



**HAL**  
open science

## A demonstration of the H3 trimethylation ChIP-seq analysis of galline follicular mesenchymal cells and male germ cells

Kaj Chokeshaiusaha, Denis Puthier, Catherine Nguyen, Thanida Sananmuang

► **To cite this version:**

Kaj Chokeshaiusaha, Denis Puthier, Catherine Nguyen, Thanida Sananmuang. A demonstration of the H3 trimethylation ChIP-seq analysis of galline follicular mesenchymal cells and male germ cells. Asian-Australasian journal of animal sciences, 2018, 10.5713/ajas.17.0744 . hal-01743254

**HAL Id: hal-01743254**

**<https://amu.hal.science/hal-01743254>**

Submitted on 10 Jan 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



18 **Title of the manuscript:** A demonstration of the H3 trimethylation ChIP-seq analysis of galline follicular  
19 mesenchymal cells and male germ cells

20

21 **ABSTRACT**

22 **Objective:** Trimethylation of histone 3 (H3) at 4<sup>th</sup> lysine N-termini (H3K4me3) in gene promoter region was the  
23 universal marker of active genes specific to cell lineage. On the contrary, coexistence of trimethylation at 27<sup>th</sup>  
24 lysine (H3K27me3) in the same loci—the bivalent H3K4m3/H3K27me3 was known to suspend the gene  
25 transcription in germ cells, and could also be inherited to the developed stem cell. In galline species, throughout  
26 example of H3K4m3 and H3K27me3 ChIP-seq analysis was still not provided. We therefore designed and  
27 demonstrated such procedures using ChIP-seq and mRNA-seq data of chicken follicular mesenchymal cells and  
28 male germ cells.

29 **Methods:** Analytical workflow was designed and provided in this study. ChIP-seq and RNA-seq datasets of  
30 follicular mesenchymal cells and male germ cells were acquired and properly preprocessed. Peak calling by  
31 MACS2 was performed to identify H3K4m3 or H3K27me3 enriched regions (Fold-change  $\geq 2$ , FDR  $\leq 0.01$ ) in  
32 gene promoter regions. Integrative Genomics Viewer (IGV) was utilized for CRAB1, GDF10, and GREM1 gene  
33 explorations.

34 **Results:** The acquired results indicated that follicular mesenchymal cells and germ cells shared several unique  
35 gene promoter regions enriched with H3K4me3 (5,704 peaks) and also unique regions of bivalent  
36 H3K4m3/H3K27me3 shared between all cell types and germ cells (1,909 peaks). Subsequent observation of  
37 follicular mesenchyme-specific genes—CRAB1, GDF10, and GREM1 correctly revealed vigorous transcriptions  
38 of these genes in follicular mesenchymal cells. As expected, bivalent H3K4m3/H3K27me3 pattern was  
39 manifested in gene promoter regions of germ cells, and thus suspended their transcriptions.

40 **Conclusions:** According the results, an example of chicken H3K4m3/H3K27me3 ChIP-seq data analysis was  
41 successfully demonstrated in this study. Hopefully, the provided methodology should hereby be useful for  
42 galline ChIP-seq data analysis in the future.

43 **Keywords:** ChIP-seq; Follicular mesenchymal cell; Galline; Germ cell; Histone 3 trimethylation

44 **INTRODUCTION**

45

46 Functional cells of multi-cellular organisms maintain their mutual state by adopting the specific gene  
47 expression patterns. They hereby fix their transcriptional activities by imprinting particular chromatin structure  
48 into genome to allocate and control target gene regions—also recognized as the chromatin packaging [1]. In  
49 eukaryotic cell, a chromatin consisted of several units of 147 base pairs DNA twining around nucleosome  
50 formed by histone octamer—which is the molecule responsible for the chromatin packaging process. A histone  
51 octamer comprised of four basic core histones—H2A, H2B, H3 and H4 of which are targets for histone  
52 modification—methylation, phosphorylation and acetylation. When a precursor cell committed a lineage-specific  
53 development, it modified particular N-termini amino acids of specific core histones. This event thus results in  
54 proper chromatin packaging to control gene transcription pattern specific to its lineage-committed cell [1,2].

55 Different histone modification can relax or condense chromatin structures differently. Since each  
56 specific chromatin structures can affect the accessibility of gene transcription factors and RNA polymerase II  
57 differently at gene promoter region, they will finally result in different gene transcription pattern specific to the  
58 lineage-committed cell [1–3]. Among well-recognized histone modifications, trimethyl modification of histone 3  
59 (H3) at 4th lysine N-termini (H3K4me3) is very common in active gene transcriptions of several lineage-  
60 committed cells [4,5]. This phenomenon, however cannot conclusively comply with several genes in stem cells  
61 and germ cells, of which H3K4me3 regions are presented without transcription. Interestingly, such genes always  
62 demonstrate not only trimethylation of H3 on 4th lysine (H3K4me3), but also the 27th lysine (H3K27me3) in the  
63 same loci of N-termini—the bivalent H3K4m3/H3K27me3 chromatin or poised chromatin [2,6,7].

64 In the embryonic stem cells, the presence of H3K4me3 in the poised gene help marking the gene as the  
65 active one. However, the concurrent H3K27me3 repressed the transcription and thus guarantee the cells'  
66 pluripotent state [6,7]. Actually, such characteristics were captivantly proved to be inherited from germ cells  
67 even prior to fertilization. According to this reason, the study of H3K4me3/H3K27me3 bivalent (poised)  
68 epigenetic state in male germ cells became a very popular topic in epigenetic evolution study [2,8].

