Biol Invasions
DOI 10.1007/s10530-012-0323-9

ORIGINAL PAPER

Cryptic genetic diversity in *Rattus* of the San Francisco Bay region, California

Chris J. Conroy · Kevin C. Rowe · Karen M. C. Rowe · Pauline L. Kamath · Ken P. Aplin · Lucia Hui · David K. James · Craig Moritz · James L. Patton

Received: 3 October 2011 / Accepted: 21 August 2012

Abstract Invasive species can have complex invasion histories, harbor cryptic levels of diversity, and pose taxonomic problems for pest management authorities. Roof rats, *Rattus rattus* sensu lato, are common invasive pests of the San Francisco Bay Area in California, USA. They are a significant health risk and pest management efforts impose a large financial investment from public institutions and private individuals. Recent molecular genetic and taxonomic studies of black rats in their native range in Asia have shown that the species is a complex of two karyotypic forms and four mitochondrial genetic lineages that may represent four distinct species. We used mtDNA sequences and nuclear microsatellite variation to identify which mitochondrial lineages of the *R. rattus* group are present in the San Francisco Bay Area and to test for gene flow among them. We recovered specimens with mtDNA sequences representing two of the major mtDNA lineages of the *R. rattus* group. Microsatellite variation, however, was not structured in concordance with mtDNA lineages, suggesting a more complex history involving hybridization and introgression between these lineages. Although Aplin et al. (2011) and Lack et al. (2012) reported *R. rattus* Lineage II in North America, this is the first detailed examination of possible gene flow amongst lineages in this region.

Keywords *Rattus* · Gene flow · Rodent · *Tanezumi* · Microsatellite · mtDNA

Introduction

The genus *Rattus* (Rodentia: Muridae) is native to southern and eastern Asia but several species and

Electronic supplementary material The online version of this article (doi:10.1007/s10530-012-0323-9) contains supplementary material, which is available to authorized users.

C. J. Conroy (✉) · K. C. Rowe · K. M. C. Rowe · C. Moritz · J. L. Patton
Museum of Vertebrate Zoology, University of California, Berkeley, Berkeley, CA 94720, USA
e-mail: ondatra@berkeley.edu

K. C. Rowe · K. M. C. Rowe
Sciences Department, Museum Victoria, Melbourne, Australia

P. L. Kamath
Department of Environmental Science, Policy and Management, University of California, Berkeley, Berkeley, CA 94720, USA

Present Address:
P. L. Kamath
Northern Rocky Mountain Science Center, US Geological Survey, Bozeman, MT 59715, USA

K. P. Aplin
National Museum of Natural History, Division of Mammals, Smithsonian Institution, Washington, DC 20013-7012, USA

L. Hui · D. K. James
Alameda County Vector Control Services District, 1131 Harbor Bay Parkway, STE 166, Alameda, CA 94502, USA

Published online: 12 September 2012

mitochondrial DNA (mtDNA) lineages have spread to other continental areas and occur both in human-modified habitats as commensal species and in natural habitats as feral invaders. The roof rat (*R. rattus*), sometimes also known as the black rat or ship rat, has achieved the broadest distribution, with populations established on all continents and most island groups. In the Americas it was documented in South America as early as 1544 (Pennant 1781, cited in Donaldson 1915), but was not recorded in eastern North America until the "beginning of the eighteenth century" (Lantz 1909). The Norway rat, *R. norvegicus*, was probably first introduced around 1775 (Lantz 1909; Silver 1927), and roof rat populations declined or disappeared from the northern and inland areas of North America as Norway rat populations increased. Today, both species of *Rattus* are known from nearly every major port in North America, and continue to move about by rail, ship, truck and plane.

The roof rat in North America is largely restricted to the southern East Coast, Gulf States and along the Pacific Coast (Jackson 1982). Roof rats invading natural habitats are uncommon within their range (Ecke 1964). Roof rats are found in riparian areas (Stroud 1982; Zeiner et al. 1990), often in association with Himalayan Blackberry, which had been introduced in the 1890s (Dutson 1974; Grinnell et al. 1930). With the suburbanization of many rural areas from the 1960s, the population and distribution of *R. rattus* has increased dramatically (Brothers 1972; Claffey et al. 1986; Nolte et al. 2003).

Invasive *Rattus* species present a number of problems where they have been introduced. *Rattus* are major crop pests creating or exacerbating food shortages for millions of humans in developing nations (Meerburg et al. 2009). Another major concern, especially on islands, is the decline or extinction of native fauna, through predation and/or competition (Atkinson 1985; Chiba 2010; Goodman 1995; Towns et al. 2006). Because roof rats invade native habitats and are not confined to commensal associations they also may replace native species of rodents, such as woodrats (*Neotoma*; Rodentia; Cricetidae), with which they can compete (Dutson 1974). The roof rat is perhaps best known as a key vector in the transmission of bubonic plague (Scott and Duncan 2001; Boisier et al. 2002) and invasive *Rattus* remain major vectors of diseases such as cowpox (Wolfs et al. 2002), murine typhus (Azad 1990), *Bartonella* (Ellis et al. 1999), Seoul hantavirus (Plyusnina et al. 2004), among others (see Aplin et al. 2011). Roof rats also may spread parasites to native rodent species increasing their impact on wildlife (Smith and Carpenter 2006).

Although at least five species of *Rattus* are commensal with humans and have invaded non-native ranges, *Rattus rattus* and *R. norvegicus* are the only two species that have established permanent populations in North America. This is somewhat surprising given the regularity and volume of ship and air transport between Asia and North America over many decades. However, *Rattus* is highly complex taxonomically (Musser and Carleton 2005) and many species are difficult to distinguish on gross morphology. One species that might easily have been overlooked is the Asian roof rat, *R. tanezumi*, which differs from Indian and most introduced populations worldwide in having 42 rather than 38 chromosomes (Yosida 1980; see discussion in Musser and Carleton 2005). More precise definition of the *Rattus* species present in North America is important for many fields including rodent disease studies and pest management. In both fields proper identification of *Rattus* species is essential because they potentially have distinct ecological behaviors and carry at least partially discrete sets of pathogenic organisms (Aplin et al. 2003).

There are a number of ways to identify rats encountered in the field (e.g. body measurements, behavior) but complementary genetic approaches have only recently become available. Robins et al. (2007) explored the utility of mtDNA as a species level marker in *Rattus*. They encountered a poor match between mtDNA phylogenetic structure and presumed taxonomic identity but whether this was due to specimen misidentification, incomplete gene sorting, or ongoing gene flow is unclear from their work. Pagès et al. (2010) also found great complexity in the phylogenetic diversity of southeast Asian *Rattus*, with the possibility of at least one undescribed species of the roof rat group. Chinen et al. (2005) reported apparent admixture of 'Asian' (=*R. tanezumi,* 2N = 42) and 'Oceanian' (=*R. rattus,* 2N = 38) forms in Japan and hybrids between these forms have been detected elsewhere, and created in the laboratory (Yosida et al. 1971). In a situation where hybridization is possible or even likely, a single maternally inherited marker such as mtDNA is insufficient for describing a population's taxonomic identity. However, molecular tools such as mitochondrial

DNA have proven highly effective for describing large-scale patterns of invasion by *Rattus* (Bastos et al. 2011; Kaleme et al. 2011; Tollenaere et al. 2010).

In this study we utilized specimens of *R. rattus* sensu lato collected by the Alameda County Vector Control District (California, United States) and others to assess their genetic diversity in the region. We combined these specimens with Asian and Madagascan *Rattus* specimens in the Museum of Vertebrate Zoology (MVZ) and other natural history museums. We included these non-North American samples to have representatives with distinct histories that would be apparent in both mtDNA and nuclear DNA (nDNA). Below we describe the use of molecular techniques to mitochondrial DNA lineages of *Rattus* in the San Francisco Bay Area and conduct preliminary analysis of their genetic variation. Given the distribution of multiple roof rat mtDNA lineages elsewhere around the world (Aplin et al. 2011), we hypothesized that our SF Bay Area sample might also contain genetic material derived from multiple mtDNA lineages of *R. rattus*, and possibly now mixed as a consequence of interbreeding. Accordingly, we used techniques that would not only allow us to identify rats of diverse genetic origin but also assess the degree of genetic admixture. We use the term *R. rattus* Complex (RrC) from Aplin et al. (2011) to describe the roof rat group as a whole and also use their somewhat conservative language in referring to Lineages and clades rather than species units. In this terminology *R. rattus* Complex Lineage I has as its primary association *R. rattus* sensu stricto, while Lineage II is associated with populations commonly referred to as *R. tanezumi* (Bastos et al. 2011; Chinen et al. 2005; Musser and Carleton 2005; Pagès et al. 2010; Robins et al. 2007).

The *R. rattus* group provides an opportunity to investigate multiple occurrences of admixture between divergent lineages over broad geographic and time scales. Our sampling area is unique as a relatively recent recipient area of *Rattus* diversity, and different from the Asian and Madagascan regions we sampled for contrast. We expect to find some differences between the regions given the length of time *Rattus* have existed there. The demographic history of the populations from which we utilized distant samples also is unknown, thus limiting our ability to make direct comparisons. There have been few studies documenting wild *Rattus* genetic diversity at the scale that we report here, especially with the markers we employ. Aplin et al. (2011) also recognize the possibility that admixture within the native ranges of *R. rattus* may have occurred and they also use caution with respect to inferences about population level histories. Thus, we restrict our analyses to the SF Bay Area where we have a much larger sample.

