Expansion of voltage-dependent Na^+ channel gene family in early tetrapods coincided with the emergence of terrestriality and increased brain complexity.

Research Article

Harold H. Zakon[1,2], Manda C. Jost[3] & Ying Lu[1]

[1]Section of Neurobiology, The University of Texas, Austin, TX 78712

[2]Bay Paul Center for Comparative and Molecular Biology, The Marine Biological Laboratory, Woods Hole, MA 02543.

[3]Department of Natural Sciences, Western New Mexico University, Silver City, NM 88061

Corresponding author: Harold Zakon, address as above; h.zakon@mail.utexas.edu, (512)-619-5523 (cell); (512)-471-9651 (fax)

Key words: Sodium channel, tetrapods, amniotes, terrestriality, gene duplication, brain

Running head: Evolution of Na^+ channel genes in tetrapods

Abstract

Mammals have 10 voltage-dependent sodium (Nav) channel genes. Nav channels are expressed in different cell types with different sub-cellular distributions and are critical for many aspects of neuronal processing. The last common ancestor of teleosts and tetrapods had four Nav channel genes presumably on four different chromosomes. In the lineage leading to mammals a series of tandem duplications on two of these chromosomes more than doubled the number of Nav channel genes. It is unknown when these duplications occurred, whether they occurred against a backdrop of duplication of flanking genes on their chromosomes, or as an expansion of ion channel genes in general. We estimated key dates of the Nav channel gene family expansion by phylogenetic analysis using teleost, elasmobranch, lungfish, amphibian, avian, lizard, and mammalian Nav channel sequences, as well as chromosomal synteny for tetrapod genes. We tested, and exclude, the null hypothesis that Nav channel genes reside in regions of chromosomes prone to duplication by demonstrating the lack of duplication or duplicate retention of surrounding genes. We also find no comparable expansion in other voltage dependent ion channel gene families of tetrapods following the teleost-tetrapod divergence. We posit a specific expansion of the Nav channel gene family in the Devonian and Carboniferous periods when tetrapods evolved, diversified, and invaded the terrestrial habitat. During this time the amniote forebrain evolved greater anatomical complexity and novel tactile sensory receptors appeared. The duplication of Nav channel genes allowed for greater regional specialization in

Nav channel expression, variation in sub-cellular localization, and enhanced processing of somatosensory input.

Introduction

Voltage-dependent sodium (Nav) channels are critical for electrical excitability and neuronal computation. Mammals have 10 Nav channels with distinct biophysical properties, types of modulation by neurotransmitters, and tissue and sub-cellular distributions (Angelino & Brenner 2007). For example, a distinct Nav channel (Nav1.4) is predominantly expressed in skeletal muscle, another (Nav1.5) predominantly in cardiac muscle. Different Nav channels are expressed in unmyelinated axons (Nav1.2) and at the nodes of Ranvier in myelinated axons (Nav1.6) (Westenbroek et al 1989; Caldwell et al 2000;). Specific Nav channels (Nav1.7, 1.8, 1.9) are highly expressed in nociceptors (Akopian et al 1996; Cummins et al 1999; Dib-Hajj et al 2002) or may be up-regulated specifically in neurons in the nociceptive pathway following injury (Nav1.3) (Hains et al 2003). Some cells types, such as fast-firing parvalbumin-positive inhibitory neurons, mainly express one type of Nav channel (Nav1.1), whereas another Nav channel is expressed in neighboring pyramidal neurons (Nav1.6) (Ogiwara et al 2007; Lorincz & Nusser 2010). Different Nav channels may even be expressed in different sub-cellular domains in neurons: distinct Nav channels are responsible for initiating the action potential at the axon initial segment (Nav1.6) and for back-propagation of the action potential into the soma

(Nav1.2), a critical function for activity-dependent synaptic plasticity (Hu et al 2009).

Recent studies have clarified the evolutionary relationships among and timing of the origin of vertebrate Nav channels (Okamura et al 1994; Plummer & Meisler 1999; Lopreato et al 2001; Goldin 2002; Piontkivska & Hughes 2003; Novak et al 2006; Hill et al 2008). Early in vertebrate evolution a single Nav channel gene of early chordates (Okamura et al 1994) duplicated twice, presumably during two consecutive whole genome duplication (WGD) events, giving rise to four Nav channel genes, each presumed on a different chromosome (Plummer & Meisler 1999; Lopreato et al 2001; Novak et al 2006). In teleosts, this number jumped to eight Nav channel genes *via* a third teleost-specific WGD (Lopreato et al 2001; Novak et al 2006), whereas a series of tandem duplications on two of these chromosomes at unknown times in the lineage leading to mammals, resulted in a total of 10 Nav channel genes in rodents and humans and presumably other mammals (Plummer & Meisler 1999). A major goal of this study was to determine the timing and significance of these tandem duplications for tetrapod evolution. Additionally, we wished to investigate whether the duplication and retention was unique to Nav channel genes and, therefore, possibly adaptive, or merely the result of passive factors such as chromosomal “hotspots” for duplication. Finally, we asked whether the expansion of the Nav channel gene family was part of a general expansion of other ion channel gene families or a unique event.

Materials and Methods

Genomic Sequences

We obtained the whole complement of Nav amino acid sequences from human (*Homo sapiens*) and rat (*Rattus norvegicus*) from GenBank. Using a BLAT search with human and rodent Nav channel genes, we derived and translated nucleotide sequences from the Ensemble genome databases for western clawed frog (*Xenopus tropicalis,* v4.1, August 2005*)*, green anole lizard (*Anolis carolinensis,* AnoCar1, assembly 2007*)*, platypus (*Ornithorhynchus anatinus, v*5.0, assembly January 2007), gray short-tailed opossum (*Monodelphis domestica*; MonDom5, October 2006), chicken (*Gallus gallus,* v2.1, May 2006), and elephant shark (*Callorhinchus milii*; v1.0, 2007). About half of the Nav channel genes from these species had already been deposited in GenBank, but the other half had not yet been annotated. Additional sequences from lamprey (*Petromyzon marinus*), newt (*Cynops pyrrhogaster*), and the African clawed frog (*Xenopus laevis*) were also used for phylogeny estimation (Table. S1).

Because *Xenopus* is extensively used as a developmental biological model, very good *Xenopus* EST databases are available from the TIGR database (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/Blast/index.cgi). We utilized this EST database to confirm *Xenopus* genomic sequences (Fig. S1).

Na^+ channels comprise four repeating domains (DI-DIV), interconnecting extra- and intracellular loops, and N and C termini. Sequences from all mammals, *Xenopus*, *Anolis*, and *Gallus* were mostly full length (N to C termini)

with the occasional small region of missing sequence due to gene assembly problems. Due to the low coverage (1.4x) of the *Callorhinchus* genome, contigs were short and typically contained only one or a few exons.

Sequences derived by RT PCR

We cloned additional Nav channel sequence by RT-PCR from various tissues of a few key species for which sequenced genomes were unavailable. Lungfish (Dipnoi) and coelocanths are the most basal living tetrapods so we cloned Nav channel transcripts from heart, muscle, brain, and spinal cord of the South American lungfish (*Lepidosiren paradoxica*). Additionally, we cloned Nav channel genes from genomic DNA of lungfish to attempt to capture any genes that might be expressed in low levels or not expressed in the tissues from which we extracted RNA.

The Chondrichthyes (e.g.-sharks, rays, skates, chimeras) diverged from the lineage leading to tetrapods ~525 MYA in the mid-Cambrian period (Hedges 2009), presumably following the second vertebrate WGD (Kuraku 2008). We were able to obtain muscle and heart from a Horn Shark (*Heterodontus francisci*) and the brain of an Atlantic stingray (*Dasyatis Sabina*) and cloned Nav channel transcripts from them to estimate the number of Nav channel genes in vertebrates well before the emergence of tetrapods.

We used nested primer sets developed in our laboratory for cloning Nav channels from a range of vertebrate species. These primarily targeted domains I, II and/or III, and variable interconnecting intracellular loops. The RT reaction

either utilized the lower primer of step one to attempt to target Nav channel genes, or a polyTTT primer to insure that we had not missed any transcripts. For lungfish the primers were as follows: 1st PCR (Upper: CCRTGGAAYTGKCTKGATTT; Lower: RTRAARRADGABCCRAADRTGATG); 2nd PCR: (Upper: ATGRCGTAYVYYACVGAGTT; Lower: TACATDATNYCCATCCANCCTTT). We used the following primer sets for: a) shark heart: 1st PCR (Upper: TGYGGYGARTGGATYGARAC; Lower: RTRAARRADGABCCRAADRTGATG); 2nd PCR: (Upper: ATGTGGGAYTGYATGGARGT; Lower: TACATDATNYCCATCCANCCTTT); b) shark muscle: 1st PCR (Upper: TCYMGAGGBTTCTGYDTTGG; Lower: RTRAARRADGABCCRAADRTGATG); 2nd PCR: (Upper: CCRTGGAAYTGKCTKGATTT; Lower: TACATDATNYCCATCCANCCTTT). For skate brain we used a similar primer set as in lungfish muscle. Temperatures and times were 53^{o}C for annealing and 94^{o}C for denaturing steps (30 - 45 seconds), and 74^{o}C for extension steps (extension time dependent on the length of predicted PCR products (1 min/1000bp) for a total of 35 cycles.

Nav channel Phylogeny

Nav channel amino acid sequences were aligned in CLUSTALX using default parameters, and poorly aligned regions, mainly long intracellular loops, were removed. Final alignments were output into NEXUS files. To reconstruct Nav gene phylogeny, two independent Bayesian analyses (Mr. Bayes, version 3) were conducted, each with 5 MCMC chains and run for 200,000 generations,

assuming 6 substitution types and a gamma distribution of rate variation among sites. For each analysis, posterior probabilities were calculated using majority-rule consensus of all trees saved after the log likelihood asymptote (burn-in). Lamprey Nav channel gene sequences were specified for rooting.

Analysis of non-Nav channel genes

We wished to determine whether duplications of non-Nav channel genes occurred within the early tetrapod lineage following the teleost-tetrapod divergence. Human nucleotide sequences were initially used to search the NCBI (nr/nt) database for orthologous or paralogous sequences from chicken, frog, and teleosts (Table S2). Then, all of the above nucleotide sequences were used for BLAT searches of Ensemble genome databases of zebrafish and chicken to assure that no unannotated genes were missed. No attempt was made to reconstruct the history of these genes within teleosts, so extensive searches were not made of teleost genomes. However, if no teleostean ortholog was initially recovered from GenBank or from the zebrafish genome, further searches were made in the genomes of stickleback (*Gasterosteus aculeatus*), medaka (*Oryzias latipes*), and pufferfish (*Takifugu rubripes*, *Tetraodon nigroviridis*) in case a putative teleostean ortholog might have been lost in zebrafish but retained in other teleost lineages. Finally, the teloestean sequences were used for BLAT searches of the NCBI and Ensemble databases to ensure that all tetrapod orthologs had been uncovered. If a gene had duplicated in tetrapods, each duplicate would be equally likely to be identified by its teleostean ortholog.

