PLOS BIOLOGY

SHORT REPORTS

Rapid evolution of a voltage-gated sodium channel gene in a lineage of electric fish leads to a persistent sodium current

Ammon Thompson[1,2☯¤]*, Daniel T. Infield[3☯], Adam R. Smith[4], G. Troy Smith[4], Christopher A. Ahern[3], Harold H. Zakon[1,2]

1 Department of Integrative Biology, The University of Texas, Austin, Texas, United States of America, **2** Department of Neuroscience, The University of Texas, Austin, Texas, United States of America, **3** Department of Molecular Physiology and Biophysics, Iowa Neuroscience Institute, The University of Iowa, Iowa City, Iowa, United States of America, **4** Department of Biology and Center for the Integrative Study of Animal Behavior, Indiana University, Bloomington, Indiana, United States of America

☯ These authors contributed equally to this work.
¤ Current address: Department of Evolution and Ecology, The University of California, Davis, Davis, California, United States of America
* ammthompson@ucdavis.edu

Citation: Thompson A, Infield DT, Smith AR, Smith GT, Ahern CA, Zakon HH (2018) Rapid evolution of a voltage-gated sodium channel gene in a lineage of electric fish leads to a persistent sodium current. PLoS Biol 16(3): e2004892. https://doi.org/10.1371/journal.pbio.2004892

Academic Editor: Laurence Hurst, University of Bath, United Kingdom of Great Britain and Northern Ireland

Received: November 10, 2017

Accepted: February 21, 2018

Published: March 27, 2018

Data Availability Statement: RNA sequencing reads and assembled transcriptomes available at NCBI under Bioproject accession PRJNA432820.

Funding: NSF (grant number DEB-1311521). NSF (grant number IOS-1557857). NSF (grant number IOS-0950721). NIH (grant number HL007121). NIH (grant number GM106569). NIH (grant number GM087519). NIH (grant number EIA22180002). NIH (grant number T32049336). The funders had no role in study design, data

Abstract

Most weakly electric fish navigate and communicate by sensing electric signals generated by their muscle-derived electric organs. Adults of one lineage (Apteronotidae), which discharge their electric organs in excess of 1 kHz, instead have an electric organ derived from the axons of specialized spinal neurons (electromotorneurons [EMNs]). EMNs fire spontaneously and are the fastest-firing neurons known. This biophysically extreme phenotype depends upon a persistent sodium current, the molecular underpinnings of which remain unknown. We show that a skeletal muscle–specific sodium channel gene duplicated in this lineage and, within approximately 2 million years, began expressing in the spinal cord, a novel site of expression for this isoform. Concurrently, amino acid replacements that cause a persistent sodium current accumulated in the regions of the channel underlying inactivation. Therefore, a novel adaptation allowing extreme neuronal firing arose from the duplication, change in expression, and rapid sequence evolution of a muscle-expressing sodium channel gene.

Author summary

The electrical properties of excitable cells, such as those in muscle and nervous tissue, were enabled in large part by the evolution of voltage-gated ion channel genes. The regulated conduction of ions through these channels results in the propagation of electrical signals, facilitating communication between cells. Here, we investigated how voltage-gated sodium (Na_v) channels contributed to the evolution of a novel electric organ system in the Apteronotids—a lineage of weakly electric fish. This organ is developmentally derived from motor neurons and used for communication between individual fish, as well as for probing their nocturnal environment. We used transcriptomic data to show that the gene encoding a

collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: BSR, Branch-sites REL; C, carboxy-terminus; COI, cytochrome c oxidase subunit I; CryoEM, cryo-electron microscopy; CytB, cytochrome B; D, aspartate; dN/dS, nonsynonymous replacements per nonsynonymous site divided by the synonymous changes per synonymous site; EMN, electromotorneuron; EO, electric organ; F, phenylalanine; H, histidine; hNa_v, human voltage-gated sodium; I, isoleucine; I_{NAP}, persistent sodium current; L, leucine; M, methionine; miRNA, micro-RNA; MYA, million years ago; N, amino-terminus; Na_v, voltage-gated sodium; PAML, phylogenetic analysis by maximum likelihood; Phe, phenylalanine; qPCR, quantitative PCR; RAG2, recombination activating gene 2; REL, random effects likelihood; TPM, transcripts per million; WT, wild-type.

broadly conserved muscle-specific sodium channel was duplicated in an ancestral fish. One duplicated gene copy subsequently gained expression in the spinal cord, where the electric organ is located. Through evolutionary analysis and biophysical experiments, we demonstrate that sequence changes in this new sodium channel transformed its function to cause novel electrical properties that can facilitate spontaneous high-frequency action potentials. This study shows that duplicate genes can gain highly novel expression patterns and quickly adapt to contribute to the phenotypic evolution of novel organ systems.

Electrocommunication convergently evolved within the distantly related South American and African weakly electric fishes approximately 100 million years ago (MYA) [1], enabling them to detect their environment and communicate with each other through the generation and sensation of electric signals. These species exhibit a highly derived phenotypic trait—the electric organ—which in both clades is developmentally derived from muscle cells. Electric organs have repurposed muscle action potential genes (voltage-gated ion channels) to generate the electric signal [2–4]. Electric organ signals have subsequently been shaped by both natural and sexual selection [5–10], resulting in diverse signaling patterns in both clades of electric fish [8,10].

The voltage-gated ion channels in electric organs are an ideal system for investigating the molecular mechanisms underpinning the evolution of a phenotypic trait. Voltage-gated ion channels are the primary molecular machinery producing electrical signals in excitable tissues. Well-developed electrophysiology assays make it possible to gain a mechanistic understanding of ion channels at single amino acid resolution. There is a clear and direct link between ion channel physiology and the attributes of electric organ signals that can be measured in the field and in the lab. Therefore, the diversity of electric organ signals provides an opportunity to study—in detail—natural experiments whereby ion channels evolve to directly influence a rapidly evolving phenotype.

Some species emit electric organ signals that are biophysically extreme [9,11–13]. One family within South American electric fish, the Apteronotidae, generates the highest electric organ discharge frequencies of any electric fish (depending on species, 650–1,800 Hz) [14–16]. Sustained firing rates of 1 kHz or more are difficult for neurons to attain, yet these rates are maintained continuously throughout the fish's lifetime. Studies of two Apteronotid species (*Apteronotus albifrons* and *A. leptorhynchus*) illustrate the unique anatomy that underlies the family's extreme physiology. Early in development, Apteronotids lose their muscle-derived (myogenic) electric organ [17], and the axons of neurons that innervate the electric organ (electromotorneurons [EMNs]) undergo substantial morphological transformation to form a neuronally derived (neurogenic) electric organ [18] (S1 Fig). The EMNs of *Apteronotus* are the fastest-firing neurons known [16] and can generate these signals spontaneously, independent of EMN inputs originating in the brainstem [19].

Specializations of voltage-gated ion channels likely contributed to the evolution of neurogenic electric organs found in *Apteronotus*. Spontaneously firing neurons can be driven by a persistent sodium (Na^+) current (I_{NAP}) [20]. I_{NAP} derives from sodium channels that fail to fully inactivate after each action potential—inactivation referring to the rapid closing of sodium channels, which halts the inward Na^+ current and terminates the action potential. A persistent sodium current can cause a strong depolarization sufficient to rapidly elicit another action potential. Electrophysiological and pharmacological experiments suggest that high spontaneous rates of *Apteronotus* EMN firing are driven by I_{NAP} [20], implicating modification of voltage-gated sodium (Na_v) channels in the production of this extreme physiology.

Gene duplication within the Na_v channel family appears to have played an important role in the evolution and diversification of electrocommunication and electrolocation. Early in teleost evolution, because of a whole genome duplication, the vertebrate muscle–expressing sodium channel gene (*scn4a*) duplicated into *scn4aa* and *scn4ab*, both expressed in muscle in most teleost fish [2,21]. In electric fish, *scn4aa* was independently co-opted as an electric organ–specific sodium channel in the South American and African clades, where its amino acid sequence rapidly diverged, mirroring the diversity of electric signals exhibited within each clade [3,4,22,23].

Here, we show that an additional duplication at the same muscle-expressing sodium channel locus (*scn4a*) occurred within a sublineage of Apteronotidae that includes the genus *Apteronotus*. Within approximately 2 million years, one of the resulting duplicates began expressing in the spinal cord, where the neurogenic electric organ is located. This duplicate rapidly evolved several amino acid substitutions in two structurally proximate domains of the channel that interact to mediate inactivation and that have otherwise been conserved in all vertebrate sodium channels. When combinations of these substitutions are introduced into a human sodium channel, they produce a significant I_{NAP}. This is the first sodium channel that generates a physiological persistent current without the aid of auxiliary proteins [24–26]. These findings indicate two phenomena: (1) paralogs of the same muscle-type Na_v channel have been implicated in the evolution of three separate electrocommunication systems within teleost fishes and (2) a molecular adaptation for the extremely high firing frequencies observed in an electric organ derived from neural cells was in part mediated by a new gene that originated from the duplication of a muscle-expressing gene.

Results

A novel Apteronotid Na_v channel gene duplication

To identify Na_v channels expressed in the EMNs, we generated transcriptomes from the posterior spinal cords (where EMNs are abundant) and trunk muscle of adults of 3 species of Apteronotids (*A. albifrons*, *A. leptorhynchus*, and *Parapteronotus hasemani*), one myogenic electric gymnotiform (*Eigenmannia virescens)*, and a catfish (*Ictalurus punctatus*) as a nonelectric outgroup. We also obtained expression data from another myogenic gymnotiform, the electric eel (*Electrophorus electricus*), from a previous study [27].

Data reported here confirm previous results; both *scn4a* paralogs are expressed in catfish muscle, while only *scn4ab* is present in the muscle of the myogenic electric fish *E. virescens* and in *E. electricus* (Figs 1 and S2). Unlike myogenic gymnotiforms but like nonelectric teleosts, *scn4aa* is significantly expressed in the muscle of Apteronotids (also confirmed here by quantitative PCR [qPCR], S3 Fig). We detected virtually no expression of *scn4aa* in *E. virescens* muscle, which confirms previous research that showed that *scn4aa* is electric organ specific and is not expressed in the muscle of myogenic gymnotiforms [2–4,27].

