MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2009) **9**, 1369–1372 doi: 10.1111/j.1755-0998.2009.02696.x

MOLECULAR DIAGNOSTICS AND DNA TAXONOMY

Species identification and sibship assignment of sympatric larvae in the yucca moths *Tegeticula synthetica* and *Tegeticula antithetica* (Lepidoptera: Prodoxidae)

CHRISTOPHER S. DRUMMOND,* CHRISTOPHER I. SMITH† and OLLE PELLMYR*
*Department of Biological Sciences, University of Idaho, PO Box 443051, Moscow, ID 83844, USA, †Department of Biology, Willamette University, 900 State Street, Salem, OR 97301, USA

Abstract

Ecological interactions between yucca moths (*Tegeticula*, Prodoxidae) and their host plants (*Yucca*, Agavaceae) are exemplary of obligate plant–pollinator mutualism and co-evolution. We describe a multiplex microsatellite DNA protocol for species identification and sibship assignment of sympatric larvae from *Tegeticula synthetica* and *Tegeticula antithetica*, pollinators of the Joshua tree (*Yucca brevifolia*). Bayesian clustering provides correct diagnosis of species in 100% of adult moths, with unambiguous identification of sympatric larvae. Sibship assignments show that larvae within a single fruit are more likely to be full-sibs or half-sibs than larvae from different fruit, consistent with the hypothesis that larval clutches are predominantly the progeny of an individual female.

Keywords: assignment, microsatellite, Prodoxidae, sibship, *Tegeticula*, *Yucca*

Received 7 January 2009; revision accepted 27 January 2009

Joshua trees (*Yucca brevifolia*, Agavaceae) are pollinated exclusively by yucca moths of the genus *Tegeticula* (*Tegeticula antithetica*, *Tegeticula synthetica*, Prodoxidae). Conversely, yucca moths reproduce by laying eggs in Joshua tree flowers, which are deliberately pollinated using specialized mouthparts, thereby ensuring that the flower will develop and provide food for the larvae. This interdependence has made yucca moths and yuccas a principal example of obligate mutualism. While *T. synthetica* and *T. antithetica* are parapatrically distributed in the western and eastern portions of the range for *Y. brevifolia* (Pellmyr & Segraves 2003), the two species occur in sympatry at Tikaboo Valley, Nevada (Smith *et al.* 2008). To facilitate ongoing research related to the yucca moth/Joshua tree system, we report the development of novel microsatellite DNA markers for *Tegeticula* and demonstrate a reliable protocol for species identification and sibship assignment of moth larvae, which cannot be identified to taxon by morphology.

Genomic DNA was extracted from: (i) 114 adult moths from across the range of *Y. brevifolia* (*T. antithetica*, $n = 53$; *T. synthetica*, $n = 61$); and (ii) 73 larvae from the Tikaboo Valley contact zone. Extractions followed the standard protocol for DNeasy tissue kits (QIAGEN). An enriched DNA library for *T. synthetica* was obtained from Ecogenics GmbH using size-selected genomic DNA ligated into SAULAA/SAULB linker oligonucleotides and hybridized with magnetic streptavidin beads and biotin-labelled oligonucleotide repeats: $(GT)_{13}$, $(CT)_{13}$, $(TAC)_{10}$, $(ATC)_{10}$ (e.g. Armour *et al.* 1994). Hybridization screening of 352 recombinant colonies revealed 216 positive clones, and 48 positive clones were sequenced. Primers were designed for 16 loci, and polymerase chain reaction (PCR) products were tested on high-resolution agarose for successful amplification yielding a total of nine candidate loci.

MultiPLX 1.2 (Kaplinski *et al.* 2004) was used to predict suitable combinations of primer pairs for two multiplex PCR comprising four and five reactions, respectively. Multiplex PCR panels with fluorescent dye-labelled primers were optimized using the reaction conditions shown in Table 1. Genotyping was performed on an ABI 3130 capillary instrument (Applied Biosystems) with 1 μL of multiplex PCR product in 10 μL HiDi formamide and 0.15 μL Genescan 500 LIZ size standard (Applied Biosystems).

Correspondence: Christopher S. Drummond, Fax: (208) 885 7905; E-mail: cdrummon@uidaho.edu

Table 1 Summary statistics from nine microsatellite DNA loci genotyped for 114 adult moths of *Tegeticula synthetica* (n = 61) and *Tegeticula antithetica* (n = 53) including number of alleles (N_A), expected heterozygosity (H_E), observed heterozygosity (H_O), null allele frequencies, and error rates.

