bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Convergent evolution of gene expression in two high-toothed stickleback populations

James C. Hart[1], Nicholas A. Ellis[1], Michael B. Eisen[1,2], and Craig T. Millller[1]

[1]Department of Molecular and Cell Biology, University of California-Berkeley, CA, USA

[2] Howard Hughes Medical Institute, University of California, Berkeley, CA, USA

*address correspondence to ctmiller@berkeley.edu

short title: Convergent evolution of gene expression

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

Changes in developmental gene regulatory networks enable evolved changes in morphology. These changes can be in *cis* regulatory elements that act in an allele-specific manner, or changes to the overall *trans* regulatory environment that interacts with *cis* regulatory sequences. Here we address several questions about the evolution of gene expression accompanying a convergently evolved constructive morphological trait, increases in tooth number in two independently derived freshwater populations of threespine stickleback fish (*Gasterosteus aculeatus*). Are convergently evolved *cis* and/or *trans* changes in gene expression associated with convergently evolved morphological evolution? Do *cis* or *trans* regulatory changes contribute more to the evolutionary gain of a morphological trait? Transcriptome data from dental tissue of ancestral low-toothed and two independently derived high-toothed stickleback populations revealed significantly shared gene expression changes that have convergently evolved in the two high-toothed populations. Comparing *cis* and *trans* regulatory changes using phased gene expression data from F1 hybrids, we found that *trans* regulatory changes were predominant and more likely to be shared among both high-toothed populations. In contrast, while *cis* regulatory changes have evolved in both high-toothed populations, overall these changes were distinct and not shared among high-toothed populations. Together these data suggest that a convergently evolved trait can occur through genetically distinct regulatory changes that converge on similar *trans* regulatory environments.

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Author Summary

Convergent evolution, where a similar trait evolves in different lineages, provides an opportunity to study the repeatability of evolution. Convergent morphological evolution has been well studied at multiple evolutionary time scales ranging from ancient, to recent, such as the gain in tooth number in freshwater stickleback fish. However, much less is known about the accompanying evolved changes in gene regulation during convergent evolution. Here we compared evolved changes in gene expression in dental tissue of ancestral low-toothed marine fish to fish from two independently derived high-toothed freshwater populations. We also partitioned gene expression changes into those affecting a gene's regulatory elements (*cis*), and those affecting the overall regulatory environment (*trans*). Both freshwater populations have evolved similar gene expression changes, including a gain of expression of putative dental genes. These similar gene expression changes are due mainly to shared changes to the *trans* regulatory environment, while the *cis* changes are largely population specific. Thus, during convergent evolution, overall similar and perhaps predictable transcriptome changes can evolve despite largely different underlying genetic bases.

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Introduction

Development is controlled by a complex series of interlocking gene regulatory networks. Much of this regulation occurs at the level of transcription initiation, where *trans* acting factors bind to *cis* regulatory elements to control their target gene's expression [1,2]. Evolved changes in an organism's morphology are the result of changes in this developmental regulatory landscape. It has been proposed that the genetic bases of many of these evolved changes are mutations within the *cis*-regulatory elements of genes [3–5]. Indeed, recent work in evolutionary genetics suggests the molecular bases of a diverse array of traits from *Drosophila* wing spots [6] to mouse pigmentation [7] to stickleback armored plate number [8,9] and size [10] are changes in the activity of *cis*-regulatory elements.

Evolved changes in gene expression can be divided into two broad regulatory classes. *Cis* regulatory changes occur within the proximal promoter [11], distal enhancer [12], or the gene body itself [13]. *Trans* regulatory changes modify the overall regulatory environment [14,15], but are genetically unlinked to the expression change. The total evolved gene expression differences can be partitioned into changes in *cis* and *trans* by quantifying expression differences between two populations and also testing for expression differences between alleles in F1 hybrids between the two populations [16]. Several studies have attempted to characterize evolved *cis* and *trans*-regulatory changes at a transcriptome-wide level [17–21]. Though the relative contribution of *cis* and *trans* regulatory changes varies extensively among studies, *cis* changes have been found to dominate [17,18,21] or at least be approximately equivalent [19,20] to *trans* changes [22]. Additionally, compensatory changes (*cis* and *trans* changes in opposing directions) have been found to be enriched over neutral models [17,18], showing evidence for selection for stable gene expression levels. However, none of these studies examined contribution of *cis* and *trans* gene expression

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

97 changes during convergent morphological evolution.

98 Populations evolve new traits following a shift to a novel environment, due to a mixture
99 of drift and selection. Truly adaptive traits can often be repeatedly observed in multiple
100 populations following a similar ecological shift. Threespine sticklebacks are an excellent
101 system for the study of evolved changes in phenotypes, including gene expression [23–27].
102 Marine sticklebacks have repeatedly colonized freshwater lakes and streams along the coasts
103 of the Northern hemisphere [28]. Each of these freshwater populations has independently
104 adapted to its new environment; however, several morphological changes, including a loss in
105 armored plates and a gain in tooth number, are shared among multiple newly derived
106 populations [29,30]. The repeated evolution of lateral plate loss is due to repeated selection of
107 a standing variant regulatory allele of the *Eda* gene within marine populations [8,9] and
108 genome sequencing studies found over a hundred other shared standing variant alleles
109 present in geographically diverse freshwater populations [31]. These studies suggest the
110 genetic basis of freshwater adaptation might typically involve repeated reuse of the same
111 standing variants to evolve the same adaptive freshwater phenotype.

112 However, more recent evidence has shown that similar traits have also evolved
113 through different genetic means in freshwater stickleback populations. A recent study which
114 mapped the genetic basis of a gain in pharyngeal tooth number in two independently derived
115 freshwater populations showed a largely non-overlapping genetic architecture [30]. Another
116 study using three different independently derived benthic (adapted to the bottom of a lake)
117 populations showed that, even when adapting to geographically and ecologically similar
118 environments, the genetic architecture of evolved traits is a mix of shared and unique
119 changes [32]. Even in cases where the same gene is targeted by evolution in multiple
120 populations (the loss of *Pitx1* expression resulting in a reduction in pelvic spines), the

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

121 individual mutations are often independently derived [33,34]. All of these genomic scale
122 studies have looked at the genetic control of morphological changes, while the extent and
123 nature of genome-wide gene expression changes has been less studied. It remains an open
124 question as to whether similar gene expression patterns evolve during the convergent
125 evolution of morphology, and if so, to what extent those potential shared gene expression
126 changes are due to shared *cis* or *trans* changes.

127 Teeth belong to a class of vertebrate epithelial appendages (including mammalian hair)
128 that develop from placodes, and have long served as a model system for studying
129 organogenesis and epithelial-mesenchymal interactions in vertebrates [35]. Odontogenesis is
130 initiated and controlled by complex interactions between epithelial and mesenchymal cell
131 layers, and involves several deeply conserved signaling pathways [36–38]. Sticklebacks retain
132 the ancestral jawed vertebrate condition of polyphyodonty, or continuous tooth replacement,
133 and offer an emergent model system for studying tooth replacement. Previous work has
134 supported the hypothesis that two independently derived freshwater stickleback populations
135 have evolved an increase in tooth replacement rate, potentially mediated through differential
136 odontogenic stem cell dynamics [30] (Cleves et al 2018 under review, see Supplementary
137 Data File 2). Recent studies have found teeth and taste bud development to be linked, with
138 one study supporting a model where teeth and taste buds are copatterned from a shared oral
139 epithelial source [39], and another study supporting a model where teeth and taste buds
140 share a common progenitor stem cell pool [40].

141 We sought to examine the evolution of the regulatory landscape controlling stickleback
142 tooth development and replacement. Using high-throughput RNA sequencing (RNA-seq), we
143 found that two independently derived high-toothed freshwater populations display highly
144 convergent gene expression changes, especially in orthologs of known tooth-expressed

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

145 genes in other vertebrates, likely reflecting the convergently evolved tooth gain phenotype
146 and the deep homology of teeth across all jawed vertebrates. We also quantitatively
147 partitioned these evolved gene expression changes into *cis* and *trans* regulatory changes
148 [16,19] in both populations at a transcriptome-wide level using RNA-seq on F1 marine-
149 freshwater hybrids. We found that *trans* regulatory changes predominate evolved changes in
150 gene expression in dental tissue. Additionally, we found that the *trans* regulatory changes are
151 more likely to be shared between the freshwater populations than the *cis* regulatory changes.
152 Thus, similar downstream transcription networks controlling tooth development and
153 replacement have convergently evolved largely through different upstream genetic regulatory
154 changes.
155

156 ## Results

157

158 **Convergent evolution of tooth gain in two freshwater populations**

159 To further test whether multiple freshwater populations have evolved increases in tooth
160 number compared to multiple ancestral marine populations [30,41], we quantified total ventral
161 pharyngeal tooth number of lab reared sticklebacks from four distinct populations: (1) a
162 marine population from the Little Campbell river (LITC) in British Columbia, Canada, (2) a
163 second marine population from Rabbit Slough (RABS) in Alaska, (3) a benthic freshwater
164 population from Paxton Lake (PAXB) in British Columbia, Canada, USA, and (4) a second
165 freshwater population from Cerrito Creek (CERC) in California, USA (Fig 1A, 1B). Freshwater
166 fish from both populations had more pharyngeal teeth than marine fish at this 35-50mm
167 standard length (SL) stage, consistent with previous findings [30,41] of increases in tooth
168 number in freshwater sticklebacks (Fig 1B, 1C, Table S1).

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

To estimate the genomic relatedness of these populations, we resequenced the genomes of three marine and six freshwater sticklebacks from the four different populations (Table S2). We aligned the resulting reads to the stickleback reference genome [31] using Bowtie2 [42], and called variants using the Genome Analysis Toolkit (GATK) [43–45]. As it has been previously shown that Pacific marine stickleback populations are an outgroup to freshwater populations from Canada and California [31], we hypothesized the two high-toothed populations would be more related to each other genomically than either marine population. A maximum-likelihood phylogeny constructed using genome-wide variant data cleanly separated freshwater populations from each other and from marine fish (Fig S1A). Principal component analysis of the genome-wide variants revealed that the first principle component explains nearly half (41.4%) of the overall variance, and separates benthic sticklebacks from both creek and marine fish (Fig S1B). The second principal component separated both freshwater populations from marine populations. These results further support the model that populations of freshwater sticklebacks used a combination of shared and independent genetic changes [31,32] when evolving a set of similar morphological changes in response to a new environment.