69 Chromatin immunoprecipitation (ChIP) assays with sequencing or ChIP-Seq is a sequencing technology  
70 generally utilized to study genome-wide histone 3 (H3) trimethylation—H3K4me3 and H3K27me3. In ChIP-  
71 Seq, the chromatin regions presented with modified histones of interest are immunoprecipitated and sequenced to  
72 study epigenetic control of gene expressions in several cell types [2,5,9,10]. Not only mammals, ChIP-Seq of H3  
73 trimethylation were also conducted in chicken (*Gallus gallus*) as the representative of galline species [2,11–13].  
74 These studies provided biologists and veterinarians the valuable insights into evolution and cell development of

75 galline species. Of note, an increase of H3K4me3 and H3K27me3 ChIP-seq databases in chicken also allowed  
76 both biologists and veterinarians a new opportunity to observe and meta-analyze ChIP-seq data of various cell  
77 types acquired from different experiments.

78 Analytical guidelines are generally concerned in human and mouse ChIP-Seq studies [10]. On the  
79 contrary, the issue was limitedly described in H3 trimethylation ChIP-Seq data analysis in galline species  
80 [2,12,13]. Since the procedures could affect the analytical results differently, a practical demonstration of such  
81 approach should prove beneficial, especially for those who are not acquainted with ChIP-Seq data analysis. Due  
82 to such necessity, we therefore demonstrated a chicken ChIP-seq meta-analysis in this study. The H3K4me3  
83 ChIP-Seq data of chicken follicular mesenchymal cells and male germ cells were acquired, and we also included  
84 H3K27me3 ChIP-Seq data of chicken germ cells to demonstrate the effect of bivalent H3K4me3/H3K27me3 on  
85 gene transcription.

86

## 87 **MATERIALS AND METHODS**

88

### 89 **Sample datasets**

90 As far as we knew, follicular mesenchymal cell was the only lineage-committed cell type with H3  
91 trimethylation ChIP-seq data available in public database. List of follicular mesenchymal cell and germ cell  
92 datasets used in this study was provided (Table 1). Follicular mesenchymal cells used in this study consisted of  
93 two populations isolated from median (Med) and lateral mesenchyme (Lat) of feather follicles, accordingly. Male  
94 germ cells used in this study derived from two time point during development: pachytene spermatocyte (PS)  
95 acquired during prophase of meiosis I, and round spermatid (RS) acquired after completing meiosis. Both germ  
96 cell types were representatives of male germ cells to illustrate the poised genes—at which their poised states  
97 were already known to retain throughout spermatogenic development [2]. For ChIP-seq data, the sample file and  
98 its corresponding control file were provided together. Of note, the control file was not required for RNA-seq data  
99 analyses, and thus not provided.

100

### 101 **Analytical workflow**

102 In this study, we categorized our analytical processes into 3 steps—data preprocessing, peak calling  
103 with visualization, and Integrative Genomics Viewer (IGV) exploration (Fig 1). In brief, the sample datasets  
104 would be aligned to chicken genome in order to acquire ChIP files and RNA-seq files in the data preprocessing  
105 step (Fig. 1A). The ChIP files were then applied in peak calling step to identify genome regions enriched with

106 H3 trimethylation—the significant peaks (Fig 1B). Only peaks overlapped with gene promoter region would be  
107 considered in IGV exploration of which, the enriched peaks along with the mapped sequences ChIP and RNA-  
108 seq files were integrated and visualized together in IGV platform (Fig 1C).

#### 109 *Data preprocessing*

110 The Sequence Read Archive (SRA) files of germ cells and follicular mesenchymal ChIP-Seq and  
111 RNA-Seq datasets were retrieved from SRA database (<https://www.ncbi.nlm.nih.gov/sra>) (Table 1),  
112 subsequently extracted, and preprocessed as previously described [14]. The ChIP files were aligned to chicken  
113 genome (*Gallus gallus* 5.0) by Bowtie2 aligner [15], while STAR aligner was applied with the RNA-seq files  
114 [14,16]. The aligned files were converted into “bam” format, indexed by gene annotation, removed for  
115 duplicated sequences [17], and bias corrected for GC base balance [18] to acquire ChIP files and RNA-seq files.  
116 The qualities of both raw and preprocessed datasets were then determined by FastQC as previously described  
117 [14]. For ChIP quality determination, Normalized Strand Cross-correlation coefficient (NSC), Relative Strand  
118 Cross-correlation coefficient (RSC) and DNA fragment length were additionally included. In more details, NSC  
119 is genome-wide correlation between positive and negative strand read counts when shifted by half of the  
120 fragment length relative to background. NSC value thus represents enrichment of clustered ChIP fragments  
121 around target sites. RSC is the relative enrichment of fragment-length cross-correlation to read-length cross-  
122 correlation. The RSC thus represented the clustering of relatively fixed sized fragment around target sites.

123

#### 124 *Peak calling with visualization*

125 Peak calling was, in case of this study, the computational method to identify genome-wide chromatin  
126 regions with targeted H3 trimethylation significant to the other regions. The peaks therefore represented the  
127 genome area with significant presences of H3K4me3 or H3K27me3 when compared to the same region  
128 presented in the corresponding control datasets (Table 1). Model-based analysis of ChIP-seq 2 (MACS2)  
129 algorithm [19] was used in this study. Since the broad peak regions were suggested in the previous quality  
130 analysis, we applied broad peak calling function in MACS2 with all ChIP files. Further details for such function  
131 calling was already described [19]. Significant peaks were those with False Discovery Rate (FDR)  $\leq 0.01$  and  
132 Fold-change  $\geq 2$  presented with chicken gene promoter regions (2000 base pairs upstream to 200 base pairs  
133 downstream of gene promoter) (Peak files in Fig 1B). Peak files were also utilized for counting aggregated DNA  
134 within peak region, and counting the DNA fragments distributed from the transcription start site of chicken genes  
135 (Fig 2B).

