

The shadow enhancer of *short gastrulation* also directs its expression in the ventral midline of the *Drosophila* embryo

Dong-Hyeon Shin¹ · Jung-Woo Hong¹ 

Received: 2 May 2015 / Accepted: 31 May 2015 / Published online: 4 June 2015
© The Genetics Society of Korea and Springer-Science and Media 2015

Abstract The *sog*, a fly homologue of the vertebrate *chordin*, is expressed in the presumptive neurogenic ectoderm and also the ventral midline during *Drosophila* embryogenesis. Previously, two enhancers, called the primary and shadow enhancer, respectively, were found to direct *sog* expression in the neurogenic ectoderm. However, efforts to identify an enhancer that directs *sog* expression in the ventral midline have failed. Here, we present evidence that the *sog* shadow enhancer has dual activities to direct *sog* expression in the neurogenic ectoderm as well as the ventral midline. Systematic truncation analyses employing transgenic embryos revealed that about a 680-bp region within the shadow enhancer is necessary and sufficient for midline enhancer activity. Intriguingly, midline enhancer activity does not appear to depend on binding sites for Sim, because distal and proximal elements critical for *sog* midline expression within the 680-bp region do not have any canonical Sim binding sites.

Keywords *Drosophila* · Transgenic embryo · *Short gastrulation* · Enhancer · Ventral midline

Introduction

Each cell comprising a developing embryo ultimately differentiates into a particular tissue or organ, and its developmental identity is determined by the positional

information of that cell within an embryo (Wolpert 1969, 2011). The positional information of an individual cell is established when spatially restricted expression patterns of developmental genes are reliably formed from initial crude instructive waves [for example (Hong et al. 2008a, b)]. The gene *short gastrulation* (*sog*) is included in a group of seven zygotically active genes [*decapentaplegic* (*dpp*), *zerknüllt* (*zen*), *sog*, *tolloid* (*tld*), *twisted gastrulation* (*tsg*), *screw* (*scw*) and *shrew* (*srw*)] required for dorsal–ventral (DV) pattern formation within the ectoderm of the *Drosophila* early embryo (Francois et al. 1994). Initially, *sog* was identified as one of three X-linked zygotic genes required for specific morphogenetic events of gastrulation (Wieschaus et al. 1984). It subsequently has been shown that *sog* and *dpp* are functionally equivalent to the respective *Xenopus* homologs of *chordin* and *bone morphogenetic protein-4* (*BMP-4*) (Holley et al. 1995). Previous genetic, transgenic, and biochemical analyses indicated seemingly contradictory roles for Sog in modulating Dpp activity. The secreted protein Sog delineates the dorsal border of the neurogenic ectoderm by locally inhibiting the anti-neutralizing activity of Dpp (Ashe and Levine 1999; Biehs et al. 1996), while it induces amnioserosa differentiation in the dorsal-most region by maximizing Dpp signaling at a distance (Ashe and Levine 1999; Shimmi et al. 2005).

The *sog* gene exhibits dynamic expression patterns during embryogenesis (Francois et al. 1994). As early as nuclear cleavage cycle 13, *sog* is first expressed in a broad lateral stripe in the presumptive neurogenic ectoderm, the dorsal limit of which subsequently abuts the ventral border of the *dpp* domain in the dorsal ectoderm. At least by germ-band extension, *sog* expression is restricted to the ventral midline, which is composed of specialized glial cells secreting signals critical for nerve cord patterning (Menne et al. 1997).

Electronic supplementary material The online version of this article (doi:10.1007/s13258-015-0302-5) contains supplementary material, which is available to authorized users.

✉ Jung-Woo Hong
jwhong46@khu.ac.kr

¹ Graduate School of East-West Medical Science, Kyung Hee University, Yongin 446-701, South Korea

The early broad stripe of *sog* expression depends on the DV determinants, dorsal (Dl), twist (Twi) and snail (Sna). Computational analyses of the genome-wide distribution of the Dl recognition sequence have identified an ~400-bp genomic fragment that functions as an enhancer to direct the broad pattern of *sog* expression in the neurogenic ectoderm (Markstein et al. 2002). Subsequent chromatin immunoprecipitation (ChIP)-chip assays suggested that many of the Dl target genes contain two separate enhancers that direct the same or similar expression patterns (Zeitlinger et al. 2007). Recent transgenic studies have showed that the *sog* locus contains a secondary enhancer to direct its expression in the neurogenic ectoderm (Hong et al. 2008a, b). The expression pattern directed by the secondary enhancer is very similar to that of endogenous *sog* expression. The two enhancers are referred to as the “primary” and “shadow” enhancer, respectively, based on the chronological order of their identification rather than any functional differences. However, the later pattern of *sog* expression in the ventral midline appears to closely follow *single-minded* (*sim*) expression. The *sim* gene encodes a basic helix-loop-helix-PAS (bHLH-PAS) transcription factor (TF) and functions as a master regulatory gene controlling differentiation of the ventral midline of the *Drosophila* central nervous system (CNS). Although previous studies have suggested that *sim* is critically involved in *sog* expression in the ventral midline (Zinzen et al. 2006), intense efforts to identify the *sog* midline enhancer have been unsuccessful to date.