						Tegeticula synthetica				*Tegeticula antithetica*				Error rates		
Locus	Primer sequences (5′–3′)	Repeat motif	Size range (bp)	Primer amount (pmol)	Genbank Accession	N_A	H_E	H_O	Null allele frequency	N_A	H_E	H_O	Null allele frequency	Replicate PCR	Allelic dropout	False alleles
Multiplex #1																
Tegsyn02	F: NED-TTCCGAAATATGTGCCTGTG R: ACGACGGCGCTCTAGTAGTC	$(ACT)_{12}$	131–281	1.0	FJ655888	17	0.864	0.811	0.000	32	0.958	0.926	0.033	0.323	0.000	0.000
Tegsyn03	F: VIC-CCTTCATACGGTGAGGCAAC R: TCACACCGTGACTCATGAAAG	$(CAT)_9$	94–122	4.0	FJ655889	10	0.834	0.811	0.011	9	0.697	0.754	0.020	0.323	0.013	0.011
Tegsyn06	F: VIC-CCGTGCCCATTCTAGAAGAC R: CGTAGCCCTGGGTAAAGGAG	$(CAT)_9$	218–233	0.7	FJ655891	6	0.597	0.132*	0.245	9	0.657	0.321*	0.205	0.323	0.089	0.000
Tegsyn13	F: 6FAM-CCGTGCTACAACATTCGTTC R: AAGCTAACCTTATTCATCGCAATC	$(GT)_{14}$	180–239	1.0	FJ655893	18	0.776	0.661	0.037	18	0.865	0.868	0.000	0.303	0.004	0.005
Tegsyn14	F: PET-GTAGAGCCTGCCTTGACCTG R: GGCATTCCAAGACATTATTGC	$(CA)_{17}$	146–218	1.0	FJ655894	25	0.944	0.925	0.017	28	0.911	0.792	0.057	0.303	0.000	0.000
Multiplex #2																
Tegsyn01	F: VIC-ACAGCCATACAAAA CAAATCATTA R: TTGGGAAGCAAAAACTTTACTG	$(TGA)_9$	87–114	1.5	FJ655887	7	0.741	0.404*	0.160	2	0.368	0.068*	0.389	0.206	0.009	0.023
Tegsyn05	F: 6FAM-GCTACTCAAGGAGT CAAAGCAAC R: TTCCTGGTCGACTTCCAGAC	$(TCA)_8$	151–179	2.0	FJ655890	5	0.642	0.519	0.123	10	0.754	0.623	0.088	0.202	0.041	0.000
Tegsyn09	F: PET-ATTGTACCGCTCCAACGAAC R: ATAAAGTGGCACGTGTTTGC	$(AC)_{12}$	178–236	2.0	FJ655892	12	0.731	0.500	0.114	21	0.962	0.919	0.000	0.202	0.035	0.015
Tegsyn15	F: VIC-TAAAGGGGCAGAACAAATGG R: GCCATATTTGCCGCTAGG	$(AC)_{13}$	219–307	1.5	FJ655895	35	0.964	0.865	0.042	29	0.960	0.725*	0.138	0.202	0.000	0.021

Significant deviations from HWE after Bonferroni correction for multiple tests ($P < 0.005$) are indicated by an asterisk (*) next to values for H_O. Maximum likelihood estimates of null allele frequencies, allelic dropout rates and false allele rates are based on replicate genotyping of 20.2–32.3% of individuals. Genotype data were obtained from two sets of multiplex PCR performed in 10 μL volumes (5 μL 2X Qiagen multiplex PCR master mix, 1 μL Qiagen Q solution, 1 μL genomic DNA, fwd/rev primer amounts listed below) using the following thermal profile: (i) 95 °C for 15 min; (ii) 27X 95 °C for 60 s, 55 °C for 90 s, 72 °C for 90 s; (iii) 72 °C for 30 min.

Because microsatellite DNA genotypes may be prone to error (Bonin *et al.* 2004; Dakin & Avise 2004; Hoffman & Amos 2005; Pompanon *et al.* 2005), we used the following approach to design an objective and reproducible format for scoring alleles. Provisionary bins were developed in GeneMapper 4.0 (Applied Biosystems) to identify true peaks vs. PCR and/or electrophoresis artefacts. Raw fragment sizing data were exported to a spreadsheet and used to compile cumulative frequency plots of estimated size distributions in FlexiBin 2.0 (Amos *et al.* 2007). New bins for the inferred number of repeats were then constructed around these distributions where discrete breaks in periodic size classes were evident. Blind replicate genotypes were obtained from a random sample of individuals (20.2–30.3%).

