

Developmental Cell, Volume 47

Supplemental Information

**A Developmental Program Truncates Long Transcripts
to Temporally Regulate Cell Signaling**

Jeremy E. Sandler, Jihyun Irizarry, Vincent Stepanik, Leslie Dunipace, Henry Amrhein, and Angelike Stathopoulos

SUPPLEMENTAL INFORMATION

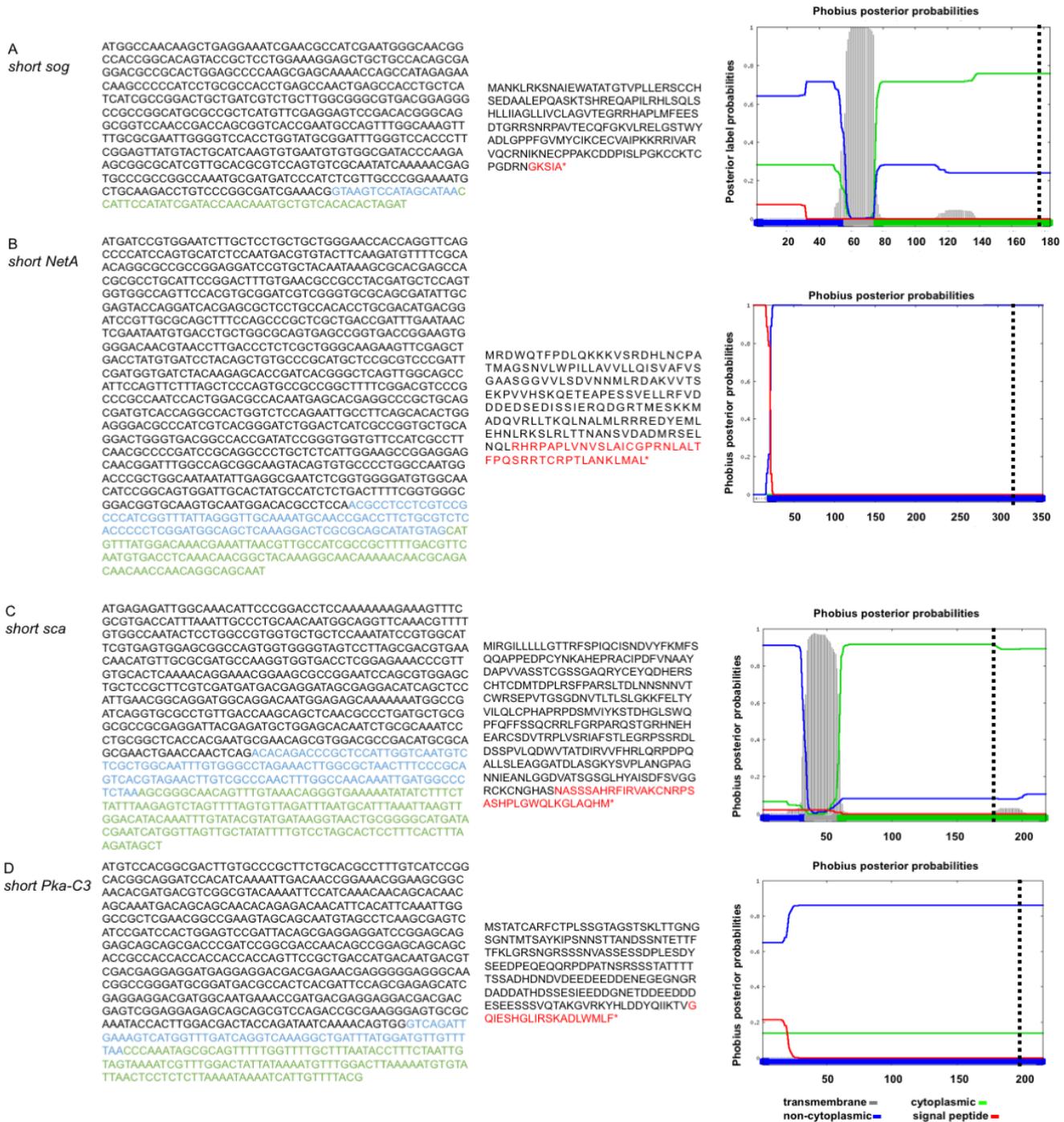
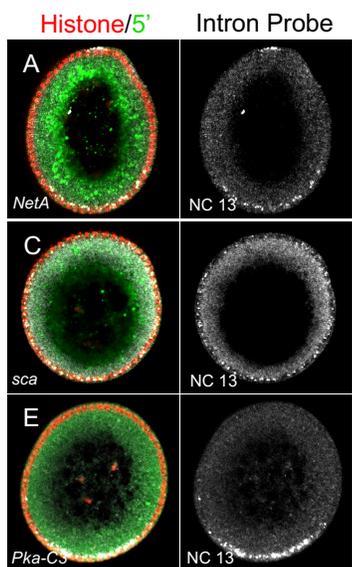
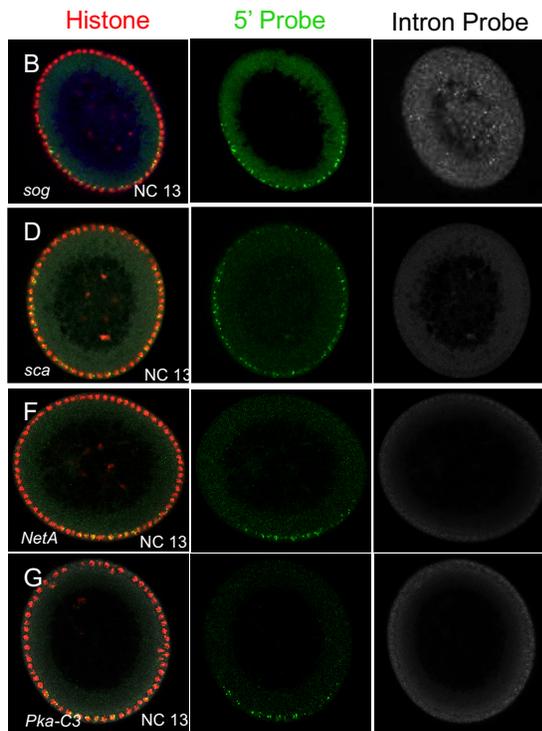
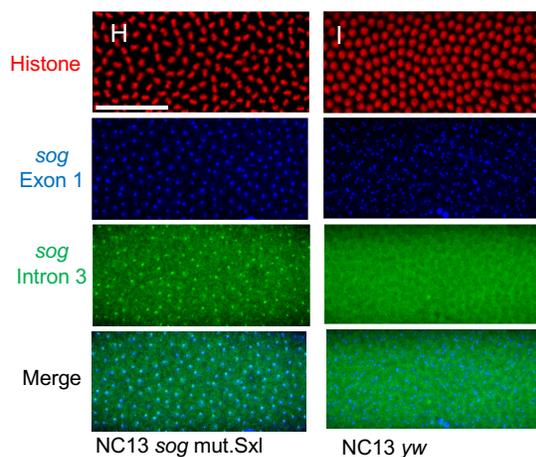


