Doctoral Dissertation

Studies on multifunctional genome elements functioning in mouse spermatogenesis

（マウス精子形成で機能する多機能性ゲノムに関する解析）

Atthanayake Mudiyanselage Thusitha Kosala Bandara

Graduate School of Life Science, Hokkaido University

2021 March

Contents

Abbreviations……01

General introduction…….02

Chapter1……11

Chapter2……59

Chapter3……88

General discussion……118

Acknowledgements……...124

Reference……...125

Abbreviations

ncRNAs	non-coding RNAs
piRNAs	PIWI-interacting RNAs
lncRNAs	long non-coding RNAs
RNAPII	RNA polymerase II
ChIP-seq	Chromatin immunoprecipitation sequencing
DPE	Dual promoter-enhancer
DES	Dual enhancer-silencer
TSS	Transcription Start site
FBS	Fetal Bovine Serum
TK	Thymidine Kinase
gRNA	Guide RNA
qRT-PCR	Quantitative real-time polymerase chain reaction
H3K4me1	Histone H3 lysine 4 monomethylation
H3K4me3	Histone H3 lysine 4 trimethylation
H3K27ac	Histone H3 lysine 27 acetylation
H3K9	Histone H3 lysine 9
H3K27	Histone H3 lysine 27
Dpp	Days postpartum

General introduction

General introduction

In an abstract definition the genome is the conveyer of genetic information which composed of single dimensional units with linear DNA sequences. In reality, it is vastly complex. Most genomes consist of the DNA helix that is wrapped around octameric histone protein complexes that eventually form a chromosome. This spatial effective compacted arrangement of the chromatin fibers dramatically affects the functional regulation of genomic activities. The well-acknowledged key functions of the genome are to store and propagate the genetic material and to regulate the expression and repression of the genes, the hereditary information encoded by DNA. Most of the organisms are made up of numerous types of cells that contain similar sets but differentially activated protein-coding genes which carry both coding and non-coding DNA sequences in the genome. The on-demand regulation process of genes corresponding to different cellular environments requires sophisticated functional machinery which involves a myriad of proteins including transcription factors and RNA-processing factors [1].

The diversity in gene expression is regulated by a variety of regulatory elements or factors present inside the cell. Traditionally, there are several types of gene regulatory elements in the genome such as promoters, enhancers, and silencers [2]. Promoters are the genomic regions that are associated with RNA polymerase to start gene transcription [3], and enhancers are the genomic regions that facilitate the promoter activity in distance and play major roles in controlling the specificity of gene expression. In some cases, one gene can be controlled by multiple enhancers and in other cases, a single enhancer can regulate multiple genes [4]. Silencers are the negative regulatory element that represses gene

expression, although the knowledge about this element remains less due to a limited number of functional assays [5].

Besides, some genome regions are known to have multifunction. For example, a genome element that possesses both promoter and enhancer is called dual promoter-enhancer (DPE) or epromoter. One of the first reported DPE region is related to the chicken lysozyme gene which is around 1.9 kb [6]. This was originally identified as a hormone responsive enhancer element and later discovered that functionality as a promoter for a long non-coding RNA (lncRNA). Later research identified that promoters exhibit enhancer activity in cancer cells and they observed transcription start site (TSS) overlapping regions that show enhancer activity and termed as epromoters. They claim that 2% -3% of regions of the human genome represent this type of epromoters, virtually the same as DPEs [7].

The concurrent process of regulating the gene expression patterns without changing the DNA sequence is known as an epigenetic mechanism. Epigenome is the cumulation of such modifications and the epigenomic modifications lead to activation or inactivation of genes by controlling the chromatin structure [8]. These changes can have short or long-term effects and could be trans-generational [9]. Therefore epigenetic alteration can be crucial when determining which and when genes are expressed in cells. As an example even though, genetic code is similar in every cell of the organism, they show tissue-specific and stimuli specific responses differently. This may be due to the dynamic epigenetic code in the cells.[10].

The epigenetic machinery is organized with heritable but potentially reversible chemical or physical modifications of chromatin which include three interconnected processes *viz.* DNA methylation, histone modification, and RNA-based methods [11,12]. Methylation of DNA is a primary epigenetic mechanism in which a methyl or hydroxymethyl group is added to the C5 position of cytosine nucleotide in the CpG (cytosine–guanine) dinucleotide island. CpG islands are the genomic regions that contain very condensed CG nucleotides, more than 55% [13]. Many CpG sequences are found in promoter regions in mammals [14], and the methylation inheritably silences gene expression at the transcriptional level. The CpG methylation is generated by DNA methyltransferase to prevent transcription factors from binding to promoter regions, which leads to the suppression of gene expression [15].

In the nucleus, DNA is condensed into chromatin and the building blocks of the chromatin are nucleosomes. These nucleosomes form a “beads-on-a-string” structure which composed of 146 base pairs of DNA wrapped around an octamer of core histones, including each two molecules of H2A, H2B, H3, and H4. As the basic component in chromatin composition, these nucleosomes participate in altering the structure of chromatin in response to specific stimuli [11]. Histones contain protein motifs rich in lysine and arginine. By such simple chemical modifications, histone changes the DNA binding capacity and the interaction with other regulatory factors. This type of changes leads to the alteration in gene activation. In general, acetylation of H3 and H4 leads to open chromatin configuration and facilitates the binding of transcription factors to activate gene transcription [16]. Histone acetylation can be found in the region where transcripts are actively generated [17]. Conversely, methylation of histone H3 lysine 9 (H3K9) and histone

H3 lysine 27 (H3K27) is associated with gene inactivation [18] while H3K4 methylation is associated with activating promoter and enhancer regions for active genes.

RNA-based epigenetic mechanisms are interconnected with the other two epigenetic processes via noncoding RNAs (ncRNA). Escalating progression in genome mining technologies has identified roles of ncRNAs including microRNAs, PIWI-interacting RNAs (piRNAs), endogenous small-interfering RNAs, and long non-coding RNAs (lncRNAs) as regulators of gene expression at transcriptional and post-transcriptional levels in various biological processes [19,20]. lncRNAs compose the large majority of ncRNAs that characteristically lack an open reading frame with more than 200 nucleotides and many of them are transcribed by RNA polymerase II (RNAPII) [21]. Though the number of functionally well-characterized lncRNAs is still limited, a recent publication has categorized the lncRNAs based on their regulatory mechanisms: 1) as “decoys” that bind to target DNA forming RNA-DNA complex and sequestering the target protein, 2) as “guides” that bind the RNA-binding proteins to a target genomic region to regulate gene expression and 3) as “scaffolds” that assemble two or more proteins using multiple effector components with different domains [22].

It is the common understanding that an underlying pathological complication of a biological process can be understood by understanding the biological process itself [23]. Male infertility has become a complex pathological problem [23] and there are much to be understood in the mechanism of action behind the process of spermatogenesis. Recently, many research groups are focusing on the molecular level basis of the male infertility [24] but still could not explain the causality due to a lack of proper understanding of the

mechanisms behind spermatogenesis. One of the promising approaches to understanding the molecular mechanisms behind spermatogenesis is the use of epigenetics [25]. Recent research suggests that some epigenetic modifications such as DNA methylation and histone modifications are vital in the regulation of gene expression during spermatogenesis.

The procurement of DNA methylation in male germ cells progresses after birth in the mitotic and meiotic germ cells until the pachytene phase of meiosis. It is vital to maintain the acquired methylation patterns in the mitotic spermatogonia and transition from spermatogonia to spermatocytes in which both loss and gain of methylation take place [26,27]. It is reported that male germ cells of adult mice contain highly distinguishable genome-wide DNA methylation patterns [28,29]. In addition, testicular DNA consists of an eight times higher number of hypermethylated loci compared to the somatic tissue [30,31]. DNA methyltransferases are known to be responsible for achieving this methylation pattern during spermatogenesis. Indeed, recent studies have shown that in sperms, promoter regions of genes for development were hypomethylated compared to somatic cells, and repeat sequences were highly methylated while transposon was weekly methylated [25,32–35].

During the post-meiotic maturation of male haploid germ cells, histones are replaced with transition proteins and then with protamines. Based on their properties there are two major types of these small basic proteins as TP1 and TP2 [36,37]. Though the complete mechanism yet to be explored, the published research proposes that the different types of histone modifications including synthesizing, removing, and replacing are carried out during spermatogenesis [38,39]. Current research indicates that histone plays a crucial

role in the determination of epigenomic position and they do not distribute in the genome randomly [39,40]. Due to this, studying histone modifications related to spermatogenesis would reveal the recurrent gene activation and silencing mechanisms of the genome.

Advancement of novel techniques has accelerated the identification of novel regulatory elements in genome-wide scale. One such innovation is chromatin immunoprecipitation followed by massively parallel sequencing (Chromatin immunoprecipitation-sequencing, ChIP-seq) [41]. This technique can be used to determine many binding sites of a transcription factor at a time, and by using antibodies against specific types of modified histones, can determine the global epigenetic patterns.

With these novel advancements, our group found a DPE at the mouse *Tcam1* locus, which drives *Smarcd2* and *lncRNA-Tcam1* transcription and enhances *Tcam1* transcription in primary spermatocytes [42]. This DPE is marked with both Histone H3 lysine 4 monomethylation (H3K4me1) and Histone H3 lysine 4 trimethylation (H3K4me3) and showed promoter and enhancer activity by reporter gene assay as well as in the chromatin context. Our group also found that the corresponding genome region in human possesses DPE activity too, although its promoter activity is unidirectional and tissue-specificity of enhancer activity was weaker [43].

Spermatogenesis is regulated by many genes especially when producing primary spermatocytes. During this stage, it was identified that significantly increased number of genes are express compared to the other stages of cell proliferation during the mammalian life span. Although the mechanism of transcriptional regulation during this stage remains

elusive. Due to this elusive regulatory mechanism, it requires extensive research to understand this transcription regulatory mechanism. As a result of identifying the transcription regulatory mechanism previously, our group was able to identify multifunctional genome regulatory element which exerts both promoter and enhancer activity. In this research our major attempt to identify such regions in genome-wide relation to spermatogenesis and to check the significance relation towards the especially when producing primary spermatocyte. As well as enhancer silencer elements are reported as the type of transcription regulatory element. Here I attempt to identify such regions which can be another type of regulatory element that possibly related to spermatogenesis. This was done based on the previously reported histone modification mark and through genome edition hence genome editing has become a promising method to identify the silencer activity.

Ensuing, as described in the first two chapters, I tried to identify the genomic region which can act as both promoter and enhancer related to spermatogenesis. For this purpose, I used the epigenomic profile marker arrangement in primary spermatocytes, spermatogonia, and GC-2spd(ts) cells. At the first attempt, I checked the genomic region which contains both H3K4me1 and H3K4me3 responsible for promoter and enhancer. However, I could only detect enhancer activity. Due to this, I assayed the genomic region which contains H3K4me1 within the 500 bp of TSS. According to my results, I conclude that such regions related to spermatogenesis can act as both promoters and enhancers.

In the third chapter, I discuss a genomic region identified in association with both H3K4me1 and H3K4me3 marks and found it to be a novel type of regulatory element ‘dual

enhancer-silencer'. Initially, I was able to observe only enhancer activity in this region, but subsequently, by using the genome editing technology in GC-2spd(ts) cells, I observed that this region also possesses silencer activity. Interestingly, I identified that this region is rich with histone modification patterns linked to silencer activity.

In conclusion, during this research, I identified two different types of multifunctional genome elements related to spermatogenesis and revealed the significance of gene regulation by those multifunctional elements during spermatogenesis. I suggest these findings paves the way to possible update towards the traditional definition of promoter and enhancer dichotomy and understanding the molecular mechanism underlying spermatogenesis.

Chapter-1

A genome-wide identification of dual promoter–enhancer candidates during Spermatogenesis

Abstract

Mammalian spermatogenesis is a sequential continuation of cellular differentiation to generate functional sperm from an initially undifferentiated germ cell. This process is regulated by different types of genes especially in producing primary spermatocytes, but the molecular mechanism of gene activation during spermatogenesis is not fully understood. Our group found a dual promoter-enhancer (DPE) element which has both promoter and enhancer activity in mouse primary spermatocytes. Here I hypothesized that more DPEs function in spermatogenesis and performed genome-wide analyses to identify DPE candidates. By chromatin immunoprecipitation-sequencing (ChIP-seq), I initially identified three genome regions marked with both H3K4me1 and H3K4me3, but they did not show promoter activity. Then, I searched for genome regions around transcriptional start sites (TSSs) that were associated with H3K4me1 by ChIP-seq and RNA-sequencing and identified 337, 2176, and 69 DPE candidates in spermatogonia, primary spermatocytes, and GC-2spd(ts) cells, respectively. I selected 13 candidates and checked transcriptional activity using reporter gene analysis. The reporter gene assay revealed that 8 out of 13 candidates showed significant enhancer activity while 10 showed promoter activity.

Keywords: dual promoter-enhancer, spermatogenesis, histone modification, GC-2spd(ts) cell, ChIP-sequencing

Introduction

Spermatogenesis is the process by which haploid spermatozoa develop from immature germ cells in the testis. This process is regulated by many meiotic stage-specific genes [44]. Initially, diploid spermatogonia mitotically proliferate, and some of them become primary spermatocytes to subsequently undergo meiosis I which results in the production of two haploid secondary spermatocytes [45]. Following meiosis I, two secondary spermatocytes become spermatids and through spermiogenesis after undergoing many chemical and macromorphological changes, they become mature spermatozoa [46,47]. Spermatogenesis involves several important events, but the molecular mechanism underlying this complex process is not completely understood, especially when considering the production of primary spermatocytes. One reason is the lack of understanding of transcriptional regulators during this process [48]. Identification of different and novel types of gene regulatory elements may assist in the proper understanding of the mechanism of this process [42,43,49], which will aid the development of new strategies to treat male infertility and novel contraceptive methods [48].

In general, cell type-specific gene regulation is driven by promoters, enhancers, insulators, and other *cis*-regulatory elements. Novel techniques based on high throughput sequencing methods have revealed the process of identification of novel transcriptional regulatory elements. These modifications include DNA methylation, histone modification, and chromatin remodeling. DNA methylation leads to gene silencing, and expression according to the methylation status can be seen at hypomethylated promoter regions [50]. Histone modification alters gene expression probably by changing the binding ability of

transcription factors, and some specific modifications are known to be linked to *cis*-regulatory genome elements [51,52]. For example, mono-methylation of histone H3 lysine 4 (H3K4me1) and acetylation of histone H3 lysine 27 (H3K27ac) are found in enhancer regions. Tri-methylation of histone H3 lysine 4 (H3K4me3) is rich in promoters [51,53].

According to recent findings, testicular cells undergo extensive epigenetic modifications and histone methylation plays a crucial role during spermatogenesis. For example, super-enhancers are one such regulatory element that is composed of several enhancers to drive strong gene transcription during mitosis-to-meiosis transition in the germline [54]. Three-dimensional chromatin organization identified high compartmentalization of chromatin domains, which leads to appropriate gene expression patterns in testicular germ cells [55]. Another finding is that retrotransposon silencing in male germ cells is linked to DNA methylation, histone modification, and piwi pathways [56]. These indicate the dynamic epigenetic regulation of meiotic genes.

Besides, recent studies suggest the presence of promoters with enhancer function and enhancer elements that work as alternative promoters [57,58]. These findings challenge the original definitions of promoter and enhancer while opening a novel spectrum of discoveries to dual active elements that may be major regulatory elements in gene regulation. Consistently, our group previously found a multifunctional genome element, a dual promoter-enhancer (DPE), at the mouse *Tcam1* locus. This DPE, which is marked with both H3K4me1 and H3K4me3, drives *Smarcd2* and *lncRNA-Tcam1* transcription and enhances *Tcam1* transcription in primary spermatocytes [42]. Interestingly, a

corresponding human genome sequence was found that potentially function as a DPE, although the molecular mechanism was partially different from the mouse [43].

Since our discovery of DPE, several studies have revealed the presence of DPEs or in some cases called ‘epromoter’ in various cell types [7,31,59], which suggests that DPE elements play major roles in many different biological processes. However, the significance of DPE in spermatogenesis is still not well understood. Therefore, in this chapter, I propose a common strategy to identify DPEs related to mouse spermatogenesis in a genome-wide manner, especially when producing primary spermatocytes. The proposed model was based on the high throughput sequencing results and the histone modification patterns that exert promoter and enhancer activity. I identified 337, 2176, and 69 DPE candidate regions in spermatogonia, primary spermatocytes, and GC-2spd(ts) cells respectively, and analyzed 13 selected DPE candidate regions through reporter gene assay.

Materials and Methods

Animals

C57Bl/6 mice were maintained on 14 hours light/10 hours dark cycles at 25℃ and free access to the food and water. The experimental procedures used in this study were approved by the Institutional Animal Use and Care Committee at Hokkaido University.

Isolation of spermatogonia

Testes were collected from 7-8 days old mice, decapsulated and treated with 0.1% collagenase in DMEM at 32 ºC for 30 min. During the incubation, seminiferous tubules were loosened by occasional gentle swirling and left for sedimentation by gravity. After discarding the supernatant, the tubules were treated with 0.1% collagenase in DMEM containing 0.1 mg/ml DNase I at 32 ºC for 15 min. Then, the tubules were washed with DMEM/F12 containing 10% fetal bovine serum (FBS), cut with scissors, and suspended in DMEM/F12 containing 10% FBS. The cell suspension was filtered through 50μm mesh, spread onto 10 cm dishes that were coated with collagen, and cultured overnight at 32 ºC with 5% CO_2 in the air. The floating cells were collected as spermatogonia.

Isolation of spermatocytes

Testes were collected from adult male mice and treated with collagenase and DNase I as described above. The tubules were washed with Phosphate buffered saline (PBS) containing 1 mM EDTA, cut with scissors, and suspended in the same buffer. After filtering through 50μm mesh, the cells were stained with Hoechst33342 for 2 hours and subsequently with propidium iodide for 5 min. This cell suspension was subjected to fluorescent activated cell sorter and collected the primary spermatocyte as previously reported [60]. A representative result showing the spermatocyte fraction is presented in Fig. 1.01.

Cell culture

Mouse spermatocyte-derived GC-2spd(ts) cells were cultured in Dulbecco's Modified Eagle Medium (Fujifilm Wako) containing 10% FBS supplemented with 100 U/ml penicillin, 100 µg/ml streptomycin, and 292 µg/ml L-glutamine (Invitrogen, Carlsbad, CA, USA). Culture cells on a 6-cm dishes were used for each analysis of RNA-seq and ChIP-seq.

RNA-sequencing analysis

Total RNA was isolated from each purified cell type (spermatocyte, spermatogonia, GC-2spd(ts) cells) using ISOGEN II (Nippon Gene, Tokyo, Japan) according to the manufacturer's instructions. After treatment with TurboDNase (Thermo Fisher Scientific, Waltham, MA, USA), first-strand cDNAs were synthesized using the oligo(dT) primer and Superscript III (Thermo Fisher). RNA was removed from the DNA-RNA hybrid using RNase H and double-stranded cDNA was synthesized using DNA polymerase I. Obtained double-stranded cDNA was fragmented using DNase I to generate small fragments of cDNA suitable for sequencing using an Illumina Genome Analyzer. T4 DNA polymerase was used to perform end repair of cDNA fragments. The phenol-chloroform extraction method was used to purify the fragmented blunt cDNA fragments. Overhanging adenine (A) base is added to the 3' end of the blunt DNA fragments by Klenow fragment. The Illumina adapters were ligated by using T4 DNA ligase (Thermo Fisher). Using gel purification, I obtained the DNA fragment sized around 150-350 bp and PCR was performed using Illumina PCR primer pairs. DNA sequencing was performed for the ethanol precipitated PCR product and sequence reads were aligned with the mouse genome

for further analysis [61]. The resultant reads and mapping rate to the mouse cDNA library are summarized in Table 1.01.

Chromatin immunoprecipitation-sequencing (ChIP-Seq)

Chromatin immunoprecipitation was performed as previously described with slight modifications [62]. I used spermatogonia, primary spermatocytes, and GC-2spd(ts) cells prepared as above for the ChIP-seq analysis. I have treated the cells with 1% formaldehyde in Opti-MEM (Invitrogen) for 10 min at room temperature with gentle agitation for cross-linking the chromatin, and the reaction was stopped by incubation with 0.125 M glycine for 5 min at room temperature. After cell suspensions were washed twice with ice-cold PBS, the final cell solution was suspended in 1 ml ice-cold PBS with 1× proteinase inhibitor cocktail (Roche Molecular Biochemicals) and centrifuged at 700 g at 4°C for 5 min. Then, SDS Lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris–HCl, pH 8.1) containing a 1× proteinase inhibitor cocktail was used to lyse the cells by incubation on ice for 10 min. The cross-linked chromatin was sonicated using Tomy Ultrasonic Disrupter UD-201 (Tomy, Tokyo, Japan) with 20% output for 15 sec 10 times. The chromatin solution was centrifuged at 15,000 rpm at 4°C for 10 min and the supernatant was mixed with 9 volumes of ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris–HCl, 167 mM NaCl, pH 8.1) containing 1× proteinase inhibitor cocktail. Then, the chromatin was precleared with protein G sepharose (GE Healthcare) and incubated at 4 °C overnight with monoclonal antibodies specific to H3K4me1 and H3K4me3 that were kindly gifted by Dr. Hiroshi Kimura at Tokyo Institute of Technology [63]. The immunocomplexes were collected with protein G sepharose, and the bound chromatin was washed with low salt wash buffer (150 mM NaCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl,

pH 8.1), high salt wash buffer (500 mM NaCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl, pH 8.1), LiCl wash buffer (0.25 M LiCl, 1% NP40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris–HCl, pH 8.1), and TE buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8.0) in this order. The chromatin was eluted twice with freshly prepared elution buffer (1% SDS, 0.1 M $NaHCO_3$) and subject to reverse cross-linking with 0.2 M NaCl at 65 °C for 4 hours. DNA was purified by phenol/CIAA extraction and ethanol precipitation and used for high throughput sequencing. Next-generation sequencing was done by Hokkaido System Science. The resultant reads and mapping rate to the mouse genome sequence are summarized in Table 1.01.

Plasmid constructs

Selected candidate dual promoter-enhancer regions (DPEs) were amplified by PCR using KOD FX Neo (Toyobo) with genomic DNA isolated from the C57BL/6 mouse liver and the primer pair listed in Table 1.02. To show the promoter activity, the DPE fragments were ligated at the *Sma*I site upstream of the luciferase gene in a pGL3-Basic vector. To show enhancer activity, I prepared the construct using the pGL-TK vector. The pGL-TK vector was prepared using thymidine kinase (TK) promoter which was taken from a pEBMulti-Neo vector (Wako) by *Nhe*I and *Bgl*II restriction digestion, blunted by T4 DNA polymerase (Takara) and inserted into a pGL3 basic vector (Promega) at the *Sma*I site. The amplified candidate sequences were inserted downstream of the luciferase gene at the blunted *Bam*HI site. I named all the constructs respective to the name of the gene promoter together with the DPE abbreviation. All constructs were subjected to the sequencing analysis.

Transfection and reporter gene assay

To perform reporter gene assay, GC-2spd(ts) cells that were cultured on 24 well-plates were transfected with GeneJuice transfection reagent according to the manufacturer's instructions (Merck, Darmstadt, Germany), and reporter gene assay was performed two days later, as previously described [42,62]. The pRL-CMV vector (Promega) was co-transfected to normalize the transfection efficiency.

Results

Expression profiles in spermatogonia, primary spermatocytes, and GC-2spd(ts) cells by RNA-seq

To evaluate the quality of my RNA-seq data, I checked the expression pattern of genes reported to be specific for each cell line. *Stra8*, *Nanos3*, *Kit*, and *Ndrg4* genes are known to be specially expressed in spermatogonia [64,65]. In my data, all the above genes are highly expressed in spermatogonia (Fig. 1.02A), which confirmed the good quality of my RNA-seq data and the spermatogonia cell sample.

Then, I checked the expression of *Sycp2*, *Tcam1*, *Prss41*, and *Rhcg* genes related to primary spermatocytes [65,66] in my RNA-seq data and observed all the genes expressed at high levels in primary spermatocytes (Fig. 1.02B). This confirmed the good quality of RNA-seq data and the primary spermatocytes sample.

Next, I checked the expression of *Rsbn1l*, *Spata21*, and *Steap1* genes that are germ cell-specific [67,68] and observed all the genes expressed at high levels in GC-2spd(ts) cells (Fig. 1.02C). *Rsbn1l* and *Spata21* genes were also expressed in both spermatogonia and primary spermatocytes [67,68] claiming that GC-2spd(ts) cells showed the characteristic of both cell types even though they were derived from primary spermatocytes. *Tfcp2* and *Esyt3* are two genes that are expressed in Leydig cells and *Ptprv* and *Cyp21a1* are markers for Sertoli cells. According to reports, these genes are expressed at higher levels in Leydig and Sertoli cells [69]. I checked the expression levels of these genes in my RNA-seq data and found that the expression levels were lower compared to the marker genes of respective cell types in whole testis RNA-seq data. Additionally, sertoli cell-specific markers *Cyp21a1* and *Ptprv* were highly expressed based on whole testis RNA-seq data suggesting the purity as it does not contain Sertoli cells as contaminant although Leydig cell-specific germ cell markers show the same trend as whole testis contain only around 3-5% of Leydig cells (Fig. 1.02D). Additionally, Table 1.03 shows the collection of cell type-specific genes based on the obtained RNA-seq data. Collectively, these data strongly suggest that my RNA-seq data reflect genuine expression patterns in GC-2spd(ts) cells, spermatogonia, and primary spermatocytes.

First attempt to identify the DPEs by chip-seq analysis

I selected the genomic regions based on specific epigenomic markers, H3K4me3 peak for promoter regions and H3K4me1 peaks which are markers for enhancer regions in a genome-wide manner. In the basic analysis, I selected 3 genomic sites but could not detect any promoter and enhancer activity as expected. Interestingly, I was able to detect

significant enhancer activity for the site6 and later named it as DES-K16. Based on these results I assumed that even though the genomic region contains specific makers for both promoter and enhancer may not act as a duel promoter enhancer (further discussed in chapter 3).

More DPEs candidates in spermatocytes by ChIP-seq

Next, I analyzed the cell type-specific genome-wide H3K4me1 distribution using ChIP-seq analysis. Here I hypothesized that a genomic region which was associated with H3K4me1 around a transcription start site (TSS) of an expressed gene shows enhancer activity in addition to promoter activity. Thus, I selected H3K4me1 rich regions that were present within 500 bp to TSSs as candidates for DPE regions (Fig. 1.03A).

Based on the ChIP-seq and RNA-seq data, I have identified 337 regions in spermatogonia, 2176 regions in primary spermatocytes, and 69 regions in GC-2spd(ts) cells as DPE candidates (Fig.1.03B). Among these, 98 DPE candidates were shared in spermatogonia and primary spermatocytes, 7 were common in spermatogonia and GC-2spd(ts) cells, and 24 were found in both primary spermatocytes and GC-2spd(ts) cells (Fig. 1.03B). Five DPE candidate regions were found in all the cell types. These findings suggested that the number of DPEs was increased from spermatogonia to primary spermatocytes and that this dynamic epigenetic change has a significant effect on transcriptional regulation during the onset of meiosis.

DPE activity by the in vitro reporter gene assay

To assess DPE activity, I selected 13 DPE candidate regions which were marked with a combination of methylation patterns representing each cell type. The selected regions were named according to the corresponding gene name hyphened with DPE in Capital letters (e.g. RIN2-DPE), in which the promoter of the respective gene is enclosed within the selected DPE region. The selection was made based on our ChIP-seq data which was also congruent with publicly available data. All the selected 13 candidates displayed H3K4me1 methylation mark in spermatocytes. PIH1D3-DPE, KCNE3-DPE, and PRSS44-DPE show the H3K4me1 pattern only in spermatogonia while GM17644-DPE, ITCH-DPE, RIN2-DPE, PLIN3-DPE, and DNMT1-DPE show H3K4me1 methylation only in GC-2spd(ts) cells. ARHGAP30-DPE and RIMBP3-DPE do not show the H3K4me1 peaks in both spermatogonia and GC-2spd(ts) cells. Additionally, SP110, GCFC2, and SPEER4COS-DPEs show the H3K4me3 in all 3 cell types which are spermatogonia, spermatocyte, and GC-2spd(ts) cells. Apart from that, I checked the H3K27ac mark related to these DPEs and around that the half of the selected candidates do not show this enhancer specific histone methylation mark including ARHGAP30-DPE, PIH1D3-DPE, RIN2-DPE, PLIN3-DPE, DNMT1-DPE, and SPEER4COS-DPE (Table 1.04).

The 13 regions were obtained by PCR as 200-bp to 1000-bp sequences and the *in vitro* reporter gene assays were performed (Table 1.05). I prepared two types of constructs for each sequence, one for promoter activity and the other for enhancer activity. To check the promoter activity, I prepared a construct in which a DPE candidate sequence was inserted upstream of the luciferase gene. I also checked enhancer activity using the constructs in

which each sequence was inserted into the downstream of luciferase gene driven by the thymidine kinase (TK) promoter in either direction.

The followings are detailed descriptions of results for each DPE candidate.

ARHGAP30-DPE: This DPE region is an actual promoter of *Arhgap30* which is responsible for producing regulatory proteins related to GTPase activator activity according to the sequence similarities. According to the result of *in vitro* reporter gene analysis, this region had high promoter activity (Fig. 1.04A) and showed the orientation-dependent enhancer activity (Fig. 1.05A). While significant enhancer activity was detected for the reverse direction, for the forward direction it showed only a slight activity which was not significant.

RIMBP3-DPE: This region is a promoter of the *Rimbp3* protein-coding gene which plays a key role in sperm head morphogenesis during the late stages of sperm development and is important for male fertility [70,71]. The region showed significant enhancer activity in the reverse direction while the forward direction showed only slight activity (Fig. 1.05A). In contrast, I could not detect any promoter activity (Fig. 1.04A) even though the region should be an actual promoter of the *Rimbp3* gene in terms of its position.

PIH1D3-DPE: This genomic region is a promoter of the *Pih1d3*/*Dnaaf6* gene, which plays a role in cytoplasmic pre-assembly of axonemal dynein in mouse sperm [72]. By reporter

assay, I observed slightly higher promoter activity than the control vector (Fig. 1.04A) and no enhancer activity in either direction (Fig. 1.05A).

KCNE3-DPE: According to the sequence similarities, this gene is related to the potassium ion channel-related peptide production. Similarly, to PIH1D3-DPE, this region did not show any significant enhancer activity in either direction (Fig. 1.05A) but I observed slight promoter activity by reporter gene analysis (Fig. 1.04A).

PRSS44-DPE: This DPE region is related to the promoter of the *Prss44/Tessp-4* gene which encodes a testis-specific serine protease. The knockout mice of this region are fertile, and the actual function is unclear [60,73]. This DPE showed significant enhancer activity in the forward direction while the reverse direction did not (Fig. 1.05A). The promoter activity was higher than the control (Fig. 1.04A).

GM16744-DPE: This region is linked to a predicted lncRNA which is expressed especially in the mouse testis [37]. While the region showed some promoter activity by reporter assay (Fig. 1.04B), it did not show any enhancer activity (Fig. 1.05B). Even though this is a promoter region of the respective gene and it shows some negative effect over the promoter activity of the TK promoter by reporter gene analysis suggesting the possibility of acting as a silencer element. I look for the H3K27me3 marks using publicly available ChIP-seq data, but it does show a very slight peak.

SPEER4COS-DPE: This region belongs to the promoter of *Speer4cos* which is a lncRNA functioning as a spermatogenesis-associated glutamate (E)-rich protein 4C, opposite strand transcript [74]. This DPE region showed slightly higher promoter activity than the control (Fig. 1.04B) while significant enhancer activity without any orientation-dependency was shown (Fig. 1.05B).

ITCH-DPE: This is a promoter region of the *Itch* gene which is responsible for protein ubiquitination [75,76]. I detected high promoter activity via reporter gene analysis (Fig. 1.04C). Significant enhancer activity was observed in the construct which contained the ITCH-DPE sequence in the forward direction, but not in the reverse direction (Fig. 1.05C).

RIN2-DPE: This is a promoter region of the *Rin2* gene which is required for cell cycle control [77]. I detected higher promoter activity than the control (Fig. 1.04C) and significant enhancer activity in the forward direction but not in the reverse direction (Fig. 1.05C).

PLIN3-DPE: While this region showed high promoter activity (Fig 1.04C), I could not detect enhancer activity in either direction (Fig. 1.05C). My results only showed lower enhancer activity than the control.

DNMT-DPE: This is a promoter region of the DNA methyltransferase gene which regulates and maintains DNA methylation patterns [78]. This region showed high promoter activity (Fig. 1.04C), and reporter construct with this region in the forward direction

showed significant enhancer activity (Fig. 1.05C). The construct in the reverse direction showed lower activity than the control.

GCFC2-DPE: This is a promoter region of the *Gcfc2* gene and did not show any promoter activity in reporter gene assay (Fig. 1.04D). The construct with this region in the reverse direction showed significant enhancer activity while the construct in the forward direction did not show any activity (Fig. 1.05D).

SP110-DPE: This DPE is a promoter region of the *Sp110* gene which is presumed to encode a transcription factor and play a role in innate immunity [79]. Even though this is an actual promoter for the *Sp110* gene, it did not show any promoter activity (Fig. 1.04D), and enhancer activity was not significantly detected either (Fig. 1.05E).

In summary, I observed significant enhancer activity in 8 out of 13 selected DPE candidates used for reporter analysis (Table 1.06): ARHGAP-DPE, RIMBP-DPE, PRSS44-DPE, SPEER4COS-DPE, ITCH-DPE, RIN2-DPE, DNMT-DPE, and GCFC2-DPE. Most of these regions showed the forward direction specificity in enhancer activity while SPEER4COS-DPE did not show the directional specificity (Fig. 1.05A, B, C). On the other hand, 10 selected DPE regions showed higher promoter activity than the control, even though these regions were within 500 bp of expressed genes. RIMBP3-DPE, PLIN3-DPE, SP110-DPE, and GCFC2-DPE showed low levels of promoter activity by my reporter gene assay in GC-2spd(ts) cells (Fig. 1.05A, B, C).

Discussion

The evolvement of the novel techniques based on high throughput DNA sequencing has widened the study of genome-wide identification of enhancer and promoter regions and evaluation of their regulatory mechanisms [80–82]. Recent research findings suggest that there are many similarities between the promoter and enhancer regions including similarities between overall chromatin structure, sequence, and core promoter architecture [83]. One of the recent research has suggested that putative ancestral enhancers turned into promoters during evolution [84]. There are some findings that the same genomic region exerts both enhancer and promoter activity such as dual enhancer/promoter regulatory element for human CD69 expression, transcriptional activation of chicken lysozyme, and DPE activity for human *TCAM1P* pseudogene. These suggest the possibility of a gradual shift from the original clear-cut promoter enhancer dichotomy. [7,31,43,85,86].

As the second attempt with the revised hypothesis, I have performed reporter gene analysis for 13 genomic regions that were selected based on the association of H3K4me1 marker with TSS. These histone modification patterns show a similar trend as in publicly available data for spermatocytes (GEO accession numbers: GSM1202711) (Fig. 1.06A-M). On the other hand, H3K27ac is another well-known histone modification marker related to enhancer regions [87], and I analyzed H3K27ac patterns at the DPE candidate regions in spermatocytes using the public data (GEO accession numbers: GSM1202714). However, while RIMBP3-DPE, KCNE3-DPE, PRSS44-DPE, GM17644-DPE, ITCH-DPE, SP110-DPE, and GCFC2-DPE regions showed clear marks (Fig. 1.06A-M), the other five regions

were not marked with H3K27ac. This suggests that the H3K27ac mark is not a prerequisite for DPE activity.

In general, TSS-associated regions are considered as promoters [88]. However, the reporter assay did not show promoter activity for three DPE regions (GCFC2-DPE, RIMBP3-DPE, and SP110-DPE). This could be attributed to the characteristics of GC-2spd(ts) cells for reporter analysis [89]. My RNA-seq data showed that these three genes were expressed at low levels in GC-2spd(ts) cells, whereas ARHGAP30-DPE, PRSS44-DPE, and SPEER4COS-DPE showed high promoter activity even though these genes were not expressed. In the case of DNMT-DPE and PLIN3-DPE, they were expressed at high levels and showed very high promoter activity in GC-2spd(ts) cells. These data suggest that the actual activity of promoters may be influenced by the internal process of the GC-2spd(ts) cells. Consistently, both the Genome Reference Consortium Mouse Build 39 of NCBI assembly database [90–92] and Berkeley Drosophila Genome Project Searches Neural Network Promoter Prediction web tool [93,94] predicted the high probability that promoter sequences of *Rimbp3*, *Sp110*, and *Gcfc2* genes function as actual promoters. Based on these findings, the three DPEs should intrinsically possess promoter activity, but there may be other factors (external stimuli, etc.) in play that influenced the results of the *in vitro* assay. In a conclusion, my data suggest the candidates I picked could actually function as DPEs during spermatogenesis.

