MOLECULAR ECOLOGY

Invasion genetics of the introduced black rat (*Rattus rattus*) in Senegal, West Africa

ADAM KONEČNÝ,*†‡ ARNAUD ESTOUP,* JEAN-MARC DUPLANTIER,§ JOSEF BRYJA,†
KHALILOU BÂ,¶ MAXIME GALAN,* CAROLINE TATARD* and JEAN-FRANCOIS COSSON*

INRA, CBGP, Campus international de Baillarguet, CS 30016, Montferrier-sur-Lez cedex F-34988, France, †Institute of Vertebrate Biology, Academy of Sciences of the Czech Republic, Květná 8, Brno 603 65, Czech Republic, ‡Department of Biodiversity and Molecular Ecology, Fondazione E. Mach, Research and Innovation Centre, Via E. Mach 1, San Michele all' Adige (TN) 38010, Italy, §IRD, CBGP, Campus international de Baillarguet, CS 30016, Montferrier-sur-Lez cedex F-34988, France, ¶IRD, CBGP, BP 1386, Dakar, Senegal

Abstract

An understanding of the evolutionary history and dynamics of invasive species is required for the construction of predictive models of future spread and the design of biological management measures. The black rat (*Rattus rattus*) is a major vertebrate invader with a worldwide distribution. Despite the severe ecological, economic and health impacts of this species, its evolutionary history has been little studied. We carried out extensive specimen sampling in Senegal, West Africa, and used microsatellite markers to describe the pattern and processes of invasion in this large continental area. The genetic data obtained were combined with historical knowledge concerning the presence of this species in Senegal. Data were analysed by a combination of Bayesian clustering and approximate Bayesian computation methods. The invasion pathways closely paralleled the history of human trade routes in Senegal. In several places, we detected the occurrence of multiple introductions from genetically different sources. Long-distance migration between towns and villages was also observed. Our findings suggest that genetic bottlenecks and admixture have played a major role in shaping the genetics of invasive black rats. These two processes may generate genetic novelty and favour rapid evolution along the invasion pathways.

Keywords: approximate Bayesian computation, Bayesian clustering, founder effects, genetic admixture, ioinvasion, microsatellites, multiple introductions

Received 14 July 2012; revision received 19 September 2012; accepted 24 September 2012

Introduction

The rate of human-mediated biological invasions has accelerated in the last few centuries, largely due to the expansion of international trade and transport (Mack *et al.* 2000). The spread of non-native species is one of the most important effects of human activities, with many negative consequences for global and local biodiversity, ecosystem functioning, agriculture, fisheries and public health (Vitousek *et al.* 1997; Wilcove *et al.* 1998). Despite the increasing importance of bioinvasions, we still know little about the evolutionary mechanisms

Correspondence: Jean-François Cosson, Fax: + 33 4 99 62 33 45; e-mail: cosson@supagro.inra.fr

underlying the success of invasive species (Facon *et al.* 2006). Evolutionary processes seem to play a crucial role during the various stages of invasion: opportunity/ transport, establishment and spread (Suarez & Tsutsui 2008). An understanding of these evolutionary processes is critical for both the theoretical aspects of evolutionary and population biology and, from a practical standpoint, the construction of predictive models of future spread and the design of biological control measures (Lee 2002).

When individuals are introduced into a new range, they are likely to experience ecological conditions very different from those in the region from which they originate. Natural selection and adaptation may therefore play a crucial role in the outcome of the invasion

process (Facon *et al.* 2006). Genetic variability is a key factor determining the capacity of invasive populations to respond to selection (Lee 2002; Dlugosch & Parker 2008; Suarez & Tsutsui 2008). Substantial genetic variability for ecologically important traits would be expected to favour adaptive evolution in response to new environmental conditions (Reed & Frankham 2001). Alternatively, good combinations of genotypes for some specific genes may also facilitate invasion, even if genetic variation is low (Facon *et al.* 2006). In turn, the genetic features of invasive populations (in terms of both variability and composition) depend on the geographic and demographic characteristics of their introduction. These characteristics include the number of introduction events, the number of individuals introduced and the demographic dynamics of invading populations. For instance, successive founder events and random genetic drift along an introduction route may cause a substantial loss of genetic variation within an invasive population. However, this loss may be readily counterbalanced by the genetic admixture of populations introduced from different sources (Kolbe *et al.* 2004; Lavergne & Molofsky 2007). A reconstruction of invasion pathways is thus required to determine and test the relative importance of such demographic features, and of other environmental and evolutionary factors in biological invasions (reviewed in Estoup & Guillemaud 2010).

Invasion routes have traditionally been reconstructed from historical and observational data, which are often sparse and incomplete. Recent advances in population genetics have made it possible to develop statistical inferences concerning invasion dynamics, the direction of invasion pathways and source populations, through the combination of genetic data with other sources of information (e.g. historical records). Population genetic analyses, including the Bayesian genetic clustering analyses implemented in STRUCTURE (Pritchard *et al.* 2000) in particular, are necessary when samples are taken from many sites (Ascunce *et al.* 2011). The model-based approaches underlying approximate Bayesian computation (ABC) can be used to consider complex evolutionary scenarios that cannot be treated by methods based entirely on maximum likelihood equations (Beaumont *et al.* 2002, 2010). ABC substitutes an algorithmic solution (a simulation) for the explicit likelihood equation, using summary statistics to compare observation with simulations. With the ABC method, it is possible to calculate the relative probabilities of complex competing models (i.e. scenarios) and to estimate the demographic parameters underlying a given model (reviewed in Bertorelle *et al.* 2010 and Csilléry *et al.* 2010).

Black rat (*Rattus rattus*) is one of the best-known invasive vertebrate species. Over a period of several hundred years, populations of this species have invaded large parts of the world via commercial transport (Aplin *et al.* 2011). This ubiquitous species decreases the diversity of the indigenous fauna by predation or competition for resources (Goodman 1995; Courchamp *et al.* 2003; Towns *et al.* 2006; Stokes *et al.* 2009) and through the introduction of new pathogens and parasites into native populations (e.g. spread of *Trichuris muris* to native *Peromyscus maniculatus*; Smith & Carpenter 2006). The spread of rats has had substantial economic and human health consequences, including the spread of haemorrhagic fever viruses (Mills & Childs 1998), plague (Tollenaere *et al.* 2010a), rickettsiosis (Azad *et al.* 1997), trypanosomiasis (Dobigny *et al.* 2011) and schistosomiasis (Prugnolle *et al.* 2005). However, despite the considerable impact of invasive rat populations on various aspects (biodiversity, the economy and health), little is known about the evolutionary history and dynamics of these populations, as the available data were obtained in highly focal (Abdelkrim *et al.* 2005, 2009, 2010; Prugnolle *et al.* 2005; Gilabert *et al.* 2007) and historical biogeographic studies (Matisoo-Smith & Robins 2009; Tollenaere *et al.* 2010b; Aplin *et al.* 2011; Lack *et al.* 2012). In particular, no study has yet specifically addressed the issues of the history of the invasion process, the geographic pathway of the invasion and the origin and genetic composition of the invading populations. However, such information might be useful for the development of effective strategies for managing black rat populations and preventing new, accidental invasions. Black rats have an ability to invade new environments, which have proved exceptional among vertebrates, from prehistoric times to the present day. In this respect, a description of the pathways and demographic processes underlying the black rat invasion could increase our general understanding of the crucial factors associated with invasion success.

We used a large sampling scheme and a combination of genetic, demographic and historical information to investigate the evolutionary history of invasive black rat populations in a large continental area (Senegal, West Africa). We aimed: (i) to bring together and summarize historical knowledge on the presence of the black rat in Senegal; (ii) to depict the spatial structure of neutral genetic variation in the black rat population; (iii) to infer the invasion pathways of *R. rattus* into Senegal from these two types of information; and (iv) to characterize the demographic features (founder effects, bottlenecks and multiple introductions) prevailing during the invasion, from the initial site of introduction to the colonization front. We characterized the genetic structure of the black rat population by Bayesian clustering methods (i.e. STRUCTURE; Pritchard *et al.* 2000) and traditional population genetics (e.g. calculation of heterozygosity and

F_{ST}). We then performed ABC, to compare various invasion scenarios covering the whole of Senegal, taking into account recent and historical records and evolutionary processes, such as mutation, genetic drift and admixture.

