Evolutionary changes in the chromatin landscape contribute to reorganization of a developmental gene network during rapid life history evolution 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 role 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 the function of putative *cis*-regulatory elements 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.

INTRODUCTION

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

Among the most common and well-studied life history shifts in the animal kingdom is the evolution of feeding or non-feeding larval development (Flatt and Heyland 2011). In marine invertebrates, planktotrophy (feeding larval development) is characterized by species that produce many, relatively small eggs and have a longer planktonic larval period, whereas lecithotrophic species (non-feeding larval development) invest more maternal resources into a

65 fewer number of large eggs that develop through an abbreviated larval stage (Strathmann 1985).
66 Importantly, the ecological circumstances associated with each strategy are predictable and well-
67 studied across diverse marine clades (Vance 1973; Strathmann 1985; Pechenik 1999; Marshall,
68 et al. 2012; Morgan 2020), providing natural systems to analyze how selection shapes often rapid
69 phenotypic divergence in developmental regulation and physiology. For example, within
70 echinoderms, lecithotrophy has independently evolved from the ancestral planktotrophic state
71 more than 20 times (Raff 1987; Wray 1996; McEdward and Miner 2001; Hart 2002), implying
72 that recurring sets of environmental conditions (e.g. limited availability of phytoplankton) and
73 selection can enact evolutionary modifications to even highly conserved developmental
74 processes. The consequences include changes throughout development, ranging from egg
75 composition (Prowse, et al. 2009; Falkner, et al. 2015; Byrne and Sewell 2019; Davidson, et al.
76 2019) and zygotic genome activation (Wray and Raff 1990; Israel, et al. 2016; Wang, et al.
77 2020), cell type specification and embryonic patterning (Wray and Raff 1989; Love and Raff
78 2006), to construction of larval structures such as the skeleton and imaginal adult rudiment
79 (Emlet 1995; Koop, et al. 2017). Still, the impact selection exerts on the evolution of *cis*-
80 regulatory interactions underlying alternative developmental programs and contrasting life
81 history strategies is not known. This gap in knowledge is manifested more broadly in our current
82 sparse understanding of the biological relationships governing regulatory element function,
83 evolutionary constraints inherent to development and gene network architecture, and phenotypic
84 diversification.
85 Perhaps the best studied example of the life history switch from planktotrophic to
86 lecithotrophic development is the sea urchin *Heliocidaris erythrogramma*. Prior work has found
87 dramatic modifications to morphogenesis and embryonic patterning in *H. erythrogramma*
88 relative to planktotrophic sea urchin species (for reviews see Raff and Byrne 2006; Wray 2022).
89 For instance, dorso-ventral and left-right axis establishment occurs considerably earlier in this
90 species (Henry, et al. 1990), concordant with accelerated rudiment development and
91 metamorphosis (Williams and Anderson 1975). At other embryonic domains, specification of
92 cell lineages such as those responsible for mesenchymal tissue types have been delayed, lost, or
93 spatially rearranged (Wray and Raff 1990; Klueg, et al. 1997; Wilson, Andrews, Rudolf Turner,
94 et al. 2005; Davidson, et al. 2021). Recent transcriptomic studies have shown that expression
95 profiles of key developmental gene regulatory network (dGRN) genes in this species have

96 diverged from the highly conserved expression patterns in planktotrophic species (Israel, et al.
97 2016), and many of these regulatory changes are predicted to be based in *cis* (Wang, et al. 2020).
98 However, the role of evolutionary changes in the sequence and configuration of *cis*-regulatory
99 elements to the rapid evolution of these derived phenotype in *H. erythrogramma* is not known.
100 To address this, we compared chromatin accessibility, regulatory element sequence, and
101 gene expression across six developmental stages in three sea urchin species spanning this life
102 history switch: the lecithotroph *H. erythrogramma*, a closely-related (~ 5 my divergence)
103 planktotroph *H. tuberculata*, and a distantly-related (~ 40 my divergence) planktotroph,
104 *Lytechinus variegatus* (**fig. 1a**). Planktotrophic (feeding) larvae represent the *ancestral* life
105 history mode within echinoderms relative to the *derived* lecithotrophic state (Strathmann 1978,
106 Raff 1987, Wray 1996, McEdward and Miner 2001), so recent and dramatic changes in
107 developmental phenotypes in *H. erythrogramma* provide an exceptional natural system to
108 analyze the consequences of strong shifts in selection and life history on the evolution of gene
109 regulatory mechanisms (Wray 2022).
110 In this study, we investigate the potential contribution of changes in regulatory element
111 accessibility and sequence, as well as their interactions, on the evolution of developmental gene
112 expression at a genome-wide scale. We consider our results in the context of two simple
113 mechanistic models for an evolutionary change in the function of an individual regulatory
114 element during development. (1) *local mutation first*: mutation that occurs within the regulatory
115 element that alters the recruitment of a transcription factor, histone remodeler, or other local
116 chromatin modifier (Clapier and Cairns, 2009; Zaret and Carroll, 2011), thereby altering the
117 accessibility of the regulatory element and, in turn, possibly the transcription of a nearby gene.
118 (2) *regulatory factor first*: a change elsewhere in the genome alters the expression of a pioneer
119 transcription factor, histone remodeler, or other transcription factor that directly alters the
120 accessibility of regulatory element in question and, in turn, possibly the transcription of a nearby
121 gene. Both models accommodate an evolutionary increase or decrease in accessibility, and the
122 accessibility change could alter interactions with other regulatory proteins, by masking or
123 unmasking additional existing protein:DNA binding sites. Finally, in both models, the
124 precipitating mutation(s) could be visible to selection through their impact on gene expression.
125 Within this framework, we identify distinct epigenetic and mutational signatures of
126 evolutionary modifications in the noncoding genome of *H. erythrogramma* relative to species

127 with the ancestral life history. Furthermore, we find that these changes have accumulated non-
128 uniformly across the genome, and are notably enriched near genes with derived expression that
129 are involved in cell type specification, morphogenesis, and development of other derived traits.
130 The patterns of *cis*-regulatory modifications characterized here highlight the contribution of the
131 epigenome to the evolution of transcriptional regulation and trait evolution more broadly.

132 **RESULTS AND DISCUSSION**

133 We performed ATAC-seq on six developmental stages of *H. erythrogramma* (*He*), *H.*
134 *tuberculata* (*Ht*), and *L. variegatus* (*Lv*) in order to identify open chromatin regions (OCRs)
135 representing putative *cis*-regulatory elements (CREs) during embryogenesis. In this study, we
136 use the term "OCR" when reporting analysis results because ATAC-seq is an assay of open
137 chromatin. We apply the more specific term "CRE" when discussing the biological interpretation
138 of our results because *cis*-regulatory elements are the operational units putatively identified
139 within open chromatin peaks used to understand mechanisms of gene regulatory evolution.
140 Importantly, however, while OCRs are known to be enriched for CRE sequences (Thurman et
141 al., 2012), no strict relationship governs the presence of CREs based on occurrence of open
142 chromatin peaks alone. Furthermore, even if an OCR is an authentic CRE, an evolutionary
143 change in chromatin state may or may not have impact on the expression of a nearby gene.
144 Indeed, many OCRs in sea urchin embryos open hours before any nearby gene is transcribed,
145 including experimentally validated CREs (Shashikant et al. 2018; Arenas-Mena 2021); this
146 finding suggests that the initial opening of many OCRs during early development makes a
147 permissive rather than determinative contribution to activation of the zygotic transcriptome.

148 The developmental stages analyzed here include $4^{th}/5^{th}$, $6^{th}/7^{th}$, and $8^{th}/9^{th}$ mixed cleavage
149 stage embryos, as well as hatched blastula, mid-gastrula, and early larva (**fig. 1b**), enabling
150 characterization of regulatory element dynamics from the onset of zygotic transcription through
151 the specification and morphogenesis of complex tissue types. We then carried out alignments of
152 chromosome-scale genome assemblies to find evolutionarily conserved OCRs between species to
153 compare accessibility and sequence composition of putative CREs. These analyses provide a
154 simultaneous comparison of gene regulatory mechanisms in two ways: (1) availability of CREs
155 to transcription factor binding associated with differential chromatin configuration (accessibility)
156 and (2) affinity of transcription factor binding to CREs via changes in nucleotide sequence, as
157 evidenced by motif analysis and signatures of positive selection. Importantly, the genomes of

158 these sea urchin species were assembled and annotated using an identical set of sequencing
159 technology and methodology (Davidson, et al. 2020; Davidson, et al. 2021), minimizing
160 technical bias and permitting high resolution identification of changes in regulatory element
161 activity.

162 Of 42,679 OCRs identified across the three species, we found 35,788 1:1:1 orthologous,
163 high confidence OCRs proximal to 11,679 gene models for cross-species comparison (**fig. 2a**).
164 Note that inclusion in this set of 1:1:1 orthologues does not require that chromatin be open in all
165 three species, but rather that it be open in at least one species at one time point surveyed and that
166 the orthologous genomic sequence is present in all three species. This set of orthologous OCRs
167 forms the basis of most analyses presented here. In addition to orthologous OCRs, we also
168 identified sets of species-specific gains or losses of OCRs occurring on entire genomic segments
169 present or lost in a single species (**fig. 2a**). Because these indel-based differences in OCRs are
170 only present in at most two of the three sea urchin species compared in this study, they are
171 treated as distinct sets of putative CREs. While the changes in regulatory element function
172 examined in this study are not applicable to species-specific OCRs, it is important to note that
173 many insertions containing species-specific OCR gains occur nearby genes with differential
174 expression levels in *H. erythrogramma*, suggesting gain or loss of entire CREs may be an
175 important component of the evolution of the derived life history strategy in this species
176 (**supplementary fig. 1**).

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

178 First, we assessed which biological factors may significantly contribute to variation in
179 chromatin accessibility at a genomic scale. Principal component analysis (PCA) suggests
180 species-level differences (PC 1: 20.3%) and the interaction of development and life history (PC
181 2: 18.5%) are the most important factors explaining variation in chromatin accessibility when all
182 orthologous OCRs are considered (**fig. 1c**). Within each species, chromatin accessibility
183 differences between stages roughly mirrors developmental progression as early and late stages
184 cluster more closely to one another. Strikingly, though, there is far greater variation in
185 accessibility through development in both species with the ancestral life history relative to *He*,
186 suggesting the developmental chromatin landscape of the derived state is less dynamic during
187 embryogenesis. This finding is consistent with delayed blastomere specification (Wray and Raff
188 1989) and simplified larval morphology in *He* (Emlet 1995), suggesting chromatin state changes

189 normally associated with establishing cell identity and tissue differentiation have been modified
190 during the evolution of its derived developmental life history.
191 Interestingly, when considering only OCRs near dGRN genes, the primary driver of
192 variation in chromatin accessibility shifts to developmental life history rather than species-level
193 differences (**fig. 1d**; PC 1: 25.5%), a finding corroborated by variance partitioning of the dataset
194 (**supplementary fig. 2**). This result suggests chromatin configuration of critical developmental
195 network genes is associated with developmental strategy in these species, as accessibility of
196 regulatory elements controls the flow and deployment of embryonic programs, ultimately driving
197 cell fate specification and morphogenesis. Therefore, changes in accessibility to CREs have the
198 potential to modify the function and construction of novel developmental gene networks via
199 altered regulatory interactions.
200 ***Loss of regulatory element accessibility within the ancestral dGRN***
201 We next quantified how the accessibility of 35,788 OCRs varies between life history
202 strategies in order to better understand how chromatin configuration at specific CREs may
203 influence the evolution and function of alternative embryonic programs. Across each cleavage
204 stage examined (timepoints spanning 16-cell through ~500-cell stage embryos), we found an
205 average of 5.7% of OCRs were differentially accessible between ancestral and derived life
206 histories, including a modest difference in sites more accessible in *Ht* and *Lv* (3.1%) or *He*
207 (2.6%) (**fig. 3a,b; top**). Strikingly, starting at blastula stage, the proportion of differentially
208 accessible OCRs across the genome rises, with most sites increasing in accessibility ("opening")
209 in the ancestral life history, but not in *He*. For example, by the larval stage, 9.5% of OCRs are
210 differentially accessible between life histories, nearly three quarters of which (73.9%) are more
211 accessible in *Ht* and *Lv*. Furthermore, when all OCRs are considered, including both orthologous
212 and non-orthologous sites across species, a much greater proportion of the *He* genome is closed
213 compared to *Ht* and *Lv*, suggesting that, at a genome-wide scale, *He* is characterized by zygotic
214 quiescence during early development relative to ancestral life history (**fig. 2e**; **supplementary**
215 **fig. 3**). Our results are consistent with the prediction that, as embryonic development proceeds,
216 more regulatory sites open and become available to regulatory factor binding as zygotic
217 transcription activates, cell lineages differentiate, tissue types are specified, and morphological
218 structures are constructed. However, this genome "awakening" and reconfiguration of chromatin
219 state across the embryo unfolds at a much grander scale in *Ht* and *Lv* relative to *He*. This finding

220 supports the idea that embryonic fate specification in *He* is delayed relative to the ancestral state
221 (Wray and Raff 1989, 1990) and collectively, is concordant with a morphologically simplified
222 larva (Williams and Anderson 1975) and derived gene expression patterning in this species
223 (Wilson, Andrews and Raff 2005; Wilson, Andrews, Rudolf Turner, et al. 2005; Smith, et al.
224 2008; Israel, et al. 2016; Koop, et al. 2017).
225 This delayed zygotic activation hypothesis is further supported when OCR accessibility
226 of dGRN genes is compared. Near these critical developmental regulatory genes, we find a
227 remarkable asymmetry in putative CRE accessibility between life history modes from the onset
228 of development: on average, 0.5% of OCRs near dGRNs are more accessible in *He* during
229 development, whereas this figure ranges from 2.8% (early cleavage) to 15.1% (larva) for species
230 with the ancestral life history (**fig. 3a.b; bottom**). Therefore, modifications to CRE accessibility
231 in *He* are not uniformly distributed across the genome, as would be expected of neutral processes
232 such as drift, but instead are strikingly enriched near dGRN genes (Fisher's exact test, two-sided:
233 $p < 1.31 \times 10^{-11}$). An example of possible CRE activation in the ancestral state that is not found
234 in *He* is the chromatin landscape proximal to the gene *patched* (*ptc*) (**fig. 4**). This gene encodes a
235 receptor of the Hedgehog ligand; its misexpression results in skeletal defects and abnormal
236 mesodermal patterning in the ancestral sea urchin developmental program (Walton, et al. 2009).
237 Given that certain ecological conditions likely favor the evolution of non-feeding
238 development (Strathmann, 1978, 1985; Marshall, et al. 2012), the finding that differences in
239 chromatin accessibility are concentrated around developmental regulatory genes suggests natural
240 selection has sculpted the epigenetic landscape nearby the sea urchin dGRN during life history
241 evolution, generating potentially adaptive changes to transcriptional regulation and expression
242 patterning during embryogenesis. Furthermore, our results indicate that the evolutionary changes
243 in putative CRE accessibility have functional consequences — if they had no impact on gene
244 expression or some other trait, natural selection would not influence their status during
245 development.

