

陈娜,刘闻,秦锐,等. 西藏小型猪多组织组蛋白修饰谱研究 [J]. 中国比较医学杂志, 2019, 29(5): 17-22.

Chen N, Liu W, Qin R, et al. Profiling of histone modifications in different tissues of the Tibet minipig [J]. Chin J Comp Med, 2019, 29(5): 17-22.

doi: 10.3969/j.issn.1671-7856.2019.05.003

西藏小型猪多组织组蛋白修饰谱研究

陈娜^{1#}, 刘闻^{2, 3#}, 秦锐¹, 于欣鑫¹, 顾为望^{2, 3*}, 高飞^{1*}

(1. 中国农业科学院深圳农业基因组研究所, 广东 深圳 518116;

2. 南方医科大学实验动物中心, 广州 510515;

3. 东莞松山湖明珠实验动物科技有限公司, 广东 东莞 523808)

【摘要】 目的 对西藏小型猪全基因组范围内的组蛋白修饰位点进行注释, 并比较不同组织间组蛋白修饰的差异。方法 取三月龄的雄性非同胎西藏小型猪两只, 分别取其血液、肝、小肠(回肠段)和背最长肌的组织样本, 进行染色质免疫共沉淀实验, 捕获 DNA 片段进行高通量测序。下机数据经过滤后比对到参考基因组上, 并使用 MACS 软件统计 reads 在基因组上的富集情况。结果 不同组蛋白在基因启动子区域分布的偏好位置不同, 分布规律大致分为两类, 一类如 H3K4me3 组蛋白修饰主要分布在转录起始位点附近的两侧, 并在距离 TSS 较近的下流位置出现一个宽度较窄的明显的峰值, 另一类如 H3K9me3 组蛋白修饰在整个启动子区域内分布较为平均, 整体呈现出中间低两边高的宽峰形状; 组织特异性组蛋白修饰影响的基因存在差异, 对应的功能通路比较广泛, 但都参与了 RNA 聚合酶 II 启动子的转录调控。结论 本研究提供了西藏小型猪四种组织的组蛋白修饰注释数据, 为今后以西藏小型猪为生物医学模型的转录调控研究提供了基础数据。

【关键词】 西藏小型猪; 动物模型; ChIP-Seq; 组蛋白修饰

【中图分类号】 R-33 【文献标识码】 A 【文章编号】 1671-7856(2019) 05-0017-06

Profiling of histone modifications in different tissues of the Tibet minipig

CHEN Na^{1#}, LIU Wen^{2, 3#}, QIN Rui¹, YU Xinxin¹, GU Weiwang^{2, 3*}, GAO Fei^{1*}

(1. Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518116, China.

2. Laboratory Animal Center, Southern Medical University, Guangzhou 510515.

3. Pearl Lab Animal Sci & Tech Co., Ltd., Dongguan 523808)

【Abstract】 **Objective** To perform a genome-wide histone modification profiling of Tibet minipigs and to compare with the differences in histone modifications in different tissues. **Methods** Tissue samples of blood, liver, ileum, and muscle of two three-month-old male Tibet minipigs were collected and used for chromatin immunoprecipitation sequencing. Through quality control, raw data were filtered and compared with the reference genome, then we used MACS software to enrich clean reads in the genome. **Results** The enrichment positions of different histone modifications differed in the gene promoter region, and the distribution patterns were roughly divided into two types. One type, H3K4me3, was mainly distributed on both sides near the transcription start site, exhibiting a peak with a narrow range at the downstream position.

【基金项目】 中国农业科学院科技创新工程。

【作者简介】 陈娜(1992—), 女, 博士研究生, 研究方向: 基因组学与蛋白质组学。E-mail: chennvnu@163.com

刘闻(1986—), 男, 硕士研究生, 研究方向: 实验动物学。E-mail: 18675849929@qq.com *共同第一作者

【通信作者】 顾为望(1956—), 男, 教授, 研究方向: 比较医学和人类疾病动物模型。E-mail: guww100@163.com

高飞(1980—), 男, 研究员, 研究方向: 动物发育、健康及与环境互作中的表观遗传学调控机制。E-mail: flys828@gmail.com

* 共同通信作者

The second type, H3K9me3, was distributed more evenly throughout the gene promoter, exhibiting a broad peak shape with a low middle and two high sides. Genes affected by the tissue-specific histone modifications differed, and the corresponding functional pathways were relatively extensive, but they were all involved in the transcriptional regulation of the RNA polymerase II promoter. **Conclusions** Our study provides a dataset of histone modification annotations for four types of different tissues of Tibet minipigs, which will serve as a foundation for future studies using transcriptional regulation of the Tibet minipig as a biomedical animal model.

