Challenges of microsatellite development in Lepidoptera: *Euphydryas aurinia* (Nymphalidae) as a case study

Melthide SINAMA[1],*, Vincent DUBUT[1],*, Caroline COSTEDOAT[1], André GILLES[1], Marius JUNKER[2], Thibaut MALAUSA[3], Jean-François MARTIN[4], Gabriel NÈVE[1], Nicolas PECH[1], Thomas SCHMITT[2], Marie ZIMMERMANN[5] and Emese MEGLÉCZ[1]

[1]Aix-Marseille Université, CNRS, IRD, UMR 6116 – IMEP, Equipe Evolution Génome Environnement, Case 36, 3 place Victor Hugo, 13331 Marseille Cedex 3, France; e-mails: melthide.sinama@gmail.com, vincent.dubut@univ-provence.fr, caroline.costedoat@univ-provence.fr, andre.gilles@univ-provence.fr, gabriel.neve@univ-provence.fr, nicolas.pech@univ-provence.fr, emese.meglecz@univ-provence.fr

[2]Trier University, Department of Biogeography, 54286 Trier, Germany; e-mails: marius.junker@gmx.de, thsh@uni-trier.de

[3]INRA, UMR 1301 – IBSV (INRA / UNSA / CNRS), Equipe Biologie des Populations en Interaction, 400 route des Chappes, BP 167, 06903 Sophia-Antipolis Cedex, France; e-mail: thibaut.malausa@sophia.inra.fr

[4]Montpellier SupAgro, UMR CBGP (INRA / IRD / CIRAD / Montpellier SupAgro), Campus International de Baillarguet, CS 30016, 34988 Montferrier-sur-Lez Cedex, France; e-mail: martinjf@supagro.inra.fr

[5]Université de Tours, CNRS, UMR 6035 – IRBI, Avenue Monge, Parc Grandmont, 37200 Tours, France; e-mail: marie.zimmermann@univ-tours.fr

Key words. Lepidoptera, microsatellite, null allele, 454 sequencing, QDD, *Euphydryas aurinia*

Abstract. Currently it remains difficult to obtain robust microsatellite markers for Lepidoptera. In an attempt to overcome the problems associated with developing microsatellite markers for this insect order we combined (i) biotin-enrichment protocol, (ii) next generation pyrosequencing (through 454 GS-FLX Titanium technology) and (iii) the use of individuals collected from eight geographically distant European populations representing three subspecies of *Euphydryas aurinia*. Out of 96 stringently designed primer pairs, 12 polymorphic microsatellite loci amplified without obvious evidence of null alleles in eight individuals from different subspecies. Between five and seven of these loci showed full within population applicability and three revealed to be robust and transferable between populations and sub-species, providing a first step towards the development of a valuable and robust tool for studying conservation issues and evolution in *E. aurinia* populations. Nevertheless, as in most studies dealing with Lepidoptera microsatellites, null alleles were detected in most of the developed markers. Our results emphasize the need for further research in order to better understand the complex evolution and organization of Lepidopteran genomes.

INTRODUCTION

Microsatellites are the most widely used molecular markers in ecological genetics, due to their genetic information content, heritability and the repeatability of their results (Selkoe & Toonen, 2006). However, it has so far proved difficult to isolate microsatellite markers for most of the Lepidoptera species assayed (e.g. Antony et al., 2001; Péténian et al., 2005; Fauvelot et al., 2006; van't Hof et al., 2007; Habel et al., 2008). This likely lies in mutually non-exclusive causes including flanking region similarity among loci (Meglécz et al., 2004, 2007), microsatellite association with transposable elements (Coates et al., 2009; Tay et al., 2010) and high frequencies of null alleles (Keyghobadi et al., 1999; Harper et al., 2003; Meglécz et al., 2004; Zhang, 2004; van't Hof et al., 2005; Mikheyev et al., 2010). As a consequence of high null allele frequencies, the transferability between species of the same genus (Flanagan et al., 2002; Nowak et al., 2009) and also between populations of the same species (Jiggins et al., 2005) is limited. Nevertheless, Daly et al.

(2004) suggest that "while problems may be encountered in some species, the microsatellite approach should not be discounted as an inefficient marker strategy in Lepidoptera". In the era of high throughput sequencing technologies, when it is possible to produce tens of thousands of genomic fragments at reasonable cost and time, the question arises: Can a large amount of sequence data counterbalance the low success rate of robust microsatellite isolation in Lepidoptera? In fact, this approach appears promising, since large shotgun genomic libraries have recently proved to be an efficient, cost-effective and time saving tool for isolating microsatellites in other taxa, such as Aves (e.g. *Hymenolaimus malacorhynchos*; Abdelkrim et al., 2009), Reptilia (e.g. *Agkistrodon contortrix*; Castoe et al., 2010), Ascidiacea (e.g. *Didemnum vexillum*; Abbott et al., 2010) and Liliopsida (e.g. *Typha minima*; Csencsics et al., 2010), especially when combined with a biotin-enrichment protocol as in the case of *Zingel asper* (Actinopterygii) (Dubut et al., 2010).

