

HSP90 α 的表达均高于假手术组,因此推测 HSP90 α 对应的潜在 miRNA 表达量在缺血再灌注组应该降低,笔者的 miRNA - 144 实时定量 PCR 结果正好与这一推测吻合。所以 HSP90 α 可能是 miRNA - 144 调控的靶蛋白之一。这也与最近一项研究结果一致,miRNA - 144 在心脏缺血再灌注后表达降低^[11]。Pan 等^[6]研究观察到,miRNA - 1 转基因鼠心脏缺血再灌注模型中,miRNA - 1 表达增高抑制了 PKC ϵ 和 HSP60 等保护蛋白的表达,使心肌梗死面积增大,加重了心肌损伤。同样的研究还有 miRNA - 320 能通过抑制 HSP20 的表达而加重心脏缺血再灌注损伤^[7]。由此可以推断,miRNA - 144 表达下调是肾组织应对缺血再灌注损伤使 HSP90 α 表达上调的内源性保护机制,故敲除 miRNA - 144 可能成为治疗肾缺血再灌注损伤的新策略。

综上所述,HSP90 α 表达上调和 miRNA - 144 表达下调在缺血再灌注中都扮演着重要的角色,对于 miRNA - 144 是否真正调控 HSP90 α ,笔者将做进一步的研究验证,从而为临床上治疗肾缺血再灌注损伤提供新的思路和理论基础。

参考文献

- 1 Chen YT, Tsai TH, Yang CC, *et al.* Exendin - 4 and sitagliptin protect kidney from ischemia - reperfusion injury through suppressing oxidative stress and inflammatory reaction [J]. *J Transl Med*, 2013, 11 (11):270
- 2 Fujino T, Muhib S, Sato N, *et al.* Silencing of p53 RNA through transarterial delivery ameliorates renal tubular injury and downregulates GSK - 3 β expression after ischemia - reperfusion injury [J].

- Am J Physiol Renal Physiol*, 2013, 305 (11): F1617 - F1627
- 3 McClellan AJ, Xia Y, Deutschbauer AM, *et al.* Diverse cellular functions of the Hsp90 molecular chaperone uncovered using systems approaches [J]. *Cell*, 2007, 131 (1): 121 - 135
- 4 Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function [J]. *Cell*, 2004, 116 (2): 281 - 297
- 5 Wu K, Xu W, You Q, *et al.* Increased expression of heat shock protein 90 under chemical hypoxic conditions protects cardiomyocytes against injury induced by serum and glucose deprivation [J]. *Int J Mol Med*, 2012, 30 (5): 1138 - 1144
- 6 Pan Z, Sun X, Ren J, *et al.* miR - 1 exacerbates cardiac ischemia - reperfusion injury in mouse models [J]. *PLoS One*, 2012, 7 (11): e50515
- 7 Ren XP, Wu J, Wang X, *et al.* MicroRNA - 320 is involved in the regulation of cardiac ischemia/reperfusion injury by targeting heat - shock protein 20 [J]. *Circulation*, 2009, 119 (17): 2357 - 2366
- 8 Biermann J, Lagreze WA, Schallner N, *et al.* Inhalative preconditioning with hydrogen sulfide attenuated apoptosis after retinal ischemia/reperfusion injury [J]. *Mol Vis*, 2011, 17 (1): 1275 - 1286
- 9 Jha S, Calvert JW, Duranski MR, *et al.* Hydrogen sulfide attenuates hepatic ischemia - reperfusion injury: role of antioxidant and anti-apoptotic signaling [J]. *Am J Physiol Heart Circ Physiol*, 2008, 295 (2): H801 - H806
- 10 Barrera - Chimal J, Perez - Villalva R, Ortega J A, *et al.* Intra - renal transfection of heat shock protein 90 alpha or beta (Hsp90alpha or Hsp90beta) protects against ischemia/reperfusion injury [J]. *Nephrol Dial Transplant*, 2014, 29 (2): 301 - 312
- 11 Li J, Rohailla S, Gelber N, *et al.* MicroRNA - 144 is a circulating effector of remote ischemic preconditioning [J]. *Basic Res Cardiol*, 2014, 109 (5): 423

(收稿日期:2014 - 10 - 17)

(修回日期:2014 - 10 - 30)

埃博拉病毒包膜糖蛋白的密码子偏爱性分析

赵玉娇 黄新伟 李 多 邱丽娟 孙强明

摘 要 目的 对从 2014 年西非埃博拉暴发地区感染者分离的 4 个埃博拉病毒株基因组包膜糖蛋白编码基因进行密码子偏爱性分析,为埃博拉病毒包膜蛋白基因表达中宿主系统选择和密码子优化提供参考。**方法** 通过 GeneBank 搜索近期公布的分离自西非地区的 4 个埃博拉病毒株全基因组序列,获得编码包膜糖蛋白 GP1 和 GP2 的基因序列及蛋白编码序列。采用 EMBOSS 在线预测埃博拉病毒包膜糖蛋白的密码子偏爱性,并通过与 Codon Usage Database 中大肠杆菌、酵母及人的密码子使用频率进行比较,筛选最适埃博拉病毒 GP 蛋白体外表达的宿主物种,并提供相应密码子优化方案。**结果** 埃博拉病毒基因组、GP1 和 GP2 的 Nc 值均在 57 左右,CAI 值均在 0.7 左右,表明其密码子偏爱性较均一,且基因的表达水平相对较高;编码 GP1 蛋白 F、

