Characterization of 16 novel microsatellite loci for Ephippiger diurnus (Orthoptera: Tettigoniidae) using pyrosequencing technology and cross-species amplification

Yareli Esquer Garrigos, Michael D. Greenfield, Virginie Party, Rejane Streiff

To cite this version:
Yareli Esquer Garrigos, Michael D. Greenfield, Virginie Party, Rejane Streiff. Characterization of 16 novel microsatellite loci for Ephippiger diurnus (Orthoptera: Tettigoniidae) using pyrosequencing technology and cross-species amplification. European Journal of Entomology, Czech Entomological Society, 2016, 113, pp.302-306. 10.14411/eje.2016.037. hal-01595437

HAL Id: hal-01595437
https://hal.archives-ouvertes.fr/hal-01595437
Submitted on 26 Sep 2017

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

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Characterization of 16 novel microsatellite loci for *Ephippiger diurnus* (Orthoptera: Tettigoniidae) using pyrosequencing technology and cross-species amplification

**YARELI ESQUER-GARRIGOS**1,2, MICHAEL D. GREENFIELD3, VIRGINIE PARTY2 and RÉJANE STREIFF1,2

1 INRA, UMR 1062 CBGP, Campus International de Baillarguet, 34988 Montferrier-sur-Lez, France; e-mails: esquergarrigos@gmail.com, streiff@supagro.inra.fr
2 INRA, UMR 1333 DGIMI, Université de Montpellier, Place Eugène Bataillon, 34095 Montpellier cedex 5, France
3 Institut de Recherche sur la Biologie de l’Insecte (IRBI), CNRS UMR 7261, Parc de Grandmont, Université François Rabelais de Tours, 37200 Tours, France; e-mails: michael.greenfield@univ-tours.fr, virginie.party@univ-tours.fr

**Key words.** Orthoptera, Tettigoniidae, *Ephippiger diurnus*, calling song, microsatellite polymorphism, katydid, population genetics

**Abstract.** A novel panel of 16 microsatellite markers, obtained by pyrosequencing of enriched genomic libraries, is reported for the flightless European bushcricket *Ephippiger diurnus* (Dufour) (Orthoptera: Tettigoniidae). Five multiplex and one simplex PCR protocols were optimized, and the polymorphism at the 16 loci was assessed in two natural populations from southern France. The mean allele number and (expected mean heterozygosity) were 8.94 (0.71) and 6.57 (0.70), respectively, in each population. Several loci were at Hardy-Weinberg disequilibrium (HWD), possibly due to the incidence of null alleles. The occurrence of null alleles has been previously reported for this species, and it is a common feature of microsatellite loci in Orthoptera. Cross-amplification tests demonstrated the transferability of some of these loci to other ephippigerine species. The microsatellite loci reported here substantially increase the number of available loci for this species and will afford an accurate picture of *E. diurnus* phylogeography, the genetic structure of its populations, and an improved understanding of the evolution of male song and other sexually-selected traits in this highly variable species.

**INTRODUCTION**

The European bushcricket *Ephippiger diurnus* (Orthoptera: Tettigoniidae) has attracted considerable attention among behavioral and evolutionary biologists because of its diverse calling songs (Duijm, 1990; Ritchie, 1996), large spermatophore (Barbosa et al., 2016), and a strong population genetic structure (Spooner & Ritchie, 2006). *E. diurnus* are flightless, do not migrate, and have specific habitat preferences, and previous studies showed that they are distributed in geographically isolated, genetically differentiated populations throughout their range in southern France and northeastern Spain (Party et al., 2015). These geographically separate populations generally exhibit distinctive male songs that are characterized by a specific number of syllables per call (Ritchie, 1991, 1996), and some attempts have been made to relate the song trait to phylogeography by evaluating mitochondrial DNA (*COI*) divergence (Party et al., 2015). The various populations can be cross in the laboratory (Ritchie, 2000), but the full potential of such hybridization is unknown. To determine the phylogeography of *E. diurnus* with greater precision and to explore the evolution of song diversity, genetic markers that afford reliable, fine-level resolution of population differences are needed.

