

## Genetic structure of black rat populations in a rural plague focus in Madagascar

Aude Gilabert, A. Loiseau, Jean-Marc Duplantier, S. Rahelinirina, L. Rahalison, S. Chanteau, Carine Brouat

► **To cite this version:**

Aude Gilabert, A. Loiseau, Jean-Marc Duplantier, S. Rahelinirina, L. Rahalison, et al.. Genetic structure of black rat populations in a rural plague focus in Madagascar. Canadian Journal of Zoology, NRC Research Press, 2007, 85 (9), pp.965-972. <10.1139/Z07-083>. <ird-00261411>

**HAL Id: ird-00261411**

**<http://hal.ird.fr/ird-00261411>**

Submitted on 6 Mar 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

This file is an author-produced PDF of an article accepted for publication:  
*Canadian Journal of Zoology*, 2007, 85 (9), p. 965-972

### **Genetic structure of black rat populations in a rural plague focus in Madagascar**

**A. Gilabert<sup>1a</sup>, A. Loiseau<sup>1</sup>, J.-M. Duplantier<sup>1</sup>, S. Rahelinirina<sup>2</sup>, L. Rahalison<sup>2</sup>, S. Chanteau<sup>2b</sup>, and C. Brouat<sup>1\*</sup>.**

<sup>1</sup> IRD (UR022) Centre de Biologie et de Gestion des Populations (CBGP), Campus International de Baillarguet, CS 30016, 34988 Montferrier-sur-Lez, cedex, France. [Aude.Gilabert@rennes.inra.fr](mailto:Aude.Gilabert@rennes.inra.fr); [loiseau@ensam.inra.fr](mailto:loiseau@ensam.inra.fr); [duplanti@ird.sn](mailto:duplanti@ird.sn); [brouat@mpl.ird.fr](mailto:brouat@mpl.ird.fr)

<sup>2</sup> Institut Pasteur de Madagascar, Unité Peste, BP 1274, Antananarivo, Madagascar. [rahelinirina@yahoo.fr](mailto:rahelinirina@yahoo.fr); [lrahalison@pasteur.mg](mailto:lrahalison@pasteur.mg)

\* Corresponding author: Carine Brouat, IRD-CBGP, Campus International de Baillarguet, CS 30016, 34988 Montferrier-sur-Lez, cedex, France. Phone : (33) 4 99 62 33 09. Fax : (33) 4 99 62 33 45. E-mail: [brouat@mpl.ird.fr](mailto:brouat@mpl.ird.fr)

---

<sup>a</sup> INRA-UMR Bio3P, Domaine de la Motte, 35653 Le Rheu Cedex, France; tel : 33 2 23 48 52 06; [Aude.Gilabert@rennes.inra.fr](mailto:Aude.Gilabert@rennes.inra.fr)

<sup>b</sup> Centre de Recherches Scientifiques et Sanitaires, BP 10887, Niamey, Niger. [chanteau@cermes.org](mailto:chanteau@cermes.org)

**A. Gilabert, A. Loiseau, J.-M. Duplantier, S. Rahelinirina, L. Rahalison, S. Chanteau, and C. Brouat**

**Genetic structure of black rat populations in a rural plague focus in Madagascar**

**Abstract:** Genetic structure of reservoir populations is a key characteristic in understanding the persistence of infectious diseases in natural systems. In the Central Highlands of Madagascar where plague persists since 1920, the black rat *Rattus rattus* L. (1758) is the sole species susceptible to act as a reservoir of the disease. Ecological surveys have shown a clear correlation between the location of the plague persistence area in Madagascar (above 800 m high) and the distribution area of one endemic plague vector, the flea *Synopsyllus fonquerniei*, which is exclusively found on rats living outdoors. This clear habitat segregation has led to the suggestion that *R. rattus* from Central Highlands might be structured into indoor and outdoor populations. Using eight microsatellite markers, we analysed the genetic structure of black rat populations living inside a human plague focus, in relationship with habitat and geographic distance. We found that habitat by itself was not a structuring factor, against geographic distance. Nevertheless, the significant genotypic differentiation of *R. rattus* populations found at a fine spatial scale might relate to different population dynamics between indoor and outdoor habitats.

**Résumé:** La structure génétique des populations de réservoirs est une caractéristique clef pour comprendre la persistance des maladies infectieuses en milieu naturel. Sur les Hautes Terres malgaches, où la peste persiste depuis 1920, le rat noir *Rattus rattus* L. (1758) est la seule espèce potentiellement réservoir de la maladie. Des études écologiques ont montré une corrélation claire entre la localisation de la zone de persistance de la peste à Madagascar (au-dessus de 800 m d'altitude) et l'aire de distribution d'une puce endémique vecteur de la peste, *Synopsyllus fonquerniei*, qui parasite exclusivement les rats vivant à l'extérieur des maisons. Une telle ségrégation en terme d'habitat a conduit à formuler l'hypothèse que les rats des Hautes Terres seraient structurés en populations d'extérieur et d'intérieur. En utilisant huit marqueurs microsatellites, nous avons analysé la structure génétique des populations naturelles de rats d'un foyer humain de peste en relation avec l'habitat et la distance géographique. Nous avons trouvé que l'habitat n'est pas un facteur expliquant la structure génétique, au contraire de la distance géographique. La différenciation génotypique significative des populations à une échelle spatiale fine pourrait en revanche indiquer des dynamiques de populations différentes des rats vivant à l'intérieur ou à l'extérieur des maisons.

## Introduction

The black rat *Rattus rattus* L. (1758) originated in Asia and now has a widespread global distribution, being arguably the most damaging invasive rodent on islands worldwide, with wide impacts on native faunas (Amori and Clout 2003). It is also known to cause serious problems in agriculture and public health (Gratz 1997).