136

### 137 *Integrative Genomics Viewer (IGV) exploration*

138           The Peak files along with their corresponding ChIP files and RNA-seq files were integratively observed  
139 together in Integrative Genomics Viewer (IGV) platform (Robinson et al., 2011) (Fig 1C). The Gal\_gal5.0  
140 genome (UCSC version galGal5 released in December 2015) was administrated as the reference genome. The  
141 gal\_Gal5.0 consisted of 34 chromosomes, 1 linkage group and 15,411 unplaced scaffolds. Further details of IGV  
142 platform could be acquired online (<http://software.broadinstitute.org/software/igv/>). In brief, DNA aggregations  
143 along the genome of sample ChIP files were expressed by blue heatmap. Similarly, the density of transcript  
144 fragments from RNA-seq files acquired from the corresponding cell sample were expressed by red heatmap.  
145 Peak regions of each cell type were also included in the figure the aggregated DNA fragments from ChIP files.  
146 To demonstrate bivalent H3K4me3/H3K27me3 chromatin of germ cell samples, CRAB1, GDF10 and GREM1  
147 genes were chosen for illustration (Fig 4). We select such genes due to the fact that transcription of these genes  
148 was crucial in topological control of chicken feather's generation of follicular mesenchyme [11].

149

## 150 **RESULTS**

151

### 152 **All sample datasets were successfully preprocessed and aligned to chicken genome.**

153           The raw datasets of all cell types were retrieved and preprocessed according to the methodology (Fig  
154 1A). All preprocessed datasets (both ChIP-seq and RNA-seq data in Table 1) acquired quality scores across all  
155 bases  $\geq 30$  without presence of contaminated adapters. Alignment percentages of ChIP-seq and RNA-seq  
156 datasets were  $\geq 85\%$  and  $\geq 70\%$ , accordingly. Normalized Strand Cross-correlation coefficient (NSC), Relative  
157 Strand Cross-correlation coefficient (RSC) and DNA fragment length of ChIP-seq samples were provided in  
158 Table 2.

159

### 160 **Follicular mesenchymal cells and germ cells shared several gene promoter regions with H3K4me3 and** 161 **H3K27me3.**

162           As previously described in materials and methods, the peaks were the chromatin loci in gene promoter  
163 regions enriched with target H3 trimethylations (H3K4me3 or H3K27me3). When categorized the acquired  
164 peaks according to their localized chromosomes, chromosome 1 contained the largest number of peaks  
165 comparing to the others (chr1 in Fig 2A). Most of aggregated DNA fragments of ChIP files were found  
166 distributed within 0-1 kilobases from the transcription start sites in all cell types (Fig 2B). The results indicated  
167 that that most H3K4me3 peaks (5,704 peaks) were uniquely shared among follicular mesenchymal cells and

168 germ cells (lat\_K4, med\_K4, PS\_K4 and RS\_K4), and the second largest intersected peaks were uniquely shared  
169 among all H3K4me3 and H3K27me3 (1,909 peaks) (Fig 2A, Fig 3).

170

171 **CRAB1, GDF10 and GREM1 genes of chicken germ cells adopted the bivalent H3 trimethylation**  
172 **chromatin pattern with presence of transcription's inhibition.**

173 Follicular mesenchymal cells (Lat RNA-seq and Med RNA-seq in Fig 4) expressed CRAB1, GDF10  
174 and GREM1 genes higher than germ cells (PS RNA-seq and RS RNA-seq in Fig 4). Of note, the bivalent  
175 H3K4me3/H3K27me3 pattern near transcription start site of these gene in germ cells was demonstrated by the  
176 significant peak regions (PS\_K4 Peak, PS\_K27 Peak, RS\_K4 Peak and RS\_K27 Peak in Fig 4) with presence of  
177 remarkable ChIP-seq DNA aggregations (PS\_K4 ChIP-seq, PS\_K27 ChIP-seq, RS\_K4 ChIP-seq, and RS\_K27  
178 ChIP-seq in Fig 4).

179

## 180 **DISCUSSION**

181

182 Data quality was among the most critical issue in ChIP-seq analysis due to its later effect on  
183 measurement and comparison. According to this reason, we preprocessed the ChIP-seq data to acquire adequate  
184 quality and alignment percentages. Another crucial issue in ChIP-seq quality was the determination of signal-to-  
185 noise ratio represented by NSC and RSC values. ChIP-seq data with few genuine binding site regions usually  
186 rendered high NSC ( $\geq 1.5$ ) and RSC ( $\geq 0.8$ ), accordingly. However, marks that tend to be enriched at repeat-like  
187 regions with diffused genome-wide patterns like H3 trimethylation usually show lower NSC and RSC values—  
188 by which also presented in our study [10,11] (Table 2).

189 Since H3 trimethylation (H3K4me3 or H3K27me3) at gene promoter regions were strongly associated  
190 with gene transcription control [4,5,12], only peaks within gene promoter region were hereby considered in this  
191 study. Inheritance mechanism of H3K4me3/H3K27me3 bivalent state from germ cells to embryonic stem cells,  
192 and activation mechanism of lineage-specific genes in lineage-committed cells were comprehensively described  
193 elsewhere [20], and were not included in this discussion. According to the result, most H3 trimethylation peaks  
194 in both follicular mesenchymal cells and germ cells were located in chromosome 1—agreeing with the largest  
195 gene numbers in the chromosome (2,162 coding genes and 524 non-coding genes) (chr1 in Fig 2). Interestingly,  
196 the largest number of intersected H3K4me3 peaks among ChIP-seq samples (orange bar in Fig 3) implied the  
197 conserved role of H3K4me3 among chicken germ cells and mesenchymal cells. Of note, shared presented among  
198 H3K4me3 and H3K27me3 (red bar in Fig 3) could only partially imply several unique genes controlled by

199 H3K4me3/H3K27me3 bivalent state in chicken germ cells due to the lack of H3K27me3 ChIP-Seq datasets from  
200 analysis.