We also discovered a novel gene duplication in *A. albifrons*: 2 *scn4ab* paralogs—*scn4ab1* and *scn4ab2*—that are about 95% identical in amino acid sequence, suggesting a recent duplication event. We confirmed that these are not assembly artifacts by cloning them by PCR from muscle mRNA and amplifying them from genomic DNA across several exons (equivalent to zebrafish scn4ab exons 22–24), with the intervening introns showing even greater divergence (S4 Fig). These 2 paralogs were also found in *A. leptorhynchus* and *P. hasemani*. *Scn4ab2* showed expression in muscle, and *scn4ab1* expressed in muscle and spinal cord. In contrast to other Gymnotiform electric fish species which only express 1 subtype in muscle (*scn4ab*), these 3 Apteronotid species show significant expression of 3 sodium channels (*scn4aa*, *scn4ab1*, and *scn4ab2*).

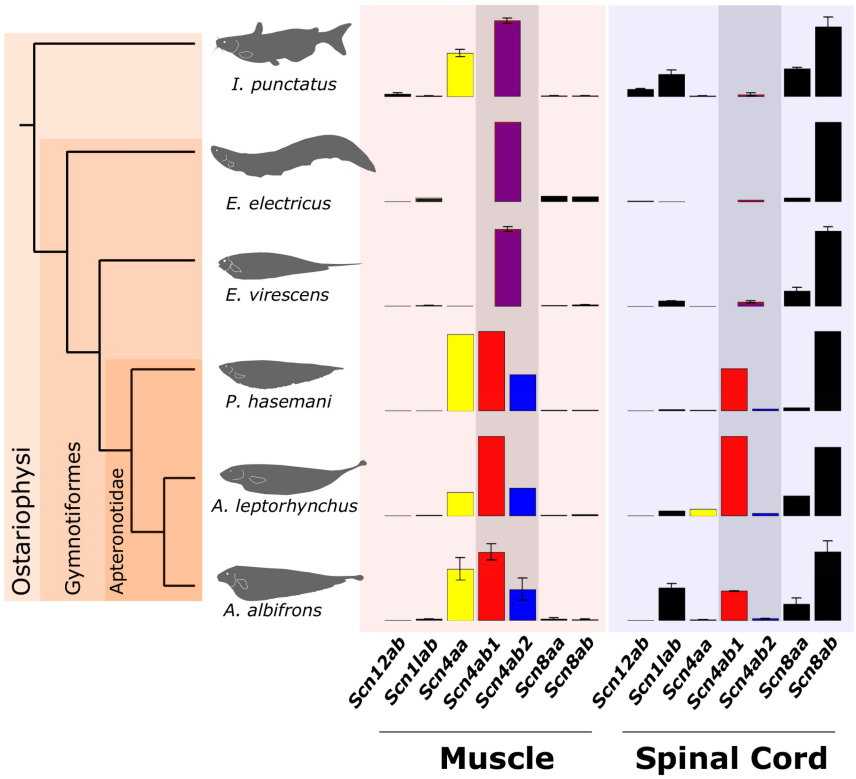

Fig 1. Expression of Na_v channel genes in muscle and spinal cord of Apteronotids (*A. albifrons* N = 3, *A. leptorhynchus* N = 1, *P. hasemani* N = 1), other Gymnotiforms (*E. virescens* N = 3, *E. electricus* N = 1), and the channel catfish (*I. punctatus* N = 3) as an outgroup, determined by NextGen sequencing. Data shown are the mean TPM of transfrags with the highest TPM for each gene. Whiskers show 1 standard deviation for species with 3 biological replicates. *scn4aa* (yellow) is expressed in muscle in most fishes except for those with myogenic electric organs, where *scn4aa* expression is lost from muscle and compartmentalized in the EO [25]. *Scn4ab* (purple) is expressed in muscle in all teleosts, including electric fish. *Scn4ab* duplicated into *scn4ab1* (red) and *scn4ab2* (blue) in the Apteronotids. *Scn4b1* expresses in muscle and spinal cord, which is a novel site of expression for the *scn4a* type gene in vertebrates. *Scn4ab2* still expresses in muscle only. The tree is a consensus tree based on Fig 2. All data points shown in S2 Fig. Figure data included in S1 Data. EO, electric organ; Na_v, voltage-gated sodium; TPM, transcripts per million.

https://doi.org/10.1371/journal.pbio.2004892.g001

Rapid evolution of new Nav channel gene

In catfish, *E. virescens* and *E. electricus*, spinal cord expression is dominated by *scn8ab* and, to a lesser degree, *scn8aa* and *scn1lab*, as reported for zebrafish [21,28] and their orthologs in mammalian spinal motorneurons [29,30]. While those genes are also expressed in the Apteronotid spinal cord, *scn4ab1* makes up 21%–45% of total sodium channel expression there. The expression of *scn4ab1* in *A. albifrons* spinal cord was confirmed by qPCR (S3 Fig). This is the first observation of a muscle-typical Na_v channel gene expressing in spinal cord in any vertebrate.

A gene tree inferred from a nucleotide alignment of gymnotiform and non-gymnotiform *scn4ab* sodium channels with additional sequences from 2 more basal and widely separated Apteronotids ("*A.*" *bonapartii*, *Adontosternarchus devenanzi*) (S5 Fig) shows strong support for the *scn4ab1* and *scn4ab2* paralogs forming a monophyletic clade in a derived lineage within Apteronotidae called the Apteronotini [31]. A time-calibrated phylogeny of the Apteronotidae (Fig 2A) estimates the duplication of *scn4ab* at the divergence of Apteronotini from other Apteronotids at approximately 14.5 MYA (min = 5.36, max = 23.66) with most amino acid substitutions fixed by approximately 12.4 MYA (min = 3.66, max = 21.21),

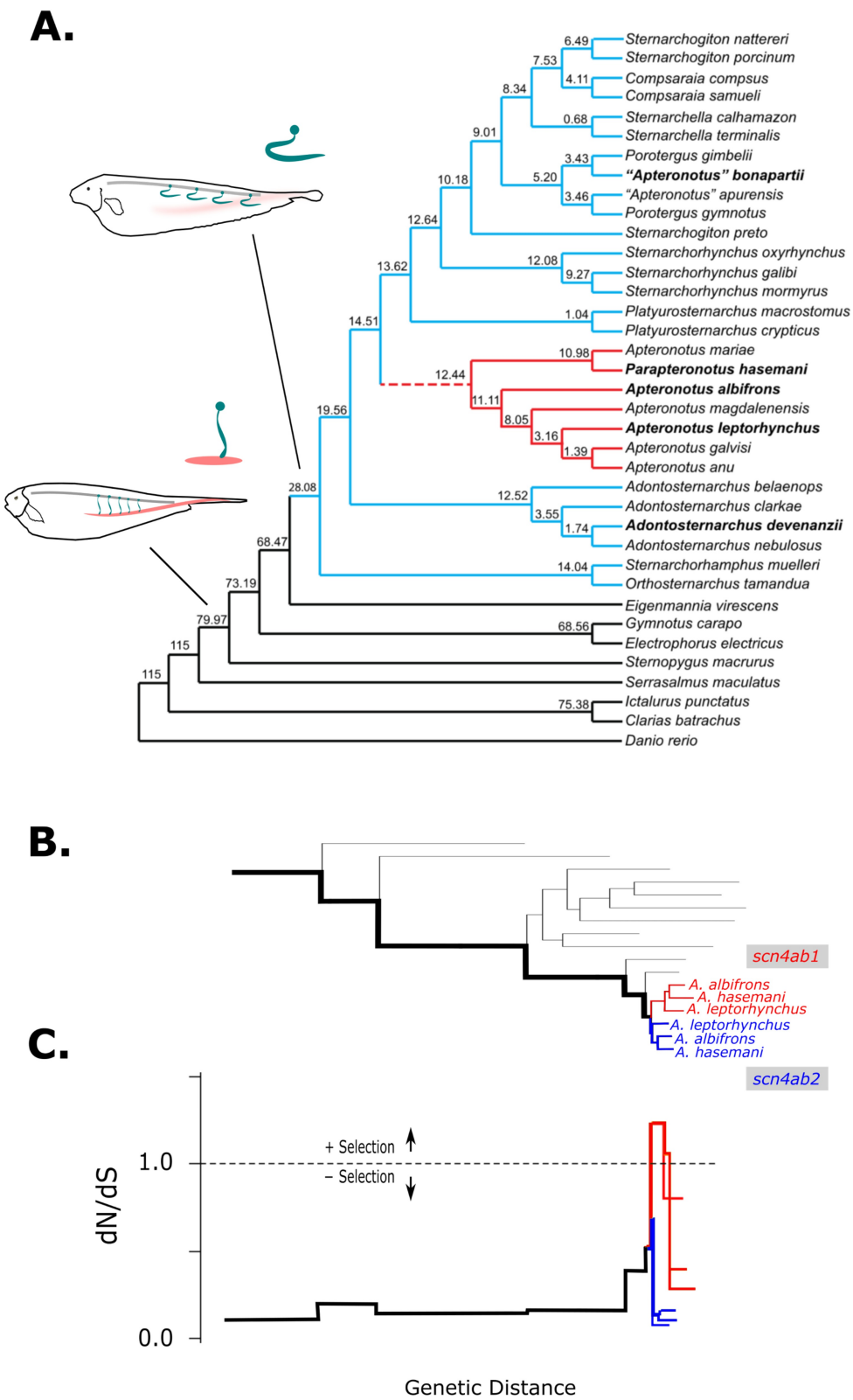

Fig 2. Timing of duplication and neofunctionalization of *scn4ab* paralogs. (A) Chronogram of electric species and relatives with diagrams of myogenic and neurogenic fish, and when they evolved, indicated. *Scn4ab* gene duplication and amino acid divergence occurred in a 2 MY window (dotted red line). This phylogeny is based on COI, CytB, and

RAG2 genes. Molecular clock analysis shows that the *scn4ab* gene duplication occurred within the family Apteronotidae (light blue) after the last common ancestor of *A. devenanzi* and *"A" bonapartii* (approximately 14.5 MYA) and before the divergence of the species within the clade Apteronotini (red, approximately 12.4 MYA), which show the amino acid substitutions in the domain 4 S4–S5 linker. Species from which *scn4ab* genes have been cloned (and illustrated in S6 Fig) are in bold. Following gene duplication, *scn4ab1* shows an episodic burst of positive selection. (B) Gene tree for *scn4ab* in Gymnotiforms and closest available relatives. The red branches indicate *scn4ab1* clade and blue for *scn4ab2* clade in the 3 Apteronotini species. (C) Branchwise maximum likelihood estimates of dN/dS using the BSR method [32] from the most recent common ancestor with the zebrafish, *Danio rerio*, leading to the inferred duplication event (thick black branches in panel B). Following duplication, the *scn4ab1* clade (red) shows elevated dN/dS relative to its muscle-specific paralog (blue) in all Apteronotini species, with the branch immediately following the duplication event having a dN/dS >1. Gene accessions are in S1 Table. Figure data included in S1 Data. BSR, Branch-sites REL; COI, cytochrome c oxidase subunit I; CytB, cytochrome B; dN/dS, nonsynonymous replacements per nonsynonymous site divided by the synonymous changes per synonymous site; MY, million-year; MYA, million years ago; RAG2, recombination activating gene 2; REL, random effects likelihood.

https://doi.org/10.1371/journal.pbio.2004892.g002

preceding the divergence of species within the Apteronotini. Therefore, the duplication and divergence of this gene likely occurred within an approximate 2 million–year window. Branch-specific nonsynonymous replacements per nonsynonymous site divided by the synonymous changes per synonymous site (dN/dS) ratios estimated by maximum likelihood [32] support an episodic burst of positive selection on *scn4ab1* immediately after duplication followed by an elevated dN/dS ratio within that clade (Fig 2B and 2C). The other paralog, *scn4ab2*, which shows muscle-specific expression like other vertebrate *scn4a* sodium channels, shows coding sequence patterns consistent with purifying selection, suggesting that this gene maintained its ancestral function along with ancestral expression in muscle.