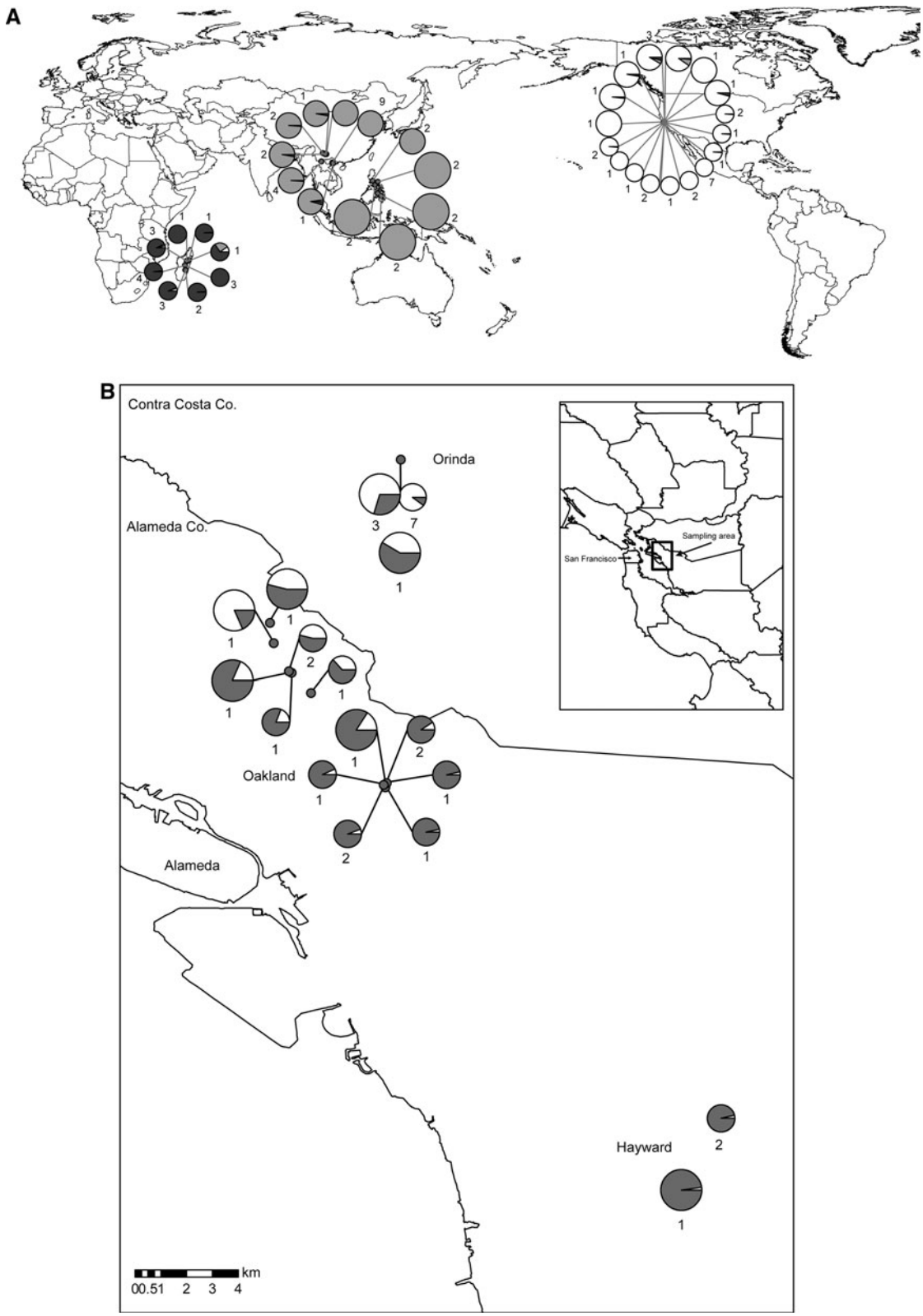
We hypothesize that if *R. rattus* mtDNA lineages can remain distinct when in contact, we should be able to pick up a signal in both mitochondrial and nuclear markers. If we have multiple mtDNA lineages of *R. rattus* in the SF Bay Area, and they can remain distinct while sympatric, then some rats in the region should align with RrC Lineage I or II rats from more native parts of their range across markers. If gene flow is occurring, we should not see a coherent signal of RrC Lineages in the SF Bay Area.

Methods and materials

Specimen acquisition

We utilized MVZ specimens originating from non-US locations (Fig. 1a), and additional tissues on loan from the Field Museum of Natural History (FMNH) and the Museum of Southwestern Biology (MSB). We analyzed *R. rattus* Complex specimens newly trapped in Alameda County, California, as well as *Rattus* samples already preserved at the MVZ collection, at the University of California at Berkeley (Fig. 1b). We limited our work only to rats identified in the field or museum as *R. rattus* based on gross morphology (e.g. tail to total length ratio and color). For a complete specimen list, see Supplementary Table S1. While we refer to the San Francisco Bay Area (SF Bay Area), our samples are confined to the east side of the San Francisco Bay, primarily in Alameda and Contra Costa Counties (Fig. 1b).

We included 214 partial or complete mtDNA cytochrome b (cyt *b*) sequences, among which are 90 individuals of *Rattus* from California, including 87 from the SF Bay Area (Table 1). To obtain an approximation of worldwide diversity we included 36 non-North American *Rattus* RrC Lineages I and III (sensu Aplin et al. 2011; of which we sequenced 19, 17 others from Genbank) and 55 non-North American RrC Lineage II and IV (this paper: 20, Genbank: 35). To better root this analysis, we included 33 sequences from *Maxomys*, *Niviventer*, other species of *Rattus*,


A
B
Contra Costa Co.
Orinda
Alameda Co.
Sampling area
San Francisco
Oakland
Alameda
Hayward
km
00.51 2 3 4

◀ **Fig. 1** **a** Range-wide distribution of *Rattus* samples showing mitochondrial haplotype lineages and average likelihood values for each of the three microsatellite clusters within each locality. *Small*, *medium*, and *large circles* represent mitochondrial Lineage I, Lineage II, and Lineage IV, respectively. Each pie slice represents the mean likelihood of assignment to microsatellite cluster 1 (*white*), cluster 2 (*gray*), or cluster 3 (*black*) of each population. Sample sizes are given near each circle. **b** SF Bay Area *Rattus* samples showing mitochondrial haplotype lineages and average likelihood values for each of the two microsatellite clusters within each locality. *Small* and *large circles* represent mitochondrial Lineage I and Lineage II, respectively. Each pie slice represents the mean likelihood of assignment to microsatellite cluster 1 (*white*) or cluster 2 (*dark gray*) of each population. Sample sizes are given below each circle

and RrC Lineages V and VI (this paper: 7, Genbank: 26).

We attempted to use diverse non-North American samples where possible. The samples from Madagascar are all from the east side of Madagascar, but are spread out over nearly 480 km. The Vietnamese samples come largely from Tam Dao National Park, with one other sample from Hanoi, about 48 km away. The samples from the Philippines come from the islands Luzon, Midanao, Sibuyan and Siquijor. The samples from China are spread over three general areas about 480 km apart. The two Pakistan specimens originate about 190 km apart. The two samples from Iran are from the same locality. The two samples from Cambodia also came from one locality in that country.

DNA extraction and mitochondrial sequencing

Genomic DNA was extracted from tissue samples using either a modified salt extraction method (Miller et al. 1988) or Qiagen DNeasy extraction kits (Valencia, CA). We amplified, through polymerase chain reaction (PCR), 803 bp of the cyt *b* mitochondrial locus using the following primers: MVZ-05 (5′-CGAAGCTTGATATGAAAAACCATCGTTG-3′) (S. Pääbo, MVZ) and MVZ-16 (5′-AAATAGGAAR TATCAYTCTGGTTTRAT-3′) (C. Orrego, MVZ). PCRs were carried out in a total volume of 15 μL and contained 25 ng of DNA, 2 μL of 10 × PCR buffer (100 mM Tris–Cl, pH 8.3, 500 mM KCl, 15 mM $MgCl_2$, 0.01 % (w/v) gelatin) 400 μM deoxynucleotide triphosphates (dNTPs), 1U Ampli*Taq* Gold DNA polymerase (Applied Biosystems), 15 μg bovine serum albumin (New England BioLabs), 0.33 μM of each primer. The following thermocycling conditions were used: an initial denaturation at 95 °C for 6 min; 40 cycles of 94 °C for 30 s, 54 °C for 1 min, and 72 °C for 1 min 30 s; final extension at 72 °C for 15 min. All PCRs were performed using BioRad thermocyclers. Amplified PCR products were purified using Exo-SAP-IT (USB Corporation) and cycle-sequenced in both forward and reverse directions using the Big Dye® Terminator v.3.1 kit and run on an ABI 3730 sequencer (Applied Biosystems).

Additional sequences of previously published cyt *b* for RrC and additional species of *Rattus* from throughout their global distribution were downloaded from GenBank (Table S1). Sequences were aligned using CodonCode v2.0.6 (CodonCode Corporation) and manually inspected in MacClade v.4.08 (Maddison and Maddison 2005).

Phylogenetic analysis of mtDNA sequences

Phylogenetic analyses were conducted on the complete data set using both partitioned Bayesian methods as implemented in MrBayes ver. 3.1.2 (Huelsenbeck and Ronquist 2001) and maximum likelihood methods

Table 1 For California populations, sample size (N), major mtDNA Lineage detected, and haplotypes

Locality	N	Lineage I	Lineage II	Haplotypes detected
Berkeley, CA	2	2	0	44
Castro Valley, CA	2	2	0	44, 57
Dublin, CA	4	4	0	44
Hayward, CA	3	2	1	47, 40
La Grange, CA	1	1	0	44
Oakland, CA	47	43	4	44, 47, 42, 63
Orinda, CA	12	8	4	44, 46
San Lorenzo, CA	17	17	0	47
Snelling, CA	2	2	0	44

For individual data, see table Supplementary Table S1

as implemented in RAxML (Stamatakis 2006). For both analyses we reduced the dataset to unique haplotypes but retained haplotypes that differed based on length or base ambiguities. In both analyses, *Niviventer* was used as the outgroup. For Bayesian analyses, sequences were partitioned by codon positions into two partitions, with the third codon position separated from the first and second positions. For each partition, we calculated likelihood scores for 24 hierarchical substitution models in PAUP v.b4.10 (Swofford 2002) using the "modelblock" file in MrModeltest (Nylander 2004) and selected model parameters based on the Akaike information criteria implemented in MrModeltest. In our MrBayes runs, we unlinked parameters for each partition and allowed branch lengths to vary proportionately across partitions using the ratepr = variable setting. We ran two independent sets of four chains for 10 million generations with trees and parameters recorded every 500 generations. Convergence and stationarity of MrBayes runs were estimated by examination of likelihood plots, split frequencies and by means of diagnostics from AWTY (Wilgenbusch et al. 2004). We also used the web-based RAxML Black Box software (Stamatakis 2006) to estimate an ML topology and to provide ML bootstrap support values. We used DnaSP (Rozas et al. 2003) to generate nucleotide and haplotype diversity (Hd) statistics for each of the geographic regions. We included all sequences, but defined haplotypes in a method that ignored base ambiguities and length differences.

