EMBRYO DEVELOPMENT

A cysteine-clamp gene drives embryo polarity in the midge *Chironomus*

Jeff Klomp,[1] Derek Athy,[1] Chun Wai Kwan,[1] Natasha I. Bloch,[1]* Thomas Sandmann,[2]† Steffen Lemke,[1]‡ Urs Schmidt-Ott[1]§

In the fruit fly *Drosophila*, head formation is driven by a single gene, *bicoid*, which generates head-to-tail polarity of the main embryonic axis. Bicoid deficiency results in embryos with tail-to-tail polarity and no head. However, most insects lack *bicoid*, and the molecular mechanism for establishing head-to-tail polarity is poorly understood. We have identified a gene that establishes head-to-tail polarity of the mosquito-like midge, *Chironomus riparius*. This gene, named *panish*, encodes a cysteine-clamp DNA binding domain and operates through a different mechanism than *bicoid*. This finding, combined with the observation that the phylogenetic distributions of *panish* and *bicoid* are limited to specific families of flies, reveals frequent evolutionary changes of body axis determinants and a remarkable opportunity to study gene regulatory network evolution.

The *bicoid* gene of *Drosophila melanogaster* is involved in a variety of early developmental and biochemical processes. Many studies have examined its activity as a morphogen. *bicoid* mRNA is maternally deposited into the egg and transported to the anterior side, forming a protein gradient that activates transcription of genes in a concentration-dependent manner (*1*–*3*). The *bicoid* gene represents an intriguing case of molecular innovation. It is related to Hox-3 genes of other animals but appears to be absent in most insects, including mosquitoes and other "lower" flies (Diptera) (*4*–*6*) (Fig. 1). Bicoid-deficient embryos cannot develop a head or thorax and instead develop a second set of posterior structures that become a second abdomen ("double abdomen") when the activity of another gene, *hunchback*, is disrupted simultaneously (*7*). Likewise, ectopic expression of *bicoid* in the posterior embryo prevents abdomen development and induces a "double head" (*8*). Although other genes have been found to play a role in anterior development in beetles (*9*, *10*) and wasps (*11*, *12*), a gene responsible for anterior-posterior (AP) polarity has not been found. Nearly 30 years after the identification of *bicoid* in *Drosophila*, we have identified a gene that is necessary for the symmetry breaking and long-range patterning roles of *bicoid* in the harlequin fly *Chironomus riparius*. Further, we reexamined *bicoid* in several fly families and conclude that *bicoid* has been lost from genomes of some higher flies, including two lineages of agricultural and public health concern, the tephritid and glossinid flies (Fig. 1, figs. S1 and S2, and table S1). These observations raise the possibility that *bicoid* has been frequently lost or substantially altered during radiations of dipterans.

Ultraviolet light irradiation of anterior chironomid fly embryos induces double-abdomen formation, providing evidence of anterior localized RNA (*13*, *14*). Therefore, we conducted gene expression profiling of AP-bisected early *C. riparius* embryos to search for asymmetrically distributed maternal mRNA transcripts. All of the 6604 identified transcripts were ranked according to the magnitude of their differential expression scores and *P* values (Fig. 2A). Those most enriched in the posterior embryo were primarily homologs of known germ cell or germ plasm components (Fig. 2A, right). This was anticipated because the germ plasm of *Chironomus* is located at the posterior pole. One transcript was highly biased in the anterior end of the early embryo (Fig. 2A, left). We confirmed localized expression in early embryos for the two most biased transcripts (Fig. 2B and fig. S3).

The anteriorly biased transcript contains an open reading frame (ORF) encoding 131 amino acids. This predicted protein possesses a cysteine-clamp domain (C-clamp, residues 63 to 92) with similarity to the C-clamp of the Wnt signaling effector Pangolin/Tcf (Fig. 2C and fig. S4) (*15*) and was therefore given the name *panish* (for "pan-ish"). However, neither the high-mobility group (HMG) domain nor the β-catenin interaction domain of Pangolin is conserved in the protein sequence encoded by *panish*. Notably, we also identified a distinct *pangolin* ortholog expressed later in development during blastoderm cellularization at the anterior pole (fig. S5). Duplication of a portion of the ancestral *pangolin* locus is a possibility, given the strong similarity of their C-clamp domains. The *panish* C-clamp region appears to encode a bipartite nuclear localization signal (*16*); hence, *panish* may be involved in transcriptional regulation. The 5′ end of the *panish* transcript (27/131 predicted residues) overlapped with an unrelated *Chironomus* transcript with homology to *Drosophila ZAP3*, a conserved nucleoside kinase gene. We mapped all transcripts onto genomic *Chironomus* sequence containing *panish* and determined that *Chironomus ZAP3* (*Cri-zap3*) overlaps mostly with the large second *panish* intron (Fig. 2D) but was not differentially expressed between the anterior and posterior halves ($P = 0.34$).

