A mitochondrial DNA phylogeny of African clawed frogs: phylogeography and implications for polyploid evolution

Ben J. Evans, Darcy B. Kelley, Richard C. Tinsley, Don J. Melnick, and David C. Cannatella

Abstract

The African clawed frogs (Silurana and Xenopus), model organisms for scientific inquiry, are unusual in that allopolyploidization has occurred on multiple occasions, giving rise to tetraploid, octoploid, and dodecaploid species. To better understand their evolution, here we estimate a mitochondrial DNA phylogeny from all described and some undescribed species. We examine the timing and location of diversification, and test hypotheses concerning the frequency of polyploid speciation and taxonomy. Using a relaxed molecular clock, we estimate that extant clawed frog lineages originated well after the breakup of Gondwana, about 63.7 million years ago, with a 95% confidence interval from 50.4 to 81.3 million years ago. Silurana and two major lineages of Xenopus have overlapping distributions in sub-Saharan Africa, and dispersal–vicariance analysis suggests that clawed frogs originated in central and/or eastern equatorial Africa. Most or all extant species originated before the Pleistocene; recent rainforest refugia probably acted as "lifeboats" that preserved existing species, rather than "species pumps" where many new successful lineages originated. We estimate that polyploidization occurred at least six times in clawed frogs.

Keywords: Xenopus; Silurana; Pipidae; Genome duplication; Hybridization; Polyploidy; Rainforest refugia

1. Introduction

The clawed frogs Xenopus laevis and Silurana tropicalis are the most intensively studied amphibians; they are readily maintained in captivity, have large oocytes that facilitate investigation and manipulation of embryogenesis, transgenesis is now possible, and the DNA sequence of the complete genome of Silurana tropicalis soon will be completed. However, a comprehensive phylogeny is still unavailable because of the challenge of sampling these animals throughout their distribution in sub-Saharan Africa, and also because of an unusual mode of evolution by allopolyploidization, which complicates phylogeny estimation. A phylogeny is central to understanding the timing and placement of diversification of this interesting group, and can be used to tease apart details of allopolyploid speciation.

1.1. Natural history and systematics

Clawed frogs comprise Xenopus and Silurana, which are sister taxa in the subfamily Xenopodinae, family Pipidae (Cannatella and de Sá, 1993; de Sá and Hillis, 1990). The other pipid subfamily, Pipinae, includes the New World Pipa (7 species), and the African sister genera Hymenochirus (4 species), and Pseudhymenochirus (1 species; Cannatella and Trueb, 1988a,b). Pipoid frogs include the Pipidae and the New World sister
family Rhinophrynidae (Ford and Cannatella, 1993) among living forms.

Pipid frogs are aquatic anurans that live principally in still or slow moving water. A derived morphology facilitates aquatic life, and is characterized by a lateral line system that persists through metamorphosis, laterally positioned limbs and dorsoventral compression of the body, pelvic modifications that facilitate swimming but impede terrestrial jumping, and modifications of the head and cranium including the loss of a tongue (Trueb, 1996). The larynx is exceptionally sexually dimorphic (larger in males; Sassoon and Kelley, 1986) and produces underwater vocalizations using a unique mechanism of sound production (Yager, 1996). At least one species has an elaborate repertoire of communication characterized by multiple call types in males and in females, and dueting between the sexes and multiple male call types (Tobias et al., 1998, 2004). Clawed frogs probably aestivate underground during dry periods and may migrate over land and through flowing water (McCoid and Fritts, 1980; Tinsley and McCoid, 1996).

*Silurana* includes one diploid species (2N = 20) and one tetraploid species (2N = 40); additionally one diploid (referred to here as *S. cf. tropicalis*) and two tetraploids are suggested by divergent mtDNA, and are undescribed (see below). *Xenopus* includes 10 tetraploid species (2N = 36), 5 octoploid species (2N = 72), and 2 dodecaploid species (2N = 108) (Schmid and Steinlein, 1991; Tymowska, 1991). One of the 10 tetraploids, "*Xenopus muelleri* west" (Kobel et al., 1996, 1998), is referred to as *X.* "new tetraploid" in this paper. We identified two tetraploids similar to *Xenopus fraseri*, and refer to them as *X. cf. fraseri* 1 and 2. *Xenopus* has been divided into five species groups based on similarities in morphology, advertisement call, and/or molecular characters: (1) the *laevis* group, which includes the tetraploids *X. laevis*, *X. gilli*, and *X. largeni*; (2) the *muelleri* group, which includes the tetraploids *X. muelleri*, *X. borealis*, and *X. clivii*; (3) the *vestitus–wittei* group, which includes the octoploids *X. wittei* and *X. vestitus*, (4) the *fraseri*-like group, which includes the tetraploids *X. fraseri*, *X. pygmaeus*, the octoploids *X. amieti*, *X. andreii*, *X. bombaensis*, and the dodecaploid *X. ruwenzoriensis*, and (5) the *longipes* group which includes only the dodecaploid *X. longipes* (Kobel et al., 1996). Additional information is necessary to test whether some of the divergent mtDNA lineages do constitute separate species; we view species assignments in this study that are based only on mtDNA as working hypotheses (Evans et al., 2003).

A mitochondrial DNA phylogeny of 7 species was estimated and reanalyzed from mapped restriction endonuclease sites of 11 enzymes in the complete mtDNA genome (Cannatella and de Sá, 1993; Carr et al., 1987). Another mtDNA phylogeny for 11 species was estimated from about 600 bp of mitochondrial ribosomal DNA (rDNA) sequence (Kobel et al., 1998). Relationships among some species have also been estimated from protein data (Bisbee et al., 1977; Graf, 1996; Mann et al., 1982), and among extant and fossil taxa from morphology (Trueb and Baez, 1997).

### 1.2. Allopolyploid speciation and mitochondrial DNA

Speciation of clawed frogs can occur without genome duplication, as is common in other organisms; an ancestor gives rise to two descendants with the same number of chromosomes. Alternatively, speciation of clawed frogs can occur by reticulation via allopolyploidy; here two species hybridize and give rise to a descendant species that inherits the complete nuclear genome of both ancestors. Allopolyploidization rather than autopolyploidization is probably the principal mode of genome duplication in clawed frogs (Kobel, 1996a).

Some polyploid species, for example, are morphologically similar to one species but carry mtDNA more closely related to another (Carr et al., 1987; Kobel et al., 1998). Hybrids are also known to occur in nature. The widely distributed *X. laevis* hybridizes with three other sympatric species: *X. gilli* in Cape Province, South Africa (Evans et al., 1997, 1998; Picker et al., 1996), *X. muelleri* in northern South Africa (Fischer et al., 2000), and *X. borealis* in Kenya (Yager, 1996). Hybrid crosses have yielded polyploid progeny in the laboratory (Kobel, 1996a; Kobel and Du Pasquier, 1986) and functional diploidization of some polyploid species is suggested by the formation of bivalents during meiosis (Kobel et al., 1996).

In *Xenopus* a diploid ancestor with 18 chromosomes is suggested by extant tetraploid karyotypes (2N = 36), but presumed extinct. Thus, polyploidization in *Xenopus* is inferred to have occurred at least once from 18 to 36 chromosomes, at least once from 36 to 72 chromosomes, and at least once from a combination of genomes with 36 and 72 chromosomes to 108 chromosomes. In *Silurana*, polyploidization is inferred to have occurred at least once from 20 to 40 chromosomes.

As an initial step to unraveling the evolutionary history of this allopolyploid group, we have constructed a phylogeny based on mitochondrial DNA. In allopolyploid speciation, a polyploid descendant inherits the complete nuclear genome of two ancestors, but mitochondrial DNA is inherited from only one ancestor. Because mitochondrial DNA is maternally inherited and haploid, a mitochondrial DNA phylogeny only reflects a portion of the ancestry of a species that originated from allopolyploidization (species C, Figs. 1a and b). A mitochondrial DNA phylogeny also offers only a minimum estimate of the number of times polyploidization has occurred, because a single mtDNA lineage might occur in polyploids that have multiple independent origins (Figs. 1c and d). A mitochondrial DNA phylogeny also
offers insight into phylogeographic patterns of differentiation, irrespective of ploidy level, over the large range of this group.

