

DOI: 10.11931/guihaia.gxzw201509024

段永波,赵丰兰,姚磊,等. 重组人 *TNFR-Fc* 基因水稻种子特异表达载体构建 [J]. 广西植物, 2016, 36(5):589-594

DUAN YB, ZHAO FL, YAO L, et al. Construction of rice seed-specific expression vector expressing rice-preferred codon-optimized *TNFR-Fc* gene [J]. *Guihaia*, 2016, 36(5):589-594

## 重组人 *TNFR-Fc* 基因水稻种子特异表达载体构建

段永波, 赵丰兰, 姚磊, 滕井通, 盛玮, 张爱民, 薛建平\*

( 淮北师范大学 生命科学学院/资源植物生物学安徽省重点实验室, 安徽 淮北 235000 )

**摘要:** 重组人 II 型肿瘤坏死因子受体-抗体融合蛋白(TNFR-Fc)在治疗系统性自身免疫疾病中有很好的功效。目前 TNFR-Fc 主要通过动物细胞系生产,高昂的成本限制了其大规模应用。使用水稻种子作为生物反应器有望大幅度降低 TNFR-Fc 的生产成本,提供优质足量的目标产品。为使 *TNFR-Fc* 基因能够在水稻中高效表达,该研究拟构建经水稻偏好密码优化的 *TNFR-Fc* 基因的种子特异表达载体。通过对水稻和人全基因组密码子使用偏好性进行比较分析并按照水稻最优密码对 TNFR-Fc 氨基酸序列进行逐一优化,命名为 *RfTNFR-Fc*,通过全基因合成法制备该基因片段;同时采用 PCR 法扩增水稻种子特异表达启动子 *Glu-4*,构建种子特异表达启动子驱动的 *RfTNFR-Fc* 基因表达载体。结果表明:水稻与人多数氨基酸中密码子使用偏好性较为一致,但在 L、S、P、R 这 4 种氨基酸的密码子使用偏好性差异较大,优化时对这些密码子进行逐一替换,优化后 30.8% 的氨基酸密码子发生了变化;PCR 扩增获得种子特异启动子,并连接至 pCAMBIA1381 载体,同时将全基因合成法制备的 *RfTNFR-Fc* 基因连入该载体,PCR、双酶切验证及测序分析表明载体成功构建。该研究构建的 *RfTNFR-Fc* 基因种子特异表达载体,为在水稻种子中大规模生产 TNFR-Fc 奠定了基础。

**关键词:** 水稻, 重组人 II 型肿瘤坏死因子受体-抗体融合蛋白(TNFR-Fc), 偏好密码优化, 种子特异表达, 载体构建

中图分类号: Q943.2, Q78, S18 文献标识码: A 文章编号: 1000-3142(2016)05-0589-06

## Construction of rice seed-specific expression vector expressing rice-preferred codon-optimized *TNFR-Fc* gene

DUAN Yong-Bo, ZHAO Feng-Lan, YAO Lei, TENG Jing-Tong,  
SHENG Wei, ZHANG Ai-Min, XUE Jian-Ping\*

( Key Laboratory of Resource Plant Biology of Anhui Province, College of Life Sciences, Huaibei Normal University, Huaibei 235000, China )

**Abstract:** Recombinant human  $\alpha$  receptor antibody- protein of tumor necrosis factor Type II fusion protein (TNFR-Fc) is an effective mean to treat the systemic autoimmune diseases. This protein is currently produced via animal cell lines. The high cost of bioreactor maintenance brings an economic burden for common patients, which largely limits the application of TNFR-Fc. Rice seed offers a good expression system for production of recombinant proteins. To efficiently express *TNFR-Fc* in rice seed, we constructed seed-specific expression vector of rice-preferred codon optimized *TNFR-Fc*

收稿日期: 2015-09-26 修回日期: 2015-12-29

**基金项目:** 国家自然科学基金(31501368, 81573518);安徽省自然科学基金(1408085MC58, 1608085MC52);安徽省教育厅省级高校自然科学基金重点项目和重大项目(KJ2014A226, KJ2015ZD35)[Supported by the National Natural Science Foundation of China (31501368, 81573518), Natural Science Foundation of Anhui Province, China (1408085MC58, 1608085MC52); Key Project of Natural Science Research of Colleges and Universities in Anhui Province, China (KJ2014A226, KJ2015ZD35)].