Figure S1. Short form sequences. Relates to Figures 1 and 4.

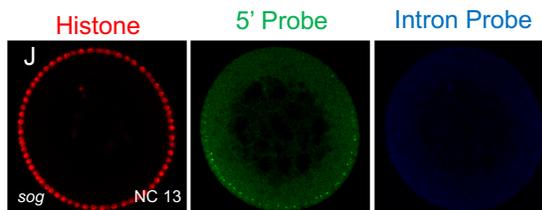
(A-D) Coding sequences, including new exons and 3' UTRs, of the four short forms (*sog*, *NetA*, *sca*, and *Pka-C3*) generated using 3' RACE. Black sequence is shared by both short and long forms, blue is new intron-derived coding sequence, and green is new 3' UTR. Amino acid sequences are also shown, with common sequence shared by full length and short form in black, and intron-derived amino acids predicted by 3' RACE in red. Plots show amino acid properties for short form proteins Sog, NetA, sca, and Pka-C3 respectively, with novel amino acids specific to the short forms after the dashed line. Blue indicates extracellular domain, green indicates cytoplasmic, red indicated signal peptide, and grey indicates transmembrane domain. Plots were generated using the Phobius tool for amino acid property prediction (Käll et al., 2004). Novel amino acids from the short forms retain the same or highly similar properties of preceding canonical sequence.



