ORIGINAL ARTICLE

Geographic patterns of inversion polymorphisms in a wild African rodent, *Mastomys erythroleucus*

G Dobigny[1], J Catalan[2], P Gauthier[1], PCM O'Brien[3], C Brouat[1], K Bâ[4], C Tatard[1], MA Ferguson-Smith[3], JM Duplantier[1], L Granjon[4] and J Britton-Davidian[2]

[1]*Institut de Recherche pour le Développement, Centre de Biologie pour la Gestion des Populations (UMR IRD/INRA/CIRAD/ Montpellier SupAgro), Montferrier-sur-Lez, France;* [2]*Institut des Sciences de l'Evolution (UM2, CNRS), Laboratoire Génétique et Environnement, Montpellier Cedex, France;* [3]*Centre for Veterinary Science, FISH Lab, Cambridge, UK and* [4]*IRD, CBGP (UMR IRD/ INRA/CIRAD/Montpellier SupAgro), Dakar, Senegal*

By suppressing recombination and reducing gene flow, chromosome inversions favor the capture and protection of advantageous allelic combinations, leading to adaptive polymorphisms. However, studies in non-model species remain scarce. Here we investigate the distribution of inversion polymorphisms in the multimammate rat *Mastomys erythroleucus* in West Africa. More than 270 individuals from 52 localities were karyotyped using G-bands and showed widespread polymorphisms involving four chromosome pairs. No significant deviations from Hardy–Weinberg equilibrium were observed either through space or time, nor were differences retrieved in viability or sex contribution between cytotypes. The distribution of chromosomal variation, however, showed perfect congruence with that of mtDNA-based phylogeographic clades. Thus, inversion diversity patterns in *M. erythroleucus* appeared more related to historical and/or demographic processes than to climate-based adaptive features. Using cross-species chromosome painting and G-banding analyses to identify homologous chromosomes in related out-group species, we proposed a phylogenetic scenario that involves ancestral-shared polymorphisms and subsequent lineage sorting during expansion/contraction of West African savannas. Our data suggest that long-standing inversion polymorphisms may act as regions in which adaptation genes may accumulate (nucleation model).
Heredity (2010) **104,** 378–386; doi:10.1038/hdy.2009.119; published online 7 October 2009

Introduction

In their earliest works dealing with 'supergenes' (see references quoted in Alvarez and Zapata, 1997), Dobzhansky (1948) underscored the function of recombination suppression during evolution, especially adaptation. Renewed interest in such processes has led to an increasing amount of studies focusing on its consequences from DNA to species levels (reviewed in Butlin, 2005). Chromosomal inversions are of particular interest because they are known to be major suppressors of recombination (reviews in Hoffman *et al.*, 2004; Ayala and Coluzzi, 2005), thus potentially providing important adaptive opportunities. Suppression of recombination may facilitate the capture of favorable combinations of alleles and may lead to the maintenance of inversion polymorphisms by selection (Kirkpatrick and Barton, 2006). Theoretical models (Alvarez and Zapata, 1997; Caceres *et al.*, 1999; Neiman and Linksvayer, 2006) as well as recent experimental evidences (see Ayala and Coluzzi, 2005, and references therein) have clearly supported such a view.

These processes have been advocated to explain several adaptive gradients that usually parallel geographic clines associated with inversion clines. One of the most striking examples may be the 'replicated time series of evolution in action' in *Drosophila subobscura* (Prevosti *et al.*, 1988; Gilchrist *et al.*, 2004), but other fascinating cases have also been identified in other animal species (for example, grasshoppers in Shaw *et al.*, 1985; *Drosophila* in Hoffman *et al.*, 2004; *Anopheles* in Coluzzi *et al.*, 2002; humans in Stefansson *et al.*, 2005). In some instances, adaptive gradients may result in considerable divergence such that they finally lead to speciation (for example, *Drosophila*: Noor *et al.*, 2001; Machado *et al.*, 2002; *Anopheles*: Coluzzi *et al.*, 2002; *Rhagoletis*: Feder *et al.*, 2005).

Because of their ability to provide adaptive opportunities, polymorphic inversions are not expected to be rare, even when they induce slight meiotic perturbations (Alvarez and Zapata, 1997). In effect, they have been extensively described in several animal groups that benefit either from favorable cytogenetic approaches, such as polytene chromosomes in diptera, or from whole sequenced genomes as is the case for humans (Kidd *et al.*, 2008). In other taxa, however, and especially in wild mammals (for example, Paresque *et al.*, 2007), wide-scale studies are quite rare, most probably due to the lack of reliable biological models as well as technical time-consuming aspects. Hence, the function of inversions in evolutionary processes still needs to be investigated in a larger range of lineages (Rieseberg, 2001; Navarro and Gazave, 2005).

Multimammate rats of the genus *Mastomys* (Rodentia, Murinae) range throughout most of Africa south of the Sahara desert, except in dense forest areas (Wilson and Reeder, 2005). They are known as major pests for crops and are reservoirs for several pathogenic agents (for example, Lassa fever virus: Lecompte *et al*., 2006). Pericentric inversions were proposed to have been pivotal in the evolution of the genus (Britton-Davidian *et al*., 1995; Volobouev *et al*., 2001), potentially driving speciation (Volobouev *et al*., 2001, 2002). Inversion polymorphism has been documented in several species (Britton-Davidian *et al*., 1995; Volobouev *et al*., 2001), but wide-scale studies are lacking. Here we provide a population cytogenetic analysis in *Mastomys erythroleucus*, a species that is distributed from Senegal to Ethiopia and south to Uganda, with a relic population in Morocco (Wilson and Reeder, 2005). We characterized pericentric inversion polymorphisms involving four chromosomal pairs in populations from West to Central Africa. The distribution pattern of chromosomal diversity was established and examined for signatures of several microevolutionary forces: (1) balancing selection by testing for heterozygote excesses and viability differences between cytotypes; (2) stochastic processes by comparing chromosome population structure to a molecular-based phylogeographic survey updated from Brouat *et al*. (2009) and (3) long-term persistence by identifying the ancestral chromosomal forms in the genus.

Materials and methods

Sampling effort

Accessibility to the field and live specimens are the main obstacles for wide-scale cytogenetic surveys in wild mammals. To karyotype a large number of individuals from as many localities as possible, we organized several field trips in West and Central Africa. Rodents were live-trapped using wire-mesh traps, brought to the nearest station of the 'Institut de Recherche pour le Développement' (Bamako, Mali; Dakar, Senegal; Niamey, Niger; Maroua, Cameroon) where they were killed and karyotyped. Preparations were then treated and analyzed in our laboratories in France. In total, 52 localities in 7 countries (Senegal, Burkina-Faso, Benin, Mali, Niger, Cameroon and Chad) were sampled, and 272 *M. erythroleucus* were karyotyped and analyzed using G-banding (see Supplementary Table). Two localities in Mali were sampled repeatedly for other demographic studies: Baguinéda in May and November 2003, May 2005 and June 2006; and Sélingué in June and November 2003. These temporal surveys allowed us to analyze changes in inversion frequencies over two (Sélingué) to at least five (Baguinéda) successive generations and several seasons.

