

Urban population genetics of slum-dwelling rats (*Rattus norvegicus*) in Salvador, Brazil

BRITTNEY KAJDACSI,* FEDERICO COSTA,†‡ CHAZ HYSENI,* FLEUR PORTER,†
JULIA BROWN,* GORETE RODRIGUES,§ HELENA FARIAS,§ MITERMAYER G. REIS,‡
JAMES E. CHILDS,† ALBERT I. KO†‡ and ADALGISA CACCONE*

*Department of Ecology and Evolutionary Biology, Yale University, 165 Prospect Street, New Haven, CT 06520-8106, USA,

†Department of Epidemiology of Microbial Diseases, Yale School of Public Health, 60 College Street, New Haven, CT 06520-8034, USA, ‡Centro de Pesquisas Goçalo Moniz, Fundação Oswaldo Cruz, Ministério da Saúde, Rua Waldemar Falcão, 121,

Candeal, CEP: 40296-710, Salvador, Brazil, §Centro de Controle de Zoonoses, Secretaria Municipal de Saúde, Ministério da Saúde, Rua do Mocambo, S/N, Trobogi, CEP: 41720010, Salvador, Brazil

Abstract

Throughout the developing world, urban centres with sprawling slum settlements are rapidly expanding and invading previously forested ecosystems. Slum communities are characterized by untended refuse, open sewers and overgrown vegetation, which promote rodent infestation. Norway rats (*Rattus norvegicus*) are reservoirs for epidemic transmission of many zoonotic pathogens of public health importance. Understanding the population ecology of *R. norvegicus* is essential to formulate effective rodent control strategies, as this knowledge aids estimation of the temporal stability and spatial connectivity of populations. We screened for genetic variation, characterized the population genetic structure and evaluated the extent and patterns of gene flow in the urban landscape using 17 microsatellite loci in 146 rats from nine sites in the city of Salvador, Brazil. These sites were divided between three neighbourhoods within the city spaced an average of 2.7 km apart. Surprisingly, we detected very little relatedness among animals trapped at the same site and found high levels of genetic diversity, as well as structuring across small geographical distances. Most F_{ST} comparisons among sites were statistically significant, including sites <400 m apart. Bayesian analyses grouped the samples in three genetic clusters, each associated with distinct sampling sites from different neighbourhoods or valleys within neighbourhoods. These data indicate the existence of complex genetic structure in *R. norvegicus* in Salvador, linked to the heterogeneous urban landscape. Future rodent control measures need to take into account the spatial and temporal linkage of rat populations in Salvador, as revealed by genetic data, to develop informed eradication strategies.

Keywords: population genetics, *Rattus norvegicus*, rodent control, urban ecology

Received 15 May 2013; revision received 7 July 2013; accepted 8 July 2013

Introduction

Norway rats (*Rattus norvegicus*) are geographically widespread, reaching maximal population densities in urban centres where abundant food resources and harborage are present (Glass *et al.* 1989). This commensal

species is responsible for damage to the environment, infrastructure and agriculture (Macdonald *et al.* 1999; Singleton *et al.* 1999; Howald *et al.* 2007), causing significant financial costs across different ecological settings. In addition to their impact as a pest, Norway rats are reservoir hosts for the transmission many infectious diseases (Webster & Macdonald 1995; Himsworth *et al.* 2013), including leptospirosis (Ko *et al.* 2009), bartonellosis (Kosoy *et al.* 2010), bubonic plague (Chanteau *et al.* 1998), murine typhus (Civen & Ngo 2008) and salmonellosis (Nkogwe *et al.* 2011).

Correspondence: Brittney Kajdacsi, Department of Biological Sciences, Louis Calder Center-Biological Field Station, Fordham University, 53 Whipponwill Road, Armonk, NY 10504, Fax: +1 914 273 2167; E-mail: brittney.kajdacsi@aya.yale.edu

The dramatic growth of urban slum settlements, most of which has occurred in tropical regions of the world in resource poor countries over the past 50 years, has transformed the urban ecology creating new habitats for Norway rats. To date, more than one billion of the world's inhabitants reside in slum settlements or shantytowns. This population will double by 2025 (UN-HABITAT 2010). In slum communities, lack of access to services and poor housing and sanitation promote infestation by *R. norvegicus* (Glass *et al.* 1989, 2005; Childs *et al.* 1991), promoting conditions for epidemic transmission of infectious diseases in humans (Ko *et al.* 1999). Leptospirosis, a spirochetal infection, which causes severe life-threatening disease in humans, has emerged as a cause of large epidemics in urban slum communities throughout the developing world (Ko *et al.* 1999, 2009; Sarkar *et al.* 2002; Amilasan *et al.* 2012; Dechet *et al.* 2012). In these outbreaks, *R. norvegicus*, as well as *R. rattus*, are the main reservoirs for human transmission (Pereira *et al.* 2000; de Faria *et al.* 2008; Matthias *et al.* 2008).

Despite the ecological and public health impacts of Norway rats in shantytowns worldwide, there is a paucity of data describing the population dynamics of Norway rats in urban landscapes. Ecological studies evaluating rodent demography have been limited to temperate urban environments of the Northern Hemisphere, most notably in Baltimore, Maryland, USA, where rodent ecological research has been continually conducted for over a half-century (Davis *et al.* 1948; Davis 1953; Orgain & Schein 1953; Calhoun 1962; Farhang-Azad & Southwick 1979; Glass *et al.* 1989, 2005; Childs *et al.* 1991). To date, only one study has used a population genetic approach on urban *R. norvegicus* populations (Gardner-Santana *et al.* 2009), while a handful of others have looked at island invasion dynamics, mostly in *R. rattus* (Calmet *et al.* 2001; Abdelkrim *et al.* 2005; Russell *et al.* 2009b, 2010). It is unclear how applicable results from studies of rat population dynamics in temperate cities in developed countries will be to populations inhabiting urban slums, especially those situated in tropical environments.

Microsatellite loci have been used to study disease vector movements and help define eradication units consisting of discrete populations (Slotman *et al.* 2006; Gilabert *et al.* 2007; McCoy 2008; Solano *et al.* 2010; Beadell *et al.* 2010; Brown *et al.* 2011) for vectors such as insect and rodent reservoirs of zoonotic pathogens. Previous studies involving *R. norvegicus* in temperate environments indicate that microsatellite loci provide a valuable tool for assessing population genetic structure at spatial scales of several kilometres (Robertson & Gemmell 2004; Abdelkrim *et al.* 2007, 2010; Russell *et al.* 2009b). However, microsatellite-based approaches have

not been applied to investigate *R. norvegicus* populations at the finer spatial resolution required to elucidate genetic structure and gene flow to inform delineation of eradication units for this reservoir host.

Herein, we present the findings of a study, which investigated the population genetic structure of *R. norvegicus* from slum settlements (*favelas*) within Salvador, situated in northeast Brazil. This urban centre has undergone recent rapid growth and expansion, which has led to a approximately 280% increase in its population since 1970 (Brazilian Institute of Geography & Statistics 1971) and establishment of densely populated slum settlements in formerly uninhabited Atlantic forested regions of the city (Reis *et al.* 2008). At present, 58% of the population of Salvador resides in *favelas* (Moreira & Pereira 2008). Large outbreaks of leptospirosis occur in the same *favela* communities each year during seasonal periods of heavy rainfall (Ko *et al.* 1999; Sarkar *et al.* 2002; Reis *et al.* 2008). In Salvador, *R. norvegicus* is the principal reservoir for leptospirosis (Ko *et al.* 1999; Pereira *et al.* 2000; Sarkar *et al.* 2002; de Faria *et al.* 2008; Reis *et al.* 2008), as in all major urban centres in Brazil (Pereira & Andrade 1988; Pereira *et al.* 2000).

We captured Norway rats from nine sites located at the base of the valleys of three different neighbourhoods, each of which had a documented history of ongoing rodent infestations and moderate-to-high annual incidence of leptospirosis (Fig. 1). Because most sites were only a few hundred metres apart, we had the opportunity to examine the patterns of population genetic structure at a very fine spatial resolution. Our goal was to identify a panel of microsatellite loci, which would generate robust inferences at microspatial scales, and use them to quantify the genetic variation at each site, evaluate the patterns of genetic differentiation within and between sites and assess the extent of genetic admixture between sites. Because *R. norvegicus* dispersal and recolonization following eradication helps sustain population abundance over time, understanding these dynamics will inform rodent control efforts by elucidating the spatial and temporal scope of eradication units and identifying potential barriers or corridors for dispersal.