基金项目:国家自然科学基金资助项目(面上项目)(81171946);云南省科技厅社会发展科技计划项目(2011CA016);云南省自然科学基金资助项目(2009ZC187M,2012FB188);中国医学科学院医学生物学研究所重大专项基金资助项目(2014MB05ZD)

作者单位:650118 昆明,中国医学科学院/北京协和医学院医学生物学研究所、云南省重大传染病疫苗研发重点实验室

通讯作者:孙强明,电子信箱:msun08@yahoo.com

H、M、N、W 等氨基酸和 GP2 蛋白 D、K、N、Q、Y 等氨基酸中不同密码子使用频率有较大差异;GP1 和 GP2 蛋白密码子偏爱性有共同之处也存在明显差异。计算 $\sigma_{GP/E. coli}$ 、 $\sigma_{GP/Yeast}$ 、 $\sigma_{GP/Human}$ 的比值,结果显示 GP1 中 >2.0 或 <0.5 的数量分别为 24、20 和 19 个;GP2 中的数量分别为 26、30 和 19,表明埃博拉病毒 GP1 和 GP2 蛋白与人的密码子使用频率较为接近。**结论** 以密码子偏爱性分析作为辅助手段,分析出此次流行于西非的埃博拉病毒包膜糖蛋白的密码子偏爱性及最适表达物种及密码子优化方案,为体外表达 GP1 和 GP2 蛋白,从而研发埃博拉病毒检测试剂、亚单位疫苗及基因工程抗体提供参考。

关键词 埃博拉病毒 包膜糖蛋白 密码子偏爱性

中图分类号 R3

文献标识码 A

DOI 10.3969/j.issn.1673-548X.2015.03.010

Analysis of the Codon Bias of Ebolavirus Glycoprotein. Zhao Yujiao, Huang Xinwei, Li Duo, et al. Chinese Academy of Medical Sciences, Peking Union Medical College, Institute of Medical Biology, Yunnan Key Laboratory of Vaccine Research & Development on Severe Infectious Diseases, Yunnan 650118, China

Abstract Objective To provide the reference for the selection of gene expression host system of Ebola virus envelope protein, codon bias were analyzed according to envelope glycoproteins encoding gene of 4 Ebola virus strains isolated in west Africa in 2014. **Methods** Envelope glycoprotein GP1 and GP2 Gene sequences and related protein coding sequences were obtained, according to Gene Bank searching results of complete genome of 4 Ebola virus strains isolated from west Africa region recently. Envelope glycoproteins codon RMS values (Nc) and codon adaptation index (CAI) were calculated by EMBOS the CHIPS and the CAI modules online. The fraction and using frequency (frequency) of different codes in same amino acids of GP1 and GP2 were investigated by CUSP module analysis. In addition, the results were compared with the using frequency of *E. coli*, yeast and human. It will be helpful for selecting the optimal expression host system of Ebola virus envelope protein. **Results** The Nc value of Ebola virus genome, GP1 and GP2 was around 57, CAI value was around 0.7. It showed that the codon preference was uniform, and the gene expression level was relatively high. There was great difference in codon bias of amino acid, such as F, H, M, N, W in GP1 and D, K, N, Q, Y in GP2. The results of $\sigma_{GP/E. coli}$, $\sigma_{GP/Yeast}$, $\sigma_{GP/Human}$ showed that the codon usage frequency of Ebola virus GP1 and GP2 protein was more close to human. **Conclusion** Codon bias and gene expression host system of Ebola virus (isolated in west Africa in 2014) GP1 and GP2 were analyzed by codon bias analysis as an auxiliary means. This study may help selecting gene expression host system and contributes further to development of subunit vaccine, antibody and diagnostic reagents.

