Dynamic in vivo binding of transcription factors to cis-regulatory modules of cer and gsc in the stepwise formation of the Spemann–Mangold organizer

Norihiro Sudou1,2,3, Shinji Yamamoto1, Hajime Ogino2,3 and Masanori Taira1,*

SUMMARY

How multiple developmental cues are integrated on cis-regulatory modules (CRMs) for cell fate decisions remains uncertain. The Spemann–Mangold organizer in Xenopus embryos expresses the transcription factors Lim1/Lhx1, Otx2, Mix1, Siamois (Sia) and VegT. Reporter analyses using sperm nuclear transplantation and DNA injection showed that cerberus (cer) and goosecoid (gsc) are activated by the aforementioned transcription factors through CRMs conserved between X. laevis and X. tropicalis. ChIP-qPCR analysis for the five transcription factors revealed that cer and gsc CRMs are initially bound by both Sia and VegT at the late blastula stage, and subsequently bound by all five factors at the gastrula stage. At the neurula stage, only binding of Lim1 and Otx2 to the gsc CRM, among others, persists, which corresponds to their co-expression in the prechordal plate. Based on these data, together with detailed expression pattern analysis, we propose a new model of stepwise formation of the organizer, in which (1) maternal VegT and Wnt-induced Sia first bind to CRMs at the blastula stage; then (2) Nodal-inducible Lim1, Otx2, Mix1 and zygotic VegT are bound to CRMs in the dorsal endodermal and mesodermal regions where all these genes are co-expressed; and (3) these two regions are combined at the gastrula stage to form the organizer. Thus, the in vivo dynamics of multiple transcription factors highlight their roles in the initiation and maintenance of gene expression, and also reveal the stepwise integration of maternal, Nodal and Wnt signaling on CRMs of organizer genes to generate the organizer.

KEY WORDS: ChIP-qPCR, Spemann–Mangold organizer, Xenopus, Gene regulation

INTRODUCTION

One of the main questions in the field of developmental biology is how are multiple developmental cues integrated stepwise with the transcriptional regulation of cell fate decisions. To date, a large number of cis-regulatory modules (CRMs) have been identified by promoter/enhancer analyses in many organisms, and conceptual gene regulatory networks have been created for various developmental processes (Davidson and Levine, 2008). However, little is known about the in vivo status of transcriptional regulation or the stepwise integration of developmental cues, at least in vertebrate embryogenesis.

The formation of the Spemann–Mangold organizer in Xenopus is thought to be established by two major signaling pathways: Nodal/activin for mesendoderm induction and Wnt/β-catenin for dorsal determination. The initial state consists of maternal VegT (mVegT), a T-box transcriptional factor in the vegetal hemisphere (Lustig et al., 1996; Zhang and King, 1996), and nuclear β-catenin downstream of canonical Wnt signaling in the dorsal region (Blythe et al., 2010). At the blastula stage, mVegT and Wnt signaling coordinately upregulate nodal expression (Xnr5 cluster genes and Xnr6, called Xnr5 hereafter), which forms a Nodal gradient in a dorsoventral direction in the presumptive endoderm (Agius et al., 2000; Takahashi et al., 2000; Takahashi et al., 2006).

In the dorsal region, high Nodal signaling and Wnt/β-catenin signaling are integrated to form the gastrula organizer. During this process, Nodal signaling induces endoderm- and mesoderm-specific transcriptional activators, such as Mix1 (Rosa, 1989); zygotic VegT (zVegT) and organizer-specific transcription factors, such as Lim1/Lhx1 (Taira et al., 1992; Taira et al., 1994); FoxA4/XFKH1 (Dirksen and Jamrich, 1992); and Otx2/Otx2 (Blitz and Cho, 1995; Pannese et al., 1995; Yamamoto et al., 2003). Wnt/β-catenin signaling induces expression of the organizer-specific transcriptional activators Siamois (Sia) and its paralog Twin (Lemaire et al., 1995; Laurent et al., 1997). In addition, the combination of these signaling proteins induces the production of a transcriptional repressor, Goosecoid (Gsc) (Watabe et al., 1995; Wessely et al., 2001). While the organizer is being formed, the secreted BMP antagonists Noggin, Chordin, Follistatin, Xnr3 and Cerberus emanate from the organizer region to induce neuroectoderm in the dorsal animal ectoderm and also dorsalize the mesoderm to further establish and maintain the organizer. In addition, Cerberus, which is expressed in the head organizer region, also inhibits Nodal and Wnt signaling to prevent their posteriorizing effects (Piccolo et al., 1999). Thus, the outline of organizer formation and neural induction has been established.

