POPULATION GENETIC ANALYSIS OF INVASIVE *RATTUS*: IMPLICATIONS FOR EVOLUTIONARY BIOLOGY, DISEASE ECOLOGY AND INVASION BIOLOGY

By

JUSTIN B. LACK

Bachelor of Science in Biology
University of Central Oklahoma
Edmond, OK
2004

Master of Science in Biology
University of Central Oklahoma
Edmond, OK
2007

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 2012

POPULATION GENETIC ANALYSIS OF INVASIVE *RATTUS*: IMPLICATIONS FOR EVOLUTIONARY BIOLOGY, DISEASE ECOLOGY AND INVASION BIOLOGY

Dissertation Approved:

Dr. Ronald A. Van Den Bussche

Dissertation Adviser

Dr. Janet Braun

Dr. Meredith Hamilton

Dr. Andy Dzialowski

Dr. Andrew Doust

Outside Committee Member

Dr. Sheryl A. Tucker

Dean of the Graduate College

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

CHAPTER I

INTRODUCTION

As human populations have increased and the difficulty of "globally-spanning" transport and commerce has decreased, the frequency with which exotic species have invaded novel ecosystems has increased dramatically. As a result, the field of invasion ecology has grown in importance, with an increasing need for answers to the questions surrounding species invasions. What are the main factors determining an ecosystems' vulnerability to invasion and/or a species' ability to invade and establish? How many individuals are needed to establish a viable population? Which species are invading? What vectors and pathways are being utilized by individual invasive species? How much transport/migration is actually occurring on different geographical scales (i.e., within and among states, inter- and intracontinentally, etc.)? Although answers to these questions have genetic components, the genetics and evolution of invasive species have been neglected in the majority of research in favor of ecology (Lee 2002; Sakai et al. 2001). Coalescent predictions, phylogeographical principles and population genetics can provide a framework for making inferences concerning all of the unknowns mentioned above.

Invasive species are second only to habitat destruction in terms of the most significant ecological destructive forces in the world (Wilson 1997). There are numerous examples in the literature of species invasions having negative impacts on native flora and fauna, often

contributing to local extinctions (Elton 1958; Ricciardi et al. 1998; Wilcove et al. 1998; Williamson 1996), and there is evidence that this trend may continue until essentially all natural ecosystems have been affected by invasive species (Sala et al. 2000). Furthermore, although eradication of invasive species is extremely difficult, especially on large geographical scales (Myers et al. 2000), population genetics provides a powerful tool for improving the effectiveness and sustainability of species eradication plans (Abdelkrim et al. 2005a). If small-scale eradication plans (i.e., for individual cities) are to be effective, it is vital that we understand from where invasive species are entering these cities so that repeat invasions can be prevented, and exterminators can deal with individual isolated populations.

In terms of invasive species in the U.S., the Norway rat (*Rattus norvegicus*) and the ship rat (*R. rattus*) sit atop the list of the most successful. Of the $130 billion dollars damage done each year by invasive species, a number exceeding the annual cost of the first Iraq War (http://nationalpriorities.org/index.php?option=com_wrapper&Itemid=182), *R. norvegicus* and *R. rattus* are responsible for approximately $19 billion (Pimentel et al. 2000). Therefore, understanding the transportation vectors and pathways of two of the most common and destructive invasive species in the U.S. is an important issue both environmentally and economically, and can have implications in developing effective eradication plans (Abdelkrim 2005a).

The historic role of *Rattus* species in the spread of human disease is unparalleled among vertebrates. Some of the deadliest pathogens in human history are carried and transmitted by *R. rattus* and *R. norvegicus*, including plague, *Salmonella*, schistosomiasis, and trichinosis (Gratz 1984). Historically, bubonic plague (*Yersinia pestis*) has been the most deadly of the rat borne pathogens, with multiple waves of infection having occurred across the globe in the last millennium (Curson and McCracken 1986; Grzmik 1975; Kingdon 1974; Lowery 1974; Scott et al. 1996; Shrewsbury 1970). The most well known of the plague epidemics occurred in Europe

during the 1300s, killing 20 to 30 million people, approximately 25% of Europe's population at the time (Slack 1989). And, although plague has been regarded as a historical disease no longer of contemporary importance, recent surges of infection have occurred in India (Dennis 1994) and Madagascar (Rasolomaharo et al. 1995), indicating the risk of plague infection and spread still exists. Furthermore, with the risk of terrorism in the U.S. at an elevated level, there has been concern about the possibility of resistant strains of *Y. pestis* being used as bioterrorism weapons, with rats acting as vectors (Inglesby et al. 2000). Therefore, knowledge of migration patterns of two major plague vectors (*R. rattus* and *R. norvegicus*) would be extremely useful in not only predicting the pattern of spread of the disease, but also in preventing further spread.

Due to a close association with humans, both *R. rattus* and *R. norvegicus* have been prolific invaders, with *R. rattus* occurring on all continents and *R. norvegicus* occurring on every continent except Antarctica (Nowak 1999; Musser and Carleton 2005). In North America, *R. rattus* is restricted to the south-central and coastal contiguous 48 states, NE Canada, and the provinces of British Columbia and Alberta in SW Canada, while *R. norvegicus* occurs throughout the country, including Alaska and Hawaii, and much of Canada. *R. rattus* are native to south-central Asia, on the Indian Peninsula (Niethammer 1975), and are believed to have come to the New World with Christopher Columbus in 1492 (Armitage 1993). Their numbers in coastal areas can grow quite large due to a close association with ships and, in the U.S., are more common in upper stories of buildings and in trees, rarely occurring on the ground or in open fields (Caire et al. 1989; Schwartz and Schwartz 1981). *Rattus norvegicus,* on the other hand, is a habitat generalist, occurring essentially anywhere humans are found. Suitable habitat includes the ground floors of buildings and dwellings, around sewers and garbage dumps, and can even be found in open fields around rural areas (Caire et al. 1989; Schwartz and Schwartz 1981). The native range for the Norway rat is northern central Asia (Dobson 1994; Jones and Johnson 1965; Kawamura 1989; Kowalski and Hasegawa 1976), and it is believed to have invaded the U.S. at approximately the time of the American Revolution (Armitage 1993; Caire et al. 1989).

Due to the extensive use of *R. norvegicus* in laboratory experimentation, there is an abundance of information available regarding their physiology and genomics. However, in spite of this and the fact that *R. rattus* and *R. norvegicus* have had such a dramatic impact on the U.S. economy, world health, and natural ecosystems, relatively little is known about their genetic structure at a large geographic scale. Even though population genetic studies have been conducted, the majority of these examined small populations occurring on islands (Abdelkrim et al. 2005a; Abdelkrim 2005b; Cheylan 1998; Chinen et al. 2005; Hingston et al. 2005; Patton et al. 1975; Voigt et al. 2000), and I am unaware of any geographically large-scale studies in the primary literature. These studies are useful in understanding invasion and establishment over small geographical scales, but the results may not be applicable to managing invasive *Rattus* across large, complex landscapes in countries with intricate transportation infrastructures, such as the U.S.

There is little known of the frequency with which each *Rattus* species is entering the U.S. In terms of eradication, information concerning which species has/is/are invading the U.S. most frequently would be valuable in developing plans for removal directed at the appropriate species. A species-specific eradication plan would be especially useful in managing invasive *Rattus* due to differences in ecology and habitat affinities. Furthermore, there are several additional species of *Rattus* that have successfully invaded other areas of the world, most notably the Polynesian rat (*R. exulans*) in the south Pacific--including Hawaii--and the Oriental house rat (*R. tanezumi*) in southeast Asia and, most recently, California (Aplin et al. 2011). Due to extensive trade between the U.S. and eastern Asian countries, these species may have had the opportunity to invade the U.S. but have either failed to establish reproducing populations or have successfully invaded but have gone undetected.

Invasive *Rattus*, like most invasive species, present a unique opportunity to study evolution, adaptation, and natural selection in novel environments (Ellstrand and Schierenbeck 2000; Lee 2002; Quinn et al. 2000; Sax et al. 2005). Furthermore, understanding the levels of

diversity present in established *Rattus* populations may be an important factor when one considers the principle that a species' rate of change under varying selection pressures is directly proportional to the standing genetic variation (Fisher 1930). Based on this concept, knowledge about the amount of genetic diversity present within invasive species may provide some basis upon which to make predictions concerning invasibility as well as establishment success. Also, many species invasions are accompanied by a genetic bottleneck due to the number of initial colonists typically being small and from a single subpopulation within the source population (Nei et al. 1975). There are multiple examples where invasive species have significantly reduced genetic variation relative to their source populations (Abdelkrim et al. 2005a; Baker & Moeed 1987; Merila et al. 1996). Such a reduction in overall genetic diversity can result in inbreeding depression (Ellstrand and Elam 1993; Newman and Pilson 1997; Nieminen et al. 2001) and/or a loss in the ability of the species to evolve and adapt to a novel environment (Nei et al. 1975; Sakai et al. 2001). Furthermore, it has been suggested that the lag time preceding establishment for many invasive species may reflect the time period needed to acquire sufficient levels of additive genetic variance, and not necessarily the time required for reaching a critical population size (Ellstrand and Schierenbeck 2000). Goodwin et al. (1994) showed that only following the introduction of novel genetic strains of the fungus was *Phytophthora infestans* able to expand their non-native ranges. However, as more population genetic studies are being performed on invasive species, there appear to be numerous exceptions to this trend of reduced genetic diversity relative to the source population (Novak and Mack 2005; Wares et al. 2005). There are a number of conditions that could be responsible for actually elevating diversity in invasive species. If invasion occurs from multiple, genetically diverse and somewhat distinct populations, introductions maintain a high frequency over time, and invasion is followed by a rapid range expansion, diversity can be equal to or even greater than that of the source population (for examples see Novak & Mack 1993; Martel et al. 2004). Furthermore, hybridization with native can similarly result in a rapid accumulation of genetic diversity following a genetic bottleneck.

References

Abdelkrim. J, Pascal M, Calmet C, Samadi S (2005a) Importance of assessing population genetic structure before eradication of invasive species: examples from insular Norway rat populations. *Conservation Biology*, **19**, 1509–1518.

Abdelkrim. J, Pascal M, Calmet C, Samadi S (2005b) Island colonization and founder effects: the invasion of the Guadeloupe islands by ship rats (*Rattus rattus*). *Molecular Ecology*, **14**, 2923–2931.

Aplin KP, Suzuki H, Chinen AA *et al.* (2011) Multiple genetic origins of commensalism and complex dispersal history of black rats. *PLoS One*, **6**, e26357.

Armitage PL (1993) Commensal rats in the New World, 1492-1992. *Biologist*, **40**, 174–178.

Baker AJ, Moeed A (1987) Rapid genetic differentiation and founder effect in colonizing populations of common mynas (*Acridotheres tristis*). *Evolution*, **4**, 525–538.

Caire W, Tyler JD, Glass BP, Mares MA (1989) *Mammals of Oklahoma*. University of Oklahoma Press, Norman, Oklahoma.

Cheylan G (1998) Distribution of genetic diversity within and between Mediterranean island populations of the black rat *Rattus rattus* (L. 1758). *Biological Journal of the Linnean Society*, **63**, 393–408.

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 and Genetic Systems*, **80**, 367–375.

Curson P, McCracken P (1986) *Plague in Sydney. The Anatomy of an Epidemic*. New South Wales University Press, Kensington, New South Wales.

Dennis DT (1994) Plague in India. *British Medical Journal*, **309**, 893–894.

Dobson M (1994) Patterns of distribution in Japanese land mammals. *Mammal Review*, **24**, 91–111.

Ellstrand NC, Elam DR (1993) Population genetic consequences of small population size: implications for plant conservation. *Anuual Review of Ecology and Systematics*, **24**, 217–242.

Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 7043–7050.

Elton CS (1958) *The ecology of invasions by animals and plants*. Methuen, London, England.

Fisher RA (1930) *The genetical theory of natural selection*. Clarendon, Oxford, England.

Goodwin SB, Cohen BA, Fry WA (1994) Panglobal distribution of a single clonal lineage of the Irish potato famine fungus. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 11591–11595.

Gratz NG (1984) The global public health importance of rodents. Pp. 413-435 In *Proceedings of a conference on the organization and practice of vertebrate pest control* (A. C. Dubock, ed.). World Health Organization, Geneva, Switzerland.

Grzmik B (1975) *Grzmik's encyclopedia: mammals, I-IV*. Nan Nostrand Reinhld, New York.

Hingston M, Goodman SM, Ganzhorn JU, Sommer S (2005) Reconstruction of the colonization of southern Madagascar by *Rattus rattus*. *Journal of Biogeography*, **32**, 1549–1559.

Inglesby TV, Dennis DT, Henderson DA, Bartlett JG, Ascher MS, Eitzen E, Fine AD, Friedlander AM, Hauer J, Koerner JF, Layton M, McDade J, Osterholm MT, O'Toole T, Parker G, Perl TM, Russell PK, Schoch-Spana M, Tonat K (2000) Plague as a Biological Weapon. *The Journal of the American Medical Association*, **283**, 2281–2290.

Jones JK Jr., Johnson DH (1965) Synopsis of the lagomorphs and rodents of Korea. *University of Kansas Publications, Museum of Natural History*, **16**, 357–407.

Kawamura Y (1989) Quaternary rodent faunas in the Japanese Islands (Part 2). *Memories of the Faculty of Science, Kyoto University, Series of Geology and Mineralogy*, **54**: 1–235.

Kingdon J (1974) *East African mammals: an atlas of evolution in Africa. II(B). Hares and rodents*. Academic Press, London, England.

Kowalski K, Hasegawa Y (1976) Quaternary rodents from Japan. *Bulletin of the National Science Museum (Tokyo), ser. C (Geology and Paleontology)*, **2**, 31–66.

Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology and Evolution*, **17**, 386–391.

Lowery GH Jr (1974) *The mammals of Louisiana and its adjacent waters*. Louisiana State University Press, Baton Rouge, Louisiana.

Martel C, Viard F, Bourguet D, Garcia-Meunier P (2004) Invasion by the marine gastropod *Ocinebrellus orinata* in France I. Scenario for the source of introduction. *Journal of Experimental Marine Biology and Ecology*, **305**, 155–170.

Merila J, Bjorkland M, Baker A (1996) The successful founder: genetics of introduced Carduelis chloris (greenfinch) populations in New Zealand. *Heredity*, **77**, 410–422.

Musser GG, Carleton MD (1993) Family Muridae. Pp. 501–755 In *Mammal species of the world, a taxanomic and geographic reference* (D. E. Wilson and D. M. Reeder, ed.). Smithsonian Institute Press, Washington D.C.

Myers JH, Simberloff D, Kuris AM, Carey JR (2000) Eradication revisited: dealing with exotic species. *Trends in Ecology and Evolution*, **15**, 316–320.

Nei M, Maruyama T, Chakraborty R (1975) The genetic bottleneck effect and genetic variability in populations. *Evolution*, **29**, 1–10.

Newman D, Pilson D (1997) Increased probability of extinction due to decreased genetic effective population size: experimental populations of *Clarkia pulchella*. *Evolution*, **51**, 354–362.

Nieminen M, Singer MC, Fortelius W, Schops K, Hanski I (2001) Experimental confirmation that inbreeding depression increases extinction risk in butterfly populations. *American Naturalist*, **157**, 237–244.

Niethammer J (1975) Zur Taxonomie und Ausbreitungsgeschichte der Hausratte (Rattus rattus). *Zoologischer Anzeiger*, **194**, 405–415.

Novak SJ, Mack RN (1993) Genetic variation in *Bromus tectorum* (Tectaceae): comparison between native and introduced populations. *Heredity*, **71**, 167–176.

Novak SJ, Mack RN (2005) Genetic bottlenecks in alien plant species: influence of mating systems and introduction dynamics. Pp. 201–228 In *Species invasions: insights into ecology, evolution, and biogeography* (D. F. Sax, J. J. Stachowicz and S. D. Gaines, ed.). Sinauer Associates, Sunderland, Massachusetts.

Nowak R (1999) *Walker's Mammals of the World (6th Edition)*. John's Hopkins University Press, Baltimore, Maryland.

Patton JL, Yang SY, Myers P (1975) Genetic and morphologic divergence among introduced rat populations (*Rattus rattus*) of the Galapagos Archipelago, Ecuador. *Systematic Zoology*, **24**, 296–310.

Pimentel D, Lach L, Zuniga R Morrison D (2000) Environmental and economic costs of nonindigenous species in the United States. *BioScience*, **50**, 53–65.

Quinn TP, Unwin MJ, Kinnison MT (2000) Evolution of temporal isolation in the wild: genetic divergence in timing of migration and breeding by introduced Chinook salmon populations. *Evolution*, **54**, 1372–1385.

Ricciardi A, Neves RJ, Rasmussen JB (1998) Impending extinctions of North American freshwater mussels (Unionoidae) in North America. *American Zoologist*, **36**, 300–310.

Rasolomaharo M, Rasoamanana B, Adrianirina Z, Buchy P, Rakotoarimanana N, Chanteau S (1995) Plague in Majunga, Madagascar. *Lancet*, **346**, 1234.

Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With KA, Baughman S, Cabin RJ, Cohen JE, Ellstrand NC, McCauley DE, O'Neil P, Parker IM, Thompson JN, Weller SG (2001) The population biology of invasive species. *Annual Review of Ecology and Systematics*, **32**, 305–332.

Sala OE, Chapin FS III, Armesto JJ, Berlow E, Bloomfield J, *et al*. (2000) Global biodiversity scenarios for the year 2100. *Science*, **287**, 1770–1774.

Sax DF, Stachowics JJ, Gaines SD (2005) *Species invasions: insights into ecology, evolution, and biogeography*. Sinauer Associates, Inc., Sunderland, Massachusetts.

Schwartz CW, Schwartz ER (1981) *The wild mammals of Missouri*. University of Missouri Press, Columbia.

Scott S, Duncan CJ, Duncan SR (1996) The plague in Penrith, Cumbria, 1597/8: its causes, biology, and consequences. *The Annals of Human Biology*, **23**, 1–21.

Shrewsbury JFD (1970) *A History of Bubonic Plague in the British Isles*. Cambridge University Press, Cambridge, England.

Slack P (1989) The black death past and present. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **83**, 461–463.

Voigt B, Kitada K, Kloting I, Serikawa T (2000) Genetic comparisons between laboratory rats and Japanese and German wild rats. *Mammalian Genome*, **11**, 789–790.

Wares JP, Hughes AR, Grosberg RK (2005) Mechanisms that drive evolutionary change: insights from species introductions and invasions. In *Species invasions: insights into ecology, evolution, and biogeography* (D. F. Sax, J. J. Stachowicz and S. D. Gaines, ed.). Sinauer Associates, Sunderland, Massachusetts.

Wilcove DS, Rothstein D, Dubow J, Phillips A, Losos E (1998) Quantifying threats to imperiled species in the United States. *BioScience*, **48**, 607–615.

Williamson M (1996) *Biological Invasions*. Chapman and Hall, New York.

CHAPTER II

COMPARATIVE PHYLOGEOGRAPHY OF INVASIVE *RATTUS*

Abstract— Invasive *Rattus* are arguably the most costly and destructive invasive species on the planet, but little is known concerning their invasion history and population structure in the U.S. We utilized both nuclear microsatellites and mitochondrial DNA sequences (mtDNA) to compare the colonization history, patterns of gene flow, and levels of genetic diversity of *Rattus rattus* and *R. norvegicus* in the U.S. Analyses of mtDNA suggest *R. rattus* is characterized by a single rapid expansion into the U.S. from one or two very closely related mtDNA lineages or geographic sources. For *R. norvegicus*, mtDNA analyses suggest at least four invasions distinct in space and/or time have occurred to establish its distribution in the U.S. Microsatellite analyses indicate for *R. rattus* that dispersal is characterized by an isolation-by-distance pattern, suggesting a relatively low frequency of long distance dispersal, and low levels of establishment for novel propagules. In contrast, microsatellite analysis of *R. norvegicus* in the U.S. suggests high frequencies of long distance dispersal and essentially panmixia among nearly all sampled populations, as well as a high frequency of novel propagules entering at the east and west coasts and assimilating into established populations. We discuss these results in the context of invasive *Rattus* management in the U.S. and their implications for invasive species in general, as well as the implications for managing the spread of rat-borne pathogens.

Introduction

In studying biological invasions, population genetic analyses can provide valuable information concerning colonization history, population demographics, and patterns of gene flow in both the native and invaded range (Lee 2002; Le Roux & Wieczorek 2009). These parameters are useful for identifying source populations of the original colonization and for contemporary dispersal into the invaded range, and are typically used to model a species' invasion, allowing for extrapolations that can be used to predict routes of dispersal and future spread. If management plans are to be effective, it is vital that we understand from where invasive species are entering, as well as understand the ecological factors affecting population connectivity. Population genetics provides a powerful tool for improving the effectiveness and sustainability of these plans (Abdelkrim *et al*. 2005a).

Although the use of population genetic analyses to indirectly quantify critical components of invasions has become a mainstay of invasion biology, genetic diversity is itself a vital component of the colonization process. The adaptability of a population to a novel environment is largely a function of the standing genetic diversity. The invasion process typically results in a genetic bottleneck from founder effects (Prentis *et al.* 2008), and reduced genetic diversity and small population size leads to a significant increase in extinction risk (Frankham & Ralls 1998; Allendorf & Lundquist 2003). In spite of this, many introduced species overcome the initial colonization to establish and spread (termed the "genetic paradox" of invasions; Sakai *et al.* 2001; Allendorf & Lundquist 2003). Studies have shown that this is likely a result of multiple geographically distinct source populations undergoing admixture in the invaded range (Kolbe *et al*. 2004, 2008), resulting in novel genetic combinations and elevated levels of genetic diversity. The time required for admixture to occur (and genetic diversity to rise) is often invoked as an

explanation for the time lag typically observed between the initial invasion and the onset of exponential population growth and geographic spread.

Rattus rattus and *R. norvegicus* are arguably the most successful invasive species on the planet. Through their commensal relationship with humans they have spread to almost every corner of the globe, with *R. rattus* occurring on all continents and *R. norvegicus* excluded only from Antarctica (Nowak 1999; Musser & Carlton 2005). In the U.S., *R. rattus* are restricted to the south-central and coastal lower 48 states with a single known population in Alaska. Their numbers in coastal areas can grow quite large due to a close association with ships. At least in the U.S., *R. rattus* are more common in the upper stories of buildings and in trees, rarely occurring on the ground or in open fields (Schwartz & Schwartz 1981; Caire *et al.* 1989). The generalist *R. norvegicus* occurs throughout the country, including Alaska and Hawaii. Suitable habitat includes the ground floor of buildings and dwellings, around sewers and garbage dumps, and can even be found in open fields around rural areas (Schwartz & Schwartz 1981; Caire *et al.* 1989). Where *R. rattus* and *R. norvegicus* co-occur, it has been noted that the more aggressive *R. norvegicus* excludes *R. rattus* from favorable habitat (Nowak 1999). In essentially every habitat invaded, *Rattus* have had severe negative impacts on natural diversity. Of the approximately 123 island groups worldwide, about 82% have been invaded by *R. norvegicus, R. rattus*, or the Polynesian rat (*R. exulans*; Courchamp *et al*. 2003), and recent reports estimated that introduced rats have been responsible for 40–60% of all bird and reptile extinctions since 1600 (Island Conservation 2006).

In addition to damaging natural ecosystems, a more universal concern is the economic and human health impacts of invasive *Rattus*. Within the U.S., invasive rats are responsible for approximately 19 billion dollars in annual economic loss through the transmission of disease, structural damage to buildings, and contamination and destruction of food supplies (Pimentel *et al*. 2000). From an epidemiological standpoint, *Rattus* are known to spread many zoonoses

including bubonic plague, murine typhus, rat-bite fever, *Salmonella* food poisoning, leptospirosis, listeriosis, chagas, trichinosis, tularemia, hantavirus and schistosomiasis (Gratz 1984). Over the last millennium, rat-borne pathogens have been estimated to have killed more people than all wars and revolutions combined (Nowak 1999; Meerburg *et al*. 2009), and rat-borne pathogens are still a serious concern. As an example, plague (*Yersinia pestis*) is often thought of as a pathogen relevant only to the dark ages, but the recent discovery of antibiotic resistant strains and the increased incidence of infections suggests otherwise (Galiman 1997; Keeling and Gilligan 2000). Drug-resistant *Y. pestis* has even been suggested as a potential bioterrorism weapon (Inglesby et al. 2000).

Despite the substantial economic and human health impacts of invasive *Rattus*, essentially nothing is known of their colonization history and population structure in the U.S. Information concerning the colonization history, genetic diversity, and gene flow of *R. rattus* and *R. norvegicus* within the U.S. can be used to inform eradication efforts, but more importantly can be used to assess the relative risk each species poses in terms of the import and spread of infectious disease. Invasive *Rattus* are notorious for their ability to utilize human transportation vectors (i.e., ships) to disperse (Sullivan 2004). Given this ability and the high volume of international shipping entering at the U.S. coastlines (6 million shipping containers annually; Frittelli *et al*. 2005), it is possible that propagules of both species are entering the U.S. in high numbers and assimilating with already established populations.

We utilized both mitochondrial and nuclear DNA to assess colonization history, partitioning of genetic diversity, and patterns of gene flow in the U.S. for *R. rattus* and *R. norvegicus*. Prior genetic analyses outside the U.S. revealed several divergent mitochondrial DNA lineages with geographic structure for *R. rattus* (Hingston *et al*. 2005; Tollenaere *et al*. 2010; Aplin *et al.* 2011) and *R. norvegicus* (Bastos *et al*. 2011), although there is less data available for *R. norvegicus*. Therefore, we can predict that mitochondrial diversity will reflect the

diversity in source populations for colonizing individuals. In addition, if individuals are continuing to enter the U.S. from international localities, we predict genetic diversity will be highest in coastal localities. Finally, if individuals of either species are dispersing throughout the U.S. using human transportation vectors (i.e., trucks and trains), we predict patterns of genetic differentiation will reflect frequent long-distance dispersal and will therefore not fit a model of isolation-by-distance (IBD) across a relatively continuous landscape.

Materials and methods

Population sampling

We obtained tissues from museum collections in addition to individuals collected through trapping efforts for localities across the U.S. and from international localities. For *R. rattus* we were able to obtain tissues from 231 individuals from 24 localities (18 localities in the U.S.). For *R. norvegicus* we obtained 212 tissues from 26 localities (23 localities in the U.S.). Sample sizes for populations varied significantly for both species, ranging from one to 51 collected individuals. A complete list of all collected individuals, source localities, population sample sizes, and loaning institutions are given in Table S2.1.

MtDNA sequence generation and analyses

The complete mitochondrial cytochrome *b* gene (1140 bp) was amplified using primers (*Rattus*CytbF 5'-TGACATGAAAAATCATCGTTGTAAT-3'; *Rattus*CytbR 5'-GGTTTACAAGACCAGAGTAAT-3') designed from an alignment of the complete mtDNA genome sequences of *R. rattus* (NC012374), *R. norvegicus* (AY769440), *R. tanezumi* (NC011638), and *Mus musculus* (NC006915) obtained from GenBank. PCR amplifications were

carried out in 30µl reactions containing 200–500 ng of DNA, 0.12 µl of 5U/µl GoTaq Flexi DNA polymerase, 0.50 µl of each 10 µM primer, 2.4 µl Bovine Serum Albumin (0.01 g/ml), 2.4 µl of 25µM $MgCl_2$, 6.0 µl 5X Green GoTaq Flexi Buffer, 4.2 µl of a 10µM nucleotide mixture, and 8.88 µl of double distilled water (ddH_2O). The thermal profile consisted of an initial denaturation of 94°C for 4 minutes, followed by 35 cycles of 94°C, 42°C, and 72°C for 1 minute each. A final elongation at 72°C for 7 minutes ensured all reactions went to completion. Double-stranded products were purified using the Wizard SV Gel and PCR Clean-Up System (Promega Corporation, Madison, WI).

Both strands of the purified PCR products were sequenced using BigDye chain terminators following the manufacturer's protocol (Applied Biosystems) using the PCR amplification primers as well as two internal primers (*Rattus*CytbIntF 5'-GGCTTCTCAGTAGACAAAGC-3'; *Rattus*CytbIntR 5'- TTTGATCCTGTTTCGTGGAGGAA-3') designed using the mtDNA genome alignment generated above. DNA sequencing reactions were electrophoresed on a 3130 Genetic Analyzer (Applied Biosystems). Contigs were assembled and edited using Geneious v 5.5 (Drummond *et al.* 2010).

In addition to the complete cytochrome *b* sequences generated above, we obtained all available *R. rattus* and *R. norvegicus* cytochrome *b* sequences from GenBank (a complete list of sequences obtained from GenBank with sampling localities is available in Table S2.2), resulting in 275 and 229 cytochrome *b* sequences for *R. rattus* and *R. norvegicus*, respectively. The sequences obtained from GenBank ranged from 713 to 1140 basepairs. Initial phylogenetic analysis (conducted in Lack *et al*. In Review) and haplotype network analyses (conducted herein) were performed on both the full-length sequences only and the total dataset, and results were identical concerning clade assignment and haplotype network structure, and were also consistent with previous mtDNA analyses (Robins *et al.* 2007; Aplin *et al*. 2011). Therefore, subsequent mtDNA analyses included all sequences and only these results are presented. Sequences were

aligned using the Geneious v5.5 aligner (Drummond *et al*. 2010) and edited using MacClade v4.08 (Madison and Madison 2000).

To verify field identifications and GenBank records, a Bayesian phylogenetic analysis of all cytochrome *b* sequences was previously conducted by Lack *et al*. (In Review). This analysis revealed all individuals previously identified as *R. norvegicus* to be correctly identified. However, several individuals obtained from the San Francisco Bay Area, CA and Panama City, FL were members of the cryptic species *R. tanezumi*, and therefore excluded from further mtDNA analysis. After confirming species identifications, we generated unrooted haplotype networks for each species (*R. rattus* and *R. tanezumi*) using TCS (Clement *et al*. 2000).

We also conducted several population level analyses for samples from within the U.S. We estimated haplotype diversity (h) and nucleotide diversity (π) with the software package DnaSP (Rozas *et al*. 2003). To investigate the demographic history of each species in the U.S. we generated mismatch distributions of pairwise differences among all individuals using DnaSP and assessed the fit of the observed distribution to a model of sudden population expansion using 1,000 bootstrap replicates (Rogers & Harpending 1992). In addition, we estimated Fu's F_S, D^*, and F^* (Fu 1997) as a measure of selective neutrality and population expansion, and assessed significance using coalescent simulations in DnaSP. A significant F_S in the absence of significance for D^* and F^* suggests recent population expansion, while the opposite scenario suggests background selection is responsible for the observed pattern of genetic variation (Fu 1997).

Microsatellite data generation and analyses

We genotyped all individuals collected from localities in the U.S. at nine microsatellite loci (loci names and primers given in Table S2.3). Microsatellites were amplified by PCR in 15 μl reactions containing 9μl of True Allele PCR Premix (Applied Biosystems, Inc., Foster City, CA),

4 μl ddH_2O, 0.5 μl of each primer (10 μM), and 1 μl template DNA with the following conditions: an initial denaturation of 95°C for 12 minutes, 35 cycles of 94°C for 40 seconds, 57°C for 40 seconds, and 72°C for 30 seconds; and a final elongation of 72°C for 4 minutes. Then 0.5 μl of product was added to 9.5 μl of loading mix containing a size standard (ROX 400HD; Applied Biosystems, Inc., Foster City, CA) and this mixture was analyzed using an ABI 3130 Genetic Analyzer and GeneMapper 3.7 to visualize microsatellite alleles and determine genotypes. All genotypes were scored twice, and anonymously randomized for the second scoring to ensure no bias was present in the final dataset.

As mentioned above, several individuals from the San Francisco Bay and Panama City populations possessed the mtDNA of the cryptic species *R. tanezumi* (Lack *et al*. In Review). While these individuals were excluded from mtDNA analyses, previous analysis of nuclear data for these populations indicated that *R. tanezumi* and *R. rattus* have undergone extensive hybridization with introgression, and measures of population structure indicated nuclear genome panmixia among these mtDNA lineages at each population (Lack *et al*. In Review; Conroy *et al.* In Review). Therefore, we conducted initial microsatellite analyses (measures of diversity and clustering analysis in Structure) excluding individuals with *R. tanezumi* mtDNA as well as using the entire dataset. Consistent with previous analyses, measures of diversity (e.g., expected and observed heterozygosity, gene diversity) and clustering analysis gave essentially the same results for both datasets (an identical optimal *K*, K=4, was selected for both datasets). Therefore, all analyses and results presented here were conducted on the total dataset.

Genepop v4.0 (Raymond and Rousset 1995) was used to test for deviations from Hardy-Weinberg equilibrium (HWE) by conducting global heterozygosity excess and deficit tests for each locality and species. Each test was run for 10,000 dememorization steps followed by 100 batches of 5,000 steps each. Additionally, we used Genepop to conduct a composite linkage disequilibrium (LD) test (Weir 1996). Significance was assessed with the same MCMC settings

as was used for the heterozygosity tests, and a Bonferroni correction was used to correct for multiple comparisons. To assess genetic diversity, we calculated the corrected gene diversity with rarefaction correction for unequal sample sizes in FSTAT v2.9.3 (Goudet 2001), excluding populations with less than 5 genotyped individuals. We estimated observed and expected heterozygosity (H_O and H_E, respectively) in Genalex v6.41 (Peakall & Smouse 2006). Mantel tests (Mantel 1967) of significance of regression between pairwise genetic distance (D_{EST}, calculated using SMOGD; Crawford 2010) and straight-line geographic distance were conducted using Genalex v6.41, and this analysis excluded populations with fewer than 5 genotyped individuals. A significant positive correlation is taken to indicate an IBD pattern of divergence.

To examine fine-scale structure we used two approaches, both of which included all sampled individuals. We conducted a principle coordinate analysis (PCoA) on a pairwise genetic distance matrix for individuals and populations in Genalex v6.41. In addition, we used the Bayesian model-based clustering approach implemented in Structure v2.3 (Pritchard et al. 2000) to infer population structure. This approach gives the probability of assignment of each individual to postulated clusters independent of the location they were sampled, allowing for the identification of individuals with ancestry attributable to multiple populations. For each value of K (we ran independent analyses from K=1 to K=N, where N equals the number of sampled populations in the microsatellite dataset for each species), we ran five independent analyses of 500,000 generations following 250,000 generations of burn-in under the admixture model and with the assumption that allele frequencies among populations are correlated. Convergence was checked by plotting likelihoods throughout the run and comparing likelihood values and population assignments between duplicate runs. We calculated the optimal number of clusters for our data using the ΔK statistic (Evanno *et al*. 2005). The five replicates for each K were then combined using Clumpp (Jakobsson & Rosenberg 2007), which identifies common modes among replicates runs, and resulting outputs were then used to construct graphs of Structure results in Distruct (Rosenberg 2004).

To estimate recent migration rates among sampled populations of each species, we used BayesAss v3.0.1 (Wilson & Rannala 2003). This analysis estimates proportions of non-migrants and the source of migrants for each sampled population over the last several generations (Wilson & Rannala 2003), which is an ideal approach for invasive species because invasions are such recent events (hundreds of years or less). For this analysis, only populations with at least five genotyped individuals were included for each species. Initial runs consisted of 10 million generations sampled every thousand generations and a burnin of 1 million generations. To ensure parameter space is adequately explored, Wilson and Rannala (2003) suggest acceptance rates between 20% and 40%. Following the preliminary analyses, proposal step length was increased to 0.3 for the migration parameter (*m*), the allele frequencies (*a*), and inbreeding coefficients (*f*) in order to reduce acceptance rates into the recommended range. Using the updated search settings, we ran 5 replicate runs of 100 million generations sampling every 1,000 generations and 10 million burnin generations. Convergence was assessed by comparing migration estimates across replicate runs and by examining the log probability of each analysis in Tracer v1.5.

Results

Due to differences in the quantity and quality of tissue and DNA extractions, we were unable to generate both cytochrome *b* sequence and microsatellite genotypes for several individuals. Therefore, sample sizes differed slightly for several localities between the mtDNA and microsatellite datasets for each species. Sample sizes for each species, population, and dataset are given in Tables 2.1–2.4.

MtDNA analyses

Despite collection localities spanning essentially the entire U.S., mtDNA diversity for individual populations was unexpectedly low for both species (Tables 2.1 and 2.2). For *R. rattus*, we detected 21 haplotypes from 163 individuals collected from 18 U.S. localities (Table 2.1), but

nearly all haplotypes were very closely related (π =0.00180). Within populations of *R. rattus*, we recovered either 1 or 2 haplotypes for the majority of localities, and this appeared to be independent of sample size. In addition, mtDNA diversity for *R. rattus* did not appear to exhibit any geographic pattern, with relatively high and comparable haplotype diversity values detected in both coastal (e.g., San Francisco Bay, *h*=0.791) and central U.S. localities (e.g., Austin, TX, *h*=0.769 and San Angelo, TX, *h*=0.681). For *R. norvegicus* the pattern was very similar in terms of haplotype diversity, with the vast majority of localities consisting of 1 or 2 haplotypes (Table 2.2). In contrast, nucleotide diversity was higher for *R. norvegicus* (π range was 0.00018–0.00560 for populations, total π=0.00294) than *R. rattus* (π range was 0.00016–0.00327 for populations, total π=0.00180).

Haplotype networks (Figs. 2.1 and 2.2) revealed distinct colonization histories for *R. rattus* and *R. norvegicus*. For *R. rattus*, nearly all haplotypes recovered in the U.S. (and Mexico, Central America, and Argentina as well) were no more than 2 mutations removed from the 2 most common and geographically widespread haplotypes (labeled Rr1 and Rr2 in Fig. 2.1). The only exceptions to this were two divergent haplotypes recovered in south Florida. One of these haplotypes was shared with multiple localities in South Africa and was closely related to several other haplotypes recovered in southeast Africa and the Middle East. The other haplotype was recovered from no other locality, but was closely related to several haplotypes recovered from South Africa, southeast Africa, the Middle East, India, and Indonesia. Haplotype Rr1 was recovered from the Aleutian Islands, San Francisco Bay, CA, central Texas, and multiple localities in Florida within the U.S., and outside the U.S. was recovered from Mexico, Central America, Egypt, the Lesser Antilles, East Asia, and several localities in Africa. Haplotype Rr2 was recovered in Washington, two California localities (San Francisco Bay and San Diego), central Texas, Louisiana, Arkansas, and all four Florida localities within the U.S., and outside the U.S. was recovered from East Asia and South Africa.

For *R. norvegicus* we recovered 4 divergent but relatively common and widespread haplotypes in the U.S. (labeled Rn1–Rn4 in Fig. 2.2), with the remaining haplotypes consisting primarily of singletons clustered around one of the widespread haplotypes (Fig. 2.2). Haplotype Rn1 was recovered within the U.S. from the Aleutian Islands, Washington, southern California, Arkansas, northern Illinois, and Pennsylvania, and was recovered outside the U.S. from East Asia and Central America. Haplotype Rn1 gave rise to haplotypes found in Arkansas, southern California, Mexico, and the Lesser Antilles. Haplotype Rn2 was recovered within the U.S. from the Alexander Archipelago (southeast Alaska) and mainland Alaska, southern California, New Mexico, Memphis TN, Indiana, Chicago, IL, Baltimore, MD, and West Virginia, and was recovered outside the U.S. from the Lesser Antilles, East Asia, and Argentina. Haplotype Rn2 gave rise to haplotypes found in West Virginia, Memphis, TN, Baltimore, MD, South Africa, Argentina, and the Lesser Antilles. Haplotype Rn3 was recovered within the U.S. from the Aleutian Islands, Oregon, Oklahoma, central Texas, and Indiana, and was recovered outside the U.S. from East Asia. Haplotype Rn3 gave rise to haplotypes detected only in East Asia. Haplotype Rn4 was recovered within the U.S. from Chicago, IL, Oregon, and New York, and was not recovered outside the U.S. Haplotype Rn4 gave rise to haplotypes from Chicago, IL and Denmark. The remaining haplotypes were recovered from East Asia.

