Molecular Phylogenetics and Evolution 125 (2018) 85–92

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

The *Microtus* voles: Resolving the phylogeny of one of the most speciose mammalian genera using genomics

Soraia Barbosa[a,b,c,*], Joana Paupério[a], Svetlana V. Pavlova[d], Paulo C. Alves[a,b,1], Jeremy B. Searle[a,b,c,1]

[a] CIBIO/InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto/InBIO Laboratório Associado, 4485-661 Vairão, Portugal
[b] Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre s/n, 4169-007 Porto, Portugal
[c] Department of Ecology and Evolutionary Biology, Corson Hall, Cornell University, Ithaca, NY 14853-2701, USA
[d] A. N. Severtsov Institute of Ecology and Evolution RAS, Moscow 119071, Russia

ARTICLE INFO

Keywords:
Small mammals
Adaptive radiation
Genotyping-by-sequencing
Mitogenomes
Targeted capture

ABSTRACT

Sequential rapid radiations pose some of the greatest difficulties in phylogenetics, especially when analysing only a small number of genetic markers. Given that most of the speciation events occur in quick succession at various points in time, this creates particular challenges in determining phylogenetic relationships, i.e. branching order and divergence times. With the development of high throughput sequencing, thousands of markers can now readily be used to tackle these issues. *Microtus* is a speciose genus currently composed of 65 species that evolved over the last 2 million years. Although it is a well-studied group, there is still phylogenetic uncertainty at various divergence levels. Building upon previous studies that generally used small numbers of mitochondrial and/or nuclear loci, in this genomic-scale study we used both mitochondrial and nuclear data to study the rapid radiation within *Microtus*, using partial mitogenomes and genotyping-by-sequencing (GBS) on seven species representing five *Microtus* subgenera and the main biogeographic ranges where this group occurs. Both types of genome (mitochondrial and nuclear) generated similar tree topologies, with a basal split of the Nearctic (*M. ochrogaster*) and Holarctic (*M. oeconomus*) species, and then a subdivision of the five Palearctic species into two subgroups. These data support the occurrence of two European radiations, one North American radiation, and a later expansion of *M. oeconomus* from Asia to both Europe and North America. We further resolved the positioning of *M. cabrerae* as sister group of *M. agrestis* and refute the claim that *M. cabrerae* should be elevated to its own genus (*Iberomys*). Finally, the data support ongoing speciation events, especially within *M. agrestis*, with high levels of genetic divergence between the three Evolutionarily Significant Units (ESUs) previously identified. Similar high levels of divergence were also found among ESUs within *M. oeconomus* and *M. arvalis*.

1. Introduction

Diversification often occurs through rapid radiations, where geographical, climatic or ecological changes drive ancestral forms to subdivide into many new taxa with distinct ecological niches (Simões et al., 2016; Stroud and Losos, 2016). The elucidation of the driving factors of rapid radiations would greatly benefit from detailed phylogenies, but the synchrony of each radiation creates challenges in deciphering the branching order and the time since divergence (Degnan and Rosenberg, 2009; Giarla and Esselstyn, 2015; Whitfield and Lockhart, 2007). Until recently most phylogenetic studies have been limited to one or two mitochondrial loci and sometimes few nuclear loci, which fail to provide enough resolution for such complex situations (Jeffroy et al., 2006). Sole use of mitochondrial DNA (mtDNA) has drawbacks and there has been a tendency for a combined approach using different types of markers reflective of different evolutionary trajectories (Alves et al., 2006; Melo-Ferreira et al., 2012; Morin et al., 2004). Fast evolving mitochondrial genes provide enough resolution to distinguish between recently diverged taxa but are also more prone to effects of genetic drift and haplotype fixation given their smaller effective population size, long branch attraction (LBA), high chance of mutational saturation and difference in evolutionary rates among lineages

* Corresponding author at: CIBIO/InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto/InBIO Laboratório Associado, 4485-661 Vairão, Portugal.
E-mail address: soraiabarbosa@cibio.up.pt (S. Barbosa).
[1] Contributed equally.

