Evolution of gene regulatory network topology and dorsal-ventral axis specification in early development of sea urchins (Echinoidea)

Short Title: Evolution of gene regulatory network topology in sea urchins

Eric M. Erkenbrack\textsuperscript{1,*,#a}

\textsuperscript{1}Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California, United States of America

\#aCurrent address: Department of Ecology and Evolutionary Biology, Yale University, West Haven, Connecticut, United States of America

* Corresponding Author

Email: erkenbra@caltech.edu (EME), eric.erkenbrack@yale.edu (EME)
Abstract

Developmental gene regulatory networks (dGRNs) are assemblages of interacting regulatory factors that direct ontogeny of animal body plans. The hierarchical topology of these networks predicts that their nodes will evolve at different rates and consequently will bias the trajectories of embryonic evolution. To test this, detailed, comparative analyses of dGRNs that specify early, global embryonic domains are required. The most extensively detailed dGRNs have been documented for one of the two subclasses of extant sea urchins, the euechinoids. Remarkably, euechinoid dGRNs operating in early development show little appreciable change even though they diverged approximately 90 million years ago (mya). Therefore, to better understand the evolutionary dynamics of dGRNs, comparative microdissection must be undertaken for sea urchins that diverged deeper in geological time. Recent studies of cidaroids, the sister clade of euechinoid sea urchins, suggest that comparative analyses of their embryonic domain specification may prove insightful for understanding the evolutionary dynamics of dGRNs. Here, I report the spatiotemporal dynamics of 19 regulatory factors involved in dorsal-ventral patterning of non-skeletogenic mesodermal and ectodermal domains in the early development of *Eucidaris tribuloides*, a cidaroid sea urchin. Multiple lines of evidence indicate that deployment of ectodermal regulatory factors is more impervious to change than mesodermal regulatory factors in the sea urchin lineage and are supported by multiple lines of experimental evidence. Additionally, endogenous spatiotemporal expression data, intra-class reporter microinjections, and perturbation analyses of Nodal and Notch signaling allow the enumeration of numerous alterations to regulatory factor deployment since the divergence of echinoids. These
results provide a global view of early embryonic developmental processes in two clades that diverged at least 268.8 mya and show that the dGRNs controlling embryonic specification exhibit differential lability, supporting the hypothesis that the topologies of dGRNs bias rates of evolutionary change and alter the developmental evolutionary trajectories of embryogenesis.

**Author Summary**

Early in the development of an embryo, networks of genes are initiated to differentiate the rapidly dividing cells into distinct territories that will later serve specific functions. Sea urchins have revealed much about how this process unfolds. Recent studies have focused on one of the two modern lineages of sea urchins and have shown that these processes have not appreciably changed over the past 90 million years. I sought to determine if this trend extends over even larger evolutionary distances by investigating similar processes in a sea urchin from the second modern lineage, which is removed by 268 million years of evolution. By revealing where and when these genes are expressed and interfering with common mechanisms of development in a distantly related sea urchin, I show that changes to these networks of genes have occurred at all levels of the network. Additionally, I present data that suggests that changes to these networks of genes occur disproportionately in certain embryonic territories, which may be true for early development for other groups of organisms as well.

**Introduction**

From egg to embryo, early bilaterian development is the transformation of a single cell, the fertilized egg, into a dynamic gastrulating embryo with multiple cell types
and embryonic domains. Integral to early development of a triploblastic bilaterian is the
delineation of embryonic domains—endoderm, ectoderm, mesoderm—and their
subdomains—dorsal, ventral, anterior, posterior, mesenchymal, etc. This partitioning
sets the stage for specification of morphological features of the larva and/or adult.
Asymmetrically distributed RNA and proteins in the egg provide the initial inputs into this
process and thereby determine the spatial coordinates of domain formation [1, 2]. In the
context of these maternal factors, zygotic transcription is initiated, and the interplay
between the genomically encoded regulatory program and its output of regulatory
factors, e.g. transcription factors and cell signaling pathways, delineates embryonic
domains [3]. The deployment of evolutionarily conserved cohorts of transcription factors,
or regulatory states, is the spatial readout of developmental gene regulatory networks
dGRNs) and provides each embryonic domain with its molecularly distinct and
functional identity [4, 5].

The trajectories of change that can occur to developmental programs during
evolution are affected both by the sequential unfolding of embryonic development and
the hierarchical structure or topology of GRNs [6]. For example that certain nodes in
GRNs will evolve at different rates would seem to follow from their inherent hierarchical
architecture and would provide a powerful mechanistic explanation as to why constraint
occurs in some developmental processes and evolutionary change has occurred in
others [7]. However, despite the overt importance of the structure of developmental
GRNs to effect change in developmental evolution in predictable ways, illustrative
examples are scant in the literature. To address questions of the frequency and nature
of change to dGRNs, the taxa sampled must be phylogenetically diverged enough to
have undergone significant change to dGRNs and phylogenetically close enough so that similarity of developmental programs will afford meaningful comparisons. Due to the cascading nature of early specification events and the rapid establishment of embryonic domains, early development is attractive in so far that it promises to provide fundamental insight into both its lineage-specific evolution and hierarchical change in developmental GRNs.

Sea urchins (class Echinodermata) provide an excellent model system to study mechanisms of evolutionary change in early development. Specification of cell lineages and embryonic domains in sea urchin embryos depends on the canonical cleavage positions of their blastomeres [8, 9], thereby facilitating interpretation of mechanisms of spatial change. Also, a well-studied fossil record constrains the dating of evolutionary events [10] and has established that the sister subclasses of sea urchins—cidaroids and euechinoids—diverged from one another at least 268.8 million years ago (mya) [11]. And yet, relative to their conspicuously diverged adult body plans, early embryonic development in these two clades is strikingly similar [12]. This geologically ancient expanse combined with copious change of life history strategies in multiple sea urchin lineages provide a convenient framework, with experimental replicates, to investigate evolution and mechanisms of developmental programs [13]. For indirect developing sea urchins (taxa with feeding larval forms), morphological and developmental heterochronies exhibited by cidaroids and euechinoids have long been a topic of interest, but only recently have become the subject of molecular research [14-20]. Research on the early development of euechinoids has brought into high resolution the players and molecular logic directing the global embryonic developmental GRN that
encompasses the varied embryonic domains and subdomains of the purple sea urchin *Strongylocentrotus purpuratus* [21-32]. Additionally, abundant comparative evidence exists for other euechinoid taxa, including *Lytechinus variegatus* [33-39] and *Paracentrotus lividus* [40-44]. Remarkably, although these three indirect-development euechinoid sea urchins diverged from one another approximately 90 mya [10, 45], very little appreciable change to developmental GRNs has been observed in their early development [46-48]. Two questions arise from this observation: (1) how deep in geological time does this early developmental constraint extend, and (2) does this apparent calcification of GRN circuitry extend to specification of all embryonic domains or merely to some? Answers to these questions would obtain fundamental insight into the lability and evolutionary dynamics of GRN topology and whether certain embryonic domains or subdomains have a greater propensity to change in early development than others. Such an analysis might also reveal the precise locations of and frequency in changes to GRN architecture over evolutionary time and would yield a more thorough understanding of the interplay of constraint and evolvability of early developmental programs.

Recently, studies of the cidaroid sea urchin *Eucidaris tribuloides* revealed that mesoderm specification in this clade is markedly different from that observed in euechinoids [12, 20, 49]. Spatiotemporal and perturbation analyses of endomesodermal formation in *E. tribuloides* arrived at the conclusion that deployment of mesodermal regulatory factors has diverged more than deployment of endodermal regulatory factors since the cidaroid-euechinoid divergence. These studies provide insight into developmental process at the vegetal pole and bring within reach a global embryonic
perspective that would afford a glimpse into rates of change to whole apparatus of
developmental GRN throughout the early embryo. Here, I surveyed spatial and temporal
expression patterns of non-skeletogenic mesodermal (NSM) and ectodermal regulatory
factors in the cidaroid sea urchin *E. tribuloides* (Table 1). This study focused on dorsal-
ventral (D-V; also called Aboral-Oral) patterning, which has consequences for both
ectoderm and mesoderm. D-V axis specification is a well-documented process in the
euechinoid GRN [48] and is a highly conserved developmental mechanism in
deuterostomes [50, 51]. I present evidence that deployment of the primary regulatory
factors specifying the sea urchin mesoderm have diverged substantially in indirect-
developing echinoids. These alterations are overrepresented in specification of
mesodermal SM and NSM subdomains. However ectodermal and endodermal domains
and subdomains show a high degree of constraint relative to mesodermal domains.
Spatiotemporal dynamics of regulatory factors involved in *E. tribuloides* D-V axis
specification are essentially congruent with that of euechinoids, suggesting constraint
on deployment of these factors for sea urchin taxa with indirect-developing, feeding
larval life strategies. Thus, I argue that in early development of indirect-developing sea
urchins unequal rates of change exist at specific nodes of early developmental GRNs. I
enumerate specific examples of these changes at every level of GRN architecture. The
lability of developmental GRNs supports the notion that change can occur at all levels of
their hierarchy in early development and offers an in principle mechanistic explanation
for observations of rapid change to nearly all components of developmental process in
the development of direct developing, nonfeeding sea urchins [52-55]. These results
suggest that, while early development is dependent on and constrained by cascading,
sequential specification events, deployment of early developmental GRNs in bilaterian lineages may be biased towards alterations to specific embryonic domains or developmental programs.