Detailed information for each individual (external measurements, sex, reproductive state) and locality (GPS coordinates) were collected and compiled in the Sahelo-Sudanian Rodent Database (BDRSS; http://www.bdrss.ird.fr/bdrsspub_form.php). Voucher specimens are deposited in the collection of the Centre de Biologie et de Gestion des Populations (CBGP, Montpellier, France) and are available upon request.

Karyotyping and mtDNA typing

Each individual was karyotyped following the air-drying method (Evans *et al*., 1963). Briefly, animals were yeast-stimulated overnight, and then injected with vinblastine sulfate (10 mg in 40 ml NaCl 8.5‰) 45–60 min before killing. Bone marrow was collected, placed in a hypotonic solution (KCl 0.75 M) for 20 min at ∼37 °C, fixed using successive hand centrifugation and renewal of fixative solution (acetic acid/methanol 1:3) and finally kept at 4 to −20 °C when possible. G-bands were obtained following a slightly modified protocol from Seabright (1971): slides were aged overnight at 65 °C, rehydrated in 2 × SSC at 60 °C for 60–90 min, denatured in a trypsin solution (0.025 M) at 11–13 °C for 75–90 s, rinsed in distilled water and finally stained using buffered 4% Giemsa. For each specimen, at least five G-banded metaphases were captured using a Zeiss microscope coupled to the Genus software (CytoVision), and used to prepare a minimum of two karyotypes.

Brouat *et al*. (2009) recently performed a phylogeographic analysis of *M. erythroleucus* based on complete cytochrome *b* (*cytb*) sequences. In total, 22 localities that they used are also included in this study (Supplementary Table). In localities common to their and our own surveys but represented by different individuals, we considered all specimens trapped in the same locality (usually during the same field trip, often within the same trap lines) as belonging to the same *cytb* lineage *sensu* Brouat *et al*. (2009). In the localities ($n = 27$) not included in their molecular study, we sequenced at least one specimen per locality for the complete *cytb* mitochondrial gene (Supplementary Table). To assign the latter specimens to a phylogeographic group, we then added these sequences to data set of Brouat *et al*. (2009) and performed phylogenetic analyses following the same procedures (*M. natalensis* and *M. huberti* as out-groups; neighbor-joining with Tamura–Nei distances, maximum parsimony and maximum-likelihood analyses performed using PAUP version 4.0b (Sunderland, MA, USA); see Brouat *et al*., 2009, for details). Three localities (Gainsoa and Sotuba in Mali; Ile Madeleine in Senegal), not included in the previous phylogeographic survey, could not be molecularly investigated due to the unavailability of tissue samples.

Population cytogenetics

Chromosomal variability was investigated by considering each inversion as a locus and its state (A for acrocentric or M for sub/metacentric) as an allele. Statistics that rely on population data were performed on localities that included at least six individuals ($n = 13$; Table 1). The software GENEPOP version 3.2a (Raymond and Rousset, 1995) was used to estimate population parameters (inversion frequencies, observed and expected heterozygosities) and to test for departure from Hardy–Weinberg (HW) expectations as well as for linkage disequilibrium among all pairs of inversions. In all cases, default options were implemented. Genotypic differentiation between localities was assessed by two methods: Fisher's exact test (GENEPOP v3.2a) and Wright's F_{ST} index, the significance of which was evaluated by permutation tests ($n = 5000$) using GENETIX version 4.05 (Belkhir *et al*., 1996–2004). Differences between observed and expected heterozygosities over all

Table 1 Chromosome variability parameters for samples of six or more individuals

Country	N	Locality	n	Chromosomal inversion								mtDNA clade	He	Ho
				2		9		10		17				
				A	M	A	M	A	M	A	M			
Senegal	1	Dielmo	8	0.12	0.88	0.37	0.63	1		0.90	0.1	A	0.29	0.15
	2	Fadiga	12	0.12	0.88	0.54	0.46	1		0.75	0.25	A	0.36	0.39
Mali	3	Soukoutali	10	0.05	0.95	0.60	0.40	1		0.9	0.1	A	0.25	0.30
	4	Sélingué	25	0.12	0.88	0.58	0.42	1		1		A	0.35	0.42
	5	Baguinéda	63	0.04	0.96	0.60	0.40	1		0.93	0.07	A	0.23	0.21
	6	Farabougou	7		1	0.69	0.31	1		1		A	0.43	0.37
	7	Niono	6		1	0.25	0.75	1		0.92	0.08	A	0.27	0.17
Burkina-Faso	8	Nazinon	7	0.36	0.64	1		1		0.93	0.07	B1	0.30	0.43
	9	Tapoa	6	0.58	0.42	1		1		0.90	0.10	B1	0.33	0.35
Niger	10	La Tapoa	9	0.89	0.11	1		1		1		B1	0.20	0.22
	11	Kojimairi	15	1		1		1		0.1	0.9	B2	0.18	0.20
Cameroon	12	Meskine	7	1		1		1		0.36	0.64	B2	0.46	0.43
Chad	13	Zakouma	6		1	1			1	0.40	0.60	C	0.48	0.80
			Mean	0.33	0.67	0.74	0.26	0.92	0.08	0.78	0.22			

Abbreviations: A, acrocentric form; M, submetacentric form; *N*, number of specimens.
mtDNA phylogeographic clades are indicated (see text). He and Ho refer to mean expected and observed heterozygosity levels over the polymorphic inversions, respectively.

samples were tested by the Wilcoxon's signed-rank test. Viability and sex differences between genotypes were estimated using all available individuals. Age was assessed according to a regression of weight and head and body length taking sex into account (data not shown; Duplantier and Granjon, unpublished data). Where appropriate, the FDR (false discovery rate) correction for multiple tests was implemented following the procedure detailed in Verhoeven *et al*. (2005, p. 644). Population genetic structure was assessed using the Bayesian clustering algorithm implemented in the software BAPS4.14 (Corander and Marttinen, 2006), which partitioned the populations into an optimal number of clusters with associated values of the Bayes factor.

