Evolutionary changes in the chromatin landscape reshape a developmental gene regulatory network during rapid life history divergence in sea urchins

Phillip L. Davidson[1], Maria Byrne[2], Gregory A. Wray[1*]

1. Department of Biology, Duke University, Durham, NC, USA
2. School of Life and Environmental Sciences, University of Sydney, Sydney, NSW, AU

*** Corresponding Author:** gwray@duke.edu

ABSTRACT

Chromatin configuration is highly dynamic during embryonic development in animals, exerting an important point of control in transcriptional regulation. Yet there exists remarkably little information about the contribution of evolutionary changes in chromatin configuration to the evolution of gene expression and organismal traits. Genome-wide assays of chromatin configuration, coupled with whole-genome alignments, can help address this gap in knowledge in several ways. In this study we present a comparative analysis of regulatory element sequences and accessibility throughout embryogenesis in three sea urchin species with divergent life histories: a lecithotroph *Heliocidaris erythrogramma*, a closely related planktotroph *H. tuberculata*, and a distantly related planktotroph *Lytechinus variegatus*. We identified distinct epigenetic and mutational signatures of evolutionary modifications to *cis*-regulatory element function in *H. erythrogramma* that have accumulated non-uniformly throughout the genome, suggesting selection, rather than drift, underlies many modifications associated with the derived life history. Specifically, regulatory elements composing the sea urchin developmental gene regulatory network are enriched for signatures of positive selection and accessibility changes which may function to alter binding affinity and access of developmental transcription factors to these sites. Furthermore, regulatory element changes often correlate with divergent expression patterns of genes involved in cell type specification, morphogenesis, and development of other derived traits, suggesting these evolutionary modifications have been consequential for phenotypic evolution in *H. erythrogramma*. Collectively, our results demonstrate that selective pressures imposed by changes in developmental life history rapidly reshape the *cis*-regulatory landscape of core developmental genes to generate novel traits and embryonic programs.

35 **INTRODUCTION**

36 *Cis*-regulatory elements are significant contributors to phenotypic evolution and trait
37 diversification by modulating transcriptional regulation via gene network interactions (Davidson
38 and Erwin 2006; Wray 2007; Wittkopp and Kalay 2012). Extensive empirical evidence
39 demonstrates that the function of *cis*-regulatory elements depends on distinct molecular
40 characteristics such as the element's sequence and epigenetic landscape that collectively provide
41 a dynamic and heritable genomic substrate affecting inter- and intraspecific diversification.
42 Enabled by advances in genomic technology and reduced sequencing costs, more recent work
43 has demonstrated the importance of *cis*-regulatory elements in diverse biological contexts across
44 a variety of animal systems at a genome-wide scale; for example: origination of human-specific
45 traits (Shibata, et al. 2012; Swain-Lenz, et al. 2019), convergent evolution of flight loss in birds
46 (Sackton, et al. 2019), population-level adaptation (Lewis and Reed 2019), genetic assimilation
47 (van der Burg, et al. 2020) and color identity (Livraghi, et al. 2021) in butterfly wing patterning,
48 embryonic regulation of gene expression in the crustacean *Parhyale hawaiensis* (Sun and Patel
49 2021), limb bud development in chicken and mice (Jhanwar, et al. 2021), and single-cell
50 resolution of *Drosophila* cell lineages (Cusanovich, et al. 2018). However, direct integration of
51 large-scale *cis*-regulatory analysis and transcription in a defined evolutionary-developmental
52 (evo-devo) context is lacking. For example, relationships governing divergence in chromatin and
53 gene expression, the role of such changes in trait evolution, and the action of natural selection on
54 these processes remain largely unexplored within an interrogable phylogenetic system. In this
55 study, we address these important questions by comparing the function of *cis*-regulatory
56 elements in embryonic gene expression of sea urchin species with divergent developmental life

57 history strategies to assess the extent *cis*-regulatory elements play in the evolution of gene
58 expression, developmental systems, and by extension, trait variation and origination.
59 Among the most common and well-studied life history shifts in the animal kingdom is
60 the evolution of feeding or non-feeding larval development (Flatt and Heyland 2011). In marine
61 invertebrates, planktotrophy (feeding development) is characterized by species that produce
62 many, relatively small eggs and have a longer planktonic larval period, whereas species with
63 lecithotrophic (non-feeding) development invest more maternal resources into a fewer number of
64 large eggs that develop through an abbreviated larval stage (Strathmann 1985). Importantly, the
65 ecological circumstances associated with each strategy are predictable and well-studied across
66 diverse marine clades (Vance 1973; Strathmann 1985; Pechenik 1999; Marshall, et al. 2012;
67 Morgan 2020), providing natural systems to analyze how selection shapes often rapid phenotypic
68 divergence in developmental regulation and physiology. For example, within Echinoderms,
69 lecithotrophy has independently evolved from the ancestral planktotrophic state more than 20
70 times (Raff 1987; Wray 1996; McEdward and Miner 2001; Hart 2002), implying that recurring
71 sets of environmental conditions and strong selection can enact evolutionary modifications to
72 even highly conserved developmental processes. The consequences include changes at every
73 developmental stage ranging from egg composition (Prowse, et al. 2009; Falkner, et al. 2015;
74 Byrne and Sewell 2019; Davidson, et al. 2019) and zygotic genome activation (Wray and Raff
75 1990; Israel, et al. 2016; Wang, et al. 2020), cell type specification and embryonic patterning
76 (Wray and Raff 1989; Love and Raff 2006), to construction of larval structures such as the
77 skeleton and rudiment (Emlet 1995; Koop, et al. 2017), the latter being the region from which
78 the adult emerges during metamorphosis. Still, the impact selection exerts on the evolution of
79 *cis*-regulatory interactions underlying alternative developmental programs and contrasting life

80 history strategies is not known. This gap in knowledge is manifested more broadly in our current
81 sparse understanding of the biological relationships governing regulatory element function,
82 evolutionary constraints inherent to development and gene network architecture, and phenotypic
83 diversification.

84 Perhaps the best studied example of the life history switch from planktotrophic to
85 lecithotrophic development is the sea urchin *Heliocidaris erythrogramma*. Prior work has found
86 dramatic modifications to morphogenesis and embryonic patterning in *H. erythrogramma*
87 relative to planktotrophic sea urchin species (for reviews see Raff and Byrne 2006; Wray 2021).
88 For instance, dorso-ventral and left-right axis establishment occurs considerably early in this
89 species (Henry, et al. 1990), concordant with accelerated rudiment development and
90 metamorphosis (Williams and Anderson 1975). At other embryonic domains, specification of
91 cell lineages such as those responsible for mesenchymal tissue types have been delayed, lost, or
92 spatially rearranged (Wray and Raff 1990; Klueg, et al. 1997; Wilson, Andrews, Rudolf Turner,
93 et al. 2005; Davidson, et al. 2021). Recent transcriptomic studies have shown that expression
94 profiles of key developmental gene regulatory network (dGRN) genes in this species have
95 diverged from the highly conserved expression patterns in planktotrophic species (Israel, et al.
96 2016), and many of these regulatory changes are predicted to be based in *cis* (Wang, et al. 2020).
97 The contribution of evolutionary changes in the sequence and configuration of *cis*-regulatory
98 elements to the rapid evolution of these derived phenotype in *H. erythrogramma* is not known.

99 To address this, we compared chromatin accessiblity, regulatory element sequence, and
100 gene expression across six developmental stages in three sea urchin species spanning this life
101 history switch: the lecithotroph *H. erythrogramma*, a closely-related (~ 5 my divergence)
102 planktotroph *H. tuberculata*, and a distantly-related (~ 40 my divergence) planktotroph,

103 *Lytechinus variegatus* (**fig. 1a**). Planktotrophic (feeding) larvae represent the *ancestral* life
104 history mode within echinoderms relative to the *derived* lecithotrophic state (Strathmann 1978,
105 Raff 1987, Wray 1996, McEdward and Miner 2001), so recent and dramatic changes in
106 developmental phenotypes in *H. erythrogramma* provide an exceptional natural system to
107 analyze the consequences of strong shifts in selection and life history on the evolution of gene
108 regulatory mechanisms. We emphasize the distinction between two molecular mechanisms,
109 chromatin accessibility and sequence evolution, in generating change to *cis*-regulatory element
110 function and developmental gene expression. We identify distinct epigenetic and mutational
111 signatures of evolutionary modifications in the *H. erythrogramma* genome relative to species
112 with the ancestral life history. Furthermore, these regulatory changes have accumulated non-
113 uniformly throughout the genome and often correlate with derived expression patterns of genes
114 involved in cell type specification, morphogenesis, and development of other derived traits.
115 Importantly, the patterns of *cis*-regulatory modifications and their impact on the evolution of
116 transcriptional regulation characterized here may be broadly applied to biological principles
117 underlying the integration of developmental and gene regulatory mechanisms in lineage
118 diversification.

119 **RESULTS AND DISCUSSION**

120 We performed ATAC-seq on six developmental stages of *H. erythrogramma* (*He*), *H.*
121 *tuberculata* (*Ht*), and *L. variegatus* (*Lv*) in order to identify open chromatin regions (OCRs)
122 representing putative *cis*-regulatory elements (CREs) during embryogenesis. In this study, we
123 use the term "OCR" when reporting analysis results because ATAC-seq is an assay of open
124 chromatin. However, we apply the more specific term "CRE" when discussing the biological
125 interpretation of our results because *cis*-regulatory elements are the operational units identified

126 by open chromatin peaks used to understand mechanisms of gene regulatory evolution. The
127 developmental stages analyzed here include $4^{th}/5^{th}$, $6^{th}/7^{th}$, and $8^{th}/9^{th}$ mixed cleavage stage
128 embryos, as well as hatched blastula, mid-gastrula, and early larva (**fig. 1b**), enabling
129 characterization of regulatory element dynamics from the onset of zygotic transcription through
130 the specification and morphogenesis of complex tissue types. We then carried out alignments of
131 chromosome-scale genome assemblies to find evolutionarily conserved OCRs between species to
132 compare accessibility and sequence composition of CREs. These analyses provide a
133 simultaneous comparison of gene regulatory mechanisms in two ways: 1) availability of CREs to
134 transcription factor binding by modulating chromatin configuration (accessibility) and 2) affinity
135 of transcription factor binding to CREs via changes in nucleotide sequence, as evidenced by
136 motif analysis and signatures of positive selection. Importantly, the genomes of these sea urchin
137 species were assembled and annotated using an identical set of sequencing technology and
138 methodology (Davidson, et al. 2020; Davidson, et al. 2021), minimizing technical bias and
139 permitting high resolution identification of changes in regulatory element activity.

140 Of 42,679 OCRs identified across the three species, we found 35,788 1:1:1 orthologous,
141 high confidence OCRs proximal to 11,679 gene models for cross-species comparison (**fig. 2a**).
142 Note that inclusion in this set of 1:1:1 orthologues does not require that chromatin be open in all
143 three species, but rather that it be open in at least one species at one time point surveyed and that
144 the orthologous sequence is present in all three species. This set of orthologous OCRs forms the
145 basis of most analyses presented here, including comparisons of developmental chromatin
146 configuration (**fig. 2b, 3**), analyses of OCR sequence changes associated with positive selection
147 and transcription factor binding (**fig. 2d, 4**), discerning how these changes are distributed among
148 OCRs located proximally or distally to protein coding sequences (**fig. 2c, 5**), and determining the

149 extent to which these regulatory modifications are associated with derived gene expression
150 patterns (**figs. 5-7**). In addition to orthologous OCRs, we also identified sets of species-specific
151 gains or losses of OCRs occurring on entire genomic segments present or lost in a single species
152 (**fig. 2a**). Because these indel-based differences in OCRs are only physically present in at most
153 two of the three sea urchin species compared in this study, they are treated as distinct sets of
154 putative CREs. While the changes in regulatory element function examined in this study are not
155 applicable to species-specific OCRs, it is important to note that many physical OCR gains occur
156 nearby genes with differential expression levels in *H. erythrogramma*, suggesting gain or loss of
157 entire CREs may be an important component of the evolution of this species' derived life history
158 strategy (**supplementary fig. 1**).

