

引用格式:

赵为民, 曹静, 邢菲, 王丽, 任守文, 付言峰, 李碧侠, 方晓敏. 猪蛋白编码基因 3'UTR 中 IRPRE1 元件的鉴定及其基因特征分析[J]. 湖南农业大学学报(自然科学版), 2020, 46(3): 336–342.

ZHAO W M, CAO J, XING F, WANG L, REN S W, FU Y F, LI B X, FANG X M. Identification and analysis of IRPRE1 elements of 3'UTR of protein-coding genes in porcine[J]. Journal of Hunan Agricultural University (Natural Sciences), 2020, 46(3): 336–342.

投稿网址: <http://xb.hunau.edu.cn>



## 猪蛋白编码基因 3'UTR 中 IRPRE1 元件的鉴定及其基因特征分析

赵为民<sup>1,2,3</sup>, 曹静<sup>4</sup>, 邢菲<sup>5</sup>, 王丽<sup>1,2,3</sup>, 任守文<sup>1,2,3</sup>, 付言峰<sup>1,2,3</sup>, 李碧侠<sup>1,2,3</sup>, 方晓敏<sup>1,2,3\*</sup>

(1.江苏省农业科学院畜牧研究所,江苏 南京 210014 2.江苏省农业种质资源保护与利用平台,江苏 南京 210014;  
3.江苏省农业科学院动物品种改良和繁育重点实验室,江苏 南京 210014; 4.江苏农科传媒有限公司,江苏 南京 210014;  
5.南京农业大学动物科技学院,江苏 南京 210095)

**摘要:** 为鉴定猪全基因组范围内蛋白编码基因 3'UTR(3'-untranslated region)中反向重复 PRE1(inverted repeated PRE1, IRPRE1)元件,对猪全基因组的 22 342 个蛋白编码基因的 3'UTR 序列进行重复序列元件的生物信息学分析。结果表明:猪蛋白编码基因的 3'UTR 序列中短散在重复序列与简单重复序列在重复序列的类别中所占比例较高,分别为 27.58%与 31.08%,在 SINE/tRNA 的重复元件中,Pre0\_SS 和 PRE1x 元件所占的比例较高,分别为 41.83%和 37.51%;共有 1 094 个候选蛋白编码基因的 3'UTR 中含有 IRPRE1 元件;GO 分析发现这些候选基因主要参与 mRNA 经由剪接体的剪接、细胞分裂、RNA 通过酯交换反应发生的剪接、T 细胞激活、RNA 剪接、T 细胞受体信号通路、对糖苷反应、甘油三酯稳态、外源性凋亡信号通路的正调节及胆固醇合成等生物过程;KEGG pathway 分析发现这些候选基因参与了缬草碱、亮氨酸和异亮氨酸降解途径、TNF 信号通路、T 细胞受体信号通路、RIG-I 样受体信号通路、吞噬体、剪接体、胆汁分泌、甲状腺激素合成和凋亡;对 3 个蛋白编码基因的 3'UTR 中的 IRPRE1 元件进行鉴定发现,其在多个组织中广泛表达。

**关键词:** 猪;蛋白编码基因;3'UTR;IRPRE1 元件;特征分析

中图分类号: S828.1 文献标志码: A 文章编号: 1007-1032(2020)03-0336-07

## Identification and analysis of IRPRE1 elements of 3'UTR of protein-coding genes in porcine

ZHAO Weimin<sup>1,2,3</sup>, CAO Jing<sup>4</sup>, XING Fei<sup>5</sup>, WANG Li<sup>1,2,3</sup>, REN Shouwen<sup>1,2,3</sup>,

FU Yanfeng<sup>1,2,3</sup>, LI Bixia<sup>1,2,3</sup>, FANG Xiaomin<sup>1,2,3\*</sup>

(1.Institute of Animal Science, Jiangsu Academy of Agricultural Sciences, Nanjing, Jiangsu 210014, China; 2.Jiangsu Germplasm Resources Protection and Utilization Platform, Nanjing, Jiangsu 210014, China; 3.Key Laboratory of Crop and Livestock Integration Ministry of Agriculture, Nanjing, Jiangsu 210014, China; 4.Jiangsu Agricultural Media Co. Ltd, Nanjing, Jiangsu 210014, China; 5.College of Animal Science and Technology, Nanjing Agricultural University, Nanjing, Jiangsu 210095, China)

**Abstract:** In order to identify the inverted repeating PRE1 element in the 3'-untranslated region(3'UTR) of protein-coding