1.3. Goals

To better understand the evolution of African clawed frogs here we: (1) estimate phylogenetic relationships among mitochondrial DNA (mtDNA) of all known and some undescribed species, (2) elucidate phylogeographic patterns of this group over sub-Saharan Africa, (3) use a molecular clock calibrated by geology and fossil pipids to estimate the age of mtDNA clades of clawed frogs, and (4) evaluate the minimum number of polyploidization events needed to account for the mtDNA phylogeny.

2. Methods

2.1. Genetic samples

We sequenced a broad taxonomic sample of 52 in-group individuals and 1 outgroup individual, including all known species of clawed frog and intraspecific samples from multiple localities (Fig. 2 and Table 1). Although some of the recognized subspecies of *X. laevis* may merit full species status (Kobel et al., 1996; Measey and Channing, 2003), we treated them all as subspecies of *X. laevis*.

Most samples in this study were obtained from the live collection at the Institute of Zoology at the University of Geneva (Table 1). These individuals were collected and identified by B. Colombelli, L. Du Pasquier, M. Fischberg, H. Kobel, C. Loumont, J. Perret, D. Rungger, E. Rungger-Brandle, C. Thiebaud, J. Tymowska, and R.C.T. Species identifications were verified using characters compiled in Tinsley and Kobel (1996). Additional samples were collected by B.J.E., R.C.T., D.B.K., D.C.C., M. Picker, and M. Tobias, and obtained from the tissue archives of CAS, AMNH, TNHC, and ROM (museum abbreviations follow Leviton et al., 1985). Sequences from other pipoids (*Pipa pipa*, *Pipa parva*, *Hymenochirus* sp., and *Rhinophrynus dorsalis*) were included, and a sequence from a pelobatid frog, *Scaphiopus hurterii* was used as an outgroup (Table 1). We also analyzed a portion of the complete mtDNA sequence of *S. tropicalis* published online by the Joint Genome Institute (genome.jgi-psf.org/xenopus0/X.tropicalisMtGenome.gb). Because this sequence was identical to our sequence of *S. tropicalis* from Nigeria, we did not include it in subsequent analyses.

2.2. Ploidy

For many species, ploidy level was ascertained previously by karyotyping or measuring DNA content of related individuals (e.g., Thiebaud and Fischberg, 1977; Tymowska, 1991). To assess ploidy level of all samples in this study, alleles of the nuclear gene RAG-1 were co-amplified with PCR, cloned, and independently sequenced (B.J.E., unpublished data). Because it was not possible to obtain a karyotype or to measure DNA content from two ethanol preserved tissues (*S. cf. tropicalis* and *S. new tetraploid 2), we inferred the ploidy level of these individuals solely by cloning and counting the number of divergent RAG-1 alleles. We view these
ploidy levels as provisional until karyotypes of other conspecific individuals can be performed.

2.3. Molecular data

Data for this study include 2335 bp of DNA (including gaps that were introduced through alignment but not including regions of ambiguous homology that were excluded from analysis) from portions of the 12S gene, tRNA\textsuperscript{val}, and 16S rDNA. The gene order of the sequenced region (5’–3’) is 12S rDNA, tRNA\textsuperscript{val}, and 16S rDNA. These sequences include the 12S rDNA gene beginning approximately 106 bp after the 5’ start of the gene, the entire tRNA\textsuperscript{val} gene, and most of the 16S rDNA gene ending about 246 bp before the 3’ end. This region is homologous to positions 2816–5056 in the \textit{Rana nigromaculata} GenBank Accession No. AB043889. To sequence this region, we used ABI Prism 3100 and 3730XL automated capillary sequencers and primers described elsewhere (Evans et al., 2003). Alignment was done with ClustalX (Thompson et al., 1997) and then adjusted by eye using MacClade, version 4.06 (Maddison and Maddison, 2000). Autapomorphies such as single site insertions in a species were verified by examining the chromatograms. Alignment was generally unambiguous with the exception of two regions that were excluded: one region of \(~13\) bp begins about 80 bp upstream of the 3’ end of the 12S gene; the other region of \(~10\) bp begins at base 18 in the 16S gene. These data have been submitted to GenBank (Table 2).

2.4. Phylogenetic analyses

Phylogenetic analyses used maximum parsimony (MP), maximum likelihood (ML), and Bayesian methods. For MP analysis we performed a heuristic search for the most parsimonious tree with 1000 replicates of random taxon-addition and TBR branch swapping, with gaps treated as missing characters, using PAUP* version 4.0 (Swofford, 2002). Non-parametric bootstrap values were obtained with 2000 replicates, each with a single replicate of random taxon-addition. Decay index values were calculated for each node with Autodecay, version 4.0.2 (Eriksson, 1999).