Materials and methods

Study area and sampling

Salvador is a Brazilian city of about 2.7 million residents (Brazilian Institute of Geography & Statistics 2011) spread over 313 km² (12°55'34" southern latitude and 38°31'12" western longitude; Fig. 1). The city has a subtropical climate characterized by average annual temperatures (the mean of the 365 days of each year)

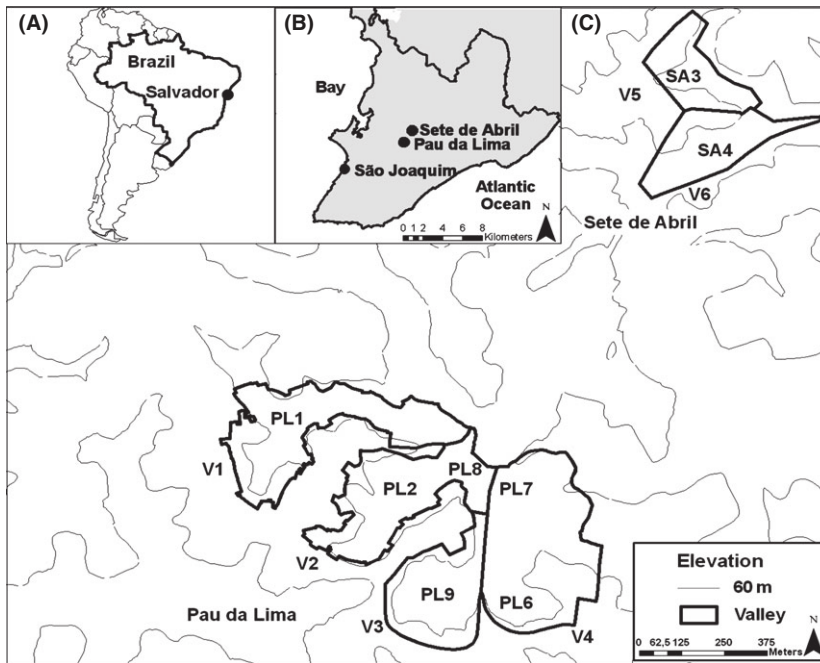


Fig. 1 Map of Salvador, Brazil, showing rat sampling locations. (A) Map of South America identifying Brazil and the city of Salvador. (B) Boundaries of Salvador city (dark outline) and neighbourhoods where rodent capture was performed (dark dots). (C) Topographical map of Sete de Abril and Pau da Lima valleys (black outlines, named as V1, V2, V3, V4 and V5) and rodent trapping sites (PL1, PL2, PL8, PL6-9, SA3 and SA4) in 2010 and 2011.

ranging from 21 to 37 °C (mean 25.3 °C) and heavy rainfall during winter. The three neighbourhoods selected for study included two residential slum areas, Pau da Lima and Sete de Abril, with documented high annual incidence of leptospirosis cases (in 2010, 10.4 and 23.3 per 100 000 people, respectively) (F. Costa, *personal communication*). The third neighbourhood, São Joaquim, although experiencing lower incidence of leptospirosis, was included to broaden the spatial extent of the study (Fig. 1), because of its location near the city port offers the potential sampling of rats originating from different geographical regions. The two residential slum communities include a series of valleys (mean valley depth, 60 m), where the highest human population density and worst slum conditions are concentrated below 30 m from the valley floor.

Six sampling sites (PL1, PL2, PL6, PL7, PL8 and PL9) were located at the base of four valleys (V1–V4) in Pau da Lima, and two additional sites (SA3 and SA4) in two valleys (V5–V6) were located in Sete de Abril (Table 1, Fig. 1). The distances between sampling sites averaged 2.7 km, although most were only 300 m–1.7 km apart (Table S1, Supporting information). Intersite distances fall within the documented range of *R. norvegicus* migration/dispersal distances of for both urban (25–150 m, Davis *et al.* 1948; Glass *et al.* 1989; Traweger *et al.* 2006) and rural settings (260–2000 m, Taylor 1978; Taylor & Quay 1979; Macdonald & Fenn 1995). The ninth sampling site (SJ5) was located at a local market area in a valley (V7) located in São Joaquim.

In both 2009 and 2010 between January and April, the months immediately preceding the wet season

(May–July), the annual peak of leptospirosis cases, rodenticide control was implemented in the four Pau da Lima valleys (V1–V4), which include six of the sampling sites in this study (PL1, PL2, PL6, PL7, PL8 and PL9). In these sites, *R. norvegicus* was by far the predominant rodent, infesting 85% of rodent-infested households. Approximately 70 000 houses were treated with both chemical intervention (second-generation anticoagulants as brodifacoum, bromadiolone and coumatetralyl) and environmental and educational activities to eliminate possible rat habitats and make the resident community aware of how to reduce exposure to the pathogen. Following this rodenticide control, there was a decrease in the proportion of rodent-infested houses in both years, from 25.3% to 7.8% (2009) and from 23.2% to 14.6% (2010; Costa, *personal communication*).

Sampling was carried out in June, July and August of 2010 and 2011, using either live Tomahawk traps (2010) or snap traps (2011). Three to five traps were placed at each of six to eight contiguous households defining an individual site. Each site contained multiple, separate burrow systems, presumably containing colonies of rats of varying kinship (Calhoun 1962). Traps were set before sundown and collected at sunrise the following morning. Because rats are known carriers of Seoul virus (genus *Hantavirus*), all procedures were conducted outdoors using approved biosafety measures (Mills *et al.* 1995). After double bagging and transport to an outdoor processing area, live-trapped rats were euthanized with thiopental solution, then weighed, sexed and 1 cm of tail tissue was removed. In total, 150

Table 1 Number of individuals (*N*) sampled per year at each collection site. Refer to Fig. 1 for site locations

Neighbourhood	Site	Valley	<i>N</i> (2010)	<i>N</i> (2011)	<i>N</i> (Total)	Male/Female	Age class composition (1/2/3)
Pau da Lima	PL1	V1	15	12	27	14/13	8/14/5
Pau da Lima	PL2	V2	15	15	30	16/14	7/20/3
Pau da Lima	PL6	V4	3	15	18	10/8	9/7/2
Pau da Lima	PL7	V4	3	—	3	2/1	—/3/—
Pau da Lima	PL8	V2	7	16	23	13/10	7/9/7
Pau da Lima	PL9	V3	—	11	11	7/4	7/4/—
Sete de Abril	SA3	V5	32	—	32	14/18	10/17/5
Sete de Abril	SA4	V6	3	—	3	2/1	1/—/2
São Joaquim	SJ5	V7	3	—	3	2/*	—/2/1
Total	—	—	81	69	150	—	—

*The sex of one individual from SJ5(V7) is unknown.

N refers to sample size for each site (year of collection in parenthesis).

The last two columns list for each sample the number of males and female and their age class (1 = juveniles; 2 = sexually mature young adult and 3 = large adult).

tail tips were collected and stored at -80°C until tested. Table 1 lists the samples obtained by year and site together with capture effort, sex and age class. Body mass was used as a proxy for age, and rats were classified into in three groups (1 ≤ 200 g, juveniles, non-breeding; 2 = 201–400 g, young adults, breeding; 3 ≥ 400 g, adults, breeding and senescent). Body mass has been shown to be a reliable indicator of age and reproductive maturity (Farhang-Azad & Southwick 1979; Glass *et al.* 1989; Iskjaer *et al.* 1989; F.H. Porter, F. Costa, G. Rodrigue, H. Farias, G. Glass, G. R. Mite-mayer, A. I. Ko & J. Childs, In review).

Genetic methods

DNA was extracted from 1 to 3 mm of tail tips using the *prepGEM* Tissue kit (ZyGEM Corp Ltd, Hamilton, New Zealand) according to the manufacturer's protocol.

Following a literature review of studies utilizing microsatellites in *Rattus*, 30 microsatellite loci were selected based on high allelic richness, absence of null alleles, lack of deviation from Hardy–Weinberg equilibrium (HWE) and no evidence of linkage to other loci in the set. All microsatellite primer sequences and references are given in the supplementary material (Table S2, Supporting information). A touchdown PCR (5 min of initial denaturation at 95°C ; 10 cycles of denaturing at 95°C , annealing at temperatures from 62°C to 53°C , decreasing by 1°C per cycle and extension at 72°C ; 40 cycles of 95°C , 52°C and 72°C ; and a final 20-min. extension at 72°C) in $13.5\ \mu\text{L}$ reaction mixture containing $0.4\ \mu\text{M}$ fluorescently labelled forward primer (6-FAM and HEX, Sigma-Aldrich), $0.4\ \mu\text{M}$ reverse primer (Eurofins Mwg/Operon), $1 \times$ buffer (Promega), $0.8\ \text{mM}$

dNTPs (NEB), $2.0\ \text{mM}$ MgCl_2 (Promega), $0.2\ \text{mg/mL}$ BSA (NEB) and $0.5\ \text{U}$ GoTaq polymerase (Promega) was used on a subset of individuals (12–14) to optimize amplification and assess the individual locus contribution out of the full complement of 30 loci to population genetic variation. PCR amplicons were analysed on an ABI 3730 \times 1 automated sequencer. PCR products were loaded in groups of two loci with nonoverlapping size ranges with a GS600 LIZ size standard. Alleles were scored using the program GENEMARKER version 1.95 (Soft-Genetics, State College, PA) with manual editing of the automatically scored peaks.

