GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Comparative analyses of histone H3K9 trimethylations in the heart and spleen of normal humans

W. Sui[1], C. Cao[1,2], W. Che[1], J. Chen[1], W. Xue[1], P. Liu[1], L. Guo[2] and Y. Dai[3]

[1]Nephrology Department of Guilin 181st Hospital,
Guangxi Key Laboratory of Metabolic Diseases Research,
Guilin, Guangxi, China
[2]College of Life Science, Guangxi Normal University, Guilin, Guangxi, China
[3]Clinical Medical Research Center,
The Second Clinical Medical College of Jinan University,
Shenzhen People's Hospital, Shenzhen, Guangdong, China

Corresponding author: Y. Dai
E-mail: daiyong2222@gmail.com

Genet. Mol. Res. 13 (1): 1697-1706 (2014)
Received June 27, 2013
Accepted October 15, 2013
Published January 14, 2014
DOI http://dx.doi.org/10.4238/2014.January.14.5

ABSTRACT. The global features of trimethylations of histone 3 at lysine 9 (H3K9me3) have been well studied in recent years; however, most of these studies were performed in mammalian cell lines. In this study, we generated genome-wide maps of H3K9me3 of the human heart and spleen using chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) technology. We examined the global patterns of H3K9me3 in both tissues and found that modifications were closely associated with tissue-specific expression, function, and development. In addition, we found that 169 genes displayed significant H3K9me3 differences between the heart and spleen. Among these genes, 64 were heart-H3K9me3-specific, 87 genes were spleen-H3K9me3-specific, and 18 were shared in both heart- and spleen-H3K9me3. In conclusion, we observed significant differences in H3K9me3 in the heart and spleen, which may help to explain epigenetic differences between these tissues.

Such novel findings highlight the significance of H3K9me3 as a potential biomarker or promising target for epigenetic-based disease treatment.

Key words: ChIP-seq; Epigenetic; H3K9me3

INTRODUCTION

Histone modifications represent one of the most common epigenetic mechanisms for controlling gene expression, and thus cell proliferation, fate determination, and survival, during development. Among the different types of histone modifications, the trimethylation of histone 3 at lysine 9 (H3K9me3) has been widely studied and is often associated with heterochromatin formation, gene repression, and transcriptional elongation in different tissue types and organisms (Lessard and Crabtree, 2010).

Epigenetics is the study of inherited modifications in gene expression and phenotypes that are caused by mechanisms other than changes in the underlying DNA sequence. Unlike genetic polymorphisms, non-genetic partial and reversible DNA modifications can occur in epigenetic alterations (Elie et al., 2012).

Methylation occurs at five major lysine residues located within the amino-terminal histone tails (H3K4, H3K9, H3K27, H3K36, and H4K20) and at one lysine residue found within the globular histone domain (H3K79); these lysines can be mono-, di-, or tri-methylated (Ciccone and Chen, 2009). Among the various histone lysine methylation patterns, this study focused on H3K9 owing to its association with condensed and repressed chromatin (Park et al., 2011). H3K9me3 is a posttranslational modification that is commonly linked with both facultative and constitutive heterochromatin formation and transcriptional repression (Shinkai and Tachibana, 2011).

High-throughput techniques are very efficient methods for simultaneously analyzing large-scale gene expressions to investigate the complex molecular basis of pathological processes on a genomic scale (Wu et al., 2011). One of these approaches, chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq), enables the genome-wide identification of binding sites of transcription factors (TFs) and other DNA-binding proteins (Raha et al., 2010). With the emergence of high-throughput sequencing platforms, such as Illumina Genome Analyzer and SOLiD, and the availability of ChIP-grade antibodies, ChIP-seq has become one of the most widely used methods for determining functional elements in the genome (Wacker and Kim, 2009). In general, when compared with ChIP-chip, ChIP-seq has a higher signal/noise ratio, is less expensive, and requires lower amounts of ChIP DNA for genome-wide analysis (Euskirchen et al., 2007).

To date, very few comparative analyses of H3K9me3 have been performed among mammalian tissues. Hence, in this study, we adopted ChIP-seq technology to profile and compare variations in H3K9me3 at the genome-wide level in the human heart and spleen to gain a better understanding of the epigenetic differences of these tissues. By combining analyses of both datasets, we globally described the tissue-specific expression, function, and development.

