

DOI: 10.3724/SP.J.1005.2012.01304

基于高密度 SNP 标记的肉牛人工选择痕迹筛查

刘喜冬^{1,2}, 王志鹏², 樊惠中¹, 李俊雅¹, 高会江¹

1. 中国农业科学院北京畜牧兽医研究所, 中国农业科学院肉牛研究中心, 农业部畜禽遗传资源与利用重点开放实验室, 北京 100193;

2. 东北农业大学动物科学技术学院, 哈尔滨 150030

摘要: 近年来随着遗传改良工作的实施, 人工选择大大提高了肉牛的生产性能并使其遗传基础发生巨大改变。

文章基于 Illumina BovineSNP50(54K)和BovineHD(770K)两款芯片数据, 采用 F_{ST} 检验方法分析牛群的遗传分化, 并筛查人工选择在牛的基因组留下的印记。通过全基因组范围内的扫描, 共发现 47 104 个“离群”位点和 3064 个群体特异的人工选择“候选基因”, 如 *CLIC5*、*TG*、*CACNA2D1*、*FSHR* 等。通过基因注释对基因的生物学过程和分子功能进行富集分析。文章构建了我国肉牛的全基因组的选择信号图谱, 为深入研究人工选择和理解生物进化提供线索, 且研究结果也显示人工选择对基因组的影响在牛品种遗传改良中发挥了重要作用。

关键词: 牛的基因组; 人工选择; 单核苷酸多态性; 基因注释

Artificial selection for cattle based on high-density SNP markers

LIU Xi-Dong^{1,2}, WANG Zhi-Peng², FAN Hui-Zhong¹, LI Jun-Ya¹, GAO Hui-Jiang¹

1. Key Laboratory of Farm Animal Genetic Resources and Utilization of Ministry of Agriculture, Beef Cattle Research Center, Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, Beijing 100193, China;

2. College of Animal Science and Technology, Northeast Agricultural University, Harbin 150030, China

Abstract: With the implementation of genetic improvement in recent years, artificial selection has greatly improved beef cattle production performance and its genetic basis has been dramatically changed. In this study, based on the Illumina BovineSNP50 (54K) and BovineHD (770K) BeadChip and the F_{ST} value, we analyzed the genetic differentiation of cattle and screened the imprints of selection in bovine genome. Finally, we found 47104 OUTLIER SNP loci and 3064 candidate genes, for example, *CLIC5*, *TG*, *CACNA2D1*, and *FSHR* etc. The biological processes and molecular functions of genes were analyzed through gene annotation. The results of this study established a genome-wide map of selection footprints in beef cattle genome and a clue for in-depth study of artificial selection and understanding of biological evolution. Our results indicate that artificial selection has played an important role in cattle breed genetic improvement.

Keywords: bovine genome; artificial selection; single nucleotide polymorphism; gene annotation

收稿日期: 2012-05-25; 修回日期: 2012-08-13

基金项目: 优质肉牛新品种(系)选育与关键技术研究与示范项目(编号: 2011BAD28B04), 基于高密度 SNP 建立中国肉牛群体全基因组选择优化方案(编号: 2010jc-2)和现代农业(肉牛)产业技术体系岗位科学家项目(编号: CARS-38)资助

作者简介: 刘喜冬, 硕士研究生, 专业方向: 数量遗传学。E-mail: liuxdchn@gmail.com

通讯作者: 高会江, 博士, 副教授, 研究方向: 数量遗传学。E-mail: gaohj111@sina.com

网络出版时间: 2012-9-10 03:16:35

URL: <http://www.cnki.net/kcms/detail/11.1913.R.20120910.1516.005.html>

牛在大约 8 千到 1 万年前被驯化用来提供牛奶和肉类产品^[1], 从那时起, 广泛的自然和人为因素的影响如地理、环境、文化和定向人工选择导致了牛品种的多样性。自牛基因组测序完成以来, 牛结构基因组学与功能基因组学研究发展迅速, 尤其在牛重要经济性状的结构基因组学研究领域中取得了大量的成果。2008 年牛全基因组高密度 SNP 芯片的问世加快了从分子水平上评估人工选择对基因组进化的影响。在基因组水平上研究选择的作用以及寻找受到选择的基因已成为进化生物学界的一个研究热点。其研究一方面为了验证分子进化的中性理论与选择学说, 另一方面则是寻找受到选择的基因。一系列研究表明选择在 DNA 分子水平上的进化起到了重要的作用。与此同时, 相关的统计分析方法也有了显著发展。

F_{ST} 检验已经被证实是最简单且有效的群体遗传指标之一^[2], 基于该策略的 outlier 方案能有效排除混杂因素确定人工选择的候选位点, 这就为在基因组水平上进行人工选择痕迹筛查的研究提供了可能。 F_{ST} 检验是 1984 年 Weir 和 Cockerham 提出的^[3], 2002 年 Akey 等^[4]采用该方法率先分析了 26530 个 SNP 位点在人群间的频率差异, 发现了 156 个正选择“候选基因”。2008 年 Hayes 等^[5]在奶牛和肉牛群体中通过分析 9 323 个 SNP 证明了该方法的功效。本研究利用最新的肉牛高密度 SNP 芯片, 以内蒙古乌拉盖地区杂交肉牛为研究对象进行群体基因组学研究, 选用基于群体遗传结构的 F_{ST} 值筛查人工选择在牛的基因组留下的选择信号, 研究人工选择对牛基因组进化的影响, 利用基因功能富集分析和通过文献挖掘发现本研究所找到的可能受到人工选择影响的基因中存在与牛肉质和繁殖性状相关的基因, 为深入研究人工选择和理解生物进化提供线索。

1 材料和方法

1.1 实验动物

实验动物选自内蒙古锡林格勒盟乌拉盖管理区构建的西门塔尔牛资源群体, 通过自然交配与冷配, 分别获得育肥公牛 307 头和 332 头, 所有个体的系谱均完整且经过校正。采集动物血液, 提取 DNA, 使用 Illumina Bovine SNP50(54K) 和 BovineHD(770K)

