bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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.

1 Title: Osmoregulation in freshwaters: Gene expression in the gills of a Neotropical cichlid in
2 contrasting pH and ionic environments
3
4 Authors: Stuart C. Willis[1,5]*, David E. Saenz[2], Gang Wang[3,4], Christopher M. Hollenbeck[1],
5 Luiz A. Rocha[5], David S. Portnoy[1], James J. Cai[3], & Kirk O. Winemiller[2]
6
7 1 Department of Biological Sciences, Texas A&M University-Corpus Christi, Corpus Christi,
8 TX 78412
9
10 2 Department of Wildlife and Fisheries Sciences and Program of Ecology and Evolutionary
11 Biology, Texas A&M University, College Station, TX 77843
12
13 3 Veterinary Integrative Biosciences, Texas A&M University, College Station, TX 77843
14
15 4 Department of Bioinformatics and Computational Biology, The University of Texas MD
16 Anderson Cancer Center, Houston, TX 77230
17
18 5 Department of Ichthyology, California Academy of Sciences, 55 Music Concourse, San
19 Francisco, CA 94118
20
21 *Corresponding author's email: swillis4@gmail.com
22

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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.

24

25 ## Abstract

26

27 Freshwater habitats of the Neotropics exhibit a gradient from relatively neutral, ion-rich
28 whitewater to acidic, ion-poor blackwater. Closely related species often show
29 complementary distributions among ionic habitats, suggesting that adaptation to divergent
30 osmoregulatory environments may be an important driver of Neotropical fish diversity.
31 However, little is known about the evolutionary tradeoffs involved in ionoregulation across
32 distinct freshwater environments. Here, we surveyed gill mRNA expression of *Cichla*
33 *ocellaris* var. *monoculus*, a Neotropical cichlid, to examine cellular and physiological
34 responses to experimental conditions mimicking whitewater and blackwater.
35 Gene ontology enrichment of expressed genes indicated that the gills were remodeled
36 during both forms of environmental challenge, with changes biased towards the cellular
37 membrane. We observed expression of signaling pathways from both the acute and
38 extended response phases, including evidence that growth hormone (GH) may mediate
39 osmoregulation in whitewater through paracrine expression of insulin-like growth factor I
40 (IGF-I), but not through the GH receptor, which instead showed correlated up-regulation
41 with the prolactin receptor and insulin-like growth factor II in blackwater.
42 Differential expression of genes related to paracellular tight junctions and transcellular ion
43 transport showed responses similar to euryhaline fishes in fresh versus seawater, with
44 some exceptions, suggesting that relaxed ion retention via the gills, possibly mediated by
45 the GH/IGF-I axis, is a strong candidate for evolutionary modification in whitewater and
46 blackwater endemic populations. In each osmoregulatory domain, we saw examples of

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

47 contrasting differential expression of paralogs of genes that are single copy in most
48 terrestrial vertebrates, indicating that adaption by fishes to diverse physicochemcial
49 environments has capitalized on diversification of osmoregulatory gene families.

50

51 Keywords: osmoregulation, ionoregulation, acidic, ion-poor, blackwater, whitewater,
52 Amazon, expression, RNA-seq

53

54 Running title: gene expression across an Amazon ionic gradient

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

55

56 ## Introduction

57

58 Gill surfaces of fishes must be thin in order to facilitate gas exchange, but this also
59 promotes rapid gain or loss of ions and water with the environment – a functional tradeoff
60 called the 'osmo-respiratory compromise' [1]. Management of osmotic and ionic stress is
61 therefore a major biological challenge for fishes and has been estimated to account for 2-
62 20% of resting metabolic rate [2]. A great deal has been learned about mechanisms
63 involved in osmoregulation from studies of euryhaline fishes exposed to salinity gradients
64 [3]; however, these fishes represent a small fraction of fish diversity, as the vast majority
65 are stenohaline. Among these, freshwater fishes are challenged to maintain homeostasis in
66 the face of wide variation in solute concentration, pH, and other chemical factors, but are
67 less well-studied than their euryhaline counterparts [4]. Moreover, it is clear that
68 numerous fish lineages have undergone evolutionary transitions between marine-
69 freshwater habitats, frequently resulting in modified regulation and utilization of the
70 proteins and genes involved in osmo- and ionoregulation [3]. Thus, although euryhaline
71 fishes provide useful systems for understanding osmoregulatory mechanisms broadly, they
72 may be poor models for understanding the evolutionary pathways by which stenohaline
73 lineages invade new habitats [5] or the osmoregulatory tradeoffs that produce
74 macroecological patterns in freshwater lineages [6].

75 Freshwaters of the Neotropics, which probably harbor more than a fourth of all fish
76 diversity [7], exhibit a well-known contrast between habitats that have ion-rich
77 "whitewater" or ion-poor "blackwater" [8] . Whitewater rivers flow over geologically young

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

78 soils and are circum-neutral in pH (6-7.5), exhibiting relatively high concentrations of
79 biologically important inorganic ions (major ions 1-10 mg/L; conductivity >50 μS/cm).
80 Blackwater rivers generally flow over low-gradient terrain with sandy, infertile soils
81 (podzols) and are relatively meager in free essential inorganic ions (<1 mg/L) [9] but have
82 high concentrations of dissolved organic carbon (DOC) that produces a low pH (3-5) and
83 low conductivity (<30 μS/cm) [10]. The ionic concentration of blackwater rivers is so low
84 they have been described as "slightly contaminated distilled water" [8]. Despite the
85 ionoregulatory challenges posed by blackwaters, over a thousand species are known to
86 inhabit the Negro sub-basin of the Amazon [11], the largest blackwater drainage basin in
87 the world (Supplemental Figure 1). Approximately two thirds of Amazon species seem to
88 be found in only a single water type [12,13], but many closely related species show
89 complementary distributions across water types, indicating that adaptation to water type
90 may be an important driver of fish diversity [14]. Most research on osmoregulation in
91 Neotropical fishes has focused on broadly or seasonally eurytopic fishes (those that occur
92 in both water types) or contrasts among species endemic to a single water type
93 (stenotopic). These studies have inferred that tolerance of blackwater requires greater
94 resistance to ion loss and acidification, and that native blackwater fishes respond
95 differently to acid challenges or ion supplementation than their whitewater or eurytopic
96 counterparts [15,16]. Moreover, the observation that most species are not distributed
97 across habitat types, a necessary intermediate stage for colonization of new habitats,
98 implies that there are costs to eurytopy such that ionoregulatory strategies for blackwater
99 versus whitewater are often evolutionarily mutually exclusive [e.g. ,17]. To understand
100 such tradeoffs, research is needed on stenotopic lineages with sub-populations adapted to

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

101 different habitat types, or that vary in their degree of habitat tolerance [18]. Identifying
102 these target species requires a robust understanding of species delimitation, population
103 structure, and ecology, information that currently is not available for most Neotropical
104 freshwater fishes.

105 One important exception are the South American tucunarés or peacock basses, large,
106 piscivorous cichlids in the genus *Cichla*. Following analysis of extensive morphological [19]
107 and molecular data [20,21], most *Cichla* species appear closely associated with a subset of
108 available environmental conditions [22]. Two notable exceptions are *Cichla orinocensis* and
109 *C. ocellaris* var. *monoculus*, the latter an evolutionary significant unit of the *C. ocellaris*
110 species complex. Both of these meta-populations are distributed across water types in
111 heterogeneous regions that are primarily whitewater but also punctuated by blackwater
112 rivers (*C. orinocensis* in the Orinoco Basin, *C. oc. monoculus* in the Amazonas Basin). In
113 addition, both occur within the Negro sub-basin of the Amazonas, a large area that is
114 almost exclusively blackwater and connects the Amazonas and Orinoco basins together (i.e.
115 the Casiquiare River, a Negro tributary). In both species, phylogeographic data suggest that
116 Negro populations were derived from populations in the heterogeneous regions [23,24].
117 However, the Negro populations have failed to expand into neighboring
118 heterogeneous/whitewater regions, suggesting that occupation of the blackwater habitats
119 in the Negro sub-basin has promoted adaptations that have made these Negro fishes less
120 tolerant of the ancestral, heterogeneous-whitewater environment.

121 Given the macroecological patterns exhibited by *Cichla*, these species provide an
122 ideal system for investigating adaptation of osmoregulatory mechanisms to new
123 physicochemical stress and associated evolutionary trade-offs. Currently, few specific

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

124 details are available about the molecular and regulatory pathways involved in
125 osmoregulation by *Cichla* in these contrasting freshwater habitats. Therefore, we employed
126 massively parallel sequencing of the transcriptome of the primary osmoregulating organ,
127 the gill, to identify candidate genes and genetic pathways involved in modifying
128 ionoregulation in *Cichla oc. monoculus* for blackwater and whitewater conditions. We
129 conducted an experiment where young fishes were exposed to water with chemistry
130 simulating blackwater versus whitewater conditions, followed by massively parallel
131 sequencing of mRNA from the gill filaments. This experiment was designed to identify
132 genes that exhibit strongly different patterns of expression between these conditions while
133 minimizing the changes in expression involved in a generalized stress response, for the
134 purpose of generating new hypotheses regarding habitat-specific physiological strategies.
135 Broadly, we expected that strongly and differentially expressed genes would implicate the
136 osmoregulatory mechanisms previously identified in fishes, while also indicating how
137 those may be modified to meet the disparate physicochemical conditions observed in the
138 Neotropics. Our results provide an informative context for further investigating the
139 evolutionary tradeoffs in osmoregulatory adaptation, and we discuss a number of
140 knowledge gaps that remain to be addressed.
141
142 Methods
143
144 We obtained 18 juvenile *Cichla oc. monoculus* from Colombia, where they had been
145 in pond culture in local water for one or more generations. To confirm their identity and
146 geographic origin, we sequenced the mitochondrial control region using previously

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

147 published primers and conditions [25]. These fishes exhibited the most common haplotype
148 from a *C. oc. monoculus* mtDNA clade exclusive to the western Amazon (including the
149 Colombian Amazon), a heterogeneous but primarily whitewater region [data available
150 upon request; see Supplemental Figure 1 and ref. ,20]. These fish, approximately 5 cm at
151 the time of the experiment, were kept in a 90 gallon (340 L) glass aquarium for 34 days
152 prior to the exposure, with water circulation/filtration by an Eheim brand canister filter
153 with ceramic media, daily 20% water changes, 12 hr light/dark cycle, and daily feedings of
154 locally-collected live *Gambusia affinis* prophylactically treated for bacterial and protozoan
155 parasites with kanamycin and metronidazole several weeks before feeding. Holding water
156 consisted of reverse osmosis (RO) with a salt mixture added to mimic an intermediate of
157 whitewater and blackwater habitats [9], and added to a conductivity of 30±4 µS/cm:
158 ~2.3:1:2.3:1.7 of NaCl, K_2SO_4, $MgSO_4 \bullet 7H_2O$, and $CaCl_2 \bullet 2H_2O$ by mass, respectively (from
159 Fisher Scientific). DOC (humic acid from Sigma Aldrich) was added at a rate of 10 mg/L,
160 assuming that this humic acid is 40% DOC [26] . The pH was adjusted to 6.2 using H_2SO_4
161 and maintained ±0.6 pH units with water changes during the holding period (Figure 1).
162 These conditions elicited no obvious signs of stress, and the fish increased in size during
163 the holding period. Temperature was maintained at 25±1°C throughout the experiment.
164 Temperature and conductivity were tracked with a YSI meter (YSI Inc.), and pH was
165 monitored using an accumet AP71 (Fisher Scientific).
166
167 Figure 1. Experimental conditions for holding (days) and exposure (hours) periods.
168 Holding period: dark gray; whitewater: green; blackwater: orange. Dashed lines: pH; solid
169 lines: conductivity.

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

170

171 Fish were not fed for 48 hours prior to the exposure experiment, and a 50% water
172 change was made 24 hours prior. Exposures were conducted over 12 hours during daylight
173 hours on day 35 (Figure 1). Nine fish were transferred into each of two 76 L aquariums that
174 were filled to 3 cm (~6 L) with water from the holding tank. Over the first hour, water from
175 either the simulated "black" or "white" water was transferred slowly so that the
176 experimental chamber was only filled at the end of the hour. All but 3 cm was removed
177 from each chamber, and over the second hour the chambers were again filled with
178 simulated "black" or "white" water. Simulated whitewater consisted of RO water with the
179 same salt combination but added to a conductivity of 100±6 μS/cm, pH of 6.7±0.2, and 2
180 mg/L DOC. Simulated blackwater consisted of RO water saturated with DOC (between 20
181 and 30 mg/L) and adjusted to pH 4.2±0.3 for a final conductivity of 23±1 μS/cm. No
182 additional salts were added to simulated blackwater, so the final chamber contained only
183 diluted salts from the initial 3 cm of holding water [minor additional salts are also added
184 with the Sigma Aldrich humic acid; see ,26]. Water was circulated in each chamber using a
185 submersible pump to maintain normoxia (dissolved oxygen ≥5.5 mg/L). We acknowledge
186 that this experimental design has no technical control, since both exposure conditions
187 differ from holding. However, this arrangement provides for a contrast that emphasizes
188 acclimation to novel ionic conditions in each treatment, since both experimental groups are
189 likely to experience a degree of generalized stress response. After 12 hours, fish were
190 euthanized by severing the spinal cord at the skull, and the gill arches excised and
191 preserved in RNAlater (Invitrogen). Gill filaments were separated from the arches, and
192 total RNA was extracted from the filament tissue using Trizol reagent (Invitrogen), treated

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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 with DNase (New England Biolabs), and quantified with a Nanodrop (Thermo Scientific).
194 Total RNA was pooled equally by mass in sets of three within treatment into six total
195 samples, and libraries were prepared with the Illumina Truseq Stranded kit (with A-tail
196 selection) at the Texas A&M AgriLife Genomics and Bioinformatics Service. Indexed
197 libraries were sequenced on a single lane of Illumina HiSeq 2500 with paired-end 125 bp
198 reads. Experimental procedures were approved by the Texas A&M University Institutional
199 Animal Care and Use Committee, 2013-0099.

