

1 **The chromatin factor Gon4l regulates embryonic axis extension by promoting mediolateral cell polarity**
2 **and notochord boundary formation through negative regulation of cell adhesion**

3
4 Margot L K Williams¹, Atsushi Sawada^{1,2}, Terin Budine¹, Chunyue Yin^{2,3}, Paul Gontarz¹, and Lilianna Solnica-
5 Krezel^{1,2*}

6
7 ¹Department of Developmental Biology, Washington University School of Medicine, Saint Louis, MO 63110,
8 USA.

9 ²Department of Biological Sciences, Vanderbilt University, Nashville, TN 37235, USA

10 ³Current address: Division of Pediatric Gastroenterology, Hepatology, and Nutrition, Cincinnati Children's
11 Hospital, Cincinnati, OH 45229, USA.

12
13 **Anteroposterior axis extension during vertebrate gastrulation requires cell proliferation, embryonic**
14 **patterning, and morphogenesis to be spatiotemporally coordinated, but the underlying genetic**
15 **mechanisms remain poorly understood. Here we define a role for the conserved chromatin factor**
16 **Gon4l, encoded by *ugly duckling* (*udu*), in coordinating tissue patterning and axis extension during**
17 **zebrafish gastrulation. Although identified as a recessive enhancer of short axis phenotypes in planar**
18 **cell polarity (PCP) mutants, we found that Gon4l functions in a genetically independent, partially**
19 **overlapping fashion with PCP signaling to regulate mediolateral cell polarity underlying axis extension**
20 **in part by promoting notochord boundary formation. We identified direct genomic targets of Gon4l and**
21 **found that it acts as both a positive and negative regulator of gene expression, including limiting**
22 **expression of the cell-cell and cell-matrix adhesion molecules EpCAM and Integrin α 3b. Excess *epcam***
23 **or *itga3b* in wild-type gastrulae phenocopied notochord boundary defects of *udu* mutants, while**
24 **downregulation of *itga3b* suppressed them. By promoting formation of this anteroposteriorly aligned**
25 **boundary and associated cell polarity, Gon4l cooperates with PCP signaling to coordinate**
26 **morphogenesis with the anteroposterior embryonic axis.**

27
28 Gastrulation is a critical period of animal development during which the three primordial germ layers -
29 ectoderm, mesoderm, and endoderm - are specified, patterned, and shaped into a rudimentary body plan.
30 During vertebrate gastrulation, mesoderm and endoderm become internalized to underlie ectoderm and the
31 three germ layers thin and expand through epibolic movements. The hallmark of the vertebrate body plan is an
32 elongated anteroposterior (AP) axis that emerges as the result of convergence and extension (C&E), a
33 conserved set of gastrulation movements characterized by the concomitant AP elongation and mediolateral
34 (ML) narrowing of the germ layers¹⁻⁴. C&E is accomplished by a combination of polarized cell behaviors,
35 including directed migration and ML intercalation behavior (MIB)⁵⁻⁷. During MIB, cells elongate and align their
36 bodies and protrusions in the ML dimension and intercalate preferentially between their anterior and posterior
37 neighbors⁷. This polarization of cell behaviors with respect to the AP axis is regulated by the planar cell polarity
38 (PCP) and other signaling pathways⁸⁻¹². Because these pathways are essential for MIB and C&E but do not
39 affect cell fates¹²⁻¹⁴, other mechanisms must spatiotemporally coordinate morphogenesis with embryonic
40 patterning to ensure normal development. BMP, for example, coordinates dorsal-ventral axis patterning with
41 morphogenetic movements by limiting expression of PCP signaling components and C&E to the embryo's

1 dorsal side¹⁵. In general, though, molecular mechanisms that coordinate gastrulation cell behaviors with axial
2 patterning are poorly understood, and remain one of the key outstanding questions in developmental biology.
3
4 Epigenetic regulators offer a potential mechanism by which broad networks of embryonic patterning and
5 morphogenesis genes can be co-regulated, conceivably by altering chromatin state within embryonic cells to
6 regulate which subsets of genes are available for transcription. Epigenetic modifiers like Histone acetyl
7 transferases (HATs), deacetylases (HDACs), and methyltransferases associate within protein complexes
8 containing chromatin factors that are thought to regulate their binding at specific genomic regions in context-
9 specific ways¹⁶. The identities, functions, and specificity of chromatin factors with roles during embryogenesis
10 are only now being elucidated, and often have described roles in cell fate specification and embryonic
11 patterning¹⁷⁻¹⁹. The contribution of epigenetic regulation to gastrulation cell movements, however, is particularly
12 poorly understood.

13
14 Here, we describe the chromatin factor Gon4l, encoded by *ugly duckling* (*udu*), as a novel regulator of
15 embryonic axis extension during zebrafish gastrulation. *udu* was identified in a forward genetic screen for
16 enhancers of short axis phenotypes in PCP mutants, but we found it functions in parallel to PCP signaling.
17 Instead, complete maternal and zygotic *udu* (*MZudu*) deficiency produced a distinct set of morphogenetic and
18 cell polarity phenotypes that implicate the notochord boundary in ML cell polarity and cell intercalation during
19 C&E. Extension defects in *MZudu* mutants were remarkably specific, as internalization, epiboly and
20 convergence gastrulation movements occurred normally. Gene expression profiling revealed that Gon4l
21 regulates expression of a large portion of the zebrafish genome, including housekeeping genes, patterning
22 genes, and many with known or potential roles in morphogenesis. Furthermore, Gon4l-associated genomic loci
23 were identified by DNA adenine methyltransferase identification^{20, 21} paired with high throughput sequencing
24 (DamID-seq), revealing both positive and negative regulation of putative direct targets by Gon4l.
25 Mechanistically, we found that increased expression of *epcam* and *itga3b*, direct targets of Gon4l-dependent
26 repression during gastrulation, were largely causative of notochord boundary defects in *MZudu* mutants. This
27 report thereby defines a critical role for a chromatin factor in the regulation of gastrulation cell behaviors in
28 vertebrate embryos: by ensuring proper formation of the AP-aligned notochord boundary and associated ML
29 cell polarity, Gon4l cooperates with PCP signaling to coordinate morphogenesis that extends the AP
30 embryonic axis.

31 32 **RESULTS**

33 34 **Gon4l is a novel regulator of axis extension during zebrafish gastrulation**

35 To identify novel regulators of C&E, we performed a three generation synthetic mutant screen^{22, 23} using
36 zebrafish carrying a hypomorphic allele of the planar cell polarity (PCP) gene *knypek* (*kny*)/*glypican 4*, *kny*^{m818}
37 ¹⁰. F0 wild-type (WT) males were mutagenized with *N*-ethyl, *N*-nitrosourea (ENU) and outcrossed to WT
38 females. The resulting F1 fish were outcrossed to *kny*^{m818/818} males (rescued by *kny* RNA injection) to generate
39 F2 families whose F3 offspring were screened during early segmentation (11-14 hours post fertilization (hpf))

1 and late embryogenesis (24hpf) for short axis phenotypes (Fig.1a). Screening nearly 100 F2 families yielded
2 eight recessive mutations that enhanced axis elongation defects in *kny*^{m818/818} embryos. *vu68* was found to be a
3 new L227P *kny* allele, and *vu64* was a new Y219* nonsense allele of the core PCP gene *trilobite(tri)/vangl2*,
4 mutations in which exacerbate *kny* mutant phenotypes²⁴, demonstrating effectiveness of our screening
5 strategy. We focused on *vu66/vu66* mutants, which displayed a pleiotropic phenotype at 24hpf, including
6 shortened AP axes, reduced tail fins, and heart edema (Fig.1d, Supplemental Fig.1b).

7
8 Employing simple sequence repeat mapping strategy^{25, 26}, we mapped the *vu66* mutation to a small region on
9 Chromosome 16 that contains the *ugly duckling (udu)* gene (Fig.1b). *udu* encodes a conserved chromatin
10 factor homologous to *gon4* in *C. elegans* and the closely related *Gon4l* in mammals²⁷. The previously
11 described *udu* mutant phenotypes resemble those of homozygous *vu66* embryos, including a shorter body axis
12 and fewer blood cells^{27, 28}, making *udu* an excellent candidate for this novel PCP enhancer. Sequencing cDNA
13 of the *udu* coding region from 24hpf *vu66/vu66* embryos revealed a T to A transversion at position 2,261
14 predicted to change 753Y to a premature STOP codon (Fig.1g). Furthermore, *vu66* failed to complement a
15 known *udu*^{sq1} allele²⁷ (Supplemental Fig.1c). Together, these data establish *vu66* as a new *udu* allele and
16 identify it as a recessive enhancer of axis extension defects in *kny* PCP mutant gastrulae.

18 **Complete loss of maternal and zygotic *udu* function impairs axial extension**

19 A majority of zebrafish genes are expressed maternally^{29, 30}, and their expression can mask the effect of
20 zygotic mutations on embryo patterning and gastrulation^{31, 32}. Because *udu* is highly maternally expressed²⁷,
21 we employed germline replacement³³ to generate otherwise WT females carrying *udu*^{vu66/vu66} (*udu*^{-/-}) germline,
22 and therefore lacking maternal *udu* expression. Females harboring *udu*^{vu66/vu66} germline were crossed to
23 *udu*^{vu66/+} or *udu*^{vu66/vu66} germline males to produce 50% or 100% embryos lacking both maternal (M) and zygotic
24 (Z) *udu* function, respectively, hereafter referred to as MZ*udu* mutants. Strict maternal loss of *udu* produced no
25 obvious phenotypes (Fig.4a). MZ*udu*^{-/-} embryos appeared outwardly to develop normally until mid-gastrula
26 stages (Fig.1i), specified the three germ layers (Fig.1m, Supplemental Fig.2), formed an embryonic shield
27 marked by *gsc* expression (Supplemental Fig.2c,d), and completed epiboly on schedule (Fig.1i). However, they
28 exhibited clear abnormalities at the onset of segmentation, as somites were largely absent in mutants
29 (Fig.1k,o). Although *myoD* expression was observed within adaxial cells by whole mount *in situ* hybridization
30 (WISH), its expression was not detected in nascent somites (Fig.1o). Formation of adaxial cells is consistent
31 with normal expression of their inducer *shh* in the axial mesoderm³⁴ (Supplemental Fig.2h), and *ntl/brachyury*
32 expression in the axial mesoderm was also largely intact (Supplemental Fig.2j). Importantly, MZ*udu*^{-/-} embryos
33 were markedly shorter than age-matched WT controls throughout segmentation (Fig.1k) and at 24hpf
34 (Supplemental Fig. 3e,h, Fig.4b). Although increased cell death was observed in MZ*udu* and *Zudu* mutants³⁵,
35 inhibiting apoptosis within MZ*udu* embryos via injection of RNA encoding the anti-apoptotic mitochondrial
36 protein Bcl-xL³⁶ did not suppress their short axis phenotype (Supplemental Fig.3a-f). MZ*udu* mutants displayed
37 a number of other phenotypes, including blood^{27, 28} and cardiac deficiencies, which will be described further

1 elsewhere (Budine, Williams, LSK, unpublished). These results demonstrate a role for Gon4l in axial extension
2 during gastrulation.

3

4 **Gon4l regulates formation of the notochord boundary**

5 Time-lapse Nomarski (Fig.2a-b) and confocal (Fig.2c-d) microscopy of dorsal mesoderm in *MZudu*^{-/-} gastrulae
6 revealed reduced definition and regularity of the boundary between axial and paraxial mesoderm, hereafter
7 referred to as the notochord boundary, compared to WT (Fig.2a-b, arrowheads). The ratio of the total/net
8 length of notochord boundaries was significantly higher in *MZudu*^{-/-} than in WT gastrulae at all time points
9 (Fig.2e), indicative of decreased straightness. Interestingly, Laminin was detected by immunostaining at the
10 notochord boundary of both WT and *MZudu*^{-/-} embryos (Fig.2f-g), indicating that *MZudu* mutants form a *bona*
11 *fide* boundary, albeit an irregular one. These results demonstrate that Gon4l is necessary for proper formation
12 of the notochord boundary during gastrulation.

13

14 **Mediolateral polarity and intercalation of axial mesoderm cells are reduced in *MZudu*^{-/-} gastrulae**

15 In vertebrate gastrulae, C&E is achieved chiefly through ML intercalation of polarized cells that elongate and
16 align their cell bodies with the ML embryonic axis^{5, 7, 9, 11}. To determine whether cell polarity defects underlie
17 reduced axis extension in *MZudu*^{-/-} gastrulae, we measured cell orientation (the angle of a cell's long axis with
18 respect to the embryo's ML axis; Fig.3b,e) and cell body elongation (length-to-width or aspect ratio (AR);
19 Fig.3c,f) in confocal time-lapse series of *MZudu*^{-/-} and WT gastrulae expressing membrane Cherry fluorescent
20 protein (mCherry). Given the irregular notochord boundaries observed in *MZudu*^{-/-} gastrulae (Fig.2d-e), we
21 examined the time course of cell polarization according to a cell's position with respect to the boundary: i.e.
22 boundary-adjacent "edge" cells were analyzed separately from those one or two cell diameters away (hereafter
23 -1 and -2, respectively), and so on (See Fig.3). Most WT axial mesoderm cells at 80% epiboly were largely ML
24 oriented and somewhat elongated, but boundary-adjacent "edge" cells were less well oriented (median angle=
25 24.6°) than cells within any of the internal rows (median angles= 15.7°, 16.6°, 21.1°; Fig.3a-c). However, at the
26 end of the gastrula period 90 minutes later, edge cells became highly aligned and elongated (median
27 angle=13.6°, mean AR=2.2) similar to internal cell rows (median angles= 12.1°, 18.4°; Fig.3a-c). All *MZudu*^{-/-}
28 axial mesoderm cells exhibited severely reduced ML alignment and elongation at 80% epiboly (Fig.3d-f), but
29 90 minutes later only the edge cells remained less aligned than WT (median angle=17.0°, mean AR=1.9)
30 (Fig.3e), although aspect ratio of the edge and -1 cells remained reduced (Fig. 3f). These results indicate
31 significantly reduced ML orientation of axial mesoderm cells in *MZudu*^{-/-} gastrulae, a defect that persisted only
32 in boundary-adjacent cells at late gastrulation. Importantly, this reduction in ML cell polarity was accompanied
33 by significantly fewer cell intercalation events within the axial mesoderm of *MZudu* mutants compared to WT
34 (Fig.3g-i). As ML intercalation is the key cellular behavior required for vertebrate C&E^{5, 7}, we conclude that this
35 is likely the primary cause of axial mesoderm extension defects in *MZudu* mutants. We also observed
36 significantly fewer mitoses in *MZudu*^{-/-} gastrulae (Fig.3j-l). Because decreased cell proliferation and the
37 resulting reduced number of axial mesoderm cells has been demonstrated to impair extension in zebrafish³⁷,
38 this could also be a contributing factor. Together, reduction of ML cell polarity, ML cell intercalations, and cell

1 proliferation provide a suite of mechanisms resulting in impaired axial extension in *MZudu*^{-/-} gastrulae.
2 Combined with irregular notochord boundaries observed in *MZudu* mutants, we hypothesize that this boundary
3 provides a ML orientation cue that contributes to ML polarization of axial mesoderm cells, and that this cue is
4 absent or reduced in *MZudu*^{-/-} gastrulae.

