MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2012) **12**, 354–362 doi: 10.1111/j.1755-0998.2011.03087.x

A single nucleotide polymorphism assay for the identification of unisexual *Ambystoma* Salamanders

KATHERINE R. GREENWALD*† and H. LISLE GIBBS*
*Department of Evolution, Ecology and Organismal Biology, Ohio State University, 300 Aronoff Laboratory, Columbus, OH 43210, USA, †Department of Biology, Eastern Michigan University, 316 Mark Jefferson, Ypsilanti, MI 48197, USA

Abstract

Unisexual (all female) salamanders in the genus *Ambystoma* are animals of variable ploidy (2N-5N) that reproduce via a unique system of 'leaky' gynogenesis. As a result, these salamanders have a diverse array of nuclear genome combinations from up to five sexual species: the blue-spotted (*A. laterale*), Jefferson (*A. jeffersonianum*), smallmouth (*A. texanum*), tiger (*A. tigrinum*) and streamside (*A. barbouri*) salamanders. Identifying the genome complement, or biotype, is a critical first step in addressing a broad range of ecological and evolutionary questions about these salamanders. Previous work relied upon genome-related differences in allele size distributions for specific microsatellite loci, but overlap in these distributions among different genomes makes definitive identification and ploidy determination in unisexuals difficult or impossible. Here, we develop the first single nucleotide polymorphism assay for the identification of unisexual biotypes, based on species-specific nucleotide polymorphisms in noncoding DNA loci. Tests with simulated and natural unisexual DNA samples show that this method can accurately identify genome complement and estimate ploidy, making this a valuable tool for assessing the genome composition of unisexual samples.

Keywords: *Ambystoma*, polyploidy, single nucleotide polymorphism, species identification, unisexual

Received 14 February 2011; revision received 29 September 2011; accepted 10 October 2011

Introduction

Unisexual (all female) populations of ambystomatid salamanders are widely distributed across northeastern North America (Bogart & Klemens 1997; Bogart *et al.* 2007) and represent the oldest known unisexual vertebrate lineage (Bi & Bogart 2010). These salamanders are generally triploid, but can be diploid, tetraploid and even, rarely, pentaploid (Bogart *et al.* 2007). Unisexual females require sperm from a co-occurring sexual male to initiate reproduction. They can then reproduce either gynogenetically (without sperm genome incorporation) or sexually, by incorporating the sperm genome into the resulting zygotes via genome replacement or ploidy elevation (Bogart *et al.* 2007; Ramsden 2008). Males from any of the five sexual species can act as the 'sperm donor': the blue-spotted salamander (*A. laterale*), Jefferson salamander (*A. jeffersonianum*), smallmouth salamander (*A. texanum*), tiger salamander (*A. tigrinum*) and streamside salamander (*A. barbouri*). These genomes, when found in the unisexual salamanders, are abbreviated as L, J, T, Ti and B, respectively (Lowcock *et al.* 1987; Bogart *et al.* 2009). This results in a large number of possible biotypes (genome combinations) in the unisexuals across their range (e.g. LJJ, LLJ, LJJJ, LTTT and LTTi).

Correspondence: Katherine R. Greenwald, Fax: 734 487 9235; E-mail: katherine.greenwald@emich.edu

Distinguishing unisexual salamanders from sympatric sexual species is straightforward using mitochondrial DNA. Because the complex is virtually entirely female and mtDNA is maternally inherited, unisexual mtDNA is clearly distinguishable from that of the sexual species involved in the complex (Hedges *et al.* 1992; Noël *et al.* 2008; Rhoads *et al.* 2009; Bi & Bogart 2010). In particular, fixed substitutions in the mitochondrial D-loop and intergenic spacer region distinguish unisexual salamanders from sexual species, as well as clearly separating all five sexual species involved in the complex (Shaffer & McKnight 1996; McKnight & Shaffer 1997; Bogart *et al.* 2007, 2009).

Although it is straightforward to distinguish sexual species from the unisexual complex as a whole using mtDNA, diagnosing the nuclear genome composition and ploidy (biotype) of the unisexuals is more difficult. Historically, allozymes were used but required sacrificing animals and thus were impractical for ecological studies requiring large sample sizes of genotyped individuals (Bogart *et al.* 1987; Bogart & Klemens 1997). More recently, researchers have used

fluorescent cytogenetic techniques (genomic in situ hybridization, GISH; Bi & Bogart 2006; Bogart *et al.* 2009) and microsatellite loci for genome identification in unisexual salamanders. While GISH results are clear and easily interpreted, the technique is too labour-intensive and costly for large-scale ecological studies. Microsatellite analysis based on loci isolated from sexual species potentially provides a more rapid and cost-effective way to determine unisexual biotype (Ramsden *et al.* 2006; Bogart *et al.* 2007). The approach is to identify a suite of loci with differential amplification of the sexual species' genomes (either failing to amplify in certain species or producing different allele size ranges in different species). It works well when a limited number of sexual species are present and involved in the complex as, for example, in northern and eastern parts of the unisexuals' range where only *A. laterale* and *A. jeffersonianum* act as sperm donors. However, in the southern and western parts of the range (e.g. Ohio, Indiana), all five potential sperm donor species may be found in close proximity. In these regions, the microsatellites currently available amplify overlapping allele size ranges in the sexual species, making it difficult or impossible to separate all potential genomes (Bogart *et al.* 2009; K. R. Greenwald and H. L. Gibbs, unpublished results). Therefore, for this particular application, the intraspecific variability of microsatellite loci is in fact a liability. The ideal assay would include markers that are variable among species, but fixed within species.