To elucidate the molecular basis of organizer formation in Xenopus, several groups, including ours, have investigated the regulation of organizer genes such as (1) gsc by Nodal and Wnt signaling (Watabe et al., 1995), and (2) lim1 and hhex by Nodal signaling (Watanabe et al., 2002; Rankin et al., 2011). In addition, the regulation of target genes for Gsc and Lim1 have been studied, such as (3) wnt8 and brachyury (bra/Xbra) by Gsc (Latinkic et al., 1997; Latinkic and Smith, 1999; Yao and Kessler, 2001) and (4) gsc and cer by Lim1 and Mix1 (Latinkic and Smith, 1999; Mochizuki et al., 2000; Yamamoto et al., 2003). However, these

Accepted 28 February 2012
experiments used mainly plasmid DNA injection for luciferase reporter assays. This approach sometimes does not reflect endogenous gene regulatory mechanisms, compared with transgenic reporter assays using nuclear transplantation (Kroll and Amaya, 1996). Furthermore, the dynamics of in vivo binding of transcription factors to their regulatory elements have not been investigated, particularly in terms of the timing and the order of their binding to target elements for the initiation, maintenance and cessation of gene expression.

The 5' region of gsc is one of the best-studied transcriptional regulatory regions. This contains the distal element (DE), the proximal element (PE) and the upstream element (UE) (Watabe et al., 1995; Mochizuki et al., 2000). The DE and PE are involved in the induction of gsc by Nodal and Wnt signaling, respectively (Watabe et al., 1995). It has been suggested that both regions are crucial for transcription factors to bind to the gsc gene as maintenance factors throughout the UE and DE for Lim1 (Mochizuki et al., 2000) or through the DE and PE for Mix1 (Latinkic and Smith, 1999).

Regarding the regulation of the cer gene, we have shown that the five consecutive homeodomain core binding sites A, B, C, D and E (named the 5×TAAT element or ABCDE element) in the 5' region are crucial for cer expression, that a complex of Mix1, Sia and Lim1 binds to the region containing sites A, B and C (named 3×TAAT element or ABC element), and that Otx2 binds to sites E (Yamamoto et al., 2003). These studies suggest that the DE, PE and UE of gsc and the 5×TAAT/ABCDE element of cer can integrate Nodal and Wnt signaling for organizer expression. However, the question remains as to how gsc and cer are regulated in vivo by organizer-expressing transcription factors through these CRMs.

In this paper, we first investigated the regulatory element of the cer gene by using comparative transgenic reporter and sequence analyses. We found significant variations in the 5×TAAT/ABCDE element in four Xenopus species and identified conserved T-box-binding sites that respond to VegT. We also found VegT-response elements near the DE and PE of the gsc gene. We next performed chromatin immunoprecipitation coupled with quantitative PCR (ChIP-qPCR) analysis with anti-Lim1, Mix1, Sia, Otx2 and VegT antibodies. This analysis revealed dynamic in vivo binding of these transcription factors to the cer and gsc regulatory modules – named cer-U1 and gsc-U1, respectively. Finally, we analyzed the overlapping expression domains of organizer genes, as well as endodermal and mesodermal genes at the blastula to neurula stages, showing two distinct origins of organizer gene expression. These data show, for the first time to our knowledge, the dynamics of in vivo binding of multiple transcription factors in early vertebrate development, leading to a new model of stepwise formation of the Spemann–Mangold organizer.

**MATERIALS AND METHODS**

**Embryos and manipulation**

Xenopus embryos were obtained by artificial fertilization, dejellied and incubated in 0.1×Steinberg’s solution (Peng, 1991). Embryos were staged according to the criteria of Nieuwkoop and Faber (Nieuwkoop and Faber, 1967).

**Cloning of cer regulatory regions**

An about 2 kb DNA fragment of the X. tropicalis cer gene (from –2005 to +37) was obtained by PCR amplification with genomic DNA and specific primers (supplementary material Table S1). The amplified fragment was inserted into the KpnI and Xhol sites of pGL3-EGFP. About 1 kb genomic DNA fragments of X. borealis and X. mulleri cer genes were isolated by PCR with degenerated primers designed based on the X. laevis cer sequence: from –288 to +958 (supplementary material Table S1). Adult X. borealis and X. mulleri frogs were purchased from the Watanabe breeder (Japan).

**mRNA synthesis and DNA injection reporter assays**

Plasmids for mRNA synthesis and reporter constructs are shown in supplementary material Table S2 and Fig. S9. Synthesis of mRNA and luciferase assays were carried out as described (Yamamoto et al., 2003). Reporter DNA was injected at 50 pg/embryo. The statistical significance (P-value) was calculated using Student’s t-test.

**Transgenic and co-transgenic reporter assays using the nuclear transplantation technique**

Transgenesis was performed as described previously (Murray, 1991; Kroll and Amaya, 1996) with some modifications (Tanaka et al., 2003). Injection experiments used mainly plasmid DNA injection for luciferase reporter assays. This approach sometimes does not reflect endogenous gene regulatory mechanisms, compared with transgenic reporter assays using nuclear transplantation (Kroll and Amaya, 1996). Furthermore, the dynamics of in vivo binding of transcription factors to their regulatory elements have not been investigated, particularly in terms of the timing and the order of their binding to target elements for the initiation, maintenance and cessation of gene expression.