Convergent evolution of gene expression

As morphological changes are often the result of changes in gene expression patterns and levels, we sought to identify evolved changes in gene expression during tooth development at stages soon after the evolved differences emerge [41]. We quantified gene expression in ventral pharyngeal dental tissue in the two high-toothed freshwater and an Alaskan low-toothed marine population using RNA-seq (Fig 2A, Table S3-S4). Principal component (PC) analysis of the resulting gene expression matrix showed a clustering of gene expression by

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

193 population, with the first PC separating benthic samples, and the second PC separating both
194 benthic and creek samples from marine, similar to the PC analysis of the genome-wide
195 variants (Fig 2B) [46].

196 Given the convergently evolved morphological change of increases in tooth number,
197 we hypothesized that convergent evolution has occurred at the gene expression level in
198 freshwater dental tissue. To test this hypothesis, we compared the evolved change in gene
199 expression in benthic dental tissue (benthic vs marine) to the evolved change in creek dental
200 tissue (creek vs marine). At a genome-wide level, correlated changes in gene expression
201 levels have evolved in the two high-toothed freshwater populations (Fig 2C, Spearman's r =
202 0.43). We next asked if orthologs of genes implicated in tooth development in other
203 vertebrates showed an increase in correlated evolved expression changes. We compared the
204 gene expression changes of stickleback orthologs of genes in the BiteIt (http://bite-
205 it.helsinki.fi/) [47] or ToothCODE (http://compbio.med.harvard.edu/ToothCODE/) [36]
206 databases (hereafter referred to as the “BiteCode” gene set, Table S5), two databases of
207 genes implicated in mammalian tooth development. Consistent with the conserved roles of
208 gene regulatory networks regulating mammalian and fish teeth [48–51] and the major evolved
209 increases in tooth number in both freshwater populations (Fig 1C), these predicted dental
210 genes showed an increase in their correlated evolved gene expression change (Fig 2C red
211 points, Spearman's r = 0.68), and tended to have an overall increase in gene expression (Fig
212 S2, *P* = 7.36e-6, GSEA, see methods). We also examined the expression levels of genes
213 whose orthologs are annotated as being expressed in zebrafish pharyngeal teeth
214 (www.zfin.org). Within this gene set, 27 of 40 genes were significantly more highly expressed
215 in at least one freshwater population, with no genes expressed significantly higher (as
216 determined by cuffdiff2 [52–55], see Materials and Methods) in marine samples than either

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

217 freshwater population (Fig 2D).

218

219 **Increased freshwater expression of stem cell maintenance genes**

220 Tooth development is controlled by several deeply conserved developmental signaling
221 pathways [49,51]. To test whether expression changes in the components of specific
222 developmental signaling pathways have evolved in the two high-toothed freshwater
223 populations, we next analyzed the expression levels of stickleback orthologs of genes
224 implicated in mammalian tooth development and annotated as components of different
225 signaling pathways [36]. When comparing gene expression levels in freshwater dental tissue
226 to marine dental tissue, genes annotated as part of the TGF-ß signaling pathway displayed
227 significantly increased expression in freshwater dental tissue (Fig S3A-F).

228 Since these two freshwater populations have a largely different developmental genetic
229 basis for their evolved tooth gain [30], we next asked whether any pathways were upregulated
230 or downregulated specifically in one freshwater population. When comparing the expression
231 of genes in benthic dental tissue to expression in creek or marine dental tissue, genes not
232 only in the TGF-ß pathway, but also in the WNT signaling pathway, displayed significantly
233 increased expression, consistent with the differing genetic basis of tooth gain in these
234 populations (Fig S3B). Genes upregulated in freshwater dental tissue were enriched for Gene
235 Ontology (GO) terms involved in anatomical structure development, signaling, and regulation
236 of cell proliferation (Fig S4A, Table S6). Genes upregulated in benthic dental tissue over
237 marine were enriched for GO terms involved in cell proliferation, division and cell cycle
238 regulation, as well as DNA replication (Fig S4B, Table S7), while genes upregulated in creek
239 over marine were enriched for GO terms involved in cell locomotion, movement, and
240 response to lipids (Fig S4C, Table S8).

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

As teeth are constantly being replaced in polyphyodont adult fish, potentially due to the action of dental stem cells [40], we hypothesized that genes involved in stem cell maintenance have evolved increased expression in freshwater tooth plates, given the higher rate of newly forming teeth previously found in adults [30], and the possibly greater number of stem cell niches in high-toothed fish (Cleves et al 2018 under review, see Supplemental Data File 2). We further hypothesized that since teeth are developmentally homologous to hair, perhaps an ancient genetic circuit regulating vertebrate placode replacement controls both fish tooth and mammalian hair replacement. For example, the *Bmp6* gene, previously described as expressed in all stickleback teeth [41] was significantly upregulated in creek fish, consistent with the evolved major increases in tooth number in this population (Table S4). In contrast, no such significant upregulation was observed in the expression of benthic *Bmp6* (Table S4), consistent with the observed evolved *cis*-regulatory decrease in benthic *Bmp6* expression [41]. Further supporting this hypothesis, the expression of the stickleback orthologs of a previously published set of mouse hair follicle stem cell (HFSC) signature genes [56] were significantly upregulated in freshwater dental tissue (Fig S3A). Creek dental tissue displayed a small but significant increase in expression of this set of HFSC orthologs relative to both benthic and marine samples (Fig S3C).

In cichlid fish, pharmacology experiments revealed that reductions in tooth density can be accompanied by concomitant increases or decreases in taste bud density [39]. To begin to test whether derived high-toothed stickleback populations have also evolved significantly altered levels of known taste bud marker gene expression, we examined the expression levels of known taste bud markers *Calbindin2* and *Phospholipase Beta 2* [57], as well as taste receptors such as *Taste 1 Receptor Member 1*, *Taste 1 Receptor Member 3*, and *Polycystin 2 Like 1* [58]. Although four of these five genes had detectable significant expression changes

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

between different populations, no consistent freshwater upregulation or downregulation of taste bud marker genes was seen (Fig S5).

Cis and *trans* regulatory changes in gene expression

Evolved changes in gene expression are due to a combination of *cis* acting changes that are linked to the genes they act on, and *trans* acting changes which usually are genetically unlinked to the gene or genes they regulate. Since the genetic basis of freshwater tooth gain mapped to non-overlapping intervals in these two populations [30] (Cleves et al 2018 under review, see Supplemental Data File 2), we hypothesized that the observed shared freshwater gene expression changes were the result of a similar *trans* environment, but a largely different set of *cis* changes. To test this hypothesis, we measured evolved *cis* expression changes in marine-freshwater F1 hybrids, which have marine and freshwater alleles present in the same *trans* environment. We raised both creek-marine and benthic-marine F1 hybrids to the late juvenile stage, dissected their ventral pharyngeal tooth plates, then generated and sequenced five barcoded RNA-seq libraries per population (10 total). We then quantified the *cis* expression change as the ratio of the number of reads mapping uniquely to the freshwater allele of a gene to the number of uniquely mapping marine reads (Fig 3A, Table S9-11). *Trans* expression changes were calculated by factoring the *cis* change out from the overall parental expression change [19].

We found 11,832 and 8,990 genes in benthic and creek F1 hybrids, respectively, that had a fixed marine-freshwater sequence difference which had more than 20 total reads mapping to it. We observed no significant bias towards either the marine or freshwater allele in either set of F1 hybrids (Fig 3B). We next classified genes into one of four categories (*cis* change only, *trans* change only, concordant *cis* and *trans* changes, discordant *cis* and *trans*

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

289 changes). We found 1640 and 1116 benthic (Fig. 3C) and creek (Fig. 3D) genes, respectively,
290 with only significant *cis* changes, and 1873 and 1048 genes, respectively, with only significant
291 *trans* changes. We also found 478 and 359 genes with significant *cis* and *trans* changes in
292 the same direction, which we term concordant changes in gene expression. Conversely, we
293 found 772 and 607 genes with significant *cis* and *trans* changes in opposing directions, which
294 we termed discordant changes. Thus, overall, *trans* regulatory changes are more common
295 than *cis* changes in the evolution of dental tissue gene expression in both freshwater
296 populations. Additionally, discordant *cis* and *trans* changes were more common in both
297 populations, suggesting selection for stable levels of gene expression.
298
299 **_Trans_ regulatory changes dominate**
300 We next wanted to determine the relative contribution of *cis* and *trans* gene expression
301 changes to evolved changes in gene expression. We restricted our analysis to differentially
302 expressed genes (as determined by cuffdiff2 [52]) to examine only genes with a significant
303 evolved difference in gene expression and quantifiable (i.e. genes with transcripts containing
304 a polymorphic variant covered by at least 20 reads) *cis* and *trans* expression changes. When
305 evolving a change in gene expression, the *cis* and *trans* regulatory basis for this change can
306 be concordant (*cis* and *trans* effects both increase or decrease expression) or discordant (*cis*
307 effects increase and *trans* decrease or vice versa). We hypothesized that genes would tend to
308 display more discordant expression changes, as stabilizing selection has been found to buffer
309 gene expression levels [17,22,59]. To test this hypothesis, we binned genes into a 2x2
310 contingency table, with genes classified as *cis* or *trans* based on which effect controlled the
311 majority of the evolved expression change, and discordant or concordant based on the
312 direction of the *cis* and *trans* changes (Fig 4A, B). In the creek population, significantly more

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

313 discordant changes than expected by a neutral model (*P* =1.35e-7, binomial test) have
314 evolved. In both populations, we found increased discordant changes when the *trans* effect is
315 larger than the *cis* effect (*P* =1.29e-7, 1.44e-13, benthic and creek respectively, binomial test).
316 In both populations, we observe the opposite (an enrichment of concordant changes) when
317 the *cis* effect is stronger, relative to the ratio when the *trans* effect is dominant (*P* =1.34e-36,
318 8.2e-11 benthic and creek respectively, binomial test). When considering all (not just
319 differentially expressed) genes with quantifiable *cis* and *trans* expression changes, discordant
320 changes dominated regardless of the relative strength of the *cis* effect (Fig S6).
321 If all gene expression changes were due to changes only in *cis*, we would expect to
322 see the measured *cis* ratios in the hybrids match the parental expression ratios. Instead, in
323 both cases of evolved change, we saw parental expression ratios of a greater magnitude than
324 F1 hybrid ratios, indicating a stronger contribution of *trans* changes to overall gene expression
325 changes (Fig 3C-D). Indeed, when we examined the overall percentage of expression
326 changes of differentially expressed genes that were due to changes in *cis*, we observed
327 median per gene values of only 25.2% and 32.5% of benthic and creek gene expression
328 changes, respectively (Fig 4C). Comparing the expression levels of orthologs of known
329 dentally expressed genes from the BiteIt [47] and ToothCODE [36] databases revealed a
330 similarly small number of gene expression changes explained by changes in *cis*, relative to
331 the genome-wide average (Fig 4D). Evolved changes in creek gene expression were more
332 due to changes in *cis* than benthic genes (Fig 4D, *P* = 1.25e-22, Mann-Whitney U test). Thus,
333 *trans* effects on gene expression dominate the evolved freshwater gene expression changes.
334
335 **_Trans_ regulatory changes are more likely to be shared between freshwater populations**
336 We next wanted to test the hypothesis that the shared freshwater gene expression changes

