MOLECULAR ECOLOGY

Molecular Ecology (2014) **23**, 1749–1763 doi: 10.1111/mec.12687

Demographic processes shaping genetic variation of the solitarious phase of the desert locust

MARIE-PIERRE CHAPUIS,* CHRISTOPHE PLANTAMP,† LAURENCE BLONDIN,‡ CHRISTINE PAGÈS,‡ JEAN-MICHEL VASSAL‡ and MICHEL LECOQ‡

CIRAD, UMR CBGP, F-34398 Montpellier, France, †Laboratoire de Biométrie et Biologie Evolutive, Université Lyon 1, CNRS, UMR 5558, 69622 Villeurbanne, France, ‡CIRAD, UPR B-AMR, F-34398 Montpellier, France

Abstract

Between plagues, the solitarious desert locust (*Schistocerca gregaria*) is generally thought to exist as small populations, which are particularly prone to extinction events in arid regions of Africa and Asia. Given the high genetic structuring observed in one geographical area (the Eritrean coast) by former authors, a metapopulation dynamics model involving repeated extinction and colonization events was favoured. In this study, we assessed the validity of a demographic scenario involving temporary populations of the solitarious phase of the desert locust by analysing large-scale population genetic data. We scored 24 microsatellites in 23 solitarious population samples collected over most of the species range during remission. We found very little genetic structuring and little evidence of declining genetic diversity. A Bayesian clustering method distinguished four genetically differentiated units. Three groups were largely consistent with three population samples which had undergone recent bottleneck events. Nevertheless, the last genetically homogeneous unit included all individuals from the remaining 18 population samples and did not show evidence of demographic disequilibrium. An approximate Bayesian computation treatment indicated a large population size for this main genetic group, moderately reduced between plague and remission but still containing tens of thousands of individuals. Our results diverge from the hypothesis of a classical metapopulation dynamics model. They instead support the scenario in which large populations persist in the solitarious phase of the desert locust.

Keywords: Bayesian inference, bottleneck, microsatellite, Orthoptera, pest, population dynamics

Received 17 July 2013; revision received 28 January 2014; accepted 29 January 2014

Introduction

The desert locust, *Schistocerca gregaria*, is one of the most infamous and widespread pests globally, with a huge potential invasion area, spanning 31 million km^2 from West Africa to southwest Asia (Fig. 1A). Locusts at high densities live gregariously and actively form marching bands of hoppers or swarms of winged adults. These highly migratory groups are devastating for pastures and crops, as shown during the last two major desert locust invasions in Sahel-Saharan Africa in 1987–1988 and 2003–2005 (Fig. 1B).

Correspondence: Marie-Pierre Chapuis, CIRAD—BIOS, CBGP, Campus International de Baillarguet, CS 30016, 34988 Montferrier-sur-Lez Cedex, France, Fax: +33 (0)4 99 62 33 45; E-mail: marie-pierre.chapuis@cirad.fr

Between plagues, desert locust populations persist at low density in a harmless solitarious phase in the central part of the species range, where annual rainfall is <250 mm. During remission, large areas of the Sahel-Saharan deserts are believed to be clear of locusts (Waloff 1966), but only scarce data exist on the distribution, census sizes and migration rates, precluding the construction of a suitable model of solitarious desert locust population dynamics (Uvarov 1977). Yet, information on functioning of populations is crucial for designing effective management strategies, particularly in the early stages of gregarization, when small impacts on population size can have dramatic effects and populations can still be locally contained (Magor *et al.* 2008; Sword *et al.* 2010).

The most common view of solitarious population dynamics is that suitable habitats are spatially restricted

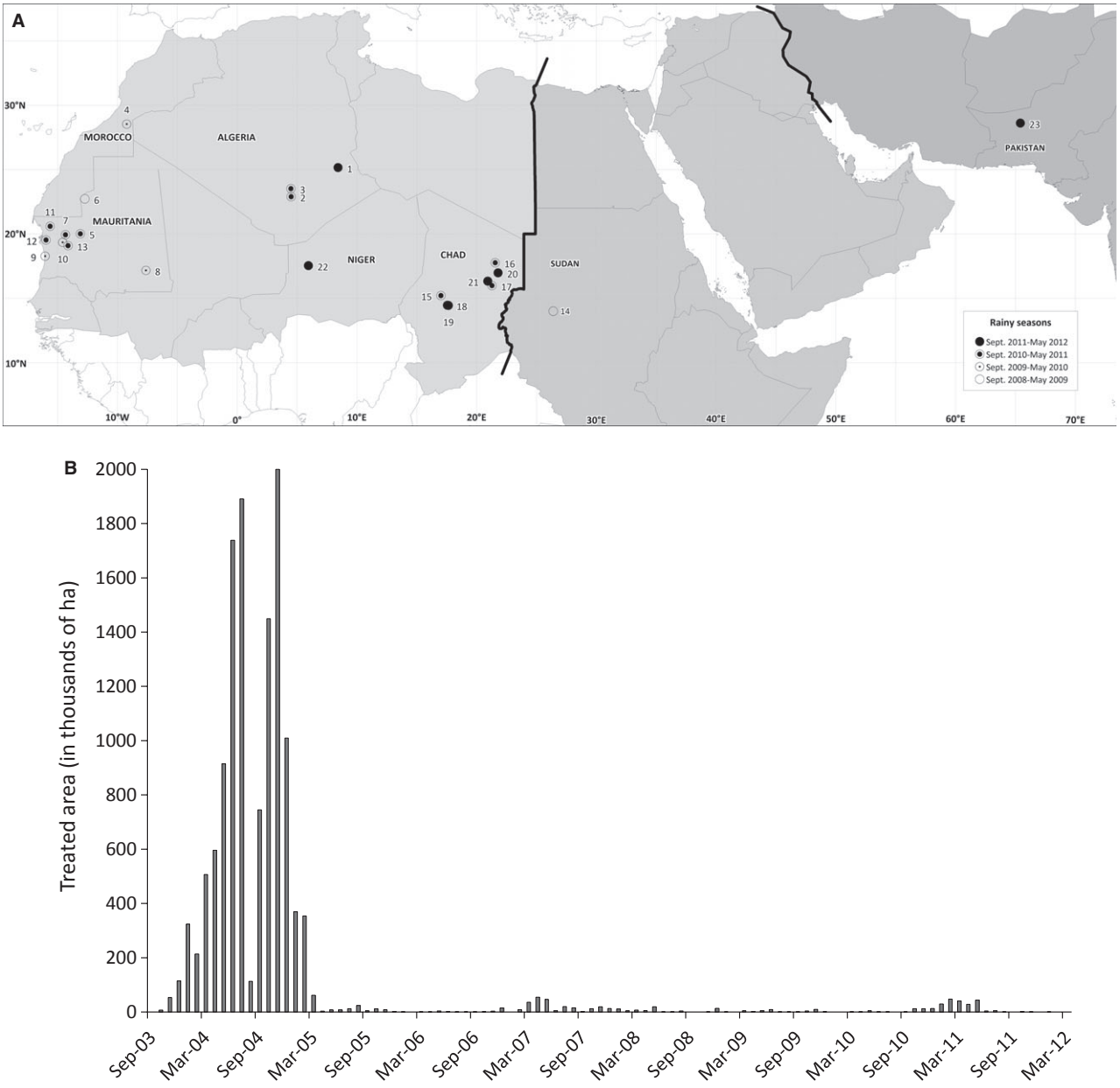

Fig. 1 Geographical distribution of the 23 *Schistocerca gregaria* population samples (A) and areas treated with pesticides during the last 10 years (B). (A) The three grey areas delimit the regions of FAO commissions for desert locust control: from left to right, Western region (population samples 1–13 and 15–22), Central region (population sample 14) and southwest Asian region (population sample 23). Populations were sampled by the following collectors: [1–3]Institut National de la Protection des Végétaux, Algeria, [4]National Unit of Locust Control of Morocco, [5,7,10–13]National Unit of Locust Control of Mauritania, [6,8–9]Centre de Coopération Internationale en Recherche Agronomique pour le Développement, France, [14]ICIPE—African Insect Science for Food and Health, Kenya,[15–21]National Unit of Locust Control of Chad, [22]National Unit of Locust Control of Niger, [23]National Unit of Locust Control of Pakistan. (B) Data are from Desert Locust Bulletins from September 2003 to March 2012 (see http://www.fao.org/ag/locusts/en/archives/archive/index.html) and include a plague period (September 2003 to March 2005) followed by the current remission. Sampling occurred from Jan-09 to Mar-12.

and ephemeral (reviewed in Uvarov 1977; Latchininsky & Launois-Luong 1997; Vesey-Fitzgerald 1957; Chapman 1976). Dispersal behaviour of solitarious desert locusts is thought to be conditioned by erratic desert rains that influence the abundance of plants for food and bare moist soil suitable for egg laying (Rao 1937, 1942). Under this view, the solitarious phase of *S. gregaria* is expected to follow metapopulation dynamics in a strict sense, with climate-driven extinction of ephemeral local populations compensated by recolonization events (Slatkin 1977).