201 To demonstrate the effect of H3K4me3/H3K27me3 bivalent state on transcription inhibition of germ  
202 cells [2,7], we chose CRAB1, GDF10 and GREM1 as candidate genes due to their redundant lineage-specific  
203 transcriptions in follicular mesenchymal cells [11]. Visualization by IGV properly demonstrated such lineage-  
204 specific characteristics in follicular mesenchymal cells. On the contrary, inhibition of the gene transcription was  
205 obvious in germ cells with significant presences of H3K4me3/H3K27me3 (poised) in the gene promoter regions  
206 as expected (Fig 4). As described in the introduction, such poised state of genes in germ cells allowed them to  
207 inherit pluripotency to the later developed embryonic stem cells. The observatory procedure in this study hereby  
208 successfully demonstrated the inhibitory effect of H3K4me3/H3K27me3 bivalent (poised) state on follicular  
209 mesenchyme-specific genes in chicken germ cells. By mean of this, other poised genes of interest could also be  
210 observed by adopting the same approach.

211 In conclusion, the current study provided an example for H3 trimethylation ChIP-seq analysis of  
212 chicken (*Gallus gallus*) cells. Since the methodology required no specific process, ChIP-seq data of other histone  
213 modifications could also be analyzed using similar approach. There were, however some important limitations  
214 that should be noted in this study. These included the lack of H3K27me3 ChIP-seq data of follicular  
215 mesenchymal cells to confirm abortion of poised genes, and limited samples per cell type available for valid  
216 differential analysis [21]. Hopefully, the rapid growth of galline ChIP-seq database would allow us to  
217 demonstrate such analysis in our future study.

218

#### 219 **CONFLICT OF INTEREST**

220 We certify that there is no conflict of interest with any financial organization regarding the material and  
221 methods discussed in the manuscript.

222

#### 223 **ACKNOWLEDGEMENTS**

224 The current study was fully supported by Thailand Research Fund (TRF) through New Research  
225 Scholar Program (Grant No. TRG5880003)

226

#### 227 **REFERENCES**

- 228 1. Feil R, Khosla S. Genomic imprinting in mammals: An interplay between chromatin and DNA  
229 methylation? Trends Genet, 1999; 15(11): 431–5.

- 230 2. Lesch BJ, Silber SJ, McCarrey JR, Page DC. Parallel evolution of male germline epigenetic poisoning and  
231 somatic development in animals. *Nat Genet*, 2016; 48(8): 888–94.
- 232 3. Waclawska A, Kurpisz M. Key functional genes of spermatogenesis identified by microarray analysis.  
233 *Syst Biol Reprod Med*, 2012; 58(5): 229–35.
- 234 4. Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. A Chromatin Landmark and Transcription  
235 Initiation at Most Promoters in Human Cells. *Cell*, 2007; 130(1): 77–88.
- 236 5. Zhang B, Zheng H, Huang B, Li W, Xiang Y, Peng X i wsp. Allelic reprogramming of the histone  
237 modification H3K4me3 in early mammalian development. *Nature*, 2016; 537(7621): 553–7.
- 238 6. Azuara V, Perry P, Sauer S, Spivakov M, Jørgensen HF, John RM i wsp. Chromatin signatures of  
239 pluripotent cell lines. *Nat Cell Biol*, 2006; 8(5): 532–8.
- 240 7. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G i wsp. Genome-wide maps of  
241 chromatin state in pluripotent and lineage-committed cells. *Nature*, 2007; 448(7153): 553–60.
- 242 8. Villar D, Berthelot C, Aldridge S, Rayner TF, Lukk M, Pignatelli M i wsp. Enhancer evolution across 20  
243 mammalian species. *Cell*, 2015; 160(3): 554–66.
- 244 9. Eissenberg JC, Shilatifard A. Histone H3 lysine 4 (H3K4) methylation in development and  
245 differentiation. *Dev Biol*, 2010; 339(2): 240–9.
- 246 10. Landt S, Marinov G. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia.  
247 *Genome ...*, 2012(Park 2009): 1813–31.
- 248 11. Li A, Figueroa S, Jiang T-X, Wu P, Widelitz R, Nie Q i wsp. Diverse feather shape evolution enabled by  
249 coupling anisotropic signalling modules with self-organizing branching programme. *Nat Commun*, 2017;  
250 8: ncomms14139.
- 251 12. Jahan S, Xu W, He S, Gonzalez C, Delcuve GP, Davie JR. The chicken erythrocyte epigenome.  
252 *Epigenetics Chromatin*, 2016; 9(1): 19.
- 253 13. Luo J, Mitra A, Tian F, Chang S, Zhang H, Cui K i wsp. Histone methylation analysis and pathway  
254 predictions in chickens after MDV infection. *PLoS One*, 2012; 7(7).
- 255 14. Chokeshaiusaha K, Thanawongnuwech R, Puthier D, Nguyen C. Inspection of C-type lectin superfamily  
256 expression profile in chicken and mouse dendritic cells. *Thai J Vet Med*, 2016; 46(3): 443–53.
- 257 15. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*, 2012; 9(4): 357–9.
- 258 16. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S i wsp. STAR: Ultrafast universal RNA-  
259 seq aligner. *Bioinformatics*, 2013; 29(1): 15–21.
- 260 17. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N i wsp. The Sequence Alignment/Map  
261 format and SAMtools. *Bioinformatics*, 2009; 25(16): 2078–9.
- 262 18. Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. DeepTools: A flexible platform for exploring  
263 deep-sequencing data. *Nucleic Acids Res*, 2014; 42(W1).
- 264 19. Feng J, Liu T, Qin B, Zhang Y, Liu XS. Identifying ChIP-seq enrichment using MACS. *Nat Protoc*,  
265 2012; 7(9): 1728–40.
- 266 20. Vastenhouw NL, Schier AF. Bivalent histone modifications in early embryogenesis. *Curr Opin Cell Biol*,  
267 2012; 24(3): 374–86.
- 268 21. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ i wsp. Differential  
269 oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature*, 2012.

