Unexpected mutual regulation underlies paralogue functional diversification and promotes epithelial tissue maturation in *Tribolium*

Daniela Gurska1, Iris M. Vargas Jentzsch1 & Kristen A. Panfilio1,2✉

Insect *Hox3/zen* genes represent an evolutionary hotspot for changes in function and copy number. Single orthologues are required either for early specification or late morphogenesis of the extraembryonic tissues, which protect the embryo. The tandemly duplicated *zen* paralogues of the beetle *Tribolium castaneum* present a unique opportunity to investigate both functions in a single species. We dissect the paralogues’ expression dynamics (transcript and protein) and transcriptional targets (RNA-seq after RNAi) throughout embryogenesis. We identify an unexpected role of Tc-Zen2 in repression of Tc-zen1, generating a negative feedback loop that promotes developmental progression. Tc-Zen2 regulation is dynamic, including within co-expressed multigene loci. We also show that extraembryonic development is the major event within the transcriptional landscape of late embryogenesis and provide a global molecular characterization of the extraembryonic serosal tissue. Altogether, we propose that paralogue mutual regulation arose through multiple instances of *zen* sub-functionalization, leading to their complementary extant roles.
C hange over macroevolutionary time scales can produce new gene functions, with the Hox3/zen genes of insects representing a striking example. Across the bilaterian animals, Hox genes are conserved in genomic organization, expression, and function, with roles in tissue specification along the anterior–posterior body axis of the developing embryo. Instead, the Hox3 genes in winged insects, known as zen, are prone to genomic microinversions and the dipteran-specific Tc-zen1 has undergone a tandem duplication. In particular, the serosa is capable of innate immune responses and it secretes a thick chitin-based cuticle that mechanically reinforces the eggshell and provides desiccation resistance. Although the Hox3/zen genes of insects are conserved in genomic organization, expression, and function, with roles in tissue specification along the anterior–posterior body axis of the developing embryo, the Hox3 locus is prone to lineage-specific duplication, to date a single EEM function—specification or morphogenesis (late tissue remodeling for withdrawal)—is known per species in bugs and flies. This is even true in the derived case of the fruit fly Drosophila melanogaster, which has three functionally distinct paralogues: zen itself is involved in EEM specification, the duplicate z2 is not required for embryogenesis, and the dipteran-specific bicoid encodes a maternal determinant with no extrablastic function. Furthermore, secondary tissue simplification of the EEMs in Drosophila obviated the requirement for the late withdrawal function. Thus, the original role of zen within the extrablastic domain has been obscured by ongoing evolutionary changes in both zen and the EEMs.

There is a notable exception to the pattern of a single EEM role of zen per species. In the red flour beetle, Tribolium castaneum, zen has undergone a tandem duplication. Tc-zen1 was first cloned from cDNA, while Tc-zen2 was later identified by sequencing the Hox cluster directly. The paralogues are striking for their compact, shared gene structure and for their proximity: within the 58-kb region between Hox2/mxp and Hox4/Dfd, the paralogues occupy a <3-kb interval, with only 216 bp between the 3’ UTR of Tc-zen1 and the initiation codon of Tc-zen2. Nonetheless, subsequent functional diversification has equipped the paralogues with either of the two known EEM functions: early-acting Tc-zen1 specifies the serosal tissue, while Tc-zen2 is required for late EEM withdrawal morphogenesis. We thus undertook a detailed molecular characterization of the beetle paralogues to elucidate the evolutionary history of zen functional changes, and more generally to shed light on the regulatory implications of paralog retention and diversification after a gene duplication event.

Here we present differences in the regulation of Tc-zen1 and Tc-zen2 as well as in their own transcriptional signatures as homeodomain transcription factors, providing the first detailed functional dissection of insect zen duplicates outside of Drosophila. Surprisingly, peak expression does not coincide with the time of primary function—detectable morphologically and transcriptionally—for Tc-zen2, which despite its late role has strong early expression like Tc-zen1. Yet, instead of a lack of function or shadow redundancy to Tc-Zen1, we uncover a distinct early role of Tc-Zen2 in the regulation of a key subset of genes, including the rapid repression of Tc-Zen1. RNA-seq data also reveal subtle aspects of temporal variability (heterochrony) after Tc-zen2 RNA interference (RNAi) that affect late morphogenesis. Our validation of specific transcriptional targets opens new avenues into serosal tissue biology and identifies a novel, paralogue-based regulatory circuit at the developmental transition from specification to maturation of the serosa. This now raises the question of how species with a single zen gene compare for the precision and progression of EEM development, and whether their molecular phenotypes support early Tc-Zen2 function as the consequence of iterative subfunctionalization events.

Results
Recent tandem duplication of zen in the Tribolium lineage. We first surveyed Tribolium beetle genomes to assess sequence conservation at the Hox3 locus. Using the T. castaneum paralogues as BLASTn queries, we find that the tandem duplication of zen is conserved across three closely related congeners: Tribolium freemanii, Tribolium madens, and Tribolium confusum (Fig. 1a, 18–61 million years divergence). Consistent with a recent event, phylogenetic analysis supports a single duplication at the base of the Tribolium lineage, and sequence alignments show particularly strong conservation in the homeobox, encoding the DNA-binding homeodomain (Fig. 1b; Supplementary Figs. 1 and 2).

Next, we investigated levels of coding sequence conservation between the T. castaneum zen (hereafter “Tc-zen”) paralogues. Strongest nucleotide conservation occurs within the homeobox, where three conservation peaks correspond to the three encoded α-helices (Fig. 1c: >80% identity). In fact, within the coding sequence for the third α-helix there is a 20-bp stretch with 100% nucleotide identity (Fig. 1c), which is roughly the effective length of sequence for achieving systemic knockdown by RNAi. Indeed, Tc-zen1-specific double-stranded RNA (dsRNA) that spans the homeobox is sufficient to effect significant cross-paralogue knockdown of Tc-zen2 (Fig. 1d: 19% of wild-type expression with the long dsRNA fragment compared to only 35% with the short fragment; beta regression, odds ratio of the short fragment vs. long fragment = 2.36, 95% CI = 1.65–3.37, p = 2.38e−6). As Tc-zen2 is expressed in the tissue domain specified by Tc-Zen1, it is in fact surprising that any residual Tc-zen2 transcript is detected in the Tc-zen1 RNAi background, an aspect we revisit in the “Discussion”. Meanwhile, for our experimental design we find that indeed a short fragment alone is sufficient to strongly knock down Tc-zen1 itself (no significant change in knockdown efficiency between the long and short fragments: 10 and 11% expression, respectively; beta regression, odds ratio of the short fragment vs. long fragment = 1.12, 95% CI = 0.75–1.66, p = 0.577). For all subsequent functional testing we thus designed our dsRNA fragments to exclude the homeobox, thereby avoiding off-target effects and ensuring paralogue-specific knockdown (Fig. 1c: Tc-zen1 short fragment: yellow; Tc-zen2: green; no ≥20-bp stretches of nucleotide identity in these regions).

