



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/141927/>

Version: Accepted Version

Article:

Wielstra, B., McCartney-Melstad, E., Arntzen, J.W. et al. (2019) Phylogenomics of the adaptive radiation of Triturus newts supports gradual ecological niche expansion towards an incrementally aquatic lifestyle. *Molecular Phylogenetics and Evolution*, 133. pp. 120-127. ISSN: 1055-7903

<https://doi.org/10.1016/j.ympev.2018.12.032>

Article available under the terms of the CC-BY-NC-ND licence
(<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

1 Running head: **Phylogenomics of *Triturus newts***

2

3 **Phylogenomics of the adaptive radiation of *Triturus newts* supports gradual ecological**
4 **niche expansion towards an incrementally aquatic lifestyle**

5

6 B. Wielstra^{a,b,c,d,1,*}, E. McCartney-Melstad^{a,e,1}, J.W. Arntzen^c, R.K. Butlin^{b,f}, H.B. Shaffer^{a,e}

7

8 ^a *Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA*
9 *90095, USA.*

10 ^b *Department of Animal and Plant Sciences, University of Sheffield, S10 2TN Sheffield, UK.*

11 ^c *Naturalis Biodiversity Center, 2300 RA Leiden, The Netherlands.*

12 ^d *Institute of Biology Leiden, Leiden University, 2300 RA, Leiden, The Netherlands.*

13 ^e *La Kretz Center for California Conservation Science, Institute of the Environment and*
14 *Sustainability, University of California, Los Angeles, CA 90095, USA.*

15 ^f *Department of Marine Sciences, University of Gothenburg, Gothenburg 405 30, Sweden.*

16 ¹ *These authors contributed equally to this work.*

17 ^{*} *Corresponding author; e-mail: ben.wielstra@naturalis.nl*

18 **Abstract**

19 Newts of the genus *Triturus* (marbled and crested newts) exhibit substantial variation in the
20 number of trunk vertebrae (NTV) and a higher NTV corresponds to a longer annual aquatic
21 period. Because the *Triturus* phylogeny has thwarted resolution to date, the evolutionary
22 history of NTV, annual aquatic period, and their potential coevolution has remained unclear.
23 To resolve the phylogeny of *Triturus*, we generated a c. 6,000 transcriptome-derived marker
24 data set using a custom target enrichment probe set, and conducted phylogenetic analyses using:
25 1) data concatenation with RAxML, 2) gene-tree summary with ASTRAL, and 3) species-tree
26 estimation with SNAPP. All analyses produce the same, highly supported topology, despite
27 cladogenesis having occurred over a short timeframe, resulting in short internal branch lengths.
28 Our new phylogenetic hypothesis is consistent with the minimal number of inferred changes in
29 NTV count necessary to explain the diversity in NTV observed today. Although a causal
30 relationship between NTV, body form, and aquatic ecology has yet to be experimentally
31 established, our phylogeny indicates that these features have evolved together, and suggest that
32 they may underlie the adaptive radiation that characterizes *Triturus*.

33

34 **Keywords:** morphology; phylogeny; sequence capture; systematics; target enrichment;
35 transcriptome

36 **1. Introduction**

37 Accurately retracing the evolution of phenotypic diversity in adaptive radiations requires well-
38 established phylogenies. However, inferring the true branching order in adaptive radiations is
39 hampered by the short time frame over which they typically unfold, which provides little
40 opportunity between splitting events for phylogenetically informative substitutions to become
41 established (resulting in low phylogenetic resolution; Philippe et al., 2011; Whitfield and
42 Lockhart, 2007) and fixed (resulting in incomplete lineage sorting and discordance among
43 gene-trees; Degnan and Rosenberg, 2006; Pamilo and Nei, 1988; Pollard et al., 2006).
44 Resolving the phylogeny of rapidly multiplying lineages becomes even more complicated the
45 further back in time the radiation occurred, because the accumulation of parallel substitutions
46 along terminal branches can lead to long-branch attraction (Felsenstein, 1978; Swofford et al.,
47 2001). A final impediment is reticulation between closely related (and not necessarily sister-)
48 species through past or ongoing hybridization, resulting in additional gene-tree/species-tree
49 discordance (Kutschera et al., 2014; Leaché et al., 2014; Mallet et al., 2016).

50 Phylogenomics, involving the consultation of a large number of markers spread
51 throughout the genome, has proven successful in resolving both recent (e.g. Giarla and
52 Esselstyn, 2015; Leaché et al., 2016; Léveillé-Bourret et al., 2018; Meiklejohn et al., 2016;
53 Nater et al., 2015; Scott et al., 2018; Shi and Yang, 2018) and more ancient (e.g. Crawford et
54 al., 2012; Irisarri and Meyer, 2016; Jarvis et al., 2014; McCormack et al., 2012; Song et al.,
55 2012) evolutionary radiations. The appeal of greatly increasing the amount of data available
56 for any given phylogenetic problem is that it often (but not always; see Philippe et al., 2011)
57 provides informative characters to resolve short branches in the tree of life. Advances in
58 laboratory and sequencing techniques, bioinformatics, and tree-building methods all facilitate
59 phylogenetic reconstruction based on thousands of homologous loci for a large number of
60 individuals, and promise to help provide the phylogenetic trees necessary to interpret the

61 evolution of eco-morphological characters involved in adaptive radiations (Alföldi et al., 2011;
62 Stroud and Losos, 2016). In this study, we conduct a phylogenomic analysis of an adaptive
63 radiation that moderately-sized multilocus nuclear DNA datasets (Arntzen et al., 2007;
64 Espregueira Themudo et al., 2009; Wielstra et al., 2014) have consistently failed to resolve: the
65 Eurasian newt genus *Triturus* (Amphibia: Urodela: Salamandridae), commonly known as the
66 marbled and crested newts.

67 One of the most intriguing features of *Triturus* evolution is the correlation between
68 certain aspects of their ecology and the number of trunk vertebrae (NTV; Fig. 1). Species
69 characterized by a higher modal NTV (which translates into a more elongate body build with
70 proportionally shorter limbs) are associated with a more aquatic lifestyle. Empirically, the
71 number of months a *Triturus* species spends in the water (defined at the population level as the
72 peak date of emigration, leaving a breeding pond, minus the peak in immigration, entering it)
73 roughly equals NTV minus 10 (Arntzen, 2003; Arntzen and Wallis, 1999; Slijepčević et al.,
74 2015). The intrageneric variation in NTV shown by *Triturus*, ranging from 12 to 17, is
75 unparalleled in the family Salamandridae (Arntzen et al., 2015; Lanza et al., 2010) and a causal
76 relationship between NTV expansion and an increasingly aquatic lifestyle has been presumed,
77 but never adequately placed into a phylogenetic comparative analysis (Arntzen, 2003; Arntzen
78 et al., 2015; Arntzen and Wallis, 1999; Govedarica et al., 2017; Slijepčević et al., 2015;
79 Urošević et al., 2016; Vukov et al., 2011; Wielstra and Arntzen, 2011). A well-established
80 *Triturus* species-tree is required to accurately retrace NTV evolution and assess the
81 concordance between aquatic lifestyle and NTV across the genus.

82 Our goal is to obtain a genome-enabled phylogeny for *Triturus* and use it to reconstruct
83 the eco-morphological evolution of NTV and aquatic/terrestrial ecology across the genus. As
84 the large size of salamander genomes hampers whole-genome sequencing (but see Elewa et al.,
85 2017; Nowoshilow et al., 2018; Smith et al., 2018), we employ a genome-reduction approach

86 in which we capture and sequence a set of transcriptome-derived markers using target
87 enrichment, an efficient technique that affords extremely high resolution at multiple taxonomic
88 levels (Abdelkrim et al., 2018; Bi et al., 2012; Bragg et al., 2016; Gnirke et al., 2009;
89 McCartney-Melstad et al., 2016; McCartney-Melstad et al., 2018). Using data concatenation
90 (with RAxML), gene-tree summarization (with ASTRAL) and species-tree estimation (with
91 SNAPP), we fully resolve the *Triturus* phylogeny and place the extreme body shape and
92 ecological variation observed in this adaptive radiation into an evolutionary context.

93

94 **2. Materials and Methods**

95

96 *2.1 Target capture array design*

97 Nine *Triturus* newts (seven crested and two marbled newt species) and one banded newt
98 (*Ommatotriton*) were subjected to transcriptome sequencing. Transcriptome assemblies for
99 each species were generated using Trinity v2.2.0 (Grabherr et al., 2011), clustered at 90% using
100 usearch v9.1.13 (Edgar, 2010), and subjected to reciprocal best blast hit analysis (Bork et al.,
101 1998; Camacho et al., 2009; Tatusov et al., 1997) to produce a set of *T. dobrogicus* transcripts
102 (the species with the highest quality transcriptome assembly) that had putative orthologues
103 present in the nine other transcriptome assemblies. These transcripts were then annotated using
104 blastx to *Xenopus tropicalis* proteins, retaining one annotated transcript per protein. We
105 attempted to discern splice sites in the transcripts, as probes spanning splice boundaries may
106 perform poorly (Neves et al., 2013), by mapping transcripts iteratively to the genomes of
107 *Chrysemys picta* (Shaffer et al., 2013), *X. tropicalis* (Hellsten et al., 2010), *Nanorana parkerii*
108 (Sun et al., 2015) and *Rana catesbeiana* (Hammond et al., 2017). A single exon ≥ 200 bp and
109 ≤ 450 bp was retained for each transcript target. To increase the ability of the target set to
110 capture markers across all *Triturus* species, orthologous sequences from multiple species were

111 included for targets with > 5% sequence divergence from *T. dobrogicus* (Bi et al., 2012). We
112 generated a target set of 7,102 genomic regions for a total target length of approximately 2.3
113 million bp. A total of 39,143 unique RNA probes were synthesized as a MyBaits-II kit for this
114 target set at approximately 2.6X tiling density by Arbor Biosciences (Ann Arbor, MI, Ref#
115 170210-32). A detailed outline of the target capture array design process is presented in
116 Supplementary Text S1.

117

118 *2.2 Sampling scheme*

119 We sampled 23 individual *Triturus* newts (Fig. 2; Supplementary Table S1) for which tissues
120 were available from previous studies (Wielstra et al., 2017a; Wielstra et al., 2017b; Wielstra et
121 al., 2013). Because the sister-group relationship between the two marbled and seven crested
122 newts is well established (Fig. 1), while the relationships among the crested newt species have
123 defied resolution, we sampled the crested newt species more densely, including three
124 individuals per species to include intraspecific differentiation and to avoid misleading
125 phylogenies resulting from single exemplar sampling (Spinks et al., 2013). Because *Triturus*
126 species show introgressive hybridization at contact zones (Arntzen et al., 2014), we aimed to
127 reduce the impact of interspecific gene flow by only including individuals that originate away
128 from hybrid zones and have previously been interpreted as unaffected by interspecific genetic
129 admixture (Wielstra et al., 2017a; Wielstra et al., 2017b). The reality of phylogenetic distortion
130 by interspecific gene flow was underscored in a test for the phylogenetic utility of the
131 transcripts used for marker design which included a genetically admixed individual (details in
132 Supplementary Text S1).

133

134 *2.3 Laboratory methods*

135 DNA was extracted from samples using a salt extraction protocol (Sambrook and Russell,
136 2001), and 10,000ng per sample was sheared to approximately 200bp-500bp on a BioRuptor
137 NGS (Diagenode) and dual-end size selected (0.8X-1.0X) with SPRI beads. Dual-indexed
138 libraries were prepared from 375-2000ng of size selected DNA using KAPA LTP library prep
139 kits (Glenn et al., 2017). These libraries were pooled (with samples from other projects) into
140 batches of 16 samples at 250ng per sample (4,000ng total) and enriched in the presence of
141 30,000ng of c0t-1 repetitive sequence blocker (McCartney-Melstad et al., 2016) derived from
142 *T. carnifex* (casualties from a removal action of an invasive population (Meilink et al., 2015))
143 by hybridizing blockers with libraries for 30 minutes and probes with libraries/blockers for 30
144 hours. Enriched libraries were subjected to 14 cycles of PCR with KAPA HiFi HotStart
145 ReadyMix and pooled at an equimolar ratio for 150bp paired-end sequencing across multiple
146 Illumina HiSeq 4000 lanes (receiving an aggregate of 18% of one lane, for a multiplexing
147 equivalent of 128 samples per lane).

148

149 *2.4 Processing of target capture data*

150 A total of 3,937,346 read pairs from the sample receiving the greatest number of reads were
151 used to *de novo* assemble target sequences for each target region using the assembly by reduced
152 complexity (ARC) pipeline (Hunter et al., 2015). A single assembled contig was selected for
153 each original target region by means of reciprocal best blast hit (RBBH) (Rivera et al., 1998),
154 and these were used as a reference assembly for all downstream analyses. Adapter
155 contamination was removed from sample reads using skewer v0.2.2 (Jiang et al., 2014), and
156 reads were then mapped to the reference assembly using BWA-MEM v0.7.15-r1140 (Li, 2013).
157 Picard tools v2.9.2 (<https://broadinstitute.github.io/picard/>) was used to add read group
158 information and to mark PCR duplicates, and HaplotypeCaller and GenotypeGVCFs from
159 GATK v3.8 (McKenna et al., 2010) were used jointly to genotype the relevant groups of

160 samples (either crested newts or crested newts + marbled newts depending on the analysis; see
161 below). SNPs that failed any of the following hard filters were removed: $QD < 2$, $MQ < 40$,
162 $FS > 60$, $MQRankSum < -12.5$, $ReadPosRankSum < -8$, and $QUAL < 30$ (Poplin et al., 2017).
163 We next attempted to remove paralogous targets from our dataset with a Hardy Weinberg
164 Equilibrium (HWE) filter for heterozygote excess. Heterozygote excess p-values were
165 calculated for every SNP using vcfTools 0.1.15 (Danecek et al., 2011), and any target containing
166 at least one SNP with a heterozygote excess p-value < 0.05 was removed from downstream
167 analysis. More detail on the processing of the target capture data can be found in Supplementary
168 Text S2.

169

170 *2.5 Phylogenetic analyses*

171 A concatenated maximum likelihood phylogeny was inferred with RAxML version 8.2.11
172 (Stamatakis, 2014) based on an alignment of 133,601 SNPs across 5,866 different targets. We
173 included all 23 *Triturus* individuals in this analysis. For gene-tree summary, ASTRAL v5.6.1
174 (Zhang et al., 2017) was used to estimate the crested newt species-tree from 5,610 gene-trees
175 generated in RAxML. The 21 crested newt samples were assigned species membership, and no
176 marbled newts were included because estimating terminal branch lengths is not possible for
177 species with a single representative. For species-tree estimation, SNAPP v1.3.0 (Bryant et al.,
178 2012) within the BEAST v2.4.8 (Bouckaert et al., 2014) environment was used to infer the
179 crested newt species-tree from single biallelic SNPs randomly selected from each of 5,581
180 post-filtering targets. All three individuals per crested newt species were treated as a single
181 terminal and marbled newts were again excluded given our single exemplar sampling of both
182 species. We also estimated divergence times in SNAPP for the crested newts. The split between
183 *T. carnifex* and *T. macedonicus*, assumed to correspond to the origin of the Adriatic Sea at the
184 end of the Messinian Salinity Crisis 5.33 million years ago, was used as a single calibration

185 point (Arntzen et al., 2007; Wielstra and Arntzen, 2011) to produce a rough estimate of the
186 timing of cladogenesis. A detailed description of our strategy for phylogenetic analyses is
187 available in Supplementary Text S3.