270 *Table 1 Raw data files used in this study*

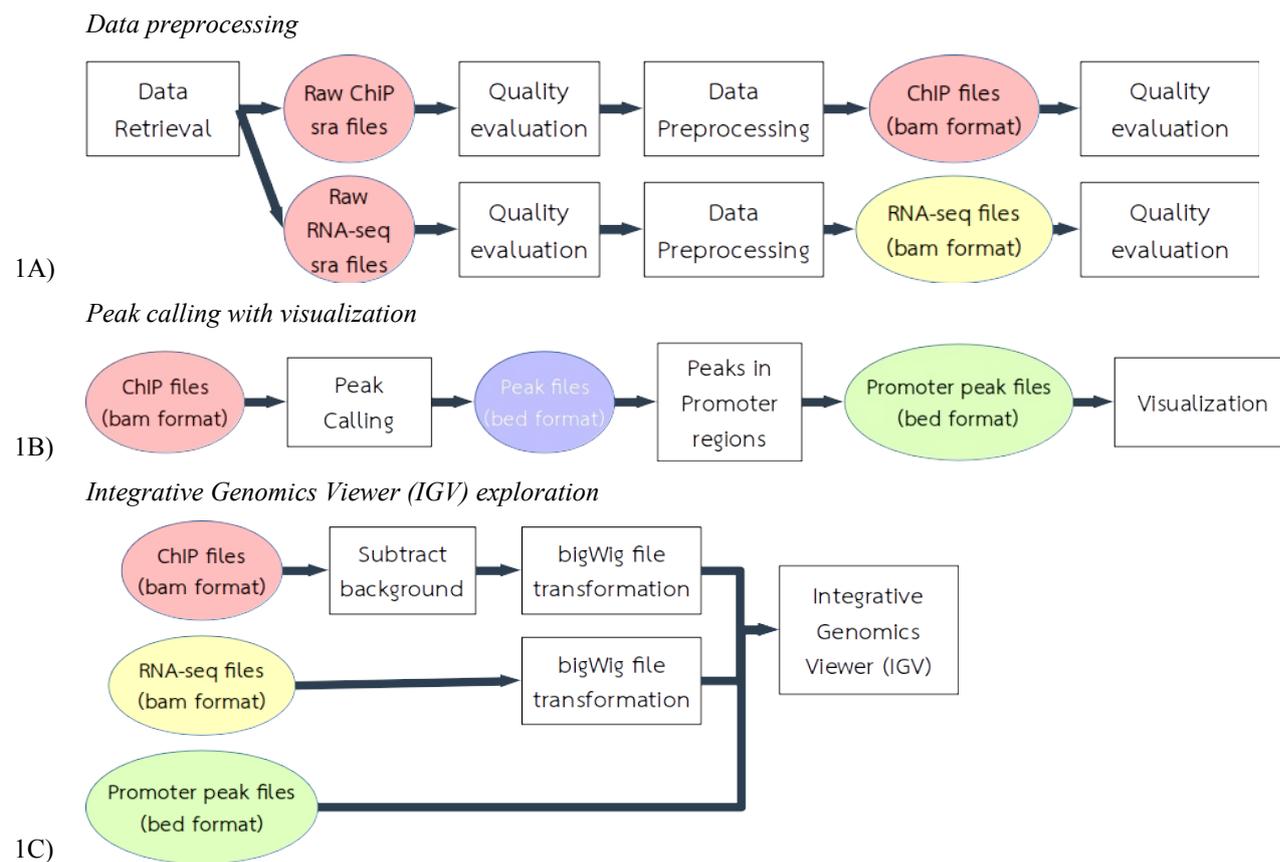
<i>Cell type's abbreviation</i>	<i>Description</i>	<i>Sample file</i>	<i>Control file</i>
PS_K4	ChIP-seq of H3K4me3 of pachytene spermatocyte	SRR1977505	SRR1977507
PS_K27	ChIP-seq of H3K27me3 of pachytene spermatocyte	SRR1977506	SRR1977507
PS_RNA-seq	RNA-seq of pachytene spermatocyte	SRR1977509	-
RS_K4	ChIP-seq of H3K4me3 of round spermatid	SRR1977535	SRR1977537
RS_K27	ChIP-seq of H3K27me3 of round spermatid	SRR1977536	SRR1977537
RS_RNA-seq	RNA-seq of round spermatid	SRR1977540	-
Lat_K4(1)	ChIP-seq of H3K4me3 of lateral mesenchyme	SRR4125205	SRR4125207
Lat_RNA-seq(1)	RNA-seq of lateral mesenchyme	SRR4125190	-
Med_K4(1)	ChIP-seq of H3K4me3 of medial mesenchyme	SRR4125206	SRR4125208
Med_RNA-seq(1)	RNA-seq of medial mesenchyme	SRR4125192	-
Lat_K4(2)	ChIP-seq of H3K4me3 of lateral mesenchyme	SRR4125209	SRR4125211
Lat_RNA-seq(2)	RNA-seq of lateral mesenchyme	SRR4125191	-
Med_K4(2)	ChIP-seq of H3K4me3 of medial mesenchyme	SRR4125210	SRR4125212
Med_RNA-seq(2)	RNA-seq of medial mesenchyme	SRR4125193	-