5

6 **Gon4l regulates axial mesoderm cell polarity and extension independent of PCP signaling**

7 Molecular control of ML cell polarity underlying C&E movements in vertebrate embryos is largely attributed to
8 planar cell polarity (PCP) and G α 12/13 signaling⁸⁻¹². While zygotic loss of *udu* enhanced axial extension
9 defects in *kny*^{m818/m818} PCP mutants (Fig.1) and axial mesoderm cells in *MZudu*^{-/-} gastrulae exhibited impaired
10 ML cell polarity and intercalation (Fig.3), it was unclear whether *udu* functions within or parallel to the PCP
11 network. To address this, we generated compound *MZudu*;*Zkny*^{fr6/fr6} (a nonsense/null *kny* allele¹⁰) mutants
12 utilizing germline replacement as described above. Strikingly, these compound mutant embryos were
13 substantially shorter than single *MZudu* or *kny*^{fr6/fr6} mutants (Fig.4d). Likewise, interference with *vangl2/tri*
14 function in *MZudu* mutants by injection of MO1-*vangl2* antisense morpholino oligonucleotide³⁸ also
15 exacerbated axis extension defect of *MZudu* mutants (Supplemental Fig.4a-d). That reduced levels of PCP
16 components *kny* or *tri* enhanced short axis phenotype resulting from complete *udu* deficiency provides
17 evidence that Gon4l affects axial extension via a parallel pathway. Furthermore, expression domains of genes
18 encoding Wnt/PCP signaling components *kny*, *tri*, *wnt5*, and *wnt11* in *MZudu*^{-/-} gastrulae were comparable to
19 WT (Supplemental Fig.4e-l). Finally, we found that the asymmetric intracellular localization of Prickle (Pk)-GFP,
20 a core PCP component, to anterior cell membranes in WT gastrulae during C&E movements^{32, 39} was not
21 affected in *MZudu*^{-/-} gastrulae (Fig.4e-f). This is consistent with intact PCP signaling in *MZudu* mutant
22 gastrulae, and provides further evidence that Gon4l functions largely in parallel to PCP to regulate ML cell
23 polarity and axis extension during gastrulation.

24

25 **A Gon4l-dependent boundary cue and PCP signaling cooperate to polarize axial mesoderm cells**

26 In addition to PCP signaling, notochord boundaries are required for proper C&E of the axial mesoderm in
27 ascidian embryos⁴⁰ and are involved in the polarization of intercalating cells during *Xenopus* gastrulation¹.
28 Boundary defects observed in *MZudu*^{-/-} gastrulae are not a common feature of mutants with reduced C&E,
29 however, as notochord boundaries in *kny*^{fr6/fr6} gastrulae were straighter than in WT siblings (Fig.4g). Consistent
30 with cell polarity defects previously reported in *kny* mutants¹⁰, all axial mesoderm cells failed to align ML within
31 *kny*^{-/-} embryos at 80% epiboly regardless their position relative to the notochord boundary (Fig.4h-j). After 90
32 minutes, however, *kny*^{-/-} cells in the edge (and to a lesser extent, -1) position attained distinct ML orientation
33 (median angle= 22.1°). Indeed, the nearer a *kny*^{-/-} cell was to the notochord boundary, the more ML aligned it
34 was likely to be (Fig.4i). This suggests that the notochord boundary provides a ML orientation cue that is
35 independent of PCP signaling and functions across approximately two cell diameters. Furthermore, this
36 boundary-associated cue appears to operate later in gastrulation, whereas PCP-dependent cell polarization is
37 evident by 80% epiboly. Importantly, distinct cell polarity phenotypes observed in *MZudu* and *kny*^{-/-} gastrulae
38 provide further evidence that Gon4l functions in parallel to PCP signaling.

1
2 To assess how PCP signaling interacts with the proposed Gon4l-dependent boundary cue, we examined the
3 polarity of axial mesoderm cells in compound zygotic *kny*^{-/-};*udu*^{-/-} mutant gastrulae (Fig.4k-m). As observed in
4 single *kny*^{-/-} mutant gastrulae¹⁰ (Fig.4h-j), both ML orientation and elongation of all axial mesoderm cells were
5 disrupted in double mutants at 80% epiboly, regardless of position with respect to the notochord boundary. By
6 late gastrulation however, edge cells in double mutant gastrulae failed to attain ML alignment as observed in
7 *kny*^{-/-} mutants, but instead remained largely randomly oriented (median angle= 38.0°) (Fig.4l). This
8 exacerbation of cell orientation defects was correlated with stronger axis extension defects in compound *kny*^{-/-}
9 ;*udu*^{-/-} compared to single *kny*^{-/-} mutants (Fig.1e). This supports our hypothesis that a Gon4l-dependent
10 boundary cue regulates ML alignment of axial mesoderm cells independent of PCP signaling, and that these
11 two mechanisms function in partially overlapping spatial and temporal domains to cooperatively polarize all
12 axial mesoderm cells and promote axial extension.

13

14 **Loss of Gon4l results in large-scale gene expression changes in zebrafish gastrulae**

15 As a nuclear-localized chromatin factor^{27, 41} (Supplemental Fig.3j), Gon4l is unlikely to influence
16 morphogenesis directly. To identify genes regulated by Gon4l with potential roles in morphogenesis, we
17 performed RNA sequencing (RNA-seq) in MZ*udu*^{-/-} and WT tailbud-stage embryos. Analysis of relative
18 transcript levels revealed that more than 11% of the genome was differentially expressed ($p_{adj} < 0.01$, ≥ 2 -fold
19 change) in MZ*udu* mutants (Fig. 5a). Of these ~2,950 differentially expressed genes, 1,692 exhibited increased
20 expression in MZ*udu* compared to 1,259 with decreased expression (Supplemental Table 1). Functional
21 annotation analysis revealed that genes downregulated in MZ*udu*^{-/-} gastrulae were enriched for ontology terms
22 related to chromatin structure, transcription, and translation (Fig.5e), while upregulated genes were enriched
23 for terms related to biosynthesis, metabolism, and protein modifications (Fig.5f). Notably, many of these
24 misregulated genes are considered to have “house-keeping” functions, in that they are essential for cell
25 survival. We further examined genes with plausible roles in morphogenesis, including those encoding signaling
26 and adhesion molecules, and found the majority of genes in both classes were upregulated in MZ*udu*^{-/-}
27 gastrulae (Fig.5c-d). This putative increase in adhesion was of particular interest given the tissue boundary
28 defects observed in MZ*udu* mutants. These findings reveal that although *udu* is ubiquitously expressed in
29 zebrafish embryos²⁷, its role in regulating gene expression during gastrulation exhibits some specificity for
30 different classes of encoded proteins, many with known or expected roles in morphogenesis during
31 gastrulation.

32

33 **DamID-seq identifies putative direct targets of Gon4l during gastrulation**

34 To determine genomic loci with which Gon4l protein associates, and thereby distinguish direct from indirect
35 targets of Gon4l regulation, we employed DNA adenine methyltransferase (Dam) identification paired with next
36 generation sequencing (DamID-seq)^{20, 21}. To this end, we generated Gon4l fused to *E. coli* Dam, which
37 methylates adenine residues within genomic regions in its close proximity²⁰. Small equimolar amounts of RNA
38 encoding a Myc-tagged Gon4l-Dam fusion or a Myc-tagged GFP-Dam control were injected into one-celled

1 embryos, and genomic DNA was collected at tailbud stage (see Material and Methods). In support of this
2 Gon4I-Dam fusion being functional, it localized to the nuclei of zebrafish embryos and partially rescued MZudu
3 embryonic phenotypes (Supplemental Fig.5). Methylated, and therefore Gon4I-proximal, genomic regions were
4 then selected using methylation specific restriction enzymes and adaptors, amplified to produce libraries, and
5 subjected to next generation sequencing. Because Dam is highly active, even an untethered version
6 methylates DNA within open chromatin regions²⁰, and so libraries generated from embryos expressing GFP-
7 Dam served as controls.

8
9 Unbiased genome-wide analysis of DamID reads revealed a significant enrichment of Gon4I-Dam over GFP-
10 Dam in promoter regions, and a significant underrepresentation of Gon4I within intergenic regions (Fig.6a).
11 Although no global difference was detected within gene bodies (Fig.6a), examination of individual loci revealed
12 approximately 4,500 genes and over 2,300 promoters in which at least one region was highly Gon4I-enriched
13 ($P_{adj} \leq 0.01$, ≥ 4 -fold enrichment over GFP controls) (Fig.6c-g, Supplemental Table 1). Of these, approximately
14 1,000 genes were co-enriched for Gon4I in both the promoter and gene body (Fig.6c,g). Levels of Gon4I
15 enrichment across a gene and its promoter were significantly correlated (Supplemental Fig.5f), indicating co-
16 enrichment or co-depletion for Gon4I at both gene features of many loci. Within gene bodies, we found robust
17 enrichment specifically within 5' untranslated regions (UTRs) (Fig.6b), consistent with association of Gon4I at
18 or near transcription start sites (Fig.6d). Of the ~2,950 genes differentially expressed in MZudu mutants,
19 approximately 28% (812) were also enriched for Gon4I at the gene body (492), promoter (170), or both (150)
20 (Fig.6c-g), and will hereafter be described as putative direct Gon4I targets. *histh1*, for example, was among the
21 most downregulated genes by RNA-seq in MZudu mutants and was highly enriched for Gon4I at both its
22 promoter and gene body (Fig.6c). By contrast, *tbx6* expression was also reduced in MZudu mutants, but
23 exhibited no enrichment of Gon4I over GFP controls (Fig.6d). Approximately 35% and 53% of genes enriched
24 for Gon4I in only the gene body or only the promoter, respectively, were positively regulated (i.e.
25 downregulated in MZudu mutants), as were 50% of genes co-enriched at both features. Furthermore, among
26 these positively regulated genes, differential expression levels (the degree to which WT expression exceeded
27 MZudu) correlated positively and significantly with Gon4I enrichment levels across both gene bodies and
28 promoters (Fig.6h). A similar correlation was not observed for negatively regulated genes, hence, the highest
29 levels of Gon4I enrichment were associated with positive regulation of gene expression. These results
30 implicate Gon4I as both a positive and negative regulator of gene expression during zebrafish gastrulation.

31 32 **Gon4I regulates notochord boundary straightness by limiting *epcam* and *itga3b* expression**

33 We next examined our list of putative direct Gon4I target genes for those with potential roles in tissue boundary
34 formation and/or cell polarity. The *epcam* gene, which encodes Epithelial cell adhesion molecule (EpCAM),
35 stood out because it was not only enriched for Gon4I by DamID (Fig.7a) and upregulated in MZudu^{-/-} gastrulae
36 (Fig.7c), but was also identified in a *Xenopus* overexpression screen for molecules that disrupt tissue
37 boundaries⁴². Furthermore, EpCAM negatively regulates Cadherin-based adhesion^{43, 44} and non-muscle
38 Myosin activity⁴⁵, making it a compelling candidate. We also chose to examine *itga3b*, which encodes

1 Integrin α 3b, because as a component of a Laminin receptor⁴⁶, it is an obvious candidate for molecules
2 involved in formation of a tissue boundary at which Laminin is highly enriched (Fig.2f). *itga3b* expression was
3 increased in *MZudu*^{-/-} gastrulae by both RNA-seq and qRT-PCR (Fig. 7d), and DamID revealed a region within
4 the *itga3b* promoter at which Gon4l was highly enriched (Fig. 7b-b'). We found that overexpression of either
5 *epcam* or *itga3b* by RNA injection into WT embryos recapitulated some aspects of the *MZudu*^{-/-} gastrulation
6 phenotype, including irregular notochord boundaries (Fig.7e) and reduced ML orientation and elongation of
7 axial mesoderm cells (Fig.7h-m). Conversely, injection of a translation-blocking *itga3b* MO (MO1-*itga3b*⁴⁷) at a
8 dose phenocopying the fin defects of *itga3b* mutants⁴⁷ (Supplemental Fig.6b) significantly improved boundary
9 straightness in *MZudu* mutants compared with control injected siblings (Fig.7f). However, injection of an *epcam*
10 MO (MO2-*epcam*⁴⁸) at a dose phenocopying the loss of otoliths in *epcam* mutants⁴⁸ (Supplemental Fig.6d) did
11 not suppress boundary defects in *MZudu* mutants (Fig.7g), indicating that excess *itga3b* is largely responsible
12 for irregular notochord boundaries in *MZudu*^{-/-} gastrulae. To determine whether excess *itga3b* was sufficient to
13 phenocopy the *kny*^{-/-};*udu*^{-/-} double mutant phenotype identified in our synthetic screen (Fig.1), we
14 overexpressed *itga3b* in *kny*^{-/-} embryos and found that it exacerbated the short axis phenotype of these PCP
15 mutants at 24hpf (Fig.7n-p), an effect not produced by injection of control *GFP* RNA. Together, these results
16 indicate that negative regulation of *itga3b* and *epcam* expression by Gon4l contributes to proper notochord
17 boundary formation in zebrafish gastrulae. Interestingly, ML cell polarity defects were not suppressed in
18 *MZudu*^{-/-} embryos injected with *itga3b* (or *epcam*) MO compared to control injected mutant siblings
19 (Supplemental Fig.6e-p), despite the improvement in boundary straightness. This implies that Gon4l has
20 boundary-dependent and independent roles in cell polarization, and that additional gene expression changes
21 likely contribute to this aspect of the mutant phenotype.