Distinct paralogue roles at different developmental stages. EEM development has been well characterized morphologically in the beetle, including the Tc-zen paralogues’ roles. Briefly, the first differentiation event distinguishes the serosa from the rest of the blastodermal cell sheet (Fig. 2a, at ~10% embryonic development). Tissue reorganization then involves serosal expansion and internalization of the embryo and amnion (EEM formation: subdivision into the “primitive pit” and “serosal window” stages). This topology is later reversed when the EEMs actively rupture and contract (“withdrawal”), coordinated with expansion of the embryo’s flanks for dorsal closure of the body (Fig. 2c, at ~75% development). After Tc-Zen1 RNAi, presumptive serosal cells are respecified to other anterior fates, leading to an early enlargement of the head and amnion and loss of serosal identity (Fig. 2b).
Tc-zen2 RNAi impairs or wholly blocks late EEM withdrawal\textsuperscript{12,26}, confusing the embryonic flanks such that the epidermis encloses the embryo’s own legs instead of closing the back, leading to an everted (inside out) configuration (Fig. 2d\textsuperscript{12,32}). Here, we were able to reproduce the morphological phenotypes after RNAi for each Tc-zen paralogue (Fig. 2a′–d′\textsuperscript{12}). RNAi is particularly efficient for Tc-zen1 (98.8% knockdown, Fig. 2e\textsuperscript{12,32}). Specific phenotypes after Tc-zen2 RNAi (73.8% knockdown) include complete eversion (20.5%, Fig. 2d\textsuperscript{12}), as well as milder defects in EEM withdrawal (53.3%, Fig. 2f; Supplementary Fig. 3). Furthermore, we newly explored how the paralogues’ functions relate to their transcript expression profiles across embryogenesis. Consistent with their functions, Tc-zen1 has early expression while only Tc-zen2 persists until the membrane rupture stage (Fig. 2g). Unexpectedly, late-acting Tc-zen2 has peak transcript expression during early development.

**Subtle expression differences during early development.** To gain insight into Tc-zen gene regulation and to determine the developmental stages of primary transcription factor function for each paralogue, we undertook a fine-scale spatiotemporal
**Fig. 2** Tc-zen paralogue roles in early specification (Tc-zen1) or late EEM morphogenesis (Tc-zen2). Phenotypic consequences of RNAi for the Tc-zen paralogues are depicted: schematically at the time of primary defect (a–d, upper row) and with micrographs for the resulting phenotypes (a’–d’, lower row: DAPI nuclear stain (a’, b’) and cuticle preparations (c’, d’); see also Supplementary Fig. 3). The dashed lines in the schematics in a, b denote the anterior embryonic border. Anatomical abbreviations: a antenna; h head; l leg; t telson; t3 third thoracic segment. Scale bars are 100 μm. e, f RNAi phenotypic penetrance per paralogue. Sample sizes are numbers of embryos. g Expression profiles of Tc-zen1 and Tc-zen2 in early (6–14 hAEL) and late (42–52 hAEL) development (RT-qPCR). Mean expression levels are shown from four biological replicates; error bars represent one standard deviation. For clarity, mean values are stated here, with the individual source data values given in Supplementary Data 5. Staging abbreviations: BF blastoderm formation/uniform blastoderm; DB differentiated blastoderm; PP primitive pit; SW serosal window; GBE extended germband; pre-R pre-rupture; MR extraembryonic membrane rupture. Time is hours after egg lay (hAEL).
characterization of Tc-zen1 and Tc-zen2 expression for both transcript and protein (RT-qPCR, in situ hybridization, western blotting, immunohistochemistry).

As both paralogues are strongly expressed in early development (Fig. 2g), we first examined these stages in detail. Tc-zen1 transcript arises in an anterior gradient during blastoderm formation (4–6 h after egg lay, hAEL), peaks at the differentiated blastoderm stage with uniform expression throughout the presumptive serosa (6–10 hAEL), and then becomes patchy and retracts to a narrow region at the tissue’s border during EEM formation (10–14 hAEL; Fig. 3a–f). After the EEMs have fully enclosed the early embryo, Tc-zen1 transcript is no longer detected (Figs. 2g and 3a). Peak Tc-zen1 transcript expression is followed shortly by detectable protein for Tc-Zen1, although this, too, only occurs during early development (Fig. 4a; Supplementary Figs. 4 and 5).

Tc-zen2 expression begins slightly later, at the differentiated blastoderm stage (6–8 hAEL), with peak levels occurring during EEM formation (10–14 hAEL; Fig. 3a). We also observed spatial differences between the paralogues. Tc-zen2 is first detected only in an anterior subset of the serosa when Tc-zen1 is expressed in the entire tissue (compare Fig. 3c with 3h). Then, Tc-zen2 transcript expands throughout the serosa while Tc-zen1 transcript retracts, concomitant with the expansion of the entire serosal tissue during EEM formation (compare Fig. 3d–f with 3i–k; Supplementary Fig. 5). Notably, the Tc-zen paralogues are expressed consecutively, but not concurrently, at the rim of the serosa. It is only during late EEM formation that we first observe Tc-zen2 expression throughout the entire serosal tissue (Fig. 3k). By this time, Tc-Zen2 protein is also strongly expressed and persists (Fig. 4a, b; Supplementary Figs. 4, 5, and see below), while Tc-zen2 transcript wanes gradually (from 14 hAEL; Fig. 3a).

Fig. 3 Transcript expression dynamics of the Tc-zen paralogues during early embryogenesis. a Quantification of mean transcript levels in 2-h intervals (RT-qPCR), from four biological replicates; error bars represent one standard deviation. For clarity, mean values are stated here, with the individual source data values given in Supplementary Data 5. Whole mount in situ hybridization for Tc-zen1 (b–f) and Tc-zen2 (g–k), with nuclear counterstains for morphological staging (b–k′, arrows label the expanding serosal border). All micrographs are oriented with anterior left and shown in lateral aspect with dorsal up (except in (b) and (g), which depict stages before this can be determined). Scale bars in (b) and (g) are 100 µm and apply to (b–f′) and (g–k′), respectively.
Early transcriptional impact of Tc-zen1 and Tc-zen2.