Assessment of the ancestral states in sister *Mastomys* species

Chromosomal homologies between several *Mastomys* species (that is, *M. coucha*, *M. kollmannspergeri*, *M. erythroleucus*, *M. huberti* and *M. natalensis*) have previously been established using G-bands (Britton-Davidian *et al*., 1995; Volobouev *et al*., 2002). However, some uncertainties may remain that cannot be resolved by this method. To unambiguously assess the chromosome homologies of the polymorphic chromosomal pairs between *M. erythroleucus* and congeneric West African species, we performed a cross-species chromosome painting (Zoo-FISH) survey. To do so, we established a fibroblast cell line from an intercostal biopsy of a male *M. erythroleucus* and used to sort chromosomes using a dual laser cell sorter (FAC-Star Plus; Becton Dickinson, Haarlem, Netherlands) as previously described (Yang *et al*., 2003). Chromosome-specific painting probes were obtained through degenerate oligonucleotide-primed PCR (DOP-PCR) amplification of the flow-sorted chromosomes (Telenius *et al*., 1992). They were then biotin-labeled, and hybridized onto *M. erythroleucus* (ERY) metaphases to identify the paints that corresponded to the polymorphic chromosomal pairs. These chromosome-specific paints were then hybridized onto at least

five metaphases of *M. huberti*, *M. kollmannspergeri* and *M. natalensis* that were used as out-groups. All DOP-PCR amplifications, hybridization and detection experiments as well as image processing followed Yang *et al*. (2003). Once the homologies between the four species were unambiguously established, the ancestral status (inverted versus noninverted) of these chromosomes was ascertained by comparison of G-banding patterns in the three congeneric out-group species.

Results

Chromosomal diversity and inversion frequencies

As expected from previous data (Britton-Davidian *et al*., 1995; Granjon *et al*., 1997; MER-1 *sensu*; Volobouev *et al*., 2001), the diploid number was $2N = 38$ for all 272 *M. erythroleucus* except for five individuals (three from Senegal, one from Mali and one from Burkina-Faso) which possessed 1–2 B chromosomes ($2N = 39–40$; data not shown). The number of autosomal arms (NFa), however, varied from 50 to 56. G-banded karyotypes were prepared for all specimens following the nomenclature described by Britton-Davidian *et al*. (1995). Variation in NFa was attributed to the presence of pericentric inversions on four chromosome pairs: ERY2, ERY9, ERY10 and ERY17 that exhibited an A and/or an M morphology (Figure 1). The state of these four chromosome pairs was scored in all individuals except nine for which the morphology and banding pattern of the short pair ERY17 could not be unambiguously assessed (see Supplementary Table).

Inversion frequencies are presented in Table 1 (populations with $n > 6$) and Supplementary Table (all populations). Diversity levels varied between chromosome pairs, the two extremes being ERY17, which showed a widespread polymorphism, and ERY10, for which no intrapopulation variability was recorded (see Figures 2a–d). Deviations from HW equilibrium were tested per chromosome pair and per sample, and no significant F_{IS} was observed (all $P > 0.06$; with
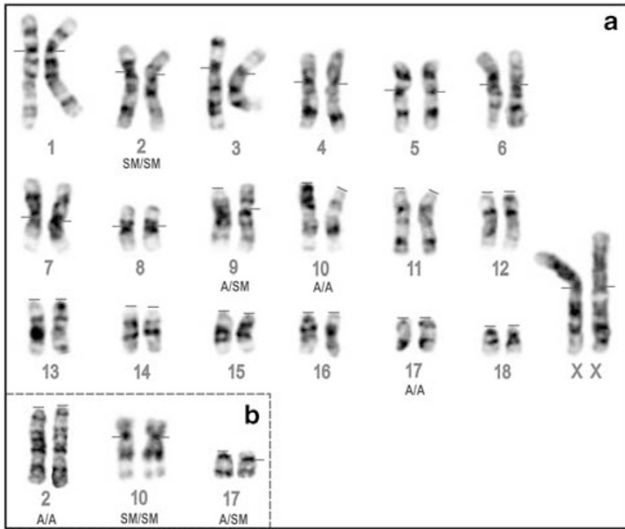
Figure 1 G-banded karyotype (**a**) of *Mastomys erythroleucus* female from Sélingué (Mali) showing homozygous acrocentric AA (pairs 10 and 17), homozygous submetacentric MM (pair 2) and heterozygous AM (pair 9) states. The inset (**b**) shows the alternative variants for pairs 2, 10 and 17. Hyphens indicate centromeres.

H1 = heterozygote excess, $P > 0.26$). In addition, no significant linkage disequilibrium was detected between chromosomes within all samples except in one case in Fadiga (ERY2 and ERY9, $P = 0.04$), which was no longer significant once the FDR correction was applied. Similarly, testing for departures from HW expectations among the different temporal sessions (season or year) in Baguinéda and Sélingué revealed no significant deviations (all $P > 0.3$). Exact tests of differentiation between these different sessions clearly showed no significant differences (each inversion, all $P > 0.05$; all chromosomes, $P > 0.1$). Expected heterozygosity (He) per locality averaged over polymorphic inversions varied from 0.18 to 0.48 (Table 1). The distribution of He versus Ho (observed heterozygosity for polymorphic inversions) over all populations showed no significant skews (Wilcoxon's signed-rank test value, all inversions, $P = 0.32$; per inversion, all $P > 0.10$). Two further tests were carried out to detect a contribution of balancing selection through overdominance in maintaining the inversion polymorphisms. The first one involved testing for differential viability by comparing the number of homozygotes and heterozygotes in juvenile ($n = 64$, 26 and 65 for ERY2, ERY9 and ERY17, respectively) versus adult ($n = 179$, 146 and 177 for ERY2, ERY9 and ERY17, respectively) individuals for each clade where polymorphic inversions had been observed. In the second one, a differential allelic contribution between the sexes (121 females and 132 males) was tested by estimating genotypic frequencies per inversion combined over all individuals or per clade. Neither test yielded significant results: no significant differences were observed between age classes (χ^2-test; ddl = 3, $P > 0.05$), or between sexes (G-test of independence, all $P > 0.17$)

Geographic patterns of chromosomal differentiation

As recapitulated in Figures 2a–d, the distribution of chromosomal diversity varied geographically in a non-random manner. In particular, chromosomal polymorphism showed several west-eastern trends of variation. For instance, the ERY9M and ERY17^A forms were widespread in the west, but disappeared or became rarer eastward. Likewise, the frequency of the M form of ERY2 was very high in the westernmost and easternmost populations but lower in the central populations. A clear discontinuity was evident in the easternmost population (Chad), which showed fixation of alternative chromosome forms for ERY2 and ERY10.

