Molecular Phylogenetics and Evolution 139 (2019) 106551

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Resolving evolutionary relationships among six closely related taxa of the horseshoe bats (*Rhinolophus*) with targeted resequencing data

Xiuguang Mao[a,*], Georgia Tsagkogeorga[b], Vu Dinh Thong[c,d], Stephen J. Rossiter[b,*]

[a] Institute of Estuarine and Coastal Research, East China Normal University, Shanghai 200062, China
[b] School of Biological and Chemical Sciences, Queen Mary University of London, London E1 4NS, UK
[c] Institute of Ecology and Biological Resources, Vietnam Academy of Science and Technology (VAST), 18 Hoang Quoc Viet Road, Cau Giay District, Hanoi, Viet Nam
[d] Graduate University of Science and Technology, VAST, 18 Hoang Quoc Viet Road, Cau Giay District, Hanoi, Viet Nam

ARTICLE INFO

Keywords:
Sequence capture
Introgression
Hybridization
Mito-nuclear discordance
Rhinolophus

ABSTRACT

Recently diverged taxa are often characterised by high rates of introgressive hybridization and incomplete lineage sorting, both of which can complicate phylogenetic reconstructions of species histories. Here we use a sequence capture approach to obtain genome-wide data to resolve the evolutionary relationships, and infer the extent and timescale of hybridization and introgression events, among six recently diverged taxa of the horseshoe bat species complexes *Rhinolophus sinicus* and *R. thomasi*. We show that two different methods of species tree reconstruction applied to a set of ~1500 nuclear loci all recover species trees with similar topologies, differing from the previous phylogeny based on two nuclear loci. By comparing the tree topology obtained from the nuclear loci with that inferred from the mitochondrial genome, we observed at least three cases of conflict, each of which likely results from past introgression. Of these, the occurrence of a highly similar mitogenome sequence shared by individuals of two taxa in a sympatric region points to very recent mtDNA introgression. The other cases are characterised by greater divergence and strong phylogeographic structure among putative introgressed individuals and their source populations, and thus likely reflect more ancient hybridization events. These results also suggest that two of the subspecies (*R. s. septentrionalis* and the undescribed taxon *R. s. ssp*) are likely to represent full species, warranting full taxonomic descriptions. This work adds a growing number of studies showing the potential problems of relying solely on mitochondrial sequences, or a limited number of loci, to infer phylogenetic relationships among recently diverged taxa.

1. Introduction

Hybridization and its evolutionary consequences have concerned evolutionary biologists since Darwin (e.g. Abbott et al., 2013; Payseur and Rieseberg, 2016). With the advance of DNA sequencing technology, hybridization - once considered rare in animals - has been documented in many groups (Arnold, 1997; Pennisi, 2016). In particular, past hybridization events (i.e. introgression) have often been interpreted from mito-nuclear discordance (e.g. Chan and Levin, 2005; Toews and Brelsford, 2012). However, characterising the extent and timescale of introgression events relies on having a reliable species tree, and is thus complicated by the fact that conflicts among gene trees are pervasive (Degnan and Rosenberg, 2006, 2009). Apart from introgression, incomplete lineage sorting also commonly leads to gene tree discordance, especially in recently diverged species group (e.g. Zink and Barrowclough, 2008; McKay and Zink, 2010).

In recent years, high throughput sequencing (HTS) has provided the means to generate genome-wide datasets for non-model species for which genomic resources were not previously available (Ekblom and Galindo, 2011; Van Dijk et al., 2014). Yet whole genome sequencing remains prohibitively expensive for studying population-level patterns and processes, including gene flow and phylogeographic history. Instead studies of such phenomena have tended to apply HTS to a fraction of the genome, and methods for genomic enrichment of homologous loci have included RNA-seq (Wang et al., 2009), restriction site associated DNA sequencing (RAD-seq, Baird et al., 2008) and targeted sequence capture (Briggs et al., 2009). Of these, RNA-seq has become an important tool in evolutionary research (Alvarez et al., 2015) including phylogenomics (e.g. Fernández et al., 2014; Davies et al., 2015), however, the need for high quality samples with non-degraded RNA has precluded RNA-seq from most studies. In contrast, RAD-seq has been widely used in population genetic and phylogeographic studies (Narum

* Corresponding authors.
E-mail addresses: xgmao@sklec.ecnu.edu.cn (X. Mao), s.j.rossiter@qmul.ac.uk (S.J. Rossiter).

https://doi.org/10.1016/j.ympev.2019.106551
Received 25 February 2019; Received in revised form 11 June 2019; Accepted 1 July 2019
Available online 02 July 2019
1055-7903/

et al., 2013) as well as in phylogenetic inference (Díaz-Arce et al., 2016). To date, multiple RAD protocols have been developed, each with its relative merits and weaknesses, and therefore choosing a suitable approach is not always straightforward (Puritz et al., 2014).

In most targeted gene resequencing approaches, exons or other conserved sequences have been examined across multiple taxa for reconstructing phylogenetic relationships (e.g. Leaché et al., 2015) and phylogeographic histories (e.g. Westram et al., 2016). Recently, Harvey et al. (2016) demonstrated the usefulness of targeted gene resequencing for comparing across datasets and species because of its ability to generate reliable orthologous markers, compared to methods such as RAD-seq. As such, this approach is likely to be particularly promising for uncovering evolutionary processes that operate over shallow evolutionary timeframes (Harvey et al., 2016; Raposo do Amaral et al., 2018).

The horseshoe bats (family Rhinolophidae) have radiated into more than 100 species within about 20 million years (Csorba et al., 2019). To date, hybridization and introgression have been reported to occur in several closely related species or subspecies of horseshoe bats (Mao et al., 2010a, 2010b, 2013a, 2014, Sun et al., 2016). In particular, multiple cases of hybridization and introgression have been detected in five divergent taxa/clades of *R. sinicus* (East *R. s. sinicus*, Hainan *R. s. sinicus*, Central *R. s. sinicus*, *R. s. septentrionalis*, and *R. s. ssp*) and *R. thomasi* (a closely related taxon to *R. sinicus*) (Mao et al., 2013b; c). In most of these cases, inferences of introgression have been made by comparing mitochondrial DNA tree and nuclear tree based on a few loci generated by Sanger sequencing, however, reconstruction of the true species relationships among recent diverged taxa may not be possible with a limited number of loci. Indeed, two nuclear loci produced conflicting and unresolved phylogenetic relationships among the five taxa of *R. sinicus* and *R. thomasi* (Mao et al., 2013c). More recently large RNA-seq datasets from a single individual of each of the three *R. sinicus* taxa (East *R. s. sinicus*, Central *R. s. sinicus*, and *R. s. septentrionalis*) were used to infer species relationships, and revealed similar proportions of loci supporting a sister relationship between Central *R. s. sinicus* and either East *R. s. sinicus* or *R. s. septentrionalis* (Mao et al., 2017), thus illustrating the difficulty of inferring correct species relationships when using small numbers of individuals. In addition, due to the challenges of RNA-based methods (see above), high-throughput datasets from the other three taxa (Hainan *R. s. sinicus*, *R. s. ssp* and *R. thomasi*) have not yet been obtained.

