

牦牛 *CART* 基因克隆、单核苷酸多态性检测及生物信息学分析

刘 建^{1,2}, 刘文博^{1,2}, 吴晓云^{1,2}, 阎 萍^{1,2*}, 梁春年^{1,2}

(1. 中国农业科学院兰州畜牧与兽药研究所, 兰州 730050; 2. 甘肃省牦牛繁育工程重点实验室, 兰州 730050)

摘要: 旨在克隆牦牛可卡因-苯丙胺调节转录肽基因(Cocaine- and amphetamine-regulated transcript, *CART*), 获得基因全长序列并分析基因结构和相关遗传变异, 检测单核苷酸多态性(SNP)位点在不同品种牦牛群体中的分布, 通过生物信息学相关方法分析预测 SNP 位点对 *CART* 基因表达的影响, 为牦牛 *CART* 基因功能研究奠定基础。以 243 头大通牦牛、208 头天祝牦牛、208 头甘南牦牛和 53 头帕里牦牛血样为材料, 提取基因组 DNA, 克隆出牦牛 *CART* 基因全长序列; 运用高分辨率熔解曲线分析技术(High-resolution melting, HRM)检测 *CART* 基因 4 个 SNPs 在 4 个牦牛品种中的分布, 进行遗传统计分析, 并应用生物信息学分析 SNPs 可能对基因功能产生的影响。本研究发现, 牦牛 *CART* 基因 4 个 SNPs, 1 个位于基因上游序列、1 个位于内含子、2 个位于 3'-UTR。χ² 检验表明, 除大通牦牛 V-221 位点和甘南牦牛 V1816 位点外, 其他所有群体各个位点都处于哈代-温伯格平衡($P > 0.05$)。群体遗传学分析显示, V168 和 V1816 位点属于中度多态, V-221 和 V1791 属于低度多态, 单倍型 C 属于优势单倍型。生物信息分析发现, V1791 和 V1816 位点碱基变化导致 RNA 二级结构发生变化, 可能会对 *CART* 基因的表达产生影响。

关键词: 牦牛; *CART* 基因; 克隆; 多态性; RNA 二级结构

中图分类号: S813.3

文献标志码: A

文章编号: 0366-6964(2013)08-1219-10

Cloning, Single Nucleotide Polymorphism and Bioinformatics Analysis of *CART* Gene in Yak

LIU Jian^{1,2}, LIU Wen-bo^{1,2}, WU Xiao-yun^{1,2}, YAN Ping^{1,2*}, LIANG Chun-nian^{1,2}

(1. Lanzhou Institute of Husbandry and Pharmaceutical Sciences of Chinese Academy of Agricultural Sciences, Lanzhou 730050, China; 2. Key Laboratory of Yak Breeding Engineering, Lanzhou 730050, China)

Abstract: This study aimed to clone the complete length sequence of Cocaine- and Amphetamine-Regulated Transcript gene (*CART*), analyze its structure, and reveal the distribution of SNPs (single nucleotide polymorphisms) in different yak breeds. Possible impact of the SNPs on *CART* gene expression was also predicted by bioinformatics analysis. This study could provide some basic information for studying the function of yak *CART* gene in future. The blood samples of 243 Datong yak, 208 Tianzhu yak, 208 Gannan yak, and 53 Pali yak were used to isolate the genomic DNA. Then the full length of *CART* gene was cloned and sequenced. Genetic variations and the distribution of 4 SNPs in the 4 yak breeds were detected by the high-resolution melting system followed by genetic statistical analysis. In this study, 4 novel SNPs located in the yak *CART* gene were identified, and these SNPs were located in upstream, intron, and 3'-UTR of the *CART* gene. χ² tests

收稿日期: 2013-01-22

基金项目: 现代农业(肉牛牦牛)产业技术体系专项资金(CARS-38); 甘肃省科技重大专项计划(1102NKDA027)

作者简介: 刘 建(1987-), 男, 四川宣汉人, 硕士生, 主要从事动物遗传育种与繁殖研究, E-mail: afenxiaoliu@yahoo.com.cn

* 通信作者: 阎 萍, 研究员, 博士生导师, 主要从事动物遗传育种研究, E-mail: pingyan@sohu.com

showed that besides the V-221 in Datong population and V1816 in Gannan population, all the other loci were in Hardy-Weinberg equilibrium ($P > 0.05$) in the 4 yak breeds. Through population genetics analysis, the results showed that V168 and V181 loci were at intermediate polymorphic status, but V-221 and V1791 loci were at low polymorphic status. The haplotype C was predominant. Bioinformatics analysis showed that different alleles of V1791 and V1816 might change the RNA secondary structure, which could influence *CART* gene's expression. The nucleotide variation of V1791 and V1816 loci reduced RNA stability that might produce negative influence on *CART* gene expression.