This metapopulation hypothesis is supported by a study based on a single nuclear DNA marker, which revealed some degree of genetic diversity and high genetic differentiation between solitarious populations of the desert locust (Ibrahim *et al.* 2000). The populations were distributed along the Red Sea coast, an area infested with swarms 4 years earlier. With caution due to the limited sample size, the authors disfavoured the hypothesis of recent fragmentation as the single process explaining their data, which would contribute to a more drastic loss of genetic diversity. Instead, a simulation study showed that the observed genetic pattern was compatible with propagule-pool metapopulation dynamics (Slatkin 1977), which assumed a low number of colonists coming from the nearest occupied patch and no migration between extant populations (Ibrahim 2001).

In the alternative view, *S. gregaria* occupies permanent habitats where the presence of adults has been reported every month of the year (reviewed in Latchininsky & Launois-Luong 1997; Waloff 1963; Popov 1997). In the persistent areas studied by Waloff (1963), the author observes that although the majority of solitary adults move locally, migratory flights over great distances may also occur at night, even when the local vegetation is green and lush. Under this view, the population structure during the solitarious phase of *S. gregaria* should be consistent with a stepping-stone model (migrants only come from neighbouring populations; Kimura & Weiss 1964) or an island model (random migration between local populations; Wright 1951), depending on the spatial scale of migration between persistent populations. However, if desert locust dispersal is primarily driven by prevailing winds (Rainey 1963) rather than rainfall, the pattern of dispersal that governs solitarious population connectivity probably occurs on a large scale and exhibits high levels of both asymmetry and stochasticity. The resulting genetic patterns may thus be weak and lack simple geographical trends.

In addition to local drift effects due to high climatic stochasticity of desert habitats, inflated genetic differentiation (and low genetic diversity) may follow plague events as large populations crash and become fragmented. Indeed, at the end of plagues, pesticide controls are generally intensive. For example, 26 and 13 million hectares were treated against desert locusts in the last two major invasions of 1987–1988 and 2003–2005, respectively (Lecoq 2001, 2005). Additionally, newly colonized environments may be unfavourable to sustained reproduction or may even lead to mass mortality, when, for instance, winds carry desert locust swarms out of their usual ecological niche, including to southern humid regions, oceans and high mountain ranges (reviewed in Uvarov 1977; Rosenberg & Burt 1999).

In this study, we used population genetics tools and methods to help elucidate population dynamics in the solitarious desert locust. We genotyped 24 microsatellite loci from 23 solitarious populations of locusts sampled across the species range in Sahel-Saharan Africa. We first computed traditional summary statistics and used a Bayesian clustering method to evaluate the levels of genetic diversity in *S. gregaria* and the extent of the population structure. We then inferred whether populations were experiencing demographic disequilibrium by using several tests on summary statistics for departure from their theoretical distribution under equilibrium. Using an approximate Bayesian computation (ABC) method, we finally provided estimates of the effective population sizes and of the magnitude of the variation from plague to remission.

Materials and methods

Population sampling

We collected 23 population samples composed of 11–30 individuals. We preferentially collected adults to avoid risk of relatedness between individuals (M.P. Chapuis & M. Lecoq, pers. obs.); only population samples 5, 6, 8 and 10 included a few late-instar larvae. Because winged adults may disperse until they find suitable conditions for breeding, this sampling scheme was at the cost of including some individuals that may not reproduce locally. The sampling area covered much of the distribution of *S. gregaria* and included 21 population samples from the northwestern African region, one population sample from the Central region (sample 14) and one from the southwestern Asian region (sample 23) (Fig. 1A). These subdivisions encompass the international management areas for desert locust control, within the framework of three desert locust commissions coordinated by the Food and Agriculture Organization of the United Nations (FAO) (Lecoq 2003). These subdivisions also delineate the three main outbreak areas (areas at the origin of invasions) of the desert locust: Indo-Pakistan border, Red Sea shores and Aden Gulf, and mountains of Central Sahara and Mauritania. The minimum and maximum distances between two population samples were 9 km (between samples 18 and 19) and 8250 km (between samples 9 and 23), respectively.

The sampling reflects the temporal heterogeneity of the desert environment as insects were sampled from January 2009 to March 2012, that is, during four different rainy seasons (Fig. 1A and Table 1). The sampling took place 3–6 years after the last inter-regional desert locust plague event (Fig. 1B). Accordingly, national locust centre staffs, following protocols for locust population mon-

Table 1 Genetic diversity of 23 *Schistocerca gregaria* population samples considering the 21 microsatellite markers at selective neutrality

Population sample	Genetic unit	Country	Sampling date	Habitat	Phase	Density	*N*	R_s	H_E	*V*	*M*-ratio	*ΔH*	*Rel*
1		Algeria	03/03/2012	W	Transiens	75 000	30	9.58	0.86	28.56	0.83	−1.20	−0.01
2		Algeria	14/11/2010	W	Solitarious	0.3	30	9.64	0.86	29.81	0.85	−0.94	0.00
3	A	Algeria	02/12/2010	W	Solitarious	1.25	30	3.14***	0.52***	17.68***	0.48***	0.55*	0.52***
4		Morocco	07/05/2010	W	Solitarious	400	30	9.56	0.86	27.99	0.80	−0.75	−0.01
5		Mauritania	17/11/2010	W	Solitarious	100	30	9.47	0.85	27.85	0.81	−1.10	0.00
6	B	Mauritania	26/01/2009	P	Solitarious	250	20	7.75***	0.79***	25.81	0.65*	−0.77	0.11***
7		Mauritania	05/01/2011	W	NA	NA	27	9.36	0.86	26.75	0.88	−0.76	−0.01
8		Mauritania	13/09/2009	ID	Solitarious	250	29	9.16	0.85	22.20	0.88	−1.14	0.00
9		Mauritania	25/11/2009	P	Solitarious	250	29	9.49	0.85	29.40	0.80	−1.36	0.00
10		Mauritania	02/02/2010	P	Solitarious	200	30	9.57	0.86	29.85	0.81	−0.56	−0.02
11		Mauritania	14/03/2011	P	Solitarious	1.7	30	9.51	0.85	29.72	0.83	−1.14	0.00
12		Mauritania	22/01/2011	P	Transiens	1500	30	9.28	0.86	24.08	0.78	−1.11	−0.01
13		Mauritania	04/01/2011	NA	Transiens	700	30	9.18	0.85	25.43	0.81	−0.81	0.00
14		Sudan	01/01/2009	NA	Solitarious	NA	11	9.25	0.84	27.40	0.68	−0.65	0.02
15		Chad	06/11/2010	ID	Solitarious	1.5	26	9.37	0.85	27.48	0.75	−1.05	0.00
16		Chad	19/10/2010	P	Solitarious	500	29	9.37	0.85	26.76	0.80	−1.17	0.00
17	C	Chad	21/10/2010	P	Solitarious	550	28	8.54*	0.83*	24.83	0.81	−0.67	0.03***
18		Chad	16/11/2011	P	Solitarious	0.17	10	9.05	0.85	23.49	0.69	−0.18	0.01
19		Chad	16/11/2011	ID	Solitarious	0.07	15	9.29	0.86	28.82	0.71	−0.27	−0.01
20		Chad	28/11/2011	P	Solitarious	1.35	11	9.05	0.85	25.78	0.68	−0.33	0.00
21		Chad	15/11/2011	P	Solitarious	55	17	9.38	0.84	29.87	0.77	−0.85	0.00
22		Niger	03/11/2011	ID	Solitarious	NA	25	9.52	0.84	27.36	0.74	−1.08	0.00
23		Pakistan	01/12/2011	NA	NA	NA	30	9.29	0.85	26.49	0.80	−1.04	−0.01
Main	D	—	—				499	9.48	0.85	27.34	1.08†	−3.33	0.00

P, plain; W, wadis; ID, interdune depression; *N*, number of diploid genotypes; R_S, allelic richness computed for nine diploid genotypes; H_E, expected heterozygosity (Nei 1987); *V*, allele size variance; *M*-ratio, statistic of Garza & Williamson (2001), which equals k/r, where k is the number of alleles and r the range in allele size; *ΔH*, standardized statistic of Cornuet & Luikart (1996), which equals $(H_E - H_{EQ})/SD$, where H_{EQ} is the heterozygosity expected on the basis of the observed number of alleles and sample size in a similar population at mutation–drift equilibrium and SD the standard deviation of H_{EQ}; *Rel*, relatedness index of Queller & Goodnight (1989); NA, no data.
See Fig. 1 for details of population sample abbreviations. Densities are in number of individuals per hectare.
***Highly significant ($P \leq 0.001$) and *significant ($P \leq 0.05$).
†Note that *M*-ratio is influenced by sample size.

itoring campaigns, reported a solitarious phase and estimated the density of the populations at a level generally below the threshold inducing gregarization, that is, a change from a cryptic solitarious phase to a mass-migrating gregarious phase (i.e. 500 individuals/ha; see Duranton & Lecoq 1990; Cissé *et al.* 2013; Table 1). There were, however, three exceptions with insects collected in a transiens phase, which marks the first stages of the gregarization process (population samples 1, 12 and 13). In these cases, locusts were collected in loose and isolated aggregates during the last year of our study, in areas where local infestations had been reported (from January 2011 to March 2012; Fig. 1B).

