Gene Conversion Facilitates the Adaptive Evolution of Self-Resistance in Highly Toxic Newts

Kerry L. Gendreau[1]*, Angela D. Hornsby[1,2], Michael T. J. Hague[3], and Joel W. McGlothlin[1]

[1]Department of Biological Sciences, Virginia Tech, Blacksburg, United States
[2]Philip L. Wright Zoological Museum, Division of Biological Sciences, University of Montana, Missoula, United States
[3]Division of Biological Sciences, University of Montana, Missoula, MT, United States

***Corresponding author**: kerryg@vt.edu

Abstract

Reconstructing the histories of complex adaptations and identifying the evolutionary mechanisms underlying their origins are two of the primary goals of evolutionary biology. *Taricha* newts, which contain high concentrations of the deadly toxin tetrodotoxin (TTX) as an antipredator defense, have evolved resistance to self-intoxication, which is a complex adaptation requiring changes in six paralogs of the voltage-gated sodium channel (Na_v) gene family, the physiological target of TTX. Here, we reconstruct the origins of TTX self-resistance by sequencing the entire Na_v gene family in newts and related salamanders. We show that moderate TTX resistance evolved early in the salamander lineage in three of the six Na_v paralogs, preceding the proposed appearance of tetrodotoxic newts by ~100 million years. TTX-bearing newts possess additional unique substitutions across the entire Na_v gene family that provide physiological TTX resistance. These substitutions coincide with signatures of positive selection and relaxed purifying selection, as well as gene conversion events, that together likely facilitated their evolution. We also identify a novel exon duplication within $Na_v1.4$ encoding an expressed TTX-binding site. Two resistance-conferring changes within newts appear to have spread via nonallelic gene conversion: in one case, one codon was copied between paralogs, and in the second, multiple substitutions were homogenized between the duplicate exons of $Na_v1.4$. Our results demonstrate that gene conversion can accelerate the coordinated evolution of gene families in response to a common selection pressure.

Key words: gene conversion, adaptation, coevolutionary arms races, molecular evolution, newts, toxin resistance

Introduction

Fitting evolutionary models to molecular sequences in a phylogenetic context can help piece together the key steps in adaptive evolution and uncover the relative contributions of selection and other evolutionary mechanisms to adaptive phenotypic evolution (Smith et al. 2020). Comparative studies of convergence, or the repeated evolution of characters within different lineages undergoing the same environmental challenges, provide powerful evidence of both adaptation and connections between genetic and

41 phenotypic change (Losos 2011). Investigations into the molecular basis of convergence have revealed
42 multiple occurrences of parallelism, where different lineages have evolved changes within the same
43 proteins, and occasionally at the same amino acid sites, in response to shared selective pressures, such as
44 insects that have evolved the ability to feed on toxic plants (Zhen et al. 2012) and populations of ducks
45 and humans living at high elevations (Graham and McCracken 2019). Such patterns support important
46 roles for both positive selection and constraint in the origin of complex adaptations (reviewed by Storz
47 2016).

48 Resistance to tetrodotoxin (TTX), a potent neurotoxin, has evolved convergently in several
49 distantly related organisms, including pufferfish, snakes, and newts (reviewed by Soong and Venkatesh
50 2006; Toledo et al. 2016), and therefore offers an ideal system to investigate the molecular basis of
51 adaptive evolution (Arbuckle et al. 2017). The genetic basis of TTX resistance, which is well established
52 in tetrodotoxic puffer fish (Jost et al. 2008) and in snakes that consume TTX-bearing prey (Geffeney et al.
53 2002; Feldman et al. 2012; McGlothlin et al. 2014; McGlothlin et al. 2016), involves amino acid
54 substitutions in the toxin's target, voltage-gated sodium channels. While voltage-gated sodium channels
55 are commonly abbreviated as *SCNA* genes and Na_v proteins, hereafter we simplify this nomenclature by
56 abbreviating both as Na_v. Na_v channels are responsible for the initiation and propagation of action
57 potentials in excitable cells and are composed of four domains (DI–DIV), each comprising six
58 transmembrane helices and a pore-loop region (the P-loop; Fux et al. 2018). The four P-loops, one in each
59 domain, form a pore within the membranes of excitable cells to selectively allow sodium ions to cross
60 when the channel is open. TTX exerts its effects by binding to the P-loops of sensitive channels and
61 preventing sodium entry into cells, thus blocking action potentials. Gene duplication events have resulted
62 in six Na_v paralogs, each with tissue-specific expression, that are shared across all tetrapods (table 1),
63 with additional lineage-specific duplications occurring in amniotes (Widmark et al. 2011; Zakon 2012).
64 Because the structure of these paralogs is highly conserved, each has the potential to be blocked by TTX
65 if it lacks resistance-conferring substitutions.

66 Species that possess or consume TTX must either have a full complement of resistant paralogs or
67 otherwise shield sodium channels from contact with the toxin. Indeed, resistant substitutions are present
68 in all eight of the Na_v genes within the genomes of multiple species of TTX-bearing pufferfish (from the
69 family Tetraodontidae; Jost et al. 2008) and in six of the nine Na_v genes in *Thamnophis sirtalis* snakes
70 that consume TTX-bearing *Taricha* newts (McGlothlin et al. 2014; Perry et al. 2018). The three brain
71 channels $Na_v1.1$, $Na_v1.2$, and $Na_v1.3$ of snakes remain TTX sensitive, but presumably are protected from
72 TTX by the blood-brain barrier (McGlothlin et al. 2014). In snakes, the evolution of extreme TTX
73 resistance appears to follow a predictable, stepwise substitution pattern across TTX-exposed members of
74 the Na_v gene family, with substitutions in heart and peripheral nerve channels preceding those in the

Table 1. Nomenclature for Voltage-Gated Sodium Channel Genes

Gene name	Protein name	Tissue expression[a]
SCN1A	$Na_V1.1$	Brain
SCN2A	$Na_V1.2$	Brain
SCN3A	$Na_V1.3$	Brain
SCN4A	$Na_V1.4$	Muscle
SCN5A	$Na_V1.5$	Heart
SCN8A	$Na_V1.6$	Brain / peripheral nervous system

[a] Patterns of tissue expression are inferred from studies of gene orthologs in mammals (reviewed in Yu and Catterall 2003).

75 muscle channel gene, $Na_V1.4$, which evolves resistance only in snakes locked in coevolutionary arms
76 races with highly tetrodotoxic amphibian prey (Feldman et al. 2012; McGlothlin et al. 2016; Perry et al.
77 2018).

78 Less is known about the evolutionary history of TTX resistance in *Taricha* newts, the highly toxic
79 coevolutionary partner of *Thamnophis* (Brodie and Brodie 1990; Brodie et al. 2002; Williams et al. 2010;
80 Hague et al. 2020). The extreme toxicity of *Taricha*, which has been elaborated by the ongoing
81 coevolutionary arms race with garter snakes, builds upon lower levels of toxicity that evolved ~30 million
82 years ago (mya) within "modern" newts (tribe Molgini; Hanifin and Gilly 2015; divergence date
83 estimated by Hime et al. 2021). The evolution of toxicity necessitates the evolution of toxin
84 autoresistance so that a prey species is not incapacitated by its own antipredator defense (Jost et al. 2008;
85 Toledo et al. 2016; Tarvin et al. 2017; Márquez et al. 2019). Understanding the timing and details of this
86 autoresistance can shed light on the genetic processes underlying the predator-prey arms race. Hanifin and
87 Gilly (2015) compared the sequences of one sodium channel gene, the muscle paralog $Na_V1.4$, across
88 several salamander species and identified substitutions in the P-loops of DIII and DIV that provide
89 extreme TTX resistance to the muscles of TTX-bearing newts. Importantly, the sister group of these toxic
90 newts had substitutions in the same gene providing more moderate resistance, indicating that the
91 evolution of autoresistance in a common ancestor paved the way for the evolution of extreme toxicity.
92 More recently, Vaelli et al. (2020) used transcriptome sequencing to characterize the genetic basis of
93 physiological resistance to TTX in *Taricha granulosa* and identified substitutions within TTX-binding
94 regions in the other five Na_V paralogs, many of which occur in within the P-loop of DI. However, because
95 it is unknown whether other salamander species possess TTX resistance in these paralogs, the order and
96 timing of the evolutionary events leading to autoresistance in toxic newts is still unknown. Furthermore,
97 no studies to date have applied evolutionary models to test for the relative importance of mechanisms
98 such as positive selection, relaxed constraint, and interlocus gene conversion in the evolution of newt
99 TTX resistance.

100 Here, we trace the evolutionary history of the entire Na_v gene family across the salamander
101 phylogeny to show the order in which resistant substitutions appeared. Using published genome
102 sequences and newly generated sequence data, we characterize the genomic structure of Na_v genes in
103 newts and their relatives, inferring the timing of resistant substitutions leading to the extreme TTX
104 resistance observed across all Na_v paralogs in *Taricha* newts (Vaelli et al. 2020). We estimate rates of
105 synonymous and nonsynonymous substitution to identify positive selection. In addition, we assess the
106 potential of nonallelic gene conversion, a process by which sequence is copied from one paralog to
107 another (Chen et al. 2007), to act as a source of adaptive variation. Combining these data provides insight
108 into the evolutionary mechanisms underlying the origin of a uniquely potent chemical defense.
109

110 # Results

111 ## Genomic Structure of Voltage-Gated Sodium Channels

112 We used targeted sequence capture to characterize Na_v sequences from the genomes of five salamander
113 species (order Urodela), including three TTX-bearing newts (family Salamandridae, subfamily
114 Pleurodelinae, tribe Molgini), *Notophthalmus viridescens*, *Taricha torosa*, and *Taricha granulosa* (n = 3
115 diploid individuals of each species), and two salamanders that do not possess tetrodotoxin,
116 *Cryptobranchus alleganiensis* (Crypotobranchidae) and *Plethodon cinereus* (Plethodontidae, n = 2 each).
117 We also identified Na_v sequences within two publicly available salamander genome sequences:
118 *Ambystoma mexicanum* (Ambystomatidae; Smith et al. 2019; AmexG.v6 assembly) and *Pleurodeles waltl*
119 (Salamandridae; Elewa et al. 2017) and a full-body transcriptome from the fire salamander *Salamandra*
120 *salamandra* (Salamandridae; Goedbloed et al. 2017; BioProject accession PRJNA607429), all three of
121 which appear to lack tetrodotoxin (Hanifin 2010). The split between *Cryptobranchus* (suborder
122 Cryptobranchoidea) and all the other salamanders in our study (members of suborder Salamandroidea)
123 represents the most ancient division in the phylogeny of extant salamanders (~160 mya; Hime et al.
124 2021).

125 We identified six Na_v genes in the genomes of all salamander species, which is consistent with
126 observations in other amphibians (Zakon et al. 2011). Hereafter, we use the exon delineation introduced
127 in Widmark et al. (2011), where domain I (DI) is encoded by exons 2-9, domain II (DII) is encoded by
128 exons 13-16, domain III (DIII) is encoded by exons 18-23, and domain IV (DIV) is encoded by exons 25-
129 26. We obtained near full-length assemblies for all paralogs (table S1); however, a few exons containing
130 TTX-binding sites, including exon 15 (encoding the DII P-loop) of $Na_v1.2$ from *N. viridescens* and exon
131 22 (encoding part of the DIII P-loop) of $Na_v1.2$ for several newt species, were missing from our
132 assemblies. Polymorphism was rare in our assemblies and we observed few nonsynonymous mutations
133 within the newt genomes, but we found slightly elevated polymorphism in *N. viridescens* relative to other

Table 2. Locations of Na_v Genes in the *Ambystoma mexicanum* AmexG.v6 Genome Assembly.

Gene	Chromosome	Start	End	Strand	Length (bp)
$Na_v1.1$	9q	503,107,904	503,797,933	+	690,029
$Na_v1.2$	9q	507,685,503	508,688,108	-	1,002,605
$Na_v1.3$	9q	509,830,827	510,576,927	+	746,100
$Na_v1.4$	13q	113,343,821	115,707,581	+	2,363,760
$Na_v1.4$ exon 26b	13q	115,891,071	115,892,251	+	1,180
$Na_v1.5$	2p	562,159,042	563,030,267	-	871,225
$Na_v1.6$	3q	465,212,114	466,692,047	-	1,479,933

species (table S2). No heterozygosity or nonsynonymous polymorphisms were observed within any of the known TTX-binding P-loop regions within any of the species sequenced for this study.

