DIVISION FOR SEX DIFFERENTIATION

Professor:	*MOROHASHI, Ken-ichirou*
Research Associates:	*FUKUI, Yuko*
	OGAWA, Hidesato
NIBB Research Fellow:	*TSUCHIYA, Megumi*
Postdoctoral Fellows:	*MOHAMAD, Zubair*
	BABA, Takashi
	KOMATSU, Tomoko
	SHIMA, Yuichi
Graduate Students:	*KUSAKA, Masatomo*
	MATSUYAMA, Makoto
	FATCHIYAH
	SATO, Yuko
	MIYABAYASHI, Kanako
	YAMADA, Makoto
	KOUTA, Ryuki
Technical Assistants:	*OWAKI, Akiko*
	ISHIKAWA, Azusa
	KOIKE, Yukari
	GOUNO, Hiroe
	TOMIDA, Sanai
	DODO, Yukiko
Secretary:	*SUGIURA, Mio*

Sexual dimorphism manifests most obviously in the gonads (testis and ovary), and is thereafter observed in other parts of the body such as the external genitalia, muscle, and brain. This process of sex differentiation is divided into three steps. The first step occurs at fertilization, during which the sexes of fertilized eggs are determined genetically according to a combination of sex chromosomes. During the second step, individuals carrying XY and XX sex chromosomes develop the testis and ovary, respectively, in the case of mammals. This gonadal sex differentiation usually proceeds during fetal stages, and subsequently sex steroids synthesized in the sexually differentiated gonads control the sexes of the other tissues. Therefore, the gonadal sexes are quite important for the sex differentiation of animals.

A number of transcription factors are known to play crucial roles in the process of gonadal differentiation. Some of these genes such as *SRY*, *WT1*, *DAX-1*, *SOX9* and *ARX* were identified as the genes responsible for human diseases that display structural and functional defects in the gonads. Functions of the other genes such as *Ad4BP/SF-1*, *Emx2*, *M33*, and *Lhx9* were elucidated by the phenotypes of the gene-disrupted mice. In addition, their expression profiles in the sexually differentiating gonad strongly suggested their functional significance at the early stage of gonadal differentiation. However, it remains to be elucidated how the genes are expressed by upstream regulators. Studies from the aspect above are quite important to define the gene regulatory cascade and the molecular mechanisms mediating sex differentiation of the gonad.

Tentatively, we have hypothesized that the sexually indifferent gonads determine their sexes under the control of two opposite signals, one of which is the signal for male (testicular) differentiation and the other is the signal for female (ovarian) differentiation. Possibly the nature of the signals is transcriptional activities of the transcription factors expressed in the sexually differentiating gonads, or otherwise activities of certain types of growth factors. The study in this division has been performed mainly from the aspect of the transcriptional control of the genes implicated in gonadal sex differentiation.

I. Boundary of *Ad4BP/SF-1* gene locus

The spatial and temporal control of gene expression is essential for establishment of cell fates in multicellular organisms and the transcriptional control requires appropriate interaction between enhancers and basal promoter. In certain genes, this interaction is enhanced by additional regulatory elements such as the locus control region (LCR).

The LCR at the *β-globin* locus is one of the most intensively studied elements, and in the case of the chicken *β-globin* LCR consists of four DNase I hypersensitive (HS) sites, suggesting an open chromatin structure that, unlike closed chromatin, can be easily accessed by transcription factors. Because many regulatory fragments including insulators show nuclease hypersensitivity, we examined whether DNase I HS sites are present between *GCNF* and *Ad4BP/SF-1* gene loci. Nuclei prepared from Y-1 cells, which are derived from mouse adrenal cortex and have robust expression of Ad4BP/SF-1, were treated with DNase I, and three DNase I HS sites (adHS1, 2, and 3) were found between the last exon of *GCNF* and the first exon of *Ad4BP/SF-1* (Figure 1). CTCF, a ubiquitous nuclear protein with multiple zinc finger motifs, is regarded as an insulator-binding protein solely known to date in vertebrates. In order to examine whether CTCF contributed to the nuclease hypersensitivity on adHS sites, we performed chromatin immunoprecipitation (ChIP) assays using anti-CTCF antibody. CTCF-attached DNA fragments were amplified by PCR using primers designed against regions around the adHS sites. Interestingly, CTCF was found to bind with only a region located downstream of adHS1 but not located in the proximity of adHS2 and 3 (CTCF in Figure 1). These data suggest that the DNase I hypersensitivity of adHS1 is most likely due to CTCF-mediated changes in chromatin architecture while the hypersensitivity at adHS2 and 3 results from chromatin structures generated by proteins other than CTCF.