Materials and methods

Historical records

Historical records of the presence of the black rat in Senegal were compiled from previous publications. Furthermore, longitudinal sampling programs carried out in Senegal since the 1980s provided information about the recent inland advance of the species in Senegal. These new data were extracted from the Sahelo-Soudanian Rodent Database: http://www.bdrss.ird.fr/bdrsspub_form.php (Granjon & Duplantier 2009).

Sampling

Between 2004 and 2007, we sampled rats at 24 villages and towns throughout Senegal, to cover the entire known distribution range of *R. rattus* in this country (Table 1). We set no more than two traps per hut or house, to prevent the sampling of closely related individuals.

Genetic diversity and structure

Genomic DNA was extracted from toe or spleen tissue samples with the QIAGEN DNeasy Tissue Kit or Gentra Puregene DNA Purification Kit, according to the manufacturer's instructions. We genotyped 17 microsatellite loci, seven of which were originally developed for the brown rat (*Rattus norvegicus*) and had already been used in studies of black rat populations (D5Rat83, D7Rat13, D10Rat20, D11Mgh5, D11Rat56, D16Rat81 and D18Rat75; Prugnolle *et al.* 2005). The other 10 microsatellite loci were developed by Loiseau *et al.* (2008), specifically for the black rat (Rr14, Rr17, Rr21, Rr22, Rr54, Rr67, Rr68, Rr93, Rr107 and Rr114).

Genetic diversity was estimated over all loci, for each site, by calculating unbiased expected heterozygosity in GENETIX version 4.0.5 (Belkhir *et al.* 1996–2004). Mean allelic richness over loci was determined with the rarefaction procedure implemented in FSTAT version 2.9.3 (Goudet 2001), to estimate the expected number of alleles in subsamples of 10 genes. We assessed the deviation from Hardy–Weinberg equilibrium for each locus separately and for all loci considered together, for each site, by exact tests based on Markov chain methods in GENEPOP version 4.0.7 (Rousset 2008). We corrected for multiple testing by the false discovery rate approach, in QVALUE version 1.0 (Storey & Tibshirani 2003). Close inspection of the results obtained revealed high frequencies of null alleles at three loci (D10Rat20, Rr14, and Rr21), which were therefore excluded from subsequent analyses. Pairwise F_{ST} values and F_{IS} values were calculated with FSTAT.

The spatial genetic structure over the entire sampling area was investigated by the Bayesian approach implemented in STRUCTURE version 2.2 (Pritchard *et al.* 2000). The method is based on an explicit evolutionary model for genetic variation, and its statistical inferences are based on individual data. It aims to subdivide the data into clusters of individuals maximizing Hardy–Weinberg equilibrium and minimizing linkage disequilibrium. We assumed an admixture model, in which the programme assigns proportions of individual genotypes to each of K clusters. We performed 100 independent runs for each K value (from 1 to 10), with different values of the Dirichlet α parameter for each assumed cluster. Each run included 200 000 burn-in iterations and 600 000 iterations. The best models were selected on the basis of likelihood, and ΔK was calculated as described by Evanno *et al.* (2005).

We used the Greedy algorithm implemented in CLUMPP version 1.1.1 (Jakobsson & Rosenberg 2007) to calculate the pairwise 'symmetric similarity coefficient' (SSC) between pairs of runs, for the identification of potentially different modes from the results of the 100 STRUCTURE runs for each K value. As recommended by the authors, we considered pairs of runs with an SSC > 0.9 to belong to the same mode. We then averaged individual genotype membership proportions for all runs in the same mode, for a given K value. We used these averaged results to generate summary barplots with DISTRUCT version 1.1 (Rosenberg 2004).

Inference of colonization pathways

Based on historical knowledge of the presence of the black rat in Senegal and the results obtained with Bayesian clustering methods, we developed and compared different colonization scenarios by approximate Bayesian computation (ABC; Beaumont *et al.* 2002). ABC analyses were carried out with DIYABC version 0.7.3 (Cornuet *et al.* 2008) and the model checking option of version 1.0.4 (Cornuet *et al.* 2010). Given the huge number of possible scenarios when all possible source populations are considered, we adopted a stepwise procedure, in which we focused our ABC analysis on four questions dealt with in succession. These four questions concerned the invasion history of the four genetically different groups identified in the STRUCTURE analysis, taking into account invasion chronology (a list of scenarios can be found in Table S2, Supporting

Table 1 Site and sampling information for genotyped black rat individuals and grouping of the material during analyses

Sites	Coordinates	Date of sampling	Sample size	Sample code	Sample grouping based on STRUCTURE	Sample grouping for DIYABC
Dakar	14°41′04.1″N, 017°25′12.5″W	May and July 2007	8	DAK	WEST	WEST
Sikilo	14°06′13.9″N, 015°29′22.4″W	April 2006	11	SIK	WEST	WEST
Pete	14°06′15.8″N, 015°26′50.6″W	April 2006	1	SIK	WEST	WEST
Oussouye	12°28′59.4″N, 016°32′58.9″W	December 2005	11	OUS	WEST	WEST
Baila	12°53′29.9″N, 016°20′14.2″W	December 2005	15	BAI	WEST	WEST
Diattacounda	12°32′38.8″N, 015°39′57.9″W	December 2005	10	DIA	WEST	WEST
Sare	12°50′57.4″N, 014°36′31.5″W	December 2005	17	SAR	WEST	WEST
Soutouré	12°56′33.2″N, 014°07′40.9″W	November 2005	3	STR	ADMIXED	-
Kounkané	12°55′48.5″N, 014°04′46.2″W	December 2005	3	STR	ADMIXED	-
Altou Fonkola	12°22′17.0″N, 013°27′35.0″W	March 2005	5	ALT	ADMIXED	-
Ndiambour	13°50′10.6″N, 014°29′04.0″W	January 2006	10	NDM	CENTER	CENTER
Méréto	13°48′57.0″N, 014°26′12.5″W	January 2006	8	NDM	CENTER	CENTER
Touba Sine	13°51′37.2″N, 014°20′34.5″W	April 2006	13	TBS	CENTER	CENTER
Malem Niani	13°56′25.3″N, 014°17′46.2″W	April 2006	1	TBS	CENTER	CENTER
Diala Koto	13°18′56.8″N, 013°17′00.7″W	August 2006, September 2007	38	DKT	NKNP	NKNP
Dienoudiala	13°12′47.9″N, 013°06′38.6″W	August 2006	31	DND	NKNP	NKNP
Tambacounda	13°46′30.0″N, 013°39′43.4″W	September 2007	7	TAM	EAST	EAST
Tata	14°06′23.5″N, 013°24′47.8″W	November 2004	9	KAL	EAST	EAST
Kaldem	14°08′45.6″N, 013°24′32.3″W	November 2004	13	KAL	EAST	EAST
Talibadji	14°04′21.1″N, 012°59′49.3″W	November 2004	5	TAL	EAST	EAST
Boynguel Bamba	14°05′44.7″N, 012°55′56.8″W	December 2004	10	TAL	EAST	EAST
Alagui Coré	14°08′29.9″N, 013°01′31.1″W	December 2004	4	TAL	EAST	EAST
Ndia	14°33′22.3″N, 012°44′58.0″W	December 2004, July 2005, January 2006	12	NDI	EAST	EAST
Kédougou	12°33′40.1″N, 012°10′35.7″W	January 2004, May 2006	12	KED	EAST	KED

information). Based on the results obtained for each of the four questions analysed separately, a best scenario was identified from the set of scenarios put forward in the first step and used to design a set of competing scenarios for subsequent steps.