Tests for historical population expansion indicate the mtDNA diversity of *R. rattus* in the U.S. is the result of a single rapid expansion. Fu's F_S was significantly negative (F_S = -8.997, P<0.001) while D^* (D^* = -0.8296, P>0.1) and F^* (F^* = -1.5018, P>0.1) were both non-significant, indicating the deviation from selective neutrality observed for this species is the result of expansion and not background selection. The mismatch distribution also supported expansion for *R. rattus* (Fig. 2.3A), with an essentially unimodal distribution of pairwise differences and low raggedness index (r = 0.0757) and *R2* (*R2* = 0.0317) values, which indicate the observed data fit the sudden expansion model. In contrast, demographic analyses for *R. norvegicus* reject the expansion model. All measures of selective neutrality were non-significant (F_S = 1.845, P =

0.088; D^* = -1.147, >0.10; F^* = -0.937, P >0.10), and the mismatch distribution (Fig. 2.3B) exhibited a multimodal distribution of pairwise differences (r = 0.288; $R2$ = 0.0893).

Microsatellite analyses

For the microsatellite data, no locality exhibited significant deviation from HWE or LD for either species and all loci were polymorphic for both species. For *R. rattus*, measures of diversity exhibited no obvious patterns among populations (Table 2.3). For populations with at least 5 sampled individuals, gene diversity values ranged from 0.488 to 0.781. For *R. norvegicus*, measures of diversity appeared to be highest in coastal populations and lowest in populations in the center of the country (Table 2.4). For populations with at least 5 samples, gene diversity values ranged from 0.362 to 0.746, with the highest values in populations in close proximity to either coast; the only exception to this was the rural Pennsylvania population.

The Mantel test of IBD revealed significantly different patterns for each species in the U.S. (Fig. 2.4). The Mantel test for *R. rattus* (Fig. 2.4A) indicated a significant positive relationship between geographic and genetic distance (P = 0.049). This suggests a pattern of gene flow conforming to the IBD model. In contrast, we detected no significant relationship between geographic and genetic distance (P = 0.353) for *R. norvegicus* (Fig. 2.4B), rejecting a model of IBD and indicating frequent long-distance dispersal events.

For the individual-based analyses of genetic structure, both species exhibited relatively little population structure. The PCoA of *R. rattus* (Fig. 2.5) indicated the Shemya Is., AK and Gainesville, FL populations were relatively distinct from all other sampled populations. Interestingly, the single individual obtained from Great Sitkin Is., AK grouped with the main cluster of samples, and not with the more distinct Shemya Is., AK population. A second PCoA analysis of *R. rattus* with the Gainesville, FL and Shemya Is., AK populations removed revealed no additional samples or populations were distinct (results not shown), and all samples formed essentially a single cluster. For the *R. norvegicus* PCoA (Fig. 2.5), the Aleutian Is., AK

populations were distinct from all other sampled populations, as in the *R. rattus* analysis. Also, the Alexander Archipelago (southeast Alaska) samples and single individual from mainland Alaska (Fairbanks, AK) grouped relatively close to the main cluster of samples. The rural Pennsylvania population also was distinct from all other sampled individuals. In a subsequent analysis with the Aleutian Is. and Pennsylvania populations removed, no additional populations became distinct (results not shown).

For the Bayesian assignment approach implemented in Structure, the results were largely congruent with the PCoA, but with increased resolution. For *R. rattus*, the ΔK statistic (Fig. S2.1) suggested the number of clusters present in the data to be four (Fig. 2.6). As in the PCoA, the Shemya Is., AK and Gainesville, FL populations each formed distinct groups. A third cluster of individuals (blue in Fig. 2.6) consisted primarily of the Gulf Coast populations (Louisiana and Florida), but with intergradation into central Texas and the San Diego, CA and San Francisco Bay, CA populations. The fourth cluster (green in Fig. 2.6) consisted primarily of western U.S. populations (with the exception of Shemya IS., AK), but with intergradation into central Texas, gulf coast Texas, and Arkansas. The Greater Sitkin Is., AK individual grouped with this western cluster and not with the Shemya Is., AK population. For *R. norvegicus*, the ΔK statistic (Fig. S2.2) suggested the number of clusters present in the data to be three (Fig. 2.6). The first cluster (green in Fig. 2.6) consisted almost entirely of Aleutian Islands populations, as well as some mixed assignments in the Alexander Archipelago, AK. The second distinct cluster (green in Fig. 2.6) consisted entirely of the rural Pennsylvania population. The third cluster (red in Fig. 2.6) consisted of all remaining populations with little intergradation from the other two clusters. The Alexander Archipelago, AK and the single Fairbanks, AK (mainland Alaska) individual also grouped more closely with the large cluster (red in Fig. 2.6) than with the Aleutian Islands populations.

Assessment of pairwise population migration rates revealed several differences between *R. rattus* and *R. norvegicus*. Duplicate runs for each species produced essentially identical results, and Tables 2.5 and 2.6 present the mean across the 5 replicates for *R. rattus* and *R. norvegicus*, respectively. The overall mean pairwise rate between sampled populations was higher for *R. norvegicus* (0.0327) than for *R. rattus* (0.0195). For *R. rattus*, pairwise migration rates (Table 2.5) appeared to support the results of the Mantel test (Fig. 2.4), with the majority of populations exhibiting higher rates of gene flow between geographically more proximate populations. As an example, the Austin, TX and Houston, TX populations exhibited an approximately 10-fold higher migration rate with San Angelo, TX than all other sampled populations. Moreover, the majority of gene flow among these Texas populations was occurring into the more rural San Angelo, TX population and not in the other direction. For Seattle, WA migration again was highest for the relatively close San Francisco Bay, CA and Tahoma, CA populations. There were two exceptions to this pattern, where the highest migration rates into San Francisco Bay, CA were from Key Largo, FL and Baton Rouge, LA, illustrating the dispersal potential of this species. For *R. norvegicus*, we did not observe a pattern of migration rates correlated with geographic proximity (Table 2.6). For example, the highest rate of gene flow into the Chicago, IL population originated from the San Diego, CA population, while the highest rate of gene flow into the Baltimore, MD population originated from the more rural Spencer, IN population. The only clear exception of this was the elevated migration rate from the Monroe, WV population into the nearby Baltimore, MD population.

Discussion

Rattus colonization history in the U.S.

Through archeological evidence and historical record, we have some understanding of the early colonization of the U.S. by *R. rattus* and *R. norvegicus*. *R. rattus* appears to have been the first to arrive in North America, with archeological evidence placing them on the island of Hispaniola

with Columbus in 1492, and with established populations on the east coast of the continental U.S. by the mid 1500s (Armitage 1993). In contrast, *R. norvegicus* arrived in the U.S. in the mid 1700s with the massive wave of British immigrants that continued into the late 1700s (Armitage 1993). By the early 1800s the larger and more aggressive *R. norvegicus* had caused a drastic reduction in *R. rattus* numbers (MacGillivray 1838). This scenario is strikingly similar to the spread of *R. rattus* from its native range on the Indian subcontinent beginning in the first millennium BC, and reaching essentially every corner of the Old World inhabited by humans by the second century AD. In the 18th century, *R. norvegicus* rapidly expanded out of central Asia into Europe, displacing *R. rattus* in much of the newly invaded range (Twigg 1975).

Although information gleaned from the historical record is important in understanding the early introduction of *Rattus* in the U.S. and elsewhere, it typically cannot deliver insight at the resolution required to understand the biological properties of the invasion, providing little concerning the source and diversity of propagules that went on to establish and produce the massive invasive rat populations present today. However, the combination of these data sources can be used to draw more powerful inferences, as has been shown in the study of *R. rattus* in the Mediterranean (Ruffina & Vidal 2010) and the commensal relationship between humans and the Polynesian rat (*R. exulans*) in the South Pacific (Matisoo-Smith *et al.* 1998; Matisoo-Smith & Robins 2004). Our mtDNA analysis provides insight into the distinct differences between *R. rattus* and *R. norvegicus* in their colonization of the U.S., even in the absence of a comprehensive global sampling.

It is clear from our haplotype network (Fig. 2.1) as well as previous studies (Hingston *et al*. 2005; Tollenaere *et al*. 2010; Aplin *et al.* 2011) that considerable mtDNA diversity exists for *R. rattus* at the global scale. In spite of this, we detected only a subsampling of this genetic diversity within the U.S., with an overall nucleotide diversity for *R. rattus* that was nearly half that of *R. norvegicus* and essentially all U.S. haplotypes forming a single star-shaped cluster (Fig. 2.1); the only exceptions were two divergent haplotypes that were recovered in coastal south

Florida. In addition, the mismatch distribution (Fig. 2.3A) and neutrality statistics suggest a single rapid expansion best explain these data. This pattern suggests several possible scenarios for the colonization history of the U.S. by *R. rattus*.

First, it is possible our geographic sampling was too sparse to be representative of the overall mtDNA diversity of *R. rattus* in the U.S., but this is unlikely. However, what we know concerning the current distribution of *R. rattus* in the U.S. suggests it has been reduced to populations only in the southeast, the Gulf coast, and the Pacific coast (Jackson 1982), and we have sampled multiple localities in each of these regions. Alternatively, it is possible *R. rattus* mtDNA diversity in the U.S. was historically much higher, and representative of the mtDNA diversity of the native range. Subsequent invasion and competitive exclusion by *R. norvegicus* did lead to a widespread and documented bottleneck for *R. rattus* in the U.S. (MacGillivray 1838; Armitage 1993), eliminating much of the mtDNA diversity of the original *R. rattus* populations. However, the bottleneck scenario is an unlikely explanation for the observed pattern because, if divergent mtDNA haplotypes had been present at a significant frequency, we would not expect the same haplogroup to persist in *R. rattus* populations from Alaska to Florida. In addition, evidence suggests many coastal populations of *R. rattus* were never adversely affected by the later invasion of *R. norvegicus* (Silver 1927; Ecke 1954; Landon 1989). Also, *R. rattus* and *R. norvegicus* are sympatric over most of their global range, including South Africa, where Bastos *et al*. (2011) conducted comparative mtDNA analyses. Bastos *et al.* (2011) found considerable mtDNA diversity for *R. rattus*, suggesting a complex colonization history with multiple global sources of divergent mtDNA lineages and indicating interspecific interactions were most likely not obscuring colonization history.

As the most likely explanation for the distribution of *R. rattus* mtDNA diversity in the U.S., we suggest colonization occurred by two closely related mtDNA lineages (Rr01 and Rr02), and subsequent colonizers from divergent mtDNA lineages have not invaded and spread. In an examination of global *R. rattus* mtDNA diversity, Aplin *et al.* (2011) noted that the initial

expansion out of the Indian subcontinent appeared to originate from a single mtDNA lineage, which they termed the "ship rat" lineage, that then spread across the planet. The main cluster of U.S. *R. rattus* haplotypes that we recovered are members of this "ship rat" lineage (Aplin *et al*. 2011), indicating the initial colonization of the U.S. and any subsequent invasions were essentially all members of this group. The presence of two common and widespread haplotypes in the U.S. (Rr1 and Rr2) suggests either two distinct waves of invasion or a single invasion by both mtDNA lineages spread across the country, giving rise to locally common haplotypes. Initial introductions of *R. rattus* into the U.S. undoubtedly came from Europe onto the east coast and spread west following human colonization (Armitage 1993). In contemporary times, as international trade and shipping have rapidly expanded, the vast majority of incoming shipping on the west coast originates in Asian ports, while the majority of cargo entering the U.S. on the east coast originates in Europe, Africa, and the Middle East (Kaluza *et al*. 2010). It is evident in this and previous analyses of black rat mtDNA (Aplin *et al*. 2011; Bastos *et al*. 2011), that the source localities for much of the U.S. cargo possess *R. rattus* mtDNA haplotypes divergent from the U.S. haplogroup, but these haplotypes are simply not entering the U.S. While Rr1 and Rr2 are each found outside of the Americas at a high frequency, the only haplotypes derived from Rr1 and Rr2 that were detected outside the Americas were three haplotypes detected in South Africa and a single haplotype from Reunion, an island near Madagascar. The two divergent haplotypes recovered in coastal Florida are members of an mtDNA haplogroup recovered in East Asia, Africa, the Middle East, and India. The presence of these haplotypes in coastal Florida suggests mtDNA lineages distinct from the "ship rat" lineage are being readily spread, but are not being incorporated into the majority of U.S. populations and do not appear to be spreading from the coastal localities where they are initially introduced.

The distribution of mtDNA diversity for *R. norvegicus* suggests a different colonization history relative to that of *R. rattus*. The mismatch distribution (Fig. 2.3B) and measures of selective neutrality all reject a single rapid expansion for this species and instead support a more

complex scenario of multiple introductions from multiple mtDNA stocks. We detected four relatively widely distributed and frequent haplotypes in the U.S., suggesting at least four distinct invasions. Moreover, the geographic distribution of these four haplotypes in the U.S. give some indication of where these invasion originated, although without a more extensive global sampling any conclusions are speculative. For haplotype Rn1, the distribution included both western and eastern U.S. populations, but it and the haplotypes it produced were more frequently detected in the west, and it was also detected in East Asia. This suggests an introduction from Asia that then spread east. Haplotype Rn2 was most frequently detected on the east coast and appeared to decrease in frequency moving west across the U.S. In addition, the more rare haplotypes closely related to Rn2 were all located in eastern U.S. localities, in addition to being recovered in the Lesser Antilles, Argentina, and South Africa. This suggests Rn2 may represent an east coast invasion. Haplotype Rn3 was recovered in the Aleutian Islands, AK, Oregon, central Texas, Oklahoma, and West Virginia in the U.S., as well as from East Asia. This haplotype distribution in the U.S. appears equivocal in supporting an east or west coast invasion, and therefore requires further sampling, although its high frequency in the relatively isolated Aleutian Islands populations suggests invasion from the west. Finally, haplotype Rn4 was recovered almost exclusively on the east coast, with one individual from Oregon, and the remaining nine individuals from Chicago (8 individuals) and New York (1 individual), suggesting an origin on the east coast of the U.S. The presence of a closely related haplotype recovered from Denmark and another closely related singleton from Chicago, IL also supports an east coast colonization for this lineage.

While the lack of a thorough global sampling for either *R. rattus* and *R. norvegicus* hinders our ability to make strong inferences concerning some aspects of their invasion history, we can make relative comparisons between the two species. It is clear that *R. norvegicus* in the U.S. have originated from a diversity of source populations, with at least four divergent mtDNA lineages detected at high frequency (Fig. 2.2). Furthermore, the significantly higher gene diversity

detected for *R. norvegicus* nuclear microsatellites in the coastal localities relative to the localities in the center of the U.S. suggests there is still a high influx of individuals that are integrating into the established coastal populations, a pattern also observed for invading cheat grass and Japanese oyster drill (Martel *et al.* 2004; Novak & Mack 2001). In contrast, *R. rattus* in the U.S. appear to be almost entirely derived from the same mtDNA lineage (the "ship rat" lineage; Aplin *et al*. 2011) that initially expanded out of India to colonize most of the planet. Furthermore, the lack of a clear difference in nuclear gene diversity between coastal populations and more inland populations, as well as the presence of divergent mtDNA haplotypes in coastal Florida but nowhere further inland (including the other two Florida localities), suggests that individuals that arrive at coastal populations from localities outside North America are not likely to integrate into already established *R. rattus* populations.

Genetic structure and dispersal in the U.S.

For species dispersing without the aid (or inhibition) of humans, it is expected that geography will be the overriding factor and that gene flow across a continuous landscape will roughly fit a pattern of isolation by distance (Skellman 1951; Kimura 1953). For organisms that are commensal with humans, the frequency of long-distance dispersal events increases dramatically (Suarez *et al.* 2001). The patterns detected for dispersal and population structure for *R. rattus* and *R. norvegicus* have implications in both rat management and zoonotic disease epidemiology, and suggest distinct patterns for each species.

For *R. rattus*, the Mantel test supported an overall pattern of IBD (Fig. 2.4A), suggesting long-distance dispersal events are rare and gene flow is likely occurring through more natural mechanisms of dispersal as opposed to the utilization of human transportation vectors. In contrast, we detected no IBD for *R. norvegicus*, with the relationship between genetic and geographic distance essentially flat (Fig. 2.4B). A similar lack of IBD has been detected for the Quagga mussel (*Dreissena bugensis*) in the Great Lakes area of the U.S. and this was similarly

attributed to a high frequency of long-distance dispersal due to human-aided dispersal (boat transportation; Wilson *et al*. 1999). This suggests that *R. norvegicus* may be utilizing human transportation vectors for frequent long-distance dispersal events much more frequently than *R. rattus*. Although life-history similarities between species can result in similarities in their population biology (e.g., population size, dispersal rate, etc.), the striking similarities in life-history strategies for *R. rattus* and *R. norvegicus* may be driving competitive interactions and therefore resulting in distinct patterns in their invaded range. The Norway rat has long been regarded as the much more aggressive of the pair, displacing *R. rattus* in most of the localities they co-inhabit (although for exceptions see Silver 1927; Ecke 1954), and typically forcing *R. rattus* into less desirable habitat (Nowak 1999). With the possible exception of southern Florida (Armitage 1993) and a few islands, *R. norvegicus* has invaded all U.S. localities where *R. rattus* occurs or previously occurred. The absence of *R. norvegicus* in coastal Florida and the fact that the Key Largo, FL and Miami, FL populations were the only U.S. populations where *R. rattus* mtDNA haplotypes divergent from the main cluster were detected lends further support to the proposed role of *R. norvegicus* in limiting establishment and dispersal of *R. rattus* in the U.S.

The overall degree of genetic differentiation between populations appears to be very low. The PCoA (Fig. 2.5) indicated the Aleutian Islands populations for both species are genetically distinct from all other poopulations, which is not surprising given the extreme isolation of these islands. However, for *R. rattus* in Alaska, all but one sample was obtained from a single island, Shemya Is., that is uninhabited by humans, and these samples represented the distinct cluster of Alaskan samples in both the PCoA (Fig. 2.5) and the Structure analysis (Fig. 2.6). The single individual from Great Sitkin Is., an island located closer to the Alaskan mainland and occurring on the Great Circle trading route that has high ship activity (approximately 3,100 ships annually) between East Asia, Alaska, and the west coast of the contiguous 48 states, actually grouped within the main cluster of individuals. Similarly, *R. norvegicus* collected from the Alexander Archipelago in southeastern Alaska and the single individual from Fairbanks, AK on the Alaskan

mainland grouped more closely with the main *R. norvegicus* cluster than the vast majority of the Aleutian Is. *R. norvegicus* in the PCoA and Structure analyses. This suggests that, while the isolation of Alaska has led to some distinction between both species in the contiguous 48 states and their conspecifics in Alaska, international trade and travel is still transporting these species and allowing for a detectable level of gene flow.

In addition to the divergent Alaskan populations, both species exhibited additional population structure, but diversity estimates suggest these may be the result of sampling artifact. For *R. rattus*, the Gainseville, FL population was distinct in both the PCoA (Fig. 2.5) and the Structure analysis (Fig. 2.6), which was unexpected given its location in close proximity to the other three Florida populations we sampled, its central location in the state with the most significant north/south interstate in Florida passing through it, and the population size (>250,000) of the Gainesville metropolitan area. The *R. rattus* collected from Gainesville were all taken from a single locality and may represent a unique population within the city, with an origin distinct from other Florida *R. rattus*. It is interesting to note that this population had low mtDNA haplotype diversity (Table 2.1) and the lowest microsatellite gene diversity (Table 2.3) of all sampled *R. rattus* populations, further attesting to its uniqueness and justifying further study. For *R. norvegicus*, the only population found to be distinct from the main U.S. cluster (aside from the Alaskan individuals) was collected from a single farm in rural Pennsylvania. This population consisted of a single haplotype for 42 individuals (Table 2.2), and the lowest microsatellite gene diversity of any sampled *R. norvegicus* population. The presence of a single mtDNA haplotype among such a large sample size and the very low nuclear genetic diversity suggests this population may represent a single family group, and its distinction from the remainder of the U.S. *R. norvegicus* is likely a sampling artifact. However, it is important to note that neither the Gainesville, FL *R. rattus* population nor the Union, PA *R. norvegicus* population exhibited significant deviations from HWE.

With the exception of the populations mentioned above, our analyses suggested very little genetic distinction among all other sampled populations. The PCoA of either species with the aforementioned populations removed resulted in a single clump of individuals with no discernible structure (Not Shown). In the model-based approach of Structure, the optimal *K* for *R. norvegicus* was three (Fig. 2.6), with near complete uniformity of assignment for all individuals to a single cluster (red in Fig. 2.6) to the exclusion of the Aleutian Islands populations and the rural Pennsylvania population. The Structure analysis of *R. rattus* microsatellites did detect some genetic structure not obvious in the PCoA. The optimal *K* for *R. rattus* was four with the Gainesville, FL and Aleutian Islands populations distinct from all others, and with the remaining two clusters (green and blue in Fig. 2.6) corresponding roughly to an east/west gradient, but with significant intergradation in the Texas and Arkansas populations. This is again consistent with the IBD pattern suggested by the Mantel analysis, with relatively natural, stepwise dispersal occurring from coastal populations into the interior of the U.S.

Rattus management implications

Given the tremendous ecological, economical, and human health impacts of invasive *Rattus*, management of these costly invasive species is extremely important. Successful management and eventual eradication has been achieved on many islands, and these techniques are increasing in efficacy allowing for success at expanding geographic scales for rats as well as other organisms (Howald *et al.* 2007; Simberloff 2009). Our sampling regime did not permit a high-resolution examination of complex gene flow scenarios (i.e., Guillemaud *et al.* 2010), which is important information for understanding the role of various dispersal pathways in maintaining population connectivity and identifying eradication units (Robertson & Gemmell 2004). Nonetheless, our analysis does provide insight into the overall extent of gene flow among the sampled populations, as well as relative differences in the extent and diversity of incoming propagules. In terms of propagule pressure from international localities, our data suggests significantly higher rates of

establishment and genetic diversity of new individuals for *R. norvegicus* than for *R. rattus*. Therefore, a management plan for *R. rattus* would be less likely to require a strong focus on monitoring incoming ships and freight, and focus the majority of resources on simply eliminating the already established populations, while an *R. norvegicus* plan would likely need to focus on eliminating incoming propagule pressure prior to any attempt to eliminate the already established populations. Recolonization has been a common issue in past rat eradication programs for the primary reason that adequate attention is not always devoted to eliminating the source of colonizing individuals (Abdelkrim *et al.* 2005a). In terms of dispersal within the U.S., *R. norvegicus* appears to be exhibiting a relatively high frequency of long-distance movements and the lower 48 states essentially represents a single panmictic unit. For *R. rattus*, patterns of gene flow fit an IBD model and clustering analysis detected slight differentiation between the west and the east, but gene flow again appears to be high among populations in the lower 48 states. With these high levels of gene flow at such a large geographic scale, a concerted nation-wide effort will be necessary to effectively manage invasive *Rattus* in the U.S. and stop recolonization of eradicated areas from unmanaged populations.

In addition to overall propagule pressure and dispersal, interspecific interactions and ecological context must also be considered in the management of *Rattus* in the U.S. If interaction between *R. rattus* and *R. norvegicus* is the key factor in limiting the ongoing establishment of new *R. rattus* individuals, as has been previously suggested (Nowak 1999), then the patterns of propagule pressure and gene flow detected for *R. rattus* in this study are only applicable in a context where these two species co-occur. Management plans in the U.S. should therefore be developed jointly for these two species. Also, because *R. norvegicus* appear to be stifling *R. rattus* colonization and dispersal, management efforts should target *R. rattus* first, and only target *R. norvegicus* after successful removal of *R. rattus*. Removal of *R. norvegicus* prior to *R. rattus* may allow for a significant influx of *R. rattus* propagules.

Disease ecology implications

Because of their unparalleled role in the spread of zoonotic disease (Gratz 1984), it is important to consider our analysis in the context of rat-borne pathogens. In the study of an infectious disease, the ability to predict the spread of the causative agent is extremely valuable in stopping disease transmission and minimizing negative impacts. Population genetic analysis of the host can provide insight into the rate and geographic extent of transmission for a pathogen (Holmes 2008; Biek and Real 2010), and these approaches have been effective in understanding the epidemiology and transmission history of many zoonotic pathogens including rabies in raccoons (*Procyon lotor*; Cullingham *et al*. 2009) and dengue fever and its mosquito vector, *Aedes aegypti* (Urdaneta-Marquez & Failloux 2010).

For the U.S., our analyses suggest *R. norvegicus* presents a greater risk than *R. rattus* in their ability to bring a diversity of pathogens from various international sources and spread them across the U.S. The mean pairwise migration rate among our sampled localities was higher for *R. norvegicus* than for *R. rattus*, suggesting gene flow among populations is in general higher for the Norway rat, and migration rates, cluster analyses, and Mantel test suggest long-distance dispersal events are more frequent for *R. norvegicus*. Under this scenario, an infectious pathogen found in *R. norvegicus* can not only spread rapidly, but it is extremely difficult to predict patterns of dispersion when long-distance dispersals are frequent, as has been observed for some bat-borne pathogens (Breed *et al*. 2010). Another issue with rat-borne pathogens in the U.S. is the potential diversity of source locations from which incoming rats may originate. Whereas our analyses suggest *R. rattus* are not entering the U.S. and assimilating with the established populations, *R. norvegicus* does appear to be entering the U.S. on both coasts, and likely from multiple distinct source localities. A major concern in disease epidemiology is the potential for increased virulence to arise through recombination among distinct strains or genotypes (e.g., Gibbs *et al.* 2001; Grigg *et al.* 2001; He *et al.* 2009). If *R. norvegicus* are entering the U.S. from distinct source localities

on opposite coasts and then spreading across the U.S., as our analyses suggest, potentially distinct pools of rat-borne zoonotic pathogens may be coming into contact within the borders of the U.S. Given the human health risk, this certainly warrants further study.

Conclusions

This study represents a first approximation of colonization history and contemporary dispersal for invasive *Rattus* at a large geographic scale and in a country with a complex transportation infrastructure. While these scenarios for the origin of each haplogroup are speculative, they represent starting hypotheses for future study and further geographic sampling. We detected clear differences in colonization history and contemporary patterns of gene flow and propagule pressure for invasive *Rattus* in the U.S., suggesting differences in the ability of these species to spread zoonotic pathogens. Moreover, we found evidence that ecological interactions between *R. rattus* and *R. norvegicus* may be driving the contemporary distribution of genetic diversity for *R. rattus*, illustrating the importance of considering both population genetics and ecological parameters in modeling species invasions and developing effective management apporoaches. Our analyses, while informative, require a much more thorough sampling of international localities to understand the importance of various potential sources in generating current patterns of genetic diversity; this is especially true for European populations, where historical records indicate both the *R. rattus* and *R. norvegicus* lineages that invaded the U.S. originated. Finally, a more thorough sampling within the U.S. is needed to increase resolution in terms of migration rates, allowing us to identify the major routes of dispersal for these species (e.g., the Mississippi River waterway vs the I-35 transportation corridor for north/south human-aided dispersal).

References

Abdelkrim J, Pascal M, Calmet C, Samadi S (2005a) Importance of assessing population genetic structure before eradication of invasive species: examples from insular Norway rat populations. *Conservation Biology*, **19**, 1509–1518.

Abdelkrim J, Pascal M, Calmet C, Samadi S (2005b) Island colonization and founder effects: the invasion of the Guadeloupe islands by ship rats (*Rattus rattus*). *Molecular Ecology*, **14**, 2923–2931.

Allendorf FW, Lundquist LL (2003) Introduction: population biology, evolution, and control of invasive species. *Conservation Biology*, **17**, 24–30.

Aplin KP, Suzuki H, Chinen AA *et al.* (2011) Multiple genetic origins of commensalism and complex dispersal history of black rats. *PLoS One*, **6**, e26357.

Armitage PL (1993) Commensal rats in the New World, 1492–1992. *Biologist*, **40**, 174–178.

Bastos AD, Nair D, Taylor PJ *et al.* (2011) Genetic monitoring detects an overlooked cryptic species and reveals the diversity and distribution of three invasive *Rattus* congeners in South Africa. *BMC Genetics*, **12**, 26.

Biek R, Real LA (2010) The landscape genetics of infectious disease emergence and spread. *Molecular Ecology*, **19**, 3515–3531.

Breed AC, Field HE, Smith CS, Edmonston J, Meers J (2010) Bats without borders: long-distance movements and implications for disease risk management. *EcoHealth*, **7**, 204–212.

Caire W, Tyler JD, Glass BP, Mares MA (1989) Mammals of Oklahoma. University of Oklahoma Press, Norman, OK.

Cheylan G (1998) Distribution of genetic diversity within and between Mediterranean island populations of the black rat *Rattus rattus* (L. 1758). *Biological Journal of the Linnean Society*, **63**, 393–408.

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 and Genetic Systems*, **80**, 367–375.

Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1660.

Courchamp F, Chapuis JL, Pascal M (2003) Mammal invaders on islands: impacts, control, and control impact. *Biological Reviews*, 78:347–383.

Crawford NG (2010) SMOGD: software for the measurement of genetic diversity. *Molecular Ecology Resources*, **10**, 556–557.

Cullingham CI, Kyle CJ, Pond BA, Rees EE, White BN (2009) Differential permeability of rivers to raccoon gene flow corresponds to rabies incidence in Ontario, Canada. *Molecular Ecology*, **19**, 3592–3602.

Drummond AJ, Ashton B, Cheung M *et al.* (2010) Geneious, version 5.0. http://www.geneious.com.

Drummond AJ, Rambaut A (2003) Tracer v1.5. http://tree.bio.ed.ac.uk/software/tracer/.

Ecke DH (1954) An invasion of Norway rats in southwest Georgia. *Journal of Mammalogy*, **35**, 521–525.

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

Frankham R, Ralls K (1998) Conservation biology: inbreeding leads to extinction. *Nature*, **392**, 441–442.

Frittelli JF, Lee MR, Medalia J, O'Rourke R, Perl R (2005) Port and maritime security: background and issues for congress. Novinka Books, New York.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **133**, 693–709.

Galimand M (1997) Multidrug resistance in *Yersinia pestis* mediated by a transferable plasmid. *New England Journal of Medicine*, **337**, 677–680.

Gardner-Santana LC, Norris DE, Fornadel CM, Hinson ER, Klein SL, Glass GE (2009) Commensal ecology, urban landscapes, and their influence on the genetic characteristics of city-dwelling Norway rats (*Rattus norvegicus*). *Molecular Ecology*, **18**, 2766–2778.

Gibbs MJ, Armstrong JS, Gibbs AJ (2001) Recombination in the hemagglutinin gene of the 1918 "Spanish flu". *Science*, **293**, 1842–1845.

Goudet J (2001) *FSTAT, a Program to Estimate and Test Gene Diversities and Fixation Indices (version 2.9.3)*. Department of Ecology and Evolution, University of Lausanne, Switzerland, available at: http://www2.unil.ch/popgen/softwares/fstat.htm. Accessed 10 January 2012.

Gratz NG (1984) The global public health importance of rodents. Pp. 413–435 in, Proceedings of a conference on the organization and practice of vertebrate pest control (A. C. Dubock, ed.). World Health Organization, Geneva, Switzerland.

Greenwood PJ (1980) Mating systems, philopatry and dispersal in birds and mammals. *Animal Behavior*, **28**, 1140–1162.

Grigg ME, Bonnefoy S, Hehl AB, Suzuki Y, Boothroyd JC (2001) Success and virulence in *Toxoplasma* as the result of sexual recombination between two distinct ancestries. *Science*, **294**, 161–165.

Guillemaud T, Beaumont MA, Ciosi M, Cornuet JM Estoup A (2010) Inferring introduction routes of invasive species using approximate Bayesian computation. *Heredity*, **104**, 88–99.

He CQ, Xie ZX, Han GZ, *et al.* (2009) Homologous recombination as an evolutionary force in the avian influenza A virus. *Molecular Biology and Evolution*, **26**, 177–187.

Hingston M, Goodman SM, Ganzhorn JU, Sommer S (2005) Reconstruction of the colonization of southern Madagascar by *Rattus rattus*. *Journal of Biogeography*, **32**, 1549–1559.

Holmes EC (2008) Evolutionary history and phylogeography of human viruses. *Annual Reviews of Microbiology*, **62**, 307–328.

Howald G, Donlan CJ, Galvan JP, *et al.* (2007) Invasive rodent eradication on islands. *Conservation Biodiversity*, **21**, 1258–1268.

Inglesby TV, Dennis DT, Henderson DA *et al*. (2000) Plague as a biological weapon. *Journal of the American Medical Association*, **283**, 2281–2290.

Island Conservation (2006) Evaluation of primary and secondary exposure risks to land bird species for: experimental use permit application for field efficacy trial of 0.0025% brodifacoum broadcast bait to eradicate introduced rats from Aleutian Islands in the Alaska Maritime National Wildlife Refuge, July 20, 2006. Island Conservation, Center for Ocean Health, Santa Cruz, California.

Jackson WB (1982) Norway rat and allies. In: *Wild Mammals of North America* (eds Chapman JA, Feldhammer GA). pp. 1077–1088, Johns Hopkins University Press, Baltimore, Maryland.

Jakobsson M and 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.

Kaluza P, Kolzsch, Gastner MT, Blasius B (2010) The complex network of global cargo ship movements. *Journal of the Royal Society Interface*, **7**, 1093–1103.

Keeling MJ, Gilligan CA (2000) Metapopulation dynamics of bubonic plague. *Nature*, **407**, 903–906.

Kimura M (1953) "Stepping-stone" model of population. *Annual Report of the National Institute of Genetics Japan*, **3**, 62–63.

Kolbe JJ, Glor RE, Schettino LR, Lara AC, Larson A, Losos JB (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, **431**, 177–181.

Kolbe JJ, Larson A, Losos JB, de Queiroz K (2008) Admixture determines genetic diversity and population differentiation in the biological invasion of a lizard species. *Biology Letters*, **4**, 434–437.

Lack JB, Greene DU, Conroy CJ, Hamilton MJ, Braun JK, Mares MA, Van Den Bussche RA (In Review) Invasion facilitates extensive hybridization with introgression between three black rat lineages in the United States and Asia. *Molecular Ecology*.

Landon DB (1991) *Zooarchaeology and urban foodways: a case study from eastern Massachusetts*. Unpublished Dissertation, Boston University, Massachusetts.

Le Roux J, Wieczorek AM (2009) Molecular systematics and population genetics of biological invasions: towards a better understanding of invasive species management. *Annals of Applied Biology*, **154**, 1–17.

Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology and Evolution*, **17**, 386–391.

MacGillivray W (1838) *A history of British quadrupeds*. WH Lizars, Edinburgh, Scotland.

Maddison DR, Maddison WP (2000) MacClade 4: analysis of phylogeny and character evolution. Sinauer Associates, Inc., Sunderland, Massachusetts.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209–220.

Martel C, Viard F, Bourguet D, Garcia-Meunier P (2004) Invasion by the marine gastropod *Ocinebrellus orinata* in France I. Scenario for the source of introduction. *Journal of Experimental Marine Biology and Ecology*, **305**, 155–170.

Matisoo-Smith E, Robins JH (2004) Origins and dispersals of Pacific people: evidence from mtDNA phylogenies of the Pacific rat. *Proceedings of the National Academies of Sciences of the United States of America*, **101**, 9167–9172.

Matisoo-Smith E, Roberts RM, Irwin GJ, Allen JS, Penny D, Lambert DM (1998) Patterns of prehistoric human mobility in Polynesia indicated by mtDNA from the Pacific rat. *Proceedings of the National Academies of Sciences, USA*, **95**, 15145–15150.

Meerburg BG, Singleton GR, Kijlstra A (2009) Rodent-borne diseases and their risks for public health. *Critical Reviews in Microbiology*, **35**, 221–270.

Musser GG, Carleton MD (2005) Family Muridae. In *Mammal species of the world, a taxonomic and geographic reference* (eds Wilson DE, Reeder DM), pp. 1189–1531. Smithsonian Institute Press, Washington D.C.

Novak SJ, Mack RN (2001) Tracing plant introduction and spread: genetic evidence from *Bromus tectorum* (cheatgrass). *BioScience*, **51**, 114–122.

Nowak, R. 1999. Walker's Mammals of the World (6th Edition). John's Hopkins University Press, Baltimore, Maryland.

Peakall R, Smouse PE (2006) Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Pimentel D, Lach L, Zuniga R, Morrison D (2000) Environmental and economic costs of nonindigenous species in the United States. *BioScience*, **50**, 53–65.

Prentis PJ, Wilson JRU, Dormontt EE, Richardson DM, Lowe AJ (2008) Adaptive evolution in invasive species. *Trends in Ecology and Evolution*, **13**, 288–294.

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. *Journal of Heredity*, **86**, 248–249.

Robertson BC, Gimmell NJ (2004) Defining eradication units in pest control progammes. *Journal of Applied Ecology*, **41**, 1032–1041.

Robins JH, Hingston M, Matisoo-Smith E, Ross HA (2007) Identifying *Rattus* species using mitochondrial DNA. *Molecular Ecology Notes*, **7**, 717–729.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.

Rosenberg NA (2004) Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes*, **14**, 4299–4312.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP. DNA polymorphism analyses by coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Ruffina L, Vidal E (2010) Early colonization of Mediterranean islands by *Rattus rattus*: a review of zooarcheological data. *Biological Invasions*, **12**, 2389–2394.

Russell JC, Abdelkrim J, Fewster RM (2009) Early colonization population structure of a Norway rat island invasion. *Biological Invasions*, **11**, 1557–1567.

Sakai AK, Allendorf FW, Holt JS *et al.* (2001) The population biology of invasive species. *Annual Reviews of Ecology and Systematics*, 32, 305–332.

Schwartz CW, Schwartz ER (1981) The wild mammals of Missouri. University of Missouri Press, Columbia.

Silver J (1927) The introduction and spread of house rats in the United States. *Journal of Mammalogy*, **8**, 58–60.

Simberloff D (2009) We can eliminate invasions or live with them. Successful management projects. *Biological Invasions*, **11**, 149–157.

Skellman JG (1951) Random dispersal in theoretical populations. *Biometrika*, **38**, 196–218.

Suarez AV, Holoway DA, Case TJ (2001) Patterns of spread in biological invasions dominated by long-distance jump dispersal: insights from Argentine ants.*Proceedings of the National Academy of Sciences USA*, **98**, 1095–1100.

Sullivan R (2004) Rats: observations on the history and habitat of the city's most unwanted inhabitants. Bloomsbury Publishing, New York City, New York.

Taylor DR, Keller SR (2007) Historical range expansion determines the phylogenetic diversity introduced during contemporary species invasion. *Evolution*, **61**, 334–345.

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. *Journal of Biogeography*, **37**, 398–410.

Twigg G (1975) *The brown rat*. David and Charles, London, England.

Urdaneta-Marquez L, Failloux AB (2010) Population genetic structure of *Aedes aegypti*, of the principle vector of dengue viruses. *Infection, Genetics and Evolution*, **11**, 253–261.

Weir BS (1996) Genetic data analysis II: methods for discrete population genetic data. Sinauer Associate, Inc., Sunderland, Massachusetts.

Wilson AB, Naish KA, Boulding EG (1999) Multiple dispersal strategies of the invasive quagga mussel (*Dreissena bugensis*) as revealed by microsatellite analysis. *Canadian Journal of Fisheries and Aquatic Species*, **56**, 2248–2261.

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*, **163**, 1177–1191.

Figure Captions

Figure 2.1. Statistical parsimony haplotype network generated from cytochrome b sequences for *R. rattus*. Colored nodes on haplotype networks correspond to unique haplotypes, and the size of the circle corresponds to the frequency of the haplotype. Color and size of the pie for each haplotype corresponds to the sampling locality indicated in the map by a colored dot and the frequency of that haplotype at that locality, respectively. The small black nodes represent extinct or unsampled haplotypes, and each uninterrupted straight line (independent of line length) corresponds to a single mutational step.