Of the 30 microsatellite loci tested, 17 were selected to screen all individuals based on high variability (4–7 alleles/locus) and reliable amplification. All 17 microsatellite loci chosen for analysis were tested for within-population deviations from HWE, and all pairs of loci across populations were tested for linkage disequilibrium (LD) using the software ARLEQUIN 3.5 (Excoffier & Lischer 2010). Markov chain parameters were set at 10 000 dememorizations, 1000 batches and 10 000 iterations per batch for both HWE and LD. Significance values were adjusted using the sequential Bonferroni method (Holm 1979).

Genetic diversity and population structure

ARLEQUIN 3.5 (Excoffier & Lischer 2010) was used to calculate summary statistics including observed (H_o) and expected (H_e) heterozygosities and allele frequencies for all loci in all populations. Allelic richness (A_r) was calculated with rarefaction correction for sample size in the program HP-Rare version 1.0 (Kalinowski 2004, 2005), and multilocus averages of allelic richness and private

(unique to a population) allelic richness were used to examine genotypic diversity differences between sites.

Genetic differentiation among populations was assessed by pairwise F_{ST} in the program ARLEQUIN 3.5 (Excoffier & Lischer 2010). The significance of these F_{ST} values was tested with 10 000 permutations. To determine which hierarchical level of population structure (neighbourhood, valley, site and nested combinations thereof) explained the largest significant proportion of variance in allele frequencies, we performed an analysis of molecular variance (AMOVA) in ARLEQUIN 3.5 (Excoffier & Lischer 2010). The AMOVA was run using four different groupings of individuals: (i) by site; (ii) by valley; (iii) by site within neighbourhood; and (iv) by valley within a neighbourhood. Prior to grouping samples from the same site from two different years, we used AMOVA and pairwise F_{ST} to test for significant temporal differentiation between samples originating from rats captured at the same site. Although little is known about the generation times of Norway rats in tropical climates, as many as 4–8 generations could have passed between the two samplings in 2010 and 2011, as Norway rats have several litters per year (Davis 1953).

Using the Bayesian clustering method implemented in STRUCTURE version 2.3 (Pritchard *et al.* 2000), patterns of population structure and individual assignment to sampling localities were tested by identifying genetic clusters without using a priori sampling information (Appendix S1, Supporting information). To identify the most likely number of clusters (K), three independent runs for each $K = 1–15$ were carried out. For all runs, the admixture model with independent allele frequencies was used with a burn-in value of 10 000 steps followed by 100 000 replications. The most likely K was determined by following the guidelines of Pritchard *et al.* 2000; and using the second-order rate of change in the likelihood distribution (ΔK , Evanno *et al.* 2005). Discriminant analysis of principal components (DAPC, Jombart *et al.* 2010) was also used to assess genetic clustering. DAPC is a two-step multivariate ordination analysis that makes no assumptions about deviations from HWE and LD, unlike the Bayesian approach implemented in STRUCTURE, which assumes HWE and linkage equilibrium (Appendix S1, Supporting information).

Furthermore, because many trap sites were separated by distances compatible with Norway rat migration capabilities, patterns of genetic differentiation were examined using a Mantel test of isolation by distance (IBD) implemented in Genepop (Raymond & Rousset 1995; Rousset 2008), comparing the linearized pairwise F_{ST} values with the Euclidean distances between capture sites using 10 000 permutations. Due to the noncontiguous nature of our sampling, analyses were run using only the six sites within the Pau da Lima neighbourhood.

Relatedness, dispersal and migration

To determine whether relatedness of individuals within sampling sites influenced genetic divergence patterns between those sites, we estimated relatedness through maximum-likelihood methods and examined relationships for all individual pairs using ML-Relate (Kalinowski *et al.* 2006), running the analyses only for sites with six or more individuals. Additionally, Bayesian approaches implemented in STRUCTURE (Pritchard *et al.* 2000) and GENECLASS 2.0 (Piry *et al.* 2004) were used to identify first-generation migrants (individuals putatively born at a site different from their capture site). We reran STRUCTURE (Pritchard *et al.* 2000) with the USEPOPINFO option (i.e. prior sample site information) for each individual included in the analysis. Burn-in and run lengths were the same as above, and migration rate was set at 0.05. STRUCTURE detects migrants by quantifying individual membership to a site and cluster by calculating Q values, the proportion of that individual's ancestry from a population, for each site in each individual. If an individual is a putative migrant, the Q value for its site of origin/capture (Q_o) is significantly lower than the Q value for its site of assignment (Q_a). In GeneClass 2.0, we utilized the 'detect migrants' function, which, in combination with Bayesian resampling methods, calculates likelihood-based statistics, including L_h , the likelihood of finding an individual in the population in which it was trapped, and L_{max} , the maximum likelihood of assignment to any sampled population. We used both L_h and L_h/L_{max} to detect first-generation migrants. L_h is the most appropriate statistic to use when all potential source populations have not been sampled. L_h lacks power, however, compared with other estimators, so we also used the L_h/L_{max} ratio, despite it being more informative when all source populations have been sampled (Paetkau *et al.* 2004). Because migrants from unsampled populations can be misclassified as residents, we selected the Rannala and Mountain (1997) criterion with the resampling method of Paetkau *et al.* (2004) to determine the critical value of L_h and L_h/L_{max} , using 1000 simulated individuals and the default $\alpha = 0.01$.

We used FSTAT 2.9.3 (Goudet 1995; Goudet *et al.* 2002) to test whether the observed population structure could be attributed to differences in dispersal between the sexes and calculated four different statistics and their significance values: F_{ST} , mAIC and vAIC (the corrected mean assignment index and its variance, Favre *et al.* 1997) and mPr (mean pairwise relatedness). The estimates were obtained for the full data set excluding the poorly sampled localities ($N < 4$ rats) of PL7(V4), SA4 (V6), SJ5(V7) (Appendix S1, Supporting information).

Population size (N_E and N_C)

We calculated both effective (N_E) and census (N_C) population sizes for the three genetic clusters identified by the clustering analysis, under the assumption that they form units within which random mating occurs. N_E was computed using Waples and Do's LD method implemented in LDNe (Waples & Do 2008). Because LD is affected by migration, the low migration detected between genetic clusters should reduce the bias in the N_E estimation. N_C was computed based on N_E estimates and the ratio of the realized (observed genotypes) and possible (expected genotypes) diploid combinations of all sampled alleles (Hyseni *et al.* 2012). This analysis used a Bayesian method (Petit & Valiere 2006) adapted from Gazey & Staley (1986; Appendix S1, Supporting information).

Results

Marker validation, genetic diversity and relatedness

We successfully genotyped 146 of 150 samples across 17 microsatellite loci. After sequential Bonferroni corrections for multiple tests, only one locus, D8Mgh7, from one sampling site, PL1(V1), showed a significant probability of deviation from HWE, and no pairwise comparisons of loci across sampling sites showed significant signatures of LD. This reaffirmed our choice of loci from previous studies based on HWE and absence of linkage (data available on request).

Both the AMOVA (Table S3, Supporting information) and the F_{ST} (Table S4, Supporting information) results on the temporally separated samples obtained from the same sites suggested that the 2010 and 2011 collections were not significantly different with the exception of the PL8(V2) site ($F_{ST} = 0.13$, $P < 0.05$). Thus, diversity statistics were calculated by site, with temporal collections combined for each site, except for the PL8(V2) site. Genetic variation was high, despite the small spatial scale of the sampling effort and the small sample sizes from a few sites (Table 2). Corrected allelic richness (A_r) ranged from 2.71 to 3.29, while private (unique to a population) allelic richness (A_p) ranged from 0.06 to 0.30, and H_e ranged from 0.57 to 0.72 (Table 2).

Results of the relatedness analyses within each site with at least 11 samples per site indicated that most of the samples (87%) came from unrelated individuals with approximately 12% sharing a half-sibling relationship and approximately 1% being either full-siblings or parent-offspring (Table 3). To further explore this result, we repeated the analysis on a subset of samples from four households in Pau da Lima with samples size >6 where traps were placed close to a single burrow (Table 3). Even for this subset of samples that was more

likely to include family groups than the whole data set, most (>90%) of trapped individuals, comprised of adult and young rats of both sexes, were unrelated (Table 3).

Genetic differentiation, sex-biased dispersal and population sizes

The AMOVA analyses show that most of the variation was attributable to intrasite differences (89.4%) with intersite variation contributing 8.2% (Table 4). The majority of the pairwise F_{ST} values for intersite comparisons were also significantly different and ranged from 0.06 to 0.28 (Table S5, Supporting information), even though the geographical distance between sampling sites was small, ranging from 79 m to a few hundred metres (Fig. 1). The four nonsignificant F_{ST} values were all associated with comparisons that included sites with small samples sizes [PL7(V4), SA4(V6) and SJ5(V7), $N = 3$, Table 1, Table S5, Supporting information].