Summary statistics for each locus (Table 1) were calculated for known adults of each species using GENEPOP 4.0 (Rousset 2008), MSA 4.05 (Dieringer & Schlötterer 2003), FreeNA (Chapuis & Estoup 2006) and PEDANT 1.0 (Johnson & Haydon 2007a, b). Exact tests (Guo & Thompson 1992) for Hardy–Weinberg equilibrium (HWE) indicated three loci (Tegsyn01, Tegsyn06, Tegsyn15) with significant deviations from HWE after Bonferroni correction for multiple tests ($P < 0.005$). Rejection of HWE may be symptomatic of: (i) biological processes such as population substructure and nonrandom mating; or (ii) genotyping artefacts resulting from null alleles, PCR dropout and other types of scoring error (Bonin *et al.* 2004; Pompanon *et al.* 2005). Maximum likelihood (ML) estimation of error rates (Johnson & Haydon 2007a, b) and null allele frequencies (Chapuis & Estoup 2006) suggested that these three loci are associated with elevated frequencies of null alleles (13.8–38.9%), whereas the remaining loci

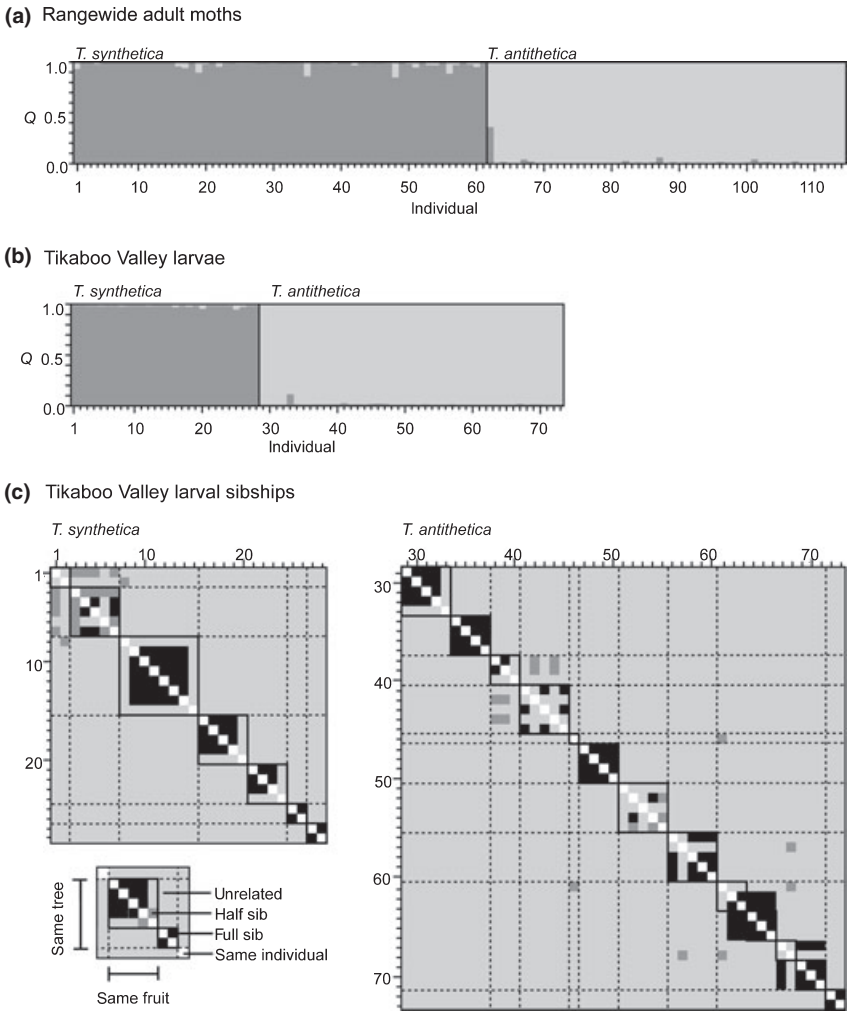

Fig. 1 Species identification and sibship assignment of *Tegeticula synthetica* and *Tegeticula antithetica* from microsatellite DNA genotype data: (a) Bayesian clustering of adults across the range of *Yucca brevifolia*, with bar plot showing ancestry coefficients (Q) for $k = 2$; (b) Bayesian clustering of larvae at Tikaboo Valley; (c) maximum likelihood estimates of pairwise sibship (black = full-sib, dark grey = half-sib, light grey = unrelated) among larvae isolated from different trees (dashed lines) and fruit (solid boxes) at Tikaboo Valley.

recovered low to moderate levels of null alleles (0.0–12.3%), allelic dropout (0.0–4.1%) and false peaks (0.0–2.3%). Exact tests (Raymond & Rousset 1995) for linkage disequilibrium showed that no pairs of loci exhibited significant linkage after Bonferroni correction for multiple tests ($P < 0.001$), consistent with effectively independent segregation.