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

were primarily due to shared *trans* changes, rather than shared *cis* changes. We first compared the overall expression levels of genes called differentially expressed between benthic and marine as well as creek and marine. Similar to the genome-wide comparison, we found a highly significant non-parametric correlation coefficient (Spearman's r = 0.62, *P* =1.2e-132) for the expression change of these shared differentially expressed genes (Fig 5A). When comparing the benthic *cis* changes to the creek *cis* changes, however, we found a much lower (though still significant) correlation coefficient (Spearman's r = 0.13, *P* =5.1e-6) (Fig 5B). When comparing the calculated *trans* changes for these shared differentially expressed genes, we observed much higher correlation coefficient (Spearman's r = 0.51, *P* =1.2e-80) (Fig 5C). When comparing all, not just differentially expressed, genes, *trans* changes are still likely to be more shared than *cis* (Fig S7). Additionally, 35/38 of the shared differentially expressed putative dental genes have shared regulatory increases or decreases in both freshwater populations relative to marine in overall expression difference, with 32/38 in *trans*, but only 25/38 in *cis* (Fig 5D-I). Thus, the *trans* effects on evolved gene expression are more likely to be shared by both freshwater populations than the *cis* changes.

Discussion

We sought to test the relative contribution of *cis* and *trans* gene regulatory changes during convergent evolution of tooth gain, as well as to ask whether the same or different regulatory changes underlie evolved changes in gene expression during this case of convergent evolution. We quantified the overall regulatory divergence, as well as the specific contribution of *cis* and *trans* changes, between ancestral low-toothed marine and two different independently derived populations of high-toothed freshwater sticklebacks. Similar overall changes in gene expression have evolved in both freshwater populations, especially in

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

orthologs of known dental regulators in mammals. In this system, *trans*-regulatory changes play a larger role than *cis* changes in both populations. Furthermore, *trans* acting changes were much more likely to be shared between freshwater populations than *cis* changes, suggesting the two high-toothed populations evolved their similar gene expression patterns through independent genetic changes.

Convergent evolution of dental gene expression

Convergent evolution at the gene expression level occurs when similar gene expression levels evolve in different populations. Both the creek and benthic stickleback populations have adapted from an ancestral marine form to their current freshwater environments. The genomic nature of their derived changes appears largely divergent, with major axis of variation separating benthic genomes from the geographically proximal marine populations (LITC), as well as the more distant marine (RABS) and creek populations. However, when looking at the gene expression basis of their convergently evolved gain in tooth number, orthologs of genes implicated in mammalian dental development showed strong correlated freshwater gains in expression. This correlation suggests both that sticklebacks deploy conserved genetic circuits regulating tooth formation during tooth replacement, but also that both populations have convergently evolved changes to similar downstream transcriptional circuits resulting in a gain of tooth number.

Though both freshwater populations showed strongly correlated changes in evolved gene expression at the *trans* regulatory level, the *cis* changes were largely not shared across populations. This was especially true for putative dentally expressed genes with evolved expression changes – the vast majority of the *trans* but not *cis* expression changes were shared between both freshwater populations. This suggests that the similar freshwater gene

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

385 expression patterns evolved through independent genetic changes. It is possible that the
386 small number of shared *cis* changes are sufficient to drive the observed changes to the
387 overall *trans* regulatory environments. However previous work has shown that the genetic
388 basis of tooth gain in these two populations is distinct [30] (Cleves et al 2018 under review,
389 see Supplemental Data File 2), and it seems parsimonious that the genetic basis of a gain in
390 dental gene expression is also mostly independent. Thus, convergent freshwater gene
391 expression changes appear to be largely due to distinct, independent population-specific
392 regulatory changes. This finding suggests that there are many regulatory alleles that are
393 accessible during the evolution of an adaptive trait.
394
395 **_Trans_ effects dominate**
396 Other studies have used RNA-seq to compare the relative contribution of *cis* and *trans*-
397 regulatory changes in the evolution of gene expression. In mice, evolved gene expression
398 changes in the liver [18] and the retina [60] were driven primarily by *cis*-regulatory changes. In
399 *Drosophila*, work on organismal-wide evolved gene expression changes on the genome-wide
400 level has shown the opposite, with *trans*-regulatory effects playing a larger role in the
401 evolution of gene expression [19,22]. Other studies have found *trans* effects contribute more
402 to intraspecific comparisons, while *cis* effects contribute more to interspecific comparisons
403 [61]. Consistent with this, we observe *trans* effects dominating in both of our intraspecific
404 comparisons.
405 Another key distinction could be that *cis*-regulatory effects dominate when looking at
406 more cellularly homogenous tissues, while *trans*-regulatory effects dominate when looking at
407 more heterogeneous tissues. Stickleback tooth plates likely fall into an intermediate category,
408 less heterogenous in cell type composition than a full adult fly or fly head, but more

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

409 heterogeneous than a specialized tissue such as the mouse retina. Overall, freshwater tooth
410 plates are more morphologically similar to each other than marine, with freshwater tooth
411 plates possessing a larger area, increased tooth number, and decreased intertooth spacing
412 [30,41]. Freshwater tooth plates likely have more similar cell type abundances and
413 compositions (e.g. more developing tooth germs with inner and outer dental epithelia, and
414 odontogenic mesenchyme) compared to each other than to marine tooth plates. Similar cell
415 types tend to have similar gene expression patterns, even when compared across different
416 species [62]. Much of the shared freshwater increase in dental gene expression could be due
417 to an increase in dental cell types in both freshwater populations. As other evolved changes
418 to stickleback morphology have been shown to be due to *cis* regulatory changes to key
419 developmental regulatory genes [8,33,41,63], this *trans* regulatory increase in cell type
420 abundance could be due to a small number of *cis* regulatory changes. These initially evolved
421 developmental regulatory changes could result in similar downstream changes in the
422 developmental landscape, resulting in the shared increase in dental cell types. Consistent
423 with this interpretation, stickleback orthologs of genes known to be expressed during
424 mammalian tooth development were found here to have a much greater incidence of
425 convergently evolved increase in *trans* regulatory gene expression.
426
427 **Compensatory *cis* and *trans***
428 Previous studies [17,18] have shown compensatory *cis* and *trans* changes are essential for
429 the evolution of gene expression. These findings are consistent with the idea that the main
430 driving force in the evolution of gene expression is stabilizing selection [59] where
431 compensatory changes to regulatory elements are selected for to maintain optimal gene
432 expression levels. In both benthic and creek dental tissue, when considering all genes with a

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

quantifiable (i.e. polymorphic and covered by ~20 reads, see Methods) *cis* effects, discordant compensatory *cis* and *trans* changes were far more common than concordant ones. This trend could be driven by some initial selection on pleiotropic *trans* changes, followed by selection for compensatory *cis* changes to restore optimal gene expression levels [17,18,22]. However, the *trans*, but not the *cis*, evolved changes in gene expression were highly shared among the two freshwater populations. Thus, collectively our data support a model where two independently derived populations have convergently evolved both similar genome-wide expression levels as well as ecologically relevant morphological changes through different genetic means.

Potential parallels between teeth and hair regeneration

Creek and benthic sticklebacks have an increased rate of new tooth formation in adults relative to their marine ancestors [30]. In constantly replacing polyphyodonts, it has been proposed that teeth are replaced through a dental stem cell intermediate [37,38]. A strong candidate gene underlying a large effect benthic tooth quantitative trait locus (QTL) is the secreted ligand *Bone Morphogenetic Protein 6* (*Bmp6*) [41] (Cleves et al 2018 under review, see Supplemental Data File 2), which is also a key regulator of stem cells in the mouse hair follicle [56]. Freshwater dental tissue displayed significantly increased expression of known signature genes of mouse hair follicle stem cells, perhaps reflecting more stem cell niches supporting the higher tooth numbers in freshwater fish. Genes upregulated in freshwater dental tissue also were significantly enriched for GO terms involved in the cell cycle and cell proliferation. Together these findings suggest that both freshwater populations have evolved an increased tooth replacement rate through an increased activity or abundance of their dental stem cells, and also suggest the genetic circuitry regulating mammalian hair and fish

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

457 tooth replacement might share an ancient, underlying core gene regulatory network.

458

459 ## Materials and Methods

460

461 ### Stickleback husbandry

462 Fish from all populations were raised in 110L aquaria in brackish water (3.5g/L Instant Ocean
463 salt, 0.217mL/L 10% sodium bicarbonate) at 18°C in 8 hours of light per day. Young fry
464 [standard length (SL) < 10 millimeters (mm)] were fed a diet of live *Artemia,* early juveniles
465 (SL ~10 - 20 mm) a combination of live *Artemia* and frozen *Daphnia*, and older juveniles (SL >
466 ~20 mm) and adults a combination of frozen bloodworms and *Mysis* shrimp. Experiments
467 were approved by the Institutional Animal Care and Use Committee of the University of
468 California-Berkeley (protocol # R330).

469

470 ### Skeletal staining and imaging

471 Sticklebacks were fixed in 10% neutral buffered formalin overnight at 4°C. Fish were washed
472 once with water and then stained in 1% KOH, 0.008% Alizarin Red for 24 hours. Following a
473 water rinse, fish were cleared in 0.25 % KOH, 50% glycerol for 2-3 weeks. Branchial
474 skeletons were dissected as previously described [64]. Pharyngeal teeth were quantified with
475 fluorescent illumination using a TX2 filter on a Leica DM2500 microscope. Representative
476 tooth plates were created using montage z-stacks on a Leica M165 FC using the RhodB filter.
477 Adult fish were imaged using a Canon Powershot S95. Some tooth count data from the
478 CERC, RABS, and PAXB populations; n = 11, 13, 29, respectively, (see Table S1) have been
479 previously published [30].