收稿日期: 2019-08-14

修回日期: 2019-09-15

基金项目: 国家自然科学基金项目(31601928); 国家生猪现代产业技术体系南京综合试验站(CARS-35); 江苏省农业重大新品种创制项目(PZCZ201733)

作者简介: 赵为民(1983—), 男, 湖北钟祥人, 博士, 副研究员, 主要从事猪的抗病育种研究, zhaoweimin1983@aliyun.com; \*通信作者, 方晓敏, 副研究员, 主要从事猪的抗病育种与生产研究, fxmw2000@163.com

gene porcine at whole genome-wide, the bioinformatics analysis of the repeating sequence elements was performed on the 22 342 protein-coding genes of the pig genome. The results showed that the short interspersed nuclear element(SINE) and the simple\_repeat in the 3'UTR sequence of the pig protein-coding gene accounted for the highest proportion of the repeat sequence, which were 27.58% and 31.08%, respectively. Among the repetitive elements of SINE/tRNA, Pre0\_SS and PRE1x components accounted for the highest proportion, which were 41.83% and 37.51%, respectively. After layer analysis, a total of 1 094 candidate protein-coding genes contained an IRPRE1 element in the 3'UTR. Gene Ontology analysis indicated that these candidate genes were mainly involved in mRNA splicing, via spliceosome, cell division, RNA splicing, via transesterification reactions, T cell activation, RNA splicing, T cell receptor signaling pathway, response to glycoside, triglyceride homeostasis, positive regulation of extrinsic apoptotic signaling pathway and cholesterol biosynthetic process. KEGG pathways showed that these candidate genes were mainly involved in Valine, leucine and isoleucine degradation, TNF signaling pathway, T cell receptor signaling pathway, RIG-I-like receptor signaling pathway, Phagosome, Spliceosome, Bile secretion, Thyroid hormone synthesis and Apoptosis. In addition, the IRPRE1 element in the 3'UTR of the three protein-coding genes was identified and was found to be widely expressed in multiple tissues.

**Keywords:** porcine; protein-coding gene; 3'-untranslated region; Inverted repeated PRE1 element; feature analysis

蛋白编码基因的3'端非翻译区(untranslated region, UTR)是指从终止密码子到转录终止位点的这段区域。该区域虽然不编码蛋白质,但却含有丰富的调控元件,以调控本身 mRNA 的表达水平,从而丰富基因调控的多样性与复杂性。如蛋白编码基因 3'UTR 含有 microRNA 结合的种子序列,使 microRNA 通过调节基因 mRNA 的稳定性来调控各种生理活动<sup>[1-3]</sup>。

Alu 是短散在核重复序列(SINEs)中最丰富的 1 种重复元件,约占人基因组的 10%,广泛存在于灵长类生物基因的内含子、5'和 3'端非翻译区(UTR)<sup>[4-5]</sup>。研究表明,Alu 元件在调控基因的选择性剪接<sup>[5]</sup>、基因的转录与翻译<sup>[6-7]</sup>、RNA 的 A-to-I 编辑<sup>[5,8-10]</sup>中具有重要作用。此外,当蛋白编码基因的 3'UTR 存在反向重复 Alu(IRAlu)时,可导致该基因的 mRNA 滞留于细胞核,而不能运出到细胞质进行翻译,从而实现对该基因蛋白水平的调控,进而影响细胞的各种生理活动<sup>[11-13]</sup>。

Alu 重复元件虽然特异存在于灵长类动物<sup>[14]</sup>,然而在小鼠和猪中已发现与 Alu 元件相对应的 B1、B2、B4 元件和 PRE1 元件。有研究<sup>[15-16]</sup>表明,这些相对应元件的结构和功能非常类似于灵长类动物的 Alu。这说明存在于猪蛋白编码基因 3'UTR 中的 IRPRE1 元件可能与 IRAlu 具有类似调控基因的蛋白表达水平的功能。鉴于此,本研究中,对猪全基因组范围内蛋白编码基因 3'UTR 中的 IRPRE1 重复元件进行鉴定,并对其对应基因的特征进行分析,旨在为研究这些含有 IRPRE1 元件的蛋白编码基因的功能提供依据。

## 1 材料与方法

### 1.1 材料

DH5a 感受态菌株购于北京全式金生物技术有限公司。苏紫猪组织样品来源于江苏省农业科学院六合动物实验基地。采集 3 头健康成年母猪的心、肝、脾、肺、肾、小肠、背肌组织,投入液氮迅速冷冻,于-80 °C 保存,备用。