Microsatellites genotyping

We developed an approach using co-dominant nuclear genetic markers to assess genetic structure and gene flow among rats in the SF Bay Area. We genotyped 12 microsatellite markers in 29 individuals from each of two major mtDNA haplogroups which correspond with Aplin et al.'s (2011) Lineage I (N = 20) and Lineage II (N = 9). All samples genotyped were also sequenced at the cyt *b* locus. We restricted our analysis of California samples to test the hypothesis of gene flow between rats of different mtDNA types. *Rattus* samples were also genotyped from non-US locations, including China ($n = 11$), Vietnam ($n = 10$), Philippines ($n = 10$) and Madagasacar ($n = 18$), to determine the ability of the marker set to assess diversity, gene flow and differentiation among rats from distinct geographic regions. We expected that if the markers were to be useful discerning any structure within the SF Bay Area, we should examine how they function at a larger scale. If they were insufficient for detecting variation or structure at this level, we would lack confidence in their utility for addressing our more local question.

We used six markers that were originally isolated from *R. rattus* (Rr14, Rr17, Rr22, Rr54, Rr68, Rr114—Loiseau et al. 2008), and six that were isolated from *R. norvegicus* (D10Rat20, D12Rat76, D15Rat77, D16Rat81, D5Rat83, D2Rat234—Jacob et al. 1995). An M13 fluorescent labeling protocol (Schuelke 2000) was used to amplify five of the microsatellite loci: Rr22, Rr68, D10Rat20, D12Rat76, D2Rat234. This protocol used an M13 primer tagged with a HEX fluorescent label (5′ HEX-TGTAAAACGACGGCCAGT-3′) in addition to the standard locus-specific forward and reverse primers. Forward primers were 5′-end tagged with a universal long tag (5′-CGAGTTTTCCCAGTCACGAC-3′) capable of binding to the M13-HEX primer, and reverse primers were pigtailed (5′-GTTTCT-3′) at the 5′-end for stability of the reaction. PCR mixes consisted of 25 ng DNA, 1.33 μL GeneAmp 10 × PCR buffer, 0.7 U Ampli*Taq* Gold DNA polymerase (Applied Biosystems), 0.2 mM dNTPs, 10 μg bovine serum albumin (New England BioLabs) and 0.12 μM of the three primers (i.e. M13 flourescently tagged, forward universal tagged, and reverse pigtailed) in a total reaction volume of 10 μL. Thermocycling profiles for amplification of loci following the M13 protocol were: initial denaturation at 94 °C for 6 min; 30 cycles of 94 °C for 50 s, 60 °C for 1 min, and 72 °C for 1 min; 10 cycles of 94 °C for 50 s, 50 °C for 1 min, and 72 °C for 1 min; final extension of 72 °C for 30 min.

For the remaining seven loci, we directly tagged the forward primers at the 5′-end with HEX or 6-FAM fluorescent labels (Sigma-Aldrich) for visualization of allele fragments and determination of individual genotypes. We used a standard protocol for amplification of these loci with PCR mixes as described above, with the exception of increased concentrations of primers (0.5 μM of each forward and reverse primer) and dNTPs (0.4 mM). PCRs for Rr14, Rr17, Rr114, D15Rat77 and D16Rat81 were carried out under a "touch-down" thermocycling profile: initial denaturation at 95 °C for 10 min; 2 cycles of 94 °C for 1 min, 60 °C for 1 min, and 70 °C for 35 s; 18 cycles

of 93 °C for 45 s, 59 °C for 45 s, and 70 °C for 45 s, with the annealing temperature decreasing by 0.5 °C with each cycle; 20 cycles of 92 °C for 30 s, 50 °C for 30 s, and 70 °C for 1 min; final extension at 72 °C for 10 min. Whereas, PCRs for Rr54 and D5Rat83 loci used the following thermocycling profile: initial denaturation at 95 °C for 15 min; 30 cycles of 94 °C for 1 min, 60 °C for 1 min 30 s, and 72 °C for 1 min; final extension at 72 °C for 30 min. Negative controls were run with all PCR reactions.

Allele fragments were scored for size against the LIZ-500 size standard through electrophoresis using an ABI3730 DNA Sequencer followed by visualization with GeneMapper v.4.0 (Applied Biosystems) software. See Table S2 for a descriptive summary of the microsatellites and corresponding protocols that were used.

Statistical analyses

For all microsatellite loci, we tested for departures from Hardy–Weinberg (H–W) equilibrium and genotypic linkage disequilibrium (LD) between pairs of loci using Genepop 4.0 on the web (Raymond and Rousset 1995). Significance of exact tests was estimated with a Markov chain (MC) algorithm (Guo and Thompson 1992) with default parameters and correcting for multiple comparisons through a sequential Bonferroni procedure (Holm 1979). Diversity statistics including mean number of alleles (A), observed heterozygosity (H_O), expected heterozygosity (H_E) and the inbreeding coefficient (F_{IS}) were calculated in Genalex v.6.2 (Peakall and Smouse 2006).

Genetic differentiation among samples (both population and "species") was assessed by computing overall and pair-wise F_{ST} values in Arlequin 3.5 (Excoffier and Lischer 2010). Differentiation among population pairs was assessed using a weighted average F_{ST} over all microsatellite loci, by computing a distance matrix between sample groupings (based on Weir and Cockerham 1984). An Analysis of Molecular Variance (AMOVA) was conducted among population samples, using non-parametric permutation procedures (10,000 permutations), to determine the significance of the covariance components associated with the different levels of population structure and estimate overall levels of differentiation (Φ-statistics; Excoffier et al. 1992). The frequency of null alleles was estimated per locus using the expectation maximum algorithm (EMA) implemented in FreeNA (Chapuis and Estoup 2007), and was accounted for in estimates of F_{ST} (Weir and Cockerham 1984).

We used the software Structure v2.3 (Pritchard et al. 2000), which uses a Bayesian model-based genotypic clustering approach, to address whether structure exists between mtDNA RrC Lineages in the SF Bay Area, and assess whether admixture has been occurring. We also examined clustering in roof rats across a larger regional scale. Individuals were grouped based on genotypic frequencies, under models that incorporated admixture and assumed alleles are correlated, but differ in the number of populations (K). We conducted five independent runs, using different random number starting seeds, for each value of K (where $K = 1$–10). Runs were carried out for a total length of 1,000,000 Markov Chain Monte Carlo (MCMC) iterations with a burn-in period of 100,000 steps. Convergence was assessed by consistency among runs and inspection of plots. The "true" number of clusters (K) was estimated by both identifying the K value with the maximum log likelihood and by ΔK, the rate of change in the log probability of the data between successive K values (Evanno et al. 2005). This latter approach maximizes the mean posterior probability of the data for a given K value, Pr ($X|K$), returned for a run set whilst minimizing variance between runs. Outputs from replicate analyses for the most likely K were aligned using the algorithm implemented in CLUMPP v1.1 (Jakobsson and Rosenberg 2007) and visualized using *Distruct* (Rosenburg 2004).

Results

MtDNA diversity

MtDNA sequences submitted to Genbank have accession numbers JQ814144 to JQ814283 and JQ824369 (Table S1). We detected 63 mtDNA haplotypes overall using the "ignore ambiguities" option in DnaSP (Table S1, Table 2). Many of these are highly divergent as they include outgroups from other species, but also many minor variants of the major *R. rattus* RrC Lineages. Haplotype diversity (Hd) differed across major geographic or Lineage groups, ranging from 0.432 in SF Bay Area Lineage I rats up to 0.876 in Lineage II rats from China. Nucleotide diversity also varied across groups in similar ways. Unfortunately, we had little

Table 2 Diversity statistics by *Rattus* mtDNA-defined Lineages in the SF Bay Area and over all regions

Population	µsat N	cyt *b* N	Hd	π	A	H_O	H_E	F_{IS}
SF Bay Area	29	87	–	–	6.79 (± 0.54)	0.591 (± 0.049)*	0.682 (± 0.042)	0.121 (± 0.047)
Lineage I	20	78	0.432 (± 0.046)	0.00113	7.25 (± 0.91)	0.567 (± 0.069)*	0.675 (± 0.069)	0.139 (± 0.064)
Lineage II	9	9	0.667 (± 0.105)	0.00633	6.33 (± 0.60)	0.616 (± 0.071)*	0.689 (± 0.051)	0.103 (± 0.072)
Madagascar	18	18 (I)	0.111 (± 0.096)	0.00015	6.42 (± 0.82)	0.645 (± 0.08)	0.664 (± 0.048)	0.079 (± 0.077)
China	11	15 (II)	0.876 (± 0.070)	0.03596	10.33 (± 0.38)	0.669 (± 0.086)*	0.837 (± 0.023)	0.221 (± 0.089)
Vietnam	10	13 (II)	0.833 (± 0.086)	0.00518	9.00 (± 0.92)	0.728 (± 0.100)*	0.796 (± 0.056)	0.143 (± 0.119)
Philippines	10	10 (IV)	0.533 (± 0.180)	0.00333	9.33 (± 0.43)	0.603 (± 0.073)*	0.823 (± 0.020)	0.280 (± 0.078)

N sample size, *Hd* cyt *b* haplotype diversity (±Standard Deviation), π cyt *b* nucleotide diversity, *A* number of alleles, H_O observed heterozygosity, H_E expected heterozygosity, F_{IS} inbreeding coefficient estimated using the Weir and Cockerham (1984) method. Standard errors of estimates are reported (±SE). In regions outside the SF Bay Area, samples for mtDNA statistics were sufficient for only one mtDNA Lineage, noted in parentheses

*Significant departures from H–W equilibrium

control over the finer scale sampling of rats from the various regions and these statistics might be difficult to interpret.

Both Bayesian and maximum likelihood approaches produced phylogenies that resolve major lineages in the *R. rattus* complex (Fig. 2a, b). These topologies agree with those also uncovered by Aplin et al. (2011). That is, we found monophyly for Clades A, B, and I to VI as depicted in Aplin et al. (2011), as well as the same ordering of those clades. *Rattus* specimens from both RrC Lineage I and Lineage II were found in California in several locations across the east San Francisco Bay region (Fig. 1b).