Since protein expression follows shortly after peak transcript expression for both paralogues (Figs. 3a and 4a), we used the high sensitivity of our RT-qPCR survey (Fig. 3a) to inform our staging for functional testing by RNAi. To identify transcriptional targets for each zen gene, our RNA-seq after RNAi approach assessed differential expression (DE) between age-matched wild-type and knockdown samples. We focused specifically on the time windows of peak gene expression: 6–10 hAEL for Tc-zen1 and 10–14 hAEL for Tc-zen2 (curly brackets in Fig. 3a). These 4-h windows were chosen to maximize the number of identified target genes while prioritizing direct targets for Zen transcription factor binding.

The RNA-seq data are consistent with a priori expectations based on the morphological consequences of RNAi for each zen gene (Fig. 2a–d). That is, Tc-zen1 has a clear early role in tissue specification, and its knockdown at these stages has a strong transcriptional impact, wherein principal component analysis (PCA) clearly distinguishes experimental treatments (Fig. 5a). In contrast, Tc-zen2 has an early expression peak but its manifest role in late EEM withdrawal occurs nearly 2 days later (56% development later). Accordingly, we find a negligible effect on the early egg's total transcriptome after Tc-zen2 RNAi (Fig. 5a), despite verification of efficient knockdown (Fig. 2f). RNAi efficiency was also confirmed directly with the RNA-seq data:
both Tc-zen1 and Tc-zen2 exhibit DE reduction after their respective knockdown (Supplementary Data 1A, B: significant fold change reductions of $-4.86$ for Tc-zen1 and $-3.54$ for Tc-zen2). Overall, we obtained 338 DE genes after Tc-zen1 RNAi compared to only 26 DE genes after Tc-zen2 RNAi, while global transcriptional changes affect nearly 12% of all genes during early embryogenesis (2221 DE genes; threshold of $P_{\text{adj}} \leq 0.01$ and $|\text{FC}| \geq 2$ for all DE genes: Fig. 5f bar chart elements a, c, d; Supplementary Data 1A–C).

Given the recent nature of the duplication, which is evident in the similarity of the Tc-zen paralogues’ DNA-binding homeodomains and early expression profiles, we asked whether there is

---

**Developmental stages:**
- early 1: 6-10 h AEL (zen1 peak)
- early 2: 10-14 h AEL (zen2 peak)
- late 1: 48-52 h AEL (pre-rupture)
- late 2: 52-56 h AEL (withdrawal)
a legacy of shared early function. If this is the case, Tc-zen2 might exhibit a subtle regulatory profile similar to Tc-zen1. However, even with relaxed thresholds for DE, we find few shared targets between the paralogues, particularly when the direction of regulation is considered (Fig. 5b; Supplementary Data 2A, B). Thus, we conclude that Tc-zen2 has a minimal impact on early development, and that this does not constitute a transcriptional “echo” of co-regulation with Tc-zen1 due to common ancestry. Why, then, is Tc-zen2 strongly expressed during early development?

The Tc-zen paralogues are mutual regulatory targets. We next considered the Tc-zen paralogues as factors necessary for defining the serosal tissue, as indicated by their specific transcriptional targets. Tc-zen1 is strictly required for serosal tissue identity26, Differentiation of the serosa involves an early switch from mitosis to the endodcyt29,30, resulting in polyploidy13. Consistent with this, we identified a homolog of the endodcyt factor fizzly-related33,34 among DE genes upregulated by Tc-zen1 (Supplementary Data 1A). From known targets of Tc-zen1, we also recovered Dorsocross and hindsight, involved in EEM formation35, and chitin synthase 1, required for production of the protective cuticle.10 In addition, we hypothesized that the slight offset whereby Tc-zen1 expression precedes Tc-zen2 is consistent with Tc-zen1 activating Tc-zen2. We could confirm this regulatory interaction both by RNA-seq and RT-qPCR after Tc-zen1 RNAi (Fig. 6a, b). Thus, Tc-zen1 as a serosal specifier upregulates factors for definitive tissue differentiation, including Tc-zen2 as a candidate (Fig. 6i).

Are there Tc-zen2 transcriptional targets that could support an early role in the serosa for this paralogue? Among the few genes with strong DE (Fig. 5f bar chart element d), we validated several as probable targets. These genes are expressed in the early serosa and/or their transcript levels are first strongly upregulated during peak Tc-zen2 expression (12–14 hAE; e.g., Supplementary Fig. 6). Their putative functions as enzymes or structural components for chitin-based cuticle (Cpr proteins) and as signaling molecules support a role for Tc-zen2 in the physiological maturation of the serosa, providing complementary regulatory control to Tc-zen1.

In performing reciprocal validation assays, we then uncovered an unexpected early function of Tc-zen2 in the repression of its own paralogous activator. After Tc-zen2 RNAi, Tc-zen1 transcript is still detectable, a qualitative observation that previously had been taken as evidence for a lack of paralogue feedback26. The sensitivity of our quantification and spatial analyses across the dynamic early stages of development now allows us to correct this. We detect an upregulation of Tc-zen1 that was only weakly suggested by our RNA-seq data but then strongly supported in RT-qPCR assays (Fig. 6a, b). We also confirmed this observation by in situ hybridization. After Tc-zen2 RNAi, Tc-zen1 transcript is expressed at higher levels than in wild type (compare Fig. 6c, d with 6f, g). Tc-zen1 also remains strongly expressed throughout the serosa at stages when its wild-type expression is restricted to low levels at the tissue rim (compare Fig. 6e with 6h). In fact, the abrupt reduction in Tc-zen1 transcript and protein levels in wild type correlates with increasing Tc-zen2 levels, and spatially their dynamic expression is largely complementary, if not outright mutually exclusive (Figs. 3 and 4). Together, these results suggest that Tc-Zen1 upregulates Tc-zen2 in its wake, and that in turn early Tc-Zen2 represses Tc-zen1. Thus, the Tribolium paralogues function as mutual regulatory targets, comprising an integrated regulatory module for early serosal development (Fig. 6i).

Tc-Zen2 is exclusively serosal and persistently nuclear. To complete our analysis of Tc-zen2, we also examined its activity at later stages. We could detect both transcript (weakly, Figs. 2g and 3a) and protein (particularly strongly in mid-embryogenesis, Fig. 4a) continuously until the stage of EEM withdrawal, spanning 14–75% of development (10–54 hAE, assayed in 2-h intervals; see also Supplementary Fig. 4). Moreover, we find that Tc-zen2 is persistently localized to the nucleus, demonstrated by fluorescent immunohistochemistry on cryosectioned material of selected stages (Fig. 4b–e, g, h). This contrasts with some species’ orthologues, which show stage-specific exclusion of Zen protein from the nucleus39. We could also refine the spatial scope of Tc-zen2 activity: in contrast to earlier reports26, we found no evidence for Tc-zen2 transcript or protein in the amnion (Fig. 4d–h’y”), indicating that this factor is strictly serosal.

