Heliyon 6 (2020) e05583

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Markers for genetic change

Giovanni Forcina [a,b,*], Miguel Camacho-Sanchez [a,b], Fred Y.Y. Tuh [c], Sacramento Moreno [d], Jennifer A. Leonard [a,*]

[a] Conservation and Evolutionary Genetics Group, Estación Biológica de Doñana (EBD-CSIC), 41092 Seville, Spain
[b] CIBIO/InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal
[c] Sabah Parks, Lot 45 & 46 KK Times Square Coastal Highway, 88100 Kotah Kinabalu, Sabah, Malaysia
[d] Department of Ethology and Biodiversity Conservation, Estación Biológica de Doñana (EBD-CSIC), 41092 Seville, Spain

ARTICLE INFO

Keywords:
Conservation
Genetics
Intron
High-throughput sequencing NGS
Multiplex
Wildlife management
Agricultural science
Environmental science
Earth sciences
Biological sciences
Veterinary medicine

ABSTRACT

Background and aims: Wildlife conservation has focused primarily on species for the last decades. Recently, popular perception and laws have begun to recognize the central importance of genetic diversity in the conservation of biodiversity. How to incorporate genetic diversity in ongoing monitoring and management of wildlife is still an open question.

Methods: We tested a panel of multiplexed, high-throughput sequenced introns in the small mammal communities of two UNESCO World Heritage Sites on different continents to assess their viability for large-scale monitoring of genetic variability in a spectrum of diverse species. To enhance applicability across other systems, the bioinformatic pipeline for primer design was outlined.

Results: The number of loci amplified and amplification evenness decreased as phylogenetic distance increased from the reference taxa, yet several loci were still variable across multiple mammal orders.

Conclusions: Genetic variability found is informative for population genetic analyses and for addressing phylogeographic and phylogenetic questions, illustrated by small mammal examples here.

1. Introduction

Genetic diversity is important for the long-term survival of species, and thus is of high conservation concern (Frankham, 2002; Spielman et al., 2004; O'Grady et al., 2006). This is even more true in an epoch of unprecedented human-caused disturbance and environmental change, where the preservation of adaptive potential and functional diversity are key for populations to respond to changing environments (Hoelzel et al., 2019; Razgour et al., 2019). Nevertheless, laws and policy have long focused primarily on more visible elements of biodiversity, such as species or populations, with genetics being most often neglected (Laikre, 2010). Genetic diversity is increasingly taking an important role in the international political agenda. It has been included in well-known international mandates such as the Habitats Directive 92/43/EEC and the Birds Directive 79/409/EEC, which apply across the European Union (EU). More recently, the United Nations (UN) has declared 2010–2020 the Decade of Biodiversity and, in the *Convention on Biological Diversity* negotiated in Nagoya, Japan, delegates adopted a strategic plan on global biodiversity. The preservation of genetic diversity has been listed among the Aichi Biodiversity Targets (Strategic Goal C, Target 13) and in the Post 2020 Biodiversity Framework (OECD) becoming an obligation for the EU member states as well as other signatories. The EU has the stated objective to slow the loss of biodiversity and ecosystem services by 2020, and restore, as possible, the environment in order for its member states to contribute to the maintenance of global biodiversity. One of the official criteria to fulfill this task is the maintenance of genetic diversity.

However, there is no consensus on how this diversity should be measured. In the context of research, a number of studies have looked at changes in genetic diversity through time, especially before and after certain events, often periods of overexploitation (e.g., Baker et al., 2000; Pinsky and Palumbi, 2013; Sanchez-Donoso et al., 2014) or pollution exposure (for a review, see van Straalen and Timmermans, 2002). These studies have been based primarily on population level analyses of mitochondrial DNA markers such as the control region or cytochrome *b* (cyt *b*), or autosomal markers such as microsatellite loci. Mitochondrial DNA (mtDNA) markers are generally variable within populations, easy to amplify with standard protocols, applicable across a wide taxonomic range, yield sequences comparable between studies, and there is a lot of

* Corresponding author.
E-mail addresses: giovanni.forcina@cibio.up.pt (G. Forcina), jleonard@ebd.csic.es (J.A. Leonard).

https://doi.org/10.1016/j.heliyon.2020.e05583
Received 1 June 2020; Received in revised form 23 September 2020; Accepted 19 November 2020
2405-8440/© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

comparative data available (e.g., Koblmüller et al., 2012, 2016). They are also easier than nuclear markers to type from degraded material, such as feces and old bones, because of their higher per-cell copy number (Templeton et al., 2013). The downside of using mtDNA to track changes through time is that it is a single marker with necessarily limited power, and also it only informs about the history of the female lineage (Heled and Drummond, 2009). These constraints, among others, mean that mtDNA diversity does not necessarily correlate with nuclear diversity (Teske et al., 2018) and the historical demography of the species being studied (Bazin et al., 2006; Nabholz et al., 2008).

Microsatellites are other popular markers for population level studies. Even though these loci are more difficult to isolate and optimize, the system generally yields data from multiple, independently inherited genetic markers useful for population level metrics (Pinsky and Palumbi, 2013). Generally very abundant in the genomes of most species, now with next generation sequencing (NGS) methods they are easier to isolate than before (Yang et al., 2014). Their loci are often very variable within populations, and so offer higher power to identify changes in genetic diversity through time (Haasl and Payseur, 2011; Putman and Carbone, 2014). The primary drawback of microsatellite studies is that the results are generally not comparable between laboratories (Moran et al., 2006), or even between different projects in the same laboratory. Moreover, ascertainment bias from the marker discovery approaches (Dufresnes et al., 2014; de Groot et al., 2016) and the low number of loci in a typical microsatellite dataset (Fischer et al., 2017) can limit their power for reliable estimates of genetic diversity and population structure (Camacho-Sanchez et al., 2020) as well as the many assumptions associated with their analysis and interpretation (Putman and Carbone, 2014). Overall, this makes it very difficult to build on previous batches of data, which is fundamental to ongoing genetic monitoring. Ideally there would be a panel of markers that could be applied to a variety of non-model organisms in a way such that genotypes produced at different times and/or in different laboratories, and/or with different methods would be comparable.

Another molecular tool used in some studies and partly meeting these criteria is intron sequences. These have some of the beneficial characters of mtDNA, such as being sequence-based markers that can be easily shared and compared between projects, and for the applicability of methods across different species as exemplified by Transcriptome Ortholog Alignment Sequence Tools (TOASTs: Jiang et al., 1998; Wcisel et al., 2020) and Comparative Anchor Tagged Sequences (CATs: Lyons et al., 1997). Likewise, introns share some of the advantages of microsatellites in that they are biparentally inherited, numerous, and evenly distributed throughout the genome. Similar to other neutral nuclear markers, introns may represent a good proxy to assess functional adaptive potential or functional diversity (Vilas et al., 2015). This combination of features offers great potential for monitoring genetic diversity in wildlife. Although generally much less variable than nuclear microsatellites or mitochondrial control region sequences (but more than respective exons: Igea et al., 2010), introns are abundant, so their power to evaluate intra-specific genetic diversity can be increased by genotyping more loci. Multilocus PCR panels can be easily established profiting from the generally conserved flanking exonic regions, which enables the use of the same primers across different taxa. However, amplification and sequencing of multiple loci can quickly become expensive and logistically complicated. Perhaps for this reason, panels of introns have mostly been used in systematic studies that look at the relationships between different species and sample few individuals (e.g., Hailer et al., 2012; Igea et al., 2013). With recent methodological advances such as molecular indexing and next generation sequencing, some of these problems can be overcome (Meyer and Kircher, 2010). The large number of publicly available genomes has enabled the development of panels of possibly suitable loci (Igea et al., 2010; Rodríguez-Prieto et al., 2014). Recently, panels of introns have been used in population level studies (Pons et al., 2010; Tollis et al., 2012; Cordero et al., 2014; Kuchta et al., 2016; Camacho-Sanchez et al., 2018).

Small mammals (defined as terrestrial mammals weighing 5 kg or less: Merritt, 2010) play a pivotal role in their ecosystems as they account for a considerable biomass proportion and include both primary and secondary consumers, seed dispersers, and predators (Ostfeld et al., 1996; Aschwanden et al., 2007). By being easy to collect and handle, and mostly habitat specialists, small mammals represent useful bioindicators and are good candidates for comparing geographically distant study areas (Talmage and Walton, 1991; Smith et al., 2002). Moreover, with their short lifespan and rapid life-history responses to environmental changes, taxa falling into this category are particularly suitable for understanding environmental effects on animal population dynamics (Rowe and Terry, 2014; Hope et al., 2017) and for singling out specific anthropogenic drivers (Byrom et al., 2015). Since *in situ* species conservation is primarily achieved by protected areas, the preservation of genetic diversity is also likely to be dependent on them. Documenting the genetic diversity in these sites will therefore be of key concern in implementing the new mandate to preserve genetic diversity.

In this study we assess a panel of nuclear markers for utility in characterizing genetic diversity in the small mammal communities (Table 1) of two UNESCO World Heritage Site national parks located at opposite ends of Eurasia (Figure 1). This panel of markers (Table 2, Table S1) is amplified in a single multiplexed PCR reaction, decreasing lab time and reagent cost. The indexed PCR products are pooled and sequenced using high-throughput sequencing platforms. We test the applicability of this panel in different taxa with increasing phylogenetic distance from the species for which it was developed, the brown rat (*Rattus norvegicus*), and evaluate the power for population genetics purposes. We also test the comparability of data generated through different library preparation protocols and on different NGS platforms. In order to facilitate the applicability of this method in other systems, the bioinformatics pipeline for primer development is described in detail. Finally, the phylogenetic utility of the panel is evaluated by a comparison with phylogenies based on mtDNA.