The strength of nodes in our analyses were not as strong as those reported by Aplin et al. (2011). However, we recovered Bayesian posterior probabilities of 1.00 and 1.00 for RrC lineages I and II, respectively, and 0.99 for their sister relationship. Lower probabilities, between 0.58 and 0.98, were found for nodes supporting relationships among more distant lineages. This may simply reflect the greater sample sizes used in their study. They also used a different set of sequences, so random differences in node strength are expected. The phylogenies we present may suggest higher diversity than expected from the number of haplotypes shown in Table S1 and Table 2. This is due to the slight differences in sequence length and inclusion of ambiguities in the sequences, whereas we chose not to include such vagaries in determining the number of unique haplotypes in the dataset with DnaSP. We are confident that the specimens we examined from the SF Bay Area are from RrC Lineages I and II, sensu Aplin et al. (2011).

Of the 87 specimens from the SF Bay Area, 78 had mitochondrial haplotypes from Lineage I and 9 individuals had mitochondrial haplotypes from Lineage II (Fig. 2). These specimens were distributed across Alameda and Contra Costa Counties, with some specimen clustering in the north, and also derive from a mixture of urban and rural locations. Due to variation in tissue quality and PCR success, there are slight differences among the cyt *b* haplotype read length and quality in the SF Bay Area. Despite this, we found 3 haplotypes (haplotypes 44, 47 and 57) among Lineage I rats in the SF Bay Area (Table 1). Two of these haplotypes were detected in samples from Iran, French Polynesia, New Zealand, Samoa, Japan and Papua New Guinea (haplotype 44), as well as Brazil (haplotype 47). We detected 4 haplotypes (haplotypes 40, 42, 46 and 63) among Lineage II rats in the SF Bay

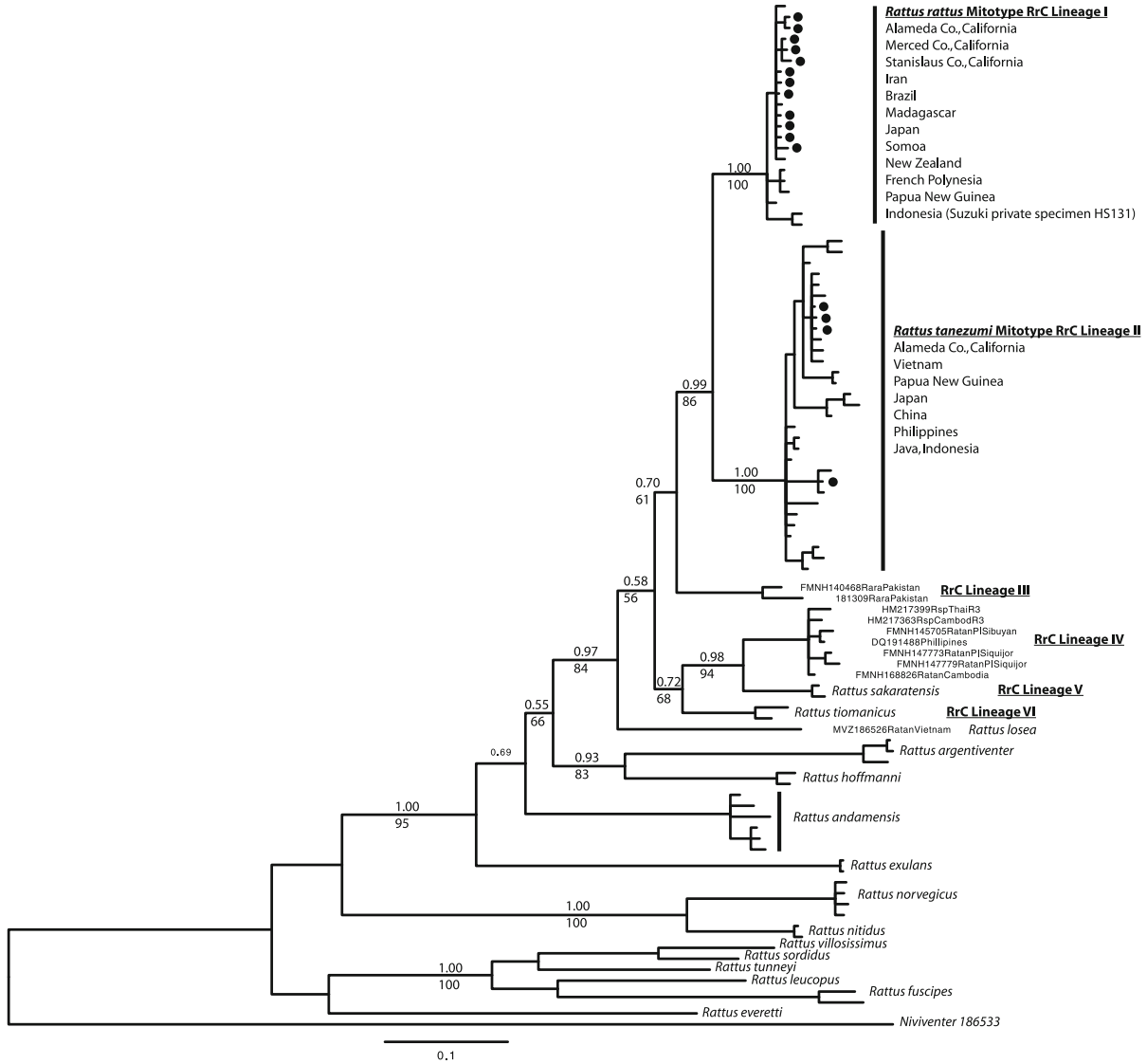


Fig. 2 **a** MtDNA phylogeny showing only unique haplotypes. *Bold tips* indicate haplotypes detected in the SF Bay region. These tips also include haplotypes distinguished by length variation and base ambiguities. Scores above branches are Bayesian posterior probabilities and those below ML bootstrap support values. **b** MtDNA phylogeny showing only unique haplotypes of *R. rattus*, showing only Lineages I and II. As in Fig. 2a, *bold tips* indicate haplotypes detected in the SF Bay region

Area, whereas Aplin et al. (2011) reported two of these from two rats. The haplotypes we detected were also found in the Philippines (haplotype 40), China, Japan and Vietnam (haplotype 42), and Indonesia (haplotype 46). Within the SF Bay Area there was sharing of mtDNA haplotypes across localities among Lineage I haplotypes, but not among Lineage II haplotypes. In Lineage II, haplotype 40 was found only in Hayward, haplotype 42 only in Oakland, haplotype 46 only in Orinda and haplotype 63 was represented only by one rat from Oakland (Table 1). Again, our sampling here was quite low and more sampling might allow greater detection of Lineage II haplotypes elsewhere.

Three localities (Hayward, Oakland, and Orinda) included Lineage I and II haplotypes (Table 1). Individuals from outside North America also fell into one or another of the known mitochondrial Lineages (Fig. 1b, Table S1).

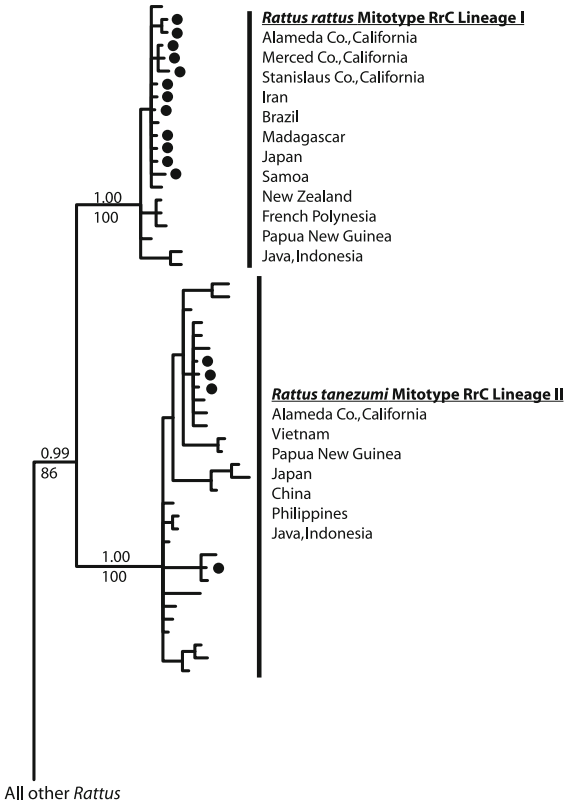


Fig. 2 continued

The three distinct Lineage I haplotypes identified in the SF Bay Area all belong to the 'ship rat' clade identified by Aplin et al. (2011) as representing the local mtDNA diversification that occurred in Europe following dispersal out of India in prehistoric times. The Lineage II haplotypes found in the SF Bay Area are members of sublineage IIB as identified by Aplin et al. (2011). This sublineage has its primary distribution in the western Pacific including the Philippines, and is less well represented on mainland Asia. It should be noted that some haplotypes are spread quite broadly. Aplin et al. (2011) noted that the Lineage II haplotype they detected in California was also found in Australia and South Africa. It seems likely that many other haplotypes, and possibly other *R. rattus* Lineages, will be found with further sampling.

Exact tests for H–W and linkage equilibrium

One locus, Rr14, exhibited significant departures from Hardy–Weinberg equilibrium in SF samples ($p < 0.001$, after Bonferroni correction). Exact tests conducted by locus in other regions (China, Vietnam, Philippines, and Madagascar) were also significant for the following combinations of loci: Rr14, Rr68 and Rr17 in China; Rr14 and Rr68 in Vietnam; Rr17, Rr54 and Rr68 in the Philippines; D12Rat76 in Madagascar. Averaged over all loci the SF Bay Area, China, Vietnam and Philippines rat populations showed significant ($p < 0.05$) heterozygote deficiency (Table 2). We hypothesize that the inconsistent patterns of departures from H–W across loci are the result of population specific demographic phenomenon, such as population sub-structure, rather than locus-specific effects. Given that there is likely widespread mixing of *Rattus* lineages, departures from H–W would be expected. We found no evidence for significant LD between locus pairs over all populations after Bonferroni corrections. Estimates of mean null allele frequency (NAF) was moderate (>10 %) at the following loci: Rr14, Rr17, Rr54 and Rr68 (Supplementary, Table S3). However, variability in these estimates was high across populations, again suggesting that these observations are likely driven by demographic or sampling effects, rather than due to genotyping error.