For ML and Bayesian analysis, we used Modeltest version 3.06 (Posada and Crandall, 1998) to evaluate
<table>
<thead>
<tr>
<th>Species</th>
<th>Museum catalog</th>
<th>Country of origin</th>
<th>Specific locality</th>
<th>Coordinates (lat long)</th>
<th>GenBank</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outgroup</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaphiopus hurterii</td>
<td>TNHC 61954</td>
<td>United States</td>
<td>Aransas County, Texas</td>
<td>28.05 –97.05</td>
<td>AY581619</td>
</tr>
<tr>
<td><strong>Silurana</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>S. tropicalis</em></td>
<td>—</td>
<td>Nigeria</td>
<td>Uyere, near Benin</td>
<td>6.72 5.77</td>
<td>AY581668</td>
</tr>
<tr>
<td><em>S. tropicalis</em></td>
<td>—</td>
<td>Nigeria</td>
<td>Unknown</td>
<td>6.72 5.77</td>
<td>AY581669</td>
</tr>
<tr>
<td><em>S. tropicalis</em></td>
<td>MHNG 2644.55</td>
<td>Sierra Leone</td>
<td>Near Freetown</td>
<td>8.50 –13.24</td>
<td>AY581670</td>
</tr>
<tr>
<td><em>S. tropicalis</em></td>
<td>—</td>
<td>Ivory Coast</td>
<td>Adiopou Doume</td>
<td>5.32 –4.12</td>
<td>AY581671</td>
</tr>
<tr>
<td><em>S. cf. tropicalis</em></td>
<td>ROM 19161</td>
<td>Liberia</td>
<td>Sinoe Province, Sapo National Park, Ghabone Creek</td>
<td>5.01 –9.04</td>
<td>AY581666</td>
</tr>
<tr>
<td><em>S. epiptropicalis</em></td>
<td>MHNG 2644.56</td>
<td>DRC</td>
<td>Kinshasa</td>
<td>–4.30 15.30</td>
<td>AY581665</td>
</tr>
<tr>
<td><em>S. new tetraploid 1</em></td>
<td>—</td>
<td>Republic of the Congo (Brazzaville)</td>
<td>Malemba</td>
<td>–5.83 12.57</td>
<td>AY581660</td>
</tr>
<tr>
<td><em>S. new tetraploid 1</em></td>
<td>MHNG 2644.58</td>
<td>Cameroon</td>
<td>Longyi</td>
<td>2.95 9.92</td>
<td>AY581661</td>
</tr>
<tr>
<td><em>S. new tetraploid 1</em></td>
<td>—</td>
<td>Gabon</td>
<td>Nkoemvone</td>
<td>2.88 11.15</td>
<td>AY581664</td>
</tr>
<tr>
<td><em>S. new tetraploid 1</em></td>
<td>—</td>
<td>Gabon</td>
<td>Makokou</td>
<td>0.57 12.87</td>
<td>AY581662</td>
</tr>
<tr>
<td><em>S. new tetraploid 1</em></td>
<td>CAS 207759</td>
<td>Equatorial Guinea</td>
<td>Bioko Island, Arena Blanca Rd.</td>
<td>3.53 8.58</td>
<td>AY581667</td>
</tr>
<tr>
<td><em>S. new tetraploid 1</em></td>
<td>—</td>
<td>Cameroon</td>
<td>Galim (type locality)</td>
<td>5.63 10.33</td>
<td>AY581634</td>
</tr>
<tr>
<td><em>S. new tetraploid 1</em></td>
<td>CAS 207765</td>
<td>Equatorial Guinea</td>
<td>Bioko Id. Arena Blanca Rd.</td>
<td>3.53 8.58</td>
<td>AY581632</td>
</tr>
<tr>
<td><em>S. new tetraploid 1</em></td>
<td>—</td>
<td>Cameroon</td>
<td>Yaounde</td>
<td>3.87 11.52</td>
<td>AY581631</td>
</tr>
<tr>
<td><em>S. new tetraploid 2</em></td>
<td>AMNH 158377</td>
<td>Chad</td>
<td>Préfecture du Borkou-Ennedi-Tibesti, Sub-Préfecture Ennedi, Guelia d’Archei</td>
<td>16.91 21.77</td>
<td>AY581651</td>
</tr>
<tr>
<td><em>X. amieti</em></td>
<td>MHNG 2644.54</td>
<td>Cameroon</td>
<td>Galim (type locality)</td>
<td>5.63 10.33</td>
<td>AY581634</td>
</tr>
<tr>
<td><em>X. cf. fraseri 1</em></td>
<td>CAS 207765</td>
<td>Equatorial Guinea</td>
<td>Bioko Id. Arena Blanca Rd.</td>
<td>3.53 8.58</td>
<td>AY581632</td>
</tr>
<tr>
<td><em>X. cf. fraseri 1</em></td>
<td>—</td>
<td>Cameroon</td>
<td>Yaounde</td>
<td>3.87 11.52</td>
<td>AY581631</td>
</tr>
<tr>
<td><em>X. new tetraploid 1</em></td>
<td>MHNG 2644.60</td>
<td>Nigeria</td>
<td>near Jos</td>
<td>8.7–9.0 11.2–12.0</td>
<td>AY581652</td>
</tr>
<tr>
<td>X. andrei</td>
<td>MHNG 2644.51</td>
<td>Cameroon</td>
<td>Longyi (type locality)</td>
<td>2.95 9.91</td>
<td>AY581627</td>
</tr>
<tr>
<td>X. borealis</td>
<td>—</td>
<td>Kenya</td>
<td>Marsabit</td>
<td>2.33 38.01</td>
<td>AY581653</td>
</tr>
<tr>
<td>X. borealis</td>
<td>—</td>
<td>Kenya</td>
<td>Kimbu</td>
<td>–1.17 38.01</td>
<td>AY581654</td>
</tr>
<tr>
<td>X. borealis</td>
<td>MHNG 2644.64</td>
<td>Kenya</td>
<td>Samburu Range</td>
<td>0.66 37.50</td>
<td>AY581656</td>
</tr>
<tr>
<td>X. borealis</td>
<td>—</td>
<td>Kenya</td>
<td>Maralal</td>
<td>1.08 36.68</td>
<td>AY581655</td>
</tr>
<tr>
<td>X. boumbaensis</td>
<td>MHNG, uncataloged</td>
<td>Cameroon</td>
<td>Yaounde</td>
<td>3.87 11.52</td>
<td>AY581635</td>
</tr>
<tr>
<td>X. boumbaensis</td>
<td>MHNG 2644.57</td>
<td>Cameroon</td>
<td>Moloundou (type locality)</td>
<td>2.02 15.13</td>
<td>AY581633</td>
</tr>
<tr>
<td>X. clivic</td>
<td>MHNG 2644.50</td>
<td>Ethiopia</td>
<td>Near Addis Ababa</td>
<td>9.02 38.70</td>
<td>AY581637</td>
</tr>
<tr>
<td>X. gilli</td>
<td>—</td>
<td>South Africa</td>
<td>Cape Province, Cape Point Nature Reserve</td>
<td>–34.25 18.41</td>
<td>AY581649</td>
</tr>
<tr>
<td>X. gilli</td>
<td>—</td>
<td>South Africa</td>
<td>Cape Province, Betty’s Bay</td>
<td>–34.37 18.92</td>
<td>AY581650</td>
</tr>
<tr>
<td>X. laevis</td>
<td>MHNG 2644.61</td>
<td>Malawi</td>
<td>Near Blantyre</td>
<td>–15.78 35.00</td>
<td>AY581648</td>
</tr>
<tr>
<td>X. laevis</td>
<td>MHNG 2644.67</td>
<td>Republic of the Congo (Brazzaville)</td>
<td>Between Loubono and Pointe de Noire near town of Kouilla</td>
<td>–4.30 12.40</td>
<td>AY581638</td>
</tr>
<tr>
<td>X. laevis</td>
<td>—</td>
<td>South Africa</td>
<td>Cape Province, Cape Point Penninsula, Lewis Gay Dam</td>
<td>–34.20 18.40</td>
<td>AY581639</td>
</tr>
<tr>
<td>X. laevis</td>
<td>—</td>
<td>Botswana</td>
<td>Okavango</td>
<td>–18.97 22.57</td>
<td>AY581642</td>
</tr>
<tr>
<td>X. laevis</td>
<td>MHNG 2644.53</td>
<td>Zambia</td>
<td>Near Lusaka</td>
<td>–15.50 28.20</td>
<td>AY581644</td>
</tr>
<tr>
<td>X. laevis</td>
<td>—</td>
<td>Cameroon</td>
<td>Ngaoundere</td>
<td>7.32 13.58</td>
<td>AY581640</td>
</tr>
<tr>
<td>X. laevis</td>
<td>MHNG 2644.52</td>
<td>Nigeria</td>
<td>Jos</td>
<td>9.90 8.90</td>
<td>AY581641</td>
</tr>
<tr>
<td>X. laevis</td>
<td>CAS 168711</td>
<td>Tanzania</td>
<td>Tanga Region, Muheza Dist., East Usambara Mts, Amani, Amani Pond</td>
<td>–5.07 38.72</td>
<td>AY581647</td>
</tr>
<tr>
<td>X. laevis</td>
<td>—</td>
<td>Rwanda</td>
<td>Shama</td>
<td>–1.95 30.06</td>
<td>AY581645</td>
</tr>
<tr>
<td>X. laevis</td>
<td>—</td>
<td>Uganda</td>
<td>Kitanga</td>
<td>–1.12 30.05</td>
<td>AY581646</td>
</tr>
<tr>
<td>X. largeni</td>
<td>MHNG 2644.59</td>
<td>Ethiopia</td>
<td>Sidamo Province, 400 km S of Addis Ababa, near Kibre Mengist</td>
<td>5.60 38.90</td>
<td>AY581643</td>
</tr>
<tr>
<td>X. longipes</td>
<td>AMNH A168447</td>
<td>Cameroon</td>
<td>Lake Oku (type locality)</td>
<td>6.21 10.47</td>
<td>AY581625</td>
</tr>
<tr>
<td>X. muelleri</td>
<td>—</td>
<td>Swaziland</td>
<td>Nkambeni area</td>
<td>–26.05 31.66</td>
<td>AY581657</td>
</tr>
<tr>
<td>X. muelleri</td>
<td>—</td>
<td>Malawi</td>
<td>Near Blantyre</td>
<td>–15.78 35.00</td>
<td>AY581658</td>
</tr>
<tr>
<td>X. muelleri</td>
<td>MHNG 2644.63</td>
<td>Tanzania</td>
<td>Ifakara</td>
<td>–8.13 36.68</td>
<td>AY581659</td>
</tr>
</tbody>
</table>
Maximum likelihood analysis was performed with PAUP* (Swofford, 2002) under the model selected by Modeltest (GTR + I + C, see below), with two rounds of successive approximation, each with five replications of random taxon-addition. Initial parameters were set to those estimated from a neighbor-joining topology, and then re-estimated from the resulting topology and set to these new values for the second iteration. Only two rounds of successive approximation were performed because the topology recovered from both rounds was identical.

Bayesian analysis was performed with MrBayes, version 3.0b4 (Huelsenbeck and Ronquist, 2001), under the same model as the ML analysis. We ran four Metropolis coupled Markov chain Monte Carlo analyses starting with random trees for each of four simultaneous chains, with flat Dirichlet prior distributions set to 1.0 for each rate substitution type, and the differential heating parameter set to 0.2. Four runs were performed, each for 5,000,000 generations. Post-run analysis suggested that parameter estimates reached stationarity well before 200,000 generations; trees recovered from these initial rounds were discarded as burn-in. The joint posterior probability distributions and parameter estimates of each of the independent runs were similar, suggesting that the chains were run for a sufficient number of generations to sample the posterior probability landscape adequately. Posterior probabilities were calculated from post-burn-in trees from all runs.