The scenarios considered included the following demographic parameters: the times at which split or admixture events occurred, expressed in numbers of generations ago (Ti); stable effective population size (Ni); the effective number of founders in introduced

populations (*NFi*); the duration of the bottleneck occurring during colonization, in generations (*TDi*); and the rate of admixture (*Ri*). The number of demographic parameters differed between scenarios, as a function of scenario complexity. Known dates of first observation were used to fix the times (expressed in numbers of generations) at which population splitting and admixture events had occurred. ABC analyses were performed with parameter values drawn from the prior distributions (Prior Set 1) described in Table S4 (Supporting information). For microsatellite loci, we followed a generalized stepwise model (GSM) in which a mutation increases or decreases the number of repeated motifs by one or several units, this number of units being drawn from a geometric distribution (Cornuet *et al.* 2008). The default values proposed by DIYABC for microsatellite markers were kept for the priors of the mutation model parameters (see DIYABC documentation manual available from http://www1.montpellier.inra.fr/CBGP/diyabc/).

The ABC method is based on summary statistics calculated from the data to represent the maximum amount of information in the simplest possible form. The within- and between-population genetic variation of the populations was summarized with statistics equivalent to those traditionally used by population geneticists (NAL, HET, VAR, MGW, FST, LIK, N2P, H2P, V2P and AML). A full description of these statistics, with appropriate references, is provided in the DIYABC documentation manual. We compared the different scenarios by calculating their relative posterior probabilities by polychotomous logistic regression (Cornuet *et al.* 2008) from the 0.1% of simulated data sets most closely resembling the observed data set (in terms of the calculated Euclidean distance between the target and observed summary statistics). The posterior distributions of parameters were then estimated under the most likely scenario by the logit transformation of parameters and linear regression on the 0.5% of simulated data sets most closely resembling the observed data set.

We evaluated the sensitivity of the DIYABC inference concerning population history, expressed as a scenario choice, with another set of prior distributions (Prior Set 2), which differed from the previous set (Prior Set 1) in using normal distributions rather than uniform distributions for all demographic parameters other than event times (*Ti*), bottleneck duration (*TDi*) and mutation parameters (Mean mutation rate, Mean coefficient *p*), for which gamma distributions were used.

We evaluated the power of our ABC method to discriminate between our sets of competing scenarios, by analysing simulated pseudo-observed data sets (pods)

with the same number of loci and individuals as our real data set. One hundred such pods were simulated for each scenario, using parameter values drawn from the same prior distributions as for previous ABC analyses (Prior Set 1; Table S4). The relative posterior probabilities of each competing scenario were estimated for each pod. These probabilities were used to calculate type I and II errors for the scenario chosen with our real data set (i.e. the focal scenario). Type I error is the likelihood of excluding the focal scenario when it is actually the true scenario, and type II error is the likelihood of selecting the focal scenario when it is not the true scenario (mean type II error was calculated over the competing scenarios).

Model checking

We evaluated the goodness-of-fit of the best scenario obtained at the end of the selection procedure, using the model checking option in DIYABC version 1.0.4 (Cornuet *et al.* 2010). This option can be used to evaluate the consistency of the observed data with the posterior predictive distribution of the model for the best scenario. We ranked the summary statistics for the observed data against those for the simulated data. The observed statistics had to lie within the distributions of the simulated statistics for the fit of the model to be considered good. As recommended by Cornuet *et al.* (2010), we used summary statistics that had not been used for model selection or parameter estimation in previous ABC treatments.

Results

History of the black rat in Senegal

The black rat began to expand out of its original area in the Indian Peninsula (Musser & Carleton 2005) about 10 000 years ago. It was first introduced into Europe about 2500 years ago, and its populations on this continent were long confined to trading routes and ports (Audoin-Rouzeau & Vigne 1994, 1997; Aplin *et al.* 2003). It reached Western Africa during the 15th century, with the first Portuguese seafarers (Rosevear 1969). The arrival of black rat populations was favoured by the establishment of permanent commercial ports along the Atlantic coast (Fig. 1). The first West African port was founded in 1659, at Saint-Louis (Senegal), and rapidly became the most important trade centre in West Africa, retaining this role until the construction of Dakar in 1860. Although less important, the ports of Banjul and Ziguinchor have played non-negligible roles in trade since 1816 and 1888, respectively (Sinou *et al.* 1989). Black rats remained restricted to coastal areas until the

1659: First European settlement

1800–1890: Other European ports
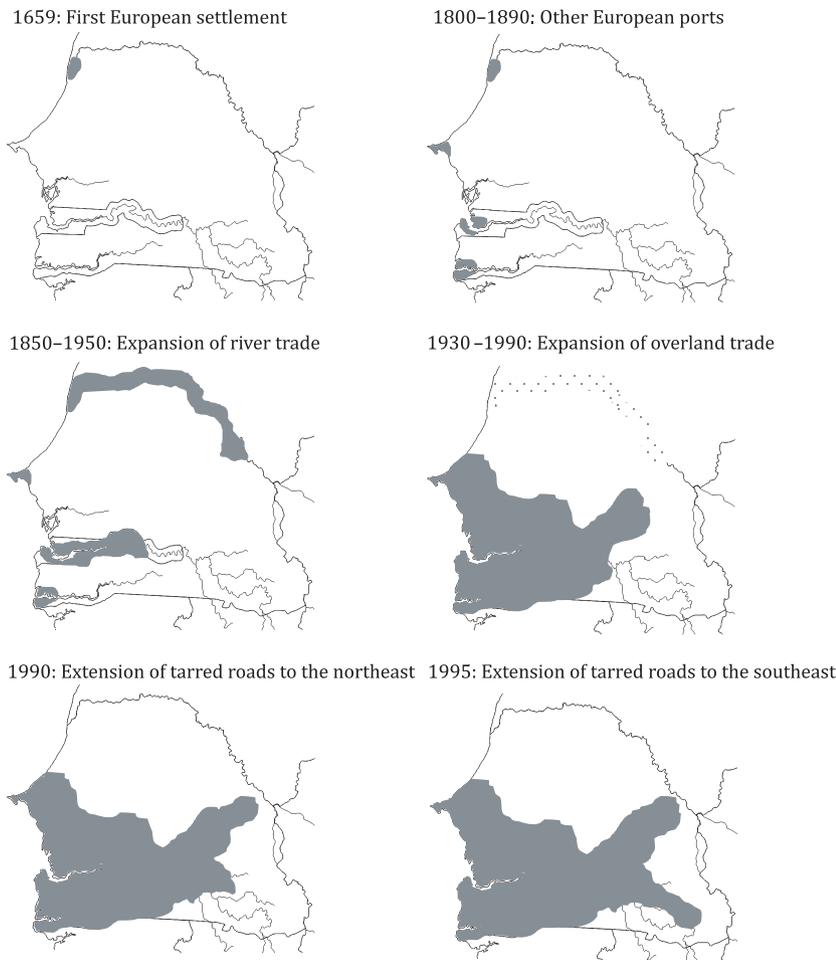
Fig. 1 Colonization of Senegal by the black rat, based on historical data. Gray areas represent the approximate distribution over time (see main text for dating and references). Note the disappearance of black rats along the Senegal River following the decrease in river-based trade after the 1930s (shaded area).