246 ***Developmental regulatory elements have accumulated mutational signatures of positive***
247 ***selection***

248 Prior studies have found enrichment of signals of positive selection in OCRs near genes
249 that are differentially expressed between species (Edsall et al. 2019; Swain-Lenz et al. 2019;
250 Shibata et al. 2021). To test whether a similar relationship exists in our study clade and whether

251 it was accentuated during the evolution of lecithotrophy, we next analyzed sequence evolution of
252 orthologous regulatory loci to identify instances of mutational pileup as a signal of positive
253 selection within these sites. Substitution rates of 35,788 OCRs on the *He* and *Ht* branches (table
254 S1) were compared to 10 randomly selected neutrally evolving, non-coding reference sequences,
255 and a likelihood ratio test was implemented to fit substitution rates across the tree to a branch-
256 specific positive selection model, using *Lv* as the outgroup species (see **Methods and Materials**
257 for details). Our goal was to identify cases of gene regulatory evolution via nucleotide variation
258 within CRE sequences, potentially acting to modify transcription factor binding affinity of these
259 elements and associated changes to gene network expression and connectivity.
260 In total, we identified 648 OCRs with evidence of positive selection exclusively in the *Ht*
261 genome, compared to 512 exclusively in the *He* genome (**fig. 5a**: green and orange points,
262 respectively; see also **fig. 2d**). This result thus suggests that, while the life history of these two
263 sea urchin species dramatically differs, many potential changes in gene regulation associated
264 with other aspects of their biology have accumulated within each species. At a genome-wide
265 scale, these putatively adaptive changes have occurred perhaps even more so in *Ht*, the species
266 with ancestral developmental life history.
267 Partitioning OCR selection results by the function of the nearest gene reveals only limited
268 enrichment of positive selection within the putative regulatory elements of most functional
269 classes of genes (**fig. 5b**, **supplementary fig. 4**). A prominent exception to this finding involves
270 OCRs near genes in the sea urchin dGRN or near genes encoding other transcription factors. In
271 these cases, approximately twice as many OCRs exhibit evidence of positive selection on the *He*
272 branch relative to *Ht*. This result is especially striking considering *Ht* has 26.3% more OCRs
273 with branch-specific evidence of positive selection across the entire genome (previous paragraph
274 and **fig. 5b**). Moreover, nucleotide sequences of the 23 dGRN OCRs under selection in *He* are
275 highly enriched for binding motifs of critical developmental transcription factors, including three
276 SOX, three zinc finger, and one homeobox protein (**fig. 5c**). The SOX binding motifs are
277 interesting because many transcription factors in this family act as "pioneer factors" that are
278 capable of initiating chromatin opening (Zaret and Carroll 2011). Another interesting example
279 involves NRF2: the binding motif for this transcription factor is overrepresented in OCRs under
280 selection in *He* (**fig. 5c**) and the gene is expressed at much higher levels in *He* embryos relative
281 to those in the ancestral life history (**fig. 5d**). This protein has gene regulatory functions in

282 diverse biological contexts ranging from antioxidant response and cellular defense to metabolism
283 and stem cell quiescence (Tonelli, et al. 2018), which in this case, could be related to derived
284 embryonic detoxification and physiological systems (Tsushima, et al. 1995; Davidson, et al.
285 2019) or early cell lineage specification programs (Wray and Raff 1989, 1990; Davidson, et al.
286 2021) of *He*.

287 Taken together, putative regulatory elements tasked with wiring the ancestral sea urchin
288 dGRN have accumulated a disproportionate number of mutations on the *He* branch. These likely
289 play a significant role in regulation of divergent gene expression patterns and its highly derived
290 developmental program. However, determining which mutations drive the signal of positive
291 selection and their contribution to gained or lost transcription factor binding motifs remains a
292 challenge. More broadly, our results provide strong evidence that natural selection can efficiently
293 generate heritable changes concentrated within the *cis*-regulatory landscape of specific subsets of
294 genes or gene networks to rapidly produce changes in transcriptional regulation.

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

296 The previous two sections describe how two distinct modifications to regulatory element
297 function, namely accessibility and sequence, have accumulated across the *He* genome, and
298 particularly so near critical developmental network genes. We next asked whether there exists
299 any interaction between these two evolutionary changes to better understand the extent to which
300 both epigenetic and sequence changes affect the same regulatory element. At all three post-
301 cleavage stages, we found that differentially accessible OCRs contain a significantly greater
302 proportion of sites with evidence of selection on the *He* branch relative to non-differentially
303 accessible OCRs, including as much as an ~50 % proportional increase at the blastula stage
304 (Fisher's exact test, two-sided: $p < 5 \times 10^{-3}$) (**fig 6a**: top). This same trend is observed at early
305 cleavage (16/32-cell embryos), although not statistically significant. In contrast, we detected no
306 statistical enrichment for positive selection within differentially accessible OCRs in the *Ht*
307 genome; at early cleavage this relationship does approach significance, but no such trend exists
308 at the later developmental stages (**fig. 6a**: bottom) (Fisher's exact test, two-sided: $p > 0.05$).

309 Significant overlap in differential accessibility and positive selection in the *He* genome
310 suggests a possible biological interaction between the two evolutionary mechanisms that may
311 operate in tandem to alter regulatory function. It therefore becomes useful to consider the order
312 in which these molecular events could take place (**fig. 7**). Measuring chromatin accessibility in

313 hybrids provides a way to distinguish *cis* from *trans* genetic contributions to differences in
314 individual OCRs (Connelly et. al 2014), using an approach originally developed for investigating
315 the genetic basis for evolutionary changes in transcription (Wittkopp, et al. 2008; Crowley, et al.
316 2015). In the present case, the disproportionate occurrence of positive selection within
317 differentially accessible OCRs only in the *He* genome suggests a potentially synergistic set of
318 gene regulatory processes may operate to drive developmental change during the evolution of its
319 derived life history. Because the tests for positive selection identify regions containing multiple
320 derived mutations, additional *cis* mutations likely accumulated after the accessibility change,
321 regardless of which molecular mechanism originally opened the region (**fig. 7**, t=3). The
322 observation that this association is restricted to the *He* branch suggests that this interaction may
323 often produce adaptive changes in transcription during development. More broadly, tandem
324 accessibility and mutational changes may be a common gene regulatory method to efficiently
325 modify gene expression for organismal adaptation and innovation across organismal lineages.

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

327 Next, we asked whether evolutionary changes in OCR sequence and chromatin
328 accessibility differ between core promoters and more distal elements such as enhancers. We
329 quantified changes in proximal (< 500 bp to translational start site of a gene) and distal (≥ 500 bp
330 to translational start site of a gene) OCRs to test if either change covaries with promoter or
331 enhancer evolution between species. While the transcription start site would be a more
332 appropriate genomic reference for defining putative promoter regions, the reference genomes of
333 these species currently lack annotation of 5' untranslated regions for all genes. Thus, the
334 translation start site was used instead to provide a consistent, albeit less sensitive, reference point
335 of comparison of all genes for each species. Although many proximal and distal OCRs show
336 evidence of positive selection on the *Ht* and *He* branches, we found no significant enrichment
337 between the two categories of putative CREs (**fig. 6b**). This finding suggests mutations in both
338 enhancers and promoters are likely important contributors in driving gene expression change, via
339 altered binding and recruitment of regulatory factors and transcriptional machinery.

340 In contrast, accessibility changes show a much greater asymmetry between proximal and
341 distal elements (**fig. 6c**). While accessibility distributions are comparable at early cleavage
342 (16/32-cell embryos), many more accessibility differences accumulate in proximal OCRs relative
343 to distal ones later in development, after zygotic activation and tissue specification has

344 commenced. For example, by the gastrula and larval stages, more than 20% of proximal
345 elements are significantly more accessible in *Lv* and *Ht*, whereas this figure is ~5% for distal
346 elements.

347 There are several possible biological and technical explanations for this result. Unlike
348 enhancers, whose activity is often restricted to a single tissue or stage (Thurman, et al. 2012),
349 promoters need to be accessible whenever the gene is transcribed, regardless of space and time.
350 This has important implications for interpreting differences in chromatin configuration among
351 species. At a biological level, greater promoter accessibility in the embryos of species with the
352 ancestral life history (*Ht* and *Lv*) may merely reflect overall greater transcriptional activity and
353 more changes in expression during development **(fig. 1,3; supplementary fig. 5)** (Israel, et al.
354 2016; Wang, et al. 2020). Secondly, for a given gene, the accessibility of a promoter region is
355 normally greater than that of nearby enhancers (Thurman, et al. 2012), which may lend increased
356 statistical power in detecting accessibility differences between species and contribute to an
357 apparent enrichment of promoter accessibility changes. In addition, these chromatin
358 configuration data were collected from whole embryos, so cell or lineage-specific differences in
359 accessibility will be averaged within the sample, a technical factor that may be especially
360 important for detecting accessibility changes of enhancer elements with tissue-specific dynamics.
361 Lastly, promoter accessibility changes may simply be more often selected for when large
362 changes in transcription are favored, at least for some genes. Adjusting enhancer accessibility
363 can certainly induce changes in gene expression, but activating or shutting down a promoter may
364 be a much more potent mechanism for large-scale changes in transcriptional regulation, such as
365 those involved in rapid developmental evolution.

366 **Relationships between magnitude of gene expression and chromatin accessibility changes**

367 Multiple lines of evidence presented above suggest that changes in the sequence or
368 accessibility of *cis*-regulatory elements contributed to the evolution of development in *He*. The
369 association between these changes and divergence in gene expression, however, is unclear. To
370 understand this relationship, we compared mRNA expression fold-change of genes with
371 accessibility fold-change of nearby OCRs at four developmental stages to measure instances of
372 correlated changes in the magnitude of expression and chromatin configuration. A diagram of
373 possible relationships between OCR accessibility and gene expression differences is provided in
374 **fig. 6d**, to which datapoints are mapped and shown as density plots in **fig. 6e**. We partitioned the

375 data into three categories: (1) all OCRs and genes (no filtering), (2) only differentially expressed
376 genes and their nearby OCRs, and (3) differentially expressed genes and their nearby
377 differentially accessible OCRs.
378 When all genes and OCRs are considered (**fig. 6e**: top row), there is no correlation
379 between the magnitude and direction of accessibility and expression differences across
380 developmental life histories. This result is not surprising at a global scale, given that expression
381 levels and accessibility are conserved between species for most genes and OCRs, respectively.
382 However, a notable set of proximal sites have greater accessibility in species with the ancestral
383 life history at the blastula, gastrula, and larval stages, consistent with overall greater promoter
384 accessibility in *Ht* and *Lv* (see previous section). A similar result occurs when only differentially
385 expressed genes and nearby OCRs are considered (**fig. 6e**: middle row). One possible
386 explanation for the lack of an obvious relationship is that, while only genes that are differentially
387 expressed between life histories are plotted, each of these genes has an average of 3.1 OCRs
388 assigned to it (based on proximity). Each of these OCRs may not be active at a given stage of
389 development, or they may function antagonistically to one another (i.e., regulatory elements can
390 enhance or repress transcription), thereby collectively dampening a correlative signal between
391 gene expression and accessibility.
392 Lastly, we considered only differentially expressed genes and nearby OCRs with
393 differential accessibility. This analysis reveals interesting relationships between transcription and
394 chromatin configuration (**fig. 6e**: bottom row). As development progress, the relatively diffuse
395 relationship between OCR accessibility and gene expression quickly tightens (compare 16/32-
396 cell stage with larva). This likely reflects synchronization of zygotic genome activation with the
397 embryo's *cis*-regulatory landscape as maternal mRNA reserves are depleted and nascent
398 transcription and cell differentiation commences. Greater accessibility at the majority of
399 proximal sites in *Ht* and *Lv* is still found in this subset of genes and OCRs, though smaller sets of
400 putative promoters with greater accessibility in *He* and corresponding expression increases are
401 detected (faint gray spots). Among distal elements, we observed a considerably more diffuse
402 relationship between accessibility, expression, and life history, possibly reflecting the potential
403 activating or repressive functions of enhancers and downstream gene network interactions. At
404 each developmental stage, the most common relationship was increased gene expression and