The *panish* transcript was tightly anteriorly localized in freshly laid eggs but was expressed more broadly in an anterior-to-posterior gradient by the beginning of the blastoderm stage (Fig. 2B). The *panish* transcript was not evident after blastoderm cellularization. To test whether the *panish* transcript was necessary for the AP axis, we conducted a series of loss- and gain-of-function experiments using double-stranded RNA (dsRNA) and capped mRNA injections. Early *Chironomus* embryos injected with dsRNA against the *panish* ORF or 3′ untranslated region (3′UTR) developed double abdomens (Fig. 3, A to C, and fig. S6A), with similar survival rates between *panish* RNAi (RNA interference) and controls. Notably, *Cri-zap3* RNAi did not cause any obvious cuticle defects (Fig. 3C). Injection of *panish* dsRNA at the later blastoderm cellularization stage also had no effect, indicating that *panish* mRNA is dispensable at later stages ($N = 112/112$ wild type).

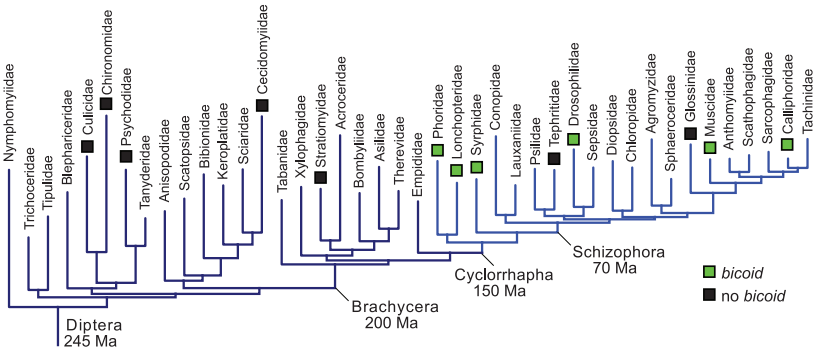


Fig. 1. *Bicoid* in dipteran families. Indicated instances of missing *bicoid* orthologs are based on genome sequences; tree is based on molecular phylogeny [see (*22*) and species list (*23*)], and cyclorrhapha clade, with *bicoid*, is indicated (light blue). Ma, millions of years ago.

[1]Department of Organismal Biology and Anatomy, University of Chicago, Chicago, IL 60637, USA. [2]Division of Signaling and Functional Genomics, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany.
*Present address: Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637, USA. †Present address: Department of Bioinformatics and Computational Biology, Genentech Inc., South San Francisco, CA 94080, USA. ‡Present address: Centre for Organismal Studies, University of Heidelberg, 69120 Heidelberg, Germany. §Corresponding author. E-mail: uschmid@uchicago.edu

To confirm the requirement for *panish* mRNA in establishing the anterior domain, we performed rescue experiments by co-injecting either wild-type or out-of-frame mutated *panish* coding mRNA in combination with *panish* 3′UTR dsRNA. Double-abdomen formation was suppressed in more than 40% of the embryos in which wild-type *panish* mRNA was injected into the anterior third, whereas double-abdomen formation was not suppressed after injection of mutated *panish* mRNA, injection buffer, or *bicoid* mRNA into the anterior third of the embryos (Fig. 3C and fig. S6, B to D) or *panish* mRNA injection into the posterior third of the embryos (130/131 wild type, $P < 0.0001$).

