

引用格式:

李梓锋, 殷利夺, 刘德武. 牛和山羊基因组共同加速进化区域与角形成的潜在关系[J]. 湖南农业大学学报(自然科学版), 2022, 48(5): 578–584.

LI Z F, YIN L D, LIU D W. The potential relationship between co-accelerated evolution regions in the genomes of *Bos taurus* and *Capra hircus* and horn formation[J]. Journal of Hunan Agricultural University(Natural Sciences), 2022, 48(5): 578–584.

投稿网址: <http://xb.hunau.edu.cn>



牛和山羊基因组共同加速进化区域与角形成的潜在关系

李梓锋¹, 殷利夺², 刘德武^{1*}

(1.华南农业大学动物科学学院, 广东 广州 510642; 2.中国科学院昆明动物研究所, 云南 昆明 650223)

摘 要: 对 14 个物种的全基因组序列进行比对, 识别其基因组中的同源区域, 进而鉴定牛和山羊基因组的共同加速进化区域(以下简称加速区域), 并对加速区域进行注释及组蛋白修饰位点的富集分析; 运用生物信息学方法, 筛选编码区出现加速进化的基因, 并对其进行 GO 和 KEGG 通路分析。结果表明: 在牛和山羊基因组中共检测到 44 794 个加速区域; 加速区域在基因区与非基因区均有分布, 分别占总数的 54.80%与 45.20%; 加速区域显著富集了 25 个不同的组蛋白标记; 鉴定出 2703 个候选基因的编码区出现了加速区域; GO 条目分析发现, 候选基因主要富集的生物过程包括轴突形成、腺体发育、肌肉组织发育等, 细胞组成包括突触膜、轴突部分、突触后致密等; KEGG 通路分析发现, 这些候选基因参与了 cAMP 信号通路、轴突导向、钙离子信号通路、Rap1 信号通路、神经活性配体-受体互作等信号通路。这揭示在牛角形成过程中, 突触和轴突的产生可能发生了独特的变化, 影响信号传递与神经结构组成并导致转录调控和基因表达发生变化, 从而导致牛科动物出现形态特异的洞角。

关 键 词: 牛; 山羊; 同源区域; 共同加速进化区域; 角; 突触; 轴突

中图分类号: S823.2; S827.2 文献标志码: A 文章编号: 1007-1032(2022)05-0578-07

The potential relationship between co-accelerated evolution regions in the genomes of *Bos taurus* and *Capra hircus* and horn formation

LI Zifeng¹, YIN Liduo², LIU Dewu^{1*}

(1.College of Animal Science, South China Agricultural University, Guangzhou, Guangdong 510642, China; 2.Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China)

Abstract: To identify co-accelerated evolution regions(accelerated regions for short) of *Bos taurus* and *Capra hircus* at whole genome-wide, and reveal the potential molecular mechanism under the evolution of the unique phenotype of bovids, we performed homologous genomic regions of 14 species the whole genome sequencing and bioinformatics analysis of genes with accelerated evolution events in coding regions, the enriched GO term, KEGG pathways and histone modification regions. The results showed that a total of 44 794 accelerated regions were detected in genome of *Bos taurus* and *Capra hircus*. These acceleration regions were distributed in both the gene region and the non-gene region, accounting for 54.80 % and 45.20 % of the total, respectively. In addition, accelerated regions were significantly enriched in 25 active sites of histone. A total of 2703 candidate genes occurred accelerated evolution events in coding regions. Gene Ontology analysis indicated that these candidate genes were mainly involved in biological processes like axon genesis, gland development, muscle tissue development, cellular components like synaptic membrane, axon membrane, axon part, post-synaptic density. KEGG pathways showed that these candidate genes were mainly involved in cAMP signaling pathways, axon guidance signaling pathways, calcium signaling pathway and Neuroactive

收稿日期: 2021-03-27

修回日期: 2022-09-30

基金项目: 第二次青藏高原综合科学考察研究任务五动物专题(2019QZKK0501)

作者简介: 李梓锋(1995—), 男, 广东广州人, 硕士研究生, 主要从事动物遗传育种与繁殖研究, 837934075@qq.com; *通信作者, 刘德武, 博士, 教授, 主要从事分子遗传与动物育种、动物健康养殖与安全生产研究, dwliu@scau.edu.cn

Ligand-Receptor Interaction. It indicated that during the formation of the horn in bovids, unique changes occur in the production of synapses and axons, which affects signal transmission and leads to the changes in transcription regulation and gene expression, resulting to the appearance of morphologically specific horns in bovids.

