Molecular Phylogenetics and Evolution 203 (2025) 108237

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Phylogenomics resolves the puzzling phylogeny of banded newts (genus *Ommatotriton*)

Konstantinos Kalaentzis [a,b,c,1], Stephanie Koster [a,b,1], Jan W. Arntzen [a,b], Sergé Bogaerts [d], James France [a,b], Michael Franzen [e], Christos Kazilas [a,b], Spartak N. Litvinchuk [f], Kurtuluş Olgun [g], Manon de Visser [a,b], Ben Wielstra [a,b,*]

[a] *Institute of Biology Leiden, Leiden University, P.O. Box 9505, 2300 RA Leiden, the Netherlands*
[b] *Naturalis Biodiversity Center, P.O. Box 9517, 2300 RA Leiden, the Netherlands*
[c] *Hydrobiological Station of Rhodes (HCMR), 85131 Rhodes, Greece*
[d] *Strijperstraat 55a, NL55995GB Leende, the Netherlands*
[e] *Zoologische Staatssammlung München (ZSM-SNSB), Münchhausenstraße 21, 81247 München, Germany*
[f] *Institute of Cytology, Russian Academy of Sciences, Tikhoretsky pr. 4, 194064 St. Petersburg, Russia*
[g] *Department of Biology, Faculty of Sciences, Adnan Menderes University, 09010 Aydın, Turkey*

ARTICLE INFO

Keywords:
Gene-tree discordance
Phylogenomics
Sequence capture
Species-tree inference

ABSTRACT

Resolving the order of speciation events that occurred in rapid succession is inherently hard and typically requires a phylogenomic approach. A case in point concerns the previously unresolved phylogeny of the three species of banded newt (genus *Ommatotriton*). We obtain c. 7k nuclear DNA markers using target enrichment by sequence capture and analyze the dataset using maximum likelihood inference of concatenated data with RAxML, summary multi-species coalescent analysis with ASTRAL and Bayesian species tree inference using a diffusion model with SNAPPER, and use TreeMix and PhyloNet to test for interspecific gene flow. All analyses recover three distinct species with no evidence of interspecific gene flow. All analyses retrieved the topology (*O. nesterovi*, (*O. ophryticus*, *O. vittatus*)), with high support. SNAPPER did show the tendency to get stuck in a local optimum, resulting in a different but still highly supported topology. Furthermore, we notice that fewer SNAPPER runs get stuck in a local optimum when we include an outgroup. Therefore, we recommend the exploration of multiple independent runs and the use of an outgroup with this approach. The banded newt radiation illustrates the use of genome-wide data to tackle formerly unresolved phylogenies.

1. Introduction

The phylogenetic relationships of lineages that speciated a relatively long time ago, over a relatively short time period, are inherently problematic to resolve (Philippe et al., 2011; Whitfield and Lockhart, 2007). While non-informative substitutions tend to accumulate along terminal branches, only a few substitutions along internal branches inform the actual branching order (leading to conflicted genealogies; Marques et al., 2019). Furthermore, both ancient and recent gene flow may give rise to incongruent gene genealogies (Degnan and Rosenberg, 2009; Eckert and Carstens, 2008; Kutschera et al., 2014). As a consequence, successive speciation events that occurred in a brief time window cannot typically be resolved using one or a few genetic markers and rather require a phylogenomic approach (e.g. Irisarri and Meyer, 2016; McCormack et al., 2012).

Salamanders possess some of the largest vertebrate genomes, with an average size of c. 36 Gb (Calboli et al., 2011; Gregory, 2021). So far, two chromosome-level assemblies of salamander genomes have been produced: the Mexican axolotl (*Ambystoma mexicanum*) (32 Gb; Nowoshilow et al., 2018; Smith et al., 2019) and the Iberian ribbed newt (*Pleurodeles waltl*) (Brown et al., 2022). The gigantic genome of salamanders is a major impediment to phylogenomic studies; besides the time and money required, compiling such a large genome also poses a considerable bioinformatic challenge, in terms of data storage and computational power (Andermann et al., 2020). To avoid these challenges, it is advisable to utilize genome reduction methods that yield a

* Corresponding author at: Institute of Biology Leiden, Leiden University, P.O. Box 9517, 2300 RA Leiden, the Netherlands.
E-mail address: b.m.wielstra@biology.leidenuniv.nl (B. Wielstra).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ympev.2024.108237
Received 6 April 2024; Received in revised form 10 September 2024; Accepted 10 November 2024
Available online 17 November 2024

summary representation of the genome.

Several genome sub-sampling methods have been developed that can lower the cost and complexity of phylogenetic studies (Andermann et al., 2020; Davey et al., 2011). These can be categorized into two main groups, the first of which being non-targeted methods, i.e. using random sequences from across the genome (RAD-seq and similar methods; e.g. Baird et al., 2008; Miller et al., 2007). Even though non-targeted methods have successfully been used in phylogenetic studies of amphibians (e.g. Ambu et al., 2023; Dufresnes et al., 2019; Gamble et al., 2015; McCartney-Melstad et al., 2018; Rancilhac et al., 2019), issues may arise, such as the unknown orthology of markers, missing data due to evolutionary changes in restriction sites and unknown genetic-linkage relationships among loci (Andermann et al., 2020; Rubin et al., 2012).

Conversely, targeted methods focus on pre-selected genomic regions such as ultraconserved elements or anchor-regions (anchor hybrid enrichment; Lemmon et al., 2012) or molecular inversion probes (Faircloth et al., 2012; Niedzicka et al., 2016). Target enrichment by sequence capture is a particularly promising genome sub-sampling method for ecological and evolutionary genomic studies (Andermann et al., 2020; Jones and Good, 2016). This procedure can generate standardized datasets of species with moderate sequence divergence. It can be applied to relatively degraded samples (e.g. museum specimens; Andermann et al., 2020; Jones and Good, 2016) because the procedure involves DNA fragmentation (e.g. by sonication or enzymatic shearing). Target enrichment by sequence capture employs custom RNA bait sequences, complementary to specific genomic regions, which are hybridized to prepared genomic libraries (Gnirke et al., 2009). After washing away the off-target genomic regions, the libraries containing the DNA of interest are enriched by PCR and sequenced. By focusing sequencing effort on several thousands of loci instead of the whole genome, sequencing costs are considerably reduced.

The increase in phylogenomic datasets has gone hand in hand with the development of computational methods that can handle such large datasets. The implementation of fast algorithms facilitated the analysis of large datasets of concatenated alignments under maximum likelihood (Stamatakis, 2014). Methods that take incomplete lineage sorting into account and are statistically consistent under the multi-species coalescent model have also been developed but are computationally demanding (Rannala and Yang, 2017). Summary multi-species coalescent analyses aim to return the true species tree based on a summary of individual gene trees (Zhang et al., 2017), while other methods employ biallelic sites to compute the species tree without inferring individual gene trees (Bryant et al., 2012; Stoltz et al., 2021). Furthermore, studies report on quantifying the degree of gene flow between evolutionary lineages, which could potentially disrupt phylogenetic inference (Pickrell and Pritchard, 2012; Than et al., 2008).