We also estimated dN/dS using phylogenetic analysis by maximum likelihood (PAML) [33]. We compared a model in which the root branch of the *scn4ab1* clade following the duplication had a unique dN/dS ratio versus a model in which all branches in the tree had the same ratio. A likelihood ratio test supports the more complex model with 2 ratios, which estimates that *scn4ab1* evolved by positive selection soon after duplication (dN/dS = 1.58), while the rest of the branches in the tree show more conservative evolution with dN/dS = 0.090 ($2\Delta L = 38.8$, df = 1, $p < 0.0001$). We were unable to find statistically significant evidence of evolution by positive selection in the root branch of the *scn4ab2* clade using this approach. This analysis further supports *scn4ab1* neofunctionalized soon after duplication.

Na_v channels comprise 4 repeating domains (D1–D4), each of which has 6 membrane-spanning helices (S1–S6) (Figs 3A–3C and S6) [34]. The S4 helices in each domain are displaced by membrane depolarization, resulting in a conformation change in the channel that allows an inward flow of Na^+ ions [34]. Within milliseconds of activation, channel conduction is spontaneously terminated via fast inactivation [34]. The molecular mechanism of this process is unresolved; however, functional experiments have identified several critical molecular features. A hydrophobic triplet of the amino acids isoleucine, phenylalanine, and methionine (IFM) in the intracellular loop between D3 and D4 [35] (the so-called "inactivation particle") is required for fast inactivation; this motif likely moves in response to membrane depolarization [36], and in the inactivated state, it may interact with a "receptor" containing amino acid side chains in the intracellular S4–S5 linkers of D3 [37] and D4 [38]. Naturally occurring mutations in these regions produce an I_{NAP} that is implicated in neurological and muscular diseases [39–41].

Fig 3 shows the relative density of amino acid substitutions of the Apteronotid *A. albifrons scn4ab* paralogs and a distant relative, the *scn4ab* ortholog, in the zebrafish, *D. rerio*. Using a minimum mutation parsimony criterion, amino acid changes and indels were mapped onto each of the 3 branches connecting the genes (Fig 3A). For example, if zebrafish *scn4ab* and *A. albifrons scn4ab2* have the same amino acid residue at a coding position and *scn4ab1* is

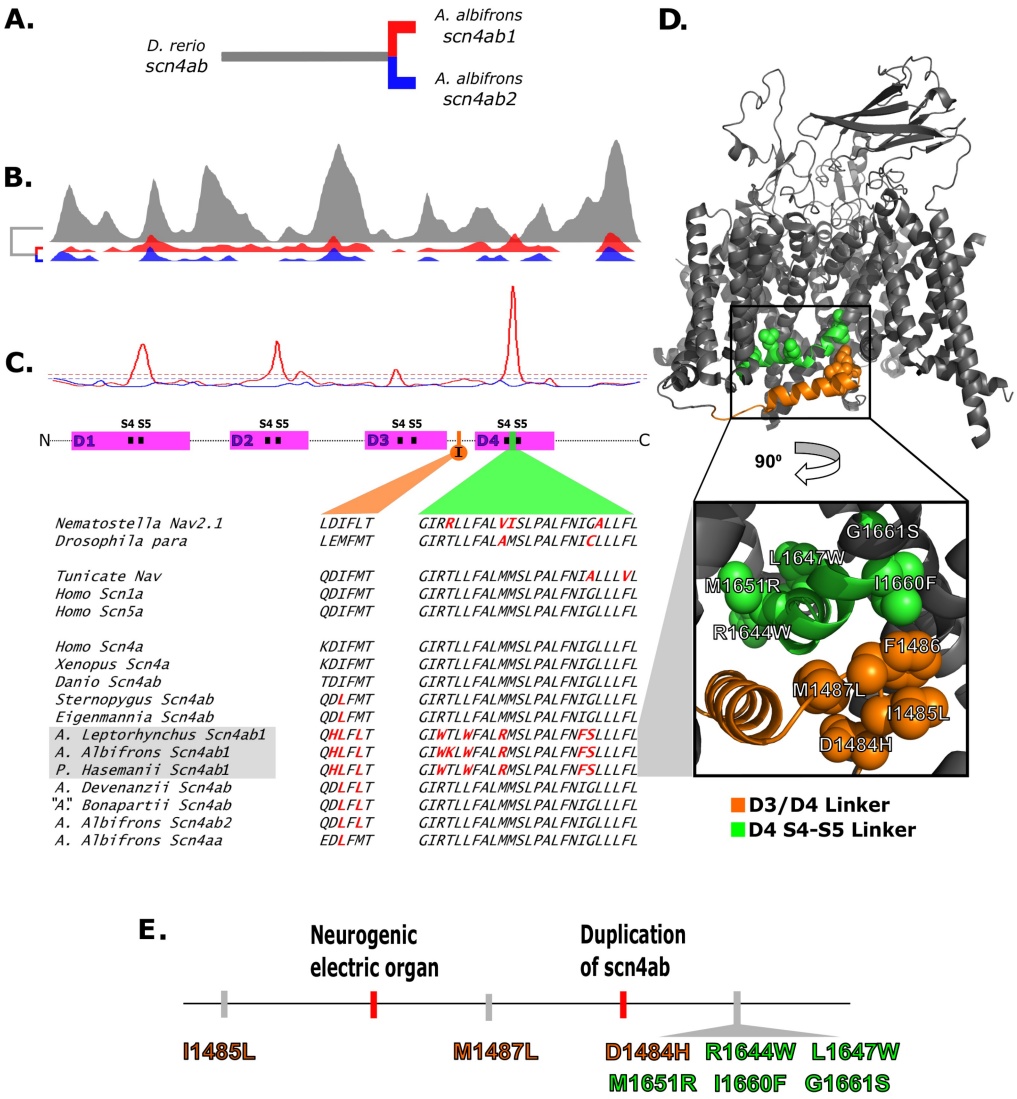

Fig 3. Relative density of amino acid replacements in *scn4ab1* and *scn4ab2*. (A) The orthologs used in the analysis, where the 2 paralogs in *A. albifrons* were compared to the much more distantly related ortholog in *D. rerio*. (B) Density plots of amino acid substitutions along the sodium channel (N to C) that are specific to each gene. (C) The relative density of substitutions specific to *scn4ab1* and *scn4ab2* (red and blue, respectively). The relative density is the ratio of substitution density at each position in the protein to the added density in the other 2 branches. The dotted lines show conservative significance thresholds for the same relative density measure from simulating random distributions of the amino acid substitutions in the 3 gene sequences. The 4 domains of a typical sodium channel are shown as magenta rectangles, with the transmembrane domains near the highest peak densities indicated in the intracellular linker between S4 and S5 (S4–S5) and the inactivation particle indicated with an orange circle and denoted "I." The bottom protein alignments are of the inactivation particle (orange) and one of the inactivation receptors in domain 4 (green) that shows the highest relative substitution signal in *scn4ab1*. *Scn4ab1* exhibits several amino acid substitutions that are highly conserved among chordates in both protein structures. Alignments of all human and zebrafish Na_v channels are shown in S6 Fig. (D) Structural view of the $Na_v1.4$-β1 complex from electric eel with positions of Apteronotid substitution represented as spheres [42]. Residue numbering is from the human cardiac sodium channel, $hNa_v1.5$, the background into which mutations were made for electrophysiology in this study. The inset is a magnified view of the putative binding site for the IFM inactivation particle in the D3–D4 loop. (E) Evolutionary sequence of amino acid replacements in the inactivation machinery. C, carboxy-terminus; hNa_v, human voltage-gated sodium; N, amino-terminus; F, phenylalanine; I, isoleucine; M, methionine.

https://doi.org/10.1371/journal.pbio.2004892.g003

different, then an amino acid substitution is assumed to have happened in *scn4ab1* at that position after duplication. With this procedure, the density of substitutions was measured along

the length of the 3 proteins (Fig 3B). The relative density of amino acid replacements was measured as the ratio of substitution density in an ortholog to the combined density of replacements of the other 2 proteins at the same position (Fig 3C). Simulations in which substitutions were randomly distributed along sequences were used to generate empirical null distributions of branch-specific substitution densities. A threshold that is reached in <99% of simulations was derived for each protein. Every site in the *scn4ab2* sequence is well below threshold. However, *scn4ab1* greatly exceeds this threshold in 3 locations, each around the S4 and S5 transmembrane regions within domains 1, 2, and 4 (Fig 3C). The strongest signal shows a nearly 10-fold higher density of amino acid substitutions than threshold in the aforementioned D4 S4–S5 linker within the putative inactivation particle receptor. These substitutions accumulated within approximately 2 million years of duplication (Fig 2) despite the strong conservation of these sites across the Na_v channel gene family of vertebrates, spanning approximately 550 million years of evolution [43] (S6 Fig).

While this paper was in preparation, a cryo-electron microscopy (CryoEM) structure of the skeletal muscle Na_v channel of *E. electricus*, the electric eel, was reported [42]. This is the first high-resolution structure of a canonical eukaryotic Na_v channel. The state of the channel is not conclusively determined, but it may represent a pre-inactivated conformation wherein the pore is trapped open by an unresolved detergent molecule. This structure suggests extensive interactions between the D3–D4 loop and D4 S4–S5 linker (Fig 3D, green and orange regions). Specifically, the structure predicts intimate association of Apteronotini substitutions in the hydrophobic inactivation particle with those within the D4 S4–S5 linker (Fig 3D inset). Of the substitutions in the D4 S4–S5 linker, 3 are located within a leucine zipper-like motif that has been previously proposed to interact with some part of the D3–D4 loop during fast inactivation [38]. The other two, a highly conserved isoleucine/glycine pair (I1660/G1661, which are F1660/S1661 in the Apteronotini), directly appose the critical phenylalanine (Phe) in the IFM inactivation particle (Fig 3D, inset).

