JOURNAL OF APPLIED ENTOMOLOGY

J. Appl. Entomol.

ORIGINAL CONTRIBUTION

# Characterization and comparison of microsatellite markers derived from genomic and expressed libraries for the desert locust

L. Blondin[1], L. Badisco[2], C. Pagès[1], A. Foucart[1], A-M. Risterucci[3], C. S. Bazelet[4,5], J. Vanden Broeck[2], H. Song[6], S. Ould Ely[7] & M-P. Chapuis[1]

1 UPR Bioagresseurs analyse et maîtrise du risque, CIRAD, Montpellier, France
2 Research Group of Molecular Developmental Physiology and Signal Transduction, Department of Animal Physiology and Neurobiology, University of Leuven, Leuven, Belgium
3 UMR Amélioration génétique et adaptation des plantes méditerranéennes et tropicales, CIRAD, Montpellier, France
4 Department of Conservation Ecology and Entomology, University of Stellenbosch, Matieland, South Africa
5 Steinhardt Museum, Department of Zoology, Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel
6 Department of Biology, University of Central Florida, Orlando, FL, USA
7 Centre National de Lutte Antiacridienne, Nouakchott, Mauritania

**Keywords**
cross-amplification, expressed sequence tags, null alleles, Orthoptera, pest, *Schistocerca gregaria*

**Correspondence**
Marie-Pierre Chapuis (corresponding author), CIRAD, UPR Bioagresseurs analyse et maîtrise du risque, F-34398, Montpellier, France.
E-mail: marie-pierre.chapuis@cirad.fr

Received: November 30, 2012; accepted: March 4, 2013.

doi: 10.1111/jen.12052

## Abstract

The desert locust (*Schistocerca gregaria*) has been feared agricultural pest since early civilization, with plagues documented in ancient texts. Population genetic studies of the desert locust are needed to determine genetic variation and movement pattern for efficient control of the pest. In this study, we complemented the limited available microsatellite collection for the desert locust with 34 new polymorphic and multiplexed microsatellite loci. To this aim, we screened an expressed sequence tags library and constructed a partial genomic library enriched for dinucleotide repeats to develop high-throughput and high-quality genotyping assays. We then paid particular attention to quality control and carefully validated 26 of these novel microsatellites and six previously described loci for the absence of null alleles in Western African field populations. This large panel of high-quality microsatellite markers provides new opportunity to infer dispersal rates between populations of the desert locust and help prioritize early monitoring and control. Furthermore, high potential for cross-taxa utility of markers was observed within *Schistocerca* genus, which includes other locust pest species, with reliable amplification achieved for at least ten of loci per species. Microsatellite markers developed from transcriptome resources were largely devoid of null alleles and were conserved across species compared with those derived from traditional genomic libraries. Accordingly, the number of highly reliable microsatellite markers was greatly improved compared with that of previous studies on Orthopteran species, and this strategy might be broadly applied in other insect species prone to null alleles.

## Introduction

Locust species usually exist at low population densities in a harmless solitarious phase, but occasionally form dense bands of marching juveniles and winged adult swarms that pose threats to agriculture across very large areas (Uvarov 1966). This migratory gregarious phase differs dramatically in many traits including morphology, behaviour, physiology and ecology from the solitarious locusts (Pener and Simpson

2009). Only a few population genetics studies have been achieved to better understand (i) the evolutionary relationships of closely related populations or subspecies in these taxa with usually broad geographic range (Chapuis et al. 2008a; Ma et al. 2012) and (ii) movement behaviour and dynamics of solitarious populations, which is of critical importance to preventive locust management (Chapuis et al. 2009; Zhang et al. 2009). This is because the isolation of a large number of microsatellite markers from traditional enriched genomic libraries is difficult in Orthopteran genomes due to high levels of redundancy and multiple banding and the prevalence of null alleles (e.g. Zhang et al. 2003; Chapuis et al. 2005, 2008b). Yet, large numbers of microsatellite markers are optimal for (i) building phylogenetic trees (e.g. $\geq$ 30 in Takezaki and Nei 2008) and (ii) investigating the patterns of contemporary dispersal in species with weak genetic structure (Anderson et al. 2010), as often observed at least in threatened areas of locust species as a result of large population sizes and long distances travelled by swarms during plagues (Chapuis et al. 2009, 2011).

Screening of large collections of expressed sequence tags (EST) is a rich source for isolating hundreds of microsatellite-containing sequences, thereby allowing a more stringent selection of markers (Ellis and Burke 2007). In addition, primers designed in flanking coding sequences are more likely to be conserved across populations, and hence, occurrence of null alleles is reduced (e.g. Kim et al. 2008). Similarly, duplication events might be rarer in such selectively more constrained regions of the genome, thereby enhancing the quality of amplification patterns (e.g. Grillo et al. 2006). In addition to higher amplification quality, EST-derived microsatellites are more widely transferable between species than microsatellites derived from enriched genomic libraries (Bouck and Vision 2007).

The infamous desert locust, *Schistocerca gregaria*, is one of the most widespread locust species with a huge potential invasion area, spanning West Africa to south-west Asia. Desert locust distribution and movements of solitarious phase locusts throughout the recession area are essentially unknown. Population genetics will be a useful tool to infer dispersal rates from the distribution of genetic variation within and amongst populations. A study based on a single nuclear DNA marker revealed significant genetic differentiation between solitarious desert locust populations sampled along the Red Sea coast despite the area being infested with swarms 4 years earlier (Ibrahim et al. 2000). Evidence of such levels of spatial genetic structure is promising and suggests that molecular markers can be used to elucidate the phylogeography and population dynamics in this species. However, it cannot be ruled out that the inferred genetic structure is specific to the single studied sequence, which may be situated in a region of the genome under selection, and the development of a large set of markers is needed. Eighteen microsatellite markers have been recently identified for the desert locust (Yassin et al. 2006; Kaatz et al. 2007). Nine of them have previously been selected for the quality of their amplification patterns in laboratory strains and multiplexed for quick and high-throughput genotyping (Berthier et al. 2010). Whether these 9 microsatellites are suitable for typing field populations remains to be determined.

Here, our objective is to identify large numbers of new microsatellite loci free of null alleles for *S. gregaria* and compare efficiency of two different strategies: a traditional genomic DNA microsatellite-enriched library and the attractive alternative of in silico mining of expressed sequences recently developed for *S. gregaria* (Badisco et al. 2011). We characterized a novel set of 34 microsatellites, which we tested for occurrence of null alleles, along with the previously published microsatellites, by genotyping 84 individuals from three field populations from Mauritania. In the end, we developed 10 multiplex PCR assays, pooled in six sets for sequencing, from 32 null alleles–free microsatellite loci previously published (6) or presently identified from genomic (9) or expressed (17) libraries.

