Key words: Rana pipiens, Rana blairi, Rana palustris, antimicrobial peptide, balancing selection, fluctuating selection

E-mail: tenness@science.oregonstate.edu.

doi:10.1093/molbev/msn208
Advance Access publication September 17, 2008

© 2008 The Authors
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
be classified as either within-lineage variation or between-lineage variation (Innan and Tajima 1999). Under stable balancing selection, Tajima’s $D$ (Tajima 1989) for within-lineage variation should be close to 0. Under fluctuating selection, Tajima’s $D$ should be negative within lineages that have recently increased in frequency, which are analogous to growing populations, and positive for lineages that have recently decreased in frequency, which are analogous to populations going through a bottleneck. Similarly, stable balancing selection predicts highly similar allele frequencies among populations, high synonymous variation due to the antiquity of alleles at that locus, a correlation between the frequency of allelic lineages and the neutral variation they harbor, and a low probability of fixation for any allele. In contrast, fluctuating selection permits divergent allele frequencies among populations and predicts low synonymous variation due to periodic bottlenecks for every allelic lineage, no correlation between the frequency of allelic lineages and the neutral variation they harbor, and the occasional fixation of particular alleles.

Antimicrobial peptides (AMPs) of the innate immune system have only rarely been observed to be under pathogen-driven balancing selection (Tennessen 2005b). These cationic, amphipathic mature peptides are cleaved off of a larger protein and then bind to the cell membranes of bacterial, fungal, and enveloped viral pathogens, killing them (Hancock 2001). Research on AMPs has been substantial in recent years, due in part to an interest in developing them for therapeutic application (Hancock 2001). Positive selection on AMP genes is very common and has resulted in an enormous functional diversity of these molecules among species and among loci in many taxa (Tennessen 2005b). Some human and mussel AMP loci appear to be under balancing selection (Hollox and Armour 2008; Pahdi and Verghese 2008), but other studies of intraspecies genetic diversity at AMP genes have revealed no evidence for balancing selection (Clark and Wang 1997; Lazzaro and Clark 2003; Tennessen and Blouin 2007).

The AMPs of leopard frogs (genus Rana; in this paper, we ignore the recent proposal by Frost et al. [2006] to revise the genus to Lithobates) are among the most well studied (Conlon et al. 2004). Most of them consist of an α-helix with a disulfide bridge forming a loop at the C-terminal end. They are functionally diverse, frequently with activity against both amphibian and human pathogens (Goraya et al. 2000; Chinchar et al. 2004; Rolls-Smith and Conlon 2005). Given the global crisis of amphibian population declines mediated by emerging infectious diseases and the immunological importance of AMPs, genetic diversity at AMP loci could be an important determinant of amphibian population stability (Daszak et al. 2003; Woodhams, Rolls-Smith, et al. 2006; Woodhams, Voyles, et al. 2006). Of the four AMP families secreted by the northern leopard frog, Rana pipiens (brevinin-1, ranatuerin-2, temporin-1, and esculentin-2), the brevinin-1 family is both the most diverse and the most active against microbes (Goraya et al. 2000; Tennessen and Blouin 2007). Previously, we investigated allelic variation at the Ranatuerin2 AMP locus in R. pipiens and found that a single haplotype had been fixed in the species by a positive selective sweep (Tennessen and Blouin 2007). We also found substantial diversity among AMP sequences at five loci of the brevinin-1 family in R. pipiens, but we were unable to assess how much of this variation was within versus between loci (Tennessen and Blouin 2007). In this paper, we test whether some of the brevinin-1 variants are allelic by designing primers that amplify a single brevinin-1 locus. We examine patterns of genetic diversity in over 400 individuals of R. pipiens. In order to assess the generality of our observations across species, we also sequence this locus in a smaller number of plains leopard frogs (Rana blairi) and pickerel frogs (Rana palustris). Several highly divergent allelic lineages are segregating at this locus, and we present evidence that the balanced maintenance of these alleles is due to a dynamic process of fluctuating natural selection.

Materials and Methods

Tissues

Samples of R. pipiens were collected from 13 sites throughout its range across northern North America as described previously ($n = 20–46$ individuals per site; supplementary table 1, Supplementary Material online; fig. 1; Hoffman et al. 2006). For one of the sites in New York, we also obtained a historical sample of R. pipiens collected in 1971 (approximately 15 generations earlier) as described previously ($n = 25$; Hoffman et al. 2006). We collected samples of R. palustris from two sites in Wisconsin and Michigan ($n = 23$ and 40, respectively) and samples of R. blairi from a single site in Illinois ($n = 27$; supplementary table 1, Supplementary Material online; fig. 1). Further samples of R. blairi were obtained from the Museum of Vertebrate Zoology (MVZ Herp 240161–240184), having all been originally collected from a single site in Missouri ($n = 24$; supplementary table 1, Supplementary Material online; fig. 1). The sample size at every site was at least 20 frogs (mean number of frogs per population = 28.8). All tissues consisted of toe clips preserved by desiccation in 1.5-ml tubes filled with Drierite desiccant (W. A. Hammond Drierite Co., Xenia, OH), except for the MVZ samples, which had been frozen. DNA was extracted as described previously (Hoffman and Blouin 2004) or using DNeasy Blood and Tissue Kits (Qiagen Inc., Valencia, CA).

We carried out 25 μl polymerase chain reactions (PCRs) using standard buffer conditions, 1.5 mM MgCl2, 0.2 mM each deoxyribonucleotide triphosphate, approximately 100 ng DNA, and 0.5 Units Taq DNA polymerase.
We visualized PCR products under ultraviolet light, purified them with the MoBio Ultraclean PCR cleanup kit (Solana Beach, CA), and sequenced them through the Nevada Genomics Center (Reno, NV).