Genotyping, equilibrium of linkage and Hardy–Weinberg

DNA was extracted from a 2-mm section of the hind femur using DNeasy tissue kit (Qiagen). Twenty-four microsatellite loci were genotyped (SgM41, SgM51, SgM66, SgM74, SgM86, SgM87, SgM88, SgM92, SgM96, diEST-2, diEST-6, diEST-8, diEST-11, diEST-12, diEST-13, diEST-16, diEST-28, diEST-29, diEST-30, diEST-35, diEST-37 and diEST40 from Blondin *et al.* (2013); and DL01 and DL06 from Yassin *et al.* (2006)). All loci were shown to be free of null alleles in six Western African population samples that had been previously analysed by Blondin *et al.* (2013; population samples 4, 8–10, 15 and 16). Microsatellite loci were genotyped, using fluorescently labelled polymerase chain reaction (PCR) primers and an ABI 3130 DNA sequencer (Applied Biosystems), as described in Blondin *et al.* (2013).

We tested for linkage disequilibrium between each pair of microsatellite loci and within each population sample using G-exact tests in GENEPOP 4.0 (Rousset 2008). Hardy–Weinberg equilibrium (HWE) was tested for each locus and within each population sample using exact tests implemented in Genepop 4.0 (Rousset 2008). Corrections for multiple tests were performed using the false discovery rate approach (Benjamin & Hochberg 1995) as implemented in R-package QVALUE (Storey & Tibshirani 2003; R Development Core Team 2012).

Detection of loci under selection

Because 13 of our microsatellite markers were derived from an expressed sequence tag (EST) library, we detected whether loci were under selection based on patterns of genetic variation between population samples. Briefly, the method detects outliers exhibiting significantly high or low F_{ST} values, controlled for within-population heterozygosities at the loci considered. We applied the method of Beaumont & Nichols (1996) which assumes an infinite island model to obtain the null F_{ST} distribution. We used ARLEQUIN 3.5.1 (Excoffier & Lischer 2010) because F_{ST} levels are scaled with the average heterozygosity within populations ($\hat{h}_0$) rather than the average heterozygosity between populations ($\hat{h}_1$) (Excoffier *et al.* 2009). Because F_{ST}-based tests of selection are biased if different samples are drawn from the same population, we pooled population samples that were genetically similar, thereby considering four populations only: population samples 3, 6, 17 and the remaining samples, which made up the fourth genetic population (Results and Fig. 1). Analyses were carried out with the one-step strict stepwise-mutation model (SMM; Ohta & Kimura 1973) assuming the presence of 100 demes with 100 000 simulations.

Genetic differentiation and clustering

The level of differentiation between population samples was quantified by computing pairwise estimators of F_{ST} (Weir 1996), and 95% confidence intervals were determined by bootstrapping 2000 replicates over all loci using FreeNA (Chapuis & Estoup 2007). Levels of genotypic differentiation between population samples were tested using Fisher's exact tests (GENEPOP 4.0; Rousset 2008) and a significant level of 0.001. We chose a more stringent level than the usual 0.05; otherwise, significance would often have been associated with confidence intervals including zero.

The clustering approach implemented in STRUCTURE version 2.3.4 (Pritchard *et al.* 2000) was used to infer the number of genetic units. We used the admixture model and the option of correlated allele frequencies between populations and, because our sampling scheme involved the collection of many individuals from discrete distant locations, we used the sampling location as prior information (Hubisz *et al.* 2009). All other settings were left at default values. Following the recommendations of Gilbert *et al.* (2012), we carried out 20 replicate runs for each value of the number of clusters. Each run consisted of 1.5×10^6 Markov chain Monte Carlo (MCMC) iterations, with a burn-in period of 5×10^5. This MCMC parameterization checked for convergence and stationary distribution in values of the model likelihood and admixture parameter.

The natural logarithm of the likelihood of the data [ln P(D)] was calculated: it is expected to be highest with a low variance for true K (Pritchard *et al.* 2000). The number of clusters (K^*) was also determined by using the ΔK metric of Evanno *et al.* (2005). Because the ΔK method was shown to detect the uppermost level of genetic structure, we followed the hierarchical approach described by Coulon *et al.* (2008). These STRUCTURE outputs were computed with STRUCTURE HARVESTER (Earl & vonHoldt 2011). We set the number of genetic clusters

from $K = 1$–8, because ΔK and lnP(D) values always decreased when approaching $K = 8$. CLUMPP software version 1.1.2 (Jakobsson & Rosenberg 2007) was then used to align the posterior estimates of cluster memberships from the 20 replicate runs with $K = K^*$ using Greedy algorithm with 10 000 random input orders, and the results were graphically displayed with DISTRUCT 1.1 (Rosenberg 2004).

Detection of demographic disequilibrium

We first estimated, over all loci, the mean values of statistics traditionally used to summarize within-population genetic variation: allelic diversity corrected for subsamples of nine diploid genotypes (R_S; El Mousadik & Petit 1996), expected heterozygosity (H_E; Nei 1987) and allele size variance (V). We used Wilcoxon signed-rank tests applied to single-locus values of the statistics of interest to determine whether the level of genetic variation within population samples (and genetic units) differed significantly between population samples (and genetic units).

In a population experiencing size reduction, rare alleles are lost quickly and then lead to transient heterozygosity excess (Cornuet & Luikart 1996). We used BOTTLENECK software 1.2.02 (Piry *et al.* 1999), which compares the observed H_E values with expected values in an isolated population at mutation–drift equilibrium with same numbers of alleles and sample sizes as observed in the data set. We performed 10 000 iterations using a generalized stepwise-mutation model (GSM; Zhivotovsky *et al.* 1997) with variance equal to 0.36 (Estoup *et al.* 2001). Significance was tested using a two-tailed Wilcoxon signed-rank test (Piry *et al.* 1999).

We also used the M-ratio = k/r, where k is the number of alleles and r the range in allele size, which is expected to be smaller in recently reduced populations than in equilibrium populations (Garza & Williamson 2001). We computed a critical value of the M-ratio (Mc) so that 95% of the simulations of an isolated population at mutation–drift equilibrium had M-ratio $>M_c$, following Garza & Williamson (2001). The only exception was the assumption of a GSM model with mean variance equal to 0.36 in order to specify a mutational model that was realistic for microsatellite loci of grasshoppers (Chapuis *et al.* 2011) and identical to BOTTLENECK analyses. To this aim, we used the simulation option of DIYABC software version 2.0 (Cornuet *et al.* 2014). Because the sample size varied across population samples, we used 10, 20 and 30 diploid individuals. We used different effective population sizes so that populations had a mean H_E equal to those observed in our samples (Table 2).

We finally computed genetic relatedness among individuals within each population sample following the method of Queller & Goodnight (1989) using GROUPRELATE (Valsecchi *et al.* 2002) and reference allele frequencies computed from the entire data set. Randomizations were conducted by replacing original genotypes with alleles drawn randomly from the observed allele frequency distributions, thereby simulating a null distribution where no individual was related to any other. This index measures how much higher the probability of recent coalescence is for a pair relative to the average probability for all pairs considered (Rousset 2002). Therefore, when averaged across individuals within a population, it may reflect drift effects and it measures the increase in homozygosity in a population.

Table 2 Critical M-ratio values computed with different sample sizes and demographic models

Simulation	N	N_e/H_E	M_C
1	30	50 000 (0.85)	0.71
2	30	40 000 (0.83)	0.72
3	30	25 000 (0.79)	0.74
4	30	3000 (0.52)	0.80
5	20	50 000 (0.85)	0.66
6	20	40 000 (0.83)	0.67
7	20	25 000 (0.79)	0.70
8	20	3000 (0.52)	0.79
9	10	50 000 (0.85)	0.56
10	10	40 000 (0.83)	0.58
11	10	25 000 (0.79)	0.61
12	10	3000 (0.52)	0.77

N, number of diploid genotypes.
We simulated an isolated population at mutation–drift equilibrium and 21 microsatellite markers with the same characteristics as those in our real data set (microsatellite size ranges and mutation parameters). The demographic history had been simplified with a constant effective population (N_e). According to simulation sets, populations had mean expected heterozygosity (H_E) equal to that of the main population, population 17, population 6 or population 3 of our real data set (see Table 1). For each set of parameter values, 10 000 data sets were simulated.

