Development and characterization of 13 new microsatellite markers in the pine processionary moth, Thaumetopoea pityocampa (Lepidoptera: Notodontidae)

Christian Burban, Emmanuelle Magnoux, Jérôme Rousselet, Carole Kerdelhue

Development and characterization of 13 new microsatellite markers in the pine processionary moth, *Thaumetopoea pityocampa* (Lepidoptera: Notodontidae)

C. Burban[1*], E. Magnoux[2], J. Rousselet[2] and C. Kerdelhué[3]

[1]*INRA, UMR1202 BIOGECO (INRA/Université de Bordeaux), F-33610 Cestas, France*

[2]*INRA, UR633 Zoologie forestière, 45075 Orléans cedex 2, France*

[3]*INRA, UMR CBGP (INRA/IRD/CIRAD/Montpellier Supagro), F-34988 Montferrier-sur-Lez, France*

Keywords: microsatellites, Lepidoptera, pine processionary moth, *Thaumetopoea pityocampa*

* to whom correspondence should be addressed:

Christian Burban

INRA, UMR BIOGECO

69, route d'Arcachon

F-33610, Cestas, France

e-mail: christian.burban@pierroton.inra.fr

tel: +33 (0)5 57 12 27 23

fax: +33 (0)5 57 12 28 81

Short title: **Microsatellites for *Thaumetopoea pityocampa***

Abstract

We describe the development of 13 microsatellite markers in the pine processionary moth *Thaumetopoea pityocampa*. In a single Portuguese population tested, the number of alleles per locus ranged from 2 to 12, and observed heterozygosity from 0.07 to 0.83. Cross amplifications gave satisfactory results for most loci in the *T. pityocampa/T. wilkinsoni* complex but proved to be of low interest for other species. These markers will be valuable to investigate the population genetic structure of this forest pest, particularly for the expanding populations at the northern limits of its range, and for a new phenological race discovered in Portugal.

The isolation of microsatellite markers in Lepidoptera is known to be difficult, and therefore only few loci are available for most species (Zhang 2004). The pine processionary moth *Thaumetopoea pityocampa* is a forest pest, responsible for heavy loss of production in pine forests in Southern Europe through the consumption of needles by its gregarious larvae. It is also known from public-health and veterinary services because of its urticating hairs that can induce allergic reactions. The distribution area of this species is clearly limited both by the availability of suitable hosts and by climatic conditions during larval development, i.e. high summer temperatures in the south and low winter temperatures in the northern areas. Its distribution range is expanding northward and in higher altitudes due to the ongoing climate warming, and in particular to the improvement of winter conditions (Battisti *et al*. 2005). Another matter of interest is the presence in Portugal of a peculiar, unique population that exhibits a shifted phenology, with reproduction in May and larval development in summer (summer population) while the classical life cycle implies a reproduction in autumn and larval development during winter. The summer population is supposed to have originated from the local winter population (Santos *et al*. 2011a) and seems to be adapted to higher temperature during larval development (Santos *et al*. 2011b). The shift in phenology is such that it prevents crosses between the sympatric winter and summer populations (Santos *et al*. 2011a).

A first set of 6 loci was previously developed for this species (Rousselet *et al*. 2004; Santos *et al*. 2007). Yet, this was far too few to develop landscape genetics approaches and investigate the genetic structure and origin of newly settled populations around the expansion fronts, either at the northern limits of the natural range or at higher elevation in mountains. It was also insufficient to test scenarios about the foundation and origin of the allochronic population in Portugal. We thus decided to develop new polymorphic microsatellite markers, and we describe here the development of 13 new loci.

An enriched library was made by Ecogenics GmbH (Zurich, Switzerland) from size selected genomic DNA ligated into SNX forward/SNX reverse linker (Hamilton *et al*. 1999) and enriched by magnetic bead selection with biotin-labelled $(CT)_{13}$, $(GT)_{13}$, $(GTAT)_7$ and $(GATA)_7$ oligonucleotide repeats (Gautschi *et al*. 2000a, b). Of 528 recombinant colonies screened, 242 gave a positive signal after

hybridization. Plasmids from 126 positive clones were sequenced and primers were designed for 24 microsatellite inserts, of which 21 were tested for polymorphism. 13 primer pairs were finally selected that gave satisfying experimental results, and grouped in 4 multiplex.