Here, we present a novel method of determining unisexual biotype, using the SNaPshot assay (Applied Biosystems, Foster City, CA), in which species-specific single nucleotide polymorphisms (SNPs) are interrogated with fluorescently labelled probes (Norton *et al.* 2002; Hurst *et al.* 2009). This approach circumvents many of the issues encountered in microsatellite analysis, as intraspecific variation is virtually absent and interpretation of results is straightforward. Presence/absence of particular SNPs determines which sexual species' genomes are present in the unisexual salamanders (genome complement). Importantly, ploidy can be estimated from the ratio of peak areas, which we show are proportional to the relative quantity of DNA present. The SNaPshot assay is sensitive to detection of DNA variants down to 5% of total DNA present (Hurst *et al.* 2009), much lower than the minimum quantity that is biologically possible in this system (20%, for one unique genome out of five present in a pentaploid). Based on our results described below, we conclude that this assay is a rapid and reliable method for identifying and differentiating sexual species and unisexual biotypes present in this complicated system.

Materials and methods

Samples and locus selection

To identify loci with nucleotide polymorphisms that were fixed within species but variable among species, we screened nuclear loci (see below) using a panel of 57 geographically diverse samples from sexual species whose genomes are found in the unisexual lineage. Our focus is on developing assays for unisexuals found in the Midwestern part of the unisexual range and so our reference samples for sexual species generally but not exclusively come from this region. These samples included: *A. jeffersonianum* from Indiana, Ohio and Pennsylvania ($n = 11$); *A. laterale* from Indiana, Minnesota, Ohio and Québec ($n = 12$); *A. tigrinum* from Illinois, Indiana and Ohio ($n = 10$); *A. texanum* from Indiana, Ohio, Nebraska and West Virginia ($n = 14$); and *A. barbouri* from Indiana, Ohio and West Virginia ($n = 10$). Samples from the same state were selected to be as geographically diverse as possible; for example, *A. jeffersonianum* from Ohio are from four different sites in southwest, southeast and northeast Ohio. A number of these samples were provided to us as extracted DNA (J. Bogart: PA *A. jeffersonianum*, QU *A. laterale*, IN *A. texanum*; W. K. Savage: MN *A. laterale*). For the remaining samples, we extracted DNA from tail-tissue samples using Qiagen DNeasy kits (Qiagen, Valencia, CA). All reference samples were confirmed as belonging to the sexual species by mtDNA sequencing (Primers F-THR and R-651, which amplify the D-loop, intergenic spacer region, $tRNA^{Pro}$, $tRNA^{Phe}$ and part of $tRNA^{Thr}$; Bogart *et al.* 2007; Shaffer & McKnight 1996).

We screened 10 anonymous nuclear loci (E05A08, E07C01, E12C11, E12G03, E12G10, E13E02, E14G07, E20C02, E21E04 and E25G07) generated from the Salamander Genome Project (Smith *et al.* 2005; http://www.ambystoma.org), as well as three coding genes (RAG1, BDNF and POMC; Vietes *et al.* 2006), with the goal of identifying diagnostic SNPs. Most loci either failed to cleanly amplify in all species or showed no fixed substitutions separating reference samples from particular species. In the end, we focused on two loci (E14G07 and E12G03) for designing useful SNP probes. E14G07 amplified cleanly for all five sexual species except for a subset of *A. tigrinum* samples (discussed below), and had unique combinations of polymorphisms for each genome type within 350 bp (Table 1; Fig. 1). Following discovery of this amplification issue with some *A. tigrinum* sequences, we developed a single SNP on locus E12G03 to act as a diagnostic Ti SNP for cases in which E14G07 failed to amplify correctly in samples including Ti genomes.

Table 1 A unique combination of SNPs on loci E14G07 (TBTi172, T27 and L313) and E12G03 (Ti92) identifies genomes from each sexual species. Intraspecific polymorphisms are indicated by slashes (e.g. C/T indicates that C and T are both present within the species)

Species	TBTi172	T27	L313	Ti92	Combination
Ambystoma jeffersonianum	A	T	T	C	ATTC
Ambystoma laterale	A	T	A	C	ATAC
Ambystoma texanum	G	C/T*	T	C	G(C/T)TC
Ambystoma barbouri	G	T/C*	T	C	G(T/C)TC
Ambystoma tigrinum	G/T	T	C	G	(G/T)TCG

SNP, single nucleotide polymorphism.
*This SNP appears to be diagnostic in the majority of the geographic range we assayed, with *A. texanum* fixed for a C and *A. barbouri* fixed for a T. However, we obtained the opposite result in individual *A. texanum* from Nebraska and *A. barbouri* from northcentral Ohio, and so it is not useful for distinguishing between T and B genomes in these regions (see text).

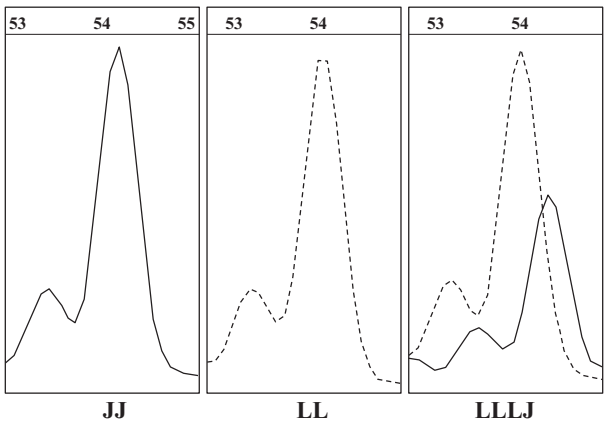

Fig. 1 An example of SNaPshot results from single nucleotide polymorphism (SNP) locus L313, in which *Ambystoma jeffersonianum* have a thymine (T, solid line), and *Ambystoma laterale* have an adenine (A, dashed line). The panel at right shows a sample identified by Dr. James Bogart as LLLJ using microsatellites; our SNP assay corroborates his results once peak areas are adjusted using the appropriate regression equation (JPB#32255, Table 4).