Keywords: *Bos taurus*; *Capra hircus*; homologous region; co-accelerated evolution region; horn; synapses; axons

反刍动物是陆生草食动物中一个重要的群体, 包含长颈鹿科、叉角羚科、麋鹿科、麝科、鹿科、牛科^[1], 包括至少 200 个现存物种。其中, 牛科中现存物种最多, 至少有 143 种, 包括重要的家养动物(牛和山羊)。反刍动物具有特异的解剖学特征, 如多腔胃、高冠齿及角(颅附外骨)^[2-4]。这些特征被认为是反刍动物在物种多样性、地域丰富度和地理分布范围方面进化成功的原因。尽管反刍动物具有生物学上的突出地位及对人类文明的重要价值, 但关于反刍亚目动物, 尤其是牛科动物的了解仍然很有限。羊亚科中, 家山羊与其他物种在全基因组水平的分歧度相当低, 从 0.21%(野山羊与家山羊)到 2.18%(绵羊与家山羊)^[5]。笔者选取牛科中牛亚科牛属的普通牛(*Bos taurus*)、羊亚科山羊属的家山羊(*Capra hircus*) 2 个物种作为牛科动物中的代表物种进行研究。

确定特定的基因组区域与特殊表型的形成是否存在联系, 是一项具有挑战性的工作, 其研究方法缺乏普适性。其中较为优越的策略之一是通过比较基因组学的方法来识别功能序列^[6-8]。由于不同物种基因组序列间存在进化上同源保守的区域, 这部分保守同源区域与动物体的生长发育相关^[9-11]。在这些保守同源序列中出现相对较高的序列替换率, 会造成物种特异的形态变化, 如在人类基因组的研究中发现加速进化区域都是高度保守的^[9,12], 至少 30%的人类基因组加速区域位于负责生长发育的基因与调控元件上, 其中在 1 个 RNA 基因 *HAR1F* 上鉴定出加速区域, 该基因在人类新皮层发育的 7~19 个妊娠周中的 Cajal-Retzius 神经元中特异表达(皮质神经元迁移的关键时期), 此阶段的发育特征被认为是造成人类与黑猩猩形态特征差异的原因之一^[13]; 因此, 基于进化上保守同源的区域, 锚定其中进化速率较快的区域, 既可用于识别在编码蛋白基因区域的信号, 也可识别非编码区域中的信号^[14-15], 从而对物种的表型变异进行分子层面的解析。

本研究中, 通过比较基因组分析, 在全基因组

范围内检测反刍亚目牛科动物牛、山羊共同的加速进化信号, 探索共同加速进化区域(以下简称加速区域)在全基因组中的分布, 找到出现加速进化信号的功能区域, 揭示角的形成机制, 旨在为探究反刍亚目牛科的特异表型形成机理提供依据。

1 材料与方法

1.1 材料

14 种具有代表性的哺乳动物: 人(*Homo sapiens*, GRCh37/hg19), 恒河猴(*Macaca mulatta*, MGSC Merged 1.0/rheMac3), 树鼩(*Tupaia belangeri*, Broad Institute tupBel1(NCBI project 13971, AAPY010000 00)), 小鼠(*Mus musculus*, NCBI37/mm10), 豚鼠(*Cavia porcellus*, Broad/cavPor3), 兔(*Oryctolagus cuniculus*, Broad/oryCun2), 牛(*Bos taurus*, Bos_taurus_UMD_3.1/bosTau7), 家山羊(*Capra hircus*), 马(*Equus caballus*, Broad/equCab2), 狗(*Canis lupus familiaris*, Broad/canFam3), 鼯鼠(*Sorex araneus*, Broad/sorAra2), 象(*Loxodonta africana*, Broad/loxAfr3), 马岛猬(*Echinops telfairi*, Broad/echTel2), 负鼠(*Monodelphis domestica*, Broad/monDom5)。从 NCBI 网站下载山羊(*Capra hircus*)基因组序列, 从 UCSC 基因组浏览器网站下载其他物种基因组序列和以人类基因组为参考序列的序列比对结果。

