Parallel Evolution of Tetrodotoxin Resistance in Three Voltage-Gated Sodium Channel Genes in the Garter Snake Thamnophis sirtalis

Joel W. McGlothlin,*1,2 John P. Chuckalovcak,2,3 Daniel E. Janes,4,5 Scott V. Edwards,4 Chris R. Feldman,6 Edmund D. Brodie Jr,7 Michael E. Pfrender,8 and Edmund D. Brodie III2,9

1Department of Biological Sciences, Virginia Tech, Blacksburg, VA
2Department of Biology, University of Virginia
3Bio-Rad Laboratories, Hercules, CA
4Department of Organismic and Evolutionary Biology, Harvard University
5Division of Genetics and Developmental Biology, National Institutes of Health, Bethesda, MD
6Department of Biology, University of Nevada, Reno
7Department of Biology, Utah State University
8Department of Biological Sciences and Environmental Change Initiative, University of Notre Dame
9Mountain Lake Biological Station, University of Virginia

*Corresponding author: E-mail: joelmcg@vt.edu.

Abstract

Members of a gene family expressed in a single species often experience common selection pressures. Consequently, the molecular basis of complex adaptations may be expected to involve parallel evolutionary changes in multiple paralogs. Here, we use bacterial artificial chromosome library scans to investigate the evolution of the voltage-gated sodium channel (Nav) family in the garter snake Thamnophis sirtalis, a predator of highly toxic Taricha newts. Newts possess tetrodotoxin (TTX), which blocks Nav’s, arresting action potentials in nerves and muscle. Some Thamnophis populations have evolved resistance to extremely high levels of TTX. Previous work has identified amino acid sites in the skeletal muscle sodium channel Nav1.4 that confer resistance to TTX and vary across populations. We identify parallel evolution of TTX resistance in two additional Nav paralogs, Nav1.6 and Nav1.7, which are known to be expressed in the peripheral nervous system and should thus be exposed to ingested TTX. Each paralog contains at least one TTX-resistant substitution identical to a substitution previously identified in Nav1.4. These sites are fixed across populations, suggesting that the resistant peripheral nerves antedate resistant muscle. In contrast, three sodium channels expressed solely in the central nervous system (Nav1.1–1.3) showed no evidence of TTX resistance, consistent with protection from toxins by the blood–brain barrier. We also report the exon–intron structure of six Nav paralogs, the first such analysis for snake genes.

Our results demonstrate that the molecular basis of adaptation may be both repeatable across members of a gene family and predictable based on functional considerations.

Key words: adaptation, coevolution, gene families, molecular evolution, predator–prey interactions, toxins.

Introduction

One of the primary goals of evolutionary biology is to understand the genetic basis of adaptive phenotypic evolution. In recent years, great progress has been made toward this goal, particularly for traits that have a relatively simple genetic architecture (Rosenblum et al. 2004; Hoekstra et al. 2006; Hoekstra and Coyne 2007; Steiner et al. 2007; Stern and Orgogozo 2008), but dissecting the genetic basis of adaptation for complex polygenic phenotypes has proved somewhat elusive (Rockman 2012; Travisano and Shaw 2013). Multimember gene families, which consist of paralogs created by segmental or whole-genome duplication, are abundant across the diversity of life (Thornton and DeSalle 2000) and provide a potential middle ground between studies of single-gene and polygenic evolution. Paralogs within gene families may have tissue-specific expression but often retain sufficiently similar functional roles. Consequently, their evolutionary response to common selection pressures may be somewhat predictable (Losi et al. 2008). If the functions of paralogs are well-characterized, then candidate genes underlying adaptation may be readily identified and investigated, providing empirical tractability similar to that of studies of single-gene evolution. At the same time, the evolution of multiple paralogs may be necessary for adaptations at the level of the whole organism. As such, studies of gene families may provide simplified models of polygenic adaptation.
Voltage-gated sodium channels (Na\textsubscript{v}) are an excellent model system for studying both the evolution of gene families and the genetics underlying phenotypic adaptation (Liebeskind et al. 2011; Zakon 2012). These channels are formed by an α subunit consisting of four domains (D1–D4), which together form a sodium-selective pore in the plasma membranes of excitable tissue such as nerves and muscle, and an optional β subunit that may modify its activity (Yu and Catterall 2003; Yu et al. 2005). Although invertebrate species typically have a single α subunit, vertebrates have at least four (Widmark et al. 2011; Zakon et al. 2011); these α subunits (Na\textsubscript{v}1) are encoded by genes in the SCNA family (Catterall et al. 2005). The original expansion of the SCNA family appears to derive from the two genome duplications early in vertebrate evolution (Dehal and Boore 2005; Blomme et al. 2006), but the family has expanded further in some lineages (Widmark et al. 2011; Zakon et al. 2011). In teleost fishes, a third whole-genome duplication has expanded the family to eight members, whereas in tetrapods, duplication along chromosomes has increased the family size to six (amphibians), nine (reptiles including birds), or ten (mammals) (Zakon et al. 2011).