ABC estimation of effective population sizes

We investigated how the levels of genetic variation observed in *S. gregaria* translated in terms of the effective population size and variation between remission and plague periods. We used a simple demographic model in which a single population experienced instantaneous change in effective population size. The effective population size varied from remission level (N_r) to plague level (N_p) following the plague history reported by Magor *et al.* (2008) from 1910. Assuming three generations per year (see Roffey & Magor 2003), we simulated the population size variation of approximately 300 generations. We also set constant and at remission level

(N_r) the effective population size of former years. In populations recurrently fluctuating in size, the long-term effective size, which determines the overall amount of genetic drift, is expected to correspond approximately to the harmonic mean size over time and should thus be closer to the size during remission than during plague (Motro & Thomson 1982). We defined the intensity of the population size decline α as the ratio of the plague population size (N_p) to the remission population size (N_r). Higher α values correspond to more severe reductions in population size between plague and remission periods. This ratio between the effective population sizes during plague and remission periods is not affected by mutation rate estimates and should provide accurate indication of changes in population size.

We used the ABC method (Beaumont *et al.* 2002) implemented in DIYABC version 2.0.1 program (Cornuet *et al.* 2014) and statistics summarizing genetic variation, which included the mean number of alleles (A), mean expected heterozygosity (H_E), M-ratio (M) and variance of allele size in base pairs (V). To compute the observed values of these summary statistics, we pooled population samples that were genetically similar (i.e. genetic group D; Results section) and excluded the three peculiar population samples 3, 6 and 17 (Results section). We produced a reference table containing five million simulated data sets. We estimated the posterior probability density for each parameter using 10% of the simulated data sets closest to the observed data set (50 000). Mutations in the repeat regions of each locus followed a symmetric GSM model without allele size constraints (as the maximum number of continuous allelic states observed in our data set was 136 for locus SgM41). Prior values for any mutation model settings were drawn independently for genomic microsatellites and microsatellites of transcriptome origin, but in similar probability distributions. Prior values for mean mutation rates were then drawn from a uniform distribution bounded between [10^{-5}, 10^{-3}] to allow for large microsatellite mutation rate uncertainty. We also considered mutations that inserted a single nucleotide into or deleted it from the microsatellite sequence, which we actually suspected in most loci (see allele frequency distributions in Fig. S2 in Supporting Information) and used a uniform distribution bounded between [10^{-8}, 10^{-5}]. We used default values for all other mutation model settings (Table S1, Supporting Information, description of prior probability distributions).

To define prior probability distribution of both effective population sizes N_p and N_r, we used a uniform distribution bounded between [10^4, 10^6]. Because the value range covers several orders of magnitude, we also performed inferences with log-uniform prior distributions. The inferences obtained by drawing effective population sizes from uniform probability distributions showed lower bias and dispersion measures (Table S2, Supporting Information). Furthermore, setting a prior distribution of one type or another resulted in posterior estimations that were relatively consistent. Hereafter, we present results from the uniform distribution only (details on results from the log-uniform distribution are presented in Table S3 and Fig. S1 in Supporting Information).

Results

Tests for conformity to basic assumptions

Independence of microsatellite loci was confirmed for each pair of loci and within each population sample. The only exception was population sample 17 with 12 significant tests of 276 pairs of loci (4%). We found 30 significant deviations from Hardy–Weinberg equilibrium for all 552 combinations of microsatellite loci and population samples (5%). The majority of these values were not associated with any locus or population sample. However, we detected two deficiencies and three excesses of heterozygotes of 24 tests for population sample 17.

Tests of selection detected diEST12, diEST35 and diEST37 as outlier microsatellite loci at 0.05 level of significance ($P = 0.019$, $P = 0.046$ and $P = 0.017$; Fig. 2). Allele frequency distributions were similar among population samples for diEST12 and diEST35 only ($P > 0.11$; Fig. S2, Supporting Information). F_{ST} values

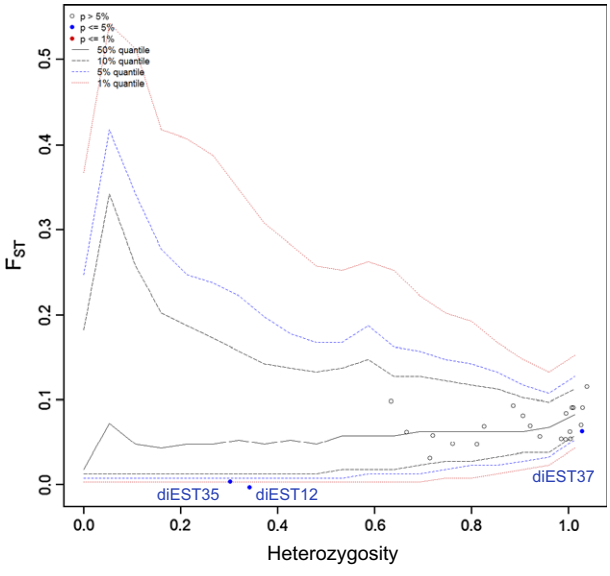

Fig. 2 F_{ST}-based detection of microsatellite loci under selection. Graphic output from ARLEQUIN 3.5.1 (Excoffier & Lischer 2010).

at these loci were not different from zero and associated with extremely low levels of heterozygosity, suggesting purifying selection. Locus diEST37 showed very few (but significant) differences between populations and the highest level of genetic diversity within populations, suggesting balancing selection ($P = 0.017$; Fig. 2). Levels of population differentiation and diversity were approximately the same whether or not the three suspicious markers were included. Hereafter, we showed results with the 21 remaining loci (Tables S4 and S5, Supporting Information, present results with all 24 microsatellite markers).

Population structure

Nonsignificant differentiation was found among many pairwise comparisons (Fig. 3 and Table S6, Supporting Information). Significant genotypic differentiation was only observed in comparisons involving population samples 3, 6 and 17. The levels of differentiation of these three population samples with other population samples were high to low with F_{ST} value ranges of [0.168–0.223], [0.016–0.030] and [0.004–0.027], respectively. The global level of population differentiation was low with an F_{ST} equal to 0.020 (95% confidence interval = [0.018–0.023]).

STRUCTURE analyses (Pritchard *et al.* 2000) provided consistent results for the 20 runs tested for each K value.

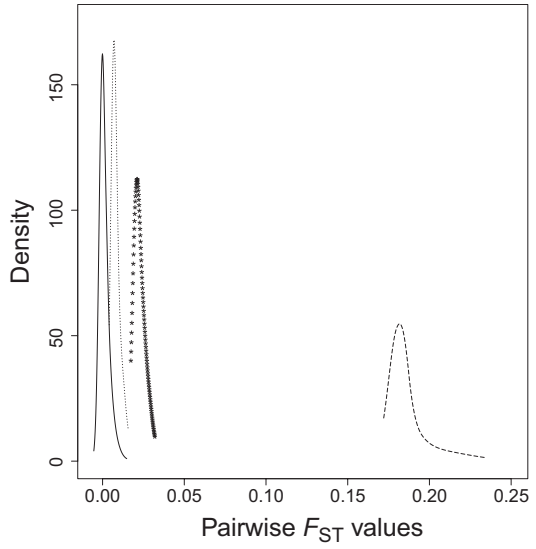

Fig. 3 Density estimation of pairwise F_{ST} values computed between 23 *Schistocerca gregaria* population samples genotyped at 21 microsatellite loci at selective neutrality. Dashed, starred and dotted lines represent pairs including samples 3, 6 and 17, respectively. The solid line represents other pairs of population samples. We used the locfit function (Loader 1996) implemented in version 3.0.0 of the R-package (Ihaka & Gentleman 1996; http://cran.r-project.org).

The natural logarithm of the likelihood of the data increased from $K = 1$ to $K = 2$ and then from $K = 3$ to $K = 4$ for which it was maximal and presented low variance (Fig. 4A). Postprocess analyses for $K = 4$ (Fig. 4B) enabled us to assign with high probability (≥0.86) all individuals from sample 3 to a genetic unit A. Population samples 6 and 17 partly corresponded to two additional genetic units, clusters B and C, respectively, and all the remaining samples corresponded to cluster D. Individuals from sample 6 were assigned in a 50:50 proportion to cluster B or to main cluster D with high confidence level (≥0.82 and 0.86, respectively), although four individuals were of mixed ancestry. Twenty-five per cent of individuals from sample 17 had a high proportion of ancestry from cluster C (≥0.99), whereas the remaining individuals of the sample belonged with high confidence to main cluster D (≥0.88). Metric ΔK peaked at $K = 2$ showing that the highest hierarchical level of the genetic structure included both units A and D. A second hierarchical level of division was identified with $K = 3$, which revealed genetic units B and C. Finally, K^* and the assignment patterns estimated with the hierarchical use of metric ΔK were similar to those inferred with the highest lnP(D) method (Fig. S4, Supporting Information).