https://doi.org/10.1016/j.ympev.2018.03.017
Received 11 July 2017; Received in revised form 3 January 2018; Accepted 14 March 2018
Available online 21 March 2018
1055-7903/

(Bergsten, 2005; Nabholz et al., 2008; Su and Townsend, 2015; Yang and Rannala, 2012). Thus mtDNA often becomes unsuitable for the study of older rapid radiations but has proven to be very helpful for recent divergences (Hurst and Jiggins, 2005) including applying contemporary calibrations (Herman and Searle, 2011). Slower evolving markers, namely nuclear exons and introns, have the advantage of avoiding many of the problems of using mitochondrial markers alone (such as mitochondrial introgression and LBA) but they often do not display enough genetic variation to resolve recent divergences, being more prone to effects of incomplete lineage sorting (ILS) (Giarla and Esselstyn, 2015). With the arrival of high throughput sequencing, genome-wide single nucleotide polymorphism (SNP) variation and whole genome sequences are revolutionising evolutionary studies, including solving difficult issues in phylogenetics (Andrews et al., 2016; Ree and Hipp, 2015; Wagner et al., 2012). Application of high throughput sequencing methods in phylogenetics ('phylogenomics', Jeffroy et al., 2006) can address deep evolutionary events like the eutherian radiation (Murphy et al., 2001; Song et al., 2012) and more recent diversifications, like the African cichlids (Brawand et al., 2014; Seehausen, 2006). Phylogenomics is increasingly going to replace studies with single or low numbers of markers but it is still desirable to include both nuclear and mitochondrial data in phylogenetics, even at a genomic scale (Duchêne et al., 2011; Filipi et al., 2015; Leaché et al., 2015; Moore, 1995).

The genus *Microtus* is part of the rodent family Cricetidae, and more specifically the subfamily Arvicolinae, representing one of the most rapid documented radiations in extant mammals. The *Microtus* radiation was proposed to have originated in central Asia with expansion both westwards and eastwards – to Europe and North America, respectively, *c.* 1 million years ago (Mya) – followed by multiple speciation events in both regions at various points in time (Chaline et al., 1999; Conroy and Cook, 2000, 1999; Jaarola et al., 2004; Weksler et al., 2010). A later expansion of *M. oeconomus* to Europe and North America is thought to have occurred *c.* 450 thousand years ago (kya), given the timing estimates of the second opportunity to cross over Beringia and supported by both fossil and molecular data leading to the current Holarctic distribution of *M. oeconomus* (Brunhoff et al., 2003; Galbreath and Cook, 2004; Lance and Cook, 1998). There are 65 species currently recognised in the genus *Microtus* (IUCN, 2016). From these, 44 species belong to 5 subgenera that are found in the Palearctic; 20 species belong to 7 subgenera representing the first wave of subgenera that colonised the Nearctic; and one species, *Microtus oeconomus*, has a Holarctic distribution in a subgenus that, considering all other species, is otherwise only found in the Palearctic.

Many molecular phylogenetic studies on *Microtus* have used mitochondrial cytochrome-*b* (*CYTB*) and found that most subgenera represent true monophyletic clades (Bannikova et al., 2010; Conroy and Cook, 2000; Jaarola et al., 2004). Other studies have used both mitochondrial and nuclear markers (single/low numbers of gene sequences and AFLPs) which have added support to the idea of a rapid radiation and two to three independent expansions from central Asia to Europe and two to North America, with associated bursts of speciation (Fink et al., 2010; Galewski et al., 2006; Martínková and Moravec, 2012; Robovský et al., 2008). These studies have provided the first steps towards a greater understanding of the *Microtus* radiation, but the high mitochondrial mutation rate observed in this genus is likely a limiting factor when aiming to resolve the initial splits of this rapid adaptive radiation (due to LBA), and the slower mutation rate of nuclear loci does not allow the effects of ILS to be parsed out (Giarla and Esselstyn, 2015; Whitfield and Lockhart, 2007). As a result, these earlier studies are impaired by phylogenetic inconsistencies across loci, conspecifics splitting into different groups, and low support for the more basal nodes, which is expected when using a small number of loci in groups that have undergone rapid radiations (Fink et al., 2010; Giarla and Esselstyn, 2015; Martínková and Moravec, 2012).

