Unique genetic variants underlie parallel gene expression within a young adaptive radiation despite specialization on highly divergent resources

Joseph A. McGirr[1], Christopher H. Martin[1,]

[1]Department of Biology, University of North Carolina at Chapel Hill, 1250 South Rd., Chapel Hill, NC 27514-3280

Running Title: Parallel expression in divergent species

Keywords: RNA-seq, adaptive radiation, trophic specialization, selective sweep, transcriptome,

Correspondence: Joseph A. McGirr. Department of Biology, University of North Carolina at Chapel Hill, Genome Science 2244, 250 Bell Tower Drive., NC, 27514, USA

Email: jmcgirr@email.unc.edu

Word count abstract: 190 words; Word count main text: 5,014; Figures: 4.

Abstract

There are many cases of parallel gene expression underlying the evolution of convergent niche specialization, but parallel expression could also underlie divergent specialization. We investigated divergence in gene expression and genetic variation across three sympatric *Cyprinodon* pupfishes endemic to San Salvador Island, Bahamas. This recent radiation consists of a generalist and two derived specialists adapted to novel niches – a 'scale-eater' and a 'snail-eater.' We sampled total mRNA from all three species at two early developmental stages and compared gene expression with whole-genome genetic differentiation between all three species. 82% of genes that were differentially expressed between snail-eaters and generalists were up or downregulated in the same direction between scale-eaters and generalists; however, there were no shared fixed variants underlying this parallel expression. These genes showing parallel expression did not exhibit increased developmental constraints, but were enriched for effects on metabolic processes. Alternatively, genes showing divergent expression were enriched for effects on cranial skeleton development and pigment biosynthesis, reflecting the most divergent phenotypes observed between specialist species. Our findings reveal that convergent adaptation to higher trophic levels between divergent niche specialists through shared genetic pathways is governed by unique genetic variants.

Background

Abundant research on the genetic basis of adaptive traits has revealed an overarching pattern in nature – when species are faced with similar selective pressures, they often respond with the same genetic solutions (1). Parallel gene expression underlying convergent traits is a well-documented evolutionary phenomenon, with examples from experimental evolution studies imposing uniform selection pressures on replicate populations (2,3), studies in natural systems between closely related taxa (4–10), and distantly related taxa (11–13). This work has shown that parallelism at the level of gene expression is common in many cases of phenotypic convergence, particularly when divergence time between species is short (1,14).

However, few studies have investigated the extent of parallel gene expression contributing to species divergence, largely because most expression studies focus on only two species and are either concerned with divergent expression giving rise to divergent phenotypes (15–17) or parallel expression of specific loci (11,18,19) (but see 19,20). Furthermore, while there many genetic and demographic factors thought to influence the probability of parallel evolution (22,23), there are no theoretical expectations for the amount of parallel genetic variation contributing to parallel gene expression during ecological speciation (24,25).

Here we ask how both parallel and divergent expression patterns influence novel phenotypes by measuring transcriptomic and genomic divergence between three sympatric species of *Cyprinodon* pupfishes endemic to hypersaline lakes on San Salvador Island, Bahamas. This recent radiation consists of a dietary generalist species (*C. variegatus*) and two novel specialists: a ‘snail-eater’ (*C. brontotheroides*) and a ‘scale-eater’ (*C. desquamator*). Scale-eaters have large jaws and elongated bodies, whereas snail-eaters have short, thick jaws and a protruding nasal region that may function in crushing hard-shelled mollusks. These specialized species have evolved within the past 10,000 years based on the most recent drying of San Salvador lakes (26,27) and occupy unique niches within atherinomorph fishes (28). Scale-eaters and snail-eaters most likely diverged from a generalist common ancestor based on phylogenetic analyses of outgroup species and surveys of pupfish populations on neighboring Bahamian islands (28,29).

We performed total mRNA sequencing to examine gene expression in lab-reared individuals of all three species from different lake populations at two developmental stages. We

also searched for SNPs unique to each specialist and determined whether fixed variants near differentially expressed genes showed signs of hard selective sweeps (30). We found divergent genetic variation underlying trophic specialization, yet significant parallelism at the level of gene expression in specialists. We tested whether this counterintuitive result of parallel expression between divergent species may be due to 1) increased developmental constraint for genes showing parallelism or that 2) specialists experience parallel selective environments and adapt to higher trophic levels using similar genetic pathways.

Materials and Methods

Study system and sample collection

Individuals were caught from hypersaline lakes on San Salvador Island, Bahamas using a hand net or seine net in 2011, 2013, and 2015. Whole genome resequencing was performed for wild-caught individuals from a total of nine isolated lakes on San Salvador (Great Lake, Stout's Lake, Oyster Lake, Little Lake, Crescent Pond, Moon Rock, Mermaid's Pond, Osprey Lake, and Pigeon Creek). 14 scale-eaters were sampled from six populations; 11 snail-eaters were sampled from four populations; and 13 generalists were sampled from eight populations on San Salvador. Outgroup samples included one *C. laciniatus* from Lake Cunningham, New Providence Island, Bahamas, one *C. bondi* from Etang Saumautre lake in the Dominican Republic, one *C. diabolis* from Devil's Hole in California, and captive-bred individuals of *C. simus* and *C. maya* from Laguna Chicancanab, Quintana Roo, Mexico. Sampling is further described in (31,32) Fish were euthanized in an overdose of buffered MS-222 (Finquel, Inc.) following approved protocols from the University of California, Davis Institutional Animal Care and Use Committee (#17455) and University of California, Berkeley Animal Care and Use Committee (AUP-2015-01-7053) and stored in 95-100% ethanol.

RNA sequencing and alignment

Juvenile pupfish were derived from individuals raised in a common laboratory environment (25-27 C, 10-15 ppt salinity, pH 8.3) with identical diets. F_1 and F_2 fry were collected at two

developmental stages after hatching: 8-10 days post-fertilization and 17-20 dpf. They were euthanized in an overdose of buffered MS-222, and stored in RNA later (Ambion, Inc.) at 4 C for one day, followed by long-term storage at -20 C for up to one year. We extracted whole-body RNA using RNeasy kits (Qiagen) from 15 larvae (8-10 dpf) (Three generalists and snail-eaters from Crescent Pond and Little Lake and three scale-eaters from Crescent Pond; Table S1). We also dissected 14 larvae (17-20 dpf) to isolate RNA from the anterior craniofacial region containing the dentary, angular articular, maxilla, premaxilla, palatine, and associated craniofacial connective tissues using fine-tipped tweezers washed with RNAase-away (Three generalists and snail-eaters from Crescent Pond and Little Lake and two scale-eaters from Crescent Pond; Table S1). Our decision to sample RNA at 8-10 dpf and 17-20 dpf affects how we interpret patterns of gene expression. We chose to sample a range of days within each developmental stage to increase noise, thereby increasing the chance of identifying transcriptional variation that robustly influences larval development.

Libraries were prepared using the KAPA stranded mRNA-seq kit (KAPA Biosystems 2016) at the High Throughput Genomic Sequencing Facility at UNC Chapel Hill. Stranded sequencing on one lane of Illumina 150PE Hiseq4000 resulted in 677 million raw reads. We filtered raw reads using Trim Galore (v. 4.4, Babraham Bioinformatics) to remove Illumina adaptors and low-quality reads (mean Phred score < 20). We mapped these reads to the *Cyprinodon* reference genome using the RNA-seq aligner STAR (v. 2.5 (33)). We used the featureCounts function of the Rsubread package (34) requiring paired-end and reverse stranded options to generate read counts across previously annotated features. We assessed mapping and count quality using MULTIQC (35).

Differential expression analyses

We quantified differences in gene expression between all three species at two developmental stages. Our raw counts determined by featureCounts were normalized and used to identify differentially expressed transcripts with DEseq2 (v. 3.5 (36)). We performed four pairwise tests pooling species across lakes to identify differentially expressed transcripts among generalists *vs.* snail-eaters and generalists *vs.* scale-eaters at 8-10 dpf and 17-20 dpf (Table. S1). Genes with fewer than two read counts were discarded from all analyses (n = 1,570), along with genes

showing low normalized counts at a threshold determined by DEseq2 (using Cook's distance, 0.99 quantile (36)). Wald tests determined significant differences in expression between species by comparing normalized posterior log fold change estimates and correcting for multiple testing using the Benjamini–Hochberg procedure with a false discovery rate of 0.05 (37).

