A new species of leopard frog (Anura: Ranidae) from the urban northeastern US

Catherine E. Newman, Jeremy A. Feinberg, Leslie J. Rissler, Joanna Burger, H. Bradley Shaffer

1. Introduction

Leopard frogs of the Rana pipiens (=Lithobates pipiens) complex are widespread and common throughout much of the United States, but species delimitation and the associated taxonomy of the group have been challenging and contentious (Brown, 1973; Pace, 1974; Moore, 1975; Brown et al., 1977, 1990; Zug et al., 1982; Hillis, 1988; Frost et al., 2006, 2008, 2009; Pauly et al., 2009). While studies of range-wide phylogeography and systematics at the genus and species level are common (e.g., Pace, 1974; Hillis et al., 1983; Pytel, 1986; Hoffman and Blouin, 2004; Hillis and Wilcox, 2005; Oláh-Hemmings et al., 2010; Newman and Rissler, 2011), relatively little attention has been focused on taxonomic status and conservation needs of local or regional populations or subspecies (but see Di Candia and Routman, 2007; Hekkala et al., 2011). As is true for any group, appropriate conservation measures cannot be identified and implemented in the face of uncertain taxonomy (Köhler et al., 2005).

The species composition of leopard frogs in parts of the mid-Atlantic and northeastern US—hereafter the Tri-State area, including New Jersey (NJ), New York (NY), and Connecticut (CT)—has been questioned by biologists over the past several decades (Kaufeld, 1937; Yeaton, 1968; Schlauch, 1971; Pace, 1974; Klemens et al., 1987; Klemens, 1993). Currently, two species are recognized in the region (Conant and Collins, 1998). Rana pipiens, the northern leopard frog, is widely distributed across New England and the Great Lakes region, including the western two-thirds of CT and central and northern NY. From NJ, Long Island (NY), and southern mainland NY to the south, it is replaced by R. sphenocephala (=L. sphenocephalus), the southern leopard frog. While natural history collection data suggest the two species have a narrow zone of overlap in southern NY (Fig. 1), no area of sympatry has been directly identified. Some earlier studies based on morphological data suggested the possibility of intergradation (Schlauch, 1971), whereas others speculatively discussed a putative third species in this region (Kaufeld, 1937; Klemens, 1993).

Although widespread and often common at the continental scale (Fig. 1), leopard frog populations have been severely declining in certain regions, resulting in extirpation from some portions

© 2012 Elsevier Inc. All rights reserved.
of their historical range (Lannoo, 2005), including coastal regions and islands north and east of Long Island, NY (Ditmars, n.d.; Lattham, 1971; Klemens, 1993; Feinberg, et al., unpublished data). Leopard frogs are also believed to be extirpated from highly developed areas including Long Island, NY (Kiviat, 2010; Feinberg et al., unpublished data); New Haven, CT; and Providence, Rhode Island (Klemens, 1993). While the exact causes of these declines are unclear, environmental pesticides and endocrine disruptors (Hayes et al., 2003; Lannoo, 2008), disease (Carey et al., 1999; Greer et al., 2005; Davis et al., 2007; Searle et al., 2011), habitat loss and alteration (Lannoo, 2005), and over-harvesting for use as laboratory specimens (Hillis, 1988; Klemens, 1993; Lannoo, 2005) have all been identified as contributing factors, particularly regarding *R. pipiens*. *Rana sphenocephala*, in contrast, remains relatively abundant throughout most of its range to the south, including coastal islands south of Long Island. However, near its northern range limit, it is listed as a Species of Special Concern in NY (NY Department of Environmental Conservation) and as endangered in Pennsylvania (PA) (Pennsylvania Fish and Boat Commission).

To gain a better understanding of the status and distributions of leopard frog populations in the Tri-State area, we analyzed mitochondrial and nuclear gene sequences from four focal populations of unknown leopard frog species composition in northern NJ, southeastern mainland NY (two populations), and Staten Island, NY (one of the five boroughs of New York City). Direct observations by one of us (JAF) showed that these four populations exhibited several unique characteristics, including an advertisement call distinct from both *R. pipiens* and *R. sphenocephala*. We also analyzed three CT populations from localities within the traditionally accepted geographic range of *R. pipiens*. We evaluated three possible interpretations of the status of leopard frogs in the Tri-State area: (1) the four focal populations are conspecific with either *R. pipiens* or *R. sphenocephala*, (2) the populations are hybrids between *R. pipiens* and *R. sphenocephala*, or (3) the populations represent a previously undescribed leopard frog lineage distinct from *R. pipiens* and *R. sphenocephala*.