1.2 基因组处理与序列比对

运用 UCSC Kent 公用程序(http://hgdownload.soe.ucsc.edu/admin/exe/linux.x86_64/)对山羊基因组数据进行预处理, 计算基因组大小并进行格式转换; 以人类基因组的各常染色体为目标序列, 运用 lastz (v1.04.03)^[16]将山羊基因组序列分别与人类基因组各常染色体序列进行跨物种的两两比对; 运用 axtChain(v302.0.0)、chainNet(v302.0.0)和 axtToMaf (v302.0.0)等进行 gap 过滤和格式转换, 得到 maf 格式的两两比对结果文件; 使用 Multiz(v11.2)^[17]获得以人类常染色体为参考的多物种基因组序列比对结果。

1.3 保守同源区域与加速区域识别

基于 UCSC 的 hg19 基因组注释,以 1.2 的比对结果作为输入文件,采用 msa_view(PHAST v1.4)与 phyloFit(PHAST v1.4)^[18]获取 4 倍简并位点,并估计中性进化模型;运用 PHAST 软件包^[18]并根据多序列比对结果及中性进化速率模型,识别保守同源元件,进一步检测牛和山羊的基因组中的加速区域。

1.4 加速区域的基因组分布分析

人类基因组的功能研究较为丰富,其基因组注释(包括功能元件的注释)均较为齐全^[19],是其他非模式生物难以企及的;因此,将加速区域映射到人类基因组同源区域进行后面的注释和分析。采用 R 包 ChIPseeker^[20]对加速区域进行注释,将加速区域落在基因组中的位置,如启动子、内含子、外显子等区域注释;根据 VISTA 增强子浏览器中已经通过实验验证的增强子注释,查看加速区域在增强子区域的分布^[21]。

1.5 加速区域在基因组组蛋白活性位点的富集分析

运用 ChIP-Atlas 数据库,识别组织特异的调控元件^[19]。检测加速区域是否位于人类基因组的候选功能元件,并描述这些元件的生物化学活性;使用 ChIP-Atlas 中公开可用的 ChIP-Seq 数据集^[19];运用 silico ChIP 并规定 100 为默认显著性阈值,将加速区域的 bed 格式文件在人类不同组织细胞中进行组蛋白位点富集分析。

1.6 加速区域分配到邻近的基因及其功能富集分析

运用 GREAT^[22]基于 hg19 注释将加速区域分配到附近的基因,使用基础设置加上 1000 kb 的扩展默认设置,基于 R 包 ClusterProfiler^[23]内置数据库,将编码区出现加速进化的基因分别进行 GO 条目富集分析与 KEGG 通路富集分析,采用超几何分布计算出基因显著富集的 GO 条目和 KEGG 通路。

1.7 加速区域与基因统计

统计基因邻近的加速区域数,展示富集加速区域最多的前 10 个基因、加速区域数及前 10 个基因涉及的 GO 分析结果。

2 结果与分析

2.1 序列比对结果与系统发育关系

根据 14 个物种的系统发育关系(图 1),以有袋类哺乳动物的负鼠作为外群,并将羊膜哺乳动物的灵长目、啮齿目、树鼩目、偶蹄目、奇蹄目、食虫目、长鼻目和猬形目中具代表性的物种进行全基因组序列比对,得到以人类(hg19)为目标序列的共 14 个物种的基因组比对结果。

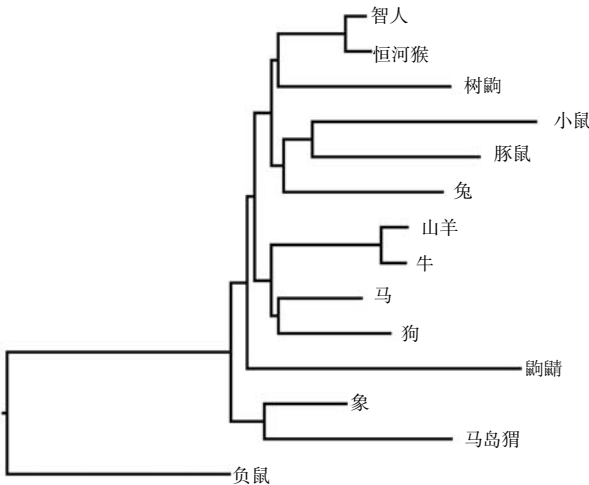


图 1 基于 4 倍简并位点构建的系统发育树

Fig.1 Phylogenetic tree based on the 4-fold degenerate sites

2.2 保守同源区域鉴定与加速区域检测结果

对多物种基因组序列比对结果进行保守同源性分析,发现 12 105 596 个哺乳动物保守同源的区域分布于常染色体上(表 1)。基于似然比检验的算法^[18],对获得的同源区域进行加速进化检测,通过与中性模型比较,识别在牛和山羊序列中具有核苷酸替代