Na\textsubscript{v} channels are the targets of numerous neurotoxins used as either predator venoms or prey defenses (Daly 1995; Cestele and Catterall 2000; Fry et al. 2009). Both toxic species and their predators are expected to be under strong selection for resistance to these neurotoxins, making Na\textsubscript{v} a model system for the predators are expected to be under strong selection for resistance to these neurotoxins, making Na\textsubscript{v} a model system for the predators–prey interactions (Yoshida 1994; Geffeney et al. 2005; Jost et al. 2008; Rowe et al. 2013). One of the best characterized neurotoxins is tetrodotoxin (TTX), which acts by binding to the Na\textsubscript{v} outer pore, which is formed by the “P-loop” region of each domain, and blocking the conductance of sodium ions across the membrane (Hanifin 2010; Moczydlowski 2013). Pufferfishes (Tetradontidae), which employ TTX as an antipredator defense, have evolved resistance to the toxin in each of their eight Na\textsubscript{v} paralogs through a variety of amino acid substitutions that are often shared across paralogs (Jost et al. 2008). Resistive sodium channels have also evolved in predator species that consume tetrodotoxic prey. For example, some populations of the garter snake Thamnophis sirtalis have evolved resistant skeletal muscle sodium channels (Na\textsubscript{v}1.4) as a consequence of the ongoing coevolutionary arms race with Pacific newts (Taricha spp.) (Brodie and Brodie 1990; Brodie et al. 2002; Geffeney et al. 2005; Hanifin et al. 2008). Similar arms races have independently led to the convergent evolution of resistance in Na\textsubscript{v}1.4 in a number of other snake species that also consume tetrodotoxic prey (Feldman et al. 2009, 2012).

Although the evolution of TTX resistance in Na\textsubscript{v}1.4 has been well-characterized, little is known about how the rest of the Na\textsubscript{v} family has evolved in other predators. In this study, we used a bacterial artificial chromosome (BAC) library and 454 sequencing to obtain the sequences of six SCNA genes from an individual Th. sirtalis collected from a TTX-resistant population. Three of these channels, Na\textsubscript{v}1.1, Na\textsubscript{v}1.2, and Na\textsubscript{v}1.3 (encoded by the genes SCN1A, 2A, and 3A, respectively), are found primarily in the central nervous system (CNS) in mammals (Goldin 2001, 2002; Catterall et al. 2005). As TTX does not appear to cross the blood–brain barrier (Zimmer 2010), these channels should be sheltered from TTX, even when predatory snakes consume toxic prey. Thus, these channels should not be under selective pressure to evolve TTX resistance, and we do not expect to detect TTX resistance-conferring replacements in these loci. Three others, Na\textsubscript{v}1.4, Na\textsubscript{v}1.6, and Na\textsubscript{v}1.7 (encoded by the genes SCN4A, 8A, and 9A), are found additionally or exclusively in the muscles or peripheral nervous system (PNS) (Caldwell et al. 2000; Thornton and DeSalle 2000; Goldin 2001, 2002; Trimmer and Rhodes 2004; Dib-Hajj et al. 2013). These tissues will be exposed to TTX when snakes consume tetrodotoxic prey, and should be under selective pressure to evolve TTX resistance. We predict that these channels will contain amino acid substitutions necessary to support TTX resistance at the organismal level and that, as in pufferfishes, TTX resistance should evolve in parallel across paralogs; that is, paralogs are expected to evolve resistance independently through similar or identical amino acid substitutions. Because our methods generated high-quality genomic sequence for most paralogs, we also report their exon–intron structure, the first such analysis for snake genes.

**Results**

**Gene Structure**

Sequences were obtained from a BAC library constructed using an individual Th. sirtalis from Benton County, OR, a population with medium-high resistance to TTX that coexists with the most highly tetrodotoxic population of Taricha newts known (Brodie et al. 2002; Hanifin et al. 2008). Complete or nearly complete genomic sequences were obtained for three paralogs: SCN4A (at least 65.1 kb from start codon to stop codon), SCN8A (73.8 kb), and SCN9A (56.8 kb) (fig. 1). Despite the recent publication of two snake genomes (Castoe et al. 2013; Vonk et al. 2013), to our knowledge these sodium channel genes represent the first snake genes to be presented and analyzed at the genomic sequence level. Full coding sequences (CDSs) were obtained for SCN3A (6,009 bp, 2,003 amino acids), 4A (5,625 bp, 1,875 amino acids), 8A (5,961 bp, 1,987 amino acids), and 9A (5,892 bp, 1,964 amino acids) (supplementary table S1, Supplementary Material online). The CDS obtained for SCN4A was identical to the previously published sequence of the same gene from Benton County, OR, except for a single nucleotide substitution (A3427C) that led to an amino acid difference (I1143L) (Geffeney et al. 2005). Other published SCN4A sequences of Th. sirtalis from other populations have C3427/L1143 as well (Feldman et al. 2010). The CDS lengths of these four genes were comparable to their orthologs in Rattus and Anolis. Partial CDSs were obtained for SCN1A (5,542 bp, 1,847 amino acids) and 2A (5,747 bp, 1,915 amino acids). Alignment of the SCN1A and 2A CDS with Rattus and Anolis orthologs suggested that 483–486 bp (161–162 amino acids) and 279 bp (93 amino acids), respectively, were missing from the 3′-ends of the snake version of these genes. The exon–intron structure of these six genes was similar to that found in other species (fig. 1). SCN8A and 9A showed the
26-exon structure typical for this gene family (Widmark et al. 2011). The regions of SCN1A, 2A, and 3A for which we obtained genomic sequence also show typical exon–intron structure. As in most nonmammalian vertebrates, SCN4A has 25 exons, with two exons in place of the exons numbered 10–12 in other paralogs (Widmark et al. 2011). We have named these exons, which encode the intracellular loop linking DI and DII, 10a and 10b. Other SCN4A exons are numbered 1–9 and 13–26 by their homology to other paralogs following Widmark et al. (2011).