405 accessibility in *Ht* and *Lv*, consistent with the preponderances of activating (as opposed to
406 repressive) interactions within the ancestral dGRN.
407 Despite these relationships, our results highlight the difficulty in measuring correlations
408 between the magnitude of cross-species changes in gene expression and regulatory element
409 accessibility at a genomic scale. Prior studies have also reported a weak relationship between
410 evolutionary changes in OCR accessibility and the expression of nearby genes between species
411 (Shibata et al. 2012; Pizzollo et al. 2018). Several factors likely contribute a weak correlation,
412 including: (1) a change in chromatin accessibility is just one of several possible molecular
413 mechanisms that can change a gene's expression; (2) individual CREs can have antagonistic
414 effects on transcriptional regulation as activators or repressors, thus cancelling out effects; and
415 (3) CRE accessibility is permissive rather than directive of regulatory factor binding and
416 downstream gene expression, such that "openness" does not necessarily indicate the CRE is
417 actively regulating transcription.
418 **Developmental trajectories of chromatin configuration and gene expression have evolved**
419 **most frequently along the *H. erythrogramma* branch**
420 To address the challenges associated with understanding the relationship between
421 expression and accessibility changes of putative regulatory elements, we applied the "jump
422 score" method to measure the magnitude of evolutionary changes in developmental trajectories,
423 which are represented as distance in principal component space (Israel, et al. 2016). This metric
424 is particularly useful because the jump score summarizes patterns of conservation or divergence
425 in gene expression across development, rather than considering only a single stage. Here we
426 apply this approach to evolutionary changes in the accessibility of putative CREs, as well as
427 expression of genes, across developmental trajectories (**fig. 8)**.
428 We compared the expression jump score of 9,423 genes and the mean jump score of
429 nearby OCRs to test how developmental changes in transcriptional regulation and chromatin
430 status are distributed along the *He* and *Ht* branches. For example, expression of *insm*, a
431 transcription factor involved in neurogenesis in sea urchins (McClay, et al. 2018), is conserved
432 across *He*, *Ht*, and *Lv* (**fig. 8a**). This is reflected by a low difference in jump scores between *He-*
433 *Lv* and *Ht-Lv,* and tight clustering of expression profiles in principal component space (**fig. 8b**).
434 Similar expression profiles are accompanied by conserved *cis*-regulatory landscape dynamics
435 through development in each species, exemplified by a near-zero mean OCR jump score

436 difference and accessibility profile of ten nearby OCRs (**fig. 8c,d**). In contrast, *brn 1/2/4*, a
437 regulator of midgut patterning (Yuh, et al. 2005), has a divergent developmental gene expression
438 trajectory in *He* relative to the ancestral state and large jump score differences in *He-Lv* and *Ht-*
439 *Lv* branch comparisons (**fig. 8e,f**). This large expression change is accompanied by significant
440 reduction in chromatin accessibility at the blastula through larval stages in *He*, represented by a
441 large mean difference in OCR jump scores along the *He* branch (**fig. 8g,h**). This locus is
442 especially interesting considering *He* larvae do not feed and have lost development of a
443 functional through-gut.

444 Given the highly derived developmental program of *He* relative to the ancestral life
445 history, we predicted that the most frequent changes in transcriptional regulation would occur
446 along the *He* branch. Indeed, comparing jump scores of all genes and OCRs identifies concerted
447 modifications to OCR accessibility and gene expression profiles during *He* development as the
448 most frequent set of changes among genes and OCRs with divergent expression and chromatin
449 accessibility trajectories (1,582 genes) (**supplementary fig. 6**; **Data S7**). Furthermore, the least
450 common combination of changes includes both expression and accessibility changes occurring
451 on the *Ht* branch (1,114 genes), consistent with an overall more conserved embryonic gene
452 expression program in this species. Thus, we find that local patterns of CRE accessibility have
453 evolved in concert with divergent developmental expression patterns during life history
454 evolution in *He*. Collectively, our results suggest that modified transcription factor access to
455 regulatory elements is associated with rapid and adaptive evolutionary changes to gene
456 expression, and by extension, dGRN functionality. However, it remains to be determined which
457 of changes evolved as a cause or consequence (or neither) of life history evolution in *He* — a
458 challenge potentially addressable via empirical validation of candidate enhancer function.

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

460 Lastly, we tested whether changes in nucleotide sequence and accessibility within OCRs
461 are distributed near genes that are differentially expressed between developmental life histories.
462 We found a small, though insignificant, increase in differential expression of genes with nearby
463 differentially accessible OCRs relative to genes lacking differentially accessible OCRs at
464 blastula, gastrula and larval stages (+6.2-9.9% change). No such increase exists during early
465 cleavage (-1.8% change), which may be expected given the transcriptome is composed almost

466 entirely of maternally-derived transcripts during early cleavage (Fisher's exact test, two-sided: p
467 > 0.05) (**fig. 9a**).
468 Similarly, we detected no significant difference in the proportion of differentially
469 expressed genes and the presence of positive selection within OCRs of *He* when compared to
470 OCRs lacking evidence of positive selection or OCRs with evidence of positive selection in *Ht* at
471 each developmental stage (Fisher's exact test, two-sided: p > 0.05) (**fig. 9b**). Still, we note that
472 throughout development, incidences of differential expression are higher when there exists one
473 or more nearby OCRs with evidence of positive selection relative to OCRs lacking evidence of
474 positive selection (+5.7-11.5% change; **fig. 9b**: orange and gray bars) or OCRs with evidence of
475 positive selection in *Ht* (+14.0-27.0% change; **fig. 9b**: orange and green bars). This finding is
476 concordant with our predictions of evolutionary mechanisms of gene expression change, as
477 selective changes in CRE sequences in *He* are most often coupled with concordant divergence in
478 developmental gene expression possibly associated with their derived life history strategy.
479 Positive selection within CRE sequences of *Ht*, on the other hand, are less often linked to
480 changes in gene expression associated with life history, as *Ht* has retained the ancestral sea
481 urchin developmental program. Lastly, while the absolute number of accessibility changes
482 between developmental life histories is typically greater than detected instances of positive
483 selection (**figs. 3,5**), mutations within putative CRE sequences are more often associated with
484 changes in gene regulation and expression relative to a no-change state (mean +9.3% change)
485 than changes in accessibility (mean +5.3% change).
486 These results may reflect distinct characteristics of the molecular mechanisms that
487 modify *cis*-regulatory function. In particular, an evolutionary change in the chromatin
488 accessibility of CRE might only manifest at certain developmental stages or in certain cell types,
489 whereas a change in the DNA sequence of a CRE is present throughout development and in
490 every cell type. For this reason, mutations that influence chromatin status might in general have
491 lower pleiotropy than mutations within OCRs. This possibility is an extension of the idea that
492 mutations in CREs that drive tissue-specific expression may have lower pleiotropy than those in
493 CREs that influence expression throughout development and in all cell types (Wray et al. 2003;
494 Stern 2010; Singh and Yi 2021; but see Preger Ben-Noon et al. 2018). Evolutionary changes in
495 chromatin configuration are an additional molecular mechanism with the potential to limit

496 pleiotropy, by physically masking a mutation within a CRE such that it can only exert an effect
497 on transcription when chromatin is open.
498 Conversely, opening chromatin in a larger range of conditions could expose cryptic
499 genetic variation to natural selection. Although there is good empirical evidence that cryptic
500 genetic variation exists within natural populations (Gibson & Dworkin, 2004; Paaby and
501 Rockman, 2014), it is less clear how it becomes visible. One possibility (among others) is that
502 mutations accumulate within a cell type-specific CRE that have no consequence because they do
503 not influence interactions with any of the *trans*-acting factors in that particular cell type. An
504 evolutionary change that opens the local chromatin in a different cell type might, however,
505 expose those new mutations to a different suite of trans-acting factors, potentially altering their
506 interactions with the CRE and thus transcription. Deconvolving the interplay and evolutionary
507 consequences of epigenetic and sequence changes within CREs remains an important challenge
508 for understanding the evolution of transcriptional regulation.
509
510 **Conclusions**
511 In this study, we leveraged a recent life history switch in sea urchins that was
512 accompanied by elevated rates of adaptive changes in developmental gene expression (Israel et
513 al. 2016; Wang et al. 2020; Wray 2022) to investigate how evolutionary mechanisms act on
514 mechanisms of transcriptional regulation and embryonic gene expression. In particular, we
515 distinguish two distinct molecular mechanisms within *cis*-regulatory elements that could alter
516 gene expression: sequence changes that might influence transcription factor:DNA interactions
517 and epigenetic changes that might limit access of transcription factors to DNA. We find that
518 evolutionary changes in chromatin are common near differentially expressed genes, especially on
519 the branch leading to the derived life history and especially so near critical developmental gene
520 regulatory network (dGRN) genes (**fig. 10**, left column). Elevated accumulation of mutations
521 within OCRs indicative of positive selection is not enriched on the branch leading to the derived
522 life history when considering all genes, but strikingly enriched near dGRN genes specifically
523 (**fig. 10**, central column). These distinct molecular mechanisms may also act synergistically, as
524 seen by enrichment for both changes within putative dGRN CREs on the branch leading to the
525 derived life history (**fig. 10**, right column). Although extensive prior evidence indicates that
526 changes in protein:DNA binding contributes to the evolution of gene expression, much less

527 empirical support exists regarding the contribution of epigenetic changes. Our results suggest
528 that evolutionary changes in local chromatin configuration during early development can
529 contribute to the evolution of embryonic gene expression, while branch- and dGRN-specific
530 enrichments suggest that many of these changes are adaptive and may be tied to the evolution of
531 the derived life history. More broadly, the comparative approaches developed in this study offer
532 new ways to analyze multimodal functional genomic datasets to investigate the evolutionary
533 mechanisms that operate on *cis*-regulatory elements and their consequences for gene expression
534 and organismal traits.
535

536 **MATERIALS AND METHODS**

537 ***Embryo Rearing and Sample Preparation***

538 *H. erythrogramma* and *H. tuberculata* adult animals were collected near Sydney, NSW,
539 Australia and maintained at 21°C in aquaria with circulating natural seawater. *L. variegatus*
540 individuals were collected near the Duke University Marine Lab in Beaufort, NC, USA and
541 maintained at 23°C in aquaria at Duke University, Durham, NC, USA. Each sea urchin species
542 was induced to spawn via injection of 0.5 M KCl solution into the coelom. Eggs were washed
543 twice in 20 µm filtered seawater (FSW). Sperm was collected dry and stored at 4°C until usage.
544 Washed eggs were fertilized with 1 µL of sperm, and after 10 minutes, were washed three times
545 in FSW. For each species, three unique male-female pairs were crossed to produce three
546 biologically-independent replicates of sea urchin embryos. Embryos were reared at 23°C in large
547 glass culture dishes supplied with FSW, and water was changed every six hours. Because these
548 species exhibit different developmental rates, developmental milestones, rather than absolute
549 developmental time, were selected to maximize developmental synchrony within cultures and
550 between species for comparison. These stages include 1) $4^{th}/5^{th}$ (EC), 2) $6^{th}/7^{th}$ (MC), and 3)
551 $8^{th}/9^{th}$ (LC) cleavage stage embryos, as well as 4) hatched blastula (B), 5) mid-gastrula (G), and
552 6) early larva (L) (two-armed pluteus for *H.* tuberculata and *L. variegatus*; 32 hours post-
553 fertilization for *H. erythrogramma*). Quickly distinguishing embryonic stages is difficult for
554 cleavage stage embryos, so we closely monitored development of each culture and used known
555 division rates for each species and water temperature to estimate when the culture would be most
556 enriched for the targeted developmental stage. Once a culture reached the appropriate stage, a
557 subset of the culture was retrieved for processing, stage confirmed by visual inspection, arrested

558 or delayed embryos removed by hand, and the remaining embryos transferred into a 1.5 mL
559 Eppendorf tube for ATAC-seq processing. Microscopic examination confirmed that most
560 embryos had reached the target stage.

561 ***ATAC-seq Protocol and Sequencing***

562 ATAC sample preparation was carried out according to the Omni-ATAC-seq protocol
563 (Corces, et al. 2017). For each replicate, embryos were washed once in 1 µm filtered seawater,
564 lysed, and ~50,000 nuclei were isolated for the transposition reaction using the Illumina TDE1
565 enzyme and tagmentation (TD) buffer (Cat. No. 20034197 and 20034198) (San Diego, CA,
566 USA). Sequencing libraries for each replicate were amplified via PCR after separately
567 determining the optimal number of amplification cycles for each sample using qPCR as
568 described in (Buenrostro, et al. 2015), and libraries were purified and size selected using Ampure
569 XP Beads at a 1.8:1 bead volume:library volume (Beckman Coulter, Brea, CA, USA). Library
570 quality and transposition efficiency was accessed using a Fragment Analyzer and PROSize 2.0
571 (Agilent, Santa Clara, CA) at the Duke University Center for Genomic and Computational
572 Biology. *H. erythrogramma* and *L. variegatus* libraries were sequenced on an Illumina HiSeq
573 4000 instrument using 50 bp SE sequencing. *H. tuberculata* libraries were sequenced on an
574 Illumina NovaSeq 6000 instrument using 50 bp PE sequencing (only SE were used for data
575 analysis).

576 ***ATAC-seq Data Analysis***

577 Raw ATAC-seq reads were trimmed for quality and sequencing adapters using
578 "cutadapt" (Martin 2011) v2.3 with the following parameters: -a CTGTCTCTTATACACATCT
579 -q 20 --trim-n -m 40. Trimmed reads were then aligned to each species' respective genome via
580 "stampy" (Lunter and Goodson 2011) v.1.0.28 using the "—sensitive" set of parameters. The *Lv*
581 genome was retrieved from Davidson et al. 2020 (NCBI BioProject: PRJNA657258). The
582 *Heliocidaris* genome assembles were retrieved from Davidson et al. 2021 (NCBI BioProject: *He*-
583 PRJNA827916; *Ht*- PRJNA827769). ATAC-seq alignments were filtered for mitochondrial
584 sequences and required an alignment quality score of at least 5 using "samtools" (Li, et al. 2009)
585 v.1.9. Filtered alignments are available in BAM and Bedgraph format (see "Data Availability"
586 below).

