ORGANISMAL BIOLOGY

Divergent cis-regulatory evolution underlies the convergent loss of sodium channel expression in electric fish

Sarah LaPotin[1]†, Mary E. Swartz[2], David M. Luecke[3], Savvas J. Constantinou[3,4], Jason R. Gallant[3,4], Johann K. Eberhart[2], Harold H. Zakon[1,5]*

South American and African weakly electric fish independently evolved electric organs from muscle. In both groups, a voltage-gated sodium channel gene independently lost expression from muscle and gained it in the electric organ, allowing the channel to become specialized for generating electric signals. It is unknown how this voltage-gated sodium channel gene is targeted to muscle in any vertebrate. We describe an enhancer that selectively targets sodium channel expression to muscle. Next, we demonstrate how the loss of this enhancer, but not trans-activating factors, drove the loss of sodium channel gene expression from muscle in South American electric fish. While this enhancer is also altered in African electric fish, key transcription factor binding sites and enhancer activity are retained, suggesting that the convergent loss of sodium channel expression from muscle in these two electric fish lineages occurred via different processes.

INTRODUCTION

Weakly electric fish provide excellent study systems to analyze how molecular processes underlie phenotypic evolution because of the convergent evolution of numerous phenotypes (1, 2). The two most studied groups of electric fish (African: Mormyroidea and South American: Gymnotiformes) generate electric fields from electric organs (EOs), an evolutionary novelty that arose independently from muscle (3). Electric fish detect their species-specific electric organ discharges (EODs) with specialized electrosensory cells that evolved from lateral line hair cells (4). Weakly electric fish are nocturnally active and use this "radar-like" system to identify objects and conspecifics in the dark or in muddy waters. One well-known member of the gymnotiforms, the electric eel (Electrophorus electricus), has additionally evolved a strong EO for defense and prey capture (5, 6).

The EODs of both groups are generated by sodium-based action potentials. The voltage-gated sodium channel gene, Scn4a, is expressed in vertebrate muscle (7). Scn4a function is essential for proper muscle physiology (8, 9), making its expression in muscle evolutionarily constrained. As a result of a whole-genome duplication event at the origin of teleosts, Scn4a duplicated into scn4aa and scn4ab. In most teleosts, both paralogs are expressed in muscle (10, 11). In both groups of weakly electric fish, this duplication event facilitated the evolution of EOs for signaling as one paralog, scn4ab, retained expression in muscle, while the other paralog, scn4aa, shifted its expression from muscle to the evolutionarily novel EO. There, it evolved at elevated rates underlying variations in species-specific EOD waveforms (3, 12, 13). We used the independent loss of expression of scn4aa from muscle to determine whether convergent molecular mechanisms underlie the convergent evolution of novelty.

Gymnotiformes include five families with myogenic (muscle-derived) EOs (Fig. 1) (14). However, the Apteronotidae only have a myogenic EO for the first few weeks of life (15). It then degenerates and is replaced with a neurogenic (neuron-derived) EO that develops from the axons of specialized spinal electromotor neurons (Fig. 1). Perhaps because of the fleeting appearance of a myogenic EO in this family, the apteronotids have not lost expression of scn4aa in skeletal muscle unlike the other gymnotiforms.

The superfamily mormyroidea includes one family (Gymnarchidae) with a single species (Gymnarchus niloticus) and a second family with well over 200 species (Mormyridae). scn4aa expression is lost from muscle in mormyroids (3, 16).
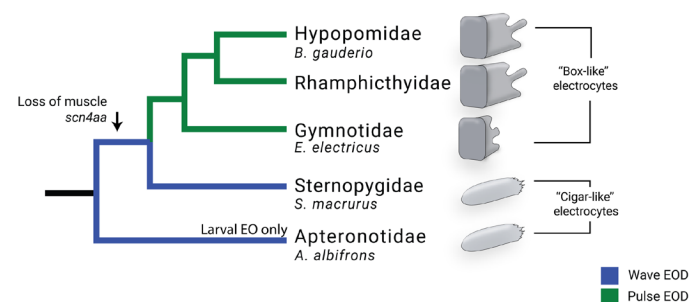

Fig. 1. Gymnotiform electric fish and their EOs. Some lineages generate sinusoidal EODs (wave EOD); others generate brief pulsatile EODs (pulse EOD). The structure of the electrocytes, the single cells that comprise the EO, is schematically represented to the right. Wave fish have simple, cylindrical electrocytes that resemble elongated muscle fibers, whereas the pulse fish have more complex, derived electrocytes with box-like shapes [the most extreme variant is observed in the electric eel (E. electricus), in which the electrocytes are thin, pancake-like cells]. The gymnotiform species used in this study are identified in their relative phylogenetic position according to (14).

[1]Department of Neuroscience, The University of Texas, Austin, TX 78712, USA. [2]Department of Molecular Biosciences, The University of Texas, Austin, TX 78712, USA. [3]Department of Integrative Biology, Michigan State University, East Lansing, MI 48824, USA. [4]Ecology, Evolution, and Behavior Program, Michigan State University, East Lansing, MI 48824, USA. [5]Department of Integrative Biology, The University of Texas, Austin, TX 78712, USA.
*Corresponding author. Email: h.zakon@austin.utexas.edu
†Present address: Department of Human Genetics, University of Utah, Salt Lake City, UT 84112-5330, USA.

RESULTS

A conserved enhancer drives scn4a gene expression in muscle

Because little was known about the transcriptional targeting of Scn4a channels to muscle, we performed a bioinformatic analysis to identify conserved noncoding elements (CNEs) of *Scn4a* and *scn4aa* in vertebrates. We aligned introns and intergenic regions, both 5′ and 3′, of *scn4aa* of nonelectric teleosts as well as the *Scn4a* of other vertebrates. We identified three CNEs within the large first intron that is flanked on both sides by exons containing the 5′ untranslated region (5′UTR; Fig. 2A). CNE1 and CNE2 are teleost specific, and their presence is labile across teleosts, whereas CNE3 is conserved across shark, teleosts, gar, chicken, and mammals (Figs. 2C and 3A, figs. S1 and S2, and table S1). We did not detect CNE3 in lamprey or hagfish.