Analysis of teleostean and tetrapod non-Nav channel amino acid trees was done using the neighbor-joining algorithm with default values in ClustalX with 1,000 bootstrap replications.

Determination of Synteny

Chromosomal location of human and chicken Nav channel and flanking genes was available from NCBI. Synteny was determined for lizard and frog Nav and flanking genes by manually assembling genes from scaffolds from Ensemble, determining gene identify by BLAT with the NCBI database, and by their locations in comparison with human and chicken chromosomes.

Results

Description of Data

From BLAT searches on archived genomes we recovered six Nav channel genes from *Xenopus*, nine from lizard, six from platypus (plus two smaller fragments that were not used for analysis), eight from opossum, and nine from chicken. Where possible predicted *X. tropicalis* gene sequences were confirmed by ESTs and in two cases (xt236, xt464b) incomplete genomic sequences was filled in by overlap with ESTs. (Fig. S1).

We derived pieces of Nav channel genes from elephant shark. Due to the low coverage, each contig had one or at most a few exons and contigs could not be unambiguously connected. However, we identified three to four distinct pieces corresponding to most exons of the Nav channel gene that suggested a total of

four Nav channel genes. We recovered four contigs with all or part of the most 3' exon (the longest exon in all vertebrate Nav channel genes); one contig included two other exons. This gave us sequences of from ~200 to 500 amino acids.

By RT-PCR we cloned three Nav channel genes from lungfish, one each from brain and spinal cord, muscle, and heart tissue; two from horn shark, one from muscle and one from heart, and one from skate brain. These were not complete sequences but covered domains II-III (lungfish muscle and skate brain), or domains I-III (shark and lungfish heart, lungfish brain, lungfish muscle). Given the limitations of our tissues samples and the RT-PCR method, we do not claim that this provides the full complement of Nav channel genes of these species.

As we show below, we have a strong case for homologizing the amniote Nav channel genes; therefore, we use the mammalian gene nomenclature (SCNxA, where x = a number) for Nav channel genes from mammals, birds, and lizard (protein designation = Nav; gene designation = SCN). But because not all of the non-amniote genes are orthologs of amniote genes, we did not use the mammalian nomenclature for these.

Expansion of Nav channel genes occurred in the Devonian and Carboniferous periods

We generated phylogenetic trees without (alignment1: Fig. 1) and with (alignment2: Fig. S2) the four elephant shark sequences. Resulting tree topologies of alignment1 were identical for the two runs with minor variation in the posterior probabilities for a few branches. Alignment2 was run only once due

to the low support values in the parts of the tree including the elephant shark sequences (Fig. S2). We will focus on alignment1.

SCN8A has a simple history with no duplications tracing back to an ancestral gene that is also represented in elasmobranchs (skatebrain), lungfish (lungfishbrain), and amphibians (xt67) (Fig. 1: shaded in light blue). There are two duplicates in zebrafish due to a teleost-specific WGD. As in humans, *SCN8A* orthologs of frog, chicken, and lizard reside alone on a chromosome (Table SI).

SCN4A shows a similar history with amniote orthologs grouping with genes from frog (xt43), lungfish (lungfishmuscle), and shark (sharkmuscle) and a pair of duplicate genes in teleosts (Fig. 1: shaded in yellow). The orthologs of SCN4A also reside singly on a chromosome (Table SI).

Orthologs of the three Nav channel genes on human chromosome 3 (*SCN5A, SCN10A, SCN11A*) are found in other mammals, chicken, and lizard and these three genes have shared synteny (Fig. 1: shaded in light red). They derive from a single gene represented in our sample by orthologs in shark (sharkheart), zebrafish (SCN5Laa, SCN5Lab), lungfish (lungfishheart), and frog (xt28). The single frog gene is syntenically related to the amniote genes (Fig. 2). The presence of a single gene at the amphibian-amniote split and of three genes before the synapsid (mammals)-diapsid (reptiles and birds) divergence means that two duplications of the ancestral gene occurred in a 30 MY window at the origin of amniotes in the lower- to mid-Carboniferous periods.

The history of the Nav channel genes on human chromosome 2 is more complex. These genes derive from a fourth ancestral gene (Novak et al 2006) although we were unable to detect an ortholog of this gene in lungfish tissues or skate brain despite extensive attempts to amplify it from RNA and genomic DNA. As expected, there are two gene duplicates in zebrafish (SCN1Laa, SCN1Lab). The gene tree suggests that a single ancestral gene underwent independent duplications in amphibians and amniotes (Fig. 1: shaded in light green). In this scenario multiple duplications of a putative ancestral Nav channel gene would have generated SCN1A, SCN2A, SCN3A and SCN7A/SCN9A (the precursor to separate SCNA7a and SCN9A genes). Because these all have orthologs in mammals, lizard, and chicken, these duplications would have occurred within the same 30 MY window as the triplicated genes on human chromosome 3 (Fig. 1). The final duplication of SCN7A/SCN9A into SCN7A and SCN9A likely occurred after the divergence of monotreme and therian mammals (220 MYA) preceding the marsupial-placental split (175 MYA). However, given the low values of the posterior probabilities in our trees, there is some uncertainty about the timing of this duplication.

On the other hand, the most parsimonious interpretation of the synteny (Fig. 3) is that two duplications had already occurred in the common ancestor of amphibians and amniotes. This is because the Nav channel and other genes in this region of the amphibian and amniote chromosomes have the same syntenic relationships. Additionally, one Nav channel gene in each lineage and in the same relative chromosomal position (amphibian xt464b and amniote SCN2A) is

oppositely oriented on the chromosome to all the other Nav channel genes (arrows, Fig. 3). This suggests that most of the duplications occurred at a slower rate over 130 MY.

What might account for the discrepancy in interpretation between the gene tree and synteny data? It is unlikely to be due to *Xenopus* Nav channel genes that were missed or misassembled in genome sequencing. First, all the ESTs that we uncovered uniquely matched a specific gene (in most cases multiple ESTs were mapped to each gene) and all *Xenopus* Nav channel genes were represented in the EST database (Fig. S1). Second, the few additional amphibian Nav channel genes available from GenBank of sufficient length to align (e.g.-newt, xlaev1.2) appeared to be orthologs of genes that we had already uncovered in *X. tropicalis*. Finally, the *Xenopus* scaffolds were assembled from overlapping reads of shotgun sequence *de novo* so that the apparent synteny is not an artifact (Hellsten et al 2010). It is possible, but seems unlikely, that independent duplications in amphibians and amniotes could have resulted in identical patterns of synteny. If the duplications had occurred in the common ancestor of amphibians and amniotes as suggested by synteny, then the non-overlapping clustering of amphibian and amniote genes in the tree might be explained by some amount of gene conversion within the amniote and/or amphibian lineages, as sometimes occurs following gene duplications (Kellis et al 2004).

Alignment2 included the fragments from elephant shark (Fig. S2). As expected, the inclusion of these short sequences resulted in low posterior

probabilities for them and their neighboring branches; these were too low to trust their exact positioning in the tree. However, each of the four sequences grouped with one of the four clades of Nav channel genes with extremely strong (posterior probabilities = 100) support. Furthermore, a BLAT search of GenBank with the nucleotide sequences of each of the four 3' exons had top matches with sequences in one specific clade. The inclusion of these sequences supports our contention (Novak et al 2006) that the ancestor of teleosts and tetrapods had four Nav channel genes and gives an indication of the "missing" ancestor of the fourth Nav channel gene clade.

In sum, we suggest that after the second WGD (estimated at ~550 MYA, Meyer & Schartl 1999; Dehal & Boore 2005; Panopoulou & Poustka 2005; Blomme et al 2006), brought the number of Nav channel genes to four there was a period of stasis. Then a series of tandem Nav channel gene duplications occurred in a 30-130 MY period during early tetrapod evolution, following which the Nav channel gene family remained largely stable for another 300 MY (Fig. 5).

Genes flanking the Nav channel genes did not duplicate or were not retained

We posit that the retention of the Nav channel gene paralogs was due to selection. The null hypothesis is that the Nav channel gene expansion was simply a consequence of instability in the regions of these two chromosomes in which the Nav channel genes reside that led to duplication and retention of all the genes in this region. We tested the null hypothesis by examining whether

flanking genes show a similar history of duplication and retention. We sampled only those flanking genes that were located on chromosomes or scaffolds of the species for which we had synteny information. In our sample (Fig. 2,3) 14 of 15 genes showed no tandem duplications on these chromosomes, with SLC22A13 and SLC22A14 being a duplicate pair. Given that the Nav channel genes on these two chromosomes have both duplicated (2/2: this analysis ignores the fact that the Nav channel genes underwent multiple duplications), this is significantly greater than expected given the number of duplications of the flanking non-Nav channel genes on these two chromosomes (1/14) ($p<0.0001$, two tailed chi square).

However, it is also possible that these flanking genes duplicated but were then dispersed throughout the genome whereas the Nav channel genes were retained where they duplicated. We derived sequences for the 16 flanking genes in figures 2 and 3 from GenBank or searches from genome databases, and constructed trees for them. We found that most showed indications of the initial two rounds of WGD, but no duplications in the 450 MY following the actinopterygian-sarcopterygian split (Fig. 4A). The only exception, mentioned above, was the duplicate pair of SLC22A members13 and 14. This gene family has a history of extensive duplication (>25 members) with the “center of gravity” of duplication elsewhere in the genome.

We conclude that the genes flanking the Nav channel genes either did not duplicate or, if they did so, were not retained. This supports the hypothesis that the Nav channel gene duplications were retained as a result of selection.