## Material and Methods

### Genomic DNA microsatellite-enriched library

A partial DNA library of 288 recombinant clones was obtained from a single locust, following the enrichment protocol from Kijas et al. (1994) and screened with $(GA)_{15}$ and $(GT)_{15}$ probes following Billotte et al. (1999). We sequenced all the 86 recombinant clones, which showed a strong hybridization signal and inserts >350 bp. Forty-three of the microsatellite-containing sequences satisfied the following criteria and were suitable for PCR amplification: (i) uniqueness (which excluded 18 of 86 sequences; see Meglecz et al. 2004), (ii) non-repetitive flanking regions (which excluded 11 of 68 remaining sequences), (iii) maximum length of 40 repeat units and an array interruption limited to two bases (which excluded 10 of 57 remaining sequences), and (iv) adequate flanking sequence size (which excluded 4 of 47 remaining sequences).

### Unigenes data mining for microsatellites

We screened for microsatellites a total 12 709 unigenes, representing ~ 10.8-Mb expressed genome of the desert locust. They correspond to 34 672 raw ESTs from cDNA library clones obtained from neural tissues of isolated- and crowded-reared larval and adult desert locust males and females (Badisco et al. 2011). A detailed description of the unigenes used for this study is given by Badisco et al. (2011). The assembled data were queried for the presence of dinucleotide microsatellites using MISA software (Thiel et al. 2003) and sequences of at least eight perfect repeats. This search yielded 331 microsatellite-containing sequences, and we subselected the 122 sequences with a single microsatellite of a maximum length of 12 repeat units. Primers were successfully designed for 90 of these remaining sequences, using Primer3 (Rozen and Skaletsky 2000) run in a batch modus by MISA (with default parameters). Fifty of these primer pairs were tested in PCR.

### Primer validation

DNA extractions of 23 individuals from three localities (Mauritania, Morocco and Sudan) were performed from a 2-mm section of hind femur following a standard CTAB-based protocol (Doyle and Doyle 1987). PCR amplifications were performed with a thermocycler TC-512 (Techne) and followed a Touchdown procedure (Don et al. 1991) with an annealing temperature decrease of −0.5°C per cycle during the first 10 cycles of the PCR (i.e. from 60°C to 55°C). PCR steps consisted of an initial denaturation of 15 min at 94°C; 35 cycles of denaturation for 30 s at 94°C, annealing for 90 s, extension at 72°C for 75 s; and a 20-min final extension step. PCR products were loaded on an ABI 3130 DNA sequencer (Applied Biosystems) for capillary electrophoresis, and profiles were analysed with ABI GENEMAPPER (version 4.1) software. Of the 93 candidates, loci were selected for further population studies based on their reproducible amplification, a low number of aspecific and/or stutter bands, no obvious deficit in heterozygous genotypes and polymorphism. Once loci selected, we developed multiplex PCR assays that may be loaded on the automated DNA sequencer in panels of two–four markers.

### Population study

From September 2009 to February 2010, we collected a total of three population samples (25–30 individuals per sample) distributed in Mauritania and distant from 190 km to 900 km (Coordinates: 17.14N 7.61W, 18.25N 16.02W, 19.32N 14.57W). Mauritanian populations are located within outbreak areas, which are limited regions where gregarization is most likely to occur. We analysed the presently developed multiplex panels and the multiplex panel of 9 microsatellites described by Berthier et al. (2010; DL01, DL06, DL07, DL09 and DL13 from Yassin et al. 2006; Sg53, Sg36, Sg56 and Sg61 from Kaatz et al. 2007). Extraction and PCR amplification were performed as described above and repeated twice. One negative and eight positive controls (samples with known genotypes) were included in each run of 96 PCRs to check for potential contamination and standardize genotyping across experiments. The level of polymorphism and allelic distribution were estimated with GENEPOP v.4 (Rousset 2008). We tested for and estimated the null allele prevalence as previously described by Chapuis et al. (2008a). We designed new multiplex PCR assays that exclude loci with a frequency of null alleles averaged over populations $\geq$ 0.05 and with at least one population at Hardy–Weinberg disequilibrium.

### Cross-species transferability

The conservation of the primer sequences and the level of polymorphism of the selected loci were examined in the South African subspecies *S. gregaria flaviventris*, in seven other species from the *Schistocerca* genus (*S. americana*, *S. lineata*, *S. literosa*, *S. melanocera*, *S. pallens*, *S. piceifrons* and *S.vaga*), in the closely related species *Halmenus robustus* and in five other grasshopper pest species of the same subfamily Cyrtacanthacridinae (*Anacridium aegyptium*, *Austracris guttulosa* and *Nomadacris septemfasciata*) and of the subfamily Oedipodinae (*Chortoicetes terminifera* and *Locusta migratoria*). Extraction and PCR amplification were performed on 3–8 individuals per species as described above and repeated twice.

## Results

### Primer validation

From the 43 genomic-derived microsatellites, we selected 16 polymorphic loci for further population studies based on their reproducible amplification (9 of 43 loci excluded), a low number of aspecific and/or stutter bands (10 of 34 remaining loci excluded), no obvious deficit in heterozygous genotypes (8 of 24 remaining loci excluded). From the 50 unigene-derived microsatellites, we selected 18 polymorphic loci for further population studies based on their

 

reproducible amplification (16 of 50 remaining loci excluded), a low number of aspecific and/or stutter bands (6 of 34 remaining loci excluded), no obvious deficit in heterozygous genotypes (6 of 28 remaining loci excluded) and allele variation (4 of 22 remaining loci excluded).

Primer sequences, accession numbers and PCR conditions are given for all 34 loci in tables 1 and 2. Only four of the 18 unigene-derived microsatellites had a known insect ortholog (diEST5, diEST30, diEST37, diEST40, respectively, in genes coding for an aldehyde-lyase, a transcriptional factor, a protein–protein linker and a membrane-trafficking protein). According to annealing temperatures, eight sets of two, three or four loci were independently co-amplified using a multilocus amplification kit (Qiagen) in a 10-$\mu$l volume containing 1× Qiagen Multiplex Master Mix (+Q), 0.05–0.3 $\mu$M of each primer and 2 $\mu$l of genomic DNA ($\leq$ 1 $\mu$g DNA).