As an example of a butterfly species, the marsh fritillary, *Euphydryas aurinia*, (Rottemburg, 1775; Lepido-

---

* These authors contributed equally to this work.

ptera: Nymphalidae) was selected. This species has a Palaearctic distribution, is found in a wide range of habitats and its larvae feed on at least ten different host species in southwest Europe (Mazel, 1986). Populations of *E. aurinia* vary in their wing pattern and three subspecies are recognized based on morphological characters (Higgins, 1950, 1955). In spite of its wide distribution and host plant range, this species is declining in abundance and considered as vulnerable in Europe (Van Swaay & Warren, 1999). These characteristics make this species an excellent model for the study of (i) the evolution of insect-host plant adaptation, (ii) the detection of glacial refuges using phylogeographical methods, (iii) issues in conservation biology and (iv), as in this study, to develop robust microsatellite markers for Lepidoptera by combining biotin-enrichment protocol and high throughput sequencing technology.

This paper is an attempt to develop "robust" microsatellite markers for Lepidoptera that (i) do not have null alleles, (ii) have perfect repeat motifs (i.e. not interrupted or compound microsatellites), (iii) are PCR multiplexable and (iv) are transferable between geographically distant populations and/or subspecies. Multiplexing a set of loci makes genotyping faster and cost effective, which is important if a large number of samples have to be analysed. Moreover, multiplex genotyping minimizes the possibility of erroneous compilation of genotypes from different individuals. Perfect microsatellites were chosen because they make microsatellites easily usable not only for frequency based analyses but also in a coalescent theory framework for demographic inferences.

In this context, 5 Mbp of microsatellite enriched DNA were sequenced using 454 GS-FLX Titanium technology followed by a stringent selection of sequences for primer design. For PCR amplification and genotyping tests of the markers, several *E. aurinia* samples from geographically and genetically distant and disconnected populations, which cover a wide geographical range and represent three subspecies, were used. Moreover, the three populations screened for microsatellite polymorphism are known to be at Hardy Weinberg equilibrium based on 8–10 allozyme loci (Zimmermann, 2000; Junker & Schmitt, unpubl.). This procedure is likely to discard markers that lack transferability because of null alleles.

MATERIAL AND METHODS

Individuals from eight European populations of *E. aurinia* representing the three European sub-species were used for the different steps in the microsatellite marker development (details in Table 1). The pairwise F_{st} (Weir & Hill, 2002) of the three populations used for population genetics analyses (from France, Portugal and Estonia) ranged from 0.24 to 0.49 based on allozyme markers. The 104 individuals were kept frozen and/or in 96% alcohol until used for DNA extraction.

Microsatellites were isolated from four individuals following a biotin-enrichment protocol adapted from Kijas et al. (1994) at Genoscreen (Lille, France) (Table 1). Genomic DNA was extracted from 25 mg of tissue using the DNeasy Tissue Kit (QIAGEN) followed by RNase treatment. The obtained genomic DNA (2 μg) was digested with *Rsa*I (Fermentas) for 1 hour at 37°C, following the manufacturer's recommendations.

DNA was further ligated to standard oligonucleotide adapters as described elsewhere (Dubut et al., 2009). Eight biotin-labelled oligonucleotides, $(AG)_{10}$, $(AC)_{10}$, $(AAC)_8$, $(AGG)_8$, $(ACG)_8$, $(AAG)_8$, $(ACAT)_6$ and $(ATCT)_6$, were simultaneously hybridized to the ligated DNA for 10 min after initial denaturation of the ligation products. The choice of screened motifs was based on microsatellite motif frequencies of the published genomes of 12 insect and vertebrate species (Martin et al., 2010). The enrichment step was completed using Dynabeads (Invitrogen) following a standard procedure. The resulting enriched DNA was amplified using primers corresponding to the standard adapters, with 25 cycles (20 s at 95°C, 20 s at 60°C and 1 min 30 s at 72°C) and a final elongation step of 30 min at 72°C. The PCR product was immediately purified using QIAquick PCR Purification Kit (QIAGEN). Then, two libraries were prepared. The first from one individual (collected from Le Puits de Rians) and the second from three individuals (Table 1), each of them labelled with a different MID (Multiplex IDentifier; MID1-ACGAGTGCGT, MID2-ACGCTCGACA, MID3-AGACGCACTC) and then pooled. Each purified enriched library was used in the 454 GS-FLX Titanium library preparation (Roche Applied Science) following the manufacturer's protocols at Genoscreen (Lille, France).

Emulsion PCR was carried out at a ratio of one copy per bead and subsequently disrupted using isopropanol. Beads containing amplified DNA fragments were enriched and recovered for sequencing to finally provide 50,000 to 70,000 enriched beads for each library. The recovered ssDNA beads were packed onto 1/16 region of a 70 × 75 mm Titanium PicoTiter plate and sequenced with 200 cycles. Sample preparation and analytical processing, such as base calling, were performed at Genoscreen (Lille, France) using the manufacturer's protocol for Titanium series.