480

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

481 **DNA preparation and genome resequencing**

482 Caudal fin tissue was placed into 600µl tail digestion buffer [10mM Tris pH 8.0, 100mM NaCl,
483 10mM EDTA, 0.05% SDS, 2.5µl ProK (Ambion AM2546)] for 12 hours at 55°C. Following
484 addition of 600 µl of 1:1 phenol:chloroform solution and an aqueous extraction, DNA was
485 precipitated with the addition of 1ml 100% ethanol, centrifuged, washed with 75% ethanol,
486 and resuspended in water. 50ng of purified genomic DNA was used as input for the Nextera
487 Library prep kit (Illumina FC-121-1031), and barcoded libraries were constructed following the
488 manufacturer's instructions. Library quality was verified using an Agilent Bioanalyzer.
489 Libraries were pooled and sequenced on an Illumina HiSeq 2000 (see Table S2 for details).
490
491 **RNA purification and creation of RNA-seq libraries**
492 Late juvenile stage female sticklebacks (SL ~40mm) were euthanized in 0.04% Tricaine.
493 Dissected [64] bilateral ventral pharyngeal tooth plates were placed into 500µl TRI reagent,
494 then incubated at room temperature for 5 minutes. Following addition of 100µl of chloroform,
495 a further 10 minute incubation and centrifugation, the aqueous layer was extracted. Following
496 addition of 250µl isopropyl alcohol and 10 minute incubation, RNA was precipitated by
497 centrifugation, washed with 75% EtOH, and dissolved in 30ul of DEPC-treated water. RNA
498 integrity was assayed by an Agilent Bioanalyzer. 500ng of RNA from each fish was used as
499 input to the Illumina stranded TruSeq polyA RNA kit (Illumina RS-122-2001), and libraries
500 were constructed following the manufacturer's instructions. Library quality was analyzed on
501 an Agilent Bioanalyzer, and libraries were pooled and sequenced on an Illumina HiSeq2000
502 (see Table S3).
503
504 **Gene expression quantification and analysis**

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

RNA-seq reads were mapped to the stickleback reference genome [31] using the STAR aligner [65] (version 2.3, parameters = --alignIntronMax 100000 --alignMatesGapMax 200000 --outFilterMultimapNmax 20 --outFilterMismatchNmax 999 --outFilterMismatchNoverLmax 0.04 --outFilterType BySJout), using ENSEMBL genes release 85 as a reference transcriptome. The resulting SAM files were sorted and indexed using Samtools version 0.1.18 [66], PCR duplicates were removed, read groups added and mate pair information fixed using Picard tools (version 1.51) (http://broadinstitute.github.io/picard/) with default settings. Gene expression was quantified with the Cufflinks suite (v 2.2.1) [52–55] using ENSEMBL genes as a reference transcriptome, with gene expression quantified with cuffquant (-u --library-type fr-firststrand) and normalized with cuffnorm. Differentially expressed genes were found using cuffdiff2, with parameters (-u --FDR .1 --library-type fr-firststrand, using the reference genome for bias correction). Genes with a mean expression less than 0.1 FPKM were filtered from further analysis.

Gene set and gene ontology enrichment

The BiteCode gene set was generated by combining all genes in the BiteIt (http://bite-it.helsinki.fi/) or ToothCODE (http://compbio.med.harvard.edu/ToothCODE/) [36] databases. Stickleback orthologs or co-orthologs were found using the annotated names of ENSEMBL stickleback genes. Gene set expression change statistical enrichment was done as previously described [67]. Briefly, a t-test was performed for each gene to test for a difference in mean expression between the two treatments. The resulting t-values were subject to a 1-sample t-test, with the null model that the mean of the t-values was 0. Cutoffs were validated using 10,000 bootstrapped replicate gene sets drawn from the same gene expression matrix. Stickleback orthologs of mouse or human genes were determined using annotated ENSEMBL

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

orthologs. Sorted lists of genes, ranked by $\log_2$ expression change in benthic dental tissue relative to marine, creek relative to marine, or the mean of creek and benthic relative to marine, were generated using the measured gene expression data. Gene Ontology enrichment was done using Gorilla [68,69], and results were visualized using REVIGO [70].

Detection of genomic and transcriptomic variants

Genomic resequencing reads were aligned to the stickleback reference genome [31] using the bwa aln and bwa sampe modules of the Burrows-Wheeler Alignment tool (v 0.6.0-r85) [71]. Resulting SAM files were converted to BAM files, sorted and indexed by Samtools version 0.1.18 [66], with PCR duplicates removed by Picard tools. GATK's (v3.2-2) IndelRealigner (parameter: '-LOD 0.4'), BaseRecalibrator, and PrintReads were used on the resulting BAM files. BAM files from the above RNA-seq alignment were readied for genotype calling using GATK's SplitNCigarReads, BaseRecalibrator, and PrintReads. Finally, the UnifiedGenotyper was used to call variants from the RNA-seq and DNA-seq BAM files, with parameters (-stand_call_conf 30 -stand_emit_conf 30 -U ALLOW_N_CIGAR_READS --genotype_likelihoods_model BOTH) [43,45].

Following final variant calling and detection, pseudo-transcriptomes were created for each F1 hybrid. The pseudo-transcriptomes consist of the predicted sequence for each allele within an F1 hybrid, with all predicted splicing variants of a gene collapsed to a single transcript. A variant was added to the pseudo-transcriptome if and only if it was homozygous in the sequenced parents (or parent's sibling in the case of the Alaskan marine parent of the Cerrito creek x Alaskan marine F1 hybrids) and called heterozygous in the F1 hybrid.

***Cis* and *trans* regulatory divergence quantification**

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

RNA-seq reads from F1 hybrid sticklebacks were aligned to the individual's pseudo-transcriptome using STAR (v 2.3) with the parameters: --outFilterMultimapNmax 1 and --outFilterMultimapScoreRange 1. By only looking at uniquely aligning reads, we ensured we only considered reads which overlapped a heterozygous variant site. Counting these unique reads minimizes double counting a single read that supports two different variant positions. Total *cis* divergence in each F1 hybrid was quantified by comparing the number of reads mapping uniquely to each allele in the pseudo-transcriptome.

Following *cis* divergence quantification in all F1 hybrids, we considered the overall *cis* change in the different freshwater populations. Genes which only had 20 or fewer uniquely mapping reads across all replicates were filtered from further analysis. We excluded genes with more than a 32-fold change, as a manual inspection revealed these to be either genotyping errors or mitochondrial genes. Reported *cis* ratios were calculated by comparing the ratio of uniquely mapped freshwater reads to uniquely mapped marine reads. Evolved *trans* changes were quantified as the difference between the log of the overall gene expression change between the freshwater and marine parents and the log of measured *cis* freshwater expression change. Percent *cis* change was calculated as the absolute value of the log of the *cis* change divided by the sum of the absolute value of the log of the *cis* change and the absolute value of the log of the *trans* change. Statistical significance of *cis* changes was determined by a binomial test comparing overall reads mapping to the freshwater allele to a null model of no *cis* divergence, with a false discovery rate of 1% applied using the Benjamini-Hochberg method. Statistical significance of *trans* changes was determined by a G-test, comparing the expected (based on the measured *cis* change) and observed ratios of marine and freshwater, with a 1% false discovery rate.

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Data Availability

All sequencing reads are available on the Sequence Read Archive (XXXXXX). All scripts used for analysis are available on GitHub (xxxxx).

Acknowledgements

We thank Nikunj Donde for collecting a subset of tooth count data.

Funding statement

This work was supported by NIH R01 DE021475 (to C.T.M.), NIH genomics training grant 5T32HG000047-15 (to J.C.H.), and NSF Graduate Research Fellowship (to N.A.E.). This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References

1. Banerji J, Rusconi S, Schaffner W. Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell. 1981;27: 299–308.

2. Ong C-T, Corces VG. Enhancer function: new insights into the regulation of tissue-specific gene expression. Nat Rev Genet. 2011;12: 283–293. doi:10.1038/nrg2957

3. Stern DL. Evolutionary developmental biology and the problem of variation. Evolution. 2000;54: 1079–1091.

4. Britten RJ, Davidson EH. Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. The Quarterly Review of Biology. 1971;46: 111–138.

5. King MC, Wilson AC. Evolution at two levels in humans and chimpanzees. Science. 1975;188: 107–116.

6. Gompel N, Prud'homme B, Wittkopp PJ, Kassner VA, Carroll SB. Chance caught on the wing: *cis*-regulatory evolution and the origin of pigment patterns in *Drosophila*. Nature. 2005;433: 481–487. doi:10.1038/nature03235

7. Manceau M, Domingues VS, Mallarino R, Hoekstra HE. The developmental role of *Agouti* in color pattern evolution. Science. 2011;331: 1062–1065. doi:10.1126/science.1200684

8. O'Brown NM, Summers BR, Jones FC, Brady SD, Kingsley DM. A recurrent regulatory change underlying altered expression and WNT response of the stickleback armor plates gene *Eda*. eLife. 2015;4: e05290. doi:10.7554/eLife.05290

9. Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G, Dickson M, Grimwood J, et al. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. Science. 2005;307: 1928–1933. doi:10.1126/science.1107239

10. Indjeian VB, Kingman GA, Jones FC, Guenther CA, Grimwood J, Schmutz J, et al. Evolving new skeletal traits by *cis*-regulatory changes in bone morphogenetic proteins. Cell. 2016;164: 45–56. doi:10.1016/j.cell.2015.12.007

11. Schor IE, Degner JF, Harnett D, Cannavò E, Casale FP, Shim H, et al. Promoter shape varies across populations and affects promoter evolution and expression noise. Nature Genetics. 2017;49: 550. doi:10.1038/ng.3791

12. McCoy RC, Wakefield J, Akey JM. Impacts of Neanderthal-introgressed sequences on the landscape of human gene expression. Cell. 2017;168: 916–927.e12. doi:10.1016/j.cell.2017.01.038

13. Ritter DI, Dong Z, Guo S, Chuang JH. Transcriptional enhancers in protein-coding exons of vertebrate developmental genes. PLOS ONE. 2012;7: e35202. doi:10.1371/journal.pone.0035202

14. Gibbons TC, Metzger DCH, Healy TM, Schulte PM. Gene expression plasticity in response to salinity acclimation in threespine stickleback ecotypes from different salinity habitats. Mol Ecol. 2017;26: 2711–2725. doi:10.1111/mec.14065

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

15. Yao C, Joehanes R, Johnson AD, Huan T, Liu C, Freedman JE, et al. Dynamic role of *trans* regulation of gene expression in relation to complex traits. The American Journal of Human Genetics. 2017;100: 571–580. doi:10.1016/j.ajhg.2017.02.003

16. Wittkopp PJ, Haerum BK, Clark AG. Evolutionary changes in *cis* and *trans* gene regulation. Nature. 2004;430: 85–88. doi:10.1038/nature02698

17. Mack KL, Campbell P, Nachman MW. Gene regulation and speciation in house mice. Genome Res. 2016; gr.195743.115. doi:10.1101/gr.195743.115