The development of a minimal number of valuable microsatellite loci for population genetics (i.e. polymorphic, easily scorable, free of null alleles and conforming to Hardy-Weinberg expectations) has proven challenging in *E. diurnus*. Indeed, a set of 16 loci had been reported for this species before our study (Hockham et al., 1999; Hamill et al., 2006). According to the authors themselves, these loci displayed strong heterozygote deficit, and the incidence of null alleles was considerable for some of them. In a preliminary trial we tested 13 of these available 16 loci on samples from highly divergent populations previously characterized for mitochondrial DNA *COI* variation (Party et al., 2015). Most loci failed to amplify and/or presented complex allelic patterns impeding their scoring. This situation significantly reduced the number of available markers to only five, which is a minimum value for population genetic analyses. We therefore applied high-throughput (pyrosequencing) technology to a partial genomic library enriched in microsatellite motifs in order to increase the number of loci and filter out those of low quality according to criteria detailed below.

**MATERIAL AND METHODS**

**Sample collection and DNA extraction**

Fifty-one specimens of *E. diurnus* were collected from nine localities in southern France between 2011 and 2014 (Fig. 1). Hind femora were dissected and preserved in 95% ethanol for DNA
Microsatellite isolation

Five μg of DNA were obtained by pooling individual DNA extracts from eight insects sampled in eight of nine localities (Fig. 1). The DNA pool was sent to Genoscreen, Lille, France (www.genoscreen.fr) for microsatellite isolation through 454 GS-FLX Titanium pyrosequencing of enriched DNA libraries following the approach described by Malaua et al. (2011). Briefly, enriched libraries were constructed using eight microsatellite probes (TG, TC, AAC, AAG, AGG, ACG, ACAT, ACTC), and the resulting library was sequenced on a GsFLX PTP. The resulting 72,447 reads were analyzed using the program QDD (Meglécz et al., 2010) and sorted according to the following criteria: number of microsatellite repeats ≥ 5, microsatellite motif not interrupted by any other bases or sequences, fragment size ≥ 80 bp. A fasta file with 5,027 reads containing a microsatellite repeat and a list of optimized primer pairs for 503 reads (size range: 90–319 bp) was provided by Genoscreen. Within these reads, 323 primer pairs with expected fragment sizes ≥ 120 bp were chosen. Special attention was paid to homologous sequences shared among distinct reads: short, repeated sequences in the vicinity of microsatellites are frequently shared among distinct loci and impede consistent single locus PCR amplifications if primers overlap them (Meglécz et al., 2007). To avoid this problem, sequences homologous among different reads identified after an “all-against-all” BLASTn analysis (http://blast.ncbi.nlm.nih.gov/Blast.cgi) were masked before primer design. In the same line, all sequences were checked and masked for the presence of known annotated repeated elements in the flanking regions by the RepeatMasker software (http://www.repeatmasker.org/) and the related domestic silkworm database. Finally, 100 loci were chosen to proceed with the first PCR screening using unlabeled primers (Eurofins Genomics, Ebersberg, Germany). For some of these loci new primers were designed when the expected amplification size was not adequate for the posterior multiplexing PCR procedure and or when primers provided by Genoscreen were located in zones of high homology according to our BLASTn analysis (Table S1).

Because of the high level of divergence among mtDNA COI clades (see Fig. 4 in Party et al., 2015) we tested PCR amplification of the 100 loci in four specimens of E. diurnus collected in Mireval, Sode, Port de Lers, and Col de Mantet (Fig. 1), and belonging to the two main COI clades (Fig. 1 in Party et al., 2015). Sequences and primers for the 100 loci are given in Table S1. Thus, we could retain only those loci amplifying unambiguously in all clades (Table S1). All amplifications were achieved with an ABI GeneAmp PCR System 2700 thermal cycler. PCR reactions were carried out in a 10-μl solution containing the following: 2 mM MgCl2, 2 mM dNTPs, 1 × PCR Buffer, 0.5 unit of GoTag G2 polymerase (Promega, Charbonnieres, France), 2 μM of each forward and reverse primer and ~10 ng of template DNA. PCR cycling conditions were as follows: 95°C 3 min, followed by 30 cycles at 94°C 30 s, 60 s at 60°C, 72°C for 45 s, and a final extension step of 10 min at 72°C. PCR products were resolved in a 2% agarose gel. Twenty-one primers not amplifying in all four specimens (i.e. partial PCR amplification) were not considered. Forty-two primer pairs that showed clear, reproducible and unique fragments in the four specimens were retained for further analysis. Among them, eight loci showed an incidence of smear or amplified nonspecific bands, thereby justifying a +3°C increase of the annealing temperature (Table S1). Lastly, 37 loci did not amplify at 60°C and were tested at 52°C using the same PCR conditions as above but were not tested in the following steps described below (Table S1).