The nuclear matrix is the fraction that is not solubilized by treatment with high salt buffer, and is thought to comprise the nuclear architecture that provides a platform for nuclear events such as transcription and have suggested that the nuclear matrix defines a transcriptional unit by establishing a chromatin loop. In addition, a collection of DNA fragments known as a matrix attachment region (MAR) is reported to be able to protect a transgene from position effects. It is believed that this phenomenon results from insulation by the MAR to block expansion of silent heterochromatin. We therefore examined whether a MAR exists around the adHS sites for the insulator activity on *Ad4BP/SF-1* locus. Interestingly, a fragment corresponding to adHS2 and adHS3 exclusively bound to nuclear matrix (MAR in Figure 1).

Posttranslational modifications of amino terminal tails of histones are thought to be epigenetic marks governing control of gene expression. Acetylation at lysine residues of the tail is known to correlate with activation of transcription, and chromatin associated with an active promoter contains mostly acetylated histones. ChIP assays with anti-AcH3 and AcH4 demonstrated that acetylated boundary was formed around adHS1 and proximal to the transcription start site (AcH3/H4 boundary in Figure 1).

Taken together, we have demonstrated currently the existence of an insulator defining the transcriptional boundary between *Ad4BP/SF-1* and *GCNF* gene loci by DNase I hypersensitive assays. Our further characterization indicated that these HS sites correspond to regions containing potential binding sites for CTCF and nuclear matrix. Importantly, these sites are almost coincident with a divergence of acetylation pattern of histone H3 and H4 tails and constitute a boundary between active and inactive chromatin domains. The correlation between the binding of CTCF and nuclear matrix attachment sites to a functional boundary for transcription is consistent with previous reports that CTCF and the nuclear matrix are required for insulator activities.

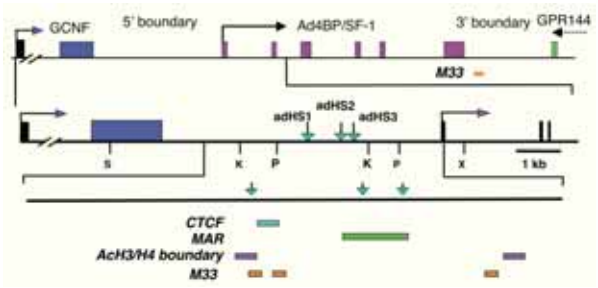

Figure 1. Structure of boundary between *Ad4BP/SF-1* and *GCNF* gene loci. *Ad4BP/SF-1* gene is localized at the downstream of another nuclear receptor gene, *GCNF*, and at the upstream of uncharacterized *GPR144* gene. DNase I hypersensitive sites, adHS1, adHS2, and adHS3, CTCF binding site, nuclear matrix binding site (MAR), acetyl histone H3 and H4, and M33 binding sites are schematically presented. All these sites are localized between *GCNF* and *Ad4BP/SF-1* gene.

II. Function of M33 as a factor bound to the boundary of *Ad4BP/SF-1* locus

In mammals, multiple orthologues of *PcG* components members have been identified. *Polycomb (Pc)* has three mouse counterparts, *M33 (Mpc1)*, *Mpc2*, and *Mpc3*. Likewise, other *PcG* members have multiple counterparts. Therefore, mammalian PcG complexes are thought to be comprised of distinct sets of constituent group members. In fact, mammalian PcG complexes were recently sub-divided into two distinct types according to their biochemical and functional properties. The first complex, containing Eed and Eenhancer of Zeste (Enx1 and Enx2), appears to be required to initiate repression of target gene expression in early development, whereas a second complex containing M33, Mpc2, rae28/Mph1, Bmi1, and Mel18 appears to be required for stable maintenance of the repression state.