Here, we present evidence that the shadow enhancer can also direct *sog* expression in the ventral midline. An approximately 680-bp region was determined to be the minimal sequence for midline enhancer activity based on transgenic embryo analyses. Intriguingly, the midline enhancer activity does not appear to depend on Sim binding sites, because distal and proximal elements within the 680-bp region, which are required for *sog* midline expression, do not have any canonical Sim binding sites.

Materials and methods

Fly stock

Strain *yw*^{67c23} was used for P-element transformation and *in situ* hybridization in *Drosophila melanogaster*. A mutant strain of *sim* (*sim*^{2/H9} *kar*¹/*TM3*, P[*ftz-lacZ*] *SC1*, *Sb*¹ *ry*^{RK}; stock number 2055) was obtained from the Bloomington Stock Center. The *sim*^{2/H9} allele is an amorphic allele, and mutant embryos homozygous for the allele are characterized by the absence of all midline cells after migration of only a few peripheral glial cells (Schmidt et al. 2011). Mutant embryos homozygous for

sim^{2/H9} were distinguished from heterozygotes by the lack of *lacZ* expression in *fushi tarazu* (*ftz*) expression domains mediated by the *ftz-lacZ* fusion gene of the *TM3* balancer chromosome.

Determination of Sim-Tgo-binding consensus sequences and identification of Sim-Tgo sites within the *sog* midline enhancer

In general, it has been accepted that the Sim-Tgo heterodimer prefers to bind the consensus sequence RWACGTG (Wharton et al. 1994). The newly calculated position frequency matrix (PFM) of the Sim-Tgo DNA binding sequence was obtained from a bacterial 1 hybrid (BIH) library (Zhu et al. 2011) (<http://pgfe.umassmed.edu/ffs/>) and used to produce a WebLogo (Crooks et al. 2004) version of the consensus sequence (Fig. 4a; Supplemental Fig. S1), which is highly similar to the typical consensus RWACGTG. In vitro electrophoretic mobility shift assay (EMSA) studies have shown that DDRC (5'-half-sites) and GTG (3'-half-sites) are high-affinity half-site recognition sequences for Sim-Tgo, respectively (Swanson et al. 1995). The second PFM of Sim-Tgo was obtained from the results of the EMSA and also used to generate another consensus sequence of Sim-Tgo (Fig. 4b, Supplemental Fig. S1). These two types of Sim-Tgo consensus sequences were used, one at a time, to search for Sim-Tgo sites within the 0.68-kb enhancer. The search for the Sim-Tgo sites was carried out by the ClusterDraw algorithm (<http://line.bioinfolab.net/webgate/submit.cgi>) fed with the 0.68-kb enhancer sequence and either PFM of Sim-Tgo. IUPAC nucleic acid codes for W, D, and R represent A or T, A, T or G and A or G, respectively.

Plasmid construction, mutagenesis, and P-element-mediated germline transformation

Genomic DNA was isolated from *yw*^{67c23} embryos aged 2–4 h after egg deposition (AED). All genomic regions used for P-element-mediated transformation were prepared by genomic polymerase chain reaction (PCR) amplification (Supplemental Table S1). PCR-amplified genomic fragments were cloned into the PromegaTM pGEM[®]-T Easy vector, and sequences of the cloned fragments were verified by DNA sequencing. Cloned fragments were digested by NotI and inserted into a [(-42)-*eveP-lacZ*]-pCaSpeR vector (Small et al. 1992) that was modified to contain a unique NotI site upstream of the *even-skipped* (*eve*) promoter (*eveP*). Enhancer sequences were all oriented in a 5'–3' direction relative to the chromosomal transcription start-site. Transformation constructs were introduced into the germline of *Drosophila melanogaster*, as described previously (Rubin and Spradling 1982). At least five

independent lines were generated and tested for each construct.

Whole-mount in situ hybridization

Whole-mount in situ hybridization was performed as described in a previous study (Hong et al. 2013). Briefly, embryos were collected 2–10 h AED, dechorinated, fixed, and hybridized with digoxigenin (DIG) UTP-labeled antisense RNA probes. To examine the pattern of endogenous *sim* and *sog* expression in wild-type or homozygous *sim* mutant embryos, antisense *sim* and *sog* RNA probes were produced by in vitro transcription with DNA templates generated by PCR. We used Campos-Ortega and Hartenstein's definitions of developmental stages during embryogenesis (Campos-Ortega and Hartenstein 1985).

Results

The “remote secondary” enhancer or shadow enhancer can also direct *sog* expression in the neurogenic ectoderm

Previous ChIP-chip assays performed with Df, Twi and Sna antibodies resulted in the prediction that many developmental genes may have secondary enhancers (Zeitlinger et al. 2007). For example, whole-genome ChIP-chip assays identified two clusters of Df, Twi, and Sna in the *sog* locus, each of which coincides with either the primary or shadow enhancer (Fig. 1a; Supplemental Table S2) (Hong et al. 2008a, b). The primary and shadow enhancers are located within the first intron and 20 kb 5' of the *sog* transcription start site, downstream of a neighboring gene, *CG8117*, respectively (Fig. 1a).