22

23 **Tissue tension is reduced at the notochord boundary of *MZudu* embryos**

24 In WT zebrafish and *Xenopus* embryos, the notochord boundary straightens over time (Fig.2e) and
25 accumulates myosin⁴⁹, implying that it is under tension. Mechanical forces within the embryo (like tension) can
26 instruct specific cell behaviors such as polarization, intercalation, and PCP protein localization⁵⁰⁻⁵⁴, and we
27 hypothesized that increased notochord boundary tension may act as an instructive cell polarity cue that is
28 reduced in *MZudu*^{-/-} gastrulae. We therefore laser-ablated interfaces between axial mesoderm cells in live WT
29 and *MZudu*^{-/-} gastrulae and recorded recoil of adjacent cell vertices as a measure of tissue tension^{55, 56}
30 (Fig.8a-b). Cell interfaces were classified according to convention described in (^{55, 57}) as V junctions (actively
31 shrinking to promote cell intercalation), T junctions (not shrinking) or Edge junctions (falling on and thus
32 comprising the notochord boundary). We found that recoil distances at WT Edge junctions were significantly
33 greater than that of WT V (p=0.0001) or T junctions (p=0.0026), demonstrating that the notochord boundary is
34 under greater tension than the rest of the tissue (Fig.8c-e). Consistent with our hypothesis, recoil distance was
35 significantly smaller in *MZudu*^{-/-} than in WT gastrulae for all classes of junctions (Fig.8c-e), and this decrease
36 was largest and most significant in Edge junctions, especially at 80% epiboly (Fig.8c). These results point to
37 reduced tension as a possible cause of notochord boundary and cell polarity defects in *MZudu* mutants.
38 Because excess *epcam* and *itga3b* were sufficient to disrupt notochord boundaries and ML cell polarity (Fig.7),

1 we next tested whether *epcam* and *itga3b* overexpression could likewise affect boundary tension. Injection of
2 WT embryos with *epcam* RNA was sufficient to disrupt tissue tension at the notochord boundary and
3 throughout the axial mesoderm (Fig.8c-e). Curiously, although we found *itga3b* to be largely causative of
4 boundary irregularity in *MZudu* mutants (Fig.7), its overexpression did not consistently produce a similar
5 reduction in tension (Fig.8c-e). Together these results implicate excess EpCAM and Integrin α 3b as key
6 molecular defects underlying reduced boundary tension and reduced boundary straightness, respectively,
7 observed in *MZudu*^{-/-} gastrulae. We propose a model whereby Gon4l negatively regulates expression of these
8 adhesion molecules to ensure proper formation of the notochord boundary, which together with additional
9 boundary-independent roles of Gon4l cooperates with PCP signaling to promote ML cell polarity underlying
10 C&E gastrulation movements (Fig.8f).

11

12 **DISCUSSION**

13

14 Significant advances have been made in defining signaling pathways instructing gastrulation cell behaviors that
15 shape the vertebrate body plan, but epigenetic control of these morphogenetic processes remains largely
16 unexplored. Here we have described the conserved chromatin factor Gon4l as a regulator of polarized cell
17 behaviors underlying axis extension during zebrafish gastrulation. We identified a large number of genes
18 regulated by Gon4l, including many genes with known or predicted roles in morphogenesis, and linked
19 misregulation of a subset of them to specific morphogenetic cell behaviors. Because Gon4l does not bind DNA
20 directly, we predict that Gon4l-enriched genomic loci are direct targets of chromatin modifying protein
21 complexes with which Gon4l associates⁴¹. Only a fraction of the thousands of Gon4l-enriched genes and
22 promoters exhibited corresponding changes in gene expression during gastrulation, which may reflect the
23 following: 1) Gon4l may not alter expression of all loci with which it associates, 2) loci recently occupied by
24 Gon4l may not yet reflect changes in transcript levels, and 3) because DamID provides a “history” of Gon4l
25 association, our experiment revealed both past and currently occupied loci. We also identified loci at which
26 Gon4l was depleted compared to GFP-Dam controls (Supplemental Fig.5), and speculate they represent open
27 chromatin regions with which Gon4l does not associate. Surprisingly, although a larger number of putative
28 Gon4l direct target genes were negatively regulated (i.e. upregulated in *MZudu* mutants), the highest levels of
29 Gon4l enrichment correlated with positive regulation of gene expression (Fig.6). This demonstrates that Gon4l
30 does not act strictly as a negative regulator of gene expression, a role assigned to it based on *in vitro* evidence
31 and thought to be mediated by its interactions with Histone deacetylases⁴¹. Our data instead indicate that
32 Gon4l acts as both a positive and negative regulator of gene expression during zebrafish gastrulation, implying
33 context-specific interactions with multiple epigenetic regulatory complexes.

34

35 Phenotypes caused by complete *udu* deficiency are conspicuously pleiotropic, but our studies point to
36 remarkably specific roles for *udu* in gastrulation morphogenesis. Loss of *udu* function reduced tissue extension
37 without impairment of other gastrulation movements, including mesendoderm internalization, epiboly,
38 prechordal plate migration, and convergence (see below). Moreover, the dorsal gastrula organizer and all three
39 germ layers were specified (Fig.1, Supplemental Fig.2), indicating that *MZudu* mutants do not suffer from a

1 general delay or arrest of development; rather Gon4l regulates a specific subset of gastrulation cell behaviors
2 including ML cell polarity and intercalation in the axial mesoderm. Importantly, the role of Gon4l in these
3 processes is independent of PCP signaling as supported by several lines of evidence, including distinct
4 morphogenetic defects in PCP versus *MZudu* mutants. Whereas PCP mutants exhibit reduced convergence
5 movements^{9, 10, 14}, evidenced by a larger number of cell rows in *kny*^{-/-} axial mesoderm (Fig.4), the axial
6 mesoderm of *MZudu* mutants contained a normal number of cell rows (Fig.3), indicating no obvious
7 convergence defect. We also observed an intact notochord boundary in PCP mutants, as well as boundary-
8 associated cell polarity (Fig.4). Despite these apparently parallel functions, expression of some PCP genes,
9 including *wnt11*, *wnt11r*, *prickle1a*, *prickle1b*, *celsr1b*, *celsr2*, and *fzd2* are regulated by Gon4l as revealed by
10 RNA-seq and DamID-seq experiments (Supplemental Table 1). However, functional redundancies within the
11 PCP network^{58, 59} likely allow for intact PCP signaling in *MZudu* mutants despite misregulation of some PCP
12 genes. Notably, gastrula morphology and cell polarity defects in *MZudu* mutants were also distinct from
13 ventralized and dorsalized patterning mutants that exhibit abnormal C&E^{13, 15, 60}.
14
15 Modulation of adhesion at tissue boundaries has been implicated as a driving force of cell intercalation^{49, 61},
16 and indeed genes annotated as encoding adhesion molecules tended to be expressed at higher levels in
17 *MZudu*^{-/-} gastrulae (Fig.5). Two of these, *itga3b* and *epcam*, exhibited increased expression in *MZudu*^{-/-}
18 gastrulae by RNA-seq and qRT-PCR, were identified as putative direct Gon4l targets by DamID (Fig.7), and
19 were each sufficient to disrupt notochord boundary straightness and ML cell orientation and elongation when
20 overexpressed in WT embryos (Fig.7). Furthermore, decreasing levels of Integrin α 3b (but not EpCAM)
21 improved boundary straightness in *MZudu* mutants (Fig.7), while overexpression of *epcam* (but not *itga3b*) was
22 sufficient to reduce tissue tension at the notochord boundary of WT embryos (Fig.8). These observations
23 suggest that excesses of each of these molecules contribute to overlapping and distinct cellular defects. In
24 *Xenopus* gastrulae, EpCAM negatively regulates non-muscle Myosin contractility⁴⁵, and experimental
25 perturbation of Myosin activity disrupts notochord boundary formation⁴⁹. Myosin accumulation also increases
26 tension at tissue boundaries^{62, 63}, which biases cell intercalations^{52, 54}. This likely explains why excess *epcam*
27 was sufficient to disrupt boundary tension and straightness in WT embryos, although modulation of Integrin-
28 mediated cell-matrix adhesion at the notochord boundary was also necessary to cause *MZudu* boundary
29 phenotypes. Notably, restoration of boundary straightness in *MZudu*^{-/-} gastrulae was not sufficient to improve
30 ML cell orientation. This could reflect at least two possibilities: 1) the suppression of boundary phenotypes by
31 the *itga3b* MO was partial and therefore not sufficient to fully restore the boundary-associated cell polarity
32 signal; or 2) the boundary-associated polarity signal was restored, but *MZudu*^{-/-} axial mesoderm cells were
33 unable to respond to it. Either of these scenarios assumes that additional gene expression changes contribute
34 to reduced ML cell polarity in *MZudu* mutants, and indeed, several other candidate genes with plausible or
35 known roles in C&E and/or boundary formation were also misregulated in *MZudu*^{-/-} gastrulae. For example,
36 *epha7*, which encodes an Eph receptor, a class of signaling molecules with well-described roles in tissue
37 boundary formation⁶⁴, was expressed at reduced levels in *MZudu* mutants. Expression of *daam1*, which
38 encodes a critical link between PCP signaling and Actomyosin contractility⁶⁵, and genes encoding several

1 Myosin light and heavy chain isoforms were also reduced in *MZudu* mutants (Supplemental Table 1). In
2 addition to excess *itga3b* and *epcam* expression, misregulation of these molecules could also contribute to
3 impaired ML cell polarity in *MZudu*^{-/-} gastrulae.

4

5 In this study of zebrafish Gon4l, we have begun to dissect the logic of epigenetic regulation of gastrulation
6 morphogenesis, and revealed a key role for this chromatin factor in limiting expression of specific genes during
7 gastrulation. This role has important developmental implications, as C&E gastrulation movements are sensitive
8 to both gain and loss of gene function^{9, 12}. We propose that by negatively regulating Integrin α 3b and EpCAM
9 levels, Gon4l promotes development of the anteroposteriorly aligned notochord boundary and influences ML
10 polarity and intercalation of axial mesoderm cells. In cooperation with PCP signaling, this cue coordinates ML
11 cell polarity with embryonic patterning to drive anteroposterior embryonic axis extension (Fig.8f).

12

13

14 **ACKNOWLEDGEMENTS**

15 We thank Dr. Bas Van Steensel for EcoDam plasmids, Drs. Christine and Bernard Thisse for WISH probes,
16 Dr. Bo Zhang and Dr. Scott Higdon for bioinformatics help, Dr. Matthew Hass for DamID advice, Bisiayo
17 Fashemi for assistance with image analysis, and the Washington University Genome Technology Access
18 Center for library preparation and sequencing services.

19

20 **AUTHOR CONTRIBUTIONS**

21 M.W., A.S., and L.S.K. designed the study. M.W., A.S., and T.B. performed experiments. C.Y. participated in
22 the initial forward genetic screen. P.G. performed bioinformatic analysis. M.W. and L.S.K. wrote the
23 manuscript. National Institutes of Health grants R01GM55101 and R35GM118179 to L.S.K. and
24 F32GM113396 to M.W., and a W.M. Keck Foundation Fellowship to M.W. in part supported this study.

25

26 **METHODS**

27

28 **Zebrafish strains and embryo staging**

29 Adult zebrafish were raised and maintained according to established methods⁶⁶ in compliance with standards
30 established by the Washington University Animal Care and Use Committee. Embryos were obtained from
31 natural mating and staged according to morphology as described⁶⁷. All studies on WT were carried out in AB*
32 or AB*/Tübingen backgrounds. Additional lines used include *kny*^{m818}, *kny*^{fr6 10}, *udu*^{sq1 27}, and *udu*^{vu66} (this work).
33 Embryos of these strains generated from heterozygous intercrosses were genotyped by PCR after completion
34 of each experiment. Germline replaced fish were generated by the method described in (³³). Briefly, donor
35 embryos from *udu*^{vu66/+} intercrosses or females with *udu*^{vu66/vu66} germline were injected with synthetic RNA
36 encoding *GFP-nos1-3'UTR*⁶⁸, and WT host embryos were injected with a morpholino oligonucleotide against
37 *dead end1* (*MO1-dnd1*)³³ to eliminate host germ cells. Cells were transplanted from the embryonic margin of
38 donor blastulae to the embryonic margin of hosts at sphere stage, and both hosts and donors were cultured in

1 agarose-coated plates. Host embryos were screened for GFP+ germ cells at 36-48 hpf, and the genotype of
2 corresponding donors was determined by phenotype. All putative *udu*^{*vu66/vu66*} germline hosts were raised to
3 adulthood and confirmed by crossing to *udu*^{*vu66/+*} animals prior to use in experiments.

4

5 **Synthetic mutant screening**

6 WT male fish were mutagenized by the chemical mutagen *N*-ethyl-*N*-nitrosourea (ENU) as described²², then
7 outcrossed to WT females to produce F1 families. F2 families were obtained by crossing F1 fish with fish
8 homozygous for the hypomorphic *knypek* allele *kny*^{*m818*} (rescued by injection with synthetic *kny* WT RNA). F3
9 embryos obtained from F2 cross were screened by morphology at 12 and 24hpf to identify recessive
10 enhancers of the *kny*^{*m818/m818*} short axis mutant phenotype¹⁰.