**作者简介:** 段永波(1981-),男,贵州贵阳人,博士,讲师,主要从事植物生物技术研究,(E-mail)yboduan@163.com。

\*通讯作者: 薛建平,博士,教授,主要从事药用植物生物技术研究,(E-mail)xuejp@163.com。

*Fc* (*RfTNFR-Fc*), by analyzing the whole genomic difference in codon preference and replacing bases with rice-preferred codons. Rice seed specific promoter Glu-4 was amplified via PCR and ligated into pCAMBIA 1381, followed by the ligation with full length synthetic *RfTNFR-Fc*. The construct was identified by PCR, double enzyme digestion and sequencing. The preferred codons of L, S, P and R differed between rice and human, though most of the amino acids shared the same preferred codons. In the optimized *RfTNFR-Fc*, 30.8% of the amino acids changed compared with the original version. PCR, double enzyme digestion and sequencing identified that pCAMBIA Glu -RfTNFR-Fc was successfully constructed. This study lays foundation for the expression of TNFR-Fc in rice seed, further facilitating its commercial production.

**Key words:** *Oryza sativa*, recombinant human  $\alpha$  receptor antibody-protein of tumor necrosis factor type II fusion protein (TNFR-Fc), preferred codon optimization, seed-specific expression, vector construction

人 II 型肿瘤坏死因子受体-抗体融合蛋白 (TNFR-Fc) 由 467 个氨基酸组成, 对系统性自身免疫疾病, 如类风湿关节炎、银屑病、强直性脊柱炎等都有很好的治疗效果 (Rothe et al, 1992; Stübgen, 2011)。临床上极低剂量的 TNFR-Fc 蛋白 (推荐成人使用剂量为每次注射 25 mg) 即可发挥功能。目前 TNFR-Fc 商业化生产主要依靠中国仓鼠卵巢细胞 (CHO) 表达体系 (Xu et al, 2011), 高昂的反应器维持费用及难以迅速扩大的生产规模使得该药用蛋白价格居高不下, 急需降低 TNFR-Fc 生产成本以满足巨大的市场需求。植物表达体系生产成本低、容易规模化、产品质量高、安全性高 (无动物病毒污染) 等优点 (Horn, 2012; Stoger et al, 2014), 具有很好的应用前景。

在诸多植物表达体系中, 水稻种子是生产药用蛋白最具潜力的系统之一 (Ou et al, 2014)。水稻遗传转化技术已较为成熟 (Mishimura et al, 2007; Duan et al, 2012), 为药用蛋白生产的“工厂” (转基因植株) 的建立奠定了基础。水稻种子已被用于特异表达白细胞介素 (Fujiwara et al, 2010)、螨螨过敏原 (Suzuki et al, 2011) 和人溶酶体酸性  $\beta$ -葡萄糖脑苷脂酶 (Patti et al, 2012) 等多种外源重组蛋白, 表达量介于  $7.0 \sim 60.0 \text{ mg} \cdot \text{g}^{-1}$  之间。采用水稻种子特异的强启动子 Gt13a 及其信号肽驱动人血清白蛋白基因表达, 目标蛋白表达量达到  $2.75 \text{ mg} \cdot \text{g}^{-1}$  糙米, 且理化结构及生化特征与人血浆中提取的蛋白基本一致 (He et al, 2011)。该体系能大幅度提高人碱性成纤维细胞生长因子表达水平 (An et al, 2013)。因此, 使用种子特异表达强启动子, 使蛋白在特定时间和空间内表达, 减少不必要的能量和原料消耗, 有利于提高目标蛋白表达水平。目前, 尚未发现有关在植物中表达 TNFR-Fc 的报道。本研究按照水稻遗

传密码使用偏好性对 *TNFR-Fc* 基因进行设计和优化, 构建水稻谷蛋白启动子驱动的种子特异表达载体, 以期发展为水稻种子 TNFR-Fc 表达系统及其大规模生产奠定基础。

## 1 材料与amp;方法

### 1.1 材料

供试的水稻材料日本晴种子为安徽省农业科学院水稻研究所李浩博士惠赠。植物表达载体 pCAMBIA1381 及大肠杆菌 (*Escherichia coli*) 菌株 DH5 $\alpha$  为资源植物生物学安徽省重点实验室保存。