Previously we showed that disruption of the murine *M33* gene, a *PcG* member, displayed posterior transformation of the vertebral columns and sternal ribs. In addition, hypoplasia and sex-reversal of the gonads has been observed. In the current study, we identified defects in the splenic and adrenal formation of *M33*-knockout (KO) mice on a C57BL/6 genetic background. Histological examination revealed that the vascular endothelium and its surrounding structures were disorganized. Immunohistochemical analyses demonstrated disturbances in vascular formation and colonization of immature hematopoietic cells. These splenic phenotypes observed in the *M33*-KO mice were quite similar to those seen in *Ad4BP/SF-1* knockouts. Moreover, we observed that the adrenal gland of *M33*-KO and *Ad4BP/SF-1* heterozygous KO mice are smaller than that of wild-type. These phenotypic similarities strongly suggested the presence of genetic and functional correlations between the two genes. In fact, western-blot, immunohistochemistry, and RT-PCR analyses of the *M33* knockouts all indicated significantly decreased expression of Ad4BP/SF-1, indicating that M33 is an essential upstream regulator of *Ad4BP/SF-1* gene.

Upstream stimulatory factors 1 (USF-1 and USF-2), Sox9, and Wt1(-KTS form) have all been identified as direct-acting positive regulators of *Ad4BP/SF-1* gene transcription. These factors bind recognition sites in the 5' upstream region of the *Ad4BP/SF-1* gene and thereby activate gene transcription. However, the present genetic analyses could not distinguish whether *Ad4BP/SF-1* is regulated directly or indirectly by M33. Therefore, we tested whether M33 directly binds to *Ad4BP/SF-1* gene locus by ChIP assay using antibody to M33. Interestingly, the assays with Y-1 adrenocotical cells revealed that target sites of M33-containing PcG complex are present at the upstream region of the first exon and the immediately downstream region of the last exon of *Ad4BP/SF-1* gene (Figure 1). As described above, our recent study demonstrated that the intergenic region between *Ad4BP/SF-1* and *GCNF* gene contains functional architectures such as DNase I hypersensitive sites, a nuclear matrix attachment region (MAR), and CTCF binding sites. Moreover, ChIP assays showed a discontinuous pattern of histone H3 and H4 acetylation over the intergenic region. These observations strongly suggested that this region forms a boundary, the so-called insulator, between the two transcriptional units, *Ad4BP/SF-1* and *GCNF*. This assumption was consistent with a previous observation that CTCF and nuclear matrix are required for insulator activity. In this regard, it seemed interesting to examine whether M33-containing PcG complex binds to this intergenic region because PcG complex is known to function in formation of inactive chromatin arrays. Expectedly, our ChIP assays demonstrated that M33-containing PcG complex binds to sites adjacent to and overlapped with the CTCF binding region and MAR, respectively. Similarly, it is noteworthy

that M33-containing PcG complex bound to region adjacent to the last exon of *Ad4BP/SF-1* gene. Although further examination is required, it is likely that M33-containing PcG complex is implicated in formation of intergenic boundary. This presumptive function of PcG complex is supported by previous *Drosophila* studies. Mihaly *et al.* (1997) showed that a *cis*-regulatory element, Fab-7, consists of a boundary element and PRE, both of which are implicated in regulation of parasegmental-specific expression of bithorax genes. In the case of *gypsy* insulator, Gerasimova and Corces (1998) demonstrated that a *gypsy* insulator binding protein, mod(mdg4), requires Pc, a PcG component, for the insulator function. Considering these observations, the current studies strongly suggested that PcG complex together with CTCF and nuclear matrix forms intergenic boundary between *Ad4BP/SF-1* and *GCNF* genes, although functional and physical correlation of PcG complex with CTCF and/or nuclear matrix remains to be elucidated (Figure 2).