There is no voucher for some samples and GPS coordinates (in decimal degree format) of some localities are approximate. Taxa demarcated by an asterisk were analyzed by Kobel et al. (1998).

### Table 1 (continued)

<table>
<thead>
<tr>
<th>Species</th>
<th>Museum catalog</th>
<th>Country of origin</th>
<th>Specific locality</th>
<th>Coordinates (lat long)</th>
<th>GenBank</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>X. pygmaeus</em></td>
<td>—</td>
<td>DRC</td>
<td>Boende</td>
<td>−0.25 20.60</td>
<td>AY581626</td>
</tr>
<tr>
<td><em>X. ruwenzoriensis</em></td>
<td>—</td>
<td>Uganda</td>
<td>Rain forest near Bundibugyo in the Semliki Valley at the foot of the Ruwenzi Mountains (type locality)</td>
<td>0.70 30.00</td>
<td>AY581624</td>
</tr>
<tr>
<td><em>X. vestitus</em></td>
<td>—</td>
<td>Uganda</td>
<td>Lake Mutanda, near Kisoro</td>
<td>−1.22 29.71</td>
<td>AY581636</td>
</tr>
<tr>
<td><em>X. wittei</em></td>
<td>CAS 201664</td>
<td>Uganda</td>
<td>Rukungriri Dist., Bwindi, Impenetrable National Park, Kabale-Kayonza Rd.</td>
<td>−0.98 29.69</td>
<td>AY581629</td>
</tr>
<tr>
<td><em>X. wittei</em></td>
<td>—</td>
<td>Rwanda</td>
<td>Cyamudongo Forest, Nyakabuye</td>
<td>−2.57 28.98</td>
<td>AY581628</td>
</tr>
<tr>
<td><em>X. wittei</em></td>
<td>MHNG 2644.62</td>
<td>Uganda</td>
<td>Chelima Forest, near Echuya and Mulehe (type locality)</td>
<td>−1.06 29.90</td>
<td>AY581630</td>
</tr>
<tr>
<td>Other pipoids</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhinophrynus dorsalis</td>
<td>LACM 129913</td>
<td>Mexico</td>
<td>24 mi. WNW Coyuca de Benitez, 0.6 mi. NW El Papayo</td>
<td>17.03 −100.07</td>
<td>AY581620</td>
</tr>
<tr>
<td>Hymenochirus sp.</td>
<td>—</td>
<td>Africa, country unknown</td>
<td>Unknown</td>
<td>— —</td>
<td>AY581623</td>
</tr>
<tr>
<td>Pipa parva</td>
<td>—</td>
<td>Venezuela</td>
<td>Merida</td>
<td>8.60 −71.15</td>
<td>AY581622</td>
</tr>
<tr>
<td>Pipa pipa</td>
<td>KU 205801</td>
<td>Peru</td>
<td>Madre de Dios, Cusco Amazonico, 15 km E of Puerto Maldonado</td>
<td>−12.60 −70.08</td>
<td>AY581621</td>
</tr>
</tbody>
</table>

There is no voucher for some samples and GPS coordinates (in decimal degree format) of some localities are approximate. Taxa demarcated by an asterisk were analyzed by Kobel et al. (1998).

### Table 2

Estimated divergence times of nodes in millions of years and 95% confidence interval (min–max)

<table>
<thead>
<tr>
<th>Node</th>
<th>Age</th>
<th>95% confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pipidae</td>
<td>129.8</td>
<td>114.5–146.1</td>
</tr>
<tr>
<td>Pipinae</td>
<td>112.0</td>
<td>(Fixed)</td>
</tr>
<tr>
<td>Xenopodinae</td>
<td>63.7</td>
<td>50.4–81.3</td>
</tr>
<tr>
<td>Pipa</td>
<td>55.7</td>
<td>46.5–65.7</td>
</tr>
<tr>
<td>Xenopus</td>
<td>41.6</td>
<td>31.8–54.6</td>
</tr>
<tr>
<td>Silurana</td>
<td>17.6</td>
<td>12.1–24.8</td>
</tr>
<tr>
<td>Clade A</td>
<td>32.3</td>
<td>22.7–45.3</td>
</tr>
<tr>
<td>Clade B</td>
<td>19.4</td>
<td>13.7–25.2</td>
</tr>
<tr>
<td>Clade C</td>
<td>12.0</td>
<td>7.8–17.7</td>
</tr>
<tr>
<td>MRCMA <em>X. longipes</em>, <em>X. ruwenzoriensis</em></td>
<td>6.0</td>
<td>2.8–8.3</td>
</tr>
<tr>
<td>MRCMA <em>X. longipes</em>, <em>X. amieti</em></td>
<td>3.7</td>
<td>0.9–15.0</td>
</tr>
<tr>
<td>MRCMA <em>X. pygmaeus</em>, <em>X. cf. fraseri I</em></td>
<td>6.8</td>
<td>4.2–10.5</td>
</tr>
<tr>
<td>MRCMA <em>X. pygmaeus</em>, <em>X. wittei</em></td>
<td>10.0</td>
<td>2.3–13.8</td>
</tr>
<tr>
<td>MRCMA <em>X. pygmaeus</em>, <em>X. andreii</em></td>
<td>9.1</td>
<td>6.0–13.4</td>
</tr>
<tr>
<td>MRCMA <em>X. largeni</em>, <em>X. pygmaeus</em></td>
<td>20.1</td>
<td>13.3–30.1</td>
</tr>
<tr>
<td>MRCMA <em>X. vestitius</em>, <em>X. pygmaeus</em></td>
<td>25.5</td>
<td>17.7–36.1</td>
</tr>
<tr>
<td>MRCMA <em>X. laevis</em>, <em>X. gilli</em></td>
<td>16.7</td>
<td>9.6–26.2</td>
</tr>
<tr>
<td>MRCMA <em>S. epitropicalis</em>, <em>S. new tetraploid 1</em></td>
<td>7.7</td>
<td>4.3–12.0</td>
</tr>
</tbody>
</table>

Age of the root (Pipoidea) is not shown because of inaccuracy of method in determining age of root. Some nodes are referred to as the most recent common maternal ancestor (MRCMA) of two species. The average rate of evolution was 0.00249 substitutions per site per million years, $s = 0.00029$, with a range of 0.00184–0.00299.

Different models of evolution. Maximum likelihood analysis was performed with PAUP* (Swofford, 2002) under the model selected by Modeltest (GTR + I + C, see below), with two rounds of successive approximation, each with five replications of random taxon-addition. Initial parameters were set to those estimated from a neighbor-joining topology, and then re-estimated from the resulting topology and set to these new values for the second iteration. Only two rounds of successive approximation were performed because the topology recovered from both rounds was identical.

Bayesian analysis was performed with MrBayes, version 3.0b4 (Huelsenbeck and Ronquist, 2001), under the same model as the ML analysis. We ran four Metropolis coupled Markov chain Monte Carlo analyses starting with random trees for each of four simultaneous chains, with flat Dirichlet prior distributions set to 1.0 for each rate substitution type, and the differential heating parameter set to 0.2. Four runs were performed, each for 5,000,000 generations. Post-run analysis suggested that parameter estimates reached stationarity well before 200,000 generations; trees recovered from these initial rounds were discarded as burn-in. The joint posterior probability distributions and parameter estimates of each of the independent runs were similar, suggesting that the chains were run for a sufficient number of generations to sample the posterior probability landscape adequately. Posterior probabilities were calculated from post-burn-in trees from all runs.
2.5. Molecular clock

In order to estimate divergence times from molecular data, we used a likelihood-ratio test (Goldman, 1993) to determine whether the likelihood of a tree without a molecular clock enforced was significantly greater than one with a molecular clock enforced under the GTR +I+Γ model of evolution, in which degrees of freedom are equal to the number of terminals–2 (Huelsenbeck and Rannala, 1997). Because the data did not conform to expectations of clock-like evolution (see below), we used a semi-parametric method to estimate divergence times of various clades using r8s, version 1.60 (Sanderson, 1997, 2002). This approach, penalized likelihood, permits a different rate of evolution on every branch of the tree, but imposes a cost on the model if rates among branches change quickly.