188

189 **3. Results**

190 Samples received a mean of 2,812,980 read pairs (s.d. = 585,815). Enrichment was highly
191 efficient, especially given the large genome size of *Triturus*, with an average of 44.5% of raw
192 reads mapping to the assembled target sequences (s.d. = 2.6%). After removing PCR duplicates,
193 which accounted for an average of 22.6% of mapped reads, the unique read on target rate was
194 34.4% (s.d. = 1.9%). The 23 samples in the final RAxML alignment contained an average of
195 10.1% missing data (min = 3.2%, max = 31.8%) after setting genotype calls with GQ scores of
196 less than 20 to missing.

197 The concatenated analysis with RAxML supports a basal bifurcation in *Triturus* between
198 the marbled and crested newts (Fig. 3), consistent with the prevailing view that they are
199 reciprocally monophyletic (Arntzen et al., 2007; Espregueira Themudo et al., 2009; Wielstra
200 et al., 2014). RAxML also recovers each of the crested newt species as monophyletic,
201 validating our decision to collapse the three individuals sampled per species in a single terminal
202 in ASTRAL and SNAPP. Furthermore, all five *Triturus* body builds are recovered as
203 monophyletic (cf. Arntzen et al., 2007; Espregueira Themudo et al., 2009; Wielstra et al., 2014).
204 The greatest intraspecific divergence is observed in *T. carnifex* (Supplementary Text S1;
205 Supplementary Fig. S1; Supplementary Table S2).

206 Phylogenetic inference based on data concatenation with RAxML (Fig. 3), gene-tree
207 summary with ASTRAL (Fig. 4a) and species-tree estimation with SNAPP (Fig. 4b) all recover
208 the same crested newt topology, with a basal bifurcation between the *T. karelinii*-group (NTV
209 = 13; *T. ivanbureschi* as the sister taxon to *T. anatolicus* + *T. karelinii*) and the remaining taxa,

210 which themselves are resolved into the species pairs *T. carnifex* + *T. macedonicus* (NTV=14;
211 the *T. carnifex*-group), and *T. cristatus* (NTV=15) + *T. dobrogicus* (NTV=16/17). Despite the
212 rapidity of cladogenesis, we obtain strong branch support for every internal node. Even with
213 the uncertainty in dating given a single biogeographically-derived calibration date, the
214 bifurcation giving rise to the four crested newt species groups (cf. Fig. 1) must have occurred
215 over a relatively short time frame (Fig. 5), reflected by two particularly short, but resolvable
216 internal branches (Fig. 3; Fig. 4).

217 The phylogenomic analyses suggest considerable gene-tree/species-tree discordance in
218 *Triturus*. The normalized quartet score of the ASTRAL tree (Fig. 4a), which reflects the
219 proportion of input gene-tree quartets consistent with the species-tree, is 0.63, indicating a high
220 degree of gene-tree discordance. Furthermore, the only node in the SNAPP tree with a posterior
221 probability below 1.0 (i.e. 0.99) is subtended by a very short branch (Fig. 4b). Consistent with
222 the high level of gene-tree/species-tree discordance, we also found that the full mtDNA-based
223 phylogeny of *Triturus* produced a highly supported, but topologically different, phylogeny
224 (Supplementary Text S3; Supplementary Fig. S2; Wielstra and Arntzen, 2011).

225 Considering an NTV count of 12, as observed in the marbled newts as well as the most
226 closely related newt genera, as the ancestral state for *Triturus* (Arntzen et al., 2015; Veith et
227 al., 2018), three sequential single-vertebral additions to NTV along internal branches, and one
228 or two additions along the terminal branch leading to *T. dobrogicus* (in which NTV = 16 and
229 NTV = 17 occur at approximately equal frequency; Arntzen et al., 2015; Wielstra et al., 2016)
230 are required under a parsimony criterion (with either ACCTRAN or DELTRAN optimization)
231 to explain the present-day variation in NTV observed in *Triturus* (Fig. 3). This is the minimum
232 possible number of inferred changes in NTV count required to explain the NTV radiation
233 observed today (Supplementary Fig. S3; Supplementary Text S5). No NTV deletions or

234 reversals are required, implying a linear, stepwise, single-addition scenario for NTV expansion
235 in *Triturus*.

236

237 **4. Discussion**

238 We use a large, transcriptome-derived phylogenomic dataset to construct a phylogenetic
239 hypothesis and study the evolution of ecological and phenotypic diversity within the adaptive
240 radiation of *Triturus* newts. In contrast to previous attempts to recover a multilocus species-
241 tree (Arntzen et al., 2007; Espregueira Themudo et al., 2009; Wielstra et al., 2014), we recover
242 full phylogenetic resolution with strong support across the tree. Despite cladogenesis having
243 occurred in a relatively brief time window (Fig. 5), resulting in a high degree of gene-
244 tree/species-tree discordance, independent phylogenetic approaches based on data
245 concatenation (RAxML), gene-tree summarization (ASTRAL) and species-tree estimation
246 (SNAPP), all recover the same, highly supported topology for *Triturus* (Fig. 3; Fig. 4). Our
247 *Triturus* case study underscores that sequence capture by target enrichment is a promising
248 approach to resolve the phylogenetic challenges associated with adaptive radiations,
249 particularly for taxa with large and complicated genomes where other genomic approaches are
250 impractical, including salamanders (McCartney-Melstad et al., 2016).

251 Our new phylogenetic hypothesis allows us to place the eco-morphological
252 differentiation shown by *Triturus* into a coherent evolutionary context. Over time, *Triturus*
253 expanded its range of NTV to encompass higher counts (Fig. 3). The *Triturus* tree is consistent
254 with a maximally parsimonious scenario, under which four to five character state changes are
255 required to explain the radiation in NTV observed today. Any other possible phylogenetic
256 relationship among *Triturus* body builds would require a higher number of inferred NTV
257 changes (Supplementary Fig. S3). Three of these inferred changes are positioned along internal
258 branches, of which two are particularly short, suggesting that changes in NTV count can evolve

259 over a relatively short time. The fourth and fifth inferred change are situated on the external
260 branch leading to *T. dobrogicus*, the only *Triturus* species with substantial intraspecific
261 variation in NTV count (Arntzen et al., 2015; Wielstra et al., 2016).

262 Newts annually alternate between an aquatic and a terrestrial habitat, and the functional
263 trade-off between adaptation to life in water or on land likely poses contrasting demands on
264 body build (Fish and Baudinette, 1999; Gillis and Blob, 2001; Gvoždík and van Damme, 2006;
265 Shine and Shetty, 2001). Considering the observed relationship between one additional trunk
266 vertebra and an extra month annually spent in the water (Fig. 1), the extraordinary NTV
267 variation observed in *Triturus* may reflect the morphological mechanism by which more
268 efficient exploitation of a wider range in hydroperiod (i.e. the annual availability of standing
269 water) evolved. Despite the evolvability of NTV count (Arntzen et al., 2015), NTV evolution
270 has been phylogenetically constrained in *Triturus*. Apparently the change in NTV was
271 directional and involved the addition of a single trunk vertebra at a time (Fig. 3; Supplementary
272 Fig. S3). Species with a more derived body build, reflected in a higher NTV, have a relatively
273 prolonged aquatic period and, because species with transitional NTV counts remain extant, the
274 end result is an eco-morphological radiation.

275 *Triturus* newts show a slight degree of intraspecific variation in NTV today. Such
276 variation is partially explained by interspecific hybridization (emphasizing the genetic basis of
277 NTV count; Arntzen et al., 2014), but there is standing variation in NTV count within all
278 *Triturus* species (Slijepčević et al., 2015). This suggests that, during *Triturus* evolution, there
279 has always been intraspecific NTV count polymorphism that could be subjected to natural
280 selection. Whether there is a causal relationship between the directional, parsimonious
281 evolution of higher NTV and the equally parsimonious evolutionary increase in aquatic
282 lifestyle, and, if so, which of these two may be the actual target of selection, remain important
283 open questions. A proper understanding of the functional relationship between NTV, body

284 build and fitness in aquatic/terrestrial environments in *Triturus* is still lacking (Gvoždík and
285 van Damme, 2006), and functional studies exploring this fitness landscape across intra and
286 interspecific variation in NTV is an important next step in establishing a firm causal
287 relationship between variation, performance and fitness. The recent availability of the first
288 salamander genomes (Elewa et al., 2017; Nowoshilow et al., 2018; Smith et al., 2018) finally
289 offers the prospect of sequencing the genome of each *Triturus* species and exploring the
290 developmental basis for NTV and its functional consequences in the diversification of the
291 genus.

292

293 **Acknowledgements**

294 Andrea Chiochio, Daniele Canestrelli, Michael Fahrbach, Ana Ivanović, Raymond van der
295 Lans, and Kurtuluş Olgun helped obtain samples for transcriptome sequencing. Permits were
296 provided by the Italian Ministry of the Environment (DPN-2009-0026530), the Environment
297 Protection Agency of Montenegro (no. UPI-328/4), the Ministry of Energy, Development and
298 Environmental Protection of Republic of Serbia (no. 353-01-75/2014-08), and TÜBİTAK,
299 Turkey (no. 113Z752). RAVON & Natuurbalans-Limes Divergens provided the *T. carnifex*
300 used to create c0t-1. Tara Luckau helped in the lab. Peter Scott provided valuable suggestions
301 on methodology. Wiesław Babik and Ana Ivanović commented on an earlier version of this
302 manuscript. This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC
303 Berkeley, supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303.
304 Computing resources were provided by XSEDE (Towns et al., 2014) and the Texas Advanced
305 Computing Center (TACC) Stampede2 cluster at The University of Texas at Austin.

306

307 **Funding**

308 This project has received funding from the European Union’s Horizon 2020 research and
309 innovation programme under the Marie Skłodowska-Curie grant agreement No. 655487.

310

311 **Data availability**

312 Raw sequence read data for the sequence capture libraries of the 23 *Triturus* samples and the
313 12 transcriptome libraries are available at SRA (PRJNA498336). Transcriptome assemblies,
314 genotype calls (VCF) for the 21- and 23-sample datasets, input files for the RAxML, ASTRAL
315 and SNAPP analyses, and synthesized target sequences are available at Zenodo
316 (<https://doi.org/10.5281/zenodo.1470914>). Supplementary data associated with this article can
317 be found, in the online version, at <https://doi.org/10.1016/j.ympcv.2018.12.032>.

318

319 **References**

320 Abdelkrim, J., Aznar-Cormano, L., Fedosov, A., Kantor, Y., Lozouet, P., Phuong, M., Zaharias,
321 P., Puillandre, N., 2018. Exon-capture based phylogeny and diversification of the venomous
322 gastropods (Neogastropoda, Conoidea). *Mol. Biol. Evol.* 35, 2355–2374.

323 Alföldi, J., Di Palma, F., Grabherr, M., Williams, C., Kong, L., Mauceli, E., Russell, P., Lowe,
324 C.B., Glor, R.E., Jaffe, J.D., Ray, D.A., Boissinot, S., Shedlock, A.M., Botka, C., Castoe, T.A.,
325 Colbourne, J.K., Fujita, M.K., Moreno, R.G., ten Hallers, B.F., Haussler, D., Heger, A.,
326 Heiman, D., Janes, D.E., Johnson, J., de Jong, P.J., Koriabine, M.Y., Lara, M., Novick, P.A.,
327 Organ, C.L., Peach, S.E., Poe, S., Pollock, D.D., de Queiroz, K., Sanger, T., Searle, S., Smith,
328 J.D., Smith, Z., Swofford, R., Turner-Maier, J., Wade, J., Young, S., Zadissa, A., Edwards,
329 S.V., Glenn, T.C., Schneider, C.J., Losos, J.B., Lander, E.S., Breen, M., Ponting, C.P.,
330 Lindblad-Toh, K., 2011. The genome of the green anole lizard and a comparative analysis with
331 birds and mammals. *Nature* 477, 587.

332 Arntzen, J.W., 2003. *Triturus cristatus* Superspecies - Kammolch-Artenkreis (*Triturus*
333 *cristatus* (Laurenti, 1768) - Nördlicher Kammolch, *Triturus carnifex* (Laurenti, 1768) -
334 Italienischer Kammolch, *Triturus dobrogicus* (Kiritzescu, 1903) - Donau-Kammolch, *Triturus*
335 *karelinii* (Strauch, 1870) - Südlicher Kammolch). In: Grossenbacher, K., Thiesmeier, B. (Eds.),
336 Handbuch der Reptilien und Amphibien Europas. Schwanzlurche IIA. Aula-Verlag,
337 Wiebelsheim, pp. 421-514.

338 Arntzen, J.W., Beukema, W., Galis, F., Ivanović, A., 2015. Vertebral number is highly
339 evolvable in salamanders and newts (family Salamandridae) and variably associated with
340 climatic parameters. *Contrib. Zool.* 84, 85-113.

341 Arntzen, J.W., Espregueira Themudo, G., Wielstra, B., 2007. The phylogeny of crested newts
342 (*Triturus cristatus* superspecies): nuclear and mitochondrial genetic characters suggest a hard
343 polytomy, in line with the paleogeography of the centre of origin. *Contrib. Zool.* 76, 261-278.

344 Arntzen, J.W., Wallis, G.P., 1999. Geographic variation and taxonomy of crested newts
345 (*Triturus cristatus* superspecies): morphological and mitochondrial data. *Contrib. Zool.* 68,
346 181-203.

347 Arntzen, J.W., Wielstra, B., Wallis, G.P., 2014. The modality of nine *Triturus* newt hybrid
348 zones, assessed with nuclear, mitochondrial and morphological data. *Biol. J. Linn. Soc.* 113,
349 604–622.

350 Bi, K., Vanderpool, D., Singhal, S., Linderoth, T., Moritz, C., Good, J.M., 2012.
351 Transcriptome-based exon capture enables highly cost-effective comparative genomic data
352 collection at moderate evolutionary scales. *BMC Genomics* 13, 403.

353 Bork, P., Dandekar, T., Diaz-Lazcoz, Y., Eisenhaber, F., Huynen, M., Yuan, Y., 1998.
354 Predicting function: from genes to genomes and back. *J. Mol. Biol.* 283, 707-725.

355 Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A.,
356 Rambaut, A., Drummond, A.J., 2014. BEAST 2: a software platform for Bayesian evolutionary
357 analysis. *PLoS Comput. Biol.* 10, e1003537.

358 Bragg, J.G., Potter, S., Bi, K., Moritz, C., 2016. Exon capture phylogenomics: efficacy across
359 scales of divergence. *Mol. Ecol. Resour.* 16, 1059-1068.

360 Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N.A., RoyChoudhury, A., 2012. Inferring
361 species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent
362 analysis. *Mol. Biol. Evol.* 29, 1917-1932.

363 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L.,
364 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10, 421-421.

365 Crawford, N.G., Faircloth, B.C., McCormack, J.E., Brumfield, R.T., Winker, K., Glenn, T.C.,
366 2012. More than 1000 ultraconserved elements provide evidence that turtles are the sister group
367 of archosaurs. *Biol. Lett.* 8, 783-786.

368 Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker,
369 R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R., Genomes Project Analysis,
370 G., 2011. The variant call format and VCFtools. *Bioinformatics* 27, 2156-2158.

371 Degnan, J.H., Rosenberg, N.A., 2006. Discordance of species trees with their most likely gene
372 trees. *PLoS Genet.* 2, e68.

373 Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST.
374 *Bioinformatics* 26, 2460-2461.

375 Elewa, A., Wang, H., Talavera-López, C., Joven, A., Brito, G., Kumar, A., Hameed, L.S.,
376 Penrad-Mobayed, M., Yao, Z., Zamani, N., Abbas, Y., Abdullayev, I., Sandberg, R., Grabherr,
377 M., Andersson, B., Simon, A., 2017. Reading and editing the *Pleurodeles waltl* genome reveals
378 novel features of tetrapod regeneration. *Nat. Commun.* 8, 2286.