**RESULTS**

**Comparison of the cer genes of Xenopus species**

To identify evolutionarily conserved regulatory elements, we compared nucleotide sequences of the cer gene of X. laevis (Xl_cer) with that of X. tropicalis (Xt_cer). Both genes have a highly conserved region at –300/-70 bp upstream of the transcription start site (supplementary material Fig. S1A). This conserved 5’ region includes the 5×TAAT/ABCDE element, but no other conserved TAAT sites are apparent in 2 kb of the 5’ region (supplementary material Fig. S1B). We further examined evolutionary conservation of the ABCDE element between four species, X. tropicalis (Xt), X. laevis (Xl), X. borealis (Xb) and X. mulleri (Xm) (Fig. 1A; supplementary material Fig. S1C). X. tropicalis is diploid, whereas the other species are allotetraploid and have two ‘homoeologs’, called a and b genes, which were supposedly derived from the two distinct parental species, thereby designated cer-a and cer-b.

**Whole mount in situ hybridization**

Whole-mount in situ hybridization was carried out as described (Harland, 1991). For hemissections, rehydrated embryos were embedded in 2% low melting agarose gel in 0.4×Marc’s Modified Ringer (MMR). Co-transgenesis was performed as described previously (Ogino et al., 2008). The –62Xt_cer/EGFP construct was used as a basal promoter, which was digested with PstI at the 5’ end of the promoter. PCR fragments digested with BsmI and the digested –62Xt_cer/EGFP construct were mixed at a 4:1 molar ratio.

**Antibodies**

Anti-Lim1, Otx2, Sia and VegT rabbit polyclonal antibodies were generated using the glutathione-S-transferase (GST) gene fusion system, and specificities of affinity-purified antibodies were verified with western blotting and immunostaining (supplementary material). Plasmids for fusion constructs are listed in supplementary material Table S2.

**ChIP-qPCR**

ChIP experiments were carried out with a chromatin immunoprecipitation kit (Upstate) as described previously (Kato et al., 2002) with some modifications. An aliquot of fragmented chromatin preparation (13 μg DNA equivalent to 50–100 embryos) was incubated with 10 μg of anti-Lim1, Otx2, Sia, Mix1 or VegT antibodies or preimmune immunoglobulin G at a 4:1 molar ratio.
Because the ABCDE element is important for gene expression in reporter assays (Yamamoto et al., 2003), these differences between the homoeologs could cause differences in gene expression. Therefore, we compared expression levels between $\textit{Xt}_c\text{er-a}$ and $\textit{Xt}_c\text{er-b}$ using expressed sequence tag (EST) search with their overlapping coding sequences (supplementary material Fig. S2A). The data show that 92% (66/72) of ESTs are $\textit{Xt}_c\text{er-a}$, suggesting that this gene is transcribed preferentially in the embryo. Because the ABC element of $\textit{Xt}_c\text{er-a}$ (designated $\textit{Xt}_c\text{er-ABC}$) is most deviated from the type of $\textit{Xt}_c\text{er}$ (designated $\textit{Xt}_c\text{er-ABC}$), we next examined whether $\textit{Xl}_c\text{er-ABC}$ and $\textit{Xt}_c\text{er-ABC}$ are equally functional for dorsal expression. To test this, we generated luciferase reporter constructs fused to the SV40 minimal promoter (Fig. 1B). The SV40 promoter alone did not show much difference in expression levels between the dorsal marginal zone (DMZ) and the ventral marginal zone (VMZ). By contrast, the $(\textit{Xla}_c\text{er-ABC})^5$ but not $(\textit{Xt}_c\text{er-ABC})^5$ construct showed dorsal-specific expression (Fig. 1B). This implies that $\textit{Xla}_c\text{er}$ evolved from the ancestral type (supposedly $\textit{Xt}_c\text{er}$) to exert such expression by changing, in part, the nucleotide sequences of both sites B and C. This raised the question as to which elements other than $\textit{Xt}_c\text{er-ABC}$ might confer dorsal-specific expression of $\textit{Xt}_c\text{er}$.

**Transgenic reporter assays for the ABCDE element and T-box sites of the cer gene**

To compare the 5′ regulatory regions of $\textit{Xla}_c\text{er}$ (designated as $\textit{Xl}_c\text{er}$ hereafter) and $\textit{Xt}_c\text{er}$, we used transgenic reporter assays, because DNA injection reporter analysis in $\textit{Xenopus}$ does not always reflect endogenous gene regulatory mechanisms. We first observed that 2 kb of the 5′ regions of both $\textit{Xl}_c\text{er}$ and $\textit{Xt}_c\text{er}$ recapitulate the endogenous cer expression pattern (Fig. 2A-C). To narrow down responsive regions, we used co-transgenesis assays (Ogino et al., 2008) (supplementary material Fig. S3A) and also examined several deletion constructs (Fig. 2). The data show that the −452/−62 region of $\textit{Xt}_c\text{er}$ was necessary and that the −229 region was sufficient (both −229$\textit{Xl}_c\text{er}$ and −229$\textit{Xt}_c\text{er}$) for exhibiting cer-like expression. Both the −229/+166 region and the ABCDE element were required for full activity (−229–Aabcde, −229–Mabcde, and −166–w of $\textit{Xl}_c\text{er}$ and $\textit{Xt}_c\text{er}$), suggesting that the −229/+166 region upstream of the ABCDE element has additional CRMs for both $\textit{Xl}_c\text{er}$ and $\textit{Xt}_c\text{er}$ expression.