The four genetic clusters largely corresponded to the four groups of population samples resolved by F_{ST} analyses. Therefore, the set of all population samples, except samples 3, 6 and 17, was considered as a single homogeneous genetic pool (i.e. genetic unit D) and hereafter referred to as the main population. Accordingly, after excluding samples 3, 6 and 17, global population differentiation (0.0006) was no longer significant at 0.001 level ($P = 0.004$) and associated with a 95% confidence interval including zero ([−0.0005 to 0.0016]). Global differentiation between the four genetic clusters was 0.072 ($P < 0.001$).

Founder effects and effective population sizes

Population samples 3, 6 and 17 (interchangeable with genetic units A, B and C) harboured lower levels of genetic diversity than samples from the main genetic unit D (Table 1). Population sample 3 in particular only had one-third of the number of alleles and 63% of the expected heterozygosity found in the main population. Population sample 3 also exhibited a highly significant positive value of ΔH. Values of the M-ratio were lower than expected under mutation–drift equilibrium for population samples 3 and 6 (Tables 1 and 2). The average relatedness across individuals differed significantly from zero within population samples 3, 6 and 17 (Table 1). Figure S3 (Supporting Information) shows that this pattern was not due to related pairs of individ-

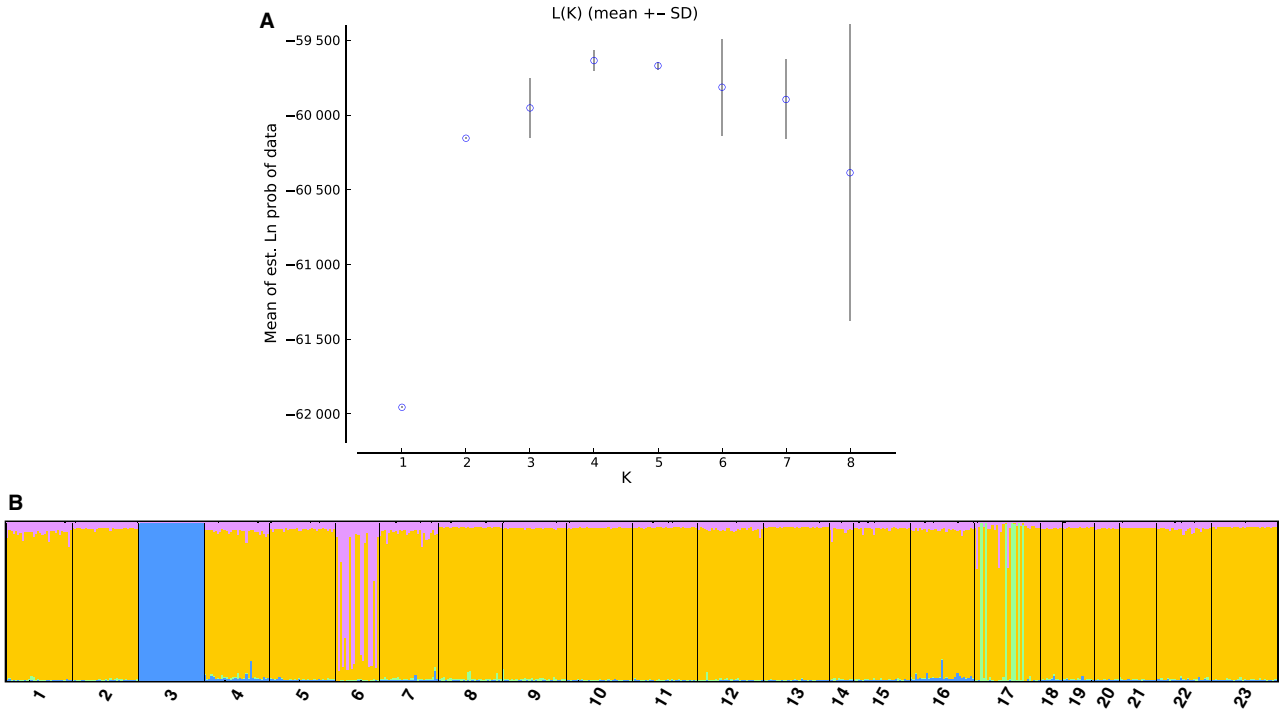

Fig. 4 Structure membership coefficients for 579 *Schistocerca gregaria* individuals assigned to four genetic clusters. (A) The plot was generated using the value of *K* associated with the maximum natural logarithm of the likelihood of the data [lnP(D)] (i.e. $K^* = 4$). Results were similar using the maximum of the metric ΔK (Fig. S4, Supporting Information). (B) Each individual included in the analysis is represented by a vertical bar partitioned into four coloured segments corresponding to the admixture proportions in four genetic clusters. Each colour represents a different genetic cluster: cluster A is in blue, cluster B in pink, cluster C in green and cluster D in orange. Black lines separate individuals from different population samples, and population samples are labelled at the bottom of the plot.

uals such as siblings, in particular in genetic unit 6 that included a few late-instar larvae. Each sample from the main genetic unit did not show evidence of demographic disequilibrium (Table 1).

In the main population, however, the ABC framework estimated a decrease in the effective population size following the end of plagues (i.e. $\alpha \geq 1$; Fig. 5). The effective population sizes during remission decreased moderately though, with a 18-fold median estimate and large uncertainty [CI: 1–60]). The ABC analysis estimated precisely the mutation-scaled remission population sizes ($N_r \times \mu$) (Fig. 6C,D). Table 3 shows that the current size of the main population during remission may be around 30 000 effective individuals. However, there was less information on mutation-scaled plague population sizes ($N_p \times \mu$), with asymmetrical distributions and large confidence intervals (Table 3 and Fig. 6A,B). Table S7 (Supporting Information) shows that similar posterior distribution samples of the mutation-scaled effective population sizes were obtained when including the 3000 closest simulated data sets and when using log or log-tangent transformation of parameters as proposed in Estoup *et al*. (2004), and Hamilton *et al*. (2005).

Table 3 Mean, median and 2.5 and 97.5% quantile estimates from DIYABC posterior distribution samples of demographic parameters

	Mean	Median	$Q_{0.025}$	$Q_{0.975}$
Original parameters				
N_p	553 000	577 000	40 000	981 000
N_r	36 000	27 000	12 000	116 000
Composite parameters				
$N_r{*}\mu g$	370	337	21	868
$N_r{*}\mu t$	120	92	5	394
$N_r{*}\mu g$	18	17	8	35
$N_r{*}\mu t$	5	4	2	12

Population sizes are given in effective number of diploid individuals.

Discussion

Evaluation of population dynamics hypotheses

In plague locusts, huge effective population sizes and long-distance dispersal of swarms probably homogenize genetic variation over large areas (Chapuis *et al*. 2009,

2011). During the short remission periods, solitarious populations are believed to differentiate through substantial genetic drift and restricted gene flow (Ibrahim *et al.* 2000; Ibrahim 2001). Various factors can reduce effective sizes of solitarious populations, for example rarity and heterogeneity of rainfall in the Sahel-Saharan environment (Vesey-Fitzgerald 1957; Latchininsky & Launois-Luong 1997), or intense control efforts and severe population crashes at the end of plagues. In our study, F_{ST} indices and Bayesian clustering showed genetic homogeneity among most of the solitarious populations sampled across a large geographical scale, from Mauritania to Pakistan, which encompassed all three desert locust commissions coordinated by FAO (Figs 3 and 4).

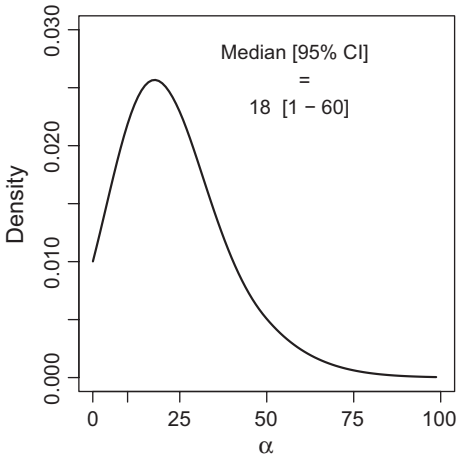

Fig. 5 Posterior probability densities of population size decline from plague to remission (α).

Furthermore, the levels of genetic diversity were large in the newly solitarized populations. Indeed, 20 of our population samples (that belonged to the same main genetic unit) showed substantial genetic diversity, with a mean expected heterozygosity of 0.85 (Table 1). This observed level of genetic diversity was higher than average compared to demographically stable populations of Orthopteran insects (Chapuis *et al.* 2012). This result also agrees with reports on other locust species that are facing different ecological challenges, for example the migratory locust, *Locusta migratoria*, a cosmopolitan species that inhabits a wide range of habitats in tropical and temperate parts of the Eastern Hemisphere. Chapuis *et al.* (2009) reported a mean expected heterozygosity of 0.83 in numerous Malagasy and Chinese solitarious populations of this migratory locust sampled 1–6 years after the last recorded outbreak.