Sequences

Previously, we had cloned two brevinin-1 sequences from a single R. pipsiens individual that were similar in the intron but quite divergent in the mature peptide region (GenBank accession numbers DQ276967 and DQ276968; Tennessen and Blouin 2007). In this study, we tested whether they were allelic by assessing whether the genotype frequencies were in Hardy–Weinberg equilibrium within populations. To do so, we designed a primer (Brev1PF4; 5′-GAT GAC CCA ATA ATA ATT TTT C-3′) that would bind to these sequences but not to any other known brevinin-1 gene in R. pipsiens. We used the primers Brev1PF4 and Brev1PR1 (Tennesen and Blouin 2007), which bind outside of the coding region, to amplify genomic DNA from R. pipsiens, R. palustris, and R. blairi. PCR amplification conditions consisted of an initial denaturation step at 94 °C for 5 min; 35 cycles of denaturation at 94 °C for 45 s, annealing at 51 °C for 30 s, and extension at 72 °C for 1 min; and a final single extension step at 72 °C for 5 min. We sequenced the resultant PCR products with an internal primer (Brev1PF5; 5′-GAA AGC TCT GTG CCA TAG-3′).

Due to difficulties amplifying some R. blairi individuals with this primer pair, we designed a forward primer specific to R. blairi. We sequenced the primer-binding site in R. blairi using a cloning procedure described previously (Tennesen and Blouin 2007). Briefly, we amplified R. blairi genomic DNA using the more distal and degenerate forward primer Brev1PF3 with Brev1PR1 in PCRs with the high fidelity enzyme Pfu DNA polymerase (Promega, Madison, WI). The resultant PCR products were incubated with Tag polymerase in order to add 3′ adenines and cloned into Escherichia coli using the TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA). The recombinant locus was amplified from screened colonies using the primers T3 and T7, and the resultant PCR products were sequenced with T7. We used these sequences to design the primer Brev1BLF (5′-TAG ATG ACC TAA TAA TAA TTT TTC-3′), which binds in the intron. We amplified and sequenced DNA from all R. blairi individuals using both primer set (Brev1PF4 and Brev1PR1) and primer set (Brev1BLF and Brev1PR1).

In order to compare genetic diversity at the brevinin-1 locus with other, putatively neutral, nuclear loci, we sequenced 1,246 bp of four nuclear genes in a subset of individuals; Arcadlin (primers ArcadF and ArcadR, 178 bp coding, 87 bp intron; Tennessen and Blouin 2007), Myosin (primers MyosinF and MyosinR, 65 bp coding, 85 bp intron; Tennessen and Blouin 2007), FIBI7 (primers FIBI7U and FIBI7L, 50 bp coding, 208 bp intron; Di Candia and Routman 2007), and Tyrosinase (primers Tyr1A and Tyr1G, 573 bp coding; Bossuyt and Milinkovitch 2000). These loci were chosen because we had preexisting protocols for amplifying them in Rana and because their polymorphisms are consistent with neutral evolution. We sequenced Arcadlin, Myosin, FIBI7, and Tyrosinase in all frogs from six populations: populations 1, 2, 4, 15, 16, and 17 (ﬁg. 1). These six populations represented all R. palustris and R. blairi individuals but only a fraction of all R. pipsiens individuals (59 frogs). In order to obtain a more representative sample of R. pipsiens, we randomly chose two frogs from each of the remaining 11 contemporary R. pipsiens populations and we sequenced Arcadlin, Myosin, FIBI7, and Tyrosinase in these 22 frogs.

Analysis

All sequences have been deposited in GenBank under accession numbers EU407141–EU407149, EU407151–EU407176, and EU769510–EU769553. All loci were nuclear and diploid, so two alleles were obtained from all individuals. Heterozygotes were identified by the presence of double peaks in electropherograms. To determine common haplotypes, we noted all homozygous genotypes and tested for recombination using the four-gamete test. Given evidence for minimal recombination, we determined the remaining haplotypes by subtracting out previously observed haplotypes from the heterozygous genotypes or, for some R. blairi genotypes, by using allele-specific PCR. To confirm that observed polymorphisms were allelic and not due to coamplification of duplicated loci, we tested whether each population was in Hardy–Weinberg equilibrium. It is highly unlikely that polymorphic sequences from multiple loci would conform to Hardy–Weinberg expectations.

We used PAUP* (version 4.0b10, Swofford 2002) to construct a maximum parsimony phylogeny of all brevinin-1 sequences. We used DnaSP (version 4.0, Rozas et al. 2003) to calculate standard statistics of selective neutrality, including ZnS (Kelly 1997), Tajima’s D (Tajima 1989), haplotype diversity or expected heterozygosity (HD), the number of haplotypes (h), genetic variation (π), and the number of segregating sites (S), and to test their significance using coalescent simulations (conditioned on S with no recombination and 1,000 replicates). To minimize the effects of population subdivision on these test statistics, we performed each test within individual populations. We used MEGA (version 2.1; Kumar et al. 2001) to test whether the ratio of nonsynonymous variation to synonymous variation (πS/πN) was significantly greater than 1. If observed πS was 0, we estimated πS/πN to be greater than what the value would be if a single synonymous polymorphism had been observed in a single sequence. Neutrality tests were conducted on the mature peptide region only because that is the region where nonneutral evolution most frequently occurs (Tennesen 2005b). We represented relationships among alleles with a haplotype network, as opposed to a phylogenetic tree, because networks show allele frequencies and because there is little evidence for recombination at this locus.