Key words Ebola virus; Glycoprotein; Codon bias

据世界卫生组织 (World Health Organization, WHO) 报道,埃博拉病毒 (Ebolavirus) 疫情在西非快速暴发,截止 2014 年 7 月 30 日已有 767 例死亡病例,最新数据似乎表明疫情正在失控^[1]。目前,为防止埃博拉病毒传入中国,边境检疫部门已加强相关检查。因此,迅速研究针对埃博拉病毒的检测试剂盒,及预防和治疗性疫苗和抗体药物,成为我国预防和控制埃博拉病毒迫在眉睫的工作。

埃博拉病毒属于丝状病毒科 (Filoviridae),是迄今为止传染性最高、致命性最强的病毒之一。病毒可透过与患者体液直接接触,或与患者皮肤、黏膜等接触而传染,能引起人类和灵长类动物产生严重的埃博拉出血热 (Ebola hemorrhagic fever, EHF),病毒潜伏期可达 2~21 天,但通常只有 5~10 天,病死率高达 90%^[2]。世界卫生组织已将 EBOV 列为生物安全第 4 级病毒,利用其毒株进行体外培养扩增具有极高风险。因此,利用基因工程技术及抗体技术等进行埃博拉病毒预防、治疗及检测制品开发是理想的选择。

本研究通过生物信息学方法,对埃博拉病毒基因

组编码包膜糖蛋白 (glycoprotein, GP) 的基因序列密码子偏爱性进行了分析,并分别与大肠杆菌、酵母及人的密码子偏爱性进行比较,为两种包膜糖蛋白进行体外表达选择合适的表达系统提供参考,为快速研发埃博拉病毒基因工程疫苗、治疗性抗体及检测试剂盒奠定基础。

材料与方法

1. 埃博拉病毒包膜糖蛋白基因序列:选取 GenBank 上公布的 2014 年暴发于西非扎伊尔,以出现出血热的埃博拉感染患者分离的 4 株病毒株基因组为研究对象,序列编号为 KM233100.1、KM233101.1、KM233114.1 和 KM233118.1。该基因组为单股负链 RNA,全长 18959bp,含 7 个开放阅读框,编码 7 种蛋白质。其中 GP 蛋白为 I 型跨膜蛋白,位于病毒颗粒表面,由 GP1 和 GP2 两个高度糖基化的亚单位形成三角裂杯状的同源三聚体,构成病毒的刺突蛋白。GP1 被认为在病毒和受体识别过程中具有重要作用,而 GP2 协助病毒跟细胞膜融合^[3~5]。GP1 与 GP2 编码基因序列见图 1。

2. 方法:(1)埃博拉病毒包膜糖蛋白密码子偏爱性分析:遗传信息在由 mRNA 到蛋白质的传递过程中同义密码子的使用概率并不相同。某一物种或某一基因通常倾向于使用一种

或几种特定的同义密码子,这些密码子被称为最优密码子,此现象被称为密码子偏爱性。本研究采用 EMBOSS 在线分析系统 (<http://emboss.toulouse.inra.fr/>) 中的 CHIPS 模块计算 GP 基因序列中的密码子有效值 (effective number of codons, Nc)。当前普遍通过比较 Nc 来确定内源基因表达量的相对高低。Nc 值反映的是 1 个基因中所用到的密码子种类的多少,其数值一般在 20 (每个氨基酸只使用 1 个密码子的极端情况) ~ 61 (各个密码子均被平均使用) 之间。已知高表达基因其密码子偏爱程度也大,从而 Nc 值较小;低表达基因则含有较多种类的稀有密码子, Nc 值也较大^[6]。采用 Codon Adaptation Index (CAI) 模块进行密码子适应指数分析。CAI 常用于基因表达水平的测量,此值为 0 ~ 1,越接近 1 表示基因的表达水平越高^[7]。随后采用 CUSP 模块进行分析。CUSP 是对 1 个或多个 CDS 进行阅读后计算出密码子频率表。该程序可计算不同密码子在同 1 个氨基酸编码中所占的比例,并通过外延法计算密码子在编码基因中出现的频率^[8,9]。(2) 埃博拉病毒包膜糖蛋白与细菌、酵母及人的密码子偏爱性比较:通过 Codon Usage Database (<http://www.kazusa.or.jp/codon/>) 获得大肠杆菌、酵母和人的密码子使用频率表^[10]。将埃博拉病毒包膜糖蛋白 GP1 和 GP2 基因的密码子使用频率分别与大肠杆菌、酵母及人的密码子使用频率相比较 (σ GP/*E. coli*、 σ GP/*Yeast*、 σ GP/*Human*), 分析这两个蛋白究竟与哪种生物类型的密码子偏爱性相近。

结 果

1. 埃博拉病毒包膜糖蛋白密码子有效值及密码子适应指数分析:将选取自 GenBank 的 4 个埃博拉病毒编码 GP1 和 GP2 的基因序列进行比对后发现,其基因编码序列一致,无突变位点。采用 CHIPS 及 CAI 计算埃博拉病毒基因组及 GP1 和 GP2 的 Nc 值和 CAI 值,结果显示,埃博拉病毒基因组、GP1 和 GP2 的 Nc 值都在 57 左右,接近 Nc 值范围的上限。这表明埃博拉病毒基因组及 GP 蛋白编码基因的密码子偏爱性较均一,各密码子编码氨基酸时的出现频率较一致。埃博拉病毒基因组、GP1 和 GP2 的 CAI 值都在 0.7 左右,表明其基因的表达水平相对较高。GP 编码基因信息和密码子有效值及密码子适应指数分析结果见表 1。