MATERIAL AND METHODS

Samples

Four hearts and spleens were obtained from four brain-dead adults who had volun-

tarily consented to donate their organs to the 181st Hospital of the Chinese People's Liberation Army, Guilin, Guangxi, China. Two groups of the samples were stored at -80°C until further analysis. The protocol for this study was approved by a suitably constituted Ethics Committee of the Guangxi Key Laboratory for Metabolic Diseases Research and the 181st Hospital of the Chinese People's Liberation Army. All organ donors signed informed consent for this study.

ChIP analysis

We collected tissue samples (pooled from four individuals) of the heart and spleen from human bodies. We carried out ChIP-seq experiments according to the published procedure (BGI). Briefly, tissue samples from the human heart and spleen were homogenized and fixed in 1% formaldehyde. Chromatins were fragmented at a size range of 100-500 bp and incubated with antibodies (against trimethylation Lys 9) at 4°C overnight. After cross-linking reversal and proteinase K treatment, DNA samples were precipitated and purified with the QIA quick PCR purification kit before library construction. Approximately 10 ng DNA was used for adaptor ligation, gel purification, and polymerase chain reaction (PCR) with 15 cycles.

Next-generation sequencing

All standard protocols for Illumina sequence preparation, sequencing, and quality control were followed in accordance with manufacturer instructions (Illumina, San Diego, CA, USA). In short, DNA recovered from a conventional ChIP procedure was quantified using the QuantiFluor fluorometer (Promega, Madison, WI, USA). The DNA integrity was verified using the Agilent Bioanalyzer 2100 (Agilent, Palo Alto, CA, USA). The DNA was then processed, including end repair, adaptor ligation, and size selection, using an Illumina sample prep kit following manufacturer instructions (Illumina). Final DNA libraries were validated and sequenced at 75 bp per sequence read using an Illumina GAIIx sequencer at a depth of approximately 30 million sequences per sample.

ChIP-seq analysis

ChIP DNA end repairing, adaptor ligation, and amplification were performed as described previously (Lin et al., 2009). Fragments of approximately 49 bp (without linkers) were isolated from agarose gel and used for sequencing using the Solexa/Illumina 2G genetic analyzer. Sequencing tags were aligned to the Hg19 February 2009 release of the reference genome using SOAP 2.21, an ultrafast memory-efficient short read aligner (Langmead et al., 2009). We only considered tags that aligned uniquely with less than two mismatches. ChIP-seq reads analysis included the ChIP sequencing distribution in the whole genome, the ChIP sequencing distribution in repeat regions, and the ChIP sequencing distribution in gene regions. For enriched-region (peak) identification (peak scanning), we used the Model-based Analysis of ChIP-seq (MACS 1.4.0) algorithm (Feng et al., 2011). A simple effective technique for the analysis of eukaryotes, MACS was designed to identify transcription factor binding sites and histone modification-enriched regions in ChIP-seq datasets, with or without control samples (Shin et al., 2012). The following parameter for MACS was applied: P value cutoff = 1.00e-05; if the candidate peak region P value <1.00e-05, the region was considered a peak.

The enriched peaks were annotated with gene annotation using AmiGO version 1.8, a web application that allows users to search, sort, analyze, and visualize ontologies and related gene product annotation data. AmiGO can be used online at the Gene Ontology (GO) website to access the data provided by the GO Consortium (Carbon et al., 2009).

RESULTS

ChIP-seq

Using specific antibodies, we performed a ChIP-seq analysis on two samples: H3K9me3 of the heart and H3K9me3 of the spleen. We used SOAP 2.21, an ultrafast memory-efficient short read aligner, to align sequencing reads to the reference genome (Hg19), and the alignment results are presented in Table 1. All ChIP-seq data aligned well with the reference genome. This process resulted in a total of 12,943,074 reads and 10,649,001 (82.28%) aligned reads of heart samples, and a total of 12,952,347 reads and 10,491,537 (91.72%) aligned reads of spleen samples.

Table 1. Results of sequence alignment.