### 1.2 主要试剂

RNA 提取试剂盒(Takara MiniBEST Universal RNA Extraction Kit)、cDNA 反转录试剂盒(PrimeScript™ RT Master Mix)、DNA Marker 和 pMD19-T 购于 Takara; Golden Star T6 Super PCR Mix(1.1×)购于北京擎科新业生物技术有限公司; DNA 纯化回收试剂盒(DNA Clean & Concentrator)购于 Zymo Research。

### 1.3 猪全基因组范围内蛋白编码基因 3'UTR 中 IRPRE1 元件的分析

从 Ensemble 网站的 BioMart 中下载猪全基因组(Sscrofa11.1, Ensemble92)范围内的蛋白编码基因的 3'UTR 序列,整理成 Fasta 格式。由于 PRE1 元件属于 SINE/tRNA,使用 RepeatMasker 在线网站对猪蛋白编码基因 3'UTR 序列先进行重复序列分析,筛选条件中 search engine 为 cross\_match, DNA source 为 pig,挑选含有 SINE/tRNA 元件的蛋白编码基因;进一步选取 SINE/tRNA 中含有 Pre0\_SS 和 PRE1x (包括 PRE1a、PRE1b、PRE1c、PRE1d、PRE1d2、PRE1e、PRE1f、PRE1f2、PRE1g、PRE1h、PRE1i、PRE1j、PRE1k)

元件的蛋白编码基因;再选择 Pre0\_SS 和 PRE1x 长度 100 bp,且其中至少有 1 对 IR 的 Pre0\_SS 和 PRE1x 元件为最终候选蛋白编码基因。

#### 1.4 GO 与 KEGG 分析

采用 DAVID(Database for Annotation, Visualization and Integrated Discovery)对候选蛋白编码基因进行 GO(Gene Ontology)和 KEGG pathway (Kyoto Encyclopedia of Genes and Genomes)分析,选择人的对应基因作为背景,研究这些蛋白编码基因参与的生物学功能与途径。筛选条件选择 Count 为 2, EASE 为 0.01, fold enrichment 不小于 1.5。

#### 1.5 候选蛋白编码基因 IRPRE1 元件的鉴定与组织表达

参照 Takara MiniBEST Universal RNA Extraction Kit 说明书提取猪组织样品的总 RNA,并进行 DNaseI 处理。参照 PrimeScript™ II 1st Strand cDNA Synthesis Kit 说明书进行反转录,对每个组织的 1 μg 总 RNA 进行反转录,反转录完后用灭菌水稀释 5 倍,于 -20 °C 保存。按照 Golden Star T6 Super PCR Mix(1.1×)说明书进行 PCR, *HPRT1* 基因为内参基因,所有引物列于表 1。引物合成与测序验证均由南京擎科生物科技有限公司完成。对 3 个蛋白编码基因 *TRMO*、*RCSD1* 和 *DBT* 的 3'UTR 中的 IRPRE1 元件进行组织表达谱分析。由于选取的 IRPRE1 元件位于 3'UTR 且都是位于单个外显子上,为了避免 RNA 样品中的 DNA 污染,利用基因组上的引物 (*Genome*)进行 PCR 检测,以检测 RNA 样品中是否含有 DNA。

表 1 引物序列信息

引物名称	引物序列(5'—3')	退火温度/°C	扩增长度/bp
<i>TRMO</i>	TGGCTTGGATCCTGTGTGAC	60	292
	AGGATCTGGTGTGCCACAG		
<i>RCSD1</i>	GGTACAGAGTCATCGCCAGG	60	531
	TACCCATTGAGAGAGGCCA		
<i>DBT</i>	AGGAACCATGAGACTGCGTG	60	562
	AGGTTGAAGACGCAGCTCAA		
<i>Genome</i>	TCCAACCTTGACGGACACTG	60	165
	TGCAGCTCTCAACTACCTGC		
<i>HPRT1</i>	CCCAGCGTCGTGATTAGTGA	60	191
	TTGAGCACACAGAGGGCTAC		