2. Results

2.1. Amplification and polymorphism across taxa

The number of loci yielding reads varied from 40 (the entire panel: Table 1, Table 3, Table S2, Appendix S1) in *Rattus* (and other representatives of the subfamily Murinae) to roughly half among Eulipotyphla (here represented by shrews and gymnures) and even less in other distantly related taxa from other orders such as Scandentia (treeshrews: only 17). Two loci (Dhcr24, Smo) amplified across all orders tested. Polymorphisms were detected in all the assembled loci, with the highest values in the subfamily Murinae. In terms of species, the maximum value was recorded in *Leopoldamys sabanus* (34 loci, 85%), while the number of alleles per species and locus ranged from one to eleven (Table 3). The lowest number of segregating sites (S) were recorded in non-rodent taxa, although the minimum value (0) was observed in a squirrel, *Sundasciurus jentinki*, and a murid, *Chiropodomys pusillus*, other than in a shrew, *Suncus etruscus* (Table 1, Table 3, Table S2 Appendix S1). Five loci (Dhcr24, P2rx1, Smo, Usp20, and Wls) were the most variable, with Dhcr24 and Smo yielding in total more than 100 and 80 alleles across 16 and 13 species, respectively. Genetic diversity indexes revealed a higher variability among murids, with values generally exceeding those inferred at the reference nuclear loci, the growth hormone receptor exon 10 (Ghr) and the retinol binding protein 3 exon 1 (Rbp3) (Appendix S1). Some of the loci that did not work in the amplicon libraries yielded sequences from shotgun libraries (i.e., Wls, Rras). Variation in coverage among loci in the same PCR and sequencing platforms did not seem to be driven by taxonomic affiliation (Figure 2). We discarded 85 out of a total of 97 (87.6 %) alleles obtained exclusively with GS Junior 454 due to sequence variation in homopolymeric regions which might be attributable to sequencing errors. Sequences at issue were not employed for tree

Table 1. List of taxa used in this study. Number of individuals sampled (n), small mammal community (DNP: Doñana National Park; IB: Iberian Peninsula; KNP: Kinabalu National Park), number of loci tested (**# T**), number of loci yielding reads (**# R**) in amplicon (PCR-based) and in-solution enriched libraries (left to right), and number of loci that were polymorphic (**# P**) are also reported. Information on polymorphic markers for shotgun libraries is reported only for sequences data covering at least 40% of the locus length, but amplification was considered successful even if a few reads mapping the reference genomes were obtained.

Order	Family	Species	n	Community	Loci		
					# T	# R	# P
Eulipotyphla	Soricidae	*Crocidura russula*	2	DNP	40	3–24	3
		Crocidura suaveolens	2	DNP	40	2–22	1
		Suncus etruscus	2	DNP	40	2–24	0
Lagomorpha	Leporidae	*Oryctolagus cuniculus*	2	DNP	40	2–31	4
Rodentia	Cricetidae	*Arvicola sapidus*	3	DNP	40	7–33	3
		Chionomys nivalis	3	IB	40	8–32	5
		Microtus cabrerae	1	IB	40	8–18	2
	Muridae	*Apodemus sylvaticus*	7	DNP	40	12–37	12
		Mus spretus	8	DNP	40	21–38	14
		Rattus norvegicus	6	DNP	40	25–40	16
		Rattus rattus	2	DNP	40	27–40	19
	Gliridae	*Eliomys quercinus*	5	DNP	40	2–26	3
	Sciuridae	*Sciurus vulgaris*	4	IB	40	4–36	2
Eulipotyphla	Erinaceidae	*Hylomys suillus*	2	KNP	34	2–11	1
Rodentia	Muridae	*Chiropodomys pusillus*	1	KNP	40	9–39	0
		Lenothrix canus	1	KNP	40	13–39	6
		Leopoldamys sabanus	5	KNP	40	18–40	34
		Maxomys ochraceiventer	1	KNP	40	19–40	5
		Maxomys whiteheadi	5	KNP	40	21–40	25
		Niviventer cremoriventer	4	KNP	40	19–40	20
		Rattus baluensis	3	KNP	40	27–40	21
		Sundamys muelleri	6	KNP	40	25–40	30
	Sciuridae	*Callosciurus prevostii*	2	KNP	40	5–30	3
		Sundasciurus jentinki	1	KNP	40	3–28	0
		Sundasciurus lowii	1	KNP	40	5–23	1
		Sundasciurus everetti	2	KNP	34	4–31	3
Scandentia	Tupaiidae	*Tupaia montana*	2	KNP	40	2–17	1

reconstruction or computation of standard genetic diversity indexes, but were scored as successful amplifications.

2.2. Phylogenetic reconstructions

The phylogenies based on nuclear markers (Figure 3) were consistent with recent work combining multiple exonic and mitochondrial markers (Rowe et al., 2019) for the subset of taxa included in both studies. However, the mtDNA phylogeny (Figure 4A and B) showed internal discordances in the clade from the Dacnomys division (*Lenothrix, Leopoldamys* and *Niviventer*). For these complicated relationships, the panel of nuclear markers seemed to resolve the evolutionary affinities more accurately than mitochondrial genome sequences.

3. Discussion

Anthropogenic activities are wiping out biological diversity at an alarming rate, raising serious concern for ecosystem functioning and the delivery of associated services in an epoch of increasing environmental stochasticity (e.g., Kremer et al., 2012; Valiente-Banuet et al., 2015; Seddon et al., 2016). Now that the broader society has started recognizing the importance of preserving genetic diversity to counteract this trend, a vital question being tackled is how to measure and compare it. Major advances in genomic technologies have heralded a new era, enabling researchers to produce massive amounts of genome-wide data on multiple individuals. Although budget, computational, and other logistical constraints still impair the large-scale adoption of whole genome approaches (Fuentes-Pardo and Ruzzante, 2017), using genomic tools to genotype targeted loci with fast, cheap and easy protocols has the potential for widespread application. There is a growing body of literature showing efforts to take stock of the plethora of molecular approaches that are being applied for biodiversity monitoring, especially among mammals (e.g., Larsen and Matoq, 2019; Forcina and Leonard, 2020). In this study, we tested a multiplexed panel of intron loci in spatially and taxonomically different small mammal communities of major ecological relevance to develop a tool for genetic monitoring which could be applied to many different taxa with limited effort.

Primers designed with the pipeline used in this study (Figure S1) and the laboratory rat genome largely worked as a single multiplex. None of the primer pairs appeared to amplify other loci or duplicated regions, suggesting that these loci will be of value across mammals, and could be worth testing in other vertebrate groups as initially suggested by Igea et al. (2010). We successfully amplified all the loci tested in at least some species, and obtained sequences across all the surveyed taxa from at least a few loci. Amplification success decreased with phylogenetic distance from the model species used for designing primers, *R. norvegicus*. The panel worked well across the entire family Muridae, whose most recent common ancestor is estimated at 11.8 MYA (Kimura et al., 2015). Two loci (Dhcr24, Smo) even worked across all the tested orders, whose most recent common ancestor is estimated at 80–95 MYA (Dos Reis et al., 2012; Foley et al., 2016; Upham et al., 2019). Given the higher rate of mutation in rodents than other groups of mammals (e.g., Cooper et al., 2004), primers may be better conserved in other orders. These results show that the genome of the target species is not necessary for primer development, as primers in evolutionary related species worked consistently well. Hence, when setting up a new project, it is recommendable to

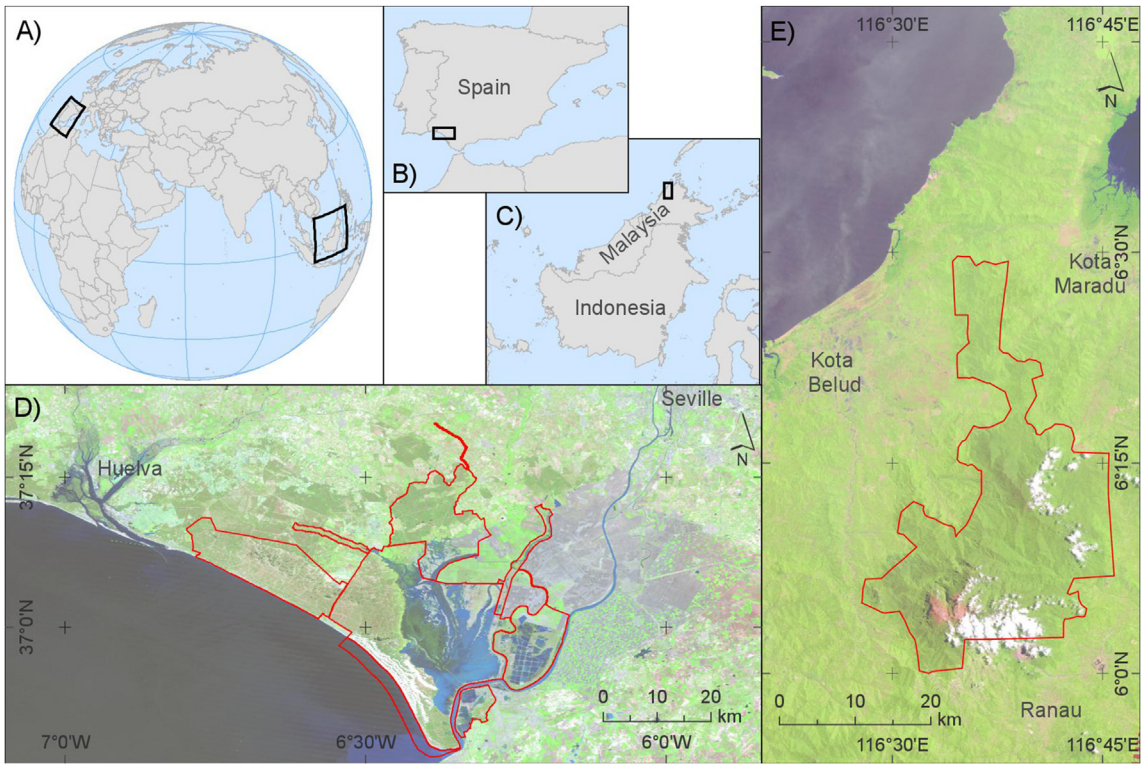

Figure 1. Study areas. A. Location of Iberia and Borneo in the world map. B, C. Doñana and Kinabalu National Parks in Iberia and Borneo, respectively. D, E. Satellite images of Doñana and Kinabalu National Parks, respectively. Park boundaries are represented by red lines.

check the fast-increasing list of newly available genomes and select the closest one to the study species for primer design. Alternatively, the primer design step could be partially avoided and the loci genotyped through sequencing of enriched shotgun libraries. This, however, will require more complex laboratory protocols, deeper sequencing (which is relatively cheap), and more importantly, will be bioinformatically more complicated to analyze.

The intron sequences contained a good amount of information as compared to those of commonly targeted nuclear loci such as Rbp3 or Ghr (Appendix S1), with fairly high genetic diversity indices in Muridae and, to a lesser extent, across other families of the order Rodentia (Table 3, Appendix S1). These results indicate that this panel may prove useful for addressing population level genetic questions (Camacho-Sanchez et al., 2018). It will still be necessary to establish appropriate species or population specific baselines from which power estimates can be made. In many cases populations of wildlife have been subject to historic changes in population size and connectivity, so it may be most appropriate to establish these baselines with historic data.