Fragment analysis of 42 loci followed the cost-effective M13 fluorescent protocol described by Schuelke (2000) with modifications described below. Each forward primer was tagged at its 5’ end with one 18–19 bp tail described in Culley et al. (2013) and one fluorescent label depending on the expected amplification size to allow posterior PCR multiplexing (Table S1). The combination of tails and fluorescent labels was as follows: M13 modA-NED, M13 modB-PET, T7 term-VIC and M13 (-21)-FAM (Applied Biosystems, Warrington, UK, see Table 1). Simplex PCR tests were performed on four to eight specimens to confirm PCR amplification with tailed primers. The 6.25 μl PCR reaction contained: 3.25 μl Multiplex PCR Master mix (Qiagen), 1 μl of a primer mix per locus containing: 2 μM of each reverse and labeled tail primer and 0.5 μM of the forward tailed primer (ratio: 1:1:1/4, see Culley et al., 2013), 1 μl H2O and 1 μl of DNA (~10 ng/μl). PCR cycling conditions followed a denaturing step of 15 min at 95°C, then 30 cycles at 94°C for 30 s, 60°C or 63°C (Table S1) for 45 s and then 8 cycles of 94°C for 30 s, 53°C for 45 s and 72°C for 45 s and a final elongation step of 10 min at 72°C. PCR products were visualized on a 2% agarose gel. Fragment analysis was conducted on a 3730 xl DNA Analyzer (Applied Biosystems) using the GeneScan 500 LIZ as internal size standard (Applied Biosystems) and 1 to 2 μl of PCR product (1:20 dilution). After visual screening of electropherogram profiles in GeneMapper version 5.0 (Applied Biosystems), 25 primers were selected for their scoring in additional specimens. Finally, 16 primer pairs showing polymorphism and unambiguous profiles were retained. For loci showing noisy electropherograms, the annealing temperature was increased by 3°C (Table S1).

Multiplex PCR amplification

We used Multiplex Manager version 1.0 (Holleley & Geerts, 2009) to determine the best combination of loci in a multiplexed PCR amplification protocol. Five multiplex PCR amplification reactions and one simplex PCR were defined for the final set of 16 loci, and these were amplified in two populations of E. diurnus (Vias and Peyriac de Mer; Table 2) belonging to each of the two main mtDNA COI clades described in Party et al. (2015, see
Table 1. Characteristics of 16 microsatellite loci from *Ephippiger diurnus*. Abbreviations: F – the forward primer sequences; R – the reverse primer sequences; Ta – annealing temperature.