Bayesian clustering analyses evaluating ΔK (Evanno *et al.* 2005) identified three genetic clusters (Fig. 2A). Cluster 1 grouped several of the Pau da Lima sites [PL1 (V1), PL6(V4), PL7(V4), PL8-2010(V2) and PL9(V3)] with the São Joaquim site, SL5(V7). Cluster 2 included only the two sets of samples obtained a year apart from Pau da Lima site 2 [PL2-2010(V2) and PL2-2011(V2)]. Cluster 3 grouped the samples from PL8 obtained in 2011 with the two Sete de Abril sites [SA3(V5) and SA4(V6)] (Fig. 2A). The multivariate analyses (DAPC) also detected three genetic clusters with most individuals assigned with >95% posterior probability to the same three groups identified by the clustering analysis (Fig. 2B,C). The Mantel test between pairwise F_{ST} values and geographical distance sites did not detect any IBD

Table 2 Genetic diversity by site for 146 individuals. The first two columns report the sampling sites with the respective valley in parentheses and samples size (N). Measures include allelic richness (A_r) and private allelic richness (A_p) averaged over all loci, and mean observed (H_o) and expected (H_e) heterozygosity

Site (Valley)	N	A_r	A_p	H_o	H_e
PL1(V1)	25	3.29	0.17	0.67	0.70
PL2(V2)	30	3.05	0.10	0.67	0.68
PL6(V4)	18	3.12	0.06	0.65	0.69
PL7(V4)	3	3.00	0.12	0.62	0.72
PL8_2010(V2)	7	3.07	0.22	0.61	0.67
PL8_2011(V2)	16	2.99	0.11	0.70	0.67
PL9(V3)	11	3.07	0.11	0.67	0.69
SA3(V5)	30	3.00	0.14	0.63	0.64
SA4(V6)	3	2.71	0.21	0.55	0.57
SJ5(V7)	3	2.71	0.30	0.71	0.59

Table 3 Percentage of pairwise comparisons of individuals that fell into each relatedness category (e.g. unrelated, half-sibling, full-sibling) as calculated in ML-Relate (Kalinowski *et al.* 2006). Comparisons were made among all individuals and among individuals within sites and households with $N > 6$

	All Sites	PL1 (V1)	PL2 (V2)	PL6 (V4)	PL8 (V2)	PL9 (V3)	SA3 (V5)	House 5 (PL1)	House 15 (PL2)	House 16 (PL9)	House 19 (PL8)
Unrelated (%)	87.1	92.0	87.6	92.2	85.4	89.1	82.8	100	90.5	95.2	90.9
Half-sibling (%)	11.6	7.3	9.2	6.5	9.1	7.3	12.4	0.0	4.8	4.8	6.1
Full-sibling (%)	1.1	0.7	2.1	0.0	3.9	3.6	2.5	0.0	4.8	0.0	1.5
Parent-Offspring	0.2	0.0	1.1	1.3	1.6	0.0	2.3	0.0	0.0	0.0	1.5

Table 4 Analysis of molecular variance (AMOVA) results by neighbourhood and collection site for all rat samples (geographical structure)

Source of variation	Variance component	% of variation	<i>P</i> -value
Among groups (neighbourhoods)	$V_a = 0.11$	2.43	0.05
Among sites within groups (sites)	$V_b = 0.37$	8.18	0.00
Within sites (individuals)	$V_c = 4.03$	89.39	0.00

for the contiguous sites in Pau de Lima ($P = 0.957 > 0.05$; Fig. S1, Supporting information).

Table 5 shows the results from the migrant analyses. Although most individuals (93.2%) were correctly assigned to their site of capture, the two GENECLASS methods identified a total of ten migrants, with the majority (9/10) being adults and all but one male. The L_h method detected six migrants, while the other GENECLASS method (L_h/L_{max}) detected four of the same individuals plus 3 additional individuals (Table 5). Four of the migrants identified by the GENECLASS methods were also identified as migrants on the estimated membership coefficients for the capture site and assigned site (Q_o and Q_a , respectively) calculated using STRUCTURE (Table 5). These analyses also suggested that these four individuals have mixed ancestry (Q_a from 0.358 to 0.875), possibly deriving from ancestral matings across multiple clusters.

Male-biased dispersal was inferred by three (F_{ST} , $vAIC$ and mPr) and one (mPr) estimate in the two-sided and one-sided tests, respectively (Table 6). Neither a one-sided nor two-sided test showed evidence for female-biased dispersal. These results complement the finding from the assignment tests (Table 5) that indicate that the majority of migrants are males.

Computations based on linkage disequilibrium (LDNe, a measure of N_e based on nonrandom associations of linked alleles, Waples & Do 2008) for the three

genetic clusters identified by Structure and multivariate analyses (Fig 2) resulted in mean estimates of N_E ranging from 34 to 94 with corresponding N_C estimates ranging from 273 to 897 (Table S6, Supporting information). Cluster 2, which only included the two temporally separated samples from PL2, had the smallest N_E and N_C estimates (N_E : 34, with 95% CI ranging from 10 to 40; $N_C = 273$, 95% HPD = 167–396). Cluster 1, which grouped together five sites from 3 valleys in Pau da Lima (PL1, PL6, PL7, PL8-2010 and PL9) and the São Joaquim site, SL5, had the highest population size estimates (N_E : 94, with 95% CI ranging from 75 to 125; $N_C = 897$, 95% HPD = 357–1608). Cluster 3, comprising one of the two PL8 temporally distinct samples and the two Sete de Abril sites (SA3 and SA4), had intermediate values (N_E : 54, with 95% CI ranging from 35 to 85; $N_C = 454$, 95% HPD = 233–722).

Discussion

Patterns of genetic diversity, differentiation and relatedness

Given the relatively recent colonization of Salvador by Norway rats (coinciding with the construction of slum settlements in the last half-century, Moreira & Pereira 2008) and the small spatial scale of this study (around 300 m–1.8 km between most sites), the finding of relatively high levels of genetic variation within sampling sites and significant genetic differentiation between sites was unexpected. We found high genetic diversity (H_e ranging from 0.57 to 0.72, Table 2) even for sites represented by small samples sizes ($N = 3$, Table 2) and no difference in diversity between samples (Pau de Lima sites) that had experienced recent (2009–2010) rodenticide control and those (Sete de Abril and São Joaquim sites) that did not. The levels of genetic diversity detected are comparable to the those reported for rat samples from urban Baltimore (Maryland, USA), where H_e values ranged from 0.67 to 0.78 for 277 Norway rats trapped from 11 sampling sites (Gardner-Santana *et al.* 2009), as well as from rural populations on a much