159 ***Development and life history drive global variation patterns of accessibility***

160 First, we assessed which biological factors may significantly contribute to variation in
161 chromatin accessibility at a genomic scale. Principal component analysis (PCA) suggests
162 species-level differences (PC 1: 20.3%) and the interaction of development and life history (PC
163 2: 18.5%) are the most important factors explaining variation in chromatin accessibility when all
164 orthologous OCRs are considered (**fig. 1c**). Within each species, chromatin accessibility
165 differences between stages roughly mirrors developmental progression as early and late stages
166 cluster more closely to one another. Strikingly, though, there is far greater variation in
167 accessibility through development in both species with the ancestral life history relative to *He*,
168 suggesting the developmental chromatin landscape of the derived state is less dynamic during
169 embryogenesis. This finding is consistent with delayed blastomere specification (Wray and Raff
170 1989) and simplified larval morphology in *He* (Emlet 1995), suggesting chromatin state changes

171 normally associated with establishing cell identity and tissue differentiation have been modified
172 during the evolution of its derived developmental life history.

173 Interestingly, when considering only OCRs near dGRN genes, the primary driver of
174 variation in chromatin accessibility shifts to developmental life history rather than species-level
175 differences (**fig. 1d**; PC 1: 25.5%), a finding corroborated by variance partitioning of the dataset
176 (**supplementary fig. 2**). This result suggests chromatin configuration of critical developmental
177 network genes depends significantly on the species' developmental strategy, as access to
178 regulatory elements of these genes controls the flow and deployment of embryonic programs.
179 Therefore, changes in accessibility to CREs have the potential to modify the function and
180 construction of developmental gene networks by altering regulatory interactions, and by
181 extension, drive the evolution of novel developmental phenotypes.

182 ***Loss of regulatory element accessibility within the ancestral dGRN***

183 We next quantified how the accessibility of 35,788 OCRs vary between life history
184 strategies in order to better understand how chromatin configuration at specific CREs may
185 influence the evolution and function of alternative embryonic programs. Across each cleavage
186 stage examined (timepoints spanning 16-cell through ~500-cell stage embryos), we found an
187 average of 5.7% of OCRs were differentially accessible between ancestral and derived sea urchin
188 species, including a modest difference in sites more accessible in *Ht* and *Lv* (3.1%) or *He* (2.6%)
189 (**fig. 3a**). Strikingly, starting at the blastula stage, the proportion of differentially accessible
190 OCRs across the genome rises, with most sites increasing in accessibility ("opening") in the
191 ancestral life history but not in *He* (**fig. 3a**). For example, by the larval stage, 9.5% of OCRs are
192 differentially accessible between life histories, nearly three quarters of which (73.9%) are more
193 accessible in *Ht* and *Lv*. Furthermore, when all OCRs are considered, including both orthologous

194 and non-orthologous sites across species, a much greater proportion of the *He* genome is closed
195 compared to *Ht* and *Lv*, suggesting that, at a genome-wide scale, *He* is characterized by zygotic
196 quiescence during early development relative to ancestral life history (**fig. 2e**; **supplementary**
197 **fig. 3**). Our results support the prediction that as embryonic development proceeds, more
198 regulatory sites open and become available to regulatory factor binding as zygotic transcription
199 activates, cell lineages differentiate, tissue types are specified, and morphological structures are
200 constructed. However, this genome "awakening" and reconfiguration of chromatin state across
201 the embryo unfolds at a much grander scale in *Ht* and *Lv* relative to *He*. This finding supports the
202 idea that embryonic fate specification in *He* is delayed relative to the ancestral state (Wray and
203 Raff 1989, 1990) and collectively, is concordant with a morphologically simplified larva
204 (Williams and Anderson 1975) and derived gene expression patterning in this species (Wilson,
205 Andrews and Raff 2005; Wilson, Andrews, Rudolf Turner, et al. 2005; Smith, et al. 2008; Israel,
206 et al. 2016; Koop, et al. 2017). Reduced accessibility of tissue-specific enhancers may alter gene
207 networks tasked with the establishment of differentiated cell types in the ancestral life history,
208 generating novel interactions and developmental phenotypes in the process.

209 This hypothesis is further supported when OCR accessibility of dGRN genes are
210 compared. Near these critical developmental regulatory genes, we find a remarkable asymmetry
211 in CRE accessibility between life history modes from the onset of development – on average,
212 0.5% of OCRs near dGRNs are more accessible in *He* during development, whereas this figure
213 ranges from 2.8% (early cleavage) to 15.1% (larva) for species with the ancestral life history (**fig.**
214 **3b**). An example of CRE activation in the ancestral state that is not found in *He* is the chromatin
215 landscape proximal to the gene *patched* (*ptc*) (**fig. 3c**), a receptor of the Hedgehog ligand whose
216 misexpression results in skeletal defects and abnormal mesodermal patterning in the ancestral sea

217 urchin developmental program (Walton, et al. 2009). Therefore, modifications to CRE
218 accessibility in *He* are not uniformly distributed across the genome, as would be expected of
219 neutral processes such as drift, but instead are highly concentrated near dGRN genes (Fisher
220 exact test, two-sided: $p < 1.31 \times 10^{-11}$). This finding suggests natural selection has sculpted the
221 epigenetic landscape composing the sea urchin dGRN during life history evolution, generating
222 adaptive changes to transcriptional regulation and expression patterning during embryogenesis.
223 This finding further suggests that the evolutionary changes in putative CRE accessibility have
224 functional consequences. If they had no impact on gene expression or some other trait, natural
225 selection would not influence their status during development.

226 ***Developmental regulatory elements have accumulated mutational signatures of positive***
227 ***selection***

228 Next, we analyzed sequence evolution of orthologous regulatory loci to identify instances
229 of mutational pileup as a signal of positive selection within these sites. Substitution rates of
230 35,788 OCRs on the *He* and *Ht* branches were compared to 10 randomly selected neutrally
231 evolving, non-coding reference sequences, and a likelihood ratio test was implemented to fit
232 substitution rates across the tree to a branch-specific positive selection model, using *Lv* as the
233 outgroup species (see **Methods and Materials** for details). Our goal was to identify cases of
234 gene regulatory evolution via nucleotide variation within CRE sequences, potentially acting to
235 modify transcription factor binding affinity of these elements and associated changes to gene
236 network expression and connectivity.

237 In total, we identified 648 OCRs with evidence of positive selection in only the *Ht*
238 genome, compared to 512 in the *He* genome (**fig. 4a**: green and orange points, respectively).
239 Unlike OCR accessibility, a trait that changes during development and differs among tissues,

240 OCR sequence is an immutable genomic property within the lifetime of an individual (excluding
241 somatic mutation and gametogenesis). This result thus suggests that, while the life history of
242 these two sea urchin species dramatically differs, many potential changes in gene regulation
243 associated with other aspects of their biology have accumulated within each species. At a
244 genome-wide scale, these adaptive changes have occurred perhaps even more so in *Ht*, the
245 species with ancestral developmental life history.

246 Partitioning OCR selection results by the function of the nearest gene reveals only limited
247 enrichment of positive selection within the putative regulatory elements of any particular class of
248 genes (**fig. 4b**, **supplementary fig. 4**). A prominent exception to this finding involves OCRs near
249 genes in the sea urchin dGRN or near genes encoding other transcription factors. In these cases,
250 approximately twice as many OCRs exhibit evidence of positive selection on the *He* branch
251 relative to *Ht*. This result is especially striking considering *Ht* has 26.3% more OCRs with
252 branch-specific evidence of positive selection across the entire genome (previous paragraph and
253 **fig. 4b**). Moreover, nucleotide sequences of the 23 dGRN OCRs under selection in *He* are highly
254 enriched for binding motifs of critical developmental transcription factors, including three SOX,
255 three zinc finger, and one homeobox protein (**fig. 4c**). The SOX binding motifs are interesting
256 because many transcription factors in this family act as "pioneer factors" that are capable of
257 initiating chromatin opening (Zaret and Carroll 2011). Another interesting example involves
258 NRF2: the binding motif for this transcription factor is overrepresented in OCRs under selection
259 in *He* (**fig. 4c**) and the gene is expressed is much at much higher levels in *He* embryos relative to
260 those in the ancestral life history (**fig. 4d**). This protein has gene regulatory functions in diverse
261 biological contexts ranging from antioxidant response and cellular defense to metabolism and
262 stem cell quiescence (Tonelli, et al. 2018), which in this case, could be related to derived

263 embryonic detoxification and physiological systems (Tsushima, et al. 1995; Davidson, et al.
264 2019) or early cell lineage specification programs (Wray and Raff 1989, 1990; Davidson, et al.
265 2021) of *He*. Taken together, regulatory elements tasked with wiring the ancestral sea urchin
266 dGRN have accumulated a disproportionate number of mutations on the *He* branch. These likely
267 play a significant role in regulation of divergent gene expression patterns and its highly derived
268 developmental program. More broadly, our results provide strong evidence that natural selection
269 can efficiently generate heritable changes concentrated within the *cis*-regulatory landscape of
270 specific subsets of genes or gene networks to rapidly produce changes in transcriptional
271 regulation.

272 **Overlap of evolutionary changes in regulatory element accessibility and sequence**

273 The previous two sections describe how two distinct modifications to regulatory element
274 function, namely accessibility and sequence, have accumulated across the *He* genome, and
275 particularly so near critical developmental network genes. We next asked whether there exists
276 any interaction between these two evolutionary changes to better understand the extent to which
277 both epigenetic and sequence changes affect the same regulatory element. At all three post-
278 cleavage stages, we found that differentially accessible OCRs contain a significantly greater
279 proportion of sites with evidence of selection on the *He* branch relative to non-differentially
280 accessible OCRs, including as much as an ~50 % proportional increase at the blastula stage
281 (Fisher exact test, two-sided: $p < 5 \times 10^{-3}$) (**fig 5a**: top). This same trend continues at early
282 cleavage (16/32-cell embryos), although not statistically significant. In contrast, we detected no
283 statistical enrichment for positive selection within differentially accessible OCRs in the *Ht*
284 genome – at early cleavage this relationship does approach significance, but no such trend exists
285 at the later developmental stages (**fig. 5a**: bottom) (Fisher exact test, two-sided: $p > 0.05$).

286 Significant overlap in differential accessibility and positive selection in the *He* genome
287 suggests a possible biological interaction between the two evolutionary mechanisms that may
288 operate in tandem to alter regulatory function. We present two simple models for how this could
289 occur. 1) *Accessibility first*: an increase or decrease in accessibility of a CRE at a particular
290 developmental stage or tissue type makes the CRE more or less visible to selection. Following an
291 increase in accessibility, mutations conferring adaptive changes in gene regulation during
292 development would be selected for and increase in frequency within the population, whereas the
293 same mutations in a closed OCR would have no effect; CREs with reduced accessibility, on the
294 other hand, may accumulate mutations that reinforce its derived state, ensuring that it remains
295 closed and unable to interact with transcription factors. 2) *Mutation first*: mutations within the
296 regulatory element drive increased accessibility of the site, perhaps through recruitment of
297 pioneer transcription factors, histone remodelers, or other chromatin modifiers, thereby
298 activating the element's regulatory function; alternatively, if the regulatory element's function is
299 detrimental to regulating gene expression, then mutations that disable the element's function or
300 decrease its accessibility via changes to chromatin configuration would be selected. Regardless
301 of mechanism, the disproportionate occurrence of positive selection within differentially
302 accessible OCRs only in the *He* genome suggests a potentially synergistic set of gene regulatory
303 processes may operate to drive developmental change during the evolution of its derived life
304 history. The observation that this association is restricted to the *He* branch suggests that this
305 interaction often produces adaptive changes in transcription during development. Tandem
306 accessibility and mutational changes may be a common gene regulatory method to efficiently
307 modify gene expression for organismal adaptation and innovation across organismal lineages.
308

309 **Relationship of evolutionary regulatory mechanisms between proximal and distal elements**

310 Next, we asked whether evolutionary changes in OCR sequence and chromatin
311 accessibility differ between core promoters and more distal elements such as enhancers. We
312 quantified changes in proximal (< 500 bp to translational start site of a gene) and distal (≥ 500 bp
313 to translational start site of a gene) OCRs to test if either change covaries with promoter or
314 enhancer evolution between species. Although many proximal and distal OCRs show evidence
315 of positive selection on the *Ht* and *He* branches, we found no significant enrichment between the
316 two categories of putative CREs (**fig. 5b**). This finding suggests mutations in both enhancers and
317 promoters are likely important contributors in driving gene expression change, via altered
318 binding and recruitment of regulatory factors and transcriptional machinery.