Synteny of Na_v genes in *A. mexicanum* is conserved relative to other tetrapods (fig. S1), which allowed us to use the *A. mexicanum* sequences as a baseline to confidently identify Na_v paralogs in all species. Three of the paralogs, $Na_v1.1$, $Na_v1.2$, and $Na_v1.3$ are arrayed in tandem on *A. mexicanum* chromosome 9, with $Na_v1.2$ inverted relative to its neighboring paralogs (table 2, fig. S1). The additional three paralogs, $Na_v1.4$, $Na_v1.5$, and $Na_v1.6$, are each located on separate chromosomes (table 2). In the gene family tree built from amino acid sequences, all salamander Na_v proteins formed a monophyletic clade (bootstrap support > 98%) with the corresponding orthologs from the genomes of the frogs *Xenopus tropicalis* and *Nanorana parkeri*, which we included as outgroups (fig. 1). The gene family tree constructed from the nucleotide coding sequences of these genes yielded a similar topology, with each salamander Na_v ortholog forming a monophyletic clade. However, in the nucleotide gene family tree, the three *X. tropicalis* nerve channels $Na_v1.1$, $Na_v1.2$, and $Na_v1.3$ formed a monophyletic clade that is distinct from the salamander sequences (bootstrap support 86%; fig. S2). The same tree topology was resolved when partitioning for the third codon position (see Supplemental Data on Dryad).

Partial Duplication of $Na_v1.4$ and Evolution of TTX Resistance in Duplicated Domains

Our search of the *A. mexicanum* genome revealed a partial tandem duplication of the 3' end of the $Na_v1.4$ gene, including the full coding region of exon 26, located ~180,000 base pairs downstream of the full-length $Na_v1.4$ gene on the same DNA strand (table 2). Both exon 26 copies are similar in length to each other and to exon 26 of other paralogs, encoding open reading frames of approximately 390 base pairs without introduced stop codons. Exon 26 is the 3'-terminal exon of the $Na_v1.4$ gene and encodes the TTX-binding P-loop region of DIV. Hereafter, we refer to the duplicate exons as 26a (more proximal to exon 25) and 26b (more distal to exon 25) and duplicate P-loop regions as DIVa (more proximal to exon 25)

and DIVb (more distal to exon 25). We also found this duplicated exon in $Na_v1.4$ orthologs within the genomes of salamanders *P. cinereus*, *P. waltl*, *N. viridescens*, *T. torosa*, and *T. granulosa* and in published transcriptomes of *Tylototriton wenxianensis* and *Bolitoglossa vallecula*, but not in the transcriptome of *Hynobius retardatus* or in the genomes of *C. alleganiensis* or the frogs *X. tropicalis* or *N. parkeri*. This pattern suggests that the duplication event likely took place after the split of Cryptobranchoidea and Salamandroidea (fig. 2). Within the *S. salamandra* transcriptome, we found four

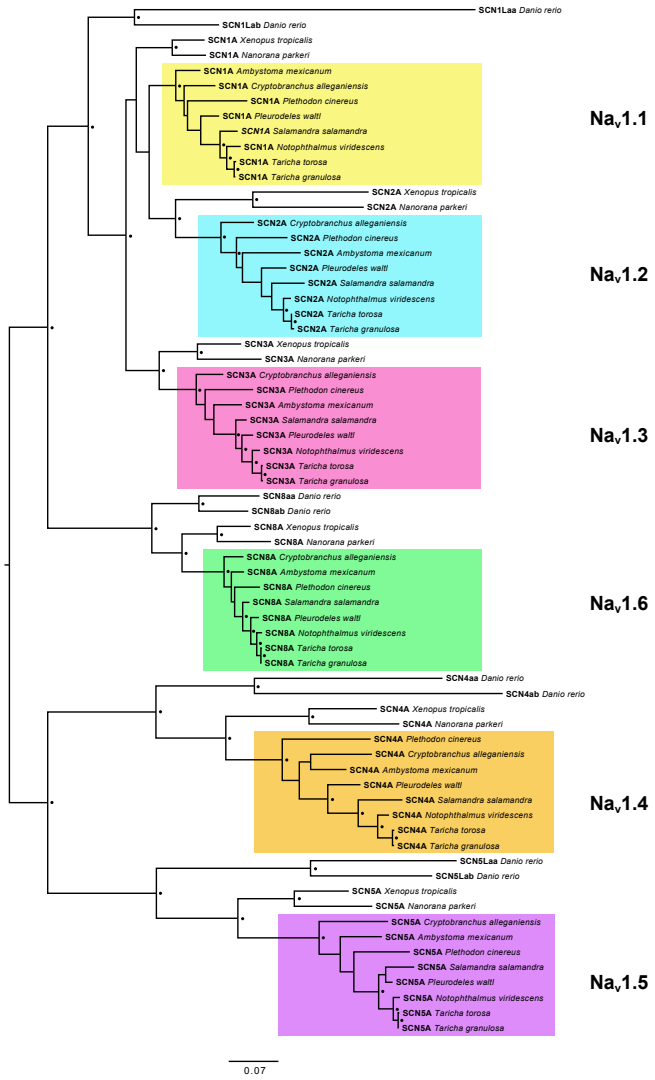

Fig. 1. Evolutionary relationships among salamander voltage-gated sodium channels. Midpoint rooted maximum likelihood tree constructed from alignment of 2188 amino acids from full coding sequence translations of salamander sodium channel genes, including sequences from two frogs (*Nanorana parkeri* and *Xenopus tropicalis*) and one fish (*Danio rerio*) as outgroups. Black circles indicate nodes with >90% bootstrap support.

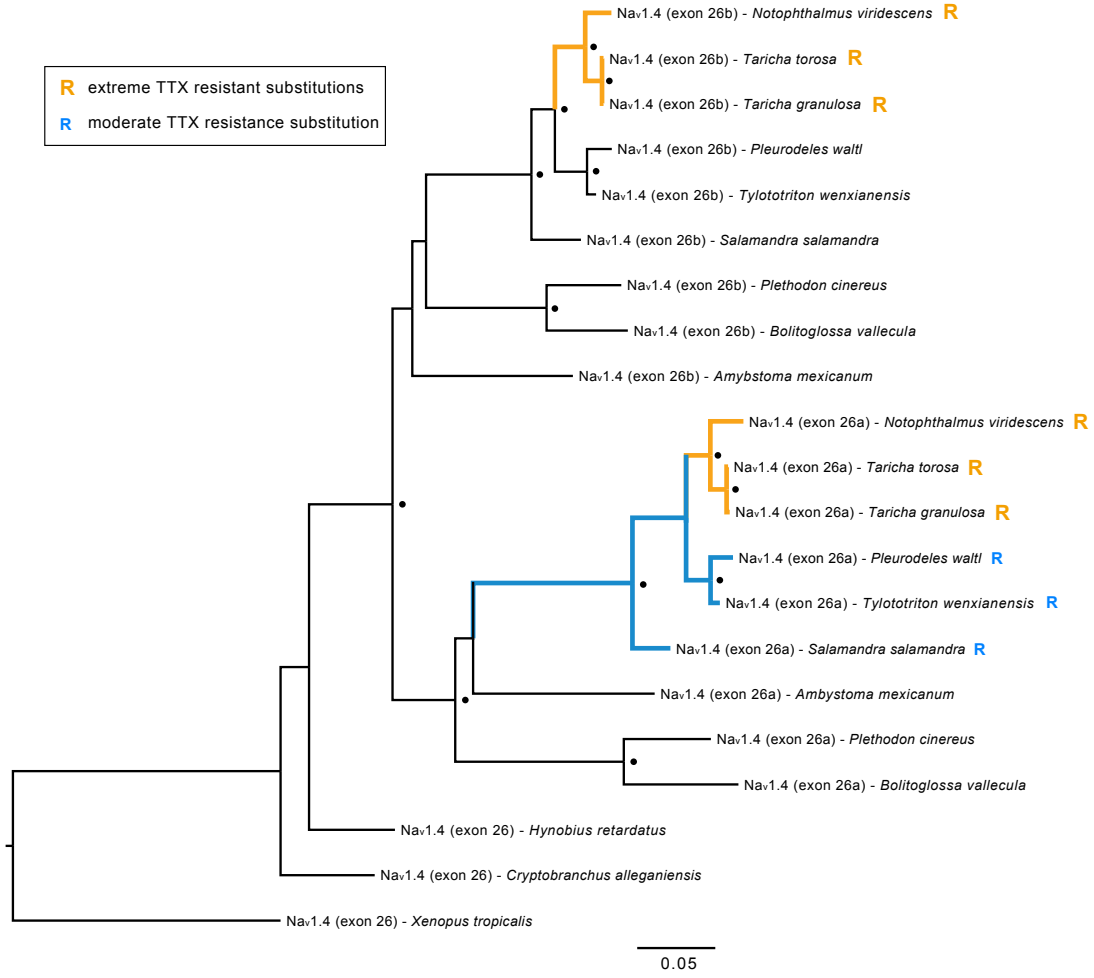

Fig. 2. Ancestral duplication and convergent evolution of TTX resistance in $Na_v1.4$ terminal exon 26. Maximum likelihood tree constructed from 1050 bp nucleotide alignment of $Na_v1.4$ exon 26 identified in salamander genomes and transcriptomes. “Exon 26a” and “exon 26b” in tip labels refer to the exon copy more proximal and more distal to exon 25, respectively. Black circles indicate node bootstrap support >95%. “R” at tips indicates the presence of substitutions conferring extreme (orange) and moderate (blue) TTX resistance. The same tree topology was resolved using an alignment of amino acids translations of these sequences.

165 unique RNA sequences transcribed from the $Na_v1.4$ locus, with alternative splicing of exon 17 and
166 alternative encoding of either exon 26a or 26b. Genome-mapped reads of multi-tissue transcriptomes of
167 *A. mexicanum* (Bryant et al. 2017; Caballero-Pérez et al. 2018; Nowoshilow et al. 2018) indicate that
168 these alternative transcripts have similar expression profiles across various tissues. Taken together, these

169 observations provide evidence that the duplication of exon 26 led to the creation of functional splice
170 variants in these salamanders.
171 While numerous nonsynonymous substitutions differentiated the duplicated $Na_v1.4$ exon 26
172 sequences relative to the original sequences within each genome, we found that identical substitutions
173 conferring extreme TTX-resistance to toxic newts (Hanifin and Gilly 2015) were present in both exons
174 from the genomes of all three TTX-bearing newts but not in other, less toxic salamander species (fig. 2).
175 Also consistent with the results of Hanifin and Gilly (2015), we found resistant substitutions conferring
176 moderate TTX resistance in exon 26a of *P. waltl*, *T. wenxianensis*, and *S. salamandra*; however, we
177 observed no resistant substitutions in exon 26b outside of the toxic newt clade.
178

179 **Evolution of TTX Resistance in Salamanders**

180 We characterized levels of TTX resistance in each Na_v paralog as extreme, moderate, and TTX-sensitive
181 based on previous site-directed mutagenesis experiments in which substitutions were introduced to TTX-
182 sensitive Na_v channels and cross-membrane Na^+ current was measured *in vitro* in the presence and
183 absence of TTX (table 3). Our results confirm that *T. granulosa* has six paralogs with extreme TTX
184 resistance (table 3, fig. 3). Our findings are consistent with those reported by Vaelli et al. (2020) with one
185 exception: we associate DIV (encoded by exon 26) substitutions A1529G, G1533V, and G1533A with
186 $Na_v1.1$ and Q1524E, G1533R, and G1533Q with $Na_v1.2$ based on synteny mapping (fig. S1), gene trees
187 (figs. 1, S2), and a phylogeny created from a coding sequence alignment of exon 26 (fig. S3), whereas the
188 previous study reversed these assignments. We also show that substitutions with extreme TTX resistance
189 are present in all six Na_v paralogs in two other species of highly toxic newt, *T. torosa* and *N. viridescens*,
190 indicating that the common ancestor of these three species possessed extreme TTX resistance. Many of
191 the substitutions in toxic newts parallel those found in TTX-bearing fish and in snakes that consume
192 tetrodotoxic amphibians (table 3).
193 No salamander species outside the clade of highly toxic newts possessed a full complement of
194 TTX-resistant Na_v paralogs, indicating that the evolution of full physiological resistance coincided with
195 the origin of extreme toxicity. However, we found at least three paralogs with moderate or extreme TTX
196 resistance in all salamander species we examined, indicating that the evolution of TTX resistance in newts
197 built upon more ancient changes that first appeared in their non-newt relatives. Substitutions conferring
198 moderate or extreme resistance were observed within the heart channel $Na_v1.5$ and brain/nerve channels
199 $Na_v1.1$ and $Na_v1.6$ of all salamander species, with additional resistance-conferring substitutions evolving
200 within TTX-bearing newts. As first shown by Hanifin and Gilly (2015), moderate resistance was present
201 in the skeletal muscle channel $Na_v1.4$ of *S. salamandra* and *P. waltl*, but not in the three other salamander
202 species we examined. Although our outgroup, the frog *X. tropicalis*, also contained a highly resistant