Hsp70-Gal4; UAS-sxl RNAi



Hsp70-Gal4; sxl RNAi HS-



Hsp70-Gal4; HS+

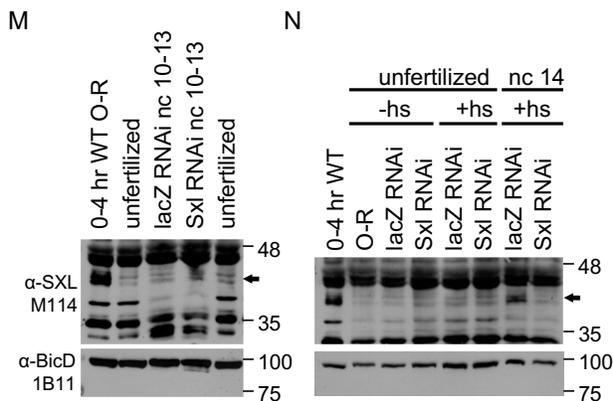
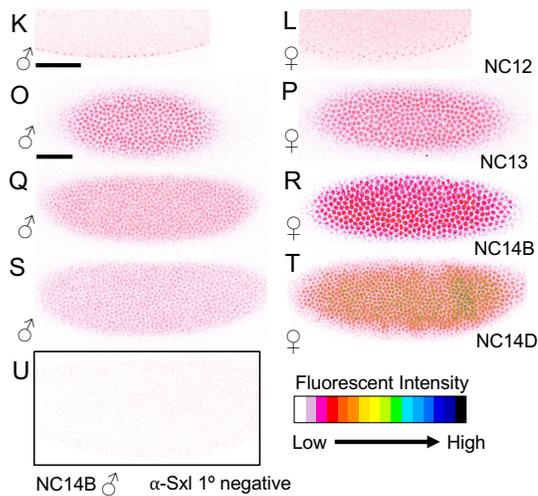


Figure S2. Transcriptional read-through in Sxl RNAi and temporal dynamics of Sxl protein expression in early embryos. Relates to Figure 2.

(A-C) Chopped embryo sections at NC13 showing RNAi against Sxl with FISH detecting the intron region downstream of the 3' RACE truncation point (white) for the genes *NetA* (A), *sca* (C), and *Pka-C3* (E). Histone H3 stained in red to stage embryos. RNAi leads to transcriptional read-through past truncation point.

(B,D,F,G) Chopped embryo sections at NC13 showing heat shock negative RNAi control for *sog* (B), *NetA* (D), *sca* (F), and *Pka-C3* (G) using 5' exon and intron probes.

(H,I) Multiplex FISH detection of *sog* Exon 1 (blue) and Intron 3 (green) with Histone H3 (red) for staging in wild type (i.e. *yw*) (H) or *sog mut.Sxl* (I) embryos of stage late NC13. When the four Sxl binding site matches in *sog* intron 3 are mutated, transcriptional read-through occurs past the site of truncation exhibited in wild type embryos at NC13. Scale bar is 50 μ m.

(J) Heat shock positive embryo containing only Hsp70-Gal4 driver. Read through into the intron of *sog* does not occur under heat shock only conditions without UAS-RNAi.

