HAL
open science

Phylotranscriptomic evidence for pervasive ancient hybridization among Old World salamanders

Loïs Rancilhac, Iker Irisarri, Claudio Angelini, Jan Arntzen, Wieslaw Babik, Franky Bossuyt, Sven Künzel, Tim Lüddecke, Frank Pasmans, Eugenia Sanchez, et al.

To cite this version:

Loïs Rancilhac, Iker Irisarri, Claudio Angelini, Jan Arntzen, Wieslaw Babik, et al.. Phylotranscriptomic evidence for pervasive ancient hybridization among Old World salamanders. Molecular Phylogenetics and Evolution, 2021, 155, pp.106967. 10.1016/j.ympev.2020.106967 . hal-03100214

HAL Id: hal-03100214

https://hal.science/hal-03100214v1

Submitted on 29 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Phylotranscriptomic evidence for pervasive ancient hybridization among Old World salamanders

Loïs Rancilhac[1], Iker Irisarri[2], Claudio Angelini[3], Jan W. Arntzen[4], Wieslaw Babik[5], Franky Bossuyt[6], Sven Künzel[7], Tim Lüddecke[8,9], Frank Pasmans[10], Eugenia Sanchez[11], David Weisrock[12], Michael Veith[13], Ben Wielstra[14], Sebastian Steinfartz[15], Michael Hofreiter[16], Hervé Philippe[17], Miguel Vences[1]

1 Zoological Institute, Technische Universität Braunschweig, Mendelssohnstr. 4, 38106 Braunschweig, Germany
2 Dept. Biodiversity and Evolutionary Biology, Museo Nacional de Ciencias Naturales, José Gutiérrez Abascal 2, 28006 Madrid, Spain
3 Via Guglielmo Marconi, 30, 04018 Sezze, Italy
4 Naturalis Biodiversity Center, 2300 RA Leiden, The Netherlands
5 Institute of Environmental Sciences, Jagiellonian University, ul. Gronostajowa 7, 30-387 Kraków, Poland
6 Amphibian Evolution Lab, Biology Department, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels Belgium
7 Max Planck Institute for Evolutionary Biology, 24306 Plön, Germany
8 Department of Bioresources, Fraunhofer Institute for Molecular Biology and Applied Ecology, Winchesterstr. 2, 35394 Gießen, Germany
9 LOEWE Centre for Translational Biodiversity Genomics (LOEWE-TBG), Senckenberganlage 25, 60325 Frankfurt, Germany
10 Department of Pathology, Bacteriology and Avian Diseases, Faculty of Veterinary Medicine, Ghent University, 9820 Merelbeke, Belgium
11 Department of Biology, Stanford University, Stanford, CA 94305, USA
12 Department of Biology, University of Kentucky, Lexington, KY 40506, USA
13 Biogeography Department, Trier University, 54286 Trier, Germany
14 Institute of Biology Leiden, Leiden University, 2300 RA Leiden, the Netherlands
15 University of Leipzig, Institute of Biology, Molecular Evolution and Systematics of Animals, Talstrasse 33, 04103, Leipzig, Germany
16 University of Potsdam, Faculty of Mathematics and Natural Sciences, Institute for Biochemistry and Biology, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany
17 Centre for Biodiversity Theory and Modelling, UMR CNRS 5321, Station of Theoretical and Experimental Ecology, 2 route du CNRS, 09200 Moulis, France

Corresponding author : Loïs Rancilhac (loisrancilhac@gmail.com)

Abstract

Hybridization can leave a signature in the genome of the parental populations through introgression, and genomic approaches are finding this phenomenon to be more frequent than commonly thought. If genomic regions of one of the populations are assimilated by the other, it can confound phylogenetic inference methods that aim to represent evolution as a strictly bifurcating tree. We apply a phylotranscriptomic approach to study the evolutionary history, and test for inter-lineage introgression in the Salamandridae, a prominent Holarctic group of salamanders of considerable interest in studies of toxicity and aposematism, courtship behavior, and molecular evolution. Although the relationships between the 21 currently recognized salamandrid genera have been the subject of numerous molecular phylogenetic studies, some nodes have remained controversial and discordances between mitochondrial vs. nuclear trees have been reported. In order to resolve the phylogeny of this family, and understand the source of the aforementioned mitonuclear incongruences, we generated new transcriptomic (RNAseq) data for 20 salamandrids and use these along with published data, including 28 mitochondrial genomes, to obtain a comprehensive nuclear and mitochondrial perspective on salamandrid evolution. Our final phylotranscriptomic data set included 5,455 gene alignments for 40 species representing 17 of the 21 salamandrid genera. Using concatenation and species tree phylogenetic methods, we find *Salamandrina* sister to the clade of the "True Salamanders" (consisting of *Chioglossa*, *Mertensiella*, *Lyciasalamandra* and *Salamandra*), *Ichthyosaura* sister to the Near Eastern genera *Neurergus* and *Ommatotriton*, *Triturus* sister to *Lissotriton*, and *Cynops* paraphyletic with respect to *Paramesotriton* and *Pachytriton*. Combining introgression tests and phylogenetic networks, we found evidence for introgression among taxa within the clades of "Modern Asian Newts" and "Modern European Newts" but the available data and currently available methods could not unambiguously identify the number, position and direction of introgressive events. Combining the evidence from nuclear gene analysis with the observed mito-nuclear phylogenetic discordances, we formulate a hypothetical scenario with hybridization and mitochondrial capture among ancestral lineages of (1) *Lissotriton* into *Ichthyosaura* and (2) *Triturus* into *Calotriton*, plus introgression of nuclear genes from *Triturus* into *Lissotriton*; furthermore, both mitochondrial capture and nuclear introgression may have occurred among lineages today assigned to *Cynops*. More comprehensive genomic data will, in the future, allow testing this against alternative scenarios involving hybridization with other, extinct lineages of newts.

Key Words: Amphibia, Salamandridae, RNAseq, Phylogenomics, Introgression, Phylogenetic Network

1 Introduction

In phylogenetics, evolutionary relationships are typically represented by bifurcating evolutionary trees. In their simplest form, trees represent the split of population-level lineages that diverge, and remain independent. However, a bifurcating tree in many cases oversimplifies evolutionary histories by neglecting that lineages can hybridize upon secondary contact. In some cases, such hybrid populations follow independent evolutionary trajectories, e.g. after allopolyploidization (Evans 2008), and result in new hybrid species, but hybridization can also leave a signature in the genome of the parental populations through introgression, where part of the genetic material of one of the lineages is assimilated by the other (Arnold 1997, Mallet 2005). The amount of introgressed genomic material, as well as its persistence through time, depends on many factors, including selective pressures. The importance of these processes has been widely investigated in plants (Soltis & Soltis 2009) where the possibility of new species arising through hybridization was already discussed by Linnaeus (Baack & Rieseberg 2007). Mallet (2005) estimated that 25% of plant species are involved in hybridization, mostly between young and closely related species, but sometimes between more distant lineages (Whitney et al. 2010). By contrast, inter-species hybridization and introgression have been traditionally disregarded by zoologists (Mallet 2005). The development of genomic methods in the past decades provided new tools to identify hybrids and understand the role of introgressive hybridization in lineages diversification (Baack & Riesenberg 2007; Payseur & Riesenberg 2016; Irisarri et al. 2018; Palkopoulou et al. 2018). Subsequently, many studies reported inter-species hybridization in various animal taxa, emphasizing that hybridization in natural populations might be much more common than usually thought (Mallet et al. 2016). Thus, although hybrids are rarely found in the wild (Mallet 2005; Mallet et al. 2007), gene flow between distinct taxa should be considered as an important process in shaping genomic diversity.

Genetic exchange between distinct species can confound phylogenetic inference methods that aim to represent evolution as a strictly bifurcating tree. Depending on the gene sampling and the sorting of introgressed alleles, traditional phylogenetic approaches can yield an accurate representation of the relationships followed by the majority of the genome, although they cannot inform on events of gene flow (Leaché et al. 2013). The incongruence among gene trees produced by introgression can be confounded with that of incomplete lineage sorting (ILS), a naturally occurring population process (Degnan & Rosenberg 2009). Hence, distinguishing their respective

contribution remains a major challenge. In addition, ancient introgression can involve now extinct species and thus be more difficult to detect. While the application of phylogenetic inference methods that account for ILS is commonplace, mainly in the framework of the multispecies coalescent (MSC), introgression has been widely ignored in large scale phylogenetic studies (Eckert & Carstens 2008). The extension of the MSC into the Multi-Species Network Coalescent (Degnan 2018) allowed the development of models accounting for both ILS and introgression as source of variation among gene trees. Although phylogenetic network estimation under this model remains a statistical challenge, particularly with large data sets, several implementations are already available based on maximum parsimony, maximum likelihood or Bayesian approaches. Alternative methods to detect ancient hybridization events include Patterson's D statistic (also named "ABBA/BABA" introgression tests) that is based on alignment site patterns (Green et al. 2010; Durand et al. 2011). Alongside this, the development of high throughput sequencing techniques makes it now possible to sample hundreds, or even thousands, of independent loci in non-model organisms for molecular analysis. This provides an unprecedented opportunity to study the variation of gene genealogies at the genomic scale and thus get insights into potential past hybridization events. As a result, phylogenomic studies have increasingly reported introgression in a wide range of taxa (Folk et al. 2018).

In this study, we apply a phylotranscriptomic approach to study the evolutionary history, and test for inter-lineage introgression, in a a diverse clade of salamanders distributed in Europe, Northern Africa, Asia and North America, classified in the family Salamandridae (Sparreboom 2014). The currently recognized 118 salamandrid species and 21 genera (Amphibiaweb 2019) comprise aquatic and semiaquatic newts with sexual dimorphism and complex nuptial displays during their aquatic phase (Arnold 1977) as well as semi- and fully terrestrial salamanders with complex life histories and reproductive modes (Buckley 2012, Lourenço et al. 2019) involving specialized courtship pheromones (Van Bocxlaer et al. 2016). Equally intriguing are their potential anti-predator defensive systems that include biosynthesis of steroidal alkaloids (Lüddecke et al. 2018), accumulation of and resistance to tetrodotoxins (Geffeney et al. 2002, Hanifin & Gilly 2014), and putatively aposematic, highly contrasting coloration in several taxa. Studying the evolution of these traits (Veith et al. 1998; Steinfartz et al. 2006; Wiens et al. 2011; Kieren et al. 2018) requires a well resolved evolutionary history.

Molecular phylogenies, mainly based on mitochondrial genes, have provided numerous surprising insights into the evolution of salamandrids (Titus and Larson, 1995; Veith et al. 1998; Weisrock et al. 2006; Steinfartz et al. 2006; Zhang et al. 2008; Veith et al. 2018), such as the non-monophyly of the "Modern European Newts" in the former genus *Triturus*, now separated into *Triturus*, *Ichthyosaura*, *Lissotriton*, and *Ommatotriton*, despite striking similarities in their aquatic-

phase courtship and crest ornamentation (Steinfartz et al. 2006). Similarly, the former genus *Euproctus* of European mountain stream newts found on the islands of Corsica and Sardinia as well as in the Pyrenees of the European mainland turned out to be non-monophyletic, resulting in the recognition of the genus *Calotriton* for the Pyrenean species (Carranza & Amat, 2005). Neither Asian nor European salamandrids formed reciprocally monophyletic groups, suggesting a convoluted evolutionary history stemming from multiple vicariant and dispersal events (Kieren et al. 2018). However, despite high support for many nodes of the salamandrid tree, several branches remain poorly supported or discordant among different studies, as reviewed by Veith et al. (2018). A recent tree inferred from sequences of selected nuclear genes (Veith et al. 2018) revealed discordances with mitochondrial phylogenies regarding the position of some European (*Calotriton*, *Lissotriton*, *Ichthyosaura*) and Asian (*Cynops*) newts. Although these results may have been biased due to the limited amount of nuclear DNA data included, biological processes such as ILS (McKay & Zink 2010) or introgression (of either nuclear genes or the mitochondrial genome, Wallis et al. 2017) could underlie the observed discordances. Indeed, hybridization has been described in the wild in several genera of Salamandridae, such as *Lissotriton* (Babik et al. 2003), *Lyciasalamandra* (Johannesen et al. 2006), *Triturus* (Arntzen et al. 2014) and *Pleurodeles* (Escoriza et al. 2016). Additionally, hybridization between genera has been reported under experimental conditions, including *Ichthyosaura*, *Lissotriton*, *Ommatotriton*, and *Triturus* (e.g. Pariser 1932; Mancino et al. 1978; Macgregor et al. 1990), or *Pleurodeles* and *Tylototriton* (Ferrier & Beetschen 1973). Thus, the use of a high number of independent nuclear loci and methods that account for introgression may be necessary to resolve the disputed nodes in the salamandrid tree.