This panel of markers was also sufficiently informative to construct a high confidence evolutionary phylogeny of the Rattini in our dataset. In comparison to the whole mitochondrial genome phylogeny with the same set of taxa, we show that the nuclear loci were better able to resolve the difficult nodes in the clade containing *Leopoldamys, Lenothrix,* and *Niviventer*.

Overall, this protocol should work in any animal taxa. Additionally, these particular loci should work at least across most mammals, and these primers in most Rattini and many rodents. Nevertheless, it is best to design taxon-specific primers for the target species for each particular study to ensure the highest amplification success, and this will not impair the comparability across studies. Indeed, the primers tested here worked best in the target group, and success dropped off substantially in other orders. To apply this panel in other groups of mammals, it would be advisable to design new, taxon specific primers, if possible. The other way to get around this- targeting the loci trough enrichment instead of PCR - required more sequencing and is bioinformatically more complex. Also, power calculations should be made for each species and question in order to ensure that the number of loci sequenced is sufficient to achieve the level of taxonomic resolution needed. Further, in this test we have used high quality samples. If low quality samples such as feces are used, PCR replicates may be necessary.

In conclusion, here we propose an efficient and effective molecular tool for the genetic screening of small mammal communities. We describe a multiplexing strategy which, in combination with the barcoding of multiple individuals, represents a cost- and time-efficient as well as easy-to-implement procedure for use by academics, governmental agencies, and wildlife managers. This has a major advantage over mitochondrial DNA alone because it relies on many independent markers. Unlike microsatellites, the sequence-based markers proposed here generate data which can be compared between projects and/or laboratories making them particularly suited to ongoing management of genetic diversity. The panel tested shows its utility across evolutionary scales, from population genetics to intra-species phylogenies, from the same genotyped dataset. The protocol applied is easily transferable to other study systems, thus making a substantial contribution to the establishment of standardized monitoring strategies for counteracting the harmful biodiversity loss in this epoch of accelerating global change.

4. Materials and methods

4.1. Study sites and species

The study focuses on two important national parks, Doñana National Park (DNP, Spain) and Kinabalu National Park (KNP, Malaysia) (Figure 1). Located within the Mediterranean Basin and Sundaland biodiversity hotspots (Myers et al., 2000), these protected areas were awarded World Heritage Site by UNESCO in 1994 and 2000, respectively. The parks contain a diversity of small mammals with no species shared, the only exceptions being the invasive black (*R. rattus*) and brown (*R. norvegicus*)

Table 2. Summary of the nuclear loci and associated primers used in this study. For the sake of clarity, amplicon size incorporates target sequence plus both primers. int.: intron; ex.: exon.

Rat gene	Intron	Size - bp (*R. norvegicus*)	Forward primer 5'-3'	Reverse primer 5'-3'	Reference
Abcb9	2 (int.)	584	GCATYGTSATCCAGAARAGCAYGGA	GCTGTGCGRTTCTCRTCRAARAAGCT	Rodríguez-Prieto et al., 2014
Abcg8	9 (int.)	430	TTTCCAATGACTTCCGGGAC	GGCAAAGAAATAAGGACCAGCA	Camacho-Sanchez et al., 2018
Agxt	10 (int.)	560	GGCTACAACTGGAGGGACATC	GTGCAGGGCCTCCYTCAGGGCCT	Rodríguez-Prieto et al., 2014
Alkbh7	3 (int.)	429	GCTGGAGGTGGCTCTTCTG	CTGGCCTTTCCCTGTTGTCT	Camacho-Sanchez et al., 2018
Apeh	14 (int.)	447	KGACACCCATGACACAGACT	CCCAGTTCTCCACACCCA	″
Apeh	17 (int.)	456	GAAAGGATGCTGTCTTGGCC	GGGGTGGCCTTGGTTGTATA	″
Catsper3	5 (int.)	521	TGCTKGCMTCSTTCATCTT	GAGRATYAYYTGCTTCTYCTCC	Rodríguez-Prieto et al., 2014
Cd27	5 (int.)	419	CAGGCTCRGGTTTCCGGT	TCCGGATCTTTGTGACCTTCT	Camacho-Sanchez et al., 2018
Chrna9	1 (int.)	447	TTATCTGGGAGAGCGTGACC	TTGGGAAARGATGAACCGGC	″
Dhcr24	7 (int.)	444	CAGGACATGCTGGTGCCCATGAA	GCCTGGCTGGCTGGGCAGGATGAA	Rodríguez-Prieto et al., 2014
Dhrs3	3 (int.)	436	CTCCTCAAGTCCCAGCATGT	GCACRGAATTGAGGCACACA	Camacho-Sanchez et al., 2018
Fancg	9 (int.)	485	CCTTTAGTTGTGACCAGGCC	GAGCATTACCTGGACCTGCT	″
Fetub	1 (int.)	434	ACAGAGAKCCCATGTCTTCC	GCCCTGCAGAACATCAACAG	″
Fnrkrp	5 (int.)	444	AGATGGACATGGTGGAGAAGA	AGTGKCCATAGAAGGATGCT	″
Gabrp	1 (int.)	434	TCTGCTGACCTCCACATTGA	AGCTACAGYCTCTATTTGGCCT	″
Gadd45g	1 (int.)	442	GACCTCCAAGTCCCAGCTG	GGATACAGTTCCGGAAAGCAC	″
Il34	3 (int.)	436	GGTACTCAGAGTGGCCAACA	CCAGCAATGTCTGAACCTCC	″
Irf5	7 (int.)	423	AAACCCCGAGAGAAGAAGCT	CTGGACCATGGGCTGCAA	″
Ivd	8 (int.)	637	CTGGACCTRGARCGCCTGGT	GCTGRAAKTGSCCRATYTTCT	Rodríguez-Prieto et al., 2014
Klc2	10 (int.)	428	AAAGCCCTACCTGTTTGCG	TCAGGATAAGCGCCGGGA	Camacho-Sanchez et al., 2018
Mmp9	2 (int.)	460	GATGATGGGAGAGAAGCAGTC	GTCTCGCGGCAAGTCTTC	″
Ms4a2	5 (int.)	430	ACACCAGTTCCTGTCAAACA	CTYCGCTTATATGAACWACTGCA	″
Mycbpap	11 (int.)	481	GGCAGAATCACACCTGGGA	GGTCAATAACGGCACKGTGG	″
Nadsyn1	4 (int.)	643	GTYCGYTACAAYTGCAGAGT	GTCCTKSHCCAKGGGGTRAACCA	Rodríguez-Prieto et al., 2014
Nfkbia	5 (int.)	533	GCCTCCAAACACACAGTCAT	TGAGGAGAGCTATGACACGG	Camacho-Sanchez et al., 2018
Npr2	10 (int.)	480	TGAACTCAAACACGTACGTACT	TGGTTGAACTGRACATCTCTCA	″
P2rx1	3 (int.)	461	CATTGTGCAGAGGTGAGGAC	TCTGCTTTTCCTGGAGTGCA	″
Pipox	5 (int.)	424	TCTGAGAAGGTTTTGGGGCA	CCCACCACATCTAYGGACTG	″
Ptgs2	7 (int.)	467	GTGTATCCYCCCACAGTCAAA	TGAGTTTGAAGTGGTAACCGC	″
Rabac1	1 (int.)	453	AATACTCCACGTTGCGWACC	CAGAAGGACCAGCAGAAGGA	″
RGD735029	5 (int.)	416	CTTCGGAGGCATGTTCTTCC	CCTTTGCCTGGGATGYGAAG	″
Rogdi	7 (int.)	435	AGAARCCGGCTCACTACCC	GAGGCACAGCTTGTTGAGG	″
Rras	4 (int.)	642	ACWCAGATCCTCMGRGTYAAGGA	GAGTTTGGCDGAKGCCTCRAAGTA	Rodríguez-Prieto et al., 2014
Sfrs5	1 (int.)	482	TCAAGGGTTACGGACGGATC	TCATCTGCATCCCTTGGGTC	Camacho-Sanchez et al., 2018
Smo	9 (int.)	456	GCCACCCTGCTCATCTGGAGGCG	GTTGGCRATCATCTTGCTYTTCTTGA	Rodríguez-Prieto et al., 2014
Ssfa2	13 (int.)	423	ACCCTCATATGACAGAGGAGG	ATTCGGACAGAGTTCCGCA	Camacho-Sanchez et al., 2018
Tmem87a	16 (int.)	447	CTGCTTGGTACTTCTCATTTTCA	TGTCAGAGGAAGATGARGAGGA	″
Trpv4	8 (int.)	401	TTACCRBACCACVGYGGACTACCT	GCTGGAAGGAGCCRTCGAYGAAGA	Rodríguez-Prieto et al., 2014
Usp20	17 (int.)	443	AACGTGATCAATGGGCAGTG	AGGAAGGTGTGGTTGGTGAT	Camacho-Sanchez et al., 2018
Wls	7 (int.)	506	AAYCACATYGCMGGSTAYTGGAA	GTCYGTKCCAACRTCYGTRGTCCA	Rodríguez-Prieto et al., 2014
Ghr	10 (ex.)	460	GGRAARTTRGAGGAGGTGAACACMATCTT	GTTGGTGGGTTGAYTCAGTTTC	Pagès et al., 2010
Ghr	10 (ex.)	436	GATCTCTTGTGCCTTGACCAG	TAAATGTCCTCCTGGTTAAAG	″
Ghr	10 (ex.)	310	CCTACTTCTGTGAGTCAGATGC	GATTTTGTTCAGTTGGTCTGTGCT	″
Rbp3	1 (ex.)	389	ATTGAGCAGGCTATGAAGAG	GGGATCCCAGAGACRTGRCC	Pagès et al., 2010; Fabre et al., 2013
Rbp3	1 (ex.)	450	TCCTTGGTGCTAGATCTCCG	TAGGGCTTGCTCTGCAGG	″
Rbp3	1 (ex.)	441	CAGACATGGGAAGGCAGTGG	GCAGGTAGCCCACATTGCC	″

rats (Wells et al., 2006). However, KNP species richness (ca. 60 species encompassing the Orders Eulipotyphla, Rodentia, and Scandentia: Nor, 2001; Phillips and Phillips, 2016) is much higher than in DNP (ca. 10 species encompassing the Orders Eulipotyphla, Rodentia, and Lagomorpha: Palomo et al., 2007). Hence, to get more comparable panels of taxa and avoid biases associated with taxonomic composition of either one or the other site, we employed only a subset of the species in KNP while complementing the DNP set with other Iberian species (Table 1). Field samples were collected according to the guidelines of the American Society of Mammalogists (Sikes and Gannon, 2016), as approved by institutional animal care and use committees (*Estación Biológica de Doñana* Proposal Numbers CGL 2010-21524 and 433/2016). Field work in KNP is detailed in Camacho-Sanchez et al. (2019) and was undertaken with permission from Sabah Parks (TS/PTD/5/4Jld. 45 (33) and TS/PTD/5/4 Jld. 47 (25)) the Economic Planning Unit (100-24/1/299), and the Sabah Biodiversity Council (JKM/MBS.1000-2/2 (104)). Biological samples were exported with permissions from the Sabah Wildlife Department (JHL.600-3/7 Jld.7/19 and JHL.600-3/7 Jld.8/) and Sabah Biodiversity Council (Ref: TK/PP:8/8Jld.2). Field work in DNP was undertaken with permission from the *Junta de Andalucía* (SGYB/AFR/DBP).