(K,L,O,P,Q,R,S,T) Embryos of indicated stages immunostained with α -Sxl antibody M114 assayed as male (K,O,Q,S) or female (L,P,R,T) sex using *sog* intronic riboprobe by in situ hybridization (as described for Fig. 2). Stainings show clear nuclear localization in both males and females, with intensity decreasing from NC14B to NC14D males while it increases during the same time period in female embryos. Number of embryos assayed as follows: K: 3, L:3, O:3, P:5, Q:6, R:2, S:4, T:8, and U:4. Scale bar is 50 μ m. NC12 embryos in K and L are from a different staining experiment than embryos O-T.

(M,N) Western analysis of extracts from staged embryos subjected to RNAi for Sxl using maternal heat-shock protocol (see Methods), or the indicated controls at specific stages. No loss of the bands seen in the MW range of zygotic Sxl is detectable in samples NC10-13 (left panel) or unfertilized eggs (right panel), but loss of zygotic Sxl in the NC14 sample is apparent (right panel). Extract equivalent to 16 embryos was loaded for all samples, except the 0-4hr wild-type (WT), which was loaded with 20 embryos.

(S) 2° antibody only control immunostaining without 1° α -Sxl, showing lack of nuclear background signal without α -Sxl antibodies.

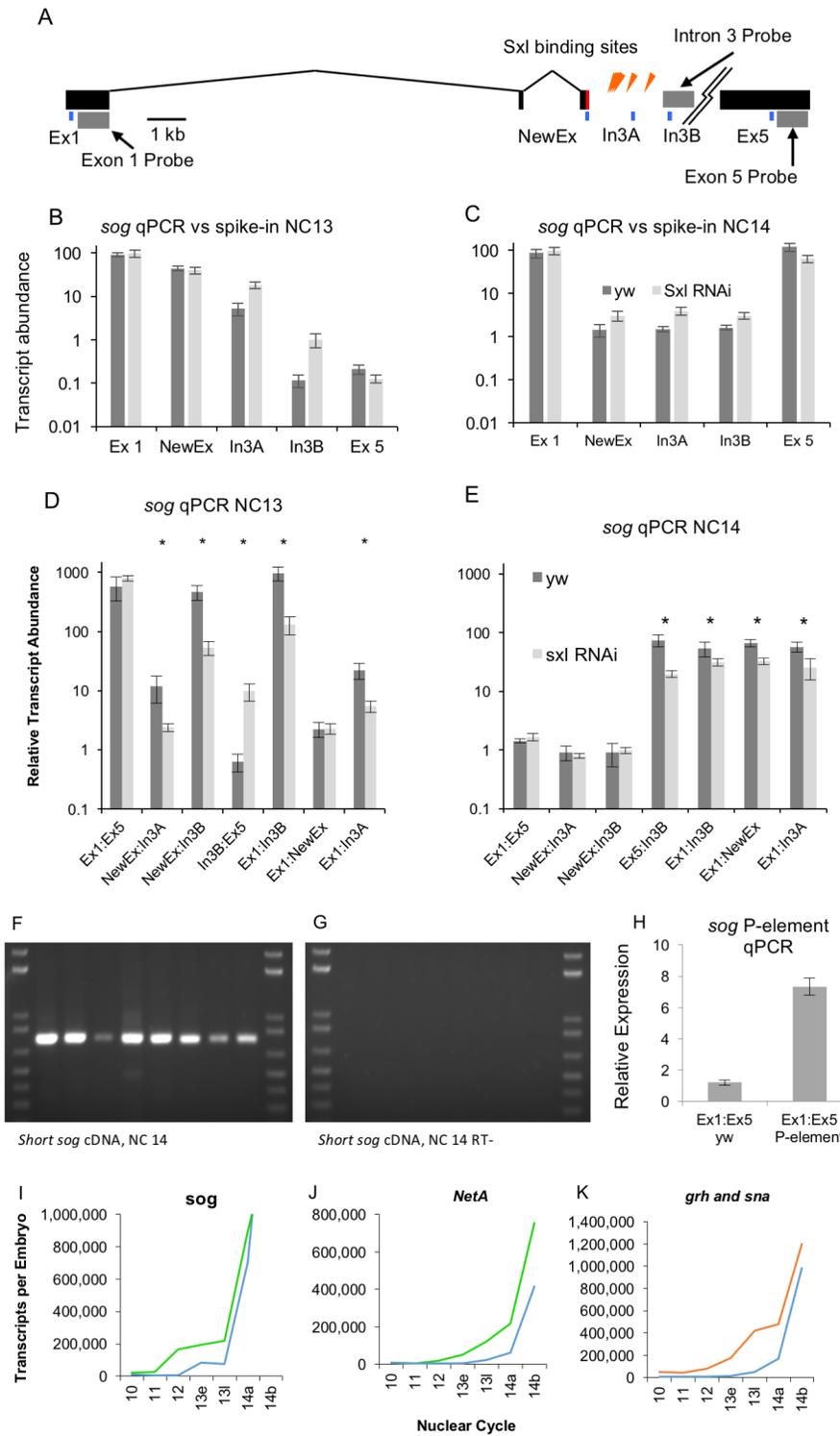


Figure S3. Experiments to assay levels of short and long transcripts. Relates to Figures 2 and 5.

(A) A diagram of the *sog* locus with in situ probe locations (grey boxes), qPCR primer sets (blue boxes), novel short form coding sequence (red box), and Sxl binding site locations (orange arrowheads).

(B,C) qPCR of RNA obtained from six individually assayed NC13 (B) or NC14 (C) embryos comparing expression levels using primers shown in (A, blue markers) to in vitro transcribed RNA spike-ins (GeneChip Eukaryotic Poly-A RNA Control Kit, Affymetrix, see Methods). Wild type embryos in dark grey, *Sxl* RNAi embryos in light grey.