18. Goncalves A, Leigh-Brown S, Thybert D, Stefflova K, Turro E, Flicek P, et al. Extensive compensatory *cis-trans* regulation in the evolution of mouse gene expression. Genome Res. 2012;22: 2376–2384. doi:10.1101/gr.142281.112

19. McManus CJ, Coolon JD, Duff MO, Eipper-Mains J, Graveley BR, Wittkopp PJ. Regulatory divergence in *Drosophila* revealed by mRNA-seq. Genome Res. 2010;20: 816–825. doi:10.1101/gr.102491.109

20. Lemmon ZH, Bukowski R, Sun Q, Doebley JF. The role of *cis* regulatory evolution in maize domestication. PLOS Genet. 2014;10: e1004745. doi:10.1371/journal.pgen.1004745

21. Osada N, Miyagi R, Takahashi A. *cis*- and *trans*-regulatory effects on gene expression in a natural population of *Drosophila* melanogaster. Genetics. 2017; 117.201459. doi:10.1534/genetics.117.201459

22. Coolon JD, McManus CJ, Stevenson KR, Graveley BR, Wittkopp PJ. Tempo and mode of regulatory evolution in *Drosophila*. Genome Res. 2014;24: 797–808. doi:10.1101/gr.163014.113

23. Peichel CL, Marques DA. The genetic and molecular architecture of phenotypic diversity in sticklebacks. Phil Trans R Soc B. 2017;372: 20150486. doi:10.1098/rstb.2015.0486

24. Pritchard VL, Viitaniemi HM, McCairns RJS, Merilä J, Nikinmaa M, Primmer CR, et al. Regulatory architecture of gene expression variation in the threespine stickleback *Gasterosteus aculeatus*. G3. 2017;7: 165–178. doi:10.1534/g3.116.033241

25. Ishikawa A, Kusakabe M, Yoshida K, Ravinet M, Makino T, Toyoda A, et al. Different contributions of local- and distant-regulatory changes to transcriptome divergence between stickleback ecotypes. Evolution. 2017;71: 565–581. doi:10.1111/evo.13175

26. Di Poi C, Bélanger D, Amyot M, Rogers S, Aubin-Horth N. Receptors rather than signals change in expression in four physiological regulatory networks during evolutionary divergence in threespine stickleback. Mol Ecol. 2016;25: 3416–3427. doi:10.1111/mec.13690

27. Huang Y, Chain FJJ, Panchal M, Eizaguirre C, Kalbe M, Lenz TL, et al. Transcriptome profiling of immune tissues reveals habitat-specific gene expression between lake and river sticklebacks. Mol Ecol. 2016;25: 943–958. doi:10.1111/mec.13520

28. Bell MA, Foster SA. The evolutionary biology of the threespine stickleback. Oxford University Press; 1994.

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

29. Hagen DW, Gilbertson LG. Geographic variation and environmental selection in *Gasterosteus aculeatus* L. in the Pacific Northwest, America. Evolution. 1972;26: 32–51. doi:10.2307/2406981

30. Ellis NA, Glazer AM, Donde NN, Cleves PA, Agoglia RM, Miller CT. Distinct developmental genetic mechanisms underlie convergently evolved tooth gain in sticklebacks. Development. 2015;142: 2442–2451. doi:10.1242/dev.124248

31. Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, Johnson J, et al. The genomic basis of adaptive evolution in threespine sticklebacks. Nature. 2012;484: 55–61. doi:10.1038/nature10944

32. Erickson PA, Glazer AM, Killingbeck EE, Agoglia RM, Baek J, Carsanaro SM, et al. Partially repeatable genetic basis of benthic adaptation in threespine sticklebacks. Evolution. 2016;70: 887–902. doi:10.1111/evo.12897

33. Chan YF, Marks ME, Jones FC, Villarreal G, Shapiro MD, Brady SD, et al. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. Science. 2010;327: 302–305. doi:10.1126/science.1182213

34. Shapiro MD, Marks ME, Peichel CL, Blackman BK, Nereng KS, Jónsson B, et al. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature. 2004;428: 717. doi:10.1038/nature02415

35. Pispa J, Thesleff I. Mechanisms of ectodermal organogenesis. Dev Biol. 2003;262: 195–205.

36. O'Connell DJ, Ho JWK, Mammoto T, Turbe-Doan A, O'Connell JT, Haseley PS, et al. A WNT-BMP feedback circuit controls intertissue signaling dynamics in tooth organogenesis. Sci Signal. 2012;5: ra4. doi:10.1126/scisignal.2002414

37. Jernvall J, Thesleff I. Tooth shape formation and tooth renewal: evolving with the same signals. Development. 2012;139: 3487–3497. doi:10.1242/dev.085084

38. Tucker AS, Fraser GJ. Evolution and developmental diversity of tooth regeneration. Seminars in Cell & Developmental Biology. doi:10.1016/j.semcdb.2013.12.013

39. Bloomquist RF, Parnell NF, Phillips KA, Fowler TE, Yu TY, Sharpe PT, et al. Coevolutionary patterning of teeth and taste buds. PNAS. 2015;112: E5954–E5962. doi:10.1073/pnas.1514298112

40. Martin KJ, Rasch LJ, Cooper RL, Metscher BD, Johanson Z, Fraser GJ. *Sox2*+ progenitors in sharks link taste development with the evolution of regenerative teeth from denticles. Proc Natl Acad Sci. 2016;113: 14769–14774. doi:10.1073/pnas.1612354113

41. Cleves PA, Ellis NA, Jimenez MT, Nunez SM, Schluter D, Kingsley DM, et al. Evolved tooth gain in sticklebacks is associated with a *cis*-regulatory allele of Bmp6. PNAS. 2014;111: 13912–13917. doi:10.1073/pnas.1407567111

42. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Meth. 2012;9: 357–359. doi:10.1038/nmeth.1923

43. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20: 1297–1303. doi:10.1101/gr.107524.110

44. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43: 491–498. doi:10.1038/ng.806

45. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics. 2013;43: 11.10.1-33. doi:10.1002/0471250953.bi1110s43

46. FactoMineR: an R package for multivariate analysis | Lê | Journal of Statistical Software [Internet]. [cited 22 Apr 2016]. Available: https://www.jstatsoft.org/article/view/v025i01

47. Gene expression in tooth (WWW database), http://bite-it.helsinki.fi. Developmental Biology Programme of the University of Helsink; 1996.

48. Fraser GJ, Hulsey CD, Bloomquist RF, Uyesugi K, Manley NR, Streelman JT. An ancient gene network is co-opted for teeth on old and new jaws. PLoS Biol. 2009;7: e1000031. doi:10.1371/journal.pbio.1000031

49. Fraser GJ, Bloomquist RF, Streelman JT. Common developmental pathways link tooth shape to regeneration. Dev Biol. 2013;377: 399–414. doi:10.1016/j.ydbio.2013.02.007

50. Stock DW. Zebrafish dentition in comparative context. J Exp Zool B Mol Dev Evol. 2007;308: 523–549. doi:10.1002/jez.b.21187

51. Jackman WR, Draper BW, Stock DW. FGF signaling is required for zebrafish tooth development. Dev Biol. 2004;274: 139–157. doi:10.1016/j.ydbio.2004.07.003

52. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotech. 2013;31: 46–53. doi:10.1038/nbt.2450

53. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotech. 2010;28: 511–515. doi:10.1038/nbt.1621

54. Roberts A, Pimentel H, Trapnell C, Pachter L. Identification of novel transcripts in annotated genomes using RNA-Seq. Bioinformatics. 2011; btr355. doi:10.1093/bioinformatics/btr355

55. Roberts A, Trapnell C, Donaghey J, Rinn JL, Pachter L. Improving RNA-Seq expression estimates by correcting for fragment bias. Genome Biology. 2011;12: R22. doi:10.1186/gb-2011-12-3-r22

56. Kandyba E, Leung Y, Chen Y-B, Widelitz R, Chuong C-M, Kobielak K. Competitive balance of intrabulge BMP/Wnt signaling reveals a robust gene network ruling stem cell homeostasis and cyclic activation. Proc Natl Acad Sci USA. 2013;110: 1351–1356. doi:10.1073/pnas.1121312110

57. Yasuoka A, Aihara Y, Matsumoto I, Abe K. Phospholipase C-beta 2 as a mammalian taste

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

735 signaling marker is expressed in the multiple gustatory tissues of medaka fish, Oryzias latipes.
736 Mechanisms of Development. 2004;121: 985–989. doi:10.1016/j.mod.2004.03.009

737 58. Bachmanov AA, Beauchamp GK. Taste Receptor Genes. Annu Rev Nutr. 2007;27: 389–414.
738 doi:10.1146/annurev.nutr.26.061505.111329

739 59. Gilad Y, Oshlack A, Rifkin SA. Natural selection on gene expression. Trends in Genetics.
740 2006;22: 456–461. doi:10.1016/j.tig.2006.06.002

741 60. Shen SQ, Turro E, Corbo JC. Hybrid mice reveal parent-of-origin and *cis*- and *trans*-regulatory
742 effects in the retina. PLOS ONE. 2014;9: e109382. doi:10.1371/journal.pone.0109382

743 61. Wittkopp PJ, Haerum BK, Clark AG. Regulatory changes underlying expression differences
744 within and between *Drosophila* species. Nat Genet. 2008;40: 346–350. doi:10.1038/ng.77

745 62. Brawand D, Soumillon M, Necsulea A, Julien P, Csárdi G, Harrigan P, et al. The evolution of
746 gene expression levels in mammalian organs. Nature. 2011;478: 343–348.
747 doi:10.1038/nature10532

748 63. Chen H, Capellini TD, Schoor M, Mortlock DP, Reddi AH, Kingsley DM. Heads, shoulders,
749 elbows, knees, and toes: modular *Gdf5* enhancers control different joints in the vertebrate
750 skeleton. PLOS Genetics. 2016;12: e1006454. doi:10.1371/journal.pgen.1006454

751 64. Ellis NA, Miller CT. Dissection and flat-mounting of the threespine stickleback branchial
752 skeleton. Journal of Visualized Experiments. 2016; doi:10.3791/54056

753 65. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal
754 RNA-seq aligner. Bioinformatics. 2012; bts635. doi:10.1093/bioinformatics/bts635

755 66. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map
756 format and SAMtools. Bioinformatics. 2009;25: 2078–2079. doi:10.1093/bioinformatics/btp352

757 67. Irizarry RA, Wang C, Zhou Y, Speed TP. Gene set enrichment analysis made simple. Stat
758 Methods Med Res. 2009;18: 565–575. doi:10.1177/0962280209351908

759 68. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and
760 visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009;10: 48.
761 doi:10.1186/1471-2105-10-48

762 69. Eden E, Lipson D, Yogev S, Yakhini Z. Discovering motifs in ranked lists of DNA sequences.
763 PLOS Comput Biol. 2007;3: e39. doi:10.1371/journal.pcbi.0030039

764 70. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO summarizes and visualizes long lists of gene
765 ontology terms. PLOS ONE. 2011;6: e21800. doi:10.1371/journal.pone.0021800

766 71. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform.
767 Bioinformatics. 2009;25: 1754–1760. doi:10.1093/bioinformatics/btp324

768
769

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fig 1. Evolved tooth gain in two freshwater populations. (A) Stickleback population locations. (B) Representative Alizarin red stained adult lab-reared sticklebacks (top, scale bars = 1 cm) and dissected ventral pharyngeal tooth plates (scale bars = 100μm). (C) Total ventral pharyngeal tooth number of 35-50 millimeter standard length lab-reared adult fish from each population.