587 In this study, we aimed to compare evolution of orthologous regulatory elements so to
588 accomplish this, we performed a series of liftOvers (Hinrichs, et al. 2006) to convert ATAC-seq

589 alignments between genomic coordinates of each sea urchin species. We took an iterative,
590 reciprocal strategy described below to minimize possible reference bias associated with
591 converting between genome assemblies, requiring a minMatch of 0.75 when lifting alignments
592 between species pairs: 1) *H. erythrogramma*: *He* → *Lv* → *He*; 2) *H. tuberculata*: *Ht* → *Lv* → *He*;
593 3) *L. variegatus*: *Lv* → *He* → *Lv* → *He*. At the end of filtering and coordinate conversion, all
594 ATAC-seq alignments were referenced to the *H. erythrogramma* genome. LiftOver files are
595 available in (Davidson, et al. 2021).
596 ATAC-seq peaks were called from these filtered, converted alignments using the
597 "MACS2" v2.1.2 (Zhang, et al. 2008) callpeak function (parameters: –nomodel, –keep-
598 dup=auto, –shift 100, –extsize 200) for each developmental stage separately at an FDR threshold
599 of 5%. Peak coordinates were merged using the "bedtools" (Quinlan and Hall 2010) v2.25 *merge*
600 function requiring a peak overlap of at least 200 bp to be merged into a single peak. Lastly, for
601 each sample, accessibility of each peak was measured with the "bedtools" *multiBamCov* function
602 and imported into R v4.0.2 for statistical analysis (Data S1). While not a primary focus of this
603 study, all ATAC-seq peaks and raw accessibility counts for each species, independent of
604 sequence conservation between species, is additionally provided in Data S8, as are ATAC-seq
605 peaks specific to *Lv* (not found in *Heliocidaris*).
606 ***Estimates of positive selection***
607 To measure branch-specific signatures of positive selection in the non-coding genome,
608 the "adaptiPhy" (Berrio, et al. 2020) pipeline (https://github.com/wodanaz/adaptiPhy) for global
609 tests of natural selection was followed. First, orthologous sequences for non-coding sites of
610 interest (ATAC-seq peaks) were extracted from each species' genome into FASTA format.
611 Sequences were trimmed to include only contiguous DNA sequence using the *prunning.py* script
612 and filtered using the *filtering.py* script, requiring a minimum alignment length of 75 base pairs.
613 These trimmed and filtered alignments serve as "query" sequences of tests for selection. To
614 generate a neutral reference for comparison, ten orthologous, neutral sites were randomly
615 selected from a curated set of ~88,000 putatively neutral sequences (Davidson, et al. 2021) (Data
616 S2) and concatenated into a single neutral reference sequence. A global neutral reference, rather
617 than a local reference, was implemented to more conservatively estimate signatures of positive
618 selection as recent work has demonstrated a local neutral reference may more likely overestimate
619 positive selection – an especially acute issue in smaller, gene-dense genomes (Berrio et al.,

620 2020). Then, substitution rates of both the query and randomly concatenated neutral reference
621 were estimated using phyloFit (Siepel and Haussler 2004), and the zeta score was calculated as
622 the ratio of the query substitution rate to the neutral reference substitution rate. In addition, p-
623 values of likelihood ratio tests for significant levels of branch-specific positive selection were
624 calculated with adaptiPhy (Berrio, et al. 2020) pipeline using HyPhy (Pond, et al. 2005). For
625 each query site, tests for positive selection were repeated 10 times against a unique neutral
626 reference to ensure reproducible estimates of zeta and significance values for each non-coding
627 site of interest. P-values and substitution rates for all query and neutral sites were then imported
628 in R v4.0.2. For each query site, the median zeta score and p-value for each test of positive
629 selection across the ten replicate calculations were used for downstream analyses.

630 ***OCR Filtering***

631 After accessibility and estimates of positive selection were calculated for each ATAC-seq
632 peak, herein referred to as an "open chromatin region" (OCR), a series of filtering and quality
633 control metrics were carried out to ensure only high confidence and quality peaks were compared
634 between species. These filtering steps are as follows: 1) each OCR is required to have at least
635 75bp of contiguous, single copy sequence (see previous section) for accurate estimations of
636 selection ; 2) a local composition complexity (LCC) (Konopka 2001) value of 1.9 or more was
637 required of the OCR for each species to remove repetitive or other low-complexity sequences
638 that may generate inaccurate estimations of selection (module:
639 biopython.org/docs/1.75/api/Bio.SeqUtils.lcc.html); 3) a CPM of 3 or more was required in at
640 least 9 samples to remove OCRs with extremely low accessibility across samples; 4) the
641 midpoint of the OCR must lie within 25 Kb (in either direction) of the translational start site of a
642 gene model. The 25 Kb window for assigning OCRs to genes was selected on the basis that the
643 median and mean intergenic distance within the *He* genome are 21.4 Kb and 31.9 Kb,
644 respectively. See Supplementary Table 1 for a summary of these filtering steps and resulting
645 number of OCRs. Given nearly no prior knowledge is known of the *cis*-regulatory landscape for
646 these sea urchin species, these stringent filtering methods were carried out in order to maximize
647 confidence in comparisons of non-coding sequence evolution and function. This resulted in a
648 final set 35,788 high-confidence OCRs for cross-species analysis (Data S3).

649 ***ATAC-seq Statistical Analysis***

650 Principal component analysis and variance partitioning was performed on the regularized
651 log-transformed (rlog) data (Love, et al. 2014) of filtered OCR accessibility counts via the
652 *prcomp* function and the "variancePartition" v1.18.3 package (Hoffman and Schadt 2016) in R
653 using the formula: ~ (1 | Stage) + (1 | Species) + (1 | Life History) + (1 | Genus) + (1 | Stage: Life
654 History). Raw accessibility counts of the filtered OCRs were loaded into DESeq2 (Love, et al.
655 2014) v1.30 to calculate differential accessibility between sample groups. Differentially
656 accessible sites between life history strategies were classified as having a 2-fold accessibility
657 difference and supported by an FDR of 10% in *He* vs *Ht* and well as *He* vs *Lv* comparisons
658 (including the sign of the accessibility difference being in the same direction) (see
659 **supplementary fig. 7** for all significance cutoff combinations of differential accessibility
660 results). For graphical purposes, the average fold-change in the *He* vs *Ht* and *He* vs *Lv*
661 comparisons is displayed within figures. Significant levels of positive selection in either the *He*
662 or *Ht* branch were classified as OCRs having a zeta value greater than 1.25 and supported by a
663 median false-discovery rate less than 10% in one species, and failing to meet both criteria in the
664 other. Motif enrichment analyses were carried out with "Homer" v4.11 (Heinz, et al. 2010) via
665 the findMotifsGenome.pl script using all OCRs as the background set and regions size of 200 bp.
666 Sea urchin GRN genes are provided in Data S4, retrieved from the Biotapestry project of the
667 Institute of Systems Biology, Seattle, WA, USA (www.biotapestry.org: accessed June 27, 2017).
668 ***RNA-seq Expression Analyses***
669 Raw RNA-seq reads from eggs, 4-cell, 16-cell + 32-cell (collectively EE) , blastula (B),
670 gastrula (G), and early larva (L) stage embryos of *He*, *Ht*, and *Lv* were retrieved from (Israel, et
671 al. 2016), trimmed and filtered for low quality bases and reads with Trimmomatic (Bolger, et al.
672 2014), and aligned to each species respective genomes and gene models with STAR (Dobin, et
673 al. 2013). From these alignments, mRNA expression was estimated with Salmon (Patro, et al.
674 2017) and loaded to R for statistical analysis. Read counts for summed to each gene's best match
675 to the *S. purpuratus* v4.2 gene models to generate a common reference for expression
676 comparisons between species as described in (Israel, et al. 2016) (Data S5). Differentially
677 expressed genes between life histories were called as having a fold-change in expression > 2 and
678 supported by an FDR of 10% or less in *He* vs *Ht* and well as *He* vs *Lv* comparisons (including
679 the sign of the accessibility difference being in the same direction) (Data S6). These criteria
680 were chosen to make RNA and ATAC-seq analyses between species comparable as possible.

681 ***Principal Component Distances of Developmental Profile***

682 To identify OCRs and genes whose developmental regulatory profiles diverged in *H.*
683 *erythrogramma* or *H. tuberculata*, we calculated principal component distances of chromatin
684 accessibility and gene expression profile shapes between species throughout development as
685 described in (Israel, et al. 2016). In this method, developmental time course data is plotted onto
686 principal component space for each species, and pairwise distances between species are
687 calculated for each gene or OCR (see below). Large differences in these distances provide
688 evidence that the shape of the developmental trajectory for a given gene or OCR has changed in
689 one of the species (the species with the larger pairwise distance). Mean accessibility and
690 expression profiles of regularized log-transformed (rlog) OCR and genes counts were
691 standardized using the *standardize* function of the "mfuzz" v2.48.0 package (Kumar and M
692 2007). From these data, scaled distance measures ("jump scores") between orthologous non-
693 coding sites or genes were calculated using loadings of the first two principal components (Data
694 S7): $d_{HeLv} = \sqrt{PC1(x_{Lv} - x_{He})^2 + PC2(y_{Lv} - y_{He})^2}$;
695 $d_{HtLv} = \sqrt{PC1(x_{Lv} - x_{Ht})^2 + PC2(y_{Lv} - y_{Ht})^2}$. A significant change in the developmental
696 trajectory of gene expression or OCR accessibility was determined to have a jump score
697 difference of at 0.15 on one branch over the other .

698 **DATA AVAILIBILITY**

699 Raw ATAC-seq reads are available in FASTQ format on NCBI's Short Read Archive
700 (PRJNA828607). Filtered alignments of ATAC-seq reads before and after between-species
701 liftOver files are available on Dryad (https://doi.org/10.5061/dryad.jsxksn0cp). Results files of
702 ATAC-seq analyses are available as Supplementary Data files.

703 **AUTHOR CONTRIBUTIONS**

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

707 **ACKNOWLEDGEMENTS**

708 We would like to thank the Sydney Institute of Marine Sciences and their staff for assistance
709 with animal care, laboratory set up, and carrying out experiments. We also thank Allison Edgar
710 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.

716 **REFERENCES**

717 Arenas-Mena C, Miljovska S, Rice EJ, Gurges J, Shashikant T, Wang Z, Ercan S, Danko CG.
718 2021. Identification and prediction of developmental enhancers in sea urchin embryos.
719 *BMC Genomics* 22(1):751.

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

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

724 Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. 2015. ATAC-seq: A Method for Assaying
725 Chromatin Accessibility Genome-Wide. *Current Protocols in Molecular Biology* 109:21
726 29 21-29.

727 Byrne M, Sewell MA. 2019. Evolution of maternal lipid provisioning strategies in echinoids with
728 non-feeding larvae: selection for high-quality juveniles. *Marine Ecology Progress Serie*s
729 616:95-106.

730 Clapier, CR., and Bradley RC. 2009. The biology of chromatin remodeling complexes. *Annual*
731 *Review of Biochemistry* 78: 273-304.

732 Connelly CF, Wakefield J, Akey JM. 2014. Evolution and genetic architecture of chromatin
733 accessibility and function in yeast. *PLoS Genetics* 10(7):e1004427.

734 Crowley, J.J., Zhabotynsky, V., Sun, W., Huang, S., Pakatci, I.K., Kim, Y., Wang, J.R., Morgan,
735 A.P., Calaway, J.D., Aylor, D.L. and Yun, Z., 2015. Analyses of allele-specific gene
736 expression in highly divergent mouse crosses identifies pervasive allelic imbalance.
737 *Nature Genetics* 47(4): 353-360.

738 Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S,
739 Satpathy AT, Rubin AJ, Montine KS, Wu B, et al. 2017. An improved ATAC-seq
740 protocol reduces background and enables interrogation of frozen tissues. *Nature Methods*
741 14:959-962.