(D,E) qPCR of RNA obtained from six individually assayed NC13 (D) or NC14 (E) embryos comparing expression levels between primer sets within the *sog* locus shown in (A, blue markers). Wild type embryos in dark grey, *Sxl* RNAi embryos in light grey. RNAi against *Sxl* leads to a read-through of intronic sequence past the truncation point, but still no full-length transcript. Data are presented as means \pm SEM. Asterisks signify $p < 0.0001$, by two-tailed Student's t-test. (D) During NC13, Ex1 is expressed ~ 600 -fold higher than Ex5 (Ex1:Ex5), indicating that Ex5 is not transcribed, consistent with FISH observations (see Figure 1C). Also during NC13, the novel coding region of truncated *sog* (NewEx) and the Sxl binding site cluster (In3A) are expressed ~ 2 -fold and ~ 20 -fold lower than Ex1 respectively, compared to a ~ 65 -fold decrease for the same sequences during NC14 (Ex1:NewEx and Ex1:In3A, Figures 3D and E), indicating this section of the intron is retained during NC13. There is a marked difference between the three intronic probes during NC13; probe In3B decreased ~ 500 -fold compared to NewEx, though they are equivalent during NC14 (NewEx:In3B, compare D and E). This decrease is similar to the difference between Ex1 and Ex5 (Ex1:Ex5), suggesting a truncation or lack of active transcription after the Sxl binding sites. In NC13, RNAi embryos retain intronic sequence past the Sxl binding site cluster, specifically primer set In3B, expressed ~ 10 -fold more

than wild type (Ex1:In3B, D *Sxl* RNAi). During NC13, Exon 5 was transcribed 600-fold less than the Exon 1 in *Sxl* RNAi embryos, indicating that the gene is not completely transcribed even when truncation is defective (Ex1:Ex5, D *Sxl* RNAi). (E) Within *sog*, the 5' and 3' exons expressed at approximately equivalent levels in NC14 (see Ex1:Ex5), while introns are expressed an average of ~65-fold lower, representing their excision (Ex5:In3B, etc.). *Sxl* RNAi also had an effect in NC14, with the intron spliced out at a slightly lower rate, but no differences between the primer sets within the intron (e.g. see Ex5:In3B and NewEx:In3A, Figure 3E *Sxl* RNAi).