SNaPshot probe design

Probe design for the SNaPshot analysis is similar to primer design for conventional PCR applications, but with additional considerations. The assay is based on a single nucleotide primer extension using a fluorescently-tagged ddNTP for sites on an amplicon that show species-specific polymorphism. As such, the end of a probe must be immediately adjacent to the SNP which means that when optimizing primers only length and directionality can be varied. We selected primers after using OLIGO Primer Analysis Software (Molecular Biology Insights, Inc., Cascade, CO) to assess issues such as melting temperature compatibility and formation of primer dimers or hairpins. We initially designed probes of 18–22 bp in length; however, to multiplex the SNPs, the probes must each be separated by 7–8 bp in length so that the fluorescent peaks can be individually interpreted. To achieve this separation, we added variable-length poly-T tails to the 5′ end of the probes (Hurst *et al.* 2009) so that their final lengths were 18, 32 and 46 bp long (Table 2).

PCR amplification and SNaPshot processing

For the initial amplification of both loci, we used a 10 μL PCR mix that contained the following: 5 μL of 2× BioMix Red enzyme and buffer mix (Bioline USA Inc., Taunton, MA); 2.5 μL ddH_20; and 1 μL of a 10-mM stock for each of the forward and reverse primers:

E14G07 F: AATTATATTCCCTTTCCATGTCTGTC
E14G07 R: TGCAGAAATATTTACGATTCTAGCAC
E12G03 F: TAAATACAGAATGTGTTCGACTGTCC
E12G03 R: GAGTGCATTACATGGTATACTGCTG

We used a PCR protocol consisting of initial denaturation of 3 min at 94 °C; followed by 35 cycles of 20 s at 94 °C, 20 s at 53 °C, 1 min at 72 °C and a final extension of 5 min at 72 °C on an Eppendorf MasterCycler 5341. We followed the initial PCR amplification with an ExoSAP-IT (Affymetrix, Santa Clara, CA) clean-up to remove unincorporated dNTPs and extra primers, as recommended in the SNaPshot protocol. We used 2 μL ExoSAP-IT per 5 μL PCR volume, incubated the mixture at 37 °C for 15 min and then heated the solution at 80 °C for 15 min to denature the enzyme.

PCR amplification of SNaPshot probes followed the protocol presented in the Applied Biosystems manual. Briefly, for each reaction, we used 5 μL SNaPshot reaction mix, 1 μL pooled 1 μM primer mix, 1 μL ddH_20 and 3 μL PCR product from the initial amplification. The reaction was then cycled 25 times for 10 s at 96 °C, 5 s at 50 °C and 30 s at 60 °C. We used a final SAP clean-up step, adding 1 μL SAP per 10 μL PCR volume, incubating at 37 °C for 60 min and then heating at 65 °C for 15 min. Samples were electrophoresed on an ABI 3370 using GeneScan-120 LIZ size standard by the Ohio State University Plant-Microbe Genome Facility. We used GeneMarker software (SoftGenetics, State College, PA) to quantify the size and area of each SNP peak.

Tests with simulated unisexuals

We tested whether the SNaPshot assay was able to accurately assess ploidy level in unisexual salamanders with

Table 2 SNP loci and sequences

SNP	Locus	Sequence	Length	SNP
TBTi172	E14G07	CTAAAGACACCCAGTCTC	18	A/G/T
T27	E14G07	$(T)_{17}$GAGAATGAGCAAGAA	32	C/T
L313	E14G07	$(T)_{26}$CAGGTTAGGACTGAGGGGAA	46	A/C/T
Ti92	E12G03	TCTGACCCAACAGGCAGACG	20	C/G

SNP, single nucleotide polymorphism.

Table 3 Expected and observed peak area ratios for simulated unisexual biotypes. Expected ratios were calculated *a priori* based on the information from Table 1; for example, for TBTi172 *Ambystoma laterale* has an A and *Ambystoma texanum* has a G, so the peak ratio for an LT unisexual should be 1:1 while the ratio for an LTT should be 1:2. Observed ratios are from simulated biotypes made by mixing known quantities of DNA. The relationship between expected and observed values was used to generate regression equations for analysis of unknown samples

SNP	Ploidy type	Biotypes	Expected ratio	Observed ratio (Mean ± SE)
TBTi172	Diploids	LJ	Undefined	31.73
		LT	1	1.03
	Triploid	LJJ, LLJ	Undefined	34.56 ± 4.47
		LTT	0.5	0.67
		LJB, LLT, LTJ	2	1.93 ± 0.23
	Tetraploid	LJJJ, LLLJ	Undefined	38.87 ± 0.06
		LTTT	0.33	0.53
		LJBB, LTTJ	1	1.42 ± 0.14
		LJJB, LLLT, LTJJ	3	2.66 ± 0.21
T27	Diploid	LJ	0	0.09
		LT	1	0.45
	Triploid	LJB, LJJ, LLJ	0	0.03 ± 0.03
		LLT, LTJ	0.5	0.29 ± 0.06
		LTT	2	0.59
	Tetraploid	LJBB, LJJB, LJJJ, LLLJ	0	0.04 ± 0.02
		LLLT, LTJJ	0.33	0.21 ± 0.04
		LTTJ	1	0.37
		LTTT	3	0.77
L313	Diploid	LJ, LT	1	0.71 ± 0.02
	Triploid	LJB, LJJ, LTJ, LTT	0.5	0.52 ± 0.05
		LLJ, LLT	2	1.07 ± 0.06
	Tetraploid	LJBB, LJJB, LJJJ, LTJJ, LTTJ, LTTT	0.33	0.39 ± 0.02
		LLLJ, LLLT	3	1.5 ± 0.06
Ti92	Triploid	LTTi	0.5	9.85
	Tetraploid	LTTTi	0.33	6.91
		LTTiTi	1	13.08