319 In contrast, accessibility changes shows a much greater asymmetry between proximal and
320 distal elements (**fig. 5c**). While accessibility distributions are comparable at early cleavage
321 (16/32-cell embryos), many more accessibility differences accumulate in proximal OCRs relative
322 to distal ones later in development, after zygotic activation and tissue specification has
323 commenced. For example, by the gastrula and larval stages, more than 20% of proximal
324 elements are significantly more accessible in *Lv* and *Ht*, whereas this figure is ~5% for distal
325 elements.

326 There are several possible biological and technical explanations for this result. Unlike
327 enhancers, whose activity is often restricted to a single tissue or stage (Thurman, et al. 2012),
328 promoters need to be accessible whenever the gene is transcribed, regardless of space and time.
329 This has important implications for interpreting differences in chromatin configuration among
330 species. At a biological level, greater promoter accessibility in the embryos of species with the
331 ancestral lie history (*Ht* and *Lv*) may merely reflect overall greater transcriptional activity and

332 more changes in expression during development (**fig. 1,3; supplementary fig. 5)** (Israel, et al.
333 2016; Wang, et al. 2020). For a given gene, the accessibility of a promoter region is normally
334 greater than that of nearby enhancers (Thurman, et al. 2012), which may lend greater statical
335 power in detecting accessibility differences between species and life histories. In other words,
336 smaller effect sizes associated with overall reduced accessibility of enhancers relative to
337 promoters may contribute to the apparent enrichment of promoter accessibility changes. In
338 addition, these chromatin configuration data were collected from whole embryos, so cell or
339 lineage-specific differences in accessibility will be averaged within the sample, a technical factor
340 that may be especially important for detecting accessibility changes of enhancer elements with
341 tissue-specific dynamics. Lastly, promoter accessibility changes may simply be the most
342 efficient way to rapidly modulate large changes in transcription, at least for some genes.
343 Adjusting enhancer accessibility can certainly induce changes gene expression, but activating or
344 shutting down a promoter may be a much more potent mechanism for large-scale changes in
345 transcriptional regulation, such as those involved in rapid developmental evolution.

346 **Relationships governing magnitude of gene expression and chromatin accessibility changes**

347 Multiple lines of evidence presented above suggest that changes in the sequence or
348 accessibility of *cis*-regulatory elements contributed to the evolution of development in *He*. The
349 association between these changes and divergence in gene expression, however, is unclear. To
350 understand this relationship, we compared mRNA expression fold-change of genes with
351 accessibility fold-change of nearby OCRs at four developmental stages to measure instances of
352 correlated changes in the magnitude of expression and chromatin configuration (**fig. 5d-e**).
353 When all genes and OCRs are considered (**fig. 5e**: top row), there is no clear association between
354 the magnitude and direction of accessibility and expression differences across developmental life

355 histories. This result is not surprising at a global scale, given that expression levels and
356 accessibility are conserved between species for most genes and OCRs, respectively. However, a
357 notable set of proximal sites have greater accessibility in species with the ancestral life history at
358 the blastula, gastrula, and larval stages, consistent with overall greater promoter accessibility in
359 *Ht* and *Lv* (see previous section). A similar result occurs when only differentially expressed
360 genes and nearby OCRs are considered (**fig. 5e**: middle row). One possible explanation for the
361 lack of an obvious relationship is that, while only genes that are differentially expressed between
362 life histories are plotted, each of these genes has an average of 3.1 OCRs assigned to it (based on
363 proximity). Each of these OCRs may not be active at a given stage of development, or they may
364 function antagonistically to one another (i.e. regulatory elements can enhance or repress
365 transcription), thus collectively dampening a correlative signal between gene expression and
366 accessibility.

367 Lastly, we considered only differentially expressed genes and nearby OCRs with
368 differential accessibility. This analysis reveals interesting relationships between transcription and
369 chromatin configuration (**fig. 5e**: bottom row). As development progress, the relatively diffuse
370 relationship between CRE accessibility and gene expression quickly tightens (compare 16/32-
371 cell stage with larva). This likely reflects synchronization of zygotic genome activation with the
372 embryo's *cis*-regulatory landscape as maternal mRNA reserves are depleted and nascent
373 transcription and cell differentiation commences. Greater accessibility at the majority of
374 proximal sites in *Ht* and *Lv* is still found in this subset of genes and OCRs, though smaller sets of
375 promoters with greater accessibility in *He* and corresponding expression increases are detected
376 (faint gray spots). Among distal elements, we observed a considerably more diffuse relationship
377 between accessibility, expression, and life history, possibly reflecting the potential activating or

378 repressive functions of enhancers and downstream gene network interactions. At each
379 developmental stage, the most common relationship was increased gene expression and
380 accessibility in *Ht* and *Lv*, consistent with the preponderances of activating (as opposed to
381 repressive) interactions within the ancestral dGRN. Despite these relationships, our results
382 highlight the difficulty in correlating the magnitude of cross-species changes in gene expression
383 and regulatory element accessibility at a genomic scale considering: **i)** Expression and
384 accessibility of most genes and CREs are conserved; **ii)** CRE activity can have antagonistic
385 effects on transcriptional regulation as activators or repressors, dependent on both the CRE
386 sequence and *trans* environment specific to cell types or stages of development; **iii)** The activity
387 and quiescence of multiple CREs regulating the same gene dampens regulatory signals when
388 analyzed at the whole-embryo level; and **iv)** CRE accessibility is permissive rather than directive
389 of regulatory factor binding and downstream gene expression such that "openness" does not
390 necessarily indicate the CRE is actively regulating transcription.

391 **Developmental trajectories of chromatin configuration and gene expression have evolved**
392 **most frequently along the *H. erythrogramma* branch**

393 To address the challenges associated with understanding the relationship between
394 expression and accessibility changes of putative regulatory elements, we applied the "jump
395 score" method to measure the magnitude of evolutionary changes in developmental trajectories,
396 which are represented as distance in principal component space (Israel, et al. 2016). This metric
397 is particularly useful because the jump score summarizes patterns of conservation or divergence
398 in gene expression across development, rather than a single stage. Here we apply this approach
399 to evolutionary changes in the accessibility of CREs across developmental trajectories.

400 We compared the expression jump score of 9,423 genes and the mean jump score of
401 nearby OCRs to test how developmental changes in transcriptional regulation and chromatin
402 status are distributed along the *He* and *Ht* branches. For example, expression of *insm*, a
403 transcription factor involved in neurogenesis in sea urchins (McClay, et al. 2018), is conserved in
404 across *He*, *Ht*, and *Lv* (**fig. 6a**), reflected by a low difference in jump scores between *He-Lv* and
405 *Ht-Lv* and tight clustering of expression profiles in principal component space (**fig. 6b**). Similar
406 expression profiles are accompanied by conserved *cis*-regulatory landscape dynamics through
407 development in each species, exemplified by a near-zero mean OCR jump score difference and
408 accessibility profile of ten nearby OCRs (**fig. 6c,d**). However, *brn 1/2/4*, a regulator of midgut
409 patterning (Yuh, et al. 2005), has a divergent developmental gene expression trajectory in *He*
410 relative to the ancestral state and large jump score differences in *He-Lv* and *Ht-Lv* branch
411 comparisons (**fig. 6e,f**). This expression change is accompanied by significant reduction in
412 chromatin accessibility at the blastula through larval stages in *He*, represented by a large mean
413 difference in OCR jump scores along the *He* branch (**fig. 6g,h**). This locus is especially
414 interesting considering *Hery* larvae do not feed and have lost development a functional through-
415 gut.
416 Given the highly derived developmental program of *He* relative to the ancestral life
417 history, we predicted that the most frequent changes in transcriptional regulation would occur
418 along the *He* branch. Indeed, concerted modifications to OCR accessibility and gene expression
419 profiles during *He* development are the most frequent set of changes among genes and OCRs
420 with divergent expression and chromatin accessibility trajectories (1,582 genes) (**supplementary**
421 **fig. 6**; **Data S7**). Furthermore, the least common combination of changes includes both
422 expression and accessibility changes occurring on the *Ht* branch (1,114 genes), consistent with

423 this species' more conserved embryonic gene expression program. Thus, we find that local
424 patterns of CRE accessibility have evolved in concert with divergent developmental expression
425 patterns during life history evolution in *He*. Collectively, our results suggest that modified
426 transcription factor access to regulatory elements can generate rapid and adaptive evolutionary
427 changes to gene expression, and by extension, dGRN functionality.

428 **Distinct *cis*-regulatory modifications differentially effect evolution of gene expression**

429 Lasty, we tested whether changes in nucleotide sequence and accessibility within OCRs
430 are distributed near genes that are differentially expressed between developmental life histories
431 in order to better understand what evolutionary mechanisms may drive evolutionary changes in
432 transcriptional regulation. We found a modest increase in differential expression of genes with
433 nearby differentially accessible OCRs relative to genes lacking differentially accessible OCRs at
434 blastula, gastrula and larval stages (~5-10% increase), but no such relationship exists during
435 early cleavage, which may be expected given the transcriptome is composed almost entirely of
436 maternally-derived transcripts during early cleavage (**fig. 7a**). Interestingly, this relationship
437 between differential expression and *cis*-regulatory modification is magnified when comparing
438 the frequency of signatures of positive selection within OCRs of *He* to OCRs lacking evidence of
439 positive selection (~6-14% increase) and especially to OCRs with evidence of positive selection
440 in *Ht* (15-33% increase) (**fig. 7b**). This finding is concordant with our predictions of evolutionary
441 mechanisms of gene expression change, as selective changes in CRE sequences in *He* most
442 frequently result in concordant divergence in developmental gene expression possibly associated
443 with their derived life history strategy. Positive selection within CRE sequences of *Ht*, on the
444 other hand, are much less often linked to changes in gene expression associated with life history,
445 as *Ht* has retained the ancestral sea urchin developmental program. Thus, while the absolute

446 number of accessibility changes between developmental life histories is typically greater than
447 detected instances of positive selection (**figs. 3-4**), mutations within CRE sequences are more
448 likely to be associated with significant changes in gene regulation and expression relative to a
449 no-change state.

450 This result may be attributable to the distinct characteristics of the two modes of *cis*-
451 regulatory modification. CRE accessibility changes are transient and can be regulated within an
452 organism across developmental stages and tissue types. Thus, relatively frequent changes in
453 chromatin accessibility may be detected within a species at any given time as altered access to
454 CREs may be less consequential for *fixed* expression differences within a species. Though
455 speculative, frequent changes in chromatin accessibility may be an important mechanism for
456 accumulating mutations within regulatory elements, as the "visibility" of mutations to natural
457 selection can fluctuate between cell types, developmental stages, or across generations.
458 Mutations, on the other hand, represent hardwired alterations to the *cis*-regulatory landscape of
459 an individual as heritable sets of sequence changes. Although it may take longer for mutations to
460 accumulate within CREs (and as a result may be relatively less common during divergence
461 between species), their impact on gene expression may be detected more frequently as selection
462 has favored the regulation change imparted by the CRE's altered function. It is important to
463 highlight that our conclusions regarding patterns of gene regulatory evolution are drawn from
464 summarizing thousands of genes, so regulatory evolution and constraints associated with any
465 specific gene will potentially include a unique arrangement and sequence of changes.