Figure 2.2. Statistical parsimony haplotype network generated from cytochrome b sequences for the *R. norvegicus*. Colored nodes on the haplotype networks correspond to unique haplotypes, and the size of the circle corresponds to the frequency of the haplotype. The color and size of the pie for each haplotype corresponds to the sampling locality indicated in the map by a colored dot and the frequency of that haplotype at that locality, respectively. Small black nodes represent extinct or unsampled haplotypes, and each uninterrupted straight line (independent of line length) corresponds to a single mutational step.

Figure 2.3. Mismatch distribution showing frequency of pairwise differences in cytochrome *b* sequence for all sampled *R. rattus* (A) and *R. norvegicus* (B) in the U.S. Observed distributions are represented by the black line, and the expected distribution under the sudden expansion model represented by the grey line.

Figure 2.4. Resulting Mantel tests of a significant relationship between genetic distance (D_{EST}) based on the microsatellite dataset and straight-line geographic distance for *R. rattus* (A) and *R. norvegicus* (B). Only populations from the U.S. and with ≥5 individuals sampled were included.

Figure 2.5. Principle coordinate analysis (PCoA) of genetic variation based on the microsatellite dataset for *R. rattus* and *R. norvegicus*.

Figure 2.6. Results of the Bayesian clustering analysis implemented in Structure under the admixture model for *R. rattus* and *R. norvegicus* based on the microsatellite dataset. Vertical columns represent assignment probabilities to each of the inferred clusters identified for the optimal value of *K* (*K* = 4 for *R. rattus* and *K* = 3 *R. norvegicus*).

Table 2.1. Descriptive statistics and population sample sizes (n) for the *R. rattus* mtDNA dataset, illustrating the number of haplotypes (*h*) and nucleotide diversity (π) for population with multiple samples.

Locality	**n**	**Haplotypes**	***h* (SD)**	**π (SD)**
Shemya Is., AK	23	1	--	--
Great Sitkin Is., AK	1	1	--	--
Little Rock, AR	6	2	0.333 (0.215)	0.00058 (0.00038)
San Francisco Bay Area, CA	23	8	0.791 (0.063)	0.00128 (0.00021)
Tehama, CA	17	1	--	--
Gainesville, FL	21	2	0.257 (0.110)	0.00023 (0.00010)
San Diego, CA	2	1	--	--
Panama City, FL	4	1	--	--
Miami, FL	3	1	--	--
Key Largo, FL	10	2	0.533 (0.095)	0.00327 (0.00058)
Baton Rouge, LA	11	2	0.182 (0.144)	0.00016 (0.00013)
Jefferson Davis, MS	1	1	--	--
Brownsville, TX	1	1	--	--
Houston, TX	7	1	--	--
Weatherford, TX	1	1	--	--
San Angelo, TX	14	6	0.681 (0.132)	0.00120 (0.00028)
Austin, TX	13	4	0.769 (0.072)	0.00157 (0.00015)
Seattle, WA	5	4	0.900 (0.161)	0.00246 (0.00058)
Total	163	21	0.806 (0.019)	0.00180 (0.00023)

Table 2.2. Descriptive statistics and population sample sizes (n) for the *R. norvegicus* mtDNA dataset, illustrating the number of haplotypes (*h*) and nucleotide diversity (π) for population with multiple samples.

Locality	**n**	**Haplotypes**	***h* (SD)**	**π (SD)**
Adak Is., AK	41	2	0.290 (0.078)	0.00204 (0.00055)
Attu Is., AK	2	1	--	--
Great Sitkin Is., AK	5	1	--	--
Sedanka Is. AK	1	1	--	--
Revillagigedo Is., AK	2	1	--	--
Fairbanks, AK	1	1	--	--
Little Rock, AR	3	2	0.667 (0.314)	0.00058 (0.00028)
San Diego, CA	11	4	0.491 (0.175)	0.00108 (0.00053)
Chicago, IL	11	3	0.473 (0.162)	0.00159 (0.00067)
Spencer, IN	10	1	--	--
Baltimore, MD	20	2	0.1 (0.088)	0.00018 (0.00015)
Albuquerque, NM	2	1	--	--
New York City, NY	1	1	--	--
Oklahoma City, OK	2	1	--	--
Corvalis, OR	4	2	0.500 (0.265)	0.00395 (0.00209)
Union, PA	42	1	--	--
Memphis, TN	16	3	0.342 (0.140)	0.00049 (0.00021)
Austin, TX	1	1	--	--
San Angelo, TX	1	1	--	--
Seattle, WA	1	1	--	--
Monroe, WV	7	5	0.905 (0.103)	0.00560 (0.00537)
Total	184	11	0.638 (0.028)	0.00294 (0.00020)

Table 2.3. Descriptive statistics and population sample sizes (n) for the *R. rattus* microsatellite dataset, illustrating the mean number of alleles across all loci (*Na*), the number of private alleles (*NP*), the observed (*HO*) and expected (*HE*) heterozygosity, and gene diversity for each population.

Sampling Locality	**n**	***Na***	***NP***	***HO***	***HE***	**Gene Diversity**
Austin, TX	13	6.667	3	0.624	0.723	0.757
Shemya Is. AK	24	4.778	7	0.486	0.541	0.554
San Francisco Bay, CA	29	8.667	6	0.678	0.76	0.775
San Diego, CA	2	2.667	7	0.611	0.556	--
Panama City, FL	31	7.111	0	0.616	0.673	0.685
Key Largo, FL	10	5.444	0	0.711	0.699	0.737
Miami, FL	3	1.889	0	0.407	0.265	--
Gainesville, FL	21	3.556	2	0.574	0.469	0.488
Houston, TX	7	4.222	2	0.46	0.59	0.649
Little Rock, AR	6	3.667	8	0.519	0.557	0.617
Baton Rouge, LA	9	5.444	1	0.704	0.662	0.702
Jefferson Davis, MS	1	1.444	0	0.444	0.222	--
Tehama, CA	17	4.111	1	0.531	0.61	0.632
Brownsville, TX	1	1.667	0	0.667	0.333	--
San Angelo, TX	14	6.333	1	0.635	0.67	0.697
Weatherford, TX	1	1.444	0	0.444	0.222	--
Seattle, WA	5	5.222	1	0.711	0.693	0.781

Table 2.4. Descriptive statistics and population sample sizes (n) for the *R. norvegicus* microsatellite dataset, illustrating the mean number of alleles across all loci (*Na*), the number of private alleles (*NP*), the observed (*HO*) and expected (*HE*) heterozygosity, and gene diversity for each population.

Sampling Locality	**n**	***Na***	***NP***	***HO***	***HE***	**Gene Diversity**
Austin, TX	1	1.444	0	0.444	0.222	--
Adak Is., AK	51	6.333	8	0.621	0.696	0.704
Attu Is., AK	2	1.889	0	0.556	0.375	--
Douglas Is., AK	1	1.111	1	0.111	0.056	--
Fairbanks, AK	1	1.778	1	0.778	0.389	--
Kagalaska Is., AK	1	1.556	0	0.556	0.278	--
Revillagigedo Is., AK	2	2.333	1	0.5	0.444	--
Sedanka Is., AK	1	1.333	0	0.333	0.167	--
Sitkin Is., AK	4	1.889	0	0.278	0.26	--
Baltimore, MD	30	7.889	7	0.522	0.684	0.699
San Diego, CA	11	6.111	9	0.525	0.7	0.744
Chicago, IL	11	3.667	2	0.434	0.459	0.483
Spencer, IN	10	2.333	1	0.589	0.425	0.439
Little Rock, AR	3	1.556	0	0.259	0.216	--
Memphis, TN	16	3.111	3	0.417	0.401	0.414
Albuquerque, NM	2	2	0	0.389	0.403	--
New York City, NY	1	1.444	0	0.444	0.222	--
Oklahoma City, OK	2	1.778	0	0.278	0.264	0.389
Corvalis, OR	4	2.667	0	0.444	0.458	0.537
Union, PA	42	3.111	0	0.329	0.358	0.362
San Angelo, TX	1	1.333	0	0.333	0.167	--

Seattle, WA	1	1.444	1	0.444	0.222	--
Monroe, WV	6	5	4	0.444	0.659	0.746

Table 2.5. Pairwise proportion of ancestry estimates based on the analysis of the microsatellite dataset in BayesAss. Only populations with ≥5 individuals sampled were included.

	Austin, TX	Shemya Is., AK	San Francisco, CA	Panama City, FL	Key Largo, FL	Gainesville, FL	Houston, TX	Little Rock, AR	Baton Rouge, LA	Tehama, CA	San Angelo, TX	Seattle, WA
Austin, TX	0.6809	0.0136	0.0148	0.0142	0.0135	0.0135	0.0136	0.0136	0.0135	0.0136	0.1817	0.0136
Shemya, AK	0.0102	0.8951	0.0088	0.0087	0.0103	0.0087	0.0101	0.0103	0.0102	0.0086	0.0087	0.0104
San Francisco, CA	0.0079	0.0078	0.8802	0.0081	0.0079	0.0156	0.0078	0.0078	0.0079	0.0083	0.0327	0.0079
Panama City, FL	0.0073	0.0073	0.0086	0.9161	0.0073	0.0087	0.0073	0.0074	0.0073	0.0075	0.0078	0.0074
Key Largo, FL	0.015	0.0151	0.1508	0.0309	0.6827	0.0151	0.015	0.0151	0.0151	0.015	0.0152	0.015
Gainesville, FL	0.0094	0.0094	0.0096	0.0098	0.0095	0.8957	0.0094	0.0095	0.0095	0.0094	0.0094	0.0095
Houston, TX	0.0174	0.0293	0.0176	0.0173	0.0174	0.0174	0.6852	0.0175	0.0174	0.0176	0.1285	0.0174
Little Rock, AR	0.0184	0.0185	0.0232	0.0185	0.0185	0.0213	0.0185	0.6866	0.0185	0.0185	0.1211	0.0185
Baton Rouge, LA	0.0158	0.0158	0.1425	0.0297	0.0158	0.0158	0.0159	0.0159	0.6834	0.0158	0.0177	0.0159
Tehama, CA	0.0108	0.0108	0.0109	0.0108	0.0109	0.0109	0.0109	0.0108	0.0109	0.8804	0.0109	0.0109

San Angelo, TX	0.0126	0.0128	0.0179	0.021	0.0126	0.0128	0.0126	0.0126	0.0126	0.0144	0.8453	0.0128
Seattle, WA	0.0198	0.0196	0.0405	0.0209	0.0196	0.0196	0.0196	0.0198	0.0197	0.065	0.0196	0.6886

Table 2.6. Pairwise proportion of ancestry estimates based on the analysis of the microsatellite dataset in BayesAss. Only populations with ≥5 individuals sampled were included.

	Adak, AK	Baltimore, MD	San Diego, CA	Chicago, IL	Spencer, IN	Memphis, TN	Corvalis, OR	Pennsylvania	West Virginia
Adak, AK	0.9545	0.0058	0.0057	0.0057	0.0057	0.0057	0.0056	0.0057	0.0056
Baltimore, MD	0.0091	0.9175	0.0102	0.0092	0.0111	0.0121	0.0112	0.0085	0.0111
San Diego, CA	0.0166	0.0165	0.7135	0.1476	0.0237	0.0166	0.0245	0.0166	0.0245
Chicago, IL	0.0165	0.0181	0.0165	0.8659	0.0164	0.0169	0.0165	0.0169	0.0164
Spencer, IN	0.0168	0.1974	0.0168	0.0167	0.6854	0.0167	0.0167	0.0168	0.0167
Memphis, TN	0.0133	0.0138	0.0133	0.0134	0.0133	0.8928	0.0133	0.0133	0.0133
Corvalis, OR	0.0257	0.0259	0.026	0.1208	0.0259	0.0265	0.6976	0.0258	0.0259
Pennsylvania	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.0067	0.9464	0.0067
West Virginia	0.0235	0.05	0.0227	0.1053	0.0227	0.0373	0.0227	0.0227	0.693

Fig. 2.1

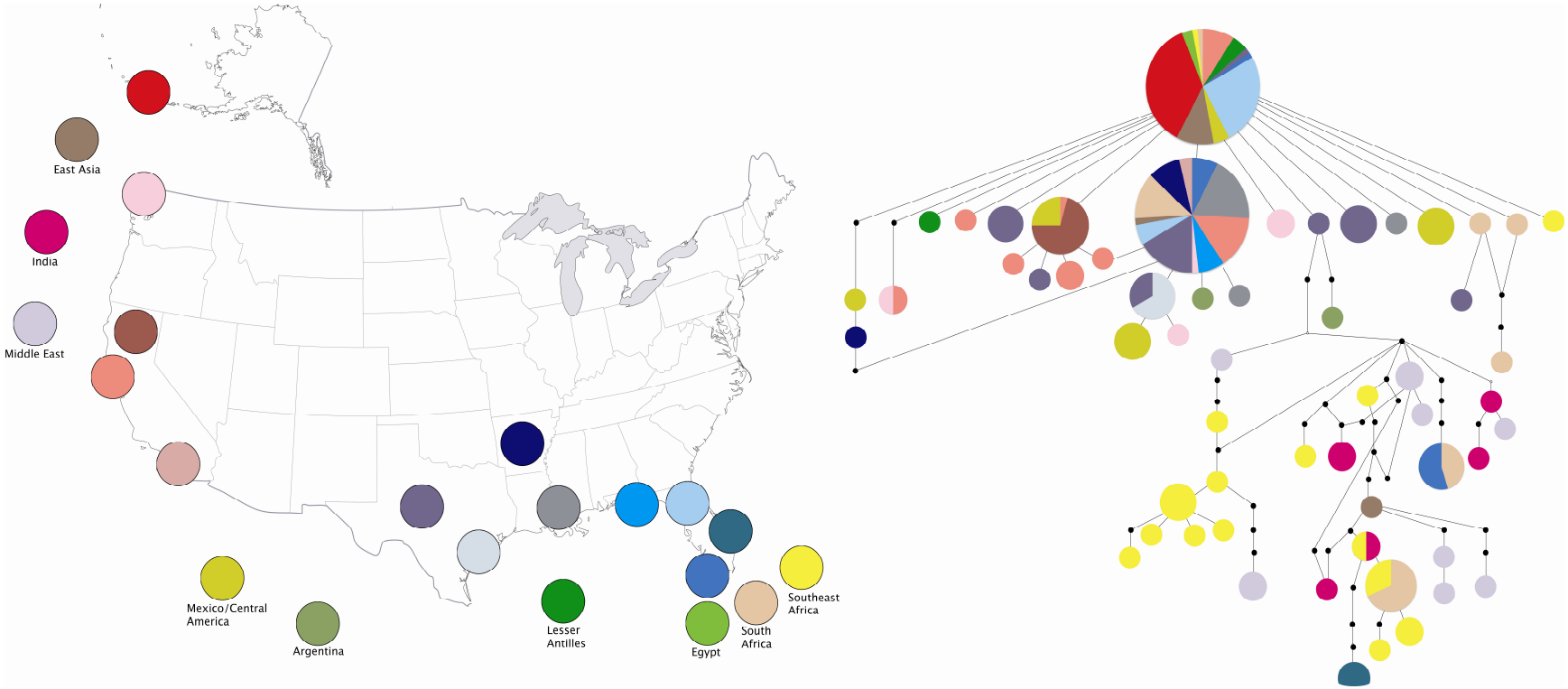

East Asia
India
Middle East
Mexico/Central America
Argentina
Lesser Antilles
Egypt
South Africa
Southeast Africa

Fig. 2.2

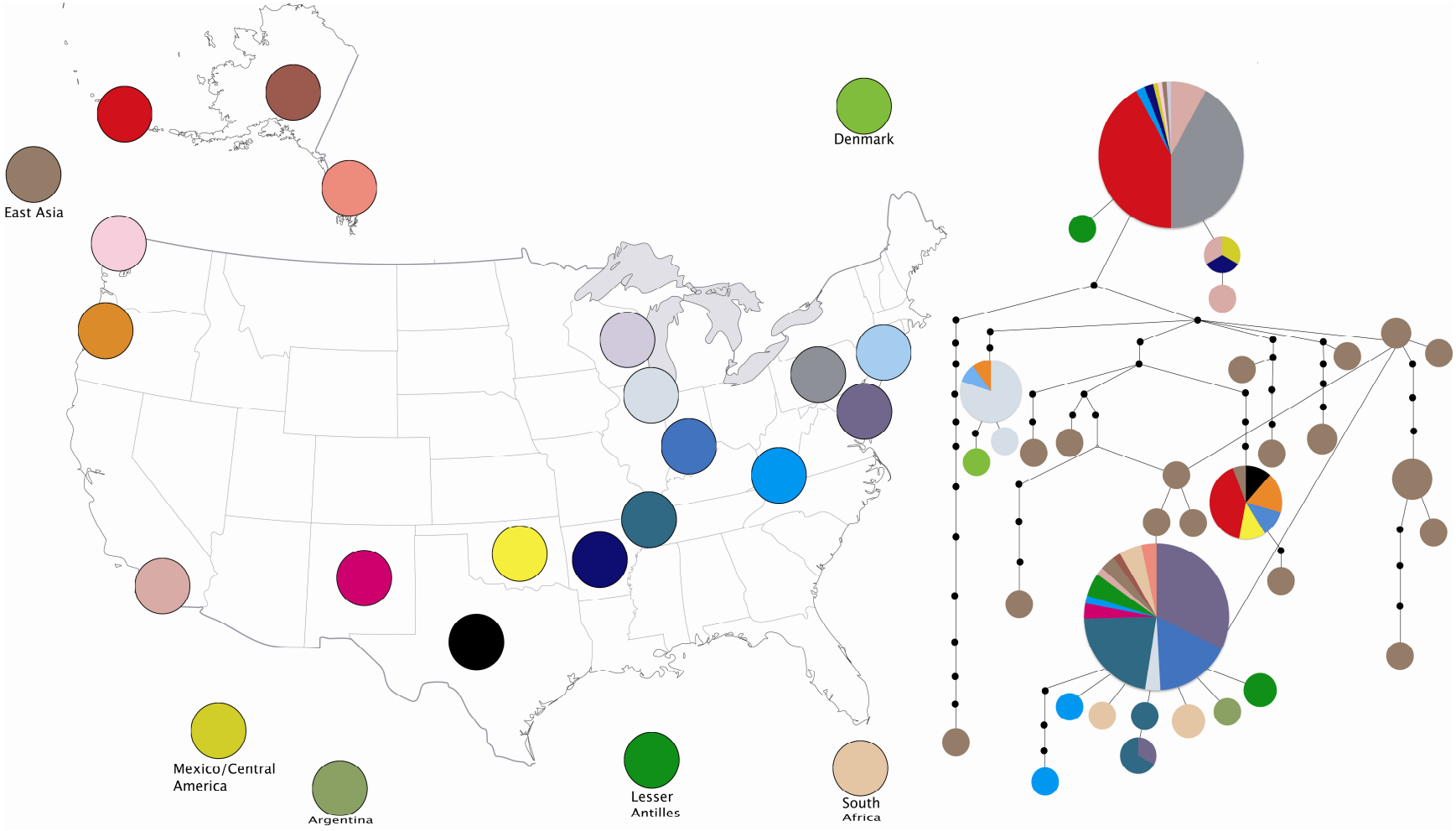

Denmark
East Asia
Mexico/Central America
Argentina
Lesser Antilles
South Africa

Fig. 2.3

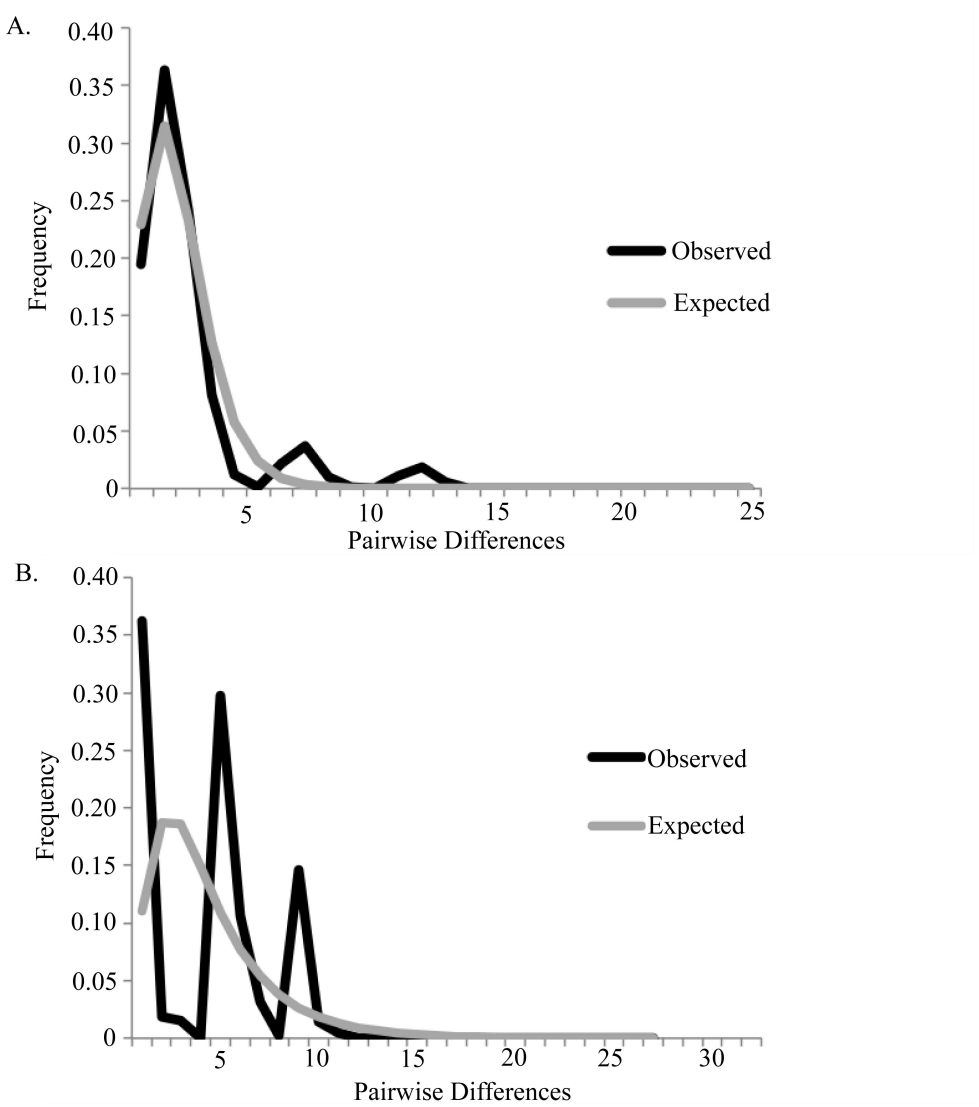

A.
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0
Frequency
5
10
15
20
25
Pairwise Differences
Observed
Expected
B.
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0
Frequency
5
10
15
20
25
30
Pairwise Differences
Observed
Expected

Fig. 2.4

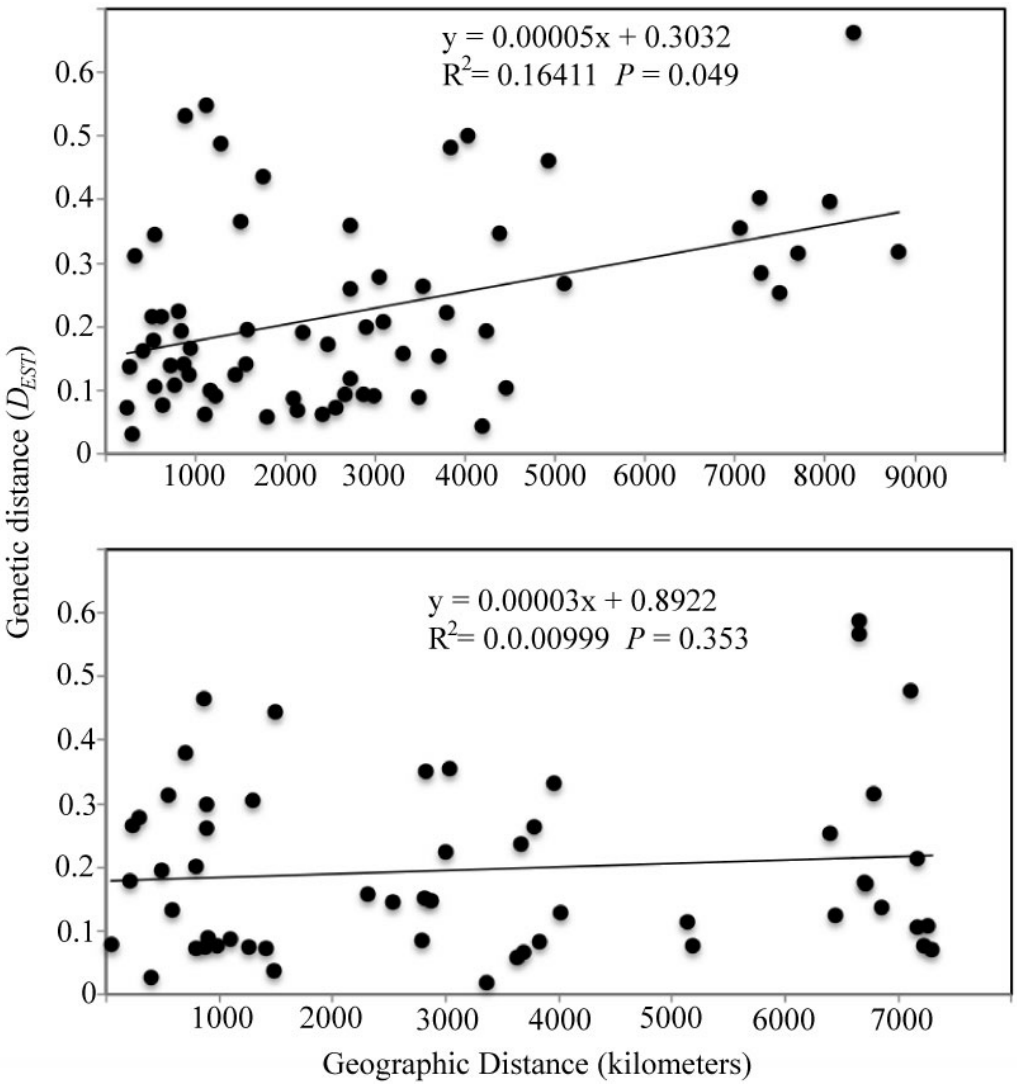

y = 0.00005x + 0.3032
R2= 0.16411 P = 0.049
y = 0.00003x + 0.8922
R2= 0.0.00999 P = 0.353
Genetic distance (DEST)
Geographic Distance (kilometers)
0
0.1
0.2
0.3
0.4
0.5
0.6
1000
2000
3000
4000
5000
6000
7000
8000
9000

Fig. 2.5

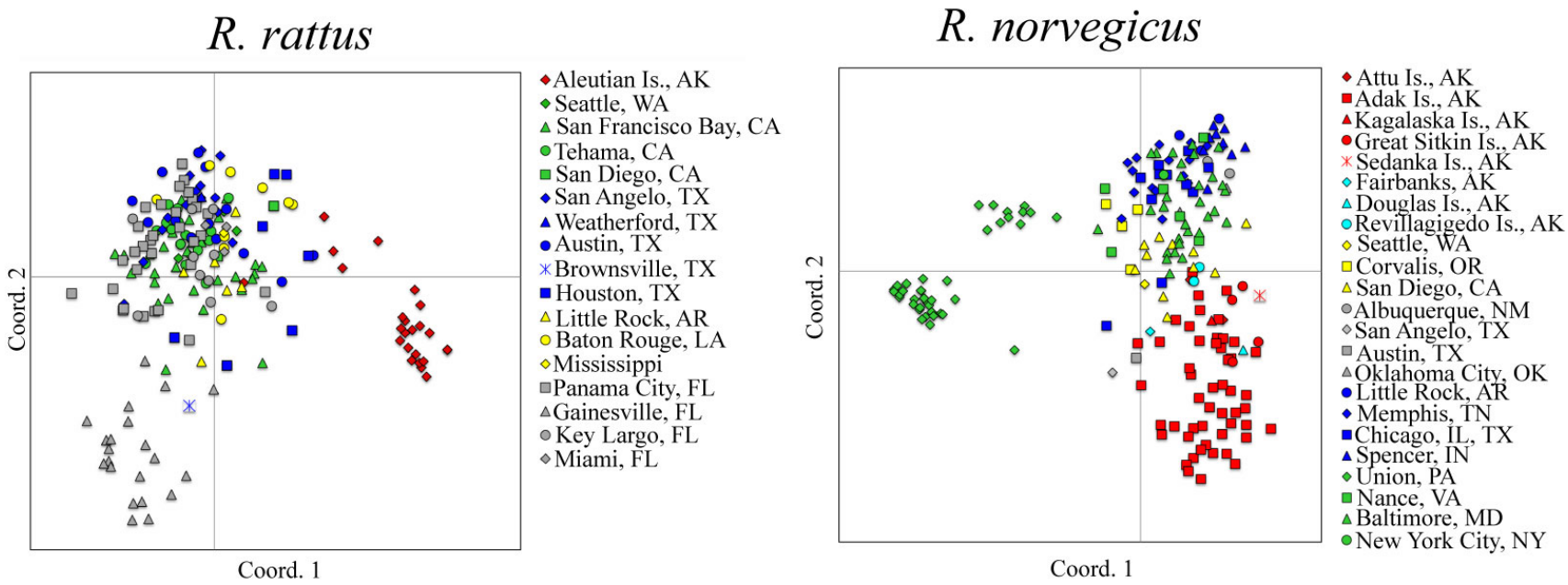
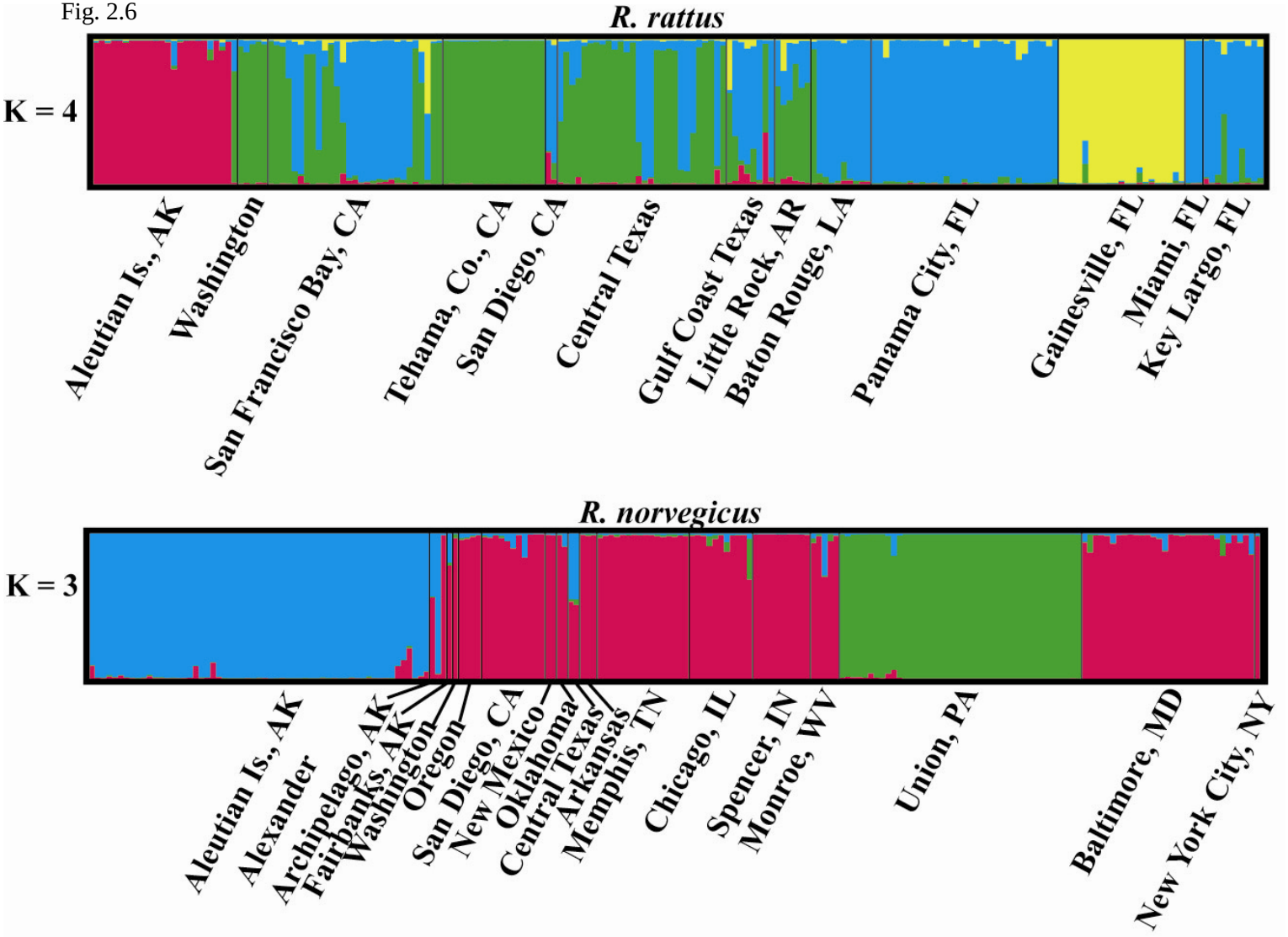

R. rattus
Coord. 2
Coord. 1
Aleutian Is., AK
Seattle, WA
San Francisco Bay, CA
Tehama, CA
San Diego, CA
San Angelo, TX
Weatherford, TX
Austin, TX
Brownsville, TX
Houston, TX
Little Rock, AR
Baton Rouge, LA
Mississippi
Panama City, FL
Gainesville, FL
Key Largo, FL
Miami, FL
R. norvegicus
Coord. 2
Coord. 1
Attu Is., AK
Adak Is., AK
Kagalaska Is., AK
Great Sitkin Is., AK
Sedanka Is., AK
Fairbanks, AK
Douglas Is., AK
Revillagigedo Is., AK
Seattle, WA
Corvalis, OR
San Diego, CA
Albuquerque, NM
San Angelo, TX
Austin, TX
Oklahoma City, OK
Little Rock, AR
Memphis, TN
Chicago, IL, TX
Spencer, IN
Union, PA
Nance, VA
Baltimore, MD
New York City, NY

Fig. 2.6


R. rattus
K = 4
Aleutian Is., AK
Washington
San Francisco Bay, CA
Tehama, Co., CA
San Diego, CA
Central Texas
Gulf Coast Texas
Little Rock, AR
Baton Rouge, LA
Panama City, FL
Gainesville, FL
Miami, FL
Key Largo, FL
R. norvegicus
K = 3
Aleutian Is., AK
Alexander Archipelago, AK
Fairbanks, AK
Washington
Oregon
San Diego, CA
New Mexico
Oklahoma
Central Texas
Arkansas
Memphis, TN
Chicago, IL
Spencer, IN
Monroe, WV
Union, PA
Baltimore, MD
New York City, NY

CHAPTER III

IDENTIFICATION OF ZOONOTIC HEPATITIS E IN WILD *RATTUS*

Abstract

The role of rodents in the epidemiology of zoonotic hepatitis E virus (HEV) has been a subject of considerable debate. Seroprevalence studies suggest widespread HEV infection in commensal *Rattus*, but transmissions have been largely unsuccessful and the recovery of zoonotic genotype 3 HEV RNA from wild *Rattus* has never been confirmed. We surveyed *R. rattus* and *R. norvegicus* from across the U.S. and several international populations using a hemi-nested RT-PCR approach. We recovered HEV RNA in liver tissue from 35 individuals out of 446 examined. All but one of these isolates was relegated to the zoonotic HEV genotype 3, and the remaining sequence represented the recently discovered rat genotype from the U.S. and Germany. Positive individuals were detected in both urban and remote localities. Genetic analyses suggest all HEV genotype 3 infection we detected for *Rattus* is the result of a single strain.

Introduction

The hepatitis E virus (HEV) is traditionally considered an important cause of acute hepatitis in developing countries, where outbreaks arise most often through the fecal contamination of drinking water or following flooding (Aggarwal 2011). Major outbreaks have been reported in India, SE Asia, Africa, and Mexico, and mortality rates are considerable among pregnant women (20–30%) (Aggarwal 2011). In industrialized countries, HEV infections are reported sporadically and contamination of drinking water is an unlikely source, but, cases are increasing as diagnostic tests are being performed more frequently (Miyamura 2011). Moreover, zoonotic transmission of HEV through the consumption of undercooked pork and deer meat has been confirmed (Tei *et al.* 2003; Yazaki *et al.* 2003), and the detection of HEV in a range of mammalian hosts suggests the potential for multiple zoonotic sources of HEV infection in industrialized countries (Meng 2010).

There are currently at least four genotypes of HEV known to infect humans. Genotypes 1 and 2 have been identified only from humans, and are responsible for the majority of outbreaks in developing countries (Lu & Hagedorn 2005). Genotypes 3 and 4 are those thought to be involved in zoonotic transmission, and have been isolated from swine (both domesticated pig and wild boar), deer, mongoose, rabbits, cattle, and humans (Meng 2010). Additional strains not known to infect humans have also been identified from rats and chickens, and it is likely the genetic diversity of HEV is only beginning to be understood.

Within the U.S., isolated HEV infections have been identified in travelers who have visited developing countries (Bader *et al.* 1991) and it is clear that for several at-risk groups in the U.S. (i.e., swine veterinarians and farmers) the high number of reported seropositives is the result of swine-human contact (Meng *et al.* 2002; Meng 2011). However, seroepidemiological examinations of blood banks in the U.S. and other industrialized countries have revealed high proportions of samples positive for antibodies to HEV (excluding individuals that have travelled

to HEV-endemic countries), and this was true even in urban areas where swine-human contact is essentially absent (Mast *et al.* 1997; Thomas *et al.* 1997; Meng *et al*. 2002). Also, the consumption of raw pork and wild game is uncommon in the U.S., although it is a common practice in other industrialized nations where high HEV seroprevalence has been reported (i.e., France) (Mansuy *et al.* 2011). This suggests that in addition to travel to HEV-endemic regions and swine-human contact additional reservoirs of HEV infection exist in the U.S. and evidence has accumulated indicating rodents as a potential HEV reservoir (Kasbrane-Lazizi *et al.* 1999; Favorov *et al.* 2000; Arankalle *et al.* 2001; Hirano *et al.* 2003; Easterbrook *et al.* 2007). In a survey of 26 rodent species in the U.S., Favorov *et al*. (2000) found individuals seropositive for anti-HEV antibodies in 14 species. Urban populations possessed approximately twice the proportion of seropositive individuals relative to rural populations, and the commensal *Rattus* species (*R. rattus* and *R. norvegicus*) exhibited the highest proportion of positives (Favorov *et al.* 2000).

The role of invasive *Rattus* as reservoirs in the epidemiology and transmission of HEV is unclear, but their ubiquity in urban environments and unparalleled propensity to carry zoonotic pathogens makes them an obvious target of investigation. Multiple studies have documented the presence of IgG and IgM anti-HEV antibodies from *R. norvegicus* and *R. rattus* populations across the U.S. and Asia (Kasbrane-Lazizi *et al.* 1999; Favorov *et al.* 2000; Arankalle *et al.* 2001; Hirano *et al.* 2003; Easterbrook *et al.* 2007). Shukla *et al*. (2011) were able to successfully infect cell lines derived from *Mus musculus*—a murid rodent closely related to *Rattus*—with HEV genotype 3. In addition, Maneerat *et al*. (1996) were able to experimentally infect laboratory *R. norvegicus* with HEV isolated from infected humans, although it is unclear which genotype. Following inoculation with the virus, the human strain was able to effectively replicate in multiple tissues and HEV RNA was detectable in the feces and serum for over 30 days post exposure, suggesting *R. norvegicus* are capable of replicating and transmitting human strains of

HEV. However, the recent discovery of a rat-specific strain of HEV not known previously to infect humans (Johne *et al.* 2010a, 2010b; Purcell *et al.* 2011) suggests the high anti-HEV seroprevalence may be a result of cross-reactivity instead of widespread infection with a human-infecting HEV genotype.