2. 埃博拉病毒包膜糖蛋白密码子频率表:采用 EMBOSS 中的 CUSP 模块分析埃博拉病毒包膜糖蛋白的密码子在同一个氨基酸编码中所占的比例,同时计算密码子在编码基因中出现的频率。如表 2 所示,阴影标志的密码子表示在编码该氨基酸中出现频率较高 (超出其他密码子频率 10%), 为埃博拉病毒 GP 蛋白的偏爱密码子;表 2 中方框标志的密码子则

表 1 GP 蛋白编码基因相关信息和密码子 Nc 值及 CAI 值分析 ($\bar{x} \pm s$)

名称	基因组	GP1 蛋白	GP2 蛋白
基因位置	1 ~ 18959	5918 ~ 6802	6802 ~ 7947
核苷酸长度 (bp)	18959	885	1146
氨基酸长度 (aa)	—	294	382
Nc 值	55.019	56.826	57.893
CAI 值	0.659	0.700	0.706

表示出现频率较低 (使用频率 < 10%)。结果显示,编码埃博拉病毒 GP1 蛋白的苯丙氨酸 (phenylalanine, F)、组氨酸 (histidine, H)、蛋氨酸 (methionine, M)、天冬酰胺 (asparagine, N)、色氨酸 (tryptophan, W) 等氨基酸不同密码子使用频率有较大差异 (某种密码子使用频率 $\geq 70\%$), GP2 蛋白中编码天冬氨酸 (tryptophan, D)、赖氨酸 (lysine, K)、天冬酰胺、谷氨酰胺 (glutamine, Q)、酪氨酸 (tyrosine, Y) 等氨基酸的不同密码子使用频率有较大差异。如在编码甘氨酸 (glycine, G) 的 4 种密码子中 GP1 和 GP2 的偏爱密码子均为 GGA, 比例达 41.7% 和 41.4%; 在编码赖氨酸的 2 种密码子中 GP1 和 GP2 的偏爱密码子均为 AAA, 比例达 62.5% 和 73.3%; 而编码丙氨酸 (alanine, A) 的 4 种密码子中 GCG 只有 6.2% 和 3.0% 的比例。GP1 和 GP2 蛋白的偏爱性密码子也存在差异, 如天冬酰胺的偏爱密码子在 GP1 中为 AAT, 出现频率为 76.9%; 在 GP2 中则为 AAC, 出现频率为 68.0%。GP 编码基因的密码子频率分析结果见表 2。

3. 埃博拉病毒包膜糖蛋白与细菌、酵母及人的密码子偏爱性比较:采用 Codon Usage Database 比较埃博拉病毒 GP1 和 GP2 基因的密码子使用频率与大肠杆菌、酵母及人的密码子使用频率,通过计算 σ GP/*E. coli*、 σ GP/*Yeast*、 σ GP/*Human* 的比值来分析 GP 蛋白与 3 种生物类型的密码子偏爱性远近关系。生物密码子使用频率比值为 0.5 ~ 2.0 的表示对该密码子的偏爱性较为接近, <0.5 及 >2.0 之间则表示对该密码子的偏爱性较差。结果显示,GP1 蛋白与大肠杆菌、酵母及人的密码子使用频率比值 >2.0 或 <0.5 的数量分别为 24、20 和 19 个; GP2 的数量分别为 26、30 和 19 个。表明埃博拉病毒 GP1 和 GP2 蛋白与人的密码子使用频率较为接近。