None of the other spinal cord–expressing Na_v channels in the Apteronotini (scna8aa, scn8ab, scn1Lab) have amino acid substitutions in these parts of the channel.

Amino acid substitutions in the channel generate a persistent sodium current

Because the mechanism of inactivation and its structural determinants are highly conserved among animal Na_v channels, we made use of an available human $Na_v1.5$ sodium channel expression construct, engineered the amino acid substitutions into this channel, and expressed these channels in *Xenopus* oocytes. We predicted that the Apteronotini substitutions would generate an I_{NAP}. Wild-type (WT) $hNa_v1.5$ channels had no apparent I_{NAP} (Fig 4A). Compared to vertebrates in general, Apteronotid *scn4ab* (protein name; $Na_v1.4b$) channels have substitutions in the inactivation particle (IFM → LFL), with that of the Apteronotini *scn4ab1* having 1 further substitution (DIFM → HLFL). None of these substitutions generated an I_{NAP} on their own (Fig 4C, blue and orange traces). The substitution of the complete set of the 5 amino acids observed in the Apteronotini *scn4ab1* channel D4 S4–S5 linker into the WT $hNa_v1.5$ linker (Fig 3) caused a large I_{NAP} (Fig 4B). At none of these positions do single alanine mutations generate persistent currents in the neuronal sodium channel $Na_v1.2$ [38]. Therefore, the specific chemical nature and/or combination of mutations in this domain of *scn4ab1* are necessary for the physiological persistent current.

The addition of the partial (LFL, found in all Apteronotids) or complete (HLFL, found in Apteronotini) inactivation particle substitutions to the Apteronotini D4 S4–S5 linker substitutions modified the amount of I_{NAP} such that the combination that most resembled the

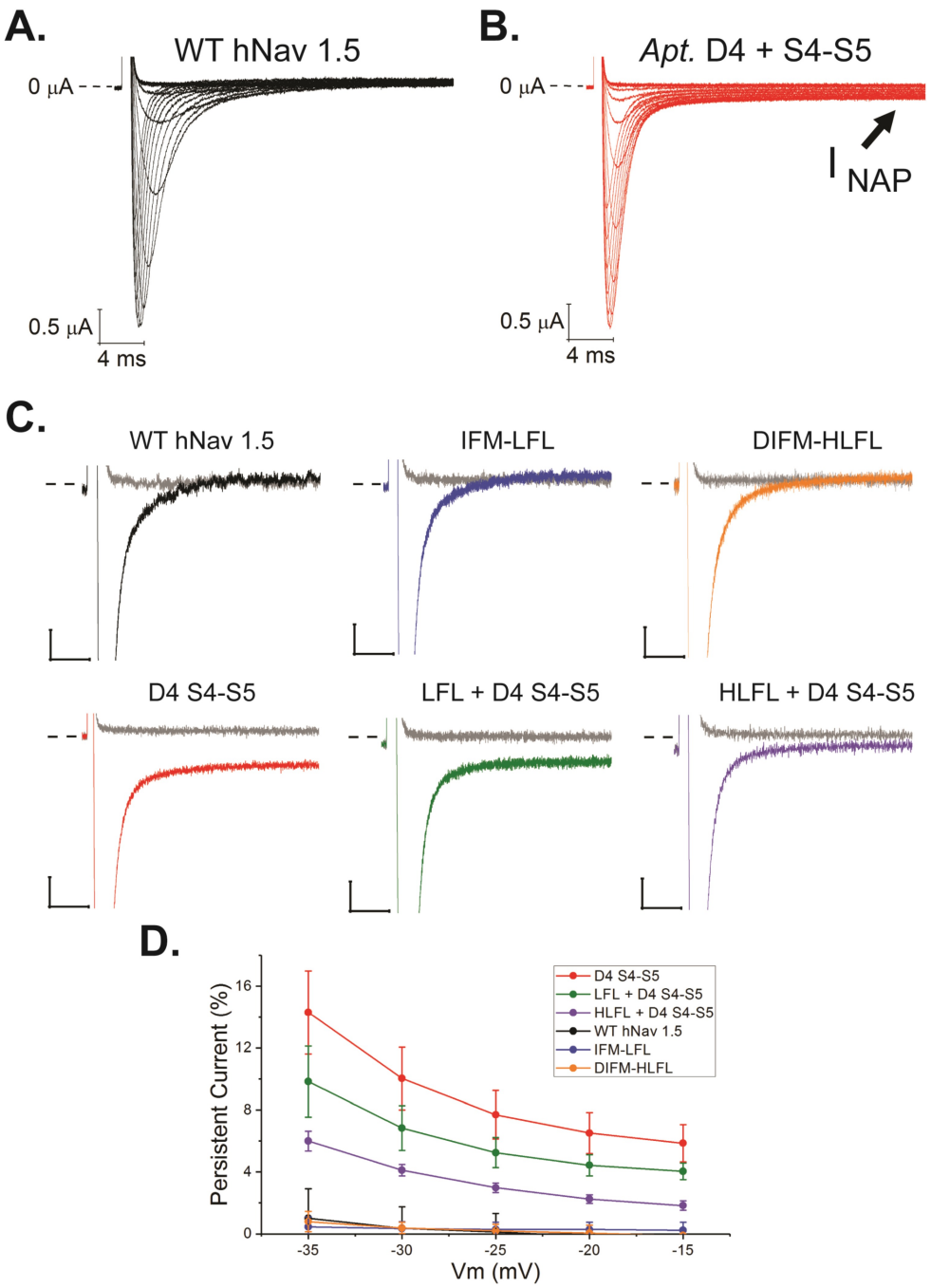

Fig 4. Apteronotini-specific $Na_v1.4b1$ (*scn4ab1*) substitutions generate a persistent current in the human cardiac $Na_v1.5$. (A) WT $hNa_v1.5$ currents elicited by depolarizations in 5 mV steps from −80 mV to +15 mV. (B) Currents generated by an $hNa_v1.5$ variant with all 5 Apteronotini-specific substitutions (see Fig 3) in the D4 S4–S5 linker. Note the persistent (non-inactivating) current at the end of the pulse. (C) Normalized exemplar traces of late currents for $hNa_v1.5$ mutants. To facilitate comparisons among variants, current responses to steps to −60 mV (gray) and −20 mV (colors) are shown. Scale bars are 5% of peak current (vertical) and 5 milliseconds of time (horizontal), and dashed lines indicate the point of 0 current. (D) Quantification of percentage of persistent current for each variant over a range of membrane potentials. Error bars are ± 1 standard deviation. $N \geq 5$ for each variant, in S4 Table. For all D4 S4–S5 substitution-containing variants, percentage of persistent current was significantly different ($p < 0.01$) as compared to WT $hNa_v1.5$. Figure data included in S2 Data. D4 S4–S5, R1644W/L1647W/M1651R/I1660F/G1661S; DIFM-HLFL, D1484H/I1485L/M1487L; hNa_v, human voltage-gated sodium; IFM-LFL, I1485L/M1487L; WT, wild-type.

https://doi.org/10.1371/journal.pbio.2004892.g004

Apteronotid *scn4ab1* channel generated about 6% maximum I_{NAP} (at −35 mV). Importantly, I_{NAP} activated at more hyperpolarized voltages (approximately −60 mV) and was a larger relative percentage of the total current (S7 Fig) in the range from −60 to −40 mV. The channel also showed less steady-state inactivation than the WT $hNa_v1.5$ current in this voltage range (S8 Fig). Electrophysiological recordings of spontaneously firing Apteronotid EMNs show that these neurons never completely return to resting potential between action potentials but, instead, briefly reach −40 to −50 mV before the next action potential [19]. This is the voltage range in which I_{NAP} produced by the inactivation apparatus of the Apteronotini *scn4ab1* channels is maximal. Furthermore, the recovery from inactivation, which influences a neuron's maximum firing frequency, is faster in $hNa_v1.5$ (HLFL + D4 S4–S5) than the WT $hNa_v1.5$, but only at the earliest (submillisecond) recovery interval (S9 Fig).

Discussion

Just as the diversification of voltage-gated ion channels contributed to diversification of animal nervous systems, the evolution of novelty in the Na_v channel gene family led to diversification of electric organ signals [24,26,44]. In both major clades of electric fish (Gymnotiformes, Mormyroidea), a muscle-specific Na_v channel gene (*scn4a*) that duplicated in the whole-genome duplication at the origin of teleosts retained muscle expression for approximately 100 million years before it was convergently recruited into the electric organ at the origin of both groups of electric fish. A clade of electric fish within the South American Gymnotiformes exhibits another more recent duplication at one of the muscle-expressing Na_v channel loci; this time, *scn4ab* duplicated to *scn4ab1* and *scn4ab2*. This lineage of electric fish is specialized for high-frequency discharges from a unique electric organ composed of the axons of EMNs. Shortly after duplication, *scn4ab1* gained expression in the spinal cord, where the EMN electric organ is located, and it shows comparative sequence patterns consistent with having evolved under positive selection (Fig 2).

The EMNs of the neurogenic electric organ of some Apteronotids, including those that contain *scn4ab1*, fire spontaneously and at the fastest rate known in any animal neuron. Our molecular evolution (Fig 3) and biophysical (Fig 4) analyses suggest that several amino acid substitutions within the inactivation particle and particle receptor lead to the generation of I_{NAP} in *scn4ab1*. The magnitude of I_{NAP} we measured in $hNa_v1.5$—that was specifically attributable to the Apteronotini substitutions—was similar in magnitude to the I_{NAP} observed in spontaneously firing mammalian Purkinje neurons, wherein I_{NAP} equal to only a small percentage of the peak current [45,46] is sufficient to drive spontaneous firing [47,48]. Because pharmacological and electrophysiological data from Apteronotus EMNs indicate the presence of a persistent sodium current, this strongly supports the origination of *scn4ab1* as an important molecular contributor to the evolution of EMN spontaneous high-frequency electric signals in the Apteronotini [3]. It is, however, important to note that the I_{NAP} measured in spontaneously firing Purkinje neurons is elicited by the interaction of the pore-forming Na_v subunit with an auxiliary subunit protein [49]. We propose that the Apteronotid EMNs use a distinct mechanism, wherein substitutions within *scn4ab1* are sufficient to generate I_{NAP}.