We note that Kraner *et al.* (*17*, *18*) identified a region that regulates *Scn4a* expression (which we refer to as CNEx) in a partial survey

of the 5′UTR of rat Scn4a (*17*, *18*). We located CNEx ~14 kb upstream of CNE3 in the rat genome. BLAST search demonstrates that CNEx is conserved in tetrapods but absent in other vertebrates (fig. S3). Because CNE1 and CNE2 are teleost specific and CNEx is tetrapod specific, we propose that CNE3 is a major, ancient element for targeted expression of Scn4a in vertebrate muscle because it is conserved widely in gnathostomes.

To test whether CNE3 regulates *scn4aa* expression, we deleted it from zebrafish using CRISPR-Cas9 mutagenesis (fig. S4), and compared the expression of *scn4aa* in the trunks of CNE-deleted and wild-type (WT) zebrafish embryos with reverse transcription quantitative polymerase chain reaction (RT-qPCR). In CNE3-deleted fish, we found that the expression of *scn4aa* was largely eliminated (>62-fold decrease compared with WT siblings; Fig. 2B).

Next, we used a green fluorescent protein (GFP) expression assay to determine whether CNE3 can drive gene expression in zebrafish
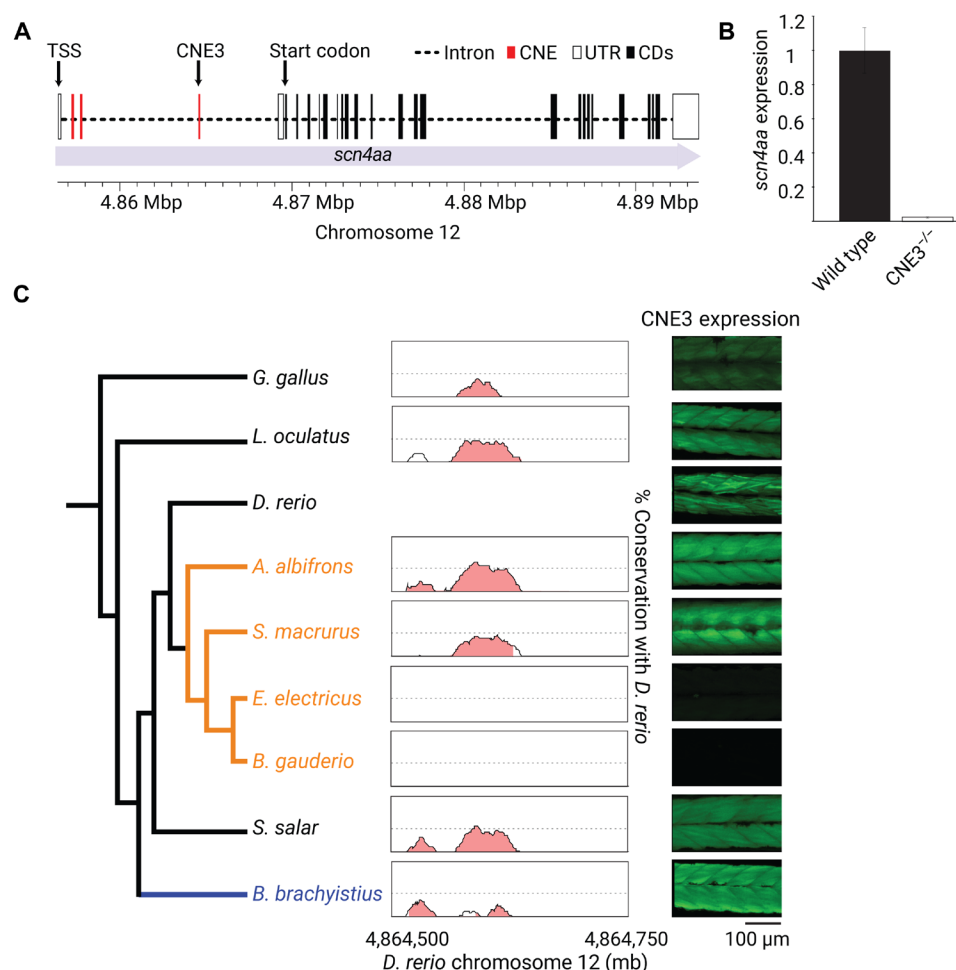