One case that could significantly benefit from target enrichment by sequence capture is the currently unresolved phylogeny of banded newts (genus *Ommatotriton*). Banded newts belong to the family Salamandridae and are endemic to the Anatolian peninsula, the Caucasus region and the Levant (Fig. 1). Three genetically and morphologically distinct species have recently been recognized (Bülbül and Kutrup, 2013; Kalaentzis et al., 2023; Üzüm et al., 2019; van Riemsdijk et al., 2022, 2017), namely the Anatolian (*Ommatotriton nesterovi*), the Caucasian (*O. ophryticus*) and the southern (*O. vittatus*) banded newt (Fig. 1). These are estimated to have radiated c. 25 million years ago over a time span of a few million years (with wide and overlapping confidence intervals for the relevant nodes; van Riemsdijk et al., 2017). A previous study, despite including 177 nuclear DNA markers, failed to resolve the phylogenetic relationships among the three banded newt species (van Riemsdijk et al., 2022). This banded newt trichotomy might represent a 'hard' trichotomy, i.e. a truly simultaneous three-way split. However, it could also reflect a 'soft' trichotomy, that simply has not been solved yet because insufficient data was consulted so far.

Fig. 1. Geographical distribution of banded newts (genus *Ommatotriton*). Colors correspond to species. The sampling localities are indicated with a dot for which the number corresponds to Table 1.

To try and break the banded newt trichotomy, we apply a phylogenomic approach, based on c. 7k nuclear DNA markers obtained with target enrichment by sequence capture. We analyze the data under distinct methods of phylogenetic inference, namely maximum likelihood inference of concatenated data with RAxML, summary multispecies coalescent analysis with ASTRAL and Bayesian species tree inference using a diffusion model with SNAPPER. Furthermore, we test if interspecific gene flow could act as a potential confounding factor in our phylogenetic inference. Our study provides a much-improved picture of the evolutionary history of banded newts.

2. Materials & methods

2.1. Array design and sampling

A probe set targeting 7,102 genomic regions in newts of the genus *Triturus*, previously designed by Wielstra et al. (2019) as a MyBaits-II kit (Arbor Biosciences, MI, USA; Ref# 170210-32), is here applied to *Ommatotriton* newts. This probe is anticipated to work in related newt genera on account of the method's ability to capture relatively diverged genomic regions (Bi et al., 2012). For each banded newt species we included five individuals from five different localities (Table 1; Fig. 1), available from van Riemsdijk et al. (2017). To reduce the potential negative effect of recent admixture on phylogeny reconstruction, we only included samples taken well away from the hybrid zone between *O. nesterovi* and *O. ophryticus* (Kalaentzis et al., 2023). Genomic data of three marbled newts (*Triturus marmoratus*), three Italian crested newts (*T. carnifex*) and three Macedonian crested newts (*Triturus macedonicus*), available from Wielstra et al. (2019) and Kazilas et al. (2023), were used as outgroup.

2.2. Genomic library preparation and sequence capture

To obtain genomic data we use "NewtCap": a target enrichment by sequence capture workflow (de Visser et al., 2024). We prepared libraries using the NEBNext Ultra™ II FS DNA Library Prep Kit for Illumina (New England Bioscience, MA, USA), following the manufacturer's protocol, with all volumes divided by 4 and a fragmentation time of 6:30 min. Following enzymatic fragmentation, NEB adapters were ligated onto the sheared DNA fragments, and cleaved with USER enzyme (New England Bioscience, MA, USE). NucleoMag™ magnetic separation beads (Macherey-Nagel, Düren, Germany) were used for size selection, targeting an insert size of 300bp. For indexing, unique combinations of i5 and i7 indices, customed ordered from IDT (Integrated DNA Technologies, Leuven, Belgium), were attached to each sample via PCR. NucleoMag™ beads were again employed for the final library clean-up. Finally, the fragment-size distribution of libraries was assessed with the Agilent 2200 TapeStation system (Agilent, Santa Clara, CA, USA).

Libraries were equimolarly pooled for sequence capture, resulting in a final pool with a total mass of DNA of 3,750 ng (250 ng per sample). Libraries were enriched with the MyBaits-V4 kit (Arbor Biosciences, MI, USA), following the manufacturer's protocol, with the following deviations: blocks C and O were replaced with 30,000 ng of *Triturus*-derived C_0t-1 repetitive sequence blockers (in 5 μl), the libraries were incubated with the blocking solution for 30 min before addition of the probes, and hybridization time was 30 h. Tissue to produce C_0t-1 DNA was available from a removal action of an invasive population of *T. carnifex* (Meilink et al., 2015). Following washing to remove unhybridized DNA, a final PCR step (14 cycles) was performed to amplify the enriched libraries, and NucleoMag™ beads were employed for cleanup. The enriched pool was quantified with the Agilent 2200 TapeStation system, and 15 Gbp of 150 bp paired-end sequencing was requested to be performed on the NovaSeq 6000 platform (Illumina Inc., San Diego, CA, USA) by BaseClear B.V. (Leiden, the Netherlands).

Table 1
Sampling details for the three species of banded newt (genus *Ommatotriton*). Locality ID corresponds to Fig. 1.

Number	Species	Sample ID	Locality	Latitude	Longitude
1	*O. nesterovi*	vit007	Turkey: Subaşi	40.257	28.538
2	*O. nesterovi*	vit031	Turkey: Abant Gölu	40.612	31.288
3	*O. nesterovi*	vit172	Turkey: Cebeli	41.121	35.330
4	*O. nesterovi*	vit206	Turkey: Çibanköy	41.201	34.037
5	*O. nesterovi*	vit581	Turkey: Levent	40.965	30.433
6	*O. ophryticus*	vit299	Turkey: Gökçekent	40.287	38.132
7	*O. ophryticus*	vit377	Turkey: Denizgören	40.968	40.380
8	*O. ophryticus*	vit491	Turkey: Aşağıkoyunlu	41.303	42.489
9	*O. ophryticus*	vit690	Georgia: Mtskheta	41.809	44.498
10	*O. ophryticus*	vit676	Georgia: Lidzava	43.183	40.350
11	*O. vittatus*	vit612	Turkey: Kulak	36.778	34.861
12	*O. vittatus*	vit610	Turkey: Yolbaşi	36.845	36.653
13	*O. vittatus*	vit654	Turkey: Ceylanpınar	36.869	39.991
14	*O. vittatus*	vit619	Syria: Homs	34.716	36.598
15	*O. vittatus*	I1	Israel: Tel Aviv	32.083	34.767

2.3. Analysis of sequence-capture data

Sequence-capture data were analyzed as follows: trimming (i.e. removal of adapter contamination) using Skewer v0.2.2 (Jiang et al., 2014), mapping reads to a *Triturus* reference assembly (Wielstra et al., 2019) using BWA-MEM 0.7.17 (Li, 2013), adding read groups with Picard tools v2.25.1 (https://broadinstitute.github.io/picard/), and calling variants and combining GVCFs using HaplotypeCaller and GenotypeGVCFs (integrated in GATK v4.1.3.0; McKenna et al., 2010). Three datasets were assembled and used for downstream analyses; a 15-sample dataset containing only the banded newt individuals, an 18-sample dataset containing the three *T. marmoratus* as outgroup, and a 24-sample dataset, which included all nine *Triturus* newts as outgroup.