Late transcriptional dynamics are largely Tc-zen2-dependent. The early RNA-seq after RNAi experiment examined the time of peak Tc-zen2 expression. Complementing this, we used the same approach to examine the stage of known Tc-zen2 function in late EEM withdrawal. Withdrawal begins with rupture of the EEMs, at 52.1 ± 2.3 hAE as determined by live imaging31. Here, we assayed the 4-h intervals just before (48–52 hAE) and after (52–56 hAE) rupture, to assess Tc-zen2 transcriptional regulation that precedes and then accompanies withdrawal. Consistent with Tc-zen2’s known role, we detect >16x more DE genes after Tc-zen2 RNAi in late development (>430 DE genes, compare Fig. 5f bar chart elements e, f with 5f bar chart element d). PCA also clearly separates knockdown and wild-type samples at late stages (Fig. 5c).

Our staging helps to contextualize Tc-zen2 and EEM-specific processes relative to concurrent embryonic development. We thus evaluated DE in pairwise comparisons not only between wild-type and RNAi samples, but also over time in both backgrounds (Fig. 5d, f bar chart elements b, e–g). Comparisons across
consecutive stages in early and late development (Fig. 5f bar chart elements a, b) reveal two general changes in the wild-type transcriptional landscape. There is far less dynamic change in gene expression in late development (5.8x fewer DE genes), consistent with steady state and ongoing processes in later embryogenesis compared to the rapid changes of early development. Also, whereas early development shows a fairly even balance between up- (48%) and downregulation (52%), late development is predominantly characterized by increasing expression levels over time (79%).

Against this backdrop, the transcriptional impact of Tc-zen2 is quite pronounced. Most genes with changing expression over time in the late wild-type background are also affected by Tc-zen2 RNAi (Fig. 5d: 77%, 293/383 DE genes from green Venn diagram).
paralogues comprise an integrated view of the full names and descriptions in the main text. Subfunctionalization. Primary subfunctionalization separated early species of Tribolium. Progressive partitioning of the functions occurred through stepwise acquisition of paralogue mutual regulation, resulting in iterative subfunctionalization. Primary subfunctionalization separated early specialization (sp) and late morphogenesis (mo) functions. Secondly, the original specification function has now been finely subdivided into the initial, substantial role that Tc-Zen1 continues to play (“sp1”) and the slightly later, subtle functions of Tc-Zen2 that we reveal here (“sp2”).

set). We detect this strong effect even though Tc-zn2 is restricted to the serosa (Fig. 4), a tissue that ceased mitosis (Fig. 6i) and comprises only a small cell population within our whole-egg samples. This suggests that most dynamic transcription in late development pertains to EEM morphogenesis, with the global transcriptional impact of Tc-zn2 at these stages even greater than for Tc-zen1 in early development (compare Fig. 5f bar chart elements e, f with 5f bar chart element c). Most candidate Tc-zn2 targets are differentially expressed at a single stage (72%), although a substantial fraction (26%) exhibits consistent activation or repression, while an intriguing handful of genes shows changing, stage-specific regulation (Fig. 5e). These patterns imply that the persistent nuclear localization of Tc-Zen2 (Fig. 4) reflects active transcriptional control, not merely localization to the nucleus or DNA binding in a paused, nonfunctional state.

To characterize late Tc-Zen2 activity, we functionally annotated and validated candidate transcriptional targets. Gene ontology (GO) enrichment tests confirmed that ongoing cuticle regulation is a primary role, including remodeling as the serosa detaches from its own cuticle in preparation for withdrawal (Supplementary Fig. 7, Supplementary Data 4A, B). For validation, we selected a dozen genes based on known biological processes for tissue remodeling (e.g., cytoskeleton and morphogenesis), prominent GO categories (e.g., transmembrane transporters), and evidence of dynamic regulation (Fig. 5e; Supplementary Fig. 8, Supplementary Tables 1 and 2). All tested candidates were confirmed by RT-qPCR (Supplementary Fig. 8). This included two of the genes that are first activated and then repressed by Tc-Zen2, where both genes encode proteins with conserved domains of unknown function (Supplementary Table 2). Notably, such dynamic targets include a subset of the Osiris multigene locus (Supplementary Fig. 8). Lastly, we evaluated Tc-Zen2 regulation of serosal immune genes. Although our samples were not pathogen challenged, we could detect expression for 83% of these genes (89 of 107 genes), with 20% showing DE after Tc-zn2 RNAi (Supplementary Data 3A, B). Thus, while Tc-Zen2 is not a global effector, it may regulate subsets of immune genes. Notably, transcripts of most serosal immune genes (87 genes) continue to be detected during withdrawal, supporting their expression as an inherent feature of the serosa—even when it is no longer a protective layer enclosing the embryo.

Several observations are consistent with a delay. As noted above, all late RNA-seq biological replicates cluster by treatment in PCA. Interestingly, the older Tc-zn2RNAi samples (52–56 hAEL) have intermediate component scores compared to the younger Tc-zn2RNAi and younger wild-type samples (48–52 hAEL, Fig. 5c). Similarly, DE comparisons identify noticeably fewer DE genes between the older Tc-zn2RNAi sample and either of the younger samples (Fig. 5f bar chart elements g, h; Supplementary Data 3D, E). In fact, the very low number of DE genes implies that there is virtually no difference in the transcriptional profile of the older Tc-zn2RNAi sample compared to the younger wild-type sample (Fig. 5f bar chart element h).

Furthermore, nearly all genes that change in expression over time in the Tc-zn2RNAi background are also candidate targets of Tc-Zen2 at the rupture stage (95%, Fig. 5d: inset Venn diagram). In other words, Tc-zn2RNAi eggs generally require an additional 4 h (5.6% development) to attain a transcriptional profile comparable to the wild-type pre-rupture stage, and this is achieved by belated activation of Tc-Zen2 target genes. However, only a subset of genes exhibit delayed recovery (34%, Fig. 5d inset). These target genes may thus be independently activated by other factors, in addition to activation by Tc-Zen2.