We utilized an RT-PCR approach to survey *R. rattus* and *R. norvegicus* for the presence of HEV RNA. Our analysis detected HEV RNA in liver tissues from both *R. rattus* and *R. norvegicus* at multiple localities across the U.S. Sequencing of RT-PCR positives revealed geographically widespread infection with the zoonotic HEV genotype 3, and a single individual in California positive for the rat-specific strain. This analysis suggests wild invasive *Rattus* are competent hosts for genotype 3 HEV.

Material and Methods

Rat Tissues

We obtained liver tissue samples from 446 *R. rattus* and *R. norvegicus* from museum collections (see Appendix for list of specimens and source collections), covering localities primarily in the U.S. (15 states) plus additional samples from China, Honduras, Madagascar, Mexico, Nicaragua, Peru, Russia, and Vietnam (Table 3.1). To maximize the likelihood of intact viral RNA, all liver samples selected for this study had been dissected from recently euthanized specimens, immediately frozen, and maintained at -80 ºC until thawed for extraction.

Hemi-nested Reverse Transcription PCR

Total RNA was extracted from about 30 mg of liver tissue by using the RNeasy Mini Kit (Qiagen). We used a modification of the broad-spectrum RT-PCR approach outlined by Johne et al. (2010b) to amplify a 334-basepair fragment of ORF1, and all primer sequences are given in Johne et al. (2010b). Initial attempts to amplify the ORF1 fragment from total RNA extractions

resulted in amplification of a portion of an unidentified transcript in all *R. rattus* samples. Upon sequencing the amplicon, it was clear that the spurious amplification was the result of non-specific binding of primer HEV-cas. To circumvent this, we used a hemi-nested approach, with the initial RT-PCR using the HEV-cs/HEV-casN primer combination, and the nested PCR using the HEV-csN/HEV-casN primer combination. With the exception of the change in primer combinations, all other aspects of the amplification followed the protocol of Johne *et al*. (2010b). Positive PCR amplicons (verified by gel electrophoresis) were purified using the Wizard SV Gel PCR Clean-Up System (Promega), and sequenced in both directions using the nested PCR primers (HEVcsN/HEVcasN).

Given the high sensitivity of a nested PCR approach, contamination can be a significant issue, and has been cited as problematic in previous investigations of HEV in rodents (He *et al.* 2002). Exceptional effort was made to ensure no contamination occurred. All PCR steps were conducted in a sterile environment, under a laminar flow hood, with all surfaces, tubes, and equipment UV irradiated between each PCR setup. Furthermore, this entire project was conducted in a newly constructed laboratory where no HEV samples (or any other animal samples) had been previously handled, extractions were conducted in a completely separate room from all PCR amplifications, and all steps (extraction, RT-PCR and nested PCR) included negative controls. In addition, a single HEV genotype 3 isolate was used as a positive control in PCRs, and we sequenced this isolate for the same locus targeted for the *Rattus* samples. Any *Rattus* HEV isolate exhibiting 100% nucleotide identity to this positive control sequence was excluded as contamination.

Phylogenetic Analyses

In addition to the sequences generated above, we downloaded all complete HEV genome sequences from GenBank (Accession numbers are given in Fig. 3.1), and extracted the

approximately 334-basepair homologous portion of ORF1 from each genome. The total sequences were aligned using the MAFFT aligner (Katoh *et al.* 2002) implemented in Geneious v5.5 (Drummond *et al.* 2011). We conducted Bayesian, maximum parsimony (MP), and maximum likelihood (ML) phylogenetic analyses on the combined alignment using the avian HEV strain as an outgroup. For the Bayesian analysis conducted in MrBayes v3.2 (Huelsenbeck & Ronquist 2001), we partitioned the alignment by codon position and utilized a GTR + I + Γ substitution model, which Modeltest v3.7 (Posada & Crandall 1998) indicated to be most appropriate. The analysis was run for 15,000,000 generations sampled every 1,000 generations, and burnin values were determined empirically by evaluating likelihood scores. For the MP analysis, we used tree-bisection reconnection branch-swapping, 25 random additions of input taxa, and 1,000 bootstrap replicates to assess node support. For the ML analysis, we utilized a GTR + I + Γ substitution model as indicated above, nearest-neighbor interchange branch-swapping, and 500 bootstrap replicates to assess node support. Finally, we generated a haplotype network for the sequences generated in this study using TCS (Clement *et al.* 2000).

Results

We identified 35 PCR-positive individuals out of 446 *Rattus* examined (7.85%). The majority of positive samples were from California (15 individuals), but there were also positives from Tennessee, Florida, Oklahoma, Pennsylvania, Texas, and Alaska (Table 3.1). Phylogenetic analysis placed 34 of these positives in a very closely related group within the HEV genotype 3 clade (Fig. 3.1), and this placement was supported in all analyses. Mean pairwise uncorrected genetic distance between the HEV genotype 3 sequences and the other known HEV genotypes was 36.19%, 24.12%, 24.91%, 24.05%, and 33.52% compared to the avian genotype, genotype 2, genotype 1, genotype 4, and the rat genotype, respectively. The network analysis revealed the HEV genotype 3 sequences obtained from *Rattus* form a tight cluster (Fig. 3.2), differing by only a few mutations, and therefore represent a single strain. Mean pairwise sequence divergence

within the *Rattus* HEV genotype 3 sequences was 0.51%. A single previously published sequence (AF082843) isolated from a HEV-infected pig was also in this group.

The single sequence not nested within the genotype-3 clade was isolated from a *R. norvegicus* from the San Francisco Bay area of California. Phylogenetic analyses placed it in a strongly supported clade with two other sequences isolated from *R. norvegicus* in Germany (Fig. 3.1). Uncorrected genetic distances indicate that the California rat HEV sequence is approximately twice as divergent from the two sequences recovered from Germany (California vs GU345042 = 13.98%; California vs GU345043 = 14.86%) as the two Germany sequences are from each other (GU345042 vs GU345043 = 7.78%), suggesting some degree of distinction between the U.S. and European rat HEV strains.

Discussion

Significant conflict has surrounded the role of *Rattus* (and other rodents) in HEV epidemiology since seroprevalence studies in the 1990s identified widespread positives for HEV antibodies in multiple species in the U.S. and Asia (Kasbrane-Lazizi *et al.* 1999; Arankalle *et al.* 2001; Hirano *et al.* 2003; Easterbrook *et al.* 2007). Maneerat *et al*. (1996) reported successful infection of three Wistar laboratory rats (*R. norvegicus*) with HEV (viral RNA was detected intermittently in the feces for at least 30 days), but it is unclear which genotype was used and this result has not been duplicated (Meng 2011). In addition, He *et al.* (2002) reported the isolation of HEV genotype 1 from *R. rattus* and *Bandicota bengalensis*, but the study was later retracted because the authors were unable to rule out contamination as a source of the detected viral RNA. More recently, Shukla *et al*. (2011) were able to successfully infect multiple *M. musculus* cell cultures (in addition to infecting cow, rabbit, cat, dog, and chicken cultures) with HEV genotype 3, supporting the hypothesis that rodents may be competent hosts. However, there was substantial variation among different strains of HEV genotype 3 in the ability to infect cells derived from

different hosts, including swine and human. A recent attempt to infect adult Sprague-Dawley laboratory rats (*R. norvegicus*) with HEV genotypes 1, 2, and 3 failed (Purcell *et al.* 2011). In this same study, infection of laboratory rats with the divergent rat genotype had limited success, with only 25% of intravenously infected Sprague-Dawley rats seroconverting and only 15.8% of nude rats seroconverting. This is unexpected given that approximately 80% of wild-caught *R. norvegicus* from Los Angeles, California, where the study was conducted, were positive for anti-HEV IgG and/or IgM antibodies, suggesting infection readily occurs in the wild (Purcell *et al.* 2011). Johne *et al*. (2010b) were similarly unable to successfully infect rat liver cell lines with rat genotype HEV isolated from wild-caught *R. norvegicus* from Germany. To add to this narrative, we now present evidence for HEV genotype 3 infection in wild-caught *R. rattus* and *R. norvegicus*.

The geographic spread of positive individuals indicates infection in wild *Rattus* is not isolated to any portion of the U.S. or to urban areas. Our positive samples included both *Rattus* species, and included both the relatively remote Aleutian Islands, Alaska population and the urban San Francisco Bay, California population. Given the commensal nature of invasive *Rattus* and their unparalleled ability to utilize human transportation vectors in dispersal (i.e., commercial shipping), the prevalence of HEV in even remote populations is not necessarily surprising. Recent work examining the genetic structure of *R. rattus* has shown that two different mtDNA haplotypes have each been rapidly spread from their origin in India to every continent, with the exception of Antarctica (Aplin *et al.* 2011*;* Lack *et al.* In Press). Our data suggest HEV has similarly been dispersed with their rat hosts. Furthermore, *R. rattus* and *R. norvegicus* are sympatric over essentially their entire contemporary range, so the lack of genetic distinction between the strains infecting these two species is not unexpected (Figs. 3.1 and 3.2).

In terms of infection rates, it is important to consider the variation in the handling of tissues in field-collected animals. While we attempted to limit our analysis to the most well-

preserved tissues, there is undoubtedly considerable variation among collection protocols and collectors in the length of time between euthanasia and dissection, the time between dissection and freezing, the number of times tissues are thawed and frozen (i.e., in sorting, subsampling, shipping, etc.), and the consistency of storage temperature. All of these factors can lead to nucleic acid degradation and negatively impact our ability to detect viral RNA. Therefore, our infection rate is likely not indicative of HEV infection rates in wild *Rattus* populations.

As stated above, recent studies have indicated significant variation in terms of the diversity of competent mammalian hosts for various strains of HEV genotype 3 (Shukla *et al.* 2011). Although seroprevalence studies have suggested infection rates as high as 80% for HEV in U.S. *Rattus* populations (Kabrane-Lazizi *et al.* 1999*;* Purcell *et al.* 2011), attempts to infect different lab strains of *R. norvegicus* with a genotype isolated from wild-caught *R. norvegicus* have experienced limited success, with the majority of attempts failing even in immunocompromised nude rats (Johne *et al.* 2010b*;* Purcell *et al.* 2011). The above patterns, as well as the low genetic diversity of the positive samples detected in this study (Fig. 3.2), suggest only a limited number of HEV genotype 3 strains may be capable of infecting *Rattus* and other rodents (i.e. *Mus*). Also, the difficulty in transmitting virus from infected wild-caught *R. norvegicus* into laboratory strains indicates certain life-history and/or genetic characteristics may be important in dictating infection success. Purcell *et al*. (2011) noted the positive correlation between antibody prevalence and age in their examination of *Rattus* seroprevalence, suggesting that rats are readily infected in the wild, and that infection occurs as juveniles. This pattern is consistent with HEV infection in humans and swine as well (Arankalle *et al.* 1995; Meng *et al.* 1999), and suggests infections should be attempted on both wild-caught and laboratory strain juvenile rats. Lending further support to this, the only report of significant long-term infection (>30 days) of rats with HEV used weanling rats (*19*), while all other attempts we are aware of have used only adult rats (Purcell *et al.* 2011). In addition, the extreme variation in host

specificity that Shukla *et al*. (2011) observed among different HEV genotype 3 strains indicates the need for future transmission studies to include as many strains as possible.

References

Aggarwal R (2011) Hepatitis E: historical, contemporary and future perspectives. Gastroenterol *Hepatology*, **26**, 72–82.

Aplin KP, Suzuki H, Chinen AA, Chesser RT, Have JT, Donnellan SC, *et al*. (2011) Multiple geographic origins of commensalism and complex dispersal history of black rats. *PLoS One*, **6**, e26357.

Arankalle VA, Tsarev SA, Chadha MS, Alling DW, Emerson SU, Banerjee K, *et al.* (1995) Age-specific prevalence of antibodies to hepatitis A and E viruses in Pune, India. *Journal of Infectious Disease*, **171**, 447–450.

Arankalle VA, Joshi MV, Kulkarni AM, Gandhe SS, Chobe LP, Rautmare SS, Mishra AC, Padbidri VS (2001) Prevalence of anti-hepatitis E virus antibodies in different Indian animal species. *Journal of Viral Hepatitis*, **8**, 223–227.

Bader TF, Krawczynski K, Polish LB, Favorov MO (1991) Hepatitis E in a U.S. traveller to Mexico. *New England Journal of Medicine*, **325**, 1659.

Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1660.

Drummond AJ, Ashton B, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A (2011) Geneious, version 5.0. Available from http://www.geneious.com.

Easterbrook JD, Kaplan JB, Vanasco NB, Reeves WK, Purcell RH, Kosoy MY, Glass GE, Watson J, Klein SL (2007) A survey of zoonotic pathogens carried by Norway rats in Baltimore, Maryland, USA. *Epimediology and Infection*, **135**, 1192–1199.

Favorov MO, Kosoy MY, Tsarev SA, Childs JE, Margolis HS (2000) Prevalence of antibody to hepatitis E virus among rodents in the United States. *Journal of Infectious Disease*, **181**, 449–455.

He J, Innis BL, Shrestha MP, Clayson ET, Scott RM, Linthicum KJ, Musser GG, Gigliotti SC, Binn LN, Kuschner RA, Vaughn DW (2002) Evidence that rodents are a reservoir of hepatitis E virus for human in Nepal. *Journal of Clinical Microbiology*, **40**, 4493–4498.

Hirano M, Ding X, Li TC, Takeda N, Kawabata H, Koizumi N, Kadosoka T, Goto I, Masuzawa T, Nakamura M, Taira K, Kuroki T, Tanikawa T, Watanabe H, Abe K (2003) Evidence for widespread infection of hepatitis E virus among wild rats in Japan. *Hepatology Research*, **27**, 1–5.

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

Johne R, Heckel G, Plenge-Bonig A, Kindler E, Maresch C, Reetz J, Schielke A, Ulrich R (2010a) Novel hepatitis E virus genotype in Norway rats, Germany. *Emerging Infectious Diseases*, **16**, 1452–1455.

Johne R, Plenge-Bonig A, Hess M, Ulrich RG, Reetz J, Schielke A (2010b) Detection of a novel hepatitis E-like virus in faeces of wild rats using a nested broad-spectrum RT-PCR. *Journal of General Virology*, **91**, 750–758.

Kasbrane-Lazizi Y, Fine JB, Elm J, Glass GE, Higa H, Diwan A, Gibbs Jr. CJ, Meng XJ, Emerson SU, Purcell RH (1999) Evidence for widespread infection of wild rats with

hepatitis E virus in the United States. *American Journal of Tropical Medicine and Hygiene*, **61**, 331–335.

Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acid Research*, **30**, 3059–3066.

Lack JB, Greene D, Conroy C, Hamilton MJ, Braun JK, Mares MA, Van Den Bussche RA (In Press) Invasion facilitates extensive hybridization between three black rat species in the United States and Asia. *Molecular Ecology*.

Lu L, Li C, Hagedorn CH (2005) Phylogenetic analysis of global hepatitis E virus sequences: genetic diversity, subtypes and zoonosis. *Reviews in Medical Virology*, **16**, 5–36.

Maneerat Y, Clayson ET, Myint KSA, Young GD, Innis BL (1996) Experimental infection of the laboratory rat with the hepatitis E virus. *Journal of Medical Virology*, **48**, 121–128.

Mansuy JM, Bendall R, Legrand-Abravanel F, Saune K, Miedouge M, Ellis V, Rech H, Destruel F, Kamar N, Dalton HR, Izopet J (2011) Hepatitis E virus antibodies in blood donors, France. *Emerging Infectious Diseases*, **17**, 2309–2312.

Mast EE, Kuramoto IK, Favorov MO, Schoening VR, Burkholder BT, Shapiro CN, Holland PV (1997) Prevalence of and risk factors for antibody to hepatitis E virus seroreactivity among blood donors in northern California. *Journal of Infectious Disease*, **176**, 34–40.

Meng XJ (2010) Hepatitis E virus: animal reservoirs and zoonotic risk. *Veterinary Microbiology*, **140**, 256–265.

Meng XJ (2011) From barnyard to food table: the omnipresence of hepatitis E virus and risk for zoonotic infection and food safety. *Virus Research*, **161**, 23–30.

Meng XJ, Dea S, Engle RE, Friendship R, Lyoo YS, Sirinarumitr T, *et al* (1999) Prevalence of antibodies to the hepatitis E virus in pigs from countries where hepatitis E is common or is rare in the human population. *Journal of Medical Virology* **59**, 297–302.

Meng XJ, Wiseman B, Elvinger F, Guenette DK, Toth TE, Engle RE, Emerson SU, Purcell RH (2002) Prevalence of antibodies to hepatitis E virus in veterinarians working with swine and in normal blood donors in the United States and other countries. *Journal of Clinical Microbiology*, **40**, 117–122.

Miyamura T (1998) Hepatitis E virus infection in developed countries. *Virus Research*, **161**, 40–46.

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Purcell RH, Engle RE, Rood MP, Kabrane-Lazizi Y, Nguyen HT, Govindarajan S, St. Claire M, Emerson SU (2011) Hepatitis E virus in rats, Los Angeles, California, USA. *Emerging Infectious Diseases*, **17**, 2216–2222.

Shukla P, Nguyen HT, Torian U, Engle RE, Faulk K, Dalton HR, Bendall RP, Keane FE, Purcell RH, Emerson SU. Cross-species infections of cultured cells by hepatitis E virus and discovery of an infectious virus-host recombinant. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 2438–2443.

Tei S, Kitajima N, Takahashi K, Mishiro S (2003) Zoonotic transmission of hepatitis E virus from deer to human beings. *Lancet*, **362**, 371–373.

Thomas DL, Yarbough PO, Vlahov D, Tsarev SA, Nelson KE, Saah AJ, Purcell RH (1997) Seroreactivity to hepatitis E virus in areas where the disease is not endemic. *Journal of Clinical Microbiology*, **35**, 1244–1247.

Yazaki Y, Mizuo H, Takahashi M, Nishizawa T, Sasaki N, Gotanda Y, Okamoto H (2003) Sporadic acute or fulminant hepatitis E in Hokkaido, Japan, may be food-borne, as suggested by the presence of hepatitis E virus in pig liver as food. *Journal of General Virology*, **84**, 2351–2357.

Figure Captions

Figure 3.1. Bayesian phylogram resulting from analysis of a 334-basepair fragment of HEV ORF1. Node labels represent Bayesian posterior probabilities/ML bootstrap/MP bootstrap values, respectively, and are given only for the nodes at the base of each genotype and nodes uniting the genotypes. For the sequences generated in this study, the terminal taxa labels correspond to tissue accession numbers given in the Appendix, followed by the species from which the tissue originated and the source population. For all other HEV genotype 3 sequences, the GenBank accession number is given, followed by the species from which the viral RNA was isolated. For all other sequences included, only the GenBank accession number is given.

Figure 3.2. Genetic network illustrating the relationship among all HEV genotype 3 sequences isolated from rats in this study. Each continuous line represents a single mutational event and the smallest unlabeled circles represent extinct or unsampled sequences. Each large circle represents a single unique sequence, and the label corresponds to the tissue number given in the Appendix and the general sampling locality.

Table 3.1. Sample sizes for each *Rattus* species at each general locality, as well as the total number of individuals positive for HEV RNA. More detailed specific locality information for each individual included in this study is given in the Appendix.

	Sample Size		
Locality	***R. norvegicus***	***R. rattus***	**Positives**
U.S.A			
Aleutian Islands, Alaska	18	7	6
San Francisco Bay Area, California	19	112	12
Gainesville, Florida		21	4
Oklahoma City, Oklahoma	1		1
Memphis, Tennessee	16		6
San Angelo, Texas	2	11	2
Little Rock, Arkansas	2	6	0
San Diego, California	8	5	3
Panama City, Florida		24	0
Key Largo, Florida		5	0
Spencer, Indiana		10	0
Baton Rouge, Louisiana		12	0
Prentiss, Mississippi		1	0
Bernallio, New Mexico	2		0
Union Co., Pennsylvania	40		1
Corvalis, Oregon	4		0
Houston, Texas		8	0
Austin, Texas		14	0
Kerns, West Virginia	1		0

Seattle, Washington	1	5	0
Vietnam		18	0
China		5	0
Honduras		2	0
Madagascar		5	0
Mexico	1	2	0
Niacaragua	1	11	0
Peru		16	0
Russia	30		0

Figure 3.1.

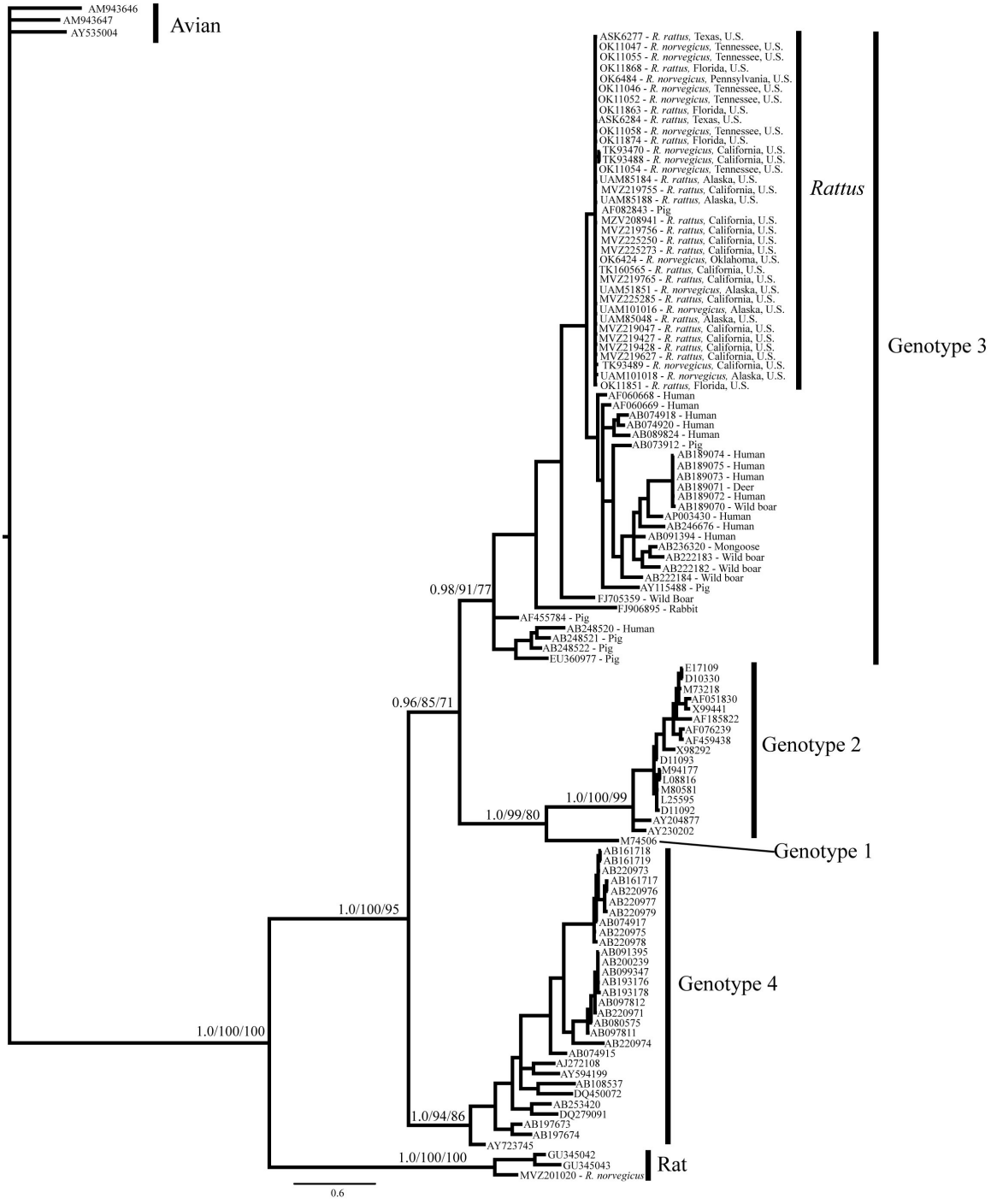
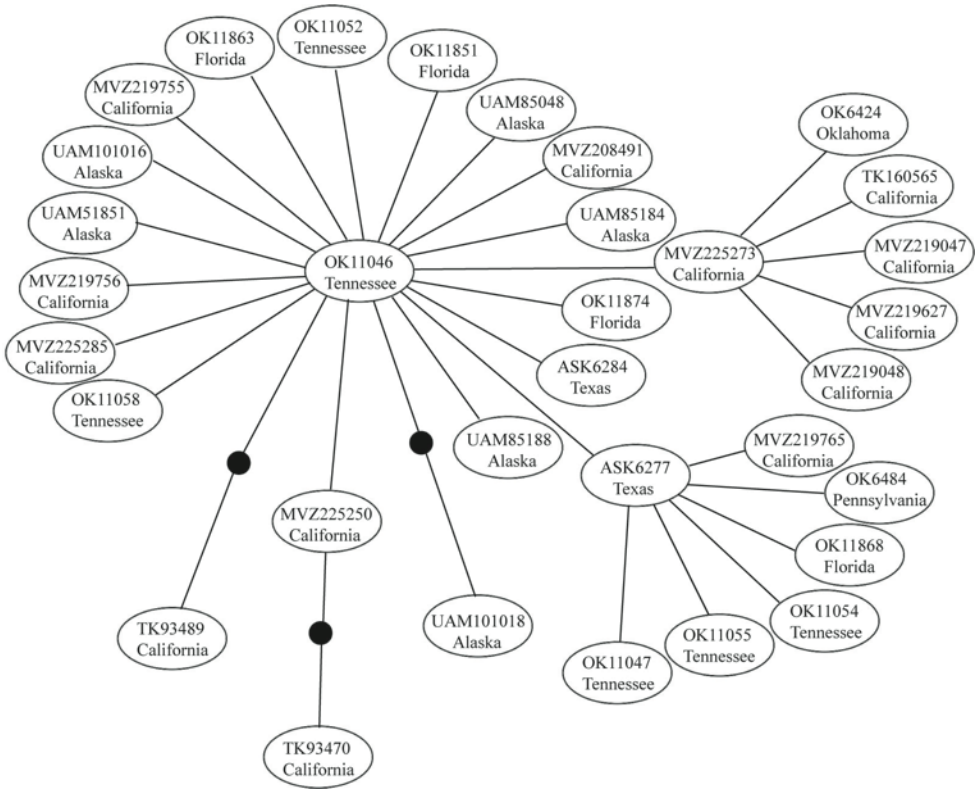

AM943646
AM943647
AY535004
Avian
ASK6277 - R. rattus, Texas, U.S.
OK11047 - R. norvegicus, Tennessee, U.S.
OK11055 - R. norvegicus, Tennessee, U.S.
OK11868 - R. rattus, Florida, U.S.
OK6484 - R. norvegicus, Pennsylvania, U.S.
OK11046 - R. norvegicus, Tennessee, U.S.
OK11052 - R. norvegicus, Tennessee, U.S.
OK11863 - R. rattus, Florida, U.S.
ASK6284 - R. rattus, Texas, U.S.
OK11058 - R. norvegicus, Tennessee, U.S.
OK11874 - R. rattus, Florida, U.S.
TK93470 - R. norvegicus, California, U.S.
TK93488 - R. norvegicus, California, U.S.
OK11054 - R. norvegicus, Tennessee, U.S.
UAM85184 - R. rattus, Alaska, U.S.
MVZ219755 - R. rattus, California, U.S.
UAM85188 - R. rattus, Alaska, U.S.
AF082843 - Pig
MZV208941 - R. rattus, California, U.S.
MVZ219756 - R. rattus, California, U.S.
MVZ225250 - R. rattus, California, U.S.
MVZ225273 - R. rattus, California, U.S.
OK6424 - R. norvegicus, Oklahoma, U.S.
TK160565 - R. rattus, California, U.S.
MVZ219765 - R. rattus, California, U.S.
UAM51851 - R. norvegicus, Alaska, U.S.
MVZ225285 - R. rattus, California, U.S.
UAM101016 - R. norvegicus, Alaska, U.S.
UAM85048 - R. rattus, Alaska, U.S.
MVZ219047 - R. rattus, California, U.S.
MVZ219427 - R. rattus, California, U.S.
MVZ219428 - R. rattus, California, U.S.
MVZ219627 - R. rattus, California, U.S.
TK93489 - R. norvegicus, California, U.S.
UAM101018 - R. norvegicus, Alaska, U.S.
OK11851 - R. rattus, Florida, U.S.
Rattus
Genotype 3
AF060668 - Human
AF060669 - Human
AB074918 - Human
AB074920 - Human
AB089824 - Human
AB073912 - Pig
AB189074 - Human
AB189075 - Human
AB189073 - Human
AB189071 - Deer
AB189072 - Human
AB189070 - Wild boar
AP003430 - Human
AB246676 - Human
AB091394 - Human
AB236320 - Mongoose
AB222183 - Wild boar
AB222182 - Wild boar
AB222184 - Wild boar
AY115488 - Pig
FJ705359 - Wild Boar
FJ906895 - Rabbit
AF455784 - Pig
AB248520 - Human
AB248521 - Pig
AB248522 - Pig
EU360977 - Pig
0.98/91/77
E17109
D10330
M73218
AF051830
X99441
AF185822
AF076239
AF459438
X98292
D11093
M94177
L08816
M80581
L25595
D11092
AY204877
AY230202
Genotype 2
0.96/85/71
1.0/100/99
1.0/99/80
M74506
Genotype 1
AB161718
AB161719
AB220973
AB161717
AB220976
AB220977
AB220979
AB074917
AB220975
AB220978
AB091395
AB200239
AB099347
AB193176
AB193178
AB097812
AB220971
AB080575
AB097811
AB220974
Genotype 4
1.0/100/95
1.0/100/100
AB074915
AJ272108
AY594199
AB108537
DQ450072
AB253420
DQ279091
1.0/94/86
AB197673
AB197674
AY723745
GU345042
GU345043
MVZ201020 - R. norvegicus
Rat
1.0/100/100
0.6

Figure 3.2.


OK11863
Florida
OK11052
Tennessee
OK11851
Florida
MVZ219755
California
UAM85048
Alaska
OK6424
Oklahoma
UAM101016
Alaska
MVZ208491
California
TK160565
California
UAM51851
Alaska
UAM85184
Alaska
MVZ219047
California
OK11046
Tennessee
MVZ225273
California
MVZ219756
California
OK11874
Florida
MVZ219627
California
MVZ225285
California
ASK6284
Texas
MVZ219048
California
OK11058
Tennessee
UAM85188
Alaska
MVZ219765
California
ASK6277
Texas
OK6484
Pennsylvania
MVZ225250
California
OK11868
Florida
TK93489
California
UAM101018
Alaska
OK11054
Tennessee
OK11055
Tennessee
OK11047
Tennessee
TK93470
California

CHAPTER IV

EVOLUTIONARY GENETIC IMPACTS OF A GLOBAL BLACK RAT INVASION

Abstract

Biological invasions result in novel species interactions, which can have significant evolutionary impacts on both native and invading taxa. One evolutionary concern with invasions is hybridization among lineages that were previously isolated, but make secondary contact in their invaded range(s). Black rats, consisting of several morphologically indistinguishable but genetically distinct taxa that collectively have invaded all seven continents, are arguably the most successful invaders on the planet. We used mitochondrial cytochrome *b* sequences, two nuclear gene sequences, and nine microsatellite loci to examine the distribution of three invasive black rat lineages (*R. tanezumi*, *R. rattus* I, and *R. rattus* IV) in the U.S. and Asia, and determine the extent of hybridization among these taxa. Our analyses revealed two mitochondrial lineages that have spread to multiple continents, including a previously undiscovered population of *R. tanezumi* in the U.S., whereas the third lineage (*R. rattus* IV) has spread throughout Southeast Asia, but no further. Analyses of nuclear DNA (both sequences and microsatellites) suggested significant hybridization is occurring among *R. tanezumi* and the other two lineages (*R. rattus* I and *R. rattus* IV) in the U.S. and Asia, with unidirectional introgression from both *R. rattus* I and *R. rattus* IV into *R. tanezumi*. Furthermore, introgression is occurring to such a pronounced extent that we

were unable to detect any nuclear genetic signal for *R. tanezumi*. This is the first example of invasion leading to genetic swamping at a global scale in a mammalian species complex.

Introduction

Traditional dogma has held that hybridization is evolutionarily significant for plants and microbes, but less important in the evolution of animals. However, examples of hybridization and introgression among vertebrates are increasing, and introgression has been suggested as a major contributor to evolutionary innovation in animals (Seehausen 2004; Mallet 2007). Examples even exist of hybrid speciation in extant mammalian taxa (Larsen *et al*. 2010). and as comparative genome-scale analyses become more common for non-model organisms, additional examples will likely be discovered. In addition, complex patterns of hybridization among multiple species, where gene flow is indirectly occurring through an intermediate species, are illuminating novel mechanisms by which genetic material crosses species boundaries (McDonald *et al*. 2008; Nevado *et al*. 2011). This suggests that restricting analyses of hybridization and introgression to pairs of sister taxa—as is intuitive and most common—may underestimate the frequency of interspecific gene flow and its significance in evolution; expanding analyses to species groups will result in a more complete understanding of interspecific gene flow and its role in evolution.

In addition to acting as a source of evolutionary innovation, hybridization can also lead to genetic homogenization and even extinction (Rhymer and Simberloff 1996; Mooney and Cleland 2001; Wolf et al. 2001; Olden et al. 2004). As human alteration of natural habitats has become pervasive, biotic homogenization is an obvious issue from environmental and ecological perspectives, and the genetic consequences of homogenization can be significant, where invading organisms hybridize with native taxa (as well as other invaders) potentially resulting in the complete swamping out of the native taxa's genome. For vertebrates, most examples of this

phenomenon are known in fish, where external fertilization and weak reproductive isolating mechanisms place relatively reduced limits on reproductive compatibility (Hubbs 1955). For mammals, genetic swamping through hybridization is relatively rare, but documented examples exist (e.g., Sitka deer and red deer in Great Britain, Abernethy 1994; domestic/feral cats and wildcats, Hubbard et al. 1992; coyotes, feral dogs, and multiple canid species, Butler 1994; Hope 1994). However, these are typically geographically localized interactions between a single invader and an already rare native species, which facilitates hybridization and genetic swamping due to the decreased frequency of conspecific mating encounters for the rarer species. Much less common are examples of genetic swamping among species complexes at a global scale, but with the rapid increase in the frequency and geographic dispersion of invasive species due to a multitude of factors ranging from increasingly complex and voluminous human transportation mechanisms to large-scale habitat modification, examples may exist but remain to be discovered.

Rattus are arguably the most costly and destructive invasive species on earth. In terms of effects on native diversity, *Rattus* have had severe negative impacts in essentially every habitat they have invaded. Of the approximately 123 island groups worldwide, roughly 82% have been invaded by *R. norvegicus*, *R. rattus*, and/or *R. exulans* (Courchamp et al. 2003), and reports have estimated that introduced rats have been responsible for 40–60% of all bird and reptile extinctions since 1600 (Island Conservation 2006). In addition, *Rattus* are known to spread many zoonoses including bubonic plague, murine typhus, rat-bite fever, *Salmonella* food poisoning, leptospirosis, listeriosis, chagas, trichinosis, tularemia, and schistosomiasis (Gratz 1984). Over the last millennium, rat-borne pathogens are estimated to have killed more people than all wars and revolutions combined (Nowak 1999; Meerburg et al. 2009). Within the U.S., invasive rats are responsible for approximately 19 billion dollars in annual economic loss (nearly 15% of the total annual cost of all invasive species combined for the U.S.) through the transmission of disease,

structural damage to buildings, and contamination and destruction of food supplies (Pimentel *et al*. 2000).

Although multiple species of *Rattus* have been successful invaders, the black rats (*Rattus rattus* species group) have been particularly successful, spreading to all seven continents through its commensal relationship with humans (Nowak 1999; Musser and Carlton 2005). The taxonomic history of this species group is complex (Robins *et al*. 2007; Aplin *et al*. 2011). Initial chromosomal analyses suggested the existence of two species within the *Rattus rattus* species group, with a globally dispersed "ship rat" species (*R. rattus*) possessing a karyotype of 2n = 38–40, and a more geographically restricted Asian species (*R. tanezumi*) possessing a karyotype of 2n = 42 (Braverstock *et al*. 1983). Subsequent mitochondrial DNA (mtDNA) analyses indicated this two-species classification was an oversimplification, and that the *R. rattus* group potentially consisted of six distinct species with at least four of these taxa having a commensal relationship with humans and invading beyond their native range (Robins *et al*. 2007; Aplin *et al*. 2011). Moreover, the four commensal lineages exhibit statistically supported paraphyly in a mtDNA phylogenetic analysis, but are morphologically indistinguishable from one another (Aplin et al. 2011). As an additional layer of complexity, analysis of a single nuclear locus for individuals from Japan (Chinen *et al*. 2005) and the existence of intermediate karyotypes (Yosida 1980) confirmed that hybridization occurs between at least two of the mtDNA lineages, but it is unclear if hybridization is geographically widespread and/or is accompanied by genetic introgression.

We utilized mtDNA and nuclear DNA analyses to: 1) identify the black rat mtDNA lineages that have invaded the U.S., where invasive *Rattus* are geographically widespread and numerous; 2) examine the geographic and genomic extent as well as the directionality of hybridization among any detected members of the *R. rattus* species group within the U.S. and at international localities; and 3) determine the extent to which hybridization among mtDNA lineages is leading to genetic introgression. We identify multiple populations of both *R. tanezumi*

and *R. rattus* in the U.S., including a novel locality for *R. tanezumi* in Florida, and sample three species lineages globally. In addition, extensive hybridization is occurring between *R. tanezumi* and both *R. rattus* and a third unnamed species lineage at all localities. Finally, we suggest genetic introgression is unidirectional, with both *R. rattus* and the unnamed *Rattus* lineage swamping out the *R. tanezumi* nuclear genome at all localities where these species co-occur. This represents the first instance where hybridization among multiple members of a species complex as a result of human-aided dispersal has occurred at a global scale, potentially leading to widespread genomic extinction.

Materials and methods

Taxon sampling

We attempted to sample three of the four commensal putative species lineages identified in previous mtDNA studies by targeting localities where these lineages have been previously documented (Aplin et al. 2011). For *R. rattus* (the globally distributed "ship rat"; Lineage I in Aplin et al. 2011 and referred to as *R. rattus* I henceforth), all global localities where *Rattus* species have been collected represent potential members of this lineage. *R. rattus* I likely originated in India and spread across the globe primarily through human-aided dispersal. We focused on the U.S., where all specimens collected and obtained from museum collections are typically assumed to be a member of this lineage; however, any collected specimen identified morphologically as *R. rattus* is potentially a member of this lineage. *R. tanezumi* (Lineage II in Aplin et al. 2011 and referred to as *R. tanezumi* henceforth) is likely native to the SE Asia mainland and, through human-mediated dispersal, it (or at least its mtDNA genome) has spread to Japan, Malaysia, Indonesia, the Philippines, and New Guinea, as well as to South Africa and the San Francisco Bay area of California. Through collecting efforts and museum loans, we were able

to obtain tissues from the Philippines, Vietnam, and California populations. To target a third unnamed and widespread species lineage (Lineage IV in Aplin et al. 2011 and referred to as *R. rattus* IV henceforth), we obtained individuals from the Philippines, where this lineage has spread via human transport from its native range on the SE Asia mainland. A complete list of collected individuals, sampling localities, and source museum collections is given in Supplementary Table S4.1. Total DNA was extracted from either liver or muscle using the Qiagen DNeasy kit following the manufacturer's protocol.