PCR 扩增试剂为生工生物工程 (上海) 股份有限公司产品; 限制性内切酶 *Bam*HI、*Hind* III、*T*<sub>4</sub> DNA 连接酶均购自 New England BioLabs (NEB) 公司; DNA 凝胶回收试剂盒购自 AxyGen Biosciences 公司; pMD18-T vector 购自 TaKaRa 生物工程 (大连) 有限公司; 卡那霉素和氨苄青霉素钠盐等购自生工生物工程 (上海) 股份有限公司。引物和基因合成以及测序由生工生物工程 (上海) 股份有限公司完成。

### 1.2 方法

1.2.1 水稻与人遗传密码使用偏好性分析 水稻 (*Oryza sativa*) 和人 (*Homo sapiens*) 的密码子信息从密码子使用数据库 [www.kazusa.or.jp/codon](http://www.kazusa.or.jp/codon) 获得。水稻包括 34、132、283 个密码子, 编码 92188 CDS (coding DNA sequences); 而人类使用 40 662 582 个密码子, 编码 93487 CDS。对水稻和人的编码各氨基酸的同义密码子使用频率进行调查比较, 使用频率最高的密码子认定为偏好密码子。

1.2.2 *RfTNFR-Fc* 基因设计与合成 重组蛋白 TNFR-Fc 由人肿瘤坏死因子可溶性受体的细胞外配体结合部分 (TNFR, 235 AA) 与人类 IgG1 Fc 段 (232

AA)融合而成。基于水稻密码子使用偏好性,根据 *TNFR-Fc* 氨基酸序列设计重组 *TNFR-Fc* 基因(*RfT-NFR-Fc*),长度1 407 bp。同时在其5'端添加醇溶蛋白信号肽序列(Os07g0206400),人工合成后克隆至 pUC57 载体,酶切位点为 *Bam*HI/*Nhe*I,获得 *Bam*HI-sp-*Spe*I-*TNFR-Fc*-*Nhe*I。

1.2.3 水稻种子特异表达启动子克隆 采用谷蛋白 *GluB4* 基因(AK107238)启动子特异引物(Patti et al,2012),以日本晴基因组 DNA 为模板进行 PCR 扩增。引物序列为 *GluB4*F-aagcttcggaattctacagggttcctcgcgtaaga(下划线表示 *Hind*III 酶切位点)和 *GluB4*R-ggatccgggatccagctatttgaggatgtattgg(下划线表示 *Bam*HI 酶切位点)。PCR 扩增体系如下:5  $\mu$ L 10  $\times$  PCR buffer,4  $\mu$ L 25 mmol  $\cdot$  L<sup>-1</sup> MgCl<sub>2</sub>,4  $\mu$ L 2.0 mmol  $\cdot$  L<sup>-1</sup> dNTPs,2  $\mu$ L 10  $\mu$ mol  $\cdot$  L<sup>-1</sup> 引物(上、下游引物各 1  $\mu$ L),0.4  $\mu$ L 5 U  $\cdot$   $\mu$ L<sup>-1</sup> *Taq* DNA 聚合酶,100 ng DNA 模板,ddH<sub>2</sub>O 至 50  $\mu$ L。PCR 扩增条件为 94  $^{\circ}$ C 预变性 5 min;94  $^{\circ}$ C 30 s,60  $^{\circ}$ C 30 s,72  $^{\circ}$ C 90 s 进行 30 个循环;72  $^{\circ}$ C 延伸 7 min。PCR 产物连接至 pMD18-T vector,挑取 PCR 阳性克隆送生工生物工程(上海)股份有限公司测序。

1.2.4 *RfTNFR-Fc* 种子特异表达载体构建 提取含有 *Bam*HI-sp-*Spe*I-*TNFR-Fc*-*Nhe*I 片段的 pUC57 质粒进行 *Bam*HI/*Nhe*I 双酶切,回收1 482 bp 片段;同时采用 *Bam*HI/*Nhe*I 对 pCAMBIA1391 质粒进行线性化处理。用 TaKaRa T4DNA 连接酶将目标片段与线性化载体进行连接重组,提取阳性菌落质粒进行 PCR 扩增和 *Bam*HI/*Nhe*I 双酶切鉴定,鉴定正确的克隆送生工生物工程(上海)股份有限公司测序。