Using a phylotranscriptomic pipeline involving extensive quality controls (Irisarri et al. 2017; Simion et al. 2017), as well as new analyses of mitochondrial sequences, we here provide rigorous tests of the topologies proposed by previous studies. We particularly focused on testing the conflicting inter-generic relationships recovered by mitochondrial genomes (Zhang et al. 2008) *versus* nuclear genes (Veith et al. 2018). We investigate the presence of introgression among salamandrid lineages using phylogenetic network reconstruction and introgression tests based on gene trees topologies. By combining these different markers and approaches, we aim to provide a more comprehensive understanding of the evolutionary history of the Salamandridae.

2 Methods

2.1 Phylogenetic analysis based on mitochondrial genomes

In order to get an alternative view of the mitochondrial phylogeny of the Salamandridae, and assess the potential effect of the employed substitution model on the inferred topology, we analyzed the available data under the CAT-GTR model. By accounting for site-specific frequency profiles, CAT-GTR can outperform site-homogeneous substitution models (Lartillot & Philippe 2004), particularly in the case of non-recombining fragments. Thus, it is a relevant approach for the analysis of miochondrial genes.

Protein sequences of 13 mitochondrial genes were recovered from GenBank for 30 salamandrid species (Table S1) and aligned using mafft (Katoh & Standley 2013). Phylogenetic inference was then performed using PhyloBayes (Lartillot & Philippe 2004, 2006; Lartillot et al. 2007), with 1500 cycles and a burnin of 500, and the consensus tree was recovered using the command *bpcomp*. To further assess the support of the branches, 100 bootstrap replicates were generated with seqboot (http://evolution.genetics.washington.edu/phylip/doc/seqboot.html) and analyzed in PhyloBayes, with the same settings as above.

2.2 Transcriptome sequencing, data set filtering and assembly

We carried out RNA sequencing (RNAseq) of transcriptomes of representative salamandrid taxa on different Illumina platforms (as specified in Table S2) and at varying sequencing depths, partly in the context of other projects (Stuglik & Babik 2016; Maex et al. 2018; Wielstra et al. 2019). Specifically for the present project we newly sequenced 8 transcriptomes from 100 mg of tissue samples per specimen, consisting of combined or separate skin, muscle, or liver samples preserved in RNAlater and frozen at -80°C. RNA extraction was carried out using standard trizol protocols (for detailed protocol see Supplementary Methods). Sequencing was carried out with a High Output 150 cycle kit, in runs that combined ca. 15 samples each. Illumina reads were quality-trimmed and filtered using Trimmomatic v. 0.32 (Bolger et al. 2014) with default settings and later filtered for rRNA sequences with SortMeRNA (Kopylova et al. 2012). Paired and unpaired filtered reads were used for *de novo* transcriptome assembly using Trinity v. 2.1.0 (Grabherr et al. 2011) following published protocols (Haas et al. 2013). The same assembly approach was used for various RNAseq data sets downloaded as raw reads from the NCBI Sequence Read Archive (SRA, Table S2); all new RNAseq data were submitted to SRA under Bioproject PRJNA607429 and assemblies and alignments are available from Figshare under DOI 10.6084/m9.figshare.11778672.

As a basis to extract orthologous genes from the transcriptomes, we employed an alignment covering all jawed vertebrate classes, assembled by Irisarri et al. (2017). We used the software “42” (D. Baurain, https://metacpan.org/release/Bio-MUST-Apps-FortyTwo) to align our transcripts to

this reference data set, which controls for orthology relationships using multiple references in strict three-way reciprocal best BLAST hit tests. For details of the pipeline, see Irisarri et al. (2017).

To remove possible remaining paralogs and contaminations from the data set, we followed the pipeline described in Irisarri et al. (2017). Briefly, (1) putative contaminations were first identified as significant BLAST hits against a custom database of ribosomal proteins containing a large diversity of eukaryotic species; (2) in cases where several homologous transcripts per taxa were present in single gene alignments, redundant (i.e. > 95% of length overlap) or too divergent sequences were removed; (3) putative cross-contaminations were identified by comparing patristic distances between sequences in gene trees (inferred with RAxML under a GTRCAT model; Stamatakis et al. 2018) vs in a concatenation tree; (4) gene alignments containing ancient paralogs were split using newly-inferred gene trees (as above) and dividing the sequences into two groups based on the tree branch separating the paraloguous clades (in practice, a long branch that maximized the taxonomic diversity in the two sub-alignments). These steps were done at the level of amino acids and including all the species in the original data set, after which the salamandrids and several outgroups were extracted and the original nucleotide sequences were retrieved using leel (D. Baurain, https://metacpan.org/release/Bio-MUST-Apps-FortyTwo) for further analyses.

For eight species, RNAseq data were available for two individuals each. In preliminary analyses, these conspecific samples were kept separate to investigate possible events of contamination or hybridization that would confound phylogenetic inference. First, we selected the individuals so that none of the conspecifics were sampled in potential hybrid zones. Then, we assembled a preliminary concatenated alignment using ScaFos v1.25 (Roure et al. 2007) without merging conspecifics and by building chimeras within individuals when several transcripts were available at the same gene. A phylogenetic tree was inferred using RAxML (under a GTR+Γ substitution model) with 50 rapid bootstrap replicates and a subsequent maximum-likelihood (ML) search. This confirmed that conspecific individuals consistently formed monophyletic groups and were separated by short branches from each other (Fig. S1). Thus, to maximize the amount of data for these species, we assembled a final data set by merging conspecifics using ScaFos, with the same setting as above. The final data set contained 31 taxa and 5,455 gene alignments (a total of 9,546,906 aligned bp). To control for the effect of low-covered taxa on the phylogenetic reconstructions, we assembled a second data set after omitting two species that were present in <50% of gene alignments, *Pachytriton brevipes* and *Ommatotriton ophryticus.* Sequences of these two species were present in only 9% and 29% of the alignments, respectively, either because the respective data were recovered from a gut metatranscriptome with only a limited number of host reads (in the case of *Pachytriton*), or due to sequencing failure (*Ommatotriton*).

2.3 Phylotranscriptomic analyses

We inferred a ML tree from the concatenated matrix with IQ-TREE v1.6.8 (Nguyen et al. 2015; Chernomor et al. 2016) using the best-fitting substitution models and gene-partitions selected with BIC in ModelFinder (Kalyaanamoorthy et al. 2017), as implemented in IQ-TREE. The support of the branches was assessed using the SH-like approximate likelihood ratio test (aLRT) with 1,000 pseudoreplicates. To further assess branch support with a more stringent criterion and to characterize the amount of data required to stabilize each branch in the tree, we performed a gene jackknifing analysis (Delsuc et al. 2008; Irisarri et al. 2017). For this, we randomly sampled gene alignments without replacement and produced sets of 100 concatenated matrices of increasing lengths of approximately 10 Kbp; 50 Kbp; 100 Kbp; 500 Kbp; 1,000 Kbp and 5,000 Kbp. For the 100 pseudoreplicates of each of these lengths, ML trees were inferred using RAxML rapid hill-climbing algorithm (GTR+Γ). For each matrix length, we calculated jackknife support values as the number of times a given bipartition is recovered among the 100 ML trees. These analyses were performed on both the full data set as well as the data set without the two low-coverage species.

In order to account for the effect of incomplete lineage sorting (ILS) on phylogenetic inference, we also inferred a tree from the full data set using ASTRAL-II (Mirarab & Warnow 2015), a summary-tree method statistically consistent with the Multi-Species Coalescent (MSC). Gene trees were inferred using PhyloBayes (Lartillot and Philippe 2004, 2006; Lartillot et al. 2007) under a CAT-GTR model, with 1100 cycles and a burnin of 100. Branch support was assessed using local posterior probabilities and quartet scores (i.e. the proportion of the most common quartet in the gene trees pool supporting a given branch, Sayyari et al. 2016).

2.4 Phylogenetic network inference and introgression tests

The presence of reticulations among salamandrid lineages was tested using two different approaches. First, we inferred a phylogenetic network using the coalescent-based approach implemented into PhyloNet 3.7.1 (Than et al. 2008; Wen et al. 2018). We ran the maximum partial-likelihood algorithm (InferNetwork_MPL command, Yu & Nakleh 2015) on the gene trees previously used for the ASTRAL analysis. According to the mito-nuclear discordances, as well as quartet scores from the nuclear transcriptomic data (c.f. results below), putative hybridization events were located within the “modern newts” group. Arguably, the discordances could also have been caused by introgression events involving genera outside of the “modern newts” (e.g. between *Calotriton* and *Euproctus*, or even a more distant genus), but this hypothesis was not supported in preliminary introgression tests (table S3), and is very unlikely given the present distribution and

ecological characteristics of the taxa. Thus, we pruned the gene trees to keep only the "modern newts", as well as *Pleurodeles* as outgroup, to reduce the computational burden of the analyses. As we focused on hybridization events between genera prior to their diversification, the following species (the ones with the highest coverage) were selected as representatives of their genus when several where available: *Triturus marmoratus*, *Lissotriton montandoni* and *Ommatotriton nesterovi*. This decision introduces new limitations to our analyses, as more recent inter-genera introgression events involving unsampled taxa could influence the results, but compromises had to be made to optimize both the computational burden and interpretability of the results. Since the monophyly of *Cynops* was not recovered in the concatenation analyses (c.f. results below), both species were kept.

In order to determine the best-fitting number of reticulations, we ran PhyloNet assuming maximum numbers of reticulations ranging from 0 to 10, with 100 independent runs performed for each value. For each of these 11 analyses, the five models with the highest likelihood were kept, and the 55 retained models were ranked using the Akaike Information Criterion (AIC) as in Yu et al. (2014). We finally retained the 10 best models for further investigations, and used Dendroscope 3 (Huson & Scornavacca 2012) to visualize them.

Another popular approach to detect introgression is Patterson's D statistic, also known as the ABBA/BABA test (Green et al. 2010; Durand et al. 2011). In brief, this method tests for introgression in rooted asymmetric four-taxa trees (outgroup,(H1,(H2,H3)) using unlinked bi-allelic markers (usually SNPs). By comparing the proportion of site patterns, this test allows to determine the presence of introgression between H1 and H2 (excess of ABBA pattern) or between H1 and H3 (excess of BABA), distinguishing it from random processes (i.e. ILS, generating similar proportions of ABBA and BABA patterns). Applying the original version of this test to our transcriptomic data would require to sample a single SNP per gene alignment, in order to meet the non-linkage requirement. Since we are working with coding regions with low variability, this would highly reduce the informativeness of the data, and thus reduce the statistical power of the test by introducing stochastic errors. To overcome this problem, we decided to use gene trees, calculated from our gene alignments, as markers for introgression tests rather than SNPs (in a comparable way to Węcek et al. 2016 and Barlow et al. 2018). Under ILS alone, discordant branching patterns (equivalent to ABBA and BABA patterns in SNPs) should be present in four-taxa gene trees in even proportions, while introgression would generate excess in one of them (assuming the inferred gene trees are correct). To test that, we applied the following procedure on the gene trees used for the ASTRAL analysis (Fig. S3): (1) we selected only the gene-trees containing all the three focal taxa (noted H1, H2 and H3) and an outgroup, and collapsed the nodes with a Posterior Probability (PP)<70 to control for phylogenetic uncertainty; (2) we pruned gene trees to keep the four relevant taxa and discarded those with polytomies (i.e. nodes with PP<70 collapsed at the previous step); (3)

we rooted the 4-taxa trees using the outgroup; (4) we counted the number of occurrences of the two discordant patterns (respectively (H2, (H1, H3)) and (H3, (H1, H2))) and (5) we used these counts to calculate a statistic similar to Patterson's D, as follow $\frac{\left(N(H2(H1,H3))-N(H3,(H1,H2))\right)}{\left(N(H2(H1,H3))+N(H3,(H1,H2))\right)}$. As for the Patterson's D, a statistic significantly different from zero would indicate introgression. The significance of the statistic was assessed using 1000 bootstrap replicates (see Fig. S4 for a graphical summary of the whole procedure). This test was implemented in a custom R function (available at https://github.com/rancilhac/Introgression-tests-from-gene-trees) based on the ape package (Paradis et al. 2004). As in the network analyses, we focused on the two clades of "modern newts", with *Pleurodeles* as an outgroup. In both these clades, the "Modern Asian Newts" and "Modern European Newts" separately, the test was performed for every four-taxon combination that fitted the model's requirement (summarized in Fig. 4, Fig. 5 and Fig. S5), with the results of the IQTREE analysis used as a reference topology.