The inactivation particle and receptor evolved on different evolutionary timescales (Fig 3E). The substitutions observed within the putative *scn4ab1* inactivation receptor (D4 S4–S5 linker) happened soon after the duplication of *scn4ab*; when substituted into the human $Na_v1.5$ channel, they cause a large I_{NAP} (Fig 4C left). The inactivation particle, however, evolved over a longer timescale, with 2 of the 3 substitutions preceding the duplication. One of the amino acid substitutions (IFM to LFM) occurred before the evolution of Gymnotiformes. The basal Apteronotid then gained another substitution (LFM to LFL). When substituted alone into human

$Na_v1.5$, these substitutions have no discernable impact on inactivation. When coupled with the 5 mutations in the putative receptor, however, they have a moderate mitigation of the persistent current (Fig 4C middle). This suggests that the initial substitutions in the gating particle may have been benign to start but later interacted with and possibly facilitated substitutions in the receptor when *scn4ab* duplicated. One more substitution found near the canonical inactivation particle, DLFL to HLFL, occurred around the same time as the mutations found in the receptor. The addition of the HLFL-to-DLFL substitution further mitigates the persistent current (Fig 4C right). The central F in the IFM particle remains conserved, which is consistent with data indicating that it has a much larger impact on inactivation than its flanking amino acids [35].

The impact of mutations from different evolutionary time periods suggests that *scn4ab1* arose and played a role in the electric organ rather than a spinal cord Na_v channel because its inactivation particle had acquired mutations that dampened the impact of receptor mutations on the electric organ. Despite the extensive changes in the 2 interacting domains, our experiments indicate that the functional interaction between inactivation particle and receptor is maintained in *scn4ab1*, suggesting great evolutionary lability post duplication in these usually highly conserved parts of the channel.

The gain of expression of *scn4ab1* in the spinal cord may have caused or was caused by the changes in the inactivation gate specific to the Apteronotini. Alternatively, the gain in expression might have occurred because of the duplication itself. Genomic and transcriptomic investigations into more basal Apteronotini species will illuminate the coevolution of expression and the inactivation machinery in this gene.

One puzzling observation in our data is that, despite *scn4ab1* displaying an I_{NAP}, this gene still maintains abundant expression in the skeletal muscle, where presumably a persistent current would be disruptive to swimming behavior. Mutations in the *scn4a* gene in humans that results in I_{NAP} in muscles cause muscle diseases [50]. A critical question is this: what prevents an I_{NAP} and spontaneous firing in muscle? There are many posttranscriptional mechanisms for compensating for this expression. I_{NAP} may be inhibited by other muscle-expressing Na_v channel–associated proteins, known as beta subunits [3,4] (which differ in their expression in the Apteronotid muscle and spinal cord [S10 Fig]). Translation of *scn4ab1* may be suppressed by muscle-specific micro-RNAs (miRNAs), or there may exist compensatory increases in hyperpolarizing ionic currents that are specific to the Apteronotini. Finally, it is possible that there is an I_{NAP} in Apteronotini muscle and that the current is part of the unique swimming patterns exhibited by these fish. Most electric fish species, and especially those that maintain high-frequency signaling, remain rigid during swimming and rely on the undulation of a long, derived anal fin. Perhaps a persistent current in these fishes' muscle contributes to their muscle rigidity. These explanations are highly speculative at this time but may inspire further study.

While the duplication of an scn4ab gene and its evolution likely underlies the ability of the Apteronotini to spontaneously fire their electric organs at high frequencies, some other Apteronotid species generate high-frequency electric organ signals often exceeding 1,000 Hz, although whether they fire spontaneously is unknown. It is also possible that another spinal cord–expressing channel (e.g., *scn8ab*) has evolved the ability to support high-frequency firing in other species. Exploration of the sodium channel gene families of other Apteronotids will likely be a rich vein for future research. A complete understanding of EMN firing rates in the Apteronotini must also consider the activities of the other Na_v channels expressed in EMNs, such as *scn8aa* and *scn8ab*, as well as the voltage-gated potassium channels expressed in these neurons. It will be interesting to know whether any of these other channels have also evolved mechanisms for rapid spontaneous firing and/or whether there are other examples of "cryptic duplication"

in these channels. Identifying such highly similar paralogs as *scn4ab1* and *scn4ab2* requires a level of granularity in transcriptomic analysis that has historically been difficult to achieve.

In this study, we have presented evidence that the continued divergence and diversification of electric organ signals is driven in part by the repeated duplication and neofunctionalization of a Na_v channel gene. In 3 separate fish lineages, a duplicate originating from a muscle-type sodium channel was co-opted by a novel electric organ, twice by a muscle-derived organ and once by a neurogenic organ. It is surprising to see *scn4a*-type sodium channels involved in neurogenic electric organs rather than one of the sodium channel types normally expressed in spinal neurons, namely *scn8ab*. The recurring neofunctionalization of *scn4a* genes suggests that this type of sodium channel may be relatively free of selective constraints to evolve novel expression and structural innovations [22]. Phylogenetic placement of the many mutations involved in *scn4ab1* inactivation elicits an interesting hypothesis that perhaps some preadaptive mutations in the inactivation machinery primed this channel for evolution in Apteronotid EMNs.

Perhaps the genetic substrate for new adaptive phenotypes will be found in gene types and families not necessarily expressed in the tissues from which novel phenotypes are derived but more from types of genes whose duplication and neofunctionalization are less detrimental to the organism. Further research may find that the scn4a-type sodium channel displays this remarkable tendency to contribute to electric fish evolution because, compared to other sodium channels, it is the least disruptive to the organism when it duplicates and neofunctionalizes.

Methods

Sample preparation

Animals were acquired through the aquarium trade. Fish were euthanized according to ethical guidelines set by IACUC at UT Austin and Indiana University Bloomington. Skeletal muscle samples were taken from the midtrunk hypaxial location and the spinal cord from the mid to tail location of 3 adult *A*. *albifrons*, *I*. *punctatus*, and *E*. *virescens*, as well as 1 adult *A*. *leptorhynchus* and *P*. *hasemanii*. Tissue samples were flash frozen in liquid nitrogen, and total RNA was extracted and DNA removed, following previously described protocols [23]. Total RNA samples were submitted to the University of Texas at Austin core genomics facility, where ribosomal RNA was removed; paired-end 100 bp RNA-seq (*A*. *albifrons* and *I*. *punctatus*) or paired-end 150 bp RNA-seq (*E*. *virescens*, *A*. *leptrohynchus*, and *P*. *hasemanii*) was performed on an Illumina HiSeq 2000 to produce between 30 and 34 million reads per sample.

De novo transcriptome assembly and sodium channel family annotation

Raw reads were quality filtered and adapter trimmed with Trimmomatic v 0.35. Muscle and spinal cord reads were combined and transcripts assembled from a single biological replicate from each species using Trinity [51] v. 2.0.6. Alpha and beta sodium channel transfrags were extracted and annotated using reciprocal blast. All 8 sodium channel alpha subunits and 5 beta subunits from the *D*. *rerio* were downloaded from GenBank and used to blast sodium channel sequence from each of the transcriptomes with e-value threshold e-6. This yielded 38–71 (alpha) and 17–45 (beta) transfrags per species. CDS's of transfrags were found using Transdecoder v. 3.0.0 with minimum CDS of 100 amino acids. CD-hit-est v. 4.6.4 was used to cluster all coding sequences with >99% sequence similarity and >90% of the sequence overlaps with the longest sequence in each cluster. This resulted in 12–25 transfrags for the alpha subunits and 7–9 for the beta subunits in each species. These sequences were then blasted against the *D*. *rerio* proteome (assembly GRCz10) with blastx (BLAST+ V 2.2.28) where the top hit was designated the transfrag's ortholog.

Scn4ab1 and *scn4ab2* assembly and annotation

The quality of transfrags was visually inspected by mapping reads (RSEM; see below) and using Integrative Genome Browser v 1.3.1 to inspect patterns among mapped reads. The quality was generally good, except for the scn4ab transfrags in the Apteronotid species. Reads from muscle showed high-frequency base mismatches distributed uniformly along the length of the transfrags. Inspection of the reads showed that all high-frequency mismatched bases existed within the same reads. This same pattern occurred in all 3 Apteronotids investigated. However, in all 3 species, these "polymorphisms" did not appear in mapped reads from the spinal cord samples, suggesting the presence of a duplicate of *scn4ab* that did not assemble well and is only expressed in muscle. The facts that more than half of the gene sequence was assembled in each species and that there were very few polymorphisms of reads mapping in the spinal cord together indicated that, while one paralog did not assemble very well, the other did and contained little or no chimeric sequence with its paralog. There were small fragments (<1 kb) that corresponded to the second duplicate in some of the Apteronotid species. To better assemble this gene, we extracted all reads that mapped to *scn4ab* in muscle with at least 2 mismatches and used SOAPdenovo-Trans [52] with kmer size of 55 to assemble large transfrags of this gene. After adding this gene to the transcriptome and remapping with RSEM, the polymorphisms largely disappeared from scn4ab in skeletal muscle in each species. The presence of duplicate scn4ab genes was confirmed with PCR and Sanger sequencing of DNA samples from *A*. *albifrons*.

Expression analysis

All sodium channel transfrags were extracted from the transcriptome and analyzed with RSEM [53] v. 1.2.28 to estimate the relative expression in transcripts per million (TPM) for each transfrag. To estimate gene-level expression, alternative splice forms as estimated by Trinity (_i notation in Trinity V. 1.2.28 notation) were summed for each gene (_g in Trinity notation). When multiple loci (_g) were annotated as orthologs of the same gene, they were assumed to be fragments of the same locus at different, mostly nonoverlapping locations in the gene. Expression levels were alternatively estimated as the averaged TPM or maximum TPM across all these transfrags for each gene in each replicate. These 2 approaches yielded very similar relative expression patterns to the averaging approach just described (S2 Fig). *E*. *electricus* expression was in units of FPKM. To make relative expression comparable between species, TPM or FPKM values were scaled by total sodium channel expression in each tissue.

qPCR analysis

Three muscle and spinal cord total RNA samples from *A*. *albifrons* were reverse transcribed to cDNA with random hexamers and oligo-dT20 primers using the superscript III kit and standard protocol (Life Technologies, Grand Island, NY). Primers and hydrolysis probes (IDT, Coralville, IA) specific to *scn4ab1*, *scn4aa*, and the housekeeping gene *RPL13a* (which was sequenced using degenerate primers, PCR, and Sanger sequencing) were used with the TaqMan Universal Master Mix NO UNG (Applied Biosystems, Branchburg, NJ) to perform qPCR reactions in the 2 tissues of 3 biological replicates. Specificity of primers and probes was confirmed through PCR and Sanger sequencing. qPCR amplicons were designed to span multiple exons, and negative controls were performed where the reverse transcriptase was not added to the reaction mix. qPCR reactions were run on a Viia7 Real-Time PCR machine (Applied Biosystems).