We searched the −229/+166 region for transcription factor binding sites, which are conserved between the six orthologs and homoeologs of the four $\textit{Xenopus}$ species, and found three putative T-box recognition sequences (TNNCAC) (Erives and Levine, 2000; Oda-Ishii et al., 2005), named T1, T2 and T3 (supplementary material Fig. S1C). Therefore, we examined T-box mutant constructs using transgenic assays. Fig. 2D shows that any one of the three T-box sites is functional in responding to the endogenous T-box transcription factor(s).

**VegT response of the cer reporter gene**

Because vegt is co-expressed with cer in the endoderm and mesoderm regions during the blastula to gastrula stages (see Fig. 5), and the T3 sites are in good agreement with a VegT consensus binding site (Conlon et al., 2001), we examined whether VegT is involved in regulating the cer gene. We first showed that the endogenous cer gene was activated by VegT in the animal ectoderm (supplementary material Fig. S3B). We next tested...
constructs. 

Fig. 2. Cis-element analysis of X. laevis and X. tropicalis cer promoter regions using transgenic reporter genes. (A) Deletion and point mutations of cer promoter-EGFP reporter constructs and their dorsal expression at the gastrula stage. Constructs of Xl_cer and Xt_cer promoter regions using transgenic reporter genes. X. laevis

Typical expression patterns of (B) or (C) reporter constructs are shown. (D) Reporter gene expression of T-box mutant constructs.

whether luciferase reporter assays using DNA injection could be applied to analyze the responsiveness of cer reporter constructs to endogenous and exogenous factors in the embryo. Both –229Xl_cer/Luc and –229Xl_cer/EGFP were specifically activated in the DMZ (Fig. 3A) and furthermore –229Xl_cer/EGFP was upregulated by VegT in the animal ectoderm (supplementary material Fig. S3C), indicating that DNA injection reporter assays can be used for such analyses.

To examine the synergy between VegT and other transcription factors, we used a lower dose of VegT, which was not sufficient to activate the reporter. As shown in Fig. 3B, –229Xl_cer/Luc was activated synergistically by VegT and a mixture of Mix1, Sia, Lim1 and Otx2 (called 4 mix). To examine the necessity of VegT-binding sites T1 to T3 in the –229 region (supplementary material Fig. S1C), one or all of them were mutated (designated MT1, MT2, MT3 and MT123). Whereas MT1 and MT2 appeared to change the synergistic activation levels, MT3 and MT123 completely lost them (Fig. 3B), indicating that T3 is the VegT response site. Furthermore, among the four, Sia, and to a lesser extent Lim1, enhanced activation of –229Xl_cer/Luc by VegT (Fig. 3C), suggesting functional interactions between VegT and Sia/Lim1. It should be noted that activation of the reporter gene by Mix1 was not enhanced by VegT, which is in contrast to Sia (Yamamoto et al., 2003). These data suggest that the T3 site is a main site for VegT to collaborate with Sia and Lim1 in activating the cer promoter.

Although we have shown a clear difference in dorsal-specific expression between (Xl_ABC)/Luc and (Xt_ABC)/Luc (Fig. 1B), both –229Xl_cer and –229Xt_cer have sufficient activities to drive such expression in both transgenic (Fig. 2A) and DNA injection (Fig. 3A) reporter assays. These data suggest that the combination of the T-box sites and the ABCDE element forms a robust functional unit as a whole that tolerates nucleotide changes in some of its elements. Therefore, we defined a region containing the T-box sites and the ABCDE element as a CRM, referred to as cer-U1 (standing for 'upstream regulatory module 1'; supplementary material Fig. S1C).

T-box sites and VegT response of the gsc reporter gene

Because gsc is directly induced by VegT (Messenger et al., 2005), we examined whether gsc has conserved VegT-binding sites near the UE, DE and PE. Comparison of the 5’ flanking sequences of the gsc genes of X. laevis and X. tropicalis (designated Xl_gsc and Xt_gsc, respectively) identified three conserved consensus T-box sites, T1, T2 and T3, in which T2 (TCACC) is perfectly conserved but the UE is not (supplementary material Fig. S4). Thus, gsc is likely to be regulated by the same set of transcription factors, Sia, Lim1, Otx2, Mix1 and VegT, as cer is.
Dynamics of in vivo binding of Lim1, Otx2, Sia, Mix1 and VegT to the cer and gsc regulatory modules

To examine the dynamics of in vivo binding of transcription factors to cer-U1 and gsc-U1, we performed ChIP-qPCR analysis with X. laevis embryos using rabbit polyclonal antibodies, which were raised against recombinant X. laevis Lim1, Mix1, Sia, Otx2 and VegT proteins and verified for their specificities (supplementary material Fig. S5). The anti-VegT antibodies recognize both mVegT and zVegT, and the anti-Lim1 and Mix1 antibodies did not crossreact with their paralogous proteins Lim5 and Mixer, respectively (Toyama et al., 1995; Henry and Melton, 1998). As shown in Fig. 4A, ChIP-qPCR primer sets 1 through 4 were designed for cer-U1 (primer set 1) and a region within cer intron 1 (primer set 2; a negative control) and for gsc-U1 (primer set 3) and gsc exon 3 (primer set 4; a negative control). Chromatin was prepared from late blastula, early gastrula and neurula embryos.