3 Results

3.1 New mitogenome analyses recover Salamandrina *sister to the "True Salamander" clade*

Analysis of the mitogenomic data set yielded a topology that was overall similar to that of Zhang et al. (2008) (Fig. 1a), recovering all the main clades defined in previous studies. However, one notable difference to previous studies (Veith et al. 2018) was the position of *Salamandrina* which in our analysis is sister to the "True Salamanders" (i.e. *Salamandra*, *Lyciasalamandra* and *Chioglossa*), although with low support (bootstrap probability, BP=0.6). It is also worth noting that the monophyly of *Cynops* and the position of *Ichthyosaura* as sister lineage to *Lissotriton* both received a very low support (respectively BP=0.42 and BP=0.53) in our analysis. All the other branches received BP values >0.9, except for the intrageneric relationships of both *Triturus* and *Salamandra*.

3.2 Phylotranscriptomic analyses confirm mito-nuclear discordances

The final RNAseq-derived sequence matrix included 31 salamandrid and two outgroup species, representing all salamandrid genera except for *Mertensiella*, *Taricha*, *Liangshantriton* and *Laotriton*. Our alignment consisted of 5,455 genes of a length ranging from 300 bp to 15,846 bp each, and a total of 9,546,906 bp, including 2,887,639 (30.25%) variable and 1,498,535 (15.70%) parsimony-informative sites. Analysis with IQTREE yielded a strongly supported topology where

all branches received 100% SH-like aLRT support (Fig. 1b). The exclusion of the two taxa with the highest proportion of missing data (91% and 71%) did not change the topology nor the support values (Fig. S2). Species-tree inference performed with ASTRAL also recovered the same topology (Fig. S3), with every node having local posterior probabilities of 1.0. This topology recovers the monophyly of the main mitochondrial clades, and confirms the new position of *Salamandrina* as sister to the "True Salamanders". However, the nuclear tree also confirms the mito-nuclear discordances suggested by Veith et al. (2018) among the "Modern European Newt" clade. The nuclear tree places *Ichthyosaura* nested within a group formed by *Neurergus* and *Ommatotriton*, *Lissotriton* as the sister lineage to *Triturus* and *Calotriton* as sister lineage to the remainder of the group. In contrast, in the mitochondrial tree, *Ichthyosaura* is sister to *Lissotriton* and *Calotriton* sister to *Triturus*. Within the "Modern Asian Newts", *Cynops cyanurus* forms a monophyletic group with *Paramesotriton* and *Pachytriton* in the nuclear tree, resulting in the paraphyly of the genus *Cynops* while the mitochondrial tree recovers it as monophyletic. The relationships within the genus *Salamandra* also differ between the two trees, with *S. infraimmaculata* being alternatively sister species to *S. salamandra*, or to all the other species.

Gene jackknifing analyses further confirmed that the RNAseq-derived topology is well resolved, most branches receiving high support even with relatively little data: with only 10 Kbp sampled (~6 genes), the average jackknife proportion among all the branches is 84%, and it is >95% with 100 Kbp (~57 genes) and onwards (Fig. 2a). However, a few nodes required more data to resolve, as shown in Figure 2b-f. These were the nodes within the genus *Salamandra*, the deep nodes of the "Modern European Newts", and the nodes of the "Modern Asian Newts". Interestingly, those nodes are also supported by rather low proportions of quartet in the ASTRAL tree, although with posterior probability of 1.0 (Fig. S3), underlining high variation among gene-tree topologies.

3.3 Phylogenetic networks and introgression tests suggest pervasive introgression in modern newts

We inferred a total of 55 phylogenetic models that recovered 0–6 reticulations through 11 PhyloNet analyses, with maximum numbers of reticulations allowed ranging from 0 to 10. AIC scores supported models with at least one reticulation as substantially more likely than those with none (Fig. 3a). However, the AIC values did not allow to clearly favor a specific number of reticulations, and besides the best model with 5 reticulations (Fig. 3b), we therefore decided to also consider the 9 sub-optimal networks (Table 1, Fig. 3c, Fig. S6) with 3–6 reticulations. While the limit to 10 networks is arbitrary, we selected it as a compromise to capture the variation within and among runs of PhyloNet, while narrowing our sample to sufficiently few models with low AIC.

The best ranking network (Fig. 3b) differs sensibly from the phylotranscriptomic tree. In both, *Neurergus*, *Ichthyosaura* and *Ommatotriton* form a monophyletic group, and the two *Cynops* species are paraphyletic in regard to *Paramesotriton*. However, in the best ranking network (Fig. 3b), *Triturus* is sister lineage to the *Neurergus*, *Ichthyosaura* and *Ommatotriton* clade. *Calotriton* is placed as sister to the aforementioned taxa, and *Lissotriton* is sister to all the "Modern European Newts". *Paramesotriton* and the two species of *Cynops* also form a monophyletic group, with relationships similar to the phylotranscriptomic tree. In the 9 sub-optimal networks the topology is variable, caused by unstable positions of *Calotriton* and *Lissotriton*. The most common alternative branching for these taxa (4 out of 10 networks, fig. 3c) places *Calotriton* as a sister lineage to all other taxa, while *Lissotriton* is the sister lineage to the *Paramesotriton*/*Cynops* clade.

The best ranking network features five reticulations. One of them links the ancestor of all the ingroup to the ancestor of the "Modern European Newts", while three others involve a putative extinct lineage sister to *Ommatotriton*. Three of the hybridization branches, however, have inheritance values that deviate strongly from the expected 0.50, with one as low as 0.002. Furthermore, the number (from 3 to 6) and position of the reticulations are very variable when considering the nine sub-optimal networks (Fig. 3b, 3c, S6). On the other hand, the fifth reticulation, placing *Lissotriton* as the descendant of a hybridization event between *Triturus* and an extinct lineage, was recovered in the 10 considered networks, with inheritance probabilities comprised between 0.4 and 0.6.

For "Modern European Newts", we performed introgression tests for a total of 20 combinations of four taxa, based on the topologies of 1257–2961 gene trees. Out of these, 11 yielded a significant signal of introgression (i.e. the statistic differed from 0 with $p \leq 0.05$, summarized in Fig. 4). The combinations that did not yield significant signals for introgression are summarized in Figure S3. The test including *Triturus*, *Lissotriton* and *Calotriton* yielded significant signal of introgression between *Triturus* and *Calotriton* (Fig. 4a). Tests including *Calotriton*, *Lissotriton* and either *Ommatotriton*, *Ichthyosaura* or *Neurergus* all yielded significant introgression between *Calotriton* and the three latter (Fig. 4b), and the same pattern was observed when testing *Triturus* instead of *Calotriton* (Fig. 4c). However, when including *Calotriton*, *Triturus* and either *Ommatotriton*, *Ichthyosaura* or *Neurergus*, no introgression could be detected (Fig. S4). Similarly, testing *Calotriton* against each possible taxon pair of the *Ommatotriton* /*Ichthyosaura*/*Neurergus* group did not yield any introgression signal (Fig. S5). Contrary to *Calotriton*, *Triturus* consistently showed introgression with *Ommatotriton*, relative to both *Neurergus* and *Ichthyosaura* (Fig. 4d). Finally, both *Lissotriton* and *Neurergus* yielded significant signals of introgression with *Ommatotriton* relative to to *Neurergus* and *Ichthyosaura* respectively (Fig. 4e).

For “Modern Asian Newts”, only four combinations of four taxa could be tested, three of which resulted in significant signals for introgression (summarized in Fig. 5). These tests were performed on only 139 to 2072 gene trees, the lowest numbers being due to the very low coverage of the *Pachytriton brevipes* transcriptome. Introgression tests between the two *Cynops* species relative to *Paramesotriton* and *Pachytriton*, respectively (Fig. 4a) gave significant positive signals, as did tests between *C. pyrrhogaster* and either *Pachytriton* or *Paramesotriton* (Fig. 4b). *C. cyanurus*, on the other hand, did not show significant introgression with any of these two taxa (Fig. S5).

4 Discussion

4.1 Phylotranscriptomic data resolve salamandrid phylogeny and confirm mito-nuclear discordances

Our analyses of both mitochondrial sequences and nuclear phylotranscriptomic data provide the most comprehensive phylogenomic assessment of the evolutionary relationships of salamandrids to date, and confirms discordances between mitochondrial and nuclear genomes. The two data sets yielded overall similar topologies which also agreed with those inferred in previous studies. The family is divided into two major clades, the “True Salamanders” and the “Newts”. Most of the previously defined major groups within these clades were also confirmed by our data. This is particularly relevant for the newts, in which the Asian and European taxa did not form reciprocally monophyletic groups, confirming their complex biogeographical history implying several events of vicariance and dispersal (Zhang et al. 2008; Kieren et al. 2018). However, some other relationships recovered by our phylotranscriptomic analysis differed from previous topologies. Perhaps most significantly, *Salamandrina*, a taxon most often found as sister lineage to all the other salamandrids (Veith et al. 2018), forms a monophyletic group with the “True Salamanders” in both the mitochondrial and the nuclear transcriptomic trees. While this position received a rather low support in the mitochondrial tree (BP=0.6), it is fully supported in the nuclear data set by the IQTREE and ASTRAL analyses, and by more than 90% of the short gene-jackknifing replicates already with 50 Kbp (Fig. 2b). This suggests that the previously recovered position of *Salamandrina* was the result of an artifact in phylogenetic reconstruction, possibly due to its long branch. Both the low quartet score of this branch (q=0.51) in the ASTRAL analysis and the short internal branch in the ancestor of *Salamandrina* + “True Salamanders” suggest the presence of ILS, which would cause difficulties for phylogenetic inference (Degnan & Rosenberg 2006). However, this low quartet score could also be the result of errors in the gene trees, as such deep and short branches can be difficult to recover.

Our analyses confirmed most of the previously detected mito-nuclear incongruencies. Firstly, the genus *Cynops* is paraphyletic in our phylotranscriptomic trees. The position of *C. cyanurus* as sister to *Pachytriton* and *Paramesotriton* is strongly supported in our trees, but 500 Kbp are necessary to recover them with high support (Fig. 2f). When repeating this analysis after removing *Pachytriton brevipes* which has a very low gene coverage, the support for a *C. cyanurus* + *Paramesotriton* clade increased substantially (from 74% to 90% at 100 Kbp). The monophyly of *Cynops* has been consistently supported by complete mitogenome analyses (Kieren et al. 2018; Zhang et al. 2008), and our mitochondrial analysis also supports it, although with only low support (BP = 0.43). A previous study focusing on the *Cynops* phylogeny also found this genus to be paraphyletic with respect to *Paramesotriton* and *Pachytriton* (Tominaga et al. 2013), although based only on one mitochondrial gene. While some authors have already placed *C. cyanurus* and the other Chinese species of *Cynops* into a different genus, *Hypselotriton* (Dubois & Raffaëlli 2009), a more extensive taxon sampling would be necessary for reliable conclusions on the evolutionary relationships and taxonomy of these newts.

Other important mito-nuclear incongruencies affect the clade of "Modern European Newts". In the mitochondrial topologies, including the one inferred in the present study, *Calotriton* is sister to *Triturus*, and *Lissotriton* to *Ichthyosaura*, albeit poorly supported. In the phylotranscriptomic tree however, *Triturus* and *Lissotriton* are sister taxa while *Ichthyosaura* is nested within *Neurergus* and *Ommatotriton*, and *Calotriton* is sister to all these five taxa. These results are partially concordant with the four nuclear gene phylogeny from Veith et al. (2018), the only difference being that *Ommatotriton* is sister to *Ichthyosaura* rather than to *Neurergus*. All relevant branches received full SH-like aLRT support in the ML analysis of the phylotranscriptomic data, as well as maximum local posterior probability in the ASTRAL tree. While some of the bipartitions received a low support from the shorter jackknife pseudo-replicates when few positions were sampled (Fig. 2 c-d), they all stabilized at full (100%) support with ≥500 Kbp on. Interestingly, most of these contentious nodes received a rather low quartet score in the ASTRAL analysis (Fig. S3), suggesting a relatively high level of variation among gene genealogies.

One possible explanation for the observed discordances between the phylogenetic signal of mitochondrial and nuclear markers could be ILS affecting the nuclear markers. In this scenario, the mitochondrial genome could be a more reliable marker than nuclear genes due to its smaller effective population size and faster sorting (Harrison 1989), that make it less sensitive to ILS. However, the concordance between the concatenation and ASTRAL analyses of the nuclear markers suggests this is not the case, as the later approach is theoretically able to recover the correct topology even under high levels of ILS (Mirarab & Warnow 2015). This is concordant with previous studies where concatenation and ASTRAL approaches were found to converge to very

similar topologies in the presence of ILS, when using large amounts of molecular data (e.g. Irisarri et al. 2018).