*Rattus rattus* may have reached Madagascar with the first human immigrants about 2000 years ago, although its presence is only recognized between the 11th and 14th centuries in archaeological excavations (Rakotozafy 1996). Competition with the black rat is a major cause of extinction threats for Madagascar's endemic rodents (Goodman 1995). Several outbreaks have led the Malagasy State to declare *R. rattus* a public calamity, and agricultural pest managements of rodents are now state-controlled. With respect to public health, the major disease problem in Madagascar is the plague (caused by the bacillus *Yersinia pestis*), which appeared during the third pandemic (1898) and became endemic from 1921 in the central part of the island, the central Highlands, above 800 meters (Brygoo 1966; Chanteau et al. 1998). *Rattus rattus* is the unique species that can act as a plague reservoir in Madagascar (Brygoo 1966), representing more than 95% of rodent catches inside or outside human villages of the Highlands (Duplantier and Rakotondravony 1999).

Analysing the genetic population structure of a target species may provide useful data to identify management strategies (e.g., Krebs 1998). Spatial structure of reservoir populations was shown to play a role in the invasion and persistence of diseases (Hagenaars et al. 2004; Hess 1996; Park et al. 2001; Swinton et al. 1998), especially for plague epizootics (Keeling 1999; Keeling and Gilligan 2000). Indeed, intermediate levels of patchiness (or of genetic differentiation) are generally associated with a higher disease persistence at the metapopulation scale.

Despite the large distribution and the detrimental impact on invaded ecosystems and human health of the black rat, very few genetic studies have focused on natural populations of this species (Abdelkrim et al. 2005). The only existing population genetic analyses using microsatellite markers were performed in the context of Guadeloupe islands (Abdelkrim et al. 2005; Prugnolle et al. 2005). Ecological surveys of *R. rattus* populations in the Central Highlands of Madagascar have shown that rats living outdoor and indoor are not parasitized by the same flea species (Brygoo 1966; Duplantier et al. 2005). *Xenopsylla cheopis*, the oriental rat flea most frequently involved throughout the world in plague transmission (Gage and Kosoy 2005) is present only on rats living in houses, whereas the endemic flea *Synopsyllus fonquerniei* is exclusively found on rats living outdoors. The clear correlation between the distribution area of *S. fonquerniei* and the plague endemic area in Madagascar (above 800 m high) may suggest a key role for this flea in plague persistence. As fleas mostly spend their entire adult life on their host or in its burrow (Ioff 1941), the clear habitat segregation of fleas has led to the suggestion that *R. rattus* might be structured into indoor and outdoor populations (Duplantier et al. 2003). In this context and using microsatellite markers, the goal of this study was thus to estimate genetic structure relative to habitat or geographic distance in natural populations of *R. rattus* inside the plague endemic area of Madagascar.

## Materials and methods

### Study area and sample collection

The study area covers 120 km<sup>2</sup> in the Middle West of Madagascar, near the town of Mandoto (19°34'S; 46°17'E). In these central highlands (between 850 and 1100 m high), human plague epidemics were reported in several villages between 1994 and 1996 (Rahelinirina and Duplantier 1999). Sampling was carried out in six localities having comparable human and ecological characteristics (Rahelinirina and Duplantier 1999),

between July 1996 and January 1997 for most localities, and in June 1997 for Antsahalava (Fig. 1).

For each of the six localities (except in Mok where only houses were sampled), three trap sites corresponding to different habitat types were considered: inside the Houses (H) of the village, in Sisal fences (S) around enclosures at close proximity of houses (around 20 m) and in surrounding irrigated rice fields (R) distant of the houses from 100 to 500 meters (Fig. 1). Outdoors, lines of 20 traps were set with a 10 m interval between traps. In villages, only two traps per house were set, to minimise the capture of closely related individuals (parent–young or full siblings). Using wire-meshed live traps, 10 to 22 rats were collected per trap site, representing a total of 325 individuals. Animals were treated in a humane manner, and in accordance with guidelines of the American Society of Mammalogists (Animal Care and Use Committee 1998).

Trap sites, hereafter referred to as "subpopulations", were distant from each other's by at least 20 m (distance between H and S population inside the same locality) and at most 29 km (distance between the localities Mok and Aha). Considering 20-m distant samples as belonging to different subpopulations was justified by ecological data on fleas and by the need to test the hypothesis of a genetic structure into indoor and outdoor populations. The minimum distance between localities was 5.7 km (between the localities Ato and Bit). Geographic coordinates were only taken for village sites.

### **Molecular methods**

DNA was extracted from ear tissue using the PUREGENE DNA purification kit. Quantification of genetic variation was performed using eight microsatellites (D11Rat56, D5Rat83, D10Rat20, D16Rat81, D11Mgh5, D2Mit14, D7Rat13 and D18Rat75) characterised for *R. norvegicus* genome mapping (Jacob et al. 1995), and already used for *R. rattus* population genetics (Abdelkrim et al. 2005; Prugnolle et al. 2005). Two Multiplex Polymerase Chain Reactions (PCR) respectively amplified the D11Rat56, D5Rat83, and D10Rat20 loci and the D16Rat81, D11Mgh5, D2Mit14 and D7Rat13 loci. We used the following touchdown PCR conditions (using the QIAGEN Multiplex PCR master Mix): 15 min of initial denaturation at 95°C, 30 s of denaturation at 94°C, 1 min 30 s of annealing at 67°C (temperature decreased by 2°C each cycle for eight cycles then at 53°C for 33 cycles), 30 s of elongation at 72°C, and after the 41 cycles, 30 min of final elongation at 60°C. D18Rat75 was amplified alone by simple PCR (Taq QIAGEN) with an annealing temperature of 55°C and 34 cycles. All PCR products were pooled together in a single run, since fluorescent dyes were chosen to avoid overlapping of the different loci, on a MegaBace 1000 DNA sequencer (Amersham). Gel files were analysed using GENETIC PROFILER (Amersham Biosciences).