All introns with complete sequence displayed standard GT–AG ends, with exceptions in two locations. All paralogs had an AT–AC intron between exons 2 and 3 except SCN4A, which had an AT–AA intron in this location. There was also an AT–AC intron between exons 23 and 24 in SCN4A, 8A, and 9A. (We obtained only cDNA sequence for this region in the other paralogs; see Materials and Methods.) Both AT–AC introns are found in vertebrate orthologs and are thought to have an ancient origin (Widmark et al. 2011). The AT–AA intron in SCN4A has not been previously noted, although it is also present in Anolis. The lengths of completely sequenced introns ranged from 89 to 11,139 bp (median = 1,391 bp, mean = 1,837 bp). Although these lengths are typical for vertebrate genes (Hong et al. 2006), extremely long introns may have overlapped contig boundaries, biasing the mean downward slightly. The intron between exons 16 and 17 appears to be quite long in at least four paralogs (SCN1A, 3A, 4A, and 9A). Although we did not detect exon 17 in SCN1A, contig 88 suggests that this intron may be at least 23 kb in length (fig. 1).

Putative splice variation was also present. Except for SCN4A, all paralogs had two copies of exon 5 separated by a short (95–147 bp) intron (fig. 1). These copies, which we call 5a and 5b, differed by 1–2 amino acids depending on the paralog. The amino acid sequence of 5b is identical across all paralogs. This exon encodes the extracellular loop between transmembrane segments S3 and S4 in domain I, and the duplication of the exon allows for alternative splicing (Copley 2004). In one paralog (SCN5A), this splice variation has been shown to lead to alternative neonatal and adult channels with different physiological properties (Onkal et al. 2008). Critically, this region does not seem to affect TTX sensitivity (Onkal et al. 2008).

In addition, three paralogs (SCN1A, 8A, and 9A) have a 33-bp segment that may or may not be spliced from the end of exon 11 (fig. 1). Both types of splice variation appear to influence interaction with the β subunit (Farmer et al. 2012).

The GC content for all paralogs was approximately 40% (fig. 1), which is close to the genome-wide mean base

![Figure 1](https://example.com/fig1.png)

**Fig. 1.** Gene structure for the six SCNA paralogs detected using BAC library scanning. To facilitate comparison, genes are aligned by the start codon in exon 1. Total base pair length represents the concatenated length of the contigs shown and may include untranscribed regions. (A) Concatenated contigs from 454 sequencing. (* indicates that contigs share an edge; † indicates that two contigs share an edge with a third small contig.) (B) Exon/intron structure. Partial CDS was obtained for three paralogs. (**) indicates an incomplete exon. (C) GC content using a 500-bp sliding window. (D) CpG islands (green) predicted by Geneious. (E) Simple repeats or TEs identified by Censor (green = satellite; red = DNA transposon; pink = endogenous retrovirus; orange = LTR retrotransposon; blue = non-LTR retrotransposon).
composition for several reptiles (Shedlock et al. 2007; Alföldi et al. 2011). However, all paralogs contained at least one CpG island, usually within introns (fig. 1). At least four paralogs (SCN1A, 3A, 4A, and 9A) had a CpG island in the long intron between exons 16 and 17. CpG islands were also detected upstream of exon 1 in two paralogs (SCN4A and 8A), suggesting a possible role in gene regulation (Deaton and Bird 2011).

Transposable elements (TEs) were common in introns, making up 8–19% of noncoding regions (fig. 1, table 1). The TE profile across paralogs was comparable to that of published reptile genomes (Shedlock 2006; Alföldi et al. 2011; Castoe et al. 2013). Paralogs differed substantially in TE content (table 1). SCN4A had the highest TE content, which was primarily due to an abundance of non-LTR (long terminal repeat) retrotransposons, particularly CR1, L2, and R4, the last of which was absent from other paralogs (table 1). DNA transposons, particularly hAT and Mariner/Tc1, were most common in SCN9A, which showed the second highest TE content (table 1).

