HHS Public Access

Author manuscript

Mol Ecol Resour. Author manuscript; available in PMC 2016 February 22.

Published in final edited form as:

Mol Ecol Resour. 2012 September ; 12(5): 909–917. doi:10.1111/j.1755-0998.2012.03150.x.

A suite of molecular markers for identifying species, detecting introgression, and describing population structure in spadefoot toads (Spea spp.)

Karin S. Pfennig[1,5], **Ashley Allenby**[1], **Ryan A. Martin**[1,4], **Anaïs Monroy**[1], and **Corbin D. Jones**[1,2]

[1]Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

[2]Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

[4]Current address: Department of Biology, North Carolina State University, Raleigh, NC 27695, USA

Abstract

Two congeneric species of spadefoot toad, Spea multiplicata and S. bombifrons, have been the focus of hybridization studies since the 1970s. Because complex hybrids are not readily distinguished phenotypically, genetic markers are needed to identify introgressed individuals. We therefore developed a set of molecular markers (AFLP, PCR – RFLP, and SNP) for identifying pure species, F1 hybrids, and more complex introgressed types. To do so, we tested a series of markers across both species and known hybrids using populations in both allopatry and sympatry. We retained those markers that differentiated the two pure species and also consistently identified known species hybrids. These markers are well suited for identifying hybrids between these species. Moreover, those markers that show variation within each species can be used in conjunction with existing molecular markers in studies of population structure and gene flow.

Keywords

Keywords: S. multiplicata; S. bombifrons; AFLP; RFLP; SNP; hybridization; speciation; reinforcement; local adaptation

Introduction

Spadefoot toads, particularly those in the genus Spea, are an emerging model system for addressing problems ranging from ecotoxicology and wetlands ecology to evolutionary development, sexual selection, and speciation (reviewed in Ledon-Rettig & Pfennig 2011; Pfennig 2000; Gray *et al.* 2004; Banbury & Maglia 2006; Arendt 2009; Martin & Pfennig 2009; McMurry *et al.* 2009). The interactions of two Spea species in particular, S.

[5]Correspondence to: Karin Pfennig, <kpfennig@unc.edu>.

Data Accessibility: DNA sequences can be found in supplemental sequences text file; GenBank accession numbers are provided in Table 2. Genotype data are provided in supplemental text files.

multiplicata and S. bombifrons, have been studied since at least the 1970s (Forester 1973; Pierce 1976; Sattler 1985). These species co-occur in some parts of the southwestern USA (Stebbins 2003) where they show a mosaic distribution: some populations consist of one or both species depending on local conditions (Pfennig *et al.* 2006). Although the most divergent of the Spea genus (Wiens & Titus 1991; García-Paris *et al.* 2003), S. multiplicata and S. bombifrons naturally hybridize and produce viable hybrid offspring (Forester 1975; Simovich 1985; Simovich & Sassaman 1986; Pfennig & Simovich 2002). Hybrid females are partially fecund and can backcross with males of the parental species to produce complex backcross hybrid offspring (Forester 1975; Sattler 1985; Simovich 1985; Pfennig & Simovich 2002). Although F1 hybrid adults can be identified reliably via adult morphology, tadpoles and complex hybrids must be identified by genotype (Sattler 1985; Simovich & Sassaman 1986).

Because of the nature of the distribution of these two species, and because they hybridize, genetic markers are often necessary to determine species composition of a population and the degree to which introgression is taking place. However, only a limited number of allozyme markers have been available for distinguishing S. multiplicata, S. bombifrons, and their hybrids (Sattler 1985; Simovich & Sassaman 1986). Although allozymes are generally adequate for identifying hybrids, their use can be difficult given the restrictive conditions of sample preservation (i.e., freezing tissue). Allozymes are, therefore, not amenable for fieldwork or for genotyping ethanol-preserved specimens.

Recently, Rice *et al.* (2008) developed nine polymorphic microsatellite markers that amplified in both S. multiplicata and S. bombifrons. Although initially developed with the intent of differentiating the two species and their hybrids, these microsatellites do not reliably distinguish the two species, let alone their hybrids. Thus, our goal was to develop a set of marker loci for S. multiplicata and S. bombifrons that would enable researchers to identify the two species and their hybrids. In particular, we sought to identify markers that would be diagnostic for both species and their hybrids, reliable for different sample types, and easy to genotype. Because microsatellites are expensive to develop and those already developed were found to be too highly variable across species to be diagnostic (Rice *et al.* 2008), we focused on developing alternative diagnostic molecular markers.

Our approach was to develop diagnostic amplified fragment length polymorphisms (AFLPs) and polymerase chain reaction-restriction fragment length polymorphisms (PCR–RFLPs). Although high throughout sequencing has made it possible to identify species-specific differences across the genome (e.g., by using restriction site associated makers, i.e., RAD-tags), the cost per sample for such analyses can be prohibitive even though the cost per marker identified is low. Furthermore, the technology to sequence RAD-tags and the expertise to analyze the results are not yet universally available. RFLPs and AFLPS, by contrast, are technologically accessible, easy to interpret, and can be used to genotype large numbers of individuals. Thus, our approach remains a general alternative to next generation methods.