11

12 **Positional cloning**

13 We employed the positional cloning approach using a panel of CA simple-sequence length polymorphism
14 markers representing 25 linkage groups²⁵ to map the *vu66* mutation to chromosome 16 between the markers
15 of z17403 and z15431. Given phenotypic similarities between *vu66/vu66* and the *udu*^{*sq1/sq1*} mutant
16 phenotype²⁷, we sequenced *udu* cDNA from 24hpf *vu66/vu66* mutant embryos revealing a T2261A
17 transversion that is predicted to create Y753STOP nonsense mutation. We designed a dCAPS⁶⁹ marker for the
18 *vu66* mutation and confirmed that no recombination occurred in 810 *vu66/vu66* homozygous embryos.

19

20 **Microinjection**

21 One-celled embryos were aligned within agarose troughs generated using custom-made plastic molds and
22 injected with 1-3 pL volumes using pulled glass needles. Synthetic mRNAs for injection were made by *in vitro*
23 transcription from linearized plasmid DNA templates using Invitrogen mMessage mMachine kits. Doses of RNA
24 per embryo were as follows: 100pg *membrane Cherry*, 50pg *membrane GFP*, 25pg *udu-gfp*, 200pg *epcam*,
25 200pg *itga3b*, 1pg *gfp-dam-myc*, 3pg *udu-dam-myc* for DamID experiments, 20pg *gfp-dam-myc* for MZudu
26 rescue, and 150pg *GFP-nos1-3'UTR* for germline transplantation. To assess Pk-GFP localization, embryos
27 were injected at one cell stage with *membrane Cherry* RNA, then injected with 15pg *Drosophila prickle-GFP*
28 and 20pg *H2B-RFP* RNAs into a single blastomere at 16-cell stage as described^{32, 39}. Injections of MOs were
29 carried out as for synthetic RNA. Doses of MOs per embryo were as follows: 3ng MO1-*dnd1*³³, 4ng MO1-
30 *tri/vangl2*³⁸, 1ng MO2-*epcam*⁴⁸, 2ng MO1-*itga3b*⁴⁷.

31

32 **Whole-mount in situ hybridization**

33 Antisense riboprobes were transcribed using NEB T7 or T3 RNA polymerase and labeled with digoxigenin
34 (DIG) (Roche). Whole-mount *in situ* hybridization (WISH) was performed as described⁷⁰.

35

36 **Immunofluorescent staining**

37 Embryos were fixed in 4% paraformaldehyde (PFA) in phosphate buffered saline (PBS), rinsed in PBS + 0.1%
38 tween (PBT), digested briefly in 10 mg/mL proteinase K, re-fixed in 4% PFA, rinsed in PBT, and blocked in 2

1 mg/mL bovine serum albumin + 2% goat serum in PBT. Embryos were then incubated overnight in rabbit anti-
2 Laminin (Sigma L9393) at 1:200, mouse anti-Myc (Cell Signaling 2276) at 1:1000, or rabbit anti-phospho
3 Histone H3 (Upstate 06-570) at 1:500 in blocking solution, rinsed in PBT, and incubated overnight in Alexa
4 Fluor 488 anti-Rabbit IgG, 568 anti-Rabbit, or 568 anti-Mouse (Invitrogen) at 1:1000 in PBT. Embryos were co-
5 stained with 4',6-Diamidino-2-Phenylindole, Dihydrochloride (DAPI) and rinsed in PBT prior to mounting in
6 agarose for confocal imaging.

7

8 **TUNEL staining**

9 Terminal deoxynucleotidyl transferase (TdT) dUTP Nick-End Labeling (TUNEL) staining to detect apoptosis
10 was carried out according to the instructions for the ApopTag Peroxidase *in situ* apoptosis detection kit
11 (Millipore) with modifications. Briefly, embryos were fixed in 4% PFA, digested with 10 µg/mL proteinase K, re-
12 fixed with 4%PFA, and post-fixed in chilled ethanol:acetic acid 2:1, rinsing in PBT after each step. Embryos
13 were incubated overnight with TdT and rinsed in stop/wash buffer, then blocked, incubated with anti-DIG
14 antibody, and stained in Roche BM Purple staining solution.

15

16 **Microscopy**

17 Live embryos expressing fluorescent proteins or fixed embryos subjected to immunofluorescent staining were
18 mounted in 0.75% low-melt agarose in glass bottomed 35-mm petri dishes for imaging using a modified
19 Olympus IX81 inverted spinning disc confocal microscope equipped with Voltran and Cobolt steady-state
20 lasers and a Hamamatsu ImagEM EM CCD digital camera. For live time-lapse series, 60 µm z-stacks with a
21 2µm step were collected every three or ten minutes (depending on the experiment) for three hours using a 40x
22 dry objective lens. Embryo temperature was maintained at 28.5°C during imaging using a Live Cell Instrument
23 stage heater. When necessary, embryos were extracted from agarose after imaging for genotyping. For
24 immunostained embryos, 200 µm z-stacks with a 1 or 2µm step were collected using a 10x or 20x dry objective
25 lens, depending on the experiment. Bright field and transmitted light images of live embryos and *in situ*
26 hybridizations were collected using a Nikon AZ100 macroscope.

27

28 **Laser ablation tension measurements**

29 Embryos were injected at one cell stage with *mCherry* mRNA and mounted at 80% epiboly for imaging (as
30 described above) on a Zeiss 880 Airyscan 2-photon inverted confocal microscope. An infrared laser tuned to
31 710 nm was used to ablate fluorescently labeled cell interfaces, immediately followed by a quick time-lapse
32 series of ten images with no interval to record recoil after each ablation event. Images were collected using a
33 40x water immersion objective lens, and a Zeiss stage heater was used to maintain embryo temperature at
34 28.5°C. Approximately 8-12 cell interfaces were ablated per embryo, and no fewer than 32 embryos per
35 genotype from no fewer than four independent experiments were analyzed due to high variability of recoil
36 distances. Image series were analyzed using ImageJ to determine the inter-vertex distance of each cell
37 interface prior to and immediately after ablation and used to calculate recoil distance.

38

1 **Image analysis**

2 ImageJ was used to visualize and manipulate all microscopy data sets. For immunostained embryos, multiple
3 z-planes were projected together to visualize the entire region of interest. For live embryo analysis, a single z-
4 plane through the length of the axial mesoderm was chosen for each time point. When possible, embryo
5 images were analyzed prior to genotyping. Cell polarity was analyzed in no fewer than six live embryos per
6 genotype due to high variability of these measurements. To measure cell orientation and elongation, the
7 anteroposterior axis in all embryo images was aligned prior to manual outlining of cells. A fit ellipse was used
8 to measure orientation of each cell's major axis and its aspect ratio. The TissueAnalyzer ImageJ package⁷¹
9 was used to automatically segment time-lapse series of axial mesoderm and detect T1 transitions. Boundary
10 straightness was measured by manually tracing the notochord boundary to determine total length, then dividing
11 it by the length of a straight line connecting the ends of the boundary (net length). To assess Pk-GFP
12 localization, isolated cell expressing Pk-GFP were scored according to whether GFP was present in puncta
13 localized to the anterior side of the cell, in puncta localized elsewhere, or in the cytoplasm. Pk-GFP localization
14 was analyzed in multiple cells from each of no fewer than three embryos per genotype.

15

16 **Quantitative RT-PCR**

17 Total RNA was isolated from tailbud stage WT and *MZudu*^{-/-} embryos homogenized in Trizol (Life
18 Technologies), 1 μ g of which was used to synthesize cDNA using the iScript kit (BioRad) following
19 manufacturer's protocol. SYBR green (BioRad) qRT-PCR reactions were run in a CFX Connect Real-Time
20 PCR detection system (BioRad). Primers used are as follows:

21 *epcam*: F- TGAGGACGGGGATTGAGAAC
22 R- GAGCCTGCCATCCTTGTCAT

23 *itga3b*: F- CCGGTGTTGGGAGAAGAGAC
24 R- CTTGAAGAAACCACACGAAGGG

25 *EF1a*: F- CTGGAGGCCAGCTCAAACAT
26 R- ATCAAGAAGAGTAGTACCGCTAGCATTAC

27 *Rpl13a*: F- TCTGGAGGACTGTAAGAGGTATGC
28 R- AGACGCACAATCTTGAGAGCAG

29

30 **RNA sequencing and analysis**

31 RNA for sequencing was isolated from 50 WT or *MZudu* embryos per sample at tailbud stage according to
32 instructions for the Dynabeads mRNA direct kit (Ambion). Embryos from two clutches per genotype (e.g. WT A
33 & B) were collected, then divided in two (e.g. WT A1, A2, B1, and B2) to yield four independently prepared
34 libraries representing two biological and two technical replicates per genotype. Libraries for were prepared
35 according to instructions for the Epicentre ScriptSeq v2 RNA-seq Library preparation kit (Illumina). Briefly, RNA
36 was enzymatically fragmented prior to cDNA synthesis. cDNA was then 3' tagged, purified using Agencourt
37 AMPure beads, and PCR amplified, at which time sequencing indexes were added. Indexed libraries were then
38 purified and submitted to the Washington University Genome Technology Access Center (GTAC) for
39 sequencing using an Illumina HiSeq 2500 to obtain single-ended 50bp reads. Raw reads were mapped to the

1 zebrafish GRCz10 reference genome using STAR (2.4.2a)⁷² with default parameters. FeatureCounts (v1.4.6)
2 from the Subread package⁷³ was used to quantify the number of uniquely mapped reads (phred score \geq 10) to
3 gene features based on the Ensembl annotations (v83). Significantly differentially expressed genes were
4 determined by using DESeq2 in the negative binomial distribution model⁷⁴ with a cutoff of adjusted p-
5 value \leq 0.01 and fold-change \geq 2.0. Heatmaps were built using the heatmaps2 package in R, and other plots
6 were built using the ggplot2 package in R.

7

8 **DamID-seq**

9 *E. coli* DNA adenine methyltransferase (EcoDam) was cloned from the pIND(V5)EcoDam plasmid²¹ (a kind gift
10 from Dr. Bas Van Steensel, Netherlands Cancer Institute) by Gibson assembly into a 3' Gateway entry vector
11 containing 6 Myc tags and an SV40 polyA signal (p3E-MTpA). The resulting p3E-EcoDam-MTpA vector was
12 then Gateway cloned into PCS2+ downstream of *udu* cDNA or *eGFP* to produce C terminal fusions. The
13 resulting *udu-dam-myc* and *gfp-dam-myc* plasmids were linearized by KpnI digestion and transcribed using the
14 mMessage mMachine SP6 *in vitro* transcription kit (Ambion). WT AB* embryos were injected at one cell stage
15 with 3pg mRNA encoding Gon4I-Dam-Myc or with 1pg encoding GFP-Dam-Myc as controls. The remainder of
16 the protocol was carried out largely as in (²¹) with modifications. Genomic (g)DNA was collected at tailbud
17 stage using a Qiagen DNeasy kit with the addition of RNase A. gDNA was digested with DpnI overnight,
18 followed by ligation of DamID adaptors. Un-methylated regions were destroyed by digestion with DpnII, and
19 then methylated regions were amplified using primers complementary to DamID adaptors. Two identical 50 μ l
20 PCR reactions were performed and pooled for each sample to reduce amplification bias, and three biological
21 replicates were collected per condition. Amplicons were purified using a Qiagen PCR purification kit, then
22 digested with DpnII to remove DamID adaptors. Finally, samples were purified using Agencourt AMPure XP
23 beads and submitted for library preparation and sequencing at the Washington University GTAC using an
24 Illumina HiSeq 2500 to obtain single-ended 50bp reads.

25

26 **Analysis of DamID-seq data**

27 Raw reads were aligned to zebrafish genome GRCz10 by using bwa mem (v0.7.12) with default parameters⁷⁵,
28 then sorted and converted into bam format by using SAMtools (v1.2)⁷⁶. The zebrafish genome was divided into
29 continuous 1000bp bins, and FeatureCounts (v1.4.6) from the Subread package⁷³ was used to quantify the
30 number of uniquely mapped reads (phred score \geq 10) in each bin. Significantly differentially Gon4I-associated
31 bins were determined by using DESeq2 in the negative binomial distribution model⁷⁴ with stringent cutoff:
32 adjusted p-value \leq 0.01 and fold-change \geq 4.0. Promoter regions were defined as 2kb upstream of the
33 transcription start site of a gene based on Ensembl annotations (v83). Wiggle and bigwig files were created
34 from bam files using IGVtools (v2.3.60) (<https://software.broadinstitute.org/software/igv/igvtools>) and
35 wigToBigWig (v4)⁷⁷, respectively. BigWig tracks were visualized using IGV (v2.3.52)⁷⁸.

36

37

38

1 **Data Availability**

2 Processed RNA-seq and DamID-seq data are available in Supplemental Table1 of this publication, and raw
3 data have been deposited in the Gene Expression Omnibus (GEO) with the following accession numbers:

4 RNA-seq: GSE96575

5 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=ipwzmaoqjpajbqd&acc=GSE96575>

6 DamID-seq: GSE96576

7 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=apkloooyvxsxij&acc=GSE96576>

8

9 **Statistical analysis**

10 Graphpad Prism 6 and 7 software was used to perform statistical analyses and generate graphs of data
11 collected from embryo images. The statistical tests used varied as appropriate for each experiment and are
12 described in the text and figure legends. All tests used were two-tailed. Differential expression and differential
13 enrichment analysis of RNA and DamID sequencing data were completed as described above. Panther⁷⁹ was
14 used to classify differentially expressed genes and to produce pie charts; DAVID Bioinformatic Resources⁸⁰
15 was used for functional annotation analysis.

16

17 **Subcloning**

18 The full-length *udu* open reading frame was subcloned from cDNA using following primers:

19 *udu* cacc F1: CACCATGGGATGGAAACGCAAGTCTTC

20 *udu* TGA R: TCAGTCCTGCTCTTCATCAGTGGC

21 *udu* R: GTCCTGCTCTTCATCAGTGGCCGAC

22 *udu* cDNAs with or without a stop codon were cloned into the pENTR/D-TOPO (Thermo Fisher) vector, which
23 were then Gateway cloned into PCS2+ upstream of polyA signal or *eGFP* to produce a C terminal fusion.