200 A detailed description of bioinformatic and statistical procedures is available in the
201 Supplemental Information. Briefly, *de novo* transcriptome assembly was made with several
202 assemblers and a range of *kmer* values; various merged assemblies were also made. These
203 assemblies were evaluated for completeness, redundancy, fragmentation, and read
204 mapping efficacy. Quality-trimmed reads were then mapped to the optimal assembly. Even
205 given a perfect assembly, the inherent similarity between splice variants and alleles of the
206 same genes (hereafter, isoforms) means that mapping of reads to the correct isoform
207 transcript can often not be done unambiguously. In addition, the analysis of differential
208 expression (DE) of isoforms from the same gene (DIE) has less power and is more artifact-
209 prone than length-corrected summation across isoforms from the same gene [27]. Having
210 no reference genome, to mediate this we clustered contigs in the optimal assembly based
211 on reads that map in common, testing for DE at a higher hierarchical level (clusters) than
212 individual transcripts. We employed two programs that cluster transcripts based on co-
213 mapping, CORSET [28] and RAPCLUST [29], and with CORSET, we clustered transcripts with
214 two read co-mapping thresholds, ~30% and ~70%. These clusters were filtered to those
215 with a sufficient expression level (length-standardized read counts per million mapped

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

216 reads ≥ 1 in at least 3 samples), and tested for differential expression between treatments
217 using three statistical packages. We employed 27 combinations of read mapping,
218 quantification, isoform clustering, and statistical testing; to eliminate false positives, the
219 union (overlap) of transcripts from all these combinations determined the final sufficiently
220 expressed (hereafter SE) and differentially expressed (DE) sets. By taking as our SE and DE
221 sets only those transcripts that were identified by the union of all methods, our
222 bioinformatic pipeline was designed to avoid anomalies based on mapping, quantification,
223 clustering, or statistical procedures.
224 The SE set (containing the DE set) was annotated using BLAST (*blastx*), with the
225 search constrained to curated proteins (NCBI Refseq) for Nile tilapia (*Oreochromis*
226 *niloticus*; hereafter *Onil*), supplemented with other cichlids (*Neolaprologus brichardi*,
227 *Haplochromis burtoni*, *Maylandia zebra*, *Pundamilia nyererei*) and *Danio rerio* (hearafter
228 "cichlid+"). Annotated transcripts were filtered to unique genes (loci) based on NCBI gene
229 symbols, and the longest contig corresponding to each "cichlid+" gene was further BLAST
230 against human and *Danio rerio* Refseq proteins (separately). Using the human accessions of
231 the "cichlid+" genes, we obtained gene ontology (GO) annotations and tested for GO term
232 enrichment using Fisher's exact test with FDR set to 0.05 [in Blast2GOv4.1.9; ,30].
233 Enrichment for GO terms was tested for DE vs. SE, up-regulated vs. down-regulated, up-
234 regulated vs. SE, and down-regulated vs. SE comparisons. We also identified which genes in
235 the SE or DE sets were part of several well-described osmoregulatory pathways based on
236 the gene sets annotated in the following Pathcards [31]: *prolactin signaling pathway*,
237 *growth hormone receptor pathway*, *aquaporin mediated transport*, *epithelial tight junctions*
238 (Qiagen), *tight junctions* (KEGG), and *transport of glucose and other sugars, bile salts and*

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

239 *organic acids, metal ions and amine compounds*. Where additional genes were hypothesized
240 to be functionally relevant for osmoregulation (see Discussion), but were not among the
241 annotated SE set, we obtained transcripts of *Onil* from Ensembl, *tblastx* searched this
242 against the raw transcriptome assembly, and confirmed new annotations by *blastx* search
243 of the identified contig against Ensembl *Onil* proteins, only accepting reciprocal best
244 matches.
245
246 Results
247
248 It was apparent during the experimental exposure (Figure 1), following a month in
249 common, intermediate conditions, that fishes in experimental blackwater conditions
250 experienced greater stress than their counterparts in experimental whitewater conditions.
251 During the second hour, fishes in the blackwater treatment moved to the corners of the
252 aquarium and increased their ventilation rate; by the end of the experiment, a few
253 individuals (<50%) were exhibiting a loss of equilibrium. No change in behavior from the
254 holding period was apparent for the fishes in the whitewater treatment.
255 After quality trimming, the sequencing reads consisted of 190.8 million read pairs,
256 ranging from 28.2 to 33.5 million per sample (mean 31.8). From the available *de novo*
257 assemblies, the TRANSFUSE [32] merger of two BINPACKER [33] assemblies was considered
258 optimal because it exhibited comparable scores and mapping rates to other top ranked
259 assemblies but contained fewer transcripts with higher N50, implying lower redundancy
260 and fewer mis-assemblies (Table 1, Supplemental Table 1). This transcriptome assembly of
261 185,480 contigs larger than 200bp (available upon request from the corresponding author),

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

262 contained 281.9 Mbp and had a GC content of 44%. Mapping to these transcripts was
263 similar across samples (92.16 to 92.71%) and mapping programs (BOWTIE2 92.36%;
264 SALMON 92.4%). Clustering transcripts based on mapping yielded between 18,203-22,997
265 sufficiently expressed (SE) clusters, depending on clustering algorithm and co-mapping
266 threshold used. Principal components analyses of regularized log data from all SE clusters,
267 with primary axes that explained 80-81% of variation, clearly separated the treatments,
268 while the secondary axes, which explained 6-7% of variation, separated samples within
269 treatments (Figure 2). Of the SE clusters, 7,753-9,041 were significantly differentially
270 expressed (DE) between treatments with an FDR ≤ 0.05 (Supplemental Figure 2).

271

272 Table 1. Statistics and scores from Transfuse-merged *de novo* transcriptome assemblies. Top score in bold. For unmerged
273 assembly scores, see Supplemental Table 1.
274
275

***De novo* Assembler**	***kmer* Range (no. *kmers*)[a]**	**Transcripts >200bp[b]**	**Total Length (bp)[c]**	**N50 Length (bp)[d]**	**Mean Contig Length (bp)[e]**	**Mapped Reads[f]**	**Detonate Score[g]**	**TransRate Score[h]**	**BUSCO genes[i]**
oases	21-99 (7)	344,662	$2.35x10^8$	835	682	$1.79x10^8$	$-4.44x10^{10}$	0.380	2,552
SOAPd.-Tr.	21-99 (11)	312,862	$3.67x10^8$	2,486	1,171	$1.87x10^8$	$-1.55x10^{10}$	0.540	4,162
Binpacker	25, 32	185,480	$2.82x10^8$	**2,648**	**1,519**	**$1.87x10^8$**	$-1.02x10^{10}$	0.548	4,189
Trinity	25	251,811	$2.40x10^8$	1,887	951	$1.85x10^8$	$-1.17x10^{10}$	0.542	4,108
transABYSS	32-96 (5)	348,367	$3.57x10^8$	1,676	1,024	$1.86x10^8$	**$-8.66x10^9$**	0.549	4,183
ALL	*n/a*	376,024	$4.34x10^8$	2,186	1,155	**$1.87x10^8$**	$-1.87x10^{10}$	**0.562**	**4,192**

276
277 [a] size range of kmers used in assembly, and where indicated, number of kmers applied
278 [b] number of transcripts larger than 200 base pairs
279 [c] combined length of assembly
280 [d] smallest contig above which 50% of the length of the assembly is found
281 [e] mean length of contigs in assembly
282 [f] total number of mapped reads (by Salmon)
283 [g] likelihood score for each assembly by Detonate program (smaller is better)
284 [h] combined score for each assembly by TransRate program (larger is better)
285 [i] number of complete conserved genes identified by BUSCO program (more is better)

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

286 Figure 2. Principal components analysis of regularized log expression data from mapping
287 and quantification with SALMON, clustering with CORSET at ~30% read co-mapping. Green
288 circles, whitewater treatment; orange circles, blackwater treatment.
289
290 The SE clusters from each mapping-quantification-clustering-testing combination
291 contained 70,062 transcripts in common, of which 21,378 were DE in all combinations.
292 After clustering transcripts at 98% sequence similarity with CD-HIT [34], resulting in
293 60,567 representative transcripts, 81% (49,079) had significant hits to Refseq proteins for
294 Nile tilapia, other cichlids, or *Danio rerio*, which corresponded to 17,379 unique protein-
295 coding loci (hereafter, the "cichlid+" set; Supplemental Table 2). The success of blast
296 annotation was partially correlated to the length of the contig (Pearson's product-moment
297 = 0.30, $P < 2.2 \times 10^{-16}$). From these "cichlid+" annotations, the DE set was found to contain
298 6,783 protein-coding genes, but 110 genes were found to be both up and down regulated;
299 these were removed from the DE set (but see Discussion), which subsequently included
300 3,273 up- and 3,400 down-regulated genes in blackwater relative to whitewater. When
301 BLAST against human proteins, 15,618 (90%) of the representative "cichlid+" transcripts
302 had significant matches to 11,312 human genes (72% unique) (Supplemental Table 2).
303 When blast against *Danio rerio* proteins, a greater percentage of representative "cichlid+"
304 transcripts had significant hits (16,477 or 95%), which also came from a higher percentage
305 of unique loci (13,707 or 83%) (Supplemental Table 2). The human proteins in the SE and
306 DE sets are shown in Supplemental Table 3 along with the surveyed pathways with which
307 they are associated. Interestingly, 242 human genes were hits for two or more "cichlid+"
308 contigs that exhibited contrasting DE (Supplemental Table 3). Mapping of gene ontology

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

309 (GO) terms for the human protein hits was successful for 15,276 of 16,477 transcripts with
310 human annotations (98%), with a similar percentage for DE contigs (5,904 of 6,011, or
311 98%). Enriched GO terms are listed in Supplemental File 2. Search results for genes not
312 identified among the SE set (and related homologs) are described below. Following
313 common practice, below italic type refers to the gene (locus) coding for a protein (e.g.
314 *AQP3*) and normal (Roman) type refers to the protein or protein complex (e.g. NKA).
315
316 Discussion
317
318 *RNA-seq and Assignment of Homology*
319
320 RNA-seq has tremendous potential to illuminate the ecological genomics of
321 organisms beyond those with well-annotated genome sequences, which currently
322 represent a tiny fraction of species and functional groups [35,36]. Meaningful results are
323 not obtained without significant challenges, however, and our bioinformatic pipeline was
324 designed to mediate many of the problems associated with RNA-seq data in non-model
325 organisms [37]. Moreover, our interpretations of the biological patterns exhibited by the
326 present data are not based on the DE or SE status of only a few genes, and so are robust to
327 some artifacts in RNA-seq data. That being said, the interpretation of any expression
328 pattern for both model and non-model organisms, that is, expectations of the similarity of
329 function, regulation, etc., is dependent on the establishment of homology. In practice, this is
330 usually based on sequence similarity to reference genes at the nucleotide or amino acid
331 level (e.g. BLAST searches), the efficacy of which diminishes with phylogenetic distance

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

332 between query and reference [38]. In addition, the presence of homologs created by
333 duplication within the genome (paralogs) of subject or reference means these matches
334 often show one-to-many or many-to-one relationships, obscuring the transfer of identify,
335 function, etc. between genomes. These types of relationships are especially common
336 between fishes and tetrapod vertebrates because of the whole genome duplication(s) early
337 in the evolution of euteleost fishes, including cichlids [39].

338 Our determination of identity and function of our SE and DE transcripts was based
339 on homology with two model organisms: Nile tilapia (*Onil*) and humans. While *Onil* is in the
340 same family as *Cichla* and is the closest reference with a robust genome sequence and
341 annotation available, it is still likely tens of millions of years divergent [>50; ,40]. We used
342 the *Onil* annotations to assign our SE and DE transcripts to putative orthologous genes (one
343 copy per haploid genome) based on amino acid similarity, a process that is subject to
344 changes in homology (orthology/paralogy) in the intervening history between *Onil* and
345 *Cichla*. For example, 110 out of the 6,783 *Onil* genes matched by DE transcripts were
346 corresponded to *Cichla* transcripts in clusters that were DE in different directions. Without
347 a full genome for *Cichla* it cannot be clarified if these result from artifacts in,clustering or
348 DE testing, inaccurate BLAST identification, or paralogy among cichlids, and for clarity
349 these genes were resigned to the SE set. While the percentage of genes showing this
350 pattern is small (1.6%), in the case of paralogy, the result would mean that a greater
351 number of unique genes (loci) are SE or DE in the present data than have been currently
352 recognized based on assumed orthology with *Onil*.