466 **Conclusions**

467 In this study, we utilized a recent life history switch in sea urchins to better understand
468 the relationships between evolutionary change in cis-regulatory mechanisms, evolutionary

469 mechanisms, and divergence in embryonic gene expression. We show that two distinct molecular
470 mechanisms within cis-regulatory elements contribute to evolutionary divergence in gene
471 expression; changes in their DNA sequence and changes in the accessibility of those sequences.
472 The former influences transcription factor binding while the latter limits access of transcription
473 factors to DNA and prevent interactions between distal elements and proximal promoters. We
474 document an elevated rate of evolutionary changes in both mechanisms in the derived relative to
475 the ancestral life history, and particularly so near critical developmental network genes whose
476 regulatory connections have been conserved for 10s to 100s million years. These concentrated
477 evolutionary modifications may operate independently or synergistically to have rapidly
478 reshaped transcriptional dynamics during developmental evolution in *H. erythrogramma* through
479 alterations to cell fate specification and ultimately morphogenesis.

480 Collectively, our results indicate that natural selection can efficiently alter the influence
481 of regulatory elements on transcription, through changes in both chromatin accessibility and
482 transcription factor binding. These results provide support for the idea that evolutionary changes
483 in chromatin configuration can contribute to adaptive changes in life history through their
484 influence on developmental gene expression. More generally, the comparative approaches
485 developed in this study present new ways of analyzing multimodal genomic datasets to
486 understand the evolutionary mechanisms that operate on regulatory elements and their
487 consequences for gene expression and organismal traits.

488 **MATERIALS AND METHODS**

489 ***Embryo Rearing and Sample Preparation***

490 *H. erythrogramma* and *H. tuberculata* adult animals were collected near Sydney, NSW,
491 Australia and maintained at 21°C in aquaria with circulating natural seawater. *L. variegatus*

492 individuals were collected near the Duke University Marine Lab in Beaufort, NC, USA and
493 maintained at 23°C in aquaria at Duke University, Durham, NC, USA. Each sea urchin species
494 was induced to spawn via injection of 0.5 M KCl solution into the coelom. Eggs were washed
495 twice in 20 µm filtered seawater (FSW). Sperm was collected dry and stored at 4°C until usage.
496 Washed eggs were fertilized with 1 µL of sperm, and after 10 minutes, were washed three times
497 in FSW. For each species, three unique male-female pairs were crossed to produce three
498 biologically-independent replicates of sea urchin embryos. Embryos were reared at 23°C in large
499 glass culture dishes supplied with FSW, and water was changed every six hours. Because these
500 species exhibit different developmental rates, developmental milestones, rather than absolute
501 developmental time, were selected to maximize developmental synchrony within cultures and
502 between species for comparison. These stages include 1) $4^{th}/5^{th}$ (EC), 2) $6^{th}/7^{th}$ (MC), and 3)
503 $8^{th}/9^{th}$ (LC) cleavage stage embryos, as well as 4) hatched blastula (B), 5) mid-gastrula (G), and
504 6) early larva (L) (two-armed pluteus for *H.* tuberculata and *L. variegatus*; 32 hours post-
505 fertilization for *H. erythrogramma*). Once a culture reached the appropriate stage, a subset of the
506 culture was collected into a 1.5 mL Eppendorf tube for ATAC-seq processing.

507 ***ATAC-seq Protocol and Sequencing***

508 ATAC sample preparation was carried out according to the Omni-ATAC-seq protocol
509 (Corces, et al. 2017). For each replicate, embryos were washed once in 1 µm filtered seawater,
510 lysed, and 50,000 nuclei were isolated for the transposition reaction using the Illumina TDE1
511 enzyme and tagmentation (TD) buffer (Cat. No. 20034197 and 20034198) (San Diego, CA,
512 USA). Sequencing libraries for each replicate were amplified via qPCR as described in
513 (Buenrostro, et al. 2015), and libraries were purified and size selected using Ampure XP Beads at
514 a 1.8:1 bead volume:library volume (Beckman Coulter, Brea, CA, USA). Library quality and

515 transposition efficiency was accessed using a Fragment Analyzer and PROSize 2.0 (Agilent,
516 Santa Clara, CA) at the Duke University Center for Genomic and Computational Biology. *H.*
517 *erythrogramma* and *L. variegatus* libraries were sequenced on an Illumina HiSeq 4000
518 instrument using 50 bp SE sequencing. *H. tuberculata* libraries were sequenced on an Illumina
519 NovaSeq 6000 instrument using 50 bp PE sequencing (only SE were used for data analysis).

520 ***ATAC-seq Data Analysis***

521 Raw ATAC-seq reads were trimmed for quality and sequencing adapters using
522 "cutadapt" (Martin 2011) v2.3 with the following parameters: -a CTGTCTCTTATACACATCT
523 -q 20 --trim-n -m 40. Trimmed reads were then aligned to each species' respective genome via
524 "stampy" (Lunter and Goodson 2011) v.1.0.28 using the "—sensitive" set of parameters. ATAC-
525 seq alignments were filtered for mitochondrial sequences and required an alignment quality score
526 of at least 5 using "samtools" (Li, et al. 2009) v.1.9. Filtered alignments are available in BAM
527 and Bedgraph format (see "Data Availability" below).

528 In this study, we aimed to compare evolution of orthologous regulatory elements so to
529 accomplish this, we performed a series of liftOvers (Hinrichs, et al. 2006) to convert ATAC-seq
530 alignments between genomic coordinates of each sea urchin species. We took an iterative,
531 reciprocal strategy described below to minimize possible reference bias associated with
532 converting between genome assemblies: 1) *H. erythrogramma*: *He* → *Lv* → *He*; 2) *H.*
533 *tuberculata*: *Ht* → *Lv* → *He*; 3) *L. variegatus*: *Lv* → *He* → *Lv* → *He*. At the end of filtering and
534 coordinate conversion, all ATAC-seq alignments were referenced to the *H. erythrogramma*
535 genome. LiftOver files are available in (Davidson, et al. 2021).

536 ATAC-seq peaks were called from these filtered, converted alignments using the
537 "MACS2" v2.1.2 (Zhang, et al. 2008) callpeak function (parameters: –nomodel, –keep-

538 dup=auto, –shift 100, –extsize 200) for each developmental stage separately at an FDR threshold
539 of 5%. Peak coordinates were merged using the "bedtools" (Quinlan and Hall 2010) v2.25 *merge*
540 function requiring a peak overlap of at least 200 bp to be merged into a single peak. Lastly, for
541 each sample, accessibility of each peak was measured with the "bedtools" *multiBamCov* function
542 and imported into R v4.0.2 for statistical analysis (Data S1).

543 ***Estimates of positive selection***

544 To measure branch-specific signatures of positive selection in the non-coding genome,
545 the "adaptiPhy" (Berrio, et al. 2020) pipeline (https://github.com/wodanaz/adaptiPhy) for global
546 tests of natural selection was followed. First, orthologous sequences for non-coding sites of
547 interest (ATAC-seq peaks) were extracted from each species' genome into FASTA format.
548 Sequences were trimmed to include only contiguous DNA sequence using the *prunning.py* script
549 and filtered using the *filtering.py* script, requiring a minimum alignment length of 75 base pairs.
550 These trimmed and filtered alignments serve as "query" sequences of tests for selection. To
551 generate a neutral reference for comparison, ten orthologous, neutral sites were randomly
552 selected from a curated set of ~88,000 putatively neutral sequences (Davidson, et al. 2021) (Data
553 S2) and concatenated into a single neutral reference sequence. Then, substitution rates of both
554 the query and randomly concatenated neutral reference were estimated using phyloFit (Siepel
555 and Haussler 2004), and the zeta score was calculated as the ratio of the query substitution rate to
556 the neutral reference substitution rate. In addition, p-values of likelihood ratio tests for
557 significant levels of branch-specific positive selection were calculated with adaptiPhy (Berrio, et
558 al. 2020) pipeline using HyPhy (Pond, et al. 2005). For each query site, tests for positive
559 selection were repeated 10 times against a unique neutral reference to ensure reproducible
560 estimates of zeta and significance values for each non-coding site of interest. P-values and

561 substitution rates for all query and neutral sites were then imported in R v4.0.2. For each query
562 site, the median zeta score and p-value for each test of positive selection across the ten replicate
563 calculations were used for downstream analyses.

564 ***OCR Filtering***

565 After accessibility and estimates of positive selection were calculated for each ATAC-seq
566 peak, herein referred to as an "open chromatin region" (OCR), a series of filtering and quality
567 control metrics were carried out to ensure only high confidence and quality peaks were compared
568 between species. These filtering steps are as follows: 1) each OCR is required to have at least
569 75bp of contiguous, single copy sequence (see previous section) for accurate estimations of
570 selection ; 2) a local composition complexity (LCC) (Konopka 2001) value of 1.9 or more was
571 required of the OCR for each species to remove repetitive or other low-complexity sequences
572 that may generate inaccurate estimations of selection (module:
573 biopython.org/docs/1.75/api/Bio.SeqUtils.lcc.html); 3) a CPM of 3 or more was required in at
574 least 9 samples to remove OCRs with extremely low accessibility across samples; 4) the
575 midpoint of the OCR must lie within 25 Kb (in either direction) of the translational start site of a
576 gene model. The 25 Kb window for assigning OCRs to genes was selected on the basis that the
577 median and mean intergenic distance within the *He* genome are 21.4 Kb and 31.9 Kb,
578 respectively. Given nearly no prior knowledge is known of the *cis*-regulatory landscape for these
579 sea urchin species, these stringent filtering methods were carried out in order to maximize
580 confidence in comparisons of non-coding sequence evolution and function. This resulted in a
581 final set 35,788 high-confidence OCRs for cross-species analysis (Data S3).

582 ***ATAC-seq Statistical Analysis***

583 Principal component analysis and variance partitioning was performed on the regularized
584 log-transformed (rlog) data (Love, et al. 2014) of filtered OCR accessibility counts via the
585 *prcomp* function and the "variancePartition" v1.18.3 package (Hoffman and Schadt 2016) in R
586 using the formula: ~ (1 | Stage) + (1 | Species) + (1 | Life History) + (1 | Genus) + (1 | Stage: Life
587 History). Raw accessibility counts of the filtered OCRs were loaded into DESeq2 (Love, et al.
588 2014) v1.30 to calculate differential accessibility between sample groups. Differentially
589 accessible sites between life history strategies were classified as having a 2-fold accessibility
590 difference and supported by an FDR of 10% in *He* vs *Ht* and well as *He* vs *Lv* comparisons
591 (including the sign of the accessibility difference being in the same direction) (see
592 **supplementary fig. 7** for all significance cutoff combinations of differential accessibility
593 results). For graphical purposes, the average fold-change in the *He* vs *Ht* and *He* vs *Lv*
594 comparisons is displayed within figures. Significant levels of positive selection in either the *He*
595 or *Ht* branch were classified as OCRs having a zeta value greater than 1.25 and supported by a
596 median false-discovery rate less than 10% in one species, and failing to meet both criteria in the
597 other. Motif enrichment analyses were carried out with "Homer" v4.11 (Heinz, et al. 2010) via
598 the findMotifsGenome.pl script using all OCRs as the background set and regions size of 200 bp.
599 Sea urchin GRN genes are provided in Data S4

600 ***RNA-seq Expression Analyses***

601 Raw RNA-seq reads from eggs, 4-cell, 16-cell + 32-cell (collectively EE) , blastula (B),
602 gastrula (G), and early larva (L) stage embryos of *He*, *Ht*, and *Lv* were retrieved from (Israel, et
603 al. 2016), trimmed and filtered for low quality bases and reads with Trimmomatic (Bolger, et al.
604 2014), and aligned to each species respective genomes and gene models with STAR (Dobin, et
605 al. 2013). From these alignments, mRNA expression was estimated with Salmon (Patro, et al.