271

272 *Table 2 NSC values, RSC values and DNA fragment length of ChIP-seq samples*

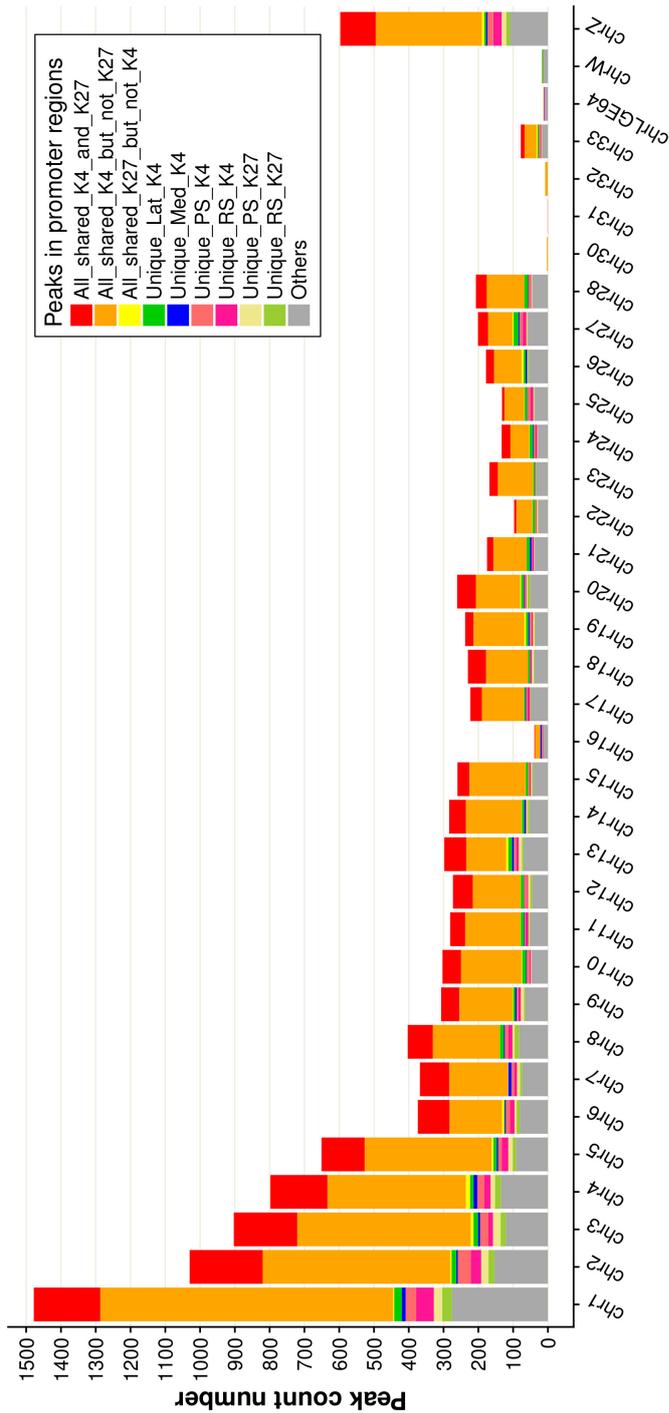
<i>Cell type</i>	<i>Sample file</i>	<i>NSC</i>	<i>RSC</i>	<i>Fragment</i>
				<i>Length</i>
PS_K4	SRR1977505	1.10	1.05	170
PS_K27	SRR1977506	1.02	1.28	170
RS_K4	SRR1977535	1.13	1.10	170
RS_K27	SRR1977536	1.03	1.25	170
Lat_K4_1	SRR4125205	1.02	0.72	210
Med_K4_1	SRR4125206	1.01	0.64	200
Lat_K4_2	SRR4125209	1.02	0.71	240
Med_K4_2	SRR4125210	1.01	0.66	195

273

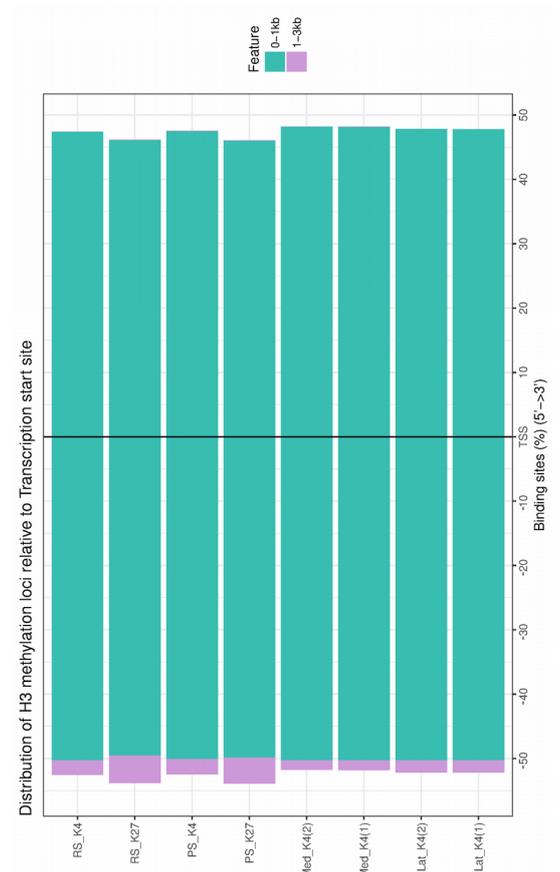


274

275 **Figure 1.** Analytical workflow for this study. Analytical processes were categorized into 3 steps—data  
 276 preprocessing (1A), peak calling with visualization (1B), and Integrative Genomics Viewer (IGV) exploration  
 277 (1C). The files required and outcome for each analytical step were provided.