表 1 保守区域与加速区域在染色体上的分布

Table 1 The number distribution of conserved regions and accelerated regions on chromosomes

染色体	保守区域数	加速区域数	染色体	保守区域数	加速区域数	染色体	保守区域数	加速区域数
Chr1	679 415	5279	Chr9	649 200	980	Chr17	531 675	1130
Chr2	653 289	6552	Chr10	656 754	2042	Chr18	485 307	503
Chr3	656 002	5030	Chr11	635 550	2318	Chr19	266 322	255
Chr4	618 133	5930	Chr12	664 194	1550	Chr20	419 118	243
Chr5	614 380	2966	Chr13	561 958	594	Chr21	198 283	94
Chr6	651 725	3576	Chr14	576 900	367	Chr22	225 503	297
Chr7	633 752	2268	Chr15	561 821	139	总计	12 105 596	44 794
Chr8	641 137	2116	Chr16	525 178	565			

率共同增加(FDR<1%)的区域,各常染色体均有加速区域分布,共检测到 44 794 个加速区域。

2.3 加速区域注释结果

对加速区域的注释发现,加速区域在基因区(包括非翻译区、外显子区与内含子区)与非基因区(包括启动子区、远端基因间区和下游区域)均有分布,分别占总数的 54.80%与 45.20%(表 2)。1957 个加速区域分布在外显子区域(占总加速区域的 4.37%)中,有 636 个加速区域分布在第一个外显子区域(占 1.42%);有 21 890 个加速区域分布在内含子区域(占

表 2 牛和山羊基因组加速进化区域的注释结果

Table2 Annotation results of the accelerated regions of genomes of cow and goat		
基因组区域	加速区域占比/%	加速区域数
启动子	3.03	1356
5'端非翻译区	0.20	89
3'端非翻译区	1.36	612
第一个外显子	1.42	636
其他外显子	2.95	1321
第一个内含子	12.69	5684
其他内含子	36.18	16 206
远端基因间区	0.98	439
下游区域	41.19	18 451

48.87%),其中,有 5684 个加速区域分布在第一个内含子区域(占 12.69%);有 439 个加速区域分布在远端基因间区(占 0.98%);有 89 个加速区域分布在 5'端非翻译区(占 0.20%);有 612 个加速区域分布在 3'端非翻译区域(占 1.36%);有 1356 个加速区域分布在启动子区域(占 3.03%)。另外,通过 VISTA 增强子实验验证数据集对加速区域的注释发现,有 108 个加速区域分布在增强子区域。

2.4 加速区域在组蛋白标记的富集分析结果

将加速区域映射到人类基因组中,通过组蛋白标记区域富集分析发现,加速区域显著富集了 25 个不同的组蛋白标记,图 2 展示的是其中 8 个组蛋白标记的结果(相对于随机抽样的加速区域 FDR<5%),如组蛋白变体(H2A.Z)、增强子活性位点(H3K27ac 和 H3K4me1)、启动子活性位点(H3K4me3 和 H3K4me2)等。从图 2 可知,加速区域在肾脏细胞、血液细胞、消化道细胞、骨细胞中的 H2A.Z 标记的富集倍数高达 3~4 倍;加速区域在神经细胞、上皮细胞、骨细胞中的 H3K27ac 与 H3K4me1 标记的富集倍数达 2~3 倍。

