The transcriptional correlates of divergent electric organ discharges in *Paramormyrops* electric fish

Mauricio Losilla and Jason R. Gallant*

Department of Integrative Biology, Michigan State University, East Lansing, MI 48824, USA.
Graduate Program in Ecology, Evolutionary Biology and Behavior, Michigan State University, East Lansing, MI 48824, USA.
BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, MI 48824, USA.

*Author to whom correspondence should be addressed: jgallant@msu.edu

Abstract

Background: Understanding the genomic basis of phenotypic diversity can be greatly facilitated by examining adaptive radiations with hypervariable traits. In this study, we focus on a rapidly diverged species group of mormyrid electric fish in the genus *Paramormyrops*, which are characterized by extensive phenotypic variation in electric organ discharges (EODs). The main components of EOD diversity are waveform duration, complexity and polarity. Using an RNA-sequencing based approach, we sought to identify gene expression correlates for each of these EOD waveform features by comparing 11 specimens of *Paramormyrops* that exhibit variation in these features.

Results: Patterns of gene expression among *Paramormyrops* are highly correlated, and 3,274 genes (16%) were differentially expressed. Using our most restrictive criteria, we detected 71-

144 differentially expressed genes correlated with each EOD feature, with little overlap between them. The predicted functions of several of these genes are related to extracellular matrix, cation homeostasis, lipid metabolism, and cytoskeletal and sarcomeric proteins. These genes are of significant interest given the known morphological differences between electric organs that underlie differences in the EOD waveform features studied.

Conclusions: In this study, we identified plausible candidate genes that may contribute to phenotypic differences in EOD waveforms among a rapidly diverged group of mormyrid electric fish. These genes may be important targets of selection in the evolution of species-specific differences in mate-recognition signals.

Introduction

Understanding the genomic basis of phenotypic diversity is a major goal of evolutionary biology [1]. Adaptive radiations and explosive diversification of species [2] are frequently characterized by interspecific phenotypic differences in divergence of few, hypervariable phenotypic traits [3–6]. Such systems offer exceptional advantages to study the genomic bases of phenotypic diversity: they can provide replication under a controlled phylogenetic framework [7], and couple ample phenotypic differentiation with relatively "clean" genomic signals between recently diverged species [8]. Study of the genomic mechanisms underlying hypervariable phenotypic traits has identified, in some cases relatively simple genetic architectures [9–13]. More often, the genetic architecture underlying such traits can be complex and polygenic [14–17]. It has long been recognized that changes in gene expression can affect phenotypic differences between species [18], and RNA-seq based approaches have greatly facilitated the study of this relationship [19]. A growing number of studies have examined differences in gene expression in phenotypic evolution (e.g., [19–27]). While these studies do not implicate mutational causes, analysis of differential gene expression (DGE) can be a useful approach in examining the genomic basis of divergent phenotypes.

African weakly electric fish (Teleostei: Mormyridae) are among the most rapidly speciating groups of ray-finned fishes [28,29]. This is partly due to the diversification of the genus *Paramormyrops* [30,31] in the watersheds of West-Central Africa, where more than 20 estimated species [32] have evolved within the last 0.5-2 million years [30]. Extensive evidence has demonstrated that electric organ discharges (EODs) exhibit little intraspecific variation, yet differ substantially among mormyrid species [33–35]. This pattern is particularly evident in

54 *Paramormyrops* [30,36], in which EOD waveforms evolve much faster than morphology, size,
55 and trophic ecology [37].

56 Mormyrid EODs are a behavior with a dual role in electrolocation [38,39] and
57 intraspecific communication [40,41]. EOD waveforms vary between species principally in terms
58 of their complexity, polarity, and duration [30,42], and all three dimensions of variation are
59 evident among *Paramormyrops* (Fig. 1). Furthermore, recent discoveries of intraspecific
60 polymorphism in EOD waveform in *P. kingsleyae* [43] and polarity among *P.* sp. 'magnostipes'
61 [35] present a unique opportunity to study the genomic basis of phenotypic traits within a rapidly
62 diverging species group.

63 EODs have a well-understood morphological (Fig. 1) and neurophysiological basis
64 [44,45]. EODs are generated by specialized cells (electrocytes) that constitute the electric organ
65 (EO), located in the caudal peduncle [46]. Mormyrid EOs are comprised of 80-360 electrocytes
66 [34], and an individual EOD is produced when the electrocytes discharge synchronously. EODs
67 are multiphasic because they result from action potentials produced by two excitable membranes:
68 the two large phases of the EOD, called P1 and P2, are produced by spikes generated by the
69 posterior and anterior electrocyte faces, respectively [47]. There is a relationship between EODs
70 of longer duration and increased surface membrane area [48], likely mediated at least in part by
71 an increase in membrane capacitance [49,50]. The duration of EODs is highly variable within
72 mormyrids-- some EODs are extremely long (>15 ms) and others are very brief (0.2 ms) [32].

73 Within the Mormyridae, triphasic EODs evolved early from biphasic EODs; however,
74 there have been multiple parallel reversions to biphasic EODs across mormyrids and within
75 *Paramormyrops* [36,43]. Triphasic (P0-present) EODs are produced by electrocytes that are

innervated on the anterior face and have penetrating stalks (*Pa*, P-type), whereas biphasic (P0-absent) EODs are produced by electrocytes innervated on the posterior face and lack penetrating stalks (*NPp*, N-type) (for more details see [42,43,47,48,51,52]). We refer to triphasic EODs as more 'complex' than biphasic EODs. In some cases, triphasic EODs display an unusually large P0 phase, which gives the appearance of an 'inverted' polarity. This is exemplified by the type I EODs of *P.* sp. 'magnostipes' (Fig. 1) [35]. The number [47] and diameter [34,43] of stalk penetrations are positively correlated with the magnitude of P0. We refer to individuals with large penetrations as 'inverted' polarity and individuals with small penetrations as 'normal' polarity.

Recent studies in mormyrids [53–57] have adopted a candidate gene approach to examine the molecular basis of variation in EOD duration on macroevolutionary scales, implicating voltage gated sodium channels (e.g. *scn4aa*) and potassium channels (e.g. *kcna7a*) as key targets of selection during EOD evolution. Beyond this recent attention to ion channels, several studies have described the importance of structural differences between EOs as an important component of EOD variation [43,48,50]. In this study, we took a transcriptome-wide approach to characterizing the molecular basis of electric signal diversity in *Paramormyrops* species divergent for EOD complexity, duration and polarity. We used RNA-sequencing to comprehensively examine DGE in the adult EOs of five *Paramormyrops* operational taxonomic units (OTUs), leveraging a recently sequenced and annotated genome assembly from the species *P. kingsleyae* (N-type) [58], and identify gene expression correlates of each of the three main EOD waveform features of electric signal diversity in *Paramormyrops*. Our results emphasize genes that influence the shape and structure of the electrocyte cytoskeleton, membrane and

98 extracellular matrix (ECM) to exhibit predictable differences between *Paramormyrops* species
99 with divergent EOD phenotypes.

100 **Methods**

101 *Sample collection*

102 We captured 11 *Paramormyrops* individuals from Gabon, West Central Africa in 2009:
103 five *P. kingsleyae* (n=3 N-type and n=2 P-type), four *P.* sp. 'magnostipes' (n=2 Type I and n=2
104 Type II), and two *P.* sp. 'SN3'. Within 1-12 hours of capture, individual specimens were
105 euthanized by overdose with MS-222. The caudal peduncle was excised and skinned, and
106 immediately immersed in RNA-later for 24h at 4˚C, before being transferred to -20˚C for long-
107 term storage. As two of these species (*P.* sp. 'magnostipes', *P.* sp. 'SN3') are presently
108 undescribed, we note that these specimens were identified by their EOD waveform, head
109 morphology and collecting locality [30,31,35,59]. All specimens, including vouchers materials,
110 are deposited in the Cornell University Museum of Vertebrates. Collection information and the
111 phenotypes per EOD feature of each sample are detailed in Table 1.

112 *RNA extraction, cDNA library preparation and Illumina Sequencing*

113 Total RNA was extracted from EOs using RNA-easy Kit (Qiagen, Inc) after
114 homogenization with a bead-beater (Biospec, Inc.) in homogenization buffer. mRNA was
115 isolated from total RNA using a NEBNext mRNA Isolation Kit (New England Biolabs, Inc.).
116 Libraries for RNA-seq were prepared using the NEBNext mRNA Sample Prep Master Mix Set,
117 following manufacturer's instructions. Final libraries after size selection ranged from 250-367
118 bp. Libraries were pooled and sequenced by the Cornell University Biotechnology Resource
119 Center Genomics Core on an Illumina HiSeq 2000 in a 2x100bp paired end format. Raw
120 sequence reads were deposited in the NCBI SRA (Supplemental File 1).

Read processing and data exploration

FastQC v0.11.3 (Babraham Bioinformatics) was used to manually inspect raw and processed reads. We used Trimmomatic v.0.32 [60] to remove library adaptors, low quality reads, and filter small reads; following the suggested settings of MacManes [61]: 2:30:10 SLIDINGWINDOW:4:5 LEADING:5 TRAILING:5 MINLEN:25. After trimming, reads from each specimen were aligned to the predicted transcripts of the NCBI-annotated (Release 100) *P. kingsleyae* (N-type) genome [58] using bowtie 2 v2.3.4.1 [62]. Expression quantification was estimated at the gene level using RSEM v1.3.0 [63], followed by exploration of the data with a gene expression correlation matrix based on Euclidean distances and Pearson's correlation coefficient (for genes with read counts >10, Trinity's default parameters). All these steps were executed using scripts included with Trinity v2.6.6 [64,65].

Data Analysis

We began by examining DGE between all possible pairwise comparisons of OTUs (n = 10, Table 2) using edgeR v3.20.9 [66] through a script provided with Trinity. We restricted our consideration of genes to those where CPM-transformed counts were > 1 in at least two samples for each comparison (edgeR default parameters). We modified this to use the function estimateDisp() instead of the functions estimateCommonDisp() and estimateTagwiseDisp(). For each comparison, we conservatively considered genes to be differentially expressed with a minimum fold change of 4 and p-value of 0.001 after FDR correction. We compiled a non-redundant list of genes that were differentially expressed in at least one comparison based on these criteria (Fig. 2, Set A).

For each of the differentially expressed genes (DEGs) in Set A, we TMM normalized, log2(TMM +1) transformed, and mean-centered their expression values. We used the

144 transformed values to compare gene expression between groups of OTUs with alternative EOD
145 waveform phenotypes (i.e. *long duration EOD vs. short duration EOD*, *biphasic vs. triphasic*
146 and *small penetrations vs. large penetrations,* see Table 1. Note that waveform polarity
147 phenotypes only apply to triphasic individuals). For each of the three phenotype pairs, we
148 extracted the genes that were on average more than four times more highly expressed in one
149 phenotype than the other. This resulted in six lists of upregulated genes, one for each EOD
150 feature across all OTUs and samples (Fig. 2, Set B).

151 In order to assess enrichment of particular gene pathways, biological functions, and
152 cellular locations using a controlled vocabulary, we performed Gene Ontology (GO) [67,68]
153 enrichment tests on every list of upregulated genes from (1) the ten pairwise comparisons (n=20,
154 two per comparison) and (2) Set B (n=6), for each of the three ontology domains: Biological
155 Process, Cellular Component, and Molecular Function. First, we identified homologous proteins
156 predicted from the *P. kingsleyae* (N-type) reference genome and those predicted from *Danio*
157 *rerio* (GRCz11) by blastp (BLAST+ v2.6, [69]). For each protein, the top hit (e-value ≤ 1e-10)
158 was used for annotation. Next, we used mygene v1.14.0 [70,71] to match the *D. rerio* proteins to
159 *D. rerio* genes and extract their GO annotations (zebrafish Zv9). This resulted in GO annotations
160 for each of the three ontology domains for *P. kingsleyae* (N-type) genes. Finally, we carried out
161 the GO enrichment tests using topGO v2.30.1 [72] and the following parameters: nodeSize = 10,
162 statistic = fisher, algorithm = weight01, p-value ≤ 0.02. The 'universe' for each enrichment test
163 on gene lists from the pairwise comparisons was all the genes deemed expressed in the
164 respective comparison, whereas the non-redundant list of genes in these ten 'universes' was the
165 'universe' for all enrichment tests on the gene lists from Set B.