RPL13a normalized expression levels were estimated for *scn4aa* and *scn4ab1* in muscle and spinal cord using previously published protocol [23]. In brief, the raw amplification data were baseline corrected, and linear regression on the log-linear phase of amplification for each individual reaction was used to verify close to 100% doubling efficiency and select a common threshold for each sodium channel gene and the housekeeping gene. Cq and doubling efficiency was estimated and confirmed using the LinRegPCR software package [54–56]. Normalized expression was estimated using $2^{-\Delta\Delta Cq}$ method [57], which assumes 100% doubling efficiency.

Phylogenetic analysis

Scn4ab sequences were trimmed to their longest coding sequence with stop codons removed for codon alignment. A codon alignment was created on the GUIDANCE2 server. A GUIDANCE alignment quality threshold of 0.93 was selected, which removed 20% of the codon positions. We used MRBayes version 3.2.5 [58] to estimate the gene tree topology. We estimated the gene tree by model averaging over the space of all possible GTR models with gamma-distributed rates and data partitioned by the 3 codon positions (lset nst = mixed rates = gamma; mcmc ngen = 1,000,000). This analysis produced a consensus tree that was identical in topology and nearly identical in branch lengths to a tree generated without partitioning the codon positions. The consensus tree was very well resolved. All branches had posterior probability >0.99 except one, which had posterior probability of 0.75. Two independent runs, each with 4 chains, converged for 2 million iterations after 1 million burn-in on identical marginal posterior distributions for all model parameters. Posterior distributions were found to be of good quality when visualized with Tracer. Maximum likelihood estimates of branch dN/dS were estimated for the majority rule consensus tree generated from unpartitioned codon alignment using the BSR method [32] on the Datamonkey server. Sequences used in this analysis are given in S1 Table. Maximum likelihood dN/dS was also estimated using PAML [33] version 4.9e with F3X4 for codon frequencies. The following two different models were evaluated: (1) all branches share the same dN/dS ratio and (2) the root branch of the *scn4ab1* clade having a unique ratio. Models were compared by the likelihood ratio test.

Amino acid substitution density analysis

To locate regions where the duplicate sodium channels may have diverged in function, the density of amino acid substitutions along the peptide sequence was analyzed. The amino acid sequences of *scn4ab1* and *scn4ab2* were most complete in *A. albifrons*. These sequences were aligned against *scn4ab* of *D. rerio*. Substitutions or indels were mapped to each of the 3 branches of the resultant tree by a minimum mutation parsimony criterion. At positions where all 3 sequences differed, a substitution was assumed to have occurred on each of the 3 branches. This slightly inflates the density of substitutions at these regions for the whole tree but does not impact investigation of the relative density of amino acid substitutions. For each of the branches connecting the 3 genes to the internal node, the density of substitutions at each residue was calculated using the R function *density* (package *stats*) with binwidth set to 15 and then weighted by that gene's share of the total number of substitutions inferred by parsimony. The relative density at position x on sequence y was calculated as the density of position x on sequence y divided by the sum of the density of the other 2 sequences at the same position. For example, if *scn4ab1* has a relative density of 2 at a position in the sequence, then the density of amino acid substitutions is twice that of the other 2 sequences combined. We generated 10,000 simulations in which the number of amino acid substitutions mapped to each branch were randomly distributed along each of the 3 sequences, and the same relative density statistic was measured to

get a NULL distribution for relative density and to determine if the clustering pattern of amino acid substitutions observed is unlikely to have arisen by chance.

Molecular clock analysis

The dataset included sequences from 37 taxa for 3 genes: cytochrome c oxidase subunit I (COI), cytochrome B (CytB), and recombination activating gene 2 (RAG2). Most sequences were derived from concatenated consensus data (see S2 Table of the manuscript for a full list of individuals, source IDs, and accession numbers), with supplemental Apteronotid sequences coming from [31] and outgroup sequences taken from NCBI (accession numbers for supplemental Apteronotid and outgroup sequences listed in S2–S3 Tables). Sequences were concatenated and aligned using the L-INS-I protocol in MAFFT [59]. Maximum likelihood trees were estimated using a GTR + gamma substitution model in RAxML v. 7.4.2 [60] with *D. rerio* as the outgroup. Dating estimates were performed using the RelTime method [61] and a Tamura-Nei model [62] in the MEGA7 software package [63]. The model included 5 discrete gamma estimates, allowed for invariant sites, included all codon positions, and discarded alignment gaps. The tree was calibrated using 2 dating ranges estimated from [64]: (i) the split between *I. punctatus* and *Serrasalmus maculatus* was assigned a potential range of 115–150 MYA, and (ii) the split between *S. maculatus* and *Orthosternarus tamandua* was assigned a potential range of 75–115 MYA.

Electrophysiology

Two electrode voltage clamp recordings in *Xenopus* oocytes were performed at 20–22˚C using a Turbotec 03X amplifier (NPI). Intracellular recording electrodes had resistances of 0.2–0.4 MΩ when backfilled with 3 M KCl. Voltage protocols are described in detail in the figure legends. For all experiments, the holding potential was −120 mV. Mutations were made into human cardiac sodium channel $Na_v1.5$ in pcDNA 3.1 using mini-gene synthesis (Biobasic, Markham, Ontario, Canada) in conjunction with Gibson Assembly. All constructs were verified by sequencing through the entire open reading frame. RNA was transcribed using the Mmessage Mmachine T7 Ultra Kit (Thermofisher, Waltham, MA) after linearization with Not1. Approximately 12.5–50 ng of cRNA from the alpha subunit was co-injected with 6.25–25 ng of cRNA from the β1 auxillary subunit, which is expressed ubiquitously in excitable cells and enhances sodium channel surface expression [65,66]. Recordings were conducted 24 to 48 hours later in oocyte Ringer's solution containing (in mM): 116 NaCl, 2 KCl, 1.8 $CaCl_2$, 2 $MgCl_2$, 5 mM HEPES, pH 7.4. Traces were acquired at 50 Khz and filtered at 10 Khz for display in the figures. For calculation of persistent current, pClamp 9.2 was used to average the last 0.5 ms of each 30-ms pulse, and then this value was divided by the peak of the transient component, after linear leak subtraction. All statistical comparisons were by unpaired Student *t* test with two-tailed distribution.

Supporting information

S1 Fig. The axons of the brainstem pacemaker neurons run down the spinal cord (grey line) and synapse on the EMNs (teal). Top: in most adult Gymnotiforms, the axons of the EMNs synapse on the muscle-derived cells of the electric organ (salmon). In Apteronotids, the muscle-derived electric organ degenerates (former position faded), and the axons of the EMNs extend and form a new, neurogenic electric organ. EMN, electromotorneuron.
(TIF)

S2 Fig. The relative expression of sodium channels in muscle and spinal cord. Plots show the relative expression of all transfrags of each paralog in each species. For most paralogs, each transfrag has similar expression. The unduplicated *scn4ab* ortholog in non-Apteronotids is highlighted in purple, while the duplicate *scn4ab* parlaogs in Apteronotids are highlighted in red and blue. Figure data included in S1 Data.
(TIF)

S3 Fig. Relative expression of *scn4aa* and *scn4ab1* in muscle and spinal cord derived from qRT-PCR in samples from *A. albifrons*. qRT-PCR, quantitative reverse transcription PCR. Expression was normalized to the housekeeping gene RPL13a.
(TIF)

S4 Fig. Sequence alignment of *A. albifrons* muscle-specific *scn4ab2* (top) and EMN/muscle–expressing *scn4ab1* (bottom). Sequences derived from molecular cloning and Sanger sequencing. Dashed vertical lines indicate exon–intron boundaries. Exon numbering is according to alignment with *D. rerio scn4ab*. EMN, electromotorneuron.
(TIF)

S5 Fig. Gene tree for *scn4ab* of Gymnotiforms and other teleosts rooted with scn4aa. Note that scn4ab is in single copy in 2 basal Apteronotids ("*A*" *bonapartii* and *A. devenanzi*) but duplicated into *scn4ab1* and *scn4ab2* before the divergence of the 3 members of the Apteronotini. Key events in the evolution of myogenic and neurogenic electric organs are noted. Posterior probabilities given for duplication of *scn4ab* in the Apteronotini. Gene accessions included in S1 Table.
(TIF)

S6 Fig. The domain 3–4 intracellular loop and the domain 4 S4–S5 linker from Na_v channels of *A. albifrons*, *D. rerio*, and *Homo sapiens* to illustrate the strong conservation of these motifs across channels and species. Some amino acid substitutions occur in all Apteronotid *scn4ab* channels (blue dot), some in all Apteronotini (red dots), and 1 only in *A. albifrons* (green dot). Na_v, voltage-gated sodium.
(TIF)

S7 Fig. Normalized current–voltage relationship for the transient and persistent current for the cell recorded in Fig 4B. Note that the persistent current activates at a more negative voltage than the peak current and that it comprises a relatively greater fraction of the peak current in the range from −60 to −40 mV. Figure data included in S2 Data.
(TIF)

S8 Fig. Effects of Apteronotinti Na_v1.4ab2 substitutions on equilibrium gating of human cardiac Na_v1.5. A. Inactivation particle substitutions (see Figs 3 and S6) did not affect channel activation. (B) Apteronitini-specific domain 4 S4–S5 substitutions were associated with a 7–10 mV depolarizing shift in channel activation either alone or in conjunction with substitutions in the inactivation particle. (C) Effects of Apteronotini (DIFM to HLFL) inactivation particle substitutions on steady-state inactivation. *Xenopus* oocytes expressing sodium channel variants were subjected to a 500-millisecond conditioning pulse at a given voltage, followed by a 1-millisecond step at −100 mV and a 20-millisecond test pulse at −20 mV. (D) Same as panel C but showing effects of Apteronotini domain 4 S4–S5 substitutions on steady-state inactivation alone and in conjunction with substitutions in the inactivation particle. Note the nonzero asymptotes for steady-state inactivation in the D4 S4–S5 variants. $N \geq 5$ for each variant; quantification in S4 Table. Figure data included in S2 Data. D, aspartate; F, phenylalanine; H,

histidine; I, isoleucine; M, methionine; Na_v, voltage-gated sodium.
(TIF)