We also did not observe double-head formation (214/214 wild type) when we injected *panish* mRNA into the posterior third of wild-type embryos. This observation suggests that Panish activity is constrained to the anterior embryo, potentially because of missing anterior components or the presence of anterior program inhibitors in the posterior embryo. To distinguish between these possibilities, we examined the expression and function of genes associated with embryonic axis specification in other insects (*17*). Candidate genes included orthologs of the anterior inhibitor *nanos* (*Cri-nos*), the anterior pattern organizers *hunchback* (*Cri-hb*) and *orthodenticle/ocelliless* (*Cri-oc*), and the posterior pattern organizers *caudal* (*Cri-cad*) and *tailless* (*Cri-tll*).

Maternal *Cri-nos* transcript was enriched at the posterior pole (fig. S3), but neither *Cri-nos* RNAi nor ectopic *Cri-nos* expression affected axial patterning. *Cri-hb* and *Cri-oc* were present in the anterior blastoderm, but RNAi against these genes only caused homeotic and gap phenotypes, respectively (fig. S7, A to D). *Cri-cad* was expressed in the posterior embryo, and *Cri-cad* RNAi resulted in abdomen truncation (Fig. 4A and fig. S7E). Therefore, these genes do not appear critical for embryonic AP polarity. Unlike *Drosophila tailless*, *Cri-tll* was expressed in a posterior-to-anterior gradient in early blastoderm stages (Fig. 4B). After *panish* RNAi, both *Cri-cad* and *Cri-tll* were no longer expressed on one side, but instead were expressed symmetrically, consistent with their critical roles in abdomen development (Fig. 4, C and D). *Drosophila tailless* encodes a nuclear receptor required for terminal structures of the abdomen and brain development but not the AP axis (*18*). In contrast, *Cri-tll* RNAi embryos lacked tail segments, and ~70% of them developed malformed, often symmetrical, double heads with duplicated mandible and labrum structures and eye spots ($N = 45$; Fig. 4, E to G, and fig. S8). This result was confirmed in independent RNAi experiments using nonoverlapping dsRNAs (24/87 and 27/88 double heads).

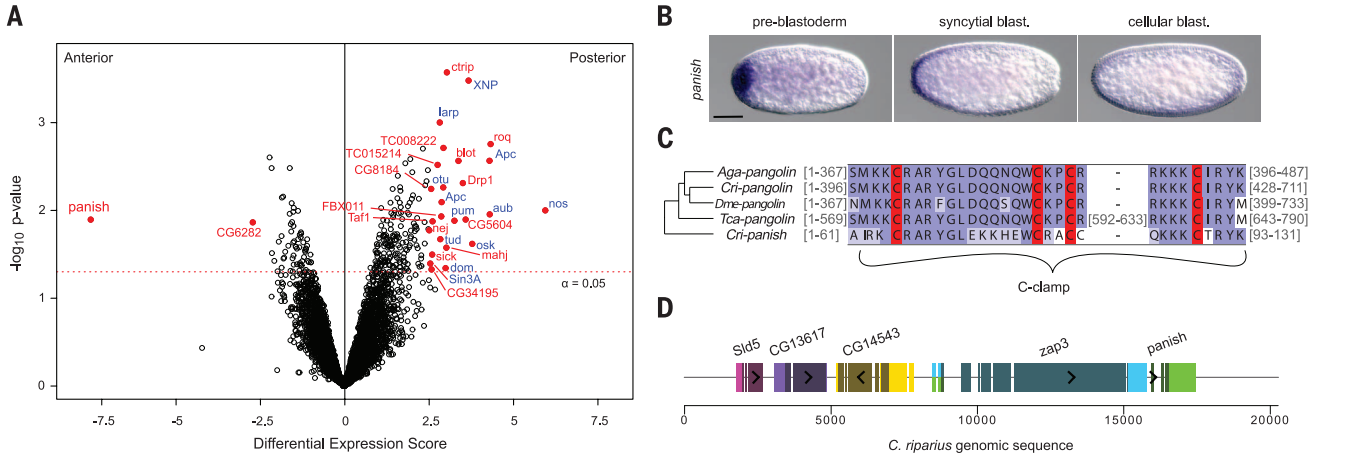
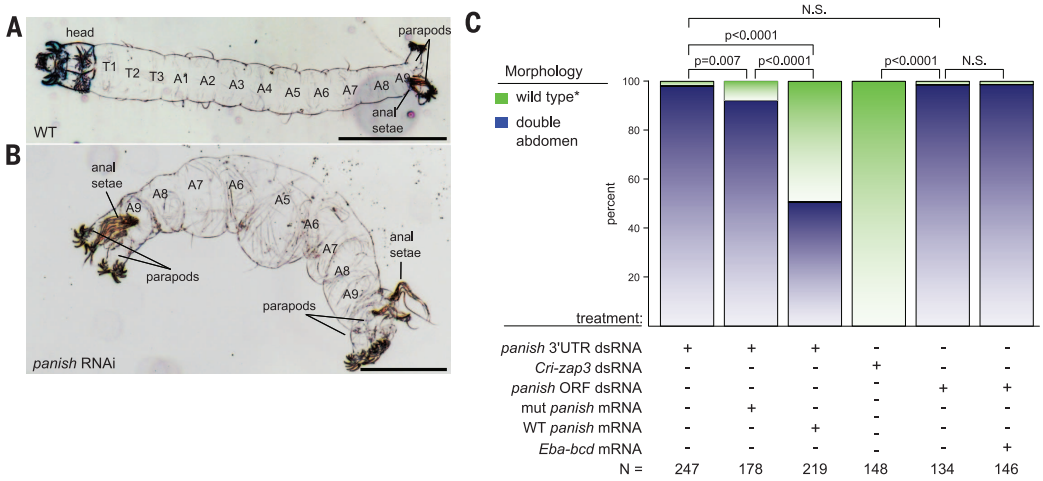