353 Similarly, our identification of gene ontology and inclusion in pathways related to
354 osmoregulation depended on homology with human proteins. While humans are

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

355 significantly more distant from *Cichla* than is *Onil* (or *Danio*), the level of annotation of
356 human proteins is unsurpassed. As with *Onil*, 242 of 6,011 DE human genes with matches
357 to *Cichla* transcripts (4%) were found to be hits for two or more "cichlid+" genes that had
358 contrasting DE expression ("Both" in Supplemental Table 3); a greater (unquantified)
359 number corresponded to both "cichlid+" genes that were SE and others that were DE in a
360 single direction. Although a small percentage of these many-to-one hits that we inspected
361 result from erroneous BLAST identification, the vast majority appear to result from
362 changes in homology along the human and cichlid lineages. This is substantiated by our
363 parallel BLAST search of the "cichlid+" contigs against *Danio* which resulted in a greater
364 percentage of unique genes compared to humans (83% vs. 72%), reflecting greater
365 paralogy in fishes. The result is that in many cases the functional role or regulation of a
366 human gene/protein cannot be unambiguously assigned to a single *Cichla* "cichlid+" gene.
367 However, rather than a failing of our pipeline, we interpret this as evidence of evolutionary
368 innovation in fishes. Following duplication, paralogs often take on functions different from
369 the original single locus (neo- or sub-functionalization). An example of this is the prolactin
370 receptor (PRLR) which is known to be important in euryhaline fishes for remodeling gill
371 tissues in hyposmotic (freshwater) transition. Humans have a single *PRLR* gene which
372 mediates the effects of prolactin on cellular signaling, but a euryhaline cichlid model for
373 osmoregulation (*Oreochromis mossambicus*; hereafter *Omoss*) has at least two paralogs of
374 this gene expressed in osmoregulatory tissues [41]. While *PRLR* is highly expressed in
375 hyposmotic conditions, *PRLR2*, which contains fewer signaling domains and occurs in at
376 least two functionally distinct isoforms, is strongly but transiently up-regulated upon
377 *hyperosmotic* challenge, and may play a role in diminishing the effects of lingering prolactin

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

378 or could have another still-unresolved role [42]. Similarly, in Ensembl *Onil* has three
379 paralogs of *PRLR*: *PRLR*, *PRLR2*, and an unnamed third paralog (ENSONIG00000003653),
380 each of which corresponded to transcripts in our *Cichla* dataset that also showed hits to the
381 single human *PRLR* (Supplemental Table 2). Moreover, we observed that *PRLR* was
382 significantly up-regulated in blackwater relative to whitewater, *PRLR2* was significantly
383 down-regulated, and the unnamed paralog was SE, making the human *PRLR* DE in both
384 directions (Supplemental Table 3). Thus, while we urge caution in usage of our human
385 annotations, even these apparent idiosyncrasies of our pipeline provide new insight into
386 the manner in which *Cichla* responds to ionic challenge at the biochemical level. We discuss
387 additional examples in the next section.

388 Finally, for simplicity, we refer to genes with significantly higher and lower
389 expression in the blackwater treatment as "up-regulated" or "down-regulated",
390 respectively, although it should be noted that the opposite expression pattern (down or up
391 in whitewater, respectively) could result in the same pattern. Similarly, because our
392 experimental design did not include a control group (i.e. expression in holding water was
393 not assessed), genes with increases in expression in one condition relative to holding may
394 appear as "down-regulated" if expression was nonetheless higher in the opposing condition,
395 while genes that exhibited expression changes in the same direction in both experimental
396 groups relative to holding may appear in the SE set but not be identified as DE. While we
397 expect the broad patterns of differences between conditions described here to be robust,
398 our interpretations may be considered hypothetical even where consistent with previous
399 results, and we look forward to future studies that test the patterns of expression in
400 specific genes and gill cell types.

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

401

402 *Expression Responses to Contrasting Ionic Stress*

403

404 The response to osmotic stress in fishes occurs in two phases, an acute phase and
405 extended phase [43]. The acute phase (lasting minutes to hours) includes the activation
406 and insertion of ion transporters and other membrane proteins in epithelial tissues as well
407 as systemic responses like increased blood flow to osmoregulatory organs. The extended
408 phase (lasting hours to days) includes re-modeling of osmoregulatory tissues through cell
409 proliferation, differentiation, and selective apoptosis. These phases are activated by
410 osmotic sensing in relevant tissues (e.g. epithelia, pituitary) and transmitted through
411 autocrine, paracrine, endocrine, and intracellular signaling networks [42,44,45]. Our
412 experiment, at twelve hours duration, is likely to have captured the transition between
413 these two phases, although, by definition, it surveyed transcription-dependent responses,
414 which are thought to make up a larger portion of the extended response. For example,
415 among the top (ranked by difference in proportion) over-abundant specific GO terms for
416 the DE vs. SE genes were *positive* and *negative regulation of cell proliferation*, *negative*
417 *regulation of apoptotic process*, *positive regulation of cell adhesion*, *positive regulation of cell*
418 *cycle*, *epithelial cell migration*, *positive regulation of cell activation*, and *positive regulation of*
419 *apoptotic signaling pathway*, which indicates that the gill epithelia were being extensively
420 remodeled (Supplemental File 2). Not surprisingly, these GO terms also indicated that gill
421 epithelial remodeling was in response to external chemical stimulus, including *response to*
422 *drug*, *response to hypoxia*, *cellular response to acid chemical*, and *response to toxic substance*.
423 Changes in gill cells apparently were biased towards the cellular interface with the

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

424 surroundings, given that the genes involved were enriched for the *integral component of*
425 *plasma membrane*, *extracellular space*, and *cell surface* terms. Comparison of terms
426 enriched in genes up-regulated vs. down-regulated (in blackwater) showed a pattern
427 suggesting that fishes in the blackwater treatment were undergoing more transcription-
428 based remodeling (e.g. proliferation and differentiation), especially an increase in
429 mitochondria-rich cells that are known to carry out ion uptake/exchange. On the other
430 hand, fishes in the whitewater treatment showed a pattern indicating a more targeted
431 reduction in certain cell types. This was manifest by terms such as *mitochondrial*
432 *translational elongation*, *mitochondrial translational termination*, *mRNA export from nucleus*,
433 *protein targeting to mitochondrion*, *regulation of mRNA stability*, *gene silencing by RNA*,
434 *regulation of translational initiation* being enriched in blackwater up-regulated genes, while
435 the term *positive regulation of apoptotic signaling pathway* was enriched in genes with
436 higher expression in the whitewater treatment.
437
438 Acute Phase Response
439
440 Acute phase responses are understood to be mediated by a number of peptide and
441 steroid hormones, prostaglandins, leukotrienes, and catecholamines, some of which act in a
442 localized fashion (auto/paracrine), while others act systemically (endocrine) [43,44]. We
443 interrogated our DE, SE, and raw transcriptome for evidence of activity of these signaling
444 processes and found many (Table 2), but we limit our discussion to a few. For example, the
445 nonapeptides vasopressin and oxytocin ("vasotocin" and "isotocin" in fish, respectively) are
446 endocrine hormones known to regulate kidney function and have antidiuretic effects in

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

447 most vertebrates [43]. However, we found two paralogs of vasotocin receptor (*AVPR*) up-
448 regulated, while two others, as well as the isotocin rector (*OXTR*), were SE, indicating that
449 these hormones may also act to modulate salt retention in gill cells, perhaps mediated by
450 the numerous receptor paralogs. Moreover, isotocin is also known to induce cell
451 proliferation and differentiation in response to osmotic challenge [44]. Like vasopressin,
452 natriuretic peptides (NPP) are known to have antidiuretic effects, but appear to more often
453 act in paracrine fashion [43]. While *Onil* appears to have four NPP paralogs, we observed
454 only one (*NPPC*) among our *Cichla* SE set. On the other hand, we observed up-regulation of
455 *Cichla* transcripts matching both *Onil* copies of the NPP receptor 1 (*NPPR1*), while receptor
456 2 (*NPPR2*) was down-regulated (*NPPR3* was SE). Finally, endothelin, stanniocalcin, and
457 calcitonin are known to mediate responses to acid or hyposmotic stress by regulating the
458 activity or transcription of ion-transporters including the V-ATPase H^+ pump (*ATP6V*) or
459 epithelial calcium channels (*TRPV*) [44]. We observed transcripts for some of these
460 signaling molecules or precursors in the SE but not the DE data sets, while one or more of
461 the various receptor paralogs were DE, often in contrasting directions. Broadly, it appears
462 that among signaling molecules associated with the acute phase, the effects on gill tissues
463 may be mediated through coordinated expression of different receptor paralogs depending
464 on the form and severity of osmotic stress.

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

465 Table 2. Search results for selected signaling molecules, their receptors, and accessory
466 proteins known to moderate osmoregulation.
467

Hormone	Molecule	Query/Hit[a]	Expression[b]
Cortisol			
	Glucocorticoid receptor a	100534398	DE down
	Glucocorticoid receptor b	100534588	DE down
	Mineralocorticoid receptor	100712208	DE down
Prolactin			
	Prolactin, long form	E-06469	raw
	Prolactin, short form	E-06467	*NP*
	Prolactin-like	E-14183	raw
	Prolactin receptor	100534586	DE up
	Prolactin receptor 2	100534417	DE down
	Prolactin receptor-like	100691673	SE
Growth Hormone			
	Growth hormone	E-09191	*NP*
	Somatolactin	E-05357	*NP*
	GH receptor 1 (somatolactin)	100534400	SE
	GH receptor (growth hor.)	100534534	DE up
Insulin-like growth factor			
	Insulin-like growth factor I	100534565	DE down
	Insulin-like growth factor II	100534429	DE up
	IGF 1 receptor	100711993	SE
	IGF 2 receptor	100703210	SE
	IGF-like family receptor 1	100700407	DE down
Vasopressin			
	Vasopressin	E-15218	raw
	Vasopressin receptor 1a	100711102	DE up
	Vasopressin receptor 1b	100695840	SE
	Vasopressin receptor 2	100707194	DE up
	Vasopressin receptor 2-like	E-16572	raw
	Vasopressin receptor 2-like	E-20274	raw
	Vasopressin receptor 2-like	E-19049	*NP*
Isotocin			
	Oxytocin	E-15235	*NP*
	Oxytocin receptor	E-18982	raw
	Oxytocin receptor-like	100702815	SE
Angiotensin II			
	Angiotensinogen	E-00537	*NP*
	Renin	E-17285	*NP*
	Angiotensin-converting enz.	100712518	DE down
	Angiotensin II receptor 1	100702961	DE up
	Angiotensin II receptor 2	100702180	DE up

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

469
470 Table 2. *continued*
471

Hormone	Molecule	Query/Hit[a]	Expression[b]
Urotensin II			
	Urotensin 2	E-16900	*NP*
	Urotensin 2 receptor	E-19667	raw
	Urotensin 2 receptor-like	E-16919	raw
	Urotensin 2 receptor-like	E-06491	*NP*
	Urotensin 2 receptor-like	E-21202	*NP*
Natriuretic Peptides			
	Atrial natriuretic peptide	E-17953	*NP*
	C-type natriuretic peptide	100712558	SE
	C-type npp-like 1	E-12915	raw
	C-type npp-like 2	E-07496	*NP*
	NPP receptor 1a	100695733	DE up
	NPP receptor 1b	100692491	DE up
	NPP receptor 2	100699899	DE down
	NPP receptor 3	100704651	SE
	NPP receptor-like	E-13704	*NP*
	Corin	100703941	DE up
Vasoactive Intestinal Peptide			
	Vasoactive intestinal peptide	101475318	DE down
	VIP receptor 1a	100697561	SE
	VIP receptor 1b	E-05778	*NP*
	VIP receptor 2	100707237	DE down
Insulin			
	Insulin	E-00343	raw
	Insulin b	E-17585	*NP*
	Insulin receptor a	100697854	DE up
	Insulin receptor b	100696191	SE
	Insulin receptor substrate 1	100707127	SE
	Insulin receptor substrate 2a	100704325	DE up
	Insulin receptor substrate 2b	100699855	DE up
	Insulin receptor substrate 4	E-11878	raw
Glucagon			
	Glucagon	E-18307	raw
	Glucagon b	E-08919	*NP*
	Glucagon receptor	100700989	DE down
	Glucagon receptor b	100692999	SE
	Glucagon-like 2 receptor	E-19565	*NP*

472

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

474
475 Table 2. *continued*
476

Hormone	Molecule	Query/Hit[a]	Expression[b]
Parathyroid Hormone			
	Parathyroid hormone	E-11680	*NP*
	PTH receptor 1a	E-09467	raw
	PTH receptor 1b	E-11572	*NP*
	PTH receptor 2	E-04451	*NP*
Vitamin D			
	Vitamin D3 receptor A	100696631	DE up
	Vitamin D3 receptor B	100695813	SE
Stanniocalcin			
	Stanniocalcin 1	100692284	SE
	Stanniocalcin 1-like	E-12623	*NP*
	Stanniocalcin 2a	100709602	SE
	Stanniocalcin 2b	106098364	DE up
Calcitonin			
	Calcitonin a	E-06147	raw
	Calcitonin b	E-15534	raw
	Calcitonin receptor	100711460	DE up
	Calcitonin receptor-like 1a	100707330	DE down
	Calcitonin receptor-like 1b	100696897	SE
	Calcitonin receptor-like a	E-04617	*NP*
	Calcitonin receptor component	100691006	SE
Endothelin			
	Endothelin 1	100689848	SE
	Endothelin 2	102080384	SE
	Endothelin 3	E-02242	raw
	Endothelin receptor aa	100698766	DE down
	Endothelin receptor ab	100706592	DE up
	Endothelin receptor ba	100704989	DE up
	Endothelin receptor bb	E-16062	*NP*
	Endothelin receptor-like	100690331	SE
	Edn-converting enzyme 1	100690201	SE
	Edn-converting enzyme 2a	100698809	SE
	Edn-converting enzyme 2b	100700204	DE down
	Edn convertin enzyme-like 1	E-10817	raw

477
478 [a] NCBI GeneID or Ensembl ID (the latter abbreviated from ENSONIG000000XXXXX).
479 [b] Present in the differentially expressed (DE) set (up, up-regulated in blackwater; down;
480 down-regulated), sufficiently expressed (SE) set, raw unannotated transcript assembly
481 (raw), or not present/identifiable by Blast (NP).