2. Materials and methods

2.1. Study area and sample collection

Our study region was focused on the Tri-State area of the northeastern US, including NY, NJ, and CT—a total area of roughly 40,000 km² (Fig. 1). The region includes an area of putative range overlap between *R. sphenocephala* and *R. pipiens* according to range maps downloaded from the IUCN [IUCN Red List of Threatened Species 2011.1 (http://www.iucnredlist.org)]. Our study included four focal populations of unknown leopard frog species composition: Great Swamp (NJ), Staten Island (NY), Putnam County (NY), and Orange County (NY) (Fig. 1). The Great Swamp and Staten Island sites fall within the geographic range of *R. sphenocephala* and outside the range of *R. pipiens*, whereas the Putnam and Orange sites fall in the overlap zone of the two species’ ranges. Leopard frog species composition in CT has also been questioned (Klemens, 1993), so we collected samples from three sites across CT to include in the analyses (Fig. 1).
Toe clips were taken from 3 to 10 individual frogs at each of the four focal sites and three populations in CT, as well as control sites for R. sphenoecephala in southern NJ and R. pippins in northeastern mainland NY (Fig. 1; Table 1). In addition, three morphologically ambiguous specimens from Long Island were included to determine if they represented an isolated relict population of leopard frogs, or if they were instead the pickerel frog R. palustris (=L. palustris) (Table 1). Three CT R. palustris specimens from the Yale Peabody Museum were also included as reference samples. Tissues were stored in 98% ethanol, and source frogs were measured, photographed, and subsequently released, or collected as vouchers to be deposited in either the Yale Peabody Museum or the University of Alabama Herpetological Collection.

2.2. DNA extraction and gene amplification

Genomic DNA was extracted at the University of California, Davis, using a standard salt extraction protocol. We amplified the ND2 and 12S–16S regions of the mitochondrial genome, including the intervening tRNA-Valine and partial flanking tRNA-Tryptophan sequences, for a total of 1444 bp. We also amplified the neurotrophin-3 (NTF3, 599 bp), tyrosinase (Tyr, 557–585 bp), Rag-1 (647–683 bp), seven-in-absentia (SIA, 362–393 bp), and chemokine receptor 4 (CXCR4, 550 bp) regions of the nuclear genome. All primer references and sequences are provided in Supplementary Table S1.

PCR amplification was performed in 18 µL reactions, consisting of 1.5 µL PCR Buffer II (10X, Applied Biosystems), 2.4 µL MgCl2 (25 mM), 0.6 µL each primer (5 mM), 2.4 µL dNTP solution (5 mM), 1U AmpliTaq (Applied Biosystems), and 10–30 ng genomic DNA. All gene regions except 12S–16S were amplified using the following PCR protocol: initial denaturation at 95°C for 10 min.; 38 cycles of 94°C for 30 s, 63–65°C (see Supplementary Table S1) for 45 s, 72°C for 1 min.; and a final extension at 72°C for 10 min. The amplification protocol for the 12S–16S region was as follows: initial denaturation at 94°C for 30 s; 35 cycles of 94°C for 45 s, 52°C for 30 s, 72°C for 1 min. and a final extension at 72°C for 7 min. PCR products were sequenced in the forward and reverse directions at Beckman Coulter Genomics (Danvers, MA, USA). Contigs were assembled in Geneious v.5.3.6 (Drummond et al., 2011). Sequence fragments were trimmed to minimize missing data.

