A timer gene network is spatially regulated by the terminal system in the *Drosophila* embryo

Erik Clark*[1,2,3], Margherita Battistara[1,4], and Matthew A. Benton†[1,5]

[1]Department of Zoology, University of Cambridge, UK
[2]Department of Systems Biology, Harvard Medical School, USA
[3]Present address: Department of Genetics, University of Cambridge, UK
[4]Department of Physiology, Development and Neuroscience, University of Cambridge, UK
[5]Present address: Departmental Biology Unit, EMBL, Heidelberg, Germany

February 13, 2022

Abstract

In insects, patterning of the anteroposterior axis relies on exquisite coordination of segmentation gene expression in space and time. And yet, there is also wide variation across species in the timing of segmentation and the geometry of the embryonic fate map. The majority of described species exhibit sequential segmentation during posterior growth. Other species, such as the fly *Drosophila melanogaster*, specify almost all of their segments simultaneously prior to gastrulation. The dynamics of both sequential segmentation and simultaneous segmentation correlate with the spatiotemporal expression of the "timer" genes *caudal*, *Dichaete* and *odd-paired*, which are expressed sequentially within segmenting tissues. These genes have critical roles in the development of many species; however, even in the extensively studied *Drosophila* embryo, the regulatory interactions responsible for shaping their expression remain poorly understood. In particular, it is not known whether the timer genes cross-regulate each other, nor why they are differentially expressed between the broad region of the blastoderm that gives rise to the gnathal, thoracic and abdominal segments and the small posterior region that gives rise to the embryonic terminalia. In this work, we investigate these questions using multiplexed fluorescent in situ hybridisation and high resolution confocal imaging of wild-type and mutant *Drosophila* embryos. First, we re-examine segmentation gene expression in the posterior of the embryo, and discover that two parasegment-like boundaries are generated sequentially from the terminal region during germband extension and posterior growth. Next, we find that *caudal*, *Dichaete*, and *odd-paired* dynamically cross-regulate each other, and also that they are differentially spatially regulated by the posterior terminal genes. Finally, we formalise the regulatory network we inferred from our data as a logical computational model. Our model qualitatively recapitulates both wild-type development and the mutant phenotypes we examined. Our findings resolve a decades-long ambiguity over the number of segments formed during *Drosophila* embryogenesis. In addition, they reveal how dynamic spatial inputs from an extrinsic signalling centre modulate the intrinsic dynamics of a gene regulatory network to generate an essential developmental pattern.

Introduction

Across insects, a relatively conserved gene regulatory network including gap genes, pair-rule genes, segment-polarity genes, and Hox genes regionalises and segments the anteroposterior (AP) axis in the developing embryo (reviewed in Nasiadka et al., 2002; Hughes and Kaufman, 2002; Clark et al., 2019).

Within and across species, axis patterning depends on these network components being activated at the right times and in the right places. Firstly, maturation of any given segment involves a specific sequence of patterning steps that requires different subsets of segmentation genes to be activated at different times (Akam, 1987; Baumgartner and Noll, 1990; Schroeder et al., 2011; Clark and Akam, 2016). Secondly, the relative timing of patterning across the AP axis varies widely across insects, from predominantly sequential, germband-based patterning in the cricket *Gryllus bimaculatus* or the beetle *Tribolium castaneum*, to more-or-less simultaneous, blastoderm-based patterning in the fruit fly *Drosophila melanogaster* (reviewed in Davis and Patel, 2002; Clark et al., 2019). Understanding how segmentation gene expression is coordinated over time and across the embryo is therefore a question of significant developmental and evolutionary importance.

Previously, we have proposed that such coordination results from an underlying framework of "timer gene" (alternatively, "timing factor") expression, which broadly regulates segmentation gene expression in time and space (Clark and Peel, 2018; Clark et al., 2019). We identified the timer genes (not necessarily exhaustively) with the genes *caudal* (*cad*) (Mlodzik et al., 1985; Macdonald and Struhl, 1986), *Dichaete* (*D*) (Russell et al.,

*ec491@cam.ac.uk
†matthewabenton@gmail.com

1996; Nambu and Nambu, 1996), and *odd-paired* (*opa*) (Benedyk et al., 1994), all of which code for transcription factors. The expression dynamics of these genes correlate with the progression of segmentation: in *Drosophila* they are expressed sequentially within the blastoderm, while in *Tribolium* the same expression sequence occurs in cells emerging from the segment addition zone into the segmented germ band (Schulz et al., 1998; Copf et al., 2004; El-Sherif et al., 2014; Clark and Peel, 2018). In addition, the protein products of these genes are known to directly regulate many segmentation genes in *Drosophila* (Rivera-Pomar et al., 1995; Schulz and Tautz, 1995; La Rosee, 1997; Häder et al., 1998; Ma et al., 1998; Clark and Akam, 2016; Vincent et al., 2018; Soluri et al., 2020).

However, we currently do not understand how the timer genes themselves are spatiotemporally regulated within the embryo. This key question underpins many other important developmental and evolutionary problems. For example, understanding how the timer genes are expressed in a specific temporal sequence would help us to explain the stereotypical progression of segment maturation. And understanding how the timer genes are differentially regulated along the AP axis would help us to explain the different modes of segment patterning and the evolution of widespread heterochrony in embryonic development.

Here, we investigate these issues in the *Drosophila* embryo, exploiting the fact that segmentation in this model species is not quite so simultaneous as it is often described. Although most of the *Drosophila* blastoderm is patterned simultaneously before gastrulation, the most posterior part of the segmental ectoderm is not patterned until germband extension (Kuhn et al., 2000). This "tail" region (see note on terminology below) is located posterior to abdominal segment 8 (A8) and anterior to the prospective hindgut, and eventually gives rise to a set of ectodermal structures known as the embryonic terminalia (Turner and Mahowald, 1979; Sato and Denell, 1986; Jürgens, 1987). Consistent with the timer gene hypothesis, the tail exhibits *cad*, *D*, and *opa* expression dynamics which differ from those in the rest of the trunk (Macdonald and Struhl, 1986; Russell et al., 1996; Clark and Akam, 2016; Clark and Peel, 2018).

The patterning of the tail region is dependent on the posterior terminal system (reviewed in Perkins and Perrimon, 1991), and, in particular, on its downstream effector, Tailless (Tll) (Strecker et al., 1986; Pignoni et al., 1990). Tll has well-characterised effects on gap gene expression (Jaeger, 2011; Janssens et al., 2013), but its contribution to timer gene regulation is relatively unexplored. As a consequence, the specific regulatory interactions that mediate tail patterning remain unknown (Casanova, 1990; Wu and Lengyel, 1998; Smits and Shvartsman, 2020).

In this study, we first re-examine segment patterning dynamics within the *Drosophila* tail, showing that this region gives rise to two sets of parasegment-like boundaries after gastrulation. Second, we use multiplexed hybridisation chain reaction in situ hybridisation (HCR ISH) (Choi et al., 2016; Trivedi et al., 2018; Choi et al., 2018) to characterise the relative spatiotemporal expression dynamics of the timer genes in wild-type embryos, timer gene mutants, and terminal system mutants, inferring a network of gene interactions that pattern *cad*, *D*, and *opa* expression in space and time. Finally, we formalise these interactions as a logical model, and show that it explains the essential features of posterior terminal patterning, including recapitulating mutant phenotypes. We conclude by discussing this newly-fleshed out patterning system within its evolutionary context.

A note on terminology

Morphological segments are offset from the initial metameric subdivisions of the embryo, the **parasegments**, by about 2/3 of a segment repeat (Martinez-Arias and Lawrence, 1985; Lawrence et al., 1985; Ingham et al., 1985; also see Figure 1C). The nth **parasegment boundary** (PSBn) refers to the anterior boundary of parasegment n.

Segment-polarity stripes are conventionally numbered according to the parasegment they are located within (Baker, 1987; also see Figure 1A,C). Thus, the first *en* stripe is en1, because it marks the anterior of parasegment 1, and the 14th *en* stripe is en14. The first *wingless* (*wg*) stripe, expressed just anterior to en1, is wg0, and the 14th wg stripe, expressed just anterior to en14, is wg13.

The term **telson** has been used to refer to the posterior region of the *Drosophila* embryo/larva (usually everything posterior to A8, sometimes everything posterior to A7) (Lohs-Schardin et al., 1979; Sato and Denell, 1986; Nusslein-Volhard et al., 1987; Perkins and Perrimon, 1991). As "telson" generally refers to a terminal non-segmental region of an animal (Snodgrass, 1935), or at least its most posterior segment, it is non-standard to use this word to refer to a region that contains more than one segment. We therefore use the more neutral term **tail** (Jürgens, 1987) to refer to the region posterior to PSB14 and anterior to the hindgut.

Results

Two parasegment-like boundaries form sequentially from the *Drosophila* tail region after gastrulation

The *Drosophila* embryo is well-known for its simultaneous mode of trunk segmentation, in which a segmental pattern is laid down at the end of the blastoderm stage, just as the first morphogenetic movements (cephalic furrow formation, mesoderm invagination, and germband extension) are about to begin. 14 prospective parasegment boundaries appear at this stage, marked by segmental stripes of segment-polarity gene expression (DiNardo, 1985; Baumgartner et al., 1987; Baker, 1988; Lee et al., 1992; Grossniklaus et al., 1992). The last 11 of these boundaries correspond to the 3 thoracic and 8 abdominal denticle belts that are conspicuous features of

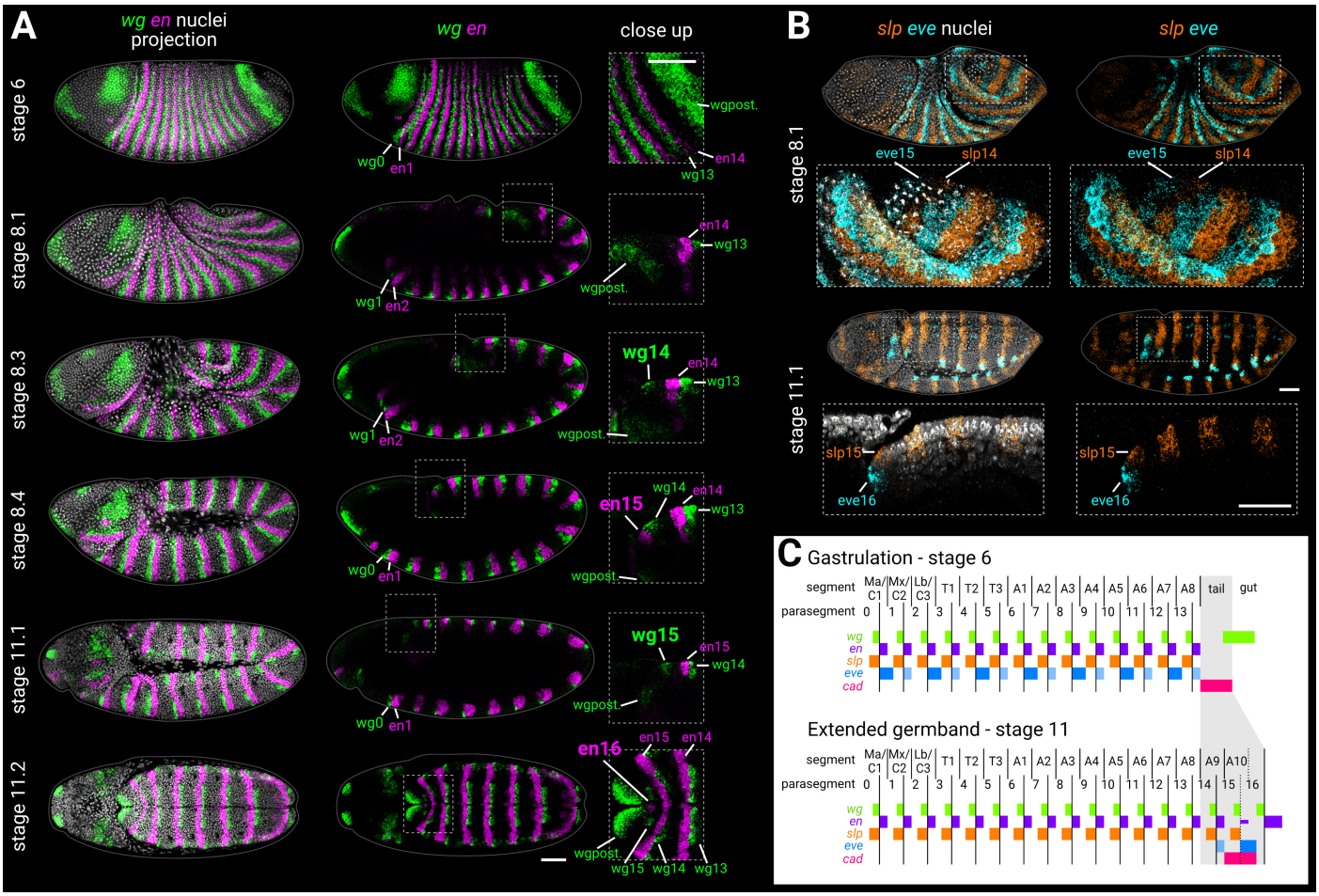

Figure 1: segmentation of the tail region after gastrulation. A: *wg* and *en* expression from gastrulation to extended germband. Left column shows merged maximum projections of *wg*, *en*, and DAPI. Middle column shows merged *wg* and *en* expression, either maximum projections (stage 6, stage 11.2), or sagittal sections (stage 8.1 to stage 11.1). Enlarged close-ups of the boxed regions are shown in the right column. Key expression domains are annotated with labels; newly-established domains are shown in large font; wgpost = *wg* posterior domain. Stages 6-11.1 show lateral views, stage 11.2 is a "dorsal" view which actually mainly shows the ventral side of the posterior germband, due to germband extension. **B**: *slp1* (*slp*) and *eve* expression during the division of mitotic domain 4 (stage 8.1) and at extended germband (stage 11.1). Both stages show dorsolateral views. Left column shows a merge with DAPI; right column shows gene expression alone. Enlarged close-ups of the boxed regions are shown below the whole embryo views; see Methods: Figure preparation for details of how the close-up for stage 11.1 was re-sliced. Key expression domains are annotated with labels. **C**: schematic diagram showing the expression of key segmentation genes before tail segmentation (stage 6) and after tail segmentation (stage 11). The tail region is shaded in grey; note the expansion of the region due to morphogenesis, and the refinement of the *cad* domain. PSB16 is shown as a dotted line due to its vestigial nature; en16 is also depicted as narrower than the other domains. Lighter shading for *eve* domains represents weaker or decaying expression. C1-3 = gnathal segments; T1-3 = thoracic segments; A1-10 = abdominal segments; Ma = mandibular segment; Mx = maxillary segment; Lb = labial segment. All embryos are anterior left, dorsal up. Scale bars = 50 μm; grey lines show embryo outlines.

the *Drosophila* larval cuticle and have been the target of several genetic screens (reviewed in St Johnston, 2002).

Sandwiched in between PSB14 and the broad posterior domain of *wg* (thought to correspond to prospective hindgut, Baker, 1988) are about 4 cell rows of ectodermal tissue, which remain unpatterned by segment-polarity genes at the end of the blastoderm stage (Figure 1A, stage 6). This "tail" region undergoes polarised cell divisions during germband extension (Foe, 1989; da Silva and Vincent, 2007) and later goes on to form the most terminal structures of the larva (Turner and Mahowald, 1979; Jürgens, 1987). These terminal structures include a 15th parasegment boundary (Kuhn et al., 1995; Kuhn et al., 2000), various sensory organs (Sato and Denell, 1986; Jürgens, 1987; Kuhn et al., 1992) and the anal pads (external organs involved in ion transport; Jarial, 1987).

The segmental nature of the tail tissue posterior to PSB15 (A10+) is unclear. No 16th parasegment boundary has been identified, but some authors consider this region to contain a cryptic 11th abdominal segment (Jür-

gens, 1987; Baumgartner et al., 1987; Gutjahr et al., 1993; Birkholz et al., 2013). To investigate this issue, we used multiplexed HCR ISH to re-examine the expression of the parasegment boundary markers *wg* (Baker, 1987) and *en* (Kornberg et al., 1985; Fjose et al., 1985) during germband extension and extended germband stages (Figure 1A; Supplementary Figure 1A).

Consistent with previous descriptions of gene expression, we observed that the *wg* and *en* stripes associated with PSB15 emerged during the fast phase of germband extension, with wg14 appearing slightly earlier than the posteriorly adjacent en15 (Figure 1A, stage 8.3-8.4). However, in contrast to published reports of *wg* expression (Baker, 1987; Baker, 1988), we also observed a further *wg* stripe, wg15, which appeared after germband extension and was located 1-2 cell rows posterior to en15 (Figure 1A, stage 11.1). During subsequent development, a small medial patch of *en* expression appeared posteriorly adjacent to the wg15 domain (Figure 1A, stage 11.2). This "en16" domain is clearly not a full stripe as found in parasegment boundaries 1-15. However, it corresponds to the median neuroblast lineage of abdominal segment 10 (Birkholz et al., 2013). As median neuroblasts always originate from posterior segment compartments (Bate, 1976; Doe, 1992; Biffar and Stollewerk, 2014), this implies it is located in the anterior of parasegment 16. wg15 and en16 could therefore correspond to a vestigial 16th parasegment boundary within the *Drosophila* embryo.

In the simultaneously-segmenting region of the embryo (here, termed the "trunk"), segment-polarity domains are patterned by stripes of pair-rule gene expression, which are established and refined prior to segment-polarity gene activation (DiNardo and O'Farrell, 1987; Jaynes and Fujioka, 2004; Clark, 2017). In particular, the odd-numbered parasegment boundaries are preceded by abutting stripes of *sloppy-paired* (*slp*) and *even-skipped* (*eve*) (Lawrence et al., 1987; Cadigan et al., 1994), an arrangement that is also found across a wide variety of sequentially segmenting species (reviewed in Clark et al., 2019). (Additional *slp* stripes abutting additional "minor" *eve* stripes later appear at even-numbered boundaries, as shown in Figure 1C, although these boundaries are initially patterned differently.) In the tail, PSB15 is prefigured by a new pair of stripes, slp14 and eve15, which appear after gastrulation but before the end of germband extension (Macdonald et al., 1986; Grossniklaus et al., 1992; Kuhn et al., 2000). (eve15 is described by some authors (e.g. Sackerson et al., 1999) as the "8th stripe" of *eve*, not counting the seven minor *eve* stripes that appear at even-numbered parasegment boundaries just before gastrulation.)

We re-examined the tail expression of *slp* and *eve* (Figure 1B; Supplementary Figure 1B), and found that slp14 and eve15 emerged simultaneously early in germband extension, at around the same time as the polarised cell divisions of mitotic domain 4 (Foe, 1989) (Figure 1B, stage 8). Later, after germband extension but before germband retraction, we found that an additional set of abutting *slp* and *eve* stripes, slp15 and eve16, emerged posterior to PSB15 (Figure 1B, stage 11.1). This finding increases our confidence that the wg15 and en16 domains, which we had observed in the same region, are segmental in nature. To the best of our knowledge, the slp15 domain has not been described previously. Persistent *eve* expression in a ring around the anal pads is well-known, although it has generally been described as a remnant of eve15 (Macdonald et al., 1986; Frasch et al., 1987; Sackerson et al., 1999; Kuhn et al., 2000) or, in one case, as a remnant of the 7th *eve* pair-rule stripe (Singer et al., 1996) rather than a separate domain.

In summary, we propose that two parasegment-like boundaries form sequentially from the tail region of the *Drosophila* embryo after gastrulation (Figure 1C). In both cases, segment-polarity gene expression is preceded by a template of abutting *slp* and *eve* expression, similar to the odd-numbered parasegment boundaries of the trunk. Unlike in the trunk, however, the resolved segmental *eve* stripes appear de novo and are not preceded by a pair-rule phase of expression.

Timer gene expression differs between the trunk and the tail

As described earlier, the timer genes (*cad*, *D* and *opa*) may regulate the differential patterning dynamics of simultaneous and sequential segmentation. To analyse this possibility in greater depth, we here focus on the expression and regulation of these genes in the trunk, tail, and posterior gut, from their earliest expression (stage 4) to early germband extension (stage 7). To aid with fine scale staging of embryos, we have divided stage 5, which lasts $\sim$40 minutes at 25° C, into 5 timeclasses based on the expression of several genes and the advancement of cellularisation (see Methods: Embryo staging and Supplementary Figure 2).