The selection of sequences for primer design was done with the program QDD (Meglécz et al., 2010). Enrichment adaptors and tags were removed from sequences. Sequences longer than 80 base pairs (bp) and containing perfect microsatellites of at least five repetitions for any motif of 2–6 bp were selected for further analyses. Sequence similarities were detected through an "all against all" BLAST (Altschul et al., 1997; e-value 1E-40) in which microsatellite motifs were soft masked. Sequences for which pairwise similarity along the flanking regions was over 95% were grouped into contigs and a 2/3 majority rule consensus sequence was created from each contig. Sequences that had significant BLAST hit to other sequences with an overall similarity among the flanking region bellow 95% were discarded to avoid potential intra-genomic multicopy sequences. All of the unique sequences (with no BLAST hit to any other "454" reads) or consensus sequences were checked for the presence of short repetitions in the flanking regions. PCR primers were designed using QDD with the following stringent criteria: (i) target microsatellites have at least seven repetitions, (ii) length of PCR product between 100 and 300 bp, (iii) flanking regions did not contain a mononucleotide stretch of more than four bases and any di-hexa base pair motifs of more than two repetitions, (iv) annealing temperatures of primer pairs were optimised to fall between 57°C and 63°C, and (v) microsatellites are not compound or interrupted. Following this procedure, 96 sequences with target microsatellites were retained.

Amplifications were initially performed on four individuals (from Estonia and Portugal, see Table 1) for each of the 96 primer pairs in a total volume of 10 μL containing 0.5 μL of 1 : 10 diluted DNA extract using the QIAGEN Multiplex PCR Kit (QIAGEN). Thermocycling was performed on a Mastercycler® gradient (Eppendorf) with the following protocol: 95°C for 15 min, followed by 30 cycles (94°C for 1 min, 56°C for 1 min,

TABLE 1. The samples of *E. aurinia* used to develop the microsatellites.

Subspecies	Sites sampled			Number of individuals used for:			
	Country	Locality	Coordinates	454 sequencing[1]	non-labelled primers (n = 96) amplifiability[2]	labelled primers (n = 23) geno-typing tests[3]	population polymorphism analyses[4]
E. a. aurinia	Estonia	Salevere	58°40′N, 23°36′E	–	2	3	29
	France	Col de Castillon	43°50′N, 7°28′E	1	–	–	–
	France	La Coutronne	43°19′N, 5°41′E	–	–	–	30
	France	Puits de Rians	43°38′N, 4°49′E	1	–	–	–
	France	Saint-Savin	45°7′N, 0°29′E	2	–	–	–
	Greece	Pissoderi	41°10′N, 21°40′E	–	–	1	–
E. a. beckeri	Portugal	Vila Verde de Ficalho	37°58′N, 7°17′W	–	2	3	39
E. a. debilis	Switzerland	Pontresina	46°30′N, 9°54′E	–	–	1	–

[1] Number of individuals (n = 4) used for biotin-enrichment and 454 pyrosequencing steps; [2] Number of individuals (n = 4) used to test the PCR amplification of the 96 primer pairs on agarose gel; [3] Number of individuals (n = 8) used for testing genotype profiles of the 23 primer pairs that passed the previous tests; [4] Individuals used for screening polymorphism, testing HW and LD equilibrium at the population level for the 12 loci that passed the previous genotyping tests.

72°C for 1 min), and 60°C for 45 min. Out of the 96 primer pairs, 23 gave clean PCR products on agarose gel electrophoresis, i.e. discrete single bands or at most two when there were large difference in size between alleles. The remaining 73 primer pairs either did not amplify in some of the 4 individuals or produced multiple bands, smears, or severe ladder-like patterns. The reverse primer of Eau52 was redesigned at this step, since it gave clear but weak bands.

All the 23 markers were then tested on eight individuals collected from distant populations and representing three subspecies (*E. a. aurinia* from Estonia and Greece, *E. a. beckeri* from Portugal and *E. a. debilis* from Switzerland; see Table 1). The loci were amplified separately using forward primers labelled with the fluorescent dyes 6-FAM (Eurogentec), PET, NED or VIC (Applied Biosystems). The amplicons were visualized using an ABI 3130XL Genetic Analyzer (Applied Biosystems). Allele sizes were scored against an internal GeneScan-500 LIZ® Size Standard (Applied Biosystems) and genotypes obtained using GeneMapper® 3.7 (Applied Biosystems). Twelve markers, which amplified in all eight individuals and showed unambiguous genotype patterns, were kept and combined into two PCR multiplex kits (Table 2). The multiplex kits were tested using the amplification protocol described above on three populations (from Estonia, France and Portugal; see Table 1). These populations were selected based on three criteria: (i) they cover a wide geographical range and represent two subspecies, (ii) they are geographically and genetically distant and are disconnected and (iii) all of them display Hardy-Weinberg (HW) equilibrium based on 8–10 polymorphic allozyme loci (Zimmerman, 2000; Junker & Schmitt, unpubl.).

GENEPOP 4.0 (Rousset, 2008) was used to: (i) test HW equilibrium, (ii) estimate the heterozygosity for all loci and populations and (iii) test linkage disequilibrium (LD) among loci within populations. Micro-Checker ver. 2.2.3 (van Oosterhout et al., 2004) was used to analyse the causes of departures from HW equilibrium: real HW disequilibrium, null alleles or scoring errors (often resulting from stuttering).