The discovery of shadow enhancers was the first evidence that two or more enhancers may direct the same or very similar expression patterns of a particular gene in the same time and space. Thus, it was speculated that shadow enhancers might share a common principle to create the specific expression pattern of a gene. Initial purpose of the current study was to elucidate the *cis*-regulatory code, synonymously called “grammar”, shared by the two *sog* enhancers. To do this, transgenic embryos containing either the primary or shadow enhancer were produced and used to test if both enhancers recapitulated the endogenous *sog* lateral expression pattern in our hands. The intronic enhancer produced broad lateral stripes of lacZ expression in transgenic embryos (Fig. 1b), similar to those directed by the newly identified shadow enhancer (Fig. 1c). These results suggested that the two clusters of Df, Twi and Sna function as authentic enhancers to direct *sog* expression in the neurogenic ectoderm.

The shadow enhancer also directs *sog* expression in the ventral midline of the late embryo

While documenting the expression profile of the lacZ transgene directed by the ~0.9-kb shadow enhancer, we observed that the expression pattern of the lacZ transgene in the transgenic embryos recapitulates the endogenous pattern of *sog* expression in both the neurogenic ectoderm and the ventral midline (Fig. 2a–d). The early broad stripe (Fig. 2a) of the lacZ fusion gene narrowed down to the ventral part (Fig. 2b) of the neurogenic ectoderm during gastrulation. Late expression of the lacZ fusion in the ventral midline was first detected immediately after completion of gastrulation (Fig. 2c) and remained during germ band elongation (Fig. 2d). This unexpected pattern of lacZ expression by the shadow enhancer was almost indistinguishable from that of endogenous *sog* expression during at least the first 9 h of embryogenesis (refer to Fig. 2i–l). These results suggested that the shadow enhancer region can direct *sog* expression in the ventral midline as well as in the neurogenic ectoderm, and led us to abruptly change the focus of this study to characterize extended shadow enhancer activity in the ventral midline.

The observed lacZ pattern directed by the shadow enhancer was not only similar to *sog* expression in the ventral midline, but also reminiscent of *sim* expression. The *sim* is a master regulatory gene that directly or indirectly influences the expression of many developmental genes in the ventral midline, thereby governing CNS development (Nambu et al. 1990). Initial transcription of *sim* was restricted to a single line of cells, called mesectoderm, identified on either side of the presumptive mesoderm (Fig. 2e). The symmetric lines of mesectodermal cells converged at the ventral midline following gastrulation (Fig. 2f); these cells eventually form specialized glial cells that secrete signals critical for nerve cord patterning (Menne et al. 1997). Once activated, *sim* expression was maintained via autoregulation during germ band elongation (Fig. 2g, h) and later stages of embryogenesis. Early and transient expression of *sog* gene was observed in broad regions of the presumptive neurogenic ectoderm (Fig. 2i). Expression was restricted to narrow regions of the ventral neurogenic ectoderm (Fig. 2j) and eventually to the ventral midline (Fig. 2k) after gastrulation. The expression of *sog* after the onset of germ band elongation (Fig. 2k, l) was indistinguishable from that of *sim*.

The coincident pattern of *sog* and *sim* expression in the ventral midline raised the question of whether Sim may be required for *sog* expression. To answer this question, *sog* expression was tested in mutant embryos homozygous for *sim* (*sim*^{-/-}) (Fig. 2m–p). As expected, early expression of *sog* was not disturbed in the *sim* mutant embryo (Fig. 2m, n), consistent with previous reports that early