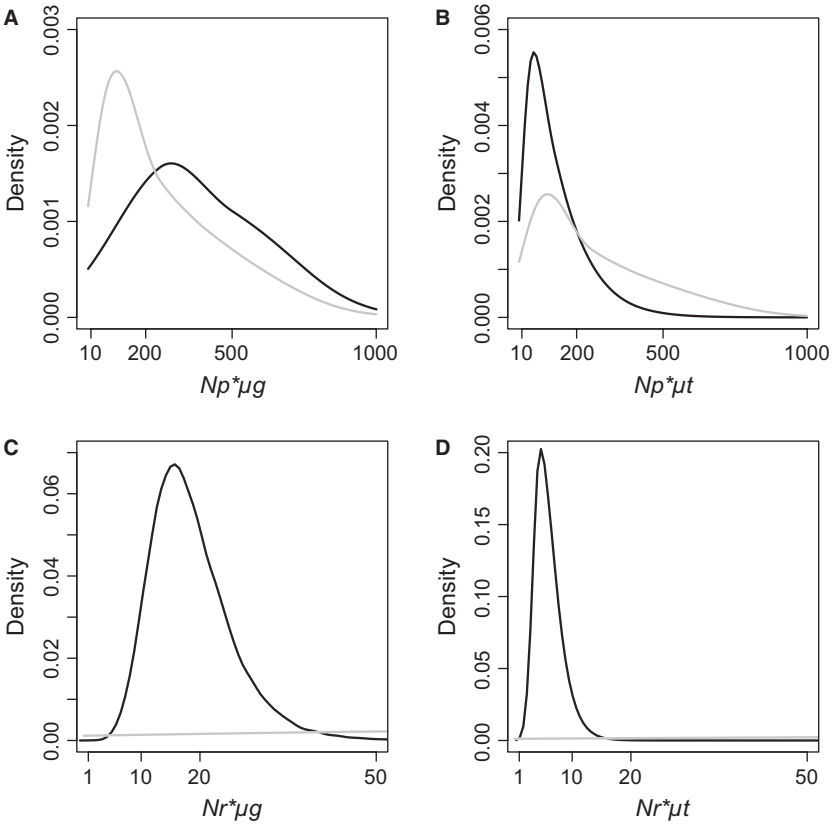

Fig. 6 Posterior probability densities of mutation-scaled effective population sizes during plague (A, B) and remission (C, D). Posterior distributions considered mutation rates either from genomic-derived microsatellite markers (μ*g*; A and C) or from transcriptomic-derived microsatellite markers (μ*t*; B and D). Prior distributions for effective population size parameters were uniform with parameters (2×10^4–2×10^6) (Table S1, Supporting Information). Point values of posterior probability distributions are presented in Table 3. Grey lines show prior distributions.

Our results showed that changes in effective population size from plague to remission were moderate and thus had little impact on *S. gregaria* genetic variation within and between populations in the short time of remission. We used several standard tests but did not detect a departure from a model of demographic equilibrium for the main population. This suggests that the action of genetic drift is slow and the effective population size remains large. Accordingly, the DIYABC program estimated a total remission effective size of at least tens of thousands and possibly hundreds of thousands of individuals (Table 3). Although solitarious populations remain large, decreases in effective size from plague to remission were inferred from the ABC analysis. However, the analysis conveyed little information on the magnitude, as the lower limit approached the value of $\alpha = 1$, implying that the population size was constant, and a more severe upper limit reached $\alpha = 60$ (Fig. 5). The low confidence in α is probably related to the low level of information on the effective number of individuals before the end of the last plague. Indeed, the posterior probability density approximately follows the same distribution as the prior probability density (Fig. 6 and Table 3). Furthermore, additional computer simulations showed that such a decrease in effective population size was not enough to differentiate separate populations genetically in the short duration of remission (~16 generations in our case), even in the absence of dispersal, due to shared ancestral polymorphism (Table 4A). A hundred generations would be necessary to produce significant genetic differentiation among many isolated populations of tens of thousands of individuals, whereas the longest remission reported in the last century was 14 years (i.e. 42 generations; Magor *et al.* 2008). Conversely, when considering the 16 generations since the last plague, a decline in effective population size to thousands of individuals or more (i.e. $\alpha \geq 100$) would be necessary for populations to differentiate in the case of no gene flow (Table 4B). These results also highlight that any rate of dispersal could explain the absence of genetic differentiation observed in *S. gregaria* solitarious populations. In the ABC analysis, neglecting migration should not have resulted to consequent overestimation of effective population sizes. It has indeed been shown that migration significantly contributes to high effective population sizes only if large numbers of genetically well-differentiated sources contribute to the migrant pool (e.g. island models) (Wright 1951; Whitlock & Barton 1997). However, such models of migration are unlikely at the 8250-km scale of the present study and most of (source or recipient) populations share same allele frequencies in the solitarious desert locust.

These results showed that high mortality associated with plague decline did not lead to severe bottleneck events and did not prevent the successful establishment of solitarious populations. In addition, these results did not support the assumption that solitarious *S. gregaria* exist as small populations threatened by extinction during remission. Because the arid climate of Sahel-Saharan deserts does not have a drastic effect on the genetic diversity and structuring of the desert locust, our results support a demographic model of persistent populations in the solitarious phase of the species. The persistence of local populations may be explained by habitats that are persistent in time (reviewed in Latchininsky & Launois-Luong 1997; Waloff 1963; Popov 1997). These habitats are presumably confined to areas near reliefs, where run-off is coupled with rainfall and which may offer more persistent favourable ecological conditions (Popov 1997). They may then approximately correspond to the 250 000 km^2 of outbreak areas, where the first occurrences of the locust transition from the

Table 4 Length of remission (A) and effective population sizes (B) needed to differentiate from the end of a plague event

(A)			
dr	F_{ST}	*P* $\alpha = 0.001$	*P* $\alpha = 0.05$
16	0.0003 ± 0.0002	16	48
42	0.0007 ± 0.0002	72	95
100	0.0016 ± 0.0002	100	100
500	0.0071 ± 0.0003	100	100
(B)			
N_r (α)	F_{ST}	*P* $\alpha = 0.001$	*P* $\alpha = 0.05$
30 000 (16)	0.0003 ± 0.0002	16	48
10 000 (50)	0.0008 ± 0.0002	84	99
5000 (100)	0.0016 ± 0.0003	100	100
1000 (500)	0.0079 ± 0.0004	100	100

We simulated 20 isolated populations and 21 microsatellite markers with the same characteristics as those in our real data set (sampling sizes and microsatellite size ranges). We considered ABC median estimates to set mutation parameters. Demographic history was simplified to a single plague event of 2 years (with $N_p = 500\ 000$), *dr* generations ago and with a current remission population size N_r. *P* is the power (in %) of the global exact G-test of heterogeneity of allelic frequencies, as computed with GENEPOP 4.0 (Rousset 2008). One hundred data sets were simulated for each set of parameter values. The first model corresponds to the actual situation: the time between *Schistocerca gregaria* last plague and the study sampling, revealing 16 generations; the intensity of effective population size decline α is very close to our median estimate. For this set of parameter values, populations had an average of nine alleles for 18 diploid genotypes and a mean expected heterozygosity equal to 0.85. (A) $N_r = 30\ 000$. Note that the second model corresponds to the longest remission recorded in the last century according to Magor *et al.* (2008). (B) $dr = 16$.

solitarious to the gregarious phase are generally observed at the onset of events leading to an outbreak and an invasion (Fig. 2 in Sword *et al.* 2010).

In more ephemeral habitats, reproductive dormancy, which delays sexual maturation up to several months, may be a key life history trait that enables populations to survive in their solitarious form when conditions are unsuitable for successful breeding. In desert locust adults, sexual maturation is delayed by various environmental factors (Uvarov 1966), mainly drought (in Uvarov 1977) and vegetation senescence (Ellis *et al.* 1965), but also low temperature and long photoperiods (Norris 1957). In some insect species, variation in life cycle duration decreases the risk of extinction caused by unpredictable catastrophic events (Menu *et al.* 2000). Immature adults of the solitary phase can also fly at night moving downwind in search of green vegetation. When temperature is above 20 °C, migrating solitarious insects can cover a few tens of kilometres per night with an average speed of 4 m/s and an average duration of 2 h (Roffey & Magor 2003). Night-flying adults are thought to settle selectively in the green areas they encounter, then mature rapidly, but the range covered is unknown (Roffey & Magor 2003). The frequency of the migratory flights is also unknown, although it might be high according to reports on the migratory locust (Davey 1955, 1959; Farrow 1975; Lecoq 1975), and the red locust *Nomadacris septemfasciata* (Lecoq *et al.* 2011).