2A)



2B)

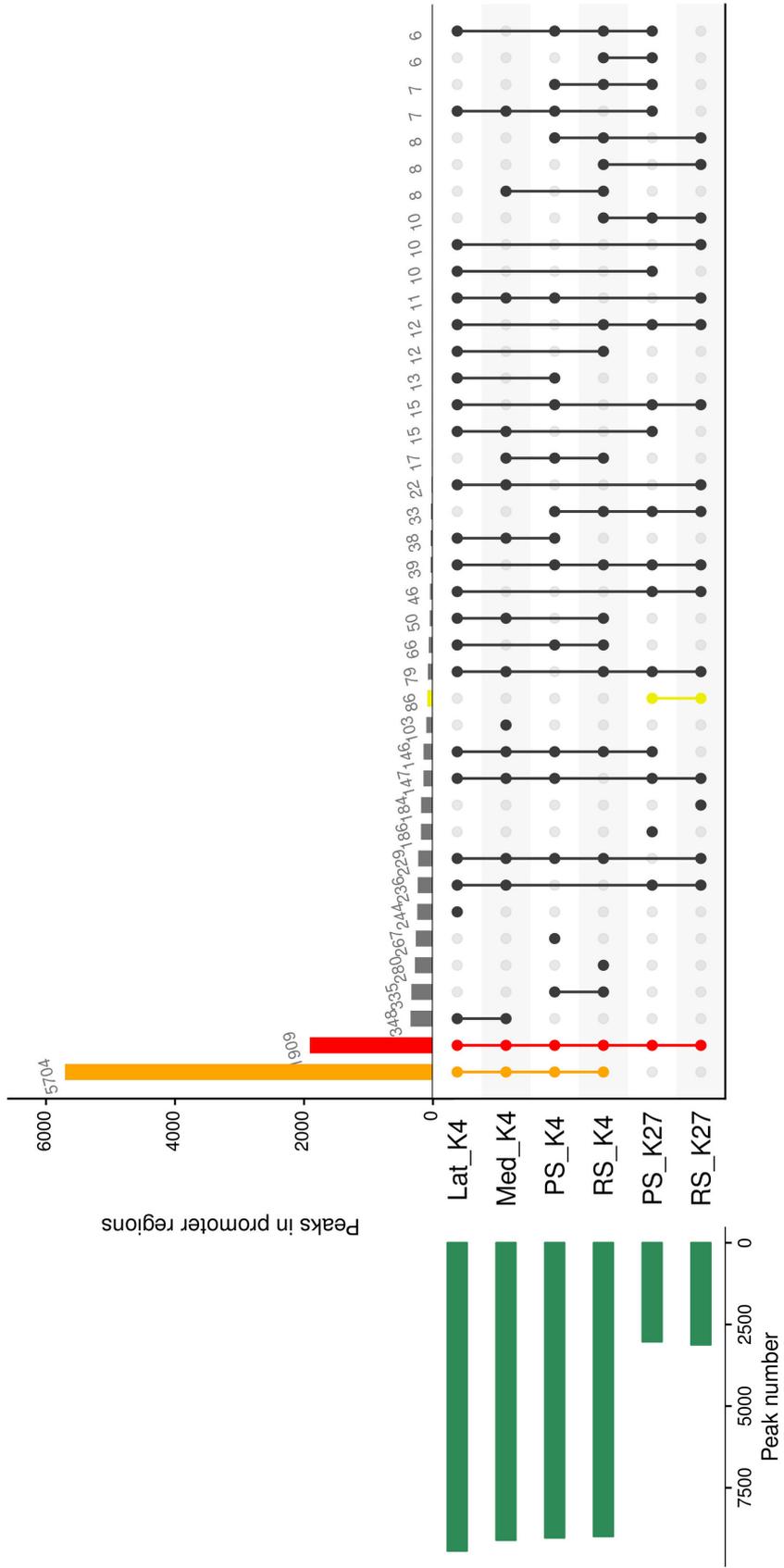
278

279 **Figure 2.** H3 trimethylation loci. The total number of peaks significantly identified were categorized according

280 to their regional chromosomes (2A). Each color bar presented in the legend was described as following: ■

281 Peaks shared among all samples with either K4 or K27 trimethylation (All\_shared\_K4\_and\_K27); ■-Peaks

282 presented only among samples with K4 trimethylation (All\_shared\_K4\_but\_not\_K27); -Peaks presented only  
283 among samples with K27 trimethylation (All\_shared\_K27\_but\_not\_K4); -Peaks presented only in Lat\_K4  
284 (Unique\_Lat\_K4); -Peaks presented only in Med\_K4 (Unique\_Med\_K4); -Peaks presented only in PS\_K4  
285 (Unique\_PS\_K4); -Peaks presented only in RS\_K4 (Unique\_RS\_K4); -Peaks presented only in PS\_K27  
286 (Unique\_PS\_K27); -Peaks presented only in RS\_K27 (Unique\_RS\_K27), and -Other peaks not in the  
287 previously mentioned groups. Distribution of precipitated DNA fragments acquired from sample ChIP-seq data  
288 according to chicken gene transcription start site was also demonstrated (2B). Aggregation of fragments within  
289 0-1 kilobases () and 1-3 kilobases () were presented.