We used permutation tests to determine whether specialist species exhibited nonrandom patterns of parallel gene expression. We performed 100,000 permutations randomly sampling transcript IDs from the combined number of transcripts differentially expressed between generalists and each specialist followed by randomly sampling transcripts as either up or downregulated relative to generalists.

Genomic variant discovery and population genetic analyses

SNP variants were called using previously outlined methods (31,32). Briefly, 42 individual DNA samples extracted from muscle tissue were fragmented, barcoded with Illumina indices, and quality checked using a Fragment Analyzer (Advanced Analytical Technologies, Inc.). Sequencing on four lanes of Illumina 150PE Hiseq4000 resulted in 2.8 billion raw reads that were mapped from 42 individuals to the *Cyprinodon* reference genome (NCBI, *C. variegatus* Annotation Release 100, total sequence length = 1,035,184,475; number of scaffold = 9,259, scaffold N50, = 835,301; contig N50 = 20,803; (38)). We followed Genome Analysis Toolkit (v 3.5) best practices and hard filter criteria to call and refine our SNP variant dataset (QD < 2.0; FS < 60; MQRankSum < -12.5; ReadPosRankSum < -8 (39)). Our final SNP dataset contained 16 million variants and a mean sequencing coverage of 7× per individual (range: 5.2–9.3×).

We identified SNPs that were fixed in each specialist species. We calculated genome wide F_{st} using VCFtools' 'weir-fst-pop' function for two different population comparisons involving samples collected from San Salvador: generalists (n = 13) *vs.* snail-eaters (n = 11) and generalists (n = 13) *vs.* scale-eaters (n = 9). Our SNP dataset included 14 scale-eaters, however, we split our scale-eater population into two groups (large-jawed scale-eaters, n = 9 and small-jawed scale-eaters, n = 5) based on previous evidence that these two populations are genetically distinct (31,32). This allowed us to identify SNPs unique to large-jawed scale-eaters (i.e. *C. desquamator* (40)), which were the only type of scale-eater we sampled for RNA-seq. We

identified which of these SNPs resided in gene regions (either exonic, intronic, or within 10kb of the first or last exon) for genes showing differential expression. We determined whether these regions showed signatures of hard selective sweeps using SweeD ((30); methods previously described in (31)). Briefly, SweeD sections scaffolds into 1000 windows of equal size and calculates a composite likelihood ratio (CLR) using a null model where the site frequency spectrum of each window does not differ from that of the entire scaffold. Various demographic histories can shift the neutral site frequency spectrum making it difficult to infer signatures of selection (41,42). We previously estimated ancestral effective population sizes of San Salvador pupfishes using MSMC (31,43) and used these estimates to correct the expected neutral site frequency spectrum for the inferred recent population bottleneck in Caribbean pupfishes using SweeD. Windows with fixed SNPs that showed CLRs above the 95th percentile across their respective scaffolds (>10,000bp) under the assumptions of a population decrease determined by MSMC were interpreted as regions that recently experienced a hard sweep.

Measuring evolutionary constraint for differentially expressed genes

We performed gene ontology (GO) enrichment analyses for differentially expressed genes using GO Consortium resources available at geneontology.org (44). We used BlastP (v. 2.6 (45)) to identify zebrafish protein orthologs with high similarity (*E*-value < 1) to NCBI protein accessions for genes that we identified as differentially expressed between *Cyprinodon* species. We used these orthologs to determine if any gene ontology categories were enriched across differentially expressed genes. We grouped enriched GO categories into similar representative terms using the REVIGO clustering algorithm (46).

We also used these orthologs to estimate pleiotropy for differentially expressed genes based on the number of associated GO biological processes, protein-protein interactions (PPIs), and developmental stages when they are known to be expressed (47). We again used the GO Consortium (44) to determine the number of biological processes associated with each gene. We examined biological process annotations only for genes from ZFIN (zfin.org) with experimental evidence (GO evidence code EXP). The String protein database (v. 10; (48)) calculates a combined score measuring confidence in protein interactions by considering known interactions (experimentally determined and from manually curated databases) and predicted interactions.

We used the String database to quantify PPIs for protein products of differentially expressed genes, focusing only on interactions with experimental evidence (i.e. non-zero experimental evidence scores).

Next, we determined the number of developmental stages where a gene is known to be expressed using the Bgee expression call database for zebrafish (v. 14.0 (49)). We considered eight developmental stages from larval day five to juvenile day 89 from the Zebrafish Stage Ontology (ZFS) that were deemed 'gold quality,' meaning there was no contradicting call of absence of expression for the same gene, in the same developmental stage (49). We tested whether levels of gene pleiotropy were different in genes showing parallel expression in specialists versus divergently expressed genes by fitting a generalized linear model (negative binomial family; *glm.nb* function in the R library "MASS") on count data for pleiotropy estimates. We did not measure pleiotropy for genes expressed at 17-20 dpf due to the low number of zebrafish orthologs matched for genes with parallel expression in craniofacial tissues (11 out of 23).

Finally, we used the duplicated genes database (50) to identify paralogs in our differentially expressed gene dataset. We calculated whether genes showing parallel expression in specialists had more paralogs than divergently expressed genes using Pearson's Chi-square test in R.

Results

Differential expression between generalists and specialists

Total mRNA sequencing across all 29 samples resulted in 677 million raw reads, which was reduced to 674 million reads after quality control and filtering. 81.2% of these reads successfully aligned to the reference genome and 75.5% of aligned reads mapped to annotated features with an average read depth of 309× per sample. We used this dataset to identify differences in gene expression associated with specialist phenotypes at two developmental stages and within two groups of tissues.

We compared expression across 22,183 genes that showed greater than two read counts out 24,383 total genes annotated in the *Cyprinodon variegatus* assembly (NCBI, *C. variegatus* Annotation Release 100, (38)). These analyses revealed 1,014 genes (1,119 isoforms) differentially expressed between generalists *vs.* snail-eaters and 5,982 genes (6,776 isoforms) differentially expressed between generalists *vs.* scale-eaters (Fig. 1) (Benjamini and Hochberg adjusted $P < 0.05$). 833 genes (923 isoforms) showed the same direction of expression (up or down regulated) in both specialists relative to generalists, indicating parallel expression in specialists (Fig. 2). Specifically, 497 differentially expressed isoforms showed lower expression in both specialist species compared to generalists, while 424 showed higher expression in specialists (Fig. S1). This is significantly more parallel expression than would be expected by chance for genes (n = 6,996; 100,000 permutations; $P < 1.0 \times 10^{-5}$) and gene isoforms (n = 8,998; 100,000 permutations; $P < 7.0 \times 10^{-4}$) (Fig. S2).

Craniofacial morphology is the most rapidly diversifying trait in the San Salvador radiation (51). In order to detect genes expressed during jaw development, we compared expression within craniofacial tissue at the 17-20 dpf stage. We discovered 120 genes (123 isoforms) differentially expressed between generalists *vs.* snail-eaters and 1,903 genes (2,222 isoforms) differentially expressed between generalists *vs.* scale-eaters (Benjamini and Hochberg adjusted $P < 0.05$). We observed a lower proportion of parallel expression in specialists for craniofacial tissue compared to whole-body tissue. 23 genes showed the same direction of expression in specialists relative to generalists, with 11 differentially expressed isoforms showing lower expression in both specialist species and 14 showing higher expression. This is significantly less parallel expression than would be expected by chance for genes (n =2,023; 100,000 permutations; $P < 1.0 \times 10^{-5}$) and gene isoforms (n = 2,345; 100,000 permutations; $P < 1.0 \times 10^{-5}$) (Fig. S2).