24 The full-length *epcam* open reading frame was subcloned from WT cDNA using the following primers:

25 *epcam* F: GGATCCCATCGATTCGATGAAGGTTTTAGTTGCCTTG

26 *epcam* R: ACTCGAGAGGCCTTGTTAAGAAATTGTCTCCATCTC

27 The 5' portion of the *itga3b* open reading frame was cloned from WT cDNA, and the 3' portion was cloned from
28 a partial cDNA clone (GE/Dharmacon) using the following primers:

29 5' *itga3b* F: TTTGCAGGCGCGCCGGATCCCATCGATTCGATGGCCGAAAGTCTCTG

30 5' *itga3b* R: ATTTGAGTGAGTATGGAATGGAGATGTTGAGCG

31 3' *itga3b* F: TCAACATCTCCATTCCATACTCACTCAAATACTCAGG

32 3' *itga3b* R: GTTCTAGAGGTTTAAACTCGAGAGGCCTTGTCAGAACTCCTCCGTCAG

33 The resulting amplicons were Gibson cloned⁸¹ into PJS2 (a derivative of PCS2+) linearized with EcoRI.

FIGURES

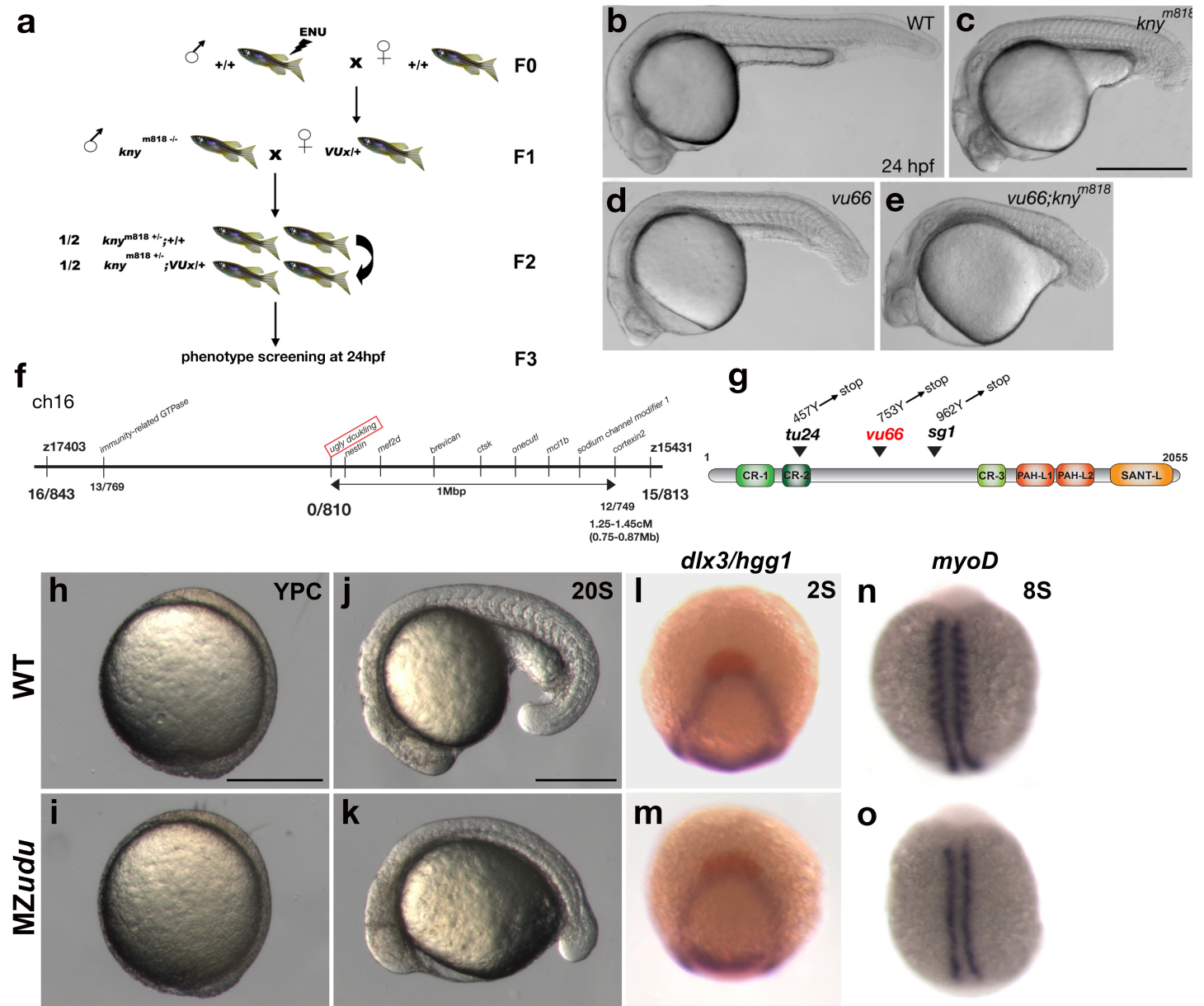


Figure 1: A forward genetic screen identifies *ugly duckling (udu)/gon4l* as a novel regulator of axis extension in zebrafish embryos.

a) Schematic of a synthetic screen to identify enhancers of the short axis phenotype in $kny^{m818/m818}$ zebrafish mutants. **b-e)** Phenotypes at 24 hpf: wild-type (WT) (b), $kny^{m818/m818}$ (c), $vu66/vu66$ (d), $kny^{m818/m818}; vu66/vu66$ compound mutants (e). **f)** Diagram of mapping $vu66$ mutation to Chromosome 16. Bold numbers below specify the number of recombination events between $vu66$ and the indicated loci. **g)** Diagram of the Gon4l protein encoded by the *ugly duckling (udu)* locus. Arrowheads indicate residues mutated in $vu66$ and other described *udu* alleles. **h-k)** Live WT (h,j) and maternal zygotic (MZ)*udu* (i,k) embryos at yolk plug closure (YPC) (h,i) and 20 somite stage (j,k). **l-m)** Whole mount *in situ* hybridization (WISH) for *dlx3* (purple) and *hgg1* (red) in WT (l) and MZ*udu* (m) embryos at 2 somite stage. **n-o)** WISH for *myoD* in WT (n) and MZ*udu* (o) embryos at 8 somite stage. Anterior is to the left in b-e, j-k; anterior is up in h-i, l-o. Scale bar is 500 μ m in b-e and 300 μ m in h-o.

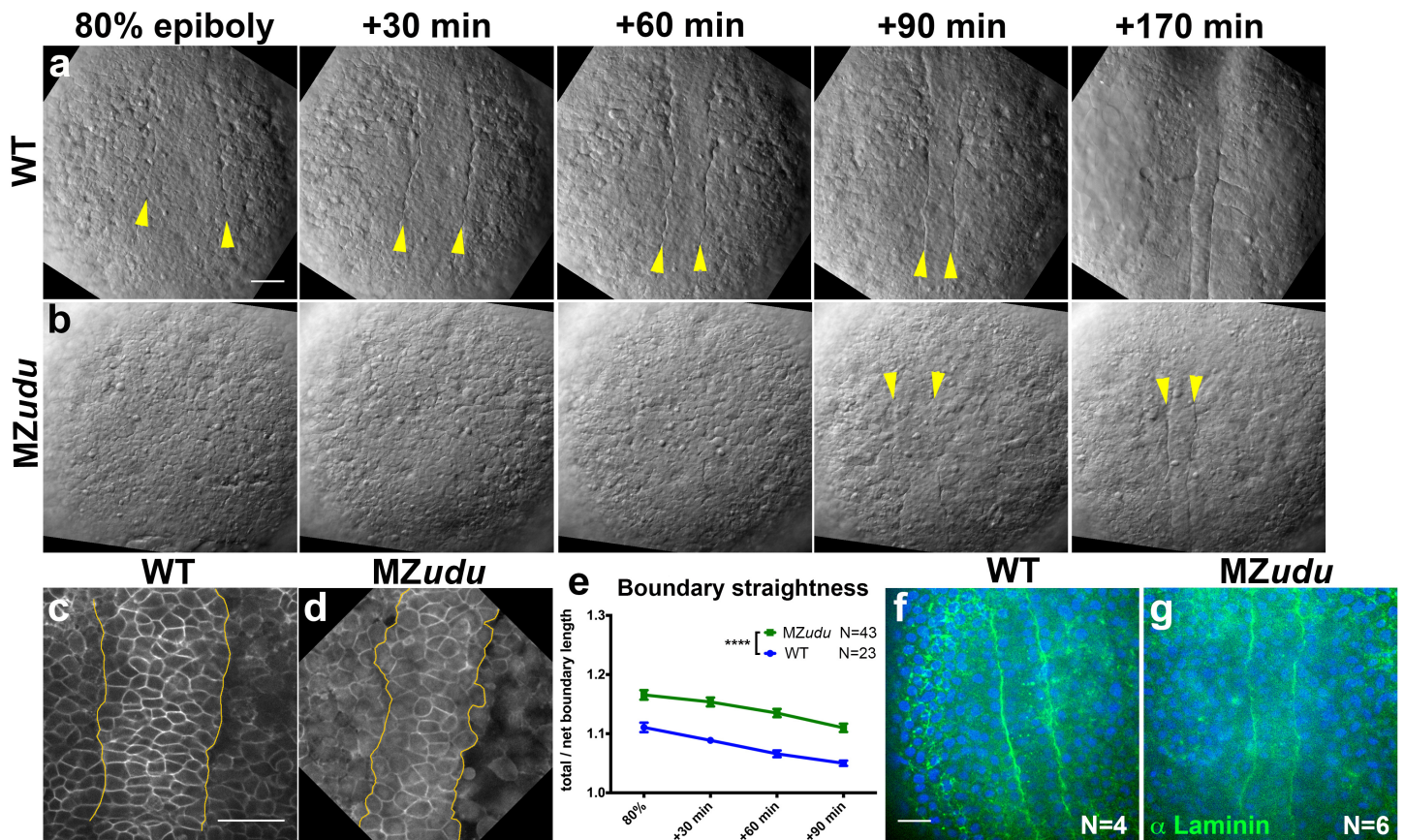


Figure 2: MZudu mutant gastrulae exhibit irregular notochord boundaries.

a-b Still images from live Nomarski time-lapse series of the dorsal mesoderm in WT (a) and MZudu embryos (b) at the time points indicated. Arrowheads indicate notochord boundaries. **c-d** Live confocal microscope images of representative WT (c) and MZudu (d) embryos expressing membrane Cherry. Yellow lines demarcate notochord boundaries. **e** Quantification of notochord boundary straightness in live WT and MZudu-/- gastrulae throughout gastrulation. Symbols are means with SEM (2-way ANOVA, ****p<0.0001). **f-g** Confocal microscope image of immunofluorescent staining for pan-Laminin in WT (f) and MZudu (g) embryos at 2 somite stage. Scale bars are 50 μ m. Anterior is up in all images.

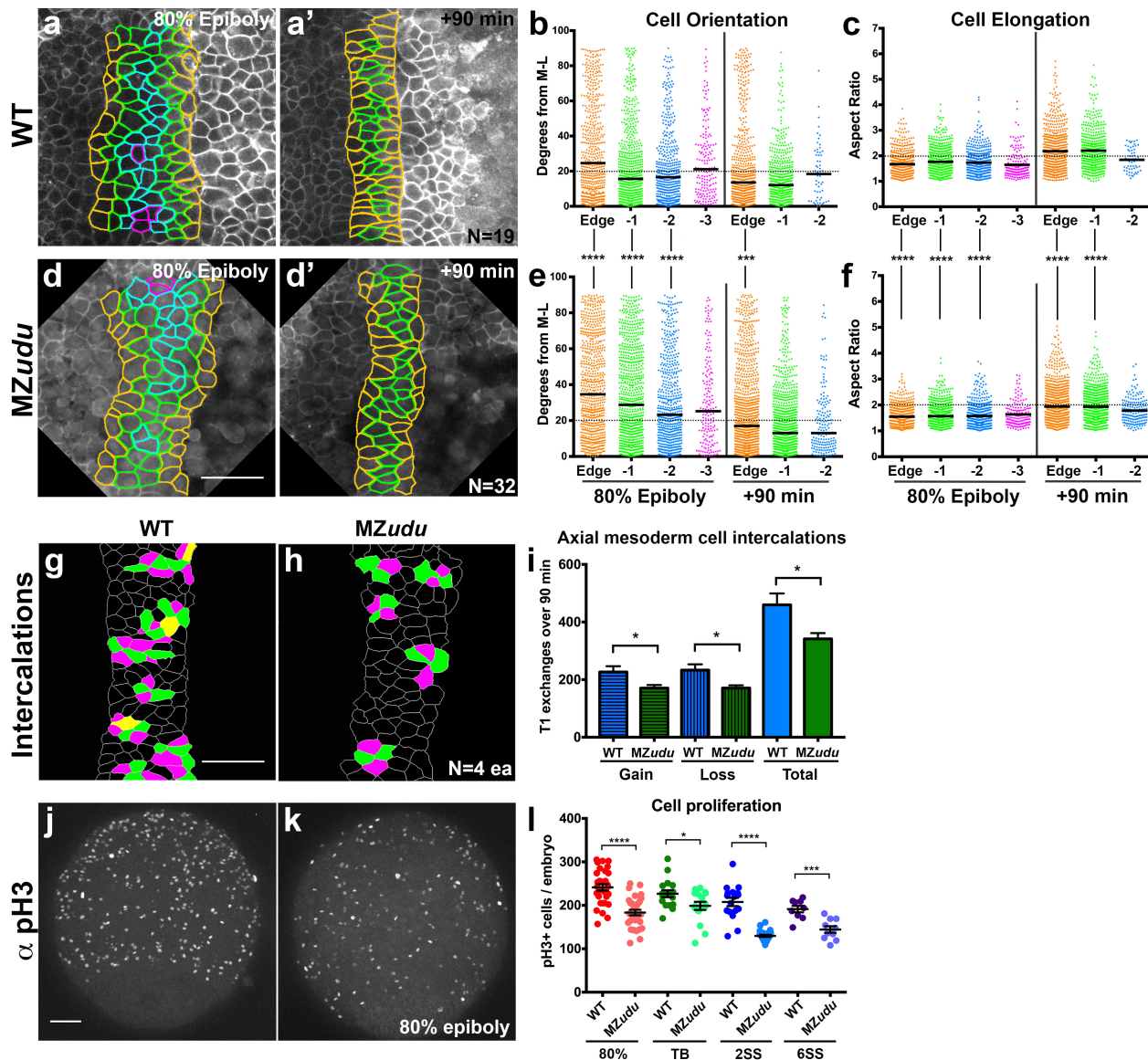


Figure 3: Mediolateral cell polarity and cell intercalations are reduced in the axial mesoderm of *MZudu*^{-/-} embryos.