## 2 结果与分析

### 2.1 水稻与人密码子使用偏好性比较

对水稻与人全基因组密码子使用偏好性分析表明(图 1),水稻 CDSs 中 GC 含量为 55.3%,比人使用 GC 的频率(52.3%)高出 3.0%。这 2 个物种在密码子 1、2、3 位使用 GC 的频率也有差异。水稻在 1、2、3 位使用 GC 频率为 58.2%、46.0%和 61.6%,分别比人类高 2.5%、3.5%和 3.0%。

这 2 个物种全基因组范围内的密码子使用频率统计情况见表 1。对各氨基酸的简并密码子进行分析,各氨基酸使用频率最高的密码子作为偏好密码子。水稻偏好密码子:TTC (F)、CTC(L)、TCC(S)、

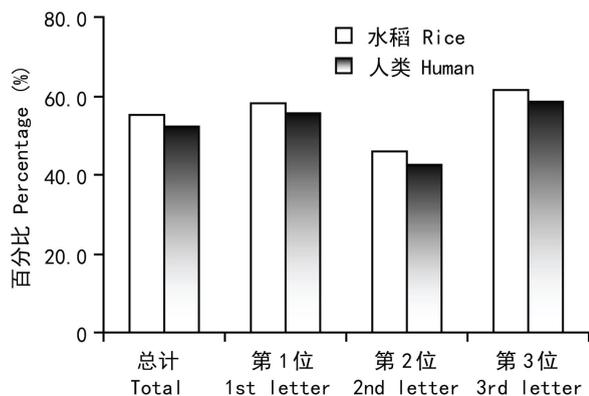


图 1 水稻与人全基因组编码区及各密码子位 GC 含量比较

Fig. 1 Comparison of GC contents in genetic codons and various positions between rice and human

TAC(Y)、TGC(C)、TGG(W)、CCG(P)、CAC(H)、CAG(Q)、CGC(R)、ATC(I)、ATG(M)、ACC(T)、AAC(N)、AAG(K)、GTG(V)、GCC(A)、GAC(D)、GAG(E)、GGC(G)、TGA(stop);人类偏好密码子:TTC(F)、CTG(L)、AGC(S)、TAC(Y)、TGC(C)、TGG(W)、CCC(P)、CAC(H)、CAG(Q)、AGA(R)、ATC(I)、ATG(M)、ACC(T)、AAC(N)、AAG(K)、GTG(V)、GCC(A)、GAC(D)、GAG(E)、GGC(G)、TGA(stop)。分析表明,水稻及人类编码 F、Y、C、W、H、Q、I、M、T、N、K、V、A、D、E、G 及终止密码的偏好密码子相同。L 在水稻的偏好密码为 CTG(使用频率 25.8%),而人类偏好密码为 CTC(39.6%),比水稻(21.0%)中高出 18.6%;S 在水稻的偏好密码为 TCC(使用频率 16.3%)而人类偏好密码为 AGC(19.5%);P 在水稻的偏好密码为 CCG(使用频率 18.0%),该密码子在人类中使用频率仅为 6.9%,而人类偏好密码为 CCC(19.8%);R 在水稻的偏好密码为 CGC(使用频率 16.1%),而人类偏好密码为 AGA(12.2%)。如直接以人源的原始核苷酸序列在水稻中进行基因表达,可能导致表达量极低。而因此有必要按照水稻密码子偏好性对 *TNFR-Fc* 进行逐一优化,以提高其在水稻中的表达水平。

### 2.2 *RfTNFR-Fc* 设计与合成

*TNFR-Fc* 氨基酸序列包括 235 aa 的 *TNFR* 和 232 aa 的 Ig gamma-1 chain C 片段,采用水稻偏好密码子对整个氨基酸序列进行逐一编码,获得 1 407 bp 长的基因片段 *RfTNFR-Fc*(图 2)。共有 144 个氨基酸的密码子发生改变,占全氨基酸序列的 30.8%。

表 1 水稻与人全基因组各密码子使用频率比较  
Table 1 Comparison of various genetic codons in rice and human