【Keywords】 Tibet minipigs; animal model; ChIP-Seq; histone modification

与啮齿类小型实验动物相比,猪在解剖生理学、血液生化值、疾病发生和器官发育等方面具有与人类更加相似的特点,且无灵长类实验动物可能存在的伦理问题,因此成为人体生长发育和疾病发生等研究的优良实验模型。小型猪与人体重相似,具有体型小便于实验操作、饲养成本低的优点,在生物医药领域有着更加广阔的应用前景^[1]。特别小型猪在疾病动物模型制作、新药安全性评价、异种器官移植等领域应用更广,已成为生物医学研究中重要的实验动物^[2]。

基因组(genome)是分子生物学研究的基础,基因组学(genomics)的研究对生物医学模型的持续开发具有重要意义。2012年,猪基因组测序协会发布第一个完整测序的家猪基因组,其大小约为2.6 G^[3]。比较基因组学研究结果显示家猪基因组与人类基因组在序列和染色体结构上具有极高的同源性,其基因集注释已经比较完整,最新的参考基因注释集包括30173个基因^[4]。另外一方面,对基因组的染色质结构、表观遗传修饰、转录因子结合等功能注释与序列注释具有同等的重要性,这些注释可以将基因组序列与物种外在的生物学特性有效的对应起来。表观遗传学(epigenetics)包括DNA甲基化、组蛋白修饰、染色质重塑及非编码RNA调控等研究内容。表观遗传信息的变化对包括人在内的哺乳动物基因组有广泛而重要的调控效应,如转录抑制、基因组印记、细胞凋亡、染色体灭活等^[5]。目前,针对人类基因组的功能注释已经广泛开展并获得了重要成果。1998年,欧盟启动了人类“表观基因组学计划”和“基因组的表观遗传可塑性”研究计划,旨在解析人类DNA甲基化模式,并阐明其建立和维持的机制。而美国则发起并主导了人类及小鼠基因组大百科全书计划(ENCODE和mENCODE),提供最为全面的基因组功能注释^[6-7]。针对其他哺乳动物基因组的注释还相对缺少,较为知名的有FANNG项目,旨在构建饲养动物的基因组功能元件注释图谱,但目前获得的数据和成果

有限^[8]。

目前一些主要物种的基因组功能注释信息定期更新,如人、小鼠、斑马鱼等,而猪的基因组功能注释相对滞后,且不同类型的猪基因组有所差别。西藏小型猪目前尚缺乏完善的表观基因组注释数据,限制其作为生物医学模型的广泛应用。本研究采用染色质免疫共沉淀(chromatin immunoprecipitation, ChIP)结合高通量测序技术,完成了西藏小型猪四种组织的组蛋白修饰谱,并进行了深入的基因组表观遗传修饰谱鉴定及不同组织间的比较分析。本研究的数据结果将为西藏小型猪作为生物医学模型的应用奠定基础。

1 材料和方法

1.1 实验动物

普通级西藏小型猪2头(3月龄,雄性,非同胞),体重(3.0±0.3) kg,于东莞松山湖明珠实验动物科技有限公司购买并进行组织取样[SCXK(粤)2017-0030][SYXK(粤)2017-0123]。饲养和实验过程中遵循了实验动物使用的3R原则。本研究经东莞松山湖明珠实验动物科技有限公司实验动物管理和使用委员会批准实施[IACUC决议编号S-20170611-01]。

1.2 主要试剂与仪器

组蛋白修饰抗体购自艾博抗(Abcam)上海贸易有限公司。微量移液器和离心管购自艾本德(Eppendorf)中国有限公司;PCR扩增仪购自德国耶拿分析仪器股份有限公司(型号:Analytikjena Easycycler);高通量测序平台为Illumina Hiseq X Ten(PE150)。

1.3 实验方法

1.3.1 组织取样

实验动物安乐死前采集静脉血液(≥60 mL)于肝素钠抗凝管中制备血液样品;安乐死后分离出肝器官、回肠肠段、背最长肌等(≥20 g),分别分离成小块组织(≥1 g)于冻存管中,液氮速冻后置于干冰