In developing these markers, we also identified a number of potentially useful species-specific single nucleotide polymorphisms (SNPs) that we have evaluated for their quality of

being diagnostic for each species and their hybrids. Ultimately, our final set consisted of 10 nuclear markers that could be used to distinguish pure species and their hybrids (see Results and Discussion). This number is a compromise between too few markers, which, on the one hand, could result in the misidentification of introgressed individuals as pure species types (Simovich & Sassaman 1986), and too many markers, which, on the other hand, could make genotyping large numbers of individuals prohibitively expensive in terms of both money and time.

Materials and methods

Tissue samples were obtained from lab-reared and field-caught adult and tadpole specimens (both fresh and preserved in ethanol). Samples from tadpoles consisted of ∼14.5 mg tail tissue, whereas tissue from adults consisted of ∼7.3 mg of tissue from a toe clip.

To identify markers that differentiated the two species, we initially used samples from allopatric populations where no introgression between the species has occurred. For S. multiplicata, we drew samples from populations in western Arizona outside of S. bombifrons' range. For S. bombifrons, we drew samples from populations in Colorado outside of S. multiplicata's range. Markers that were potentially diagnostic of the two species were then tested using tissue from toads collected in sympatric populations from Texas, Arizona, and New Mexico. Moreover, we also tested the markers on known hybrids from our lab colony that had been generated from experimental crosses of pure-species parents (e.g., Pfennig & Simovich 2002; Pfennig *et al.* 2007).

DNA isolation was performed using the Qiagen kit spin-column protocol. Our only variation from the protocol was that we eluted the samples twice with 100 ul Buffer AE (instead of twice with 200 ul) to increase the final concentration of DNA. Doing so was particularly important with the tadpoles to maximize the resulting amount of DNA. Typically we recovered 60-100ug (300-350ng/uL in 200ul per sample) for the tadpoles whereas toeclips yielded 2.5ug (35ng/uL in 70uL per sample), but the concentrations varied depending on the quality of the tissue.

We used the Applied Biosystems AFLP® Plant Mapping Protocol to develop AFLP markers from three selective primer combinations (Table 1). PCR products from the selective amplifications were submitted for genotyping on an 3730xl Genetic Analyzer (Applied Biosystems) at the UNC-CH Genome Analysis Facility. Amplified fragments between 50 and 500 base pairs (bp) were scored based on an internal size standard (GeneScan500 ROX; Applied Biosystems Inc.) using GeneMarker software version 1.85 (SoftGenetics), which were then checked by visual inspection for the presence or absence of peaks. Only distinct peaks were scored as present, and the manual scoring procedure was repeated on a separate occasion to reduce any inconsistencies in scoring. Additionally, we repeated the entire process, from initial amplification to manual scoring, at least once for each sample to evaluate the repeatability of the AFLP markers. We used 24 individuals of each species from allopatric populations outside of the other species' range. Species-specific loci were identified as those that were fixed for one species in allopatry, and totally absent in the other. The ability to detect hybrids was verified with a sample of 12 known hybrids.

For PCR–RFLP development, we sequenced a collection of cDNAs isolated from S. bombifrons. The cDNAs were annotated by homology comparisons to the Xenopus genome using BLAST. From these cDNAs sequences, primers were developed to amplify fragments of 100 to 200 base pairs. These primers were then tested using S. multiplicata and S. bombifrons samples from the allopatric populations describe above. The set of initial primers was pruned to only those that amplified well and at the same temperature in both species. These PCR fragments were sequenced. Those fragments showing a single nucleotide polymorphism were then targeted for RFLP development. We confirmed that only one species PCR product would cut with that restriction enzyme and that the resulting fragments are clearly visible on an agarose gel (2%). Markers passing this initial filter were then tested on samples from sympatric populations and seven known hybrids from our lab colony (described above). Using this process, we ultimately identified 10 nuclear markers that reliably distinguished the two species and their hybrids (see Results and Discussion below). To verify that these markers were species-specific, we tested 38 *S. multiplicata* and 33 *S. bombifrons* samples from allopatric populations. Of these samples, 100% were scored as the appropriate species for all markers, indicating that no intraspecies polymorphisms had been missed. Moreover, to further validate that these makers reliably identified hybrids, we tested them with 28 known hybrids. These markers were then used to measure frequency of hybridization in sympatric populations from a region where hybridization has been previously described (see below).

As a byproduct of our PCR–RFLP development we also identified a number of potentially informative species-specific SNPs that were not targeted by restriction endonucleases (Table 2). In Table 3, we have identified those SNPs that are unlikely to be sequencing errors and resided in regions suitable for Taqman probe development (Kalinina *et al.* 1997; Vos *et al.* 1995).