To distinguish between fluctuating and stable balancing selection, we performed several tests of allelic stasis. A model of stable balancing selection predicts that the silent variation at the selected locus should be higher than elsewhere in the genome because allelic lineages have been maintained at intermediate frequency for a very long time.
and many mutations have been able to accumulate over this long coalescence period (Charlesworth 2006). In contrast, under fluctuating selection each allelic lineage spends some time being rare when selection favors the other allelic lineage. Thus, the effective population size of the locus, estimated as the sum of the harmonic means of the effective sizes of all allelic lineages over time, is small. As a result, silent variation will be purged by genetic drift, making silent variation lower than silent variation at neutral loci (Tiffin et al. 2004). Thus, we calculated genetic variation (π) at silent sites (synonymous and noncoding) for Arcadinlin, Myosin, FIB17, Tyrosinase, and the brevinin-1 locus. We used equation 12.63 of Nei and Kumar (2000) to calculate the variance in π, and we used the square root of this variance as the standard error in t-tests comparing mean π values. Similarly, stable balancing selection predicts within-lineage Tajima’s D values near 0, while fluctuating selection predicts high or low Tajima’s D values for lineages that have been shrinking or growing, respectively. Thus, we calculated Tajima’s D within each allelic lineage.

Our five easternmost populations of R. piciens are the “Eastern” populations used by Hoffman et al. (2006) to generate a neutral distribution of the expected Fg ST value, based on microsatellite markers and interspersed sequence repeat loci, using the method of Beaumont and Nichols (1996). Under stable balancing selection, Fg ST would be lower than the neutral expectations. Under fluctuating selection, Fg ST could be higher than the neutral expectation if selection coefficients varied substantially among populations or if the effective population size at the selected locus is substantially lower than at neutral loci. Therefore, we compared Fg ST at the brevinin-1 locus with this distribution. Furthermore, we used the previously described genotype data from the R. piciens individuals at nine microsatellite loci (Hoffman et al. 2006) to construct phylogenies of the R. piciens populations east of the Mississippi River based on both the microsatellites and the brevinin-1 locus. We used the Fitch method based on the genetic distance of Reynolds et al. (1983) in PHYLIP (version 3.65; Felsenstein 1989). We only used the eastern populations because the high genetic divergence between eastern and western R. piciens could result in substantial homoplasy at the microsatellite markers (Hoffman and Blouin 2004; Hoffman et al. 2006). Although microsatellites have a higher mutation rate than coding sequences, among closely related populations genetic distances based on both marker types should be correlated under neutrality (Richard and Thorpe 2001), and therefore differences between the brevinin-1 phylogeny and the microsatellite phylogeny could be due to selection. To formally evaluate the difference between interpopulation divergences at the brevinin-1 locus and at the microsatellite loci, we performed a Mantel test of pairwise genetic distances in FSTAT version 2.9.3 (Goudet 1995).

Populations with low genetic diversity at the brevinin-1 locus could have recently experienced a selective sweep, or they could simply have low genome-wide variation due to a low effective population size. To distinguish between these two hypotheses, we calculated the mean expected heterozygosity (Hg) for each R. piciens population using the previously described genotypes at the nine microsatellite loci (Hoffman et al. 2006). This analysis was restricted to R. piciens because we lacked microsatellite data for the other species. We compared population Hg values at the brevinin-1 locus with the population Hg values at the microsatellite loci, which we assumed to be selectively neutral. If populations with low brevinin-1 diversity also have low microsatellite diversity, genetic drift is probably responsible in both cases. If not, the low brevinin-1 diversity could be due to a recent or ongoing selective sweep.

**Results**

**Brevinin-1 Locus**

We consistently obtained readable brevinin-1 sequences 236 bp in length, which consisted of a 98-bp partial intron followed by a 138-bp partial exon. The exon contained 61 bp of the C-terminal end of the propiece, the entire 72-bp mature peptide region, a stop codon, and 2 bp of postcoding sequence. Out of 517 frogs, 344 were homozygous. All common haplotypes (seen more than five times) were observed as homozygotes and showed no evidence of recombination. All remaining haplotypes (1% of the total) could easily be resolved, either because there was only one way to resolve the heterozygous genotypes such that at least one allele in every genotype matched an allele previously observed in that population or by using allele-specific PCR. It remains possible that some of these nonhomozygous haplotypes were erroneously resolved, which could bias our results, but they are so rare that they are unlikely to affect our conclusions substantially. All populations were found to be in Hardy–Weinberg equilibrium, and there was no evidence of more than two alleles in any one individual, confirming that only a single locus was amplified, here named the Brevinin1.1 locus.

Within R. piciens, four common, highly divergent alleles were observed at the Brevinin1.1 locus: alleles Rp1, Rp2, Rp3, and Rp4 (fig. 2). These four alleles accounted for 97% of the sequences; all remaining minor variants were one or two steps away from one of them (fig. 3). Allele Rp1 encoded the previously described peptide brevinin-1Pa, and allele Rp2 encoded the previously described peptide brevinin-1Pg (Goraya et al. 2000; Tennessen and Blouin 2007). Alleles Rp3 and Rp4 both encoded a peptide, brevinin-1PLa, which previously had only been described in R. palustris (Basir et al. 2000). The distances among these four major R. piciens alleles ranged from 1 to 13 substitutions (fig. 3). The species overall showed an excess of homozygotes, probably owing to population subdivision, because every individual population was in Hardy–Weinberg equilibrium (supplementary table 1, Supplementary Material online) and because this species shows substantial population structure at neutral markers (Hoffman and Blouin 2004; Hoffman et al. 2006). We observed similar patterns of diversity in R. blairi. A single R. blairi sample, MVZ Herp 240161, was homozygous for unique alleles at both Brevinin1.1 and Arcadinlin and carried a unique allele at Myosin. Because this individual could be mislabeled or a migrant, we excluded it from all further analyses. Allele Rb1 accounted for 86% of the R. blairi sequences at the Brevinin1.1 locus (figs. 2 and 3). The remaining R. blairi alleles differed from allele Rb1 by 1–13 substitutions,
including a single-nucleotide indel in the intron. In contrast to the other species, in *R. palustris*, we observed only two alleles which were separated by a single mutational step in the intron. Both coded for the peptide brevinin-1PLa (figs. 2 and 3). The more common allele (*Rp4*) was identical to an allele observed in *R. pipiens*.