Sample ID	Number of total reads (pair)	Number of aligned reads (pair)	Number of unique aligned reads (pair)	Aligned rate	Unique aligned rate
heart-H3K9me3	12943074	10649001	9358082	82.28%	72.30%
spleen-H3K9me3	12952347	11879735	10491537	91.72%	81.00%

Genome-wide distributions of H3K9me3 of samples

We used MACS for enriched region (peak) identification with a P-value cutoff of 1.00e-05 (if the candidate peak region P value <1.00e-05, the region is a peak). The ChIP-seq for the heart sample generated 187 enriched regions (peaks) with a total length of 155,759 bp, and the spleen sample generated 217 enriched regions (peaks) with a total length of 180,635 bp. The peak distribution of the ChIP-seq reads of H3K9me3 in samples exhibited a typical priority distribution pattern (Figure 1).

GO analysis of peaks relative to annotated genes

To further understand the functions of the annotated genes related to peaks, they were functionally categorized according to GO annotation terms. In terms of the GO database, the differentially expressed proteins encoded by these genes were divided into 3 categories: biological process, cellular component, and molecular function. Among the annotated genes, 104 genes of the heart with annotation terms were linked to the GO cellular component (37 genes), molecular function (37 genes), and biological process (30 genes) categories, and 137 genes of the spleen with annotation terms were linked to the GO cellular component (52 genes), molecular function (45 genes), and biological process (40 genes) categories. Within the cellular component categories, we found that of the annotated genes of the heart, 34 were annotated as cell parts and 19 located on organelles, whereas of the annotated genes of the spleen, 46 were annotated as cell parts and 21 located on organelles. The molecular function ontology showed that the numbers of binding and catalytic activity genes of the heart were 33 and 17, respectively,

and the numbers of binding and catalytic activity genes of the spleen were 36 and 16, respectively. The proportion of genes showing enzyme regulator activity of the heart and spleen were 8.1 and 8.9% respectively. With respect to the biological process, 27 and 20 genes of the heart were involved in cellular processes and in biological regulation, respectively, and 36 and 25 genes of the spleen were involved in cellular processes and biological regulation, respectively.

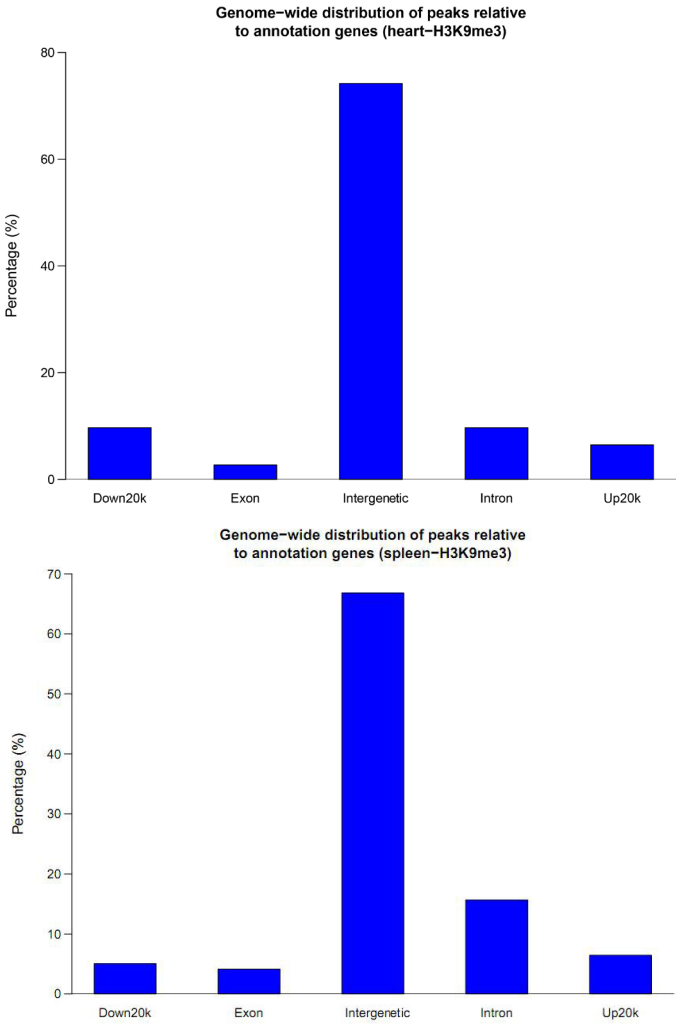

Figure 1. Distribution of the peak at different gene function element.