Analysis of chromosomal population structure by the Bayesian approach partitioned the populations into four geographic clusters (Table 2). These different groups gathered populations along a west–east axis and were supported by a strong assignment value (log marginal likelihood = −410.872, $P = 0.84$). Samples from Senegal and Mali were clustered into one group, those from Burkina-Faso and the westernmost one from Niger into a second one, the easternmost sample from Niger clustered with the one from Cameroon in a third group and finally the sample from Chad formed the fourth group. This longitudinal gradient of differentiation closely resembled the phylogeographical structure recently established using mitochondrial *cytb* sequences (Brouat *et al.*, 2009). To further investigate this similarity, we sequenced the *cytb* gene of representatives from all (but three) newly sampled localities (GenBank accession numbers ranging between GQ409938 and GQ409966) and included in Brouat *et al.*'s (2009) data set, thus allowing the unambiguous attribution of each locality to one of the phylogeographic lineages (see Supplementary Figure and Supplementary Table). As expected, the newly sampled localities that fell within the geographic range of the phylogeographic lineages *sensu*, Brouat *et al.* (2009) , were all shown to belong to the expected clade (see Supplementary Figure, Supplementary Table and Figure 2) whatever the methods used (NJ, MP or ML; data not shown). The mtDNA sequences indicated that three major mtDNA lineages were present in our data (namely clades A, B and C in Supplementary Table, Supplementary Figure and Figure 2; cf. Brouat *et al.*, 2009). The phylogeographic structure established using *cytb* was thus fully concordant with the partitioning of chromosomal diversity by the Bayesian analysis which further retrieved two subclades within clade B (B$_1$, represented here by populations 8, 9 and 10, and B$_2$, represented here by populations 11 and 12; see Tables 1 and 2). This geographic structure of chromosomal variation was also supported by the mean values of karyotypic differentiation within and between clades (Table 3). Mean F_{ST} was lowest within clade A (0.033 ± 0.003; none of the values was significantly different from zero), and reached only slightly larger values within subgroups of clade B (B$_1$, 0.191 ± 0.116; B$_2$, 0.153). Between-cluster chromosomal differentiation was much higher in all cases ($0.435 \pm 0.038 < F_{ST} < 0.842 \pm 0.046$) and followed a west-east trend of increasing values (that is, A versus B$_1$, 0.435 ± 0.038; B$_2$ versus C, 0.842 ± 0.046).

Chromosome painting and assessment of ancestral states between A and M forms

The flow-sort karyotype of *M. erythroleucus* could be resolved into 19 different peaks, each of which was
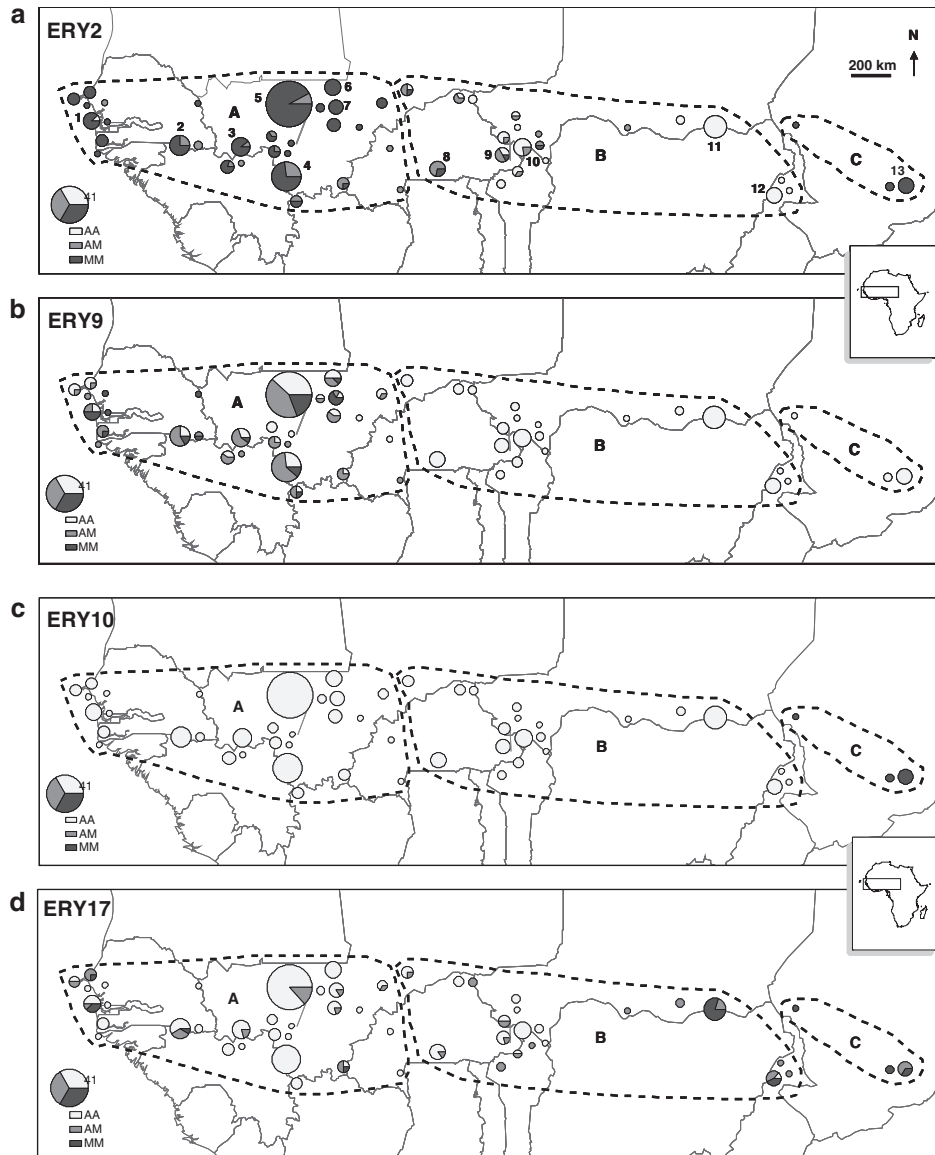
Figure 2 Distribution maps of cytotype frequencies observed for (**a**) ERY2, (**b**) ERY9, (**c**) ERY10 and (**d**) ERY17 across West Africa. 'AA', 'AM' and 'MM' correspond to the acrocentric homozygote, heterozygote and meta/submetacentric homozygote forms, respectively. See text for details. Circle sizes are proportional to sample sizes. The three mtDNA phylogeographic clades (**a–c**) are indicated in dashed lines. Numbers refer to localities in Tables 1 and 2.