Here we set out to reconstruct the history of introgression among six recently diverged clades of horseshoe bat using large genomic datasets obtained through sequence capture. We generated and assembled sequence data for over 1500 homologous nuclear loci from multiple individuals of East *R. s. sinicus*, Hainan *R. s. sinicus*, Central *R. s. sinicus*, and *R. s. septentrionalis*, *R. s. ssp* and *R. thomasi*. Using this dataset, we first estimated the phylogenetic relationship among the six focal taxa. We then compared the species tree with the tree based on the mitogenome, and re-examined the cases of mito-nuclear discordances among the six recently diverged taxa, which can be used to assess the extent of introgression. Finally, by analysing the phylogeographic structure of introgressed individuals (i.e. whether introgressed individuals are mixed with their source populations, or are separated from their source populations), we aimed to infer the probable timescale of introgression events.

2. Materials and methods

2.1. Sampling

For targeted gene resequencing, we sampled 37 individuals (Table 1 and Fig. 1). Of these, 28 were identified as *R. sinicus* (six East *R. s. sinicus*, seven Central *R. s. sinicus*, three Hainan *R. s. sinicus*, five *R. s. ssp* and seven *R. s. septentrionalis*), seven as *R. thomasi*, and two from the outgroups (one *R. affinis* and one *R. pearsoni*). Our study taxa show variation in morphology and call frequency. Briefly, *R. s. septentrionalis* has a larger body size and lower echolocation call frequency than both *R. s. sinicus* and *R. thomasi* (see details in Mao et al., 2013b; c), and distinct skull and nose-leaf size have been recorded between *R. thomasi* and *R. sinicus* (Bates et al., 2004).

From each bat, a biopsy of wing tissue was collected using a 3-mm dermatological punch, before the bat was released *in situ*. Our sampling procedure was approved by the National Animal Research Authority, East China Normal University (approval ID 20080209). Samples were stored in 95% ethanol at −20 °C until DNA extraction using DNeasy kits (Qiagen). Genomic DNA was quantified using a Qubit 2.0 Fluorometer.

2.2. Procedures for targeted gene resequencing and data analysis

For sequence capture, we targeted 13 mitochondrial protein-coding (MPC) genes and 1919 nuclear single-copy genes. To monitor any cross contamination between samples during the procedures of library construction and sequencing, we included the sex determining region (*sry*) gene in our targeted gene set. To test for the reliability of gene sequences generated by targeted resequencing strategy here, we obtained the complete cytochrome *b* (*cytb*) sequences from 12 individuals using Sanger sequencing. Details of primers, PCR reaction and the thermal profile for *cytb* are given in Mao et al. (2013b). New sequences of *cytb* were deposited in GenBank (Accession numbers: MK976903-MK976914). In addition, we compared all *cytb* sequences from targeted resequencing with those previously published for the same individuals (Mao et al., 2014).

We performed targeted sequence capture following the approach set out in Bailey et al. (2015). First, baits were designed based on the targeted 1932 coding sequences (CDSs) of *R. pearsoni* which were generated via homology searches of the complete CDSs from human (see below) against the transcriptome of *R. pearsoni* (Mao et al., 2017) using BLASTN (the E-value $< 10^{-6}$). The draft genome sequence of *R. ferrumequinum* (GenBank accession: GCA_000465495.1) was used as the reference during the process of baits design. Bait sequences were divided into 80 nucleotide fragments that overlapped by 40 nucleotides (2× tiling density). Baits design and synthesis were performed by Arbor Biosciences. Second, DNA libraries with insert fragments of ~300 bp were constructed for each sample individually using the NEB Next Ultra DNA Library Prep kit (New England Biosciences). Baits were then hybridized with libraries according to the standard protocol for the MYbaits kit (Arbor Biosciences). After hybridization, the baited libraries were recovered with Streptavidin C1 magnetic beads and were washed in order to remove the off-target sequences. Finally, post-hybridization samples were indexed during PCR amplification using the Agilent SureSelect^{XT} 96 index primers. For sequencing, all libraries were pooled and 150 bp pair-end sequences were generated on an Illumina Hieq2500 using the fast mode.

Raw reads for each sample were cleaned by trimming unknown or low-quality bases (Phred quality score < 20) at the end of the reads using NGSQCToolkit_v2.3.3 (Dai et al., 2010). Only reads longer than 120 bp after trimming were retained for further analysis. Filtered reads were deposited to the NCBI Sequence Read Archive database (SRA) (Accession number SRR157708, SRR9157945-47, SRR9176954-86). To remove untargeted sequences, reads with homology to the targeted genes were first identified by performing BLASTN searches with the expectation value of 1E-6. Reads with blastn hits were extracted and used for *de novo* assembly using *Trinity* v2.2.0 (Grabherr et al., 2011) with default parameters. Redundancy of assembled contigs was reduced using CD-HIT-EST (Li and Godzik, 2006) with 95% sequence identity threshold and only contigs with an average sequencing depth over 10× were retained for further analysis. In addition, we estimated the sequencing coverage of the final assembled contigs using BWA-0.5.7 (Li and Durbin, 2009) and RSEM with default settings.

Table 1
Summary of samples used in targeted resequencing in this study.

ID	Taxon	Sex	Locality	Locality (No.)
FKD001	East *R. s. sinicus*	Male	Wuyishan mountain, Fujian, China	9
FKD002	East *R. s. sinicus*	Male	Wuyishan mountain, Fujian, China	9
SYD04	East *R. s. sinicus*	Male	Jinggang mountain, Jiangxi, China	13
WLB009	East *R. s. sinicus*	Female	Guilin, Guangxi, China	21
WLB010	East *R. s. sinicus*	Male	Guilin, Guangxi, China	21
VN19	East *R. s. sinicus*	Male	Muong Do NR, Vietnam	43
YGL416	Hainan *R. s. sinicus*	Male	Yinggeling, Hainan, China	35
CX07	Hainan *R. s. sinicus*	Female	Qiongzhong, Hainan, China	37
DL161	Hainan *R. s. sinicus*	Male	Lingshui, Hainan, China	39
JJ05	Central *R. s. sinicus*	Female	Zhangjiajie, Hunan, China	24
JJ06	Central *R. s. sinicus*	Male	Zhangjiajie, Hunan, China	24
GZ06	Central *R. s. sinicus*	Female	Anlong, Guizhou, China	29
EM49	Central *R. s. sinicus*	Male	Emeishan, Sichuan, China	30
EM189	Central *R. s. sinicus*	Female	Emeishan, Sichuan, China	30
ZS01	Central *R. s. sinicus*	Male	Shangnuo, Shanxi, China	49
ZS03	Central *R. s. sinicus*	Male	Shangnuo, Shanxi, China	49
YBG18	*R. s. septentrionalis*	Female	Huize, Yunnan, China	31
SHC001	*R. s. septentrionalis*	Male	Fumin, Yunnan, China	33
SHC002	*R. s. septentrionalis*	Male	Fumin, Yunnan, China	33
WM01	*R. s. septentrionalis*	Male	Yongde, Yunnan, China	34
WM04	*R. s. septentrionalis*	Female	Yongde, Yunnan, China	34
WM06	*R. s. septentrionalis*	Female	Yongde, Yunnan, China	34
WM16	*R. s. septentrionalis*	Male	Yongde, Yunnan, China	34
VN37	*R. s. ssp*	Female	Xuan Son, Ninh Binh, Vietnam	47
VN39	*R. s. ssp*	Male	Xuan Son, Ninh Binh, Vietnam	47
VN41	*R. s. ssp*	Male	Xuan Son, Ninh Binh, Vietnam	47
VN44	*R. s. ssp*	Male	Xuan Son, Ninh Binh, Vietnam	47
VN46	*R. s. ssp*	Male	Xuan Son, Ninh Binh, Vietnam	47
MN05	*R. thomasi*	Male	Longling, Yunnan, China	44
MN06	*R. thomasi*	Female	Longling, Yunnan, China	44
MN21	*R. thomasi*	Female	Longling, Yunnan, China	44
MN23	*R. thomasi*	Female	Longling, Yunnan, China	44
MN27	*R. thomasi*	Female	Longling, Yunnan, China	44
B0067	*R. thomasi*	Female	Jinuo, Yunnan, China	45
B0068	*R. thomasi*	Female	Jinuo, Yunnan, China	45