We previously surveyed timer gene expression in wild-type *Drosophila* embryos using enzymatically amplified double fluorescent ISH (Clark and Peel, 2018). Here, we refine that initial characterisation, by carrying out multiplexed HCR ISH and extracting quantitative AP expression traces from the resulting imaging data (Figure 2; see also Methods: Image analysis). Note that there are modest nuclear movements (of up to 20 μm / 3 cell diameters) away from the poles of the embryo over the course of the blastoderm stage (Keränen et al., 2006), which should be factored in when interpreting expression dynamics relative to the AP axis over time; more dramatic morphogenetic movements then occur at gastrulation. Accordingly, one particularly useful fiducial marker when assessing gene expression patterns in the posterior of the embryo is the posterior *wg* domain, which starts to be transcribed at stage 5.2, and remains around 3 nuclei wide throughout stages 5 and 6 (Figure 2; Supplementary Figure 2; Supplementary Figure 3). Once established, the posterior *wg* domain seems to be stable relative to nuclei, as nuclear transcription foci are not offset anteriorly or posteriorly relative to cytoplasmic transcripts. *wg* is therefore included as a co-stain in most of our experimen-

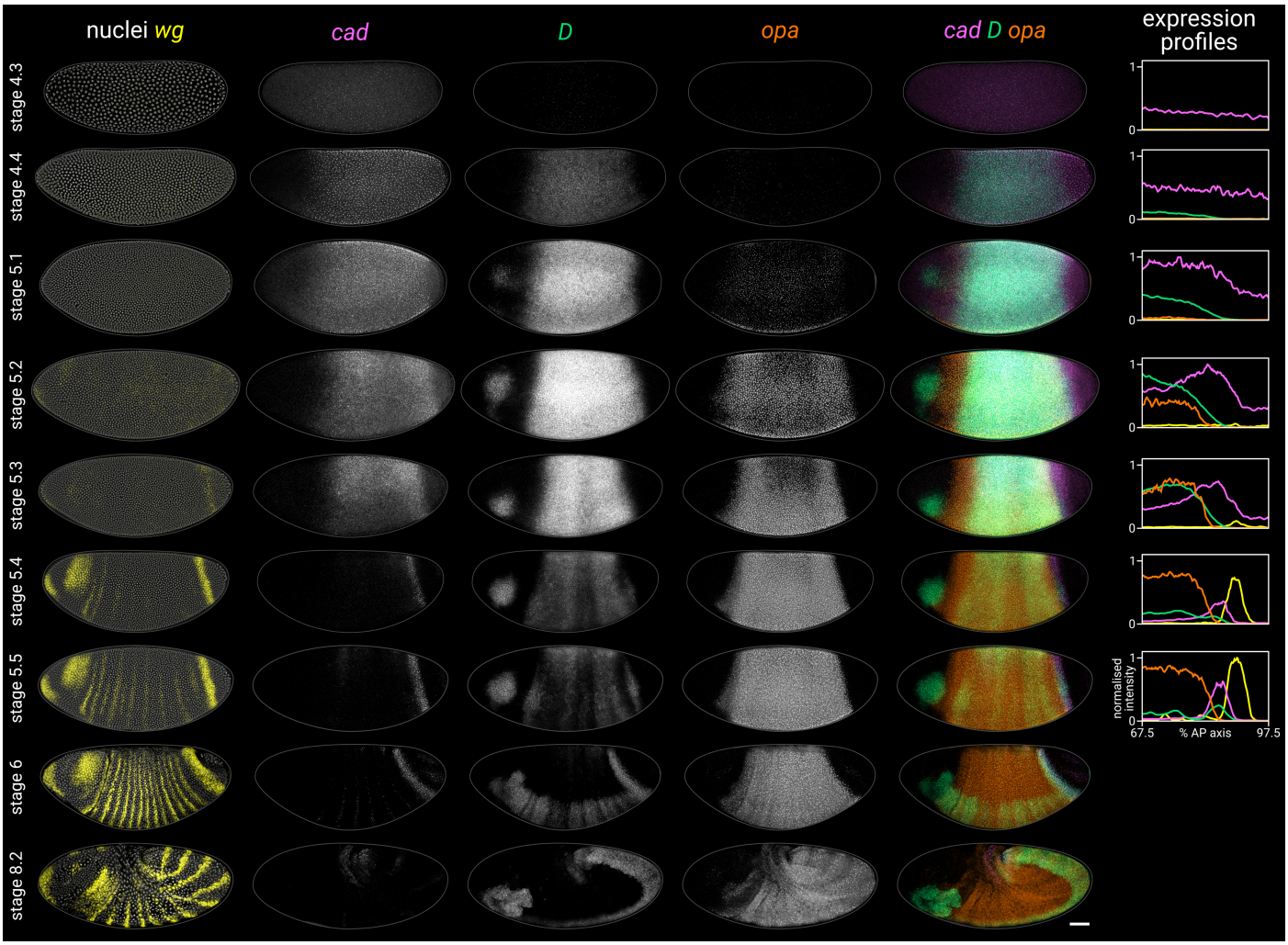

Figure 2: timer gene expression dynamics in wild-type embryos. Column 1 shows a two-channel *wg* and DAPI merge for embryos of gradually increasing age; columns 2-4 show *cad*, *D*, and *opa* channels from the same embryos; column 5 shows a three-channel *cad* / *D* / *opa* merge. The plots at the right show quantitative expression traces (67.5-97.5% AP axis; all measurements from the anterior pole) for all four genes, extracted from the embryos pictured to the left. The stage 4.3, stage 4.4 and stage 5.1 embryos are from a different scanning session compared to the rest of the figure. All embryos are anterior left, dorsal up. Stage 4.3-6 show lateral views; stage 8.2 is dorsolateral. Scale bar = 50 μm; grey lines show embryo outlines.

tal data.

In the trunk, *cad*, *D*, and *opa* transcripts are expressed sequentially over stages 4-6; first *cad*, then *D*, then *opa* (Figure 2; Supplementary Figures 4-6). Despite some intensity modulation across the AP axis (presumably downstream of gap and pair-rule genes), similar temporal dynamics are present across the whole region, consistent with its simultaneous mode of segmentation. *cad*, which is maternally deposited and then zygotically expressed (Levine et al., 1985; Mlodzik et al., 1985; Hoey et al., 1986; Macdonald and Struhl, 1986; Mlodzik and Gehring, 1987a; Schulz and Tautz, 1995), clears from the trunk by stage 5.4. *D*, which is detectable from stage 4.1 (nuclear cycle 10) (Russell et al., 1996; Nambu and Nambu, 1996), reaches appreciable levels at stage 4.4 (nuclear cycle 13), rapidly reaches a very high peak at stage 5.2, then declines sharply, with residual pair-rule modulated expression clearing by stage 6, replaced ventrally by persistent expression in the neuroectoderm. Finally, *opa* (Benedyk et al., 1994; Clark and Akam, 2016) appears at stage 5.1, rapidly builds to high levels, then tapers off during germband extension.

Antibody data, when compared to the transcript dynamics described above, show that Cad, D, and Opa protein dynamics reflect time-lags for synthesis and decay (Supplementary Figures 5,6). Cad levels decrease steadily in the trunk over stage 5 (Figure 2B in Surkova et al., 2008). D levels rise and fall gradually from stage 4.4 to stage 6, peaking at mid stage 5 (Supplementary Figure 5, Supplementary Figure 6C). Finally, Opa levels increase throughout stage 5 and into stage 6 (Supplementary Figure 6B; see also the live quantification of llama-tagged Opa in Soluri et al., 2020). Segmentation stages in the trunk are therefore characterised temporally by decreasing Cad levels, increasing Opa levels, and a pulse of D expression in between (Supplementary Figure 6D).

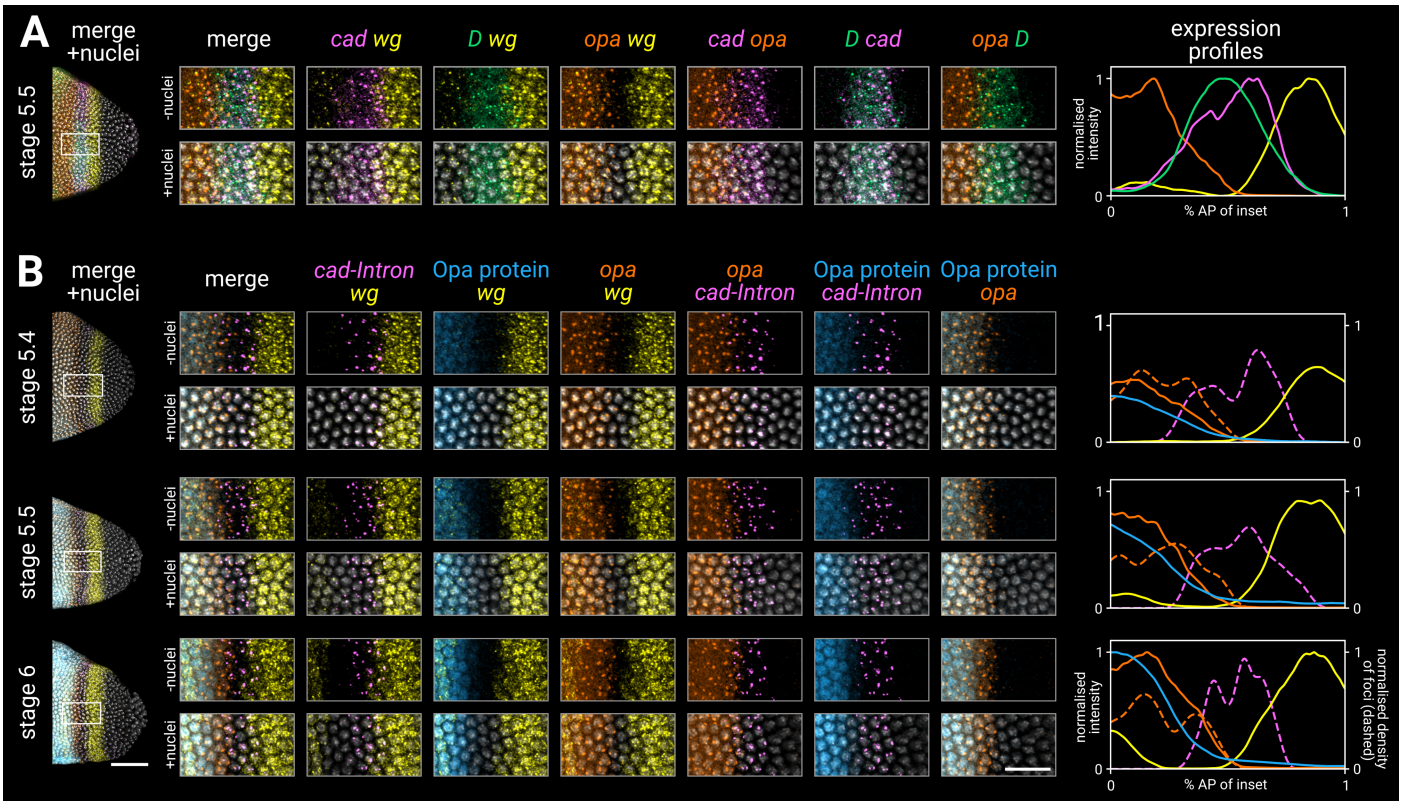

Figure 3: timer gene expression in the tail region of wild-type embryos at high resolution. **A**, **B**: leftmost column shows the posterior ends of the selected embryos, each with a boxed region of interest in the tail; middle columns show high resolution close-ups of the boxed region without and with DAPI signal ("-nuclei" vs "+nuclei"); rightmost column shows quantitative expression traces along the x-axis of the boxed region. **A**: timer gene expression, as in Figure 2. **B**: *wg* and *opa* expression as in A, combined with a *cad* intronic probe (*cad-Intron*, showing intranuclear transcription foci) and an antibody stain for Opa protein. Solid lines in the expression plots show the average intensity of *wg*, *opa*, and Opa protein; dashed lines show the normalised density of *cad* and *opa* transcription foci (see Methods: Image analysis). Note the staggered AP distributions of Opa protein, *opa* transcript, and *opa* transcription foci, the shrinking gap between the posterior *wg* domain and the *opa*/Opa signal, and the refinement of the *cad-Intron* domain over time. All embryos are anterior left, dorsal up, lateral view. Scale bars = 50 μm (embryo posteriors), 20 μm (boxed close-ups). For the high resolution close-ups, the curvature of the tissue was straightened prior to z-projection (see Methods: Figure preparation)

In the tail (see expression traces in Figure 2), a similar *cad*/*D*/*opa* expression sequence is evident, but markedly delayed. *cad* is expressed continuously in the tail region throughout stage 5 and into germband extension, long after it has been repressed in the trunk to the anterior and in the presumptive hindgut and midgut to the posterior. In contrast, *D* and *opa* expression in the tail region remains either low (*D*) or absent (*opa*) through most of stage 5. At stage 5.4, a *D* tail domain emerges within the lateral part of the *cad* tail domain, rapidly strengthening and extending dorsoventrally. D protein becomes prominent in the tail domain at stage 6 (Supplementary Figure 5, Supplementary Figure 6C), again reflecting a modest time-lag for protein synthesis. Finally, as will be described below, *opa* expression gradually expands posteriorly into the tail region from late stage 5 onwards, producing subtle shifting dynamics within the tail region that are most apparent at single cell resolution.

As shown in the high resolution close-ups in Figure 3 (see also Supplementary Figure 7 for single channels), the *cad* tail domain lies just anterior to the *wg* posterior domain, which it overlaps by one cell row (Figure 3A *cad*/*wg* merge). At stage 5.4, the domain of active *cad* transcription is 3-4 cells wide, but it shrinks to 2-3 cells wide by stage 6, with transcription ceasing at the anterior edge (*cad* intronic probe, Figure 3B). Throughout this period, the domain of active *opa* transcription, marked by prominent intranuclear foci, extends about one cell row posterior to the Opa protein domain (Figure 3B Opa/*opa* merge), and also overlaps the *cad* domain by about one cell row, reaching almost to the posterior *wg* domain by stage 6 (Figure 3A *cad*/*opa* merge, Figure 3B *opa*/*cad-Intron* merge). This suggests that *opa* transcription gradually invades the *cad* tail domain from the anterior edge, with *cad* transcription then ceasing in these cells as Opa levels increase (Figure 3B Opa/*cad-intron* merge). This posterior shift of the Opa posterior boundary can be seen directly in the movie of llama-tagged Opa mentioned above (Soluri et al., 2020).

In summary, we find that timer gene expression dif-

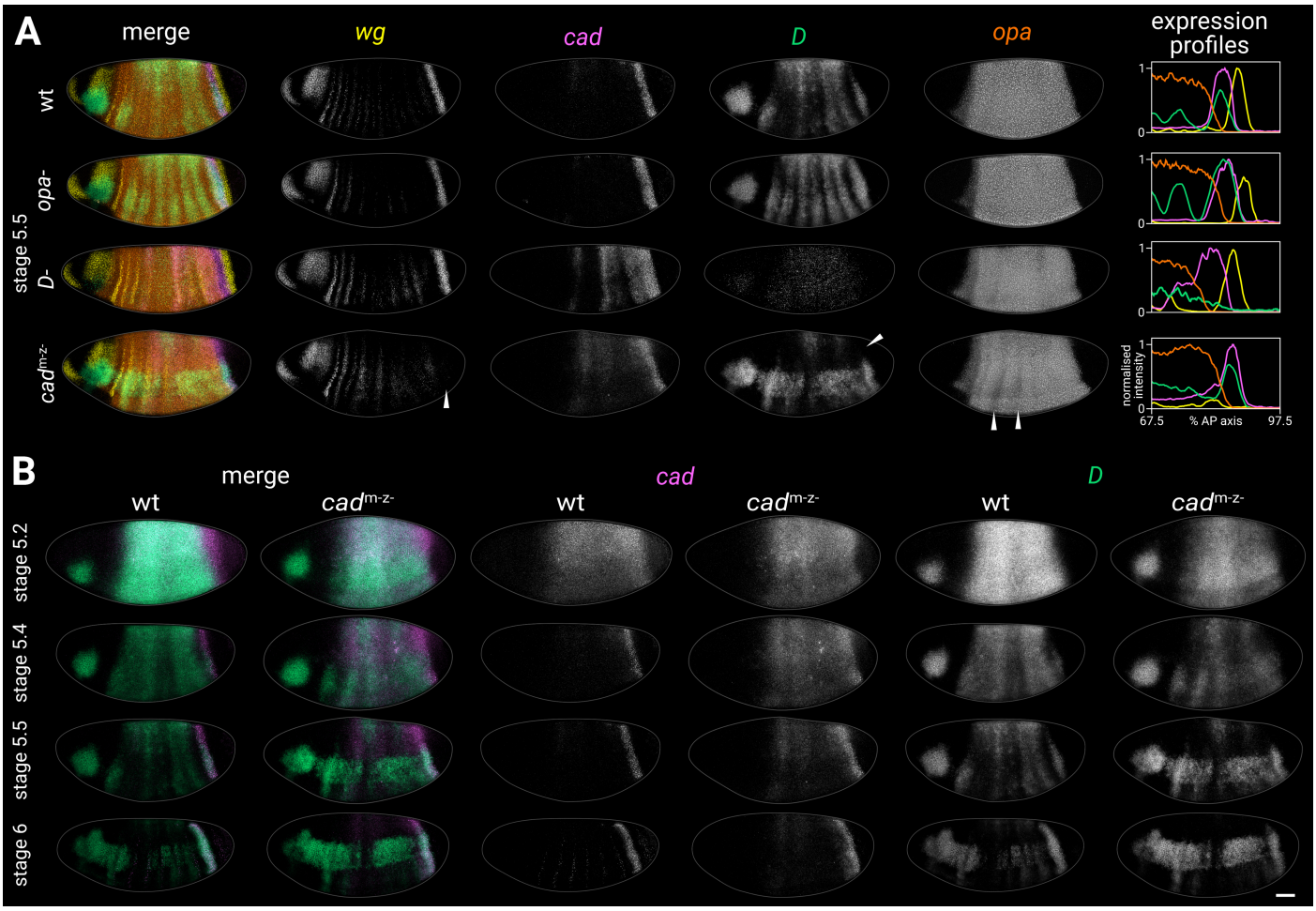

Figure 4: timer gene expression in timer gene mutants. **A**: Timer gene expression in wild-type, *opa*⁻ mutants, *D*⁻ mutants, and *cad*$^{m-z-}$ mutants at stage 5. The leftmost column shows a four-channel merge and the rightmost column shows quantitative expression traces (67.5-97.5% AP axis) from the embryos shown. In the *cad*$^{m-z-}$ embryo, note the absence of the *wg* posterior domain (arrowhead in *wg* channel), the dorsal loss of the *D* tail domain (arrowhead in *D* channel), and the anteroposterior modulation of the *opa* trunk domain (arrowheads in *opa* channel). The brightness and contrast of the *D* channel was adjusted for the *D*⁻ embryo to reveal the very weak residual signal. **B**: *cad* and *D* expression in wild-type and *cad*$^{m-z-}$ mutant embryos of gradually increasing age; left column shows a two-channel merge. In the *cad*$^{m-z-}$ embryos, note that *cad* transcript takes longer to clear from the trunk, while *D* is initially expressed at lower intensity and its neuroectodermal expression domain emerges earlier. All embryos are anterior left, dorsal up, lateral view. Scale bar = 50 μm; grey lines show embryo outlines.

fers sharply between the trunk and the tail, although both regions express *cad*, *D*, and *opa* in the same temporal sequence. The difference in timer gene expression between the trunk and the tail correlates with the difference in segmentation dynamics described in the previous section. Furthermore, the three genes' relative spatiotemporal dynamics are suggestive of cross-regulation.

The timer genes are patterned by cross-regulation

To determine whether the timer genes cross-regulate each other, we examined their expression in timer gene mutants. As *cad* is expressed both maternally and zygotically, we examined *cad* null mutants, *cad* maternal mutants, and *cad* zygotic mutants (denoted *cad*$^{m-z-}$, *cad*$^{m-z+}$, and *cad*$^{m+z-}$, respectively, see Methods: *Drosophila* husbandry and genetics), as well as *D*⁻ mutants and *opa*⁻ mutants. Because posterior Wnt signalling has strong effects on timer gene expression in sequentially segmenting species (reviewed in Williams and Nagy, 2017; Clark et al., 2019), we also checked timer gene expression in *wg*⁻ mutants, but did not observe any aberrant expression in these embryos during our stages of interest (Supplementary Figure 8).

Figure 4A compares timer gene expression between wild-type, *opa*⁻, *D*⁻, and *cad*$^{m-z-}$ embryos at stage 5.5; expression traces from additional embryonic replicates are shown in Supplementary Figures 9,10. Various changes to *cad*, *D*, and *opa* expression in the mutant genotypes indicate that the timer genes have cross-regulatory effects on each other, in addition to their previously characterised effects on segment patterning (Benedyk et al., 1994; Rus-

sell et al., 1996; Nambu and Nambu, 1996; Macdonald and Struhl, 1986).

In *opa*$^-$ mutants, *D* transcript did not clear from the trunk as quickly as normal, resulting in a more prominent stripy pair-rule pattern that persisted into stage 6, while the tail domain was stronger and wider (extending further anterior than normal) (Figure 4A, Supplementary Figure 9, and see Supplementary Figure 11, *opa*$^-$ for a stage 6 embryo). The *cad* tail domain looked similar to wild-type at stage 5.5 (Supplementary Figure 9), but was broader than normal at stage 6 (Supplementary Figure 11, *opa*$^-$), suggesting it failed to retract posteriorly as in wild-type. *opa* transcription and the posterior *wg* domain looked normal.