Datasets were tested for any paralogous sites through the Hardy Weinberg Equilibrium (HWE) filter of heterozygote excess. Heterozygote excess p-values were calculated for each SNP using BCFtools 1.15.1 (Danecek et al., 2021; Li, 2011), and any site with a p-value less than 0.05 was excluded from downstream analyses. Next, a quality filtering step was performed using VCFtools v0.1.16 (Danecek et al., 2021), where genotypes were set to a minimum quality score of 20, indels were removed and only sites with less than 50 % missing data were kept. A separate VCF file including only one SNP per target was generated using a custom Perl script for species-tree estimation with SNAPPER v1.1.3 (see below) with the added requirement of no missing data being allowed.

2.4. Principal component analysis

Principal Component Analysis was performed in R Studio (R Studio Team, 2021) using the R packages gdsfmt and SNPRelate (Zheng et al., 2012) in order to clarify the inter- and intra-specific genetic structure of banded newts. For PCA, only the 15-sample dataset (without the *Triturus* outgroup) was analyzed, for which 197,495 biallelic SNPs were ultimately used, after pruning another 144,736 SNPs that were monomorphic following the default criteria of the SNPRelate package.

2.5. Maximum likelihood inference of concatenated data with RAxML

The 24-sample dataset (i.e. including the *Triturus* outgroup) was converted into PHYLIP format, which was used as an input for RAxML, using the vcf2phylip.py Python script (Ortiz, 2019). Additionally, sites that were regarded by RAxML as monomorphic were removed using the

Python script ascbias.py (https://github.com/btmartin721/raxml_ascbias). RAxML v.8.2.12 (Stamatakis, 2014) was then used to infer a concatenated maximum likelihood phylogeny based on the alignment of thousands of SNPs from an initial amount of 6,884 target-sequences (182,480 SNPs). RAxML was run with 100 bootstrap replicates and the ASC_GTRGAMMA model with Lewis ascertainment correction for SNP analysis (Lewis, 2001).

To date the phylogenetic tree obtained from RAxML we used treePL v2.6.3 (Smith and O'Meara, 2012). Because we lack internal calibration points or markers with a known substitution rate, we make use of two calibration points in the outgroup. The split between *T. marmoratus* and the other two *Triturus* species, dated to 24 Ma (Steinfartz et al., 2007), and between *T. macedonicus* and *T. carnifex*, dated to 5.33 Ma (Wielstra and Arntzen, 2011), were used as fixed calibration points. A preliminary analysis was first performed on the RAxML tree. Based on the output of this analysis, "opt" and "optad" were both set to 3. The "moredetail" and "moredetailad" options were added, and the value for "optcvad" was set to 5. The optimal smoothing value as indicated by cross-validation was 0.00000001 (Maurin, 2020). We obtained 100 bootstrap replicates with RAxML in which the topology was fixed, but not the branch lengths, by explicitly specifying the multi-sample *Triturus* outgroup, which ensured the root was positioned correctly and effectively fixed the topology of the tree, but not the branch lengths. We assessed node age estimates by repeating the treePL analysis (with optimized settings) on the 100 bootstrap trees and summarized the trees with TreeAnnotator v2.4.7 (Bouckaert et al., 2019) to obtain 95% confidence intervals.

2.6. Summary multi-species coalescent analysis with ASTRAL

To generate a collection of gene trees SNPSift v.4.3 (Cingolani et al., 2012) was used to split the 24-sample (i.e. including the *Triturus* outgroup) VCF file into separate SNP alignments for each target. Then, each alignment of SNPs was converted into PHYLIP format with PGD Spider v.2.1.1.5 (Lischer and Excoffier, 2012) and, after monomorphic sites were again excluded as above (resulting in 5,930 markers), RAxML v.8.2.12 (Stamatakis, 2014) inferred a maximum likelihood tree for each target, using 100 bootstrap replicates and the ASC_GTRGAMMA model with Lewis ascertainment correction for SNP analysis (Lewis, 2001). Finally, ASTRAL v.5.7.7 (Zhang et al., 2017) was used to estimate the species tree based on the 'forest' of 5,930 gene trees. We constrained the six species (three *Ommatotriton* and three *Triturus*) to be monophyletic, as supported by previous knowledge as well as our RAxML and PCA analyses (see results).

2.7. Bayesian species tree inference using a diffusion model with SNAPPER

The 15-sample (without the *Triturus* outgroup) VCF file with one SNP per target for a total of 6,324SNPs was converted into binary nexus format using the vcf2phylip.py Python script (Ortiz, 2019). The resulting file with 5,922 biallelic SNPs was converted to XML format using the SNAPPER template implemented in BEAUti v2.7.5 (Bouckaert et al., 2019). Default settings were used, and chain length was set to 20,000,000. SNAPPER v1.1.3 (Stoltz et al., 2021) within BEAST v.2.7.0 (Bouckaert et al., 2019) was used to estimate the banded newt species tree. Inspired by Stervander et al. (2016), we repeated the analysis including three *T. marmoratus* individuals as an outgroup, i.e. for the 18-sample VCF file with one SNP per target for a total of 6,334 SNPs. The resulting binary nexus file contained 5,859 biallelic SNPs. We ran 12 replicates for SNAPPER both with and without outgroup and explored the output in Tracer v1.7.2 (Rambaut et al., 2018) and FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

2.8. Gene flow analysis with TreeMix and PhyloNet

To determine whether genetic introgression was signaled by the data, TreeMix and PhyloNet were used on the 15-sample *Ommatotriton*-only dataset. For TreeMix, VCFtools (Danecek et al., 2021) was first used to remove sites with missing data from the combined banded newt vcf file, and 166,402 of the 441,706 sites were retained. Subsequently, linkage pruning was performed using PLINK v1.07 (Purcell et al., 2007). The resulting output file with 25,488 sites that passed the requirements was then used with VCFtools (Danecek et al., 2021) to keep only the sites that were not in linkage disequilibrium. Population data were added to the file using the populations module of Stacks v2.64 (Catchen et al., 2013), to assign the samples to their respective species. TreeMix v1.13 (Pickrell and Pritchard, 2012) was run on the resulting file for zero to five migration edges. R Studio (R Studio Team, 2021) was then used to visualize the output files using the script provided by Pickrell and Pritchard (2012).