Our RNA-seq data also reveal increased variability after Tc-zn2 RNAi. The pre-rupture Tc-zn2RNAi biological replicates show comparably tight clustering to their age-matched wild-type counterparts (48–52 hAEL, Fig. 5c). This suggests that pre-rupture is the stage of primary Tc-Zen2 function, also supported by our detection of the greatest number of DE genes at this stage (compare Fig. 5f bar chart elements d, e, f). In contrast, the older Tc-zn2RNAi samples have a noticeably greater spread along the vectors of the first two principal components (52–56 hAEL, Fig. 5c), consistent with cumulative variability as the RNAi phenotype develops, presumably in part due to the observed partial transcriptional recovery (Fig. 5d). This variability may in itself provide explanatory power for the spectrum of end-stage Tc-zn2RNAi phenotypes (Fig. 2f, see below).

Discussion

Our analysis of regulation upstream and downstream of the beetle zen genes reveals several unexpected features regarding the evolution and biological roles of these unusual paralogues.

First, sequence conservation belies the extent of zen paralogue functional divergence. Fine-tuned transcriptional regulation is required to restrict regulatory crosstalk, and conserved non-coding regions may contribute to this. The region upstream of zen1 has particularly high conservation and was tested as an in vivo Tc-zen1 reporter in a recent study (Fig. 1a: dashed line region). This construct recapitulates expression at the rim of the serosal window, a feature common to both paralogues (as in
Our own observations that detectable, residual levels of Tc-zen2 persist after Tc-zen1 RNAi (Fig. 1d: short fragment) and that low but detectable levels of Tc-zen2 are expressed during blastodermal formation (4–6 h AEL, compare Fig. 3a with 3g) imply early Tc-zen2 regulation that is independent of Tc-Zen1, possibly involving unknown anterior terminal activators relevant for Tc-zen1 itself. Thus, regulation of the Tc-zen genes requires multiple inputs that remain to be elucidated.

Specificity of regulation by the Tc-zen genes is also elusive. In the Tc-zen homeobox, sequence similarity is particularly high in the third α-helix, which confers DNA-binding specificity (Fig. 1c, d)[30,40,41]. Yet, the paralogues’ shared ancestry is not reflected in redundant activity (Fig. 5a, b). Rather, strong conservation across Tribolium species, particularly of zen2 (Supplementary Fig. 1), may indicate not only limited divergence but also positive, purifying selection[52]. How, then, do the paralogues regulate different targets? In contrast, Hox3 proteins, DNA-binding specificity can be enhanced by the common Hox cofactor Extrathorax[11]. In contrast, insect Zen proteins have lost the hexapeptide motif required for this interaction, and no other cofactor binding motifs are known[16], deepening the long recognized “Hox specificity paradox”[43] in the case of the beetle zen paralogues.

Recent work has revealed a role for differential chromatin accessibility in conferring Hox binding specificity[44]. This highlights an intriguing direction for future research, on the extent of rapid functional genomic changes that may occur in early development. However, the stage-specific and gene-specific regulatory activity of Tc-Zen2 (Fig. 5e; Supplementary Fig. 8) argues for greater regulatory precision than expected based on opportunistic binding to accessible targets alone.

Meanwhile, our molecular dissection of the Tc-zen paralogues elucidates their functional divergence. Specifically, mutual regulation has implications for the paralogues’ network logic and confers temporal precision. The newly discovered negative feedback loop of Tc-Zen1 activation leading to repression by Tc-Zen2 constitutes a tight linkage. To what extent could Tc-zen1 overexpression by upregulation of Tc-zen2 as its target, resulting in repression of Tc-zen1 and thus canceling out the manipulation? In fact Tc-zen2 RNAi does confer overexpression of Tc-zen1 and reduced Tc-zen2 (Fig. 6; Supplementary Data 1B). However, consistently lower knockdown efficiency for Tc-zen2 than for Tc-zen1 (Fig. 2, ref. 26) may reflect a dose-limiting lack of regulatory disentanglement. Arguably, Tc-zen1 and Tc-zen2 together satisfy the criteria of a minimal gene regulatory network (GRN) kernel[45], including “recursive wiring” and the experimental challenges this entails. Alternatively, the Tc-zen paralogues could be viewed as a single unit in a serosal GRN and thus qualify as a “paradoxical component” that both activates and inhibits (Fig. 6i)[46]. Consistent with theoretical expectations, delayed inhibition produces a discrete pulse of Tc-zen1 (Figs. 3 and 4). As the pulse is non-oscillatory, this may also imply that Tc-zen2 is a positive auto-regulator[46], a property known for a number of canonical Hox genes[47,48] but thus far not known for insect zen genes.

Furthermore, Tc-Zen2 was previously implicated in a second regulatory innovation, Tc-zen2 for early specification, Tc-zen2 for late morphogenesis. A second regulatory innovation, Tc-Zen1’s activation of Tc-zen2, then conferred the temporal offset in early paralogue activity and resulted in a second, nested instance of subfunctionalization, for serosal specification (Fig. 6i), j). Thus, whereas the loss of autoregulation may have contributed to subfunctionalization of certain Hox genes among the four paralogous Hox clusters of vertebrates[66], here we infer that the acquisition of paralogue mutual regulation drove active, iterative subfunctionalization.

We have uncovered multiple roles of Tc-zen2 as a diverged Hox gene throughout the lifetime of the serosal tissue, itself a morphological innovation[5]. Early, Tc-Zen2’s repression of Tc-zen1 (Fig. 6) and Tc-caudal[19] is noteworthy. A predominantly repressive role contrasts with Hox genes typically serving as activators, as do both Tc-zen paralogues at the stages of their primary function (Fig. 5f) bar chart elements, c, e). Also, the precise mechanism and targets of potential Tc-zen2 translational repression remain open questions. Future work will clarify whether such a function arose independently in Tribolium Zen2 and dipiteran Bicoid[20,23,40] as distinct Hox3/zen derivatives. In later development, the serosa is the cellular interface with the outer environment. Our data elucidate Tc-Zen2’s roles in the known protective functions of cuticle formation[10] and innate immunity[7].
Beyond this, our DE gene sets comprise a large, unbiased sample of candidate targets, laying the foundation for investigating wider roles of Tc-zen2 in this critical tissue.

Finally, we identify Tc-zen2-dependent EEM withdrawal as the major transcriptionally regulated event in late embryogenesis and assess its precision (Fig. 5d). Temporal and molecular variability after Tc-zen2 RNAi underpins observed variability in knockdown EEM tissue structure, integrity, and morphogenetic competence, defining the broad spectrum of end-stage phenotypes (Fig. 2; Supplementary Fig. 3). This ranges from mild defects in dorsal closure after transient EEM obstruction to persistently closed EEMs that cause complete eversion of the embryo (Fig. 2; Supplementary Fig. 3, ref. 12). The unifying feature is a heterochronic shift of extraembryonic compared to embryonic developmental processes (delayed EEM withdrawal compared to epidermal outgrowth for dorsal closure).