<table>
<thead>
<tr>
<th>Repeat Class</th>
<th>SCN1A</th>
<th>SCN2A</th>
<th>SCN3A</th>
<th>SCN4A</th>
<th>SCN8A</th>
<th>SCN9A</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. Length</td>
<td>No. Length</td>
<td>No. Length</td>
<td>No. Length</td>
<td>No. Length</td>
<td>No. Length</td>
<td>No. Length</td>
</tr>
<tr>
<td>Simple repeat</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>33</td>
</tr>
<tr>
<td>Satellite</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SAT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TE</td>
<td>30</td>
<td>5,199</td>
<td>26</td>
<td>5,770</td>
<td>35</td>
<td>4,979</td>
</tr>
<tr>
<td>DNA transposon</td>
<td>11</td>
<td>2,277</td>
<td>8</td>
<td>948</td>
<td>11</td>
<td>986</td>
</tr>
<tr>
<td>Dada</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>189</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Harbinger</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mariner/Tc1</td>
<td>5</td>
<td>1,811</td>
<td>1</td>
<td>73</td>
<td>5</td>
<td>473</td>
</tr>
<tr>
<td>Polinton</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>hAT</td>
<td>6</td>
<td>466</td>
<td>6</td>
<td>686</td>
<td>4</td>
<td>416</td>
</tr>
<tr>
<td>Endogenous retrovirus</td>
<td>2</td>
<td>181</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>80</td>
</tr>
<tr>
<td>ERV1</td>
<td>2</td>
<td>181</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>80</td>
</tr>
<tr>
<td>ERV2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ERV3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>80</td>
</tr>
<tr>
<td>LTR retrotransposon</td>
<td>1</td>
<td>93</td>
<td>1</td>
<td>38</td>
<td>3</td>
<td>571</td>
</tr>
<tr>
<td>DIRS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>38</td>
<td>1</td>
<td>102</td>
</tr>
<tr>
<td>Gypsy</td>
<td>1</td>
<td>93</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>469</td>
</tr>
<tr>
<td>Non-LTR retrotransposon</td>
<td>16</td>
<td>2,648</td>
<td>17</td>
<td>4,784</td>
<td>19</td>
<td>3,304</td>
</tr>
<tr>
<td>CR1</td>
<td>3</td>
<td>598</td>
<td>5</td>
<td>3,666</td>
<td>2</td>
<td>484</td>
</tr>
<tr>
<td>L1</td>
<td>3</td>
<td>282</td>
<td>2</td>
<td>99</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>L2</td>
<td>2</td>
<td>353</td>
<td>4</td>
<td>403</td>
<td>3</td>
<td>197</td>
</tr>
<tr>
<td>Penelope</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>208</td>
<td>4</td>
<td>808</td>
</tr>
<tr>
<td>R4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>693</td>
</tr>
<tr>
<td>RTE</td>
<td>3</td>
<td>557</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>RTEX</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Rex1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>SINE</td>
<td>5</td>
<td>858</td>
<td>5</td>
<td>408</td>
<td>7</td>
<td>1,122</td>
</tr>
<tr>
<td>SINE2/tRNA</td>
<td>4</td>
<td>549</td>
<td>4</td>
<td>327</td>
<td>4</td>
<td>632</td>
</tr>
<tr>
<td>SINE3/SS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>80</td>
</tr>
<tr>
<td>Tx1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Vingi</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>5,199</td>
<td>26</td>
<td>5,770</td>
<td>34</td>
<td>4,979</td>
</tr>
<tr>
<td>% of total sequence</td>
<td>9.8</td>
<td>14.2</td>
<td>8.4</td>
<td>19.5</td>
<td>8.0</td>
<td>15.4</td>
</tr>
</tbody>
</table>

Putative TTX Resistance and Interpopulation Variation

Sodium channels expressed predominantly in the CNS (Na1.1–1.3) showed no evidence of TTX resistance evolution (fig. 2). Across all three channels, the amino acid sequence of the four P-loops was identical to the sequence from either rat or lizard. In contrast, all three channels expressed in the periphery (Na1.4, 1.6, and 1.7) showed evidence of TTX resistance. There was no significant evidence of gene conversion between resistant paralogs, suggesting that TTX resistance initially arose through independent mutational events. However, gene conversion does appear to occur among SCN1–3/9A, which are found on the same chromosome in other amniotes (Zakon et al. 2011). Across these four paralogs, 13 CDS fragments found in seven distinct locations and averaging 120 bp in length showed evidence of gene conversion. Only one of these gene conversion events occurred in a P-loop region, the DI P-loop of SCN2/3A.
As previously described, the skeletal muscle channel Nav1.4 of Benton County *Th. sirtalis* contains two substitutions in DIV: an Ile–Val (I1561V) and a Gly–Asp (G1566A). Together, these substitutions decrease the binding affinity of TTX approximately 10-fold, providing moderate TTX resistance (Geffeney et al. 2002, 2005) (fig. 2).

PNS channels Nav1.6 and 1.7 showed evidence of parallel evolution of TTX resistance. In Nav1.6, DIV contained an Ile–Val substitution (I1709V) identical to the one seen in Nav1.4 (fig. 2). On its own, this substitution is expected to confer 2-fold resistance to TTX (Geffeney et al. 2005). In garter snakes from Warrenton, OR, Nav1.4 displays this Ile–Val and no other TTX-resistance substitutions, providing low to moderate levels of resistance to skeletal muscle (Geffeney et al. 2002, 2005). This finding suggests that PNS axons in *Th. sirtalis* should also display low to moderate TTX resistance.