ChIP-qPCR with primer sets 1 and 3, compared with negative controls with primer sets 2 and 4, revealed that VegT and Sia, but not Mix1, Lim1 or Otx2, were already bound to the cer-U1 and gsc-U1 regions at the late blastula stage (Fig. 4B). These results are consistent with the order of expression of these five transcription factors (supplementary material Fig. S6A), suggesting that both VegT and Sia are involved in the initiation of cer and gsc gene expression at this stage. At the early gastrula stage, all five factors were bound to the cer-U1 and gsc-U1 regions (Fig. 4B). These in vivo binding data strongly support the idea that these five transcription factors are involved in regulating cer and gsc in the organizer through cer-U1 and gsc-U1, and also suggest maintenance roles for Mix1, Lim1 and Otx2 in cer and gsc expression at the gastrula stage.

At the early neurula stage, the binding of all five factors to the cer-U1 region decreased dramatically to basal levels. In the gsc gene, binding of Sia, VegT and Mix1 to the gsc-U1 region also decreased as the expression levels of sia, vegt and mix1 decreased by the early neurula stage. By contrast, the binding of Lim1 and Otx2 to the gsc-U1 region persisted and if anything increased (Fig. 4B). At this stage, because lim1 mRNA gradually disappears in an anteroposterior direction (Taira et al., 1994), lim1 mRNA is not present in the anterior-most region (Fig. 5E; supplementary material Fig. S6C, arrow in the lim1 panel at...
stage 15). However, the Lim1 protein persisted in the prechordal plate and notochord until the tailbud stage (supplementary material Fig. S5N’). Thus, our ChIP-qPCR analysis strongly support the idea that the five transcription factors differentially and coordinately regulate the cer and gsc genes in the formation and functioning of the organizer.

**Detailed comparison of expression domains of organizer-expressing genes**

Given that the cer and gsc genes are regulated by the five transcription factors, as suggested above, their expression domains should overlap each other. The expression domains of these seven genes have already been reported individually (Taira et al., 1992; Bouwmeester et al., 1996; Lemaire et al., 1998; Rankin et al., 2011), but some of their initial or early expression domains have not been reported, such as gsc, mix1, cer, lim1 and otx2. Therefore, we performed whole-mount in situ hybridization using bisected embryos at stages 8-15 for 11 genes (supplementary material Figs S6, S7), and summarized their expression domains on a reported plate and notochord until the tailbud stage (supplementary material Fig. S5N). Thus, our ChIP-qPCR analysis strongly support the idea that the five transcription factors differentially and coordinately regulate the cer and gsc genes in the formation and functioning of the organizer.

**Detailed comparison of expression domains of organizer-expressing genes**

Given that the cer and gsc genes are regulated by the five transcription factors, as suggested above, their expression domains should overlap each other. The expression domains of these seven genes have already been reported individually (Taira et al., 1992; Bouwmeester et al., 1996; Lemaire et al., 1998; Rankin et al., 2011), but some of their initial or early expression domains have not been reported, such as gsc, mix1, cer, lim1 and otx2. Therefore, we performed whole-mount in situ hybridization using bisected embryos at stages 8-15 for 11 genes (supplementary material Figs S6, S7), and summarized their expression domains on a reported plate and notochord until the tailbud stage (supplementary material Fig. S5N). Thus, our ChIP-qPCR analysis strongly support the idea that the five transcription factors differentially and coordinately regulate the cer and gsc genes in the formation and functioning of the organizer.

**Detailed comparison of expression domains of organizer-expressing genes**

Given that the cer and gsc genes are regulated by the five transcription factors, as suggested above, their expression domains should overlap each other. The expression domains of these seven genes have already been reported individually (Taira et al., 1992; Bouwmeester et al., 1996; Lemaire et al., 1998; Rankin et al., 2011), but some of their initial or early expression domains have not been reported, such as gsc, mix1, cer, lim1 and otx2. Therefore, we performed whole-mount in situ hybridization using bisected embryos at stages 8-15 for 11 genes (supplementary material Figs S6, S7), and summarized their expression domains on a reported plate and notochord until the tailbud stage (supplementary material Fig. S5N). Thus, our ChIP-qPCR analysis strongly support the idea that the five transcription factors differentially and coordinately regulate the cer and gsc genes in the formation and functioning of the organizer.

**Detailed comparison of expression domains of organizer-expressing genes**

Given that the cer and gsc genes are regulated by the five transcription factors, as suggested above, their expression domains should overlap each other. The expression domains of these seven genes have already been reported individually (Taira et al., 1992; Bouwmeester et al., 1996; Lemaire et al., 1998; Rankin et al., 2011), but some of their initial or early expression domains have not been reported, such as gsc, mix1, cer, lim1 and otx2. Therefore, we performed whole-mount in situ hybridization using bisected embryos at stages 8-15 for 11 genes (supplementary material Figs S6, S7), and summarized their expression domains on a reported plate and notochord until the tailbud stage (supplementary material Fig. S5N). Thus, our ChIP-qPCR analysis strongly support the idea that the five transcription factors differentially and coordinately regulate the cer and gsc genes in the formation and functioning of the organizer.