203 substitution in $Na_v1.2$, we found no evidence for resistance in this paralog in any salamanders outside of
204 tetrodotoxic newts.
205 Based on our ancestral sequence reconstructions, the most recent common ancestor of all
206 salamanders had three TTX resistant sodium channels: $Na_v1.1$ (brain, moderately resistant), $Na_v1.5$ (heart,
207 highly resistant), and $Na_v1.6$ (brain/peripheral nerves, moderately resistant; figs. 3, S4-S9). Moderate
208 resistance in the muscle channel $Na_v1.4$ appeared between 75–130 mya, after the divergence of
209 Ambystomatidae and Salamandridae (the family consisting of true salamanders and newts; Hanifin and
210 Gilly 2015). This gain in muscle resistance coincided with the appearance of two highly resistant
211 substitutions in DI of $Na_v1.5$, which are present in all Salamandridae. Extreme TTX resistance across all
212 Na_v paralogs evolved more recently, occurring approximately 30 mya, after the split between primitive
213 newts, which include *Pleurodeles*, and modern newts (Hime et al. 2021), which include all TTX-bearing
214 species. Over this time period, TTX resistance evolved in $Na_v1.2$ and $Na_v1.3$, and multiple additional
215 resistant mutations appeared and became fixed in $Na_v1.1$, $Na_v1.4$, and $Na_v1.6$.
216

Table 3. List of TTX Resistance-Conferring Substitutions Observed in Salamanders

Substitution[a]	Tetrodotoxic newts	Non-tetrodotoxic salamanders	Pufferfish and snakes[b]	Fold-change in TTX sensitivity	Resistance category	Citation
Y401C (DI)	$Na_v1.1$ $Na_v1.2$	$Na_v1.5$	Pufferfish: $Na_v1.4a$ $Na_v1.5La$ $Na_v1.5Lb$	~2500×	Extreme	Venkatesh et al. (2005); Jost et al. (2008)
Y401A (DI)	$Na_v1.1$ $Na_v1.3$ $Na_v1.6$	$Na_v1.1$	Pufferfish: $Na_v1.1La$ $Na_v1.6b$	>600×	Extreme	Jost et al. (2008); Vaelli et al. (2020)
Y401S (DI)	$Na_v1.5$	$Na_v1.5$	-	~7000×	Extreme	Leffler et al. (2005)
D1532S (DIV) G1533D (DIV)	$Na_v1.4$		Snake: $Na_v1.4$	-	Extreme[c]	Feldman et al. (2012); Hanifin and Gilly (2015)
E758D (DII)	-	$Na_v1.5$	Pufferfish: $Na_v1.4b$	~3000×	Extreme	Bricelj et al. (2005)
M1240T (DIII)	$Na_v1.1$ $Na_v1.4$	$Na_v1.1$ $Na_v1.5$	Pufferfish: $Na_v1.1La$ $Na_v1.1Lb$ $Na_v1.4a$ $Na_v1.4b$	~15×	Moderate	Jost et al. (2008)
V1233I (DIII)	$Na_v1.6$	$Na_v1.6$	Snake: $Na_v1.6$	~2×	Moderate	McGlothlin et al. (2014); Vaelli et al. (2020)
I1525V (DIV)	$Na_v1.6$	$Na_v1.5$ $Na_v1.6$	Snake: $Na_v1.6$	~2×	Moderate	Geffeney et al. (2005); McGlothlin et al. (2014); Vaelli et al. (2020)
I1525S (DIV)	$Na_v1.4$	$Na_v1.4$		-	Moderate[c]	Hanifin and Gilly (2015)
I1525T (DIV)	-	$Na_v1.4$ $Na_v1.5$		~7.7×	Moderate	Du et al. (2009)

[a] Amino acid sites are in reference to rat $Na_v1.4$ (accession number AAA41682)
[b] Parallel substitutions in Na_v channels of TTX-bearing pufferfish and in snakes that consume TTX-bearing prey
[c] Resistance category determined from action potentials recorded from salamander muscle fibers

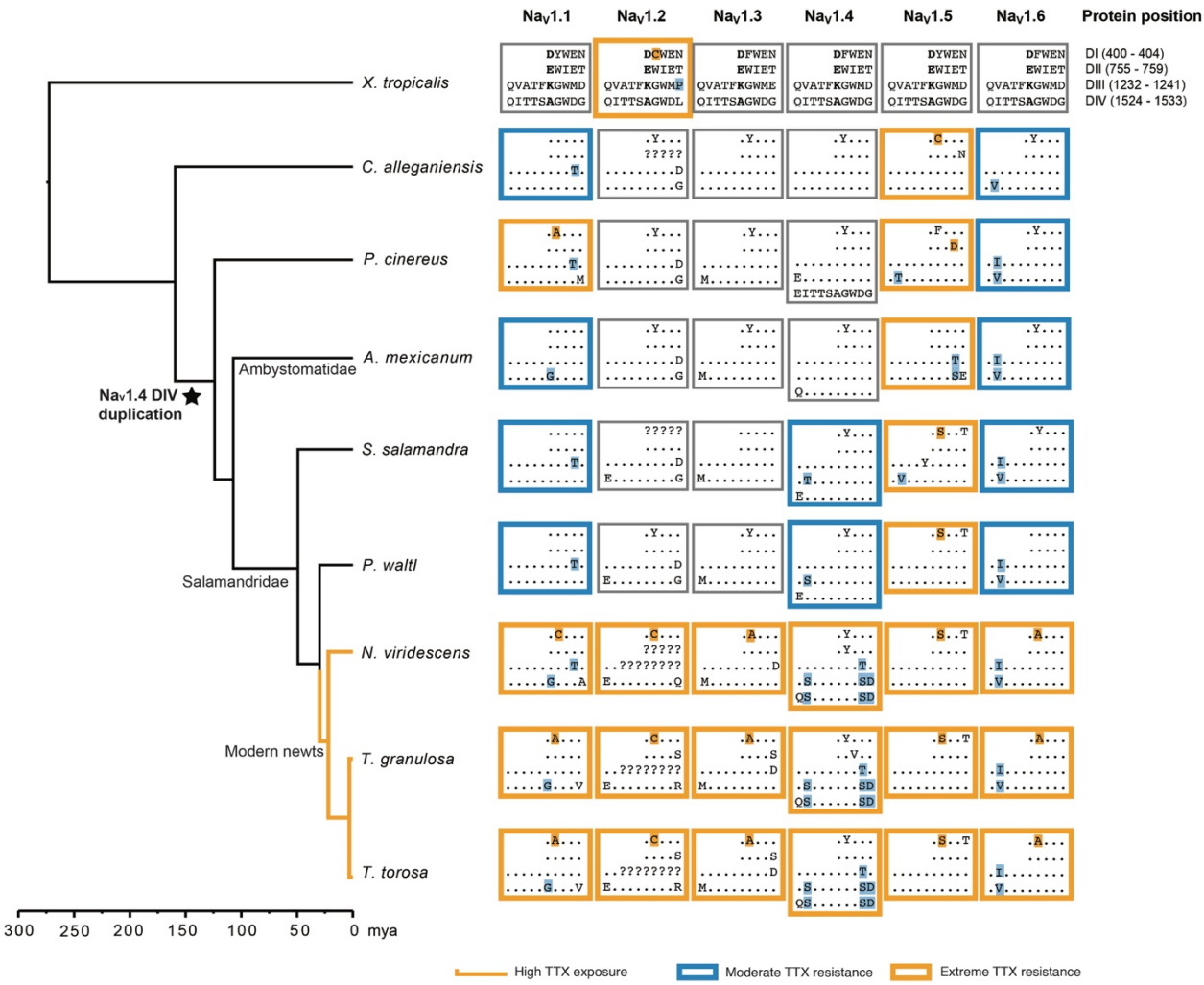

Fig. 3. Distribution of TTX-resistance conferring substitutions in salamander voltage-gated sodium channels. Toxic newt species are indicated with orange branches ("High TTX exposure"). Amino acids associated with Na^+ selectivity (400D, 755E, 1237K, 1529A) are shown in bold. Dots indicate identity with consensus sequences. Blue boxes indicate paralogs inferred to have moderate (~2-15-fold) resistance, orange boxes indicate paralogs with extreme (>300-fold) resistance, and grey boxes indicate paralogs without resistance or with insufficient sequence data. Substitutions known to confer TTX resistance are highlighted with respective colors. Extreme resistance in a paralog can result from the presence of one highly resistant substitution or the combination of multiple moderately resistant substitutions. Exon duplication has led to an additional TTX-binding domain (DIVb) in $Na_v1.4$ of all salamanders except *C. alleganiensis*. We did not identify the sequence encoding DIII of $Na_v1.2$ for any of the newts, however Vaelli et al. (2020) report that, in *T. granulosa*, this domain is identical in amino acid sequence to other salamanders. Amino acid sites are in reference to rat $Na_v1.4$ (accession number AAA41682).

217 **Selective Regimes and Evolutionary Rates**

218 In order to characterize the selective regimes acting on Na_v genes, we used the codeml program in PAML
219 (Yang 2007) to fit models of selection to Na_v codon alignments and compared nested models using
220 likelihood ratio tests (LRTs). We tested for site-specific positive selection within all amphibians by
221 comparing two sets of nested models: one set using a discrete distribution of ω values either with or
222 without positive selection (M1a vs. M2a; Yang et al. 2005); and another set fitting a continuous
223 distribution of ω values under purifying selection only, or adding categories of unconstrained evolution
224 and positive selection (M7 vs. M8 and M8a vs. M8; Yang et al. 2000). Parameter estimates and LRT
225 results for site models are summarized in table S3. To test for selection within toxic newts, we fit branch
226 (Yang 1998) and branch-site models (Zhang et al. 2005) to our datasets, which allowed ω to vary both
227 among codon sites and between toxic newts and other amphibians (summarized in table S4).

228 All models estimated relatively low ratios of nonsynonymous to synonymous substitution (d_N/d_S,
229 or ω ratios) for all Na_v paralogs (average ω ratios from branch models ranged from 0.05 to 0.25),
230 indicating pervasive purifying selection. Based on LRTs comparing one ω-value models (which allow
231 only for one ω value across the entire phylogeny) to branch models (allowing for a different ω ratio in the
232 toxic newt clade relative to other amphibians), we found that ω ratios were significantly ($P < 0.01$) higher
233 in toxic newts for $Na_v1.1$, $Na_v1.3$, and $Na_v1.4$ and non-significant ($P > 0.01$) for $Na_v1.2$, $Na_v1.5$, and
234 $Na_v1.6$ (fig. 4A, table S4). The largest difference in ω ratios in the branch test was observed for the
235 muscle channel $Na_v1.4$ (newt ω = 0.23; all salamanders ω = 0.10), which appears to be due to both an
236 increase in the proportion of unconstrained sites as well as a larger number of estimated sites undergoing
237 positive selection (fig. 4B). However, the posterior probability support for positive selection at many of
238 these sites was low, and LRTs from branch-site models indicated significant evidence for a shift in
239 positive selection only in paralog $Na_v1.3$ (tables S3, S4).

240 Our site and branch-site models identified a number of TTX-binding sites with elevated ω ratios
241 (tables 4, S5, S6). Because of the small number of species in our study, we had low power to detect
242 statistically significant positive selection (Yang et al. 2000; Anisimova et al. 2001; Kosakovsky Pond and
243 Frost 2005), and the posterior probabilities provided low-to-moderate support for positive selection at
244 most sites. We also note that the false positive rates of site selection models can be inflated due to gene
245 conversion events (Casola and Hahn 2009), among-site variation in d_S (Kosakovsky Pond and Frost
246 2005), and multinucleotide mutations (Venkat et al. 2018). However, results were consistent between the
247 M2a and M8 models, and the codons that were identified suggest that positive selection may have been
248 important for observed substitutions in TTX-binding regions.

249 Within the brain channels $Na_v1.1$, $Na_v1.3$, and $Na_v1.6$, putative positive selection was detected by
250 site models at site 401 within the DI P-loop, indicating selection acting across all salamanders rather than

251 specifically within toxic newts (table 4). Replacement of the aromatic amino acid at site 401 with a non-
252 aromatic amino acid can substantially impact TTX binding capacity (Leffler et al. 2005; Venkatesh et al.
253 2005; Vaelli et al. 2020), and we observed non-aromatic substitutions at this site within all brain channels
254 of highly toxic newts (fig. 3, table 3). The codon sequence for site 401 was variable across many
255 salamanders lacking TTX; however, almost all of the nonsynonymous changes observed outside of TTX-
256 bearing newts were biochemically conservative (both phenylalanine and tyrosine are aromatic and do not
257 affect TTX binding; Sunami et al. 2000), with the exception of the 401A observed in *P. cinereus* $Na_v1.1$.
258 This conservative variation likely contributes to the signal of diversifying selection acting on this codon
259 in less toxic salamander lineages. Site models also suggested positive selection acting on site 1533 in DIV
260 of $Na_v1.1$ and $Na_v1.2$ (table 4). Although the substitutions present at site 1533 in these newt paralogs have
261 not been tested experimentally for their effects on TTX binding, Maruta et al. (2008) showed that a
262 G1533T substitution at this site led to a moderate (~2-3-fold) decrease in TTX binding affinity, and
263 substitutions at this site are common in TTX-resistant channels (Geffeney et al. 2005; Jost et al. 2008;
264 Feldman et al. 2012; McGlothlin et al. 2016), suggesting that these non-synonymous changes likely also
265 reduce TTX binding affinity in newts.