Other ion channel gene families did not duplicate or were not retained

Voltage-dependent ion channels are fundamental to the electrical activity of the brain. We next asked whether there was a general expansion of other six trans-membrane (6TM) voltage-dependent ion channel gene families during tetrapod evolution or whether this expansion was specific to the Nav channel gene family. We addressed this question using both published literature (Saito & Shingai 2006; Hoegg & Meyer 2007; Jackson et al 2007) and gene trees that we constructed with sequences from teleost, human, frog, and chicken databases. The channel that we investigated included the major depolarizing (Ca^{2+}, TRP, HCN) and hyperpolarizing (Kv, ERG, slo) channels. We sampled all 17 members of the voltage-dependent K^+ channel (Kv), four members of the hyperpolarization-activated cyclic nucleotide-gated ion channel (HCN), seven members of the canonical transient receptor potential (TRPC), six members of the transient receptor potential, vanilloid-sensitive (TRPV), four members of the transient receptor potential, melastatin (TRPM), one member of the transient receptor potential ankyrin repeat (TRPA), eight members of the ether-a-go-go related (ERG), three members of the large-conductance calcium-activated K^+ family (slo), and 10 members of the voltage-dependent calcium (CACNA1 or Cav) gene families (Fig. 4B).

Most (54/60) of the ion channel genes in our sample showed no duplications following the teleost-tetrapod divergence. There is a mammalian-

specific duplication in the slo gene family (slo3). A number of TRP channels do not have teleost orthologs (TRPV1/TRPV2, TRPV3, TRPV5/TRPV6, TRPM6, TRPM8). However, the absence of TRPM8 and TRPV3 in teleosts is likely due to losses in teleosts rather than duplications in amniotes (Saito & Shingai 2006). TRPV5/TRPV6 are a pair that clearly duplicated in amniotes (Saito & Shingai 2006). The timing of the TRPV1/TRPV2 duplication is not resolvable (Saito & Shingai 2006), but it may represent a tetrapod-specific duplication. Even assuming that all apparent duplications in tetrapods are real rather than reflecting losses of teleost genes, Nav channel genes duplicated (2/4: again, this analysis ignores repeated duplications of each ancestral paralog) significantly more than other ion channel genes (6/60) ($p < 0.007$, two tailed chi square).

Discussion

Relationship to previous Nav channel phylogenies

Based on chromosomal location and phylogeny of the Nav channel family in mammals, Plummer & Meisler (1999) proposed that the 10 mammalian Nav channels resulted from a single ancestral chordate Nav channel gene that underwent two rounds of genome duplication early in vertebrate evolution. These duplications ended in a single Nav channel gene on each of four chromosomes, followed by a series of tandem duplications on two of those chromosomes. Inclusion of teleosten Nav channel sequences in subsequent analyses supported

this idea, and further demonstrated that: the initial two rounds of duplications preceded the teleost-tetrapod divergence; teleosts have eight Nav channel genes likely as the result of a WGD; the duplications in teleosts and tetrapods occurred independently (Lopreato et al 2001; Novak et al 2006). Novak et al (2006), which also included chicken sequences, did not attempt to resolve the timing of the tandem duplications within tetrapods. The trees from that study would suggest that many of the Nav channel genes from chicken and mammals duplicated independently. However, the sequences then available from the chicken were likely misassembled (and some have been discontinued by NCBI) thereby confusing the relationship of mammalian and avian genes. In the current study we addressed this issue by determining the chicken sequences manually from a newer Ensemble release of the chicken genome (version 2.1). Furthermore, we added sequence from key taxa such as amphibians, lizard, and monotreme and marsupial mammals as well as including sequences cloned in our laboratory from lungfish and elasmobranchs. A strong conclusion from the current study is that, with the exception of a single mammalian-specific duplication, all mammalian Nav channel genes were present in an early amniote ancestor (Fig. 5).

In addition, the current phylogeny supports our and others' previous conclusion that four Nav channel genes existed in the common ancestor of teleosts and tetrapods. A novel result in the current study is that we identified orthologs of these four genes in lungfish and elasmobranchs. Orthologs to three of these four were easily cloned from lungfish and elasmobranch tissues. Despite numerous attempts, we were unable to amplify the fourth gene, the ancestor of

the SCNA1, SCN2A etc complex (which are mainly neural-expressing genes), from lungfish brain, spinal cord or skate brain or lungfish genomic DNA. However, we identified a likely ortholog in the elephant shark genome. We suggest that this gene is not expressed, or at least not at high levels, in the CNS of elasmobranchs and lungfish.

Nav Channel gene expression and tetrapod evolution

The Nav channel gene family expansion occurred in two chromosomal regions. It was not accompanied by duplication and/or retention of flanking genes nor by widespread expansion of other ion channel genes; it was specific and presumably advantageous. The Nav channel gene tandem duplications were concurrent with the origin of tetrapods and their invasion of the terrestrial habitat. It largely preceded the diversification of amniotes into synapsid (mammal) and diapsid (reptiles/birds) lineages (Hedges 2009; Shedlock & Edwards 2009).

As tetrapods took to land, they evolved new modes of locomotion, coped with loss of buoyancy, confronted a novel sensory environment, and exploited new food resources (Glenner et al 2006). Meeting these challenges was facilitated by the evolution of new sensory receptors in their skin and muscles. For example, early tetrapods evolved muscle spindles (Maeda et al 1983; Ross et al 2007) and different lineages of tetrapods later evolved other kinds of somatosensory receptors (e.g.—lamellated Pacinian/Herbst corpuscles in amniotes, dome pressure receptors in crocodilians, etc) (Hunt 1961; Proske

1969; Dorward & Macintyre 1971; von Düring & Miller 1979; Soares 2002). This was accompanied by greater anatomical and physiological complexity in the dorsal root ganglion system (Sneddon *2002;* Sneddon et al 2003). Little is known about the expression of Nav channel genes in the dorsal root ganglia of non-mammalian tetrapods. In mammals, a number of different Nav channels are expressed in dorsal root ganglion neurons and axons--Nav1.6, Nav1.3, Nav1.8, Nav1.9, Nav1.7--and these contribute to the systematic variation in conduction velocity and other biophysical properties of different classes of dorsal root ganglion neurons (Herzog et al 2003). Of these channels, all but Nav1.6 derive from the tetrapod-specific duplications. Thus, the duplications of Nav channel genes enabled the diversity of Nav channels types in dorsal root ganglion cells.

Besides the obvious advantages of possessing a repertoire of Nav channels that can be called upon for performing various coding "jobs", matching Nav channel biophysical properties with signaling requirements also results in metabolic efficiency (Hasenstaub et al 2010; Schmidt-Hieber & Bischofberger 2010). A more expansive repertoire of Nav channel genes might also have been selected on the basis of energy savings.

The Nav channel gene duplications also occurred at a time when the amniote brain, especially the forebrain, was robustly expanding and adding new anatomical regions (Northcutt 2002). We do not believe that the increase in number of Nav channel genes was causal to the increase in forebrain complexity. Rather, these likely happened in parallel, both as manifestations of increasing brain complexity. Also, since we have only examined 6TM ion channel gene

families, we cannot comment on the possible role of other brain-expressing genes, such as neurotransmitters and their receptors, etc in the evolution of the amniote brain.

Parallel expansion of Nav channel gene family in teleosts

Global cataclysmic events in the late Devonian initiated a period of mass extinction for most vertebrate lineages (Sallan & Coates 2010). The main survivors of this event were the tetrapods (the other sarcopterygian taxa mostly went extinct and are represented today solely by lungfishes and coelacanths), on land, and the actinopterygii (and chondrichthyes), in water. Similar to the evolutionary success of tetrapods, one group of actinopterygii, the teleosts, eventually came to dominate the marine and aquatic habitats.

Interestingly, the Nav channel gene family independently expanded in teleosts to almost the same number of genes (eight Nav channel genes) as in amniotes (Lopreato et al 2001; Novak et al 2006). In contrast to tetrapods, this duplication arose suddenly as a result of a third teleost-specific WGD (Meyer & Schartl 1999; Hurley et al 2007) and no tandem duplications occurred over the next ~250-300MY. As in tetrapods, all the Nav channel gene duplicates have been retained. It will be interesting to reconstruct the detailed histories of other ion channel gene families in teleosts and determine if these are all retained as the Nav channel genes have been or whether there has been greater loss of other ion channel genes back to a “baseline” pre-WGD number. In other words, is there a relative increase in Nav over other ion channel genes in teleosts as in

tetrapods? In a further parallel to the amniotes, the teleostean forebrain also increased in complexity compared to that of non-teleost actinopterygian fishes (e.g.--bowfin, gar, sturgeon) (Northcutt 2002).

Recent Nav channel gene duplicates are differentially expressed

It has been proposed that recently duplicated genes show more restricted expression or greatest sequence divergence than those that duplicated in the distant past (Farré & Alba 2009; Milinkovitch et al 2010). *SCN4A*, with no history of duplication since the last WGD is predominantly expressed in mammalian muscle; in lungfish its ortholog was only expressed in muscle. *SCN8A*, also with no history of duplication since the WGD, is expressed in brain (and not heart or muscle) in lungfish and is expressed at uniform levels throughout the mammalian brain (Whitaker et al 2000; Whitaker et al 2001) (no data from reptiles and birds; although SCN8A is also expressed in mammalian heart, Maier et al 2004). On the other hand, two of the complex of triplicated genes represented by human chromosome 3 and one from the complex localized to human chromosome 2 are expressed in neurons of the peripheral somatosensory system, and some have "unusual" biophysical properties (Akopian et al 1996; Cummins et al 1999; Cummins et al 2001; Dib-Hajj et al 2002).

The complex of genes that shows the greatest duplication (human chromosome 2) are mainly expressed in brain and show the greatest variation in regional patterns of expression (telencephalon *vs* brainstem), or in sub-cellular distribution (axons *vs* somata) in the mammalian brain (Westenbroek et al 1989;

Akopian et al 1996; Cummins et al 1999; Caldwell et al 2000; Whitaker et al 2000; Cummins et al 2001; Whitaker et al 2001; Dib-Hajj et al 2002; Jarnot & Corbett 2006; Ogiwara et al 2007; van Wart et al 2007; Duflocq et al 2008; Hu et al 2009). The final gene duplication that occurred before the origin of therian mammals, gave rise to a unique Nav channel with a highly derived sequence that has lost its voltage-sensitivity but is still permeable to Na^+ ions (Na_x) (Watanabe et al 2006). This channel is more involved in Na^+ ion regulation than neural computation.

Conclusions

The Nav channel gene family of tetrapods underwent a series of duplications 300-450 MYA largely during their early evolution (Fig. 5). This wave of duplications did not involve the duplication or retention of flanking genes or other ion channel genes. We speculate that the rapid expansion of the Nav channel gene family accommodated greater complexity in neural processing and was a seminal event in the evolution of the amniote brain.

Supplementary Material

Supplementary table I: Gene names and chromosomal/scaffold/contig locations, and Accession number of amino acid sequences used in the phylogenetic analysis.

Supplementary table II: Information on other gene sequences used in this study.

Supplementary figure 1: Schematic overlay of *X. tropicalis* ESTs (short lines) on Nav channel genes (long bottom line).