1850–1950: Expansion of river trade

1930–1990: Expansion of overland trade

1990: Extension of tarred roads to the northeast   1995: Extension of tarred roads to the southeast

development of commercial transport along the Senegal and Gambia rivers during the 18th and 19th centuries (Sinou 1981; Fig. 1). This inland penetration along the rivers is confirmed by the presence in the British Museum of *R. rattus* specimens originating from Bakel on the Senegal River, and from Kuntaur and Maka-Colibentan on the Gambia River (Rosevear 1969). River trade progressively declined after the 1950s, and black rats are no longer found along the Senegal River (Fig. 1; J.-M. Duplantier & K. Bâ, unpublished). However, an alternative route for the invasion of inland Senegal appeared in the 1930s, with the construction of roads and the development of road transport. The distribution of black rats in Senegal at the end of the 20th century was described in detail by Duplantier *et al.* (1991). Rats are abundant in the villages and towns of the Casamance region (south of the Gambia), around Tambacounda (east of the Gambia) and along the Atlantic coast (south of Dakar). However, they are absent from the eastern part of Senegal bordering Mali and from the northern half of the country (Fig. 1). The black rat is currently invading south-eastern Senegal. It was captured

for the first time at several guard camps at the entrance of the Niokolo Koba National Park (NKNP) in the 1980s and in the town of Kédougou in 1998 (Bâ 2002).

Spatial genetic structure of the black rat in Senegal

Between 2004 and 2007, we sampled 257 rats from 24 sites scattered over the entire distribution range of the species in Senegal (Table 1). We pooled together sampling sites that were physically close together and had small sample sizes, for the final analysis of a set of 17 samples (Table 1, Fig. 2). Genetic variability varies from six to 21 alleles per locus (mean, 12.93), depending on the site considered. Standardized mean allelic richness (A) varied from 1.86 to 3.99, and expected heterozygosity over all loci (H_E) varied from 0.32 to 0.67 (Table S1, Supporting information). Both these measurements of genetic diversity decreased significantly from the west to the east of Senegal ($F = 11.68$, $P = 0.004$ for A; $F = 10.11$, $P = 0.006$ for H_E), following the direction of spread of the invasion (inland from the coast). Mean genetic differentiation between sites, for all loci, was
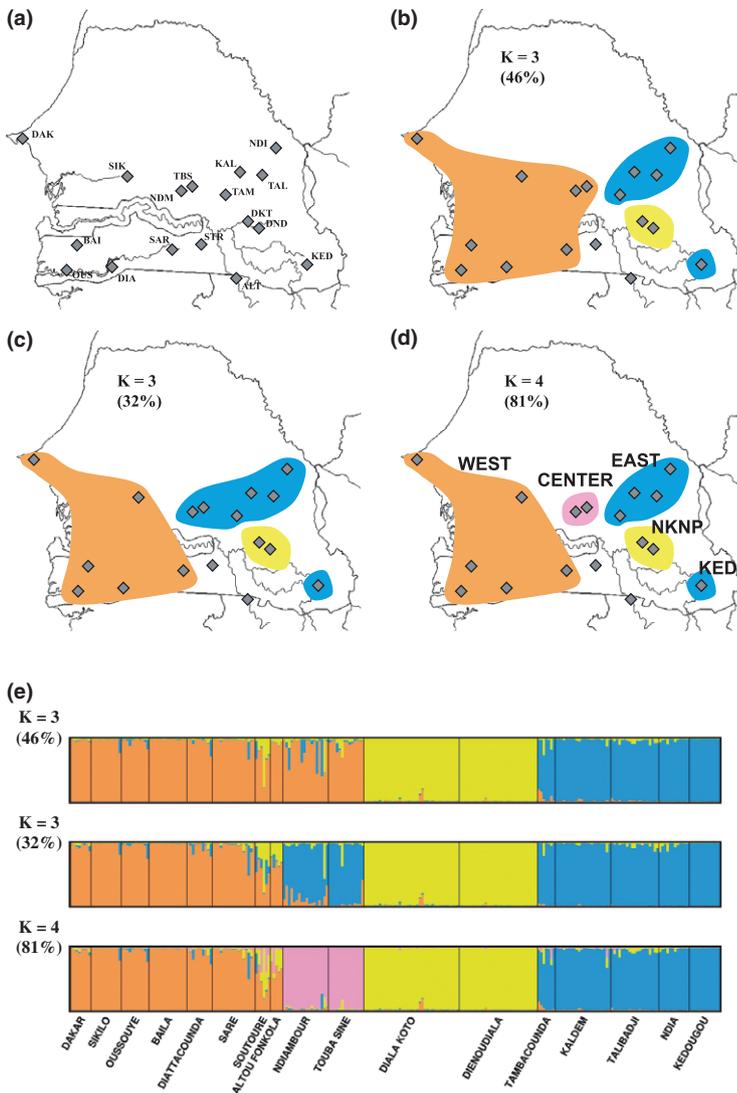
Fig. 2 Sampling and genetic structure of the black rat population in Senegal. (a) Map of sites sampled for the black rats in Senegal (including one site in northern Guinea). Code names are given in Table 1. (b-d) Mapping of the three genetic clustering that are more or less equally likely based on the Bayesian analysis processed with STRUCTURE version 2.2 (Pritchard *et al.* 2000). (b) The first clustering pattern for $K = 3$, which represents 46% of the runs for $K = 3$. (c) The second clustering pattern for $K = 3$, which represents 32% of the runs for $K = 3$. (d) The clustering pattern for $K = 4$, which represents 81% of the runs for $K = 4$. WEST, CENTER, NKNP and EAST clusters (orange, pink, yellow and blue, respectively). Note that the blue cluster is formed by two separate geographic areas: KED in the southeast (Kédougou site) and the other sites of the EAST group. Two sites, Soutouré (STR) and Altou Fonkola (ALT), were consistently found to be admixed. (e) For each of the three genetic clustering, we give the estimated membership of each cluster, for each individual, indicated by a single vertical line for each individual divided into K colours, according to the STRUCTURE analysis. Black lines separate individuals from different sites. The percentage indicates the proportion of the data set corresponding to the genetic structure mode presented, for each number of potential clusters (K).

$F_{ST} = 0.25$ (95% CI: 0.22–0.28). Pairwise F_{ST} values (for pairs of sites) ranged from 0.07 (TAL vs. NDI) to 0.51 (SIK vs. KED).

Bayesian clustering analysis revealed that the population was genetically structured into separate clusters (Fig. 2). However, it was difficult to determine the optimal value for K, because the log-likelihood values and ΔK values for $K = 3$ and $K = 4$ were similar (Fig. S1, Supporting information). Moreover, for $K = 3$, CLUMPP analysis gave two different modes, supported by similar numbers of runs (46% and 32%). These two modes differed only in the assignment of the NDM and TBS samples, which are assigned either to the western cluster or to the eastern cluster. For $K = 4$, NDM and TBS form the fourth cluster. We thus end-up with a total of three spatial structures that are more or less equally likely. This uncertainty was taken into account for specifying the invasion scenarios within an ABC framework. As a

matter of fact, each one of the three spatial structures could be translated into a specific invasion scenario (Fig. S2, Supporting information). The ABC analyses then allowed formal discrimination among these three scenarios (see next section).

Most individuals (236 of 257) were unambiguously assigned to only one of these four clusters (estimated membership for each individual greater than 80%). The remaining individuals, particularly those from ALT and STR, displayed a high degree of admixture. For the sake of simplicity, and given the small number of individuals in the ALT and STR samples, we excluded both these samples from subsequent ABC analyses. Pairwise F_{ST} values between the four clusters ranged from 0.13 to 0.23, and F_{IS} varied from 0.13 to 0.25 (Table 2). These high F_{IS} values result from substantial residual genetic differentiation between sites within clusters.