MtDNA sequence generation and analyses

The complete mitochondrial cytochrome b gene (1,140 bp) was amplified using primers (*Rattus*CytbF 5'-TGACATGAAAAATCATCGTTGTAAT-3'; *Rattus*CytbR 5'-GGTTTACAAGACCAGAGTAAT-3') designed from an alignment of the complete mtDNA genome sequences of *R. rattus* (NC012374), *R. norvegicus* (AY769440), *R. tanezumi* (NC011638) and *Mus musculus* (NC006915) obtained from GenBank. PCR amplifications were carried out in 30μl reactions containing 200–500 ng of DNA, 0.12 μl of 5U/μl GoTaq Flexi DNA polymerase, 0.50 μl of each 10 μM primer, 2.4 μl Bovine Serum Albumin (0.01 g/ml), 2.4 μl of 25μM $MgCl_2$, 6.0 μl 5X Green GoTaq Flexi Buffer, 4.2 μl of a 10μM nucleotide mixture, and 8.88 μl of double distilled water (ddH_2O). The thermal profile consisted of an initial denaturation of 94°C for 4 minutes, followed by 35 cycles of 94°C, 42°C, and 72°C for 1 minute each. A final elongation at 72°C for 7 minutes ensured all reactions went to completion. Double-stranded products were purified using the Wizard SV Gel and PCR Clean-Up System (Promega Corporation, Madison, WI).

Both strands of the purified PCR products were sequenced using BigDye chain terminators following the manufacturer's protocol (Applied Biosystems) using the PCR amplification primers as well as two internal primers (*Rattus*CytbIntF 5'-

GGCTTCTCAGTAGACAAAGC-3'; *Rattus*CytbIntR 5'- TTTGATCCTGTTTCGTGGAGGAA-3') designed using the mtDNA genome alignment generated above. DNA sequencing reactions were electrophoresed on a 3130 Genetic Analyzer (Applied Biosystems). Contigs were assembled and edited using Geneious v 5.5 (Drummond et al. 2010).

In addition to the complete cytochrome b sequences generated above, we obtained all available *Rattus* cytochrome b sequences from GenBank (a complete list of sequences, sampling localities, and accession numbers is available in Tables S4.1 and S4.2), resulting in 631 cytochrome b sequences. Sequences were aligned using the Geneious v5.5 aligner (Drummond et al. 2010) and edited using MacClade v4.08 (Madison and Madison 2000). We included sequences from all available *Rattus* species to aid in species identification (and validate field identifications), because many species of *Rattus* are extremely difficult to differentiate based on morphology and a phylogenetic approach has been shown to be an effective method of species identification (Robins et al. 2007). The sequences obtained from GenBank ranged from 713 basepairs to the complete 1,140 basepairs. Initial phylogenetic analyses were conducted on both the full-length sequences only and the total dataset, and results were identical concerning clade assignment and support and consistent with previous mtDNA analyses (Robins et al. 2007; Aplin et al. 2011). Therefore, subsequent mtDNA analyses included all sequences and only these results are presented.

To generate a total phylogeny of all collected and downloaded *Rattus* cytochrome b sequences, we conducted a Bayesian phylogenetic analysis in MrBayes with the alignment partitioned by codon position and utilizing the HKY + I + Γ substitution model, which was indicated to be the best-fit model by AIC implemented in jModeltest (Posada 2008). Complete cytochrome b sequences for *Mus musculus musculus* (AY057804) and *M. musculus castaneus* (AY057805) were used as outgroups, and the analysis was run for 15,000,000 generations sampling every 1,000 generations, resulting in a posterior sample of 15,000 phylogenies. The Bayesian analysis was checked for sufficient mixing and stable convergence on a unimodal

posterior for all parameters using Tracer v1.5 (Drummond and Rambaut 2003). Burn-in was determined empirically by evaluating likelihood scores. The analysis reached stationarity at approximately 1,200 sampled phylogenies, so we conservatively discarded 2,000 phylogenies as burn-in. For each of the three *R. rattus* species group lineages recovered in the total mtDNA phylogeny (Fig. 4.1), we generated haplotype networks using TCS (Clement et al. 2000).

Nuclear sequence generation and analysis

For a subset of individuals, we sequenced two unlinked nuclear loci (Atp5a1 and DHFR) previously useful in resolving phylogenetic relationships within *Rattus* (Rowe et al. 2011). From localities where *R. tanezumi* [four localities; San Francisco Bay, CA (29 individuals), Panama City, FL (31), the Philippines (22), and Vietnam (8)] or the unnamed species lineage (one locality; the Philippines) were detected based on mtDNA sequence, we attempted to sequence both nuclear loci from all individuals, regardless of mtDNA species identity; however, we were unable to successfully amplify both loci from several individuals (see Table S4.1 for a list of sample sizes from each locality for each locus). In addition, we sequenced the two nuclear loci for individuals from several U.S. localities where only *R. rattus* was detected [Gainesville, FL (15 individuals) and Key Largo, FL (10 individuals)].

PCR amplification primers for both loci are given in Rowe et al. (2011). PCR recipes were identical for both loci, and carried out in 30μl reactions containing 200–500 ng of DNA, 0.12 μl of 5U/μl GoTaq Flexi DNA polymerase, 0.50 μl of each 10 μM primer, 2.4 μl Bovine Serum Albumin (0.01 g/ml), 2.4 μl of 25μM $MgCl_2$, 6.0 μl 5X Green GoTaq Flexi Buffer, 4.2 μl of a 10μM nucleotide mixture, and 8.88 μl of double distilled water (ddH_2O). The thermal profile was also the same for both nuclear loci, and consisted of an initial denaturation of 94°C for 5 minutes, followed by 35 cycles of 94°C for 30 seconds, 50°C for 1 minute, and 72°C for 1 minute 30 seconds. A final elongation at 72°C for 10 minutes ensured all reactions went to completion.

Double-stranded products were purified using the Wizard SV Gel and PCR Clean-Up System (Promega Corporation, Madison, WI). Both strands of the purified PCR products were sequenced using BigDye chain terminators following the manufacturer's protocol (Applied Biosystems, Inc., Foster City, CA) using only the PCR amplification primers, producing approximately 900 basepairs of high quality bi-directional sequence from the original approximately 1,100 basepair amplicons.

For each nuclear locus, we conducted a maximum likelihood (ML) phylogenetic analysis in PAUP* (Swofford 2003). The most appropriate model of nucleotide substitution for each locus was determined using AIC in jModeltest (Posada 2008), followed by a tree-search using 25 replicates of nearest neighbor interchange branch-swapping and random addition of taxa. For each locus, we used two *R. fuscipes* individuals as outgroups (GenBank accession numbers: HQ334330 and HQ334331 for ATP5A1, HQ334805 and HQ334811 for DHFR). This analysis was used to identify the most probable relationship among sequences, with the expectation that the mtDNA and nuclear gene trees would be largely congruent in the absence of gene flow. However, due to the relatively close relationship among the species examined here, it is possible that incongruence is a result of incomplete lineage sorting (Maddison 1997). Because lineage sorting is a stochastic process, we would expect similar estimates of gene flow for pairwise comparisons of sympatric and allopatric populations of different species (e.g., Grant et al. 2005; Nevado et al. 2011). In contrast, if gene flow has occurred among previously isolated lineages, we would expect estimates of gene flow among species to correlate geographically, with the highest amount of gene flow among sympatric populations. To test this hypothesis for the sampled mtDNA lineages, we used the Bayesian coalescent approach implemented in Migrate-n v3.2.16 (Beerli and Felsenstein 1999; Beerli 2006), which generates pairwise estimations of bi-directional gene flow. One assumption of Migrate-n is the absence of intralocus recombination. We tested for recombination events within the alignment for each locus using the GARD test (Kosakovsky Pond et al. 2006) implemented in HyPhy (Kosakovsky Pond et al. 2005), and no

recombination breakpoints were detected for either locus. Preliminary runs in Migrate-n (100,000 burn-in steps followed by 100,000 steps sampled every 100 steps) were conducted for each pairwise comparison to establish upper and lower bounds for parameter prior distributions. In the full-length analyses, each pairwise estimation of gene flow among populations was estimated three times independently by conducting 200,000 burn-in steps followed by 1,000,000 steps sampled every 100 steps. Convergence for migration estimates was checked by examining the posterior distribution for each parameter using Tracer v1.5, as well as by comparing posterior estimates between duplicate runs.

Microsatellite data generation and analyses

Because the nuclear loci we sequenced are gene regions, it is possible that their patterns of introgression may be affected by natural selection and do not accurately reflect patterns of hybridization. To provide neutral (or approximately neutral) estimates of population and interspecific divergence and gene flow, we genotyped all individuals collected in the United States, the Philippines, and Vietnam (253 individuals from 16 populations; see Table S4.1 for the detailed list of individuals and localities) at nine microsatellite loci (loci names and primers given in Table S4.3). Microsatellites were amplified by PCR in 15 μl reactions containing 9μl of True Allele PCR Premix (Applied Biosystems, Inc., Foster City, CA), 4 μl ddH_2O, 0.5 μl of each primer (10 μM), and 1 μl template DNA with the following conditions: an initial denaturation of 95°C for 12 minutes, 35 cycles of 94°C for 40 seconds, 57°C for 40 seconds, and 72°C for 30 seconds; and a final elongation of 72°C for 4 minutes. Then 0.5 μl of product was added to 9.5 μl of loading mix containing a size standard (ROX 400HD; Applied Biosystems, Inc., Foster City, CA) and this mixture was analyzed using an ABI 3130 Genetic Analyzer and GeneMapper 3.7 to visualize microsatellite alleles and determine genotypes. All genotypes were scored twice, and anonymously randomized for the second scoring to ensure no bias was present in the final dataset.

Genepop v4.0 (Raymond and Rousset 1995) was used to test for deviations from Hardy-Weinberg equilibrium (HWE) by conducting global heterozygosity excess and deficit tests for each locality and species. Each test was run for 10,000 dememorization steps followed by 100 batches of 5,000 steps each. Additionally, we used Genepop to conduct a composite linkage disequilibrium test (Weir 1996). Significance was assessed with the same MCMC settings as was used for the heterozygosity tests. We used Structure v2.3 (Pritchard et al. 2000) to test the groups recovered in the mtDNA phylogenetic analysis (Fig. 4.1) and identify the optimal number of clusters in our data. In the absence of introgression, we would expect microsatellite clusters generated at $K = 3$ to match the three clades recovered in the mtDNA phylogenetic analysis. For each value of K (we ran the analysis at K from 1–16), we ran five independent analyses of 500,000 generations following 250,000 generations of burn-in under the admixture model and with the assumption that allele frequencies among populations are correlated. Convergence was checked by plotting likelihoods throughout the run and comparing likelihood values and population assignments between duplicate runs. We calculated the optimal number of clusters for our data using the ΔK statistic (Evanno et al. 2005).

In addition to clustering analyses, we attempted to use Migrate-n to calculate pairwise migration rates among sympatric and allopatric populations of our three species, as was conducted for the nuclear sequence data above. This analysis employed a Brownian motion approximation of the ladder (stepwise) mutation model. Preliminary runs (100,000 burn-in steps followed by 100,000 steps sampled every 100 steps) were conducted for each pairwise comparison to establish upper and lower bounds for parameter prior distributions. In the full-length analyses, each pairwise estimation of gene flow among populations was estimated twice independently by conducting 500,000 burn-in steps followed by 5,000,000 steps sampled every 100 steps. Convergence for migration estimates was checked by examining the posterior distribution for each parameter using Tracer v1.5, as well as by comparing posterior estimates between duplicate runs.

Results

MtDNA analyses

Cytochrome b sequences were generated for 272 putative black rats. The Bayesian phylogenetic analysis of cytochrome b placed all of the sampled *R. rattus* in three distinct and statistically supported clades (Fig. 4.1) corresponding to lineages I, II, and IV from Aplin et al. (2011). With the exception of a subset of individuals from the San Francisco Bay, CA and Panama City, FL, all individuals collected from the U.S. belonged to the *R. rattus* I clade (highlighted in blue in Fig. 4.1). The Panama City, FL population represents only the second documented population of *R. tanezumi* in the U.S. From this population, 27 of 31 individuals collected clustered into the *R. tanezumi* clade (highlighted in green in Fig. 4.1), and all of these individuals had the same mtDNA haplotype. The phylogenetic analysis placed the remaining four individuals from that locality in the *R. rattus* I clade. The individuals collected from the Philippines fell into two clades in the mtDNA phylogenetic analysis, with the majority (42 individuals) found in the *R. rattus* IV clade (highlighted in red in Fig. 4.1), and the six remaining individuals nested within the *R. tanezumi* clade. Finally, all tissues obtained from Vietnam clustered into the *R. tanezumi* clade.

For the network analyses, the *R. tanezumi* clade consisted of 38 unique haplotypes. The three unique California *R. tanezumi* haplotypes were either shared with eastern Asian localities or differed by only a single mutation from eastern Asian haplotypes, and were far removed from both the South African and Florida *R. tanezumi* haplotypes (Fig. 4.2). The single *R. tanezumi* haplotype recovered from Panama City, FL is relatively isolated in the haplotype network, differing by five mutations from a Vietnamese haplotype and an Indonesian haplotype, and was nine mutations removed from the most common South African haplotype. The four *R. tanezumi* haplotypes recovered from the Philippines were not shared outside of the Philippines and formed a closely related terminal cluster that also contained two of the three California *R. tanezumi* haplotypes.

Relative to the *R. tanezumi* and *R. rattus* IV haplotype networks, the 55 haplotypes recovered for *R. rattus* I formed multiple, very closely related clusters with some geographic structure (Fig. 4.2). With the exception of two haplotypes recovered from Key Largo, FL, all U.S. haplotypes (as well as Caribbean, Mexican, Central American, and South American haplotypes) formed a tight cluster around two globally distributed haplotypes. This cluster of haplotypes was largely isolated to the western hemisphere. The two divergent haplotypes detected in Key Largo, FL were nested within a cluster of African and Asian haplotypes. In addition, three haplotypes detected in southern Africa were nested within the western hemisphere cluster of haplotypes.

The *R. rattus* IV haplotype network consisted of 29 unique haplotypes isolated to eastern Asia (Fig. 4.2). Relative to the *R. tanezumi* and *R. rattus* I haplotype networks, the *R. rattus* IV network exhibited considerable divergence among haplotypes, but without any clear geographic associations. For the Philippines (and *R. rattus* IV in general), haplotypes were highly divergent from one another, and all but one of these haplotypes was isolated to the Philippines. The single shared haplotype was recovered from Malaysia, Indonesia, and Vietnam in addition to the Philippines.

Nuclear sequence analyses

For both *Atp5a1* and *DHFR*, the maximum likelihood gene tree differed significantly from the mtDNA topology (Fig. 4.1B). Identical sequences for each locus were recovered in individuals from multiple mtDNA clades (Appendix), suggesting hybridization was occurring. The three supported monophyletic mtDNA clades were not evident for either nuclear locus, with a complete lack of distinction between *R. tanezumi* and *R. rattus* I. *R. rattus* IV conflicted less with the mtDNA topology, with sequences falling out relatively basal, but not in a monophyletic clade. In addition, there appeared to be a clear association between topology and locality, in that the majority of U.S. individuals clustered together regardless of mtDNA clade. The same was true for the Asian *Rattus*, where samples from the Philippines and Vietnam were relatively more

closely related to each other than to members of the same mtDNA clade from different geographic localities. This overall pattern is indicative of gene flow among sympatric (or geographically proximate) populations of the different mtDNA lineages.

This pattern was confirmed by the gene flow analyses using Migrate-n. For both loci, even when population sample sizes were relatively small (i.e., six *R. tanezumi* individuals from the Philippines), the average mode of the migration rate across the duplicate run was always highest between the sympatric populations of different species (Tables 4.1 and 4.2). Moreover, directional estimates indicated gene flow is largely unidirectional. For the two sympatric populations of *R. tanezumi* and *R. rattus* I (San Francisco Bay, CA and Panama City, FL), estimates of gene flow for both loci were higher in the direction of *R. tanezumi*. Similarly, for the single sympatric comparison of *R. tanezumi* and *R. rattus* IV (Philippines), the majority of gene flow was from *R. rattus* IV into *R. tanezumi* for both loci. These analyses suggest significant genetic introgression from both *R. rattus* I and *R. rattus* IV into *R. tanezumi*, but a relative lack of introgression from *R. tanezumi* into the other two species. However, sample sizes are generally too small for several of these comparisons to say definitively, and additional samples are required to confirm this unidirectional pattern.

Microsatellite analyses

We scored nine microsatellite loci for 253 individuals from 16 localities (see Table S4.1). None of the pairwise linkage disequilibrium tests were significant (not shown). The global tests of HWE revealed multiple significant deviations (Table 4.3). All instances of HWE deviation occurred in populations where two mtDNA species are sympatric and potentially hybridizing. For *R. tanezumi*, HWE was rejected due to heterozygote deficit in the Panama City, FL and San Francisco Bay, CA populations. For *R. rattus* I, HWE was rejected due to heterozygote excess in the San Francisco Bay, CA and Panama City, FL populations. For *R. rattus* IV, HWE was not rejected in the only sampled population.

The clustering analyses conducted in Structure suggested significant gene flow among sympatric populations of each species and a general lack of agreement with the mtDNA analyses (Fig. 4.3). At $K = 2$, *R. rattus* I and *R.* rattus IV were distinct. The *R. tanezumi* clustered according to geographic association, with the San Francisco Bay, CA and Panama City, FL populations grouping with the North American *R. rattus* I individuals, whereas the two Asian populations of *R. tanezumi* (Vietnam and the Philippines) grouped with the Asian population of *R. rattus* IV (Philippines). Figure 4.3 shows individual assignments through $K = 6$, and at no point do the Asian *R. tanezumi* and *R. rattus* IV populations become distinct from one another. At $K = 3$, there was only partial congruence with the mtDNA clades recovered in the phylogenetic analysis. The *R. rattus* IV individuals were distinct from *R. rattus* I, but included the Asian *R. tanezumi*, while the *R. tanezumi* from the two U.S. populations remained clustered with *R. rattus* I, with the exception of the Alaska population of *R. rattus* I which was distinct. This pattern was maintained at $K = 4$, which the ΔK statistic indicated to be the optimal number of groupings, except with the Gainesville, FL population of *R. rattus* I now distinct. At $K = 5$, the Panama City, FL *R. tanezumi* and *R. rattus* I fell into a largely distinct group, while the San Francisco Bay, CA *R. tanezumi* remained clustered with the remaining *R. rattus* I until $K = 9$ (not shown).

For the microsatellite gene flow analysis conducted using Migrate-n, we were unable to get stable convergence on posterior estimates for migrations rates. Posterior distributions for the vast majority of comparisons were seldom unimodal, and estimates fluctuated wildly between duplicate runs. This issue was not resolved by increasing runs significantly (we attempted a maximum of 10,000,000 steps sampled at a 100-step density) or by using a simpler mutation model (i.e., the infinite allele model). Convergence problems with microsatellite analyses have been reported previously with Migrate-n (see software documentation), and were attributed to the complexity of the analysis attempting to integrate over multiple loci. Due to the convergence issues, we do not report Migrate-n results for the microsatellite data.

Discussion

Species invasions are dynamic events in which individuals—typically in low numbers—enter a novel environment through human facilitation (i.e., disturbance, human dispersal, etc.). In this way, species invasions are analogous to natural founding events with the exception of the impetus for entering a new locality (Sax et al. 2007). While the ecology of these events is of obvious and critical importance, species invasions have clear evolutionary implications for both the invader and the native flora and fauna. In addition, when multiple invading species are simultaneously colonizing a single location, they will inevitably interact, adding an additional layer of complexity. Furthermore, novel invasions detected in their infancy represent "experiments" in evolution, where we can seek to understand the genetic and evolutionary consequences of founder events (both natural and unnatural) in various ecological contexts. Early detection of novel invasions, particularly those replicated in multiple localities for a given species, present the opportunity to examine the factors that are important in dictating the ability of a given species to establish and disperse. Given the ongoing invasion of multiple *Rattus* species at a global scale, they are under-utilized as models for studying the evolutionary implications of species invasions. Furthermore, the *R. rattus* species complex is of particular utility in studies of evolution and reproductive isolation due to the relatively recent diversification of the species members and their documented ability to hybridize (Yosida 1980; Chinen et al. 2005). In the U.S., we have now identified two invading black rat "species" exhibiting drastically different degrees of success in establishing and invading, and experiencing significant reproductive interaction.

Aplin et al. (2011) conducted the most thorough genetic analysis of *R. rattus* to date, utilizing mitochondrial cytochrome b sequences to examine global patterns of genetic diversity. Their analyses indicated considerable diversity that at least partially corresponded to the chromosomal races of Yosida (1980), and suggests the existence of at least six morphologically indistinguishable but genetically distinct taxa. In our analyses, we focused on the three mtDNA lineages for which widespread invasion has been documented (termed *R. rattus* I, *R. tanezumi*,

and *R. rattus* IV), with *R. rattus* I likely evolving in India and now globally dispersed, *R. tanezumi* evolving in mainland East Asia and spreading throughout the South Pacific, South Africa and localities on both the east and west coasts of the U.S., and *R. rattus* IV evolving in mainland Southeast Asia on the Indochina Peninsula and spreading throughout the South Pacific (Aplin et al. 2011). Phylogenetic analysis of mtDNA (presented both here and by Aplin et al. 2011) and chromosomal data (Yosida 1980) indicate that, prior to invasion, these three lineages were largely allopatric and evolving in isolation. Dating analyses indicated the mtDNA genomes of these lineages initially diverged about 1 million years ago (mya), followed by the *R. rattus* I/*R. tanezumi* divergence approximately 600 years ago (kya) (Robins et al. 2008; Aplin et al. 2011). These levels of isolation are certainly consistent with species-level divergences in other mammalian taxa (Bradley and Baker 2001), with maximum-likelihood corrected pairwise cytochrome b sequence divergences of 5.1% (*R. tanezumi* vs *R. rattus* I), 7.8% (*R. rattus* I vs *R. rattus* IV), and 9.2% (*R. tanezumi* vs *R. rattus* IV). In spite of this considerable divergence, our results suggest widespread invasion, which has led to sympatry for various combinations of these taxa in the entirety of the *R. tanezumi* and *R. rattus* IV ranges, and has allowed for significant hybridization with introgression on multiple continents.

Black rat mtDNA population genetics

Our analysis of the cytochrome b gene identified *R. rattus* I as being essentially globally distributed, as has been shown in previous studies (Aplin et al. 2011). Within the U.S., the vast majority of individuals examined fell into this group, with the exception of a subset of samples collected from the San Francisco Bay, CA and Panama City, FL, which belonged to the *R. tanezumi* lineage (Fig. 4.2). Whereas the California population had been identified by Aplin et al. (2011), the Panama City, FL population of *R. tanezumi* was unexpected, primarily because *R. tanezumi* was expected to be relatively isolated to East Asia (Braverstock et al. 1983), and only recently expanded to the single California population (Aplin et al. 2011). However, a recent

analysis of *Rattus* mtDNA from South Africa (Bastos et al. 2011) detected both *R. tanezumi* and *R. rattus* I in sympatry in northeastern South Africa and Swaziland, but was unable to identify the source of the animals due to sampling restrictions. We included the two haplotypes Bastos et al. (2011) recovered from South Africa in our more extensive geographic sample and found them to be most closely related to three haplotypes from Indonesia. Although our global sampling is far from adequate to identify their origin with certainty, the *R. tanezumi* haplotype network suggests an Asian origin. Nevertheless, it is quite possible that with further collection of individuals (much of the African continent and western Asia remain unsampled), novel source populations will be identified.

Within the U.S., the two populations where *R. tanezumi* has been detected are quite different from one another in terms of mtDNA diversity. For the three California *R. tanezumi* haplotypes, our analyses suggest quite convincingly that they originated from Southeast Asia. Two of the California *R. tanezumi* haplotypes also were recovered from multiple Asian localities, while the third was nested within a cluster of Asian haplotypes. Moreover, the California *R. tanezumi* appear to be the result of introductions from multiple source localities, with these haplotypes shared between California, Japan, the Philippines, Indonesia, and several localities on the Indochina peninsula. In contrast, the single *R. tanezumi* haplotype recovered from Panama City, FL was not detected in another locality, making it difficult to speculate as to the source for this population. In addition, the detection of only one haplotype for this entire population despite significant collection effort (31 individuals sampled) suggests this population may have been founded by a single introduction of one maternal lineage, with no further input from additional propagules. This Florida *R. tanezumi* haplotype appears somewhat isolated in the haplotype network, with the most closely related haplotype detected on the Indochina peninsula, but five mutational steps removed. Regardless of the specific source population, the Florida *R. tanezumi* appear distinct from the California *R. tanezumi*, suggesting different sources founded these two U.S. populations. Furthermore, the complete absence of *R. tanezumi* haplotypes at intervening

sampled populations within the U.S. suggests dispersers from one of the U.S. *R. tanezumi* populations did not found the other.

The occurrence of *R. tanezumi* throughout eastern Asia is relatively well documented (Braverstock et al. 1983; Chinen et al. 2005; Rickart et al. 2011), but the discovery of novel populations in the U.S. (in this study and Aplin et al. 2011) and South Africa (Bastos et al. 2011) suggests this is potentially a widespread but unsampled mtDNA lineage. In addition, the relative isolation of the South African and Florida *R. tanezumi* haplotypes on the haplotype network further suggests that these are either long-isolated populations that have diverged from their Asian sources, or that *R. tanezumi* actually represents a geographically widespread and genetically diverse lineage that we have simply failed to sample. We believe the latter to be the more likely explanation, and we expect additional *R. tanezumi* individuals to be detected with further sampling in western Asia and Africa.

For *R. rattus* I, we detected two geographically widespread haplotypes, which were both found spread across North America, as well as in Africa, South America, and Asia (Fig. 4.2). Aplin et al. (2011) conducted an extensive mtDNA phylogeographical analysis of black rats at the global scale, so we will focus our discussion to the U.S. The U.S. *R. rattus* I haplotype network consisted of many very closely related haplotypes clustered around the two most common haplotypes—a common pattern for invading organisms that undergo a rapid expansion following invasion (Schaal et al. 2003). These two haplotypes likely reflect the founding mtDNA lineages for the initial black rat invasion into the U.S., hypothesized to have occurred with Columbus in 1492 (Armitage 1993), followed by a rapid expansion to cover the North American continent. With only two exceptions, no U.S. *R. rattus* I haplotype was more than two mutations removed from one of the two widespread haplotypes, suggesting the vast majority of *R. rattus* I mtDNA diversity arose *in situ* following the initial establishment of this lineage in the U.S., or that any additional individuals that entered the U.S. came from essentially the same locality (or at least the same pool of mtDNA diversity) where a recent expansion had occurred. The two exceptions to

this pattern were haplotypes recovered from Key Largo, FL that were nested within the haplotypes primarily recovered from Africa and the Middle East. This suggests the *R. rattus* I in Key Largo likely arose from a different source—possibly Africa—relative to all other *R. rattus* I sampled in the U.S. Moreover, the isolation of these divergent haplotypes in the Florida Keys suggests they are not dispersing following introduction.

For the *R. rattus* IV haplotype network, we detected many divergent lineages, all isolated to East Asia (Fig. 4.2). For the samples collected from the Philippines, we detected multiple divergent haplotypes, suggesting this is a stable and genetically diverse population that has spread from its ancestral region on the Indochina peninsula (Aplin et al. 2011) throughout Indonesia, Malaysia, and the Philippines but has not spread further south to Australia or New Guinea where *R. rattus* I was detected, nor has it crossed the Pacific into the U.S. as *R. tanezumi* has, despite being found in sympatry throughout Asia.

Evidence for hybridization with introgression

That hybridization between *R. rattus* I and *R. tanezumi* is possible was established prior to our analysis, in both laboratory breeding experiments and genetic analysis of natural populations (Yosida 1980; Chinen et al. 2005). Chromosomal analysis of wild populations and laboratory breeding led Yosida (1980) to suggest introgression is rare among these species due to reduced F_1 fitness, noting that parental karyotypes (*R. rattus* I, 2n = 38; *R. tanezumi*, 2n = 42) were far more common than hybrid karyotypes (2n = 40) at localities where the two lineages were sympatric. Chinen et al. (2005) identified two localities in Japan where both hybrid and parental karyotypes existed, and these karyotypes roughly corresponded to mtDNA haplotypes. But, neither mtDNA nor karyotypic data corresponded with the nuclear phylogeny, suggesting introgression was common at these localities. Hybridization of *R. rattus* IV with *R. rattus* I and *R. tanezumi* is less well understood, because the *R. rattus* IV lineage appears to consist of individuals with 2n = 40 and 2n = 42 (Aplin et al. 2011), and has been referred to in the literature

by a variety of names, creating significant confusion. However, Yosida (1980) reported chromosomal intermediates (2n = 39) between the 2n = 40 individuals (*R. rattus* IV) and 2n = 38 (*R. rattus* I) on Sri Lanka, as well as in laboratory matings. In addition, matings were successful between Japanese 2n = 42 (*R. tanezumi*) and 2n = 42 black rats from Malaysia that Yosida (1980) referred to as *R. rattus diardi*, which has previously been a synonym for *R. rattus* IV (Robins et al. 2007). These "*R. rattus diardi*" were distinct from other 2n = 42 black rats (i.e., *R. tanezumi*) in being homozygous for a subtelocentric chromosome 9, and hybrids were heterozygous for this chromosomal arrangement. Therefore, karyotypic analyses suggest these lineages are at least capable of hybridizing. However, ours is the first analysis examining the extent of hybridization and introgression among these three taxa, and is the first to do so in both Asian and U.S. populations.

Our analysis of protein coding nuclear sequence data and nuclear microsatellites provide conclusive evidence of widespread hybridization with introgression between *R. rattus* I and *R. tanezumi*, as well as between *R. rattus* IV and *R. tanezumi*. As mentioned above, conflict between mtDNA and nuclear loci is expected given the 4-fold difference in effective population size and therefore 4-fold reduction in the rate with which nuclear polymorphisms sort relative to mtDNA polymorphisms (Neigel and Avise 1986; Maddison 1997). Given the potential for multiple mechanisms to generate conflict among mtDNA and nuclear phylogenies, it can be difficult to tease apart the effects of hybridization vs lineage sorting. Because lineage sorting proceeds randomly, and lineage reticulation due to hybridization is a result of mating in sympatry, we can, however, predict that incomplete lineage sorting will result in the random retention of ancestral polymorphisms that will differ starkly between unlinked nuclear loci. In contrast, hybridization will result in an association between geographic location and genetic relationship, as only sympatric interspecific populations can interbreed (with the possible exception of organisms exhibiting external or broadcast fertilization, such as wind-pollinated plants and broadcast-spawning fish). Therefore, we can predict gene flow among sympatric interspecific lineages will

be higher than among allopatric interspecific lineages. The direct contrast between the mtDNA and nuclear data (Figs. 4.1 and 4.3), in conjunction with the clear relationship between geographic location and measures of gene flow (Tables 4.1 and 4.2)—independent of mtDNA lineage—indicates the commensal relationship between these species and humans has had significant evolutionary implications, facilitating hybridization between *R. tanezumi* and *R. rattus* I in the U.S., and between *R. tanezumi* and *R. rattus* IV in Asia. For the migration analyses, migration rates for sympatric populations of different mtDNA lineages were always the highest pairwise comparisons, and this pattern was replicated across the two unlinked nuclear loci, suggesting the pattern is robust. In addition, Structure analysis of nine unlinked microsatellite loci similarly indicated the partitioning of nuclear genetic diversity was a result of geography, and did not correspond with mtDNA lineages. The two Asian *R. tanezumi* localities cluster with the Asian *R. rattus* IV samples and not with other members of the *R. tanezumi* mtDNA lineage. In addition, the two U.S. populations of *R. tanezumi* cluster with other U.S. localities, not with other members of the *R. tanezumi* mtDNA lineage. Moreover, these populations do not become distinct from *R. rattus* I until K=5, after the Alaska and Gainesville, FL populations have already been relegated to distinct clusters. This suggests that multiple populations of *R. rattus* I are actually more genetically unique than the admixed populations in San Francisco Bay, CA and Panama City, FL, attesting to the extent of genetic homogenization that has occurred.

The lack of morphological divergence in these species renders it difficult to comment with certainty on the magnitude and direction of introgression, because we cannot tell what, if any, individuals constitute "true" unhybridized members of each species. In addition, because *R. rattus* I is globally distributed (Nowak 1999; Musser and Carlton 2005; Aplin et al. 2011) we cannot say with certainty that any individual from one of the other two mtDNA lineages is from a population not sympatric with *R. rattus* I, or is not a migrant from a population sympatric with *R. rattus* I. As a result, we could make no *a priori* assumption about the status of any sampled individual. But, even in the absence of morphological diagnostics, we can rely on the non-

recombining mtDNA locus to at least give us an idea of the maternal history of each individual, and in the absence of admixture we can hypothesize the existence of three nuclear genetic signals. If admixture is significant and bidirectional, we would expect individuals from each mtDNA lineage to be a mixture of the two parent nuclear genomes. In contrast, if introgression is highly unidirectional, we would expect the mtDNA and nuclear genomes of the "donating" species to be consistent with one another, and the mtDNA and nuclear DNA of the "receiving" species to conflict significantly. This unidirectional introgression pattern is precisely what we see for these three species. At a K=2, the microsatellite cluster analysis (Fig. 4.3) consistently clustered all individuals with *R. rattus* I mtDNA together, and similarly clustered all individuals with *R. rattus* IV mtDNA together, whereas the *R. tanezumi* nuclear genome consistently took on the identity of the species with which the sampled population was sympatric or geographically most proximate. In addition, for sympatric mtDNA lineages in both Asia and the U.S., migration rates estimated from both nuclear loci were always higher in the direction of *R. tanezumi* (Tables 4.1 and 4.2). Although the small sample size for some of these comparisons (i.e., four *R. rattus* I individuals from Panama City, FL) may limit the power of these analyses, the consistency of this pattern across two unlinked nuclear genes and neutrally evolving microsatellite loci suggests these results are robust.

Whereas the evidence for significant unidirectional introgression in both Asia and the U.S. appears unequivocal, it is less clear what factor(s) may be driving this pattern. F_2 offspring have been produced from the hybrids of the *R. tanezumi*/*R. rattus* I cross, but with considerable difficulty, suggesting semisterility in the F_1 generation (Yosida et al. 1971; Yosida 1977). For *R. tanezumi*/*R. rattus* IV hybrids, no F_2 offspring have been produced despite considerable effort (Yosida 1977). For both of these hybrids, backcrosses to both parental species (regardless of sex) were successful and likely represent the mechanism by which hybrid lineages persist in the wild. The reason for unidirectional gene flow in the U.S. is potentially an issue of mate availability. Both populations of *R. tanezumi* (California and Florida) appear to be small, isolated populations

that were relatively recently established. This is supported by the fact that *R. tanezumi* mtDNA haplotypes were not detected outside of the San Francisco Bay area for the California population or outside of Panama City for the Florida populations, despite our sampling multiple localities in both states (Fig. 4.2). In contrast, coastal populations of *R. rattus* I are typically enormous (Sullivan 2004). If this is the case, then it is likely that hybridization is largely driven by the availability of mates (Rhymer and Simberloff 1996), where the much less common *R. tanezumi* simply encounter *R. rattus* I individuals more often, resulting in many more *R. tanezumi*/*R. rattus* I matings than conspecific matings for *R. tanezumi*, followed by mostly F_1/*R. rattus* I backcrosses. This would perpetuate the eventual swamping out of the *R. tanezumi* nuclear genome as we have observed, but maintain the *R. tanezumi* non-recombining mtDNA genome as long as there were no fitness costs (Wolf et al. 2001; Buerkle et al. 2003). However, the assertion that the California and Florida *R. tanezumi* populations are small, recent introductions relative to *R. rattus* populations is speculative and this certainly warrants further study and much more thorough geographic sampling.

In Asia, it is difficult to say whether or not the same mechanism may be driving unidirectional gene flow from *R. rattus* IV into *R. tanezumi*. This is because the relative abundance of each of these species is unclear due to their being morphologically indistinguishable. In the Philippines, we were able to obtain tissue samples from 48 individuals from several locations, but only detected six *R. tanezumi* individuals, suggesting *R. tanezumi* individuals are relatively rare in the Philippines. For Vietnam, we only obtained tissues for 17 individuals (all *R. tanezumi*) from a single coastal locality, and are therefore unable to comment on the relative frequency of each of these mtDNA lineages; however, we should note that *R. rattus* IV haplotypes downloaded from GenBank were detected in animals collected in Vietnam, Cambodia, Laos, and Thailand (Fig. 4.2, Table S4.2), suggesting *R. rattus* IV are common on the Indochina Peninsula.

Invasion biology and evolutionary implications

This system presents considerable opportunity to study the role of interspecific interactions in dictating the outcome of the invasion process, as well allowing for an evolutionary analysis of reproductive isolation. The role of evolutionary processes in biological invasions has been traditionally neglected in favor of ecology, but evolutionary data has begun to move to the forefront in the recent invasion biology literature (Prentis et al. 2008). Hybridization has been suggested as a mechanism capable of stimulating invasiveness, but essentially all of the evidence for this is found in plants (Ellstrand and Schierenbeck 2000). The three species examined in this study represent morphologically and ecologically similar "replicates" with distinct invasive histories. *Rattus rattus* I has dispersed across the globe and has an existence essentially inseparable from that of humans. *Rattus tanezumi* originated on the Southeast Asian mainland, invaded much of East Asia and the South Pacific, and has now begun to invade localities outside of Asia. These novel localities represent invasions in their infancy, and we show that in both Asia and the recent invasions in the U.S., hybridization with introgression is common, to the extent that we were unable to distinguish the nuclear genome of *R. tanezumi* from its sympatric heterospecifics. This potentially is an instance where hybridization has augmented the ability of *R. tanezumi* to disperse and establish in novel localities, and is an opportunity to investigate the role of hybridization in the spread of a mammalian invader.

In contrast to the other two species investigated here, *R. rattus* IV spread from its origin on the Indochina peninsula to occupy Malaysia, Indonesia, and the Philippines, but has not spread beyond these limits (Aplin et al. 2011). In Asia, we detected significant hybridization with introgression between this species and *R. tanezumi*, but detect no individuals outside of Asia exhibiting the nuclear genome signature of *R. rattus* IV. This suggests that hybridization between these two lineages is not sufficient in itself to lead to further successful invasions, and *R. rattus* IV has not been able to expand out of Southeast Asia. Alternatively, *R. rattus* IV has successfully spread out of Asia, but we have simply failed to detect it in our current sampling. Only additional

work will distinguish between these two scenarios, but our extensive sampling suggests the former to be more likely, and if invasion outside of Asia has occurred it is being limited to relatively few populations and/or low numbers of individuals. Hybridization is common among invasives, and it has been suggested that it elevates the rate of response to novel selective pressures (Seehausen 2004). If hybridization is an important factor in facilitating the invasion of *R. tanezumi* into the U.S. and South Africa, the genetic contributions from *R. rattus* I, arguably the world's most successful vertebrate invader, may be a critical contributing factor.

References

Abernethy K (1994) The establishment of a hybrid zone between red and sitka deer (genus *Cervus*). *Molecular Ecology*, **3**, 551-562.

Aplin KP, Suzuki H, Chinen AA *et al.* (2011) Multiple genetic origins of commensalism and complex dispersal history of black rats. *PLoS One*, **6**, e26357.

Armitage PL (1993) Commensal rats in the New World, 1492–1992. Biologist, **40**, 174–178.

Bastos AD, Nair D, Taylor PJ *et al.* (2011) Genetic monitoring detects an overlooked cryptic species and reveals the diversity and distribution of three invasive *Rattus* congeners in South Africa. *BMC Genetics*, **12**, 26.

Beerli P (2006) Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics*, **22**, 341–345.

Beerli P, Felsenstein J (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics*, **152**, 763–773.

Bradley RD, Baker RJ (2001) A test of the genetic species concept: cytochrome-*b* sequences and mammals. Journal of Mammalogy, **82**, 960–973.

Braverstock PR, Adams M, Maxson LR, Yosida TH (1983) Genetic differentiation among karyotypic forms of the black rat, *Rattus rattus*. *Genetics*, **105**, 969–983.