Haplotype networks of the *Brevinin1.1* locus alleles are shown in figure 3. There were three main allelic lineages, separated by substantial and mostly nonsynonymous divergence. There was no reciprocal monophyly between *R. pipiens* and either of the other two species, but *R. palustris* and *R. blairi* were reciprocally monophyletic with respect to each other. Although there were no shared *Brevinin1.1* alleles between *R. pipiens* and *R. blairi*, both species had alleles belonging the same divergent allelic lineages, Lineage 1 and Lineage 2 (table 1 and fig. 3).

### Other Loci

For all individuals, preexisting protocols for amplifying and sequencing *Arcadlin*, *Myosin*, *FIBI7*, and *Tyrosinase* (Bossuyt and Milinkovitch 2000; Di Candia and Routman...
2007; Tennessen and Blouin 2007) were successful. In total, we obtained 1,246 bp of nuclear sequence unlinked to Brevinin1.1, including 866 bp of coding sequence and 380 bp of intronic sequence. These loci showed moderate levels of variation (fig. 4 and table 2; supplementary table 2, Supplementary Material online). In R. pipiens and R. blairi, patterns of variation did not deviate from neutral expectations. In R. palustris, several neutrality tests indicated an excess of divergent, intermediate frequency haplotypes. Because these patterns were observed across all loci, they probably have demographic causes such as population subdivision or recent migration from genetically distinct populations and are not

Table 1
Population Genetic Parameters for the Brevinin1.1 Mature Peptide Region (72 bp) in Three Frog Species

<table>
<thead>
<tr>
<th>Species</th>
<th>2N</th>
<th>Lineage 1</th>
<th>Lineage 2</th>
<th>Lineage 3</th>
<th>S</th>
<th>h</th>
<th>Hd</th>
<th>pi</th>
<th>pi_s</th>
<th>ZnS</th>
<th>pi/pi_s</th>
<th>Tajima’s D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rana pipiens</td>
<td>848</td>
<td>60%</td>
<td>34%</td>
<td>6%</td>
<td>16</td>
<td>8*</td>
<td>0.55</td>
<td>0.06</td>
<td>0.009</td>
<td>0.32*</td>
<td>8.67**</td>
<td>2.23*</td>
</tr>
<tr>
<td>Rana blairi</td>
<td>100</td>
<td>92%</td>
<td>8%</td>
<td>0%</td>
<td>12</td>
<td>4**</td>
<td>0.24*</td>
<td>0.02</td>
<td>0.001</td>
<td>0.48*</td>
<td>25.00**</td>
<td>-0.81</td>
</tr>
<tr>
<td>Rana palustris</td>
<td>86</td>
<td>0%</td>
<td>0%</td>
<td>100%</td>
<td>0</td>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.000</td>
<td>0.00</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

NOTE.—For all statistical tests, *P < 0.05 and **P < 0.01.

2N = number of alleles = twice the number of individuals.
Lineages 1, 2, and 3 refer to the percentage of alleles in each species belonging to each of these three allelic lineages.
S, number of segregating sites.
h, number of haplotypes; tested whether h is significantly low given S.
Hd, haplotype diversity, equivalent to expected heterozygosity; tested whether Hd is significantly low given S.
pi, mean number of pairwise differences among sequences; tested whether pi is significantly extreme given S.
pi_s, mean number of pairwise differences among sequences at silent sites; tested whether pi_s is significantly extreme given silent S.
ZnS, linkage disequilibrium statistic of Kelly (1997); tested whether ZnS is significantly high given S.
pi_s/pi_s, ratio of nonsynonymous nucleotide variation to synonymous nucleotide variation; tested whether significantly different than 1.
Tajima’s D, statistic of Tajima (1989); tested whether significantly different than 0.

Fig. 4.—Maximum parsimony haplotype networks for four putatively neutral non-AMP nuclear loci unlinked to Brevinin1.1: Arcadlin, Myosin, FIBI7, and Tyrosinase. Symbols are the same as in figure 3. In contrast with the Brevinin1.1 locus (fig. 3), nonsynonymous substitutions (small black rectangular bars) are rarer than synonymous substitutions (small white rectangular bars). In both Rana pipiens (white circles) and Rana blairi (gray circles), non-AMP loci show intermediate frequency alleles and few missing transitional haplotypes (gray squares), in contrast with the Brevinin1.1 locus (fig. 3), which shows high- and low-frequency alleles forming divergent lineages separated by many missing transitional haplotypes. In Rana palustris (black circles), non-AMP loci show diverse, moderately divergent alleles, in contrast with the Brevinin1.1 locus (fig. 3), which shows very low diversity.
Table 2
Population Genetic Parameters for Four Putatively Neutral Nuclear Loci Unlinked to *Brevinin1.1* (1,246 bp) in Three Frog Species (*Arcadlin, Myosin, FIBI7*, and *Tyrosinase*)