表 2 GP 蛋白密码子频率表

密码子	编码氨基酸	比例数 - GP1	频率数 - GP1	数量 - GP1	比例数 - GP2	频率数 - GP2	数量 - GP2
GCA	A	0.312	16.949	5	0.364	31.414	12
GCC	A	0.188	10.169	3	0.303	26.178	10
GCG	A	0.062	3.390	1	0.030	2.618	1
GCT	A	0.438	23.729	7	0.303	26.178	10
TGC	C	0.400	6.780	2	0.571	10.471	4
TGT	C	0.600	10.169	3	0.429	7.853	3
GAC	D	0.692	30.508	9	0.682	39.267	15
GAT	D	0.308	13.559	4	0.318	18.325	7
GAA	E	0.444	27.119	8	0.529	23.560	9
GAG	E	0.556	33.898	10	0.471	20.942	8
TTC	F	0.700	47.458	14	0.556	13.089	5
TTT	F	0.300	20.339	6	0.444	10.471	4
GGA	G	0.417	33.898	10	0.414	31.414	12
GGC	G	0.125	10.169	3	0.138	10.471	4
GGG	G	0.208	16.949	5	0.207	15.707	6
GGT	G	0.250	20.339	6	0.241	18.325	7
CAC	H	1.000	13.559	4	0.462	15.707	6
CAT	H	0.000	0.000	0	0.538	18.325	7
ATA	I	0.214	10.169	3	0.185	13.089	5
ATC	I	0.429	20.339	6	0.370	26.178	10
ATT	I	0.357	16.949	5	0.444	31.414	12
AAA	K	0.625	33.898	10	0.733	28.796	11
AAG	K	0.375	20.339	6	0.267	10.471	4
CTA	L	0.120	10.169	3	0.074	5.236	2
CTC	L	0.080	6.780	2	0.148	10.471	4
CTG	L	0.280	23.729	7	0.333	23.560	9
CTT	L	0.240	20.339	6	0.111	7.853	3
TTA	L	0.080	6.780	2	0.185	13.089	5
TTG	L	0.200	16.949	5	0.148	10.471	4
ATG	M	1.000	3.390	1	1.000	7.853	3
AAC	N	0.231	10.169	3	0.680	44.503	17
AAT	N	0.769	33.898	10	0.320	20.942	8
CCA	P	0.438	23.729	7	0.200	10.471	4
CCC	P	0.250	13.559	4	0.450	23.560	9
CCG	P	0.188	10.169	3	0.250	13.089	5
CCT	P	0.125	6.780	2	0.100	5.236	2
CAA	Q	0.444	13.559	4	0.667	31.414	12
CAG	Q	0.556	16.949	5	0.333	15.707	6
AGA	R	0.353	20.339	6	0.375	15.707	6
AGG	R	0.176	10.169	3	0.188	7.853	3
CGA	R	0.176	10.169	3	0.250	10.471	4
CGC	R	0.000	0.000	0	0.062	2.618	1
CGG	R	0.176	10.169	3	0.000	0.000	0
CGT	R	0.118	6.780	2	0.125	5.236	2
AGC	S	0.100	6.780	2	0.259	18.325	7
AGT	S	0.250	16.949	5	0.259	18.325	7
TCA	S	0.300	20.339	6	0.111	7.853	3
TCC	S	0.200	13.559	4	0.185	13.089	5
TCG	S	0.050	3.390	1	0.037	2.618	1
TCT	S	0.100	6.780	2	0.148	10.471	4
ACA	T	0.480	40.678	12	0.370	44.503	17
ACC	T	0.160	13.559	4	0.261	31.414	12
ACG	T	0.160	13.559	4	0.109	13.089	5
ACT	T	0.200	16.949	5	0.261	31.414	12
GTA	V	0.091	6.780	2	0.214	7.853	3
GTC	V	0.318	23.729	7	0.143	5.236	2
GTG	V	0.227	16.949	5	0.214	7.853	3
GTT	V	0.364	27.119	8	0.429	15.707	6
TGG	W	1.000	20.339	6	1.000	20.942	8
TAC	Y	0.364	13.559	4	0.600	7.853	3
TAT	Y	0.636	23.729	7	0.400	5.236	2
TAA	*	0.000	0.000	0	0.000	0.000	0
TAG	*	0.000	0.000	0	1.000	2.618	1
TGA	*	0.000	0.000	0	0.000	0.000	0

方框所标密码子表示出现频率较低(使用频率 < 10%)