4.2 Inter-species gene flow as a source of incongruence?

A second possible explanation for the observed mito-nuclear incongruences is gene flow between lineages, i.e., introgression of nuclear genes and/or mitochondrial capture. In salamandrids, we hypothesize that ancient introgressive hybridization between the ancestors of different genera explains the variations in nuclear genes genealogies, as well as the differences between mitochondrial and nuclear phylotranscriptomic topologies in both the "Modern European Newts" and "Modern Asian Newts". When performing phylogenetic network analyses on a subset of taxa including all of the modern newts, introducing reticulations in the models substantially increased their likelihoods, supporting the presence of introgression within this group. However, the exact number of reticulations, as well as their position, were more difficult to assess. In that aspect, it has been shown that sequence and gene tree data are often not informative enough to distinguish between complex networks with several reticulation occuring along the same branch (Pardi et al. 2015), which might be the cause of the uncertainties in our results. The best network, according to comparison of AIC values, has five reticulations. We also considered nine other sub-optimal models for comparison, which displayed between three and six reticulations. Considering the results displayed on Figure 3a, one could hypothesize that adding reticulations to the phylogenetic models artificially increases their likelihood, regardless of the true hybridization events. This result is partially supported by the very low inheritance probabilities of some hybridization branches (Yu et al. 2012), suggesting that at least some of the recovered reticulations could be over-fitted. However, we allowed up to 10 reticulations in our analyses but never recovered more than six, which might mean that the signal in our data is strong enough to minimize over-fitting.

When comparing the best model with the nine sub-optimal ones, it appears that the topology of the recovered networks, as well as the number of hybridization branches and their locations, are quite variable. However, it is possible to identify some consistent features among all these networks. Regarding the topology, two groups seem to have unambiguous relationships. First, among the "Modern European Newts", a monophyletic group comprising *Triturus*, *Neurergus*, *Ichthyosaura* and *Ommatotriton* (branching hierarchically in that order; this group is hereafter referred to as the "TNOI" group) is supported by all the 10 networks. The monophyly of this group relative to *Lissotriton* is also consistent with our introgression tests, albeit the latter do not allow to further precise the relationships within the "TNOI" group. Among the "Modern Asian Newts", the

relationships are similar to those recovered on the phylotranscriptomic tree, with the genus *Cynops* being paraphyletic with respect to *Paramesotriton* (*Pachytriton* was excluded from this analysis because of its low coverage). Thus, it appears that the uncertainty in the networks topologies actually results from two unstable taxa, *Calotriton* and *Lissotriton*. The number and position of the reticulations in the different networks is also very variable. All the recovered hybridization branches are relatively deep and thus involve extinct populations. One possibility to explain such a pattern could be the presence of strong genetic structure within the ancestral populations of the extant species, which would leave a signal very similar to gene flow (Slatkin & Pollack 2008). Although models have been developed that are theoretically able to differentiate between gene flow and structure in ancestral populations (Theunert & Slatkin 2017), these are not yet scalable to data sets with high numbers of both loci and taxa, such as ours. A wrong phylogenetic placement of *Calotriton* and *Lissotriton* could also artificially inflate the signal for gene flow, although PhyloNet is theoretically able to address that issue by allowing topological rearrangements. Because of these inconsistencies, the network analyses do not allow us to determine unambiguously the number and positions of the reticulations in our tree. However, by comparing the results of our different analyses, including the mitochondrial-nuclear discordances, we could identify common patterns supporting several introgression events, discussed in detail thereafter. We summarized them in a hypothetical evolutionary scenario for the "modern newts", represented in Figure 6.

Although most reticulations are poorly supported, it is interesting to note that all the considered networks recovered *Lissotriton* as originating from a hybridization event between *Triturus* and an unsampled lineage. This hybridization event is further supported by the inheritance probabilities, always comprised between 0.4 and 0.6, suggesting a close-to-equal contribution of both parental populations to the genetic material of *Lissotriton* (Yu et al. 2012). As hypothesized earlier, this reticulation is matching one of the discordances between the mitochondrial and nuclear trees, the two genera being grouped together on the latter. This result supports a scenario (Fig. 6) in which *Lissotriton* inherited part of its genetic material (including its mitochondrial genome) from a lineage that diverged early within the "Modern European Newts", and the remaining nuclear genes from the *Triturus* lineage following introgressive hybridization. Although the introgression tests did not allow to directly investigate introgression between these two genera, a high proportion of the nuclear genes supported the monophyly of the "TNOI" group in respect to *Lissotriton*, which is consistent with this hypothesis. However, it is important to note that such pattern could also arise if *Lissotriton* shared a sister relationship with *Triturus*, with subsequent introgression of nuclear genes from a more distant, unsampled lineage into *Lissotriton* (Fig. 6), although such hypothesis seems to contradict the networks. *Lissotriton* species are widespread ecological generalists, often occurring in sympatry with *Triturus* species, and the two genera share an overall similar reproductive

behavior. Given that experimental hybridization between *Triturus* and *Lissotriton* (and *Lissotriton* and *Ichthyosaura*) can result in viable offspring (Pariser 1932; Mancino et al. 1978) it is reasonable to assume that their respective ancestors could have successfully hybridized in the wild. Some networks suggest that ancestral populations of *Lissotriton* have exchanged genes with other lineages. However, as discussed above, inconsistent reticulations could be the result of artifacts and should rather be taken cautiously in the absence of further evidences.

The case of *Calotriton* is more convoluted. The observed mito-nuclear discordances could be explained by mitochondrial introgression from *Triturus* into the *Calotriton* lineage, leading to their sister relationship in the mitochondrial phylogeny. From a biological perspective, their highly different mating behavior makes hybridization unlikely, although occasional occurrence in spatial proximity can be observed (e.g. of *T. marmoratus* with *C. arnoldi* and with *C. asper* in north-eastern Spain). However, it is important to keep in mind that the suspected introgression probably occurred millions of years ago, implying that both the distribution and behaviour of the considered populations might have been very different. For instance, Pleistocene fossil remaining tentatively assigned to *Euproctus* (likely referring to *Calotriton*) have been found in the Spanish region of Asturias (Sanchiz 1977), far outside the current range of *Calotriton* but climatically more suitable during glacial maxima, as revealed by modelling (Carranza and Amat 2005). This suggests that even during rather recent Pliocene-Pleistocene times, the range of *Calotriton* may have been intermittently overlapping with that of *Triturus* more widely than today. Introgression tests suggest *Calotriton* to be monophyletic with the "TNOI" group relative to *Lissotriton*, but do not support significant introgression with *Triturus* when compared to either *Neurergus*, *Ichthyosaura* or *Ommatotriton*. This result suggests introgression at the root of this group, rather than more recently between *Calotriton* and *Triturus*, which might seems contradictory to the branching in the mitochondrial tree. However it is possible that a mitochondrial capture occurred from the ancestral populations of *Triturus* to those of *Calotriton* (summarized in Fig. 6), without introgressed nuclear genes being retained (Toews & Brelsford 2012; Good et al. 2015).

Within the "Modern Asian Newts", only one network recovered a reticulation, linking *Paramesotriton* and *Cynops cyanurus*, with the two *Cynops* species forming a monophyletic group. When testing *C. cyanurus* against *Paramesotriton* and *Pachytriton*, no introgression could be detected, but introgression tests including *Pachytriton* are prone to stochastic error due to its low gene coverage. The only introgression test that could be performed without this taxon, using about 2000 gene trees, suggested introgression between *C. cyanurus* and *C. pyrrhogaster*. This result fits a scenario where the lineage represented by *C. pyrrhogaster* splits early within the "Modern Asian Newts" and latter hybridizes with the lineage represented by *C. cyanurus*, resulting in both mitochondrial capture and introgression of nuclear genes from *C. pyrrhogaster* into *C. cyanurus*

(Fig. 6). This hypothesis is also consistent with the mitochondrial topology. However, given the few tests that could be performed, the possibility that the *Cynops* paraphyly is the result of introgression of nuclear genes from the *Paramesotriton* clade into *C. cyanurus* cannot be completely ruled out. As advocated earlier, further studies with a special focus on this group of Asian newts and a more comprehensive taxon sampling (particularly including more species of *Cynops*) will be needed to clarify these relationships.

It is surprising that we did not find evidence for nuclear introgression directly involving *Ichthyosaura*, considering that this genus' placements in the mitochondrial and nuclear transcriptomic trees are highly different. However, as hypothesized for *Calotriton*, it is possible that *Ichthyosaura* captured a mitochondrial lineage similar to *Lissotriton* following an ancient hybridization event (both scenarios summarized in Fig. 6), without retaining any introgressed nuclear genes. *Ichthyosaura* also displays an extremely long branch on the mitochondrial tree which could result from an accelerated mutation rate subsequent to mitochondrial replacement, caused by the introgressed mitochondrial genes adapting to the new genomic environment of mitochondrial-related nuclear genes (Sloan et al. 2017). However, in the absence of evidence for introgression, it is also possible that *Ichthyosaura* displays a higher mitochondrial mutation rate for another, unrelated reason, and this causes difficulties to reconstruct its relationships based on mitochondrial sequences.

To conclude, we confirm that some aspects of the evolutionary history of the Salamandridae are rather difficult to solve, even with extensive genomic data sets. The deep nodes of the tree were uncontroversial in all our analyses, largely in agreement with previous studies. On the other hand, relationships within several groups are unclear, to say the least. Our results contribute evidence for introgression between lineages at different phylogenetic scales (summarized in Fig. 6), but we can only speculate about the exact number of reticulations and the branches along which they occurred. Further in-depth analyses of the contentious clades separately, using a more complete species-level taxon sampling, could help solving all the relationships within this family. Moreover, the potential confounding effect of introgression in deep phylogenetic inference (e.g. regarding the placement of *Salamandrina*) remains to be clarified, as the power of our approaches decreases at such evolutionary scales. Ultimately, studying signals of introgression based on whole genome sequences might be needed to fully understand the relationships among the Salamandridae, a goal still difficult to achieve at present due to the very large genome sizes, and high content of repetitive elements in salamandrid genomes.

4.3 Gene flow in phylogenomic studies

In line with a growing number of publications, our results point out the importance of considering introgression in phylogenetic studies. This is particularly relevant when analyzing large genomic data sets, which tend to give high support to spurious branches and hide signals of discordance among markers. Even the gene-jackknifing method, usually a reliable method to assess branch support, in our study gave strong support for relationships that most likely were influenced by gene flow, e.g. the *Lissotriton*/*Triturus* clade (Fig. 2d). On the other hand, large phylogenomic data sets also offer an unprecedented opportunity to integrate introgression into the phylogenetic paradigm. Indeed, sampling numerous loci at a genomic scale allows in-depth insights into the various evolutionary histories supported by one genome, especially using long loci sequences as in phylotranscriptomic approaches, which can yield well-resolved gene-trees.

The main remaining issue with integrating reticulations into phylogenetic studies is methodological, as the various methods available all have shortcomings. Inferring phylogenetic networks under the Multi-Species Network Coalescent might be the most promising approach, as it integrates explicit modeling of both ILS and introgression. However, such analyses are computationally intensive when including a large amount of taxa, and still have a lot of shortcomings, particularly regarding the identification of the optimal number of reticulations (Pardi et al. 2015; Solís-Lemus & Ané 2016; Wen et al. 2016). As a result, most analyses using phylogenetic networks focus on a small number of reticulations, often only one, at shallow evolutionary scales (e.g. Yu et al. 2012). Thus, the applicability of phylogenetic networks to extensive phylogenomic data sets remains limited, but further theoretical developments might improve both the computational efficiency and the accuracy of the existing algorithms. Different alternative approaches to detect introgression in phylogenetic studies have been used, including introgression tests (Green et al. 2013) or evolutionary scenarios tests based on simulations (Burbrink & Gehara 2018), but they all have limitations that might narrow their applicability to specific questions and data. However, even if these methods do not allow to fully characterize the hybridization events that occurred during the evolutionary history of a group of species, or in other words do not allow to infer the exact species network, we advocate that they should not be ignored as they can still yield very valuable insights on the presence of reticulations among the branches of a tree. Several very simple approaches, such as quartet sampling (Pease et al. 2018) or the comparison of mitochondrial and nuclear markers, can be used to define putative introgression events. Depending of the data set size, and the apparent complexity of the evolutionary scenario, one of the methods described above can then be used to get a more accurate overview of the evolutionary history. Additionally, information regarding introgression can be very useful for evolutionary studies that rely on backbone trees, to determine whether some nodes should be better considered as unresolved.