unambiguously identified through Zoo-FISH experiments onto *M. erythroleucus* metaphases. Results will be detailed elsewhere (Dobigny *et al.*, in preparation); however, for the purpose of the present study, the four paints that included ERY2, ERY9, ERY10 and ERY17 were successfully isolated and hybridized onto *M. huberti* (MHU), *M. kollmannspergeri* (MKO) and *M. natalensis* (MNA) chromosome preparations, thus allowing for the identification of homologous segments in the other *Mastomys* species (see an example in Figure 3). In essence, our Zoo-FISH results showed that ERY2, 9, 10 and 17 were all conserved as entire segments in the other *Mastomys* species: ERY2 was homologous to MHU2q dist, MKO2 dist and MNA4; ERY9 corresponded to MHU1q dist, MKO1 dist and MNA3q dist; ERY10 was homologous to MHU1q prox, MKO1 prox and MNA3p + 3q prox; ERY17 corresponded to MHU7,

MKO18 and MNA14. Observed homologies were in good agreement with previously published comparative maps based on G-banding patterns (Britton-Davidian *et al.*, 1995).

G-band comparisons of ERY2, ERY9, ERY10 and ERY17 with their homologues in the three other *Mastomys* species showed that the ancestral forms were acrocentric for both ERY2 and ERY9 because their homologues showed an A-like banding pattern (that is, collinear with the acrocentric form) in all three outgroups. In the two other cases, the ancestral form could not be identified, because the ERY10 homologues showed either an A-like or an M-like pattern in *M. huberti*, *M. kollmannspergeri* and *M. natalensis*, and the homologues of ERY17 were submetacentric in *M. kollmannspergeri* and *M. huberti*, but polymorphic in *M. natalensis*.

Table 2 Population composition of the four clusters retrieved by the Bayesian analysis of population structure (BAPS)

Bayesian group	Country	Locality	Clade A 1	Clade B 2(B1)	Clade B 3(B2)	Clade C 4
1	Senegal	Dielmo (1)	0	−21.7	−50.6	−29.7
		Fadiga (2)	0	−20.6	−52.4	−27.7
	Mali	Soukoutali (3)	0	−22.2	−61.2	−31.0
		Sélingué (4)	0	−27.3	−97.4	−46.5
		Baguinéda (5)	0	−51.9	−144.4	−56.3
		Farabougou (6)	0	−17.9	−58.5	−27.7
		Niono (7)	0	−24.7	−54.9	−26.6
2	Burkina-Faso	Nazinon (8)	−8.3	0	−25.3	−22.3
		Tapoa (9)	−16.1	0	−17.3	−25.6
	Niger	La Tapoa (10)	−36.9	0	−15.7	−36.4
3		Kojimairi (11)	−127.1	−36.4	0	−49.9
	Cameroon	Meskine (12)	−52.6	−13.4	0	−34.1
4	Chad	Zakouma (13)	−58.3	−39.6	−52.1	0

Locality numbers are indicated in brackets following the same nomenclature as on Figure 2.
Log (marginal likelihood) values associated with changes in cluster assignment of populations are indicated. The corresponding mtDNA phylogroups (that is A, B, C) and sublineages (that is B1, B2) are provided.

Table 3 F_{ST} values of chromosomal differentiation between and within mtDNA phylogeographic lineages (A–C) and sublineages (B$_1$, B$_2$)

Country	Locality		Clade A 2	A 3	A 4	A 5	A 6	B$_1$ 7	B$_1$ 8	B$_1$ 9	B$_1$ 10	B$_2$ 11	B$_2$ 12	C 13
Senegal	1	Dielmo	−0.022	−0.006	0.026	0.036	0.080	−0.082	**0.320**	**0.369**	**0.632**	**0.798**	**0.634**	**0.665**
	2	Fadiga		−0.011	0.045	0.031	0.057	0.063	**0.207**	**0.271**	**0.521**	**0.694**	**0.543**	**0.582**
Mali	3	Soukoutali			−0.008	−0.029	−0.024	0.095	**0.235**	**0.340**	**0.620**	**0.799**	**0.663**	**0.672**
	4	Sélingué				0.002	0.003	0.124	**0.235**	**0.342**	**0.564**	**0.771**	**0.666**	**0.696**
	5	Baguinéda					−0.017	0.114	**0.258**	**0.382**	**0.579**	**0.749**	**0.675**	**0.679**
	6	Farabougou						0.222	**0.270**	**0.414**	**0.725**	**0.871**	**0.754**	**0.757**
	7	Niono							**0.514**	**0.556**	**0.784**	**0.882**	**0.770**	**0.769**
Burkina-Faso	8	Nazinon								0.012	**0.408**	**0.772**	**0.558**	**0.716**
	9	Tapoa									0.153	**0.703**	**0.398**	**0.715**
Niger	10	La Tapoa										**0.803**	**0.534**	**0.871**
	11	Kojimairi											0.153	**0.887**
Cameroon	12	Meskine												**0.796**
Chad	13	Zakouma												0
		Mean intraclade F$_{ST}$			0.033 ± 0.003					0.191 ± 0.116			0.153	0.796

Significant values are indicated in bold.

Discussion

Inversion polymorphisms and their evolutionary implications in animals have been extensively investigated in dipterans, essentially *Drosophila* and *Anopheles* species (see references in Introduction). In these species, clinal variation in the frequency of inversions has provided strong evidence of natural selection and adaptation to different gradients of environmental variables. In mammals, however, studies that rely on a reasonable amount of individuals are extremely scarce. Some of the most extensive surveys have focused on humans (Stefansson et al., 2005) and, to a lesser extent, on the agile gibbon *Hylobates agilis* (Van Tuinen et al., 1999). Among rodents, several studies have been performed (for example, Dowler, 1989) although the latter often involved incipient species complexes rather than a single well-defined polymorphic species. Two cases are of note although they both mostly relied on conventional non-banded

karyotypes: the black rat in Asia ($n = 353$ individuals; Yosida et al., 1971a, b; Yosida, 1976) and the Brazilian pygmy rice rat *Oligoryzomys nigripes* (Bonvicino et al., 2001; Paresque et al., 2007; $n = 467$). In both models, clinal variation of a pericentric inversion along a north-south gradient suggested a possible adaptive value of this chromosomal variation, although alternative historical causes were also considered as compatible (Yosida et al., 1971a, b; Yosida, 1976; Bonvicino et al., 2001; Paresque et al., 2007). As such, our study is one of the most comprehensive surveys of chromosomal inversion polymorphism in wild mammals fully relying on G-banding patterns, and to the best of our knowledge, it is the first one dealing with African rodents.