To create the input file for PhyloNet, the same steps as for ASTRAL were followed for the first part; the VCF file (without *Triturus*) had heterozygote excess and missing data removed, was split into separate SNP alignments, which were converted into PHYLIP format, and RAxML was run on each alignment to infer maximum likelihood phylogenies. The RAxML bestTree gene trees, 5,008 in total, were then concatenated and numbered. A NEXUS header and instructions for PhyloNet were added. PhyloNet v3.8.2 (Than et al., 2008) was first used to infer a species tree under the "Minimize Deep Coalescence" criterion, to prepare a null hypothesis without reticulations (Yu et al., 2011). The *Ommatotriton* section of the earlier RAxML phylogeny was added to the file as gene tree 5,009 to constrain the topology of the networks. Next, PhyloNet was run for one to five reticulations, each with ten replicates, to infer networks under Maximum Parsimony. Each output file shows a number of extra lineages, and the run with the lowest number was selected to present the optimal number of reticulations. The resulting networks were visualized with Dendroscope v3.8.10 (Huson and Scornavacca, 2012).

3. Results

A total of 30.3 Gbp of raw sequence data was obtained. The 15 banded newt samples had a mean of 7,160,518 raw read pairs (s.d. = 2,325,050). An average of 66.6% of raw reads (s.d. = 1.3 %) could be mapped against the *Triturus* reference. PCR duplicates accounted for an average of 93.4% of mapped reads (s.d. = 1.1 %). After deduplication, samples had an average median peak 100 bp region coverage of 26.5 (s.d. = 11), with a minimum of 7 and a maximum of 48.

The three *Ommatotriton* species are differentiated by the principal component analysis when plotting principal component 1 (explaining 22.1 % of the variance) against principal component 2 (8.3 % of the variance) (Fig. 2). Along principal component 3 (explaining 6.6 % of the variance), *O. vittatus cilicensis* (locality 11), the only recognized subspecies of banded newt, is separated from both the nominate subspecies and the other two species (Fig. 2).

The maximum likelihood inference of concatenated data with RAxML recovered each of the three banded newts as monophyletic, supporting their treatment as such in downstream analyses. A sister-species relationship for *O. vittatus* and *O. ophryticus*, to the exclusion of *O. nesterovi*, is recovered with a bootstrap support of 100 percent (Fig. 3; Supplementary Fig. 1). The TreePL analysis dates the *Ommatotriton* crown to c. 15 Ma and the split between *O. vittatus* and *O. ophryticus* to c. 13 Ma (Fig. 3).

The same sister-species relationship as in RAxML is found under gene-tree summary with ASTRAL (Fig. 4a) with a posterior probability of 1.0. The normalized quartet score, the proportion of the quartet trees generated by all the gene trees that are present in the species tree and therefore a measure of gene-tree discordance (Sayyari and Mirarab, 2016), is 0.75, which signifies a relatively low degree of gene-tree discordance.

Exploration of the output of the Bayesian species-tree inference using a diffusion model with SNAPPER shows that part of the runs got stuck in

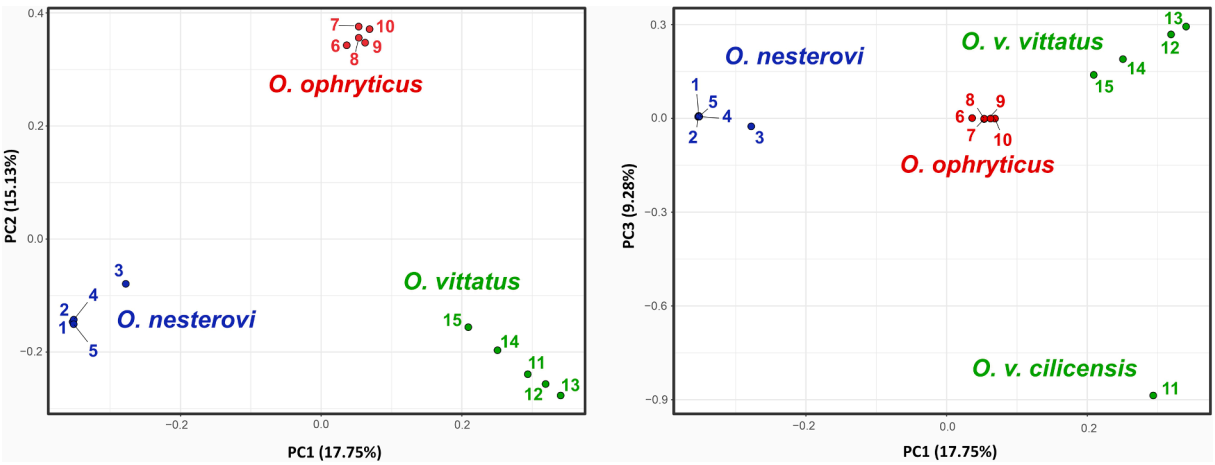

Fig. 2. Principal component analysis of 15 individual banded newts (genus *Ommatotriton*). The analysis is based on 197,495 biallelic SNPs from across 6,884 target-sequences. Percentage of variance explained by each principal component (PC) is given in parentheses on each axis. Labels correspond to localities in Table 1.

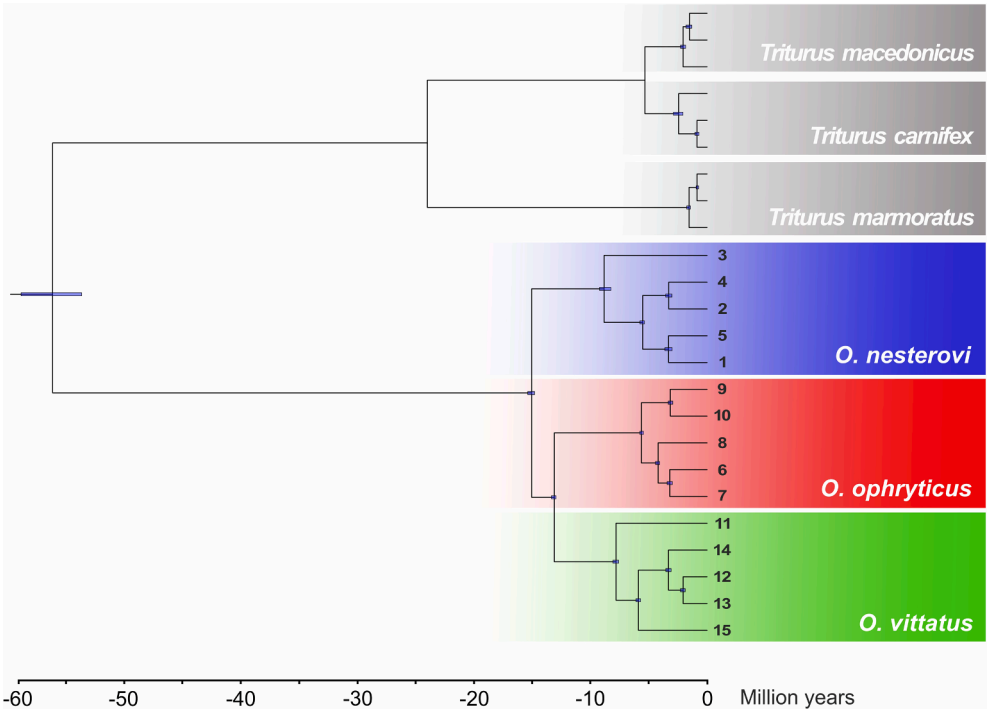

Fig. 3. Phylogeny of the three species of banded newt (genus *Ommatotriton*), based on maximum likelihood inference of concatenated data with RAxML and time calibration with TreePL. The analysis is based on 182,480 SNPs from across 6,884 target-sequences. The output of RAxML, used as input for TreePL, can be found as Supplementary Fig. 1. S1. The 95% confidence intervals are based on the 100 bootstrap trees generated with RAXML with the topology, but not branch lengths, fixed. All nodes have a bootstrap support of 100 percent (not shown). Labels correspond to Table 1.

a local optimum. The runs that converged on the highest likelihood value returned the same topology as RAxML and ASTRAL (Fig. 4b). Runs that got stuck in a local optimum with a lower likelihood rather suggested a basal split of *O. vittatus* versus the other two species with full support. Additionally, we find that less of the replicates that included an outgroup (5/12) got stuck in a local optimum than replicates that did not include an outgroup (9/12). ESS values for the combined runs that converged on the highest likelihood were ≫100.