Supplementary figure 2: Alignment2. Nav channel gene phylogeny with the inclusion of short sequences from the elephant shark genome (bolded). The inclusion of these short sequences perturbs the placement of surrounding branches and lowers posterior probability scores. However, each of the four elephant shark genes groups with high support (posteriors = 100) with one of the four clades of Nav channel genes.

Acknowledgements: The authors gratefully acknowledge Marianna Grenadier for artwork and Dr. Robert Dores, University of Denver, for supplying tissues, and Dr. Hans Hofmann for commenting on the MS. This work was funded by the National Science Foundation (IBN 0236147 to H.H.Z and M.C.J), and the National Institutes of Health (R01GM084879 to H.H.Z).

Literature Cited

Akopian A, Sivilotti L, Wood J. 1996. A tetrodotoxin-resistant voltage-gated sodium channel expressed by sensory neurons. Nature 379:257–262

Angelino E, Brenner M. 2007. Excitability constraints on voltage-gated sodium channels. PLoS Comput Biol 3:e177

Blomme T, Vandepoele K, De Bodt S, Simillion C, Maere S, Van de Peer Y. 2006. The gain and loss of genes during 600 million years of vertebrate evolution. Genome Biology 7:R43

Caldwell J, Schaller K, Lasher R, Peles E, Levinson S. 2000. Sodium channel $Na_v1.6$ is localized at nodes of Ranvier, dendrites, and synapses. Proc. Natl Acad. Sci. USA 97:5616-5620

Campbell D. 1992. Large and small vertebrate sensory neurons express different Na and K channel subtypes. Proc. Natl Acad. Sci. USA 89:9569-9573

Cummins T, Aglieco F, Renganathan M, Herzog R, Dib-Hajj S, Waxman S. 2001. Nav1.3 sodium channels: rapid repriming and slow closed-state inactivation display quantitative differences after expression in a mammalian cell line and in spinal sensory neurons. J Neurosci. 21:5952-5961

Cummins T, Dib-Hajj S, Black J, Akopian A, Wood J, Waxman S. 1999. A novel persistent tetrodotoxin-resistant sodium current In SNS-null and wild-type small primary sensory neurons. J. Neurosci. 19:1-6

Dehal P, Boore J. 2005. Two rounds of genome duplication in the ancestral vertebrate. PLoS Biology 3:e314

Dib-Hajj S, Black J, Cummins T, Waxman S. 2002. NaN/Nav1.9: a sodium channel with unique properties. Trends Neurosci. 25:253-259

Dorward P, Macintyre A. 1971. Responses from vibration-sensitive receptors in the interosseous region of the duck's hind limb. J. Physiol. 219:77-87

Duflocq A, Le Bras B, Bullier E, Couraud F, Davenne M. 2008. Nav1.1 is predominantly expressed in nodes of Ranvier and axon initial segments. Mol. Cell. Neurosci. 39:180-192

Farré D, Alba M. 2009. Heterogenous patterns of gene-expression diversification in mammalian gene duplicates. Mol. Biol. Evol. 27:325-335

Glenner H, Thomsen P, Hebsgaard M, Sørensen M, Willwerslev E. 2006. The origin of insects. Science 314:1883-1884

Hains B, Klein J, Saab C, Craner M, Black J, Waxman S. 2003. Upregulation of sodium channel Nav1.3 and functional involvement in neuronal hyperexcitability associated with central neuropathic pain after spinal cord injury. Journal of Neuroscience 23:8881-8892

Hasenstaub A, Otte S, Callaway E, Sejnowski TJ. 2010. Metbolic cost as a unifying principle governing neuronal biophysics. Proc. Natl Acad. Sci. USA 107:12329-12334

Hedges S. 2009. Vertebrates (Vertebrata). New York: Oxford University Press. 309-314 pp.

Hellsten U, Harland R, Gilchrist M, Hendrix D, Jurka J, et al. 2010. The Genome of the Western Clawed Frog *Xenopus tropicalis*. Science 328:633-636

Herzog R, Cummins T, Ghassemi F, Dib-Hajj S, Waxman S. 2003. Distinct repriming and closed-state inactivation kinetics of Nav1.6 and Nav1.7 sodium channels in mouse spinal sensory neurons. Journal of Physiology 551:741-750

Hill A, Nishino A, Nakajo K, Zhang G, Fineman J, et al. 2008. Ion channel clustering at the axon initial segment and node of Ranvier evolved sequentially in early chordates. PLoS Genet. 4:e1000317

Hoegg S, Meyer A. 2007. Phylogenomic analyses of KCNA gene clusters in vertebrates: why do gene clusters stay intact? BMC Evolutionary Biology **7**:139

Hu W, Tian C, Li T, Yang M, Hou H, Shu Y. 2009. Distinct contributions of Nav1.6 and Nav1.2 in action potential initiation and backpropagation. Nature Neuroscience 12:996–1002

Hunt C. 1961. On the nature of vibration receptors in the hind limb of the cat. J. Physiol. **155**:175-186

Hurley I, Mueller R, Dunn K, Schmidt E, Friedman M, et al. 2007. A new time-scale for ray-finned fish evolution. Proc. R. Soc. B 274:489-498

Jackson H, Marshall C, Accill E. 2007. Evolution and structural diversification of hyperpolarization-activated cyclic nucleotide-gated channel genes. Physiol Genomic 29:231-245

Jarnot M, Corbett A. 2006. Immunolocalization of NaV1.2 channel subtypes in rat and cat brain and spinal cord with high affinity antibodies. Brain Research **1107**:1-12

Kellis M, Birren B, Lander E. 2004. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. Nature 28:617-624

Kuraku S. 2008. Insights into Cyclostome Phylogenomics: Pre-2R or Post-2R. Zoological Science 25:960-968

Lopreato G, Lu Y, Southwell A, Atkinson A, Hillis D, et al. 2001. Evolution and divergence of sodium channel genes in vertebrates. Proc. Natl Acad. Sci. USA 98:7588-7592

Lorincz A, Nusser Z. 2010. Molecular identity of dendritic voltage-gated sodium channels. Science 328:906-909

Madsen O. 2009. Mammals (Mammalia). New York: Oxford University Press. 459-61 pp.

Maeda N, Miyoshi S, Toh H. 1983. First observation of a muscle spindle in fish. Nature 302:61-62

Maier SKG, Westenbroek RE, McCormick KA, Curtis R, Scheuer T, Catterall WA. 2004 Distinct subcellular localization of different sodium channel α and β subunits in single ventricular myocytes from mouse heart. Circulation 109:1421-1427

Meyer A, Schartl M. 1999. Gene and genome duplications in vertebrates: the one-to-four (-to eight in fish) rule and the evolution of novel gene function. Curr. Opin. Cell Biol. 11:699-704

Milinkovitch M, Helaers R, Tzika A. 2010. Historical Constraints on Vertebrate Genome Evolution. Genome Biology and evolution 2:13-18

Northcutt R. 2002. Understanding brain evolution. Integratie and Comparative Biology 42:743-756

Novak A, Jost M, Lu Y, Taylor A, Zakon H, Ribera A. 2006. Gene duplications and evolution of vertebrate voltage-gated sodium channels. J. Mol. Evol. 63: 208-221

Ogiwara I, Miyamoto H, Morita N, Atapour N, Mazaki E, et al. 2007. $Na_v1.1$ localizes to axons of parvalbumin-positive inhibitory interneurons: a circuit basis for epileptic seizures in mice carrying an Scn1a gene mutation. J. Neurosci. 27:5903-5914

Okamura Y, Ono F, Okagaki R, Chong J, Mandel G. 1994. Neural expression of a sodium channel gene requires cell-specific interactions. Neuron 13:937-948

Panopoulou G, Poustka A. 2005. Timing and mechanisms of ancient vertebrate genome duplications--the adventure of a hypothesis. Trends in Genetics 121:559-567

Piontkivska H, Hughes A. 2003. Evolution of vertebrate voltage-gated ion channel alpha chains by sequential gene duplication. J Mol Evol. 56:277-285

Plummer N, Meisler M. 1999. Evolution and diversity of mammalian sodium channel genes. Genomics 57:323-331

Proske U. 1969. Vibration sensitive mechanoreceptors in snake skin. Experimental Neurology 23:187-194

Ross C, Eckhardt A, Herrel A, Hylander W, Metzger K, et al. 2007. Modulation of intra-oral processing in mammals and lepidosaurs. Integr. Comp. Biol. 47:118-136

Saito S, Shingai R. 2006. Evolution of thermo TRP ion channel homologs in vertebrates. Physiol. Genomics 27:219-230

Sallan L, Coates M. 2010. End-Devonian extinction and a bottleneck in the early evolution of modern jawed vertebrates. Proc. Natl Acad. Sci. USA 107:10131-10135

Schmidt-Hieber C,Bischofberger J. (2010) Fast sodium channel gating supports localized and efficient axonal action potential initiation. J. Neurosci. 30:10233-10242

Shedlock A, Edwards S. 2009. Amniotes (amniota). New York: Oxford University Press. 375-379 pp.

Sneddon L. 2002. Anatomical and electrophysiological analysis of the trigeminal nerve in a teleost fish, *Oncorhynchus mykiss*. Neuroscience Letters 319:167-171

Sneddon L, Braithwaite V, Gentle M. 2003. Do fishes have nociceptors? Evidence for the evolution of a vertebrate sensory system. Proc. R. Soc. Lond. B 270:1115-1121

Soares D. 2002. Neurology: An ancient sensory organ in crocodilians. Nature 417:241-242

van Wart A, Trimmer J, Matthews G. 2007. Polarized distribution of ion channels within microdomains of the axon initial segment. J. Comp. Neurol. 500:339-352

von Düring M, Miller M. 1979. Sensory nerve endings in the skin and deep structures. In Biology of the Reptilia, ed. C Gans, R Northcutt, P Ulinski, pp. 407-442. New York: cademic Press

Watanabe E, Hiyama T, Shimizu H, Kodama R, Hayashi N, et al. 2006. Sodium-level-sensitive sodium channel Na(x) is expressed in glial laminate processes in the sensory circumventricular organs. Am J Physiol Regul Integr Comp Physiol. 290:R568-576.

Westenbroek R, Merrick D, Catterall W. 1989. Differential subcellular localization of the R_I and R_{II} Na^+ channel subtypes in central neurons. Neuron 3:695–704

Whitaker W, Clare J, Powell A, Chen Y-H, Faull R, Emson P. 2000. Distribution of voltage-gated sodium channel alpha-subunit and beta-subunit mRNAs in human hippocampal formation, cortex, and cerebellum. J Comp Neurol 422: 123-139

Whitaker WR, Faull RL, Waldvogel HJ, Plumpton CJ, Emson PC, Clare JJ. 2001. Comparative distribution of voltage-gated sodium channel proteins in human brain. Brain research. Molecular brain research 88:37-53

Figure Legends.