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

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

747 Davidson PL, Guo H, Swart JS, Massri AJ, Edgar A, Wang L, Berrio A, Devens HR, Koop D,
748 Cisternas P, et al. 2021. Recent reconfiguration of an ancient developmental gene
749 regulatory network in *Heliocidaris* sea urchins. *bioRxiv* 2022.03.03.482868.
750 Davidson PL, Guo H, Wang L, Berrio A, Zhang H, Chang Y, Soborowski AL, McClay DR, Fan
751 G, Wray GA. 2020. Chromosomal-level genome assembly of the sea urchin
752 *Lytechinus variegatus* substantially improves functional genomic analyses. *Genome*
753 *Biology and Evolution* 12:1080-1086.
754 Davidson PL, Thompson JW, Foster MW, Moseley MA, Byrne M, Wray GA. 2019. A
755 comparative analysis of egg provisioning using mass spectrometry during rapid life
756 history evolution in sea urchins. *Evolution & Development* 21:188-204.
757 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras
758 TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29:15-21.
759 Edsall LE, Berrio A, Majoros WH, Swain-Lenz D, Morrow S, Shibata Y, Safi A, Wray GA,
760 Crawford GE, Allen AS. 2019. Evaluating Chromatin Accessibility Differences Across
761 Multiple Primate Species Using a Joint Modeling Approach. *Genome Biology and*
762 *Evolution* 11(10):3035-3053.
763 Emlet RB. 1995. Larval spicules, cilia, and symmetry as remnants of indirect development in the
764 direct developing sea urchin *Heliocidaris erythrogramma*. *Developmental Biology*
765 167:405-415.
766 Falkner I, Sewell MA, Byrne M. 2015. Evolution of maternal provisioning in ophiuroid
767 echinoderms: characterisation of egg composition in planktotrophic and lecithotrophic
768 developers. *Marine Ecology Progress Series* 525:1-13.
769 Flatt T, Heyland A. 2011. Mechanisms of life history evolution: the genetics and physiology of
770 life history traits and trade-offs. Oxford: Oxford University Press.
771 Gibson G, Dworkin I. 2004. Uncovering cryptic genetic variation. *Nature Reviews Genetics*
772 5(9):681-90
773 Hart MW. 2002. Life history evolution and comparative developmental biology of echinoderms.
774 *Evolution & Development* 4:62-71.
775 Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass
776 CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-
777 regulatory elements required for macrophage and B cell identities. *Molecular Cell*

778 38:576-589.
779 Henry JJ, Wray GA, Raff RA. 1990. The dorsoventral axis is specified prior to 1st cleavage in
780 the direct developing sea urchin *Heliocidaris erythrogramma*. *Development* 110:875-
781 884.
782 Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, Diekhans M, Furey
783 TS, Harte RA, Hsu F. 2006. The UCSC genome browser database: update 2006. *Nucleic*
784 *Acids Research* 34:D590-D598.
785 Hoffman GE, Schadt EE. 2016. variancePartition: interpreting drivers of variation in complex
786 gene expression studies. *BMC Bioinformatics* 17:483.
787 Israel JW, Martik ML, Byrne M, Raff EC, Raff RA, McClay DR, Wray GA. 2016. Comparative
788 developmental transcriptomics reveals rewiring of a highly conserved gene
789 regulatory network during a major life history switch in the sea urchin genus
790 *Heliocidaris*. *PLoS Biology* 14:e1002391.
791 Jhanwar S, Malkmus J, Stolte J, Romashkina O, Zuniga A, Zeller R. 2021. Conserved and
792 species-specific chromatin remodeling and regulatory dynamics during mouse and
793 chicken limb bud development. *Nature Communications* 12:5685.
794 Klueg KM, Harkey MA, Raff RA. 1997. Mechanisms of evolutionary changes in timing, spatial
795 expression, and mRNA processing in the msp130 gene in a direct-developing sea urchin,
796 *Heliocidaris erythrogramma*. *Developmental Biology* 182:121-133.
797 Konopka AK. 2001. Sequence complexity and composition. *eLS*.
798 Koop D, Cisternas P, Morris VB, Strbenac D, Yang JY, Wray GA, Byrne M. 2017. Nodal and
799 BMP expression during the transition to pentamery in the sea urchin *Heliocidaris*
800 *erythrogramma*: insights into patterning the enigmatic echinoderm body plan. *BMC*
801 *Developmental Biology* 17:4.
802 Kumar L, M EF. 2007. Mfuzz: a software package for soft clustering of microarray data.
803 *Bioinformation* 2:5-7.
804 Lewis JJ, Reed RD. 2019. Genome-wide regulatory adaptation shapes population-level
805 genomic landscapes in *Heliconius*. *Molecular Biology & Evolution* 36:159-173.
806 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R,
807 Genome Project Data Processing S. 2009. The sequence alignment/map format and
808 SAMtools. *Bioinformatics* 25:2078-2079.

809 Livraghi L, Hanly JJ, Van Bellghem SM, Montejo-Kovacevich G, van der Heijden ES, Loh LS,
810 Ren A, Warren IA, Lewis JJ, Concha C, et al. 2021. Cortex cis-regulatory switches
811 establish scale colour identity and pattern diversity in *Heliconius*. *eLife* 10: e68549.
812 Love AC, Raff RA. 2006. Larval ectoderm, organizational homology, and the origins of
813 evolutionary novelty. Journal of Experimental Zoology Part B: *Molecular and*
814 *Developmental Evolution* 306:18-34.
815 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for
816 RNA-seq data with DESeq2. *Genome Biology* 15:550.
817 Lunter G, Goodson M. 2011. Stampy: a statistical algorithm for sensitive and fast mapping of
818 Illumina sequence reads. *Genome Research* 21:936-939.
819 Marshall DJ, Krug PJ, Kupriyanova EK, Byrne M, Emlet RB. 2012. The biogeography of
820 marine invertebrate life histories. *Annual Review of Ecology, Evolution, and*
821 *Systematics* 43:97-114.
822 Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads.
823 *EMBnet.journal* 17:10-12.
824 McClay DR, Miranda E, Feinberg SL. 2018. Neurogenesis in the sea urchin embryo is initiated
825 uniquely in three domains. *Development* 145(21):dev167742.
826 McEdward LR, Miner BG. 2001. Larval and life-cycle patterns in echinoderms. *Canadian*
827 *Journal of Zoology* 79:1125-1170.
828 Morgan SG. 2020. Life and death in the plankton: larval mortality and adaptation. In McEdward,
829 LR, editor. Ecology of marine invertebrate larvae. Baca Raton (FL): CRC Press. p. 279-
830 321.
831 Paaby AB, Rockman MV. 2014. Cryptic genetic variation: evolution's hidden substrate. *Nature*
832 *Reviews Genetics*. 15(4):247-58.
833 Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-
834 aware quantification of transcript expression. *Nature Methods* 14:417-419.
835 Pechenik JA. 1999. On the advantages and disadvantages of larval stages in benthic marine
836 invertebrate life cycles. *Marine Ecology Progress Series* 177:269-297.
837 Preger-Ben Noon E, Sabarís G, Ortiz DM, Sager J, Liebowitz A, Stern DL, Frankel N. 2018.
838 Comprehensive Analysis of a cis-Regulatory Region Reveals Pleiotropy in Enhancer
839 Function. *Cell Reports* 22(11):3021-3031.

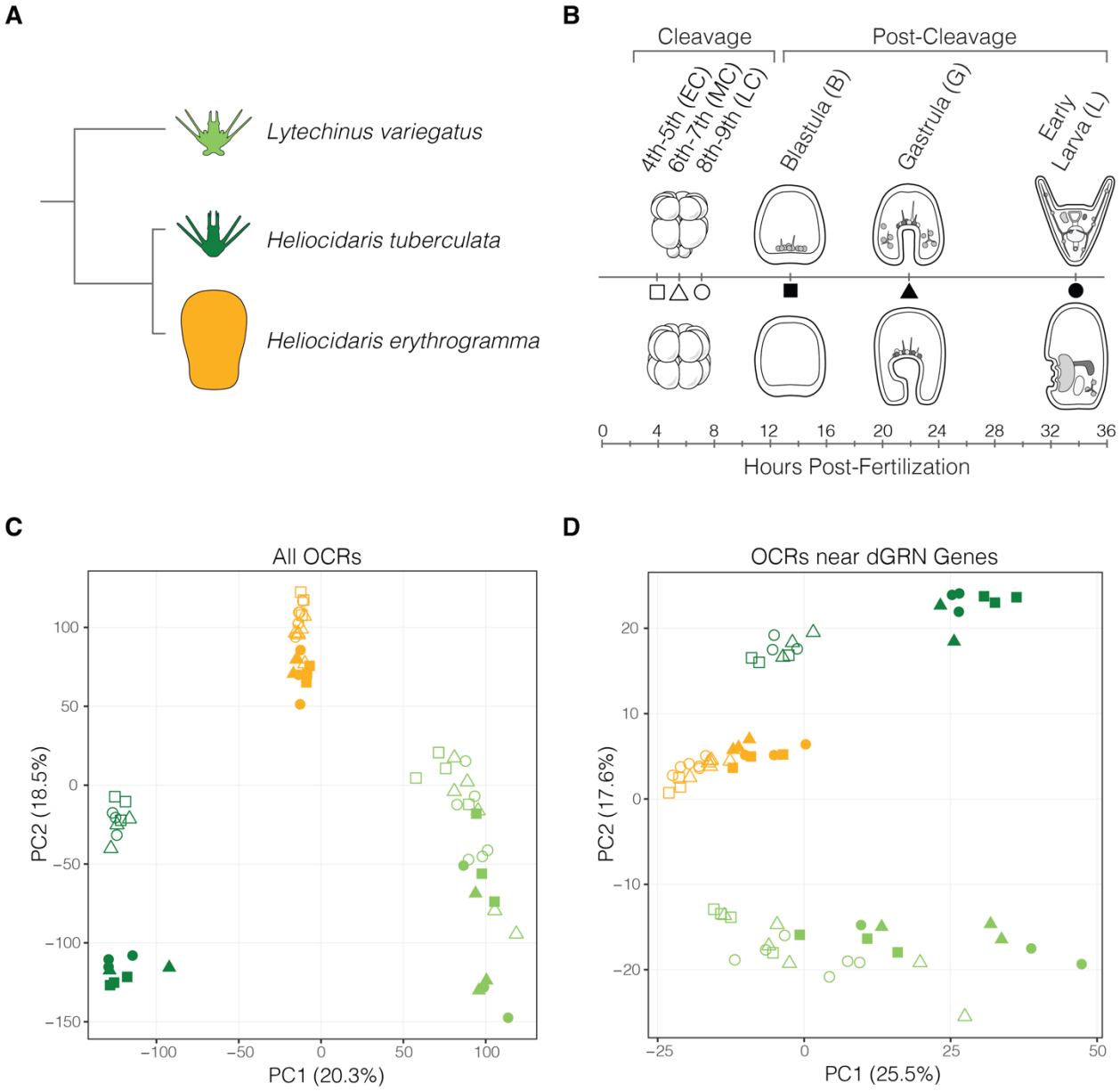
840 Pond SL, Frost SD, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies.
841 *Bioinformatics* 21:676-679.
842 Prowse TAA, Falkner I, Sewell MA, Byrne M. 2009. Long-term storage lipids and
843 developmental evolution in echinoderms. *Evolutionary Ecology Research* 11:1069-1083.
844 Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic
845 features. *Bioinformatics* 26:841-842.
846 Raff RA. 1987. Constraint, flexibility, and phylogenetic history in the evolution of direct
847 development in sea urchins. *Developmental Biology* 119:6-19.
848 Raff RA, Byrne M. 2006. The active evolutionary lives of echinoderm larvae. *Heredity* 97:244-
849 252.
850 Sackton TB, Grayson P, Cloutier A, Hu Z, Liu JS, Wheeler NE, Gardner PP, Clarke JA, Baker
851 AJ, Clamp M, et al. 2019. Convergent regulatory evolution and loss of flight in
852 paleognathous birds. *Science* 364:74-78.
853 Shashikant T, Khor JM, Ettensohn CA. 2018. Global analysis of primary mesenchyme cell cis-
854 regulatory modules by chromatin accessibility profiling. *BMC Genomics* 19(1):206.
855 Shibata Y, Sheffield NC, Fedrigo O, Babbitt CC, Wortham M, Tewari AK, London D, Song L,
856 Lee B-K, Iyer VR, et al. 2012. Extensive evolutionary changes in regulatory element
857 activity during human origins are associated with altered gene expression and
858 positive selection. *PLOS Genetics* 8:e1002789.
859 Siepel A, Haussler D. 2004. Phylogenetic estimation of context-dependent substitution rates by
860 maximum likelihood. *Molecular Biology & Evolution* 21:468-488.
861 Singh D, Yi SV. 2021. Enhancer Pleiotropy, Gene Expression, and the Architecture of Human
862 Enhancer-Gene Interactions. *Molecular Biology & Evolution* 38(9):3898-3909.
863 Smith MS, Turner FR, Raff RA. 2008. Nodal expression and heterochrony in the evolution of
864 dorsal-ventral and left-right axes formation in the direct-developing sea urchin
865 *Heliocidaris erythrogramma*. *Journal of Experimental Zoology Part B-Molecular and*
866 *Developmental Evolution* 310b:609-622.
867 Strathmann RR. 1978. The evolution and loss of feeding larval stages of marine invertebrates.
868 *Evolution* 32:894-906.
869 Strathmann RR. 1985. Feeding and nonfeeding larval development and life history evolution
870 in marine invertebrates. *Annual Review of Ecology and Systematics* 16:339-361.