266 Within the toxic newt lineage, we identified putative positive selection acting on three known
267 TTX-binding sites: 1240 and 1532 in DIII and DIV of muscle channel $Na_v1.4$ and 1529 in DIV of brain
268 channel $Na_v1.1$ (table 4). In $Na_v1.4$, sites 1240 and 1532 contain resistance-conferring substitutions
269 exclusively within toxic newts and these substitutions have been associated with extreme TTX resistance
270 in *Taricha* muscle fibers (Hanifin and Gilly 2015). In $Na_v1.1$, site 1529 encodes part of the Na^+ selectivity
271 filter (comprised of interacting amino acids DEKA – sites 400, 755, 1237, and 1529), which is highly
272 conserved across Na_v paralogs. An A1529G substitution (resulting in a DEKG filter) is present in $Na_v1.1$
273 of *A. mexicanum* and all members of the toxic newt clade. While this substitution does not appear to
274 affect Na^+ selectivity or to be sufficient in preventing TTX from binding (Jost et al. 2008), it may alter
275 channel firing properties, as this A1529G substitution resulted in substantially higher Na^+ currents relative
276 to the wild-type when introduced into a mammalian $Na_v1.4$ channel (Jost et al. 2008). The same alanine
277 to glycine substitution has been observed in Na_v channels of TTX-bearing flatworms (Jeziorski et al.
278 1997) and pufferfish (Jost et al. 2008), which suggests that it may play a role in TTX resistance in these
279 organisms.

280 Outside of TTX-binding regions, we found evidence for putative positive selection in similar
281 regions across multiple Na_v paralogs (tables S5, S6). For $Na_v1.4$, the majority of sites under positive
282 selection reside in terminal exon 26a that encodes the DIVa P-loop (exon 26b was excluded from this
283 analysis due to its absence in some species). For most other paralogs, the largest clusters of sites with

284 elevated d_N/d_S ratios were within the DI L5 turret (the extracellular loop upstream of the DI P-loop,
285 encoded by exons 6 and 7) and the DIII L5 turret upstream of the DIII P-loop encoded by exon 21.

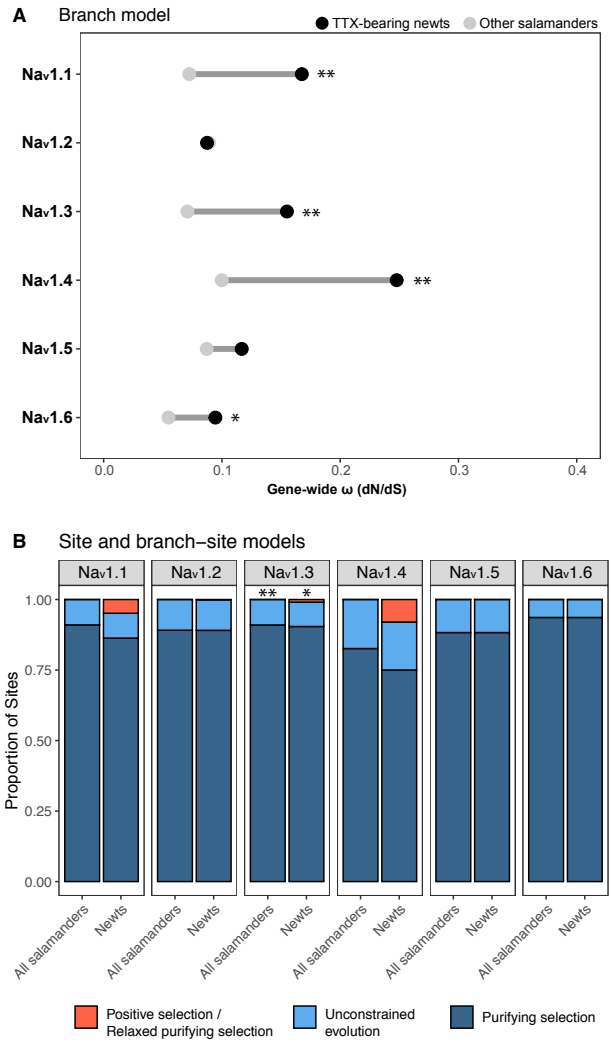

Fig. 4. Tests for shifts in selective pressure on salamander voltage-gated sodium channels. (A) PAML branch models comparing estimates of gene-wide ω (d_N/d_S) within TTX-bearing newts (black circles) and other salamanders (grey circles). (B) PAML site and branch-site model estimates of proportions of sites under purifying selection ($0 < \omega < 0.05$ in both lineages), unconstrained evolution ($\omega = 1$ in both lineages), and positive selection or relaxed purifying selection ($0 < \omega < 0.05$ in salamanders and $\omega \geq 1$ in TTX-bearing newts). Significant differences based on likelihood ratio tests are indicated with * ($p < 0.05$) and ** ($p < 0.01$). Although likelihood-ratio tests were non-significant for $Na_V1.1$ and $Na_V1.4$ (tables S3, S4), PAML identified a large proportion of amino acid coding sites within these paralogs with elevated d_N/d_S ratios in toxic newts.

Table 4. P-loop Sites with Elevated ω Values in All Salamanders and in Toxic Newts.

	$Na_V1.1$	$Na_V1.2$	$Na_V1.3$	$Na_V1.4$	$Na_V1.5$	$Na_V1.6$
Tissue expression	Brain	Brain	Brain	Muscle	Heart	Brain / PNS
P-loop sites under positive selection in all salamanders[b,d]	Y401A/C (DI)[a] G1533V (DIV)[a]	G1533R (DIV)[a]	Y401A (DI)[a]	-	739 (DII)	Y401A (DI)[a]
P-loop sites under positive selection in toxic newts[c,d]	1224 (DIII) A1529G (DIV)[a]	-	T759S (DII)	W756Y (DII) M1240T (DIII)[a] 1517 (DIV) 1519 (DIV) D1532S (DIV)*[,a]	-	-

* Empirical Bayes posterior probability > 0.95
[a] Non-conservative amino acid substitutions associated with TTX resistance in toxic newts
[b] Sites identified with site models using empirical Bayes method (posterior probability > 0.50)
[c] Sites identified with branch-site models using empirical Bayes method (posterior probability > 0.50)
[d] Because we had low power to detect positive selection at specific sites due to the small number of species in our study, these results should be interpreted with caution

286 These sites may facilitate interaction with other proteins, or alternatively, some of these sites identified by
287 the branch-site model may be selected to compensate for biochemical changes produced by TTX-resistant
288 mutations.
289
290 **Gene Conversion Events**
291 We also tested for evidence of nonallelic gene conversion as a mechanism of sequence evolution
292 contributing to adaptive evolution in Na_v genes using the program GENECONV. Nonallelic or ectopic
293 gene conversion results from an interlocus exchange of DNA that can occur between closely related
294 sequences during double-stranded break repair (Hansen et al. 2000). GENECONV uses the information in
295 a multiple sequence alignment to identify regions of similarity shared between two sequences that is
296 higher than expected by chance based on comparisons to permuted alignments (Sawyer 1999). We
297 selected this program to detect gene conversion because of its low false positive rates and robustness to
298 shifts in selective pressure (Posada and Crandall 2001; Bay and Bielawski 2011). Because nonallelic gene
299 conversion is more likely to occur between paralogs residing on the same chromosome (Semple and

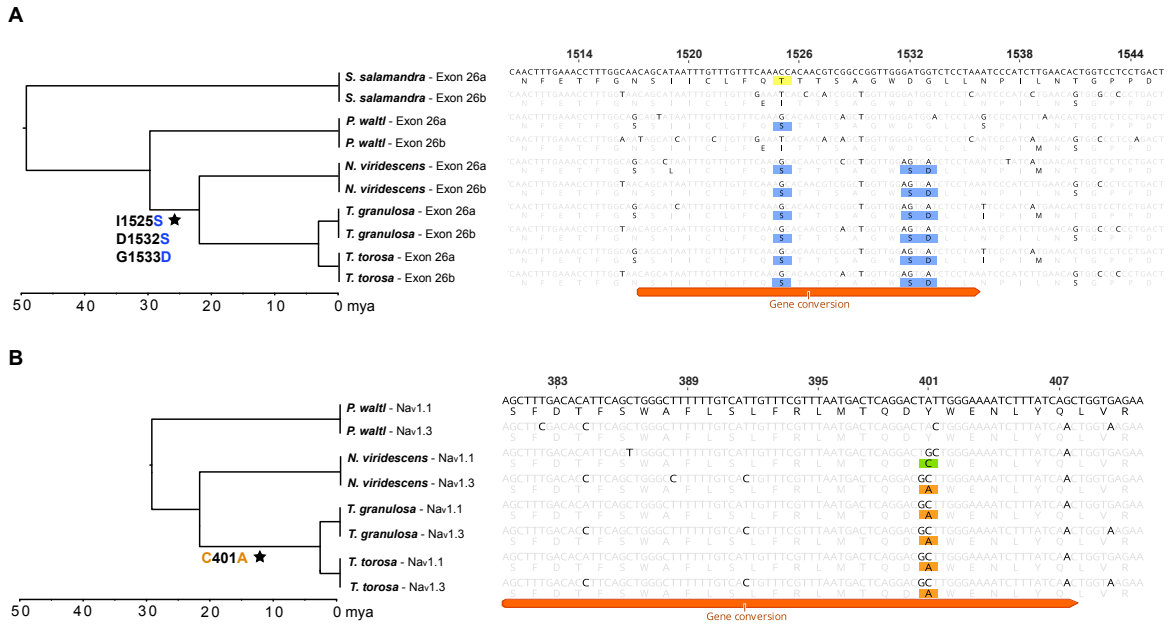

Fig. 5. Gene conversion within TTX binding regions. (A) Gene conversion was detected between DIVa and DIVb P-loops within duplicate exons 26a and 26b of $Na_v1.4$ in *T. torosa*. Substitutions conferring TTX resistance are highlighted in yellow and blue. While one moderately resistant substitution is present in DIVa of less toxic salamanders (1525T/S), three identical substitutions are present in both DIVa and DIVb of highly toxic newts. (B) Gene conversion was also detected between the DI P-loops of paralogs $Na_v1.1$ and $Na_v1.3$ in both *T. torosa* and *T. granulosa*. In *Notophthalmus*, both paralogs have different resistant substitutions (401C in $Na_v1.1$ – highlighted in green; 401A in $Na_v1.3$ – highlighted in orange). This putative gene conversion may have homogenized TTX-resistant substitutions among these paralogs within *Taricha* newts. Site numbers are in reference to amino acids of rat $Na_v1.4$ (accession number AAA41682). Stars at branches indicate the inferred transfer of substitutions associated with moderate resistance (blue text) or extreme resistance (orange text) via non-allelic gene conversion.

300 Wolfe 1999; Drouin 2002; Benovoy and Drouin 2009), we limited our search to events between the
301 tandem duplicates $Na_v1.1$, $Na_v1.2$, and $Na_v1.3$ and between exons 26a and 26b of $Na_v1.4$ within each
302 salamander genome. While few gene conversion events were detected within each species by
303 GENECONV, all of the regions that were detected within *Taricha* newts contain TTX-binding sites with
304 substitutions associated with TTX resistance, including the DIVa and DIVb P-loops of $Na_v1.4$ (fig. 5,
305 table 5). We observed three TTX resistant amino acids (I1525S, D1532S, G1533D) within the DIVa and

306 DIVb P-loops of $Na_v1.4$ in toxic newt genomes. In contrast, *P. waltl*, a closely related but non-
307 tetrodotoxic newt, contains one moderately TTX-resistant amino acid (I1525S) in the DIVa P-loop and no
308 resistant amino acids in the DIVb P-loop (fig. 5A). These differences involve four identical nucleotide
309 changes at homologous sites in both exon duplicates. Our short reads from the genomes of *N. viridescens*
310 and *Taricha* newts mapped onto each of these exon assemblies across putative recombination break points
311 with high (> 50-fold) coverage, lending support for sequence convergence rather than an assembly error.
312 We did not detect gene conversion between these exons within the genomes of *T. granulosa* or *N.*
313 *viridescens*; however, this may be due to the low power of GENECONV to detect conversion (Bay and
314 Bielawski 2011), particularly in the presence of low sequence diversity and when the conversion tract is