Table 1.01 Summary of RNA-seq and ChIP-seq Results

Cell type	Experiment	Antibody	Total reads	Overall alignment rate
GC-2spd(ts) cell	RNA-seq		32,789,204	90.79
GC-2spd(ts) cell	ChIP-seq	H3K4me1	44,405,976	96.61
GC-2spd(ts) cell	ChIP-seq	H3K4me3	47,092,972	98.08
Spermatogonia	RNA-seq		35,350,264	85.34
Spermatogonia	ChIP-seq	H3K4me1	44,268,650	98.21
Spermatogonia	ChIP-seq	H3K4me3	46,433,320	86.16
Primary spermatocytes	RNA-seq		26,081,767	83.96
Primary spermatocytes	ChIP-seq	H3K4me1	45,916,902	96.63
Primary spermatocytes	ChIP-seq	H3K4me3	45,404,280	78.53

Table 1.02 Primer Used for Cloning

Designation	Forward 5'-3'	Reverse 3'-5'
[Cloning]		
ARHGAP30-DPE	TGTTCTTCGTGGTCCTTTCC	CAAACCCGCTTCCTGTTAGA
RIMBP3-DPE	AACTCCTGCAATCCAAGCAC	GAGCTCTGAGCCCTAAGTGG
GM17644PR-DPE	TTGCTGGTCCTTTGAGTTGA	CAAACAGGTGATGGAAATGAAC
SPEER4COS-DPE	GCTGGCCTCAAATTCAGAGA	TTCCTCTAAAGAACTGGCTTG
PIH1D3PR-DPE	CTGAGCCAGCTGGGTATCTG	CTTCGGGATTGAGGAGACTG
KCNE3PR-DPE	ATGAGACCACCCAGACCATT	CAAGGTCCTCCCTCTTTTCC
PRSS44PR-DPE	CTGGCTCTCCTGGAACTCAC	ACCTGGAGTAAATGCCATCG
ITCH -DPE	GCTGTCACCCTCCTAAGAACT	GAAATGACCCCAGTCTGCCT
RIN2 -DPE	GCTGTTTACATGCGTCCCCT	TGGTCCACGCTGACATTTCC
PLIN3 -DPE	CCCAGGAAGCCTCGTTTCTC	GCTTACAGGCCAAGGGGAG
DNMT -DPE	ACCCCCAATATATGCCTCGG	GCCTCCTTAGTGTAAGGGACTG
SP110-DPE	TCTCCAGCAGAAGGTGTAGC	GAAGTGAAAGATGAGCAGGT
GCFC2-DPE	GTAGCTGTATGCACCCAACTTT	GTTTGTGAGAGCTCCGCCT

Table 1.03 Cell Type-Specific Genes by RNA-seq Analysis

Gene Symbol	Description
Spermatogonia specific genes	
Nanos2	Mus musculus nanos homolog 2 (Drosophila) (*Nanos2*), mRNA.
Nanos3	Mus musculus nanos homolog 3 (Drosophila) (*Nanos3*), mRNA.
Stra8	Mus musculus stimulated by retinoic acid gene 8 (*Stra8*), mRNA.
Kit	Mus musculus kit oncogene (*Kit*), transcript variant 2, mRNA.
Sall4	Mus musculus sal-like 4 (Drosophila) (*Sall4*), transcript variant a, mRNA.
Csf1r	Mus musculus colony stimulating factor 1 receptor (*Csf1r*), mRNA.
Nlrp4c/Rnh2	Mus musculus NLR family, pyrin domain containing 4C (*Nlrp4c*), mRNA.
Spermatocyte Specific Genes	
Sycp2	Mus musculus synaptonemal complex protein 2 (*Sycp2*), mRNA.
Dkkl1/Soggy	Mus musculus dickkopf-like 1 (*Dkkl1*), mRNA.
Dnah8	Mus musculus dynein, axonemal, heavy chain 8 (*Dnah8*), mRNA.
Pgk2	Mus musculus phosphoglycerate kinase 2 (*Pgk2*), mRNA.
Tcam1	Mus musculus testicular cell adhesion molecule 1 (*Tcam1*), mRNA.
Prss41	Mus musculus protease, serine 41 (*Prss41*), mRNA.
Prss42	Mus musculus protease, serine 42 (*Prss42*), mRNA.
Prss43	Mus musculus protease, serine 43 (*Prss43*), mRNA.
Prss44	Mus musculus protease, serine 44 (*Prss44*), mRNA.

Table 1.04 Histone Methylation Pattern Distribution Among Selected 13 DPE Candidates

DPE	**H3K4me1 Peak**			**H3k4me1 Peak**	**H3k27ac Peak**
	SG	SC	GC2	Spermatocyte	Spermatocyte
ARHGAP30	x	○	x	○	x
RIMBP3	x	○	x	○	○
PIH1D3	○	○	x	○	x
KCNE3	○	○	x	○	○
PRSS44	○	○	x	○	○
GM17644	x	○	○	○	○
ITCH	x	○	○	○	○
RIN2	x	○	○	○	x
PLIN3	x	○	○	○	x
DNMT1	x	○	○	○	x
SP110	○	○	○	○	○
GCFC2	○	○	○	○	○
SPEER4COS	○	○	○	○	x

Table 1.05 Genomic Location Length and the Distribution of Analyzed DPE

DPE	Chromosome	Start	End	Length	Amplified length
ARHGAP30	chr1	171388731	171388927	197	254
RIMBP3	chr16	17207804	17208169	366	469
PIH1D3	chr1	31222506	31222846	341	341
KCNE3	chr7	100176328	100176729	402	986
PRSS44	chr9	110813993	110814198	206	597
GM17644	chr1	12667512	12667648	137	465
SPEER4COS	chr5	15680102	15680238	137	549
ITCH	chr2	155133048	155133198	151	641
RIN2	chr2	145785768	145785905	138	647
PLIN3	chr17	56290680	56290804	125	699
DNMT1	chr9	20953140	20953276	137	745
SP110	chr1	85604352	85604603	252	694
GCFC2	chr6	81923093	81923201	109	758

Table 1.06 Summary of Duel Promoter Enhancer Activity by Reporter Gene Assay

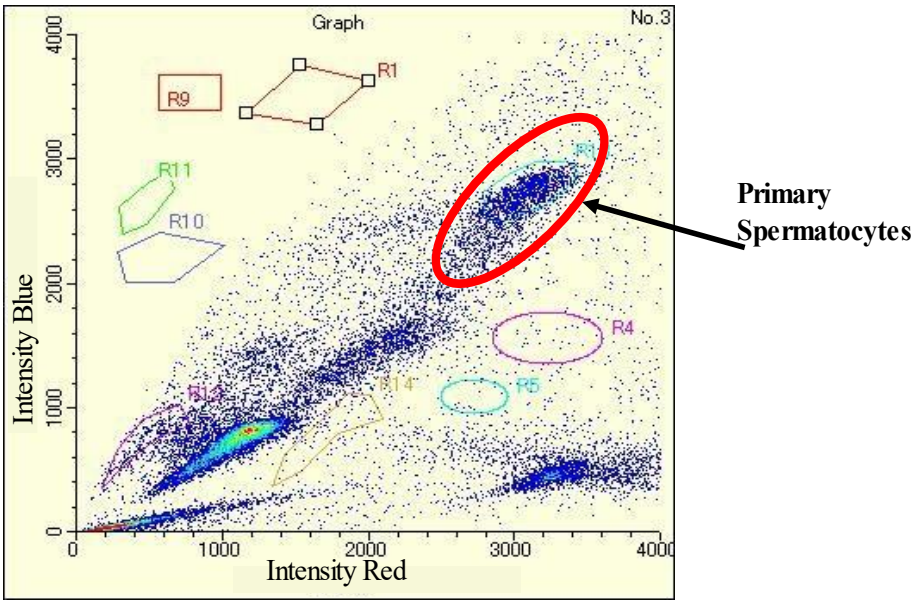
DPE	Promoter activity	Enhancer Activity
RIMBP3	x	○
PIH1D3	○	x
KCNE3	○	x
GM17644	○	x
ARHGAP30	○	○
PRSS44	○	○
ITCH	○	○
GCFC2	x	○
RIN2	○	○
DNMT1	○	○
SPEER4COS	○	○
PLIN3	○	x
SP110	x	x

Table 1.07 Transcription Factors Related to DPE Regions

Transcription factor	DPE Region
GATA2	ARHGAP30, RIMBP3, PIH1D3, KCNE3, PRSS44, GM17644, SPEER4COS, ITCH, RIN2, PLIN3, DNMT1, SP110, GCFC2
CREB	ARHGAP30, RIMBP3, PIH1D3, KCNE3, PRSS44, GM17644, SPEER4COS, ITCH, RIN2, PLIN3, DNMT1, SP110, GCFC2
AP1	ARHGAP30, RIMBP3, PIH1D3, KCNE3, GM17644,SPEER4COS,ITCH,RIN2, PLIN3, DNMT1, SP110
NRF	ARHGAP30, RIMBP3, PIH1D3, KCNE3, PRSS44, SPEER4COS,ITCH, RIN2, PLIN3, DNMT1, SP110,
SOX5	ARHGAP30, RIMBP3, PIH1D3, KCNE3, PRSS44, GM17644, SPEER4COS, ITCH PLIN3, SP110, GCFC2

Fig.1.01

A



B

Fig1.01 Germ cell isolation for ChIP-seq and RNA seq analysis. A). Flow cytometric analysis of germ cells. Cells were plotted against the color intensity of the stain Hoechst blue (y axis) and red (x axis). Circled fragment was collected as Primary spermatocyte. **B).** Photograph of sorted primary spermatocytes under the florescent microscope. According to the shape of a cells quality of the cell was accessed.

Fig. 1.02

A

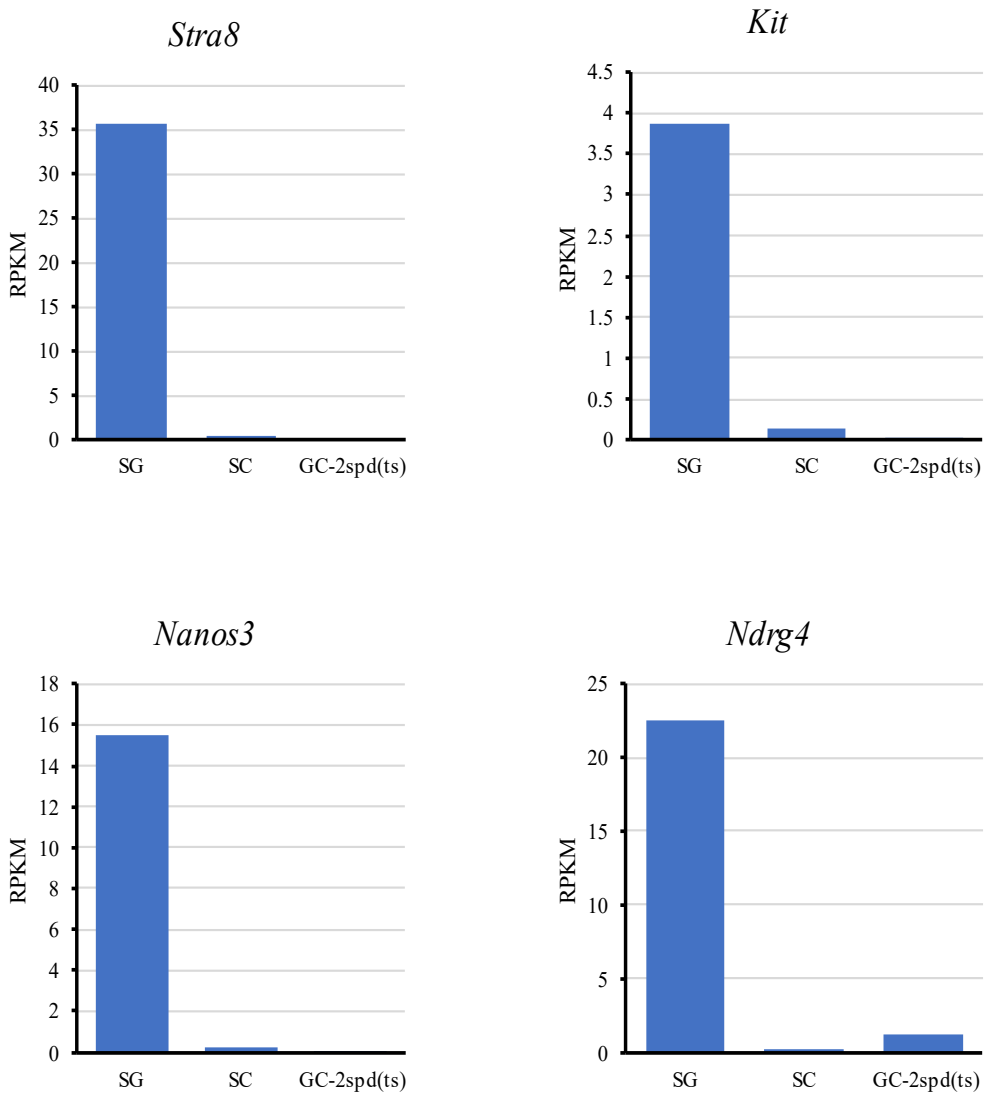


Fig. 1.02 A) Expression of marker genes for meiotic stage specific Cells and GC-2spd(ts) cells by RNA-Seq. *Stra8, Kit, Nanos3, Ndg4 as* Spermatogonia specific genes

Fig. 1.02

B

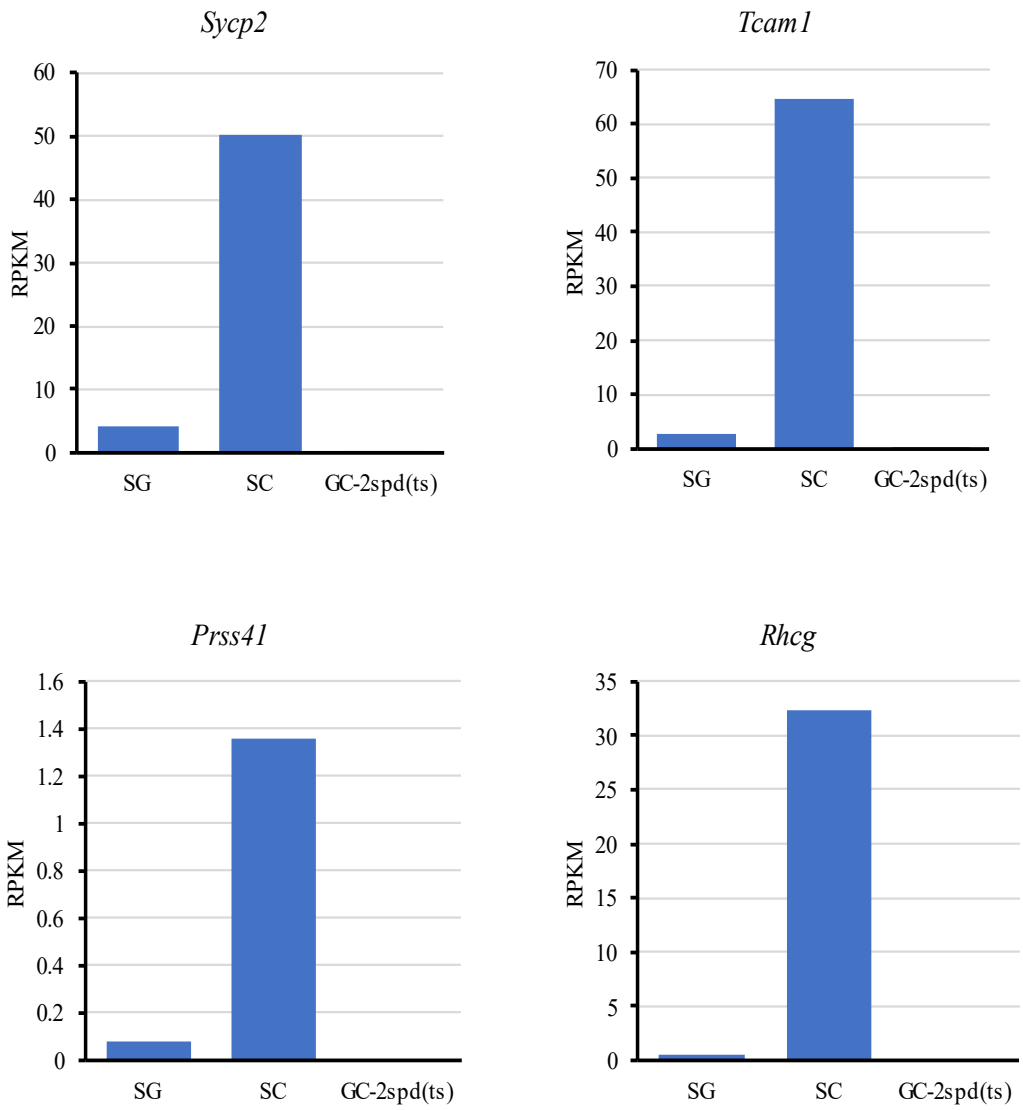


Fig. 1.02 B) Expression of marker genes for meiotic stage specific Cells and GC-2spd(ts) cells by RNA-Seq. *Sycp2, Tcam1, Prss41, Rhcg* as Spermatocyte

Fig. 1.02

C

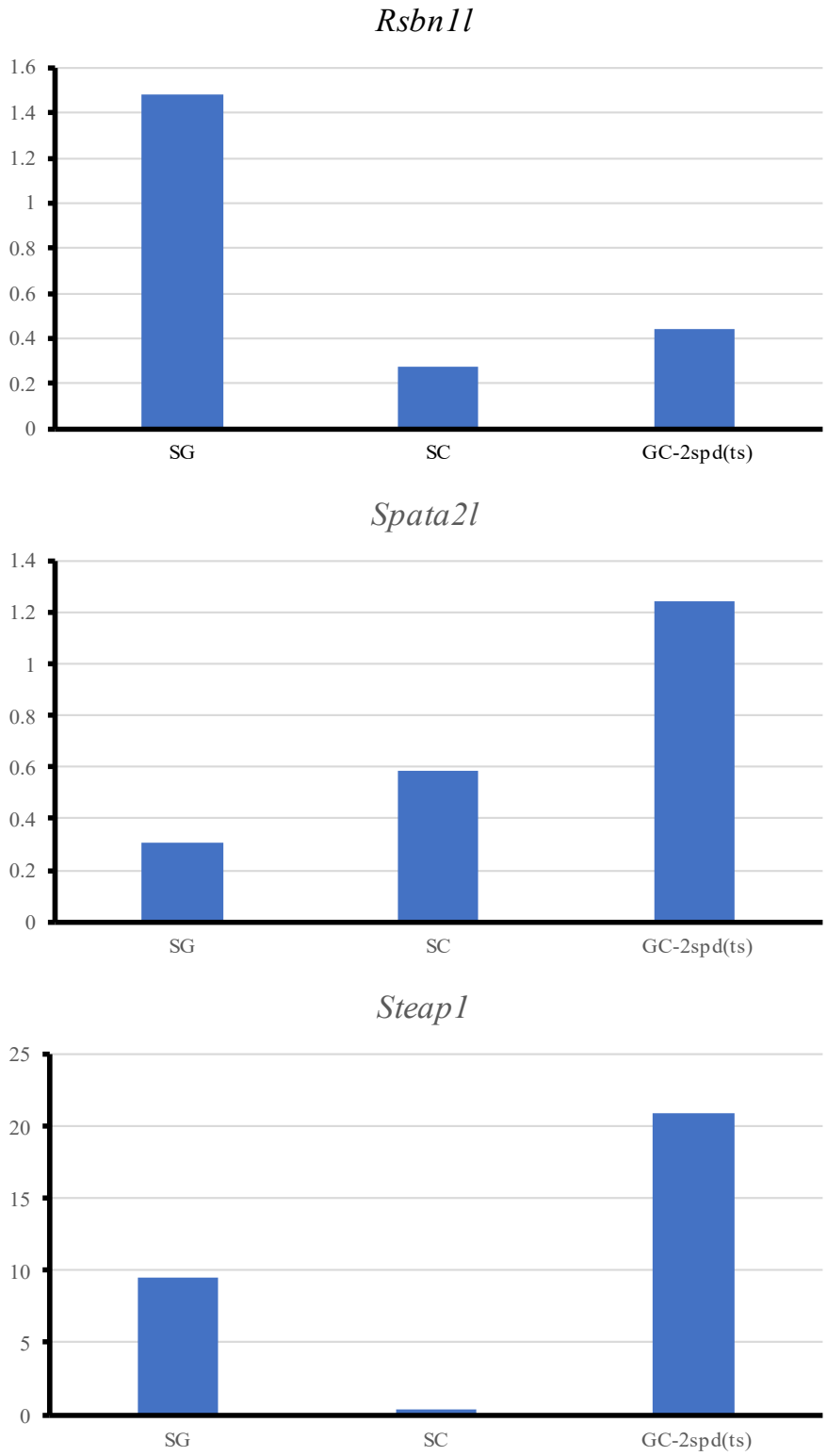


Fig. 1.02 C) Expression of marker genes for meiotic stage specific Cells and GC-2spd(ts) cells by RNA-Seq. *Rsbn1l*, *Steap1* and *Spata21* germ cell-specific

Fig. 1.02

D

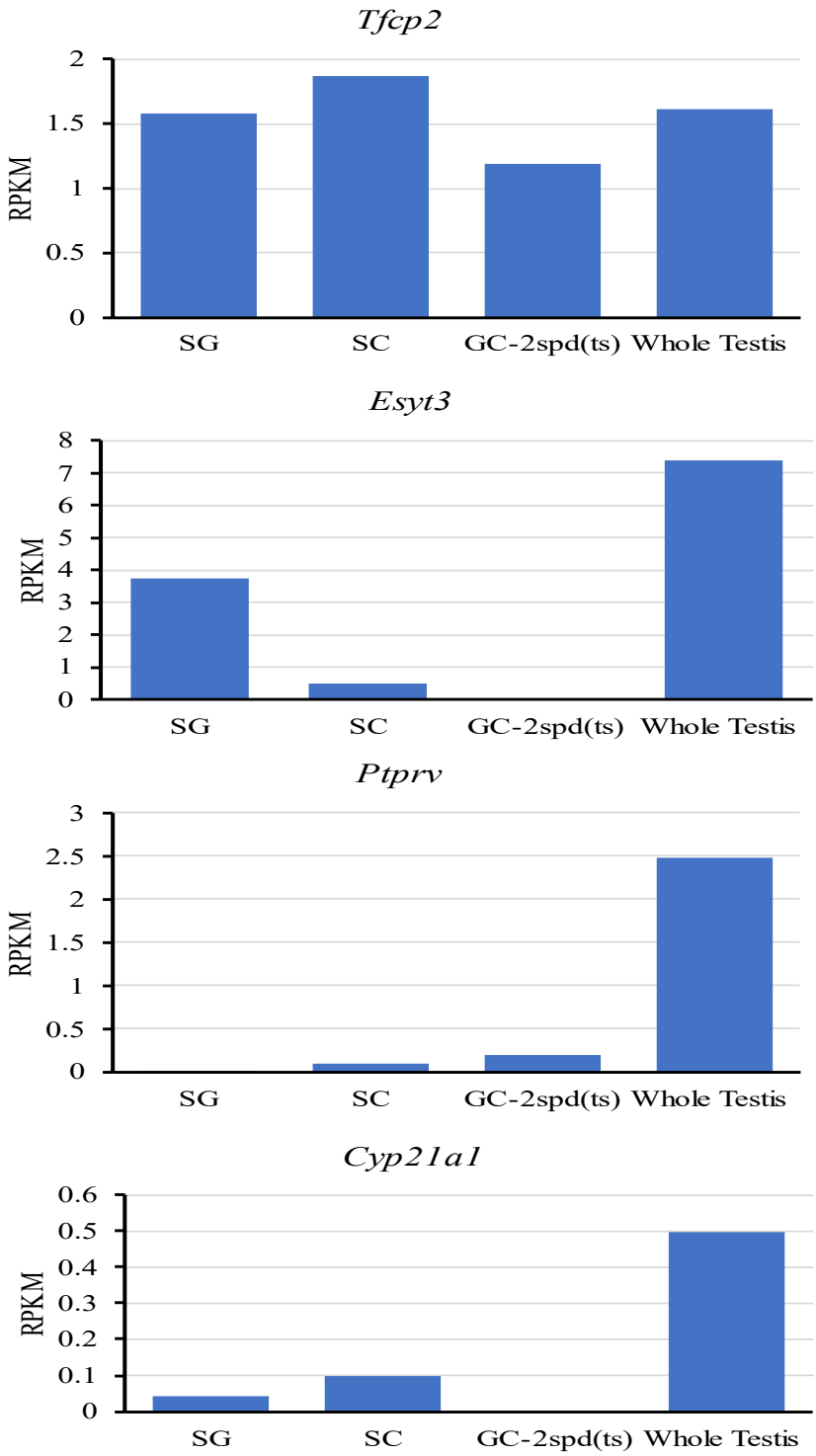
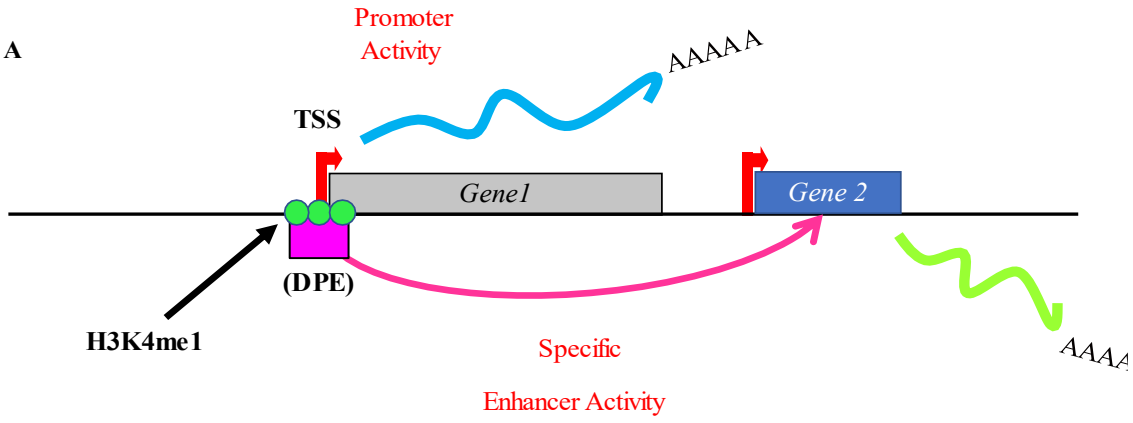


Fig1.02 D) Expression of marker genes for meiotic stage specific Cells and GC-2spd(ts) cells by RNA-Seq ***Tfcp****2, Esyt3,Ptprv, Cyp21a1* as common germ cell specific genes.

Fig. 1.03



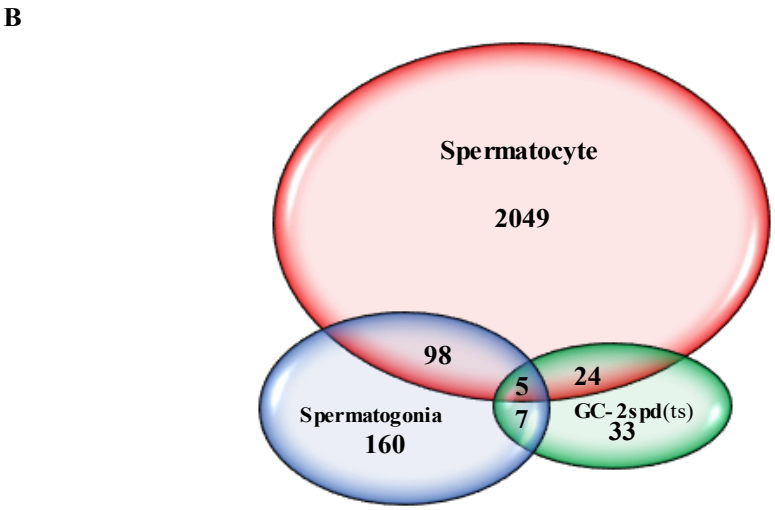


Fig. 1.03 Proposed model for identification of DPE and Venn diagram of RNA-seq and ChIP-seq. A) Transcription start site overlapped H3K4me1 methylation patterns were indicated in green circles and red arrow shows the direction of transcription of the gene. TSS overlapped H3K4me1 peaks considered as the DPE (duel promoter enhancer regions) and pink arrow shows the expected enhancer activity over other adjacent genes. B) Venn diagram indicates the number of DPE candidate region distribution among the Spermatocyte (Pink), Spermatogonia (blue), and GC-2spd(ts) cells(green) and comment DPE regions shows in the cross regions.

Fig. 1.04

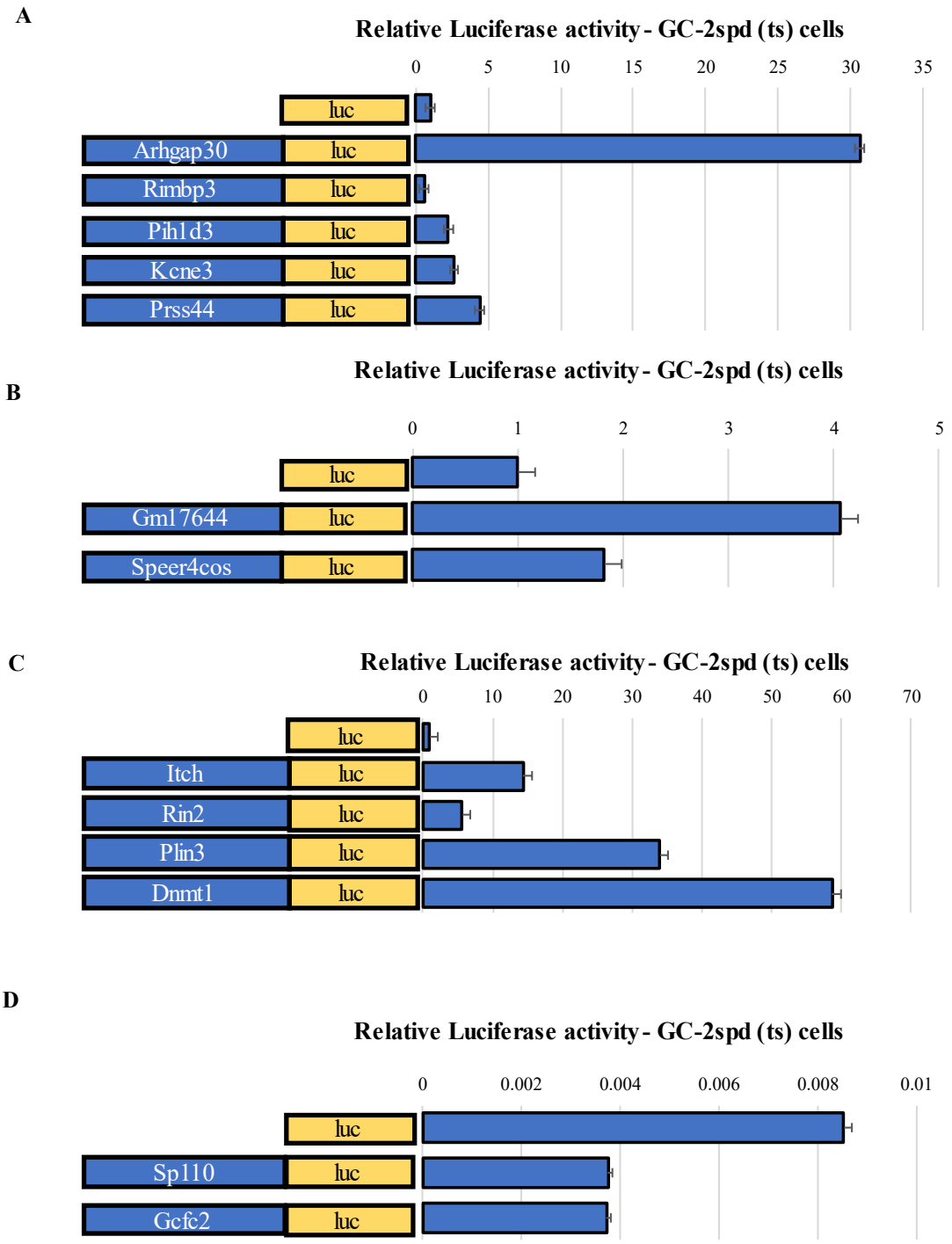
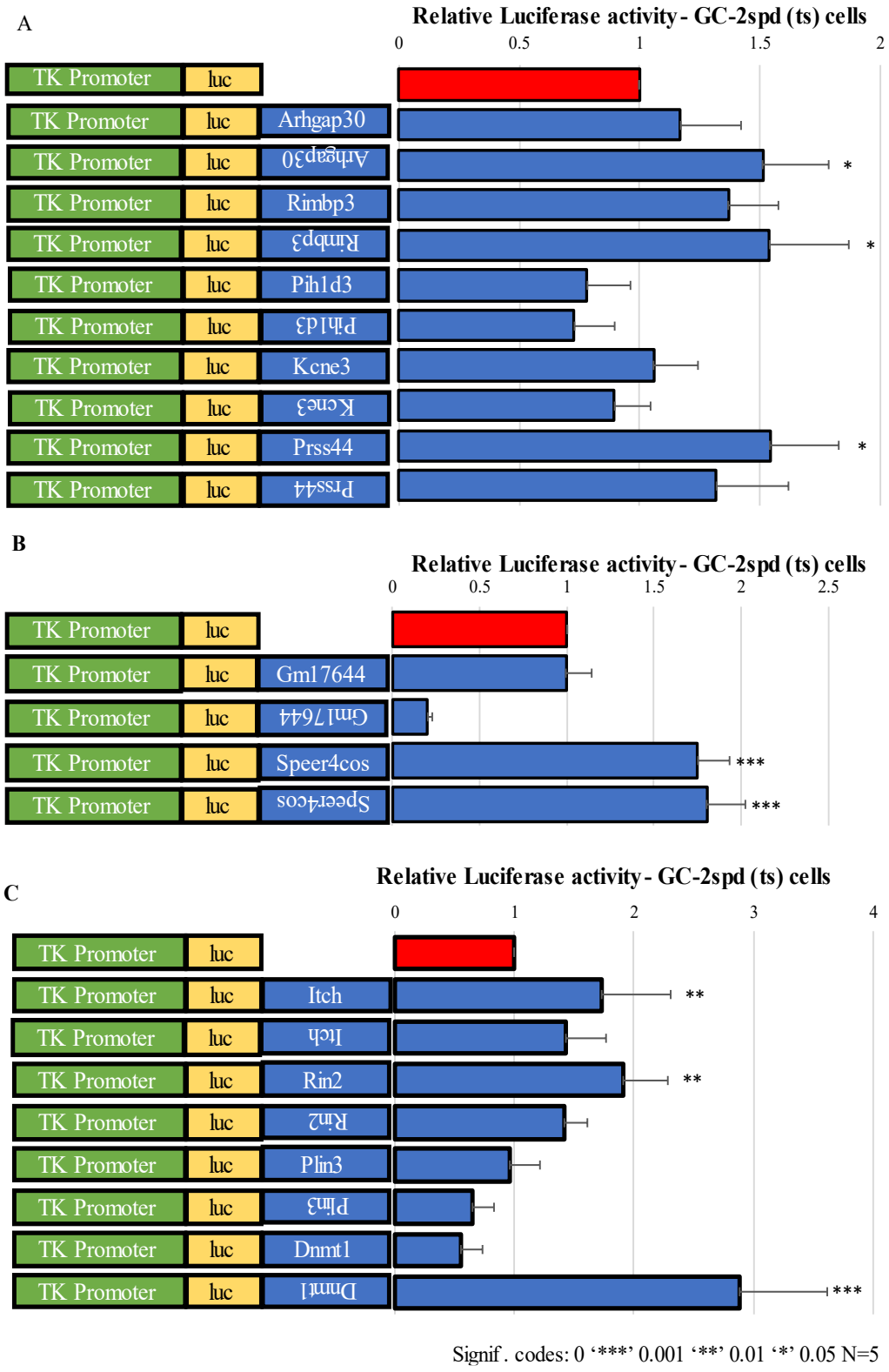


Fig. 1.04 *In vitro* reporter gene assay for transcriptional promoter activity. The reporter gene constructs were generated as indicated at the left side of the graph. Each construct was transfected into GC-2spd(ts) . (A) (B) (C) and (D) luciferase activity was measured two days later. The activity of the empty vector was set to 1.0 in each experiment, and the values from other constructs were calculated. Data are presented as means ± S.D. from five or six independent experiments。