Table 3. Nuclear diversity across species and small mammal community. Number of individuals (n), number of loci typed (# loci), allele range per locus (# alleles), number of polymorphic sites (S), number of haplotypes (h), haplotype diversity (Θ), nucleotide diversity (π). Average values inherent to the intron loci (per-locus data are given in Appendix S1) are compared with those obtained for the Ghr exon. The upper table section identifies members of Rattini tribe. DNP: Doñana National Park; IB: Iberian Peninsula; KNP: Kinabalu National Park.

Study area	Species	n	# loci	# alleles	Introns				Ghr			
					S	h	Θ	π	S	h	Θ	π
DNP	*Rattus norvegicus*	6	39	1–6	2.2	1.7	0.2	0.002	0	1	0	0
DNP	*Rattus rattus*	2	38	1–3	3	1.6	0.3	0.004	1	2	1	0.003
KNP	*Rattus baluensis*	3	40	1–4	2.2	1.7	0.3	0.01	0	1	0	0
KNP	*Sundamys muelleri*	6	39	1–5	5.5	2.9	0.5	0.003				
KNP	*Niviventer cremoriventer*	4	32	1–5	3.7	2.2	0.4	0.03	1	2	0.7	0.001
KNP	*Leopoldamys sabanus*	5	39	1–6	7.6	3.3	0.6	0.05	0	1	0	0
KNP	*Lenothrix canus*	1	18	1–2	0.9	1.3	0.4	0.003	0	1	0	0
KNP	*Maxomys ochraceiventer*	1	30	1–2	0.7	1.2	0.2	0.002				
KNP	*Maxomys whiteheadi*	5	34	1–9	8.1	3.9	0.6	0.02				
DNP	*Mus spretus*	8	25	1–7	1.9	2	0.3	0.002	0	1	0	0
DNP	*Apodemus sylvaticus*	7	14	1–11	16.9	5.7	0.7	0.01	0	1	0	0
KNP	*Chiropodomys pusillus*	1	28	1	0	1	0	0.003	0	1	0	0
DNP	*Arvicola sapidus*	3	9	1–4	7.6	1.6	0.2	0.002				
IB	*Microtus cabrerae*	1	11	1–2	0.2	1.2	0.2	3.6×10^{-4}				
IB	*Chionomys nivalis*	3	9	1–4	2.8	2.1	0.4	0.001	1	2	0.5	0.01
DNP	*Eliomys quercinus*	5	3	1–3	4.7	2	0.5	4.7	3	4	0.9	0.001
IB	*Sciurus vulgaris*	3	9	1–2	0.3	1.2	0.1	4.44×10^{-4}	0	1	0	0
KNP	*Callosciurus prevostii*	2	6	1–3	3.7	1.8	0.4	0.01				
KNP	*Sundasciurus everetti*	2	9	1–2	0.9	1.3	0.2	0.002	0	1	0	0
KNP	*Sundasciurus jentinki*	1	6	1	0	1	0	0	0	1	0	0
KNP	*Sundasciurus lowii*	1	8	1–2	0.5	1.1	0.1	0.001	0	1	0	0
DNP	*Oryctolagus cuniculus*	2	13	1–2	2.6	1.3	0.2	0.003	5	2	1	0.01
KNP	*Tupaia montana*	2	4	1–2	11.8	1.8	0.5	0.05	1	2	0.7	0.001
KNP	*Hylomys suillus*	2	2	1–2	0.5	1.5	0.25	5×10^{-4}	0	1	0	0
DNP	*Suncus etruscus*	2	8	1	0	1	0	0	0	1	0	0
DNP	*Crocidura russula*	2	8	1–4	1.1	1.6	0.3	0.002	0	1	0	0
DNP	*Crocidura suaveolens*	2	7	1–2	0.6	1.1	0.1	0.001	0	1	0	0

Samples were also obtained from the scientific collection at *Estación Biológica de Doñana*.

4.2. Loci selection and primer testing

A panel of 40 intronic loci was selected including 30 from among those identified as single-copy and potentially informative in Igea et al. (2010) for addressing the phylogeny of closely related mammals and developed in Camacho-Sanchez et al. (2018), plus another 10 developed by Rodríguez-Prieto et al. (2014) specifically for rodent phylogeny (Table 2, Table S1). The 30 loci were successfully used to study intraspecific phylogeographic patterns in the summit rat (*Rattus baluensis*; Camacho-Sanchez et al., 2018). Here we describe in detail the process of primer design so it can be applied in other non-model organisms. The brown rat was used for primer design due to the large availability of genomic resources for this model organism as well as its occurrence among the focal taxa of the present study. After inspecting the species tree available at http://www.ensembl.org/info/about/speciestree.html, the house mouse (*Mus musculus*) was found to be the closest species to our target with a fully annotated genome. In previous studies (Igea et al., 2010; Rodríguez-Prieto et al., 2014: see Table S1 for details), the target introns had been selected via multiple filtering steps including long flanking exonic regions for primer design and appropriate levels of size conservation and variation within mammals. Loci selection for the present study relied on a size criterion, namely that introns of choice were ideally between 300-400 bp long as to be completely sequenced in Roche 454 by pyrosequencing chemistry (upper read length limit of about 500 bp) and Illumina MiSeq 300PE. Lower length limit was established to ensure a satisfactory information content and because fragment size homogeneity should increase sequencing efficacy. Human gene IDs corresponding to the introns selected were obtained and the BioMart tool (Smedley et al., 2009) in Ensembl (Herrero et al., 2016) was used to find the respective orthologues in *R. norvegicus* (Rnor 5.0) and *M. musculus* (GRCm38). These sets of genes were downloaded into Geneious 8.1.5 (http://www.geneious.com: Kearse et al., 2012) using the NCBI (National Center for Biotechnology Information) plugin and aligned using the MAFFT plugin 7.453 (Katoh and Standley, 2013) with default parameters. The annotations of aligned orthologues were used to manually check for intron number and size as well as the homology of flanking regions when compared to those of the sequences from Igea et al. (2010) and Rodríguez-Prieto et al. (2014). Exonic regions complying with length and conservation criteria were targeted for primer design in Primer3 (http://primer3.ut.ee/: Koressaar and Remm, 2007; Untergasser et al., 2012). Since the panel was intended to be suitable for multiplexing in a single PCR reaction, primer pairs for the loci selected by Igea et al. (2010) were designed to have a narrow delta Tm ($< 3\ ^{\circ}C$) and roughly the same size (ca. 20–25 bp: PCR product size 400 to 500 bp). The primers for the loci from Rodríguez-Prieto et al. (2014) were used without modification. A 2 bp GC clamp was added to enhance primer stability, while for all the other parameters default values were applied. When Primer3 failed to retrieve solutions complying with desired features, either stringency criteria were relaxed and GC clamps removed or new exonic regions were selected for primer design. Candidate primer pairs were mapped against their homologous regions in Geneious alignments and those mapping in the most conserved regions were finally selected. If mismatches at the 3′ end occurred, either the respective region was excluded or ambiguities were introduced manually to a maximum of two per primer (Table 2, Table S1). The whole workflow is illustrated in Figure S1. The primers for

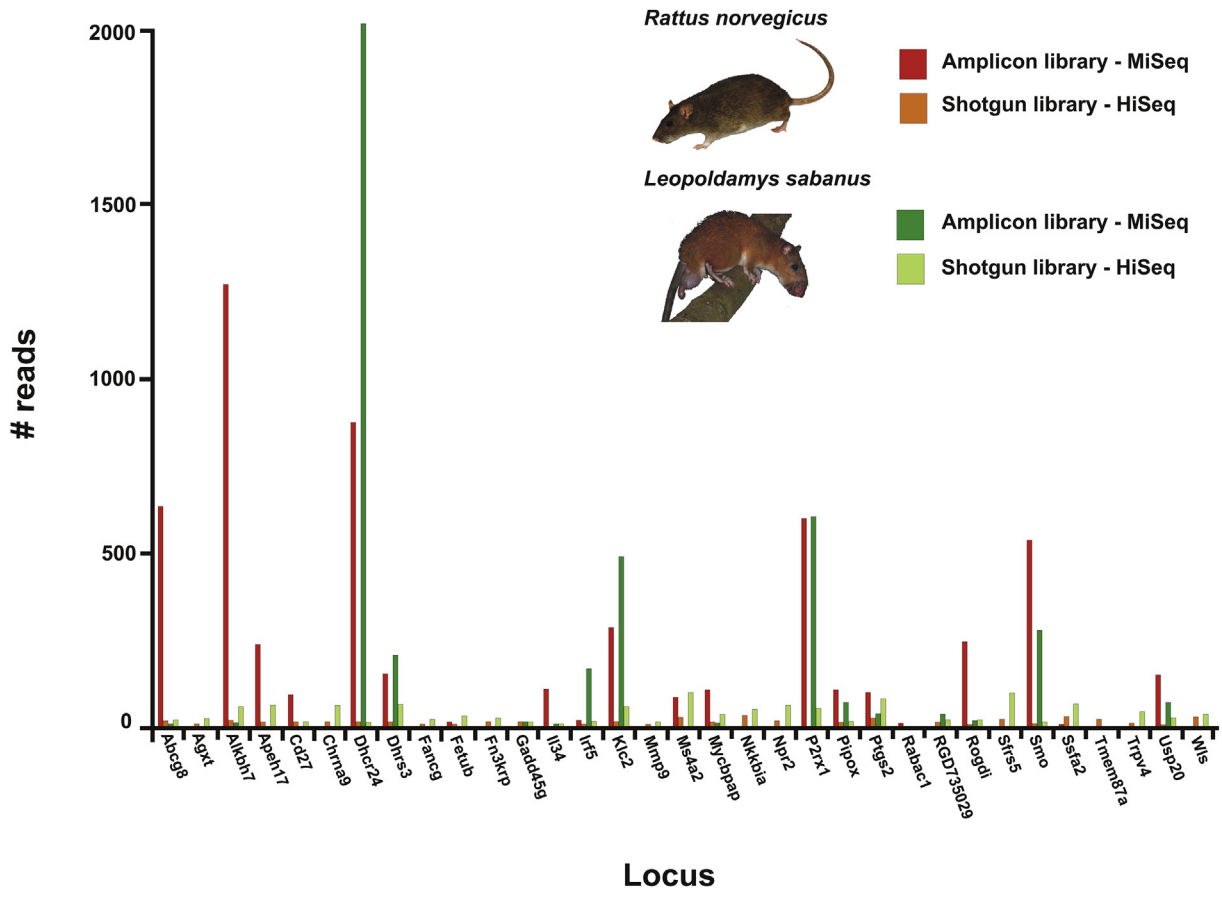

Figure 2. Histogram showing the high variance associated with read depth across different loci in *R. norvegicus* and *L. sabanus* (one PCR per species).

the loci from Rodríguez-Prieto et al. (2014) were used without modification. Additionally, two commonly sequenced nuclear genes, Rbp3 and Ghr (Table S1), were also sequenced by amplified overlapping fragments for the purpose of compared phylogeny and variability (e.g., Pagès et al. 2010; Fabre et al., 2013).