2.3. Mitochondrial sequence analysis

The 12S–16S and ND2 sequence fragments with associated tRNA fragments were concatenated and aligned using ClustalW in Geneious and manually adjusted. All sequences were uploaded to GenBank (see Supplementary Table S2 for accession numbers). Bayesian analyses were conducted in MrBayes 3.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) with five partitions: 12S plus tRNA-Val, 16S, and each of the three ND2 codon positions. Based on output from jModelTest v.0.1.1 (Guindon and Gascuel, 2003; Posada, 2008) and convergence analyses of trial runs (data not shown), the 12S and ND2 partitions were assigned a GTR model of evolution, and the 16S partition was assigned an HKY model of evolution. The 12S partition allowed across-site rate variation under a gamma distribution, and rates were allowed to vary among partitions. Bayesian analyses were run with random starting trees, two simultaneous runs of 10 million generations, and sampling from the posterior distribution of trees every 5000 generations. Tracer v.1.4.1 (Rambaut and Drummond, 2007) was used to assess convergence and to determine appropriate burn-in. The first 25% of samples were omitted as burn-in. Nodal support was further assessed with a maximum-likelihood (ML) analysis in RaxML v.7.0.3 (Stamatakis, 2006; Stamatakis et al., 2008) with 1000 bootstraps. Rana clamitans sequences (DQ347036, 12S–16S; AY206480, ND2) were downloaded from GenBank and used as an outgroup. Tajima’s D and Fu’s F3 were calculated in Arlequin v.3.5 (Excoffier and Lischer, 2010) for test selection.

2.4. Nuclear sequence analysis

For each locus, sequences were aligned using ClustalW in Geneious and manually adjusted, and sequences were uploaded to GenBank (Supplementary Table S2). Phylogenies were reconstructed for each locus individually and for the concatenated data set using unphased sequences (see below) in MrBayes. For the individual gene trees, models of evolution, based on jModelTest output and preliminary runs (data not shown), were as follows: HKY for CXCR4; HKY + G for NTF3, Rag-1, and Tyr; and JC for SIA. The concatenated data set was partitioned by locus, but a consistent lack of convergence suggested that this model was inappropriate for our data (results not shown). Bayesian analyses were thus run on the entire, unpartitioned nuclear data set, with an HKY model of evolution [based on jModelTest and trial runs (data not shown)]. All analyses were run for 10 million generations and sampled every 5000 generations. Convergence was assessed in Tracer, and the first 25% of samples were omitted as burn-in. Nodal support was further assessed with 1000 ML bootstraps in RaxML. Tests for selection were done in Arlequin, using Tajima’s D and Fu’s F3 statistics.

Table 1

<table>
<thead>
<tr>
<th>Population</th>
<th>Sample size</th>
<th>Species (a priori)</th>
<th>Map code</th>
<th>County</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focal unknown populations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staten Island</td>
<td>6</td>
<td>Unknown</td>
<td>1</td>
<td>Richmond</td>
<td>New York</td>
</tr>
<tr>
<td>GSNWR</td>
<td>5</td>
<td>Unknown</td>
<td>2</td>
<td>Morris</td>
<td>New Jersey</td>
</tr>
<tr>
<td>Putnam</td>
<td>5</td>
<td>Unknown</td>
<td>3</td>
<td>Putnam</td>
<td>New York</td>
</tr>
<tr>
<td>Orange</td>
<td>3</td>
<td>Unknown</td>
<td>4</td>
<td>Orange</td>
<td>New York</td>
</tr>
<tr>
<td>Unknown populations in connecticut</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hartford</td>
<td>5</td>
<td>Unknown</td>
<td>5</td>
<td>Hartford</td>
<td>Connecticut</td>
</tr>
<tr>
<td>Middlesex</td>
<td>5</td>
<td>Unknown</td>
<td>6</td>
<td>Middlesex</td>
<td>Connecticut</td>
</tr>
<tr>
<td>Litchfield</td>
<td>10</td>
<td>Unknown</td>
<td>7</td>
<td>Litchfield</td>
<td>Connecticut</td>
</tr>
<tr>
<td>Unknown populations on long island</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eastport</td>
<td>3</td>
<td>Unknown</td>
<td>10</td>
<td>Suffolk</td>
<td>New York</td>
</tr>
<tr>
<td>Control populations (known species)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRSP</td>
<td>5</td>
<td>R. sphenoecephala</td>
<td>8</td>
<td>Burlington</td>
<td>New Jersey</td>
</tr>
<tr>
<td>Saratoga</td>
<td>6</td>
<td>R. pippins</td>
<td>9</td>
<td>Saratoga</td>
<td>New York</td>
</tr>
<tr>
<td>Litchfield</td>
<td>1</td>
<td>R. palustris</td>
<td>7</td>
<td>Litchfield</td>
<td>Connecticut</td>
</tr>
<tr>
<td>New London</td>
<td>1</td>
<td>R. palustris</td>
<td>11</td>
<td>New London</td>
<td>Connecticut</td>
</tr>
<tr>
<td>Fairfield</td>
<td>1</td>
<td>R. palustris</td>
<td>12</td>
<td>Fairfield</td>
<td>Connecticut</td>
</tr>
</tbody>
</table>
To test our hypotheses concerning the status of the four unknown populations, we used a Bayesian approach implemented in Structure v.2.3.3 (Pritchard et al., 2000; Falush et al., 2003) with an allelic data set (6% missing data) generated from our nuclear sequence data. We used the software Phase v.2.1 (Stephens et al., 2001; Stephens and Donnelly, 2003) to infer haplotypes for each locus in the five-locus sequence data set using a Bayesian algorithm. Each allele represented a single haplotype. Input files for Phase were generated from alignment nexus files using a Perl script (RC Thomson, unpublished).