表 3 GP 蛋白密码子与细菌、酵母及人的密码子偏爱性比较

氨基酸	密码子	大肠杆菌	酵母	人类	GP1/ 大肠杆菌	GP1/ 酵母	GP1/ 人类	GP2/ 大肠杆菌	GP2/ 酵母	GP2/ 人类
		1/1000	1/1000	1/1000						
A	GCA	20.6	16.1	16.1	0.82	1.05	1.05	1.52	1.95	1.95
	GCC	25.5	12.5	28.4	0.40	0.81	0.36	1.03	2.09	0.92
	GCG	31.7	6.1	7.5	0.11	0.56	0.45	0.08	0.43	0.35
	GCT	15.6	21.1	18.6	1.52	1.12	1.28	1.68	1.24	1.41
C	TGC	6.9	4.7	12.2	0.98	1.44	0.56	1.52	2.23	0.86
	TGT	5.5	8.0	10.0	1.85	1.27	1.02	1.43	0.98	0.79
D	GAC	18.6	20.2	25.6	1.64	1.51	1.19	2.11	1.94	1.53
	GAT	32.1	37.8	21.9	0.42	0.36	0.62	0.57	0.48	0.84
E	GAA	38.2	48.5	29.0	0.71	0.56	0.94	0.62	0.49	0.81
	GAG	17.7	19.1	39.9	1.92	1.77	0.85	1.18	1.10	0.52
F	TTC	16.9	18.2	20.6	2.81	2.61	2.30	0.77	0.72	0.64
	TTT	23.2	26.1	17.1	0.88	0.78	1.19	0.45	0.40	0.61
G	GGA	9.0	10.9	16.4	3.77	3.11	2.07	3.49	2.88	1.92
	GGC	27.9	9.7	22.5	0.36	1.05	0.45	0.38	1.08	0.47
	GGG	11.3	6.0	16.3	1.50	2.82	1.04	1.39	2.62	0.96
	GGT	24.4	24.0	10.8	0.83	0.85	1.88	0.75	0.76	1.70
H	CAC	9.8	7.7	15.0	1.38	1.76	0.90	1.60	2.04	1.05
	CAT	13.6	13.7	10.5	0.00	0.00	0.00	1.35	1.34	1.75
I	ATA	5.4	17.8	7.7	1.88	0.57	1.32	2.42	0.74	1.70
	ATC	24.2	17.0	21.6	0.84	1.20	0.94	1.08	1.54	1.21
	ATT	29.8	30.4	16.1	0.57	0.56	1.05	1.05	1.03	1.95
K	AAA	33.2	42.2	24.1	1.02	0.80	1.41	0.87	0.68	1.19
	AAG	10.7	30.7	32.2	1.90	0.66	0.63	0.98	0.34	0.33
L	CTA	4.0	13.3	7.8	2.54	0.76	1.30	1.31	0.39	0.67
	CTC	11.0	5.4	19.8	0.62	1.26	0.34	0.95	1.94	0.53
	CTG	50.9	10.4	39.8	0.47	2.28	0.60	0.46	2.27	0.59
	CTT	11.7	12.1	13.0	1.74	1.68	1.56	0.67	0.65	0.60
	TTA	13.9	26.7	7.5	0.49	0.25	0.90	0.94	0.49	1.75
	TTG	14.0	27.0	12.6	1.21	0.63	1.35	0.75	0.39	0.83
M	ATG	27.0	20.9	22.2	0.13	0.16	0.15	0.29	0.38	0.35
N	AAC	21.4	24.9	19.5	0.48	0.41	0.52	2.08	1.79	2.28
	AAT	18.6	36.3	16.7	1.82	0.93	2.03	1.13	0.58	1.25
P	CCA	8.5	18.2	16.7	2.79	1.30	1.42	1.23	0.58	0.63
	CCC	5.8	6.8	20.1	2.34	1.99	0.67	4.06	3.46	1.17
	CCG	21.8	5.3	6.9	0.47	1.92	1.47	0.60	2.47	1.90
	CCT	7.3	13.6	17.3	0.93	0.50	0.39	0.72	0.39	0.30
Q	CAA	15.0	27.5	12.0	0.90	0.49	1.13	2.09	1.14	2.62
	CAG	29.5	12.1	34.1	0.57	1.40	0.50	0.53	1.30	0.46
R	AGA	2.9	21.3	11.5	7.01	0.95	1.77	5.42	0.74	1.37
	AGG	1.9	9.2	11.4	5.35	1.11	0.89	4.13	0.85	0.69
	CGA	3.9	3.0	6.3	2.61	3.39	1.61	2.68	3.49	1.66
	CGC	21.0	2.6	10.7	0.00	0.00	0.00	0.12	1.01	0.24
	CGG	6.3	1.7	11.6	1.61	5.98	0.88	0.00	0.00	0.00
	CGT	20.3	6.5	4.6	0.33	1.04	1.47	0.26	0.81	1.14
S	AGC	16.0	9.7	19.3	0.42	0.70	0.35	1.15	1.89	0.95
	AGT	9.5	14.2	11.9	1.78	1.19	1.42	1.93	1.29	1.54
	TCA	7.8	18.8	12.0	2.61	1.08	1.69	1.01	0.42	0.65
	TCC	8.9	14.2	17.6	1.52	0.95	0.77	1.47	0.92	0.74
	TCG	8.7	8.5	4.4	0.39	0.40	0.77	0.30	0.31	0.60
	TCT	8.7	23.5	14.7	0.78	0.29	0.46	1.20	0.45	0.71

续表 3

氨基酸	密码子	大肠杆菌	酵母	人类	GP1/ 大肠杆菌	GP1/ 酵母	GP1/ 人类	GP2/ 大肠杆菌	GP2/ 酵母	GP2/ 人类
		1/1000	1/1000	1/1000						
T	ACA	8.2	17.8	15.1	<u>4.96</u>	<u>2.29</u>	<u>2.69</u>	<u>5.43</u>	<u>2.50</u>	<u>2.95</u>
	ACC	22.8	12.6	19.4	0.59	1.08	0.70	1.38	<u>2.49</u>	1.62
	ACG	14.8	7.9	6.1	0.92	1.72	<u>2.22</u>	0.88	1.66	<u>2.15</u>
	ACT	9.1	20.3	13.0	1.86	0.83	1.30	<u>3.45</u>	1.55	<u>2.42</u>
V	GTA	11.1	11.8	7.2	0.61	0.57	0.94	0.71	0.67	1.09
	GTC	15.1	11.6	14.6	1.57	<u>2.05</u>	1.63	<u>0.35</u>	<u>0.45</u>	<u>0.36</u>
	GTG	25.5	10.6	28.4	0.66	1.60	0.60	<u>0.31</u>	0.74	<u>0.28</u>
	GTT	18.5	22.0	11.0	1.47	1.23	<u>2.47</u>	0.85	0.71	1.43
W	TGG	15.2	10.3	12.7	1.34	1.97	1.60	1.38	<u>2.03</u>	1.65
Y	TAC	12.1	14.6	15.5	1.12	0.93	0.87	0.65	0.54	0.51
	TAT	16.5	18.9	12.1	1.44	1.26	1.96	<u>0.32</u>	<u>0.28</u>	<u>0.43</u>
*	TAA	2.0	1.0	0.7	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>
	TAG	0.3	0.5	0.6	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>8.73</u>	<u>5.24</u>	<u>4.36</u>
	TGA	1.1	0.7	1.5	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>