A first step in estimating divergence times with r8s is to use a cross-validation procedure to empirically test various values for a smoothing parameter that determines the relative importance of the independent rates of each branch versus the cost of big changes in rates among branches (Sanderson, 1997, 2002). Larger smoothing parameters are appropriate for phylogenies that more closely approximate clock-like behavior. Using the TN algorithm, we tested a range of smoothing parameter values from 1 to about 3100. A value of 10 parameter for each bootstrap phylogram with r8s, and then analyzing each phylogram with the appropriate smoothing parameter and other settings as described above (Sanderson, 1997). For these analyses, we used Perl scripts to make input files and to parse output files. Confidence limits (95%) were estimated as the mean estimated age of a node in all bootstrap replicates ± 1.96 standard deviations of these estimates.

2.6. Phylogeographic analysis

Area cladograms were constructed from ML and MP topologies by changing the name of each sample to the location(s) of that species or, in X. laevis, the location of intraspecific clades (western, central, eastern, and/or southern Africa). These regions were delimited by three putative biogeographic barriers that may or may not have affected clawed frogs: the Rift Valley, the Dahomey Gap (where contemporary savannah habitat extends southward to the coast from Ghana, through Togo and Benin, to west Nigeria), and the southern limit of tropical rainforest. Western, central, and eastern Africa, have and/or had tropical rainforest in the past, whereas southern Africa was dominated by drier habitat during clawed frog evolution (Stokes et al., 1997). To better understand phylogeographic evolution of this group, we included the additional geographic information that X. new tetraploid is also known from western Africa, even though genetic material from this species was not available from this population (Tinsley et al., 1996). The distribution of Hymenochirus was set to western and central Africa and the distributions of Pipa, Rhinophrynus, and Scaphiopus were designated as New World. We also generated a chronogram from the ML topology using r8s.

We used dispersal–vicariance analysis to estimate the ancestral locations of various clades in all of the MP topologies, and in the ML topology, with DIVA, version 1.1 (Ronquist, 1996, 1997). The maximum number of ancestral areas was set to five (the total number of areas in the analysis), and ambiguous optimal solutions
were reported. This approach has the advantages that it does not assume a hierarchical (or any) area relationship, it permits an ancestor to occupy multiple areas, and it permits dispersal between areas without speciation (Ronquist, 1996, 1997). DIVA reconstructs ancestral areas by minimizing the number of dispersal and vicariance events needed to explain an extant distribution. Vicariance is the preferred (no cost) mode of speciation; dispersal and extinction each impose an equal cost on the optimization model. Most differences between the topologies recovered from each phylogenetic analysis did not affect reconstruction of ancestral areas because the topological differences occurred within a geographical region. One difference in the topology within Silurana in some of the MP trees affected reconstruction of ancestral areas within this genus, but not on other parts of the tree, and is thus not discussed.

2.7. Testing of phylogenetic hypotheses

To estimate the minimum number of polyploidization events that occurred in clawed frogs, we assume that genome duplication is unidirectional—that reversions to a lower ploidy level do not occur. Five phylogenetic hypotheses concerning the frequency of polyploidization were tested with parametric bootstrapping (Goldman et al., 2000; Hillis et al., 1996; Huelsenbeck et al., 1996). Hypothesis 1 postulates a single origin of octoploids. Because dodecaploids could potentially inherit mtDNA from a tetraploid ancestor rather than an octoploid ancestor (Fig. 1b), this hypothesis was tested with a backbone constraint in which dodecaploids were allowed to assume any phylogenetic position with respect to octoploids and tetraploids. Hypothesis 1 also leaves open the possibility that one or both dodecaploids were derived from the union of two hexaploid (2N = 54) genomes that are now extinct. Intermediate ploidy levels (triploid or hexaploid) may be preludes to allopolyploid speciation of clawed frogs (Kobel, 1996a). Hypothesis 2 postulates a single origin of dodecaploids and monophyly of dodecaploid mtDNA. Hypothesis 3 postulates a single origin of Silurana tetraploids and monophyly of these mtDNA sequences.

Phylogenetic analyses recovered a clade that contains all of the octoploids except X. vestitus and both dodecaploids (see below). However, ML and Bayesian analyses also nest the tetraploids (X. pygmaeus + X. cf. fraseri 1) in this clade, and MP analysis nests the tetraploid X. cf. fraseri 2 in this clade. To examine whether further episodes of polyploidization are supported, two hypotheses were tested that postulate mtDNA monophyly of both dodecaploids and monophyly of all octoploids except X. vestitus. The first one, Hypothesis 4, postulates monophyly of dodecaploid mtDNA within a clade containing all octoploid mtDNA except X. vestitus. The second, Hypothesis 5, postulates one clade containing dodecaploid mtDNA and a separate clade containing all octoploid mtDNA except X. vestitus. Both of these hypotheses are consistent with two origins of octoploids and one origin of dodecaploids.

Additionally, we tested for mtDNA monophyly of each species group with more than one species. Hypotheses 6–9 postulate monophyly of the laevis, muelleri, vestitus–wittei, and fraseri-like groups, respectively.

For parametric bootstrapping, a heuristic search for the most parsimonious tree consistent with each biogeographic (null) hypothesis was performed using 100 replications of random taxon-addition. The set of most parsimonious trees consistent with each null hypothesis differed primarily in relationships among short terminal branches. Because of this, simulation on any of these trees should produce similar results. We ranked these trees under the K2P + Γ model of evolution and selected the most likely tree for data simulation. Using this tree, which is the null hypothesis of the parametric bootstrap test (Goldman et al., 2000; Huelsenbeck and Rannala, 1997), a hierarchical likelihood-ratio test was used to select a model of evolution for data simulation using Modeltest (GTR + I + Γ was independently selected for each test). One thousand datasets were simulated under this more complex model for each null hypothesis with Seq-Gen version 1.2.6 (Rambaut and Grassly, 1997). For each simulated dataset, we calculated the treelength difference in parsimony score between searches unconstrained and constrained for the null hypothesis. The Logreader program (D. Zwickl, unpublished) was used to parse the output files. The probability (P) of the null hypothesis is equal to the proportion of treelength differences from the simulated data that is equal to or more extreme than the observed treelength difference. We applied the sequential Bonferroni procedure to adjust a critical value of 0.05 for hypothesis rejection in multiple tests (Rice, 1989).

The posterior probability of each hypothesis was estimated by filtering post-burn-in trees from Bayesian analysis with a constraint tree (Huelsenbeck et al., 2000). We also evaluated the prior probability of each hypothesis (the expected probability of each hypothesis before data collection) by generating 400,000 random trees with PAUP and computing the proportion of these trees that satisfy each hypothesis.

3. Results

3.1. Phylogeny

Results from MP, ML, and Bayesian analyses were similar, with exceptions discussed below. Maximum parsimony searches recovered 55 equally parsimonious trees of 3355 steps (CI = 0.477, RI = 0.766). For model-based phylogeny estimation, the GTR + I + Γ model was
Fig. 3. Maximum likelihood tree. Number of chromosomes of each species is indicated in parentheses. Non-parametric MP bootstrap values and posterior probabilities are separated by a forward slash above branches, and decay indexes are below them. An asterisk indicates bootstrap values or posterior probabilities greater than 95, a dash indicates values below 50, and for clarity some bootstrap and decay values of terminal branches are omitted. Scale bars indicate the number of substitutions per site to the left and to the right of the dotted line. Clades A–C referred to in text are indicated; these clades are also in all MP trees. Initial parameters estimated for the GTR + I + C model on a neighbor-joining tree were: base frequencies (A C G) = (0.3861 0.2381 0.1588), rate matrix (A–C A–G A–T C–G C–T G–T) = (4.1564 11.1392 10.1821 0.4933 40.2440 1), shape parameter = 0.4155, and proportion of invariant sites = 0.3135. Parameters estimated on a likelihood tree after a successive approximation search were: base frequencies (A C G) = (0.3848 0.2374 0.1605), rate matrix (A–C A–G A–T C–G C–T G–T) = (4.2356 11.1240 10.4404 0.4606 41.4022 1), shape parameter = 0.4103, and proportion of invariant sites = 0.3089.
selected by the hierarchical likelihood-ratio test. The hypothesis that these sequences conformed to a molecular clock was rejected ($P < 0.001$). Maximum likelihood analysis and Bayesian analyses recovered very similar topologies (Fig. 3).