SNP, single nucleotide polymorphism.

multiple copies of distinct genomes. To generate 'simulated unisexuals' of known genetic make-up, we mixed DNA of known concentration (25 μg/mL; quantified using a Qubit fluorometer) from each sexual species into 19 combinations found in unisexual salamanders (Table 3). We chose to include every possible combination for the most common biotypes across the range of the unisexuals (L and J combinations; Bogart & Klemens 1997) and in our region of interest (L and T combinations), as well as including a number of biotype combinations that are rare but have been documented as naturally occurring (i.e. biotypes including Ti and B; Bogart *et al*. 2007, 2009). For example, to simulate an LJJ, we mixed one part *laterale* DNA with two parts *jeffersonianum* DNA. We assume that this simulates the DNA make-up of a sample extracted from an individual of a matching biotype. We were then able to calculate expected peak area ratios for each SNP, given the known identity of the SNPs in each type of genome. For example, SNP L313 has an A peak for *laterale* genomes, but a T peak for all others. So for an LJ unisexual, we predict that the peak ratios should be 1A:1T (ratio = 1.0). In an LLJ or LLT unisexual, the A peak should be twice as big as the T peak (A/T ratio = 2.0), while an LJJ or LTJ unisexual should have the opposite ratio (the T peak should be twice as big as the A peak; A/T ratio = 0.5). Based on the relationship between predicted and observed peak area ratios, we generated regression equations for each SNP that would allow us to estimate ploidy in unknown samples (see below).

Comparisons with microsatellite-identified unisexuals

Finally, to assess our method on samples from wild-caught unisexuals, we obtained DNA samples from J. Bogart (University of Guelph) from 19 individuals that he had previously genotyped using a suite of microsatellite loci and applied the SNP described earlier. These samples have the advantage of coming from real animals, but the disadvantage of possibly being misidentified because of errors in interpreting their microsatellite genotypes. To infer biotype and estimate ploidy, we used the regression equations generated from the simulated unisexual DNA samples described earlier. For each test sample, we inserted the observed peak area ratios into the regression equations to adjust them to the scale of the predicted ratios, which should be either 0.33, 0.5, 1, 2 or 3 (Table 3). We then calculated the absolute value difference between the adjusted values and each of these five possible values; to assign ploidy, we used the minimum of these differences. For example, an adjusted value of 0.6 would be rounded to 0.5, while an adjusted ratio value of 0.8 would be rounded to 1. In the few cases in which individual SNPs disagreed on the ploidy assessment for a single sample, we used the ploidy inferred by the SNP with the closer of the two absolute value differences.

Results

Single nucleotide polymorphism peaks clearly indicate the presence of particular species' sexual genomes in the unisexual salamanders (Fig. 1; Table 1). TBTi127 reliably separates *A. laterale* and *A. jeffersonianum* genomes (all show an A nucleotide at this site) from *A. texanum* and *A. barbouri* genomes (all show G). From the reference samples we sequenced, we also expected it to separate *A. tigrinum* (T); however, this was only moderately successful. Some populations of *A. tigrinum* appear to have alleles with an indel mutation that causes double sequence at the E14G07 locus. If present, this mutation causes tiger salamander genomes, even in pure *A. tigrinum*, to resemble unisexuals because two peaks are present at the SNP sites. This result was not limited to certain geographic regions. We followed up our initial results by sequencing 22 *A. tigrinum* samples from three states and found indels present in some samples from each state (Illinois: 5/7 indel present; Indiana: 2/6; Ohio: 4/9). We strongly recommend sequencing a number of *A. tigrinum* individuals to determine whether the indel mutation is present if tiger salamanders are involved in the unisexual complex in the region of interest. If present, we recommend using the second locus E12G03 to diagnose presence of *tigrinum* genomes (see below).

In our initial screening ($n = 30$ individuals), T27 appeared to separate *A. texanum* (C) from all four other sexual species, including its close relative *A. barbouri* (T). These two species are both phenotypically and genotypically similar (Shaffer *et al.* 1991; Bi & Bogart 2010) and may share the same vernal pool breeding sites (Bogart *et al.* 2009; Greenwald and Gibbs, unpublished data). However, the sequencing of additional samples, particularly from northern Ohio, indicated that this SNP does not reliably diagnose mtDNA-defined samples of *A. texanum* and *A. barbouri* in this region (Gibbs and Greenwald, unpublished data). The reason for this incongruity is unknown at present and is under investigation. However, this issue appears to be geographically limited to northcentral Ohio; in other regions, this SNP reliably distinguishes these two species (see Discussion).

The polymorphism assayed by L313 distinguishes *A. laterale* (A) from all other species (T). As with TBTi127, this SNP also separates *A. tigrinum* (C) in some cases; however, the problem of unclear sequence described earlier is still present in certain populations of tiger salamanders. As mentioned above, because of these amplification issues in some *A. tigrinum* samples, we developed a SNP on a second locus (E12G03). This SNP, Ti92, cleanly amplified in all samples; *tigrinum* genomes have a G at this site, while all other species have a C. If *A. tigrinum* is present in the region of interest, and full (sexual) *A. tigrinum* samples fail to amplify cleanly, we recommend including this SNP in the assay to confirm presence of Ti genomes in unisexuals.