In *D*$^-$ mutants, there were changes to the relative expression of specific genes and also a modest anterior translation of the entire tail pattern. *cad* expression persisted abnormally in the posterior trunk, with marked anteroposterior modulation, and the *cad* tail domain was wider (again, extending further anterior than normal) (Figure 4A, Supplementary Figure 9). The *D* allele we used had very low transcript levels (presumably due to nonsense-mediated decay, S. Russell pers. comm.), but the residual expression indicated that both the clearance of *D* expression from the trunk and the appearance of the *D* tail domain may have been delayed. The posterior *wg* domain, the posterior border of the *cad* tail domain, and the posterior border of the *opa* domain were all anteriorly shifted relative to wild-type (Supplementary Figure 9A, Supplementary Figure 10); even after allowing for this shift, the gap between the *wg* domain and the *opa* domain was slightly larger in *D*$^-$ embryos than in wild-type (Supplementary Figure 9B).

Finally, in *cad*$^{m-z-}$ mutants, the AP positions of the timer gene tail boundaries were remarkably normal, but other aspects of gene expression were perturbed (Figure 4, Supplementary Figure 9A, Supplementary Figure 12A). *cad* expression cleared from the trunk more slowly than in wild-type, and was still visible at stage 6 (Figure 4B). *D* expression levels were much weaker than normal at early stage 5 (Figure 4B, stage 5.2), the *D* neuroectodermal expression domain appeared precociously (Figure 4B, stage 5.4), and the *D* tail domain was only expressed in the ventral half of the embryo, rather than forming a complete dorsoventral ring (arrowhead in Figure 4A). The posterior *wg* domain was generally absent from this genotype (arrowhead in Figure 4A), as previously described (Wu and Lengyel, 1998), although weak expression was observed in some embryos, in keeping with the marked variability of the *cad*$^{m-z-}$ phenotype at the cuticle level (Macdonald and Struhl, 1986). In addition, the *opa* domain showed strong pair-rule modulation in the anterior trunk (arrowheads in Figure 4A, Supplementary Figure 10).

Intrigued by the premature neuroectodermal expression of *D* in *cad*$^{m-z-}$ mutants, we examined the neuroectoderm patterning gene *muscle segment homeobox* (*msh*) (Lord et al., 1995) and found that it was also expressed prematurely in *cad*$^{m-z-}$ mutants, particularly in posterior parts of the embryo (Supplementary Figure 12A), raising the possibility that Cad normally inhibits neuroectoderm specification.

We additionally noticed that *cad*$^{m-z-}$ embryos appeared "wider" (larger in the y dimension) than wild-type embryos (Supplementary Figure 13). We found that the *cad*$^{m-z-}$ embryos were also shallower (shorter in the z dimension) than wild-type embryos (Supplementary Figure 13A,C), and often showed concave dips in their blastoderm surface (Supplementary Figure 13A). We hypothesise that *cad*$^{m-z-}$ embryos are more flaccid than wild-type embryos, and therefore become misshapen during fixation, staining and/or mounting for microscopy. This observation suggests that Cad, which is expressed in the *Drosophila* germline (Mlodzik and Gehring, 1987a), affects egg morphogenesis.

One copy of maternal *cad* (*cad*$^{m+z-}$ embryos) largely rescued the *cad*$^{m-z-}$ phenotype, except that the *D* tail domain was lost prematurely during germband extension (Supplementary Figure 12B). Of particular note, the posterior *wg* domain was present in this genotype, conflicting with a previous report (Wu and Lengyel, 1998).

One copy of zygotic *cad* (*cad*$^{m-z+}$ embryos) rescued the *D* tail domain fully, and partially rescued the *wg* posterior domain (Supplementary Figure 12C). The blastoderm dynamics of *D* and *cad* expression were still perturbed relative to wild-type (and *cad*$^{m+z-}$) embryos, however.

In summary, our investigation of timer gene mutant phenotypes provides strong evidence for timer gene cross-regulation. *cad* is derepressed in *D*$^-$ mutants, and *D* is derepressed in *opa*$^-$ mutants. *cad*$^{m-z-}$ embryos have a complex phenotype in which the early expression of *D* is quantitatively reduced, neuroectodermal gene expression is activated prematurely, the posterior *wg* domain is lost, and the *D* tail domain fails to activate dorsally. (Most of this phenotype, except the effect on the *D* tail domain, is mediated by maternal rather than zygotic Cad.) *opa* expression is fairly normal across all the mutants, except that its posterior border is anteriorly shifted in *D*$^-$ mutants.

This set of phenotypes, in combination with the expression dynamics described in the previous section, suggests that Opa represses *D* and *cad*, D represses *cad*, and Cad activates *D* (see Supplementary Table 1 for detailed reasoning). In addition, Cad is required for the expression of posterior *wg*, and D has a modest but concerted effect on the entire posterior fate map.

Tll and Hkb expression dynamics correlate with timer gene patterning in the posterior of the embryo

Having found evidence for cross-regulatory interactions between the timer genes, which might help explain their temporal dynamics, we next wanted to understand why timer gene expression differs between the trunk, tail, and prospective gut regions; i.e., how the timer gene network is spatially regulated. We therefore examined how timer gene expression relates to the expression domains of the zygotic terminal system genes *tll* (Jürgens et al., 1984;

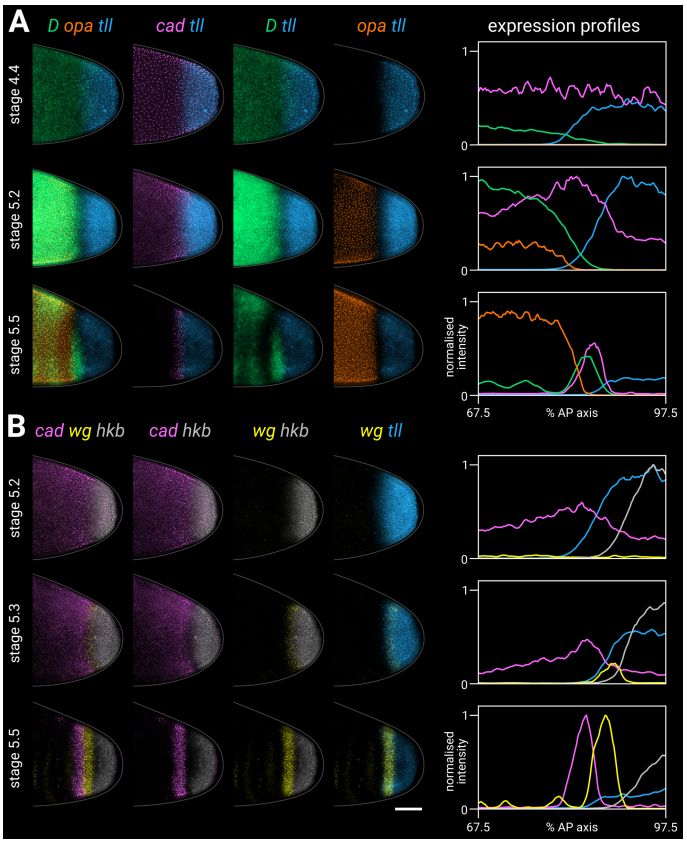

Figure 5: timer gene expression relative to posterior terminal gene expression in wild-type embryos. A, **B**: timer and terminal gene expression in embryos of increasing age; only the posterior end of each embryo is shown. Left columns show either three-channel or two-channel merges; right column shows quantitative expression traces (67.5-97.5% AP axis) of all four genes in the stain. **A**: timer gene expression relative to *tll*; note the posterior regression and changing intensity of the *tll* domain and the different spatial relationships with *opa*, *D*, and *cad*. **B**: *cad* and *wg* expression relative to *hkb* and *tll*; note how the posterior *wg* domain emerges within the *tll*-positive gap that opens up between *cad* and *hkb*. All embryos are anterior left, dorsal up, lateral view. Scale bar = 50 μm; grey lines show embryo outlines.

Strecker et al., 1986; Pignoni et al., 1990) and *hkb* (Weigel et al., 1990; Bronner and Jackle, 1991), the obvious candidates for providing this spatial information.

tll and *hkb*, which both code for repressive transcription factors, are expressed in nested domains at the posterior pole, with *tll* expression extending more anteriorly than *hkb* expression (Figure 5, Supplementary Figures 14-19; Pignoni et al., 1990; Bronner and Jackle, 1991). *tll* is transcribed at low levels from as early as nuclear cycle 9 (Pignoni et al., 1992), and we detected similar early transcription for *hkb*. Transcript levels in both domains peak at around stage 5.2 and then decline, with expression fading by stage 6 (Figure 5, Supplementary Figure 14, Supplementary Figure 16). The anterior border of the *tll* domain has previously been reported to be dynamic, retracting by about 5% egg length between stage 4.4 (nuclear cycle 13) and stage 5 (nuclear cycle 14) (Pignoni et al., 1990; Pignoni et al., 1992). Here, we noticed that this border also retracts by about 3-4 nuclear diameters over the course of stage 5 (Supplementary Table 2). Smaller magnitude shifts are apparent in the expression trace data (Supplementary Figure 16) due to the posterior flow of gene expression across nuclei being partially cancelled out by the anterior flow of nuclei away from the posterior pole (Keränen et al., 2006).

We also found that Tll and Hkb protein dynamics are spatiotemporally very similar to *tll*/*hkb* transcript dynamics, incorporating a slight time lag (Supplementary Figures 16-18), with the Tll protein border therefore lying slightly anterior to the *tll* transcript border during the second half of stage 5 (Supplementary Figure 17A). Note that our Tll antibody data closely resembles that collected by the Reinitz group, who noted that "in contrast to the posterior domains of the other gap genes, the [Tll] posterior domain does not shift position with time" (Surkova et al., 2008). We interpret the same data as providing evidence for a modest posterior retraction of the Tll domain over time, which does indeed contrast with the marked anterior shifts of the trunk gap genes, and is partially masked by anterior nuclear flow.

Through multiplexed HCR ISH, we found that the *tll* and *hkb* anterior borders correlate closely with the resolving expression boundaries of *cad*, *D*, *opa* and *wg* in the posterior of the embryo (Figure 5). At stage 4.4 (nuclear cycle 13), the graded *tll* border overlaps the graded posterior edge of the *D* domain (Figure 5A, top row). By mid stage 5, a narrow gap of low expression levels opens between the *tll* domain and the trunk domains of *D* and *opa* (Figure 5A, middle row), which is then filled by the *cad* and *D* tail domains at late stage 5 (Figure 5A, bottom row). *cad*, which is still expressed ubiquitously throughout the posterior of the embryo at stage 4.4 (Figure 5B, top row), fades from the *hkb* domain by mid stage 5 (Figure 5B, middle row), with a narrow gap of low expression levels opening up between the *cad* and *hkb* domains by late stage 5 (Figure 5B, bottom row). The *wg* posterior domain initiates at the border between *cad* and *hkb* expression present at mid stage 5 (Figure 5B, middle row), and by late stage 5 the *wg* posterior domain neatly demarcates the strip of *tll*-positive *hkb*-negative cells (Figure 5B, bottom row).

In summary, the spatiotemporal expression dynamics of Tll and Hkb make them good candidates for patterning the timer gene boundaries and the posterior *wg* domain, because they are differentially expressed across the various terminal regions. Specifically, the prospective posterior midgut experiences strong expression of both Tll and Hkb, the prospective hindgut experiences strong expression of Tll but weak/transient expression of Hkb, the tail region experiences weak/transient expression of Tll, and the trunk is consistently free of Tll and Hkb expression.

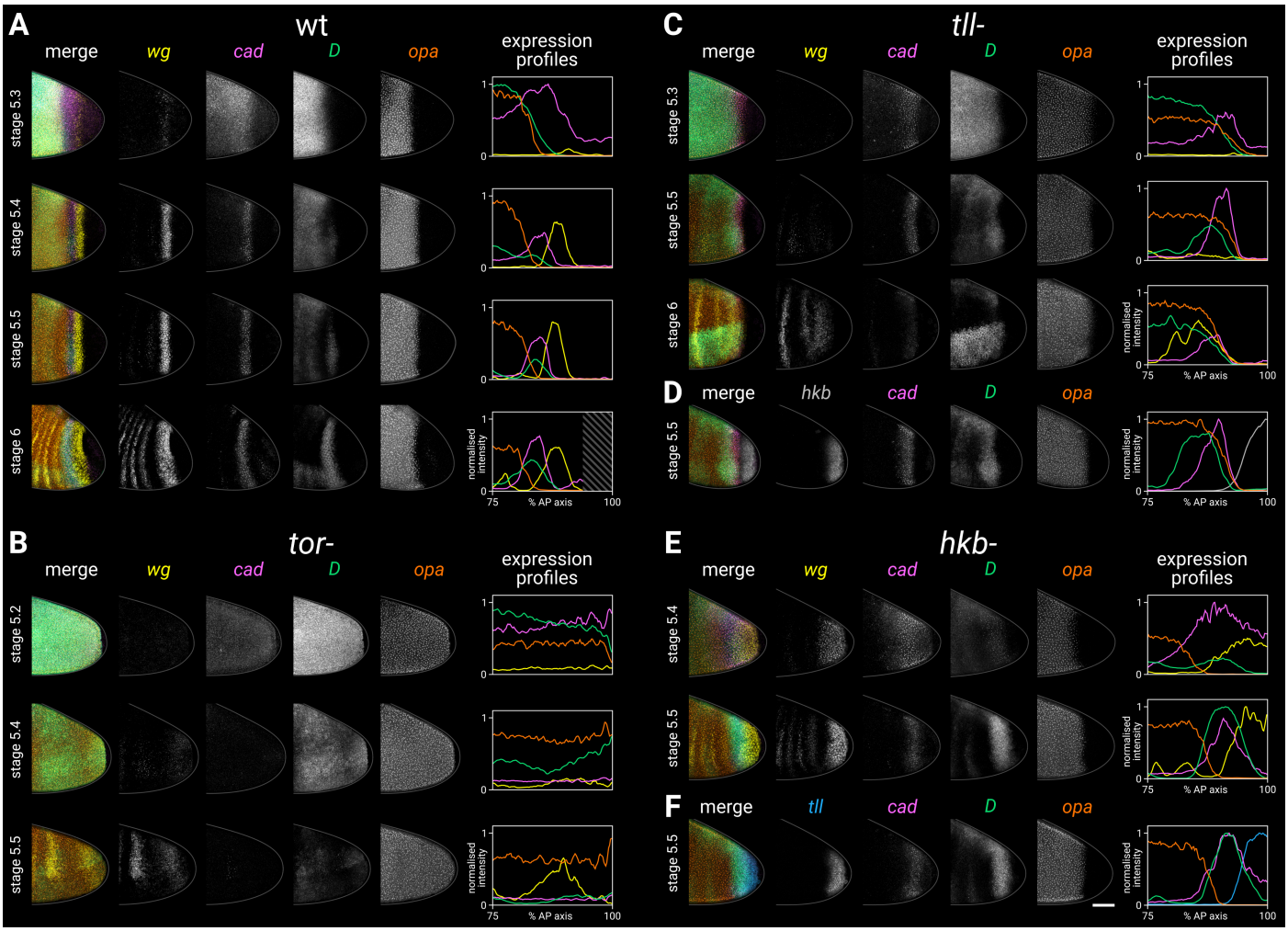

Figure 6: timer gene expression in terminal system mutants. A-F: gene expression in wild-type and mutant embryos of increasing ages. The leftmost column shows a four-channel merge; the middle columns show individual channels; the rightmost column shows quantitative expression traces (75-100% AP axis) from the embryos shown to the left. **A**: timer gene expression in wild-type. The AP axis is truncated in the expression plot for the stage 6 embryo (diagonally shaded area) due to proctodaeal invagination. **B**: timer gene expression in *tor*⁻ mutants. Note how the timer gene expression expands all the way to the posterior pole (excluding the pole cells). The broad posterior *wg* domain seen at stage 5.4-5.5 is mispatterned segmental expression; the posterior *wg* domain seen in wild-type embryos is absent. **C**: timer gene expression in *tll*⁻ mutants, relative to *wg* expression. Note that the *cad*, *D*, and *opa* domains share a similar posterior border, the fading of the *cad* domain over time, and the absence of the *wg* posterior domain. (Some mispatterned segmental *wg* expression is seen near the posterior of the embryo, similar to *tor*⁻ mutants). **D**: timer gene expression in *tll*⁻ mutants, relative to *hkb* expression. Note that the posterior borders of *cad*, *D* and *opa* all abut the *hkb* expression domain. **E**: timer gene expression in *hkb*⁻ mutants, relative to *wg* expression. Note that the repression of *cad* from the posterior pole is delayed, the posterior *wg* domain extends to the posterior pole, and the positions of all the expression domains are posteriorly shifted. **F**: timer gene expression in *hkb*⁻ mutants, relative to *tll* expression. Note how the *tll* domain is smaller than in wild-type (compare Figure 5A, stage 5.5) but preserves normal relationships with the *cad*, *D* and *opa* domains. All embryos are anterior left, dorsal up, lateral view. Scale bar = 50 μm; grey lines show embryo outlines.

The terminal system interacts with the timer gene network to pattern the posterior of the embryo

To determine whether Hkb and Tll spatially regulate the timer genes, we investigated timer gene expression in *hkb*⁻, *tll*⁻, and *torso* (*tor*⁻) mutants (Figure 6; expression traces from additional embryonic replicates are shown in Supplementary Figure 20). Tor (Klingler et al., 1988; Sprenger et al., 1989; Casanova and Struhl, 1989) is a maternally provided receptor necessary for transducing the extracellular signal-regulated kinase (ERK) signal that specifies the poles of the embryo (reviewed in Duffy and Perrimon, 1994; Li, 2005; Goyal et al., 2018); in *tor*⁻ mutants, neither *hkb* nor *tll* are expressed (Bronner and Jackle, 1991; Pignoni et al., 1992).

In *tor*⁻ mutants (Figure 6B; Supplementary Figure

20A), we found that all posterior spatial patterning of the timer genes was lost, and their temporal expression dynamics resembled those seen in the trunk of wild-type embryos. Thus *cad*, *D*, and *opa* were all expressed to the very posterior of the embryo at the beginning of stage 5, with first *cad* and then *D* expression turning off as stage 5 progressed. The posterior domain of *wg* was absent, and the region of segmental *wg* expression expanded posteriorly, as previously described (Mohler, 1995). The loss of the *cad* tail domain in *tor*$^-$ and *torso-like* (*tsl*$^-$) mutants has also been described previously (Mlodzik and Gehring, 1987b; Schulz and Tautz, 1995).

In *tll*$^-$ mutants (Figure 6C,D, Supplementary Figure 20A), the very posterior of the embryo remained clear of timer gene expression, but much of the normal spatial patterning anterior to this was lost. The posterior *wg* domain was absent, as previously reported (Wu and Lengyel, 1998). In addition, the *cad*, *D* and *opa* domains were expanded posteriorly, and all three genes shared a similar posterior boundary. We found that this common boundary abutted the posterior *hkb* domain, which looked similar to wild-type (Figure 6D, Supplementary Figure 21B). Normal expression of *hkb* in *tll*$^-$ mutants has been previously reported (Bronner and Jackle, 1991; Bronner et al., 1994; Ashyraliyev et al., 2009).

A posteriorly-shifted *cad* tail domain was transiently expressed in the *tll*$^-$ mutants, fading at stage 6 (Figure 6C, Supplementary Figure 20A). This finding conflicts with previous reports that the *cad* tail domain was either unaffected (Reinitz and Levine, 1990) or completely absent (Mlodzik and Gehring, 1987b) in *tll*$^-$ mutants. The AP modulated trunk pattern of *D* expression was abnormal, presumably caused by feedback from the segmentation genes, which are misregulated in *tll*$^-$ mutants (Mahoney and Lengyel, 1987; Casanova, 1990; Janssens et al., 2013). A persistent posterior *D* domain did not emerge (Figure 6C, stage 6, Supplementary Figure 11, *tll*$^-$). *tll*$^-$ *opa*$^-$ double mutants showed similar patterning dynamics to *tll*$^-$ single mutants, except that tail-like expression of *D* was rescued and persisted into germband extension (Supplementary Figure 11).

Finally, in *hkb*$^-$ mutants (Figure 6E,F, Supplementary Figure 20), the *wg* posterior stripe became a posterior cap, as previously reported (Mohler, 1995), and we also noticed that *cad* expression persisted longer than normal at the posterior pole, only clearing at the end of stage 5. Aside from these defects, timer gene expression on first inspection looked broadly normal, because the relative phasing of the *cad*, *D*, *opa* and *wg* domains was preserved.