Table 1. Regulatory factors examined in this study and their spatiotemporal expression in *Eucidaris tribuloides*.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Maternal/zygotic</th>
<th>Onset of zygotic activation</th>
<th>Embryonic domain/spatial expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>bra</td>
<td>zygotic</td>
<td>early blastula</td>
<td>broad in early endomesoderm, then endodermal; perianal ectoderm and ventral ectoderm by mid-gastrula</td>
</tr>
<tr>
<td>chordin</td>
<td>zygotic</td>
<td>hatching blastula</td>
<td>center of presumptive ventral ectoderm, then expanding slightly to most of presumptive ventral ectoderm</td>
</tr>
<tr>
<td>ese</td>
<td>zygotic</td>
<td>64-cell</td>
<td>broad in anterior/animal ectoderm early; then by early gastrula broadly in NSM and restricted to presumptive ANE; later restricted in archenteron by mesenchyme gastrula and in ANE</td>
</tr>
<tr>
<td>foxq2</td>
<td>maternal</td>
<td>16-cell</td>
<td>broadly in anterior/animal ectoderm early; subsequently restricted to ANE/lateral ectoderm by late blastula</td>
</tr>
<tr>
<td>gatac</td>
<td>zygotic</td>
<td>swimming blastula</td>
<td>in SM by late blastula and later in NSM as well; later asymmetrical in NSM by gastrula stage</td>
</tr>
<tr>
<td>gatae</td>
<td>zygotic</td>
<td>early blastula</td>
<td>first broadly in endomesoderm, then cleared from SM; later in endoderm and asymmetrical in NSM by gastrula stage</td>
</tr>
<tr>
<td>gcm</td>
<td>zygotic</td>
<td>early blastula</td>
<td>first in SM; then in NSM and cleared from SM; later asymmetrical in NSM by early gastrula; presumptive dorsal</td>
</tr>
<tr>
<td>gsc</td>
<td>zygotic</td>
<td>hatching blastula</td>
<td>early spatial not observed; presumptive ventral ectoderm from late blastula onwards</td>
</tr>
<tr>
<td>irxa</td>
<td>zygotic</td>
<td>late blastula</td>
<td>in dorsal ectoderm extending from border of ANE to blastopore; by gastrula stage excluded only from ventral ectoderm and pre-oral ANE</td>
</tr>
<tr>
<td>lefty</td>
<td>zygotic</td>
<td>64-cell</td>
<td>early blastula distribution not observed; presumptive ventral ectoderm by late blastula and onwards</td>
</tr>
<tr>
<td>msx</td>
<td>zygotic</td>
<td>late blastula</td>
<td>pregastrular distribution not observed; dorsal lateral ectoderm by early gastrula</td>
</tr>
<tr>
<td>nodal</td>
<td>zygotic</td>
<td>64-cell</td>
<td>early blastula distribution not observed; center of presumptive ventral ectoderm at SB, expanding slightly to most of ventral ectoderm by gastrula stage</td>
</tr>
<tr>
<td>not</td>
<td>zygotic</td>
<td>early blastula</td>
<td>early blastula distribution not observed; presumptive ventral ectoderm, then also in presumptive ventral mesoderm by mesenchyme gastrula</td>
</tr>
<tr>
<td>onecut</td>
<td>maternal</td>
<td>64-cell</td>
<td>early blastula pattern not observed, by early gastrula in post-oral ventral ectoderm and expanding anteriorly in a band encompassing ventral ectoderm</td>
</tr>
<tr>
<td>prox</td>
<td>zygotic</td>
<td>hatching blastula</td>
<td>SM early and subsequently in NSM; broadly in mesoderm by mid-gastrula</td>
</tr>
<tr>
<td>scl</td>
<td>zygotic</td>
<td>swimming blastula</td>
<td>SM early and subsequently in NSM; partially restricted in mesoderm by mesenchyme gastrula</td>
</tr>
<tr>
<td>tbx2/3</td>
<td>maternal</td>
<td>64-cell</td>
<td>early blastula distribution not observed; presumptive dorsal ectoderm by late blastula; later in dorsal lateral ectoderm as well as dorsal archenteron</td>
</tr>
</tbody>
</table>
Results

Dynamics of ectodermal D-V axis regulatory states in the cidaroid E. tribuloides

In euechinoids, numerous regulatory factors direct segregation of ectoderm into a diverse set of regulatory states [48, 56, 57]. Nodal, a member of the activin subfamily of the transforming growth factor-β (TGF-β) family of signaling molecules, is a critical factor in establishing dorsal-ventral (D-V) polarity in sea urchins [37, 40]. Nodal directly regulates, among others, nodal (itself), not, lefty and chordin [42, 58]. In E. tribuloides, zygotic transcripton of nodal, not and lefty begins by early blastula stage (Figure 1A, 1B, 1C). In contrast, transcriptional activation of chordin is delayed by at least 5 hours from this initial cohort, indicative of an intermediate regulator between nodal and chordin in E. tribuloides (Figure 1A, 1E). From 17 hpf to 40 hpf, spatial expression of nodal is observed in a well-defined region in the ventral ectoderm (VE) that expands slightly as gastrulation proceeds (Figure 1A1-1A4, Figure S1). Unlike nodal, the spatial distribution of its targets is not solely restricted to a small field of cells in the VE. Lefty (also known as Antivin), an antagonist of nodal, exhibits a broader pattern of expression that, by 50 hpf, expands into the ventral side of the archenteron (Figure 1B1-1B4, Figure S1). Similarly, chordin transcripts are detected in VE throughout early E. tribuloides development (Figure 1E1-1E4, Figure S1). The homeobox gene not, known to play a role directly downstream of nodal in euechinoid D-V ectodermal and mesodermal polarization [29, 59], was observed spatially in VE during gastrulation, and later extends
vegetally towards the perianal ectoderm and is observed in the archenteron (Figure 1C1-1C4, Figure S1). While I do not present the spatial distribution of the critical Nodal-responsive regulatory factor \textit{bmp2/4} here, qPCR timecourse data indicate that \textit{bmp2/4} is upregulated with the \textit{nodal-not-lefty} cohort (Figure S2). In euechinoids, the \textit{bmp2/4} ligand is a direct target of Nodal and is translocated across the embryo to the dorsal side, where it upregulates dorsal ectoderm (DE) specification genes such as \textit{tbx2/3} [40, 60, 61]. In \textit{E. tribuloides}, \textit{tbx2/3} is transcriptionally active very early with the \textit{nodal-not-lefty} cohort. \textit{Tbx2/3} exhibits spatial expression from late blastula stage onwards that is complementary to VE genes (Figure 1D1-1D4). By mid-gastrula stage, \textit{tbx2/3} is also expressed in the archenteron and much later, by 70 hpf, is expressed in the bilateral clusters of cells synthesizing the larval skeleton (Figure S1), which is similar to the spatial expression in two euechinoids with notably interesting heterochronic differences [62, 63]. Lastly, the Forkhead family transcription factor \textit{foxq2} is sequentially restricted to and specifically expressed in embryonic anterior neural ectoderm (ANE) territory in deuterostomes [64]. In euechinoids, \textit{foxq2} restriction to ANE is a crucial component of D-V axis specification, setting the anterior boundary of VE by restricting expression of \textit{nodal} [25, 57]. In \textit{E. tribuloides}, \textit{foxq2} exhibited an expression pattern consistent with observations in euechinoids and other deuterostomes, suggesting conserved roles for this gene in ANE and D-V specification (Figure 1F, Figure S1).

\textbf{Fig 1.} Spatiotemporal dynamics of six regulatory factors in \textit{E. tribuloides} suggest conserved deployment of ectodermal dorsal-ventral embryonic domains in echinoids. Visualization of mRNA transcripts revealed by whole mount in situ
hybridization, and estimates of absolute mRNA transcript abundance determined by qPCR during first 35 hours post fertilization (hpf). Individual data points are light grey. Blue data points represent the mean at that particular timepoint. (A1) Temporal dynamics of nodal. (A2-A5) Nodal spatial distribution is restricted to a small field of cells in the ventral ectoderm (VE) up to early-mid gastrula stage. (B1) Temporal dynamics of lefty. (B2-B5) Lefty spatial distribution is restricted to a small field of cells in VE. (C1) Temporal dynamics of not. (C2-C5) Not spatial distribution is first detected in a similar field of cells as nodal and lefty; however, the domain of not subsequently expands by 28 hpf where it is seen in the ventral side of the archenteron, where non-skeletogenic mesoderm (NSM) and endoderm are being segregated. By 40 hpf, the spatial domain of not extends from anterior neural ectoderm to the perianal ectoderm, is clearly seen in NSM, and was not detected in endodermal lineages. (D1) Temporal dynamics of tbx2/3. (D2-D5) Spatial distribution of tbx2/3 at 17 hpf is detected broadly in dorsal ectoderm and later extends from the perianal ectoderm to lateral AE, but not past the embryonic equator. (E1) Temporal dynamics of chordin. (E2-E5) Spatial distribution of chordin is first observed in a few cells in VE at 17 hpf and subsequently expands to extend from the perianal ectoderm to ANE. (F1) Temporal dynamics of foxq2. (F2-F5) Foxq2 spatial distribution is detected very early in development and is spatially restricted to anterior ectoderm by 17 hpf.