Tests with simulated unisexuals

Expected and observed peak area ratios are highly correlated with each other (all $R^2 > 0.88$; see Fig. 2 and Table 3), meaning there is a strong linear relationship between the numbers of genome copies present and the relative sizes of the SNP peaks. However, this relationship is not the expected one-to-one ratio, and so we generated regression equations to allow us to adjust an observed unknown sample to the known scale of the predicted values. This step is not absolutely necessary; peak ratio values obtained from unknown samples could be compared directly with the observed (unadjusted) values from our simulated unisexual data. However, we found it more intuitive to do the conversions and then determine which expected value (which should always be 0.33, 0.5, 1, 2 or 3) each observed value was closest to. The equations for particular SNPs are as follows:

$$\text{TBTi172}: y = 0.7558x + 0.4193;\ R^2 = 0.88$$

$$\text{T27}: y = 0.2566x + 0.0803;\ R^2 = 0.91$$

$$\text{L313}: y = 2.3975x - 0.6496;\ R^2 = 0.97$$

where y is predicted value and x is observed value from the chromatograms.

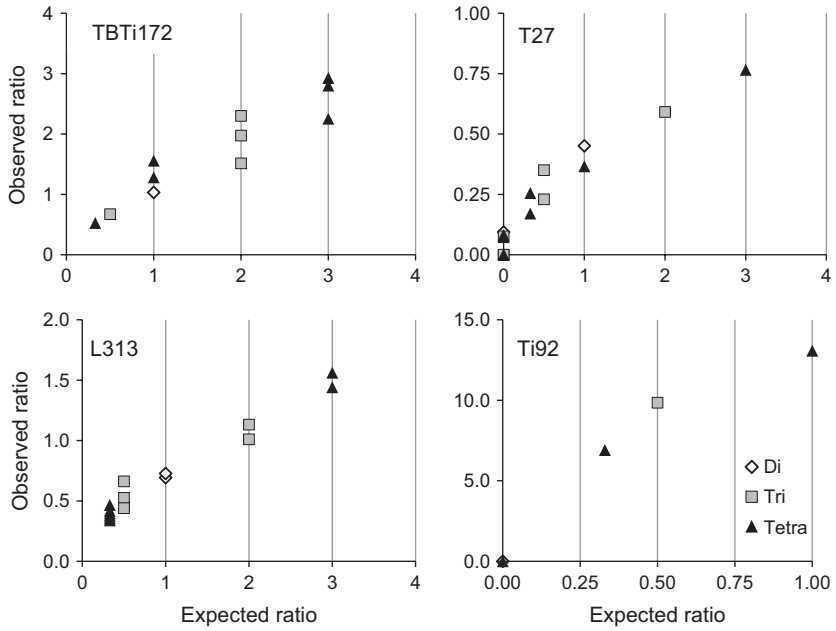

Fig. 2 Each panel represents an individual single nucleotide polymorphism (SNP) locus used in a test of ploidy determination. Expected ratios were determined *a priori* for each DNA combination (e.g. for a SNP differentiating L and J genomes, an LJJ individual would have an expected ratio of 1:2). There is a strong linear fit between the peak heights expected in the polyploid SNPs, and those observed in the SNaPshot trials (TBTi172: $r^2 = 0.88$; T27: $r^2 = 0.91$; L313 $r^2 = 0.97$; Ti92: $r^2 = 0.90$). Symbols represent ploidy as follows: diploid (◇), triploid (■) or tetraploid (▲).

For some tests with SNP TBTi172, the theoretical expected ratio was undefined (Table 3). This SNP differentiates L and J genomes from all others. Therefore, if salamanders contain only L and J combinations (e.g. LJJ, LLJ and LJJJ), the L/J (A) peak should be nonzero while the T/B/Ti (G) peak should be zero, making the A/G ratio undefined. For these cases, the observed ratios for the simulated DNA were extremely large (mean 35.72 ± 1.98 SE for five cases).

We tested Ti92 on three mixed-DNA samples containing 1 or 2 copies of Ti genomes (LTTi, LTTTi and LTTiTi), as well as numerous samples containing 0 copies. There was again a good relationship between the expected and observed peak area ratios (Fig. 2 and Table 3); the relationship was described by the equation:

$$\text{Ti92}: y = 12.715x + 1.6432;\ R^2 = 0.90$$

Comparisons with microsatellite-identified unisexuals

Single nucleotide polymorphism- and microsatellite-based identifications were generally in agreement, with no disagreement over genome complement but seven disagreements over ploidy (Table 4). In the ploidy disagreements, there were five cases in which the SNP-based ploidy estimate was lower than the corresponding microsatellite-based estimate, and two cases in which it was higher.

Unisexual salamanders with all L and J genomes have a theoretically undefined predicted value for SNP TBTi127, as explained previously. In terms of actual data processing, there is often a nonzero but extremely small value for the G peak that results in a very large A/G ratio in these cases. For eight of the test samples genotyped here, the A/G ratios averaged 50.04 ± 10.22 SE, creating a clearly bimodal pattern (all other samples had A/G ratios <4). We recommend the interpretation that A/G > 5 implies that only L and J genomes are present.