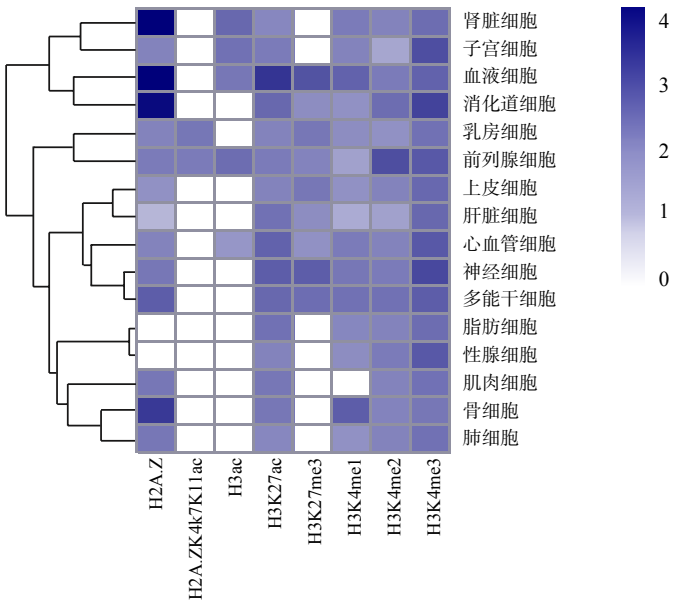


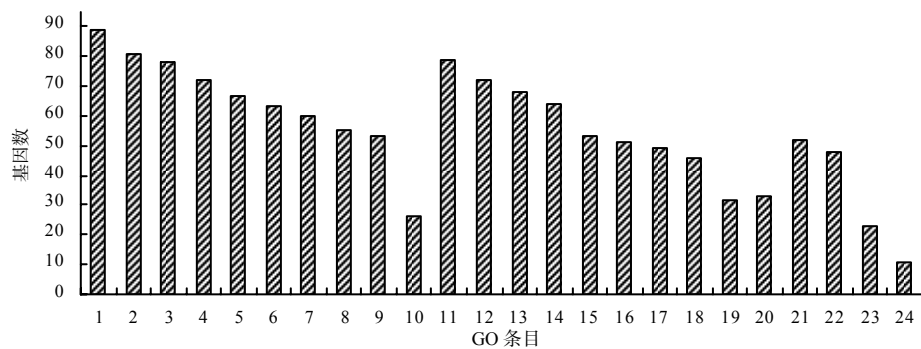
图 2 加速区域在不同组织细胞中的组蛋白标记显著富集情况(FDR<5%)

Fig.2 The significantly enrichment results of accelerated regions in histone markers of different types cells(FDR<5%)

2.5 候选基因的 GO 条目富集分析结果

运用 Great 基于 hg19 注释到了 2703 个编码区加速进化的基因。运用 GO 数据库,对 2703 个候选基因进行功能富集分析,结果如图 3 所示。共有 24 个 GO 条目显著(P<0.05)富集,其中生物过程包

括轴突形成、腺体发育、肌肉组织发育和上皮细胞管形态发生等;细胞组成包括突触膜、轴突部分、突触后致密和非对称性突触等;分子功能包括金属离子跨膜转运体活性、磷酸酯水解酶活性、钙调蛋白结合等。



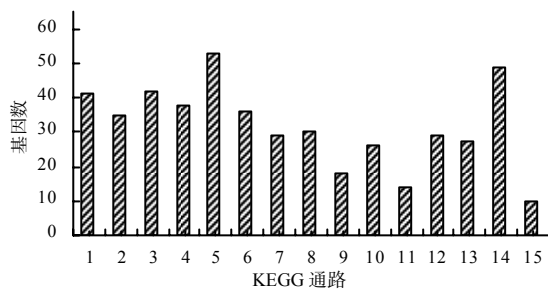
1~10 为生物过程(1 轴突形成, 2 腺体发育, 3 肌肉组织发育, 4 胚胎器官发育, 5 横纹肌发育, 6 上皮细胞管形态发生, 7 胚胎器官形态发生, 8 感觉器官形态发生, 9 心脏形态发生, 10 细胞命运决定); 11~20 为细胞组成(11 突触膜, 12 轴突部分, 13 突触后致密, 14 非对称性突触, 15 收缩纤维, 16 肌原纤维, 17 收缩纤维部分, 18 肌原纤维节, 19 I 带, 20 Z 线); 21~24 为分子功能(21 金属离子跨膜转运体活性, 22 磷酸酯水解酶活性, 23 钙调蛋白结合, 24 γ -氨基丁酸跨膜转运活性)。