Table 2 Genetic differentiation (F_{ST}) between pairs of genetic clusters and genetic differentiation (F_{IS}) within clusters

	WEST	CENTER	NKNP	F_{IS}
WEST	—	—	—	0.181
CENTER	0.132	—	—	0.129
NKNP	0.200	0.228	—	0.151
EAST	0.153	0.145	0.168	0.250

Inference of colonization pathways

For ABC analyses, we considered the following groups based on the STRUCTURE results (Table 1): WEST (the DAK, SIK, OUS, BAI, DIA and SAR sites), EAST (TAM, KAL, TAL and NDI), CENTER (NDM and TBS), and NKNP (DKT and DND). Moreover, we considered EAST and KED separately, although STRUCTURE outputs group them in the same genetic cluster, for the following reasons. EAST and KED correspond to two separate and distant areas (Fig. 2d). Besides, historical records clearly indicate that both areas, KED and EAST, were colonized at very different times (i.e. EAST was colonized well before KED). The scenarios we used in our ABC analyses take into account this temporal sequence

of the invasion. Moreover, although STRUCTURE results strongly suggest that KED most likely originated from EAST, it was interesting to use ABC analyses to: (i) check the single origin of the recent KED foundation; and (ii) assess whether such population foundation was associated with a bottleneck event. Finally, to be sure that our way of modelling did not cause a bias in data analysis, we processed other ABC analyses in which KED is including into EAST cluster, hence mimicking the STRUCTURE clustering pattern for $K = 4$. Results are displayed in Tables S2 and S3 (Supporting information). Including KED into EAST did not change anything in the selection of best scenarios (although *P*-values slightly changed).

Based both on our Bayesian clustering results and historical records, we developed and compared four sets of competing colonization scenarios, addressing four questions successively (Table S2). The first question focused on the colonization pathway of the two oldest invasive foci, corresponding to the WEST/EAST genetic partition. We developed eight competing scenarios differing in terms of whether the WEST and EAST groups were founded successively by rats from the other group or independently, from an external source, either directly from abroad or from a hypothetical unsampled
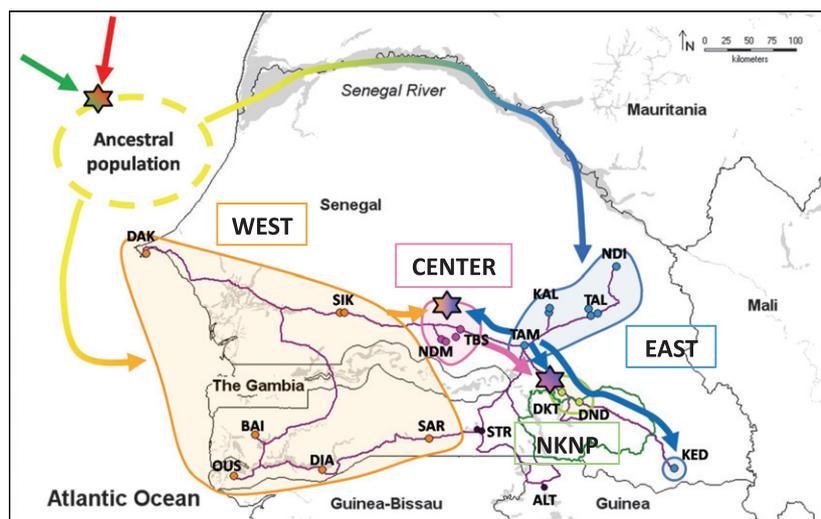
Fig. 3 Most likely colonization pathways (marked by arrows) of the black rat in Senegal, inferred by DIYABC, following STRUCTURE analysis. All sites from which rats were sampled and analysed are indicated by a circle colour-coded according to cluster membership, as revealed by Bayesian assignment analysis for $K = 4$ clusters: WEST is shown in orange, CENTER in pink, NKNP (the sites close to the Niokolo Koba National Park, which is indicated by a dark green line) in green, EAST in blue (including the Kédougou population, KED) and the admixed populations are shown in black. Stars indicate the clusters created by an admixture event. The principal roads are indicated by violet lines. An 'ancestral' (unsampled) population of *Rattus rattus* was established somewhere on the Senegalese coast by at least two independent introductions from outside Senegal, followed by an admixture event. Both the WEST and EAST groups were colonized independently from the ancestral population (rats introduced into the EAST probably followed the Senegal River, as indicated by historical records). The CENTER and NKNP groups were colonized from two sources, followed by genetic admixtures: WEST and EAST, and CENTER and EAST, respectively, whereas KED was colonized from a single source, EAST (see text for details).

Table 3 Confidence in scenario selection by DIYABC for the colonization of Senegal by black rats

Issue - invaded area	Number of competing scenarios	Description of the selected scenario	Posterior probability	95% credibility interval	Type I error	Type II error
1 - Invasion of Senegal, its western and eastern parts (groups WEST and EAST)	8	Scenario 3: multiple introductions from outside Senegal into an ancestral population from which WEST and EAST were independently colonized	0.6881	[0.6464,0.7298]	0.47	0.08
2 - Invasion of central Senegal (group CENTER)	3	Scenario 3.3: CENTER originated from both WEST and EAST (admixture event)	0.6414	[0.4967,0.7861]	0.17	0.08
3 - Invasion of sites near the Niokolo Koba National Park (NKNP group)	6	Scenario 3.3.6: NKNP originated from both CENTER and EAST (admixture event)	0.8842	[0.8197,0.9490]	0.23	0.04
4 - Invasion of the town of Kédougou (KED group)	10	Scenario 3.3.6.3: KED invaded from EAST	0.9133	[0.8529,0.9737]	0.17	0.03

population somewhere in Senegal (the 'ancestral' population). We also included the possibility of single or multiple introduction events (Table S2; question 1, scenarios 1–8). The posterior probabilities calculated for each scenario provided the strongest statistical support for scenario 3 ($P = 0.69$; Table 3). This scenario involved multiple introductions from an unknown common population in Senegal (the 'ancestral' population), followed by inland colonization of Senegal via two separate, independent colonization routes (the WEST and EAST groups).

We tackled the second question, regarding the colonization history of the CENTER group, by constructing three competing scenarios (Table S2; question 2, scenarios 3.1–3.3). Each one of these scenarios corresponds to one of the three equally likely clustering patterns inferred from STRUCTURE analyses (Fig. 2, Fig. S2). Scenario 3.1 corresponds to the first clustering pattern for $K = 3$, which represents 46% of the runs, where CENTER populations resulted of the spatial extension of the WEST cluster. Scenario 3.2 corresponds to the second clustering pattern for $K = 3$, which represents 32% of the runs, where CENTER populations resulted of the spatial extension of the EAST cluster. Scenario 3.3 corresponds to the clustering pattern for $K = 4$, where CENTER populations form a separate cluster as a result of being a mixture between WEST and EAST populations. Following ABC analysis, the most likely scenario was scenario 3.3 ($P = 0.64$; Table 3), which assumed two separate introductions from WEST and EAST followed by genetic admixture within CENTER.

Six scenarios were then designed to tackle the third question, focusing on the origin of the NKNP group (Table S2; question 3, scenarios 3.3.1–3.3.6). The strongest support was obtained for scenario 3.3.6 ($P = 0.88$; Table 3). This scenario assumed two separate introductions from CENTER and EAST followed by genetic

admixture in NKNP. Finally, we attempted to retrace the origin of the population from Kédougou (KED). Ten competing scenarios were constructed, involving colonization from WEST, CENTER, EAST or NKNP and/or admixture between each of these populations (Table S2; question 4, scenarios 3.3.6.1–3.3.6.10). The most probable scenario, by far, was scenario 3.3.6.3 ($P = 0.91$; Table 3), involving a single introduction from EAST. The final colonization scenario, scenario 3.3.6.3 (Fig. 3 and Fig. S4, Supporting information), highlights the complexity and the somewhat counterintuitive nature of the invasion pathways followed by the black rat in Senegal.

Robustness of scenario choice

We evaluated the robustness of our scenario choice conclusions with respect to prior distributions, by carrying out the same ABC analyses with a different set of prior distributions (Prior Set 2). For each of the four questions we addressed, the same scenarios were unambiguously chosen as were obtained with Prior Set 1 [scenario 3: $P = 0.63$, 95% CI = (0.58; 0.67); scenario 3.3: $P = 0.93$, 95% CI = (0.86; 1.00); scenario 3.3.6: $P = 0.99$, 95% CI = (0.97; 1.00); scenario 3.3.6.3: $P = 0.88$, 95% CI = (0.78; 0.99)]. As for Prior Set 1, the confidence intervals of the chosen intervals never overlapped with those of competing scenarios.