Several reasons may explain the contradictory results obtained by Ibrahim *et al.* (2000). The authors used the 3′-end fragment of a gene of the *antennapedia*-class homeobox family (GenBank no. X73982). *Antennapedia*-class homeobox genes are clustered in the genome and involved in the specification of the developmental patterning of the body segments along the anterior–posterior axis of the embryo (Zhang & Nei 1996). Dawes *et al.* (1994) identified this gene in *S. gregaria* as a homologue of *Drosophila* gene *ftz*, member of a family whose sequences are less constrained than those of other homeotic genes and evolve rapidly. Thus, it cannot be ruled out that the inferred genetic structure was specific to the single studied sequence, which may be located in a region of the genome under natural selection. Alternatively, if the genetic structure observed by Ibrahim *et al.* (2000) among Eritrean populations is representative of the neutral part of the genome, discrepancies between the two studies may be due either to the temporal instability of the genetic pattern or to the nonexhaustive sampling of the species range.

Bottlenecked populations

In our study, 10% of the population samples were impacted by genetic drift. We found three genetically distinct populations with reduced genetic diversity, which ranged from 0.52 to 0.83. However, it is possible that more discrete populations exist considering the large spatial scale of the remission area. Moreover, we could not infer a common trend behind the location of these three samples, as they were neither located at the periphery or core of the range (Algeria, Mauritania, Chad), nor sampled at a particular date (beginning of 2009 and end of 2010) or type of habitat (wadis and plains) (Table 1). This absence of geographical trend reinforces the likelihood that population dynamics can be driven by unpredictable catastrophic events.

Founder effects could explain the lower genetic diversity in these population samples. BOTTLENECK software (Cornuet & Luikart 1996) detected a significant excess of heterozygosity within population sample 3, which had the most severe reduction in genetic diversity. In all the other population samples, we observed negative ΔH values, which thus suggested demographic expansion (e.g. populations not yet in migration–drift equilibrium following the last plague) or deviation from the isolation hypothesis (i.e. very low migration rates between differentiated populations; Pope *et al.* 2000). This would make it difficult to identify the molecular signatures of population size reduction in samples 6 and 17, and at least bring strong support to the bottleneck signal in sample 3. The *M*-ratio detected a signal of population size reduction in samples 3 and 6. This test has more power to detect population decline than BOTTLENECK, particularly more ancient events (Girod *et al.* 2011), whereas we expected very recent reductions in our desert locust populations (i.e. ≤time of the last plague event ~16 generations ago). Application of the index of Queller & Goodnight (1989) corroborated the assumption of a lower population size for all three populations, with a significant increase in homozygosity from a few per cent up to 50%.

Climate-driven egg and hatchling mortality provides a mechanism for severe stochastic drift effects. Once eggs are fully developed, desert locust females can retain them for about only 3 days if no suitable laying sites are found (Roffey & Magor 2003). Moreover, once eggs are laid in bare moist soil, they need to absorb their own weight in water to complete their development, usually in the first days after laying, although some cases of egg dormancy have been reported (Roffey & Magor 2003). Hatchling survival will highly depend on the local abundance of host plants for food (Roffey & Magor 2003). At these stages, the species is thus sensitive to stochastic climate events, which cause drying out, for example, of temporary wadis in plains or of lakes formed after intensive rainfall in interdune depressions.

Nevertheless, the observed pattern of genetic variation was more complex for two populations. Structure

analyses showed a mixture of individuals of different genetic origin within population samples 6 and 17. The majority of their genotypes belonged to the main genetic group but the genetic composition of some of their genotypes was unique (Fig. 4). This suggested that new individuals had recently arrived from higher-quality permanent populations. In fewer instances, genotypes were of mixed ancestry, which resulted from successful interbreeding in previous generations. Positive and negative F_{IS} values and increased linkage disequilibrium observed in population sample 17 may also have occurred from the admixture of genetically distinct individuals (Wilson & Goldstein 2000). These results suggested that large numbers of colonists from the main genetic group compensated for rare and local genetic drift events. Long-range movements and/or proximity of suitable habitats would thus be necessary for separate solitarious populations to mix successfully. From an ecological point of view, this is plausible because desert locusts are strong fliers and seasonal dry-out events can easily concentrate suitable habitats.

Acknowledgements

This work was funded by CIRAD. We are grateful to Sidi Ould Ely, Antoine Foucart, Mourad Nouriddine, Akodom Bang-Neseth, Emmimerta Emmanuel, Thomas Masdengarti, Yacoub Habab, Mani Tanko, Nadjana Darry Kara, Ousmane Awada, Azam Khan, Mohamed Lazar and more generally the National Units of Locust Control and FAO Regional Commissions (CLCPRO, CRC, SWAC) for providing us with samples of desert locusts. Jean-Marc Bouvet provided access to the laboratory facilities of Research Unit 39 at CIRAD. Data used in this work were also partly produced through the technical facilities of the Centre Méditerranéen Environnement Biodiversité, Montpellier. We thank Pierre-Emmanuel Gay for assistance with maps, Antoine Foucart and Arnaud Estoup for discussions that helped improve the manuscript and Karine Berthier for constructive and insightful comments on an earlier version of the manuscript.

References

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London, Series B*, **263**, 1619–1626.

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

Benjamin Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, **57**, 289–300.

Blondin L, Badisco L, Pagès C *et al.* (2013) Characterization and comparison of microsatellite markers derived from genomic and expressed libraries for the desert locust. *Journal of Applied Entomology*, **137**, 673–683.

Chapman RF (1976) A biology of locusts. *The Institute of Studies in Biology*, vol. 71. Edward Arnold, London.

Chapuis M-P, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular and Biological Evolution*, **24**, 621–631.

Chapuis M-P, Loiseau A, Michalakis Y *et al.* (2009) Outbreaks, gene flow and effective population size in the migratory locust, *Locusta migratoria*: a regional scale comparative survey. *Molecular Ecology*, **18**, 792–800.

Chapuis M-P, Popple J-A, Berthier K *et al.* (2011) Challenges to assessing connectivity between massive populations of the Australian Plague locust. *Proceedings of the Royal Society of London, Series B*, **278**, 3152–3160.

Chapuis M-P, Streiff R, Sword GA (2012) Long microsatellites and unusually high levels of genetic diversity in the Orthoptera. *Insect Molecular Biology*, **21**, 181–186.

Cissé S, Ghaout S, Mazih A *et al.* (2013) Effect of vegetation on density thresholds of adult desert locust gregarization from survey data in Mauritania. *Entomologia Experimentalis et Applicata*, **149**, 159–165.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.

Cornuet J-M, Pudlo P, Veyssier J *et al.* (2014) DIYABC v2.0: a software to make Approximate Bayesian Computation inferences about population history using Single Nucleotide Polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, doi: 10.1093/bioinformatics/btt763.

Coulon A, Fitzpatrick JW, Bowman R *et al.* (2008) Congruent population structure inferred from dispersal behaviour and intensive genetic surveys of the threatened Florida scrub-jay (*Aphelocoma coerulescens*). *Molecular Ecology*, **17**, 1685–1701.

Davey JT (1955) A preliminary note on seasonal movements of the African Migratory Locust in the solitary phase. *Locusta*, **3**, 1–14.

Davey JT (1959) The African Migratory Locust (*Locusta migratoria migratorioides* Rch. & Frm., Orth.) in the Central Niger Delta. Part two. The ecology of *Locusta* in the semi-arid lands and seasonal movements of populations. *Locusta*, **7**, 1–180.

Dawes R, Dawson I, Falciani F *et al.* (1994) Dax, a locust Hox gene related to fushi-tarazu but showing no pair-rule expression. *Development*, **120**, 1561–1572.

Duranton JF, Lecoq M (1990) *Le criquet pèlerin au sahel, Collection Acridologie Opérationnelle*, no. 6. CILSS/DFPV, Niamey, Niger.

Earl DA, vonHoldt BM (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among population of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco]. *Theoretical and Applied Genetics*, **92**, 832–839.

Ellis PE, Carlisle DB, Osborne DJ (1965) Desert locusts: sexual maturation delayed by feeding on senescent vegetation. *Science*, **149**, 546–547.

Estoup A, Beaumont M, Sennedot F *et al.* (2004) Genetic analysis of complex demographic scenarios: spatially expanding populations of the cane toad. *Bufo marinus*. *Evolution*, **58**, 2021–2036.

Estoup A, Wilson IJ, Sullivan C *et al.* (2001) Inferring population history from microsatellite and enzyme data in serially introduced cane toads, *Bufo marinus*. *Genetics*, **159**, 1671–1687.

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

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. *Heredity*, **103**, 285–298.

Farrow RA (1975) The African Migratory Locust in its main outbreak area of the middle Niger: quantitative studies of solitary populations in relation to environmental factors. *Locusta*, **11**, 1–198.

Garza JC, Williamson E (2001) Detection of reduction in population size using data from microsatellite DNA. *Molecular Ecology*, **10**, 305–318.

Gilbert KJ, Andrew RL, Bock DG *et al.* (2012) Recommendations for utilizing and reporting population genetic analyses: the reproducibility of genetic clustering using the program structure. *Molecular Ecology*, **21**, 4925–4930.

Girod C, Vitalis R, Leblois R *et al.* (2011) Inferring population decline and expansion from microsatellite data: a simulation-based evaluation of the MSVAR method. *Genetics*, **188**, 165–179.