Fig. 2. CNE3 is an enhancer for the muscle sodium channel gene *Scn4a/scn4aa*. (**A**) Map of *scn4aa* exon and intron structure in *Danio rerio* (zebrafish) beginning at the transcription start site (TSS). The *scn4aa* 5′UTR is interrupted by a long intron with three CNEs (CNE1 to CNE3). (**B**) CRISPR-Cas9–mediated deletion of CNE3 knocks down expression of *scn4aa* in trunk muscle of zebrafish embryos by 62-fold. Average fold change in expression of *scn4aa* in WT (black) versus CNE3-deleted (white) siblings at 52 hours postfertilization. (**C**) Examination of CNE3s conservation among vertebrates including salmon (*Salmo salar*), gar (*Lepisosteus oculatus*), and chicken (*Gallus gallus*) (see figs. S1 and S2 for full set). Putative CNE3 sequences were aligned using LAGAN and displayed using mVISTA software (calculation window size = 50 bp; white = 50% homology, pink = 60% homology, dotted line = 75% homology with the reference species, *D. rerio*). Transgenic zebrafish were created using Tol2-transposase–mediated insertion (see Materials and Methods). For each species shown, CNE3 was inserted into a plasmid containing a minimal promoter and GFP gene, injected into single-cell zebrafish embryos, and then imaged under standardized conditions 50 hours postfertilization.
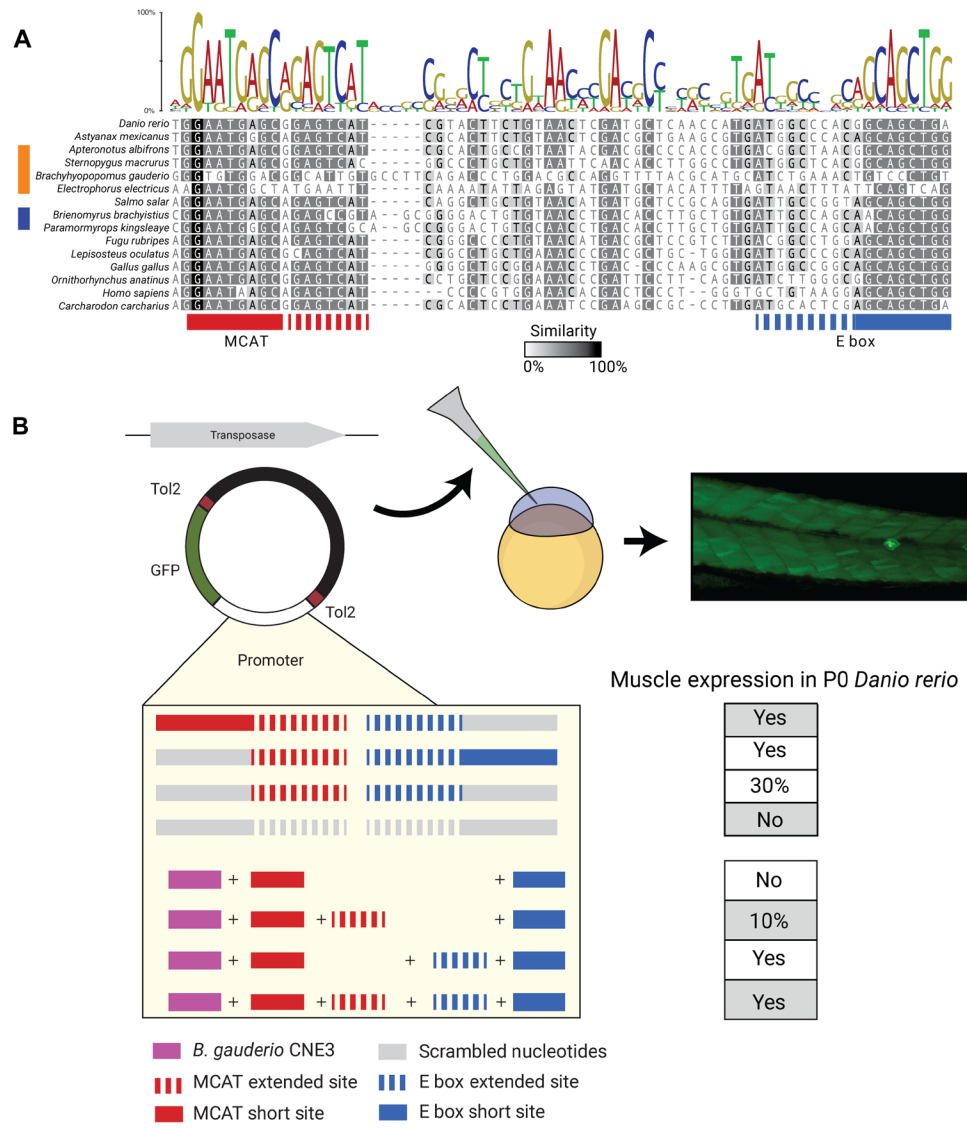
Fig. 3. TFBS in CNE3. (A) Alignment of CNE3 region across vertebrates. CNE3 has MCAT and E-Box TFBS, with gymnotiforms (orange) and mormyrids (blue) indicated. The sequence logo plot indicates strong conservation of key motifs across vertebrates. TFBS predicted by rVISTA are indicated by a bold bar (MCAT, red; E-Box, blue). Ten additional nucleotides flanking each TFBS are indicated by a broken bar (TEF1-3′-extended, red; myogenin-5′-extended, blue). **(B)** Functional assessment of TFBS in zebrafish muscle using Tol2-transposase–mediated transgenesis. Zebrafish CNE3 with scrambled TFBS for MCAT, MCAT-3′-ex, and E-Box-5′-ex alone were combined with a minimal promoter and GFP into constructs in combinations as indicated and injected into single-cell stage embryos. Alternatively, CNE3 from the gymnotiform *B. gauderio* in zebrafish muscle with intact TFBS for MCAT, MCAT-3′-ex, and E-Box-5′-ex alone were used to make constructs instead, in combinations as indicated. Resulting P0 progeny were scored for fluorescence in muscle. A minimum of 40 individuals were scored for each group.

muscle. We made GFP constructs containing CNE3 from salmon (*Salmo salar*), zebrafish (*Danio rerio*), gar (*Lepisosteus oculatus*), or chicken (*Gallus gallus*) and injected each construct into zebrafish embryos at the single-cell stage (*19*). The CNE3 sequences from each species drove GFP expression selectively in zebrafish muscle (Fig. 2C and movies S1 to S3). Collectively, these experiments demonstrate that CNE3 is a major enhancer of the *Scn4a* and *scn4aa* sodium channel genes, selectively targeting it to vertebrate muscle.

CNE3 has binding sites for muscle transcription factors

The strong conservation of CNE3 and its demonstrated enhancer activity in zebrafish suggest that it has evolutionarily conserved transcription factor binding sites (TFBS). Using rVISTA, we identified two putative TFBS within CNE3 that are highly conserved across vertebrates (Fig. 3A and fig. S2): an E box motif (CANNTG) and an MCAT element (GGAATG). These are binding sites for basic helix-loop-helix (bHLH) transcription factors (TFs) important in muscle development such as myogenin, myoD, etc. (*20*, *21*), and TEA domain transcription factor 1 (TEAD1; also known as NTEF-1 or Tef1) (*22*), respectively. We also detected conservation of the 5′ flanking region of the MCAT element TFBS and the 3′ region of the E box TFBS, although we identified no additional TFBSs in these regions.