4.3. DNA extraction

DNA was isolated with phenol-chloroform and ethanol precipitation (Maniatis et al., 1985) or with SeraMag™ SpeedBeads (Thermoscientific). DNA samples were quantified by Nanodrop spectrophotometry relying on absorbance readings at 260 nm wavelength and diluted to working concentrations of 15 ng/μl.

4.4. 454 amplicon library preparation (PCR) and sequencing

We prepared amplicon libraries following a two-step PCR. The 40 selected primer pairs were equimolar in a single PCR. Reactions included 1x Multiplex PCR Master Mix (Qiagen), 3.34 μM of primer mix and 20–50 ng of template DNA (final volume 25 μl). Primers had an M13 tail on their 5′ end (fw: 5′-GTTTTCCCAGTCACGAC; rev: 5′-AACAGCTATGACCATG). The thermal profile was 95 °C for 15 min followed by 15 cycles of touchdown: 95 °C for 30 s, 65-60 °C for 30 s, 72 °C for 30 s, and then 15 more cycles with the annealing temperature at 60 °C with a final extension at 72 °C for 10 min. PCR products were then cleaned with SPRI beads (Rohland and Reich, 2012) and diluted 1:10 before the indexing PCR. Reaction mix consisted of 0.1 μM of each Multiplexing Index sequence plus the sequencing primers, 1x PCR Gold buffer, 1.5 mM $MgCl_2$, 200 μM of each dNTP, 0.2 U AmpliTaq Gold DNA Polymerase (Applied Biosystems) and 2 μl of diluted PCR product (final volume 12 μl). PCR conditions were: 98 °C for 30 s and then 25 cycles of 98 °C for 10 s, 56 °C for 20 s, 72 °C for 45 s. PCR products were checked on a 2% agarose gel and viewed with Quantity-One software (Bio-Rad Laboratories) for relative quantification of 300–650 bp-long fragments prior to equimolar pooling. Products were cleaned with 1x SPRI beads and quantified with a fluorometer. Finally, an emulsion PCR was carried out with the Roche emPCR-a kit prior to sequencing on a GS Junior 454 sequencer following manufacturer's instructions.

4.5. Illumina amplicon library preparation (PCR) and sequencing

Library was amplified in 20 μL reaction of 1x Phusion Master Mix (New England Biolabs), 0.05 μM of each primer, and 20–50 ng of template DNA. Primers had a tail (fw: 5′-TCTTTCCCTACACGACGCTCTTCCGATCT; rev: 5′- GAGTTCAGACGTGTGCTCTTCCGATCT) complementary to indexing primers. PCR was run as follows: 98 °C for 1 min then 25 cycles of 98 °C for 10 s, 61 °C for 30 s, 72 °C for 45 s with a final denaturation at 95 °C for 3 min. PCR products were checked and cleaned with SPRI beads as above. The indexing PCR was performed in a final volume of 12 μl including 0.42 μM of each Multiplexing Index sequence plus the sequencing primers, 1x Phusion Master Mix, and 1 μl of undiluted PCR product. Thermal profile was 98 °C for 30 s, the 12 cycles of 98 °C for 10 s, 60 °C for 20 s, 72 °C for 45 s and a final denaturation at 95 °C for 3 min. Indexed PCR product were cleaned, quantified and pooled at equimolar ratios as above before sequencing of 300 bp paired end

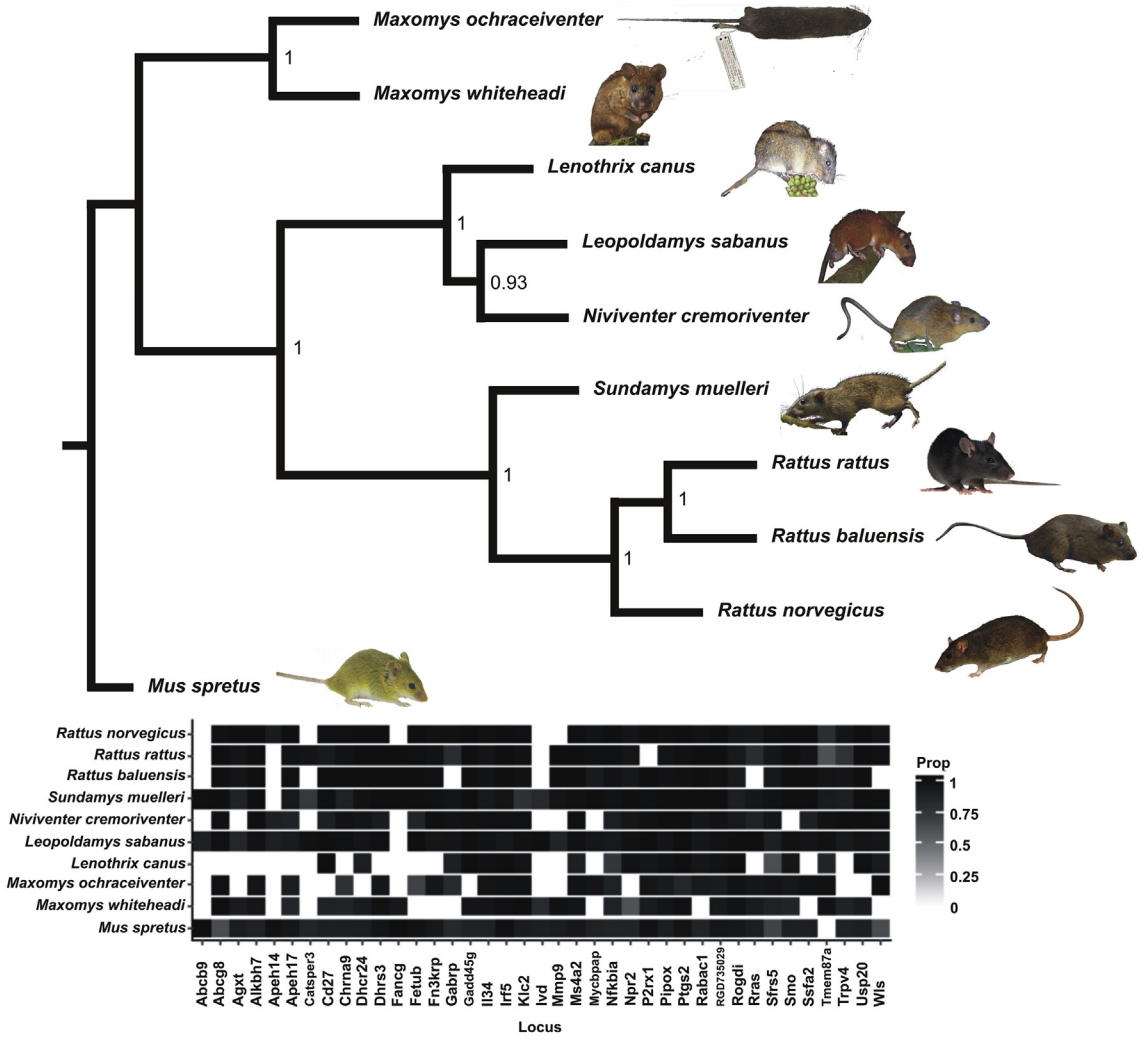

Figure 3. ML tree of Rattini tribe based on the whole set of introns amplified in the representatives tested in this study and built using the coalescent approach implemented in ASTRAL. The plot at the bottom of the figure indicates the loci amplified in each taxon and their completeness (Prop.). Animal photos are not to scale; for credits see the Acknowledgements and Appendix S2.

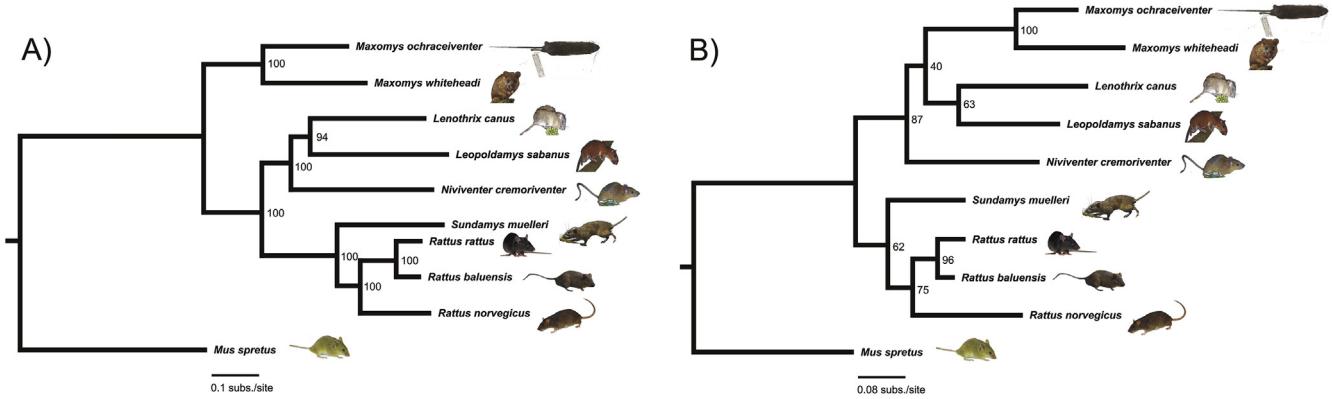

Figure 4. MtDNA-based phylogenies of Rattini representatives sampled in this study. A. ML tree reconstructed on the basis of the whole mitogenome (10,852 bp). B. ML tree reconstructed on the basis of the cytochrome *b* gene.

reads on the Illumina MiSeq platform at the Johns Hopkins University Genetic Resources Core Facility (Baltimore, MD, USA).