Diversity and differentiation among mtDNA RrC Lineages and populations

All microsatellite loci examined were polymorphic in both *Rattus* mtDNA Lineages in the study area and among geographic localities (Table S2). In our California sample, the number of alleles at each locus ranged from 2 to 12 alleles in rats with Lineage I (N = 20) mitochondrial DNA and 2 to 9 alleles in rats with Lineage II (N = 9), and diversity indices (A, H_E) among rats in each mtDNA Lineage category were not significantly different (Table 2). SF Bay Area *Rattus* exhibited significantly lower indices of genetic diversity than rats from China and the Philippines ($p < 0.05$); they were also less diverse than the Vietnam rat population sample, though not significantly. In contrast, SF Bay Area *Rattus* exhibited levels of diversity similar to that of rats from Madagascar.

We found a lack of genetic differentiation at nuclear loci between rats with mtDNA Lineages I and II within the study area ($F_{ST} = 0.001$, $p = 0.701$; Table 3) and AMOVA results revealed that the majority of genetic variation (70.6 %) is found within individuals (Table S4). Global differentiation was relatively high ($\Phi_{ST} = 0.166$, $p < 0.001$) providing evidence for limited gene flow among the regions.

Table 3 Population pair-wise F_{ST}

Weir and Cockerham's pairwise F_{ST}. values are shown below the diagonal. Significance (p values) based on 10,000 permutations are shown above the diagonal, with highly significant ($p < 0.001$) values in bold

Populations	SF Bay Area Lineage II	SF Bay Area Lineage I	Madagascar	China	Vietnam	Philippines
SF Lineage II	–	0.701	**<0.001**	**<0.001**	**<0.001**	**<0.001**
SF Lineage I	0.001	–	**<0.001**	**<0.001**	**<0.001**	**<0.001**
Madagascar	0.174	0.184	–	**<0.001**	**<0.001**	**<0.001**
China	0.153	0.192	0.155	–	**<0.001**	**<0.001**
Vietnam	0.176	0.211	0.205	0.049	–	**<0.001**
Philippines	0.176	0.210	0.194	0.066	0.093	–

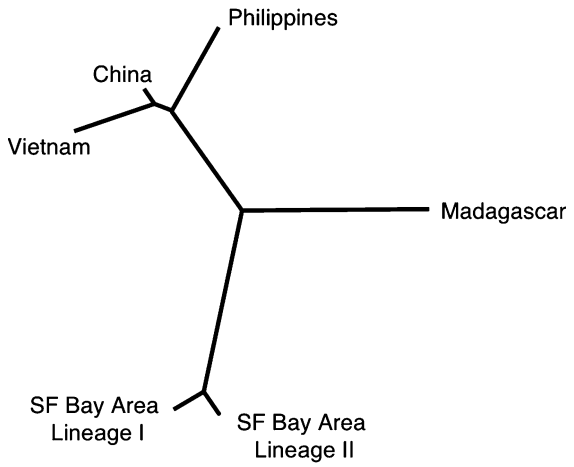


Fig. 3 F_{ST} NJ plot of inferred similarities among regional groups

Table 4 Average posterior probability of K populations in STRUCTURE

K	By Lineage		By population	
	Ln P(D)	SD	Ln P(D)	SD
1	−1066.72	0.48	−4406.26	1.34
2	−1058.78	11.11	−3938.14	10.38
3	−1175.72	48.34	−3640.54	1.96
4	−1325.72	57.68	−3680.32	21.09
5	−1304.60	45.18	−3748.02	62.23
6	−1343.22	58.86	−3685.84	52.47
7	−1262.16	29.83	−3811.92	53.78
8	−1267.10	41.83	−3968.32	55.20
9	−1266.08	13.89	−4013.50	82.40
10	−1311.78	42.00	−4127.38	119.83

Analyses conducted over $K = 1$–10 populations within SF Bay Area *Rattus* "species" and over all populations. Runs were conducted under a model of admixture and assuming alleles are correlated (burnin = 100,000; # iterations = 1,000,000). STRUCTURE output of the average log likelihood (Ln P(D)) and standard deviation (SD) for five runs at each value of K are reported

As would be expected based on geographic proximity, differentiation was lower between pairs of populations within Asia ($F_{ST} = 0.049$–0.093) in comparison to all other population pairs ($F_{ST} = 0.155$–0.211) (Table 3; Fig. 3). Results after correcting for potential null alleles were similar (Table S5).

Genotypic clustering

Only Rr14 showed significant departures from Hardy–Weinberg equilibrium over a majority (3 out of 5) of populations; therefore, structure analyses were run after removing this locus. In the study area, clustering analyses revealed the highest log-likelihood for two genetic populations ($K = 2, \text{Ln} = -1058.78$; Table 4), a result also corroborated by estimates of ΔK (Fig. S1). The ΔK approach for model selection cannot assess $K = 1$, and, thus, one genetic cluster cannot be rejected. Our results suggest that while there may be genetic sub-structure within SF Bay Area *Rattus*, individuals do not cluster by cyt *b* Lineage (Fig. 4) indicating a significant amount of mixing between the two lineages.

Structure analyses including *Rattus* samples from outside of the study area (Madagascar, China, Vietnam, and the Philippines) revealed that this marker set is capable of detecting structure among geographic regions of *Rattus* (Fig. 5). Notably, clustering did not correspond with haplotype lineage assignments based on the mtDNA phylogeny. We found the highest log-likelihood for $K = 3$ populations ($\text{Ln} = -3640.24$): (1) Asian populations, mtDNA Lineage II and IV, (Table 4; Fig. 5), (2) Madagascar populations, mtDNA Lineage I, and (3) San Francisco Bay Area populations, mtDNA Lineages I and II. Estimates of ΔK furthermore corroborated this result and provided the strongest evidence for three genetic groups (Fig. S1). When K was

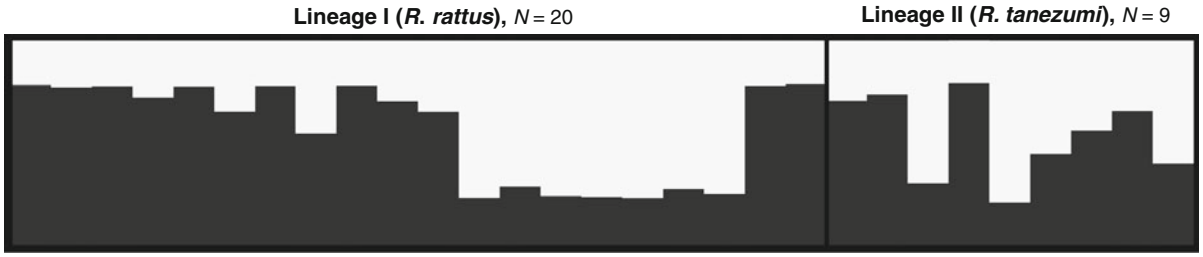


Fig. 4 STRUCTURE plot of *Rattus* from the SF Bay Area. With $K = 2$ populations having the highest posterior probability (Ln $= -1058.78$). *Black* and *white* indicate the relative probability of each specimen to belong to either $K = 2$ populations

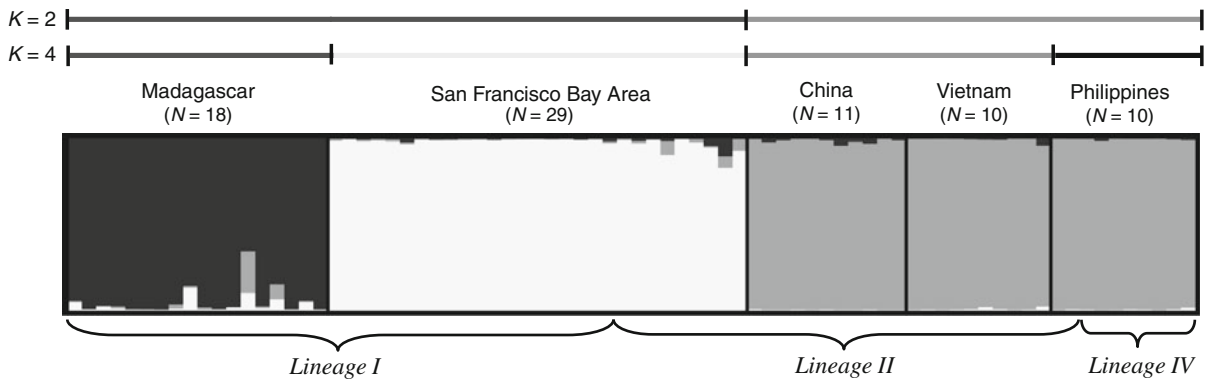


Fig. 5 STRUCTURE plot of *Rattus* across regions. Plot displays $K = 3$ populations, which has the highest log likelihood (Ln $= -3640.54$), confirmed by estimation of "true" K based on ΔK (see Fig. S1). *Black*, *grey* and *white* indicate relative probability of each specimen to each of the $K = 3$ populations. Bars above plot show the genotyping clustering by region when K is forced to equal 2 and 4. Genotypic clustering breaks down when $K > 4$. Mitochondrial haplotype lineages indicated, corresponding to cyt *b* phylogeny (Fig. 2)

forced to equal four populations, the Philippines *Rattus* clustered separately from the other Asian populations. Whereas, when K was forced to equal two, SF Bay Area rats grouped with the Madagascan population.

Discussion

Invasive species are an ongoing threat to numerous aspects of intact, native ecosystems and human health, welfare and industry (Lee 2002; Pimentel et al. 2005). The ability to detect and combat these invasive species assumes the ability to correctly identify them. Proper identification, with knowledge of their ecology in their native habitat or other colonized habitat, can aid in their eradication or control. Tracking the movement of diseases carried by invasive species also relies on proper identification of those invaders. However, some invasive species may be cryptic and our data suggest that *Rattus* are no exception.