**Detailed comparison of expression domains of organizer-expressing genes**

Given that the cer and gsc genes are regulated by the five transcription factors, as suggested above, their expression domains should overlap each other. The expression domains of these seven genes have already been reported individually (Taira et al., 1992; Bouwmeester et al., 1996; Lemaire et al., 1998; Rankin et al., 2011), but some of their initial or early expression domains have not been reported, such as gsc, mix1, cer, lim1 and otx2. Therefore, we performed whole-mount in situ hybridization using bisected embryos at stages 8-15 for 11 genes (supplementary material Figs S6, S7), and summarized their expression domains on a reported plate and notochord until the tailbud stage (supplementary material Fig. S5N). Thus, our ChIP-qPCR analysis strongly support the idea that the five transcription factors differentially and coordinately regulate the cer and gsc genes in the formation and functioning of the organizer.

**Detailed comparison of expression domains of organizer-expressing genes**

Given that the cer and gsc genes are regulated by the five transcription factors, as suggested above, their expression domains should overlap each other. The expression domains of these seven genes have already been reported individually (Taira et al., 1992; Bouwmeester et al., 1996; Lemaire et al., 1998; Rankin et al., 2011), but some of their initial or early expression domains have not been reported, such as gsc, mix1, cer, lim1 and otx2. Therefore, we performed whole-mount in situ hybridization using bisected embryos at stages 8-15 for 11 genes (supplementary material Figs S6, S7), and summarized their expression domains on a reported plate and notochord until the tailbud stage (supplementary material Fig. S5N). Thus, our ChIP-qPCR analysis strongly support the idea that the five transcription factors differentially and coordinately regulate the cer and gsc genes in the formation and functioning of the organizer.
localizations begin to overlap, which was confirmed by whole-mount in situ hybridization with a gsc or vegt probe for each half embryo (supplementary material Fig. S6A,B).

cer expression (group 2) starts at mid-stage 9 in the dorsovegetal region where mveg mRNA exits, and sia expression expands from the mesodermal region as reported (Crease et al., 1998) (Fig. 5B,B’; supplementary material Fig. S6A), Overlapping of cer and sia expression domains was confirmed by whole-mount in situ hybridization using a pair of bisected embryos (supplementary material Fig. S6B). Nodal/activin-inducible mix1, lim1 and otx2 (group 2) start to be expressed in a deep region at mid stage 9 (supplementary material Fig. S6A), but their mRNA localizations just begin to overlap with gsc and cer in the dorsal region at mid stage 9 (Fig. 5B’), accounting for no binding of Mix1, Lim1 and Otx2 to gsc-U1 and cer-U1 at this stage (Fig. 4B). At stages 10 to 10.5, all mRNA localizations of sia, zveg, mix1, lim1 and otx2 overlapped with gsc and cer in the dorsal region (Fig. 5C; supplementary material Fig. S6A, Fig. S7A). Thus, these mRNA localization patterns appear to be consistent with the ChIP-qPCR data (Fig. 4B).

To discriminate between endodermal and mesodermal expressions, DAPI staining was performed for whole-mount in situ hybridization-stained hemisections, in which nuclear density is lower in the endoderm than that in the mesoderm, and sox17b expression was further analyzed to identify the differentiated endoderm region (Hudson et al., 1997) (supplementary material Fig. S7). As a result, expression domains of lim1 and otx2 at stages 10 were assigned to both the dorsoanterior endoderm and the dorso-mesoderm. At this stage, cer expression expanded to the mesodermal region. At stage 10.5/11, cer expression was seen in both the anterior endoderm and the anterior-most dorsal mesoderm, but absent in the posterior mesoderm, whereas lim1 and otx2 expression domains were largely restricted to the dorsal mesodermal region (supplementary material Fig. S7B). Thus, expression domains of organizer genes were dynamically changed from the late blastula to gastrula stages by expansion of expressing areas from the endoderm to mesoderm (cer, lim1 and otx2) and vice versa (gsc, sia and chd).

By the neural stage, cer, vegt, sia and mix1 are greatly reduced in the midline tissues, but the expression of the Lim1 protein and otx2 and gsc mRNAs in the head organizer region persist to the neurula stage (Fig. 5E; supplementary material Fig. S5N,N’, Fig. S6C). Therefore, the Lim1, Otx2 and Gsc proteins are likely to co-exist in the anterior prechordal region, consistent with the ChIP-qPCR data for gsc-U1 at stage 13 (Fig. 4B).

**DISCUSSION**

**Integrated and robust regulation through cer-U1**

We have previously shown the importance of the 3’TAAT/ABC element in the −415 region of the *X. laevis* cer (*Xl_cer*) gene in activation by a complex of Lim1, Mix1 and Sia using DNA injection reporter assays and gel mobility shift assays (Yamamoto et al., 2003). In the present study, using the transgenic technique and sequence comparisons between *Xl* _cer_ and *X. tropicalis* cer (*Xt_cer*), as well as the *X. borealis* and *X. mulleri* cer genes, we have confirmed that the −415/−472 and −229 regions of *Xl* _cer_ and *Xt* _cer_ are sufficient for recapitulating endogenous cer expression in transgenic embryos (Fig. 2). Although DNA injection reporter assays showed the difference in responsiveness of the ABC elements between *Xl* _cer_ and *Xt* _cer_ (Fig. 1B), transgenic reporter assays showed similar activity with the −1938/2005 to −229 regions between *Xl* _cer_ and *Xt* _cer_ (Fig. 2). This implies that cer-U1,
which contains both the VegT response element and the ABCDE element, is a functionally conserved CRM, which can integrate various developmental cues and may be robust against nucleotide changes. Thus, transgenic reporter analyses led to the identification of T-box sites necessary for dorsal expression of the cer gene. This finding further let us to find the importance of T-box site for the gsc gene.