Table 5. Nonallelic Gene Conversion Events

$Na_v1.1$ – $Na_v1.2$ – $Na_v1.3$ Coding Sequences

Species	Gene 1 – Gene 2	BC p-value[a]	Protein begin[b]	Protein end[b]	P-loop	Exon	Length (bp)	Total poly (bp)[c]	Num diffs (bp)[d]	Total diffs (bp)[e]
A. mexicanum	$Na_v1.1$ - $Na_v1.2$	0.01	12	42	-	1	93	48	5	794
	$Na_v1.1$ - $Na_v1.3$	0.01	45	84	-	1	128	54	5	722
	$Na_v1.2$ - $Na_v1.3$	0.00	203	405	DI	5; 6; 7; 8 19; 20;	>626	74	2	902
	$Na_v1.1$ - $Na_v1.2$	0.00	1059	1193	-	21	>407	134	25	794
	$Na_v1.1$ - $Na_v1.3$	0.02	1104	1144	-	20	127	37	1	722
	$Na_v1.1$ - $Na_v1.2$[f]	**0.02**	**1208**	**1232**	**DIII**	**21**	**84**	**35**	**2**	**794**
C. alleganiensis	$Na_v1.1$ - $Na_v1.2$	0.00	1111	1232	DIII	20;21	>376	153	23	621
	$Na_v1.2$ - $Na_v1.3$[f]	**0.02**	**1478**	**1611**	**DIV**	**26**	**406**	**164**	**32**	**693**
	$Na_v1.1$ - $Na_v1.2$	0.00	1531	1652	DIV	26	371	177	27	621
N. viridescens	$Na_v1.1$ - $Na_v1.3$	0.00	500	512	-	10	89	41	1	733
	$Na_v1.1$ - $Na_v1.2$	0.01	1104	1199	-	16	292	109	20	796
P. cinereus	$Na_v1.1$ - $Na_v1.2$	0.00	12	32	-	1	64	31	0	838
	$Na_v1.1$ - $Na_v1.3$	0.00	362	499	DI	8; 9; 10	>576	88	13	786
	$Na_v1.2$ - $Na_v1.3$	0.02	573	599	-	13	83	29	2	897
	$Na_v1.1$ - $Na_v1.3$	0.00	610	667	-	14	177	68	8	786
	$Na_v1.1$ - $Na_v1.2$	0.00	1135	1305	DIII	20; 21; 22; 23; 24	>524	138	9	838
	$Na_v1.1$ - $Na_v1.3$	0.03	1430	1473	-	26	134	46	5	786
P. waltl	$Na_v1.1$ - $Na_v1.3$	0.03	1369	1386	-	25	56	30	0	738
T. granulosa	**$Na_v1.1$ - $Na_v1.3$[g]**	**0.01**	**374**	**408**	**DI**	**8**	**105**	**41**	**2**	**725**
T. torosa	**$Na_v1.1$ - $Na_v1.3$[g]**	**0.01**	**374**	**408**	**DI**	**8**	**105**	**41**	**2**	**725**

$Na_v1.4$ Exon 26

Species	Gene	BC p-value[a]	Protein begin[b]	Protein end[b]	P-loop	Exon	Length (bp)	Total poly (bp)[c]	Num diffs (bp)[d]	Total diffs (bp)[e]
P. cinereus	$Na_v1.4$	0.01	1509	1534	DIV	26a-26b	77	26	0	154
T. torosa	**$Na_v1.4$[g]**	**0.02**	**1518**	**1536**	**DIV**	**26a-26b**	**56**	**22**	**1**	**178**

[a] Global Bonferroni-corrected (BC) p-values from 10,000 simulations
[b] Protein begin and end sites are in reference to the rat $Na_v1.4$ channel (accession number AAA41682)
[c] Polymorphism within the converted region across all species
[d] Pairwise differences between the two paralogs within the converted region
[e] Total pairwise differences across the entire length of the two paralogs
[f] Putative gene conversion events associated with the loss of TTX resistance
[g] Putative gene conversion events associated with the gain of TTX resistance

315 shorter than ~100 bp (Posada and Crandall 2001; McGrath et al. 2009). Together, these results suggest
316 that the three resistant amino acids accumulated together in one exon copy followed by conversion of the
317 other exon ~30 mya in a toxic newt ancestor.
318 We also detected gene conversion between the DI P-loops of paralogs $Na_v1.1$ and $Na_v1.3$ within
319 the genomes of both *Taricha* species (fig. 5B, table 5). TTX-resistant substitutions are identical in the DI
320 P-loops of $Na_v1.1$ and $Na_v1.3$ in both *Taricha* species (401A, encoded by a GCT codon), while $Na_v1.1$ of
321 *N. viridescens* contains a different codon at this position (TGC, encoding 401C), consistent with gene
322 conversion occurring at this locus in *Taricha.* Both $Na_v1.1$ and $Na_v1.3$ of non-tetrodotoxic *P. waltl* newts
323 encode a TTX-sensitive tyrosine at this locus. It is unclear whether putative gene conversion event(s)
324 occurred before or after resistance evolved in both paralogs. Gene conversion may have converted a non-
325 resistant channel to a resistant channel in an ancestral *Taricha*, while *N. viridescens* independently
326 acquired a Y401A substitution, or it may have homogenized substitutions within two channels that had
327 previously evolved resistance in an ancestor of all toxic newts. In *Taricha* newts, the transition from
328 either tyrosine or cysteine to alanine required multiple nucleotide substitutions in both paralogs, making
329 gene conversion a likely explanation for the observed substitution patterns.
330 We detected additional gene conversion events that may have involved a non-resistant paralog
331 acting as a donor to a resistant paralog, leading to the loss of TTX resistance in *A. mexicanum* and *C.*
332 *alleganiensis* paralogs. A resistant substitution is present in the DIII P-loop of $Na_v1.1$ in most
333 salamanders but is absent in $Na_v1.1$ of *A. mexicanum* (fig. 3), and we detected gene conversion in an
334 adjacent region between $Na_v1.1$ and the non-resistant $Na_v1.2$ paralog within the *A. mexicanum* genome
335 (table 5, fig. S10A). Similarly, *C. alleganiensis* $Na_v1.2$ contains a non-resistant 1533G within DIV of all
336 six paralogs, while $Na_v1.2$ of *X. tropicalis* frogs encodes a putatively TTX-resistant 1533L, and gene
337 conversion between paralogs $Na_v1.2$ and $Na_v1.3$ may have facilitated the loss of this substitution in
338 salamanders (fig. S10B)
339

340 ## Discussion

341 Here we show that TTX-bearing newts have evolved resistance to their own toxicity through
342 multiple parallel changes in their Na_v genes and that, similar to snakes that consume TTX-containing prey
343 (McGlothlin et al. 2016; Perry et al. 2018), some of the resistance in this taxon is ancient, first appearing
344 in an early salamander. However, while substitutions conferring moderate TTX resistance are present in
345 some non-toxic salamander genomes, only the TTX-bearing newts have substitutions conferring high
346 resistance across all six of their Na_v paralogs, and several of these channels harbor multiple resistant
347 substitutions in more than one domain. Many of the substitutions conferring resistance to toxic newts are
348 also present in Na_v paralogs of TTX-resistant pufferfish (Jost et al. 2008) and snakes (McGlothlin et al.

349 2016). Similar to pufferfish, newts appear to require resistance within all of their brain/nerve channels in
350 addition to their hearts and muscles. This feature apparently distinguishes toxic prey from their predators,
351 whose brain channels lack resistant substitutions (McGlothlin et al. 2014). This molecular parallelism
352 emphasizes the strong structural and functional constraints on this gene family, which appear to limit the
353 evolution of TTX resistance to a small number of predictable pathways, leading to convergent and
354 parallel changes across multiple taxa (Feldman et al. 2012). We show that the evolution of extreme TTX
355 resistance is accompanied by a shift in signatures of selection in four out of six paralogs, with suggestive
356 evidence of positive selection acting directly on TTX-binding sites. Finally, the evolution of physiological
357 resistance appears to have been facilitated by at least two instances of non-allelic gene conversion, which
358 acted to introduce TTX-resistance substitutions from one paralog (or exon duplicate) to another,
359 providing a rare example of the facilitation of adaptive evolution via non-allelic gene conversion.

360 Our reconstruction of the history of TTX resistance in salamanders reveals the ancient origins of
361 moderate resistance in nerve channels $Na_v1.1$ and $Na_v1.6$ and high resistance in the heart channel $Na_v1.5$,
362 which arose ~160 mya. Resistant substitutions in the muscle channel $Na_v1.4$ evolved later, becoming
363 fixed in the clade including all newts and *S. salamandra* between 75–130 mya. This was followed by the
364 accumulation of additional substitutions in DIII, DIVa, and DIVb of $Na_v1.4$ within members of the highly
365 toxic newt clade, providing their muscles with resistance to much higher concentrations of TTX relative
366 to salamanders lacking these mutations (Hanifin and Gilly 2015). Substitutions in DI conferring extreme
367 TTX resistance to the brain/nerve channels $Na_v1.1$, $Na_v1.2$, $Na_v1.3$, and $Na_v1.6$ also evolved more
368 recently (~30 mya) and they are limited to the toxic newt clade with only one exception, $Na_v1.1$ of *P.*
369 *cinereus*, which may have arisen independently in this lineage. Toxic newts also have unique
370 substitutions in DIV of $Na_v1.1$ and $Na_v1.2$, which may provide additional resistance to their brain
371 channels or may compensate for structural or functional changes resulting from resistant substitutions in
372 DI or in other regions of the protein. The widespread presence of TTX in modern newts suggests that
373 possession of TTX evolved in the common ancestor of this clade (Hanifin 2010). This is supported by our
374 observation of shared TTX-resistance substitutions across all Na_v paralogs in highly toxic newts.
375 However, the highly toxic newts included in this study are limited to North American species and do not
376 include Asian or European newts (e.g. *Cynops* and *Triturus* spp.) known to have TTX (Hanifin 2010).
377 Sequencing the Na_v genes of these species will reveal if extreme resistance in the nerve channels is
378 specific to North American newts or arose earlier in an ancestral newt species.

379 Whether the TTX-resistance substitutions observed in newt relatives are adaptive and what
380 selective pressures act to retain them remains to be determined. In ancestral salamandrids, resistance in
381 three or four of the six Na_v proteins may have provided tolerance to low levels of TTX, facilitating the
382 evolution of extreme toxicity in the modern newts, although it is unclear whether these substitutions arose

383 in response to TTX exposure or as a side effect of selection for another aspect of channel function.
384 Environmental exposure to TTX could potentially occur from TTX-bearing prey, such as terrestrial
385 *Bipalium* flatworms (Stokes et al. 2014) or from TTX-producing bacteria (Vaelli et al. 2020). In a feeding
386 study on *B. adventitium* flatworms, Ducey et al. (1999) demonstrated that while all salamanders rejected
387 the flatworms when they were first presented, some habituated *Ambystoma* and *Plethodon* individuals
388 were able to consume them with only minor symptoms of TTX poisoning, including apparent mucus
389 production and numbing of the mouth. Although the amount of TTX present in the worms was not
390 measured, this observation supports the conjecture that these substitutions play a protective role against
391 consuming toxic prey.

392 Our power to detect sites undergoing positive selection was low due to the small number of
393 species available for our analyses and the high sequence conservation among orthologs (Yang et al. 2000;
394 Anisimova et al. 2001). However, our analyses provided evidence consistent with positive selection
395 acting on TTX-binding sites, including sites implicated in the extreme TTX resistance of salamander
396 muscle fibers. Our branch models indicate that several Na_v paralogs have higher ω ratios within the
397 tetrodotoxic newt clade relative to other salamanders. This increase in ω ratio appears to be mainly the
398 result of relaxed purifying selection rather than positive selection, as few sites were estimated to have $\omega >$
399 1 by branch-site models. Muscle channel $Na_v1.4$ had the highest ω ratio in toxic newts, coinciding with a
400 relatively high proportion (~0.06) of sites detected as undergoing positive selection in toxic newts and
401 purifying selection in other salamanders, although the likelihood ratio test was non-significant (table S4).
402 This pattern may derive from ongoing positive selection on $Na_v1.4$ resulting from the coevolutionary
403 arms race between newts and snakes. Increased ω ratios in toxic newts coincide with the appearance of
404 highly TTX-resistant substitutions within the DI P-loops of $Na_v1.1$, $Na_v1.3$, and $Na_v1.6$, DIV of $Na_v1.1$,
405 and the DIII, DIVa, and DIVb P-loops in $Na_v1.4$. While the pattern of these substitutions suggests that
406 they are adaptive changes that occurred specifically within the tetrodotoxic newt clade, our branch-site
407 models detected positive selection on TTX-binding sites only within the DIV P-loop of $Na_v1.1$ and the
408 DIII and DIVa P-loops of $Na_v1.4$, while our site models detected positive selection on site 401 within
409 paralogs $Na_v1.1$, $Na_v1.3$, and $Na_v1.6$ of all salamanders. This may be due to the statistically conservative
410 nature of the branch-site model and the tendency of PAML to detect ongoing diversifying selection rather
411 than rare positive selection events (Yang and dos Reis 2010). Furthermore, the site and branch-site
412 models do not distinguish between biochemically conservative and non-conservative amino acid changes.
413 Nevertheless, the detection of relatively high ω ratios within the toxic newt clade along with site-specific
414 positive selection at known TTX-binding sites provides strong evidence that these substitutions are
415 adaptive.

416 In addition to the six Na_v paralogs described in amphibians (Zakon 2012), our sequence data
417 revealed the presence of a partial duplication of the $Na_v1.4$ gene in salamanders that includes the entirety
418 of exon 26, encoding the DIV P-loop, which likely occurred in an ancestor of Salamandroidea. The
419 maintained open reading frame and shared expression patterns of transcripts encoding exons 26a and 26b
420 in *A. mexicanum* tissues suggests that this duplicate region is functional in salamander muscles. In some
421 insects that feed on toxic plants, the appearance of resistance-conferring substitutions is accompanied by
422 one or more duplications of the genes that the toxin targets (Zhen et al. 2012; Petschenka et al. 2017).
423 Petschenka et al. (2017) show that in at least one species, such gene duplication precedes resistance, and
424 they suggest that gene duplication may help to alleviate the potential decrease in fitness incurred by the
425 insects due to the negative pleiotropic effects of toxin-resistant mutations. Similarly, resistance in DIV of
426 salamander $Na_v1.4$ appears after the duplication of this domain, and only the TTX-bearing newts have
427 resistant substitutions in both exon copies, lending support to this hypothesis and raising the possibility
428 that the evolution of physiological resistance in salamanders may have been mediated in part by this
429 genomic novelty.