Amplification for each multiplex was optimized to be performed in a 10 µl reaction volume containing <10 ng of DNA, 5 µl 2X QIAGEN multiplex PCR master mix, distilled water, and 0.3 µM of forward and reverse primers each, 10% to 50% of the forward primer being labeled with a fluorescent dye (either 6-Fam, Hex, Ned or Vic, Applied Biosystems). We used the following thermotreatment: 35 cycles at 94°C for 30 seconds, 56°C for 90 seconds, and 72°C for 60 seconds. Before the first cycle, a prolonged denaturation step (95°C for 15 min) was included and the last cycle was followed by a 30 min extension at 72°C. Fragments were run on an ABI 3730 automatic sequencer using the size standard GeneScan-600 LIZ, and sized with GENEMAPPER 4.0 software (Applied Biosystems).

Microsatellite loci were amplified in multiplex from 30 larvae belonging to a Portuguese population exhibiting the classical winter phenology, in the forest of Leiria (39°47'N, 8°58'W). We also tested cross amplification, with the same PCR conditions, for individuals from different clades of the *T. pityocampa/T. wilkinsoni* complex, namely from the *pityocampa*, *wilkinsoni* and ENA clades (Kerdelhué *et al*. 2009) and for 5 others species of the same genus.

Null allele frequencies were estimated using FreeNA (Chapuis & Estoup 2007) following the Expectation Maximization algorithm (Dempster *et al*. 1977). Genetic diversity indices were calculated, and used to test for Hardy-Weinberg equilibrium and linkage disequilibrium with ARLEQUIN 3.11 software (Excoffier *et al*. 2005).

Polymorphism characteristics in the Leiria population are provided for each locus in Table 1. Number of alleles per locus ranged from 2 (MS-*Thpit*17 and 18) to 12 (MS-*Thpit*13). The estimated null allele frequencies were below 5% for all loci except MS-*Thpit*11. This locus was also the only one exhibiting deviation from Hardy-Weinberg equilibrium. This deficit of heterozygotes is probably linked to the 13% of null alleles estimated for this locus. No significant linkage disequilibrium was detected between any pair of loci ($p < 0.01$).

Cross amplifications proved the usefulness of these microsatellites in the *T. pityocampa/T. wilkinsoni* complex (Table 2). Most loci can be successfully amplified in all localities and exhibit polymorphism in each taxon despite the low sampling effort (but a large geographical scale). For *T. wilkinsoni*, amplification failed for all individuals when using MS-*Thpit*7 and 9, and for some individuals with MS-*Thpit*10, 14 and 19. Such partial success was also obtained for individuals from the ENA clade with MS-*Thpit*9 and 19. Considering the others species on gymnosperm, we obtained positive and clear amplification only for 2 to 5 loci, the individual tested for each species being homozygous. Finally, no amplification product was obtained for *T. solitaria*, the only species hosted by an Angiosperm.

These results mean that we now benefit from a total of 19 microsatellite loci for *T. pityocampa* population genetic studies. This will allow to more precisely investigate the origin of the allochronic population in Portugal, and to look for the existence of natural hybrids with the sympatric winter population. It should be also possible to characterize the process of latitudinal range expansion in regard to climate warming and landscape characteristics. It opens further perspectives to link the evolution of adaptive traits and neutral genetic patterns underlying the establishment of new populations.

Acknowledgements

The authors are grateful to M. Branco and H. Santos for collecting the samples from Portugal, and to A. Battisti, M.L. Ben Jamâa, D. Ghaioule, F. Goussard, K. Ipekdal, M. Kalapanida, N. Nemer and the University Omar Almukhtar for sampling of the other taxa. I. Pivotto and H. Santos are acknowledged for their help with DNA extraction. Thanks to S. Monllor and F. Salin, the genotyping was performed at the Genome-Transcriptome facility of Bordeaux. This work was supported by the ANR 07 BDIV 013-07-URTICLIM.