## 2 结果与分析

### 2.1 猪全基因组范围内蛋白编码基因 3'UTR 序列的重复元件

对 Ensemble 网站中猪的全基因组范围内蛋白编码基因分析发现,一共有 22 342 个蛋白编码基因,对应 45 788 个转录本。对 45 788 个转录本的 3'UTR 序列进行重复序列分析发现,其含有重复序列中的 2 大类型,即分散重复序列与串联重复序列。分散重复序列中的 DNA 转座子(DNA transposon)、长末端重复序列(LTR)、长散在重复序列(LINE)和短散在重复序列(SINE)分别占总数的 7.31%、5.36%、14.87%和 27.58%;串联重复序列中的简单重复序列(simple\_repeat)和低复杂度序列(low\_complexity)分别占总数的 31.08%和 4.09%;其他种类重复序列占总数的 9.71%。进一步发现 SINE 与简单重复序列种类的重复序列所占比例较高,其次为 LINE,其余的都低于 10%。

### 2.2 猪全基因组范围内蛋白编码基因 3'UTR 中 SINE/tRNA 的种类

对猪的全基因组范围内蛋白编码基因 3'UTR 中仅含有 SINE/tRNA 的序列分析发现,其亚种类繁多,每个种类都占有一定的比例,而其中 Pre0\_SS 和 PRE1x 所占的比例分别为 41.83%和 37.51%(表 2),这 2 类重复元件所占比例较高。

表 2 猪全基因组范围内蛋白编码基因 3'UTR 中 SINE/tRNA 的种类占比

Table 2 Analysis of SINE/tRNA type within 3'UTR sequences of genome-wide of protein-coding genes in porcine genome

种类	占比/%	种类	占比/%
CHR-1	3.12	PRE1g	3.78
LFSINE_Vert	0.09	PRE1h	1.54
Pre0_SS	41.83	PRE1i	3.75
PRE1_SS	0.13	PRE1j	3.43
PRE1a	0.48	PRE1k	2.11
PRE1b	1.66	SINE1_SS	2.03
PRE1c	1.65	SINE1A_SS	3.24
PRE1d	2.67	SINE1B_SS	4.04
PRE1d2	3.16	SINE1C_SS	3.67
PRE1e	3.24	SINE1D_SS	3.24
PRE1f	4.56	SINE2-1_SSc	0.95
PRE1f2	5.48	SUSINE2	0.15

### 2.3 猪全基因组范围内蛋白编码基因 3'UTR 含有 IRPRE1 元件的鉴定结果

先去除猪全基因组范围内的 22 342 个蛋白编码基因 3'UTR 中不含有 SINE/tRNA 重复元件的基因,留下 5 017 个蛋白编码基因;再去掉 SINE/tRNA 重复元件中不含 Pre0\_SS 与 PRE1x 元件的基因,留下 4 486 个蛋白编码基因;然后去除 Pre0\_SS 与 PRE1x 元件全部为同一方向、长度小于 100 bp 的 Pre0\_SS 和 PRE1x 的基因,最终得到 1 094 个候选蛋白编码基因,对应 1 636 个转录本,进一步分析发现,这 1 636 个转录本中 PRE1 家族元件的长度为 10~323 bp。如表 3 所示,1 094 个蛋白编码基因在猪不同染色体上的分布不同,其中 Y 染色体上的最少,只有 3 个。1~4 号和 6 号染色体上的较多,其中 6 号染色体的最多,达 117 个。

表 3 候选蛋白编码基因在染色体上的分布

Table 3 The number distribution of candidate protein-coding genes on chromosomes

染色体	基因数	染色体	基因数
Chr1	82	Chr11	21
Chr2	91	Chr12	60
Chr3	89	Chr13	85
Chr4	85	Chr14	55
Chr5	56	Chr15	42
Chr6	117	Chr16	17
Chr7	68	Chr17	43
Chr8	34	Chr18	33
Chr9	47	ChrX	36
Chr10	30	ChrY	3

### 2.4 候选蛋白编码基因的 GO 分析结果

候选蛋白编码基因一共参与 38 个生物学过程,22 个细胞组分和 15 个分子功能( $P < 0.05$ )。以  $P$  值显著性程度来衡量,生物过程、细胞组分和分子功能中排前十的注释列于图 1。