Most clades with high bootstrap values or high posterior probabilities are present in trees recovered from all analyses. These analyses all support mtDNA monophyly of *Silurana* with respect to *Xenopus*, and monophyly of *X. borealis*, *X. muelleri*, and *X. new tetraploid* (Clade B, Fig. 3) with respect to other *Xenopus* (Clade A, Fig. 3). In all analyses, within Clade A the sequence of *X. clivii* is sister to other sequences, sequences of *X. gilli* and *X. laevis* form a clade, and mtDNA of *X. pygmaeus*, *X. cf. fraseri* 1 and 2, all octoploids except *X. vestitus*, and both dodecaploids form a clade (Clade C, Fig. 3). All analyses support monophyly of sequences of *X. ruwenzoriensis*, *X. longipes*, *X. amieti*, and *X. boumbaensis*, and monophyly of mtDNA of *X. pygmaeus* and *X. cf. fraseri* 1. Within *X. laevis*, mtDNA from South Africa is the sister to other sequences, sequences of *X. gilli* and *X. laevis* form a clade, and mtDNA from Malawi is sister to mtDNA of *X. laevis* to the North and West (Cameroon, Nigeria, Botswana, Zambia, Rwanda, Uganda, and Tanzania). *X. laevis* mtDNA from western Africa (Cameroon and Nigeria) and the center of southern Africa (Botswana and Zambia) form a clade that is sister to *X. laevis* mtDNA from East Africa (Uganda, Rwanda, and Tanzania). Of the five species groups in *Xenopus*, the only one that includes a clade of mtDNA in our analyses is the monotypic *longipes* group.

In *Silurana*, mtDNA of two tetraploids, *X. epitropicalis* and *S. tropicalis* and a putative new tetraploid, *S. new tetraploid 2* form a clade (Fig. 3).

Topologies recovered from these analyses differ in some respects within Clades A and C, and within *Silurana* (Fig. 3). Within Clade A, the mtDNA relationship of *X. largeni* with respect to (*X. laevis + X. gilli*) is not well supported in any analysis. In the MP trees, *X. largeni* is sister to a clade containing *X. laevis*, *X. gilli*, and Clade C, but in ML and Bayesian analysis *X. largeni* is sister to Clade C only (Fig. 3).

In Clade C, MP places (*X. pygmaeus + X. cf. fraseri 1*) sister to a clade containing (*X. wittei, X. andrei, X. cf. fraseri 2, X. ruwenzoriensis, X. longipes, X. amieti, and X. boumbaensis*). Within this latter clade, *X. wittei* is sister to (*X. andrei, X. cf. fraseri 2, X. ruwenzoriensis, X. longipes, X. amieti, and X. boumbaensis*). *X. andrei* is sister to a clade within which *X. cf. fraseri 2* is sister to (*X. ruwenzoriensis, X. longipes, X. boumbaensis, and X. amieti*).

In *Silurana*, 19 out of 55 MP trees ally *S. cf. tropicalis* with (*S. epitropicalis + S. new tetraploid 1*) and the other MP trees have the same topology in *Silurana* as the ML and Bayesian analysis (Fig. 3).

### 3.2. Divergence times, rates of evolution, and ancestral areas

The most recent common ancestor of *P. pipa* and *P. parva* is thought to be the ancestor of all extant species in this genus (Cannatella and Trueb, 1988a; Trueb and Cannatella, 1986). If this is the case, we estimate that all extant members of the genus *Pipa* diversified in the New World about 55.7 mya with a 95% confidence interval of 46.5–65.7 mya (Table 2).

Extant lineages of clawed frogs, the Xenopodinae, appear to have evolved in eastern and/or central equatorial Africa well after the split of South America from Africa (Fig. 4 and Table 2). Ancestral area reconstruction of *Xenopus* is eastern Africa and of *Silurana* is central Africa. The ancestor of (*X. laevis + X. gilli*) probably originated in southern Africa. All octoploids and dodecaploids, except possibly *X. vestitus*, originated in central Africa. Under the assumptions of no extinction of ancestral tetraploids (Fig. 1d) and comprehensive sampling of extant species, all age estimates for octoploid and dodecaploid lineages in this study predate the Pleistocene. Confidence intervals of *X. amieti, X. longipes, X. boumbaensis*, and *X. wittei* have some overlap with the Pleistocene.

### 3.3. Hypothesis testing

Our analyses reject the hypothesis that octoploid clawed frogs evolved from only one polyploidization event. A parametric bootstrap test rejected Hypothesis 1 ($P < 0.001$) and the prior and posterior probabilities are near 0.0. A search for the most parsimonious trees consistent with Hypothesis 1 recovered 57 MP trees of 3395 steps in length as compared to the unconstrained tree length of 3355.

A parametric bootstrap test did not reject Hypothesis 2, a single origin of dodecaploids, ($P = 0.955$). A search for the most parsimonious trees consistent with monophyly of dodecaploid mtDNA, recovered 149 MP trees of length 3357 steps, two steps longer than the unconstrained tree. The prior probability of this hypothesis was 0.01025 and the posterior probability was 0.25313. The hypothesis of a single origin of *Silurana* tetraploids, Hypothesis 3, was rejected ($P < 0.001$) and the prior and posterior probabilities of this hypothesis are near 0.0. A search for the most parsimonious tree consistent with this hypothesis recovered 54 MP trees of length 3372, 17 steps longer than the unconstrained trees.

Under a scenario of two origins of octoploids and one origin of dodecaploids, dodecaploids could have arisen from an octoploid maternal ancestor and have mtDNA that is nested in a clade containing only octoploid mtDNA, or not, and therefore have a clade of mtDNA that is not nested within octoploid mtDNA. Both of these
scenarios were explored (Hypotheses 4 and 5, respectively). Hypothesis 4 was not rejected after Bonferroni correction ($P = 0.048$, adjusted critical value = 0.0167, $k = 9$ tests; Rice, 1989) but Hypothesis 5 was rejected ($P < 0.001$). One hundred and eighty MP trees of length 3359 steps were consistent with Hypothesis 4 and 54 MP trees of length 3370 steps were consistent with Hypothesis 5. These sets of trees are 4 and 15 steps longer than the unconstrained trees, respectively. The prior probabilities of Hypotheses 4 and 5 are about 0.0 and the posterior
probabilities are 0.02644 and 0.0000, respectively. Hypothesis 4 is consistent with a scenario in which octoploids originated twice and dodecaploids originated once from an octoploid maternal (mtDNA contributing) ancestor. It does not support the hypothesis that dodecaploids originated from a tetraploid maternal ancestor (Fig. 1b).

Hypotheses 6–8, monophyly of the laevis, muelleri, and vestittus–wittei groups, were rejected ($P = 0.007$, adjusted critical value $= 0.0125$, $k = 9$ tests; Rice, 1989; $P < 0.001$; and $P < 0.001$, respectively). The set of most parsimonious trees consistent with these hypotheses were 7, 18, and 36 steps longer, respectively, than the unconstrained tree. The prior probabilities of Hypotheses 6–8 are close to 0.0 and the posterior probabilities are 0.0019, 0.0062, and about 0.0.

A parametric bootstrap test did not reject Hypothesis 9, which postulates monophyly of the Fraseri-like group ($P = 0.556$). The most parsimonious tree consistent with this hypothesis was 3356 steps (as compared to an unconstrained length of 3355 steps). The prior probability of this hypothesis is close to 0.0 and the posterior probability is 0.0172.

4. Discussion

4.1. Phylogenetic relationships

The monophyly of Xenopus + Silurana is extremely well supported by this analysis, in accord with other studies (Cannatella and de Sá, 1993; de Sá and Hillis, 1990; Trueb and Báez, 1997), but not Cannatella and Trueb (1988a), who placed Silurana as the sister taxon of Pipinae. This last result was reconsidered by Cannatella and de Sá (1993), and there is consensus that the morphological data also supports monophyly of Xenopus + Silurana.