Acknowledgments

We are grateful to Uwe Seidel for help with obtaining samples and to Paul Hime for generating data. LR and MVen were supported by the Deutsche Forschungsgemeinschaft (grant VE247/16-1 – HO 3492/6-1) in the framework of the "TaxonOmics" priority program. II was supported by a Juan de la Cierva–Incorporación postdoctoral fellowship from the Spanish Ministry of Economy and Competitiviness (MINECO). ES was supported by the Deutscher Akademischer Austauschdienst (DAAD).

Competing interest statement

The authors declare that there is no competing interests.

References

AmphibiaWeb, 2019. Information on amphibian biology and conservation. http://amphibiaweb.org/ (accessed 11 January 2019).

Arnold, S.J., 1977. The evolution of courtship behavior in new world salamanders with some comments on Old World salamandrids. In: Taylor, D.H., Guttman, S.I. (Eds.), The Reproductive Biology of Amphibians. Springer, US, Boston, Massachusetts, pp. 141–183.

Arnold, M. L., 1997. *Natural hybridization and evolution*. Oxford University Press on Demand.

Arntzen, J. W., Wielstra, B., Wallis, G. P., 2014. The modality of nine *Triturus* newt hybrid zones assessed with nuclear, mitochondrial and morphological data. Biol. J. Linn. Soc. 113, 604–622. https://doi.org/10.1111/bij.12358

Baack, E. J., & Rieseberg, L. H., 2007. A genomic view of introgression and hybrid speciation. Curr. Opin. Genet. Dev. 17, 513–518. https://doi.org/10.1016/j.gde.2007.09.001

Babik, W., Szymura, J.M., Rafiński, J., 2003. Nuclear markers, mitochondrial DNA and male secondary sexual traits variation in a newt hybrid zone (*Triturus vulgaris*×*T. montandoni*). Mol. Ecol. 12, 1913–1930. https://doi.org/10.1046/j.1365-294X.2003.01880.x

Barlow, A., Cahill, J. A., Hartmann, S., Theunert, C., Xenikoudakis, G., Fortes, G. G., Paijmans, J.L.A., Rabeder, G., Frischauf, C., Grandal-d'Anglade, A., García-Vázquez, A., Murtskhvaladze, M., Saarma, U., Anijalg, P., Skrbinšek, T., Bertorelle, G., Gasparian, B., Bar-Oz, G., Pinhasi, R., Slatkin, M., Dalén, L., Shapiro, B., Hofreiter, M., 2018. Partial genomic survival of cave bears in living brown bears. Nature Ecol. Evol. 2, 1563. https://doi.org/10.1038/s41559-018-0654-8

Bolger, A.M., Lohse, M., Usadel B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–21120. https://doi.org/10.1093/bioinformatics/btu170

Buckley, D., 2012. Evolution of viviparity in salamanders (Amphibia, Caudata). Encyclopedia of Life Sciences. John Wiley & Sons, Ltd, Chichester, UK.

Burbrink, F. T., Gehara, M., 2018. The biogeography of deep time phylogenetic reticulation. Syst. Biol. 67, 743–755. https://doi.org/10.1093/sysbio/syy019

Carranza, S., Amat, F., 2005. Taxonomy, biogeography and evolution of *Euproctus* (Amphibia: Salamandridae), with the resurrection of the genus *Calotriton* and the description of a new endemic species from the Iberian Peninsula. Zool. J. Linn. Soc. 145, 555–582. https://doi.org/10.1111/j.1096-3642.2005.00197.x

Chernomor, O., von Haeseler, A., Minh, B.Q., 2016. Terrace aware data structure for phylogenomic inference from supermatrices. Syst. Biol. 65, 997–1008. https://doi.org/10.1093/sysbio/syw037

Degnan, J.H., Rosenberg, N.A., 2006. Discordance of species trees with their most likely gene trees. PLoS Genetics 2, e68. 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. 24, 332–340. https://doi.org/10.1016/j.tree.2009.01.009

Degnan, J.H., 2018. Modeling hybridization under the network multispecies coalescent. Syst. Biol. 67, 786–799. https://doi.org/10.1093/sysbio/syy040

Delsuc, F., Tsagkogeorga, G., Lartillot, N., Philippe, H., 2008. Additional molecular support for the new chordate phylogeny. Genesis 46, 592–604. https://doi.org/10.1002/dvg.20450

Dubois, A., Raffaëlli, J., 2009. A new ergotaxonomy of the family Salamandridae Goldfuss, 1820 (Amphibia, Urodela). Alytes 26, 1–85.

Durand, E.Y., Patterson, N., Reich, D., Slatkin, M., 2011. Testing for ancient admixture between closely related populations. Mol. Biol. Evol. 28, 2239–2252. https://doi.org/10.1093/molbev/msr048

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

Escoriza, D., Gutiérrez-Rodríguez, J., Hassine, J.B., Martínez-Solano, I., 2016. Genetic assessment of the threatened microendemic *Pleurodeles poireti* (Caudata, Salamandridae), with molecular evidence for hybridization with *Pleurodeles nebulosus*. Cons. Genet. 17, 1445–1458. https://doi.org/10.1007/s10592-016-0875-9

Evans, B.J., 2008. Genome evolution and speciation genetics of clawed frogs (*Xenopus* and *Silurana*). Front. Biosci. 13, 4687–4706. http://dx.doi.org/10.2741/3033

Ferrier, V., Beetschen, J.C., 1973. Etude des chromosomes de *Tylototriton verrucosus* Anderson et de l'hybride viable *Pleurodeles waltlii* × *Tylototriton verrucosus* (Amphibiens Urodèles, Salamandridae). Chromosoma 42, 57–69. https://doi.org/10.1007/BF00326330

Folk, R. A., Soltis, P. S., Soltis, D. E., Guralnick, R., 2018. New prospects in the detection and comparative analysis of hybridization in the tree of life. Amer. J. Bot. 105, 364–375. https://doi.org/10.1002/ajb2.1018

Geffeney, S., Brodie, E. D., Ruben, P. C., 2002. Mechanisms of adaptation in a predator-prey arms race: TTX-resistant sodium channels. Science 297, 1336-1339. https://doi.org/10.1126/science.1074310

Good, J. M., Vanderpool, D., Keeble, S., & Bi, K., 2015. Negligible nuclear introgression despite complete mitochondrial capture between two species of chipmunks. Evolution 69, 1961–1972. https://doi.org/10.1111/evo.12712

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Palma, F.D., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat. Biotech. 29, 644–652. https://dx.doi.org/10.1038%2Fnbt.1883

Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Li, H., Zhai, W., Fritz, M.H.Y., Hansen, N.F., Durand, E.Y., Malaspinas, A.S., Jensen, J.D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H.A., Good, J.M., Schultz, R., Aximu-Petri, A., Butthof, A., Höber, B., Höffner, B., Siegemund, M., Weihmann, A., Nusbaum, C., Lander, E.S., Russ, C., Novod, N., Affourtit, J., Egholm, M., Verna, C., Rudan, P., Brajkovic, D., Kucan, Ž., Gušic, I., Doronichev, V.B., Golovanova, L.V., Lalueza-Fox, C., de la Rasilla, M., Fortea, J., Rosas, A.,

Schmitz, R.W., Johnson, P.L.F., Eichler, E.E., Falush, D., Birney, E., Mullikin, J.C., Slatkin, M., Nielsen, R., Kelso, J., Lachmann, M., Reich, D., Pääbo, S., 2010. A draft sequence of the Neandertal genome. Science 328, 710–722. https://dx.doi.org/10.1126/science.1188021

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., MacManes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., LeDuc, R. D., Friedman, N., Regev, A., 2013. De novo transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. Nat. Protoc. 8, 1494–1512. https://doi.org/10.1038/nprot.2013.084

Hanifin, C.T., Gilly, W.F., 2014. Evolutionary history of a complex adaptation: tetrodotoxin resistance in salamanders. Evolution 69, 232–244. https://doi.org/10.1111/evo.12552

Harrison, R.G., 1989. Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. Trends Ecol. Evol. 4, 6–11. https://doi.org/10.1016/0169-5347(89)90006-2

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., Baurain, D., Brinkmann, H., Delsuc, F., Sire, J.Y., Kupfer, A., Petersen, J., Jarek, M., Meyer, A., Vences, M., Philippe, H., 2017. Phylotranscriptomic consolidation of the jawed vertebrate timetree. Nat. Ecol. Evol. 1, 1370–1378. https://doi.org/10.1038/s41559-017-0240-5

Irisarri, I., Singh, P., Koblmüller, S., Torres-Dowdall, J., Henning, F., Franchini, P., Fischer, C., Lemmon, A.R., Lemmon, E.M., Thallinger, G.G., Sturmbauer, C., Meyer, A., 2018. Phylogenomics uncovers early hybridization and adaptive loci shaping the radiation of Lake Tanganyika cichlid fishes. Nat. Commun. 9, 3159. https://doi.org/10.1038/s41467-018-05479-9

Johannesen, J., Johannesen, B., Griebeler, E.M., Baran, I., Tunc, M.R., Kiefer, A., Veith, M., 2006. Distortion of symmetrical introgression in a hybrid zone: evidence for locus-specific selection and uni-directional range expansion. J. Evol. Biol. 19, 705–716. https://doi.org/10.1111/j.1420-9101.2005.01064.x

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: Fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589. https://doi.org/10.1038/nmeth.4285

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010

Kieren, S., Sparreboom, M., Hochkirch, A., Veith, M., 2018. A biogeographic and ecological perspective to the evolution of reproductive behaviour in the family Salamandridae. Mol. Phylogenet. Evol. 121, 98–109. https://doi.org/10.1016/j.ympev.2018.01.006

Kopylova, E., Noé, L., Touzet, H., 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28, 3211–3217. https://doi.org/10.1093/bioinformatics/bts611

Lartillot, N., Philippe, H., 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Mol. Biol. Evol. 21, 1095–1109. https://doi.org/10.1093/molbev/msh11

Lartillot, N., Philippe, H., 2006. Computing Bayes factors using thermodynamic integration. Syst. Biol. 55, 195–207. https://doi.org/10.1080/10635150500433722

Lartillot, N., Brinkmann, H., Philippe, H., 2007. Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. BMC Evol. Biol. 7, S4. https://doi.org/10.1186/1471-2148-7-S1-S4

Leaché, A. D., Harris, R. B., Rannala, B., Yang, Z. 2013. The influence of gene flow on species tree estimation: a simulation study. Syst. Biol. 63, 17–30. https://doi.org/10.1093/sysbio/syt049

Lourenço, A., Gonçalves, J., Carvalho, F., Wang, I. J., Velo-Antón, G., 2019. Comparative landscape genetics reveals the evolution of viviparity reduces genetic connectivity in fire salamanders. Mol. Ecol. 28, 4573-4591. https://doi.org/10.1111/mec.15249

Lüddecke, T., Schulz, S., Steinfartz, S., Vences, M., 2018. A salamander's toxic arsenal: review of skin poison diversity and function in true salamanders, genus *Salamandra*. Sci. Nat. 105, 56. https://doi.org/10.1007/s00114-018-1579-4

Maex, M., Treer, D., De Greve, H., Proost, P., Van Bocxlaer, I., Bossuyt, F., 2018. Exaptation as a mechanism for functional reinforcement of an animal pheromone system. Curr. Biol. 28, 2955–2960. doi: 10.1016/j.cub.2018.06.074.

Macgregor, H. C., Sessions, S. K., Arntzen, J. W., 1990. An integrative analysis of phylogenetic relationships among newts of the genus *Triturus* (family Salamandridae), using comparative biochemistry, cytogenetics and reproductive interactions. J. Evol. Biol. 3, 329–373. https://doi.org/10.1046/j.1420-9101.1990.3050329.x

Maletzky, A., Mikulíček, P., Franzen, M., Goldschmid, A., Gruber, H.J., Horák, A., Kyek, M., 2008. Hybridization and introgression between two species of crested newts (*Triturus cristatus* and *T.*

carnifex) along contact zones in Germany and Austria: morphological and molecular data. Herpetol. J. 18, 1–15.