<table>
<thead>
<tr>
<th>Species</th>
<th>2N&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Locus</th>
<th>Length&lt;sup&gt;b&lt;/sup&gt;</th>
<th>S&lt;sup&gt;c&lt;/sup&gt;</th>
<th>h&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Hd&lt;sup&gt;e&lt;/sup&gt;</th>
<th>p&lt;sup&gt;f&lt;/sup&gt;</th>
<th>π&lt;sup&gt;g&lt;/sup&gt;</th>
<th>ZnS&lt;sup&gt;h&lt;/sup&gt;</th>
<th>πs/π&lt;sub&gt;n&lt;/sub&gt;&lt;sup&gt;i&lt;/sup&gt;</th>
<th>Tajima’s D&lt;sup&gt;j&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Rana pipiens</em></td>
<td>162</td>
<td>Arcadlin</td>
<td>265</td>
<td>6</td>
<td>9</td>
<td>0.75</td>
<td>0.01</td>
<td>0.010</td>
<td>0.06</td>
<td>0.04**</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Myosin</td>
<td>150</td>
<td>7</td>
<td>10</td>
<td>0.76</td>
<td>0.01</td>
<td>0.020</td>
<td>0.12</td>
<td>0.00*</td>
<td>1.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FIBI7</td>
<td>258</td>
<td>12</td>
<td>9</td>
<td>0.57</td>
<td>0.01</td>
<td>0.010</td>
<td>0.24</td>
<td>0.00</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tyrosinase</td>
<td>573</td>
<td>17</td>
<td>17</td>
<td>0.84</td>
<td>0.00</td>
<td>0.016</td>
<td>0.06</td>
<td>0.08*</td>
<td>−0.41</td>
</tr>
<tr>
<td>Mean ± standard deviation</td>
<td>312 ± 182</td>
<td>Arcadlin</td>
<td>10.5 ± 5.1</td>
<td>11.3 ± 3.9</td>
<td>0.73 ± 0.11</td>
<td>0.01 ± 0.00</td>
<td>0.014 ± 0.005</td>
<td>0.12 ± 0.08</td>
<td>0.03**</td>
<td>0.38 ± 0.71</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Myosin</td>
<td>162</td>
<td>2</td>
<td>3</td>
<td>0.46</td>
<td>0.00</td>
<td>0.004</td>
<td>0.02</td>
<td>0.00*</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FIBI7</td>
<td>100</td>
<td>2</td>
<td>3</td>
<td>0.50</td>
<td>0.00</td>
<td>0.006</td>
<td>0.04</td>
<td>0.00</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tyrosinase</td>
<td>100</td>
<td>3</td>
<td>4</td>
<td>0.56</td>
<td>0.00</td>
<td>0.000</td>
<td>0.00</td>
<td>0.00</td>
<td>−1.15</td>
</tr>
<tr>
<td>Mean ± standard deviation</td>
<td>312 ± 182</td>
<td>Arcadlin</td>
<td>2.3 ± 0.5</td>
<td>3.3 ± 0.5</td>
<td>0.40 ± 0.21</td>
<td>0.00 ± 0.00</td>
<td>0.002 ± 0.003</td>
<td>0.02 ± 0.02</td>
<td>0.14</td>
<td>−0.02 ± 0.79</td>
<td></td>
</tr>
<tr>
<td><em>Rana blairi</em></td>
<td>100</td>
<td>Arcadlin</td>
<td>265</td>
<td>4</td>
<td>4</td>
<td>0.56</td>
<td>0.00</td>
<td>0.009</td>
<td>0.33</td>
<td>0.00**</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Myosin</td>
<td>150</td>
<td>4</td>
<td>2*</td>
<td>0.34</td>
<td>0.01</td>
<td>0.010</td>
<td>1.00*</td>
<td>0.64</td>
<td>1.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FIBI7</td>
<td>258</td>
<td>14</td>
<td>7</td>
<td>0.72</td>
<td>0.02**</td>
<td>0.024**</td>
<td>0.37*</td>
<td>0.00*</td>
<td>2.53**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tyrosinase</td>
<td>573</td>
<td>7</td>
<td>5</td>
<td>0.75</td>
<td>0.00</td>
<td>0.011</td>
<td>0.38</td>
<td>0.07**</td>
<td>0.56</td>
</tr>
<tr>
<td>Mean ± standard deviation</td>
<td>312 ± 182</td>
<td>Arcadlin</td>
<td>7.3 ± 4.7</td>
<td>4.5 ± 2.1**</td>
<td>0.59 ± 0.19</td>
<td>0.01 ± 0.01**</td>
<td>0.014 ± 0.007**</td>
<td>0.52 ± 0.32**</td>
<td>0.13**</td>
<td>1.39 ± 0.83**</td>
<td></td>
</tr>
</tbody>
</table>

**NOTE.**—For all statistical tests, *P* < 0.05 and **P** < 0.01.

<sup>a</sup> 2N = number of alleles = twice the number of individuals.
<sup>b</sup> Length, locus length in base pairs.
<sup>c</sup> S, number of segregating sites.
<sup>d</sup> h, number of haplotypes; tested whether h is significantly low given S.
<sup>e</sup> Hd, haplotype diversity, equivalent to expected heterozygosity; tested whether Hd is significantly low given S.
<sup>f</sup> p, mean number of pairwise differences among sequences; tested whether p is significantly extreme given S.
<sup>g</sup> π, mean number of pairwise differences among sequences; tested whether π is significantly extreme given S.
<sup>h</sup> ZnS, linkage disequilibrium statistic of Kelly (1997); tested whether ZnS is significantly high given S.
<sup>i</sup> πs/π<sub>n</sub>, ratio of nonsynonymous nucleotide variation to synonymous nucleotide variation; tested whether significantly different than 1.
<sup>j</sup> Tajima’s D, statistic of Tajima (1989); tested whether significantly different than 0.
due to selection. There were no fixed differences between *R. blairi* and *R. pipiens*. There were eight fixed differences between *R. palustris* and *R. blairi*; of these, four were also fixed between *R. palustris* and *R. pipiens*.