We also identified 4 genes that showed opposite patterns of expression in specialists relative to generalists (Table S2). In 8-10 dpf whole-body tissue, 2 genes (*agxt2, si:ch211-197h24.9*) were upregulated in snail-eaters and downregulated in scale-eaters, while one gene (*plin2*) was upregulated in scale-eaters and downregulated in snail-eaters. In 17-20 dpf craniofacial tissue, one gene (*mybpc2a*) was upregulated in snail-eaters and downregulated in scale-eaters.

254

255 ***Parallel expression does not share the same genetic basis***

256 Similar to previous work on a smaller sample size (31), we identified 79 SNPs fixed between
257 generalists and snail-eaters and 1,543 SNPs fixed between generalists and scale-eaters. None of
258 these fixed variants were shared between specialists. Next, we determined which of these fixed
259 SNPs fell within gene regions (either exonic, intronic, or within 10kb of the first or last exon).
260 29% of SNPs fixed in snail-eaters overlapped with 21 gene regions, while 59% of SNPs fixed in
261 scale-eaters overlapped with 245 gene regions. We found 365 SNPs fixed in scale-eaters (24%)
262 within 68 gene regions that showed differential expression between generalists and scale-eaters
263 in whole-body tissue (8-10 dpf) and 123 SNPs (8%) within 26 gene regions differentially
264 expressed between generalists and scale-eaters in craniofacial tissue (17-20 dpf). We suspect that
265 some of these fixed variants are causal cis-regulatory variants responsible for species specific
266 expression patterns that ultimately give rise to phenotypic differences in scale-eaters.
267 Conversely, we only identified a single SNP fixed in snail-eaters within a gene (*tmprss2*) that
268 was differentially expressed between generalists and snail-eaters in whole-body tissue (8-10 dpf).
269 We did not find any fixed variants near genes differentially expressed between generalists and
270 snail-eaters in craniofacial tissue, possibly suggesting that fixed variants regulate expression
271 divergence at an earlier developmental.

272 Since we did not find any variants that were fixed between scale-eaters and generalists
273 that were also fixed between snail-eaters and generalists, we repeated our analysis with a lower
274 threshold of genetic divergence by examining SNPs with $F_{st} \geq 0.8$ between generalists *vs.* snail-
275 eaters and generalists *vs.* scale-eaters. Again, none of these SNPs were shared between
276 specialists. We found 11,491 divergent SNPs within 810 gene regions differentially expressed
277 between scale-eaters and generalists. We also found 297 SNPs within 27 gene regions
278 differentially expressed between snail-eaters and generalists. Strikingly, only three genes out of
279 833 that showed parallel expression in specialists contained SNPs with $F_{st} \geq 0.8$ between
280 generalists and each of the specialists. This suggests that parallel expression is not a result of the
281 same genetic variation in specialists.

282

Pleiotropic constraint and functional redundancy do not explain parallel expression

We predicted that genes showing parallel expression patterns in specialists may be constrained by negative pleiotropy. High gene pleiotropy is correlated with participation in more protein-protein interactions (PPIs), which in turn effects multiple biological processes (52,53). Genes that act across multiple developmental stages are also more pleiotropic (54). We estimated pleiotropy for differentially expressed genes based on the number of protein-protein interactions (PPIs), associated GO biological processes, and developmental stages when they are known to be expressed. However, we did not find evidence for higher pleiotropy for genes showing parallel expression patterns in specialists compared to genes showing divergent expression using any of these three metrics (GLM; biological processes: $P = 0.67$; PPIs: $P = 0.09$; developmental stages: $P = 0.89$) (Fig. 3).

Next, we predicted that genes showing divergent expression patterns in specialists might be under less developmental constraint because they share redundant functions with paralogous genes. We searched the duplicated genes database (50) and found that datasets for genes with divergent expression patterns in specialists showed similar numbers of paralogs to genes with parallel expression patterns in specialists (χ^2, 8-10 dpf: 48 duplicated genes out of 627 genes showing parallel expression and 270 duplicated genes out of 3,491 divergently expressed genes, $P = 0.95$; 17-20 dpf: 2 duplicated genes out of 11 genes showing parallel expression and 41 duplicated genes out of 663 divergently expressed genes, $P = 0.11$).

Genes showing parallel expression are enriched for metabolic processes

We performed GO enrichment analyses with zebrafish orthologs for genes showing parallel expression patterns in specialists (n = 620) and genes showing divergent expression patterns in scale eaters (n = 3,349) and snail-eaters (n = 102). We restricted these analyses to genes expressed at 8-10 dpf because the number of genes showing parallel expression in specialists at 17-20 dpf (n = 23) was low and did not show enrichment for any biological process.

Genes showing parallel expression in trophic specialists were enriched for metabolic processes (20% of GO terms). In contrast, genes with divergent expression patterns in specialists were enriched for cranial skeletal development and pigment biosynthesis (7% and 3% of terms,

312 respectively) while only 11% of enriched categories described metabolic processes (Fig. 4, Table
313 S3). We identified genes with divergent expression between specialists enriched for effects on
314 viscerocranium, striated muscle, liver, erythrocyte, and the development of the heart, optic cup,
315 pancreas, and brain. 41 divergently expressed genes were annotated for viscerocranium
316 development, which is particularly interesting because this tissue gives rise to the most divergent
317 craniofacial structures in specialists (55,56).

318

319 ***The genetic basis of extreme craniofacial divergence***

320 We previously described 30 candidate gene regions containing variants fixed between trophic
321 specialist species associated with variation in jaw length (31). These candidates also showed
322 signatures of a recent hard selective sweep (31). Encouragingly, we found eleven of these genes
323 differentially expressed between generalists and specialists (seven at 8-10 dpf and four at 17-20
324 dpf) (Table S4).

325 We searched for variants fixed in specialists within gene regions across all 7,394 genes
326 that were differentially expressed between either generalists and snail-eaters or generalists and
327 scale-eaters. We discovered fixed SNPs in 81 of these gene regions (either intronic or within
328 10kb of the first or last exon), potentially implicating cis-acting variants regulating gene
329 expression. Further testing this hypothesis, we searched for signatures of hard selective sweeps
330 in specialists. Interestingly, 95% of these gene regions containing fixed SNPs showed signs of
331 experiencing a selective sweep (estimated by SweeD; CLR > 95^{th} percentile across their
332 respective scaffolds) (Table S5). All of these genes regions contained SNPs fixed between
333 generalists and scale-eaters and showed differential expression in this same comparison. Five of
334 these 81 genes were indicated as strong craniofacial candidates in our previous study (31).
335 Finally, we compared this list of genes experiencing selection to those annotated for cranial
336 skeletal system development (GO:1904888) and muscle organ development (GO:0007517). This
337 revealed three genes containing fixed variation in scale-eaters that likely influence craniofacial
338 divergence through cis-acting regulatory mechanisms (*loxl3b* (annotated for cranial effects)*;*
339 *fbxo32* and *klhl40a* (annotated for muscle effects)).

340

Discussion

We combined RNA sequencing with genome-wide divergence scans to study the molecular evolution of two trophic specialist species that rapidly diverged from a generalist common ancestor within the last 10,000 years. We examined how gene expression and SNP variation influence snail-eater and scale-eater niche adaptations using comparisons between each specialist and their generalist sister species. We found a significant amount of parallelism at the level of gene expression yet no parallelism at the level of fixed genetic variation in specialists. Specifically, 82% of genes that were differentially expressed between snail-eaters and generalists were up or downregulated in the same direction when comparing expression between scale-eaters and generalists (Fig. 2). We show that this parallel expression is not the result of the same underlying genetic variation.

We explored two possible explanations for this pattern: 1) developmental constraints resulting from pleiotropy or functional redundancy favor parallel expression in specialist species or 2) parallel selective environments act on similar genetic pathways needed for adaptation to higher trophic levels.