606 2017) and loaded to R for statistical analysis. Read counts for summed to each gene's best match
607 to the *S. purpuratus* v4.2 gene models to generate a common reference for expression
608 comparisons between species as described in (Israel, et al. 2016) (Data S5). Differentially
609 expressed genes between life histories were called as having a fold-change in expression > 2 and
610 supported by an FDR of 10% or less in *He* vs *Ht* and well as *He* vs *Lv* comparisons (including
611 the sign of the accessibility difference being in the same direction) (Data S6). These criteria
612 were chosen to make RNA and ATAC-seq analyses between species comparable as possible.

613 ***Principal Component Distances of Developmental Profile***

614 To identify OCRs and genes whose developmental regulatory profiles diverged in *H.*
615 *erythrogramma* or *H. tuberculata*, we calculated principal component distances of chromatin
616 accessibility and gene expression profile shapes between species throughout development as
617 described in (Israel, et al. 2016). Mean accessibility and expression profiles of regularized log-
618 transformed (rlog) OCR and genes counts were standardized using the *standardize* function of
619 the "mfuzz" v2.48.0 package (Kumar and M 2007). From these data, scaled distance measures
620 ("jump scores") between orthologous non-coding sites or genes were calculated using loadings
621 of the first two principal components (Data S7): $d_{HeLv} =$
622 $\sqrt{PC1(x_{Lv} - x_{He})^2 + PC2(y_{Lv} - y_{He})^2}$;
623 $d_{HtLv} = \sqrt{PC1(x_{Lv} - x_{Ht})^2 + PC2(y_{Lv} - y_{Ht})^2}$. A significant change in the developmental
624 trajectory of gene expression or OCR accessibility was determined to have a jump score
625 difference of at 0.15 on one branch over the other .

626 **DATA AVAILIBILITY**

627 Raw ATAC-seq reads are available in FASTQ format on NCBI's Short Read Archive (accession
628 number pending publication). Filtered alignments of ATAC-seq reads before and after between-

species liftOver files are available on Dryad (accession number pending publication). Results files of ATAC-seq analyses are available as Supplementary Data files.

AUTHOR CONTRIBUTIONS

PLD and GAW conceived and designed the study. PLD and MB performed sample collection. PLD performed the data analysis. PLD and GAW wrote the manuscript. All authors revised the manuscript.

ACKNOWLEDGEMENTS

We would like to thank the Sydney Institute of Marine Sciences and their staff for assistance with animal care, laboratory set up, and carrying out experiments. We also thank Allison Edgar and Hannah Devens for assistance preparing sea urchin embryo cultures associated with this work. This work was supported by the National Science Foundation Division of Integrative Organismal Systems (award no. 1929934 to GAW).

COMPETING INTERESTS

The authors have no competing interests to declare.

REFERENCES

Berrio A, Haygood R, Wray GA. 2020. Identifying branch-specific positive selection throughout the regulatory genome using an appropriate proxy neutral. BMC Genomics 21:359.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-2120.

Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. 2015. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr Protoc Mol Biol 109:21 29 21-29.

Byrne M, Sewell MA. 2019. Evolution of maternal lipid provisioning strategies in echinoids with non-feeding larvae: selection for high-quality juveniles. Marine Ecology Progress Series 616:95-106.

Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, Satpathy AT, Rubin AJ, Montine KS, Wu B, et al. 2017. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods 14:959-962.

Cusanovich DA, Reddington JP, Garfield DA, Daza RM, Aghamirzaie D, Marco-Ferreres R, Pliner HA, Christiansen L, Qiu X, Steemers FJ, et al. 2018. The cis-regulatory dynamics of embryonic development at single-cell resolution. Nature 555:538-542.

Davidson EH, Erwin DH. 2006. Gene regulatory networks and the evolution of animal body plans. Science 311:796-800.

Davidson PL, Guo H, Swart JS, Massri AJ, Edgar A, Wang L, Berrio A, Devens HR, Koop D, Cisternas P, et al. 2021. Recent reconfiguration of an ancient developmental gene regulatory network in *Heliocidaris* sea urchins.

Davidson PL, Guo H, Wang L, Berrio A, Zhang H, Chang Y, Soborowski AL, McClay DR, Fan G, Wray GA. 2020. Chromosomal-level genome assembly of the sea urchin *Lytechinus variegatus* substantially improves functional genomic analyses. Genome Biology and Evolution 12:1080-1086.

Davidson PL, Thompson JW, Foster MW, Moseley MA, Byrne M, Wray GA. 2019. A comparative analysis of egg provisioning using mass spectrometry during rapid life history evolution in sea urchins. Evolution & Development 21:188-204.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15-21.

Emlet RB. 1995. Larval spicules, cilia, and symmetry as remnants of indirect development in the direct developing sea urchin *Heliocidaris erythrogramma*. Dev Biol 167:405-415.

Falkner I, Sewell MA, Byrne M. 2015. Evolution of maternal provisioning in ophiuroid echinoderms: characterisation of egg composition in planktotrophic and lecithotrophic developers. Marine Ecology Progress Series 525:1-13.

Flatt T, Heyland A. 2011. Mechanisms of life history evolution: the genetics and physiology of life history traits and trade-offs: Oxford University Press.

Hart MW. 2002. Life history evolution and comparative developmental biology of echinoderms. Evolution & Development 4:62-71.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38:576-589.

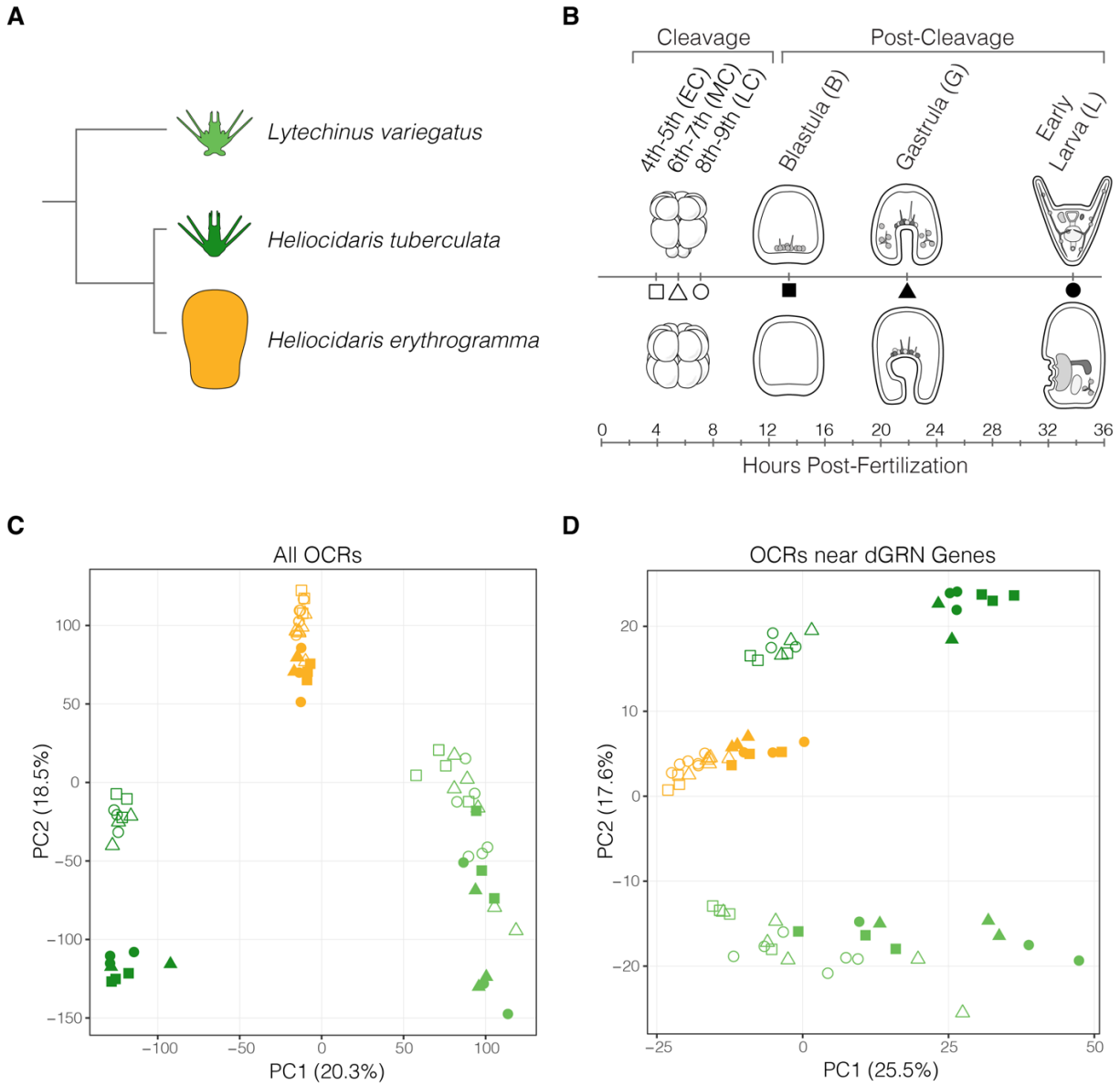
Henry JJ, Wray GA, Raff RA. 1990. The dorsoventral axis is specified prior to 1st cleavage in the direct developing sea urchin *Heliocidaris erythrogramma*. Development 110:875-

690 884.
691 Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, Diekhans M, Furey
692 TS, Harte RA, Hsu F. 2006. The UCSC genome browser database: update 2006. Nucleic
693 acids research 34:D590-D598.
694 Hoffman GE, Schadt EE. 2016. variancePartition: interpreting drivers of variation in complex
695 gene expression studies. BMC Bioinformatics 17:483.
696 Israel JW, Martik ML, Byrne M, Raff EC, Raff RA, McClay DR, Wray GA. 2016. Comparative
697 developmental transcriptomics reveals rewiring of a highly conserved gene
698 regulatory network during a major life history switch in the sea urchin genus
699 *Heliocidaris*. PLoS Biol 14:e1002391.
700 Jhanwar S, Malkmus J, Stolte J, Romashkina O, Zuniga A, Zeller R. 2021. Conserved and
701 species-specific chromatin remodeling and regulatory dynamics during mouse and
702 chicken limb bud development. Nat Commun 12:5685.
703 Klueg KM, Harkey MA, Raff RA. 1997. Mechanisms of evolutionary changes in timing, spatial
704 expression, and mRNA processing in the msp130 gene in a direct-developing sea urchin,
705 *Heliocidaris erythrogramma*. Developmental Biology 182:121-133.
706 Konopka AK. 2001. Sequence complexity and composition. e LS.
707 Koop D, Cisternas P, Morris VB, Strbenac D, Yang JY, Wray GA, Byrne M. 2017. Nodal and
708 BMP expression during the transition to pentamery in the sea urchin *Heliocidaris*
709 *erythrogramma*: insights into patterning the enigmatic echinoderm body plan. BMC Dev
710 Biol 17:4.
711 Kumar L, M EF. 2007. Mfuzz: a software package for soft clustering of microarray data.
712 Bioinformation 2:5-7.
713 Lewis JJ, Reed RD. 2019. Genome-wide regulatory adaptation shapes population-level
714 genomic landscapes in *Heliconius*. Mol Biol Evol 36:159-173.
715 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R,
716 Genome Project Data Processing S. 2009. The sequence alignment/map format and
717 SAMtools. Bioinformatics 25:2078-2079.
718 Livraghi L, Hanly JJ, Van Bellghem SM, Montejo-Kovacevich G, van der Heijden ES, Loh LS,
719 Ren A, Warren IA, Lewis JJ, Concha C, et al. 2021. Cortex cis-regulatory switches
720 establish scale colour identity and pattern diversity in *Heliconius*. eLife 10.
721 Love AC, Raff RA. 2006. Larval ectoderm, organizational homology, and the origins of
722 evolutionary novelty. Journal of Experimental Zoology Part B: Molecular and
723 Developmental Evolution 306:18-34.
724 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for
725 RNA-seq data with DESeq2. Genome Biol 15:550.
726 Lunter G, Goodson M. 2011. Stampy: a statistical algorithm for sensitive and fast mapping of
727 Illumina sequence reads. Genome research 21:936-939.
728 Marshall DJ, Krug PJ, Kupriyanova EK, Byrne M, Emlet RB. 2012. The biogeography of
729 marine invertebrate life histories. Annual Review of Ecology, Evolution, and
730 Systematics, Vol 43 43:97-+.
731 Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads.
732 EMBnet. journal 17:10-12.
733 McClay DR, Miranda E, Feinberg SL. 2018. Neurogenesis in the sea urchin embryo is initiated
734 uniquely in three domains. Development 145.
735 McEdward LR, Miner BG. 2001. Larval and life-cycle patterns in echinoderms. Canadian

