MOLECULAR ECOLOGY

Molecular Ecology (2011) **20**, 4216–4230 doi: 10.1111/j.1365-294X.2011.05262.x

The Rift Valley is a major barrier to dispersal of African clawed frogs (*Xenopus*) in Ethiopia

BEN J. EVANS,* SHIREEN M. BLISS,* SIMONE A. MENDEL*
and RICHARD C. TINSLEY†
*Department of Biology, McMaster University, Life Sciences Building Room 328, 1280 Main Street West, Hamilton, ON, L8S 4K1 Canada, †School of Biological Sciences, University of Bristol, Bristol, BS81UG UK

Abstract

The Ethiopian highlands – home to striking species diversity and endemism – are bisected by the Rift Valley, a zone of tectonic divergence. Using molecular data we examined the evolutionary history of two co-distributed species of African clawed frog (*Xenopus clivii* and *X. largeni*) that are endemic to this region. Our field collections substantially extend the known distribution of *X. largeni*, a species formerly known from highlands southeast of the Rift, but that also occurs to the northwest. In both species, analysis of mitochondrial DNA and 19 autosomal loci identifies significant population structure, suggests little or no recent migration across the Rift Valley, and provides divergence time estimates across the Rift of ~1–3.5 million years. These results indicate that the Ethiopian Rift Valley is a major obstacle to dispersal of highland-adapted amphibians.

Keywords: coalescence, demography, endemicity, migration, vicariance

Received 8 May 2011; revision received 16 July 2011; accepted 1 August 2011

Introduction

About half of Africa above 2000 m in elevation is in Ethiopia (Yalden 1983). Perhaps not surprisingly, species endemism in this country is also striking. As part of the Eastern Afromontane ecosystem, the Ethiopian highlands are considered a biodiversity hotspot (Mittermeier *et al.* 2004). These highlands include two Endemic Bird Areas (Stattersfield *et al.* 1998) and are home to numerous endemic species such as the Ethiopian wolf, the Mountain Nyala, and the Gelada. Of ~63 amphibian species that occur in Ethiopia, about 40% are endemic and the majority of these are endemic to the highlands (Largen, 2001). Uplift of the Ethiopian highlands began about 20 Ma and continued until the end of the Pliocene, and is probably due to a massive mantle plume (Ebinger & Sleep 1998; Pik *et al.* 2003; Sepulchre *et al.* 2006). The ecology of these highlands is divided into the Afro-alpine region over 3200 m above sea level (m asl) and Afro-montane regions between 900 and 3200 m asl (Egziabher 1988).

Correspondence: Ben Evans, Fax: +1 905 522 6066;
E-mail: evansb@mcmaster.ca

The Great Rift Valley is a major geological spreading zone between multiple continental plates. The Rift runs south from Lebanon, under the Red Sea, through the Ethiopian highlands, splits into an Eastern and Western Rift just north of the equator, and then converges again and continues south into central Mozambique. The portion of the Great Rift Valley that bisects the central Ethiopian highlands is called the Main Ethiopian Rift, which is ~1550–1700 m asl (Fig. 1). In southern Ethiopia the Main Ethiopian Rift intersects with the Eastern Rift in a 'broadly rifted zone' (Ebinger *et al.* 2000), which is somewhat lower (~1250 m asl). Northeast of the Main Ethiopia Rift is the Afar Depression, a low-lying region (<1000 m asl) with desert scrubland ecology. Climate in the Main Rift Valley is generally hotter and drier than in the adjacent highlands. Pliocene and Pleistocene climate change associated with glacial and inter-glacial periods had a pronounced impact on regional ecology, with montane-adapted vegetation and associated fauna periodically being lower in elevation than the present (Flenley 1979; Leakey *et al.* 1996; WoldeGabriel *et al.* 2001; Sepulchre *et al.* 2006).