### **Population genetic structure**

Deviations from Hardy-Weinberg Equilibrium (HWE) and genotypic linkage disequilibrium were tested by the Markov chain method implemented in GENEPOP 3.4 (Raymond and Rousset 1995). In every case of multiple tests, significance levels were adjusted by using the method described in Kinnison et al. (2002). For each locus and each plot, we estimated the likelihood ( $L$ ) of getting as many or more significant tests as actually obtained by chance using the binomial likelihood function (Chapman et al. 1999)  $L = \sum_{i=r}^n (n! / (r! (n-r)!)) (1-\alpha)^{n-i} (\alpha)^i$ , where  $n$  is the total number of tests and  $r$  is the number of significant tests at a given  $\alpha$  level of significance. When  $L$  was  $< 0.05$ , the null hypothesis of such a result by chance was rejected, and the overarching hypothesis was considered to be supported.

As null genotypes were found for some individuals, the presence of null alleles was suspected. Every individual that was successfully genotyped at some loci but not at some others was re-amplified once by simple PCR (to avoid primer competition) for each failed

locus. We used MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004) to evaluate whether heterozygote deficiencies may be explained by the existence of null alleles. We then used the software FREENA (available at <http://www.montpellier.inra.fr/URLB>; Chapuis and Estoup 2007) to estimate null allele frequencies for each population and locus following Dempster et al. (1977), and to correct our genotype dataset for null alleles.

Mean numbers of alleles per locus, observed ( $H_O$ ) and expected ( $H_E$ ) heterozygosities (Nei 1987) were calculated over all loci at each sampling location using the software GENECLASS 2 (Piry et al. 2004) on the original data set, excluding null genotypes. The allelic richness ( $r$ , a measure of the number of alleles independent of sample size) was calculated using the rarefaction procedure implemented in FSTAT 2.9.3.2 (Goudet 2001) for a minimum sample size of 10 diploid individuals. Null alleles can result in an underestimation of statistics traditionally used to summarize genetic variation within populations. However,  $H_E$  and  $r$  are little affected by mean null allele frequencies below 0.15 (Chapuis 2006) such as those that we obtained (see results).

Genotypic divergence among populations for all loci and population pairs was tested using Markov chain methods in GENEPOP 3.4 (Raymond and Rousset 1995) on the original dataset. Null allele frequencies may conduct to an overestimation of population differentiation (Chapuis and Estoup 2007).  $F_{ST}$  were thus estimated following Weir (1996) using FREENA, with the so-called ENA (for Excluding Null Alleles) method described in Chapuis and Estoup (2007). This method was found to efficiently correct for the bias induced by null alleles and provide unbiased estimates of  $F_{ST}$ , whatever the mean null allele frequency.  $F_{ST}$  estimated with FREENA will be called hereafter  $F_{ST}^{ENA}$ . Ninety-five percent confidence intervals (CI) for mean  $F$ -statistics were generated by bootstrap resampling across loci.

Population structure was also investigated using the Bayesian model-based clustering approach implemented in STRUCTURE v. 2.2 (Pritchard et al. 2000). On the basis of multilocus data, this approach places individuals into  $K$  populations. Individuals can be assigned to multiple clusters with the membership coefficients of all the clusters summing up to one. The 2.2 release of STRUCTURE can accommodate for genotypic ambiguities in studies of population structure such as null alleles (Falush et al. 2007). We thus used the original dataset for these analyses, with null genotypes coded as homozygous for a recessive allele. We used the admixture model assuming correlated allele frequencies among populations. In each case STRUCTURE was run with three repetitions of 700,000 iterations following a burn-in period of 70,000 iterations, and with values of  $K$  from 1 to 16. The most appropriate number of populations was determined by comparing log-likelihoods ( $\ln \Pr(X/K)$ , where  $X$  is the number of genotypes). At their maximum, the corresponding value of  $K$  encompassed the main structure embodied in the data.

To estimate the part of the genetic variability that can be explained by localities and habitat types, we performed an AMOVA [analysis of molecular variance, (Excoffier et al. 1992)] considering the following two structures. The first one (Locality structure) was defined according to the six sampled localities. The second structure (Habitat structure) was defined according to the three habitat types (S, R and H). The variance components were tested using randomisation tests (1 000 permutations). These analyses were performed with the program ARLEQUIN version 2.000 (Schneider et al. 2000) using the locus-by-locus AMOVA option. Two four-level hierarchical  $F$ -statistics were also carried out using HIERFSTAT (Goudet 2005). For the first analysis, the four levels considered were: individual, subpopulation, locality and overall the sampling area. For the second analysis, the four levels considered were: individual,

subpopulation, habitat type and overall the sampling area. Significance of these different levels of differentiation was tested using randomisation tests (1 000 permutations per test).

Under a model of isolation by distance, genetic distance between subpopulations is expected to increase with geographical distance. Isolation by distance (IBD) was analysed by regressing pairwise estimates of  $F_{ST}^{ENA} / (1 - F_{ST}^{ENA})$  against  $\ln(\text{distance})$  between sampled sites (Rousset 1997). Mantel tests were performed to test the correlation between matrices of genetic differentiation and geographical distance between sampled subpopulations using GENEPOP 3.4. (10 000 permutations) (Raymond and Rousset 1995), excluding intra-locality comparisons.

## Results

Only three loci (D10Rat20, D11Rat56 and D7Rat13) were at HWE. The others showed significant heterozygote deficiencies in several subpopulations, probably due to null alleles (Table 1). Estimated null allele frequencies per subpopulation and locus ranged from 0 to 0.83 (Table 1), the mean frequency of null alleles per subpopulation and locus being 0.12.

Of the 448 exact tests performed for genotypic disequilibria, 43 were significant at the 0.05 level. Only six tests remained significant after correction for multiple tests. Significant values involved different pairs of loci and occurred in different subpopulations.