Neutrality Tests

To test whether the high allelic variation at the *Brevinin1.1* locus in *R. pipiens* and *R. blairi* was maintained by balancing selection, we conducted multiple tests of selective neutrality (table 1; supplementary table 1, Supplementary Material online). Neutrality tests were not conducted on *R. palustris* because that species had no variation in the mature peptide region. The ratio of mature peptide region nonsynonymous variation to synonymous variation \((\pi_d/\pi_s)\) was significantly greater than 1 in most populations as well as in *R. pipiens* overall and *R. blairi* overall (table 1; supplementary table 1, Supplementary Material online; fig. 5). For both species, nonsynonymous variation in the mature peptide region was also significantly higher than synonymous variation at *Arcadlin, Myosin, FIBI7*, and *Tyrosinase* (fig. 5). Coalescent simulations of several population genetic parameters in DnaSP demonstrated that polymorphisms were structured into divergent allelic lineages, with skewed frequencies, more so than would be expected under neutrality given the number of segregating sites (table 1; supplementary table 1, Supplementary Material online; fig. 3). Given the number of segregating sites, the number of haplotypes \((h)\) was too low for all populations that had any variation and the haplotype diversity \((Hd)\) was too low for nearly all populations. Linkage disequilibrium \((ZnS)\) was significantly higher than neutral expectations for all populations that had any variation. In *R. pipiens*, the presence of several common, divergent alleles made pairwise variation \((\pi)\) significantly higher than expected given the number of segregating sites, resulting in significantly positive values of Tajima’s *D*. In contrast, Tajima’s *D* was negative in *R. blairi*, owing to rare but divergent alleles.

Multiple tests rejected neutrality in both *R. pipiens* and *R. blairi*, providing strong evidence for balancing selection. In *R. pipiens*, neutrality was also rejected if the sequences from any one allelic lineage were removed (significant ZnS and *h* values, \(\pi_d/\pi_s > 1\)). This result suggested that all three allelic lineages are adaptively divergent *R. pipiens*. Similarly, in *R. blairi*, if Lineage 2 sequences were removed, multiple tests still rejected neutrality (significant ZnS and *h* values, \(\pi_d/\pi_s > 1\)). This result suggested that alleles *Rb1* and *Rb2*, while both in Lineage 1, might be adaptively divergent.

Distinguishing Stable from Fluctuating Selection

The first indication about the type of balancing selection acting was that allele frequencies were not intermediate but skewed. In *R. pipiens*, Lineage 1 alleles were 10 times more common than Lineage 3 alleles. In *R. blairi*, allele frequencies were even more skewed, with 86% of the sequences consisting of the same allele, and other alleles being quite rare. In *R. palustris*, only a single allelic lineage was observed. The other allelic lineages could be so rare they were not sampled in *R. palustris* or they might not exist in this species. No species showed a pattern of equally frequent allelic lineages consistent with symmetrical models of stable balancing selection. The observed skew was more consistent with fluctuating selection or highly asymmetrical overdominance.

To formally distinguish between stable balancing selection and fluctuating selection, we compared silent variation at four non-AMP nuclear loci (*Arcadlin, Myosin, FIBI7*, and *Tyrosinase*) with silent variation at the *Brevinin1.1* locus (fig. 5). Silent variation at the *Brevinin1.1* locus was lower than at all the non-AMP loci for both *R. pipiens* (Brevinin1.1 *\(\pi_s = 0.005; P = 0.08\)) and *R. palustris* (Brevinin1.1 *\(\pi_s = 0.002; P = 0.02\)*). Silent variation was extremely low at all loci for *R. blairi* (Brevinin1.1 *\(\pi_s = 0.002; P > 0.1\)*). These results were consistent with fluctuating selection, which predicts low synonymous diversity, but not with stable balancing selection, which predicts high synonymous diversity. Although demographic processes appeared to have caused nonneutral patterns at the non-AMP loci *R. palustris*, these processes could not explain the difference in genetic diversity between *Brevinin1.1* and the other loci, which therefore is more likely to be due to selection.

We calculated values of Tajima’s *D* within allelic lineages (table 3). In *R. pipiens*, Tajima’s *D* was significantly negative within the common Lineage 1 \((D = −1.84; P < 0.01)\). Because we did not sequence non-AMP loci in all *R. pipiens* individuals, we cannot completely understand possible demographic factors that could affect Tajima’s *D*, so this result should be interpreted with caution. In both *R. pipiens* and *R. blairi*, there was a nonsignificant trend of other common allelic lineages having negative Tajima’s *D* values and rare allelic lineages having...
positive Tajima’s $D$ values consistent with fluctuating selection (table 3). Because population subdivision can affect Tajima’s $D$, we also calculated within-lineage Tajima’s $D$ values for individual populations of $R$. pipiens and $R$. blairi. For all populations, including those shown to have neutral Tajima’s $D$ values at non-AMP loci (populations 1, 2, 14, and 15; supplementary table 2, Supplementary Material online), the trend was the same as for the whole species (table 3). Only one allelic lineage with a single polymorphism occurred in $R$. palustris, so within-lineage Tajima’s $D$ could not be meaningfully estimated.