下划线所示表示对该密码子的偏爱性较差

讨 论

埃博拉病毒的生物安全等级为 4 级,至今仍没有能够有效预防或治疗埃博拉病毒感染的疫苗或药物获准上市。这也为深入了解埃博拉病毒,从而研发预防、治疗及检测相关制品带来困难^[11]。因此,通过基因工程手段进行相关制品开发是理想的途径。不同种属间使用同义密码的频率有很大差异,所以外源基因在宿主系统中的表达往往有强弱之分。因此对埃博拉 GP 蛋白的密码子偏爱性进行分析,比较出较为接近的物种作为表达载体,并参考表达载体的密码子偏爱性进行目的基因的密码子优化,是 GP 蛋白基因工程表达的前期工作。

本研究选取了 GeneBank 中公布的 4 个 2014 年西非流行株埃博拉病毒全基因组序列,将编码 GP1 和 GP2 的基因序列进行比对,发现 4 株病毒的该基因序列密码子均相同,无差异碱基。这为 2014 西非流行株 GP 蛋白体外表达提供了便利。随后采用 EMBOSS 在线系统对从 2014 年西非埃博拉流行地区感染病人分离到的 4 个埃博拉病毒株包膜糖蛋白进行了密码子偏爱性分析。为确保结果的可靠性,我们同时采用了另一密码子偏爱性分析软件 CodonW 进行分析和结果比对,两种方法的分析结果一致,证明 EMBOSS 在线系统的分析数据具有可靠性,可进一步作为依据进行基因表达中的密码子优化。埃博拉病毒基因组、GP1 和 GP2 的 Nc 值都在 57 左右,而 CAI 值都在 0.7 左右,表明埃博拉病毒基因组及 GP 蛋白编码基因的表达水平相对较高,同时又存在大量稀有密码子的使用^[12]。在进行优化时应充分考虑二者的

平衡,结合宿主系统的密码子偏好,修改 GP 蛋白基因序列的密码子组成。

研究表明,基因密码子的使用与基因编码蛋白的结构和功能,以及基因表达的生理功能有着密切的联系。有研究发现密码子的使用与蛋白质的三级结构有很大的相关性,从而证明 DNA 的一级信息中含有蛋白质的三级结构相关信息^[12]。从密码子偏爱性分析结果中可以看出,埃博拉病毒的包膜糖蛋白的氨基酸密码子有较为明显的偏爱性,而且同样是包膜糖蛋白,GP1 和 GP2 基因的密码子偏爱性还是存在许多差异。例如天冬酰胺的偏爱密子在 GP1 中为 AAT,出现频率为 76.9%;在 GP2 中则为 AAC,出现频率为 68.0%。这与我们前期对 GP1 和 GP2 进行结构预测的结果有相关性。在下一步的工作中,将进一步明确对 GP1 和 GP2 三级结构有影响的氨基酸位点,在密码子优化中既满足宿主系统的密码子偏爱性,又保留部分影响蛋白结构的埃博拉原始密码子,以期获得表达量高且保持原始生物学功能的基因工程蛋白。

参考文献

1 WHO. Ebola Virus Disease Outbreak Response Plan in West Africa [R]. 2014, 07 - 12

2 Geisbert TW, Feldmann H. Recombinant vesicular stomatitis virus - based vaccines against Ebola and Marburg virus infections[J]. J Infect Dis, 2011, 204(3):1075 - 1081

3 Leroy E M, Becquart P, Wauquier N, et al. Evidence for Elbola virus superantigen activity[J]. J Virol, 2011, 85(8):4041 - 4042

4 Nanbo A, Imai M, Watanabe S, et al. Ebolavirus is internalized into host cells via macropinocytosis in a viral glycoprotein - dependent manner[J]. PLoS Pathog, 2010, 6(9):e1001121

5 丁国永,王志玉,高璐,等. 埃博拉病毒包膜糖蛋白研究进展[J].