379 Espregueira Themudo, G., Wielstra, B., Arntzen, J.W., 2009. Multiple nuclear and
380 mitochondrial genes resolve the branching order of a rapid radiation of crested newts (*Triturus*,
381 Salamandridae). *Mol. Phylogenet. Evol.* 52, 321-328.

382 Felsenstein, J., 1978. Cases in which parsimony or compatibility methods will be positively
383 misleading. *Syst. Biol.* 27, 401-410.

384 Fish, F.E., Baudinette, R.V., 1999. Energetics of locomotion by the Australian water rat
385 (*Hydromys chrysogaster*): a comparison of swimming and running in a semi-aquatic mammal.
386 *The Journal of Experimental Biology* 202, 353.

387 Giarla, T.C., Esselstyn, J.A., 2015. The challenges of resolving a rapid, recent radiation:
388 empirical and simulated phylogenomics of Philippine shrews. *Syst. Biol.* 64, 727-740.

389 Gillis, G.B., Blob, R.W., 2001. How muscles accommodate movement in different physical
390 environments: aquatic vs. terrestrial locomotion in vertebrates. *Comparative Biochemistry and*
391 *Physiology Part A: Molecular & Integrative Physiology* 131, 61-75.

392 Glenn, T.C., Bayona-Vásquez, N.J., Kieran, T.J., Pierson, T.W., Hoffberg, S.L., Scott, P.A.,
393 Louha, S., Bentley, K.E., Finger Jr., J.W., Troendle, N., Díaz-Jaimes, P., Mauricio, R.,
394 Faircloth, B.C., 2017. Adapterama III: Quadruple-indexed, triple-enzyme RADseq libraries
395 for about \$1USD per Sample (3RAD). *BioRxiv*.

396 Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., LeProust, E.M., Brockman, W., Fennell, T.,
397 Giannoukos, G., Fisher, S., Russ, C., Gabriel, S., Jaffe, D.B., Lander, E.S., Nusbaum, C., 2009.
398 Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted
399 sequencing. *Nat. Biotechnol.* 27, 182.

400 Govedarica, P., Cvijanović, M., Slijepčević, M., Ivanović, A., 2017. Trunk elongation and
401 ontogenetic changes in the axial skeleton of *Triturus* newts. *J. Morphol.* 278, 1577-1585.

402 Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X.,
403 Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind,
404 N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011.
405 Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat.*
406 *Biotechnol.* 29, 644-652.

407 Gvoždík, L., van Damme, R., 2006. *Triturus* newts defy the running-swimming dilemma.
408 *Evolution* 60, 2110-2121.

409 Hammond, S.A., Warren, R.L., Vandervalk, B.P., Kucuk, E., Khan, H., Gibb, E.A., Pandoh,
410 P., Kirk, H., Zhao, Y., Jones, M., Mungall, A.J., Coope, R., Pleasance, S., Moore, R.A., Holt,
411 R.A., Round, J.M., Ohora, S., Walle, B.V., Veldhoen, N., Helbing, C.C., Birol, I., 2017. The
412 North American bullfrog draft genome provides insight into hormonal regulation of long
413 noncoding RNA. *Nat. Commun.* 8, 1433.

414 Hellsten, U., Harland, R.M., Gilchrist, M.J., Hendrix, D., Jurka, J., Kapitonov, V., Ovcharenko,
415 I., Putnam, N.H., Shu, S., Taher, L., Blitz, I.L., Blumberg, B., Dichmann, D.S., Dubchak, I.,
416 Amaya, E., Detter, J.C., Fletcher, R., Gerhard, D.S., Goodstein, D., Graves, T., Grigoriev, I.V.,
417 Grimwood, J., Kawashima, T., Lindquist, E., Lucas, S.M., Mead, P.E., Mitros, T., Ogino, H.,
418 Ohta, Y., Poliakov, A.V., Pollet, N., Robert, J., Salamov, A., Sater, A.K., Schmutz, J., Terry,
419 A., Vize, P.D., Warren, W.C., Wells, D., Wills, A., Wilson, R.K., Zimmerman, L.B., Zorn,

- 420 A.M., Grainger, R., Grammer, T., Khokha, M.K., Richardson, P.M., Rokhsar, D.S., 2010. The
421 genome of the western clawed frog *Xenopus tropicalis*. *Science* 328, 633.
- 422 Hunter, S.S., Lyon, R.T., Sarver, B.A.J., Hardwick, K., Forney, L.J., Settles, M.L., 2015.
423 Assembly by Reduced Complexity (ARC): a hybrid approach for targeted assembly of
424 homologous sequences. *bioRxiv*.
- 425 Irisarri, I., Meyer, A., 2016. The identification of the closest living relative(s) of tetrapods:
426 phylogenomic lessons for resolving short ancient internodes. *Syst. Biol.* 65, 1057-1075.
- 427 Jarvis, E.D., Mirarab, S., Aberer, A.J., Li, B., Houde, P., Li, C., Ho, S.Y.W., Faircloth, B.C.,
428 Nabholz, B., Howard, J.T., Suh, A., Weber, C.C., da Fonseca, R.R., Li, J., Zhang, F., Li, H.,
429 Zhou, L., Narula, N., Liu, L., Ganapathy, G., Boussau, B., Bayzid, M.S., Zavidovych, V.,
430 Subramanian, S., Gabaldón, T., Capella-Gutiérrez, S., Huerta-Cepas, J., Rekepalli, B., Munch,
431 K., Schierup, M., Lindow, B., Warren, W.C., Ray, D., Green, R.E., Bruford, M.W., Zhan, X.,
432 Dixon, A., Li, S., Li, N., Huang, Y., Derryberry, E.P., Bertelsen, M.F., Sheldon, F.H.,
433 Brumfield, R.T., Mello, C.V., Lovell, P.V., Wirthlin, M., Schneider, M.P.C., Prosdocimi, F.,
434 Samaniego, J.A., Velazquez, A.M.V., Alfaro-Núñez, A., Campos, P.F., Petersen, B., Siche-
435 Ponten, T., Pas, A., Bailey, T., Scofield, P., Bunce, M., Lambert, D.M., Zhou, Q., Perelman,
436 P., Driskell, A.C., Shapiro, B., Xiong, Z., Zeng, Y., Liu, S., Li, Z., Liu, B., Wu, K., Xiao, J.,
437 Yinqi, X., Zheng, Q., Zhang, Y., Yang, H., Wang, J., Smeds, L., Rheindt, F.E., Braun, M.,
438 Fjeldsa, J., Orlando, L., Barker, F.K., Jönsson, K.A., Johnson, W., Koepfli, K.-P., O'Brien, S.,
439 Haussler, D., Ryder, O.A., Rahbek, C., Willerslev, E., Graves, G.R., Glenn, T.C., McCormack,
440 J., Burt, D., Ellegren, H., Alström, P., Edwards, S.V., Stamatakis, A., Mindell, D.P., Cracraft,
441 J., Braun, E.L., Warnow, T., Jun, W., Gilbert, M.T.P., Zhang, G., 2014. Whole-genome
442 analyses resolve early branches in the tree of life of modern birds. *Science* 346, 1320.
- 443 Jiang, H., Lei, R., Ding, S.-W., Zhu, S., 2014. Skewer: a fast and accurate adapter trimmer for
444 next-generation sequencing paired-end reads. *BMC Bioinformatics* 15, 182.
- 445 Kutschera, V.E., Bidon, T., Hailer, F., Rodi, J.L., Fain, S.R., Janke, A., 2014. Bears in a forest
446 of gene trees: phylogenetic inference is complicated by incomplete lineage sorting and gene
447 flow. *Mol. Biol. Evol.* 31, 2004-2017.
- 448 Lanza, B., Arntzen, J.W., Gentile, E., 2010. Vertebral numbers in the Caudata of the Western
449 Palearctic (Amphibia). *Atti Mus. Civ. Stor. Nat. Trieste* 54, 3-114.
- 450 Leaché, A.D., Banbury, B.L., Linkem, C.W., de Oca, A.N.-M., 2016. Phylogenomics of a rapid
451 radiation: is chromosomal evolution linked to increased diversification in north american spiny
452 lizards (Genus *Sceloporus*)? *BMC Evol. Biol.* 16, 63.
- 453 Leaché, A.D., Harris, R.B., Rannala, B., Yang, Z., 2014. The influence of gene flow on species
454 tree estimation: a simulation study. *Syst. Biol.* 63, 17-30.
- 455 Lévillé-Bourret, É., Starr, J.R., Ford, B.A., Moriarty Lemmon, E., Lemmon, A.R., 2018.
456 Resolving rapid radiations within angiosperm families using anchored phylogenomics. *Syst.*
457 *Biol.* 67, 94-112.
- 458 Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
459 *arXiv preprint arXiv:1303.3997*.

- 460 Mallet, J., Besansky, N., Hahn, M.W., 2016. How reticulated are species? *Bioessays* 38, 140-
461 149.
- 462 McCartney-Melstad, E., Mount, G.G., Bradley Shaffer, H., 2016. Exon capture optimization
463 in amphibians with large genomes. *Mol. Ecol. Resour.* 16, 1084-1094.
- 464 McCartney-Melstad, E., Vu, J.K., Shaffer, H.B., 2018. Genomic data recover previously
465 undetectable fragmentation effects in an endangered amphibian. *Mol. Ecol.*,
466 <https://doi.org/10.1111/mec.14892>.
- 467 McCormack, J.E., Faircloth, B.C., Crawford, N.G., Gowaty, P.A., Brumfield, R.T., Glenn, T.C.,
468 2012. Ultraconserved elements are novel phylogenomic markers that resolve placental
469 mammal phylogeny when combined with species-tree analysis. *Genome Res.* 22, 746-754.
- 470 McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella,
471 K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The Genome Analysis Toolkit:
472 A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.*
473 20, 1297-1303.
- 474 Meiklejohn, K.A., Faircloth, B.C., Glenn, T.C., Kimball, R.T., Braun, E.L., 2016. Analysis of
475 a rapid evolutionary radiation using ultraconserved elements: evidence for a bias in some
476 multispecies coalescent methods. *Syst. Biol.* 65, 612-627.
- 477 Meilink, W.R.M., Arntzen, J.W., van Delft, J.J.C.W., Wielstra, B., 2015. Genetic pollution of
478 a threatened native crested newt species through hybridization with an invasive congener in the
479 Netherlands. *Biol. Conserv.* 184, 145-153.
- 480 Nater, A., Burri, R., Kawakami, T., Smeds, L., Ellegren, H., 2015. Resolving evolutionary
481 relationships in closely related species with whole-genome sequencing data. *Syst. Biol.* 64,
482 1000-1017.
- 483 Neves, L.G., Davis, J.M., Barbazuk, W.B., Kirst, M., 2013. Whole-exome targeted sequencing
484 of the uncharacterized pine genome. *The Plant Journal* 75, 146-156.
- 485 Nowoshilow, S., Schloissnig, S., Fei, J.-F., Dahl, A., Pang, A.W.C., Pippel, M., Winkler, S.,
486 Hastie, A.R., Young, G., Roscito, J.G., Falcon, F., Knapp, D., Powell, S., Cruz, A., Cao, H.,
487 Habermann, B., Hiller, M., Tanaka, E.M., Myers, E.W., 2018. The axolotl genome and the
488 evolution of key tissue formation regulators. *Nature* 554, 50-55.
- 489 Pamilo, P., Nei, M., 1988. Relationships between gene trees and species trees. *Mol. Biol. Evol.*
490 5, 568-583.
- 491 Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T.J., Manuel, M., Wörheide, G.,
492 Baurain, D., 2011. Resolving difficult phylogenetic questions: why more sequences are not
493 enough. *PLoS Biol.* 9, e1000602.
- 494 Pollard, D.A., Iyer, V.N., Moses, A.M., Eisen, M.B., 2006. Widespread discordance of gene
495 trees with species tree in *Drosophila*: evidence for incomplete lineage sorting. *PLoS Genet.* 2,
496 e173.
- 497 Poplin, R., Ruano-Rubio, V., DePristo, M.A., Fennell, T.J., Carneiro, M.O., Van der Auwera,
498 G.A., Kling, D.E., Gauthier, L.D., Levy-Moonshine, A., Roazen, D., Shakir, K., Thibault, J.,

- 499 Chandran, S., Whelan, C., Lek, M., Gabriel, S., Daly, M.J., Neale, B., MacArthur, D.G., Banks,
500 E., 2017. Scaling accurate genetic variant discovery to tens of thousands of samples. bioRxiv.
- 501 Rivera, M.C., Jain, R., Moore, J.E., Lake, J.A., 1998. Genomic evidence for two functionally
502 distinct gene classes. Proc. Natl. Acad. Sci. U.S.A. 95, 6239-6244.
- 503 Sambrook, J., Russell, D.W., 2001. Molecular cloning: a laboratory manual 3rd edition.
504 ColdSpring-Harbour Laboratory Press, UK.
- 505 Scott, P.A., Glenn, T.C., Rissler, L.J., 2018. Resolving taxonomic turbulence and uncovering
506 cryptic diversity in the musk turtles (*Sternotherus*) using robust demographic modeling. Mol.
507 Phylogenet. Evol. 120, 1-15.
- 508 Shaffer, H.B., Minx, P., Warren, D.E., Shedlock, A.M., Thomson, R.C., Valenzuela, N.,
509 Abramyan, J., Amemiya, C.T., Badenhorst, D., Biggar, K.K., Borchert, G.M., Botka, C.W.,
510 Bowden, R.M., Braun, E.L., Bronikowski, A.M., Bruneau, B.G., Buck, L.T., Capel, B., Castoe,
511 T.A., Czerwinski, M., Delehaunty, K.D., Edwards, S.V., Fronick, C.C., Fujita, M.K., Fulton,
512 L., Graves, T.A., Green, R.E., Haerty, W., Hariharan, R., Hernandez, O., Hillier, L.W.,
513 Holloway, A.K., Janes, D., Janzen, F.J., Kandoth, C., Kong, L., de Koning, A.J., Li, Y.,
514 Litterman, R., McGaugh, S.E., Mork, L., O'Laughlin, M., Paitz, R.T., Pollock, D.D., Ponting,
515 C.P., Radhakrishnan, S., Raney, B.J., Richman, J.M., St John, J., Schwartz, T., Sethuraman,
516 A., Spinks, P.Q., Storey, K.B., Thane, N., Vinar, T., Zimmerman, L.M., Warren, W.C., Mardis,
517 E.R., Wilson, R.K., 2013. The western painted turtle genome, a model for the evolution of
518 extreme physiological adaptations in a slowly evolving lineage. Genome Biol. 14, R28.
- 519 Shi, C.-M., Yang, Z., 2018. Coalescent-based analyses of genomic sequence data provide a
520 robust resolution of phylogenetic relationships among major groups of gibbons. Mol. Biol.
521 Evol. 35, 159-179.
- 522 Shine, R., Shetty, S., 2001. Moving in two worlds: aquatic and terrestrial locomotion in sea
523 snakes (*Laticauda colubrina*, Laticaudidae). J. Evol. Biol. 14, 338-346.
- 524 Slijepčević, M., Galis, F., Arntzen, J.W., Ivanović, A., 2015. Homeotic transformations and
525 number changes in the vertebral column of *Triturus* newts. PeerJ 3, e1397.
- 526 Smith, J.J., Timoshevskaya, N., Timoshevskiy, V.A., Keinath, M.C., Hardy, D., Voss, S.R.,
527 2018. A chromosome-scale assembly of the enormous (32 Gb) *Axolotl* genome. bioRxiv.
- 528 Song, S., Liu, L., Edwards, S.V., Wu, S., 2012. Resolving conflict in eutherian mammal
529 phylogeny using phylogenomics and the multispecies coalescent model. Proc. Natl. Acad. Sci.
530 U.S.A. 109, 14942-14947.
- 531 Spinks, P.Q., Thomson, R.C., Pauly, G.B., Newman, C.E., Mount, G., Shaffer, H.B., 2013.
532 Misleading phylogenetic inferences based on single-exemplar sampling in the turtle genus
533 *Pseudemys*. Mol. Phylogenet. Evol. 68, 269-281.
- 534 Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
535 large phylogenies. Bioinformatics 30, 1312-1313.
- 536 Stroud, J.T., Losos, J.B., 2016. Ecological opportunity and adaptive radiation. Annu. Rev. Ecol.
537 Evol. Syst. 47, 507-532.