Fig. 2. *Panish* mRNA is enriched in the anterior embryo and encodes a C-clamp protein. (**A**) Differential expression of transcripts based on RNA sequencing data from anterior and posterior embryo halves. Red, score > |2.5| and $P < 0.05$; blue, putative germ cell or germ plasm components. (**B**) RNA in situ hybridizations of early embryos for *panish*. Anterior is to the left; scale bar, 10 μm. (**C**) C-clamp region of *panish* aligned with Pangolin sequences from *C. riparius* (*Cri*), *D. melanogaster* (*Dme*), *Anopheles gambiae* (*Aga*), and *Tribolium castaneum* (*Tca*). Gray numbers, residues not shown; red, conserved cysteine residues; blue, residue similarity. Amino acid abbreviations: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; W, Trp; Y, Tyr. (**D**) The *panish* locus. Homology is based on reciprocal-best-BLAST between *C. riparius* transcripts and *D. melanogaster* genes. Longest ORFs (shaded) and orientation (arrowheads) are indicated.

Fig. 3. *Panish* is required to establish AP polarity in *C. riparius*. (**A**) Inverted dark-field image of wild-type first-instar larval cuticle. A, abdominal segment; T, thoracic segment. (**B**) *Panish* RNAi cuticle of symmetrical double-abdomen larva. Scale bars, 30 μm. (**C**) Comparison of *panish* and *Cri-zap3* RNAi phenotypes and rescue of the *panish* RNAi phenotype by anterior injection of *panish* mRNA. *Note: In the third column, "wild type" includes 13 partial rescues (deformed head structures; fig. S6, B to D). N.S., $P > 0.05$; *Eba-bcd* from (*6*).