箱内冷冻保存。

1.3.2 染色质免疫共沉淀实验

用上述组织样品进行染色质免疫共沉淀实验,选择 6 种组蛋白修饰抗体: H3K4me3、H3K9me3、H3K9ac、H3K27me3、H3K27ac、H3K36me3,经甲醛固定、细胞核提取、染色质破碎、免疫沉淀、解交联等步骤,收集特异性捕获的 DNA 片段。将 DNA 片段超声打断为 150~200 bp 片段后补齐末端、加测序接头并进行 PCR 扩增,构建文库用于上机测序。

1.3.3 ChIP-Seq 及其数据处理

文库质量检测合格后,经 Illumina 高通量测序平台测序,下机后的原始数据使用软件 Trimmomatic (Version 0.30) 进行过滤,得到的 clean reads 数据使用 BWA 软件比对到家猪参考基因组 (*Sscrofa* 11.1) 上,使用 MACS (Version 1.4.2) 软件统计 reads 在基因组的富集情况,并结合家猪基因组注释信息,统计 reads 富集区间在基因组各区域的分布情况。

2 结果

2.1 数据产出及质控

利用 Illumina 测序平台对西藏小型猪四个组织的 ChIP 样本进行双端测序,经统计,原始数据过滤后每个样品的 Q30 均大于 95%,GC 含量分布曲线接近正态分布,说明碱基质量稳定、序列组成均一,能保证后续分析的正常进行。

2.2 西藏小型猪组蛋白修饰的整体特征

基因启动子区含有转录必需的关键调控元件,在功能基因的表达与调控方面起着重要作用。为了从全基因组水平上展示四种不同组织各种组蛋白修饰类型的特征,我们选取了基因启动子区域并统计在这一区域内的组蛋白修饰值。定义每个基因的转录起始位点 (transcription start site, TSS) 上、下游各 2000 bp 为启动子区域,并将其平均分为 400 个 bin,每个 bin 的长度为 10 bp。组蛋白修饰标准化值的计算参考基因表达量 RPKM 的计算,具体公式如下:

$$N_i = \frac{n_i}{n * l_i} * 10^9$$

式中 n_i 为落在此区域内的 reads 数, n 为测序深度, l_i 为区域长度,每个 bin 的组蛋白修饰水平为全部基因在此区域内组蛋白修饰值的平均值。根据计算结果绘制了 TSS 附近区域组蛋白修饰的分布图 (图 1),横坐标为距离 TSS 的位置,纵坐标为标准化