1. Developmental constraints on the evolution of parallel gene expression

We tested two hypotheses considering how different forms of developmental constraint could promote parallel expression patterns in specialist species. One mechanism constraining the probability of parallelism is pleiotropy – when the expression of one transcript influences multiple phenotypic characteristics (22,47,57). Higher gene pleiotropy is correlated with participation in more protein-protein interactions (PPIs), which can cause cascading downstream effects on multiple biological processes (52,53). We predicted that genes showing parallel expression might be under higher constraint due to negative pleiotropy. We estimated three measures of gene pleiotropy (number of associated GO biological processes, protein-protein interactions (PPIs), and developmental stages when they are known to be expressed) and found no significant difference in any measure for genes showing parallel versus divergent expression patterns in specialists (Fig. 3). This finding is consistent with some empirical evidence and

theoretical models of gene expression evolution that found pleiotropy constrains the variability of gene expression within species, but does not hinder divergence between species (15,58).

Gene sequence similarity can also affect the probability of parallelism (22). Genes that are similar due to duplication events (paralogs) may be less constrained and therefore more likely to exhibit divergent expression (7,22,59–61). Given that some paralogs are known to share molecular functions even when they have had much time to diverge (62–64), we predicted that the set of genes showing divergent expression might be under less constraint because it was enriched for paralogs exhibiting functional redundancy. However, we did not find any difference in the proportion of paralogs for genes showing parallel and divergent expression in specialists (χ^2; 8-10 dpf: $P = 0.95$; 17-20 dpf $P = 0.11$). Overall, we find that parallel expression in specialists is not the result of increased developmental constraints.

2. Parallel gene expression underlies convergent metabolic adaptations to a higher trophic level in each specialist

While the specialists are more morphologically diverged from one another than either is from generalist species, particularly in their craniofacial phenotype and male reproductive coloration (40,65) (Fig. 4), dietary isotope analyses show that they occupy a higher trophic level than generalists (66). Fish scales and mollusks contribute to more nitrogen-rich diets in specialists compared to generalist species that primarily consume algae and detritus (66). Perhaps the same metabolic processes required for this type of diet are adaptive at higher trophic levels for both scale-eaters and snail-eaters, which might explain patterns of parallel expression. Thus, we predicted that genes showing parallel expression would affect metabolic processes that may be similar between specialists, whereas genes showing divergent expression in specialists would affect developmental processes.

GO enrichment analyses support both hypotheses. We found that 20% of GO categories enriched for genes showing parallel expression described metabolic processes, and zero described cranial skeletal development or pigment biosynthesis (Fig. 4A). In contrast, 10% of terms showing enrichment in the divergently expressed gene set described developmental processes (cranial skeletal development and pigment biosynthesis) and only 11% described

metabolic processes (Fig 4A, Table S4). These results may suggest that parallel expression in specialists is important for adapting to a higher trophic level, in which snail-eating and scale-eating present similar metabolic requirements relative to the lower trophic level of generalists. This is consistent with the high macroalgae content of generalist diets relative to both specialist species (51) and the shorter intestinal lengths observed in both specialists relative to the generalist (CHM and JAM personal observation). In contrast, genes showing divergent expression in specialists are more important for shaping divergent skeletal and pigmentation phenotypes between species (Fig. 4B, C and D).

Parallel gene expression despite divergent genetic variation

We find significant parallel gene expression across genes that are annotated for effects on metabolism, yet shared expression patterns do not seem to be driven by the same fixed variants. This is surprising in this young radiation given that the probability of parallel genetic variation underlying phenotypic convergence increases with decreasing divergence time (1,24,67). Although 95% of differentially expressed gene regions containing fixed SNPs show signs of experiencing a selective sweep, and almost none of these variants in these SNPs were in exons, it is still possible that fixed alleles do not regulate parallel expression of metabolic genes. Instead, we reasoned that segregating variants shared between specialists might be responsible for shared expression patterns. However, we still failed to find any shared variants at lower levels of differentiation ($F_{st} \geq 0.8$). Another possibility is that parallel expression is not controlled by shared SNPs, but rather a more complex type of genetic variation (copy number variants, inversions, insertions, deletions). However, preliminary results suggest that there are no shared insertion/deletions between specialists (JAM unpublished data). Many studies on parallel adaptation through gene reuse describe convergence within pigmentation and skeletal development pathways (6,18,23,68). Perhaps the architecture of metabolic adaptation is more flexible, having more mutational targets or employing more late-acting developmental regulatory networks that are less constrained than early-acting networks (67,69–73). Our findings highlight the importance of understanding the probability of genetic convergence across different biological processes.

Caveats to gene expression analyses and the robustness of parallel expression

We chose to sample RNA at 8-10 dpf and 17-20 dpf to identify transcriptional variation that influences larval development, however, some activation of parallel gene networks is likely specified at pre-hatching developmental stages (69,70). It is also possible that we did not have the power to identify subtle differences in expression for genes that showed high divergence between specialists and generalists. Detecting differential expression of transcripts is notoriously difficult when read counts are low and variance within treatment groups is high (74,75). We were able to detect differential expression for genes with a mean normalized count as low as 1.6 (median = 150) and $\log_2$ fold change as low as 0.2 (median = 1.11). Thus, we could not detect any variation causing very slight fold changes in expression below 0.2. Furthermore, our scale-eater sample sizes (8-10 dpf n = 3; 17-20 dpf n = 2) were lower than that of generalists and snail-eaters (n = 6 at both stages) (Fig. S1), which likely resulted in more false positives for scale-eater comparisons.

Finally, our results are robust in a recently published independent analysis of gene expression in San Salvador pupfishes that identified many of the same genes we found divergently expressed between specialists (38). We examined this dataset using the same significance thresholds for differentially expressed genes as described in Lencer *et al.* for mRNA extracted from all three species at 8 dpf and 15dpf ($p < 0.1$ and $|\text{Log}_2 \text{ fold change}| > 0.2$). We found that 40% of genes divergently expressed between specialists in this dataset were divergently expressed in our own dataset. Importantly, Lencer *et al.* only examined cranial tissues at both of these developmental stages, and they did not investigate parallel expression between each specialist. Next we searched for evidence of parallel expression for mRNA extracted from all three species at 8 dpf. 28.8% of genes that were differentially expressed between snail-eaters and generalists were up or downregulated in the same direction between scale-eaters and generalists. This is a lower proportion of parallel expression than we identified (Fig. 2), but this is most likely because Lencer *et al.* only sampled RNA from cranial tissues (38). Thus, the majority of parallel expression between specialists likely occurs in non-cranial tissues, consistent with our shared metabolic hypothesis.

Conclusion

Here we find that divergent genetic variation underlies significant parallel gene expression in two novel specialist species. While there are many cases of parallel expression underlying parallel ecological specialization, to our knowledge, this represents the first case of parallel expression accompanying divergent specialization. Our findings will be useful in developing theoretical predictions of parallelism contributing to species divergence at both the gene and nucleotide level. Numerous studies have shown that the probability of parallel genetic variation underlying phenotypic convergence is higher when divergence time between species is short (1,24,67). Scale-eating and snail-eating species have evolved rapidly within the last 10,000 years, yet parallel expression patterns are not driven by the same underlying genetic variants, arguing that some forms of phenotypic convergence between young species do not exhibit parallelism at the nucleotide level. We show that parallel expression is not a result of developmental constraint imposed by negative pleiotropy or functional redundancy, but most likely reveals convergent adaptation to a higher trophic level in each trophic specialist, despite their specialization on highly divergent resources.

Competing interests

We declare we have no competing interests.

Funding

This study was funded by the University of North Carolina at Chapel Hill and the Miller Institute for Basic Research in the Sciences.

Acknowledgements

We thank Jelmer Poelstra and Sara Suzuki for valuable discussions and computational assistance; the High Throughput Genomic Sequencing Facility at UNC Chapel Hill for performing RNA library prep and Illumina sequencing; the Gerace Research Centre for accommodation; and the Bahamian government BEST Commission for permission to conduct

485 this research. All datasets used for this study will be deposited in the NCBI Short Read Archive
486 associated with BioProject PRJNA391309.

487 **Author Contributions**

488 JAM wrote the manuscript. JAM extracted the RNA samples, and conducted all bioinformatic
489 and population genetic analyses. Both authors contributed to the conception and development of
490 the ideas and revision of the manuscript.