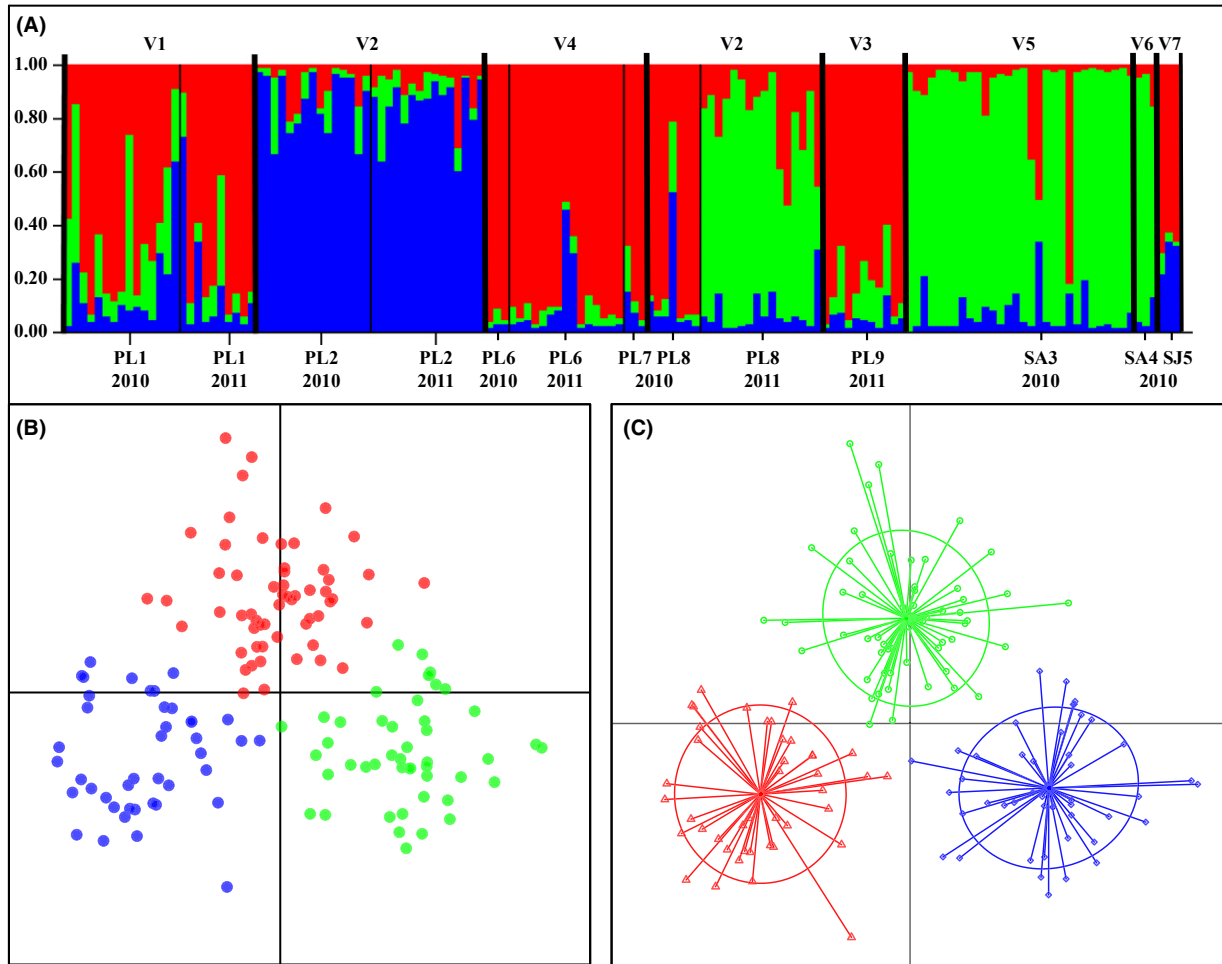


Fig. 2 Genetic clusters and population structure. In all panels, cluster 1 (red) = Pau da Lima sites 1, 6, 7, 8 (2010 only) and 9, cluster 2 (blue) = Pau da Lima site 2, cluster 3 (green) = Sete de Abril sites 3 and 4, Pau da Lima site 8 (2011 only), (A) STRUCTURE bar plot ($K = 3$) with individuals organized by site, year and valley. Each vertical bar represents a single individual. The height of each colour represents the probability of assignment to that genetic cluster, (B) principal components analysis, dots of different colours identify individuals from different genetic clusters, colours as in A, (C) discriminant analysis of principal components with 3 K-means clusters. Circles indicate 95% confidence interval of assignment, dots identify individuals, and colours of the individuals are as in A.

larger spatial scale ($H_e = 0.42$ Abdelkrim *et al.* 2005; $H_e = 0.59$ Abdelkrim *et al.* 2009; $H_e = 0.72$ – 0.82 Abdelkrim *et al.* 2010; $H_e = 0.55$ Russell *et al.* 2009b). In neither case had these populations been subjected to rodenticide treatment in times just prior to sampling.

Our results highlight the existence of complex population genetic structure among Norway rats in Salvador with moderate genetic admixture (Fig. 2A, Table 6). There was a considerable amount of genetic differentiation between geographically proximate sites, and in one case, even between samples obtained at the same site but in two consecutive years [PL8(V2), Fig. 2]. These observations indicate that rapid population turnover, although uncommon, does happen.

The results of the clustering analyses (Fig. 2) suggest the existence of three genetic groups, which differentiate rats captured from sampling sites between the two

neighbourhoods of Pau da Lima and Sete de Abril, as well as sites within Pau da Lima. The three clusters correspond to the sites with the highest sample size (approximately 30 individuals). For instance, sites with low sample sizes, including SJ5(V7), grouped with PL1 (V1), a site located in a different neighbourhood more than 8 km away, in cluster 1. This result suggests that inadequate sampling might be responsible for our inability to discriminate between some of the sites, highlighting the need for further studies with increased sample sizes.

The results of the Mantel test and F_{ST} analyses show moderate but significant pairwise genetic differentiation without a pattern of IBD (examined for the more densely sampled Pau da Lima neighbourhood). This parallels the results of the clustering analyses and suggests little correlation ($P = 0.957$, $R^2 = 0.342$) between genetic

Table 5 First-generation migrant analyses for all the sites and sampling times, except PL8(V2) which is separated by year. The first three columns report the individual identification number followed by its age class and sex. The following two columns show the origin and the assigned sites. The next two columns report the results of the GENECLASS 2.0 analyses, listing P -values for L_h and L_h/L_{max} . The last two columns show the assignment values (Q) using STRUCTURE for the site of origin (Q_o) and the assigned site (Q_a). Individuals in bold were detected as first-generation migrants by two or more methods

Sample	Age	Sex	Site (origin)	Site (assigned)	L_h P -value	L_h/L_{max} P -value	Q_o	Q_a
5	3	M	PL8_2010(V2)	PL1(V1)	0.005	0.003	0.110	0.358
62	3	M	SA4(V6)	SA3(V5)	0.001	0.002	—	—
47	3	M	SA3(V5)	PL1(V1)	0.000	0.000	0.000	0.875
9	3	M	PL8_2010(V2)	PL9(V3)	0.006	0.003	0.092	0.708
101	2	M	SJ5(V7)	SJ5(V7)	0.004	—	—	—
13	2	M	PL7(V4)	SA4(V6)	0.000	0.000	—	—
131	1	M	PL8_2011(V2)	SA3(V5)	—	0.004	—	—
134	3	F	PL8_2011(V2)	PL1(V1)	—	0.004	—	—
146	2	M	PL2(V2)	PL1(V1)	—	0.008	—	—
42	2	M	SA3(V5)	PL1(V1)	—	0.007	0.157	0.554

Table 6 Results from tests for sex-biased dispersal using four different statistics reported for both the two- and the one-sided tests for females and males. Number of individuals for each sex is listed in parenthesis. The second columns show F_{ST} values followed by the correct mean assignment index (mAIC) and its variance (vAIC), using Favre *et al.*'s (1997) method. The last column reports the mean pairwise relatedness (mPr). Significance values (P -values) for these measures are reported in the last row

	F_{ST}	mAIC	vAIC	mPr
Two sided				
F (67)	0.104	0.585	16.03	0.191
M (70)	0.076	-0.560	27.38	0.135
P -value	0.059	0.152	0.192	0.034
One sided				
F (67)	0.104	0.585	16.03	0.191
M (70)	0.076	-0.560	27.38	0.135
P -value	0.046	0.081	0.026	0.027

and geographical distance. In addition to the nonsignificant P -value in the Mantel test, the plotted distribution in Fig. S1 (Supporting information) shows widespread clumped points, connected by a negatively sloped trend line, rather than a scatter of points, suggesting a negative correlation of genetic and geographical distance. While the F_{ST} value for PL6 and PL7, located 317 m apart in V4, is not statistically significant, the F_{ST} values for localities in different valleys [PL7(V4) and PL8(V2)], separated by just 79 m, are significant, as are F_{ST} values for sites separated by geographical distances comparable to the one between PL7(V4) and PL8(V2) (393–263 m; Tables 5, Fig. S1, Supporting information).

More comprehensive sampling, involving placing more traps at sites already included in this study, may improve the resolution of the genetic differentiation

between valleys. In the only study directly comparable to ours, Gardner-Santana *et al.* 2009 analysed rats captured from 11 sites in Baltimore, separated by an average of 4.7 km and detected three to four genetic clusters across the city. In this study, we found genetic differentiation between sites separated by <100 m. This discrepancy could be influenced by the topographical complexity of Salvador, as evidenced by the presence of numerous semi-isolated valleys, compared with the relatively flat and uniform topography of Baltimore. Interestingly, in a 3-year-long capture–mark–recapture study in Denmark, there was no movement of Norway rats between adjacent, but separated, sewer systems; rats were genetically clustered within the same sewer system, moving an average maximal distance of approximately 200 m, suggesting the existence of separate population units with only limited mixing between them (Heiberg *et al.* 2012).

As dispersal patterns and social structures have been shown to influence the distribution of genetic variation in populations (Chesser 1991a, 1991b; Matocq *et al.* 2000), we initially hypothesized that the genetic differentiation observed between these sites could be partly due to the social behaviour and ecology of this species, as related extended families of Norway rats tend to inhabit burrow systems proximal to one another (Calhoun 1962). However, the observed patterns of genetic divergence among geographically close groups of sampling sites cannot be attributed to high relatedness among individuals, because our sampling overwhelmingly (approximately 90% pairwise relationships) consisted of unrelated individuals, even from traps located around the same household (Table 3). No information is available for tropical urban settings, but temporal and seasonal fluctuations, that is, colony instability that only increased sampling efforts will elucidate, at local

scales may rapidly change allele frequencies (Balloux & Lugon-Moulin 2002; Gileva *et al.* 2006). The presence of unrelated juveniles and adults captured from the same household could reflect colony instability that for the Pau de Lima sites, could have been aggravated by rodenticide poisoning in the 2 years before the samples were collected and could have permitted burrow recolonization by unrelated rats from other neighbourhoods.