There may also be species-specific differences in the timing of Zen function for withdrawal morphogenesis. The sole zen orthologue in the milkweed bug Oncopeltus fasciatus has a similar persistent expression profile and a spatiotemporal role in withdrawal, termed “katatogenesis” in this and other hemimetabolous insects37. We previously observed a number of Of-zen-dependent, long-term morphological changes prior to rupture65, contrasting with the more proximate effect of Tc-zen2 (discussed above). Taking the work forward, it will be interesting to compare Tc-zen2 and Of-zen transcriptional changes. Evaluating conserved regulatory features of EEM withdrawal across the breadth of the insects will clarify macroevolutionary patterns of change in the very process of epithelial morphogenesis.

Methods

*Tricholoma castaneum* stock husbandry. All experiments were conducted with the San Bernardino wild-type strain, maintained under standard culturing conditions at 30 °C and 40–60% relative humidity66.

In silico analyses. Draft genome assemblies for *T. freemani*, *T. madens*, and *T. confusum* were obtained as assembled scaffolds in FASTA format (version 26 March 2013 for each species), accessed from the BeadleBase.org FTP site at Kansas State University ([ftp://bioinformatics.k-state.edu/pub/BeadleBase/]). Transcripts for Tc-zen1 [TC000921-RA] and Tc-zen2 [TC000922-RA] were obtained from the *T. castaneum* official gene set 3 (OGS)47. These sequences were used as queries for BLASTn searches in the other species’ genomes (BLAST + 2.2.30)16,67. Sequences were extracted to comprise the *Hox3-zen* genomic loci, spanning the interval from 5 kbp upstream of the BLAST hit to the 5’ UTR of Tc-zen1 to 5 kbp downstream of the BLAST hit for the 3’ UTR of Tc-zen2. These genomic loci were then aligned with the mVista tool70,71 using default parameters. Nucleotide identities were calculated for a sliding window of 100 bp. The maximum likelihood phylogenetic tree (Fig. 1b) was constructed based on an alignment of full-length Zen proteins, with gaps permitted, using the Phylogeny.fr default pipeline settings, with MUSCLE v3.8.1 alignment and PhyML v3.1.
RNA-seq after RNAi. RNAi for RNA-seq used the short, paralogue-specific dsRNAs (depicted schematically in Fig. 1e). For transcriptomic profiling by RNA-seq, three separate Tc-zen1RNAi experiments were conducted, each with a separate cohort of injected females. A total of seven Tc-zen2RNAi experiments were conducted: one for each biological replicate at each developmental stage. Samples chosen for sequencing were assessed by RT-qPCR for level of knockdown in RNAi samples, with Tc-zen1 reduced to <10% of wild-type levels and Tc-zen2 to 24% across biological replicates. For early development (6–14 AEL), three biological replicates were sequenced for each experimental treatment, with 100-bp paired end reads on an Illumina HiSeq2000 machine. For late development (48–56 AEL), four biological replicates were sequenced with 75-bp paired end reads on a HiSeq4000 machine. All sequencing was performed at the Cologne Center for Genomics, with six (HiSeq2000) or eight (HiSeq4000) multiplexed samples per lane yielding ≥6 Gbp per sample. The complete dataset of all 56 RNA-seq samples is deposited in GenBank (NCBI) under BioProject accession number PRJNA645519.

The quality of raw Illumina reads was examined with FastQC38, and all RNA-seq samples were retained for analysis. The adapter sequences and low quality bases were removed with Trimmomatic v0.3639. Trimmomatic was also used to shorten 100-bp reads from the 3′ end to 75-bp reads to increase mapping efficiency (Supplementary Table 4)41. The overrepresented sequences of mitochondrial and ribosomal RNA were filtered out by mapping to a database of 1266 T. castaneum mitochondrial and ribosomal sequences extracted from the NCBI nucleotide database (accessed 21 October 2016, search query "trituberculis (organism) AND (ribosomal OR mitochondrial OR mitochondrion) NOT (whole genome shotgun) NOT (Karoozra purpurea)") with Bowtie2 v2.2.931. Trimmed and filtered reads were mapped to the T. castaneum OGS34 (file name: Tcars2_GenBank_corrected.v5.renamed.mrna.fa) with RSEM45. The raw read count output from RSEM was compiled into count tables. Both principal component and DE analyses were performed in R using the package DESeq2 v1.14.146 with default parameters. For PCA, raw (unfiltered) read counts were used. For DE analyses, to eliminate noise all genes with very low read counts were filtered out by sorting in Microsoft Excel, following recommendations47. Specifically, genes were excluded from DE analysis if read counts ≤10 in ≥1 biological replicates for both the knockdown and wild-type samples. For serosal immune genes, a given gene was considered to be expressed in the late serosa if we detected ≥100 RNA-seq reads in each of the four biological replicates for our wild-type samples. Throughout, our reporting of "DE gene" refers to analyses across all isoforms (18,536 isoforms models) in the T. castaneum official gene set OGS37.

Gene ontology (GO) analyses. GO enrichment analysis was performed with Blast2GO49 using two-tailed Fisher’s exact test with a threshold false discovery rate of 0.05. GO term analysis was performed by Blast2GO against the Drosophila database (accessed 9 June 2017). Only GO terms from the level 5 were considered. Next, GO terms were grouped into categories of interest based on similarity in function (Supplementary Table 1). Afterwards a unique count of T. castaneum gene sequences was calculated for each category of interest and the percentage was compared to the rest of the GO terms in the level 5 for each GO domain (Supplementary Fig. 8).