Treemix and PhyloNet do not indicate introgression between *Ommatotriton* species.

4. Discussion

By utilizing a large genomic dataset, we finally resolve the phylogeny of banded newts (genus *Ommatotriton*). All three phylogenetic analyses find strong support for the topology: (*O. nesterovi*, (*O. ophryticus*, *O. vittatus*)). Although our new temporal estimates for the banded newt radiation, based on genome-wide data, are not as old as those based on mtDNA (van Riemsdijk et al., 2017), they are still of Miocene age, with the key nodes flanking the relatively short internal branch dated to c. 15 and c. 13 Ma. We do not find evidence for ancient introgression between banded newt species – a potential worry, given that rampant signatures of ancient introgression characterize the phylogeny of the Salamandridae family that includes the newts (Rancilhac et al., 2021) – which increases our confidence in the new banded newt phylogeny. Although the banded newt phylogeny was previously considered a trichotomy (van Riemsdijk et al., 2022), with genome-wide data we show that it was, in hindsight, a 'soft' trichotomy (i.e. reflecting a lack of

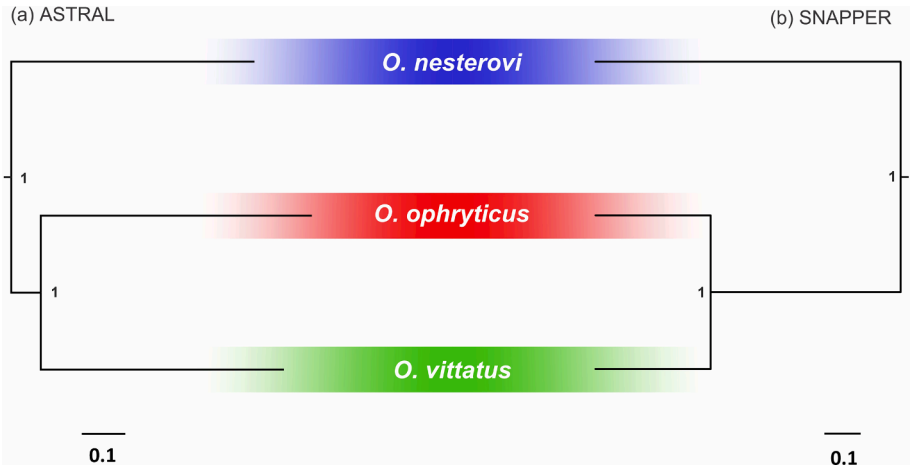

Fig. 4. Phylogeny of the three species of banded newt (genus *Ommatotriton*) based on summary multi-species coalescent analysis with ASTRAL (a) and full-likelihood multi-species coalescent analysis with SNAPPER (b). The ASTRAL tree is inferred from 5,930 gene trees. The outgroup is not shown. Branch lengths represent coalescent units and numbers at nodes local quartet support posterior probabilities. The SNAPPER tree is based on the analysis without outgroup, inferred from 6,324 unlinked SNPs. Branch lengths represent the number of expected substitutions per site, and numbers at nodes posterior probabilities.

information) and not a 'hard' trichotomy (i.e. an actual three-way split).

Maximum likelihood inference of concatenated data is commonly used because it is computationally tractable (Stamatakis, 2014). Data concatenation has been criticized as statistically inappropriate, because it oversimplifies the complexity of the evolutionary process and ignores biological phenomena such as gene flow, incomplete lineage sorting and gene duplication or gene loss (Kubatko and Degnan, 2007; Roch and Steel, 2015; Wang et al., 2014). While several studies claim that concatenation methods are less effective (compared with approaches that consider the multi-species coalescent model) at uncovering the 'true' species tree (e.g. Stervander et al., 2015; Weisrock et al., 2012), others have found that they converge on the outcome of 'more complicated' phylogenetic analyses (e.g. Stervander et al., 2016; Wielstra et al., 2019).

Summary-coalescent methods use the original sequence data and reconstruct gene-tree estimates for individual markers, the summary over which is then interpreted as the species tree (a two-step approach). These methods are computationally fast and easy to implement (Zhang et al., 2017). Summary-coalescent methods have been criticized for being susceptible to the anomaly zone (i.e. where the majority of gene trees do not reflect the topology of the true species tree) when the species tree has short internal branches (e.g. Degnan and Rosenberg, 2006; Degnan and Salter, 2005; Xu and Yang, 2016). However, summary coalescence has been successfully used to infer the phylogeny of complicated cases (e.g. Eaton and Ree, 2013; McCormack et al., 2013).

Methods that incorporate the full-likelihood multi-species coalescent model (i.e. assuming that gene trees evolve under unique coalescent processes because of recombination between the genes) are considered most reliable for species-tree inference (Bryant et al., 2012; Xu and Yang, 2016). They are considered to be particularly suited for rapid radiations because they are less susceptible to the anomaly zone (Rannala and Yang, 2017; Xu and Yang, 2016). A large downside is that these analyses are computationally burdensome, a burden that increases drastically with sample size. A solution is offered by the use of biallelic sites in combination with a diffusion-based model (Stoltz et al., 2021). The computational burden for this analysis does not increase exponentially with sample size, so there is effectively no limit on the number of samples (Stoltz et al., 2021).

While all three phylogenetic analyses converge on the same topology, we notice that the Bayesian species tree inference using a diffusion model with SNAPPER regularly gets trapped in a local optimum for our banded newt dataset. Adding an outgroup seems to reduce this problem, but comes with the downside that the number of informative SNPs within the ingroup is reduced. We recommend that future studies using SNAPPER explore if multiple replicates converge on the same answer, as well as whether the inclusion of an outgroup influences the outcome of the analysis. Banded newts are an excellent example of speciation that happened in rapid succession, resulting in a phylogeny that has not been straightforward to resolve.