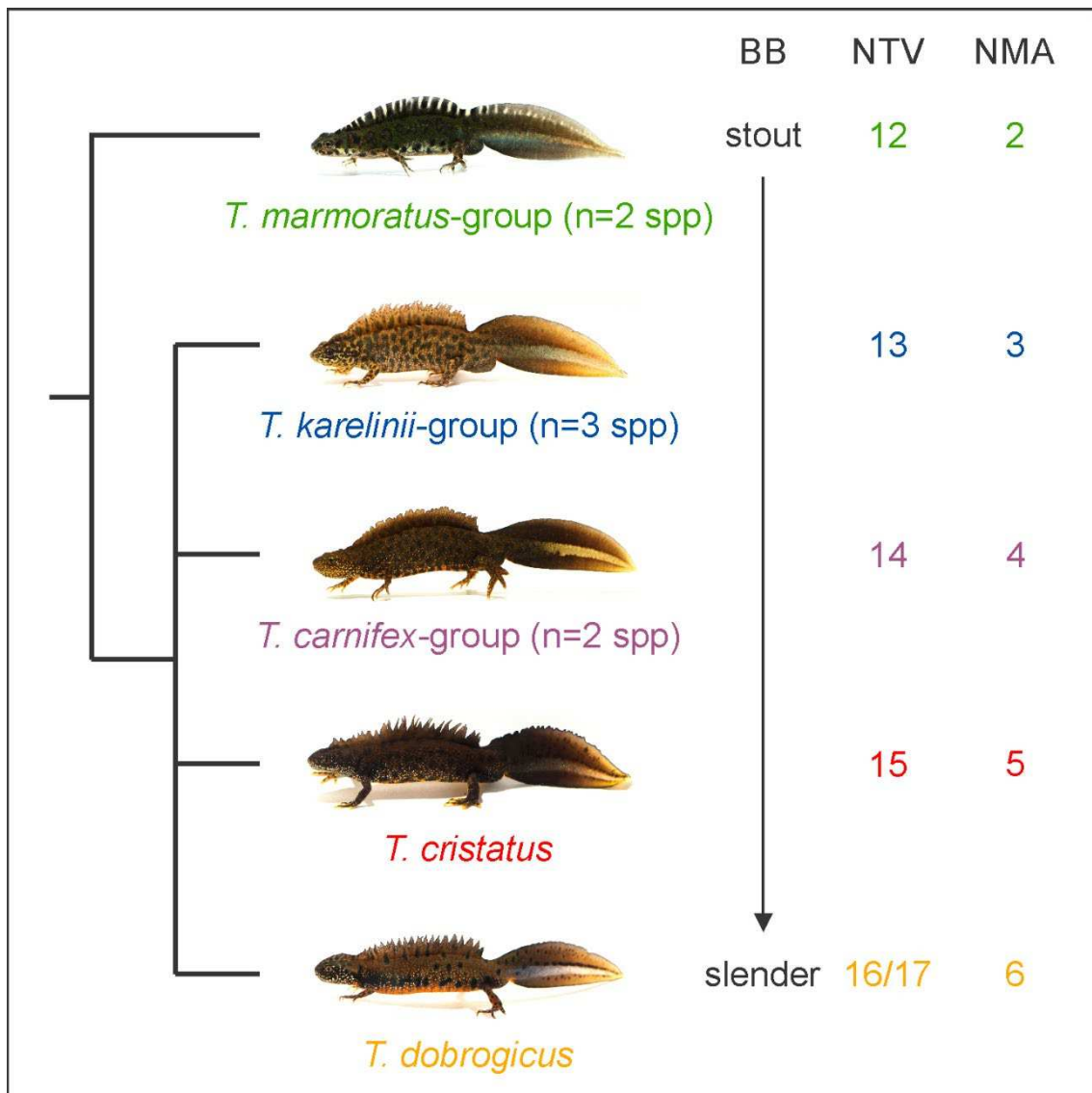
- 538 Sun, Y.-B., Xiong, Z.-J., Xiang, X.-Y., Liu, S.-P., Zhou, W.-W., Tu, X.-L., Zhong, L., Wang,
539 L., Wu, D.-D., Zhang, B.-L., Zhu, C.-L., Yang, M.-M., Chen, H.-M., Li, F., Zhou, L., Feng,
540 S.-H., Huang, C., Zhang, G.-J., Irwin, D., Hillis, D.M., Murphy, R.W., Yang, H.-M., Che, J.,
541 Wang, J., Zhang, Y.-P., 2015. Whole-genome sequence of the Tibetan frog *Nanorana parkeri*
542 and the comparative evolution of tetrapod genomes. *Proc. Natl. Acad. Sci. U.S.A.* 112, E1257.
- 543 Swofford, D.L., Waddell, P.J., Huelsenbeck, J.P., Foster, P.G., Lewis, P.O., Rogers, J.S., 2001.
544 Bias in phylogenetic estimation and Its relevance to the choice between parsimony and
545 likelihood methods. *Syst. Biol.* 50, 525-539.
- 546 Tatusov, R.L., Koonin, E.V., Lipman, D.J., 1997. A genomic perspective on protein families.
547 *Science* 278, 631.
- 548 Towns, J., Cockerill, T., Dahan, M., Foster, I., Gaither, K., Grimshaw, A., Hazlewood, V.,
549 Lathrop, S., Lifka, D., Peterson, G.D., Roskies, R., Scott, J.R., Wilkins-Diehr, N., 2014.
550 XSEDE: accelerating acientific discovery. *Computing in Science & Engineering* 16, 62-74.
- 551 Urošević, A., Slijepčević, M.D., Arntzen, J.W., Ivanović, A., 2016. Vertebral shape and body
552 elongation in *Triturus* newts. *Zoology* 119, 439-446.
- 553 Veith, M., Bogaerts, S., Pasmans, F., Kieren, S., 2018. The changing views on the evolutionary
554 relationships of extant Salamandridae (Amphibia: Urodela). *PLoS ONE* 13, e0198237.
- 555 Vukov, T.D., Sotiropoulos, K., Wielstra, B., Džukić, G., Kalezić, M.L., 2011. The evolution
556 of the adult body form of the crested newt (*Triturus cristatus* superspecies, Caudata,
557 Salamandridae). *Journal of Zoological Systematics and Evolutionary Research* 49, 324-334.
- 558 Whitfield, J.B., Lockhart, P.J., 2007. Deciphering ancient rapid radiations. *Trends Ecol. Evol.*
559 22, 258-265.
- 560 Wielstra, B., Arntzen, J.W., 2011. Unraveling the rapid radiation of crested newts (*Triturus*
561 *cristatus* superspecies) using complete mitogenomic sequences. *BMC Evol. Biol.* 11, 162.
- 562 Wielstra, B., Arntzen, J.W., van der Gaag, K., Pabijan, M., Babik, W., 2014. Data
563 concatenation, Bayesian concordance and coalescent-based analyses of the species tree for the
564 rapid radiation of *Triturus* newts. *PLoS ONE* 9, e111011.
- 565 Wielstra, B., Burke, T., Butlin, R.K., Arntzen, J.W., 2017a. A signature of dynamic
566 biogeography: enclaves indicate past species replacement. *Proc. Royal Soc. B* 284, 20172014.
- 567 Wielstra, B., Burke, T., Butlin, R.K., Avci, A., Üzümlü, N., Bozkurt, E., Olgun, K., Arntzen,
568 J.W., 2017b. A genomic footprint of hybrid zone movement in crested newts. *Evolution Letters*
569 1, 93-101.
- 570 Wielstra, B., Crnobrnja-Isailović, J., Litvinchuk, S.N., Reijnen, B.T., Skidmore, A.K.,
571 Sotiropoulis, K., Toxopeus, A.G., Tzankov, N., Vukov, T., Arntzen, J.W., 2013. Tracing
572 glacial refugia of *Triturus* newts based on mitochondrial DNA phylogeography and species
573 distribution modeling. *Front. Zool.* 10, 13.
- 574 Wielstra, B., Vörös, J., Arntzen, J.W., 2016. Is the Danube crested newt *Triturus dobrogicus*
575 polytypic? A review and new nuclear DNA data. *Amphib.-Reptil.* 37, 167-177.

576 Zhang, C., Sayyari, E., Mirarab, S., 2017. ASTRAL-III: Increased Scalability and Impacts of
577 Contracting Low Support Branches. In: Meidanis, J., Nakhleh, L. (Eds.), *Comparative*
578 *Genomics*. Springer International Publishing, Cham, pp. 53-75.

579

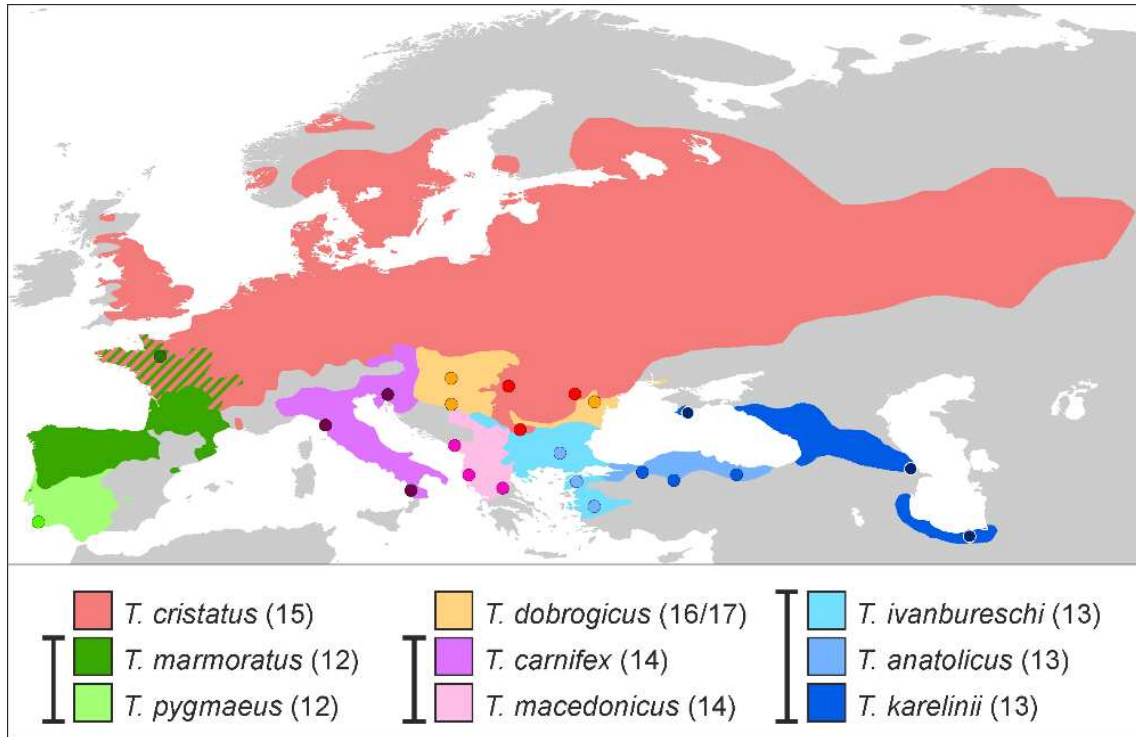
580 **Figures**

581



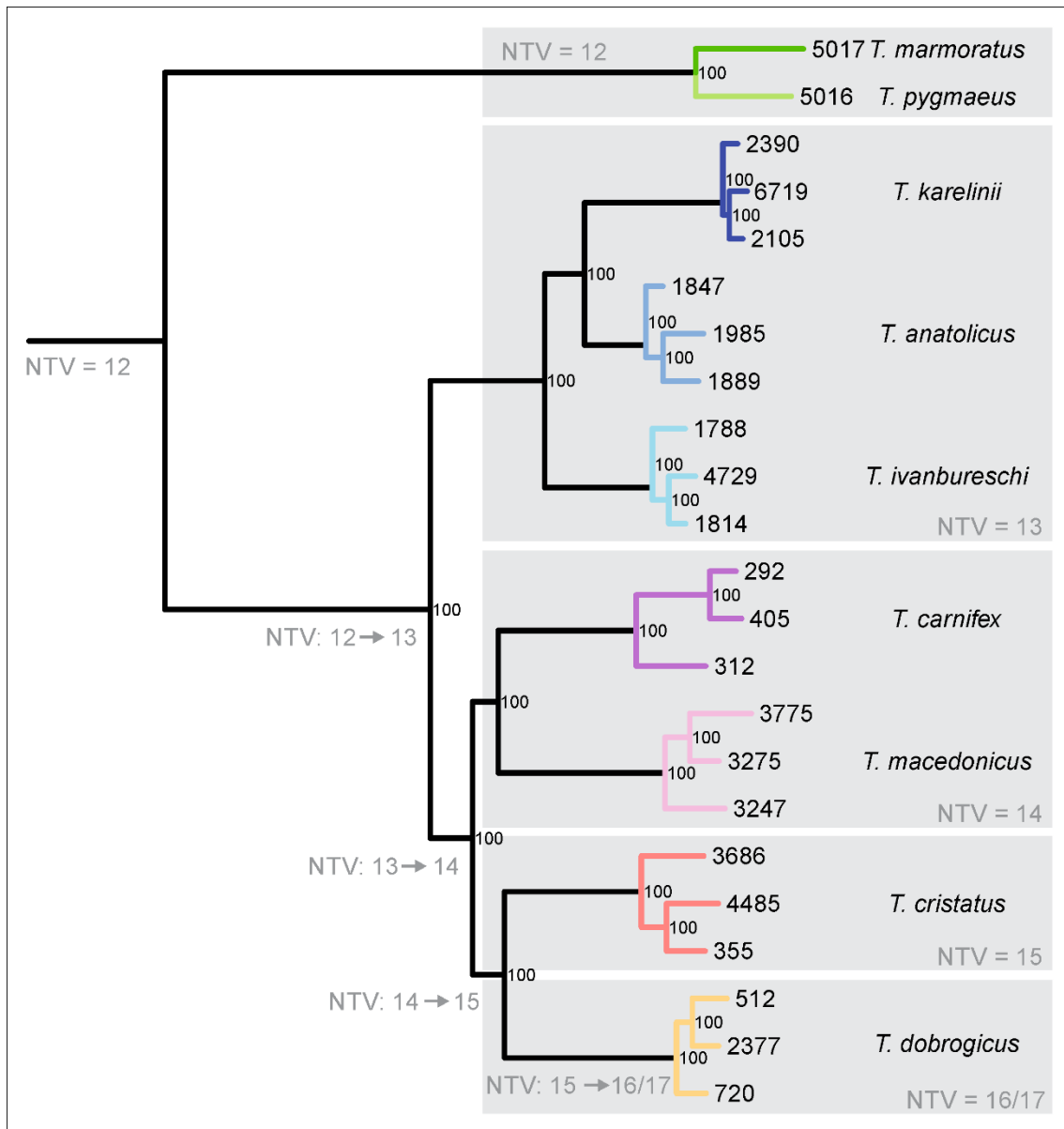
582

583 **Fig. 1. The adaptive radiation of *Triturus* newts.** Five body builds (BB) from stout to slender
 584 are observed in *Triturus* that are also characterized by an increasing number of trunk vertebrae
 585 (NTV) and number of annual aquatic months (NMA). The marbled newts (*T. marmoratus*-
 586 group) and crested newts (remaining four BBs) are sister clades. Relationships among the
 587 crested newts are not yet resolved and are the main focus of the present study.