Fig. 1.05



Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 N=5

Fig. 1.05

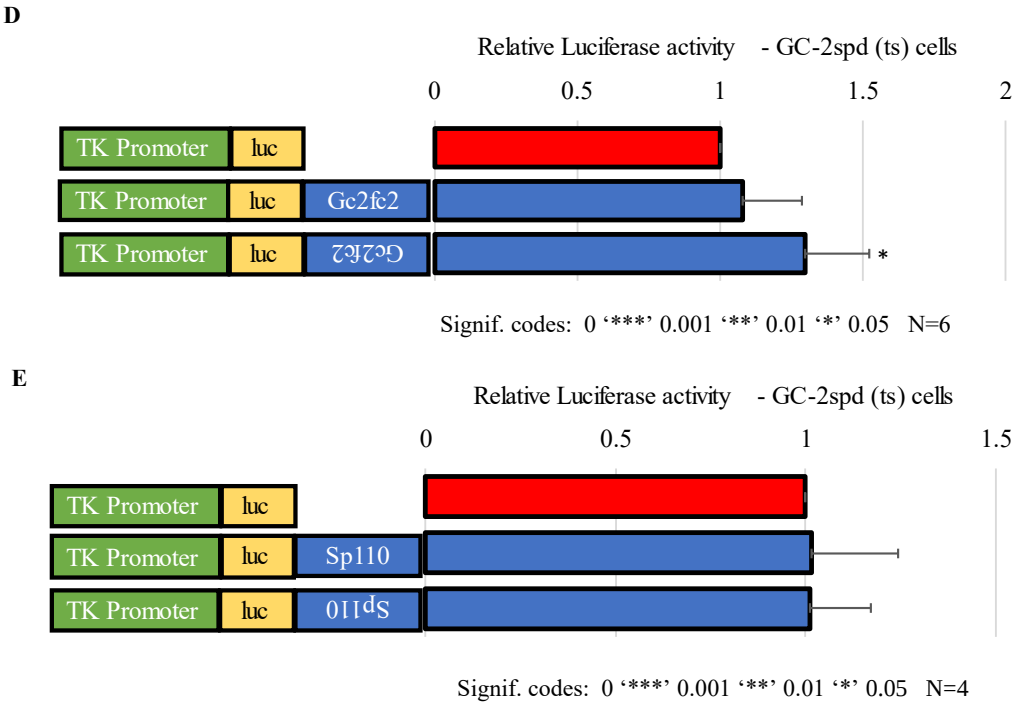


Fig. 1.05 *In vitro* reporter gene assay for transcriptional Enhancer activity The reporter gene constructs were generated as indicated at the left side of the graph. Each construct was transfected into GC-2spd(ts) (A) (B) (C) (D) and (E) luciferase activity was measured two days later. The activity of the vector containing thymidine kinase promoter was set to 1.0 in each experiment, and the values from other constructs were calculated. Data are presented as means ± S.D. from four independent experiments, and statistical significance was assessed by Dunnett's test to compare the construct which containing only thymidine kinase promoter with construct contain both thymidine kinase promoter and DPE candidate with it. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

Fig. 1.06A

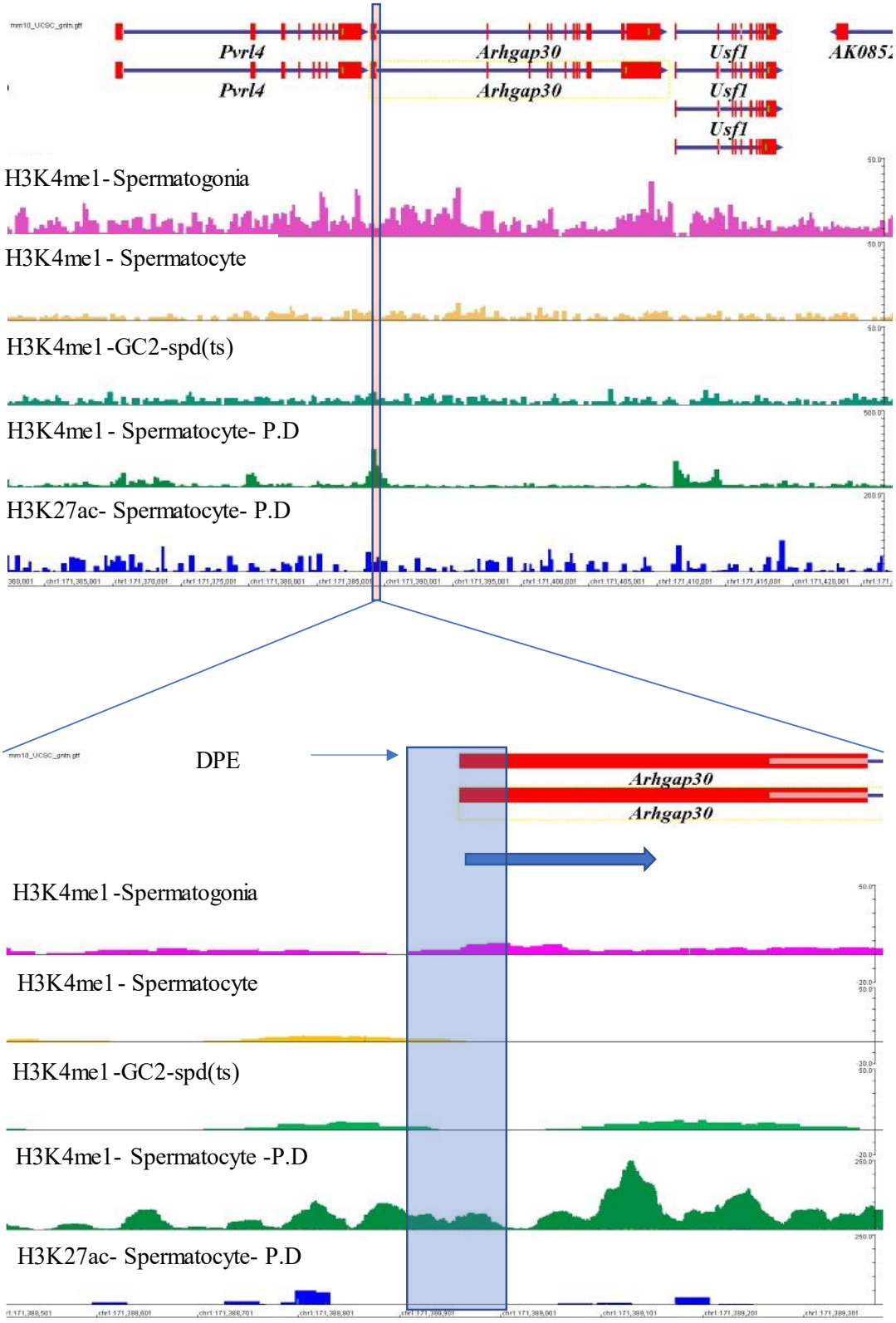


Fig. 1.06A Histone modification patterns ARGAP30-DPE region (P.D- Public Data)

Fig. 1.06B

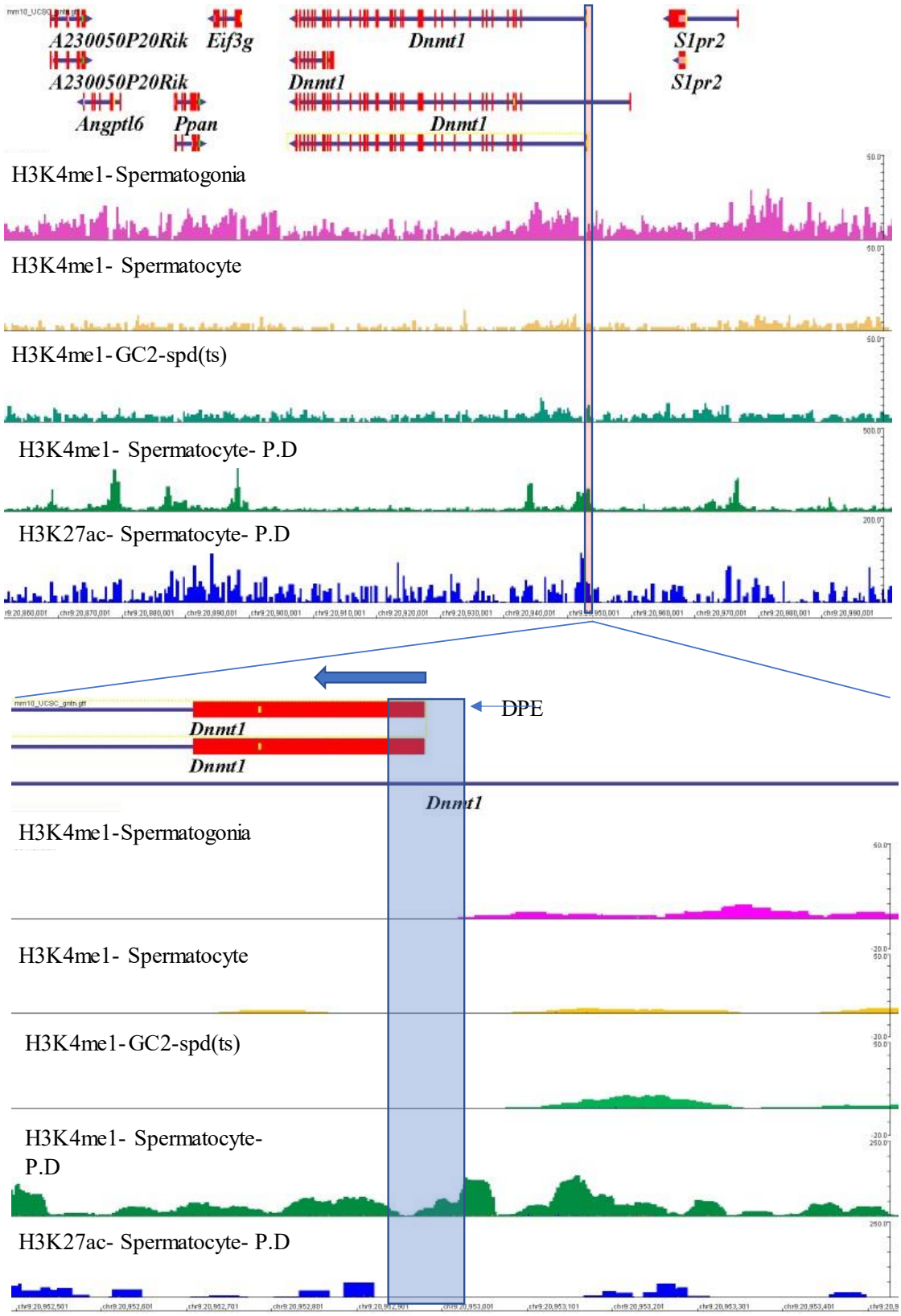


Fig. 1.06B Histone Methylation patterns DNMT- DPE region (P.D- Public Data)

Fig. 1.06C

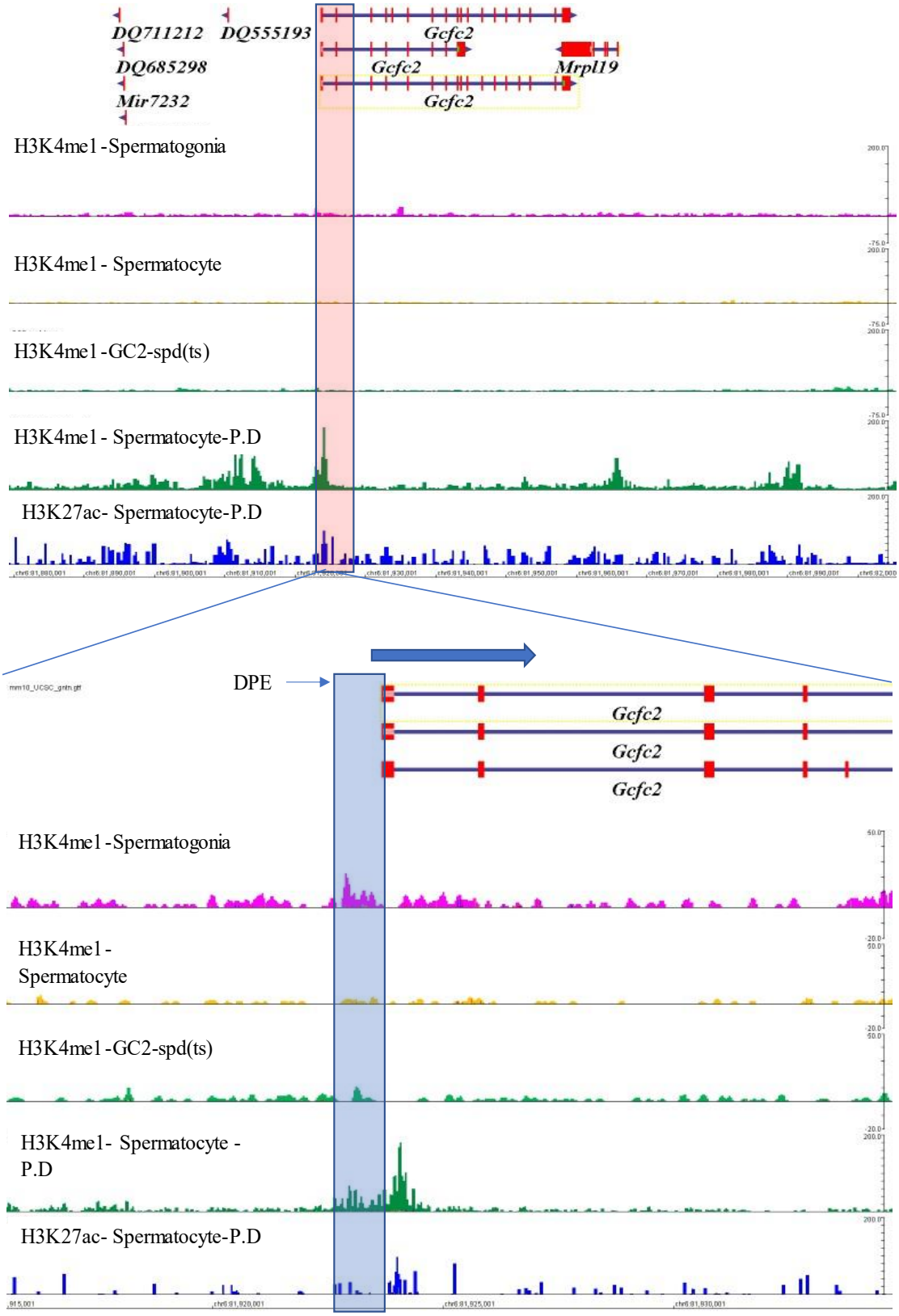


Fig. 1.06C Histone modification patterns GCFC2- DPE region (P.D- Public Data)

Fig. 1.06D

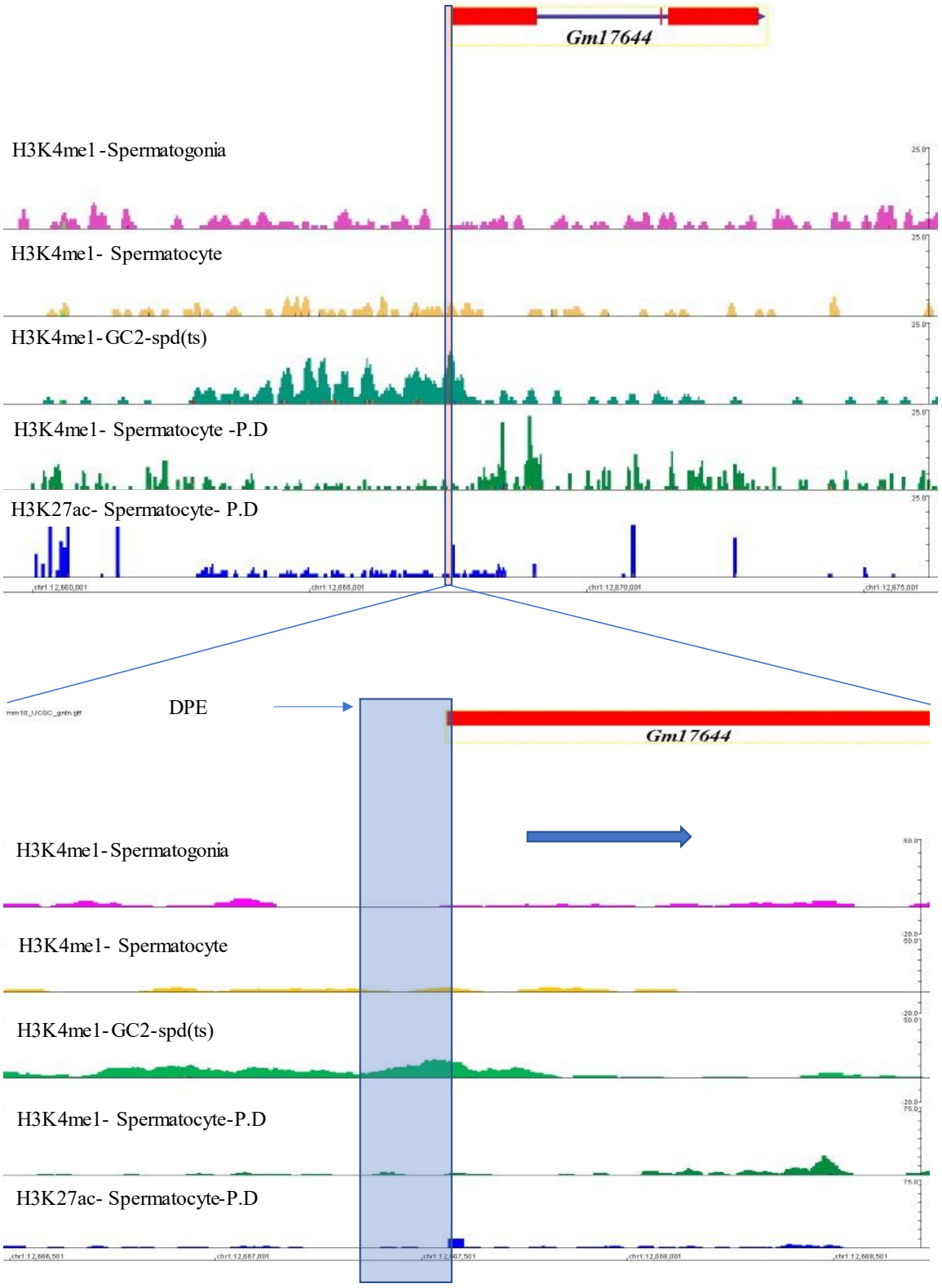


Fig. 1.06D Histone modification patterns GM17644- DPE region (P.D- Public Data)

Fig. 1.06E

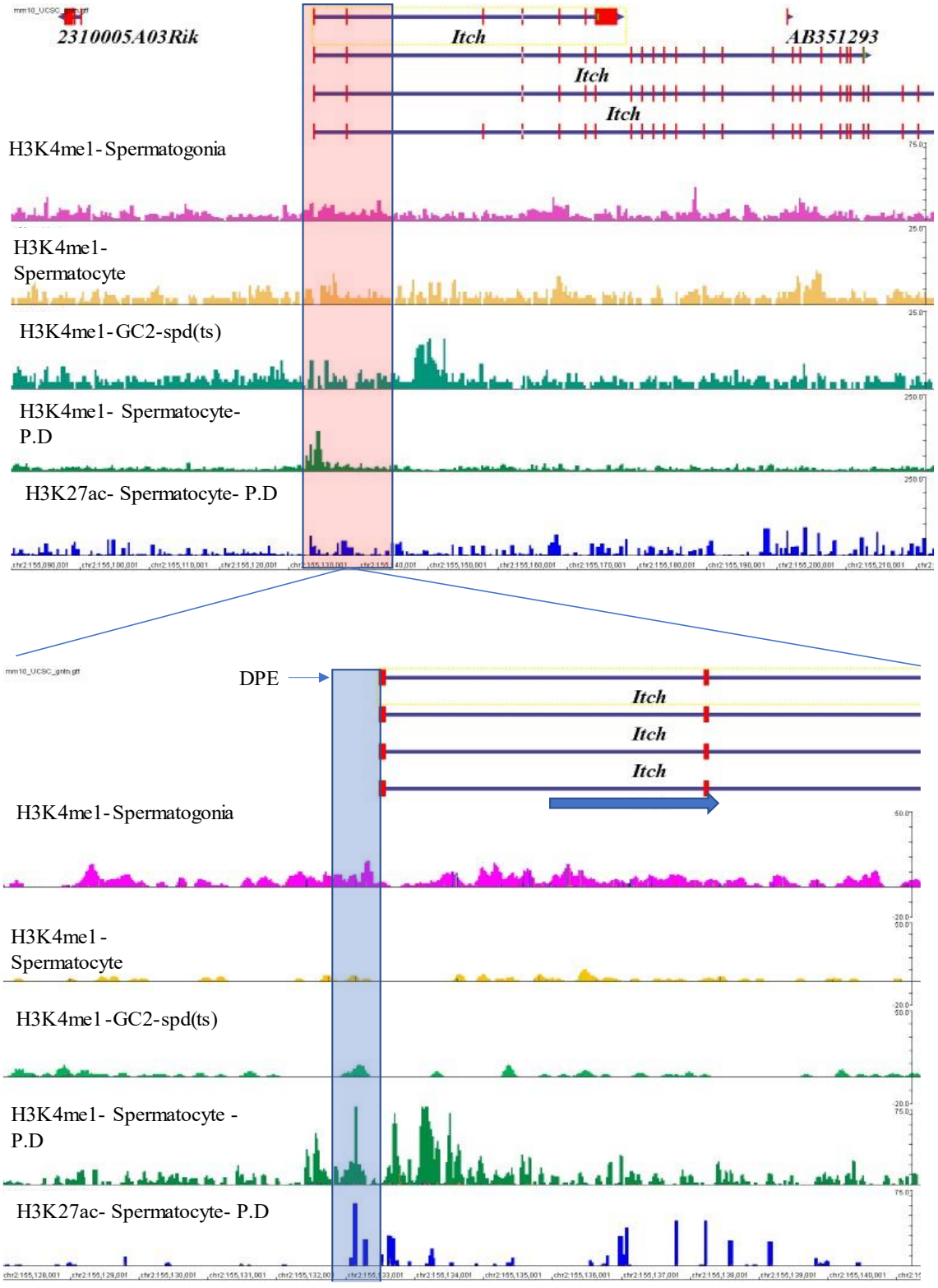


Fig. 1.06E Histone modification patterns ITCH- DPE region (P.D- Public Data)

Fig. 1.06F

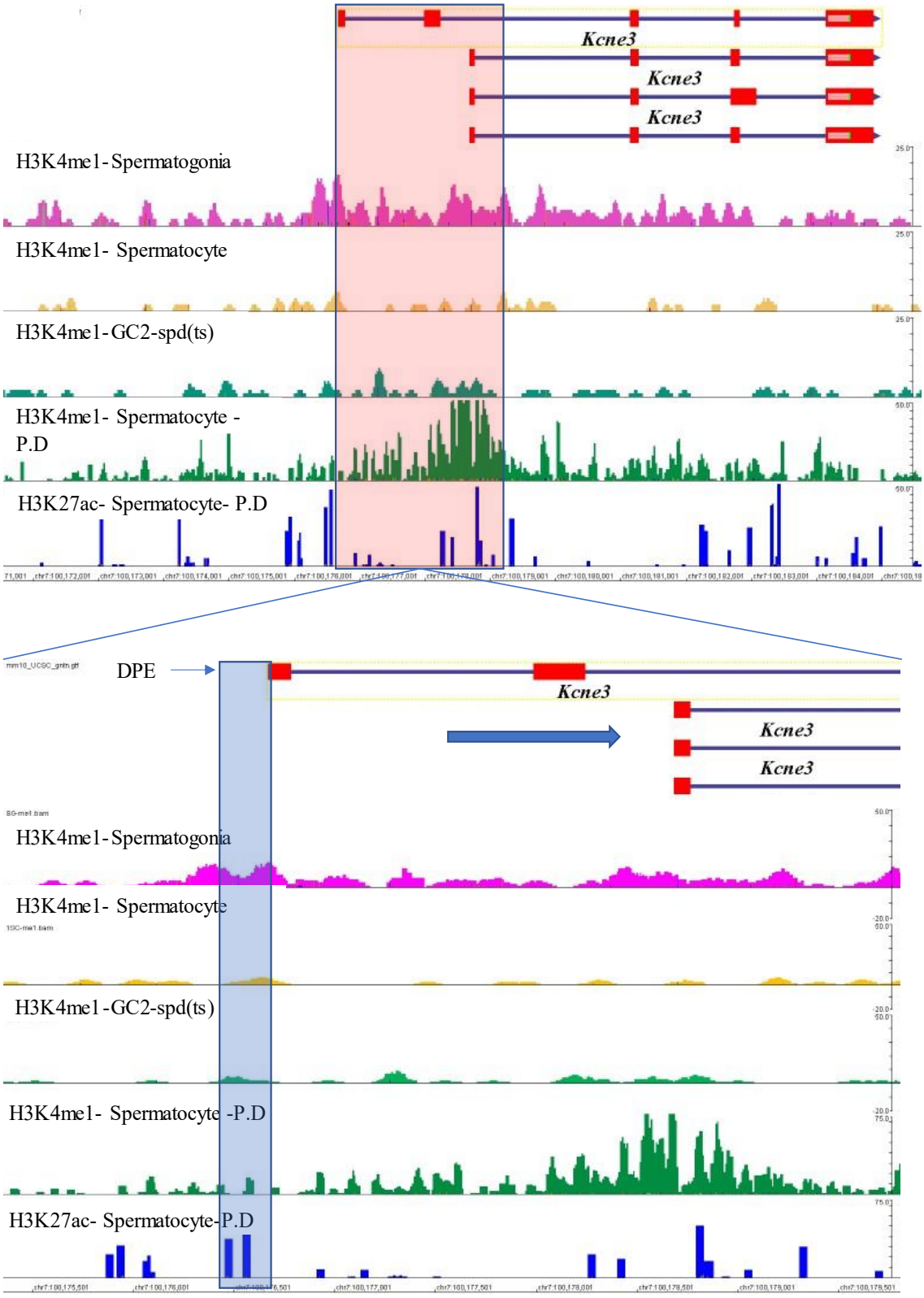


Fig. 1.06F Histone modification patterns KCNE3- DPE region (P.D- Public Data)

Fig. 1.06G

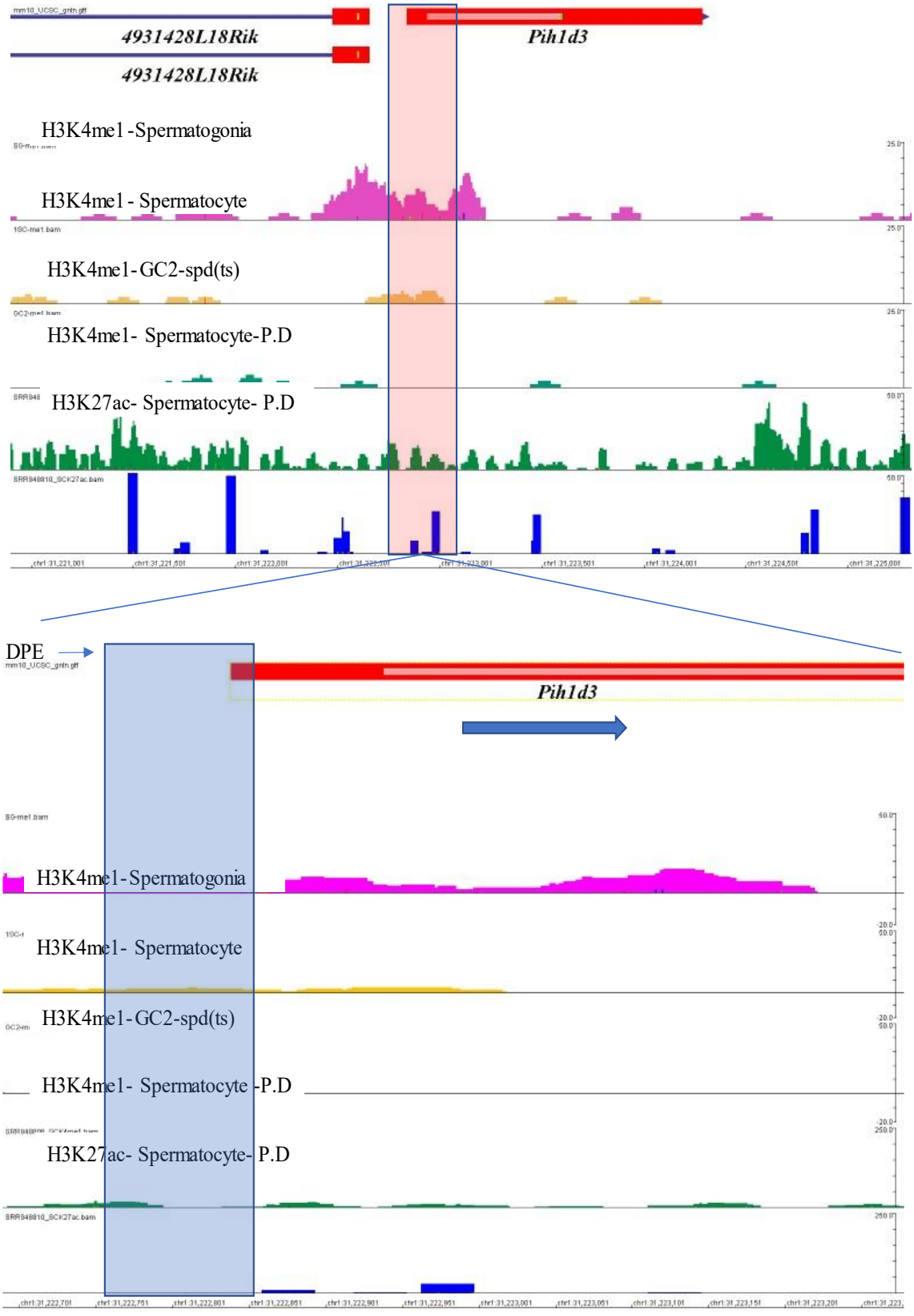


Fig. 1.06G Histone modification patterns PIH1D3- DPE region (P.D- Public Data)

Fig. 1.06H

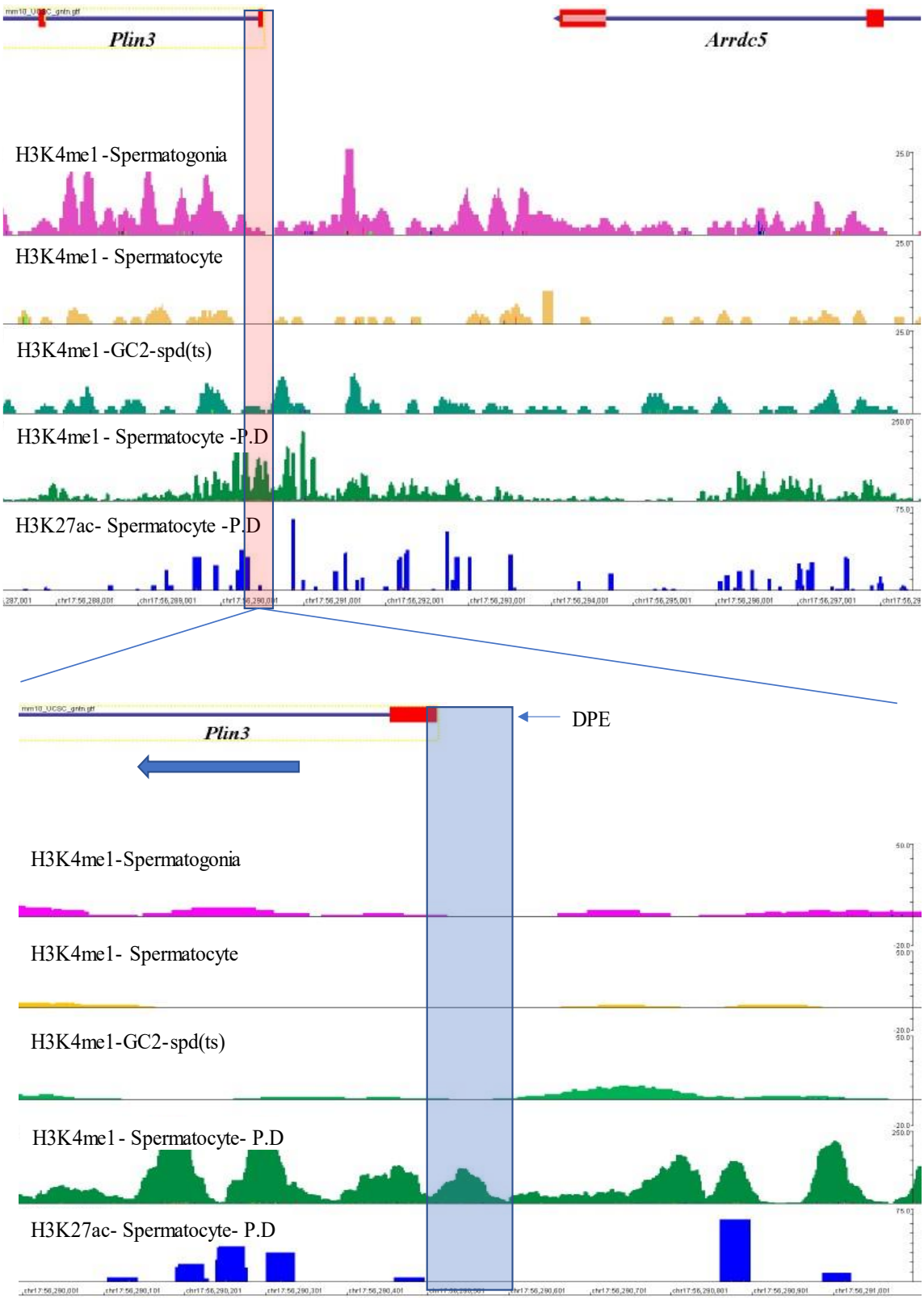


Fig. 1.06H Histone modification patterns PLIN3- DPE region

Fig. 1.06I

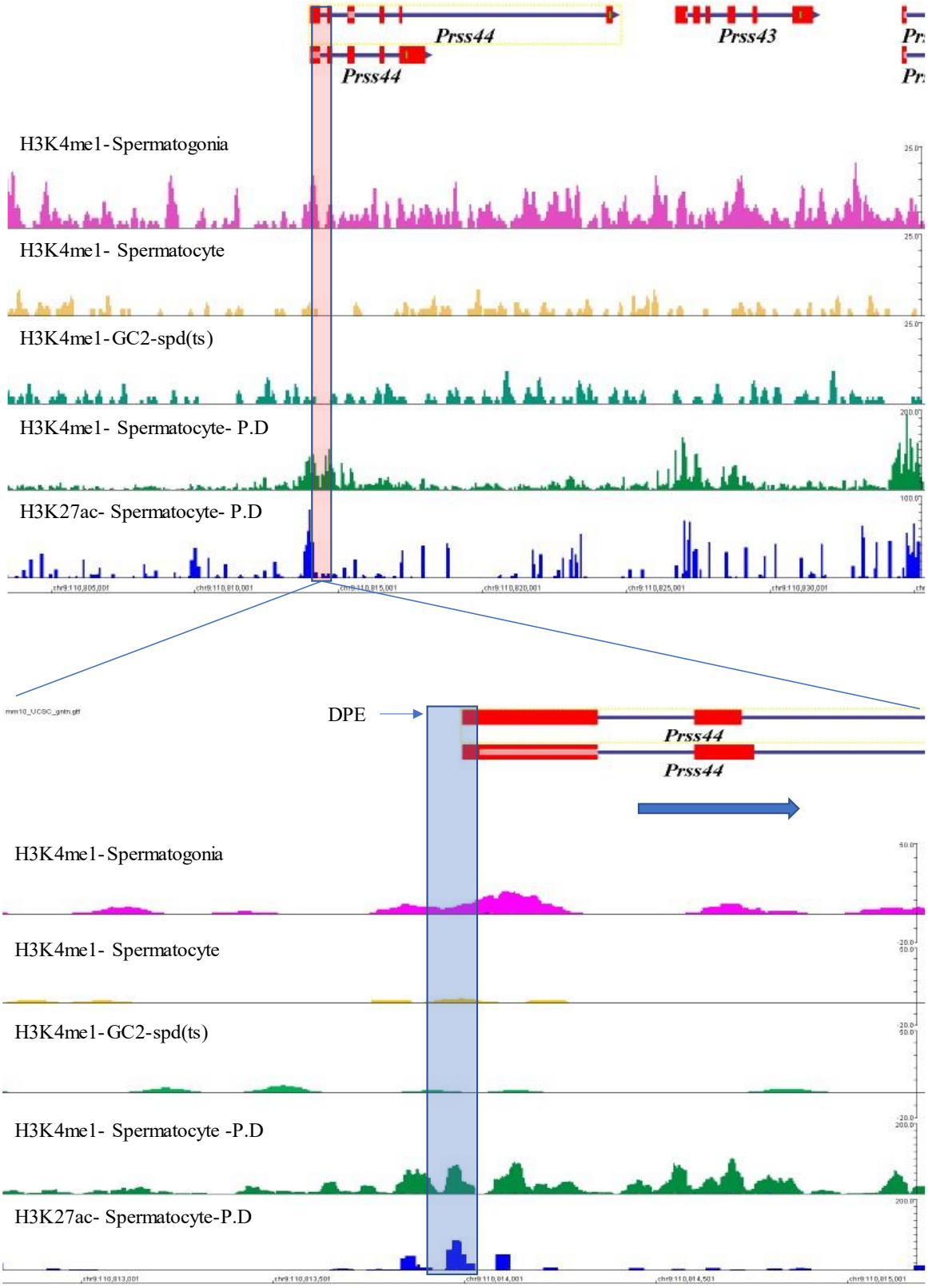


Fig. 1.06I Histone modification patterns PRSS44- DPE region (P.D - Public Data)