### Intrapopulation genetic diversity

All loci were polymorphic: the number of alleles per locus ranged from 9 to 23. Allelic richness per locus ranged from 2 to 8.2 (mean allelic richness per locus and subpopulation =  $4.30 \pm 1.17$ ). Observed and expected heterozygosities ranged from 0.11 to 1 for  $H_O$  (mean value per locus and subpopulation =  $0.63 \pm 0.21$ ) and from 0.2 to 0.96 for  $H_E$  (mean value per locus and subpopulation =  $0.70 \pm 0.14$ ) (see table 2 for mean values per subpopulation).

### Population structure

Microsatellites revealed significant genotypic differentiation among subpopulations ( $P < 0.0001$  for each locus). Ninety-three % (and within localities, 80 %) of the subpopulations pairs differed by more loci than expected by chance (Table 3). Pairwise  $F_{ST}^{ENA}$  estimates ranged from 0.004 (AhiS and AhiR, belonging to the same locality) and 0.13 (AhiH and AhaH, separated by 21.38 km). The overall  $F_{ST}^{ENA}$  estimate was 0.057 (CI = [0.042; 0.072]). Within localities, mean  $F_{ST}^{ENA}$  ranged from  $0.03 \pm 0.01$  between S and H habitats to  $0.04 \pm 0.02$  between R and H habitats.

Using the program STRUCTURE, five different genetic clusters could be identified (i.e.  $\ln \text{Pr}(X/K)$  was maximum with  $K = 5$  clusters). Subpopulations were heterogeneous regarding these five clusters (Fig. 1), with individuals consisting of a mixture of five clusters, thus representing pools of diverse genetic ancestries. Generally, one of the clusters obtained for a given value of  $K$  is split into two clusters for  $K + 1$  and genetic differentiation was more or less organized along a geographical gradient from the East to the West. In their order of appearance, one first cluster (in white on Fig. 1) contributed from 40 % to 75 % to the easternmost subpopulations of Aha, and nevermore than 22 % to the subpopulations of other localities. A second cluster (in yellow on Fig. 1) contributed  $> 50$  % to subpopulations of Ato. One third cluster (in black on Fig. 1) contributed to 65 % to the westernmost subpopulation of MokH, and nevermore than 18% to the easternmost subpopulations of Aas, Bit, Ato and Aha. The fourth and fifth clusters (in orange and red on Fig. 1) contributed together  $> 50$  % to the central subpopulations of Aas, Ahi and Bit.

The AMOVA indicated that populations of rats were more structured according to locality than according to habitat type (Table 4). In agreement with these results, hierarchical  $F$ -statistics analyses showed that significant genetic differentiation was observed between localities ( $F_{\text{locality/total}} = 0.048$ ,  $p = 0.001$ ) and between subpopulations within localities ( $F_{\text{population/locality}} = 0.037$ ,  $p = 0.001$ ) but that subpopulations were not structured by habitat type ( $F_{\text{habitat/total}} = -0.006$ ,  $p = 0.95$ ).

A weak but significant isolation-by-distance pattern was observed throughout the whole sampling area (Fig. 2; Mantel test,  $p = 0.007$ ).

## Discussion

The goal of this study was to examine the pattern of genetic structure across populations of *R. rattus* in a human plague focus of Madagascar, and to evaluate the occurrence of a structure into indoor and outdoor populations. We found a significant genotypic differentiation of *R. rattus* populations at a relatively small spatial scale. However, habitat by itself was not a structuring factor, against geographic distance.

Null alleles were found in our dataset. Their occurrence can be accorded to the fact that microsatellites were first developed for *Rattus norvegicus*. Null alleles are frequent in cross-priming experiments, because of divergence time between species, leading to mutations in the flanking microsatellite regions and thus poor primer annealing (Paetkau and Strobeck 1995). None of the authors who studied population genetics of *R. rattus* using the same microsatellite markers (Abdelkrim et al. 2005; Prugnolle et al. 2005) found any null alleles. In our case, null alleles could be attributed to PCR failure due to inconsistent DNA template quality, as the samples that we used were not conserved in ideal conditions since 1996. Alternatively, they may illustrate the divergence between Malagasy and Guadeloupean *R. rattus* natural populations. Null alleles may overestimate population differentiation by reducing the estimates of genetic diversity within populations (e.g., Paetkau and Strobeck 1995; Chapuis and Estoup 2007). We have thus carefully taken into account all the possible bias relative to null alleles, particularly by using recent methods developed to account for null alleles in genetic analyses.

Levels of genetic diversity observed for *R. rattus* in Madagascar were a little higher than those obtained by Prugnolle et al. (2005) (between five and twelve alleles per locus) and especially by Abdelkrim et al. (2005) (between two and six alleles per locus) in small and large islands of Guadeloupe. These differences were expected, according to different factors such as the size of the considered islands, or their colonization time by *R. rattus*, expected to be more recent for the Guadeloupe Archipelago (probably during the 17<sup>th</sup> century: see references in Abdelkrim et al. 2005) than for Madagascar.

The overall level of differentiation observed in the Mandoto area was similar with that found by Prugnolle et al. (2005) at an equivalent spatial scale in Guadeloupe ( $F_{\text{ST}} = 0.058$ ). Demographic analyses have already shown limited dispersal distances and home ranges in other parts of the world for *R. rattus* (e.g., Dowding and Murphy 1994; Cox et al. 2000). The genetic differentiation observed in the plague endemic area between populations distant from 29 km at the most is in accordance with the expectations of theoretical models explaining disease persistence (Hagenaars et al. 2004; Hess 1996; Keeling and Gilligan 2000). Intermediate levels of genetic structure such as those that we found should allow disease transmission between populations (disease persistence), but also recolonisation of extinct host population patches via healthy rats (host persistence).