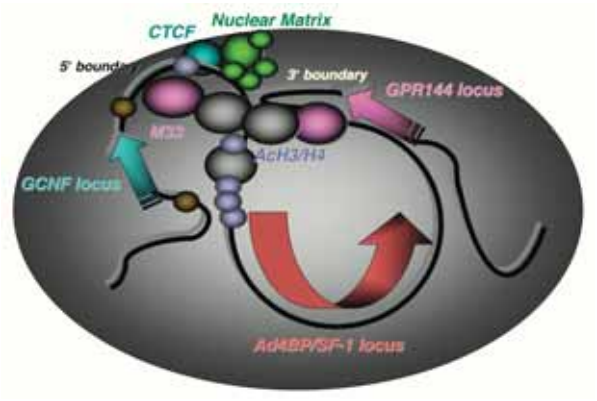

Figure 2. Intranuclear structure of *Ad4BP/SF-1* gene locus. Through binding of the multiple factors at the upstream and downstream of *Ad4BP/SF-1* gene, the gene locus appears forms structurally and functionally a single transcription unit.

It has been well accepted that PcG complex directly binds to target gene loci to keep silent chromatin conformation. Nevertheless, in our current study, multiple binding regions of the M33-containing PcG complex were found to lie at the 5' boundary region of the transcriptionaly active *Ad4BP/SF-1* gene. Although our findings seemed contradictory to the accepted notion, recent ChIP analyses with antibodies to PcG components have accumulated observations similar to the present study, showing that PcG complex binds to not only inactive but active gene loci. Accordingly, it is reasonable to assume that PcG complex is implicated in the formation of transcriptionaly active chromatin unit. Based on these observations together with our current results, *Ad4BP/SF-1* gene is thought to form a transcriptional unit through forming complexes M33 containing polycomb, CTCF, and nuclear matrix. Considering that the complex formation is tightly coupled with the gene transcriptional regulation, it might be noteworthy that structural and thus functional properties of this region might be differentially modified according to developmental stages, tissues and possibly sexes.

Publication List:

Original papers

Baba, T., Mimura, J., Nakamura, N., Harada, N.,Yamamoto, M., Morohashi, K.,* and Fujii-Kuriyama, Y.* (*; The last two authors equally contributed to this work.) (2005). Ah (dioxin) receptor as a key factor in the regulation of female reproduction. Mol. Cell. Biol. in press.

Ishihara, S., and Morohashi, K. (2005). Identification of the boundary for histone acetylation between nuclear receptor genes, Ad4BP/SF-1 and GCNF, aligned in tandem. Biochem. Biophys. Res. Commun. *329*, 554-562.

Kamei, Y., Aoyama, Y., Fujimoto, T., Kenmotsu, N., Kishi, C., Koushi, M., Sugano, S., Morohashi, K., Kamiyama, R., and Asakai, R. (2005). A steroidogenic cell line with differentiation potential from mouse granulosa cells, transfected with Ad4BP and SV40 large T antigen genes. J. Endocrinol. *185*, 187-195.

Katoh-Fukui, Y., Owaki, A., Sotoyama, Y., Kusaka, M., Shinohara, Y., Maekawa, M., Toshimori, K., and Morohashi, M. (2005). Mouse *Polycomb M33* is required for splenic vascular and adrenal gland formation through regulating *Ad4BP/SF-1* expression. Blood *106*, 1612-1620.

Matsuyama, M., Mizusaki, H., Shimono, A., Mukai, T., Okumura, K., Abe, K., Shimada, K., and Morohashi, K. (2005). Novel isoform of vinexin, vinexin γ regulates *Sox9* gene expression through activating MAPK cascade in mouse fetal gonad. Genes Cells *10*, 421-434.

Shima, Y., Mohamad, Z., Ishihara, S., Shinohara, Y., Oka, S., Kimura, S., Suita, S., and Morohashi, K. (2005). VMH specific enhancer of *Ad4BP/SF-1* gene. Mol. Endocrinol. *19*, 2812-2823.

Toda, K., Hayashi, Y., Okada, T., Morohashi, K., and Saibara, T. (2005). Expression of the estrogen-inducible EGFP gene in aromatase-deficient mice reveals differential tissue-response to estrogenic compounds. Mol. Cell. Endocrinol. *229*, 119-126.

Yoshioka, H., Ishimaru, Y., Sugiyama, N., Tsunekawa, N., Noce, T., Kasahara, M., and Morohashi, K. (2005). Mesonephric *FGF* signaling is associated with the development of sexually indifferent gonadal primordium in chick embryos. Dev. Biol. *280*, 150-161.