To determine whether these putative TFBSs facilitate gene expression in muscle, we selectively scrambled them in our CNE3 GFP constructs, scrambling either the predicted binding site alone (MCAT-sc; E box-sc) or along with 10 to 12 nucleotides extended on one

side (MCAT+3'ex-sc; E box+5'ex-sc; Fig. 3B). We tested the CNE3 GFP constructs singly or in combination to assess whether these sequences are necessary for enhancer activity. The zebrafish CNE3 GFP constructs with either MCAT-sc or E box-sc alone consistently drove GFP expression in muscle, suggesting a level of redundancy of these TFBS. The combination of MCAT-sc and E box-sc still drove GFP expression, but in only about 30% of cases, suggesting that some of the conserved regions flanking MCAT and E box binding sites or other regions of CNE3 have residual ability to drive transcription (Fig. 3B). When the two TFBS along with their flanking extensions were all scrambled (MCAT-3'ex-sc and E box-5'ex-sc), GFP muscle expression was never observed (Fig. 3B). We conclude that the MCAT and E box TFBSs and their extended regions are necessary for enhancer activity.

Evolutionary change in CNE3 in both groups of electric fish
We next introduced sequences from four gymnotiforms and two mormyrids into the CNE3 alignments (Figs. 2C and 3A and fig. S1). In the gymnotiforms, the black ghost (*Apteronotus albifrons*), *scn4aa* is robustly expressed in muscle (*3, 12, 23*). In three other species—the gold-lined knifefish (*Sternopygus macrurus*), the pintail knifefish (*Brachyhypopomus gauderio*), and the electric eel (*E. electricus*)—*scn4aa* expression is lost from muscle (*3, 12, 24–26*). Mormyrid muscle does not express *scn4aa*.

The electric fish sequences were well anchored at both ends within the alignments by the sequence from the genomic regions that correspond to the 5′UTR of the mRNA transcript. In contrast, the presence and integrity of CNE3 were variable across gymnotiforms in a pattern generally consistent with the expression of *scn4aa* in muscle. *A. albifrons*, a member of an earlier branching lineage (*14, 27*) [but see (*28*)] (Fig. 1) in which *scn4aa* is highly expressed in muscle, has an CNE3 similar to other teleosts. In the remaining three species, *scn4aa* expression is lost in muscle. *S. macrurus* shows a recognizable but reduced CNE3. In contrast, *B. gauderio* and *E. electricus* have no clear CNE3 despite repeated attempts at alignment with a range of window sizes and thresholds of homology; at most, a few remnants remain (Fig. 3A). Furthermore, these regions showed little sequence similarity between *B. gauderio* and *E. electricus*, suggesting that once this region began degrading in a common ancestor, it degraded rapidly and independently in each lineage. Considering these findings in light of gymnotiform phylogeny suggests that CNE3 began to degrade following the divergence of the Apteronotidae from the rest of the gymnotiforms and was completely lost at the base of the more derived pulse EOD lineages (Fig. 2B). Last, we note that CNE3 is present in mormyrids but is substantially altered in comparison with other teleosts (fig. S1). It still has E box and MCAT motifs (Fig. 3A) but has a slightly degraded 3′ flanking region. Thus, extensive changes occurred in the CNE3 in both electric lineages, most markedly, its complete loss in gymnotiforms with derived EOs.

We next made GFP constructs containing CNE3 from gymnotiforms (*A. albifrons*, *S. macrurus*, *B. gauderio*, and *E. electricus*) and one mormyroid (*Brienomyrus brachyistius*). CNE3 from the gymnotiforms *A. albifrons* and *S. macrurus* drove gene expression in zebrafish muscle. The GFP constructs made from the best-aligned regions from *B. gauderio* and *E. electricus* did not drive GFP expression (Fig. 2C). Loss of CNE3 would likely contribute to the loss of expression of *scn4aa* from trunk muscle, as seen in derived gymnotiforms. However, the CNE3 from the mormyrid, *B. brachyistius*, did drive GFP expression. Thus, loss of enhancer activity of CNE3

is not an explanation for the loss of *scn4aa* expression in muscle of *S. macrurus* or mormyrids.

Sufficiency of TFBS in CNE3 to drive gene expression
While the MCAT and E box TFBSs are greatly degraded in the derived gymnotiforms, it remains possible that other nucleotide changes in CNE3 contribute to the loss of enhancer activity in these species. To determine whether either of these two TFBS is sufficient for enhancer activity, we assembled chimeric constructs with the zebrafish TFBSs inserted into the CNE3 sequence from *B. gauderio* at approximately the same location as seen in other vertebrates. We first tested the activity of the minimal MCAT+E box sites (MCAT+E box-bg) and found that this combination was insufficient to drive GFP expression (Fig. 3B). We next tested each extended binding site and found that MCAT-3'ex-bg drove expression 10% of the time, whereas E box-5'ex-bg alone and the combination of MCAT-3'ex-bg with E box-5'ex-bg consistently drove muscle expression. Collectively, our results demonstrate that the extended MCAT and E box binding sites are necessary and sufficient to drive gene expression in muscle, and the loss of these binding sites, as occurs in derived gymnotiforms, is likely responsible for the loss of *scn4aa* expression in muscle.