a-a', d-d') Still images from live time-lapse confocal movies of the axial mesoderm in WT (**a**) and *MZudu*^{-/-} (**d**) gastrulae at the time points indicated. Cell outlines are colored according to a cell's position with respect to the notochord boundary. **b, e**) Quantification of axial mesoderm cell orientation at 80% epiboly (left side) and +90 minute (right side) time points. Each dot represents the orientation of the major axis of a single cell with respect to the embryonic ML axis and is colored according to that cell's position with respect to the notochord boundary (colors match those of the images to the left). Horizontal black lines indicate median values. Asterisks indicate significant differences between WT and *MZudu* (Kolmogorov-Smirnov test, *** $p < 0.001$, **** $p < 0.0001$). **c, f**) Quantification of axial mesoderm cell elongation at 80% epiboly (left side) and +90 minute (right side) time points. Each dot represents the aspect ratio of a single cell and is color-coded as in (**b**). Horizontal black lines indicate mean values. Asterisks indicate significant differences between WT and *MZudu* (Mann-Whitney test, **** $p < 0.0001$). **g-h**) Cell intercalations detected in the axial mesoderm of WT (**g**) and *MZudu*^{-/-} gastrulae (**h**). Cells gaining contacts with neighbors are green, cells losing contacts are magenta, and cells that both gain and lose contacts are yellow. **i**) Quantification of cell intercalation events (T1 exchanges) in WT (blue bars) and *MZudu*^{-/-} gastrulae (green bars) over 90 minutes. Bars are means with SEM (T test, * $p < 0.05$). **j-k**) 200 μ m confocal Z projections of immunofluorescent staining for phosphorylated Histone H3 (pH3) in WT (**j**) and *MZudu*^{-/-} gastrulae (**k**) at 80% epiboly. **l**) Quantification of pH3⁺ cells/embryo at the stages indicated. Each dot represents a single embryo, dark lines are means with SEM (T test, * $p < 0.05$, *** $p < 0.001$, **** $p < 0.0001$). Scale bars are 50 μ m. N indicates the number of embryos analyzed. Anterior/animal pole is up in all images.

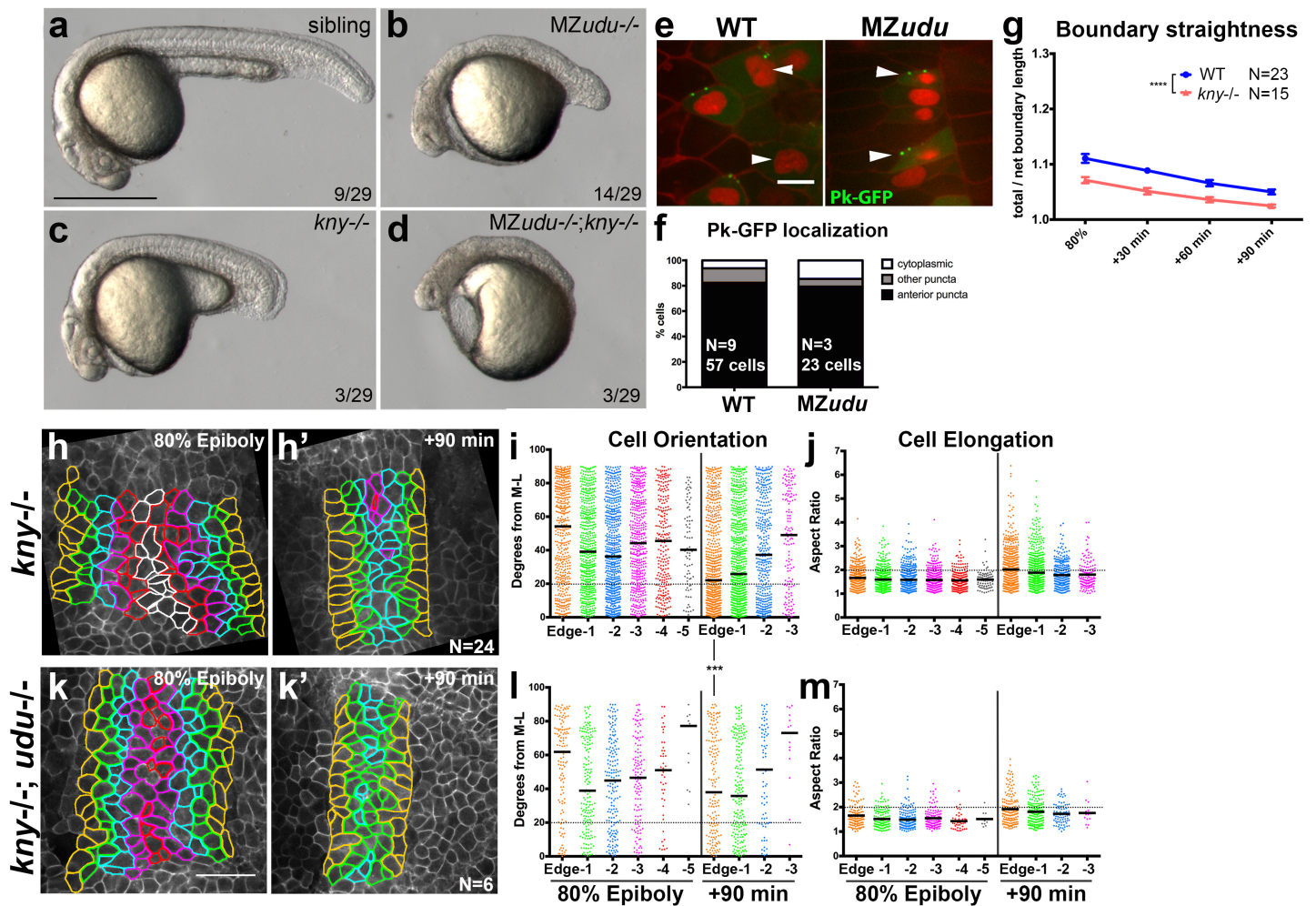


Figure 4: Gon4l regulates axis extension independent of PCP signaling.

a-d) Live embryos at 24hpf resulting from a cross between a germline-replaced *udu*^{-/-}; *kny*^{fr6/+} female and an *udu*^{+/-}; *kny*^{fr6/+} male. Genotypes are indicated in the upper right corner, the fraction of embryos in the clutch with the pictured phenotype is provided at the bottom. **e)** Mosaically expressed Prickle(Pk)-GFP in WT and *MZudu*^{-/-} gastrulae. Arrowheads indicate anteriorly localized Pk-GFP puncta. Membrane Cherry demarcates cell membranes, nuclear-RFP marks cells injected with *pk-gfp* RNA. **f)** Quantification of Pk-GFP localization shown in e. Distribution is not significantly different (chi-square, $p=0.07$). **g)** Quantification of notochord boundary straightness in WT and *kny*^{-/-} gastrulae. Symbols are means with SEM (2-way ANOVA, **** $p<0.0001$). **h,h',k,k')** Still images from live time-lapse confocal movies of the axial mesoderm in *kny*^{-/-} (h) and *kny*^{-/-}; *udu*^{-/-} (k) gastrulae at the time points indicated. Cell outlines are colored as in Figure 3. **i,i)** Quantification of axial mesoderm cell orientation as in Figure 3 (Kolmogorov-Smirnov test, *** $p=0.0003$). **j,m)** Quantification of axial mesoderm cell elongation as in Figure 3. Scale bar is 500 μ m in a-d, 10 μ m in e, 50 μ m in h-k.

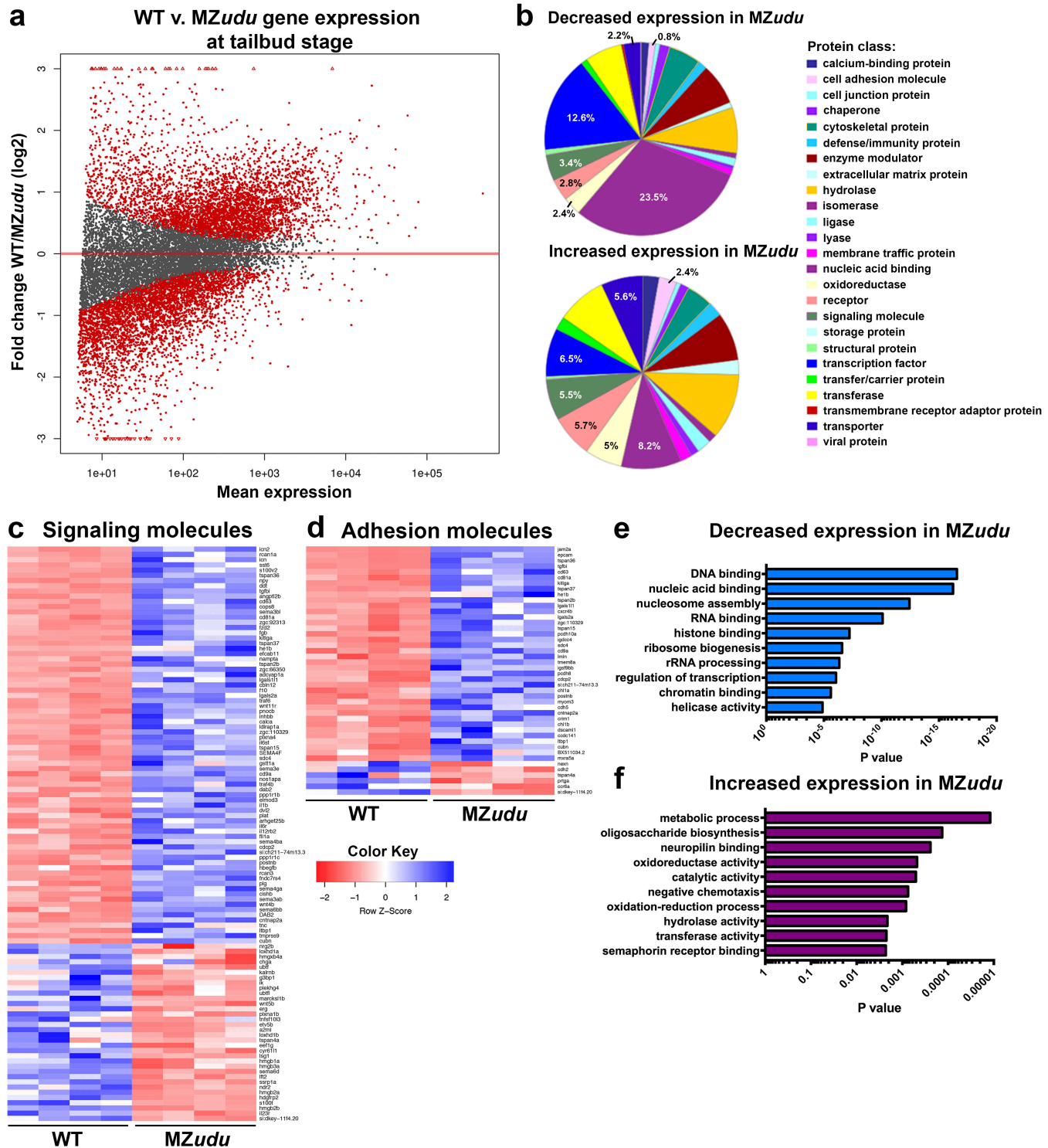


Figure 5: Loss of Gon4l results in large-scale gene expression changes.

a) Plot of gene expression changes in MZudu mutants compared to WT at tailbud stage as assessed by RNA-sequencing. Red dots indicate genes expressed at significantly different levels than WT ($p \leq 0.01$, at least 2-fold change in transcript level). **b**) Protein classes encoded by differentially expressed genes with decreased (top graph) or increased (bottom graph) expression in MZudu mutants. Percentages indicate the number of genes within a given class/total number of genes with increased or decreased expression, respectively. **c-d**) Heat maps of differentially expressed genes annotated as encoding signaling (c) or adhesion molecules (d). The four columns represent two biological and two technical replicates for each WT and MZudu. **e-f**) P-values of the top ten most enriched gene ontology (GO) terms among genes with decreased (e) or increased (f) expression in MZudu-/- embryos compared to WT.