For those nuclear markers that we developed in this study, we performed a search of the *S. bombifrons* and *S. multiplicata* genes targeted by the markers using xenbase.org. From here, we determined which scaffold each gene was on and, using information obtained from tropmap.biology.uh.edu, the linkage group was identified. Doing so allowed us to determine whether or not the markers were closely linked and therefore independent assays of species identity.

To develop a PCR-RFLP marker for cyt-b we used previously sequenced haplotypes for each species from the allopatric locations described above (GenBank accession nos. EU285613, EU285616, EU285617, EU285643; Rice and Pfennig 2008). We analyzed these haplotype sequences with the online NEBcutter v. 2.0 (Vincze *et al.* 2003) to choose a restriction enzyme that would cut the PCR product from only one species. We then used previously published primers developed for *Spea* (Rice & Pfennig 2008) to amplify the cyt-b fragment and confirm that only one species PCR product would cut with that restriction enzyme and that the resulting fragments were clearly visible on an agarose gel (2%). Finally, we tested samples from sympatric populations and known hybrids from our lab colony as described previously for the nuclear PCR-RFLP development.

Using the 10 nuclear markers and the one mitochondrial marker, we then successfully genotyped 39-93 tadpoles from each of 12 ponds following different breeding events (spadefoots breed explosively on a single night following a rainstorm). Of these breeding events, three occurred at the same pond site in three different years. All aggregations were found near Portal, Arizona, USA, and were at sites where introgression between *S. multiplicata* and *S. bombifrons* has been previously observed (Simovich 1985; Simovich & Sassaman 1986). Not all samples were successfully genotyped across the entire suite of markers (in some cases only one marker worked for a given sample), and some ponds exhibited higher failure rates than others, possibly due to the quality of sample preservation. Nevertheless, we were able to calculate the percent of tadpoles that exhibited introgressed genotypes. Where F1 hybrids were produced, we estimated their frequency.

Results and Discussion

AFLP markers

We identified 12 AFLP loci distinct to S. multiplicata and 16 AFLP loci distinct to S. bombifrons (Table 1). These markers were species-specific and potentially could be used to diagnose hybrids. Their utility, however, varied with the type of sample and quality of sample preservation. When using fresh tissue or well-preserved samples in ethanol, the AFLPs worked well. However, older or poorly preserved specimens that had low concentrations or degraded DNA often failed or provided mis-leading results. Thus, these markers were not useful for tracking historical patterns of introgression from older samples. Indeed, because of the variability in outcome, the use of AFLPs for diagnosing these species and their hybrids might best be restricted to fresh tissue.

PCR – RFLP markers

We identified 10 PCR–RFLP nuclear markers, that could distinguish both species and their hybrids (Tables 2 & 3). We tested these markers on known hybrids, and the markers reliably identified these known hybrids. We also noted a bias as to which species tended to harbor the allele with the restriction site. Spea bombifrons tended to harbor more “cut” alleles. We found that these 10 markers generally map to different scaffolds of the Xenopus genome (Supplemental Table). Although the mean scaffold size is only 76,000 bp in *Xenopus*, half the genome is in scaffolds of 1.56 megabases or more. Thus, our finding that the markers are on separate scaffolds indicates that our markers likely serve as independent identifiers of species identity.

SNP markers

As a byproduct of PCR – RFLP development, we identified a total of 28 potential SNP markers, which varied in their ability to distinguish pure-species and hybrid genotypes (Tables 2 & 3). Although only 10 of the nuclear markers proved useful for distinguishing the *Spea* species and their hybrids, 14 additional markers are suitable for TaqMan probes and will be useful for anyone studying the natural ecology of members of the genus *Spea* (Tables 3). In particular, these markers can be combined with other within-species markers to measure population structure and differentiation within either *S. multiplicata* or *S. bombifrons* (e.g., Rice *et al.* 2008; Rice & Pfennig 2010).

Measuring introgression in the field

When we applied the 10 RFLP markers along with our species-specific mitochondrial marker to genotyping individuals from natural populations, we found levels of introgression that were similar to previously published values using allozyme studies. In particular, we found that the frequency of introgressed tadpoles (i.e., individuals that were either identified as a hybrid at one or more markers or that showed mixed species assignment across markers) arising from 12 different breeding events ranged from 0%, in a pond where only *S. multiplicata* was present among the samples, to 51%. By comparison, previously published accounts from these same populations, using a smaller set of four allozyme markers to estimate introgression, reported frequencies of introgressed individuals ranging from 0.8% to 42.5% (see Table 1, pp. 82-83 in Simovich 1985).