获得 SNP 分型数据。本研究采用的分型数据质量控制标准为:

(1) 个体的筛选: SNP 基因型缺失率小于 5%; 孟德尔遗传错误率小于 5%。

(2) SNP 位点的筛选: 分型成功率大于 95%; 最小等位基因频率大于 5%; 达到哈迪-温伯格平衡检验 ($P > 10^{-6}$)。

1.2 基因型填充

由于实验样本分别采用 Illumina Bovine SNP50 (54K) 和 BovineHD(770K) 两款芯片, SNP 位点数目和 SNP 信息不同。为了扩大样本量和利用高密度 SNP 芯片信息, 本研究选取 Mach 软件将来自芯片 Bovine SNP50 的 2008 年的 136 头 54K 芯片数据填充成 770K 芯片数据, 根据 Pei 等^[6]的研究结果, 在几种常用的基因型填充方法中 (MACH、IMPUTE、fastPHASE、PLINK、Beagle), Mach 方法无论在不同 SNP 密度还是在不同连锁程度下, 其准确性均最高。

1.3 F_{ST} 检验

F_{ST} 值是检测种群间选择作用差异最广泛的一种方法, 本研究按照 Weir 等^[7,8]描述的方法计算了 F_{ST} 值的无偏估计值, 每个 SNP 在自然交配与冷配群体间的等位基因频率差异采用单位点配对 F_{ST} (Pair-wise F_{ST}) 衡量, 具体见公式 (1)~(4)。

$$F_{st} = \frac{MSP - MSG}{MSP + (n_c - 1)MSG} \quad (1)$$

$$MSG = \frac{1}{\sum_{i=1}^s (n_i - 1)} \sum_i n_i p_{Ai} (1 - p_{Ai}) \quad (2)$$

$$MSP = \frac{1}{s-1} \sum_i n_i (p_{Ai} - \bar{p}_A)^2 \quad (3)$$

$$n_c = \frac{1}{s-1} \left(\sum_{i=1}^s n_i - \frac{\sum_i n_i^2}{\sum_i n_i} \right) \quad (4)$$

其中, MSP 表示检测到的群体间的平均误差平方, MSG 表示检测到的群体内的位点的均方误差, p_{Ai} 表示第 i 个群体中 A 的 SNP 位点的等位基因频率, n_i 表示群体 i 的平均样本大小, $\bar{p} = n_i p_{Ai} / \sum_i n_i$ 表示所有群体 P_A 的加权平均值, n_c 表示校正过的整个样本的平均样本大小。

1.4 鉴定离群值

为了确定单位点 F_{ST} 的“离群值”，采用Boxplot策略^[9]确定各指标的上、下分界值(Threshold)。首先计算各染色体 F_{ST} 经验分布的四分位区间(Interquartile range, IQR)：

$$Q = F^U - F^L \quad (5)$$

其中, F^U 和 F^L 分别是上、下四分位值。则上、下分界值为：

$$UL = F^U + 1.5Q \quad (6)$$

$$LL = F^L - 1.5Q \quad (7)$$

所有大于上分界值(UL)的位点被定义为“离群”位点。

1.5 基因注释

本研究基于基因本体数据库(Gene ontology, GO)和KEGG的注释数据对“候选基因”进行基因注释,采用David6.7 分析平台^[9]对获得基因进行分子功能(Molecular function)和生物学过程(Biological process)的GO富集分析和KEGG通路富集分析。

本研究的数据处理分析,统计作图以及文本信息挖掘分别使用了 R 和 Perl。

2 结果与分析

2.1 SNP 筛选及质量控制结果

将 2008 年的 136 头牛的 54K 芯片利用基因型填充成 770K, 达到了增加样本量的目的。在对全部 639 头西门塔尔杂交牛的 SNP 分型数据进行质量控制后,发现常染色体区域共有 658 889 个位点和 631 个个体通过了质控标准。本研究所用到的实验群体(通过质控)在各个年度的分布见表 1。

表 1 质量控制后自然交配与冷配个体数

来源	2008 年 (头)	2009 年 (头)	2010 年 (头)	合计(头)
自然交配	100	132	72	304
冷配	36	160	131	327
合计	136	292	203	631

2.2 常染色体区域的人工选择信号

2.2.1 全基因组 F_{ST} 值的分布

本研究利用 F_{ST} 衡量 SNP 在自然交配与冷配个体

间的等位基因频率差异,构建了 F_{ST} 的经验分布(图 1),经验分布图的右上角为按照超出总体特征特定法确定的阈值之上的 F_{ST} 分布。由图可得,在 2008~2010 年度自然交配与冷配(人工授精)个体比较中,可发现在高 F_{ST} 值区域存在明显拖尾,提示这些标记所在的区域可能受到正选择的影响。