Journal of Zoology 79:1125-1170.
Morgan SG. 2020. Life and death in the plankton: larval mortality and adaptation. Ecology of marine invertebrate larvae: 279-321.
Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 14:417-419.
Pechenik JA. 1999. On the advantages and disadvantages of larval stages in benthic marine invertebrate life cycles. Marine Ecology Progress Series 177:269-297.
Pond SL, Frost SD, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics 21:676-679.
Prowse TAA, Falkner I, Sewell MA, Byrne M. 2009. Long-term storage lipids and developmental evolution in echinoderms. Evolutionary Ecology Research 11:1069-1083.
Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841-842.
Raff RA. 1987. Constraint, flexibility, and phylogenetic history in the evolution of direct development in sea urchins. Developmental Biology 119:6-19.
Raff RA, Byrne M. 2006. The active evolutionary lives of echinoderm larvae. Heredity 97:244-252.
Sackton TB, Grayson P, Cloutier A, Hu Z, Liu JS, Wheeler NE, Gardner PP, Clarke JA, Baker AJ, Clamp M, et al. 2019. Convergent regulatory evolution and loss of flight in paleognathous birds. Science 364:74-78.
Shibata Y, Sheffield NC, Fedrigo O, Babbitt CC, Wortham M, Tewari AK, London D, Song L, Lee B-K, Iyer VR, et al. 2012. Extensive evolutionary changes in regulatory element activity during human origins are associated with altered gene expression and positive selection. PLOS Genetics 8:e1002789.
Siepel A, Haussler D. 2004. Phylogenetic estimation of context-dependent substitution rates by maximum likelihood. Mol Biol Evol 21:468-488.
Smith MS, Turner FR, Raff RA. 2008. Nodal expression and heterochrony in the evolution of dorsal-ventral and left-right axes formation in the direct-developing sea urchin *Heliocidaris erythrogramma*. Journal of Experimental Zoology Part B-Molecular and Developmental Evolution 310b:609-622.
Strathmann RR. 1978. The evolution and loss of feeding larval stages of marine invertebrates. Evolution 32:894-906.
Strathmann RR. 1985. Feeding and nonfeeding larval development and life history evolution in marine invertebrates. Annual Review of Ecology and Systematics 16:339-361.
Sun DA, Patel NH. 2021. Identification and classification of cis-regulatory elements in the amphipod crustacean *Parhyale hawaiensis*. bioRxiv:2021.2009.2016.460328.
Swain-Lenz D, Berrio A, Safi A, Crawford GE, Wray GA. 2019. Comparative analyses of chromatin landscape in white adipose tissue suggest humans may have less beigeing potential than other primates. Genome Biology and Evolution 11:1997-2008.
Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B, et al. 2012. The accessible chromatin landscape of the human genome. Nature 489:75-82.
Tonelli C, Chio IIC, Tuveson DA. 2018. Transcriptional Regulation by Nrf2. Antioxid Redox Signal 29:1727-1745.
Tsushima M, Byrne M, Amemiya S, Matsuno T. 1995. Comparative biochemical studies of carotenoids in sea urchins. 3. Relationship between developmental mode and

782 carotenoids in the Australian echinoids *Heliocidaris erythrogamma* and *H. tuberculata*
783 and a comparison with japanese species. Comparative Biochemistry and Physiology B-
784 Biochemistry & Molecular Biology 110:719-723.
785 van der Burg KRL, Lewis JJ, Brack BJ, Fandino RA, Mazo-Vargas A, Reed RD. 2020. Genomic
786 architecture of a genetically assimilated seasonal color pattern. Science 370:721-725.
787 Vance RR. 1973. On the reproductive strategies of marine benthic invertebrates. The American
788 Naturalist 107:339-352.
789 Walton KD, Warner J, Hertzler PH, McClay DR. 2009. Hedgehog signaling patterns mesoderm
790 in the sea urchin. Dev Biol 331:26-37.
791 Wang L, Israel JW, Edgar A, Raff RA, Raff EC, Byrne M, Wray GA. 2020. Genetic basis for
792 divergence in developmental gene expression in two closely related sea urchins. Nat Ecol
793 Evol 4:831-840.
794 Williams DHC, Anderson DT. 1975. Reproductive system, embryonic development, larval
795 development and metamorphosis of sea urchin H*eliocidaris erythrogramma* (Val)
796 (Echinoidea-Echinometridae). Australian Journal of Zoology 23:371-403.
797 Wilson KA, Andrews ME, Raff RA. 2005. Dissociation of expression patterns of homeodomain
798 transcription factors in the evolution of developmental mode in the sea urchins
799 *Heliocidaris tuberculata* and *H. erythrogramma*. Evolution & Development 7:401-415.
800 Wilson KA, Andrews ME, Rudolf Turner F, Raff RA. 2005. Major regulatory factors in the
801 evolution of development: the roles of goosecoid and Msx in the evolution of the direct-
802 developing sea urchin *Heliocidaris erythrogramma*. Evolution & Development 7:416-
803 428.
804 Wittkopp PJ, Kalay G. 2012. Cis-regulatory elements: molecular mechanisms and evolutionary
805 processes underlying divergence. Nature Reviews Genetics 13:59-69.
806 Wray GA. 2007. The evolutionary significance of cis-regulatory mutations. Nat Rev Genet
807 8:206-216.
808 Wray GA. 2021. Extreme phenotypic divergence and the evolution of development. In. Current
809 Topics in Developmental Biology.
810 Wray GA. 1996. Parallel evolution of nonfeeding larvae in echinoids. Systematic Biology
811 45:308-322.
812 Wray GA, Raff RA. 1989. Evolutionary modification of cell lineage in the direct-developing
813 sea urchin *Heliocidaris erythrogramma*. Developmental Biology 132:458-470.
814 Wray GA, Raff RA. 1990. Novel origins of lineage founder cells in the direct-developing sea
815 urchin *Heliocidaris erythrogramma*. Dev Biol 141:41-54.
816 Yuh CH, Dorman ER, Davidson EH. 2005. Brn1/2/4, the predicted midgut regulator of the
817 endo16 gene of the sea urchin embryo. Dev Biol 281:286-298.
818 Zaret KS, Carroll JS. 2011. Pioneer transcription factors: establishing competence for gene
819 expression. Genes Dev 25:2227-2241.
820 Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM,
821 Brown M, Li W. 2008. Model-based analysis of ChIP-Seq (MACS). Genome biology
822 9:R137.
823

824 **FIGURE and LEGENDS**



825

826 **Figure 1: Global variation in chromatin accessibility between sea urchin embryos is largely**

827 **driven by developmental stage and life history strategy. A)** Three focal sea urchin species of

828 this study, including two planktotrophic species representing the ancestral life history (light

829 green: *Lytechinus variegatus*; dark green: *Heliocidaris tuberculata*) and one lecithotrophic

830 species representing the derived life history (orange: *Heliocidaris erythrogramma*). This color

831 scheme is consistent throughout the study. **B)** Developmental stages assayed in this study for

832 each species. The ancestral life history is illustrated above, with corresponding developmental
833 stages of the derived life history below. Developmental times are shown for *Heliocidaris*
834 *tuberculata* at 21°C. Note, however, samples were collected on the basis of developmental stage,
835 not hours post-fertilization, for all three species. **C-D)** Principal component analysis of variation
836 in normalized chromatin accessibility among **C)** all orthologous OCRs and **D)** only OCRs near
837 dGRN genes through embryogenesis. OCR: open chromatin region; dGRN: developmental gene
838 regulatory network.

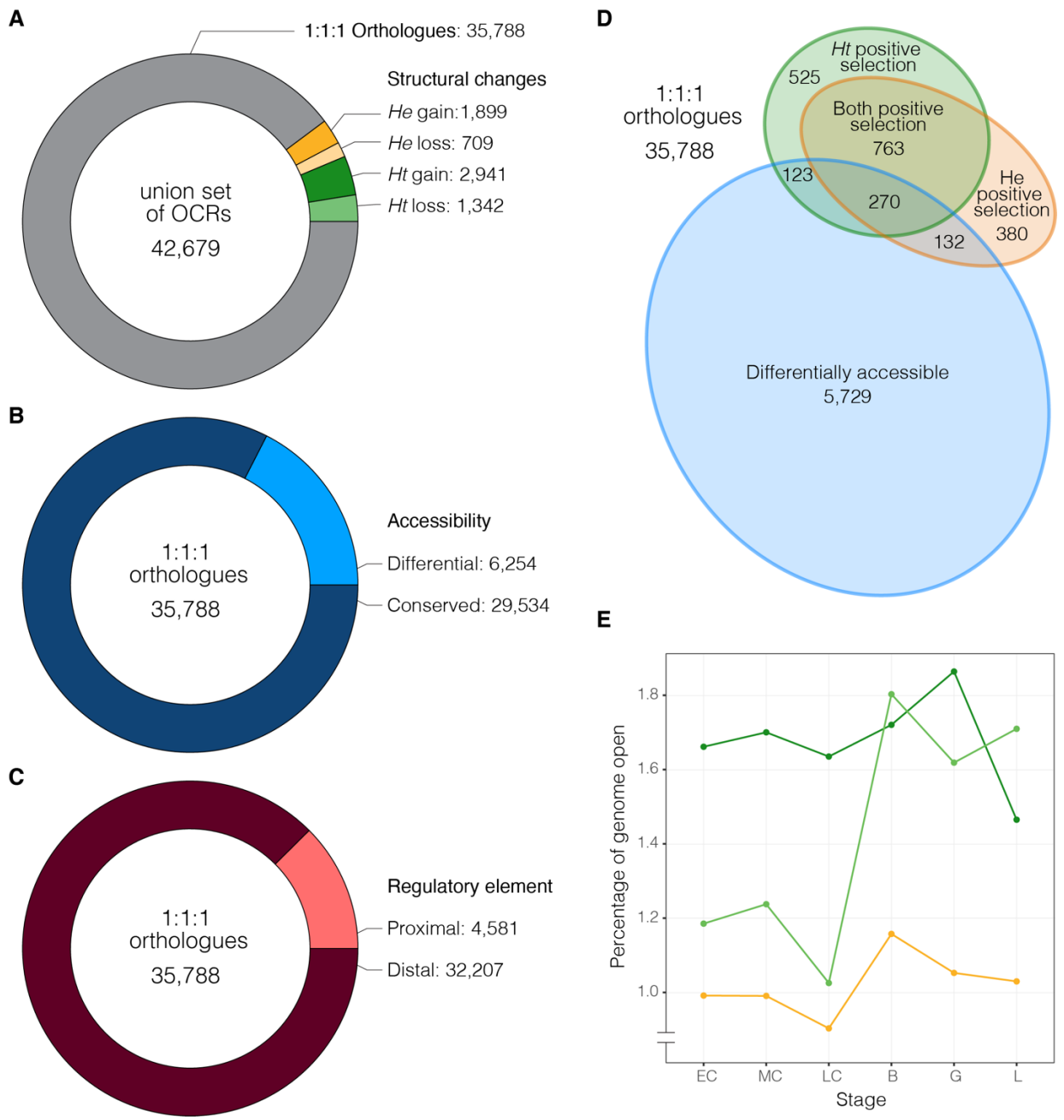


Figure 2: Summary statistics of regulatory element gains, losses, accessibility, and positive selection. A) Union set of all OCRs, including structural differences between species such as lineage-specific gains and losses as well as 35,788 1:1:1 orthologous OCRs shared between all three species. These orthologous OCRs are the primary focus of this study. Breakdown of **B)** OCR differentially accessibility between life history strategies and **C)** OCR distance to gene

845 models for these orthologous sites. **D)** Overlap of OCR differential accessibility and evidence of
846 positive selection on the *He*, *Ht*, or both branches. **E)** Percentage of each species genome that is
847 "open" at each stage of development. "Open" is defined as the cumulative portion of the genome
848 represented by all OCRs (orthologous and lineage-specific) with significant levels of
849 accessibility for a given stage and species (see **supplementary fig. 3** for additional information).
850 OCR: open chromatin region; *He*: *Heliocidaris erythrogramma*; *Ht*: *Heliocidaris tuberculata*.