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510 **References**

511 1. Conte GL, Arnegard ME, Peichel CL, Schluter D. The probability of genetic parallelism
512 and convergence in natural populations. 2012;(October):5039–47.
513 2. Cooper TF, Rozen DE, Lenski RE. Parallel changes in gene expression after 20 , 000
514 generations of evolution in Escherichia coli. 2003;2002(Track II).
515 3. Riehle MM, Bennett AF, Lenski RE, Long AD, Michelle M, Bennett AF, et al.
516 Evolutionary changes in heat-inducible gene expression in lines of Escherichia coli
517 adapted to high temperature. 2003;47–58.
518 4. Derome N, Bernatchez L. The Transcriptomics of Ecological Convergence between 2
519 Limnetic Coregonine Fishes (Salmonidae). 2004;
520 5. Chan YF, Marks ME, Jones FC, Jr GV, Shapiro MD, Brady SD, et al. Adaptive Evolution
521 of Pelvic Reduction of a Pitx1 Enhancer. 2010;327(January):302–6.
522 6. Reed RD, Papa R, Martin A, Hines HM, Kronforst MR, Chen R, et al. optix Drives the
523 Repeated. 2011;(August).
524 7. Nagai H, Terai Y, Sugawara T, Imai H, Nishihara H, Hori M, et al. Reverse Evolution in
525 RH1 for Adaptation of Cichlids to Water Depth in Lake Tanganyika Research article
526 Determination of Sequences Upstream. 2011;28(6):1769–76.
527 8. Manousaki T, Hull PM, Kusche H, Machado-schiaffino G, Franchini P, Harrod C. Parsing
528 parallel evolution : ecological divergence and differential gene expression in the adaptive
529 radiations of thick-lipped Midas cichlid fishes from Nicaragua. 2013;22:650–69.
530 9. Reid NM, Proestou DA, Clark BW, Warren WC, Colbourne JK, Shaw JR, et al. The
531 genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild
532 fish.
533 10. Zhao L, Wit J, Svetec N, Begun DJ. Parallel Gene Expression Differences between Low
534 and High Latitude Populations of Drosophila melanogaster and D . simulans. 2015;1–25.
535 11. Shapiro MD, Bell MA, Kingsley DM. Parallel genetic origins of pelvic reduction in
536 vertebrates. 2006;103(37):13753–8.
537 12. Miller CT, Swartz ME, Khuu PA, Walker MB, Eberhart JK, Kimmel CB. mef2ca is

required in cranial neural crest to effect Endothelin1 signaling in zebrafish. 2007;308:144–57.

13. Shen Y, Liang L, Li G, Murphy RW, Zhang Y. Parallel Evolution of Auditory Genes for Echolocation in Bats and Toothed Whales. 2012;8(6).

14. Losos JB. CONVERGENCE , ADAPTATION , AND CONSTRAINT. 2011;1827–40.

15. Bolivar P, Burri R, Dutoit L, Mugal CF, Nater A, Aken B, et al. Divergence in gene expression within and between two closely related flycatcher species. 2016;2015–28.

16. Davidson JH, Balakrishnan CN. Gene Regulatory Evolution During Speciation in a Songbird. 2016;6(May):1357–64.

17. Poelstra JW, Vijay N, Bossu CM, Lantz H, Ryll B, Müller I, et al. The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. Science [Internet]. 2014 Jun 20 [cited 2016 Mar 9];344(6190):1410–4. Available from: http://science.sciencemag.org/content/344/6190/1410.abstract

18. Miller CT, Beleza S, Pollen AA, Schluter D, Kittles RA, Shriver MD, et al. cis - Regulatory Changes in Kit Ligand Expression and Parallel Evolution of Pigmentation in Sticklebacks and Humans. 2007;1179–89.

19. Quin KEO, Hofmann CM, Hofmann HA, Carleton KL. Parallel Evolution of Opsin Gene Expression in African Cichlid Fishes Research article. 2010;27(12):2839–54.

20. Ahi EP, Kapralova KH, Pálsson A, Maier VH, Gudbrandsson J, Snorrason SS, et al. Transcriptional dynamics of a conserved gene expression network associated with craniofacial divergence in Arctic charr. 2014;1–19.

21. Enard W, Khaitovich P, Klose J, Heissig F, Giavalisco P, Nieselt-struwe K, et al. Intra- and Interspecific Variation in Primate Gene Expression Patterns. 2002;296(April):340–4.

22. Rosenblum EB, Parent CE, Brandt EE. The Molecular Basis of Phenotypic Convergence.

23. Conte GL, Arnegard ME, Peichel CL, Schluter D. The probability of genetic parallelism and convergence in natural populations. Proc R Soc B Biol Sci [Internet]. 2012 Oct 17 [cited 2015 Oct 5];279(1749):5039–47. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3497250&tool=pmcentrez&rendertype=abstract

24. Schluter D, Clifford EA, Nemethy M, Mckinnon JS. Parallel Evolution and Inheritance of Quantitative Traits. 2004;163(6).

25. Pavey SA, Nosil P, Rogers SM. The role of gene expression in ecological speciation. 2010;1206:110–29.
26. Mylroie, J.E, Hagey FM. Terrestrial and Shallow Marine Geology of the Bahamas and Bermuda. Boulder, CO: Geological Society of America; 1995. 77-90 p.
27. Turner BJ, Duvernell DD, Bunt TM, Barton MG. Reproductive isolation among endemic pupfishes (Cyprinodon) on San Salvador Island, Bahamas: Microsatellite evidence. Biol J Linn Soc. 2008;95(3):566–82.
28. Martin CH, Wainwright PC. Trophic novelty is linked to exceptional rates of morphological diversification in two adaptive radiations of Cyprinodon pupfish. Evolution [Internet]. 2011 Aug [cited 2016 Jan 14];65(8):2197–212. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21790569
29. Martin CH. Context-dependence in complex adaptive landscapes□: frequency and trait-dependent selection surfaces within an adaptive radiation of Caribbean pupfishes. Evolution (N Y). 2016;(Simpson 1944):1–57.
30. Pavlidis P, Živković D, Stamatakis A, Alachiotis N. SweeD: Likelihood-based detection of selective sweeps in thousands of genomes. Mol Biol Evol. 2013;30(9):2224–34.
31. Mcgirr JA, Martin CH. Novel Candidate Genes Underlying Extreme Trophic Specialization in Caribbean Pupfishes. 2016;34(4):873–88.
32. Richards EJ, Martin CH. Adaptive introgression from distant Caribbean islands contributed to the diversification of a microendemic adaptive radiation of trophic specialist pupfishes. 2017. 1-35 p.
33. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. Sequence analysis. 2013;29(1):15–21.
34. Liao Y, Smyth GK, Shi W. Sequence analysis featureCounts□: an efficient general purpose program for assigning sequence reads to genomic features. 2014;30(7):923–30.
35. Ewels P, Lundin S, Max K. Data and text mining MultiQC□: summarize analysis results for multiple tools and samples in a single report. 2016;32(June):3047–8.
36. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. 2014;1–21.
37. Society RS. Controlling the False Discovery Rate□: A Practical and Powerful Approach to Multiple Testing Author (s): Yoav Benjamini and Yosef Hochberg Source□: Journal

of the Royal Statistical Society . Series B (Methodological), Vol . 57 , No . 1 Published by□: Wiley for the Royal Statistical Society Stable URL□: http://www.jstor.org/stable/2346101. 2017;57(1):289–300.

38. Lencer ES, Warren WC, Harrison R, Mccune AR. The Cyprinodon variegatus genome reveals gene expression changes underlying differences in skull morphology among closely related species. 2017;1–33.