Our finding of a higher upper range of introgression could be accounted for in two ways. First, our higher measure of introgression may reflect the additional number of markers at our disposal relative to that in previous studies. With fewer markers, complex backcrosses are more likely to be assigned as pure species. Thus, previous estimates of introgression using fewer markers may have been more conservative (Simovich & Sassaman 1986).

Second, and perhaps more critically, the higher rates of introgression may reflect genuinely higher rates of introgression in those ponds where hybridization has been observed. Indeed, a single site accounted for some of the highest rates of introgressed individuals. Following three separate breeding aggregations at this site (each in a different year), the frequency of introgressed tadpoles was 11%, 33% and 51%. Such variation would be generated by year-to-year variation in the types of adults present at the breeding aggregation. Interestingly, however, no F1 hybrids were detected at this site in the years sampled, suggesting that introgression stemmed from an historical hybridization event(s). In the absence of this site, our range of observed introgressed individuals was 0% to 30%, which is more similar to the range previously observed using allozymes (Simovich 1985).

When we looked specifically at the frequency of F1 hybrids, we found that F1 hybrids were relatively rare, and occurred in only one of the 12 ponds sampled. In the one pond where F1 hybrids did occur, however, F1 hybrids represented 4.3% of the tadpoles sampled at that site. This result is consistent with previously published findings showing that, although hybridization has declined between these two species (Pfennig 2003), facultative hybridization in any given year could generate “bursts” of hybridization that contribute to introgression between these species (Pfennig 2007). As indicated above, these bursts of hybridization could contribute to the on-going detection of complex hybrids, even in the absence of F1s in any given year.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to Sophia Shih, Sara Garnett, Caitlin Angell, Laura Exline, Christina Lebonville, and Max Lundberg for their assistance in the lab. We also thank Travis Glenn and two anonymous reviewers for comments

that greatly improved the manuscript. This work was supported by a NSF grant (NSF DEB 0542566) and a NIH Office of the Director's New Innovator award (NIH 1DP2OD004436-01) to K. S. P. and grants from the NIH (NIH RC2 GM092501) and NSF (NSF MCB 0920196) to C. D. J.

Literature Cited

Arendt JD. Influence of sprint speed and body size on predator avoidance in New Mexican spadefoot toads (*Spea multiplicata*). Oecologia. 2009; 159:455–461. [PubMed: 18987891]

Banbury B, Maglia AM. Skeletal development of the Mexican spadefoot, *Spea multiplicata* (Anura : Pelobatidae). Journal of Morphology. 2006; 267:803–821. [PubMed: 16572407]

Forester DC. Mating call as a reproductive isolating mechanism between *Scaphiopus bombifrons* and *S. hammondii*. Copeia. 1973; 1973:60–67.

Forester DC. Laboratory evidence for potential gene flow between two species of spadefoot toads, *Scaphiopus bombifrons* and *Scaphiopus hammondii*. Herpetologica. 1975; 31:282–286.

García-Paris M, Buchholz DR, Para-Olea G. Phylogenetic relationships of Pelobatidae re-examined using mtDNA. Molecular Phylogenetics and Evolution. 2003; 28:12–23. [PubMed: 12801468]

Gray MJ, Smith LM, Brenes R. Effects of agricultural cultivation on demographicsof Southern High Plains amphibians. Conservation Biology. 2004; 18:1368–1377.

Kalinina O, Lebedeva I, Brown J, Silver J. Nanoliter scale PCR with TaqMan detection. Nucleic acids research. 1997; 25:1999–2004. [PubMed: 9115368]

Ledon-Rettig CC, Pfennig DW. Emerging model systems in eco-evo-devo: the environmentally responsive spadefoot toad. Evolution & Development. 2011; 13:391–400. [PubMed: 21740512]

Martin RA, Pfennig DW. Disruptive selection in natural populations: the roles of ecological specialization and resource competition. American Naturalist. 2009; 174:268–281.

McMurry ST, Smith LM, Dupler KD, Gutierrez MB. Influence of Land Use on Body Size and Splenic Cellularity in Wetland Breeding Spea spp. Journal Of Herpetology. 2009; 43:421–430.

Pfennig DW, Rice AM, Martin RA. Ecological opportunity and phenotypic plasticity interact to promote character displacement and species coexistence. Ecology. 2006; 87:769–779. [PubMed: 16602305]

Pfennig KS. Female spadefoot toads compromise on mate quality to ensure conspecific matings. Behavioral Ecology. 2000; 11:220–227.

Pfennig KS. A test of alternative hypotheses for the evolution of reproductive isolation between spadefoot toads: support for the reinforcement hypothesis. Evolution. 2003; 57:2842–2851. [PubMed: 14761062]

Pfennig, KS. Science. Vol. 318. Washington, D. C.: 2007. Facultative mate choice drives adaptive hybridization; p. 965-967.

Pfennig KS, Chunco AJ, Lackey ACR. Ecological selection and hybrid fitness: hybrids succeed on parental resources. Evolutionary Ecology Research. 2007; 9:341–354.