Key words: yak; *CART* gene; clone; polymorphism; RNA secondary structure

可卡因-苯丙胺调节转录肽(Cocaine- and amphetamine-regulated transcript, CART)是单胎物种在 1 次发情周期中限量排卵的关键因子,尤其是对优势卵泡的选择具有潜在的调控作用^[1-4]。目前已对人、鼠、牛和绵羊的 *CART* 基因有了较为深入的研究,而牦牛这一基因的特性尚不清楚。因此,克隆牦牛 *CART* 基因全长,研究该基因结构,分析其遗传变异,将为探究牦牛 *CART* 基因的功能提供理论基础。*CART* 是一种在体内广泛分布的神经肽类物质,最早由 J. Spiess 等^[5]从绵羊下丘脑提取物中分离出来的,是一种类生长抑制素多肽。研究发现, *CART* 在动物机体内参与摄食^[6]及体内能量代谢调节^[7]、机体应激反应^[8]、自主神经和感觉传导^[9-10]、抗精神兴奋^[11]等多种生理代谢过程。E. Valle 等^[12]利用人 *CART* 同源基因的参考序列以及牛上该基因的表达序列标签,设计引物对牛 *CART* 进行了测序和定位,发现该基因位于牛 20 号染色体上,获得了 1 769 bp 包含全长编码区的 DNA 序列,其中有 3 个外显子和 2 个内含子。通过对瘤牛、美洲野牛和普通牛 *CART* 基因进行多态性分析,发现 SNP 位点主要集中在基因的内含子 1^[12]和启动子区^[13]。K. Yamada 等^[14]克隆了人 *CART* 基因的启动子序列,探讨了该区域多态性与肥胖的相关性,认为在转录起点上游-156 bp 处由 A 转换为 G 可能导致肥胖症发生几率增加^[15]。除了卵泡发育外,还有研究^[13]对牛 *CART* 基因 4 个 SNPs 位点基因型与生长性状进行关联分析,发现 1 个多态性位点(C1)A1A1 基因型个体的体重显著性高于其他基因型个体($P < 0.05$),而另 1 个多态性位点(C2)A2A2 基因型个体的体重显著性低于其他基因型个体($P < 0.05$)。

本研究参考牛的 *CART* 基因序列,克隆牦牛的 *CART* 基因,分析甘南牦牛、天祝白牦牛、帕里牦牛

和大通牦牛 *CART* 基因的多态性位点,并进一步对其进行了生物信息学分析,从而阐明牦牛 *CART* 基因的特性。

1 材料与方法

1.1 材料收集

试验材料为青海大通牦牛血样 243 份、甘肃天祝牦牛血样 208 份、甘南牦牛血样 208 份和西藏帕里牦牛血样 53 份。

1.2 基因组 DNA 提取

利用 RelaxGene 血液基因组 DNA 提取系统(天根,北京)提取基因组 DNA,采用 1% 的琼脂糖凝胶电泳检测 DNA 的质量,使用 NANO DROP 2000(Thermo,美国)对提取 DNA 浓度进行测量并稀释到 $18 \sim 22 \text{ ng} \cdot \mu\text{L}^{-1}$ 作为 PCR 扩增模板。

1.3 基因组 DNA 池构建

分别取 4 种牦牛稀释 DNA 样各 20 份,每个样品各取 $1 \mu\text{L}$,然后混合均匀,构成 4 个独立牦牛基因组 DNA 池。

1.4 引物设计和 PCR 扩增

参考 NCBI 数据库 GenBank 序列(登录号:NT_186496),采用交叉重叠原理^[16]设计 3 对引物(表 1)获得牦牛 *CART* 基因的全长序列。由于 *CART* 基因序列中存在 2 段 Poly(T),*CART*-2 引物不能进行 PCR 扩增产物直接测序,所以分别从 4 种牦牛 DNA 样中各选 1 个,以 *CART*-2 为引物扩增,然后将扩增产物送往上海生工生物工程技术有限公司进行克隆测序。

PCR 扩增体系($50 \mu\text{L}$):包括约 $20 \text{ ng} \cdot \mu\text{L}^{-1}$ 的基因组 DNA $3 \mu\text{L}$, 10 pmol 的上下游引物各 $1 \mu\text{L}$, $2 \times \text{Taq PCR MasterMix}$ (内含 *Taq* DNA 聚合酶、 Mg^{2+} 、dNTPs 等)(北京,天根) $25 \mu\text{L}$,灭菌超纯水 $20 \mu\text{L}$ 。

表 1 扩增牦牛 *CART* 基因全长序列所用引物信息表Table 1 Information of primer sequences for amplifying complete yak *CART* gene

引物 Primer	片段长度/bp Size of products	退火温度/℃ T _m	引物序列(5'-3') Sequence	扩增片段在 GenBank 中位置 (GenBank: NT_186496)
CART-1	1 432	52.8	F: TTTTATGGACTGTATTTGC R: GCCAATCTATTCTTAACTT	(nt 59 766~61 197) 包括部分启动子、5'UTR、外显子 1、内含子 1、外显子 2 及部分内含子 2
CART-2	614	54.3	F: GTGGCTATTCATTCAGATAA R: TTGGGCTTGTTATTTGAG	(nt 61 092~61 705) 包括部分内含子 2
CART-3	1 229	57.4	F: GAAAAGTTTGGTTGACT R: TAATGTGACTAAAGGCTGA	(nt 61 564~62 792) 包括外显子 3、3'-UTR、及部分基因下游序列

PCR 反应程序:95 ℃ 预变性 5 min;94 ℃ 变性 50 s,T_m(表 1)退火 40 s,72 ℃ 延伸 1 min,35 个循环;72 ℃ 延伸 5 min;15 ℃ 保存。

1.5 DNA 测序分析

分别对 CART-1 和 CART-3 引物进行 PCR 扩增,将扩增产物通过琼脂糖凝胶回收试剂盒(天根,北京)回收纯化送往大连宝生物工程有限公司进行测序,测序结果通过 MEGA5.0 软件对比分析。

1.6 生物信息学分析

通过使用 NCBI 中 Spidey^[17] 分析扩增出的牦牛 *CART* 基因序列,并使用 MEGA5.0^[18] 将测序结

果与牛 *CART* 基因对比分析,找出其差异位点。通过生物信息学软件(RNAfold 2.0^[19]、Match-1.0 Public)分析、预测 SNP 位点是否影响基因的功能。