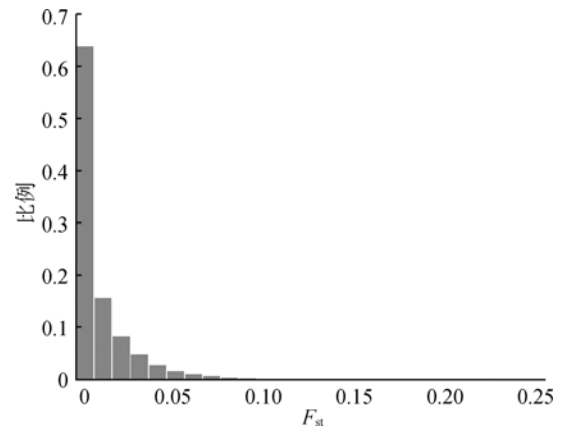


图 1 2008~2010 年自然交配与冷配个体的单位点配对 F_{ST} 在所有常染色体区的经验分布

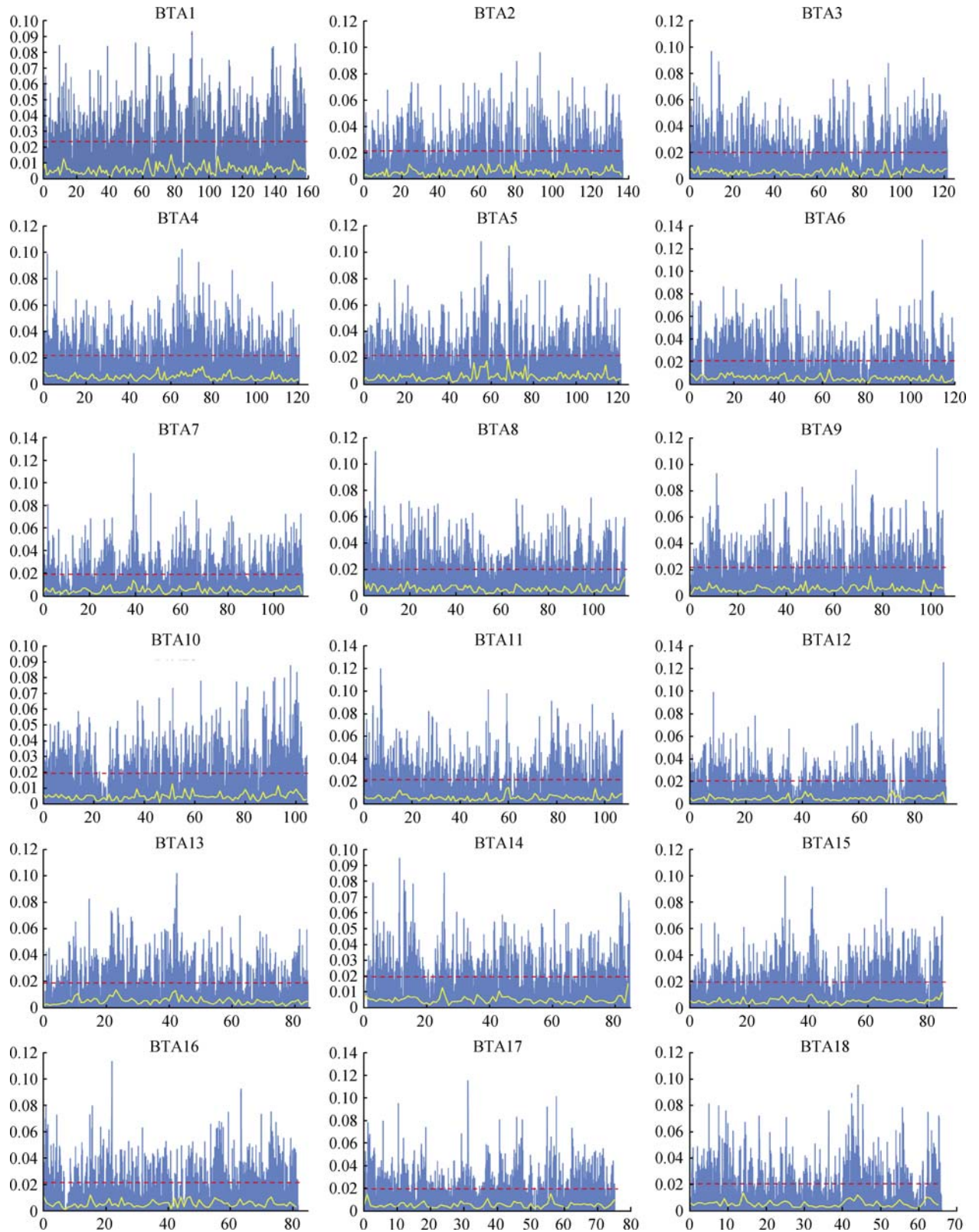
2.2.2 常染色体 F_{ST} 值的分布

为了进一步验证选择信号所在的区域,同时由于单位点 F_{ST} 的经验分布的分布趋势与自由度为 2(或 3)的卡方分布相类似,为了消除偏倚,得到稳健的结果,采用Boxplot策略鉴定人工选择的“离群”位点。该策略考虑了单个位点配对的 F_{ST} 值在各染色体与整个常染色体的分布之间存在显著差异,因此在上、下分界值的设定上采用了各染色体特定的分界值来确定相应的“离群”位点。

图 2 中超过阈值的部分即为受到人工选择作用的“离群”位点,可发现间隔越小的位点间的 F_{ST} 数值越相近,邻近位点间的 F_{ST} 值尤其是“离群”位点对应的 F_{ST} 值呈现出聚集的趋势,这可能提示邻近位点间隔更远的位点共享了更多的进化历史,这些位点均不同程度地受到了选择作用的影响。

2.3 受到选择的 SNP 位点

本研究在常染色体区域共检测到 47 104 个“离群”位点,约占总位点数的 7%。这些位点在基因组上分布见表 2,从表 2 可看出受到选择影响的位点在基因组上的分布不均匀,例如最少的是 12 号染色体