166 Interpretation of lists of genes from Set A and Set B each suffered limitations for the
167 overall goals of this analysis, which is to identify the DEGs most strongly associated with each
168 waveform feature (duration, complexity, and polarity). The ten comparisons made to construct
169 Set A were not equally informative for two primary reasons: (1) the OTUs in this analysis vary
170 in terms of their phylogenetic relatedness (see [30,31]) and (2) several OTU comparisons varied
171 in more than one waveform characteristic (Table 2). As such, we elected to focus on the most
172 informative comparison for each EOD feature: the comparison that contrasted only the given
173 feature and that minimized phylogenetic distance between OTUs. Of the ten pairwise
174 comparisons, we classified three as the most informative comparisons, one per EOD feature
175 (Table 2). The six lists of upregulated genes from these three comparisons constitute Set A'.

176 Comparisons in Set A'; however, lack biological replication. In contrast, interpretations
177 of Set B were potentially limited in that many of the OTUs in this analysis differed in more than
178 one EOD feature. To circumvent the limitations of Sets A' and B within the limits of our study
179 design, we constructed a third set (Set C). Set C is defined as the intersection of the upregulated
180 genes and their enriched GO terms from Sets A' and B, for each phenotype. Since there were six
181 phenotypes in our study, Set C encompasses six lists of upregulated genes and their respective
182 enriched GO terms (Fig. 2). Therefore, Set C represents the DEGs that are (1) differentially
183 expressed between closely-related OTUs that vary in a single waveform characteristic, and (2)
184 are consistently differentially expressed among all OTUs that share that waveform feature. We
185 focus our attention on Set C: We retrieved GO term definitions from QuickGO [73] and
186 descriptions of gene function of the functional annotations from UniProt [74]; and to facilitate
187 the discussion, we classified the more interesting genes in Set C into "general" functional

188 classes, or themes. All source code necessary to perform the methods described here is provided
189 in a GitHub repository: http://github.com/msuefishlab/paramormyrops_rnaseq.

190 **Results**

191 *Overall Results*

192 Overall alignment rates to the *Paramormyrops kingsleyae* reference transcriptome ranged
193 from 28-74% (>375 million sequenced reads in total, 50% aligned), with no clear differences
194 among OTUs (Supplemental File 1). On inspection, we concluded that these rates are a
195 consequence of the presence of overrepresented sequences from rRNA, mtDNA and bacterial
196 contamination in the RNA-seq reads.

197 Fig. 3 shows a heatmap of pairwise correlations of gene expression for 24,960 genes
198 across all 11 samples. Our ten DGE comparisons detected a range of 16,420-19,273 expressed
199 genes. Intersection of these lists resulted in a non-redundant list of 20,197 genes expressed in EO
200 across all DGE comparisons. We found that 3,274 (16%) were differentially expressed in at least
201 one comparison, and expression patterns across all OTUs were highly correlated (Pearson's r >
202 0.89, Fig. 3). Despite this, correlation values were higher among recognized OTUs, except for
203 the *P.* sp. 'magnostipes type II' 6768 sample (Fig. 3). Thus, we excluded comparisons with the
204 *P.* sp. 'magnostipes type II' OTU from the informative comparisons for Set A'.

205 *Set A: Differential Expression Analysis*

206 We found between 489-1542 DEGs (50-128 enriched GO terms) in every comparison
207 except *P.* sp. 'magnostipes type I' vs *P.* sp. 'magnostipes type II', which had only nine DEGs
208 with seven enriched GO terms (Table 2). Supplemental File 2 provides a tabular list of DEGs for

each comparison, and Supplemental File 3 provides a tabular list of enriched GO terms for each comparison.

We chose the phylogenetically most informative comparisons (see methods) to construct Set A', which are indicated in Table 2. We found: 507 DEG and 69 enriched GO terms comparing *P. kingsleyae* (N-type) vs *P.* sp. 'SN3' (EOD duration); 1322 DEG and 77 enriched GO terms comparing *P. kingsleyae* (P-type) vs *P.* sp. 'magnostipes type I' (waveform polarity); and 530 DEG and 75 enriched GO terms comparing *P. kingsleyae* (N-type) vs *P. kingsleyae* (P-type) (waveform complexity).

Set B: Expression Based Clustering

For each EOD feature (n=3), we grouped OTUs by phenotype (Table 1), and calculated normalized expression values for Set A genes (n = 3,274). For each EOD feature, we selected genes exhibiting a greater than four-fold difference in averaged, normalized expression between phenotypes to construct Set B. The expression profiles of the genes in the clusters for each EOD feature, along with the enriched GO terms for Biological Process and Cellular Component, are shown in Figs. 4-6. Supplemental File 4 lists the identities of these DEG and Supplemental File 5 lists their enriched GO terms for all three GO ontologies.

Contrast of waveform duration identified 181 DEG and 40 enriched GO terms. 121 of the DEG were upregulated in the short EOD phenotype (Fig. 4A, purple lines). These genes were enriched with GO terms that include 'extracellular space,' 'extracellular region,' 'muscle contraction,' and 'phosphatidylinositol phosphorylation' (Fig. 4B, purple bars), whereas 61 genes were upregulated in samples with long EODs (Fig. 4A, yellow lines). These genes were enriched with GO terms like 'negative regulation of cation transmembrane transport,' 'regulation

of voltage-gated calcium channel activity,' and 'neuropeptide hormone activity' (Fig. 4B, yellow bars).

Contrast of waveform polarity identified 147 DEG and 35 enriched GO terms. We found 40 upregulated genes (Fig. 5A, grey lines) in individuals with small penetrations. These genes were enriched with GO terms such as 'response to mechanical stimulus,' 'extracellular matrix organization,' and 'collagen trimer' (Fig. 5B, grey bars). In large penetrations phenotype, we found 107 upregulated genes (Fig. 5A, salmon lines). These genes were enriched with GO terms that included 'sarcomere organization,' 'plasma membrane bounded cell projection morphogenesis,' 'calcium ion binding,' and 'structural constituent of cytoskeleton' (Fig. 5B, salmon bars).

Finally, contrast of waveform complexity identified 174 DEG and 16 enriched GO terms. We detected 82 upregulated genes in individuals with biphasic EODs (Fig. 6A, blue lines). These genes were enriched with GO terms like 'positive regulation of canonical Wnt signaling pathway,' 'mucopolysaccharide metabolic process,' and 'integral component of membrane' (Fig. 6B, blue bars). We detected 92 upregulated genes in individuals with triphasic EODs (Fig. 6A, orange lines). These genes were enriched with GO terms that included 'cell-matrix adhesion,' 'regulation of ion transmembrane transport,' and 'neuronal cell body' (Fig. 6B, orange bars).

Set C: Intersection of Phylogenetically Informative Comparisons and Expression Based Clustering

We were motivated to obtain the DEGs and enriched GO terms that were most likely to be associated with divergent EOD phenotypes. To obtain this list, we constructed Set C, which is the intersection of Set A' and Set B described above.

253 Contrast of waveform duration identified 105 DEG and 11 enriched GO terms. 79 of the
254 DEG were upregulated in the short EOD phenotype, and 26 genes were upregulated in samples
255 with long EODs. Contrast of waveform polarity identified 144 DEG and 9 enriched GO terms.
256 We found 38 upregulated genes in individuals with small penetrations and 106 upregulated genes
257 in individuals with large penetrations. Finally, contrast of waveform complexity identified 71
258 DEG and 4 enriched GO terms. We detected 46 upregulated genes in individuals with biphasic
259 EODs and 25 upregulated genes in individuals with triphasic EODs. These results are further
260 detailed in Table 3. The enriched GO terms in Set C for Biological Process and Cellular
261 Component are emphasized in boldface for waveform duration (Fig. 4B), polarity (Fig. 5B), and
262 complexity (Fig. 6B). The individual genes in Set C are listed in Supplemental File 6 and their
263 associated GO terms are listed in Supplemental File 7.

264 **Discussion**

265 It has long been recognized that changes in gene expression can affect phenotypic
266 differences between species [18], and RNA-seq has facilitated the study of this relationship [19].
267 The goal of this study was to determine DEGs associated with divergent EOD features within
268 *Paramormyrops*. Expression patterns across all OTUs were highly correlated (Pearson's $r > 0.89$,
269 Fig. 3) and we detected differential expression of only 3,274 (16%) genes between any two
270 OTUs. Thus, a major finding of this study is that EO gene expression is overall quite similar
271 across *Paramormyrops* species with divergent EODs, and relatively few genes are associated
272 with phenotypic differences in EOD waveform between OTUs. Given generally high levels of
273 genetic distances observed between geographically proximate populations of these
274 *Paramormyrops* species [35,75] this was a somewhat unexpected finding.

Despite the relatively small number of DEGs compared to the total number of genes expressed in the EO, we constructed our analysis to extract genes that were highly associated with particular phenotypes. Set A' represents a formal statistical test that contrasted OTUs. Each comparison contrasted samples from OTUs that were divergent in only one EOD feature, while minimizing phylogenetic distance. The tradeoff of this approach is small sample sizes without biological replication and potentially confounding variables, such as collection sites. Set B followed the opposite approach by minimizing the problem of biological replication at the expense of confounding phylogenetic relatedness and phenotypic heterogeneity. To mitigate this, we constructed Set C, which represents genes and GO terms that are differentially expressed/enriched between closely related OTUs divergent in only one phenotypic character and that are also consistently differentially expressed/enriched among representatives with similar EOD phenotypes. As such, we focus our discussion on the results of Set C. We classified the genes in Set C into "general" functional classes, or themes; and focus our attention on the ones that relate to the known morphological underpinnings of waveform duration (Table 4), polarity (Table 5), and complexity (Table 6). These functional classes were genes related to the ECM, cation homeostasis, lipid metabolism, and cytoskeletal and sarcomeric genes.

Waveform Duration

Several researchers have implicated the role of ion channels in the evolution of duration changes in mormyrid signals [53–57]. We did not find evidence of large changes in expression of potassium or sodium channels between short-duration *P.* sp. 'SN3' and other *Paramormyrops* species. We note that, while no differential expression of potassium or sodium channels was discovered, the GO term 'regulation of voltage-gated calcium channel activity' was enriched in long EOD phenotypes, although there were no genes annotated to this GO term common to the

298 lists of DEGs from Sets A' and B. In particular, individuals with short EODs upregulate two
299 calcium-binding proteins: *parvalbumin-2*, and *parvalbumin-2-like*. Parvalbumins are highly
300 expressed in skeletal muscle where they sequester calcium after contraction, thus facilitating
301 relaxation. Frequently, muscles with fast relaxation rates express higher levels of parvalbumins
302 [76]. The upregulated parvalbumin genes we detected may somehow be related to shorter EODs
303 by sequestering calcium at a faster rate, which could affect action potentials directly or indirectly
304 through calcium-activated ion channels.

305 Previous studies have demonstrated that changes in EOD duration result from changes in
306 electrocyte ultrastructure. The two major phases of the EOD waveform are caused by action
307 potentials generated by the anterior and posterior faces [47]. Bennett [48] demonstrated a
308 relationship between EOD duration and increased surface membrane area, and Bass et al. [50]
309 showed that differences in surface area are more readily noticeable on the anterior face.
310 Membrane surface area is increased by folding the electrocyte membrane into papillae and other
311 tube-like invaginations [77]. Testosterone can induce increases in EOD duration in several
312 mormyrids [49,50,78,79], and it also increases membrane surface area, either particularly on the
313 anterior face [50] or on both anterior and posterior faces [80]. A larger surface area may increase
314 the capacitance of the membrane, thus delaying spike initiation [49,50]. Consequently, genes
315 involved in the synthesis of membranes could influence EOD duration.

316 We found the most prominent differences in gene expression between the EOD duration
317 phenotypes in genes that code for cytoskeletal, sarcomeric, and lipid metabolism proteins (Table
318 4). We emphasize the last group: no lipid metabolism genes were upregulated in individuals with
319 long EODs, whereas samples with short EODs upregulated *protein EFR3 homolog B-like* (a
320 regulator of phosphatidylinositol 4-phosphate synthesis), *retinoic acid receptor responder*

protein 3-like (hydrolysis of phosphatidylcholines and phosphatidylethanolamines), *PTB domain-containing engulfment adapter protein 1-like* (modulates cellular glycosphingolipid and cholesterol transport), *phosphatidylinositol 3-kinase regulatory subunit gamma-like*, (PI3K, which phosphorylates phosphatidylinositol), and *proto-oncogene c-Fos-like* (can activate phospholipid synthesis), and showed enrichment of the GO term 'phosphatidylinositol phosphorylation.' We hypothesize that these genes are involved in the surface proliferation of the electrocytes membranes.