Comparison of H3K9me3 status between the heart and spleen

Analysis of differences in samples was based on the peak-related gene list of different samples to determine the numbers of unique and shared genes between the heart and spleen based on gene ID (Figure 2). The total number of sites of unique genes of samples affected by

peak was used as the cardinal number to calculate the proportion of gene-related sites (Figure 3). GO functional clustering was calculated after variance analysis. The enrichment of shared peak regions was compared in the two samples using heat maps (Figure 4).

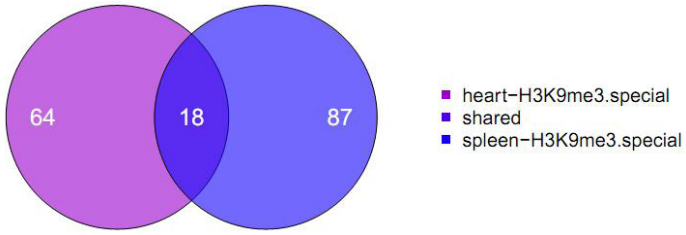

Figure 2. Venturi of different analysis of related genes of peak.

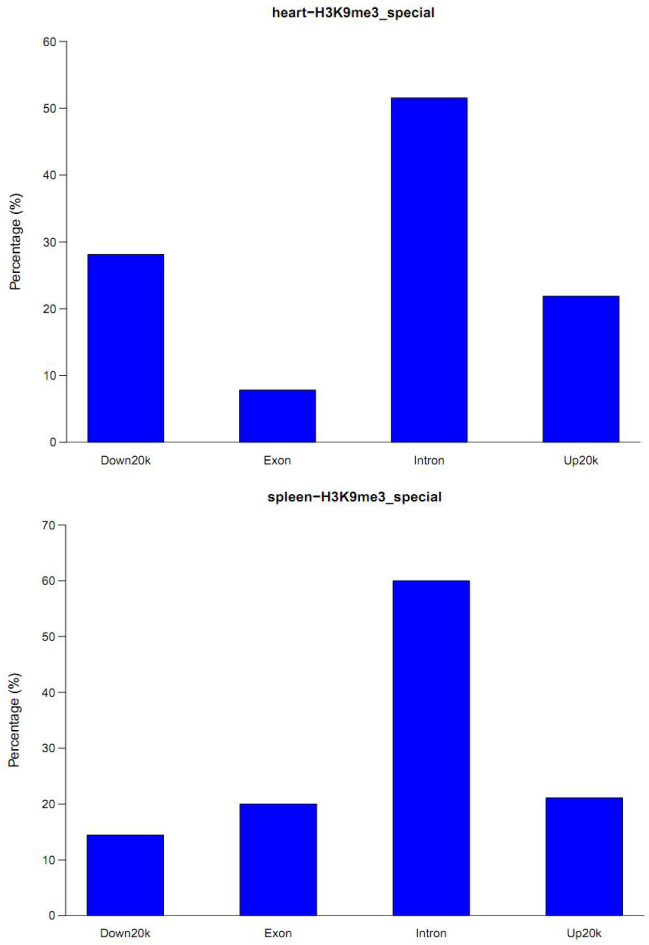

Figure 3. Proportion of distribution of each sample specific gene reads.