图 3 候选基因的 GO 条目富集分析结果

Fig.3 GO term enrichment analysis results of candidate genes

2.6 候选基因的 KEGG 通路富集分析结果

基于 KEGG 数据库,KEGG 富集分析结果中前 15 个 KEGG 富集通路如图 4 所示,包括 cAMP 信



1 cAMP 信号通路; 2 轴突导向; 3 钙离子信号通路; 4 Rap1 信号通路; 5 神经活性配体-受体互作; 6 黏着斑; 7 胃癌; 8 Hippo 信号通路; 9 致心律失常性右室心肌病; 10 信号通路调控干细胞多能性; 11 其他类型的 O-聚糖生物合成; 12 Wnt 信号通路; 13 细胞粘连分子; 14 人乳头状瘤病毒感染; 15 黏蛋白 O-聚糖生物合成。

图 4 候选基因的 KEGG 通路富集分析结果

Fig.4 KEGG pathway enrichment analysis results of candidate genes

号通路、轴突导向、钙离子信号通路、Rap1 信号通路、神经活性配体-受体互作等信号通路显著 ($P<0.05$)富集。

2.7 加速进化基因与 GO 条目联合分析结果

经统计发现,部分基因附近检测出较多的加速区域。表 3 展示了加速区域总数排名前 10 的基因。其中, *PALLD*、*OPCML*、*NTM* 基因在神经元投射导向、神经元迁移、轴突形成、神经元识别的生物过程等 GO 条目中显著富集(表 4)。另外,在 *NTM* 基因附近区域检测到了 143 个加速区域(图 5),该基因编码 IgLON 家族中含有免疫球蛋白(Ig)结构域的糖基磷脂酰肌醇(GPI)锚定的细胞黏附分子。

表 3 基因加速区域总数排名前 10 的结果

Table 3 Top ten results for the total number of accelerated regions in genes

基因 ID	基因名称	功能注释	加速区域数
ENSG00000145439	<i>CBR4</i>	羧基还原酶 4	581
ENSG00000129116	<i>PALLD</i>	含钼蛋白, 细胞骨架相关蛋白	390
ENSG00000154447	<i>SH3RF1</i>	SH3 域环指蛋白 1	227
ENSG00000183715	<i>OPCML</i>	阿片类结合蛋白/细胞黏附分子样	204
ENSG00000109586	<i>GALNT7</i>	多肽 N-乙酰氨基半乳糖转移酶 7	185
ENSG00000174473	<i>GALNTL6</i>	多肽 N-乙酰氨基半乳糖转移酶样 6	179
ENSG00000205129	<i>C4orf27</i>	染色体 4 开放阅读框 47	158
ENSG00000108018	<i>SORCS1</i>	分拣蛋白相关 VPS10 结构域受体 1	143
ENSG00000182667	<i>NTM</i>	Neurotrimin	143
ENSG00000145416	<i>MARCH1</i>	膜相关环指蛋白 1	135

表 4 加速区域总数排名前 10 的基因涉及的 GO 条目

Table 4 The GO terms involved in the top ten genes in the total accelerated regions				
GO ID	GO 条目	基因数	分类	<i>P</i>
GO:0097485	神经元凸起导向	60	生物过程	1.76×10^{-8}
GO:0007411	轴突导向	59	生物过程	3.31×10^{-8}
GO:0001764	神经元迁移	41	生物过程	3.79×10^{-8}
GO:0007409	轴突新生	83	生物过程	5.42×10^{-8}
GO:0008038	神经元识别	11	生物过程	3.06×10^{-2}
GO:0033267	轴突部分	59	细胞组成	3.45×10^{-4}
GO:0030018	Z 板	29	细胞组成	1.36×10^{-4}
GO:0031674	I 带	30	细胞组成	2.31×10^{-4}
GO:0004653	多肽 N-乙酰氨基半乳糖转移酶活性	9	分子功能	3.19×10^{-3}
GO:0008376	乙酰氨基半乳糖转移酶活性	14	分子功能	6.86×10^{-3}