### Genetic diversity

Nine of the 2709 tests for linkage disequilibrium between the 43 loci (i.e. 34 validated and 9 published primer pairs) were significant after false discovery rate correction (Storey and Tibshirani 2003; package QVALUE in statistical software R 2.12.0; R Development Core Team 2010). Because these pairs of loci were never significant in more than a single population, all microsatellite loci were considered unlinked.

The level of polymorphism and allelic distribution were high amongst the genomic-derived microsatellite markers (see Figure S1 for allele frequency distributions). Means across populations of the number of alleles ($N_a$) and the expected heterozygosities ($H_e$) ranged between 14 and 27, and 0.590 and 0.958, respectively (table 1). Ranges for unigene-derived microsatellite markers were similar, with $N_a$ varying from 5 and 33 and $H_e$ from 0.269 and 0.972, respectively (table 2; and see Figure S2 for allele frequency distributions). Averages over loci tended to be slightly lower in these markers, with $N_a$ = 14.2 (vs. 17.9; i.e. ~20% of decrease) and $H_e$ = 0.768 (vs. 0.877; ~12%), but differences were not significant (Mann–Whitney U-tests; P > 0.09). Figure 1 showed that lower levels of genetic diversity in expressed microsatellites were attributable to shorter microsatellites. However, two microsatellite markers from translated regions (diEST-12 and diEST-35) displayed low levels of polymorphism, with $H_e$ less than 0.4 and 4 allelic states, that might be unusual and explained by selective processes.

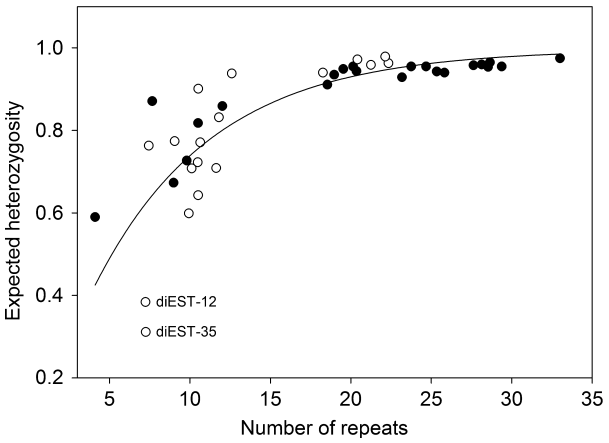


**Fig. 1** Level of expected heterozygosity as a function of number of repeats in dinucleotide microsatellites derived from genomic (filled circles) or expressed (open circles) libraries. Number of repeat units was computed for each microsatellite marker from the average allele size in the three Mauritanian populations and the number of repeat units of the cloned allele (see tables 1 and 2). The line represents a fit by an exponential curve (see Lai and Sun 2003 and Chapuis et al. 2012).

### Null alleles

Differences between observed and expected heterozygosities were observed at some of the loci (tables 1 and 2). Null alleles may have caused the observed Hardy–Weinberg disequilibrium at these loci, which all showed an excess of homozygotes evenly distributed across allele size classes (MICRO-CHECKER; van Oosterhout et al. 2004). Estimated frequencies of null alleles per locus ranged from 0 to 0.26 (tables 1 and 2). Seven genomic-derived loci (SgM23, SgM35, SgM36, SgM37, SgM40, SgM67 and SgM68) and one EST-derived locus (diEST-4) showed a mean frequency of null alleles $\geq$ 0.05. Prevalence of null alleles was verified with (i) Hardy–Weinberg tests in three other Western African natural populations collected at the end of 2010 (Coordinates: 28.52N 9.22W in Morocco, 15.19N 17.02E and 17.75N 21.56E in Chad) and (ii) presence of homozygote–homozygote mismatches in parent–offspring segregation data from a laboratory Mauritanian population. All eight loci were further discarded, for a total of 26 novel high-quality microsatellite markers.

Figure S3 shows the new multiplex PCR assays excluding these eight loci and including six of the nine markers multiplexed by Berthier et al. (2010). The three excluded loci from this set of markers showed multiple banding in a few individuals (Sgr61 and DL07) or harboured a high prevalence of null alleles in natural populations (i.e. 34%, Sgr56). These 10 PCR assays can be pooled in only

**Table 1** Sixteen microsatellite loci developed for the desert locust *S. gregaria* from a genomic DNA library. The observed size range, number of alleles ($N_a$) and observed ($H_o$) and expected ($H_e$) heterozygosities were averaged over 3 populations from Mauritania for a total of 84 individuals