In *M. erythroleucus*, karyotypic variation is evidenced for four chromosome pairs, all involving pericentric inversions. Maintenance of polymorphisms within populations can result from balancing selection due to heterozygote advantage, differential selection along
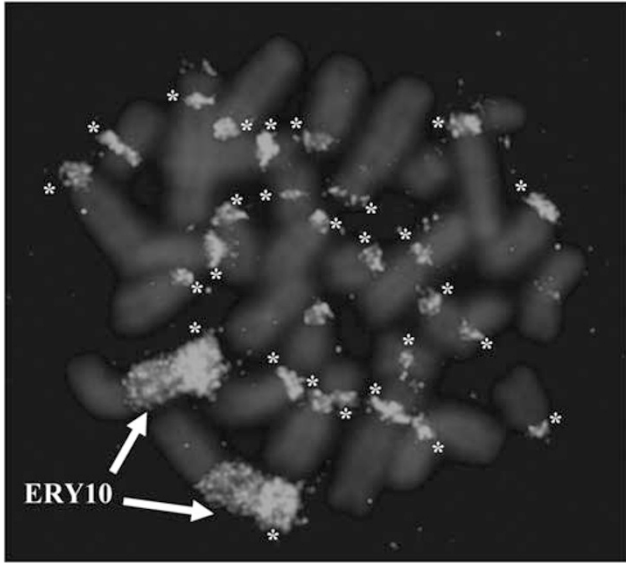
Figure 3 Example of interspecific chromosome painting: here, hybridization of a *M. erythroleucus* paint that contains the pair 10-sorted chromosome onto a *M. huberti* metaphase. Arrows indicate the hybridized homologous region, that is, region 1q prox, in *M. huberti* (see text for details). Asterisks correspond to pericentromeric repeats that are shared between *Mastomys* species, and so cross-hybridize (Dobigny, unpublished data).

environmental gradients or from stochastic population processes if the inversions have little or no effect on fitness. Tests for departures from HW proportions indicated that no significant deviation was present, nor were significant differences detected in genotypic frequencies between sexes, age classes, seasons or generations. Thus, our study provides no evidence supporting a selective maintenance of the inversion polymorphisms. However, it should be mentioned that the small size of most of the populations studied precluded the detection of low levels of significance, although the tests were diversified in consequence.

Heterozygosity for chromosomal inversions is expected to affect the fertility of carriers, because recombination within the inversion will generate unbalanced gametes with deletions or duplications of genetic material. Data from the literature indicate that the reproductive consequences vary from severe to nonexistent (Gabriel-Robez *et al.*, 1988; Jaarola *et al.*, 1998; Bonvicino *et al.*, 2001; Anton *et al.*, 2006; Paresque *et al.*, 2007). Two features are known to lessen the fitness disadvantage: the length of the inversion and the type of meiotic pairing. Several studies have highlighted a positive correlation between reproductive impairment and the size of the inversion (Hale, 1986; Navarro and Ruiz, 1997; Anton *et al.*, 2006). In addition, meiotic analyses in different species have evidenced nonhomologous pairing of the inverted region through meiotic heterosynaptic adjustment, thus suppressing loop formation and subsequent crossing-over in this region, so that no recombinant gametes are formed (Hale, 1986; Gabriel-Robez *et al.*, 1988; cf. Fagundes and Yonenaga-Yassuda, 1998; Bonvicino *et al.*, 2001). In *M. erythroleucus*, although no meiotic analyses are available, the inversion sizes are less than 50% of that of the chromosomes

involved, which is the limit in *Peromyscus* below which there is low incidence on the fertility of heterozygotes (Hale, 1986). However, the reproductive cost of these inversions would deserve to be assessed through relevant meiotic analyses and fertility screening in wild populations and/or experimental crosses.

The geographic distribution of chromosomal variability in *M. erythroleucus* was shown to vary nonrandomly along a west-east gradient closely matching the molecular signatures of *cytb* sequences (Brouat *et al.*, 2009; Supplementary Figure). Of the three main mtDNA lineages described, the westernmost phylogroup A (Senegal and Mali) shows the highest level of polymorphism involving three (ERY2, ERY9 and ERY17) of the four inversions. The transition to phylogroup B in Mali is perfectly retrieved by the fixation of ERY9^A (subclade B_1), and further east by that of ERY2^A (subclade B_2). The latter also coincides with the limit between two sublineages identified within clade B by *cytb* sequences. Finally, the molecular B and C lineages that are separated by Lake Chad and the Logone/Chari rivers are clearly distinguished by the fixation of alternative chromosomal forms of ERY2 and ERY10 (Figures 2a and c). Although chromosomal diversity was not tested against an environmental gradient, the West African climatic regime is clearly characterized by marked latitudinal clines, suggesting that if the distribution of inversion polymorphism were shaped by adaptation to such a gradient, latitudinal variation in inversion frequencies would have been expected. Rather, the strong congruence between the cytogenetic and the molecular patterns suggests that the partitioning of inversion diversity and polymorphisms in *M. erythroleucus* may be more related to historical and/or demographic processes rather than to climate-based selective pressures. Such an evolutionary pattern is in excellent agreement with Brouat *et al.* (2009) biogeographic hypothesis which suggests that *M. erythroleucus* populations were isolated in savanna fragments whereas West African forests were expanding northward during Pleistocene humid periods (De Menocal, 2004). During subsequent drier periods, savannas expanded again, allowing the once isolated multimammate rat populations to disperse and differentiate into the extant phylogeographic clades (Brouat *et al.*, 2009). These demographic bottlenecks within refuge populations are fully congruent with the proposed lineage sorting and geographic structuring of the cytotypes. Such a scenario implies that drift, rather than selection or migration, appears as the major evolutionary force that has shaped chromosomal inversion distribution in *M. erythroleucus*.