Structure was used to determine the number of genetically distinct clusters (K) of samples. We implemented the admixture model (Pritchard et al., 2000), assumed correlation of allele frequencies among clusters (Falush et al., 2003), and assumed no other a priori population information. We tested values of K from 1 to 10. For each K, 20 iterations were run, each consisting of 100,000

**Fig. 2.** Bayesian phylogeny for concatenated mtDNA (12S–16S and ND2). Nodal support: Bayesian posterior probabilities/maximum-likelihood bootstrap values. Tip labels correspond to Supplementary Table S2. Clade symbols correspond to Fig. 1.
generations after a burn-in of 10,000 generations. The appropriate value of $K$ was determined by assessing the posterior probabilities (Pritchard et al., 2000) and $\Delta K$ values following Evanno et al. (2005). An individual was considered of mixed ancestry if its cluster membership probability $q$ was between 0.10 and 0.90 (Vähä and Primmer, 2005).

For both mitochondrial and nuclear loci, measures of sequence divergence (uncorrected $p$), nucleotide diversity ($\pi$) and haplotype diversity (Hd) were determined at the species level using either DnaSP v.5.10.01 (Librado and Rozas, 2009) or Arlequin. Pairwise $F_{ST}$ values were calculated in Arlequin from the concatenated, phased nuclear sequence data set.

3. Results

3.1. mtDNA phylogenetic analyses

The concatenated mtDNA data set consisted of 1461 bp and 15 unique haplotypes. Bayesian analyses of mtDNA revealed four distinct clades, three of which correspond to the known species *R. sphenophala*, *R. pipiens*, and *R. palustris* (Fig. 2). The three samples from Long Island fell out with the *R. palustris* reference samples, rejecting the hypothesis that those frogs represented a relict population of leopard frogs on Long Island. All specimens from the four focal populations and three of five specimens from Middlesex, CT, formed a clade (hereafter *Rana* sp. nov.) distinct from *R. sphenophala*, *R. pipiens*, and *R. palustris*. All other CT specimens grouped with *R. pipiens*. All four clades were strongly supported with Bayesian posterior probabilities (all 1.0) and ML bootstraps (all >99). Importantly, the sister group to the *R. sp. nov. clade, with reasonably strong statistical support (Bayesian posterior probability = 1.0, ML bootstrap support = 0.78) is the pickerel frog *R. palustris* rather than *R. pipiens* or *R. sphenophala*. Species-level $\pi$ and Hd values are listed in Table 2. Pairwise sequence divergence between *R. sp. nov. and the three recognized species were 6.79% (*R. palustris*), 11.0% (*R. sphenophala*), and 12.5% (*R. pipiens*), and pairwise divergence between the latter three described species ranged from 11.1% to 13.4% (Table 3). These data indicate that a differentiated lineage, distinct from *R. sphenophala*, *R. pipiens*, and *R. palustris*, occurs in the region and may represent a previously unrecognized species if additional data confirm these mtDNA results.