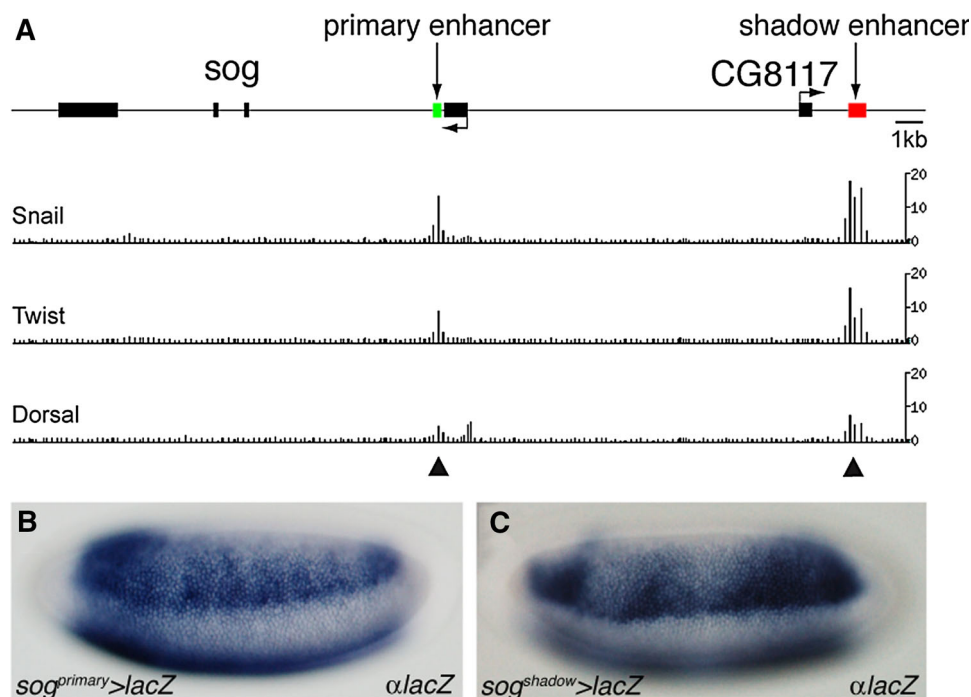


Fig. 1 Primary and shadow enhancers act as authentic enhancers to control lateral stripes of *short gastrulation* (*sog*) in the presumptive neurogenic ectoderm of the *Drosophila* early embryo. **a** The *sog* genomic locus is shown schematically. Black boxes represent exons of *sog* and *CG8117*. Angled arrows indicate start sites and orientations of their transcription. The 0.4-kb primary and ~0.9-kb shadow enhancers are represented by green and red boxes, respectively (Supplemental Table S2). The three horizontal lines below the gene models show the distributions of snail (Sna), twist (Twi), and dorsal (Dl) based on whole-genome ChIP-chip assays (http://flybuzz.berkeley.edu/cgi-bin/gbrowse/fly4_3/) (Zeitlinger et al. 2007). The ChIP-chip assays identified two clusters (black triangles) of Sna, Twi, and Dl around the *sog* locus, each of which coincides with either the primary or shadow enhancer. The primary (**b**) and shadow (**c**) enhancers direct broad stripes of *lacZ* reporter gene expression, a pattern similar to that of endogenous *sog* expression. LacZ expression was visualized by whole-mount in situ hybridization with an antisense RNA probe. Embryos are oriented with anterior to the left and dorsal up

b and **c** enhancers direct broad stripes of *lacZ* reporter gene expression, a pattern similar to that of endogenous *sog* expression. LacZ expression was visualized by whole-mount in situ hybridization with an antisense RNA probe. Embryos are oriented with anterior to the left and dorsal up

sog expression depends on the maternal TFs Dl and Zelda (Zld) (Liang et al. 2008; Markstein et al. 2002). In contrast, the late *sog* pattern in the ventral midline was absent in the same mutant embryo (Fig. 2o, p). These results indicated that *sog* expression depends on the presence of Sim protein in the ventral midline and thus suggested that the shadow enhancer might also contain Sim-Tango (Tgo) binding sites that are found in enhancers of Sim target genes.

A 0.68-kb region within the shadow enhancer is sufficient to direct *sog* expression in the ventral midline

Next, we mapped the minimal region of the shadow enhancer required to control *sog* expression in the ventral midline by employing P element-mediated germline transformation and whole-mount in situ hybridization (Fig. 3). The first set of four constructs (0.732, 0.616, 0.40 and 0.24 kb) was produced by sequentially truncating the

5' region of the 0.88-kb full-length shadow enhancer. Deleting an ~150-bp region (0.732 kb) of the 0.88-kb construct led to complete loss of *lacZ* expression in the ventral midline, suggesting that the 5' ~150-bp fragment plays a critical role in midline enhancer activity. Other constructs that were truncated to a greater extent (0.616, 0.40 and 0.24 kb) could not direct *lacZ* transcription. The second set of three constructs (0.731, 0.54 and 0.37 kb) consisted of a series of 3' truncations. The 3' ~150-bp deletion (0.731 kb) did not have an effect on *lacZ* expression in the ventral midline of a transgenic embryo. More 3' truncation (0.54 kb), however, gave significantly reduced *lacZ* expression and further truncation (0.37 kb) almost failed to direct *lacZ* expression. These results suggested that in addition to the 5' ~150-bp region of the full-length construct, a 200-bp region at the 3' end of the 0.731-kb construct is also required for midline enhancer activity and that the 3' end of the 0.731-kb construct delineates the 3' limit of the minimal region of the *sog* midline enhancer. The map was further refined using three more constructs

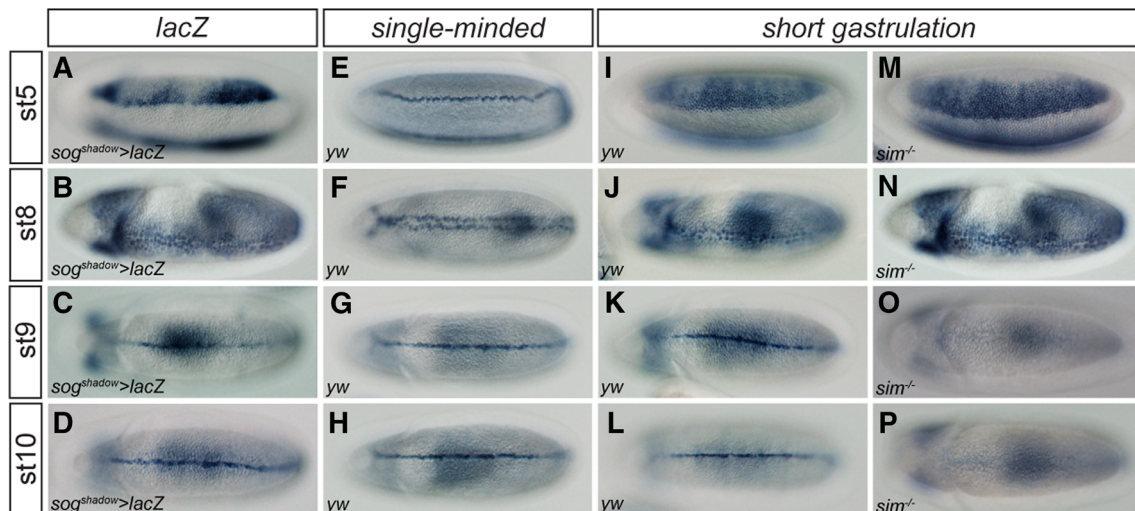


Fig. 2 The shadow enhancer also directs *sog* expression in the ventral midline of the late embryo. About 2–10 h after egg deposition (AED) embryos were collected, dechorinated, and fixed. Whole-mount in situ hybridization was performed with fixed embryos and digoxigenin (DIG)-UTP labeled antisense RNA probes complementary to *lacZ*, *single-minded* (*sim*) and *sog*. Each probe used in the individual in situ hybridization experiment is shown on the top of each column. Expression of a *lacZ* fusion gene directed by the ~0.9-kb shadow enhancer in a transgenic embryo recapitulated the