**Subdivision of the organizer region**

We defined here the organizer region as a domain expressing *chd, gsc, lim1* and *otx2* (Fig. 5D; supplementary material Fig. S8). This organizer region includes the BCNE center-derived mesoderm, called here the ‘M’ region, and the dorsal endoderm, called here the ‘E’ region (Fig. 5D). These ‘M’ and ‘E’ regions were verified by whole-mount in situ hybridization for *sox17b* and DAPI staining (supplementary material Fig. S7). Tracing back to the blastula stage, the ‘M’ region is derived from the BCNE region that expresses *sia, chd* and *gsc*, but not *mvegt, lim1* or *otx2* (Fig. 5A,B). By contrast, the ‘E’ region is derived from the dorsoventral region that expresses *mvegt, mix1, cer, lim1* and *otx2*, but not *sia* or *gsc* (Fig. 5A,B’). This means that expression domains of group 1 genes (*sia, chd* and *gsc*) expand from the ‘M’ region to the ‘E’ region during the late blastula to early gastrula, and vice versa for group 2 genes (*cer, lim1* and *otx2*). The group 1 genes have characteristics that can be induced in animal caps by Wnt signaling (*sia and chd*) (Lemaire et al., 1995; Ishibashi et al., 2008) or by Wnt and Nodal/activin signaling (*gsc*) (Crease et al., 1998). The group 2 genes start to be expressed in the dorsoventral region at mid stage 9, and are directly induced by Nodal/activin signaling in animal caps (Taira et al., 1992; Yamamoto et al., 2003). Thus, the initial expression domains of group 1 and group 2 genes are well correlated with the distribution of nuclear β-catenin in the BCNE region (Schohl and Fagotto, 2002) and *Xnr5* expression in the presumptive dorsal endoderm (Takahashi et al., 2006), respectively.

It should be noted that *gsc* expression in the BCNE region at stages 8.5-9 (Fig. 5A,B) as well as mesodermal expression of *lim1* and *otx2* at mid stage 9 (Fig. 5B’) is likely to be induced by *Xnr5* secreted from the presumptive dorsal endoderm, the so-called Nieuwkoop center.

Among organizer genes, *cer* is exceptional, because this gene is not directly induced by activin in animal caps (Yamamoto et al., 2003), in spite of its initial expression in the dorsoventral region. It is probable that *cer* expression is initiated by maternal VegT and Sia, which is suggested by activation of a *cer-U1* reporter by VegT and Sia (Fig. 3C). Thus, the mechanisms of gene induction of *cer* and *gsc* are different from each other at the blastula stage, but appear to become similar at the late blastula to gastrula stage, probably because *cer* and *gsc* are regulated by the same set of transcription factors in the organizer.

**Dynamics of in vivo binding of transcription factors to cer-U1 and gsc-U1**

Conventional reporter analysis can show which transcription factors can bind to regulatory elements and activate or repress reporter genes, but it does not tell when or how long transcription factors bind to regulatory elements. In this study, we performed ChIP-qPCR for *cer-U1* and *gsc-U1* to analyze the five transcription factors that were expected to regulate the *cer* and *gsc* genes (Fig. 4). Based on the developmental expression patterns of *cer* and *gsc*, and their transcriptional regulator genes *vegt, sia, mix1, lim1* and *otx2* (Fig. 5; supplementary material Fig. S6), we interpret the ChIP-qPCR data as follows. First, both mVegT and Sia bind to *cer-U1* and *gsc-U1* at the late blastula stage (mid stage 9) before Lim1, Otx2 and Mix1 do (Fig. 4B). As we have shown overlapping mRNA localizations of *cer, mvegt* and *sia*, as well as those of *gsc, sia* and *zvegt*, at mid-stage 9 (supplementary material Fig. S6A), it is possible that both Sia and VegT are directly involved in the expression of *cer* and *gsc* in prospective ‘E’ and ‘M’ regions, respectively, at this stage (Fig. 6A,D, indicated by asterisks). However, because *vegt* and *sia* mRNAs still stay in the nucleus and their expression domains just expand from the endodermal to mesodermal regions (*vegt*) and vice versa (*sia*) during the blastula stage, accumulation of VegT and Sia may not be enough to regulate *gsc* and *cer*, respectively. If this is the case, what is the significance of the binding of Sia to *cer-U1* (*cer* in Fig. 6D) and of mVegT to *gsc-U1* (*gsc* in Fig. 6A) at stage 9? It has been proposed that some transcription factors bind to tissue-specific enhancers before actual transcription starts and initiate chromatin marking for transcriptional competence by modifying histones. Such transcription factors with these kinds of roles are called ‘pioneer factors’ (Smale, 2010). Therefore, it is possible to speculate that the T-box protein mVegT binds to *gsc-U1* as well as to *cer-U1* as a pioneer factor in the dorsoventral region to maintain *gsc-U1* in a poised state (Fig. 6A) until *gsc* expression starts in this region at the gastrula stage (Fig. 6B). In fact, the T-box sites in the *cer-U1* are necessary for dorsal expression in transgenic analysis (Fig. 2A,B), in which reporter DNA is integrated into chromatin, whereas the TAAT elements in *Xla ABC* are sufficient for dorsal expression in DNA injection reporter assays (Fig. 1B), in which reporter DNA is not integrated into chromatin. Thus, it is possible that mVegT binds to *cer-U1* as a maternal factor to maintain it in a poised state until mid-stage 9 when *cer* expression begins. Likewise, Sia may bind to *cer-U1*, as a pioneer factor, in the prospective mesodermal region (Fig. 6D), leading to the expression of *cer* in the dorsal mesoderm at the gastrula stage (Fig. 6E).