Figure 1—Phylogenetic tree for amino acid translation of tetrapod Nav channel genes determined by Bayesian analysis. Amniote Nav channel genes are homologized with and named according to human gene nomenclature. Genes from the frog, *Xenopus tropicalis*, are named after their scaffold location, genes from lungfish, skate, and shark according to the tissue from which they were amplified. Names of the mammalian Nav channel proteins are given on the right side of the figure. Human Nav channel genes that are on the same chromosome and their orthologs from other species are within a block of color.

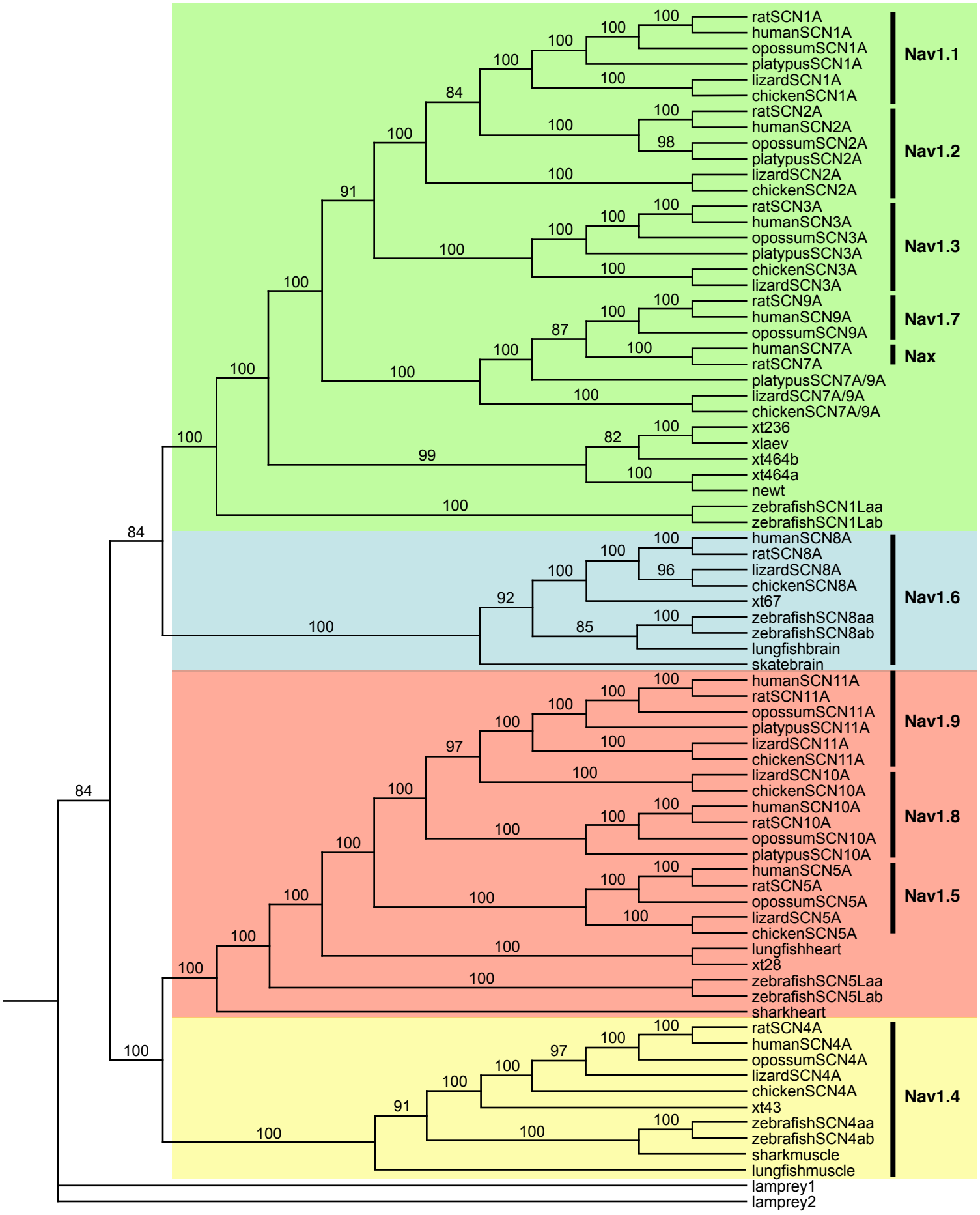
Figure 2—Synteny for tetrapod Nav channel genes referenced to Nav channel genes on human chromosome 3. Black boxes represent Nav channel genes, gray boxes represent other genes.

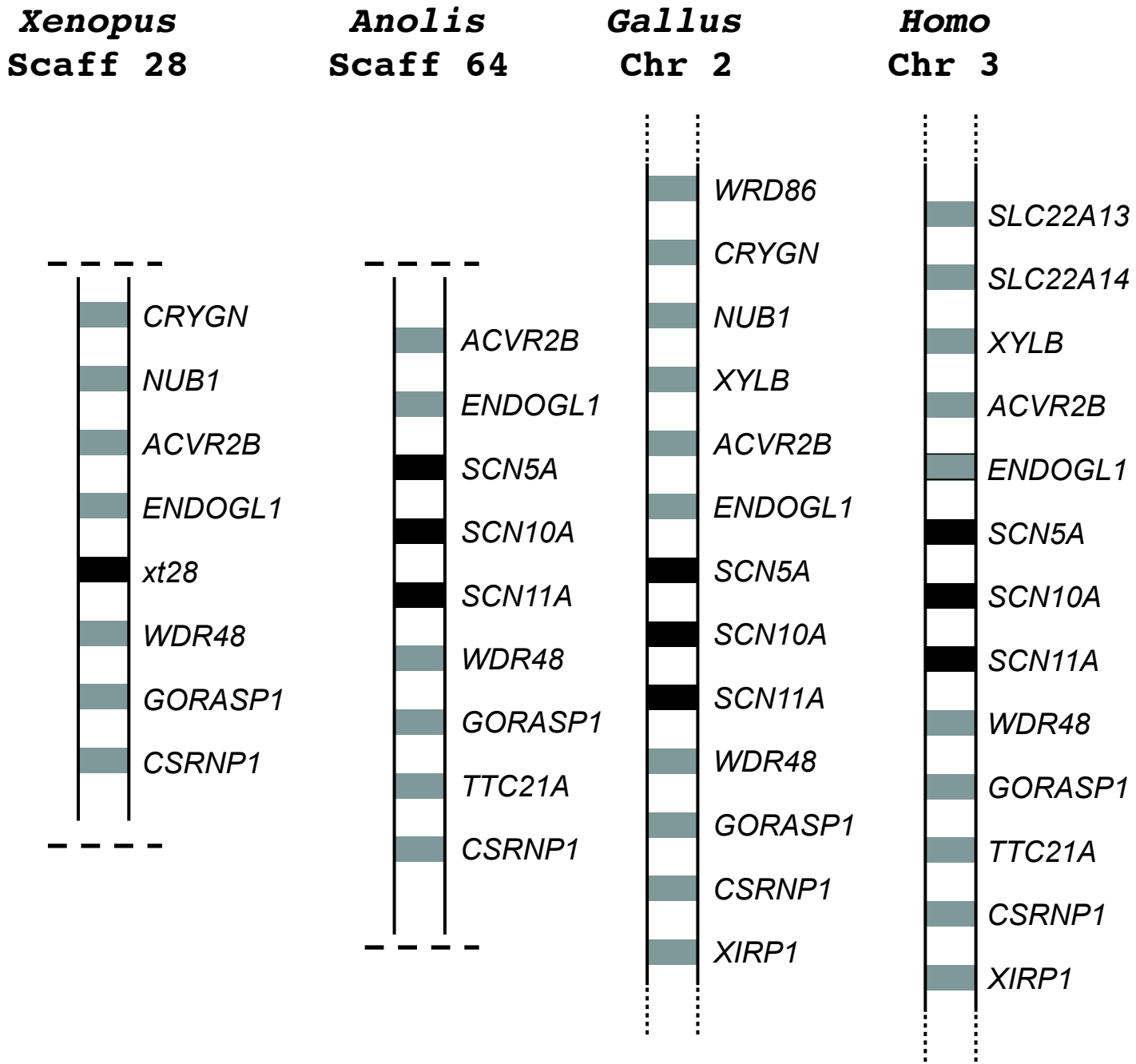
Figure 3--Synteny for tetrapod Nav channel genes referenced to Nav channel genes on human chromosome 2. Black boxes represent Nav channel genes, gray boxes represent other genes. Arrows represent Nav channel gene whose chromosomal orientation is opposite all the other Nav channel genes.