endogenous *sog* expression in the neurogenic ectoderm and the ventral midline (a–d). Expression patterns of *sim* (e–h) were visualized in wild-type (*yw*) *Drosophila* embryos. An antisense *sog* RNA probe was used to target endogenous *sog* transcripts in both wild-type (*yw*) (i–l) and *sim* mutant (*sim*^{-/-}) (m–p) embryos. The *sim*^{-/-} embryos were homozygous for the *sim*^{2/H9} allele. ‘st’ indicates the developmental stage of *Drosophila* embryogenesis. Developmental stages were defined according to previously established criteria (Campos-Ortega and Hartenstein 1985)

(0.68, 0.610, and 0.50 kb). The 0.68-kb construct, which was the same as the 0.731-kb construct except that the 5' end was about 45 bp shorter, drove *lacZ* expression to the same extent as the full-length enhancer. 3' 70- (0.610 kb) and additional 5' 110-bp (0.50 kb) truncations of the 0.68-kb construct produced barely detectable amounts of *lacZ* transcripts and no *lacZ* transcript, respectively. These results suggested that the 0.68-kb enhancer is the minimal region of the *sog* midline enhancer and that the 5' ~110- and 3' ~70-bp regions of the minimal enhancer are required for the midline enhancer activity of the shadow enhancer. We refer to the 5' ~110- and 3' ~70-bp regions as the distal and proximal elements (Supplemental Table S3), respectively, in terms of their positions relative to the transcription start site of the *sog* gene.

The distal and proximal elements contain no canonical Sim-Tgo binding sites

Studies on *sog* expression in *sim* mutant embryos (Fig. 2m–p) together with a previous *sog* misexpression experiment (Zinzen et al. 2006) have shown that Sim protein is required for *sog* expression in the ventral midline, implying that the 0.68-kb minimal midline enhancer contains canonical Sim-Tgo binding sites. To test this hypothesis, the 0.68-kb enhancer region was searched for two different consensus sequences of Sim-Tgo binding

sites (Fig. 4a, b). One was obtained from a database of TF binding sites built from a bacterial 1-hybrid (B1H) library (Fig. 4a; Supplemental Fig. S1) (Zhu et al. 2011), while the other was produced by recalculating Sim-Tgo binding sequences retrieved from systematic evolution of ligands by exponential enrichment (SELEX) analyses (Fig. 4b; Supplemental Fig. S1) (Swanson et al. 1995). The ClusterDraw algorithm (Papatsenko 2007) fed with position frequency matrices (PFMs) (Supplemental Fig. S1) of the two consensus sequences identified three Sim-Tgo sites in the 0.68-kb enhancer, all of which are located in-between the distal and proximal elements (Fig. 4c; Supplemental Table S4). These Sim-Tgo sites do not appear to be critically involved in the midline enhancer activity for the following three reasons. First, the 0.50-kb construct containing the three Sim sites completely failed to activate *sog* in the ventral midline (Fig. 3), suggesting that those sites are nonfunctional. Second, the distal and proximal elements, identified as critical sequences for *sog* midline expression, do not have any canonical Sim-binding sites. Finally, even if the identified sites are functional, they do not also seem to be indispensable. DNA sequences of the three Sim-Tgo sites are a subset of the consensus sequence containing GCGTG as a core (Fig. 4d, compare with a, b). Previous studies on several enhancers of Sim target genes showed that all midline enhancers tested so far contain at least one ACGTG-core consensus sequence (Hong et al.