871 Stern, DL. 2010. Evolution, Development, and the Predictable Genome. Austin, TX:
872 Roberts and Company Publishers.
873 Sun DA, Patel NH. 2021. Identification and classification of cis-regulatory elements in the
874 amphipod crustacean *Parhyale hawaiensis*. *bioRxiv* 2021.09.16.460328.
875 Swain-Lenz D, Berrio A, Safi A, Crawford GE, Wray GA. 2019. Comparative analyses of
876 chromatin landscape in white adipose tissue suggest humans may have less
877 beigeing potential than other primates. *Genome Biology and Evolution* 11:1997-2008.
878 Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC,
879 Stergachis AB, Wang H, Vernot B, et al. 2012. The accessible chromatin landscape of the
880 human genome. *Nature* 489:75-82.
881 Tonelli C, Chio IIC, Tuveson DA. 2018. Transcriptional Regulation by Nrf2. *Antioxidants &*
882 *Redox Signaling* 29:1727-1745.
883 Tsushima M, Byrne M, Amemiya S, Matsuno T. 1995. Comparative biochemical studies of
884 carotenoids in sea urchins. 3. Relationship between developmental mode and
885 carotenoids in the Australian echinoids *Heliocidaris erythrogamma* and *H. tuberculata*
886 and a comparison with japanese species. *Comparative Biochemistry and Physiology B-*
887 *Biochemistry & Molecular Biology* 110:719-723.
888 van der Burg KRL, Lewis JJ, Brack BJ, Fandino RA, Mazo-Vargas A, Reed RD. 2020. Genomic
889 architecture of a genetically assimilated seasonal color pattern. *Science* 370:721-725.
890 Vance RR. 1973. On the reproductive strategies of marine benthic invertebrates. *The American*
891 *Naturalist* 107:339-352.
892 Walton KD, Warner J, Hertzler PH, McClay DR. 2009. Hedgehog signaling patterns mesoderm
893 in the sea urchin. *Developmental Biology* 331:26-37.
894 Wang L, Israel JW, Edgar A, Raff RA, Raff EC, Byrne M, Wray GA. 2020. Genetic basis for
895 divergence in developmental gene expression in two closely related sea urchins. *Nature*
896 *Ecology & Evolution* 4:831-840.
897 Williams DHC, Anderson DT. 1975. Reproductive system, embryonic development, larval
898 development and metamorphosis of sea urchin H*eliocidaris erythrogramma* (Val)
899 (Echinoidea-Echinometridae). *Australian Journal of Zoology* 23:371-403.
900 Wilson KA, Andrews ME, Raff RA. 2005. Dissociation of expression patterns of homeodomain
901 transcription factors in the evolution of developmental mode in the sea urchins

902 *Heliocidaris tuberculata* and *H. erythrogramma*. *Evolution & Development* 7:401-415.
903 Wilson KA, Andrews ME, Rudolf Turner F, Raff RA. 2005. Major regulatory factors in the
904 evolution of development: the roles of goosecoid and Msx in the evolution of the direct-
905 developing sea urchin *Heliocidaris erythrogramma*. *Evolution & Development* 7:416-
906 428.
907 Wittkopp PJ, Kalay G. 2012. Cis-regulatory elements: molecular mechanisms and evolutionary
908 processes underlying divergence. Nature Reviews Genetics 13:59-69.
909 Wittkopp, P. J., Haerum, B. K. & Clark, A. G. 2008. Independent effects of *cis*- and *trans*-
910 regulatory variation on gene expression in *Drosophila melanogaster*. *Genetics* 178,
911 1831–1835.
912 Wray GA, Hahn MW, Abouheif E, Balhoff JP, Pizer M, Rockman MV, Romano LA. 2003. The
913 evolution of transcriptional regulation in eukaryotes. *Molecular Biology & Evolution*
914 20(9):1377-419.
915 Wray GA. 2007. The evolutionary significance of cis-regulatory mutations. *Nature Reviews*
916 *Genetics* 8:206-216.
917 Wray GA. 2022. Extreme phenotypic divergence and the evolution of development. *Current*
918 *Topics in Developmental Biology* 146:79-112.
919 Wray GA. 1996. Parallel evolution of nonfeeding larvae in echinoids. *Systematic Biology*
920 45:308-322.
921 Wray GA, Raff RA. 1989. Evolutionary modification of cell lineage in the direct-developing
922 sea urchin *Heliocidaris erythrogramma*. *Developmental Biology* 132:458-470.
923 Wray GA, Raff RA. 1990. Novel origins of lineage founder cells in the direct-developing sea
924 urchin *Heliocidaris erythrogramma*. *Developmental Biology* 141:41-54.
925 Yuh CH, Dorman ER, Davidson EH. 2005. Brn1/2/4, the predicted midgut regulator of the
926 endo16 gene of the sea urchin embryo. *Developmental Biology* 281:286-298.
927 Zaret KS, Carroll JS. 2011. Pioneer transcription factors: establishing competence for gene
928 expression. *Genes & Development* 25:2227-2241.
929 Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM,
930 Brown M, Li W. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biology*
931 9:R137.
932

933 **FIGURE LEGENDS**

934



935

936 **Figure 1: Global variation in chromatin accessibility between sea urchin embryos is largely**
937 **driven by developmental stage and life history strategy.** (*A*) Three focal sea urchin species of
938 this study, including two planktotrophic species representing the ancestral life history (light
939 green: *Lytechinus variegatus*; dark green: *Heliocidaris tuberculata*) and one lecithotrophic
940 species representing the derived life history (orange: *Heliocidaris erythrogramma*). This color
941 scheme is consistent throughout the study. (*B*) Developmental stages assayed in this study for
942 each species. The ancestral life history is illustrated above, with corresponding developmental

943 stages of the derived life history below. Developmental times are shown for *Heliocidaris*
944 *tuberculata* at 21°C. Note, however, samples were collected on the basis of developmental stage,
945 not hours post-fertilization, for all three species. (*C-D*) Principal component analysis of variation
946 in normalized chromatin accessibility among (*C*) all orthologous OCRs and (*D*) only OCRs near
947 dGRN genes through embryogenesis. OCR: open chromatin region; dGRN: developmental gene
948 regulatory network.
949

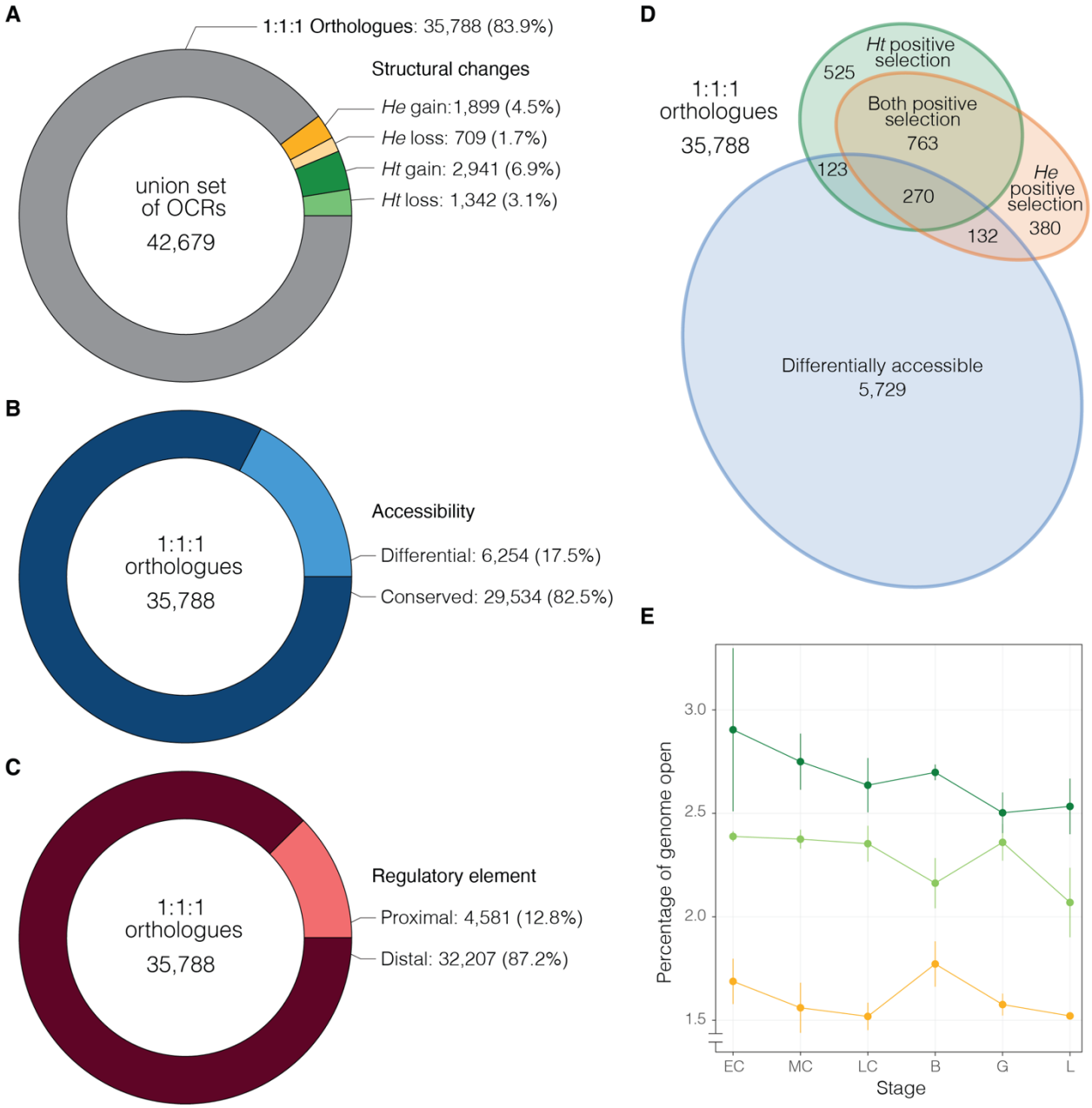


Figure 2: Summary statistics of regulatory element gains, losses, accessibility, and positive selection. (*A***)** Union set of all OCRs, including structural differences between species such as lineage-specific gains and losses as well as 35,788 1:1:1 orthologous OCRs shared between all three species. 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 models for these orthologous sites. (*D*) Overlap of OCR differential accessibility and evidence of positive selection on the *He*, *Ht*, or both branches. (*E*) Percentage of each species genome that is "open" at each stage of development. "Open" is defined as the cumulative portion of the genome

959 represented by all OCRs (orthologous and lineage-specific) with significant levels of
960 accessibility for a given stage and species (see **supplementary fig. 3** for additional information).
961 Error bars indicate standard deviation of % openness across biological replicates for each species
962 and developmental stage. OCR: open chromatin region; *He*: *Heliocidaris erythrogramma*; *Ht*:
963 *Heliocidaris tuberculata*.
964

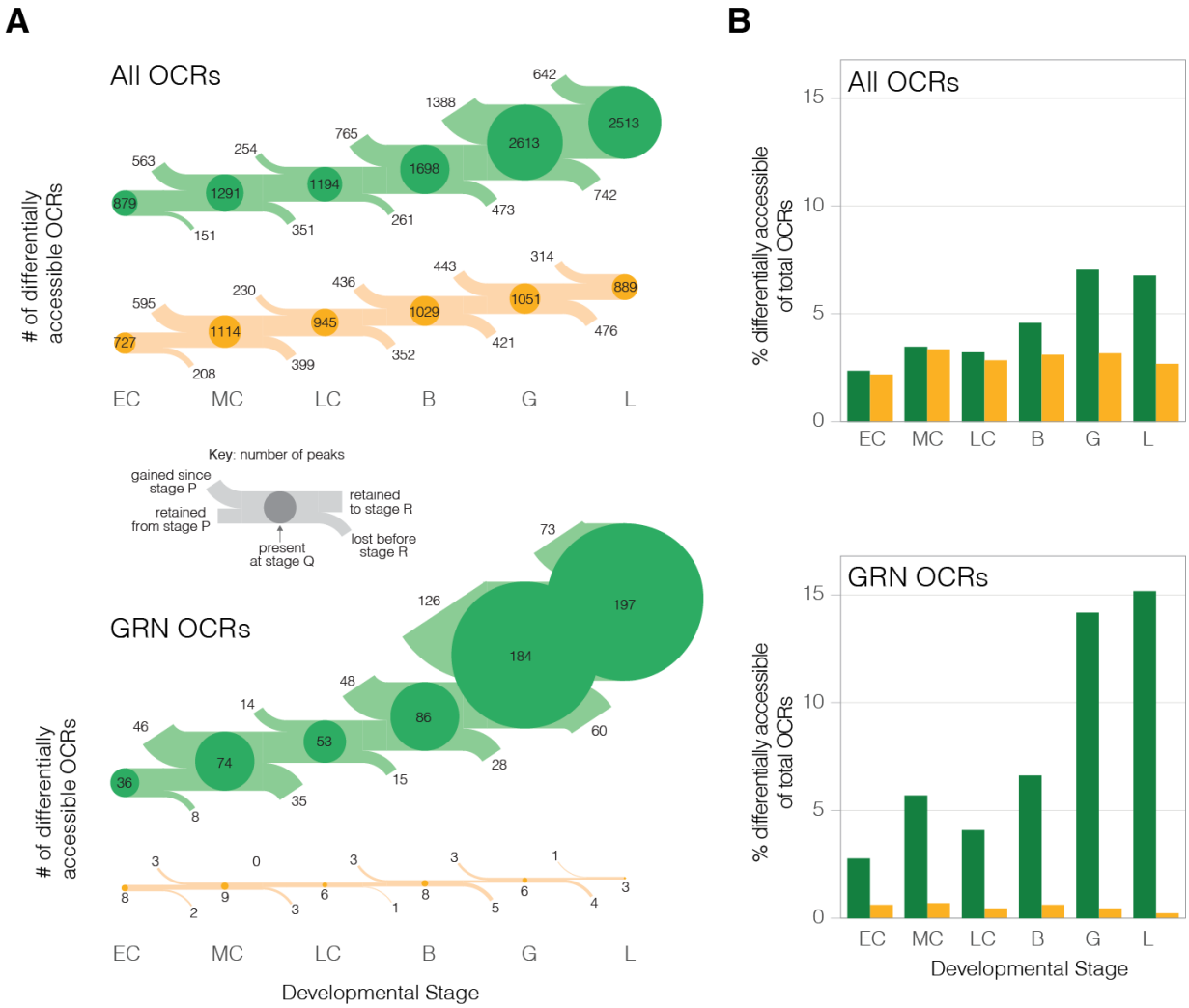


Figure 3: The chromatin landscape of the derived life history is far less accessible than that of species with the ancestral life history mode, particularly near core dGRN genes. (*A*) Number (flow diagram) and (*B*) percentage (barplots) of OCRs that are significantly differentially accessible in *H. erythrogramma* (orange) or species with the ancestral life history strategy (green) for all orthologous OCRs (top) and only OCRs near dGRN genes (bottom). If a differentially accessible OCR is more open in species with the ancestral developmental mode or in *H. erythrogramma*, it is included in the green or orange OCR sets, respectively. A key for flow diagrams is provided to explain how the number of differentially accessible OCRs are gained, retained, and lost between developmental stages. Circle size is proportional to the percentage of each set of OCRs (those near all genes or only dGRN genes) that are differentially accessible at a given stage. OCR: open chromatin region; dGRN: developmental gene regulatory network.