Fig S1. Spatial expression of regulatory factors involved in dorsal-ventral (D-V) axis formation in E. tribuloides, including the dorsal ectoderm transcription factor msx. Additional whole mount in situ hybridization images of selected timepoints for (A1-
A6) bra, (B1-B8) chordin, (C1-C8) foxq2, (D1-D6) lefty, (E1-E4) msx, (F1-F3) nodal,

(G1) not, and (H1-H4) tbx2/3.

Fig S2. Temporal expression dynamics of four ectodermal regulatory factors in E. tribuloides. qPCR timecourses of (A) bmp2/4, (B) emx, (C) hesC and (D) msx.

Dynamics of ciliated band regulatory states in the cidaroid E. tribuloides

Free-feeding, indirect-developing sea urchins possess a single neurogenic ciliated band (CB) early in development that circumnavigates the larval ventral face and facilitates feeding and locomotion [65]. This structure has undergone frequent modification in the lineages leading to modern sea urchins, viz. in planktotrophic larvae [66]. In euechinoids goosecoid (gsc), onecut and irxa contribute to the geometric patterning of CB formation [30, 42, 67]. In euechinoids, gsc is expressed in VE and is directly downstream of nodal signaling on the ventral side of the embryo [42]. onecut (also known as hnf6) is a ubiquitous, maternally deposited factor that is restricted to the boundary of VE and DE, at which lies progenitor CB territory; and irxa is expressed exclusively in DE downstream of tbx2/3 [42, 68]. In the cidaroid E. tribuloides, gsc is zygotically expressed with the nodal-not-lefty cohort by 12 hpf and is specifically expressed in VE (Figure 2A, Figure S3). Onecut is also a maternally deposited factor in E. tribuloides; early onecut spatial expression was difficult to interpret, as staining was only observed much later in development in a restricted band of cells encircling the VE. The spatial dynamics of onecut in E. tribuloides is quite remarkable, however, insofar
that whole-mount *in situ* hybridization (WMISH) timecourse revealed that its activation unfolds slowly and in a sequential manner that begins in the progenitor field of post oral CB and subsequently extends in a narrow band of 4-8 cell diameters towards progenitor pre oral CB (Figure 2B, Figure S3). This observation is in stark contrast to that in euechinoids, in which *onecut* is observed to be ubiquitously expressed early and later delimited to the CB territory by transcriptional repressors in the VE and DE [30, 69]. *Irxa* initiates zygotic expression at mid-blastula stage (~14 hpf) in *E. tribuloides*, and by 28 hpf is observed broadly in DE (Fig 2C). Unlike in euechinoids, *irxa* is broadly distributed in DE—much more so than *tbx2/3*—indicating that it is likely broadly activated in the ectoderm and repressed in VE and ANE. The spatial distributions of *gsc*, *onecut* and *irxa* are highly suggestive of a conserved regulatory apparatus that spatially restricts CB to the boundary of VE and DE. To test for this conservation, I assayed a series of endogenous and site-directed mutagenesis *onecut* BACs from *S. purpuratus* by microinjection [67]. Remarkably, a BAC that has been shown to recapitulate the endogenous *S. purpuratus onecut* expression pattern faithfully expressed reporter GFP in the CB of *E. tribuloides* (Figure S4). Further, a BAC harboring mutated repressor sites for the ventral repressor *gsc* repeatedly exhibited ectopic expression in VE of *E. tribuloides* (Figure S4). Taken together, the early specification of CB regulatory factors suggests divergence of initial activation and spatial distributions of *onecut* and *irxa* and is consistent with conserved circuitry of *gsc*. Later, *E. tribuloides* CB patterning exhibits congruence with spatial expression patterns and circuitry observed in euechinoids, suggesting stage-specific constraint during larval morphogenesis.
Fig 2. Spatiotemporal dynamics of three regulatory factors in *E. tribuloides* suggest conserved spatial distribution of ciliary band embryonic domain in euechinoids. Visualization of mRNA transcripts revealed by whole mount in situ hybridization, and estimates of absolute mRNA transcript abundance determined by qPCR during first 35 hours post fertilization (hpf). Individual data points are light grey. Blue data points represent the mean at that particular timepoint. (A1) Temporal dynamics of *goosecoid* (*gsc*). (A2,A3) Spatial distribution of *gsc* at 22 hpf to 32 hpf is observed exclusively in ventral ectoderm (VE). (B1) Temporal dynamics of *onecut*. (B2,B3) By 40 hpf *onecut* is detected in the future post oral ciliary band and is initiated in a band moving from the posterior to the anterior. (C1) Temporal dynamics of *irxa*. (C2,C3) At 28 hpf, *irxa* is detected in dorsal ectoderm (DE) and extends from the vegetal endodermal domains to anterior neural ectoderm. By 40 hpf *irxa* is seen extending anteriorly at the boundary of DE and VE.

Fig S3. Spatial expression in *E. tribuloides* of regulatory factors involved in euechinoid ciliary band restriction. Additional whole mount in situ hybridization images of selected timepoints for (A1-A9) *gsc*, (B1-B9) *onecut*, and (C1-C5) *irxa*.

Fig S4. Expression dynamics in *E. tribuloides* of GFP reporter BACs harboring the regulatory locus of *S. purpuratus* *onecut* suggest conservation of geometric positioning circuitry of ciliary band in echinoids. (A) Table showing reporter analysis of spatial expression of five different *S. purpuratus* GFP BACs in *E. tribuloides*. The BACs were previously utilized to analyze the cis-regulatory dynamics of *onecut*. 
spatial distribution in *S. purpuratus* (Barsi and Davidson, 2016). (B) BAC reporter 1, which harbors mutations to all known cis-regulatory modules (CRMs), showed an absolute reduction of all reporter expression. (C, D) BAC reporters 2 and 3, which harbor mutated enhancer CRMs, showed reduced reporter expression both in terms of percent embryos exhibiting reporter activity and in the number of cells with reporter. (E) BAC Reporter 4, which harbors mutations to the dorsal and ventral ectodermal repression CRMs, increases the percent of embryos that show reporter expression in the ventral ectoderm, as well as normal reporter expression in the ciliary band. (F) BAC Reporter 5, which harbors the unperturbed, wild-type locus of *S. purpuratus onecut*, faithfully exhibits reporter GFP in the ciliary band domain of *E. tribuloides*.

**Dynamics of non-skeletogenic mesoderm regulatory states**

in the cidaroid *E. tribuloides*

Non-skeletogenic mesoderm (NSM) in euechinoids arises at the vegetal plate from early cleavage endomesodermal precursors and gives rise to four different cell types: blastocoelar cells, pigment cells, circumesophageal cells and coelomic pouch cells [70]. Experimental observations indicate that euechinoids completely rely on presentation of Delta ligand in the adjacent SM to upregulate NSM regulatory factors in veg2 endomesodermal cells [34, 71, 72]. As gastrulation begins, euechinoid NSM has already become segregated into dorsal NSM and ventral NSM in response to Nodal signaling from VE [29, 73]. Mesodermal patterning in *E. tribuloides* also depends on Notch signaling, though by restricting SM fate to the micromere-descendants and, strikingly, not affecting the early expression of *gcm*, a regulatory factor involved in NSM
segregation and pigment cell specification [20]. In *E. tribuloides*, *ese* and *gcm* are early euechinoid NSM regulatory factors that are zygotically activated at late cleavage/early blastula stage (Figure 3A, 3D). In contrast to *S. purpuratus* spatial distribution, *ese* in *E. tribuloides* is observed both in the ANE and the NSM simultaneously (Figure 3A1-3A4, Figure S5) [74]. Indeed, very early in development *ese* is exclusively in animal blastomeres and later becomes zygotically expressed in NSM progenitors at the vegetal pole (Figure S5). In NSM, *ese* expression first occurs broadly just prior to the onset of gastrulation and is subsequently restricted to one side of the archenteron (Figure 3A1-3A4). After gastrulation begins, *gcm* is expressed transiently in ventral and dorsal NSM, and by 28 hpf is restricted to a cluster of cells just below the tip of the archenteron (Figure 3D1-3D4). Later this expression is seen solely on one side of the archenteron as *gcm*-positive cells ingress rapidly into the blastocoel at 36 hpf (Figure S5). In contrast to its spatial expression in euechinoids and similar to its expression in asteroids [75, 76], *gcm* in *E. tribuloides* is upregulated in the ectoderm at late blastula/early gastrula stage (Figure S5). While I cannot definitively preclude the possibility that these *gcm*-positive cells are mesodermal in origin, all observations of and experimental data on *E. tribuloides* supports the notion that SM is the first mesodermal lineage to ingress at 28 hpf. The data presented here are at least 6 hours prior to this initial ingression event and are highly supportive of the hypothesis that *gcm* is activated in the ectoderm at the onset of gastrulation. Directly downstream of *gcm* in euechinoids is *gatae* [29]. In *S. purpuratus*, *gatae* is observed in endomesoderm early in development [77]. In *E. tribuloides* NSM, *gatae* is expressed throughout the endomesoderm at the time of SM ingestion (~28 hpf) and later is observed restricted to one side near the tip of the
archenteron, as well as in the second wave of ingressing mesenchyme (Figure 3C1-3C4). *Gatac* (*gata1/2/3*), *prox* and *scl*, all of which are ventral NSM genes in euechinoids [29], come off the baseline at similar times in *E. tribuloides* and are detectable in a few cells at the base of the vegetal pole by 18 hpf by WMISH (Figure S5). Of these three genes, *scl* was the first to show D-V NSM polarity followed by *gatac* (Figure 3B1-3B4, 3F1-3F4). Surprisingly, by 36 hpf *prox* did not exhibit an expression pattern that clearly indicated D-V polarity (Figure 3E1-3E4; Figure S5), suggesting that either *prox* is a general mesodermal regulatory factor in *E. tribuloides* or it is spatially restricted later in its development.