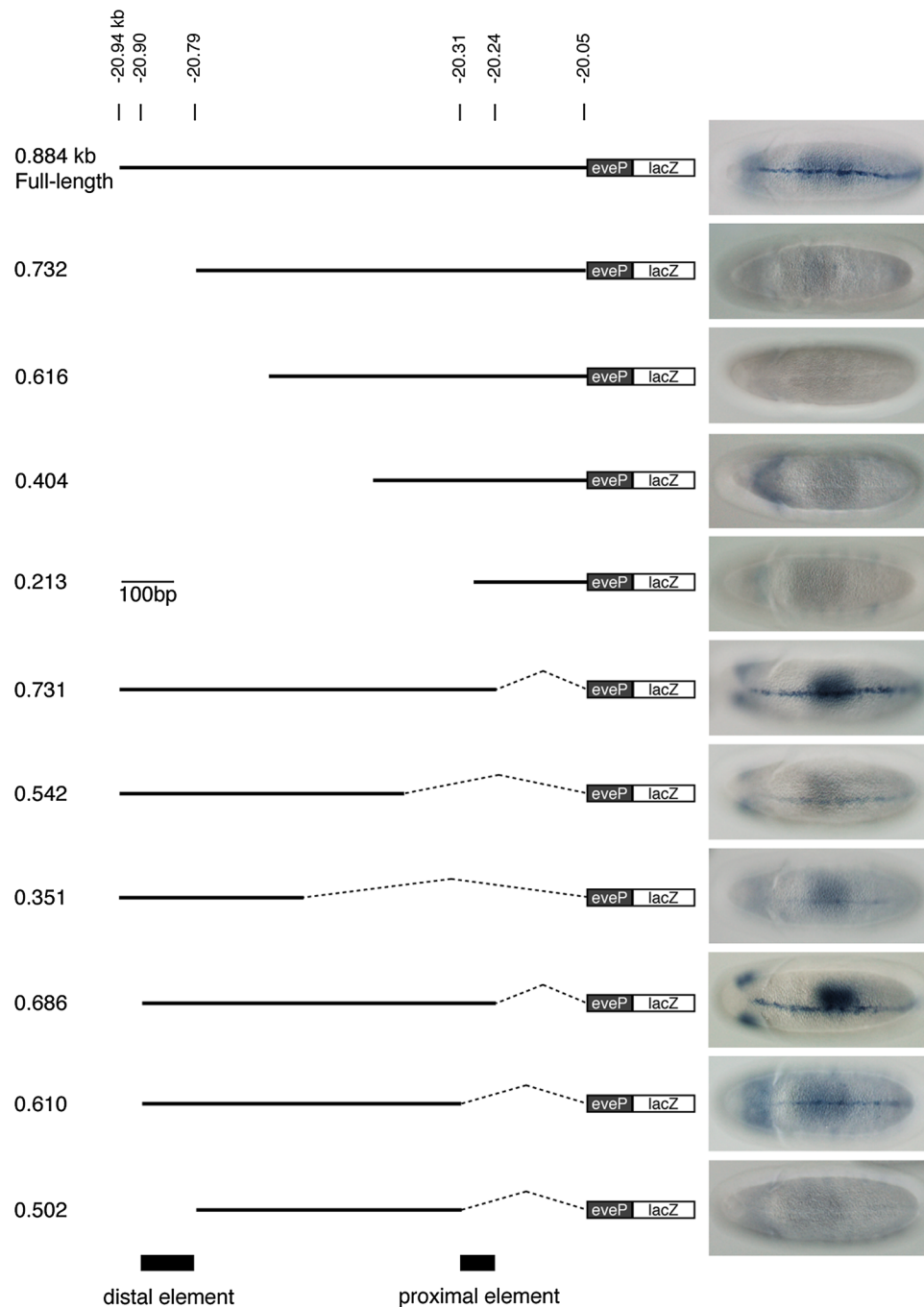


Fig. 3 Two separate DNA elements within the 0.68-kb minimal enhancer are required to direct *sog* expression in the ventral midline. *Top vertical lines* indicate distances from the transcription start site of the *sog* gene. LacZ expression in the transgenic embryos was visualized by in situ hybridization with an antisense lacZ RNA probe. The lacZ patterns in the ventral midline directed by fragments of the *sog* shadow enhancer are shown at the *right* of each construct. The 0.732-kb construct completely failed to direct lacZ expression in the ventral midline, implying that the removed ~150-bp fragment contains an essential element(s) for the midline enhancer activity. Only the full-length construct among the first five constructs could activate lacZ transcription in the ventral midline. The 0.731-kb

construct gave a lacZ expression pattern comparable that mediated by the full-length construct. In contrast, the 0.54- and 0.37-kb constructs gave significantly lower or barely detectable levels of expression, respectively. The 0.68-kb construct directed lacZ expression in a pattern comparable to that directed by the full-length and 0.731-kb constructs. The 0.610-kb construct gave a very low level of expression, while the 0.50-kb construct completely failed to activate the lacZ fusion gene. These data indicate that the 0.68-kb fragment is the minimal region that contains midline enhancer activity and that the 5' 110 bp and 3' 70 bp fragments are indispensable for the midline enhancer activity of the shadow enhancer

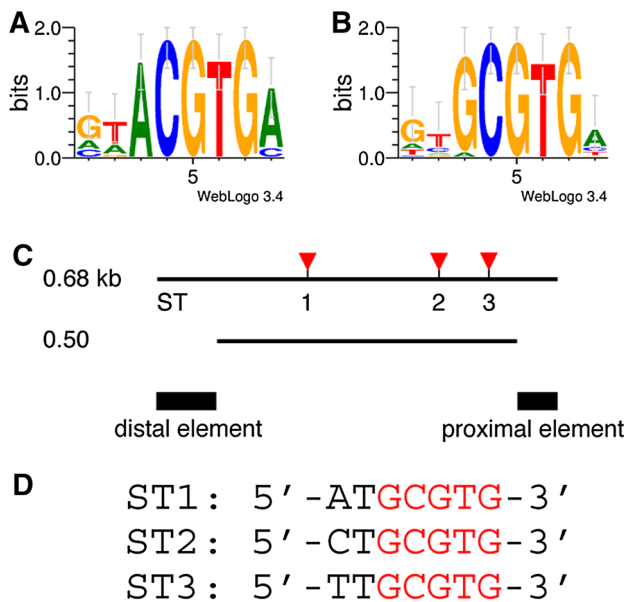


Fig. 4 Expression of *sog* in the ventral midline does not appear to depend on canonical Sim binding sites. Two consensus sequences for Sim-Tgo binding site were used to search the sequence of the 0.68-kb construct. One was obtained from a database of TF binding sites built from a bacterial 1-hybrid (B1H) library (a) while the other was produced by recalculating Sim-Tgo binding sequences retrieved from systematic evolution of ligands by exponential enrichment (SELEX) analyses (b). The collection of those sequences was visualized using the WebLogo 3.4 algorithm (Crooks et al. 2004). The 0.68-kb region contains three putative Sim binding sites (c), all of which contain “GCGTG” as a core sequence (d). ST stands for the binding site of the Sim-Tango (Tgo) heterodimer. Core sequences of the consensus sequence (b) are shown in red (d)