RESULTS

From the two libraries, 19879 raw sequences of at least 80 bp were obtained, and 3413 were (i) unique sequences (with no BLAST hit to any other "454" reads) or consensus sequences and (ii) contained a perfect microsatellite of at least 5 repetitions. The flanking regions of all of these sequences were checked and 263 with a flanking region of at least 30 bp without homopolymers and short repetitions (i.e. 3–5 repeats) were retained. For 206 sequences primer pairs could be designed. Of them, 110 sequences with microsatellites of only 5 and 6 repetitions were discarded (maximizing chance of being polymorphic for target microsatellites).

Out the 96 sequences with a perfect microsatellite motif and at least seven repeats, 23 displayed clean PCR products and amplified all four individuals tested on agarose gels and 12 loci amplified in all eight individuals (representative of three subspecies) that were used for testing the labelled primers. The total number of alleles of these 12 loci (based on three population samples) ranged from 4 to 34 per locus. The expected heterozygosity for the different loci (based on 98 individuals from three populations) ranged from 0.10 to 0.87 (Table 2).

After applying Benjamini & Hochberg's (1995) False Discovery Rate (FDR) controlling procedure, between five and seven markers displayed HW equilibrium within each of the tested *E. aurinia* populations. Three markers (Eau59, Eau64 and Eau73) displayed HW equilibrium in all three populations and four other markers (Eau21, Eau45, Eau71 and Eau81) were at HW equilibrium in two populations. Regarding the loci displaying a significant heterozygote deficiency, Micro-Checker analyses indicated the presence of null alleles in all cases. Moreover, loci Eau32, Eau52, Eau79 and Eau81 were not amplified in more than four individuals out of the 98 analyzed, which indicates homozygotes with null alleles.

Additionally, no significant LD was detected between any of the twelve loci after applying the FDR controlling procedure.

DISCUSSION

Although over the last decade an increasing number of microsatellite markers have been developed for Lepidoptera, with over 20 species for which there are at least 10 markers, it is still difficult to develop microsatellite markers for this insect order. Currently, high numbers of

TABLE 2. Characteristics of the twelve polymorphic microsatellite loci in *E. aurinia*. Abbreviations: N_A – number of alleles; *Ho* – observed heterozygosity; *He* – expected heterozygosity; *N* – number of individuals amplified; *P* – Probability value of the HW test; * – HW disequilibrium significant at 0.05 level after FDR controlling procedure.

Locus (GenBank ID)	Primers (5'–3')	Core motif	Allele size range (N_A)	Multiplex PCR Kit	Primer conc.	France (n = 30) He/Ho	P / N	Estonia (n = 29) He/Ho	P / N	Portugal (n = 39) He/Ho	P / N
Eau21 (HQ235030)	F-VIC: ACGGATGGTTGACTGAATAATAA R: AAATTACACGAAACGACATCGAG	(TC)₁₀	93–201 (34)	1	150 nM	0.858/ 0.567**	0.000 30	0.619/ 0.621	0.504 29	0.835/ 0.692	0.086 39
Eau32 (HQ235031)	F-VIC: ACTTGCAGAGGGTCTTGGTG R: ACACTCCTCCTTTGACGTGG	(AG)₇	261–275 (13)	1	100 nM	0.755/ 0.500**	0.001 30	0.631/ 0.621	0.189 29	0.528/ 0.211**	0.001 19
Eau45 (HQ235032)	F-NED: TTGTGATCGCGATTGATACTG R: GCGTGCGACTAGAGGAAGAT	(AG)₁₂	139–166 (10)	2	100 nM	0.836/ 0.448**	0.000 29	0.458/ 0.517	0.877 29	0.435/ 0.410	0.721 39
Eau52 (HQ235033)	F-6FAM: TTGTCAGTGTCAAAGCAGGAA R: TGATTAAGTCTCTATACTGTGTTCCG	(AG)₁₂	63–79 (9)	1	100 nM	0.682/ 0.467**	0.002 30	0.639/ 0.444	0.027 27	0.000/ 0.000	NA 36
Eau59 (HQ235034)	F-6FAM: CATAATCATCAATTTGTATGTGGG R: CAACGTACATACGCAAACCAA	(TG)₇	97–108 (7)	2	50 nM	0.709/ 0.500	0.064 30	0.708/ 0.586	0.062 29	0.626/ 0.667	0.747 39
Eau64 (HQ235035)	F-PET: GATCACGGAATGAGGCAAA R: CCTGTGAGAGTAGTTGTCGAAGAA	(AC)₉	136–153 (8)	1	100 nM	0.650/ 0.600	0.476 30	0.654/ 0.690	0.622 29	0.563/ 0.487	0.526 39
Eau71 (HQ235036)	F-NED: CCCTCGACGTCAATTATGTT R: ACGGGAAAGGAATGAGTTGT	(AG)₇	105–120 (10)	1	50 nM	0.792/ 0.933	0.639 30	0.586/ 0.500	0.396 28	0.362/ 0.211*	0.011 38
Eau72 (HQ235037)	F-NED: TTGTACAGGACGCGACTTTG R: GAGTGGGGAAGGCCTACATT	(AC)₈	147–205 (18)	1	100 nM	0.873/ 0.667**	0.000 30	0.745/ 0.724	0.582 29	0.783/ 0.333**	0.000 39
Eau73 (HQ235038)	F-6FAM: TTCATAATAACGCGCAATCAA R: AATCAAATCACGGAACATCG	(TC)₈	149–162 (4)	1	100 nM	0.098/ 0.100	1.000 30	0.000/ 0.000	NA 29	0.410/ 0.410	1.000 39
Eau79 (HQ235041)	F-6FAM: CTTAATCTCACTGATGTTTGAAGTTTT R: TGGAGAAGAGACCTTCGCTC	(TATG)₇	165–194 (12)	2	150 nM	0.853/ 0.433**	0.000 30	0.729/ 0.150**	0.000 20	0.000/ 0.000	NA 39
Eau81 (HQ235039)	F-PET: CCTTCTGCAACAATAGAGCTAGA R: CTTGTGTCTTAGTAATAGGATTGCG	(TG)₈	151–176 (13)	2	50 nM	0.739/ 0.556	0.031 30	0.576/ 0.200**	0.000 26	0.759/ 0.795	0.725 35
Eau88 (HQ235040)	F-VIC: TTGTCATCGAAAGATCACAGAAG R: GCTTCGTCTTCGAAAATTGG	(CA)₇	149–159 (6)	2	100 nM	0.711/ 0.333**	0.000 30	0.101/ 0.034*	0.018 29	0.414/ 0.308	0.066 39