Sex-biased dispersal

The relatively high level of genetic structuring in our data could also be influenced by Norway rat behavioural traits (Brouat *et al.* 2007). *R. norvegicus*, like the majority of rodents, tend to exhibit female philopatry and male dispersal (Calhoun 1962), although social aggression (Davis 1951a,b) and environmental disturbances can precipitate long-distance movement by both sexes (Taylor 1978). Despite using similar methods, genetic studies in Baltimore (Gardner-Santana *et al.* 2009) failed to find evidence for philopatry and sex-biased dispersal among urban Norway rats, suggesting that both sexes might move in high-density urban populations because adult males may not be able to establish and hold breeding territories (Childs *et al.* 1991; Macdonald & Fenn 1995). It is possible that these rat populations, or at least those from the Pau de Lima sites, were not subject to these density-dependent effects on social hierarchies and dispersal because populations were below carrying capacity due to the previous rodent control efforts. Gardner-Santana *et al.* did not mention any rodent control activities during their study. Population reductions could result in conditions favourable for recolonization of burrow systems by migrating or dispersing animals from adjacent populations, especially if adult male-dominated breeding territories were disrupted in populations destabilized by rodenticide control interventions (Calhoun 1962).

Male-biased migration among study sites in Salvador, as assessed by *F*_{STAT}, *STRUCTURE* and *GENECLASS*, provides support for the hypothesis that population disruption increases genetic diversity and admixture between sites. Of the ten individuals detected as migrants, all but one were male, and among the males, nine were sexually mature adults or subadults (>200 g). Moderate genetic structuring could result as subordinate males migrate/disperse and, in areas where highly structured dominant male-based territories are less common, could have greater access to females (Macdonald *et al.* 1999; Hinson *et al.* 2006). The influx of new genes would be amplified by the high fecundity and short generation time typical of this species.

It is also important to note the promiscuous breeding habits of the Norway rat, where numerous copulations

occur nightly among receptive females, and the presence of a female in oestrus entices male rats from other burrow systems to travel outside their accustomed ranges (Calhoun 1962). Although dominant males may decrease access of subdominant males to adult females residing in the same burrow, this protection is always incomplete, and numerous mountings of dominant male-associated females with other males are assured (Calhoun 1962). Although not every mating attempt results in ejaculation, the potential for mixed parentage among litters is great and will reduce the degree of relatedness among extended families. Additionally, the ability of male rats to control the amount of ejaculate based on the presence of other males in the breeding environment (sperm allocation) may increase the probability that males lower down in the succession of males mounting females have an advantage in contributing their sperm/genes to receptive females (Pound & Gage 2004).

Our migrant analyses identified putative dispersal/migration routes both within and between valleys and neighbourhoods, which are keys for implementation of pest management programmes targeting the eradication unit (Russell *et al.* 2009a). Although further investigations with a denser sampling scheme are needed to clarify 'source' and 'sink' population dynamics, our results suggested that PL1(V1) could be a source population, providing migrants to other sites within and beyond the same neighbourhood [*i.e.* PL8(V2) and SA3 (V5)].

The shape and adjacency of two of the valley systems sampled (Fig 1) could facilitate dispersal/migration, as no obvious impermeable barriers exist. During the rainy season, especially if flooding occurs, rats can disperse to neighbouring valleys riding on flowing water (Cottam 1948). PL8, located at the bottom of valley V2, seems to represent a sink, receiving migrants from sites in other valleys (V1 and V3) and from different neighbourhoods (Sete de Abril, SA3). The dynamic nature of the genetic variation at this site was emphasized by the finding of significant pairwise *F*_{ST} values between rat samples from successive years (Table 5) and their genetic assignment to different clusters (Fig. 2). This result is probably due to the geographical position of PL8 in valley 2, at the entrance to the other valleys in Pau da Lima, and the existence of potential corridors between Pau da Lima and Sete de Abril (Fig. 1). At this time, we can only conjecture about the mechanisms contributing to the genetic variability present at this site, and only future efforts to acquire additional samples will provide the means to understand source-sink dynamics within and between valleys.

The São Joaquim market site and two sites in Pau da Lima [PL6(V4) and PL7(V4)], although showing signs of

genetic admixture (Fig. 2), did not show strong evidence of recent migration from other sampled sites, possibly because of the greater distance between the São Joaquim site and all other sites (approximately 8.2–10.1 km). São Joaquim is located near the international sea port, which could be a source of additional genetic diversity, but sites as yet unsampled could also be the source of the migrants to SJ5(V7), PL6(V4) and PL7(V4), leading to the detection of genetic admixture. The potential contribution of as-yet-unsampled sites is a complicating factor precluding firm conclusions regarding the processes leading to the observed genetic admixtures and thus dictates the need for increased sampling.

The population size estimates were obtained using the genetic clusters identified by the Bayesian and multivariate analyses (Fig. 2), which include multiple sampling sites within each genetic cluster. As expected cluster 2, which included only one site with two temporal samplings, had the lowest population size estimates, while cluster 3, which groups most of the other sites, had the largest (Table S6, Supporting information). The sparse sampling scheme of the present study does not support robust population size estimates, as we suspect our ability to identify additional genetic clusters may have been hindered by the small number of sampling sites both within and around the valleys. A denser sampling scheme will allow a better spatial definition of the genetic units and thus lead to more accurate population estimates.

However, it is reassuring to find that the N_C estimates based on genetic rather than traditional mark-recapture data are higher than the corresponding N_E estimates because N_C estimates are usually higher than N_E estimates. In this study, we report a $1/10 N_E/N_C$ ratio, which is similar to other estimates (Frankham 1995; Luikart *et al.* 2010). In further studies, we aim to confirm the validity of this approach and provide a solid understanding of both the stability of the N_E/N_C ratio and the factors that could impact its change in this study system. Calculating N_E estimates from genetic data could provide a cost-effective method to assess the number of rats in Salvador at different spatial and temporal scales with obvious benefits for designing more targeted control interventions simultaneously targeting genetically linked populations that provide rats a corridor to disperse/migrate between neighbourhoods.

Implications for control

Currently, the eradication unit for rodent control programmes in Salvador, as well as many urban centres in Brazil and worldwide, is the city 'block', a unit based on artificial boundaries. The finding of genetic

differentiation between sampling sites in neighbouring valleys only a few hundred metres apart, coupled with evidence of genetic admixture between distant sites, and genetic discontinuities between temporal samplings at some sites suggests that eradication unit sizes should vary between sites. Our data suggest that, given the complex urban landscape, an eradication unit based on the city block might not be necessary or advisable, as geographically based eradication units can be equally efficient, resulting in less expensive in the long run and easier to sustain city-wide control and monitoring measures.

The temporal comparisons of samples from the same sites collected in successive years provide further insight into the re-emergence dynamics of these rat populations that is relevant for developing control interventions and their monitoring through time. In some sites, individuals belonging to the same genetic cluster were captured in the same location in successive years, while in others, we observed a population replacement. These observations suggest that colony re-establishment can happen either by *in situ* re-emergence or by recolonization from surrounding areas. The detection of first-generation migrants in a few sites implies that rats can move to new sites quite readily and underscores the importance of understanding the factors that enable such movements, especially in geographical areas at the base of each valley, which act at the same time as dispersal corridors between valleys and sources of migrants for slum areas located within a valley.

Acknowledgments

We thank the staff of the *Centro de Controle de Zoonoses* (CCZ) in Salvador for their assistance in conducting the study. This work was supported by the Secretariat of Health Surveillance and Oswaldo Cruz Foundation, Brazilian Ministry of Health; CAPES (Coordination for the Improvement of Higher Education Personnel), Brazilian Ministry of Education; National Institutes of Health (grants R01 TW009504, U01 AI088752, R01 AI052473, R24 TW007988 and D43 TW00919); and Yale Institute for Biospheric Studies, Molecular Systematics and Conservation Genetics Center and Wilbur Downs Fellowship Program. We thank Danielle Edwards, Sarah Federman and Ryan Garrick for their assistance during data collection and analysis, as well as The Yale Institute for Biospheric Studies, Molecular Systematics and Conservation Genetics Center for use of laboratory facilities. Brittney Kajdacs was supported by a Science Scholarship from The Graduate Center, City University of New York (CUNY) during the writing of this manuscript.