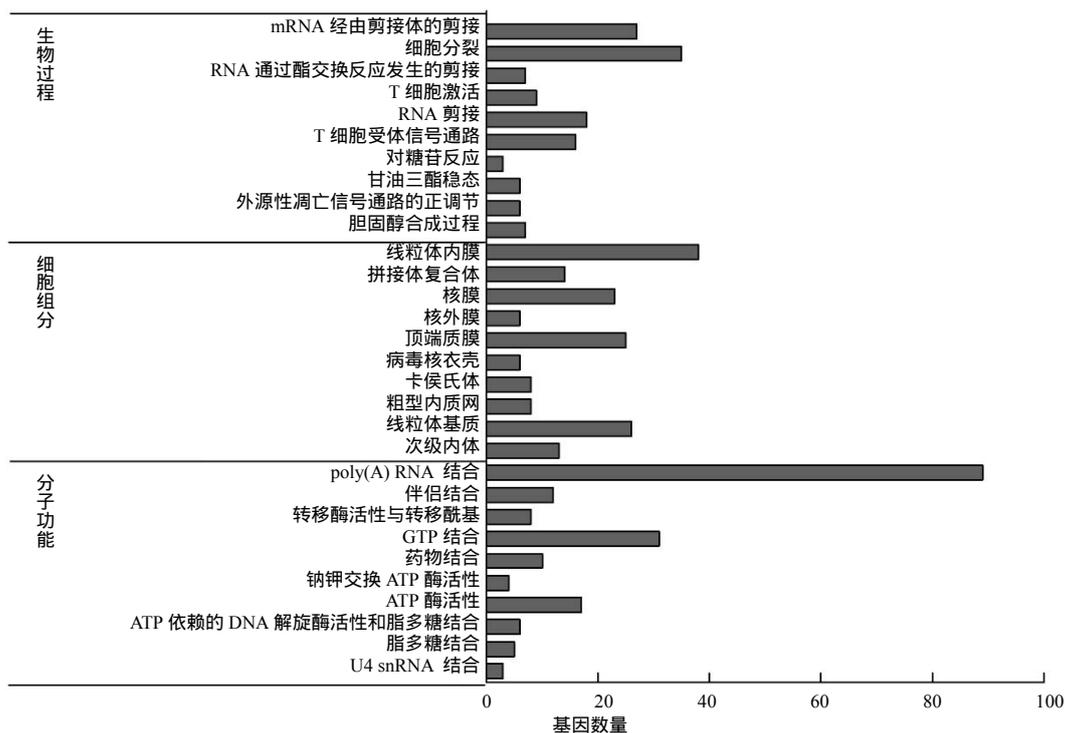


图 1 候选蛋白编码基因的 GO 注释

Fig.1 GO analysis of candidate protein-coding genes

### 2.5 候选蛋白编码基因的 KEGG pathway 分析结果

如表 4 所示,以  $P$  值显著性程度来衡量,候选

蛋白编码基因一共参与 9 个生物学通路( $P < 0.05$ )。

表 4 KEGG 通路中基因相关信息

**Table 4 Gene information involved in KEGG pathway**

KEGG term ID	通路	基因数	基因	P
hsa00280	缬草碱、亮氨酸和异亮氨酸降解途径	10	<i>ACAA2 ,DBT ,ACADSB ,HMGCS2 ,JVD ,MCEE ,OXCT1 , AOX1 , HIBCH , ACAT2</i>	0.000 804
hsa04668	TNF 信号通路	10	<i>IL18R1 , CXCL2 , CASP10 , PGAM5 , RIPK1 , CASP8 , CREB3L2 , CREB3L4 , CHUK , AKT2</i>	0.004 620
hsa04660	T 细胞受体信号通路	13	<i>PTPN6 , CD3G , CD8A , CD3D , PRKCQ , CBLC , PAK2 , MAP3K8 , LCK , PPP3CC , GRAP2 , CHUK , AKT2</i>	0.009 556
hsa04622	RIG-I 样受体信号通路	10	<i>MAVS , RNF125 , CASP10 , TMEM173 , SIKE1 , RIPK1 , CASP8 , TBKBP1 , AZI2 , CHUK</i>	0.012 802
hsa04145	吞噬体	14	<i>RAB7B , MSR1 , TLR2 , ITGB5 , C1R , ATP6V1G1 , CANX , MARCO , LAMP2 , TUBA8 , CORO1A , ATP6V1E1 , TUBAL3 , ATP6V1G3</i>	0.018 573
hsa03040	剪接体	14	<i>PRPF31 , DDX39B , CWC15 , SNRPD1 , PRPF3 , HNRNPA1 , PRPF4 , SF3A3 , PRPF19 , SRSF4 , DDX46 , SRSF9 , DHX16 , HSPA8</i>	0.027 904
hsa04976	胆汁分泌	9	<i>ABCG8 , SULT2A1 , ADCY8 , AQP8 , SLC22A8 , ATP1A3 , ATP1A1 , ATP1A2 , ABCG2</i>	0.032 295
hsa04918	甲状腺激素合成	9	<i>SLC26A4 , ADCY8 , CREB3L2 , ATP1A3 , ATP1A1 , CREB3L4 , ATP1A2 , LRP2 , CANX</i>	0.034 810
hsa04210	凋亡	7	<i>CASP10 , TNFSF10 , RIPK1 , CASP8 , CSF2RB , CHUK , AKT2</i>	0.049 468