Fig. 1.06J

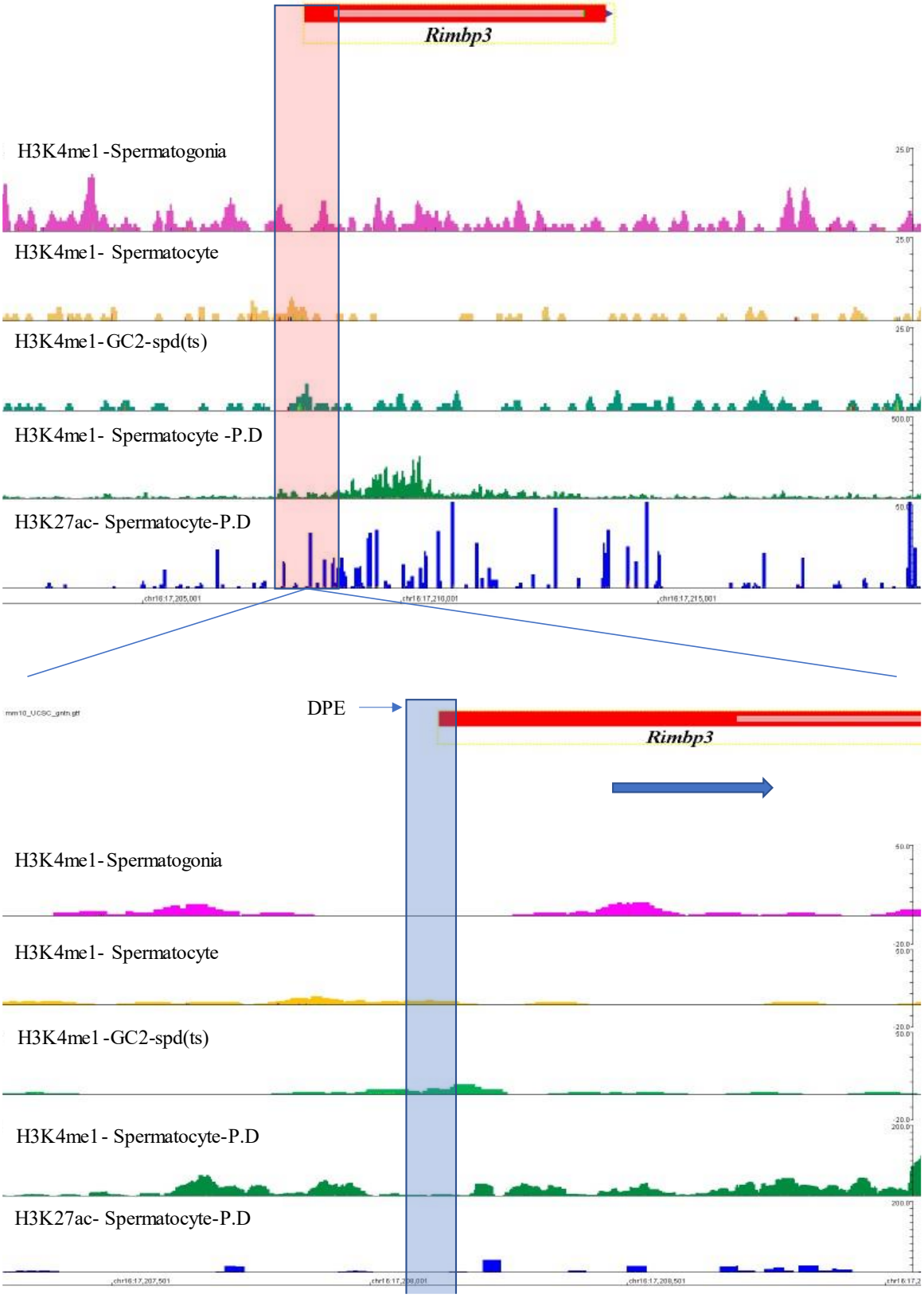


Fig. 1.06J Histone modification patterns RIMBP3- DPE region (P.D - Public Data)

Fig. 1.06K

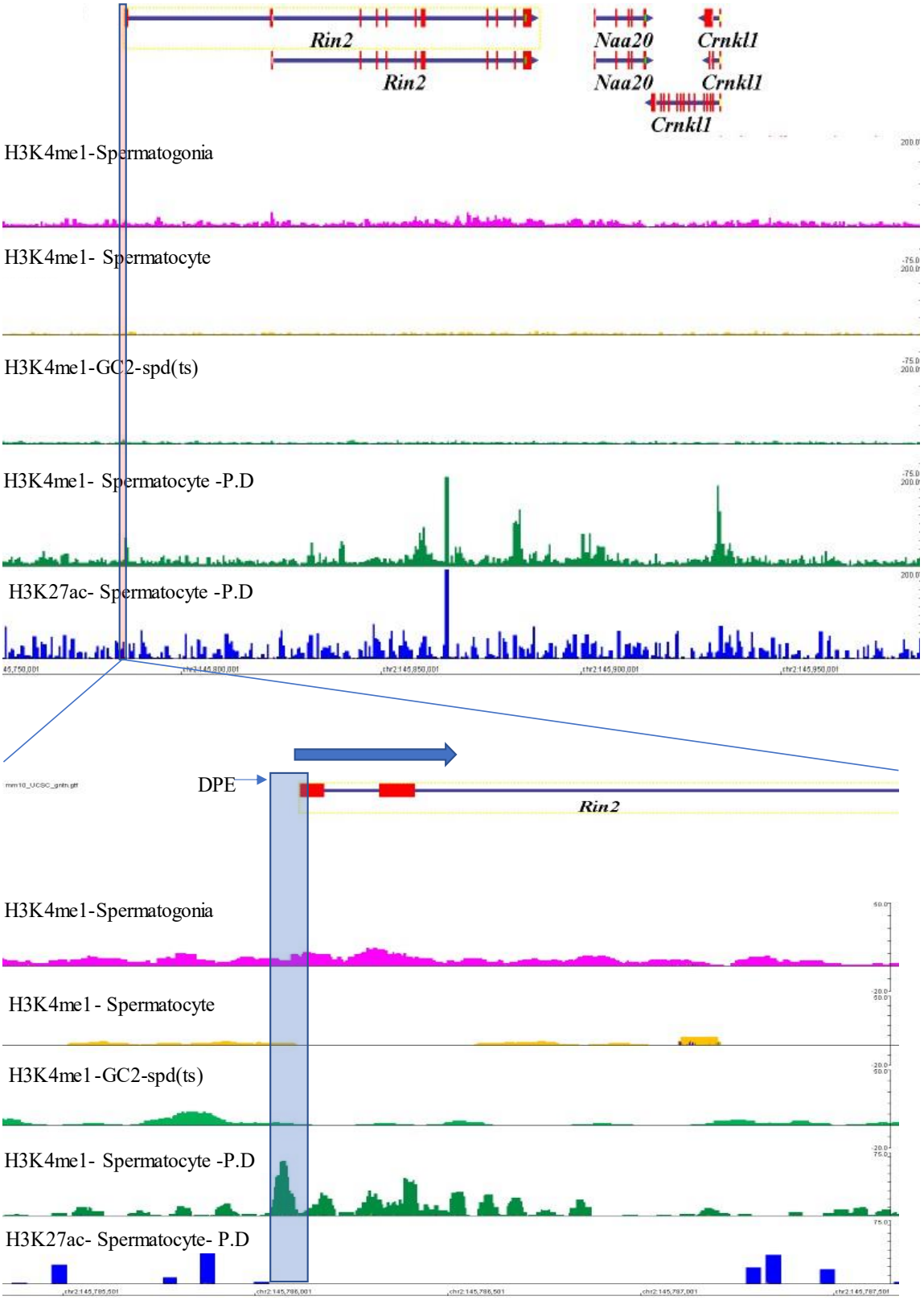


Fig. 1.06K Histone modification patterns RIN2- DPE region (P.D - Public Data)

Fig. 1.06L

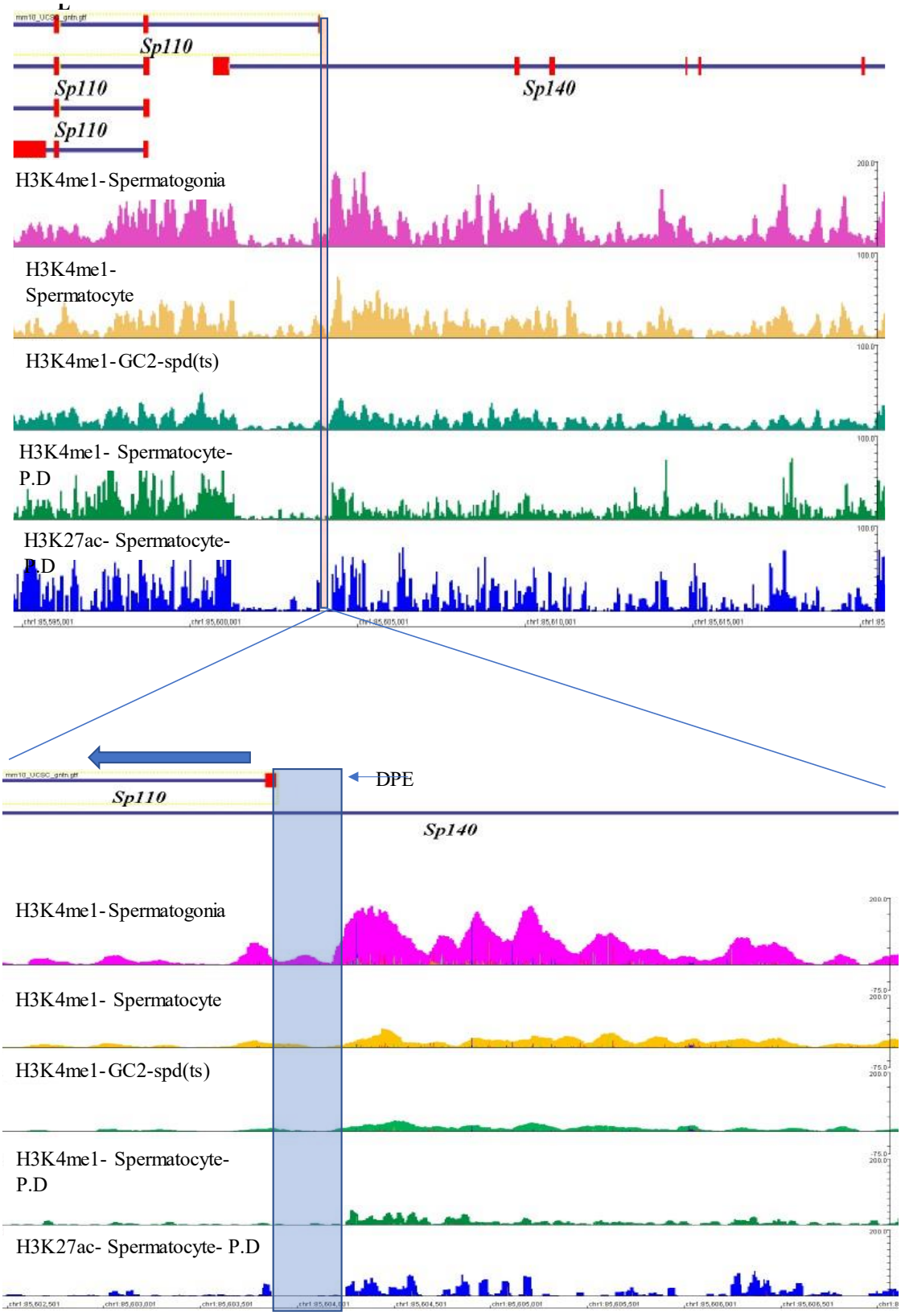


Fig. 1.06L Histone modification patterns SP110- DPE region (P.D - Public Data)

Fig. 1.06M

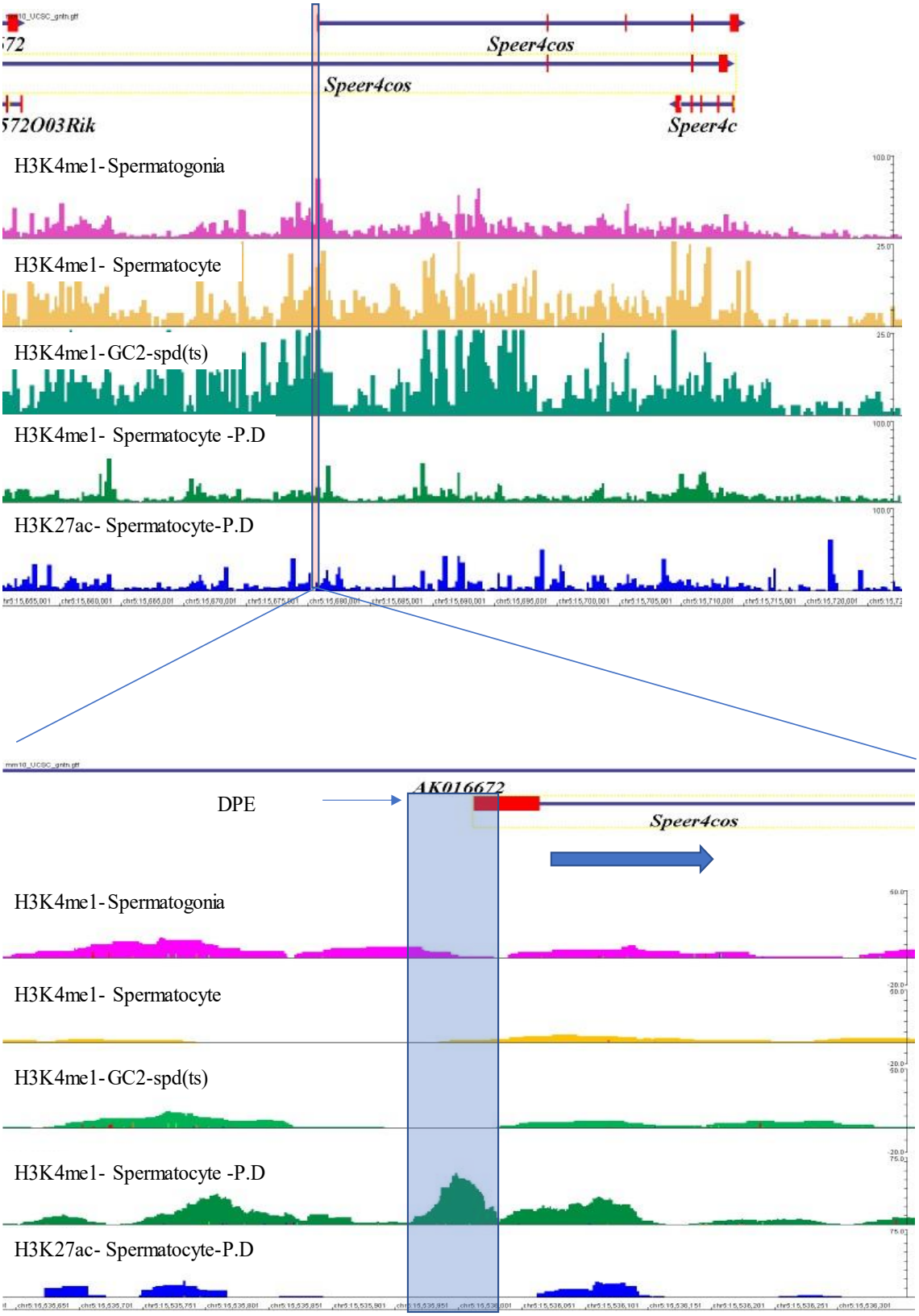


Fig. 1.06M Histone modification patterns SPEER4COS- DPE region (P.D - Public Data)

Chapter 2

Verification of dual promoter–enhancers by genome editing points to their significance in gene activation during spermatogenesis

Abstract

The ability of genome edition in any living organism gives many advantages over understanding transcription regulatory mechanisms of specific genes and specific biological processes on the molecular basis. Especially if you use mammalian cell lines specified for a different process, this may advance the process in a time-efficient manner to analyze and identify the mechanism of action. In this chapter, I targeted transcriptional regulatory elements which have dual regulatory action during spermatogenesis which are promoter and enhancer activity. In order to identify the *in vivo* mechanism of these regions, I have deleted 5 DPE sequences in GC-2spd(ts) cells by CRISPR/Cas9 genome editing. Then I used quantitative real-time polymerase chain reaction (qRT-PCR) with the deleted and control cells and demonstrated that all of them function as promoters of respective genes and 3 of them enhance the expression of a neighboring gene. Notably, one region that did not exhibit enhancer activity by reporter gene assay acted as an enhancer with the results of genome editing. Therefore, my current results strongly suggest that genome regions with H3K4me1 around TSSs of expressed genes are DPEs and that the number of DPEs dramatically increases on the onset of meiosis from spermatogonia to primary spermatocytes. These point to the significance of DPE in transcriptional activation during spermatogenesis.

Keywords: Duel promoter -enhancer, spermatogenesis, histone modification, qRT-PCR, CRISPR/Cas9

Introduction

Spermatogenesis is a lifelong cyclic process in which sperms are generated from spermatogonia stem cells. This is a multistep development process which is controlled by waves of gene activation, but the mechanism of action is not yet clearly understood. Recent research has identified the involvement of multifunctional genome elements as transcription regulators related to spermatogenesis [43]. Other than the traditional reporter analysis method, deletion of the suspected genomic region will help to understand the regulatory process and the relation of different genes toward the deleted elements.

Advances in the capability of editing the genomic information in biological organisms have elevated the molecular level understanding of the transcriptional regulatory mechanism of specific genes and specific biological processes. CRISPR/Cas 9 system has become a promising approach to achieve editing genome in DNA site-specific manner. CRISPR (clustered regularly interspaced short palindromic repeats) is a family of DNA sequences found in the organisms such as bacteria and archaea. These sequences are derived from the bacteriophages previously infected to those organisms. Later these fragments are used to detect and destroy DNA from similar bacteriophages during subsequent infections as immune response system in prokaryotes. Recent studies have used RNA guided endonuclease Cas9 (CRISPR-associated) proteins which recognize and introduce DNA double stranded break to the DNA at the site guided by the introduced RNA sequences. Therefore, CRISPR/Cas system can be programmed by changing the gRNA sequence. However, it is a critical issue to avoid the cleavage of the unintended off-target genes. In general, this is a double stranded breakage in DNA and, this modification

becomes heritable in the genome[95]. In this chapter, I have used the CRISPR direct web tool (http://crispr.dbcls.jp/), which provides an efficient selection of CRISPR/Cas target sites with reduced numbers of potential off-target candidates make specific cleavage in the intended site.

In this chapter, I have used GC-2spd(ts) cells to study the molecular mechanism involved in the production of primary spermatocytes related to meiosis. GC-2spd(ts) cell line has been widely used for the identification of fundamentals related to reproduction and development of males[89,96]. This cell line was established by co-transfecting the primary mouse testicular germ cell with the Simian virus 40 large tumor antigen gene and a gene coding for a temperature sensitive mutant of P53 making this cell line immortal. Especially this cell line showed the characteristics of early stages of round spermatids at 32°C[89]. I have deleted 5 selected DPE candidate regions and identified the promoter activity and target genes of the enhancer activity of some regions. My results suggest a dramatic increase in the number of DPE in spermatogenesis and point to the significance of DPEs in transcriptional activation during spermatogenesis.

Materials and Method

Cell culture

Mouse spermatocyte-derived GC-2spd(ts) cells were cultured in Dulbecco's Modified Eagle Medium (Fujifilm Wako) containing 10% FBS supplemented with 100 U/ml penicillin, 100 µg/ml streptomycin, and 292 µg/ml L-glutamine (Invitrogen, Carlsbad,

CA, USA). Culture cells on a 6-cm dishes were used for each analysis of RNA-seq and ChIP-seq.

Genome editing

Guide RNA (gRNA) sequences were designed for five selected DPE candidates using the CRISPRdirect web tool [64]. Synthesized complementary oligonucleotides (Table 2.01) were annealed and ligated into a BbsI-digested pX330 vector [65], and the completed constructs were confirmed by DNA sequencing. The constructs were transfected into GC-2spd(ts) cells together with pKO-Select-Puro and selected with 1 μg/μl puromycin for 7 days. The resulting cells were used for DNA and RNA isolation. The isolated DNA was used for PCR to see whether the DPE region was deleted with a primer pair listed in Table 2.02. The isolated RNA was used for qRT-PCR.

Quantitative RT-PCR

Total RNA was isolated from the cells selected as above using ISOGEN II (Nippongene, Tokyo, Japan) according to the manufacturer's instructions. After treatment with TurboDNase (Thermo Fisher Scientific, Waltham, MA, USA), cDNAs were synthesized using the oligo(dT) primer and Superscript III (Thermo Fisher). Quantitative PCR (qPCR) reactions were performed using Power SYBR Green Master Mix (Thermo Fisher) in the 7300 real-time PCR system (Applied Biosystems, Foster City, CA, USA). Primers used for this qPCR are listed in Table 2.03. Expression levels were normalized to that of the housekeeping *Gapdh* gene.

Statistical analysis

The results were expressed as means ± standard deviation (S.D.). Student's t-test was performed using Microsoft Excel statistical analysis functions (Microsoft Corp., Redmond, WA, USA). $P < 0.05$ was considered statistically significant.

Result

Native DPE activity in GC-2spd(ts) cells by genome editing

First, I attempted to check the native DPE activity in GC-2spd(ts) cells. As this method was somewhat time consuming, I had to focus on few regions based on the results of the reporter assay namely RIN2-DPE, PLIN3-DPE, DNMT-DPE, SP110-DPE, and SPEER4COS-DPE (nomenclature of the selected regions were explained in chapter 1). RIN2-DPE (Fig. 2.01) DNMT-DPE (Fig. 2.03) and SPEER4COS-DPE (Fig. 2.06) regions were selected in order to confirm the promoter and enhancer activity as exhibited by the results of reporter gene assay. PLIN3-DPE (Fig. 2.02) region was selected as the region only with the enhancer activity and SP110-DPE (Fig. 2.04) region without any promoter and enhancer activity. Furthermore, these

DPEs were considered based on the availability of the H3K4me1 mark in all the cell types. All the selected DPE regions contain the H3K4me1 mark in both the spermatocytes and GC-2spd(ts) cells. H3K4me1 marks were not present in RIN2-DPE (Fig. 2.01), DNMT-DPE (Fig. 2.03), and PLIN3-DPE (Fig. 2.02) regions in spermatogonia.

I used the CRISPR/Cas9 system to delete these regions and checked the effect on gene expression. As described in the materials and method section, I designed two guide RNAs complementary to both ends of each DPE candidate region. As a control, I transfected a vector without any guide RNA sequence. After selection with an antibiotic, genomic DNAs were purified to confirm the deletion of each candidate DPE sequence. Subsequently, I accessed the effect of deletion of the selected regions on the expression of adjacent genes by qRT-PCR.

For three DPE candidates, PCR with genomic DNA exhibited the expected band size at the lower position in the deleted cells compared to the control. DNA sequencing of the PCR product confirmed the deletion of respective DPEs by 688 bp in RIN2-DPE (Fig. 2.01C), 309 bp in PLIN3-DPE (Fig. 2.02C), and 344 bp in DNMT-DPE (Fig. 2.03C). On the other hand, I was unable to detect the deletion of the SP110-DPE and SPEER4COS-DPE by genomic PCR, presumably because of the miniature population size of the deleted cells.

The following are detailed descriptions of the result for each DPE candidate.

RIN2-DPE: As shown in the results of genomic PCR, a band was detected at a lower position in the deleted cells than in the control cells (Fig. 2.01C). DNA sequencing of the PCR product confirmed the deletion of respective DPE by 688 bp overlapping with RIN2-DPE. qRT-PCR results for the deleted RIN2-DPE region showed significant downregulation of the *Rin2* gene (Fig. 2.01D), confirming the promoter activity of this DPE. However, the two genes at proximity to the *Rin2* gene did not show any significant

reduction of the expression levels (Fig.2.01E, F), even though RIN2-DPE showed significant enhancer activity by reporter gene analysis. The enhancer target may be located far downstream or upstream.

PLIN3-DPE: As shown in Fig. 2.02, PCR with genomic DNA exhibited the expected band size at the lower position in the deleted cells than the control (Fig. 2.02C), and DNA sequencing indicated that a 309-bp sequence overlapping with this DPE was successfully deleted. The deletion resulted in the downregulation of the *Plin3* gene itself (Fig.2.02D), providing evidence of successful removal of the *Plin3* promoter. Concurrently, expression of the *Fem1* gene, which is upstream of the *Plin3* gene, was found to be significantly decreased (Fig.2.02E), while *Ticam1* and *Unrf1* genes were not affected (Fig. 2.02F, G). These findings provide strong evidence of DPE activity in the PLIN3-DPE region.

DNMT-DPE: My next candidate region to check DPE activity was DNMT-DPE (Fig. 2.03 A, B). DNA sequencing of the product of genome PCR confirmed the successful deletion of a 344-bp sequence of respective DPE (Fig 2.03C). qRT-PCR showed significant downregulation of the *Dnmt1* gene expression level, supporting its promoter activity (Fig. 2.03D). However, I could not identify the enhancer target which should be downregulated by the deletion of the DNMT-DPE sequence (Fig. 2.03E-H), although this region showed highly significant enhancer activity by the reporter gene analysis. Interestingly, I observed the upregulation of a gene, *A230050p20Rik (Shfl)* (Fig. 2.03F). This gene codes for Shiftless antiviral inhibitor of ribosomal frameshifting protein which inhibits programmed-1 ribosomal frameshifting (-1PRF) of a variety of mRNAs from viruses and cellular genes according to the sequence similarities [97]. The web tool Sequence Manipulation Suite:

CpG Islands [98] predicted a 6666-bp sequence starting with "AACCAGCATG" in the 1000 bp of 5'UTR region to be a CpG island. Therefore, the upregulation of *A230050p20Rik (Shfl)* could be a result of changes in the DNA methylation status in the highly CpGs rich promoter of this gene as that the *Dnmt* gene is vital for the regulation of DNA methylation.

SP110-DPE: Followingly, I checked the effect of the deletion of SP110-DPE (Fig. 2.04A, B). Somehow, I could not successfully amplify the genome region presumably due to the smaller number of deleted cells. Even though the SP110-DPE region showed lower promoter activity than the control by reporter assay, the deletion of this region in GC-2spd(ts) cells by the CRISPR/Cas9 system significantly downregulated the *Sp110* gene, calming the promoter activity of SP110-DPE (Fig. 2.04C). Then, I checked the expression level of the adjacent genes by qRT-PCR and observed significant downregulation of the *A630001g21rik* gene but not *Sp140* (Fig. 2.04D, E). These results suggest that the Sp110-DPE can be an enhancer of the *A630001g21rik* gene.

SPEER4COS-DPE: I could not detect the expected size of deleted products by genome PCR. My RNA-seq data showed that the expression level of the *Speer4cos* gene was low in GC-2spd(ts) cells, and indeed qRT-PCR results demonstrated the low level of expression. Nonetheless, deleted cell lines showed a slight downregulation of the *Speer4cos* gene, although it was not statistically significant (Fig. 2.05C). The deletion of SPEER4COS-DPE (Fig. 2.05A, B) led to the downregulation of *Cacna2d1*, one of the neighboring genes (Fig. 2.03D), showing the possible effect as an enhancer region. For the other linked genes,

4930519h02Rik was not expressed in GC-2spd(ts) cells by my RNA-seq data and I could not detect the expression of *Gm21190* by qRT-PCR in either deleted or control cells.

Discussion

The CRISPR-Cas9 system is a widely used genome editing tool available to manipulate any genomic sequence, enabling the creation of isogenic cell lines and animal models for identification of transcription regulation. In this study, I described editing genomic regions which were identified as candidates for dual promoter enhancers as described in chapter 1. I used GC-2spd(ts) cell line for this analysis which was derived from the spermatocytes. Guide RNA sequence was selected as all the 20 basses can bind with high specificity with only one target. However, when considering 12 bases or 8 bases there can be possible off-targets, and they did not directly interfere with the affected genes. I checked this result for different sets of cell groups and the reproducibility of the result suggested that the DPE activity is constant, confirming the specificity of the selected guide RNA sequences.

Genotyping was performed after selection of the cell following the successful transfection of guide RNA carried vectors which require for CRISPR/Cas9 based genome editing. I could not detect the expected PCR product for the deleted SP110-DPE region and SPEER4COS-DPE regions. qRT-PCR results demonstrated that the significant downregulation of the *Sp110* gene and slight downregulation of the *Sperr4cos* gene suggesting the loss of transcription regulatory elements related to the activity of these genes. I used CRISPR/Cas9 system to delete the promoter region of these genes and the qRT-PCR

results confirmed the deletion of promoter region which were the DPE candidate regions. These results indicate that genomic PCR may not be sensitive enough to detect the deleted region as a smaller number of deleted cells may lead to a reduced amount of template or due to difficulty in amplifying these regions.

Selected 5 DPE candidate regions were deleted by the CRISPR/Cas9 system. Especially, PLIN3-DPE, SPEER4COS-DPE, and SP110-DPE showed both native enhancer and promoter activity. I was able to identify *Fem1a* and *A630001G21Rik* genes as target genes for enhancer activity of PLIN3-DPE and SP110-DPE regions, respectively. However, these DPE regions did not show enhancer activity by reporter gene assay (Chapter 1). This indicates the possibility that even if a DPE did not show enhancer activity by reporter gene assay, it can be an actual DPE in native cells. In contrast, in the case of DNMT-DPE and RIN2-DPE regions, while I observed the downregulation of genes (*Dnmt1* and *Rin2*) after the deletion of the promoter regions I could not identify the target genes that affected by enhancer activity. This could be due to the localization of their target genes were far upstream or downstream of the DPE region, which agrees with previous findings that enhancer activity could exert over the genes far from it [85]. Therefore, my results strongly suggest that most of the DPE candidates I detected can actually function as DPEs during spermatogenesis (Table 2.04).

Finally, it is interesting to note that transcription start site associated H3K4me1 rich genome elements related to spermatogenesis can act as dual promoter enhancer regions. The advent of the CRISPR/Cas9 system allows the investigation of the gene interaction to

identify their promoter and enhancer region much easier. Apart from that, this result suggests that *in vivo* activity can differ from the *in vitro* analysis results.

Interestingly, my current data identified 337, 2176, and 69 DPE candidate regions in spermatogonia, primary spermatocytes, and GC-2spd(ts) cells, respectively (Table 2.05). These results show an increment of genomic regions containing H3K4me1 at the promoter by 7 folds during spermatogenesis from spermatogonia to primary spermatocytes. Thus, I conclude that the dramatic variations in the number of DPE candidate regions in different stages reveal the physiological significance of DPEs in spermatogenesis, especially in transcriptional activation in primary spermatocytes.

Table 2.01 Guide RNA Oligonucleotide Used for Genome Editing

Designation	**Forward 5'-3'**	**Reverse 3'-5'**
[Guide RNA]		
Itch Start 1	CACCTCTAATCCTACCGAATTAAT	AAACATTAATTCGGTAGGATTAGA
Itch End 1	CACCAGGAGTGTGGCGCTCGAGTC	AAACGACTCGAGCGCCACACTCCT
Rin2 Start 1	CACCGATAGTCTCGGCGATGATTC	AAACGAATCATCGCCGAGACTATC
Rin2 End 1	CACCTCCTCAGGCTGCCGCGCAGA	AAACTCTGCGCGGCAGCCTGAGGA
Plin3 Start 1	CACCGTAGGGTTTCCTAGCGTGCA	AAACTGCACGCTAGGAAACCCTAC
Plin3 End 1	CACCTGTTAACCTCAGACGGCTGC	AAACGCAGCCGTCTGAGGTTAACA
Dnmt1 Start1	CACCCAGCTAAGGGCACGCGCGAC	AAACGTCGCGCGTGCCCTTAGCTG
Dnmt1 End 1	CACCGGTCAGCCAAGTCTATACAG	AAACCTGTATAGACTTGGCTGACC
Spear4cos Start	CACCAGAGTCTGTCTATGTAGCCTTGG	AAACCCAAGGCTACATAGACAGACTCT
Spear4cos End	CACCTAATTCTATAATTTATAAAAGGG	AAACCCCTTTTATAAATTATAGAATTA
Sp110 Start	CACCAAAGGCATAGCCCCGATTAAAGG	AAACCCTTTAATCGGGGCTATGCCTTT
Sp110 End	CACCCACACACACGATAGGAGAGATGG	AAACCCATCTCTCCTATCGTGTGTGTG

Table 2.02 Primers Used for Genotyping

Designation	Forward 5'-3'	Reverse 3'-5'
[Genotyping]		
ITCH-DPE	GTTTGCTAATTATCACGTGCCTCT	GGCTCGGCTCTCTCCCAAC
RIN2-DPE	ACTTCCTACTTTCTGTGTTGCTTGT	CCTGAACAGCACACCACTTC
PLIN3-DPE	AATCCACCTTTTTGGACGCTC	GGTAACTCAGTAAGGGTGGGTC
DNMT1- DPE	TCCCTCAAGCTCCCAGTCAAT	TGGAGTTACAGGGCTTCGAGT
SPEER4COS- DPE 1	GCTGTGGTTTGCTTGGGAAC	ACACAGGAGGGACACCACT
SPEER4COS- DPE 2	TCTGAGAGTGCTATGGGCCT	ACAGGAGGGACACCACTCAT
SPEER4COS- DPE 3	ACATACACACATACACACACAC	CAGCCAACCCATTTCAGAC
SP110 -DPE 1	CCTCAAGGACACACTTGGC	TAATGCAGCCATCCCTACGC
SP110 -DPE 2	GATTCTCCTCAAGGACACAC	GACACAGCACACTACTGGTATC
SP110 -DPE 3	ACACAACTCCTTCCTTTCATTC	CACACACCTTCGTCTCATC

Table 2.03 Primers Used for qRT-PCR

Designation	Forward 5'-3'	Reverse 3'-5'
Rin2	GGGAGCAACTAAAGAGGATGGA	AGAATGCTGAGCCTGTTGGA
Crnkl1	ATCGTGAGCAAACGGAGGTT	TGTCAGCTTCCGCATCACTT
Naa20	CCTGTTCCGCTTCAACAACA	GCTTTGCCCATAATATAACCCATT
Plin3	GCCGAAGCTCAAGCTGCTAT	CTGACGTGGGGGTAGTTCTC
Ticam1	ACCAGCTCAAGACCCCTACA	TGGAGACATGTTCCGAACCG
Uhrf1	TGACATCATGTACCATGTCAAGT	CACGAGCACGGACATTCTTG
Fem1a	GGGGATGGGCTTCTAACCAA	GTGGAGTCATAGGCAGCTGTT
Dnmt1	TGGCGTCATAGCCCATAAGC	TTGCCTTCTGCACAGGAACA
Ppan	ACCTGTTCCCGTCCATCAAC	AGTCCAGCTCTTGGGAGTCA
A230050P20Rik	AATCGGAGAGAGCCCCATGT	GGGAGGTGGTTCTGCTTACG
S1pr2	TCCTGCCGCCTCCTTTTTAG	GCCCGGTCTGCATAGAAAGT
Mrpl4	CCAGGGCTGAGCTACTGAAC	CAGGGACAGTGGCGTATGAG
Sp110	AGCTCTTGCTCCTTAGCCT	GGGAGTCAGGATCTCCAAGG
A630001G21Rik	CTCCTTCATCACAGACAA	TACACCACTTCCACTATG
9930111H07Rik	CTGCCTGCACCTCCAGATAC	GCCTTGGGTCCTTGGATGAT
Sp140	CTCAGAGGGGGTTCGGTCTA	GGATTGTCCTGCTGCTGAGT
Pclo	TTCTCTAAACGCAATGCCTACA	TCTGATATCACTAGCAAGAAAGCA
Speer4cos	ATAGCCATTGAGCACATC	CACAGAGTTAGATACATAGGAA
Cacna2d1	GCCGTGAATAACATTACAG	CTCCATCCGTGAATAACA

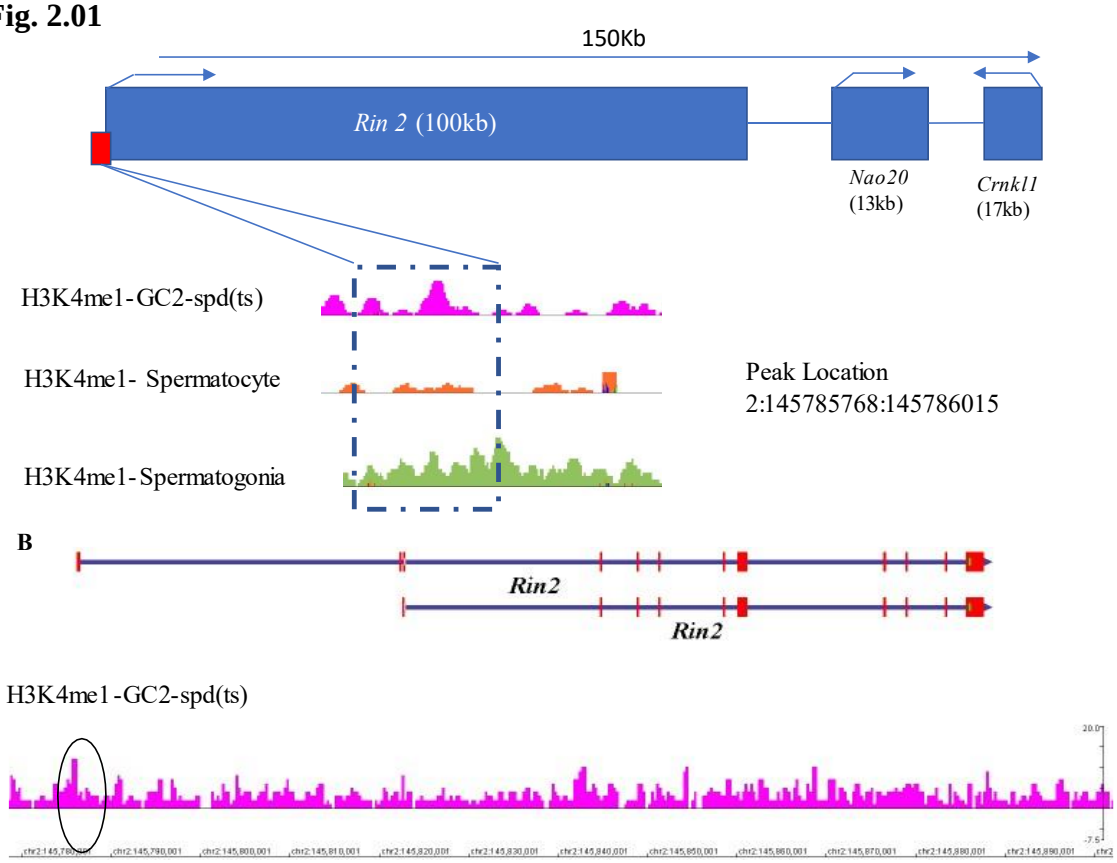
Table 2.04 Summary of DPE Activity by Deletion of DPE Region