2.3. Generation of exon sequences

To perform gene tree reconstructions, we first identified and partitioned the sequences into exons. While genes with multiple exons may offer greater resolving power, they are also more susceptible to recombination among exons, potentially violating the multispecies coalescent model and leading to erroneous phylogenetic reconstruction (Castillo-Ramirez et al., 2010). This possibility was recently supported by Scornavacca and Galtier (2017), who showed that two exons from the same gene did not yield more similar trees than two exons from distinct genes, and thus exons should be viewed as genomic units for gene-tree based phylogenetic analysis.

To generate a reference exon database, the complete CDSs of 1919 targeted nuclear genes from human were obtained from Ensembl BioMart (http://www.ensembl.org/biomart/). We then used BLASTN (the E-value $< 10^{-6}$) to search for these CDSs in the published *R. sinicus* genomic sequence (GenBank accession: GCA_001888835.1), yielding a total of 12,227 exon sequences from 1693 of the target set of genes. For each individual bat, the assembled contig sequences were used to BLAST against the reference with BLASTN (the E-value $< 10^{-6}$). Exon sequences of targeted genes for each individual were identified and retrieved using custom Perl and bash scripts. Exon-wise sequence alignments were performed in MAFFT v5 (Katoh et al., 2007) with default parameters, and alignments were then trimmed using Gblocks_0.91b (Castresana, 2000) with the default settings except for treating gap positions to remove ambiguously aligned and uninformative positions (i.e. b = h). To increase the number of informative sites in each exon, we filtered out exons < 150 bp. To reduce the effect of missing data on phylogenetic reconstruction, only exons sampled for at least 35 individuals (i.e. 10% missing individuals) were retained. Previous studies have shown that datasets with more individuals sequenced at low genomic coverage might be advantageous relative to larger datasets with fewer individuals screened at greater coverage (e.g. Roure et al., 2013).

To check for the presence of collapsed paralogous genes in our alignments we screened for high levels of variability. For this, we estimated the mean sitewise Shannon entropy for each exon alignment using the Shannon Heterogeneity using a custom script (available on request from J. Parker). Alignments with mean sitewise Shannon entropy values of > 0.02 were checked manually using MEGA7 (Kumar et al., 2016) for potential paralogous copies that may be more divergent with other ingroups than with outgroups. This approach is described in Mao et al. (2017).

2.4. Inference of species tree

Our previous studies have shown evidence of hybridization between *R. s. septentrionalis* and Central *R. s. sinicus* (Mao et al., 2013b, 2017), as well as between *R. s. septentrionalis* and *R. thomasi* (Mao et al., 2013c). In addition, the process of incomplete lineage sorting (ILS) cannot be ruled out in explaining the phylogenomic discordance among taxa from *R. sinicus* (Mao et al., 2017). To test for the effect of ILS and/or introgressive hybridization on species tree inference among six closely related taxa, we used two different methods (concatenation and coalescent methods). First, the concatenation method was applied by reconstructing a maximum-likelihood tree using RAxML 7.2.8 (Berger et al., 2011) based on a super nucleotide matrix from concatenated exon alignments. This analysis was conducted with the GTRGAMMA model and bootstrap support was estimated from 1000 pseudoreplicate searches (the –f a command line option was used).

Second, the species tree was inferred using coalescent methods for the set of complete 3914 exon alignments. Recent studies suggest that

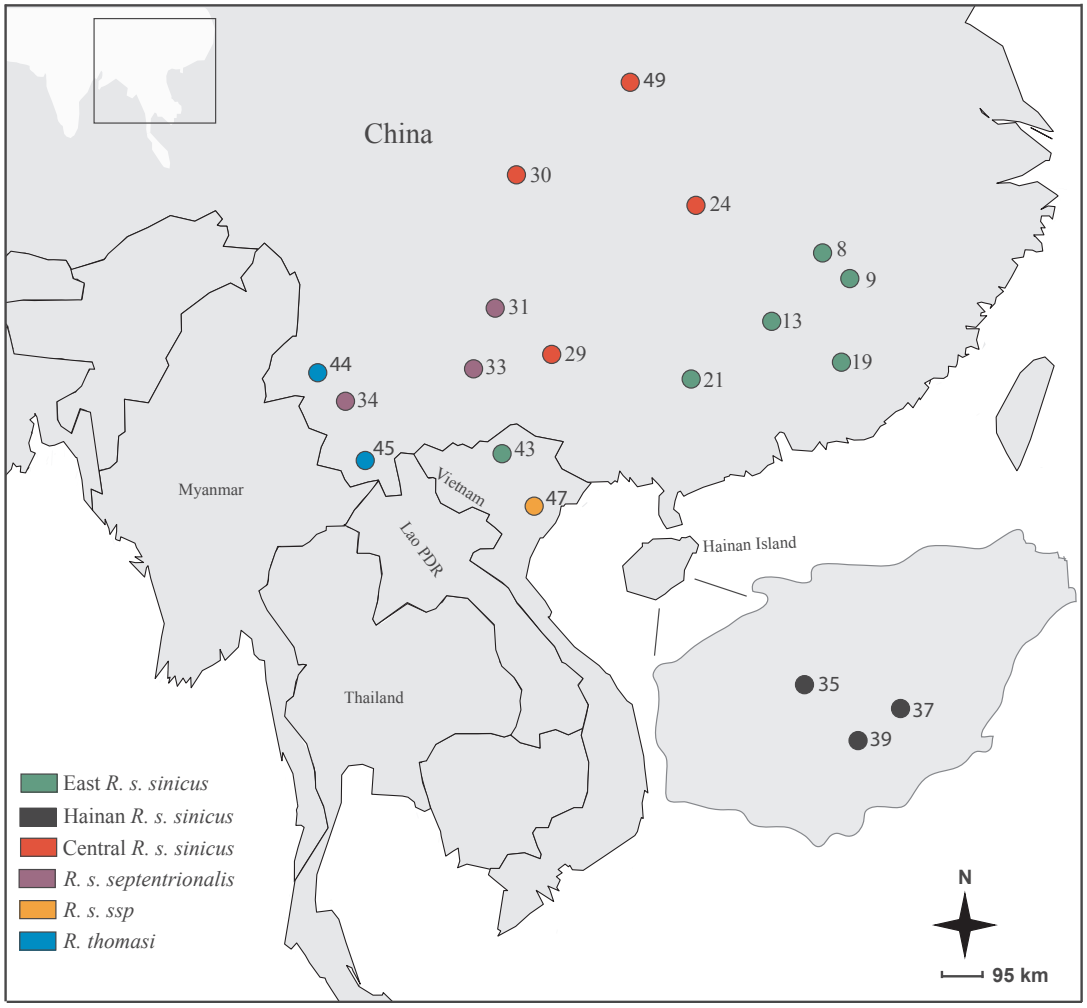

Fig. 1. Sampling map of the six focal taxa in this study (modified from Mao et al., 2013c). The numbers noted for each sampling locality correspond to the ones in Table 1.