| Locus | Core repeat (cloned allele) | Size (bp) | Size range (bp) | $N_a$ | $H_o$ | $H_e$ | HW test | $r_D$ | Primer ($\mu M$) | Primer sequence (5′–3′) | GenBank Accession No. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| PCR multiplex set 1 | | | | | | | | | | | |
| *SgM51* | $(TG)_{24}(GT)_6$ | 226 | 194–258 | 23 | 0.845 | 0.954 | 0/3 | 0.055 | 0.1 | F: *Ned*-AATCTGTGCTGCCGGAGA<br>R: AGGGTAGATTGCTTCCCATGT | JX101522 |
| *SgM92* | $(AC)_8$ | 174 | 170–211 | 16 | 0.777 | 0.859 | 0/3 | 0.038 | 0.3 | F: *Fam*:TTTTAATGGTTGATTGCACTT<br>R: CTCTGTAAGGCAACTGTTCA | JX101523 |
| PCR multiplex set 2 | | | | | | | | | | | |
| *SgM41* | $(TG)_{23}$ | 213 | 162–282 | 27 | 1.000 | 0.965 | 0/3 | 0.000 | 0.2 | F: *Pet*-ACAAGTGGCATTGACGGAAT<br>R: GCAGGATGAAATACGTGGGTA | JX101524 |
| *SgM23* | $(AC)_6$ | 243 | 230–242 | 4 | 0.370 | 0.590 | 1/3 | 0.122 | 0.1 | F: *Fam*-GAAGGAGCCGTAGGAGGTAGA<br>R: AGCCCTTCGTGTCTGATGTC | JX101532 |
| *SgM36* | $(TG)_8$ | 153 | 152–161 | 5 | 0.458 | 0.673 | 1/3 | 0.125 | 0.1 | F: *Ned*-CCATTTGTATCGCAACCTCA<br>R: GTGAGGGCGTCCTTAATCTG | JX101533 |
| *SgM74* | $(CA)_{10}CG\ (CA)_8$ | 245 | 210–268 | 14 | 0.702 | 0.727 | 0/3 | 0.043 | 0.1 | F: *Vic*-CGTCTATGCAGCGTGGTAGTT<br>R: CCACAAATCGCTGAAACACA | JX101525 |
| PCR multiplex set 3 | | | | | | | | | | | |
| *SgM37* | $(GT)_{22}$ | 236 | 215–268 | 23 | 0.549 | 0.955 | 3/3 | 0.263 | 0.2 | F: *Pet*-TCTTCTGAATGTGGTTGCACTT<br>R: TTCATACCGTCTAAGTGTGGATGT | JX101535 |
| *SgM66* | $(GT)_{19}GC(GT)_{10}$ | 162 | 125–188 | 20 | 0.865 | 0.929 | 0/3 | 0.023 | 0.2 | F: *Fam*-CCTAAATCAGGATGGCTGGA<br>R: TAGGCCACCTAGCACCACTT | JX101530 |
| *SgM35* | $(CA)_{28}$ | 303 | 270–320 | 24 | 0.769 | 0.955 | 2/3 | 0.089 | 0.2 | F: *Fam*-CGATGCATTACAGAGCTAGAG<br>R: GTTGTGCACCCTTAATATGG | JX101534 |
| PCR multiplex set 4 | | | | | | | | | | | |
| *SgM96* | $(AG)_{22}$ | 191 | 177–221 | 19 | 0.859 | 0.940 | 0/3 | 0.027 | 0.2 | F: *Pet*-ACTACCATTGGGCTTCGTGA<br>R: CAGCTCCAACACCACTCCAT | JX101531 |
| *SgM40* | $(TG)_{14}$ | 103 | 92–144 | 18 | 0.775 | 0.935 | 1/3 | 0.073 | 0.2 | F: *Fam*-GGTGTAGCAATTTTAATGACG<br>R: TGGAAAGCACCACAAGAC | JX101526 |
| *SgM67* | $(CA)_{12}$ | 214 | 202–273 | 24 | 0.730 | 0.955 | 2/3 | 0.106 | 0.2 | F: *Fam*-AAGGAGACAAGCCACTGGAA<br>R: ACCCAATCCACCTGTTGAAA | JX101536 |
| *SgM68* | $(AC)_{12}$ | 207 | 199–226 | 12 | 0.522 | 0.818 | 2/3 | 0.218 | 0.2 | F: *Ned*-ACCTCAGCACTCCGACATCT<br>R: TCATCCTAAGCCCAGTCACAC | JX101537 |
| PCR multiplex set 5 | | | | | | | | | | | |
| *SgM87* | $(CA)_7$ | 201 | 193–220 | 15 | 0.885 | 0.871 | 0/3 | 0.000 | 0.1 | F: *Vic*-TGCCCACAGACTTCTTGAAA<br>R: GTATTCGGCGACACATTGC | JX101528 |
| *SgM88* | $(AC)_{10}$ | 217 | 211–267 | 19 | 0.879 | 0.944 | 1/3 | 0.027 | 0.2 | F: *Ned*-GGAAACACCAAGGCAATGAG<br>R: TGCAAGAGCGTTTATCAGGA | JX101527 |
| *SgM86* | $(TG)_{23}$ | 298 | 274–338 | 24 | 0.910 | 0.958 | 0/3 | 0.025 | 0.2 | F: *Vic*-ACTGCACGGACTTTGCTTT<br>R: GACGGGACAAATGTGGAGA | JX101529 |
| *Total* | – | – | – | 17.9 | 0.743 | 0.877 | 0.75/3 | 0.077 | – | | |

HW test: number of significant probability Hardy–Weinberg exact tests (Rousset 2008) after false discovery rate correction (Storey and Tibshirani 2003; package QVALUE in statistical software R 2.12.0; R Development Core Team 2010), $r_D$: mean null allele frequency computed with the method of Dempster et al. (1977) (FREENA software, Chapuis and Estoup 2007), Primer: primer concentration for both forward and reverse primers.

six sequencing assays for high-throughput genotyping (Figure S3).

### Cross-species transferability

Twenty-four of the 26 microsatellites developed herein from *S. g. gregaria* were successfully analysed using the above PCR conditions in the Southern African subspecies *S. g. flaviventris*. The exceptions were diEST-5, which gave allelic patterns of low quality and diEST-29, which showed a low level of polymorphism.

Table 3 shows that 80% of the 26 new primer pairs produced a reliable amplicon in at least one other species in the Schistocerca genus. Ten to 16 microsatellite markers amplified per *Schistocerca* species, five of them being common for the eight species. Amplification across species within genus

 

**Table 2** Eigtheen microsatellite loci developed for the desert locust *S. gregaria* from an unigene data set. The observed size range, number of alleles ($N_a$) and observed ($H_o$) and expected ($H_e$) heterozygosities were averaged over 3 populations from Mauritania for a total of 84 individuals