接下页

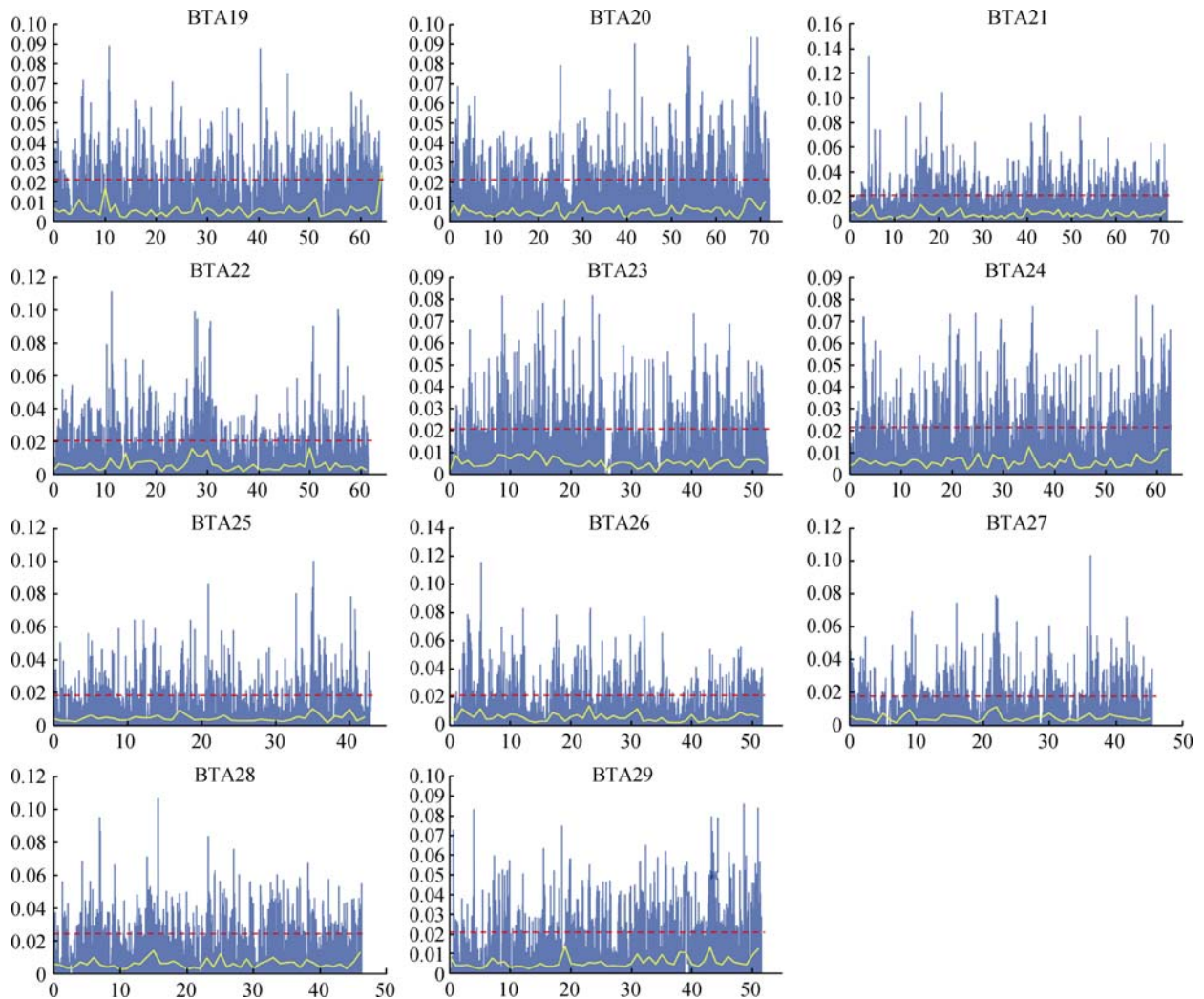


图 2 牛基因组各常染色体(BTA1~BTA29) F_{ST} 值分布图

黄色实线表示各染色体不重叠的 1Mb 区间范围内 F_{ST} 值做平均值; 红色虚线表示 Boxplot 策略确定的阈值。

每 Mb 仅有 15 个 SNP 位点, 最多的是 24 号染色体, 每 Mb 有 21.52 个位点。

2.4 确定候选基因

本研究利用 UCSC 数据库中牛基因组注释信息, 对检测到的 47 104 个受到人工选择的遗传标记进行了基因定位, 若受到选择的 SNP 位于某个基因内, 则将该基因定义为受到选择的“候选基因”。本研究共筛选到 3 064 个“候选基因”。通过文本挖掘和基因组注释等信息, 本研究对筛选到的“候选基因”进行了详细的注释。限于篇幅, 表 3 罗列了部分重要的“候选基因”, 主要包括该基因的物理位置和功能。在筛选到的候选基因中, 本研究发现了一些熟知的受到人工选择的基因, 如 *LIPC* 基因、*TCF12* 基因等^[10]。

2.5 候选基因的富集分析

本研究利用在线工具 DAVID(<http://david.abcc.ncifcrf.gov/>)对获得的 3 064 个“候选基因”进行富集分析, 并将 FDR 阈值设为 0.05。从表 4 可看出本研究筛选到受到人工选择的“候选基因”主要集中于离子转运、金属离子转运、钙离子转运和囊泡介导的转运等 4 个相关生物学过程类, 这些生物学过程共包括了 196 个受到人工选择的“候选基因”。在分子功能富集分析方面共发现有 710 个受到人工选择的“候选基因”富集在通道活性、基板的特异通道活性、金属离子的跨膜转运活性和门控通道活性等 8 个相关分子功能类, 相关信息见表 5。在 KEGG 富集分析发现有 54 个受到人工选择的“候选基因”富集于钙信号通路(FDR=6.34E-03, $P=5.11E-06$)。

表 2 上分界值和“离群”位点数

染色体	物理距离 (Mb)	上分界值	“离群”位点个数	“离群”位点密度
1	158	0.023	2565	16.23
2	137	0.021	2495	18.21
3	121	0.020	2238	18.50
4	121	0.022	2141	17.69
5	121	0.022	2317	19.15
6	119	0.021	2299	19.32
7	113	0.019	2171	19.21
8	113	0.020	2084	18.44
9	106	0.022	1974	18.62
10	104	0.019	2024	19.46
11	107	0.022	2089	19.52
12	91	0.021	1365	15
13	84	0.019	1650	19.64
14	85	0.019	1503	17.68
15	85	0.019	1790	21.06
16	82	0.021	1575	19.21
17	75	0.019	1588	21.17
18	66	0.021	1374	20.82
19	64	0.021	1160	18.13
20	72	0.021	1299	18.04
21	72	0.021	1459	20.26
22	61	0.021	1225	20.08
23	53	0.021	912	17.21
24	63	0.021	1356	21.52
25	43	0.018	822	19.12
26	52	0.021	939	18.06
27	45	0.017	858	19.07
28	46	0.025	796	17.30
29	52	0.021	1036	19.92
合计	2511	0.025	47104	18.76