Brown and Simpson (1981) were perhaps the first to examine genetic diversity in *Rattus* across North America. These authors were surprised by the degree of mtDNA diversity they uncovered, but at that time our understanding of mtDNA evolution was rudimentary. Although they included *R. rattus* from Alameda County (specific individuals not identified) they detected only one restriction enzyme profile in 11 rats which, most likely, corresponds to the numerically dominant mtDNA type (Lineage I) that we found. However, their work suggested divergent mtDNA haplogroups among roof rats from other geographic areas and they called for further examination of the genetic diversity of rats across North America to better understand their movement patterns.

Our main finding is that roof rats in the SF Bay Area contain genetic material derived from two distinct sources—*R. rattus* Complex Lineages I and II, which have been equated to *R. rattus* and *R. tanezumi* of recent authors (e.g. Bastos et al. 2011, Chinen et al. 2005;

Musser and Carleton 2005; Pagès et al. 2010; Robins et al. 2007). In contrast, the results of the microsatellite analyses indicate that roof rats in the SF Bay Area comprise a single genetic population, with weak evidence of geographic subdivision but no indication of subdivision by the mtDNA Lineages. Elsewhere in the geographic range of roof rats we did find significant nuclear genetic substructuring (i.e. Madagascar populations distinct from Asian populations), hence the most likely explanation of the lack of agreement between mtDNA and nDNA microsatellite variation in the SF Bay Area is a history involving hybridization and genetic introgression. This finding has implications for various issues in the wider biology of roof rats.

Wider taxonomic issues

Evaluation of single genetic markers like mtDNA to identify animals to species (i.e. genetic barcoding) can often be efficient and highly informative. However, in groups where hybridization is common, reliance on a single marker can be misleading. In the case of the RrC, previous work demonstrates that populations with different chromosome numbers are capable of interbreeding, albeit with greatly impaired fertility at the F_2 generation (Yosida 1980).

Although other studies have detected likely instances of hybridization and introgression between chromosomally distinct populations of *R. rattus* (Yosida 1980; Chinen et al. 2005; Bastos et al. 2011), the extent of genetic introgression in any one case has not been documented. Our study demonstrates that a population of roof rats in the SF Bay Area carries mtDNA of two historically divergent lineages (Lineages I and II) which elsewhere show close association with different karyotypes—2N = 38 and 2N = 42, respectively (Yosida 1980). By using microsatellite DNA markers, we have also demonstrated that this population shows no genetic subdivision, in contrast with microsatellite DNA differentiation between populations of roof rats from other parts of the world, including animals from within the known ranges of Lineages I and II (Aplin et al. 2011). The clear implication is that Californian roof rats are the product of introgression between two genetically divergent populations of roof rats, with the result that they now constitute a single gene pool.

Outside of North America, the microsatellite DNA analysis revealed only three distinct genetic groups, distinguishing Madagascan roof rats (mtDNA Lineage I) from those of Asian origin (mtDNA Lineages II and IV), or where one more group is defined, further separating Philippine rats (mtDNA Lineages II and IV) from mainland Asian rats (China and Vietnam—mtDNA Lineage II). Madagascan roof rat populations show limited mitochondrial diversity with the majority of haplotypes belonging to a discrete sublineage of RrC Lineage I, with Hd = 0.111 and cyt b nucleotide diversity = 0.00015 (Table 2). This sublineage shares a common ancestral haplotype with the European 'ship rat' cluster (Aplin et al. 2011) and both populations are probably derived from a common stock that invaded the Middle East from western India in the late Pleistocene or early Holocene (Aplin et al. 2011). Haplotypes of the 'ship rat' clade were spread worldwide on board ships during the Age of Exploration, and in more recent times by numerous means. Genetic differentiation of the Californian and Madagascan populations might be due to strong filtering and founder effects operating during the early dispersal history of the émigré roof rat population. Alternatively, it might be due to a novel admixture of Asian and ship rat genetic components in the case of the Californian population. In contrast, Asian roof rats appear to show genetic uniformity among a large geographic area including China, Vietnam, Indonesia and the Philippines. Mitochondrial diversity within Lineage II is high in the East Asian region (Aplin et al. 2011) but similar to the situation in Lineage I, the Asian lineage includes several mtDNA clusters that have undergone major geographic range expansions into the western Pacific. Our sampling of Lineage II is dominated by populations carrying one of the sublineages (IIA of Aplin et al. 2011), hence the low genetic diversity suggested by the microsatellite DNA is not surprising given the inferred population history. Based on these considerations, we suspect that the pattern of genetic variation revealed by the present study of microsatellite DNA will prove to be unrepresentative of broader patterns of diversity. Further to this, we note that a recent investigation of house mouse phylogeography in western Europe (Jones et al. 2010) also found a poor geographic correspondence between mitochondrial and microsatellite DNA, with the former apparently tracking the historical routes of introductions and the latter showing regional population clusters only. Jones et al. (2010) suggest a possible dichotomy in the genetic components between original colonisation and subsequent genetic

contamination but also note that low natural dispersal capacity and infrequent long distance translocations of house mice could generate the strong regionalism in microsatellite DNA alleles.

For the present, we interpret our microsatellite DNA results as an effective demonstration of introgression between RrC Lineage I (aka *R.rattus*) and RrC Lineage II (aka *R. tanezumi*) but of little value for inferring the wider population genetic history of the RrC. This situation may change with more extensive genotyping of microsatellite DNA, nDNA SNP's, and including populations within the natural ranges as well as regions of prehistoric to recent introduction.

The taxonomy of SF Bay Area roof rats

Our results demonstrate that there is *only one species of roof rat in the SF Bay Area*. However, they also demonstrate that *this population is of mixed genetic origin*. What taxonomic name should be applied to this population is not immediately clear. The International Commission on Zoological Nomenclature (2000 and earlier editions) explicitly avoids the issue of animals of hybrid origin and various auxillary systems have been proposed to deal with this thorny issue. In the present case, what is clear is that designation of SF Bay Area rats to RrC Lineage I and II (or to *R. rattus* and *R. tanezumi*) based solely on mtDNA gene sequences is inaccurate and misleading, as both mtDNA lineages are present within a single, freely interbreeding gene pool. One option is to refer to them as hybrids (e.g. RrC Lineage I X RrC Lineage II); however, as noted above, the currently available genetic data do not allow us to estimate the degree of genetic admixture. At one extreme, the genetic composition of SF Bay Area rats might prove to be near identical with 'ship rats' of European origin and the presence of the RrC Lineage II haplotypes might be an example of 'allele surfing' sensu Excoffier and Lischer (2010). Alternatively, the SF Bay Area rats might be genuinely mixed in their genetic composition, with significant components derived from immigrant rats of Asian origin. In the former case, we might be justified in calling them *R. rattus* and treating them as little different (except in their mtDNA composition) from other populations derived from ship rats. In the latter case, it might be preferable to employ a terminology that better reflects their chimeric genetic identity.

As demonstrated by recent genetic work on the house mouse—a species with a remarkably parallel history of allopatric differentitation, prehistoric range expansion and dispersal, and genetic admixture (Boursot et al. 1993)—it is possible to estimate both the genetic composition (Abe et al. 2004; Sakai et al. 2005; Frazer et al. 2007) and the hybridization history (Nunome et al. 2010) of admixed populations given sufficient genetic data. Until such time as this becomes available for the roof rat group, we recommend that the SF Bay Area rats be referred to as members of the *R. rattus* Complex (sensu Aplin et al. 2011) or *R. rattus* sensu lato.

Invasive species history

Our results suggest that there is still much to learn regarding the nature and timing of historical rodent invasions into North America. In the case of the roof rat group, further work is now required to clarify the historical events that lead to the genetically mixed SF Bay population of roof rats. The data currently in hand tell us that SF Bay Area roof rats are of mixed ancestry but do not allow us to discriminate among various possible scenarios, including (1) genetic admixture prior to the arrival of roof rats in California; and (2) historically separate invasions of Lineage I and II rats with distinct genetic profiles, followed by in situ hybridization and backcrossing, resulting in a single effective gene pool. To some extent, these alternative scenarios might be tested by investigation of historical museum specimens of roof rats from the SF Bay Area and other localities in the US. However, we caution that the true history could well be more complex than suggested by this simple dichotomy—potentially involving multiple introductions at various times over the past 200 years, and multiple phases of genetic introgression. Indeed, given the proven success of roof rats in stowaway dispersal on board ships and potentially also on airplanes, a more complex history intuitively seems more likely than a less complex one. Further work on this problem obviously needs to include detailed genetic studies of contemporary populations, using highly variable genetic markers (e.g. linked and unlinked nDNA SNP's), larger sample sizes and more dispersed populations (c.f. Bronnenhuber et al. 2011). Other *R. rattus* populations in the US that include both Lineages I and II have recently been discovered in Florida (Lack et al. 2012) and southern

California (Thiemann et al. 2012). However, it should also extend to genotyping of historical specimens, both as a way of testing hypotheses derived from the population genetics and suggesting others that might otherwise have escaped consideration.

Another murid rodent with a potentially parallel invasion history in California is the house mouse. While most *Mus* in North America are *M. musculus domesticus* of western European origin, a small pocket of *M. m. castaneus* exists in southern California at Lake Casitas. The *castaneus* subspecies of house mice is of Asian origin and extends today from the Middle East through to Japan and the Philippines (Yonekawa et al. 2003). The Californian population has been studied extensively for its genomic retroviruses (Gardner et al. 1991) but its origin within the broad Asian distribution of *M. m. castaneus* has not been pinpointed. This population is known to be hybridizing with the more common *M. m. domesticus* (Orth et al. 1998). The obvious parallels between Californian populations of *R. rattus* and *M. musculus*, each of which support populations of European and Asian origin, invites a comparative study of their invasion histories and their outcomes.