African clawed frogs of the genus *Xenopus* are primarily aquatic, occasionally migrate overland during

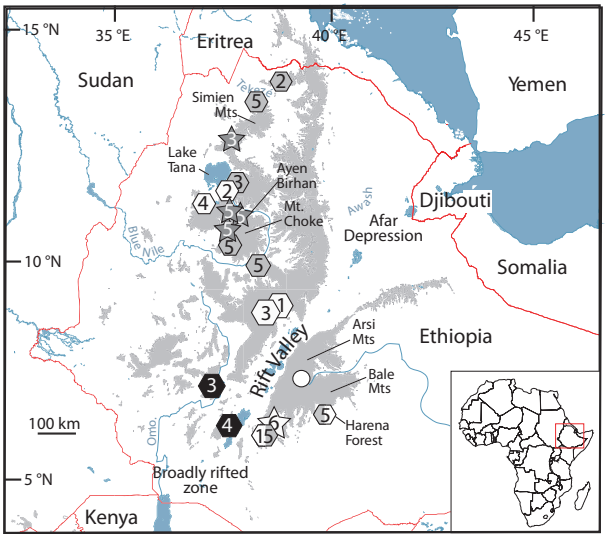


Fig. 1 Geographic setting and sample localities. Pentagons and stars indicate sample localities of *X. clivii* and *X. largeni*, respectively; numbers inside these symbols refer to the number of individuals from which mtDNA was sequenced from each locality and shading corresponds to Fig. 2. As detailed in Table S2 (Supporting information), four additional *X. clivii* individuals and one additional *X. largeni* individual were included for sequencing of some autosomal loci. Regions >2000 m asl are indicated in gray and water is blue. An unsampled locality of *X. largeni* is indicated with an open circle.

rainy periods, and are distributed over much of sub-Saharan Africa (Tinsley & Kobel 1996). Two *Xenopus* species occur in Ethiopia. *Xenopus largeni* is endemic to the Ethiopian highlands and, until now, was known from only two localities southeast of the Rift Valley (Fig. 1; Tinsley 1995; Tinsley *et al.* 1996; Largen 2001). *Xenopus clivii* is also found in the Ethiopian highlands but has a broader range that extends to Eritrea and possibly parts of Sudan and northern Kenya, although the latter two localities have not been confirmed by recent collections (Tinsley 1995). Both species are generally restricted to the Afro-montane region including, for *X. clivii*, two sites in the Main Rift Valley. *X. clivii* also is known from one location that is lower than the Afro-montane region (Godare at 820 m asl; Tinsley 1995; Largen 2001). Neither species has been reported from the Afar Depression. Evolutionary relationships between *X. clivii*, *X. largeni*, and other *Xenopus* species are not fully resolved, although available information points to divergence of both species from other extant lineages of *Xenopus* a few dozen million years ago (Evans *et al.* 2004). Estimates based on a relaxed molecular clock suggest that divergence of *X. clivii* probably occurred before uplift of the Ethiopian highlands, and that divergence of *X. largeni* probably occurred during the uplift (Evans *et al.* 2004).

In light of the dynamic geological and palaeoecological history of Ethiopia, we set out to test whether the Rift Valley is associated with significant population structure in these two frog species and to further characterize their evolutionary history. Based on their known distributions, these frog species appear to have different ecological specificities: *X. clivii* is known from a greater diversity of altitudes and a broader spectrum of habitat types. Despite this, these species do appear to have broadly similar ecological requirements (standing or slow moving water, wet conditions for overland dispersal) and they are known to co-occur in the same pond. Conceivably, past climate change and associated ecological oscillations could have affected the ranges and population structure of both species, but perhaps in different ways due to differences in ecological specificity. To further explore this, we sampled these species from widespread sites in Ethiopia, including many parts of their known ranges and also previously uncharacterized portions of their distributions. We collected sequence data from mitochondrial DNA and 19 autosomal genes and used phylogenetic and population genetic methods to test for population structure, and for evidence of ongoing migration across the Rift Valley.

Materials and methods

Fieldwork and molecular data

In September and October of 2008, *X. clivii* and *X. largeni* were collected by BJE. *Xenopus clivii* samples are composed of 41 individuals collected northwest of the Rift Valley and 21 from the southeast. *X. largeni* samples are composed of 19 individuals collected northwest of the Rift and six from the southeast, including the type locality (Fig. 1). Genetic samples are deposited at the Museum of Comparative Zoology at Harvard University and morphological voucher specimens are deposited at the University of Addis Ababa.