Pfennig KS, Simovich MA. Differential selection to avoid hybridization in two toad species. Evolution. 2002; 56:1840–1848. [PubMed: 12389729]

Pierce JR. Distribution of two mating call types of the plains spadefoot, *Scaphiopus bombifrons*. Southwestern Naturalist. 1976; 20:578–582.

Rice AM, Pearse DE, Becker T, et al. Development and characterization of nine polymorphic microsatellite markers for Mexican spadefoot toads (*Spea multiplicata*) with cross-amplification in Plains spadefoot toads (*S. bombifrons*). Molecular Ecology Resources. 2008; 8:1386–1389. [PubMed: 21586053]

Rice AM, Pfennig DW. Analysis of range expansion in two species undergoing character displacement: why might invaders generally “win” during character displacement? Journal of Evolutionary Biology. 2008; 21:696–704. [PubMed: 18341542]

Rice AM, Pfennig DW. Does character displacement initiate speciation? Evidence of reduced gene flow between populations experiencing divergent selection. Journal of Evolutionary Biology. 2010

Sattler PW. Introgressive hybridization between the spadefoot toads *Scaphiopus bombifrons* and *Scaphiopus multiplicatus* (Salientia: Pelobatidae). Copeia. 1985; 1985:324–332.

Simovich, MA. Ph D. University of California; 1985. Analysis of a hybrid zone between the spadefoot toads Scaphiopus multiplicatus and Scaphiopus bombifrons.

Simovich MA, Sassaman CA. Four independent electrophoretic markers in spadefoot toads. Journal of Heredity. 1986; 77:410–414. [PubMed: 3559167]

Stebbins, RC. A field guide to western reptiles and amphibians. Houghton Mifflin; Boston, MA: 2003.

Vincze T, Posfai J, Roberts RJ. NEBcutter: a program to cleave DNA with restriction enzymes. Nucleic acids research. 2003; 31:3688–3691. [PubMed: 12824395]

Vos P, Hogers R, Bleeker M, et al. AFLP - a new technique for DNA-fingerprinting. Nucleic acids research. 1995; 23:4407–4414. [PubMed: 7501463]

Wiens JJ, Titus TA. A phylogenetic analysis of Spea (Anura: Pelobatidae). Herpetologica. 1991; 47:21–28.

Table 1

Summary of species-specific AFLP markers from three selective primer combinations. Diagnostic markers of *S. multiplicata* (Sm, in table) or *S. bombifrons* (Sb, in table) were identified as loci fixed in allopatric populations of one species, and absent in the other.

Primer combination EcoRI/MseI	Sm fragment size (bp)	Sb fragment size (bp)
aac/cac	63	104
	121	108
	160	147
	247	153
	410	157
	—	276
	—	484
aag/caa	93	68
	107	84
	173	149
	284	163
	300	182
	—	325
	—	489
acg/cag	158	214
	189	496
Total species specific loci	12	16

Table 2

Summary of PCR-RFLP and SNP markers for *S. multiplicata* and *S. bombifrons*. Of these markers, a subset reliably distinguishes both species and their hybrids (see Table 3).