**Fig 3. Spatiotemporal dynamics of six regulatory factors in *E. tribuloides* suggest divergent deployment and specification of non-skeletogenic mesoderm domains in echinoids.** Visualization of mRNA transcripts revealed by whole mount in situ hybridization, and estimates of absolute mRNA transcript abundance determined by qPCR during first 35 hours post fertilization (hpf). Individual data points are light grey. Blue data points represent the mean at that particular timepoint. (A1) Temporal dynamics of *ese*. (A2-A5) *Ese* is detected primarily in non-skeletogenic mesoderm (NSM) but also in anterior neural ectoderm. By 28 hpf, *ese* is observed at the tip of the archenteron and is asymmetrically polarized in NSM. (B1) Temporal dynamics of *gatac*. (B2-B5) At 28 hpf *gatac* is expressed throughout the mesoderm and does not show polarity. By 36 hpf, *gatac* is detected in ingressing cells and is polarized at the tip of the archenteron. (C1) Temporal dynamics of *gatae*. (C2-C5) *Gatae* is detected throughout the endomesoderm at 28 hpf. Later at 36 hpf *gatae* is cleared from progenitor foregut
endodermal domains and is expressed at the blastopore, in ingressing mesenchymal cells and at the tip of the archenteron, where it is polarized. (D1) Temporal dynamics of gcm. (D2-D5) As gastrulation begins, gcm is expressed broadly in NSM; however by 28 hpf it exhibits stark polarity in a field of cells that resides basal to the tip of the archenteron. Gcm is also observed in a few ectodermal cells at the time of primary mesenchymal ingression. (E1) Temporal dynamics of prox. (E2-E5) Prox spatial distribution is observed throughout NSM at 28 hpf. By 36 hpf it is expressed in ingressing mesenchymal cells and polarity is not yet observed. (F1) Temporal dynamics of scl. (F2-F5) Spatial distribution of scl is observed throughout NSM at 22 hpf. At 28 hpf it is expressed in ingressing mesenchyme and throughout NSM, where it is polarized. Orange asterisks denote position of archenteron.

Fig S5. Spatial expression of euechinoid non-skeletogenic mesodermal regulatory factors in *E. tribuloides*, including pregastrular timepoints. Additional whole mount in situ hybridization (WMISH) images of selected timepoints for (A1-A12) ese, (B1-B6) gatac, (C1-C4) gatae, (D1-D6) gcm, (E1-E6) prox, (F1-F4) scl. Additionally, double WMISH is reported for (G1,G2) gcm and ese, (H1,H2) gcm and alx1, and (I1,I2) ets1 and tbrain.

These data on spatial dynamics of NSM regulatory factors suggest that there exist numerous regulatory states in the anterior archenteron. To provide some clarity, double fluorescent WMISH (dfWMISH) indicated this was indeed the case. Previous observations suggested that *E. tribuloides* mesodermal domains broadly express ets1/2
and tbrain [20], and dfWMISH confirmed this result (Figure S5I1, S5I2). Within this broad ets1/2-tbrain domain, three regulatory states are identified (Figure S5G-S5I): (1) ventrally localized ets1/2, tbrain and ese; (2) dorsally localized ets1/2, tbrain, and gcm; and (3) an anteriorly localized micromere-descendant regulatory state at the tip of the archenteron of ets1/2, tbrain, ese and alx1. These early mesodermal partitions are superficially consistent with those seen in euechinoids, as ese is restricted to ventral NSM and gcm to dorsal NSM [29]. However, the regulatory states expressed in these mesodermal domains are very different, and in E. tribuloides it is clear from the preceding data that the archenteron harbors multiple NSM regulatory states, the sequential of which is markedly different from that in euechinoids.

Effects of perturbation of D-V axis specification on ectodermal regulatory factors in E. tribuloides

The spatiotemporal data presented thus far are highly suggestive that D-V axis specification, as well as gastrular CB formation, in E. tribuloides is consistent with similar processes in euechinoids and that NSM specification has ostensibly diverged. To establish differences in the topology of these developmental GRNs, perturbation experiments disrupting initial inputs into D-V axis specification were conducted. In euechinoids, the primary molecular event responsible for animal-vegetal (A-V) axis polarity is nuclearization of β-catenin in micromere nuclei at the vegetal pole, and, unexpectedly, these experiments showed that perturbation of A-V axis formation disrupted D-V axis specification [33, 35]. One mechanism underlying the crosstalk of these two deuterostome specification events was found to be restriction of foxq2 to ANE, as its presence in VE blocked nodal transcription [25]. To test for this GRN
linkage, I overexpressed dn-Cadherin RNA in *E. tribuloides* to block nuclearization of β-catenin at the vegetal pole. As in euechinoids, this perturbation led to upregulation of *foxq2*, whereas *nodal* and its euechinoid downstream components of D-V axis GRN circuitry—e.g. *bmp2/4, not, and tbx2/3*—were strongly downregulated (Figure 4A). This result suggests that the molecular crosstalk between and GRN topology of β-catenin/TCF, *foxq2* and *nodal* are conserved between euechinoid and cidaroid echinoids.

*Fig 4. Perturbation of dorsal-ventral axis formation in *E. tribuloides* reveals conserved and divergent aspects of regulatory factor deployment in echinoids.*

Disruption of dorsal-ventral (D-V) specification was achieved by overexpression of cadherin mRNA (MOE) and culturing embryos in the presence of the alk4/5/7 small molecule inhibitor SB431542. (A1) Caderhin MOE affects Nodal and its downstream targets. Change in Ct (ddCt) values relative to internal control is listed on the y-axis. When cadherin is overexpressed in *E. tribuloides, foxq2* is not substantially cleared and *nodal* and its targets are strongly downregulated. (B1-B3) Effect of SB431542 on embryonic morphology of *E. tribuloides*. At 120 hpf, *Eucidaris* shows two triradiate skeletal rods extending anteriorly and ventrally. When cultured in the presence of 15 μM SB431542, *E. tribuloides* embryos show dorsal radialization and exhibit serial loci of spiculogenesis (black arrows). (C1) Quantitative effect of SB431542 on expression of *E. tribuloides* regulatory factors as revealed by qPCR. Change in Ct (ddCt) is show on the y-axis. Two timepoints from two independent replicates are shown. Regulatory factors are listed on the x-axis and font color designates their embryonic domain: black,
anterior neural ectoderm; blue, dorsal ectoderm; green, ventral ectoderm; yellow, endoderm; red, mesoderm.

Next, I aimed to determine the spatiotemporal effects of perturbation of D-V specification by culturing *E. tribuloides* embryos in the presence of SB43152, a small molecule antagonist of the TGF-β (Nodal) receptor Alk4/5/7 [78]. At four days post fertilization, these embryos exhibited strong dorsalization, archenterons that failed to make contact with VE, and supernumerary skeletal elements (Figure 4B). Quantitative PCR (qPCR) analysis at four different timepoints in *E. tribuloides* development showed strong downregulation of VE regulatory factors *chordin, gsc, lefty, nodal* and *not* (Figure 4C). This result was confirmed spatially by WMISH for mRNA transcripts of *chordin, nodal* and *not* (Figure 5A, 5B). Another critical VE regulatory factor is the secreted TGF-β ligand *bmp2/4*. This gene was clearly not affected to the same degree as the aforementioned cohort of VE factors (Figure 4C). This result is strikingly different from the strong downregulation of *bmp2/4* observed in the euechinoid *P. lividus* when it was cultured in the presence of SB431542 or when injected with Nodal morpholino (MASO) [40, 42]. Lastly for VE, this quantitative assay does not indicate disturbance in the regulation of *brachyury (bra)* and *foxa*, two euechinoid stomodeum (larval mouth) regulatory factors strongly downregulated upon Nodal perturbation in euechinoids (Figure 4C) [42, 79]. However, there is a clear heterochrony in the onset of *bra* and *foxa* in VE of *E. tribuloides* as stomodeum-specific genes such as these are not activated in VE until, at least for *brachyury*, around 36 hpf in *E. tribuloides* development (Figure S1).
Notably, foxa expression was never observed in the *E. tribuloides* stomodeum up to 40 hpf.