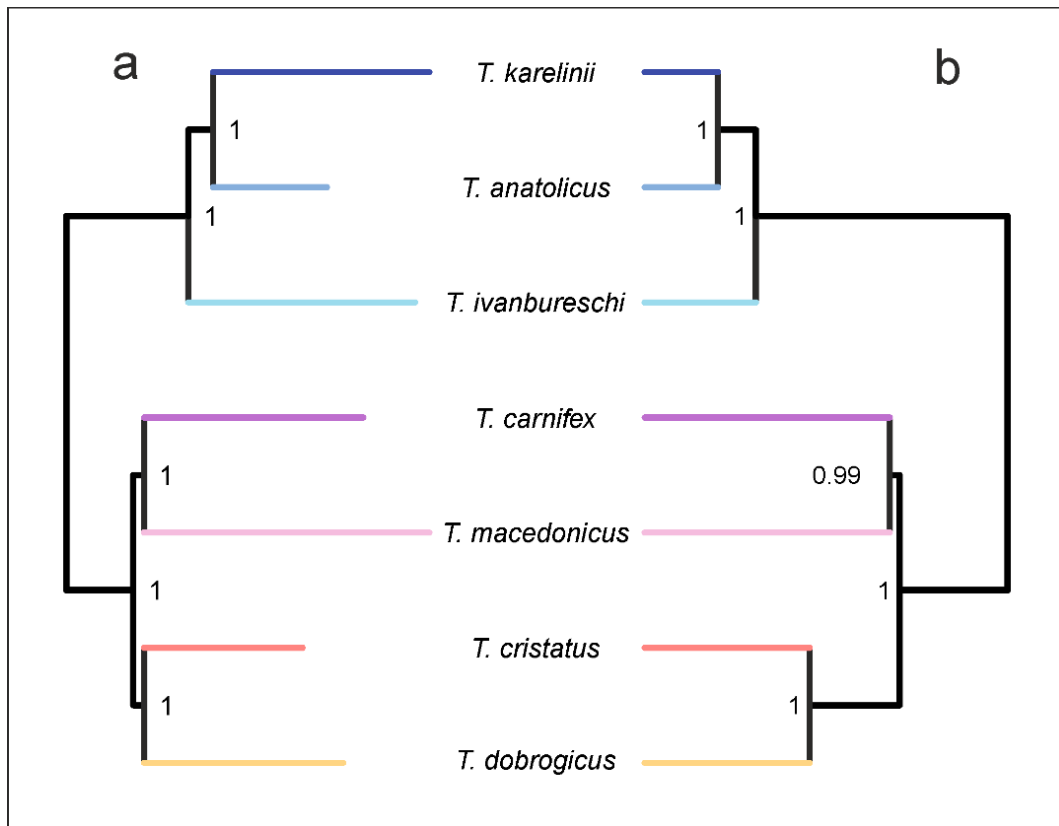
588

589 **Fig. 2. Distribution and sampling scheme for *Triturus*.** Dots represent sample localities
 590 (details in Supplementary Table S1). For the marbled newts (in green) a single individual is
 591 sampled for each of the two species and for the crested newts (other colours) three individuals
 592 are sampled for all seven species. The number in parentheses reflects each species'
 593 characteristic number of trunk vertebrae and whiskers link species that possess the same body
 594 build (see Fig. 1).



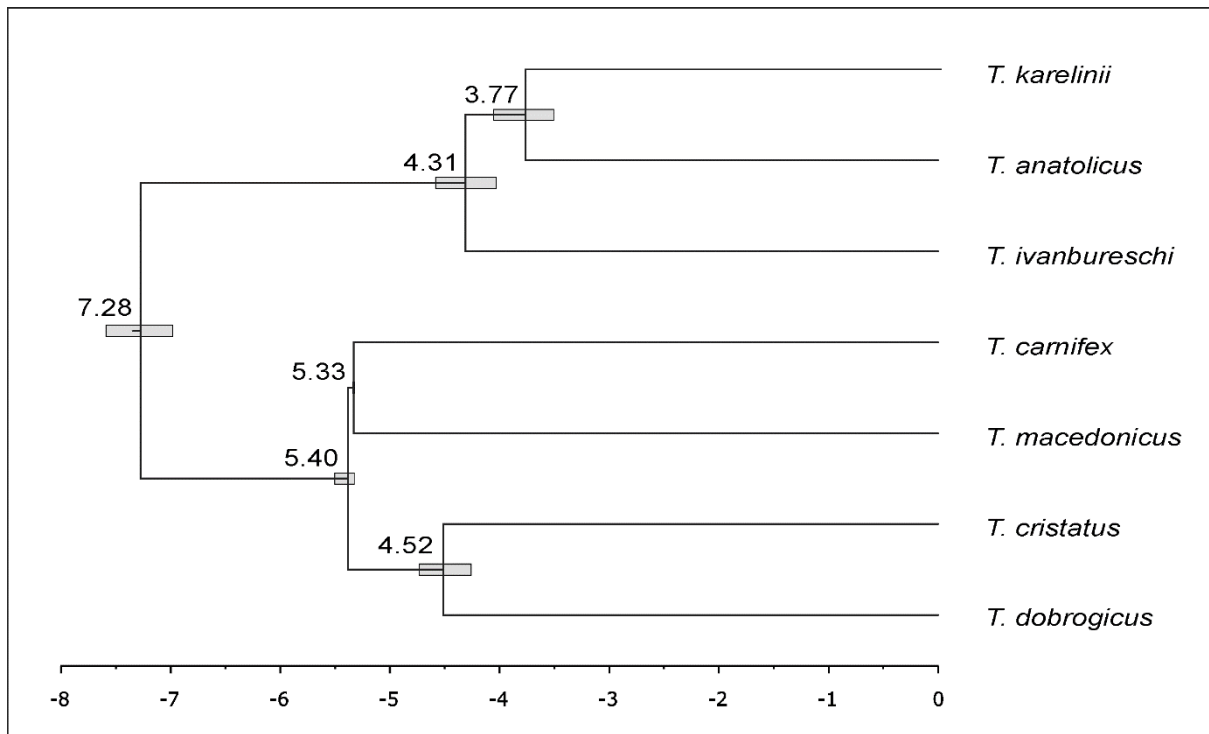
595

596 **Fig. 3. *Triturus* newt phylogeny based on data concatenation with RAxML.** This maximum
 597 likelihood phylogeny is based on 133,601 SNPs derived from 5,866 nuclear markers. Numbers
 598 at nodes indicate bootstrap support from 100 rapid bootstrap replicates. The five *Triturus* body
 599 builds (see Fig. 1) are delineated by grey boxes, with their characteristic number of trunk
 600 vertebrae (NTV) noted. Inferred changes in NTV under the parsimony criterion are noted along
 601 branches. Colours reflect species and correspond to Fig. 2. Tip labels correspond to
 602 Supplementary Table S1.



603

604 **Fig. 4. Crested newt phylogeny based on gene-tree summary with ASTRAL and species-**
 605 **tree estimation with SNAPP.** The ASTRAL tree (a) is based on 5,610 gene-trees. Numbers
 606 at nodes indicate local quartet support posterior probabilities. The SNAPP tree (b) is based on
 607 single biallelic SNPs taken from 5,581 nuclear markers. Numbers at nodes indicate posterior
 608 probabilities. Colours reflect species and correspond to Fig. 2. Note that both topologies are
 609 identical to the phylogeny based on data concatenation (Fig. 3).



610

611 **Fig. 5. Dated species-tree for the crested newts.** Divergence times were determined with
 612 SNAPP, using a single *T. carnifex*–*T. macedonicus* inferred split date of 5.33 million years ago
 613 as a calibration point. Numbers at nodes reflect median divergence times in millions of years
 614 ago and bars the 95% credibility interval around the median.

615

616 **Phylogenomics of the adaptive radiation of *Triturus* newts supports gradual ecological niche**
617 **expansion towards an incrementally aquatic lifestyle**

618

619 B. Wielstra, E. McCartney-Melstad, J.W. Arntzen, R.K. Butlin, H.B. Shaffer

620

621 **SUPPLEMENTARY TEXT, FIGURES AND TABLES**

622

623 **SUPPLEMENTAL TEXT S1-S3**

624

625 **Text S1: Array Design**

626

627 *Transcriptome sequencing* – Liver tissue samples in RNAlater from ten newts (one each of *Triturus*
628 *anatolicus*, *T. carnifex*, *T. cristatus*, *T. dobrogicus*, *T. ivanbureschi*, *T. karelinii*, *T. macedonicus*, *T.*
629 *marmoratus*, *T. pygmaeus*, and *Ommatotriton nesterovi*; Supplementary Table S1) were sent to ZF-
630 Genomics (Leiden, The Netherlands) for RNA extraction and sequencing on a HiSeq 2500. Samples
631 received an average of 43,810,415 clusters (SD=9,744,176) in 150bp paired-end configuration.

632

633 *QC and Assembly* – Paired-end sequencing reads were trimmed for adapter contamination and sequence
634 quality using a 4-bp sliding window in Trimmomatic 0.32 (Bolger, et al. 2014), clipping the 3' ends of
635 reads when the average sequence quality within the window dropped below 20. Leading bases with a
636 quality score less than 5 and trailing bases with a quality score less than 15 were also removed, and
637 reads shorter than 40bp after trimming were discarded.

638 A median of 38,575,204 read pairs were input into the Trinity assembler for each of the ten
639 species (min=27,572,854, max=54,993,188, sd=8,916,227), and a median of 18.6% of these were

640 retained after *in silico* normalization (min=15.8%, max=22.8%, sd=2.3%). Each transcriptome was
641 individually assembled using Trinity 2.2.0 with read coverage normalized to a maximum of 50
642 (Grabherr, et al. 2011). Individual Trinity assemblies were clustered at 90% identity using usearch
643 v9.1.13 to reduce redundancy (Edgar 2010). Assemblies contained a median of 157,608 contigs after
644 clustering at 90% similarity (min=80,803 for *T. karelinii*, and max=182,488 for *T. carnifex*).

645 These clustered assemblies were then used for pairwise comparison between *T. dobrogicus* and
646 the other nine species using *blastn* v2.2.30 (Camacho, et al. 2009). The reciprocal best blast hits (RBBH)
647 method was used to determine presumptive orthology between the assembled transcripts for each
648 pairwise species comparison (Tatusov, et al. 1997; Bork, et al. 1998). *T. dobrogicus* transcripts that
649 returned reciprocal best blast hits to all of the nine other species were retained and all other
650 transcripts were discarded.

651

652 *Transcriptome comparison* – The remaining set of 10,333 *T. dobrogicus* transcripts was self-blasted to
653 attempt to reduce redundancy, which may help reduce the inclusion of multiple isoforms of the same
654 gene, chimeric transcripts assembled by Trinity, and transcripts with truly similar regions that may
655 complicate downstream bioinformatics. As a conservative measure, both the subject and query
656 transcript were discarded if any transcript showed significant similarity (blast e-value < 0.001) to a
657 different transcript or to different regions of itself.

658

659 *Annotation* – The remaining set of 9,214 *T. dobrogicus* transcripts were annotated using a translated
660 blastx search to known *X. tropicalis* proteins with an e-value cutoff of 0.1 (Hellsten, et al. 2010).
661 Transcripts that did not have a positive blastx hit to the *Xenopus* protein database were discarded, and
662 only a single transcript annotating to a particular *Xenopus* protein was retained.

663

664 *Splice site prediction* – For the remaining set of 7,228 *T. dobrogicus* transcripts we attempted to infer
665 splice sites in the candidate targets to avoid designing baits that span such boundaries, as these baits
666 may perform poorly (Neves, et al. 2013) and because targeting a single exon for each transcript
667 simplifies downstream analyses. Splice sites were predicted by attempting to map each transcript to the
668 *Chrysemys picta* genome (Shaffer, et al. 2013) using exonerate’s est2genome model (Slater and Birney
669 2005) with a DNA word length of 10. Approximately 93% of all transcripts (n=6,758) successfully
670 mapped to the *C. picta* genome, and for regions that mapped, the longest contiguous section of the
671 mapped transcript was harvested. If the longest contiguous segment was less than 200bp, the first high-
672 scoring segment pair (HSP) was extended towards the 5’ end until reaching 200bp, followed by
673 extending the final HSP towards the 3’ end until reaching 200bp if necessary. Of the 6,758 transcripts
674 that mapped to *C. picta*, 69 transcripts did not have an HSP longer than 200bp and could not be extended
675 to 200bp in the 5’ or 3’ direction and were dropped as prospective targets.

676 The 470 transcripts that did not align to the *C. picta* genome were sequentially aligned to the
677 genomes of *X. tropicalis* (Hellsten, et al. 2010), *Nanorana parkerii* (Sun, et al. 2015), and *Rana*
678 *catesbeiana* (Hammond, et al. 2017) to attempt to find splice sites, taking the first successful species
679 alignment from the list. Of these 470 transcripts, 125 mapped to *X. tropicalis*, 39 mapped to *N. parkerii*,
680 and 36 mapped to *R. catesbeiana*. Again the longest contiguous aligned segment of each transcript
681 was retained as a possible target, and transcripts with no aligned HSP of at least 200bp had their
682 alignments extended in the 5’ then 3’ directions to attain targets of at least 200bp. For the 270
683 transcripts that did not map to any genome, the first (leftmost) 300bp of the assembled transcript was
684 selected for a target region (except for the one transcript that was only 231bp long—for this target

685 the entire 231bp transcript was used). It is possible that this leftmost orientation may enrich these
686 targets for UTR sequence, assuming that the transcript was fully assembled by Trinity.

687 All exon targets were trimmed to a maximum of 450bp (from the 3' edge) and checked again
688 for complementarity using a self BLAST in blastn. The first qualifying target from each unique Trinity
689 cluster-gene identifier was retained, and any other targets that arose from the same Trinity gene
690 identifier were discarded (n=19). This target set contains sub-sequences from 7,139 different
691 transcripts for a total length of 2,272,851bp (mean of each sub-sequence=318bp, min=200bp,
692 max=400bp, median=300bp).

693 As we are interested in capturing these loci from all *Triturus* taxa, including both crested and
694 marbled newts, we decided to include probes designed from multiple species for the same target if
695 divergence between representative species in the two main clades was greater than 5% (Bi, et al.
696 2012). Since the bulk of the target sequences were designed from *T. dobrogicus*, which together with
697 *T. carnifex*, *T. cristatus* and *T. macedonicus* encompasses one of two main clades in the crested newts
698 (Wielstra and Arntzen 2011; Wielstra, Arntzen, et al. 2014), the three remaining species of crested
699 newts encompassing the other clade (*T. karelinii*, *T. anaticus*, and *T. ivanbureschi*) were used to
700 determine if greater than 5% divergence existed between the two major clades for that target. First,
701 the *T. dobrogicus* targets were blasted against *T. karelinii*, enforcing a full-length HSP with respect to
702 the query sequence, yielding 2,850 hits; 30 of these were found to have a divergence greater than 5%
703 and were added to the 7,139 *T. dobrogicus* targets. Then the remaining 4,289 *T. dobrogicus* targets
704 were blasted to *T. anaticus*, yielding 2,883 hits and an additional 35 targets. Finally the remaining
705 1,406 *T. dobrogicus* targets were blasted to *T. ivanbureschi*, yielding 631 hits and 10 more targets.
706 Subsequently the process was repeated for the marbled newts *T. pygmaeus* and *T. marmoratus*, which
707 constitute the sister lineage of the two crested newt clades, yielding an additional 222 and 27 targets
708 after positive hits for 5,544 of 7,139 targets and 440 of 1,595 residual targets, respectively. Overall, an
709 additional 324 orthologous targets that were more than 5% divergent between *T. dobrogicus* and

710 other *Triturus* species were added to attempt to generate a set of probes that would perform well
711 across the genus.