Mallet, J., 2005. Hybridization as an invasion of the genome. Trends Ecol. Evol. 0, 229-237. https://doi.org/10.1016/j.tree.2005.02.010

Mallet, J., Beltrán, M., Neukirchen, W., Linares, M., 2007. Natural hybridization in heliconiine butterflies: the species boundary as a continuum. BMC Evol. Biol. 7, 28. https://doi.org/10.1186/1471-2148-7-28

Mallet, J., Besansky, N., Hahn, M. W., 2016. How reticulated are species? BioEssays 38, 140–149. https://doi.org/10.1002/bies.201500149

Mancino, G., Ragghianti, M., Bucci-Innocenti, S., 1978. Experimental hybridization within the genus *Triturus* (Urodela: Salamandridae). I. Spermatogenesis of F1 species hybrids, *Triturus cristatus carnifex* female x *T. vulgaris meridionalis* male. Chromosoma 69, 27–46. https://doi.org/10.1080/00087114.1979.10796777

McKay, B.D., Zink, R.M., 2010. The causes of mitochondrial DNA gene tree paraphyly in birds. Mol. Phylogenet. Evol. 54, 647–650. https://doi.org/10.1016/j.ympev.2009.08.024

Mirarab, S., Warnow, T., 2015. ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics 31, i44–i52. https://doi.org/10.1093/bioinformatics/btv234

Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. Mol. Biol. Evol. 32, 268–274. https://doi.org/10.1093/molbev/msu300

Palkopoulou, E, Lipson, M, Mallick, S, Nielsen, S, Rohland, N, Baleka, S, Karpinski, E, Ivancevic, AM, To, TH, Kortschak, RD, Raison, JM, Qu, Z, Chin, TJ, Alt, KW, Claesson, S, Dalén, L, MacPhee, RDE, Meller, H, Roca, AL, Ryder, OA, Heiman, D, Young, S, Breen, M, Williams,C, Aken, BL, Ruffier, M, Karlsson, E, Johnson, J, Di Palma, F, Alfoldi, J, Adelson, DL, Mailund, T, Munch, K, Lindblad-Toh, K, Hofreiter, M, Poinar, H, Reich, D., 2018. A comprehensive genomic history of extinct and living elephants. Proc. Natl. Acad. Sci. U.S.A. 115, E2566–E2574.

Paradis, E., Claude, J., Strimmer, K., 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20, 289–290. https://doi.org/10.1093/bioinformatics/btg412

Pardi, F., Scornavacca, C., 2015. Reconstructible phylogenetic networks: do not distinguish the indistinguishable. PLoS Comput. Biol. 11, e1004135. https://doi.org/10.1371/journal.pcbi.1004135

Pariser, K., 1932. Verschiebung des Geschlechtsverhältnisses bei künstlich erzeugten Tritonbastarden. Biol. Zbl. 52, 654–659.

Payseur, B. A., Rieseberg, L. H., 2016. A genomic perspective on hybridization and speciation. Mol. Ecol. 25, 2337–2360. https://doi.org/10.1111/mec.13557

Pease, J.B., Brown, J.W., Walker, J.F., Hinchliff, C.E., Smith, S.A., 2018. Quartet Sampling distinguishes lack of support from conflicting support in the green plant tree of life. Amer. J. Bot. 105, 385–403. https://doi.org/10.1002/ajb2.1016

Roure, B., Rodriguez-Ezpeleta, N., & Philippe, H., 2007. SCaFoS: a tool for selection, concatenation and fusion of sequences for phylogenomics. BMC Evol. Biol. 7, S2. https://doi.org/10.1186/1471-2148-7-S1-S2

Sanchiz, B. 1977. Catálogo de los anfibios fósiles de España. Acta Geol. Hispánica 12, 103–107.

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

Simion, P., Philippe, H., Baurain, D., Jager, M., Richter, D.J., Di Franco, A., Roure, B., Satoh, N., Quéinnec, É., Ereskovsky, A., Lapébie, P., Corre, E., Delsuc, F., King, N., Wörheide, G., Manuel, M., 2017. A large and consistent phylogenomic dataset supports sponges as the sister group to all other animals. Curr. Biol. 27, 958–967.

Slatkin, M., Pollack, J. L., 2008. Subdivision in an ancestral species creates asymmetry in gene trees. Mol. Biol. Evol. 25, 2241–2246. https://doi.org/10.1093/molbev/msn172

Sloan, D. B., Havird, J. C., Sharbrough, J., 2017. The on-again, off-again relationship between mitochondrial genomes and species boundaries. Mol. Ecol. 26, 2212–2236. https://doi.org/10.1111/mec.13959

Solís-Lemus, C., Ané, C., 2016. Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. PLoS Genet. 12, e1005896. doi:10.1371/journal.pgen.1005896

Soltis, P. S., Soltis, D. E., 2009. The role of hybridization in plant speciation. Annu. Ev. Plant Biol. 60, 561–588. https://doi.org/10.1146/annurev.arplant.043008.092039

Sparreboom, M., 2014. Salamanders of the Old World: The Salamanders of Europe, Asia and Northern Africa. KNNV Publishing, Zeist.

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., 2006. A Bayesian approach on molecules and behavior: reconsidering phylogenetic and evolutionary patterns of the Salamandridae with emphasis on *Triturus* newts. J. Exper. Zool. B: Mol. Devel. Evol. 308, 139–162. https://doi.org/10.1002/jez.b.21119

Stuglik, M.T., Babik, W., 2016. Genomic heterogeneity of historical gene flow between two species of newts inferred from transcriptome data. Ecol. Evol. 6, 4513–4525.

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

Theunert, C., Slatkin, M., 2017. Distinguishing recent admixture from ancestral population structure. Genome Biol. Evol. 9: 427–437. https://doi.org/10.1093/gbe/evx018

Titus, T.A., Larson, A., 1995. A molecular phylogenetic perspective on the evolutionary radiation of the salamander family Salamandridae. Syst. Biol. 44, 125–151. https://doi.org/10.1093/sysbio/44.2.125

Toews, D. P., Brelsford, A., 2012. The biogeography of mitochondrial and nuclear discordance in animals. Mol. Ecol. 21, 3907–3930. https://doi.org/10.1111/j.1365-294X.2012.05664.x

Tominaga, A., Matsui, M., Yoshikawa, N., Nishikawa, K., Hayashi, T., Misawa, Y., Tanabo, S., Ota, H., 2013. Phylogeny and historical demography of *Cynops pyrrhogaster* (Amphibia: Urodela): taxonomic relationships and distributional changes associated with climatic oscillations. Mol. Phylogenet. Evol. 66, 654–667. https://doi.org/10.1016/j.ympev.2012.10.015

Van Bocxlaer, I., Maex, M., Treer, D., Janssenswillen, S., Janssens, R., Vandebergh, W., Proust, P., Bossuyt, F., 2016. Beyond sodefrin: evidence for a multi-component pheromone system in the model newt *Cynops pyrrhogaster* (Salamandridae). Sci. Rep. 6, 21880. https://doi.org/10.1038/srep21880

Veith, M., Steinfartz, S., Zardoya, R., Seitz, A., Meyer, A., 1998. A molecular phylogeny of ‘true’ salamanders (family Salamandridae) and the evolution of terrestriality of reproductive modes. J. Zool. Syst. Evol. Res. 36, 7–16. https://doi.org/10.1111/j.1439-0469.1998.tb00774.x

Veith, M., Bogaerts, S., Pasmans, F., Kieren, S., 2018. The changing views on the evolutionary relationships of extant Salamandridae (Amphibia: Urodela). PLoS ONE. 13, e0198237. https://doi.org/10.1371/journal.pone.0198237

Wallis, G.P., Cameron-Christie, S.R., Kennedy, H.L., Palmer, G., Sanders, T.R., Winter, D.J., 2017. Interspecific hybridization causes long-term phylogenetic discordance between nuclear and mitochondrial genomes in freshwater fishes. Mol. Ecol. 26: 3116–3127. https://doi.org/10.1111/mec.14096

Węcek, K., Hartmann, S., Paijmans, J. L., Taron, U., Xenikoudakis, G., Cahill, J. A., Heintzman, P.D., Shapiro, B., Baryshnikov, G., Bunevich, A.N., Crees, J. J., Dobosz, R., Manaserian, N., Okarma, H., Tokarska, M., Turvey, S.T., Wójcik, J.M., Żyła, W., Szymura, J.M., Hofreiter, M., Barlow, A., 2016. Complex admixture preceded and followed the extinction of wisent in the wild. Mol. Biol. Evol. 34, 598–612. https://doi.org/10.1093/molbev/msw254

Wen, D., Yu, Y., Nakhleh, L., 2016. Bayesian inference of reticulate phylogenies under the multispecies network coalescent. PLoS Gene. 12, e1006006. https://doi.org/10.1371/journal.pgen.1006006

Wen, D., Yu, Y., Zhu, J., Nakhleh, L., 2018. Inferring phylogenetic networks using PhyloNet. Syst. Biol. 67, 735–740. https://doi.org/10.1093/sysbio/syy015

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. doi: 10.1016/j.ympev.2018.12.032. Epub 2019 Jan 7.

Whitney, K. D., Ahern, J. R., Campbell, L. G., Albert, L. P., King, M. S., 2010. Patterns of hybridization in plants. Persp. Plant Ecol. Evol. Syst. 12, 175–182. https://doi.org/10.1016/j.ppees.2010.02.002

Weisrock, D.W., Papenfuss, T.J., Macey, J.R., Litvinchuk, S.N., Polymeni, R., Ugurtas, I.H., Zhao, E., Jowkar, H., Larson, A., 2006. A molecular assessment of phylogenetic relationships and lineage accumulation rates within the family Salamandridae (Amphibia, Caudata). Mol. Phylogenet. Evol. 41, 368–383. https://doi.org/10.1016/j.ympev.2006.05.008

Wiens, J.J., Sparreboom, M., Arntzen, J.W., 2011. Crest evolution in newts: implications for reconstruction methods, sexual selection, phenotypic plasticity and the origin of novelties. J. Evol. Biol. 24, 2073–2086. https://doi.org/10.1111/j.1420-9101.2011.02340.x

Yu, Y., Degnan, J. H., Nakhleh, L., 2012. The probability of a gene tree topology within a phylogenetic network with applications to hybridization detection. PLoS Genet. 8, e1002660. https://doi.org/10.1371/journal.pgen.1002660

Yu, Y., Dong, J., Liu, K.J., Nakhleh, L., 2014. Maximum likelihood inference of reticulate evolutionary histories. Proc. Natl. Acad. Sci. U.S.A. 111, 16448–16453. https://doi.org/10.1073/pnas.1407950111

Yu, Y., Nakhleh, L., 2015. A maximum pseudo-likelihood approach for phylogenetic networks. BMC Genomics 16, S10. https://doi.org/10.1186/1471-2164-16-S10-S10

Zhang, P., Papenfuss, T.J., Wake, M.H., Qu, L., Wake, D.B., 2008. Phylogeny and biogeography of the family Salamandridae (Amphibia: Caudata) inferred from complete mitochondrial genomes. Mol. Phylogenet. Evol. 49, 586–597. https://doi.org/10.1016/j.ympev.2008.08.020

Table 1. Details of the 10 best networks inferred with PhyloNet, ranked according to their Akaike Information Criterion (AIC). The last column indicates which figure illustrates the relationships corresponding to the respective network.

Maximum number of reticulations	Inferred number of reticulations	Ln(L)	AIC	Figure
7	5	-143047.0	286170.0	3b
10	4	-143051.9	286171.7	S5a
10	5	-143058.9	286193.8	S5b
7	3	-143072.1	286204.3	S5c
10	5	-143060.3	286204.5	S5d
8	4	-143080.0	286228.0	S5e
8	5	-143076.5	286229.0	3c /S5f
7	6	-143074.2	286232.4	S5g
8	5	-143082.6	286235.1	S5h
8	3	-143087.8	286235.7	S5i

Figure 1. (a) Phylogenetic tree inferred from protein (amino acid) sequences of 13 mitochondrial genes of 28 Salamandrids representatives under a CAT-GTR model. The numbers at the nodes are Bayesian Posterior Probabilities (black dots represent BPP=1.0). (b) Maximum likelihood tree of phylotranscriptomic relationships among the Salamandridae, inferred with IQ-TREE, using a by-genes partitioned analysis with best-fitting models and partitions on DNA sequences of 5,455 genes. Numbers at the nodes represent the approximate likelihood ratio test support and the gene jackknife proportion for 500 Kbp sampled, respectively in percent (shown only when <100%; all other nodes received 100% support in both analyses, represented by a black dot). Clade names follow Weisrock et al. (2006), Steinfartz et al. (2006), and Zhang et al. (2008). The red branches denote the taxa whose position differ between the two topologies; the blue branches show the new position of *Salamandrina*. Both trees were rooted with sequences of *Ambystoma laterale* and either *Necturus beyurus* (mitochondrial tree) or *Andrias davidianus* (nuclear tree) included as hierarchical outgroup (both removed from the graphs to improve graphical resolution within Salamandridae).