As well as the interest in the overall *Microtus* phylogeny and its biogeographic history, there is an interest in the phylogenetic position of certain species. In particular, the Cabrera vole (*Microtus cabrerae*) has one of the most unresolved positions within the genus and, although a European species, it has often been found to cluster with the North American species (Fink et al., 2010; Jaarola et al., 2004; Robovský et al., 2008). The Cabrera vole is classified in the subgenus *Agricola* together with the field vole (*Microtus agrestis*) based on its karyotype (Zagorodnyuk, 1990), however many still prefer its earlier classification by Chaline (1974) as the sole member of a separate subgenus, *Iberomys* (López-García and Cuenca-Bescós, 2012; Pita et al., 2014). Given the combination of numerous differentiating morphological and biological features – e.g. molar morphology, body size, reproductive strategy, chromosomes (Pita et al., 2014) – Cuenca-Bescós et al. (2014) have recently proposed the elevation of the Cabrera vole to its own genus, *Iberomys*. However, the elevation of this species to its own genus needs to be considered carefully and justified on phylogenetic grounds.

The aim of this study was to clarify the phylogeny of the genus *Microtus* using mitochondrial and nuclear genomes, giving particular relevance to the positioning of certain European species in a biogeographical context. We compare seven species from five subgenera defined primarily on morphology following the species description of Wilson and Reeder (2005): *Agricola* (*M. agrestis* and *M. cabrerae*), *Alexandromys* (*M. oeconomus*), *Microtus* (*M. arvalis* and *M. levis*), *Pedomys* (*M. ochrogaster*) and *Terricola* (*M. subterraneus*). These subgenera are particularly important to the biogeographic history of the genus and include all European subgenera, one Nearctic subgenus and the Holarctic subgenus. Our phylogenomic approach using a partial mitogenome and genotyping-by-sequencing (GBS) SNP data demonstrates that increasing the number of loci and their representation of the genome can improve phylogenetic inferences for recent radiations. We were confidently able to identify all species, subgenera and their relations, but the most basal branching order was difficult to resolve, indicating that rapid radiations are still a challenge even with increased genomic power.

2. Material and methods

2.1. Sources of mitochondrial and nuclear genomes

We obtained mitochondrial and nuclear genomes using tissue samples from seven *Microtus* species distributed across the entire northern hemisphere as listed in Table 1 and mapped in Fig. S1 (Appendix A). Up to four individuals per species were analysed to account for intra-specific variation (Table 1). Whenever possible, conspecific individuals were chosen from geographically distant localities (Fig. S1, Appendix A). For outgroups we also analysed two individuals of *Arvicola sapidus*, given its close relationship to *Microtus* (Buzan et al., 2008; Martínková and Moravec, 2012), generating a total of 18 specimens analysed separately for their mitochondrial and nuclear genomes (see Table 1 for details). Single *M. ochrogaster* and *M. levis* mitochondrial genomes were obtained from GenBank. We used the full genome of *M. ochrogaster* from GenBank (assembly GCA_000317375) as the reference for the SNP calling pipeline of all *Microtus*. We used a different *M. ochrogaster* individual as representative of the species in the multi-species SNP analysis (Table 1). This analysis used the TASSEL-GBS pipeline, a methodology known to be very robust against ascertainment bias (Glaubitz et al., 2014).

2.2. Mitochondrial data

2.2.1. Mitochondrial capture

All samples were amplified for the complete *CYTB* gene following Barbosa et al. (2013) to verify species identification (accession numbers in Table 1). For mitogenomes we performed a mitochondrial capture technique on two *Arvicola* and 14 *Microtus* samples, following Fu et al. (2013), that consisted of two main steps: (1) bait preparation and (2)