References

- Abdelkrim J, Pascal M, Calmet C, Samadi S (2005) 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, Samadi S (2007) Establishing causes of eradication failure based on genetics: case study of ship rat eradication in Ste. Anne archipelago. *Conservation Biology*, **21**, 719–730.
- Abdelkrim J, Pascal M, Samadi S (2009) Genetic structure and functioning of alien ship rat populations from a Corsican micro-insular complex. *Biological Invasions*, **11**, 473–482.
- Abdelkrim J, Byrom AE, Gemmell NJ (2010) Fine-scale genetic structure of mainland invasive *Rattus rattus* populations: implications for restoration of forested conservation areas in New Zealand. *Conservation Genetics*, **11**, 1953–1964.
- Amilasan AS, Ujiie M, Suzuki M *et al.* (2012) Outbreak of leptospirosis after flood, the Philippines, 2009. *Emerging Infectious Diseases*, **18**, 91–94.
- Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology*, **11**, 155–165.
- Beadell JS, Hyseni C, Abila PP *et al.* (2010) Phylogeography and population structure of *Glossina fuscipes fuscipes* in Uganda: implications for control of tsetse. *PLoS Neglected Tropical Diseases*, **4**, e636.
- Brazilian Institute of Geography and Statistics (1971) Brazilian Ministry of Health, estimated population retrospective statistical series 1960–70.
- Brazilian Institute of Geography and Statistics (2011) Summary of the 2010 demographic census.
- Brouat C, Loiseau A, Kane M, Bâ K, Duplantier JM (2007) Population genetic structure of two ecologically distinct multimammate rats: the commensal *Mastomys natalensis* and the wild *Mastomys erythroleucus* in southeastern Senegal. *Molecular ecology*, **16**, 2985–2997.
- Brown JE, McBride CS, Johnson P *et al.* (2011) Worldwide patterns of genetic differentiation imply multiple “domestications” of *Aedes aegypti*, a major vector of human diseases. *Proceedings. Biological Sciences/The Royal Society*, **278**, 2446–2454.
- Calhoun JB (1962) *The Ecology and Sociology of the Norway Rat*. US Department of Health, Bethesda, MD.
- Calmet C, Pascal M, Samadi S (2001) Is it worth eradicating the invasive pest *Rattus norvegicus* from Molène archipelago? *Biodiversity and Conservation*, **10**, 911–928.
- Chanteau S, Ratsifasoamanana L, Rasoamanana B *et al.* (1998) Plague, a reemerging disease in Madagascar. *Archives de l'Institut Pasteur de Madagascar*, **64**, 12–14.
- Chesser RK (1991a) Gene diversity and female philopatry. *Genetics*, **127**, 437–447.
- Chesser RK (1991b) Influence of gene flow and breeding tactics on gene diversity within populations. *Genetics*, **129**, 573–583.
- Childs JE, Glass GE, LeDuc JW (1991) Rodent sightings and contacts in an inner-city population of Baltimore, Maryland, U.S.A. *Bulletin of the Society of Vector Ecologists*, **16**, 245–255.
- Civen R, Ngo V (2008) Murine typhus: an unrecognized suburban vectorborne disease. *Clinical Infectious Diseases*, **46**, 913–918.
- Cottam C (1948) Aquatic habitats of the Norway rat. *Journal of Mammalogy*, **29**, 299.
- Davis DE (1951a) The relation between level of population and size and sex of Norway rats. *Ecology*, **32**, 462–464.
- Davis DE (1951b) The characteristics of global rat populations. *American Journal of Public Health and the Nations Health*, **41**, 158–163.
- Davis DE (1953) The characteristics of rat populations. *The Quarterly Review of Biology*, **28**, 373–401.
- Davis D, Emlen J, Stokes A (1948) Studies on home range in the brown rat. *Journal of Mammalogy*, **29**, 207–225.
- Dechet AM, Parsons M, Rambaran M *et al.* (2012) Leptospirosis outbreak following severe flooding: a rapid assessment and mass prophylaxis campaign; Guyana, January–February 2005. *PLoS ONE*, **7**, e39672.
- 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.
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.
- Farhang-Azad A, Southwick CH (1979) Population ecology of Norway rats in the Baltimore Zoo and Druid Hill Park, Baltimore, Maryland. *Annals of Zoology*, **15**, 1–42.
- de Faria MT, Calderwood MS, Athanazio DA *et al.* (2008) Carriage of *Leptospira interrogans* among domestic rats from an urban setting highly endemic for leptospirosis in Brazil. *Acta tropica*, **108**, 1–5.
- Favre L, Balloux F, Goudet J, Perrin N (1997) Female-biased dispersal in the monogamous mammal *Crocodyrus russula*: evidence from field data and microsatellite patterns. *Proceedings. Biological Sciences/The Royal Society*, **264**, 127–132.
- Frankham R (1995) Effective population-size: adult-population size ratios in wildlife – A review. *Genetical Research*, **66**, 95–107.
- 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.
- Gazey WJ, Staley MJ (1986) Population estimation from mark-recapture experiments using a sequential Bayes algorithm. *Ecology*, **67**, 941–951.
- Gilbert A, Loiseau A, Duplantier JM *et al.* (2007) Genetic structure of black rat populations in a rural plague focus in Madagascar. *Canadian Journal of Zoology*, **85**, 965–972.
- Gileva EA, Rakitin SB, Cheprakov MI (2006) Genomic instability in the bank vole: population-ecological aspects. *Russian Journal of Ecology*, **37**, 271–277.
- Glass GE, Childs JE, Korch GW, LeDuc DE (1989) Comparative ecology and social interactions of Norway rat (*Rattus norvegicus*) populations in Baltimore, Maryland. *Occasional Papers of the Museum of Natural History, University of Kansas, Lawrence, Kansas*, **130**, 1–33.
- Glass GE, Klein SL, Shields T, Easterbrook J (2005) The Norway rat population in Baltimore. *Maryland Vector-Borne and Zoonotic Diseases*, **5**, 296–299.
- Goudet J (1995) FSTAT (Version 1.2): a computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485–486.
- Goudet J, Perrin N, Waser P (2002) Tests for sex-biased dispersal using bi-parentally inherited genetic markers. *Molecular Ecology*, **11**, 1103–1114.
- Heiberg A, Sluydts V, Leirs H (2012) Uncovering the secret lives of sewer rats (*Rattus norvegicus*): movements, distribution and population dynamics revealed by a capture-mark-recapture study. *Wildlife Research*, **39**, 202–219.
- Himsworth CG, Parsons KL, Jardine C, Patrick DM (2013) Rats, cities, people, and pathogens: a systematic review and narra-

- tive synthesis of literature regarding the ecology of rat-associated zoonoses in urban centers. *Vector-Borne and Zoonotic Diseases*, **13**, 349–359.
- Hinson ER, Hannah MF, Norris DE, Glass GE, Klein SL (2006) Social status does not predict responses to Seoul virus infection or reproductive success among male Norway rats. *Brain, Behavior, and Immunity*, **20**, 182–190.
- Holm S (1979) A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, **6**, 65–70.
- Howald G, Donlan JC, Galván JP *et al.* (2007) Invasive rodent eradication on Islands. *Conservation Biology*, **21**, 1258–1268.
- Hyseni C, Kato AB, Okedi LM *et al.* (2012) The population structure of *Glossina fuscipes fuscipes* in the Lake Victoria basin in Uganda: implications for vector control. *Parasites & Vectors*, **5**, 222.
- Iskjaer C, Slade NA, Childs JE, Glass GE, Korch GW (1989) Body mass as a measure of body size in small mammals. *Journal of Mammalogy*, **70**, 662–667.
- Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, **11**, 94.
- Kalinowski ST (2004) Counting alleles with rarefaction: private alleles and hierarchical sampling designs. *Conservation Genetics*, **5**, 539–543.
- Kalinowski ST (2005) Hp-Rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, **5**, 187–189.
- Kalinowski ST, Wagner AP, Taper ML (2006) ML-relate: a computer program for maximum likelihood estimation of relatedness and relationship. *Molecular Ecology Notes*, **6**, 576–579.
- Ko AI, Reis MG, Ribeiro Dourado CM, Johnson WD, Riley LW (1999) Urban epidemic of severe leptospirosis in Brazil. Salvador Leptospirosis Study Group. *Lancet*, **354**, 820–825.
- Ko AI, Goarant C, Picardeau M (2009) Leptospira: the dawn of the molecular genetics era for an emerging zoonotic pathogen. *Nature reviews. Microbiology*, **7**, 736–747.
- Kosoy M, Bai Y, Sheff K *et al.* (2010) Identification of Bartonella infections in febrile human patients from Thailand and their potential animal reservoirs. *The American Journal of Tropical Medicine and Hygiene*, **82**, 1140–1145.
- Luikart G, Ryman N, Tallmon DA, Schwartz MK, Allendorf FW (2010) Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conservation Genetics*, **11**, 355–373.
- Macdonald DW, Fenn MPG (1995) Rat ranges in arable areas. *Journal of Zoology (London)*, **236**, 349–353.
- Macdonald DW, Mathews F, Berdoy M (1999) The behaviour and ecology of *Rattus norvegicus*: from opportunism to kamikaze tendencies. In: *Ecologically-Based Rodent Management* (eds Singleton GR, Hinds L, Leirs H, Zhang Z), pp. 49–80. Australian Centre for International Agricultural Research, Canberra.
- Matocq MD, Patton JL, da Silva MNF (2000) Population genetic structure of two ecologically distinct Amazonian spiny rats: separating history and current ecology. *Evolution*, **54**, 1423–1432.
- Matthias MA, Ricaldi JN, Cespedes M *et al.* (2008) Human leptospirosis caused by a new, antigenically unique leptospira associated with a rattus species reservoir in the peruvian Amazon. *PLoS Neglected Tropical Diseases*, **2**, e213.
- McCoy KD (2008) The population genetic structure of vectors and our understanding of disease epidemiology. *Parasite*, **15**, 444–448.
- Mills JN, Childs JE, Ksiazek TG, Peters CJ, Velleca WM (1995) *Methods for Trapping and Sampling Small Mammals for Virologic Testing*. US Dept. of Health & Human Services, Centers for Disease Control and Prevention, Atlanta, GA.
- Moreira IM, Pereira GC (2008) *Salvador and Metropolitan Region*, pp. 228. Edufba, Salvador, Brazil.
- Nkogwe C, Raletobana J, Stewart-Johnson A, Suepaul S (2011) Frequency of detection of *Escherichia coli*, *Salmonella spp.*, and *Campylobacter spp.* in the faeces of wild rats (*Rattus spp.*) in Trinidad and Tobago. *Veterinary Medicine International*, **2011**, 1–7.
- Orgain H, Schein MW (1953) A preliminary analysis of the physical environment of the Norway rat. *Ecology*, **34**, 467–473.
- Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology*, **13**, 55–65.
- Pereira MM, Andrade J (1988) Epidemiological aspects of leptospirosis in a slum area in the city of Rio de Janeiro, Brazil. Search for leptospire and specific antibodies in rodents. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **82**, 768–770.
- Pereira MM, Matsuo MGS, Bauab AR *et al.* (2000) A clonal subpopulation of *Leptospira interrogans sensu stricto* is the major cause of leptospirosis outbreaks in Brazil. *Journal of Clinical Microbiology*, **38**, 450–452.
- Petit E, Valiere N (2006) Estimating population size with non-invasive capture-mark-recapture data. *Conservation Biology*, **20**, 1062–1073.
- Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. *The Journal of Heredity*, **95**, 536–539.
- Pound N, Gage MJG (2004) Prudent sperm allocation in Norway rats, *Rattus norvegicus*: a mammalian model of adaptive ejaculate adjustment. *Animal Behaviour*, **68**, 819–823.
- Pritchard J, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences, USA*, **94**, 9197–9201.
- Raymond M, Rousset F (1995) GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. *The Journal of Heredity*, **86**, 248–249.
- Reis RB, Ribeiro GS, Felzemburgh RDM *et al.* (2008) Impact of environment and social Gradient on *Leptospira* infection in urban slums. *PLoS Neglected Tropical Diseases*, **2**, e228.
- Robertson BC, Gemmel NJ (2004) Defining eradication units to control invasive pests. *Journal of Applied Ecology*, **41**, 1042–1048.
- Rousset F (2008) Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.
- Russell JC, Mackay JWB, Abdelkrim J (2009a) Insular pest control within a metapopulation context. *Biological Conservation*, **142**, 1404–1410.