Invasive species biology

Invasive species by their nature exist in a non-natural environment and they can show novel biological responses to unfamiliar environmental pressures. Where multiple introductions of closely related taxa occur, as in the case of the roof rats, there is a further novel juxtaposition of genetic groups that, within their native ranges, do not normally have the opportunity to interact either genetically or ecologically. Novel outcomes are to be anticipated, including higher than usual levels of genetic diversity, unprecedented recombinants, and new adaptations in response to novel environmental and community pressures (Gomulkiewicz et al. 2010). Mixing of genetic components from divergent evolutionary lineages of roof rats might produce a blending of their ecological capabilities. Although several studies have been conducted on SF Bay Area *Rattus* ecology (e.g. Brooks 1966; Brothers 1972; Dutson 1973; Recht 1988; Stroud 1982), *Rattus* community ecology has been relatively under-studied in the recent literature. The distribution of *Rattus* in some parts of the SF Bay Area is surely influenced by that of the native woodrats (*Neotoma fuscipes*), most likely by resource competition. JLP and others (W. Z. Lidicker, pers. comm.) noted the increase of *R. rattus* in Tilden Regional Park in Alameda County during annual trapping for a field course over the past two decades. The possibility that genetic components derived from RrC Lineage II (aka *R. tanezumi*) play a role in this *Rattus-Neotoma* community interaction warrants examination.

On a more practical level, higher genetic diversity due to unique combinations of alleles or novel recombinants among roof rats could enhance a population's ability to withstand rodenticides. For example, it is well documented that there is a genetic basis for resistance to anticoagulant compounds such as warfarin by repression of the vitamin K reductase reaction (VKOR; Rost et al. 2004). Mutations at this locus can lead to resistance through structural changes in the VKORC1 protein, but may also be due to other mechanisms for blood clotting (Pelz et al. 2005; Rost et al. 2004). Mutations in cytochrome P450 are also known to effect metabolism of warfarin through increased CYP3A2 expression in warfarin-resistant rats (Ishizuka et al. 2007). While we have not tested SF Bay Area rats for these kinds of functional mutations, we suggest that invasive species with complex genetic backgrounds may possess unique combinations of genes that potentially confer resistance to a number of rodenticides. Conversely, gene flow between these disparate groups may also break up gene combinations that in the past may have been selected for such properties. Gene flow between rats from regions without a prior selective background with those exhibiting resistance to rodenticides can result in a selective disadvantage.

Disease issues

Rattus species are known to carry a wide variety of pathogenic micro-organisms (e.g., *Bartonella,* Ellis et al. 1999). Aplin et al. (2011) postulated that the exceptional diversity of such organisms in the roof rat group might be due in part to its originally wide natural distribution that included both South and East Asia, and Indochina. However, given the close evolutionary relationship between the various lineages of roof rats and apparent ability to interbreed in the wild, it is likely that disease strains that formed their initial association with one lineage may be transferrable to another, or to hybrids between the two. This may

obscure our understanding of the co-evolution of microorganisms and their hosts, which might already be complicated by host transfer in evolutionary time scales (e.g. hantaviruses, Plyusnina et al. 2009). In the SF Bay Area, some of the rats we analyzed were tested for *Bartonella* and *Borrelia*. Several individuals were found to harbor *Bartonella coopersplainsensis*, only recently described from *Rattus leucopus* in Australia (Gundi et al. 2009), and several rats harbored *Borrelia bissetti* (R. Lane, UC Berkely, pers. comm.). Rats in this area are also known to carry *Borrelia burgdorferi* (Peavey et al. 1997). More basic distribution information might suggest a connection between rats in Australia and the SF Bay Area, perhaps through a common source area. More generally, the complex genetic history and composition of roof rats in California raises interesting questions regarding the possible historical interchange of disease agents, and the possible ongoing role of genetic diversity among roof rats in modulating disease ecology.

Acknowledgments This work benefitted from support from Denise Bonilla of the California Department of Public Health, and fieldwork by Tia Reddick, an intern with the Alameda County Vector Control District. Michael Kosoy (CDC) tested tissues for *Bartonella* strains, and Robert Lane (UC Berkeley) tested tissues for *Borrelia*. Cherilyn Acorda conducted some of the early laboratory work. This work was improved by discussions with Steve Donnellan, Ron Van Den Bussche, Hitoshi Suzuki, Jean-François Cosson and others at the first *R. rattus* consortium meeting in Bangkok in 2010. Thanks to curators and collection managers Joseph Cook and Jon Dunnum (Museum of Southwestern Biology), Larry Heaney, Bruce Patterson and Bill Stanley (Field Museum of Natural History) for access to worldwide samples of *Rattus* tissues. This work would not have been possible without those loans. Many thanks to Anita Pearson who personally donated many of the rats included here, especially where Lineages I and II are sympatric. Funding for laboratory analyses was kindly provided by the MVZ, and the Alameda County Vector Control District directly funded PK.

References

Abe K, Noguchi H, Tagawa K, Yuzuriha M, Toyoda A, Kojima T, Ezawa K, Saitou N, Hattori M, Sakaki Y, Moriwaki K, Shiroishi T (2004) Contribution of Asian mouse subspecies *Mus musculus molossinus* to genomic constitution of strain C57BL/6 J, as defined by BAC-end sequence-SKIP analysis. Genome Res 14:2439–2447

Aplin KP, Brown PR, Jacob J, Krebs C, Singleton GR (2003) Field methods for rodent studies in Asia and the Pacific. Australian Centre for International Agricultural Research, Canberra, p 223

Aplin KP, Suzuki H, Chinen AA, Chesser RT, ten Have J, Donnellan SC, Austin J, Frost A, Gonzalez JP, HerbreteauV, Catzeflis F, Soubrier J, Fang Y-P, Robins J, Matisoo-Smith E, Bastos ADS, Maryanto I, Sinaga MH, Denys C, Van Den Bussche RA, Conroy C, Rowe K, Cooper A (2011) Multiple geographic origins of commensalism and complex dispersal history of black rats. PLoS One 6(11):e26357

Atkinson IAE (1985) The spread of commensal species of *Rattus* to oceanic islands and their effects on island avifaunas. In: Moors PJ (ed) Conservation of island birds. ICBP Technical Publication no. 3, pp 35–81

Azad AF (1990) Epidemiology of murine typhus. Ann Rev Entomol 35:553–569

Bastos AD, Nair D, Taylor PJ, Brettschneider H, Kirsten F, Mostert E, von Maltitz E, Lamb JM, van Hoof P, Belmain SR, Contrafatto G, Downs S, Chimimba CT (2011) Genetic monitoring detects an overlooked cryptic species and reveals the diversity and distribution of three invasive *Rattus* congeners in South Africa. BMC Genet 12:26

Boisier P, Rahalison L, Rasolomaharo M, Ratsitorahina M, Mahafaly M, Razafimahefa M, Duplantier JM, Ratsifasoamanana L, Chanteau S (2002) Epidemiologic features of four successive annual outbreaks of bubonic plague in Mahajanga, Madagascar. Emerg Infect Dis 8:311–316

Boursot P, Auffray JC, Britton-Davidian J, Bonhomme F (1993) The evolution of house mice. Annu Rev Ecol Syst 24: 119–152

Bronnenhuber J, Dufour B, Higgs D, Heath D (2011) Dispersal strategies, secondary range expansion, and invasion genetics of the non-indigenous Round goby, *Neogobius melanostomus*, in Great Lakes tributaries. Mol Ecol 20:1845–1859

Brooks JE (1966) Roof rats in residential areas—the ecology of invasion. Calif Vector Views 13:69–74

Brothers DR (1972) An innovation in Roof Rat control. In: Proceedings of the 5th Vertebrate Pest Conference. Fresno, p 161

Brown GG, Simpson MV (1981) Intra- and interspecific variation of the mitochondrial genome in *Rattus norvegicus* and *Rattus rattus*: restriction enzyme analysis of variant mitochondrial DNA molecules and their evolutionary relationships. Genetics 97:125–143

Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. Mol Biol Evol 24:621–631

Chiba S (2010) Invasive rats alter assemblage characteristics of land snails in the Ogasawara Islands. Biol Conserv 143:1558–1563

Chinen AA, Suzuki H, Aplin KP, Tsuchiya K, Suzuki S (2005) Preliminary genetic characterization of two lineages of black rats (*Rattus rattus* sensu lato) in Japan, with evidence for introgression at several localities. Genes Genet Syst 80:367–375

Claffey DP, Madon MB, Smith RT (1986) An integrated pest management approach to roof rat control in oceanfront riprap, Ventura County, California. In: Salmon TP (ed) Proceedings of the 12th Vertebrate Pest Conference. University of California, Davis, p 60

Donaldson HH (1915) The rat: reference tables and data for the albino rat (*Mus norvegicus albinus*) and the Norway rat (*Mus norvegicus*). Wistar Institute, Philadelphia

Dutson VJ (1973) Use of the himalayan blackberry, *Rubus discolor*, by the roof rat, *Rattus rattus*, in California. Calif Vector Views 20:59–68

Dutson VJ (1974) The association of the roof rat (*Rattus rattus*) with the Himalayan Blackberry (*Rubus discolor*) and Algerian Ivy (*Hedera canariensis*) in California. In: Proceedings of the 6th Vertebrate Pest Conference, Anaheim, pp 41–48

Ecke DH (1964) Roof rat populations in Santa Clara County, California. In: Proceedings of the 2nd Vertebrate Pest Control Conference, Anaheim, p 99

Ellis BA, Regnery RL, Beati L, Bacellar F, Rood M, Glass GG, Marston E, Ksiazek TG, Jones D, Childs JE (1999) Rats of the genus *Rattus* are reservoir hosts for pathogenic *Bartonella* species: an Old World origin for a New World disease? J Infect Dis 180:220–224

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Res 10:564–567

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Frazer KA, Eskin E, Kang HM, Bogue MA, Hinds DA, Beilharz EJ, Gupta RV, Montgomery J (2007) A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. Nature 448:1050–1053

Gardner MB, Kozak CA, O'Brien SJ (1991) The Lake Casitas wild mouse: evolving genetic resistance to retroviral disease. TIG 7:22–27

Gomulkiewicz R, Holt RD, Barfield M, Nuismer SL (2010) Genetics, adaptation, and invasion in harsh environments. Evol Appl 3:97–108

Goodman SM (1995) *Rattus* on Madagascar and the dilemma of protecting the endemic rodent fauna. Conserv Biol 9:450–453

Grinnell J, Dixon J, Linsdale JM (1930) Vertebrate natural history of a section of northern California through the Lassen Peak region. UC Publ Zool 35:1–526

