 t 检验,得出了马铃薯的高表达优越密码子 4 个。根据马铃薯密码子用法和高表达优越密码子分析结果,对 t-PA 基因序列进行密码子改造,设计出具有马铃薯密码子使用特点的 t-

PA 基因序列,下一步将通过人工合成方法获得该序列。从而为以马铃薯为生物反应器高效生产 t-PA 奠定分子基础。

参 考 文 献 (References):

- [1] Fiedler U, Conrad U. High-level production and long-term storage of engineered antibodies in transgenic tobacco seeds. *Biotechnology (N Y)*, 1995, 13(10):1090~1093.
- [2] Hockem A A, Kastellin R A, Vasser M, de Boer H A. Codon replacement in the PGK1 gene of *S. cerevisiae*: Experimental approach to study the role of biased codon usage in gene expression. *Mol Cell Biol*, 1987, 7:2914~2924.
- [3] ZOU He-Chang. The development and application of. *Chin Pharm J*, 1997, 5(32):163~267.
邹和昌. 溶栓剂的发展与应用. 中国药学期刊, 1997, 5(32): 163~267.
- [4] Frank W. The effective number of codon's used in a gene. *Gene*, 1990, 87:23~29.
- [5] ZHAO Xiang, LI Zhi, LU Shen-Feng, HUO Ke-Ke, LI Yu-Yang. Synonymous codon usage in *Yarrowia Lipolytica*. *Journal of Fudan University (Natural Science)*, 1999, 38(5):510~516.
赵 翔, 李 至, 陆身枫, 霍克克, 李育阳. 酵母 *Yarrowia Lipolytica* 的密码子用法分析. 复旦学报(自然科学版), 1999, 38(5):510~516.
- [6] GUO Zheng, LI Xia, LI Jing. Computational Molecular Biology and Genome Informatics. Harbin: Science and Technology Publishing Company of Heilongjiang Province, 1998, 152~153.
郭 政, 李 霞, 李 晶. 计算分子生物学与基因组信息学. 哈尔滨:黑龙江科学技术出版社, 1998, 152~153.
- [7] <http://www.bioperl.org>
- [8] Harris T J, Pate T, Marston F A, Little S, Emtage J S, Opdenakker G, Volckaert G, Rombauts W, Billiau A, de Somer P. Cloning of cDNA coding for human tissue-type plasminogen activator and its expression in *Escherichia coli*. *Mol Biol Med*, 1986, 3(3):279~292.
- [9] Ichiro Mitsuahara, Masashi Ugaki, Hirohiko Hirochika, et al. Efficient promoter cassettes for enhanced expression of foreign genes in dicotyledonous and monocotyledonous plants. *Plant Cell Physiology*, 1996, 37(1):49~59.
- [10] Joshi C P, Zhou H, Huang X. Context sequences of translation initiation codon in plants. *Plant Molecular Biology*, 1997, 35: 993~1001.
- [11] Mason H S, Haq T A, Clements Arntxen C J. Edible vaccine protects mice against *Escherichia coli* heat-labile enterotoxin (LT): potatoes expressing a synthetic LT-B gene. *Vaccine*, 1998, 16(13):1336~1343.
- [12] Sijmons P C, Dekker B M M, Schrammeijer B, Verwoerd T C, Van den Elzen P J, Hoekema A. Production of correctly processed human serum albumin in transgenic plants. *Biotechnology (N Y)*, 1990, 8(3):217~221.
- [13] Wandelt C I, Khan M R I, Craig S, Schroeder H E, Spencer D, Higgins T J. Vicilin with carboxy-terminal KDEL is retained in the endoplasmic reticulum and accumulates to high levels in the leaves of transgenic plants. *Plant Journal*, 1992, 2:181~192.
- [14] <http://www.kazusa.or.jp/codon/>
- [15] ZHAO Xiang, HUO Ke-Ke, LI Yu-Yang. Synonymous codon usage in *Pichia Pastoris*. *Chinese Journal of Biotechnology*, 2000, 16(3):308~311.
赵 翔, 霍克克, 李育阳. 毕赤酵母的密码子用法分析. 生物工程学报, 2000, 16(3):308~311.
- [16] Murray E E, Lotzer J, Eberle M. Codon usage in plant gene. *Nucl Acides Res*, 1989, 17:477~498.

科学出版社新书介绍

结构生物学与药学研究 杨 铭 主编 2003年11月出版 定价:48.00元

结构生物学是以生物大分子三维结构及其运动性的研究为基础,定量阐明生命现象的学科。药物的合理设计、新药的发现都以结构生物学的研究成果为基础。本书概述了结构生物学的研究现状和发展趋势,从分子水平上探讨主要生物大分子的三维结构与生物功能的关系及药学研究前沿领域中的一些重要科学问题,最后介绍了结构生物学研究中的主要方法。

实验室生物安全手册 马文丽 郑文岭 主编 2003年11月出版 定价:24.00元

本书综述了生物安全1~4级(biological safety level, BSL)标准和特殊微生物学操作规程、安全设备和设施,目的是帮助实验室将研究生物病原体的过程严格控制在生物安全标准范围内有序进行。

鲢、鳙与藻类水华控制 谢 平 著 2003年11月出版 定价:38.00元

本书是一部论述滤食性鱼类——鲢、鳙与藻类水华控制的专著。作者从鲢、鳙的自然分布和向国外的引种历史,鲢、鳙养殖的起源及在中国池塘混养系统中的地位入手,然后过渡到关于用鲢、鳙控制池塘养殖系统、富营养水库及小型人工湖中的藻类以改善水质的国内外的实验研究,并详细地介绍了在武汉东湖用鲢、鳙控制蓝藻水华的成功实践,仔细地分析了国内外关于鲢、鳙的食性及鲢、鳙对藻类消化机制方面的研究。

邮购地址:100717 北京东黄城根北街16号 科学出版社科学分社 联系人:阮 志 电话/传真:010-64034622