1.7 HRM 小片段法基因分型^[20-21]

针对所发现的 SNP 位点,利用 LightScanner Primer Design Software(Idaho 公司,美国)设计用于 HRM 检测 SNP 位点基因型的引物(表 2),对所有试验群体样本进行 SNP 位点的基因型分析。合成 2 对高低温内标^[22](表 3)对熔解曲线进行校正,提高分型的精确度。

表 2 牦牛 *CART* 基因 HRM 分析 SNPs 的引物信息表Table 2 Information of primer sequences for genotyping mutations by HRM within yak *CART* gene

引物 Primer	检测位点/ bp Locus	片段长度/bp Size of products	退火温度/℃ T _m	引物序列(5'-3') Primer sequence
CARTPT-1	-221	55	59.3	F: GGGGCTGACAGCACTTG R: CCGGAACGAGGTTCTGC
CARTPT-2	168	84	59.3	F: GGATGCCTCCCATGAGAAG R: GCGTACAAAGCAAGGAATATCG
CARTPT-3	1 791	61	57.8	F: TGCTTCCTGCTGTTTCCAAATA R: CACACCATAACAAGGCATTACC
CARTPT-4	1 816	94	64.3	F: GTCTTCCTGCTGTTTCCAAATAA R: GAACACTTTGTACACAGAGCCA

表 3 高低温内标序列信息表

Table 3 Information of low and high calibrator sequence

高低温内标 Low and high calibrator	T _m 值/℃ T _m	内标序列(5'-3') Calibrator
Low1	62	TTAAATTATAAAATATTTATAATATTAATTATATATATATAAAATAT AATAC-C3-3'
High1	92	GCGCGGCCGGCACTGACCCGAGACTCTGAGCGGCTGCTGGAGGTGCGGAAG CGGAGGGCGGGC-C3-3'
Low2	69.85	GCGGTCAGTCGGCCTAGCGGTAGCCAGCTGCGGCACTGCGTGACGCTCAG-C3-3' ATCGTGATTTCTATAGTTATCTAAGTAGTTGGCATTAAATAATTT
High2	87.92	CATTTT-C3-3'

HRM-PCR 扩增体系(11 μL):包括约 20 $\text{ng} \cdot \mu\text{L}^{-1}$ 的基因组 DNA 1 μL , 10 pmol 的上下游引物各 0.2 μL , $2 \times \text{Taq PCR MasterMix}$ (内含 TaqDNA 聚合酶、 Mg^{2+} 、dNTPs 等)(北京,天根)5 μL , 灭菌超纯水 3.6 μL , $10 \times \text{LC Green}$ 饱和染料(美国,Idaho) 1 μL 。

HRM-PCR 反应程序:95 $^{\circ}\text{C}$ 预变性 5 min;94 $^{\circ}\text{C}$ 变性 20 s, T_m 退火 20 s, 72 $^{\circ}\text{C}$ 延伸 20 s, 35 个循环;72 $^{\circ}\text{C}$ 延伸 5 min; 15 $^{\circ}\text{C}$ 保存。

高低温内标稀释方法:1OD 内标单链 + 400 μL 水, 退火体系:饱和氯化钠 1 μL , 内标互补双链各 1 μL , 补水至 10 μL , 95 $^{\circ}\text{C}$ 水浴 3 min, 然后缓慢降至室温。

HRM 荧光信号采集:将稀释好的高低温内标各 1 μL 加入 HRM-PCR 产物单孔中, 95 $^{\circ}\text{C}$ 水浴 30 s; 25 $^{\circ}\text{C}$ 水浴 30 s; 放入 LightScanner 96(美国,Idaho) 中收集荧光信号, 荧光信号温度收集区间设为 60~95 $^{\circ}\text{C}$, 运用 LightScanner 软件进行熔解曲线分析。

1.8 数据统计

根据群体遗传学理论直接计算得到基因型和等位基因频率。分别统计不同种牦牛群体 CART 基因多态性位点的纯合度(H_o)、杂合度(H_e)、有效等位

基因数(N_e)和多态信息含量(PIC)^[23-25], 并进行单倍型分析。利用 SPSS17.0 软件对 CART 基因多态性位点在不同种牦牛群体中的基因型分布进行卡方检测。

2 结果

2.1 牦牛 CART 基因序列和多态性分析

本研究使用 CART-1 和 CART-3 引物, 利用 DNA 混合池技术, 结合 HRM 共检测发现牦牛 CART 基因 4 个 SNPs 位点, 分别命名为 V-221(C>G)、V168(C>T)、V1791(G>A)和 V1816(C>T)(图 1)。V-221 在牦牛 CART 基因的上游序列, V168 位于第 1 内含子, V1791 和 V1816 位于 3'-UTR(图 2)。

对 CART-2 引物扩增产物进行克隆测序和比对发现, 大通牦牛和帕里牦牛存在缺失突变, 命名为 M1153-1190, 而甘南牦牛和天祝牦牛中没有发现。

使用 MEGA5.0 拼接出牦牛 CART 基因序列, 并运用 Spidey 软件分析发现:牦牛 CART 基因的外显子、内含子及蛋白编码区与牛 CART 基因一致。