However, comparing expression traces from *hkb*$^-$ mutant and wild-type embryos (Supplementary Figure 20) revealed that the whole terminal pattern was posteriorly shifted and expanded into embryonic territory that would normally express *hkb*. We examined *tll* expression in these mutants and found that the *tll* domain was markedly reduced in size, preserving the correlation between *tll* levels and timer gene expression boundaries that we described in the previous section (Figure 6F, Supplementary Figure 20B). This finding conflicts with previous reports that *tll* expression is unaffected in *hkb*$^-$ mutants (Bronner and Jackle, 1991; Bronner et al., 1994; Bronner and Jackle, 1996).

In summary, by examining *tor*$^-$ mutants, we found that all posterior spatial patterning of the timer genes is dependent on the terminal system. Expression boundaries associated with the tail and hindgut are perturbed in *tll*$^-$ mutants, while expression boundaries associated with the posterior midgut are perturbed in *hkb*$^-$ mutants. In addition, there is a concerted posterior shift of the fate map in *hkb*$^-$ mutants, which we attribute to the reduced size of the *tll* domain in this genotype.

Our observations from this section and the previous section suggest that Tll strongly represses *D* and *opa* and weakly represses *cad*, while Hkb represses *wg*, *cad*, *D*, and *opa* (see Supplementary Table 1 for detailed reasoning). Hkb is also necessary for activation of *tll* at normal levels, an interaction that is presumably indirect, since Hkb acts as a repressor (Goldstein, 1999).

Fkh demarcates the tail / hindgut border and activates posterior *wg*

Having found that Tll is necessary for patterning both the tail region and the posterior *wg* domain (prospective hindgut), we next wanted to understand how these regions are distinguished from each other. Therefore, we turned our attention to Forkhead (Fkh), a zygotic transcription factor that is required for the specification of hindgut identity (Jürgens and Weigel, 1988; Weigel et al., 1989; Kuhn et al., 1995; Hoch and Pankratz, 1996). *fkh* is expressed in the posterior of the embryo from stage 4.4 (nuclear cycle 13), with expression strengthening over the course of stage 5 and persisting into germband extension (Weigel et al., 1989; Weigel et al., 1990). This expression is dependent on the terminal system, being absent in *tor*$^-$ mutants (Weigel et al., 1990).

First, we examined the expression of *fkh* relative to other terminal genes in wild-type embryos and in mutant genotypes in which tail or hindgut patterning is perturbed (Figure 7A; expression traces from additional embryonic replicates are shown in Supplementary Figure 21). In wild-type embryos at stage 5.4, the posterior *fkh* domain had a fairly sharp border, which lined up with the anterior border of the posterior *wg* domain and the posterior border of the *cad* tail domain (Figure 7A top row, Supplementary Figure 20). In *cad*$^{m-z-}$ mutants, *fkh* expression was strongly reduced (Figure 7A second row, Supplementary Figure 21), as previously reported (Wu and Lengyel, 1998), contrasting with the *tll* and *hkb* domains in these embryos, which looked normal (Supplementary Figure 21; Wu and Lengyel, 1998; Olesnicky et al., 2006). In *hkb*$^-$ mutants, the *fkh* domain was reduced in size, as previously reported (Weigel et al., 1990; Gaul and Weigel, 1991), which correlated with the reduced size of the *tll* domain and the posteriorly shifted *wg* and *cad* borders in this genotype (Figure 7A third row, Supplementary Figure 21A). The *fkh* domain was also reduced in size in *tll*$^-$

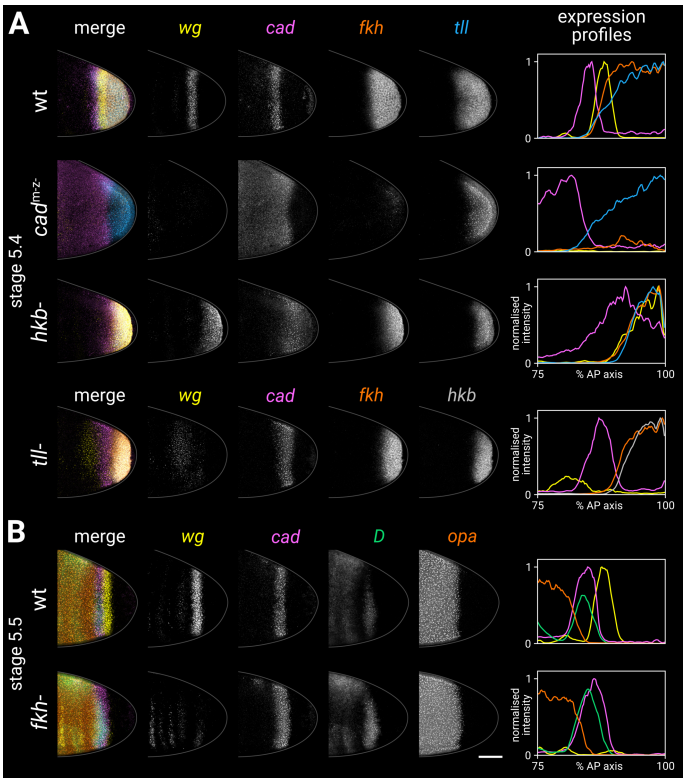

Figure 7: spatial regulation of *fkh*, and timer gene expression in *fkh*⁻ mutants. **A**: terminal gene expression (*wg*, *cad*, *fkh*, and *tll*/*hkb*) in wild-type and mutant embryos at stage 5.4. The leftmost column shows a four-channel merge; middle columns show individual channels; the rightmost column shows quantitative expression traces (75-100% AP axis) from the embryos shown to the left. In the cad^{m-z-} mutant, note the loss of *wg* and *fkh* expression. In the hkb^{-} mutant, note the posterior fate map shift and the delayed repression of posterior *cad*. In the tll^{-} mutant, note the loss of the posterior *wg* domain, the posteriorly shifted *cad* domain, and the reduced size of the *fkh* domain. **B**: timer gene expression in wild-type and fkh^{-} mutant embryos at stage 5.5. Note the extremely reduced posterior *wg* domain in the fkh^{-} mutant. All embryos are anterior left, dorsal up, lateral view. Scale bar = 50 μm; grey lines show embryo outlines.

mutants (Figure 7A bottom row, Supplementary Figure 21), as previously reported (Weigel et al., 1990; Gaul and Weigel, 1991). The reduced *fkh* domain in tll^{-} mutants was the same size as the *hkb* domain, and it abutted the posteriorly shifted *cad* tail domain, which we described in the previous section.

Second, we examined timer gene expression in fkh^{-} mutants (Figure 7Be; xpression traces from additional embryonic replicates are shown in Supplementary Figure 22). We confirmed that the posterior *wg* domain was largely absent in these mutants, as previously reported (Wu and Lengyel, 1998), although there was some residual posterior *wg* expression, particularly in ventral tissue. *cad*, *D*, and *opa* expression looked largely normal throughout stage 5, although the *cad* posterior border appeared to be slightly posteriorly expanded relative to the *D* tail domain. After gastrulation, new *cad* transcription appeared at the posterior edge of the *cad* tail domain, whereas in wild-type embryos new *cad* transcription appeared just posterior to the *wg* domain, and was separated from the existing *cad* expression domain by a clear gap (Supplementary Figure 23). Our findings contrast with a previous report, which described *cad* expression as being normal in fkh^{-} mutants (Jürgens and Weigel, 1988).

We also found that morphogenesis was abnormal in fkh^{-} mutants, in that proctodaeal invagination was delayed until after stage 7 (Supplementary Figure 24). This finding contrasts with previous reports that morphogenesis in fkh^{-} mutants is normal until the end of the extended germband stage (Weigel et al., 1989; Wu and Lengyel, 1998). cad^{m-z-} mutants, which as described above have severely reduced levels of *fkh* expression, show a similar morphogenetic delay (Supplementary Figure 24). While we could not find a previous description of this specific delay in cad^{m-z-} mutants, it has been previously noted that the hindgut primordium fails to internalise in cad^{m-z-} mutants, although the posterior midgut and Malpighian tubule primordia do (Wu and Lengyel, 1998). Aberrant morphogenesis in cad^{m-z-} mutants is thought to be mediated by a reduction in posterior Fog signalling (Wu and Lengyel, 1998). This reduction in Fog signalling may be mediated by the reduced Fkh levels in cad^{m-z-} mutants, as Fkh is known to activate Fog signalling in other developmental contexts (Chung et al., 2017).

In summary, we found a consistent pattern across wild-type, cad^{m-z-}, hkb^{-}, and tll^{-} genotypes, in which the *fkh* border abutted the posterior border of the *cad* tail domain, and posterior *wg* was only expressed in *fkh*-positive *hkb*-negative territory. Accordingly, in fkh^{-} mutants, the posterior *wg* domain was largely lost.

These results are consistent with previously proposed regulatory interactions: that Fkh activates *wg* (Wu and Lengyel, 1998), that Cad activates *fkh* (Wu and Lengyel, 1998), and that Tll and Hkb indirectly enable *fkh* to be expressed (Weigel et al., 1990; Casanova, 1990; Goldstein, 1999; Moran and Jimenez, 2006). Accordingly, the activation of *wg* by Cad (Figure 4A, Supplementary Figure 24; Wu and Lengyel, 1998) appears to be indirect, via Fkh (see Supplementary Table 1 for detailed reasoning). In addition, it is possible that Fkh represses *cad*, but current evidence is inconclusive (see Supplementary Table 1).

Inferred regulatory interactions collectively form a network that can be formalised and simulated

Thus far, we have looked at how gene expression is affected in various mutant genotypes, and have inferred a network of regulatory interactions between the timer genes and the posterior terminal genes (Figure 8A; Supplementary Table 1). Most (11/18) of these proposed interactions originate from this study, although we also find support for previously proposed interactions related

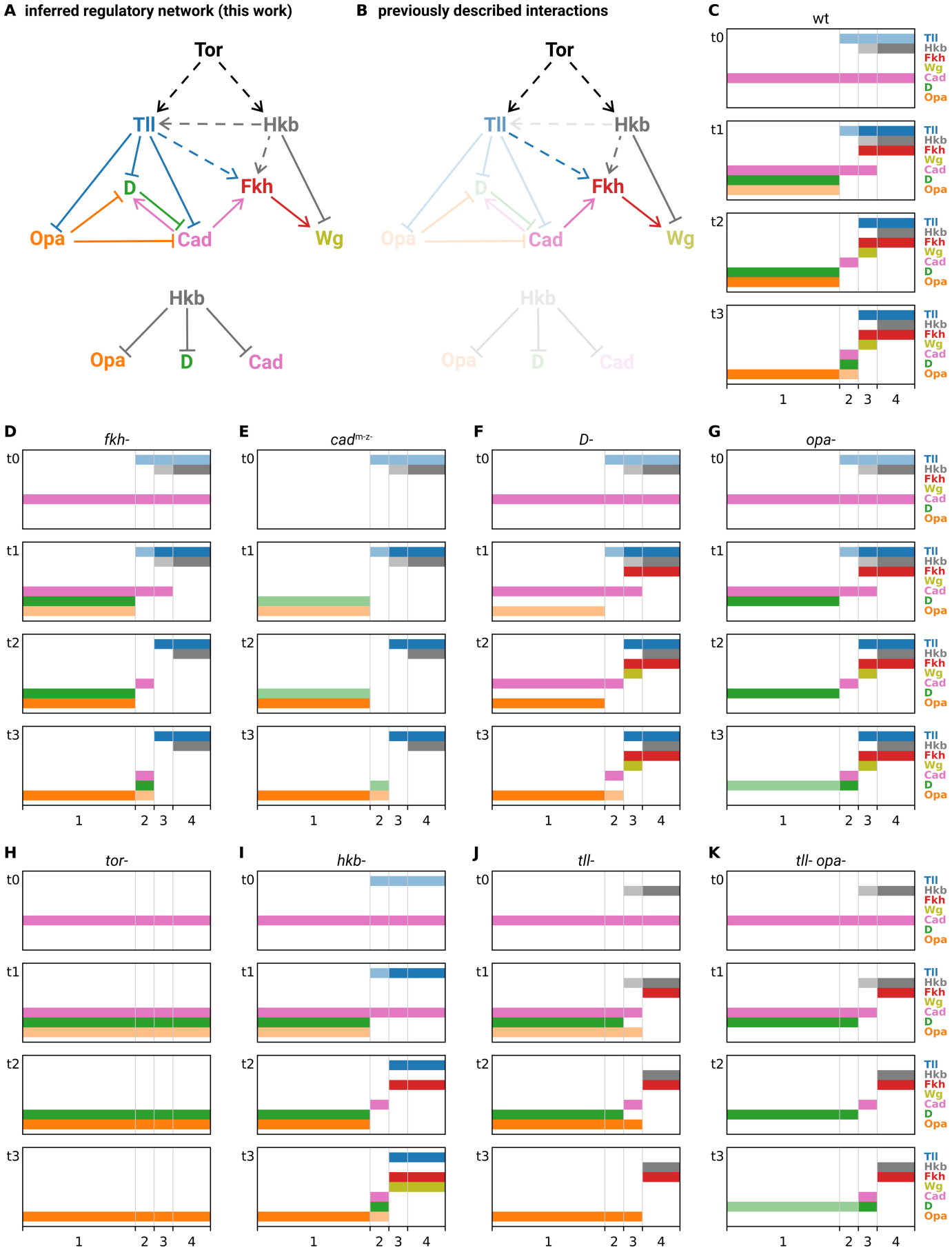

Figure 8: inferred regulatory network for posterior terminal patterning and output of resulting model. (Legend continued on next page.)

Figure 8: A: arrow diagram showing the regulatory interactions we have inferred from the experiments described in this work. Pointed arrowheads indicate activation; flat arrowheads indicate repression. Solid lines indicate interactions that are presumed to be direct; dashed lines indicate interactions that are presumed to be indirect. The diagram is laid out so that the factors are arranged in approximately the same order left to right as their expression along the AP axis, and causation mainly flows from top to bottom (with exceptions for Opa and Cad). To avoid arrow crossovers, the repression of Opa, D and Cad by Hkb is shown separately from the main network. **B**: the same network as in A, highlighting the interactions described in the existing literature. **C-K**: simulation output for a logical model of posterior terminal patterning, for wild-type and eight mutant genotypes (see main text for details). Each set of plots shows the expression patterns of the logical variables `Tll`, `Hkb`, `Fkh`, `Wg`, `Cad`, `D` and `Opa` (y-axis) across AP regions 1-4 (x-axis), at timepoints $t0$-$t3$. For `Tll`, `Hkb`, `D`, and `Opa`, a light colour shade represents weak expression and a dark colour shade represents strong expression. Mutant genotypes never express the relevant protein; *tor*⁻ mutants were simulated as *tll*⁻ *hkb*⁻ double mutants.

to the patterning of *tll*, *hkb*, *fkh* and *wg* (Figure 8B).

We now formalise this regulatory network as a logical model, and see whether it reproduces the patterning dynamics that we observed in the embryo. This exercise is important for multiple reasons. First, by determining whether we can work backwards from interactions to observations, the model tests the soundness of the original reasoning from observations to interactions. Second, by testing the sufficiency of the network to explain the original data, the model indicates whether we are likely to have overlooked a major component of the system. Third, by mechanically reproducing the patterning phenomena we are interested in explaining, the model clarifies how the system is likely to work in the real embryo.

For the purposes of this study, we are aiming for a minimal, qualitative explanation of timer gene patterning, commensurable with the essentially qualitative developmental genetic paradigm we have been working within. We are interested in the relative ordering of gene expression domains in time and space, abstracted away from specific domain sizes, expression levels or expression kinetics. To the extent that the model is able to recapitulate the essential features of both wild-type and mutant genotypes, our confidence in the network will be accordingly increased.

The modelling framework we have chosen is very simple (for a full description see Methods: Model details). Briefly, Hkb and Tll are assumed to be extrinsic inputs to the system (we ignore the cross-regulation of *tll* by Hkb), and we model how Fkh, Wg, Cad, D, and Opa are expressed in response. Each of these 7 factors is modelled as a logical variable, some of which (`Hkb`, `Tll`, `D`, `Opa`) may take one of three levels of expression (off/weak/strong), while the others (`Fkh`, `Wg`, `Cad`) may take only two (off/on). The AP axis is modelled as four discrete regions, 1-4 (corresponding to trunk, tail, hindgut, and posterior midgut, respectively), which differ in their hardcoded `Hkb` and `Tll` inputs. Each simulation consists of 4 time points, $t0$-$t3$ (corresponding to nuclear cycle 13, early stage 5, mid stage 5, and stage 6, respectively). At $t0$, `Cad` is on in all regions, and the other output factors are off. Expression at subsequent timepoints is computed from expression at $t(n-1)$, according to factor-specific logical rules (which remain the same for all timepoints). Mutants are simulated by keeping the relevant factor(s) turned off for all timepoints.

The regulatory network explains the patterning dynamics of each genotype

We simulated the patterning model for the wild-type condition (Figure 8C) and eight mutant genotypes examined in this study (*fkh*⁻, *cad*$^{m\text{-}z\text{-}}$, *D*⁻, *opa*⁻, *tor*⁻, *hkb*⁻, *tll*⁻, and *tll*⁻ *opa*⁻) (Figure 8D-K). Allowing for the simplicity of the model, it was remarkably accurate at recapitulating the observed patterning dynamics of each genotype (see Supplementary Table 3 to cross-reference empirical and in silico observations). Below, we describe the overall logic of patterning for the wild-type condition, and the causes of the deviations that arise in the mutants. We also flag two behaviours of the real system that are not recapitulated by the model, suggesting an interesting avenue for future research.

Wild-type (Figure 8C): At $t1$, all three timer genes have begun to be expressed, but they are differentially repressed by the terminal gap genes; `D` and `Opa` are more sensitive to `Tll` and so are repressed everywhere but region 1, while `Cad` is only repressed in region 4, due to the strong `Hkb` expression there. `Fkh` has been activated in regions 3 and 4 due to strong combined `Hkb` and `Tll` expression, together with activation from `Cad`. At $t2$, the `Cad` expression domain has refined from both the anterior and the posterior. In region 1 it has been repressed by `D`, and in region 3 it has been repressed by the strengthening of `Tll` expression. `Wg` has been activated by `Fkh` in region 3, but remains repressed in region 4 by strong `Hkb` expression. At t3, `D` has been repressed in region 1 by the strong `Opa` expression that has built up over time. Finally, `D` and `Opa` have been de-repressed in region 2, due to the previous clearance of `Tll`.

fkh⁻ (Figure 8D): due to the absence of `Fkh`, `Wg` is never activated in region 3.

cad$^{m\text{-}z}$ (Figure 8E): due to the absence of `Cad`, `Fkh` is never activated in regions 3-4, and `Wg` in turn is never activated in region 3. `D` is also expressed less strongly, both in region 1 and in region 2. (We modelled mutants as deficiencies, so `Cad` expression is not shown. However, we can explain the delayed *cad* repression in *cad*$^{m\text{-}z\text{-}}$ embryos as a knock-on effect of reduced D expression, making Cad

an indirect autorepressor.)

D⁻ (Figure 8F): due to the absence of D, Cad expression persists longer in region 1, although it is later repressed by Opa. Real *D*⁻ embryos also show a modest anterior shift and expansion of the fate map, which is not recapitulated by the model. A potential explanation for this phenomenon would be if D repressed *tll*, and was therefore involved in shaping the *tll* domain early in the patterning process, when *tll* and *D* are expressed in opposing gradients. The *tll* domain would thus be expected to expand in *D*⁻ embryos, with cascading downstream effects. Although we have not tested this prediction as part of this project, there are hints in the literature that *tll* and *hkb* may be cross-regulated by other AP patterning genes (Casanova et al., 1994; Greenwood and Struhl, 1997; de las Heras and Casanova, 2006).

opa⁻ (Figure 8G): due to the absence of Opa, D is not repressed completely in region 1. The residual D expression in region 1 is weaker than in region 2, because only region 2 receives activation from Cad.

tor⁻ (Figure 8H, modelled as a *hkb*⁻ *tll*⁻ double mutant): in the absence of Tll and Hkb input, all regions behave exactly like region 1.

hkb⁻ (Figure 8I): due to the absence of Hkb, Cad expression persists for longer in region 4 and Wg is derepressed. In the simulation, we also see a delay in Fkh and (therefore) Wg expression, which does not affect the final expression pattern, but was not something we noticed in real embryos. In real embryos, we instead saw a marked reduction in the sizes of the *tll* and (therefore) *fkh* domains, along with a posterior shift of the resulting fate map. As with the anterior fate map shift in *D*⁻ embryos, this phenomenon suggests that *tll* expression is shaped by zygotic cross-regulation in addition to terminal signalling.

tll⁻ (Figure 8J): due to the absence of Tll, the expression of all three timer genes is posteriorly expanded and the size of the Fkh domain is reduced. An assumption of graded early Hkb expression that represses D more anteriorly than Cad and Opa is necessary to explain the transient Cad expression in region 3: Cad is first repressed by D in regions 1 and 2, and only later by Opa in region 3. Because Fkh is not expressed outside the Hkb domain, Wg is never expressed.

tll⁻ *opa*⁻ (Figure 8K): patterning resembles the *tll*⁻ mutant through $t2$, but diverges at $t3$ due to the absence of Opa. Specifically, Cad expression in region 3 is allowed to persist, and D expression is de-repressed in region 3 after the clearance of Hkb. Weak D expression also persists in regions 1 and 2, similar to region 1 in *opa*⁻ mutants. In real embryos, this emergence of "tail-like" *cad* and *D* patterning posterior to the normal tail region, and in the absence of the normal tail-patterning signal (Tll), is a striking phenotype (Supplementary Figure 11 *tll*⁻ *opa*⁻), which highlights the potential for any dynamic gradient that interacts with the timer gene network to generate a similar effect.