Power analyses revealed that the type I errors (i.e. false positives) associated with the best scenarios were high (Table 3). Thus, in some cases, when the data were simulated with the best scenario, there was another scenario with a non-negligible probability that could have been used. This was particularly true for the first step in the modelling procedure (type I error = 0.47). The alternative scenario differed only slightly from the best scenario in that the best scenario assumed multiple

introductions in Senegal, whereas the alternative scenario assumed a single introduction. We therefore conclude that this question has not yet been fully resolved. Data selection for the other three steps gave smaller type I errors. The (mean) type II errors (i.e. false negative) associated with the scenarios chosen on the basis of the real data set were all low (0.027–0.085), indicating that the probability of selecting the best scenario when data were simulated with an alternative scenario was low.

Estimation of demographic parameters

We inferred the posterior distributions of demographic parameters under the final colonization scenario (scenario 3.3.6.3) only. A close comparison of prior and posterior (Tables S3 and S4, Supporting information) distributions indicated that the estimated posterior distributions for effective sizes and the timing of split/ admixture events were both relatively flat and, hence, non-informative. By contrast, the confidence intervals of the posterior distributions for parameters measuring the strength of the bottleneck (i.e. TDi/NFi) and for admixture rates (Ri) were consistently narrower than those for prior distributions, because the posterior distributions were more pointy and, hence, informative. Bottlenecks were weak at the sites initially colonized by rats: 0.07 (95% CI: 0.00–0.22) for WEST, 0.32 (0.02–0.89) for EAST and 0.32 (0.00–0.83) for CENTER. By contrast, strong bottlenecks were detected for more recently founded populations: 0.50 (0.11–1.00) for NKNP and to 1.75 (1.00 –2.67) for KED. Admixture rates were fairly balanced, with a gene contribution of about 0.40 (0.16–0.75) from CENTER and 0.44 (0.16–0.80) from NKNP (Table S5, Fig. S3, Supporting information).

Model checking

We used 20 summary statistics that had not previously been used for model selection as test statistics for model checking (Table S6, Supporting information). None of the summary statistics used as test statistics had low probability values when the model was checked (Table S6). We therefore concluded that the selected scenario fitted the observed data well.

Discussion

This study highlights the contribution of novel approaches based on model-based population genetics modelling that can be used for any kind of invasive species. Using both historical and genetic data, analysed initially by a clustering method and then by ABC methods, we were able to reconstruct the historical pathways following during the invasion of a large continental area in West Africa by

black rats. Our data support the occurrence of two separate invasion routes extending inland from coastal sites. Human trade routes had a major impact on the invasion pathways used by rats, and some long-dispersal events were found to have occurred. Finally, our findings suggest that genetic bottlenecks and genetic admixture have played a major role in shaping the genetic characteristics of invasive black rat populations in Senegal. These demographic processes may have generated genetic novelty and favoured rapid adaptive evolution along the invasion pathways, as described elsewhere (see also reviews by Novak & Mack 2005; Dlugosch & Parker 2008).

A complex invasion pathway that parallels the history of human trade routes

Black rats colonized Senegal following the development of commercial trade between West Africa and Europe in the 15th century (Rosevear 1969). Historical records suggest that these rats spread through Senegal from west (the Atlantic coast) to east (inland Senegal). This historical pattern is supported by the gradient of decreasing genetic diversity from the west to the east. At first sight, the genetic data thus appeared to be consistent with the simplest model of a single invasion wave spreading throughout the country in a west–east direction. However, closer analysis of the genetic structure of black rat populations within the invaded area revealed a patchy spatial structure, calling this simple invasion scenario into question. Model-based methods, such as STRUCTURE and ABC, then further highlighted (and resolved) the complexity of the invasion pathways followed by rat populations, from the area of initial introduction towards the colonization front.

The year 1659, when the port of Saint-Louis was founded, may be considered the first possible date of black rat introduction into West Africa. Multiple introductions probably occurred from different European ports (i.e. from different gene pools). These introductions would have led to the establishment of an ancestral population. As our analyses with STRUCTURE and DIYABC were not spatially explicit, we cannot determine where this ancestral population first settled and developed or the precise geographic route of invasion without reference to historical records of the presence of this species. Historical data suggest that the ancestral population was probably established in the ports at which black rat sightings were first recorded. The region of Saint-Louis (north-western Senegal) is a particularly good candidate. The use of two independent colonization routes within Senegal probably then gave rise to two main genetic groups in inland Senegal (Fig. 3). The WEST branch has remained close to the

coast, spreading in the western half of the country, with river trade along the Gambia River favouring the colonization of the Casamance region (southern Senegal). The EAST branch may correspond to a separate invasion route along the Senegal River in the 18th and 19th centuries, when there trade flourished along this river. The recorded presence of several black rats in Bakel on the Senegal River (NE Senegal; Rosevear 1969) suggests that north-eastern Senegal may have been colonized via this fluvial route. From the region of Bakel, the black rat may have continued south to the region of Tambacounda (TAM), a major trading centre in eastern Senegal. The WEST and EAST branches subsequently met up in central Senegal, where they merged and genetic admixture gave rise to the CENTER group (Fig. 3). During the 1980s, black rats founded new populations in villages at the entrance of the NKNP (Duplantier *et al.* 1991). Data modelling provides evidence to suggest that these populations resulted from multiple introduction events from several sources, in this case from the EAST and CENTER groups.

A particularly unusual situation was observed for Kédougou (KED) in south-eastern Senegal. The presence of black rats in the town of Kédougou was first recorded in November 1998 (Bâ 2002) despite the devotion of considerable effort to sampling before this date (Duplantier *et al.* 1991). The rat population of Kédougou is very isolated, as no other populations have been found in the villages between Dienoudiala (DND) and Kédougou, despite intensive sampling (A. Konečný, unpublished). The black rat thus appears to have colonized Kédougou recently. The simplest scenario would be that the black rats in this town originate from Dienoudiala, the closest site known to harbour black rats, belonging to the NKNP genetic group. However, Bayesian clustering analysis revealed that the black rats from Kédougou belonged to another genetic group, the EAST group. ABC analysis was entirely consistent with clustering analyses, strongly supporting the recent colonization of Kédougou by rats immigrating from the EAST group. This would necessarily involve long-distance dispersal, because the EAST cluster site closest to Kédougou (Tambacounda) is about 150 km away. This long-distance colonization may have been favoured by the increase in road transport and trade between Tambacounda and Kédougou following the construction of tarmac road connecting Kédougou to the road network in 1995. Trade between Tambacounda and Kédougou involves the use of large trucks that do not stop in the small towns and villages on the way. These trucks loaded with food may have been responsible for transporting *R. rattus* to Kédougou. All these findings strongly suggest that the colonization of Senegal by black rats is closely related to human activity, including transport along rivers and overland trade routes, as previously reported in other areas of the world by Aplin *et al.* (2003).

Evolutionary processes occurring during the invasion of Senegal by black rats

The demographic and evolutionary processes underlying invasion have been the object of considerable attention in recent years (e.g. Lee 2002; Dlugosch & Parker 2008; Suarez & Tsutsui 2008). From a practical standpoint, an understanding of these processes facilitates the design of strategies for preventing invasion. From an academic perspective, such an understanding is crucial to appreciate the importance of neutral and nonneutral evolution in invasive populations. The most important demographic processes occurring during invasions are bottleneck events and multiple introductions (Dlugosch & Parker 2008).