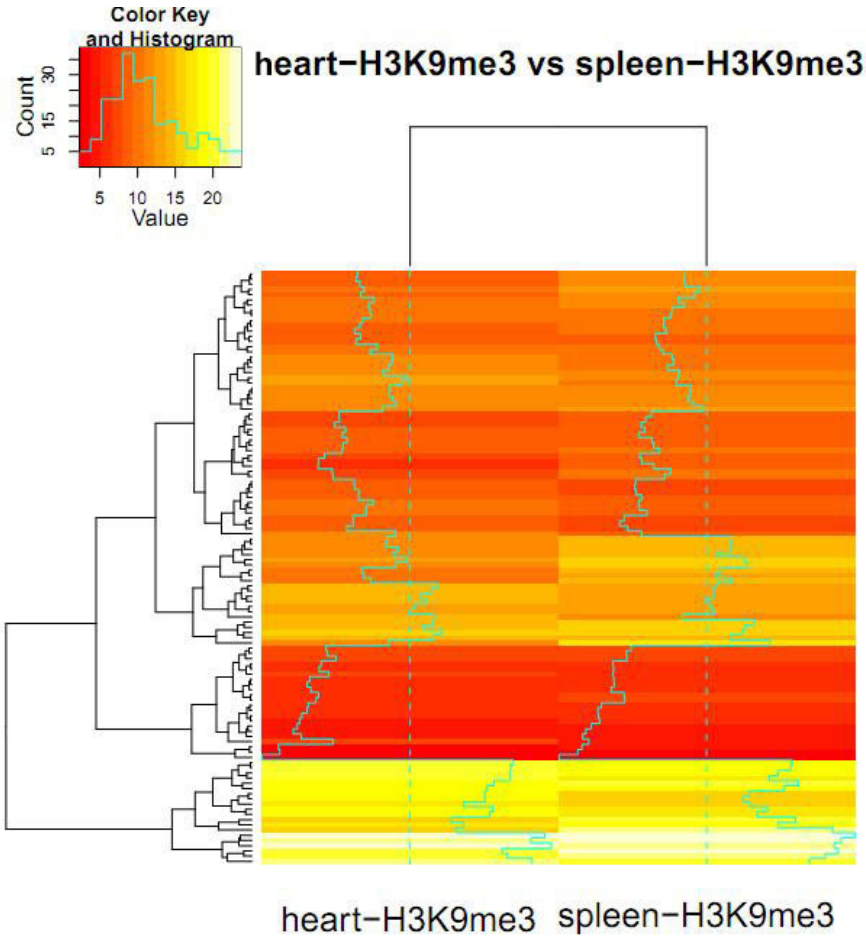

Figure 4. Cluster analysis of the enrichment of shared peak.

By applying the above analysis procedure to the sequence results, we found that 169 genes displayed significant H3K9me3 differences between the heart and spleen. Among these genes, 64 were heart-H3K9me3-specific, 87 were spleen-H3K9me3-specific, and 18 genes were shared in both heart- and spleen-H3K9me3.

The H3K9me3 alterations of 8 selected genes are *DLG2*, *ANO6*, *NFIA*, *PTN3*, *C11orf4*, *ITSN2*, *LPP*, and *RBMS3*. These selected genes showed the largest differences in the heart and spleen, respectively. Among these genes, *PTPN3* and *RBMS3* were further selected, which showed the largest differences in introns of the heart and spleen, respectively.

DISCUSSION

Modifications of histone tails are believed to specify a code that regulates the expression of genes. Trimethylated H3K9 (H3K9me3) is a highly conserved histone posttransla-

tional modification that is commonly linked with both facultative and constitutive heterochromatin formation and transcriptional repression. H3K9 methylation was initially studied as a marker of epigenetic silencing and constitutive heterochromatin, which is found close to the centromeres. To cover these regions, in this study, we used ChIP-seq, which enables the genome-wide identification of binding sites of TFs and other DNA-binding proteins. Although H3K9me3 was first observed several years ago, very little is known about its subtle interrelationships with other epigenetic modifications and its potential functional significance in human organs. Therefore, in this study, we selected H3K9me3 as the target, performed investigations using the ChIP-seq strategy, and explored the hypothesis that H3K9me3 are associated with the heart and spleen.

In the present study, we mainly analyzed the trimethylation status of H3K9 in the heart and spleen. The identified candidate genes included genes associated with immunity, cell signal transduction, protein transcription, and synthesis of ion channels, transporters, and the extracellular matrix, among others.