430 We observed a rare case of the generation of adaptive variants via non-allelic gene conversion,
431 which occurred both between the duplicated exons of $Na_v1.4$ and between paralogs $Na_v1.1$ and $Na_v1.3$.
432 Gene conversion is often thought to play a role in constraint, preserving the core functions of gene
433 families (Chapman et al. 2006) and reducing deleterious mutation loads (Ohta 1989; Khakhlova and Bock
434 2006), and has also been associated with the diversification of major histocompatibility complex genes in
435 mammals (Kuhner et al. 1991; Go et al. 2003) and the introduction of deleterious nonsynonymous
436 mutations into different parts of the genome (Galtier et al. 2009; Casola et al. 2012). However, the
437 potential for gene conversion to facilitate adaptation is less widely appreciated. Theory suggests that gene
438 duplication and subsequent gene conversion may allow for movement between adaptive peaks via the
439 accumulation of beneficial mutations in one gene copy that can be transferred to a favorable genetic
440 background (Hansen et al. 2000). This process has been implicated in the adaptive evolution of hypoxia
441 tolerance in high-altitude Tibetan wolves (Signore et al. 2019) and in heavy metal tolerance in
442 *Arabidopsis* (Hanikenne et al. 2013). We observed a rare case of the generation of adaptive variants via
443 non-allelic gene conversion, which occurred both between the duplicated exons of $Na_v1.4$ and between
444 paralogs $Na_v1.1$ and $Na_v1.3$. In the brain channel genes $Na_v1.1$ and $Na_v1.3$, this likely occurred between
445 two genes that had previously evolved resistance, homogenizing the substitutions between the two
446 paralogs in *Taricha* newts. In the muscle channel gene $Na_v1.4$, three identical amino acids were present
447 within both duplicated DIV TTX-binding domains in toxic newts, while only one resistant amino acid
448 was present in a single DIV domain in their non-tetrodotoxic salamander relatives, suggesting that
449 resistant substitutions accumulated in one exon copy in a toxic newt ancestor and subsequently spread to

450 the other exon copy via gene conversion. Both Hanifin and Gilly (2015) and Du et al. (2009) have shown
451 that the single resistant amino acid observed in less toxic salamanders confers only low levels of TTX
452 resistance, while the combination of the three amino acids found in highly toxic newts confers extreme
453 resistance. As splice variants encoding the alternative exons appear to share similar expression patterns in
454 salamanders, both exon copies should require extreme resistance in species exposed to high TTX
455 concentrations. The concerted evolution of TTX resistance among homologous P-loop domains of $Na_v1.4$
456 may have expedited the evolution of extreme resistance in newt muscles, requiring new resistant
457 mutations to appear in only one of these domains before being transferred to the other copy.
458 The degree of parallel molecular evolution among members of the Na_v gene family and across
459 lineages provides insight into the constraints on Na_v nucleotide sequence as well as the evolvability of the
460 TTX-resistance phenotype. Our results reveal that similar to their coevolutionary partners, *Thamnophis*
461 garter snakes, *Taricha* newts evolved extreme TTX resistance through a stepwise process that built upon
462 ancient changes that were in place millions of years before the arms race began. However, the pattern of
463 TTX resistance evolution in newts also displays important differences from that of their predators. First,
464 perhaps because of the constitutive presence of TTX, newts display extreme levels of resistance even in
465 channels that are expressed in the central nervous system, which are protected by the blood-brain barrier
466 in species that encounter TTX in their diet. Second, our analysis indicates that many substitutions may
467 have become fixed relatively close to one another in evolutionary time within the clade of modern newts.
468 This is in contrast to snakes, where key changes were separated by millions of years. Due to our lack of
469 sampling of newt species outside of North America, however, further work is necessary to understand the
470 timing of these changes on a finer scale. We also show that while positive selection appears to be a strong
471 driving force of the evolution of TTX auto-resistance in newts, gene conversion may have sped up the
472 process of adaptive evolution in some Na_v paralogs, and constraints have limited the possible locations
473 and types of resistant substitutions to a small subset of realized genetic changes. Taken together, our
474 results emphasize the interplay among selection, constraint, and historical contingency in the evolution of
475 complex adaptations.
476

477 # Methods

478 ## Sequencing and Annotation of Voltage-Gated Sodium Channel Paralogs

479 We identified Na_v genes in the two publicly available salamander genome assemblies, *A. mexicanum*
480 (Smith et al. 2019; AmexG.v6 assembly) and *P. waltl* (Elewa et al. 2017), and in one full-body
481 transcriptome from the fire salamander *S. salamandra* (Goedbloed et al. 2017; BioProject accession
482 PRJNA607429) using the reciprocal best BLAST hit method (Moreno-Hagelsieb and Latimer 2007) with
483 queries of Na_v sequences from *X. tropicalis* (Hellsten et al. 2010) and salamanders (Hanifin and Gilly

484 2015). We confirmed assignments of each amphibian Na_v paralog based on nucleotide alignments of the
485 coding sequences with *Xenopus* sequences, as well as synteny of the chromosomal segments containing
486 Na_v genes (fig. S1). These Na_v annotations were then used to design targeted sequencing probes and to
487 subsequently assign paralog identity to our de novo salamander assemblies. We used Geneious v10.2.3
488 for sequence visualization and to create DNA and protein alignments (Kearse et al. 2012). We also
489 performed BLASTn searches of published transcriptome assemblies of *Tylototriton wenxianensis*
490 (PRJNA323392), *Bolitoglossa vallecula* (PRJNA419601), and *Hynobius retardatus* (Matsunami et al.
491 2015; PRJDB2409). However, we were unable to identify the full set of Na_v paralogs within these three
492 assemblies, possibly due to their tissue specificity. Therefore, we used the sequences for probe design and
493 to identify exon duplications in $Na_v1.4$ but excluded them from PAML analyses.
494 In order to design targeted sequencing probes, we compiled partial and complete amphibian Na_v
495 sequences obtained from NCBI databases and additional published sources (table S7) into a single
496 FASTA file. Each individual FASTA entry included a single exon and, when available, up to 200 bp of
497 intron sequence upstream and downstream of each exon to aid in paralog assignment. Using
498 RepeatMasker (Smit et al. 2013-2015), we replaced transposable elements and other sequence repeats
499 with Ns, and subsequently filtered out sequences < 120 bp in length, as well as those with more than 25%
500 missing or ambiguous characters. We submitted this masked and filtered file, which was 465 kb in total
501 length, to Agilent Technologies (Santa Clara, CA, USA) for custom probe design using the SureSelect
502 tool, resulting in 7518 unique 120-mer probes. Probe sequences are available in our Dryad repository.
503 We obtained DNA samples from adult individuals of three TTX-bearing species (n=3 for each
504 species): *T. torosa* from Hopland, CA, *T. granulosa* from Benton, OR, and *N. viridescens* from Mountain
505 Lake, VA, and from two additional salamander species presumed to lack TTX (n = 2 for each species): *P.*
506 *cinereus* collected in Mountain Lake, VA and *C. alleganiensis* collected in southwestern VA. We
507 extracted genomic DNA using the DNeasy Blood & Tissue kit (Qiagen Inc., Valencia, CA) and prepared
508 sequencing libraries using the SureSelect^{XT} Target Enrichment system for Illumina paired-end
509 multiplexed sequencing from Agilent Technologies (Santa Clara, CA, USA), following the protocol for
510 low input (200 ng) DNA samples. We used a Covaris M220 Focused-ultrasonicator to shear 200 ng of
511 purified whole genomic DNA from each sample into ~250 bp fragments using the following settings:
512 duty factor 10%, peak incident power 75 w, 200 cycles per burst, and treatment time of 160 seconds. We
513 followed the Agilent SureSelect^{XT} Target Enrichment kit protocol for end repair, adaptor ligation,
514 amplification, hybridization and bead capture (using the custom SureSelect probes described above),
515 indexing, and purification. We quantified the resulting enriched, indexed libraries with qPCR and
516 combined them in equimolar concentrations into one final library pool, which was submitted to the
517 Genomics Sequencing Center at Virginia Tech for sequencing on an Illumina MiSeq 300-cycle v2 with

518 150 bp paired-end reads. Prior to alignment, we trimmed Illumina reads of TruSeq3 adapter sequences,
519 removed bases with a phred64 quality score less than 3, and filtered out subsequent reads shorter than 100
520 bp using Trimmomatic version 0.33 (Bolger et al. 2014). We used SPAdes 3.6.0 (Bankevich et al. 2012)
521 to create de novo assemblies of the trimmed and filtered reads. Each Na_v paralog had > 10-fold sequence
522 coverage, with an average of 32-fold coverage across all species and paralogs (table S1). After filtering,
523 all reads had Phred Q-scores >30 for >99% of sites.

524 Because we designed the sequencing probes to capture only small portions of the Na_v intronic
525 regions flanking exons (regions more likely to be conserved across species), each exon was assembled
526 into a separate scaffold. For each individual from our sequencing trial, we created BLAST databases from
527 the assembled de novo scaffolds and performed BLAST searches using single exons from the *Ambystoma*
528 Na_v genes as queries. We created 26 separate nucleotide alignments, one for each individual Na_v exon,
529 including sequences from *Ambystoma*, *Pleurodeles*, and our de novo assemblies using MAFFT v 7.450
530 (Katoh and Standley 2013) and created consensus neighbor-joining trees with a Tamura-Nei genetic
531 distance model using the Geneious Tree Builder (Kearse et al. 2012) with 1000 bootstrap replicates and
532 an 80% support threshold. The resulting tree topologies were used to assign paralog identity to each of the
533 exons. When necessary, we included exon and intron sequences from additional species in the alignments
534 to resolve the topology of the trees.

535 We concatenated all exons from each paralog into full coding sequences. Based on alignment
536 with full-length *Xenopus* sequences, all salamander Na_v coding sequences collected for this study were
537 >90% complete with the exception of two sequences from the *Salamandra salamandra* transcriptome
538 (paralogs $Na_v1.1$ and $Na_v1.2$, which were 68.8% and 70.7% complete respectively; table S1). For each
539 species, we used Na_v paralogs sequenced from the genome of a single individual for our downstream
540 analyses, based on completeness of the assembly.

541 **Determination of TTX Resistance Levels**

542 TTX sensitivity is commonly measured *in vitro* by using site-directed mutagenesis to introduce mutations
543 of interest into a TTX-sensitive Na_v channel, followed by expression in *Xenopus* oocyte or HEK 293 cells
544 and the application of patch-clamp whole-cell recordings to measure channel current in the presence of
545 TTX. The fold-change in TTX sensitivity is then calculated by taking the ratio of the IC_{50} values, or the
546 TTX concentration at which 50% of the Na_v channels are blocked, of mutated and wild-type channels (see
547 table 3 for references).

548 Another line of evidence for resistance in salamander muscle channels ($Na_v1.4$) comes from
549 Hanifin and Gilly (2015), who estimated TTX sensitivity by recording action potentials generated from
550 salamander muscle fibers and estimating the amount of TTX required to diminish the rise of the action
551 potential, associating these relative changes with the presence and absence of substitutions in TTX-

552 binding sites. They associated moderate TTX resistance (reduced sensitivity to 0.010 μM TTX) with the
553 presence of DIII substitution M1240T and extreme resistance (low sensitivity to 300 μM TTX) with the
554 presence of DIII and DIV substitutions M1240T, D1532S, and G1533D. We categorize levels of TTX
555 resistance conferred by Na_v substitutions as extreme or moderate based on the results of Hanifin and Gilly
556 (2015) as well as the data summarized in table 3.