In summary, we found that the genetic interactions we uncovered in this study were able to explain the qualitative aspects of timer gene patterning across a wide array of mutant genotypes. In addition, modelling the regulatory system emphasised the key role played by Tll spatiotemporal dynamics in patterning the tail. The graded nature of the Tll domain was important for differentially repressing Cad relative to Cad's repressors, Opa and D, while the dynamic nature of the Tll domain was important for de-repressing these repressors later on. Thus a single input signal was able to delineate both the anterior and posterior boundaries of the tail region, and also influence temporal aspects of patterning within it.

Discussion

In this study we have used mutants, multiplexed imaging and modelling to elucidate how the blastoderm expression dynamics of the *Drosophila* timer genes *cad*, *D*, and *opa* arise from a combination of cross-regulatory interactions and spatially localised inputs from the posterior terminal system. This work has four main implications. First, we have demonstrated that timer gene expression is partially driven by intrinsic network dynamics. Second, we have added to the evidence that this network has broad effects on developmental timing, in our discovery that *cad*$^{m-z-}$ embryos precociously express genes associated with neural differentiation. Third, we have produced a coherent model for the patterning of the posterior terminal region. Fourth, we have clarified the segmental nature of the *Drosophila* tail. These findings increase our understanding of *Drosophila* development and have evolutionary import for the mechanisms of axial patterning in other species.

Timer gene expression is regulated by intrinsic network dynamics and extrinsic spatiotemporal inputs

This work provides evidence for a set of cross-regulatory interactions between *cad*, *D*, and *opa* that helps generate dynamic, sequential expression. In particular, we find that Cad activates *D* (i.e. promotes the expression of the next gene in the sequence), while D represses *cad* and Opa represses *cad* and *D* (i.e. both inhibit the previous genes in the sequence). Opa is not cross-regulated, however, making it an "input-only" component of the three gene network (at least in the blastoderm context).

Timer gene expression is also shaped by extrinsic spatiotemporal regulation. In this work, we show how the timer gene network interacts with the posterior terminal system: most notably, Tll differentially represses *cad*, *D* and *opa* in the tail region, indirectly allowing *cad* expression to be maintained. In addition, the inputs from the posterior terminal system are overlaid on global temporal regulation provided by the nuclear:cytoplasmic ratio (which is particularly important for regulating the onset of *opa* transcription; Lu et al., 2009) as well as the levels of maternal factors such as Tramtrack (Harrison and Travers, 1990; Brown et al., 1991; Read et al., 1992), Zelda (Liang

et al., 2008; Harrison et al., 2011; Nien et al., 2011; McDaniel et al., 2019), Stat92e (Yan et al., 1996; Hou et al., 1996; Tsurumi et al., 2011) and GAGA Factor/Trithorax-like (Farkas et al., 1994; Bhat et al., 1996; Moshe and Kaplan, 2017; Gaskill et al., 2021). Ironically, precisely because these maternal factors are so crucial to development, their patterning roles remain less well understood than those of the zygotic patterning genes, which are less pleiotropic and therefore easier to study.

Timer gene expression has broad effects on developmental timing

Recent work in the *Drosophila* blastoderm has demonstrated the extensive effects of timer genes on developmental gene expression. Opa has been shown to act as a pioneer factor, reshaping gene expression genome-wide by opening chromatin at hundreds of target enhancers (Soluri et al., 2020; Koromila et al., 2020). Cad and D are also known to regulate expression across the genome (Li et al., 2008; MacArthur et al., 2009; Aleksic et al., 2013). Here, we have found that early Cad expression appears to be necessary for the correct timing of later developmental events, because neuroectodermal gene expression turns on precociously in *cad*$^{m-z-}$ embryos. The vertebrate Cad ortholog Cdx4 has also been shown to temporally regulate neural differentiation, in the developing spinal cord (Joshi et al., 2019), a tissue in which D and Opa orthologs play key developmental roles (reviewed in Graham et al., 2003; Merzdorf, 2007; Houtmeyers et al., 2013; Stevanovic et al., 2021). More generally, comparative evidence suggests that Cad/Cdx plays a deeply conserved role in the formation of the posterior body and the patterning of the posterior gut (Copf et al., 2004; Wu and Lengyel, 1998; van Rooijen et al., 2012; Zhong et al., 2020). In this context, *Drosophila cad*$^{m-z-}$ mutants offer a rare opportunity to study the genome-wide effects of a total loss of Cad/Cdx function without also catastrophically perturbing early developmental events.

A revised picture of posterior terminal patterning in *Drosophila*

In this work we have investigated blastoderm gene expression downstream of the posterior terminal system, revisiting a patterning network that was most intensely studied in the late 1980s and early 1990s (Strecker et al., 1986; Mahoney and Lengyel, 1987; Mlodzik and Gehring, 1987b; Strecker et al., 1988; Jürgens and Weigel, 1988; Weigel et al., 1990; Casanova, 1990; Bronner and Jackle, 1991; Wu and Lengyel, 1998). The modern availability of marked balancers and multiplexed imaging techniques has allowed us to clarify the topology and spatiotemporal dynamics of this network, as well as incorporate genes (*D* and *opa*) that had not been cloned at the time most of the original work was completed. This study also includes the first formalised model (to our knowledge) for how posterior terminal patterning downstream of Tll and Hkb occurs, providing a foundation for future quantitative analyses of this system.

Although simple, our model provides new insights into how the tail and hindgut regions are specified in the early embryo. Both regions, along with segment A8, have long been known to depend on Tll expression (Strecker et al., 1986; Diaz et al., 1996). *tll* alleles can be arranged into a coherent phenotypic series in which the most posterior structures within the Tll-dependent region are the most sensitive to *tll* perturbation and the most anterior structures are the least (Strecker et al., 1986; Diaz et al., 1996), suggesting that this part of the blastoderm fate map is patterned by a gradient of Tll activity (Casanova, 1990). However, it has not been clear at the network level how graded Tll activity would be transduced into a specific series of boundaries and domains.

We found that *tll* expression was strong and persistent within the hindgut region, but weaker and transient in the tail region, with the anterior border of the expression domain retracting posteriorly across nuclei over time. We additionally found that Tll effectively patterned both the anterior and posterior boundaries of the tail region, by differentially repressing *D* and *opa* relative to *cad*. Crucially, *D* and *opa* were repressed even where Tll expression was transient and weak, but *cad* was not repressed (and *fkh* was not activated) unless Tll expression was stronger, helping explain the transition from tail fate to hindgut fate as Tll levels increase. Furthermore, the retraction of the Tll domain over time explains the posterior shifting dynamics we found for the timer genes within the tail region, which contrasts with the anterior shifting dynamics previously described for the pair-rule and gap genes (Jaeger et al., 2004; Keränen et al., 2006; Surkova et al., 2008; Lim et al., 2018).

We also discovered, to our surprise, that there is a concerted posterior fate map shift in *hkb*$^{-}$ embryos, apparently mediated by a reduction in the size of the *tll* domain. (A subtle anterior fate map shift additionally occurs in *D*$^{-}$ embryos, which might also be mediated by *tll*.) Although further research is necessary to determine the mechanism by which Hkb cross-regulates *tll*, the phenotype implies that the size of the *tll* domain is not an unmediated response to terminal signalling. This finding may complicate the interpretation of recent studies that have characterised this particular input:output relationship using optogenetics (Johnson and Toettcher, 2019; Johnson et al., 2020; Keenan et al., 2020).

The segmental character of the *Drosophila* tail

The ancestral insect body plan has 11 true abdominal segments plus the periproct/telson, but this number has been reduced in many extant insect lineages (Snodgrass, 1935; Demerec, 1950; Matsuda, 1976; Chapman et al., 2013). In *Drosophila*, the most common view has been that the embryo makes 10 abdominal segments (i.e. 15 parasegment boundaries), with the anal pads located in PS15/A10 (Turner and Mahowald, 1979; DiNardo, 1985; Sato and Denell, 1986; Perkins and Perrimon, 1991; Kuhn

et al., 1992; Schmidt-Ott et al., 1994). In particular, territories corresponding to A8, A9, and A10 are visible at the morphological level during embryogenesis (Turner and Mahowald, 1979), and surveys of *en*, *wg*, *hh* and *slp* staining have found evidence for (at most) 15 parasegment boundaries (DiNardo, 1985; Baker, 1987; Baker, 1988; Kuhn et al., 1992; Grossniklaus et al., 1992; Mohler and Vani, 1992; Tabata et al., 1992; Lee et al., 1992; Tashiro et al., 1993; Kuhn et al., 1995). However, fate mapping experiments (Jürgens, 1987) and surveys of *gsb* expression (Baumgartner et al., 1987; Gutjahr et al., 1993) have suggested that the embryo makes 16 parasegment boundaries, with the anal pads located in PS16/A11. There is also some evidence for A11 from patterns of gene expression in adult genital discs (Freeland and Kuhn, 1996).

Given the small size of the tail region within the embryo, the fact that it is covered by amnioserosa during key stages of patterning, and the fact that it later undergoes complicated morphogenetic rearrangements and fusions that obscure its metameric nature, it is perhaps unsurprising that the number of *Drosophila* segments has not been unambiguously resolved. In this study, we present evidence for a vestigial 16th parasegment boundary in the embryo, by identifying additional domains of *slp* and *wg* expression and reinterpreting previously described domains of *eve* and *en*. These observations suggest that the anal pads are located in PS16. (Whether the tissue between PSB16 and the anus should be classified as a true 11th abdominal segment or a non-segmental periproct/telson is beyond the scope of this paper.) However, PSB16 appears extremely dorsoventrally restricted and may have little functional significance in the organism. As the number of abdominal segments varies across insects (Matsuda, 1976), the mechanistic basis of this evolutionary reduction would be interesting to study within a comparative developmental framework.

Our findings suggest that the *Drosophila* embryo sequentially patterns two parasegment boundaries after gastrulation, and that in both cases the new boundary is patterned by abutting stripes of *slp* and *eve*. In PS15 and PS16, the relative arrangement of *slp*, *eve*, *wg* and *en* expressing cells is the same, conserved pattern that is found at parasegment boundaries in the *Drosophila* trunk and throughout the arthropod phylum (reviewed in Clark et al., 2019). However, tail segmentation differs from trunk segmentation in that resolved, stable *eve* stripes emerge de novo and with single-segmental periodicity, rather than from a dynamic and double-segmental phase of pair-rule gene expression.

Intriguingly, a remarkably similar switch from double-segment to single-segment periodicity occurs towards the end of segmentation in the centipede *Strigamia maritima*, where stable, resolved *eve* stripes start appearing de novo in the anterior segmentation zone instead of emerging from posterior oscillatory expression (Brena and Akam, 2013). A possible switch from double-segmental to single-segmental patterning has also been reported for terminal segments in the beetle *Tribolium* (Janssen, 2014). These observations hint that terminal and trunk segments may be homonomous at the level of segment-polarity gene expression but derived from distinct ontogenetic programs. More work is needed to determine how such a developmental switch - if present - is controlled, as well as its relationship to the more general problem of terminating axial development.

Comparative analysis and evolutionary implications

We end this study by assessing the relevance of our findings from *Drosophila* to the development of other insect species. Which aspects of the *Drosophila* network are likely to be conserved in other insect species that have been used to study segmentation, such as *Tribolium*, *Nasonia vitripennis*, and *Oncopeltus fasciatus*? And how might the *Drosophila* network differ from that of its sequentially segmenting ancestors?

The cross-regulatory interactions that we found between the timer genes might be quite widely conserved in insect segmentation. Activation of *D* by Cad, repression of *cad* by Opa, and repression of *D* by Opa are all consistent with a segment addition zone that is subdivided into a posterior region that expresses Cad and D and an anterior region that expresses Opa, as seen for example in *Tribolium* (Clark and Peel, 2018). However, repression of *cad* by D would need to be reconciled with the sustained expression of both *cad* and *D* in the posterior segment addition zone. Intriguingly, some of the timer gene cross-regulatory interactions may even be important for regulating expression dynamics in completely different developmental contexts, given that Opa has recently been found to repress *D* during the temporal patterning of *Drosophila* intermediate neural progenitors (Abdusselamoglu et al., 2019).

The different components of the *Drosophila* terminal system seem to have acquired their posterior patterning roles at various different times: posterior *tll* expression is found in diverse holometabolan species (Schroder et al., 2000; Lynch et al., 2006; Wilson and Dearden, 2009; García-Solache et al., 2010; Lemke et al., 2010; Klomp et al., 2015) although not in hemipterans (Weisbrod et al., 2013; Bickel et al., 2013), whereas *hkb* and *tor* appear to have been recruited to terminal patterning roles more recently (García-Solache et al., 2010; Kittelmann et al., 2013; Duncan et al., 2013). In *Tribolium*, *tll* is expressed downstream of *tor* (as in *Drosophila*), and *tor* RNAi embryos fail to express *cad* and *wg* in the posterior of the embryo, resulting in anteroposterior truncation (Schoppmeier and Schröder, 2005). In *Nasonia*, *tll* RNAi results in a reduction of posterior *cad*, as well as in gap gene misregulation that disrupts much of abdominal segmentation (Lynch et al., 2006). It will be instructive to test whether these losses of *cad* expression in *Tribolium* and *Nasonia* are mediated by ectopic expression of Opa, as we found for *tll*$^-$ and *tor*$^-$ mutants in *Drosophila*. If so, it would suggest that the initial spatial regulation of the timer gene network by Tll in the posterior blastoderm might be conserved across holometabolan embryos, de-

spite their varying modes of development.

So, how *does* timer gene regulation differ between sequentially segmenting embryos (which establish a persistent segment addition zone) and simultaneously segmenting embryos like *Drosophila*? One key difference is likely to be the role of a posterior Wnt signalling centre: there is evidence from many different sequentially segmenting species that Wnt signalling is important for activating *cad* expression and maintaining the segment addition zone (reviewed in Williams and Nagy, 2017; Clark et al., 2019), whereas we found that timer gene expression was unaffected in *Drosophila wg*$^{-}$ mutants, at least during our stages of interest. In addition, it seems probable that timer gene cross-regulation of *opa* is important in sequentially segmenting species, with this having been lost from the *Drosophila* lineage during the evolution of simultaneous patterning.

If we modify the *Drosophila* timer gene network to incorporate these additional features, we can see how appropriate segment addition zone dynamics might naturally emerge (Supplementary Figure 25). It therefore seems plausible that the cross-regulatory interactions between the *Drosophila* timer genes may represent an evolutionary vestige of a "dynamical module" that was originally involved in axial elongation (Clark and Peel, 2018; Clark, 2021). Functional experiments in sequentially segmenting species will be necessary to test this hypothesis.

Methods

Drosophila husbandry and genetics

Stock maintenance and embryo fixation (20 minutes with 4% formaldehyde in PBS) was performed as described in Sullivan et al., 2000). "Wild-type" flies were Oregon R. The mutant alleles used were *wg*$^{I-8}$ (Bloomington #5351), *cad*^{3} (gift from H. Skaer), *cad*^{2} (Bloomington #7091), *D*^{r72} (gift from S. Russell), *opa*^{8} (Bloomington #5340), *tor*^{XR1} (gift from T. Johnson), *hkb*$^{[A321R1]}$ (Bloomington #2059), Df(3R)Exel6217 (Bloomington #7695, a deficiency covering the *tll* locus), and *fkh*^{6} (gift from K. Roeper). Mutant lines obtained from the Bloomington *Drosophila* Stock Centre were verified by cuticle preparations as described in Sullivan et al., 2000. The *tll*$^{-}$ *opa*$^{-}$ double mutant was generated by the Cambridge Fly Facility by recombining Df(3R)Exel6217 and *opa*^{8}. Mutants were balanced over marked balancer chromosomes expressing *lacZ* during early embryogenesis: *CyO*, *hb-lacZ* (Bloomington #6650) for the 2nd chromosome and *TM6c*, *twi-lacZ Sb*^{1} *Tb*^{1} (Bloomington #7251) or *TM3*, *hb-lacZ Sb*^{1} (gift from S. Russell) for the 3rd.

cad germline clones were generated using the heatshock induced FLP/FRT system as described in Selva and Stronach, 2007. Briefly, 8 vials of 30 *cad*^{2} *FRT40A* / *CyO* virgin females (Bloomington #7091) were each crossed with 10 *hsFLP w*; *ovoD1 FRT 40A* / *CyO* males (constructed by crossing Bloomington #6 *hsFLP w*; *Adv* / *CyO* females with Bloomington #2121 *ovoD1 FRT 40A* / *CyO*, but note that #2121 is no longer listed in Bloomington). Adults were flipped to new vials every two days, resulting in a total of ∼100 vials. When crawling L3 larvae were visible, vials were heatshocked at 37° C in a waterbath for 1 hour, allowed to recover at 25° C for 24 hours, then heatshocked again at 37° C for 1 hour. Approximately 600 non-*CyO* virgin females (all presumably with *cad*^{2} / *cad*^{2} ovaries) were collected from the heatshocked vials and crossed with ∼300 *cad*^{3} / *CyO*, *hb-lacZ* males. Resulting embryos without *lacZ* expression lacked both maternal and zygotic *cad* (*cad*$^{m-z-}$), while embryos with *lacZ* expression were paternal rescues (*cad*$^{m-z+}$). Zygotic *cad* mutants (*cad*$^{m+z-}$) were offspring from *cad*^{3}/ *CyO*, *hb-lacZ* parents that lacked *lacZ* expression; note that this genotype is also heterozygous for maternal *cad*.

Opa antibody generation

Clone FI01113 containing Opa coding sequence was obtained from the Drosophila Genomics Resource Center. Gateway attB primers were designed to express 386 amino acids from the N-terminus of Opa (amino acids 3 to 389), spanning the zinc finger region in the centre of the protein. The forward primer included a Shine-Dalgarno sequence; the reverse primer was designed to be in-frame with the C-terminal fusion of the Gateway expression vector pET-DEST42 (ThermoFisher Scientific). A two stage PCR procedure was used to obtain a final amplicon carrying the attB-sequences at each end of the N-terminal opa sequence.

Primers for the first amplification were:

opaDM-F: AAAAAGCAGGCTTCGAAGGAGATAGAACCATGAACGCCTTCATTGAGC

opaA-R: AGAAAGCTGGGTTGTCGTAGCCGTGGGATG

Overlapping primers for the second amplification to complete the attB regions were:

attB1adap-F: GGGGACAAGTTTGTACAAAAAAGCAGGCT

attB2adap-R: GGGGACCACTTTGTACAAGAAAGCTGGGT

The attB-opa amplicon was obtained by PCR with Phusion proofreading polymerase (ThermoFisher Scientific), using primers opaDM-F/opaA-R. This first amplicon was diluted 1000-fold, then Phusion PCR was repeated with primers attB1adap-F/attB2adap-R. This attB-opa amplicon was recombined into Gateway donor vector pDONR (ThermoFisher Scientific), using the BP Clonase II kit (ThermoFisher Scientific). Plasmid DNA from a sequence-verified clone was then recombined into pET-DEST42, using the LR Clonase II kit (ThermoFisher Scientific).

For expression of the fusion protein, plasmid DNA was transformed into One Shot BL21 Star (DE3) Chemically Competent E. coli (ThermoFisher Scientific). Opa protein was expressed in two ways, firstly by IPTG induction of exponentially growing cells (0.75 mM IPTG for 2.75 hours), secondly by overnight culture in TB Overnight Express (Novagen). The Opa fusion protein in pET-DEST42 had a C-terminal 6-His tag. Protein was purified from bacterial pellets, each from 100 ml of cells induced in IPTG

or TB Overnight Express. Purification was carried out using Ni-NTA Agarose (Qiagen), under 8 M urea denaturing conditions according to the manufacturer's protocol. Purified protein was dialysed against water then concentrated using an Amicon Ultra-Ultracel 5k centrifugal filter (Millipore). Antibodies were raised in two guinea pigs by Eurogentec.

HCR in situ hybridisation and antibody staining

Prior to staining, fixed embryos stored in methanol were put through a rehydration series of 5 minutes each at 75%, 50%, and 25% methanol in PBS + 0.1% Tween-20, then washed 3 times with PBS + 0.1% Tween-20.