| Locus | Core repeat (cloned allele) | Size (bp) | Size range (bp) | $N_a$ | $H_o$ | $H_e$ | HW test | $r_D$ | Primer ($\mu$M) | Primer sequence (5′–3′) | Unigene ID |
|---|---|---|---|---|---|---|---|---|---|---|---|
| PCR multiplex set 1 | | | | | | | | | | | |
| *diEST-29* | $(GT)_{10}$ | 238 | 230–267 | 9 | 0.705 | 0.708 | 0/3 | 0.001 | 0.1 | F: *Vic*-TGGATAGAATAATTT ATGGGTGCTT<br>R: CAGAGCTCACAATTAAGGCACA | LC01023A 2F07.f1 |
| *diEST-30* | $(GT)_{11}$ | 133 | 127–145 | 7 | 0.788 | 0.771 | 0/3 | 0.011 | 0.2 | F: *Ned*-TCTTGCGTCAACAAATCGAG<br>R: CCACTTCCTTATTGTGCGGT | LC.3018.C1. Contig3172 |
| *diEST-16* | $(CA)_{12}$ | 240 | 232–267 | 12 | 0.811 | 0.832 | 0/3 | 0.013 | 0.2 | F: *Pet*-CCCGGATTTTATCACTTCCA<br>R: CAGCTGTCCACACGTCCTTA | LC01024A2 A05.f1 |
| *diEST-4* | $(AC)_8$ | 271 | 267–318 | 20 | 0.827 | 0.940 | 1/3 | 0.095 | 0.1 | F: *Fam*-GTCCCGAGTTCGAGTCTCAC<br>R: GCTTTCTGCCACAATGTTGA | LC02007B1 G07.f1 |
| PCR multiplex set 2 | | | | | | | | | | | |
| *diEST-47* | $(TG)_8$ | 260 | 253–291 | 10 | 0.755 | 0.763 | 0/3 | 0.002 | 0.1 | F: Fam-TGTGTCCTGTGTCGAATCGT<br>R: TGTCACGAGAAATGCAGCTT | LC.2018.C1. Contig2180 |
| *diEST-20* | $(CT)_{11}$ | 130 | 110–180 | 33 | 0.962 | 0.972 | 0/3 | 0.000 | 0.1 | F: *Vic*-GTCTTGCCCATATTTCAGCC<br>R: GATCACAGAAGTGACATGCGA | LC.3905.C1. Contig4033 |
| *diEST-42* | $(TC)_{10}$ | 279 | 271–321 | 11 | 0.725 | 0.723 | 0/3 | 0.000 | 0.1 | F: *Vic*-GCTGCAGTGGAAACGGTAA<br>R: CTTCTCTTTCACAGCGTCCC | LC03017 B2F09.f1 |
| PCR multiplex set 3 | | | | | | | | | | | |
| *diEST-11* | $(AT)_3$ $AA(AT)_8$ | 177 | 171–191 | 8 | 0.630 | 0.643 | 0/3 | 0.005 | 0.2 | F: *Ned*-AAGCAACTCAGAAAGTGGCAA<br>R: AAAAGATAGACCGTCGTCGC | LC.161.C1. Contig225 |
| *diEST-5* | $(AC)_9$ $AT(AC)_2$ | 270 | 252–346 | 26 | 0.872 | 0.963 | 1/3 | 0.033 | 0.1 | F: *Ned*-GACAAGATGGATGGACGGAG<br>R: ACAGGCATGGTGTCACTGAG | LC.1901.C1. Contig2058 |
| *diEST-40* | $(TA)_9$ | 263 | 254–266 | 7 | 0.749 | 0.774 | 0/3 | 0.012 | 0.3 | F: *Pet*-TTTCTTCATGTGGGAAAGCC<br>R: AACATTCTGTTTACAAGGTTGCAG | LC.568.C1. Contig667 |
| *diEST-8* | $(AT)_{10}$ | 252 | 242–263 | 14 | 0.855 | 0.901 | 0/3 | 0.015 | 0.2 | F: *Fam*-TGCGATACTCAACAGTAGCTT<br>R: CAGCTGCAGTATGGTTGCCT | LC01005 A1A08.f1 |
| PCR multiplex set 4 | | | | | | | | | | | |
| *diEST-6* | $(AG)_8$ $(ACAG)_4$ | 170 | 167–175 | 5 | 0.641 | 0.709 | 0/3 | 0.019 | 0.1 | F: *Fam*-ATCAGGTTCTGCATAGGGCT<br>R: GTGGCTTTCCGCTAACTACG | LC.382.C1. Contig464 |
| *diEST-28* | GTAT $(GT)_{10}$ | 189 | 179–231 | 21 | 0.895 | 0.938 | 1/3 | 0.025 | 0.1 | F: *Vic*-TGGACAGTTTGTTGACGGAA<br>R: AGATTTCATCCCCTTCCTCC | LC.2033.C1. Contig2195 |
| *diEST-13* | $(AT)_8$ | 127 | 127–157 | 7 | 0.617 | 0.599 | 0/3 | 0.006 | 0.2 | F: *Pet*-TCGGGCTTTGTGTGTAAGTG<br>R: GACAAATCTCAAAATCTTTCATGC | LC01067 B1H12.f1 |
| PCR multiplex set 5 | | | | | | | | | | | |
| *diEST-37* | $(GT)_9$ | 263 | 259–334 | 33 | 0.976 | 0.979 | 0/3 | 0.000 | 0.2 | F: *Pet*-TTGTACGATAGCTGCATGGC<br>R: GCTGTAGGTCACCCAATGGT | LC.1999.C1. Contig2161 |
| *diEST-12* | $(AT)_8$ | 247 | 239–247 | 4 | 0.401 | 0.384 | 0/3 | 0.000 | 0.1 | F: *Ned*-ATACGACAGGAAGCGCAGTT<br>R: TGCTGTCTTGTTTGGGAAGA | LC01057 B1D11.f1 |
| *diEST-2* | $(AC)_{12}$ | 227 | 215–289 | 24 | 0.903 | 0.959 | 2/3 | 0.018 | 0.2 | F: *Fam*-GTTGCTGACAAGCACAATGG<br>R: TGCCAGAAAGCAGTTCTTCA | LC03005 A2H07.f1 |
| *diEST-35* | $(GT)_8$ | 105 | 102–106 | 4 | 0.257 | 0.269 | 0/3 | 0.022 | 0.1 | F: Fam-GCATAGGCCGAAAGACATTG<br>R: TCACACCACATCCTTAGAGTTCA | LC01018 B2E05.f1 |
| Total | – | – | – | 14.2 | 0.742 | 0.768 | 0.28/3 | 0.015 | – | | |

HW test: number of significant probability Hardy–Weinberg exact tests (Rousset 2008) after false discovery rate (Storey and Tibshirani 2003; package QVALUE in statistical software R 2.12.0; R Development Core Team 2010), $r_D$: mean null allele frequency computed with the method of Dempster et al. (1977) (FREENA software, Chapuis and Estoup 2007), Primer: primer concentration for both forward and reverse primers. Please see the hyperlink to the desert locust expressed sequence tags project for further details on unigenes (http://titan.biotec.uiuc.edu/locust/).

was about 5-fold higher for EST-derived microsatellite markers. Despite the limited number of study individuals per species, multiple alleles were found in 90% of successful amplifications, with an average of 12 polymorphic microsatellite markers per *S. gregaria* close relative.