后的组蛋白修饰值。

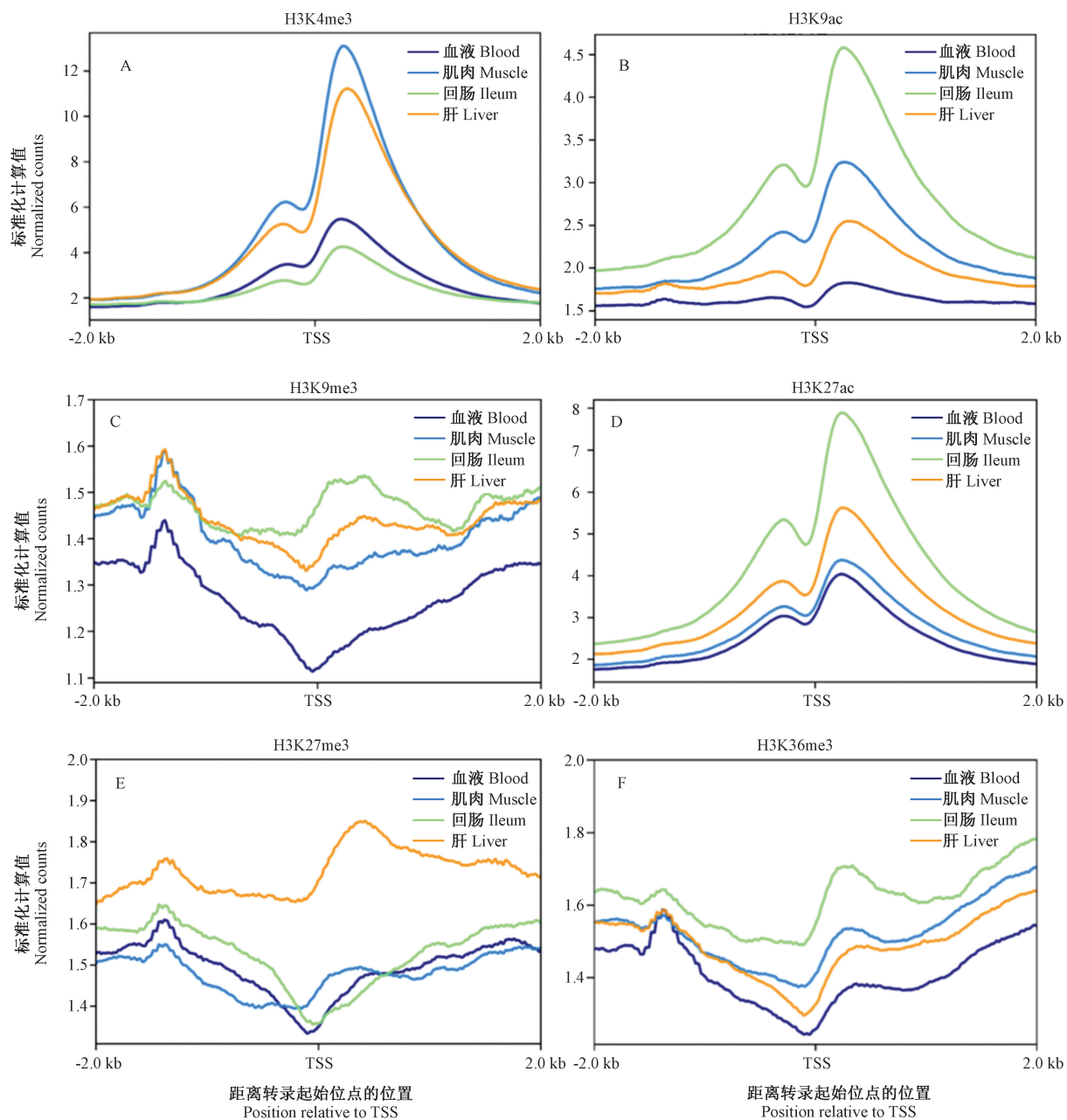
从图中可以看出,不同组蛋白修饰在基因启动子区域中分布的位置不同,各组蛋白修饰的分布规律大致可以分为两类,第一类如 H3K4me3 (图 1A), H3K9ac (图 1B) 和 H3K27ac (图 1D),组蛋白修饰信号富集于 TSS 两侧,并在距离 TSS 较近的下流位置出现一个宽度较窄的明显的峰值;另一类如 H3K9me3 (图 1C), H3K27me3 (图 1E) 和 H3K36me3 (图 1F),组蛋白修饰值较低且在整个启动子区域内分布较为平均,在 TSS 位点处出现明显的下降,在 -1500 bp 处有一个不明显的小峰,整体呈现出中间低两边高的宽峰形状。这与文献中报导的人类及其他哺乳类动物基因组的组蛋白修饰特征相似^[9]。组蛋白修饰峰值分布位置的不同显示了其具有的不同基因转录调控功能和特征。

2.3 西藏小型猪组织间组蛋白修饰分布特征

为了进一步探讨组蛋白修饰的结合位点特征及组织特异性,理解组蛋白修饰对基因调控的机制,我们结合了物种的基因组注释信息,对组蛋白修饰富集区间在整个基因组范围内的分布特征进行了分析。首先利用 MACS 软件进行组蛋白修饰的富集区间分析,得到每种组蛋白在基因组不同位置、不同数量的富集区间,然后将富集区间的信息与基因组注释信息结合,得到富集区间在基因组六个功能区域内的分布情况。如图 2 所示,横坐标表示不同的组蛋白修饰类型,纵坐标表示富集区间在各个区域内所占的百分比。可以看到四种组织中组蛋白修饰分布有一定的相似性,所有类型的修饰在基因间区 (intergenic region) 和内含子 (intron) 区的分布比例均高于其他功能区,特别是 H3K9me3;而同一组蛋白位点的乙酰化修饰 H3K9ac 在不同组织中的分布模式差异很大,具有组织特异性,在血液和肌肉组织中基因间区所占比例较大,而在肝和回肠组织中各功能区分布比例较为平均;H3K4me3 在各组织的分布模式相似,且启动子区所占的比例高于其他修饰类型。组蛋白修饰在各组织间的不同,造成特异性基因的表达发生改变,影响了细胞分化、组织形成等过程。