Gene name (abbreviation)	Forward primer (5′ -> 3′)	Reverse primer (5′ -> 3′)	Restriction enzyme	*S. multiplicata* eequence (5′ -> 3′)	*GenBank accession number*[2]	*S. bombifrons* sequence (5′ -> 3′)	*GenBank accession number*[2]
hypothetical Loc496414	GCATTAC ACGAAAA GCAAACC	CGGAAGG CCAACTC ATGTAT	BtsCI	CATG__CTT*GGATGAA*C N		CATGCACTTG**AAG**GAAC A	
hypothetical Loc496744	GTTCCAA GCGTGAC TTGGTT	ATTGCCA TCTCCAT CAGACC	TaqI	TTCCT*TCG*AGACACGTG AGTT	JQ707931	TTCCTTC**A**AGACACGTG AGTTT	JQ707932
creatine kinase, muscle (ckm)	TAACCTC CGTGACG GTCTTC	AAGTGGC TTCACCC TTGATG	N/A[1]	GTTGAGTTT___G**T**C**A**G TGCTA**G**_____ __GATG_____ _____ _____ _____GTGTA**C**TA ATACGC	JQ707917	GTTGAGTTTTCAG**A**C**T**G TGCTA**A**GACGCATTGTT AGGATGGAATCCTTAGA TAGTTCTTTCTTTGGAA CGGCATGGTGGGAAATA AGAGGACAGGTGTA**G**TA ATACGC	JQ707918
DEAD box polypeptide 6 (ddx6)	CTGCGTG CTCAGAT AGTTGC	TGGACTC CAAAATG CAATGA	Hpy188I	TGTTCA**A**AAGACAAGCC AAGCAGTTTTTTTT	JQ707921	TGT*TCAG*AAGACAAGCC AAGCAGTTTTTTTT	JQ707922
desmin, gene 1 (des.1)	CAAAGCC ATCTTGA CGTTGA	AGGGACC TGGAGGA GAGGTA	CviQI	GGCAGCTTCCCCGCC**A**T ACCTCTCCTCCAG		GGCAGCTTCCCCGCC*GT* *A*CCTCTCCTCCAG	
dual specificity phosphatase 22 (dusp22)	GCTGGGA GTTCTGG AAGACA	GTGCCAA TCACCAA CTTTGA	BfuCI	CTGACACTCTACAAA**T**A TCACCGGTCAAAG		CTGACACTCTACAAA*GA* *TC*ACCGGTCAAAG	
glia maturation factor gamma (gmfg)	GATCAGG AACACCA GCGACT	CTGACCC ACAGAGC GTACAA	BstUI	ATGGCTGCG**T**GGATTAA CCCCACAAA		ATGGCTG*CGCG*GATTAA CCCCACAAA	
hydroxyacyl-CoA dehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoA hydratase (trifunctional protein), alpha subunit (hadha)	TGGTAGC CAAAGGG TTTTTG	GACGCAT CTGGATG TCTTCA	BtsCI	TACTGA_____G CCTAGGGTTCT	JQ707929	TACTG**G**ACTGTC*GGATG* *CC*TAGGGTTCT	JQ707930
SRY_box 2 (sox2)	CAGCTCC CAGACGT ACATGA	TGGTACT TCTGCCC CAGGTA	AvaI	ACCTC**A**TCTTCCCACTC **G**CGGGCTCCATGT		ACCTC**T**TCTTCCCACT*C* ***T****CGGG*CTCCATGT	
cytochrome c, testis_specific protein (cyct)	TATTCCC GGAACCA AAATGA	AGGGACC GACACAA GTGAAC	HpyCH4IV	AAAGCCAC**A**TCGGTAAT GATACA		AAAGCC*ACGT*CGGCAAT GATACA	

Gene name (abbreviation)	Forward primer (5′ -> 3′)	Reverse primer (5′ -> 3′)	Restriction enzyme	*S. multiplicata* sequence (5′ -> 3′)	*GenBank accession number*[2]	*S. bombifrons* sequence (5′ -> 3′)	*GenBank accession number*[2]
core promoter element binding protein (copeb)	GATCGAT GGCTGTC CCAATA	GTGCTAC ACCTGGC GTCTTC	Hpy188I	ACGAGCCGTTTTTGAAGA TAAAGCACCACTAGTGC CGGGG	JQ707915	ACGAGCCGT*TCAG*AAGA TAAAGCACCACTAGTGC CGGGG	JQ707916
BCL2_associated transcription factor 1 (bclaf)	TTTCCTG CACGATG ACAGAG	AGGCCAT CTCCTTG GAACTT	Hpy166II	GGGGGCGTTTTACCTTT AAAAAATCTGGAAGCAG CCCAAAATGGACACATG ATA		GGGGGCG*GTTTA*CCTTT AAAAAATCTGGAAGCAG CCCAAAATGGACACA**T**G ATA	
ets domain transcription factor (elf1)	CCGAATC AACTAGC CCTGAA	TTCACAT GGAACCC TTCTCC	BtgI	ACGGCCGCT**G**CGGTG**G**G CG**C**GNAAGCCCCG	JQ707923	ACGGCCGC**T**GCGGTG**G**G CG**C**GGAAGCCCCG	JQ707924
GDP_mannose pyrophosphorylase A (gmppa)	TTAGGAT TGGGGTC ACTTGG	TATTGGA GCTGGGG TGACAG	BspDI	TTGCTCTGTTGGCTGCA ***T****CGATG*GGACTGCT		TTGCTCTGGTGGCTGCA **G**CGATGGGACTGCT	
glia maturation factor gamma (gmfg)	GATCAGG AACACCA GCGACT	CTGACCC ACAGAGC GTACAA	HpaII	TT**C**ACTGAA…C**G**T**G**GA… ATTCACCCTTTAGTG*C* ***CG****G*AAGGGTTAAATAC**C** CTG	JQ707925	TT**A**ACTGAA…C**G**C**G**GA… ATTCACCCTT**G**TAGTGC CA**G**AAGGGTTAAATAC**A** CTG	JQ707926
ATG2 autophagy related 2 homolog B (atg2b)	ACCATCC CATGCAT ACAGGT	TCCATGC AAACTGT CTGAGC	HinfI	GTGTGAGCCAAGCTG**AA** GTCACCTGTCATG		GTGTGAGCCAAGCTG***GA*** *GTC*ACCTGTCATG	
GDP_mannose pyrophosphorylase A (gmppa)	TTAGGAT TGGGGTC ACTTGG	TATTGGA GCTGGGG TGAGAG	Taq I	TGGCTGCA*T*C*G*ATGGGA CTGCT		TGGCTGCA***G***CGATGGGA CTGCT	
Glyceraldehyde-3-phosphate dehydrogenase (gapdh)	GTTGGTG TGAACCA CGAGAA	CTGTGAA AGCGTGG ACAGTG	Hpy188I	TTGGTC*TCT__GA*GACT TGGCTG**A**GATTAAAGCT TTTT**G**GTT	JQ707927	TTGGTCTCTTTGAGACT TGGCTG**G**GATTAAAGCT TTTT**T**GTT	JQ707928
hmgb2;high-mobility group box 2 (hmgb2)	CTCGAGT GCAGCTC ATTTTG	AATGCCA GGCTACC CTTAGAA	MseI	AGT***T****TA*AGAGCTTATGT GGGATTTCTCTGCATAT *TTAA*TGG		AGT**GT**AAGAGCTTATGT GGGATTTCTCTGCATAT TTAG**T**GG	
hypothetical Loc496690	GGCTTGG TGTACGC TCTCTC	GAAGGCT CTGCAGG ACTCTG	DdeI	ATCTTCCACACCATCGC CTACC*TGAG****T****C*CACTTC CGC		ATCTTCCACACCATCGC CTACCTGA**CT**CC**G**CTTC CGC	
acidic ribosomal phosphoprotein P0 (arbp/ rplp0)	TGGAAGC ACTGACA AGATGC	CAGGTGA CCACGGA TAGCTT	HpyCH4IV	AGTTTTGGGGATTTGTG AGAGTTG**T**CGTTG**TT**	JQ707909	AGTTTTGGGGATTTGTG AGAGTTG***A****CGTTG***C**N	
aldolase A, fructose_bisphosphate (aldoa)	TCCACGA GACCCTC TACCAC	CAGACCC TGAGTGG TGGTTT	HaeIII	GGGCGGC**G**CTGCTGCTN …CGANAAAGGTGTCN TCCNNCTGGC**T**GN		GGGC*GGC****C****C*TGCTGCTT …CGACAAAGGTGTCG TCCCCCT*GGC*CGG	