Nav1.7 contained four substitutions that are likely to confer resistance to TTX: An Asp–Glu (D1393E) in DIII and an Ala–Gly (A1681G), an Asp–Asn (D1684N), and a Gly–Tyr (G1685Y) in DIV (fig. 2). The Asp–Glu substitution has been previously described in Nav1.4 from two snake species, the resistant *Th. atratus* (Feldman et al. 2009) and the putatively resistant *Amphiesma pryeri* (Feldman et al. 2012), and replacements at this site lead to minor changes in TTX-binding affinity (Terlau et al. 1991; Choudhary et al. 2003), but this specific replacement has not been empirically tested for its effects on TTX binding. The Ala–Gly substitution occurs at the channel’s selectivity filter, and is also found in four different paralogs in TTX-bearing pufferfishes, where it provides 1.5-fold resistance to TTX and 11-fold resistance to saxitoxin (Jost et al. 2008). Despite its location, this substitution does not affect ion selectivity, but delays recovery from slow inactivation (Wu et al. 2013), suggesting the possibility that it evolved to compensate for other changes to the channel’s pore. The Asp–Asn substitution occurs in a highly TTX-resistant (600X) isoform of Nav1.4 found in California *Th. sirtalis* (Geffeney et al. 2005) and on its own should confer approximately 30- to 40-fold resistance to TTX (Penzotti et al. 1998; Choudhary et al. 2003). The Gly–Tyr substitution is previously undescribed and untested, but substitutions in this location are often found in TTX-resistant species (Geffeney et al. 2005; Jost et al. 2008; Feldman et al. 2012). This residue is located in the outer portion of the pore, and the replacement of glycine’s hydrogen with tyrosine’s large aromatic side chain may conceivably interfere with the ability of TTX to access its binding site in the pore. Taken together, these substitutions suggest that Nav1.7 in *Th. sirtalis* should be highly resistant to TTX, protecting the function of the various tissues in which this channel is expressed (Dib-Hajj et al. 2013).

**Fig. 2.** (A) Amino acid substitutions in the pore regions (or “P-loops”) for six Na+, in *Thamnophis sirtalis*. All sequences are shown relative to rat Na+, with residues involved in TTX binding underlined. Substitutions putatively conferring TTX resistance are shown in red. (B) Locations of putative resistance-conferring substitutions in Na1.4, 1.6, and 1.7.
Amino acid sequences of the P-loops from Na\textsubscript{v}1.6 DIV, and Na\textsubscript{v}1.7 DIII and DIV obtained from 16 individual snakes from two additional populations were uniformly identical to those obtained from BAC library sequencing. These two populations differ markedly in the levels of TTX exposure individuals may experience. At Willow Creek, CA, Th. sirtalis coexists with Taricha newts of moderate toxicity (Hanifin et al. 2008) and snake resistance phenotypes are a mixture of high resistance (64%) and low resistance (36%) (Feldman et al. 2009, 2010), whereas at Mountain Lake, VA, Notophthalmus viridescens newts are mildly toxic (Yotsu-Yamashita and Mebs 2001) and snakes have low-resistance phenotypes (Feldman et al. 2009). Based on these P-loop sequences, all snakes were predicted to have resistant peripheral nerves regardless of their geographical origin (eastern or western North America) or their level of organismal resistance to TTX, suggesting that resistant peripheral nerves may be fixed in Th. sirtalis. In contrast, various alleles of Na\textsubscript{v}1.4 associated with distinct resistance levels are found within and among populations (Geffeney et al. 2005; Feldman et al. 2009, 2010).

Incidentally, Anolis carolinensis Na\textsubscript{v}1,7 also shows evidence of TTX resistance. Differences putatively conferring TTX resistance occur in two of the same sites as in Th. sirtalis, but with different amino acid substitutions. The Asp–Ala substitution in A. carolinensis DIV has previously been shown to provide approximately 150-fold resistance to TTX (Choudhary et al. 2003), and the Gly–Glu substitution adds a negative charge in the adjacent site, a feature often seen in TTX-resistant channels (Feldman et al. 2012).

**Discussion**

Our results indicate that the evolution of organismal resistance to ingested TTX in Th. sirtalis involves substitutions in at least three genes, SCN4A, SCN8A, and SCN9A, which encode three Na\textsubscript{v} paralogs (Na\textsubscript{v}1.4, 1.6, and 1.7, respectively) expressed in the periphery. Two peripheral nerve channels, Na\textsubscript{v}1.6 and 1.7, show evidence of TTX-resistance evolution that parallels that previously described for skeletal muscle channel Na\textsubscript{v}1.4 (Geffeney et al. 2005; Feldman et al. 2009, 2010). In two cases, the amino acid substitution leading to putative TTX resistance is identical across paralogs. Na\textsubscript{v}1.7 has additional substitutions not seen in either of the other paralogs that are also expected to interfere with TTX binding. In a limited sample of individuals from across a wide geographic range (OR, CA, and VA), we found no evidence of variation in the TTX-binding region of either Na\textsubscript{v}1.6 or Na\textsubscript{v}1.7, suggesting either dramatically coincident replacements across populations in these genes, or more likely, that the substitutions in these two genes are fixed across the species’ range. Because Na\textsubscript{v}1.4 varies both within and among populations (Geffeney et al. 2005; Feldman et al. 2009, 2010), the latter scenario suggests that resistance in peripheral nerves is of more ancient origin than resistance in muscle. Three Na\textsubscript{v} paralogs found primarily in the CNS, Na\textsubscript{v}1.1–1.3, showed no evidence of TTX resistance, consistent with protection of these channels from toxins by the blood–brain barrier.