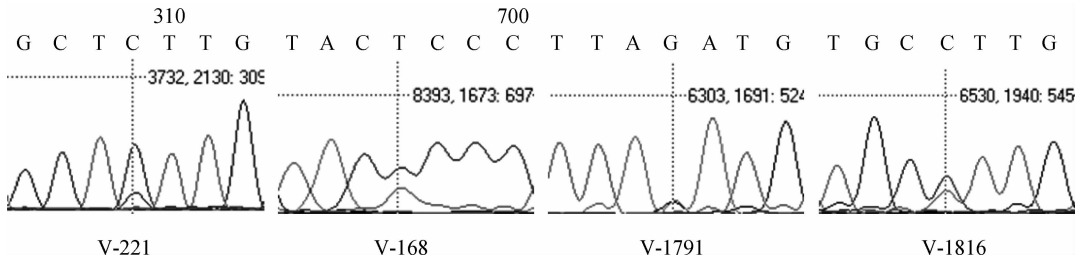


图 1 4 个 SNPs 的 DNA 片段测序图谱

Fig. 1 The sequencing chromatograms of the fragments carrying the 4 SNPs

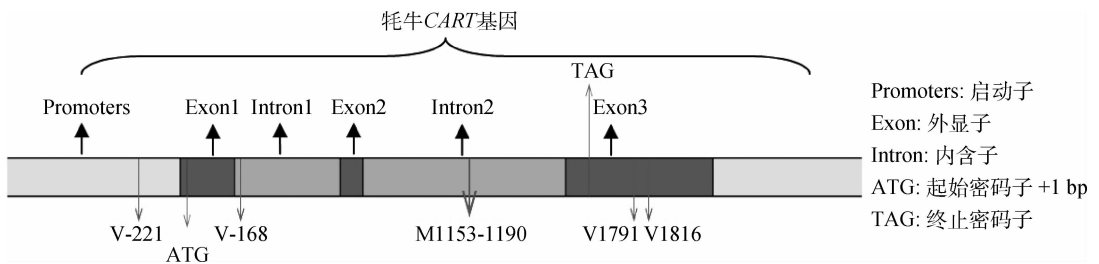


图 2 牦牛 CART 基因结构图

Fig. 2 The structural map of yak CART gene

V-221 基因型判定: 利用 HRM-PCR 引物 SNP1 扩增包含多态性位点在内的 55 bp 的 DNA

片段, 然后将扩增通过 Light Scanner(美国,Idaho) 扫描, 通过熔解曲线分析出有 2 种基因型: CC 和

CG 基因型(图 3a)。统计分析发现,CC 基因型属于优势基因型,在所有群体中并未能找到 GG 基因型,Hardy-Weinberg 平衡检验发现,除大通牦牛外($P < 0.05$,表 4),其他 3 种牦牛都处于哈代-温伯格平衡状态($P > 0.05$,表 4)。

V168 基因型判定:利用引物 SNP2 扩增包含 V168 位点在的 84 bp 的 DNA 片段,通过熔解曲线分析有 3 种基因型:CC、CT 和 TT 基因型(图 3b)。TT 基因型属于劣势基因型,该位点所有品种群体都处于 Hardy-Weinberg 平衡状态($P > 0.05$,表 4)。

V1791 基因型判定:利用引物 SNP3 扩增包含多态性位点在的 61 bp 的 DNA 片段,通过熔解曲线分析有 3 种基因型:GG、GA 和 AA 基因型(图 3c)。GG 基因型属于优势基因型,个体数明显多于其他基因型。所有牦牛群体都处于 Hardy-Weinberg 平衡状态($P > 0.05$,表 4)。

V1816 基因型判定:利用引物 SNP4 扩增包含

多态性位点在的 94 bp 的 DNA 片段,通过熔解曲线分析有 3 种基因型:CC、CT 和 TT 基因型(图 3d)。除甘南牦牛外($P < 0.05$,表 4),其他 3 种牦牛都处于 Hardy-Weinberg 平衡状态($P > 0.05$,表 4)。

2.2 牦牛群体参数分析

为了探索 *CART* 基因 SNPs 位点在 4 个不同品种牦牛群体中的遗传特性,对其纯合度(H_o)、杂合度(H_e)、有效等位基因数(N_e)和多态信息含量(PIC)指标进行统计,结果显示:4 种牦牛 *CART* 基因 4 个多态性位点的杂合度为 0.037 3~0.479 2,有效等位基因数为 1.038 7~1.920 1。V-221 位点在 4 种牦牛都属于轻度多态性($PIC < 0.25$);V168 位点除在帕里牦牛属于轻度多态性($PIC < 0.25$),其他牦牛都属于中度多态性($0.25 < PIC < 0.50$);V1791 位点 4 种牦牛都属于轻度多态性($PIC < 0.25$);V1816 位点 4 种牦牛都属于中度多态性($0.25 < PIC < 0.50$)(表 5)。