Understanding how genetic differentiation between populations varies with geographical distance can help to determine whether genetic differentiation is primarily due to limited dispersal or to more complex demographic processes (e.g. Leblois et al. 2000). At mutation–migration–drift equilibrium, and for species with limited dispersal in space, genetic differentiation is expected to increase with geographical distance (Slatkin 1993; Rousset 1997). *Rattus rattus* Malagasy populations conformed to these theoretical expectations. The observed pattern of isolation by distance suggests that gene flow is limited to small distances in this species, explaining the relationship between genetic structure and locality. The points representing subpopulation pairs were however clearly scattered (Fig. 2), suggesting a relative importance of genetic drift, together with migration, to explain genetic differentiation (Hutchison and Templeton 1999).

Preliminary results based on parasitological data had suggested that *R. rattus* might be structured into indoor and outdoor subpopulations. We showed here that habitat type was not by itself a structuring factor for *R. rattus* populations. Nevertheless, genetic differences between trap sites reflected the existence of different genetic demes at the scale of the locality. Differentiation may be transient and produced by extinction–recolonisation dynamics through founder events (McCaughley 1993; Squire and Newman 2002). Density differences between indoor and outdoor populations might explain why they were genetically differentiated even nearby, as spatial variability in abundance may lead to local heterogeneity in gene flow (Berthier et al. 2005). Indeed, densities are expected to be higher in indoor populations of rodents (Pocock et al. 2005), due to resource permanence and environmental stability that lead to continuous reproduction all over the year inside houses (Duplantier and Rakotondravony 1999). Also, habitat specific demographic rates (Kreuzer and Huntly 2003) might be responsible for a source-sink population structure (Pulliam 1988), with stable populations of *R. rattus* inside the villages, and seasonally unstable populations outdoor that persist by continuing immigration from source populations that offset local mortality, or even repeated founder effects. Plague epizootics might reinforce the system as some indices led us to think that the disease would be more present in outdoor populations: ecological and serological surveys showed for example that rats from sisal fences had more fleas and were more often plague seropositive than those from houses (Rahelinirina and Duplantier 1999). As the ability to detect unidirectional dispersal using genetic markers, particularly under non-equilibrium conditions, is limited (Rousset 2001; see Pope et al. 2005), temporal surveys are needed to ascertain the demographic links between outdoor and indoor populations. For ethical reasons, capture-mark-recapture studies are not possible regarding the risk for human health to release potentially plague-infected rodents, but follow-up studies using marked baits are currently initiated, and will also permit to better quantify dispersal between habitats.

## Acknowledgments

The authors thank the technicians of the Plague and Entomology Laboratories of the Institut Pasteur of Madagascar who provided help in the fields, and M.-P. Chapuis and J. Foucaud for their advices concerning some statistical analyses. Development of microsatellite markers on *R. rattus* were performed by M. Galan and J.-F. Cosson. The manuscript benefited from the helpful comments of K. Berthier and J. Foucaud.

## References

- Abdelkrim, J., Pascal, M., and Samadi, S. 2005. Island colonization and founder effects: the invasion of the Guadeloupe islands by ship rats (*Rattus rattus*). *Mol. Ecol.* **14**: 2923–2931.
- Amori, G., and Clout, M. 2003. Rodents on islands: a conservation challenge. *In* Rats, mice and people: rodent biology and management. *Edited by* G.R. Singleton, L.A. Hinds, C.J.

- Krebs and D.M. Spratt. Australian Centre for International Agricultural Research, Canberra, Australia pp. 63–68.
- Animal care and use committee. 1998. Guidelines for the capture, handling, and care of mammals as approved by the American Society of Mammalogists. *J. Mammal.* **79**:1416–1431.
- Berthier, K., Galan, M., Foltete, J.C., Charbonnel, N., and Cosson, J.F. 2005. Genetic structure of the cyclic fossorial water vole (*Arvicola terrestris*): landscape and demographic influences. *Mol. Ecol.* **14**: 2861-2872.
- Brygoo, E.R. 1966. Epidémiologie de la peste à Madagascar. *Arch. Inst. Pasteur Mad.* **35**: 9-149.
- Chanteau S, Ratsifasoamanana L., Rasoamanana B., Rahalison L., Randriambeloso J., Roux J., and Rabeson D. 1998. Plague, a re-emerging disease in Madagascar. *Emerg. Infect. Dis.* **4**: 101-104.
- Chapman, R.W., Sedberry, G.R., Koenig, C.C., and Eleby, B.M. 1999. Stock Identification of Gag, *Mycteroperca microlepis*, along the Southeast Coast of the United States. *Mar. Biotech.* **1**: 137-146.
- Chapuis, M.P. 2006. Génétique des populations d'un insecte pullulant, le criquet migrateur, *Locusta migratoria*. PhD thesis, Montpellier II University, Montpellier.
- Chapuis, M.P., and Estoup, A. 2007. Microsatellite null alleles and estimation of population differentiation. *Mol. Biol. Evol.* **24**: 621-631.
- Cox, M. P. G., Dickman, C.R., and Cox, W.G. 2000. Use of habitat by the black rat (*Rattus rattus*) at North Head, New South Wales: an observational and experimental study. *Austral. Ecol.* **5**: 375–385.
- Dempster, A.P., Laird, N.M., and Rubin, D.B. 1977. Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc. B* **39**: 1-38
- Dowding, J.E., and Murphy, E.C. 1994. Ecology of Ship Rats (*Rattus-Rattus*) in a Kauri (*Agathis-Australis*) Forest in Northland, New-Zealand. *N. Z. J. Ecol.* **18**:19–28.
- Duplantier, J.-M., and Rakotondravony, D. 1999. The rodent problem in Madagascar: agricultural pest and threat to human health. *In Ecologically-based management of rodent pests. Edited by G. R. Singleton, L. A. Hinds, H. Leirs and Z. ZhiBin.* ACIAR, Canberra, pp. 441-459.
- Duplantier, J.-M., Duchemin, J.-B., Chanteau, S., and Carniel, E. 2005. From the recent lessons of the Malagasy foci towards a global understanding of the factors involved in plague reemergence. *Vet. Res.* **36**: 437-453.
- Duplantier, J.-M., Catalan, J., Orth, A., Grolleau, B., and Britton-Davidian, J. 2003. Systematics of the black rat in Madagascar: consequences for the transmission and distribution of plague. *Biol. J. Linn. Soc.* **78**: 335-341.
- Excoffier, L., Smouse, P.E., and Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**: 479-491.
- Falush, D., Stephens, M., and Pritchard, J.K. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol. Ecol. Notes*, onlineEarly Article: doi:10.1111/j.1471-8286.2007.01758.x.
- Gage, K.L., and Kosoy, M.Y. 2005. Natural History of plague: perspectives from more than a century of research. *Annu. Rev. Entomol.* **50**: 505-528.
- Goodman, S.M. 1995. *Rattus* on Madagascar and the dilemma of protecting the endemic rodent fauna. *Conserv. Biol.* **9**: 450-453.
- Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from <http://www.unil.ch/izea/software/fstat.html>.