| Locus       | Repeat | Primer sequence (5’–3’) | Expected size (bp) | Dye   | Tail * | Ta (°C) | PCR reaction | Primer mix (µM) |
|-------------|--------|-------------------------|--------------------|-------|--------|---------|--------------|----------------|
| Ediur_3     | (GA)_9 | F: TTG CAA TGA AAC GTT CTT CCT R: TGA CCA GCA TTG TGT CCT GG | 216         | FAM   | M13 (-21) | 60      | Multiplex 1 | F tailed + tail = 0.7 |
| Ediur_4     | (ATAC)_8 | F: TCA GAG GGG TTT CCT TCC CT R: ACG GAT ACC AGC ACG AT | 235         | VIC   | T7 term | 63       | Simplex     | F tailed + tail = 0.7 |
| Ediur_6     | (TGT)_5 | F: GAC ACT CTT CGG GAG CTT CA R: TGA TCC TCG ATT AGT CGG AA | 205         | VIC   | T7 term | 60       | Multiplex 1 | F tailed + tail = 0.7 |
| Ediur_8     | (TGA)_8 | F: ACT GAG TGC GTA CAC AGC AT R: CGA ATC GAG CGG ATC TTA TT | 175         | VIC   | T7 term | 60       | Multiplex 2 | F tailed + tail = 0.7 |
| Ediur_20    | (ACA)_5 | F: CAC AAT CAT CTT GGT GTC GAA R: GGC GCC TAA TCG ACT GTA | 236         | FAM   | M13 (-21) | 60       | Multiplex 2 | F tailed + tail = 0.7 |
| Ediur_21    | (TTG)_8 | F: AGG AAT GAG AAA ACT GCC GA R: CAG AGG CAG CTG CAA ACA TA | 227         | PET   | M13 modB | 60       | Multiplex 4 | F tailed + tail = 1 |
| Ediur_22    | (AC)_12 | F: CCC CTC AAA TAT CCC AAC AC R: CGC AAT GTC GAA CAC ATT TT | 189         | FAM   | M13 (-21) | 63       | Multiplex 5 | F tailed + tail = 0.7 |
| Ediur_25    | (AAG)_8 | F: AAA GCT ATT GGG TTT GTG GAA R: CCA GCA ACT GTA CAG TGT CCT | 214         | PET   | M13 modB | 60       | Multiplex 1 | F tailed + tail = 1 |
| Ediur_47    | (TGA)_10 | F: TGG GAA ACA TGG AAA GGT GT R: CCT GTC ATT TGC TGC TTC CT | 149         | VIC   | T7 term | 63       | Multiplex 5 | F tailed + tail = 0.7 |
| Ediur_48    | (AC)_6  | F: AAC CCA CCG GCC TAT TAA CT R: GGG AAA CTC CAG TCA TCC AT | 227         | VIC   | T7 term | 60       | Multiplex 4 | F tailed + tail = 0.7 |
| Ediur_55    | (ATAG)_10 | F: GCA CCG CAG CCA TAG ATA AG R: ATG TGA GTT AGG AGG GAA GC | 135         | PET   | M13 modB | 60       | Multiplex 2 | F tailed + tail = 1 |
| Ediur_57    | (CA)_14 | F: TGA ACA AAG CAT AAG GCC AGC A R: ATT TCT GTC GCG TGA TGT GA | 123         | VIC   | T7 term | 60       | Multiplex 3 | F tailed + tail = 0.7 |
| Ediur_59    | (TGAT)_5 | F: TTC GGC CAT ATT AAG GGA AT G R: TGA CAT GGA GTG ATG GAT GG | 237         | PET   | M13 modB | 60       | Multiplex 3 | F tailed + tail = 1 |
| Ediur_67    | (AC)_4  | F: TGC GTG AGT GCC AGG CAG R: CTA CAG GCA CAT CAG TTA GA | 160         | NED   | M13 modA | 60       | Multiplex 4 | F tailed + tail = 1 |
| Ediur_85    | (ATG)_6 | F: TGG CCG TCT AAG GCC CAT AC R: CAA TTA ACC TAA TGA CAG TCA | 240         | FAM   | M13 (-21) | 60       | Multiplex 4 | F tailed + tail = 0.7 |
| Ediur_86    | (TTG)_9 | F: AAT GCA TCC AAC GCA CTA CC R: TGC TAA CTT ATT CCG GTG GC | 297         | VIC   | T7 term | 60       | Multiplex 2 | F tailed + tail = 0.7 |

* Tail primer sequences from Culley et al. (2013): M13 (-21) TGT AAA ACG ACG GCC AGT, M13 modA TAG GAG TGC AGC AAG CAT, M13 modB CAC TGC TTA GAG CGA TGG T, T7 term CTA GTT ATT GTT CAG CGG T. Modified F and R primers are highlighted (see Table S1).

Table 2. Characteristics of 16 microsatellite loci from two populations of *Ephippiger diurnus*. Abbreviations: N – number of genotyped specimens; Na – number of alleles; Ho – observed heterozygosity; He – expected heterozygosity.