Although *mix1* expression starts before gastrulation, the binding of *Mix1* to *cer-U1* and *gsc-U1* is mainly seen at the gastrula stage. As *Mix1* forms a complex with *Lim1* and *Sia* on the 3×TAAT/ABC element (Yamamoto et al., 2003), binding of *Mix1* may require both *Lim1* and *Sia* or other co-factors to form a complex on *cer-U1* and *gsc-U1* at the gastrula stage (Fig. 6B,E). Because *Mix1* is expressed ubiquitously in the endoderm and mesoderm, its role may be to determine the level of *cer* expression, whereas *Lim1* and *Sia* determine where *cer* is expressed.

The *gsc* reporter gene is activated through the distal element (DE) by *Lim1* and its co-factor *Ldb1* (Mochizuki et al., 2000), and more strongly with Sspd1 (the same as Sspb3) (Nishioka et al., 2005), whereas the *cer* reporter gene is not activated by *Lim1* and *Ldb1* (Yamamoto et al., 2003). Therefore, the *Lim1/Ldb1/Sspb3* complex, together with *Otx2* bind to DE to activate *gsc* but not *cer* in the prospective prechordal plate (i.e. the anterior part of the ‘M’ region) from the gastrula to neurula stages (Fig. 6E,F). In the anterior endoderm at the neural stage, none of the five transcription factors genes is expressed (Fig. 5; supplementary material Fig. S6C), thereby showing no expression of *cer* and *gsc* (Fig. 6C). Thus, all data shown in this and previous papers are consistent and support our model of gene regulations in the organizer.

**Stepwise formation of the Spemann–Mangold organizer by three developmental cues**

We now discuss what the developmental cues are that induce the organizer in the dorsal blastopore lip. In our model for organizer formation, the first cue is vegetally localized maternal factors,
including mVegT, in the presumptive endoderm; the second cue is Sia and Twin, which are directly activated by canonical Wnt signaling in the BCNE center; the third cue involves transcription factors, including Lim1, Otx2, Mix1 and FoxA4, which are directly induced by Nodal/activin signaling. Xnr5 is upregulated in the interior dorsovegetal region by maternal VegT and Wnt signaling (Takahashi et al., 2000). Subsequently, Xnr1 and Xnr2 are induced by Xnr5 in the dorsally tilted vegetal surface (Takahashi et al., 2006), in which the endodermal marker \textit{sox17b} is induced at early stage 9 (supplementary material Fig. S6A). Xnr5 also induces \textit{gsc} and maintain \textit{chd} at stage 8.5 in the BCNE-derived region in combination with Wnt signaling (Fig. 6D), and induces \textit{mix1}, \textit{lim1} and \textit{otx2} in the dorsovegetal region at mid stage 9. At late stage 9 to stage 10, Nodal signaling induces z\textit{vegT}, \textit{mix1}, \textit{lim1} and \textit{otx2} in the ‘M’ region, and \textit{gsc} and \textit{chd} in the ‘E’ region (Fig. 5C,D, Fig. 6B,E). In the ‘M’ region, a combination of z\textit{VegT}, Mix1, Lim1 and Otx2 together with Sia induces the \textit{cer} gene (Fig. 6E). Thus, the order of developmental cues for organizer formation is (1) VegT (vegetal cue), (2) Wnt/Sia (dorsal cue) and (3) Nodal (endoderm and mesoderm cue), which are combined in a stepwise fashion to induce the organizer-expressing transcription factors in the dorsal mesodermal and endodermal regions, and finally mediate organizer-specific transcriptional regulation (Fig. 6). Thus, our detailed expression patterns and in vivo binding data of transcription factors have uncovered the stepwise formation of the Spemann–Mangold organizer.

Acknowledgements
We thank for Toshiaki Tanaka, Mari Itoh and Yumi Izutsu for the transgenic technique; Shoji Tajima for protein syntheses and purification of polyclonal antibodies; Yoichi Kato for a protocol for ChIP; K. Mise for pCS2+2FT-4HA-