- 病毒学报, 2013, 29(3): 233 - 237
- 6 王艳, 马文丽, 郑文岭. SARS 冠状病毒的密码子偏爱性分析[J]. 生命科学研究, 2003, 7(3): 219 - 223
- 7 Rice P, Longden I, Bleasby A. EMBOSS: The European molecular biology open software suite[J]. Trends in Genetics, 2000, 16(6): 276 - 277
- 8 Nakamura Y, Gojobori T, Lkmura T. Codon usage tabulate form the international DNA sequencedatabases: status for the year 2000 [J]. Nucleic Acids Research, 2000, 28(1): 292
- 9 November JA. Accounting for background nucleotide composition when measuring codon usage bias[J]. Mol Biol Evol, 2002, 19(8): 1390 - 1394
- 10 Wright F. The effective number of codons' used in a gene[J]. Gene, 1990, 87(1): 23 - 29
- 11 Geisbert TW, Feldmann H. Recombinant vesicular stomatitis virus based vaccines against Ebola and Marburg virus infections[J]. J Infect Dis, 2011, 204(3): 1075 - 1078
- 12 王世峰, 阮力. 病毒密码子使用频率研究进展[J]. 生物技术通讯, 2003, 14(1): 42 - 47
- (收稿日期: 2014 - 08 - 17)
- (修回日期: 2014 - 09 - 17)

重组人肠三叶因子对结肠癌 HT - 29 细胞移行能力的影响及其机制研究

李 腾 彭 曦

摘 要 **目的** 观察重组人肠三叶因子对 HT - 29 细胞移行能力的影响并探讨其作用机制。**方法** 采用重组表达的人肠三叶因子 (rhITF) 作为细胞刺激药物, 以传代培养的人结肠癌 HT - 29 细胞株为研究模型。用不同浓度 (10、25 和 50 $\mu\text{g/ml}$) rhITF 刺激 HT - 29 细胞, 采用 Transwell 法观察 HT - 29 细胞移行能力的变化; 用 50 $\mu\text{g/ml}$ rhITF 在不同时相点分别刺激 HT - 29 细胞, 分为 4、8、12 和 24h 组。通过 Western blot 法观察黏附蛋白 β - catenin、E - cadherin 和磷酸化 β - catenin 的蛋白表达变化。**结果** rhITF 促细胞移行能力随其浓度的增加而增强, 50 $\mu\text{g/ml}$ rhITF 组细胞数明显高于阴性对照组、10 $\mu\text{g/ml}$ rhITF 组和 25 $\mu\text{g/ml}$ rhITF 组, 差异有统计学意义 ($P < 0.01$); β - catenin 和 E - cadherin 的蛋白表达均受到 rhITF 抑制, 与其他组比较, 12h 组蛋白表达明显减弱 ($P < 0.05$); 12h 组磷酸化 β - catenin 的蛋白表达明显增强 ($P < 0.05$)。**结论** ITF 具有促进 HT - 29 细胞移行的能力, ITF 能促使 β - catenin 磷酸化并抑制 β - catenin 与 E - cadherin 的表达。

关键词 肠三叶因子 HT - 29 细胞 细胞移行 黏附分子 β - catenin E - cadherin

中图分类号 R735 **文献标识码** A **DOI** 10.3969/j.issn.1673-548X.2015.03.011

Effects of Recombinant Human Intestinal Trefoil Factor on Migration in HT - 29 Cells and its Mechanism. Li Teng, Peng Xi. Department of Nursing, Xian Medical University, Shaanxi 710021, China

Abstract **Objective** To observe the effects of recombinant human intestinal trefoil factor (rhITF) on HT - 29 cell migration, and explore its possible mechanism. **Method** Recombinant human intestinal trefoil factor (rhITF) was used as stimulation drugs, and subcultured human colon cancer cell (HT - 29) for research model. With different concentrations (10, 25 and 50 $\mu\text{g/ml}$) rhITF stimulate HT - 29 cells, the change of cell migration ability was observed by Transwell. At different time points with 50 $\mu\text{g/ml}$ rhITF were used to stimulate HT - 29 cells, and they were divided into 4h, 8h, 12h and 24h group. The change of β - catenin, E - cadherin and phosphorylated β - catenin were observed by Western blot. **Results** RhITF could promote HT - 29 migration. The ability of cell migration was along with the concentration increase. The number of cell in 50 $\mu\text{g/ml}$ rhITF group was more than negative control group and 10 $\mu\text{g/ml}$ rhITF group and 25 $\mu\text{g/ml}$ rhITF group, and the difference was statistically significant ($P < 0.01$). Protein expression of β - catenin and E - Cadherin were inhibited by rhITF. Compared with normal control group, the protein expression of 12h group decreased obviously ($P < 0.05$); the protein expression of phosphorylation β - catenin in 12h group increased significantly ($P < 0.05$). **Conclusion** ITF could promote HT - 29 cell migration. ITF could promote β - catenin phosphorylation, and inhibit protein expression of β - catenin and E - cadherin.

基金项目: 国家自然科学基金面上项目 (81372049)

作者单位: 710021 西安医学院护理学院 (李腾); 重庆, 中国人民解放军第三军医大学西南医院烧伤研究所 (彭曦)

通讯作者: 彭曦, 教授, 电子信箱, pxlrmm@163.com