HCR in situ hybridisation v3.0 (Choi et al., 2018) was performed using probes and hairpins produced by Molecular Instruments, following the protocol for whole-mount fruit fly embryos included in Choi et al., 2016, adapted for v3.0 probes as described in Choi et al., 2018. In addition, the percentage of dextran sulphate in the probe hybridisation and amplification buffers was reduced from 10% w/v to 5% w/v, to reduce viscosity and allow the embryos to settle more easily in the tube. A 20 minute postfix step (4% formaldehyde in 5x SSC + 0.1% Tween-20) was added at the end of the protocol to stabilise the signal.

For antibody staining following HCR, embryos were incubated for 30 minutes in blocking solution (5% Normal Goat Serum (Vector Laboratories) in 5x SSC + 0.1% Triton X-100), at room temperature with rocking. Embryos were then incubated overnight in preabsorbed primary antibody diluted in blocking solution, at 4° C with rocking. Embryos were washed 4 times for 15 minutes in 5x SSC + 0.1% Triton X-100, at room temperature with rocking, then incubated for 30 minutes in blocking solution, at room temperature with rocking. Embryos were then incubated for 2 hours with fluorescently labeled secondary antibody diluted in blocking solution, at room temperature with rocking. Embryos were washed 4 times for 15 minutes then 1 time for 30 minutes with 5x SSC + 0.1% Triton X-100, at room temperature with rocking. Antibody staining without prior HCR was performed as above with the exception that PBS was used instead of 5x SSC. Primary antibodies were guinea pig anti-Opa (this work) at 1:5000, rabbit anti-Dichaete (Soriano and Russell, 1998) at 1:10, rabbit anti-Tll (Kosman et al., 1998) at 1:100, and rat anti-Hkb (Ashyraliyev et al., 2009) at 1:100. Secondary antibodies were goat anti-guinea pig Alexa Fluor 647, goat anti-rabbit Alexa Fluor 488, goat anti-rabbit Alexa Fluor 555, and goat anti-rat Alexa Fluor 488 (all ThermoFisher Scientific/Invitrogen), used at 1:1000.

Following HCR and/or antibody staining, embryos were incubated for 30 minutes with 1 ng/μl DAPI (ThermoFisher Scientific) in 5x SSC + 0.1% Tween-20, at room temperature with rocking, then washed 3 times for 30 minutes in 5x SSC + 0.1% Tween-20, at room temperature with rocking. Prior to mounting, embryos were stored in 1.5 mL tubes in SlowFade Gold Antifade Mountant (ThermoFisher Scientific).

Microscopy

Embryos were mounted in SlowFade Gold Antifade Mountant (ThermoFisher) on glass microscope slides (Thermo Scientific) with #1.5 coverslips (Corning). #1.5 coverslips were used as bridges to prevent embryos from being squashed. Clear nail varnish was used to seal the edges of the slide.

Microscopy was performed on an Olympus FV3000 confocal microscope at the Department of Zoology Imaging Facility (University of Cambridge). Acquired images were 12-bit, with a 1024x768 scan format and a 2 μs/pixel dwell time. Whole embryo images were acquired using an Olympus UPlanSApo 30x 1.05 NA silicon immersion oil objective, a physical pixel size of 0.47 μm x 0.47 μm, and a Z-stack step size of 1.5 μm. The close-ups in Figure 1 and Figure 3 were acquired using an Olympus UPlanSApo 60x 1.3 NA silicon immersion oil objective, a physical pixel size of 0.21 μm x 0.21 μm, and a Z-stack step size of 0.8 μm. Each Z-stack was specified so as to span from just above the top surface of the focal embryo through to the middle of its yolk.

In each experiment, embryos had been stained for up to four transcripts and/or proteins of interest plus nuclei, generally using Alexa Fluor 488, Alexa Fluor 546, Alexa Fluor 594, Alexa Fluor 647, and DAPI. (For mutant experiments, a *lacZ* probe or a probe to a gene covered by a deficiency was additionally labelled with one of these same fluorophores, so that homozygous mutant embryos could be easily identified.) All imaging channels were acquired sequentially, to minimise cross-talk. The laser lines and collection windows were: 405 laser and 443-472 nm window for DAPI; 488 laser and 500-536 nm window for Alexa Fluor 488; 561 laser and 566-584 nm window for Alexa Fluor 546 or Alex Fluor 555; 594 laser and 610-631 nm window for Alexa Fluor 594; 640 laser and 663-713 nm window for AlexaFluor 647. Alexa Fluor 514 (514 laser and 519-540 nm window) was used in place of Alexa Fluor 488 for a round of HCR experiments carried out when the 488 laser was awaiting repair. When necessary, a transmitted light channel was also collected, to allow for embryo staging based on the progress of cellularisation.

Embryo staging

Embryo staging was based on Bownes stages (Bownes, 1975; Campos-Ortega and Hartenstein, 1997), with subdivision of particular stages into substages where necessary. Substages and staging criteria are described in Supplementary Table 4. Briefly, embryos younger than stage 5 were staged to a nuclear cycle based on their nuclear density (stage 4.1 = nuclear cycle 10; stage 4.2 = nuclear cycle 11; stage 4.3 = nuclear cycle 12; stage 4.4 = nuclear cycle 13), while embryos older than stage 5 were staged by the progress of morphogenesis, the presence of mitotic domains (Foe, 1989), and/or the appearance of terminal segment-polarity stripes. Stage 5 itself was divided into 5 substages, stage 5.1 to stage 5.5, which can be differentiated from one another on the basis of *wg*

expression, *D* expression, *eve* expression, or the progress of cellularisation (Supplementary Figure 2). (For most stains, we relied on *wg* and/or *D* expression.) Supplementary Table 4 also notes how our stage 5 classification scheme maps onto the 8 "temporal equivalence" classes used in Surkova et al., 2008, and the 4 "phases" used in Schroeder et al., 2011.

Image analysis

Fiji (Schindelin et al., 2012) was used for routine inspection of imaging data. Image processing scripts were written in Python 3 (www.python.org), using the libraries `NumPy` (Harris et al., 2020), `SciPy` (Virtanen et al., 2020), `pandas` (McKinney, 2010) and `scikit-image` (van der Walt et al., 2014). Initial processing of the raw images was carried out to detect, rotate, mask, and crop each focal embryo (Supplementary Figure 26A-E). Briefly, a "height map" of a given z-stack was built up by applying a Gaussian filter to the DAPI channel to smooth it, thresholding a maximum projection of z range $0 : i$ for increasing values of i, and summing these together to produce an image showing the topography of any embryos within the field of view (Supplementary Figure 26B). Local peaks within a truncated version of this image were detected and then used as the seeds for a watershed segmentation to separate touching embryos, while "low-lying" background areas were masked (Supplementary Figure 26C). Because all images were centred on a specific embryo of interest, the convex hull of the central segmented region was used as the embryo mask (Supplementary Figure 26D). The major axis of the embryo mask was used to determine the orientation of this focal embryo, and the image was rotated accordingly so as to align its AP axis with the horizontal. The embryo mask was then dilated slightly, before being used to crop the image and mask non-embryo background (Supplementary Figure 26E). This process was first applied to all images in batch, and then the resulting masks were inspected for accuracy. Any images with unsatisfactory masks were reprocessed individually, with manual parameter adjustment at the image segmentation step to correct the mask. Processed images were then flipped horizontally and/or vertically as necessary, to yield a consistent "anterior left, dorsal up" orientation.

Laterally oriented embryos of the appropriate stages and genotypes were then selected for the extraction of quantitative AP expression traces (Supplementary Figure 26F-L). Previous studies have tended to use % egg length to quantify AP expression profiles (e.g. Pignoni et al., 1990; Surkova et al., 2008; Janssens et al., 2013), but % egg length is not a perfect proxy for the AP axis, due to the embryo's curvilinear intrinsic coordinate system (Spirov et al., 2000; Luengo Hendriks et al., 2006; Spirov et al., 2013). % egg length measurements for expression domains near the poles are also potentially unreliable because they depend on the degree of flattening of a mounted embryo, given that a z-projection of a squashed embryo will exaggerate the size of the termini compared to a z-projection of an unsquashed embryo, due to the different curvature in z. We therefore decided to use a heuristic approach to approximate a curved trace along the lateral surface of the embryo, using guidance from embryo morphology.

Briefly, DAPI-derived height maps were used to define thin embryo "shells" ($\sim$18 μm thick) that tracked the blastoderm surface in 3D and contained most of the gene expression signal (Supplementary Figure 26F,G). A mean z-projection of the voxels within this shell region was then saved as a multichannel 2D image (Supplementary Figure 26H). Next, the dorsal and ventral borders of the embryo mask were used to create a "DV map" for this image by interpolation (Supplementary Figure 26I), and 30% of the DV axis, corresponding to the mid-lateral part of the embryo, was selected for quantification (Supplementary Figure 26J). (As each embryo had a slightly different DV orientation on the slide, the selected DV range had to be adjusted manually for each image, so that the centre of the selected region consistently intersected with a DV position corresponding to the centre of the D head domain. This DV adjustment was important because the positions and expression intensities of most AP expression domains vary along the DV axis (Luengo Hendriks et al., 2006).) A 3D spline was fitted along the middle of this curved DV strip, using z values from the height map. To improve the consistency of the traces, the posterior endpoint of the spline was anchored close to a pixel coordinate marking the transition between the posterior midgut primordium and the pole cells, which was selected manually for each image. Cumulative distance along the spline was calculated in 3D using the Pythagorean theorem, accounting for the anisotropy of the z axis relative to the x and y. The total AP distance along the spline was normalised to 1, where 0 = the anterior tip of the embryo mask, and 1 = the beginning of the pole cells. Expression intensity traces were extracted for each channel by running a sliding window of 1% AP length (roughly 1 nuclear diameter) along the spline, with each window angled normal to the xy orientation of the spline (so as to avoid generating artificial expression overlaps from slanted posterior domains), and bounded dorsally and ventrally by the DV strip specified earlier (Supplementary Figure 26K). Each extracted trace (Supplementary Figure 26J) consisted of 500 measurements separated by a distance of 0.2% AP length.

The expression intensity traces in Figure 3 (solid plotted lines) were calculated by moving a sliding window with a width of 25 pixels ($\sim$1 nuclear diameter) across the x axis of the rectangular region of interest and measuring the average intensity at 1 pixel intervals. Nuclear foci for *opa* and the *cad* intronic probe were identified by detecting local peaks above a threshold intensity; the dashed plotted lines in Figure 3B are density plots for the x coordinates of the detected foci.

Normalised expression plots

Expression plots were generated using `matplotlib` from the expression intensity traces described above. When

comparing traces from embryos of different stages to examine the dynamics of gene expression, all traces from a particular experimental sample were normalised to the range 0-1 as a group (i.e., for each channel, normalised values = (original values - min(group)) / (max(group) - min(group)). When comparing traces from individual embryos of the same stage to examine the positioning of expression domains within and between genotypes, each trace was normalised to the range 0-1 individually (i.e., normalised values = (original values - min(individual) / (max(individual) - min(individual)). In *D*⁻ mutants, expression levels were severely reduced across the entire AP axis, and so the normalised expression traces were multiplied by a small constant to dampen them. In most cases, expression traces are presented without any further adjustments. In Supplementary Figure 9B, additional plots show "aligned traces", in which each trace has been shifted anteriorly or posteriorly by a small amount so that the position of the anterior border of the *wg* posterior domain coincides in all traces. The aligned plots are useful for assessing any changes to the relative positioning of particular domains (as opposed to their absolute positional variation across different embryos).

Figure preparation

Unless otherwise stated, all images of embryos are maximum intensity projections of confocal z-stacks of the upper half of the embryo. Fiji was used to adjust image brightness and contrast, in accordance with guidelines presented by Schmied and Jambor, 2021. Image gamma was adjusted to 0.1 for all *opa* transcript stains, due to the extremely bright transcriptional foci. Embryos from the same round of staining and imaging are presented using the same brightness and contrast values; unless otherwise noted, this holds for any embryos within a given figure that share the same genotype and combination of stains. To correct for uneven illumination from the 405 laser, the DAPI signal from each blastoderm stage embryo was flattened by applying a Gaussian filter with $\sigma = 6$, and then dividing the original image by the new blurred image.

In Figure 1B; stage 8.1, the inset shows a maximum intensity projection from the surface to the midline of the embryo. In Figure 1B; stage 11.1, the inset shows a single section of a z-stack that was rotated -45° around the x axis using the ImageJ plug-in `TransformJ` (Meijering et al., 2001) using the "Quintic B-Spline" method for interpolation. In Figure 3, the curved surface of the embryo was flattened in Fiji by reslicing each channel along the long axis of the embryo (output spacing 0.206 μm), manually masking the region of interest with a segmented line ("spline fit" checked) of width 130 pixels, using the "Straighten" tool to process the entire stack, then reslicing the stack (output spacing 0.206 μm) and re-merging the channels to return to the original view. Average projections (DAPI and Opa channels) or maximum projections (HCR channels) were then generated for a z-range spanning from the top of the embryo to just below the nuclei.

Figures were assembled in Affinity Designer (Serif Europe). Embryo outlines were drawn manually in Affinity. Image look-up tables (LUTs) were either chosen from the "ChrisLUTs" LUT package for ImageJ (Christophe Leterrier and Scott Harden; github.com/cleterrier/ChrisLUTs; "NeuroCyto LUTs" update site in Fiji), or generated for custom colours using a macro provided by Nicolás De Francesco (github.com/ndefrancesco).

Model details

Models were implemented in python using `NumPy`, and outputs were plotted using `matplotlib`. In Figure 8, the *Drosophila* AP axis is modelled as four discrete regions, where region 1 represents the trunk, region 2 represents the tail, region 3 represents the hindgut primordium, and region 4 represents the posterior midgut primordium. Each region receives hard-coded inputs from `Tll` and `Hkb`, logical variables that can take the values 0 (no expression), 1 (weak expression), or 2 (strong expression). Region 1 remains free of both `Tll` and `Hkb` expression $(0, 0, 0, 0)$. Region 2 experiences weak, transient `Tll` expression $(1, 1, 0, 0)$ and no `Hkb` expression $(0, 0, 0, 0)$. Region 3 experiences rapidly established strong `Tll` expression $(1, 2, 2, 2)$ and transient weak `Hkb` expression $(1, 1, 0, 0)$. Region 4 experiences rapidly established strong `Tll` expression $(1, 2, 2, 2)$ and persistent strong `Hkb` expression $(2, 2, 2, 2)$.

In addition to `Tll` and `Hkb`, each region can express `Fkh`, `Wg`, `Cad`, `D`, and `Opa`, logical variables that can take one of either three (0, 1, 2) or two (0, 1) possible values, as defined by their regulatory logic:

```
Fkh = 1 if (Tll + Hkb) > 1 and (Cad + Fkh) > 0;
    else Fkh = 0.
Wg = 1 if Fkh > 0 and Hkb < 2; else Wg = 0.
Cad = 1 if D < 2 and Opa < 2 and Hkb < 2 and Tll
    < 2; else Cad = 0.
D = 2 if Opa < 2 and Tll < 1 and Hkb < 1 and Cad
    > 0; D = 1 if Opa < 2 and Tll < 1 and Hkb < 1
    and Cad < 1; else D = 0.
Opa = 2 if Hkb < 2 and Tll < 1 and Opa > 0; Opa
    = 1 if Hkb < 2 and Tll < 1 and Opa < 1; else
    Opa = 0.
```

Thus, `Fkh` is only expressed when combined `Tll` and `Hkb` levels are high, and `Cad` must initially be present for `Fkh` expression to become established. `Wg` is expressed when `Fkh` is present but `Hkb` levels are low. `Cad` is on by default but repressed by strong `D`, strong `Opa`, strong `Tll` or strong `Hkb`. `D` can be repressed by strong `Opa` or any amount of `Hkb` or `Tll`, and `Cad` must be present for `D` to be expressed strongly. Finally, `Opa` can only be repressed by `Tll` or strong `Hkb`, but it must transit through weak expression before it reaches high levels. This last condition represents the observation that Opa protein is synthesised relatively slowly (Supplementary Figure 6; Clark and Akam, 2016; Soluri et al., 2020).

Each simulation begins at $t0$ with `Cad` ubiquitously expressed, and then proceeds through 3 iterations ($t1$-$t3$) in

which the expression of Fkh, Wg, Cad, D, and Opa is synchronously updated based on the current state of the region. $t0$ represents stage 4, $t1$ represents early stage 5, $t2$ represents mid stage 5, and $t3$ represents stage 6. Over the course of a simulation, expression dynamics within each region are shaped both by the (potentially dynamic) inputs from Tll and Hkb, and by cross-regulation between the other factors. The limited number of expression updates reflects the rapid development of the *Drosophila* blastoderm, which limits the number of regulatory links (i.e. temporally distinct rounds of protein synthesis or decay) within any particular dynamical causal chain (Nasiadka and Krause, 1999). Mutant genotypes are simulated by keeping the relevant factor(s) turned off for all timepoints.

In Supplementary Figure 26, the AP axis of a sequentially segmenting species is modelled as a growing array of "cells" with a Wg signalling centre at the posterior end, as in Clark, 2021. The domain starts at one cell long at $t0$, then adds a cell each iteration by duplicating the most posterior cell. The range of effective Wg signalling is finite (in this case, 8 cells from the posterior signalling centre), so the zone of Wg signalling moves posteriorly with time. Each cell may express Cad, D, and Opa, which are Boolean variables with the following regulatory logic:

```
Cad = 0 if Opa > 0 or (D > 0 and Wg < 1); else
   Cad = 1.
D = 0 if Opa > 0; else D = 1.
Opa = 0 if Cad > 0; else Opa = 1.
```

Thus, Cad is repressed by Opa and D but can be co-expressed with D in the presence of Wg signalling, D is repressed by Opa, and Opa is repressed by Cad. At each iteration, expression in each cell is updated synchronously, based on the current state of the cell.

Supplementary information

- **Supplementary Figure 1**: single channel images associated with Figure 1.
- **Supplementary Figure 2**: substaging scheme for stage 5 embryos.
- **Supplementary Figure 3**: additional single channel images associated with Figure 2.
- **Supplementary Figure 4**: additional early stage embryos associated with Figure 2.
- **Supplementary Figure 5**: D antibody staining in embryos from stage 4 to stage 6.
- **Supplementary Figure 6**: expression dynamics of timer gene transcripts and proteins.
- **Supplementary Figure 7**: single channel images associated with Figure 3.
- **Supplementary Figure 8**: timer gene expression in *wg*$^-$ mutants at stage 6.
- **Supplementary Figure 9**: timer gene expression traces from timer gene mutants, relative to wild-type.
- **Supplementary Figure 10**: full length *opa* expression traces from timer gene mutants.
- **Supplementary Figure 11**: timer gene expression in *opa*$^-$, *tll*$^-$, and *tll*$^-$ *opa*$^-$ embryos at stage 6.
- **Supplementary Figure 12**: additional characterisation of *cad*$^{m-z-}$, *cad*$^{m-z+}$, and *cad*$^{m+z-}$ mutants.
- **Supplementary Figure 13**: morphological differences between wild-type and *cad*$^{m-z-}$ blastoderms.
- **Supplementary Figure 14**: expression of *tll* and *hkb* in wild-type embryos from stage 2 to stage 6.
- **Supplementary Figure 15**: single channel images associated with Figure 5.
- **Supplementary Figure 16**: expression dynamics of *tll* and *hkb* transcripts.
- **Supplementary Figure 17**: relative expression dynamics of Tll protein and *tll* transcript.
- **Supplementary Figure 18**: relative expression dynamics of Hkb, Tll and Opa proteins.
- **Supplementary Figure 19**: examples of source imaging data for Supplementary Figures 17 and 18.
- **Supplementary Figure 20**: timer gene expression traces from terminal mutants, relative to wild-type.
- **Supplementary Figure 21**: terminal gene expression traces from *cad*$^{m-z-}$, *tll*$^-$ and *hkb*$^-$ mutants, relative to wild-type.
- **Supplementary Figure 22**: timer gene expression traces from *fkh*$^-$ mutants, relative to wild-type.
- **Supplementary Figure 23**: *wg* and *cad* expression in *fkh*$^-$ mutants at stages 6 and 7.
- **Supplementary Figure 24**: abnormal morphogenesis in *cad*$^{m-z-}$ and *fkh*$^-$ mutants.
- **Supplementary Figure 25**: simulation of a hypothetical timer gene network for sequential segmentation.
- **Supplementary Figure 26**: illustration of image processing steps.
- **Supplementary Table 1**: evidence for proposed cross-regulatory interactions between Tll, Hkb, Fkh, Wg, Cad, D, and Opa.
- **Supplementary Table 2**: size of the posterior *tll* expression domain at stage 5.2 versus 5.5.
- **Supplementary Table 3**: cross-references for simulation output and corresponding expression data.
- **Supplementary Table 4**: embryo staging and substaging criteria.

Data availability

All necessary data are included in the main text and supplementary information. The confocal imaging dataset on which this study is based will be uploaded to an open access data repository after appropriate curation and organisation; this section will be updated with the relevant details once this has been carried out.