concatenation methods can yield results that differ from those obtained by coalescent-based methods, especially in the presence of ILS (Degnan and Rosenberg, 2006; Liu et al., 2015). We implemented the coalescent methods in MP-EST (Liu et al., 2010) and STAR (Liu and Edwards, 2009), which are robust to ILS. For both of these analyses, input ML trees were obtained for each exon alignment using RAxML with the GTRGAMMA model, with bootstrap support estimated from 100 pseudoreplicate searches. All trees were rooted with *R. pearsoni*. Aside from ILS, introgression can also alter the topology of the estimated species tree. Thus to incorporate both processes, we used another coalescent method implemented in a Bayesian concordance analysis (Ané et al., 2007) with the software BUCKy (Larget et al., 2010). This method has been widely used to infer the species tree in recently diverged groups with phylogenetic discordance (e.g. White et al., 2009; Cui et al., 2013). To reduce the computation time, we included two or three individuals from each focal taxon and one from the outgroup *R. affinis*. For BUCKy analysis, the posterior probability of each gene tree for each exon was first obtained using MrBayes 3.1.2 (Ronquist et al., 2012). Two Metropolis-coupled Markov chain Monte Carlo (MCMC) runs were conducted, each comprising four chains and 10 million generations. Trees were sampled every 100 generations, and the first 25% of the sampled trees was discarded as a burn-in. BUCKy was then used to combine all genes and estimate the concordance factor (CF) for each group with different values of the Dirichlet parameter (α = 1, 2 and 5). Output files were visualized with bucky-tools (www. stat.wisc.edu/~ane/bucky/).

2.5. Phylogenetic analysis of mtDNA

We generated the sequences of 13 mitochondrial protein-coding genes (MPCs) applying the same procedure as the above for exon sequences, but using the 13 MPCs of East *R. s. sinicus* (Mao et al., 2017) as the reference. None of the retrieved mtDNA sequences contained stop codons and were thus considered to be genuine mitochondrial genes (i.e. not nuclear copies). Sequences from the 13 MPC genes were aligned using MAFFT with default parameters and were concatenated into a final mitochondrial nucleotide matrix. Based on this mtDNA matrix, we generated trees using Bayesian inference (BI) implemented in MrBayes 3.1.2 and maximum-likelihood (ML) method in RAxML. The best-fit substitution model was determined to be GTR + I + G [I = 0.6096; G = 1.1195] using jModelTest (Posada, 2008). For BI analysis, we conducted two simultaneous runs of Metropolis-coupled Markov chain Monte Carlo (MCMC) analysis using the substitution model parameters above, each comprising four chains and 10 million generations. Trees and parameters were sampled every 100 generations, and the first 25% of the sampled trees were discarded as a burn-in. Convergence of the analysis was determined by checking the average standard deviation of split frequencies that were below 0.01. For ML tree, we used the same parameters as the above for the concatenated nuclear matrix. Trees were rooted with *R. pearsoni*.

3. Results

3.1. Tests for the procedures in targeted gene resequencing

The assembled contigs ranged in average sequencing depth from 111× to 174× per individual bat. After sequence alignment and trimming, an average of 9469 exon sequences were captured from 12,227 targeted exon sequences, ranging from 8766 to 10,132 exons per individual. After filtering out exons that were short (< 150 bp) or sampled in < 35 individuals, we retained 3942 exon sequences for further analysis.

To test for potential cross contamination between samples we examined the male-specific sex determining region gene (*sry*). Of 19 bats for which the *sry* was retrieved (> 120×) we recorded two mismatches with field morphological records. Specifically, no *sry* sequence was obtained from FKD001, assigned as a male in the field, whereas a *sry* sequence was recorded from EM189, assigned as a female in the field. Given that our *cytb* sequences suggested no cross contamination between these two individuals (see below), these mismatches are likely caused by a labelling error in the field.

We used Sanger sequencing to obtain complete *cytb* sequences for 12 of the focal individuals, of which 11 matched the respective sequence obtained by targeted capture. The exception was individual VN46, for which the two sequences differed by 48 nucleotides and one amino acid. To reduce the effect of possible inclusion of nuclear copy of mitochondrial DNA, sequences of 13 MPCs from this individual were excluded from subsequent phylogenetic analyses for mtDNA. In addition, we also compared a section of the *cytb* sequence in the remaining 22 individuals generated here with those published previously from the same individuals, including both EM189 and FKD001 (Mao et al., 2014). This comparison revealed identical respective sequences from the two methods, implying that our targeted gene capture is reliable for phylogenetic analyses.

We estimated heterogeneity (i.e. the mean sitewise Shannon entropy) across all alignments, and found that 319 exon alignments showed mean sitewise Shannon entropy values of > 0.02. Of these, 27 alignments were removed due to misalignments and 28 were edited manually by deleting putative paralogous copies that were highly divergent, even with respect to the outgroup sequences. In the remaining 264 alignments, high heterogeneity was mainly caused by the inclusion of outgroup individuals. In total, 3914 exon alignments were retained, each ≥150 bp long, and sampled for > 34 individuals.

3.2. Inference of species tree

We used four different methods to infer the species tree of six closely related taxa. First, we conducted a ML tree reconstruction based on the concatenated sequence alignments of 3914 exons. The resulting tree revealed one outlier individual from *R. s. septentrionalis* (i.e. SHC002) that fell outside of all other ingroup individuals, and instead showed a closer relationship with the outgroups (Fig. S1). To test for the effects of this outlier on tree topology, we repeated the ML trees after removing SHC002 and observed changes in the topologies of both trees. In particular, the relationships among the major clades differed, especially with respect to the position of *R. s. ssp* (Fig. 2a and S1). We confirmed the correct identification of this individual (SHC002) based on results from mtDNA tree in which it was classified with other *R. s. septentrionalis* individuals (see below). We propose that this outlier is likely an artefactual error, introduced during sequencing and/or reads assembly. In subsequent analyses based on nuclear datasets, we therefore focus on the trees without SHC002.

In the ML tree based on concatenated nucleotide matrix (Fig. 2a) all individuals were classified into the following five maximally-supported (BPP = 1) major clades: *R. thomasi*, Central *R. s. sinicus*, East + Hainan *R. s. sinicus*, *R. s. septentrionalis*, and *R. s. ssp*. The sole exception of individual GZ60707 was from Central *R. s. sinicus*, which was classified with East + Hainan *R. s. sinicus*. Examining the relationships among these five clades revealed high support for (*R. thomasi* + *R. s. septentrionalis*) and ((Central *R. s. sinicus* + (East + Hainan *R. s. sinicus*)), with node support of 93 and 100, respectively. *R. s. ssp* was basal with respect to the other four clades, which were grouped together with the support value of 79.

To complement the concatenation method, which does not account for ILS, the species tree was also inferred using three coalescent-based methods. These were implemented in MP-EST and STAR, which both incorporate ILS (Fig. S2), and BUCKy, which accounts for both ILS and introgression (Fig. S3). These methods recovered almost identical topologies as the concatenated tree produced by RAxML. It was notable that the support rate based on bootstrap analysis from both MP-EST and STAR were very low for each clade or the combination of clades (< 50, data not shown). For BUCKy analysis, we only present results from $\alpha = 1$ because no difference for concordance factor (CF) was observed among different α values.