Hamilton G, Stoneking M, Excoffier L (2005) Molecular analysis reveals tighter social regulation of immigration in patrilocal populations than in matrilocal populations. *Proceedings of the National Academy of Sciences USA*, **102**, 7476–7480.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

Ibrahim KM (2001) Plague dynamics and population genetics of the desert locust: can turnover during recession maintain population genetic structure? *Molecular Ecology*, **10**, 581–591.

Ibrahim KM, Sourrouille P, Hewitt GM (2000) Are recession populations of the desert locust (*Schistocerca gregaria*) remnants of past swarms? *Molecular Ecology*, **9**, 783–791.

Ihaka R, Gentleman R (1996) R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, **5**, 299–314.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program with label switching and multimodality in analysis of population structure. *Bioinformatics*, **3**, 1801–1806.

Kimura M, Weiss GH (1964) The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, **49**, 561–576.

Latchininsky AV, Launois-Luong MH (1997) *Le Criquet pèlerin (Schistocerca gregaria Forskäl, 1775) dans la partie nord orientale de son aire d'invasion*. *Les Acridiens*, vol. 29. CIRAD, Montpellier/VIZR Saint Petersbourg, 192 p.

Lecoq M (1975) *Les déplacements par vol du criquet migrateur malgache en phase solitaire: leur importance sur la dynamique des populations et la grégarisation*. Thèse de Doctorat d'Etat ès Sciences (Université Paris XI, Orsay). Ministère de la coopération, Paris.

Lecoq M (2001) Recent progress in Desert and Migratory Locust management in Africa. Are preventive actions possible? *Journal of Orthoptera Research*, **10**, 277–291.

Lecoq M (2003) Desert locust threat to agricultural development and food security and FAO/international role in its control. *Arab Journal of Plant Protection*, **21**, 188–193.

Lecoq M (2005) Desert locust management: from ecology to anthropology. *Journal of Orthoptera Research*, **14**, 179–186.

Lecoq M, Andriamaroahina TRZ, Solofonaina H *et al.* (2011) Ecology and population dynamics of solitary Red locusts in Southern Madagascar. *Journal of Orthoptera Research*, **20**, 141–158.

Loader CR (1996) Local likelihood density estimation. *Annals of Statistics*, **24**, 1602–1618.

Magor JL, Lecoq M, Hunter DM (2008) Preventive control and desert locust plagues. *Crop Protection*, **27**, 1527–1533.

Menu F, Roebuck JP, Viala M (2000) Bet-hedging diapauses strategies in stochastic environments. *American Naturalist*, **155**, 724–734.

Motro U, Thomson G (1982) On heterozygosity and the effective size of populations subject to size changes. *Evolution*, **36**, 1059–1066.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Norris MS (1957) Factors affecting the rate of sexual maturation of the Desert Locust (*Schistocerca gregaria* Forsk.) in the laboratory. *Anti-Locust Bulletin*, **28**, 25 p.

Ohta T, Kimura M (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genetical Research*, **22**, 201–204.

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, **90**, 502–503.

Pope LC, Estoup A, Moritz C (2000) Phylogeography and population structure of an ecotonal marsupial, *Bettongia tropica*, determined using mtDNA and microsatellites. *Molecular Ecology*, **9**, 2041–2053.

Popov GB (1997) *Atlas of Desert Locust Breeding Habitats*. Food and Agriculture Organization, Rome.

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

Queller DC, Goodnight KF (1989) Estimating relatedness using genetic markers. *Evolution*, **43**, 258–275.

R Development Core Team (2012) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rainey RC (1963) Meteorology and the migration of Desert Locusts. Applications of synoptic meteorology in locust control. *Anti-Locust Memoirs*, **7**, 125 p.

Rao YR (1937) A study of migration among the solitaries of the desert locust (*Schistocerca gregaria* Forsk.). Proceedings of the IV International Locust Conference, Cairo.

Rao YR (1942) Some results of the studies on the desert locust (*Schistocerca gregaria* Forsk.) in India. *Bulletin of Entomological Research*, **33**, 241–265.

Roffey J, Magor JI (2003) Desert Locust population parameters. *Desert Locust Field Research Stations, Technical Series*, **30**, 29 p. FAO, Rome, Italy.

Rosenberg NA (2004) Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.
Rosenberg J, Burt PJA (1999) Windborne displacements of Desert Locusts from Africa to the Caribbean and South America. *Aerobiologia*, **15**, 161–175.
Rousset F (2002) Inbreeding and relatedness coefficients: what do they measure? *Heredity*, **88**, 371–380.
Rousset F (2008) GenePop'007: a complete re-implementation of the GenePop software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.
Slatkin M (1977) Gene flow and genetic drift in a species subject to frequent local extinctions. *Theoretical Population Biology*, **12**, 253–262.
Storey JD, Tibshirani R (2003) Statistical significance for genome-wide studies. *Proceedings of the National Academy of Sciences, USA*, **100**, 9440–9445.
Sword GA, Lecoq M, Simpson SJ (2010) Phase polyphenism and preventative locust management. *Journal of Insect Physiology*, **56**, 949–957.
Uvarov BP (1966) *Grasshoppers and Locusts*, vol. 1. Cambridge University Press, Cambridge, UK.
Uvarov BP (1977) *Grasshoppers and Locusts*, vol. 2. Centre for Overseas Pest Research, London, UK.
Valsecchi E, Hale P, Corkeron P *et al*. (2002) Social structure in migrating humpback whales *(Megaptera novaeangliae)*. *Molecular Ecology*, **11**, 507–518.
Vesey-Fitzgerald DF (1957) The vegetation of central and eastern Arabia. *Journal of Ecology*, **45**, 779–798.
Waloff Z (1963) Field studies on solitary and *transiens* desert locusts in the Red Sea area. *Anti-Locust Bulletin*, **40**, 93 p.
Waloff Z (1966) The upsurges and recessions of the desert locust plague: an historical survey. *Anti-Locust Memoir*, **8**, 111 p.
Weir BS (1996) *Genetic Data Analysis II*. Sinauer Associates, Sunderland, Massachusetts.
Whitlock MC, Barton NH (1997) The effective size of a subdivided population. *Genetics*, **146**, 427–441.
Wilson JF, Goldstein DB (2000) Consistent long-range linkage disequilibrium generated by admixture in a Bantu-Semitic hybrid population. *American Journal of Human Genetics*, **67**, 926–935.
Wright S (1951) The genetical structure of populations. *Annuals of Eugenics*, **15**, 323–354.
Yassin YA, Heist EJ, Ibrahim KM (2006) PCR primers for polymorphic microsatellite loci in the Desert locust, *Schistocerca gregaria* (Orthoptera: Acrididae). *Molecular Ecology Notes*, **6**, 784–786.
Zhang J, Nei M (1996) Evolution of antennapedia-class homeobox genes. *Genetics*, **142**, 295–303.
Zhivotovsky LA, Feldman MW, Grishechkin SA (1997) Biased mutations and microsatellite variation. *Molecular Biology and Evolution*, **14**, 926–933.

M.P.C. conceived the study, performed statistical analyses and wrote the manuscript. L.B. produced the molecular work. C.Pl. contributed to the production of data and to the analyses and provided valuable comments that improved the manuscript. C.Pa. contributed to the production of data. J.M.V. helped coordinate field population sampling and provided valuable comments that improved the manuscript. M.L. conceived the study, coordinated field population sampling and contributed to the discussion of the results. All authors read and approved the final version.

Data accessibility

Genotype data set: Dryad doi: 10.5061/dryad.k6qb1.

Supporting information

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

Fig. S1 Posterior probability densities for the mutation-scaled effective population sizes during plague (A, B) and remission (C, D) periods, when considering log-uniform distributions of the effective population size priors.

Fig. S2 Allele frequency distributions for the 24 microsatellites.

Fig. S3 Inbreeding coefficient distributions.

Fig. S4 Hierarchical Structure clustering analyses based on the Coulon *et al*.'s (2008) approach.

Table S1 Prior distributions for DIYABC demographic and mutational parameters.

Table S2 Bias and precision of effective population size estimation using DIYABC.

Table S3 Mean, median and 2.5 and 97.5% quantile estimates from DIYABC posterior distribution samples of demographic parameters, when considering log-uniform distributions of the effective population size priors.

Table S4 Genetic diversity for 23 *S. gregaria* population samples at all 24 microsatellite markers.

Table S5 Pairwise F_{ST} values (Weir 1996) for 23 *S. gregaria* population samples (above the diagonal) and confidence intervals based on 2000 bootstrap samples (below diagonal) for all 24 microsatellite markers.

Table S6 Pairwise F_{ST} values (Weir 1996) 23 *S. gregaria* population samples for the 21 microsatellite markers at selective neutrality (above the diagonal) and confidence intervals based on 2000 bootstrap samples (below diagonal).

Table S7 Robustness of DIYABC inferences on the intensity of the population size decline associated with the end of plagues (α) to rejection and estimation procedures.