(F,G) *short sog* transcript is detected by RT-PCR using cDNA from eight individual NC14 embryos. cDNA was transcribed using oligo-dT and then amplified using a reverse primer located in novel coding region of *short sog* (F). Reverse Transcriptase negative control on RNA from the same embryos, demonstrating that signal is cDNA-dependent (G). DNA Ladder is 1kb Plus Ladder (ThermoFisher).

(H) qPCR on *sog* comparing the relative expression of Exons 1 and 5 in wild type and the P-element insertion line. In *yw* embryos, Exons 1 and 5 are equivalently transcribed, but in the P-element insertion line, Exon 5 is transcribed 7-fold less than Exon 1.

(I-K) Quantification of transcripts (y-axis individual transcript counts per embryo) within single *Drosophila* embryos of the indicated stages (x-axis) using NanoString (Sandler and Stathopoulos, 2016b). Levels of expression for long genes *sog* (I) and *NetA* (J) were probed at the 5' (green) and 3' (blue) end of genes. *grh* (K) was identified as truncated only after 3' RNA-seq data was obtained, and therefore had only been probed by NanoString using a 3' probe (blue), and is compared to short gene *sna* (orange). Nevertheless, all three long genes show vast upregulation of expression of 3' ends from NC 13 late (13l) to early-mid NC 14 (14a and 14b,

respectively). The trajectories of short genes, 5' probes for *sog* and *NetA*, as well as *sna* are more similar. (See also Star Methods)