2.4 组织间差异修饰基因的功能富集分析

为探究组蛋白的不同修饰对组织差异性的影响,我们对在 4 种组织间启动子区组蛋白修饰有显著差异的基因进行了统计。图 3 显示了各类组蛋白修饰对应的差异基因的重叠情况,可以看出各类修



注:A:组蛋白修饰 H3K4me3 在 TSS 附近的分布;B:组蛋白修饰 H3K9ac 在 TSS 附近的分布;C:组蛋白修饰 H3K9me3 在 TSS 附近的分布;D:组蛋白修饰 H3K27ac 在 TSS 附近的分布;E:组蛋白修饰 H3K27me3 在 TSS 附近的分布;F:组蛋白修饰 H3K36me3 在 TSS 附近的分布。

图 1 TSS 附近组蛋白修饰的分布

Note. A, Histone modification H3K4me3 near transcription start sites. B, Histone modification H3K9ac near transcription start sites. C, Histone modification H3K9me3 near transcription start sites. D, Histone modification H3K27ac near transcription start sites. E, Histone modification H3K27me3 near transcription start sites. F, Histone modification H3K36me3 near transcription start sites.

Figure 1 Histone modifications near the transcription start sites

饰实际影响的基因存在差异,对于每种修饰约有半数差异基因与其他修饰没有交叉,为此种修饰类型特有,如 H3K9ac。接着我们将差异基因名称输入到 DAVID 数据库 (<https://david.ncifcrf.gov/tools.jsp>) 中进行了 GO (gene ontology) 功能分析,结果显示,

尽管组织特异性组蛋白修饰所覆盖的功能通路比较广泛,但几乎都参与了 RNA 聚合酶 II 启动子的转录调控。转录的正、负调控及转录抑制、转录激活的差异导致了不同组织之间多种基因的表达差异。

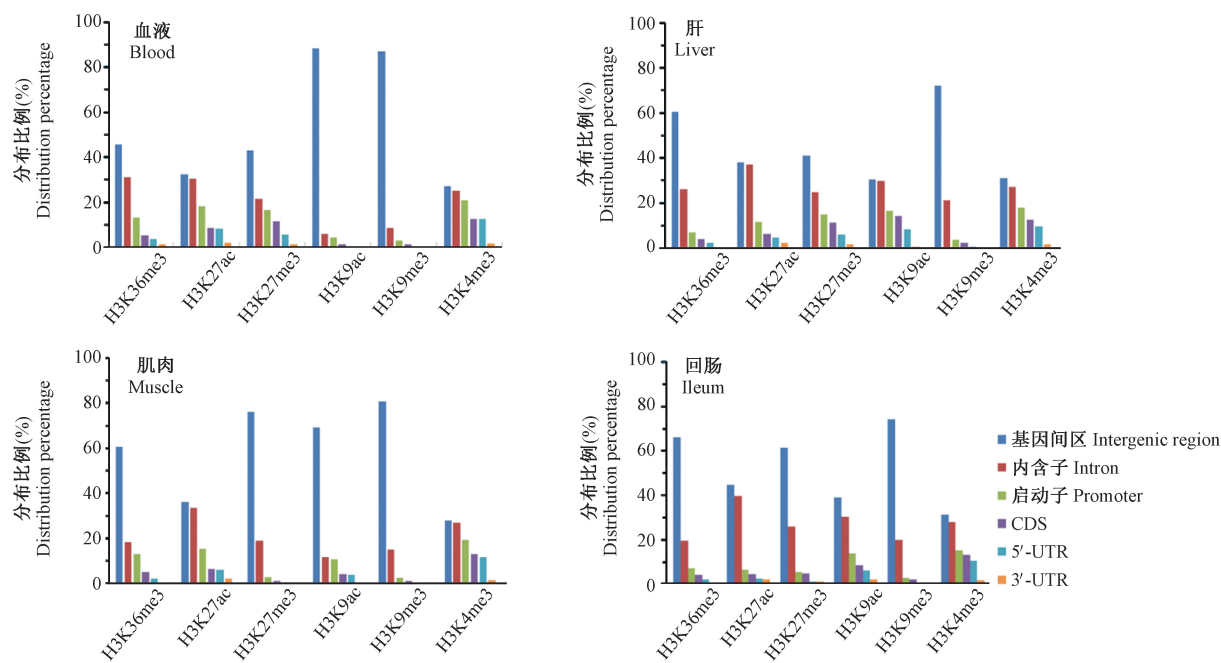


图 2 全基因组六个功能区域的组蛋白修饰分布
Figure 2 Distribution of histone modifications in six functional regions of the genome