---

**Fig 5. Spatial effect of perturbation of dorsal-ventral axis formation on expression of selected ectodermal, mesodermal and ciliary band regulatory factors.** (A) At 28 hpf expression of *chordin*, *nodal* and *not* are completely extinguished. Whereas *gcm* is regularly restricted to one side of the archenteron, in the presence of SB431542 it exhibits expression throughout the archenteron. In the ectoderm, expression *tbx2/3* expands from dorsal ectoderm (DE) into ventral ectoderm in the presence of the inhibitor. (B) At 40 hpf, *chordin*, *nodal* and *not* are not detected. *Gcm* fails to be restricted to one side of the archenteron. *Tbx2/3*, which is normally expressed in DE and the dorsal side of the archenteron, is now expressed in a concentric band nearer the blastopore than the equator. (C) The ciliary band marker *onecut* is normally observed in a band of cells between the boundaries of DE and VE. However, in the presence of SB431542, *onecut* is expressed in an equatorial band that is 6-10 cell diameters across.

On the dorsal side, a striking difference is the effect of this treatment on regulatory factor *tbx2/3*. In euechinoids, *tbx2/3* is downstream of *bmp2/4* ligand, which diffuses from VE to DE [60, 61, 80]. Treatment of *P. lividus* embryos with SB431542 inhibitor completely and specifically extinguishes *tbx2/3* in DE while not interfering with its SM expression [42]. In *E. tribuloides*, qPCR data suggest SB431542 inhibitor has no effect on *tbx2/3* regulation (Figure 4C). However, when I assayed *tbx2/3* by WMISH, its
spatial distribution expanded into VE (Figure 5A, 5B). Similarly, whereas in *E.
tribuloides* qPCR data indicate strong downregulation of the DE regulatory factor *irxa*
(Figure 4C), its domain of expression expanded into VE in *P. lividus* embryos upon
SB431542-treatment [42]. These results suggest distinct GRN topologies exist
immediately downstream of the initial *nodal* and *bmp2/4* circuitry in echinoids.

The preceding results detailing the effect of SB431542 on specification of VE and
DE in *E. tribuloides* suggest that cidaroids and euechinoids share multiple
transcriptional targets directly downstream of Nodal in VE. However, the notable
exception in the euechinoid VE cohort is *bmp2/4*, the spatial expression of which has
not been detailed in *E. tribuloides* and was not detailed in this study. In DE it would
appear that multiple euechinoid GRN linkages are different in *E. tribuloides*, including
the spatial regulation of *tbx2/3* and *irxa*. Taken together these results suggest that the
initial specification of regulatory factors immediately downstream of Nodal in euechinoid
VE exhibit similar deployment than regulatory interactions that are immediately
downstream of the ventral to dorsal signal.

**Effects of perturbation of D-V axis specification on**

**mesodermal regulatory factors in *E. tribuloides***

In euechinoids studied thus far, polarity in NSM (D-V) lineages is also regulated
by regulatory factors downstream of Nodal signaling [29, 73]. In *E. tribuloides*, qPCR
data did not indicate consistent differences in mRNA abundance for NSM regulatory
factors (Figure 4C). However, WMISH assays revealed that embryos treated with
SB431542 failed to restrict *gcm* to the dorsal side (Figure 5A, 5B). This observation is
consistent with the euechinoid GRN linkage of the Nodal-responsive *not* repressing
dorsal NSM in the ventral-facing region of the archenteron [29]. Indeed, in *E. tribuloides*, not can be seen observed in the archenteron throughout gastrulation (Figure 1C2-1C4). However, upon disruption of the Nodal signal, not expression is extinguished and *gcm* is not properly restricted (Figure 5A, 5B). These observations are consistent with a conserved role for Nodal signaling in NSM segregation in the archenteron of *E. tribuloides*.

Lastly, CB formation in euechinoids is dependent on repression of *gsc* in VE and *irxa* in DE [42, 67]. While little is known about CB formation in *E. tribuloides*, recent work indicated that Onecut is expressed in CB and that disruption of endomesoderm formation by treatment with zinc resulted in embryos exhibiting a ring of highly concentrated proneural synaptotagmin-B positive cells at the equator of the embryo [81]. This result is remarkably similar to that shown in Figure 5C, which shows *onecut* mRNA transcripts detected by WMISH in an equatorial band in *E. tribuloides* embryos cultured with SB431542. Thus, by blocking D-V axis specification in *E. tribuloides*, embryos produce a single proneural CB encircling the embryo at the equator. However, this perturbation is drastically different in euechinoids, where treatment with SB431542 or injection of Nodal MASO markedly increases *onecut* expression throughout the ectoderm [80]. However, this is not the case in cidaroids, as perturbation data presented here and elsewhere [81] suggest that, in the absence of proper D-V patterning, a proneural CB appears only at the equator in the cidaroid sister-clade. These conflicting results are consistent with the hypothesis that there are anteriorly positioned regulatory factors repressing *onecut* in cidaroids.
Comparative analysis of global developmental GRN dynamics in early echinoid embryos

Next I undertook a statistical comparative analysis between *E. tribuloides* and two euechinoids that would inform hypotheses on correlation of transcriptional activity of GRN regulatory factors and global developmental GRN topology. While there are multiple datasets published with timecourse data of transcript abundance in *S. purpuratus* [82, 83], until recently there were no large datasets for other euechinoids. However, a high density timecourse dataset of temporal expression dynamics and initiation times was recently published for early regulatory factors operating in *P. lividus*, and their inclusion with *S. purpuratus* data provided the foundation for a comparative analysis between three species [84]. To conduct this analysis, distinct ontogenetic rates between the species were corrected for by comparing the timing of major developmental events, e.g. gastrulation, between the species, and relative transcript abundance in each species combined with Spearman’s rank correlation coefficients ($\rho$) were compiled for orthologues. Previous analyses had already posited the absence of a double-negative gate in cidaroids [19, 20], an observation that even without additional data supports the notion of large scale rewiring at the top of the SM GRN hierarchy. To determine if altered deployment of early GRN topologies is the rule and not the exception for early patterning of embryonic territories in echinoids, an analysis of regulatory factors in *E. tribuloides, P. lividus* and *S. purpuratus* was conducted. Plotting relative mRNA transcriptional dynamics for the three species were indicative of compelling correlation for ectodermal and endodermal regulatory factors and supported
the notion of poor correlation for regulatory factors driving mesoderm specification (Figure 6).

**Fig 6. Comparative gene expression analysis suggests conserved deployment of ectodermal and endodermal regulatory factors and divergent deployment of mesodermal regulatory factors in echinoids.** *Strongylocentrotus purpuratus*, purple line; *Paracentrotus lividus*, green dashed line; *Eucidaris tribuloides*, black dashed line.

Transcripts per embryo for each gene were normalized to their maximal expression over the first 30 hours of development and are plotted against *E. tribuloides* development on the x-ordinate. Comparative developmental staging for each species is listed in Table S4. Each analysis is accompanied by a matrix of Spearman correlation coefficients (marked as greek rho).

To provide further support for the hypothesis of domain-specific change to GRN topology, a two-species comparison between *E. tribuloides* and *S. purpuratus* was conducted to analyze an increased sample size of 34 regulatory genes, the spatial distributions of which are all known in *S. purpuratus* and *E. tribuloides*. Spearman’s rank correlation coefficients (ρ) were calculated pairwise for each orthologue. Values for ρ were then binned by their embryonic domain of expression in *S. purpuratus*. Comparison of the domain-specific ρ of regulatory factors expressed in each of the three canonical bilaterian embryonic domains (germ layers) against the mean of all ρ values, regulatory factors expressed in both *S. purpuratus* and *E. tribuloides* endoderm and ectoderm exhibited significantly higher ρ relative to the mean of all ρ values,
suggesting strong conservation of transcriptional dynamics of these factors in echinoids (Figure 7A). However, regulatory factors expressed in *S. purpuratus* and *E. tribuloides* mesodermal germ layers did not depart significantly from the mean ρ, suggesting transcriptional dynamics of mesodermal regulatory factors have changed markedly since the cidaroid-euechinoid divergence (Figure 7A). To determine whether mesodermal subdomains had undergone changes to GRN deployment, regulatory factors were further binned into embryonic subdomains. This finer-scale analysis revealed that whereas both SM and NSM regulatory factors showed significant variation in their transcriptional dynamics relative to the mean of all ρ values, the SM showed significantly more variation than the NSM (Figure 7B). In contrast to this, deployment of subdomains of ectodermal and endodermal regulatory factors exhibit statistically significant departures from the mean of all ρ values (Figure 7B).

**Fig 7. Distribution plots of Spearman’s rank correlation coefficients (ρ) for temporal dynamics in *E. tribuloides* and *S. purpuratus* reveals domain-specific alterations to deployment of regulatory factors.** Genes were binned by embryonic domain of *S. purpuratus* expression. Boxplot boundaries show interquartile range, means and standard deviation. Asterisks mark statistical significance as determined by a two-tailed t-test. (A) Boxplots for statistical distribution of endodermal, ectodermal and mesodermal regulatory factors in *E. tribuloides* and *S. purpuratus*. Mean ρ for endodermal and ectodermal regulatory factors were significantly higher than the mean ρ. Mesodermal regulatory factors did not significantly vary from the mean. (B) Boxplots for statistical distribution of subdomains of endodermal, ectodermal and mesodermal
regulatory factors in *E. tribuloides* and *S. purpuratus*. Whereas veg2 endoderm and

dorsal and ventral ectodermal domains showed statistically significant differences from

the mean, both skeletogenic and non-skeletogenic regulatory factors did not differ

significantly from the mean $p$.