DPE	**H3K4me1 Peak**			**By Deletion of DPE**	
	SG	SC	GC2	Promoter activity	Enhancer Activity
RIN2	X	○	○	○	?
DNMT1	X	○	○	○	?
PLIN3	X	○	○	○	○
SP110	○	○	○	○	○
SPEER4COS	○	○	○	Δ	○

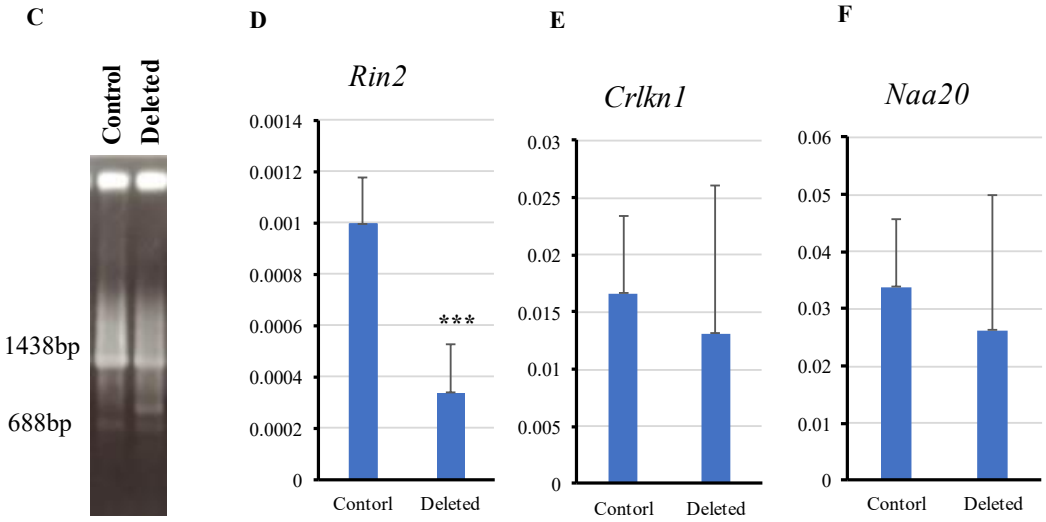
Table 2.05. **Summary of DPE Candidates by ChIP-seq Results**

Types of Cells	The Number of Peaks
Spermatogonia	337
Primary Spermatocytes	2176
GC-2spd(ts) Cells	69

Fig. 2.01



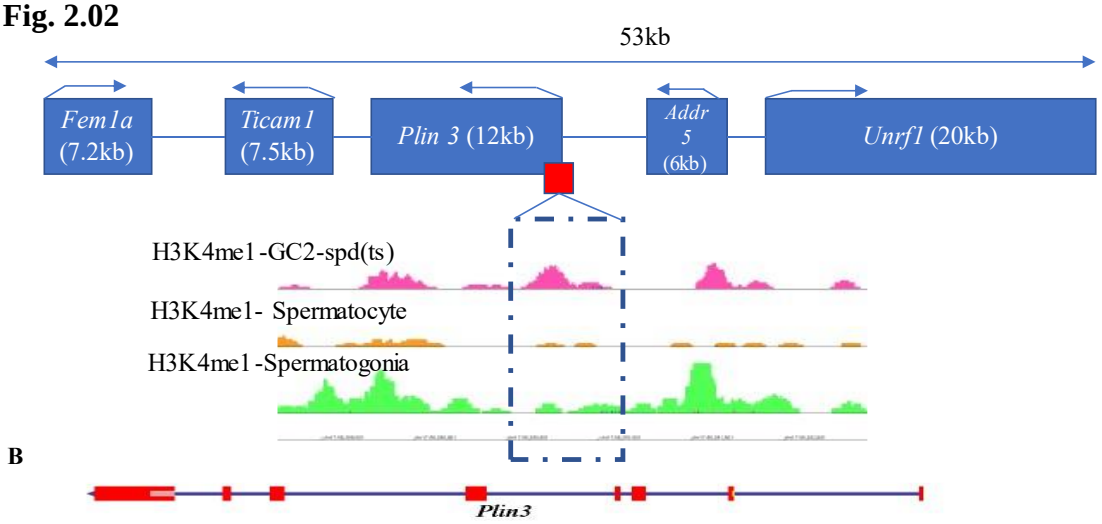
Amplified Region 2:145785668:145786420



Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 N= 8

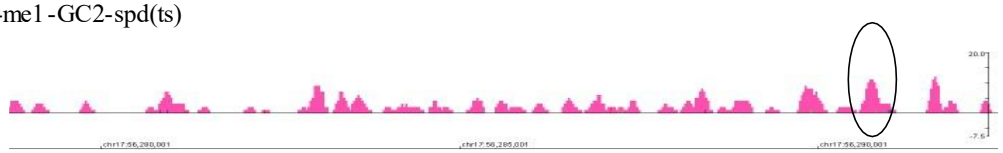
Fig. 2.01 Identification of a dual promoter activity of RIN2-DPE in mouse spermatocytes/GC-2spd(ts) by genome editing. A) Three protein-coding genes which are closed to the RIN2-DPE were indicated in blue boxes and the transcription direction of each gene is indicated in arrowhead depicted by blue boxes with the name and length. The position of RIN2-DPE is shown by a red square at the TSS of the *rin2* gene. ChIP-seq data for three cell types were shown below the RIN2-DPE. In the blue dash line, the epigenetic mark of histone H3 lysine 4 monomethylating peak was indicated. B) Two transcript variants of *Rin2* are enlarged and drawn by indicating the exon-intron structure and red indicate the exons, and the parallel ChIP-seq data for H3K4Me1 modification patterns along the Rin2 gene region of GC-2spd(ts) was shown and around the representative peak for the RIN2-DPE was circled. C) A representative result of PCR for genotyping. GC-2spd(ts) cells were transfected with the CRISPR/Cas9 vectors containing and not containing gRNA sequences (control), and DNAs were isolated after selection. and the products were analyzed by electrophoresis and ethidium bromide staining. The band size confirmed by DNA sequencing is indicated. D) qRT-PCR for *Rin2* gene expression in GC-2spd(ts) cells edited by the CRISPR/Cas9 system. Total RNAs were isolated from the cells selected in (C) and qRT-PCR was performed for the *Rin2* gene. *Gapdh* was used as an internal control. Data are presented as means ± S.D. from eight independent experiments, and statistical significance was assessed by Student's t-test. $^{**}P < 0.001$. (E, F) qRT-PCR for *Crlkn1* gene and *Naa20* expression in GC-2spd(ts) cells edited by the CRISPR/Cas9 system. Total RNAs were isolated and qRT-PCR was performed as in (C). In the control sample, and test sample no significant difference was identified in the cell expression levels.

Fig. 2.02

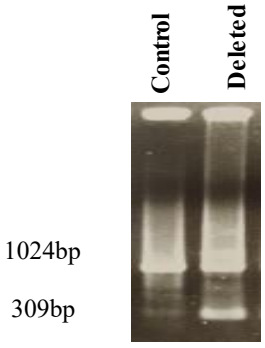


B

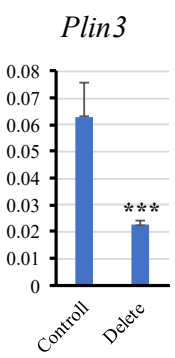
H3K4me1-GC2-spd(ts)



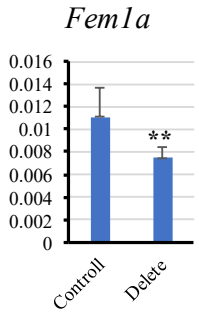
C



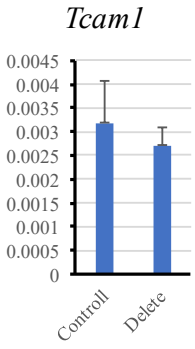
D



E



F



G

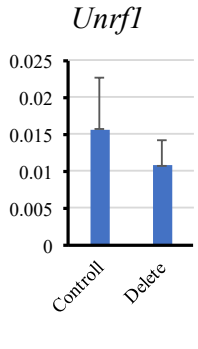
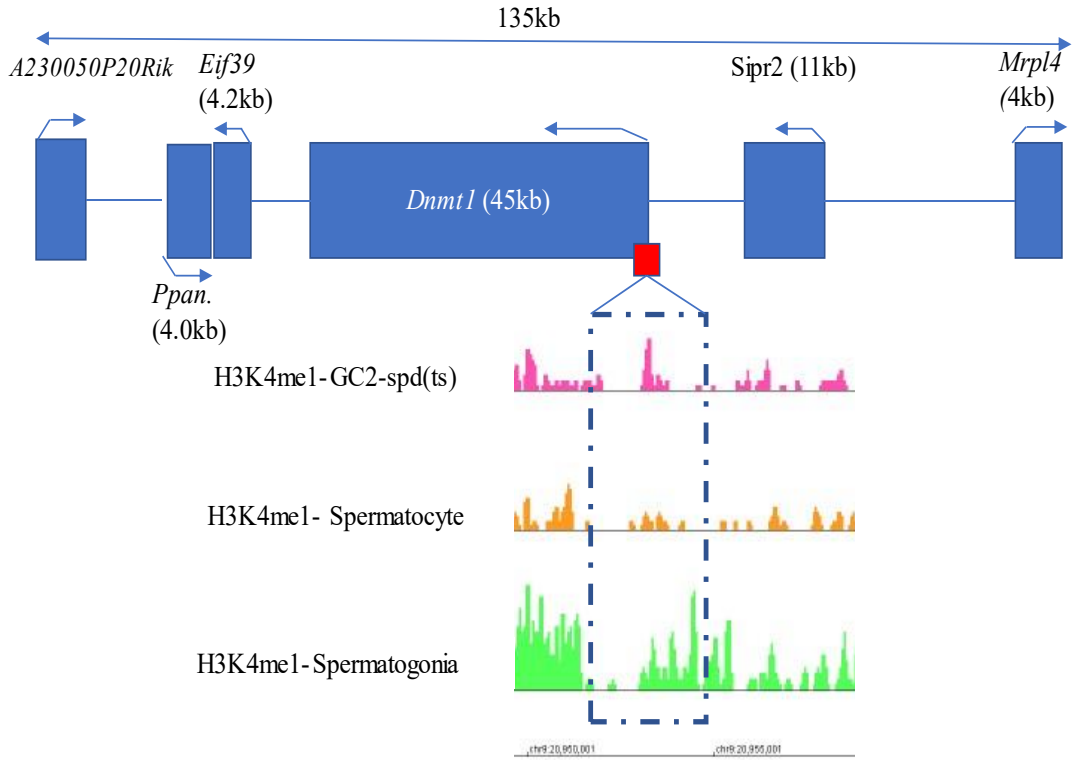


Fig. 2.02 Identification of a dual promoter activity of PLIN3-DPE in mouse spermatocytes/GC-2spd(ts) by genome editing. A) Four protein-coding genes which are closed to the RIN2-DPE upstream and downstream were indicated in blue boxes and the transcription direction of each gene is indicated in the arrowhead showed by blue boxes with the name and length. The position of PLIN32-DPE is shown by a red square at the TSS of the *Plin3* gene. ChIP-seq data for three cell types were shown below the PLIN3-DPE. In the blue dash line, the epigenetic mark of histone H3 lysine 4 monomethylating peak of PLIN3-DPE was indicated. B) *Plin3* protein-coding gene transcripts were enlarged and drawn by indicating the exon-intron structure, red indicate the exons, and the parallel ChIP-seq data for H3K4me1 modification patterns along the *Plin3* gene region of GC-2spd(ts) as shown below and representative peak for the PLIN3-DPE was circled. C) A representative result of PCR for genotyping. GC-2spd(ts) cells were transfected with the CRISPR/Cas9 vectors containing and not containing gRNA sequences (control), and DNAs were isolated after selection. and the products were analyzed by electrophoresis and ethidium bromide staining. The band size confirmed by DNA sequencing is indicated. D) qRT-PCR for *Plin3* gene expression in GC-2spd(ts) cells edited by the CRISPR/Cas9 system. Total RNAs were isolated from the cells selected in (C) and qRT-PCR was performed for the *Plin3* gene. *Gapdh* was used as an internal control. Data are presented as means ± S.D. from eight independent experiments, and statistical significance was assessed by Student's t-test. $^{**}P < 0.01$. (E) qRT-PCR for *Fem1a* gene expression in GC-2spd(ts) cells edited by the CRISPR/Cas9 system. Total RNAs were isolated and qRT-PCR was performed as in (C). Data are presented as means ± S.D. from four to eight independent experiments, and statistical significance was assessed by Student's t-test. $^{**}P < 0.01$. (F, G) qRT-PCR for *Tcam1* gene and *Unrf1* expression in GC-2spd(ts) cells edited by the CRISPR/Cas9 system. Total RNAs were isolated and qRT-PCR was performed as in (C).

In the control sample, and test sample no significant difference was identified in the cell expression levels.

Fig. 2.03


135kb
A230050P20Rik
Eif39
(4.2kb)
Sipr2 (11kb)
Mrpl4
(4kb)
Dnmt1 (45kb)
Ppan.
(4.0kb)
H3K4me1-GC2-spd(ts)
H3K4me1- Spermatocyte
H3K4me1-Spermatogonia

B

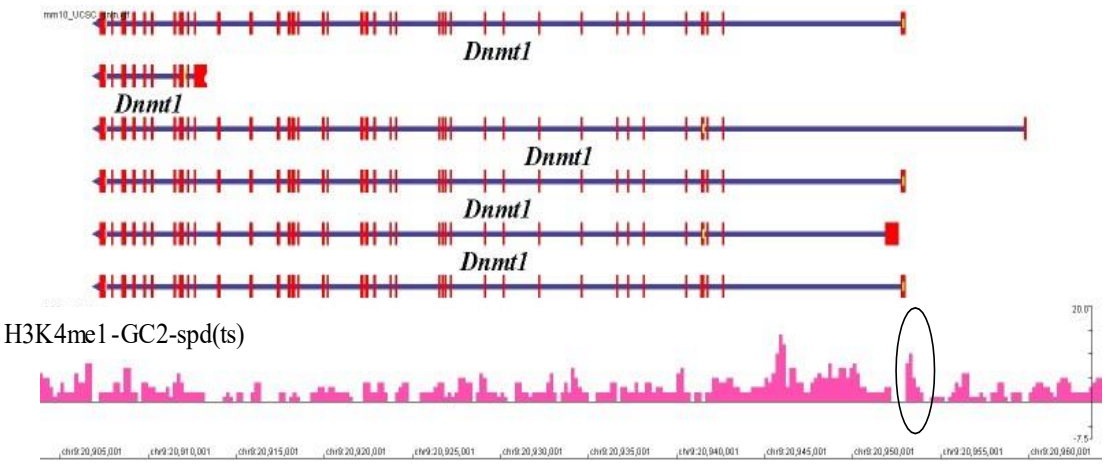

Dnmt1
Dnmt1
Dnmt1
Dnmt1
Dnmt1
H3K4me1-GC2-spd(ts)

Fig. 2.03

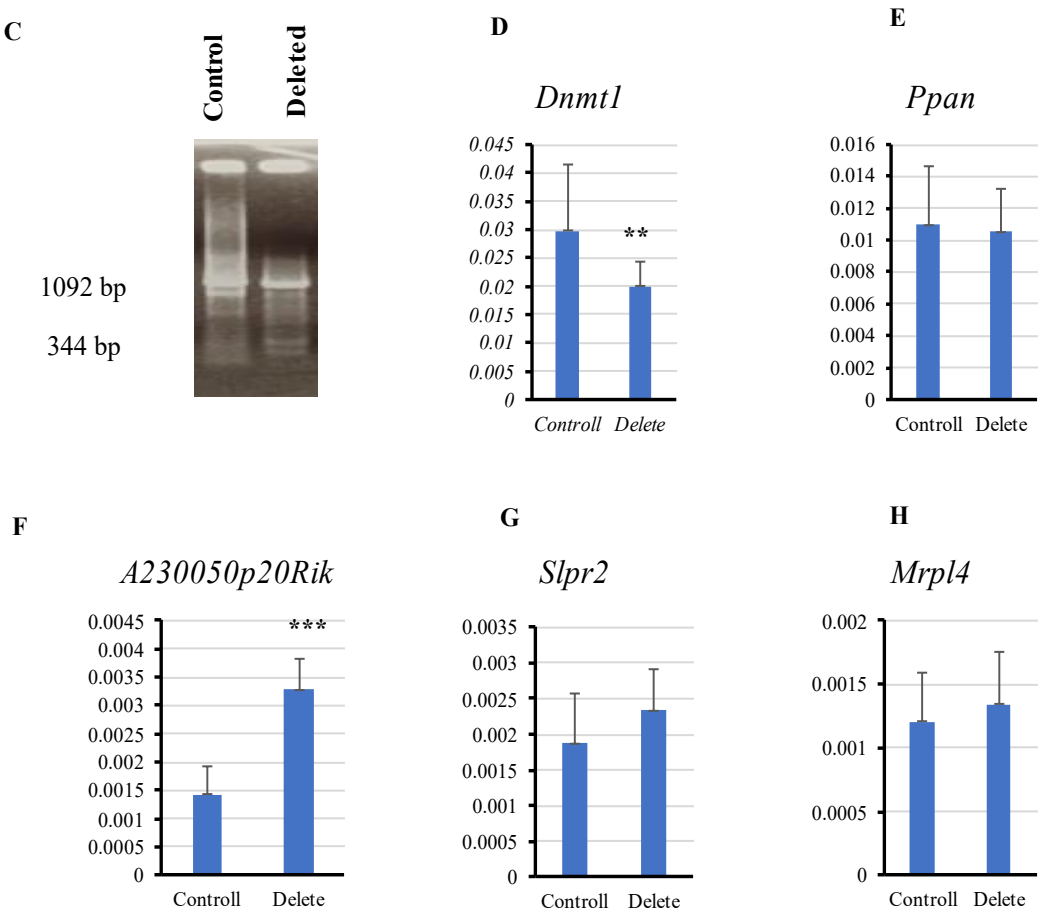


Fig. 2.03 Identification of a dual promoter activity of DNMT-DPE in mouse spermatocytes/GC-2spd(ts) by genome editing. A) Four protein-coding genes which are closed to the DNMT-DPE upstream and downstream were indicated in blue boxes and the transcription direction of each gene is indicated in the arrowhead showed by blue boxes with the name and length. The position of DNMT-DPE is shown by a red square at the TSS of the *Dnmt1* gene. ChIP-seq data for three cell types were shown below the DNMT-DPE. In the blue dash line, the epigenetic mark of histone H3 lysine 4 monomethylating peak of PLIN3-DPE was indicated. B) *Dnmt1* protein-coding gene transcripts were enlarged and drawn by indicating the exon-intron structure, red indicate the exons, and the parallel ChIP-seq data for H3K4Me1 modification patterns along the *Dnmt1* gene region of GC-2spd(ts) as shown below and representative peak for the DNMT-DPE was circled. C) A representative result of PCR for genotyping. GC-2spd(ts) cells were transfected with the CRISPR/Cas9 vectors containing and not containing gRNA sequences (control), and DNAs were isolated after selection. and the products were analyzed by electrophoresis and

ethidium bromide staining. The band size confirmed by DNA sequencing is indicated. D) qRT-PCR for *Plin3* gene expression in GC-2spd(ts) cells edited by the CRISPR/Cas9 system. Total RNAs were isolated from the cells selected in (C) and qRT-PCR was performed for the *Dnmt1* gene. *Gapdh* was used as an internal control. Data are presented as means ± S.D. from eight independent experiments, and statistical significance was assessed by Student's t-test. $^{**}P < 0.01$. (E) qRT-PCR for the *Ppan* gene expression in GC-2spd(ts) cells edited by the CRISPR/Cas9 system. Total RNAs were isolated and qRT-PCR was performed as in (C). (F) qRT-PCR for *A230050p20Rik* gene expression in GC-2spd(ts) cells edited by the CRISPR/Cas9 system. Total RNAs were isolated and qRT-PCR was performed as in (C). Data are presented as means ± S.D. from four to eight independent experiments, and statistical significance was assessed by Student's t-test. $^{**}P < 0.01$. (G, H) qRT-PCR for *Slpr2* gene and *Mrpl4* expression in GC-2spd(ts) cells edited by the CRISPR/Cas9 system. Total RNAs were isolated and qRT-PCR was performed as in (C). In the control sample, and test sample no significant difference was identified in the cell expression levels.

Fig. 2.04

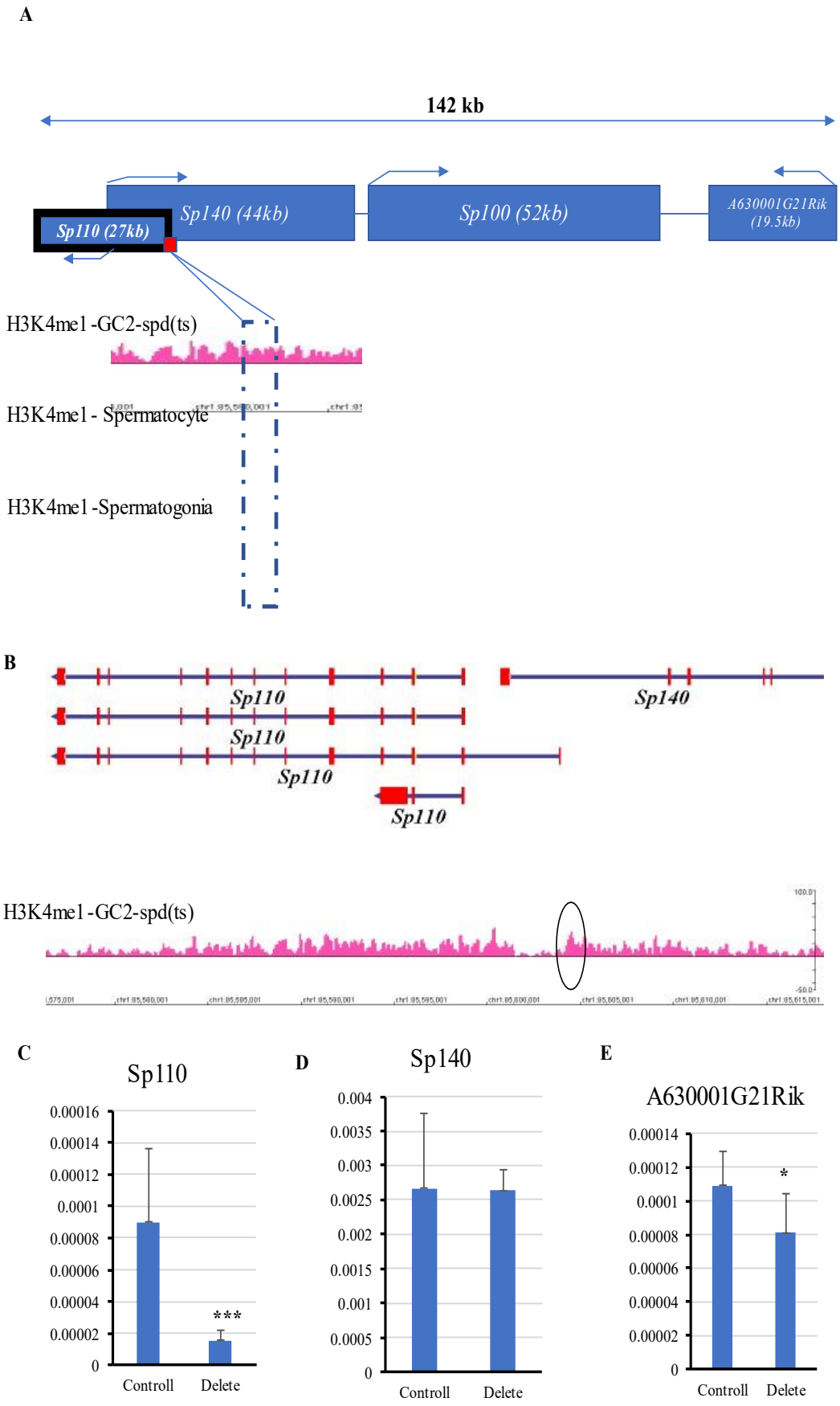
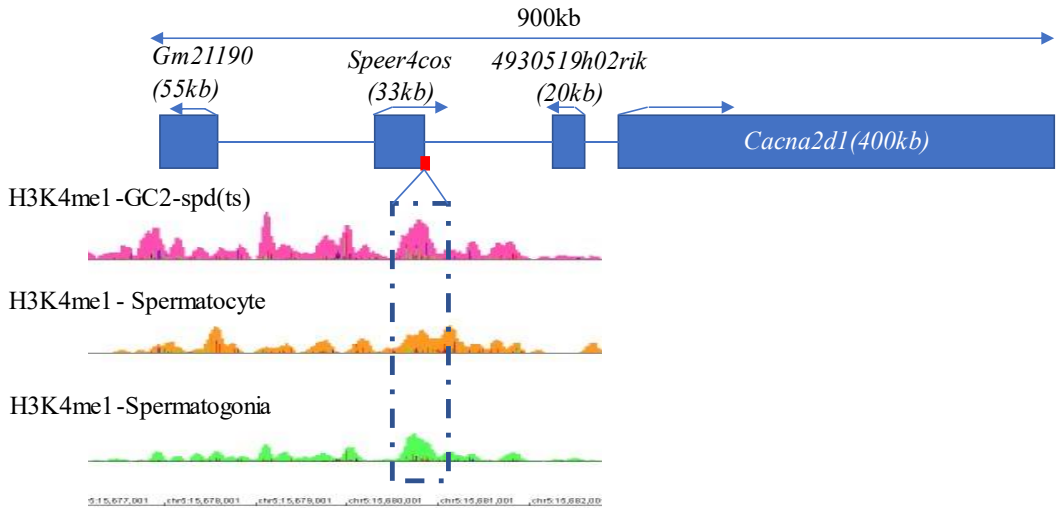

A
142 kb
Sp110 (27kb)
Sp140 (44kb)
Sp100 (52kb)
A630001G21Rik (19.5kb)
H3K4me1 -GC2-spd(ts)
H3K4me1 - Spermatocyte
H3K4me1 -Spermatogonia
B
Sp110
Sp110
Sp110
Sp110
Sp140
H3K4me1 -GC2-spd(ts)
C
Sp110
0.00016
0.00014
0.00012
0.0001
0.00008
0.00006
0.00004
0.00002
0
***
Controll
Delete
D
Sp140
0.004
0.0035
0.003
0.0025
0.002
0.0015
0.001
0.0005
0
Controll
Delete
E
A630001G21Rik
0.00014
0.00012
0.0001
0.00008
0.00006
0.00004
0.00002
0
*
Controll
Delete

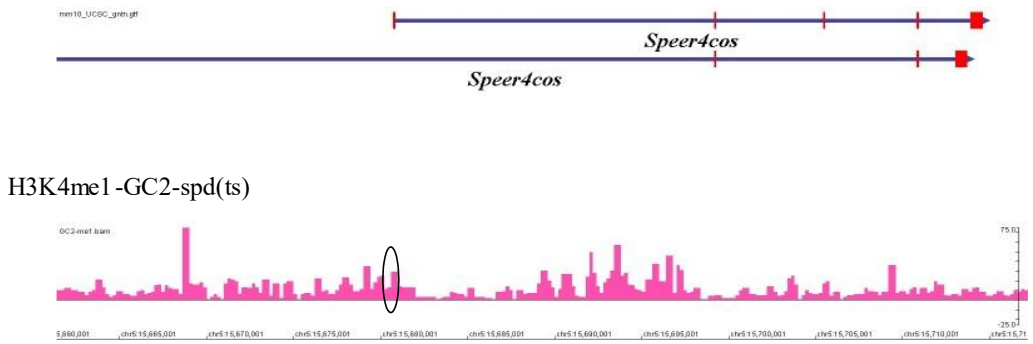
Fig. 2.04 Identification of a dual promoter activity of SP110-DPE in mouse spermatocytes/GC-2spd(ts) by genome editing. A) Three protein-coding genes which are closed to the SP110-DPE were indicated in blue boxes and the transcription direction of each gene is indicated in arrowhead depicted by blue boxes with the name and length. The position of SP110-DPE is shown by a red square at the TSS of the *Sp110* which is the intron region of the *Sp140* gene. ChIP-seq data for three cell types were shown below the SP110-DPE. In the blue dash line, the epigenetic mark of histone H3 lysine 4 monomethylating peak was indicated. B) Three transcript variants of *Sp110* are enlarged and drawn by indicating the exon-intron structure and red indicate the exons, and the parallel ChIP-seq data for H3K4Me1 modification patterns along the Rin2 gene region of GC-2spd(ts) was shown and around the representative peak for the SP110-DPE was circled. C) qRT-PCR for *Sp110* gene expression in GC-2spd(ts) cells edited by the CRISPR/Cas9 system. Total RNAs were isolated from the puromycin selected cells selected and qRT-PCR was performed for the *Sp110* gene. *Gapdh* was used as an internal control. Data are presented as means ± S.D. from eight independent experiments, and statistical significance was assessed by Student's t-test. $^{**}P < 0.001$. (E,) qRT-PCR for *Sp140* gene expression in GC-2spd(ts) cells edited by the CRISPR/Cas9 system. Total RNAs were isolated and qRT-PCR was performed as in (C). E) qRT-PCR for *Plin3* gene expression in GC-2spd(ts) cells edited by the CRISPR/Cas9 system. Total RNAs were isolated from the cells selected with puromycin and qRT-PCR was performed for the *A630001G21Rik* gene. *Gapdh* was used as an internal control. Data are presented as means ± S.D. from eight independent experiments, and statistical significance was assessed by Student's t-test. $^{*}P < 0.05$.

Fig. 2.05

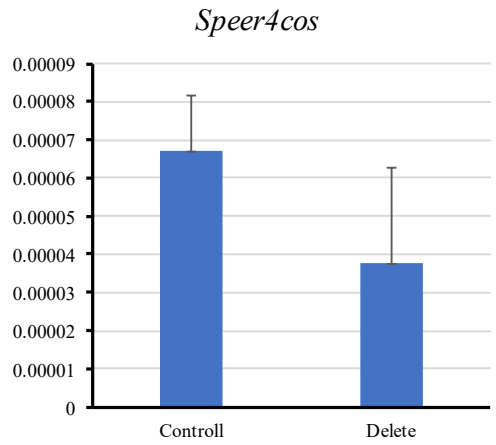
A


900kb
Gm21190
(55kb)
Speer4cos
(33kb)
4930519h02rik
(20kb)
Cacna2d1(400kb)
H3K4me1-GC2-spd(ts)
H3K4me1 - Spermatocyte
H3K4me1-Spermatogonia

B


Speer4cos
Speer4cos
H3K4me1-GC2-spd(ts)

C


Speer4cos
0.00009
0.00008
0.00007
0.00006
0.00005
0.00004
0.00003
0.00002
0.00001
0
Controll
Delete

D

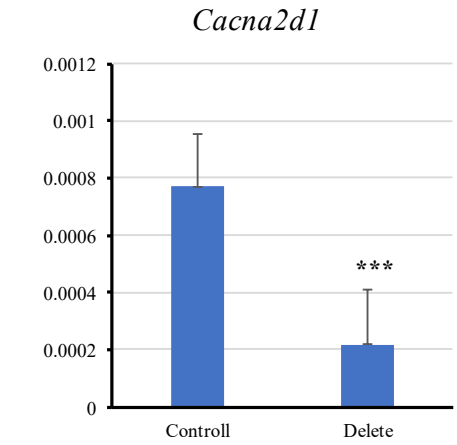

Cacna2d1
0.0012
0.001
0.0008
0.0006
0.0004
0.0002
0
***
Controll
Delete

Fig. 2.05 Identification of a dual promoter activity of SPEER4COS-DPE in mouse spermatocytes/GC-2spd(ts) by genome editing. A) Three protein-coding genes which are closed to the SPEER4COS-DPE were indicated in blue boxes and the transcription direction of each gene is indicated in arrowhead depicted by blue boxes with the name and length. The position SPEER4COS -DPE is shown by a red square at the TSS of the *Speer4cos*. ChIP-seq data for three cell types were shown below the SPEER4COS-DPE. In the blue dash line, the epigenetic mark of histone H3 lysine 4 monomethylating peak was indicated. B) Two transcript variants of the *Speer4cos* gene are enlarged and drawn by indicating the exon-intron structure and red indicate the exons, and the parallel ChIP-seq data for H3K4me1 modification patterns along the *Speer4cos* gene region of GC-2spd(ts) was shown and around the representative peak for the *SPEER4COS* -DPE was circled. C) qRT-PCR for *Speer4cos* gene expression in GC-2spd(ts) cells edited by the CRISPR/Cas9 system. Total RNAs were isolated from the puromycin selected cells selected and qRT-PCR was performed for the *Speer4cos* gene. *Gapdh* was used as an internal control. E) qRT-PCR for *Cacna2d1* gene expression in GC-2spd(ts) cells edited by the CRISPR/Cas9 system. Total RNAs were isolated from the cells selected with puromycin and qRT-PCR was performed for the *Cacna2d1* gene. *Gapdh* was used as an internal control. Data are presented as means ± S.D. from eight independent experiments, and statistical significance was assessed by Student's t-test. ***$P < 0.001$.

Chapter 3

A dual enhancer-silencer element, DES-K16, in mouse spermatocyte- derived GC-2spd(ts) cells

Abstract

The multifunctionality of the genome is suggested at some loci in different species but not well understood. Here I identified a DES-K16 region in an intron of the *Kctd16* gene as the chromatin highly marked with epigenetic modifications of both enhancers (H3K4me1 and H3K27ac) and silencers (H3K27me3) in mouse spermatocytes. *In vitro* reporter gene assay demonstrated that DES-K16 exhibited significant enhancer activity in spermatocyte-derived GC-2spd(ts) and hepatic tumor-derived Hepa1-6 cells, and deletion of this sequence in GC-2spd(ts) cells resulted in a decrease and increase of expression levels *Yipf5* and *Kctd16* genes, respectively. This was consistent with increased and decreased expression of *Yipf5* and *Kctd16*, respectively, in primary spermatocytes during testis development. While known dual enhancer-silencers exert each activity in different tissues, my data suggest that DES-K16 functions as both enhancer and silencer in a single cell type, GC-2spd(ts) cells. This is the first report on a dual enhancer- silencer element which activates and suppresses gene expression in a single cell type simultaneously.

Keywords: Dual enhancer-silencer, Multifunctional genome, Spermatocyte, Histone modification, GC-2spd(ts) cell, ChIP-sequencing

Introduction

The genome consists of different types of *cis*-regulatory elements such as transcriptional enhancers and silencers [87,99]. Transcriptional enhancers play a major role in activating gene expression and can be located upstream or downstream of transcriptional start sites of target genes or within gene bodies [87]. A silencer is a *cis*-regulatory element which negatively regulates the transcription of its target gene. Since its discovery in 1985, various silencers have been identified in different species [100–104], but compared to enhancers, detailed studies about silencers are limited.

Epigenomic analyses have revealed the association of epigenetic marks with *cis*-regulatory elements. H3K4me1 has been identified as an important and enriched mark at active and primed enhancers, and possibly finetunes, rather than tightly controls, the activity and function of enhancers [53,105,106]. H3K27ac distinguishes active enhancers from inactive/poised enhancer elements that contain H3K4me1 alone [51,105]. H3K27 trimethylation (H3K27me3) was recently reported to be a mark of silencers when it coincides with DNase I hypersensitivity [107]. Thus, epigenetic marks are useful for exploration of the functional genomic regions.

In addition, the multifunctionality of genome and its association with epigenetic marks have been revealed by recent studies. A dual promoter-enhancer (DPE) or epromoter is associated with H3K4me1, H3K27ac, and a typical mark of promoter, H3K4me3, and possesses both promoter and enhancer activity to control the transcription of two or more genes at a time [6,31,42,43,82,108–110]. A dual enhancer-silencer element possesses both enhancer and silencer activity [111,112], and a recent study with *Drosophila* indicated its

association with H3K4me1, H3K27ac, and H3K27me3 [113]. However, while dual promoter-enhancers mostly function as both enhancer and promoter in a single cell type, enhancer and silencer activity of a dual enhancer-silencer can be exerted in different tissues [5,111–113]. This raises a question whether a dual enhancer-silencer element can function as both enhancer and silencer in a single cell type or not.

I pursued the function of genome regions with both H3K4me1 and H3K4me3 identified in chapter 1 and conclude that promoter activity and enhancer activity will not exert concurrently even though these genomic regions are rich with common histone modification mark for enhancer and promoters. One such region, which exerts enhancer activity via reporter assay that originally named site 6 in an intron of the *Kctd16* gene, was renamed DES-K16 in this chapter. DES-K16 was marked also with H3K27ac and H3K27me3 and my data demonstrate that DES-K16 exhibits unprecedented dual enhancer-silencer activity in the spermatocyte-derived GC-2spd(ts) cell line. To the best of my knowledge, this is the first demonstration of the simultaneous enhancer and silencer activity in a single cell type.

Materials and Methods

Animals

C57Bl/6 mice were maintained on 14 hr light/10 hr dark cycles at 25℃ and given food and water Ad libitum. The experimental procedures in this study were approved by the Institutional Animal Use and Care Committee at Hokkaido University.

ChIP-sequencing data analysis

ChIP-sequencing (ChIP-seq) data were collected and analyzed as previously described [42].