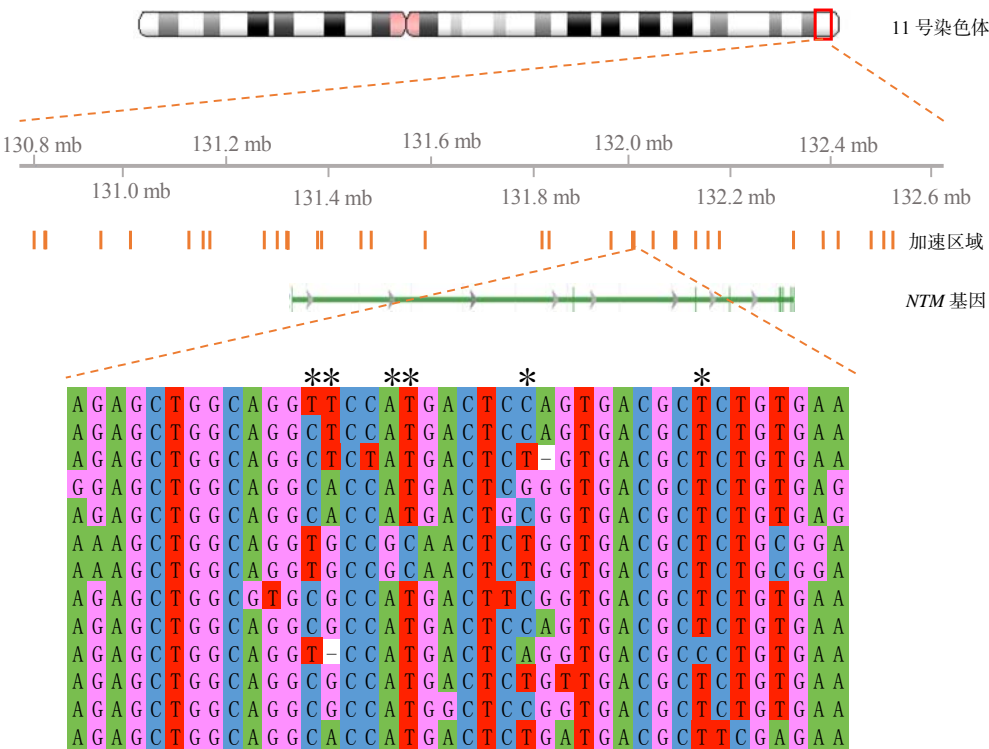


图 5 NTM 基因周围的加速区域分布情况与其中 1 个加速区域

Fig.5 The distribution of accelerated regions around NTM gene and an example of accelerated regions

3 结论与讨论

以往的研究^[24]表明，与下游的其他内含子相比，人类基因组大部分转录本的第一个内含子富集于受进化选择的主动转录调控信号，如 H3K4me1，H3K4me3 增强子活性标记，这些信号对调控复杂的基因表达模式尤为重要。本研究中，对反刍亚目牛科的牛、山羊 2 个物种全基因组深入研究发现，有 5684 个加速区域位于基因组中的第一个内含子，可能使得该区域增强子结合活性改变，从而在转录调控水平发生改变，影响相关表型的形成，并且在非

翻译区等其他调控元件区域也发现加速进化的信号，需要后续进一步的研究验证其功能有何具体影响。

组蛋白变体 H2A.Z 的核小体比常规核小体展开程度更高^[25]。加速区域在染色体活性区域有显著高的富集倍数，包括在增强子活性区域，启动子活性区域及组蛋白变体 H2A.Z，表明绝大多数加速区域具有生物学功能，参与多种调控活动，但具体的调控活性功能有待进一步的验证。

在反刍亚目动物关于角起源的研究^[5]中，对牛科洞角和鹿角共同高表达基因的研究结果表明，这 2 种颅附外骨的发育都依赖于相似的基因表达谱，

主要来自神经、骨骼和皮肤组织;GO 功能富集分析表明,这些共同高表达基因在神经嵴细胞迁移、发育的生物过程显著富集;KEGG 通路分析发现,共同高表达基因在神经嵴分化通路、轴突导向等信号通路显著富集,认为反刍亚目动物的角可能在神经嵴干细胞中具有相同起源。本研究中,对 2703 个编码区加速进化的候选基因进行生物学功能富集,GO 结果表明,候选基因在轴突形成、上皮细胞管形态发生等生物过程显著富集,在突触膜、轴突部分、突触后致密、非对称性突触等神经元的组成结构显著富集;KEGG 通路分析发现,候选基因在轴突导向通路显著富集。表明可能从神经嵴干细胞分化产生颅面软骨和颅面骨时,伴随着神经元的分化,在突触结构和轴突形成过程出现特异变化,在神经调控通路发生牛科动物特异的信号传导等变化,从而导致牛科动物洞角的形成^[26]。