Overall, both concatenation and coalescent methods supported the following relationships: (i) Central *R. s. sinicus* is sister to East + Hainan *R. s. sinicus*, (ii) *R. thomasi* is sister to *R. s. septentrionalis*, (iii) *R. s. ssp* is basal with respect to the other taxa.

3.3. Phylogenetic analysis on mtDNA

Complete sequences for all 13 MPC genes were obtained for all individuals except for VN46. Based on concatenated 13 MPC genes, phylogenetic analyses using both Bayesian inference (BI) and ML methods generated the same tree topologies in which three major clades were resolved with extremely high support. Among them, the first clade is *R. thomasi*, the second includes *R. s. septentrionalis* and Central *R. s. sinicus*, and the third includes East *R. s. sinicus*, Hainan *R. s. sinicus* and *R. s. ssp* (Fig. 2b). In this tree topology, *R. thomasi* is the basal taxon with respect to the other taxa, thus differing from the ncDNA trees. It was notable that the mitogenomes of two individuals of *R. thomasi* (MN06 and MN21) were almost identical to that of *R. s. septentrionalis*. Additionally, two individuals from Central *R. s. sinicus* (ZS01 and ZS03) were classified with the major clade of (East + Hainan *R. s. sinicus* + *R. s. ssp*) with support rates of 69 and 0.9 in ML and BI trees, respectively. We also noted that the mtDNA sequence of SHC002 was correctly classified with other *R. s. septentrionalis* individuals.

4. Discussion

4.1. Inference of the species tree

Based on a large dataset of 3914 nuclear loci, four different analytical approaches of species tree reconstruction (i.e. concatenation and coalescent methods) recovered almost identical phylogenetic topologies. Specifically, each tree showed sister relationships between Central *R. s. sinicus* and East + Hainan *R. s. sinicus*, and also between *R. thomasi* and *R. s. septentrionalis*, with *R. s. ssp* positioned as basal relative to these former taxa (Fig. 2a). These results, including mtDNA and ncDNA conflicts (see below), provide insights into their evolutionary history and respective species boundaries among the six focal taxa. It is noteworthy that the monophyletic grouping of *R. thomasi* and *R. s. septentrionalis*, and the position of *R. s. ssp*, were not resolved in earlier study based on just two nuclear loci (Mao et al., 2013b), thus highlighting the importance of using genome-wide datasets for determining phylogenetic relationships among recently diverged taxa.

Our finding that the concatenation and coalescent approaches applied to nuclear loci gave consistent results was perhaps surprising. Indeed theoretical and simulation studies indicate that concatenation approaches may sometimes yield misleading estimates of the species tree in the presence of ILS, and in such instances may be outperformed by coalescent methods (Degnan and Rosenberg, 2009; DeGiorgio and Degnan, 2010; Liu et al., 2015). Our current study also indicated that

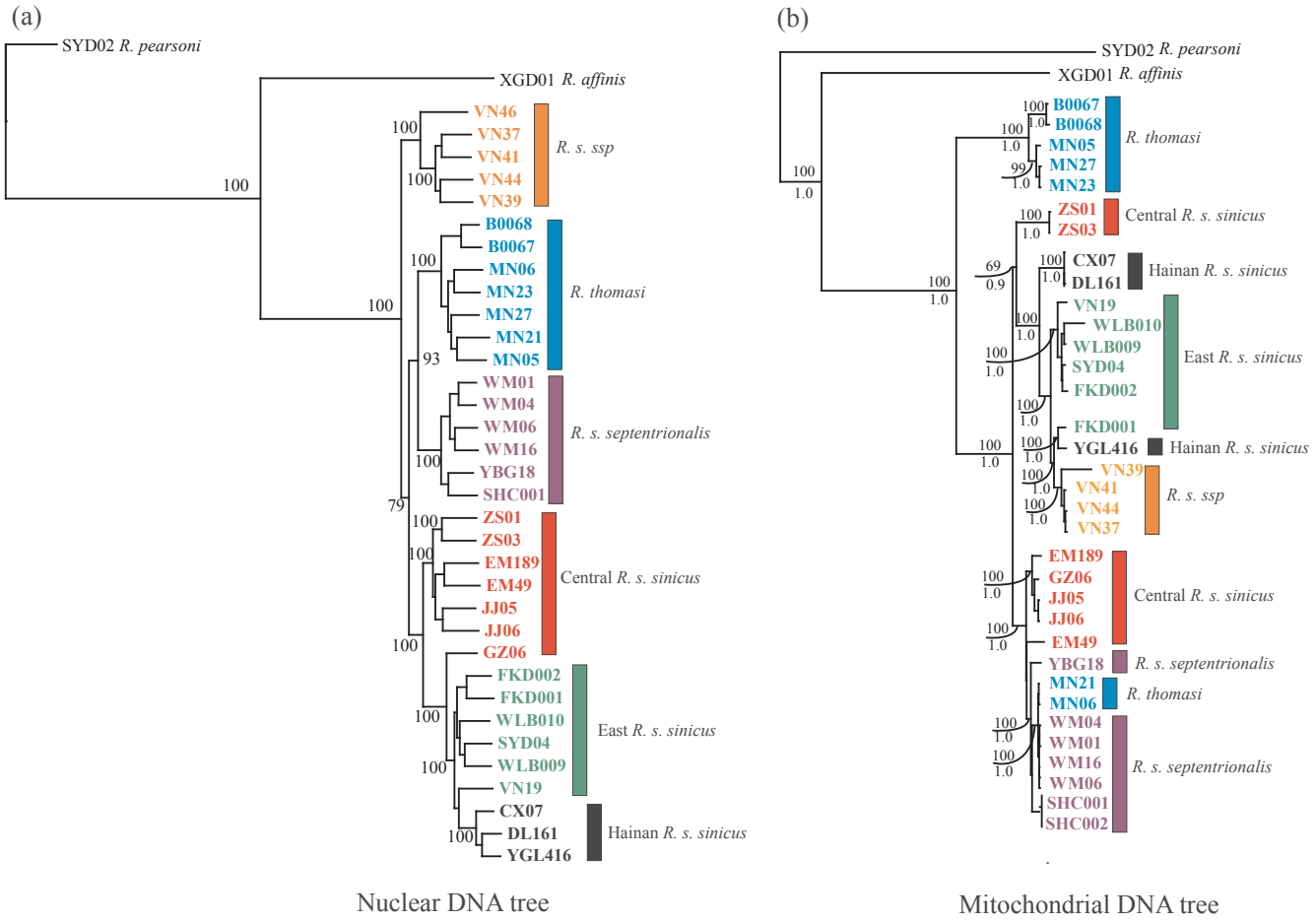

Fig. 2. ML trees reconstructed based on nuclear and mitochondrial genomes, respectively. (a) ML-tree based on a concatenated dataset of 3914 exon alignments. Node support is indicated with ML bootstrap values; (b) ML-tree based on a concatenated dataset of 13 mitochondrial protein-coding genes. Node support is indicated with ML bootstrap values and Bayesian posterior probabilities (given above and below the branch, respectively).

phylogenetic reconstruction based on concatenation was more sensitive than coalescent methods to the presence of an outlier (Fig. S4). On the other hand, concatenation methods continue to be commonplace in the field of systematics (e.g. Fernández et al., 2014; Davies et al., 2015), and coalescent methods are not always found to be superior to concatenation analyses (e.g. in mammalian phylogenomics, Scornavacca and Galtier, 2017), with conflicts likely reflecting the nature of the empirical data. In the future, whole-genome resequencing will be used to investigate the effects of ILS and/introgression on the species tree reconstruction in our current system.