Additionally, each mormyrid electrocyte stands embedded in a gelatinous mucopolysaccharide matrix (the ECM) separated from neighboring electrocytes by connective tissue septa (Fig. 1) [34], and the membrane surface invaginations are coated by the same ECM that surrounds the electrocytes [50,77]. Hence, differences in surface invaginations could also be reflected in differences in the expression of genes whose products interact with the ECM. We detected none of these genes upregulated in individuals with long EODs, whereas those with short EODs upregulated three: *hyaluronidase-5-like* (breaks down hyaluronan), *collagenase 3-like* (plays a role in the degradation of ECM proteins), and *fibroblast growth factor binding protein 1* (acts as a carrier protein that releases fibroblast-binding factors from the ECM storage).

Overall, our results identify genes that may affect EOD duration through membrane rearrangements, which could be coupled with changes in the interaction with the ECM and the expression of cytoskeletal and sarcomeric genes. Since this waveform feature is modulated by testosterone, this androgen could facilitate the study of these suggested genetic underpinnings under more rigorously controlled circumstances.

342 *Waveform Polarity*

343 The number [47] and diameter [34,43] of stalk penetrations are positively correlated with
344 the magnitude of P0. This phenomenon is exemplified by *P.* sp. 'magnostipes type I', which has
345 the largest P0 in the OTUs examined in this study, giving the EOD the appearance that it
346 'inverted' relative to other EODs. This OTU has numerous, large diameter penetrations, whereas
347 *P. kingsleyae* (P-type) has relatively fewer, small diameter penetrations (Fig. 1). These large
348 structural differences may influence the electrocyte's connection with the surrounding ECM, and
349 our results support this: the phenotypes of waveform polarity exhibited differences in the
350 expression of genes that interact with the extracellular space. We found no such genes
351 upregulated in individuals with large penetrations, whereas in samples with small penetrations
352 two GO terms were enriched: 'extracellular matrix organization' and 'response to mechanical
353 stimulus,' and three genes were upregulated: *collagen alpha-1(I) chain-like*, *collagen alpha-1(X)*
354 *chain-like*, and *hyaluronidase-5-like* (breaks down hyaluronan).

355 OTUs with large penetrations also exhibited higher expression of genes related to
356 cytoskeletal, sarcomeric, and lipid metabolism proteins than do individuals with smaller
357 penetrations (Table 5). This includes the GO terms 'heart contraction,' 'sarcomerogenesis,' and
358 'cardiac myofibril assembly,' representative of the genes *myosin XVB*, *heat shock protein HSP*
359 *90-alpha 1* (has a role in myosin expression and assembly), *tubulin beta chain*, *tubulin beta 2A*
360 *class IIa*, *fibroblast growth factor 13* (a microtubule-binding protein which directly binds tubulin
361 and is involved in both polymerization and stabilization of microtubules), *spindle and*
362 *kinetochore associated complex subunit 3* (a component of the kinetochore-microtubule
363 interface), *troponin C, slow skeletal and cardiac muscles*, *protein tilB homolog* (may play a role
364 in dynein arm assembly), and *cysteine and glycine rich protein 3* (codes for the Muscle LIM

Protein (MLP), which is implicated in various cytoskeletal and sarcomeric macromolecular complexes [81–83] and is a positive regulator of myogenesis [84]). In contrast, samples with small penetrations only show upregulation of the genes *myosin light chain 3-like* and *desmin-like.*

We hypothesize that the differences in the number and diameter of penetrations that drive variation in EOD waveform polarity require changes to the electrocyte's cytoskeletal and membrane properties. These arrangements may be necessary for the electrocytes body to adjust to the increased volume displacements imposed by larger penetrations; or alternatively, they may be a prerequisite for penetrating stalks to enlarge. Our observations support and elaborate on the hypothesis that sarcomeric proteins (which are non-contractile in mormyrids) may function as a means of cytoskeletal support and structural integrity in mormyrid electrocytes [85].

Waveform Complexity

Waveform complexity refers to the number of phases present in an EOD, and mormyrid EODs vary in the presence of a small head negative phase (P0). The presence or absence of P0 in the EOD depends on the anatomical configuration of the electrocytes: P0-present (or triphasic) EODs are produced by electrocytes that are innervated on the anterior face and have penetrating stalks (Pa), whereas P0-absent (or biphasic) EODs are produced by electrocytes innervated on the posterior face and lack penetrating stalks (NPp) [42,43,47,48,51,52]. Developmental studies of the adult EO suggest that Pa electrocytes go through a NPp stage before developing penetrations [86,87]. This motivated the hypothesis that penetrations develop by the migration of the posteriorly innervated stalk system (NPp stage) through the edge of the electrocyte, and that the interruption of this migration represents a mechanism for Pa-to-NPp reversals [88,89].

Our data indicates several DEGs that implicate specific cytoskeletal and ECM reorganizations between triphasic and biphasic EODs (Table 6). We observed differential expression of several genes associated with the polymerization of F-actin. In triphasic individuals, we observe upregulation of the gene *capping protein regulator and myosin 1 linker 3* (CARMIL3); although this gene is little studied, its paralog CARMIL2 enhances F-actin polymerization. Also upregulated is the gene *cysteine and glycine rich protein 3* (MLP protein, see Waveform Polarity). In contrast, the biphasic phenotype upregulated the genes *protein-methionine sulfoxide oxidase mical2b-like* (promotes F-actin depolymerization), *transmembrane protein 47-like* (may regulate F-actin polymerization), 5'-*AMP-activated protein kinase subunit gamma-2-like* (could remodel the actin cytoskeleton), and *FYVE, RhoGEF and PH domain-containing protein 4-like* (regulates the actin cytoskeleton). Thus, biphasic and triphasic EODs display several DEG, with potentially diverging outcomes, that influence the cellular internal structure.

We hypothesize that electrocytes with penetrating stalks (which produce triphasic EODs) require cytoskeletal arrangements to produce penetrations, perhaps related to increasing F-actin, to maintain their structural integrity. Similar to what we propose under waveform polarity, these arrangements may be necessary for the electrocyte body to adjust to the penetrations; or alternatively, they may be a prerequisite for penetrations to occur.

We also observed differential expression in a number of proteins expressed in the ECM. In biphasic OTUs, we found the GO term ‘mucopolysaccharide metabolic process’ to be enriched, and two upregulated copies of the gene *inter-alpha-trypsin inhibitor heavy chain 3*, which may act as a binding protein between hyaluronan and other ECM proteins. In triphasic individuals, we found the GO term ‘cell-matrix adhesion’ enriched, and the upregulated genes

410 *epiphycan-like*, which may play a role in cartilage matrix organization, and two copies of
411 *ependymin-like* (these paralogs are ortholog to the zebrafish ependymin-like gene *epdl2*).

412 The two ependymin-like genes are among the most differentially expressed genes
413 between biphasic and triphasic OTUs (500-fold more highly expressed in triphasic individuals in
414 the comparison from Set A', Supplemental File 2). Although expressed in many tissues and with
415 little amino acid similarity, all ependymin-related proteins are secretory, calcium-binding
416 glycoproteins that can undergo conformational changes and associate with collagen in the ECM.
417 They have been involved in regeneration, nerve growth, cell contact, adhesion and migration
418 processes [90]. We hypothesize that ependymin-related proteins, and potentially some of the
419 other ECM proteins highly expressed in triphasic individuals, are part of the "fibrillar substance"
420 that lies between the stalk and the electrocyte body in individuals with penetrating electrocytes
421 [50]. Notably, the *P. kingsleyae* genome assembly, which is based on a biphasic individual,
422 contains three paralogs of *epdl2*, whereas the osteoglossiform *Scleropages formosus* only has
423 one, suggesting the intriguing possibility that this gene may have been duplicated in
424 *Paramormyrops* or in mormyrids. Ependymin-related paralogs have been proposed as suitable
425 targets to experimentally test gene subfunctionalization [91].

426 Altogether, our results for EOD waveform complexity suggest that the conformation of
427 the cytoskeleton and the expression of proteins secreted to the ECM are important elements of
428 the stalk penetrations, which generate triphasic EODs.

429 *Concluding Remarks*

430 Two previous studies focused on DGE between EOs in another mormyrid species
431 adaptive radiation/explosive diversification (genus *Campylomormyrus*). Both focused on

comparisons between the species *C. tshokwe* (long duration) and *C. compressirostris* (short duration). The first study performed a canditate gene approach to quantify the expression patterns of 18 sodium and potassium homeostasis genes between the EOs of the two species [55], whereas Lamanna et al. [92] used RNA-seq to simultaneously compare gene expression between EOs of these species. While we did not observe differences in expression of any of the potassium channels reported by Nagel et al. (2017), we note that Lamanna et al. (2015) reported differential expression of metabolic pathways related genes, particularly fatty acid metabolism, and ion transport and neuronal function (subcluster 4). While we found no overlap in the identities of any specific genes in our study, we note that our analysis also detected differential expression of lipid metabolism related genes when comparing EODs of different duration.

The widespread differential expression within *Paramormyrops* of calcium-related genes (Supplemental File 6) emphasizes a much-needed area of future research. Calcium is known to be necessary for the proper electrocyte repolarization in some gymnotiform species [93], but it may not be as important in others [94]. Few studies have addressed calcium physiology in mormyrids: calcium-related proteins have been reported as differentially expressed in EO vs skeletal muscle in *Campylomormyrus* [92] and in *Brienomyrus brachyistius* [85]. As electrocytes do not contract, calcium may act in electrocytes as an important second messenger or cofactor, participate in interactions with the ECM, and/or to contribute to the electrocyte's electrical properties through interaction with voltage gated ion channels.

A second notable pattern in our results is the unusual degree to which mormyrid electrocytes retain expression of some sarcomeric genes, which has been noted in several studies [58,85,92,95,96]. The role these proteins serves in electrocytes is presently unknown; however, results indicate that they are highly differentially expressed between *Parmormyrops* with

different EOD waveforms. This strongly suggests that sarcomeric proteins could play an important role in the conformational changes required to develop and sustain penetrations.

Finally, the biochemical composition and function of the ECM in electrocytes is poorly understood. Our analysis identifies differential expression in ECM-related genes across the *Paramormyrops*, associated with each of the three EOD features studied. At least two of these genes (*inter-alpha-trypsin inhibitor heavy chain 3* and *hyaluronidase-5-like*), distributed across all three EOD features, interact with hyaluronan. Hyaluronan is a type of mucopolysaccharide and a major component of some soft tissues and fluids [97]. Therefore, we propose that hyaluronan is an important constituent of the ECM in mormyrid fish. In addition, the electrocyte-ECM interactions should be an important area of future investigation, as they are likely to influence electrocyte shape, electrical properties, and potentially the morphology of penetrations and surface membrane invaginations.

To conclude, this study examined the expression correlates of a hyper-variable phenotype in a rapidly diversified genus of mormyrid electric fish. We examined DGE between taxa exhibiting variability along three major axes of variation that characterize EOD differences within *Paramormyrops* and among mormyrids: duration, polarity, and complexity. We found that gene expression in EOs among closely related species is largely similar, but patterns of DGE between EOs is primarily restricted to four broad functional sets: (1) cytoskeletal and sarcomeric proteins, (2) cation homeostasis, (3) lipid metabolism and (4) proteins that interact with the ECM. Our results suggest specific candidate genes that are likely to influence the size, shape and architecture of electrocytes for future research on gene function and molecular pathways that underlie EOD variation in mormyrid electric fish.

Competing interests

The authors declare they have no competing interests.

Ethics approval and consent to participate

All research protocols involving live fish were approved by the Michigan State University and Cornell University Institutional Animal Care and Use Committees.

Acknowledgements

This work was supported by the National Science Foundation (1455405: JRG, PI) and a grant from the Cornell University Center for Vertebrate Genomics. The authors thank Bruce Carlson, Matt Arnegard, Carl Hopkins, Roger Afene, Jean Danielle Mbega and Marie-Francois Eva for assistance with specimen collection in Gabon. The authors also acknowledge the Michigan State University Institute for Cyber-Enable Research for use of their high-performance computing infastructure, and CENREST Gabon for logistical support and collection permits.

Author Contributions

JRG designed the experiment, collected and identified specimens, performed library construction and sequencing, and oversaw data analysis. ML performed QC analysis, designed and implemented the data analysis procedure, and performed examination of gene ontology and gene function. Both authors contributed to writing the manuscript.