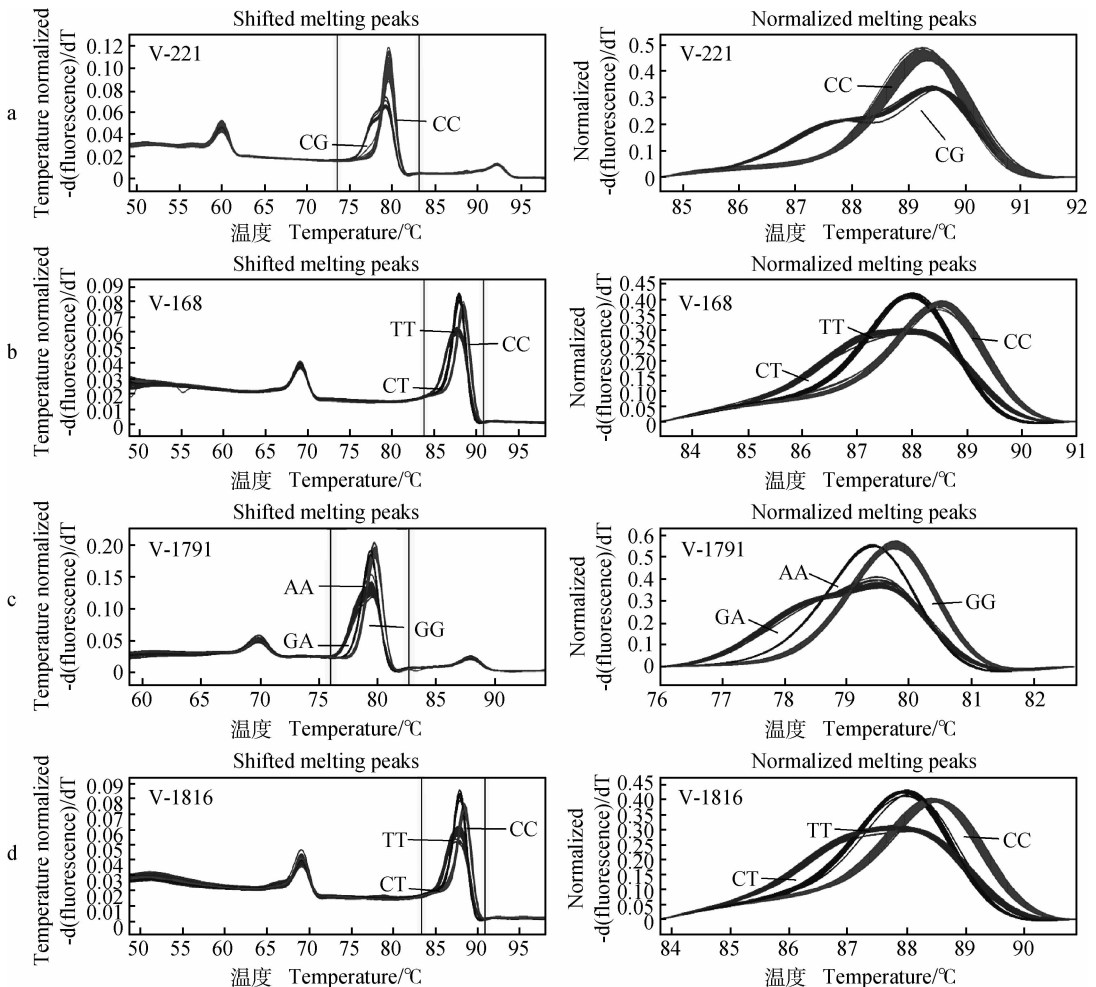


图 3 4 个 SNPs 位点熔解曲线和正常化熔解曲线
Fig. 3 The HRM analysis of the 4 SNPs

表 4 4 种牦牛 *CART* 基因 4 个 SNPs 位点的基因型分布和等位基因频率

Table 4 The genotype distributions and allelic frequencies of the 4 SNPs in *CART* gene in 4 breeds of yak

位点 Loci	观察基因型 Observed genotype			等位基因频率 Allele frequency		哈代-温伯格平衡 HWE	
V-221	CC	CG	GG	C	G	χ^2	<i>P</i> value
大通牦牛 Datong	184	59	0	0.879	0.121	4.639	<i>P</i> <0.05
甘南牦牛 Gannan	136	16	0	0.947	0.053	0.469	<i>P</i> >0.05
帕里牦牛 Pali	51	2	0	0.981	0.019	0.020	<i>P</i> >0.05
天祝牦牛 Tianzhu	176	27	0	0.933	0.067	1.030	<i>P</i> >0.05
V168	CC	CT	TT	C	T		
大通牦牛 Datong	133	94	16	0.741	0.259	0.012	<i>P</i> >0.05
甘南牦牛 Gannan	71	64	12	0.701	0.299	0.212	<i>P</i> >0.05
帕里牦牛 Pali	38	15	0	0.858	0.142	1.440	<i>P</i> >0.05
天祝牦牛 Tianzhu	77	93	22	0.643	0.357	0.588	<i>P</i> >0.05
V1791	GG	GA	AA	G	A		
大通牦牛 Datong	210	27	0	0.943	0.057	0.865	<i>P</i> >0.05
甘南牦牛 Ganna	151	57	2	0.855	0.145	1.824	<i>P</i> >0.05
帕里牦牛 Pali	41	8	1	0.9	0.1	0.617	<i>P</i> >0.05
天祝牦牛 Tianzhu	142	19	0	0.941	0.059	0.633	<i>P</i> >0.05
V1816	CC	CT	TT	C	T		
大通牦牛 Datong	117	78	14	0.746	0.254	0.042	<i>P</i> >0.05
甘南牦牛 Gannan	49	103	15	0.602	0.398	13.744	<i>P</i> <0.01
帕里牦牛 Pali	27	12	1	0.65	0.35	0.061	<i>P</i> >0.05
天祝牦牛 Tianzhu	66	69	9	0.698	0.302	2.679	<i>P</i> >0.05