Fig 2. Convergent evolution of gene expression in dental tissue. (A) Ventral pharyngeal tooth plates from three different populations were dissected and gene expression quantified by RNA-seq. (B) Principal component analysis of dental tissue gene expression shows population specific expression profiles. (C) Freshwater dental tissue exhibited correlated gene expression changes for all genes (blue), with increased correlation observed for orthologs of genes known to be expressed during mammalian tooth development (red). (D) Expression of genes annotated as expressed in zebrafish teeth (zfin.org) which were significantly upregulated in one or both freshwater populations.

Fig 3. Evolved changes in *cis*-regulation (A) Ventral pharyngeal tooth plates from marine-creek and marine-benthic lake F1 hybrids were dissected and *cis* regulatory changes assayed using phased RNA-seq reads. (B) Density plot showing the measured *cis*-regulatory changes. Neither population displayed a significant allelic bias, as measured by a Wilcoxon signed-rank test. (C-D) Gene expression changes in both parental and hybrid dental tissue – genes are color-coded based on the role of *cis* and/or *trans* change in benthic (C) or creek (D) dental tissue. Dashed line indicates the first principal component axis.

Fig 4. *Trans* changes predominate evolved dental gene expression changes. (A-B)

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

794 Proportion of differentially expressed genes displaying opposing and concordant *cis* and *trans*
795 changes in benthic (A) or creek (B) dental tissue. Genes whose expression differences were
796 mostly explained by *cis* changes tended to be more concordant (*P* =5.0e-17, 0.002 for benthic
797 and creek, respectively) than those mostly explained by *trans* changes. (C) Density of the
798 relative percentage of gene expression differences which are explained by *cis* changes in
799 benthic and creek dental tissue. (D) Cumulative percentage of percentage of gene expression
800 due to *cis* changes. Genes in creek samples display a higher percentage *cis* change than
801 genes in benthic samples (*P* = 1.25e-22, Mann-Whitney U test).
802
803 **Fig 5. *Trans* changes are more likely to be shared across populations.** (A) Genes with
804 significantly different evolved expression in both freshwater populations relative to marine
805 fish, showing significantly correlated changes in gene expression in benthic and creek dental
806 tissue. (B) Freshwater dental tissue had a significant but small number of shared *cis*-
807 regulatory changes. (C) Freshwater dental tissue showed significantly correlated changes in
808 *trans* expression changes. A-C show genes with significant expression changes between
809 populations and quantifiable (i.e. genes with transcripts containing a polymorphic SNP
810 covered by at least 20 reads) *cis*-regulatory changes in both populations. Density (color) was
811 estimated with a Gaussian kernal density estimator. BiteCode genes (see Methods) are
812 indicated with black stars. D-F Bar graphs show the number of genes with shared or divergent
813 expression patterns from the above panels. G-I are similar to A-C, but show only genes in the
814 BiteCode gene set.
815
816 **Fig S1. Independent freshwater evolutionary history.** (A) Genome-wide maximum-
817 likelihood phylogeny created from genomic resequencing data. Wild-caught fish are non-

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

818 italicized. All nodes have 100% bootstrapping support. (B) Principal component analysis of
819 genome-wide genotypes separates marine and creek populations from the benthic lake
820 population, with the 2nd PC separating marine and freshwater populations.
821
822 **Fig S2. Freshwater upregulation of putative dental genes.** (A) Benthic upregulation of
823 BiteCode genes (*P* = 9.8e-3, GSEA). (B) Creek upregulation of BiteCode genes (*P* = 2.1e-5,
824 GSEA). (C) Benthic and creek upregulation of BiteCode genes (*P* = 5.1e-6, GSEA).
825
826 **Fig S3. Concerted changes in stem cell markers and signaling pathways.** (A-F) Changes
827 in gene expression changes of genes annotated as components of the indicated signaling
828 pathways (BMP, FGF, SHH, WNT, ACT, TGFB, NOTCH, or EDA) [36] or orthologs of a
829 described set of mouse hair follicle stem cell signature genes (HFSC) [56]. Violin plots show
830 the mean expression change of genes in the pathway. (A) Change in freshwater (benthic +
831 creek) relative to marine. (B) Benthic specific changes (benthic relative to creek + marine). (C)
832 Creek specific changes (creek relative to benthic + marine). (D) Benthic evolved changes
833 (benthic relative to marine) (E) Creek evolved changes (Creek relative to marine) (F) Benthic
834 vs creek changes (benthic relative to creek).
835
836 **Fig S4. Gene ontology of freshwater upregulated genes.** (A-C) GO enrichment of genes
837 upregulated in benthic (A), creek (B), or both (C). GO analysis was preformed using Gorilla
838 [68], with the results visualized with Revigo [70].
839
840 **Fig S5. Expression of taste bud marker genes.** Expression levels of known taste bud
841 marker genes in marine, benthic and creek tooth plates as assayed by RNA-seq. * indicates

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

842 differentially expressed genes. Error bars are standard error of the mean.

843

844 **Fig S6. Compensatory changes dominate genes with no significant evolved gene**
845 **expression difference.** (A-B) Proportion of genes with quantifiable (i.e. genes with
846 transcripts containing a polymorphic SNP covered by at least 20 reads) hybrid expression
847 displaying opposing and concordant *cis* and *trans* changes in benthic (A) or creek (B) dental
848 tissue. Similar to Fig. 5, but here showing all genes, not just genes with significantly different
849 expression levels compared to marine. *Trans* regulatory changes predominate, as do
850 opposing over concordant changes. (C) Density plot of the percentage of gene expression
851 changes explained by *cis*-regulatory changes.

852

853 **Fig S7. *Trans* changes are more likely to be shared across populations.** (A) Expression
854 changes of genes with quantifiable (i.e. genes with transcripts containing a polymorphic SNP
855 covered by at least 20 reads) hybrid expression in both freshwater populations relative to
856 marine fish, showing significantly correlated changes in gene expression in benthic and creek
857 tooth plates. (B) *cis* regulatory changes of genes with quantifiable hybrid expression
858 expression in freshwater dental tissue overall do not display correlated evolved changes. (C)
859 *trans* regulatory changes of genes with quantifiable hybrid expression in freshwater dental
860 tissue. Density (color) was estimated with a Gaussian kernal density estimator. (D-F) Bar
861 graphs show the number of genes with shared or divergent expression patterns from A-C. G-I
862 are similar to A-C, but show only genes in the BiteCode gene set, revealing that these
863 orthologs have evolved highly convergent changes in the two freshwater populations (G),
864 despite non-convergent *cis* regulatory changes (H).

865

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

866 **Table S1. Population ventral pharyngeal tooth counts**

867 For each fish, the population, ecotype (freshwater or marine), total ventral pharyngeal tooth
868 number (TVTP), total length (TL), standard length (SL), and whether data has been published
869 [30] is shown.

870

871 **Table S2. Genomic DNA sequencing reads**

872 For each fish, population and biological replicate number (Fish), the total number of barcoded
873 reads from each fish (reads), and number of reads that mapped and passed all filters (final
874 mapped) is listed.

875

876 **Table S3. RNA-seq reads**

877 For each fish, population of parents and biological replicate number (sample), standard length
878 (SL), total reads (generated by HiSeq2000 over two different runs (run1 and run2)), mapped
879 reads (reads that mapped to the genome), and final reads (excludes reads filtered due to low
880 quality or PCR duplication) is listed.

881

882 **Table S4. Overall gene expression in tooth plate**

883 Estimated abundance in in fragments per kilobases per million reads (FPKM) of ENSEMBL
884 genes (rows) in ventral pharyngeal dental tissue from three individual fish from three
885 populations (in columns). Mean expression (in FPKM) is shown after the 3 replicates.
886 Log_2(Pop1/Pop2) shows the fold-change in log_2 of the estimated mean expression between
887 the two populations. IsSig(Pop1/Pop2) indicates whether the difference was significant as
888 reported by cuffdiff2.

889

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Table S5. BiteCode genes in sticklebacks

A list of stickleback orthologs in the BiteIt [47] (http://bite-it.helsinki.fi/) or ToothCODE (http://compbio.med.harvard.edu/ToothCODE/) [36] databases.

Table S6. GO process upregulated in freshwater

Gene Ontology (GO) term category and name are given in GO term and description, with the p-value, q-value, and relative enrichment within genes upregulated in freshwater dental tissue reported by GOrilla [68].

Table S7. GO process upregulated in benthic

Gene Ontology (GO) term category and name are given in GO term and description, with the p-value, q-value, and relative enrichment within genes upregulated in benthic dental tissue reported by GOrilla [68].

Table S8. GO process upregulated in creek

Gene Ontology (GO) term category and name are given in GO term and description, with the p-value, q-value, and relative enrichment within genes upregulated in creek dental tissue reported by GOrilla [68].

Table S9. F1 hybrid RNA-seq reads

For each ventral pharyngeal tooth plate (VTP), population of parents and biological replicate number (sample), standard length (SL), total reads (generated by HiSeq2000), mapped reads (reads that mapped to the genome), final reads (excludes reads filtered due to low quality or PCR duplication), and unique reads (reads that mapped uniquely to one haplotype) is listed.

bioRxiv preprint doi: https://doi.org/10.1101/259499; this version posted February 3, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

914

915 **Table S10. Benthic vs marine *cis* divergence**

916 Estimated gene expression change in *cis* in $\log_2$, benthic vs marine. Name is the reported

917 ENSEMBL gene name. Log_2(F/M) is the $\log_2$ of the ratio of freshwater vs marine reads

918 mapping uniquely to the gene.