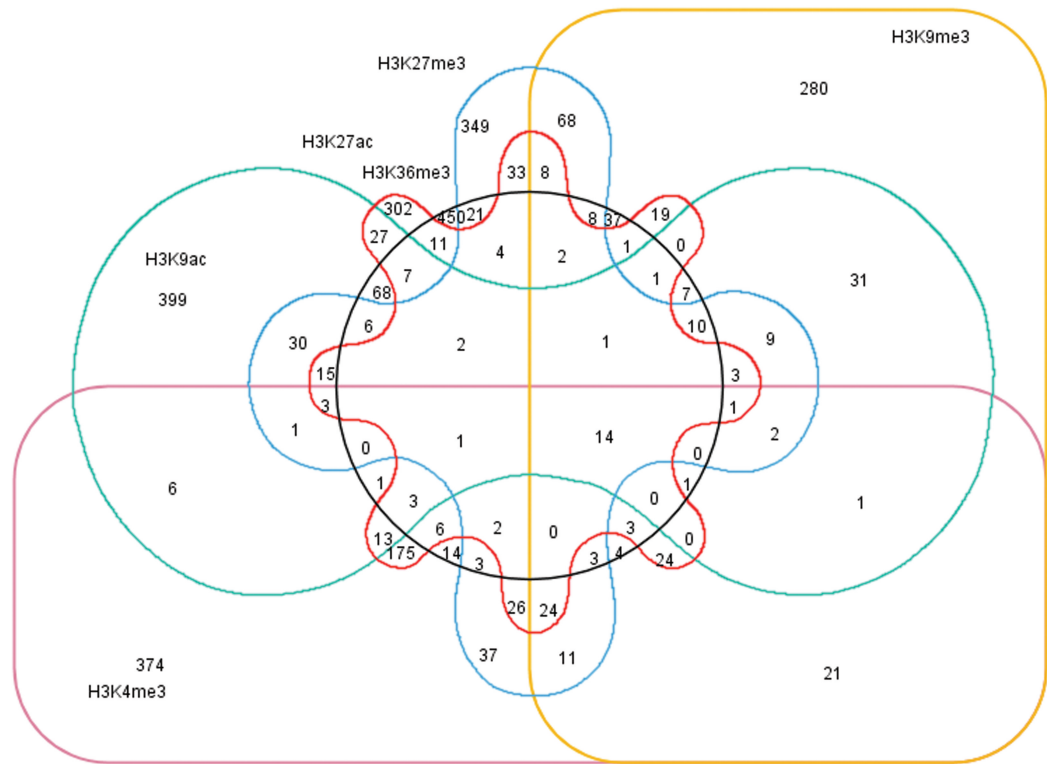


图 3 组蛋白差异修饰基因的重叠
Figure 3 Overlapping of the differential histone modification genes

3 讨论

西藏小型猪原产于我国藏区高原,是体型较小

的实验用猪种之一,2004 年由南方医科大学实验动物中心将其从西藏林芝地区引进广东省,并完成了风土驯化及实验动物化培育。免疫学、遗传学和病

原组织学等方面的研究表明,西藏小型猪具有独特的免疫指标和种质资源特性,其血液生理生化指标与普通猪差异较大,与人类水平接近,是一种优良的人类疾病实验动物模型^[10]。包括组蛋白修饰在内的生物表观遗传学修饰能够激活或抑制基因转录,调节基因表达,已成为遗传学研究的一个重要分支,与 DNA 序列水平上的变化同样重要,特定位点的组蛋白修饰已经成为肿瘤检查标志物^[11]。目前对西藏小型猪开展的研究主要围绕相关动物模型、药物实验及转基因克隆猪等方面,但是对其基因组功能及表观组注释的研究还是空白。本研究从全基因组水平初步解析了西藏小型猪组蛋白修饰的特征,各种组蛋白修饰在启动子区的分布模式有所差别,这与此前在的报道结果一致^[12-13]。此前对于组蛋白各位点修饰的功能研究表明,转录起始位点的 H3K4me3、H3K9ac 和 H3K27ac 修饰有利于激活基因转录,活性转录基因的 gene body 区富集 H3K36me3,而基因的转录抑制与 H3K9me3 和 H3K27me3 有关^[14-15]。对其基因组的注释解析将有利于西藏小型猪在转基因克隆发育等方面的研究 and 应用。