712 A set of 7,463 target sequences (average length=317bp, min=175bp, max=474bp) was sent to
713 Arbor Biosciences for probe tiling and synthesis. After removing any probes softmasked by
714 RepeatMasker and the Amphibia database, 39,143 unique 120 bp RNA probes were synthesized at
715 approximately 2.6X tiling density across 7,418 target sequences by Arbor Biosciences (Ann Arbor, MI)
716 as a MyBaits-II kit.

717

718 *Test for phylogenetic utility* – The phylogenetic utility of the genomic transcript markers was validated
719 by building a phylogeny from the transcript sequences with RAxML. Trinity-assembled transcriptomes
720 were clustered at 90% identity using usearch v9.1.13 (Edgar 2010), and the sequence capture targets
721 were aligned to these clusters using blastn v2.2.31 (Camacho, et al. 2009). The sequences corresponding
722 to each target were extracted for each sample and aligned using mafft v7.313 (Kato and Standley 2013)
723 and all 7,139 sequence alignments (1 per target) were concatenated. RAxML v8.2.11 (Stamatakis 2014)
724 was used to generate a maximum likelihood phylogeny using 100 rapid bootstrap replicates and the
725 GTRCAT model of sequence evolution. Results suggested sufficient phylogenetic resolution, but one
726 unexpected finding was the placement of *T. carnifex* as the sister lineage to *T. dobrogicus*
727 (Supplementary Fig. S1a). Yet, in our main experiment, *T. carnifex* was more closely related to *T.*
728 *macedonicus* (see Results). The fact that the *T. carnifex* sample used for transcriptome sequencing
729 originated from close to the documented hybrid zone with *T. dobrogicus* (Arntzen, et al. 2014; Wielstra,
730 Sillero, et al. 2014) suggests that substantial interspecific gene flow might underlie this relationship. To
731 further explore this scenario we obtained transcriptomes from two additional *T. carnifex* individuals,
732 sampled away from the hybrid zone with *T. dobrogicus*, representing the distinct Balkan and Italian
733 mtDNA clades (Canestrelli, et al. 2012; Wielstra, et al. 2013). We processed these two individuals as
734 above and reran RAxML, replacing the *T. carnifex* sample from the hybrid zone, and found that *T.*
735 *carnifex* was recovered as the sister lineage to *T. macedonicus* (Supplementary Fig. S1b). Assuming
736 that the *T. carnifex-T. macedonicus* relationship is correct, this phylogenetic shift reflects both the

737 general risk of single-exemplar sampling (Spinks, et al. 2013) and the distorting influence that
738 interspecific gene flow can have on phylogenetic inference (Leaché, et al. 2014). These findings support
739 our decision to include multiple samples per species and to exclude samples from near known hybrid
740 zones in our main experiment.

741

742 **Text S2: Processing of Sequence Capture Data**

743

744 *Reference assembly* – Sequence reads from the sample with the most reads (*T. carnifex* 292 with
745 3,937,346 read pairs) were used *de novo* to assemble target sequences for each target region.
746 Trimmomatic v0.36 (Bolger, et al. 2014) was first used to remove adapter contamination and to trim
747 leading bases with scores < 5, trailing bases with scores < 15, also employing a 4bp sliding window
748 from 5' to 3', trimming the window and downstream sequence when the average quality of the
749 window dropped < 20. Reads < 40bp were discarded. Trimmed reads were input into PEAR v0.9.10
750 (Zhang, et al. 2014) to merge overlapping paired end reads into longer single-end fragments with the
751 following settings: p-value = 0.01, minimum assembly length = 50, statistical method = OES, using
752 empirical frequencies = YES, quality score threshold = 0, minimum overlap = 10, and scaling method =
753 scaled score.

754 Unmerged reads and merged read pairs were input into the assembly by reduced complexity
755 (ARC) pipeline (Hunter, et al. 2015), which performs alternating tasks of mapping reads to target
756 sequences, followed by per-target *de novo* assembly of mapped reads, replacing the original target
757 sequences with the target assembly at each iteration. Six iterations were performed to generate a set
758 of reference contigs assembled from reads relevant to each target region. A single assembled contig
759 was then selected for each original target region by means of reciprocal best blast hit (RBBH) (Rivera,
760 et al. 1998). These RBBHs were then blasted against one another to determine self-complementary
761 regions, which can indicate chimeric assembly regions, and regions found to be similar to other target

762 regions were trimmed to the nearest terminus of the contig (McCartney-Melstad, et al. 2016). This set
763 of chimera-trimmed RBBHs was used as a target reference assembly for all downstream analyses.

764

765 *QC, SNP calling and genotyping* – Adapter contamination from library DNA inserts < 150bp was
766 removed from reads using skewer v0.2.2 (Jiang, et al. 2014). Reads were mapped to the reference
767 assembly using BWA-MEM v0.7.15-r1140 (Li 2013). Picard tools v2.9.2
768 (<https://broadinstitute.github.io/picard/>) was used to add read group information and mark PCR
769 duplicates, and GATK v3.8 was used to generate gVCFs for each sample using HaplotypeCaller.
770 GenotypeGVCFs was used for groups of samples (crested newts or crested + marbled newts, depending
771 on the analysis) to call SNPs/genotypes, removing SNPs flagged by the following hard filters: QD < 2,
772 MQ < 40, FS > 60, MQRankSum < -12.5, ReadPosRankSum < -8, QUAL < 30 (DePristo, et al. 2011;
773 Poplin, et al. 2017).

774 The *de novo* assembly followed by RBBH approach is susceptible to the inclusion of paralogous
775 loci as putatively single-copy targets. Because fixed differences between paralogues will appear as
776 consistently heterozygous SNPs, we next attempted to remove paralogous targets from our dataset
777 through the use of a Hardy Weinberg Equilibrium (HWE) filter for heterozygote excess. Heterozygote
778 excess p-values were calculated for every SNP using vcftools 0.1.15 (Danecek, et al. 2011), and any
779 target containing at least one SNP with a heterozygote excess p-value less than 0.05 was removed
780 from downstream analysis.

781

782 *Reference assembly and genotyping* – A total of 4,932,636 reads (including 2,579,319 merged read
783 pairs with an average length of 196bp) were used as input in the ARC assembly pipeline. After six
784 iterations of mapping and assembly, 6,970 targets finished with an average of 295 reads apiece
785 (median=167, sd=1,152), and 6,686 of the original targets had RBBHs to the assembly. After self-blast
786 and trimming to remove potentially chimeric assemblies, a reference assembly of 5,593,497bp was
787 generated for subsequent read mapping and SNP calling.

788 A median of 44.1% of trimmed reads aligned to the reference assembly (min=41.0%,
789 max=50.5%), and an average of 22.6% of mapped reads were flagged as PCR duplicates, yielding a
790 median unique reads on target of 34.2% (min=31.3%, max=39.4%). For the 23-sample dataset
791 including the two marbled newt species, a total of 370,007 SNPs were recovered that passed hard
792 filters. Of the 6,686 starting targets, 798 were found to contain at least 1 SNP with a HWE heterozygote
793 excess p-value less than 0.05 and were removed. For the 21-sample dataset that did not contain the
794 marbled newts, a total of 286,691 SNPs passed the hard filters and 814 targets were removed because
795 they failed the HWE filter. Pairwise F84 divergences calculated with Phylip 3.697 (Felsenstein 1989)
796 and based on the 23-sample dataset (including all *Triturus* species) are provided in Supplementary
797 Table S2. The highest intraspecific divergence was observed between the Italian and Balkan clades
798 comprising *T. carnifex*.

799

800 **Text S3: Phylogenetic Analyses**

801

802 *Data concatenation with RAxML* – RAxML version 8.2.11 (Stamatakis 2014) was used to infer
803 phylogenies from concatenated alignments of SNPs. All biallelic SNPs in the 23-sample dataset that
804 had genotype qualities of at least 20, that were present in at least 50% of the samples, and that fit
805 RAxML's definition of variable (133,601 SNPs total across 5,866 different targets) were used for
806 maximum likelihood phylogenetic analysis. 100 rapid bootstrap replicates and 20 maximum likelihood
807 searches were conducted with the ASC_GTRGAMMA model with Lewis ascertainment correction for
808 SNP analysis (Lewis 2001). The resulting phylogeny with bootstrap support values was plotted in R
809 using phytools (Revell 2011).

810 The mean depth of passing genotype calls across all samples was 42.4X, and median per-site
811 missingness was 4.3%, which corresponds to one sample out of 23 missing data for a site (mean=10.1%,
812 sd=14.0%). All crested newt species (for which three individuals were included) were recovered as
813 monophyletic, and all bootstrap values on the tree were 100 (Fig. 3). The longest branch was between

814 the marbled and crested newts and was used to root the tree. Within the crested newts, *T. ivanbureschi*
815 was the sister lineage to a clade consisting of *T. anatolicus* and *T. karelinii*. The remaining four species
816 were the sister-group to this assemblage, with *T. carnifex* most closely related to *T. macedonicus* and *T.*
817 *cristatus* most closely related to *T. dobrogicus*. Since the monophyly of all species was strongly
818 supported, species designations were fixed for subsequent species-tree inference.

819

820 *Gene-tree summary with ASTRAL* – ASTRAL v5.6.1 was used to estimate the crested newt phylogeny
821 and to explore gene-tree discordance, presumably derived primarily from incomplete lineage sorting
822 from a collection of gene-trees (Mirarab, et al. 2014; Sayyari and Mirarab 2016; Zhang, et al. 2017).
823 No marbled newts were included because estimating terminal branch lengths is not possible for species
824 with a single representative (note that the reciprocal monophyly of crested and marbled newts is well
825 established (Arntzen, et al. 2007; Espregueira Themudo, et al. 2009; Wielstra, Arntzen, et al. 2014) and
826 also strongly supported by our concatenated RAxML analysis). Separate polymorphic SNP alignments
827 were first generated for each target using SnpSift 4.3 (Ruden, et al. 2012) and PGDSpider 2.1.1.2
828 (Lischer and Excoffier 2012), omitting SNPs with > 50% missing data across the 21 crested newt
829 samples and removing targets that contained one or more samples with 100% missing data across the
830 target using trimal v1.4.1 (Capella-Gutiérrez, et al. 2009). RAxML v8.2.11 (Stamatakis 2014) was used
831 to infer a maximum likelihood gene-tree for each target with the ASC_GTRGAMMA model and Lewis
832 ascertainment bias correction (Lewis 2001).

833 After setting genotypes with quality scores less than 20 to missing data and filtering out sites
834 with > 50% missing data, a total of 143,571 SNPs remained across 5,861 targets to build gene-trees.
835 After removing targets that contained samples with 100% missing data and removing sites that RAxML
836 determined to be monomorphic, maximum likelihood gene-trees were built for 5,610 targets. These
837 gene-trees were used as input into ASTRAL, constraining the seven crested newt species to be
838 monophyletic (as supported by our concatenated RAxML analysis) and outputting local posterior
839 probabilities and inferring terminal branch lengths. Midpoint rooting was used to determine the root.
840 ASTRAL yielded a final normalized quartet score of 0.63. The same topology as in the concatenated

841 RAxML analysis was recovered, with local posterior probabilities of 1 for all nodes (Fig. 4a). Branch
842 lengths in ASTRAL are measured in coalescent units and indicate the degree of discordance among
843 gene-trees (within taxa for terminal branches and among taxa for internal branches). The longest
844 terminal branch was recovered for *T. macedonicus*, and the shortest belonged to *T. anatolicus*. The
845 shortest internal branches were those separating the sister lineages *T. carnifex* + *T. macedonicus* from
846 *T. cristatus* + *T. dobrogicus*.

847

848 *Species-tree estimation with SNAPP* – The coalescent species-tree inference method SNAPP v1.3.0 was
849 used to infer the crested newt species-tree from biallelic SNPs (Bryant, et al. 2012). Marbled newts
850 were not included because they introduce a long internal branch that can render parameter estimation
851 inaccurate and splits between them and crested newts is not a primary goal of our paper. Polymorphic
852 biallelic SNPs with genotype phred scores ≥ 20 across all 21 crested newts were first collected. Then,
853 a single SNP from each of the 5,581 remaining loci was randomly selected to reduce the impacts of
854 physical genetic linkage. These SNPs were used as input into SNAPP within the BEAST v2.4.8
855 environment (Bouckaert, et al. 2014) with the following parameters: species assignment=7 respective
856 species, mutation rate U=1.0, mutation rate V=1.0, coalescence rate=10.0 (and sampled), use log
857 likelihood correction=True, lambda prior=Gamma (initial=10[0.0,inf]) with alpha=2.0 and beta=200.0,
858 snapprior.alpha=1.0, snapprior.beta=250.0, snapprior.kappa=1.0, snapprior.lambda=10.0 (and
859 sampled), snapprior.rateprior=gamma, chain length=10,000,000, store every=1000 (and logging every
860 1000), and pre burnin=0. A 10% burnin was used and convergence and mixing were assessed with
861 Tracer v1.7.1 (Rambaut, et al. 2018). ESS values for all parameters were > 400 . A maximum clade
862 credibility tree was constructed with common ancestor heights using TreeAnnotator v2.4.8 (Bouckaert,
863 et al. 2014). Note that BEAST infers the root as part of the analysis. The same topology as in the
864 RAxML and ASTRAL analyses was recovered (Fig. 4b). All posterior probabilities were 1, except for
865 the node subtending *T. carnifex* + *T. macedonicus*, which was 0.99.

866

867 *Molecular dating with SNAPP* – A time-calibrated phylogeny was estimated with SNAPP using the same
868 input SNP file as above. For calibration we interpreted the origin of the Adriatic Sea at the end of the
869 Messinian Salinity Crisis at 5.33 million years ago (Krijgsman, et al. 1999) as the vicariance event
870 causing the *T. carnifex* versus *T. macedonicus* split (Arntzen, et al. 2007; Wielstra and Arntzen 2011)
871 and set the age of their most recent common ancestor to a uniform distribution between 5.32 and
872 5.34 million years ago (Stange, et al. 2018). Input XML files for divergence time estimation were
873 prepared using `snapp_prep.rb` (https://github.com/mmatschiner/snapp_prep). We recognize that
874 this is only a rough approximation given a single, biogeographically-informed date calibration point,
875 and use it primarily to estimate the closeness in time of the crested newt radiation. The output tree
876 from the original, undated SNAPP analysis was used as a starting tree, scaling the entire tree so that
877 the starting age of the calibration node was 5.33 million years ago. The topology was fixed to that
878 recovered by the original SNAPP analysis, and dates of remaining nodes were estimated using
879 1,000,000 MCMC steps, sampling every 500 steps and removing a 10% burn-in. ESS values for
880 parameters were confirmed > 400 with Tracer. A maximum clade credibility tree with median node
881 heights was generated with TreeAnnotator (Fig. 5).