| Locus       | Size range (bp) | Na | Ho | He | Null alleles frequency | Size range (bp) | Na | Ho | He | Null alleles frequency |
|-------------|-----------------|----|----|----|------------------------|-----------------|----|----|----|------------------------|
| Ediur_3     | 226–232         | 4  | 0.04 | 0.59** | 0.34 | 228–238 | 4  | 0.50 | 0.61 |
| Ediur_4     | 229–284         | 14 | 0.50 | 0.91** | 0.21 | 225–280 | 11 | 0.58 | 0.83** | 0.14 |
| Ediur_6     | 220–247         | 6  | 0.22 | 0.61** | 0.24 | 226–238 | 4  | 0.76 | 0.65 |
| Ediur_8     | 143–261         | 20 | 0.61 | 0.95** | 0.17 | 161–241 | 13 | 0.62 | 0.86** | 0.12 |
| Ediur_20    | 247–271         | 7  | 0.35 | 0.82** | 0.26 | 234–259 | 8  | 0.25 | 0.82** | 0.31 |
| Ediur_21    | 205–247         | 8  | 0.65 | 0.82 | 0.09 | 219–244 | 5  | 0.45 | 0.45 |
| Ediur_22    | 186–207         | 10 | 0.61 | 0.77 | 0.09 | 186–195 | 5  | 0.70 | 0.64** |
| Ediur_25    | 207–237         | 6  | 0.68 | 0.71 | 0.09 | 210–240 | 6  | 0.81 | 0.79 |
| Ediur_47    | 160–190         | 11 | 0.54 | 0.85** | 0.16 | 160–200 | 7  | 0.38 | 0.56 | 0.12 |
| Ediur_48    | 236–245         | 7  | 0.58 | 0.73 | 0.11 | 235–243 | 5  | 0.67 | 0.68 |
| Ediur_55    | 120–260         | 18 | 0.65 | 0.85 | 0.11 | 121–204 | 6  | 0.37 | 0.72** | 0.20 |
| Ediur_57    | 137–142         | 4  | 0.30 | 0.33 | 0.11 | 137–145 | 3  | 0.43 | 0.41 |
| Ediur_59    | 244–305         | 9  | 0.12 | 0.87** | 0.40 | 244–305 | 7  | 0.41 | 0.79** | 0.21 |
| Ediur_75    | 157–182         | 9  | 0.27 | 0.78** | 0.28 | 150–174 | 9  | 0.76 | 0.82 |
| Ediur_85    | 245–291         | 9  | 0.91 | 0.82 | 0.09 | 242–272 | 7  | 0.86 | 0.80 |
| Ediur_86    | 301             | 1  | 0   | 0   | 0.10 | 301–339 | 5  | 0.30 | 0.76** | 0.26 |
| Mean        | 8.94            | 0.44 | 0.71 | 6.57 | 0.55 | 0.70 |
| SE          | 1.25            | 0.06 | 0.06 | 0.67 | 0.05 | 0.03 |

** Significant deviations from Hardy-Weinberg equilibrium after Bonferroni correction (P = 0.003).
Table 3. Cross-amplification of 16 microsatellite loci isolated for E. diurnus. The INRA collection ID code for each voucher specimen is given within parentheses. Abbreviations: Lat – Latitude; Long – Longitude.

| Taxon (INRA collection ID code) | Sampling locality | Locus |
|---------------------------------|-------------------|-------|
|                                 | Ediur_3 Ediur_4 Ediur_6 Ediur_8 Ediur_20 Ediur_21 Ediur_22 Ediur_25 |       |
| E. provicensis (JSTR02143_0101) |                   |       |
|                                 | – + – + – + – + + |       |
| E. provicensis (JSTR02144_0101) |                   |       |
|                                 | – + – + – + – + + |       |
| E. terrestris (JSTR02148_0101)  |                   |       |
|                                 | – + – + – + – + + |       |
| E. terrestris (JSTR02149_0101)  |                   |       |
|                                 | – + – + – + – + + |       |
| U. rugosicollis (JSTR02145_0101)|                   |       |
|                                 | – + – + – + – + + |       |
| U. rugosicollis (JSTR02146_0101)|                   |       |
|                                 | – + – + – + – + + |       |

acknowledgements. We thank J.-C. Streito and A. Foucart for collecting specimens for the cross-amplification tests. F. Cerqueira, E. Desmarais and T. Cantinelli (Centre Méditerranéen Environnement Biodiversité; Labex CEMEB) provided technical support with genotyping. The DGIMI (Diversité, génome & interactions microorganismes-insectes, UMR INRA 1333 – Université de Montpellier) research unit made its laboratory facilities available for our analyses. This study was funded by grant ANR-11-BSV7-025-01 (EVOLCHOR) from the Agence Nationale de la Recherche de France.

REFERENCES

Barbosa F., Rebar D. & Greenfield M.D. 2016: Reproduction and immune trade-offs constrain mating signals and nuptial gift size in a bushcricket. — Behav. Ecol. 27: 109–117.

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

Chapuis M.P., Loseau A., Michalakis Y. & Estoup A. 2005: Characterization and PCR multiplexing of polymorphic microsatellite loci for the locust Locusta migratoria. — Mol. Ecol. Notes 5: 554–557.

Culley T.M., Stamper T.I., Stokes R.L., Brzyski J.R., Hardiman N.A., Klooster M.R. & Merritt B.J. 2013: An efficient technique for primer development and application that integrates fluorescent labeling and multiplex PCR. — Appl. Plant Sci. 1: 1300027, 10 pp.