microsatellites (n > 20) are published only for some economically important species (e.g. *Cydia pomonella*: Franck et al., 2005; Zhou et al., 2005; *Ostrinia nubilalis*: Dopman et al., 2004; Coates et al., 2005; Dalecky et al., 2006; Kim et al., 2008; *Bombyx mori*: Prasad et al., 2005; Nagaraja et al., 2005) or for model species (e.g. *Bicyclus anynana*: Beldade et al., 2009). Furthermore, apart from the few studies where all the markers were found at Hardy-Weinberg equilibrium (MERPDC et al., 2010; Zhou et al., 2005; Endersby et al., 2005; Torres-Leguizamon et al., 2009), in almost all the other reports of novel microsatellites in Lepidoptera, heterozygote deficiencies and the presence of null alleles are reported. It appears that the major problem with microsatellite development in Lepidopera is the frequent presence of null alleles. It is also noteworthy that next generation pyrosequencing, through 454 GS-FLX Titanium technology, was recently used to develop microsatellite markers from an *E. editha* transcriptome library (Mikheyev et al., 2010). Nevertheless, the markers displayed HW disequilibrium in many of the populations screened for polymorphism and a non-negligible number of individuals could not be amplified. Most of these markers therefore are non-robust and have limited transferability. Furthermore, markers designed from transcriptome libraries might be subject to selection (by linkage) and could result in biased estimates of population differentiation.

To overcome these problems, a combination of a biotin-enrichment protocol and 454 GS-FLX Titanium technology was used to develop robust microsatellite markers for *E. aurinia*. This method produces 5–10 Mbp genomic sequences more cheaply than the few hundred sequences of 2-700 bp produced by cloning-based methods (ca. 1200 EUR for 454 GS-FLX vs. ca. 4000 EUR). Thus it is possible to stringently select sequences for primer design and still leave a high number of potential markers that can be tested later, if necessary. In addition, in order to avoid microsatellites with null alleles, individuals from three *E. aurinia* sub-species collected from eight geographically and genetically distant populations were included in the microsatellite development protocol. In this way, it was expected that the number of markers that are transferable between sub-species and/or distant populations for population analyses, would be maximized.

Although five to seven of the microsatellites developed showed applicability within each population and subspecies tested, only three of the initial 96 loci (Eau59, Eau64 and Eau73) showed evidence of transferability among *E. aurinia* sub-species and populations (i.e. HW equilibrium in all the three populations tested and therefore no null alleles), and thus appear suitable for analyses of population genetics over the entire European range of the species. Admittedly, more robust microsatellites are needed for carrying out studies on population genetics on *E. aurinia*. Nevertheless, the combination of 454 pyrosequencing of a microsatellite enriched library and a sequential testing of *E. aurinia* sub-species and geographically distant populations during the microsatellite development protocol appears to be an effective way of isolating robust microsatellites for Lepidoptera, especially when time and cost are considered. We suggest this strategy should be systematically applied in a population and conservation genetics framework in order to validate microsatellites. When reporting novel microsatellite

markers it is especially important not to test only one population for HW equilibrium, since it is uninformative regarding the transferability of markers.

A high success rate of microsatellite isolation obviously is an important factor for microsatellite users. However, the low success rate reported here emphasizes earlier hypotheses that stress the complexity of genome organisation and evolution in Lepidoptera. The high frequency of null alleles suggest that mutation rate in the flanking regions of most of the microsatellite loci is unusually high in Lepidoptera. A better understanding of the dynamics shaping the genomes of this insect order depends on future genomic studies focusing on this problem.