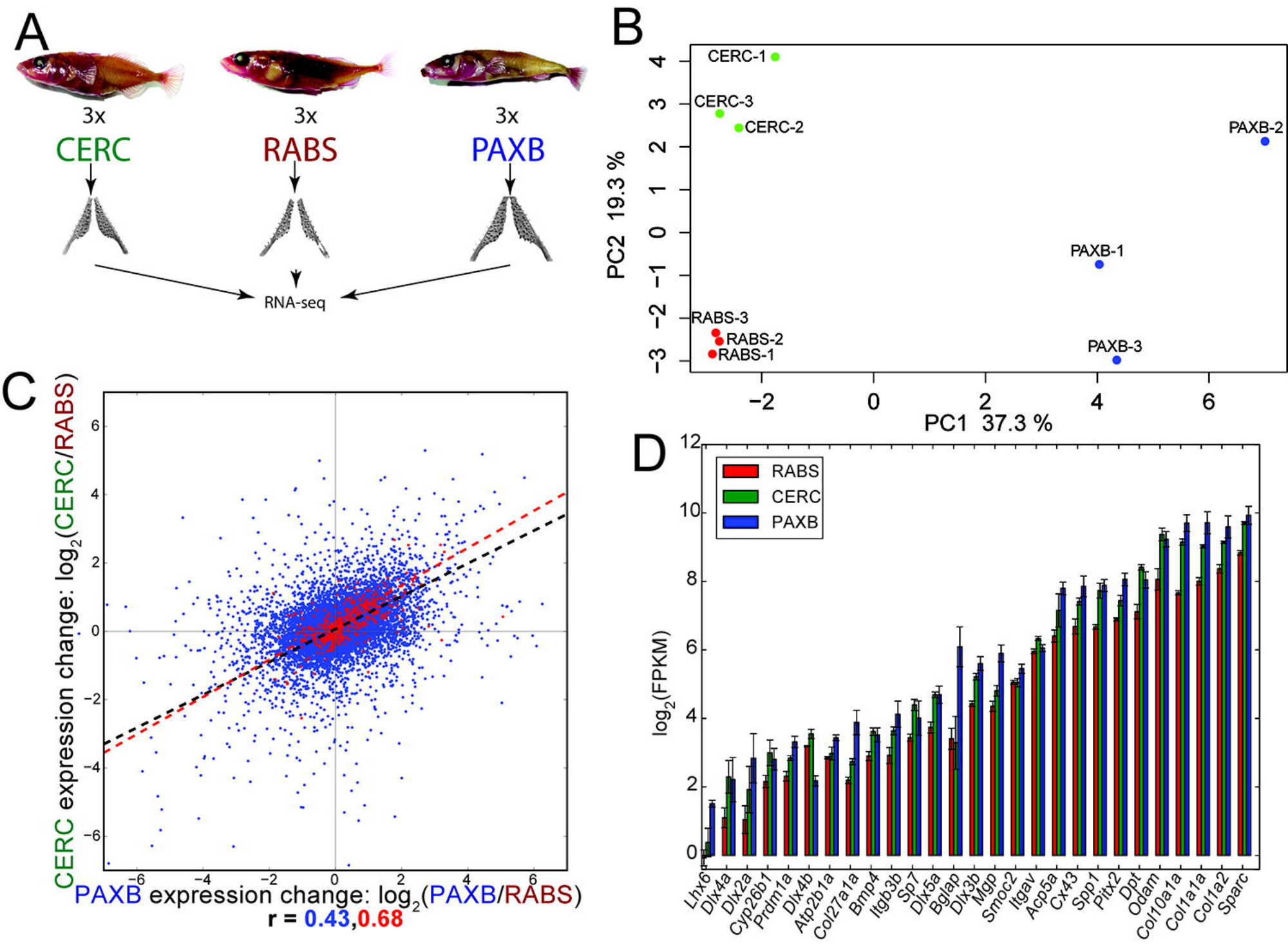
919

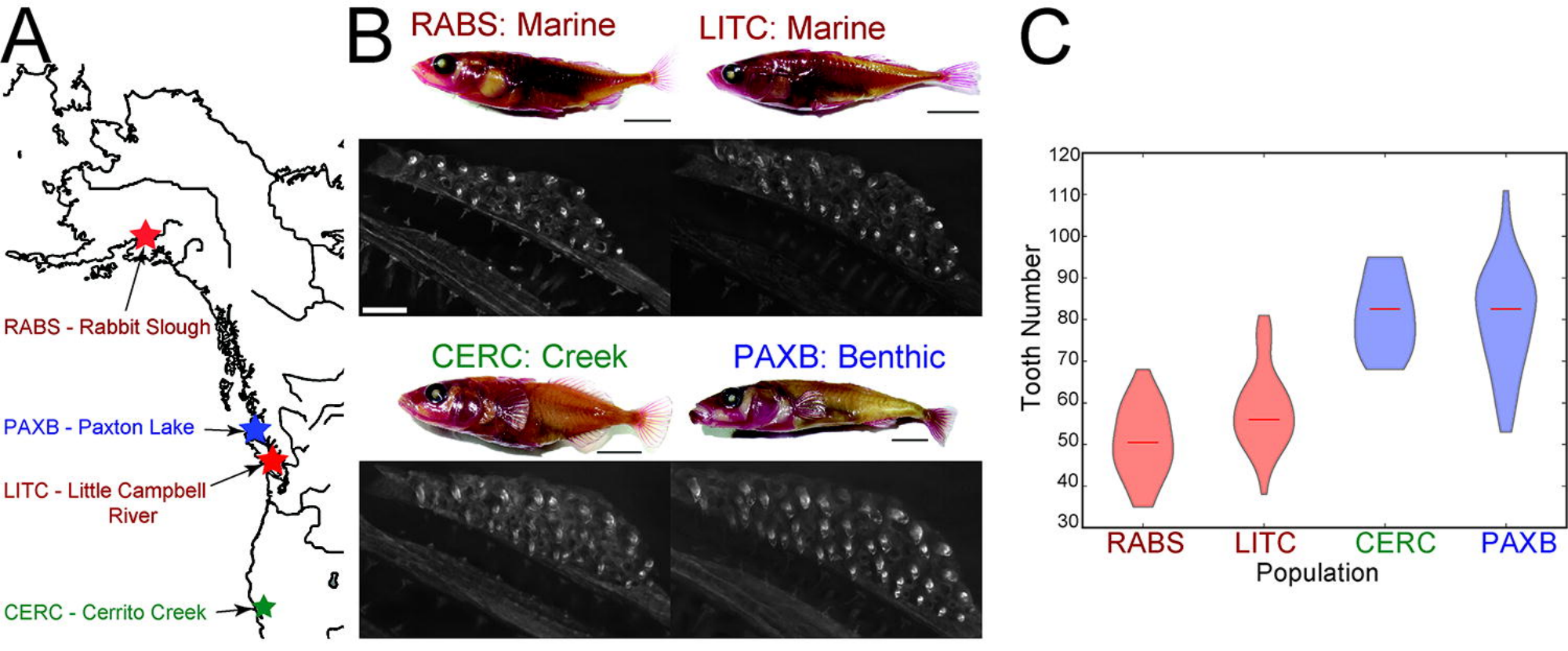
920 **Table S11. Creek vs marine *cis* divergence**

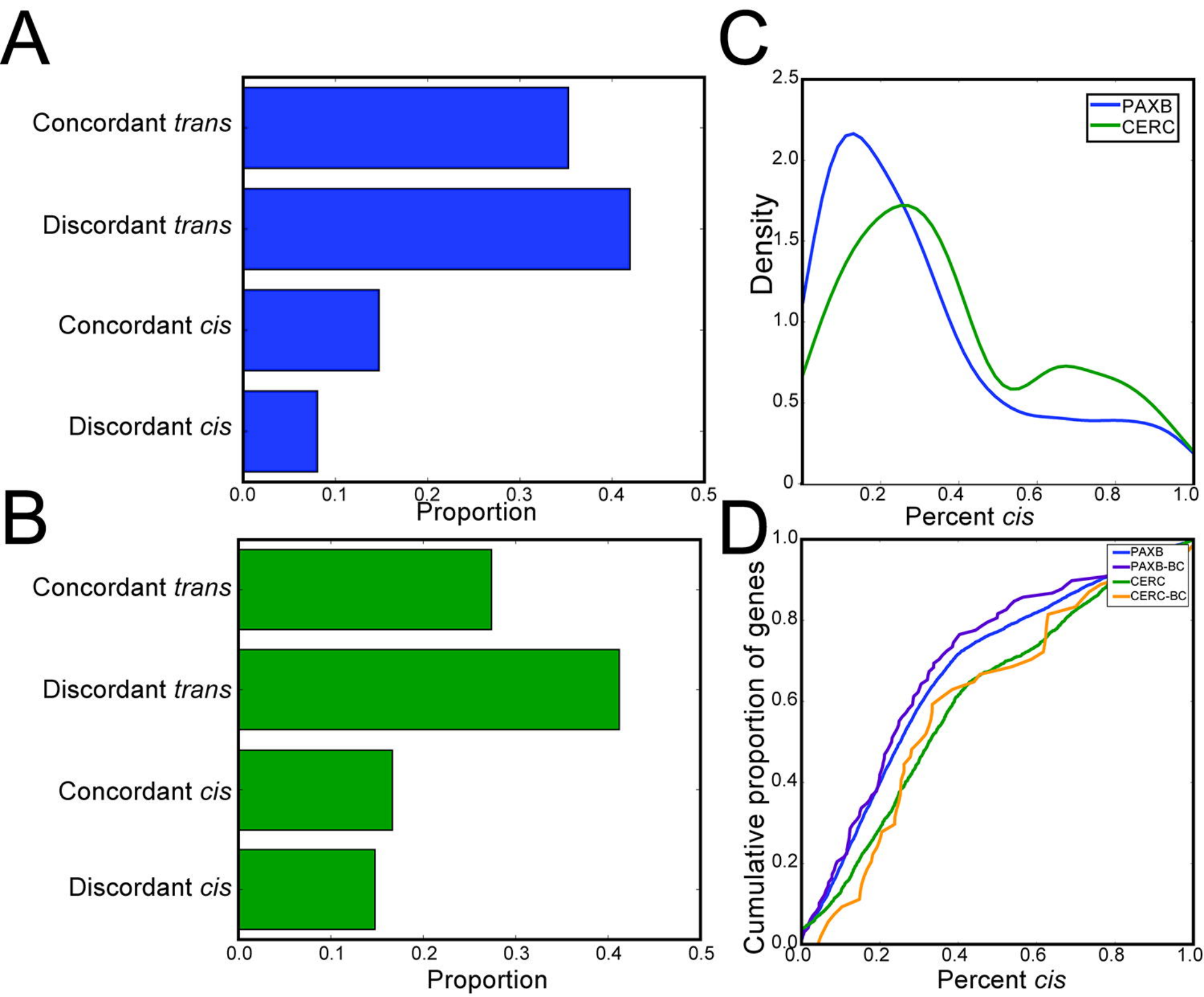
921 Estimated gene expression change in *cis* in $\log_2$, creek vs marine. Name is the reported