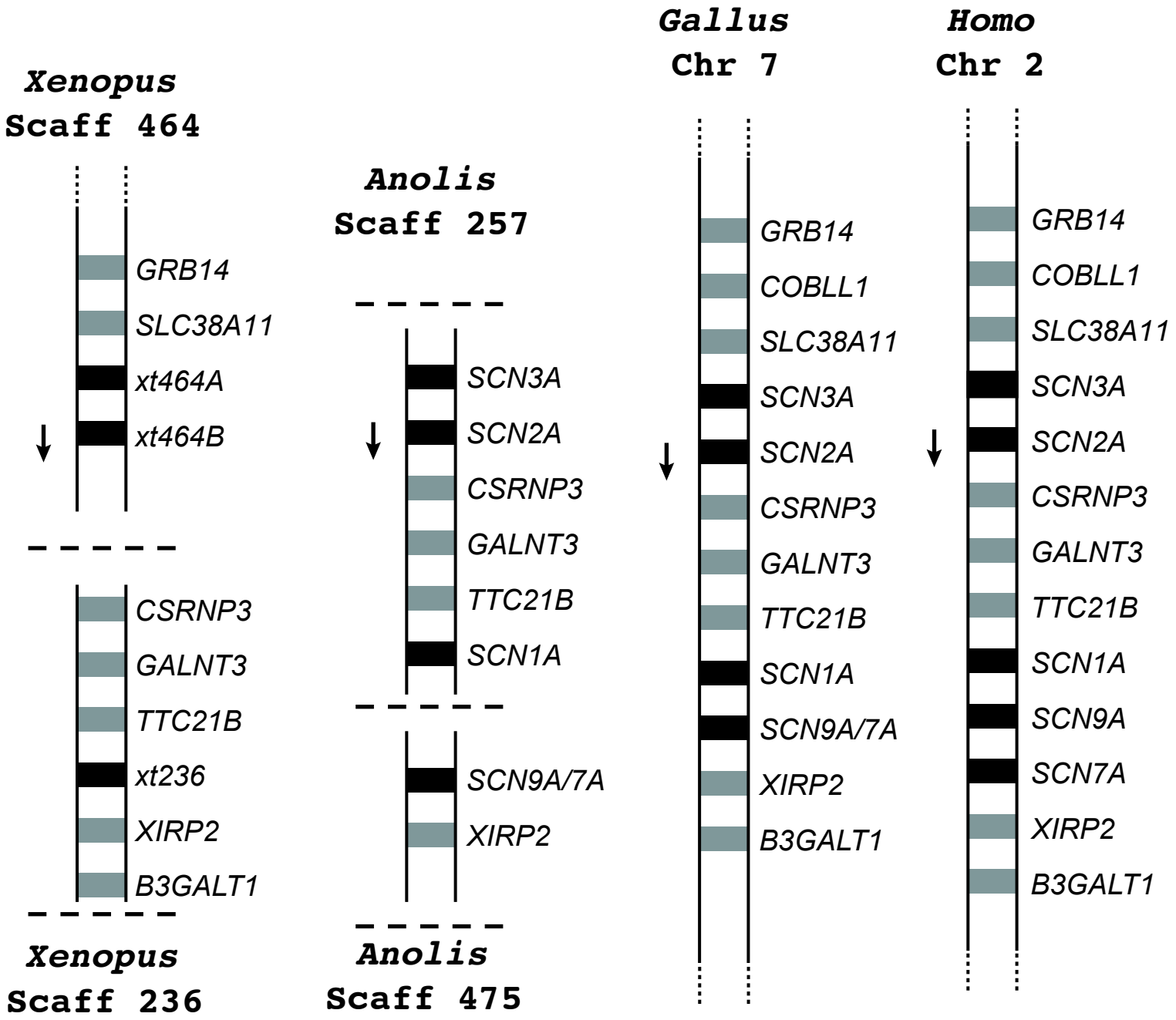
Figure 4—Representative non-Nav channel gene trees. (A) CSNRP, a gene flanking the Nav channel genes, which was studied to test for duplication and retention of neighboring non-Nav genes. (B) CACNA1, the calcium channel gene

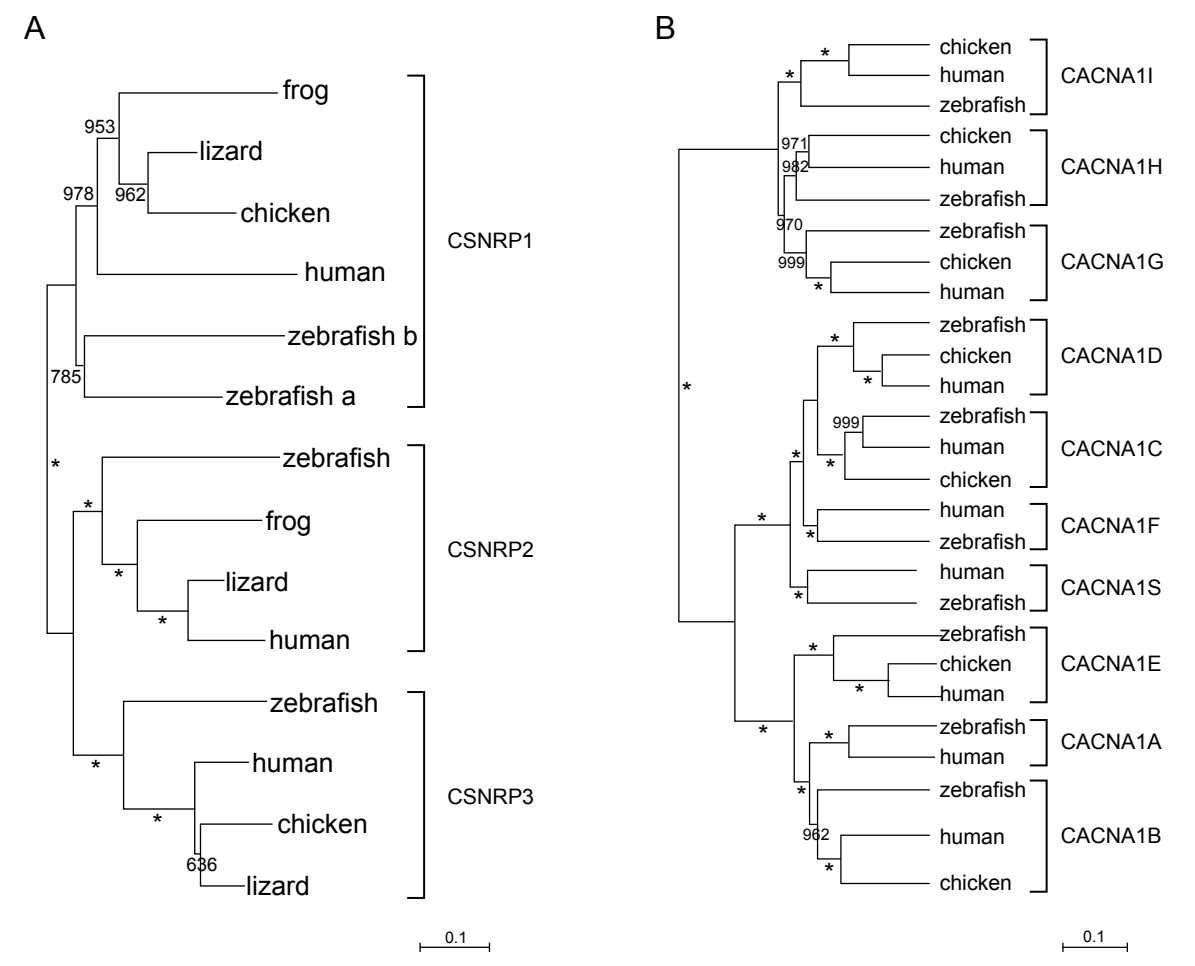
family, an example of another ion channel family. No attempt was made to systematically sample teleostean orthologs, therefore, these trees do not represent detailed phylogenies of teleostean genes. Teleost orthologs were only used to establish the number of duplications that occurred in tetrapods following the teleost-tetrapod divergence. Asterisks = bootstrap values of 100.

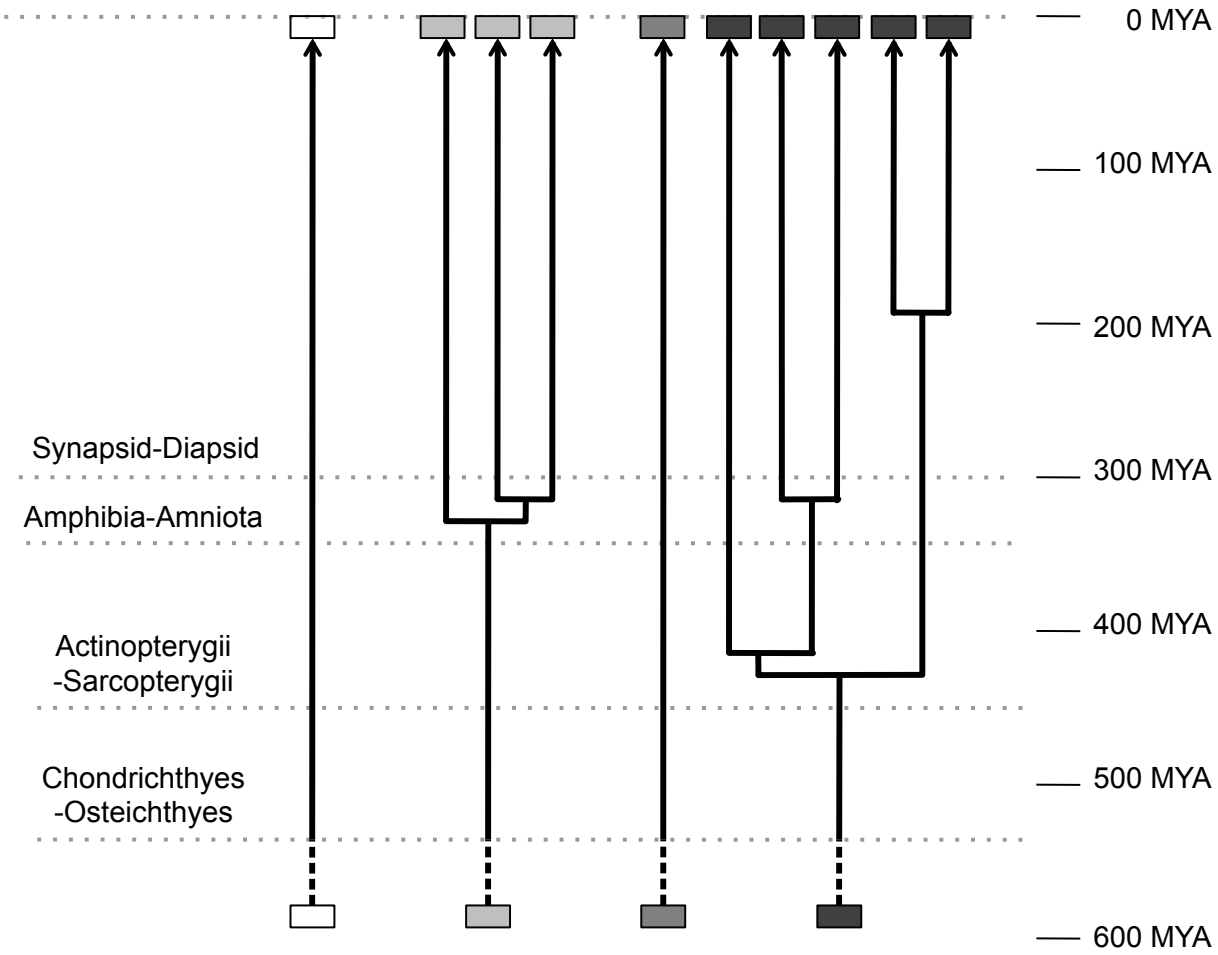
Figure 5—Schematic timeline for Nav channel gene duplications. Each set of boxes represents the lineage of four ancestral genes. Timing of the duplication of the orthologs of human chr 2 (the darkest boxes) is conservatively estimated according to the most parsimonious interpretation offered by synteny. Four Nav channel genes were present in the last common ancestor of teleosts and tetrapods (actinopterygian-sarcopterygian divergence, ~450 MYA) and likely also in the common ancestor of chondrichthyes and osteichthyes (~525MYA). The vertical dotted lines imply that these four genes resulted from the second of two vertebrate WGD events estimated at ~550 MYA. Divergence times from (Hedges 2009; Madsen 2009; Shedlock & Edwards 2009).