Under asymmetrical stable balancing selection, such as asymmetrical overdominance, the more favored allelic lineage will be more common and harbor most of the neutral genetic variation, owing to its larger effective size (Innan and Tajima 1999; Stahl et al. 1999). We therefore compared levels of silent variation among allelic lineages. Within $R$. pipiens, values of silent $\pi$ were 0.0004 ± 0.0011 for Lineage 1, 0.0002 ± 0.0008 for Lineage 2, and 0.0038 ± 0.0037 for Lineage 3. Within $R$. blairi, values of silent $\pi$ were 0.0018 ± 0.0018 for Lineage 1 and 0.0053 ± 0.0042 for Lineage 2. Thus, for both species, genetic variation was relatively low in the more common allelic lineages and much higher in the rarest allelic lineages.

We tested for excessive differentiation among populations, a signature of fluctuating selection. The $F_{ST}$ value among the five easternmost populations of $R$. pipiens was 0.20, and the expected neutral value of $F_{ST}$ at the same level of heterozygosity was approximately 0.06 (fig. 2A in Hoffman et al. [2006]). Thus, the observed $F_{ST}$ value at the Brevinin1.1 locus was higher than the mean neutral expectation but not significantly so. The high $F_{ST}$ value is primarily due to population 2, where allele Rp1 predominates, in contrast to the other four populations, where allele Rp2 predominates (supplementary table 1, Supplementary Material online). In addition, $F_{ST}$ between populations 1 and 2 was 0.02 at the four non-AMP loci sequenced in this study and 0.04 at the microsatellites, but 0.34 at Brevinin1.1. The high divergence of population 2 at Brevinin1.1 but not at microsatellites is shown in figure 6. The matrix of pairwise microsatellite differences among eastern populations was not correlated with the corresponding matrix of Brevinin1.1 pairwise differences (Mantel test; $P > 0.1$). Therefore, the high genetic divergence between population 2 and the other eastern populations is inconsistent with neutral expectations and could be due to spatially differing selective pressures.

The two $R$. pipiens populations with the lowest microsatellite $H_e$ values were fixed for the allele Rp1, in contrast to the high genetic variation seen in other populations (supplementary table 1, Supplementary Material online). This result was consistent with the hypothesis that these two populations had lower effective sizes than the others, in which case genetic drift would be more likely to fix alleles in these two populations. After these two fixed populations, the $R$. pipiens population with the lowest Brevinin1.1 $H_d$ was population 2 (supplementary table 1, Supplementary Material online), which had slightly higher than average microsatellite $H_e$ (0.84), and therefore its effective population size did not appear to be substantially lower than the other populations. Because population 2 had a very different allele frequency pattern than its closest geographic neighboring populations (fig. 6), and because it showed an unusually low Brevinin1.1 $H_d$ value for its microsatellite $H_e$ value, population 2 was the population most likely to have undergone a recent shift in allele frequencies at the Brevinin1.1 locus due to selection. However, allele frequencies were nearly identical between the 1971 sample and the 2001 sample, both at the Brevinin1.1 locus ($F_{ST} = 0.0$; supplementary table 1, Supplementary Material online).
Material online) and at the nine microsatellite loci ($F_{ST} = 0.0$).

**Discussion**

Patterns of genetic diversity at the *Brevinin1.1* locus are highly unusual in three species of leopard frogs, and multiple lines of evidence suggest that strong balancing selection is maintaining adaptive variation in the mature peptide region. Evidence for nonneutrality includes high ratios of nonsynonymous/synonymous variation, unusual values of $Hd$, $h$, and $ZnS$ given the number of segregating sites, and significant values of Tajima’s $D$ (table 1; supplementary table 1, Supplementary Material online). Because we did not sequence non-AMP loci from all *R. pipiens* individuals, we cannot rule out the possibility that demographic factors have contributed to some nonneutral patterns; however, in populations 1 and 2, neutrality is strongly rejected the *Brevinin1.1* locus but not at other loci (supplementary tables 1 and 2, Supplementary Material online), suggesting that our results are not an artifact of demography. Likewise, we cannot rule out the possibility that the high variation at the *Brevinin1.1* locus arose not merely through point mutations but also through gene conversion from other loci, which can contribute to a pattern of strikingly divergent alleles (Storz et al. 2007). Several Lineage 1 alleles, including allele *Rp1*, are identical in the mature peptide region to other brevinin-1 loci that exist in leopard frog genomes (Tennessen and Blouin 2007). Lineage 2 and Lineage 3 sequences have not been observed at other brevinin-1 loci (Tennessen and Blouin 2007). However, even if demographic processes or gene conversion have enhanced the diversity at this locus, they must have acted in concert with selection as this is the only explanation for the elevated mutation rate and/or positive selection on synonymous sites, which would tend to cause the opposite pattern (Tennessen 2005a). Fourth, the allelic lineages with the most variation are not the most common, again suggesting that allele frequencies were quite different in the recent past. Finally, genetic distance at *Brevinin1.1* is exceptionally high among some *R. pipiens* populations, and stable balancing selection would predict a trend in the opposite direction (fig. 6).