Statistics and reproducibility. All reported results were reproducible in our hands and consistent with published results with these methods and genes12,13,30,35. In addition to the use of robust biological replicates for any one technique, expression data were corroborated across RT-qPCR, RNAbinding methods (mRNA transcript) or across western blot and immunohistochemistry (protein). RT-qPCR and RNA-seq analyses are based on 3–4 biological replicates, as indicated in the specific "Results" and "Methods" sections and associated figure legends (Figs. 1d, 2g, 3a and 6b; Supplementary Fig. 8B), and the source data values for these are available in Supplementary Data 5. Each biological replicate represents an independent sample and was derived from a different egg collection, with RNAi experiments conducted to obtain material from at least three distinct maternal cohorts. As noted above, RT-qPCR data were evaluated using LinRegPCR v12.1652,56, and RNA-seq data were evaluated and processed with FastQC51 and Trimmomatic v0.3657. To evaluate knockdown efficiency of long and short dsRNA fragments, beta regression analyses were performed with betareg v3.1-074, as noted above. DE statistical analyses were conducted with DESeq2 v1.14.186, and GO enrichment was determined with Blast2GO58, as noted above.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability. All processed data and analyses generated during this study are included in this published article and its Supplementary Information files, including source data: Supplementary Data 1–3: tables of differentially expressed genes from all comparisons; Supplementary Data 4: gene ontology (GO) terms for differentially expressed genes after Tc-zen2 RNAi in late development; Supplementary Data 5: source values and dual plotting (means, individual values) for bar charts in figures (Figs. 1d, 2g, 3a and 6b; Supplementary Fig. 8B). The RNA-seq raw read data generated during the current study are available in GenBank (NCBI), under BioProject accession number PRJNA645519. The paralogue-specific peptide antibodies are available on request from the corresponding author or from the source laboratory59.

Received: 19 September 2019; Accepted: 21 August 2020;
Published online: 05 October 2020

References
1. Preumont, R. Evolution of the vertebrate Hox homeobox genes. BioEssays: News Rev. Mol., Cell. Dev. Biol. 14, 245–252 (1992).
2. Arnisotso, D. et al. The genome of the water strider Gerris bilineatus reveals expansions of gene repertoires associated with adaptations to life on the water. BMC Genom. 19, 832 (2018).
3. McKenna, D. D. et al. Genome of the Asian longhorned beetle (Anoplophora glutinata), a globally significant invasive species, reveals key functional and evolutionary innovation at the beetle–plant interface. Genome Biol. 17, 227 (2016).
4. Negre, B. & Ruiz, A. HOM-C evolution in Drosophila: is there a need for Hox gene clustering? Trends Genet. 23, 55–59 (2007).
5. Panfilio, K. A. Extraembryonic development in insects and the acrobatics of blastokinesis. Dev. Biol. 381, 431–491 (2008).
6. Chen, G., Handel, K. & Roth, S. The maternal NF-κB/Dorsal gradient of Tribolium castaneum: dynamics of early dorsoventral patterning in a short-germ beetle. Development 127, 5145–5156 (2000).
7. Jacobs, G. C. G., Spank, H. P. & van der Zee, M. The extraembryonic serosa is a further epithelium protecting the insect egg with a full range innate immune response. eLife 3, e01114 (2014).
8. Cobben, R. H. Evolutionary Trends in the Heteroptera, Part I, Eggs. Architecture of the Shell, Gross Embryology and Elosion. (Centre for Agricultural Publishing and Documentation, 1968).
9. Rezende, G. L. et al. Embryonic desiccation resistance in Aedes aegypti: presumptive role of the chitinized serosal cuticle. BMC Dev. Biol. 8, 82 (2008).
10. Jacobs, G. C. G., Rezende, G. L., Lamers, G. E. M. & van der Zee, M. The extraembryonic serosa protects the insect egg against desiccation. Proc. R. Soc. B 280, 20131082 (2013).
11. Farnesi, L. C., Menna-Barreto, R. F. S., Martins, A. J., Valle, D. & Rezende, G. L. Physical features and chitin content of eggs from the mosquito vectors Aedes aegypti, Anopheles aquasalis and Culex quinquefasciatus: connection with distinct levels of resistance to desiccation. J. Insect Physiol. 83, 43–52 (2015).
12. Hilbrant, M., Horn, T., Koelzer, S. & Panfilio, K. A. The beetle amnion and serosa functionally interact as apposed epithelia. eLife 5, e13384 (2016).
13. Panfilio, K. A., Oberhofer, G. & Roth, S. High plasticity in epithelial morphogenesis during insect dorsal closure. Biol. Open 2, 1108–1118 (2013).
14. Horn, T., Hilbrant, M. & Panfilio, K. A. Evolution of epithelial morphogenesis: phenotypic integration across multiple levels of biological organization. Front. Genet. 6, 303 (2015).
15. Ferguson, L. et al. Ancient expansion of the Hox cluster in Lepidoptera generated four homeobox genes implicated in extra-embryonic tissue formation. PLoS Genet. 10, e1004698 (2014).
16. Panfilio, K. A. & Akam, M. A comparison of Hox3 and Zen protein coding sequences in taxa that span the Hox3/zen divergence. Dev. Genes Evol. 217, 323–329 (2007).
17. Panfilio, K. A., Liu, P. Z., Akam, M. & Kaufman, T. C. Oncopeltus fasciatus zen is essential for serosal tissue function in katydidres. Dev. Biol. 292, 226–243 (2006).
18. Rafaq, A. M., Lemke, S., Ferguson, S., Staubner, M. & Schmidt-Ott, U. Evolutionary origin of the amnioserosa in cyclorrhaphan flies correlates with spatial and temporal expression changes of zen. Proc. Natl Acad. Sci. USA 105, 234–239 (2008).
19. Wakimoto, B. T., Turner, F. R. & Kaufman, T. C. Defects in embryogenesis in mutants associated with the Antennapedia Complex of Drosophila melanogaster. Dev. Biol. 102, 147–172 (1984).
20. McGregor, A. P. How to get ahead: the origin, evolution and function of bicoid. BioEssays: News Rev. Mol., Cell. Dev. Biol. 27, 904–913 (2005).
21. Pultz, M. A., Diederich, R. J., Cribbs, D. L. & Kaufman, T. C. The proboscipedia locus of the Antennapedia complex: a molecular genetic analysis. Genes Dev. 2, 901–920 (1988).
22. Rushlow, C. & Levine, M. Role of the zerknullt gene in dorsal-ventral pattern formation in Drosophila. In Advances in Genetics: Genetic Regulatory
Dearden, P., Grbic, M., Falciani, F. & Akam, M. Maternal expression and early development of Drosophila is a derived class 3 Hox gene. *Proc. Natl Acad. Sci. USA* 96, 3786–3789 (1999).

Falciani, F. et al. Class 3 Hox genes in insects and the origin of zen. *Proc. Natl Acad. Sci. USA* 93, 8479–8484 (1996).

Birkenbosch, A. et al. Sequence of the Tribolium castaneum homeotic complex: the region corresponding to the Drosophila melanogaster Antennapedia complex. *Genetics* 160, 1067–1074 (2002).

van der Zee, M., Berns, N. & Roth, S. Distinct functions of the Tribolium zen/kermit genes in serosa specification and dorsal closure. *Curr. Biol. 15*, 624–636 (2005).