4.2. Discordance between mtDNA and ncDNA datasets

Compared with the species tree, phylogenetic reconstruction based on the complete mitogenome recovered a discordant topology in which *R. thomasi* was the basal taxon, and *R. s. ssp* was classified with East and Hainan *R. s. sinicus* (Fig. 2b). In line with our findings, large numbers of studies have reported extensive discordance between mtDNA and ncDNA datasets (e.g. Weigand et al., 2017; Wallis et al., 2017; Thielsch et al., 2017). Such cases are commonly explained by hybridization among lineages, leading to introgression of the mitochondrial genome, which has a high propensity to cross species boundaries (Chan and Levin, 2005; Toews and Brelsford, 2012). Such introgression can occur via either neutral demographic processes, such as resulting from range expansion during past climatic changes (e.g. Cahill et al., 2013; Marques et al., 2017), or through adaptive retention of mtDNA, such as in species that harbour deleterious mtDNA mutations (Llopart et al., 2014; Hulsey et al., 2016). Introgression of mtDNA among lineages can be commonly inferred by the observation that mtDNA haplotypes are mixed (or shared) in their current or former contact regions. In contrast, related or shared haplotypes among lineages tend to be more geographically scattered under the scenario of ILS (Funk and Omland, 2003).

In our study, two mtDNA haplotypes of *R. thomasi* (MN06 and MN21) (Fig. 2b), shared with *R. s. septentrionalis* and sampled in the current contact region of the two taxa, might result from a recent mitochondrial introgression with *R. s. septentrionalis* (see also Mao et al., 2013c). However, no nuclear introgression was detected between these two taxa, suggesting there was almost complete reproductive isolation between them. In the future, detailed morphological and gene flow analysis will be needed to test for the species status of *R. s. septentrionalis*. Another case of mtDNA introgression was observed between Central *R. s. sinicus* and *R. s. septentrionalis*, with close relationships between most of Central *R. s. sinicus* and *R. s. septentrionalis* in mtDNA tree, compared with close relationships between Central *R. s. sinicus* and East + Hainan *R. s. sinicus* in the ncDNA tree (Fig. 2, see also Mao et al., 2013b, 2017). Unlike the case in the two individuals of *R. thomasi* (MN06 and MN21), the almost complete differentiation of mtDNA observed between Central *R. s. sinicus* and *R. s. septentrionalis* implies that the mtDNA introgression was an ancient event.

Further discordance was seen in the phylogenetic positions of *R. thomasi* and *R. s. ssp*, with *R. thomasi* either basal among the six taxa in the mtDNA tree, or sister to *R. s. septentrionalis* in the ncDNA trees (with *R. s. ssp* as the ancestor). Clear divergence between *R. s. ssp* and other

focal taxa was also reflected by an amino acid change in the *sry* gene, which has been suggested to be associated with reproductive isolation and thus less likely to introgress between taxa (Geraldes et al., 2008; Melo-Ferreira et al., 2009). Thus, our current data support the species status of *R. s. ssp*, which must be verified in the future by detailed morphological analyses. Like Central *R. s. sinicus* and *R. s. septentrionalis*, the classification of *R. s. ssp* and East *R. s. sinicus* in the current mtDNA tree might also result from ancient mtDNA introgression. Such historical events are much harder to detect than recent ones, where mtDNA is often still shared or highly similar (as seen for *R. s. septentrionalis* and *R. thomasi*). It follows that older events are more problematic in the context of phylogenetic inference.

4.3. Applications of sequence capture

This study demonstrates that sequence capture provides a powerful alternative to other some more commonly-used approaches, such as RAD-seq, for obtaining orthologous sequences (also see Harvey et al., 2016). As such, capture approaches are likely to become increasingly adopted in studies of the structure, divergence and phylogeographic history of populations and closely related species. Importantly, sequence capture also has the advantage of generating data that can be combined with other datasets, both current and in the future. Indeed, the large set of loci generated here for several related horseshoe bat taxa of varying phylogenetic relatedness will be useful for comparing with similar datasets of other taxa in this highly speciose family (e.g. Bailey et al., 2015). In the near future, even when whole-genome resequencing becomes commonly applicable, sequence capture will continue to contribute the development of many evolutionary disciplines due to its broad implementation (Jones and Good, 2016).

Yet despite these advantages of sequence capture, our current results also highlight several potential caveats. First, even with due care, cross-contamination between samples may still occur during library construction and sequencing, necessitating additional quality control steps such as the inclusion of sex-specific genes in baits. Second, the presence of mitochondrial genes in baits can give rise to uneven pull-down due to the relatively high copy number of mitochondria, and may also cause the untargeted pull-down of nuclear copies of mitochondrial sequence, potentially leading to misinterpretations of results. Third, where baits are designed to only include orthologous genes, the possibility that paralogous copies are captured cannot be ruled out. Informatics and analytical steps to identify and remove outliers, for example on the basis of high divergence, can reduce these problem.

5. Conclusions

We show that the large-scale sequence capture of nearly 4000 orthologous exons provides a powerful means of resolving phylogenetic and phylogeographic histories at shallow evolutionary timescales. All concatenation and coalescent methods used here produced a single species tree topology in which *R. s. ssp* was positioned as basal, and *R. thomasi* and *R. s. septentrionalis* showed a sister relationship. These results suggest that two of the subspecies (*R. s. septentrionalis* and the undescribed taxon *R. s. ssp*) are likely to represent full species, warranting further taxonomic investigation. In addition, comparisons of phylogenetic reconstructions based on large-scale ncDNA and mtDNA datasets provided compelling evidence of both recent and ancient introgressive hybridization.

Acknowledgements

We thank Zhang Junpeng and Zhu Guangjian for assistances with the field collection and Monika Struebig for her help in the lab work. We also thank two anonymous reviewers whose comments improved the manuscript.

Funding

This work was supported by a Marie Curie International Incoming Fellowship to X. Mao, and a European Research Council Starting grant (ERC-2012-StGr310482) awarded to SJR. X. Mao was also supported by grants from the National Natural Science Foundation of China (No. 31570378) and the Shanghai Pujiang Talent Program Foundation. Vu D. Thong was supported by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 106.05-2017.35, and VAST under the project ID VAST04.10/17-18.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2019.106551.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R., Butlin, R.K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S.H., Hermansen, J.S., Hewitt, G., Hudson, A.G., Jiggins, C., Jones, J., Keller, B., Marczewski, T., Mallet, J., Martinez-Rodriguez, P., Möst, M., Mullen, S., Nichols, R., Nolte, A.W., Parisod, C., Pfennig, K., Rice, A.M., Ritchie, M.G., Seifert, B., Smadja, C.M., Stelkens, R., Szymura, J.M., Väinölä, R., Wolf, J.B., Zinner, D., 2013. Hybridization and speciation. J. Evol. Biol. 26, 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x.

Alvarez, M., Schrey, A.W., Richards, C.L., 2015. Ten years of transcriptomics in wild populations: what have we learned about their ecology and evolution? Mol. Ecol. 24, 710–725. https://doi.org/10.1111/mec.13055.

Ané, C., Larget, B., Baum, D.A., Smith, S.D., Rokas, A., 2007. Bayesian estimation of concordance among gene trees. Mol. Biol. Evol. 24, 412–426. https://doi.org/10.1093/molbev/msm107.