Bottlenecks have been found to play an important role in the invasion of Senegal by black rats. The populations of the most recently founded groups (NKNP and KED) were associated with intense bottlenecks, greatly limiting genetic diversity. Such genetic depletion, due to consecutive bottlenecks and founder effects, has been documented for various invasive organisms (Ciosi *et al.* 2008; Peacock *et al.* 2009) and in simulation studies (Austerlitz *et al.* 1997; Hallatschek & Nelson 2008). Theoretically, genetic depletion should reduce the ability of the species to adapt to new environments, thereby slowing the invasion. However, empirical observations are often at odds with this theoretical prediction, resulting in the so-called genetic paradox (Sax & Brown 2000). One possible explanation for this discrepancy is that the consequences of genetic drift may not necessarily be negative. In a theoretical study, Glémin (2003) showed that a combination of genetic drift and nonrandom mating in small populations may purge the population of deleterious mutations. Thus, under certain conditions (i.e. effective population size not too small and a sufficiently long bottleneck), wild populations may be purged of deleterious mutations during invasion, resulting in at least a partial decrease in mutation load (the accumulation of deleterious recessive mutations during evolution). Such an effect would not only help to compensate for the decrease in genetic variation, but would also facilitate invasion, as recently suggested for invasive populations of the ladybird *Harmonia axyridis* (Facon *et al.* 2011).

Recent studies have also shown that multiple introductions and admixture events may be common occurrences during invasions by most species. These processes may help to limit the loss of genetic variability (Kolbe *et al.* 2004; Suarez & Tsutsui 2008; Wilson

et al. 2009). Our findings strongly suggest that multiple introductions have made a major role in black rat invasions in Senegal. During our selection procedure, scenarios including multiple introductions were consistently preferred over scenarios including a single introduction only (all other parameters being equal). The hypothesis of multiple introductions followed by genetic admixture was unambiguously accepted, and the alternative hypothesis (unique introduction) rejected for the rat populations of central Senegal (i.e. for the genetic groups CENTER and NKNP). In both these cases, estimated genetic admixture was substantial (about 0.40), indicating that each cluster was likely to have resulted from the admixture of two sources in almost equal proportions. The occurrence of multiple introductions during the founding of the ancestral population in Senegal was also the preferred scenario (posterior probability = 0.69). However, test results were less conclusive in this particular case, the scenario involving a single introduction event also remaining a viable alternative (posterior probability = 0.18). The data studied may not have been entirely suitable for distinguishing unambiguously between these two scenarios. For example, we had no access to samples collected from the various putative sources of the ancestral Senegal population (Europe and India). We dealt with this issue indirectly, by including nonsampled populations in our colonization scenarios. The ability to model such so-called ghost populations, which contribute to the invasion scenario but are not sampled and analysed (Guillemaud *et al.* 2010), is an important characteristic of DIYABC software. However, in the absence of rigorous testing, we cannot assess the power for distinguishing between a single and multiple introduction events when ghost populations are included in the models. Power analyses with our data set showed type I errors to be fairly high for the first step in the modelling procedure, which included ghost populations. However, type II errors were markedly lower for all the selected scenarios.

Multiple introductions are thought to increase genetic diversity in introduced areas by mixing gene pools that were originally differentiated, increasing effective population size and, subsequently, population growth rate. Increases in genetic variation and evolutionary potential are particularly likely when multiple introductions result in the importation of novel genetic variants and combine genotypes from populations of different origin. Such processes may facilitate spread and establishment at new sites and/or may allow a breakthrough or stimulate already established populations (Suarez & Tsutsui 2008). Together with multiple introductions and admixture events, recurrent migration from the area of origin to already established populations, and/or migration

between sites within the introduced area, may preserve high levels of genetic variation (Suarez & Tsutsui 2008; Lombaert *et al.* 2010; Turgeon *et al.* 2011). These interacting processes have all had an impact on rat populations in Senegal, potentially increasing their invasive potential.

This mitigation of the loss of genetic diversity is time-dependent. For example, the genetic variation of the most recently founded population, in Kédougou (KED), has not yet been increased by admixture. At the time of sampling (May 2006), the level of genetic variation at this site was the lowest anywhere in Senegal. This may reflect a relatively low adaptive potential because, since the first recorded sighting in January 1998, the increase in the size of the black rat population at Kédougou has been limited (L. Granjon, personal communication). We might anticipate that KED population may be enriched by new alleles in the future, through new introductions either from the same source area (EAST) or from other sources (WEST, CENTER or NKNP), and that this increase in genetic variation may enhance population dynamics. Recent advances in population genetics suggest that invasions may favour the occurrence of genetic revolutions within species (Excoffier *et al.* 2009), and the consequences of which for adaptive genetics (particularly in the context of host–parasite interactions) remain to be explored.

Conclusion

The invasion of Senegal by black rats has been strongly influenced by human settlements and transport. The invasion pathways were more complex than initially anticipated, due, in particular, to the occurrence of long-distance dispersal, multiple introductions and admixture events between differentiated gene pools. The frequent occurrence of multiple introductions and admixture events may be of prime importance for understanding the invasion success of black rats throughout the world (see also Lack *et al.* 2012). Further experimental and empirical studies are required to determine the precise consequences of such an invasion pattern. However, this pattern may have helped to maintain high levels of genetic diversity along the invasion pathways and to enhance genetic novelty, adaptation and invasive potential.

Acknowledgements

References

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

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.

Aplin KP, Chesser T, Ten Have J (2003) Evolutionary biology of the genus *Rattus*: profile of an archetypal rodent pest. In: *Rats, Mice and People: Rodent Biology and Management* (eds Singleton GR, Hinds LA, Krebs CJ, Spratt DM), pp. 487–498. Australian Centre for International Agricultural Research, Canberra.

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

Ascunce MS, Yang CC, Oakey J *et al.* (2011) Global invasion history of the fire ant *Solenopsis invicta*. *Science*, **331**, 1066–1068.

Audoin-Rouzeau F, Vigne JD (1994) The colonization of Europe by the black rat (*Rattus rattus*). *Révue de Paléobiologie*, **13**, 125–145.

Audoin-Rouzeau F, Vigne JD (1997) The black rat (*Rattus rattus*) in Roman and Medieval Europe: the commercial roads and the spread of plague. *Anthropozoologica*, **25–26**, 399–404.

Austerlitz F, Jung-Muller B, Godelle B, Gouyon PH (1997) Evolution of coalescence times, genetic diversity and structure during colonization. *Theoretical Population Biology*, **51**, 148–164.

Azad AF, Radulovic S, Higgins JA, Noden BH, Troyer JM (1997) Flea-borne rickettsioses: ecologic considerations. *Emerging Infectious Diseases*, **3**, 319–327.

Bâ K (2002) *Systematics, Ecology and Dynamics of Small Rodents Populations Potentially Reservoirs or Hosts of Viruses in Senegal.* Mémoire Ecole Pratique des Hautes Etudes, Montpellier, France.

Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics*, **162**, 2025–2035.

Beaumont MA, Nielsen R, Robert C *et al.* (2010) In defence of model-based inference in phylogeography REPLY. *Molecular Ecology*, **19**, 436–446.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996–2004) *GENETIX 4.05, Logiciel sous Windows TM pour la génétique des Populations.* Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France.

Bertorelle G, Benazzo A, Mona S (2010) ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular Ecology*, **19**, 2609–2625.

Ciosi M, Miller NJ, Kim KS, Giordano R, Estoup A, Guillemaud T (2008) Invasion of Europe by the western corn rootworm, *Diabrotica virgifera virgifera*: multiple transatlantic introductions with various reductions of genetic diversity. *Molecular Ecology*, **17**, 3614–3627.

Cornuet JM, Santos F, Beaumont MA *et al.* (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics*, **24**, 2713–2719.

Cornuet JM, Ravigné V, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (version 1.0). *BMC Bioinformatics*, **11**, 401.

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

Csilléry K, Blum MGB, Gaggiotti OE, François O (2010) Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology and Evolution*, **25**, 410–418.

Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, **17**, 431–449.