密码子 Codon	氨基酸 Amino acid	水稻 Rice		人 Human		密码子 Codon	氨基酸 Amino acid	水稻 Rice		人 Human	
		密码子 数目 Codon No.	使用频率 Use frequency (%)	密码子 数目 Codon No.	使用频率 Use frequency (%)			密码子 数目 Codon No.	使用频率 Use frequency (%)	密码子 数目 Codon No.	使用频率 Use frequency (%)
TTT	F	446063	13.1	714298	17.6	AGG		544515	16.0	486463	12.0
TTC		763386	<b><u>22.4</u></b>	824692	<b><u>20.3</u></b>	ATT	I	483941	14.2	650473	16.0
TTA	L	209911	6.1	311881	7.7	ATC		662207	<b><u>19.4</u></b>	846466	<b><u>20.8</u></b>
TTG		500913	14.7	525688	12.9	ATA		300085	8.8	304565	7.5
CTT		517919	15.2	536515	13.2	ATG	M	812432	<b><u>23.8</u></b>	896005	<b><u>22.0</u></b>
CTC		880959	<b><u>25.8</u></b>	796638	19.6	ACT	T	363451	10.6	533609	13.1
CTA		263871	7.7	290751	7.2	ACC		508156	<b><u>14.9</u></b>	768147	<b><u>18.9</u></b>
CTG		716525	21.0	1611801	<b><u>39.6</u></b>	ACA		394803	11.6	614523	15.1
TCT	S	433684	12.7	618711	15.2	ACG		388036	11.4	246105	6.1
TCC		557258	<b><u>16.3</u></b>	718892	17.7	AAT	N	515761	15.1	689701	17.0
TCA		424426	12.4	496448	12.2	AAC		633017	<b><u>18.5</u></b>	776603	<b><u>19.1</u></b>
TCG		420825	12.3	179419	4.4	AAA	K	544476	16.0	993621	24.4
AGT		300879	8.8	493429	12.1	AAG		1101342	<b><u>32.3</u></b>	1295568	<b><u>31.9</u></b>
AGC		545529	16.0	791383	<b><u>19.5</u></b>	GTT	V	529509	15.5	448607	11.0
TAT	Y	339638	10.0	495699	12.2	GTC		685971	20.1	588138	14.5
TAC		517042	<b><u>15.1</u></b>	622407	<b><u>15.3</u></b>	GTA		231761	6.8	287712	7.1
TGT	C	211609	6.2	430311	10.6	GTG		828681	<b><u>24.3</u></b>	1143534	<b><u>28.1</u></b>
TGC		422851	<b><u>12.4</u></b>	513028	<b><u>12.6</u></b>	GCT	A	667854	19.6	750096	18.4
TGG	W	472543	<b><u>13.8</u></b>	535595	<b><u>13.2</u></b>	GCC		1050723	<b><u>30.8</u></b>	1127679	<b><u>27.7</u></b>
CCT	P	463459	13.6	713233	17.5	GCA		591267	17.3	643471	15.8
CCC		411848	12.1	804620	<b><u>19.8</u></b>	GCG		908634	26.6	299495	7.4
CCA		486283	14.2	688038	16.9	GAT	D	863983	25.3	885429	21.8
CCG		613159	<b><u>18.0</u></b>	281570	6.9	GAC		959498	<b><u>28.1</u></b>	1020595	<b><u>25.1</u></b>
CAT	H	385174	11.3	441711	10.9	GAA	E	738891	21.6	1177632	29.0
CAC		470960	<b><u>13.8</u></b>	613713	<b><u>15.1</u></b>	GAG		1315826	<b><u>38.6</u></b>	1609975	<b><u>39.6</u></b>
CAA	Q	460234	13.5	501911	12.3	GGT	G	505609	14.8	437126	10.8
CAG		709469	<b><u>20.8</u></b>	1391973	<b><u>34.2</u></b>	GGC		1005701	<b><u>29.5</u></b>	903565	<b><u>22.2</u></b>
CGT	R	244821	7.2	184609	4.5	GGA		542264	15.9	669873	16.5
CGC		550575	<b><u>16.1</u></b>	423516	10.4	GGG		583357	17.1	669768	16.5
CGA		219662	6.4	250760	6.2	TAA	STOP	22360	0.7	40285	1.0
CGG		458602	13.4	464485	11.4	TAG		28508	0.8	32109	0.8
AGA		358226	10.5	494682	<b><u>12.2</u></b>	TGA		41361	1.2	63237	1.6