Among the identified candidate genes, we found that *PTPN3* and *RBMS3* displayed differences between introns in the heart and the spleen, respectively. *PTPN3* encodes the protein tyrosine phosphatase non-receptor type 3 (PTPN3) protein, which is a member of the protein tyrosine phosphatase (PTP) family. PTPs are known to be signaling molecules that regulate a variety of cellular processes including cell growth, differentiation, mitotic cycle, and oncogenic transformation. This protein contains a C-terminal PTP domain and an N-terminal domain homologous to the band 4.1 superfamily of cytoskeletal-associated proteins. P97, a cell cycle regulator that is involved in a variety of membrane-related functions, has been shown to be a substrate of this PTP. This PTP was also found to interact with, and be regulated by, adaptor protein 14-3-3 beta. Several alternatively spliced transcript variants encoding different isoforms have been found for this gene.

Jung et al. (2012) applied the next-generation sequencing technique to analyze the transcriptome of the non-small cell lung carcinoma (NSCLC) cell line H2228, and discovered a fusion transcript composed of multiple exons of anaplastic lymphoma receptor tyrosine kinase (ALK) and PTPN3. Detailed analysis of the genomic structure revealed that a portion of the genomic region encompassing exons 10 and 11 of ALK has been translocated into the intronic region between exons 2 and 3 of PTPN3. The key net result appears to be the null mutation of one allele of PTPN3, a gene with tumor suppressor activity. Consistently, ectopic expression of PTPN3 in NSCLC cell lines led to inhibition of colony formation (Jung et al., 2012).

The RNA binding motif, single stranded interacting protein 3 (RBMS3) gene encodes an RNA-binding protein that belongs to the c-myc gene single-strand binding protein family. These proteins are characterized by the presence of two sets of the ribonucleoprotein consensus sequence (RNP-CS) that contain conserved motifs, RNP1 and RNP2, was originally described in RNA binding proteins, and is required for DNA binding. These proteins have been implicated in such diverse functions as DNA replication, gene transcription, cell cycle progression, and apoptosis. The encoded protein was isolated by virtue of its binding to an upstream element of the alpha2 (I) collagen promoter. The observation that this protein localizes mostly in the cytoplasm suggested that it might be involved in a cytoplasmic function, such as controlling RNA metabolism, rather than transcription. Multiple alternatively spliced transcript variants encoding different isoforms have been found for this gene. Functional studies using both overexpression and suppression systems demonstrated that RBMS3 plays a strong

tumor suppressive role in nasopharyngeal carcinoma. The tumor suppressive mechanism of RBMS3 was associated with its role in cell cycle arrest at the G1/S checkpoint by upregulating p53 and p21, downregulating cyclin E and cyclin-dependent kinase 2, and subsequently inhibiting Rb-ser780. Further analysis demonstrated that RBMS3 played a pro-apoptotic role in a mitochondrial-dependent manner via activation of caspase-9 and poly-ADP ribose polymerase. Finally, RBMS3 inhibited microvessel formation, which might be mediated by the down-regulation of matrix metalloproteinase (MMP)-2 and β-catenin and inactivation of its downstream targets, including cyclin-D1, c-Myc, MMP-7, and MMP-9 (Chen et al., 2012).

In a novel zebrafish protein trap screen, Jayasena and Bronner (2012) identified an RNA-binding protein, Rbms3, which is transiently expressed in the cytoplasm of condensing neural crest cells within the pharyngeal arches. Morphants for rbms3 displayed reduced proliferation of the prechondrogenic crest and significantly altered expression of chondrogenic/osteogenic lineage markers. This phenotype strongly resembles that of cartilage/crest defects observed in Tgf-βr2:Wnt1-Cre mutants, which suggests a possible link with transforming growth factor (TGF)-β signaling. Several findings are consistent with this: a) Rbms3 stabilized a reporter transcript with the smad2 3'-untranslated region, b) RNA immunoprecipitation with full-length Rbms3 showed enrichment for smad2/3, and c) pSmad2 levels were reduced in rbms3 morphants. Overall, these results suggest that Rbms3 post-transcriptionally regulates one of the major pathways that promotes chondrogenesis, the TGFβ receptor pathway (Jayasena and Bronner, 2012).