Ingested TTX has dramatic effects on peripheral nerves (Kao and Fuhrman 1963; Kiernan et al. 2005), suggesting that both Na\textsubscript{v}1.6 and 1.7 are likely to experience strong selection for resistance in predators that consume tetrodotoxin prey. In humans exposed to pufffish toxin, sensory symptoms (numbness and paresthesia) occur in low-grade poisoning, with higher-grade poisoning leading to motor symptoms such as paralysis and slurred speech (Isbister and Kiernan 2005). Sensory symptoms in response to low doses of TTX likely derive from the blockade of Na\textsubscript{v}1.7, which sets the threshold for the initiation of action potentials in sensory neurons, particularly those involved in olfaction and nociception (Dib-Hajj et al. 2013). Later motor symptoms are likely mediated by effects on Na\textsubscript{v}1.6, which is found mainly at nodes of Ranvier, where it is involved in propagation of action potentials along axons (Caldwell et al. 2000). Human deaths from severe TTX poisoning are usually caused by respiratory failure, and evidence from laboratory mammals suggests that this effect derives from direct blockade of sodium channels in the diaphragm, which seems to be more sensitive to the toxin than most muscles (Cheng et al. 1968). However, squamates lack a diaphragm and instead power respiration with a variety of axial muscles (Rosenberg 1973; Carrier 1989). It is probable that the disruption of the motor neurons innervating this musculature contributes to respiratory failure in squamates with vulnerable Na\textsubscript{v}1.6. Another source of mortality in TTX poisoning, especially in laboratory mammals, is severe hypotension (Zimmer 2010), which is due in part to the blockade of sympathetic nerves (Feinstein and Paimre 1968). Na\textsubscript{v}1.7 is highly expressed in the sympathetic nervous system (Dib-Hajj et al. 2013), suggesting that TTX-induced hypotension may contribute to selection for TTX resistance in this channel.

Both western (OR, CA) and eastern (VA) populations of Th. sirtalis have substitutions in Na\textsubscript{v}1.6 and 1.7 that are predicted to confer TTX resistance to peripheral nerves, but as of yet, individuals with resistant Na\textsubscript{v}1.4 have only been detected on the west coast where the species overlaps with highly toxic Taricha newts (Geffeney et al. 2005). This result suggests that the origin of TTX resistance in nerves likely antedates that of TTX resistance in muscle. In general, nerves appear to be impaired by lower TTX concentrations than required to impair muscles. TTX-sensitive nerves are typically blocked by nanomolar concentrations of TTX (Kao and Fuhrman 1963), whereas blocking action potentials in muscle fibers, which tend to have diameters an order of magnitude wider than the axons that innervate them, requires concentrations an order of magnitude higher (Geffeney et al. 2002). Consequently, resistant peripheral nerves alone are likely to provide sufficient protection to allow ingestion of very small amounts of TTX, such as a garter snake might encounter when consuming the mildly toxic eastern newt N. viridescens (Yotsu-Yamashita and Mebs 2001). However, muscle resistance will be necessary to provide protection against the higher concentrations of TTX seen in many Taricha populations (Hanifin et al. 2008).

Although we suggest that resistance in Na\textsubscript{v}1.6 and 1.7 has evolved specifically in response to selection mediated by TTX in prey, it is also possible that either or both nerve channels evolved resistance as a byproduct of selection for other functional modifications. In general, TTX-resistant substitutions
are associated with a decrease in conductance and/or sodium selectivity (Lee et al. 2011; Feldman et al. 2012). Although it is unlikely that reduced conductance or selectivity would be favored by selection in the absence of another selective agent such as TTX, other properties of sodium channels are also affected by the pore region and may be more likely targets of selection. For example, one of the substitutions in garter snake Na\(_{1.7}\) is seen in pufferfishes and a number of nontetrodotoxic fishes and has been shown to delay recovery from slow inactivation (Wu et al. 2013). Mammalian Na\(_{1.7}\) is characterized by a slow recovery from inactivation when compared with Na\(_{1.6}\) (Herzog et al. 2003), so it is possible that the Ala–Gly substitution seen in garter snakes reflects a refinement of this property. The Ile–Val substitution seen in both garter snake Na\(_{1.4}\) and Na\(_{1.6}\) has also been shown to affect slow inactivation, altering its voltage sensitivity (Lee et al. 2011). Interestingly, garter snake Na\(_{1.6}\) also has a Val–Ile substitution located in the analogous site of the DIII P-loop that may be involved in compensating for this change in voltage sensitivity.

The lack of resistance evolution in CNS channels can be attributed to the effectiveness of the blood–brain barrier against TTX. The blood–brain barrier of reptiles is similar to that of mammals (Kenny and Shivers 1974; Csern and Bundgaard 1984) and prevents the transport of many molecules, particularly hydrophilic compounds and those of relatively high molecular weight. TTX has a molecular weight of 319.27 and is highly polar and thus lipid-insoluble. Based on these properties, one simple model (Fu et al. 2008) predicts a log ratio of steady-state brain to plasma concentrations (log BB) of —3.54 for TTX, meaning that concentrations of TTX should be 3–4 orders of magnitude lower in the brain than in the plasma. To our knowledge, there has been no direct empirical quantification of the log BB of TTX, although one early study found low concentrations of injected TTX in the brain as compared with other tissues (Kao 1966). In accordance with this observation, unconsciousness occurs only in the most severe cases of TTX poisoning in humans (Isbister and Kiernan 2005). The snake from which our BAC library was constructed coexisted with the most toxic newts known (Kiernan 2005). The snake from which our BAC library was constructed coexists with the most toxic newts known (Kiernan 2005). The snake from which our BAC library was constructed...
DNA probes to scan a BAC library, which had been previously constructed using an individual snake from Benton County, OR (http://www.genome.gov/11008350, last accessed April 2, 2014). This population consists of snakes with medium-high resistance to TTX and coexists with a highly toxic population of *Taricha* newts (Geffeney et al. 2005; Hanifin et al. 2008). Our approach is described only briefly below; see Janes et al. (2011) for an overview of our approach and Ross et al. (1999) for a detailed protocol.