39. DePristo MA, Banks E, Poplin R, Garimella K V, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet [Internet]. 2011 May [cited 2014 Jul 9];43(5):491–8. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3083463&tool=pmcentrez&rendertype=abstract

40. Martin CH, Wainwright PC. A remarkable species flock of Cyprinodon pupfishes endemic to San Salvador Island, Bahamas. Bull Peabody Museum Nat Hist [Internet]. 2013;54(2):231–41. Available from: file:///Users/antoniaford/Library/Application Support/Papers2/Articles/2013/Martin/Martin_Bulletin_of_the_Peabody_Museum_of_Natural_History_2013-1.pdf%5Cnhttp://www.bioone.org/doi/abs/10.3374/014.054.0201

41. Galtier N, Depaulis F, Barton NH. Detecting Bottlenecks and Selective Sweeps From DNA Sequence Polymorphism. 2000;

42. Nielsen R. Molecular signatures of natural selection. Annu Rev Genet [Internet]. 2005;39(1):197–218. Available from: citeulike-article-id:524184%5Cnhttp://dx.doi.org/10.1146/annurev.genet.39.073003.112420

43. Schiffels S, Durbin R. Inferring human population size and separation history from multiple genome sequences. Nat Publ Gr [Internet]. 2014;46(8):919–25. Available from: http://dx.doi.org/10.1038/ng.3015

44. Gene T, Consortium O. Gene Ontology□: tool for the. 2000;25(may):25–9.

45. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST +□: architecture and applications. 2009;9:1–9.

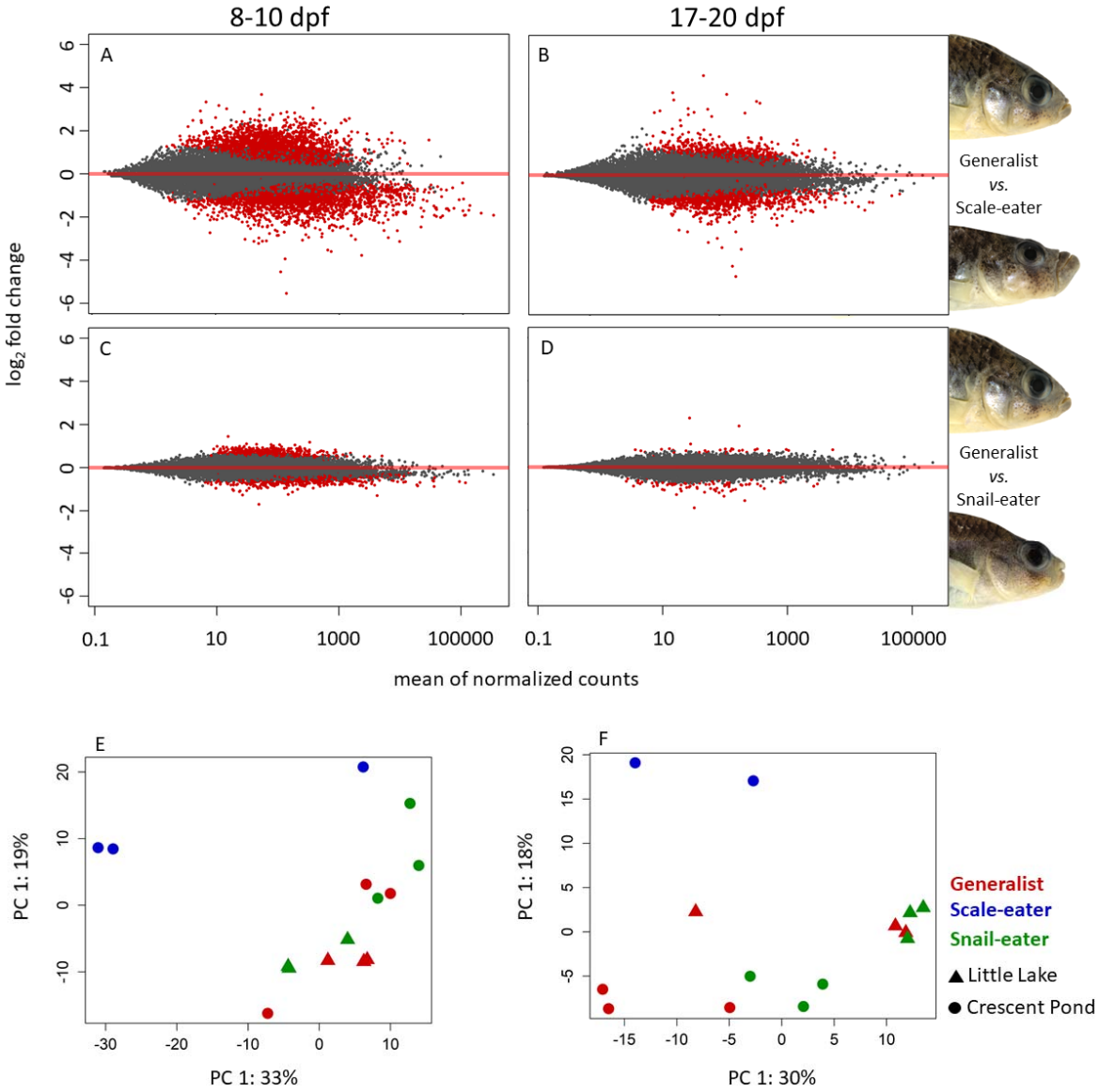
46. Tomislav S. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. 2011;6(7).

47. Papakostas S, Vøllestad LA, Bruneaux M, Aykanat T, Vanoverbeke J, Ning M, et al. conditions. 2014;(316):1–9.

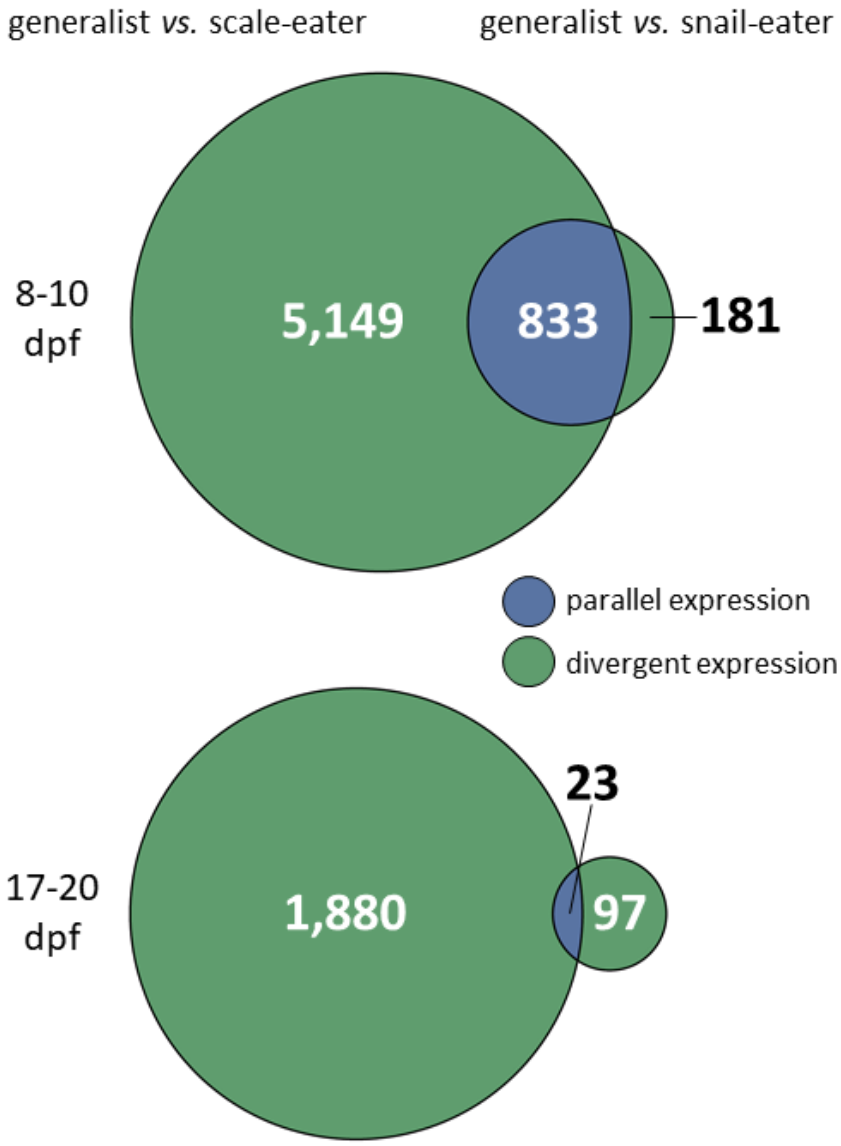
631 48. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-cepas J, et al.
632 STRING v10□: protein – protein interaction networks , integrated over the tree of life.
633 2015;43(October 2014):447–52.
634 49. Bastian F, Parmentier G, Roux J, Moretti S, Lyon U De, Génomique I De, et al. Bgee□:
635 Integrating and Comparing Heterogeneous Transcriptome Data Among Species. :124–31.
636 50. Ouedraogo M, Bettembourg C, Bretaudeau A, Sallou O, Diot C. The Duplicated Genes
637 Database□: Identification and Functional Annotation of Co-Localised Duplicated Genes
638 across Genomes. 2012;7(11):1–8.
639 51. Martin CH, Wainwright PC. On the measurement of ecological novelty: scale-eating
640 pupfish are separated by 168 my from other scale-eating fishes. PLoS One [Internet].
641 2013;8(8):e71164. Available from:
642 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3747246&tool=pmcentrez&re
643 ndertype=abstract
644 52. He X, Zhang J. Toward a Molecular Understanding of Pleiotropy.
645 2006;1891(August):1885–91.
646 53. Safari-alighiarloo N, Taghizadeh M, Rezaei-tavirani M, Goliaei B, Peyvandi AA. Protein-
647 protein interaction networks (PPI) and complex diseases. 2014;7(7):17–31.
648 54. Stern DL, Orgogozo V. THE LOCI OF EVOLUTION□: HOW PREDICTABLE IS
649 GENETIC EVOLUTION□? 2008;2155–77.
650 55. Article R. Developmental Patterning and Evolution of the Mammalian Viscerocranium□:
651 1997;155(September 1996):139–55.
652 56. Cerny R, Lwigale P, Ericsson R, Meulemans D, Epperlein H, Bronner-fraser M.
653 Developmental origins and evolution of jaws□: new interpretation of b maxillary Q and b
654 mandibular Q. 2004;276:225–36.
655 57. Wagner GP, Zhang J. The pleiotropic structure of the genotype – phenotype map□: the
656 evolvability of complex organisms. Nat Publ Gr [Internet]. 2011;12(3):204–13. Available
657 from: http://dx.doi.org/10.1038/nrg2949
658 58. Tulchinsky AY, Johnson NA, Porter AH. Hybrid Incompatibility Despite Pleiotropic
659 Constraint in a Sequence-Based Bioenergetic. 2014;198(December):1645–54.
660 59. Zhang J. Parallel Functional Changes in the Digestive RNases of Ruminants and
661 Colobines by Divergent Amino Acid Substitutions. 2001;1310–7.