The three outlier specimens from Middlesex, CT, are more difficult to interpret. Phylogenetically, they fell in the *R. sp. nov. clade, but their geographic location substantially farther east than all other *R. sp. nov. samples and, more importantly, their sympathy with *R. pipiens* at the same site made us question whether they represent a natural population of *R. sp. nov. farther east than expected or human-mediated introductions. Given this uncertainty, we omitted these individuals from population genetic calculations, pending future sampling in CT, particularly the region between Middlesex County and the closest known *R. sp. nov. population roughly 95 km due west in Putnam County, NY.

3.2. Nuclear phylogenetic analyses

Aligned sequence lengths for nuclear loci were 550 bp (CXCR4), 599 bp (NTF3), 683 bp (Rag-1), 393 bp (SIA), and 585 bp (Tyr). The concatenated data set consisted of 2810 bp of aligned, trimmed sequence. The number of variable sites for each locus ranged from 10 to 30 (Table 2). Species-level $\pi$ and Hd values are listed in Table 2,

<p>| Table 2 |
| Species-level general descriptive statistics. Length = aligned sequence length (bp), $#VS =$ number of variable sites, Hd = haplotype diversity, $\pi =$ nucleotide diversity, $h = $ number of pairwise nucleotide differences, $d_s =$ Watterson’s estimator of genetic diversity. |</p>
<table>
<thead>
<tr>
<th>Descriptive statistics</th>
<th>Tests of neutrality</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>mtDNA</strong></td>
<td><strong>Tajima’s $D$</strong></td>
</tr>
<tr>
<td>R. sphenophala</td>
<td>1.09</td>
</tr>
<tr>
<td>R. palustris</td>
<td>1.05</td>
</tr>
<tr>
<td>R. pipiens</td>
<td>0.2</td>
</tr>
<tr>
<td>Rana sp. nov.</td>
<td>0.22</td>
</tr>
<tr>
<td>CXCR4</td>
<td>1.12</td>
</tr>
<tr>
<td>R. sphenophala</td>
<td>1.3</td>
</tr>
<tr>
<td>R. palustris</td>
<td>1.066</td>
</tr>
<tr>
<td>R. pipiens</td>
<td>1.0</td>
</tr>
<tr>
<td>Rana sp. nov.</td>
<td>1.18</td>
</tr>
<tr>
<td>NTF3</td>
<td>1.15</td>
</tr>
<tr>
<td>R. sphenophala</td>
<td>0.87</td>
</tr>
<tr>
<td>R. palustris</td>
<td>0.429</td>
</tr>
<tr>
<td>R. pipiens</td>
<td>0.273</td>
</tr>
<tr>
<td>Rana sp. nov.</td>
<td>0.263</td>
</tr>
<tr>
<td>Rag-1</td>
<td>1.325</td>
</tr>
<tr>
<td>SIA</td>
<td>1.325</td>
</tr>
<tr>
<td>Tyrosinase</td>
<td>0.492</td>
</tr>
<tr>
<td>R. sphenophala</td>
<td>0.429</td>
</tr>
<tr>
<td>R. palustris</td>
<td>0.439</td>
</tr>
<tr>
<td>R. pipiens</td>
<td>0.165</td>
</tr>
<tr>
<td>Rana sp. nov.</td>
<td>0.163</td>
</tr>
<tr>
<td>1.349</td>
<td>0.793</td>
</tr>
<tr>
<td>R. sphenophala</td>
<td>0.929</td>
</tr>
<tr>
<td>R. palustris</td>
<td>0</td>
</tr>
<tr>
<td>R. pipiens</td>
<td>0.803</td>
</tr>
<tr>
<td>Rana sp. nov.</td>
<td>0.767</td>
</tr>
<tr>
<td>0.454</td>
<td>0.94</td>
</tr>
<tr>
<td>0.886</td>
<td>0.886</td>
</tr>
</tbody>
</table>
and pairwise FST in Table 4. Tajima’s D and Fu’s Fs tests for selection were non-significant for all loci (Table 2), indicating that all sampled loci were selectively neutral.