- Goudet, J. 2005. HIERFSTAT, a package for R to compute and test hierarchical  $F$ -statistics. *Mol. Ecol. Notes*, **5**: 184-186.
- Gratz, N. 1997. The burden of rodent-borne diseases in Africa South of the Sahara. *Belg. J. Zool. (suppl.)* **127**: 71-84.
- Hagenaars, T.J., Donnelly, C.A., and Ferguson, N.M. 2004. Spatial heterogeneity and the persistence of infectious diseases. *J. Theor. Biol.* **229**: 349-359.
- Hess, G. 1996. Disease in metapopulation models: implications for conservation. *Ecology*, **77**: 1617-1632.
- Hutchinson, D.W., and Templeton, A.R. 1999. Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, **53**: 1898-1914.
- Ioff, I.G. 1941. *Ecology of fleas in relevance to their medical importance*. Pyatyorsk Publishers, Pyatyorsk, Russia.
- Jacob, H. J., Brown, D.M., Bunker R.K., Daly, M.J., Dzau, V.J., Goodman, A., Koike, G., Kren, V., Kurtz, T., Lernmark, A., Levan, G., Mao, Y., Pettersson, A., Pravenec, M., Simon, J.S., Szpirer, C., Szpirer, J., Trolliet, M.R., Winer, E.S., Lander, E.S. 1995. A genetic linkage map of the laboratory rat, *Rattus norvegicus*. *Nature Genetics*, **9**: 63-69.
- Keeling, M.J. 1999. The effects of local spatial structure on epidemiological invasions. *Proc. R. Soc. Lond. B Biol. Sci.* **266**: 859-867.
- Keeling, M.J., and Gilligan, C.A. 2000. Metapopulation dynamics of bubonic plague. *Nature*, **407**: 903-906.
- Kinnison, M.T., Bentzen, P., Unwin, M.J., and Quinn, T.P. 2002. Reconstructing recent divergence: evaluating nonequilibrium population structure in New Zealand chinook salmon. *Mol. Ecol.* **11**: 739-754.
- Krebs, C.J. 1998. Wither small rodent population studies? *Res. Popul. Ecol.* **40**: 123-125.
- Kreuzer, M.P., and Huntly, N.J. 2003. Habitat-specific demography: evidence for source-sink population structure in a mammal, the pika. *Oecologia*, **134**: 343-349.
- Leblois, R., Rousset, F., Tikel, D., Moritz, C., and Estoup, A. 2000. Absence of evidence for isolation by distance in an expanding cane toad (*Bufo marinus*) population: an individual-based analysis of microsatellite genotypes. *Mol. Ecol.* **9**: 1905-1909.
- McCaughley, D. E. 1993. Evolution in metapopulations with frequent local extinction and recolonization in *In Oxford Surveys in Evolutionary Biology*. Edited by D. Futuyma and J. Antonovics. Oxford University Press, Oxford, pp. 109-134.
- Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New-York.
- Paetkau, D., and Strobeck, C. 1995. The molecular basis and evolutionary history of a microsatellite null allele in bears. *Mol. Ecol.* **4**: 519-520.
- Park, A.W., Gubbins, S., and Gilligan, C.A. 2001. Invasion and persistence of plant parasites in a spatially structured host population. *Oikos*, **94**: 162-174.
- Piry, S., Alapetite, A., Cornuet, J.-M., Paetkau, D., Baudouin, D., and Estoup, A. 2004. GENECLASS2: a software for genetic assignment and first-generation migrant detection. *J. Hered.* **95**: 536-539.
- Pocock, M.J.O., Hauffe, H.C., and Searle, J.B. 2005. Dispersal in house mice. *Biol. J. Linn. Soc.* **84**: 565-583.
- Pope, L.C., Blair, D., and Johnson, C.N. 2005. Dispersal and population structure of the rufous bettong, *Aepyprymnus rufescens* (Marsupialia: Potoroidae). *Austral Ecol.* **30**: 572-580.
- Pritchard, J.K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics*, **155**: 945-959