882

883 **Text S4: Comparison with full mtDNA-based phylogeny**

884

885 MtDNA has proven misleading at both recent (Rodríguez, et al. 2017) and deeper (Veith, et al. 2018)
886 nodes in the Salamandridae phylogeny and our genome-enabled phylogeny shows a highly supported
887 deviation with a previous full mtDNA (i.e. single marker) phylogeny as well (Wielstra and Arntzen 2011).
888 The deviation concerns the relationship among the three species constituting the '*T. karelinii*-group';
889 we here recover *T. anatolicus* as the sister lineage to *T. karelinii*, rather than to *T. ivanbureschi* as
890 suggested by mtDNA (Supplementary Fig. S2). While such gene-tree discordance could reflect
891 incomplete lineage sorting of mtDNA (Platt, et al. 2018), we consider ancient mtDNA introgression

892 more likely, as *T. ivanbureschi* and *T. anaticus* show geographically extensive introgressive
893 hybridization today (Wielstra, et al. 2017). A scenario of ancient introgression is in line with the high
894 degree of gene-tree/species-tree discordance in the nuclear genome in *T. anaticus*, as suggested by
895 the short branch in the ASTRAL tree (Fig. 4a). However, as all members of the '*T. karelinii*-group'
896 possess an identical number of trunk vertebrae, the mtDNA-nuDNA mismatch does not influence our
897 interpretation of character evolution (Supplementary Fig. S3). The calibrated nuclear DNA-based (Fig.
898 5) and mtDNA-based phylogenies agree that cladogenesis among crested newts occurred over a
899 relatively brief time window. However, mtDNA-based dates are older (cf. Table 2 in Wielstra and
900 Arntzen 2011). This could simply reflect the differences in the dating method and the (slight)
901 differences in the calibration scheme applied, but it is well-known that divergence times derived from
902 individual gene-trees, and particularly from mtDNA, can be overestimates of lineage divergence
903 (McCormack, et al. 2011).

904

905 **Text S5: Inference of changes in the number of trunk vertebrae**

906 The number of trunk vertebrae (NTV) in crested newts is characterized by a punctuated continuous
907 character state distribution, with modal values for NTV in the range of 13-16 (for convenience an NTV
908 count of 16 was used for *T. dobrogicus*, but note that NTV = 17 also occurs at roughly equal frequency
909 in this species, which does not influence our interpretation). We consider NTV = 12, as observed in the
910 sister lineage the marbled newts (the *T. marmoratus*-group), as well as the most closely related genus
911 *Lissotriton*, to be the ancestral state (Arntzen, et al. 2015; Veith, et al. 2018). We applied the parsimony
912 criterion to infer changes in NTV along all possible crested newt topologies (Supplementary Fig. S3).
913 The program PAUP* (Swofford 1998) was used to allocate character state gains and losses over the
914 tree, under ACCTRAN as well as DELTRAN optimization.

915

916 **References**

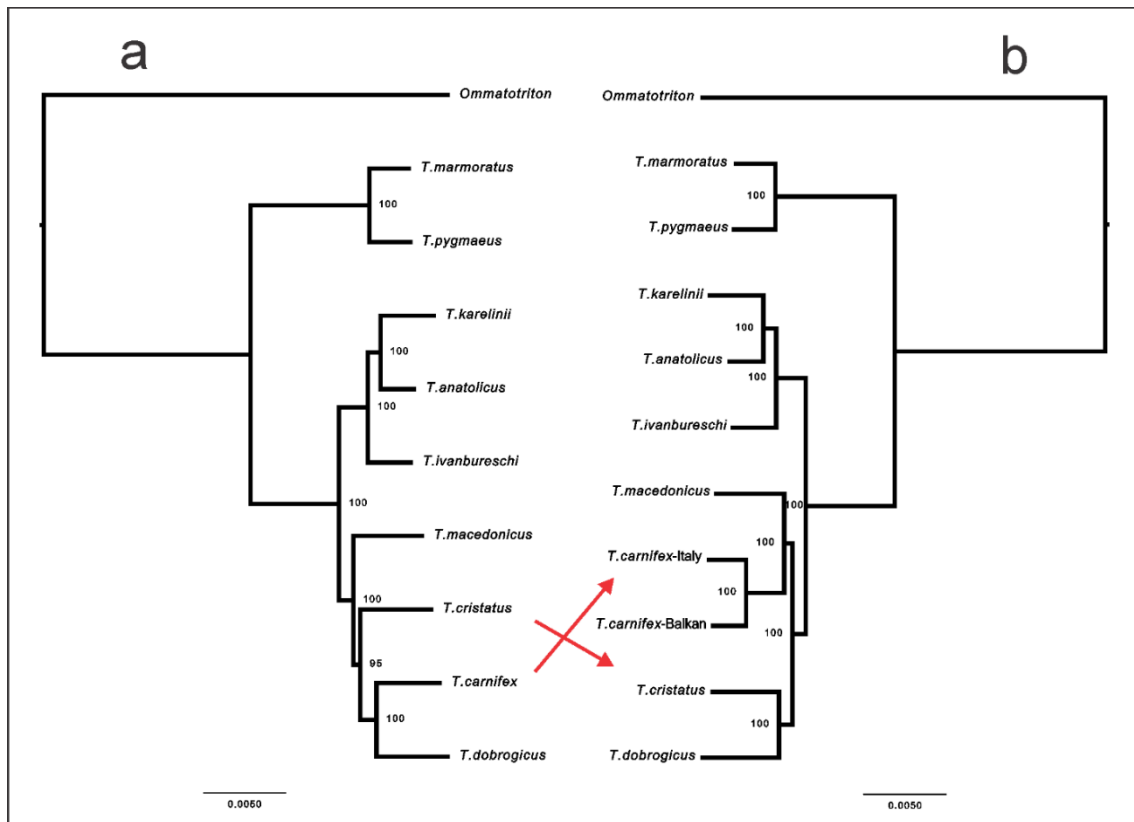
- 917 Arntzen JW, Beukema W, Galis F, Ivanović A. 2015. Vertebral number is highly evolvable in
918 salamanders and newts (family Salamandridae) and variably associated with climatic
919 parameters. *Contributions to Zoology* 84:85-113.
- 920 Arntzen JW, Espregueira Themudo G, Wielstra B. 2007. The phylogeny of crested newts
921 (*Triturus cristatus* superspecies): nuclear and mitochondrial genetic characters suggest a hard
922 polytomy, in line with the paleogeography of the centre of origin. *Contributions to Zoology*
923 76:261-278.
- 924 Arntzen JW, Wielstra B, Wallis GP. 2014. The modality of nine *Triturus* newt hybrid zones,
925 assessed with nuclear, mitochondrial and morphological data. *Biological Journal of the*
926 *Linnean Society* 113:604–622.
- 927 Bi K, Vanderpool D, Singhal S, Linderoth T, Moritz C, Good JM. (Bi2012 co-authors). 2012.
928 Transcriptome-based exon capture enables highly cost-effective comparative genomic data
929 collection at moderate evolutionary scales. *BMC Genomics* 13:403.
- 930 Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence
931 data. *Bioinformatics* 30:2114-2120.
- 932 Bork P, Dandekar T, Diaz-Lazcoz Y, Eisenhaber F, Huynen M, Yuan Y. 1998. Predicting
933 function: from genes to genomes and back. *Journal of Molecular Biology* 283:707-725.
- 934 Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A,
935 Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS*
936 *Computational Biology* 10:e1003537.
- 937 Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A. 2012. Inferring
938 species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent
939 analysis. *Molecular Biology and Evolution* 29:1917-1932.
- 940 Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009.
941 BLAST+: architecture and applications. *BMC Bioinformatics* 10:421-421.
- 942 Canestrelli D, Salvi D, Maura M, Bologna MA, Nascetti G. 2012. One Species, three
943 Pleistocene evolutionary histories: phylogeography of the Italian crested newt, *Triturus*
944 *carnifex*. *PLoS ONE* 7:e41754.
- 945 Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated
946 alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972-1973.
- 947 Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter
948 G, Marth GT, Sherry ST, et al. 2011. The variant call format and VCFtools. *Bioinformatics*
949 27:2156-2158.
- 950 DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del
951 Angel G, Rivas MA, Hanna M, et al. 2011. A framework for variation discovery and
952 genotyping using next-generation DNA sequencing data. *Nature Genetics* 43:491-498.
- 953 Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*
954 26:2460-2461.

- 955 Espregueira Themudo G, Wielstra B, Arntzen JW. 2009. Multiple nuclear and mitochondrial
956 genes resolve the branching order of a rapid radiation of crested newts (*Triturus*,
957 Salamandridae). *Molecular Phylogenetics and Evolution* 52:321-328.
- 958 Felsenstein J. 1989. PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics* 5:164-
959 166.
- 960 Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L,
961 Raychowdhury R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-Seq data
962 without a reference genome. *Nature Biotechnology* 29:644-652.
- 963 Hammond SA, Warren RL, Vandervalk BP, Kucuk E, Khan H, Gibb EA, Pandoh P, Kirk H,
964 Zhao Y, Jones M, et al. 2017. The North American bullfrog draft genome provides insight into
965 hormonal regulation of long noncoding RNA. *Nature Communications* 8:1433.
- 966 Hellsten U, Harland RM, Gilchrist MJ, Hendrix D, Jurka J, Kapitonov V, Ovcharenko I,
967 Putnam NH, Shu S, Taher L, et al. 2010. The genome of the western clawed frog *Xenopus*
968 *tropicalis*. *Science* 328:633.
- 969 Hunter SS, Lyon RT, Sarver BAJ, Hardwick K, Forney LJ, Settles ML. 2015. Assembly by
970 Reduced Complexity (ARC): a hybrid approach for targeted assembly of homologous
971 sequences. *bioRxiv*.
- 972 Jiang H, Lei R, Ding S-W, Zhu S. 2014. Skewer: a fast and accurate adapter trimmer for next-
973 generation sequencing paired-end reads. *BMC Bioinformatics* 15:182.
- 974 Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7:
975 Improvements in Performance and Usability. *Molecular Biology and Evolution* 30:772-780.
- 976 Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS. 1999. Chronology, causes and
977 progression of the Messinian salinity crisis. *Nature* 400:652-655.
- 978 Leaché AD, Harris RB, Rannala B, Yang Z. 2014. The influence of gene flow on species tree
979 estimation: a simulation study. *Systematic Biology* 63:17-30.
- 980 Lewis PO. 2001. A likelihood approach to estimating phylogeny from discrete morphological
981 character data. *Systematic Biology* 50:913-925.
- 982 Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
983 arXiv preprint arXiv:1303.3997.
- 984 Lischer HEL, Excoffier L. 2012. PGDSpider: an automated data conversion tool for connecting
985 population genetics and genomics programs. *Bioinformatics* 28:298-299.
- 986 McCartney-Melstad E, Mount GG, Bradley Shaffer H. 2016. Exon capture optimization in
987 amphibians with large genomes. *Molecular Ecology Resources* 16:1084-1094.
- 988 McCormack JE, Heled J, Delaney KS, Peterson AT, Knowles LL. 2011. Calibrating divergence
989 times on species trees versus gene trees: implications for speciation history of *Aphelocoma* jays.
990 *Evolution* 65:184-202.

- 991 Mirarab S, Reaz R, Bayzid MS, Zimmermann T, Swenson MS, Warnow T. 2014. ASTRAL:
992 genome-scale coalescent-based species tree estimation. *Bioinformatics* 30:i541-i548.
- 993 Neves LG, Davis JM, Barbazuk WB, Kirst M. 2013. Whole-exome targeted sequencing of the
994 uncharacterized pine genome. *The Plant Journal* 75:146-156.
- 995 Platt IIRN, Faircloth BC, Sullivan KAM, Kieran TJ, Glenn TC, Vandeweghe MW, Lee JTE,
996 Baker RJ, Stevens RD, Ray DA. 2018. Conflicting evolutionary histories of the mitochondrial
997 and nuclear genomes in New World *Myotis* bats. *Systematic Biology* 67:236-249.
- 998 Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, Kling
999 DE, Gauthier LD, Levy-Moonshine A, Roazen D, et al. 2017. Scaling accurate genetic variant
1000 discovery to tens of thousands of samples. *bioRxiv*.
- 1001 Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in
1002 Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*:syy032-syy032.
- 1003 Revell LJ. 2011. phytools: an R package for phylogenetic comparative biology (and other
1004 things). *Methods in Ecology and Evolution* 3:217-223.
- 1005 Rivera MC, Jain R, Moore JE, Lake JA. 1998. Genomic evidence for two functionally distinct
1006 gene classes. *Proceedings of the National Academy of Sciences of the United States of America*
1007 95:6239-6244.
- 1008 Rodríguez A, Burgon JD, Lyra M, Irisarri I, Baurain D, Blaustein L, Göçmen B, Künzel S,
1009 Mable BK, Nolte AW, et al. 2017. Inferring the shallow phylogeny of true salamanders
1010 (*Salamandra*) by multiple phylogenomic approaches. *Molecular Phylogenetics and Evolution*
1011 115:16-26.
- 1012 Ruden D, Cingolani P, Patel V, Coon M, Nguyen T, Land S, Lu X. 2012. Using *Drosophila*
1013 *melanogaster* as a model for genotoxic chemical mutational studies with a new program,
1014 SnpSift. *Frontiers in Genetics* 3.
- 1015 Sayyari E, Mirarab S. 2016. Fast coalescent-based computation of local branch support from
1016 quartet frequencies. *Molecular Biology and Evolution* 33:1654-1668.
- 1017 Shaffer HB, Minx P, Warren DE, Shedlock AM, Thomson RC, Valenzuela N, Abramyan J,
1018 Amemiya CT, Badenhorst D, Biggar KK, et al. (Bradley Shaffer2013 co-authors). 2013. The
1019 western painted turtle genome, a model for the evolution of extreme physiological adaptations
1020 in a slowly evolving lineage. *Genome Biology* 14:R28.
- 1021 Slater GSC, Birney E. 2005. Automated generation of heuristics for biological sequence
1022 comparison. *BMC Bioinformatics* 6:31.
- 1023 Spinks PQ, Thomson RC, Pauly GB, Newman CE, Mount G, Shaffer HB. 2013. Misleading
1024 phylogenetic inferences based on single-exemplar sampling in the turtle genus *Pseudemys*.
1025 *Molecular Phylogenetics and Evolution* 68:269-281.
- 1026 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
1027 large phylogenies. *Bioinformatics* 30:1312-1313.