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

482 The effects of the acute phase signaling molecules are diverse and not fully resolved.
483 Some act directly to activate target proteins or enhance transcription, while others activate
484 intracellular signaling cascades with more widespread effects [42,45]. We observed
485 evidence of several of these signaling cascades (Supplemental Tables 2 and 3). For example,
486 calcium/calmodulin and adenylate cyclase/cAMP represent signaling networks that result
487 in the activation and insertion of many membrane proteins involved in the osmoregulatory
488 response, including ion transporters and tight junction proteins, and many of these
489 signaling participants were SE or DE in the dataset. Moreover, one of the top GO terms
490 enriched in the up-regulated vs. down-regulated comparison was *response to cAMP*
491 (Supplemental File 1). A number of signaling cascades are also known to affect
492 osmoregulation through activation and transcriptional regulation leading to cell
493 proliferation, differentiation, or turnover (e.g. MAPK or PI3K-AKT), and many of the
494 molecules in these networks were in the SE or DE sets. Moreover, among the top GO terms
495 enriched in the DE vs. SE comparison were *ERK1 and ERK2 cascade* and *positive regulation*
496 *of JNK cascade* (both MAPK signaling), while in the up-regulated vs. down-regulated
497 comparison, up-regulated genes were enriched for *negative regulation of MAPK cascade*,
498 whereas down-regulated genes were enriched for *Wnt signaling pathway-calcium*
499 *modulating pathway*, *activation of JUN kinase activity*, and *non-canonical Wnt signaling*
500 *pathway via JNK cascade*. The differential expression and GO term enrichment suggest that
501 different (though not exclusive) signaling cascades were being utilized to coordinate
502 responses to these divergent hyposmotic challenges.
503
504 Extended Phase Response

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

505

506 The canonical extended response to osmotic challenge in euryhaline fishes is
507 understood to be mediated through the endocrine peptide hormones prolactin, growth
508 hormone, and insulin-like growth factor I, and the glucocorticoid hormone cortisol [46].
509 The effects of cortisol appear to be complex and context dependent, and increased cortisol
510 levels are associated with both hyper- and hyposmotic challenge. In *Omoss* and *Onil* cortisol
511 has been observed to both enhance, suppress, or act independently of the effects of
512 prolactin or growth hormone on osmoregulatory proteins [47–49]. Consistent with this,
513 cortisol appears to have been an important factor in both of our treatments, because the GO
514 term *response to corticosteroid* appeared in the DE vs. SE comparison that considered all DE
515 genes (Supplemental File 2). Interestingly, both of two paralogs of the glucocorticoid
516 receptor (*NR3C1*), as well as the mineralocorticoid receptor (*NR3C2*), all of which may bind
517 cortisol in fish [but see ,50], were down-regulated in the blackwater treatment
518 (Supplemental Tables 2 and 3). While there is general consensus that prolactin acts to
519 transform the gills of euryhaline fishes for hyposmotic conditions, it is unclear how the
520 effects of prolactin may be mediated for different types of hyposmotic challenges. In the up-
521 regulated vs. down-regulated comparison that contrasted the two treatments, the term
522 *cellular response to peptide hormone stimulus* was enriched in the up-regulated set. As
523 described above, we observed two prolactin receptors (*PRLR*, *PRLR2*), previously
524 implicated in hyper/hyposmotic transition in cichlids [41], to exhibit similar contrasting
525 expression in these two hyposmotic challenges (Figure 3). In contrast, growth hormone
526 (GH), another endocrine peptide hormone, has been suggested to promote acclimation to
527 hypertonic environments, its effects mediated through the auto/paracrine peptide

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

528 hormone insulin-like growth factor I (IGF-I). However, although GH protein and mRNA
529 levels increase in *Omoss* following fresh to seawater transition, the molecule itself has not
530 consistently been shown to have significant direct impacts on osmoregulatory effectors
531 [51,52]. Moreover, at least in *Omoss*, transcription of the GH receptor (*GHR*) is usually
532 higher following adaptation to *freshwater*, and consistent with this, we observed that *GHR*
533 expression was higher in the blackwater treatment. By contrast, transcription of IGF-I
534 (*IGF1*), usually acting in auto/paracrine fashion in *Omoss* in response to GH [53], was up-
535 regulated in whitewater, whereas transcripts of IGF-II (*IGF2*), which also promotes cell
536 proliferation/differentiation, were up-regulated in blackwater. No significant change was
537 observed in IGF receptor expression (*IGF1R*, *IGF2R*), but we did observe contrasting DE in
538 several IGF binding proteins (*IGFBP*) that mediate interactions of the IGFs and their shared
539 receptor (*IGF1R*) [54]. Notably, we also observed the GO term *positive regulation of peptide*
540 *hormone secretion* enriched in the DE vs SE comparison, possibly reflecting auto/paracrine
541 signaling (e.g. IGF) in both treatments. Thus, the actions of growth hormone, if coordinated
542 with cortisol and paracrine action of IGF-I in the gills, may be de-coupled from GHR, which
543 instead may have been co-opted by signaling involving prolactin and IGF-II. Some studies
544 have speculated that, in *Omoss*, GHR may relay the somatotropic signals of prolactin in
545 hyposmotic conditions [55,56]. Conversely, like the short form of PRLR2, some isoforms of
546 GHR may act to mitigate the effects of growth hormone in these conditions [42]. The data
547 from *Cichla*, with contrasting expression of *PRLR*, *GHR*, and *IGF2* vs. *PRLR2* and *IGF1*, are
548 consistent with either of these scenarios, and we urge caution in interpreting these results
549 until more detail expression studies with additional controls are performed. In any case,
550 the effects of the PRLR and GHR are understood to be transmitted through the JAK-STAT

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

551 signaling pathway, as well as through the PI3K-AKT and MAPK pathways [57], and many
552 constituents of the JAK-STAT pathway were found in the SE set, and some, like *JAK1* and
553 *STAT3*, were DE (Supplemental Tables 2 and 3).
554
555 Figure 3. Gene expression results for select proteins, in cellular context. Color reflects
556 expression of one or more protein isoforms (paralogs): orange, up-regulated in blackwater;
557 green, down-regulated in blackwater; gray, no change but sufficiently expressed.
558 Representation here is not intended to imply co-localization in the same cell type.
559 Membrane localization and orientation reflects *Oreochromis mossambicus* (e.g. Wilson et al.
560 2000, Furukawa et al. 2014), but is speculative for several proteins. For protein
561 descriptions, please see text.
562
563 The coordinated actions of prolactin, GH/IGF, and cortisol mediate paracellular
564 permeability by regulating tight junction proteins [49,58]. In freshwater, these junctions
565 allow little movement of ions or water ("tight"), whereas in saline conditions paracellular
566 junctions promote the selective escape of ions down transepithelial gradients ["leaky"; ,58].
567 This is understood to be accomplished through increased/decreased expression of
568 MARVEL domain-containing proteins (e.g. occludin) and selective expression of different
569 claudin proteins; both of these groups are transmembrane proteins which obstruct and
570 polarize the paracellular pathways as well as coordinate and communicate with gap and
571 adherens junctions, the cytoskeleton, membrane skeleton, and signaling networks of the
572 cell [59]. We observed contrasting DE in 25 of the 37 "cichlid+" genes identified as claudins:
573 14 up and 11 down (Figure 3; Supplemental Tables 2 and 3). Poor characterization of these

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

574 proteins in fishes, and unclear homology with human proteins (*Onil* have ~58 *CLDN* genes;
575 humans ~28) precludes us from predicting their effects on permeability [but see ,49]. By
576 contrast, four of seven MARVEL domain-containing paralogs (*OCLN*, *MARVELD2*,
577 *MARVELD3*) were up-regulated, and the others were SE, indicating that tight junctions
578 were becoming 'tighter' in the blackwater treatment, and more 'leaky' in whitewater. It is
579 possible that fishes adapted to whitewater may exhibit relaxed or altered mechanisms for
580 retention of certain ions via paracellular pathways. Those losses would ultimately need to
581 be coordinated with the transcellular transport processes discussed below.

582 In addition to the paracellular transmembrane proteins, a number of accessory and
583 intermediary proteins involved in tight junction formation were also observed in the DE or
584 SE sets. For example, the zona occludens proteins (TJP1-3), cingulin (CGN), and
585 paracingulin (CNGL1) are known to mediate the interaction between the transmembrane
586 proteins and the cytoskeleton [59], and several of these (*TJP3*, *CGN*, and *CNGL1*) were up-
587 regulated (Figure 3; Supplemental Tables 2 and 3). Intriguingly, whereas humans have only
588 a single copy of the small GTPases *CDC42* and *RAC1*, which respond to peptide hormone
589 signaling and mediate tight junction assembly and cell polarity, fishes have multiple copies
590 of each, and we observed contrasting DE among several of these. Similarly, the two
591 paralogs of PALS2 (*MPP6*), a guanylate kinase that helps mediate the interaction of tight
592 junction accessory proteins with occludin, were DE in different directions. Collectively,
593 these observations indicate that gill remodeling in response to peptide and corticosteroid
594 signaling involved proteins known to reduce paracellular permeability in the fish exposed
595 to blackwater, and that evolution of this capability may have involved functional
596 divergence of these duplicated genes.

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

597 As with tight junctions, alteration of the gill for salt uptake or secretion has been
598 shown to be regulated by prolactin, GH/IGF, and cortisol, including in *Omoss* and *Onil*
599 [e.g. ,52]. The uptake of ions from freshwater environments and creation of transepithelial
600 gradients is largely accomplished by transcellular transport via ion channels and pumps in
601 coordination with passive or facilitated diffusion of metabolic solutes [3]. One of the most
602 well documented of these is the "Na^+/NH_4^+" exchange complex [60]. In this model,
603 gradients to effectively exchange H^+ for Na^+ at the apical membrane are facilitated by the
604 efflux of NH_3 via Rhesus proteins (Figure 3). Na^+ uptake and H^+ excretion is hypothesized to
605 be via Na^+/H^+ exchanger (NHE) or by the H^+ pump (HA) coupled to epithelial Na^+ channels
606 (ENaC). As it moves across the apical gill membrane via Rhesus C (RhC), NH_3 immediately
607 ionizes to NH_4^+, reducing the external [H^+] and partial pressure of NH_3 (acid trapping). This
608 process is also dependent on carbonic anhydrase (CA), which increases intracellular [H^+]
609 relative to the surface boundary by the hydrolysis of passively-diffusing CO_2 to H^+ and
610 HCO_3^-, and basolateral excretion of the conjugate HCO_3^- by the Na^+-HCO_3^- co-transporter
611 (NBC), thereby also decreasing intracellular [Na^+] relative to the boundary layer. Less clear
612 is how NH_3 enters the cell basolaterally, via another Rhesus protein (e.g. RhB) as NH_3, or by
613 active pumping through substitution of NH_4^+ for K^+ in the Na^+/K^+ ATPase (NKA). Versions
614 of this model are supported by experimental evidence, yet many questions remain. For
615 example, in *Omoss* Na^+ and Cl^- uptake occurs through mitochondria-rich cells (MRC) that
616 express partially exclusive sets of ion transporters ("NHE" or "NCC" MRC), but many of
617 these transporters move overlapping sets of ions (Na^+, K^+, Ca^{2+}, Cl^-, HCO_3^-) to energize
618 critical gradients, and their division and coordination have not been clarified [61].
619 Moreover, it remains to be resolved if this Na^+/NH_4^+ exchange model is effective in acidic,

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

620 ion-poor environments, where the external [H^+] and [Na^+] are extremely unfavorable for
621 exchange [62]. Some evidence from fish native to blackwater habitats suggests that these
622 exchanges may be de-coupled, whereby Na^+ uptake depends on NH_3 excretion, but NH_3
623 excretion is independent of Na^+ uptake [15]. For example, Hirata et al. [18] found that dace
624 native to blackwater increased somatic NH_3 production (glutamate catalysis) to remediate
625 Na^+ losses, as well as NHE, CA, NBC, and aquaporin (AQP3) transcription (but not HA),
626 whereas non-adapted dace were not able to maintain plasma [Na^+] or pH. Similarly, when
627 exposed to acidic, ion-poor water, *Omoss* increased both NHE and NCC (Na^+-Cl^- co-
628 transporter) transcription, but not HA [63], suggesting that Na^+ and Cl^- transport may both
629 be coordinated with NH_3 excretion, but perhaps not via the HA/ENaC arrangement.
630 By contrast, *Cichla* in our acidic, ion-poor treatment exhibited up-regulation of both
631 NHE2 (*SLC9A2*) and several subunits of HA (*ATP6V*), along with one of two paralogs of NBC
632 (*SLC4A4*), five of seven paralogs or subunits of NKA (*ATP1*), one of two paralogs of RhC
633 (*Rhcg*), and *AQP3* (Figure 3; Supplemental Tables 2 and 3). We saw no significant changes
634 in expression (SE) in either of two paralogs of NCC (*SLC12A3*), NHE3 (*SLC9A3*), the other
635 Rh paralogs (*Rhbg*, two paralogs of *Rhag*), or a potential ENaC (*ASIC2*). We did, however,
636 note increased expression of aquaporin 8, *AQP8*, which may transport NH_3 as well as water
637 [64]. Interestingly, while the hydrolysis of CO_2 invoked by the Na^+/NH_4^+ exchange model is
638 generally by cytosolic (e.g. *CA2*) or membrane-associated (e.g. *CA4*) anhydrases, expression
639 of these did not change (SE). However, another membrane form (*CA12*) and a secreted
640 form (*CA6*) were down-regulated, while the mitochondrial form (*CA5b*), which was shown
641 to be critical for osmotic homeostasis in zebrafish embryos [65], was up-regulated.
642 Although these isoforms have not been previously identified as being specifically localized

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

643 or active in gill MRC, this observation is consistent with the hypothesis that MRC of fishes
644 in the blackwater treatment were acclimating to increase intracellular [H^+], or that in
645 whitewater MRC were working to increase surface boundary [H^+], to maintain gradients
646 appropriate for acid-trapping facilitation of metabolite excretion and effective ion exchange.