**Discussion**

**Divergence of embryonic domain specification in early development of echinoids**

Since the divergence of cidaroids and euechinoids at least 268.8 mya, echinoid
developmental GRNs have significantly diverged, as shown above by the large-scale
survey of regulatory factors establishing D-V polarity in mesoderm and ectoderm of *E. tribuloides*. Importantly, these networks are not so dissimilar as to be unrecognizable.

Indeed, at all levels of GRN deployment there exist commonalities. By contrasting these
observations with those in other echinoderms, we can begin to appreciate the degree to
which embryonic developmental GRNs are constrained or malleable over vast
evolutionary distances and can reconstruct the ancestral regulatory states that must
have existed in the embryos of echinoderm ancestors [49].

**Regulatory states and polarity of NSM in *E. tribuloides***

The most conspicuous morphological differences during embryogenesis of
cidaroids and euechinoids are the asymmetric cleavage of micromeres and the
heterochrony of primary mesenchymal ingression. Euechinoids exhibit asymmetric
cleavage of vegetal blastomeres at 4th and 5th cleavage to yield large micromeres and
small micromeres. Large micromeres present the Delta ligand to immediately adjacent cell layers, which give rise to mesodermal NSM anteriorly and the small micromere quartet (SMQ) posteriorly. In mesodermal NSM, \textit{gcm} is directly downstream of Notch signaling and is restricted to dorsal NSM by the time that SM ingresses into the blastocoel prior to gastrulation [39, 73, 75, 85]. In cidaroids, mesodermal polarity of \textit{gcm} occurs 4-6 hours prior to SM ingress and does not occur until after gastrulation has begun. Thus, if \textit{gcm} is near the top of the NSM GRN in cidaroids [19] as is the case in euechinoids [85], then this pregastrular NSM polarization can be viewed as a euechinoid synapomorphy. This hypothesis is supported by the observation that no significant polarity occurs in mesodermal specification in holothuroids [86]. Two observations make it likely that euechinoid regulatory linkages mediating \textit{gcm} polarization via the transcription factor \textit{not} [29] are likely to exist in \textit{E. tribuloides} as well: (1) \textit{not} expression is observed at on the ventral side of the archenteron by early gastrula stage when \textit{gcm} is spatially restricted (Figure 1C1, Figure 3D1-3D4) and (2) dorsal localization of \textit{gcm} does not obtain when D-V axis patterning is perturbed (Figure 5A). Together these observations suggest a conserved role for VE regulatory factors in patterning the NSM of echinoids and that, in the lineage leading to modern euechinoids, deployment of GRN circuitry polarizing NSM underwent a heterochronic shift in the lineage leading to euechinoids.

Intriguingly, \textit{gcm} expression is observed in the ectoderm prior to SM ingress. One hypothesis that would explain this observation in \textit{E. tribuloides} is that NSM ingresses prior to SM ingress. However, as \textit{gcm}-expressing cells have never been observed in the blastocoel prior to 30 hpf, these cells would not express \textit{gcm} until they...
intercalate into ectoderm. This scenario is very unlikely, though, given that numerous independent observations show that the primary mesenchymal ingression event in *E. tribuloides* is executed by SM and occurs only after the archenteron has extended considerably into the blastocoel [14-17, 20, 87]. A competing hypothesis is that ectodermal *gcm* expression in *E. tribuloides* is evolutionarily related to *gcm* expression seen in late blastula stages of asteroids [76]. Indeed, follow up experiments indicated that perturbation of Notch signaling increased the spatial domain of ectodermal *gcm* and resulted in supernumerary pigment cell formation (Figure S6). These observations support the hypothesis that *gcm* in *E. tribuloides* has roles both in mesodermal NSM and ectoderm. If this is the case, two more things are clear evolutionarily: (1) *gcm* was likely expressed in the ectoderm in the echinozoan ancestor at least 481 mya; and (2) the lineage leading to camaradont euechinoids lost ectodermally-derived *gcm* activity, which may have been a consequence of the endomesodermal Notch-dependent *gcm* linkage now observed. Further investigation disentangling the roles of *gcm* in cidaroids will provide insight into how the regulation and function of this gene has evolved in echinoderms.

Fig S6. Ectodermal *gcm* expression and pigment cell abundance are altered in a Delta-Notch perturbation background. (A-D) Morpholino antisense oligonucleotide targeting the Delta-Notch mediator *hesC* induces supernumerary pigment cells in the ectoderm of *E. tribuloides*. (A) Pigment cell counts for individual larvae (5 days post fertilization) either injected with *hesC* MASO or uninjected. (B) Bar graph showing mean number of pigment cells per larva in either uninjected control or injected with *hesC*
MASO. Larvae injected with MASO targeting hesc exhibited significantly more pigment cells. p value ≤ 0.005, Mann-Whitney rank-sum test. (C, D) Example larvae from both uninjected control and larvae injected with hesc MASO at 36 hpf. (E, F) Gcm expression is elevated in the ectoderm of embryos injected with hesc MASO. (G, H) Embryos treated with the Notch-inhibitor DAPT exhibit increased numbers of gcm-positive cells in the ectoderm at 26 and 40 hpf.

The data presented on D-V polarity in the NSM of E. tribuloides suggest that multiple regulatory domains unfold at and around the tip of the archenteron as gastrulation proceeds. Similar to euechinoids, this study determined that ese operates in the ventral NSM exclusive of gcm in the dorsal NSM. While the regulatory states in E. tribuloides need further refinement by two-color WMISH, for our purposes the overt disorder in its formation relative to the overt order of S. purpuratus NSM makes two salient points. First, early pregastrular or early gastrular polarity of NSM regulatory states represents an echinoid novelty, as no evidence for early mesodermal polarity exists in outgroup echinoderms [86, 88]. Second, if we take E. tribuloides as a proxy to the ancestral state for this character/regulatory state, then it is clear that the D-V polarity observed in the euechinoid NSM was shifted to occur prior to gastrulation in the lineage leading to modern euechinoids. On the other hand, an alternate evolutionary scenario is that NSM polarity manifested in these two modern echinoids is the result of two independent evolutionary trajectories with heterochronic and spatial differences, but both meeting a similar end in the diversification of NSM cell types in early development. That at least two D-V regulatory states are common to these embryos and that gcm is
downstream of Nodal and Notch signaling provide support for the first scenario. Further investigation into the developmental timing and regulatory states of cidaroid NSM will be required to parse out the most likely evolutionary scenario.

**Ectodermal regulatory states in *E. tribuloides***

Correspondence between *E. tribuloides* and euechinoids in deployment of ectodermal regulatory factors provides support to the idea that ectodermal specification is constrained and that alteration to the circuitry is nontrivial in early development. However, major alterations have occurred to ectodermal patterning pathways in regards to deployment and rewiring of circuitry during the evolution of euechinoid lineages that possess direct-developing, non-feeding larvae [54, 89-91]. These observations support the idea that the pressures of selection can overwhelm strong evolutionary constraint in early development. Of course there are very interesting differences in *E. tribuloides* ectodermal spatiotemporal dynamics and regulatory states relative to euechinoids. For instance, perturbation of Nodal signaling reveals that, while initial specification events are highly similar, alterations likely have occurred to the regulation of *bmp2/4* and *tbx2/3*. In *E. tribuloides* *tbx2/3* is expressed in DE and dorsal NSM by mid-gastrula. By late gastrula, it is expressed in the lateral clusters of skeletogenic synthesis, at the tip of the gut, in the gut endoderm, and residually in the ectoderm. This unfolding pattern of *tbx2/3* expression in *E. tribuloides* has essentially been compressed into the early stages of euechinoid development [62]. In euechinoids, perturbation of Nodal signaling with SB431542 extinguishes dorsal ectodermal *tbx2/3* specifically in *P. lividus*, while not affecting its expression in SM [42]. In *E. tribuloides* I observed the expression domain of *tbx2/3* expand into VE upon perturbation with this inhibitor (Figure 5A, 5B). This
observation combined with the result that bmp2/4 responds differently to Nodal perturbation suggests altered GRN circuitry downstream of Nodal. However, the vast evolutionary distances between cidaroids and euechinoids and the conserved spatiotemporal deployment of regulatory factors strongly argue for developmental constraint of ectodermal patterning mechanisms.

Ciliary band formation and ANE patterning in *E. tribuloides* are evolutionarily interesting evolutionarily given the fact that cidaroids lack the pan-deuterostome apical senory organ [15, 16, 92, 93]. Understanding the alterations in GRN circuitry that accompanied the loss of this embryonic structure and its downstream consequences would provide insight into the evolution of embryonic morphology and GRN architecture. Previous studies indicated that ANE patterning in *E. tribuloides* is more similar to outgroup echinoderms than it is in euechinoids; though expression of CB and anterior regulatory factors, e.g. *onecut* and *nk2.1*, exhibited spatial distributions similar to those seen in euechinoids [81]. Here, I observed patterning and regulation of CB that are consistent with the hypothesis that this process is conserved in echinoids. Additionally, I observed the sequential spatial restriction of *foxq2* to ANE, a pan-bilaterian observation driven by endomesodermal *wnt* factors [25, 94-97]. These data suggest that specification of the apical sensory organ in *E. tribuloides* is developmentally downstream of these events and that the loss of this embryonic structure had little effect on conserved patterning of CB and anterior localization of *foxq2*.