ratSCN1A
humanSCN1A
opossumSCN1A
platypusSCN1A
lizardSCN1A
chickenSCN1A
Nav1.1
ratSCN2A
humanSCN2A
opossumSCN2A
platypusSCN2A
lizardSCN2A
chickenSCN2A
Nav1.2
ratSCN3A
humanSCN3A
opossumSCN3A
platypusSCN3A
chickenSCN3A
lizardSCN3A
Nav1.3
ratSCN9A
humanSCN9A
opossumSCN9A
Nav1.7
humanSCN7A
ratSCN7A
Nax
platypusSCN7A/9A
lizardSCN7A/9A
chickenSCN7A/9A
xt236
xlaev
xt464b
xt464a
newt
zebrafishSCN1Laa
zebrafishSCN1Lab
humanSCN8A
ratSCN8A
lizardSCN8A
chickenSCN8A
xt67
zebrafishSCN8aa
zebrafishSCN8ab
lungfishbrain
skatebrain
Nav1.6
humanSCN11A
ratSCN11A
opossumSCN11A
platypusSCN11A
lizardSCN11A
chickenSCN11A
Nav1.9
lizardSCN10A
chickenSCN10A
humanSCN10A
ratSCN10A
opossumSCN10A
platypusSCN10A
Nav1.8
humanSCN5A
ratSCN5A
opossumSCN5A
lizardSCN5A
chickenSCN5A
Nav1.5
lungfishheart
xt28
zebrafishSCN5Laa
zebrafishSCN5Lab
sharkheart
ratSCN4A
humanSCN4A
opossumSCN4A
lizardSCN4A
chickenSCN4A
xt43
zebrafishSCN4aa
zebrafishSCN4ab
sharkmuscle
lungfishmuscle
Nav1.4
lamprey1
lamprey2

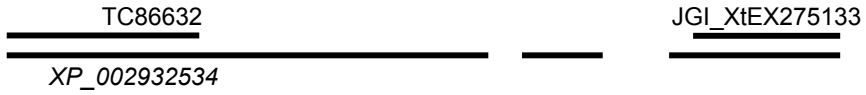

Xenopus
Scaff 28
CRYGN
NUB1
ACVR2B
ENDOGL1
xt28
WDR48
GORASP1
CSRNP1
Anolis
Scaff 64
ACVR2B
ENDOGL1
SCN5A
SCN10A
SCN11A
WDR48
GORASP1
TTC21A
CSRNP1
Gallus
Chr 2
WRD86
CRYGN
NUB1
XYLB
ACVR2B
ENDOGL1
SCN5A
SCN10A
SCN11A
WDR48
GORASP1
CSRNP1
XIRP1
Homo
Chr 3
SLC22A13
SLC22A14
XYLB
ACVR2B
ENDOGL1
SCN5A
SCN10A
SCN11A
WDR48
GORASP1
TTC21A
CSRNP1
XIRP1


Xenopus
Scaff 464
GRB14
SLC38A11
xt464A
xt464B
CSRNP3
GALNT3
TTC21B
xt236
XIRP2
B3GALT1
Xenopus
Scaff 236
Anolis
Scaff 257
SCN3A
SCN2A
CSRNP3
GALNT3
TTC21B
SCN1A
SCN9A/7A
XIRP2
Anolis
Scaff 475
Gallus
Chr 7
GRB14
COBLL1
SLC38A11
SCN3A
SCN2A
CSRNP3
GALNT3
TTC21B
SCN1A
SCN9A/7A
XIRP2
B3GALT1
Homo
Chr 2
GRB14
COBLL1
SLC38A11
SCN3A
SCN2A
CSRNP3
GALNT3
TTC21B
SCN1A
SCN9A
SCN7A
XIRP2
B3GALT1

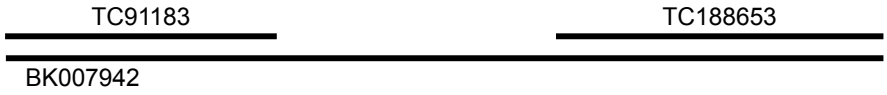

A
frog
lizard
chicken
human
zebrafish b
zebrafish a
953
962
978
785
CSNRP1
zebrafish
frog
lizard
human
CSNRP2
zebrafish
human
chicken
lizard
636
CSNRP3
0.1
B
chicken
human
zebrafish
CACNA1I
971
982
CACNA1H
970
999
CACNA1G
CACNA1D
999
CACNA1C
CACNA1F
CACNA1S
CACNA1E
CACNA1A
962
CACNA1B
0.1


0 MYA
100 MYA
200 MYA
300 MYA
400 MYA
500 MYA
600 MYA
Synapsid-Diapsid
Amphibia-Amniota
Actinopterygii
-Sarcopterygii
Chondrichthyes
-Osteichthyes

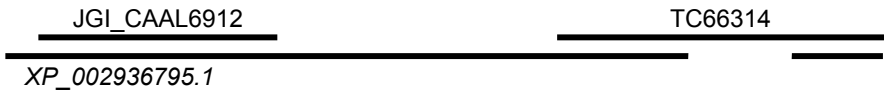
A) xt28



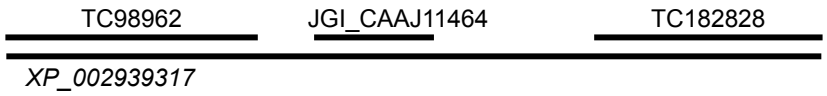
B) xt67



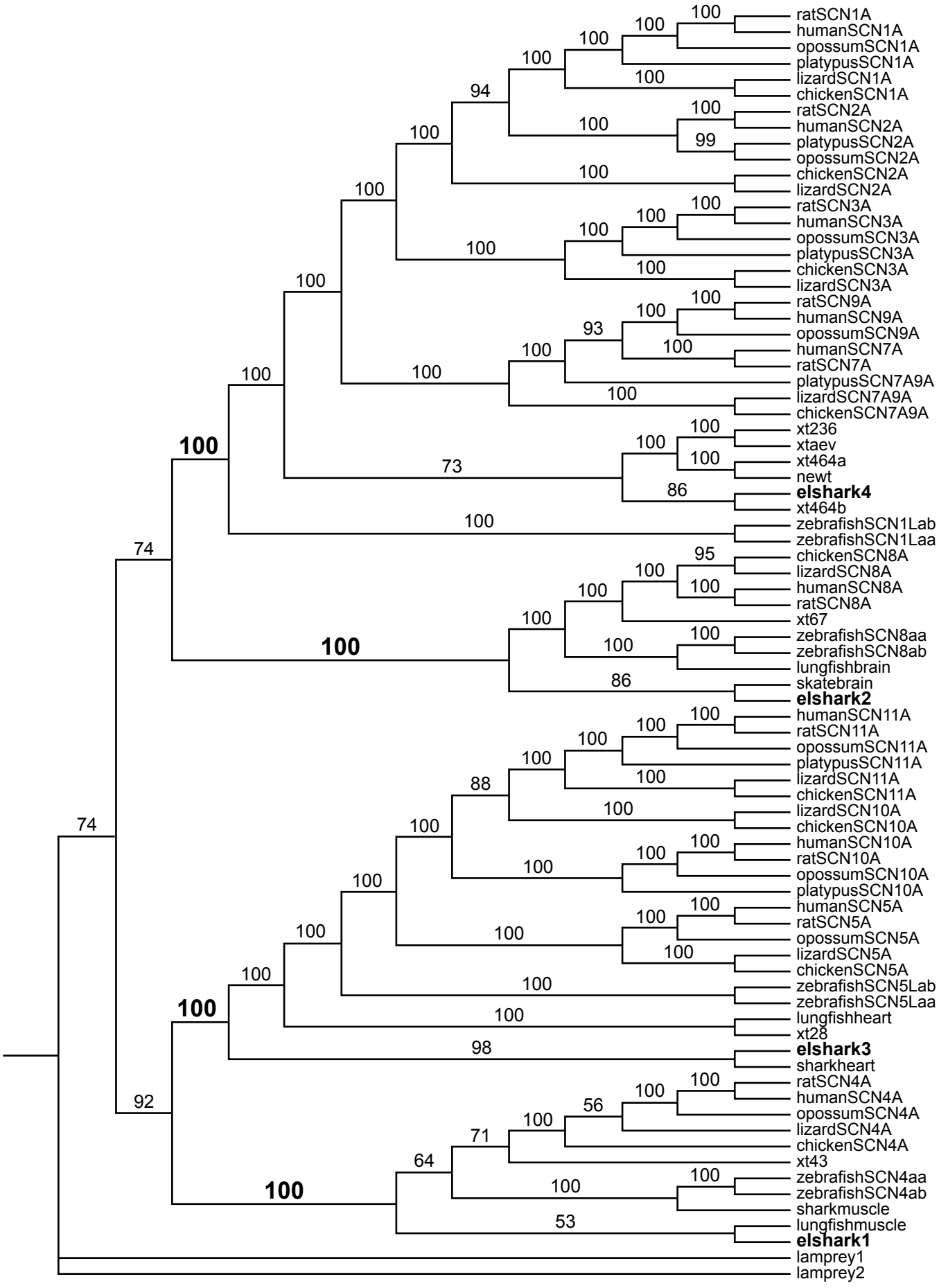
C) xt236



D) Xt464a



E) Xt464b


ratSCN1A
humanSCN1A
opossumSCN1A
platypusSCN1A
lizardSCN1A
chickenSCN1A
ratSCN2A
humanSCN2A
platypusSCN2A
opossumSCN2A
chickenSCN2A
lizardSCN2A
ratSCN3A
humanSCN3A
opossumSCN3A
platypusSCN3A
chickenSCN3A
lizardSCN3A
ratSCN9A
humanSCN9A
opossumSCN9A
humanSCN7A
ratSCN7A
platypusSCN7A9A
lizardSCN7A9A
chickenSCN7A9A
xt236
xtaev
xt464a
newt
elshark4
xt464b
zebrafishSCN1Lab
zebrafishSCN1Laa
chickenSCN8A
lizardSCN8A
humanSCN8A
ratSCN8A
xt67
zebrafishSCN8aa
zebrafishSCN8ab
lungfishbrain
skatebrain
elshark2
humanSCN11A
ratSCN11A
opossumSCN11A
platypusSCN11A
lizardSCN11A
chickenSCN11A
lizardSCN10A
chickenSCN10A
humanSCN10A
ratSCN10A
opossumSCN10A
platypusSCN10A
humanSCN5A
ratSCN5A
opossumSCN5A
lizardSCN5A
chickenSCN5A
zebrafishSCN5Lab
zebrafishSCN5Laa
lungfishheart
xt28
elshark3
sharkheart
ratSCN4A
humanSCN4A
opossumSCN4A
lizardSCN4A
chickenSCN4A
xt43
zebrafishSCN4aa
zebrafishSCN4ab
sharkmuscle
lungfishmuscle
elshark1
lamprey1
lamprey2

Supplementary Table I: Gene names and chromosomal/scaffold/contig locations, and Accession number of amino acid sequences used in the phylogenetic analysis.