From there, support for a historic origin of the inversion polymorphism patterns in *M. erythroleucus* enables us to propose a phylogenetic evolutionary scheme of their distribution within the extant clades and species (Figure 4). Such an endeavor is possible because homologous chromosomes were identified in the other *Mastomys* species that are here used as out-groups. It also requires the *ad hoc* but reasonable assumption that each of the inversions has appeared only once (no convergence). The most parsimonious scenario supports the existence of ancestral polymorphisms and lineage sorting of one or the other or both chromosome forms in the different phylogroups during phases of isolation and population contraction. The ancestral polymorphisms
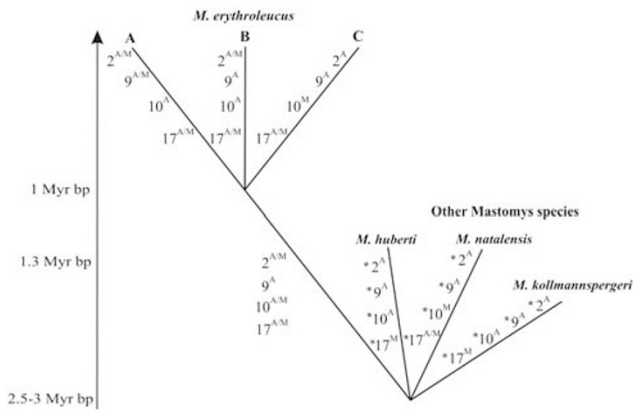
Figure 4 Summary of the proposed evolutionary scenario for ERY2, ERY9, ERY10 and ERY17. In each case, 'A' and 'M' correspond to the acrocentric and meta/submetacentric forms, respectively. Approximate dates are from Brouat *et al*. (2009) and Mouline *et al*. (2008). Asterisks refer to the chromosome or chromosome segments homologous to those in *M. erythroleucus*.

are either specific to the *M. erythroleucus* ancestral pool (for example, ERY2 and ERY9 in Figure 4) or shared with the other *Mastomys* species (for example, ERY10 and ERY17). The maintenance of polymorphisms over time is compatible with inversions conferring a neutral or near-neutral selective status, particularly in the case of ERY2 and ERY9, which are restricted to *M. erythroleucus* (that is, are more recent), and of ERY10, which shows fixation of alternative forms in the different clades and species. However, the phylogenetic scenario indicates that the shared inversion polymorphism involving ERY17 and its homologue in *M. natalensis* would have persisted for at least 2.3–3 Myr according to molecular inferences (Mouline *et al*., 2008; Brouat *et al*., 2009); such a long-term persistence through series of population contractions may require the intervention of selective processes and/or large effective population sizes (Asthana *et al*., 2005). Alternatively, the comparative analysis of inversion breakpoints in the different species may shed some light on the possibility of convergent chromosomal evolution in this genus.

In a more general manner, our results support a major role of demographic and historic factors in the current patterns observed for inversion polymorphism in *M. erythroleucus*. In the black rat and the Brazilian *O. nigripes* (see references), the lack of a significant association with ecological or environmental features suggests that the inversion polymorphisms in these rodents may be behaving in a near environmentally neutral manner at least during a transient period. The persistence of ancestral inversion polymorphisms within *M. erythroleucus* would follow such a pattern. During such an initial neutral phase, the distribution of the inversions would tend to reflect historical or demographic events rather than selective gradients. If heterosynapsis removes the reproductive cost of inversion heterozygotes by inhibiting recombination, a (near)neutral inversion polymorphism may set the stage for the capture and accumulation of adapted allelic combinations when, and if, they appear. Because no evidence supporting the function of inversions as promoters of adaptation (that is, the so-called facilitation model; Noor *et al*., 2001; Rieseberg,

2001; McAllister *et al*., 2008) was retrieved in our study, this two-step process may be more consistent with the nucleation model *sensu* McAllister *et al*. (2008) in which chromosomal rearrangements act as regions where adaptation/isolation genes accumulate.

References

Alvarez G, Zapata C (1997). Conditions for protected inversion polymorphism under supergene selection. *Genetics* **146**: 717–722.

Anton E, Vidal F, Egozcue J, Blanco J (2006). Genetic reproductive risk in inversion carriers. *Fertil Steril* **85**: 661–666.

Asthana S, Schmidt S, Sunyaev S (2005). A limited role for balancing selection. *Trends Genet* **21**: 30–32.

Ayala FJ, Coluzzi M (2005). Chromosome speciation: humans, *Drosophila*, mosquitoes. *Proc Natl Acad Sci USA* **102**: 6535–6542.

Belkhir K, Borsa P, Raufaste N, Bonhomme F (1996–2004). *GENETIX 4.05 logiciel sous Windows TM pour la génétique des populations*. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II: Montpellier, France.

Bonvicino C, D'Andrea P, Borodin P (2001). Pericentric inversion in natural populations of *Oligoryzomys nigripes* (Rodentia, Sigmodontinae). *Genome* **44**: 791–796.

Britton-Davidian J, Catalan J, Granjon L, Duplantier JM (1995). Chromosomal phylogeny and evolution in the genus *Mastomys* (Mammalia, Rodentia). *J Mammal* **76**: 248–262.

Brouat C, Tatard C, Bâ K, Cosson JF, Dobigny G, Fichet-Calvet E *et al*. (2009). Phylogeography of the Guinea multimammate mouse (*Mastomys erythroleucus*): a case study for Sahelian species in West Africa. *J Biogeogr* (doi:10.1111/j.1365-2699.2009.02184x).

Butlin R (2005). Recombination and speciation. *Mol Ecol* **14**: 2621–2635.

Caceres M, Barbadilla A, Ruiz A (1999). Recombination rate predicts inversion size in Diptera. *Genetics* **153**: 251–259.

Coluzzi M, Sabatini A, Della Torre A, Di Deco M, Tetrarca V (2002). A polytene chromosome analysis of the *Anopheles gambiae* species complex. *Science* **298**: 1415–1418.

Corander J, Marttinen P (2006). Bayesian identification of admixture events using multi-locus molecular markers. *Mol Ecol* **15**: 2833–2843.

De Menocal P (2004). African climate change and faunal evolution during the Pliocene-Pleistocene. *Earth Planet Sci Lett* **220**: 3–24.

Dobzhansky T (1948). Genetics of natural populations. XVIII. Experiments on chromosomes of *Drosophila pseudoobscura* from different geographic regions. *Genetics* **33**: 588–602.

Dowler RC (1989). Cytogenetic studies of three chromosomal races of pocket gophers (*Geomys bursarius* complex) at hybrid zones. *J Mammal* **70**: 253–266.

Evans EP, Breckon G, Ford CE (1963). An air-drying method for meiotic preparations from mammalian testes. *Cytogenetics* **3**: 289–294.

Fagundes V, Yonenaga-Yassuda Y (1998). Evolutionary conservation of whole homeologous chromosome arms in the Akodont rodents *Bolomys* and *Akodon* (Muridae, Sigmodontinae): maintenance of interstitial telomeric segments (ITBs) in recent event of centric fusion. *Chromosome Res* **6**: 643–648.

Feder J, Xie X, Rull J, Velez S, Forbes A, Leung B *et al.* (2005). Mayr, Dobzhansky, and Bush and the complexities of sympatric speciation in *Rhagoletis*. *Proc Natl Acad Sci USA* **102**: 6573–6580.