Discussion

The SNP loci described here reliably indicate the presence of particular sexual species' genomes in unisexual salamanders. With one exception (*A. texanum* vs. *A. barbouri*, discussed below), we detected no intraspecific polymorphism at these sites in our reference samples; thus, the SNPs give a clear signal of the presence of particular genomes. Furthermore, there is a strong linear relationship between predicted peak area ratios and the known 'ploidy' in simulated unisexual DNA. This relationship allows us to determine ploidy of real, unknown samples. In a test of this technique, SNPs and microsatellites generally agreed on unisexual biotype identification, with 100% agreement on genome complement and 66% agreement on ploidy assignment.

The disagreement between the SNP-based and microsatellite-based biotype identifications is difficult to interpret, as it could arise from misidentifications during either procedure. However, we are confident in the SNP protocol outlined here because of the highly consistent results with the simulated DNA, when the exact genome ratios were known *a priori*. Interestingly, the peak area ratios for observed and expected values were not 1:1 for any of the SNP loci tested on simulated unisexual DNA.

Table 4 Peak area ratios (corrected using regression equations; see text for details) and biotype inferences provided by each SNP locus. The 'SNP Biotype' column integrates the biotype information from all three loci, while the 'μsat Biotype' column gives the biotype assigned by J. Bogart based on microsatellite loci. There are seven cases of ploidy disagreement, but no disagreement in genome type presence/absence

	Ti92		T27		TBTi172		L313			
JPB#	G?	Biotype	C?	Biotype	A/G ratio	Biotype	A/T ratio	Biotype	SNP Biotype	Micro Biotype
17450	No	Ti	Yes	T	1.36	LX	1.06	LX	LT*	LTTT
30243	No	Ti	No	T	57.81	All LJ	0.05	LXXX	LJJJ	LJJJ
30602	Yes	Ti	No	T	10.39	All LJ	1.21	LX	LJTi*	LJJTi
30924	No	Ti	Yes	T	2.29	L(L)X	0.43	LXX	LTJ	LTJ
31333	No	Ti	Yes	T	71.42	All LJ	0.32	LXXX	LTJJ	LTJ†
31989	No	Ti	Yes	T	3.53	L(L)(L)X	0.48	LXX	LTJ*	LTJJ
32255	No	Ti	No	T	13.77	All LJ	3.43	LLLX	LLLJ	LLLJ
32620	Yes	Ti	No	T	Undefined	All LJ	1.07	LX	LJTi	LJTi
33254	No	Ti	Yes	T	2.65	L(L)(L)X	0.45	LXX	LTJ	LTJ
33840	No	Ti	Yes	T	1.50	LX/LLX	1.57	LLX	LLT	LT†
37034	Yes	Ti	Yes	T	1.47	LX	1.55	LLX	LTTi	LTTi
37068	No	Ti	Yes	T	0.00‡	All TB	−0.38	No L	TT	TT
37255	No	Ti	No	T	12.32	All LJ	2.24	LLX	LLJ*	LLLJ
37709	No	Ti	No	T	0.00	All TB	−0.48	No L	BB	BB
37735	No	Ti	No	T	2.98	L(L)(L)X	0.06	LXXX	LJJB	LJJB
38309	No	Ti	No	T	28.98	All LJ	−0.65	No L	JJ	JJ
38335	No	Ti	No	T	53.42	All LJ	0.05	LXXX	LJJJ	LJJJ
38388	No	Ti	No	T	45.03	All LJ	Undefined	LL	LL	LL
38413	No	Ti	No	T	24.24	All LJ	2.74	LLLX	LLLJ	LLLJ
39680	No	Ti	Yes	T	1.04	LX	0.78	LXX	LT	LT
39945	No	Ti	No	T	23.19	All LJ	0.66	LXX	LJJ*	LLJJ

SNP, single nucleotide polymorphism.
*SNP ploidy < Microsatellite ploidy.
†Microsatellite ploidy < SNP ploidy.
‡Zero values are not corrected with regression equation.

This pattern may be attributable to a number of possible reasons: variation in genome size among species, variation in DNA quality or variation in how well the SNP probes anneal to the various species' DNA. The first two reasons appear unlikely. Genome size estimates (*C*-values) based on flow cytometry vary only a small amount among species involved in the unisexual complex: *A. jeffersonianum*: 28.8 pg; *A. laterale*: 29.2 pg; *A. texanum*: 27.1 pg; and *A. tigrinum*: 30.7 pg (Licht & Lowcock 1991). DNA quality was also similar because samples were approximately the same age, were extracted in the same way and when assessed on a gel were of similar quality. Observed peak area ratios were fairly consistent across multiple samples with the same expected ratios, supporting the conclusion that deviations from the expected 1:1 ratio are likely due to inherent variation in the SNP protocol, rather than issues of DNA quantity or quality.

Using our method to discriminate between *A. texanum* and *A. barbouri* genomes remains problematic. These taxa are sister species with a complex evolutionary history that has lead to contrasting patterns of divergence based on mtDNA vs. nuclear DNA (Niedzwiecki 2005). Although our initial screening seemed to support a fixed polymorphism at T27 (with *A. texanum* having a C and *A. barbouri* having a T), this pattern did not hold up upon adding more samples to the panel. However, the only *A. texanum* individuals to have a T at this site were from Nebraska, well outside the range of *A. barbouri*; all samples from Ohio, Indiana and West Virginia were fixed for C. All *A. barbouri* samples from Indiana and West Virginia had a T at this site; however, the Ohio *A. barbouri* were split between T ($n = 3$) and C ($n = 4$). Interestingly, all of the individuals with a T at this site were from southwest Ohio, in the region that has traditionally been considered the known range of *A. barbouri*. All of the individuals with a C were from the northern half of the state, where *A. barbouri* was formerly not considered to be present (Pfingsten & Downs 1989), but where we identified *barbouri* mtDNA in populations previously thought to be *A. texanum*. It appears that the evolutionary history of these two species is more complex than was previously understood, and additional work to understand how mtDNA and nuclear DNA variation covary in these species is needed. Regardless, we recommend retaining T27

in the assay, as the presence of a C at this site can confirm the results at TBTi172 in separating the *texanum-barbouri* pair from all other species. In addition, if the polymorphism happens to be fixed in the population/region of interest (e.g. Indiana, West Virginia and southern Ohio), these two species' genomes can be successfully distinguished.