- Russell JC, Abdelkrim J, Fewster RM (2009b) Early colonisation population structure of a Norway rat island invasion. *Biological Invasions*, **11**, 1557–1567.
- Russell JC, McMorland AJC, MacKay JWB (2010) Exploratory behaviour of colonizing rats in novel environments. *Animal Behaviour*, **79**, 159–164.
- Sarkar U, Nascimento SF, Barbosa R *et al.* (2002) Population-based case-control investigation of risk factors for leptospirosis during an urban epidemic. *American Journal of Tropical Medicine and Hygiene*, **66**, 605–610.
- Singleton GR, Leirs H, Hinds LA, Zhang Z (eds) (1999) *Ecologically-Based Management of Rodent Pests*, pp. 49–80. Australian Centre for International Agricultural Research, Canberra.
- Slotman MA, Kelly NB, Harrington LC *et al.* (2006) Polymorphic microsatellites markers for studies of *Aedes aegypti* (Diptera Culicidae), the vector of dengue and yellow fever. *Molecular Ecology Notes*, **7**, 168–171.
- Solano P, Ravel S, de Meeüs T (2010) How can tsetse population genetics contribute to African trypanosomiasis control? *Trends in Parasitology*, **26**, 255–263.
- Taylor KD (1978) Range of movement and activity of common rats (*Rattus norvegicus*) on agricultural land. *Journal of Applied Ecology*, **15**, 663–677.
- Taylor KD, Quy RJ (1979) Long distance movements of a common rat (*Rattus norvegicus*) revealed by radio-tracking. *Mammalia*, **42**, 63–72.
- Traweger D, Travnitzky R, Moser C, Walzer C, Bernatzky G (2006) Habitat preferences and distribution of the brown rat (*Rattus norvegicus* Berk.) in the city of Salzburg (Austria): implications for an urban rat management. *Journal of Pest Science*, **79**, 113–125.
- UN-HABITAT (2010) Hidden cities: unmasking and overcoming health inequities in urban settings. World Health Organization, The WHO Centre for Health Development and United Nations Human Settlements Programme, .
- Waples RS, Do C (2008) LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, **8**, 753–756.
- Webster JP, Macdonald DW (1995) Parasites of wild brown rats (*Rattus norvegicus*) on UK farms. *Parasitology*, **111**, 247–255.

This project was carried out under the direction of Dr. Adalgisa Caccone in the Yale Department of Ecology and Evolutionary Biology and Dr. Albert Ko at the Department of Epidemiology of Microbial Diseases, Yale School of Public Health. F.C., F.P., G.R., H.F., M.G.R., J.E.C., A.I.K. and A.C. designed the study. F.C., F.P. and G.R. performed sample collection and rat tissue preparation in Brazil. F.C. and F.P. also contributed to manuscript preparation. J.B. and C.H. helped with analyses, performing PCA and DAPC, and N_E/N_C estimates, as well as contributing to manuscript preparation. B.K., currently a Biology PhD student at City University of New York, performed all laboratory work and most of the analysis, in addition to preparing the manuscript. Her current research focuses on urban land- and stream-scape genomics of two salamander

species in New York City. C.H. is currently a graduate student at Cornell University, and his research focuses on understanding spatial and environmental correlates of genomic variation in a variety of taxa. F.C. is an eco-epidemiologist interested in the study of urban ecology and infectious diseases in the developing with the aim to implement and evaluate disease control programmes. G.R. and H.F. are biologists interested in rodent control to prevent zoonotic diseases. F.P. is a parasite ecologist interested in how reservoir host ecology and population dynamics impact the spatial and temporal distribution of zoonotic disease risk. J.B. is an evolutionary geneticist interested in disease vectors and pathogens, and their ecological interactions with humans. M.G.R. is a molecular epidemiologist interested in human pathology and immunopathogenesis of parasitological and infectious diseases including hepatitis C, dengue, meningitis, leptospirosis, chagas disease and schistosomiasis. J.E.C. is an infectious disease epidemiologist interested in the transmission of zoonotic pathogens to humans, including directly transmitted zoonotic viruses, such as the hantaviruses, arenaviruses and rabies, and vector-borne bacteria, including rickettsia, bartonella and borrelia. AIK is an infectious disease epidemiologist whose research focuses on the dynamics of health problems that have emerged due to rapid urbanization and social inequity and the interventions that are implemented in urban slum communities. A.C. is an evolutionary biologist interested in understanding spatial and temporal patterns and levels of genomic diversity in organisms of conservation or epidemiological concern.

Data accessibility

Microsatellite genotypes and analysis input files: DRYAD entry doi: 10.5061/dryad.hd300.

Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1 Methods.

Table S1 Pairwise geographic distance in meters between sites in Salvador, Brazil. On average sites are 2.7 km apart, but most sites are between 300 m and 1.7 km apart.

Table S2 Locus information with bold names representing the final loci selected for analyses. Locus name is provided in the first column followed by the reference, the size of the locus in base pairs (bp; with allele ranges provided when available). Loci referenced as “non-referenced by Robertson & Gemmel

2004;" were used in the 2004 publication without an original source cited. The fourth columns list the paired locus for multiplex analyses (Multiplex Partner) The The last three columns report the DNA sequence of the Forward (F-Primer) and reverse (R-primer) primers used for the PCR amplifications and the fluorescent dye used to label each locus (Label).

Table S3 AMOVA results by sampling year for all rat samples (temporal structure).

Table S4 Population pairwise F_{ST} values with samples temporally separated.

Table S5 Population pairwise F_{ST} values (lower diagonal) with

pairwise geographic distance in meters (upper diagonal) between sites.

Table S6 Population size (N_E) with 95% confidence intervals (CI) and census size (N_C) estimates with 95% highest probability density (HPD) intervals for clusters 1, 2, and 3. Estimates of N_E intervals were computed using the linkage disequilibrium method implemented by LDNE, while N_C estimates were calculated with a sequential Bayesian method (Petit & Valiere 2006).

Fig. S1 Plot of linearized pairwise F_{ST} [$F_{ST}/(1-F_{ST})$] over pairwise distance (m) between sites in Pau da Lima.