Supplementary Files

Supplementary File 1. Raw reads NCBI SRA accession numbers, number of reads and alignment rates per sample, using bowtie 2 as the aligner and the Paramormyrops kingsleyae (N-type) genome as the reference.

498 **Supplementary File 2.** DEG per comparison from the 10 pairwise DGE analysis. Positive
499 values under logFC indicate genes upregulated in the OTU under sampleA, whereas negative
500 values correspond to genes upregulated in the OTU under sampleB. Values under each sample
501 are gene raw counts. Significance threshold was abs(log(base2)FC) > 2 (= 4-fold expression
502 difference) and FDR <0.001.

503 **Supplementary File 3.** Enriched GO terms per comparison, ontology and OTU in the DEG from
504 the 10 pairwise comparisons. Also listed are the DEG annotated to each GO term. The pvalue is
505 in the column weight01.

506 **Supplementary File 4.** DEG per EOD feature and phenotype identified with the Set B analysis.
507 Values under each sample are TMM normalized, log2(TMM +1) transformed, and mean-
508 centered expression values.

509 **Supplementary File 5.** Enriched GO terms per EOD feature, ontology and phenotype in the
510 DEG from the Set B analysis. Also listed are the DEG annotated to each GO term. The pvalue is
511 in the column weight01.

512 **Supplementary File 6.** DEG in Set C, per EOD feature and phenotype.

513 **Supplementary File 7.** GO terms enriched in the DEG in Set C, per EOD feature, ontology and
514 phenotype. Also listed are the DEG annotated to each GO term, and the quickGO definitions of
515 each GO term.

516 **References**

517

518 1. Butlin R, Debelle A, Kerth C, Snook RR, Beukeboom LW, Castillo Cajas RF, et al. What do
519 we need to know about speciation? Trends Ecol Evol. 2012;27:27–39.

520 2. Givnish TJ. Adaptive radiation versus “radiation” and “explosive diversification”: Why
521 conceptual distinctions are fundamental to understanding evolution. New Phytol. 2015;207:297–
522 303.

523 3. Grant RB, Grant PR. What Darwin’s Finches Can Teach Us about the Evolutionary Origin and
524 Regulation of Biodiversity. Bioscience. 2003;53:965–75.

525 4. Kocher TD. Adaptive evolution and explosive speciation: The cichlid fish model. Nat Rev
526 Genet. 2004;5:288–98.

527 5. Shaw KL. Further acoustic diversity in Hawaiian forests: two new species of Hawaiian cricket
528 (Orthoptera: Gryllidae: Trigonidiinae: Laupala). Zool J Linn Soc. 2000;129:73–91.

529 6. Mullen SP, Mendelson TC, Schal C, Shaw KL. Rapid evolution of cuticular hydrocarbons in a
530 species radiation of acoustically diverse Hawaiian crickets (Gryllidae: Trigonidiinae: Laupala).
531 Evolution (N Y). 2007;61:223–31.

532 7. Seehausen O. African cichlid fish: A model system in adaptive radiation research. Proc R Soc
533 B Biol Sci. 2006;273:1987–98.

534 8. Hulsey CD. Cichlid genomics and phenotypic diversity in a comparative context. Integr Comp
535 Biol. 2009;49:618–29.

536 9. Han F, Lamichhaney S, Grant BR, Grant PR, Andersson L, Webster MT. Gene flow, ancient
537 polymorphism, and ecological adaptation shape the genomic landscape of divergence among
538 Darwin's finches. Genome Res. 2017;27:1004–15.

539 10. Lamichhaney S, Berglund J, Almén MS, Maqbool K, Grabherr M, Martinez-Barrio A, et al.
540 Evolution of Darwin's finches and their beaks revealed by genome sequencing. Nature. Nature
541 Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.;
542 2015;518:371–5.

543 11. Lamichhaney S, Han F, Berglund J, Wang C, Almén MS, Webster MT, et al. A beak size
544 locus in Darwin's finches facilitated character displacement during a drought. Science.
545 2016;352:470–4.

546 12. O'Quin CT, Drilea AC, Roberts RB, Kocher TD. A Small Number of Genes Underlie Male
547 Pigmentation Traits in Lake Malawi Cichlid Fishes. J Exp Zool Part B Mol Dev Evol.
548 2012;318:199–208.

549 13. Magalhaes IS, Seehausen O. Genetics of male nuptial colour divergence between sympatric
550 sister species of a Lake Victoria cichlid fish. J Evol Biol. John Wiley & Sons, Ltd (10.1111);
551 2010;23:914–24.

552 14. Shaw KL, Lesnick SC. Genomic linkage of male song and female acoustic preference QTL
553 underlying a rapid species radiation. Proc Natl Acad Sci. National Academy of Sciences;
554 2009;106:9737–42.

555 15. Wiley C, Ellison CK, Shaw KL. Widespread genetic linkage of mating signals and
556 preferences in the Hawaiian cricket Laupala. Proc R Soc B Biol Sci. The Royal Society;

557 2012;279:1203–9.

558 16. Blankers T, Oh KP, Bombarely A, Shaw KL. The Genomic Architecture of a Rapid Island
559 Radiation: Recombination Rate Variation, Chromosome Structure, and Genome Assembly of the
560 Hawaiian Cricket Laupala. Genetics. Genetics; 2018;209:1329–44.

561 17. Ding B, Daugherty DW, Husemann M, Chen M, Howe AE, Danley PD. Quantitative genetic
562 analyses of male color pattern and female mate choice in a pair of cichlid fishes of Lake Malawi,
563 East Africa. PLoS One. 2014;9:1–22.

564 18. King MC, Wilson AC. Evolution at two levels in humans and chimpanzees. Science.
565 1975;188:107–16.

566 19. Alvarez M, Schrey AW, Richards CL. Ten years of transcriptomics in wild populations:
567 What have we learned about their ecology and evolution? Mol Ecol. 2015;24:710–25.

568 20. Brawand D, Soumillon M, Necsulea A, Julien P, Csárdi G, Harrigan P, et al. The evolution
569 of gene expression levels in mammalian organs. Nature. 2011;478:343–8.

570 21. Ferreira PG, Patalano S, Chauhan R, Ffrench-Constant R, Gabaldón T, Guigó R, et al.
571 Transcriptome analyses of primitively eusocial wasps reveal novel insights into the evolution of
572 sociality and the origin of alternative phenotypes. Genome Biol. 2013;14:R20.

573 22. Abolins-Abols M, Kornobis E, Ribeca P, Wakamatsu K, Peterson MP, Ketterson ED, et al.
574 Differential gene regulation underlies variation in melanic plumage coloration in the dark-eyed
575 junco (Junco hyemalis). Mol Ecol. 2018;4501–15.

576 23. Härer A, Meyer A, Torres-Dowdall J. Convergent phenotypic evolution of the visual system
577 via different molecular routes: How Neotropical cichlid fishes adapt to novel light environments.

578 Evol Lett. 2018;341–54.

579 24. Nigenda-Morales SF, Hu Y, Beasley JC, Ruiz-Piña HA, Valenzuela-Galván D, Wayne RK.
580 Transcriptomic analysis of skin pigmentation variation in the Virginia opossum (Didelphis
581 virginiana). Mol Ecol. 2018;27:2680–97.

582 25. Young RL, Ferkin MH, Ockendon-Powell NF, Orr VN, Phelps SM, Pogány Á, et al.
583 Conserved transcriptomic profiles underpin monogamy across vertebrates. Proc Natl Acad Sci.
584 2019;116:201813775.

585 26. Jeukens J, Renaut S, St-Cyr J, Nolte AW, Bernatchez L. The transcriptomics of sympatric
586 dwarf and normal lake whitefish (Coregonus clupeaformis spp., Salmonidae) divergence as
587 revealed by next-generation sequencing. Mol Ecol. 2010;19:5389–403.

588 27. Giorello FM, Feijoo M, D'Elía G, Naya DE, Valdez L, Opazo JC, et al. An association
589 between differential expression and genetic divergence in the Patagonian olive mouse (Abrothrix
590 olivacea). Mol Ecol. 2018;27:3274–86.

591 28. Carlson BA, Hasan SM, Hollmann M, Miller DB, Harmon LJ, Arnegard ME. Brain
592 evolution triggers increased diversification of electric fishes. Science. 2011;332:583–6.

593 29. Rabosky DL, Santini F, Eastman J, Smith SA, Sidlauskas B, Chang J, et al. Rates of
594 speciation and morphological evolution are correlated across the largest vertebrate radiation. Nat
595 Commun. 2013;4:1–8.

596 30. Sullivan JP, Lavoué S, Hopkins CD. Discovery and phylogenetic analysis of a riverine
597 species flock of African electric fishes (Mormyridae: Teleostei). Evolution. 2002;56:597–616.

598 31. Sullivan JP, Lavoué S, Arnegard ME, Hopkins CD. AFLPs resolve phylogeny and reveal

599 mitochondrial introgression within a species flock of African electric fish (Mormyroidea:
600 Teleostei). Evolution. 2004;58:825–41.

601 32. Lavoué S, Arnegard ME, Sullivan JP, Hopkins CD. Petrocephalus of Odzala offer insights
602 into evolutionary patterns of signal diversification in the Mormyridae, a family of weakly
603 electrogenic fishes from Africa. J Physiol Paris. 2008;102:322–39.

604 33. Hopkins CD. On the diversity of electric signals in a community of mormyrid electric fish in
605 West Africa. Am Zool. Oxford University Press; 1981;21:211–22.

606 34. Bass AH. Species differences in electric organs of mormyrids: Substrates for species-typical
607 electric organ discharge waveforms. J Comp Neurol. 1986;244:313–30.

608 35. Arnegard ME, Bogdanowicz SM, Hopkins CD. Multiple cases of striking genetic similarity
609 between alternate electric fish signal morphs in sympatry. Evolution. 2005;59:324–43.

610 36. Sullivan JP, Lavoué S, Hopkins CD. Molecular systematics of the African electric fishes
611 (Mormyroidea: teleostei) and a model for the evolution of their electric organs. J Exp Biol.
612 2000;203:665–83.

613 37. Arnegard ME, McIntyre PB, Harmon LJ, Zelditch ML, Crampton WGR, Davis JK, et al.
614 Sexual signal evolution outpaces ecological divergence during electric fish species radiation. Am
615 Nat. 2010;176:335–56.

616 38. Lissmann HW, Machin KE. The mechanism of object location in Gymnarchus niloticus and
617 similar fish. J Exp Biol. 1958;35:451–486.

618 39. von der Emde G, Amey M, Engelmann J, Fetz S, Folde C, Hollmann M, et al. Active
619 electrolocation in Gnathonemus petersii: Behaviour, sensory performance, and receptor systems.

620 J Physiol - Paris. Elsevier Ltd; 2008;102:279–90.

621 40. Kramer B. Electric organ discharge interaction during interspecific agonistic behaviour in
622 freely swimming mormyrid fish. J Comp Physiol. Springer-Verlag; 1974;93:203–35.

623 41. Möhres FP. Elektrische Entladungen im Dienste der Revierabgrenzung bei Fischen.
624 Naturwissenschaften. Springer-Verlag; 1957;44:431–2.

625 42. Hopkins CD. Signal evolution in electric communication. In: Hauser M, Konishi M, editors.
626 Des Anim Commun. Cambridge, MA: MIT Press; 1999. p. 461– 491.

627 43. Gallant JR, Arnegard ME, Sullivan JP, Carlson BA, Hopkins CD. Signal variation and its
628 morphological correlates in Paramormyrops kingsleyae provide insight into the evolution of
629 electrogenic signal diversity in mormyrid electric fish. J Comp Physiol A. 2011;197:799–817.

630 44. Carlson BA. Electric signaling behavior and the mechanisms of electric organ discharge
631 production in mormyrid fish. J Physiol. 2002;96:405–19.

632 45. Caputi AA, Carlson BA, Macadar O. Electric Organs and Their Control. In: Bullock TH,
633 Hopkins C.D., Popper A.N., Fay R.R., editors. Electroreception. Springer H. New York, NY:
634 Springer Handbook of Auditory Research, vol 21; 2005. p. 410–51.

635 46. Hopkins CD. Behavior of Mormyridae. In: Bullock TH, Heiligenberg W, editors.
636 Electroreception. New York: Wiley; 1986. p. 527–76.

637 47. Bennett MVL, Grundfest H. Studies on the morphology and electrophysiology of electric
638 organs. III. Electrophysiology of electric organs in mormyrids. In: Chagas C, de Carvalho A,
639 editors. Bioelectrogenesis. Amsterdam: Elsevier; 1961. p. 113–35.