References

Battisti A, Stastny M, Netherer S, Robinet C, Schopf A, Roques A, Larsson S (2005) Expansion of geographic range in the pine processionary moth caused by increased winter temperatures. *Ecological Applications*, **15**, 2084–2096.

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

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B Statistical Methodology*, **39**, 1–38.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Gautschi B, Tenzer I, Müller JP, Schmid B (2000a) Isolation and characterization of microsatellite loci in the bearded vulture (*Gypaetus barbatus*) and crossamplification in three Old World vulture species. *Molecular Ecology*, **9**, 2193-2195.

Gautschi B, Widmer A, & Koella J (2000b) Isolation and characterization of microsatellite loci in the dice snake (*Natrix tessellata*). *Molecular Ecology*, **9**, 2191-2193.

Hamilton MB, Pincus EL, Di Fiore A, Fleischer RC (1999) Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *Biotechniques*, **27**, 500-507.

Kerdelhué C, Zane L, Simonato M, Salvato P, Rousselet J, Roques A, Battisti A (2009) Quaternary history and contemporary patterns in a currently expanding species. *BMC Evolutionary Biology*, **9**: 220.

Rousselet J, Magnoux E, Kerdelhué C (2004) Characterization of five microsatellite loci in the pine processionary moth *Thaumetopoea pityocampa* (Lepidoptera: Notodontidae: Thaumetopoeinae). *Molecular Ecology Notes*, **4**, 213–214.

Santos H, Rousselet J, Magnoux E, Paiva MR, Branco M, Kerdelhué C (2007) Genetic isolation through time: allochronic differentiation of a phenologically atypical population of the pine processionary moth. *Proceedings of the Royal Society of London B Biological Sciences*, 274, 935–941.

Santos H, Burban C, Rousselet J, Rossi J-P, Branco M, Kerdelhué C (2011a) Incipient allochronic speciation in the pine processionary moth *Thaumetopoea pityocampa* (Lepidoptera: Notodontidae). *Journal of Evolutionary Biology*, **24**: 146-158.

Santos H, Paiva MR, Tavares C, Kerdelhué C, Branco M (2011b) Temperature niche shift observed in a Lepidoptera population under allochronic divergence. *Journal of Evolutionary Biology*, doi: 10.1111/j.1420-9101.2011.02318.x

Zhang D-X (2004) Lepidopteran microsatellite DNA: redundant but promising. *Trends in Ecology & Evolution*, **19**, 507–509.

Table 1 Microsatellite data and polymorphism characterization of the *T. pityocampa* population from Leiria