647 Although this model focuses on Na^+ uptake, this must also be coordinated with the
648 transport of other ions to maintain effective electrochemical gradients. For example, it was
649 recently discovered that *Omoss* excretes K^+ in both fresh and seawater, facilitated by apical
650 potassium channels (Kir) [61], and we observed that several paralogs of Kir (*KCNJ*) were
651 down-regulated, consistent with increased ion conservation in blackwater (Figure 3;
652 Supplemental Tables 2 and 3). These authors also speculated that K^+ may be excreted
653 through K^+-Cl^- co-transporters (KCC1, KCC4), but we noted these paralogs were up-
654 regulated (*SLC12A4, SLC12A7*), suggesting they may actually play a role in K^+/Cl^- uptake or
655 recycling. We also noted contrasting DE of several Na^+/K^+-Ca^{2+} exchangers, NCKX2
656 (*SLC24A2*), NCKX3 (*SLC24A3*), and NCKX6 (*SLC8B1*), whose roles mediating osmotic stress
657 in fishes have not been characterized, but which generally serve to convey Na^+ into the cell
658 in exchange for Ca^{2+} and K^+ [see also ,66]. Similarly, we observed contrasting DE of two
659 epithelial Ca^{2+} channel genes (ECaC), *TRPV2* and *TRPV6,* which convey Ca^{2+} down
660 electrochemical gradients. However, we did not observe changes in expression of either the
661 basolateral Na^+/Ca^{2+} exchanger, NCX (*SLC8A1*), or Ca^{2+}-ATPase, PMCA (*ATP2B*), that were
662 hypothesized to facilitate Ca^{2+} uptake in zebrafish and tilapia in coordination with ECaC
663 [3,67]. Consistent with their role in Na^+, K^+, and Cl^- secretion in *Omoss* in seawater [68], we
664 observed expression of one of two paralogs of the Na^+-K^+-Cl^- co-transporter, NKCC
665 (*SLC12A2*), down-regulated in blackwater, whereas the apical chloride channel CFTR was

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

666 only present in the raw transcriptome. Surprisingly, there was also down-regulation of the
667 basolateral Cl^- channel (CLCN1) and the Cl^-/HCO_3^- exchanger (AE1, *SLC4A1*), both of which
668 have been suggested to facilitate Cl^- uptake in freshwater *Omoss* [69,70,71; but see 59].
669 Finally, we identified many other DE or SE proteins known to be involved in transport of
670 other important inorganic ions (Mg^{2+}, Zn^{2+}, Cu^{2+}, $Fe^{2+/3+}$) as well as amino acids and other
671 organic ions (Supplemental Tables 2 and 3). Overall, expression of these ion transporters is
672 largely consistent with models describing hyposmotic stress, although the specific protein
673 isoforms and genes involved appear to vary somewhat between *Cichla*, euryhaline cichlids,
674 and other fishes, and indicate a significant amount of functional divergence among paralogs
675 depending on the nature and degree of the osmotic stress.
676
677 *Future Directions*
678
679 We observed extensive changes in expression of genes involved in gill remodeling
680 processes in response to ionic stress under conditions that mimicked natural
681 environmental gradients. Our experiment contrasted conditions in whitewater and
682 blackwater habitats, which sometimes occur in close proximity in South America and
683 between which fish of some species disperse, though the majority of Amazonian fishes
684 appear to be found at either end of this gradient [13]. Responding to these ionic challenges
685 apparently involved many of the same mechanisms and regulatory pathways as fishes in
686 hyper/hyposmotic transitions, although with some notable deviations. An RNA-seq study
687 of an Amazonian characin, *Triportheus albus*, from rivers of different water types observed
688 differential expression in a number of these same transporters [72]. While this study is not

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

689 directly comparable because experimental subjects were surveyed *in situ*, and it is unclear
690 whether observed expression differences result from plasticity or population-specific,
691 constitutive expression [i.e. ,14], these similarities suggest that acclimation to
692 physicochemical challenges, and potentially also adaptation, utilize many of the same
693 mechanisms to cope with ionic gradients of significant magnitude largely regardless of the
694 relative tonicity of environment to body fluids [73]. In addition, our observations of several
695 examples of contrasting expression of fish-specific paralogs of genes with known
696 importance in human osmoregulation indicates the importance of functional diversification
697 of these gene families in fishes for transitions among habitats with distinct
698 physicochemistry [39]. Nevertheless, it remains unclear how many of these patterns are
699 generalized among freshwater or Amazonian fishes.
700 While we observed expression of many transcripts putatively involved in osmotic
701 stress, the system requires further study. For example, studies to date have revealed that
702 the degree of correspondence between mRNA transcription and actual changes in protein
703 abundance and activity can vary widely [74]. In addition, the interactive partners and
704 cellular localization of proteins constrains their immediate function, and many of the
705 proteins we identified are unresolved with respect to localization in the tissues or
706 physicochemical conditions surveyed, though their correlated expression is indicative of
707 activity in related pathways. Furthermore, it is clear that the fish gill is not a homogenous
708 population of cells, being comprised of pavement, mucosal, neuroepithelial, and multiple
709 types of MRCs in spatially varying proportion across the lamellae, filament, and arch
710 epithelia [75]. However, our tissue samples consisted of homogenized filaments including
711 blood cells, and we can directly corroborate the derivation of observed transcripts from

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

712 osmoregulatory cells. Indeed, even the morphology of MRCs is not homogenous and plays
713 an important functional role, with follicular apical crypts or pits creating ionic
714 microenvironments at the apical membrane. This MRC morphology has been observed in
715 cichlids inhabiting acidic, ion-poor, and hypoxic waters as well as hyperosmotic
716 environments [63,76,77]. Our findings point to extensive molecular interactions and co-
717 regulation in coping with ionic stress, and resolution of other details of these processes will
718 be important for achieving a robust understanding of osmoregulatory physiology and
719 adaptation.

720 Several avenues also remain to be explored in our own data. We chose not to
721 explore differential isoform expression (DIE) because we lack splicing models for most
722 genes; nonetheless, there are well known isoforms for some of the key osmoregulatory
723 participants [e.g. the short form of PRLR2; ,41]. Many of the genes in the DE or SE sets likely
724 exhibited functionally distinct isoforms whose expression may have varied between
725 treatments, but our bioinformatic pipeline subordinated these patterns (while accounting
726 for transcript length) to gene level expression. We also acknowledge that 19% of SE
727 transcripts were not identifiable with the selected reference sets. A significant proportion
728 could represent non-coding RNA that would not be identified in our search against protein
729 databases, and indeed a casual search with some of these unidentified transcripts showed
730 significant matches to *Onil* non-coding RNA genes (results not shown). The functional role
731 of non-coding RNA is an area of active research and lies beyond the scope of the present
732 study. Finally, additional genes undoubtedly exhibited DE in our experiment but were not
733 identified due to low expression or other technical artifacts. We therefore consider our
734 findings conservative with regard to identification of all DE genes.

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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 Our findings provide an initial step for research exploring evolutionary tradeoffs in
736 adaptation to novel osmotic environments and stimulate many new questions. The fish
737 utilized here are native to a region dominated by whitewater habitats, and they were tested
738 for responses to whitewater and blackwater conditions after being acclimated to
739 conditions that were intermediate. Thus, our procedure captured only one of several
740 dimensions important in adaptation to novel physicochemical environments, including
741 population-level variation, developmental plasticity, and epigenetic effects [e.g. ,5,6].
742 Physiological observation of Negro River fishes has indicated that some species possess
743 higher affinity Na^+ uptake mechanisms and decoupling of NH_3 excretion from Na^+
744 import/H^+ export [e.g. ,15]. In addition, the DOC found in Negro habitats may have unique
745 chelating properties utilized by native fishes to minimize ion loss [78,79], and attempts to
746 recreate this with other DOC sources (including Sigma humic acid) have produced mixed
747 results [26,80,81]. The latter observation may partially explain why several Negro fishes
748 tested in native water have shown lower dependency on external Ca^{2+} to charge
749 paracellular junctions [82,83].

750 As a result, while acclimation to the hyposmotic gradients created here mimicked
751 patterns seen in euryhaline fishes transitioning between seawater and freshwater, it
752 remains to be seen if this would be true of blackwater-native *Cichla*. Based on results from
753 our experiment, we hypothesize that *Cichla* endemic to the Negro River sub-basin appear
754 to have been unable to colonize whitewater regions, in part, due to an inability to efficiently
755 regulate NH_3 excretion via boundary-layer acidification or to modulate the retention of
756 some ions through paracellular or transcellular pathways [36,84]. Possible reasons for this
757 could be isoform expression canalization, amino-acid substitutions in effector proteins (ion

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

758 transporters or tight junction regulators), insensitivity of osmoregulatory complexes to the
759 GH/IGF-I regulatory axis, or insensitivity of the axis itself to the ion concentrations
760 common in whitewater habitats [e.g. ,85,86], any of which could reflect ionoregulatory
761 adaptation to blackwater that becomes maladaptive in whitewater. However, *Cichla oc.*
762 *monoculus* is distributed across water types in the Amazon, and molecular data suggest
763 that these populations are connected by low to moderate gene flow [20,23]. It will be
764 important to assess if gene flow between proximal whitewater and blackwater habitats in
765 the central Amazon constrains ionoregulatory adaptation. If there is an antagonism
766 between adaptation and gene flow among sub-populations in different water types, this
767 could explain why fishes in homogenous regions, like the almost exclusively blackwater
768 Negro sub-basin, would be less tolerant of whitewater than their counterparts from
769 heterogeneous regions: reduced gene flow-selection antagonism facilitates fixation of
770 blackwater-adaptive alleles in the Negro [e.g. ,17]. However, dispersal across habitat types
771 in heterogeneous regions would depend first on the ability of individual fish to tolerate a
772 range of physicochemical conditions, and though little data are available on the breath of
773 physicochemical tolerance in *Cichla*, observations from the current experiment, in which
774 blackwater elicited significant stress in fish from the heterogeneous, western Amazon,
775 suggest that tolerance is not broad. However, this also highlights the unknown influence of
776 developmental plasticity and epigenetics on osmoregulation. For example, Moorman et al.
777 [87] observed that *Omoss* raised in tanks mimicking the temporal variation in ionic
778 concentration of tidal habitats successfully transitioned from fresh to seawater, while those
779 raised in freshwater-only environments could not. *Cichla* usually exhibit site fidelity, but
780 occasionally disperse over several kilometers [88], and consequently fish that encounter

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

781 environmental variation during early developmental stages may be more capable of
782 efficient ionoregulation across habitats as adults. Considerable additional data will be
783 needed to address these questions. The transcriptomic findings presented here provide a
784 foundation for research addressing both proximal and ultimate mechanisms influencing
785 biogeographic and diversification patterns in *Cichla* and other freshwater fishes.
786
787 Acknowledgements
788
789 The authors appreciate assistance by personnel managing the Texas A&M University-
790 Corpus Christi High Performance Computing Cluster where several bioinformatics
791 procedures were run. We are grateful for helpful discussions with the members of the
792 Marine Genomics Lab (TAMU-CC) and Institute for Biodiversity Science and Sustainability
793 (CalAcademy). A.L. Val and C.M. Wood graciously provided comments on a previous version
794 of this manuscript. Figure 3 was created with the skillful assistance of P. Dimens. This
795 article is publication number XX of the Marine Genomics Laboratory at Texas A&M
796 University-Corpus Christi.
797
798 Funding
799
800 This project and personnel were supported by the Estate of George and Caroline
801 Kelso via the International Sportfish Fund (KW), the TAMU diversity fellowship program
802 (DES), and the College of Science and Engineering at TAMU-CC (SCW, CMH, DSP). These

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

803 funding bodies had no role in the study design, implementation, or interpretation or
804 reporting of the results.
805
806 Authors' Contributions
807
808 SCW and KOW conceived of the study, and designed the experiment with JJC. SCW and DS
809 conducted the experiment, and GW performed laboratory procedures leading to library
810 preparation by TAMU Agrilife Genomics. SCW performed bioinformatics processing and
811 statistical analyses, with assistance from CMH. All authors contributed to interpretation of
812 the results and editing the manuscript.
813
814 Availability of Data and Materials
815
816 Raw sequence data has been deposited with the NCBI Short Read Archive as XXXXX. The
817 transcriptome assembly used for gene expression quantification is available upon request
818 from the corresponding author.
819
820 Conflicts of Interest
821
822 *The authors declare that the research was conducted in the absence of any commercial or*
823 *financial relationships that could be construed as a potential conflict of interest.*

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

825

826 References

827

828 1. Evans DH. The Multifunctional Fish Gill: Dominant Site of Gas Exchange,
829 Osmoregulation, Acid-Base Regulation, and Excretion of Nitrogenous Waste. Physiol
830 Rev. 2005;85: 97–177. doi:10.1152/physrev.00050.2003
831 2. Febry R, Lutz P. Energy partitioning in fish: the activity-related cost of
832 osmoregulation in a euryhaline cichlid. J Exp Biol. 1987;128: 63–85.
833 3. Hwang P-P, Lee T-H, Lin L-Y. Ion regulation in fish gills: recent progress in thecellular
834 and molecular mechanisms. Am J Physiol - Regul Integr Comp Physiol. 2011;301:
835 R28–R47.
836 4. Brauner CJ, Gonzalez RJ, Wilson JM. Extreme Environments: Hypersaline, Alkaline,
837 and Ion-Poor Waters. In: McCormick SD, Farrell AP, Brauner CJ, editors. Fish
838 Physiology. Elsevier, New York.; 2012. pp. 435–476. doi:10.1016/B978-0-12-
839 396951-4.00009-8
840 5. Feder ME. Evolvability of physiological and biochemical traits: evolutionary
841 mechanisms including and beyond single-nucleotide mutation. J Exp Biol. 2007;210:
842 1653–1660. doi:10.1242/jeb.02725
843 6. Chown SL, Gaston KJ. Macrophysiology - progress and prospects. Funct Ecol.
844 2016;30: 330–344. doi:10.1111/1365-2435.12510
845 7. Reis RE, Albert JS, Di Dario F, Mincarone MM, Petry P, Rocha LA. Fish biodiversity and
846 conservation in South America. J Fish Biol. 2016;89: 12–47. doi:10.1111/jfb.13016
847 8. Sioli H. The Amazon. Limnology and landscape ecology of a mighty tropical river and