组蛋白修饰通过改变染色质结构,影响转录因子与靶基因各种顺式作用元件之间的亲和性,发挥调控基因转录调控的作用,正确识别它们的富集部位对理解发育和细胞功能至关重要。启动子是关键的基因功能元件之一,启动子区域范围广泛,包含核心启动子区域和调控区域。核心启动子区域通常位于转录起始点上游 2000 bp 以内,负责基因基础水平的转录,调控区域能够对不同环境因素或信号传导作出应答,调节相应的基因表达^[16]。本研究选取了转录起始位点上下游各 2000 bp 的区域作为启动子区域,统计此区域内组蛋白修饰密度并进行标准化,结果表明组蛋白修饰在启动子区的富集模式主要有宽峰和窄峰两种,宽峰覆盖整个启动子区且无明显峰值,窄峰在转录起始位点附近有明显的富集高峰。组蛋白修饰信号的富集模式的不同,决定了其转录调控功能的差异。这几种组蛋白修饰类型的具体作用方式,后续可结合基因表达数据展开深入探究。

此外,本研究发现有大量的组蛋白修饰富集区间落在了基因间区和内含子区,意味着同人类和小鼠基因组一样,小型猪基因组存在着大量未鉴定的转录本和新基因。因此,本研究的数据结果将为后

续西藏小型猪的基于基因发掘、转录调控的生物医学模型研究奠定重要基础。

参考文献:

- [1] 袁进, 顾为望. 小型猪作为人类疾病动物模型在生物医学研究中的应用 [J]. 动物医学进展, 2011, 32(2): 108-111.
- [2] Schachtschneider KM, Madsen O, Park C, et al. Adult porcine genome-wide DNA methylation patterns support pigs as a biomedical model [J]. BMC Genomics, 2015, 16:743.
- [3] Groenen MA, Archibald AL, Uenishi H, et al. Analyses of pig genomes provide insight into porcine demography and evolution [J]. Nature, 2012, 491(7424): 393-398.
- [4] Meurens F, Summerfield A, Nauwynck H, et al. The pig: a model for human infectious diseases [J]. Trends Microbiol, 2012, 20: 50-57.
- [5] Rivera CM and Ren B. Mapping human epigenomes [J]. Cell, 2013, 155(1): 39-55.
- [6] ENCODE Project Consortium. The ENCODE (ENCyclopedia Of DNA Elements) Project [J]. Science, 2004, 306(5696): 636-640.
- [7] Stamatoiyannopoulos JA, Snyder M, Hardison R, et al. An encyclopedia of mouse DNA elements (Mouse ENCODE) [J]. Genome Biol, 2012, 13(8): 418.
- [8] Andersson L, Archibald AL, Bottema CD, et al. Coordinated international action to accelerate genome-to-phenome with FAANG, the Functional Annotation of Animal Genomes project [J]. Genome Biol, 2015, 16(1): 57.
- [9] Kolasinska-Zwierz P, Down T, Latorre I, et al. Differential chromatin marking of introns and expressed exons by H3K36me3 [J]. Nat Genet, 2009, 41(3): 376-381.
- [10] 顾为望, 刘运忠, 唐小江, 等. 西藏小型猪血液生理生化指标的初步研究 [J]. 中国实验动物学报, 2007, 15(1): 60-63.
- [11] Barski A, Cuddapah S, Cui K, et al. High-resolution profiling of histone methylations in the human genome [J]. Cell, 2007, 129(4): 823-837.
- [12] Liu X, Wang C, Liu W, et al. Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos [J]. Nature, 2016, 537(7621): 558-562.
- [13] Zhao XD, Han X, Chew JL, et al. Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells [J]. Cell Stem Cell, 2007, 1(3): 286-298.
- [14] Kimura H. Histone modifications for human epigenome analysis [J]. J Hum Genet, 2013, 58(7): 439-445.
- [15] Zhou VW, Goren A, Bernstein BE. Charting histone modifications and the functional organization of mammalian genomes [J]. Nat Rev Genet, 2011, 12(1): 7-18.
- [16] Zabidi MA, Arnold CD, Schemhuber K, et al. Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation [J]. Nature, 2015, 518(7540): 556-559.