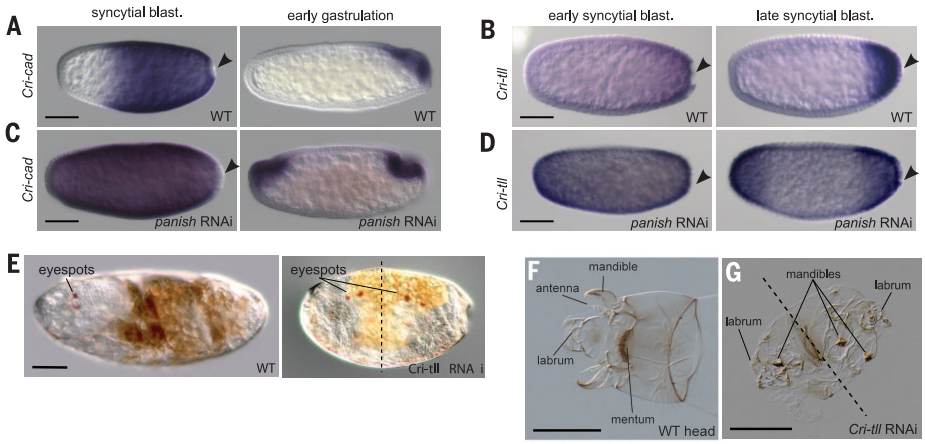


Fig. 4. *Cri-cad* and *Cri-tll* are regulated by *panish*, and *Cri-tll* is required to establish AP polarity. (**A** to **D**) Staining for *Cri-cad* [(A) and (C)] and *Cri-tll* [(B) and (D)] with RNA in situ hybridization in wild-type and *panish* RNAi embryos. Black arrowheads indicate posterior pole cells. (**E**) Eye spots indicated on live wild-type and *Cri-tll* RNAi embryos. (**F** and **G**) Cuticle preparations of a wild-type larval head (F) and a *Cri-tll* RNAi larva (G). Dashed lines indicate approximate plane of symmetry. Anterior is to the left in all panels; scale bars, 10 μm [except (G), 30 μm].

Duplication of head structures at both poles of the *Cri-tll* RNAi embryos suggests that unlike the *Drosophila* homolog, *Cri-tll* plays a role in AP polarity. Moreover, because *Cri-tll* is not expressed maternally, maternal *Cri-nos* would not appear to inhibit the formation of the anterior program. The observation that both heads of *Cri-tll* RNAi embryos develop with deformities is not surprising because *Drosophila tll* has a role in head development. *Cri-tll* may also play this part irrespective of its role in AP polarity.

A receptor tyrosine kinase gene, *torso*, controls the activation of *tailless* at the poles of the *Drosophila* blastoderm along with a second target, the zinc finger gene *huckebein* (*19*). We were unable to detect expression of the *Chironomus* homolog of *huckebein* (*Cri-hkb*) in early embryos (fig. S9A). *Cri-tor* was expressed zygotically at the poles of the blastoderm embryo (fig. S9B). Maternal *Cri-tor* transcript was detected in the RNA sequencing data but not by RNA in situ hybridization. *Cri-tor* RNAi caused tail deletions and head defects similar to those of *Cri-tll* RNAi embryos, but not double heads (fig. S9C). This finding suggests that *Cri-tll* has a role in axis polarity that is outside of its role in the terminal system driven by *torso*.

The double-head phenotype of *Cri-tll* RNAi embryos suggests a permissive role for *panish* in specifying embryonic AP polarity because the *panish* transcript was not detected in the posterior embryo. We suspected that *Cri-tll* might also have a permissive role in AP axis specification, because *Drosophila* Tailless functions as a dedicated repressor (*20*). This was confirmed by double RNAi experiments against *panish* and *Cri-tll* that resulted in perfect double abdomens (73/95 double abdomen, 20/95 intermediate, 2/95 wild type; fig. S9D). However, it raises the question of why the default developmental program establishes a double abdomen and not a double head.

One possible explanation is that Panish functions as a direct activator of head genes. This would imply that there is Panish activity in the posterior, but it seems unlikely given the lack of detectable mRNA in the posterior of embryos in both the RNA sequencing data and in situ hybridizations and the inability of *panish* mRNA to rescue the *panish* RNAi phenotype when *panish* mRNA is injected into the posterior third of the embryo. A more cogent possibility is that *panish* protein is more effective in repressing posterior genes than *Cri-tll* is in repressing anterior genes. Knockdown of *panish* would therefore result in proportionally higher levels of posterior transcripts such as *Cri-cad* and consequently inhibit head formation. This interpretation is consistent with high penetrance of the double-abdomen phenotype after *panish* RNAi.

Our results show that *Drosophila bicoid* and *Chironomus panish* encode structurally distinct DNA binding domain proteins that play similar essential roles in establishing AP polarity of the primary axis. In each case, the protein is necessary for breaking the symmetry of the primary axis and, when inactive, results in duplication of the posterior domain. Bicoid is a transcriptional activator of anterior genes. However, Panish appears to be a repressor of posterior patterning genes (fig. S10A). Moreover, maternally expressed *nanos*, which inhibits anterior programming in the posterior *Drosophila* embryo (*21*), appears to be ineffective in this regard in *Chironomus*. Two pieces of evidence argue against the existence of an additional, maternally localized, instructive factor for anterior development like *bicoid* in *Chironomus*. First, *panish* was the only transcript found strongly enriched at the anterior pole. Second, factors required for head development were also present in the posterior pole of *Cri-tll* embryos. We did not find evidence of *panish* in other dipteran genomes, even though the locus is conserved in two closely related chironomid species, *C. tentans* and *C. piger* (fig. S10B and table S2). This suggests a recent origin of *panish*. Our study shows that mechanisms of AP patterning in insects are more labile than previously acknowledged. The functionally diverse primary axis determinants of fly embryos provide a remarkable opportunity for studying molecular innovations in the context of gene regulatory networks.