**Table 3** Cross-species PCR tests for the 26 microsatellite loci developed from *S. gregaria* and free of null alleles in Western African desert locust populations

| | SgM51 | SgM92 | SgM41 | SgM66 | SgM96 | SgM88 | diEST-30 | diEST-20 | diEST-47 | diEST-11 | diEST-40 | diEST-8 | diEST-13 | diEST-28 | diEST-6 | diEST-12 | diEST-2 | diEST-16 | diEST-5 | diEST-35 | diEST-37 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Cyrtacanthacridinae* subfamily | | | | | | | | | | | | | | | | | | | | | |
| *S. americana (8)* | 220–325 | 195–216 | 180–237 | – | 167–210 | 215–253 | 123–125 | 135–180 | 251–267 | 177–181 | 258 | 240–244 | ? | 179–233 | 164–166 | 243 | 228–262 | – | – | – | – |
| *S. pallens (8)* | 190–250 | – | 187–212 | – | 166–170 | 211–249 | 127–129 | 117–162 | 250–253 | 179–191 | 256–260 | 240–242 | 147–155 | – | 162–166 | 243–247 | – | 216–220 | 306–356 | – | 259–262 |
| *S. piceifrons (8)* | – | 201–228 | 188–217 | – | 166–168 | 232–251 | 123–125 | 140–172 | 252–254 | 179–199 | – | 244–252 | 150 | – | 164–168 | 247–267 | 226–266 | 229 | – | – | – |
| *S. lineata (10)* | 184–232 | ? | 173–210 | 125–156 | ? | 211–240 | 127–133 | ? | 250–252 | 179–181 | – | 238–240 | 147–171 | – | 164–170 | 243–249 | 219–279 | 229–231 | 271–315 | 102–104 | – |
| *S. literosa (5)* | – | – | ? | – | 179–203 | – | 127–131 | – | 249–251 | 179–185 | 260–262 | 240 | – | – | 164 | 241–243 | 233–250 | – | 301–315 | 97–104 | 264–269 |
| *S. melanocera (3)* | 224–243 | – | – | – | 183–190 | – | 129–131 | ? | 251 | 171–179 | 260–262 | 240–242 | – | – | 164–168 | 243 | ? | 215 | ? | – | – |
| *S. vaga (3)* | 180–219 | 204–267 | 175–212 | 136–158 | – | 225–240 | 123 | 142–156 | 250–252 | 183–205 | 253–255 | ? | 150 | – | 162–164 | 243–247 | 212–254 | 229–230 | 293–330 | – | – |
| *Halmenus robustus (5)* | 177–262 | – | 185–212 | – | 183–217 | 215–242 | 123–127 | 140–174 | 250 | 179–199 | 257–263 | 238 | 147–151 | – | ? | 243 | – | 215 | – | – | – |
| *Austracis guttulosa (8)* | – | – | – | – | – | – | 123–125 | – | 249 | 172–180 | – | 236 | – | – | 162–166 | 242–244 | – | – | – | – | – |
| *Nomadacris septemfasciata (8)* | – | – | – | – | – | – | 120–125 | – | 249 | 177 | – | 236 | ? | – | 160–168 | 240–256 | – | – | – | – | – |
| *Anacridium aegyptium (8)* | – | – | – | – | – | 218–245 | 120 | – | 249 | 183–199 | – | 236 | – | – | | 243–251 | – | – | – | – | – |
| *Oedipodinae* subfamily | | | | | | | | | | | | | | | | | | | | | |
| *Locusta migratoria (8)* | – | – | – | – | – | – | 118 | – | – | – | – | – | – | – | 154–167 | – | – | – | – | – | – |
| *Chortoicetes terminifera (8)* | – | – | – | – | – | – | 118 | – | – | – | – | ? | – | – | 152–167 | – | – | – | – | – | – |

Numbers in brackets indicated numbers of genotyped individuals. Amplification failure is indicated by a dash. (?): specific PCR product but uncertainty on the allelic pattern. Loci that failed to amplify in all species are not indicated (SgM87, SgM86, SgM74, diEST42, diEST29).

Thirteen microsatellite markers amplified in the closely related genus *Halmenus*. For the three other Cyrtacanthacridinae species, amplification was successful for six microsatellite loci, five of them being shared and all except one being from transcribed sequences. The two same pairs of *S. gregaria* primers for EST-derived markers amplified in both Oedipodinae species tested (table 3).

## Discussion

Orthopteran genotypes are often characterized by the presence of null alleles, in spite of the optimization efforts to limit this risk (e.g. Zhang et al. 2003; Chapuis et al. 2005, 2008b). Recent studies assessing null allele effects on statistics (e.g. $F_{ST}$) and methods traditionally used to analyse microsatellite variation (e.g. phylogenetic tree; individual clustering) allow a comprehensive analysis of such genotype data sets (Chapuis and Estoup 2007; Chapuis et al. 2008a; Guillot et al. 2008). However, null allele bias remains unknown and difficult to assess in coalescent inferences of population history, because of their computationally intensive and complex algorithms (e.g. Cornuet et al. 2008; Hey 2010). We here screened a hundred of microsatellite primer pairs designed on both coding and non-coding sequences to report a final set of 32 microsatellite markers, either previously published (6) or presently identified (26), which all fulfil the criterion of absence of null alleles in six Western African field populations. Validation tests for the absence of null alleles will be required in the further genotyped populations, particularly if a high level of differentiation exists with the Western African population samples. However, these microsatellite markers will provide a promising starting point for further population genetics studies in this species.

Of the 26 novel microsatellite markers, about ten were polymorphic per congeneric species, but cross-taxa utility will depend on the level of versatility achieved. The prevalence of null alleles must also be verified on a species basis because the null allele frequency in a congeneric species has been shown to rapidly increase with increasing phylogenetic distance from a focal species (Li et al. 2003; Chapuis and Estoup 2007). However, the microsatellite markers reported herein represent a promising resource for population-level analyses of *S. gregaria*-related species and have the potential to facilitate cross-congener evolutionary studies depicting phylogenetic relationships (Mian et al. 2005) and levels of genetic variation (e.g. Ellis et al. 2006).

The development of a highly validated panel of numerous microsatellite markers for the desert locust was more easily achieved by the use of transcriptome resources. Indeed, from 24 polymorphic microsatellites with reproducible and accurate electrophoresis profiles characterized in each library, fifteen genomic-derived markers and seven expressed-derived markers, respectively, harboured null alleles in our study populations and had to be excluded (i.e. 62% vs. 29%; one-sided Fisher's exact test; $P = 0.03$). This is an important advantage for taxa in which the development of genotyping resources is particularly difficult because of high prevalence of null alleles. These include insects other than the Orthoptera (Lepidoptera, reviewed in Meglecz et al. 2004; Diptera, e.g. Lehmann et al. 1997).

The lower prevalence of null alleles (and higher cross-taxa amplification) observed at EST-derived markers may be explained by their location in conserved coding regions of the genome. Therefore and in agreement with previous studies (reviewed in Ellis and Burke 2007), this pattern was at the cost of lower levels of genetic diversity. However, a large fraction of desert locust microsatellites derived from gene sequences were found to be polymorphic, and the decrease in genetic diversity within populations was relatively small, with an average of 10% (i.e. 0.77 vs. 0.88). In addition, the lower level of genetic diversity was partly due to two microsatellite markers diEST-12 and diEST-35 presumed to be under selection (see fig. 1). When these two singular microsatellite markers are removed from the analysis, the level of genetic diversity for EST-derived microsatellites was enhanced to 0.84.