640 48. Bennett MVL. Electric Organs. In: Hoar WS, Randall DJ, editors. Fish Physiol. London:
641 Academic Press; 1971. p. 347–491.

642 49. Bass AH, Volman SF. From behavior to membranes: testosterone-induced changes in action
643 potential duration in electric organs. Proc Natl Acad Sci. 1987;84:9295–8.

644 50. Bass AH, Denizot J-P, Marchaterre MA. Ultrastructural features and hormone-dependent sex
645 differences of mormyrid electric organs. J Comp Neurol. 1986;254:511–28.

646 51. Szabo T. Les Organes Electriques des Mormyrides. In: Chagas C, de Carvalho A, editors.
647 Bioelectrogenesis. New York: Elsevier; 1961. p. 20–4.

648 52. Carlson BA, Gallant JR. From sequence to spike to spark: evo-devo-neuroethology of
649 electric communication in mormyrid fishes. J Neurogenet. 2013;27:106–29.

650 53. Zakon HH, Lu Y, Zwickl D, Hillis DM. Sodium channel genes and the evolution of diversity
651 in communication signals of electric fishes: convergent molecular evolution. Proc Natl Acad Sci
652 U S A. 2006;103:3675–80.

653 54. Arnegard ME, Zwickl DJ, Lu Y, Zakon HH. Old gene duplication facilitates origin and
654 diversification of an innovative communication system--twice. Proc Natl Acad Sci U S A.
655 2010;107:22172–7.

656 55. Nagel R, Kirschbaum F, Tiedemann R. Electric organ discharge diversification in mormyrid
657 weakly electric fish is associated with differential expression of voltage-gated ion channel genes.
658 J Comp Physiol A Neuroethol Sensory, Neural, Behav Physiol. Springer Berlin Heidelberg;
659 2017;203:183–95.

660 56. Swapna I, Ghezzi A, York JM, Markham MR, Halling DB, Lu Y, et al. Electrostatic Tuning

661 of a Potassium Channel in Electric Fish. Curr Biol. Cell Press; 2018;28:2094–2102.e5.

662 57. Paul C, Kirschbaum F, Mamonekene V, Tiedemann R. Evidence for Non-neutral Evolution
663 in a Sodium Channel Gene in African Weakly Electric Fish (Campylomormyrus, Mormyridae). J
664 Mol Evol. Springer New York LLC; 2016;83:61–77.

665 58. Gallant JR, Losilla M, Tomlinson C, Warren WC. The Genome and Adult Somatic
666 Transcriptome of the Mormyrid Electric Fish Paramormyrops kingsleyae. Genome Biol Evol.
667 Oxford University Press; 2017;9:3525–30.

668 59. Arnegard ME, Jackson BS, Hopkins CD. Time-domain signal divergence and discrimination
669 without receptor modification in sympatric morphs of electric fishes. J Exp Biol.
670 2006;209:2182–98.

671 60. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence
672 data. Bioinformatics. 2014;30:2114–20.

673 61. MacManes MD. On the optimal trimming of high-throughput mRNA sequence data. Front
674 Genet. 2014;5:1–7.

675 62. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. Nature
676 Publishing Group; 2012;9:357–9.

677 63. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or
678 without a reference genome. BMC Bioinformatics. BioMed Central; 2011;12:323.

679 64. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length
680 transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. Nature
681 Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.;

682 2011;29:644–52.

683 65. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo
684 transcript sequence reconstruction from RNA-seq using the Trinity platform for reference
685 generation and analysis. Nat Protoc. Nature Publishing Group, a division of Macmillan
686 Publishers Limited. All Rights Reserved.; 2013;8:1494–512.

687 66. Robinson MD, Oshlack A. A scaling normalization method for differential expression
688 analysis of RNA-seq data. Genome Biol. 2010;11:R25.

689 67. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology:
690 tool for the unification of biology. Nat Genet. 2000;25:25–9.

691 68. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing
692 strong. Nucleic Acids Res. 2019;47:D330–8.

693 69. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+:
694 architecture and applications. BMC Bioinformatics. BioMed Central; 2009;10:421.

695 70. Wu C, MacLeod I, Su AI. BioGPS and MyGene.info: Organizing online, gene-centric
696 information. Nucleic Acids Res. Oxford University Press; 2013;41:D561–5.

697 71. Xin J, Mark A, Afrasiabi C, Tsueng G, Juchler M, Gopal N, et al. High-performance web
698 services for querying gene and variant annotation. Genome Biol. BioMed Central; 2016;17:91.

699 72. Alexa A, Rahnenführer J, Lengauer T. Improved scoring of functional groups from gene
700 expression data by decorrelating GO graph structure. Bioinformatics. Oxford University Press;
701 2006;22:1600–7.

73. Binns D, Dimmer E, Huntley R, Barrell D, O'Donovan C, Apweiler R. QuickGO: A web-based tool for Gene Ontology searching. Bioinformatics. Oxford University Press; 2009;25:3045–6.

74. The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 2019;47:D506–15.

75. Gallant JR, Sperling J, Cheng C, Arnegard ME, Carlson BA, Hopkins CD. Microevolutionary processes underlying macroevolutionary patterns of electric signal diversity in mormyrid fish. bioRxiv.org. Cold Spring Harbor Laboratory; 2017;154047.

76. Wilwert JL, Madhoun, Nisreen M, Coughlin DJ. Parvalbumin correlates with relaxation rate in the swimming muscle of sheepshead and kingfish. J Exp Biol. 2006;209:227–37.

77. Schwartz IR, Pappas GD, Bennett MVL. The fine structure of electrocytes in weakly electric teleosts. J Neurocytol. 1975;4:87–114.

78. Bass AH, Hopkins CD. Hormonal control of sexual differentiation: changes in electric organ discharge waveform. Science. 1983;220:971–4.

79. Bass AH, Hopkins CD. Hormonal control of sex differences in the electric organ discharge (EOD) of mormyrid fishes. J Comp Physiol A. 1985;156:587–604.

80. Freedman EG, Olyarchuk J, Marchaterre MA, Bass AH. A temporal analysis of testosterone-induced changes in electric organs and electric organ discharges of mormyrid fishes. J Neurobiol. 1989;20:619–34.

81. Vafiadaki E, Arvanitis DA, Papalouka V, Terzis G, Roumeliotis TI, Spengos K, et al. Muscle lim protein isoform negatively regulates striated muscle actin dynamics and differentiation.

723 FEBS J. John Wiley & Sons, Ltd (10.1111); 2014;281:3261–79.

724 82. Papalouka V, Arvanitis DA, Vafiadaki E, Mavroidis M, Papadodima SA, Spiliopoulou CA,
725 et al. Muscle LIM protein interacts with cofilin 2 and regulates F-actin dynamics in cardiac and
726 skeletal muscle. Mol Cell Biol. American Society for Microbiology Journals; 2009;29:6046–58.

727 83. Louis HA, Pino JD, Schmeichel KL, Pomiès P, Beckerle MC. Comparison of three members
728 of the cysteine-rich protein family reveals functional conservation and divergent patterns of gene
729 expression. J Biol Chem. American Society for Biochemistry and Molecular Biology;
730 1997;272:27484–91.

731 84. Arber S, Halder G, Caroni P. Muscle LIM protein, a novel essential regulator of myogenesis,
732 promotes myogenic differentiation. Cell. Elsevier; 1994;79:221–31.

733 85. Gallant JR, Hopkins CD, Deitcher DL. Differential expression of genes and proteins between
734 electric organ and skeletal muscle in the mormyrid electric fish Brienomyrus brachyistius. J Exp
735 Biol. 2012;215:2479–94.

736 86. Szabo T. Development of the Electric Organ of Mormyridae. Nature. Nature Publishing
737 Group; 1960;188:760–2.

738 87. Denizot JP, Kirschbaum F, Westby GWM, Tsuji S. On the development of the adult electric
739 organ in the mormyrid fish Pollimyrus isidori (with special focus on the innervation). J
740 Neurocytol. 1982;11:913–34.

741 88. Alves-Gomes J, Hopkins CD. Molecular insights into the phylogeny of mormyriform fishes
742 and the evolution of their electric organs. Brain Behav Evol. 1997;49:324–50.

743 89. Hopkins CD. Design features for electric communication. J Exp Biol. Co Biol;

744 1999;202:1217–28.

745 90. Ganss B, Hoffmann W. Calcium-induced conformational transition of trout ependymins
746 monitored by tryptophan fluorescence. Open Biochem J. 2009;3:14–7.

747 91. Suárez-Castillo EC, García-Arrarás JE. Molecular evolution of the ependymin protein
748 family: a necessary update. BMC Evol Biol. 2007;7:23.

749 92. Lamanna F, Kirschbaum F, Waurick I, Dieterich C, Tiedemann R. Cross-tissue and cross-
750 species analysis of gene expression in skeletal muscle and electric organ of African weakly-
751 electric fish (Teleostei; Mormyridae). BMC Genomics. BMC Genomics; 2015;16:668.

752 93. Bartels E. Depolarization of electroplax membrane in calcium-free ringer's solution. J
753 Membr Biol. 1971;5:121–32.

754 94. Ferrari MB, Zakon HH. Conductances contributing to the action-potential of Sternopygus
755 electrocytes. J Comp Physiol A-sensory neural Behav Physiol. UNIV TEXAS,CTR DEV
756 BIOL,PATTERSON LAB,AUSTIN,TX 78712: Springer; 1993;173:281–92.

757 95. Gallant JR, Traeger LL, Volkening JD, Moffett H, Chen P-H, Novina CD, et al. Genomic
758 basis for the convergent evolution of electric organs. Science. 2014;344:1522–5.

759 96. Lamanna F, Kirschbaum F, Tiedemann R. De novo assembly and characterization of the
760 skeletal muscle and electric organ transcriptomes of the African weakly electric fish
761 Campylomormyrus compressirostris (Mormyridae, Teleostei). Mol Ecol Resour. 2014;14:1222–
762 30.

763 97. Fraser JRE, Laurent TC, Laurent UBG. Hyaluronan: its nature, distribution, functions and
764 turnover. J Intern Med. John Wiley & Sons, Ltd (10.1111); 1997;242:27–33.

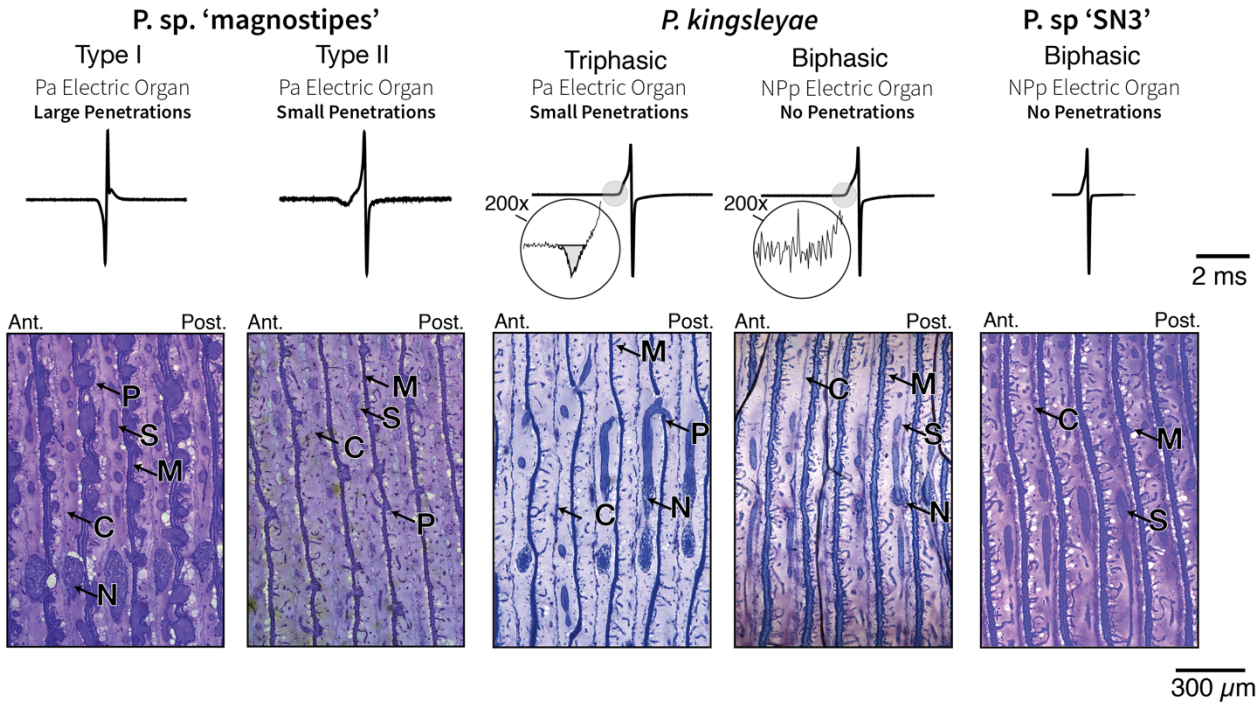

Figure 1. Electric organ discharge (EOD) diversity and electric organ anatomy in *Paramormyrops*. EOD traces from specimens in this study and representative parasagittal sections of the five *Paramormyrops* operational taxonomic units (OTUs) considered in this study. 200x magnification on *P. kingsleyae* EODs reveals a P0 phase on triphasic EODs only. Individuals with triphasic EODs all have penetrations, whereas individuals with biphasic EODs do not. OTUs with 'inverted' polarity triphasic EODs have large penetrations compared to OTUs with normal polarity triphasic EODs. Ant. = anterior, C = connective tissue septa, N = nerve, M = microstalklets (profusely branched stalks), P = penetrations, Post. = posterior, S = stalks.