有研究^[7-8,27]表明,加速区域不等比例的分布在基因组的各个区域,这就存在部分加速区域会在某些调控区域范围集中形成一个热点。本研究中,通过对加速区域分布的统计发现,加速区域分布最多的基因,如 *PALLD*、*OPCML*、*NTM* 基因在神经结构的发育和组成相关的 GO 条目中显著富集,其中, *NTM* 基因周围存在 143 个加速区域。相关研究^[5,28]表明, *NTM* 基因在山羊的角形成过程中特异高表达,且该基因可促进神经突的生长和黏附,提示可能在牛和山羊基因组中这些基因出现了加速进化,改变了周围调控区域,从而促使该基因在角形成过程中表达量升高,导致牛科动物的颅附外骨进化发育成洞角。

参考文献:

- [1] KER D F E, YANG Y P. Ruminants: Evolutionary past and future impact[J]. Science, 2019, 364: 1130–1131.
- [2] JANIS C. The evolutionary strategy of the Equidae and the origins of rumen and cecal digestion[J]. Evolution, 1976, 30(4): 757–774.
- [3] CHEN L, QIU Q, JIANG Y, et al. Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits[J]. Science, 2019, 364(6446): eaav6202.
- [4] KIM T H, SHIVDASANI R A. Stomach development, stem cells and disease[J]. Development, 2016, 143(4): 554–565.
- [5] LI R, FU W, SU R, et al. Towards the complete goat pan-genome by recovering missing genomic from the reference genome[J]. Frontiers in Genetics, 2019, 10: 1169.
- [6] BOYD J L, SKOVE S L, ROUANET J P, et al. Human-chimpanzee differences in a FZD8 enhancer alter cell-cycle dynamics in the developing neocortex[J]. Current Biology, 2015, 25(6): 772–779.
- [7] FRANCHINI L F, POLLARD K S. Human evolution: the non-coding revolution[J]. BMC Biology, 2017, 15(1): 89.
- [8] HOLLOWAY A K, BRUNEAU B G, SUKONNIK T, et al. Accelerated evolution of enhancer hotspots in the mammal ancestor[J]. Molecular Biology and Evolution, 2016, 33(4): 1008–1018.
- [9] HUBISZ M J, POLLARD K S. Exploring the genesis and functions of Human Accelerated Regions sheds light on their role in human evolution[J]. Current Opinion in Genetics & Development, 2014, 29: 15–21.
- [10] LOOTS G G, LOCKSLEY R M, BLANKESPOOR C M, et al. Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons[J]. Science, 2000, 288(5463): 136–140.
- [11] BOFFELLI D, MCAULIFFE J, OVCHARENKO D, et al. Phylogenetic shadowing of primate sequences to find functional regions of the human genome[J]. Science, 2003, 299(5611): 1391–1394.
- [12] PRABHAKAR S, NOONAN J P, PÄÄBO S, et al. Accelerated evolution of conserved noncoding sequences in humans[J]. Science, 2006, 314(5800): 786.
- [13] CAPRA J A, ERWIN G D, MCKINSEY G, et al. Many human accelerated regions are developmental enhancers[J]. Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 2013, 368(1632): 20130025.
- [14] 冯俊, 李光, 王义权. 后生动物非编码保守元件[J]. 遗传, 2013, 35(1): 35–44.
- [15] POLLARD K S, SALAMA S R, LAMBERT N, et al. An RNA gene expressed during cortical development evolved rapidly in humans[J]. Nature, 2006, 443(7108): 167–172.
- [16] HARRIS R S. Improved pairwise Alignment of genomic DNA[D]. Philadelphia: The Pennsylvania State University, 2007.
- [17] BLANCHETTE M, KENT W J, RIEMER C, et al. Aligning multiple genomic sequences with the threaded blockset aligner[J]. Genome Research, 2004, 14(4): 708–715.
- [18] HUBISZ M J, POLLARD K S, SIEPEL A. PHAST and RPHAST: phylogenetic analysis with space/time models[J]. Briefings in Bioinformatics, 2011, 12(1): 41–51.