Locus	Primer sequences (5'-3')	Repeat motif (cloned allele)	Multiplex	Genbank Accession no.	Allele range (bp)	N_A	Null allele frequency	H_O	H_E
MS-Thpit 7	F GGCTCGTGACCTAAAACTGC R ACACGCCCACATACATGC	$(TG)_{13}$	M1	JN400258	237-247	6	0.01480	0.633	0.693
MS-Thpit 8	F GAGTGGTCGCCCTGTCTG R CTCCAATGCCAAGGTAGACAC	$(CA)_{15}(CG)(CA)_{10}$	M1	JN400259	182-228	6	0.00000	0.600	0.543
MS-Thpit 9	F TTGCACAAGACAATAAAAAGAACAG R TCTCACTCTCTCACACACACTATTTC	$(AG)_{26}$	M1	JN400260	85-107	6	0.02109	0.367	0.404
MS-Thpit 10	F TGAATTTGGAGGAAACGATG R CCACTGCATGTTCGCATAAC	$(GT)_{10}$	M1	JN400261	148-164	4	0.02924	0.633	0.683
MS-Thpit 11	F GCAGGTGGAAAGAAGTGGTC R TCTCCGGAAGAGAATCCAAG	$(GT)_{15}$	M2	JN400262	198-226	6	0.12928	**0.167**	**0.275**
MS-Thpit 12	F GTTCTACGTCGGGAACAAGC R ACGGTTCCTGTGCGATATAC	$(CA)_{19}$	M2	JN400263	130-176	8	0.00001	0.567	0.625
MS-Thpit 13	F TTTTGATTGTCAGCACTCACG R AGAGCGTGGTTCGTTCTTTG	$(CA)_{10}(CG)(CA)_5(CG)(CA)_6$	M2	JN400264	99-137	12	0.00000	0.833	0.815
MS-Thpit 14	F TCACTACTGGACTTCAACGACTG R CGAGAAGTAAAAGTGGTATGTGT	$(AC)_{11}(GCAA)(AC)_{13}$	M3	JN400265	216-372	10	0.03082	0.667	0.701
MS-Thpit 15	F ACACGCGCACGCTTACTC R TCAGCAAAGTTGGAGCATAAAC	$(AC)_{25}$	M3	JN400266	150-198	11	0.00464	0.733	0.798
MS-Thpit 16	F ACACGACAGGCTTGCACA R TTTTCCTTTACTAACTCTGTCCAT	$(AG)_{22}$	M3	JN400267	51-65	3	0.02592	0.400	0.468
MS-Thpit 17	F AAAGGCCAAGTTCGTAGCAC R CCACATACAAAAGTTCAAAAATACG	$(TATG)_7$	M4	JN400268	180-184	2	0.00003	0.067	0.065
MS-Thpit 18	F TTCAAGAACAATGTGGCTTTC R GTCTACCAACCCGCACTAGG	$(TCA)_9$	M4	JN400269	144-147	2	0.00000	0.533	0.427
MS-Thpit 19	F AAGCACTCGTCCGCAGTC R GGGGTTCTAAAATGCCAGTTG	$(TG)_{11}..(TG)_{12}..(GT)_{14}..(GT)_{14}$	M4	JN400270	113-311	6	0.00001	0.633	0.662

N_A: number of alleles; H_O: observed heterozygosity; H_E: expected heterozygosity; values in bold indicate significant deviation from Hardy-Weinberg equilibrium ($P < 0.05$)

Table 2 Cross-species amplification results for *Thaumetopea* spp.: number of successful amplifications / number of alleles in the taxon

Species	host plant	Country	*N*	MS-Thpit 7	MS-Thpit 8	MS-Thpit 9	MS-Thpit 10	MS-Thpit 11	MS-Thpit 12	MS-Thpit 13	MS-Thpit 14	MS-Thpit 15	MS-Thpit 16	MS-Thpit 17	MS-Thpit 18	MS-Thpit 19
T. pityocampa	*Pinus*	Morocco, W.Turkey	2	2/3	2/4	2/4	2/4	2/4	2/3	2/3	2/3	2/3	2/3	2/1	2/3	2/4
T. p. ena clade	*Pinus*	Algeria, Libya, Tunisia	3	3/6	3/4	1/1	3/4	3/4	3/2	3/6	3/5	3/4	3/5	3/2	3/3	2/2
T. wilkinsoni	*Pinus*	Crete, Cyprus, S. & N Turkey	4	0	4/5	0	1/2	4/1	4/2	4/3	3/6	4/7	4/6	4/1	4/1	3/5
T. pinivora	*Pinus*	France	1	1/1	0	0	0	1/1	1/1	1/1	0	0	0	0	0	0
T. libanotica	*Cedrus*	Lebanon	1	1/1	0	0	0	1/1	1/1	1/1	0	0	0	0	1/1	0
T. bonjeani	*Cedrus*	Algeria	1	1/1	0	0	0	1/1	1/1	0	0	0	0	0	0	0
T. ispartensis	*Cedrus*	Turkey	1	1/1	0	0	0	1/1	0	0	0	0	0	0	0	0
T. solitaria	*Pistacia*	Greece	2	0	0	0	0	0	0	0	0	0	0	0	0	0

N: number of individuals tested