Approximately 768 base pairs of nucleotide sequence were obtained from a portion of the 16S rDNA gene of the mitochondrial genome from 57 *X. clivii* and 24 *X. largeni* samples using 16sc and 16sd primers (Evans *et al.* 2004). Nineteen autosomal loci were sequenced from 10 – 58 *X. clivii* individuals and 16 – 24 *X. largeni* individuals, drawn almost entirely from the samples for which mitochondrial sequences were obtained, as detailed in Tables 1 and S1 (Supporting information). Amplifications were performed using primers detailed in Supporting information. *X. clivii* and *X. largeni* are tetraploid species and for some loci two paralogs were co-amplified in preliminary data collection. In order to obtain polymorphism data from individual genes, when possible we re-designed paralog-specific primers from

Table 1 Molecular polymorphism data for *X. clivii* and *X. largeni* including locus name (Locus), number of base pairs of data (bp), number of silent sites (silent bp), mutation rate scalar (μ scalar), number of individuals from the northwest and southeast (nNW and nSE), private, shared, and fixed segregating sites for MIMAR analysis (s1, s2, ss, sf; see MIMAR documentation for explanation), and Genbank accession numbers (Accession)

	X. clivii										*X. largeni*									
Locus	bp	Silent bp	μ Scalar	nNW	nSE	s1	s2	ss	sf	Accession	bp	Silent bp	μ Scalar	nNW	nSE	s1	s2	ss	sf	Accession
Hypothetical protein LOC100158283	630	132	2	24	30	6	0	0	4	JN223982–JN224008	636	133	1.82	24	10	5	1	2	3	JN223965–JN223981
Sodium-dependent glucose transporter 1 (kiaa1919)	705	182	1.16	24	28	1	6	1	8	JN224181–JN224206	–	–	–	–	–	–	–	–	–	–
Protein arginine methyltransferase 6 (prmt6)	660	154	1.54	28	30	6	3	0	1	JN224275–JN224303	660	155	1.73	24	10	2	0	0	4	JN224304–JN224320
Mannosyl-oligosaccharide glucosidase (mogs)	660	161	1.31	14	28	0	2	0	7	JN224064–JN224084	660	161	1.28	24	10	3	1	2	1	JN224085–JN224101
Chromosome 7 open reading frame 25 (c7orf25)	531	117	1.03	28	24	3	3	1	1	JN223686–JN223711	531	118	0.96	24	10	2	0	0	3	JN223712–JN223728
Zinc finger, BED-type containing 4	471	110	0.74	28	26	1	1	0	4	JN224412–JN224438	471	110	0.83	24	10	1	1	0	1	JN224439–JN224455
Nuclear receptor interacting protein 1 (nrip1)	477	99	1.46	26	20	4	0	0	2	JN224207–JN224229	–	–	–	–	–	–	–	–	–	–
BTB domain protein 6 (p7e4)	522	123	0.62	30	26	3	1	0	1	JN224230–JN224257	522	122	0.6	24	10	2	1	0	2	JN224258–JN224274
Open reading frame 125 (c9orf125)	480	119	0.74	26	26	1	1	0	0	JN223729–JN223754	480	118	0.6	26	10	0	0	0	1	JN223755–JN223772
Fem-1 homolog c (fem1c)	495	113	0.79	26	22	4	1	0	3	JN223845–JN223869	495	112	0.7	24	10	3	0	0	2	JN223870–JN223886
Zinc finger protein 238, gene 2 (znf238.2)	–	–	–	–	–	–	–	–	–	–	525	117	0.57	24	10	1	1	0	1	JN224456–JN224472
B-cell CLL/lymphoma 9 (bcl9)	471	110	0.96	26	26	1	0	0	3	JN223616–JN223641	471	110	1.05	24	10	2	0	4	2	JN223599–JN223615
Androgen receptor	306	76	0.79	74	42	1	1	0	1	JN223616–JN223598	306	76	1.48	38	10	1	0	0	0	JN223517–JN223615
Recombination activation gene 2	975	220	1.2	64	42	4	12	6	0	JN224321–JN224373	975	219	1.18	36	8	3	0	2	8	JN224374–JN224394
NLRC3-like (LOC100487412)	351	351	0.93	30	30	6	8	0	0	JN224009–JN224037	–	–	–	–	–	–	–	–	–	–
DnaJ (Hsp40) homolog, subfamily C, member 6 (dnajc6)	380	380	0.78	26	30	4	6	2	5	JN223791–JN223818	380	380	0.74	26	10	1	3	0	8	JN223773–JN223790
EF-hand calcium binding domain 5 (efcab5)	157	148	1.42	4	16	0	1	2	2	JN223835–JN223844	157	148	1.41	22	10	3	0	5	2	JN223819–JN223834
Mediator complex subunit 26 (med26)	315	83	0.66	26	26	3	0	0	0	JN224038–JN224063	–	–	–	–	–	–	–	–	–	–
High-affinity lysophosphatidic acid receptor	–	–	–	–	–	–	–	–	–	–	430	103	1.34	24	10	0	0	0	2	JN223904–JN223920
BHLH transcription factor atonal homolog 2	420	100	0.34	28	26	2	0	0	0	JN223642–JN223668	420	100	0.38	24	10	1	0	0	0	JN223669–JN223685
Gravin-like (gl)	–	–	–	–	–	–	–	–	–	–	507	110	0.69	24	10	2	0	0	2	JN223887–JN223903
Potassium voltage-gated channel, Shal-related subfamily, member 3	516	114	0.52	28	26	2	0	0	1	JN223921–JN223947	516	114	0.61	24	10	0	0	0	1	JN223948–JN223964
Velo1 protein (velo1)	–	–	–	–	–	–	–	–	–	–	459	92	1.02	24	10	0	0	0	2	JN224395–JN224411