The slightly lower level of polymorphism observed at desert locust EST-derived microsatellite markers may result in a lower statistical power to detect subtle differences in genetic variation amongst populations. Therefore, the utility of expressed-derived microsatellites as population genetics markers will be verified in future studies by a comparative assessment with the set of similar numbers of genomic-derived microsatellites (Woodhead et al. 2005; Kim et al. 2008). Moreover, in our study, lower levels of genetic diversity in EST-derived microsatellites correlated with lower numbers of repeat units of the cloned allele (mean of 13 repeats) than anonymous microsatellite markers (mean of 18 repeats). This is because short microsatellites were more abundant in desert locust expressed library than genomic library (see also Ellison and Shaw 2010), and we preferentially selected short polymorphic microsatellites to avoid frequent presence of null alleles. Whether transcriptome-based

 

microsatellite markers harbour less null alleles than genome-based microsatellite markers for a same locus length will require further examination.

There is another possible concern with microsatellites derived from gene sequences. These markers, being in coding or UTR regions, are more expected to be linked to sites under selection and so to exhibit differentiation amongst populations and diversity within populations that do not reflect neutral expectation. However, microsatellite polymorphisms in non-coding regions (i.e. intronic and promoter sequences) can affect levels of gene expression and may also be selected for (Gemayel et al. 2010). Previous findings showed very similar estimates of population differentiation with traditional anonymous microsatellites (reviewed in Ellis and Burke 2007). In future studies, it will be possible to assess deviation from neutrality by statistical outlier detection on the many available microsatellite loci (Beaumont and Nichols 1996; Vitalis et al. 2001). Much attention will be paid to diEST-12 and diEST-35 that exhibit unusually low levels of polymorphism (fig. 1).

## Conclusion

The large panel of high-quality microsatellite markers provides new opportunity to infer dispersal rates from the distribution of genetic variation amongst populations of the desert locust. The microsatellite loci reported in this study will also represent an important tool for population genetic analyses at larger scale and allow to reconstruct more ancient evolutionary events (e.g. phylogeny of distant desert locust subspecies). In addition, the collections of a ten of polymorphic markers that successfully amplified in *S. gregaria* close relatives hold promises for population genetic studies in these species.

## Acknowledgements

This work was funded by CIRAD. We are grateful to Mr. Mborode Bamtoboin, Greg Sword and Wail M. Haroon for providing us with specimens of some studied species. SANParks, Cape Nature and Northern Cape Province, South Africa, issued permits for sampling of desert locust on their holdings, and their staff provided valuable administrative and field assistance. Jean-Marc Bouvet provided access to the laboratory facilities of the 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 Michel Lecoq and Jean-Michel Vassal for their support.

## References

Anderson CD, Epperson BK, Fortin M-J, Holdereger R, James PMA, Rosenberg MS, Scribner KT, Spear S, 2010. Considering spatial and temporal scale in landscape-genetic studies of gene flow. Mol. Ecol. 19, 3565–3575.

Badisco L, Huybrechts J, Simonet G, Verlinden H, Marchal E, Huybrechts R, Schoofs L, De Loof A, Vanden Broeck J, 2011. Transcriptome analysis of the Desert locust central nervous system: production and annotation of a *Schistocerca gregaria* EST database. PLoS ONE 6, e17274.

Beaumont MA, Nichols RA, 1996. Evaluating loci for use in the genetic analysis of population structure. Proc. R. Soc. Lond., B 263, 1619–1626.

Berthier K, Chapuis M-P, Simpson SJ, Ferenz HJ, Habib Kane CM, Kang L, Lange A, Ott SR, Babah Ebbe MA, Rodenburg KW, Rogers SM, Torto B, Vanden Broeck J, Van Loon JJA, Sword GA, 2010. Laboratory populations as a resource for understanding the relationship between genotypes and phenotypes: a global case study in locusts. Adv. Insect Phys. 39, 1–37.

Billotte N, Lagoda PJL, Risterucci A-M, Baurens FC, 1999. Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. Fruits 54, 277–288.

Bouck A, Vision T, 2007. The molecular ecologist's guide to expressed sequence tags. Mol. Ecol. 16, 907–924.

Chapuis M-P, Estoup A, 2007. Microsatellite null alleles and estimation of population differentiation. Mol. Biol. Evol. 24, 621–631.

Chapuis M-P, Loiseau A, Michalakis Y, Lecoq M, Estoup A, 2005. Characterization and PCR multiplexing of polymorphic microsatellite loci for the locust *Locusta migratoria*. Mol. Ecol. Notes 5, 554–557.

Chapuis M-P, Lecoq M, Michalakis Y, Loiseau A, Sword GA, Piry S, Estoup A, 2008a. Do outbreaks affect genetic population structure? A worldwide survey in Locusta migratoria, a pest plagued by microsatellite null alleles. Mol. Ecol. 17, 3640–3653.

Chapuis M-P, Popple J-A, Simpson SJ, Steinbauer M, McCulloch L, Estoup A, Martin JF, Sword GA, 2008b. Eight polymorphic microsatellite loci for the Australian plague locust, *Chortoicetes terminifera*. Mol. Ecol. Resour. 8, 1414–1416.

Chapuis M-P, Loiseau A, Michalakis Y, Lecoq M, Franc A, Estoup A, 2009. Outbreaks, gene flow and effective population size in the migratory locust, *Locusta migratoria*: a regional scale comparative survey. Mol. Ecol. 18, 792–800.

Chapuis M-P, Popple J-A, Berthier K, Simpson SJ, Deveson T, Spurgin P, Steinbauer MJ, Sword GA, 2011. Challenges to assessing connectivity between massive populations of the Australian Plague locust. Proc. R. Soc. Lond., 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 Mol. Biol. 21, 181–186.

Cornuet J-M, Santos F, Robert PC, Marin J-M, Balding DJ, Guillemaud T, Estoup A, 2008. Inferring population history with DIYABC: a user-friendly approach to Approximate Bayesian Computation. Bioinformatics 24, 2713–2719.

Dempster AP, Laird NM, Rubin DB, 1977. Maximum likelihood from incomplete data via the EM algorithm. J. Roy. Stat. Soc. B 39, 1–38.

Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS, 1991. 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res. 19, 4008.

Doyle JJ, Doyle JL, 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19, 11–15.

Ellis JR, Burke DE, 2007. EST-SSRs as a resource for population genetic analyses. Heredity 99, 125–132.