CRediT authorship contribution statement

Konstantinos Kalaentzis: Writing – original draft, Formal analysis, Conceptualization. **Stephanie Koster:** Writing – original draft, Formal analysis, Conceptualization. **Jan W. Arntzen:** Writing – review & editing, Resources. **Sergé Bogaerts:** Writing – review & editing, Resources. **James France:** Writing – review & editing, Supervision. **Michael Franzen:** Writing – review & editing, Resources. **Christos Kazilas:** Writing – review & editing, Formal analysis. **Spartak N. Litvinchuk:** Writing – review & editing, Resources. **Kurtuluş Olgun:** Writing – review & editing, Resources. **Manon de Visser:** Writing – review & editing, Supervision. **Ben Wielstra:** Writing – original draft, Supervision, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

KK was the recipient of an Onassis Foundation scholarship. Michael Fahrbach provided pictures used in the graphical abstract. This work was performed using the compute resources from the Academic Leiden Interdisciplinary Cluster Environment (ALICE) provided by Leiden University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2024.108237.

Data availability

The Illumina sequencing reads generated for this study have been

submitted to the NCBI Sequence Read Archive (SRA) and can be retrieved through BioProject PRJNA1096730. All scripts utilized can be found in the following GitHub repository: https://github.com/Wielstra-Lab/banded_newts.

References

Ambu, J., Martínez-Solano, Í., Suchan, T., Hernandez, A., Wielstra, B., Crochet, P.A., Dufresnes, C., 2023. Genomic phylogeography illuminates deep cyto-nuclear discordances in midwife toads (*Alytes*). Mol. Phylogenet. Evol. 183. https://doi.org/10.1016/j.ympev.2023.107783.

Andermann, T., Torres Jiménez, M.F., Matos-Maraví, P., Batista, R., Blanco-Pastor, J.L., Gustafsson, A.L.S., Kistler, L., Liberal, I.M., Oxelman, B., Bacon, C.D., Antonelli, A., 2020. A guide to carrying out a phylogenomic target sequence capture project. Front. Genet. 10, 1–20. https://doi.org/10.3389/fgene.2019.01407.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., Johnson, E.A., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS One 3, 1–7. https://doi.org/10.1371/journal.pone.0003376.

Bi, K., Vanderpool, D., Singhal, S., Linderoth, T., Moritz, C., Good, J.M., 2012. Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. BMC Genomics 13. https://doi.org/10.1186/1471-2164-13-403.

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., Mendes, F.K., Müller, N.F., Ogilvie, H.A., Du Plessis, L., Popinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., Suchard, M.A., Wu, C.H., Xie, D., Zhang, C., Stadler, T., Drummond, A.J., 2019. BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 15. https://doi.org/10.1371/journal.pcbi.1006650.

Brown, T., Elewa, A., Iarovenko, S., Subramanian, E., Araus, A.J., Petzold, A., Susuki, M., Suzuki, K.-I.-T., Hayashi, T., Toyoda, A., Oliveira, C., Osipova, E., Leigh, N.D., Simon, A., Yun, M.H., 2022. Sequencing and chromosome-scale assembly of the giant *Pleurodeles waltl* genome. BioRxiv. https://doi.org/10.1101/2022.10.19.512763.

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N.A., RoyChoudhury, A., 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. Mol. Biol. Evol. 29, 1917–1932. https://doi.org/10.1093/molbev/mss086.

Bülbül, U., Kutrup, B., 2013. Morphological and genetic variations of *Ommatotriton* in Turkey. Anim. Biol. 63, 297–312. https://doi.org/10.1163/15707563-00002413.

Calboli, F.C.F., Fisher, M.C., Garner, T.W.J., Jehle, R., 2011. The need for jumpstarting amphibian genome projects. Trends Ecol. Evol. 26, 378–379. https://doi.org/10.1016/j.tree.2011.04.006.

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., Cresko, W.A., 2013. Stacks: an analysis tool set for population genomics. Mol. Ecol. 22, 3124–3140. https://doi.org/10.1111/mec.12354.

Cingolani, P., Patel, V.M., Coon, M., Nguyen, T., Land, S.J., Ruden, D.M., Lu, X., 2012. Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program, SnpSift. Front. Genet. 3, 1–9. https://doi.org/10.3389/fgene.2012.00035.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H., 2021. Twelve years of SAMtools and BCFtools. GigaScience 10, 1–4. https://doi.org/10.1093/gigascience/giab008.

Davey, J.W., Hohenlohe, P.A., Etter, P.D., Boone, J.Q., Catchen, J.M., Blaxter, M.L., 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat. Rev. Genet. 12, 499–510. https://doi.org/10.1038/nrg3012.

Degnan, J.H., Rosenberg, N.A., 2006. Discordance of species trees with their most likely gene trees. PLoS Genet 2, 762–768. https://doi.org/10.1371/journal.pgen.0020068.

Degnan, J.H., Rosenberg, N.A., 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol. https://doi.org/10.1016/j.tree.2009.01.009.

Degnan, J.H., Salter, L.A., 2005. Gene tree distributions under the coalescent process. Evol 59, 24. https://doi.org/10.1554/04-385.

de Visser, M.C., France, J., McCartney-Melstad, E., Bucciarelli, G.M., Theodoropoulos, A., Shaffer, H.B., Wielstra, B., 2024. NewtCap: an efficient target capture approach to boost genomic studies in Salamandridae (True Salamanders and Newts). bioRxiv. https://doi.org/10.1101/2024.1110.1125.620290.

Dufresnes, C., Mazepa, G., Jablonski, D., Oliveira, R.C., Wenseleers, T., Shabanov, D.A., Auer, M., Ernst, R., Koch, C., Ramírez-Chaves, H.E., Mulder, K.P., Simonov, E., Tiutenko, A., Kryvokhyzha, D., Wennekes, P.L., Zinenko, O.I., Korshunov, O.V., Al-Johany, A.M., Peregontsev, E.A., Masroor, R., Betto-Colliard, C., Denoël, M., Borkin, L.J., Skorinov, D.V., Pasynkova, R.A., Mazanaeva, L.F., Rosanov, J.M., Dubey, S., Litvinchuk, S.N., 2019. Fifteen shades of green: the evolution of *Bufotes* toads revisited. Mol. Phylogenet Evol. 141, 106615. https://doi.org/10.1016/j.ympev.2019.106615.

Eaton, D.A.R., Ree, R.H., 2013. Inferring phylogeny and introgression using RADseq data: an example from flowering plants (*Pedicularis*: Orobanchaceae). Syst. Biol. 62, 689–706. https://doi.org/10.1093/sysbio/syt032.

Eckert, A.J., Carstens, B.C., 2008. Does gene flow destroy phylogenetic signal? The performance of three methods for estimating species phylogenies in the presence of gene flow. Mol. Phylogenet. Evol. 49, 832–842. https://doi.org/10.1016/j.ympev.2008.09.008.

Faircloth, B.C., McCormack, J.E., Crawford, N.G., Harvey, M.G., Brumfield, R.T., Glenn, T.C., 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Syst. Biol. 61, 717–726. https://doi.org/10.1093/sysbio/sys004.

Gamble, T., Coryell, J., Ezaz, T., Lynch, J., Scantlebury, D.P., Zarkower, D., 2015. Restriction site-associated DNA sequencing (RAD-seq) reveals an extraordinary number of transitions among gecko sex-determining systems. Mol. Biol. Evol. 32, 1296–1309. https://doi.org/10.1093/molbev/msv023.

Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., LeProust, E.M., Brockman, W., Fennell, T., Giannoukos, G., Fisher, S., Russ, C., Gabriel, S., Jaffe, D.B., Lander, E.S., Nusbaum, C., 2009. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat. Biotechnol. 27, 182–189. https://doi.org/10.1038/nbt.1523.

Gregory, T.R., 2021. Animal genome size database. http://www.genomesize.com.

Huson, D.H., Scornavacca, C., 2012. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. Syst. Biol. 61, 1061–1067. https://doi.org/10.1093/sysbio/sys062.

Irisarri, I., Meyer, A., 2016. The identification of the closest living relative(s) of tetrapods: phylogenomic lessons for resolving short ancient internodes. Syst. Biol. 65, 1057–1075. https://doi.org/10.1093/sysbio/syw057.

Jiang, H., Lei, R., Ding, S.W., Zhu, S., 2014. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinf 15, 1–12. https://doi.org/10.1186/1471-2105-15-182.

Jones, M.R., Good, J.M., 2016. Targeted capture in evolutionary and ecological genomics. Mol. Ecol. 25, 185–202. https://doi.org/10.1111/mec.13304.

Kalaentzis, K., Arntzen, J.W., Avcı, A., van den Berg, V., Beukema, W., France, J., Olgun, K., van Riemsdijk, I., Üzüm, N., de Visser, M.C., Wielstra, B., 2023. Hybrid zone analysis confirms cryptic species of banded newt and does not support competitive displacement since secondary contact. Ecol. Evol. 13. https://doi.org/10.1002/ece3.10442.

Kazilas, C., Dufresnes, C., France, J., Kalaentzis, K., Martinez-Solano, I., de Visser, M.C., Arntzen, Jan.W., Wielstra, B., 2023. Spatial genetic structure in European marbled newts revealed with target enrichment by sequence capture. Mol. Phylogenet. Evol. 194. https://doi.org/10.1016/j.ympev.2024.108043.

Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst. Biol. 56, 17–24. https://doi.org/10.1080/10635150601146041.

Kutschera, V.E., Bidon, T., Hailer, F., Rodi, J.L., Fain, S.R., Janke, A., 2014. Bears in a forest of gene trees: phylogenetic inference is complicated by incomplete lineage sorting and gene flow. Mol. Biol. Evol. 31, 2004–2017. https://doi.org/10.1093/molbev/msu186.

Lemmon, A.R., Emme, S.A., Lemmon, E.M., 2012. Anchored hybrid enrichment for massively high-throughput phylogenomics. Syst. Biol. 61, 727–744. https://doi.org/10.1093/sysbio/sys049.

Lewis, P.O., 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50, 913–925. https://doi.org/10.1080/106351501753462876.

Li, H., 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27, 2987–2993. https://doi.org/10.1093/bioinformatics/btr509.

Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997.

Lischer, H.E.L., Excoffier, L., 2012. PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics 28, 298–299. https://doi.org/10.1093/bioinformatics/btr642.

Marques, D.A., Meier, J.I., Seehausen, O., 2019. A combinatorial view on speciation and adaptive radiation. Trends Ecol. Evol. 34, 531–544. https://doi.org/10.1016/j.tree.2019.02.008.

Maurin, K.J.L., 2020. An empirical guide for producing a dated phylogeny with treePL in a maximum likelihood framework. ArXiv:2008.07054 2020.

McCartney-Melstad, E., Gidis, M., Shaffer, H.B., 2018. Population genomics of the foothill yellow-legged frog. Hered 121, 112–125. https://doi.org/10.1038/s41437-018-0097-7.

McCormack, J.E., Faircloth, B.C., Crawford, N.G., Gowaty, P.A., Brumfield, R.T., Glenn, T.C., 2012. Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. Genome Res. 22, 746–754. https://doi.org/10.1101/gr.125864.111.

McCormack, J.E., Harvey, M.G., Faircloth, B.C., Crawford, N.G., Glenn, T.C., Brumfield, R.T., 2013. A phylogeny of birds based on over 1,500 loci collected by target enrichment and high-throughput sequencing. PLoS One 8. https://doi.org/10.1371/journal.pone.0054848.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20, 1297–1303. https://doi.org/10.1101/gr.107524.110.

Meilink, W.R.M., Arntzen, J.W., van Delft, J.J.C.W., Wielstra, B., 2015. Genetic pollution of a threatened native crested newt species through hybridization with an invasive congener in the Netherlands. Biol. Conserv. 184, 145–153. https://doi.org/10.1016/j.biocon.2015.01.022.

Miller, M.R., Dunham, J.P., Amores, A., Cresko, W.A., Johnson, E.A., 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Res. 17, 240–248. https://doi.org/10.1101/gr.5681207.

Niedzicka, M., Fijarczyk, A., Dudek, K., Stuglik, M., Babik, W., 2016. Molecular Inversion Probes for targeted resequencing in non-model organisms. Sci. Rep. 6, 1–9. https://doi.org/10.1038/srep24051.

Nowoshilow, S., Schloissnig, S., Fei, J.F., Dahl, A., Pang, A.W.C., Pippel, M., Winkler, S., Hastie, A.R., Young, G., Roscito, J.G., Falcon, F., Knapp, D., Powell, S., Cruz, A., Cao, H., Habermann, B., Hiller, M., Tanaka, E.M., Myers, E.W., 2018. The axolotl genome and the evolution of key tissue formation regulators. Nature 554, 50–55. https://doi.org/10.1038/nature25458.

Ortiz, E.M., 2019. vcf2phylip v2.0: convert a VCF matrix into several matrix formats for phylogenetic analysis. https://doi.org/10.5281/zenodo.2540861.

Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T.J., Manuel, M., Wörheide, G., Baurain, D., 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol. 9. https://doi.org/10.1371/journal.pbio.1000602.

Pickrell, J.K., Pritchard, J.K., 2012. Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet 8. https://doi.org/10.1371/journal.pgen.1002967.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., De Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet 81, 559–575. https://doi.org/10.1086/519795.

R Studio Team, 2021. RStudio: Integrated Development for R.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901–904. https://doi.org/10.1093/sysbio/syy032.

Rancilhac, L., Goudarzi, F., Gehara, M., Hemami, M.R., Elmer, K.R., Vences, M., Steinfarz, S., 2019. Phylogeny and species delimitation of near Eastern *Neurergus* newts (Salamandridae) based on genome-wide RADseq data analysis. Mol. Phylogenet. Evol. 133, 189–197. https://doi.org/10.1016/j.ympev.2019.01.003.

Rancilhac, L., Irisarri, I., Angelini, C., Arntzen, J.W., Babik, W., Bossuyt, F., Künzel, S., Lüddecke, T., Pasmans, F., Sanchez, E., Weisrock, D., Veith, M., Wielstra, B., Steinfartz, S., Hofreiter, M., Philippe, H., Vences, M., 2021. Phylotranscriptomic evidence for pervasive ancient hybridization among Old World salamanders. Mol. Phylogenet. Evol. 155, 106967. https://doi.org/10.1016/j.ympev.2020.106967.