注: 使用频率指每 1 000 碱基对中该密码子出现的频率, 下划线黑体表示最优密码子。

Note: Use frequency is defined as the frequency of specific codons in every 1 000 base pairs, and the underlined bold data represent preferred codons.

ggatccatgaagatcattttcgtatttctctccttgctattgttgcgcatgcaatgcctctgcgactagtATGCTCCCTGCACAAGTTGC  
 TTTACACCCCTTACGCACCTGAGCCTGGTTCAACATGTGCGTTGCGCGAGTACTACGACCAGACCCGCACAAA  
 TGTGCTGCTCCAAGTGCTCCCCTGGACAACATGCTAAGGTGTTCTGCACCAAGACCTCCGACACCGTTTGC  
 GACTCTTGCAGGATTCCACCTACACCCAGCTTTGGAAGTGGGTGCCAGAATGCTTGTCTGCGGTTCAAG  
 ATGCTCCTCCGATCAGGTGGAAACCCAGGCTTGCACTAGAGAGCAGAACAGGATCTGCACCTGCCGGCCA  
 GGTGGTACTGTGCTTTGTCTAAGCAGGAGGGTTGCACTCTGCGCACCTTTGAGGAAGTGCCGCCCTG  
 GATTTGGAGTTGCTAGGCCTGGAAGTAACTTCGGATGTGGTGTGTAAGCCGTGTGCCCTGGAACATTC  
 TCGAACACAACGTCGTCGACGGATATCTGCAGGCCACATCAGATCTGCAACGTGGTGGCCATCCCAGGAA  
 ACGCCTCTATGGATGCGGTTTGTACGTCGACCTCGCCAACTAGGAGTATGGCGCCAGGGGCGGTACATTT  
 GCCACAACCAGTTTCAACACGCTCACAGCATAACGAGCCTACACCCAGAACATCAACGGCGCCATCAACAT  
 CATTCTCTTCCGATGGGCCCTCACCGCTGCGGAAGGCTCAACAGGCGATGAACCGAAGTCTTGTGA  
 TAAGACACATACATGTCCACCATGTCCCGCGCCCGAACTTCTTGCGGGCCCTCTGTTTTCTTTTCCCCC  
 CAAGCCCAAGGATACACTTATGATAAGTCGCACTCCCGAAGTACTTGTGTTGTGGTGGACGTTAGCCACG  
 AAGACCCCAAGTTAAGTTCAACTGGTACGTGGACGGGTACAGGTGCACAACGCGAAGACTAAGCCGC  
 GGGAGCAACAATAACAACAGCACGTACCGGGTCTGTCTGTCTCACAGTCTCCACCAAACTGGCTGGA  
 CGGGAAGGAGTATAAGTGCAAGTCTCTAATAAGGCCCTGCCCGCCCCGATTGAGAAGACTATTAGCAAA  
 GCCAAAGGGCAGCCCCGGGAGCCGCAAGTCTATACCCTGCCGCGTCTCGCGAGGAGATGACTAAAAAT  
 CAAGTCTCTGACGTGTCTGGTCAAAGGCTTTTATCCGAGCGATATTGCCGTCGAGTGGGAGAGCAATGG  
 CCAGCCGGAGAATAATTATAAGACGACGCCCGCTCTGGACAGCGACGGCAGCTTTTTCTGTATTCCA  
 AACTCACCGTCGACAAAAGCCGCTGGCAGCAGGGCAATGTCTTTCTGCGAGCGTGATGCACGAGGCCCT  
 CCACAATCACTATACCCAGAAATCCCTCAGCCTCTCGCCGGCAAATGAGCTAGC

图 2 *RfTNFR-Fc* 基因序列 小写字母表示醇溶蛋白信号肽序列; 下划线部分表示酶切位点。

Fig. 2 Sequence of synthetic *RfTNFR-Fc* Lowercase letters represent signal peptide of prolamin, and the underlined ones are the endonuclease recognition sites.