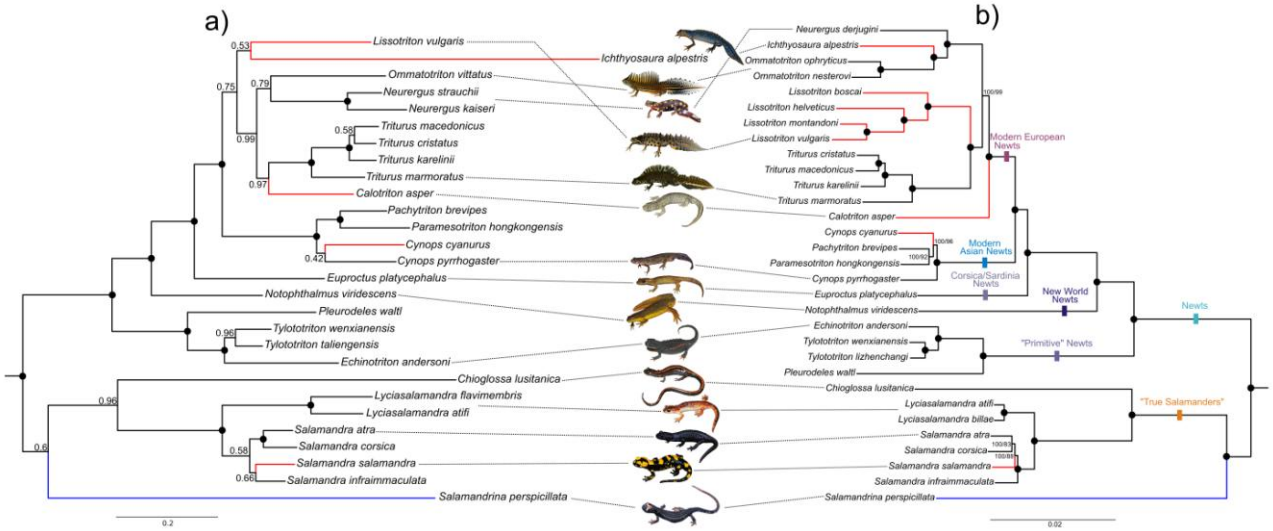
Figure 2. Proportion (%) of gene jackknifing replicates supporting selected nodes of the maximum-likelihood phylotranscriptomic tree (Fig. 1b), as a function of the number of nucleotides sampled: (a) average support for all the nodes of the tree and (b-f) supports for specific nodes discussed in the text.

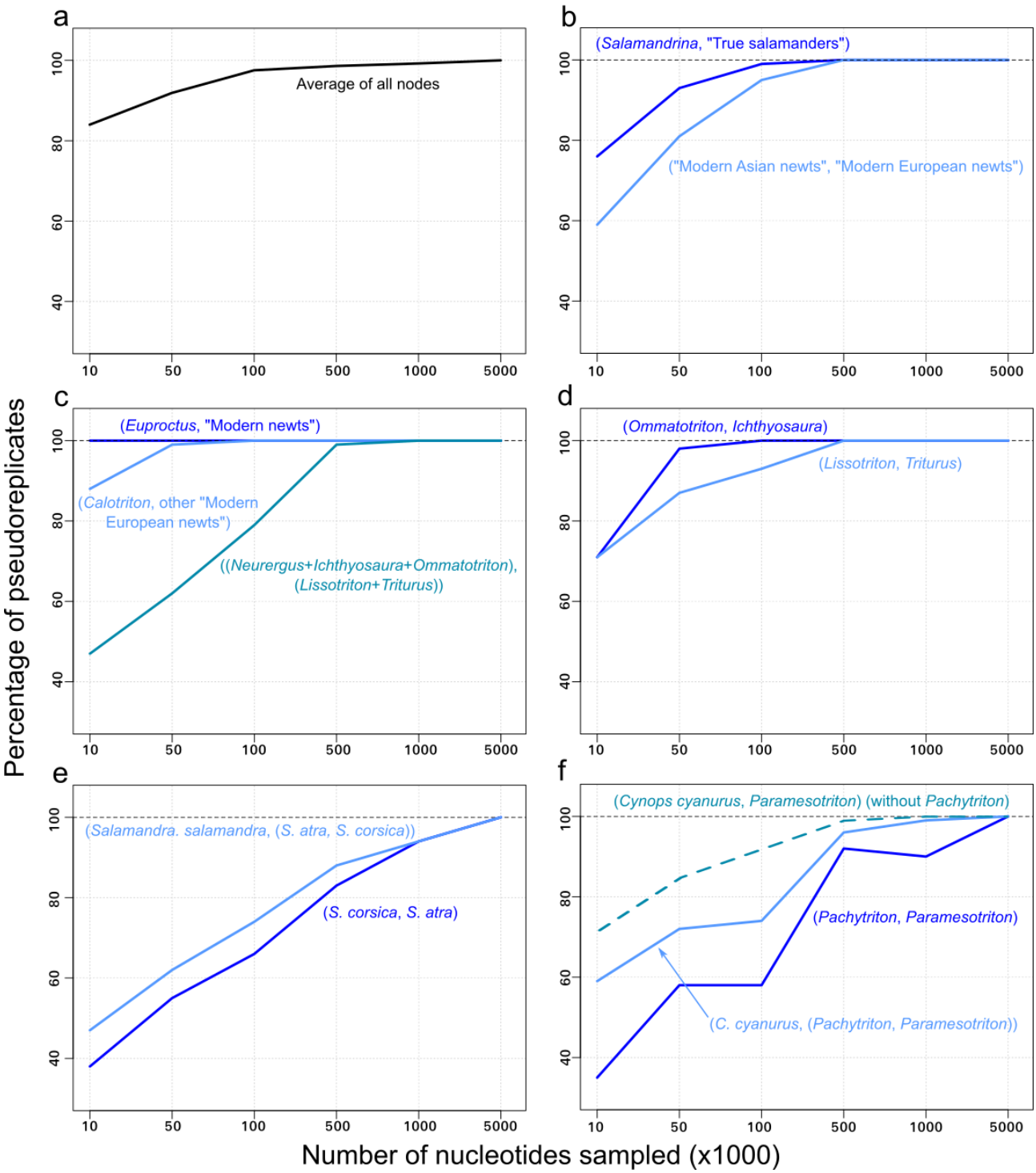
Figure 3. (a) Akaike Information Criterion (AIC) values of the 55 inferred phylogenetic networks for the “modern newt” clades, in function of the number of inferred reticulations. The line shows the average AIC for each number of reticulations, and the red crosses the 10 best models, considered for further investigations. (b) The best network according to AIC. Blue dashed lines show hybridization branches (light blue = inheritance probability < 0.20). The numbers adjacent to the hybridization branches give the inheritance probabilities from both parental lineages, as given by PhyloNet. (c) One of the sub-optimal networks (ranked 7th according to AIC, cf. Table 1) showing alternative positions for *Calotriton* and *Lissotriton*.

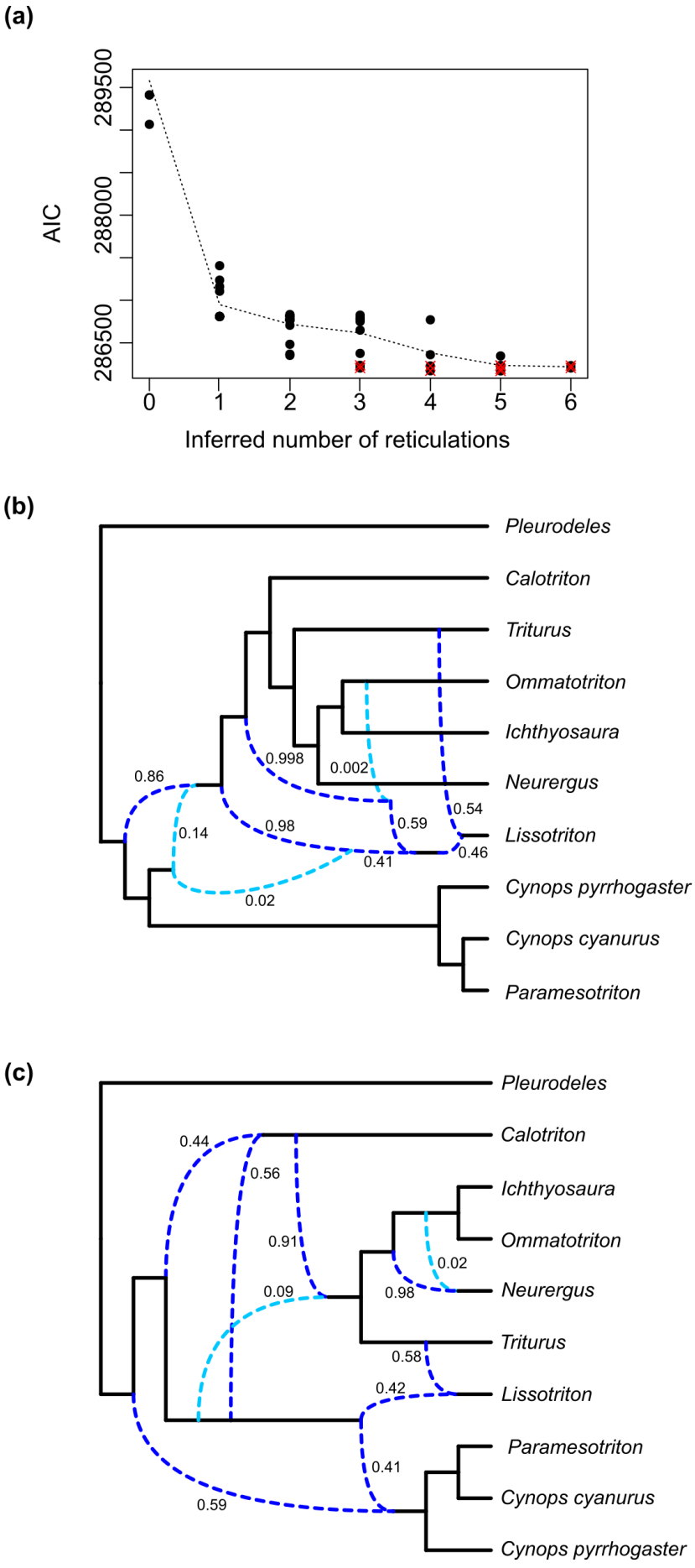
Figure 4. Summary of the significant introgression tests within the “Modern European Newts”. Each tree in the left part show the subsampling used, and the barplots represent the counts of the three alternative patterns within the gene trees. The trees on the right are schematic representations of the gene flow events inferred from the different tests (their position along the branch and width are arbitrary). **: 0.05 > p > 0.01 ; ***: p < 0.01.