In summary, we systematically evaluated the status of H3K9me3 in the heart and spleen and gained new insights into the links between key genes and histone methylation in the context of human organs. Our study confirms the utility of next-generation sequencing as a tool for the discovery of DNA methylation and has led to the identification of a novel methylation and disease that may be of diagnostic, prognostic, and therapeutic importance. Such novel findings show the significance of H3K9me3 as a potential biomarker or promising target for epigenetic-based disease treatment.

ACKNOWLEDGMENTS

The authors are deeply grateful to all the volunteers. Research supported by grants from the Key Project for Science and Technology of Guangxi (#2012GXNSFDA053017), the Construction Projects Planned Mission Statement of Guangxi Science and Technology Infrastructure (Key Laboratory, #11-031-33), and the Construction Projects Planned Mission Statement of Guangxi Key Laboratory (#12-071-32).

REFERENCES

Carbon S, Ireland A, Mungall CJ, Shu S, et al. (2009). AmiGO: online access to ontology and annotation data. *Bioinformatics* 25: 288-289.

Chen J, Kwong DL, Zhu CL, Chen LL, et al. (2012). RBMS3 at 3p24 inhibits nasopharyngeal carcinoma development via inhibiting cell proliferation, angiogenesis, and inducing apoptosis. *PLoS One* 7: e44636.

Ciccone DN and Chen T (2009). Histone lysine methylation in genomic imprinting. *Epigenetics* 4: 216-220.

Elie V, Fakhoury M, Deschenes G and Jacqz-Aigrain E (2012). Physiopathology of idiopathic nephrotic syndrome: lessons from glucocorticoids and epigenetic perspectives. *Pediatr. Nephrol.* 27: 1249-1256.

Euskirchen GM, Rozowsky JS, Wei CL, Lee WH, et al. (2007). Mapping of transcription factor binding regions in mammalian cells by ChIP: comparison of array- and sequencing-based technologies. *Genome Res.* 17: 898-909.

Feng J, Liu T and Zhang Y (2011). Using MACS to identify peaks from ChIP-Seq data. *Curr. Protoc. Bioinform.* Chapter 2: Unit.

Jayasena CS and Bronner ME (2012). Rbms3 functions in craniofacial development by posttranscriptionally modulating TGF-β signaling. *J. Cell Biol.* 199: 453-466.

Jung Y, Kim P, Jung Y, Keum J, et al. (2012). Discovery of ALK-PTPN3 gene fusion from human non-small cell lung carcinoma cell line using next generation RNA sequencing. *Genes Chromosomes Cancer* 51: 590-597.

Langmead B, Schatz MC, Lin J, Pop M, et al. (2009). Searching for SNPs with cloud computing. *Genome Biol.* 10: R134.

Lessard JA and Crabtree GR (2010). Chromatin regulatory mechanisms in pluripotency. *Annu. Rev. Cell Dev. Biol.* 26: 503-532.

Lin B, Wang J, Hong X, Yan X, et al. (2009). Integrated expression profiling and ChIP-seq analyses of the growth inhibition response program of the androgen receptor. *PLoS One* 4: e6589.

Park JA, Kim AJ, Kang Y, Jung YJ, et al. (2011). Deacetylation and methylation at histone H3 lysine 9 (H3K9) coordinate chromosome condensation during cell cycle progression. *Mol. Cells* 31: 343-349.

Raha D, Hong M and Snyder M (2010). ChIP-Seq: a method for global identification of regulatory elements in the genome. *Curr. Protoc. Mol. Biol.* Chapter 21: Unit-14.

Shin JH, Li RW, Gao Y, Baldwin R, et al. (2012). Genome-wide ChIP-seq mapping and analysis reveal butyrate-induced acetylation of H3K9 and H3K27 correlated with transcription activity in bovine cells. *Funct. Integr. Genomics* 12: 119-130.

Shinkai Y and Tachibana M (2011). H3K9 methyltransferase G9a and the related molecule GLP. *Genes Dev.* 25: 781-788.

Wacker DA and Kim TH (2009). From sextant to GPS: twenty-five years of mapping the genome with ChIP. *J. Cell Biochem.* 107: 6-10.

Wu CC, Chen JS, Huang CF, Chen CC, et al. (2011). Approaching biomarkers of membranous nephropathy from a murine model to human disease. *J. Biomed. Biotechnol.* 2011: 581928.