S9 Fig. Apteronotinti $Na_v1.4ab$ substitutions in domain 4 S4–S5 sped recovery from fast inactivation of $hNa_v1.5$, which is partially mitigated by substitutions in the inactivation particle. Sodium channel currents were activated via 20-millisecond pulses to −20 mV, which were separated by a recovery interval at −120 mV for a specified length of time. (A) Normalized fraction of current as a function of recovery interval for WT $hNa_v1.5$ and Apteronotinti inactivation particle substitutions. (B) Same as panel A but showing effects of Apteronotinti $Na_v1.4ab$ domain 4 S4–S5 substitutions on recovery from fast inactivation alone and in conjunction with knifefish substitutions in the inactivation particle. (C) Quantification of fraction of current recovered at the shortest recovery interval (0.5 milliseconds). Asterisk indicates variants with statistically significant differences ($p < 0.01$) as compared to WT $hNa_v1.5$. (D) Same as panel C but showing fraction recovered by 1.0 milliseconds. $N \geq 4$ for each variant. In panel C and D, note that the domain 4 S4–S5 substitutions (red bar) increased the fraction of current recovered as compared to $hNa_v1.5$ (indicating faster recovery from inactivation). Inclusion of the inactivation particle substitutions (green and purple bars) reduced this effect. Figure data included in S2 Data. hNa_v, human voltage-gated sodium; WT, wild-type.
(TIF)

S10 Fig. Beta subunits of muscle and spinal cord. Beta subunits are known to modify the properties of Na_v channels. Note that different subunits are expressed in the muscle and spinal cord of Apteronotids. Figure data included in S1 Data. Na_v, voltage-gated sodium.
(TIF)

S1 Table. Gene accessions for gene tree and expression analysis.
(DOCX)

S2 Table. Accession for genes used for chronogram.
(DOCX)

S3 Table. Supplemental sequences for chronogram.
(DOCX)

S4 Table. Activation and steady-state inactivation statistics.
(DOCX)

S1 Data. Data from expression and phylogeny figures.
(XLSX)

S2 Data. Data from electrophysiology figures.
(XLSX)

Acknowledgments

The authors thank Derek Vo and Ying Lu (Universty of Texas, Austin) for technical assistance and Laura Crothers (University of Texas, Austin) for comments on the manuscript. We also acknowledge the Texas Advanced Computing Center (TACC) at the University of Texas at Austin for providing HPC resources that have contributed to the research results reported in this paper. We also acknowledge Grace Galles (University of Iowa) for excellent technical assistance.

Author Contributions

Conceptualization: Ammon Thompson, Christopher A. Ahern.

Data curation: Ammon Thompson, Daniel T. Infield, Adam R. Smith.

Formal analysis: Ammon Thompson, Daniel T. Infield, Adam R. Smith, G. Troy Smith.

Funding acquisition: Ammon Thompson, Daniel T. Infield, Christopher A. Ahern, Harold H. Zakon.

Investigation: Ammon Thompson, Daniel T. Infield, Adam R. Smith, Christopher A. Ahern, Harold H. Zakon.

Methodology: Ammon Thompson, Daniel T. Infield, Adam R. Smith, G. Troy Smith, Christopher A. Ahern, Harold H. Zakon.

Project administration: Christopher A. Ahern, Harold H. Zakon.

Resources: Ammon Thompson, Daniel T. Infield, Christopher A. Ahern, Harold H. Zakon.

Software: Ammon Thompson.

Supervision: G. Troy Smith, Christopher A. Ahern, Harold H. Zakon.

Validation: Ammon Thompson, Daniel T. Infield, Adam R. Smith.

Visualization: Ammon Thompson.

Writing – original draft: Ammon Thompson, Daniel T. Infield, Christopher A. Ahern, Harold H. Zakon.

Writing – review & editing: Ammon Thompson, Daniel T. Infield, Adam R. Smith, G. Troy Smith, Christopher A. Ahern, Harold H. Zakon.

References

1. Lavoué S, Miya M, Arnegard ME, Sullivan JP, Hopkins CD, Nishida M. Comparable ages for the independent origins of electrogenesis in African and South American weakly electric fishes. PLoS ONE. 2012; 7(5):e36287. https://doi.org/10.1371/journal.pone.0036287 PMID: 22606250
2. Novak AE, Taylor AD, Pineda RH, Lasda EL, Wright MA, Ribera AB. Embryonic and larval expression of zebrafish voltage-gated sodium channel alpha-subunit genes. Dev Dyn Off Publ Am Assoc Anat. 2006 Jul; 235(7):1962–73.
3. Arnegard ME, Zwickl DJ, Lu Y, Zakon HH. Old gene duplication facilitates origin and diversification of an innovative communication system—twice. Proc Natl Acad Sci U S A. 2010 Dec 21; 107 (51):22172–7. https://doi.org/10.1073/pnas.1011803107 PMID: 21127261
4. Zakon HH, Lu Y, Zwickl DJ, Hillis DM. Sodium channel genes and the evolution of diversity in communication signals of electric fishes: convergent molecular evolution. Proc Natl Acad Sci U S A. 2006 Mar 7; 103(10):3675–80. https://doi.org/10.1073/pnas.0600160103 PMID: 16505358
5. Hopkins CD. Design features for electric communication. J Exp Biol. 1999 May; 202(Pt 10):1217–28. PMID: 10210663
6. Hopkins CD, Heiligenberg WF. Evolutionary designs for electric signals and electroreceptors in gymnotoid fishes of Surinam. Behav Ecol Sociobiol. 1978 Jun 1; 3(2):113–34.
7. Hopkins CD. On the Diversity of Electric Signals in a Community of Mormyrid Electric Fish in West Africa. Integr Comp Biol. 1981 Feb 1; 21(1):211–22.
8. Stoddard PK. Predation enhances complexity in the evolution of electric fish signals. Nature. 1999 Jul; 400(6741):254. https://doi.org/10.1038/22301 PMID: 10421365
9. Lewis JE, Gilmour KM, Moorhead MJ, Perry SF, Markham MR. Action Potential Energetics at the Organismal Level Reveal a Trade-Off in Efficiency at High Firing Rates. J Neurosci. 2014 Jan 1; 34 (1):197–201. https://doi.org/10.1523/JNEUROSCI.3180-13.2014 PMID: 24381281

10. Arnegard ME, McIntyre PB, Harmon LJ, Zelditch ML, Crampton WGR, Davis JK, et al. Sexual Signal Evolution Outpaces Ecological Divergence during Electric Fish Species Radiation. Am Nat. 2010 Sep 1; 176(3):335–56. https://doi.org/10.1086/655221 PMID: 20653442
11. Moortgat KT, Keller CH, Bullock TH, Sejnowski TJ. Submicrosecond pacemaker precision is behaviorally modulated: The gymnotiform electromotor pathway. Proc Natl Acad Sci. 1998 Apr 14; 95 (8):4684–9. PMID: 9539799
12. Salazar VL, Krahe R, Lewis JE. The energetics of electric organ discharge generation in gymnotiform weakly electric fish. J Exp Biol. 2013 Jul 1; 216(13):2459–68.
13. Shifman AR, Longtin A, Lewis JE. Ultrafast traveling wave dominates the electric organ discharge of Apteronotus leptorhynchus: an inverse modelling study. Sci Rep [Internet]. 2015 Oct 30 [cited 2018 Jan 12]; 5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4626797/
14. Szabó T, Czéh G. Sensory Physiology of Aquatic Lower Vertebrates: Satellite Symposium of the 28th International Congress of Physiological Sciences, Keszthely, Hungary, 1980. Elsevier; 2013. 297 p.
15. Turner CR, Derylo M, Santana CD de, Alves-Gomes JA, Smith GT. Phylogenetic comparative analysis of electric communication signals in ghost knifefishes (Gymnotiformes: Apteronotidae). J Exp Biol. 2007 Dec 1; 210(23):4104–22.
16. Schaefer J e, Zakon HH. Opposing actions of androgen and estrogen on in vitro firing frequency of neuronal oscillators in the electromotor system. J Neurosci Off J Soc Neurosci. 1996 Apr 15; 16(8):2860–8.
17. Kirschbaum F. Myogenic electric organ precedes the neurogenic organ in apteronotid fish. Naturwissenschaften. 1983 Apr 1; 70(4):205–7. PMID: 6855924
18. Waxman SG, Pappas GD, Bennett MVL. MORPHOLOGICAL CORRELATES OF FUNCTIONAL DIFFERENTIATION OF NODES OF RANVIER ALONG SINGLE FIBERS IN THE NEUROGENIC ELECTRIC ORGAN OF THE KNIFE FISH STERNARCHUS. J Cell Biol. 1972 Apr 1; 53(1):210–24. PMID: 5013596
19. Smith GT. Pharmacological characterization of ionic currents that regulate high-frequency spontaneous activity of electromotor neurons in the weakly electric fish, Apteronotus leptorhynchus. J Neurobiol. 2006 Jan; 66(1):1–18. https://doi.org/10.1002/neu.20202 PMID: 16187302
20. Taddese A, Bean BP. Subthreshold Sodium Current from Rapidly Inactivating Sodium Channels Drives Spontaneous Firing of Tuberomammillary Neurons. Neuron. 2002 Feb 14; 33(4):587–600. PMID: 11856532
21. Chen Y-H, Huang F-L, Cheng Y-C, Wu C-J, Yang C-N, Tsay H-J. Knockdown of zebrafish Nav1.6 sodium channel impairs embryonic locomotor activities. J Biomed Sci. 2007 Aug 9; 15(1):69–78. https://doi.org/10.1007/s11373-007-9200-4 PMID: 17687633
22. Thompson A, Zakon HH, Kirkpatrick M. Compensatory Drift and the Evolutionary Dynamics of Dosage-Sensitive Duplicate Genes. Genetics. 2016 Feb; 202(2):765–74. https://doi.org/10.1534/genetics.115.178137 PMID: 26661114
23. Thompson A, Vo D, Comfort C, Zakon HH. Expression evolution facilitated the convergent neofunctionalization of a sodium channel gene. Mol Biol Evol. 2014 Aug; 31(8):1941–55. https://doi.org/10.1093/molbev/msu145 PMID: 24782440
24. Aman TK, Grieco-Calub TM, Chen C, Rusconi R, Slat EA, Isom LL, et al. Regulation of Persistent Na Current by Interactions between β Subunits of Voltage-Gated Na Channels. J Neurosci. 2009 Feb 18; 29(7):2027–42. https://doi.org/10.1523/JNEUROSCI.4531-08.2009 PMID: 19228957
25. Ma JY, Catterall WA, Scheuer T. Persistent Sodium Currents through Brain Sodium Channels Induced by G Protein βγ Subunits. Neuron. 1997 Aug 1; 19(2):443–52. PMID: 9292732
26. Qu Y, Curtis R, Lawson D, Gilbride K, Ge P, DiStefano PS, et al. Differential Modulation of Sodium Channel Gating and Persistent Sodium Currents by the β1, β2, and β3 Subunits. Mol Cell Neurosci. 2001 Nov 1; 18(5):570–80. https://doi.org/10.1006/mcne.2001.1039 PMID: 11922146
27. Gallant JR, Traeger LL, Volkening JD, Moffett H, Chen P-H, Novina CD, et al. Genomic basis for the convergent evolution of electric organs. Science. 2014 Jun 27; 344(6191):1522–5. https://doi.org/10.1126/science.1254432 PMID: 24970089
28. Pineda RH, Svoboda KR, Wright MA, Taylor AD, Novak AE, Gamse JT, et al. Knockdown of Nav 1.6a Na+ channels affects zebrafish motoneuron development. Development. 2006 Oct 1; 133 (19):3827–36. https://doi.org/10.1242/dev.02559 PMID: 16943272
29. Garcıa KD, Sprunger LK, Meisler MH, Beam KG. The Sodium Channel Scn8a Is the Major Contributor to the Postnatal Developmental Increase of Sodium Current Density in Spinal Motoneurons. J Neurosci. 1998 Jul 15; 18(14):5234–9. PMID: 9651206
30. Schaller KL, Caldwell JH. Developmental and regional expression of sodium channel isoform NaCh6 in the rat central nervous system. J Comp Neurol. 2000 Apr 24; 420(1):84–97. PMID: 10745221