- Prugnolle, F., Théron, A., Pointier, J.-P., Jabbour-Zahab, R., Jarne, P., Durand, P., and de Meeûs, T. 2005. Dispersal in a parasitic worm and its two hosts: consequence for local adaptation. *Evolution*, **59**: 296-303.
- Pulliam, H.R. 1988. Sources, sinks, and population regulation. *Am. Nat.* **132**: 653–661.
- Rakotozafy, L.M.A. 1996. Etude de la constitution du régime alimentaire des habitants du site de Mahilaka du Xe au XIVE siècle à partir des produits de fouilles archéologiques. DPhil Thesis, Antananarivo University, Madagascar.
- Rahelinirina, S., and Duplantier, J.-M. 1999. Dynamique des populations de rats noirs (*Rattus rattus*) dans les foyers de peste de la région de Mandoto. *In Actes du Symposium national sur les rongeurs et la lutte antimurine*, DPV et GTZ, Antananarivo, Madagascar, pp. 91-105.
- Raymond, M., and Rousset, F. 1995. Genepop (version 1.2., population genetics software for exact tests and ecumenicism. *J. Hered.* **86**: 248-249.
- Rousset, F. 1997. Genetic differentiation and estimation of gene flow from  $F$ -Statistics under isolation by distance. *Genetics*, **145**: 1219-1228.
- Rousset, F. 2001. Genetic approaches to the estimation of dispersal rates. *In Dispersal. Edited by J. Clobert, E. Danchin, A.A. Dhondt and J.D. Nichols.* Oxford University Press, Oxford, pp. 18–28.
- Schneider, S., Roessli, D., and Excoffier, L. 2000. Arlequin Ver 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory, Dept. of Anthropology, University of Geneva, Switzerland.
- Slatkin, M. 1993. Isolation by distance in equilibrium and nonequilibrium populations. *Evolution*, **47**: 264–279.
- Squire, T., and Newman, R.A. 2002. Fine-scale population structure in the wood frog (*Rana sylvatica*) in a northern woodland. *Herpetologica*, **58**: 119-130.
- Swinton, J., Harwood, J., Grenfell, B.T., and Gilligan, C.A. 1998. Persistence thresholds for phocine distemper virus infection in harbour seal *Phoca vitulina* metapopulations. *J. Anim. Ecol.* **67**: 54-68.
- Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., and Shipley, P. 2004. Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes*, **4**: 535-538.
- Weir, B. S. 1996. Genetic Data Analysis II. Sinauer Associates, Sunderland, Mass.

**Table 1.** Estimates of null allele frequencies for loci having heterozygote deficiencies in *R. rattus* populations, and mean null allele deficiency ( $\bar{a}$ ) per locus.

	D5Rat83	D16Rat81	D11Mgh5	D2Mit14	D18Rat75
AtoS	0.18	0.05	0	0	0.03
AtoR	0.04	0.12	0	0	0
AtoH	0.23	0.08	0	0.11	0.09
AasS	0.27	0.01	0.09	0	0.07
AasR	0.23	0.12	0.30	0	0.09
AasH	0.29	0.06	0	0.04	0.15
AhiS	0.22	0.12	0.61	0.15	0.25
AhiR	0.24	0.13	0.28	0.19	0.08
AhiH	0.46	0.15	0.58	0.21	0.14
AhaS	0.02	0	0	0	0.12
AhaR	0.10	0.02	0.03	0.02	0.17
AhaH	0.03	0.02	0	0	0.19
BitS	0.71	0.19	0	0	0.24
BitR	0.54	0.14	0	0	0.20
BitH	0.39	0.09	0.12	0	0.10
MokH	0.83	0.08	0.23	0	0.43
$\bar{a}$	0.30	0.09	0.14	0.05	0.15

**Table 2.** Population polymorphism (mean  $\pm$  standard deviation) across all loci for the 16 subpopulations sampled.  $N$  is the number of individuals,  $n$  the number of alleles,  $r$  the allelic richness, and  $H_O$  and  $H_E$  the observed and expected heterozygosities.

Suppopulation	$N$	$n$	$r$	$H_E$	$H_O$
AtoS	20	7.0 $\pm$ 2.3	4.3 $\pm$ 1.2	0.69 $\pm$ 0.20	0.68 $\pm$ 0.16
AtoR	22	6.6 $\pm$ 2.2	4.3 $\pm$ 1.0	0.70 $\pm$ 0.21	0.70 $\pm$ 0.13
AtoH	22	6.9 $\pm$ 3.0	4.0 $\pm$ 1.2	0.57 $\pm$ 0.19	0.64 $\pm$ 0.18
AasS	22	7.4 $\pm$ 2.4	4.6 $\pm$ 1.3	0.64 $\pm$ 0.21	0.71 $\pm$ 0.14
AasR	21	6.0 $\pm$ 1.9	4.1 $\pm$ 1.1	0.54 $\pm$ 0.27	0.67 $\pm$ 0.17
AasH	22	7.8 $\pm$ 3.2	4.6 $\pm$ 1.4	0.69 $\pm$ 0.24	0.73 $\pm$ 0.13
AhiS	22	7.1 $\pm$ 1.4	4.6 $\pm$ 1.2	0.65 $\pm$ 0.20	0.73 $\pm$ 0.12
AhiR	16	7.5 $\pm$ 3.0	4.7 $\pm$ 1.6	0.66 $\pm$ 0.20	0.75 $\pm$ 0.12
AhiH	22	5.9 $\pm$ 1.7	4.0 $\pm$ 1.1	0.50 $\pm$ 0.22	0.67 $\pm$ 0.15
AhaS	22	5.6 $\pm$ 1.9	3.9 $\pm$ 1.0	0.66 $\pm$ 0.18	0.69 $\pm$ 0.11
AhaR	19	6.5 $\pm$ 1.9	4.3 $\pm$ 0.7	0.68 $\pm$ 0.18	0.73 $\pm$ 0.08
AhaH	22	6.6 $\pm$ 1.8	4.2 $\pm$ 0.9	0.67 $\pm$ 0.13	0.70 $\pm$ 0.11
BitS	22	5.6 $\pm$ 2.4	4.0 $\pm$ 1.1	0.58 $\pm$ 0.21	0.68 $\pm$ 0.12
BitR	10	6.4 $\pm$ 2.8	4.8 $\pm$ 1.6	0.64 $\pm$ 0.28	0.75 $\pm$ 0.19
BitH	22	7.8 $\pm$ 2.8	4.5 $\pm$ 1.2	0.58 $\pm$ 0.12	0.70 $\pm$ 0.13
MokH	19	6.1 $\pm$ 2.9	4.2 $\pm$ 1.6	0.58 $\pm$ 0.25	0.63 $\pm$ 0.25