Probes for BAC library scanning were designed using Primer3 (Koressaar and Remm 2007; Untergasser et al. 2012) and four sequence resources: The transcriptome of a related species, *Th. elegans* (Schwartz et al. 2010), the CDS of *SCN4A* from *Th. sirtalis* (Geffeney et al. 2005), the CDS of an unknown *SCNA* paralog identified using degenerate polymerase chain reaction (PCR) of brain cDNA (Chuckalovcak 2010), and the universal probe database UProbe (Kellner et al. 2005; http://uprobe.genetics.emory.edu, last accessed April 2, 2014). These probes fell into four categories, each of which was designed to detect one of the four major branches in the radiation of the SCNA family: SCN1/2/3/9A, SCN4A, SCN5/10/11A, and SCN8A (supplementary table S2, Supplementary Material online). Each probe consisted of two DNA oligomers, or “overgos” (22 or 24 bp) with 8 bp of complementary sequence at the 3′-end (supplementary table S2, Supplementary Material online). These overgos were combined in solution and radiolabeled following Ross et al. (1999), resulting in probes 36–40 bp in length.

The first 144 plates of the *Th. sirtalis* BAC library were gridded onto high-density nylon filters using a QBot gridding robot (Genetix, Ltd.) (Ross et al. 1999). Hybridization of labeled probes to these filters was performed as in Ross et al. (1999) in two steps. First, all 31 probes were pooled and 99 putative positives were identified using a Typhoon imager (GE Healthcare). Next, these putative positives were gridded onto new filters, which were hybridized with 18 smaller pools of probes to confirm positive hits (supplementary table S2, Supplementary Material online). Following this step, 31 BAC colonies were selected to be sequenced. We also added nine BAC colonies that had been previously identified using screens of the same library with *SCN4A*-specific probes (Chuckalovcak JP, unpublished data), for a total of 40 putative positives (supplementary table S3, Supplementary Material online).

**DNA Isolation, Sequencing, and Assembly**

These 40 colonies were cultured and BAC DNA was isolated in duplicate using a Qiagen R.E.A.L. Prep 96 Plasmid Kit. DNA concentration was quantified using a Qubit fluorometer (Invitrogen). Samples were then diluted to equal concentration (16.67 ng/µl) and pooled for sequencing. A shotgun library was prepared for 454 sequencing using standard protocols and was sequenced on two separate half-plate runs using a Roche GS-FLX with Titanium reagents at the University of Virginia (UVA) Biology Genomics Core Facility. Sequences (138 Mb) from the two half-plate runs were combined for assembly using gsAssembler 2.3 (Roche). Reads matching the *Escherichia coli* genome or the pTarBAC2.1 vector were filtered, and default assembly settings were used. The assembly produced 812 contigs with an average size of 5,096 bp and an average read depth of 31×.

**Gene Annotation**

Contigs containing *SCNA* genes were initially identified with BLASTn searches using amino acid sequences of *Th. sirtalis* Na,1.4 (Geffeney et al. 2005) and the unknown *Th. sirtalis* paralog (Chuckalovcak 2010) as queries and the 454 output as the database. A variety of other BLAST (Basic Local Alignment Search Tool) searches using publicly available SCNA sequences were used to confirm these results. In all, 32 contigs ranging from 1,471 to 39,713 bp in length (average 12,416 bp) and with an average read depth of 74× were identified as containing SCNA exons. Most contigs contained multiple exons and the intervening intronic sequence (fig. 1). Edges between some contigs were identified by gsAssembler, providing further evidence of contig affinity (fig. 1). In two cases, two exon-containing contigs shared an edge with a third small (~20 bp) contig. Identified contigs included six versions of the first half of SCNA (exons 1–16), but only three versions of the second half (exons 17–26). Note that we use the numbering scheme 1–26 regardless of paralog, following Widmark et al. (2011).

To identify the specific paralog contained in each contig, we used BLAT (Kent 2002) searches against the *A. carolinensis* genome (Alföldi et al. 2011), whose paralogs were identified by Zakon et al. (2011). The results of these searches suggested that our sequences contained exons 1–16 of *SCN1A* (three contigs; exon 9 was incomplete), exons 1–16 of *SCN2A* (four contigs), exons 1–17 of *SCN3A* (four contigs), exons 1–26 of *SCN4A* (four contigs; complete CDS), exons 1–26 of *SCN8A* (11 contigs; complete CDS), and exons 1–26 of *SCN9A* (six contigs; nearly complete CDS, exon 24 was incomplete) (fig. 1). Two versions of contigs containing *SCN8A* exons 6–8 were present; these contained minor variation in intronic regions as well as a single synonymous difference in exon 8. We were unable to detect evidence of *SCN5A*, 10A, or 11A in our sequence. Exon–intron boundaries were identified by comparing sequences to published genomes (*A. carolinensis* and *Homo sapiens*). The exon–intron structure of contigs was annotated using Geneious 6.1.7 (Biomatters). We also used algorithms within Geneious 6.1.7 to measure GC content and to detect CpG islands. To explore the genomic content of intronic sequence, contigs were screened against a database of vertebrate simple repeats and TEs (Repbase) using Censor (Kohany et al. 2006).