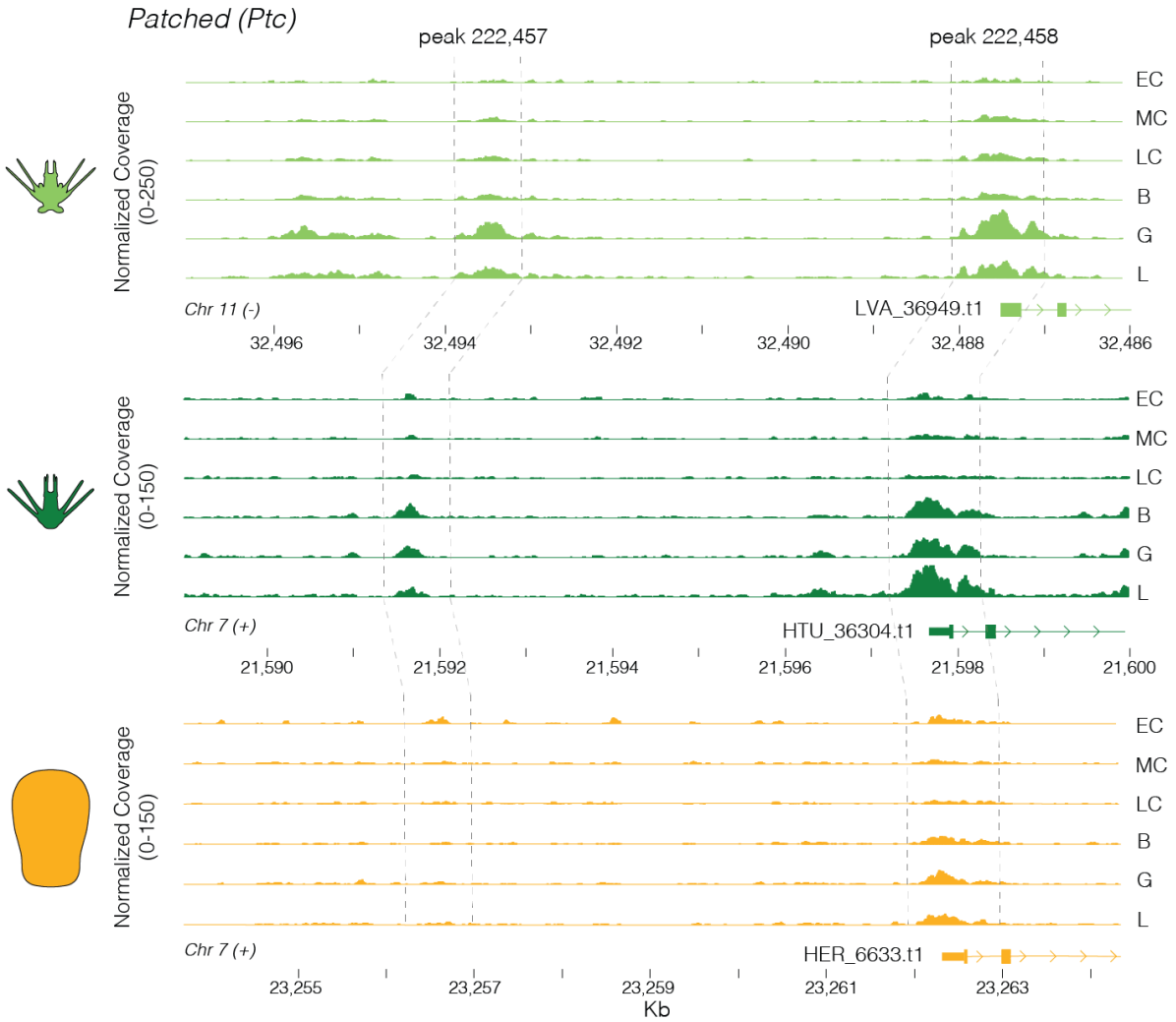


Figure 4: Example of an evolutionary change in chromatin accessibility during early development. The embryonic chromatin landscape near the gene *Patched* is shown for the three study species, with gene models on the right and dashed lines indicating orthologous noncoding regions. In the two planktotrophs, the most proximal OCR (likely promoter based on location at the 5' end of the 5' UTR) opens substantially after cleavage, and a distal OCR ~7kb upstream opens by gastrula stage. In in *H. erythrogramma*, the proximal OCR is substantially less accessible and the distal OCR is not above background accessibility levels.

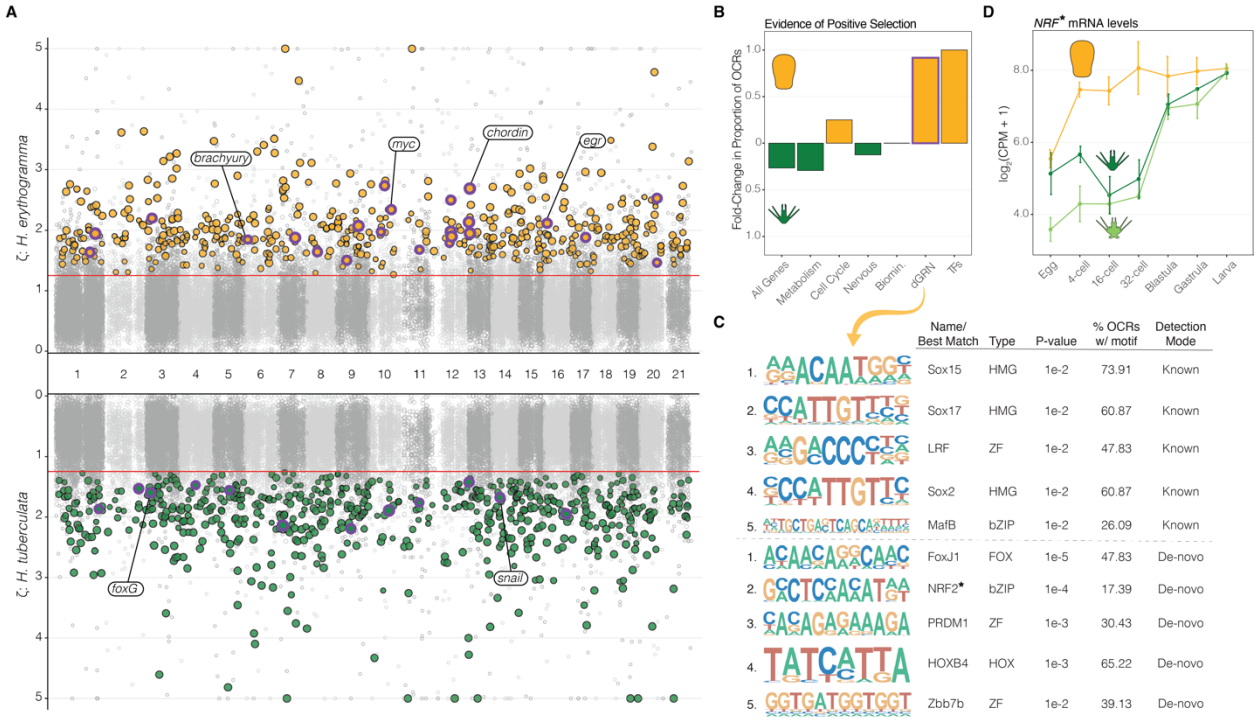


Figure 5: Sequence changes indicative of positive selection are enriched within *cis*-regulatory elements near dGRN genes and other transcription factors in the *H. erythrogramma* genome. (*A*) Zeta values (ratio of OCR substitution rate to that of neutrally evolving reference sequences) for *H. erythrogramma* (top) and *H. tuberculata* (bottom) across the genome. OCRs with significant support for lineage-specific evidence of positive selection are noted for *H. erythrogramma* (orange) and *H. tuberculata* (dark green), and those nearby dGRN genes are outlined in purple. (*B*) Fold-change in proportion of OCRs with more frequent evidence of positive selection in either *H. erythrogramma* or *H. tuberculata*, partitioned by functional categories of nearby genes (see **supplementary fig. 4** for all categories). (*C*) Transcription factor binding motif enrichment results of OCRs near dGRN genes with evidence of positive selection in *H. erythrogramma*. Detection mode indicates OCR sequences were searched against a (i) "known" set of optimized motif models or (ii) enriched motifs within the OCR were identified "de-novo" and matched to a known set of motifs. (*D*) Embryonic mRNA expression levels for *Nuclear factor erythroid 2-related factor* (*NRF*) for each species. OCR: open chromatin region; dGRN: developmental gene regulatory network.

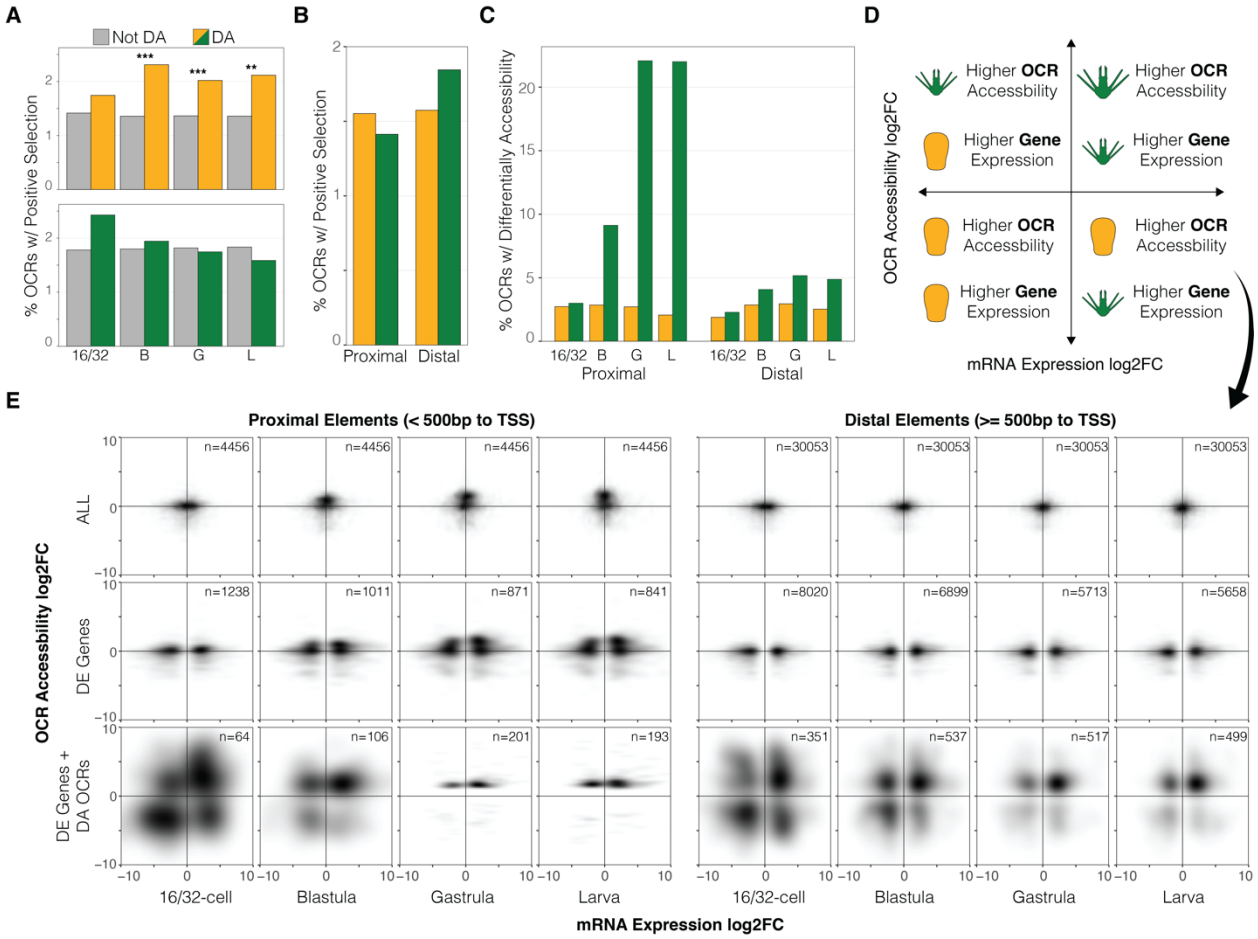


Figure 6: Independent and synergistic changes to regulatory element function have evolved during life history evolution in *H. erythrogramma*. (*A*) Proportion of differentially accessible or non-differentially accessible OCRs with evidence of positive selection on the *H. erythrogramma* branch (orange, top) or *H. tuberculata* branch (dark green, bottom). Two-sided, Fisher's exact test: ** $p \leq 5 \times 10^{-3}$; *** $p \leq 5 \times 10^{-4}$. (*B*) Proportion of proximal of distal OCRs with evidence of positive selection in either *Heliocidaris* species. (*C*) Differential accessibility of proximal and distal OCRs at four stages of development. Orange indicates significantly greater accessibility in *H. erythrogramma*, dark green indicates significantly greater accessibility in *H. tuberculata*. (*D*) Diagram describing possible correlative patterns between changes in OCR accessibility and gene expression in either life history strategy (green: ancestral, planktotroph; orange: derived, lecithotroph): the layout of this panel is a key for interpreting the density plots in panel (*E*) below. (*E*) Fold-change of mRNA expression vs nearby OCR accessibility between life history strategies for proximal (left) and distal (right) OCRs. Relationships are shown for all genes and OCRs (*E*: top row), significantly differentially expressed genes (*E*: middle row), and

1021 significantly differentially expressed genes *and* differentially accessible OCRs (*E*: bottom row).
1022 OCR: open chromatin region; DE: differentially expressed (genes); DA: differentially accessible
1023 (OCRs).
1024