No loss of trans-factors in gymnotiform electric fish
While CNE3 is necessary and sufficient for *scn4aa* expression in muscle, it remains plausible that loss of expression of *scn4aa* in gymnotiform muscle was secondary due to the loss of its trans-activation. To address this, we injected a zebrafish CNE3-driven GFP construct into newly fertilized eggs of the gymnotiform *B. gauderio*. Of 18 embryos with visible fluorescence, all showed a consistent pattern of parallel stripes along the dorsal flank (Fig. 4A). Visualizing the larvae at a time when the EO is present (18 days) (*29*), we noted robust expression of GFP in muscle, none in the electrocytes, and none in uninjected siblings (Fig. 4B). We observed an additional individual after 5 months postfertilization, which showed consistent and strong GFP expression in all skeletal muscles but not in adjacent electrocytes (Fig. 4C). The lack of GFP expression in electrocytes is consistent with our previous findings that EOs down-regulate myogenic TFs such as myogenin (*25*). This demonstrates that gymnotiforms retain the capacity to drive *scn4aa* expression in muscle, indicating that only loss of cis-regulatory elements led to reduced expression of *scn4aa* in their muscle.

DISCUSSION
CNE3 is an enhancer targeting scn4aa expression to muscle in vertebrates
Before our study, little was known about what directs *Scn4a* expression to vertebrate muscle. We identified a CNE that is widely present in vertebrates within the 5′ intron of *Scn4a* that selectively drives gene expression in muscle and whose deletion eliminates *scn4aa* expression. CNE3 has an E box that would potentially bind bHLH myogenic TFs that initiate and maintain the muscle phenotype in vertebrates and an MCAT motif that would likely bind TEAD1, known to activate a number of muscle-specific genes (*22*). Myogenic TFs, in particular MRF4, were known to regulate *Scn4a* expression in rat muscle (*17, 18, 30, 31*) at a site in the 5′UTR. However, we only observed this site in tetrapods; so, it cannot be a universal enhancer for *Scn4a* expression in vertebrates. We propose that these
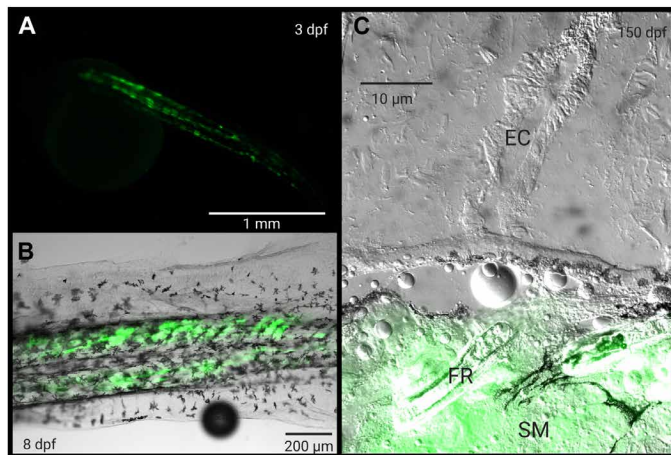
Fig. 4. Trans-environment of *B. gauderio* muscle is consistent with *Danio rerio*. Transgenic *B. gauderio* were created using Tol2-mediated transgenesis. A *D. rerio* CNE3 was combined with a minimal promoter and GFP into a plasmid construct and injected single-cell *B. gauderio* embryos. Representative P0 progeny are shown. (**A**) A newly hatched *B. gauderio* embryo 3 days postfertilization (dpf) show fluorescence in muscle under epifluorescence. (**B**) A confocal micrograph of the larval tail at 8 dpf shows individual fluorescent muscle fibers. (**C**) At 150 dpf, the EO is fully functional. Tail containing electrocytes (EC), fin rays (FR), and skeletal muscle (SM) indicates strong fluorescence in the muscle but not EO. In all images, dorsal is up and anterior is to the left.

bHLH TFs, as well as TEAD1, likely regulate *Scn4a* widely across vertebrates by their action on CNE3.

Alterations in CNE3 of electric fish

Both groups of electric fish show alterations in CNE3 compared with other teleosts (fig. S1). The most pronounced alteration of CNE3—its complete loss—occurs in the pulse gymnotiform species that comprise the majority (68%) of gymnotiforms with myogenic EOs (i.e., excluding apteronotids) as adults (*32*); *B. gauderio* and *E. electricus* are from two different families and are widely separated within the lineage of pulse fish (Fig. 1B), suggesting that the loss of CNE3 occurred in most gymnotiforms with adult myogenic EOs.

S. macrurus has a recognizable CNE3 with intact TFBS that drive GFP expression in the zebrafish assay. This species, unique among gymnotiforms, is believed to regulate gene expression in its muscle and EO primarily by posttranscriptional processes (*33*). Such a mechanism would not be detected in our transgenic method. The two mormyrids we examined have an altered CNE3, although with largely intact TFBS, and their CNE3 drives gene expression in the zebrafish muscle assay. The mormyrids lack *scn4aa* mRNA in their EO. Collectively, our findings indicate that the degradation of critical TFBS in CNE3 of the gymnotiforms with derived EOs can fully explain the loss of *scn4aa* expression in muscle. However, despite some unusual features, the CNE3 in mormyrids functions as an enhancer in the GFP assay and contains the critical TFBS. Therefore, the loss of *scn4aa* expression in muscle in these two different clades of electric fish is likely via different mechanisms (e.g., epigenetic regulation, or mormyrid-specific repressors and silencer elements or miRNAs). Our results open the way for further studies aimed at identifying how CNE3 is regulated in electric species with other modes of regulation than complete loss.

MATERIALS AND METHODS
Bioinformatics
Available genome data on the NCBI (National Center for Biotechnology Information) were BLAST searched with *scn4a* or *scn4aa* 5′UTR sequences from a number of teleosts and other vertebrates (see supplemental file 3 for complete list) to find the 5′ intron sequence. Unpublished genome data were similarly BLAST searched with *scn4aa* 5′UTR sequences to find 5′ intron sequences from the electric species *S. macrurus* (provided by G. Unguez and M. Samanta), *B. gauderio*, and *B. brachyistius* (J. Gallant, Michigan State University). The 5′ intron sequence from black ghost (*A. albifrons*) was sequenced from genomic DNA via PCR (NEB Q5 High-Fidelity DNA Polymerase, Qiagen DNeasy Blood & Tissue protocol, GSAF, The University of Texas at Austin). Alignments of the 5′ intron from all vertebrates were run using mLAGAN and visualized using mVISTA (http://genome.lbl.gov/vista/mvista/submit.shtml).