2013; Wharton et al. 1994) (Fig. 4a). These results suggest that the *sog* midline enhancer activity may depend on non-canonical Sim-Tgo sites and/or binding sites for unknown TFs.

Discussion

Numerous developmental studies have revealed that an enhancer directs expression of its target gene only at a particular time and in a particular space. Here, we present the first evidence that an enhancer of the *sog* gene, called a shadow enhancer, sequentially directs the expression of *sog* at different times and in different spaces, namely the neurogenic ectoderm of the early embryo and the ventral midline of the late embryo.

Genetic (Fig. 2), transgenic (Fig. 3), and bioinformatics (Fig. 4) analyses showed that *sog* expression in the ventral midline requires the Sim protein, even though the critical regions of the midline enhancer contain no canonical Sim-binding sites. How can these paradoxical observations be explained? The simplest interpretation is that recruitment of

Sim protein to the shadow enhancer may not depend on DNA-protein contacts, but on indirect mechanisms such as protein-protein interactions. Indeed, the postulated recruitment of TFs to a DNA element has been observed in HOT DNAs. HOT DNAs, also called HOT regions, are a novel class of genomic DNAs found in nematode, fly, and human that function as tissue-specific enhancers [reviewed in (Farley and Levine 2012) and references therein]. The newly identified HOT DNAs in the *Drosophila* genome are characterized by their ability to recruit an average of 10 different TFs without enrichment for DNA motifs recognized by the corresponding TFs (modENCODE Consortium et al. 2010). It is conceivable that Sim protein can be recruited to the midline enhancer even without interaction with Sim-Tgo binding sites if the shadow enhancer can function as HOTA DNA. Consistent with this interpretation, *Drosophila* HOTA DNAs share certain sequence features, including GAGA elements (Kvon et al. 2012) and Zld binding motifs (Satija and Bradley 2012), which are present in the 0.68-kb minimal enhancer (Supplemental Table S4). In particular, the distal and proximal elements contain two GAGA elements and a Zld-binding motif, respectively (Supplemental Tables S3, S4). Thus, we speculate that binding of GAGA and Zld to the shadow enhancer may help decondense the local chromatin structure, thereby recruiting Sim protein to the enhancer region without direct DNA contact. However, there still exists a possibility that Sim protein is recruited to the shadow enhancer through non-canonical Sim binding sites instead of the three canonical Sim sites. It would be interesting to study how the shadow enhancer functions as a midline enhancer for the *sog* gene.

The *rhomboid* (*rho*) gene, like *sog*, is expressed in the neurogenic ectoderm and the ventral midline (Hong et al. 2013). The two *rho* enhancers that control the expression of this gene at different times in different tissues are discrete and autonomous units (Ip et al. 1992). Unlike the *rho* enhancers, the midline enhancer activity of the *sog* shadow enhancer could not be physically uncoupled from its neurogenic ectoderm enhancer (NEE) activity. Transgenic studies with lacZ fusions showed that the 3' ~ 500-bp region of the 0.68-kb construct closely recapitulated endogenous *sog* expression in the neurogenic ectoderm (Foo et al. 2014), suggesting that the minimal region of the *sog* NEE is present within the 3' ~ 500-bp fragment. If this is the case, about two thirds of the minimal midline enhancer (0.68-kb enhancer) overlaps with the minimal NEE. Consistently, simultaneous removal of the distal and proximal elements from the 0.68-kb construct led to concurrent loss of lacZ expression in the neurogenic ectoderm and the ventral midline (Supplemental Fig. S2). This suggested that the 0.68-kb enhancer was able to direct *sog* expression at different times and spaces. Further investigation of NEE activity within the 0.68-kb enhancer is required to precisely determine the minimal NEE and confirm this hypothesis.

Acknowledgments This work was supported by the National Research Foundation of Korea (NRF) Grants funded by the Korean Government (MSIP and MOE) (NRF-2010-0002792 and NRF-2012R1A1 A2038502).

Conflict of interest The authors declare that there are no conflicts of interest.