Ellis JR, Pashley CH, Burke JM, McCauley DE, 2006. High genetic diversity in a rare and endangered sunflower as compared to a common congener. Mol. Ecol. 15, 2345–2355.

Ellison CK, Shaw KL, 2010. Mining non-model genomic libraries for microsatellites: BAC versus EST libraries and the generation of allelic richness. BMC Genomics 11, 428.

Gemayel R, Vinces MD, Legendre M, Verstrepen KJ, 2010. Variable tandem repeats accelerate evolution of coding and regulatory sequences. Annu. Rev. Genet. 44, 445–477.

Grillo V, Jackson F, Gilleard J, 2006. Characterisation of *Teladorsagia circumcincta* microsatellites and their development as population genetic markers. Mol. Biochem. Parasitol. 148, 181–189.

Guillot G, Santos F, Estoup A, 2008. Analysing georeferenced population genetics data with Geneland: a new algorithm to deal with null alleles and a friendly graphical user interface. Bioinformatics, 24, 1406–1407.

Hey J, 2010. The divergence of chimpanzee species and subspecies as revealed in multipopulation isolation-with-migration analyses. Mol. Biol. Evol. 27, 921–933.

Ibrahim KM, Sourrouille P, Hewitt GM, 2000. Are recession populations of the desert locust Schistocerca gregaria remnants of past swarms? Mol. Ecol. 9, 783–791.

Kaatz K-H, Ferenz H-J, Langer B, Moritz RFA, 2007. Isolation and characterization of nine polymorphic microsatellite loci from the desert locust, *Schistocerca gregaria*. Mol. Ecol. Notes 7, 1042–1044.

Kijas JMH, Fowler JCS, Garbet CA, Thomas MR, 1994. Enrichment of microsatellites from citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. Biotechniques, 16, 657–662.

Kim KS, Ratcliffe ST, French BW, Liu L, Sappington TW, 2008. Utility of EST-derived SSRs as population genetics markers in a beetle. J. Hered., 99, 112–124.

Lai YL, Sun FZ, 2003. The relationship between microsatellite slippage mutation rate and the number of repeat units. Mol. Biol. Evol. 20, 2123–2213.

Lehmann T, Besanky NJ, Hawley WA, Fahey TG, Kamau L, Collins FH, 1997. Microgeographic structure of Anopheles gambiae in western Kenya based on mtDNA and microsatellite loci. Mol. Ecol. 6, 243–253.

Li G, Hubert S, Bucklin K, Ribes V, Hedgecock D, 2003. Characterization of 79 microsatellite DNA markers in the Pacific oyster *Crassostrea gigas*. Mol. Ecol. Notes 3, 228–232.

Ma C, Yang P, Jiang F, Chapuis MP, Shali Y, Sword GA, Kang L, 2012. Mitochondrial genomes reveal the global phylogeography and dispersal routes of the migratory locust. Mol. Ecol. 21, 4344–4358.

Meglecz E, Petenian F, Danchin E, D'Acier AC, Rasplus JY, Faure E, 2004. High similarity between flanking regions of different microsatellites detected within each of two species of Lepidoptera: Parnassius apollo and Euphydryas aurinia. Mol. Ecol. 13, 1693–1700.

Mian RM, Saha MC, Hopkins AA, Wang Z, 2005. Use of tall fescue EST-SSR markers in phylogenetic analysis of cool-season forage grasses. Genome 48, 637–647.

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P, 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol. Ecol. 4, 535–538.

Pener MP, Simpson SJ, 2009. Locust phase polyphenism: an update. Adv. Insect Phys. 36, 1–286.

R Development Core Team, 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rousset F, 2008. GENEPOP '007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol. Ecol. Res. 8, 103–106.

Rozen S, Skaletsky HJ, 2000 Primer3 on the www for general users and for biologist programmers. In: Bioinformatics methods and protocols: methods in molecular biology. Ed. by Krawetz S, Misener S, Humana Press, Totowa, NJ, 365–386.

Storey JD, Tibshirani R, 2003. Statistical significance for genome-wide studies. Proc. Natl Acad. Sci. 100, 9440–9445.

Takezaki N, Nei M, 2008. Empirical tests of the reliability of phylogenetic trees constructed with microsatellite DNA. Genetics, 178, 384–392.

Thiel T, Michalek W, Varshney RK, Graner A, 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley *Hordeum vulgare*. Theor. Appl. Genet. 106, 411–422.

 

Uvarov BP, 1966 Grasshoppers and locusts. a handbook of general acridology. Vol 1. Cambridge University Press, UK.
Vitalis R, Dawson K, Boursot P, 2001. Interpretation of variation across marker loci as evidence of selection. Genetics 158, 1811–1823.
Woodhead M, Russell J, Squirrell J, Hollingsworth PM, Mackenzie K, Gibby M, Powell W, 2005. Comparative analysis of population genetic structure in *Athyrium distentifolium Pteridophyta* using AFLPs and SSRs from anonymous and transcribed gene regions. Mol. Ecol. 14, 1681–1695.
Yassin YA, Heist EJ, Ibrahim KM, 2006. PCR primers for polymorphic microsatellite loci in the Desert locust, *Schistocerca gregaria* (Orthoptera: Acrididae). Mol. Ecol. Notes 6, 784–786.
Zhang DX, Yan LN, Ji YJ, Kang L, Hewitt GM, Huang ZS, 2003. Isolation, characterization and cross-species amplification of eight microsatellite DNA loci in the migratory locust (*Locusta migratoria*). Mol. Ecol. Notes 3, 483–486.
Zhang DX, Yan LN, Ji YJ, Hewitt GM, Huang ZS, 2009. Unexpected relationships of substructured populations in Chinese *Locusta migratoria*. BMC Evol. Biol. 9, 144.

## Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Figure S1.** Allele frequency distributions for the 16 microsatellites derived from a genomic DNA library.

**Figure S2.** Allele frequency distributions for the 18 microsatellites derived from an unigene data set.

**Figure S3.** PCR multiplex assays of the 32 dinucleotide microsatellite markers free of null alleles in Western-African desert locust populations. Loci are either previously published (6; Yassin et al. 2006; Kaatz et al. 2007) or presently identified from genomic (9) or expressed (17) libraries. Identification numbers for PCR multiplex assays are indicated at left of the figure. Braces (at right) show sequencer multiplex sets (each pooling two PCR multiplex sets). Primer concentrations are same as in tables 1 and 2 except if mentioned in brackets ($\mu$M).