Monophyly of both Silurana and of Xenopus is consistent with results of other molecular and morphological studies that include fossil taxa (Cannatella and de Sá, 1993; Kobel et al., 1998; Trueb and Báez, 1997). It should be recognized that given our results, the choice between a taxonomy that recognizes only Xenopus for clawed frogs versus Xenopus and Silurana is a matter of emphasis rather than science. Although recognition of Silurana as a separate genus is not mandatory given strong support for monophyly of all clawed frogs, use of both names underscores trenchant biological and historical differences between the two clades. Each of these lineages is older than most families in the order Primates, for instance (Table 2; Yoder and Yang, 2004), and fixed chromosomal and morphological differences are present (Trueb and Báez, 1997).

Topologies recovered in this analysis correspond almost exactly to that detailed by Kobel et al. (1998) based on a smaller portion of mitochondrial rDNA and fewer species. Close relationships between X. muelleri and X. borealis and between X. laevis and X. gilli reported in our study agree with results from electrophoretic studies of sperm and serum proteins (Graf and Fischberg, 1986; Mann et al., 1982). MtDNA and nuclear DNA studies also both suggest a close relationship between X. ruwenzoriensis and X. amieti (Bürki and Fischberg, 1985; Graf and Fischberg, 1986; Mann et al., 1982)—species that have similar male vocalizations (Kobel et al., 1996).

However, few consistencies beyond monophyly of X. laevis emerge from comparisons to the unrooted mtDNA networks of Carr et al. (1987), or a reanalysis of those data (Cannatella and de Sá, 1993). To test whether inaccuracy in mapped restriction sites could explain this
difference, we examined whether the location of restriction sites inferred by Carr et al. (1987) corresponds to the locations of these sites as predicted by our sequence data. For the few comparisons that were possible, there was excellent congruence. For example, both datasets suggest two SacII sites are fixed across all taxa, and a PvuII site is present only in X. ruwenzorii. Our sequences predict more EcoRI positions (AATT) than were mapped by Carr et al. (1987), and suggest that EcoRI’ sites (GAATTC) were actually mapped. Our predictions for EcoRI’ positions relative to the first fixed SacII site correspond exactly to the map of Carr et al. (1987). Thus we dismiss differences in data quality as a source of differences in phylogeny estimation, and attribute the increased phylogenetic resolution in our study to increased sampling of taxa and polymorphic characters (Pollock et al., 2002; Zwickl and Hillis, 2002).

This mtDNA phylogeny also is not consistent with the laevis, muelleri, or vestitus–wittei species groups. These discrepancies could stem from the incomplete picture of allopolyploid evolution that is gleaned from the former (Fig. 1). For example, mtDNA paralogy of each species group could be due to the fact that group members do not share a common maternal ancestor (that contributes mtDNA) but do share a common paternal one. Alternately, the paternal ancestor of some species could be the maternal ancestor of the others. A portion of the recent shared ancestry of some species, however, can be discerned from the mtDNA phylogeny (Table 3), barring recent extinction of ancestral mtDNA lineages (Fig. 1d). Paralogy of mtDNA in species groups also could occur if characters used to define species groups are symplesiomorphic rather than synapomorphic.

4.2. Time and place of diversification

It is useful to examine the degree to which molecular estimates of the time and place of diversification are consistent with other sources of information from fossils, climatic history, geological history, and other molecules, while also avoiding circularity when some of this information was used to make the molecular estimates. Estimation of divergence times by a molecular clock is confounded by error in phylogenetic estimation and taxonomic identification of fossils, and by error and uncertainty in the age of geological events and fossils (Graur and Martin, 2004). Incomplete information in the fossil record and in the molecular phylogeny (which does not contain fossil taxa) also poses a challenge. Error in estimating ancestral distributions is expected to increase with an increasing age of a clade and this error could be compounded by potentially high dispersability of clawed frogs, some of which have close mtDNA relationships that span large geographic distances. Under the assumption of a correlation between time and molecular divergence (that is not necessarily clock-like), the wide dispersability of some species is evidenced by, for example, mtDNA of X. laevis from Nigeria and Botswana, separated by over 3500 km, that is only 0.5% divergent.

Living pipoid frogs (Rhinophrynidae and Pipidae) are distributed in North, Central, and South America, and Africa, and their ancestors were widely distributed before the breakup of Gondwana. Close relatives of the Pipidae are known from Early Cretaceous deposits of Israel (Trueb, 1999), and Middle Cretaceous strata of Argentina (Báez et al., 2000), and the age of pipoid origin is at least Late Jurassic (150 mya; Henrici, 1998). A proposed sister relationship between (Hymenochirus + Pseudhymenochirus) and the extinct African species Pachybatrachus taqueti (83.5–89.0 mya; Báez and Rage, 1998) is consistent with the hypothesis that Gondwana breakup led to the diversification of Pipinae.

We estimate that the ancestor of extant Xenopus originated between 31.8 and 54.6 mya and that all extant clawed frogs originated between 50.5 and 81.3 mya (Table 2). These time intervals predate fossils of the extinct species Xenopus arabiensis, a putative sister taxon of X. muelleri, which were deposited in the Late Oligocene of Yemen (28.5–23.8 mya; Henrici and Báez, 2001), but are well after the breakup of Gondwana about 112 mya. Other studies that assumed clock-like divergence of mtDNA sequences (Kobel et al., 1998), duplicate genes (Hughes and Hughes, 1993), or serum albumins (Bisbee et al., 1977), estimate that Xenopus is even younger. However, Estes (1975b) detailed similarities between an extinct fossil species “Xenopus” romeri from the late Paleocene of Brazil (61.0–54.8 mya; Estes et al., 1978) and fossils similar to Xenopus (=Silurana) tropicalis from Niger, implying that the ancestor of Xenopodinae existed before the separation of Africa and South America. Cannatella and de Sá (1993) placed “Xenopus” romeri as sister to either Silurana or Pipinae. A more detailed phylogenetic analysis by Trueb and Báez (1997) placed “Xenopus” romeri as sister to extant clawed frogs (Xenopus + Silurana). If “Xenopus” romeri is not nested within the Xenopodinae, it is plausible that cladogenesis of the most recent common ancestor of these taxa occurred before the breakup of Gondwana. Under the assumption that the breakup of Gondwana did lead to the divergence of Pipa from (Hymenochirus + Pseudhymenochirus), our age estimates (calibrated by this event) cast further doubt on the relationship between South American and African species proposed by Estes (1975a) because the most recent common ancestor of extant clawed frogs is younger than the formation of the Atlantic Ocean.

Similarly, Cannatella and de Sá (1993) addressed the implications of Estes’ (1975a) synonymy of the Argentine fossil Shelania pascauli (Paleogene, 47.2–43.4 mya) into Xenopus. They showed that this placement intro-
duced an inconsistency, because the split between Ar-

gentine and African *Xenopus* would then have to be

older than the Atlantic Ocean. The conundrum was re-

solved by a thorough re-analysis by Trueb and Báez

(1997), in which *Shelania pascuali* was shown to be

outside of Xenopodinae.

The following integration of postulated dates with
geological and paleoclimatological information is nec-

essarily speculative. We do not test hypotheses, but ra-

ther point to consistencies between these sources of data.

This scenario, coupled with future discoveries, may

provide testable predictions.

Reconstruction of ancestral distributions suggests that

the most recent common ancestor of extant *Xeno-
pus* is from eastern equatorial Africa and that of *Silur-

ana* is from central equatorial Africa; all clawed frogs

may be derived from an ancestor that occupied either

one or both of these areas (Fig. 4). Fifty million years

ago, Africa was south of its current position and the
equatorial rain forest was located on land that is now

the Sahara desert (Lovett, 1993). Ancient biogeographic

patterns of clawed frogs are thus best conceptualized in
terms global location relative to the equator, rather than

in terms of a fixed position on the African continent.