Gabriel-Robez O, Ratomponirina C, Croquette M, Couturier J, Rumpler Y (1988). Synaptonemal complexes in a subfertile man with a pericentric inversion in chromosome 21. *Cytogenet Cell Genet* **48**: 84–87.

Gilchrist GW, Huey RB, Balanya J, Pascual M, Serra L (2004). A time series of evolution in action: a latitudinal cline in wing size in South American *Drosophila subobscura*. *Evolution* **58**: 768–780.

Granjon L, Duplantier JM, Catalan J, Britton-Davidian J (1997). Systematics of the genus *Mastomys* (Thomas, 1915) (Rodentia, Muridae): a review. *Belg J Zool* **127**: 7–18.

Hale DW (1986). Heterosynapsis and suppression of chiasmata within heterozygous pericentric inversions of the Sitka deer mouse. *Chromosoma* **94**: 425–436.

Hoffman AA, Sgro CM, Weeks AR (2004). Chromosomal inversion polymorphisms and adaptation. *Trends Ecol Evol* **19**: 483–488.

Jaarola M, Martin RH, Ashley T (1998). Direct evidence for suppression of recombination within two pericentric inversions in human: a new sperm-FISH technique. *Am J Hum Genet* **63**: 218–224.

Kidd J, Cooper G, Donahue W, Hayden H, Sampas N, Graves T *et al.* (2008). Mapping and sequencing of structural variation from eight human genomes. *Nature* **453**: 56–64.

Kirkpatrick M, Barton N 2006)). Chromosome inversions, local adaptation and speciation. *Genetics* **173**: 419–434.

Lecompte E, Fichet-Calvet E, Daffis S, Koulema K, Sylla O, Kourouma F *et al.* (2006). *Mastomys natalensis* and Lassa fever, West Africa. *Emerg Infect Dis* **12**: 178–183.

Machado C, Kliman R, Markert J, Hey J (2002). Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and close relatives. *Mol Biol Evol* **19**: 472–488.

McAllister BF, Sheeley SL, Mena PA, Evans AL, Schlötterer C (2008). Clinal distribution of a chromosomal rearrangement: a precursor to chromosomal speciation? *Evolution* **62**: 1852–1865.

Mouline K, Granjon L, Galan M, Tatard C, Abdoullaye D, Ag Atteyine S *et al.* (2008). Phylogeography of Sahelian rodent species *Mastomys huberti*: a Plio-Pleistocene story of emergence and colonization of humid habitats. *Mol Ecol* **17**: 1036–1053.

Navarro A, Gazave E (2005). Inversions with classical style and trendy lines. *Nat Genet* **37**: 115–116.

Navarro A, Ruiz A (1997). On the fertility effects of pericentric inversions. *Genetics* **147**: 931–933.

Neiman M, Linksvayer T (2006). The conversion of variance and the evolutionary potential of restricted recombination. *Heredity* **96**: 111–121.

Noor M, Grams K, Bertucci L, Reiland J (2001). Chromosomal inversions and the reproductive isolation of species. *Proc Natl Acad Sci USA* **98**: 12084–12088.

Paresque R, de Jesus Silva M, Yonenaga-Yassuda Y, Fagundes V (2007). Karyological geographic variation of *Oligoryzomys nigripes* Olfers 1818 (Rodentia, Cricetidae) from Brazil. *Genet Mol Biol* **30**: 43–53.

Prevosti A, Ribo G, Serra L, Aguade M, Balana J, Monclus M *et al.* (1988). Colonization of America by *Drosophila subobscura*: experiment in natural populations that supports the adaptive role of chromosomal inversion polymorphism. *Proc Natl Acad Sci USA* **85**: 5597–5600.

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

Rieseberg L (2001). Chromosomal rearrangements and speciation. *Trends Ecol Evol* **16**: 351–358.

Seabright M (1971). A rapid banding technique for human chromosome. *Lancet* **2**: 971–972.

Shaw D, Coates D, Arnold M, Wilkinson P (1985). Temporal variation in the chromosomal structure of a hybrid zone and its relationship to karyotypic repatterning. *Heredity* **55**: 293–306.

Stefansson H, Helgason A, Thorleifsson G, Steinthorsdottir V, Masson G, Barnard J *et al.* (2005). A common inversion under selection in Europeans. *Nat Rev Genet* **37**: 129–137.

Telenius H, Carter NP, Bebb CE, Nordenskjöld M, Ponder BA, Tunnacliffe A (1992). Degenerate oligonucleotide-primed PCR: a general amplification of target DNA by a single degenerate primer. *Genomics* **13**: 718–725.

Van Tuinen P, Mootnick A, Kingswood S, Hale D, Kumamoto A (1999). Complex, compound inversion/translocation polymorphism in an ape: presumptive intermediate stage in the karyotypic evolution of the agile gibbon, *Hylobates agilis*. *Am J Phys Anthrop* **110**: 129–142.

Verhoeven KJF, Simonsen KL, McIntyre LM (2005). Implementing false discovery rate: increasing your power. *Oikos* **108**: 643–647.

Volobouev V, Hoffman A, Sicard B, Granjon L (2001). Polymorphism and polytypy for pericentric inversions in 38-chromosome *Mastomys* (Rodentia, Murinae) and possible taxonomic implications. *Cytogenet Cell Genet* **92**: 237–242.

Volobouev V, Aniskin A, Lecompte E, Ducroz JF (2002). Patterns of karyotype evolution in complexes of sibling species within three genera of African murid rodents inferred from the comparison of cytogenetic and molecular data. *Cytogenet Genome Res* **96**: 261–275.

Wilson DE, Reeder DAM (2005). *Mammal Species of the World: A Taxonomic and Geographic Reference*. The Johns Hopkins University Press: Baltimore.

Yang F, Alkalaeva EZ, Perelman PL, Pardini AT, Harrison WR, O'Brien PC *et al.* (2003). Reciprocal chromosome painting among human, aardvark and elephant (superorder Afrotheria) reveals the likely eutherian ancestral karyotype. *Proc Natl Acad Sci USA* **100**: 1062–1066.

Yosida TH, Tsuchiya K, Moriwaki K (1971a). Frequency of chromosome polymorphism in *Rattus rattus* collected in Japan. *Chromosoma* **33**: 30–40.

Yosida TH, Tsuchiya K, Moriwaki K (1971b). Karyotypic differences of black rats, *Rattus rattus*, collected in various localities of East and Southeast Asia and Oceania. *Chromosoma* **33**: 252–267.

Yosida TH (1976). Frequencies of chromosome polymorphism (pairs no. 1, 9 and 13) in the black rat of Japan. *Proc Jap Soc* **52**: 405–408.

Supplementary Information accompanies the paper on Heredity website (http://www.nature.com/hdy)