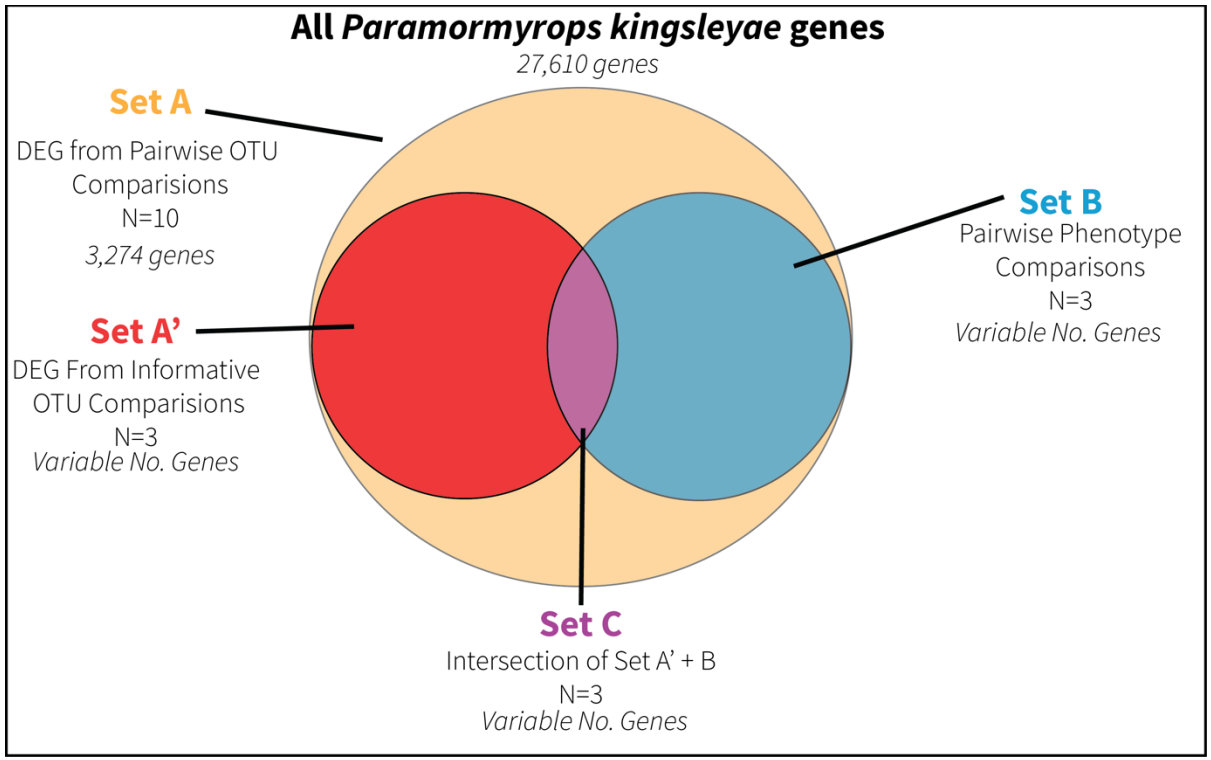

Figure 2. Diagram of how we constructed the lists of upregulated genes of Set C. We used the same approach to find the respective enriched GO terms. DEG = differentially expressed genes, N = number of comparisons made for each set, OTU = operational taxonomic unit.

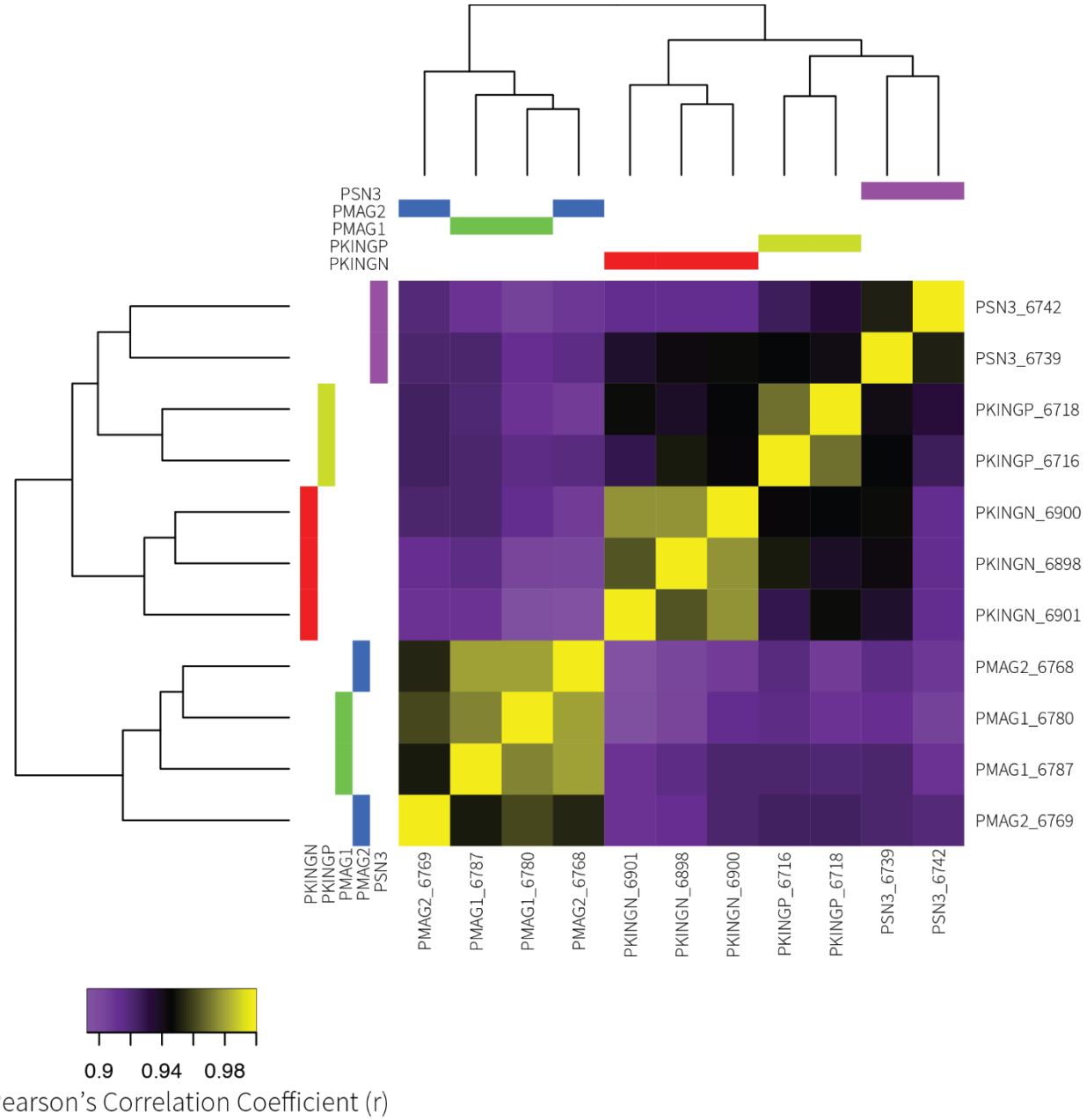

Figure 3. Heatmap of sample by sample correlations in gene expression, and the inferred phylogenetic relationships from these expression correlation values.

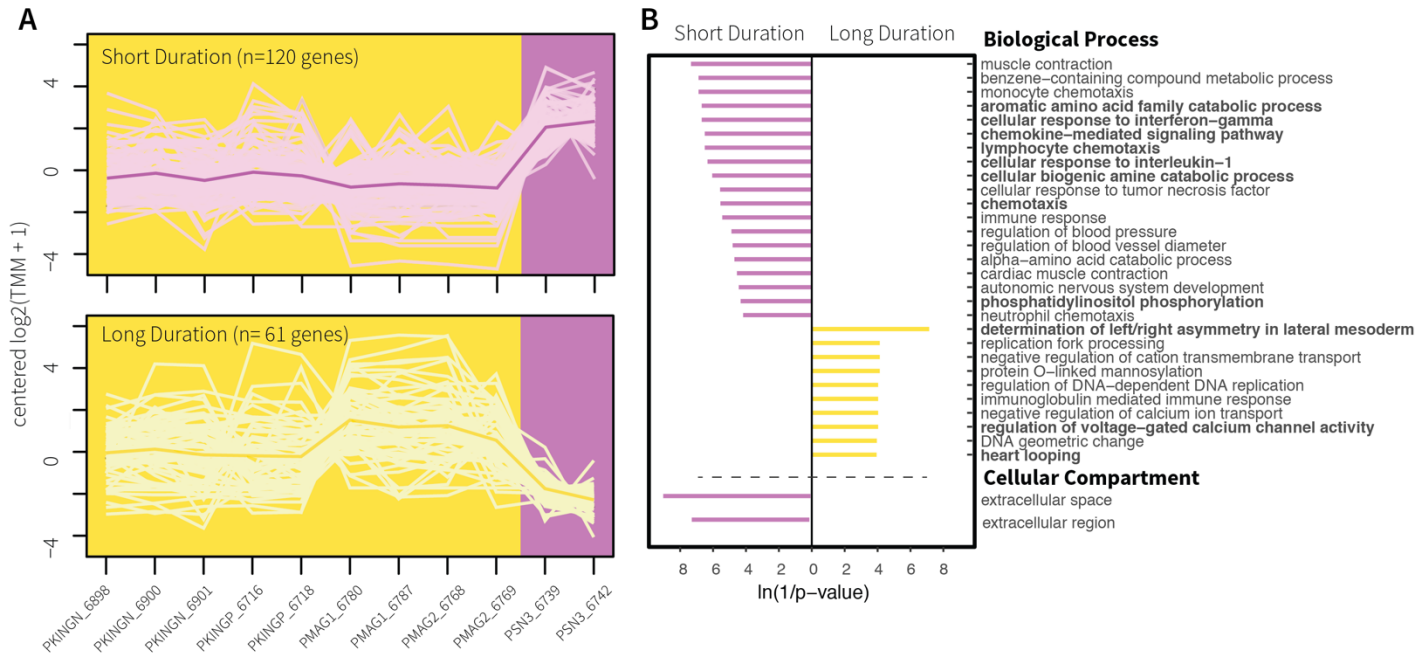

Figure 4. A) Gene clusters whose expression patterns correlate with the electric organ discharge (EOD) duration phenotypes short EODs (purple background) and long EODs (yellow background). Samples are sorted alphabetically on the X axis. The lines connect transformed gene expression values across all samples; light-color lines represent one gene, the dark-color line is the average expression pattern of all genes in the cluster. B) Gene Ontology (GO) terms for Biological Process and Cellular Component found enriched in the gene clusters from (A). The X axis shows transformed p-values, the longer a bar the smaller its p-value. The direction and color of a bar indicate the phenotype in which the GO term is enriched [same color code as (A)]. GO terms highlighted in bold also belong to Set C.