Reporter Constructs

Selected candidate regions were amplified (Table 3.01) by PCR using KOD FX Neo (Toyobo) with genomic DNA isolated from the C57BL/6 mouse liver and the primer pair listed in Table 3.02. To show the promoter activity, site2, site4, and site6 (DES-K16) fragments were ligated at the *Sma*I site upstream of the luciferase gene in a pGL3-Basic vector respectively. All constructs were subjected to the DNA sequencing analysis.

The thymidine kinase (TK) promoter was obtained from a pEBMulti-Neo vector (Fujifilm Wako Pure Chemical Corporation, Osaka, Japan) by *Nhe*I and *Bgl*II restriction digestion, blunted by T4 DNA polymerase (Takara, Kusatsu, Japan), and inserted into a pGL3-Basic vector (Promega, Madison, WI, USA) at the *Sma*I site. This construct was named pGL-TK site2, site4, and site6 (DES-K16) respectively. All constructs were subjected to the DNA sequencing analysis.

Detection of the possible presence of transcripts by PCR

Primers were designed at both ends of the selected regions to investigate the possible presence of transcripts in the selected region of the mouse testis (Table 3.02). Synthesized cDNA samples from the mouse testis and GC-2spd(ts) cells were used as PCR templates. Genomic DNA from mouse testis was used as a control template for this experiment.

Cell culture, transfection, and reporter gene assay

GC-2spd(ts) cells and Hepa1-6 cells were cultured in Dulbecco's Modified Eagle Medium (Fujifilm Wako) containing 10% fetal bovine serum supplemented with 100 U/ml penicillin, 100 µg/ml streptomycin, and 292 µg/ml L-glutamine (Invitrogen, Carlsbad, CA, USA). The cells were transfected with GeneJuice transfection reagent (Merck, Darmstadt, Germany), and reporter gene assay was performed two days later, as previously described [42,62]. The pRL-CMV vector (Promega) was co-transfected to normalize the transfection efficiency.

Genome editing by the CRISPR/Cas9 system

Guide RNA (gRNA) sequence was designed using the CRISPRdirect web tool (https://crispr.dbcls.jp/) [114]. Complementary oligonucleotides (Table 3.03) were annealed and ligated into a *Bbs*I-digested pX330 vector [115], and the completed construct was confirmed by DNA sequencing. The constructs were transfected into GC-2spd(ts) cells together with pKO-Select-Puro (Lexicon Genetics, The Woodlands, TX, USA), and the transfected cells were selected with 1-µg/µl puromycin for 7 days. The surviving cells were used for DNA and RNA isolation. The DNA was used for PCR with a primer pair listed in Table 3.03 to see whether the DES-K16 sequence was deleted. The RNA was used for quantitative reverse transcription-polymerase chain reaction (qRT-PCR).

qRT-PCR analysis

Total RNA was isolated from the cells selected as above or mouse testes using ISOGEN II (Nippongene, Tokyo, Japan) according to the manufacturer's instructions. After treatment with TurboDNase (Thermo Fisher Scientific, Waltham, MA, USA), cDNAs were synthesized using the oligo(dT) primer and Superscript III (Thermo Fisher).

Quantitative PCR (qPCR) reactions were performed using Power SYBR Green Master Mix (Thermo Fisher) in the 7300 real-time PCR system (Applied Biosystems, Foster City, CA, USA). Primers used for this qPCR are listed in Table 3.03. Expression levels were normalized to that of the housekeeping *Gapdh* gene. For the *Kctd16* gene, a transcript variant, *Kctd16-202*, which encodes a full-length protein, was detected.

Statistical analysis

The results were expressed as means ± standard deviation (S.D.). Dunnett's test and Tukey HSD test were performed by the R software and Student's t-test was performed using Microsoft Excel statistical analysis functions (Microsoft Corp., Redmond, WA, USA). One-way analysis of variance (ANOVA) was done by F-test (Microsoft Corp.). $P < 0.05$ was considered statistically significant.

Results

First attempt to identify DPEs by ChIP-seq

Initially, I looked for the cell type-specific genome-wide regions that contain both H3K4me3 (an epigenetic marker for promoters) and H3K4me1 peaks (a marker for enhancers). I selected these regions close to the transcribed region identified by RNA-seq analysis as I assumed that promoter regions should be proximal to the transcribed regions (Fig. 3.01). Based on the RNA-seq profiles and ChIP-seq data profile, I selected three regions for further analysis. Originally, I named these regions as Site2, Site4, and Site6. These sites were located at different chromosomes and lengths of these sites were varied around 200 bp to 900 bp (Table 3.01).

I first assessed the transcriptional regulatory activity of Site2, Site4, and Site6 by *in vitro* reporter gene assay. To check the promoter activity, these regions were inserted upstream of the luciferase gene as promoters. To check the enhancer activity, as a general promoter, the TK promoter was placed prior to the luciferase gene, and Site2, Site4, and Site6 were conjugated with the downstream of the luciferase gene in both directions. Primarily I have used GC2-spd(ts) cells for reporter analysis which was derived from mouse primary spermatocytes.

The followings are detailed descriptions of the results of each site

Site2: This genomic region is an intron region of the *Tmem267* which is predicted to be a membrane transporting component orthologous to human TMEM267. The length of this region is around 980 bases which contain both H3K4me1 and H3K4me3 histone modification patterns in the spermatocyte, spermatogonia, and GC2-spd(ts) cells (Fig. 3.02). I performed the reporter gene analysis for this region and did not observe any promoter activity (Fig. 3.05A), or enhancer activity (Fig. 3.05B) even though this region is rich with common histone modification marks for enhancer and promoter.

Site4: This site is a noncoding genomic region which is around 660 bp and the closest coding gene was *270054A10Rik* and which was located 20 kb downstream of the selected Site4. ChIP-seq data analysis revealed that this region is rich with H3K4me1 and H3K4me3 histone modification patterns in Spermatogonia, GC-2spd(ts) cells as well as in spermatocytes (Fig. 3.03). I performed the reporter gene analysis to check the expected promoter (Fig. 3.05A), and enhancer activity (Fig. 3.05B), but I could not detect either of the activities.

Site6: This site is located within the intron region of the *Kctd16* gene and the length of the region was around 341 bp. This region is rich with H3K4me1 and H3K4me3 histone modification patterns in Spermatogonia, GC-2spd(ts) cells as well as in spermatocytes (Fig. 3.04). Though I attempted to identify the promoter and enhancer activity using reporter gene analysis I could not obtain any promoter activity for Site 6 (Fig. 1.03A, B). Interestingly I observed significant enhancer activity for the Site 6 region and named this region as DES-K16 prior to further analysis.

Attempt for detecting transcript upstream and downstream of the DES-K16 region

Further, I tried to detect transcripts generated adjacent to the DESK-16 sequences in the testis and GC2-spd(ts) cells to check possible promoter-like activity. Here I performed the PCR reaction as described in the materials and method section. Although I could not detect any evidence of the presence of RNAs transcribed from this region using the designed primers (Table 3.02) for downstream and upstream of the DESK-16 region (Fig. 3.06).

None of these sequences shows promoter activity. I observed lower luciferase activity compared to the empty pGL3 luciferase reporter vector which is used as control. This may be due to the inserted sequences affect the native level of expression of the pGL3 vector by interfering with the Kozak consensus sequence which originally helps to start the translation, or another possibility is that cryptic RNA splicing due to newly generated point mutation by the insertion of this region.

Enhancer activity of DES-K16 by in vitro reporter gene assay

In GC-2spd(ts) cells, the activity was dependent on the direction and significantly increased only by a forward direction of DES-K16 (Fig. 3.07A). Due to this, I have transfected the construct into Hepa1-6 as Hepa1-6 cells originated from a mouse hepatic tumor but are known to express testicular germ cell-specific genes [116]. As a result, I observed the luciferase activity was significantly increased in both cell lines when the DES-K16 sequence was present. On the other hand, both directions of DES-K16 significantly enhanced TK promoter activity in Hepa1-6 cells (Fig. 3.07B). These results indicated that DES-K16 possesses significant enhancer activity.

DES-K16 with H3K4me1, H3K27ac, and H3K27me3 marks in mouse spermatocytes

To find the function of DES-K16 during spermatogenesis, I analyzed publicly available ChIP-seq data indicating H3K4me1, H3K27ac, and H3K27me3 patterns (GEO accession numbers: GSM1202711, GSM1202714, and GSM1674011) in mouse primary spermatocytes [34,117]. I detected a high peak for all of the three histone marks overlapping with an intron region of the *Kctd16* gene on chromosome 18 (Fig. 3.08). This approximately 350-bp region contained a 100-bp high peak and 250-bp small but clear peaks. The high peak of H3K4me1 in primary spermatocytes was also observed in my ChIP-seq data. Due to a recent report indicating that a genome region with these three marks as a dual enhancer-silencer element in *Drosophila* [113], I further analyzed DES-K16 for its transcriptional activity.

Effect of a deletion of DES-K16 by the CRISPR/Cas9 system

Next, I deleted the DES-K16 sequence in GC-2spd(ts) cells by genome editing using the CRISPR/Cas9 system. I transfected the vectors containing gRNAs for 5’ and 3’

regions of DES-K16 and selected the successfully transfected cells (Fig. 3.09A). As a control, the vector containing no gRNAs was transfected. I isolated genome DNAs from the selected cells to examine whether the sequence was deleted. As shown in a result of genomic PCR, a band was detected at a lower position in the deleted cells than in the control cells (Fig. 3.09B), and DNA sequencing of the PCR product confirmed the deletion of a 358-bp sequence encompassing DES-K16 (Fig. 3.10).

Results of qRT-PCR demonstrated that the *Yipf5* gene, located at 40-kb upstream of the *Kctd16* gene in the antisense direction, was significantly downregulated by the deletion (Fig. 3.09C), revealing that the DES-K16 region function as an enhancer of *Yipf5* gene in GC-2spd(ts) cells. Interestingly, the qRT- PCR result also showed significant upregulation of the *Kctd16* gene in the DES-K16-deleted cells (Fig. 3.09D). The *Kctd16* gene was not expressed in the control GC-2spd(ts) cells (Fig. 3.09D), but the deletion led to an expression of this gene. These results suggested that the DES-K16 region at as a silencer of the *Kctd16* gene in GC-2spd(ts) cells. Additionally, expression of *Prelid2* and *Sh3rf2* genes, located downstream of the *Kctd16* gene, was not affected by the deletion (Fig. 3.11), and the *Pabpc2* gene, located upstream of DES-K16, was not expressed in either control or deleted cells. Collectively, these results indicated that DES-K16 specifically functions as a "dual enhancer-silencer": an enhancer of *Yipf5* and a silencer of *Kctd16* in GC2-spd(ts) cells.

Expression patterns of Yipf5 and Kctd16 genes during mouse testis development

To examine the possibility that DES-K16 is also a dual enhancer-silencer in mouse spermatocytes, I investigated the expression of *Yipf5* and *Kctd16* genes in mouse testes at various postnatal days. The *Yipf5* expression was gradually increased from 7 days

postpartum (dpp) to 28 dpp, which coincided with the onset and progression of meiosis (Fig.3.12A). This is in good agreement with the enhancer activity of DES-K16 in spermatocytes. In contrast, the *Kctd16* expression was significantly decreased from 7 dpp to 14 dpp and thereafter (Fig. 3.12 B), suggesting that this gene was suppressed in spermatocytes, in agreement with the silencer activity of DES-K16.

Discussion

In this study, I tried to identify the DPE candidate regions based on histone modification patterns in mouse testicular cells and assessed their transcriptional activity in cultured cells. As the first attempt, I identified the genomic regions marked with H3K4me1 and H3K4me3 using ChIP-seq data. H3K4me1 is a well-known histone modification pattern that is associated with active enhancers [105] and H3K4me3 being widely considered as an active mark for promoters. Therefore, our initial hypothesis was that the genomic regions marked with both of these epigenetic markers could possess both promoter and enhancer activity. However, among the initially selected three candidates, only one region exhibited enhancer activity. This result may attribute to the differences in a cellular environment, simultaneously suggesting that the existence of H3K4me1 and H3K4me3 markers is not mandatorily an indication of both promoter and enhancer activity.

Here, I identified the site6 later which was named as the DES-K16 region based on histone modification patterns in mouse spermatocytes and assessed its transcriptional activity in cultured cells. DES-K16 showed significant enhancer activity in two cell types by reporter gene assay, and the deletion of this region resulted in both an increase and decrease in expression of linked genes in GC-2spd(ts) cells. These data suggest that DES-

K16 functions as a dual enhancer-silencer in this cell type. Hitherto known dual enhancer-silencer elements act as either silencer or enhancer depending on the cell specificity or the cellular condition [111–113]. In contrast, the present study demonstrates that DES-K16 functions as both enhancer and silencer in one cell type, GC-2spd(ts) cells, and that DES-K16 is a novel type of multifunctional genome element.

Interestingly, in GC-2spd(ts) cells, DES-K16 enhanced the *Yipf5* gene which was upregulated in spermatocytes during testis development and silenced the *Kctd16* gene which was downregulated in spermatocytes (Fig. 3.12A and 3.12B). These results showed that the dual enhancer-silencer activity of DES-K16 is consistent with the expression patterns of its target genes in the mouse testis. Combined with the presence of epigenetic marks of both enhancer and silencer (H3K4me1, H3K27ac, and H3K27me3) in spermatocytes (Fig. 3.08), the present gene expression analyses verified that DES-K16 is presumed to function as a dual enhancer-silencer element not only in GC-2spd(ts) cells but also in spermatocytes during meiosis.

In general, enhancers and silencers are bound by distinct transcription factors for their activity [118] [119], and I indeed identified potential binding sites for both types of transcription factors in the DES-K16 sequence (Fig. 3.13). Moreover, a transcription factor such as GATA1 can regulate both enhancer and silencer functions [87,119], and DES-K16 contains such a potential binding sequence. For instance, SOX5 and YY1 are reported to be related to gene activation and silencing, respectively, and expressed in spermatocytes [120–122]. This suggests that the DES-K16 region is bound by such kinds of transcription factors, leading to incorporated enhancer and silencer activity in the same cellular environment.

The *Kctd16* gene was reported to interact with GABAb receptors and control the signaling pathway in the brain [123], and the *Yipf5* gene plays a major role in innate immune responses in mammalian cells [124]. Although no specific functions of these two genes have been identified in the testis, their expression changes during testis development suggest some roles in spermatogenesis. A significant decrement in the expression of *Kctd16* from 7 dpp to 14 dpp (Fig. 3.12A) raises the possibility that *Kctd16* has a role in spermatogonia, and the increase of *Yipf5* expression (Fig. 3.12B) implies its role during meiosis. Consequently, DES-K16 may contribute to the regulation of spermatogenesis through its multifunctional gene regulatory activity.

In conclusion, I identified a novel type of dual enhancer-silencer element, DES-K16, which functioned as a simultaneous enhancer and silencer in a single cell type, GC-2spd(ts) cells. DES-K16 was marked with H3K4me1, H3K27ac, and H3K27me3 in spermatocytes and enhanced *Yipf5* and silenced *Kctd16* genes, respectively, in GC-2spd(ts) cells. The present study paves the way for understanding the multifunctionality of genome and the gene regulation during spermatogenesis.

Table 3.01 Selected Three Candidate Regions (DPE) for Further Analysis

DPE candidate	Chromosome	Start	End	Length
Site2	chr13	119598198	119599177	980
Site4	chr17	13590933	13591595	663
Site6 (DES-K16)	chr18	40308013	40308353	341

Table 3.02 Oligonucleotides Used for Cloning

Site	Forward primer	Reverse Primer
Site2	CCTCGGTTACAGCATCTTTTG	GCAGCAGCAGTCACAATGAG
Site4	CAGGGGATGTCTGTGAGTGA	CATTTACTCACATTCAGCCC
Site6 (DES-K16)	TTAAACATTAAACATGAAGT	GGCCTGGTTCTCTGAACTGC

Table 3.03 Oligonucleotides Used to Amplify the Adjacent Region of the Site6 Using cDNA

	Forward primer	**Reverse Primer**
Site6	CCTCGGTTACAGCATCTTTTG	GCAGCAGCAGTCACAATGAG
Site6 Left Side	TGCAGGTGGTCAGTTAGCA	CTGCCTGGAACTTCCTTTGT
Site6 Right Side	AGGAAGGAAGGAAGGAAGGA	TAATCCCAGCCCTAGAGAGG

Table 3.03 Oligonucleotides Used in This Study

Designation	Forward 5'-3'	Reverse 3'-5'
[Guide RNA]		
DES-K16 Start	CACCAATAAAATAAAGCTTGTTGA	AAACTCAACAAGCTTTATTTTATT
DES-K16 End	CACCAGTCTGTGAACTATTGGAGA	AAACTCTCCAATAGTTCACAGACT
[Genotyping]		
P1	ACCTCTCTAGCAATCCCCGA	
P2		AGTCACACAAGCTGTGCAGT
[qRT-PCR]		
Yipf5	AGCCAACTCAGGCCTATCCT	GGGGTGTAGCACCGTTAGTG
Kctd16	CCGAGAGCACCAATCTCTTGA	AAACCACAGCCACACACTGA
Prelid2	CCGAAAGTATCCCAATCCAA	TCTCTGGAACCACATTCTGACA
Sh3rf2	CCCAGTGTCAGGATCCACAT	GGAGGATGACATCGCCTTTA
Gapdh	TGCACCACCAACTGCTTAGC	GGCATGGACTGTGGTCATGAG

Fig. 3.01

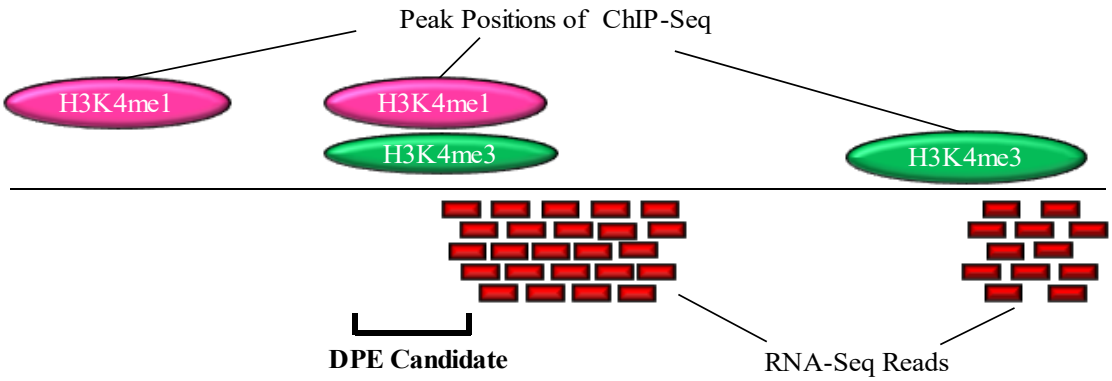


Fig. 3.01 Hypothesized model for identification of DPE as strategy 1-Transcribed region overlapped H3K4me1 and H3K4me3 methylation patterns were indicated in pink and green circles respectively and red boxes shows the RNA-seq reads. H3K4me1 and H3K4me3 close to transcribed regions peaks considered as the DPE

Fig. 3.02

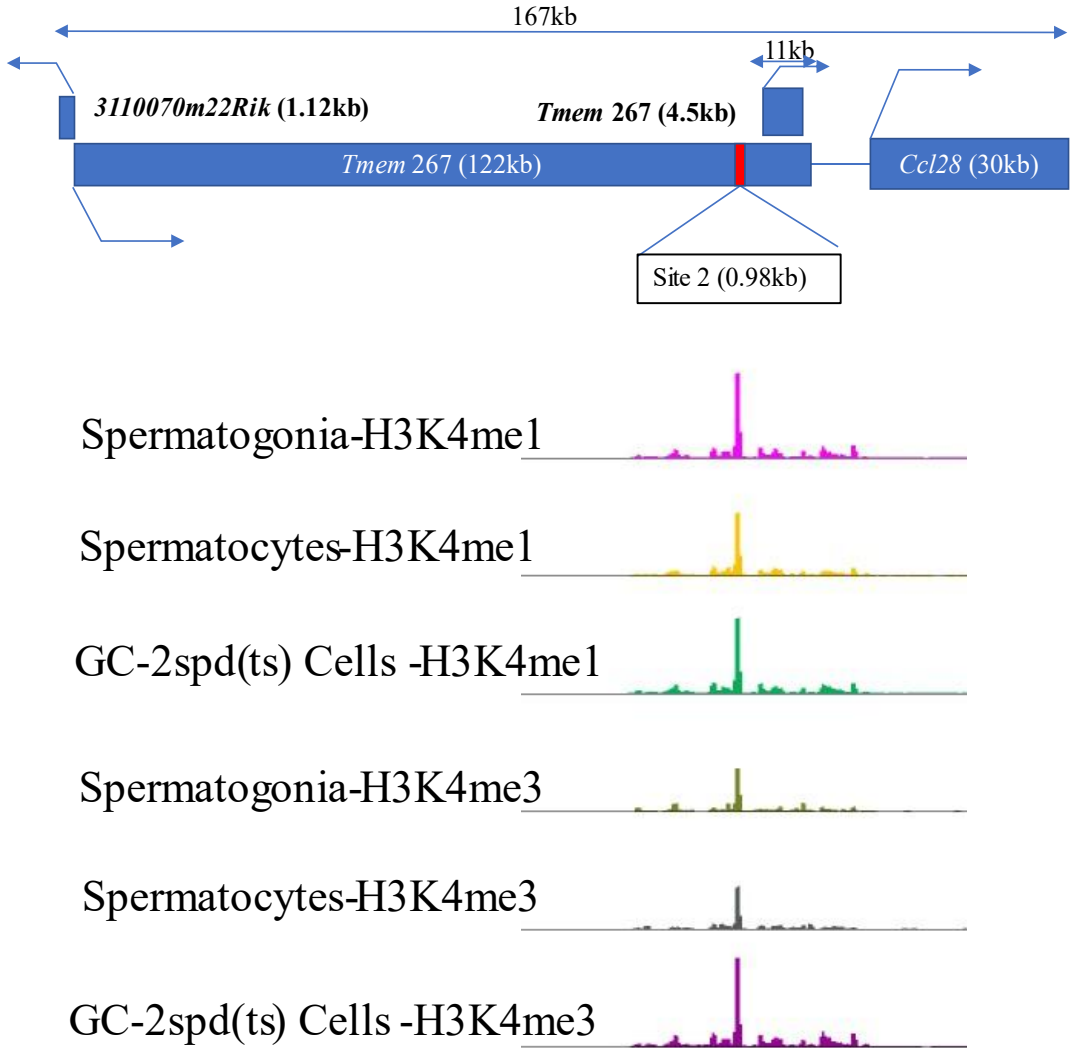


Fig. 3.02 Selected candidate Site 2 for analyzed the DPE activity according to strategy 1 Three protein-coding genes which are closed to the site 6 were indicated in blue boxes and transcription direction of each gene is indicated in arrow head depicted by blue boxes with the name and length. The position of Site 2 is shown by a red square which is intron region of *Tmem267* gene. ChIP-seq data for three cell types were shown below the Site 2.

Fig. 3.03

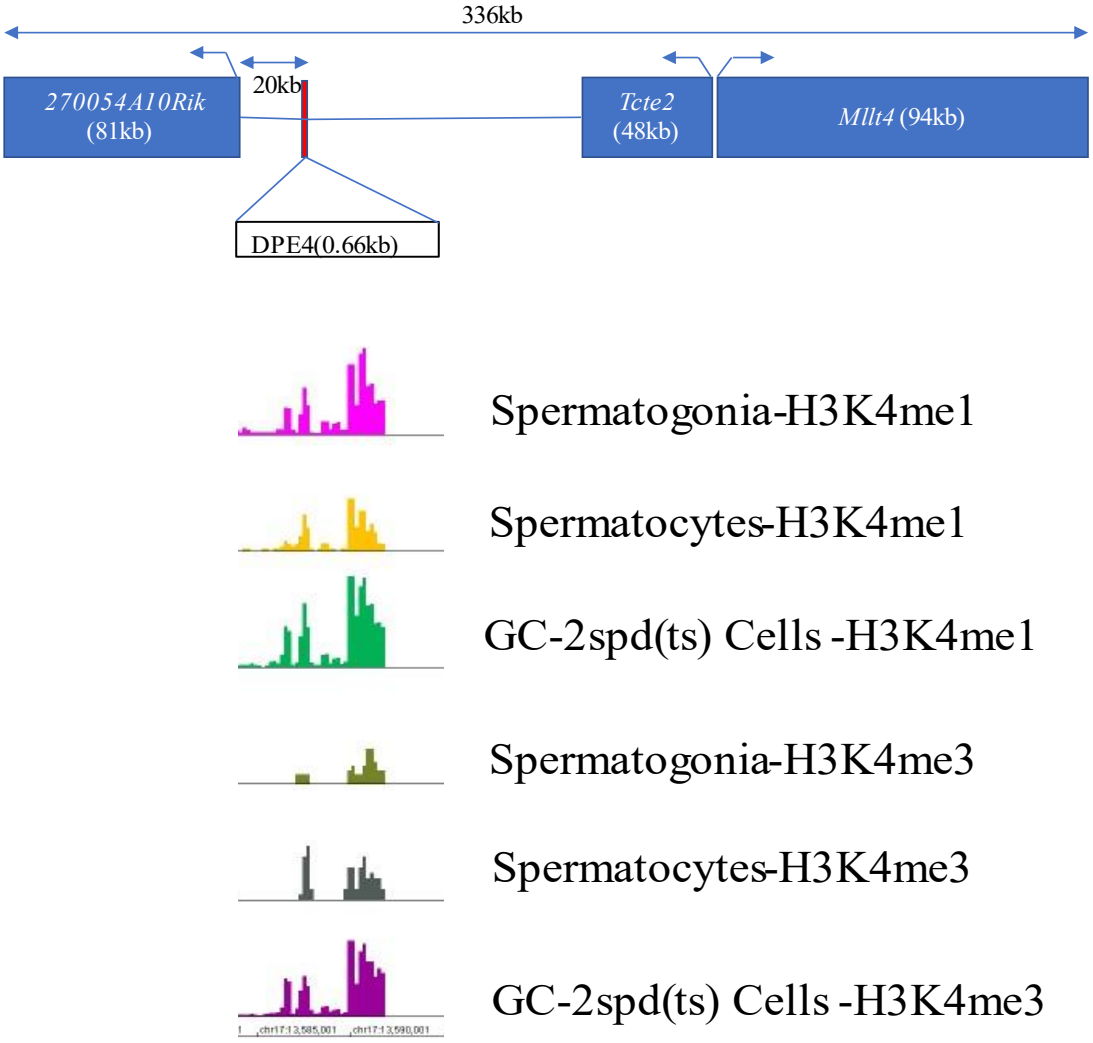


Fig. 3.03 Selected candidate site2 for analyzed the DPE activity according to strategy 1 Three protein-coding genes which are closed to the site 4 were indicated in blue boxes and transcription direction of each gene is indicated in arrowhead depicted by blue boxes with the name and length. The position of Site 4 is shown by a red square which is noncoding region. ChIP-seq data for three cell types were shown below the Site 4.

Fig. 3.04

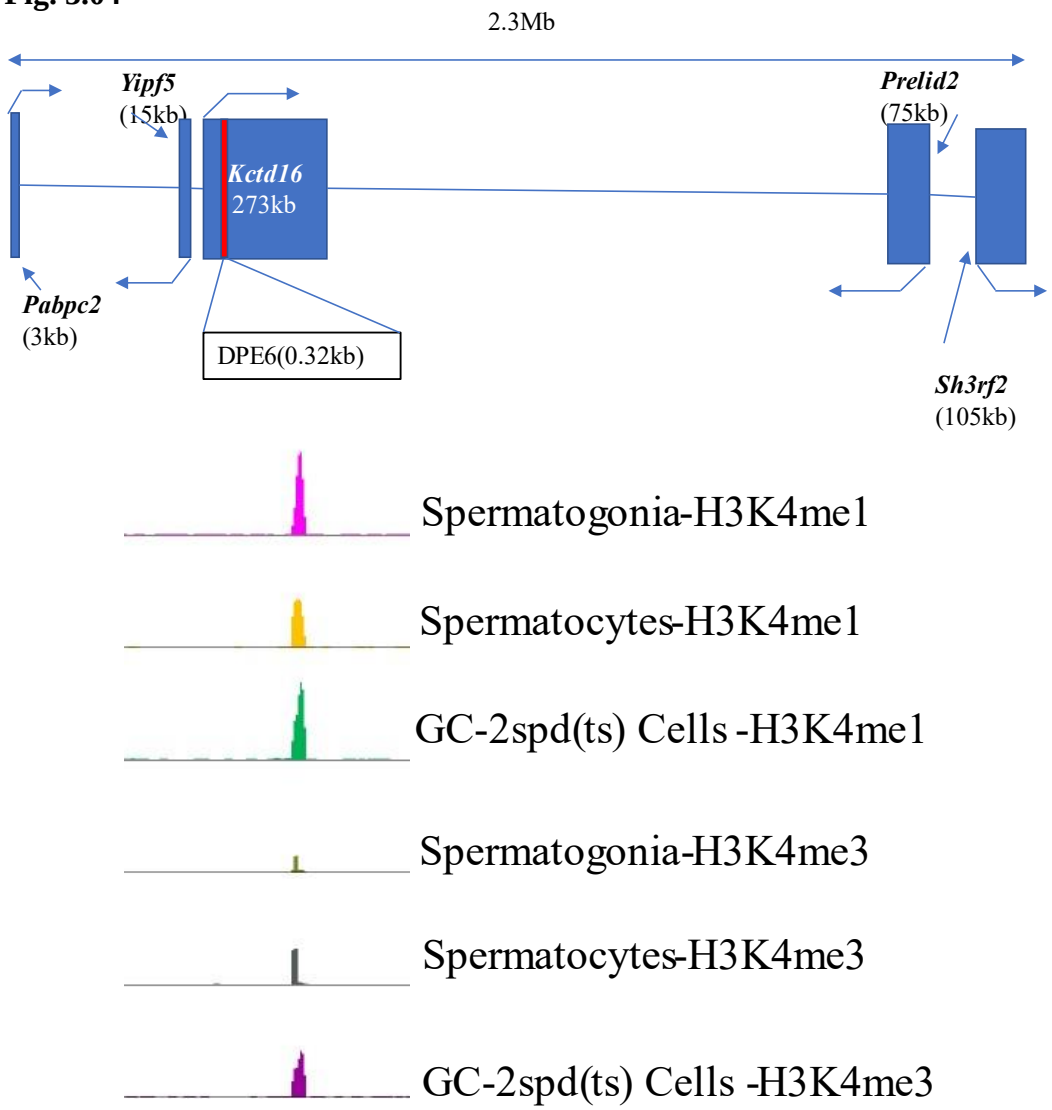


Fig3.04 Selected candidate Site6 for analyzed the DPE activity according to strategy 1 Four protein-coding genes which are closed to the site 6 were indicated in blue boxes and transcription direction of the each gene is indicated in arrow head depicted by blue boxes with the name and length. The position of Site 6 is shown by a red square which is intron region of *Kctd16* gene. ChIP-seq data for three cell types were shown below the Site 4.

Fig. 3.05

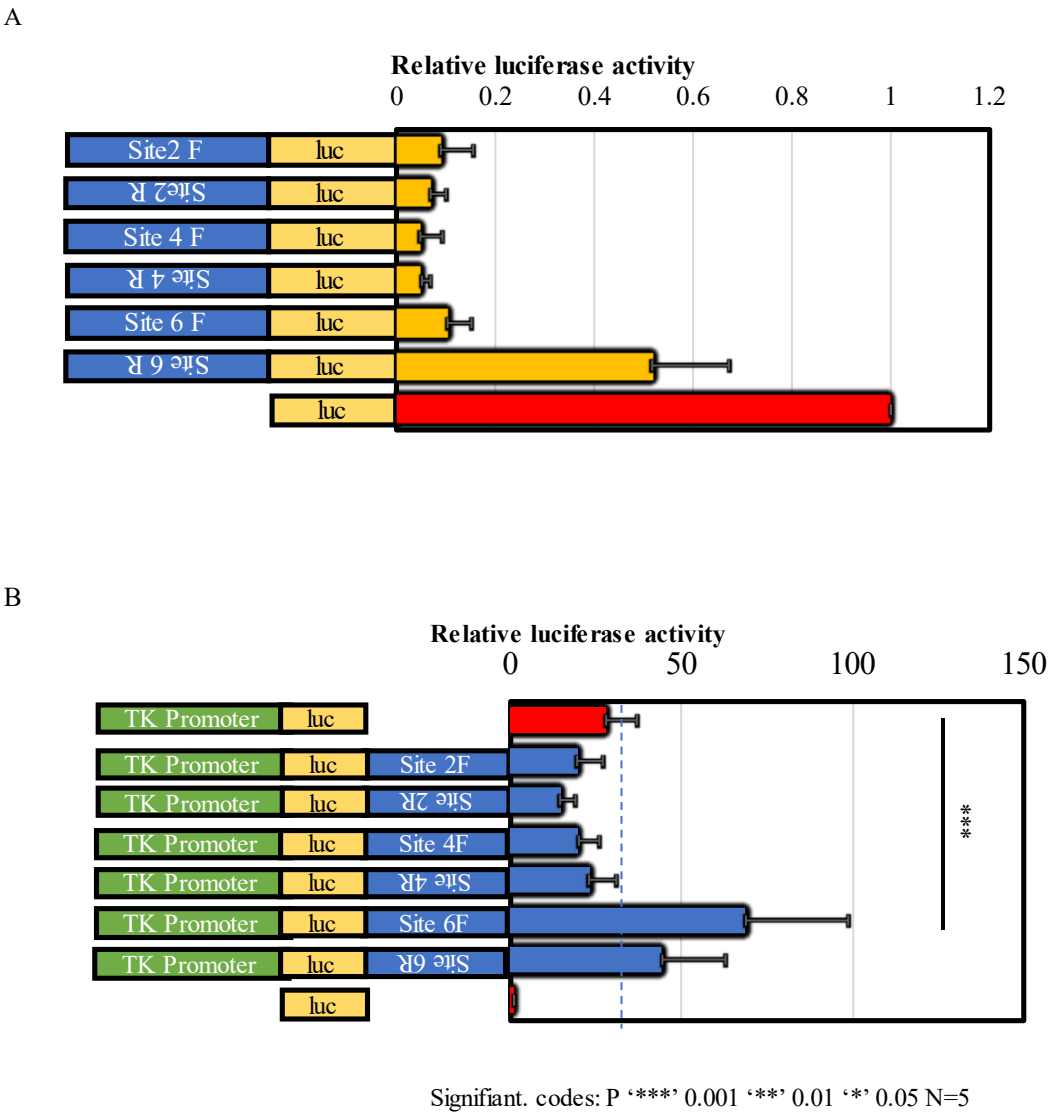


Fig. 3.05 *In vitro* reporter gene assay for transcriptional promoter and enhancer activity

The reporter gene constructs were generated as indicated at the left side of the graph. **(A)** Promoter construct **(B)** Enhancer construct. Each construct was transfected into GC-2spd(ts) cells and luciferase activity was measured two days later. The activity of the empty vector was set to 1.0 in each experiment, and the values from other constructs were calculated. Data are presented as means ± S.D. from five or six independent experiments, and statistical significance was assessed by Dunnett's test to compare the construct not containing DPE candidate with those containing it. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

Fig. 3.06

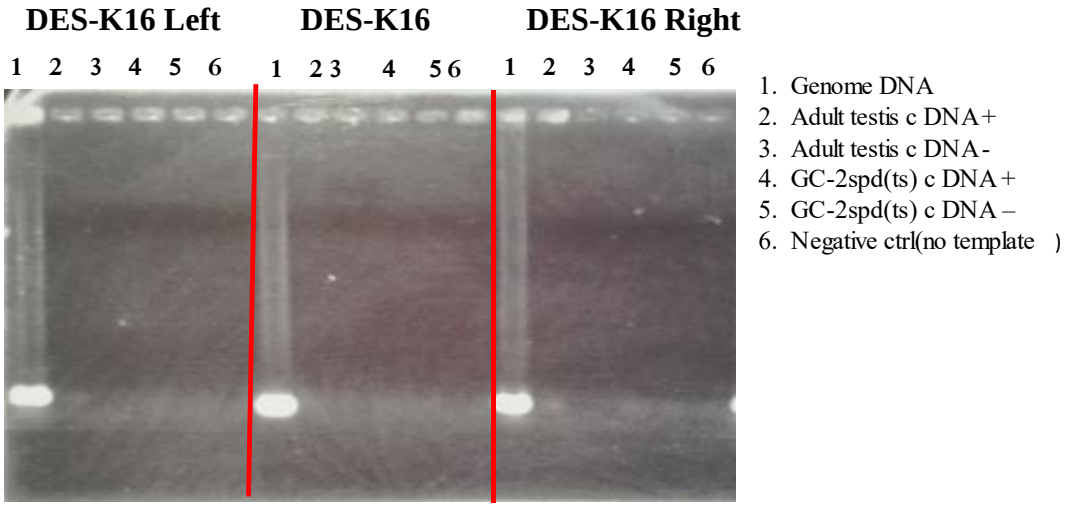


Fig 3.06 Detecting transcript upstream and downstream of the DES-K16 region

Detection of transcript adjacent to the DES-K16 region using cDNA sample as indicated.