References

- Ashe HL, Levine M (1999) Local inhibition and long-range enhancement of *dpp* signal transduction by *sog*. *Nature* 398:427–431
- Biehs B, Francois V, Bier E (1996) The *Drosophila short gastrulation* gene prevents *dpp* from autoactivating and suppressing neurogenesis in the neuroectoderm. *Genes Dev* 10:2922–2934
- Campos-Ortega JA, Hartenstein V (1985) The embryonic development of *Drosophila Melanogaster*. Springer, Heidelberg
- Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14:1188–1190
- Farley E, Levine M (2012) HOT DNAs: a novel class of developmental enhancers. *Genes Dev* 26:873–876
- Foo SM, Sun Y, Lim B, Ziukaite R, O'Brien K et al (2014) Zelda potentiates morphogen activity by increasing chromatin accessibility. *Curr Biol* 24:1341–1346
- Francois V, Solloway M, O'Neill JW, Emery J, Bier E (1994) Dorsal-ventral patterning of the drosophila embryo depends on a putative negative growth factor encoded by the *short gastrulation* gene. *Genes Dev* 8:2602–2616
- Holley SA, Jackson PD, Sasai Y, Lu B, De Robertis EM et al (1995) A conserved system for dorsal-ventral patterning in insects and vertebrates involving *sog* and *chordin*. *Nature* 376:249–253
- Hong JW, Hendrix DA, Levine MS (2008a) Shadow enhancers as a source of evolutionary novelty. *Science* 321:1314
- Hong JW, Hendrix DA, Papatsenko D, Levine MS (2008b) How the dorsal gradient works: insights from postgenome technologies. *Proc Natl Acad Sci USA* 105:20072–20076
- Hong JW, Park KW, Levine MS (2013) Temporal regulation of *single-minded* target genes in the ventral midline of the *Drosophila* central nervous system. *Dev Biol* 380:335–343
- Ip YT, Park RE, Kosman D, Bier E, Levine M (1992) The dorsal gradient morphogen regulates stripes of *rhomboid* expression in the presumptive neuroectoderm of the *Drosophila* embryo. *Genes Dev* 6:1728–1739
- Kvon EZ, Stampfel G, Yanez-Cuna JO, Dickson BJ, Stark A (2012) HOT regions function as patterned developmental enhancers and have a distinct *cis*-regulatory signature. *Genes Dev* 26:908–913
- Liang HL, Nien CY, Liu HY, Metzstein MM, Kirov N et al (2008) The zinc-finger protein *zelda* is a key activator of the early zygotic genome in *Drosophila*. *Nature* 456:400–403
- Markstein M, Markstein P, Markstein V, Levine MS (2002) Genome-wide analysis of clustered dorsal binding sites identifies putative target genes in the *Drosophila* embryo. *Proc Natl Acad Sci USA* 99:763–768
- Menne TV, Luer K, Technau GM, Klambt C (1997) CNS midline cells in *Drosophila* induce the differentiation of lateral neural cells. *Development* 124:4949–4958
- modENCODE Consortium, Roy S, Ernst J, Kharchenko PV, Kheradpour P et al (2010) Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science* 330:1787–1797
- Nambu JR, Franks RG, Hu S, Crews ST (1990) The *single-minded* gene of *Drosophila* is required for the expression of genes important for the development of CNS midline cells. *Cell* 63:63–75
- Papatsenko D (2007) ClusterDraw web server: a tool to identify and visualize clusters of binding motifs for transcription factors. *Bioinformatics* 23:1032–1034
- Rubin GM, Spradling AC (1982) Genetic transformation of *Drosophila* with transposable element vectors. *Science* 218:348–353
- Satija R, Bradley RK (2012) The TAGteam motif facilitates binding of 21 sequence-specific transcription factors in the *Drosophila* embryo. *Genome Res* 22:656–665
- Schmidt I, Franzdottir SR, Edenfeld G, Rodrigues F, Zierau A et al (2011) Transcriptional regulation of peripheral glial cell differentiation in the embryonic nervous system of *Drosophila*. *Glia* 59:1264–1272
- Shimmi O, Umulis D, Othmer H, O'Connor MB (2005) Facilitated transport of a *dpp/scw* heterodimer by *sog/tsg* leads to robust patterning of the drosophila blastoderm embryo. *Cell* 120:873–886
- Small S, Blair A, Levine M (1992) Regulation of *even-skipped* stripe 2 in the *Drosophila* embryo. *EMBO J* 11:4047–4057
- Swanson HI, Chan WK, Bradfield CA (1995) DNA binding specificities and pairing rules of the *ah* receptor, ARNT, and SIM proteins. *J Biol Chem* 270:26292–26302
- Wharton KA Jr, Franks RG, Kasai Y, Crews ST (1994) Control of CNS midline transcription by asymmetric E-box-like elements: similarity to xenobiotic responsive regulation. *Development* 120:3563–3569
- Wieschaus E, Nusslein-Volhard C, Jurgens G (1984) Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster* III. Zygotic loci on the X-chromosome and fourth chromosome. Roux's Arch Dev Biol 193:296–307
- Wolpert L (1969) Positional information and the spatial pattern of cellular differentiation. *J Theor Biol* 25:1–47
- Wolpert L (2011) Positional information and patterning revisited. *J Theor Biol* 269:359–365
- Zeitlinger J, Zinzen RP, Stark A, Kellis M, Zhang H et al (2007) Whole-genome ChIP-chip analysis of *dorsal*, *twist*, and *snail* suggests integration of diverse patterning processes in the *Drosophila* embryo. *Genes Dev* 21:385–390
- Zhu LJ, Christensen RG, Kazemian M, Hull CJ, Enameh MS et al (2011) FlyFactorSurvey: a database of *Drosophila* transcription factor binding specificities determined using the bacterial one-hybrid system. *Nucleic Acids Res* 39:D111–D117
- Zinzen RP, Cande J, Ronshaugen M, Papatsenko D, Levine M (2006) Evolution of the ventral midline in insect embryos. *Dev Cell* 11:895–902