$\chi^2_{0.05}(df=1) = 3.84$; $\chi^2_{0.01}(df=1) = 6.63$

表 5 4 种牦牛 *CART* 基因的遗传变异特征

Table 5 Genetic diversity of the *CART* gene in 4 breeds of yak

位点 Loci	品种 Breed	纯和度 <i>H_o</i>	杂合度 <i>H_e</i>	有效等位基因数 <i>N_e</i>	多态信息含量 <i>PIC</i>
V-221	大通牦牛	0.787 3	0.212 7	1.270 2	0.190 1
	甘南牦牛	0.899 6	0.100 4	1.111 6	0.095 3
	帕里牦牛	0.962 7	0.037 3	1.038 7	0.036 6
	天祝牦牛	0.875 0	0.125 0	1.142 9	0.117 2
V168	大通牦牛	0.616 2	0.383 8	1.622 9	0.310 2
	甘南牦牛	0.580 8	0.419 2	1.721 8	0.331 3
	帕里牦牛	0.756 3	0.243 7	1.322 2	0.214 0
	天祝牦牛	0.540 9	0.459 1	1.848 8	0.353 7
V1791	大通牦牛	0.892 5	0.107 5	1.120 5	0.101 7
	甘南牦牛	0.752 1	0.248 0	1.329 7	0.217 2
	帕里牦牛	0.820 0	0.180 0	1.219 5	0.163 8
	天祝牦牛	0.889 0	0.111 0	1.124 9	0.104 9
V1816	大通牦牛	0.621 0	0.379 0	1.610 2	0.307 2
	甘南牦牛	0.520 8	0.479 2	1.920 1	0.364 4
	帕里牦牛	0.545 0	0.455 0	1.834 9	0.351 5
	天祝牦牛	0.578 4	0.421 6	1.728 9	0.332 7

2.3 4 种牦牛 *CART* 基因单倍型分析

使用 SHEsis 软件^[26]对 4 个多态位点进行单倍型分析,共发现 10 个单倍型,且 4 种牦牛的 10 个单倍型频率整体趋势相同(表 6)。进一步分析发现,单倍型 C 在 4 种牦牛群体中频率都是最高,分别为大通牦牛 57.00%、甘南牦牛 52.90%、帕里牦牛 73.7%和天祝牦牛 47.3%。其次是单倍型 G、D,频

率最低的 3 个单倍型分别是 J、I 和 E。

2.4 牦牛 *CART* 基因启动子序列分析

通过使用 TRANSFAC 数据库中的 Match-1.0 Public 对扩增出的牦牛启动子序列进行分析,发现了 CRE 转录因子作用位点(图 4)。通过同源序列比对分析发现了牦牛启动子序列中的 TATA-Box(表 7)。

表 6 4 个 SNPs 位点单倍型在各群体中的频率分布

Table 6 The Haplotype distribution of SNPs in the 4 yak breeds

单倍型 Haplotype	天祝牦牛 Tianzhu frequency	大通牦牛 Datong frequency	甘南牦牛 Gannan frequency	帕里牦牛 Pali frequency	%
CCAC (A)	4.20	2.40	1.60	5.40	
CCAT (B)	4.60	1.30	2.00	3.40	
CCGC (C)	47.30	57.0	52.90	73.70	
CCGT (D)	9.40	5.40	8.70	4.10	
CTAC (E)	1.20	1.20	0.00	2.20	
CTGC (F)	7.10	3.00	5.20	2.80	
CTGT (G)	22.3	16.35	23.20	7.20	
GCGC (H)	4.00	10.60	4.20	1.20	
GTAT (I)	0.00	0.50	2.30	0.00	
GTGT (J)	0.00	2.10	0.00	0.00	

AGCTCTTGCCCATCTGTGCGCAGAACCTCGTTCCGGGGCGCCTGGAGCCCCGGCGGGCATTG
ACGTCAAAGCGGCGGCGAAGCGCTGCCTACAGACGGTTGACCCGGACCCTCCTCCACACCC
 CCTTCCTTCCTTCATCGCCCCCTCCTTCTCCCTGCGCCCCGGCTCAGGCTCGCTATAAAAG
 GTGGGAGCGCAGGGTGCCCGAGTAGCGCCGAGTTTCAGCACCATGGAGAGCCCCCGTCT

下划线部分为顺式元件; ATG 为起始密码子(+1 bp)

The sequences underlined are cis-elements; ATG is the start codon(+1 bp)

图 4 牦牛 *CART* 基因部分启动子序列

Fig. 4 The partial promoter sequence of yak *CART* gene

表 7 牦牛 *CART* 基因的部分顺式元件

Table 7 Cis-elements of *CART* gene in yak

顺式元件名称 Cis-element	位置/bp Location	序列 Sequence
TATA-Box	-50~-45	TATAAA
CRE	-166~-159	TGACGTCA

2.5 牦牛 mRNA 二级结构预测

本试验在 *CART* 基因 3'-UTR 内发现存在 2

个多态性位点,通过 Vienna RNA 软件包中的 RNA-flofold, exe 2.0 预测显示,*CART* 基因 3'-UTR 的二级结构和最低自由能值为 $-105.62 \text{ kcal} \cdot \text{mol}^{-1}$ (图 5a),发现多态性位点 V1791 碱基由 G 变为 A 时,结构发生较大改变(图 5b),最低自由能值变为 $-103.89 \text{ kcal} \cdot \text{mol}^{-1}$;多态性位点 V1816 碱基由 C 变为 T 时,结构发生轻微变化(图 5c),最低自由能值变为 $-104.12 \text{ kcal} \cdot \text{mol}^{-1}$ 。

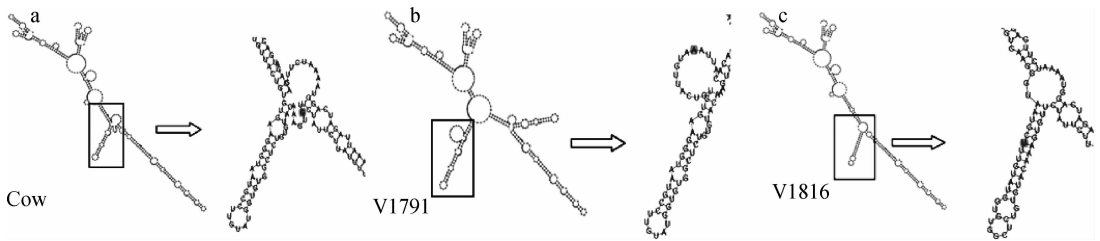


图 5 RNA 二级结构预测图

Fig. 5 The picture of RNA secondary structure prediction

3 讨论

本研究通过分段测序的方法克隆测序得到牦牛 *CART* 基因的全序列, 并发现 4 个新的 SNPs 位点为牦牛所特有, 其中 1 个位于启动子区域; 1 个位于内含子; 2 个位于 3'-UTR。Hardy-Weinberg 平衡检测表明, 除大通牦牛的 V-221 位点和甘南牦牛的 V1816 位点不处于 Hardy-Weinberg 平衡状态 ($P < 0.05$), 其他位点都处于哈代-温伯格平衡状态 ($P > 0.05$)。大通牦牛和甘南牦牛的长期选育是造成其多态性位点不处于 Hardy-Weinberg 平衡状态的主要原因之一。