Gundi VAKB, Taylor C, Raoult D, La Scola B (2009) *Bartonella rattaustraliani* sp. nov., *Bartonella queenslandensis* sp. nov. and *Bartonella coopersplainsensis* sp. nov., identified in Australian rats. Int J Syst Evol Microbiol 59:2956–2961

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361–372

Holm S (1979) A simple sequentially rejective multiple test procedure. Scand J Stat 6:65–70

Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogeny. Bioinformatics 17:754–755

International Commission on Zoological Nomenclature (1985) International code of zoological nomenclature, 4th edn. International Trust for Zoological Nomenclature

Ishizuka M, Okajima F, Tanikawa T, Min H, Tanaka KD, Sakamoto KQ, Fujita S (2007) Elevated warfarin metabolism in warfarin-resistant roof rats (*Rattus rattus*) in Tokyo. Drug Metab Dispos 35:62–66

Jackson WB (1982) Norway rat and allies. In: Chapman JA, Feldhamer GA (eds) Wild mammals of North America. John Hopkins University Press, Baltimore, pp 1077–1088

Jacob HJ, Brown DM, Bunker RK, Daly MJ, Dzau VJ, Goodman A, Koike G, Kren V, Kurtz T, Lernmark Å, Levan G, Mao Y-P, Pettersson A, Pravenec M, Simon JS, Szpirer C, Szpirer J, Trolliet MR, Winer ES, Lander ES (1995) A genetic linkage map of the laboratory rat, *Rattus norvegicus*. Nat Genet 9:63–69

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806

Jansa SA, Barker FK, Heaney LR (2006) The pattern and timing of diversification of Philippine endemic rodents: evidence from mitochondrial and nuclear gene sequences. Syst Biol 55:73–88

Jones EP, van der Kooij J, Solheim R, Searle JB (2010) Norwegian house mice (*Mus musculus musculus*/*domesticus*): distributions, routes of colonization and patterns of hybridization. Mol Ecol 19:5252–5264

Kaleme PK, Bates JM, Belesi HK, Bowie RCK, Gambalemoke M, Kerbis-Peterhans J, Michaux J, Mwanga JM, Ndara BR, Taylor PJ, van Vuuren BJ (2011) Origin and putative colonization routes for invasive rodent taxa in the Democratic Republic of Congo. Afr Zool 46:133–145

Lack JB, Greene DU, Conroy CJ, Hamilton MJ, Braun JK, Mares MA, Van den Busche RA (2012) Invasion facilitates hybridization with introgression in the *Rattus rattus* species complex. Mol Ecol 21:3545–3561

Lantz DE (1909) The brown rat in the United States. US Dept Agric Biol Surv Bull 33:1–54

Lee CE (2002) Evolutionary genetics of invasive species. TREE 17:386–391

Loiseau A, Rahelinirina S, Rahalison L, Konečný A, Duplantier JM, Brouat C (2008) Isolation and characterization of microsatellites in *Rattus rattus*. Mol Ecol Res 8:916–918

Maddison DR, Maddison WP (2005) MacClade 4: analysis of phylogeny and character evolution. Version 4.08a. http://macclade.org

Meerburg BG, Singleton GR, Leirs H (2009) The year of the rat ends—time to fight hunger! Pest Manag Sci 65: 351–352

Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16:1215

Musser GG, Carleton MD (2005) Family Muridae. In: Wilson DE, Reeder DM (eds) Mammal species of the world: a taxonomic and geographic reference. Smithsonian Institution Press, Washington, pp 894–1531

Nolte D, Bergman D, Townsend J (2003) Roof rat invasion of an urban desert island. In: Singleton GR, Hinds LA, Krebs CJ, Spratt DM (eds) Rats, mice and people: rodent biology and management. Australian Centre for International Agricultural Research, Canberra, p 481

Nunome M, Ishimori C, Aplin KP, Yonekawa H, Moriwaki K, Suzuki H (2010) Detection of recombinant haplotypes in wild mice (*Mus musculus*) provides new insights into the origin of Japanese mice. Mol Ecol 19:2474–2489

Nylander JAA (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University

Orth A, Adama T, Din W, Bonhomme F (1998) Hybridation naturelle entre deux sous-espèces de souris domestique, *Mus musculus domesticus* et *Mus musculus castaneus*, près du lac Casitas (Californie). Genome 41:104–110

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

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Peavey C, Lane R, Kleinjan J (1997) Role of small mammals in the ecology of *Borrelia burgdorferi* in a peri-urban park in north coastal California. Exp Appl Acarol 21:569–584

Pelz H-J, Rost S, Hünerberg M, Fregin A, Heiberg A-C, Baert K, MacNicoll AD, Prescott CV, Walker A-S, Oldenburg J, Müller CR (2005) The genetic basis of resistance to anticoagulants in rodents. Genetics 170:1839–1847

Pennant T (1781) History of quadrupeds, vol. 2. London. Rats, p 438

Pimentel D, Zuniga R, Morrison D (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecol Econ 52:273–288

Plyusnina A, Ibrahim I-N, Winoto I, Porter KR, Gotama IBI, Lundkvist A, Vaheri A, Plyusnin A (2004) Identification of Seoul Hantavirus in *Rattus norvegicus* in Indonesia. Scand J Infect Dis 36:356–359

Plyusnina A, Ibrahim I-N, Plyusnin A (2009) A newly recognized hantavirus in the Asian house rat (*Rattus tanezumi*) in Indonesia. J Gen Virol 90:205–209

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249

Recht MA (1988) Biology of domestic rats: telemetry yields insights for pest control. In: Crabb AC, Marsh RE (eds) Proceedings of the 13th Vertebrate Pest Conference, University of California, Davis, pp 98–100

Robins JH, Hingston J, Matisoo-Smith E, Ross HA (2007) Identifying *Rattus* species using mitochondrial DNA. Mol Ecol Notes 7:717–729

Rosenburg NA (2004) *Distruct*: a program for the graphical display of population structure. Mol Ecol Notes 4:137–138

Rost S, Fregin A, Ivaskevicius V, Conzelmann E, Hortnagel K, Pelz HJ, Lappegard K, Selfried E, Scharrer I, Tuddenham EGD et al (2004) Mutations in VKORC1 cause warfarin resistance and multiple coagulation factor deficiency type 2. Nature 427:537–541

Rowe KC, Reno ML, Richmond DM, Adkins RM, Steppan SJ (2008) Pliocene colonization and adaptive radiations in Australia and New Guinea (Sahul): multilocus systematics of the old endemic rodents (Muroidea: Murinae). Mol Phylogenet Evol 47:84–101

Rozas J, Sánchez-Delbarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496–2497

Sakai T, Kikkawa Y, Miura I, Inoue T, Moriwaki K, Shiroishi T, Satta Y, Takahata N, Yonekawa H (2005) Origins of mouse inbred strains deduced from whole-genome scanning by polymorphic microsatellite loci. Mamm Genome 16:11–19

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol 18:233–234

Scott S, Duncan CJ (2001) Biology of plagues: evidence from historical populations. Cambridge University Press, Cambridge

Silver J (1927) The introduction and spread of house rats in the United States. J Mammal 8:58–60

Smith KF, Carpenter SM (2006) Spread of exotic black rat (*Rattus rattus*) parasites to endemic deer mice (*Peromyscus maniculatus*) on the California Channel Islands. Divers Distrib 12:742–748

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690

Stroud DC (1982) Dynamics of *Rattus rattus* and *R. norvegicus* in a riparian habitat. J Mammal 63:154–155

Suzuki H, Tsuchiya K, Takezaki N (2000) A molecular phylogenetic framework for the Ryukyu endemic rodents *Tokudaia osimensis* and *Diplothrix legata*. Mol Phylogenet Evol 15:15–24

Swofford DL (2002) PAUP*. Phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland

Thiemann TC, Lemenager DA, Kluh S, Carroll BD, Lothrop HD, Reisen WK (2012) Spatial variation in host feeding patterns of *Culex tarsalis* and the *Culex pipiens* complex (Diptera: Culicidae) in California. J Med Entomol 49:903–916

Tollenaere C, Brouat C, Duplantier JM et al (2010) Phylogeography of the introduced species *Rattus rattus* in the western Indian Ocean, with special emphasis on the colonization history of Madagascar. J Biogeogr 37:398–410

Towns DR, Atkinson IAE, Daugherty CH (2006) Have the harmful effects of introduced rats on islands been exaggerated? Biol Invasions 8:863–891

Truong TT, Yoshimatsu K, Araki K, Lee BH, Nakamura I, Endo R, Shimizu K, Yasuda SP, Koma T, Taruishi M, Okumura M, Truong UN, Arikawa J (2009) Molecular epidemiological and serological studies of hantavirus infection in northern Vietnam. J Vet Med Sci 71:1357–1363

Weir BS, Cockerham CC (1984) Estimating *F*-Statistics for the analysis of population structure. Evolution 38:1358–1370

Wilgenbusch JC, Warren DL, Swofford DL (2004) AWTY: a system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference, http://king2.scs.fsu.edu/CEBProjects/awty/awty_start.php

Wolfs TFW, Wagenaar JA, Niesters HGM, Osterhaus ADME (2002) Rat-to-human transmission of cowpox infection. Emerg Infect Dis 8:1495–1496

Yonekawa H, Tsuda K, Tsuchiya K, Yakimenko L, Korobitsyna K, Chelomina GN, Spiridonova LN, Frisman L, Kryukov A, Moriwaki K (2003) Genetic diversity, geographic distribution and evolutionary relationships of *Mus musculus* subspecies based on polymorphisms of mitochondrial DNA. In: Kryukov A, Yakimenko L (eds) Problems of evolution, vol. 5 Vladivostok, Dalnauka, pp 90–108

Yosida TH (1980) Cytogenetics of the black rat: karyotype evolution and species differentiation. University of Tokyo Press, Tokyo

Yosida TH, Kato H, Tsuchiya K, Moriwaki K (1971) Karyotype and serum transferrin patterns of hybrids between Asian and Oceanian black rats, *Rattus rattus*. Chromosoma 34:40–50

Zeiner DC, Laudenslayer WF Jr, Meyer KE, White M, (eds) (1990) California's Wildlife. Volume III: Mammals. State of California. The Resources Agency, Sacramento