2.6 候选蛋白编码基因 IRPRE1 元件的鉴定与组织表达结果

图 2 为 3 个蛋白编码基因 *TRMO*、*RCSD1* 和 *DBT* 的 3'UTR 中 IRPRE1 元件方向示意图。从图 3 可知，阴性对照与 DNaseI 消化后的各个组织 RNA 均检测不到 *Genome* 扩增产物，而作为阳性对照的基因组 DNA 中能检测到 *Genome* 扩增产物，说明提取的组织 RNA 中消除了 DNA 的污染。从图 4 可

知，*TRMO*、*RCSD1* 和 *DBT* 中 IRPRE1 元件在各个组织中都有不同程度的表达；*TRMO* 基因的 IRPRE1 元件在心、肝、脾、肺、肾和肌肉中的表达较高，小肠中表达较弱；*RCSD1* 基因的 IRPRE1 元件在肌肉中表达较高，其他组织表达较弱，小肠中几乎检测不到；*DBT* 基因的 IRPRE1 元件在肌肉中表达较高，其他组织表达较弱，肺中几乎检测不到。

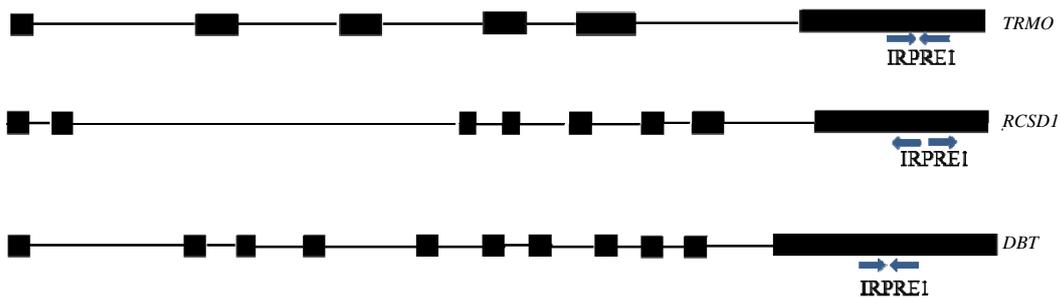
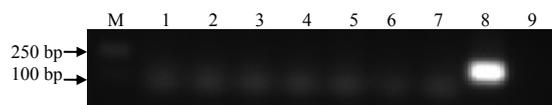


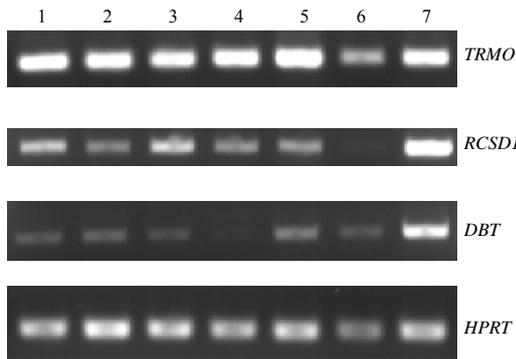
图 2 *TRMO* 和 *RCSD1* 及 *DBT* 基因的 3'UTR 中 IRPRE1 元件  
Fig.2 A diagram of IRPRE1 elements of 3'UTR of *TRMO*,*RCSD1* and *DBT* gene



M DL 2000 Marker ; 1~8 心脏、肝脏、脾脏、肺脏、肾脏、小肠、背肌、猪基因组 DNA ; 9 阴性对照。

图 3 RNA 样品的 DNA 污染检测结果

Fig.3 DNA contamination detection for the RNA sample



泳道 1~7 分别示心脏、肝脏、脾脏、肺脏、肾脏、小肠、背肌中 IRPRE1 元件的表达。

图 4 候选蛋白编码基因 IRPRE1 元件的组织表达谱

Fig.4 Tissue expression profile of IRPRE1 element of candidate protein coding genes