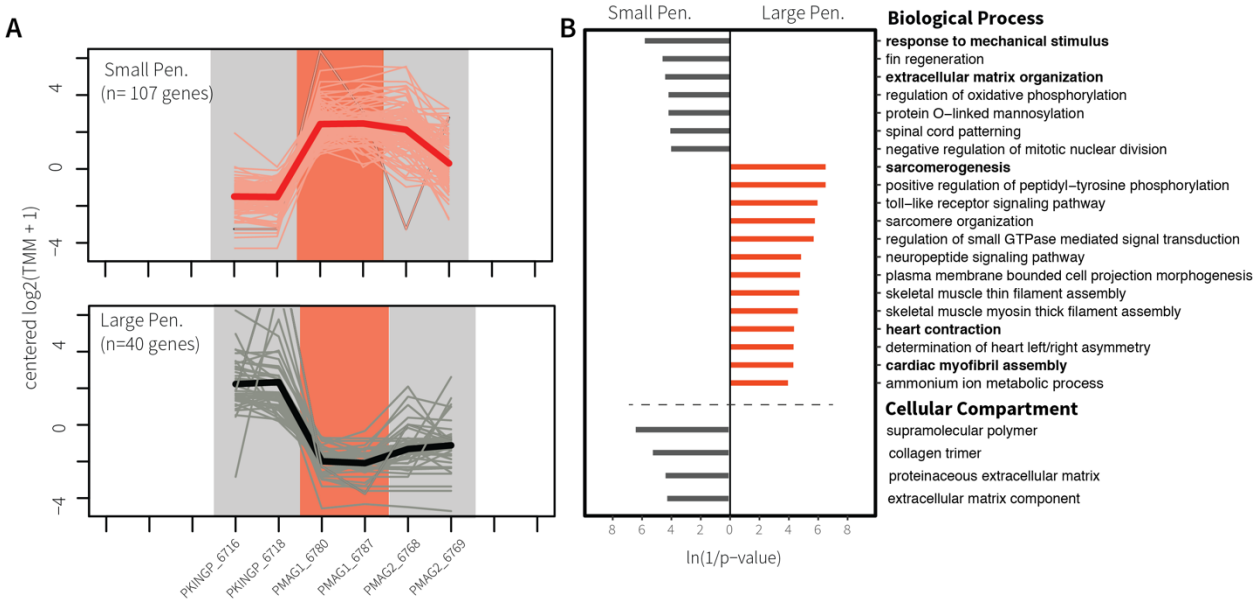

Figure 5. A) Gene clusters whose expression patterns correlate with the electric organ discharge (EOD) waveform polarity phenotypes small penetrations (grey background) and large penetrations (salmon background). Samples are sorted alphabetically on the X axis, but only samples with penetrations are considered for this EOD feature. The lines connect transformed gene expression values across all samples; light-color lines represent one gene, the dark-color line is the average expression pattern of all genes in the cluster. B) Gene Ontology (GO) terms for Biological Process and Cellular Component found enriched in the gene clusters from (A). The X axis shows transformed p-values, the longer a bar the smaller its p-value. The direction and color of a bar indicate the phenotype in which the GO term is enriched [same color code as (A)]. GO terms highlighted in bold also belong to Set C. Pen = penetrations.

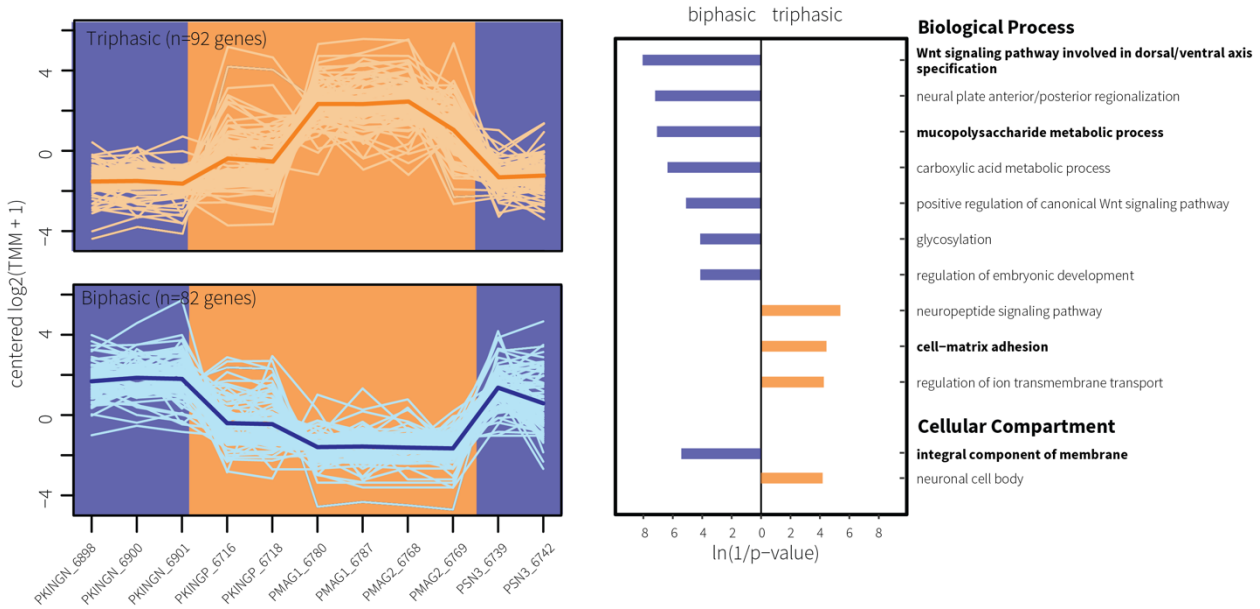

Figure 6. A) Gene clusters whose expression patterns correlate with the electric organ discharge (EOD) waveform complexity phenotypes biphasic (blue background) and triphasic (orange background). Samples are sorted alphabetically on the X axis. The lines connect transformed gene expression values across all samples; light-color lines represent one gene, the dark-color line is the average expression pattern of all genes in the cluster. B) Gene Ontology (GO) terms for Biological Process and Cellular Component found enriched in the gene clusters from (A). The X axis shows transformed p-values, the longer a bar the smaller its p-value. The direction and color of a bar indicate the phenotype in which the GO term is enriched [same color code as (A)]. GO terms highlighted in bold also belong to Set C.

Table 1. Phenotypic and collection information of the samples studied.

Tag No.	OTU	Phenotypes per EOD feature			CUMV Accession Number	Site Name	Lat, Long	SL (mm)	Sex
		Duration	Complexity	Polarity (diameter of penetrations)					
PKINGN_6898	*P. kingsleyae* (N-type)	long EOD	biphasic	NA (no penetrations)	95184	Bikagala Creek	-2.20, 11.56	131	M
PKINGN_6900	*P. kingsleyae* (N-type)	long EOD	biphasic	NA (no penetrations)	95184	Bikagala Creek	-2.20, 11.56	145	M
PKINGN_6901	*P. kingsleyae* (N-type)	long EOD	biphasic	NA (no penetrations)	95184	Bikagala Creek	-2.20, 11.56	149	M
PKINGP_6716	*P. kingsleyae* (P-type)	long EOD	triphasic	small penetrations	95183	Mouvanga Creek	-2.33, 11.69	92.5	J/F
PKINGP_6718	*P. kingsleyae* (P-type)	long EOD	triphasic	small penetrations	95183	Mouvanga Creek	-2.33, 11.69	91	J/F
PMAG1_6780	*P.* sp. 'magnostipes type I'	long EOD	triphasic	large penetrations	95155	Mouvanga Creek	-2.33, 11.69	107	M
PMAG1_6787	*P.* sp. 'magnostipes type I'	long EOD	triphasic	large penetrations	95155	Mouvanga Creek	-2.33, 11.69	97.5	F
PMAG2_6768	*P.* sp. 'magnostipes type II'	long EOD	triphasic	small penetrations	95155	Mouvanga Creek	-2.33, 11.69	93	M
PMAG2_6769	*P.* sp. 'magnostipes type II'	long EOD	triphasic	small penetrations	95155	Mouvanga Creek	-2.33, 11.69	124	M
PSN3_6739	*P.* sp. 'SN3'	short EOD	biphasic	NA (no penetrations)	Uncatalogued	Mouvanga Creek	-2.33, 11.69	73	J/F
PSN3_6742	*P.* sp. 'SN3'	short EOD	biphasic	NA (no penetrations)	95173	Mouvanga Creek	-2.33, 11.69	70	J/F

CUMV = Cornell University Museum of Vertebrates, EOD = electric organ discharge, F= female, J = juvenile, Lat = Latitude, Long = Longitude, M = male, NA = not applicable, OTU = operational taxonomic unit, SL = standard length

Table 2. All ten possible pairwise DGE comparisons with the total number of DEG and enriched GO terms for each. Also indicated is whether each comparison is informative for contrasting each EOD feature. The phenotypes for waveform polarity can only be contrasted in comparisons where both OTUs have penetrations. Informative comparisons for each EOD feature (Set A') are marked with an * in the column of the EOD feature they contrasted.

Comparison		Contrast			DEG genes	enriched GO terms		
OTU #1	OTU #2	Duration	Complexity	Polarity		BP	CC	MF
P. kingsleyae (N-type)	*P. kingsleyae* (P-type)	no	yes*	NA (no)	530	46	12	17
P. kingsleyae (N-type)	*P.* sp. 'magnostipes type I'	no	yes	NA (no)	1542	76	15	37
P. kingsleyae (N-type)	*P.* sp. 'magnostipes type II'	no	yes	NA (no)	1174	71	16	25
P. kingsleyae (N-type)	*P.* sp. 'SN3'	yes*	no	NA (no)	507	52	4	13
P. kingsleyae (P-type)	*P.* sp. 'magnostipes type I'	no	no	yes*	1322	40	12	25
P. kingsleyae (P-type)	*P.* sp. 'magnostipes type II'	no	no	no	719	47	10	24
P. kingsleyae (P-type)	*P.* sp. 'SN3'	yes	yes	NA (no)	385	33	3	14
P. sp. 'magnostipes type I'	*P.* sp. 'magnostipes type II'	no	no	yes	9	5	1	1
P. sp. 'magnostipes type I'	*P.* sp. 'SN3'	yes	yes	NA (no)	1053	43	6	27
P. sp. 'magnostipes type II'	*P.* sp. 'SN3'	yes	yes	NA (no)	489	40	9	16

BP = biological process, CC = cellular component, DEG = differentially expressed genes, DGE = differential gene expression, GO = gene ontology, NA = not applicable, MF = molecular function, OTU = operational taxonomic unit

Table 3. Total number of upregulated genes and enriched GO terms in Set C for each EOD feature, phenotype and ontology.

EOD feature	Phenotype	Upregulated genes	Protein-coding (%)	enriched GO terms		
				BP	CC	MF
Duration	short EODs	79	65 (82)	8	0	0
Duration	long EODs	26	21 (81)	3	0	0
Polarity	small penetrations	38	28 (74)	2	0	1
Polarity	large penetrations	106	87 (82)	3	0	3
Complexity	biphasic	46	34 (74)	2	1	0
Complexity	triphasic	25	23 (92)	1	0	0

BP = biological process, CC = cellular component, EOD = electric organ discharge, GO = gene ontology, MF = molecular function.

Table 4. Selected DEG in Set C for waveform duration by "general" functional class and EOD phenotype, and highlights of their predicted function.