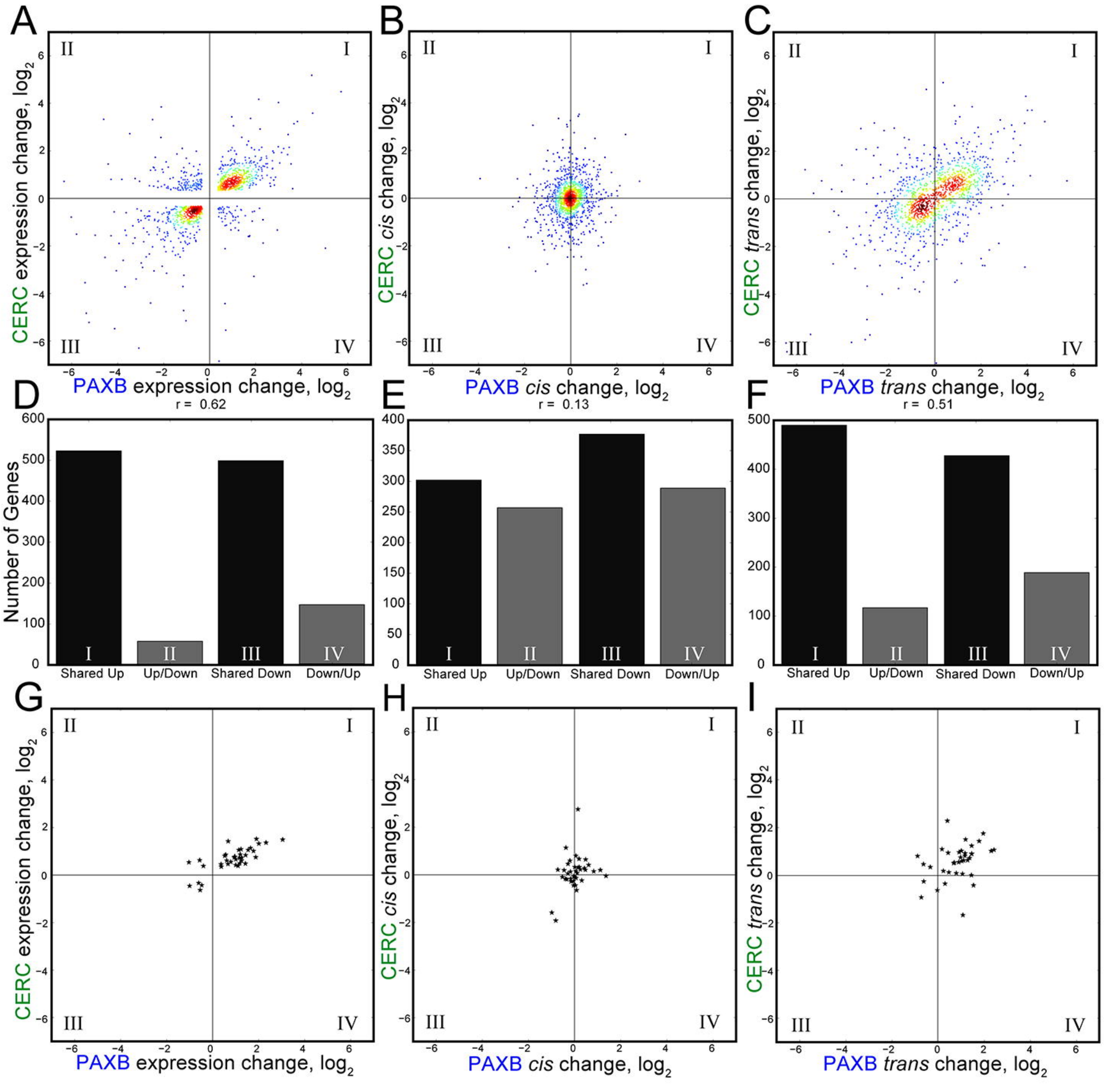
922 ENSEMBL gene name. Log_2(F/M) is the $\log_2$ of the ratio of freshwater vs marine reads

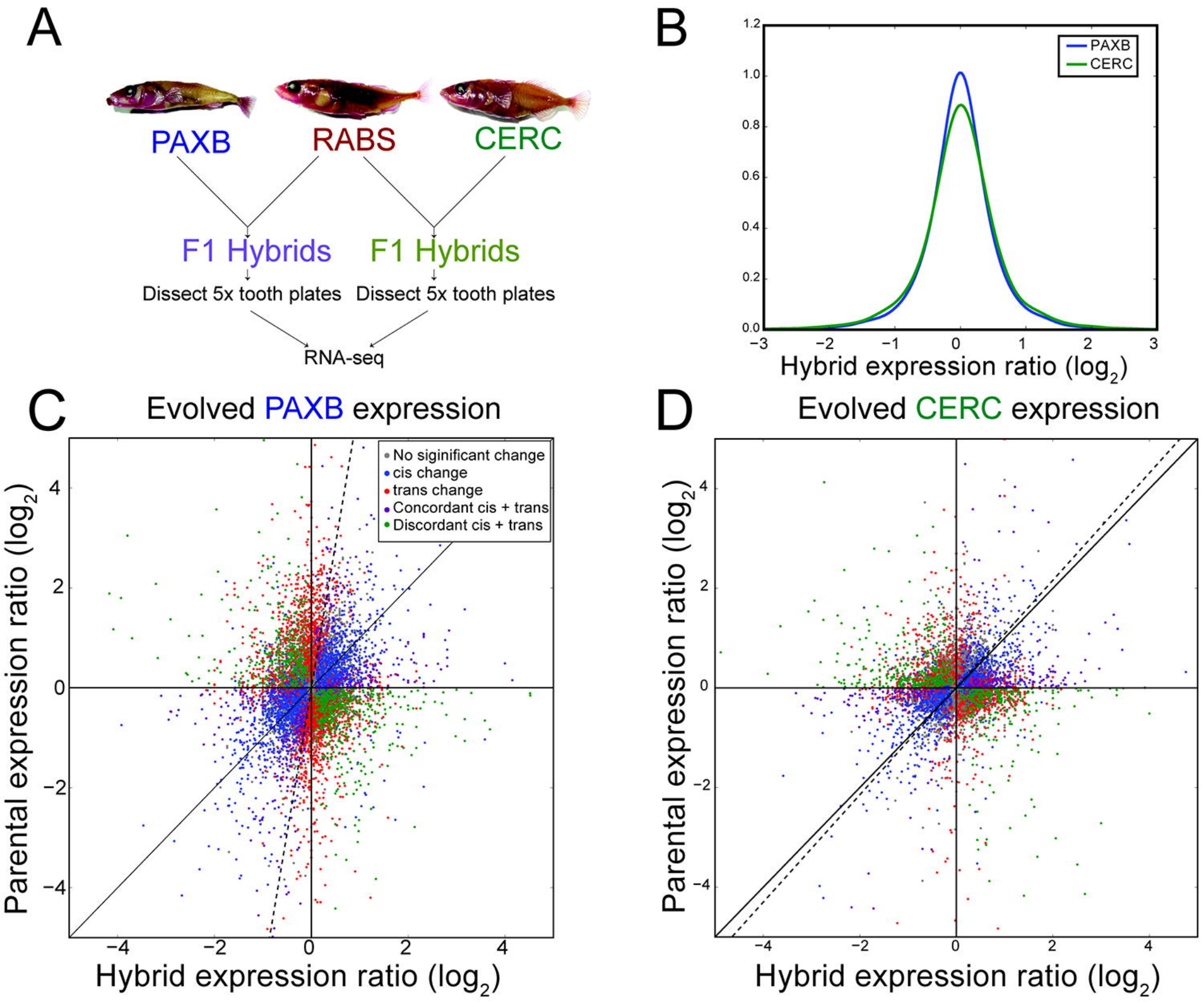
923 mapping uniquely to the gene.


A
3x
CERC
3x
RABS
3x
PAXB
RNA-seq
B
CERC-1
CERC-3
CERC-2
PAXB-2
PAXB-1
RABS-3
RABS-2
RABS-1
PAXB-3
PC2 19.3 %
PC1 37.3 %
C
CERC expression change: log2(CERC/RABS)
PAXB expression change: log2(PAXB/RABS)
r = 0.43,0.68
D
RABS
CERC
PAXB
log2(FPKM)
Lhx6
Dlx4a
Dlx2a
Cyp26b1
Prdm1a
Dlx4b
Atp2b1a
Col27a1a
Bmp4
Itgb3b
Sp7
Dlx5a
Bglap
Dlx3b
Mgp
Smoc2
Itgav
Acp5a
Cx43
Spp1
Pitx2
Dpt
Odam
Col10a1a
Col1a1a
Col1a2
Sparc


A
RABS - Rabbit Slough
PAXB - Paxton Lake
LITC - Little Campbell River
CERC - Cerrito Creek
B
RABS: Marine
LITC: Marine
CERC: Creek
PAXB: Benthic
C
Tooth Number
120
110
100
90
80
70
60
50
40
30
RABS
LITC
CERC
PAXB
Population


A
Concordant trans
Discordant trans
Concordant cis
Discordant cis
0.0
0.1
0.2
0.3
0.4
0.5
Proportion
B
Concordant trans
Discordant trans
Concordant cis
Discordant cis
0.0
0.1
0.2
0.3
0.4
0.5
Proportion
C
PAXB
CERC
Density
0
0.5
1.0
1.5
2.0
2.5
0.2
0.4
0.6
0.8
1.0
Percent cis
D
PAXB
PAXB-BC
CERC
CERC-BC
Cumulative proportion of genes
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
Percent cis


A
B
C
D
E
F
G
H
I
II
I
III
IV
CERC expression change, log2
CERC cis change, log2
CERC trans change, log2
PAXB expression change, log2
PAXB cis change, log2
PAXB trans change, log2
r = 0.62
r = 0.13
r = 0.51
Number of Genes
Shared Up
Up/Down
Shared Down
Down/Up


A
PAXB
RABS
CERC
F1 Hybrids
F1 Hybrids
Dissect 5x tooth plates
Dissect 5x tooth plates
RNA-seq
B
PAXB
CERC
Hybrid expression ratio (log2)
C
Evolved PAXB expression
No siginificant change
cis change
trans change
Concordant cis + trans
Discordant cis + trans
Parental expression ratio (log2)
Hybrid expression ratio (log2)
D
Evolved CERC expression
Parental expression ratio (log2)
Hybrid expression ratio (log2)