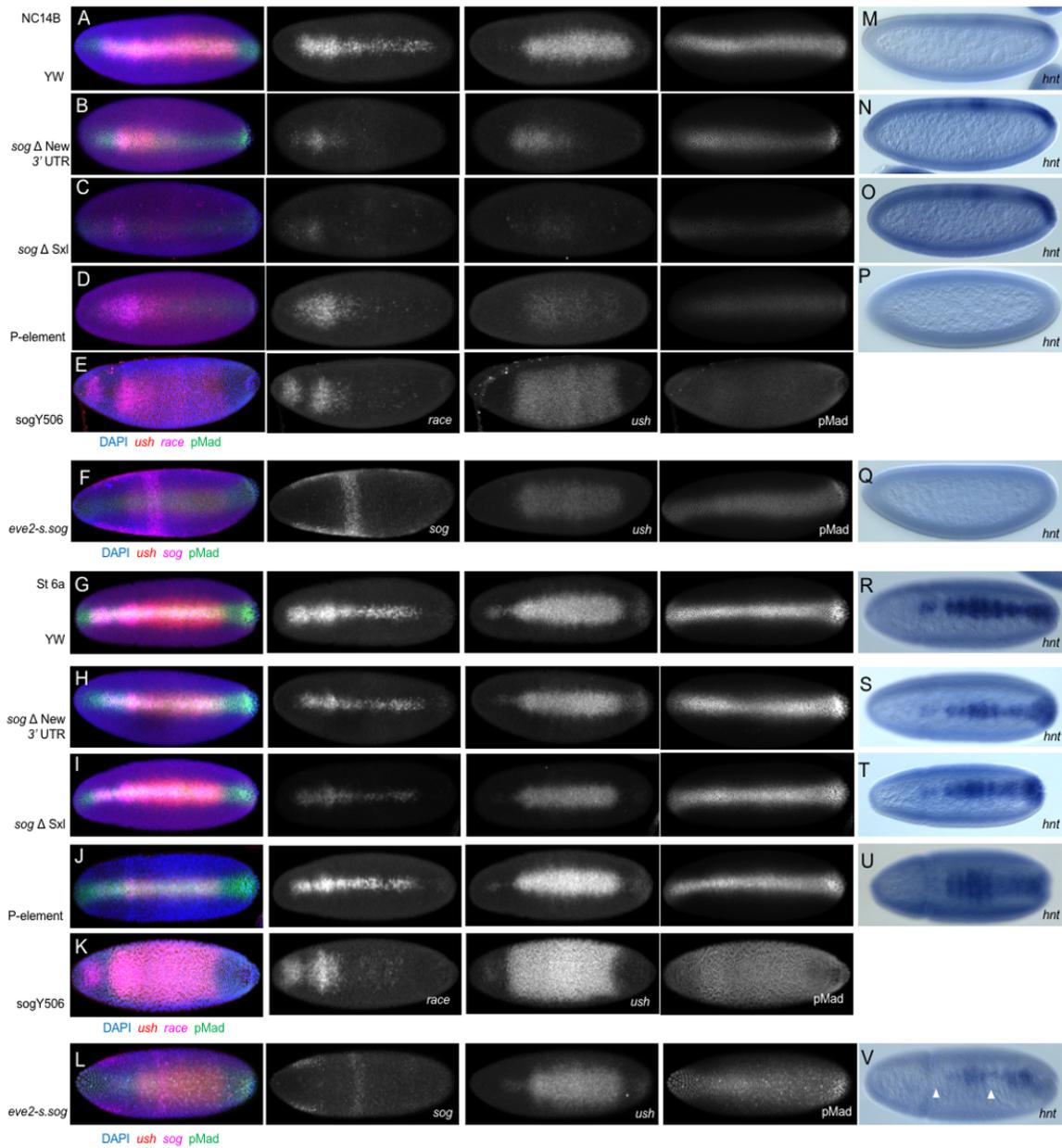


Figure S4. Short Sog supports proper spatiotemporal TGF- β signaling. Relates to Figure 3.

(A-L) *race* and *ush* gene expression patterns as well as pMad are affected by partial *sog* mutants in early nc14 embryos but partially recover by stage 6. Dorsal views of embryos at NC14B (A-F) and stage 6a (G-L) stained for *race* or *sog*, *ush*, and pMad. In *sog* Δ *New* 3' UTR (B) and *sog* Δ *Sxl* mutants (C), *race* and *ush* are retracted and weak with dots of ectopic expression throughout the embryo, and pMad is also weak compared to *yw* embryos (A). Expression

patterns return to the normal domain by late NC14 (H and I), but are somewhat weaker than *yw* embryos (G).

(M-V) *hnt* gene expression is also affected by *sog* mutants. Expression of the TGF- β target gene *hnt* in early NC14 (M) and late NC14 (R) embryos in *yw* embryos. Early expression in wild type is limited to a patch in the posterior dorsal ectoderm near the pole cells, and expression expands along the posterior $\sim 3/4$ of the dorsal ectoderm by late NC14. In *sog* Δ New 3' UTR embryos (N,S) and *sog* Δ *Sxl* embryos (O,T), *hnt* expression along the dorsal ectoderm is observed precociously in NC14 embryos (N,O) before it normally appears in wild type, at the same time *race* and *ush* have ectopic expression. *hnt* returns to mostly normal by late NC14 (S,T). In *sog* P-element embryos (P,U), *hnt* expression is not present at early NC14 (P), when the posterior patch first appears wild type (M), but expression is expanded laterally, as is seen in TGF- β target genes in *sog* hypomorphic mutants (U)(Ashe and Levine, 1999; Wharton et al., 2004). In embryos expressing *eve 2*-short *sog* (Q,V), *hnt* expression is not observed in early NC14 (Q), and is also decreased in late NC14 (V) with gaps in the expression pattern, as observed for *race* (Fig. 3H,I,O).

TABLE S2

Ac78C	CG7029	Mkp3	sdt
boi	dlg1	NetA	sog
btsz	eIF4G	oc	ssp3
CG17514	Fps85D	pan	Stlk
CG32556	grh	Pka-C3	toy
CG42314	L	sca	vn
CG42342	lt	scrib	zip
CG5594	Mbs	sd	

Table S2. Genes with putative short forms in 3' RNA-seq. Related to Figure 5.

A list of 31 manually annotated long genes expressed in 3' RNA seq data, in both NCs 13 and 14, with short forms.