REFERENCES

ABBOTT C.L., EBERT D., TABATA A. & THERRIAULT T.W. 2010: Twelve microsatellite markers in the invasive tunicate, Didemnum vexillum, isolated from low genome coverage 454 pyrosequencing reads. *Conserv. Genet. Resour.* **3**: 79–81.

ABDELKRIM J., ROBERTSON B.C., STANTON J.A.L. & GEMMELL N.J. 2009: Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *BioTechniques* **46**: 185–191.

ALTSCHUL S.F., MADDEN T.L., SCHÄFFER A.A., ZHANG J., ZHANG Z., MILLER W. & LIPMAN D.J. 1997: Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucl. Acids Res.* **25**: 3389–3402.

ANTHONY N., GELEMBIUK G., RATERMAN D., NICE C. & FFRENCH-CONSTANT R.H. 2001: Isolation and characterization of microsatellite markers from the endangered Karner blue butterfly Lycaeides melissa samuelis (Lepidoptera). *Hereditas* **134**: 271–273.

BELDADE P., VAN'T HOF A.E., JERÓNIMO M.A. & LONG A.D. 2009: Microsatellite markers associated with genes expressed in developing wings of Bicyclus anynana butterflies. *Mol. Ecol. Resour.* **9**: 1487–1492.

BENJAMINI Y. & HOCHBERG Y. 1995: Controlling the false discovery rate – a practical and powerful approach to multiple testing. *J. R. Stat. Soc. (B)* **57**: 289–300.

CASTOE T.A., POOLE A.W., GU W., DE KONING A.P.J., DANZA J.M., SMITH E.M. & POLLOCK D.D. 2010: Rapid identification of thousands of copperhead snake (Agkistrodon contortrix) microsatellite loci from modest amounts of 454 shotgun genome sequence. *Mol. Ecol. Resour.* **10**: 341–347.

COATES B.S., HELLMICH R.L. & LEWIS L.C. 2005: Polymorphic CA/GT and GA/CT microsatellite loci for Ostrinia nubilalis (Lepidoptera: Crambidae). *Mol. Ecol. Notes* **5**: 10–12.

COATES B.S., SUMERFORD D.V., HELLMICH R.L. & LEWIS L.C. 2009: Repetitive genome elements in a European corn borer, Ostrinia nubilalis, bacterial artificial chromosome library were indicated by bacterial artificial chromosome end sequencing and development of sequence tag site markers: Implications for lepidopteran genomic research. *Genome* **52**: 57–67.

CSENCSICS D., BRODBECK S. & HOLDEREGGER R. 2010: Cost-effective, species-specific microsatellite development for the endangered dwarf bulrush (Typha minima) using next-generation sequencing technology. *J. Hered.* **101**: 789–793.

DALECKY A., BOGDANOWICZ S.M., DOPMAN E.B., BOURGUET D. & HARRISON R.G. 2006: Two multiplex sets of eight and five microsatellite markers for the European corn borer, Ostrinia nubilalis Hubner (Lepidoptera: Crambidae). *Mol. Ecol. Notes* **6**: 945–947.

DALY D., WALTHAM K., MULLEY J., WATTS P.C., ROSIN A., KEMP S.J. & SACCHERI I.J. 2004: Trinucleotide microsatellite loci for the peppered moth (Biston betularia). *Mol. Ecol. Notes* **4**: 179–181.

DOPMAN E.B., BOGDANOWICZ S.M. & HARRISON R.G. 2004: Genetic mapping of sexual isolation between E and Z pheromone strains of the European corn Borer (Ostrinia nubilalis). *Genetics* **167**: 301–309.

DUBUT V., MARTIN J.F., COSTEDOAT C., CHAPPAZ R. & GILLES A. 2009: Isolation and characterization of polymorphic microsatellite loci in the freshwater fishes Telestes souffia and Telestes muticellus (Teleostei: Cyprinidae). *Mol. Ecol. Resour.* **9**: 1001–1005.

DUBUT V., GRENIER R., MEGLÉCZ E., CHAPPAZ R., COSTEDOAT C., DANANCHER D., DESCLOUX S., MALAUSA T., MARTIN J.F., PECH N. & GILLES A. 2010: Development of 55 novel polymorphic microsatellite loci for the critically endangered Zingel asper L. (Actinopterygii: Perciformes: Percidae) and cross-species amplification in five other percids. *Eur. J. Wildl. Res.* **56**: 931–938.

ENDERSBY N.M., MCKECHNIE S.W., VOGEL H., GAHAN L.J., BAXTER S.W., RIDLAND P.M. & WEEKS A.R. 2005: Microsatellites isolated from diamondback moth, Plutella xylostella (L.), for studies of dispersal in Australian populations. *Mol. Ecol. Notes* **5**: 51–53.

FAUVELOT C., CLEARY D.F.R. & MENKEN S.B.J. 2006: Short-term impact of 1997/1998 ENSO-induced disturbance on abundance and genetic variation in a tropical butterfly. *J. Hered.* **97**: 367–380.