4.6. Illumina shotgun library preparation and sequencing

Aliquots of DNA extracts were diluted to 20 ng/μl in sonication buffer (10 mM Tris, 1 mM EDTA pH 7.5–8) to 100 μl and sheared in an ultrasonic bath (Bioruptor UCD-20 0TM-EX Sonication System). Sonication cycle conditions were adjusted to the desired target size (around 500 bp): two rounds of 3 cycles of 30 s on High (H)/30 s off with a spin in the middle. Libraries were prepared with the Illumina Kapa Library Preparation Kit (Kapa Biosystems, Wilmington, Massachusetts, US) using ¼ reactions. Cleaning steps were done with SPRI beads as above. We used 10 μM Y-adapters (5′-5Phos-GATCGGAAGAGCACACGTCTGAACTCCAGTCAC and 5′-ACACTCTTTCCCTACACGACGCTCTTCCGATC*T; 5Phos, 5′ phosphorylation; *, phosphorothioate bond) previously annealed as in (Meyer and Kircher, 2010). The indexing PCR was done using a double indexing strategy as in (Kircher et al., 2012), with the PCR thermal conditions: 98 °C for 45 s then 8 cycles of 98 °C for 15 s, 60 °C for 30 s, 72 °C for 60 s with a final extension at 72 °C for 5 min. The libraries were cleaned, quantified, pooled and enriched for introns and mitochondrial DNA in independent hybridization reactions. We prepared the nuclear probes from PCR products using *R. norvegicus* as template for all the species tested. Each locus was amplified separately, pooled at equimolar ratios and cleaned. Mitogenomes were amplified by long-range PCR in two overlapping fragments of approximately 7 kb and 9 kb using primers from Sasaki et al. (2005) and sheared. Further details on primer and adapter sequences are reported in Camacho-Sanchez et al. (2017). In this case, the target regions across our sample were enriched with probes obtained from the same or related species. Specifically, we used *R. norvegicus*, *S. vulgaris* and *C. russula* PCR products as template for Muridae-Cricetidae, Sciuridae-Gliridae and Soricidae-Erinaceidae representatives, respectively. For Leporidae and Tupaiidae we used *O. cuniculus* and *T. montana* templates. Then, the nuclear and mitochondrial baiting molecules were biotinylated as in Maricic et al. (2010) and the enrichment was done accordingly, with modifications as in Camacho-Sanchez et al. (2017). The enriched libraries were quantified by qPCR and were subsequently re-amplified as above except for the initial denaturation at 98 °C for 45 s and 15 or 20 PCR cycles depending on their concentration. The sequences of PCR indexing primers, indexing oligos and reamplification primers are as in Camacho-Sanchez et al. (2017). The re-amplified libraries were cleaned and quantified using qPCR. Final libraries were pooled in equimolar ratios and sequenced with 100 bp paired end reads on the Illumina HiSeq 2500 platform at the Johns Hopkins University Genetic Resources Core Facility.

4.7. Intron genotyping - amplicon libraries

Size filtering, read demultiplexing, primer trimming, amplicon assignment and allele calling of FASTQ data obtained from amplicon libraries (454 and Illumina MiSeq sequencing platforms) were performed separately with AmpliSAS in AmpliSAT, a package specifically for automated amplicon assignment from NGS data (Sebastian et al., 2015). The tool AmpliCHECK was first used to perform exploratory assays and detect likely artefacts due to PCR or sequencing errors by using default analysis parameters for 454 and Illumina data, respectively. The lack of prior information on intron length in most of the surveyed taxa and the low number of individuals analyzed for each taxon did not allow adopting highly restrictive criteria for artefact detection at this step of the pipeline. However, the alleles labelled as putative errors in AmpliCHECK after read quality filtering were removed from further analyses. Hence, the tool AmpliSAS was used with default filtering and clustering parameters applying to each sequencing technology, while selecting 2 as both maximum number of alleles and minimum amplicon depth, and fixing 5000 as maximum number of reads per amplicon.

4.8. Intron genotyping - shotgun libraries

Adapter sequences were removed with cutadapt 1.8.3 (Martin, 2011) and output files were imported into Geneious. Paired end reads were iteratively mapped (5 cycles) with medium-low sensitivity against species-specific references represented by intron sequences obtained in the same taxon by means of amplicon libraries. When a given locus was not available, homologous sequences of the closest relative available in Ensembl were used as reference (Table S3). Consensus sequence callings were performed with a 75% threshold and two read minimum. BAM files generated in Geneious were processed with SAMtools 0.1.18 (Li et al., 2009) to remove PCR duplicates. Once re-imported in Geneious, BAM files were newly inspected and consensus sequences finally obtained by applying the same parameters as above. Intron sequences retrieved from amplicon and shotgun libraries were aligned in Geneious with MAFFT plugin for each locus and species. The alignment was carefully inspected to check for mismatches between data obtained with different sequencing platforms. When more than two alleles from the same individual were retrieved across amplicon and shotgun libraries, these were visually inspected to evaluate their possibly undetected artifactual nature and purged from multiple alignments of orthologous intronic sequences. When less than five reads where available for shotgun data, polymorphisms were confirmed only if they occurred in more than one individual and/or matched amplicon data from the same individual. However, we scored a locus as having worked in a given species (e.g., being amplified) when at least one read mapped properly to the reference.

4.9. Mitogenome assembly

FASTQ files with adapters trimmed (see above) were mapped to the phylogenetically closest available mitogenome reference in GenBank (Table S4) using the mapping tool in Geneious with medium-low sensitivity and 5 iterations. We assembled one mitogenome per species. Consensus sequence callings were performed with over 75% threshold and two reads as minimum. The mapping files were exported in BAM format and the PCR duplicates removed using SAMtools, then reimported into Geneious.

4.10. Locus amplification and variability

The alleles identified as putative errors in AmpliCHECK after read quality filtering were removed from further analyses. Likewise, alleles found at a frequency below 25% in a given individual were conservatively discarded as possibly arising from undetected sequencing errors This threshold is consistent with that proposed in recent works (e.g., O′Leary et al., 2018). The number of alleles per locus and taxon were scored to evaluate the polymorphism of loci tested.

4.11. Indices of diversity

For each species we computed number of polymorphic sites (S), number of haplotypes (h), haplotype diversity (Θ), and nucleotide diversity (π) at each locus with DnaSP 5.10.1 (Librado and Rozas, 2009).

4.12. Phylogenetic reconstructions

A phylogeny of Rattini was built relying on all intron loci yielding sequences in members of this group. We retained the 9 species from Rattini to maximize the number of loci with homologous sequences. For each species we selected the most common allele. We completed the data for *R. baluensis* with sequences from GenBank: MG424797 (Mmp9), MG425076 (Npr2), MG425817 (Tmem87a) (Camacho-Sanchez et al.,

2018). We added *Mus spretus* as outgroup using the sequences generated in this study and completing the dataset with loci from the genome assembly *M. spretus* SPRET_EiJ_v1 (www.ensembl.org), except for the locus Sfrf5, which was taken from the genome assembly of *M. musculus* GRCm38.p6 (www.ensembl.org). We excluded Nadsyn1 because of its fragmentary data from enrichment. The sequences for each locus were aligned using MAFFT. These were first filtered with Divvier 1.0 (Ali et al., 2019) to remove characters from columns with a coverage below 3 and, secondly, with TrimAI 1.4.1 (Capella-Gutiérrez et al., 2009) to remove sequences for which 40 % or more had coverage below 50 % of their total length as well as columns with gaps in more than 60% of the sites. The search for best scoring Maximum Likelihood (ML) tree was done in RAxML using the GTR + G model of sequence evolution, the rapid bootstrap analysis and alternative runs starting on 100 different trees. The locus Abcb9 was discarded since, after filtering steps, was present in only three species. A total of 38 gene trees (all except Nadsyn1 and Abcb9) from the introns were used in ASTRAL 5.6.3 (Mirarab and Warnow, 2015) to produce a single coalescent-based species tree.

We reconstructed the phylogenetic relationships among the same set of species with mitogenomes. We used the annotation module in MitoZ (Meng et al., 2019) to automatically generate annotations for all mitogenomes. All 37 mitochondrial genes (13 protein-coding genes, 2 rRNA genes and 22 tRNA genes) were successfully annotated in the mitogenomes of all species. Sequences were aligned on a per-gene basis using MAFFT. For *Lenothrix canus*, we used the already published mitogenome KY464180 (Camacho-Sanchez et al., 2017). Gblocks (Castresana, 2000) was used to trim gappy positions with the codon flag on for the protein-coding genes. The correct coding frame for all protein coding genes was confirmed in www.ebi.ac.uk/Tools/st/emboss_transeq. Alignments from protein-coding genes were split into codon positions 1, 2 and 3, using AMAS (Borowiec, 2016). All alignments were concatenated with the same software. The complete alignment contained 10,852 positions with 0.1% of ambiguous positions. The best partition scheme for the resulting 48 partitions was determined with PartitionFinder 2.1.1 (Lanfear et al., 2012). Phylogenetic reconstructions were performed in a ML framework with RAxML 8.0.0 (Stamatakis, 2014). We followed the same procedure to reconstruct a phylogenetic tree with only cyt *b* for the same taxa.

Declarations

Author contribution statement

Giovanni Forcina: Analyzed and interpreted the data; Wrote the paper.

Miguel Camacho-Sanchez: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data.

Fred Y. Y. Tuh: Contributed reagents, materials, analysis tools or data.

Sacramento Moreno: Contributed reagents, materials, analysis tools or data.

Jennifer A. Leonard: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data.

Funding statement

This work was supported by the former Spanish Ministry of Science and Innovation (now Spanish Ministry of Economy, Industry and Competition) grants CGL2010-21524, CGL2014-58793-P and CGL2017-86068-P, LifeWatch, and a “Centre of Excellence Severo Ochoa” award to EBD-CSIC (SEV-2012-0262).