557 **Phylogenetic Analyses and Identification of Site-Specific Evolutionary Rates**

558 We constructed phylogenetic trees for the entire Na_v gene family using our *de novo* assembled sequences
559 as well as sequences from the genomes of *A. mexicanum*, *P. waltl*, the whole-body transcriptome of *S.*
560 *salamandra*, two frog genomes: *X. tropicalis* (Hellsten et al. 2010) and *Nanorana parkeri* (Sun et al.
561 2015), and one fish genome: *Danio rerio* (Howe et al. 2013). We created an amino acid alignment of
562 translated coding sequences using MAFFT v 7.450 (Katoh and Standley 2013) and constructed maximum
563 likelihood trees from these alignments with RAxML v8.2.11 (Stamatakis 2014) in Geneious using a
564 GAMMA BLOSSUM62 protein model and estimated clade support with 100 bootstrap replicates. In
565 order to improve the accuracy of the nucleotide alignment, we used this amino acid alignments to guide
566 codon alignments with PAL2NAL v14 (Suyama et al. 2006). We identified the best fitting substitution
567 models for the nucleotide alignment using jModelTest 2.1.10 v20160303 (Darriba et al. 2012) and
568 constructed maximum likelihood trees of the coding sequences with RAxML v8.2.11. The two models
569 with the lowest AIC scores were (1) the transition model with unequal base frequencies, a gamma shape
570 parameter, and some proportion of invariable sites (TIM2+I+G) and (2) a general time-reversible model
571 with a gamma shape parameter and a proportion of invariable sites (GTR+I+G), which is nearly identical
572 to TIM2+I+G, but includes two additional rate parameters (Posada 2008). Because the two models are
573 nearly equivalent and both fit our data equally well, we chose the GTR+I+G model for its ease of
574 implementation across different programs. We repeated these methods for each individual Na_v paralog for
575 downstream analyses in PAML, excluding the sequences from the outgroup species *N. parkeri* and *D.*
576 *rerio*. We estimated chronograms in Figs. 3 and 5 using the chronos( ) function of the ape v5.4-1 package
577 in R using a phylogenetic tree from Pyron (2014) and divergence dates from Hime et al. (2021; see Dryad
578 repository for R code).

579 We estimated synonymous (d_S) and nonsynonymous (d_N) substitution rates, as well as the d_N/d_S
580 ratios (ω) for each paralog using codeml in PAML v4.8 (Yang 2007). In order to test for changes in
581 selective regimes in the Na_v genes between salamanders and highly toxic newts (*Notophthalmus* and
582 *Taricha* species), we fit the following models to our Na_v alignments: (1) one-ratio models (allowing for a
583 single ω ratio among all sites and all branches of the phylogeny), (2) branch models (allowing for separate
584 ω ratios for the foreground [toxic newts] and background [other salamanders]), (3) branch-site neutral
585 models (allowing ω to vary both between toxic newts and other salamanders and among sites, with two

586 possible categories: $0 < \omega < 1$ and $\omega = 1$), and (4) branch-site models (allowing ω to vary both between
587 toxic newts and other salamanders and among sites, with three possible categories: $0 < \omega < 1$, $\omega = 1$, and
588 $\omega > 1$, the latter being allowed only within toxic newts). To test for site-specific positive selection among
589 all salamanders, we fit two sets of nested models: (1) discrete ω ratio models M1a vs. M2a, which
590 allowing for ω to vary among sites, with either two possible categories: $0 < \omega < 1$ and $\omega = 1$ or three
591 possible categories: $0 < \omega < 1$, $\omega = 1$, and $\omega > 1$; and (2) continuous ω ratio models M7, M8a, and M8,
592 which fit ω ratios of sites into a beta distribution, with ncatG (the number of ω values in the beta
593 distribution) set to 5. We used F3x4 codon models, which estimate individual nucleotide frequencies for
594 each of the three codon positions, and allowed codeml to estimate ω and transition-transversion rates (κ).
595 The outputs of these models were used to estimate and compare gene-wide ω between toxic newts and
596 salamanders (one-ratio and branch models), to identify sites under positive selection in all salamanders
597 (site and neutral site models), and to identify sites under positive selection or with elevated ω in toxic
598 newts relative to other salamanders (branch-site neutral and branch-site models). We also created
599 ancestral sequence reconstructions by specifying RateAncestor = 1 within codeml configuration files. We
600 report posterior probability estimates for ancestral sequence reconstructions using models with the highest
601 likelihood scores (Table S3), which were the neutral site models (M8a) for all Na_v paralogs except
602 $Na_v1.3$, for which the site selection model M8 had the most support (Fig. S4-S9). We performed all of the
603 above analyses both including and excluding the two *S. salamandra* paralogs with a large number of gaps
604 ($Na_v1.1$ and $Na_v1.2$). Maximum likelihood parameter estimates were largely congruent using these
605 different datasets. Therefore, we present results from the gene trees including *S. salamandra* here.

606 **Detection of Gene Conversion**

607 We used the program GENECONV to detect potential nonallelic gene conversion events between Na_v
608 paralogs. We used a full codon alignment of all six Na_v paralogs from 7 of the 8 salamanders included in
609 the study (excluding *S. salamandra* due to missing data), but targeted our search to include only genes
610 $Na_v1.1$, $Na_v1.2$, and $Na_v1.3$ within each species under the assumption that gene conversion is more likely
611 to occur between closely related sequences that reside on the same chromosome. This increased the power
612 of detecting gene conversion from multiple pairwise comparisons. We also performed a separate search
613 for gene conversion events between duplicate exons 26a and 26b of $Na_v1.4$. For this analysis, we used a
614 codon alignment of $Na_v1.4$ exons 26a and 26b from 10 salamander species. We assigned a mismatch
615 penalty using gscale=1 and used corrected *p*-values from 10,000 permutations to determine significance.

616 **Acknowledgments**
617

618 We thank Miguel Vences for helpful feedback on the manuscript and for providing an early draft of the *Salamandra*
619 *salamandra* transcriptome. We also thank Vincent Farallo, Edmund Brodie III, William Hopkins, and Brian Case for

620 providing DNA samples for this study, Kaitlyn Malewicz, Mercedes Collins, and Emily Orr for assistance with lab
621 work, John Abramyan, Edmund Brodie III, Charles Hanifin, and Andrew Kern for helpful conversations and
622 Matthew Hahn for advice on analyses and comments on the manuscript. Hellbender (*Cryptobranchus*) samples were
623 collected by Brian Case and William Hopkins in cooperation with the Virginia Department of Wildlife Resources
624 (DWR) and the U.S. Forest Service, and with assistance from J.D. Kleopfer. Hellbenders were handled under
625 scientific collecting permit #060465 and approved Virginia Tech IACUC procedures (protocol #16-162). We thank
626 the Departments of Fish and Wildlife in Oregon and California and the Virginia DWR for scientific collecting
627 permits to MTJH (063-18, SC-11937 and 048104 respectively). This work was supported by awards from the
628 National Science Foundation to JWM (DEB-1457463 and IOS-1755055), and to MTJH (DEB-1601296), by an R.C.
629 Lewontin award from the Society for the Study of Evolution to KLG, and by a graduate fellowship award from the
630 Interfaces of Global Change program at Virginia Tech to KLG.

631

632 **Data Availability**

633 Newly generated sequences have been submitted to the SRA database in NCBI under BioProject accession
634 PRJNA732671. Sequence alignments and configuration files for GENECONV analyses have been deposited at
635 https://github.com/kerrygendreau/Newt-TTX-Resistance.git and archived in a Dryad repository at
636 http://doi.org/10.5061/dryad.kprr4xh4v.
637

638 **References**

639

640 Anisimova M, Bielawski JP, Yang Z. 2001. Accuracy and power of the likelihood ratio test in detecting
641 adaptive molecular evolution. *Mol Biol Evol* 18(8):1585-1592.

642 Arbuckle K, Rodríguez de la Vega RC, Casewell NR. 2017. Coevolution takes the sting out of it:
643 Evolutionary biology and mechanisms of toxin resistance in animals. *Toxicon* 140:118-131.

644 Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI,
645 Pham S, Prjbelski AD, et al. 2012. SPAdes: A new genome assembly algorithm and its applications to
646 single-cell sequencing. *J Comput Biol* 19(5):455-477.

647 Bay RA, Bielawski JP. 2011. Recombination detection under evolutionary scenarios relevant to
648 functional divergence. *J Mol Evol* 73(5-6):273-286.

649 Benovoy D, Drouin G. 2009. Ectopic gene conversions in the human genome. *Genomics* 93(1):27-32.

650 Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data.
651 *Bioinformatics* 30(15):2114-2120.

652 Bricelj VM, Connell L, Konoki K, MacQuarrie SP, Scheuer T, Catterall WA, Trainer VL. 2005. Sodium
653 channel mutation leading to saxitoxin resistance in clams increases risk of PSP. *Nature* 434(7034):763-
654 767.

655 Brodie ED, III, Brodie ED, Jr. 1990. Tetrodotoxin resistance in garter snakes: an evolutionary response of
656 predators to dangerous prey. *Evolution* 44(3):651-659.

657 Brodie ED, Jr., Ridenhour BJ, Brodie ED, III. 2002. The evolutionary response of predators to dangerous
658 prey: hotspots and coldspots in the geographic mosaic of coevolution between garter snakes and newts.
659 *Evolution* 56(10):2067-2082.

660 Bryant DM, Johnson K, DiTommaso T, Tickle T, Couger MB, Payzin-Dogru D, Lee TJ, Leigh ND, Kuo
661 T-H, Davis FG, et al. 2017. A tissue-mapped axolotl de novo transcriptome enables identification of limb
662 regeneration factors. *Cell Rep* 18(3):762-776.

663 Caballero-Pérez J, Espinal-Centeno A, Falcon F, García-Ortega LF, Curiel-Quesada E, Cruz-Hernández
664 A, Bako L, Chen X, Martínez O, Alberto Arteaga-Vázquez M, et al. 2018. Transcriptional landscapes of
665 Axolotl (*Ambystoma mexicanum)*. *Dev Biol* 433(2):227-239.

666 Casola C, Hahn MW. 2009. Gene conversion among paralogs results in moderate false detection of
667 positive selection using likelihood methods. *J Mol Evol* 68(6):679-687.

668 Casola C, Zekonyte U, Phillips AD, Cooper DN, Hahn MW. 2012. Interlocus gene conversion events
669 introduce deleterious mutations into at least 1% of human genes associated with inherited disease.
670 *Genome Research* 22(3):429-435.

671 Chapman BA, Bowers JE, Feltus FA, Paterson AH. 2006. Buffering of crucial functions by paleologous
672 duplicated genes may contribute cyclicality to angiosperm genome duplication. *P Natl Acad Sci USA*
673 103(8):2730.

674 Chen J-M, Cooper DN, Chuzhanova N, Férec C, Patrinos GP. 2007. Gene conversion: mechanisms,
675 evolution and human disease. 8(10):762-775.

676 Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and
677 parallel computing. *Nat Methods* 9(8):772-772.

678 Drouin G. 2002. Characterization of the gene conversions between the multigene family members of the
679 yeast genome. *J Mol Evol* 55(1):14-23.

680 Du Y, Nomura Y, Liu Z, Huang ZY, Dong K. 2009. Functional expression of an arachnid sodium channel
681 reveals residues responsible for tetrodotoxin resistance in invertebrate sodium channels. *Journal of*
682 *Biological Chemistry* 284(49):33869-33875.

683 Ducey PK, Messere M, Lapoint K, Noce S. 1999. Lumbricid prey and potential herpetofaunal predators
684 of the invading terrestrial flatworm Bipalium adventitium (Turbellaria: Tricladida: Terricola). *Am Midl*
685 *Nat* 141(2):305-314.

686 Elewa A, Wang H, Talavera-López C, Joven A, Brito G, Kumar A, Hameed LS, Penrad-Mobayed M,
687 Yao Z, Zamani N, et al. 2017. Reading and editing the Pleurodeles waltl genome reveals novel features of
688 tetrapod regeneration. *Nat Commun* 8(1):2286.

689 Feldman CR, Brodie ED, Jr., Brodie ED, III, Pfrender ME. 2012. Constraint shapes convergence in
690 tetrodotoxin-resistant sodium channels of snakes. *Proc Natl Acad Sci USA* 109(12):4556-4561.

691 Fux JE, Mehta A, Moffat J, Spafford JD. 2018. Eukaryotic voltage-gated sodium channels: on their
692 origins, asymmetries, losses, diversification and adaptations. *Front Physiol* 9:1406.

693 Galtier N, Duret L, Glémin S, Ranwez V. 2009. GC-biased gene conversion promotes the fixation of
694 deleterious amino acid changes in primates. *Trends in Genetics* 25(1):1-5.

695 Geffeney SL, Brodie ED, Jr., Ruben PC, Brodie ED, III. 2002. Mechanisms of adaptation in a predator-
696 prey arms race: TTX-resistant sodium channels. *Science* 297(5585):1336-1339.

697 Geffeney SL, Fujimoto E, Brodie ED, III, Brodie ED, Jr., Ruben PC. 2005. Evolutionary diversification
698 of TTX-resistant sodium channels in a predator-prey interaction. *Nature* 434(7034):759-763.

699 Go Y, Satta Y, Kawamoto Y, Rakotoarisoa G, Randrianjafy A, Koyama N, Hirai H. 2003. Frequent
700 segmental sequence exchanges and rapid gene duplication characterize the MHC class I genes in lemurs.
701 *Immunogenetics* 55(7):450-461.

702 Goedbloed DJ, Czypionka T, Altmüller J, Rodriguez A, Küpfer E, Segev O, Blaustein L, Templeton AR,
703 Nolte AW, Steinfartz S. 2017. Parallel habitat acclimatization is realized by the expression of different
704 genes in two closely related salamander species (genus *Salamandra*). *Heredity* 119(6):429-437.

705 Graham AM, McCracken KG. 2019. Convergent evolution on the hypoxia-inducible factor (HIF)
706 pathway genes EGLN1 and EPAS1 in high-altitude ducks. *Heredity* 122(6):819-832.