cloned sequences. If paralog-specific amplification was not possible for all individuals within a species, we did not include data in the analysis from that locus for that species.

Phylogenetic analysis

Phylogenetic analysis of mitochondrial DNA sequences was performed using MrBayes version 3.2 (Huelsenbeck & Ronquist 2001) including representative sequences from all known species in the genus *Xenopus*, including those undescribed. A sequence from *Silurana tropicalis* was used as an outgroup. A model of evolution was selected by the Akaike Information Criterion with MrModelTest version 2 (Nylander 2004). Two Markov Chain Monte Carlo (MCMC) runs were performed with four chains per run, each for 10 million generations, with the temperature parameter set to 0.2. To assess convergence, post-run diagnostics were performed using R (R development core team 2005) and Tracer version 1.5 (Drummond & Rambaut 2007) including overlay plots of the posterior distribution of likelihoods and model parameters from independent chains, and calculation of the effective sample size of the post burn-in likelihoods and parameters. Based on these diagnostics a burn-in of one million generations was discarded.

Structure analysis

Population structure causes nonrandom associations between alleles of different loci (linkage disequilibrium) and an excess of homozygotes (Hardy–Weinberg disequilibrium). We used an approach to evaluate population structure that has a general aim of minimizing linkage and Hardy–Weinberg disequilibrium (Structure analysis; Pritchard *et al.* 2000) and an extension of this analysis where individuals are assigned to populations under a Dirichlet process prior (Structurama; Huelsenbeck & Andolfatto 2007). These approaches are appealing because they do not require *a priori* assignment of individuals to populations. We used Structure and Structurama to evaluate population structure in *X. clivii* and *X. largeni* using the autosomal data only. Mitochondrial DNA data were not included in these analyses because many slightly divergent haplotypes were recovered and also to make the results more comparable to the MIMAR analysis described below. Sequences were collected from samples from both sides of the Rift for all loci, but missing data from each locus ranged from 0–84% for *X. clivii* and 0–35% for *X. largeni* with a mean of 51% and 28% respectively. To explore the impact of missing data, we also re-analysed the data for *X. clivii* after discarding 12 individuals to bring the average missing data from each locus down to 33%.