Analyses of individual nuclear loci (Fig. 3) revealed varying degrees of support for the four species recovered in the mtDNA analysis (Fig. 2). Monophyly of R. palustris was strongly supported by four loci, R. pipiens by two loci, and Rana sp. nov. by one locus. None of the loci supported a monophyletic R. sphenecophala. Importantly, none of the loci recovered strong clade support for non-monophyly of any of the species. In other words, no strongly supported clade contained individuals of multiple species.

Bayesian analysis of the concatenated data set recovered three strongly supported clades corresponding to the three known species (R. sphenecophala, R. pipiens, R. palustris) (Fig. 4), although their interrelationships were unresolved. The remainder of the samplesthat formed the R. sp. nov. mtDNA clade constituted an unresolved collection of samples that were excluded from all three currently recognized species. While we acknowledge the problems associated with phylogenetic analyses of concatenated nuclear data sets (e.g., Kubatko and Degnan, 2007), we emphasize the concordance among the delimitations in our mitochondrial (Fig. 2) and concatenated nuclear (Fig. 4) phylogenies, as well as the Structure analysis (Fig. 5, see below).

The number of inferred haplotypes per locus ranged from 10 to 19. Bayesian cluster analyses in Structure recovered four clusters (lnL = -504.0, AK = 224.13) consistent with the phylogenetic analyses (Fig. 5). As in the mtDNA analyses, R. sphenecophala grouped together in one cluster, R. palustris reference samples grouped with the Long Island specimens in a second cluster, all specimens from CT except three from Middlesex grouped with R. pipiens controls, and all specimens from the four focal populations grouped with three from Middlesex, CT, in a fourth cluster (R. sp. nov.).

Intraspecific and pairwise percent sequence divergence (uncorrected p, 3SD, for mtDNA).

<table>
<thead>
<tr>
<th></th>
<th>R. sphenecophala (%)</th>
<th>R. palustris (%)</th>
<th>R. pipiens (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R. sphenecophala</td>
<td>0.11</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R. palustris</td>
<td>0.08</td>
<td>11.1</td>
<td>-</td>
</tr>
<tr>
<td>R. pipiens</td>
<td>0.43</td>
<td>13.4</td>
<td>12.8</td>
</tr>
<tr>
<td>Rana sp. nov.</td>
<td>0.04</td>
<td>11.0</td>
<td>6.8</td>
</tr>
</tbody>
</table>

Table 4
Species-level pairwise FST, based on phased nuDNA.

<table>
<thead>
<tr>
<th></th>
<th>R. sphenecophala</th>
<th>R. palustris</th>
<th>R. pipiens</th>
</tr>
</thead>
<tbody>
<tr>
<td>R. sphenecophala</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R. palustris</td>
<td>0.661</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R. pipiens</td>
<td>0.463</td>
<td>0.627</td>
<td>-</td>
</tr>
<tr>
<td>Rana sp. nov.</td>
<td>0.423</td>
<td>0.695</td>
<td>0.536</td>
</tr>
</tbody>
</table>

4. Discussion

4.1. Taxonomic status and geographic distribution of Rana sp. nov

Our data strongly support the recognition of three evolutionary lineages of leopard frogs in the Tri-State area, with the four focal populations collectively forming a new, previously undescribed leopard frog species (R. sp. nov.). Phylogenetic and cluster analyses revealed the four unknown populations to be a distinct group from all locally occurring, recognized leopard frog species, rejecting the hypotheses that those populations are conspecific with one or more of the known species or that they are admixed, intergrade populations. Mitochondrial pairwise sequence divergences between R. sp. nov. and the currently recognized species ranged from 6.79% to 12.9%, consistent with or greater than divergence estimates among other ranid species (Jaeger et al., 2001; Shaffer et al., 2004; Di Candia and Routman, 2007; Funk et al., 2008; Oláh-Hemmings et al., 2010). These high levels of divergence strongly suggest a lack of gene flow between R. sp. nov. populations and other leopard frog species, and cluster analysis indicated that none of the samples were of admixed ancestry.