Gene name (abbreviation)	Forward primer (5′ -> 3′)	Reverse primer (5′ -> 3′)	Restriction enzyme	*S. multiplicata* eequence (5′ -> 3′)	*GenBank accession number*[2]	*S. bombifrons* sequence (5′ -> 3′)	*GenBank accession number*[2]
ATG2 autophagy related 2 homolog B (atg2b)	ACCATCC CATGCAT ACAGGT	TCCATGC AAACTGT CTGAGC	TaqI	TCATGTA*TCG*AAAGCCC TTATGTACTTTTATA		TCATGTAT**T**GAAAGTCC TCATGTACTTTTATA	JQ707910
cold inducible RNA binding protein (cirbp)	GCCTCAG CTTTGAA ACCAAC	CACCGGA ACCTCCT CTGTAG	SspI	AACCATTTA**T***AATATT*C TAA**A**ATGTGAT**G**TGCAG C**A**TCTGCATTATTT	JQ707911	AACCATTTA**C**A**T**T**G**TTC TAA**C**ATGTGAT**A**TGCAG C**C**TCTGCATTATTT	JQ707912
COP9 constitutive photomorphogenic homolog subunit 3 (cops3)	AATAACA TGGGCCT GGTGAA	CCATCCC GTCCTTC TGATTA	HinfI	GTAAGAC**C**CAGCAATAG AGTGAGGCTTA***GA****C**TC*C TCCAGG**G**G**A**CGGTCAA TTC**G**GGGGGTCTCGTGC T**T**CCCCGTGGCTT	JQ707913	GTAAGAC**G**CANCAATAG AGTGAGGCTTA**T**C**T**CCTCC TCCGAGG**T**G**C**CGGTCAA TTC**A**GGGGGTCTCGTGC T**C**CCCC_TGGGTT	JQ707914
D_dopachrome tautomerase (ddt)	CCGCGAT AGGATAA CGCTAA	CCAGGCC TCAATAG GAATGA	HindIII	TGGAA**A**CTTTC**CA**GAGC CGAGAATAG	JQ707919	TGG*AA**G**CTT*TCTGGAGC CGAGAATAG	JQ707920
GDP_mannose pyrophosphorylase A (gmppa)	TTAGGAT TGGGGTC ACTTGG	TATTGGA GCTGGGG TGAGAG	Hpy166II	GATCCTAG**C**GAACAAGA CGAGTCGGTGTGTAA		GATCCTA*G**T**GAAC*AAGA CGAGTCGGTGTGTAA	
huntingtin (htt)	AACTGGG TTCTGGG AAAGGT	TTTGCTG TCCCCAC ACAGTA	AvaII/ Sau96I	CTC_____ _____CTCTTTCTAT		CTCTCCCATATAAAGCA AA*GGTC*CTCTTTCTAT	
mitochondrially encoded cytochrome B (cytb)	CAATAGC ATTCTCT TCAGT	GGGGGTT ACTAGGG GGTTTG	Acc1	GCTTCTCAGTG**GA**TAAT GCCACA**T**TAACCCGTTT		GCTTCTCAG*T**AGAC**C*AAT GCCACACTAACCCGCTT	