Arnold, M.L., 1997. Natural Hybridization and Evolution. Oxford University Press, Oxford, UK.

Bailey, S.E., Mao, X., Struebig, M., Tsagkogeorga, G., Csorba, G., Heaney, L.R., Sedlock, J., Stanley, W., Rouillard, J., Rossiter, S.J., 2015. The use of museum samples for large-scale sequence capture: a study of congeneric horseshoe bats (family Rhinolophidae). Biol. J. Linn. Soc. 117, 58–70. https://doi.org/10.1111/bij.12620.

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, e3376. https://doi.org/10.1371/journal.pone.0003376.

Bates, P., Thi, M.M., New, T., Bu, S.S.H., Mie, K.M., Nyo, N., Khaing, A.A., Aye, N.N., Oo, T., Mackie, I., 2004. A review of *Rhinolophus* (Chiroptera: Rhinolophidae) from Myanmar, including three species new to the country. Acta Chiropterol. 6, 23–48. https://doi.org/10.3161/1508110042176653.

Berger, S.A., Krompass, D., Stamatakis, A., 2011. Performance, accuracy, and Web server for evolutionary placement of short sequence reads under maximum likelihood. Syst. Biol. 60, 291–302. https://doi.org/10.1093/sysbio/syr010.

Briggs, A.W., Good, J.M., Green, R.E., Krause, J., Maricic, T., Stenzel, U., Pääbo, S., 2009. Primer extension capture: targeted sequence retrieval from heavily degraded DNA sources. J. Vis. Exp. 31, 1573. https://doi.org/10.3791/1573.

Cahill, J.A., Green, R.E., Fulton, T.L., Stiller, M., Jay, F., Ovsyanikov, N., Salamzade, R., John, J.S., Stirling, I., Slatkin, M., Shapiro, B., 2013. Genomic evidence for island population conversion resolves conflicting theories of polar bear evolution. PLoS Genet. 9, e1003345. https://doi.org/10.1371/journal.pgen.1003345.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334.

Castillo-Ramirez, S., Liu, L., Pearl, D.K., Edward, S.V., 2010. Bayesian estimation of species trees: a practical guide to optimal sampling and analysis. In: Knowles, L.L., Kubatko, L. (Eds.), Estimating Species Trees: Practical and Theoretical Aspects. Wiley-Blackwell, Hoboken, NJ, pp. 15–33.

Chan, K.M.A., Levin, S.A., 2005. Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. Evolution 59, 720–729. https://doi.org/10.1554/04-534.

Csorba, G., Hutson, A., Rossiter, S.J., Burgin, C., 2019. Family Rhinolophidae (horseshoe bats). In: In: Wilson, D.E., Mittermeier, R.A. (Eds.), Handbook of the Mammals of the World, vol. 9 Lynx Edicions, Barcelona.

Cui, R., Schumer, M., Kruesi, K., Walter, R., Andolfatto, P., Rosenthal, G.G., 2013. Phylogenomics reveals extensive reticulate evolution in *Xiphophorus* fishes. Evolution 67, 2166–2179. https://doi.org/10.1111/evo.12099.

Dai, M., Thompson, R.C., Maher, C., Contreras-Galindo, R., Kaplan, M.H., Markovitz, D.M., Omenn, G., Meng, F., 2010. NGSQC: cross-platform quality analysis pipeline for deep sequencing data. BMC Genomics 11, S7. https://doi.org/10.1186/1471-2164-11-S4-S7.

Davies, K.T.J., Bennett, N.C., Tsagkogeorga, G., Rossiter, S.J., Faulkes, C.G., 2015. Family wide molecular adaptations to underground life in African mole-rats revealed by phylogenomic analysis. Mol. Biol. Evol. 32, 3089–3107. https://doi.org/10.1093/molbev/msv175.