FLANAGAN N.S., BLUM M.J., DAVISON A., ALAMO M., ALBARRÁN R., FAULHABER K., PETERSON E. & MCMILLAN W.O. 2002: Characterization of microsatellite loci in neotropical Helsonius butterflies. *Mol. Ecol. Notes* **2**: 398–401.

FRANCK P., GUÉRIN B., LOISEAU A. & SAUPHANOR B. 2005: Isolation and characterization of microsatellite loci in the codling moth Cydia pomonella L. (Lepidoptera, Tortricidae). *Mol. Ecol. Notes* **5**: 99–102.

HABEL J.C., FINGER A., MEYER M., SCHMITT T. & ASSMANN T. 2008: Polymorphic microsatellite loci in the endangered butterfly Lycaena helle (Lepidoptera: Lycaenidae). *Eur. J. Entomol.* **105**: 361–362.

HARPER G.L., MACLEAN N. & GOULSON D. 2003: Microsatellite markers to assess the influence of population size, isolation and demographic change on the genetic structure of the UK butterfly Polyommatus bellargus. *Mol. Ecol.* **12**: 3349–3357.

HIGGINS L.G. 1950: A descriptive catalogue of the Palaearctic Euphydryas (Lepidoptera: Rhopalocera). *Trans. R. Entomol. Soc. Lond.* **101**: 435–487.

HIGGINS L.G. 1955: A descriptive catalogue of the genus Mellicta Billberg (Lepidoptera: Nymphalidae) and its species, with supplementary notes on the genera Melitaea and Euphydryas. *Trans. R. Entomol. Soc. Lond.* **106**: 1–127.

JIGGINS C.D. 2005: A genetic linkage map of the mimetic butterfly Heliconius melpomene. *Genetics* **171**: 557–570.

KEYGHOBADI N., ROLAND J. & STROBECK C. 1999: Influence of landscape on the population genetic structure of the alpine

butterfly Parnassius smintheus (Papilionidae). *Mol. Ecol.* **8**: 1481–1495.

KIJAS J.M.H., FOWLER J.C.S., GARBETT C.A. & THOMAS M.R. 1994: Enrichment of microsatellites from the Citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. *Biotechniques* **16**: 656–662.

KIM K.S., COATES B.S., HELLMICH R.L., SUMERFORD D.V. & SAPPINGTON T.W. 2008: Isolation and characterization of microsatellite loci from the European corn borer, Ostrinia nubilalis (Hubner) (Insecta: Lepidoptera: Crambidae). *Mol. Ecol. Resour.* **8**: 409–411.

MARTIN J.F., PECH N., MEGLÉCZ E., FERREIRA S., COSTEDOAT C., DUBUT V., MALAUSA T. & GILLES A. 2010: Representativeness of microsatellite distributions in genomes, as revealed by 454 GS-FLX Titanium pyrosequencing. *BMC Genomics* **11**: 560.

MAZEL R. 1986: Structure et évolution du peuplement d'Euphydryas aurinia Rottemburg (Lepidoptera) dans le sud-ouest européen. *Vie Milieu* **36**: 205–225.

MEGLÉCZ E., PÉTÉNIAN F., DANCHIN E., COEUR D'ACIER A., RASPLUS J.Y. & 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.

MEGLÉCZ E., ANDERSON S.J., BOURGUET D., BUTCHER R., CALDAS A., CASSEL-LUNDHAGEN A., COEUR D'ACIER A., DAWSON D.A., FAURE N., FAUVELOT C., FRANCK P., HARPER G., KEYGHOBADI N., KLUETSCH C., MUTHULAKSHMI M., NAGARAJU J., PATT A., PÉTÉNIAN F., SILVAIN J.F. & WILCOCK H.R. 2007: Microsatellite flanking region similarities among different loci within insect species. *Insect Mol. Biol.* **16**: 175–185.

MEGLÉCZ E., COSTEDOAT C., DUBUT V., GILLES A., MALAUSA T., PECH N. & MARTIN J.F. 2010: QDD: A user-friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics* **26**: 403–404.

MIKHEYEV A.S., VO T., WEE B., SINGER M.C. & PARMESAN C. 2010: Rapid microsatellite isolation from a butterfly by de novo transcriptome sequencing: Performance and a comparison with AFLP-derived distances. *PLoS ONE* **5**: e11212.