Rannala, B., Yang, Z., 2017. Efficient Bayesian species tree inference under the multispecies coalescent. Syst. Biol. 66, 823–842. https://doi.org/10.1093/sysbio/syw119.

Roch, S., Steel, M., 2015. Likelihood-based tree reconstruction on a concatenation of aligned sequence data sets can be statistically inconsistent. Theor. Popul. Biol. 100, 56–62. https://doi.org/10.1016/j.tpb.2014.12.005.

Rubin, B.E.R., Ree, R.H., Moreau, C.S., 2012. Inferring phylogenies from RAD sequence data. PLoS One 7, 1–12. https://doi.org/10.1371/journal.pone.0033394.

Sayyari, E., Mirarab, S., 2016. Fast coalescent-based computation of local branch support from quartet frequencies. Mol. Biol. Evol. 33, 1654–1668. https://doi.org/10.1093/molbev/msw079.

Smith, S.A., O'Meara, B.C., 2012. TreePL: divergence time estimation using penalized likelihood for large phylogenies. Bioinformatics 28, 2689–2690. https://doi.org/10.1093/bioinformatics/bts492.

Smith, J.J., Timoshevskaya, N., Timoshevskiy, V.A., Keinath, M.C., Hardy, D., Voss, S.R., 2019. A chromosome-scale assembly of the axolotl genome. Genome Res. 29, 317–324. https://doi.org/10.1101/gr.241901.118.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033.

Steinfartz, S., Vicario, S., Arntzen, J.W., Caccone, A., 2007. A Bayesian approach on molecules and behavior: reconsidering phylogenetic and evolutionary patterns of the Salamandridae with emphasis on *Triturus* newts. J. Exp. Zool. B Mol. Dev. Evol. 308B, 139–162. https://doi.org/10.1002/jez.b.21119.

Stervander, M., Illera, J.C., Kvist, L., Barbosa, P., Keehnen, N.P., Pruisscher, P., Bensch, S., Hansson, B., 2015. Disentangling the complex evolutionary history of the Western Palearctic blue tits (*Cyanistes* spp.) - Phylogenomic analyses suggest radiation by multiple colonization events and subsequent isolation. Mol. Ecol. 24, 2477–2494. https://doi.org/10.1111/mec.13145.

Stervander, M., Alström, P., Olsson, U., Ottosson, U., Hansson, B., Bensch, S., 2016. Multiple instances of paraphyletic species and cryptic taxa revealed by mitochondrial and nuclear RAD data for *Calandrella* larks (Aves: Alaudidae). Mol. Phylogenet Evol. 102, 233–245. https://doi.org/10.1016/j.ympev.2016.05.032.

Stoltz, M., Baeumer, B., Bouckaert, R., Fox, C., Hiscott, G., Bryant, D., 2021. Bayesian inference of species trees using diffusion models. Syst. Biol. 70, 145–161. https://doi.org/10.1093/sysbio/syaa051.

Than, C., Ruths, D., Nakhleh, L., 2008. PhyloNet: a software package for analyzing and reconstructing reticulate evolutionary relationships. BMC Bioinf. 9. https://doi.org/10.1186/1471-2105-9-322.

Üzüm, N., Avcı, A., Olgun, K., Bülbül, U., Fahrbach, M., Litvinchuk, S.N., Wielstra, B., 2019. Cracking cryptic species: external characters to distinguish two recently recognized banded newt species (*Ommatotriton ophryticus* and *O. nesterovi*). Salamandra 55, 131–134.

van Riemsdijk, I., Arntzen, J.W., Bogaerts, S., Franzen, M., Litvinchuk, S.N., Olgun, K., Wielstra, B., 2017. The Near East as a cradle of biodiversity: a phylogeography of banded newts (genus *Ommatotriton*) reveals extensive inter- and intraspecific genetic differentiation. Mol. Phylogenet. Evol. 114, 73–81. https://doi.org/10.1016/j.ympev.2017.05.028.

van Riemsdijk, I., Arntzen, J.W., Babik, W., Bogaerts, S., Franzen, M., Kalaentzis, K., Litvinchuk, S.N., Olgun, K., Wijnands, J.W.P.M., Wielstra, B., 2022. Next-generation phylogeography of the banded newts (*Ommatotriton*): a phylogenetic hypothesis for three ancient species with geographically restricted interspecific gene flow and deep intraspecific genetic structure. Mol. Phylogenet. Evol. 167. https://doi.org/10.1016/j.ympev.2021.107361.

Wang, Y., Zhou, Y., Li, L., Chen, X., Liu, Y., Ma, Z.M., Xu, S., 2014. A new method for modeling coalescent processes with recombination. BMC Bioinf. 15, 1–12. https://doi.org/10.1186/1471-2105-15-273.

Weisrock, D.W., Smith, S.D., Chan, L.M., Biebouw, K., Kappeler, P.M., Yoder, A.D., 2012. Concatenation and concordance in the reconstruction of mouse lemur phylogeny: an empirical demonstration of the effect of allele sampling in phylogenetics. Mol. Biol. Evol. 29, 1615–1630. https://doi.org/10.1093/molbev/mss008.

Whitfield, J.B., Lockhart, P.J., 2007. Deciphering ancient rapid radiations. Trends Ecol. Evol. 22, 258–265. https://doi.org/10.1016/j.tree.2007.01.012.

Wielstra, B., Arntzen, J.W., 2011. Unraveling the rapid radiation of crested newts (*Triturus cristatus* superspecies) using complete mitogenomic sequences. BMC Evol. Biol. 11. https://doi.org/10.1186/1471-2148-11-162.

Wielstra, B., McCartney-Melstad, E., Arntzen, J.W., Butlin, R.K., Shaffer, H.B., 2019. Phylogenomics of the adaptive radiation of *Triturus* newts supports gradual ecological niche expansion towards an incrementally aquatic lifestyle. Mol. Phylogenet Evol. 133, 120–127. https://doi.org/10.1016/j.ympev.2018.12.032.

Xu, B., Yang, Z., 2016. Challenges in species tree estimation under the multispecies coalescent model. Genetics 204, 1353–1368. https://doi.org/10.1534/genetics.116.190173.

Yu, Y., Warnow, T., Nakhleh, L., 2011. Algorithms for MDC-based multi-locus phylogeny inference: beyond rooted binary gene trees on single alleles. J. Comput. Biol. 18, 1543–1559. https://doi.org/10.1089/cmb.2011.0174.

Zhang, C., Sayyari, E., Mirarab, S., 2017. ASTRAL-III: increased scalability and impacts of contracting low support branches. In: Meidanis, J., Nakhleh, L. (Eds.), Comparative Genomics. Springer International Publishing, Cham, pp. 53–75.

Zheng, X., Levine, D., Shen, J., Gogarten, S.M., Laurie, C., Weir, B.S., 2012. A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics 28, 3326–3328. https://doi.org/10.1093/bioinformatics/bts606.