遗传变异特征分析表明, V-221 和 V1791 在 4 个牦牛群体中均属于轻度多态, V168 和 V1816 位点属于中度多态, 说明在该位点的选择潜力较大。多态位点 (SNP) 和差异位点 (MS) 分析表明, V-221 位点 (位于 *CART* 基因 5' 侧翼序列) GG 型的个体数为 0, CG 型个体比较少。K. Yamada 等^[14] 克隆人 *CART* 基因启动子序列, 分析该区域多态性位点与肥胖症关联分析发现转录位点上游 -156 bp 处的多态位点 A 变为 G 可能导致机体发生肥胖症的几率增加, 说明 *CART* 基因启动子区域的碱基变化可能影响基因的功能。V168 位点 (位于内含子 1) 在所有群体中 C 碱基的频率大于 T 碱基, 说明群体中 CC 基因型可能是优势基因型。V1791 与 V1816 (位于 3'-UTR) 的碱基变化影响了所预测的 mRNA 二级结构及其最低自由能值, V1791 位点的碱基由 G 变为 A, 最低自由能值增加了 $1.73 \text{ kcal} \cdot \text{mol}^{-1}$, 且明显改变 mRNA 的稳定二级结构, 所以该位点 GG 型的结构比 AA 型的结构稳定; V1816 位点的碱基由 C 变为 T, 最低自由能值增加了 $1.5 \text{ kcal} \cdot \text{mol}^{-1}$, 且局部改变 mRNA 的二级结构, 所以该位点 CC 型的结构比 TT 型稳定。mRNA 的二级结构变化可能影响 mRNA 的表达^[27-29], 与群体中 V1791 位点 G

的频率远大于 A, V1816 位点 C 的频率大于 T 一致。单倍型分析结果表明, 单倍型 C 的频率在 4 种牦牛群体中都超过 45%, 在帕里牦牛群体中单倍型频率达到 73.7%, 说明单倍型 C 在 4 种牦牛群体中都是优势单倍型。在对牦牛 *CART* 基因的启动子序列分析并没有发现在猪上报道的 ZBPF (锌指蛋白) 结合位点^[30], 而 P. Barrett 等^[31] 研究大鼠 *CART* 基因的启动子时发现 1 个位于 -162 bp 处的调控基因表达的转录因子 CRE (cAMP 应答元件结合因子), 这与本研究发现位于 -316 ~ -309 bp 的 CRE 相同。说明 CRE 可能是 *CART* 基因各物种共有的转录因子, 而 ZBPF 则可能不是, 更说明 CRE 在 *CART* 基因表达中起的作用。

本研究以大通牦牛、甘南牦牛、天祝牦牛和帕里牦牛为试验材料, 运用 DNA 测序和 HRM 等技术克隆 *CART* 基因并对该基因序列的变异位点和其多态性分布进行了全面的探索。分析了 4 种牦牛群体遗传学特征, 并应用生物信息学推测各个位点变异可能产生的影响。结果表明: 群体遗传学分析发现 4 个 SNPs 位点的杂合度为 $0.0373 \sim 0.4792$, 多态信息含量为 $0.0366 \sim 0.3644$, 多属于轻度或中度多态性; V1791 和 V1816 位点的碱基变异对 mRNA 二级结构和最低自由能值可能产生影响。

参考文献:

- [1] LV L, JIMENEZ K F, SEN A, et al. Evidence supporting a role for cocaine- and amphetamine-regulated transcript (CARTPT) in control of granulosa cell estradiol production associated with dominant follicle selection in cattle[J]. *Biol Reprod*, 2009, 81(3): 580-586.
- [2] 张涛, 李富禄, 黄洋, 等. *CART* 对 FSH 诱导的绵羊卵巢卵泡颗粒细胞雌激素分泌影响的研究[C]. 中国畜牧兽医学动物繁殖学分会第十五届学术研讨会, 2010:304.