Lastly, perturbation of D-V patterning drastically altered the spatial distribution of CB regulatory factor *onecut* and resulted in a belt of 6-10 cells encircling the *E. tribuloides* embryo as a single dense ciliary band. A similar result was also obtained for
Synaptotagmin in *E. tribuloides* by disruption of endomesodermal specification via zinc perturbation [81]. These results are consistent with the hypothesis that anteriorly positioned ANE repressors restrict CB fate to the equator when D-V patterning is disrupted. Indeed, in *S. purpuratus*, *foxq2* restricts CB positioning anteriorly in ANE [30].

Although it is clear from work in euechinoids that ANE regulatory factors do not expand when D-V patterning is disrupted [42], this is likely not the case in *E. tribuloides*. While ANE is greatly expanded in *E. tribuloides* relative to other echinoderms [94, 98], there is no evidence to indicate that it extends to the embryonic equator. The most likely scenario is that disruption of D-V patterning expands anteriorly positioned ANE repressors of CB, e.g. candidate regulatory factors being *foxq2* and *nk2.1*, and CB positioning occurs at the equator where a pan-ectodermal driver, e.g. SoxB1 [30, 42, 61], is able to drive *onecut* expression. Elevated mRNA levels of *foxq2* in *E. tribuloides* upon disruption of D-V patterning support this hypothesis (Figure 4A,4C). Further, the sequential vegetal-to-animal zygotic activation of *onecut* seen during *E. tribuloides* early development is consistent with the hypothesis of anteriorly positioned ANE repressors that must be cleared for proper *onecut* expression.

**Evolution of global embryonic domains in early development of echinoids**

Previous analyses of embryonic domain regulatory states in *E. tribuloides* surveyed SM regulatory factors [20] and anterior neural ectoderm specification [81]. Additionally, two previous studies investigated SM and early endomesodermal micromere regulatory factors in the Pacific-dwelling cidaroid *Prionocidaris baculosa* [19, 99]. Integrating these data into this study affords an analysis of global embryonic
regulatory states and GRN linkages over 268.8 mya of evolution in indirect-developing sea urchins. From these studies, numerous alterations to deployment and GRN circuitry at all levels of GRN topology can be enumerated. Here, I enumerate 19 changes in spatiotemporal deployment or regulation of ectodermal and mesodermal embryonic regulatory factors since the cidaroid-euechinoid divergence (Table 2). Prominent among rewiring events are those that have occurred in establishing polarity in mesodermal embryonic domains. Endodermal and ectodermal specification and regulatory states also have undergone change, but to a lesser degree. One hypothesis that can accommodate these observations is that endodermal and ectodermal developmental programs may be more recalcitrant to change than mesodermal programs due to their more ancient evolutionary origin, suggesting that accretion of process over evolutionary time is a mechanism of constraint in developmental programs [5]. Indeed, in euechinoids there have been additional layers of GRN topology accrued in mesodermal specification, e.g. the pmar1-hesc double-negative gate novelty [19, 20, 24], delta-dependent NSM specification [20, 36], etc., which cidaroids do not exhibit, and which may explain the observation that little to no appreciable change has been observed in the mesodermal developmental programs of L. variegatus, P. lividus and S. purpuratus, representatives of modern euechinoid lineages that diverged approximately 90 mya.

Table 2. Enumeration of evolutionary changes to GRN deployment since the cidaroid-euechinoid divergence.

<table>
<thead>
<tr>
<th>No.</th>
<th>Regulatory factor</th>
<th>Change in spatiotemporal dynamics</th>
<th>Description of change</th>
<th>Euechinoid citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>bmp2/4</td>
<td>heterotopy</td>
<td>Altered regulation in D-V perturbation background</td>
<td>42</td>
</tr>
</tbody>
</table>
Biased rates of change to GRN topology in early development

Davidson and Erwin [7] first proposed the hypothesis that the hierarchical nature of GRN structure would manifest unequal rates of change during developmental evolution. This hypothesis was formulated from experimental observations in multiple bilaterian lineages [3, 100], and its underlying principle is to couch the systematic
Here evidence was presented that affords a first approximation of the lability of GRN deployment and circuitry underlying GRN topology in early echinoid development. I have presented a comparative analysis of developmental programs that diverged in the middle Permian and that argues for domain-specific, biased rates of change in deployment of GRN regulatory factors. While the genomic hardwired changes underlying this bias were not revealed here, the confluence of spatial, temporal and experimental evidence strongly suggests that regulatory circuitry specifying mesodermal domains in early echinoid embryonic development has undergone substantially more alteration at all levels of GRN topology than endodermal and ectodermal domains. For the early embryo it is imperative to establish canonical domains that are tasked with highly conserved processes, e.g. boundary formation and gastrulation. Thus, rates of change to GRN topology will vary during embryonic development depending on the capacity of the domain to buffer the effect of any mutation. The prediction that recursively wired, hierarchical developmental GRNs constrain the possible trajectories of change in future lineages was a prescient observation that we are only now beginning to fully appreciate.

**Materials and Methods**

**Animals and embryo culture**

Adult *E. tribuloides* were obtained from KP Aquatics (Tavernier, Florida). Eggs were collected by gravity and washed four times in Millipore filtered sea water (MFSW). Eggs were fertilized with a dilute sperm solution, and embryo cultures with less than
95% fertilization were discarded. Embryos were developed in glass pyrex dishes in a temperature-controlled setting of 22°C, and MFSW was refreshed daily.

**Cloning and gene isolation**

RNAseq and genomic databases of *E. tribuloides* reads were utilized for primer design using euechinoid sequences as seeds for BLAST searches and subsequent verification of orthology. PCR products were cloned into PGEM-T vector (Promega) and sequence verified in house using an ABI 3730xl sequencer. WMISH antisense RNA probes were synthesized from restricted plasmid vectors using T7 or SP6 RNA polymerase with digoxygenin or fluorescein dUTP incorporation (Roche). Primers for WMISH and qPCR are listed in Table S1.

**Table S1. WMISH and qPCR primer sequences. To be placed in Supporting Information or uploaded as Excel spreadsheet.**

**Whole-mount *in situ* hybridization and mRNA transcript abundance**

Transcript abundance of mRNA was estimated as described [49]. Briefly, transcripts were estimated by counting the number of embryos and spiking in an external standard of quantified synthetic XenoRNA (Power SYBR Green Cells-to-Ct Kit, Thermo-Fisher Scientific) prior to RNA isolation (RNeasy, Qiagen). Thus, to each qPCR reaction a known amount of embryos and RNA were added and the transcript number is
deduced by ddCt method. Additionally, some estimates were made with an internal standard that had been previously quantified.

Whole-mount *in situ* hybridization (WMISH) was conducted as previously described [20]. The WMISH protocol slightly modified for double fluorescent WMISH (dfWMISH) with different antibodies and probe detection. Antibodies for dfWMISH were either Anti-DIG or Anti-FLU conjugated to horseradish peroxidase (Roche) at a concentration of 0.25 μg/mL. Probes were detected with the Tyramide Signal Amplification Plus kit (Perkin Elmer) by using cyanine 5 or fluorescein conjugates at a dilution of 1:4000 in TBST. The amplification reaction was quenched by addition of 1% hydrogen peroxide. The protocol then cycled back to the blocking step and proceeded as described to detect the second probe.

**Perturbations**

**Microinjections**

Dominant-negative cadherin RNA overexpression (dnCad or Δ-cadherin) was microinjected at a concentration of 1,000 ng/μL as previously described [20]. Translation blocking morpholino (MASO) targeting *hesc* mRNA transcript (AATCACAAGGTAAGACGAGGATGGT) was purchased from Gene Tools (Philomath, OR, USA) and microinjected at a concentration of 1 mM as described [20]. BACs were microinjected at a concentration of 60 ng per mL nuclease-free water in the presence of 10 ng HindIII-digested genomic carrier DNA.

**Small molecule inhibitors**

For perturbation of D-V patterning, both timing and concentration of treatment with the Alk4/5/7-antagonist SB431542 (Cat no. 1614, Tocris Bioscience) were
determined. MFSW containing 2x concentration of the inhibitor was added to an equal volume of embryo culture in a 6-well tissue culture plate. To determine the optimum concentration, embryo cultures were reared in 5, 15 and 30 μM SB431542. Embryos reared at 5 μM showed no gross morphological deformities, whereas embryos reared at 30 μM exhibited significant developmental delays and gross deformities. To determine the sensitive period of inhibitor exposure, embryo cultures were exposed to the inhibitor at 1, 12 and 24 hpf. Embryos cultured in the inhibitor from 1 hpf or 24 hpf onwards showed significant developmental delays or no significant morphological differences, respectively. Results of these manipulations showed that treatment with 15 μM SB431542 at 12 hpf was the concentration and sensitive period at which a majority of larvae in the culture showed the characteristic phenotypes of dorsalization: multiple centers of skeletal synthesis and an hourglass phenotype. For Notch perturbation, embryos were cultured in the Notch-antagonist DAPT (Cat No. S2215, Selleck Chemicals) at a concentration of 12 μM from 1 hpf onwards.