Data availability statement

DNA sequences from complete mitochondrial genomes and nuclear loci have been deposited in GenBank with accession numbers MW209719-MW209719 and MW394671-MW396361, respectively. Details on phylogenetic methods and phylogenetic trees are available at github.com/csmiguel/community-genetics and at https://doi.org/10.5281/zenodo.4275480.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2020.e05583.

Acknowledgements

Alvaro Sebastian is acknowledged for assistance in usage of AmpliSAT package, Carles Vilà for helping with Phyton scripts and Irene Quintanilla with Anna Cornellas for their highly qualified laboratory work. Logistical support was provided by the Laboratory of Molecular Ecology (LEM-EBD), the Laboratory of GIS and Remote Sensing (LAST-EBD) and by the infrastructures offered by Doñana's Singular Scientific-Technical Infrastructure (ICTS-EBD). We thank the Malaysian institutions that allowed us to do fieldwork and export the samples: the Sabah Biodiversity Centre for issuing a research and export permit, Sabah Parks for research and collection permits, as well as support and cooperation during our time in Malaysia, and the Sabah Wildlife Department and the Economic Planning Unit for research and export permits. Materials were provided by the scientific collection at the *Estación Biológica de Doñana*. The authors are also grateful to Melissa Hawkins for her logistic assistance with fieldwork and to a list of nature photographers for granting permission to use their photos (when these were not available in open access media repositories) of the taxa surveyed in this study (see Appendix S2 for details). Special thanks go also to the authors of the following software for allowing to use their logos in Figure S1: Paul Flicek (Ensembl: https://www.ensembl.org/index.html), Steve Rozen, Andreas Untergasser, Maido Remm (Primer3web: http://primer3.ut.ee/), Alvaro Sebastian (AmpliSAS: http://evobiolab.biol.amu.edu.pl/amplisat/index.php?amplisas), and Claire Westwood (Geneious https://www.geneious.com/).

References

Ali, R.H., Bogusz, M., Whelan, S., 2019. Identifying clusters of high confidence homologies in multiple sequence. Mol. Biol. Evol. 36, 2340–2351.

Aschwanden, J., Holzgang, O., Jenni, L., 2007. Importance of ecological compensation areas for small mammals in intensively farmed areas. Wildl. Biol. 13, 150–158.

Baker, C.S., Lento, G.M., Cipriano, F., Palumbi, S.R., 2000. Predicted decline of protected whales based on molecular genetic monitoring of Japanese and Korean markets. Proc. R. Soc. B 267, 1191–1199.

Bazin, E., Glémin, S., Galtier, N., 2006. Population size does not influence mitochondrial genetic diversity in animals. Science 312, 570–572.

Borowiec, M.L., 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. PeerJ 4, e1660.

Byrom, A.E., Ruscoe, W.A., Nkwabi, A.K., Metzger, K.L., Forrester, G.J., Craft, M.E., Durant, S.M., Makacha, S., Bukombe, J., Mchetto, J., Mduma, S.A.R., Reed, D.N., Hampson, K., Sinclair, A.R.E., 2009. Small mammal diversity and population dynamics in the Greater Serengeti ecosystem. In: Sinclair, A.R.E. (Ed.), Serengeti IV: Sustaining Biodiversity in a Coupled Human–Natural System. University of Chicago Press, Chicago, pp. 323–357.

Camacho-Sanchez, M., Leonard, J.A., Fitriana, Y., Tilak, M.K., Fabre, P.H., 2017. The generic status of *Rattus annandalei* (Bonhote, 1903) (Rodentia, Murinae) and its evolutionary implications. J. Mammal. 98, 1340–1355.

Camacho-Sanchez, M., Quintanilla, I., Hawkins, M.T.R., Tuh, F.Y.Y., Wells, K., Maldonado, J.E., Leonard, J.A., 2018. Interglacial refugia on tropical mountains: novel insights from the summit rat (*Rattus baluensis*), a Borneo mountain endemic. Divers. Distrib. 24, 1252–1266.

Camacho-Sanchez, M., Hawkins, M.T.R., Tuh, Y.Y.F., Maldonado, J.E., Leonard, J.A., 2019. Endemism and diversity of small mammals along two neighboring Bornean mountains. PeerJ 7, e7858.

Camacho-Sanchez, M., Velo-Antón, G., Hanson, J.O., Veríssimo, A., Martínez-Solano, I., Marques, A., Moritz, C., Carvalho, S., 2020. Comparative assessment of range-wide patterns of genetic diversity and structure with SNPs and microsatellites: a case study with Iberian amphibians. Ecol. Evol. 1–11.