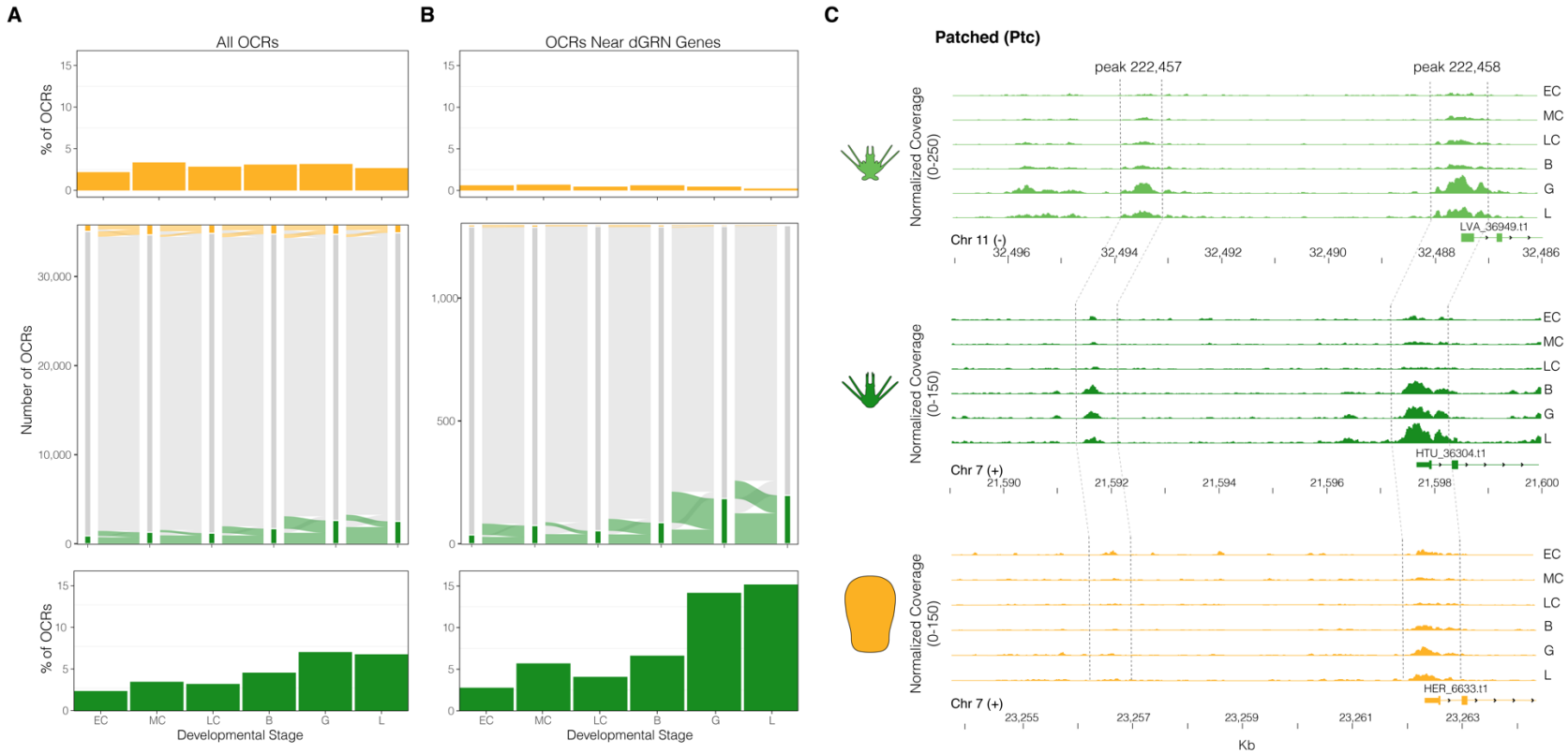


Figure 3: The chromatin landscape of the derived life history is far less accessible than that of species with the ancestral life history mode, particularly near core dGRN genes. A-B) Number and proportions of differentially accessible OCRs that are significantly more open in *H. erythrogramma* (orange, top) or species with the ancestral life history strategy (green, bottom) for **A)** all orthologous OCRs and **B)** only OCRs near dGRN genes. **C)** Example of reduced chromatin accessibility of a proximal and distal OCR in *H. erythrogramma* during embryogenesis near the gene *patched*. OCR: open chromatin region; dGRN: developmental gene regulatory network.

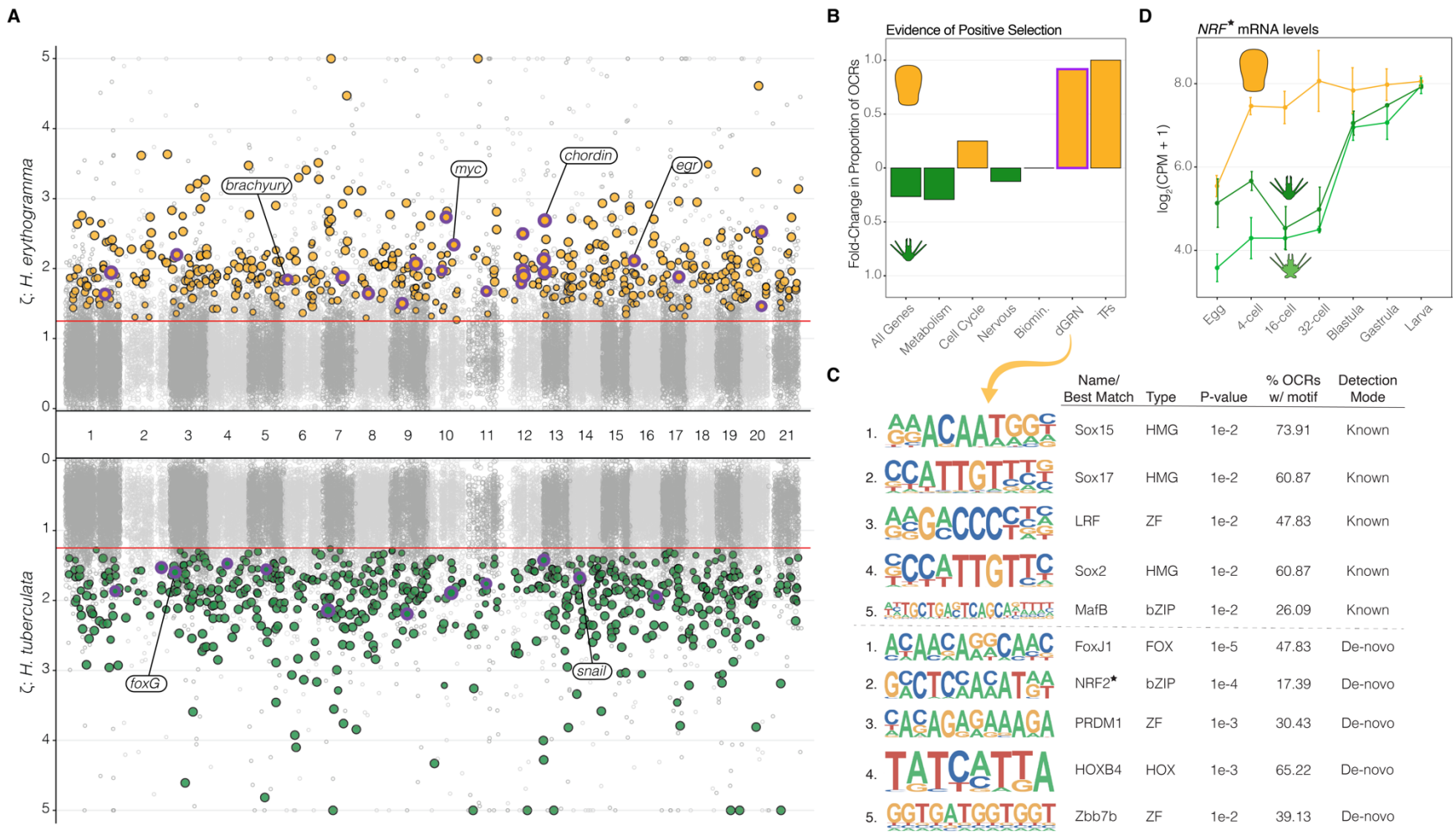
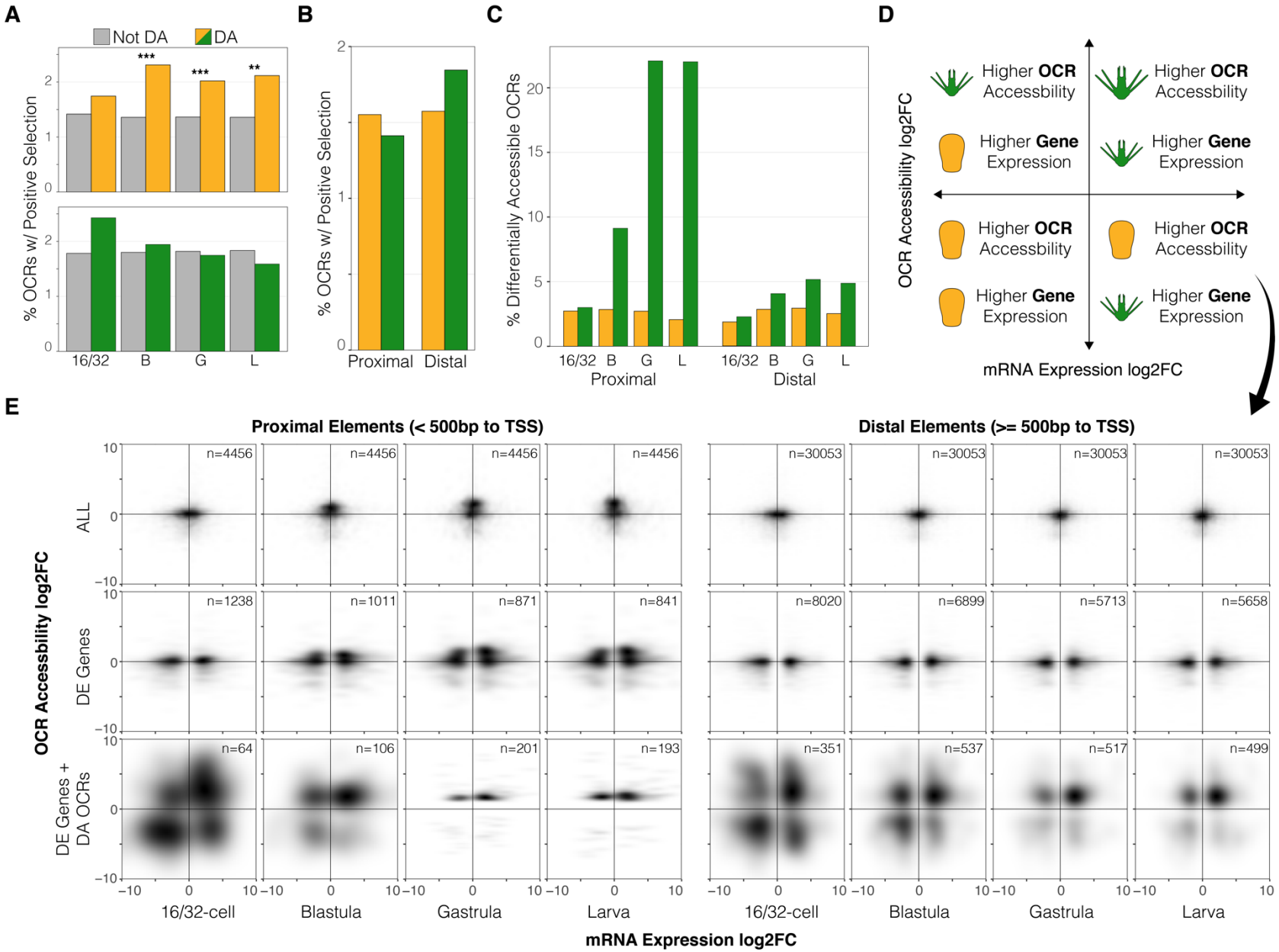