Acknowledgements

This project was made possible by Michael Akam, who provided laboratory space, resources, encouragement, and helpful feedback on the manuscript. We are grateful to Ken Siggens for generating the guinea pig anti-Opa antibody, and to Simon Collier at the Department of Genetics Fly Facility (University of Cambridge) for creating the *tll- opa*$^{-}$ double mutant. We thank the Imaging Facility at the Department of Zoology (University of Cambridge) for confocal imaging support, and members of the *Drosophila* community for various fly lines and reagents. Stocks and materials obtained from the Bloomington Drosophila Stock Center (NIH P40OD018537) and the Drosophila Genomics Resource Center (NIH 2P40OD010949) were used in this study. Information from FlyBase (Larkin et al., 2021) was invaluable.

Funding statement

EC was supported by a BBSRC Research Grant (BB/P009336/1), a Junior Research Fellowship from Trinity College, Cambridge, and an EMBO postdoctoral fellowship (EMBO ALTF 383-2018). MAB was supported by the Deutsche Forschungsgemeinschaft (DFG Research Fellowship BE 6732/1-1), the Isaac Newton Trust (Research Grant) and the University of Cambridge Department of Zoology. MB was supported by a PhD studentship from the Wellcome Trust (220019/Z/19/Z).

Competing interests statement

The authors declare no competing interests.

References

Abdusselamoglu, M. D., Eroglu, E., Burkard, T. R., & Knoblich, J. A. (2019). The transcription factor odd-paired regulates temporal identity in transit-amplifying neural progenitors via an incoherent feed-forward loop. *eLife*, *8*, e46566.

Akam, M. (1987). The molecular basis for metameric pattern in the Drosophila embryo. *Development*, *101*(1), 1–22.

Aleksic, J., Ferrero, E., Fischer, B., Shen, S. P., & Russell, S. (2013). The role of Dichaete in transcriptional regulation during Drosophila embryonic development. *BMC genomics*, *14*, 861.

Ashyraliyev, M., Siggens, K., Janssens, H., Blom, J., Akam, M., & Jaeger, J. (2009). Gene circuit analysis of the terminal gap gene huckebein. *PLoS computational biology*, *5*(10), e1000548.

Baker, N. (1988). Localization of transcripts from the wingless gene in whole Drosophila embryos. *Development*, *103*(2), 289–298.

Baker, N. E. (1987). Molecular cloning of sequences from wingless, a segment polarity gene in Drosophila: The spatial distribution of a transcript in embryos. *The EMBO Journal*, *6*(6), 1765–1773.

Bate, C. M. (1976). Embryogenesis of an insect nervous system. I. A map of the thoracic and abdominal neuroblasts in Locusta migratoria. *Journal of Embryology and Experimental Morphology*, *35*(1), 107–123.

Baumgartner, S., Bopp, D., Burri, M., & Noll, M. (1987). Structure of two genes at the gooseberry locus related to the paired gene and their spatial expression during Drosophila embryogenesis. *Genes & Development*, *1*(10), 1247–1267.

Baumgartner, S., & Noll, M. (1990). Network of interactions among pair-rule genes regulating paired expression during primordial segmentation of Drosophila. *Mechanisms of Development*, *33*(1), 1–18.

Benedyk, M. J., Mullen, J. R., & DiNardo, S. (1994). Odd-paired: A zinc finger pair-rule protein required for the timely activation of engrailed and wingless in Drosophila embryos. *Genes & Development*, *8*(1), 105–117.

Bhat, K. M., Farkas, G., Karch, F., Gyurkovics, H., Gausz, J., & Schedl, P. (1996). The GAGA factor is required in the early Drosophila embryo not only for transcriptional regulation but also for nuclear division. *Development*, *122*(4), 1113–1124.

Bickel, R. D., Cleveland, H. C., Barkas, J., Jeschke, C. C., Raz, A. A., Stern, D. L., & Davis, G. K. (2013). The pea aphid uses a version of the terminal system during oviparous, but not viviparous, development. *EvoDevo*, *4*(1), 10.

Biffar, L., & Stollewerk, A. (2014). Conservation and evolutionary modifications of neuroblast expression patterns in insects. *Developmental Biology*, *388*(1), 103–116.

Birkholz, O., Rickert, C., Berger, C., Urbach, R., & Technau, G. M. (2013). Neuroblast pattern and identity in the Drosophila tail region and role of doublesex in the survival of sex-specific precursors. *Development*, *140*(8), 1830–1842.

Bownes, M. (1975). A photographic study of development in the living embryo of Drosophila melanogaster. *Journal of Embryology and Experimental Morphology*, *33*(3), 789–801.

Brena, C., & Akam, M. (2013). An analysis of segmentation dynamics throughout embryogenesis in the centipede Strigamia maritima. *BMC biology*, *11*, 112.

Bronner, G., & Jackle, H. (1991). Control and function of terminal gap gene activity in the posterior pole region

of the Drosophila embryo. *Mechanisms of Development*, *35*(3), 205–211.

Bronner, G., & Jackle, H. (1996). Regulation and function of the terminal gap gene huckebein in the Drosophila blastoderm. *International Journal of Developmental Biology*, *40*(1), 157–165.

Bronner, G., Chu-LaGraff, Q., Doe, C. Q., Taubert, H., & Jackie, H. (1994). Spl/egr-like zinc-finger protein required for endoderm specification and germ-layer formation in Drosophila. *Nature*, *369*, 664–8.

Brown, J. L., Sonoda, S., Ueda, H., Scott, M. P., & Wu, C. (1991). Repression of the Drosophila fushi tarazu (ftz) segmentation gene. *The EMBO journal*, *10*(3), 665–674.

Cadigan, K. M., Grossniklaus, U., & Gehring, W. J. (1994). Localized expression of sloppy paired protein maintains the polarity of Drosophila parasegments. *Genes & Development*, *8*(8), 899–913.

Campos-Ortega, J. A., & Hartenstein, V. (1997). *The Embryonic Development of Drosophila Melanogaster*. Springer Berlin / Heidelberg.

Casanova, J., & Struhl, G. (1989). Localized surface activity of torso, a receptor tyrosine kinase, specifies terminal body pattern in Drosophila. *Genes & Development*, *3*(12B), 2025–2038.

Casanova, J. (1990). Pattern formation under the control of the terminal system in the Drosophila embryo. *Development*, *110*, 621–8.

Casanova, J., Llimargas, M., Greenwood, S., & Struhl, G. (1994). An oncogenic form of human raf can specify terminal body pattern in Drosophila. *Mechanisms of Development*, *48*(1), 59–64.

Chapman, R. F., Simpson, S. J., & Douglas, A. E. (2013). *The Insects: Structure and Function*. Cambridge University Press.

Choi, H. M. T., Calvert, C. R., Husain, N., Huss, D., Barsi, J. C., Deverman, B. E., Hunter, R. C., Kato, M., Lee, S. M., Abelin, A. C. T., Rosenthal, A. Z., Akbari, O. S., Li, Y., Hay, B. A., Sternberg, P. W., Patterson, P. H., Davidson, E. H., Mazmanian, S. K., Prober, D. A., ... Pierce, N. A. (2016). Mapping a multiplexed zoo of mRNA expression. *Development*, *143*(19), 3632–3637.

Choi, H. M. T., Schwarzkopf, M., Fornace, M. E., Acharya, A., Artavanis, G., Stegmaier, J., Cunha, A., & Pierce, N. A. (2018). Third-generation in situ hybridization chain reaction: Multiplexed, quantitative, sensitive, versatile, robust. *Development*, *145*(12), dev165753.

Chung, S., Kim, S., & Andrew, D. J. (2017). Uncoupling apical constriction from tissue invagination. *eLife*, *6*, e22235.

Clark, E. (2017). Dynamic patterning by the Drosophila pair-rule network reconciles long-germ and short-germ segmentation. *PLoS biology*, *15*(9), e2002439.

Clark, E. (2021). Time and space in segmentation. *Interface Focus*, *11*(3), 20200049.

Clark, E., & Akam, M. (2016). Odd-paired controls frequency doubling in Drosophila segmentation by altering the pair-rule gene regulatory network. *eLife*, *5*, e18215.

Clark, E., & Peel, A. D. (2018). Evidence for the temporal regulation of insect segmentation by a conserved sequence of transcription factors. *Development*, dev.155580.

Clark, E., Peel, A. D., & Akam, M. (2019). Arthropod segmentation. *Development*, *146*(18), dev170480.

Copf, T., Schröder, R., & Averof, M. (2004). Ancestral role of caudal genes in axis elongation and segmentation. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(51), 17711–17715.

da Silva, S. M., & Vincent, J.-P. (2007). Oriented cell divisions in the extending germband of Drosophila. *Development*, *134*(17), 3049–3054.

Davis, G. K., & Patel, N. H. (2002). Short, long, and beyond: Molecular and embryological approaches to insect segmentation. *Annual Review of Entomology*, *47*, 669–699.

de las Heras, J. M., & Casanova, J. (2006). Spatially distinct downregulation of Capicua repression and tailless activation by the Torso RTK pathway in the Drosophila embryo. *Mechanisms of Development*, *123*(6), 481–486.

Demerec, M. (1950). *Biology of Drosophila*. New York: John Wiley; Sons, Inc.

Diaz, R. J., Harbecke, R., Singer, J. B., Pignoni, F., Janning, W., & Lengyel, J. A. (1996). Graded effect of tailless on posterior gut development: Molecular basis of an allelic series of a nuclear receptor gene. *Mechanisms of Development*, *54*(1), 119–130.

DiNardo, S. (1985). Development of embryonic pattern in D. melanogaster as revealed by accumulation of the nuclear engrailed protein. *Cell*, *43*(1), 59–69.

DiNardo, S., & O'Farrell, P. H. (1987). Establishment and refinement of segmental pattern in the Drosophila embryo: Spatial control of engrailed expression by pair-rule genes. *Genes & Development*, *1*(10), 1212–1225.

Doe, C. Q. (1992). Molecular markers for identified neuroblasts and ganglion mother cells in the Drosophila central nervous system. *Development*, *116*(4), 855–863.

Duffy, J. B., & Perrimon, N. (1994). The Torso Pathway in Drosophila: Lessons on Receptor Tyrosine Kinase Signaling and Pattern Formation. *Developmental Biology*, *166*(2), 380–395.

Duncan, E. J., Benton, M. A., & Dearden, P. K. (2013). Canonical terminal patterning is an evolutionary novelty. *Developmental Biology*, *377*(1), 245–261.

El-Sherif, E., Zhu, X., Fu, J., & Brown, S. J. (2014). Caudal Regulates the Spatiotemporal Dynamics of Pair-Rule Waves in Tribolium. *PLoS Genetics*, *10*(10), e1004677.

Farkas, G., Gausz, J., Galloni, M., Reuter, G., Gyurkovics, H., & Karch, F. (1994). The Trithorax-like gene encodes the Drosophila GAGA factor. *Nature*, *371*(6500), 806–808.

Fjose, A., McGinnis, W. J., & Gehring, W. J. (1985). Isolation of a homoeo box-containing gene from the engrailed region of Drosophila and the spatial distribution of its transcripts. *Nature*, *313*(6000), 284–289.

Foe, V. E. (1989). Mitotic domains reveal early commitment of cells in Drosophila embryos. *Development*, *107*(1), 1–22.

Frasch, M., Hoey, T., Rushlow, C., Doyle, H., & Levine, M. (1987). Characterization and localization of the even-skipped protein of Drosophila. *The EMBO journal*, *6*(3), 749–759.

Freeland, D., & Kuhn, D. (1996). Expression patterns of developmental genes reveal segment and parasegment organization of D. melanogaster genital discs. *Mechanisms of Development*, *56*(1-2), 61–72.

García-Solache, M., Jaeger, J., & Akam, M. (2010). A systematic analysis of the gap gene system in the moth midge Clogmia albipunctata. *Developmental Biology*, *344*(1), 306–318.

Gaskill, M. M., Gibson, T. J., Larson, E. D., & Harrison, M. M. (2021). GAF is essential for zygotic genome activation and chromatin accessibility in the early Drosophila embryo. *eLife*, *10*, e66668.

Gaul, U., & Weigel, D. (1991). Regulation of Krüppel expression in the anlage of the Malpighian tubules in the Drosophila embryo. *Mechanisms of Development*, *33*, 57–67.

Goldstein, R. E. (1999). Huckebein repressor activity in Drosophila terminal patterning is mediated by Groucho. *Development*, *126*(17), 3747–55.

Goyal, Y., Schüpbach, T., & Shvartsman, S. Y. (2018). A quantitative model of developmental RTK signaling. *Developmental Biology*, *442*(1), 80–86.

Graham, V., Khudyakov, J., Ellis, P., & Pevny, L. (2003). SOX2 functions to maintain neural progenitor identity. *Neuron*, *39*(5), 749–765.

Greenwood, S., & Struhl, G. (1997). Different levels of Ras activity can specify distinct transcriptional and morphological consequences in early Drosophila embryos. *Development*, *124*(23), 4879–4886.

Grossniklaus, U., Pearson, R. K., & Gehring, W. J. (1992). The Drosophila sloppy paired locus encodes two proteins involved in segmentation that show homology to mammalian transcription factors. *Genes & Development*, *6*(6), 1030–1051.

Gutjahr, T., Patel, N. H., Li, X., Goodman, C. S., & Noll, M. (1993). Analysis of the gooseberry locus in Drosophila embryos: Gooseberry determines the cuticular pattern and activates gooseberry neuro. *Development*, *118*(1), 21–31.

Häder, T., La Rosée, A., Ziebold, U., Busch, M., Taubert, H., Jäckle, H., & Rivera-Pomar, R. (1998). Activation of posterior pair-rule stripe expression in response to maternal caudal and zygotic knirps activities. *Mechanisms of Development*, *71*(1-2), 177–186.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., Del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362.

Harrison, M. M., Li, X.-Y., Kaplan, T., Botchan, M. R., & Eisen, M. B. (2011). Zelda binding in the early Drosophila melanogaster embryo marks regions subsequently activated at the maternal-to-zygotic transition. *PLoS genetics*, *7*(10), e1002266.

Harrison, S. D., & Travers, A. A. (1990). The tramtrack gene encodes a Drosophila finger protein that interacts with the ftz transcriptional regulatory region and shows a novel embryonic expression pattern. *The EMBO journal*, *9*(1), 207–216.

Hoch, M., & Pankratz, M. J. (1996). Control of gut development by fork head and cell signaling molecules in Drosophila. *Mechanisms of Development*, *58*(1-2), 3–14.

Hoey, T., Doyle, H. J., Harding, K., Wedeen, C., & Levine, M. (1986). Homeo box gene expression in anterior and posterior regions of the Drosophila embryo. *Proceedings of the National Academy of Sciences*, *83*(13), 4809–4813.

Hou, X. S., Melnick, M. B., & Perrimon, N. (1996). Marelle acts downstream of the Drosophila HOP/JAK kinase and encodes a protein similar to the mammalian STATs. *Cell*, *84*(3), 411–419.

Houtmeyers, R., Souopgui, J., Tejpar, S., & Arkell, R. (2013). The ZIC gene family encodes multi-functional proteins essential for patterning and morphogenesis. *Cellular and molecular life sciences: CMLS*, *70*(20), 3791–3811.

Hughes, C. L., & Kaufman, T. C. (2002). Hox genes and the evolution of the arthropod body plan1. *Evolution & Development*, *4*(6), 459–499.

Ingham, P., Martinez-Arias, A., Lawrence, P. A., & Howard, K. (1985). Expression of engrailed in the parasegment of Drosophila. *Nature*, *317*(6038), 634–636.

Jaeger, J. (2011). The gap gene network. *Cellular and molecular life sciences: CMLS*, *68*(2), 243–274.

Jaeger, J., Surkova, S., Blagov, M., Janssens, H., Kosman, D., Kozlov, K. N., Manu, n., Myasnikova, E., Vanario-Alonso, C. E., Samsonova, M., Sharp, D. H., & Reinitz, J. (2004). Dynamic control of positional information in the early Drosophila embryo. *Nature*, *430*(6997), 368–371.

Janssen, R. (2014). Gene expression suggests double-segmental and single-segmental patterning mechanisms during posterior segment addition in the beetle Tribolium castaneum. *The International Journal of Developmental Biology*, *58*(5), 343–347.

Janssens, H., Crombach, A., Richard Wotton, K., Cicin-Sain, D., Surkova, S., Lu Lim, C., Samsonova, M., Akam, M., & Jaeger, J. (2013). Lack of tailless leads to an increase in expression variability in Drosophila embryos. *Developmental Biology*, *377*(1), 305–317.

Jarial, M. S. (1987). Ultrastructure of the anal organ of drosophila larva with reference to ion transport. *Tissue & Cell*, *19*, 559–75.

Jaynes, J. B., & Fujioka, M. (2004). Drawing lines in the sand: Even skipped et al. and parasegment boundaries. *Developmental Biology*, *269*(2), 609–622.

Johnson, H. E., Djabrayan, N. J., Shvartsman, S. Y., & Toettcher, J. E. (2020). Optogenetic Rescue of a Patterning Mutant. *Current Biology*, S0960982220309118.

Johnson, H. E., & Toettcher, J. E. (2019). Signaling Dynamics Control Cell Fate in the Early Drosophila Embryo. *Developmental Cell*, *48*(3), 361–370.e3.

Joshi, P., Darr, A. J., & Skromne, I. (2019). CDX4 regulates the progression of neural maturation in the spinal cord. *Developmental Biology*, *449*(2), 132–142.

Jürgens, G., Wieschaus, E., Nüsslein-Volhard, C., & Kluding, H. (1984). Mutations affecting the pattern of the larval cuticle inDrosophila melanogaster. *Wilhelm Roux's archives of developmental biology*, *193*(5), 283–295.

Jürgens, G. (1987). Segmental organisation of the tail region in the embryo of Drosophila melanogaster: A blastoderm fate map of the cuticle structures of the larval tail region. *Roux's Archives of Developmental Biology*, *196*(3), 141–157.

Jürgens, G., & Weigel, D. (1988). Terminal versus segmental development in the Drosophila embryo: The role of the homeotic gene fork head. *Roux's Archives of Developmental Biology*, *197*(6), 345–354.

Keenan, S. E., Blythe, S. A., Marmion, R. A., Djabrayan, N. J.-V., Wieschaus, E. F., & Shvartsman, S. Y. (2020). Rapid Dynamics of Signal-Dependent Transcriptional Repression by Capicua. *Developmental Cell*, *52*(6), 794–801.e4.

Keränen, S. V. E., Fowlkes, C. C., Luengo Hendriks, C. L., Sudar, D., Knowles, D. W., Malik, J., & Biggin, M. D. (2006). Three-dimensional morphology and gene expression in the Drosophila blastoderm at cellular resolution II: Dynamics. *Genome Biology*, *7*(12), R124.

Kittelmann, S., Ulrich, J., Posnien, N., & Bucher, G. (2013). Changes in anterior head patterning underlie the evolution of long germ embryogenesis. *Developmental Biology*, *374*(1), 174–184.

Klingler, M., Erdélyi, M., Szabad, J., & Nüsslein-Volhard, C. (1988). Function of torso in determining the terminal anlagen of the Drosophila embryo. *Nature*, *335*(6187), 275–277.

Klomp, J., Athy, D., Kwan, C. W., Bloch, N. I., Sandmann, T., Lemke, S., & Schmidt-Ott, U. (2015). A cysteine-clamp gene drives embryo polarity in the midge Chironomus. *Science*, *348*(6238), 1040–1042.

Kornberg, T., Sidén, I., O'Farrell, P., & Simon, M. (1985). The engrailed locus of Drosophila: In situ localization of transcripts reveals compartment-specific expression. *Cell*, *40*(1), 45–53.

Koromila, T., Gao, F., Iwasaki, Y., He, P., Pachter, L., Gergen, J. P., & Stathopoulos, A. (2020). Odd-paired is a pioneer-like factor that coordinates with Zelda to control gene expression in embryos. *eLife*, *9*, e59610.

Kosman, D., Small, S., & Reinitz, J. (1998). Rapid preparation of a panel of polyclonal antibodies to Drosophila segmentation proteins. *Development Genes and Evolution*, *208*(5), 290–294.

Kuhn, D. T., Chaverri, J. M., Persaud, D. A., & Madjidi, A. (2000). Pair-rule genes cooperate to activate en stripe 15 and refine its margins during germ band elongation in the D. melanogaster embryo. *Mechanisms of Development*, 4.

Kuhn, D., Sawyer, M., Packert, G., Turenchalk, G., Mack, J. A., Sprey, T. E., Gustavson, E., & Kornberg, T. B. (1992). Development of the D. melanogaster caudal segments involves suppression of the ventral regions of A8, A9 and A10. *Development*, *116*, 11–20.

Kuhn, D., Turenchalk, G., Mack, J., Packert, G., & Kornberg, T. (1995). Analysis of the genes involved in organizing the tail segments of the Drosophila melanogaster embryo. *Mechanisms of Development*, *53*(1), 3–13.