291 **Figure 3.** UpSet plot of intersected peaks among cell types. The intersected peaks shared among cell types  
292 (Lat\_K4, Med\_K4, PS\_K4, RS\_K4, PS\_K27 and RS\_K27) were presented by UpSet plot. The connected lines  
293 among cell types shown in the lower panel of the plot represented the group of intersected peaks among them.  
294 The upper panel bar graph of the plot presented the number of peaks found in each group. The shared peak  
295 number among all samples with either K4 or K27 trimethylation and only K4 trimethylation were in red (■) and  
296 orange (■) color, accordingly. The horizontal leftmost bar graph (olive green color—■) show to total number  
297 of peaks presented in each cell type.

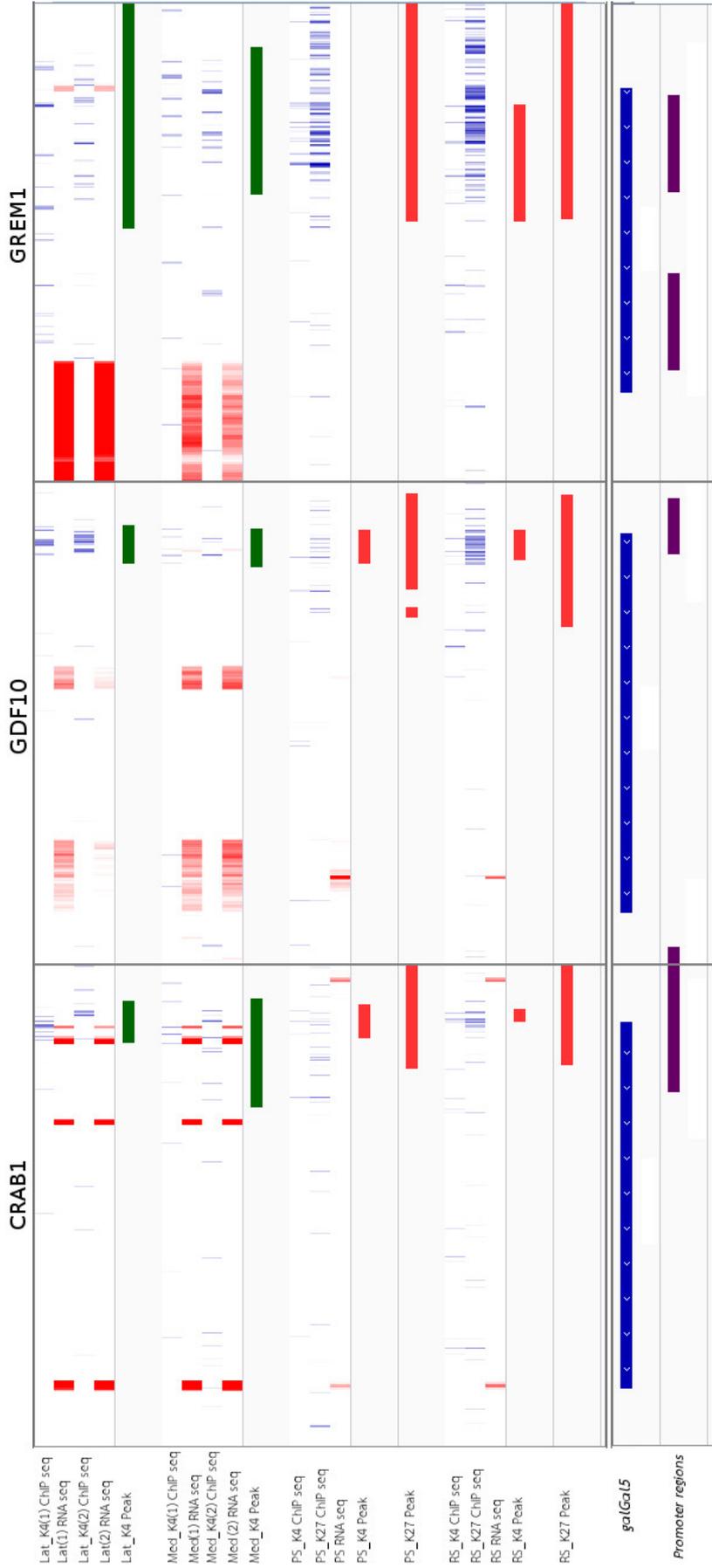


Figure 4

298

299 **Figure 4.** Chromatin H3 trimethylation state and transcription of GDF10, CRAB1 and GREM1 genes. The  
300 heatmaps of CRAB1, GDF10 and GREM1 genes were acquired from the IGV interface. The chicken full  
301 genome sequences was used to identify the chromatin loci of these genes with transcription direction indicated  
302 by the arrow heads presented in the genome along with the associated promoters in the area (galGal5 and  
303 Promoter regions at the bottom of the figure). The figure show 3 types of data—ChIP-seq, RNA-seq, and Peak  
304 files acquired from each sample cell type. In brief, ChIP-seq data show the density of DNA fragments aligned to  
305 a chromatin region (aggregated DNA fragments) in blue heatmap. The RNA-seq data show the density of  
306 transcript fragments aligned to the gene regions representing level of gene expressions by red heatmap. Finally,  
307 The Peak file of each cell sample indicated the chromatin region with significant peaks identified by peak calling  
308 analysis.