**Comparative RNA timecourse analysis and statistics**

Absolute mRNA transcript number was estimated as described above for regulatory genes in early development of *E. tribuloides* (Table S2). Comparative analyses of this dataset were based on published data from two euechinoids, *S. purpuratus* [82] and *P. lividus* [84]. As developmental timing differs between the three species, a one-to-one comparison of timecourse datapoints could not be obtained. This issue was resolved by utilizing the adjustments for *S. purpuratus* and *P. lividus* described in Gildor and Ben-Tabou de Leon [84]. The comparative timepoints used in this study are presented in Table S4. Absolute transcript number for each timepoint was
then ranked highest to lowest for each gene relative to itself. Spearman’s rank correlation coefficient ($\rho$) was chosen over Pearson’s correlation in order to reduce the influence of large differences sometimes observed in estimates of absolute mRNA transcript numbers. For each pair of orthologous genes for which data were available in two species $\rho$ was calculated; these data are presented for three species in Figure 6 and are found in Table S3. For comparative analysis of global embryonic regulatory factors shown in Figure 7, values for $\rho$ were calculated for 34 regulatory genes in *E. tribuloides* and *S. purpuratus* and were compared. Only regulatory factors for which the embryonic domain of expression is known in *E. tribuloides* were used in the analysis, though data for 55 regulatory genes are presented in Table S3. Values for $\rho$ were binned by their expression in embryonic regulatory domains in the *S. purpuratus* global developmental GRN (available at http://sugp.caltech.edu/endomes/). The standard statistical distribution is represented in Figure 7. Statistical significance was calculated for each embryonic domain using the average of all $\rho$ values (55 regulatory genes) as the expected mean. Conservation of regulatory gene deployment is then interpreted as $\rho$ values near 1, i.e. high correlation of temporal deployment between two species.

**Table S2. Spreadsheet of mRNA transcript abundance estimates in *E. tribuloides.*
To be placed in Supporting Information or uploaded as Excel spreadsheet.

**Table S3. Spreadsheet of Spearman’s rank correlation coefficients between *E. tribuloides*, *P. lividus* and *S. purpuratus*. To be placed in Supporting Information or uploaded as Excel spreadsheet.
**Table S4. Comparative developmental timepoints for *E. tribuloides, P. lividus* and *S. purpuratus.*

<table>
<thead>
<tr>
<th>Timepoint</th>
<th><em>E. tribuloides</em></th>
<th><em>P. lividus</em></th>
<th><em>S. purpuratus</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>7</td>
<td>13</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>8</td>
<td>14</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>9</td>
<td>16</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>10</td>
<td>18</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>11</td>
<td>20</td>
<td>18</td>
<td>24</td>
</tr>
<tr>
<td>12</td>
<td>22</td>
<td>20</td>
<td>26</td>
</tr>
<tr>
<td>13</td>
<td>24</td>
<td>22</td>
<td>29</td>
</tr>
<tr>
<td>14</td>
<td>26</td>
<td>24</td>
<td>31</td>
</tr>
</tbody>
</table>
References

15. Mortensen T. Contributions to the study of the development and larval forms of echinoderms. IV. Danske Vid Selsk Ser. 1938;9(7(3)):1-59.


82. Materna SC, Nam J, Davidson EH. High accuracy, high-resolution prevalence measurement for the majority of locally expressed regulatory genes in early sea urchin development. Gene Expr Patterns. 2010;10(4-5):177-84.


Figure 1: Gene expression patterns in C. elegans embryos.

A) Expression of gsc (A1: Graph; A2: LV 22h; A3: LV 32h).
B) Expression of onecut (B1: Graph; B2: LV 24h; B3: LV 40h).
C) Expression of irxa (C1: Graph; C2: LV 28h; C3: LV 40h).

mRNA transcripts per embryo vs. hours post fertilization.
28 hours post fertilization

A

<table>
<thead>
<tr>
<th>Control</th>
<th>SB431542</th>
</tr>
</thead>
<tbody>
<tr>
<td>chordin</td>
<td>chordin</td>
</tr>
<tr>
<td>gcm</td>
<td>gcm</td>
</tr>
<tr>
<td>nodal</td>
<td>nodal</td>
</tr>
<tr>
<td>not</td>
<td>not</td>
</tr>
<tr>
<td>tbx2/3</td>
<td>tbx2/3</td>
</tr>
</tbody>
</table>

AV AV LV LV LV AV

40 hours post fertilization

B

<table>
<thead>
<tr>
<th>Control</th>
<th>SB431542</th>
</tr>
</thead>
<tbody>
<tr>
<td>chordin</td>
<td>chordin</td>
</tr>
<tr>
<td>gcm</td>
<td>gcm</td>
</tr>
<tr>
<td>nodal</td>
<td>nodal</td>
</tr>
<tr>
<td>not</td>
<td>not</td>
</tr>
<tr>
<td>tbx2/3</td>
<td>tbx2/3</td>
</tr>
</tbody>
</table>

LV AV LV LV LV LV

70 hours post fertilization

C

<table>
<thead>
<tr>
<th>Control</th>
<th>SB431542</th>
</tr>
</thead>
<tbody>
<tr>
<td>onecut</td>
<td>onecut</td>
</tr>
<tr>
<td>onecut</td>
<td>onecut</td>
</tr>
<tr>
<td>onecut</td>
<td>onecut</td>
</tr>
</tbody>
</table>

LV LV LV LV LV
The figure shows various stages of development for different genes.

- **bra (B):** Stages 28h, 40h, 50h, 70h
- **chordin (C):** Stages 18h, 21h, 24h, 28h, 50h, 70h
- **foxq2 (D):** Stages 6h, 8h, 10h, 12h, 15h, 20h, 22h, 50h
- **lefty (E):** Stages 15h, 24h, 40h, 40h, 50h, 70h
- **msx (F):** Stages 24h, 30h, 40h, 50h
- **nodal (G):** Stages 18h, 21h, 40h
- **not (H):** Stages 24h, 30h, 40h, 70h

Each stage is indicated with a time label.
<table>
<thead>
<tr>
<th>SpOnecut BAC construct*</th>
<th>CRM perturbation</th>
<th>embryos scored</th>
<th>Percent expressing reporter</th>
<th>Percent expressing in ciliary band</th>
<th>Percent expressing in dorsal ectoderm</th>
<th>Percent expressing in ventral ectoderm</th>
<th>Percent expressing in gut or mesenchyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAC Reporter 1 (B)</td>
<td>all CRMs</td>
<td>77</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>BAC Reporter 2 (C)</td>
<td>initial activation</td>
<td>56</td>
<td>12.5%</td>
<td>7.1%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>BAC Reporter 3 (D)</td>
<td>post-activation maintenance</td>
<td>40</td>
<td>10.0%</td>
<td>7.5%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>BAC Reporter 4 (E)</td>
<td>DE &amp; VE clearance</td>
<td>123</td>
<td>27.6%</td>
<td>12.2%</td>
<td>2.4%</td>
<td>10.4%</td>
<td>0.0%</td>
</tr>
<tr>
<td>BAC Reporter 5 (F)</td>
<td>none / wild-type</td>
<td>102</td>
<td>17.6%</td>
<td>14.7%</td>
<td>1.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

B. Sp Onecut BAC reporter 1 in *E. tribuloides*
C. Sp Onecut BAC reporter 2 in *E. tribuloides*
D. Sp Onecut BAC reporter 3 in *E. tribuloides*
E. Sp Onecut BAC reporter 4 in *E. tribuloides*
F. Sp Onecut BAC reporter 5 in *E. tribuloides*
A Pigment cell counts

<table>
<thead>
<tr>
<th>Control</th>
<th>hesc MASO</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>61</td>
</tr>
<tr>
<td>20</td>
<td>24</td>
</tr>
<tr>
<td>28</td>
<td>65</td>
</tr>
<tr>
<td>38</td>
<td>38</td>
</tr>
<tr>
<td>38</td>
<td>69</td>
</tr>
<tr>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>26</td>
<td>52</td>
</tr>
<tr>
<td>35</td>
<td>40</td>
</tr>
<tr>
<td>36</td>
<td>42</td>
</tr>
<tr>
<td>29</td>
<td>49</td>
</tr>
<tr>
<td>26</td>
<td>72</td>
</tr>
<tr>
<td>30</td>
<td>48</td>
</tr>
<tr>
<td>38</td>
<td>46</td>
</tr>
<tr>
<td>35</td>
<td>27</td>
</tr>
<tr>
<td>40</td>
<td>29</td>
</tr>
<tr>
<td>36</td>
<td>44</td>
</tr>
<tr>
<td>38</td>
<td>36</td>
</tr>
</tbody>
</table>

B Number of pigment cells

C Uninjected control

D hesc MASO

E Uninjected control

F hesc MASO

G Control

H 15 uM DAPT