60. Carvalho LS, Cowing JA, Wilkie SE, Bowmaker JK, Hunt DM. The Molecular Evolution of Avian Ultraviolet- and Violet-Sensitive Visual Pigments. 2006;
61. Holtzinger A, Evans T. NIH Public Access. 2008;312(2):613–22.
62. Lynch M, Force A. The Probability of Duplicate Gene Preservation by Subfunctionalization. 2000;
63. Kassahn KS, Dang VT, Wilkins SJ, Perkins AC, Ragan MA. Evolution of gene function and regulatory control after whole-genome duplication : Comparative analyses in vertebrates. 2009;1404–18.
64. Qian W, Liao B, Chang AY, Zhang J. Maintenance of duplicate genes and their functional redundancy by reduced expression. Trends Genet [Internet]. 26(10):425–30. Available from: http://dx.doi.org/10.1016/j.tig.2010.07.002
65. MARTIN, P. ERICKSON CM. The genetic architecture of novel trophic specialists : larger effect sizes are associated with exceptional oral jaw diversification in a pupfish adaptive radiation. 2017;624–38.
66. Martin CH. The cryptic origins of evolutionary novelty : 1000-fold faster trophic diversification rates without increased ecological opportunity or hybrid swarm. 2016;
67. Martin A, Orgogozo V. THE LOCI OF REPEATED EVOLUTION : A CATALOG OF GENETIC HOTSPOTS OF PHENOTYPIC VARIATION. 2013;(McGhee 2011):1235–50.
68. Kronforst MR, Barsh GS, Kopp A, Mallet J, Monteiro A, Mullen SP, et al. INTERNATIONAL FEDERATION OF PIGMENT CELL SOCIETIES · SOCIETY FOR MELANOMA RESEARCH PIGMENT CELL & MELANOMA of diversity and convergence in animal pigmentation. 2012;
69. Garfield DA, Runcie DE, Babbitt CC, Haygood R, Nielsen WJ, Wray GA. The Impact of Gene Expression Variation on the Robustness and Evolvability of a Developmental Gene Regulatory Network. 2013;11(10).
70. Ferna A, Tena JJ, Gonza C, Parra-acero H, Cross JW, Rigby PWJ, et al. Comparative epigenomics in distantly related teleost species identifies conserved cis -regulatory nodes active during the vertebrate phylotypic period. 2014;1075–85.
71. Kalinka AT, Varga KM, Gerrard DT, Preibisch S, Corcoran DL, Jarrells J, et al. developmental hourglass model. Nature [Internet]. 2010;468(7325):811–4. Available

693 from: http://dx.doi.org/10.1038/nature09634
694 72. Reddiex AJ, Gosden TP, Bonduriansky R, Chenoweth SF. Sex-Specific Fitness
695 Consequences of Nutrient Intake and the Evolvability of Diet Preferences. 2013;182(1).
696 73. Comeault AA, Serrato-capuchina A, Turissini DA, Mclaughlin PJ, David JR, Matute DR.
697 A nonrandom subset of olfactory genes is associated with host preference in the fruit fly
698 Drosophila orena. 2017;1–13.
699 74. Conesa A, Madrigal P, Tarazona S, Gomez-cabrero D, Cervera A, Mcpherson A, et al. A
700 survey of best practices for RNA-seq data analysis. 2016;1–19.
701 75. Lin Y, Golovnina K, Chen Z, Lee HN, Negron YLS, Sultana H, et al. Comparison of
702 normalization and differential expression analyses using RNA-Seq data from 726
703 individual Drosophila melanogaster. BMC Genomics [Internet]. 2016;1–20. Available
704 from: http://dx.doi.org/10.1186/s12864-015-2353-z
705
706
707
708
709
710
711
712
713
714
715
716
717
718



719 **Fig. 1. Differential gene expression between generalists and trophic specialists.** Red points
720 represent genes that are differentially expressed in 8-10 dpf whole-body tissue (A, C) and 17-20
721 dpf craniofacial tissue (B, C) between generalists *vs.* scale-eaters (A, B) and generalist *vs.* snail-
722 eaters (C, D). Bottom panels show the top two principal components accounting for a combined
723 52% (8-10 dpf; E) and 48% (17-20 dpf; F) of the total variation between samples across 413
724 million reads mapped to annotated features. Triangles represent samples from Little Lake and
725 circles represent samples from Crescent Pond. Generalists are shown in red, scale-eaters in blue,
726 and snail-eaters in green.



727

728

729 **Fig. 2. Significant parallel gene expression between specialists despite divergent trophic**
730 **adaptation.** Circles illustrate genes differentially expressed in 8-10 dpf whole-body tissue (top)
731 and 17-20 dpf craniofacial tissue (bottom) for generalists *vs.* scale-eaters (left) and generalists *vs.*
732 scale-eaters (right). Genes that show the same expression patterns in specialists relative to
733 generalists are blue, and those showing divergent expression patterns unique to each specialist
734 are green. Significantly more genes show the same expression pattern (either up or
735 downregulated) in specialists relative to generalist gene expression at 8-10 dpf than expected by
736 chance (Fig. S2; 100,000 permutations; $P < 1.0 \times 10^{-5}$). Alternatively, significantly fewer genes
737 show the same expression pattern in 17-20 dpf craniofacial tissue (Fig. S2; 100,000
738 permutations; $P < 7.0 \times 10^{-4}$).