3 讨 论

3.1 统计方法筛查人工选择信号的影响

在基因组水平的研究中, 由于使用不同的统计分析方法的效率不同, 对同一个基因组处理经常产生不同的结果, 主要的原因是不同的统计分析方法的效率不同。传统的分析策略是单基因检验, 但是整个基因组的DNA变异模式会受到种群动态历史的影响, 而人工选择作用通常只特异地作用于某一个位点^[19], 为了解决这个问题, 一些研究者在全基因

组范围内针对群体的遗传指标构建了经验分布, 使之反映基因组受到特定人口动态历史的影响, 同时, 受到了选择的位点则位于分布的边缘。最近, 众多选择研究发现, 受到选择的候选位点可以通过基于该策略的outlier方案来排除混杂的因素影响。而Thornton等^[6]研究发现, F_{ST} 是群体基因组学中最简单有效的群体遗传指标之一。其原理是在中性的条件下遗传漂变(Genetic drift)决定了群体间的 F_{ST} 值, 而 F_{ST} 值增大可能是由群体特异的正选择作用在候选基因及邻近的遗传标记引起的^[3,20,21]。例如, Beaumont等^[22]发现通过人群遗传分化分析检测到的正选择具有足够的效力。因此本研究中的单位点配对 F_{ST} 能有效发现牛SNP分型数据中的选择信号, 并且在常染色体区鉴别出了较高比例的高 F_{ST} “离群”位点(约 7%), 表明牛基因组中存在广泛分布的正选择信号^[23-26]。

3.2 鉴定离群值的方法

在目前研究中, 全基因组扫描的异常值方法(Outlier approach)主要包括以下几种: 一种是“把数据点在超出平均值两个或 3 个标准差以上的位点作为离群位点”; 另一种是“把数据点在超出特定百分位数以上的位点作为离群位点”的方法^[27]; 还有一种是Boxplot策略。前两种方法一般要求数据集服从正态分布或其他对称的钟形分布的假设, 而且由于实际数据中都存在少量集中存在的离群数据, 这些均会影响参数估计的准确性^[28]。但是, 对于Boxplot方法, 则无需对数据的分布进行附加假设, 也不需要估计分布参数(均值或标准差)。对于大部分数据而言, 其经验分布并不服从正态。本研究利用Boxplot方法鉴定出 47 104 个受选择的SNP位点。

3.3 选择对基因组的影响

本研究共筛查到 47 104 个受选择的SNP位点, 遍布整个基因组上。MacEachern等^[14]使用 7 500 个SNP比较了澳大利亚安格斯和荷斯坦牛的等位基因频率差异, 发现 8 号、20 号和 24 号染色体上各有一个受到选择作用影响的区域。Saber等^[16]使用bovine 50K SNP芯片对荷斯坦牛进行全基因组正向选择痕迹的扫描, 提出了荷斯坦牛的全基因组选择图谱。Laurence等^[17]利用 42 486 个SNP对荷斯坦牛、诺曼

表 3 部分重要“候选基因”位置和功能

染色体	位置(Mb)	基因	基因名称或功能
2	23.64~23.97	<i>RAPGEF4</i>	鸟嘌呤核苷酸交换因子 4
3	3.57~3.64	<i>RXRG</i>	维甲酸 X 受体, γ
4	63.05~63.31	<i>BMPEP</i>	骨形态发生蛋白内皮调节
4	38.34~38.86	<i>CACNA2D1</i>	电压依赖性钙通道 $\alpha 2$ 或 δ 亚基 1
4	97.79~98.60	<i>EXOC4</i>	参与葡萄糖的转运
10	3.20~3.37	<i>KCNN2</i>	调节神经细胞
10	53.09~53.48	<i>TCF12</i>	转录因子 12
10	51.76~51.92	<i>LIPC</i>	肝脂酶
11	31.11~31.30	<i>FSHR</i>	卵泡刺激素受体
11	107.16~107.19	<i>SIRT3</i>	调节细胞凋亡
14	16.97~17.13	<i>MTSS1</i>	转移抑制基因 1
14	9.29~9.51	<i>TG</i>	甲状腺球蛋白基因
15	31.80~32.14	<i>GRIK4</i>	离子型谷氨酸受体-红藻氨酸 4 基因
15	23.80~24.17	<i>NCAM1</i>	神经细胞粘附分子 1
15	38.99~39.31	<i>SPON1</i>	spondin1, 细胞外基质蛋白
23	19.17~19.33	<i>CLIC5</i>	氯离子转运中发挥独特作用
26	9.93~10.21	<i>RNLS</i>	肾酶, FAD 的依赖性胺氧化酶

表 4 受选择基因在 GO 的生物学过程(BP)富集

GO 功能分类	基因数	P	FDR
离子转运	125	1.01E-07	1.84E-04
囊泡介导的转运	69	2.59E-06	4.71E-03
钙离子转运	29	3.54E-06	6.44E-03
金属离子转运	73	1.24E-05	2.25E-02

注: P 值表示显著水平, FDR 表示错误发现率。下表同。

表 5 受选择基因在 GO 的分子功能(MF)富集

GO 功能分类	基因数	P	FDR
门控通道活性	23	4.00E-08	5.58E-05
钙离子结合	35	2.94E-06	4.10E-03
离子通道的活性	23	4.48E-06	6.24E-03
基板的特异通道活性	23	5.36E-06	7.47E-03
金属离子的跨膜转运活性	19	6.30E-06	8.79E-03
通道活性	23	6.88E-06	9.59E-03
被动跨膜转运活性	23	6.88E-06	9.59E-03
阳离子通道的活性	17	1.20E-05	1.68E-02