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

848 its basin [Internet]. Dordrecht. 1984. doi:10.1007/978-94-009-6542-3
849 9. Furch K. Water chemistry of the Amazon basin: The distribution of chemical
850 elements among freshwaters. In: Sioli H, editor. The Amazon Limnology and
851 Landscape Ecology of a Mighty Tropical River and Its Basin. W. Junk, Dordrecht;
852 1984. pp. 167–199.
853 10. Köhler S, Buffam I, Jonsson A, Bishop K. Photochemical and microbial processing of
854 stream and soil water dissolved organic matter in a boreal forested catchment in
855 northern Sweden. Aquat Sci. 2002;64: 269–281. doi:10.1007/s00027-002-8071-z
856 11. Reis RE, Kullander SO, Ferraris CJ. Checklist of the Freshwater Fishes of South and
857 Central America. Porto Alegre, Brazil: EDIPUCRS; 2003. p. 729.
858 12. Winemiller KO, Willis SC. Biogeography of the Vaupes Arch and Casiquiare River:
859 Barriers and Passages between the Amazon and Orinoco. In: Albert J, Reis RE, editors.
860 Historical Biogeography of Neotrpical Freshwater Fishes. Berkeley, CA: University of
861 California Press; 2010. pp. 225–242.
862 13. Winemiller KO, López-Fernández H, Taphorn DC, Nico L, Barbarino-Duque A. Fish
863 assemblages of the Casiquiare River, a corridor and zoogeographical filter for
864 dispersal between the Orinoco and Amazon basins. J Biogeogr. 2008;35: 1551–1563.
865 14. De Queiroz LJ, Torrente-Vilara G, Quilodran C, da Costa Doria CR, Montoya-Burgos JI.
866 Multifactorial genetic divergence processes drive the onset of speciation in an
867 Amazonian fish. PLoS One. 2017;12. doi:10.1371/journal.pone.0189349
868 15. Wood CM, Robertson LM, Johannsson OE, Val AL. Mechanisms of Na+ uptake,
869 ammonia excretion, and their potential linkage in native Rio Negro tetras
870 (Paracheirodon axelrodi,Hemigrammus rhodostomus, and Moenkhausia diktyota). J

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

871 Comp Physiol B Biochem Syst Environ Physiol. 2014;184: 877–890.
872 16. Gonzalez RJ, Wilson RW, Wood CM, Patrick ML, Val AL. Diverse strategies of ion
873 regulation in fish collected from the ion-poor, acidic Rio Negro. Physiol Biochem Zool.
874 1998;75: 37–47.
875 17. Whitehead A, Roach JL, Zhang S, Galvez F. Genomic mechanisms of evolved
876 physiological plasticity in killifish distributed along an environmental salinity
877 gradient. Proc Natinal Acad Sci USA. 2011;108: 6193–6198.
878 18. Hirata T, Kaneko T, Ono T, Nakazato T, Furukawa F, Hasegawa M, et al. Mechanism of
879 acid adaptation of a fish lining in a pH 3.5 lake. Am J Physiol - Regul Integr Comp
880 Physiol. 2003;284: R1199-1212.
881 19. Kullander SO, Ferreira EJG. A review of the South American cichlid genus Cichla, with
882 descriptions of nine new species (Teleostei : Cichlidae). Ichthyol Explor Freshwaters.
883 2006;17: 289–398.
884 20. Willis SC, Macrander J, Farias IP, Ortí G. Simultaneous delimitation of species and
885 quantification of interspecific hybridization in Amazonian peacock cichlids (genus
886 cichla) using multi-locus data. BMC Evol Biol. 2012;12: 96. Available:
887 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3563476&tool=pmcent
888 rez&rendertype=abstract
889 21. Willis SC. One species or four? Yes!…and, no. Or, arbitrary assignment of lineages to
890 species obscures the diversification processes of Neotropical fishes. PLoS One.
891 2017;12. doi:10.1371/journal.pone.0172349
892 22. Winemiller KO. Ecology of peacock cichlids (Cichla spp.) in Venezuela. J Aquaric
893 Aquat Sci. 2001;9: 93–112.

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

894 23. Willis SC, Nunes MS, Montana CG, Farias IP, Orti G, Lovejoy NR. The Casiquiare River
895 acts as a corridor between the Amazonas and Orinoco River basins: biogeographic
896 analysis of the genus Cichla. Mol Ecol. 2010;19: 1014–1030.
897 24. Willis SC, Farias IP, Ortí G. Testing mitochondrial capture and deep coalescence in
898 amazonian cichlid fishes (Cichlidae: Cichla). Evolution (N Y). 2014;68: 256–268.
899 25. Willis SC, Nunes MS, Montaña CG, Farias IP, Lovejoy NR. Systematics, biogeography,
900 and evolution of the Neotropical peacock basses Cichla (Perciformes: Cichlidae). Mol
901 Phylogenet Evol. 2007;44: 291–307.
902 26. Matsuo AYO, Val AL. Acclimation to humic substances pre- vents whole body sodium
903 loss and stimulates branchial cal- cium uptake capacity in the cardinal tetras,
904 Paracheirodon axelrodi (Schultz) subjected to extremely low pH. J Fish Biol. 2007;70:
905 989–1000.
906 27. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level
907 estimates improve gene-level inferences. F1000Research. 2016;4: 1521.
908 doi:10.12688/f1000research.7563.2
909 28. Davidson NM, Oshlack A. Corset: enabling differential gene expression analysis for de
910 novoassembled transcriptomes. Genome Biol. 2014;15. doi:10.1186/s13059-014-
911 0410-6
912 29. Srivastava A, Sarkar H, Malik L, Patro R. Accurate, Fast and Lightweight Clustering of
913 de novo Transcriptomes using Fragment Equivalence Classes. bioRxiv.
914 2016;1604.03250.
915 30. Götz S, Garia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, et al. High-
916 throughput functional annotation and data mining with the Blast2GO suite. Nucleic

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

917 Acid Res. 2008;36: 3420–3435.

918 31. Belinky F, Nativ N, Stelzer G, Zimmerman S, Iny Stein T, Safran M, et al. PathCards:
919 multi-source consolidation of human biological pathways. Database (Oxford).
920 2015;2015. doi:10.1093/database/bav006

921 32. Boursnell C, Smith-Unna R. Transfuse: Merge transcriptome assemblies [Internet].
922 2014. Available: https://github.com/cboursnell/transfuse

923 33. Liu J, Li G, Chang Z, Yu T, Liu B, McMullen R, et al. BinPacker: Packing-Based De Novo
924 Transcriptome Assembly from RNA-seq Data. PLoS Comput Biol. 2016;12.
925 doi:10.1371/journal.pcbi.1004772

926 34. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: Accelerated for clustering the next-generation
927 sequencing data. Bioinformatics. 2012;28: 3150–3152.
928 doi:10.1093/bioinformatics/bts565

929 35. Rausher MD, Delph LF. Commentary: When does understanding phenotypic
930 evolution require identification of the underlying genes? Evolution. 2015. pp. 1655–
931 1664. doi:10.1111/evo.12687

932 36. Connon R, Jeffries K, Komoroske L, Todgham A, Fangue NA. The utility of
933 transcriptomics in fish conservation. J Exp Biol. 2018;221. doi:10.1242/jeb.148833

934 37. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. A
935 survey of best practices for RNA-seq data analysis. Genome Biol. 2016;17: 13–19.
936 doi:10.1186/s13059-016-0881-8

937 38. Pearson WR. An introduction to sequence similarity ("homology") searching. Curr
938 Protoc Bioinforma. 2013; doi:10.1002/0471250953.bi0301s42

939 39. Pasquier J, Cabau C, Nguyen T, Jouanno E, Severac D, Braasch I, et al. Gene evolution

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

940 and gene expression after whole genome duplication in fish: The PhyloFish database.
941 BMC Genomics. 2016;17. doi:10.1186/s12864-016-2709-z
942 40. Matschiner M, Musilová Z, Barth JMI, Starostová Z, Salzburger W, Steel M, et al.
943 Bayesian phylogenetic estimation of clade ages supports trans-Atlantic dispersal of
944 cichlid fishes. Syst Biol. 2017;66: 3–22. doi:10.1093/sysbio/syw076
945 41. Fiol DF, Sanmarti E, Sacchi R, Kültz D. A novel tilapia prolactin receptor is
946 functionally distinct from its paralog. J Exp Biol. 2009;212: 2007–2015.
947 doi:10.1242/jeb.025601
948 42. Kultz D. The Combinatorial Nature of Osmosensing in Fishes. Physiology. 2012;27:
949 259–275. doi:10.1152/physiol.00014.2012
950 43. McCormick SD, Bradshaw D. Hormonal control of salt and water balance in
951 vertebrates. General and Comparative Endocrinology. 2006. pp. 3–8.
952 doi:10.1016/j.ygcen.2005.12.009
953 44. Guh Y-J, Lin C-H, Hwang P-P. Osmoregulation in zebrafish: ion transport mechanisms
954 and functional regulation. EXCLI J. 2015;14: 627–659. doi:10.17179/excli2015-246
955 45. Fiol DF, Kültz D. Osmotic stress sensing and signaling in fishes. FEBS Journal. 2007.
956 pp. 5790–5798. doi:10.1111/j.1742-4658.2007.06099.x
957 46. Sakamoto T, McCormick SD. Prolactin and growth hormone in fish osmoregulation.
958 Gen Comp Endocrinol. 2006;147: 24–30. doi:10.1016/j.ygcen.2005.10.008
959 47. Breves JP, Inokuchi M, Yamaguchi Y, Seale AP, Hunt BL, Watanabe S, et al. Hormonal
960 regulation of aquaporin 3: Opposing actions of prolactin and cortisol in tilapia gill. J
961 Endocrinol. 2016;230: 325–337. doi:10.1530/JOE-16-0162
962 48. Breves JP, Seale AP, Moorman BP, Lerner DT, Moriyama S, Hopkins KD, et al. Pituitary

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

963 control of branchial NCC, NKCC and Na(+), K (+)-ATPase α-subunit gene expression
964 in Nile tilapia, Oreochromis niloticus. J Comp Physiol B. 2014;184: 513–23.
965 doi:10.1007/s00360-014-0817-0
966 49. Tipsmark CK, Breves JP, Rabeneck DB, Trubitt RT, Lerner DT, Grau EG. Regulation of
967 gill claudin paralogs by salinity, cortisol and prolactin in Mozambique tilapia
968 (Oreochromis mossambicus). Comp Biochem Physiol -Part A Mol Integr Physiol.
969 2016;199: 78–86. doi:10.1016/j.cbpa.2016.05.014
970 50. Prunet P, Sturm A, Milla S. Multiple corticosteroid receptors in fish: From old ideas to
971 new concepts. General and Comparative Endocrinology. 2006. pp. 17–23.
972 doi:10.1016/j.ygcen.2006.01.015
973 51. Breves JP, Watanabe S, Kaneko T, Hirano T, Grau EG. Prolactin restores branchial
974 mitochondrion-rich cells expressing Na+/Cl- cotransporter in hypophysectomized
975 Mozambique tilapia. Am J Physiol Regul Integr Comp Physiol. 2010;299: 702–710.
976 doi:10.1152/ajpregu.00213.2010
977 52. Breves JP, Fox BK, Pierce AL, Hirano T, Grau EG. Gene expression of growth hormone
978 family and glucocorticoid receptors, osmosensors, and ion transporters in the gill
979 during seawater acclimation of mozambique tilapia, Oreochromis mossambicus. J
980 Exp Zool Part A Ecol Genet Physiol. 2010;313 A: 432–441. doi:10.1002/jez.613
981 53. Reinecke M, Schmid A, Ermatinger R, Loffing-Cueni D. Insulin-like growth factor I in
982 the teleost Oreochromis mossambicus, the tilapia: gene sequence, tissue expression,
983 and cellular localization. Endocrinology. 1997;138: 3613–9.
984 doi:10.1210/endo.138.9.5375
985 54. Wood AW, Duan C, Bern HA. Insulin-like growth factor signaling in fish. International

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

986 Review of Cytology. 2005. pp. 215–285. doi:10.1016/S0074-7696(05)43004-1
987 55. Breves JP, Seale AP, Helms RE, Tipsmark CK, Hirano T, Grau EG. Dynamic gene
988 expression of GH/PRL-family hormone receptors in gill and kidney during
989 freshwater-acclimation of Mozambique tilapia. Comp Biochem Physiol - A Mol Integr
990 Physiol. 2011;158: 194–200. doi:10.1016/j.cbpa.2010.10.030
991 56. Shepherd BS, Sakamoto T, Nishioka RS, Richman NH, Mori I, Madsen SS, et al.
992 Somatotropic actions of the homologous growth hormone and prolactins in the
993 euryhaline teleost, the tilapia, Oreochromis mossambicus. Proc Natl Acad Sci U S A.
994 1997;94: 2068–72. doi:10.1073/pnas.94.5.2068
995 57. Ferlazzo A, Carvalho ESM, Gregorio SF, Power DM, Canario AVM, Trischitta F, et al.
996 Prolactin regulates luminal bicarbonate secretion in the intestine of the sea bream
997 (Sparus aurata L.). J Exp Biol. 2012;215: 3836–3844. doi:10.1242/jeb.074906
998 58. Chasiotis H, Kolosov D, Bui P, Kelly SP. Tight junctions, tight junction proteins and
999 paracellular permeability across the gill epithelium of fishes: A review. Respir
1000 Physiol Neurobiol. 2012;184: 269–281.
1001 59. Zihni C, Mills C, Matter K, Balda MS. Tight junctions: From simple barriers to
1002 multifunctional molecular gates. Nature Reviews Molecular Cell Biology. 2016. pp.
1003 564–580. doi:10.1038/nrm.2016.80
1004 60. Wright PA, Wood CM. A new paradigm for ammonia excretion in aquatic animals:
1005 role of Rhesus (Rh) glycoproteins. J Exp Biol. 2009;212: 2303–2312.
1006 61. Furukawa F, Watanabe S, Kakumura K, Hiroi J, Kaneko T. Gene expression and
1007 cellular localization of ROMKs in the gills and kidney of Mozambique tilapia
1008 acclimated to fresh water with high potassium concentration. AJP Regul Integr Comp

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

1009 Physiol. 2014;307: R1303–R1312. doi:10.1152/ajpregu.00071.2014

1010 62. Parks SK, Tresguerres M, Goss GG. Theoretical considerations underlying Na_+_
1011 uptake mechanisms in freshwater fishes. Comp Biochem Physiol C-Toxicology
1012 Pharmacol. 2008;148: 411–418. doi:10.1016/j.cbpc.2008.03.002

1013 63. Furukawa F, Watanabe S, Inokuchi M, Kaneko T. Responses of gill mitochondria-rich
1014 cells in Mozambique tilapia exposed to acidic environments (pH 4.0) in combination
1015 with different salinities. Comp Biochem Physiol Part A. 2011;158: 468–476.