MOLECULAR ECOLOGY RESOURCES PRIMER DEVELOPMENT CONSORTIUM, ABDOULLAYE D., ACEVEDO I., ADEBAYO A.A., BEHRMANN-GODEL J., BENJAMIN R.C., BOCK D.G., BORN C., BROUAT C., CACCONE A., CAO L.Z., CASADO-AMEZÚA P., CATANÉO J., CORREA-RAMIREZ M.M., CRISTESCU M.E., DOBIGNY G., EGBOSIMBA E.E., ETCHBERGER L.K., FAN B., FIELDS P.D., FORCIOLI D., FURLA P., GARCIA DE LEON F.J., GARCÍA-JIMÉNEZ R., GAUTHIER P., GERGS R., GONZÁLEZ C., GRANJON L., GUTIÉRREZ-RODRÍGUEZ C., HAVILL N.P., HELSEN P., HETHER T.D., HOFFMAN E.A., HU X., INGVARSSON P.K., ISHIZAKI S., JI H., JI X.S., JIMENEZ M.L., KAPIL R., KARBAN R., KELLER S.R., KUBOTA S., LI S., LI W., LIM D.D., LIN H., LIU X., LUO Y., MACHORDOM A., MARTIN A.P., MATTHYSEN E., MAZZELLA M.N., MCGEOCH M.A., MENG Z., NISHIZAWA M., O'BRIEN P., OHARA M., ORNELAS J.F., ORTU M.F., PEDERSEN A.B., PRESTON L., REN Q., ROTHHAUPT K.O., SACKETT L.C., SANG Q., SAWYER G.M., SHIOJIRI K., TAYLOR D.R., VAN DONGEN S., VAN VUUREN B.J., VANDEWOESTIJNE S., WANG H., WANG J.T., WANG L, XU X.L., YANG G., YANG Y., ZENG Y.Q., ZHANG Q.W., ZHANG Y., ZHAO Y. & ZHOU Y. 2010: Permanent genetic resources added to Molecular Ecology Resources database 1 August 2009 – 30 September 2009. *Mol. Ecol. Resour.* **10**: 232–236.

NAGARAJA G.M., MAHESH G., SATISH V., MADHU M., MUTHULAKSHMI M. & NAGARAJU J. 2005: Genetic mapping of Z chromosome and identification of W chromosome-specific markers in the silkworm, Bombyx mori. *Heredity* **95**: 148–157.

NOWAK C., BROWN C.M. & HELLMANN J.J. 2009: Microsatellite markers for the hybridizing tiger swallowtails, Papilio glaucus and P. canadensis, and their applicability to historic specimens and congeneric species. *Mol. Ecol. Resour.* **9**: 800–803.

PÉTÉNIAN F., MEGLÉCZ E., GENSON G., RASPLUS J.Y. & FAURE E. 2005: Isolation and characterization of polymorphic microsatellites in Parnassius apollo and Euphydryas aurinia (Lepidoptera). *Mol. Ecol. Notes* **5**: 243–245.

PRASAD M.D., MUTHULAKSHMI M., MADHU M., ARCHAK S., MITA K. & NAGARAJU J. 2005: Survey and analysis of microsatellites in the silkworm, Bombyx mori: Frequency, distribution, mutations, marker potential and their conservation in heterologous species. *Genetics* **169**: 197–214.

ROUSSET F. 2008: Genepop'007: A complete re-implementation of the Genepop software for Windows and Linux. *Mol. Ecol. Resour.* **8**: 103–106

SELKOE K.A. & TOONEN R.J. 2006: Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecol. Lett.* **9**: 615–629.

TAY W.T., BEHERE G.T., BATTERHAM P. & HECKEL D.G. 2010: Generation of microsatellite repeat families by RTE retrotransposons in Lepidopteran genomes. *BMC Evol. Biol.* **10**: 144.

TORRES-LEGUIZAMON M., SOLIGNAC M., VAUTRIN D., CAPDEVIELLE-DULAC C., DUPAS S. & SILVAIN J.F. 2009: Isolation and characterization of polymorphic microsatellites in the Potato Tuber Moth Tecia solanivora (Povolny, 1973) (Lepidoptera: Gelechiidae). *Mol. Ecol. Resour.* **9**: 1167–1169.

VAN OOSTERHOUT C., HUTCHINSON W.F., WILLS D.P.M. & SHIPLEY P. 2004: Micro-Checker: Software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* **4**: 535–538.

VAN SWAAY C.A.M. & WARREN M.S. 1999: *Red Data Book of European Butterflies (Rhopalocera).* Council of Europe Publishing, Strasbourg, 259 pp.

VAN'T HOF A.E., ZWAAN B.J., SACCHERI I.J., DALY D., BOT A.N.M. & BRAKEFIELD P.M. 2005: Characterization of 28 microsatellite loci for the butterfly Bicyclus anynana. *Mol. Ecol. Notes* **5**: 169–172.

VAN'T HOF A.E., BRAKEFIELD P.M., SACCHERI I.J. & ZWAAN B.J. 2007: Evolutionary dynamics of multilocus microsatellite arrangements in the genome of the butterfly Bicyclus anynana, with implications for other Lepidoptera. *Heredity* **98**: 320–328.

WEIR B.S. & HILL W.G. 2002: Estimating F-statistics. *Annu. Rev. Genet.* **36**: 721–750.

ZHANG D.X. 2004: Lepidopteran microsatellite DNA: Redundant but promising. *Trends Ecol. Evol.* **19**: 507–509.

ZHOU Y.H., GU H.N. & DORN S. 2005: Isolation of microsatellite loci in the codling moth, Cydia pomonella (Lepidoptera: Tortricidae). *Mol. Ecol. Notes* **5**: 226–227.

ZIMMERMANN M. 2000: *Biodiversité au Sein des Melitées (Lepidoptera: Nymphalidae): Phylogénie et Génétique des Populations chez Euphydryas aurinia.* PhD thesis, Université Aix-Marseille I, France, 316 pp.