[1]PCR product length varies (≥50 bp), so size differences can be visualized without restriction enzyme

[2]If applicable; see supplemental file for additional sequence information

Restriction Enzyme Target Site = *italicized*

Single Nucleotide Polymorphism= **Bold**

Insertions and Deletions = Underline and underscore___

Table 3

Summary of annealing temperature, band sizes, and quality for PCR-RFLP and SNP markers for *S. multiplicata* (Sm) and *S. bombifrons* (Sb). Only those markers with a grade of "A" were reliable for distinguishing between the two species and their hybrids.

Gene name(abbreviation)	Annealing temp (°C)	Sm band size(s) (bp)	Sb band size(s) (bp)	Hybrid band size(s) (bp)	Quality[1]
hypothetical Loc496414	60	100	150	100, 150	A
hypothetical Loc496744	60	100, 125	200	100, 125, 200	A
creatine kinase, muscle(ckm)	65	370	420	420, 370	A
DEAD box polypeptide 6(ddx6)	56	180, 70	100, 80, 70	180, 80, 70	A
desmin, gene 1(des.1)	56	145	105, 40	145, 105, 40	A
dual specificity phosphatase 22 (dusp22)	59	155	125, 30	155, 125, 30	A
glia maturation factor gamma(gmfg)	62	285	160, 125	285, 160, 125	A
hadha2; hydroxyacyl-Coenzyme A dehydrogenase/3-ketoacyl-Coenzyme A thiolase/ enoyl-Coenzyme A hydratase (trifunctional protein), alpha subunit (hadha), nuclear gene encoding mitochondrial protein(hadha)	60	400	300	400,300	A
SRY-box 2(sox2)	65	210	150, 60	210, 150, 60	A
cytochrome c, testis-specific protein(cyct)	58	150	100, 50	150, 100, 50	A
core promoter element binding protein(copeb)*	62	250	200, 50	250, 200, 50	B
BCL2-associated transcription factor 1(bclaf)*	62	170	100, 70	170, 100, 70	B
ets domain transcription factor(elf1)*	62	115, 100	215	215, 115, 100	B
GDP-mannose pyrophosphorylase A(gmppa)*	60	140, 110	250	250, 140, 110	B
glia maturation factor gamma(gmfg)*	62	215, 70	285	285, 215, 70	B
ATG2 autophagy related 2 homolog B(atg2b)	62	235, 40	200, 75	235, 200, 75, 40	C
GDP-mannose pyrophosphorylase A(gmppa)	60	140, 110	250	250, 140, 100	C
glyceraldehyde-3-phosphate dehydrogenase(gapdh)*	62	220, 110	330	330, 220, 110	C
hmgb2;high-mobility group box 2(hmgb2)*	60	100,50	100		C
hypothetical Loc496690*	60	105	230	105, 230	C
acidic ribosomal phosphoprotein P0(arbp/ rplp0)*	65	215, 100	115, 100, 50	215, 115, 100, 50	D
aldolase A, fructose-bisphosphate(aldoa)	65	250, 50	150, 100, 50	250, 150, 100, 50	D
ATG2 autophagy related 2 homolog B(atg2b)*	62	180, 85	265	265, 180, 85	D
cold inducible RNA binding protein(cirbp)*	65	370, 130	500	500, 370, 130	D

Gene name(abbreviation)	Annealing temp (°C)	Sm band size(s) (bp)	Sb band size(s) (bp)	Hybrid band size(s) (bp)	Quality[1]
COP9 constitutive photomorphogenic homolog subunit 3(cops3)*	56	520, 72	615	615, 520, 72	D
D-dopachrome tautomerase(ddt)	56	400	240, 160	400, 240, 160	D
GDP-mannose pyrophosphorylase A(gmppa)*	60	250	175, 75	250, 175, 75	D
huntingtin (htt)*	60	250	190	250, 190	D
mitochondrially encoded cytochrome B(cytb)	53.5	750	400, 350	*mitochondrial	N/A

[1]Markers were assigned a letter grade indicating quality of marker to discriminate the two species and/or their hybrids as follows:

"A" = reliably discriminates species and their hybrids

"B" = reliably discriminates species, but incomplete digestion makes diagnosis of hybrids difficult

"C" = reliably discriminates species, but fails to discriminate hybrids from pure species types

"D" = does not reliably distinguish between species or their hybrids

*Fragments contain SNPs suitable for TaqMan probes.