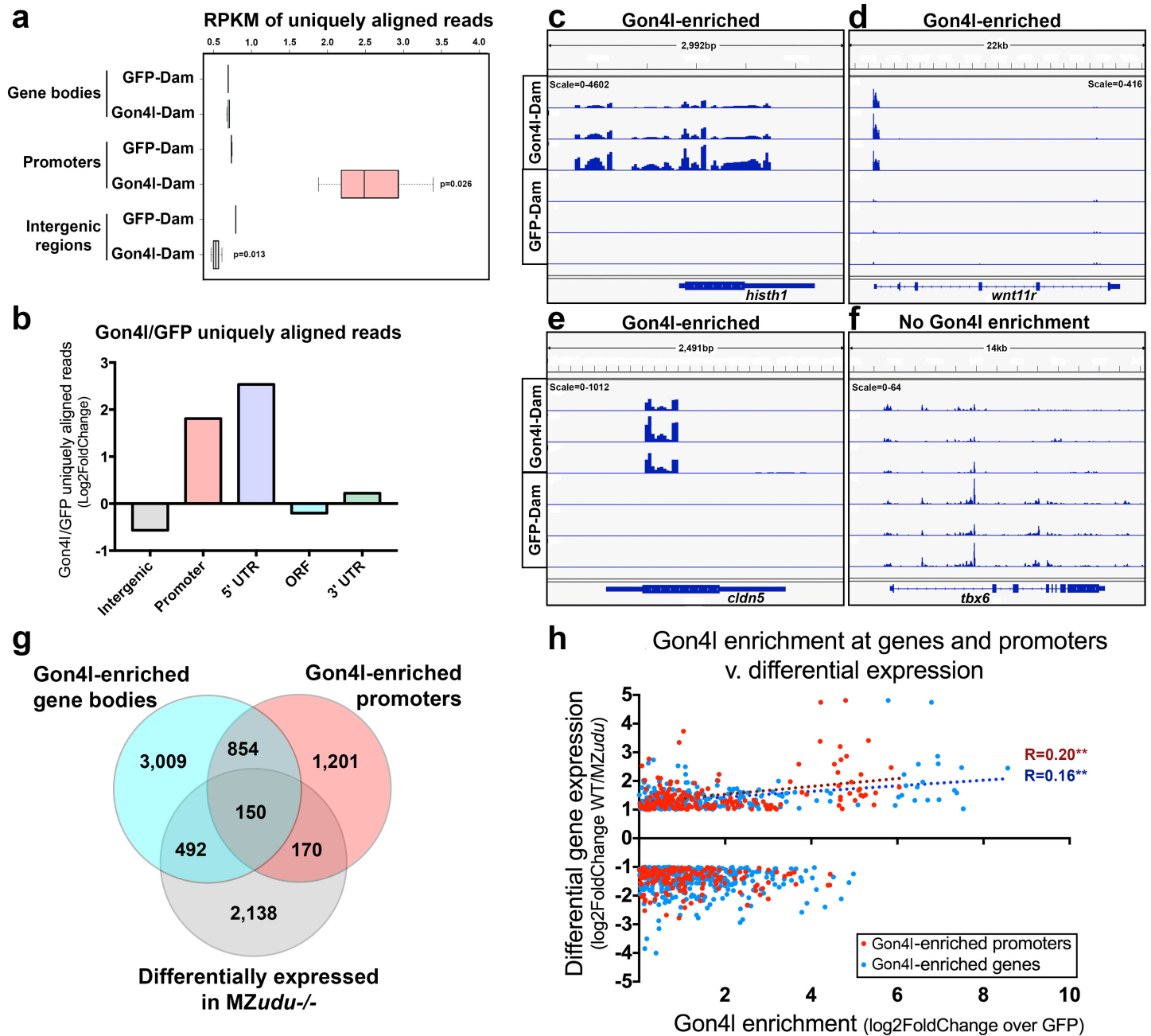


Figure 6: DamID-seq identifies putative direct targets of Gon4l during gastrulation.

a) Box plot of normalized uniquely aligned DamID reads at each of three categories of genomic regions. P values indicate significant differences between Gon4l-Dam samples and GFP-controls (T-test). **b**) Fold change (log₂) Gon4l-Dam over GFP-Dam reads (RPKM) at each of five gene features. **c-f**) Genome browser tracks of Gon4l-Dam and GFP-Dam association at the *histh1* (c), *wnt11r* (d), *cldn5* (e), and *tbx6* (f) loci. Scale is number of reads. Each track represents one biological replicate at tailbud stage. **g**) Venn diagram of genes with regions of significant Gon4l-enrichment within gene bodies (blue) or promoters (pink), and genes differentially expressed in MZudu^{-/-} gastrulae (gray). **h**) Correlation of Gon4l enrichment levels across gene bodies (blue dots) and promoters (red dots) with relative transcript levels of genes differentially expressed in MZudu mutants. A positive correlation was detected between increased expression in WT relative to MZudu mutants and Gon4l enrichment in both gene bodies (Spearman correlation $p=0.0095$) and promoters ($p=0.0098$). Dotted lines are linear regressions of these correlations.

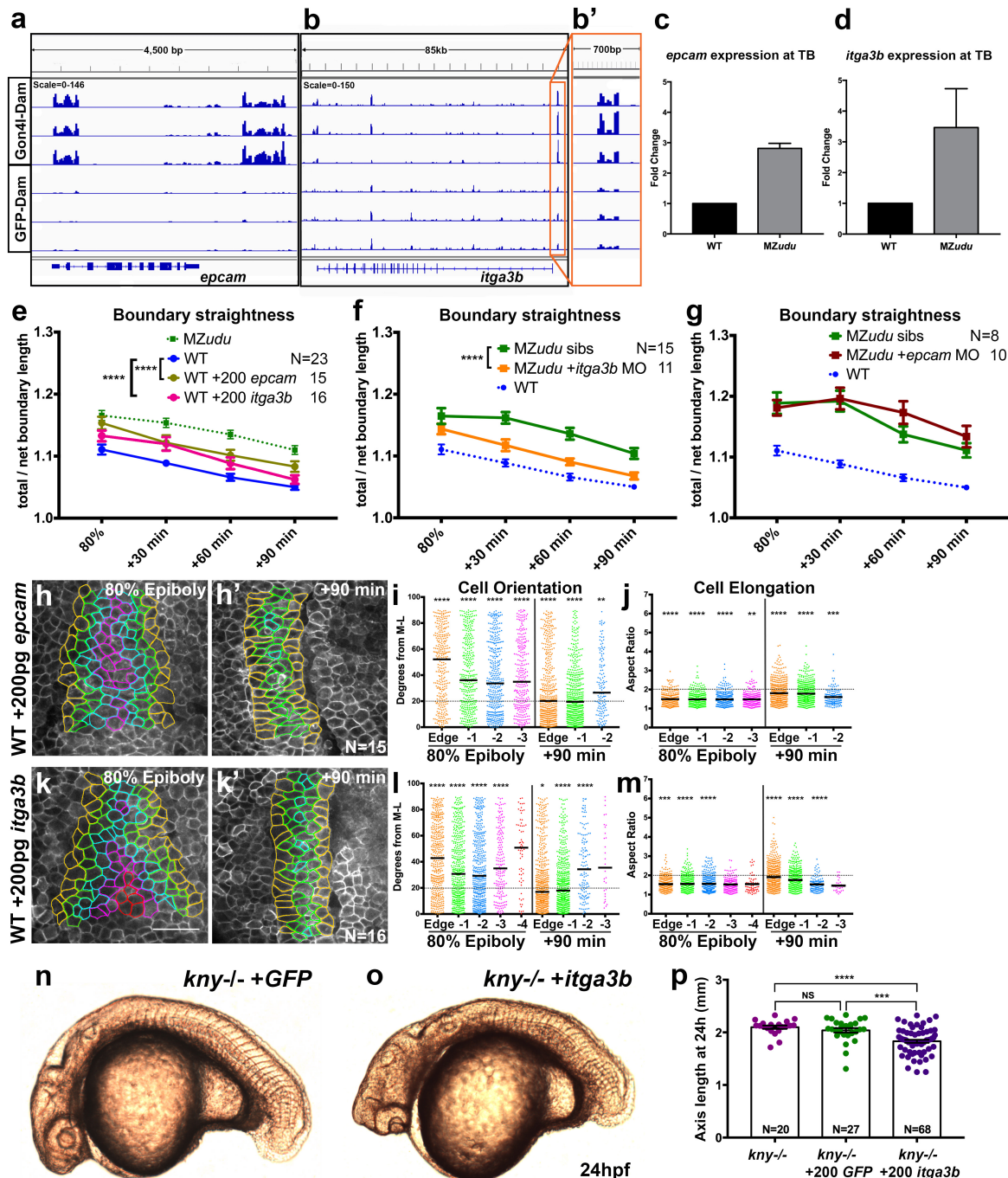


Figure 7: Gon4l regulates notochord boundary straightness by limiting *itga3b* and *epcam* expression. **a-b'**) Genome browser tracks of Gon4l-Dam and GFP-Dam association at the *epcam* (a) and *itga3b* (b) locus. An expanded view of the *itga3b* promoter is shown in b'. **c-d**) Quantitative RT-PCR for *epcam* (c) and *itga3b* (d) in WT and MZudu^{-/-} embryos at tailbud stage. Bars are means with SEM. **e**) Quantification of notochord boundary straightness in WT *epcam* and *itga3b* overexpressing embryos (2-way ANOVA, ****p<0.0001). **f-g**) Quantification of notochord boundary straightness in MZudu^{-/-} *epcam* (f) and *itga3b* (g) morphants and sibling controls (2-way ANOVA, ****p<0.0001). N indicates the number of embryos analyzed, symbols are means with SEM. **h,h',k,k'**) Still images from live time-lapse confocal movies of the axial mesoderm in *epcam* (h) and *itga3b* (k) overexpressing WT gastrulae at the time points indicated. Cell outlines are colored as in Figure 3. **i,l**) Quantification of axial mesoderm cell orientation as in Figure 3. **j,m**) Quantification of axial mesoderm cell elongation as in Figure 3. Asterisks indicate significant differences compared to WT controls. Scale bar is 50µm. **n-o**) Live images of *kny*^{-/-} embryos at 24hpf injected with 200pg GFP (n) or 200pg *itga3b* (o) RNA. **p**) Quantification of axis length of injected *kny*^{-/-} embryos at 24hpf. Each dot represents one embryo, bars are means with SEM (T-tests, ***p<0.001, ****p<0.0001).

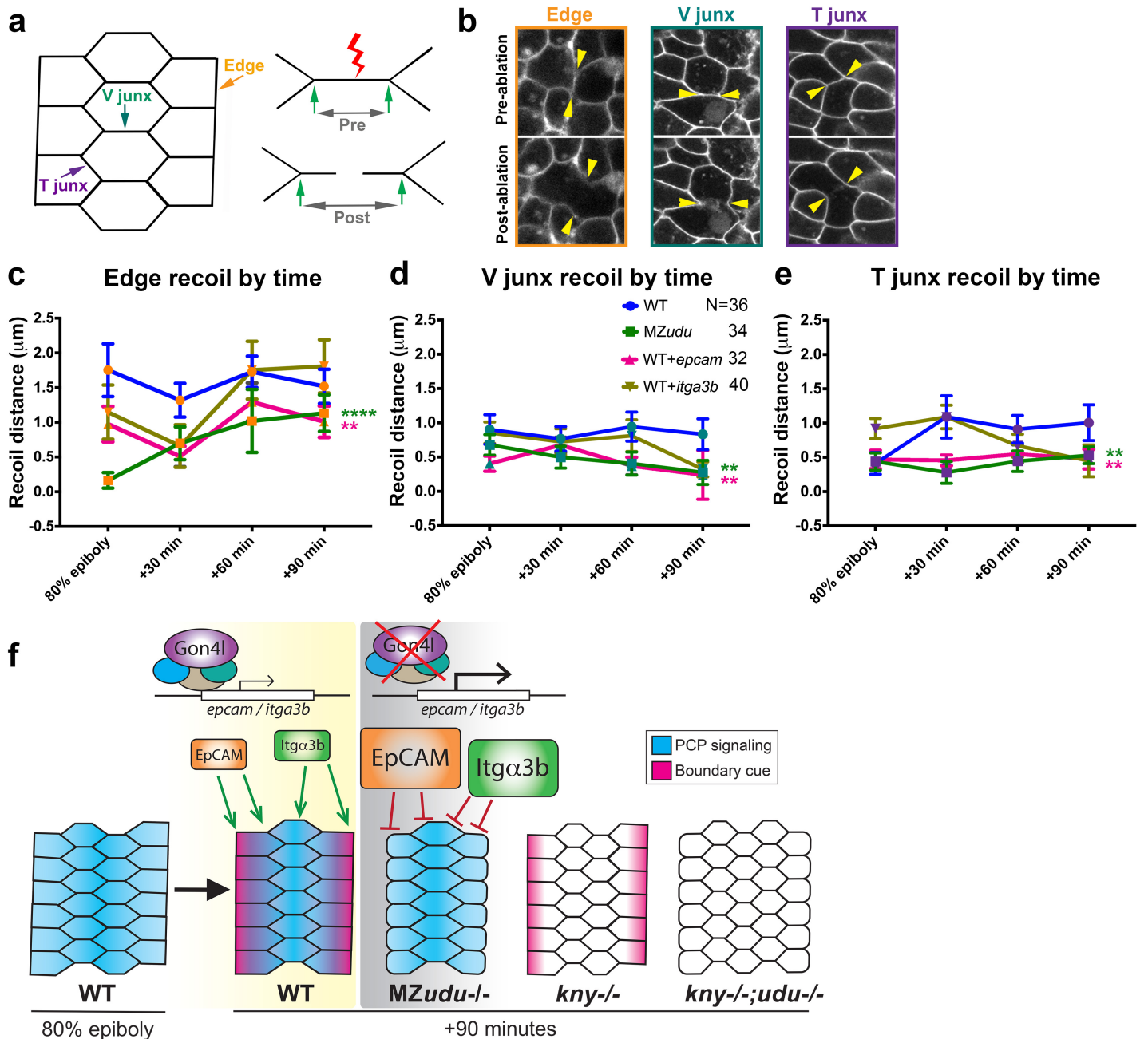


Figure 8: Tension is reduced at the notochord boundaries of MZudu mutant and *epcam* overexpressing WT gastrulae.

a) Diagram of laser ablation experiments to measure tension at axial mesoderm cell interfaces. b) Still images from confocal time-lapse movies of each of the three types of cell interfaces (Edge, V junctions, and T junctions) before and after laser ablation. Arrowheads indicate cell vertices adjacent to the ablated interface. c-e) Quantification of cell vertex recoil distance immediately after laser ablation of Edge (c), V junction (d), and T junction (e) interfaces at the time points indicated in WT, MZudu *-/-*, WT *epcam* overexpressing, and WT *itga3b* overexpressing gastrulae. Symbols means with SEM. Asterisks are colored according to key and indicate significant differences compared to WT controls (2-way ANOVA, **** $p < 0.0001$, ** $p < 0.01$). N indicates the number of embryos analyzed. f) Graphical model of the roles of Gon4l and PCP signaling in regulating ML cell polarity of axial mesoderm cells.