Buerkle CA, Wolf DE, Riesberg LH (2003) The origin and extinction of species through hybridization. In: *Population viability in plants: conservation, management, and modeling of rare plants* (eds Brigham CA, Schwartz MW). pp. 117–144, Springer, New York, New York.

Butler D (1994) Bid to protect wolves from genetic pollution. *Nature*, **360**, 497.

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 & Genetic Systems*, **80**, 367–375.

Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1660.

Courchamp F, Chapuis JL, Pascal M (2003) Mammal invaders on islands: impacts, control, and control impact. *Biological Reviews*, **78**, 347–383.

Drummond AJ, Ashton B, Cheung M *et al.* (2010) Geneious, version 5.0. http://www.geneious.com.

Drummond AJ, Rambaut A (2003) Tracer v1.5. http://tree.bio.ed.ac.uk/software/tracer/.

Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 7043–7050.

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

Giraudeau F, Apiou F, Amarger V, Kaisaki PJ, Biboreau M, Lathrop M, Vergnaud G, Gauguier D (1999) Linkage and physical mapping of rat microsatellites derived from minisatellite loci. *Mammalian Genome*, **10**, 405–409.

Grant PR, Grant BR, Petren K (2005) Hybridization in the recent past. *The American Naturalist*, **166**, 56–67.

Gratz NG (1984) The global public health importance of rodents. In: *Proceedings of a conference on the organization and practice of vertebrate pest control* (ed. Dubock, AC). Pp. 413–435, World Health Organization, Geneva, Switzerland.

Hope J (1994) A wolf in pet's clothing. *Smithsonian*, **25**, 34–45.

Hubbard AL, McOrist S, Jones TW, Boid R, Scott R, Easterbee N (1992) Is survival of European wildcats *Felis silvestris* in Britain threatened by interbreeding with domestic cats? *Biological Conservation*, **61**, 203–208.

Hubbs CL (1955) Hybridization between fish species in nature. *Systematic Zoology*, **4**, 1–20.

Island Conservation (2006) Evaluation of primary and secondary exposure risks to land bird species for: experimental use permit application for field efficacy trial of 0.0025% brodifacoum broadcast bait to eradicate introduced rats from Aleutian Islands in the Alaska Maritime National Wildlife Refuge, July 20, 2006. *Island Conservation, Center for Ocean Health*, Santa Cruz, California.

Kosakovsky Pond SL, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. *Bioinformatics*, **21**, 676–679.

Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SDW (2006) GARD: a genetic algorithm for recombination detection. *Bioinformatics*, **22**, 3096–3098.

Larsen PA, Marchan-Rivadeneira MR, Baker RJ (2010) Natural hybridization generates mammalian lineage with species characteristics. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 11447–11452.

Maddison WP (1997) Gene trees in species trees. *Systematic Biology*, **46**, 523–536.

Maddison DR, Maddison WP (2000) MacClade 4: analysis of phylogeny and character evolution. Sinauer Associates, Inc., Sunderland, Massachusetts.

Mallet JM (2007) Hybrid speciation. *Nature*, **446**, 279–283.

McDonald DB, Parchman TL, Bower MR, Hubert WA, Rahel FJ (2008) An introduced and a native vertebrate hybridize to form a genetic bridge to a second native species.

Proceedings of the National Academy of Sciences of the United States of America, **105**, 10837–10842.

Meerburg BG, Singleton GR, Kijlstra A (2009) Rodent-borne diseases and their risks for public health. *Critical Reviews in Microbiology*, **35**, 221–270.

Mooney HA, Cleland EE (2001) The evolutionary impact of species invasions. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 5446–5451.

Musser GG, Carleton MD (2005) Superfamily Muroidea. In: *Mammals of the world: a taxonomic and geographic reference, 3rd ed* (eds Wilson DE, Reeder DR). pp. 894-1531, Johns Hopkins University Press, Baltimore, Maryland.

Neigel JE, Avise JC (1986) Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. In: *Evolutionary Processes and Theory* (eds Nevo E, Karlin S). pp. 515–534, Academic Press, New York, New York.

Nevado B, Fazalova V, Backeljau T, Hanssens M, Verheyen E (2011) Repeated unidirectional introgression of nuclear and mitochondrial DNA between four congeneric Tanganyikan cichlids. *Molecular Biology and Evolution*, **28**, 2253–2267.

Nowak R (1999) Walker's Mammals of the World (6th Edition). Johns Hopkins University Press, Baltimore, Maryland.

Olden JD, Poff NL, Douglas MR, Douglas ME, Fausch KD (2004) Ecological and evolutionary consequences of biotic homogenization. *Trends in Ecology and Evolution*, **19**, 18–24.

Pimentel D, Lach L, Zuniga R, Morrison D (2000) Environmental and economic costs of nonindigenous species in the United States. *BioScience*, **50**, 53–65.

Posada D (2008) jModeltest: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

Prentis PJ, Wilson JRU, Dormontt EE, Richardson DM, Lowe AJ (2008) Adaptive evolution in invasive species. *Trends in Ecology and Evolution*, **13**, 288–294.

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. *Journal of Heredity*, **86**, 248–249.

Rhymer JM, Simberloff D (1996) Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, **27**, 83–109.

Rickart EA, Balete DS, Rowe RJ, Heaney LR (2011) Mammals of the northern Philippines: tolerance for habitat disturbance and resistance to invasive species in an endemic insular fauna. *Diversity and Distributions*, **17**, 530–541.

Robins JH, Hingston M, Matisoo-Smith E, Ross HA (2007) Identifying *Rattus* species using mitochondrial DNA. *Molecular Ecology Notes*, **7**, 717–729.

Robins JH, McLenachan PA, Phillips MJ, Craig L, Ross HA, Matisoo-Smith E (2008) Dating of divergences within the *Rattus* genus phylogeny using whole mitochondrial genomes. *Molecular Phylogenetics and Evolution*, **49**, 460–466.

Rowe KC, Aplin KP, Braverstock PR, Moritz C (2011) Recent and rapid speciation with limited morphological disparity in the genus *Rattus*. *Systematic Biology*, **60**, 188–203.

Sax DF, Stachowicz JJ, Brown JH *et al.* (2007) Ecological and evolutionary insights from species invasions. *Trends in Ecology and Evolution*, **22**, 465–471.

Schaal BA, Gaskin JF, Caicedo AL (2003) Phylogeography, haplotype trees, and invasive plant species. *Journal of Heredity*, **94**, 197–204.

Seehausen O (2004) Hybridization and adaptive radiation. *Trends in Ecology and Evolution*, **19**, 198–207.

Sullivan R (2004) Rats: observations on the history and habitat of the city's most unwanted inhabitants. Bloomsbury Publishing, New York, New York.

Swofford DL (2003) PAUP* 4.0: phylogenetic analysis using parsimony (*and other methods). Version 4.0b10. Sinauer Associates, Inc., Sunderland, Massachusetts.

Weir BS (1996) Genetic data analysis II: methods for discrete population genetic data. Sinauer Associate, Inc., Sunderland, Massachusetts.

Wolf DE, Takebayashi N, Riesberg LH (2001) Predicting the risk of extinction through hybridization. *Conservation Biology*, **15**, 1039–1053.

Yosida TH (1977) Karyological studies on hybrids between Asian, Ceylonese, and Oceanian black rats, with a note on an XO female occurring in the F_2 generation. *Cytogenetics and Cell Genetics*, **19**, 262–272.

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

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

Figure Captions

Figure 4.1. Phylogenies depicting relationships among mitochondrial cytochrome b sequences (A) for all individuals examined in this study, and the two nuclear loci DHFR (B), and ATP5A1 (C) for a subset of samples. See Tables S4.1 and S4.2 for the list of samples included in each analysis. The mtDNA phylogeny (A) is the consensus Bayesian phylogram with the *R. tanezumi* (green), *R. rattus* I (blue), and *R. rattus* IV (red) clades shown and all other taxa pruned off. The numbers at major nodes are Bayesian posterior probabilities. The DHFR (B; -ln*L* = 2500.87) and ATP5A1 (C; -ln*L* = 2047.35) phylograms are the most likely topologies resulting from a maximum likelihood analysis. Terminal taxa are highlighted with colors corresponding to their clade assignment in the mtDNA phylogeny (A), and taxon names correspond to museum accession numbers and the sampling locality for the specimen given in Table S4.1. Outgroups have been removed for presentation.

Figure 4.2. Statistical parsimony haplotype networks generated from cytochrome b sequences for the *R. tanezumi*, *R. rattus* I, and *R. rattus* IV clades identified in the mtDNA phylogenetic analysis (Fig. 4.1A). Colored nodes on the haplotype networks correspond to individual sampled haplotypes, and the size of the circle corresponds to the frequency of the haplotype. The color and size of the pie for each haplotype corresponds to the sampling locality indicated in the global map by a colored dot and the frequency of that haplotype at that locality, respectively. The small black nodes represent extinct or unsampled haplotypes, and each uninterrupted straight line (independent of line length) corresponds to a single mutational step.

Figure 4.3. Posterior assignment probabilities per individual from all sampled localities (indicated at bottom) for the combined 9 microsatellite loci and all three mtDNA lineages (indicated at top), and illustrating *K* ranging from 2 to 6. Each distinct column represents a single individual, and each color corresponds to the probability of assignment to each of a number of clusters (*K*).

Table 4.1. Pairwise gene flow estimates averaged across three duplicate runs for the DHFR nuclear locus. Gray highlighted boxes indicate comparisons between two sympatric mtDNA lineages. Gene flow estimates correspond to the migration from the population in the left column into the population along the top of the table.

	Philippines (*R. tanezumi*)	**California (*R. tanezumi*)**	**Florida (*R. tanezumi*)**	**Vietnam (*R. tanezumi*)**	**California (*R. rattus* I)**	**Gainesville (*R. rattus* I)**	**Florida (*R. rattus* I)**	**Florida Keys (*R. rattus* I)**	**Philippines (*R. rattus* IV)**
Philippines (*R. tanezumi*)		1.7	47.7	51	37	67.7	157.7	0.3	374.3
California (*R. tanezumi*)	20.3		0.3	155.7	369.7	133	0.3	27.7	74.3
Florida (*R. tanezumi*)	0.3	260.3		12.3	128.3	461.7	332.3	0.3	0.3
Vietnam (*R. tanezumi*)	239	9	10.3		0.3	0.3	155	27.7	0.3
California (*R. rattus* I)	17.7	480.3	337.7	31		17	11	22.3	7
Gainesville (*R. rattus* I)	17	260.3	363.7	0.3	128.3		11	480.3	0.3
Florida (*R. rattus* I)	0.3	217.7	487	150.3	0.3	48.3		87.7	0.3
Florida Keys (*R. rattus* I)	19	241.7	60.3	0.3	0.3	52.3	27		0.3

Philippines (*R. rattus* IV)	381	18.3	0.3	387.3	0.3	41.7	129.7	6.3

Table 4.2. Pairwise gene flow estimates averaged across three duplicate runs for the ATP5A1 nuclear locus. Gray highlighted boxes indicate comparisons between two sympatric mtDNA lineages. Gene flow estimates correspond to the migration from the population in the left column into the population along the top of the table.

	Philippines (*R. tanezumi*)	**California (*R. tanezumi*)**	**Florida (*R. tanezumi*)**	**Vietnam (*R. tanezumi*)**	**California (*R. rattus* I)**	**Gainesville (*R. rattus* I)**	**Florida (*R. rattus* I)**	**Florida Keys (*R. rattus* I)**	**Philippines (*R. rattus* IV)**
Philippines (*R. tanezumi*)		80.3	16.3	341.3	2.3	0.3	59	0.3	381.7
California (*R. tanezumi*)	41		23.7	0.3	481	22.3	180.3	0.3	0.3
Florida (*R. tanezumi*)	0.3	180.3		15	335	0.3	381	381	14.3
Vietnam (*R. tanezumi*)	307	0.3	0.3		0.3	0.3	0.3	0.3	0.3
California (*R. rattus* I)	80.3	381.7	180.3	12.3		19	177.3	161	0.3
Gainesville (*R. rattus* I)	80.3	180.3	179.7	18.3	181		274.3	219.7	0.3
Florida (*R. rattus* I)	27.7	0.3	399.7	0.3	28.3	0.3		180.3	47.7
Florida Keys (*R. rattus* I)	175	180.3	341	0.3	180.3	0.3	181		0.3

Philippines (*R. rattus* IV)	380	37.7	18.3	399.3	0.3	18.3	121.7	13

Table 4.3. Results (*P* values) Hardy-Weinberg equilibrium tests for heterozygosity deficit and excess performed for the microsatellite data for each population of each species. Asterisks indicate significant *P*-values. Mitochondrial lineage is based on the cytochrome b phylogenetic analysis (Fig. 4.1A).

mtDNA Lineage	**Excess**	**Deficit**
R. tanezumi		
Vietnam	0.9951	0.4940
San Francisco Bay, CA	0.8884	0.0023*
Panama City, FL	0.2538	0.0034*
Philippines	0.9458	0.1157
R. rattus I		
Alaska	0.9495	0.5054
San Francisco Bay, CA	0.0093*	0.3283
Panama City, FL	0.0375*	0.4436
Gainesville, FL	0.8243	0.9997
Key Largo, FL	0.8832	0.1757
Arkansas	0.9595	0.4054
Louisiana	0.9855	0.1145
Nicaragua	0.9974	0.2699
Houston, TX	0.9996	0.8877
Austin, TX	0.9917	0.8301
San Angelo, TX	0.8649	0.1351
R. rattus IV		
Philippines	0.9989	0.7725

Fig. 4.1

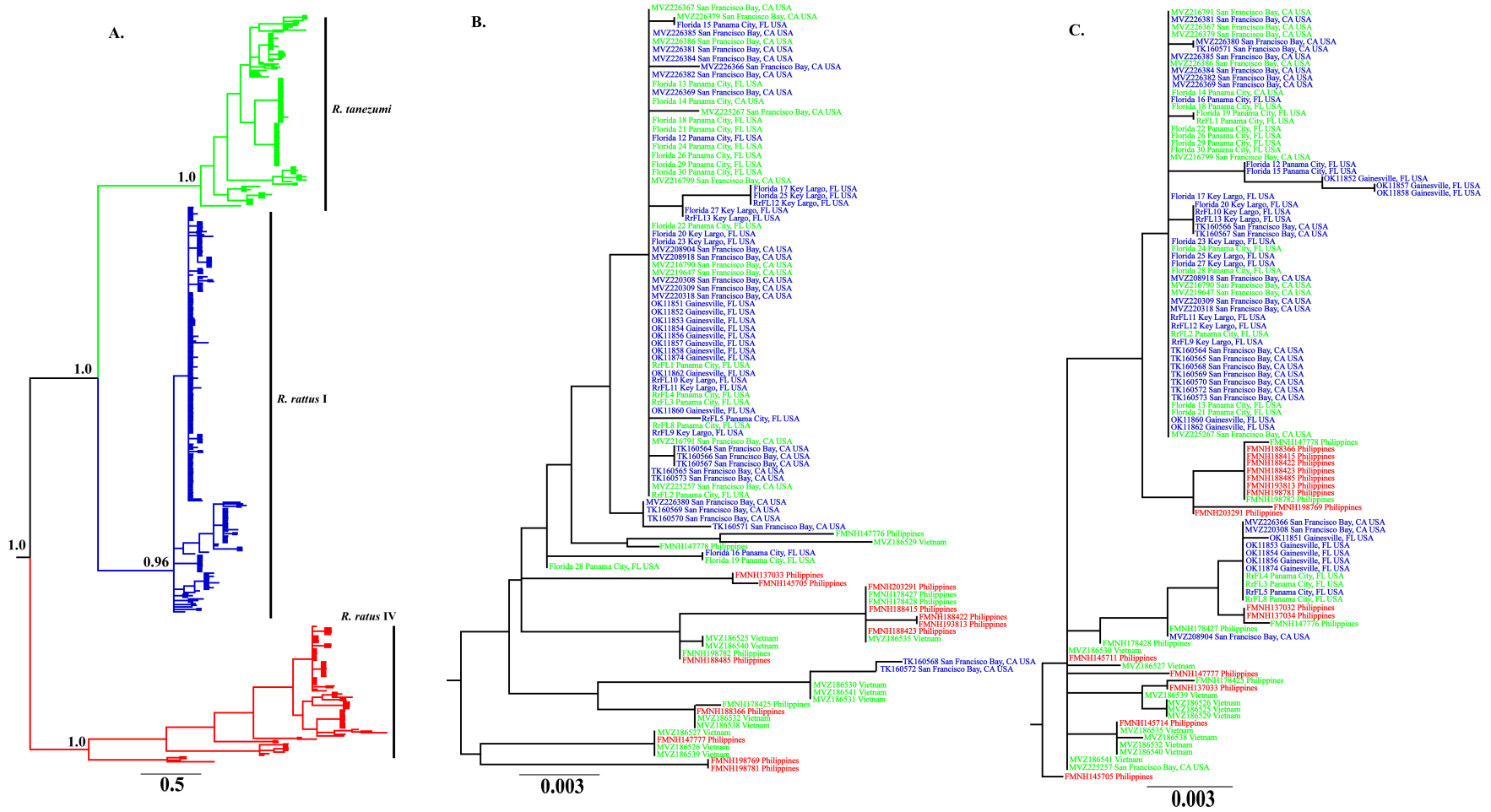

A.
R. tanezumi
R. rattus I
R. ratus IV
1.0
1.0
1.0
0.96
1.0
0.5
B.
MVZ226367 San Francisco Bay, CA USA
MVZ226379 San Francisco Bay, CA USA
Florida 15 Panama City, FL USA
MVZ226385 San Francisco Bay, CA USA
MVZ226386 San Francisco Bay, CA USA
MVZ226381 San Francisco Bay, CA USA
MVZ226384 San Francisco Bay, CA USA
MVZ226366 San Francisco Bay, CA USA
MVZ226382 San Francisco Bay, CA USA
Florida 13 Panama City, FL USA
MVZ226369 San Francisco Bay, CA USA
Florida 14 Panama City, CA USA
MVZ225267 San Francisco Bay, CA USA
Florida 18 Panama City, FL USA
Florida 21 Panama City, FL USA
Florida 12 Panama City, FL USA
Florida 24 Panama City, FL USA
Florida 26 Panama City, FL USA
Florida 29 Panama City, FL USA
Florida 30 Panama City, FL USA
MVZ216799 San Francisco Bay, CA USA
Florida 17 Key Largo, FL USA
Florida 25 Key Largo, FL USA
RrFL12 Key Largo, FL USA
Florida 27 Key Largo, FL USA
RrFL13 Key Largo, FL USA
Florida 22 Panama City, FL USA
Florida 20 Key Largo, FL USA
Florida 23 Key Largo, FL USA
MVZ208904 San Francisco Bay, CA USA
MVZ208918 San Francisco Bay, CA USA
MVZ216790 San Francisco Bay, CA USA
MVZ219647 San Francisco Bay, CA USA
MVZ220308 San Francisco Bay, CA USA
MVZ220309 San Francisco Bay, CA USA
MVZ220318 San Francisco Bay, CA USA
OK11851 Gainesville, FL USA
OK11852 Gainesville, FL USA
OK11853 Gainesville, FL USA
OK11854 Gainesville, FL USA
OK11856 Gainesville, FL USA
OK11857 Gainesville, FL USA
OK11858 Gainesville, FL USA
OK11874 Gainesville, FL USA
RrFL1 Panama City, FL USA
OK11862 Gainesville, FL USA
RrFL10 Key Largo, FL USA
RrFL11 Key Largo, FL USA
RrFL4 Panama City, FL USA
RrFL3 Panama City, FL USA
OK11860 Gainesville, FL USA
RrFL5 Panama City, FL USA
RrFL8 Panama City, FL USA
RrFL9 Key Largo, FL USA
MVZ216791 San Francisco Bay, CA USA
TK160564 San Francisco Bay, CA USA
TK160566 San Francisco Bay, CA USA
TK160567 San Francisco Bay, CA USA
TK160565 San Francisco Bay, CA USA
TK160573 San Francisco Bay, CA USA
MVZ225257 San Francisco Bay, CA USA
RrFL2 Panama City, FL USA
MVZ226380 San Francisco Bay, CA USA
TK160569 San Francisco Bay, CA USA
TK160570 San Francisco Bay, CA USA
TK160571 San Francisco Bay, CA USA
FMNH147776 Philippines
MVZ186529 Vietnam
FMNH147778 Philippines
Florida 16 Panama City, FL USA
Florida 19 Panama City, FL USA
Florida 28 Panama City, FL USA
FMNH137033 Philippines
FMNH145705 Philippines
FMNH203291 Philippines
FMNH178427 Philippines
FMNH178428 Philippines
FMNH188415 Philippines
FMNH188422 Philippines
FMNH193813 Philippines
FMNH188423 Philippines
MVZ186535 Vietnam
MVZ186525 Vietnam
MVZ186540 Vietnam
FMNH198782 Philippines
FMNH188485 Philippines
TK160568 San Francisco Bay, CA USA
TK160572 San Francisco Bay, CA USA
MVZ186530 Vietnam
MVZ186541 Vietnam
MVZ186531 Vietnam
FMNH178425 Philippines
FMNH188366 Philippines
MVZ186532 Vietnam
MVZ186538 Vietnam
MVZ186527 Vietnam
FMNH147777 Philippines
MVZ186526 Vietnam
MVZ186539 Vietnam
FMNH198769 Philippines
FMNH198781 Philippines
0.003
C.
MVZ216791 San Francisco Bay, CA USA
MVZ226381 San Francisco Bay, CA USA
MVZ226367 San Francisco Bay, CA USA
MVZ226379 San Francisco Bay, CA USA
MVZ226380 San Francisco Bay, CA USA
TK160571 San Francisco Bay, CA USA
MVZ226385 San Francisco Bay, CA USA
MVZ226386 San Francisco Bay, CA USA
MVZ226384 San Francisco Bay, CA USA
MVZ226382 San Francisco Bay, CA USA
MVZ226369 San Francisco Bay, CA USA
Florida 14 Panama City, CA USA
Florida 16 Panama City, FL USA
Florida 18 Panama City, FL USA
Florida 19 Panama City, FL USA
RrFL1 Panama City, FL USA
Florida 22 Panama City, FL USA
Florida 26 Panama City, FL USA
Florida 29 Panama City, FL USA
Florida 30 Panama City, FL USA
MVZ216799 San Francisco Bay, CA USA
Florida 12 Panama City, FL USA
Florida 15 Panama City, FL USA
OK11852 Gainesville, FL USA
OK11857 Gainesville, FL USA
OK11858 Gainesville, FL USA
Florida 17 Key Largo, FL USA
Florida 20 Key Largo, FL USA
RrFL10 Key Largo, FL USA
RrFL13 Key Largo, FL USA
TK160566 San Francisco Bay, CA USA
TK160567 San Francisco Bay, CA USA
Florida 23 Key Largo, FL USA
Florida 24 Panama City, FL USA
Florida 25 Key Largo, FL USA
Florida 27 Key Largo, FL USA
Florida 28 Panama City, FL USA
MVZ208918 San Francisco Bay, CA USA
MVZ216790 San Francisco Bay, CA USA
MVZ219647 San Francisco Bay, CA USA
MVZ220309 San Francisco Bay, CA USA
MVZ220318 San Francisco Bay, CA USA
RrFL11 Key Largo, FL USA
RrFL12 Key Largo, FL USA
RrFL2 Panama City, FL USA
RrFL9 Key Largo, FL USA
TK160564 San Francisco Bay, CA USA
TK160565 San Francisco Bay, CA USA
TK160568 San Francisco Bay, CA USA
TK160569 San Francisco Bay, CA USA
TK160570 San Francisco Bay, CA USA
TK160572 San Francisco Bay, CA USA
TK160573 San Francisco Bay, CA USA
Florida 13 Panama City, FL USA
Florida 21 Panama City, FL USA
OK11860 Gainesville, FL USA
OK11862 Gainesville, FL USA
MVZ225267 San Francisco Bay, CA USA
FMNH147778 Philippines
FMNH188366 Philippines
FMNH188415 Philippines
FMNH188422 Philippines
FMNH188423 Philippines
FMNH188485 Philippines
FMNH193813 Philippines
FMNH198781 Philippines
FMNH198782 Philippines
FMNH198769 Philippines
FMNH203291 Philippines
MVZ226366 San Francisco Bay, CA USA
MVZ220308 San Francisco Bay, CA USA
OK11851 Gainesville, FL USA
OK11853 Gainesville, FL USA
OK11854 Gainesville, FL USA
OK11856 Gainesville, FL USA
OK11874 Gainesville, FL USA
RrFL4 Panama City, FL USA
RrFL3 Panama City, FL USA
RrFL5 Panama City, FL USA
RrFL8 Panama City, FL USA
FMNH137032 Philippines
FMNH137034 Philippines
FMNH147776 Philippines
FMNH178427 Philippines
MVZ208904 San Francisco Bay, CA USA
FMNH178428 Philippines
MVZ186530 Vietnam
FMNH145711 Philippines
MVZ186527 Vietnam
FMNH147777 Philippines
FMNH178423 Philippines
FMNH137033 Philippines
MVZ186539 Vietnam
MVZ186526 Vietnam
MVZ186535 Vietnam
MVZ186529 Vietnam
FMNH145714 Philippines
MVZ186533 Vietnam
MVZ186538 Vietnam
MVZ186532 Vietnam
MVZ186540 Vietnam
MVZ186541 Vietnam
MVZ225257 San Francisco Bay, CA USA
FMNH145705 Philippines
0.003

Fig. 4.2

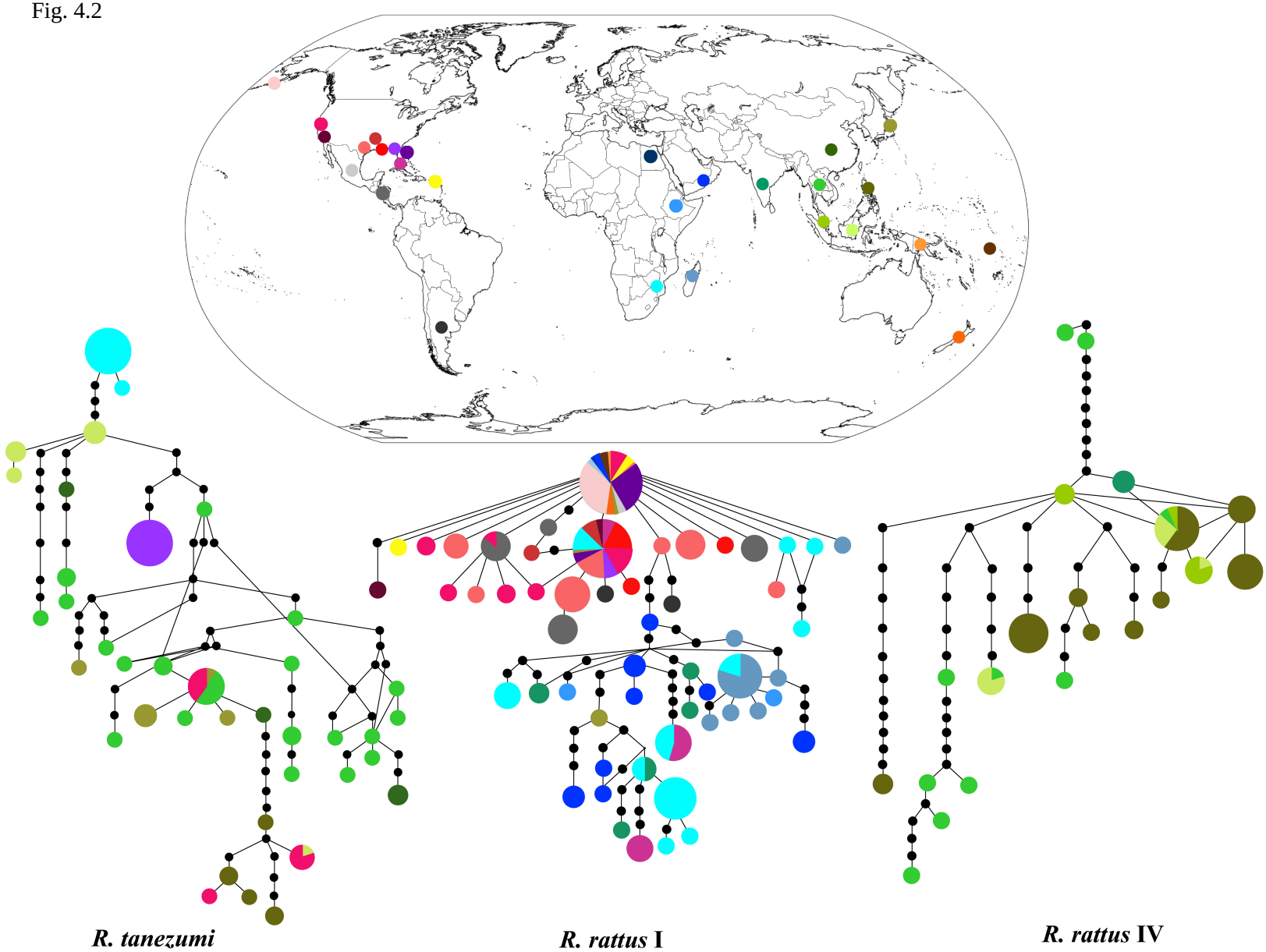
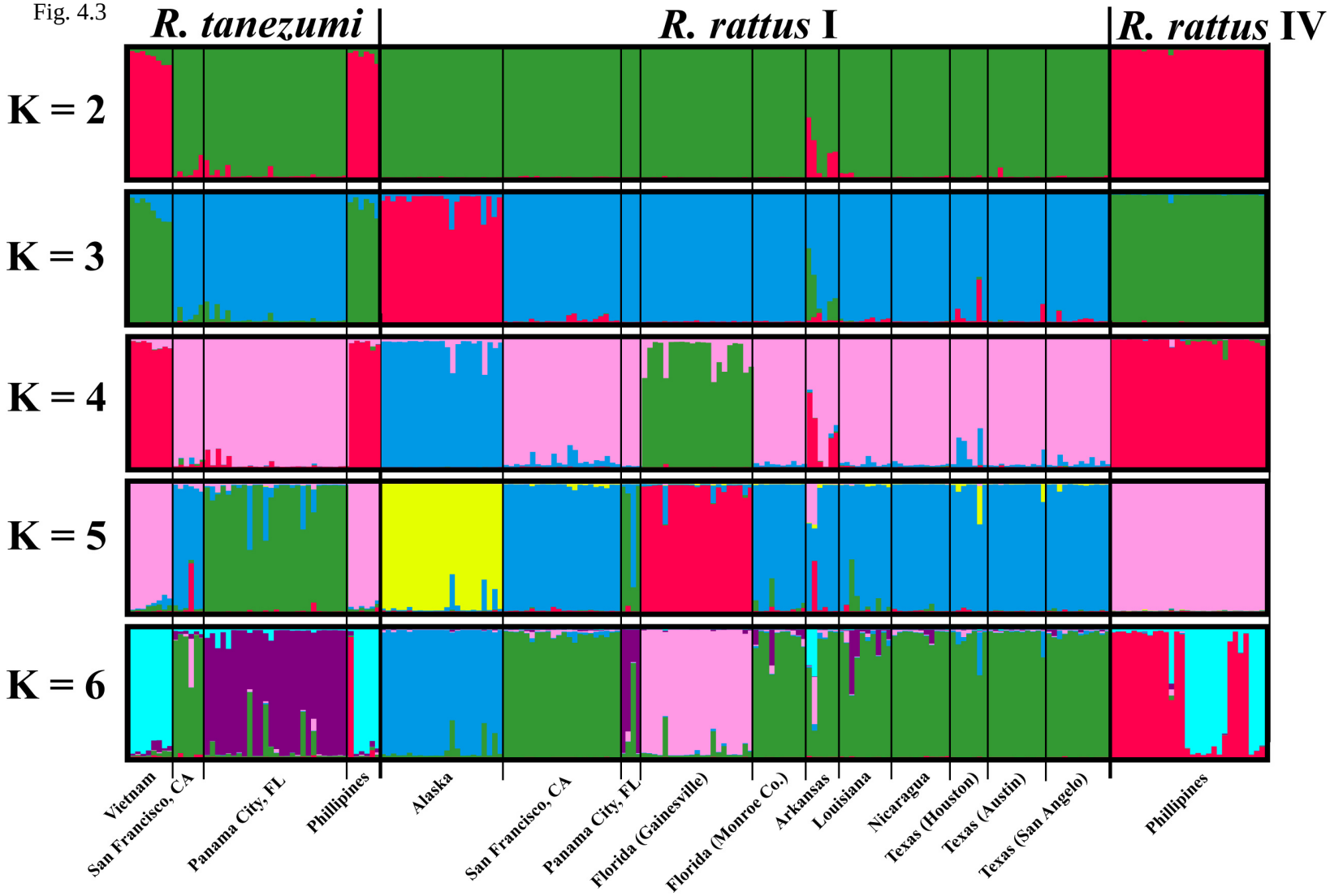


Fig. 4.3


R. tanezumi
R. rattus I
R. rattus IV
K = 2
K = 3
K = 4
K = 5
K = 6
Vietnam
San Francisco, CA
Panama City, FL
Phillipines
Alaska
San Francisco, CA
Panama City, FL
Florida (Gainesville)
Florida (Monroe Co.)
Arkansas
Louisiana
Nicaragua
Texas (Houston)
Texas (Austin)
Texas (San Angelo)
Phillipines

CHAPTER V

CONCLUSION

Invasive *Rattus* are collectively the most destructive and costly invasive vertebrate taxa on the planet. Through their commensal relationship with humans, they have simultaneously caused widespread ecological destruction, especially on islands, as well as spread some of the most deadly pathogens in human history (Gratz 1984). My study of invasive *Rattus* has highlighted multiple facets of the *Rattus* invasion and provided insight into the biological process of species invasions in general.

In my examination of the *Rattus rattus* invasion in both the U.S. and Asia, I have provided insight into the evolutionary implications of the dynamic process of colonization through invasion, during which novel interspecific interactions can result. Through the spread of multiple mitochondrial lineages, widespread hybridization with unidirectional introgression from two distinct mtDNA lineages into the *R. tanezumi* lineage has lead to a significant loss of *R. tanezumi* nuclear genome diversity, which has been replaced by the nuclear diversity of the other two taxa. Although this is monumental in the fact that it represents the rare occasion of pronounced introgression in a species complex (as opposed to among a pair of sister taxa), and at an essentially global scale, this

also represents an opportunity to study the importance of hybridization in the process of invasion.

This group may also present a rare opportunity to study the evolution of reproductive isolation in a novel model. The strong unidirectional pattern of nuclear gene flow suggests some role for genetic isolating mechanisms leading to asymmetrical fitness in back-crossed hybrids (Coyne and Orr 2004). If this is in fact the case, laboratory studies including controlled breeding, back-crossing, and genetic mapping may be possible, allowing us to identify specific loci that may be dictating the observed patterns for the nuclear genome. In addition, the plethora of genomic tools available for the closely related *R. norvegicus* make the study of *R. rattus* and other members of the black rat complex especially attractive and potentially very lucrative.

In my examination of colonization history and dispersal for *R. rattus* and *R. norvegicus* in the U.S., I detected clear differences between these two species on multiple fronts. In terms of colonization history, *R. rattus* appear to exhibit a relatively simple pattern, with essentially a single rapid expansion colonizing the U.S., with the exception of the southern tip of Florida. In contrast, *R. norvegicus* exhibited a more complex pattern of colonization, with a total of four distinct haplogroups indicating four distinct invasions either in space or time. In terms of dispersal and population structure within the U.S., my analyses suggest *R. rattus* to be dispersing more naturally and fitting a pattern of isolation by distance, while *R. norvegicus* may be exhibiting a high frequency of long-distance dispersal. Furthermore, the combination of these results and what we know about the life-history of these two species suggests competitive interactions among

these two species may be driving contemporary invasion dynamics, with *R. norvegicus* limiting the establishment and spread of *R. rattus* in the U.S.

These results have clear implications in the management of these two species, and has more broad implications for invasive species management in general. On both levels, my analysis suggests management of an invasive species must be undertaken with interspecific interactions being considered, and not in isolation, as is most often the case (Simberloff 2009). Furthermore, *R. norvegicus* clearly presents a higher risk in terms of the spread of infectious disease. With its high establishment rate from a diversity of global populations, *R. norvegicus* is more likely to bring in novel pathogens and rapidly (as well as unpredictably) spread them across the U.S.

Finally, my study of invasive *Rattus* as vectors for zoonotic hepatitis E virus (HEV) provides the first conclusive evidence for widespread infection of these human commensals with this human pathogen. This investigation represents the first step in identifying whether the high HEV infection rate reported for humans in urban areas (Mast *et al.* 1997; Thomas *et al.* 1997; Meng *et al*. 2002) is a result of rat-human interactions. Furthermore, my study sets the stage for much more rigorous examinations of invasive rodents in the epidemiology of HEV. The zoonotic HEV genotype 3 appears to exhibit a wide range of mammalian hosts, suggesting a very complex epidemiological history, which has been extremely difficult to unravel in previous examinations. Rodents may represent a key missing component of past treatments, potentially shedding light on the source(s) of human infection.

References

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Inc., Sunderland, MA.

Gratz NG (1984) The global public health importance of rodents. In: *Proceedings of a conference on the organization and practice of vertebrate pest control* (ed. Dubock, AC). Pp. 413–435, World Health Organization, Geneva, Switzerland.

Mast EE, Kuramoto IK, Favorov MO, Schoening VR, Burkholder BT, Shapiro CN, Holland PV (1997) Prevalence of and risk factors for antibody to hepatitis E virus seroreactivity among blood donors in northern California. *Journal of Infectious Disease*, **176**, 34–40.

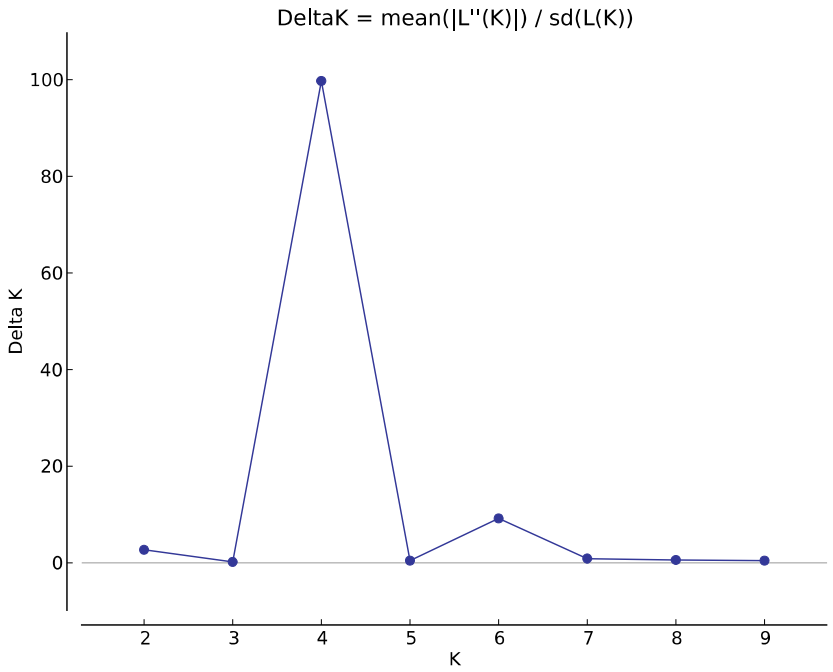
Meng XJ, Wiseman B, Elvinger F, Guenette DK, Toth TE, Engle RE, Emerson SU, Purcell RH (2002) Prevalence of antibodies to hepatitis E virus in veterinarians working with swine and in normal blood donors in the United States and other countries. *Journal of Clinical Microbiology*, **40**, 117–122.

Simberloff D (2009) We can eliminate invasions or live with them. Successful management projects. *Biological Invasions*, **11**, 149–157.