Degnan, J.H., Rosenberg, N.A., 2006. Discordance of species trees with their most likely gene trees. PLoS Genet. 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.
DeGiorgio, M., Degnan, J.H., 2010. Fast and consistent estimation of species trees using supermatrix rooted triples. Mol. Biol. Evol. 27, 552–569. https://doi.org/10.1093/molbev/msp250.
Díaz-Arce, N., Arrizabalaga, H., Murua, H., Irigoien, X., Rodríguez-Ezpeleta, N., 2016. RAD-seq derived genome-wide nuclear markers resolve the phylogeny of tunas. Mol. Phylogenet. Evol. 102, 202–207. https://doi.org/10.1016/j.ympev.2016.06.002.
Ekblom, R., Galindo, J., 2011. Applications of next generation sequencing on molecular ecology of non-model organisms. Heredity 107, 1–15. https://doi.org/10.1038/hdy.2010.152.
Fernández, R., Laumer, C.E., Vahtera, V., Libro, S., Kaluziak, S., Sharma, P.P., Pérez-Porro, A.R., Edgecombe, G.D., Giribet, G., 2014. Evaluating topological conflict in centipede phylogeny using transcriptomic data sets. Mol. Biol. Evol. 31, 1500–1513. https://doi.org/10.1093/molbev/msu108.
Funk, D.J., Omland, K.E., 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 34, 397–423. https://doi.org/10.1146/annurev.ecolsys.34.011802.132421.
Geraldes, A., Basset, P., Gibson, B., Smith, K.L., Harr, B., Yu, H.T., Bulatova, N., Ziv, Y., Nachman, M.W., 2008. Inferring the history of speciation in house mice from autosomal, X-linked, Y-linked and mitochondrial genes. Mol. Ecol. 17, 5349–5363. https://doi.org/10.1111/j.1365-294X.2008.04005.x.
Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q.D., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 297, 644. https://doi.org/10.1038/nbt.1883.
Harvey, M.G., Smith, B.T., Glenn, T.C., Faircloth, B.C., Brumfield, R.T., 2016. Sequence capture versus restriction site associated DNA sequencing for shallow systematics. Syst. Biol. 65, 910–924. https://doi.org/10.1093/sysbio/syw036.
Hulsey, C.D., Bell, K.L., García-de-León, F.J., Nice, C.C., Meyer, A., 2016. Do relaxed selection and habitat temperature facilitate biased mitogenomic introgression in a narrowly endemic fish? Ecol. Evol. 6, 3684–3698. https://doi.org/10.1002/ece3.2121.
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.
Katoh, K., Kuma, K., Toh, H., Miyata, T., 2007. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucl. Acids Res. 33, 511–518. https://doi.org/10.1093/nar/gki198.
Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874. https://doi.org/10.1093/molbev/msw054.
Larget, B., Kotha, S.K., Dewey, C.N., Ané, C., 2010. BUCKy: Gene tree/species tree reconciliation with the Bayesian concordance analysis. Bioinformatics 26, 2910–2911. https://doi.org/10.1093/bioinformatics/btq539.
Leaché, A.D., Chavez, A.S., Jones, L.N., Grummer, G.A., Gottscho, A.D., Linkem, C.W., 2015. Phylogenomics of phrynosomatid lizards: conflicting signals from sequence capture versus restriction site associated DNA sequencing. Genome Biol. Evol. 7, 706–719. https://doi.org/10.1093/gbe/evv026.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
Li, W., Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22, 1658–1659. https://doi.org/10.1093/bioinformatics/btl158.
Liu, L., Edwards, S.V., 2009. Phylogenetic analysis in the anomaly zone. Syst. Biol. 58, 452–460. https://doi.org/10.1093/sysbio/syp034.
Liu, L., Xi, Z., Davis, C.C., 2015. Coalescent methods are robust to the simultaneous effects of long branches and incomplete lineage sorting. Mol. Biol. Evol. 32, 791–805. https://doi.org/10.1093/molbev/msu331.
Liu, L., Yu, L., Edwards, S.V., 2010. A maximum pseudo-likelihood approach for estimating species trees under the coalescent model. BMC Evol. Biol. 10, 302. https://doi.org/10.1186/1471-2148-10-302.
Llopart, A., Herrig, D., Brud, E., Stecklein, Z., 2014. Sequential adaptive introgression of the mitochondrial genome in *Drosophila yakuba* and *Drosophila santomea*. Mol. Ecol. 23, 1124–1136. https://doi.org/10.1111/mec.12678.
Mao, X.G., Zhang, J.P., Zhang, S.Y., Rossiter, S.J., 2010a. Historical male-mediated introgression in horseshoe bats revealed by multi-locus DNA sequence data. Mol. Ecol. 19, 1352–1366. https://doi.org/10.1111/j.1365-294X.2010.04560.x.
Mao, X.G., Zhu, G.J., Zhang, S.Y., Rossiter, S.J., 2010b. Pleistocene climatic cycling drives intra-specific diversification in the intermediate horseshoe bat (*Rhinolophus affinis*) in Southern China. Mol. Ecol. 19, 2754–2769. https://doi.org/10.1111/j.1365-294X.2010.04704.x.
Mao, X.G., He, G., Hua, P., Jones, G., Zhang, S.Y., Rossiter, S.J., 2013a. Historical introgression and the persistence of ghost alleles in the intermediate horseshoe bat (*Rhinolophus affinis*). Mol. Ecol. 22, 1035–1050. https://doi.org/10.1111/mec.12154.
Mao, X.G., He, G.M., Zhang, J.P., Rossiter, S.J., Zhang, S.Y., 2013b. Lineage divergence and historical gene flow in Chinese horseshoe bat (*Rhinolophus sinicus*). PLoS One 8, e5678. https://doi.org/10.1371/journal.pone.0056786.
Mao, X.G., Thong, V.D., Bates, P.J., Jones, G., Zhang, S.Y., Rossiter, S.J., 2013c. Multiple cases of asymmetric introgression among horseshoe bats detected by phylogenetic conflicts across loci. Biol. J. Linn. Soc. 110, 346–361. https://doi.org/10.1111/bij.12138.
Mao, X.G., Zhu, G.J., Zhang, J.P., Zhang, S.Y., Rossiter, S.J., 2014. Differential introgression among loci across a hybrid zone of the intermediate horseshoe bat (*Rhinolophus affinis*). BMC Evol. Biol. 14, 154. https://doi.org/10.1186/1471-2148-14-154.
Mao, X.G., Tsagkogeorga, G., Bailey, S.B., Rossiter, S.J., 2017. Genomics of introgression in the Chinese horseshoe bat (*Rhinolophus sinicus*) revealed by transcriptome sequencing. Biol. J. Linn. Soc. 121, 698–710. https://doi.org/10.1093/biolinnean/blx017.
Marques, J.P., Farelo, L., Vilela, J., Vanderpool, D., Alves, P.C., Good, J.M., Boursot, P., Melo-Ferreira, J., 2017. Range expansion underlies historical introgressive hybridization in the Iberian hare. Sci. Rep. 7, 40788. https://doi.org/10.1038/srep40788.
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.
Melo-Ferreira, J., Alves, P.C., Freitas, H., Ferrand, N., Boursot, P., 2009. The genomic legacy from the extinct *Lepus timidus* to the three hare species of Iberia: contrast between mtDNA, sex chromosomes and autosomes. Mol. Ecol. 18, 2643–2658. https://doi.org/10.1111/j.1365-294X.2009.04221.x.
Narum, S.R., Buerkle, C.A., Davey, J.W., Miller, M.R., Hohenlohe, P.A., 2013. Genotyping-by-sequencing in ecological and conservation genomics. Mol. Ecol. 22, 2841–2847. https://doi.org/10.1111/mec.12350.
Posada, D., 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256. https://doi.org/10.1093/molbev/msn083.
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.
Pennisi, E., 2016. Shaking up the tree of life. Science 354, 817–821.
Puritz, J.B., Matz, M.V., Toonen, R.J., Weber, J.N., Bolnick, D.I., Bird, C.E., 2014. Demystifying the RAD fad. Mol. Ecol. 23, 5937–5942. https://doi.org/10.1111/mec.12965.
Raposo do Amaral, F., Maldonado-Coelho, M., Aleixo, A., Luna, L.W., Rêgo, P.S.D., Araripe, J., Souza, T.O., Silva, W.A., Thom, G., 2018. Recent chapters of Neotropical history overlooked in phylogeography: shallow divergence explains phenotype and genotype uncoupling in Antilophia manakins. Mol. Ecol. 27, 4108–4120. https://doi.org/10.1111/mec.14843.
Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542. https://doi.org/10.1093/sysbio/sys029.
Roure, B., Baurain, D., Philippe, H., 2013. Impact of missing data on phylogenies inferred from empirical phylogenomic data sets. Mol. Biol. Evol. 30, 197–214. https://doi.org/10.1093/sysbio/syw014.
Scornavacca, C., Galtier, N., 2017. Incomplete lineage sorting in mammalian phylogenomics. Syst. Biol. 66, 112–120. https://doi.org/10.1093/sysbio/syw082.
Sun, K., Kimball, R.T., Liu, T., Wei, X., Jin, L., Jiang, T., Lin, A., Feng, J., 2016. The complex evolutionary history of big-eared horseshoe bats (*Rhinolophus macrotis* complex): insights from genetic, morphological and acoustic data. Sci. Rep. 6, 35417. https://doi.org/10.1038/srep35417.
Thielsch, A., Knell, A., Mohammadyari, A., Petrusek, A., Schwenk, K., 2017. Divergent clades or cryptic species? Mito-nuclear discordance in a *Daphnia* species complex. BMC Evol. Biol. 17, 227. https://doi.org/10.1186/s12862-017-1070-4.
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.
Van Dijk, E.L., Auger, H., Jaszczyszyn, Y., Thermes, C., 2014. Ten years of next-generation sequencing technology. Trends Genet. 30, 418–426. https://doi.org/10.1016/j.tig.2014.07.001.
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.
Wang, Z., Gerstein, M., Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. 10, 57–63. https://doi.org/10.1038/nrg2484.
Weigand, H., Weiss, M., Cai, H., Li, Y., Yu, L., Zhang, C., Leese, F., 2017. Deciphering the origin of mito-nuclear discordance in two sibling caddisfly species. Mol. Ecol. 26, 5705–5715. https://doi.org/10.1111/mec.14292.
Westram, A.M., Panova, M., Galindo, J., Butlin, R.K., 2016. Targeted resequencing reveals geographical patterns of differentiation for loci implicated in parallel evolution. Mol. Ecol. 25, 3169–3186. https://doi.org/10.1111/mec.13640.
White, M.A., Ané, C., Dewe, C.N., Larget, B.R., Payseur, B.A., 2009. Fine-scale phylogenetic discordance across the house mouse genome. PLoS Genet. 5, e1000729. https://doi.org/10.1371/journal.pgen.1000729.
Zink, R.M., Barrowclough, G., 2008. Mitochondrial DNA under siege in avian phylogeography. Mol. Ecol. 17, 2107–2121. https://doi.org/10.1111/j.1365-294X.2008.03737.x.