1016 64. Saparov SM, Liu K, Agre P, Pohl P. Fast and selective ammonia transport by
1017 aquaporin-8. J Biol Chem. 2007;282: 5296–5301. doi:10.1074/jbc.M609343200

1018 65. Postel R, Sonnenberg A. Carbonic Anhydrase 5 regulates acid-base homeostasis in
1019 Zebrafish. PLoS One. 2012;7. doi:10.1371/journal.pone.0039881

1020 66. Lam SH, Lui EY, Li Z, Cai S, Sung WK, Mathavan S, et al. Differential transcriptomic
1021 analyses revealed genes and signaling pathways involved in iono-osmoregulation
1022 and cellular remodeling in the gills of euryhaline Mozambique tilapia, Oreochromis
1023 mossambicus. BMC Genomics. 2014;15. doi:10.1186/1471-2164-15-921

1024 67. Lin C-H, Kuan W-C, Liao B-K, Deng A-N, Tseng D-Y, Hwang P-P. Environmental and
1025 cortisol-mediated control of Ca2+ uptake in tilapia (Oreochromis mossambicus). J
1026 Comp Physiol B. 2016;186: 323–332. doi:10.1007/s00360-016-0963-7

1027 68. Li Z, Lui EY, Wilson JM, Ip YK, Lin Q, Lam TJ, et al. Expression of key ion transporters
1028 in the gill and esophageal- gastrointestinal tract of euryhaline mozambique tilapia
1029 oreochromis mossambicus acclimated to fresh water, seawater and hypersaline
1030 water. PLoS One. 2014;9. doi:10.1371/journal.pone.0087591

1031 69. Wilson JM, Laurent P, Tufts BL, Benos DJ, Donowitz M, Vogl AW, et al. NaCl uptake by

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

1032 the branchial epithelium in freshwater teleost fish: an immunological approach to
1033 ion-transport protein localization. J Exp Biol. 2000;203: 2279–96. doi:10.1016/0003-
1034 2697(76)90527-3
1035 70. Chang IC, Hwang PP. Cl- uptake mechanism in freshwater-adapted tilapia
1036 (_Oreochromis mossambicus_). Physiol Biochem Zool. 2004;77: 406–414.
1037 doi:10.1086/383505
1038 71. Tang C-H, Lee T-H. Ion-Deficient Environment Induces the Expression of Basolateral
1039 Chloride Channel, ClC-3-Like Protein, in Gill Mitochondrion-Rich Cells for Chloride
1040 Uptake of the Tilapia Oreochromis mossambicus. Physiol Biochem Zool. 2011;84:
1041 54–67. doi:10.1086/657161
1042 72. Araújo JDA, Ghelfi A, Val AL. Triportheus albus Cope, 1872 in the Blackwater,
1043 Clearwater, and Whitewater of the Amazon: A Case of Phenotypic Plasticity? Front
1044 Genet. 2017;8. doi:10.3389/fgene.2017.00114
1045 73. Kultz D. Physiological mechanisms used by fish to cope with salinity stress. J Exp Biol.
1046 2015;218: 1907–1914. doi:10.1242/jeb.118695
1047 74. Liu Y, Beyer A, Aebersold R. On the Dependency of Cellular Protein Levels on mRNA
1048 Abundance. Cell. 2016. pp. 535–550. doi:10.1016/j.cell.2016.03.014
1049 75. Perry SF. THE CHLORIDE CELL:Structure and Function in the Gills of Freshwater
1050 Fishes. Annu Rev Physiol. 1997;59: 325–347. doi:10.1146/annurev.physiol.59.1.325
1051 76. Inokuchi M, Hiroi J, Watanabe S, Lee KM, Kaneko T. Gene expression and
1052 morphological localization of NHE3, NCC and NKCC1a in branchial mitochondria-rich
1053 cells of Mozambique tilapia (Oreochromis mossambicus) acclimated to a wide range
1054 of salinities. Comp Biochem Physiol - A Mol Integr Physiol. 2008;151: 151–158.

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

1055 doi:10.1016/j.cbpa.2008.06.012
1056 77. Wood CM, Iftikar FI, Scott GR, De Boeck G, Sloman KA, Matey V, et al. Regulation of
1057 gill transcellular permeability and renal function during acute hypoxia in the
1058 Amazonian oscar (Astronotus ocellatus): new angles to the osmorespiratory
1059 compromise. J Exp Biol. 2009;212: 1949–1964. doi:10.1242/jeb.028464
1060 78. Johannsson OE, Smith DS, Sadauskas-Henrique H, Cimprich G, Wood CM, Val AL.
1061 Photo-oxidation processes, properties of DOC, reactive oxygen species (ROS), and
1062 their potential impacts on native biota and carbon cycling in the Rio Negro
1063 (Amazonia, Brazil). Hydrobiologia. 2017;789: 7–29. doi:10.1007/s10750-016-2687-
1064 9
1065 79. Gonzalez RJ, Wood CM, Wilson RW, Patrick ML, Bergman HL, Narahara a, et al.
1066 Effects of water pH and calcium concentration on ion balance in fish of the Rio Negro,
1067 Amazon. Physiol Zool. 1998;71: 15–22. doi:10.1086/515893
1068 80. Al-Reasi HA, Smith SD, Wood CM. The influence of dissolved organic matter (DOM)
1069 on sodium regulation and nitrogenous waste excretion in the zebrafish (*Danio rerio*).
1070 J Exp Biol. 2016;219: 2289–2299. doi:10.1242/jeb.139444
1071 81. Wood CM, Matsuo AYO, Gonzalez RJ, Wilson RW, Patrick ML, Val AL. Mechanisms of
1072 ion transport in Potamotrygon, a stenohaline freshwater elasmobranch native to the
1073 ion-poor blackwaters of the Rio Negro. J Exp Biol. 2002;205: 3039–3054. Available:
1074 http://jeb.biologists.org/content/205/19/3039.abstract
1075 82. Gonzalez RJ, Wood CM, Wilson RW, Patrick ML, Bergman A, Narahara A, et al. Effects
1076 of water pH and calcium concentration on ion balance in fish of the Rio Negro,
1077 Amazon. Physiol Zool. 1998;71: 15–22.

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

1078 83. Matsuo AYO, Val AL. Low pH and calcium effects on net Na+ and K+ fluxes in two
1079 catfish species from the Amazon River (Corydoras: Callichthyidae). Brazilian J Med
1080 Biol Res. 2002;35: 361–367.
1081 84. Schulte PM. What is environmental stress? Insights from fish living in a variable
1082 environment. J Exp Biol. 2014;217: 23–34. doi:10.1242/jeb.089722
1083 85. Shaw JR, Hampton TH, King BL, Whitehead A, Galvez F, Gross RH, et al. Natural
1084 selection canalizes expression variation of environmentally induced plasticity-
1085 enabling genes. Mol Biol Evol. 2014;31: 3002–3015.
1086 86. Bystriansky JS, Frick NT, Richards JG, Schulte PM, Ballantyne JS. Failure to up-
1087 regulate gill Na+,K+-ATPase α-subunit isoform α1b may limit seawater tolerance of
1088 land-locked Arctic char (Salvelinus alpinus). Comp Biochem Physiol Part A Mol Integr
1089 Physiol. 2007;148: 332–338.
1090 87. Moorman BP, Lerner DT, Grau EG, Seale AP. The effects of acute salinity challenges
1091 on osmoregulation in Mozambique tilapia reared in a tidally changing salinity. J Exp
1092 Biol. 2015;218: 731–739. doi:10.1242/jeb.112664
1093 88. Hoeinghaus DJ, Layman CA, Arrington DA, Winemiller KO. Movement of Cichla
1094 species (Cichlidae) in a Venezuelan floodplain river. Neotrop Ichthyol. 2003;1: 121–
1095 126.
1096

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

1097 Supplemental Tables

1098

1099 Supplemental Table 1. Statistics and scores from Transfuse-merged and constituent (pre-
1100 merge) *de novo* transcriptome assemblies. K: kmer; # transcripts (>200bp); number of
1101 transcripts in the resulting assembly larger than 200 base pairs; N50: smallest contig above
1102 which 50% of the length of the assembly is found; mean: mean length of contigs in
1103 assembly in base pairs; length (bp): combined length of assembly in base pairs; Detonate
1104 score: likelihood score for each assembly by Detonate program (smaller is better);
1105 TransRate score: combined score for each assembly by TransRate program (larger is
1106 better); TRate good: number of contigs in assembly in the "optimal" set; % good: percent of
1107 contigs in assembly in the "optimal" set; TR opt.score: score if only "good" contigs are
1108 considered (score is generated by iteratively adding high scoring contigs); total mappings:
1109 total number of mapped reads (by Salmon); % good mappings: acceptable mappings by
1110 Transrate criteria; BUSCO: number of complete conserved genes identified by BUSCO
1111 program (more is better); BUSCO p-mc: proportion of complete genes in represented by
1112 multiple transcripts; BUSCO mc: number of complete genes in represented by multiple
1113 transcripts; BUSCO miss.: number of genes from the BUSCO set that were not recovered;
1114 BUSCO out: "short" output from BUSCO program.

1115

1116 Supplemental Table 2. Unique Refseq protein hits to *Oreochromis niloticus*, supplemented
1117 with *Neolaprologus brichardi*, *Haplochromis burtoni*, *Maylandia zebra*, *Pundamilia nyererei*,
1118 and *Danio rerio* ("cichlid+"; see text), for the sufficiently expressed and differentially
1119 expressed transcript assemblies. Longest_contig: name of the longest contig blast-

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

1120 annotated to that gene; length(bp): length in base pairs of that contig; Expression_Change:
1121 expression change in the blackwater treatment relative to the whitewater treatment across
1122 bioinformatic combinations; Mean_Log_Expression: Mean counts per million expression
1123 across six samples from limma voom with Corset clustering at ~70% read co-mapping (-d
1124 0.3) and quantification with Bowtie2/RSEM, on log2 scale; Log_Fold_Change: log2 fold
1125 change of blackwater relative to whitewater samples; FDR: false discovery rate, i.e. p-value
1126 adjusted for multiple tests; cichlid+_ncbi_accession: NCBI protein accession for cichlid+
1127 annotation; cichlid+_ncbi_geneID: NCBI GI for gene corresponding to protein accession for
1128 cichlid+ annotation; cichlid+_nbci_symbol: NCBI gene symbol for gene corresponding to
1129 protein accession for cichlid+ annotation; cichlid+_description: NCBI gene description for
1130 gene corresponding to protein accession for cichlid+ annotation; human_ncbi_accession:
1131 NCBI GI for gene corresponding to protein accession for human annotation;
1132 human_ncbi_geneID: NCBI gene symbol for gene corresponding to protein accession for
1133 human annotation; human_symbol: NCBI gene symbol for gene corresponding to protein
1134 accession for human annotation; human_description: NCBI gene description for gene
1135 corresponding to protein accession for cichlid+ annotation; danio_ncbi_accession: NCBI GI
1136 for gene corresponding to protein accession for *Danio rerio* annotation; danio_ncbi_geneID:
1137 NCBI gene symbol for gene corresponding to protein accession for *Danio rerio* annotation;
1138 danio_symbol: NCBI gene symbol for gene corresponding to protein accession for *Danio*
1139 *rerio* annotation; danio_description: NCBI gene description for gene corresponding to
1140 protein accession for *Danio rerio* annotation.
1141

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

1142 Supplemental Table 3. Unique hits to human proteins for the sufficiently expressed and
1143 differentially expressed transcript assemblies. For differentially expressed transcripts, the
1144 expression for the blackwater treatment relative to the whitewater treatment is shown.
1145 Genes present in surveyed osmoregulatory pathways are identified: pr, *prolactin signaling*
1146 *pathway*; gh, *growth hormone receptor pathway*; aq, *aquaporin mediated transport*, tjc,
1147 *epithelial tight junctions* (Qiagen) or *tight junctions* (KEGG); smt, *transport of glucose and*
1148 *other sugars, bile salts and organic acids, metal ions and amine compounds*. For genes that
1149 are differentially expressed, the expression for the blackwater treatment relative to the
1150 whitewater treatment is indicated. * Duplicate hits show contrasting expression, most often
1151 due to many-to-one paralogy; see text.
1152
1153 Supplemental Figures
1154
1155 Supplemental Figure 1. Map showing localities from which the same mitochondrial control
1156 region haplotype as the experimental fish was sampled (red), from which the containing
1157 mtDNA clade was sampled (orange), from which other *Cichla oc. monoculus* haplotype
1158 clades were sampled (blue), and where other evolutionary significant units of *Cichla*
1159 *ocellaris* were sampled (black). The Amazonas and Orinoco basins are identified, along with
1160 the Negro sub-basin, and the Casiquiare River that connects the Amazonas and Orinoco
1161 Basins. Major blackwater areas are identified by shading; smaller blackwater rivers occur
1162 sporadically throughout the lowland Amazonas and Orinoco basins.
1163

bioRxiv preprint doi: https://doi.org/10.1101/489211; this version posted December 6, 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.