Genus	Position	Location	Accession Number
Anolis (lizard)			
SCN1A:	scaff 257	2049782-1985822	BK007953
SCN2A:	scaff 257	1678493-1692523	BK007954
SCN3A:	scaff 257	1507197-1418797	BK007955
SCN4A:	scaff 246	694801-736327	BK007956
SCN5A:	scaff 64	541646-320536	BK007957
SCN7A/9A:	scaff 475	800954-867504	BK007959
SCN8A:	scaff 42	2493172-2428296	BK007958
SCN10A:	scaff 64	695331-647637	BK007960
SCN11A:	scaff 64	713407-712707	BK007961
Gallus (chicken)			
SCN1A:	chr 7	21070912-21127897	XP_001233839
SCN2A:	chr 7	21251019- 21298038	XP_001233892
SCN3A:	chr 7	21390804-21444906	BK007950
SCN4A:	chr 27	1536757-1559549	BK007949
SCN5A:	chr 2	5384093-5546821	XP_001232818
SCN7A/9A:	chr 7	20980645-21016261	XP_422021
SCN8A:	scaffE22C19	51092-90325	XP_424477
SCN10A:	chr 2	5294563-5325480	BK007952
SCN11A:	chr 2	5235482-5272011	BK007951
Xenopus tropicalis (western clawed frog)			
xt28	scaff 28	681183-84470	XP_002932534
xt236	scaff 236	243937-311547	XP_002936795
xt464a	scaff 464	152975-190960	XP_002939317
xt464b	scaff 464	68762-69015	XP_002939316
xt67	scaff 67	2483289-2535145	BK007942
xt43	scaff 43	635601-675244	XP_002933087
Monodelphis (opossum)			
SCN1A	chr 4	4526966-1746559824	XP_001367386
SCN2A	chr 4	173638959-173744377	XP_001367245
SCN3A	chr 4	173511981-173401133	XP_001367154
SCN4A	chr 2	211255665-211188530	BK007943
SCN5A	chr 6	288151353-288257274	BK007944
SCN9A	chr 4	174856021-174960654	XP_001367438
SCN10A	chr 6	288328228-288425463	XP_001373143
SCN11A	chr 6	288504210-288640631	BK007945
Ornithorhynchus (platypus)			
SCN1A	contig 514	8978877-9047893	BK007946
SCN2A	contig 514	8565874-8635550	XP_001512785
SCN3A	contig 514	8423554-8506526	XP_001513457
SCN7A/9A	contig 514	9134386-9198701	XP_001513667
SCN10A	contig 6313	5821-36850	BK007947
SCN11A	contig 1661	44794-6148	BK007948
Homo (human)			
SCN1A:	chr 2		BAC45228

SCN2A:	chr 2	AAA18895
SCN3A:	chr 2	Q9NY46
SCN4A:	chr 17	AAO83647
SCN5A:	chr 3	BAD12084
SCN7A:	chr 2	AAA59899
SCN8A:	chr 12	AAF35390
SCN9A:	chr 2	Q15858
SCN10A:	chr 3	NP_006505
SCN11A:	chr 3	NP_054858

Rattus (rat)

SCN1A:	chr 3	NP_110502
SCN2A:	chr 3	NP_036779
SCN3A:	chr 3	NP_037251
SCN4A:	chr 10	NP_037310
SCN5A:	chr 8	NP_037257
SCN7A:	chr 3	NP_113874
SCN8A:	chr 7	AAC26015
SCN9A:	chr 3	NP_579823
SCN10A:	chr 8	Q62968
SCN11A:	chr 8	EDL76900

Lepidosiren (Lungfish)

lungfishbrain	HQ289894
lungfishheart	HQ289893
lungfishmuscle	HQ289895

Heterodontus (Horn Shark)

Sharkmuscle	HQ434339
Sharkheart	HQ434340

Dasyatis (Skate)

skatebrain	HQ434341

Petromyzon (Lamprey)

lamprey1	ABB84815
lamprey2	ABB84816

Cynops (Newt)

Newt	AAD17315

Xenopus laevis (African clawed frog)

Xlaev	AAM83131

Danio rerio (zebrafish)

SCN1Laa	ABA54918
SCNlLab	ABA54919
SCN4aa	ABA54921
SCN4ab	ABA54920
SCN5Laa	ABA54922
SCN5Lb	ABA54923
SCN8aa	NP_571703
SCN8ab	ABA54924

Callorhinchus milii (elephant shark)

elshark1	AAVX01073135.1
elshark2	AAVX01077205.1
elshark3	AAVX01425063.1
elshark4	AAVX01490561.1

Table II: Other Sequences used in this study.

A) Genes that flank the Nav channel genes

Activin A receptor (ACVR)
Homo: NP_001097.2, NP_001607; **Gallus:** NP_990698, NP_989648; **Danio:** NP_001103748, NP_571285; **Xenopus:** NP_001081479, NP_001135613.1.

β 1,3-galactosyltransferase (B3GALT)
Homo: NP_066191, NP_003774; **Gallus:** XP_426584, XP_001231894; **Danio:** XP_699646, NP_996984

Cordone bleu protein like 1 (COBLL)
Homo: NP_055715.3; **Gallus:** XP_422028; **Danio:** XP_002663463

Cysteine-serine-rich nuclear protein (CSRNP)
Xenopus: NM_001078993, 187608188; **Gallus:** XP_418530.2, XM_001235288; **Homo:** Q96S65, NP_110436, NP_079245.2; **Danio:** XP_688758.2, XP_001343412, NP_955913.1; **Anolis:** ENSACAP00000012615, ENSACAP00000009555, ENSACAP00000012185.

Endonuclease G like protein (ENDOGL)
Xenopus: NP_001017202; **Gallus:** XP_418536; **Homo:** NP_001138936; **Danio:** NP_001019385

Gamma N crystalin (CRYGN)
Homo: NP_653328; **Gallus:** XP_425967.2; **Xenopus:** AAI61267; **Danio:** NP_00100778, NP_001003428.

Golgi reassembly-stacking protein (GORASP)
Homo: Q9BQQ3.3; **Gallus:** NP_001026134; **Xenopus:** CAJ83122; **Danio:** NP_001007412.1

Growth factor receptor-bound protein 14 (GRB14)
Homo: AAC15861, BAF76353; **Gallus:** XP_422029.2, NP_001034371; **Danio:** TC437516, XP_690519.3

N-acetylgalactosaminyltransferase (GALNT)
Homo: NP_004473.2, NP_004472, NP_065207.2; **Gallus:** XP_422023.2, XP_419581.2, NP_001006381; **Xenopus:** AAI10707, NP_001083410; **Tetraodon:** CAG09349.1; **Danio:** XP_698799.3, XP_687472.2.

NEDD8 Ultimate buster (NUB)
Homo: AAK21001; **Salmo:** ACN11249; **Gallus:** XP_418538.2; **Danio:** NP_001107052; **Xenopus:** NP_001017076.2; **Platypus:** XP_001513245.

Solute carrier family 22, members 13 and 14 (SLC2213&14)
Homo: NP_004794.2, NP_004247.2; **Mus:** NP_001032838; **Gallus:** XP_418529; **Xenopus:** TC206214; **Danio:** XP_001346178.

Tetratricopeptide repeat domain 21 (TTC21)
Homo: NP_001098983.2, AAH63579; **Gallus:** XP_422022.2; **Xenopus:** TC170322, TC20670; **Danio:** TC395657, TC384177.

WD repeat domain 48 (WDR48)
Xenopus: NP_001072858; **Gallus:** NP_001026135, XP_001233535; **Homo:** NP_065890, NP_938026.2; **Danio:** NP_999874

Xin actin-binding repeat containing (XIRP)
Homo: NP_001185550; NP_001073278; **Gallus:** NP_989679.3; **Xenopus: TC216350; Danio:** NP_001012377; XP_688205.3

Xylulokinase (XYLB)
Homo: NP_005099.2; **Gallus:** XP_418537; **Danio:** NP_956673; **Xenopus:** NP_001108300.

B) Voltage-dependent Ion Channels

Ether a-go-go related K^+ channel (ERG)
Homo: NP_653234.2, NP_150375.2, NP_110406, NP_647479.2, NP_036417, NP_036416, NP_002229, NP_000229; **Gallus:** XP_418747.2, XP_422030.2, XP_418075.2, XP_421414.2, XP_001235280, XP_419440.2; **Danio:** XP_001922595, XP_688778.3, NP_998002, NP_001038263, XP_001920653, XP_001919436.1, NP_001038396, XP_001918581.

Calcium channel (Cav)
Homo: NP_001120694, NP_000709, NP_955630.2, NP_000711, NP_000712.2, NP_005174.2, NP_938199, NP_001005407, NP_001003406, NP_000060.2; **Gallus:** TC227671, NP_989624, XP_416388.2, NP_990365, XP_422255.2, XP_001232654, XP_414830.2, XP_425474.2; **Danio:** XP_690548.3; NP_001108020; NP_571975; NP_982351; XP_001920777; XP_688452.2; XP_001920550, XP_001919609, XP_694715.3, XP_699282.3.

Hyperpolarization-activateded cyclic nucleotide-gated potassium channel (HCN)
Homo: NP_066550.2, NP_001185.3, NP_065948, NP_005468; **Monodelphis:** XP_001363953, XP_001366855; **Taeniopygia:** XP_002194435; **Xenopus:** XP_002933077, TC182733; **Oncorhynchus:** NP_001117790; **Tetraodon:** CAF97159, CAG05571; **Danio:** XP_685414, CAP09378, XM_002193409; **Gallus:** XP_425050.2, chickenmdv004_b03 CF253024.

Slo K channel (BK)
Homo: NP_001014797, NP_940905.2, NP_001027006.2; **Gallus:** NP_989555, XP_426614.2, ADD16620; **Danio:** NP_001139072, XP_694050.4; **Mus:** NP_032458.3; **Monodelphis:** XP_001381830.