Figure 4: Sequence changes indicative of positive selection are enriched within *cis*-regulatory elements near dGRN genes and other transcription factors in the *H. erythrogramma* genome. A) Zeta values (ratio of OCR substitution rate to that of neutrally evolving reference sequences) for *H. erythrogramma* (top) and *H. tuberculata* (bottom) across the genome. OCRs with significant

862 support for lineage-specific evidence of positive selection are noted for *H. erythrogramma* (orange) and *H. tuberculata* (dark green),
863 and those nearby dGRN genes are outlined in purple. **B)** Fold-change in proportion of OCRs with more frequent evidence of positive
864 selection in either *H. erythrogramma* or *H. tuberculata*, partitioned by functional categories of nearby genes (see **supplementary fig.**
865 **5** for all categories). **C)** Transcription factor binding motif enrichment results of OCRs near dGRN genes with evidence of positive
866 selection in *H. erythrogramma*. Detection mode indicates OCR sequences were searched against a (i) "known" set of optimized motif
867 models or (ii) enriched motifs within the OCR were identified "de-novo" and matched to a known set of motifs. **D)** Embryonic mRNA
868 expression levels for *Nuclear factor erythroid 2-related factor* (*NRF*) for each species. OCR: open chromatin region; dGRN:
869 developmental gene regulatory network.


A
Not DA
DA
% OCRs w/ Positive Selection
16/32
B
G
L
B
% OCRs w/ Positive Selection
Proximal
Distal
C
% Differentially Accessible OCRs
16/32 B G L
Proximal
16/32 B G L
Distal
D
Higher OCR Accessbility
Higher Gene Expression
OCR Accessibility log2FC
mRNA Expression log2FC
E
Proximal Elements (< 500bp to TSS)
Distal Elements (>= 500bp to TSS)
ALL
DE Genes
DE Genes + DA OCRs
OCR Accessbility log2FC
n=4456
n=30053
n=1238
n=1011
n=871
n=841
n=8020
n=6899
n=5713
n=5658
n=64
n=106
n=201
n=193
n=351
n=537
n=517
n=499
16/32-cell
Blastula
Gastrula
Larva
mRNA Expression log2FC

870

871 **Figure 5: Independent and synergistic changes to regulatory element function have evolved during life history evolution in *H.***
872 ***erythrogramma*. A)** Proportion of differentially accessible or non-differentially accessible OCRs with evidence of positive selection
873 on the *H. erythrogramma* branch (orange, top) or *H. tuberculata* branch (dark green, bottom). Two-sided, Fisher's exact test: ** $p \leq 5$
874 $x\ 10^{-3}$; *** $p \leq 5 \times 10^{-4}$. **B)** Proportion of proximal of distal OCRs with evidence of positive selection in either *Heliocidaris* species. **C)**
875 Differential accessibility of proximal and distal OCRs at four stages of development. Orange indicates significantly greater
876 accessibility in *H. erythrogramma*, dark green indicates significantly greater accessibility in *H. tuberculata*. **D-E)** Fold-change of
877 mRNA expression vs nearby OCR accessibility between life history strategies for proximal (left) and distal (right) OCRs.
878 Relationships are shown for all genes and OCRs (E: top row), significantly differentially expressed genes (E: middle row), and
879 significantly differentially expressed genes *and* differentially accessible OCRs (E: bottom row). OCR: open chromatin region; DE:
880 differentially expressed (genes); DA: differentially accessible (OCRs).

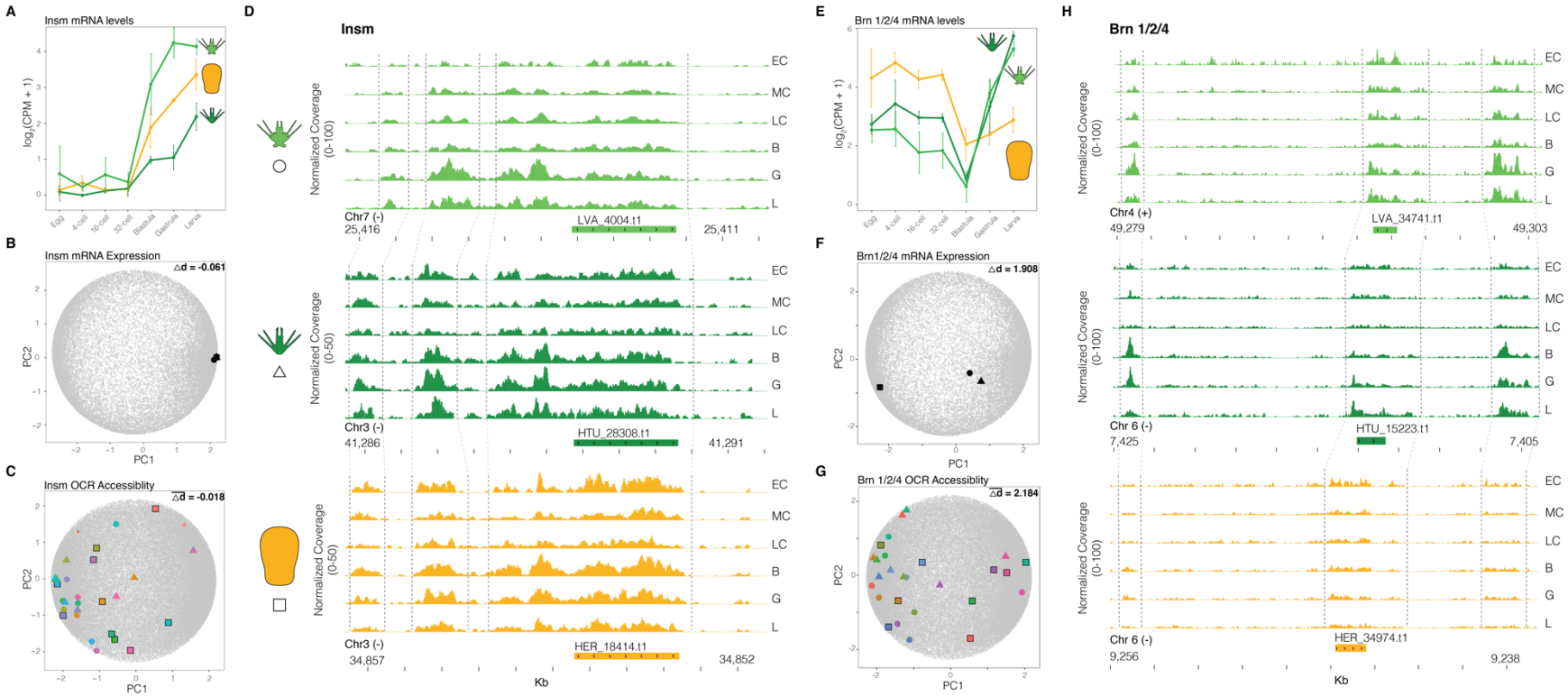


Figure 6: Quantifying developmental expression and accessibility changes in principal component space predicts divergent patterns of gene regulation. A-D) An example of conserved developmental gene expression and chromatin accessibility for *insm*. **A)** Developmental mRNA expression profile of *insm* for each sea urchin species and **B)** their coordinates in PCA space. **C)** Coordinates in PCA space of developmental accessibility of orthologous OCRs near *insm* and **D)** normalized chromatin accessibility coverage of these same OCRs through embryogenesis for each species. **E-H)** An example of divergent developmental gene expression and chromatin accessibility associated with life history for *brn 1/2/4*. **E)** Developmental mRNA expression profile of *brn 1/2/4* for each

888 sea urchin species and **F)** their coordinates in PCA space. **G)** Coordinates in PCA space of developmental accessibility of orthologous

889 OCRs near *brn 1/2/4* and **H)** normalized chromatin accessibility coverage of these same OCRs through embryogenesis for each

890 species. OCR: open chromatin region; PCA: principal component analysis.

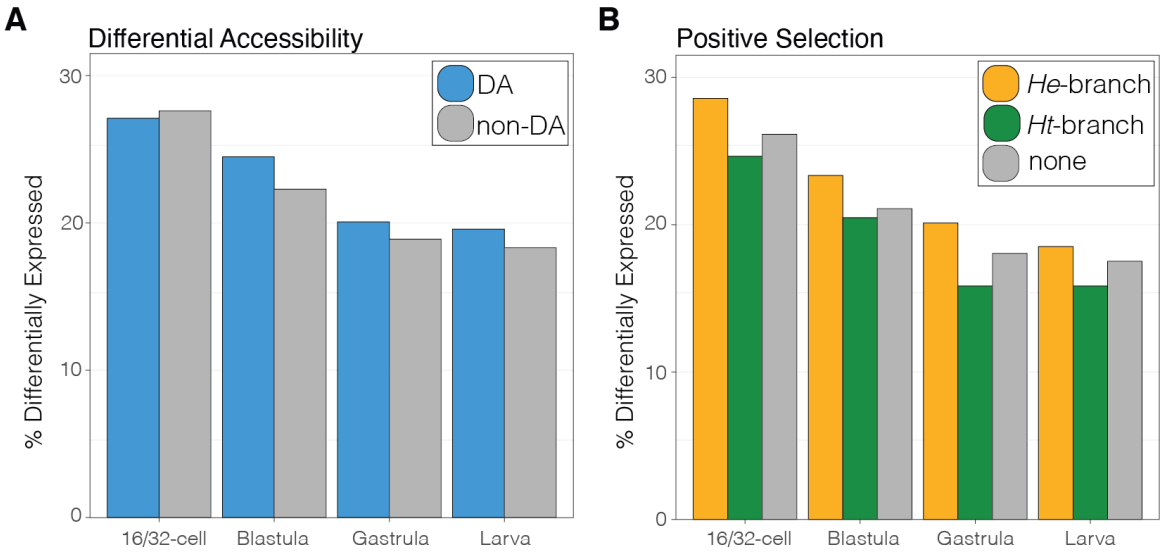


Figure 7: Evolutionary changes to *cis*-regulatory element function are frequently associated with expression changes of nearby genes underlying life history evolution in sea urchins. Percentage of genes that are differentially expressed between developmental life history strategies at four developmental stages, distinguished as having at least one nearby OCR with **A)** differential accessibility or **B)** evidence of lineage-specific positive selection. OCR: open chromatin region.