La Rosee, A. (1997). Mechanism and Bicoid-dependent control of hairy stripe 7 expression in the posterior region of the Drosophila embryo. *The EMBO Journal*, *16*(14), 4403–4411.

Larkin, A., Marygold, S. J., Antonazzo, G., Attrill, H., Dos Santos, G., Garapati, P. V., Goodman, J. L., Gramates, L. S., Millburn, G., Strelets, V. B., Tabone, C. J., Thurmond, J., & FlyBase Consortium. (2021). FlyBase: Updates to the Drosophila melanogaster knowledge base. *Nucleic Acids Research*, *49*(D1), D899–D907.

Lawrence, P. A., Johnston, P., Macdonald, P., & Struhl, G. (1987). Borders of parasegments in Drosophila embryos are delimited by the fushi tarazu and even-skipped genes. *Nature*, *328*(6129), 440–442.

Lawrence, P. A., Martinez-Arias, A., Gardner, R. L., & Lawrence, P. A. (1985). The cell lineage of segments and parasegments in Drosophila. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *312*(1153), 83–90.

Lee, J. J., von Kessler, D. P., Parks, S., & Beachy, P. A. (1992). Secretion and localized transcription suggest a role in positional signaling for products of the segmentation gene hedgehog. *Cell*, *71*(1), 33–50.

Lemke, S., Busch, S. E., Antonopoulos, D. A., Meyer, F., Domanus, M. H., & Schmidt-Ott, U. (2010). Maternal activation of gap genes in the hover fly Episyrphus. *Development*, *137*(15), 2604–2604.

Levine, M., Harding, K., Wedeen, C., Doyle, H., Hoey, T., & Radomska, H. (1985). Expression of the Homeo Box Gene Family in Drosophila. *Cold Spring Harbor Symposia on Quantitative Biology*, *50*(0), 209–222.

Li, W. X. (2005). Functions and mechanisms of receptor tyrosine kinase Torso signaling: Lessons from Drosophila embryonic terminal development. *Developmental Dynamics*, *232*(3), 656–672.

Li, X.-y., MacArthur, S., Bourgon, R., Nix, D., Pollard, D. A., Iyer, V. N., Hechmer, A., Simirenko, L., Stapleton, M., Luengo Hendriks, C. L., Chu, H. C., Ogawa, N., Inwood, W., Sementchenko, V., Beaton, A., Weiszmann, R., Celniker, S. E., Knowles, D. W., Gingeras, T., ... Biggin, M. D. (2008). Transcription factors bind thousands of active and inactive regions in the Drosophila blastoderm. *PLoS biology*, *6*(2), e27.

Liang, H.-L., Nien, C.-Y., Liu, H.-Y., Metzstein, M. M., Kirov, N., & Rushlow, C. (2008). The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. *Nature*, *456*(7220), 400–403.

Lim, B., Fukaya, T., Heist, T., & Levine, M. (2018). Temporal dynamics of pair-rule stripes in living Drosophila embryos. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(33), 8376–8381.

Lohs-Schardin, M., Cremer, C., & Nüsslein-Volhard, C. (1979). A fate map for the larval epidermis of Drosophila melanogaster: Localized cuticle defects following irradiation of the blastoderm with an ultraviolet laser microbeam. *Developmental Biology*, *73*(2), 239–255.

Lord, P. C. W., Lin, M.-H., Hales, K. H., & Storti, R. V. (1995). Normal Expression and the Effects of Ectopic Expression of the Drosophila muscle segment homeobox (msh) Gene Suggest a Role in Differentiation and Patterning of Embryonic Muscles. *Developmental Biology*, *171*(2), 627–640.

Lu, X., Li, J. M., Elemento, O., Tavazoie, S., & Wieschaus, E. F. (2009). Coupling of zygotic transcription to mitotic control at the Drosophila mid-blastula transition. *Development*, *136*(12), 2101–2110.

Luengo Hendriks, C. L., Keränen, S. V. E., Fowlkes, C. C., Simirenko, L., Weber, G. H., DePace, A. H., Henriquez, C., Kaszuba, D. W., Hamann, B., Eisen, M. B., Malik, J., Sudar, D., Biggin, M. D., & Knowles, D. W. (2006). Three-dimensional morphology and gene expression in the Drosophila blastoderm at cellular resolution I: Data acquisition pipeline. *Genome Biology*, *7*(12), R123.

Lynch, J. A., Olesnicky, E. C., & Desplan, C. (2006). Regulation and function of tailless in the long germ wasp Nasonia vitripennis. *Development Genes and Evolution*, *216*(7-8), 493–498.

Ma, Y., Niemitz, E. L., Nambu, P. A., Shan, X., Sackerson, C., Fujioka, M., Goto, T., & Nambu, J. R. (1998). Gene regulatory functions of Drosophila fish-hook, a high mobility group domain Sox protein. *Mechanisms of Development*, *73*(2), 169–182.

MacArthur, S., Li, X.-Y., Li, J., Brown, J. B., Chu, H. C., Zeng, L., Grondona, B. P., Hechmer, A., Simirenko, L., Keränen, S. V. E., Knowles, D. W., Stapleton, M., Bickel, P., Biggin, M. D., & Eisen, M. B. (2009). Developmental roles of 21 Drosophila transcription factors are determined by quantitative differences in binding to an overlapping set of thousands of genomic regions. *Genome Biology*, *10*(7), R80.

Macdonald, P. M., Ingham, P., & Struhl, G. (1986). Isolation, structure, and expression of even-skipped: A second pair-rule gene of Drosophila containing a homeo box. *Cell*, *47*(5), 721–734.

Macdonald, P. M., & Struhl, G. (1986). A molecular gradient in early Drosophila embryos and its role in specifying the body pattern. *Nature*, *324*(6097), 537–545.

Mahoney, P. A., & Lengyel, J. A. (1987). The zygotic segmentation mutant tailless alters the blastoderm fate map of the Drosophila embryo. *Developmental Biology*, *122*(2), 464–470.

Martinez-Arias, A., & Lawrence, P. A. (1985). Parasegments and compartments in the Drosophila embryo. *Nature*, *313*(6004), 639–642.

Matsuda, R. (1976). *Morphology and Evolution of the Insect Abdomen: With Special Reference to Developmental Patterns and Their Bearings Upon Systematics*. Pergamon Press.

McDaniel, S. L., Gibson, T. J., Schulz, K. N., Fernandez Garcia, M., Nevil, M., Jain, S. U., Lewis, P. W., Zaret, K. S., & Harrison, M. M. (2019). Continued Activity of the Pioneer Factor Zelda Is Required to Drive Zygotic Genome Activation. *Molecular Cell*, *74*(1), 185–195.e4.

McKinney, W. (2010). Data Structures for Statistical Computing in Python. In S. van der Walt & J. Millman (Eds.), *Proceedings of the 9th Python in Science Conference* (pp. 56–61).

Meijering, E. H., Niessen, W. J., & Viergever, M. A. (2001). Quantitative evaluation of convolution-based methods for medical image interpolation. *Medical Image Analysis*, *5*(2), 111–126.

Merzdorf, C. S. (2007). Emerging roles for zic genes in early development. *Developmental Dynamics: An Official Publication of the American Association of Anatomists*, *236*(4), 922–940.

Mlodzik, M., Fjose, A., & Gehring, W. J. (1985). Isolation of caudal, a Drosophila homeo box-containing gene with maternal expression, whose transcripts form a concentration gradient at the pre-blastoderm stage. *The EMBO Journal*, *4*(11), 2961–2969.

Mlodzik, M., & Gehring, W. J. (1987a). Expression of the caudal gene in the germ line of Drosophila: Formation of an RNA and protein gradient during early embryogenesis. *Cell*, *48*(3), 465–478.

Mlodzik, M., & Gehring, W. J. (1987b). Hierarchy of the genetic interactions that specify the anteroposterior segmentation pattern of the Drosophila embryo as monitored by caudal protein expression. *Development*, *101*, 421–435.

Mohler, J., & Vani, K. (1992). Molecular organization and embryonic expression of the hedgehog gene involved in cell-cell communication in segmental patterning of Drosophila. *Development*, *115*(4), 957–971.

Mohler, J. (1995). Spatial regulation of segment polarity gene expression in the anterior terminal region of the Drosophila blastoderm embryo. *Mechanisms of Development*, *50*(2-3), 151–161.

Moran, E., & Jimenez, G. (2006). The Tailless Nuclear Receptor Acts as a Dedicated Repressor in the Early Drosophila Embryo. *Molecular and Cellular Biology*, *26*(9), 3446–3454.

Moshe, A., & Kaplan, T. (2017). Genome-wide search for Zelda-like chromatin signatures identifies GAF as a pioneer factor in early fly development. *Epigenetics & Chromatin*, *10*(1), 33.

Nambu, P. A., & Nambu, J. R. (1996). The Drosophila fish-hook gene encodes a HMG domain protein essential for segmentation and CNS development. *Development*, *122*(11), 3467–3475.

Nasiadka, A., & Krause, H. M. (1999). Kinetic analysis of segmentation gene interactions in Drosophila embryos. *Development*, *126*(7), 1515–1526.

Nasiadka, A., Dietrich, B. H., & Krause, H. M. (2002). Anterior-posterior patterning in the Drosophila embryo. *Advances in Developmental Biology and Biochemistry* (pp. 155–204).

Nien, C.-Y., Liang, H.-L., Butcher, S., Sun, Y., Fu, S., Gocha, T., Kirov, N., Manak, J. R., & Rushlow, C. (2011). Temporal coordination of gene networks by Zelda in the early Drosophila embryo. *PLoS genetics*, *7*(10), e1002339.

Nusslein-Volhard, C., Frohnhofer, H. G., & Lehmann, R. (1987). Determination of anteroposterior polarity in Drosophila. *Science*, *238*(4834), 1675–1681.

Olesnicky, E. C., Brent, A. E., Tonnes, L., Walker, M., Pultz, M. A., Leaf, D., & Desplan, C. (2006). A caudal mRNA gra-

dient controls posterior development in the wasp Nasonia. *Development*, *133*(20), 3973–3982.

Perkins, L. A., & Perrimon, N. (1991). The molecular genetics of tail development in Drosophila melanogaster. *In Vivo*, *5*(5), 521–531.

Pignoni, F., Steingrimsson, E., & JA, L. (1992). Bicoid and the terminal system activate tailless expression in the early Drosophila embryo. *Development*, *115*, 239–251.

Pignoni, F., Baldarelli, R. M., Steingrimsson, E., Diaz, R. J., Patapoutian, A., Merriam, J. R., & Lengyel, J. A. (1990). The Drosophila gene tailless is expressed at the embryonic termini and is a member of the steroid receptor superfamily. *Cell*, *62*(1), 151–163.

Read, D., Levine, M., & Manley, J. L. (1992). Ectopic expression of the Drosophila tramtrack gene results in multiple embryonic defects, including repression of even-skipped and fushi tarazu. *Mechanisms of Development*, *38*(3), 183–195.

Reinitz, J., & Levine, M. (1990). Control of the initiation of homeotic gene expression by the gap genes giant and tailless in Drosophila. *Developmental Biology*, *140*(1), 57–72.

Rivera-Pomar, R., Lu, X., Perrimon, N., Taubert, H., & Jäckle, H. (1995). Activation of posterior gap gene expression in the Drosophila blastoderm. *Nature*, *376*(6537), 253–256.

Russell, S. R., Sanchez-Soriano, N., Wright, C. R., & Ashburner, M. (1996). The Dichaete gene of Drosophila melanogaster encodes a SOX-domain protein required for embryonic segmentation. *Development*, *122*(11), 3669–3676.

Sackerson, C., Fujioka, M., & Goto, T. (1999). The even-skipped locus is contained in a 16-kb chromatin domain. *Developmental Biology*, *211*(1), 39–52.

Sato, T., & Denell, R. E. (1986). Segmental identity of caudal cuticular features of Drosophila melanogaster larvae and its control by the bithorax complex. *Developmental Biology*, *116*(1), 78–91.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: An open-source platform for biological-image analysis. *Nature Methods*, *9*(7), 676–682.

Schmidt-Ott, U., Sander, K., & Technau, G. M. (1994). Expression of engrailed in embryos of a beetle and five dipteran species with special reference to the terminal regions. *Roux's Archives of Developmental Biology*, *203*(6), 298–303.

Schmied, C., & Jambor, H. K. (2021). *Effective image visualization for publications – a workflow using open access tools and concepts* (tech. rep. 9:1373). F1000Research.

Schoppmeier, M., & Schröder, R. (2005). Maternal Torso Signaling Controls Body Axis Elongation in a Short Germ Insect. *Current Biology*, *15*(23), 2131–2136.

Schroder, R., Eckert, C., Wolff, C., & Tautz, D. (2000). Conserved and divergent aspects of terminal patterning in the beetle Tribolium castaneum. *Proceedings of the National Academy of Sciences*, *97*(12), 6591–6596.

Schroeder, M. D., Greer, C., & Gaul, U. (2011). How to make stripes: Deciphering the transition from non-periodic to periodic patterns in Drosophila segmentation. *Development*, *138*(14), 3067–3078.

Schulz, C., Schröder, R., Hausdorf, B., Wolff, C., & Tautz, D. (1998). A caudal homologue in the short germ band beetle Tribolium shows similarities to both, the Drosophila and the vertebrate caudal expression patterns. *Development Genes and Evolution*, *208*(5), 283–289.

Schulz, C., & Tautz, D. (1995). Zygotic caudal regulation by hunchback and its role in abdominal segment formation of the Drosophila embryo. *Development*, *121*, 1023–1028.

Selva, E. M., & Stronach, B. E. (2007). Germline clone analysis for maternally acting Drosophila hedgehog components. *Methods in Molecular Biology (Clifton, N.J.)*, *397*, 129–144.

Singer, J. B., Harbecke, R., Kusch, T., Reuter, R., & Lengyel, J. A. (1996). Drosophila brachyenteron regulates gene activity and morphogenesis in the gut. *Development*, *22*, 3707–18.

Smits, C. M., & Shvartsman, S. Y. (2020). The design and logic of terminal patterning in Drosophila. *Current Topics in Developmental Biology* (pp. 193–217).

Snodgrass, R. E. (1935). *Principles of Insect Morphology*. McGraw-Hill Book Company, Incorporated.

Soluri, I. V., Zumerling, L. M., Payan Parra, O. A., Clark, E. G., & Blythe, S. A. (2020). Zygotic pioneer factor activity of Odd-paired/Zic is necessary for late function of the Drosophila segmentation network. *eLife*, *9*, e53916.

Soriano, N. S., & Russell, S. (1998). The Drosophila SOX-domain protein Dichaete is required for the development of the central nervous system midline. *Development*, *125*(20), 3989–3996.

Spirov, A. V., Timakin, D. L., Reinitz, J., & Kosman, D. (2000). Experimental Determination of Drosophila Embryonic Coordinates by Genetic Algorithms, the Simplex Method, and Their Hybrid. In G. Goos, J. Hartmanis, J. van Leeuwen, & S. Cagnoni (Eds.), *Real-World Applications of Evolutionary Computing* (pp. 97–106). Springer Berlin Heidelberg.

Spirov, A. V., Vanario-Alonso, C. E., Spirova, E. N., & Holloway, D. M. (2013). Experimental Determination of Intrinsic Drosophila Embryo Coordinates by Evolutionary Computation. In A. Ngom, E. Formenti, J.-K. Hao, X.-M. Zhao, & T. van Laarhoven (Eds.), *Pattern Recognition in Bioinformatics* (pp. 126–137). Springer Berlin Heidelberg.

Sprenger, F., Stevens, L. M., & Nüsslein-Volhard, C. (1989). The Drosophila gene torso encodes a putative receptor tyrosine kinase. *Nature*, *338*(6215), 478–483.

St Johnston, D. (2002). The art and design of genetic screens: Drosophila melanogaster. *Nature Reviews. Genetics*, *3*(3), 176–188.

Stevanovic, M., Drakulic, D., Lazic, A., Ninkovic, D. S., Schwirtlich, M., & Mojsin, M. (2021). SOX Transcription Factors as Important Regulators of Neuronal and Glial Differentiation During Nervous System Develop-

ment and Adult Neurogenesis. *Frontiers in Molecular Neuroscience*, *14*, 654031.

Strecker, T. R., Kongsuwan, K., Lengyel, J. A., & Merriam, J. R. (1986). The zygotic mutant tailless affects the anterior and posterior ectodermal regions of the Drosophila embryo. *Developmental Biology*, *113*(1), 64–76.

Strecker, T. R., Merriam, J. R., & Lengyel, J. A. (1988). Graded requirement for the zygotic terminal gene, tailless, in the brain and tail region of the Drosophila embryo. *Development*, *102*, 721–34.

Sullivan, W., Ashburner, M., & Hawley, R. S. (2000). *Drosophila protocols*. Cold Spring Harbor Laboratory Press.

Surkova, S., Kosman, D., Kozlov, K., Manu, Myasnikova, E., Samsonova, A. A., Spirov, A., Vanario-Alonso, C. E., Samsonova, M., & Reinitz, J. (2008). Characterization of the Drosophila segment determination morphome. *Developmental Biology*, *313*(2), 844–862.

Tabata, T., Eaton, S., & Kornberg, T. B. (1992). The Drosophila hedgehog gene is expressed specifically in posterior compartment cells and is a target of engrailed regulation. *Genes & Development*, *6*(12B), 2635–2645.

Tashiro, S., Michiue, T., Higashijima, S., Zenno, S., Ishimaru, S., Takahashi, F., Orihara, M., Kojima, T., & Saigo, K. (1993). Structure and expression of hedgehog, a Drosophila segment-polarity gene required for cell-cell communication. *Gene*, *124*(2), 183–189.

Trivedi, V., Choi, H. M. T., Fraser, S. E., & Pierce, N. A. (2018). Multidimensional quantitative analysis of mRNA expression within intact vertebrate embryos. *Development*, *145*(1), dev156869.

Tsurumi, A., Xia, F., Li, J., Larson, K., LaFrance, R., & Li, W. X. (2011). STAT is an essential activator of the zygotic genome in the early Drosophila embryo. *PLoS genetics*, *7*(5), e1002086.

Turner, F. R., & Mahowald, A. P. (1979). Scanning electron microscopy of Drosophila melanogaster embryogenesis: III. Formation of the head and caudal segments. *Developmental Biology*, *68*(1), 96–109.

van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., Yu, T., & scikit-image contributors. (2014). Scikit-image: Image processing in Python. *PeerJ*, *2*, e453.

van Rooijen, C., Simmini, S., Bialecka, M., Neijts, R., van de Ven, C., Beck, F., & Deschamps, J. (2012). Evolutionarily conserved requirement of Cdx for post-occipital tissue emergence. *Development*, *139*(14), 2576–2583.

Vincent, B. J., Staller, M. V., Lopez-Rivera, F., Bragdon, M. D. J., Pym, E. C. G., Biette, K. M., Wunderlich, Z., Harden, T. T., Estrada, J., & DePace, A. H. (2018). Hunchback is counter-repressed to regulate even-skipped stripe 2 expression in Drosophila embryos. *PLoS genetics*, *14*(9), e1007644.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., . . . SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272.

Weigel, D., Jurgens, G., Klingler, M., & Jackle, H. (1990). Two gap genes mediate maternal terminal pattern information in Drosophila. *Science*, *248*(4954), 495–498.

Weigel, D., Jürgens, G., Küttner, F., Seifert, E., & Jäckle, H. (1989). The homeotic gene fork head encodes a nuclear protein and is expressed in the terminal regions of the Drosophila embryo. *Cell*, *57*(4), 645–658.

Weisbrod, A., Cohen, M., & Chipman, A. D. (2013). Evolution of the insect terminal patterning system—Insights from the milkweed bug, Oncopeltus fasciatus. *Developmental Biology*, *380*(1), 125–131.

Williams, T. A., & Nagy, L. M. (2017). Linking gene regulation to cell behaviors in the posterior growth zone of sequentially segmenting arthropods. *Arthropod Structure & Development*, *46*(3), 380–394.

Wilson, M. J., & Dearden, P. K. (2009). Tailless patterning functions are conserved in the honeybee even in the absence of Torso signaling. *Developmental Biology*, *335*(1), 276–287.

Wu, L. H., & Lengyel, J. A. (1998). Role of caudal in hindgut specification and gastrulation suggests homology between Drosophila amnioproctodeal invagination and vertebrate blastopore. *Development*, *125*, 2433–42.

Yan, R., Small, S., Desplan, C., Dearolf, C. R., & Darnell, J. E. (1996). Identification of a Stat gene that functions in Drosophila development. *Cell*, *84*(3), 421–430.

Zhong, Y., Herrera-Úbeda, C., Garcia-Fernàndez, J., Li, G., & Holland, P. W. H. (2020). Mutation of amphioxus Pdx and Cdx demonstrates conserved roles for ParaHox genes in gut, anus and tail patterning. *BMC Biology*, *18*(1), 68.