REFERENCES AND NOTES

1. H. G. Frohnhöfer, C. Nüsslein-Volhard, *Nature* **324**, 120–125 (1986).
2. W. Driever, C. Nüsslein-Volhard, *Cell* **54**, 83–93 (1988).
3. W. Driever, C. Nüsslein-Volhard, *Cell* **54**, 95–104 (1988).
4. M. Stauber, H. Jackle, U. Schmidt-Ott, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 3786–3789 (1999).
5. M. Stauber, A. Prell, U. Schmidt-Ott, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 274–279 (2002).
6. S. Lemke *et al.*, *Development* **137**, 1709–1719 (2010).
7. W. Driever, in *The Development of* Drosophila melanogaster, M. Bate, A. M. Arias, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1993), vol. 1, pp. 301–324.
8. W. Driever, V. Siegel, C. Nüsslein-Volhard, *Development* **109**, 811–820 (1990).
9. J. Fu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 7782–7786 (2012).
10. K. Kotkamp, M. Klingler, M. Schoppmeier, *Development* **137**, 1853–1862 (2010).
11. J. A. Lynch, A. E. Brent, D. S. Leaf, M. A. Pultz, C. Desplan, *Nature* **439**, 728–732 (2006).
12. A. E. Brent, G. Yucel, S. Small, C. Desplan, *Science* **315**, 1841–1843 (2007).
13. K. Kalthoff, *Dev. Biol.* **25**, 119–132 (1971).
14. H. Yajima, *J. Embryol. Exp. Morphol.* **12**, 89–100 (1964).
15. F. A. Atcha *et al.*, *Mol. Cell. Biol.* **27**, 8352–8363 (2007).
16. S. Kosugi, M. Hasebe, M. Tomita, H. Yanagawa, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 10171–10176 (2009).
17. M. I. Rosenberg, J. A. Lynch, C. Desplan, *Biochim. Biophys. Acta* **1789**, 333–342 (2009).
18. T. R. Strecker, J. R. Merriam, J. A. Lengyel, *Development* **102**, 721–734 (1988).
19. G. Brönner *et al.*, *Nature* **369**, 664–668 (1994).
20. E. Morán, G. Jiménez, *Mol. Cell. Biol.* **26**, 3446–3454 (2006).
21. R. P. Wharton, G. Struhl, *Cell* **67**, 955–967 (1991).
22. B. M. Wiegmann *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 5690–5695 (2011).
23. See supplementary materials on *Science* Online.

ACKNOWLEDGMENTS

Transcriptomic data are available at the National Center for Biotechnology Information Sequence Read Archive (PRJNA229141). We thank G. K. Bergtrom for *Chironomus* culture and equipment; T. Hankeln, F. Ripp, L. Wieslander, and N. Visa for chironomid sequence data; J. E. Klomp for technical advice and discussions; M. Ludwig, S. Lott, M. Kim, C. Avila, A. Barnett, E. Davydova, T. Uchanski, C. Martinez, and Y. Yoon for technical help; J. Reinitz, M. Long, P. Rice, P. Faber, A. Ruthenberg, and C. Martinez for discussions; and C. Ferguson and M. Feder for laboratory equipment. Supported by Cell Networks Cluster of Excellence grant EXC81 (T.S.), NIH grant 1R03HD67700-01A1, NSF grants IOS-1121211 and IOS-1355057, and the University of Chicago.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6238/1040/suppl/DC1
Materials and Methods
Figs. S1 to S10
Tables S1 and S2
References (*24*–*51*)

16 January 2015; accepted 27 April 2015
Published online 7 May 2015;
10.1126/science.aaa7105

A cysteine-clamp gene drives embryo polarity in the midge *Chironomus*

Jeff Klomp *et al.*

Science **348**, 1040 (2015);
DOI: 10.1126/science.aaa7105