Fig. 3.07

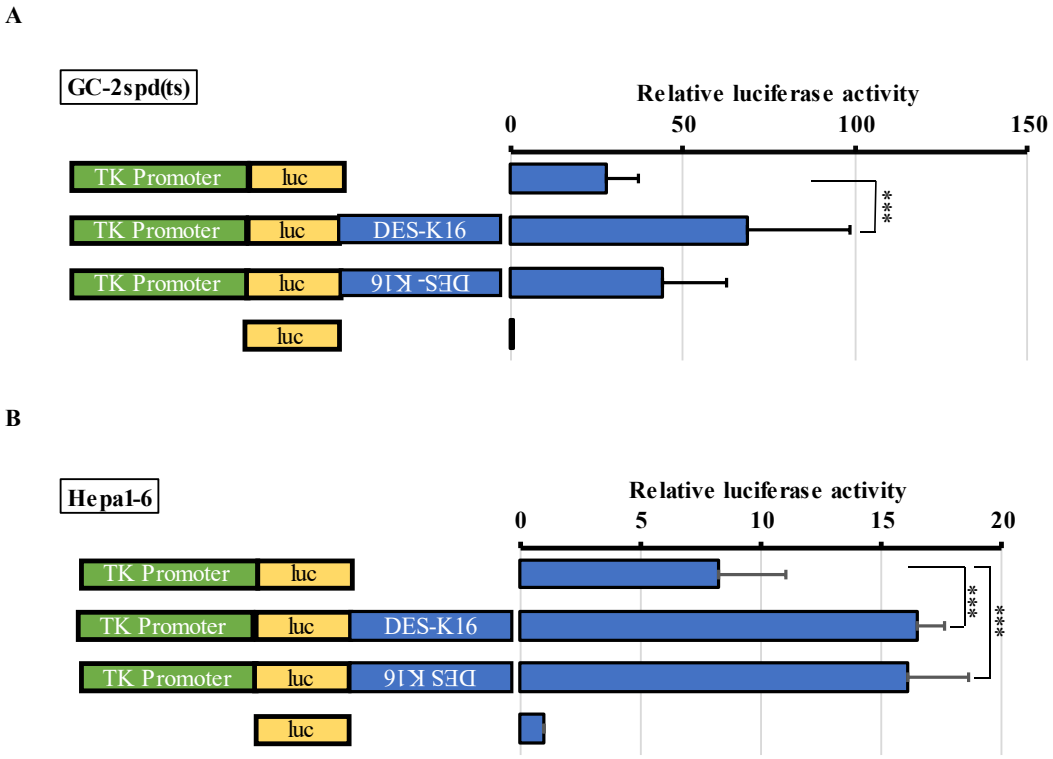


Fig. 3.07 *In vitro* reporter gene assay for transcriptional activity of DES-K16. The reporter gene constructs were generated as indicated at the left side of the graph. Each construct was transfected into GC-2spd(ts) (A) and Hepa1-6 cells (B), and luciferase activity was measured two days later. The activity of the empty vector was set to 1.0 in each experiment, and the values from other constructs were calculated. Data are presented as means ± S.D. from five or six independent experiments, and statistical significance was assessed by Dunnett's test to compare the construct not containing DES-K16 with those containing it. ***$P < 0.001$.

Fig. 3.08

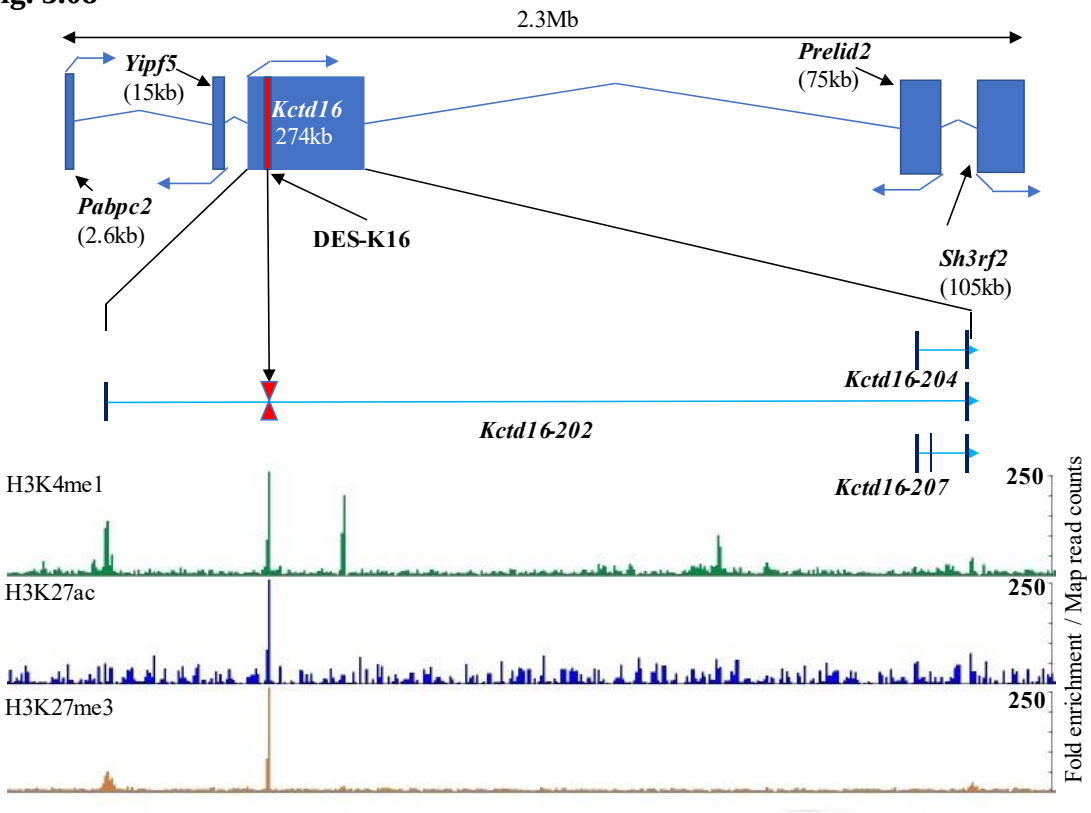


Fig. 3.08 Identification of a dual enhancer-silencer element, DES-K16, in mouse spermatocytes. Five protein-coding genes are depicted by blue boxes with the name and length. The position of DES-K16 is shown by a red line within the *Kctd16* gene. Below the gene structure, three transcript variants of *Kctd16* are enlarged and drawn by indicating the exon-intron structure, and the ChIP-seq data for three histone modification patterns in spermatocytes are also shown. The epigenetic marks are histone H3 lysine 4 monomethylation (H3K4me1), histone H3 lysine 27 acetylation (H3K27ac), and histone H3 lysine 27 trimethylation (H3K27me3). DES-K16 is a region containing a high peak of all of the three histone modifications.

Fig. 3.09

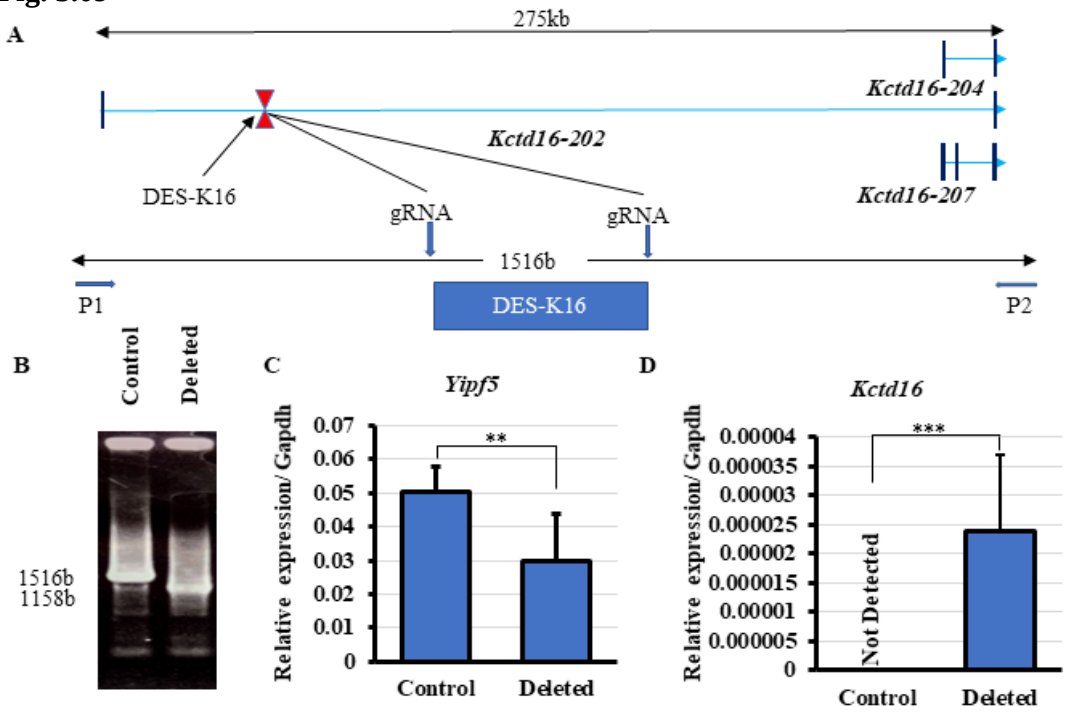


Fig. 3.09 Effect of a deletion of DES-K16 by genome editing in GC-2spd(ts) cells. (A) The exon-intron structure of *Kctd16* gene is indicated as in Fig. 1. The region amplified by PCR for genotyping is enlarged and the positions of primers (P1 and P2) and gRNA targets are shown. (B) A representative result of PCR for genotyping. GC-2spd(ts) cells were transfected with the CRISPR/Cas9 vectors containing and not containing gRNA sequences, and DNAs were isolated after selection. PCR was done with primers shown in (A), and the products were analyzed by electrophoresis and ethidium bromide staining. The band size confirmed by DNA sequencing is indicated. (C) qRT-PCR for *Yipf5* gene expression in GC-2spd(ts) cells edited by the CRISPR/Cas9 system. Total RNAs were isolated from the cells selected in (B) and qRT-PCR was performed for the *Yipf5* gene. *Gapdh* was used as an internal control. Data are presented as means ± S.D. from four to eight independent experiments, and statistical significance was assessed by Student's t-test. $^{**}P < 0.01$. (D) qRT-PCR for *Kctd16* gene expression in GC-2spd(ts) cells edited by the CRISPR/Cas9 system. Total RNAs were isolated and qRT-PCR was performed as in (C). In the control sample, *Kctd16* expression was not detected. Data are presented as means ± S.D. from eight independent experiments, and statistical significance was assessed by Student's t-test. $^{***}P < 0.001$.

Fig. 3.10

18 : 40307886 to 40308445

ATTTTGCTGAAAGGACTTAGATATATCTGTCTCTTGTGAGGCTATGCTGGGGCCTAGCAAACACAG
AAGTGGATGCTAACCGTCAGCTATTGGATGGATCACAGGGCCTTCAAGAGCTTTATTTTATTAAA
CATTAAACATGAAGTAGAAATTCCAAATCATGTAAATATTGTCCAGATTGGCTTTACAAATTCCTA
AACTGCTGGCAGGGCGAGGCTTCTAGTGAAAGTCTTCTAGTGAGTGGAGCCTTCCTCTTTCTGTTA
GAGTGTGTTTTCAACTCCATTACCATTTTCGAAGATCTAGGTATGAGCCTATTGTTTGGTTCAGTC
AGTTGCCAGAGCCAGAGCAAAGGTGGACCTGAAAGGACAGTGAGTTCAGAGAATTCACACTGAAAT
CATCACTCACGGACTTCTGGCAGAAAACTAGCCTCTGGATAGACATGGCTCAGCAGTTCAGAGAAC
CAGGCCATCTCCAATAGTTCACAGACTTGCAGGGACGACATGCAGTTCCGTGCCCCTCTCCTGGAC
ATTTGCAGTCCTTCTCTTTTCATGTTCCCTTC

Chr 18: 40308000 to 40308357
Yellow – reporter assayed region
Yellow & red – Deleted region

Fig. 3.10 The DES-K16 sequence and a deleted region. A genome sequence including DES-K16 is shown. A genomic region used for reporter gene analysis is marked with yellow, and the region deleted by the CRISPR/Cas9 system is the sequence marked with yellow and red.

Fig. 3.11

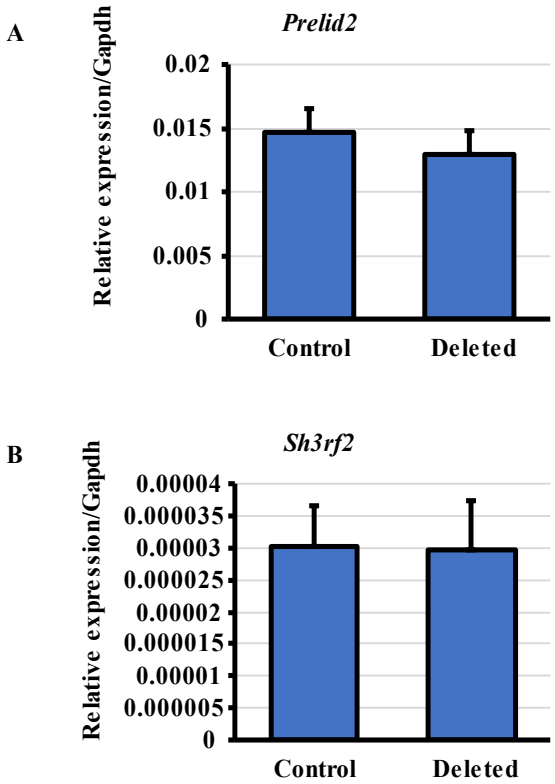


Fig. 3.11 Effect of a deletion of DES-K16 on expression of *Prelid2* and *Sh3rf2* genes in GC-2spd(ts) cells. qRT-PCR was performed with total RNAs from GC-2spd(ts) cells with or without a deletion of DES-K16. *Gapdh* was amplified as an internal control and used to normalize the expression level. Data are presented as means ± S.D. from six to eight independent experiments, and statistical significance was analyzed by Student t test. No significant differences were detected for expression levels of any of these genes.

Fig. 3.12

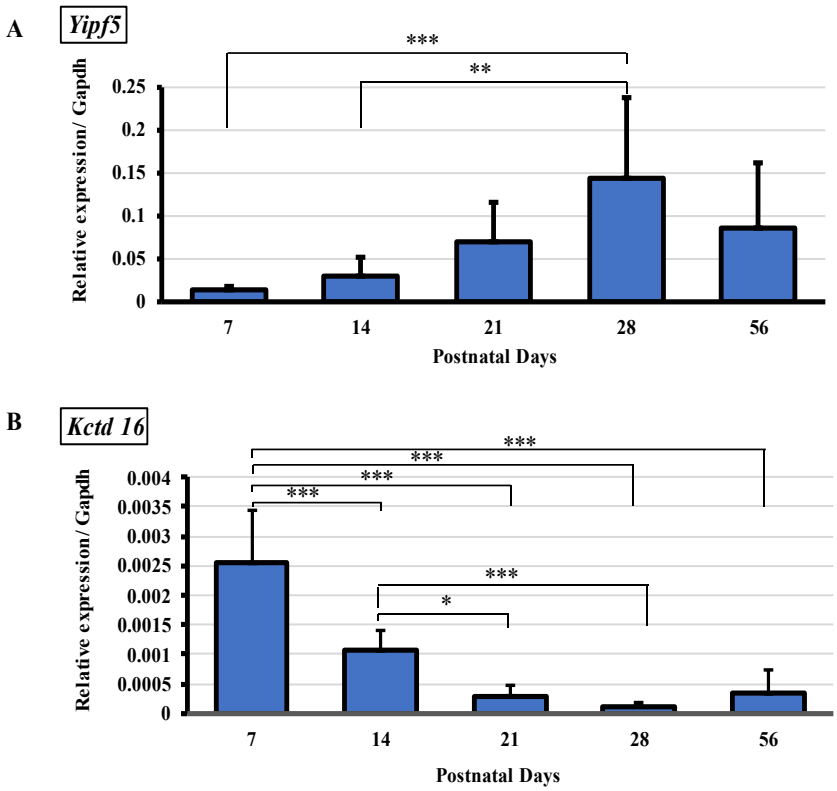


Fig. 3.12 *Yipf5* and *Kctd16* expression during mouse testis development. The expression pattern of *Kctd16* (A) and *Yipf5* (B) genes in postnatal testes by qRT-PCR. Total RNAs were isolated from whole testes at the indicated ages, and qRT-PCR was performed. *Gapdh* was amplified as an internal control and used to calculate the relative expression level. Data are presented as means ± S.D. from six (A) and eight (B) independent experiments, respectively, and statistical significance was analyzed by one-way ANOVA followed by the Tukey post hoc test. **P* < 0.05, ***P* < 0.01, ****P* < 0.001.

Fig. 3.13

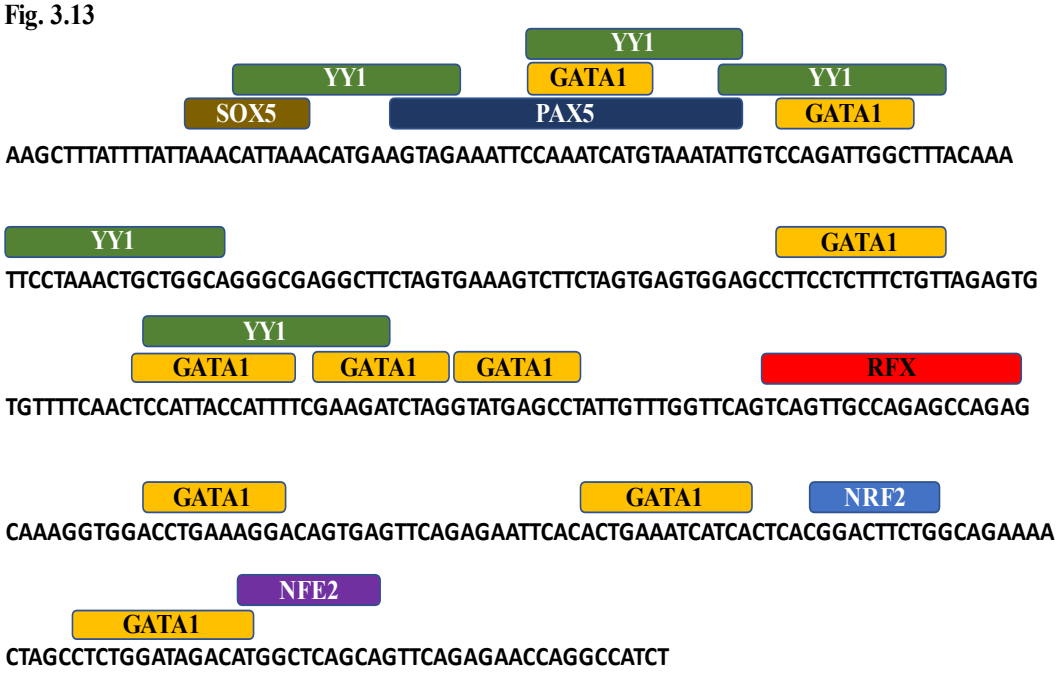


Fig. 3.13 Potential binding sequences of transcription factors related to enhancer and silencer in DES-K16. Potential transcription factor binding sites in DES-K16 were searched by the TRANSFAC software and are indicated in boxes at the upper side of each sequence. Purportedly, PAX5 and SOX5 are associated with enhancer and other factors are associated with silencer. GATA1 has a potential binding ability to both enhancer and silencer regions. All of these transcription factors are expressed in mouse spermatocytes.

General Discussion

General Discussion

The aim of this study was to investigate the multifunctional regulatory elements that regulate transcription during mammalian spermatogenesis. Using the high throughput sequencing technique, I was able to identify novel DPE regions and a novel type of dual enhancer-silencer element. My results suggest that various multifunctional genome elements play important roles in gene regulation during spermatogenesis and may be the key to the regulation of male reproduction.

Germ cell-specific genes were classically thought to be controlled solely by their proximal promoters, based on the identification of minimum promoters that were sufficient to activate genes in transgenic mice [125,126]. General transcription factors that are exclusively expressed in testicular germ cells are also known to play important roles in gene activation during meiosis [127,128]. Apart from that, there are findings which suggest many types of elements that control the gene regulation related to spermatogenesis. Some reports have identified lncRNA which acts as an enhancer for some testis-specific modulators in mammalian spermatogenesis [116,129]. In addition to these reported elements/factors, my current study is the findings of dual-acting elements that provide insight towards the mechanism of action during spermatogenesis [43].

So far, a limited number of multifunctional elements have been identified. The first two DPE elements were reported in chicken, and both drive lncRNAs that function in transcriptional activation of genes targeted by enhancer activity of the DPEs [6,59,108]. Our group identified the first mammalian DPEs in mouse and human, and the DPE was

evolutionarily conserved sequences that promote a ubiquitous gene and enhance a spermatocyte-specific gene[43,130]. Then, two groups attempted to identify DPEs or epromoters in a genome-wide fashion and reported that some hundreds of all promoters in the mouse genome show enhancer activity as well [7,31,82]. Apart from that, two independent studies performed CRISPR/Cas9-based screening of *cis*-regulatory elements and revealed that distal gene promoters control the expression of other genes [131,132]. Apart from that two different studies have shown that mouse long-non-coding RNAs whose promoter sequences were orthologous to putative human enhancers suggesting the possibility of acting as enhancers[133,134]. Another massive parallel reporter assay results related to mouse neuronal activity also shows the capability of the same sequence which can encode for both promoter and enhancer activity [82] and the recent report about architectural and functional commonalities between enhancers and promoters supports the idea of the existence of dual promoter enhancer regions among the species[135]. In addition, the presence of DPEs was also performed in *Drosophila* and the physiological significance has been discussed [109,136].

When considering dual enhancer-silencer elements, it was reported that the presence of such regions in human is related to brain nicotine receptors and L1 cell adhesion molecule gene expression during postnatal development [112]. Another finding provides evidence for a transcriptional silencer which also serves as a transcriptional enhancer in alternate cellular contexts in *Drosophila* [5]. All these findings support that DPEs and DESs play some roles in gene regulation in various biological events, which suggests that such multifunctional genome sequences can be vital regulatory elements in other tissues such as the testis. Given that many genes are simultaneously activated or silenced at each stage of spermatogenesis[34,137,138] common control mechanisms may exist during

spermatogenesis. My results showed a significant variation in the number of DPE elements before and after the onset of meiosis and discussed the regulatory actions of a dual enhancer-silencer element during spermatogenesis. These indicate the possibility of the key role of these multifunctional elements during spermatogenesis, especially in primary spermatocytes.

This simultaneous functionality by the same region gives rise to a question of whether the whole regulatory element is responsible for the dual function or different components of the element control either activity. A previous study reported the positive correlation between both activities by the same region and the requirement of distinct motifs for either type of regulation activity [82]. This suggests that even a multifunctional genome possesses both dual activities as a whole sequence, factors or elements related to each activity may be different. When considering the mechanism of action of DPE, most DPEs show the binding ability towards the variety of transcription factors. As an example, All the DPE candidate I assayed for the enhancer and promoter activity shows the binding ability towards the promoter-specific transcription factor GATA2 [4,139] and some of them have binding ability to the transcription factors like AP1, NRF and SOX5 that are commonly seen in the enhancer regions [82,106,118] (Table 1.07) Thus, DPEs may be controlled by slightly different mechanisms as they didn't show any specific binding ability towards the common transcription factor which controls the enhancer activity. Also, these factors are generally inducible transcription factors that are linked to the housekeeping and stress response. Therefore, DPEs may play an important role in regulating the housekeeping and stress-responsive genes during spermatogenesis [31].

In general, epigenetic regulation is tightly associated with gene regulatory activity in genome elements. Here, I have distinguished different types of dual-acting elements according to the distinct epigenetic modification patterns. DPEs are comprised of the H3K4me1 mark which is prominent at the enhancers. Interestingly, H3K27ac marks that are commonly used to distinguish poised and active enhancers were not necessarily a good mark for DPEs in testicular cells. When considering some of the assayed DPE candidates in this study shows both enhancer and promoter activity even without the presence of H3K27ac but all the candidates process H3K4me1 mark. These results suggest H3K4me1 mark is a useful histone modification pattern to distinguish the dual promoter enhancer region, not the H3K27ac. When considering the dual enhancer-silencer activity, I observed that this region is comprised of H3K4me1, H3K27ac, and H3K27me3 markers. This suggests that a different combination of epigenetic modifications in one genomic site can lead to creating a differently active genomic region. The actual relation of epigenetic modifications to the genome multifunctionality will require further investigation

Spermatogenesis is regulated by many meiotic stage-specific genes even though the mechanism of many individual processes is still not fully understood [116]. The function of these dual active elements may be particularly relevant to the rapid coordination of gene regulation and might be key transcription activation elements during spermatogenesis. Apart from that, identification of this type of dual-functional elements will help to uncover the mechanism which cannot be explained through traditional clear-cut promoter enhancer activity, and therefore, the present study paves the way for understanding the multifunctionality of the genome. Furthermore, it requires extensive study to understand the mechanism of action of these identified novel types of regulatory elements as it may

assist in overcoming the limited knowledge towards the mechanism of action of spermatogenesis.

Acknowledgment

First of all, I would like to express my sincere gratitude to Associate Professor Atsushi P Kimura for his consistent support and guidance throughout this project. Secondly, my deep gratitude goes to senior members of the academic staff including Prof. Yoshinao Katsu, Prof. Asato Kuroiwa, and Associate Prof. Katsueki Ogiwara for reviewing this thesis and their invaluable guidance and insight throughout this research and final examinations. My sincere thanks also go to Dr. Honoo Satake, Dr. Shin Matsubara, and Dr. Akira Shiraishi for their valuable support in Bioinformatics analysis.

I would also like to thank MEXT, the Japanese government's scholarship program, and Hokkaido University – (IGP) for letting me be a part of this global University and to achieve my academic goals. I am also thankful to the School of Life Sciences and all its staff members including the staff of OIAS for all the considerate guidance. I would also like to acknowledge and extend my gratitude to my former supervisor, Prof. Sanath Rajapakse for given recommendations, guidance, and encouragement.

My appreciation also extends to my laboratory colleagues. Dr. Yuki's guidance as the international student supporter made it easier for me to settle in and former members of the lab, Dr.Yui, Ms. Yuri, and Mr. Naoki for creating a positive atmosphere for me to blend in and carry out research. My heartfelt thanks also go to Mr. Kai Otsuka whose support and encouragement are especially valued in achieving my goal. As well as my thanks go to the rest of the junior members of the lab including Mr. Josei for all the given support during my research and altogether for the Kimura lab for making my highest study stay in Japan fruitful, enjoyable, and memorable.

To conclude, I cannot forget to thank my late mother and my wife, Menaka and all my friends for their unconditional support in this very intense academic period.

References

[1] M. Bustin, T. Misteli, Nongenetic functions of the genome, Science (80-.). 352 (2016). https://doi.org/10.1126/science.aad6933.

[2] R. Andersson, A. Sandelin, Determinants of enhancer and promoter activities of regulatory elements, Nat. Rev. Genet. 21 (2020) 71–87. https://doi.org/10.1038/s41576-019-0173-8.

[3] P. Carninci, A. Sandelin, B. Lenhard, S. Katayama, K. Shimokawa, J. Ponjavic, C.A.M. Semple, M.S. Taylor, P.G. Engström, M.C. Frith, A.R.R. Forrest, W.B. Alkema, S.L. Tan, C. Plessy, R. Kodzius, T. Ravasi, T. Kasukawa, S. Fukuda, M. Kanamori-Katayama, Y. Kitazume, H. Kawaji, C. Kai, M. Nakamura, H. Konno, K. Nakano, S. Mottagui-Tabar, P. Arner, A. Chesi, S. Gustincich, F. Persichetti, H. Suzuki, S.M. Grimmond, C.A. Wells, V. Orlando, C. Wahlestedt, E.T. Liu, M. Harbers, J. Kawai, V.B. Bajic, D.A. Hume, Y. Hayashizaki, Genome-wide analysis of mammalian promoter architecture and evolution, Nat. Genet. 38 (2006) 626–635. https://doi.org/10.1038/ng1789.

[4] R. Andersson, C. Gebhard, I. Miguel-Escalada, I. Hoof, J. Bornholdt, M. Boyd, Y. Chen, X. Zhao, C. Schmidl, T. Suzuki, E. Ntini, E. Arner, E. Valen, K. Li, L. Schwarzfischer, D. Glatz, J. Raithel, B. Lilje, N. Rapin, F.O. Bagger, M. Jørgensen, P.R. Andersen, N. Bertin, O. Rackham, A.M. Burroughs, J.K. Baillie, Y. Ishizu, Y. Shimizu, E. Furuhata, S. Maeda, Y. Negishi, C.J. Mungall, T.F. Meehan, T. Lassmann, M. Itoh, H. Kawaji, N. Kondo, J. Kawai, A. Lennartsson, C.O. Daub, P. Heutink, D.A. Hume, T.H. Jensen, H. Suzuki, Y. Hayashizaki, F. Müller, A.R.R. Forrest, P. Carninci, M. Rehli, A. Sandelin, An atlas of active enhancers across human cell types and tissues, Nature. 507 (2014) 455–461.

https://doi.org/10.1038/nature12787.

[5] M.S. Halfon, Silencers, Enhancers, and the Multifunctional Regulatory Genome, Trends Genet. 36 (2020) 149–151. https://doi.org/10.1016/j.tig.2019.12.005.

[6] P. Lefevre, J. Witham, C.E. Lacroix, P.N. Cockerill, C. Bonifer, The LPS-Induced Transcriptional Upregulation of the Chicken Lysozyme Locus Involves CTCF Eviction and Noncoding RNA Transcription, Mol. Cell. 32 (2008) 129–139. https://doi.org/10.1016/j.molcel.2008.07.023.

[7] L.T.M. Dao, S. Spicuglia, Transcriptional regulation by promoters with enhancer function, Transcription. 9 (2018) 307–314. https://doi.org/10.1080/21541264.2018.1486150.

[8] M. Aarabi, K.E. Christensen, D. Chan, D. Leclerc, M. Landry, L. Ly, R. Rozen, J. Trasler, Testicular MTHFR deficiency may explain sperm DNA hypomethylation associated with high dose folic acid supplementation, Hum. Mol. Genet. 27 (2018) 1123–1135. https://doi.org/10.1093/hmg/ddy021.

[9] M.D. Anway, A.S. Cupp, N. Uzumcu, M.K. Skinner, Toxicology: Epigenetic transgenerational actions of endocrine disruptors and male fertility, Science (80-.). 308 (2005) 1466–1469. https://doi.org/10.1126/science.1108190.

[10] Z.X. Li, X. Ma, Z.H. Wang, A differentially methylated region of the DAZ1 gene in spermatic and somatic cells, Asian J. Androl. 8 (2006) 61–67. https://doi.org/10.1111/j.1745-7262.2006.00069.x.

[11] M.S.C. Yan, C.C. Matouk, P.A. Marsden, Epigenetics of the vascular endothelium, J. Appl. Physiol. 109 (2010) 916–926. https://doi.org/10.1152/japplphysiol.00131.2010.

[12] M.A. Surani, 414122a0, 414 (2001).

[13] P.B. Talbert, S. Henikoff, Spreading of silent chromatin: Inaction at a distance, Nat. Rev. Genet. 7 (2006) 793–803. https://doi.org/10.1038/nrg1920.

[14] D. Hernandez-Cortes, I. Alvarado-Cruz, M.J. Solís-Heredia, B. Quintanilla-Vega, Epigenetic modulation of Nrf2 and Ogg1 gene expression in testicular germ cells by methyl parathion exposure, Toxicol. Appl. Pharmacol. 346 (2018) 19–27. https://doi.org/10.1016/j.taap.2018.03.010.

[15] S. V. Yi, M.A.D. Goodisman, Computational approaches for understanding the evolution of DNA methylation in animals, Epigenetics. 4 (2009) 551–556. https://doi.org/10.4161/epi.4.8.10345.

[16] Y. Liu, C. Lu, Y. Yang, Y. Fan, R. Yang, C.F. Liu, N. Korolev, L. Nordenskiöld, Influence of histone tails and H4 tail acetylations on nucleosome- nucleosome interactions, J. Mol. Biol. 414 (2011) 749–764. https://doi.org/10.1016/j.jmb.2011.10.031.

[17] L. Peng, E. Seto, Deacetylation of nonhistone proteins by HDACs and the implications in cancer, Handb. Exp. Pharmacol. 206 (2011) 39–56. https://doi.org/10.1007/978-3-642-21631-2_3.

[18] M. Werner, A.J. Ruthenburg, The United States of Histone Ubiquitylation and Methylation, Mol. Cell. 43 (2011) 5–7. https://doi.org/10.1016/j.molcel.2011.06.015.

[19] S.T. Grivna, B. Pyhtila, H. Lin, MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 13415–13420. https://doi.org/10.1073/pnas.0605506103.

[20] E. Beyret, N. Liu, H. Lin, PiRNA biogenesis during adult spermatogenesis in mice is independent of the ping-pong mechanism, Cell Res. 22 (2012) 1429–1439. https://doi.org/10.1038/cr.2012.120.

[21] F.P. Marchese, I. Raimondi, M. Huarte, The multidimensional mechanisms of long noncoding RNA function, Genome Biol. 18 (2017) 1–13. https://doi.org/10.1186/s13059-017-1348-2.

[22] A. Mukherjee, S. Koli, K.V.R. Reddy, Regulatory non-coding transcripts in spermatogenesis: Shedding light on "dark matter," Andrology. 2 (2014) 360–369. https://doi.org/10.1111/j.2047-2927.2014.00183.x.

[23] C.L.R. Barratt, L. Björndahl, C.J. De Jonge, D.J. Lamb, F.O. Martini, R. McLachlan, R.D. Oates, S. van der Poel, B.S. John, M. Sigman, R. Sokol, H. Tournaye, The diagnosis of male infertility: An analysis of the evidence to support the development of global WHO guidance-challenges and future research opportunities, Hum. Reprod. Update. 23 (2017) 660–680. https://doi.org/10.1093/humupd/dmx021.

[24] N. Alhathal, S. Maddirevula, S. Coskun, H. Alali, M. Assoum, T. Morris, H.A. Deek, S.A. Hamed, S. Alsuhaibani, A. Mirdawi, N. Ewida, M. Al-Qahtani, N. Ibrahim, F. Abdulwahab, W. Altaweel, M.J. Dasouki, A. Assiri, W. Qabbaj, F.S. Alkuraya, A genomics approach to male infertility, Genet. Med. 0 (2020). https://doi.org/10.1038/s41436-020-0916-0.

[25] S. Kishigami, N. Van Thuan, T. Hikichi, H. Ohta, S. Wakayama, E. Mizutani, T. Wakayama, Epigenetic abnormalities of the mouse paternal zygotic genome associated with microinsemination of round spermatids, Dev. Biol. 289 (2006) 195–205. https://doi.org/10.1016/j.ydbio.2005.10.026.

[26] J.M. Trasler, Epigenetics in spermatogenesis, Mol. Cell. Endocrinol. 306 (2009) 33–

36. https://doi.org/10.1016/j.mce.2008.12.018.

[27] C.C. Oakes, S. La Salle, D.J. Smiraglia, B. Robaire, J.M. Trasler, Developmental acquisition of genome-wide DNA methylation occurs prior to meiosis in male germ cells, Dev. Biol. 307 (2007) 368–379. https://doi.org/10.1016/j.ydbio.2007.05.002.

[28] S. Takashima, M. Takehashi, J. Lee, S. Chuma, M. Okano, K. Hata, I. Suetake, N. Nakatsuji, H. Miyoshi, S. Tajima, Y. Tanaka, S. Toyokuni, H. Sasaki, M. Kanatsu-Shinohara, T. Shinohara, Abnormal DNA methyltransferase expression in mouse germline stem cells results in spermatogenic defects, Biol. Reprod. 81 (2009) 155–164. https://doi.org/10.1095/biolreprod.108.074708.

[29] M. Godmann, R. Lambrot, S. Kimmins, The dynamic epigenetic program in male germ cells: Its role in spermatogenesis, testis cancer, and its response to the environment, Microsc. Res. Tech. 72 (2009) 603–619. https://doi.org/10.1002/jemt.20715.

[30] C.J. Marques, T. Francisco, S. Sousa, F. Carvalho, A. Barros, M. Sousa, Methylation defects of imprinted genes in human testicular spermatozoa, Fertil. Steril. 94 (2010) 585–594. https://doi.org/10.1016/j.fertnstert.2009.02.051.

[31] L.T.M. Dao, A.O. Galindo-Albarrán, J.A. Castro-Mondragon, C. Andrieu-Soler, A. Medina-Rivera, C. Souaid, G. Charbonnier, A. Griffon, L. Vanhille, T. Stephen, J. Alomairi, D. Martin, M. Torres, N. Fernandez, E. Soler, J. Van Helden, D. Puthier, S. Spicuglia, Genome-wide characterization of mammalian promoters with distal enhancer functions, Nat. Genet. 49 (2017) 1073–1081. https://doi.org/10.1038/ng.3884.

[32] A. Molaro, E. Hodges, F. Fang, Q. Song, W.R. McCombie, G.J. Hannon, A.D. Smith, Sperm methylation profiles reveal features of epigenetic inheritance and evolution

in primates, Cell. 146 (2011) 1029–1041. https://doi.org/10.1016/j.cell.2011.08.016.

[33] N. Song, D. Endo, T. Koji, Roles of epigenome in mammalian spermatogenesis, Reprod. Med. Biol. 13 (2014) 59–69. https://doi.org/10.1007/s12522-013-0167-8.