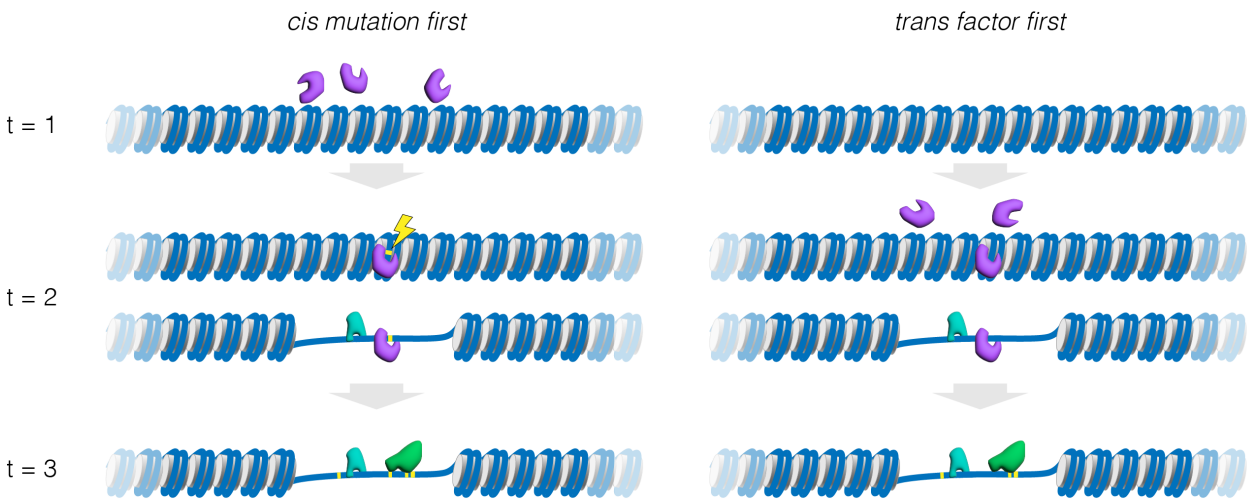


Figure 7. Mechanistic models for the evolutionary origin of an open chromatin region. Three distinct evolutionary time slices are illustrated: t=1 prior to the appearance of the open chromatin region (OCR), t=2 the event that precipitates local chromatin opening, and t=3 many generations after the origin of the OCR. Two general molecular processes could transform a given region of the genome from a relatively closed to a relatively open region of chromatin, based on the location of the genetic basis: *cis* or local and *trans* or distant. We illustrate this distinction based on a pioneer transcription factor, a class of proteins capable of sequence-specific binding to DNA in closed chromatin causing a local opening, typically on the scale of 1-2 nucleosomes (Zaret and Caroll, 2011). Note, however, that several other specific molecular mechanisms are also possible, involving chromatin remodeling complexes, histone acetylase/deacetylases, histone methylase/demethylases, and other proteins capable of acting to modify local chromatin configuration. (*A*) "*cis* mutation first" model. Here, the precipitating event is a local mutation (yellow; adjacent to the lightning bolt) that allows a pioneer transcription factor (purple protein) to bind to relatively closed chromatin. This pioneer factor was already present in the nucleus (see t=1), but can only bind locally following the new mutation that creates a binding site for it (t=2). Because pioneer factors can convert closed to open chromatin, other transcription factors also present in the nucleus (teal protein) may gain access to DNA within the newly opened region without the need for any additional mutations. (*B*) "*trans* factor first" model. Here, the pioneer factor is not originally present (t=1) in the nucleus at the developmental stage of interest. A mutation elsewhere in the genome (t=2) results in expression of the pioneer factor transcription factor, which can bind locally due to a pre-existing recognition site and thus without the requirement for a local mutation. Once present, the

1048 pioneer factor can open the local DNA, potentially allowing other transcription factors (teal
1049 protein) to bind without the need for any local mutations. Any impact on the transcription of a
1050 local gene could be susceptible to natural selection, since it has a genetic basis and a trait
1051 consequence. Over extended evolutionary timescales, additional mutations could accumulate
1052 within the OCR (t=3), potentially altering the suite of interacting transcription factors (for
1053 example, green protein now binds, purple protein no longer binds). In exceptional cases, natural
1054 selection may favor fixation of several new mutations within the OCR that alter gene expression.
1055 This figure illustrates the evolutionary origin of an OCR, but the evolutionary loss of an OCR
1056 could similarly occur through distinct molecular mechanisms based on the location of the genetic
1057 basis: for example, loss of a binding site for a transcription factor or loss of expression of a
1058 pioneer factor.
1059

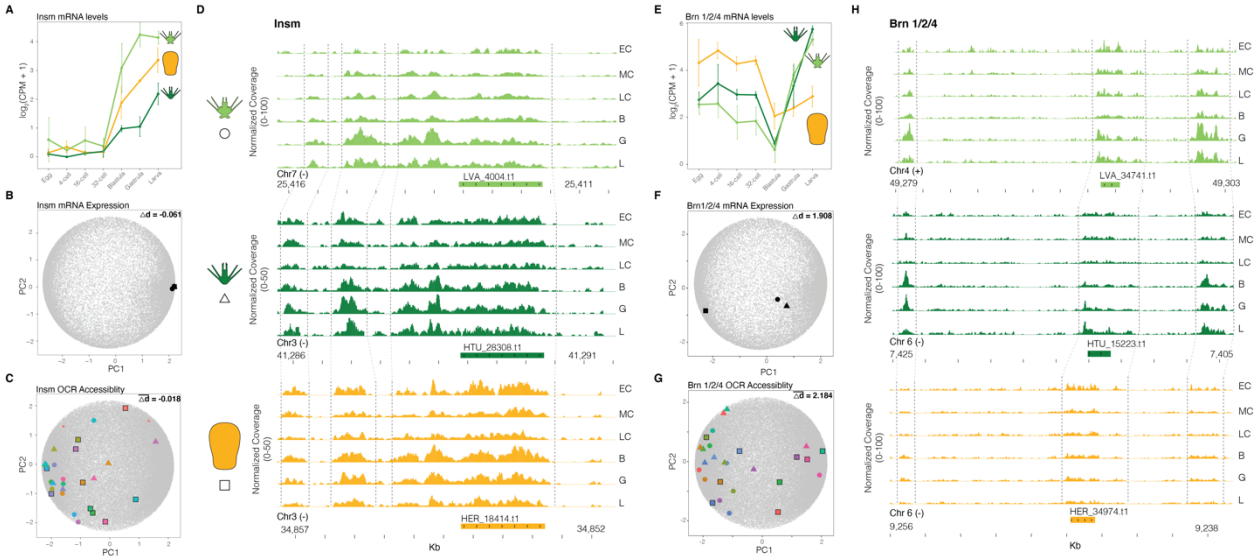
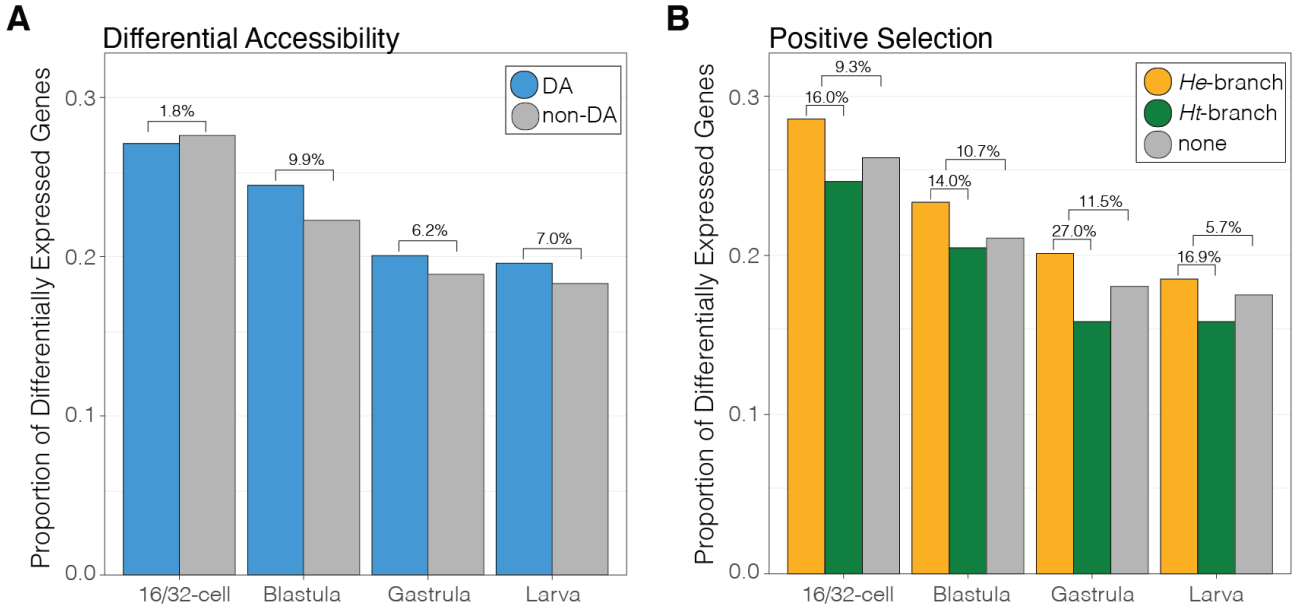


Figure 8: Quantifying developmental expression and accessibility changes in principal component space predicts divergent patterns of gene regulation. (*A-D*) An example of conserved developmental gene expression and chromatin accessibility for the gene *insm*. (*A*) Developmental mRNA expression profile 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 the gene *brn 1/2/4*. (*E*) Developmental mRNA expression profile of *brn 1/2/4* for each sea urchin species and (*F*) their coordinates in PCA space. (*G*) Coordinates in PCA space of developmental accessibility of orthologous OCRs near *brn 1/2/4* and (*H*) normalized chromatin accessibility coverage of these same OCRs through embryogenesis for each species. See Supplementary Figure 6 for a summary figure of PCA distances for all genes and OCRs. OCR: open chromatin region; PCA: principal component analysis.

1077

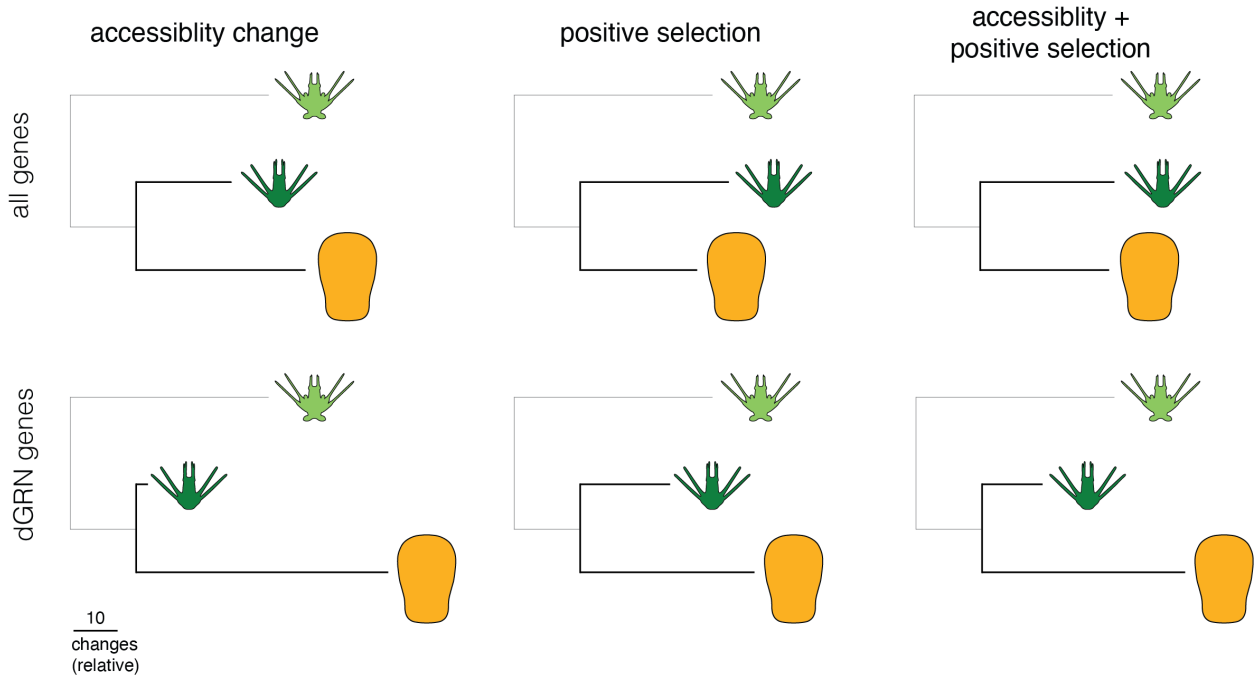


1078

1079 **Figure 9: Evolutionary changes to *cis*-regulatory element function associated with life**
1080 **history evolution in sea urchins and expression changes of nearby genes.** Proportion of genes
1081 that are differentially expressed between developmental life history strategies at four
1082 developmental stages, distinguished as having at least one nearby OCR with (*A*) differential
1083 accessibility or (*B*) evidence of lineage-specific positive selection. OCR: open chromatin region.
1084 Values above brackets indicate the percent change in proportion of genes that are differentially
1085 expressed between categories.

1086

1087



1088

1089 **Figure 10: Chromatin accessibility changes, mutational signatures of positive selection, and**
1090 **the coincidence of both evolutionary changes are abundant in the *H. erythrogramma***
1091 **genome, especially nearby dGRN genes.** Enrichment of regulatory modifications to OCRs near
1092 dGRN genes suggests selective pressures imposed by changes in developmental life history can
1093 rapidly modify *cis*-regulatory interactions underlying developmental processes. Tandem
1094 accessibility and mutational changes may act as a common gene regulatory method to efficiently
1095 modify gene expression for organismal adaptation and innovation. Branch lengths scaled to
1096 relative number of changes between *Heliocidaris* branches for each set of modifications.