**Table 3.** Pairwise  $F_{ST}^{ENA}$  values calculated for all loci (below the diagonal) and counts of significant ( $p < 0.05$ ) genotypic tests between samples (above the diagonal). With eight microsatellites, likelihood of two or more significant tests is 0.054 and likelihood of three or more is 0.0057. Significant tests are in bold.

pop	AtoS	AtoR	AtoH	AasS	AasR	AasH	AhiS	AhiR	AhiH	AhaS	AhaR	AhaH	BitS	BitR	BitH	MokH
AtoS		<b>6</b>	<b>3</b>	<b>6</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>4</b>	<b>8</b>	<b>8</b>	<b>6</b>	<b>8</b>	<b>8</b>	<b>6</b>	<b>6</b>	<b>8</b>
AtoR	0.058		<b>4</b>	<b>8</b>	<b>6</b>	<b>6</b>	<b>6</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>7</b>	<b>8</b>	<b>7</b>	<b>6</b>	<b>5</b>	<b>7</b>
AtoH	0.019	0.048		<b>7</b>	<b>6</b>	<b>5</b>	<b>5</b>	<b>3</b>	<b>7</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>8</b>	<b>4</b>	<b>4</b>	<b>8</b>
AasS	0.044	0.069	0.053		<b>4</b>	<b>2</b>	<b>5</b>	<b>3</b>	<b>8</b>	<b>6</b>	<b>6</b>	<b>7</b>	<b>7</b>	<b>2</b>	<b>5</b>	<b>8</b>
AasR	0.064	0.081	0.082	0.028		<b>3</b>	<b>4</b>	<b>3</b>	<b>7</b>	<b>8</b>	<b>6</b>	<b>8</b>	<b>7</b>	<b>4</b>	<b>6</b>	<b>7</b>
AasH	0.062	0.070	0.066	0.009	0.015		<b>2</b>	<b>2</b>	<b>7</b>	<b>7</b>	<b>7</b>	<b>6</b>	<b>6</b>	<b>2</b>	<b>4</b>	<b>6</b>
AhiS	0.060	0.070	0.063	0.019	0.016	0.014		<b>2</b>	<b>5</b>	<b>7</b>	<b>6</b>	<b>6</b>	<b>5</b>	<b>4</b>	<b>6</b>	<b>6</b>
AhiR	0.051	0.057	0.052	0.020	0.016	0.015	0.004		<b>5</b>	<b>7</b>	<b>6</b>	<b>6</b>	<b>6</b>	<b>2</b>	<b>4</b>	<b>7</b>
AhiH	0.081	0.096	0.079	0.071	0.067	0.075	0.038	0.052		<b>7</b>	<b>8</b>	<b>8</b>	<b>7</b>	<b>6</b>	<b>7</b>	<b>7</b>
AhaS	0.069	0.061	0.061	0.050	0.062	0.049	0.059	0.049	0.112		<b>7</b>	<b>6</b>	<b>7</b>	<b>6</b>	<b>7</b>	<b>8</b>
AhaR	0.067	0.062	0.072	0.046	0.061	0.051	0.038	0.039	0.091	0.050		<b>6</b>	<b>7</b>	<b>7</b>	<b>6</b>	<b>8</b>
AhaH	0.082	0.069	0.066	0.096	0.108	0.084	0.086	0.069	0.132	0.029	0.066		<b>6</b>	<b>6</b>	<b>7</b>	<b>7</b>
BitS	0.074	0.074	0.064	0.031	0.058	0.025	0.035	0.029	0.093	0.054	0.049	0.078		<b>3</b>	<b>6</b>	<b>5</b>
BitR	0.057	0.053	0.066	0.022	0.032	0.008	0.041	0.032	0.087	0.045	0.067	0.075	0.030		<b>2</b>	<b>6</b>
BitH	0.062	0.070	0.058	0.027	0.071	0.034	0.053	0.047	0.086	0.064	0.065	0.088	0.041	0.019		<b>6</b>
MokH	0.097	0.083	0.092	0.059	0.067	0.050	0.056	0.047	0.069	0.092	0.106	0.121	0.055	0.046	0.054	

**Table 4.** Results of AMOVA performed according to two different grouping factors: locality and habitat type. The variance components ( $Va$ ,  $Vb$ ,  $Vc$ ), the percentage of variation and the associated probability are provided for the three levels of each structure: between groups, between subpopulations within groups and within subpopulations.

Grouping factor	between groups			between subpopulations within groups			within subpopulations		
	$Va$	% of variation		$Vb$	% of variation		$Vc$	% of variation	
		$p$			$p$			$p$	
Locality	0.025	1.82	0.005	0.05	3.73	< 0.001	1.28	94.45	< 0.001
Habitat type	0.002	0.18	0.35	0.07	5.20	< 0.001	1.28	94.62	< 0.001



## Figure captions

**Fig. 1.** Location of the study area, of sampled localities, and pie charts of combined genetic ancestries of all individuals sampled in each subpopulation, as obtained from STRUCTURE (Pritchard *et al.* 2000). On the map of Madagascar, black lines delimited the plague endemic area. The different levels of grey indicated increasing altitude levels per steps of 200 m high, from 0-200 high to above 800 m high. The white circles represented the area of Mandoto on the map of Madagascar, and the sampled localities (with their name and their code in brackets). The black circles represented villages that were not sampled in the studied area. The box around each pie chart indicated habitat as follows: black for sisal (S), grey for rice plantations (R), and white for houses (H). The five different colours inside each pie chart corresponded to five different genetic clusters.

**Fig. 2.** Relationships between genetic differentiation [estimated as  $F_{ST}^{ENA} / (1 - F_{ST}^{ENA})$ ] and geographic distance between sampled subpopulations excluding intra-locality comparisons.

Fig. 1. Gilabert et al.