[34] S.S. Hammoud, D.H.P. Low, C. Yi, D.T. Carrell, E. Guccione, B.R. Cairns, Chromatin and transcription transitions of mammalian adult germline stem cells and spermatogenesis, Cell Stem Cell. 15 (2014) 239–253. https://doi.org/10.1016/j.stem.2014.04.006.

[35] X. Cui, X. Jing, X. Wu, M. Yan, Q. Li, Y. Shen, Z. Wang, DNA methylation in spermatogenesis and male infertility, Exp. Ther. Med. 12 (2016) 1973–1979. https://doi.org/10.3892/etm.2016.3569.

[36] T. Wang, H. Gao, W. Li, C. Liu, Essential Role of Histone Replacement and Modifications in Male Fertility, Front. Genet. 10 (2019) 1–15. https://doi.org/10.3389/fgene.2019.00962.

[37] J. Bao, K. Vitting-Seerup, J. Waage, C. Tang, Y. Ge, B.T. Porse, W. Yan, UPF2-Dependent Nonsense-Mediated mRNA Decay Pathway Is Essential for Spermatogenesis by Selectively Eliminating Longer 3'UTR Transcripts, PLoS Genet. 12 (2016) 1–24. https://doi.org/10.1371/journal.pgen.1005863.

[38] R. Oliva, Protamines and male infertility, Hum. Reprod. Update. 12 (2006) 417–435. https://doi.org/10.1093/humupd/dml009.

[39] J. Govin, C. Caron, C. Lestrat, S. Rousseaux, S. Khochbin, The role of histones in chromatin remodelling during mammalian spermiogenesis, Eur. J. Biochem. 271 (2004) 3459–3469. https://doi.org/10.1111/j.1432-1033.2004.04266.x.

[40] S. Güneş, T. Kulaç, Epigenetiğin spermatogenezdeki rolü, Turk Urol. Derg. 39

(2013) 181–187. https://doi.org/10.5152/tud.2013.037.

[41] P. Herzyk, Next-Generation Sequencing, Elsevier Inc., 2014. https://doi.org/10.1016/B978-0-12-386882-4.00008-6.

[42] M. Kurihara, A. Shiraishi, H. Satake, A.P. Kimura, A Conserved Noncoding Sequence Can Function as a Spermatocyte-Specific Enhancer and a Bidirectional Promoter for a Ubiquitously Expressed Gene and a Testis-Specific Long Noncoding RNA, J. Mol. Biol. 426 (2014) 3069–3093. https://doi.org/10.1016/j.jmb.2014.06.018.

[43] M. Kurihara, A.P. Kimura, Characterization of the human TCAM1P pseudogene and its activation by a potential dual promoter-enhancer: Comparison with a protein-coding mouse orthologue, FEBS Lett. 589 (2015) 540–547. https://doi.org/10.1016/j.febslet.2015.01.023.

[44] D.M. De Kretser, K.L. Loveland, A. Meinhardt, D. Simorangkir, N. Wreford, Spermatogenesis, 13 (1998).

[45] S. Sharma, A. Hanukoglu, I. Hanukoglu, Localization of epithelial sodium channel (ENaC) and CFTR in the germinal epithelium of the testis, Sertoli cells, and spermatozoa, J. Mol. Histol. 49 (2018) 195–208. https://doi.org/10.1007/s10735-018-9759-2.

[46] L. Rato, M.G. Alves, S. Socorro, A.I. Duarte, J.E. Cavaco, P.F. Oliveira, Metabolic regulation is important for spermatogenesis, Nat. Rev. Urol. 9 (2012) 330–338. https://doi.org/10.1038/nrurol.2012.77.

[47] S.A. Garcia-Moreno, M.P. Plebanek, B. Capel, Epigenetic regulation of male fate commitment from an initially bipotential system, Mol. Cell. Endocrinol. 468 (2018)

19–30. https://doi.org/10.1016/j.mce.2018.01.009.

[48] J.A. MacLean, M.F. Wilkinson, Gene Regulation in Spermatogenesis, Curr. Top. Dev. Biol. 71 (2005) 131–197. https://doi.org/10.1016/S0070-2153(05)71005-X.

[49] T.A.M.K. Bandara, K. Otsuka, S. Matsubara, A. Shiraishi, H. Satake, A.P. Kimura, A dual enhancer-silencer element, DES-K16, in mouse spermatocyte-derived GC-2spd(ts) cells, Biochem. Biophys. Res. Commun. (2020). https://doi.org/10.1016/j.bbrc.2020.10.049.

[50] M. Ehrlich, M. Lacey, DNA methylation and differentiation: Silencing, upregulation and modulation of gene expression, Epigenomics. 5 (2013) 553–568. https://doi.org/10.2217/epi.13.43.

[51] M.P. Creyghton, A.W. Cheng, G.G. Welstead, T. Kooistra, B.W. Carey, E.J. Steine, J. Hanna, M.A. Lodato, G.M. Frampton, P.A. Sharp, L.A. Boyer, R.A. Young, R. Jaenisch, Histone H3K27ac separates active from poised enhancers and predicts developmental state, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 21931–21936. https://doi.org/10.1073/pnas.1016071107.

[52] A. Manuscript, S.P.C. Nanostructures, Modification of enhancer chromatin: what, how and why? Eliezer, Nano. 6 (2008) 2166–2171. https://doi.org/10.1016/j.molcel.2013.01.038.Modification.

[53] A. Local, H. Huang, C.P. Albuquerque, N. Singh, A.Y. Lee, W. Wang, C. Wang, J.E. Hsia, A.K. Shiau, K. Ge, K.D. Corbett, D. Wang, H. Zhou, B. Ren, Identification of H3K4me1-associated proteins at mammalian enhancers, Nat. Genet. 50 (2018) 73–82. https://doi.org/10.1038/s41588-017-0015-6.

[54] S. Maezawa, A. Sakashita, M. Yukawa, X. Chen, K. Takahashi, K.G. Alavattam, I.

Nakata, M.T. Weirauch, A. Barski, S.H. Namekawa, Super-enhancer switching drives a burst in gene expression at the mitosis-to-meiosis transition, Nat. Struct. Mol. Biol. 27 (2020) 978–988. https://doi.org/10.1038/s41594-020-0488-3.

[55] K.G. Alavattam, S. Maezawa, A. Sakashita, H. Khoury, A. Barski, N. Kaplan, S.H. Namekawa, Attenuated chromatin compartmentalization in meiosis and its maturation in sperm development, Nat. Struct. Mol. Biol. 26 (2019) 175–184. https://doi.org/10.1038/s41594-019-0189-y.

[56] J. Dong, X. Wang, C. Cao, Y. Wen, A. Sakashita, S. Chen, J. Zhang, Y. Zhang, L. Zhou, M. Luo, M. Liu, A. Liao, S.H. Namekawa, S. Yuan, UHRF1 suppresses retrotransposons and cooperates with PRMT5 and PIWI proteins in male germ cells, Nat. Commun. 10 (2019) 1–14. https://doi.org/10.1038/s41467-019-12455-4.

[57] M.S. Kowalczyk, J.R. Hughes, D. Garrick, M.D. Lynch, J.A. Sharpe, J.A. Sloane-Stanley, S.J. McGowan, M. De Gobbi, M. Hosseini, D. Vernimmen, J.M. Brown, N.E. Gray, L. Collavin, R.J. Gibbons, J. Flint, S. Taylor, V.J. Buckle, T.A. Milne, W.G. Wood, D.R. Higgs, Intragenic Enhancers Act as Alternative Promoters, Mol. Cell. 45 (2012) 447–458. https://doi.org/10.1016/j.molcel.2011.12.021.

[58] G. Li, X. Ruan, R.K. Auerbach, K.S. Sandhu, M. Zheng, P. Wang, H.M. Poh, Y. Goh, J. Lim, J. Zhang, H.S. Sim, S.Q. Peh, F.H. Mulawadi, C.T. Ong, Y.L. Orlov, S. Hong, Z. Zhang, S. Landt, D. Raha, G. Euskirchen, C.L. Wei, W. Ge, H. Wang, C. Davis, K.I. Fisher-Aylor, A. Mortazavi, M. Gerstein, T. Gingeras, B. Wold, Y. Sun, M.J. Fullwood, E. Cheung, E. Liu, W.K. Sung, M. Snyder, Y. Ruan, Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation, Cell. 148 (2012) 84–98. https://doi.org/10.1016/j.cell.2011.12.014.

[59] J. Witham, L. Ouboussad, P.F. Lefevre, A NF-κB-Dependent Dual Promoter-Enhancer Initiates the Lipopolysaccharide-Mediated Transcriptional Activation of the Chicken Lysozyme in Macrophages, PLoS One. 8 (2013). https://doi.org/10.1371/journal.pone.0059389.

[60] R. Yoneda, T. Takahashi, H. Matsui, N. Takano, Y. Hasebe, K. Ogiwara, A.P. Kimura, Three testis-specific paralogous serine proteases play different roles in murine spermatogenesis and are involved in germ cell survival during meiosis, Biol. Reprod. 88 (2013) 1–14. https://doi.org/10.1095/biolreprod.112.106328.

[61] M. Kurihara, K. Otsuka, S. Matsubara, A. Shiraishi, H. Satake, A.P. Kimura, A testis-specific long non-coding RNA, lncRNA-Tcam1, regulates immune-related genes in mouse male germ cells, Front. Endocrinol. (Lausanne). 8 (2017) 1–12. https://doi.org/10.3389/fendo.2017.00299.

[62] S. Matsubara, T. Takahashi, A.P. Kimura, Epigenetic patterns at the mouse prolyl oligopeptidase gene locus suggest the CpG island in the gene body to be a novel regulator for gene expression, Gene. 465 (2010) 17–29. https://doi.org/10.1016/j.gene.2010.06.006.

[63] H. Kimura, Y. Hayashi-Takanaka, Y. Goto, N. Takizawa, N. Nozaki, The organization of histone H3 modifications as revealed by a panel of specific monoclonal antibodies, Cell Struct. Funct. 33 (2008) 61–73. https://doi.org/10.1247/csf.07035.

[64] Z. Wang, X. Xu, J.L. Li, C. Palmer, D. Maric, J. Dean, Sertoli cell-only phenotype and scRNA-seq define PRAMEF12 as a factor essential for spermatogenesis in mice, Nat. Commun. 10 (2019). https://doi.org/10.1038/s41467-019-13193-3.

[65] S. Suzuki, V.D. Diaz, B.P. Hermann, What has single-cell RNA-seq taught us about

mammalian spermatogenesis?, Biol. Reprod. 101 (2019) 617–634. https://doi.org/10.1093/biolre/ioz088.

[66] R. Yoneda, A.P. Kimura, A testis-specific serine protease, Prss41/Tessp-1, is necessary for the progression of meiosis during murine in vitro spermatogenesis, Biochem. Biophys. Res. Commun. 441 (2013) 120–125. https://doi.org/10.1016/j.bbrc.2013.10.028.

[67] K. Ohyama, M. Ohta, Y.Z. Hosaka, Y. Tanabe, T. Ohyama, Y. Yamano, Expression of anti-Müllerian hormone and its type II receptor in germ cells of maturing rat testis, Endocr. J. 62 (2015) 997–1006. https://doi.org/10.1507/endocrj.EJ15-0370.

[68] R.J. Vaghjiani, S. Talma, C.L. Murphy, Six-transmembrane epithelial antigen of the prostate (STEAP1 and STEAP2)-Differentially expressed by murine and human mesenchymal stem cells, Tissue Eng. - Part A. 15 (2009) 2073–2083. https://doi.org/10.1089/ten.tea.2008.0519.

[69] U.H. Saxena, L. Owens, J.R. Graham, G.M. Cooper, U. Hansen, Prolyl isomerase Pin1 regulates transcription factor LSF (TFCP2) by facilitating dephosphorylation at two serine-proline motifs, J. Biol. Chem. 285 (2010) 31139–31147. https://doi.org/10.1074/jbc.M109.078808.

[70] J. Zhou, Y.R. Du, W.H. Qin, Y.G. Hu, Y.N. Huang, L. Bao, D. Han, A. Mansouri, G.L. Xu, RIM-BP3 is a manchette-associated protein essential for spermiogenesis, Development. 136 (2009) 373–382. https://doi.org/10.1242/dev.030858.

[71] H. Okuda, K. DeBoer, A.E. O'Connor, D.J. Merriner, D. Jamsai, M.K. O'Bryan, LRGUK1 is part of a multiprotein complex required for manchette function and male fertility, FASEB J. 31 (2017) 1141–1151. https://doi.org/10.1096/fj.201600909R.

[72] F. Dong, K. Shinohara, Y. Botilde, R. Nabeshima, Y. Asai, A. Fukumoto, T. Hasegawa, M. Matsuo, H. Takeda, H. Shiratori, T. Nakamura, H. Hamada, Pih1d3 is required for cytoplasmic preassembly of axonemal dynein in mouse sperm, J. Cell Biol. 204 (2014) 203–213. https://doi.org/10.1083/jcb.201304076.

[73] R.J. Holcomb, S. Oura, K. Nozawa, K. Kent, Z. Yu, M.J. Robertson, C. Coarfa, M.M. Matzuk, M. Ikawa, T.X. Garcia, The testis-specific serine proteases PRSS44, PRSS46, and PRSS54 are dispensable for male mouse fertility, Biol. Reprod. 102 (2020) 84–91. https://doi.org/10.1093/biolre/ioz158.

[74] F. Yue, Y. Cheng, A. Breschi, J. Vierstra, W. Wu, T. Ryba, R. Sandstrom, Z. Ma, C. Davis, B.D. Pope, Y. Shen, D.D. Pervouchine, S. Djebali, R.E. Thurman, R. Kaul, E. Rynes, A. Kirilusha, G.K. Marinov, B.A. Williams, D. Trout, H. Amrhein, K. Fisher-Aylor, I. Antoshechkin, G. DeSalvo, L.H. See, M. Fastuca, J. Drenkow, C. Zaleski, A. Dobin, P. Prieto, J. Lagarde, G. Bussotti, A. Tanzer, O. Denas, K. Li, M.A. Bender, M. Zhang, R. Byron, M.T. Groudine, D. McCleary, L. Pham, Z. Ye, S. Kuan, L. Edsall, Y.C. Wu, M.D. Rasmussen, M.S. Bansal, M. Kellis, C.A. Keller, C.S. Morrissey, T. Mishra, D. Jain, N. Dogan, R.S. Harris, P. Cayting, T. Kawli, A.P. Boyle, G. Euskirchen, A. Kundaje, S. Lin, Y. Lin, C. Jansen, V.S. Malladi, M.S. Cline, D.T. Erickson, V.M. Kirkup, K. Learned, C.A. Sloan, K.R. Rosenbloom, B.L. De Sousa, K. Beal, M. Pignatelli, P. Flicek, J. Lian, T. Kahveci, D. Lee, W.J. Kent, M.R. Santos, J. Herrero, C. Notredame, A. Johnson, S. Vong, K. Lee, D. Bates, F. Neri, M. Diegel, T. Canfield, P.J. Sabo, M.S. Wilken, T.A. Reh, E. Giste, A. Shafer, T. Kutyavin, E. Haugen, D. Dunn, A.P. Reynolds, S. Neph, R. Humbert, R.S. Hansen, M. De Bruijn, L. Selleri, A. Rudensky, S. Josefowicz, R. Samstein, E.E. Eichler, S.H. Orkin, D. Levasseur, T. Papayannopoulou, K.H. Chang, A. Skoultchi, S. Gosh, C. Disteche, P. Treuting, Y. Wang, M.J. Weiss, G.A. Blobel, X. Cao, S. Zhong, T.

Wang, P.J. Good, R.F. Lowdon, L.B. Adams, X.Q. Zhou, M.J. Pazin, E.A. Feingold, B. Wold, J. Taylor, A. Mortazavi, S.M. Weissman, J.A. Stamatoyannopoulos, M.P. Snyder, R. Guigo, T.R. Gingeras, D.M. Gilbert, R.C. Hardison, M.A. Beer, B. Ren, A comparative encyclopedia of DNA elements in the mouse genome, Nature. 515 (2014) 355–364. https://doi.org/10.1038/nature13992.

[75] A. Oberst, M. Malatesta, R.I. Aqeilan, M. Rossi, P. Salomoni, R. Murillas, P. Sharma, M.R. Kuehn, M. Oren, C.M. Croce, F. Bernassola, G. Melino, The Nedd4-binding partner 1 (N4BP1) protein is an inhibitor of the E3 ligase Itch, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 11280–11285. https://doi.org/10.1073/pnas.0701773104.

[76] B.A. Azakir, G. Desrochers, A. Angers, The ubiquitin ligase Itch mediates the antiapoptotic activity of epidermal growth factor by promoting the ubiquitylation and degradation of the truncated C-terminal portion of Bid, FEBS J. 277 (2010) 1319–1330. https://doi.org/10.1111/j.1742-4658.2010.07562.x.

[77] S.M. Soond, J.L. Terry, D.W.H. Riches, TRUSS, a tumor necrosis factor receptor-1-interacting protein, activates c-Jun NH2-terminal kinase and transcription factor AP-1, FEBS Lett. 580 (2006) 4591–4596. https://doi.org/10.1016/j.febslet.2006.06.098.

[78] L. Schermelleh, A. Haemmer, F. Spada, N. Rösing, D. Meilinger, U. Rothbauer, C.M. Cardoso, H. Leonhardt, Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation, Nucleic Acids Res. 35 (2007) 4301–4312. https://doi.org/10.1093/nar/gkm432.

[79] H. Pan, B.S. Yan, M. Rojas, Y. V. Shebzukhov, H. Zhou, L. Kobzik, D.E. Higgins, M.J. Daly, B.R. Bloom, I. Kramnik, Ipr1 gene mediates innate immunity to tuberculosis, Nature. 434 (2005) 767–772. https://doi.org/10.1038/nature03419.

[80] J.J. Liu, U. Francke, Identification of cis-regulatory elements for MECP2 expression, Hum. Mol. Genet. 15 (2006) 1769–1782. https://doi.org/10.1093/hmg/ddl099.

[81] S.W. Cho, J. Xu, R. Sun, M.R. Mumbach, A.C. Carter, Y.G. Chen, K.E. Yost, J. Kim, J. He, S.A. Nevins, S.F. Chin, C. Caldas, S.J. Liu, M.A. Horlbeck, D.A. Lim, J.S. Weissman, C. Curtis, H.Y. Chang, Promoter of lncRNA Gene PVT1 Is a Tumor-Suppressor DNA Boundary Element, Cell. 173 (2018) 1398-1412.e22. https://doi.org/10.1016/j.cell.2018.03.068.

[82] T.A. Nguyen, R.D. Jones, A.R. Snavely, A.R. Pfenning, R. Kirchner, M. Hemberg, J.M. Gray, High-throughput functional comparison of promoter and enhancer activities, Genome Res. 26 (2016) 1023–1033. https://doi.org/10.1101/gr.204834.116.

[83] L.J. Core, A.L. Martins, C.G. Danko, C.T. Waters, A. Siepel, J.T. Lis, Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers, Nat. Genet. 46 (2014) 1311–1320. https://doi.org/10.1038/ng.3142.

[84] F.N. Carelli, A. Liechti, J. Halbert, M. Warnefors, H. Kaessmann, Repurposing of promoters and enhancers during mammalian evolution, Nat. Commun. 9 (2018) 1–11. https://doi.org/10.1038/s41467-018-06544-z.

[85] J. Mitchelmore, N.F. Grinberg, C. Wallace, M. Spivakov, Functional effects of variation in transcription factor binding highlight long-range gene regulation by epromoters, Nucleic Acids Res. 48 (2020) 2866–2879. https://doi.org/10.1093/nar/gkaa123.

[86] J. Redondo-Antón, M. Fontela, L. Notario, R. Torres-Ruiz, S. Rodríguez-Perales, E. Lorente, P. Lauzurica, Functional Characterization of a Dual Enhancer/Promoter

Regulatory Element Leading Human CD69 Expression, Front. Genet. 11 (2020) 1–13. https://doi.org/10.3389/fgene.2020.552949.

[87] L.A. Pennacchio, W. Bickmore, A. Dean, M.A. Nobrega, G. Bejerano, Enhancers: Five essential questions, Nat. Rev. Genet. 14 (2013) 288–295. https://doi.org/10.1038/nrg3458.

[88] J.T. Kadonaga, Perspectives on the RNA polymerase II core promoter, Wiley Interdiscip. Rev. Dev. Biol. 1 (2012) 40–51. https://doi.org/10.1002/wdev.21.

[89] M.C. Hofmann, R.A. Hess, E. Goldberg, J.L. Millán, Immortalized germ cells undergo meiosis in vitro, Proc. Natl. Acad. Sci. U. S. A. 91 (1994) 5533–5537. https://doi.org/10.1073/pnas.91.12.5533.

[90] D.M. Church, V.A. Schneider, T. Graves, K. Auger, F. Cunningham, N. Bouk, H.C. Chen, R. Agarwala, W.M. McLaren, G.R.S. Ritchie, D. Albracht, M. Kremitzki, S. Rock, H. Kotkiewicz, C. Kremitzki, A. Wollam, L. Trani, L. Fulton, R. Fulton, L. Matthews, S. Whitehead, W. Chow, J. Torrance, M. Dunn, G. Harden, G. Threadgold, J. Wood, J. Collins, P. Heath, G. Griffiths, S. Pelan, D. Grafham, E.E. Eichler, G. Weinstock, E.R. Mardis, R.K. Wilson, K. Howe, P. Flicek, T. Hubbard, Modernizing reference genome assemblies, PLoS Biol. 9 (2011) 1–5. https://doi.org/10.1371/journal.pbio.1001091.

[91] C.A. Steward, S. Humphray, B. Plumb, M.C. Jones, M.A. Quail, S. Rice, T. Cox, R. Davies, J. Bonfield, T.M. Keane, M. Nefedov, P.J. de Jong, P. Lyons, L. Wicker, J. Todd, Y. Hayashizaki, O. Gulban, J. Danska, J. Harrow, T. Hubbard, J. Rogers, D.J. Adams, Genome-wide end-sequenced BAC resources for the NOD/MrkTac{star, open} and NOD/ShiLtJ{star, open}{star, open} mouse genomes, Genomics. 95 (2010) 105–110. https://doi.org/10.1016/j.ygeno.2009.10.004.

[92] M.P. Bayona-Bafaluy, R. Acín-Pérez, J.C. Mullikin, J.S. Park, R. Moreno-Loshuertos, P. Hu, A. Perez-Martos, P. Fernández-Silva, Y. Bai, J.A. Enríquez, Revisiting the mouse mitochondrial DNA sequence, Nucleic Acids Res. 31 (2003) 5349–5355. https://doi.org/10.1093/nar/gkg739.

[93] M.G. Reese, Application of a time-delay neural network to promoter annotation in the Drosophila melanogaster genome, Comput. Chem. 26 (2001) 51–56. https://doi.org/10.1016/S0097-8485(01)00099-7.

[94] M.G. Reese, G. Hartzell, N.L. Harris, U. Ohler, J.F. Abril, S.E. Lewis, Genome annotation assessment in Drosophila melanogaster, Genome Res. 10 (2000) 483–501. https://doi.org/10.1101/gr.10.4.483.

[95] F.A. Ran, P.D. Hsu, J. Wright, V. Agarwala, D.A. Scott, F. Zhang, Genome engineering using the CRISPR-Cas9 system, Nat. Protoc. 8 (2013) 2281–2308. https://doi.org/10.1038/nprot.2013.143.

[96] M. -C Hofmann, D. Abramian, J.L. Millán, A haploid and a diploid cell cycle coexist in an in vitro immortalized spermatogenic cell line, Dev. Genet. 16 (1995) 119–127. https://doi.org/10.1002/dvg.1020160205.

[97] Y. Suzuki, W.X. Chin, Q. Han, K. Ichiyama, C.H. Lee, Z.W. Eyo, H. Ebina, H. Takahashi, C. Takahashi, B.H. Tan, T. Hishiki, K. Ohba, T. Matsuyama, Y. Koyanagi, Y.J. Tan, T. Sawasaki, J.J.H. Chu, S.G. Vasudevan, K. Sano, N. Yamamoto, Characterization of RyDEN (C19orf66) as an Interferon-Stimulated Cellular Inhibitor against Dengue Virus Replication, 2016. https://doi.org/10.1371/journal.ppat.1005357.

[98] M. Gardiner-Garden, M. Frommer, CpG Islands in vertebrate genomes, J. Mol. Biol. 196 (1987) 261–282. https://doi.org/10.1016/0022-2836(87)90689-9.

[99] S. Ogbourne, T.M. Antalis, Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes, Biochem. J. 331 (1998) 1–14. https://doi.org/10.1042/bj3310001.

[100] A.H. Brand, L. Breeden, J. Abraham, R. Sternglanz, K. Nasmyth, Characterization of a "silencer" in yeast: A DNA sequence with properties opposite to those of a transcriptional enhancer, Cell. 41 (1985) 41–48. https://doi.org/10.1016/0092-8674(85)90059-5.

[101] L. Laimins, M. Holmgren-Konig, G. Khoury, Transcriptional "silencer" element in rat repetitive sequences associated with the rat insulin 1 gene locus, Proc. Natl. Acad. Sci. U. S. A. 83 (1986) 3151–3155. https://doi.org/10.1073/pnas.83.10.3151.

[102] A. Baniahmad, M. Muller, C. Steiner, R. Renkawitz, Activity of two different silencer elements of the chicken lysozyme gene can be compensated by enhancer elements., EMBO J. 6 (1987) 2297–2303. https://doi.org/10.1002/j.1460-2075.1987.tb02504.x.

[103] K.F. Bergeron, T. Cardinal, A.M. Touré, M. Béland, D.L. Raiwet, D.W. Silversides, N. Pilon, Male-Biased Aganglionic Megacolon in the TashT Mouse Line Due to Perturbation of Silencer Elements in a Large Gene Desert of Chromosome 10, PLoS Genet. 11 (2015) 1–26. https://doi.org/10.1371/journal.pgen.1005093.

[104] W.C. Johnson, A.J. Ordway, M. Watada, J.N. Pruitt, T.M. Williams, M. Rebeiz, Genetic Changes to a Transcriptional Silencer Element Confers Phenotypic Diversity within and between Drosophila Species, PLoS Genet. 11 (2015) 1–20. https://doi.org/10.1371/journal.pgen.1005279.

[105] A. Rada-Iglesias, Is H3K4me1 at enhancers correlative or causative?, Nat. Genet. 50 (2018) 4–5. https://doi.org/10.1038/s41588-017-0018-3.

[106] D. Zhang, D. Xie, X. Lin, L. Ma, J. Chen, D. Zhang, Y. Wang, S. Duo, Y. Feng, C. Zheng, B. Jiang, Y. Ning, C. Han, The transcription factor SOX30 is a key regulator of mouse spermiogenesis, Dev. 145 (2018). https://doi.org/10.1242/dev.164723.

[107] D. Huang, H.M. Petrykowska, B.F. Miller, L. Elnitski, I. Ovcharenko, Identification of human silencers by correlating cross-tissue epigenetic profiles and gene expression, Genome Res. 29 (2019) 657–667. https://doi.org/10.1101/gr.247007.118.

[108] C. Wilczek, O. Chayka, A. Plachetka, K.H. Klempnauer, Myb-induced chromatin remodeling at a dual enhancer/promoter element involves non-coding RNA transcription and is disrupted by oncogenic mutations of v-myb, J. Biol. Chem. 284 (2009) 35314–35324. https://doi.org/10.1074/jbc.M109.066175.

[109] N.D. Tippens, A. Vihervaara, J.T. Lis, Enhancer transcription: what, where, when, and why?, Genes Dev. 32 (2018) 1–3. https://doi.org/10.1101/gad.311605.118.

[110] S. Bae, B.J. Lesch, H3K4me1 Distribution Predicts Transcription State and Poising at Promoters, Front. Cell Dev. Biol. 8 (2020) 1–11. https://doi.org/10.3389/fcell.2020.00289.

[111] A. Bessis, N. Champtiaux, L. Chatelin, J.P. Changeux, The neuron-restrictive silencer element: A dual enhancer/silencer crucial for patterned expression of a nicotinic receptor gene in the brain, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 5906–5911. https://doi.org/10.1073/pnas.94.11.5906.

[112] P. Kallunki, G.M. Edelman, F.S. Jones, The neural restrictive silencer element can act as both a repressor and enhancer of L1 cell adhesion molecule gene expression during postnatal development, Proc. Natl. Acad. Sci. U. S. A. 95 (1998) 3233–3238. https://doi.org/10.1073/pnas.95.6.3233.

[113] S.S. Gisselbrecht, A. Palagi, J. V. Kurland, J.M. Rogers, H. Ozadam, Y. Zhan, J. Dekker, M.L. Bulyk, Transcriptional Silencers in Drosophila Serve a Dual Role as Transcriptional Enhancers in Alternate Cellular Contexts, Mol. Cell. 77 (2020) 324-337.e8. https://doi.org/10.1016/j.molcel.2019.10.004.

[114] Y. Naito, K. Hino, H. Bono, K. Ui-Tei, CRISPRdirect: Software for designing CRISPR/Cas guide RNA with reduced off-target sites, Bioinformatics. 31 (2015) 1120–1123. https://doi.org/10.1093/bioinformatics/btu743.

[115] L. Cong, F.A. Ran, D. Cox, S. Lin, R. Barretto, P.D. Hsu, X. Wu, W. Jiang, L. a Marraffini, Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., … Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. Science (New York, N.Y.), Science. 339 (2013) 819–823. https://doi.org/10.1126/science.1231143.Multiplex.

[116] R. Yoneda, Y. Satoh, I. Yoshida, S. Kawamura, T. Kotani, A.P. Kimura, A genomic region transcribed into a long noncoding RNA interacts with the Prss42/Tessp-2 promoter in spermatocytes during mouse spermatogenesis, and its flanking sequences can function as enhancers, Mol. Reprod. Dev. 83 (2016) 541–557. https://doi.org/10.1002/mrd.22650.

[117] B.J. Lesch, S.J. Silber, J.R. McCarrey, D.C. Page, Parallel evolution of male germline epigenetic poising and somatic development in animals, Nat. Genet. 48 (2016) 888–894. https://doi.org/10.1038/ng.3591.

[118] F. Spitz, E.E.M. Furlong, Transcription factors: From enhancer binding to developmental control, Nat. Rev. Genet. 13 (2012) 613–626. https://doi.org/10.1038/nrg3207.

[119] N. Doni Jayavelu, A. Jajodia, A. Mishra, R.D. Hawkins, Candidate silencer elements

for the human and mouse genomes, Nat. Commun. 11 (2020) 1–15. https://doi.org/10.1038/s41467-020-14853-5.

[120] M. Mata-Rocha, J. Hernández-Sánchez, G. Guarneros, E. De La Chesnaye, A.A. Sánchez-Tusié, C.L. Treviño, R. Felix, N. Oviedo, The transcription factors Sox5 and Sox9 regulate Catsper1 gene expression, FEBS Lett. 588 (2014) 3352–3360. https://doi.org/10.1016/j.febslet.2014.07.024.

[121] S. Wu, Y.-C. Hu, H. Liu, Y. Shi, Loss of YY1 Impacts the Heterochromatic State and Meiotic Double-Strand Breaks during Mouse Spermatogenesis, Mol. Cell. Biol. 29 (2009) 6245–6256. https://doi.org/10.1128/mcb.00679-09.

[122] P. Denny, S. Swift, F. Connor, A. Ashworth, An SRY-related gene expressed during spermatogenesis in the mouse encodes a sequence-specific DNA-binding protein, EMBO J. 11 (1992) 3705–3712. https://doi.org/10.1002/j.1460-2075.1992.tb05455.x.

[123] H. Zuo, I. Glaaser, Y. Zhao, I. Kurinov, L. Mosyak, H. Wang, J. Liu, J. Park, A. Frangaj, E. Sturchler, M. Zhou, P. McDonald, Y. Geng, P.A. Slesinger, Q.R. Fan, Structural basis for auxiliary subunit KCTD16 regulation of the GABAB receptor, Proc. Natl. Acad. Sci. U. S. A. 116 (2019) 8370–8379. https://doi.org/10.1073/pnas.1903024116.

[124] Y. Ran, M. Xiong, Z. Xu, W. Luo, S. Wang, Y.-Y. Wang, YIPF5 Is Essential for Innate Immunity to DNA Virus and Facilitates COPII-Dependent STING Trafficking, J. Immunol. 203 (2019) 1560–1570. https://doi.org/10.4049/jimmunol.1900387.

[125] P.P. Reddi, C.J. Urekar, M.M. Abhyankar, S.A. Ranpura, Role of an insulator in testis-specific gene transcription, Ann. N. Y. Acad. Sci. 1120 (2007) 95–103.

https://doi.org/10.1196/annals.1411.012.

[126] S.M. Kehoe, M. Oka, K.E. Hankowski, N. Reichert, S. Garcia, J.R. McCarrey, S. Gaubatz, N. Terada, A conserved E2F6-binding element in murine meiosis-specific gene promoters, Biol. Reprod. 79 (2008) 921–930. https://doi.org/10.1095/biolreprod.108.067645.

[127] J. DeJong, Basic mechanisms for the control of germ cell gene expression, Gene. 366 (2006) 39–50. https://doi.org/10.1016/j.gene.2005.10.012.

[128] J.A. Goodrich, R. Tjian, Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation, Nat. Rev. Genet. 11 (2010) 549–558. https://doi.org/10.1038/nrg2847.

[129] B.W. Sahlu, S. Zhao, X. Wang, S. Umer, H. Zou, J. Huang, H. Zhu, Long noncoding RNAs: New insights in modulating mammalian spermatogenesis, J. Anim. Sci. Biotechnol. 11 (2020) 1–12. https://doi.org/10.1186/s40104-019-0424-8.

[130] M. Kurihara, A. Shiraishi, H. Satake, A.P. Kimura, Erratum: A conserved noncoding sequence can function as a spermatocyte-specific enhancer and a bidirectional promoter for a ubiquitously expressed gene and a testis-specific long noncoding RNA (Journal of Molecular Biology (2015) 426:17 (3069-3093) DOI: 1, J. Mol. Biol. 426 (2014) 3069–3093. https://doi.org/10.1016/j.jmb.2015.03.002.

[131] Y. Diao, R. Fang, B. Li, Z. Meng, J. Yu, Y. Qiu, K.C. Lin, H. Huang, T. Liu, R.J. Marina, I. Jung, Y. Shen, K.L. Guan, B. Ren, A tiling-deletion-based genetic screen for cis-regulatory element identification in mammalian cells, Nat. Methods. 14 (2017) 629–635. https://doi.org/10.1038/nmeth.4264.

[132] N. Rajagopal, S. Srinivasan, K. Kooshesh, Y. Guo, M.D. Edwards, B. Banerjee, T.

Syed, B.J.M. Emons, D.K. Gifford, R.I. Sherwood, High-throughput mapping of regulatory DNA, Nat. Biotechnol. 34 (2016) 167–174. https://doi.org/10.1038/nbt.3468.

[133] V.R. Paralkar, C.C. Taborda, P. Huang, Y. Yao, A. V. Kossenkov, R. Prasad, J. Luan, J.O.J. Davies, J.R. Hughes, R.C. Hardison, G.A. Blobel, M.J. Weiss, Unlinking an lncRNA from Its Associated cis Element, Mol. Cell. 62 (2016) 104–110. https://doi.org/10.1016/j.molcel.2016.02.029.

[134] J.M. Engreitz, J.E. Haines, E.M. Perez, G. Munson, J. Chen, M. Kane, P.E. McDonel, M. Guttman, E.S. Lander, Local regulation of gene expression by lncRNA promoters, transcription and splicing, Nature. 539 (2016) 452–455. https://doi.org/10.1038/nature20149.

[135] R.T. Nakayama, J.L. Pulice, A.M. Valencia, M.J. McBride, Z.M. McKenzie, M.A. Gillespie, W.L. Ku, M. Teng, K. Cui, R.T. Williams, S.H. Cassel, H. Qing, C.J. Widmer, G.D. Demetri, R.A. Irizarry, K. Zhao, J.A. Ranish, C. Kadoch, SMARCB1 is required for widespread BAF complex-mediated activation of enhancers and bivalent promoters, Nat. Genet. 49 (2017) 1613–1623. https://doi.org/10.1038/ng.3958.

[136] E.M. Camino, M.L. Weinstein, M.P. List, J.E. Vellky, M. Rebeiz, T.M. Williams, Red light/green light, a dual fluorescent protein reporter system to study enhancer-promoter specificity in Drosophila, G3 Genes, Genomes, Genet. 10 (2020) 985–997. https://doi.org/10.1534/g3.119.401033.

[137] M. Vigodner, P.L. Morris, Testicular expression of small ubiquitin-related modifier-1 (SUMO-1) supports multiple roles in spermatogenesis: Silencing of sex chromosomes in spermatocytes, spermatid microtubule nucleation, and nuclear

reshaping, Dev. Biol. 282 (2005) 480–492. https://doi.org/10.1016/j.ydbio.2005.03.034.

[138] H.S. Sin, A. V. Kartashov, K. Hasegawa, A. Barski, S.H. Namekawa, Poised chromatin and bivalent domains facilitate the mitosis-to-meiosis transition in the male germline, BMC Biol. 13 (2015) 1–15. https://doi.org/10.1186/s12915-015-0159-8.

[139] T.W. Whiteld, J. Wang, P.J. Collins, E.C. Partridge, S.F. Aldred, N.D. Trinklein, R.M. Myers, Z. Weng, Functional analysis of transcription factor binding sites in human promoters, Genome Biol. 13 (2012). https://doi.org/10.1186/gb-2012-13-9-r50.