Putative TFBS were identified with rVISTA (https://rvista.dcode.org/). To minimize the high rate of potential false positives when individual sequences are analyzed, rVISTA identifies putative TFBS in pairs of sequences. We took a conservative approach and only called a putative TFBS when it was recognized in all of the sequences submitted for a particular CNE. We then cross-checked with the JASPAR database (http://jaspar.genereg.net/) to determine whether there was a match for the predicted TFBS.

Zebrafish husbandry
Zebrafish were raised and cared for in the Eberhart laboratory facility at UT Austin. All zebrafish were derived from AB WT stocks. Institutional animal care and use committee–approved protocols were observed.

GFP assay
We used transgenic zebrafish to assess the functional importance of CNE3. Cytoplasmic GFP constructs (Eberhart laboratory, The University of Texas at Austin) were made with CNE3 sequences from a wide variety of electric and nonelectric species. Genomic DNA was obtained from *S. macrurus*, *B. brachyistius* (Zakon laboratory, UT Austin), gar (I. Braasch, Michigan State University), and zebrafish (Eberhart laboratory, UT Austin). Genomic DNA from chicken, salmon, eel, *B. gauderio*, and black ghost was extracted according to the DNeasy Blood & Tissue protocol (QIAGEN). Primers were designed to flank the CNE3 sequences that were identified for each species with restriction sites added (5'Xho I, 3'Bgl II, Integrated DNA Technologies). Each insert sequence was PCR amplified from genomic DNA using Q5 High-Fidelity DNA Polymerase (NEB) and digested with Xho I and Bgl II restriction enzymes (NEB). Ligations were performed for each insert and GFP vector (Quick Ligation Kit, NEB), cloned into One Shot TOP10 Chemically Competent *Escherichia. coli* (Thermo Fisher Scientific), and grown on LB Agar plates with ampicillin resistance. Plasmid DNA was extracted according to the QIAprep Spin Miniprep Kit (QIAGEN) protocol. All GFP constructs were sequenced for verification (GSAF, The University of Texas at Austin) before injection. In vitro synthesis of a transposase plasmid was done with the mMESSAGE mMACHINE SP6 Transcription Kit (Invitrogen).

GFP construct injections were performed in zebrafish AB WT embryos at the single-cell stage. Each embryo was injected with a 1-nl cocktail of GFP plasmid (100 ng/μl), transposase mRNA (50 ng/μl), and 0.2% phenol red. At 24 hours postfertilization, embryos were screened for GFP expression. Injected P0 fish were raised to maturity

and outcrossed to an AB WT line to generate stable transgenic lines. Confocal images of P0 embryos were taken at 50 hours postfertilization using a Zeiss LSM 710. A minimum of 40 embryos were sampled for every group. The *pTol2-HuC(elavl3)-CaMPARI2* plasmid used to generate pan-neuronal expression was obtained from Addgene (catalog #137185).

There were two predicted TFBS within CNE3, as well as highly conserved regions flanking both of these TFBS. We used a GFP assay to test the functional consequence of the loss of these predicted TFBS from zebrafish CNE3. GFP constructs were made, injected, and imaged according to the above protocol.

The GFP construct that contains the CNE3 sequence from the electric fish species *B. gauderio* does not consistently drive GFP expression in any tissue and does not contain either of the predicted TFBS seen in other vertebrates. To test whether either of these TFBS could cause the gain of muscle expression, we also made GFP constructs containing the CNE3 sequence from *B. gauderio*, with both the E box and MCAT (and extended regions) sites added in at approximately the same location seen in other vertebrates (Genscript). Site-directed mutagenesis was performed according to the QuikChange II Kit (Agilent) to add the highly conserved 5′ (E box) and 3′ (MCAT) flanking sequences. GFP constructs were made, injected, and imaged according to the same established protocol.

Generating zebrafish CRISPR mutants
CNE3 zebrafish mutants were generated using CRISPR-Cas9. We designed two guide RNAs (Integrated DNA Technologies) to flank and target CNE3 for deletion in zebrafish (extended data, Fig. 4). Injections were performed in WT AB embryos at the single-cell stage according to the Alt-R CRISP-Cas9 system guidelines (Integrated DNA Technologies). At 3 months old, fish were fin clipped to extract DNA and genotyped using PCR. A 229–base pair (bp) indel completely deleting CNE3 was identified by sequencing (GSAF, The University of Texas at Austin) and designated *scn4aa^au108*. A single heterozygous P0 founder was then backcrossed to the AB strain.

Reverse transcription quantitative polymerase chain reaction
We quantified the expression level of *scn4aa* in CNE3 mutants and WT siblings using RT-qPCR. Trunk tissue from CNE3 mutant zebrafish embryos and WT siblings at 52 hours after fertilization was homogenized and stored in TRIzol at −80°C. Heads were lysed and DNA was extracted to verify CNE3 mutations using PCR. Trunk tissues from three individual embryos were combined together before RNA extraction. Total RNA was extracted from the trunks using the TRIzol RNA isolation protocol and RNA Clean & Concentrator kit (Zymo Research). Quality of total RNA extracted was assessed via NanoDrop, and the concentration was normalized to 50 ng/µl for each sample.

cDNA was synthesized from *scn4aa^au108* and WT total RNA samples using SuperScript II First-Strand Synthesis (Thermo Fisher Scientific) with random hexamer primers. RT-qPCR was performed using the QuantStudio 3 Real-Time PCR System (Thermo Fisher Scientific) with SYBR Green PCR Master Mix (Thermo Fisher Scientific). Beta-actin and glyceraldehyde-3-phosphate dehydrogenase genes were chosen as endogenous controls because of their stable expression. Expression levels were calculated via delta delta CT relative to both endogenous controls across all groups. Five biological replicates, each with three technical replicates, of both CNE3 mutant and WT samples were assessed.