**Gap Filling**

Gaps in the CDS for SCN1A, 2A, and 3A were filled using Sanger sequencing of brain cDNA, which was obtained from a captive adult female *Th. sirtalis* that had been previously captured from the highly TTX-resistant population in Willow Creek, CA. The CDSs for exons 14–26 of *SCN1A* (44 reads, 3,114 bp, 4.7× coverage, 92% of sequence from both strands), exons 16–26 of *SCN2A* (32 reads, 2,632 bp, 4.3×, 77%
from both strands), and exons 17–26 of SCN3A (38 reads, 2,821 bp, 4.4×, 79% from both strands) were obtained using primers designed from the transcriptome of Th. elegans (Schwartz et al. 2010) (supplementary table S5, Supplementary Material online). We attempted 3′-RACE (Rapid Amplification of cDNA Ends) to obtain the sequences of the 3′-ends of transcripts, but this gave low-quality sequence in SCN1A and was unsuccessful in SCN2A; therefore, the sequences of exon 26 were incomplete for these paralogs (690 and 916 bp, respectively). The complete CDS of exon 26 (1,193 bp, including the stop codon) plus 222 bp 3′-UTR was obtained for SCN3A. The CDS for the missing segment of SCN1A exon 9 was obtained using primers designed from contigs 360 and 88 of our 454 run (supplementary table S5, Supplementary Material online). Finally, CDS for the missing segment of SCN9A exon 24 was obtained from previous degenerate PCR results (Chuckalovcak 2010), as BLAST searches against our 454 contigs showed that sequence containing exons 19–26 obtained through this method belonged to SCN9A. All PCR followed standard protocols, and post-PCR cleanup was achieved using a combination of exonuclease I and shrimp alkaline phosphatase (Exo-SAP, USB). Sequencing was carried out either on an ABI 3130 at UVA’s Genomics Core Facility or on an ABI 3730 at Yale University’s DNA Analysis Facility.

Paralog Identification
The identity of each paralog was confirmed by using BLAT searches of the entire CDS against the A. carolinensis genome. The best BLAT hit was invariably the paralog assigned using our preliminary BLAT searches of separate contigs (supplementary table S4, Supplementary Material online). We also constructed a protein phylogeny of the six Th. sirtalis paralogs and all nine A. carolinensis paralogs (Zakon et al. 2011, accession nos. BK007953–BK007961). Amino acid sequences were aligned using MUSCLE (Edgar 2004) and a maximum-likelihood phylogeny was constructed using PhyML (Guindon and Gascuel 2003). Each Th. sirtalis paralog was shown to be most similar to its assigned ortholog in A. carolinensis, and each of these nodes had a bootstrap value of 100 (supplementary fig. S1, Supplementary Material online). We tested for evidence of gene conversion among paralogs using a MUSCLE alignment of the six Th. sirtalis paralogs in GENECONV (Sawyer 1989).

Putative TTX Resistance and Interpopulation Variation
We scanned for putative TTX resistance by comparing the P-loop regions from CDS of the six paralogs we identified to reference sequence from both A. carolinensis and a mammal (the rat Rattus norvegicus). From paralogs whose sequences indicated putative TTX resistance, we also examined sequences from individuals from two populations: Willow Creek, CA, which consists of a mixture of low- and high-resistance individuals (Feldman et al. 2009, 2010) and Mountain Lake Biological Station, VA, which displays low resistance and occurs outside the range of Taricha but overlaps with the mildly toxic N. viridescens (Yotsu-Yamashita and Mebs 2001).

We used tissue samples from 16 individual snakes whose SCN4A genotypes and organismal resistance were reported by Feldman et al. (2009): CA high-resistance n = 10 (SCN4A accession nos. FJ571047–FJ571056), CA low-resistance n = 1 (GU154075), and VA n = 4 (FJ571057–FJ571060). Sample sizes for individual regions were as follows: SCN8A DIII and DIV, CA high-resistance n = 2, VA n = 2; SCN9A DIII CA high-resistance n = 10, CA low-resistance n = 1, VA n = 1; SCN9A DIV CA high-resistance n = 10, CA low-resistance n = 1, VA n = 4. Primers were designed from our 454 contigs (supplementary table S6, Supplementary Material online). PCR and cleanup were carried out as above, and sequencing was performed on an ABI 3730 at the Yale DNA Analysis Facility.

Accession Numbers
All new sequences have been deposited in GenBank under accession numbers KJ908861–KJ908938 and KM066119. The predicted CDSs for each paralog have been deposited as Third Party Annotations under accession numbers BK008860-BK008865.

Supplementary Material
Supplementary figure S1 and tables S1–S6 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments
The authors thank Michelle Sivilich and Megan Kobiela for assistance in the laboratory, Dan Sloan for advice on molecular techniques, Shana Geffeney and Charles Hanifin for helpful discussions, Nigel Delaney and Yoel Stuart for logistical support in Cambridge, and Angela Hornsby, Marjorie Matocq, and Kendra Sewall for comments on the manuscript. This material is based upon work supported by the National Science Foundation (grant numbers DEB 0922216 to E.D.B. III and DEB 1034686 to M.E.P.). Publication of this article was supported by Virginia Tech’s Open Access Subvention Fund.

References