在 *RfTNFR-Fc* 5'端前加上醇溶蛋白信号肽序列和相应酶切位点, 获 1 482 bp 片段。全基因人工合成后导入克隆载体 pTC57, 获得 pUC57-*RfTNFR-Fc*。

### 2.3 种子特异启动子 *GluB-4* 克隆

以 *GluB-4* 特异引物进行 PCR 扩增, 获得 1 447 bp 片段(图 3:a)。回收该片段进行 *HindIII/BamHI* 双酶切, 将酶切产物与线性化处理的 pCAMBIA1381 载体进行重组。提取阳性菌落质粒进行酶切验证, 获得约 1 500 bp 特异条带(图 3:b), 说明 *GluB-4* 启动子片段成功连入表达载体, 获得种子特异表达载体 pCAMBIA-*Glu*。

### 2.4 *RfTNFR-Fc* 种子特异表达载体构建与验证

用 *BamHI/NheI* 同时双酶切处理 pUC57-*RfTNFR-Fc* 和载体 pCAMBIA1381, 分别回收小片段和大片段进行重组(用 *RfTNFR-Fc* 替换 *Gus*)。对重组质粒进行 *BamHI/NheI* 双酶切验证表明, *RfTNFR-Fc* 已成功连接至 pCAMBIA *Glu* 载体(图 4), 获得 pCAMBIA *Glu-RfTNFR-Fc*。测序结果验证表明种

子特异启动子驱动的水稻偏好密码子优化的 *TNFR-Fc* 植物表达载体 pCAMBIA *Glu-RfTNFR-Fc* 已成功构建。

## 3 讨论与结论

采用异源表达方式生产药用重组蛋白已成为药物生产从化学合成到生物合成转变的必要途径, 对于免疫自缺陷类疾病更是如此。在异源表达过程中, 宿主、分泌途径、启动子特性和密码子使用偏好性等因素都可能导致目标蛋白表达水平降低(蔡海莺等, 2013)。相关研究中必须考虑上述因素, 进而提高目标蛋白表达量。目前有关基因密码子优化的研究多集中在抗病虫害基因工程领域, 如将苏云金芽孢杆菌毒蛋白基因 *Cry1* 进行作物偏好密码子优化(周宗梁等, 2012; Song et al, 2014), 以提高植物中的表达水平。此类研究通常将微生物来源的基因优化后在高等生物中表达, 常需根据密码子偏好性

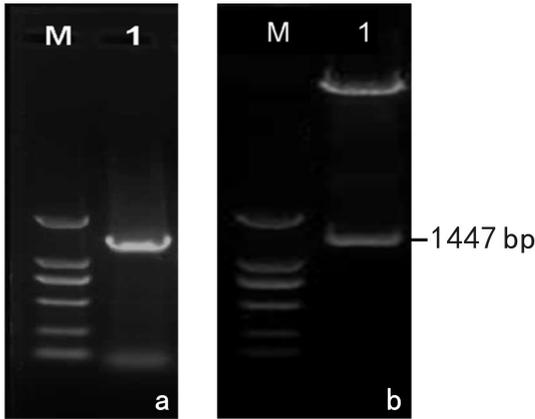


图 3 GluB-4 启动子克隆与表达载体 pCAMBIA 1381 构建 a. GluB-4 启动子 PCR 扩增结果; b. pCAMBIA-Glu 载体双酶切检测结果; M. 标准分子量 DL2000; 1. 产物。

Fig. 3 Cloning of GluB-4 promoter and ligation into expression vector pCAMBIA 1381 a. PCR product of GluB-4 promoter; b. Identification of pCAMBIA-Glu by double enzyme digestion; M. DNA marker DL2 000; 1. Product.

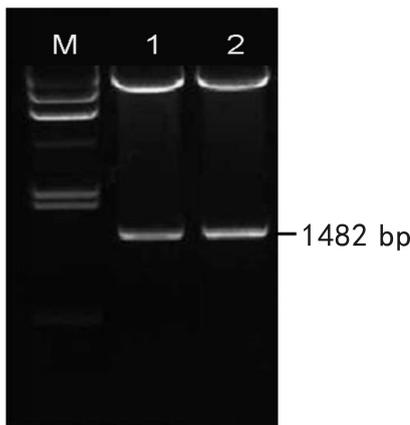


图 4 pCAMBIA Glu -RfTNFR-Fc 酶切验证 M.  $\lambda$  Hind III; 1, 2. 不同重组子克隆。

Fig. 4 Identification of pCAMBIA Glu -RfTNFR-Fc via double enzyme digestion. M.  $\lambda$  Hind III; 1, 2. Represent different clones of recombinants.