底牛和蒙贝利亚牛执行了等位基因频率差异的全基因组扫描, 发现了 13 个受到强烈人工选择的区域。Saber 等^[18]采用 Bovine 50K SNP 芯片在德国、加拿大和澳大利亚的荷斯坦牛、瑞士褐牛、澳大利亚安格斯牛、赫里福德牛和西门塔尔牛中分别找到了 19、27、9、10 和 17 个受到人工选择的区域。

利用基因组注释信息, 本研究共筛选到 3 064 个可能受到人工选择的基因。通过文本挖掘发现这些基因或与牛重要经济性状相关, 或参与一些重要的生物学功能。发现 *RXRG*、*CACNA2D1* 和 *FSHR* 等可能受到人工选择的基因与繁殖性状相关。*RXRG* 基因是组织形态发生和细胞分化的主要调节因子^[29], 而且该基因富集到离子结合、阳离子结合和金属离子结合分子功能类中。一些关联分析发现该基因的多态性与牛产胎性能相关, 其中黄荫等^[30]在鲁西牛、中国西门塔尔牛、安格斯牛和西蒙杂交牛单、二胎群体间发现该基因的 3'UTR 存在一个 SNP, 该 SNP 与双胎性能显著相关 ($P < 0.01$)。 *CACNA2D1* 基因是编码动物 L-型钙离子通道亚基的基因, 该基因富集在门控通道活性、离子通道活性及基板的特异通道活性分子功能类中。Hou 等^[31]在 8 个肉牛品种发现 *CACNA2D1* 基因与胴体重、屠宰率、净肉率和背膘厚相关。Yuan 等^[32]在 *CACNA2D1* 基因第 25、35 外显子处发现 2 个与肉色、屠宰率、背膘厚显著相关的 SNP。 *FSHR* 基因是卵泡刺激素受体基因, 该基因属于 G 蛋白偶联受体家族中的成员, 通过 cAMP 通路介导卵泡刺激素的信号转导。该基因对提高雌性哺乳动物的产仔数具有重要的作用。雷雪芹等^[33]通过研究发现在 *FSHR* 基因的第 10 外显子上秦川和荷

斯坦牛的单胎母牛和双胎母牛有明显差异的突变率。

本研究筛查到的部分候选基因与牛肉质性状相关。*TG*基因又名甲状腺球蛋白基因,是甲状腺素(*T4*)和三碘甲状腺氨酸(*T3*)的前体。研究证明受到人工选择的*TG*基因在脂肪沉积,机体代谢以及脂肪细胞生长分化中有重要作用^[34]。而且*T3*和*T4*与大理石花纹的等级的相关性已经在日本黑牛中得到了证明^[35]。Hayes等^[5]鉴别不同生产类型牛受到人工选择区域,基于9323个SNP利用滑动窗在荷斯坦牛和安格斯牛群体比较了等位基因频率的差异,发现*TG*基因在奶牛和肉牛群体间存在显著差异,且该基因已被证实影响肉牛脂肪和大理石花纹^[11,12]。Barendse等^[36]发现了在*TG*基因上C422T位点与育肥牛的大理石花纹显著相关的;同时,Thaller等^[37]发现该位点也与牛肉的肌内脂肪含量相关;Shin等^[39]发现该位点在韩牛中与大理石花纹的等级评分显著相关;Casas等^[38]在瘤牛群体中发现该位点虽然与大理石花纹的等级无关,但与眼肌面积和背膘厚有关。Hou等^[40]发现在牛的*TG*基因3'侧翼区上存在4个与大理石花纹有关的SNP位点。除此之外,本研究还发现了*GRIK4*、*KCNN2*、*EXOC4*和*CLIC5*等具有重要生物学功能的基因。其中*GRIK4*基因编码的蛋白质属于谷氨酸门控离子通道家族。谷氨酸的作用主要是作为中央神经系统的主要兴奋性神经递质通过配体门控离子通道和G蛋白偶联膜受体。Sturgill等^[41]认为在中枢神经系统中,谷氨酸通过红藻氨酸受体的信号增加B细胞的增殖和免疫球蛋白的产生。该基因富集在离子转运生物学过程类和通道活性、基板的特异通道活性、被动跨膜转运活性分子功能类中。*KCNN2*基因是调节神经细胞的重要基因,且富集在离子转运、金属离子转运等生物学过程类和通道活性、基板的特异通道活性、被动跨膜转运活性分子功能类中。*EXOC4*基因与囊泡介导的转运的生物学过程相关,其作用是将囊泡连接到细胞膜上。Laramie等^[42]研究表明*EXOC4*基因参与胰岛素刺激的葡萄糖的转运,可作为糖尿病的候选基因。*CLIC5*基因是细胞内氯离子通道蛋白家族的成员,在生物体内主要通过转运阴离子参与细胞的体积调节、细胞内pH值的维持、静息电位和兴奋的调节、细胞分泌和激素作用等各种生理过程,有研

究表明*CLIC5*与猪的脂肪性状的QTL相关^[43]。尽管如上所述本研究筛查到很多重要基因,但也有部分受到选择的基因在本研究中没有检测到,如生长激素受体基因(*GHR*)^[17]。这可能一方面是由于生*GHR*基因可能已经在本研究群体中被选择后固定下来,另一方面由于本研究的统计方法和群体大小等因素使得本研究没有筛查到此基因。

3.4 关联分析结果的对照

本研究检测到47104个受到选择的位点,这些位点覆盖了牛基因组上多个区域。利用本研究群体,本研究针对牛肉质性状开展了全基因组关联分析(数据未公布)。通过与全基因组关联分析的结果比较发现,本研究筛查到可能受到选择的*GRIK4*与*RNLS*、*NCAMI*、*SPONI*和*MTSSI*等基因与牛肉剪切力性状显著相关,*RAPGEF4*、*BMPER*、*KCNN2*和*EXOC4*等基因和牛肉大理石花纹显著相关。