We then evaluated the utility of the full panel of nine loci for: (i) species identification of adult and larval moths using Bayesian clustering in Structure 2.3.2 (Pritchard *et al*. 2000; Falush *et al*. 2007); and (ii) sibship assignment among larval clutches isolated from different fruit using ML in COLONY 2.0 (Wang 2004). Few microsatellite DNA markers are free from genotyping error. However, there are advantages to retaining marginally problematic loci when error can be quantified and modelled effectively using methods that incorporate estimates of null allele frequencies and/or genotyping error (e.g., Wang 2004; Falush *et al*. 2007).

Bayesian clustering of adult moths ($k = 2$) without a priori designation of groups assigned 100% of individuals to the correct morphospecies (Fig. 1a) based on 95% CI values for ancestry coefficients (Q). Likewise, larvae from Tikaboo Valley were assigned unambiguously to species (Fig. 1b). ML estimates of sibship indicated that larvae sampled within the same fruit were more likely to be full-sibs or half-sibs than larvae sampled from different fruit (Fig. 1c). These results are consistent with prior taxonomical studies (Pellmyr & Segraves 2003) and behavioural observations of female oviposition (O. Pellmyr & C. Smith, unpublished).

We have described the development and implementation of a multiplex microsatellite DNA protocol for species identification and sibship assignment of yucca moths in the Prodoxidae. These markers can be used to identify sister species of *Tegeticula* including larvae that cannot be identified to taxon by morphology. We also are able to assign sibship among larval clutches isolated from different fruit. Based on these results, this panel of microsatellite DNA markers is sufficient to yield reliable and biologically informative results when used with analyses that model genotyping error, providing new opportunities to ask questions about co-evolution between yucca moths and Joshua trees.

Acknowledgements

J. B. Yoder and W. Godsoe (University of Idaho) participated in field collections. A. Hite (University of Idaho) provided laboratory support. This research was funded by the National Science Foundation (DEB 0516841). Collection permission was provided by the US National Park Service, US Fish and Wildlife Service and US Department of Energy.

References

Amos W, Hoffman JI, Frodsham A, Zhang L, Best S, Hill VS (2007) Automated binning of microsatellite alleles: problems and solutions. *Molecular Ecology Notes*, **7**, 10–14.

Armour JA, Neumann R, Gobert S, Jeffreys AJ (1994) Isolation of human simple repeat loci by hybridization selection. *Human Molecular Genetics*, **3**, 599–605.

Bonin A, Bellemain E, Bronken Eidesen P, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*, **13**, 3261–3273.

Chapuis MP, Estoup A (2006) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, **24**, 621–631.

Dakin EE, Avise JC (2004) Microsatellite null alleles in parentage analysis. *Heredity*, **93**, 504–509.

Dieringer D, Schlötterer C (2003) MICROSATELLITE ANALYSER (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes*, **3**, 167–169.

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574–578.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361–372.

Hoffman JI, Amos W (2005) Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology*, **14**, 599–612.

Johnson PCD, Haydon DT (2007a) Maximum–likelihood estimation of allelic dropout and false allele error rates from microsatellite genotypes in the absence of reference data. *Genetics*, **175**, 827–842.

Johnson PCD, Haydon DT (2007b) Software for quantifying and simulating microsatellite genotyping error. *Bioinformatics and Biology Insights*, **1**, 71–75.

Kaplinski K, Andreson R, Puurand T, Remm M (2004) MultiPLX: automatic grouping and evaluation of PCR primers. *Bioinformatics*, **21**, 1701–1702.

Pellmyr O, Segraves KA (2003) Pollinator divergence within an obligate mutualism: two yucca moth species (Lepidoptera; Prodoxidae: *Tegeticula*) on the Joshua Tree (*Yucca brevifolia*; Agavaceae). *Annals of the Entomological Society of America*, **96**, 716–722.

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics*, **6**, 847–859.

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

Raymond M, Rousset F (1995) An exact test for population differentiation. *Evolution*, **49**, 1283–1286.

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.

Smith CI, Godsoe WKW, Tank S, Yoder JB, Pellmyr O (2008) Distinguishing coevolution from covicariance in an obligate pollination mutualism: asynchronous divergence in Joshua tree and its pollinators. *Evolution*, **62**, 2676–2687.

Wang J (2004) Sibship reconstruction from genetic data with typing errors. *Genetics*, **166**, 1963–1979.