31. Tagliacollo VA, Bernt MJ, Craig JM, Oliveira C, Albert JS. Model-based total evidence phylogeny of Neotropical electric knifefishes (Teleostei, Gymnotiformes). Mol Phylogenet Evol. 2016 Feb 1; 95:20–33. https://doi.org/10.1016/j.ympev.2015.11.007 PMID: 26616344

32. Pond K, L S, Murrell B, Fourment M, Frost SDW, Delport W, et al. A Random Effects Branch-Site Model for Detecting Episodic Diversifying Selection. Mol Biol Evol. 2011 Nov 1; 28(11):3033–43. https://doi.org/10.1093/molbev/msr125 PMID: 21670087

33. Yang Z. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol Biol Evol. 1998 May 1; 15(5):568–73. https://doi.org/10.1093/oxfordjournals.molbev.a025957 PMID: 9580986

34. Ahern CA, Payandeh J, Bosmans F, Chanda B. The hitchhiker's guide to the voltage-gated sodium channel galaxy. J Gen Physiol. 2016 Jan; 147(1):1–24. https://doi.org/10.1085/jgp.201511492 PMID: 26712848

35. West JW, Patton DE, Scheuer T, Wang Y, Goldin AL, Catterall WA. A cluster of hydrophobic amino acid residues required for fast Na(+)-channel inactivation. Proc Natl Acad Sci U S A. 1992 Nov 15; 89 (22):10910–4. PMID: 1332060

36. Kellenberger S, Scheuer T, Catterall WA. Movement of the Na+ channel inactivation gate during inactivation. J Biol Chem. 1996 Nov 29; 271(48):30971–9. PMID: 8940085

37. Smith MR, Goldin AL. Interaction between the sodium channel inactivation linker and domain III S4-S5. Biophys J. 1997 Oct; 73(4):1885–95. https://doi.org/10.1016/S0006-3495(97)78219-5 PMID: 9336184

38. McPhee JC, Ragsdale DS, Scheuer T, Catterall WA. A critical role for the S4-S5 intracellular loop in domain IV of the sodium channel alpha-subunit in fast inactivation. J Biol Chem. 1998 Jan 9; 273 (2):1121–9. PMID: 9422778

39. Fertleman CR, Baker MD, Parker KA, Moffatt S, Elmslie FV, Abrahamsen B, et al. SCN9A Mutations in Paroxysmal Extreme Pain Disorder: Allelic Variants Underlie Distinct Channel Defects and Phenotypes. Neuron. 2006 Dec 7; 52(5):767–74. https://doi.org/10.1016/j.neuron.2006.10.006 PMID: 17145499

40. Freilich ER, Jones JM, Gaillard WD, Conry JA, Tsuchida TN, Reyes C, et al. Novel SCN1A Mutation in a Proband With Malignant Migrating Partial Seizures of Infancy. Arch Neurol. 2011 May 9; 68 (5):665–71. https://doi.org/10.1001/archneurol.2011.98 PMID: 21555645

41. Groome J, Larsen MF, Coonts A. Differential effects of paramyotonia congenita mutations F1473S and F1705I on sodium channel gating. Channels. 2008 Jan 1; 2(1):39–50. PMID: 18690054

42. Yan Z, Zhou Q, Wang L, Wu J, Zhao Y, Huang G, et al. Structure of the Nav1.4-β1 Complex from Electric Eel. Cell. 2017 Jul 27; 170(3):470–482.e11. https://doi.org/10.1016/j.cell.2017.06.039 PMID: 28735751

43. Putnam NH, Butts T, Ferrier DEK, Furlong RF, Hellsten U, Kawashima T, et al. The amphioxus genome and the evolution of the chordate karyotype. Nature. 2008 Jun; 453(7198):1064. https://doi.org/10.1038/nature06967 PMID: 18563158

44. Liebeskind BJ, Hillis DM, Zakon HH. Convergence of ion channel genome content in early animal evolution. Proc Natl Acad Sci. 2015 Feb 24; 112(8):E846–51. https://doi.org/10.1073/pnas.1501195112 PMID: 25675537

45. Crill WE. Persistent Sodium Current in Mammalian Central Neurons. Annu Rev Physiol. 1996; 58 (1):349–62.

46. French CR, Sah P, Buckett KJ, Gage PW. A voltage-dependent persistent sodium current in mammalian hippocampal neurons. J Gen Physiol. 1990 Jun 1; 95(6):1139–57. PMID: 2374000

47. Khaliq ZM, Bean BP. Pacemaking in Dopaminergic Ventral Tegmental Area Neurons: Depolarizing Drive from Background and Voltage-Dependent Sodium Conductances. J Neurosci. 2010 May 26; 30 (21):7401–13. https://doi.org/10.1523/JNEUROSCI.0143-10.2010 PMID: 20505107

48. Zhang C, Bosch MA, Qiu J, Rønnekleiv OK, Kelly MJ. 17β-Estradiol Increases Persistent Na+ Current and Excitability of AVPV/PeN Kiss1 Neurons in Female Mice. Mol Endocrinol. 2015 Apr 1; 29 (4):518–27. https://doi.org/10.1210/me.2014-1392 PMID: 25734516

49. Lewis AH, Raman IM. Resurgent current of voltage-gated Na(+) channels. J Physiol. 2014 Nov 15; 592 (22):4825–38. https://doi.org/10.1113/jphysiol.2014.277582 PMID: 25172941

50. Zhao J, Duprè N, Puymirat J, Chahine M. Biophysical characterization of M1476I, a sodium channel founder mutation associated with cold-induced myotonia in French Canadians. J Physiol. 2012 Jun 1; 590(11):2629–44. https://doi.org/10.1113/jphysiol.2011.223461 PMID: 22250216

51. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 2013 Aug; 8(8):1494–512. https://doi.org/10.1038/nprot.2013.084 PMID: 23845962

52. Xie Y, Wu G, Tang J, Luo R, Patterson J, Liu S, et al. SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. Bioinforma Oxf Engl. 2014 Jun 15; 30(12):1660–6.

53. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011 Aug 4; 12:323. https://doi.org/10.1186/1471-2105-12-323 PMID: 21816040

54. Ramakers C, Ruijter JM, Deprez RHL, Moorman AFM. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci Lett. 2003 Mar 13; 339(1):62–6. PMID: 12618301

55. Ruijter JM, Ramakers C, Hoogaars WMH, Karlen Y, Bakker O, Hoff VD, et al. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Res. 2009 Apr 1; 37(6): e45–e45. https://doi.org/10.1093/nar/gkp045 PMID: 19237396

56. Tuomi JM, Voorbraak F, Jones DL, Ruijter JM. Bias in the Cq value observed with hydrolysis probe based quantitative PCR can be corrected with the estimated PCR efficiency value. Methods. 2010 Apr 1; 50(4):313–22. https://doi.org/10.1016/j.ymeth.2010.02.003 PMID: 20138998

57. Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. Methods. 2001 Dec 1; 25(4):402–8. https://doi.org/10.1006/meth.2001.1262 PMID: 11846609

58. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. Syst Biol. 2012 May 1; 61(3):539–42. https://doi.org/10.1093/sysbio/sys029 PMID: 22357727

59. Katoh K, Kuma K, Toh H, Miyata T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 2005 Jan 1; 33(2):511–8. https://doi.org/10.1093/nar/gki198 PMID: 15661851

60. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 2006 Nov 1; 22(21):2688–90. https://doi.org/10.1093/bioinformatics/btl446 PMID: 16928733

61. Tamura K, Battistuzzi FU, Billing-Ross P, Murillo O, Filipski A, Kumar S. Estimating divergence times in large molecular phylogenies. Proc Natl Acad Sci. 2012 Nov 20; 109(47):19333–8. https://doi.org/10.1073/pnas.1213199109 PMID: 23129628

62. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol. 1993 May 1; 10(3):512–26. https://doi.org/10.1093/oxfordjournals.molbev.a040023 PMID: 8336541

63. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016 Jul 1; 33(7):1870–4. https://doi.org/10.1093/molbev/msw054 PMID: 27004904

64. Near TJ, Eytan RI, Dornburg A, Kuhn KL, Moore JA, Davis MP, et al. Resolution of ray-finned fish phylogeny and timing of diversification. Proc Natl Acad Sci. 2012 Aug 21; 109(34):13698–703. https://doi.org/10.1073/pnas.1206625109 PMID: 22869754

65. Isom LL, Jongh KD, Patton DE, Reber BF, Offord J, Charbonneau H, et al. Primary structure and functional expression of the beta 1 subunit of the rat brain sodium channel. Science. 1992 May 8; 256 (5058):839–42. PMID: 1375395

66. Makita N, Bennett PB, George AL. Voltage-gated Na+ channel beta 1 subunit mRNA expressed in adult human skeletal muscle, heart, and brain is encoded by a single gene. J Biol Chem. 1994 Mar 11; 269 (10):7571–8. PMID: 8125980