Another difficulty with our SNP development was the problem described previously with clean amplification of locus E14G07 in some (but not all) *A. tigrinum* samples. Importantly, this issue was not limited to particular locations – there were problematic *A. tigrinum* samples from Illinois, Indiana and Ohio. Because of the double sequence that is produced in these samples, the SNPs often resemble those of unisexual salamanders (with more than one peak per locus). We therefore developed a SNP on another locus (Ti92 on E12G03) that amplifies cleanly in all populations of tiger salamanders that we assessed. Fortunately, this particular genome generates distinct phenotypes in unisexuals in the form of large size and golden-coloured eyes that signal the presence of an *A. tigrinum* genome (K. Greenwald, personal observation); thus, it should be relatively obvious whether assaying Ti92 is necessary in the population(s) of interest.

Recommendations for Implementation

We make the following two recommendations regarding the use of this SNP assay. (i) The first step should be locating sexual species from the region of interest (and confirming their species identity via mtDNA sequencing). Once identified, these sexual individuals should be sequenced at E14G07 (and E12G03 if *A. tigrinum* is present); this will allow confirmation of the species-specific polymorphisms documented here. (ii) Prior to the SNP analysis, we recommend sequencing mtDNA for all individuals of interest. If this is performed, sexual individuals can be removed from the subsequent SNP analysis. It is helpful to know that all individuals in the SNP analysis are unisexuals, as it narrows the possibilities for biotypes and also means that there will always be at least one *laterale* genome present. This approach may also be financially advantageous depending on the relative costs of sequencing vs. SNP analysis, as fewer individuals will have to be processed for SNPs once sexual individuals are removed.

The most difficult part of determining biotype in unisexual salamanders is the accurate assessment of ploidy. If feasible, it would be ideal to use multiple methods to obtain the most precise assessment. The SNP assay can be combined with microsatellite data as presented here; microsatellites have the additional advantage of providing information on differentiation within and among populations of unisexuals with the same biotypes. Alternatively, flow cytometry is an effective way to assess ploidy, as red blood cell size increases with the number of genomes present (Bogart & Licht 1986; Lowcock & Griffith 1991). The combination of SNPs and flow cytometry would work well to produce an accurate depiction of both genome complement and ploidy. This method could be especially beneficial in populations that contain high numbers of diploid and '2/2 tetraploid' biotypes (e.g. LJ with LLJJ or LT with LLTT). The SNP assay will not discriminate between these, because the same genomes are present and the peak area ratios in both cases are 1.0. However, these particular tetraploid biotypes appear to be exceptionally rare (e.g. 1 of over 600 unisexuals surveyed in Bogart & Klemens 2008), so interpreting a peak area ratio of 1.0 as a diploid appears to be a relatively safe assumption.

Additional SNPs that isolate *A. laterale* from all other species (as with our SNP L313) would be a valuable area of future work. These SNPs are especially useful for diagnosing ploidy, as all unisexual salamanders have at least one L genome (Bi *et al.* 2008), and the most common biotypes are composed of L plus one other type of genome (e.g. LLJ, LJJ and LTT) making the peak area ratios easily interpretable. While development of more SNPs on other loci would be beneficial, our study shows that this is a feasible and straightforward approach to assessing genomic composition of unisexual *Ambystoma.* Additionally, this technique may facilitate research on the sexual species involved in the complex, as the close resemblance of some biotypes and sexual species (Pfingsten & Downs 1989) necessitates genetic methods of identification.

Acknowledgements

We thank Jim Bogart for generously sharing his samples and expertise, Jose Diaz for assistance with troubleshooting lab protocols and processing samples and Carolyn Caldwell (Ohio Division of Wildlife) for her sustained support of this work. We thank the following people for contributing samples or assisting with collections: M. Austin, L. Blyth, R. Brodman, C. Ciola, B. Cosentino, J. Davis, J. Dyer, M. Graziano, T. Hetherington, G. Lipps, T. Matson, M. Micacchion, D. Oestrich, R. Pfingsten, E. Rhoads, W. Savage, J. Wooten and M. Zloba. This research was supported by the funds from the Ohio Division of Wildlife, Ohio State University and the National Geographic Society in the form of a Waitt Grant to KRG.

References

Bi K, Bogart JP (2006) Identification of intergenomic recombinations in unisexual salamanders of the genus *Ambystoma* by genomic in situ hybridization (GISH). *Cytogenetic and Genome Research*, **112**, 307–312.

Bi K, Bogart JP (2010) Time and time again: unisexual salamanders (genus *Ambystoma*) are the oldest unisexual vertebrates. *BMC Evolutionary Biology*, **10**, 238–253.

Bi K, Bogart JP, Fu J (2008) The prevalence of genome replacement in unisexual salamanders of the genus *Ambystoma* (Amphibia, Caudata) revealed by nuclear gene genealogy. *BMC Evolutionary Biology*, **8**, 158–166.