Angeli, D. R. & Jockusch, E. L. Relationships among pest flour beetles of the genus Tribolium (Tenebrionidae) inferred from multiple molecular markers. *Mol. Phylogenet. Evol.* 46, 127–141 (2008).

Svobodova, E., Kubikova, J. & Svoboda, P. Production of small RNAs by mammalian Dicer. *Pflug. Arch. Eur. J. Physiol.* 468, 1089–1102 (2016).

Benton, M. A., Akam, M. & Pavlopoulos, A. Cell and tissue dynamics during Tribolium castaneum embryogenesis revealed by versatile fluorescence labeling approaches. *Development* 140, 3210–3220 (2013).

Handel, K., Grünfelder, C. G., Roth, S. & Sander, K. Tribolium embryogenesis: a SEM study of cell shapes and movements from blastoderm to serosal closure. *Dev. Genes Evol.* 216, 177–190 (2006).

Koerner, S., Köhne, K. & Panfilio, K. A. Visualizing late insect embryogenesis: extraembryonic and mesodermal enhancer trap expression in the beetle Tribolium castaneum. *PLoS ONE* 9, e103967 (2014).

Truckenbrodt, W. The embryonic covers during blastokinesis and dorsal closure of the normal and of the actomyosin D treated egg of *Odontotermes badius* (Hav.) (Insecta, Isoptera). *Zoologische Jahrbi. (Abt. Anat. Ontog. Tierw.)* 104, 77–178 (1979).

Cohen, E., Allen, S. R., Sawyer, J. K. & Fox, D. T. Fizzy-related dictates a cell cycle switch during organ repair and tissue growth responses in the Drosophila hindgut. *eLife* 7, e38327 (2018).

Schaeffer, V., Althause, C., Schererbata, H. R., Deng, W. M. & Ruohola-Baker, H. Notch-dependent Fuzzy-related/Hex/Cd16 expression is required for the mitotic-to-endoderm transition in Drosophila follicle cells. *Curr. Biol. 14*, 630–636 (2004).

Horn, T. & Panfilio, K. A. Novel functions for Dorsocross in epithelial morphogenesis in the beetle Tribolium castaneum. *Development* 143, 3002–3011 (2016).

Dearden, P., Gebic, M., Falciani, F. & Akam, M. Maternal expression and early zygotic regulation of the Hox/zex/zen gene in the grasshopper *Schistocerca gregaria*. *Evol. Dev.* 2, 261–270 (2000).

Banks, C. J., Joshi, A. & Michoel, T. Functional transcription factor target discovery via comepinda of binding and expression profiles. *Sci. Rep.* 6, 20649 (2016).

Shiba, N., Dorer, D. R., Moriyama, E. N. & Christensen, A. C. Evolution of a Drosophila Hox gene family in the apterygote insect Thermobia domestica. *Dev. Evol. 6*, 393–401 (2004).

Force, A. et al. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151, 1531–1545 (1999).

Gavin-Smyth, J., Wang, Y. C., Butler, I. & Ferguson, E. L. A genetic network conference catalyzing a bistable patterning system in Drosophila. *Curr. Biol.* 23, 2296–2302 (2013).

Hoffmann, A., Levchenko, A., Scott, M. L. & Baltimore, D. The IkapppaB-NF-kappaB signaling module: temporal control and selective gene activation. *Nature* 406, 335–341 (2000).

Nunes da Fonseca, R. et al. Self-regulatory circuits in dorsoventral axis formation of the short-germ beetle Tribolium castaneum. *Dev. Cell* 14, 605–615 (2008).

Schmidt-Ott, U., Rafiqi, A. M. & Lemke, S. Hox/zex/zen and the evolution of extraembryonic epithelia in insects. In *Hox genes: studies from the 20th to the 21st century*, Vol. 689. *Advances in Experimental Medicine and Biology* (ed. Deutsch, Jean S.) 133–144 (Landes Bioscience, 2010).

Rafiqi, A. M., Lemke, S. & Schmidt-Ott, U. Postgastrular zen expression is required to develop distinct amniotic and serosal epithelia in the scuttle fly *Megaselia*. *Dev. Biol.* 341, 282–290 (2010).

Reim, L., Lee, H.-H. & Frasch, M. The T-box encoding Dorsocross genes function in amnioserosa development and the patterning of the dorsolateral germ band downstream of Dpp. *Development* 130, 3187–3204 (2003).

Panfilio, K. A. & Roth, S. Epithelial reorganization events during late extraembryonic development in a hemimetabolous insect. *Dev. Biol.* 340, 100–115 (2010).

Orr-Weaver, T. L. When bigger is better: the role of polyplody in organogenesis. *Trends Genet.* 31, 307–315 (2015).

Hughes, C. L., Liu, P. Z. & Kaufman, T. C. Expression patterns of the rogue Hox genes Hox/zex and zushi tarazu in the apterygote insect *Thermobia domestica*. *Dev. Evol. 6*, 393–401 (2004).

R Core Team. *R: A Language and Environment for Statistical Computing*. (Foundation for Statistical Computing, Vienna) https://www.R-project.org/ (2016).
Acknowledgements

We thank Denise Mackrodt and Michael Schoppmeier for the kind gift of the Tc-Zen1 and Tc-Zen2 peptide antibodies, Viera Kovacova for bioinformatic program recommendations, Luigi Pontieri for assistance with statistical analyses, Thorsten Horn for sharing unpublished data on cuticle gene expression, Gustavo Lazzaro Rezende for discussions on cuticle regulation, and Hilary Ashe and Chris Rushlow for insights into zen regulation in Drosophila. RNA-seq analyses (trimming, filtering, OGS mapping) were performed on the Cologne High Efficient Operating Platform for Science (CHEOPS). We also thank Milos Tsiantis and Siegfried Roth for helpful discussions and recommendations throughout the course of this research project. Siegfried Roth, Peter Heger, and Matthias Pechmann provided helpful feedback on the paper. This work was supported by funding from the German Research Foundation (Deutsche Forschungsgemeinschaft) through SFB 680 project A12 and Emmy Noether Program grant PA 2044/1-1 to K.A.P.

Author contributions

D.G. designed experiments, collected and analyzed data, established the bioinformatic pipeline for the RNA-seq data, wrote the paper. I.M.V.J. analyzed data, established the bioinformatic pipeline for the RNA-seq data, edited the paper. K.A.P. conceived the project, designed experiments, analyzed data, established the bioinformatic pipeline for the RNA-seq data, wrote the paper.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s42003-020-01250-3.

Correspondence and requests for materials should be addressed to K.A.P.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2020