- 1028 Stange M, Sánchez-Villagra MR, Salzburger W, Matschiner M. 2018. Bayesian divergence-
1029 time estimation with genome-wide Single-Nucleotide Polymorphism data of sea catfishes
1030 (Ariidae) supports Miocene closure of the Panamanian isthmus. *Systematic Biology* 67:681-
1031 699.
- 1032 Sun Y-B, Xiong Z-J, Xiang X-Y, Liu S-P, Zhou W-W, Tu X-L, Zhong L, Wang L, Wu D-D,
1033 Zhang B-L, et al. 2015. Whole-genome sequence of the Tibetan frog *Nanorana parkeri* and
1034 the comparative evolution of tetrapod genomes. *Proceedings of the National Academy of*
1035 *Sciences of the United States of America* 112:E1257.
- 1036 Swofford DL. 1998. PAUP 4.0b: Phylogenetic Analysis Using Parsimony. Sunderland: Sinauer
1037 Associates.
- 1038 Tatusov RL, Koonin EV, Lipman DJ. 1997. A genomic perspective on protein families.
1039 *Science* 278:631.
- 1040 Veith M, Bogaerts S, Pasmans F, Kieren S. 2018. The changing views on the evolutionary
1041 relationships of extant Salamandridae (Amphibia: Urodela). *PLoS ONE* 13:e0198237.
- 1042 Wielstra B, Arntzen JW. 2011. Unraveling the rapid radiation of crested newts (*Triturus*
1043 *cristatus* superspecies) using complete mitogenomic sequences. *BMC Evolutionary Biology*
1044 11:162.
- 1045 Wielstra B, Arntzen JW, van der Gaag K, Pabijan M, Babik W. 2014. Data concatenation,
1046 Bayesian concordance and coalescent-based analyses of the species tree for the rapid radiation
1047 of *Triturus* newts. *PLoS ONE* 9:e111011.
- 1048 Wielstra B, Burke T, Butlin RK, Avcı A, Üzümlü N, Bozkurt E, Olgun K, Arntzen JW. 2017. A
1049 genomic footprint of hybrid zone movement in crested newts. *Evolution Letters* 1:93-101.
- 1050 Wielstra B, Crnobrnja-Isailović J, Litvinchuk SN, Reijnen BT, Skidmore AK, Sotiropoulis K,
1051 Toxopeus AG, Tzankov N, Vukov T, Arntzen JW. 2013. Tracing glacial refugia of *Triturus*
1052 newts based on mitochondrial DNA phylogeography and species distribution modeling.
1053 *Frontiers in Zoology* 10:13.
- 1054 Wielstra B, Sillero N, Vörös J, Arntzen JW. 2014. The distribution of the crested and marbled
1055 newt species (Amphibia: Salamandridae: *Triturus*) – an addition to the New Atlas of
1056 Amphibians and Reptiles of Europe. *Amphibia-Reptilia* 35:376-381.
- 1057 Zhang C, Sayyari E, Mirarab S editors. *Comparative Genomics*. 2017 2017//: Cham.
- 1058 Zhang J, Kobert K, Flouri T, Stamatakis A. 2014. PEAR: a fast and accurate Illumina Paired-
1059 End reAd mergeR. *Bioinformatics* 30:614-620.
- 1060

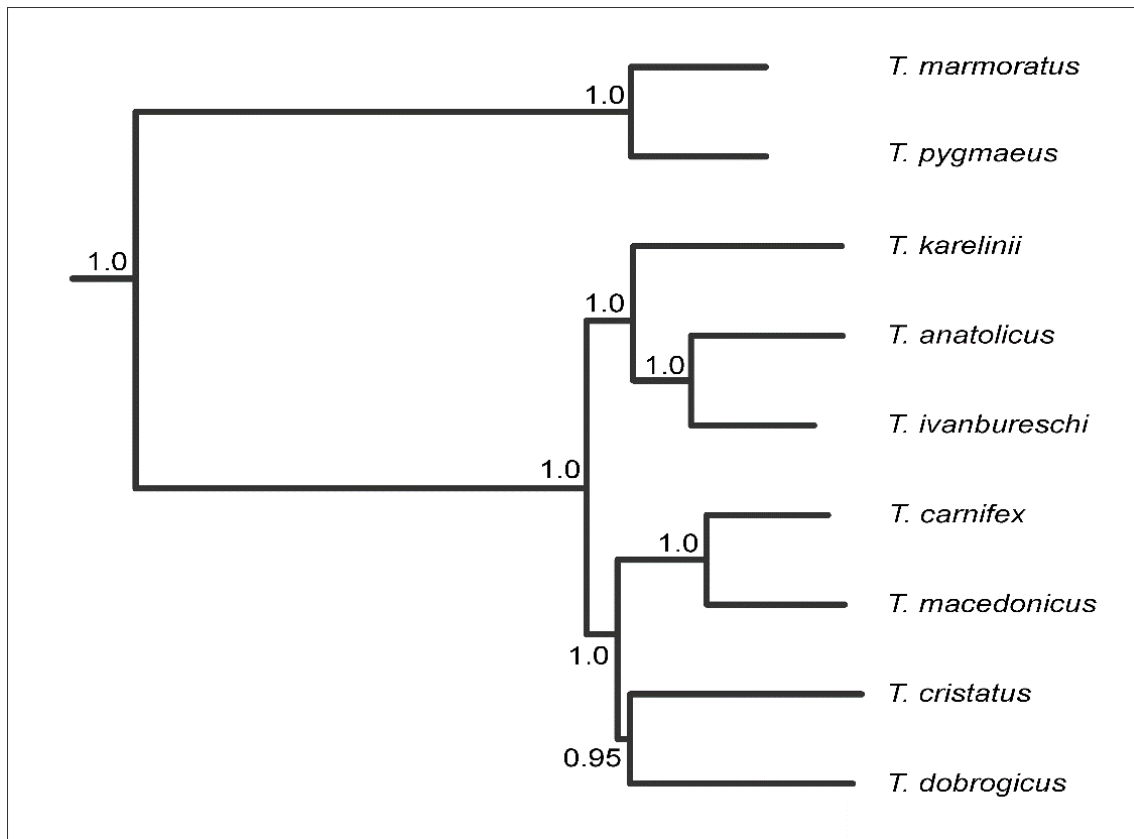
1062



1063

1064

1065 **Fig. S1. *Triturus* newt phylogenies based on data concatenation of transcriptome data**
 1066 **with RAxML.** In a) a *T. carnifex* individual is included that is suspected to be admixed with *T.*
 1067 *dobrogicus* and in b) this is replaced by two other *T. carnifex* individuals assumed to not be affected by
 1068 genetic admixture, one from the Balkans and one from Italy, away from the contact zone with *T.*
 1069 *dobrogicus*. Note the differences in sister species relationships (reflected by red arrows), with the
 1070 phylogeny in b) being in full agreement with the one based on target capture data (Fig. 3; Fig. 4).

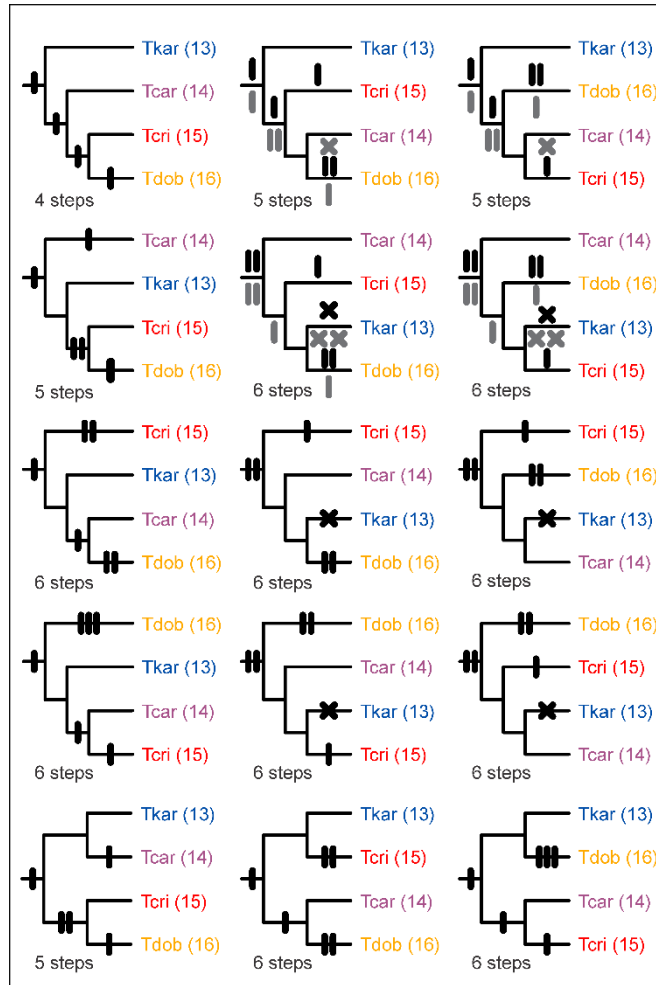


1071

1072

1073 **Fig. S2. Full mtDNA phylogeny for *Triturus*.** The genome-enabled *Triturus* phylogeny (Fig. 3; Fig.
 1074 4) deviates from the phylogeny based on full mtDNA (taken from (Wielstra and Arntzen 2011)) for the
 1075 species relationships in the *T. karelinii*-group of crested newts (with *T. anatolicus* being sister to *T.*
 1076 *karelinii* rather than *T. ivanbureschi*). Numbers at nodes indicate posterior probabilities. Note the
 1077 relatively low support for the sister relationship between *T. cristatus* and *T. dobrogicus*.

1078



1079

1080

1081 **Fig. S3. All 15 topologies possible for a fully bifurcating phylogeny of the four crested newt body**

1082 **builds.** Abbreviations: Tkar = *T. karelinii*-group; Tcar = *T. carnifex*-group; Tcri = *T. cristatus*; Tdob =

1083 *T. dobrogicus*. The number of trunk vertebrae (NTV) for each body build is provided in parentheses. A

1084 bar represents an NTV addition and a cross a deletion. NTV changes were inferred under the parsimony

1085 criterion, considering NTV = 12 as the ancestral character state for *Triturus* (see Supplementary Text

1086 S5). Results under ACCTTRAN and DELTRAN optimization were identical for 11 topologies; for the

1087 four ones that deviated, character state changes under DELTRAN optimization are in black and above

1088 and under ACCTTRAN optimization in grey and below branches. The top left topology corresponds to

1089 the *Triturus* species tree (Fig. 3; Fig. 4).

1090 SUPPLEMENTARY TABLES S1-S2

1091

1092 **Table S1. Sampling details.** Individuals are identified with a code that refers to complete specimens
 1093 (ID starting with ZMA) or tail tips (remaining samples). All material is stored at Naturalis Biodiversity
 1094 Center, Leiden, The Netherlands.

1095

Target capture

ID	Species	Locality	Latitud	Longitud
5017	<i>Triturus marmoratus</i>	France: Jublains	48.252	-0.473
5016	<i>Triturus pygmaeus</i>	Portugal: Serra de Monchique	37.335	-8.506
4729	<i>Triturus ivanbureschi</i>	Bulgaria: Ostar Kamak	41.878	25.853
1814	<i>Triturus ivanbureschi</i>	Turkey: Karakadılar	40.010	26.940
1788	<i>Triturus ivanbureschi</i>	Turkey: Bozdağ	38.367	28.103
1847	<i>Triturus anatolicus</i>	Turkey: Abanta Gölu	40.612	31.288
1889	<i>Triturus anatolicus</i>	Turkey: Gököy	40.083	33.347
1985	<i>Triturus anatolicus</i>	Turkey: Çakırlı	40.446	37.483
2105	<i>Triturus karelinii</i>	Ukraine: Nikita	44.538	34.243
6719	<i>Triturus karelinii</i>	Azerbaijan: Altiagac	40.854	48.935
RMNH RenA 46931-2390	<i>Triturus karelinii</i>	Iran: Qu'Am Shahr	36.436	52.803
ZMA9108-405	<i>Triturus carnifex</i>	Italy: Fuscaldo	39.417	16.033
ZMA9145-292	<i>Triturus carnifex</i>	Italy: Pisa	43.717	10.400
ZMA9132-312	<i>Triturus carnifex</i>	Slovenia: Kramplje	45.733	14.500
3247	<i>Triturus macedonicus</i>	Montenegro: Bjeloši	42.374	18.907
3275	<i>Triturus macedonicus</i>	Albania: Bejar	40.429	19.850
3775	<i>Triturus macedonicus</i>	Greece: Kerameia	39.562	22.081
4485	<i>Triturus cristatus</i>	Bulgaria: Montana	43.416	23.222
3686	<i>Triturus cristatus</i>	Romania: Budeni	45.768	26.839
ZMA9167-355	<i>Triturus cristatus</i>	Romania: Virfuri	46.283	22.467
ZMA9083-512	<i>Triturus dobrogicus</i>	Hungary: Alap	46.800	18.683
ZMA9172-720	<i>Triturus dobrogicus</i>	Croatia: Zupanja	45.083	18.700
2377	<i>Triturus dobrogicus</i>	Romania: Măcin	45.251	28.121

Transcriptomes

ID	Species	Locality	Latitud	Longitud
6720	<i>Triturus marmoratus</i>	Portugal: Valongo	41.168	-8.500
6721	<i>Triturus pygmaeus</i>	Portugal: Serra de Monchique	37.335	-8.506
6722	<i>Triturus karelinii</i>	Azerbaijan: Katex	41.646	46.543
6723	<i>Triturus anatolicus</i>	Turkey: Hacilar	41.495	32.088
6724	<i>Triturus ivanbureschi</i>	Turkey: Keşan	40.924	26.635
6725	<i>Triturus carnifex</i>	Croatia: Prkovic	45.569	16.094
6726	<i>Triturus carnifex</i>	Croatia: Radetići	45.146	13.842
6727	<i>Triturus carnifex</i>	Italy: Viterbo	42.703	13.325
6728	<i>Triturus macedonicus</i>	Montenegro: Ceklin	42.367	18.982
6729	<i>Triturus cristatus</i>	France: Belgeard	48.259	-0.574
6730	<i>Triturus dobrogicus</i>	Serbia: Sremski Karlovski	45.175	19.991
6731	<i>Ommatotriton nesterovi</i>	Turkey: Hürriyet	40.276	28.650

1096

Table S2. Inter- and intraspecific divergence in *Triturus* newts. Shown are pairwise F84 divergences calculated with Phylip. Intraspecific distances are in italics. IDs correspond to Supplementary Table S1.

		5017	5016	4729	1814	1788	1847	1889	1985	2105	6719	2390	405	292	312	3247	3275	3775	4485	3686	355	512	720	2377
<i>T. marmoratus</i>	5017	-																						
<i>T. pygmaeus</i>	5016	0.10	-																					
<i>T. ivanbureschi</i>	4729	0.72	0.70	-																				
	1814	0.72	0.70	0.02	-																			
	1788	0.72	0.71	0.03	0.02	-																		
<i>T. anaticus</i>	1847	0.68	0.66	0.08	0.08	0.09	-																	
	1889	0.72	0.70	0.11	0.11	0.12	0.02	-																
	1985	0.72	0.70	0.11	0.11	0.12	0.03	0.03	-															
<i>T. karelinii</i>	2105	0.71	0.69	0.14	0.14	0.14	0.09	0.11	0.11	-														
	6719	0.74	0.72	0.14	0.14	0.15	0.10	0.12	0.11	0.01	-													
	2390	0.73	0.71	0.14	0.14	0.15	0.10	0.12	0.11	0.01	0.02	-												
<i>T. carnifex</i>	405	0.74	0.72	0.25	0.25	0.26	0.23	0.26	0.25	0.27	0.27	0.27	-											
	292	0.75	0.73	0.25	0.25	0.26	0.23	0.26	0.25	0.26	0.27	0.27	0.03	-										
	312	0.71	0.69	0.23	0.23	0.24	0.21	0.24	0.23	0.25	0.25	0.25	0.08	0.07	-									
<i>T. macedonicus</i>	3247	0.74	0.72	0.24	0.24	0.25	0.22	0.25	0.25	0.25	0.26	0.26	0.22	0.21	0.19	-								
	3275	0.72	0.70	0.23	0.23	0.23	0.20	0.23	0.23	0.24	0.24	0.24	0.20	0.20	0.18	0.04	-							
	3775	0.75	0.73	0.24	0.24	0.25	0.22	0.25	0.25	0.26	0.26	0.26	0.22	0.22	0.20	0.06	0.04	-						
<i>T. cristatus</i>	4485	0.71	0.69	0.21	0.21	0.22	0.19	0.22	0.21	0.23	0.23	0.23	0.21	0.21	0.19	0.20	0.18	0.20	-					
	3686	0.72	0.70	0.22	0.22	0.22	0.20	0.22	0.22	0.23	0.24	0.24	0.21	0.21	0.19	0.20	0.18	0.20	0.05	-				
	355	0.69	0.67	0.21	0.21	0.21	0.19	0.21	0.21	0.22	0.23	0.23	0.21	0.21	0.19	0.19	0.17	0.19	0.04	0.05	-			
<i>T. dobrogicus</i>	512	0.71	0.69	0.23	0.23	0.24	0.21	0.24	0.23	0.25	0.25	0.25	0.21	0.20	0.18	0.21	0.19	0.21	0.17	0.17	0.17	-		
	720	0.69	0.67	0.21	0.21	0.22	0.20	0.22	0.22	0.23	0.23	0.23	0.19	0.18	0.17	0.19	0.17	0.19	0.16	0.16	0.16	0.03	-	
	2377	0.70	0.68	0.22	0.22	0.22	0.20	0.22	0.22	0.23	0.24	0.24	0.19	0.19	0.17	0.19	0.18	0.19	0.15	0.16	0.15	0.02	0.03	-

1097
1098
1099

1100
1101