Transgenics and microscopy methods for *B. gauderio*
The CNE3-GFP and Tol2 plasmids were produced with the Qiagen Miniprep Kit, and mRNA synthesis with the mMESSAGE mMACHINE SP6 Transcription Kit (Thermo Fisher Scientific). The transposase template was pCS2FA from the Tol2kit (*34*).

A mix of CNE3 construct (100 ng/µl), capped Tol2 mRNA (25 ng/µl), and 1% phenol red dye was injected into single-cell *B. gauderio* embryos obtained via in vitro fertilization, following the methods described by (*35*). In total, 80 embryos were injected. Initial embryo screening was done under a Leica MZ10F stereoscope: Of the 80 embryos, 18 showed a pattern of fluorescence. After 18 days, two GFP-expressing individuals were euthanized in 0.1% w/v MS-222, whole mounted in VectaShield, and imaged on an Olympus FV1000 (Fig. 3). After 5 months, another individual was euthanized; total body fluorescence was examined before samples of flank muscle and tail (containing EO) were removed for histological processing. Tissue samples were embedded in OTC and flash frozen in an isopentane bath submerged in liquid nitrogen, then cryo-sectioned to 20-µm thickness, and imaged on a Leica MZ10F stereomicroscope.

SUPPLEMENTARY MATERIALS
Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abm2970

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES
1. J. R. Gallant, L. A. O'Connell, Studying convergent evolution to relate genotype to behavioral phenotype. *J. Exp. Biol.* **223**, jeb213447 (2020).
2. S. Lavoué, M. Miya, M. E. Arnegard, J. P. Sullivan, C. D. Hopkins, M. Nishida, Comparable ages for the independent origins of electrogenesis in African and South American weakly electric fishes. *PLOS ONE* **7**, e36287 (2012).
3. H. H. Zakon, Y. Lu, D. Zwickl, D. Hillis, Sodium channel genes and the evolution of diversity in communication signals of electric fishes: Convergent molecular evolution. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 3675–3680 (2006).
4. M. S. Modrell, M. Lyne, A. R. Carr, H. H. Zakon, D. Buckley, A. S. Campbell, M. C. Davis, G. Micklem, C. V. H. Baker, Insights into electrosensory organ development, physiology and evolution from a lateral line-enriched transcriptome. *eLife* **6**, e24197 (2017).
5. K. Catania, The shocking predatory strike of the electric eel. *Science* **346**, 1231–1234 (2014).
6. C. W. Coates, R. Cox, W. Rosemblith, M. Brown, Propagation of the electric impulse along the organs of the electric eel, *Electrophorus electricus* (Linnaeus). *Fortschr. Zool.* **25**, 249–256 (1940).
7. H. Zakon, Y. Lu, M. Jost, Expansion of voltage-dependent Na+ channel gene family in early tetrapods coincided with the emergence of terrestriality and increased brain complexity. *Mol. Biol. Evol.* **28**, 1415–1424 (2011).
8. N. Patel, L. L. Smith, E. Faqeih, J. Mohamed, V. A. Gupta, F. S. Alkuraya, ZBTB42 mutation defines a novel lethal congenital contracture syndrome (LCCS6). *Hum. Mol. Genet.* **23**, 6584–6593 (2014).
9. D. Simkin, S. Bendahhou, Skeletal muscle Na channel disorders. *Front. Pharmacol.* **2**, 63 (2011).
10. G. Lopreato, Y. Lu, A. Southwell, N. S. Atkinson, D. M. Hillis, T. P. Wilcox, H. H. Zakon, Evolution and divergence of sodium channel genes in vertebrates. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7588–7592 (2001).
11. A. Novak, M. C. Jost, Y. Lu, A. D. Taylor, H. H. Zakon, A. B. Ribera, Gene duplications and evolution of vertebrate voltage-gated sodium channels. *J. Mol. Evol.* **63**, 208–221 (2006).
12. M. Arnegard, P. B. McIntyre, L. J. Harmon, M. L. Zelditch, W. G. R. Crampton, J. K. Davis, J. P. Sullivan, S. Lavoué, C. D. Hopkins, Sexual signal evolution outpaces ecological divergence during electric fish species radiation. *Am. Nat.* **176**, 335–356 (2010).
13. C. Paul, F. Kirschbaum, V. Mamonekene, R. Tiedemann, Evidence for non-neutral evolution in a sodium channel gene in African weakly electric fish (*Campylomormyrus*, Mormyridae). *J. Mol. Evol.* **83**, 61–77 (2016).
14. D. Arcila, G. Ortí, R. Vari, J. W. Armbruster, M. L. J. Stiassny, K. D. Ko, M. H. Sabaj, J. Lundberg, L. J. Revell, R. Betancur-R., Genome-wide interrogation advances resolution of recalcitrant groups in the tree of life. *Nat. Ecol. Evol.* **1**, 0020 (2017).

15. F. Kirschbaum, Myogenic electric organ precedes the neurogenic organ in apteronotid fish. *Naturwissenschaften* **70**, 205–207 (1983).

16. M. E. Arnegard, D. J. Zwickl, Y. Lu, H. H. Zakon, Old gene duplication facilitates origin and diversification of an innovative communication system—Twice. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 22172–22177 (2010).

17. S. D. Kraner, M. Rich, R. G. Kallen, R. L. Barchi, Two E-boxes are the focal point of muscle-specific skeletal muscle type 1 Na+ channel gene expression. *J. Biol. Chem.* **273**, 11327–11334 (1998).