Bogart JP, Klemens MW (1997) Hybrids and genetic interactions of mole salamanders (*Ambystoma jeffersonianum* and *A. laterale*) (Amphibia: Caudata) in New York and New England. *American Museum Novitates*, **3218**, 1–78.

Bogart JP, Klemens MW (2008) Additional distributional records of *Ambystoma laterale*, *A. jeffersonianum* (Amphibia: Caudata) and their unisexual kleptogens in northeastern North America. *American Museum Novitates*, **3627**, 1–58.

Bogart JP, Licht LE (1986) Reproduction and the origin of polyploids in hybrid salamanders of the genus *Ambystoma*. *Canadian Journal of Genetics and Cytology*, **28**, 605–617.

Bogart JP, Lowcock LA, Zeyl CW, Mable BK (1987) Genome constitution and reproductive biology of hybrid salamanders, genus *Ambystoma*, on Kelleys Island in Lake Erie. *Canadian Journal of Zoology*, **65**, 2188–2201.

Bogart JP, Bi K, Fu JZ, Noble DWA, Niedzwiecki J (2007) Unisexual salamanders (genus *Ambystoma*) present a new reproductive mode for eukaryotes. *Genome*, **50**, 119–136.

Bogart JP, Bartoszek J, Noble DWA, Bi K (2009) Sex in unisexual salamanders: discovery of a new sperm donor with ancient affinities. *Heredity*, **103**, 483–493.

Hedges SB, Bogart JP, Maxson LR (1992) Ancestry of unisexual salamanders. *Nature*, **356**, 708–710.

Hurst CD, Zuiverloon TCM, Hafner C, Zwarthoff EC, Knowles MA (2009) A SNaPshot assay for the rapid and simple detection of four common hotspot codon mutations in the PIK3CA gene. *BMC Research Notes*, **2**, 66–71.

Licht LE, Lowcock LA (1991) Genome size and metabolic rate in salamanders. *Comparative Biochemistry and Physiology*, **100B**, 83–92.

Lowcock LA, Griffith H (1991) The *Ambystoma laterale-jeffersonianum* complex in central Ontario – ploidy structure, sex ratio, and breeding dynamics in a bisexual-unisexual community. *Copeia*, **1991**, 87–105.

Lowcock LA, Licht LE, Bogart JP (1987) Nomenclature in hybrid complexes of *Ambystoma* (Urodela, Ambystomatidae) – no case for the erection of hybrid species. *Systematic Zoology*, **36**, 328–336.

McKnight ML, Shaffer HB (1997) Large, rapidly evolving intergenic spacer in the mitochondrial DNA of the salamander family *Ambystomatidae* (Amphibia: Caudata). *Molecular Biology and Evolution*, **14**, 1167–1176.

Niedzwiecki J (2005) Evolutionary history and hybridization of two mole salamander sister species from different habitats. Ph.D. Dissertation, University of Kentucky, Lexington, Kentucky.

Noël S, Dumoulin J, Ouellet M, Galois P, Lapointe FJ (2008) Rapid identification of salamanders from the Jefferson complex with taxon-specific primers. *Copeia*, **2008**, 158–161.

Norton N, Williams NM, Williams HJ *et al.* (2002) Universal, robust, highly quantitative SNP allel frequency measurement in DNA pools. *Human Genetics*, **110**, 471–478.

Pfingsten RA, Downs FL (1989) *Salamanders of Ohio*. College of Biological Sciences, The Ohio State University, Columbus, OH.

Ramsden C (2008) Population genetics of *Ambystoma jeffersonianum* and sympatric unisexuals reveal signatures of both gynogenetic and sexual reproduction. *Copeia*, **2008**, 586–594.

Ramsden C, Beriault K, Bogart JP (2006) A nonlethal method of identification of *Ambystoma laterale*, *Ambystoma jeffersonianum* and sympatric unisexuals. *Molecular Ecology Notes*, **6**, 261–264.

Rhoads EA, Krane CM, Williams PK (2009) A molecular screening method for unisexual *Ambystoma* using mitochondrial DNA. *Conservation Genetics*, **11**, 1177–1179.

Shaffer HB, McKnight ML (1996) The polytypic species revisited: genetic differentiation and molecular phylogenetics of the tiger salamander *Ambystoma tigrinum* (Amphibia: Caudata) complex. *Evolution*, **50**, 417–433.

Shaffer HB, Clark JM, Kraus F (1991) When molecules and morphology clash: a phylogenetic analysis of the North American Ambystomatid salamanders (Caudata: Ambystomatidae). *Systematic Zoology*, **40**, 284–303.

Smith JJ, Putta S, Walker JA *et al.* (2005) Sal-Site: Integrating new and existing ambystomatid salamander research resources. *BMC Genomics*, **6**, 181–186.

Vietes DR, Min M-S, Wake DB (2006) Rapid diversification and dispersal during periods of global warming by plethodontid salamanders. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 19903–19907.

Data Accessibility

Sexual species used in initial SNP screening are extracted DNA vouchers in the collection of H. Lisle Gibbs (Ohio State University). Unisexual DNA used for comparison (JPB numbered samples; Table 4) are extracted DNA vouchers in the collection of Dr. James P. Bogart (University of Guelph). All sequences generated for this study are available in FAS file format as Supporting Information; unique sequences >200 bp are additionally available on NCBI/GenBank, accession numbers JN648810–JN648825 (E14G07) and JN652592–JN652602 (E12G03).

Supporting Information

Additional supporting information may be found in the online version of this article.

Appendix S1 XLS file containing all sequences generated for this study (loci E14G07 and E12G03).

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.