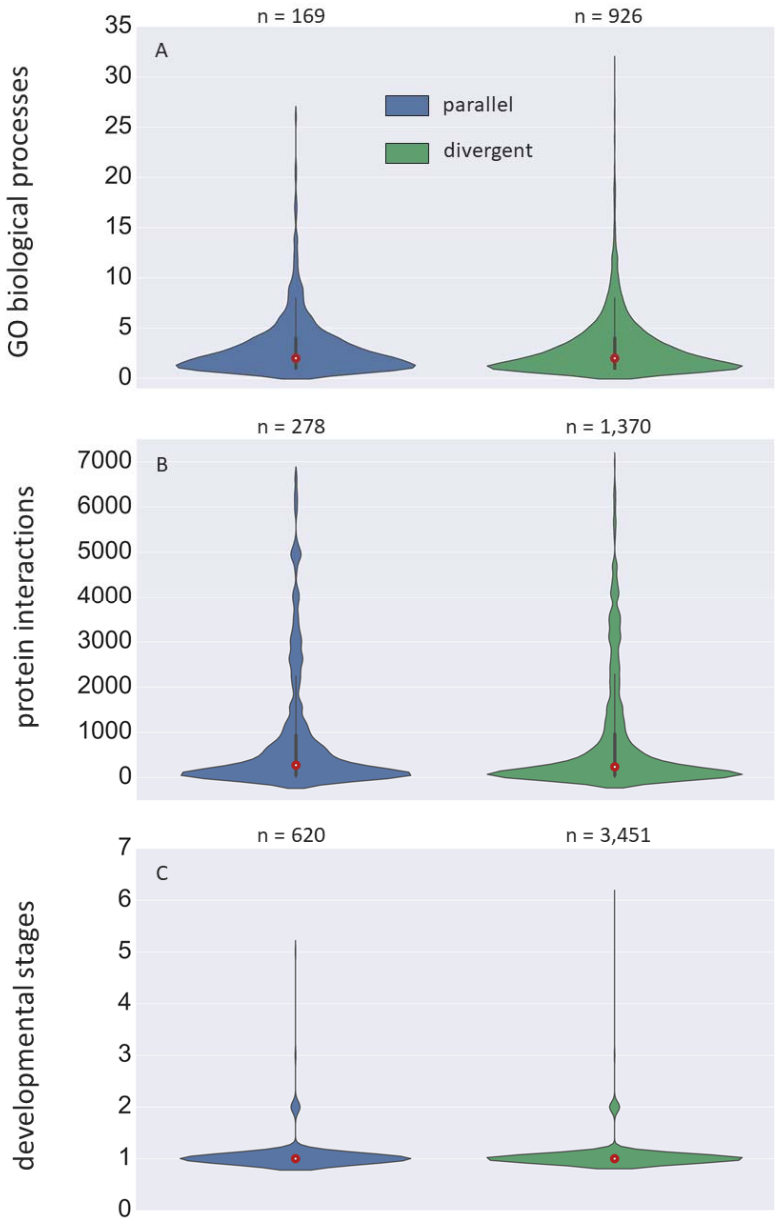
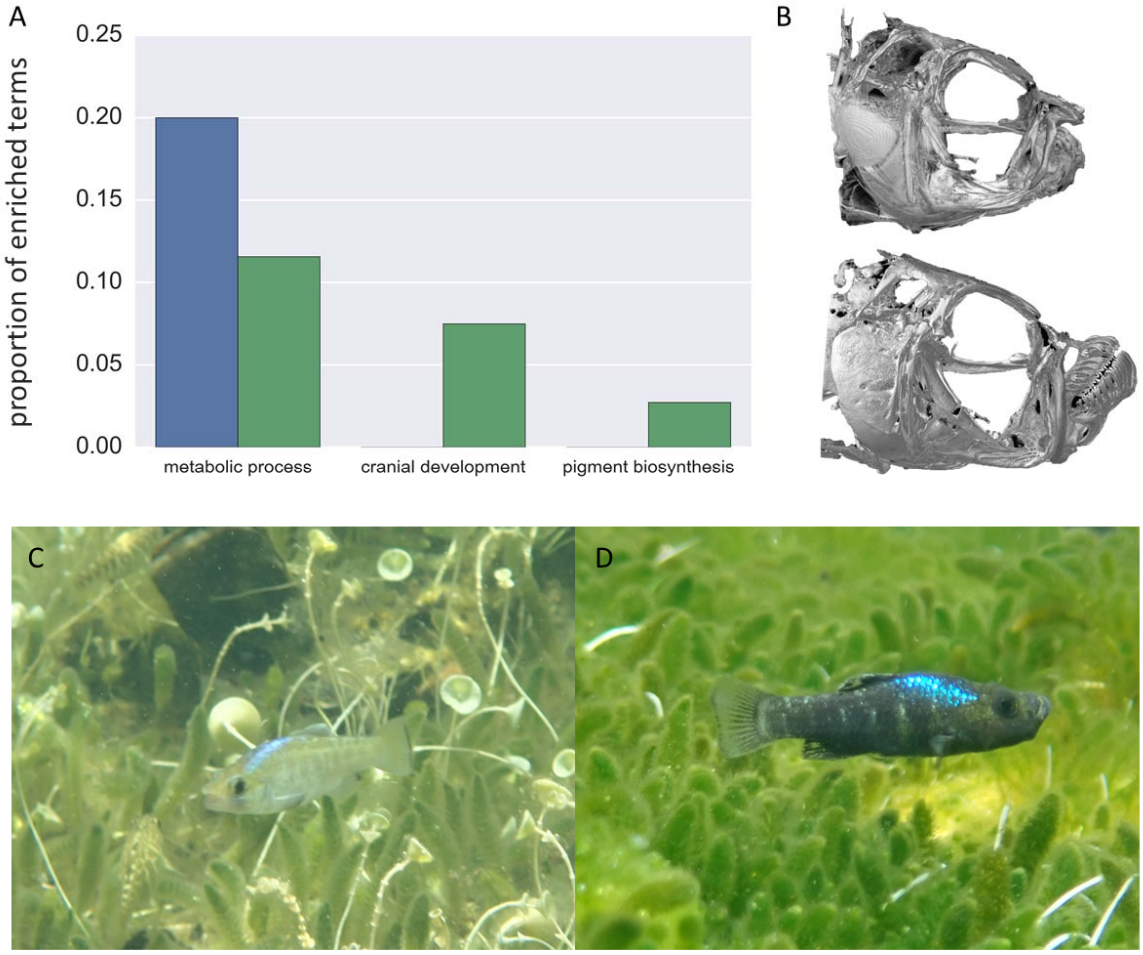


Fig. 3. Genes showing parallel expression patterns in specialists are not more pleiotropic than genes showing divergent expression. Violin plots show the distribution of pleiotropy estimates (GO biological processes (A), protein-protein interactions (B), and developmental stages expressed (C)) for genes showing parallel expression patterns (blue) and divergent expression patterns (green) in specialists relative to generalists. Red dots show the median, thick black bars show interquartile ranges and thin bars show 95% confidence intervals. Genes showing parallel expression are not significantly more or less pleiotropic than divergently expressed genes (GLM; biological processes: $P = 0.67$; PPIs: $P = 0.09$; developmental stages: $P = 0.89$).



770

771 **Fig. 4. Parallel gene expression underlies metabolic adaptations while divergent expression**
772 **underlies trophic morphology** A) Genes showing parallel expression in specialists (blue) and
773 genes showing divergent expression (green) are contrastingly enriched for terms describing
774 metabolic processes (parallel: %20; divergent: %11). Genes showing divergent expression are
775 enriched for cranial skeleton development (7% of terms) and pigment biosynthesis (3% of
776 terms). B) µCT scans show drastic craniofacial divergence between snail-eaters (top) and scale-
777 eaters (bottom). Bottom panels show male breeding coloration characteristic of light snail-eaters
778 (C) and dark scale-eaters (D).