Empirical methods for species delimitation (Sites and Marshall, 2004) could potentially add support to our conclusions. In addition, new methods have recently become available that use Bayesian analyses of multilocus sequence data to concurrently estimate the species tree and delimit species (O’Meara, 2010; Niewieler et al., 2011). We argue that such analyses are not necessary in our case, however, because species delimitation is relatively straightforward given the data herein. The older species are, the more time they have had to accumulate various evidences of lineage divergence, such as diagnosable morphological characters, reproductive isolation, or reciprocal monophyly (de Queiroz, 2007; Shaffer and Thomson, 2007). In our study, genetic data suggest monophyly of each of the four species, and the sympatry of R. pipiens and Rana sp. nov. in Middlesex, CT, suggests some extent of reproductive isolation between the two. Together, reciprocal monophyly and reproductive isolation strongly indicate the reality of independently evolving lineages, which we designate as distinct species.

Based on our current, relatively sparse sampling, R. sp. nov. is restricted to northern NJ, extreme southeastern mainland NY, and Staten Island (Fig. 1), although range limits may extend as far as CT and northeastern PA (Pace, 1974). Three samples from Middlesex County, CT, suggest that the range potentially extends into the western half of CT, where R. sp. nov. is currently sympatric with R. pipiens. Additional sampling in western CT should help to clarify the range extent of R. sp. nov. However, we reiterate that our results show no evidence of hybridization between R. sp. nov. and either of the other two leopard frog species in the region, including central CT where R. sp. nov. and R. pipiens occur in sympathy, suggesting some level of reproductive isolation.

4.2. Conservation implications and recommendations

The geographic extent of R. sp. nov. is limited to a small portion of NJ, NY, and possibly CT and PA (Fig. 1). This northeastern endemic distribution is concordant with few other amphibian taxa (but see Pseudacris kalmii; Lemmon et al., 2007) and presents a unique situation compared to more “standard” amphibian phylogeographic patterns (Rissler and Smith, 2010). Pending additional field sampling, the recognition of a distinct, geographically-restricted species suggests that conservation needs may be high, particularly in light of the tremendous human population density in this region.
Fig. 3. Bayesian phylogenies for individual nuDNA loci: (a) CXC4R, (b) NTF3, (c) Rag-1, (d) SIA, (e) Tyr. Nodal support: Bayesian posterior probabilities/maximum-likelihood bootstrap values. Outgroup root (R. catesbeiana) was removed for diagram simplicity. Tip labels correspond to Supplementary Table S2. Clade symbols correspond to Fig. 1. Colors correspond to inferred species: R. sphenocephala (green), R. pipiens (blue), R. palustris (orange), Rana sp. nov. (red). Colors are available in the online version. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Fig. 4. Bayesian phylogeny for concatenated nuDNA (CXCR4, NTF3, Rag-1, SIA, Tyr). Nodal support: Bayesian posterior probabilities. Tip labels correspond to Supplementary Table S2.
and epidemic declines and extirpations from mainland and coastal regions of the Tri-State area, including Long Island (Feinberg et al., unpublished data), an area once considered a regional stronghold for leopard frogs (Schlauch, 1978).

*Rana sphenocephala* is currently (as of 2011) listed as a Species of Special Concern in NY; it is not listed in NJ. *Rana pipiens* is not a listed species in NY and is not known to be present in NJ. Our genetic data demonstrate that all of the leopard frogs collected in southern mainland NY for this study were *R. sphenocephala*, rather than *R. sphenocephala*. Staten Island and the two populations in southern mainland NY (Orange, Putnam) are the only known extant putative *R. sphenocephala* populations in NY, suggesting that southern leopard frogs do not occur in NY, although information gaps remain regarding Long Island. Furthermore, *R. sphenocephala* is currently believed to be present throughout the entire state of NJ, but all of the samples collected in northern NJ were *R. sp. nov*. Our findings therefore have important implications for conservation and geographic range delimitation for not only *R. sp. nov.*, but also *R. sphenocephala*, which until now has likely been erroneously considered to be part of the fauna of NY and northern NJ.