### 3 结论与讨论

Ensemble 数据库对每个蛋白编码基因有完善的注释信息,包括转录起始和终止位点、5'UTR 和 3'UTR 的序列信息、cDNA 序列、CDS 序列等等,而且每隔一段时间就会发布一个新版本,非常有利于查询,很多研究对基因的注释信息都采用 Ensemble<sup>[17-18]</sup>。NCBI 数据库对蛋白编码基因的结构注释信息,尤其是 UTR 序列信息没有 Ensemble 数据库完善,且 NCBI 数据库与 Ensemble 数据库对基因的转录本数量及 UTR 的序列注释差异也较大,目前对这种差异性还不能有效地验证哪一个数据库注释的正确。在对 Ensemble 数据库中的蛋白编码基因 3'UTR 分析的同时,也查阅了相应 NCBI 数据库中的信息,发现存在很多差异,如在选取的 3 个基因 *TRMO*、*RCSD1*、*DBT* 中,*RCSD1* 和 *DBT* 基因的 3'UTR 序列在 Ensemble 和 NCBI 数据库中一致,但 *TRMO* 基因的 3'UTR 序列的 IRPRE1 元件序列在 NCBI 的注释中没有,在 Ensemble 的注释中有,且试验也验证了 *TRMO* 基因 3'UTR 序列的 IRPRE1 元件的表达;因此,在本研究中采用了 Ensemble 数据库。

重复序列元件广泛存在于哺乳动物基因组中,人的基因组中有将近一半的序列为重复元件<sup>[4]</sup>。LINE 和 SINE 作为重复元件的主要类型,在基因的多样性中发挥重要作用<sup>[14,19-20]</sup>。在猪的基因组序列中,LINE 和 SINE 分别占 17.5%和 11.4%<sup>[16]</sup>。本研究中,LINE 和 SINE 分别占 3'UTR 总重复元件的

14.87%和 27.58%,且 SINE 的占比仅仅稍低于简单重复序列(31.08%),表明 SINE 相对于 LINE 在 3'UTR 序列可能发挥更加重要的调控作用。研究<sup>[16]</sup>表明,Pre0\_SS 是非常类似于 PRE1 的元件,故本研究中把 Pre0\_SS 元件也归纳到 PRE1 元件家族中。本研究中,SINE/tRNA 中,Pre0\_SS 和 PRE1x 所占的比例分别为 41.83%和 37.51%,是 SNINE/tRNA 中最丰富的元件,这与对应人物种中 Alu 元件占 SINE 比例最高是一致的<sup>[4]</sup>。

Alu 元件的长度约为 300 bp<sup>[21]</sup>,而 PRE1 家族元件的长度约为 250 bp<sup>[16]</sup>。本文研究中,PRE1 家族元件的长度为 10~323 bp,说明有些蛋白编码基因 3'UTR 中的 PRE1 元件的长度并不完整,有些只是 PRE1 的部分序列。有研究<sup>[22-23]</sup>表明,IRAlu 中的 2 个反向的 Alu 元件长度大于 100 bp 时才能形成有效的双链 RNA,从而发挥生物学功能,故本研究中对 IRPRE1 元件的分析中也选取了蛋白编码基因 3'UTR 中长度大于 100 bp 的 PRE1。本研究中,IRPRE1 的 2 个相反的 PRE1 元件之间的间隔序列长度从几十到上千个碱基,目前尚不清楚这种长度大小对 IRPRE1 发挥的功能有何具体影响,且这种间隔长度大小变化也见于 IRAlu 中<sup>[12]</sup>,需要后续进一步研究。

KEGG pathway 分析发现,有 1 094 个蛋白编码基因参与多种生物学通路,其中包括 TNF 信号通路及 RIG-I 样受体信号等。TNF 与 RIG-I 信号通路是响应细胞内外信号的通路,可影响多种细胞的生理活动<sup>[24-25]</sup>,其相关基因的表达水平在上述通路中发挥重要功能,表明这些候选基因的 IRPRE1 元件可能在调控其基因的表达水平中发挥着作用,进而影响基因本身在这些通路中的调控作用。

#### 参考文献:

- [1] BARTEL D P. MicroRNAs: genomics, biogenesis, mechanism, and function[J]. Cell, 2004, 116(2): 281-297.
- [2] AMBROS V. The functions of animal microRNAs[J]. Nature, 2004, 431: 350-355.
- [3] GU Y, BECKER V, ZHAO Y, et al. MiR-370 inhibits the angiogenic activity of endothelial cells by targeting smoothened(SMO) and bone morphogenetic protein (BMP)-2[J]. The FASEB Journal, 2019, 33(6): 7213-7224.