本研究第一次在全基因水平对我国肉牛进行选择信号筛查研究,筛查到多个可能受到选择且与重要经济性质相关的基因,但是这些基因仍需要进一步的验证。而且随着芯片成本的降低和更高密度的SNP芯片的出现,通过扩大群体样本和采用多种统计检测方法,将会更准确的检测到受到人工选择的基因和区域。

参考文献(References):

- [1] Bradley DG, Cunningham EP. Genetic aspects of domestication. In: Fries R, Ruvinsky A, eds. The Genetics of Cattle. Wallingford: CABI Publishing, 1999: 15–31. DOI
- [2] Thornton KR, Jensen JD. Controlling the false-positive rate in multilocus genome scans for selection. *Genetics*, 2007, 175(2): 737–750. DOI
- [3] Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. *Evolution*, 1984, 38(6): 1358–1370. DOI
- [4] Akey JM, Zhang G, Zhang K, Jin L, Shriver MD. Interrogating a high-density SNP map for signatures of natural selection. *Genome Res*, 2002, 12(12): 1805–1814. DOI
- [5] Hayes BJ, Chamberlain AJ, Maceachern S, Savin K, McPartlan H, MacLeod I, Sethuraman L, Goddard ME. A genome map of divergent artificial selection between *Bos taurus* dairy cattle and *Bos taurus* beef cattle. *Anim Genet*, 2009, 40(2): 176–184. DOI

- [6] Pei YF, Li J, Zhang L, Papasian CJ, Deng HW, Heutink P. Analyses and comparison of accuracy of different genotype imputation methods. *PLoS One*, 2008, 3(10): e3551. [DOI](#)
- [7] Weir BS. Population substructure: Genetic data analysis II. Sunderland, MA: Sinauer Associates, 1996: 161–173. [DOI](#)
- [8] McGill R, Tukey JW, Larsen WA. Variations of box plots. *Am Stat*, 1978, 32(1): 12–16. [DOI](#)
- [9] Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*, 2009, 37(1): 1–13. [DOI](#)
- [10] Qanbari S, Pimentel EC, Tetens J, Thaller G, Lichtner P, Sharifi AR, Simianer H. A genome-wide scan for signatures of recent selection in Holstein cattle. *Anim Genet*, 2010, 41(4): 377–389. [DOI](#)
- [11] Barandse WJ. Assessing lipid metabolism. Patent, International Publication Number: WO 99/23248. World International Property Organization, 1999. [DOI](#)
- [12] Baud S, Donaldson N. The TG5 thyroglobulin gene test for amarbling quantitative trait loci evaluated in feedlot cattle. *Aust J Exp Agric*, 2004, 44(7): 669–674. [DOI](#)
- [13] Cohen-Zinder M, Seroussi E, Larkin DM, Loor JJ, van der Wind-Everts A, Lee JH, Drackley JK, Band MR, Hernandez AG, Shani M, Lewin HA, Weller JI, Ron M. Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. *Genome Res*, 2005, 15(7): 936–944. [DOI](#)
- [14] MacEachern S, Hayes B, McEwan J, Goddard M. An examination of positive selection and changing effective population size in Angus and Holstein cattle populations (*Bos taurus*) using a high density SNP genotyping platform and the contribution of ancient polymorphism to genomic diversity in domestic cattle. *BMC Genomics*, 2009, 10(1): 181. [DOI](#)
- [15] Prasad A, Schnabel RD, McKay SD, Murdoch B, Stothard P, Kolbehdari D, Wang Z, Taylor JF, Moore SS. Linkage disequilibrium and signatures of selection on chromosomes 19 and 29 in beef and dairy cattle. *Anim Genet*, 2008, 39(6): 597–605. [DOI](#)
- [16] Qanbari S, Pimentel EC, Tetens J, Thaller G, Lichtner P, Sharifi AR, Simianer H. A genome-wide scan for signatures of recent selection in Holsteincattle. *Anim Genet*, 2010, 41(4): 377–389. [DOI](#)
- [17] Flori L, Fritz S, Jaffrézic F, Boussaha M, Gut I, Heath S, Foulley JL, Gautier M. The genome response to artificial selection: a case study in dairy cattle. *PLoS One*, 2009, 4(8): e6595. [DOI](#)
- [18] Qanbari S, Gianola D, Hayes B, Schenkel F, Miller S, Moore S, Thaller G, Simianer H. Application of site and haplotype-frequency based approaches for detecting selection signatures in cattle. *BMC Genomics*, 2011, 12(1): 318. [DOI](#)
- [19] Black WC 4th, Baer CF, Antolin MF, DuTeau NM. Population genomics: genome-wide sampling of insect populations. *Annu Rev Entomol*, 2001, 46(4): 441–469. [DOI](#)
- [20] Cavalli-Sforza LL. Population structure and human evolution. *Proc R Soc Lond B Biol Sci*, 1966, 164(995): 362–379. [DOI](#)
- [21] Lewontin RC, Krakauer J. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, 1973, 74(1): 175–195. [DOI](#)
- [22] Beaumont MA, Balding DJ. Identifying adaptive genetic divergence among populations from genome scans. *Mol Ecol*, 2004, 13(4): 969–980. [DOI](#)
- [23] Wang ET, Kodama G, Baldi P, Moyzis RK. Global landscape of recent inferred Darwinian selection for *Homo sapiens*. *Proc Natl Acad Sci USA*, 2006, 103(1): 135–140. [DOI](#)
- [24] Voight BF, Kudaravalli S, Wen X, Pritchard JK. A map of recent positive selection in the human genome. *PLoS Biol*, 2006, 4(3): e72. [DOI](#)
- [25] Akey JM, Eberle MA, Rieder MJ, Carlson CS, Shriver MD, Nickerson DA, Kruglyak L. Population history and natural selection shape patterns of genetic variation in 132 genes. *PLoS Biol*, 2004, 2(10): e286. [DOI](#)
- [26] Akey JM, Zhang G, Zhang K, Jin L, Shriver MD. Interrogating a high-density SNP map for signatures of natural selection. *Genome Res*, 2002, 12(12): 1805–1814. [DOI](#)
- [27] Hoaglin DC, Mosteller F, Tukey JW. Understanding Robust and Exploratory Data Analysis. New York: John Wiley & Sons, 1983. [DOI](#)
- [28] Wayne WA, Stephan SR. Testing for outliers from a mixture distribution when some data are missing. *Comput Stat Data Anal*, 2003, 44(1–2): 193–210. [DOI](#)
- [29] Lohnes D, Kastner P, Dierich A, Mark M, LeMeur M, Chambon P. Function of retinoic acid receptor Y in the mouse. *Cell*, 1993, 73(4): 643–658. [DOI](#)
- [30] 黄萌, 许尚忠, 咎林森, 张路培, 高雪, 陈金宝. 牛RXRG基因遗传变异与双胎性状的关联分析. *遗传学报*, 2008, 30(2): 190–194. [DOI](#)
- [31] Hou GY, Yuan ZR, Gao X, Li JY, Gao HJ, Chen JB, Xu SZ. Genetic polymorphisms of the CACNA2D1 gene and their