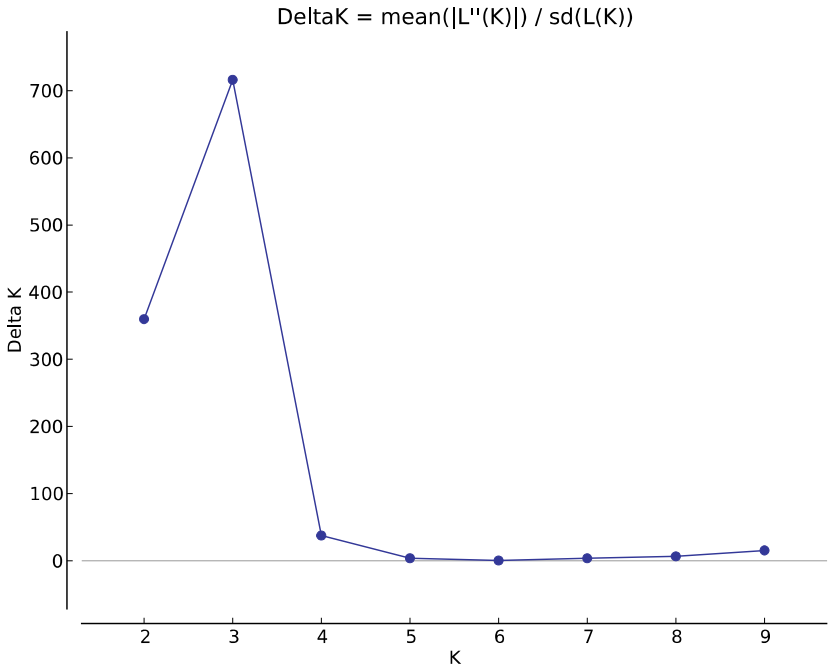
Thomas DL, Yarbough PO, Vlahov D, Tsarev SA, Nelson KE, Saah AJ, Purcell RH (1997) Seroreactivity to hepatitis E virus in areas where the disease is not endemic. Journal of Clinical Microbiology, **35**, 1244–1247.

SUPPLEMENTARY MATRIALS

Supplementary Figure S2.1



Supplementary Figure S2.2



Supplementary Table S2.1. Specimens for which molecular data was generated in this study. An "X" in the cytochrome b or microsatellites columns indicates the individuals was included in the cytochrome b or microsatellite dataset. Acronyms in the individual's names correspond to the following collections: OCGR, Sam Noble Museum, Norman, Oklahoma; MVZ, Museum of Vertebrate Zoology, Berkeley, California; RMT, Field Collection of Dr. Robert Timm, University of Kansas; SP, Special Collections of the Carnegie Museum of Natural History; TK, The Museum at Texas Tech University; FMNH, The Field Museum; EAR and LRH, Field Collections of Dr. Lawrence Heaney, The Field Museum; UAM, University of Alaska-Fairbanks Museum of Natural History; OK, Oklahoma State University Collection of Vertebrates; RrFL and MC2, Uncatalogued specimens from the Oklahoma State University Collection of Vertebrates; M-, The Louisiana State University Museum of Natural Science; ASK, Angelo State University Museum of Natural History; ACU, Abilene Christian University Natural History Collection.

Name	**City**	**County**	**State**	**Country**	**Species**	**Cytochrome *b***	**Microsatellites**
OCGR1611		1 de Mayo	Chaco	Argentina	*R. rattus*	X	
OCGR1913		Yerba Buena	Tucuman	Argentina	*R. rattus*	X	
RMT4738			Bijagual	Costa Rica	*R. rattus*	X	
RMT4796			Bijagual	Costa Rica	*R. rattus*	X	
RMT4807			Bijagual	Costa Rica	*R. rattus*	X	
RMT4812			Bijagual	Costa Rica	*R. rattus*	X	
RMT4871			Bijagual	Costa Rica	*R. rattus*	X	
RMT4884			Bijagual	Costa Rica	*R. rattus*	X	
RMT4900			Bijagual	Costa Rica	*R. rattus*	X	

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
SP10241				Egypt	*R. rattus*	X	
SP10242				Egypt	*R. rattus*	X	
TK101036				Honduras	*R. rattus*	X	
TK136842				Honduras	*R. rattus*	X	
TK144811			Lesser Antilles	Lesser Antilles	*R. rattus*	X	
TK144810			Lesser Antilles	Lesser Antilles	*R. rattus*	X	
TK161246			Lesser Antilles	Lesser Antilles	*R. rattus*	X	
TK161279			Lesser Antilles	Lesser Antilles	*R. rattus*	X	
ASK8016	Miacatlan		Morelos	Mexico	*R. rattus*	X	
TK72396	Durango		Durango	Mexico	*R. rattus*	X	
TK150539	Oaxaca		Oaxaca	Mexico	*R. rattus*	X	
TK72907				Nicaragua	*R. rattus*	X	
TK72908				Nicaragua	*R. rattus*	X	
TK72910				Nicaragua	*R. rattus*	X	

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
TK72911				Nicaragua	*R. rattus*	X	
TK72913				Nicaragua	*R. rattus*	X	
TK72919				Nicaragua	*R. rattus*	X	
TK72921				Nicaragua	*R. rattus*	X	
TK72924				Nicaragua	*R. rattus*	X	
TK72927				Nicaragua	*R. rattus*	X	
TK72933				Nicaragua	*R. rattus*	X	
TK72934				Nicaragua	*R. rattus*	X	
SP6156				South Africa	*R. rattus*	X	
UAM85048	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85172	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85173	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85174	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85175	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85176	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85177	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
UAM85179	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85184	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85188	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85191	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85192	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85193	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85194	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85195	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85196	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85198	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85199	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85200	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85201	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM85202	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM86988	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
UAM91776	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
UAM97703	Great Sitkin Is.	Aleutian Is.	Alaska	USA	*R. rattus*	X	X
OK6153	Little Rock	Pulaski	Arkansas	USA	*R. rattus*	X	X
OK6154	Little Rock	Pulaski	Arkansas	USA	*R. rattus*	X	X
OK6156	Little Rock	Pulaski	Arkansas	USA	*R. rattus*	X	X
OK6158	Little Rock	Pulaski	Arkansas	USA	*R. rattus*	X	X
OK6159	Little Rock	Pulaski	Arkansas	USA	*R. rattus*	X	X
OK6155	Little Rock	Pulaski	Arkansas	USA	*R. rattus*	X	X
TK160564	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
TK160565	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
TK160566	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
TK160567	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
TK160568	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
TK160569	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
TK160570	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
TK160571	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
TK160572	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
TK160573	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
MVZ216790	San Francisco Bay	Alameda	California	USA	*R. tanezumi*	Excluded	X
MVZ216791	San Francisco Bay	Alameda	California	USA	*R. tanezumi*	Excluded	X
MVZ220308	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
MVZ220309	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
MVZ220318	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
MVZ208904	San Francisco Bay	Contra Costa	California	USA	*R. rattus*	X	X
MVZ208918	San Francisco Bay	Contra Costa	California	USA	*R. rattus*	X	X
MVZ219647	San Francisco Bay	Contra Costa	California	USA	*R. tanezumi*	Excluded	X
MVZ225257	San Francisco Bay	Contra Costa	California	USA	*R. tanezumi*	Excluded	
MVZ225267	San Francisco Bay	Contra Costa	California	USA	*R. tanezumi*	Excluded	
MVZ216799	San Francisco Bay	Contra Costa	California	USA	*R. tanezumi*	Excluded	
MVZ226367	San Francisco Bay	Alameda	California	USA	*R. tanezumi*	Excluded	X
MVZ226369	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
MVZ226384	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
MVZ226380	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
MVZ226366	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
MVZ226385	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
MVZ226379	San Francisco Bay	Alameda	California	USA	*R. tanezumi*	Excluded	X
MVZ226381	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
MVZ226382	San Francisco Bay	Alameda	California	USA	*R. rattus*	X	X
MVZ226386	San Francisco Bay	Alameda	California	USA	*R. tanezumi*	Excluded	X
TK113626	San Francisco Bay	Marin	California	USA	*R. rattus*	X	X
TK113734	San Diego	San Diego	California	USA	*R. rattus*	X	X
TK93474	San Diego	San Diego	California	USA	*R. rattus*	X	X
MVZ219755	Tehama	Tehama County	California	USA	*R. rattus*	X	X
MVZ219756	Tehama	Tehama County	California	USA	*R. rattus*	X	X
MVZ219757	Tehama	Tehama County	California	USA	*R. rattus*	X	X
MVZ219758	Tehama	Tehama County	California	USA	*R. rattus*	X	X
MVZ219760	Tehama	Tehama County	California	USA	*R. rattus*	X	X
MVZ219761	Tehama	Tehama County	California	USA	*R. rattus*	X	X
MVZ219762	Tehama	Tehama County	California	USA	*R. rattus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
MVZ219766	Tehama	Tehama County	California	USA	*R. rattus*	X	X
MVZ219767	Tehama	Tehama County	California	USA	*R. rattus*	X	X
MVZ219624	Tehama	Tehama County	California	USA	*R. rattus*	X	X
MVZ219628	Tehama	Tehama County	California	USA	*R. rattus*	X	X
MVZ219629	Tehama	Tehama County	California	USA	*R. rattus*	X	X
MVZ219630	Tehama	Tehama County	California	USA	*R. rattus*	X	X
MVZ219631	Tehama	Tehama County	California	USA	*R. rattus*	X	X
MVZ219632	Tehama	Tehama County	California	USA	*R. rattus*	X	X
MVZ219633	Tehama	Tehama County	California	USA	*R. rattus*	X	X
MVZ219634	Tehama	Tehama County	California	USA	*R. rattus*	X	X
OK11855	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11867	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11874	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11851	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11852	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11853	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
OK11854	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11856	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11857	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11858	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11860	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11862	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11863	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11864	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11865	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11866	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11868	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11869	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11871	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11878	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
OK11879	Gainesville	Alachua	Florida	USA	*R. rattus*	X	X
RrFL5 (E)	Panama City	Bay	Florida	USA	*R. rattus*	X	X

Name	**City**	**County**	**State**	**Country**	**Species**	**Cytochrome *b***	**Microsatellites**
RrFL1 (A)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
RrFL2 (B)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
RrFL3 (C)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
RrFL4 (D)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
RrFL8 (H)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11918 (Florida 1)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11928 (Florida 2)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11891 (Florida 3)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11923 (Florida 4)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11916 (Florida 5)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11921 (Florida 6)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11926 (Florida 7)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
OK11932 (Florida 8)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11935 (Florida 9)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11924 (Florida 10)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11931 (Florida 11)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11912 (Florida 12)	Panama City	Bay	Florida	USA	*R. rattus*	X	X
OK11914 (Florida 13)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11908 (Florida 14)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11894 (Florida 15)	Panama City	Bay	Florida	USA	*R. rattus*	X	X
OK11903 (Florida 16)	Panama City	Bay	Florida	USA	*R. rattus*	X	X
OK11881 (Florida 18)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
OK11906 (Florida 19)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11897 (Florida 21)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11888 (Florida 22)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11900 (Florida 24)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11904 (Florida 26)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11890 (Florida 28)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11901 (Florida 29)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
OK11907 (Florida 30)	Panama City	Bay	Florida	USA	*R. tanezumi*	Excluded	X
MC2A	Miami	Martin	Florida	USA	*R. rattus*	X	X
MC2B	Miami	Martin	Florida	USA	*R. rattus*	X	X
MC2C	Miami	Martin	Florida	USA	*R. rattus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
RrFL11	Key Largo	Monroe	Florida	USA	*R. rattus*	X	X
RrFL12	Key Largo	Monroe	Florida	USA	*R. rattus*	X	X
RrFL9	Key Largo	Monroe	Florida	USA	*R. rattus*	X	X
RrFL10	Key Largo	Monroe	Florida	USA	*R. rattus*	X	X
RrFL13	Key Largo	Monroe	Florida	USA	*R. rattus*	X	X
OK11884 (Florida 17)	Key Largo	Monroe	Florida	USA	*R. rattus*	X	X
OK11882 (Florida 20)	Key Largo	Monroe	Florida	USA	*R. rattus*	X	X
OK11893 (Florida 23)	Key Largo	Monroe	Florida	USA	*R. rattus*	X	X
OK11886 (Florida 25)	Key Largo	Monroe	Florida	USA	*R. rattus*	X	X
OK11899 (Florida 27)	Key Largo	Monroe	Florida	USA	*R. rattus*	X	X
M-2990	Baton Rouge	Baton Rouge Parish	Louisiana	USA	*R. rattus*	X	X
M-3418	Baton Rouge	Baton Rouge Parish	Louisiana	USA	*R. rattus*	X	X
M-3419	Baton Rouge	Baton Rouge Parish	Louisiana	USA	*R. rattus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
M-3436	Baton Rouge	Baton Rouge Parish	Louisiana	USA	*R. rattus*	X	X
M-3437	Baton Rouge	Baton Rouge Parish	Louisiana	USA	*R. rattus*	X	X
M-3438	Baton Rouge	Baton Rouge Parish	Louisiana	USA	*R. rattus*	X	X
M-3439	Baton Rouge	Baton Rouge Parish	Louisiana	USA	*R. rattus*	X	X
M-3440	Baton Rouge	Baton Rouge Parish	Louisiana	USA	*R. rattus*	X	
M-1346	Baton Rouge	Baton Rouge Parish	Louisiana	USA	*R. rattus*	X	X
M-1928	Baton Rouge	Lafourche Parish	Louisiana	USA	*R. rattus*	X	
M-3420	Baton Rouge	Lafourche Parish	Louisiana	USA	*R. rattus*	X	X
M-1929		Jefferson Davis	Mississippi	USA	*R. rattus*	X	X
ASK3695	Brownsville	Cameron	Texas	USA	*R. rattus*	X	X
OK11861	Houston	Harris	Texas	USA	*R. rattus*	X	X
OK11872	Houston	Harris	Texas	USA	*R. rattus*	X	X
OK11873	Houston	Harris	Texas	USA	*R. rattus*	X	X
OK11875	Houston	Harris	Texas	USA	*R. rattus*	X	X
OK11876	Houston	Harris	Texas	USA	*R. rattus*	X	X
OK11877	Houston	Harris	Texas	USA	*R. rattus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
OK11880	Houston	Harris	Texas	USA	*R. rattus*	X	X
ASK5325	Weatherford	Parker	Texas	USA	*R. rattus*	X	X
ASK4763	San Angelo	Tom Green	Texas	USA	*R. rattus*	X	X
ASK6277	San Angelo	Tom Green	Texas	USA	*R. rattus*	X	X
ASK6283	San Angelo	Tom Green	Texas	USA	*R. rattus*	X	X
ASK6284	San Angelo	Tom Green	Texas	USA	*R. rattus*	X	X
ASK6286	San Angelo	Tom Green	Texas	USA	*R. rattus*	X	X
ASK6296	San Angelo	Tom Green	Texas	USA	*R. rattus*	X	X
ASK7213	San Angelo	Tom Green	Texas	USA	*R. rattus*	X	X
ASK7596	San Angelo	Tom Green	Texas	USA	*R. rattus*	X	X
ASK6367	San Angelo	Tom Green	Texas	USA	*R. rattus*	X	X
ASK3691	San Angelo	Tom Green	Texas	USA	*R. rattus*	X	X
ASK4753	San Angelo	Tom Green	Texas	USA	*R. rattus*	X	X
ASK5451	San Angelo	Tom Green	Texas	USA	*R. rattus*	X	X
ACU1400	San Angelo	Tom Green	Texas	USA	*R. rattus*	X	X
ACU868	San Angelo	Tom Green	Texas	USA	*R. rattus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
ASK6992	Austin	Travis	Texas	USA	*R. rattus*	X	X
ASK6993	Austin	Travis	Texas	USA	*R. rattus*	X	X
ASK6982	Austin	Travis	Texas	USA	*R. rattus*	X	X
ASK6986	Austin	Travis	Texas	USA	*R. rattus*	X	X
ASK6989	Austin	Travis	Texas	USA	*R. rattus*	X	X
ASK6990	Austin	Travis	Texas	USA	*R. rattus*	X	X
ASK6981	Austin	Travis	Texas	USA	*R. rattus*	X	X
ASK6983	Austin	Travis	Texas	USA	*R. rattus*	X	X
ASK6985	Austin	Travis	Texas	USA	*R. rattus*	X	X
ASK6991	Austin	Travis	Texas	USA	*R. rattus*	X	X
ASK6994	Austin	Travis	Texas	USA	*R. rattus*	X	X
ASK6887	Austin	Williamson	Texas	USA	*R. rattus*	X	X
ASK6888	Austin	Williamson	Texas	USA	*R. rattus*	X	X
UWBM81886	Seattle	King	Washington	USA	*R. rattus*	X	X
UWBM81800	Seattle	King	Washington	USA	*R. rattus*	X	X
UWBM79624	Seattle	King	Washington	USA	*R. rattus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
UWBM81887	Seattle	King	Washington	USA	*R. rattus*	X	X
UWBM75778	Seattle	King	Washington	USA	*R. rattus*	X	X
OCGR3101				Argentina	*R. norvegicus*	X	
TK93082	Tlaxcala			Mexico	*R. norvegicus*	X	
TK72926				Nicaragua	*R. norvegicus*	X	
TK144809				Lesser Antilles	*R. norvegicus*	X	
TK161119				Lesser Antilles	*R. norvegicus*	X	
TK161120				Lesser Antilles	*R. norvegicus*	X	
TK161121				Lesser Antilles	*R. norvegicus*	X	
TK161212				Lesser Antilles	*R. norvegicus*	X	
TK161213				Lesser Antilles	*R. norvegicus*	X	
UAM101009	Adak Is.	Aleutain Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101015	Adak Is.	Aleutain Is.	Alaska	USA	*R. norvegicus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
UAM101016	Adak Is.	Aleutain Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101017	Adak Is.	Aleutain Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101018	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101019	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101020	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101021	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101022	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101023	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101024	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101025	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101026	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101028	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101029	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101030	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101031	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101032	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
UAM101033	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101034	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101035	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101036	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101037	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101039	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101040	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101041	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101042	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101043	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101052	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM101071	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM41851	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM51851	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM52206	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM52208	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
UAM52213	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM52215	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM52217	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM52218	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM83260	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM83262	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM83263	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM51633	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*		X
UAM51898	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*		X
UAM52207	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*		X
UAM52209	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*		X
UAM52210	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*		X
UAM52211	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*		X
UAM52212	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*		X
UAM52214	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*		X
UAM52219	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*		X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
UAM83261	Adak Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*		X
UAM41895	Attu Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM44510	Attu Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM97705	Great Sitkin Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM97706	Great Sitkin Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM97707	Great Sitkin Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM97708	Great Sitkin Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM97709	Great Sitkin Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	
UAM97704	Sedanka Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*	X	X
UAM52220	Kagalaska Is.	Aleutian Is.	Alaska	USA	*R. norvegicus*		X
UAM51388	Douglas Is.	Alexander Archipelago	Alaska	USA	*R. norvegicus*		X
UAM54608	Revillagigedo Is.	Alexander Archipelago	Alaska	USA	*R. norvegicus*	X	X
UAM83264	Revillagigedo Is.	Alexander Archipelago	Alaska	USA	*R. norvegicus*	X	X
UAM86879	Fairbanks		Alaska	USA	*R. norvegicus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
OK6152	Little Rock	Pulaski	Arkansas	USA	*R. norvegicus*	X	X
OK6157	Little Rock	Pulaski	Arkansas	USA	*R. norvegicus*	X	X
OK6160	Little Rock	Pulaski	Arkansas	USA	*R. norvegicus*	X	X
TK93473	San Diego	San Diego	California	USA	*R. norvegicus*	X	X
TK93475	San Diego	San Diego	California	USA	*R. norvegicus*	X	X
TK90796	San Diego	San Diego	California	USA	*R. norvegicus*	X	X
TK93468	San Diego	San Diego	California	USA	*R. norvegicus*	X	X
TK93469	San Diego	San Diego	California	USA	*R. norvegicus*	X	X
TK93470	San Diego	San Diego	California	USA	*R. norvegicus*	X	X
TK93471	San Diego	San Diego	California	USA	*R. norvegicus*	X	X
TK93485	San Diego	San Diego	California	USA	*R. norvegicus*	X	X
TK93486	San Diego	San Diego	California	USA	*R. norvegicus*	X	X
TK93488	San Diego	San Diego	California	USA	*R. norvegicus*	X	X
TK93489	San Diego	San Diego	California	USA	*R. norvegicus*	X	X
PL7239	Chicago	Cook	Illinois	USA	*R. norvegicus*	X	X
PL7240	Chicago	Cook	Illinois	USA	*R. norvegicus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
PL7257	Chicago	Cook	Illinois	USA	*R. norvegicus*	X	X
PL7261	Chicago	Cook	Illinois	USA	*R. norvegicus*	X	X
PL7274	Chicago	Cook	Illinois	USA	*R. norvegicus*	X	X
PL7275	Chicago	Cook	Illinois	USA	*R. norvegicus*	X	X
PL7276	Chicago	Cook	Illinois	USA	*R. norvegicus*	X	X
PL7277	Chicago	Cook	Illinois	USA	*R. norvegicus*	X	X
PL7278	Chicago	Cook	Illinois	USA	*R. norvegicus*	X	X
PL7282	Chicago	Cook	Illinois	USA	*R. norvegicus*	X	X
PL7158	Chicago	Cook	Illinois	USA	*R. norvegicus*	X	X
NK53103		Spencer	Indiana	USA	*R. norvegicus*	X	X
NK53104		Spencer	Indiana	USA	*R. norvegicus*	X	X
NK53105		Spencer	Indiana	USA	*R. norvegicus*	X	X
NK53106		Spencer	Indiana	USA	*R. norvegicus*	X	X
NK53111		Spencer	Indiana	USA	*R. norvegicus*	X	X
NK53114		Spencer	Indiana	USA	*R. norvegicus*	X	X
NK53115		Spencer	Indiana	USA	*R. norvegicus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
NK53116		Spencer	Indiana	USA	*R. norvegicus*	X	X
NK53117		Spencer	Indiana	USA	*R. norvegicus*	X	X
NK53118		Spencer	Indiana	USA	*R. norvegicus*	X	X
1980s1	Baltimore		Maryland	USA	*R. norvegicus*		X
1980s2	Baltimore		Maryland	USA	*R. norvegicus*		X
1980s3	Baltimore		Maryland	USA	*R. norvegicus*		X
1980s4	Baltimore		Maryland	USA	*R. norvegicus*		X
1980s5	Baltimore		Maryland	USA	*R. norvegicus*		X
1980s6	Baltimore		Maryland	USA	*R. norvegicus*		X
1980s7	Baltimore		Maryland	USA	*R. norvegicus*		X
1980s8	Baltimore		Maryland	USA	*R. norvegicus*		X
1980s9	Baltimore		Maryland	USA	*R. norvegicus*		X
1980s10	Baltimore		Maryland	USA	*R. norvegicus*		X
1990s1	Baltimore		Maryland	USA	*R. norvegicus*	X	X
1990s2	Baltimore		Maryland	USA	*R. norvegicus*	X	X
1990s3	Baltimore		Maryland	USA	*R. norvegicus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
1990s4	Baltimore		Maryland	USA	*R. norvegicus*	X	X
1990s5	Baltimore		Maryland	USA	*R. norvegicus*	X	X
1990s6	Baltimore		Maryland	USA	*R. norvegicus*	X	X
1990s7	Baltimore		Maryland	USA	*R. norvegicus*	X	X
1990s8	Baltimore		Maryland	USA	*R. norvegicus*	X	X
1990s9	Baltimore		Maryland	USA	*R. norvegicus*	X	X
1990s10	Baltimore		Maryland	USA	*R. norvegicus*	X	X
2000s1	Baltimore		Maryland	USA	*R. norvegicus*	X	X
2000s2	Baltimore		Maryland	USA	*R. norvegicus*	X	X
2000s3	Baltimore		Maryland	USA	*R. norvegicus*	X	X
2000s4	Baltimore		Maryland	USA	*R. norvegicus*	X	X
2000s5	Baltimore		Maryland	USA	*R. norvegicus*	X	X
2000s6	Baltimore		Maryland	USA	*R. norvegicus*	X	X
2000s7	Baltimore		Maryland	USA	*R. norvegicus*	X	X
2000s8	Baltimore		Maryland	USA	*R. norvegicus*	X	X
2000s9	Baltimore		Maryland	USA	*R. norvegicus*	X	X

Name	**City**	**County**	**State**	**Country**	**Species**	**Cytochrome *b***	**Microsatellites**
2000s10	Baltimore		Maryland	USA	*R. norvegicus*	X	X
NK8768	Bernallio		New Mexico	USA	*R. norvegicus*	X	X
NK8769	Bernallio		New Mexico	USA	*R. norvegicus*	X	X
NYSM15268			New York	USA	*R. norvegicus*	X	X
OK6424	Oklahoma City	Oklahoma	Oklahoma	USA	*R. norvegicus*	X	X
OK6459	Oklahoma City	Oklahoma	Oklahoma	USA	*R. norvegicus*	X	X
OK11859	Corvalis	Benton	Oregon	USA	*R. norvegicus*	X	X
OK11870	Corvalis	Benton	Oregon	USA	*R. norvegicus*	X	X
OK6423	Corvalis	Benton	Oregon	USA	*R. norvegicus*	X	X
OK6427	Corvalis	Benton	Oregon	USA	*R. norvegicus*	X	X
OK11883		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11885		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11887		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11889		Union	Pennsylvania	USA	*R. norvegicus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
OK11892		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11895		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11896		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11898		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11902		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11905		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11909		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11910		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11911		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11913		Union	Pennsylvania	USA	*R. norvegicus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
OK11917		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11919		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11920		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11922		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11925		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11927		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11929		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11930		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11933		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11934		Union	Pennsylvania	USA	*R. norvegicus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
OK11937		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11938		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11939		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11940		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11941		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11942		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11943		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK6479		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK6483		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK6484		Union	Pennsylvania	USA	*R. norvegicus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
OK6489		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK6488		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK6481		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK6485		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK6480		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK6482		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK6486		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK6487		Union	Pennsylvania	USA	*R. norvegicus*	X	X
OK11045	Memphis	Shelby	Tennessee	USA	*R. norvegicus*	X	X
OK11046	Memphis	Shelby	Tennessee	USA	*R. norvegicus*	X	X
OK11047	Memphis	Shelby	Tennessee	USA	*R. norvegicus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
OK11048	Memphis	Shelby	Tennessee	USA	*R. norvegicus*	X	X
OK11049	Memphis	Shelby	Tennessee	USA	*R. norvegicus*	X	X
OK11050	Memphis	Shelby	Tennessee	USA	*R. norvegicus*	X	X
OK11051	Memphis	Shelby	Tennessee	USA	*R. norvegicus*	X	X
OK11052	Memphis	Shelby	Tennessee	USA	*R. norvegicus*	X	X
OK11053	Memphis	Shelby	Tennessee	USA	*R. norvegicus*	X	X
OK11054	Memphis	Shelby	Tennessee	USA	*R. norvegicus*	X	X
OK11055	Memphis	Shelby	Tennessee	USA	*R. norvegicus*	X	X
OK11056	Memphis	Shelby	Tennessee	USA	*R. norvegicus*	X	X
OK11057	Memphis	Shelby	Tennessee	USA	*R. norvegicus*	X	X
OK11058	Memphis	Shelby	Tennessee	USA	*R. norvegicus*	X	X
OK11059	Memphis	Shelby	Tennessee	USA	*R. norvegicus*	X	X
OK11060	Memphis	Shelby	Tennessee	USA	*R. norvegicus*	X	X
ASK6988	Austin		Texas	USA	*R. norvegicus*	X	X
ASK7342	San Angelo		Texas	USA	*R. norvegicus*	X	X
UWBM78688		King	Washington	USA	*R. norvegicus*	X	X

Name	City	County	State	Country	Species	Cytochrome *b*	Microsatellites
OK11936		Monroe	West Virginia	USA	*R. norvegicus*	X	
SP2323		Monroe	West Virginia	USA	*R. norvegicus*	X	X
SP8984		Monroe	West Virginia	USA	*R. norvegicus*	X	X
SP4640		Monroe	West Virginia	USA	*R. norvegicus*	X	X
SP8954		Monroe	West Virginia	USA	*R. norvegicus*	X	X
SP5680		Monroe	West Virginia	USA	*R. norvegicus*	X	X
NDM2817		Monroe	West Virginia	USA	*R. norvegicus*	X	X

Supplementary Table S2.2. Accession numbers for mitochondrial cytochrome b sequences obtained from GenBank, including the geographic locality and species.

Accession Number	City	State	Country	Species
AB355902	Haiphong port		Vietnam	*R. norvegicus*
AB355903	Haiphong port		Vietnam	*R. norvegicus*
AJ428514	Copenhagen		Denmark	*R. norvegicus*
DQ439839			South Africa	*R. norvegicus*
DQ439840			South Africa	*R. norvegicus*
DQ439841			South Africa	*R. norvegicus*
DQ439842			South Africa	*R. norvegicus*
DQ439843			South Africa	*R. norvegicus*
DQ439844			South Africa	*R. norvegicus*
DQ673916	Milwaukee	Wisconsin	USA	*R. norvegicus*
DQ673917	Tokyo		Japan	*R. norvegicus*
EF186461			French Polynesia	*R. norvegicus*
EF186462			French Polynesia	*R. norvegicus*
GU592954			China	*R.*

Accession Number	City	State	Country	Species
				norvegicus
GU592955			China	*R. norvegicus*
GU592956			China	*R. norvegicus*
GU592957			China	*R. norvegicus*
GU592958			China	*R. norvegicus*
GU592959			China	*R. norvegicus*
GU592960			China	*R. norvegicus*
GU592961			China	*R. norvegicus*
GU592962			China	*R. norvegicus*
GU592963			China	*R. norvegicus*
GU592964			China	*R. norvegicus*
GU592965			China	*R. norvegicus*
GU592966			China	*R. norvegicus*
GU592967			China	*R. norvegicus*
GU592968			China	*R. norvegicus*
GU592969			China	*R.*

Accession Number	City	State	Country	Species
				norvegicus
GU592970			China	*R. norvegicus*
GU592971			China	*R. norvegicus*
GU592972			China	*R. norvegicus*
HM217370			Thailand	*R. norvegicus*
HM217373			Thailand	*R. norvegicus*
HM217429			Thailand	*R. norvegicus*
HM217481			Cambodia	*R. norvegicus*
HM222710			China	*R. norvegicus*
AB211042	Kagoshima		Japan	*R. rattus*
AB211043	Kagoshima		Japan	*R. rattus*
AY263617			Indonesia	*R. rattus*
NC012374			New Zealand	*R. rattus*
DQ439830			South Africa	*R. rattus*
DQ439831			South Africa	*R. rattus*
DQ439832			South Africa	*R. rattus*
DQ439833			South Africa	*R. rattus*
DQ439834			South Africa	*R. rattus*
DQ439835			South Africa	*R. rattus*
DQ439836			South Africa	*R. rattus*

Accession Number	City	State	Country	Species
DQ439837			South Africa	*R. rattus*
DQ439838			South Africa	*R. rattus*
DQ439851			South Africa	*R. rattus*
DQ439852			South Africa	*R. rattus*
DQ439853			South Africa	*R. rattus*
DQ439854			South Africa	*R. rattus*
DQ439855			South Africa	*R. rattus*
DQ439856			South Africa	*R. rattus*
DQ439857			South Africa	*R. rattus*
DQ439858			South Africa	*R. rattus*
DQ439864			South Africa	*R. rattus*
EF186469			French Polynesia	*R. rattus*
EF186470			New Zealand	*R. rattus*
EF186472			Papua New Guinea	*R. rattus*
EF186473			Papua New Guinea	*R. rattus*
EF186474			French Polynesia	*R. rattus*
EF186475			Samoa	*R. rattus*
FJ842266			South Africa	*R. rattus*
FJ842267			South Africa	*R. rattus*
FJ842268			South Africa	*R. rattus*
GQ891569			India	*R. rattus*
GQ891570			India	*R. rattus*
GQ891571			India	*R. rattus*

Accession Number	City	State	Country	Species
GQ891572			India	*R. rattus*
GQ891573			India	*R. rattus*
GQ891574			Oman	*R. rattus*
GQ891575			Oman	*R. rattus*
GQ891576			Oman	*R. rattus*
GQ891577			Oman	*R. rattus*
GQ891578			Oman	*R. rattus*
GQ891579			Oman	*R. rattus*
GQ891580			Oman	*R. rattus*
GQ891581			Yemen	*R. rattus*
GQ891582			Yemen	*R. rattus*
GQ891583			Ethiopa	*R. rattus*
GQ891584			Ethiopa	*R. rattus*
GQ891585			Tanzania	*R. rattus*
GQ891586			Tanzania	*R. rattus*
GQ891587			Tanzania	*R. rattus*
GQ891588			Mozambique	*R. rattus*
GQ891589			Mozambique	*R. rattus*
GQ891590			Mozambique	*R. rattus*
GQ891591			Grand Comore	*R. rattus*
GQ891592			Grand Comore	*R. rattus*
GQ891593			Grand Comore	*R. rattus*
GQ891594			Grand Comore	*R. rattus*
GQ891595			Grand Comore	*R. rattus*

Accession Number	City	State	Country	Species
GQ891596			Grand Comore	*R. rattus*
GQ891597			Mayotte	*R. rattus*
GQ891598			Mayotte	*R. rattus*
GQ891599			Madagascar	*R. rattus*
GQ891600			Madagascar	*R. rattus*
GQ891601			Madagascar	*R. rattus*
GQ891602			Madagascar	*R. rattus*
GQ891603			Madagascar	*R. rattus*
GQ891604			Madagascar	*R. rattus*
GQ891605			Madagascar	*R. rattus*
GQ891606			Madagascar	*R. rattus*
GQ891607			Reunion	*R. rattus*
GQ891608			South Africa	*R. rattus*
HM217365			Tanzania	*R. rattus*
HM217366			Oman	*R. rattus*
HM217367			India	*R. rattus*
HM217368			Madagascar	*R. rattus*
HQ157800			South Africa	*R. rattus*
HQ157801			South Africa	*R. rattus*
HQ157802			South Africa	*R. rattus*
HQ157803			South Africa	*R. rattus*
HQ157806			South Africa	*R. rattus*
HQ157808			South Africa	*R. rattus*
HQ157809			South Africa	*R. rattus*

Supplementary Table S2.3. Locus designation, primer sequence, chromosomal location and either published citation or GenBank Accession for the nine microsatellite loci examined.

Locus	Primer Sequence (5' to 3')	Chromosome	Citation
D1Cebr4	GACCTCCTGCCCTTCACTG	1	Giraudeau et al. 1999
	TGAAAAATGAATTGCTTGTG		
D3Cebr3	CAGGGAATGCAGAAGATACAG	3	Giraudeau et al. 1999
	GCGGCTTTAGGACTCTGGAG		
D4Ceb2	TGTCAAAGAAAGCCAGTAAAAC	4	Giraudeau et al. 1999
	TTGGCAACCAGGAATAGC		
D5Cebr1	AACCGCCTGTATTTCTATTTC	5	Giraudeau et al. 1999
	GCCCAAGTTTGATCCTCAG		
D5Rat83	ACTTGGAAACAGGGAGATGG GGGTCTTCAGGATGGCAATGT	5	UniSTS:227970
D7Rat13	GACTTCTGCTACACGCCACA CAGCCCTAGAAGGAAATGCA	7	UniSTS:119018
D9Rat13	CCCATCTTTACACCTCCCAA GGAAAGGAAACTGGAGGGTC	9	UniSTS:119154
D11Mgh5	CAGCTCTAATTCCAGAAAGGTTT GAATCGATTGACAGATGTCTGTG	11	UniSTS:118224
D16Rat81	GAGCCTTAGCACAGTGGCTT GGCCCACATGTGCATGTATA	16	UniSTS:226803

Supplementary Table S4.1. Specimens for which molecular data was generated in this study. The mtDNA Clade/Species column refers to the position of the sequence in figure 4.1A, and individuals occurring in previously unnamed clades are listed as *Rattus* sp. All individuals included in this study were included in the cytochrome b phylogenetic analysis (Fig. 4.1A), and a "yes" in the Microsatellites or nuclear loci (DHFR, ATP5A1) columns indicate it was included those respective analyses. Acronyms in the individual's names correspond to the following collections: OCGR, Sam Noble Museum, Norman, Oklahoma; MVZ, Museum of Vertebrate Zoology, Berkeley, California; RMT, Field Collection of Dr. Robert Timm, University of Kansas; SP, Special Collections of the Carnegie Museum of Natural History; TK, The Museum at Texas Tech University; FMNH, The Field Museum; EAR and LRH, Field Collections of Dr. Lawrence Heaney, The Field Museum; UAM, University of Alaska-Fairbanks Museum of Natural History; OK, Oklahoma State University Collection of Vertebrates; RrFL and MC2, Uncatalogued specimens from the Oklahoma State University Collection of Vertebrates; M-, The Louisiana State University Museum of Natural Science; ASK, Angelo State University Museum of Natural History; ACU, Abilene Christian University Natural History Collection.