The extent of moist and dry habitats also varied on

Africa. Clawed frogs may have become widespread in

the Middle East and Africa when the Sahara region was

more humid (Henrici and Báez, 2001). As Africa drifted

northward, the central African plateau was uplifted

(Tertiary; 65–23 mya), and the Rift Valley opened first in

northern Africa about 30–20 mya, then in eastern Africa

by 15–12 mya, and in southern Africa by the Pliocene

(5–1.8 mya; Griffiths, 1993). During this time the equa-
torial rainforest was fragmented by savannah habitat

into the expansive western/central rainforest and the

patchy eastern rainforest (Lovett, 1993). Some divergent

species such as *X. clivii* and *X. largeni* may have had

wider ranges before the Ethiopian highlands became

isolated from other habitat by desert. A similar expla-
nation might account for the disjunct populations of *X.*

new tetraploid in Chad and Nigeria, but the relatively

low divergence between mtDNA sequences from these

regions (0.9% patristic distance, Fig. 3) suggests more

recent vicariance or dispersal is a better explanation.

During the Pleistocene, repeated changes in aridity

led to a retraction of rain forest to small refugia and an

expansion of savannah and desert habitat (Fig. 4; Ni-

chol, 1999; Stokes et al., 1997). These changes may have

affected the distributions of forest-adapted species, such

as those in the *fraseri*-like group (Kobel, 1996a). The

diverse topology and environmental conditions of the

Rift Valley and eastern Congo may have promoted in-
teraction and hybridization among species that are

adapted to different habitats and elevation (Tinsley

et al., 1996). Interestingly, octoploid and dodecaploid

species of clawed frogs are known only from areas in

Central Africa that had refugia of tropical rainforest
during the Pleistocene, although these lineages probably
originated before this period (Fig. 4 and Table 2). That

octoploids and dodecaploids (which are thought to be

allopolyploid) are endemic to, but older than, Pleisto-
cene rainforest refugia suggests that these regions acted

as “lifeboats” for some species to endure the vagaries of

climatic change, and possibly as platforms from which

lineages were able to disperse and diversify when suit-
able habitat became more widespread. This does not

support the hypothesis that rainforest refugia acted as

“species pumps” that spawned new successful lineages

(proposed for the Neotropics, Haffer, 1969; but see

Knapp and Mallet, 2003) because the diversity within

these areas predates the age of Pleistocene refugia.

We estimate that *X. laevis* originated in Southern

Africa about 6.0–19.3 mya (Fig. 4 and Table 2), al-

though the sister relationship between *X. laevis* mtDNA

from South Africa and other *X. laevis* mtDNA is not

strongly supported (Fig. 3). *Xenopus l. laevis* also has

separate mtDNA clades in the southwestern and eastern

parts of South Africa (Grohovaz et al., 1996; Measey

and Channing, 2003). Regional differences in morphol-

gy and vocal characteristics in this species may war-

rant taxonomic revision (Kobel et al., 1996; Measey

and Channing, 2003).

Marine barriers may pose formidable obstacles to
dispersal of clawed frogs. For example, populations of

*X. gilli* that are separated by only 100 km have 3.6%

patristic divergence (Fig. 3). This genetic difference is

virtually fixed in each *X. gilli* population (Evans et al.,

1997, 1998). Our estimate of the divergence times of *X.
gilli* corresponds with the Pliocene (Fig. 4 and Table 2)
and implicates a Pliocene marine inundation into the

area that is now Cape Town and the Cape Flats (Siesser

and Dingle, 1981). Patterns of differentiation similar to

*X. gilli* have also been reported in other animals

(Mounton, 1986; Stewart, 1997).

4.3. Polypliodization

Studies of HOX gene clusters suggest that complete

or partial genome duplication happened twice in all

vertebrates and once again in most fish (Holland, 1997;

Holland and García-Fernández, 1996) but subsequent

episodes of polypliodization are infrequent in animals as

compared to plants (Orr, 1990). In frogs, tetraploid

(probably mostly autopolyploid) species are known in the

Leptodactylidae, Myobatrachidae, Hylidae, Bufon-

idae, Ranidae, Arthroleptidae, and Microhylidae (Du-

ellman and Trueb, 1994; Martino and Sinsch, 2002), but

it is in the Pipidae that polypliods comprise the majority

of the species, in which species ploidy reaches one of its

highest levels (dodecaploid; Tymowska, 1991), and in

which suspected instances of allopolypliodization are

most prevalent.
Genome duplication could be advantageous if novel protein function evolves when one duplicated gene copy is released from stabilizing selection while the other carries out the duties of its single-copy precursor (Hughes, 1999; Ohno, 1970). Benefits might be gleaned from increased levels of gene expression after genome duplication (Clark, 1994). Alloployploid clawed frogs might also benefit from resistance to both sets of parasites that are specific to each of their ancestors (Jackson and Tinsley, 2003). However, polyploidization could be disadvantageous due to complications associated with the formation of multivalents during cell division, and because of difficulties associated with variation in the stoichiometry and expression levels of genes. Alloployploid speciation, which by definition must be sympatric, also poses challenges to a new alloployploid species, which potentially could be “reabsorbed” into a sympatric species via further hybridization.

Our analyses suggest polyploidization occurred at least six times in clawed frogs (Fig. 4). In Xenopus, polyploidization happened at least four times: from 18 to 36 chromosomes at least once, from 36 to 72 chromosomes at least twice, and from 72 + 36 to 108 chromosomes at least once. In Silurana, the mtDNA phylogeny supports the occurrence of two polyploidization events to generate tetraploids, though the karyotype of one tetraploid species (S. new tetraploid 2) was not available and was estimated by cloning nuclear DNA paralogs of the RAG-1 gene. A strict interpretation of the mtDNA phylogeny suggests that additional episodes of polyploidization occurred, but parametric bootstrap tests did not reject alternative hypotheses of fewer events (Hypotheses 2 and 4).

Ploidy levels of extant pipoids are higher than 20 (Duellman and Trueb, 1994), but an extinct ancestor of Xenopus is inferred with a diploid chromosome number of \(2N=18\). Thus, the most parsimonious explanation is that a pair of homologous chromosomes in a diploid ancestor of all Xenopus species fused to yield a lower fundamental number of chromosomes (\(2N=18\)). Genome duplication then yielded tetraploid, octoploid, and dodecaploid descendants.

5. Conclusions

This work supports the monophyly of clawed frogs and further division of these animals into two clades that are each descended from an ancestor with a different ploidy level: \(2N=18\) for Xenopus and \(2N=20\) for Silurana. Within Xenopus, two major mtDNA clades exist in partial sympatry and these clades are both partially sympatric with Silurana. Mitochondrial DNA phylogeography points to an origin of clawed frogs in eastern and/or central equatorial Africa, with Xenopus possibly diversifying from an ancestor in eastern Africa, and Silurana possibly diversifying from an ancestor in central Africa. This study also suggests that at least six independent polyploidization events occurred in clawed frogs; four occurred in Xenopus, and two in Silurana. Further information on nuclear genes will test whether polyploid species that share a maternal ancestor, as suggested by closely related mtDNA, are also derived from the same paternal ancestor.

Acknowledgments

We are grateful to D. Rungger, J. Montoya-Burgos, C. Thiebaud, A. Solaro, M. Picker, and M. Lebreton, and the Station de Zoologie Expérimentale, Université de Genève, for their hospitality. We also thank P. Chippindale, B. Colombelli, L. Du Pasquier, L. Du Preez, M. Fischberg, D. Foguekem, D. Hillis, M. Klemens, H. Kobel, J. Le Doux Diifo, G. Legrand, C. Lieb, C. Loumont, J. Perret, D. Rungger, M. Ruggles, E. Rungger-Brandle, C. Thiebaud, T. Titus, and J. Tymowska for their fieldwork, without which this study would have been impossible. For loans of genetic samples, we thank the following individuals and their respective institutions (museum abbreviations follow Leviton et al., 1985): J. Montoya-Burgos, S. Fisch-Muller, and A. Schmitz (MHNG), J. Vindum and A. Leviton (CAS), L. Ford (AMNH), T. LaDuc (TNHC), L. Trueb and J. Simmons (KU), and R. Murphy (ROM). We also thank M. Tobias and two anonymous reviewers for helpful comments. This work was supported by INT-01-07294 and a Columbia University Earth Institute Postdoctoral Fellowship (BJE), DEB-9981631 (DCC), NIH grant NS23684 (DBK), and the Center for Environmental Research and Conservation at Columbia University.

References


Ronquist, F., 1996. DIVA. Uppsala University, Uppsala.