Our results provide unique insight into the evolution of immunity genes. Nearly all examples of pathogen-induced balancing selection are detection and signaling proteins, such as those of the vertebrate adaptive immune system (Garrigan and Hedrick 2003). The *Brevinin1.1* locus represents a rare example of an effector gene harboring an adaptive polymorphism. Loci encoding AMPs in mussels and humans also show high variation consistent with balancing selection (Hollox and Armour 2008; Pahdi and Verghese 2008). Drosophila AMPs are not under balancing selection (Clark and Wang 1997; Lazzeri and Clark 2003), possibly because fly AMPs, which also show little evidence for positive selection, may have less specific targets than AMPs in other taxa (Sackton et al. 2007). The AMP locus *Ranatuerin2* has undergone a selective sweep in *R. pipiens* and shows no evidence of balancing selection (Tennessen and Blouin 2007), but this pattern is also consistent with fluctuating selection if the other alleles are extremely rare or have recently been eliminated. Fluctuating selection in particular has only occasionally been convincingly demonstrated for any immunity locus, effector or otherwise. Many of the best examples of fluctuating selection on defense genes are from plants (Stahl et al. 1999; Tiffin et al. 2004; Tiffin and Moeller 2006), but a few animal examples also exist (Jensen et al. 2008). Even for well-studied loci under balancing selection like the genes of the major histocompatibility complex, the importance of stable versus fluctuating selection is not clear (Pierrney and Oliver 2006); the universally observed high allelic diversity at these loci suggests that if fluctuating selection occurs, the fluctuations are nearly instantaneous and, therefore, closer to stable balancing selection than they are at the *Brevinin1.1* locus. Microbes can evolve resistance to detection mechanisms by simply substituting one or more amino acid residues on their surface proteins. In contrast, microbial resistance to AMPs involves changing the biochemistry of the cell membrane or producing AMP-degrading enzymes, adaptations which might involve several coordinated changes across multiple genes. Thus, we hypothesize that it is more difficult for microbes to evolve resistance to AMPs than to detection mechanisms, resulting in longer lag periods between host evolution and pathogen evolution and therefore greater fluctuations in allele frequencies, including occasional allele fixation. Fluctuating selection may be difficult to distinguish from positive selection if divergent, low-frequency alleles are not sampled, which may be partly why documented examples of balancing selection on effector molecules are rare. Overall, our results help to illustrate when and how balancing selection can act on effector loci.
Allele frequency differences among populations are either due to genetic drift or different selective pressures in different habitats. In *R. pipiens*, allele frequencies differ substantially between populations east and west of the Mississippi River (supplementary table 1, Supplementary Material online), but this differentiation is consistent with genetic drift, given the high east–west mitochondrial DNA divergence also observed in this species (Hoffman and Blouin 2004). The high differentiation among the eastern populations at *Brevinin1.1*, despite low differentiation at neutral markers, is more likely to be due to selection (fig. 6). Population 2 is the most likely candidate for a recent population-specific shift in allele frequencies because it is quite divergent from other populations nearby and because its haplotype diversity is low despite having normal levels of neutral genetic diversity. Allele frequencies in this population have not changed noticeably between 1971 and 2001, so if there is a difference in selective pressures between population 2 and its neighbors, it is more than a few decades old.

Variation at AMP loci is likely to be an important contributor to the ability of amphibian species to adapt to novel disease threats. Agents of emerging infectious diseases are causing precipitous declines, and in some cases extinction, in many amphibian species (Carey et al. 1999; Daszak et al. 2003). It is thought that AMPs are an important defense against these pathogens (Chinchar et al. 2004; Rollins-Smith and Conlon 2005; Woodhams, Rollins-Smith, et al. 2006; Woodhams, Voyles, et al. 2006). Several epizootics of the fungal pathogen *Batrachochytrium dendrobatidis* and iridovirus have caused declines of *R. pipiens* populations in recent decades (Carey et al. 1999; Green et al. 2002; Greer et al. 2005). Although these epizootics could be selective agents on the *Brevinin1.1* locus, we did not observe any historic allele frequency change in population 2, and allelic lineages are too divergent to have arisen within a few decades. Therefore, nonneutral patterns of genetic diversity are probably primarily due to long-term allelic fluctuations caused by unidentified native diseases, with recent disease outbreaks possibly shaping this variation.

Our study adds to the growing body of knowledge on diversifying selection among AMPs by demonstrating that AMPs encoded by alleles at the same locus can show evidence of major adaptive differences. We have shown that genes encoding immune effector proteins can harbor balanced polymorphisms comparable to many pathogen-detection genes. Researchers developing AMPs for therapeutic applications (Hancock 2001) should consider examining multiple individuals of the same species to encounter functionally novel allelic variants. Our data are consistent with a model of fluctuating selection, likely caused by pathogens. However, because the results of many of our tests were merely suggestive, we are reluctant to completely reject the hypothesis of stable balancing selection. Further study is required to determine the specific functional difference among these peptides, the particular pathogens that are driving selection, and the precise mechanism maintaining variation. We are currently examining the antimicrobial effects of these peptides in vitro (Tennessen JA, Woodhams DC, Reinert LK, Blouin MS, and Rollins-Smith LA, unpublished data). Loci encoding AMPs, such as *Brevinin1.1*, can be important sites of adaptive genetic diversity and major players in the coevolutionary arms race between hosts and pathogens.

### Supplementary Material

Supplementary tables 1 and 2 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

### Acknowledgments

The authors are grateful to Eric Hoffman, Eric Tobin, Mike Redmer, Stephan Swanson, Mike Foy, and the Museum of Vertebrate Zoology for help obtaining samples. Becky Cooper provided invaluable laboratory assistance. Hitoshi Araki, Kaitlin Bonner, and Mark Christie commented on earlier versions of the manuscript. We also thank three anonymous reviewers for their helpful comments. This research was supported by an Environmental Protection Agency Science to Achieve Results fellowship to J.A.T. and an Idea Network of Biomedical Research Excellence grant (2 P20 RR016463) to the Nevada Genomics Center.

### Literature Cited


Scott Edwards, Associate Editor

Accepted September 11, 2008