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
OCGR1611		1 de Mayo	Chaco	Argentina	*R. rattus* I			
OCGR1913		Yerba Buena	Tucuman	Argentina	*R. rattus* I			
MVZ193095			Yunnan	China	*R. andamanensis*			
MVZ176529			Yunnan	China	*R. tanezumi*			
MVZ176528			Yunnan	China	*R. tanezumi*			
MVZ176527			Yunnan	China	*R. tanezumi*			
MVZ176526			Yunnan	China	*R. tanezumi*			

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
RMT4738			Bijagual	Costa Rica	*R. rattus* I			
RMT4796			Bijagual	Costa Rica	*R. rattus* I			
RMT4807			Bijagual	Costa Rica	*R. rattus* I			
RMT4812			Bijagual	Costa Rica	*R. rattus* I			
RMT4871				Costa Rica	*R. rattus* I			
RMT4884				Costa Rica	*R. rattus* I			
RMT4900				Costa Rica	*R. rattus* I			
SP10241				Egypt	*R. rattus* I			
SP10242				Egypt	*R. rattus* I			
TK101036				Honduras	*R. rattus* I			
TK136842				Honduras	*R. rattus* I			
TK144811			Lesser Antilles	Lesser Antilles	*R. rattus* I			
TK144810			Lesser Antilles	Lesser Antilles	*R. rattus* I			
TK161246			Lesser Antilles	Lesser Antilles	*R. rattus* I			

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
TK161279			Lesser Antilles	Lesser Antilles	*R. rattus* I			
ASK8016	Miacatlan		Morelos	Mexico	*R. rattus* I			
TK72396	Durango		Durango	Mexico	*R. rattus* I			
TK150539	Oaxaca		Oaxaca	Mexico	*R. rattus* I			
TK72907				Nicaragua	*R. rattus* I	Yes		
TK72908				Nicaragua	*R. rattus* I	Yes		
TK72910				Nicaragua	*R. rattus* I	Yes		
TK72911				Nicaragua	*R. rattus* I	Yes		
TK72913				Nicaragua	*R. rattus* I	Yes		
TK72919				Nicaragua	*R. rattus* I	Yes		
TK72921				Nicaragua	*R. rattus* I	Yes		
TK72924				Nicaragua	*R. rattus* I	Yes		
TK72927				Nicaragua	*R. rattus* I	Yes		
TK72933				Nicaragua	*R. rattus* I	Yes		
TK72934				Nicaragua	*R. rattus* I	Yes		

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
FMNH198769			Benguet	Philippines	*R. rattus IV*	Yes	Yes	Yes
FMNH198770			Benguet	Philippines	*R. rattus IV*	Yes		
FMNH198771			Benguet	Philippines	*R. rattus IV*	Yes		
FMNH198775			Benguet	Philippines	*R. rattus IV*	Yes		
FMNH198781			Benguet	Philippines	*R. rattus IV*	Yes	Yes	Yes
FMNH198768			Benguet	Philippines	*R. rattus IV*	Yes		
FMNH198780			Benguet	Philippines	*R. rattus IV*	Yes		
FMNH198782			Benguet	Philippines	*R. tanezumi*	Yes	Yes	Yes
FMNH203921			Camarines Sur	Philippines	*R. rattus IV*	Yes	Yes	Yes
EAR1655			Catanduanes	Philippines	*R. rattus IV*	Yes		
EAR1485			Leyte	Philippines	*R. rattus IV*	Yes		
EAR1522			Leyte	Philippines	*R. rattus IV*	Yes		
EAR1486			Leyte	Philippines	*R. rattus IV*	Yes		
FMNH193818			Mountain	Philippines	*R. rattus IV*	Yes		
FMNH193819			Mountain	Philippines	*R. rattus IV*	Yes		

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
FMNH193820			Mountain	Philippines	*R. rattus IV*	Yes		
FMNH193821			Mountain	Philippines	*R. rattus IV*	Yes		
FMNH193825			Mountain	Philippines	*R. rattus IV*	Yes		
FMNH193977			Mountain	Philippines	*R. rattus IV*	Yes		
FMNH193979			Mountain	Philippines	*R. rattus IV*	Yes		
FMNH193812			Mountain	Philippines	*R. rattus IV*	Yes		
FMNH193813			Mountain	Philippines	*R. rattus IV*	Yes	Yes	Yes
FMNH193814			Mountain	Philippines	*R. rattus IV*	Yes		
FMNH193816			Mountain	Philippines	*R. rattus IV*	Yes		
FMNH188415			Mountain	Philippines	*R. rattus IV*	Yes	Yes	Yes
FMNH188416			Mountain	Philippines	*R. rattus IV*	Yes		
FMNH188419			Mountain	Philippines	*R. rattus IV*	Yes		
FMNH188422			Mountain	Philippines	*R. rattus IV*	Yes	Yes	Yes
FMNH188423			Mountain	Philippines	*R. rattus IV*	Yes	Yes	Yes
FMNH188366			Mountain	Philippines	*R. rattus IV*	Yes	Yes	Yes

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
FMNH188420			Mountain	Philippines	*R. rattus IV*	Yes		
FMNH188421			Mountain	Philippines	*R. rattus IV*	Yes		
FMNH188485			Mountain	Philippines	*R. rattus IV*	Yes	Yes	Yes
FMNH188414			Mountain	Philippines	*R. rattus IV*	Yes		
LRH3531			Negros Oriental	Philippines	*R. rattus IV*	Yes		
LRH3530			Negros Oriental	Philippines	*R. rattus IV*	Yes		
FMNH178425			Quezon	Philippines	*R. tanezumi*	Yes	Yes	Yes
FMNH178427			Quezon	Philippines	*R. tanezumi*	Yes	Yes	Yes
FMNH178428			Quezon	Philippines	*R. tanezumi*	Yes	Yes	Yes
FMNH137032	Sibuyan I		Romblon	Philippines	*R. rattus IV*	Yes		Yes
FMNH137033	Sibuyan I		Romblon	Philippines	*R. rattus IV*	Yes	Yes	Yes
FMNH137034	Sibuyan I		Romblon	Philippines	*R. rattus IV*	Yes		Yes
FMNH145705	Sibuyan I		Romblon	Philippines	*R. rattus IV*	Yes	Yes	Yes
FMNH145711	Sibuyan I		Romblon	Philippines	*R. rattus IV*	Yes		Yes
FMNH145714	Sibuyan I		Romblon	Philippines	*R. rattus IV*	Yes		Yes

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
FMNH147776	Siquijor I		Siquijor	Philippines	*R. tanezumi*	Yes	Yes	Yes
FMNH147778	Siquijor I		Siquijor	Philippines	*R. tanezumi*	Yes	Yes	Yes
FMNH147777	Siquijor I		Siquijor	Philippines	*R. rattus IV*	Yes	Yes	Yes
SP6156				South Africa	*R. rattus* I			
UAM85048	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85172	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85173	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85174	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85175	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85176	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85177	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85179	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85184	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85188	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85191	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
UAM85192	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85193	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85194	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85195	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85196	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85198	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85199	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85200	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85201	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM85202	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM86988	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM91776	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM97703	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
UAM97706	Shemya Is.	Aleutian Is.	Alaska	USA	*R. rattus* I	Yes		
OK6153	Little Rock	Pulaski	Arkansas	USA	*R. rattus* I	Yes		

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
OK6154	Little Rock	Pulaski	Arkansas	USA	*R. rattus* I	Yes		
OK6156	Little Rock	Pulaski	Arkansas	USA	*R. rattus* I	Yes		
OK6158	Little Rock	Pulaski	Arkansas	USA	*R. rattus* I	Yes		
OK6159	Little Rock	Pulaski	Arkansas	USA	*R. rattus* I	Yes		
OK6155	Little Rock	Pulaski	Arkansas	USA	*R. rattus* I	Yes		
TK160564		Alameda	California	USA	*R. rattus* I	Yes	Yes	Yes
TK160565		Alameda	California	USA	*R. rattus* I	Yes	Yes	Yes
TK160566		Alameda	California	USA	*R. rattus* I	Yes	Yes	Yes
TK160567		Alameda	California	USA	*R. rattus* I	Yes	Yes	Yes
TK160568		Alameda	California	USA	*R. rattus* I	Yes	Yes	Yes
TK160569		Alameda	California	USA	*R. rattus* I	Yes	Yes	Yes
TK160570		Alameda	California	USA	*R. rattus* I	Yes	Yes	Yes
TK160571		Alameda	California	USA	*R. rattus* I	Yes	Yes	Yes
TK160572		Alameda	California	USA	*R. rattus* I	Yes	Yes	Yes
TK160573		Alameda	California	USA	*R. rattus* I	Yes	Yes	Yes

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
MVZ216790		Alameda	California	USA	*R. tanezumi*	Yes	Yes	Yes
MVZ216791		Alameda	California	USA	*R. tanezumi*	Yes	Yes	Yes
MVZ220308		Alameda	California	USA	*R. rattus* I	Yes	Yes	Yes
MVZ220309		Alameda	California	USA	*R. rattus* I	Yes	Yes	Yes
MVZ220318		Alameda	California	USA	*R. rattus* I	Yes	Yes	Yes
MVZ208904		Contra Costa	California	USA	*R. rattus* I	Yes	Yes	Yes
MVZ208918		Contra Costa	California	USA	*R. rattus* I	Yes	Yes	Yes
MVZ219647		Contra Costa	California	USA	*R. tanezumi*	Yes	Yes	Yes
MVZ225257		Contra Costa	California	USA	*R. tanezumi*	Yes	Yes	Yes
MVZ225267		Contra Costa	California	USA	*R. tanezumi*	Yes	Yes	Yes
MVZ216799		Contra Costa	California	USA	*R. tanezumi*	Yes	Yes	Yes
TK113734	San Diego	San Diego	California	USA	*R. rattus* I	Yes		
TK93474	San Diego	San Diego	California	USA	*R. rattus* I	Yes		
TK113626	San Diego	San Diego	California	USA	*R. rattus* I	Yes		
MVZ226367			California	USA	*R. tanezumi*	Yes	Yes	Yes

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
MVZ226369			California	USA	*R. rattus* I	Yes	Yes	Yes
MVZ226384			California	USA	*R. rattus* I	Yes	Yes	Yes
MVZ226380			California	USA	*R. rattus* I	Yes	Yes	Yes
MVZ226366			California	USA	*R. rattus* I	Yes	Yes	Yes
MVZ226385			California	USA	*R. rattus* I	Yes	Yes	Yes
MVZ226379			California	USA	*R. tanezumi*	Yes	Yes	Yes
MVZ226381			California	USA	*R. rattus* I	Yes	Yes	Yes
MVZ226382			California	USA	*R. rattus* I	Yes	Yes	Yes
MVZ226386			California	USA	*R. tanezumi*	Yes	Yes	Yes
OK11855	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes		
OK11867	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes		
OK11874	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes	Yes	Yes
OK11851	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes	Yes	Yes
OK11852	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes	Yes	Yes
OK11853	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes	Yes	Yes

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
OK11854	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes	Yes	Yes
OK11856	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes	Yes	Yes
OK11857	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes	Yes	Yes
OK11858	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes	Yes	Yes
OK11860	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes	Yes	Yes
OK11862	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes	Yes	Yes
OK11863	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes		
OK11864	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes		
OK11865	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes		
OK11866	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes		
OK11868	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes		
OK11869	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes		
OK11871	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes		
OK11878	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes		
OK11879	Gainesville	Alachua	Florida	USA	*R. rattus* I	Yes		

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
RrFL5 (E)	Panama City	Bay	Florida	USA	*R. rattus* I	Yes	Yes	Yes
RrFL1 (A)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes	Yes	Yes
RrFL2 (B)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes	Yes	Yes
RrFL3 (C)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes	Yes	Yes
RrFL4 (D)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes	Yes	Yes
RrFL8 (H)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes	Yes	Yes
OK11918 (Florida 1)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes		
OK11928 (Florida 2)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes		
OK11891 (Florida 3)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes		
OK11923 (Florida 4)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes		
OK11916 (Florida 5)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes		
OK11921 (Florida 6)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes		

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
OK11926 (Florida 7)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes		
OK11932 (Florida 8)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes		
OK11935 (Florida 9)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes		
OK11924 (Florida 10)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes		
OK11931 (Florida 11)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes		
OK11912 (Florida 12)	Panama City	Bay	Florida	USA	*R. rattus* I	Yes	Yes	Yes
OK11914 (Florida 13)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes	Yes	Yes
OK11908 (Florida 14)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes	Yes	Yes
OK11894 (Florida 15)	Panama City	Bay	Florida	USA	*R. rattus* I	Yes	Yes	Yes
OK11903 (Florida 16)	Panama City	Bay	Florida	USA	*R. rattus* I	Yes	Yes	Yes

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
OK11881 (Florida 18)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes	Yes	Yes
OK11906 (Florida 19)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes	Yes	Yes
OK11897 (Florida 21)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes	Yes	Yes
OK11888 (Florida 22)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes	Yes	Yes
OK11900 (Florida 24)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes	Yes	Yes
OK11904 (Florida 26)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes	Yes	Yes
OK11890 (Florida 28)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes	Yes	Yes
OK11901 (Florida 29)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes	Yes	Yes
OK11907 (Florida 30)	Panama City	Bay	Florida	USA	*R. tanezumi*	Yes	Yes	Yes
MC2A		Martin	Florida	USA	*R. rattus* I	Yes		

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
MC2B		Martin	Florida	USA	*R. rattus* I	Yes		
MC2C		Martin	Florida	USA	*R. rattus* I	Yes		
RrFL11	Key Largo	Monroe	Florida	USA	*R. rattus* I	Yes	Yes	Yes
RrFL12	Key Largo	Monroe	Florida	USA	*R. rattus* I	Yes	Yes	Yes
RrFL9	Key Largo	Monroe	Florida	USA	*R. rattus* I	Yes	Yes	Yes
RrFL10	Key Largo	Monroe	Florida	USA	*R. rattus* I	Yes	Yes	Yes
RrFL13	Key Largo	Monroe	Florida	USA	*R. rattus* I	Yes	Yes	Yes
OK11884 (Florida 17)	Key Largo	Monroe	Florida	USA	*R. rattus* I	Yes	Yes	Yes
OK11882 (Florida 20)	Key Largo	Monroe	Florida	USA	*R. rattus* I	Yes	Yes	Yes
OK11893 (Florida 23)	Key Largo	Monroe	Florida	USA	*R. rattus* I	Yes	Yes	Yes
OK11886 (Florida 25)	Key Largo	Monroe	Florida	USA	*R. rattus* I	Yes	Yes	Yes
OK11899 (Florida 27)	Key Largo	Monroe	Florida	USA	*R. rattus* I	Yes	Yes	Yes
M-2990	Baton Rouge	Baton Rouge	Louisiana	USA	*R. rattus* I	Yes		

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
		Parish						
M-3418	Baton Rouge	Baton Rouge Parish	Louisiana	USA	*R. rattus* I	Yes		
M-3419	Baton Rouge	Baton Rouge Parish	Louisiana	USA	*R. rattus* I	Yes		
M-3436	Baton Rouge	Baton Rouge Parish	Louisiana	USA	*R. rattus* I	Yes		
M-3437	Baton Rouge	Baton Rouge Parish	Louisiana	USA	*R. rattus* I	Yes		
M-3438	Baton Rouge	Baton Rouge Parish	Louisiana	USA	*R. rattus* I	Yes		
M-3439	Baton Rouge	Baton Rouge Parish	Louisiana	USA	*R. rattus* I	Yes		
M-3440	Baton Rouge	Baton Rouge Parish	Louisiana	USA	*R. rattus* I	Yes		
M-1346	Baton Rouge	Baton Rouge Parish	Louisiana	USA	*R. rattus* I	Yes		
M-1928		Lafourche Parish	Louisiana	USA	*R. rattus* I	Yes		

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
M-2061		Terrebonne Parish	Louisiana	USA	*R. rattus* I	Yes		
M-3420			Louisiana	USA	*R. rattus* I	Yes		
M-1929		Jefferson Davis	Mississippi	USA	*R. rattus* I	Yes		
ASK3695	Brownsville	Cameron	Texas	USA	*R. rattus* I			
OK11861	Houston	Harris	Texas	USA	*R. rattus* I	Yes		
OK11872	Houston	Harris	Texas	USA	*R. rattus* I	Yes		
OK11873	Houston	Harris	Texas	USA	*R. rattus* I	Yes		
OK11875	Houston	Harris	Texas	USA	*R. rattus* I	Yes		
OK11876	Houston	Harris	Texas	USA	*R. rattus* I	Yes		
OK11877	Houston	Harris	Texas	USA	*R. rattus* I	Yes		
OK11880	Houston	Harris	Texas	USA	*R. rattus* I	Yes		
ASK5325	Weatherford	Parker	Texas	USA	*R. rattus* I			
ASK4763	San Angelo	Tom Green	Texas	USA	*R. rattus* I	Yes		
ASK6277	San Angelo	Tom Green	Texas	USA	*R. rattus* I	Yes		

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
ASK6283	San Angelo	Tom Green	Texas	USA	*R. rattus* I	Yes		
ASK6284	San Angelo	Tom Green	Texas	USA	*R. rattus* I	Yes		
ASK6286	San Angelo	Tom Green	Texas	USA	*R. rattus* I	Yes		
ASK6296	San Angelo	Tom Green	Texas	USA	*R. rattus* I	Yes		
ASK7213	San Angelo	Tom Green	Texas	USA	*R. rattus* I	Yes		
ASK7596	San Angelo	Tom Green	Texas	USA	*R. rattus* I	Yes		
ASK6367	San Angelo	Tom Green	Texas	USA	*R. rattus* I	Yes		
ASK3691	San Angelo	Tom Green	Texas	USA	*R. rattus* I	Yes		
ASK4753	San Angelo	Tom Green	Texas	USA	*R. rattus* I	Yes		
ASK5451	San Angelo	Tom Green	Texas	USA	*R. rattus* I	Yes		
ACU1400	San Angelo (Abilene)	Tom Green	Texas	USA	*R. rattus* I	Yes		
ACU868	San Angelo (Abilene)	Tom Green	Texas	USA	*R. rattus* I	Yes		
ASK6992	Austin	Travis	Texas	USA	*R. rattus* I	Yes		
ASK6993	Austin	Travis	Texas	USA	*R. rattus* I	Yes		

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
ASK6982	Austin	Travis	Texas	USA	*R. rattus* I	Yes		
ASK6986	Austin	Travis	Texas	USA	*R. rattus* I	Yes		
ASK6989	Austin	Travis	Texas	USA	*R. rattus* I	Yes		
ASK6990	Austin	Travis	Texas	USA	*R. rattus* I	Yes		
ASK6981	Austin	Travis	Texas	USA	*R. rattus* I	Yes		
ASK6983	Austin	Travis	Texas	USA	*R. rattus* I	Yes		
ASK6985	Austin	Travis	Texas	USA	*R. rattus* I	Yes		
ASK6991	Austin	Travis	Texas	USA	*R. rattus* I	Yes		
ASK6994	Austin	Travis	Texas	USA	*R. rattus* I	Yes		
ASK6887	Austin (Georgetown)	Williamson	Texas	USA	*R. rattus* I	Yes		
ASK6888	Austin (Georgetown)	Williamson	Texas	USA	*R. rattus* I	Yes		
MVZ186525	Vinh Yen District, Tam Dao		Vinh Phu	Vietnam	*R. tanezumi*	Yes	Yes	Yes
MVZ186526	Vinh Yen District, Tam Dao		Vinh Phu	Vietnam	*R. tanezumi*	Yes	Yes	Yes
MVZ186529	Vinh Yen District, Tam Dao		Vinh Phu	Vietnam	*R. tanezumi*	Yes	Yes	Yes

Name	City	County	State	Country	mtDNA Clade/Species	Microsatellites	DHFR	ATP5A1
MVZ186540	Vinh Yen	District, Tam Dao	Vinh Phu	Vietnam	*R. tanezumi*	Yes	Yes	Yes
MVZ186527	Vinh Yen	District, Tam Dao	Vinh Phu	Vietnam	*R. tanezumi*	Yes	Yes	Yes
MVZ186530	Vinh Yen	District, Tam Dao	Vinh Phu	Vietnam	*R. tanezumi*	Yes	Yes	Yes
MVZ186535	Vinh Yen	District, Tam Dao	Vinh Phu	Vietnam	*R. tanezumi*	Yes	Yes	Yes
MVZ186538	Vinh Yen	District, Tam Dao	Vinh Phu	Vietnam	*R. tanezumi*	Yes	Yes	Yes
MVZ186541	Vinh Yen	District, Tam Dao	Vinh Phu	Vietnam	*R. tanezumi*	Yes	Yes	Yes
MVZ186539	Vinh Yen	District, Tam Dao	Vinh Phu	Vietnam	*R. tanezumi*	Yes	Yes	Yes
MVZ186531	Vinh Yen	District, Tam Dao	Vinh Phu	Vietnam	*R. tanezumi*	Yes	Yes	
MVZ186532	Vinh Yen	District, Tam Dao	Vinh Phu	Vietnam	*R. tanezumi*	Yes	Yes	Yes
MVZ186536	Vinh Yen	District, Tam Dao	Vinh Phu	Vietnam	*R. tanezumi*	Yes		
MVZ186537	Vinh Yen	District, Tam Dao	Vinh Phu	Vietnam	*R. tanezumi*	Yes		
MVZ186528	Vinh Yen	District, Tam Dao	Vinh Phu	Vietnam	*R. tanezumi*	Yes		
MVZ186534	Vinh Yen	District, Tam Dao	Vinh Phu	Vietnam	*R. tanezumi*	Yes		
MVZ186524	Vinh Yen	District, Tam Dao	Vinh Phu	Vietnam	*R. tanezumi*	Yes		

Supplementary Table S4.2. Accession numbers for mitochondrial cytochrome b sequences obtained from GenBank, including the geographic locality and putative mtDNA species identification based on clade assignment in the cytochrome b phylogenetic analysis (Fig. 4.1A). Individuals occurring in unnamed clades are referred to as *Rattus* sp.

Name	City	State	Country	mtDNA Clade/Species
AB096841			Japan	*R. tanezumi*
AB211040	Kagoshima	Shibushi	Japan	*R. tanezumi*
AB211041	Miyazaki		Japan	*R. tanezumi*
AB211042	Kagoshima		Japan	*R. rattus* I
AB211043	Kagoshima		Japan	*R. rattus* I
AB355899	Haiphong port		Vietnam	*R. tanezumi*
AB355900	Haiphong port		Vietnam	*R. tanezumi*
AB355901	Haiphong port		Vietnam	*R. tanezumi*
AB355902	Haiphong port		Vietnam	*R. norvegicus*
AB355903	Haiphong port		Vietnam	*R. norvegicus*
AJ428514	Copenhagen		Denmark	*R. norvegicus*
AY263617			Indonesia	*R. rattus* I
NC014858			Australia	*R. lutreolus*
NC014861			Australia	*R. tunneyi*
NC014864			Australia	*R. villosissimus*
NC014867			Australia	*R. fuscipes*
NC014871			Australia	*R. sordidus*
NC014855			Australia	*R. leucopus*
NC001665			Lab Strain	*R. norvegicus*
NC011638			Japan	*R. tanezumi*
NC012374			New Zealand	*R. rattus* I

Name	City	State	Country	mtDNA Clade/Species
NC012389			Papua New Guinea	*R. exulans*
NC012461			New Guinea	*R. praetor*
DQ191485			Philippines	*R. everetti*
DQ191486			Philippines	*R. exulans*
DQ191487			New Guinea	*R. praetor*
DQ191488			Indonesia	*R. rattus* IV
DQ191486			Philippines	*R. exulans*
DQ191487			Philippines	*R. praetor*
DQ191488			Philippines	*R. rattus* IV
DQ439816			South Africa	*R. tanezumi*
DQ439817			South Africa	*R. tanezumi*
DQ439818			South Africa	*R. tanezumi*
DQ439819			South Africa	*R. tanezumi*
DQ439820			South Africa	*R. tanezumi*
DQ439821			South Africa	*R. tanezumi*
DQ439822			South Africa	*R. tanezumi*
DQ439823			South Africa	*R. tanezumi*
DQ439824			South Africa	*R. tanezumi*
DQ439825			South Africa	*R. tanezumi*
DQ439826			South Africa	*R. tanezumi*
DQ439827			South Africa	*R. tanezumi*
DQ439828			South Africa	*R. tanezumi*
DQ439829			South Africa	*R. tanezumi*
DQ439830			South Africa	*R. rattus* I

Name	City	State	Country	mtDNA Clade/Species
DQ439831			South Africa	*R. rattus* I
DQ439832			South Africa	*R. rattus* I
DQ439833			South Africa	*R. rattus* I
DQ439834			South Africa	*R. rattus* I
DQ439835			South Africa	*R. rattus* I
DQ439836			South Africa	*R. rattus* I
DQ439837			South Africa	*R. rattus* I
DQ439838			South Africa	*R. rattus* I
DQ439839			South Africa	*R. norvegicus*
DQ439840			South Africa	*R. norvegicus*
DQ439841			South Africa	*R. norvegicus*
DQ439842			South Africa	*R. norvegicus*
DQ439843			South Africa	*R. norvegicus*
DQ439844			South Africa	*R. norvegicus*
DQ439845			South Africa	*R. tanezumi*
DQ439846			South Africa	*R. tanezumi*
DQ439847			South Africa	*R. tanezumi*
DQ439848			South Africa	*R. tanezumi*
DQ439849			South Africa	*R. tanezumi*
DQ439850			South Africa	*R. tanezumi*
DQ439851			South Africa	*R. rattus* I
DQ439852			South Africa	*R. rattus* I
DQ439853			South Africa	*R. rattus* I
DQ439854			South Africa	*R. rattus* I

Name	City	State	Country	mtDNA Clade/Species
DQ439855			South Africa	*R. rattus* I
DQ439856			South Africa	*R. rattus* I
DQ439857			South Africa	*R. rattus* I
DQ439858			South Africa	*R. rattus* I
DQ439864			South Africa	*R. rattus* I
DQ673916	Milwaukee	Wisconsin	USA	*R. norvegicus*
DQ673917	Tokyo		Japan	*R. norvegicus*
EF186409			Malaysia	*R. rattus* IV
EF186410			Malaysia	*R. rattus* IV
EF186411			Malaysia	*R. rattus* IV
EF186412			Malaysia	*R. rattus* IV
EF186413			Malaysia	*R. rattus* IV
EF186414			Cook Islands (South Pacific)	*R. exulans*
EF186415			Cook Islands (South Pacific)	*R. exulans*
EF186416			Cook Islands (South Pacific)	*R. exulans*
EF186417			Fiji	*R. exulans*
EF186418		Hawaii	USA	*R. exulans*
EF186419		Hawaii	USA	*R. exulans*
EF186420			French Polynesia	*R. exulans*
EF186421			Indonesia	*R. exulans*
EF186422			French Polynesia	*R. exulans*
EF186423			French Polynesia	*R. exulans*
EF186424			New Zealand	*R. exulans*
EF186425			New Zealand	*R. exulans*

Name	City	State	Country	mtDNA Clade/Species
EF186426			New Zealand	*R. exulans*
EF186427			Papua New Guinea	*R. exulans*
EF186428			Papua New Guinea	*R. exulans*
EF186429			French Polynesia	*R. exulans*
EF186430			Samoa	*R. exulans*
EF186431			Samoa	*R. exulans*
EF186432			Thailand	*R. exulans*
EF186433			Thailand	*R. exulans*
EF186434			Thailand	*R. exulans*
EF186435			Australia	*R. fuscipes*
EF186436			Australia	*R. fuscipes*
EF186437			Australia	*R. fuscipes*
EF186438			Australia	*R. fuscipes*
EF186439			Australia	*R. fuscipes*
EF186440			China	*R. tanezumi*
EF186441			Indonesia	*R. hoffmanni*
EF186442			Indonesia	*R. hoffmanni*
EF186443			Indonesia	*R. hoffmanni*
EF186444			Sri Lanka	*R. rattus* IV
EF186445			Sri Lanka	*R. rattus* IV
EF186446			Sri Lanka	*R. rattus* IV
EF186447			Australia	*R. leucopus*
EF186448			Australia	*R. leucopus*
EF186449			Papua New Guinea	*Rattus* sp.

Name	City	State	Country	mtDNA Clade/Species
EF186450			Papua New Guinea	*Rattus* sp.
EF186451			Australia	*R. leucopus*
EF186452			Australia	*R. leucopus*
EF186453			Papua New Guinea	*Rattus* sp.
EF186454			Papua New Guinea	*R. novaeguineae*
EF186455			Papua New Guinea	*Rattus* sp.
EF186456			Papua New Guinea	*Rattus* sp.
EF186457			Papua New Guinea	*Rattus* sp.
EF186458			Papua New Guinea	*Rattus* sp.
EF186459			Papua New Guinea	*Rattus* sp.
EF186460			Papua New Guinea	*Rattus* sp.
EF186461			French Polynesia	*R. norvegicus*
EF186462			French Polynesia	*R. norvegicus*
EF186463			Papua New Guinea	*Rattus* sp.
EF186464			Papua New Guinea	*Rattus* sp.
EF186465			Papua New Guinea	*Rattus* sp.
EF186466			Papua New Guinea	*Rattus* sp.
EF186467			Papua New Guinea	*Rattus* sp.
EF186468			Papua New Guinea	*Rattus* sp.
EF186469			French Polynesia	*R. rattus* I
EF186470			New Zealand	*R. rattus* I
EF186471			Papua New Guinea	*Rattus* sp.
EF186472			Papua New Guinea	*R. rattus* I
EF186473			Papua New Guinea	*R. rattus* I

Name	City	State	Country	mtDNA Clade/Species
EF186474			French Polynesia	*R. rattus* I
EF186475			Samoa	*R. rattus* I
EF186476			Papua New Guinea	*Rattus* sp.
EF186477			Australia	*R. sordidus*
EF186478			Australia	*R. sordidus*
EF186479			Australia	*R. sordidus*
EF186480			Australia	*R. sordidus*
EF186481			Indonesia	*Rattus* sp.
EF186482			Indonesia	*Rattus* sp.
EF186483			Papua New Guinea	*Rattus* sp.
EF186484			Papua New Guinea	*Rattus* sp.
EF186485			Papua New Guinea	*Rattus* sp.
EF186486			Papua New Guinea	*Rattus* sp.
EF186487			Papua New Guinea	*Rattus* sp.
EF186488			Papua New Guinea	*Rattus* sp.
EF186489			Papua New Guinea	*Rattus* sp.
EF186490			Indonesia	*R. rattus* IV
EF186491			Indonesia	*R. tanezumi*
EF186492			Indonesia	*R. rattus IV*
EF186493			Indonesia	*R. tanezumi*
EF186494			Indonesia	*R. tanezumi*
EF186495			Indonesia	*R. tanezumi*
EF186496			Indonesia	*R. rattus* IV
EF186497			Indonesia	*R. rattus* IV

Name	City	State	Country	mtDNA Clade/Species
EF186498			Indonesia	*R. rattus* IV
EF186499			Indonesia	*R. rattus* IV
EF186500			Indonesia	*R. rattus* IV
EF186501			Indonesia	*R. rattus* IV
EF186502			Indonesia	*R. tanezumi*
EF186503			Indonesia	*R. tanezumi*
EF186504			Indonesia	*R. tanezumi*
EF186505			Indonesia	*R. tanezumi*
EF186506			Indonesia	*R. rattus* IV
EF186507			Indonesia	*R. tanezumi*
EF186508			Japan	*R. tanezumi*
EF186509			Japan	*R. tanezumi*
EF186510			Japan	*R. tanezumi*
EF186511			Indonesia	*R. rattus* IV
EF186512			Indonesia	*R. rattus* IV
EF186513			Indonesia	*R. tiomanicus*
EF186514			Indonesia	*R. tiomanicus*
EF186515			Australia	*R. tunneyi*
EF186516			Australia	*R. tunneyi*
EF186517			Australia	*R. tunneyi*
EF186518			Australia	*R. tunneyi*
EF186519			Papua New Guinea	*Rattus* sp.
EF186520			Papua New Guinea	*Rattus* sp.
EF186521			Papua New Guinea	*Rattus* sp.

Name	City	State	Country	mtDNA Clade/Species
EF186522			Papua New Guinea	*Rattus* sp.
EF186523			Papua New Guinea	*Rattus* sp.
EU349781			Australia	*R. leucopus*
EU349782			Lab Strain	*R. norvegicus*
EU349783			Australia	*R. villosissimus*
FJ842262			South Africa	*R. tanezumi*
FJ842263			South Africa	*R. tanezumi*
FJ842264			Swaziland	*R. tanezumi*
FJ842265			South Africa	*R. tanezumi*
FJ842266			South Africa	*R. rattus* I
FJ842267			South Africa	*R. rattus* I
FJ842268			South Africa	*R. rattus* I
GQ274948			Singapore	*R. rattus* IV
GQ274949			Singapore	*R. rattus* IV
GQ891569			India	*R. rattus* I
GQ891570			India	*R. rattus* I
GQ891571			India	*R. rattus* I
GQ891572			India	*R. rattus* I
GQ891573			India	*R. rattus* I
GQ891574			Oman	*R. rattus* I
GQ891575			Oman	*R. rattus* I
GQ891576			Oman	*R. rattus* I
GQ891577			Oman	*R. rattus* I
GQ891578			Oman	*R. rattus* I

Name	City	State	Country	mtDNA Clade/Species
GQ891579			Oman	*R. rattus* I
GQ891580			Oman	*R. rattus* I
GQ891581			Yemen	*R. rattus* I
GQ891582			Yemen	*R. rattus* I
GQ891583			Ethiopa	*R. rattus* I
GQ891584			Ethiopa	*R. rattus* I
GQ891585			Tanzania	*R. rattus* I
GQ891586			Tanzania	*R. rattus* I
GQ891587			Tanzania	*R. rattus* I
GQ891588			Mozambique	*R. rattus* I
GQ891589			Mozambique	*R. rattus* I
GQ891590			Mozambique	*R. rattus* I
GQ891591			Grand Comore	*R. rattus* I
GQ891592			Grand Comore	*R. rattus* I
GQ891593			Grand Comore	*R. rattus* I
GQ891594			Grand Comore	*R. rattus* I
GQ891595			Grand Comore	*R. rattus* I
GQ891596			Grand Comore	*R. rattus* I
GQ891597			Mayotte	*R. rattus* I
GQ891598			Mayotte	*R. rattus I*
GQ891599			Madagascar	*R. rattus* I
GQ891600			Madagascar	*R. rattus* I
GQ891601			Madagascar	*R. rattus* I
GQ891602			Madagascar	*R. rattus* I

Name	City	State	Country	mtDNA Clade/Species
GQ891603			Madagascar	*R. rattus* I
GQ891604			Madagascar	*R. rattus* I
GQ891605			Madagascar	*R. rattus* I
GQ891606			Madagascar	*R. rattus* I
GQ891607			Reunion	*R. rattus* I
GQ891608			South Africa	*R. rattus* I
GU570659			Australia	*R. leucopus*
GU570660			Papua New Guinea	*R. leucopus*
GU570661			Australia	*R. lutreolus*
GU570662			Australia	*R. tunneyi*
GU570663			Australia	*R. villosissimus*
GU570664			Australia	*R. fuscipes*
GU570665			Australia	*R. sordidus*
GU570671			Australia	*R. lutreolus*
GU570672			Australia	*R. lutreolus*
GU570673			Australia	*R. villosissimus*
GU570674			Australia	*R. villosissimus*
GU570675			Australia	*R. villosissimus*
GU592954			China	*R. norvegicus*
GU592955			China	*R. norvegicus*
GU592956			China	*R. norvegicus*
GU592957			China	*R. norvegicus*
GU592958			China	*R. norvegicus*
GU592959			China	*R. norvegicus*

Name	City	State	Country	mtDNA Clade/Species
GU592960			China	*R. norvegicus*
GU592961			China	*R. norvegicus*
GU592962			China	*R. norvegicus*
GU592963			China	*R. norvegicus*
GU592964			China	*R. norvegicus*
GU592965			China	*R. norvegicus*
GU592966			China	*R. norvegicus*
GU592967			China	*R. norvegicus*
GU592968			China	*R. norvegicus*
GU592969			China	*R. norvegicus*
GU592970			China	*R. norvegicus*
GU592971			China	*R. norvegicus*
GU592972			China	*R. norvegicus*
HM217362			Cambodia	*R. argiventer*
HM217363			Cambodia	*R. rattus* IV
HM217364			Cambodia	*R. argiventer*
HM217365			Tanzania	*R. rattus* I
HM217366			Oman	*R. rattus* I
HM217367			India	*R. rattus* I
HM217368			Madagascar	*R. rattus* I
HM217370			Thailand	*R. norvegicus*
HM217371			Thailand	*R. tanezumi*
HM217372			Thailand	*R. rattus* IV
HM217373			Thailand	*R. norvegicus*

Name	City	State	Country	mtDNA Clade/Species
HM217374			Thailand	*R. rattus* IV
HM217375			Thailand	*R. rattus* IV
HM217377			Thailand	*R. exulans*
HM217379			Thailand	*R. exulans*
HM217381			Thailand	*R. rattus* IV
HM217382			Thailand	*Rattus* sp.
HM217383			Thailand	*R. exulans*
HM217384			Thailand	*R. rattus* IV
HM217388			Thailand	*R. exulans*
HM217389			Thailand	*R. rattus* IV
HM217391			Thailand	*R. tiomanicus*
HM217392			Thailand	*R. exulans*
HM217393			Thailand	*R. rattus* IV
HM217394			Thailand	*R. rattus* IV
HM217395			Thailand	*R. exulans*
HM217396			Thailand	*R. andamanensis*
HM217397			Thailand	*R. rattus* IV
HM217398			Thailand	*R. tanezumi*
HM217399			Thailand	*R. rattus* IV
HM217403			Thailand	*R. andamanensis*
HM217407			Thailand	*R. tanezumi*
HM217410			Thailand	*R. tanezumi*
HM217411			Thailand	*R. exulans*

Name	City	State	Country	mtDNA Clade/Species
HM217421			Thailand	*R. rattus* IV
HM217423			Thailand	*R. rattus* IV
HM217424			Thailand	*R. exulans*
HM217426			Thailand	*R. tanezumi*
HM217428			Thailand	*R. exulans*
HM217429			Thailand	*R. norvegicus*
HM217430			Thailand	*R. tanezumi*
HM217436			Thailand	*R. tanezumi*
HM217437			Thailand	*R. exulans*
HM217438			Thailand	*R. tanezumi*
HM217440			Thailand	*R. exulans*
HM217441			Thailand	*R. exulans*
HM217442			Thailand	*R. rattus* IV
HM217443			Thailand	*R. rattus* IV
HM217446			Thailand	*R. rattus* IV
HM217452			Thailand	*R. tanezumi*
HM217454			Thailand	*R. rattus* IV
HM217456			Thailand	*R. tanezumi*
HM217457			Thailand	*R. tanezumi*
HM217458			Thailand	*R. tanezumi*
HM217466			Thailand	*R. tanezumi*
HM217467			Thailand	*R. tanezumi*
HM217470			Thailand	*R. exulans*
HM217472			Thailand	*R. exulans*

Name	City	State	Country	mtDNA Clade/Species
HM217474			Laos	*R. nitidus*
HM217475			Laos	*R. tanezumi*
HM217478			Laos	*R. nitidus*
HM217479			Laos	*R. nitidus*
HM217480			Laos	*R. tanezumi*
HM217481			Cambodia	*R. norvegicus*
HM222710			???	*R. norvegicus*
HQ157800			South Africa	*R. rattus* I
HQ157801			South Africa	*R. rattus* I
HQ157802			South Africa	*R. rattus* I
HQ157803			South Africa	*R. rattus* I
HQ157804			South Africa	*R. tanezumi*
HQ157805			South Africa	*R. tanezumi*
HQ157806			South Africa	*R. rattus* I
HQ157807			South Africa	*R. tanezumi*
HQ157808			South Africa	*R. rattus* I
HQ157809			South Africa	*R. rattus* I
J01436			Lab Strain	*R. norvegicus*

VITA

Justin B. Lack

Candidate for the Degree of

Doctor of Philosophy

Thesis: POPULATION GENETIC ANALYSIS OF INVASIVE *RATTUS*: IMPLICATIONS FOR EVOLUTIONARY BIOLOGY, DISEASE ECOLOGY AND INVASION BIOLOGY

Major Field: Type Field Here

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Zoology at Oklahoma State University, Stillwater, Oklahoma in May, 2012.

Completed the requirements for the Master of Science in Biology at The University of Central Oklahoma, Edmond, Oklahoma in 2007.

Completed the requirements for the Bachelor of Science in Biology at The University of Central Oklahoma, Edmond, OK in 2004.

Experience: Teaching Assistant: General Biology (1 semester), Microbiology (1 semester), Physiology (10 semesters). Research: Graduate Research Assistant (2 semesters, 5 summers)

Professional Memberships: Oklahoma Academy of Science, 2005-present; Central Plains Society of Mammalogists, 2005-present; Southwestern Association of Naturalists, 2006-present; American Society of Mammalogists, 2006-present; American Association for the Advancement of Science (AAAS), 2007-present; Texas Society of Mammalogists, 2008-present; Society for the Study of Evolution, 2009-present; Society for Molecular Biology and Evolution, 2009-present.

Name: Justin B. Lack Date of Degree: May, 2012

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of Study: POPULATION GENETIC ANALYSIS OF INVASIVE *RATTUS*: IMPLICATIONS FOR EVOLUTIONARY BIOLOGY, DISEASE ECOLOGY AND INVASION BIOLOGY

Pages in Study: 201 Candidate for the Degree of Doctor of Philosophy

Major Field: Zoology

Scope and Method of Study: I utilized population genetic analyses to examine the colonization history and contemporary dispersal patterns of invasive *Rattus* in the U.S., as well as identifying any evolutionary impacts of these invasions (i.e., hybridization). In addition, I used reverse-transcription PCR to examine whether invasive *Rattus* were competent hosts for zoonotic hepatitis E virus in the U.S., where the source of infections has gone largely unidentified.

Findings and Conclusions: In terms of colonization history, I found that *R. rattus* and *R. norvegicus* were characterized by distinct patterns of colonization, with *R. rattus* colonizing from a single maternal lineage and *R. norvegicus* colonizing from at least four maternal lineages. In addition, *R. rattus* do not appear to be establishing in the U.S. at a high rate, nor do they appear to be exhibiting a high frequency of long-distance dispersal. In contrast, *R. norvegicus* appears to be establishing and dispersing long distances at a very high frequency. In terms of evolutionary impacts, I found that extensive hybridization with introgression is occurring among several black rat species, and introgression is leading to widespread genomic swamping of *R. tanezumi* by two other species. Finally, I found conclusive evidence that invasive *R. rattus* and *R. norvegicus* are capable of carrying the zoonotype genotype 3 of the hepatitis E virus within the U.S., laying the groundwork for future studies investigating their role in human infection.

ADVISER'S APPROVAL: Ronald A. Van Den Bussche