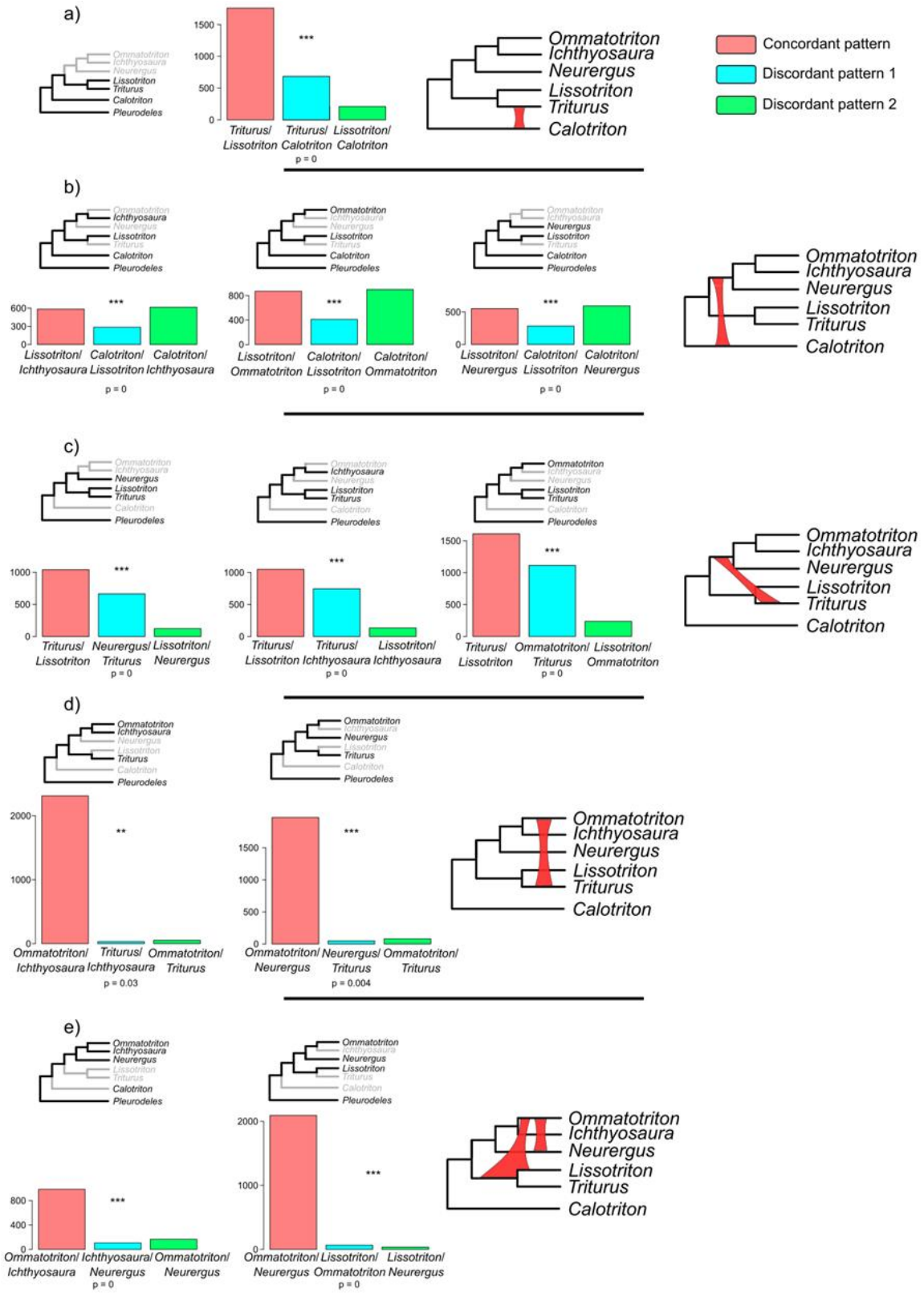
Figure 5. Summary of the significant introgression tests within the “Modern Asian Newts”. Each tree in the left part show the subsampling used, and the barplots represent the counts of the three alternative patterns within the gene trees. The trees on the right are schematic representations of the gene flow events inferred from the different tests (their position along the branch and width are arbitrary). ** : 0.05 > p > 0.01 ; *** : p < 0.01.

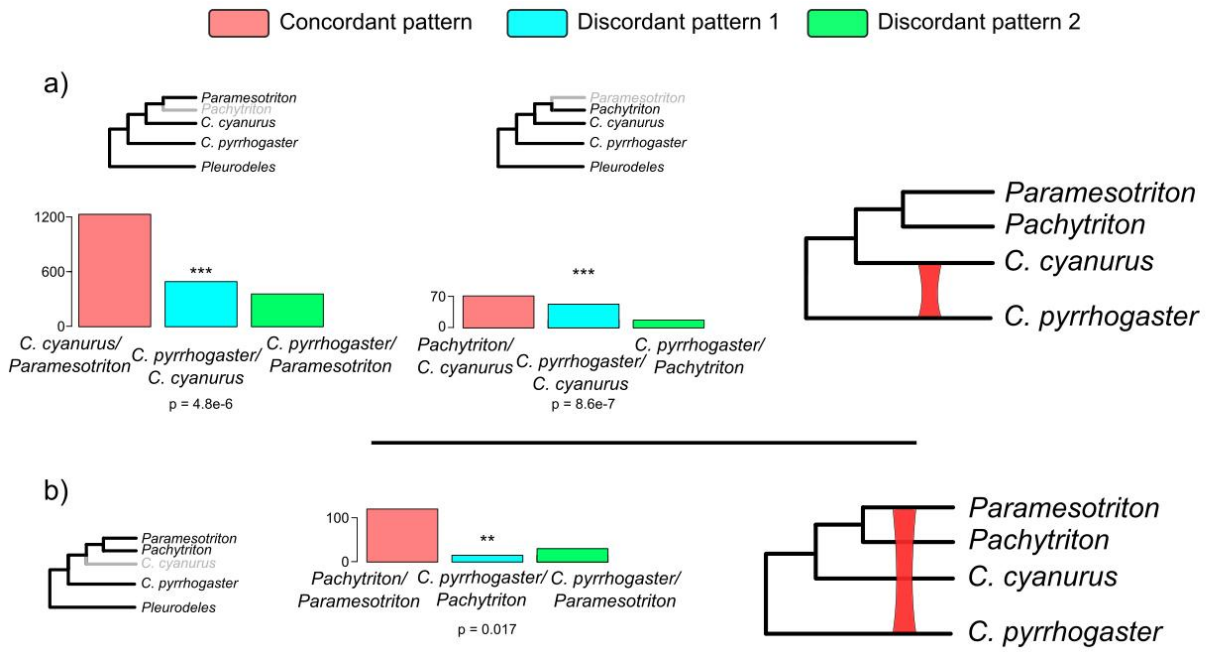
Figure 6. Schematic representations of two hypothetical evolutionary scenarios for the “modern newts”, involving several introgression events. The trees represent the splits between lineages, while the arrows represent directional introgressions (from donor to recipient lineage) of either mitochondrial genomes (red), nuclear genes (blue) or both (red with blue stripes). The position and direction of introgressions were estimated based on mito-nuclear discordances, phylogenetic networks and introgression tests. The relative order of the introgression events is putative, and the branches lengths are arbitrary.

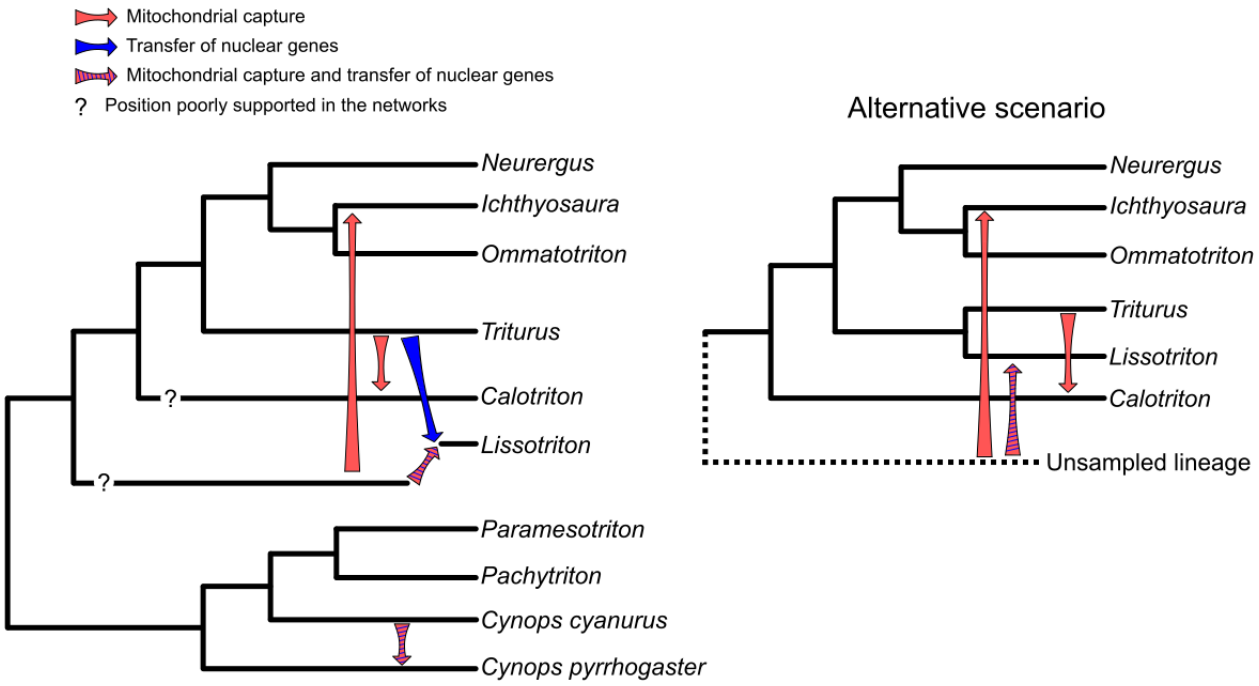



a
Average of all nodes
b
(Salamandrina, "True salamanders")
("Modern Asian newts", "Modern European newts")
c
(Euproctus, "Modern newts")
(Calotriton, other "Modern European newts")
((Neurergus+Ichthyosaura+Ommatotriton), (Lissotriton+Triturus))
d
(Ommatotriton, Ichthyosaura)
(Lissotriton, Triturus)
e
(Salamandra. salamandra, (S. atra, S. corsica))
(S. corsica, S. atra)
f
(Cynops cyanurus, Paramesotriton) (without Pachytriton)
(Pachytriton, Paramesotriton)
(C. cyanurus, (Pachytriton, Paramesotriton))
Percentage of pseudoreplicates
Number of nucleotides sampled (x1000)
10
50
100
500
1000
5000
40
60
80
100


(a)
AIC
286500
288000
289500
0
1
2
3
4
5
6
Inferred number of reticulations
(b)
Pleurodeles
Calotriton
Triturus
Ommatotriton
Ichthyosaura
Neurergus
Lissotriton
Cynops pyrrhogaster
Cynops cyanurus
Paramesotriton
0.86
0.14
0.998
0.002
0.98
0.59
0.54
0.46
0.41
0.02
(c)
Pleurodeles
Calotriton
Ichthyosaura
Ommatotriton
Neurergus
Triturus
Lissotriton
Paramesotriton
Cynops cyanurus
Cynops pyrrhogaster
0.44
0.56
0.91
0.09
0.02
0.98
0.58
0.42
0.41
0.59


a)
Ommatotriton
Ichthyosaura
Neurergus
Lissotriton
Triturus
Calotriton
Pleurodeles
1500
1000
500
0
***
Triturus/ Lissotriton
Triturus/ Calotriton
Lissotriton/ Calotriton
p = 0
Ommatotriton
Ichthyosaura
Neurergus
Lissotriton
Triturus
Calotriton
Concordant pattern
Discordant pattern 1
Discordant pattern 2
b)
600
300
0
Lissotriton/ Ichthyosaura
Calotriton/ Lissotriton
Calotriton/ Ichthyosaura
800
400
Lissotriton/ Ommatotriton
Calotriton/ Lissotriton
Calotriton/ Ommatotriton
500
Lissotriton/ Neurergus
Calotriton/ Lissotriton
Calotriton/ Neurergus
c)
1000
500
Triturus/ Lissotriton
Neurergus/ Triturus
Lissotriton/ Neurergus
Triturus/ Lissotriton
Triturus/ Ichthyosaura
Lissotriton/ Ichthyosaura
1500
Triturus/ Lissotriton
Ommatotriton/ Triturus
Lissotriton/ Ommatotriton
d)
2000
1000
**
Ommatotriton/ Ichthyosaura
Triturus/ Ichthyosaura
Ommatotriton/ Triturus
p = 0.03
1500
Ommatotriton/ Neurergus
Neurergus/ Triturus
Ommatotriton/ Triturus
p = 0.004
e)
800
400
Ommatotriton/ Ichthyosaura
Ichthyosaura/ Neurergus
Ommatotriton/ Neurergus
Ommatotriton/ Neurergus
Lissotriton/ Ommatotriton
Lissotriton/ Neurergus


Concordant pattern
Discordant pattern 1
Discordant pattern 2
a)
Paramesotriton
Pachytriton
C. cyanurus
C. pyrrhogaster
Pleurodeles
1200
600
0
***
C. cyanurus/ Paramesotriton
C. pyrrhogaster/ C. cyanurus
C. pyrrhogaster/ Paramesotriton
p = 4.8e-6
70
0
***
Pachytriton/ C. cyanurus
C. pyrrhogaster/ C. cyanurus
C. pyrrhogaster/ Pachytriton
p = 8.6e-7
b)
100
0
**
Pachytriton/ Paramesotriton
C. pyrrhogaster/ Pachytriton
C. pyrrhogaster/ Paramesotriton
p = 0.017


Mitochondrial capture
Transfer of nuclear genes
Mitochondrial capture and transfer of nuclear genes
? Position poorly supported in the networks
Neurergus
Ichthyosaura
Ommatotriton
Triturus
Calotriton
Lissotriton
Paramesotriton
Pachytriton
Cynops cyanurus
Cynops pyrrhogaster
?
?
Alternative scenario
Neurergus
Ichthyosaura
Ommatotriton
Triturus
Lissotriton
Calotriton
Unsampled lineage