- [3] SEN A, BETTEGOWDA A, JIMENEZ K F, et al. Cocaine- and amphetamine-regulated transcript regulation of follicle-stimulating hormone signal transduction in bovine granulosa cells[J]. *Endocrinology*, 2007, 148(9):4400-4410.
- [4] SEN A, LV L, BELLO N, et al. Cocaine- and amphetamine-regulated transcript accelerates termination of follicle-stimulating hormone-induced extracellularly regulated kinase 1/2 and Akt activation by regulating the expression and degradation of specific mitogen-activated protein kinase phosphatases in bovine granulosa cells[J]. *Mol Endocrinol*, 2008, 22:2655-2676.
- [5] SPIESS J, VILLARREAL J, VALE W. Isolation and sequence analysis of a somatostatin-like polypeptide from ovine hypothalamus [J]. *Biochemistry*, 1981, 20:1982-1988.
- [6] JANIUK I, MYNEK K, WYSOCKI J. Identification and location of the cocaine and amphetamine regulated transcript (CART) in the abomasum of cattle[J]. *Acta Histochemica*, 2012, pii: S0065-1281(12)00119-5.
- [7] KRISTENSEN P, JUDGE M E, THIM L, et al. Hypothalamic CART is a new anorectic peptide regulated by leptin[J]. *Nature*, 1998, 393:72-76.
- [8] STANLEY S A, MUPRHY K G, BEWIEK G A, et al. Regulation of rat pituitary cocaine- and amphetamine-regulated transcript (CART) by CRH and glucocorticoids[J]. *Am J Physiol Endocrinol Metab*, 2004, 287:E583-E590.
- [9] OHSAWA M, DUN S L, TSENG L F, et al. Decrease of hindpaw with drawal latency by cocaine- and amphetamine-regulated transcript peptide to the mouse spinal cord[J]. *Eur J Pharmacol*, 2000, 399(2-3):165-169.
- [10] WOJTKIEWICZ J, GONKOWSKI S, BLADOWSKI M, et al. Characterisation of cocaine- and amphetamine-regulated transcript-like immunoreactive (CART-LI) enteric neurons in the porcine small intestine[J]. *Acta Vet Hung*, 2012, 60(3): 371-381.
- [11] KUHAR M J, JAWORSKI J N, HUBERT G W, et al. Cocaine- and amphetamine-regulated transcript peptides play a role in drug abuse and are potential therapeutic targets[J]. *AAPS J*, 2005, 7:259-265.
- [12] VALLE E, MOORE S S, JANN O, et al. The bovine cocaine and amphetamine-regulated transcript locus: gene characterization and SNP discovery [J]. *Anim Genet*, 2005, 36(1):74-75.
- [13] ZHANG C L, CHEN H, WANG Y H, et al. The polymorphisms of bovine cocaine- and amphetamine-regulated transcripts and their associations with cattle (*Bos taurus*) growth traits[J]. *J Biosci*, 2008, 33(3):365-370.
- [14] YAMADA K, YUAN X, OTABE S. Sequencing of the putative promoter region of the cocaine- and amphetamine-regulated-transcript gene and identification of polymorphic sites associated with obesity[J]. *Int J Obes Relat Metab Disord*, 2002, 26(1):132-136.
- [15] CHALLIS B G, YEO G S, FAROOQI I S, et al. The CART gene and human obesity: mutational analysis and population genetics[J]. *Diabetes*, 2000, 49(5):872-875.
- [16] 李爱民, 马云, 杨东英, 等. 鲁西牛 ANGPTL6 基因的 3 个多态位点与其生长性状的关联性分析[J]. *中国农业科学*, 2012, 45(11): 2306-2314.
- [17] WHEELAN S J, CHURCH D M, OSTELL J M. Spidey: a tool for mRNA-to-genomic alignments[J]. *Genome Res*, 2001, 11(11):1952-1957.
- [18] KUMAR S, TAMURA K, NEI M. MEGA3: Integrated software formolecular evolutionary genetics analysis and sequence alignment[J]. *Bioinformatics*, 2004, 5: 150-163.
- [19] DENMAN R B. using RNAFOLD to predict the activity of small catalytic RNAs [J]. *Biotechniques*, 1993, 15(6):1090-1095.
- [20] LIEW M, PRYORR, PALAIS R, et al. Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons[J]. *Clin Chem*, 2004, 50(7): 1156-1164.
- [21] YE M H, CHEN J L, ZHAO G P. Associations of A-FABP and H-FABP markers with the content of intramuscular fat in Beijing-you chicken [J]. *Anim Biotechnol*, 2010, 21:14-24.
- [22] GUNDRY C N, DOBROWOLKI S F, MARTIN Y R, et al. Base-pair neutral homozygotes can be discriminated by calibrated high-resolution melting of small amplicons [J]. *Nucleic Acids Res*, 2008, 36(10): 3401-3408.
- [23] NEI M, ROYCHOUDHURY A K. Sampling variances of heterozygosity and genetic distance[J]. *Genetics*, 1974, 76(2): 379-390.
- [24] NEI M, LI W H. Mathematical model for studying genetic variation in terms of restriction endonucleases [J]. *Proc Natl Acad Sci USA*, 1979, 76(10): 5269-5273.

- [25] BOTSTEIN D, WHITE R L, SKOLNICK M, et al. Construction of a genetic linkage map in man using restriction fragment length polymorphisms[J]. *Am J Hum Genet*, 1980, 32(3): 314-331.
- [26] SHI Y Y, HE L. SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci[J]. *Cell Res*, 2005, 15(2):97-98.
- [27] CHATTERJEE S, PAL J K. Role of 5'- and 3'-untranslated regions of mRNAs in human diseases[J]. *Biol Cell*, 2009, 101(5):251-262.
- [28] ABANADES D R, RAMÍREZ L, IBORRA S, et al. Key role of the 3' untranslated region in the cell cycle regulated expression of the *Leishmania infantum* histone H2A genes: minor synergistic effect of the 5' untranslated region[J]. *BMC Mol Biol*, 2009, 10: 48.
- [29] JI Z, TIAN B. Reprogramming of 3' untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types[J]. *PLoS ONE*, 2009, 4(12):e8419.
- [30] LING F, WEI L, WANG T, et al. Cloning and characterization of the 5'-flanking region of the pig cocaine- and amphetamine-regulated transcript gene[J]. *DNA Cell Biol*, 2011, 30(2): 91-97.
- [31] BARRETT P, DAVIDSON J, MORGAN P. CART gene promoter transcription is regulated by a cyclic adenosine monophosphate response element[J]. *Obes Res*, 2002, 10(12):1291-1298.

(编辑 郭云雁)