707 Hague MTJ, Stokes AN, Feldman CR, Brodie Jr. ED, Brodie III ED. 2020. The geographic mosaic of
708 arms race coevolution is closely matched to prey population structure. *Evolution Letters* 4(4):317-332.

709 Hanifin CT. 2010. The chemical and evolutionary ecology of tetrodotoxin (TTX) toxicity in terrestrial
710 vertebrates. *Marine Drugs* 8(3):577-593.

711 Hanifin CT, Gilly WF. 2015. Evolutionary history of a complex adaptation: tetrodotoxin resistance in
712 salamanders. *Evolution* 69(1):232-244.

713 Hanikenne M, Kroymann J, Trampczynska A, Bernal M, Motte P, Clemens S, Krämer U. 2013. Hard
714 selective sweep and ectopic gene conversion in a gene cluster affording environmental adaptation. *PLOS*
715 *Genet* 9(8):e1003707.

716 Hansen TF, Carter AJR, Chiu C-H. 2000. Gene conversion may aid adaptive peak shifts. *J Theor Biol*
717 207(4):495-511.

718 Hellsten U, Harland RM, Gilchrist MJ, Hendrix D, Jurka J, Kapitonov V, Ovcharenko I, Putnam NH, Shu
719 S, Taher L, et al. 2010. The genome of the western clawed frog *Xenopus tropicalis*. *Science*
720 328(5978):633-636.

721 Hime PM, Lemmon AR, Lemmon ECM, Prendini E, Brown JM, Thomson RC, Kratovil JD, Noonan BP,
722 Pyron RA, Peloso PLV, et al. 2021. Phylogenomics reveals ancient gene tree discordance in the
723 amphibian tree of life. *Syst Biol* 70(1):49-66.

724 Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, Muffato M, Collins JE, Humphray S, McLaren
725 K, Matthews L, et al. 2013. The zebrafish reference genome sequence and its relationship to the human
726 genome. *Nature* 496(7446):498-503.

727 Jeziorski MC, Greenberg RM, Anderson PAV. 1997. Cloning of a putative voltage-gated sodium channel
728 from the turbellarian flatworm *Bdelloura candida*. *Parasitology* 115(3):289-296.

729 Jost MC, Hillis DM, Lu Y, Kyle JW, Fozzard HA, Zakon HH. 2008. Toxin-resistant sodium channels:
730 parallel adaptive evolution across a complete gene family. *Mol Biol Evol* 25(6):1016-1024.

731 Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in
732 performance and usability. *Mol Biol Evol* 30(4):772-780.

733 Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz
734 S, Duran C, et al. 2012. Geneious Basic: An integrated and extendable desktop software platform for the
735 organization and analysis of sequence data. *Bioinformatics* 28(12):1647-1649.

736 Khakhlova O, Bock R. 2006. Elimination of deleterious mutations in plastid genomes by gene
737 conversion. *Plant J* 46(1):85-94.

738 Kosakovsky Pond SL, Frost SDW. 2005. Not so different after all: a comparison of methods for detecting
739 amino acid sites under selection. *Mol Biol Evol* 22(5):1208-1222.

740 Kuhner MK, Lawlor DA, Ennis PD, Parham P. 1991. Gene conversion in the evolution of the human and
741 chimpanzee MHC class I loci. *Tissue Antigens* 38(4):152-164.

742 Leffler A, Herzog RI, Dib-Hajj SD, Waxman SG, Cummins TR. 2005. Pharmacological properties of
743 neuronal TTX-resistant sodium channels and the role of a critical serine pore residue. *Pflug Arch Eur J*
744 *Phy* 451(3):454-463.

745 Losos JB. 2011. Convergence, adaptation, and constraint. *Evolution* 65(7):1827-1840.

746 Márquez R, Ramírez-Castañeda V, Amézquita A. 2019. Does batrachotoxin autoresistance coevolve with
747 toxicity in *Phyllobates* poison-dart frogs? *Evolution* 73(2):390-400.

748 Maruta S, Yamaoka K, Yotsu-Yamashita M. 2008. Two critical residues in P-loop regions of pufferfish
749 Na^+ channels on TTX sensitivity. *Toxicon* 51(3):381-387.

750 Matsunami M, Kitano J, Kishida O, Michimae H, Miura T, Nishimura K. 2015. Transcriptome analysis of
751 predator- and prey-induced phenotypic plasticity in the Hokkaido salamander (*Hynobius retardatus*).
752 *Molecular Ecology* 24(12):3064-3076.

753 McGlothlin JW, Chuckalovcak JP, Janes DE, Edwards SV, Feldman CR, Brodie ED, Jr., Pfrender ME,
754 Brodie ED, III. 2014. Parallel evolution of tetrodotoxin resistance in three voltage-gated sodium channel
755 genes in *Thamnophis sirtalis*. *Mol Biol Evol* 31(11):2386-2846.

756 McGlothlin JW, Kobiela ME, Feldman CR, Castoe TA, Vonk FJ, Richardson MK, Brodie ED, Jr.,
757 Pfrender ME, Brodie ED, III. 2016. Historical contingency in a multigene family facilitates adaptive
758 evolution of toxin resistance. *Curr Biol* 26:1616-1621.

759 McGrath CL, Casola C, Hahn MW. 2009. Minimal effect of ectopic gene conversion among recent
760 duplicates in four mammalian genomes. *Genetics* 182(2):615-622.

761 Moreno-Hagelsieb G, Latimer K. 2007. Choosing BLAST options for better detection of orthologs as
762 reciprocal best hits. *Bioinformatics* 24(3):319-324.

763 Nowoshilow S, Schloissnig S, Fei J-F, Dahl A, Pang AWC, Pippel M, Winkler S, Hastie AR, Young G,
764 Roscito JG, et al. 2018. The axolotl genome and the evolution of key tissue formation regulators. *Nature*
765 554(7690):50-55.

766 Ohta T. 1989. The mutational load of a multigene family with uniform members. *Genet Res* 53(2):141-
767 145.

768 Perry BW, Card DC, McGlothlin JW, Pasquesi GIM, Adams RH, Schield DR, Hales NR, Corbin AB,
769 Demuth JP, Hoffmann FG, et al. 2018. Molecular adaptations for sensing and securing prey and insight
770 into amniote genome diversity from the garter snake genome. *Genome Biol Evol* 10(8):2110-2129.

771 Petschenka G, Wagschal V, von Tschirnhaus M, Donath A, Dobler S. 2017. Convergently evolved toxic
772 secondary metabolites in plants drive the parallel molecular evolution of insect resistance. *Am Nat*
773 190(S1):S29-S43.

774 Posada D. 2008. jModelTest: Phylogenetic model averaging. *Mol Biol Evol* 25(7):1253-1256.

775 Posada D, Crandall KA. 2001. Evaluation of methods for detecting recombination from DNA sequences:
776 computer simulations. *P Natl Acad Sci USA* 98(24):13757-13762.

777 Pyron RA. 2014. Biogeographic analysis reveals ancient continental vicariance and recent oceanic
778 dispersal in amphibians. *Syst Biol* 63(5):779-797.

779 Sawyer SA. 1999. GENECONV: a computer package for the statistical detection of gene conversion.
780 http://www.math.wustl.edu/~sawyer.

781 Semple C, Wolfe KH. 1999. Gene duplication and gene conversion in the Caenorhabditis elegans
782 genome. *J Mol Evol* 48(5):555-564.

783 Signore AV, Yang Y-Z, Yang Q-Y, Qin G, Moriyama H, Ge R-L, Storz JF. 2019. Adaptive changes in
784 hemoglobin function in high-altitude Tibetan canids were derived via gene conversion and introgression.
785 *Mol Biol Evol* 36(10):2227-2237.

786 Smit A, Hubley R, Green P. 2013-2015. RepeatMasker Open-4.0.

787 Smith JJ, Timoshevskaya N, Timoshevskiy VA, Keinath MC, Hardy D, Voss SR. 2019. A chromosome-
788 scale assembly of the axolotl genome. *Genome Research* 29(2):317-324.

789 Smith SD, Pennell MW, Dunn CW, Edwards SV. 2020. Phylogenetics is the new genetics (for most of
790 biodiversity). *Trends Ecol Evol* 35(5):415-425.

791 Soong TW, Venkatesh B. 2006. Adaptive evolution of tetrodotoxin resistance in animals. *Trends in*
792 *Genetics* 22(11):621-626.

793 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
794 phylogenies. *Bioinformatics* 30(9):1312-1313.

795 Stokes AN, Ducey PK, Neuman-Lee L, Hanifin CT, French SS, Pfrender ME, Brodie ED, III, Brodie ED,
796 Jr. 2014. Confirmation and distribution of tetrodotoxin for the first time in terrestrial Invertebrates: two
797 terrestrial flatworm species (*Bipalium adventitium* and *Bipalium kewense*). *PLOS One* 9(6):e100718.

798 Storz JF. 2016. Causes of molecular convergence and parallelism in protein evolution. *Nature Reviews*
799 *Genetics* 17(4):239-250.

800 Sun YB, Xiong ZJ, Xiang XY, Liu SP, Zhou WW, Tu XL, Zhong L, Wang L, Wu DD, Zhang BL, et al.
801 2015. Whole-genome sequence of the Tibetan frog *Nanorana parkeri* and the comparative evolution of
802 tetrapod genomes. *P Natl Acad Sci USA* 112(11):E1257.

803 Sunami A, Glaaser IW, Fozzard HA. 2000. A critical residue for isoform difference in tetrodotoxin
804 affinity is a molecular determinant of the external access path for local anesthetics in the cardiac sodium
805 channel. *P Natl Acad Sci USA* 97(5):2326-2331.

806 Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into
807 the corresponding codon alignments. *Nucleic Acids Research* 34(Web Server issue):W609-W612.

808 Tarvin RD, Borghese CM, Sachs W, Santos JC, Lu Y, O'Connell LA, Cannatella DC, Harris RA, Zakon
809 HH. 2017. Interacting amino acid replacements allow poison frogs to evolve epibatidine resistance.
810 *Science* 357(6357):1261-1266.

811 Toledo G, Hanifin C, Geffeney S, Brodie ED, III. 2016. Convergent evolution of tetrodotoxin-resistant
812 sodium channels in predators and prey. *Curr Top Membr* 78:87-113.

813 Vaelli PM, Theis KR, Williams JE, O'Connell LA, Foster JA, Eisthen HL. 2020. The skin microbiome
814 facilitates adaptive tetrodotoxin production in poisonous newts. *eLife* 9:e53898.

815 Venkat A, Hahn MW, Thornton JW. 2018. Multinucleotide mutations cause false inferences of lineage-
816 specific positive selection. *Nat Ecol Evol* 2(8):1280-1288.

817 Venkatesh B, Lu SQ, Dandona N, See SL, Brenner S, Soong TW. 2005. Genetic basis of tetrodotoxin
818 resistance in pufferfishes. *Curr Biol* 15(22):2069-2072.

819 Widmark J, Sundström G, Ocampo Daza D, Larhammar D. 2011. Differential evolution of voltage-gated
820 sodium channels in tetrapods and teleost fishes. *Mol Biol Evol* 28(1):859-871.

821 Williams BL, Hanifin CT, Brodie ED, Jr., Brodie ED, III. 2010. Tetrodotoxin affects survival probability
822 of rough-skinned newts (*Taricha granulosa*) faced with TTX-resistant garter snake predators
823 (*Thamnophis sirtalis*). *Chemoecology* 20(4):285-290.

824 Yang Z. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme
825 evolution. *Mol Biol Evol* 15(5):568-573.

826 Yang Z, dos Reis M. 2010. Statistical properties of the branch-site test of positive selection. *Mol Biol*
827 *Evol* 28(3):1217-1228.

828 Yang Z, Nielsen R, Goldman N, Pedersen A-MK. 2000. Codon-substitution models for heterogeneous
829 selection pressure at amino acid sites. *Genetics* 155(1):431-449.

830 Yang Z, Wong WSW, Nielsen R. 2005. Bayes empirical Bayes inference of amino acid sites under
831 positive selection. *Mol Biol Evol* 22(4):1107-1118.

832 Yang ZH. 2007. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24(8):1586-
833 1591.

834 Yu FH, Catterall WA. 2003. Overview of the voltage-gated sodium channel family. *Genome Biol*
835 4(3):207.

836 Zakon HH. 2012. Adaptive evolution of voltage-gated sodium channels: The first 800 million years. *P*
837 *Natl Acad Sci USA* 109(Supplement 1):10619-10625.

838 Zakon HH, Jost MC, Lu Y. 2011. Expansion of voltage-dependent Na^+ channel gene family in early
839 tetrapods coincided with the emergence of terrestriality and increased brain complexity. *Mol Biol Evol*
840 28(4):1415-1424.

841 Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an improved branch-site likelihood method for detecting
842 positive selection at the molecular level. *Mol Biol Evol* 22(12):2472-2479.

843 Zhen Y, Aardema ML, Medina EM, Schumer M, Andolfatto P. 2012. Parallel molecular evolution in an
844 herbivore community. *Science* 337(6102):1634-1637.
845