提高 G+C 含量,消除 attta (Nicolai et al, 2000)、A+T 富含区 (Goodall & Filipowicz, 1989) 及 polyA (Dean et al, 1986) 等易导致表达量低的因素。

通过对水稻与人全基因组密码子使用偏好性分析,发现水稻 CDSs 中 GC 含量为 55.3%, 而人使用 GC 的频率为 52.3%; 2 个物种多数氨基酸密码子使用情况较为一致,但 L、S、P、R 这 4 种氨基酸差异极大。推测高等生物间异源表达时基因优化的关键不在于消除 A+T 造成的序列不稳定,极可能是少数氨

基酸差异密码子的优化。本研究以人 II 型肿瘤坏死因子受体-抗体融合蛋白 (TNFR-Fc) 为对象,选择水稻种子特异表达系统,按照水稻密码子使用偏好进行优化,构建植物表达载体。PCR 法获得种子特异启动子并克隆至植物表达载体 pCAMBIA1381。全基因组范围分析水稻与人密码子使用偏好性并进行优化,全基因合成时在 RfTNFR-Fc 5' 端加入醇溶蛋白信号肽序列指导目标蛋白的分泌。将人工合成的 RfTNFR-Fc 连入 pCAMBIA-Glu, 酶切及测序验证表达载体成功构建,获得水稻种子特异表达载体 pCAMBIA-Glu- RfTNFR-Fc。本研究为以水稻种子为“工厂”生产 TNFR-Fc 及其理化结构和生物活性评价奠定了基础。

### 参考文献:

- AN N, OU J, JIANG D, et al, 2013. Expression of a functional recombinant human basic fibroblast growth factor from transgenic rice seeds [J]. *Int J Mol Sci*, 14(2), 3 556-3 567.
- CAI HY, LI Y, ZHANG H, et al, 2013. Effects of gene design on recombinant protein expression; a review [J]. *Chin J Biotechnol*, 29(9): 1 201-1 213. [蔡海莺,李杨,张辉,等,2013. 基因设计对重组蛋白表达的影响研究进展 [J]. *生物工程学报*, 29(9): 1 201-1 213.]
- DEAN C, TAMAKI S, DUNSMUIR P, et al, 1986. mRNA transcripts of several plant genes are polyadenylated at multiple sites in vivo [J]. *Nucl Acids Res*, 14(5): 2 229-2 240.
- DUAN Y, ZHAI C, LI H, et al, 2012. An efficient and high-throughput protocol for *Agrobacterium*-mediated transformation based on phosphomannose isomerase positive selection in *Japonica* rice (*Oryza sativa* L.) [J]. *Plant Cell Rep*, 31(9): 1 611-1 624.
- FUJIWARA Y, AIKI Y, YANG L, et al, 2010. Extraction and purification of human interleukin-10 from transgenic rice seeds [J]. *Protein Expr Purif*, 72: 125-130.
- GOODALL GJ, FILIPOWICZ W, 1989. The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing [J]. *Cell*, 58(3): 473-483.
- HE Y, NING T, XIE T, et al, 2011. Large-scale production of functional human serum albumin from transgenic rice seeds [J]. *PNAS*, 108: 19 078-19 083.
- HORN ME, 2012. Plant molecular pharming 2012 and beyond [J]. *Plant Cell Rep*, 31(3): 437-438.
- MISHIMURA A, AICHI I, MATSUOKA M, 2007. A protocol for *Agrobacterium*-mediated transformation in rice [J]. *Nat Protoc*, 1: 2 796 -2 802.
- NICOLAI S, AVIAH Z, MENACHEM K, et al, 2000. Synthetic *Bacillus thuringiensis* cry I C gene encoding insect toxin [P]: US, 6043415.2000-3-28.
- OU J, GUO Z, SHI J, et al, 2014. Transgenic rice endosperm as a bioreactor for molecular pharming [J]. *Plant Cell Rep*, 33(4): 585-594.

(下转第 581 页 Continue on page 581)