- association with carcass and meat quality traits in cattle. *Biochem Genet*, 2010, 48(9-10): 751-759. [DOI](#)
- [32] Yuan ZR, Xu SZ. Novel SNPs of the bovine CACNA2D1 gene and their association with carcass and meat quality traits. *Mol Biol Rep*, 2011, 38(1): 365-370. [DOI](#)
- [33] 雷雪芹, 陈宏, 袁志发, 雷初朝, 徐廷生, 孙维斌. 促卵泡生成素受体基因的SNP对牛双胎性状的标记研究. *云南畜牧兽医*, 2002, (4): 28-29. [DOI](#)
- [34] Cases S, Zhou P, Schillingford JM, Wiseman BS, Fish JD, Angle CS, Hennighausen L, Werb Z, Farese RV Jr. Development of the mammary gland requires DGAT1 expression in stromal and epithelial tissues. *Development*, 2004, 131(13): 3047-3055. [DOI](#)
- [35] Ailhaud G, Grimaldi P, Négrel R. Cellular and molecular aspects of adipose tissue development. *Annu Rev Nutri*, 1992, 12(7): 207-233. [DOI](#)
- [36] Barendse WJ. Assessing lipid metabolism. International Patent Application: PCT 98/00882, 1999. [DOI](#)
- [37] Thaller G, Kuhn C, Winter A, Ewald G, Bellmann O, Wegner J, Zühlke H, Fries R. DGAT1, a new positional and functional candidate gene for intramuscular fat deposition in cattle. *Anim Genet*, 2003, 34(5): 354-357. [DOI](#)
- [38] Casas E, White SN, Riley DG, Smith TPL, Brennemant RA, Olson TA, Johnson DD, Coleman SW, Bennett GL, Chase CC Jr. Assessment of single nucleotide polymorphisms in genes residing on chromosomes 14 and 29 for association with carcass composition traits in *Bos indicus* cattle. *J Anim Sci*, 2005, 83(1): 13-19. [DOI](#)
- [39] Shin SC, Chung ER. Association of SNP marker in the thyroglobulin gene with carcass and meat quality traits in Korean cattle. *Asian-Australas J Anim Sci*, 2007, 20(2): 172-177. [DOI](#)
- [40] Hou GY, Yuan ZR, Zhou HL, Zhang LP, Li JY, Gao X, Wang DJ, Gao HJ, Xu SZ. Association of thyroglobulin gene variants with carcass and meat quality traits in beef cattle. *Mol Biol Rep*, 2011, 38(7): 4705-4708. [DOI](#)
- [41] Sturgill JL, Mathews J, Scherle P, Conrad DH. Glutamate signaling through the kainate receptor enhances human immunoglobulin production. *J Neuroimmunol*, 2011, 233(1-2): 80-89. [DOI](#)
- [42] Laramie JM, Wilk JB, Williamson SL, Nagle MW, Latourelle JC, Tobin JE, Province MA, Borecki IB, Myers RH. Polymorphisms near EXOC4 and LRGUK on chromosome 7q32 are associated with Type 2 Diabetes and fasting glucose. *BMC Med Genet*, 2008, 9: 46. [DOI](#)
- [43] 张海燕, 尹靖东, 王强. CLIC5 等位基因与PIC猪胴体品质和肉质的关系. 中国畜牧兽医学会 2009 学术年会论文集(下册), 2009. [DOI](#)

•综合信息•

关于申报首届“谈家桢遗传教育奖”的通知

中国遗传学会 2011 年 8 月 10 日在新疆医科大学召开的八届五次常务理事会上, 审议通过了“谈家桢遗传教育奖奖励章程”。学会定于 2012 年 9 月 25 日至 2013 年 5 月 31 日受理首届谈家桢遗传教育奖申请和推荐。申请书、推荐书和奖项章程请登陆学会网站(<http://www.geneticsociety.cn>)中“表彰奖励”栏目下载。欢迎符合申请和推荐条件的从事遗传学教学的科技工作者积极申请该奖项。

申请者请将申请书或推荐书纸质版材料一式 2 份于 2013 年 5 月 31 日前邮寄到中国遗传学会办公室, 同时将电子版发至 geneticsociety@163.com。

邮编: 100101

地址: 北京市朝阳区北辰西路一号院 2 号

联系电话: (010)64806635; (010)64806636

联系人: 王长城