- [4] LANDER E S , LINTON L M , BIRREN B , et al . Initial sequencing and analysis of the human genome[J]. *Nature* , 2001 , 409 : 860–921 .
- [5] ATHANASIADIS A , RICH A , MAAS S . Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome[J] . *PLoS Biology* , 2004 , 2(12) : e391 .
- [6] LANDRY J R , MEDSTRAND P , MAGER D L . Repetitive elements in the 5' untranslated region of a human zinc-finger gene modulate transcription and translation efficiency[J] . *Genomics* , 2001 , 76(1/2/3) : 110–116 .
- [7] HÄSLER J , STRUB K . Alu RNP and Alu RNA regulate translation initiation in vitro[J] . *Nucleic Acids Research* , 2006 , 34(8) : 2374–2385 .
- [8] LEVANON E Y , EISENBERG E , YELIN R , et al . Systematic identification of abundant A-to-I editing sites in the human transcriptome[J] . *Nature Biotechnology* , 2004 , 22(8) : 1001–1005 .
- [9] BAZAK L , HAVIV A , BARAK M , et al . A-to-I RNA editing occurs at over a hundred million genomic sites , located in a majority of human genes[J] . *Genome Research* , 2014 , 24(3) : 365–376 .
- [10] LIN L , JIANG P , PARK J W , et al . The contribution of Alu exons to the human proteome[J] . *Genome Biology* , 2016 , 17 : 15 .
- [11] PRASANTH K V , PRASANTH S G , XUAN Z , et al . Regulating gene expression through RNA nuclear retention[J] . *Cell* , 2005 , 123(2) : 249–263 .
- [12] CHEN L L , DECERBO J N , CARMICHAEL G G . Alu element-mediated gene silencing[J] . *The EMBO Journal* , 2008 , 27(12) : 1694–1705 .
- [13] TORRES M , BECQUET D , BLANCHARD M P , et al . Circadian RNA expression elicited by 3'-UTR IRAlu-paraspeckle associated elements[J] . *eLife* , 2016 , 5 : e14837 .
- [14] BATZER M A , DEININGER P L . Alu repeats and human genomic diversity[J] . *Nature Reviews Genetics* , 2002 , 3(5) : 370–379 .
- [15] WANG J , GONG C , MAQUAT L E . Control of myogenesis by rodent SINE-containing lncRNAs[J]. *Genes & Development* , 2013 , 27(7) : 793–804 .
- [16] FUNKHOUSER S A , STEIBEL J P , BATES R O , et al . Evidence for transcriptome-wide RNA editing among *Sus scrofa* PRE-1 SINE elements[J] . *BMC Genomics* , 2017 , 18(1) : 360 .
- [17] CARPENTER S , AIELLO D , ATIANAND M K , et al . A long noncoding RNA mediates both activation and repression of immune response genes[J] . *Science* , 2013 , 341 : 789–792 .
- [18] LI B J , DONG C , LI P H , et al . Identification of candidate genes associated with porcine meat color traits by genome-wide transcriptome analysis[J] . *Scientific reports* , 2016 , 6 : 35224 .
- [19] KREHLING J , GRAVELEY B R . The origins and implications of Alu alternative splicing[J] . *Trends in Genetics* , 2004 , 20(1) : 1–4 .
- [20] MÖLLER-KRULL M , ZEMANN A , ROOS C , et al . Beyond DNA : RNA editing and steps toward Alu exonization in Primates[J] . *Journal of Molecular Biology* , 2008 , 382(3) : 601–609 .
- [21] CHEN L L , CARMICHAEL G G . Gene regulation by SINES and inosines : biological consequences of A-to-I editing of Alu element inverted repeats[J] . *Cell Cycle* , 2008 , 7(21) : 3294–3301 .
- [22] BASS B L . RNA editing by adenosine deaminases that act on RNA[J] . *Annual Review of Biochemistry* , 2002 , 71 : 817–846 .
- [23] BLOW M , FUTREAL P A , WOOSTER R , et al . A survey of RNA editing in human brain[J] . *Genome Research* , 2004 , 14(12) : 2379–2387 .
- [24] LIU Z G . Molecular mechanism of TNF signaling and beyond[J] . *Cell Research* , 2005 , 15(1) : 24–27 .
- [25] QUICKE K M , DIAMOND M S , SUTHAR M S . Negative regulators of the RIG-I-like receptor signaling pathway[J] . *European Journal of Immunology* , 2017 , 47(4) : 615–628 .

责任编辑: 邹慧玲  
英文编辑: 柳正