18. S. D. Kraner, M. M. Rich, M. A. Sholl, H. Zhou, C. S. Zorc, R. G. Kallen, R. L. Barchi, Interaction between the skeletal muscle type 1 Na+ channel promoter E-box and an upstream repressor element. Release of repression by myogenin. *J. Biol. Chem.* **274**, 8129–8136 (1999).

19. K. Asakawa, M. L. Suster, K. Mizusawa, S. Nagayoshi, T. Kotani, A. Urasaki, Y. Kishimoto, M. Hibi, K. Kawakami, Genetic dissection of neural circuits by Tol2 transposon-mediated Gal4 gene and enhancer trapping in zebrafish. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 1255–1260 (2008).

20. P. Hasty, A. Bradley, J. H. Morris, D. G. Edmondson, J. M. Venuti, E. N. Olson, W. H. Klein, Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. *Nature* **364**, 501–506 (1993).

21. Y. Nabeshima, K. Hanaoka, M. Hayasaka, E. Esuml, S. Li, I. Nonaka, Y. I. Nabeshima, Myogenin gene disruption results in perinatal lethality because of severe muscle defect. *Nature* **364**, 532–535 (1993).

22. T. Yoshida, MCAT elements and the TEF-1 family of transcription factors in muscle development and disease. *Arterioscler. Thromb. Vasc. Biol.* **28**, 8–17 (2008).

23. A. Thompson, D. T. Infield, A. R. Smith, G. T. Smith, C. A. Ahern, H. H. Zakon, Rapid evolution of a voltage-gated sodium channel gene in a lineage of electric fish leads to a persistent sodium current. *PLoS Biol.* **16**, e2004892 (2018).

24. B. Ching, J. M. Woo, K. C. Hiong, M. V. Boo, W. P. Wong, S. F. Chew, Y. K. Ip, Voltage-gated Na+ channel isoforms and their mRNA expression levels and protein abundance in three electric organs and the skeletal muscle of the electric eel *Electrophorus electricus*. *PLOS ONE* **11**, e0167589 (2016).

25. J. R. Gallant, L. L. Traeger, J. D. Volkening, H. Moffett, P. H. Chen, C. D. Novina, G. N. Phillips Jr., R. Anand, G. B. Wells, M. Pinch, R. Güth, G. A. Unguez, J. S. Albert, H. H. Zakon, M. P. Samanta, M. R. Sussman, Genomic basis for the convergent evolution of electric organs. *Science* **344**, 1522–1525 (2014).

26. L. Traeger, J. D. Volkening, H. Moffett, J. R. Gallant, P.-H. Chen, C. D. Novina, G. N. Phillips Jr., R. Anand, G. B. Wells, M. Pinch, R. Güth, G. A. Unguez, J. S. Albert, H. Zakon, M. R. Sussman, M. P. Samanta, Unique patterns of transcript and miRNA expression in the South American strong voltage electric eel (*Electrophorus electricus*). *BMC Genomics* **16**, 243 (2015).

27. C. Aguilar, M. J. Miller, J. R. Loaiza, R. Krahe, L. F. De León, Mitogenomics of Central American weakly-electric fishes. *Gene* **686**, 164–170 (2019).

28. F. Alda, V. A. Tagliacollo, M. J. Bernt, B. T. Waltz, W. B. Ludt, B. C. Faircloth, M. E. Alfaro, J. S. Albert, P. Chakrabarty, Resolving deep nodes in an ancient radiation of neotropical fishes in the presence of conflicting signals from incomplete lineage sorting. *Syst. Biol.* **68**, 573–593 (2019).

29. C. R. Franchina, Ontogeny of the electric organ discharge and the electric organ in the weakly electric pulse fish *Brachyhypopomus pinnicaudatus* (Hypopomidae, Gymnotiformes). *J. Comp. Physiol. A* **181**, 111–119 (1997).

30. S. L. Hebert, C. Simmons, A. L. Thompson, C. S. Zorc, E. M. Blalock, S. D. Kraner, Basic helix–loop–helix factors recruit nuclear factor I to enhance expression of the NaV 1.4 Na+ channel gene. *Biochim. Biophys. Acta* **1769**, 649–658 (2007).

31. A. L. Thompson, G. Filatov, C. Chen, I. Porter, Y. Li, M. M. Rich, S. D. Kraner, A selective role for MRF4 in innervated adult skeletal muscle: Na(V) 1.4 Na+ channel expression is reduced in MRF4-null mice. *Gene Expr.* **12**, 289–303 (2005).

32. R. Fricke, W. N. Eschmeyer, R. van der Laan (eds.) Eschmeyer's catalog of fishes. Genera, species, references (2021); www.calacademy.org/scientists/projects/eschmeyers-catalog-of-fishes.

33. M. Pinch, R. Güth, M. P. Samanta, A. Chaidez, G. A. Unguez, The myogenic electric organ of *Sternopygus macrurus*: A non-contractile tissue with a skeletal muscle transcriptome. *PeerJ* **4**, e1828 (2016).

34. K. M. Kwan, E. Fujimoto, C. Grabher, B. D. Mangum, M. E. Hardy, D. S. Campbell, J. M. Parant, H. J. Yost, J. P. Kanki, C. B. Chien, The Tol2kit: A multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev. Dyn.* **236**, 3088–3099 (2007).

35. S. J. Constantinou, L. Nguyen, F. Kirschbaum, V. L. Salazar, J. R. Gallant, Silencing the spark: CRISPR/Cas9 genome editing in weakly electric fish. *J. Vis. Exp.* **152**, 10.3791/60253 (2019).

Science Advances

Divergent cis-regulatory evolution underlies the convergent loss of sodium channel expression in electric fish

Sarah LaPotinMary E. SwartzDavid M. LueckeSavvas J. ConstantinouJason R. GallantJohann K. EberhartHarold H. Zakon

View the article online
https://www.science.org/doi/10.1126/sciadv.abm2970
Permissions
https://www.science.org/help/reprints-and-permissions

Use of this article is subject to the Terms of service