We strongly suspect that *R. sp. nov.* also occurred on Long Island based on historic descriptive literature and photographs (Overton, 1914a, 1914b; Villani, 1997). Leopard frogs were once abundant on Long Island (Latham, 1971) but are now presumed extirpated (Kiviat, 2010; Feinberg et al., unpublished data). The samples that we analyzed from our field collections on Long Island came from recently metamorphosed tadpoles that our genetic data indicated were *R. palustris*. *Rana palustris* is still common in many central and eastern Long Island localities, and tadpoles and recent metamorphs of this species can be morphologically very similar to leopard frogs. The most recent verified photograph of a live leopard frog on Long Island was taken between 1994 and 1995 (Villani, 1997; Villani, pers. comm.). The historical and current status of leopard frogs on Long Island reflects a distressing trend throughout this region of rapid decline of leopard frog populations (Lannoo, 2005).

The geographic range of *R. sp. nov.* is very small and likely contains only a relatively small number of individual frogs. Amphibians are sensitive to small changes in their environment, and geographically restricted species with few individuals have a reduced chance for survival in the face of rapid climate change, pesticides, and disease (Lande, 1988). *Rana* sp. nov. potentially faces all of these threats, as the pesticide atrazine (Hayes et al., 2002, 2003, 2010), the fungus *Batrachochytrium dendrobatis* (Morell, 1999; Bradley et al., 2002; Stuart et al., 2004; Greer et al., 2005; Searle et al., 2011), and *Ranavirus* outbreaks (Granoff et al., 1965) have been shown to have adverse effects on leopard frog populations in this and other regions (but see Voordouw et al., 2010).

Future studies should focus on the ecology and population genetics of *R. sp. nov.*, including breeding phenology and call structure, and incorporate more fine-scaled sampling to gain a better understanding of the distribution of, and gene flow among, existing populations. Ongoing additional work (Feinberg et al., unpublished) will address these issues and describe *R. sp. nov.* as a novel species, furthering our understanding of the *R. pipiens* species complex in this region. In light of this new systematic knowledge, the “precautionary principle” (Raffensperger et al., 1999; Georges et al., 2011) suggests that appropriate conservation measures should be considered for immediate implementation at the state and possibly federal levels. The northeastern US is generally viewed as a glacially-impacted region of low diversity compared to the southeastern US (Rissler and Smith, 2010) or California (Rissler et al., 2006), and thus this region has received relatively less scrutiny and study in recent decades compared to regions that are believed to harbor higher overall diversity (but see *Pseuderis kalmi* Lemmon et al., 2007). However, urban environments such as the northeastern US have been shown to be detrimental to anuran populations, primarily due to habitat fragmentation and isolation, road mortality, and contamination (Findlay and Houlihan, 1997; Hitchings and Beebee, 1997; Knutson et al., 1999; Gibbs et al., 2005). It is therefore likely that species endemic to the Northeast require swift management attention to preserve what biodiversity still remains in the region. Our study revealed a new leopard frog species in the midst of this highly developed region of the US, suggesting that the densely populated Northeast still harbors cryptic biodiversity that remains to be discovered.

**Acknowledgments**

We are indebted to R.C. Thomson, P.Q. Spinks, and G.B. Pauly for bioinformatics help and comments. We would also like to thank M.D. Schlesinger, B. Zarate, and J. Victoria for collecting tissue samples and providing comments, G.J. Watkins-Colwell at the Yale Peabody Museum for providing tissue samples, and J.M. Burnley, E.C. Schlauch, and E. Klaastad for providing comments and valuable natural history information. We thank J.J. Apodaca and G.B. Pauly
for designing primers. We also thank two anonymous reviewers for suggestions that improved this manuscript. Genetics laboratory work was conducted at U.C. Davis, and all research was conducted under appropriate state collecting permits for New York (Collect or Possess, #969, to M.D. Schlesinger) and New Jersey (Scientific Collecting Permit to B. Zarate). Funding for genetic analyses was provided by the National Science Foundation DEB-0817042 to H.B. Shaffer, and the U.C. Davis Agricultural Experiment Station. Additional funding was provided by the New York State Biodiversity Research Institute, Long Island Community Foundation, Rutgers University, Brookhaven National Laboratory, and the Foundation for Ecological Research in the Northeast to J.A. Feinberg.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2012.01.021.

References