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.
Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540–552.
Cooper, G.M., Brudno, M., Stone, E.A., Dubchak, I., Batzoglou, S., Sidow, A., 2004. Characterization of evolutionary rates and constraints in three Mammalian genomes. Genome Res. 14, 539-548.
Cordero, D., Peña, J.B., Saavedra, C., 2014. Phylogeographic analysis of introns and mitochondrial DNA in the clam *Ruditapes decussatus* uncovers the effects of Pleistocene glaciations and endogenous barriers to gene flow. Mol. Phylogenet. Evol. 71, 274–287.
de Groot, G.A., Nowak, C., Skrbinšek, T., Andersen, L.W., Aspi, J., Fumagalli, L., Godinho, R., Harms, V., Jansman, H.A., Liberg, O., Marucco, F., Mysłajek, R.W., Nowak, S., Pilot, M., Randi, E., Reinhardt, I., Śmietana, W., Szewczyk, M., Taberlet, P., Vilà, C., Muñoz-Fuentes, V., 2016. Decades of population genetic research reveal the need for harmonization of molecular markers: the grey wolf *Canis lupus* as a case study. Mamm Rev. 46, 44–59.
Dos Reis, M., Inoue, J., Hasegawa, M., Asher, R.J., Donoghue, P.C.J., Yang, Z.H., 2012. Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. Proc. R Soc. B 279, 3491–3500.
Dufresnes, C., Brelsford, A., Béziers, P., Perrin, N., 2014. Stronger transferability but lower variability in transcriptomic- than in anonymous microsatellites: evidence from Hylid frogs. Mol. Ecol. Resour. 14, 716–725.
Fabre, P.-H., Pagès, M., Musser, G.G., Fitriana, Y.S., Fjeldså, J., Jennings, A., Jønsson, K.A., Kennedy, J., Michaux, J., Semiadi, G., Supriatna, N., Helgen, K.M., 2013. A new genus of rodent from Wallacea (Rodentia: Muridae: Murinae: Rattini), and its implication for biogeography and Indo-Pacific Rattini systematics. Zool. J. Linn. Soc. 169, 408–447.
Fischer, M.C., Rellstab, C., Leuzinger, M., Roumet, M., Gugerli, F., Shimizu, K.K., Holderegger, R., Widmer, A., 2017. Estimating genomic diversity and population differentiation – an empirical comparison of microsatellite and SNP variation in *Arabidopsis halleri*. BMC Genom. 18, 69.
Foley, N.M., Springer, M.S., Teeling, E.C., 2016. Mammal madness: is the mammal tree of life not yet resolved? Phil. Trans. R. Soc. B 371, 20150140.
Forcina, G., Leonard, J.A., 2020. Tools for monitoring genetic diversity in mammals: past, present and future. In: Ortega, J., Maldonado, J.E. (Eds.), Conservation Genetics in Mammals: Integrative Research Using Novel Approaches. Springer, Cham, pp. 13–27.
Frankham, R., 2002. Predicting extinction risk. Nature 419, 18–19.
Fuentes-Pardo, A.P., Ruzzante, D.E., 2017. Whole-genome sequencing approaches for conservation biology: advantages, limitations and practical recommendations. Mol. Ecol. 26, 5369–5406.
Haasl, R.J., Payseur, B.A., 2011. Multi-locus inference of population structure: a comparison between single nucleotide polymorphisms and microsatellites. Heredity 106, 158–171.
Hailer, F., Kutschera, V.E., Hallström, B.M., Klassert, D., Fain, S.R., Leonard, J.A., Arnason, U., Janke, A., 2012. Nuclear genomic sequences reveal that polar bears are an old and distinct bear lineage. Science 336, 344–347.
Heled, J., Drummond, A.J., 2009. Bayesian inference of species trees from multilocus data. Mol. Biol. Evol. 27, 570–580.
Herrero, J., Muffato, M., Beal, K., Fitzgerald, S., Gordon, L., Pignatelli, M., Vilella, A.J., Searle, S.M., Amode, R., Brent, S., Spooner, W., Kulesha, E., Yates, A., Flicek, P., 2016. Ensembl comparative genomics resources. Database (Oxford) 2016, bav096. Erratum in: Database (Oxford), 2016; 2016. pii: baw053.
Hoelzel, A.R., Bruford, M.W., Fleischer, R.C., 2019. Conservation of adaptive potential and functional diversity. Conserv. Genet. 20, 1–5.
Hope, A.G., Waltari, E., Morse, N.R., Flamme, M.J., Cook, J.A., Talbot, S., 2017. Small mammals as indicators of climate, biodiversity, and ecosystem change. Alaska Park Sci. 16, 72–78.
Igea, J., Juste, J., Castresana, J., 2010. Novel intron markers to study the phylogeny of closely related mammalian species. BMC Evol. Biol. 10, 369.
Igea, J., Aymerich, P., Fernández-González, A., González-Esteban, J., Gómez, A., Alonso, R., Gosálbez, J., Castresana, J., 2013. Phylogeography and postglacial expansion of the endangered semi-aquatic mammal *Galemys pyrenaicus*. BMC Evol. Biol. 13, 115.
Jiang, Z., Priat, C., Galibert, F., 1998. Traced orthologous amplified sequence tags (TOASTs)and mammalian comparative maps. Mamm. Genome 9, 577–587.
Katoh, K., Standley, D.M., 2013. MAFFT Multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649.
Kimura, Y., Hawkins, M., McDonough, M., Jacobs, L.L., Flynn, L.J., 2015. Corrected placement of *Mus-Rattus* fossil calibration forces precision in the molecular tree of rodents. Sci. Rep. 5, 14444.
Kircher, M., Sawyer, S., Meyer, M., 2012. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Res. 40, e3.
Koblmüller, S., Wayne, R.K., Leonard, J.A., 2012. Impact of Quaternary climatic changes and interspecific competition on the demographic history of a highly mobile generalist carnivore, the coyote. Biol. Lett. 8, 644–647.
Koblmüller, S., Vilà, C., Lorente-Galdos, B., Dabad, M., Ramirez, O., Marques-Bonet, T., Wayne, R.K., Leonard, J.A., 2016. Whole mitochondrial genomes illuminate ancient intercontinental dispersals of grey wolves (*Canis lupus*). J. Biogeogr. 43, 1728–1738.
Koressaar, T., Remm, M., 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23, 1289–1291.
Kremer, A., Ronce, O., Robledo-Arnuncio, J.J., Guillaume, F., Bohrer, G., Nathan, R., Bridle, J.R., Gomulkiewicz, R., Klein, E.K., Ritland, K., Kuparinen, A., Gerber, S., Schueler, S., 2012. Long-distance gene flow and adaptation of forest trees to rapid climate change. Ecol. Lett. 15, 378–392.
Kuchta, S.R., Brown, A.D., Converse, P.E., Highton, R., 2016. Multilocus phylogeography and species delimitation in the Cumberland Plateau Salamander *Plethodon kentucki*: incongruence among data sets and methods. PloS One 11, e0150022.
Laikre, L., 2010. Genetic diversity is overlooked in international conservation policy implementation. Conserv. Genet. 11, 349–354.
Lanfear, R., Calcott, B., Ho, S.Y., Guindon, S., 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29, 1695–1701.
Larsen, P.A., Matocq, M.D., 2019. Emerging genomic applications in mammalian ecology, evolution, and conservation. J. Mammal. 100, 786–801.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. 1000 Genome Project Data Processing Subgroup, The Sequence Alignment/Map format and SAMtools,. Bioinformatics 25, 2078–2079.
Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452.
Lyons, L.A., Laughlin, T.F., Copeland, N.G., Jenkins, N.A., Womack, J.E., O'Brien, S.J., 1997. Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. Nat. Genet. 15, 47–56.
Maniatis, T., Fritsch, E.F., Sambrook, J., Engel, J., 1985. Molecular cloning? A laboratory manual. New York Cold Spring Harbor Laboratory. 1982, 545 S., 42 Volume 5, Issue 1. Acta Biotechnol. 5, 104.
Maricic, T., Whitten, M., Pääbo, S., 2010. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. PLoS One 5, e14004.
Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10–12.
Meng, G., Li, Y., Yang, C., Liu, S., 2019. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. Nucleic Acids Res. 47, e63.
Merritt, J.F., 2010. The Biology of Small Mammals. Johns Hopkins University Press, Baltimore.
Meyer, M., Kircher, M., 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. 5 pdb.prot5448-pdb.prot5448.
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.
Moran, P., Teel, D.J., LaHood, E.S., Drake, J., Kalinowski, S., 2006. Standardising multi-laboratory microsatellite data in Pacific salmon: an historical view of the future. Ecol. Freshw. Fish 15, 597–605.
Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–858.
Nabholz, B., Mauffrey, J.F., Bazin, E., Galtier, N., Glemin, S., 2008. Determination of mitochondrial genetic diversity in mammals. Genetics 178, 351–361.
Nor, S.M., 2001. Elevational diversity patterns of small mammals on Mount Kinabalu, Sabah, Malaysia. Global Ecol. Biogeogr. 10, 41–62.
O'Leary, S.J., Puritz, J.B., Willis, S.C., Hollenbeck, C.M., Portnoy, D.S., 2018. These aren't the loci you're looking for: principles of effective SNP filtering for molecular ecologists. Mol. Ecol. 27, 1–14.
Ostfeld, R.S., Jones, C.G., Wolff, J.O., 1996. Of mice and mast. Bioscience 46, 323–330.
O'Grady, J.J., Brook, B.W., Reed, D.H., Ballou, J.D., Tonkyn, D.W., Frankham, R., 2006. Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. Biol. Conserv. 133, 42–51.
Pagès, M., Chaval, Y., Herbreteau, V., Waengsothorn, S., Cosson, J.F., Hugot, J.P., Morand, S., Michaux, J., 2010. Revisiting the taxonomy of the Rattini tribe: a phylogeny-based delimitation of species boundaries. BMC Evol. Biol. 10, 184.
Palomo, L.J., Gisbert, J., Blanco, J.C., 2007. Atlas y Libro Rojo de los Mamíferos Terrestres de España (in Spanish). Direccion General para la Biodiversidad-SECEM-SECEMU, Madrid.
Phillips, Q., Phillips, K., 2016. Mammals of Borneo and Their Ecology: Sabah, Sarawak, Brunei and Kalimantan. Natural History Publications, Kota Kinabalu.
Pinsky, M.L., Palumbi, S.R., 2013. Meta-analysis reveals lower genetic diversity in overfished populations. Mol. Ecol. 23, 29–39.
Pons, J.-M., Olioso, G., Cruaud, C., Fuchs, J., 2010. Phylogeography of the Eurasian green woodpecker (*Picus viridis*). J. Biogeogr. 38, 311–325.
Putman, A.I., Carbone, I., 2014. Challenges in analysis and interpretation of microsatellite data for population genetic studies. Ecol. Evol. 4, 4399–4428.
Razgour, O., Forester, B., Taggart, J.B., Bekaert, M., Juste, J., Ibáñez, C., Puechmaille, S.J., Novella-Fernandez, R., Alberdi, A., Manel, S., 2019. Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. Proc. Natl. Acad. Sci. U.S.A. 116, 10418–10423.
Rodríguez-Prieto, A., Igea, J., Castresana, J., 2014. Development of rapidly evolving intron markers to estimate multilocus species trees of rodents. PloS One 9, e96032.
Rohland, N., Reich, D., 2012. Cost-effective high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res. 22, 939–946.
Rowe, R.J., Terry, R.C., 2014. Small mammal responses to environmental change: integrating past and present dynamics. J. Mammal. 95, 1157–1174.
Rowe, K.C., Achmadi, A.S., Fabre, P.-H., Schenk, J.J., Steppan, S.J., Esselstyn, J.A., 2019. Oceanic islands of Wallacea as a source for dispersal and diversification of murine rodents. J. Biogeogr. 46, 2752–2768.
Sanchez-Donoso, I., Huisman, J., Echegaray, J., Puigcerver, M., Rodríguez-Teijeiro, J.D., Hailer, F., Vilá, C., 2014. Detecting slow introgression of invasive alleles in an extensively restocked game bird. Front. Ecol. Evol. 2, 00015.

Sasaki, T., Nikaido, M., Hamilton, H., Goto, M., Kato, H., Kanda, N., Pastene, L., Cao, Y., Fordyce, R., Hasegawa, M., Okada, N., 2005. Mitochondrial phylogenetics and evolution of mysticete whales. Syst. Biol. 54, 77–90.

Sebastian, A., Herdegen, M., Migalska, M., Radwan, J., 2015. amplisas: a web server for multilocus genotyping using next-generation amplicon sequencing data. Mol. Ecol. Resour. 16, 498–510.

Seddon, N., Mace, G.M., Naeem, S., Tobias, J.A., Pigot, A.L., Cavanagh, R., Mouillot, D., Vause, J., Walpole, M., 2016. Biodiversity in the Anthropocene: prospects and policy. Proc. R. Soc. B 283, 20162094.

Sikes, R.S., Gannon, W.L., 2016. Animal care and use committee of the American Society of Mammalogists, 2016 guidelines of the American Society of Mammalogists for the use of wild mammals in research and education. J. Mamm. 97, 663–688.

Smedley, D., Haider, S., Ballester, B., Holland, R., London, D., Thorisson, G., Kasprzyk, A., 2009. BioMart – biological queries made easy. BMC Genomics 10, 22.

Smith, P.N., Cobb, G.P., Harper, F.M., Adair, B.M., McMurry, S.T., 2002. Comparison of white-footed mice and rice rats as biomonitors of polychlorinated biphenyl and metal contamination. Environ. Pollut. 119, 261–268.

Spielman, D., Brook, B.W., Frankham, R., 2004. Most species are not driven to extinction before genetic factors impact them. Proc. Natl. Acad. Sci. U.S.A. 101, 15261–15264.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

Talmage, S.S., Walton, B.T., 1991. Small mammals as monitors of environmental contaminants. In: Ware, G.W. (Ed.), Reviews of Environmental Contamination and Toxicology. Springer, New York, pp. 47–145.

Templeton, J.E.L., Brotherton, P.M., Llamas, B., Soubrier, J., Haak, W., Cooper, A., Austin, J.J., 2013. DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification. Investig. Genet. 4, 26.

Teske, P.R., Golla, T.R., Sandoval-Castillo, J., Arsalan Emami-Khoyi, A., van der Lingen, C.D., von der Heyden, S., Chiazzari, B., van Vuuren, B.J., Beheregaray, L.B., 2018. Mitochondrial DNA is unsuitable to test for isolation by distance. Sci. Rep. 8, 8448.

Tollis, M., Ausubel, G., Ghimire, D., Boissinot, S., 2012. Multi-locus phylogeographic and population genetic analysis of *Anolis carolinensis*: historical demography of a genomic model species. PloS One 7, e38474.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G., 2012. Primer3—new capabilities and interfaces. Nucleic Acids Res. 40, e115.

Upham, N.S., Esselstyn, J.A., Jetz, W., 2019. Inferring the mammal tree: species-level sets of phylogenies for questions in ecology, evolution, and conservation. PLoS Biol. 17, e3000494.

Valiente-Banuet, A., Aizen, M.A., Alcántara, J.M., Arroyo, J., Cocucci, A., Galetti, M., García, M.B., García, D., Gómez, J.M., Jordano, P., Medel, R., Navarro, L., Obeso, J.R., Oviedo, R., Ramírez, N., Rey, P.J., Traveset, A., Verdú, M., Zamora, R., 2015. Beyond species loss: the extinction of ecological interactions in a changing world. Funct. Ecol. 29, 299–307.

van Straalen, N.M., Timmermans, M.J.T.N., 2002. Genetic variation in toxicant-stressed populations: an evaluation of the "genetic erosion" hypothesis. Hum. Ecol. Risk Assess. 8, 983–1002.

Vilas, A., Pérez-Figueroa, A., Quesada, H., Caballero, A., 2015. Allelic diversity for neutral markers retains a higher adaptive potential for quantitative traits than expected heterozygosity. Mol. Ecol. 24, 4419–4432.

Wcisel, D.J., Howard, J.T., Yoder, J.A., Dornburg, A., 2020. Transcriptome ortholog alignment sequence tools (TOAST) for phylogenomic dataset assembly. BMC Evol. Biol. 20, 41.

Yang, Y., Xie, B., Yan, J., 2014. Application of next-generation sequencing technology in forensic science. Dev. Reprod. Biol. 12, 190–197.