Pking_Entrez_geneID	Pking_gene_name	Pking_gene_symbol	"general" functional class	upregulated in phenotype	Highlights of Predicted Function (edited from UniProt)
111832799	*stathmin domain containing 1*	*stmnd1*	Cytoskeletal & sarcomeric	long EODs	GO MF: tubulin binding; GO BP: microtubule depolymerization, regulation of cytoskeleton organization
111833088	*myosin-7-like*	LOC111833088	Cytoskeletal & sarcomeric	long EODs	Myosins are actin-based motor molecules with ATPase activity essential for muscle contraction
111856289	*pleckstrin homology-like domain family B member 1*	LOC111856289	Cytoskeletal & sarcomeric	long EODs	GO BP: regulation of microtubule cytoskeleton organization
111842483	*parvalbumin-2*	LOC111842483	Cytoskeletal & sarcomeric	short EODs	In muscle, parvalbumin is thought to be involved in relaxation after contraction. It binds two calcium ions
111846153	*troponin I, slow skeletal muscle-like*	LOC111846153	Cytoskeletal & sarcomeric	short EODs	Inhibitory subunit of troponin, the thin filament regulatory complex which confers calcium-sensitivity to striated muscle actomyosin ATPase activity
111856036	*parvalbumin-2-like*	LOC111856036	Cytoskeletal & sarcomeric	short EODs	In muscle, parvalbumin is thought to be involved in relaxation after contraction. It binds two calcium ions
111860236	*tropomyosin alpha-1 chain-like*	LOC111860236	Cytoskeletal & sarcomeric	short EODs	Binds to actin filaments in muscle and non-muscle cells. Plays a central role, in association with the troponin complex, in the calcium dependent regulation of vertebrate striated muscle contraction. In non-muscle cells is implicated in stabilizing cytoskeleton actin filaments.
111833641	*hyaluronidase-5-like*	LOC111833641	Extracellular matrix	short EODs	Catalyzes the hydrolysis of hyaluronan into smaller oligosaccharide fragments
111837392	*fibroblast growth factor binding protein 1*	*fgfbp1*	Extracellular matrix	short EODs	Acts as a carrier protein that releases fibroblast-binding factors (FGFs) from the extracellular matrix (EM) storage and thus enhance the mitogenic activity of FGFs
111860877	*collagenase 3-like*	LOC111860877	Extracellular matrix	short EODs	Plays a role in the degradation of extracellular matrix proteins
111834720	*protein EFR3 homolog B-like*	LOC111834720	Lipid metabolism	short EODs	Component of a complex required to localize phosphatidylinositol 4-kinase (PI4K) to the plasma membrane. The complex acts as a regulator of phosphatidylinositol 4-phosphate (PtdIns4P) synthesis
111840357	*PTB domain-containing engulfment adapter protein 1-like*	LOC111840357	Lipid metabolism	short EODs	Modulates cellular glycosphingolipid and cholesterol transport

111846286	*retinoic acid receptor responder protein 3-like*	LOC111846286	Lipid metabolism	short EODs	Catalyzes the calcium-independent hydrolysis of acyl groups in various phosphatidylcholines (PC) and phosphatidylethanolamine (PE)
111847640	*phosphatidylinositol 3-kinase regulatory subunit gamma-like*	LOC111847640	Lipid metabolism	short EODs	Binds to activated (phosphorylated) protein-tyrosine kinases through its SH2 domain and regulates their kinase activity
111852518	*proto-oncogene c-Fos-like*	LOC111852518	Lipid metabolism	short EODs	In growing cells, activates phospholipid synthesis, possibly by activating CDS1 and PI4K2A

BP = biological process, DEG = differentially expressed genes, EOD = electric organ discharge, GO = gene ontology, MF = molecular function.

Table 5. Selected DEG in Set C for waveform polarity, by "general" functional class and EOD phenotype, and highlights of their expected function.

Pking_Entrez_geneID	Pking_gene_name	Pking_gene_symbol	"general" functional class	upregulated in phenotype	Highlights of Predicted Function (edited from UniProt)
111838673	*Kv channel-interacting protein 1-like*	LOC111838673	Cation homeostasis	Large penetrations	Regulatory subunit of Kv4/D (Shal)-type voltage-gated rapidly inactivating A-type potassium channels. Regulates channel density, inactivation kinetics and rate of recovery from inactivation in a calcium-dependent and isoform-specific manner.
111843424	*potassium voltage-gated channel subfamily C member 2-like*	LOC111843424	Cation homeostasis	Large penetrations	Voltage-gated potassium channel that mediates transmembrane potassium transport in excitable membranes, primarily in the brain. Contributes to the regulation of the fast action potential repolarization and in sustained high-frequency firing in neurons of the central nervous system
111849794	*extracellular calcium-sensing receptor-like*	LOC111849794	Cation homeostasis	Small penetrations	G-protein-coupled receptor that senses changes in the extracellular concentration of calcium ions and plays a key role in maintaining calcium homeostasis. The activity of this receptor is mediated by a G-protein that activates a phosphatidylinositol-calcium second messenger system
111853690	*protein kinase cGMP-dependent 1*	*prkg1*	Cation homeostasis	Small penetrations	Serine/threonine protein kinase. Numerous protein targets for PRKG1 phosphorylation are implicated in modulating cellular calcium. Proteins that are phosphorylated by PRKG1 regulate platelet activation and adhesion, smooth muscle contraction, cardiac function, gene expression.
111833185	*myosin XVB*	*myo15b*	Cytoskeletal & sarcomeric	Large penetrations	Unknown, due to the absence of a functional motor domain
111837476	*troponin C, slow skeletal and cardiac muscles*	LOC111837476	Cytoskeletal & sarcomeric	Large penetrations	Troponin is the central regulatory protein of striated muscle contraction. The binding of calcium troponin abolishes its inhibitory action on actin filaments
111837614	*spindle and kinetochore associated complex subunit 3*	*ska3*	Cytoskeletal & sarcomeric	Large penetrations	Component of the SKA1 complex, which is a direct component of the kinetochore-microtubule interface and directly associates with microtubules as oligomeric assemblies
111847443	*cysteine and glycine rich protein 3*	*csrp3*	Cytoskeletal & sarcomeric	Large penetrations	Positive regulator of myogenesis. Plays a crucial and specific role in the organization of cytosolic structures in cardiomyocytes. It is essential for calcineurin anchorage to the Z line. Can directly bind to actin filaments. Isoform 2 may play a role in early sarcomere organization.
111848724	*protein tilB homolog*	LOC111848724	Cytoskeletal & sarcomeric	Large penetrations	May play a role in dynein arm assembly

111850126	*tubulin beta 2A class Iia*	*tubb2a*	Cytoskeletal & sarcomeric	Large penetrations	Tubulin is the major constituent of microtubules
111852410	*heat shock protein HSP 90-alpha 1*	LOC111852410	Cytoskeletal & sarcomeric	Large penetrations	Plays a key role in slow and fast muscle development in the embryo. Plays a role in myosin expression and assembly
111853965	*tubulin beta chain*	LOC111853965	Cytoskeletal & sarcomeric	Large penetrations	Tubulin is the major constituent of microtubules
111859691	*fibroblast growth factor 13*	LOC111859691	Cytoskeletal & sarcomeric	Large penetrations	Microtubule-binding protein which directly binds tubulin and is involved in both polymerization and stabilization of microtubules
111834243	*desmin-like*	LOC111834243	Cytoskeletal & sarcomeric	Small penetrations	Muscle-specific type III intermediate filament essential for proper muscular structure and function. Plays a crucial role in maintaining the structure of sarcomeres. May act as a sarcomeric microtubule-anchoring protein
111856797	*myosin light chain 3-like*	LOC111856797	Cytoskeletal & sarcomeric	Small penetrations	Regulatory light chain of myosin. Does not bind calcium.
111833641	*hyaluronidase-5-like*	LOC111833641	Extracellular matrix	Small penetrations	Catalyzes the hydrolysis of hyaluronan into smaller oligosaccharide fragments
111848653	*collagen alpha-1(I) chain-like*	LOC111848653	Extracellular matrix	Small penetrations	Type I collagen is a member of group I collagen (fibrillar forming collagen)
111857875	*collagen alpha-1(X) chain-like*	LOC111857875	Extracellular matrix	Small penetrations	Type X collagen is a product of hypertrophic chondrocytes and has been localized to presumptive mineralization zones of hyaline cartilage
111833176	*phospholipid-transporting ATPase IA*	LOC111833176	Lipid metabolism	Large penetrations	Involved in the transport of aminophospholipids from the outer to the inner leaflet of various membranes and ensures the maintenance of asymmetric distribution of phospholipids
111833450	*low-density lipoprotein receptor-like*	LOC111833450	Lipid metabolism	Large penetrations	Binds LDL, the major cholesterol-carrying lipoprotein of plasma, and transports it into cells by endocytosis
111844857	*ectonucleotide pyrophosphatase/ phosphodiesterase 2*	*enpp2*	Lipid metabolism	Large penetrations	Hydrolyzes lysophospholipids to produce the signaling molecule lysophosphatidic acid (LPA) in extracellular fluids. Acts as an angiogenic factor by stimulating migration of smooth muscle cells and microtubule formation.
111845574	*phospholipase A2-like*	LOC111845574	Lipid metabolism	Large penetrations	PA2 catalyzes the calcium-dependent hydrolysis of the 2-acyl groups in 3-sn-phosphoglycerides, this releases glycerophospholipids and arachidonic acid that serve as the precursors of signal molecules.
111847497	*peroxiredoxin-6-like*	LOC111847497	Lipid metabolism	Large penetrations	It has phospholipase activity

111855292	*long-chain-fatty-acid--CoA ligase 4-like*	LOC111855292	Lipid metabolism	Large penetrations	Activation of long-chain fatty acids for both synthesis of cellular lipids, and degradation via beta-oxidation
111853114	*alkaline ceramidase 2-like*	LOC111853114	Lipid metabolism	Small penetrations	Hydrolyzes the sphingolipid ceramide into sphingosine and free fatty acid

DEG = differentially expressed genes, EOD = electric organ discharge.

Table 6. Selected DEG in Set C for waveform complexity, by "general" functional class and EOD phenotype, and highlights of their expected function.

Pking_Entrez_geneID	Pking_gene_name	Pking_gene_symbol	"general" functional class	upregulated in phenotype	Highlights of Predicted Function (edited from UniProt)
111838181	*solute carrier family 9 member A7*	*slc9a7*	Cation homeostasis	Biphasic	Protein: Sodium/hydrogen exchanger 7. Gene: SLC9A7. Mediates electroneutral exchange of protons for Na+ and K+ across endomembranes
111838015	*chloride intracellular channel protein 2-like*	LOC111838015	Cation homeostasis	Triphasic	Can insert into membranes and form chloride ion channels. Inhibits calcium influx
111848312	*voltage-dependent calcium channel gamma-1 subunit-like*	LOC111848312	Cation homeostasis	Triphasic	Regulatory subunit of the voltage-gated calcium channel that gives rise to L-type calcium currents in skeletal muscle. Regulates channel inactivation kinetics
111845832	*5'-AMP-activated protein kinase subunit gamma-2-like*	LOC111845832	Cytoskeletal & sarcomeric	Biphasic	AMP/ATP-binding subunit of AMP-activated protein kinase (AMPK). Acts as a regulator of cellular polarity by remodeling the actin cytoskeleton; probably by indirectly activating myosin
111850616	*transmembrane protein 47-like*	LOC111850616	Cytoskeletal & sarcomeric	Biphasic	Regulates cell junction organization in epithelial cells. May regulate F-actin polymerization
111851223	*FYVE, RhoGEF and PH domain-containing protein 4-like*	LOC111851223	Cytoskeletal & sarcomeric	Biphasic	Plays a role in regulating the actin cytoskeleton and cell shape
111857398	*protein-methionine sulfoxide oxidase mical2b-like*	LOC111857398	Cytoskeletal & sarcomeric	Biphasic	Promotes depolymerization of F-actin
111847443	*cysteine and glycine rich protein 3*	*csrp3*	Cytoskeletal & sarcomeric	Triphasic	Positive regulator of myogenesis. Plays a crucial and specific role in the organization of cytosolic structures in cardiomyocytes. It is essential for calcineurin anchorage to the Z line. Can directly bind to actin filaments. Isoform 2 may play a role in early sarcomere organization
111854588	*capping protein regulator and myosin 1 linker 3*	*carmil3*	Cytoskeletal & sarcomeric	Triphasic	No info por CARMIL3, but CARMIL2 is a cell membrane-cytoskeleton-associated protein that plays a role in the regulation of actin polymerization at the barbed end of actin filaments. Enhances actin polymerization

111841398	*inter-alpha-trypsin inhibitor heavy chain 3*	*itih3*	Extracellular matrix	Biphasic	May act as a carrier of hyaluronan in serum or as a binding protein between hyaluronan and other matrix proteins
111841399	*inter-alpha-trypsin inhibitor heavy chain H3-like*	LOC111841399	Extracellular matrix	Biphasic	May act as a carrier of hyaluronan in serum or as a binding protein between hyaluronan and other matrix proteins
111853010	*ependymin-like*	LOC111853010	Extracellular matrix	Triphasic	GO MF: calcium ion binding. GO BP: cell-matrix adhesion
111853027	*ependymin-like*	LOC111853027	Extracellular matrix	Triphasic	GO MF: calcium ion binding. GO BP: cell-matrix adhesion
111853814	*epiphycan-like*	LOC111853814	Extracellular matrix	Triphasic	May have a role in bone formation and also in establishing the ordered structure of cartilage through matrix organization

BP = biological process, DEG = differentially expressed genes, EOD = electric organ discharge, GO = gene ontology, MF = molecular function