Dobigny G, Poirier P, Hima K *et al.* (2011) Molecular survey of rodent-borne *Trypanosoma* in Niger with special emphasis on *T. lewisi* imported by invasive black rats. *Acta Tropica*, **117**, 183–188.

Duplantier JM, Granjon L, Adam F, Bâ K (1991) Present distribution of the black rat (*Rattus rattus*) in Senegal: historical and ecological factors. *Actes du colloque « Le Rongeur et l'Espace*, **2**, 339–346.

Estoup A, Guillemaud T (2010) Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology*, **19**, 4113–4130.

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, Foll M, Petit RJ (2009) Genetic consequences of range expansions. *Annual Review of Ecology, Evolution and Systematics*, **40**, 481–501.

Facon B, Genton BJ, Shykoff J, Jarne P, Estoup A, David P (2006) A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology and Evolution*, **21**, 130–135.

Facon B, Hufbauer RA, Tayeh A *et al.* (2011) Inbreeding depression is purged in the invasive insect *Harmonia axyridis*. *Current Biology*, **21**, 424–427.

Gilabert 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.

Glémin S (2003) How are deleterious mutations purged? Drift vs. non random mating. *Evolution*, **57**, 2678–2687.

Goodman SM (1995) *Rattus* on Madagascar and the dilemma of protecting the endemic rodent fauna. *Conservation Biology*, **9**, 450–453.

Goudet J (2001) FSTAT, a programme to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www2.unil.ch/izea/softwares/fstat.html.

Granjon L, Duplantier JM (2009) *Les rongeurs de l'Afrique Sahélo-Soudanienne.* (ed. IRD/MNHN), Collection Faune et Flore Tropicale, Paris.

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

Hallatschek O, Nelson DR (2008) Gene surfing in expanding populations. *Theoretical Population Biology*, **73**, 158–170.

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

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

Lack JB, Greene DU, Conroy CJ et al. (2012) Invasion facilitates hybridization with introgression in the *Rattus rattus* species complex. *Molecular Ecology*, **21**, 3545–3561.

Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 3883–3888.

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

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

Lombaert E, Guillemaud T, Cornuet JM, Malausa T, Facon B, Estoup A (2010) Bridgehead effect in the worldwide invasion of the biocontrol harlequin ladybird. *PLoS ONE*, **5**, e9743.

Mack RN, Simberloff D, Lonsdale WM, Evans H, Clout M, Bazzaz FA (2000) Biotic invasions: causes, epidemiology, global consequences and control. *Ecological Applications*, **10**, 689–710.

Matisoo-Smith E, Robins J (2009) Mitochondrial DNA evidence for the spread of Pacific rats through Oceania. *Biological Invasions*, **11**, 1521–1527.

Mills JN, Childs JE (1998) Ecologic studies of rodent reservoirs: their relevance for human health. *Emerging Infectious Diseases*, **4**, 529–537.

Musser GG, Carleton MD (2005) Superfamily Muroidea. In: *Mammal Species of the World: A Taxonomic and Geographic Reference* (eds Wilson DE, Reeder DM). pp. 894–1531, The Johns Hopkins University Press, Baltimore.

Novak SJ, Mack RN (2005) Genetic bottlenecks in alien plant species: influences of mating systems and introduction dynamics. In: *Species Invasions: Insights into Ecology, Evolution, and Biogeography* (eds Sax DF, Stachowicz JJ, Gaines SD), pp. 201–228. Sinauer Associates, Sunderland, Massachusetts.

Peacock MM, Beard KH, O'Neill EM, Kirchoff VS, Peters MB (2009) Strong founder effects and low genetic diversity in introduced populations of Coqui frogs. *Molecular Ecology*, **18**, 3603–3615.

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

Prugnolle F, Theron A, Pointier JP et al. (2005) Dispersal in a parasitic worm and its two hosts: consequence for local adaptation. *Evolution*, **59**, 296–303.

Reed DH, Frankham R (2001) How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution*, **55**, 1095–1103.

Rosenberg NA (2004) DISTRUCT: a programme for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

Rosevear DR (1969) *The Rodents of West Africa*. Trustees of the British Museum (Natural History), London.

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.

Sax DF, Brown JH (2000) The paradox of invasion. *Global Ecology and Biogeography*, **9**, 363–371.

Sinou A (1981) Key periods in the foundation of some colonial cities. *Cahiers d'Etudes Africaines*, **81–83**, 375–388.

Sinou A, Poinsot J, Sternadel J (1989) *Les villes d'Afique noire: politiques et opérations d'urbanisme et d'habitat entre 1650 et 1960*. Ministère de la Coopération, Documentation française, Paris.

Smith KF, Carpenter SM (2006) Potential spread of introduced black rat (*Rattus rattus*) parasites to endemic deer mice (*Peromyscus maniculatus*) on the California Channel Islands. *Diversity and Distributions*, **12**, 742–748.

Stokes VL, Banks PB, Pech RP, Spratt DM (2009) Competition in an invaded rodent community reveals black rats as a threat to native bush rats in littoral rainforest of southeastern Australia. *Journal of Applied Ecology*, **46**, 1239–1247.

Storey JD, Tibshirani R (2003) Statistical significance for genome-wide experiments. *Proceedings of the National Academy of Sciences of the United States of America*, **190**, 9440–9445.

Suarez AV, Tsutsui ND (2008) The evolutionary consequences of biological invasions. *Molecular Ecology*, **17**, 351–360.

Tollenaere C, Rahalison L, Ranjalahy M et al. (2010a) Susceptibility to *Yersinia pestis* experimental infection in wild *Rattus rattus*, reservoir of plague in Madagascar. *EcoHealth*, **7**, 242–247.

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

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

Turgeon J, Tayeh A, Facon B et al. (2011) Experimental evidence for the phenotypic impact of admixture between wild and biocontrol Asian ladybird (*Harmonia axyridis*) involved in the European invasion. *Journal of Evolutionary Biology*, **24**, 1044–1052.

Vitousek PM, D'Antonio CM, Loope LL, Rejmánek M, Westbrooks R (1997) Introduced species: a significant component of human-caused global change. *New Zealand Journal of Ecology*, **21**, 1–16.

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

Wilson JRU, Dormontt EE, Prentis PJ, Lowe AJ, Richardson DM (2009) Something in the way you move: dispersal pathways affect invasion success. *Trends in Ecology and Evolution*, **24**, 136–144.

Data accessibility

Both individual microsatellite data and the scripts for DIYABC analyses have been deposited to Dryad: doi: 10.5061/dryad.bb312.

Supporting information

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

Table S1 Genetic diversity within samples.

Table S2 Description of the invasion scenarios of the black rat in Senegal that were compared using the Approximate Bayesian Computation (ABC) approach implemented in DIYABC version 0.7.3 (Cornuet *et al.* 2008).

Table S3 Same as Table S2 but with adding KED to the EAST group (see text for details).

Table S4 Prior distributions of demographic, historic and mutation parameters used for the ABC analyses.

Table S5 ABC estimations of posterior distributions of demographic, historic and mutation parameters under the final colonization scenario 3.3.6.3 and assuming the Prior Set 1 (see Table S3).

Table S6 ABC model checking using 20 summary statistics not used previously for model selection as test statistics.

Fig. S1 Graphical representation of the most appropriate K (number of clusters regrouping the individuals) estimation revealed by STRUCTURE version 2.2 (Pritchard *et al.* 2000).

Fig. S2 Correspondence between the scenarios for DIYABC (see details in Table S2) and the best solutions from STRUCTURE analyses for $K = 3$ and $K = 4$ (see also Fig. 2).

Fig. S3 Posterior distribution of the admixture parameter Rcs in the scenario 3.3 (see also Table S2 and Fig. S4). Rcs represents the admixture parameter between EAST and WEST groups to form the group CENTER.

Fig. S4 Scheme of the selected final colonization scenario 3.3.6.3 (Table 3) summarizing the origin and diversification of black rats inferred using DIYABC.