1164 Supplemental Figure 2. Determination of the union of differentially expressed among 27
1165 combinations of clustering (RAPCLUST, CORSET with -d 0.3, CORSET with -d 0.7),
1166 mapping/quantification (SALMON to SALMON, BOWTIE2 to SALMON, or BOWTIE2 to RSEM), and
1167 statistical procedure (DESEQ2, EDGER, or LIMMA). The union of clusters is evaluated for
1168 statistical procedure and mapping/quantification, while for clustering algorithms, where
1169 cluster names are not comparable, the union of transcripts is made. Determination of
1170 sufficiently expressed transcripts was similar except filtering produced the same results
1171 regardless of statistical procedure (total nine combinations).

1172

1173 Supplemental Files

1174

1175 Supplemental File 1. GO term enrichment. Tags: GO terms were "over"-enriched or "under"
1176 enriched; GO ID: standard ID number of GO term; GO Name: GO term description; GO
1177 Category: biological arena to which GO term refers; FDR: false discovery rate, i.e. p-value
1178 corrected for multiple testing; P-Value: uncorrected p-value; Nr Test: number in the test
1179 group annotated with GO term; Nr Reference number in the reference group annotated
1180 with GO term; Non Annot Test: number in the test group not annotated with GO term; Non
1181 Annot Reference: number in the reference group not annotated with GO term; difference in
1182 proportions: difference between the numbers in each test or reference group that were
1183 annotated with the GO term relative to those in each group that were not annotated with
1184 this term; difference in numbers: absolute difference in numbers in each test or reference
1185 group that were annotated with the GO term

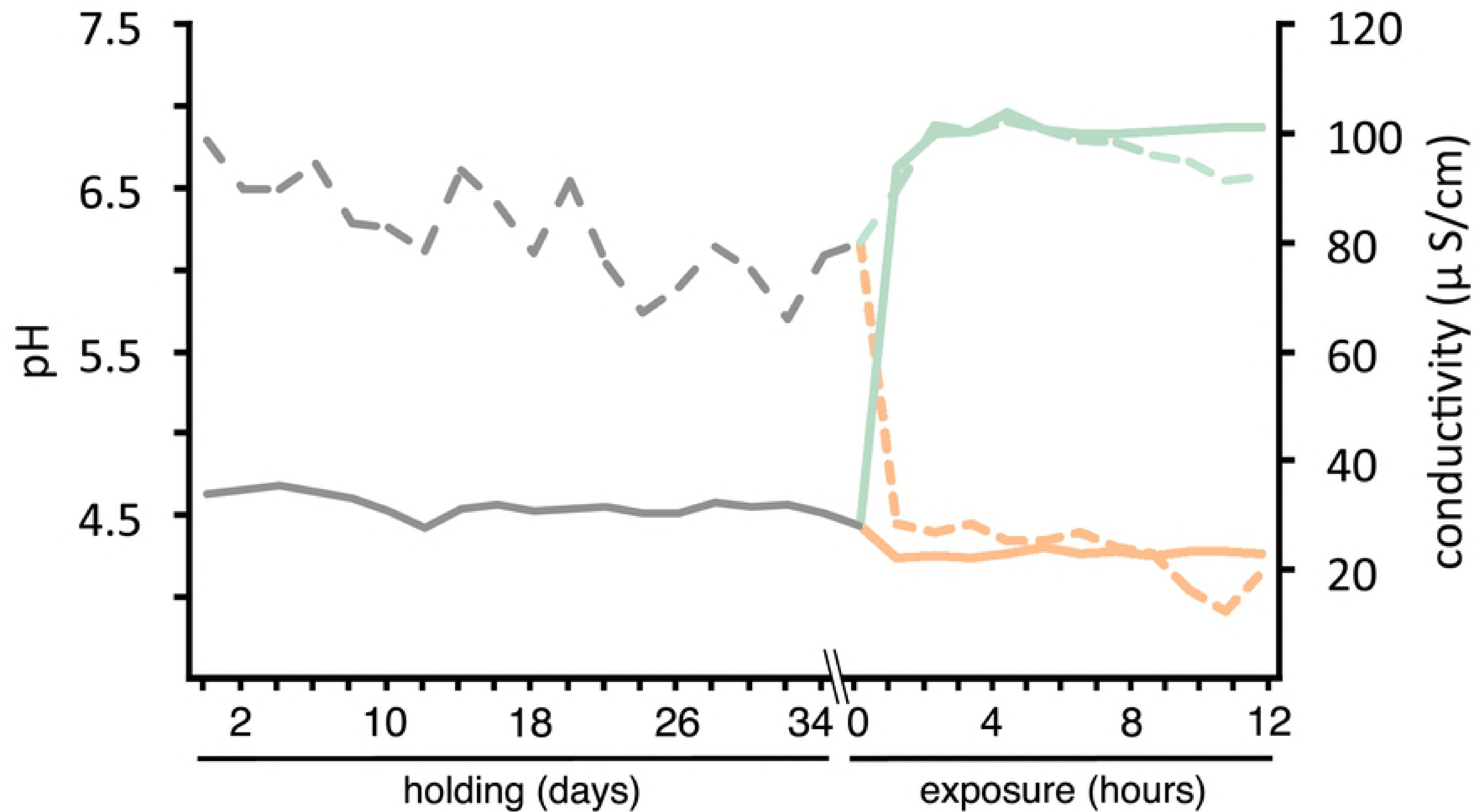

Figure 1

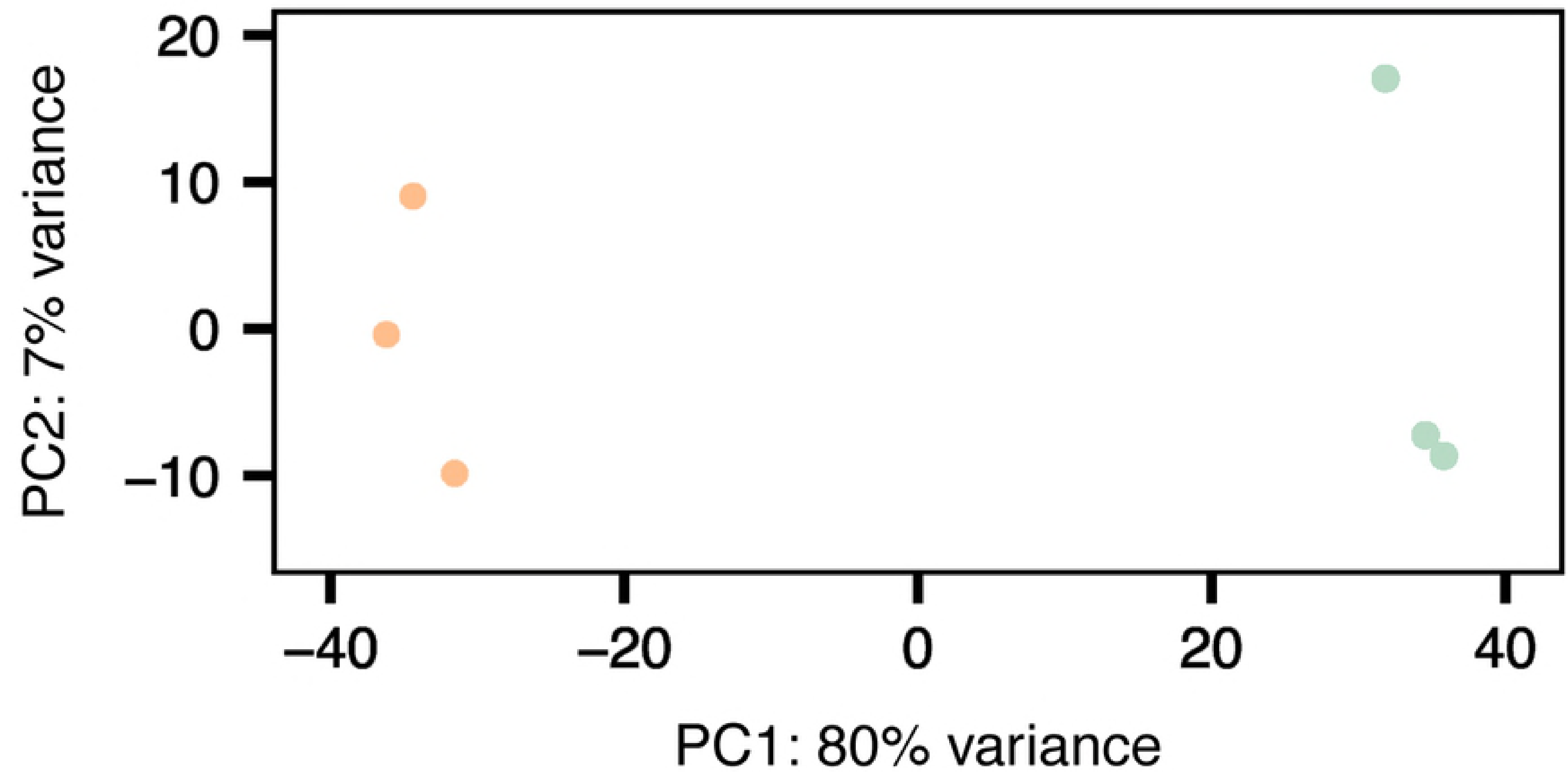

Figure 2

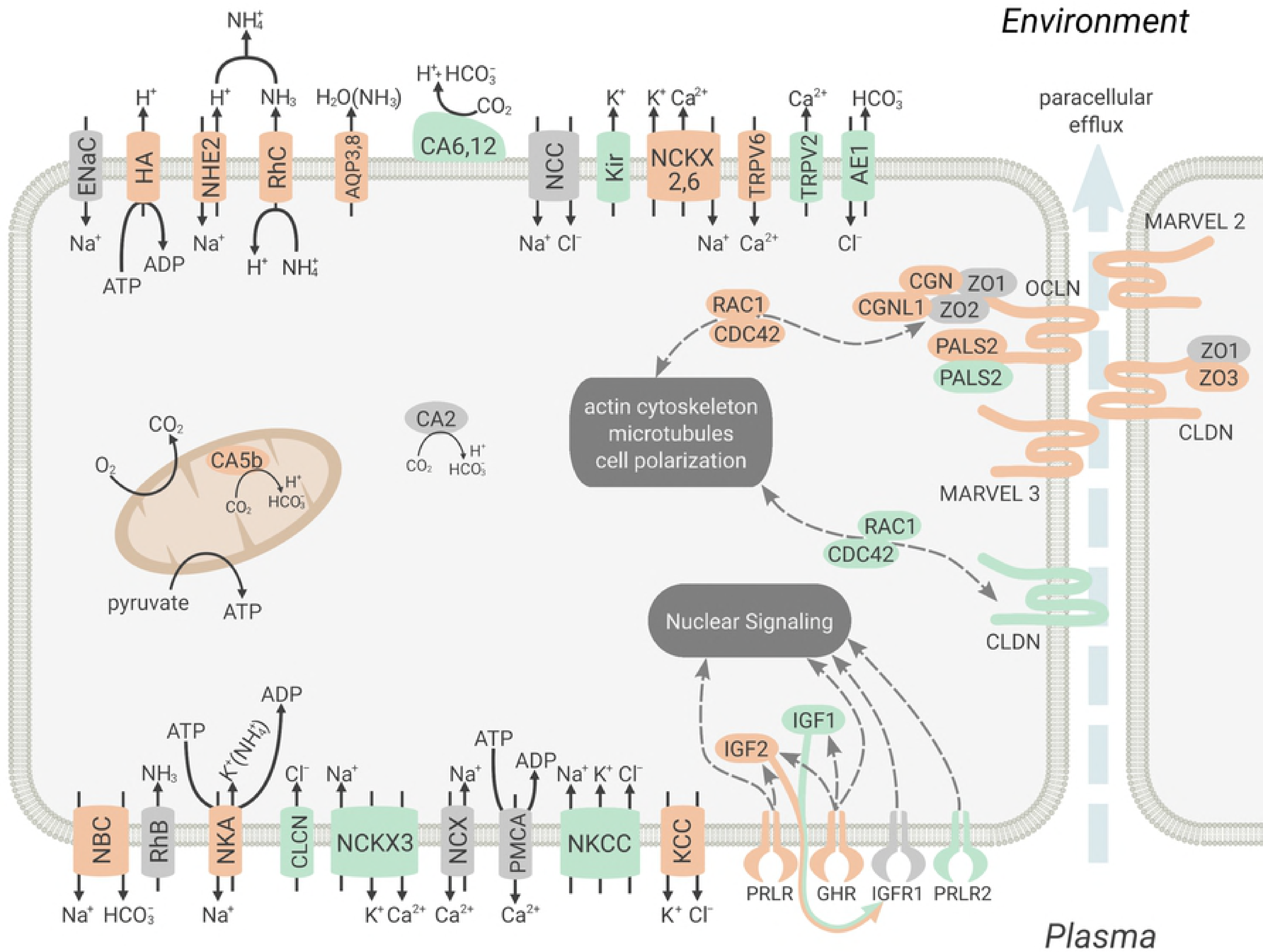

Figure 3