REFERENCES

1. Shih, J. & Keller, R. Patterns of cell motility in the organizer and dorsal mesoderm of *Xenopus laevis*. *Development* **116**, 915-930 (1992).
2. Yen, W.W. *et al.* PTK7 is essential for polarized cell motility and convergent extension during mouse gastrulation. *Development* (2009).
3. Voiculescu, O., Bertocchini, F., Wolpert, L., Keller, R.E. & Stern, C.D. The amniote primitive streak is defined by epithelial cell intercalation before gastrulation. *Nature* **449**, 1049-1052 (2007).
4. Warga, R.M. & Kimmel, C.B. Cell movements during epiboly and gastrulation in zebrafish. *Development* **108**, 569-580 (1990).
5. Shih, J. & Keller, R. Cell motility driving mediolateral intercalation in explants of *Xenopus laevis*. *Development* **116**, 901-914 (1992).
6. Sepich, D.S., Calmelet, C., Kiskowski, M. & Solnica-Krezel, L. Initiation of convergence and extension movements of lateral mesoderm during zebrafish gastrulation. *Dev Dyn* **234**, 279-292 (2005).
7. Keller, R. *et al.* Mechanisms of convergence and extension by cell intercalation. *Philos Trans R Soc Lond B Biol Sci* **355**, 897-922 (2000).
8. Wallingford, J.B. *et al.* Dishevelled controls cell polarity during *Xenopus* gastrulation. *Nature* **405**, 81-85 (2000).
9. Jessen, J.R. *et al.* Zebrafish trilobite identifies new roles for Strabismus in gastrulation and neuronal movements. *Nat Cell Biol* **4**, 610-615 (2002).
10. Topczewski, J. *et al.* The zebrafish glypican knypek controls cell polarity during gastrulation movements of convergent extension. *Dev Cell* **1**, 251-264 (2001).
11. Williams, M., Yen, W., Lu, X. & Sutherland, A. Distinct apical and basolateral mechanisms drive planar cell polarity-dependent convergent extension of the mouse neural plate. *Dev Cell* **29**, 34-46 (2014).
12. Lin, F. *et al.* Essential roles of G α 12/13 signaling in distinct cell behaviors driving zebrafish convergence and extension gastrulation movements. *J Cell Biol* **169**, 777-787 (2005).
13. Solnica-Krezel, L. *et al.* Mutations affecting cell fates and cellular rearrangements during gastrulation in zebrafish. *Development* **123**, 67-80 (1996).
14. Sepich, D.S. *et al.* Role of the zebrafish trilobite locus in gastrulation movements of convergence and extension. *Genesis* **27**, 159-173 (2000).
15. Myers, D.C., Sepich, D.S. & Solnica-Krezel, L. Bmp activity gradient regulates convergent extension during zebrafish gastrulation. *Dev Biol* **243**, 81-98 (2002).
16. Chen, T. & Dent, S.Y. Chromatin modifiers and remodellers: regulators of cellular differentiation. *Nat Rev Genet* **15**, 93-106 (2014).
17. Moreno-Ayala, R., Schnabel, D., Salas-Vidal, E. & Lomelí, H. PIAS-like protein Zimp7 is required for the restriction of the zebrafish organizer and mesoderm development. *Dev Biol* **403**, 89-100 (2015).
18. Ma, Y. *et al.* The Chromatin Remodeling Protein Bptf Promotes Posterior Neuroectodermal Fate by Enhancing Smad2-Activated wnt8a Expression. *J Neurosci* **35**, 8493-8506 (2015).
19. Nambiar, R.M. & Henion, P.D. Sequential antagonism of early and late Wnt-signaling by zebrafish colgate promotes dorsal and anterior fates. *Dev Biol* **267**, 165-180 (2004).
20. van Steensel, B. & Henikoff, S. Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. *Nat Biotechnol* **18**, 424-428 (2000).
21. Vogel, M.J., Peric-Hupkes, D. & van Steensel, B. Detection of in vivo protein-DNA interactions using DamID in mammalian cells. *Nat Protoc* **2**, 1467-1478 (2007).
22. Solnica-Krezel, L., Schier, A.F. & Driever, W. Efficient recovery of ENU-induced mutations from the zebrafish germline. *Genetics* **136**, 1401-1420 (1994).
23. Mullins, M.C., Hammerschmidt, M., Haffter, P. & Nüsslein-Volhard, C. Large-scale mutagenesis in the zebrafish: in search of genes controlling development in a vertebrate. *Current biology : CB* **4**, 189-202 (1994).
24. Marlow, F. *et al.* Functional interactions of genes mediating convergent extension, knypek and trilobite, during the partitioning of the eye primordium in zebrafish. *Dev Biol* **203**, 382-399 (1998).
25. Knapik, E.W. *et al.* A microsatellite genetic linkage map for zebrafish (*Danio rerio*). *Nat Genet* **18**, 338-343 (1998).
26. Knapik, E.W. *et al.* A reference cross DNA panel for zebrafish (*Danio rerio*) anchored with simple sequence length polymorphisms. *Development* **123**, 451-460 (1996).

27. Liu, Y. *et al.* The zebrafish *udu* gene encodes a novel nuclear factor and is essential for primitive erythroid cell development. *Blood* **110**, 99-106 (2007).
28. Hammerschmidt, M. *et al.* Mutations affecting morphogenesis during gastrulation and tail formation in the zebrafish, *Danio rerio*. *Development* **123**, 143-151 (1996).
29. Lee, M.T. *et al.* Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition. *Nature* **503**, 360-364 (2013).
30. Harvey, S.A. *et al.* Identification of the zebrafish maternal and paternal transcriptomes. *Development* **140**, 2703-2710 (2013).
31. Gritsman, K. *et al.* The EGF-CFC protein one-eyed pinhead is essential for nodal signaling. *Cell* **97**, 121-132 (1999).
32. Ciruna, B., Jenny, A., Lee, D., Mlodzik, M. & Schier, A.F. Planar cell polarity signalling couples cell division and morphogenesis during neurulation. *Nature* **439**, 220-224 (2006).
33. Ciruna, B. *et al.* Production of maternal-zygotic mutant zebrafish by germ-line replacement. *Proc Natl Acad Sci U S A* **99**, 14919-14924 (2002).
34. Blagden, C.S., Currie, P.D., Ingham, P.W. & Hughes, S.M. Notochord induction of zebrafish slow muscle mediated by Sonic hedgehog. *Genes Dev* **11**, 2163-2175 (1997).
35. Lim, C.H., Chong, S.W. & Jiang, Y.J. Udu deficiency activates DNA damage checkpoint. *Mol Biol Cell* **20**, 4183-4193 (2009).
36. Payne, E.M. *et al.* Ddx18 is essential for cell-cycle progression in zebrafish hematopoietic cells and is mutated in human AML. *Blood* **118**, 903-915 (2011).
37. Liu, Y., Sepich, D.S. & Solnica-Krezel, L. Stat3/Cdc25a-dependent cell proliferation promotes embryonic axis extension during zebrafish gastrulation. *PLoS Genet* **13**, e1006564 (2017).
38. Park, M. & Moon, R.T. The planar cell-polarity gene *stbm* regulates cell behaviour and cell fate in vertebrate embryos. *Nat Cell Biol* **4**, 20-25 (2002).
39. Yin, C., Kiskowski, M., Pouille, P.A., Farge, E. & Solnica-Krezel, L. Cooperation of polarized cell intercalations drives convergence and extension of presomitic mesoderm during zebrafish gastrulation. *J Cell Biol* **180**, 221-232 (2008).
40. Veeman, M.T. *et al.* Chongmague reveals an essential role for laminin-mediated boundary formation in chordate convergence and extension movements. *Development* **135**, 33-41 (2008).
41. Lu, P. *et al.* The developmental regulator protein Gon4l associates with protein YY1, co-repressor Sin3a, and histone deacetylase 1 and mediates transcriptional repression. *J Biol Chem* **286**, 18311-18319 (2011).
42. Maghzal, N., Vogt, E., Reintsch, W., Fraser, J.S. & Fagotto, F. The tumor-associated EpCAM regulates morphogenetic movements through intracellular signaling. *J Cell Biol* **191**, 645-659 (2010).
43. Litvinov, S.V. *et al.* Epithelial cell adhesion molecule (Ep-CAM) modulates cell-cell interactions mediated by classic cadherins. *J Cell Biol* **139**, 1337-1348 (1997).
44. Winter, M.J. *et al.* Expression of Ep-CAM shifts the state of cadherin-mediated adhesions from strong to weak. *Exp Cell Res* **285**, 50-58 (2003).
45. Maghzal, N., Kayali, H.A., Rohani, N., Kajava, A.V. & Fagotto, F. EpCAM controls actomyosin contractility and cell adhesion by direct inhibition of PKC. *Dev Cell* **27**, 263-277 (2013).
46. Belkin, A.M. & Stepp, M.A. Integrins as receptors for laminins. *Microsc Res Tech* **51**, 280-301 (2000).
47. Carney, T.J. *et al.* Genetic analysis of fin development in zebrafish identifies furin and hemicentin1 as potential novel fraser syndrome disease genes. *PLoS Genet* **6**, e1000907 (2010).
48. Slanchev, K. *et al.* The epithelial cell adhesion molecule EpCAM is required for epithelial morphogenesis and integrity during zebrafish epiboly and skin development. *PLoS Genet* **5**, e1000563 (2009).
49. Fagotto, F., Rohani, N., Touret, A.S. & Li, R. A molecular base for cell sorting at embryonic boundaries: contact inhibition of cadherin adhesion by ephrin/ Eph-dependent contractility. *Dev Cell* **27**, 72-87 (2013).
50. Shindo, A. *et al.* Tissue-tissue interaction-triggered calcium elevation is required for cell polarization during *Xenopus* gastrulation. *PLoS One* **5**, e8897 (2010).
51. Chien, Y.H., Keller, R., Kintner, C. & Shook, D.R. Mechanical strain determines the axis of planar polarity in ciliated epithelia. *Curr Biol* **25**, 2774-2784 (2015).
52. Landsberg, K.P. *et al.* Increased cell bond tension governs cell sorting at the *Drosophila* anteroposterior compartment boundary. *Curr Biol* **19**, 1950-1955 (2009).

53. Rudolf, K. *et al.* A local difference in Hedgehog signal transduction increases mechanical cell bond tension and biases cell intercalations along the *Drosophila* anteroposterior compartment boundary. *Development* **142**, 3845-3858 (2015).
54. Umetsu, D. *et al.* Local increases in mechanical tension shape compartment boundaries by biasing cell intercalations. *Curr Biol* **24**, 1798-1805 (2014).
55. Rauzi, M., Verant, P., Lecuit, T. & Lenne, P.F. Nature and anisotropy of cortical forces orienting *Drosophila* tissue morphogenesis. *Nat Cell Biol* **10**, 1401-1410 (2008).
56. Kiehart, D.P., Galbraith, C.G., Edwards, K.A., Rickoll, W.L. & Montague, R.A. Multiple forces contribute to cell sheet morphogenesis for dorsal closure in *Drosophila*. *J Cell Biol* **149**, 471-490 (2000).
57. Shindo, A. & Wallingford, J.B. PCP and septins compartmentalize cortical actomyosin to direct collective cell movement. *Science* **343**, 649-652 (2014).
58. Kilian, B. *et al.* The role of Ppt/Wnt5 in regulating cell shape and movement during zebrafish gastrulation. *Mech Dev* **120**, 467-476 (2003).
59. Nikaïdo, M., Law, E.W. & Kelsh, R.N. A systematic survey of expression and function of zebrafish frizzled genes. *PLoS One* **8**, e54833 (2013).
60. Mullins, M.C. *et al.* Genes establishing dorsoventral pattern formation in the zebrafish embryo: the ventral specifying genes. *Development* **123**, 81-93 (1996).
61. Green, J.B., Dominguez, I. & Davidson, L.A. Self-organization of vertebrate mesoderm based on simple boundary conditions. *Dev Dyn* **231**, 576-581 (2004).
62. Monier, B., Pélissier-Monier, A. & Sanson, B. Establishment and maintenance of compartmental boundaries: role of contractile actomyosin barriers. *Cell Mol Life Sci* **68**, 1897-1910 (2011).
63. Monier, B., Pélissier-Monier, A., Brand, A.H. & Sanson, B. An actomyosin-based barrier inhibits cell mixing at compartmental boundaries in *Drosophila* embryos. *Nat Cell Biol* **12**, 60-65; sup pp 61-69 (2010).
64. Cayuso, J., Xu, Q. & Wilkinson, D.G. Mechanisms of boundary formation by Eph receptor and ephrin signaling. *Dev Biol* **401**, 122-131 (2015).
65. Habas, R., Kato, Y. & He, X. Wnt/Frizzled activation of Rho regulates vertebrate gastrulation and requires a novel Formin homology protein Daam1. *Cell* **107**, 843-854 (2001).
66. Westerfield, M. *The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio)*. (Univ. of Oregon Press, 1993).
67. Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B. & Schilling, T.F. Stages of embryonic development of the zebrafish. *Dev Dyn* **203**, 253-310 (1995).
68. Köprunner, M., Thisse, C., Thisse, B. & Raz, E. A zebrafish nanos-related gene is essential for the development of primordial germ cells. *Genes Dev* **15**, 2877-2885 (2001).
69. Neff, M.M., Neff, J.D., Chory, J. & Pepper, A.E. dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in *Arabidopsis thaliana* genetics. *Plant J* **14**, 387-392 (1998).
70. Thisse, C. & Thisse, B. High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat Protoc* **3**, 59-69 (2008).
71. Aigouy, B., Umetsu, D. & Eaton, S. Segmentation and Quantitative Analysis of Epithelial Tissues. *Methods Mol Biol* **1478**, 227-239 (2016).
72. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
73. Liao, Y., Smyth, G.K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res* **41**, e108 (2013).
74. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
75. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
76. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
77. Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S. & Karolchik, D. BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* **26**, 2204-2207 (2010).
78. Thorvaldsdóttir, H., Robinson, J.T. & Mesirov, J.P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* **14**, 178-192 (2013).
79. Mi, H., Muruganujan, A., Casagrande, J.T. & Thomas, P.D. Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc* **8**, 1551-1566 (2013).

80. Huang, d.W., Sherman, B.T. & Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57 (2009).
81. Gibson, D.G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343-345 (2009).