For Structure and Structurama analysis, the phase of single nucleotide polymorphisms in autosomal DNA was inferred using Phase version 2.1.1 (Stephens *et al.* 2001; Stephens & Donnelly 2003) with the default settings including 100 iterations, a thinning interval of one, and a burn-in of 100. 'Best guess' estimates of unique alleles recovered from Phase summarize the posterior distribution of haplotypes by trying to minimize the number of differences between the guesses and the true haplotypes. These 'best guess' estimates were analysed using Structure version 2.3.3 (Pritchard *et al.* 2000) under the 'admixture' model, assuming no correlation between alleles, and using no prior information about sample localities. The 'admixture' model allows individuals to have mixed ancestry and is recommended for populations with the potential for genetic exchange and hybrid zones. Eight runs were performed for each value of K beginning with different random seeds, each for 200 000 generations, and with a burn-in of 10 000 generations discarded. The maximum likelihood of the eight runs is reported. Post-run inspection of the likelihood of the data and the value of the alpha parameter (the level of admixture) were stable, suggesting appropriate mixing of the Markov chain. For Structure, the algorithm calculates the likelihood of the data given a designated number of populations (K) and also estimates for each individual the posterior probability of membership in each of the K populations. A range of values of K are considered and, to evaluate support for alternative values of K, we followed the *ad hoc* suggestions of Pritchard *et al.* (2000).

Unlike Structure analysis, Structurama treats K as a random variable with a Dirichlet process prior (Huelsenbeck & Andolfatto 2007). The prior for the Dirichlet process has a clustering parameter (α) that influences the prior expectation for the number of populations. For each species, we performed Structurama analyses under the 'no admixture' model using α values from 1 to 5. For each value of α, two MCMC chains were run for 1 million generations with the heating parameter at the default value of 0.2 and 100 000 generations were discarded as burn-in. The 'no admixture' model assumes each individual is from only one of the K populations.

MIMAR analysis

In order to test for evidence of genetic exchange between populations on either side of the Rift, we fitted the data to two demographic models using the March 3, 2009 version of MIMAR (Becquet & Przeworski 2007). The null model (hereafter the 'no migration' model) is one in which a single ancestral population diverges into two descendant populations with no

subsequent gene flow between the descendant populations after divergence. The alternative model (hereafter the 'migration' model) permits ongoing gene flow between the descendant populations after divergence.

MIMAR is a Bayesian approach that uses MCMC to update model parameters (Becquet & Przeworski 2007). The models include polymorphism parameters for the ancestral and both descendant populations (θ_A, θ_{NW}, θ_{SE}; subscripts refer to the ancestral, northwest, and southeast populations), divergence time (τ), and (for the migration model) the magnitude of gene flow in terms of effective number of migrants per generation in each direction (from the northwest population to the southeast population and *vice versa*). Instead of directly analysing sequence data, MIMAR uses four summary statistics (s_1, s_2, s_s and s_f) that are affected by the parameters of the model. These summary statistics describe the distribution among populations of derived mutations, including the number that are fixed in one population or the other (s_f), the number that are polymorphic in one population but not the other (s_1) or *vice versa* (s_2), and the number that are polymorphic in both populations (s_s) (Wakeley & Hey 1997). We considered only synonymous sites in this analysis. Mitochondrial DNA data (which included no synonymous sites) were therefore excluded, and the analysis was restricted to the data from 19 autosomal markers from each species.

Polarization of polymorphisms as ancestral or derived was achieved using the baseml program of PAML version 4.4 (Yang 1997) with *S. tropicalis* and another *Xenopus* species as an outgroup and the ingroups fitted to a star phylogeny following Foxe *et al.* (2009). A mutation rate scalar for each locus was based on divergence From the pipid frogs *Hymenochirus curtipes* and *Pipa carvallhoi*. The number of synonymous and nonsynonymous substitutions were estimated with PAML (Yang 1997) using a free ratio model and assuming phylogenetic relationships reported in Evans *et al.* (2004) among *Rhinophrynus*, *Hymenochirus*, *Pipa*, *Xenopus*, and *Silurana*. Assuming *Hymenochirus* diverged from *Pipa* ~102 Ma as South America rifted from Africa (an average of estimates provided by Pitman III *et al.* 1993; Maisey 2000; McLoughlin 2001; Sanmartín & Ronquist 2004; Sereno *et al.* 2004; Ali & Aitchison 2008) and a generation time of 1 year, we estimated a mutation rate (μ) of 4.10×10^{-9} synonymous substitutions per synonymous site per generation. For MIMAR analysis, a variable rate of intralocus recombination was assumed to follow an exponential distribution with a mean of 0.6, in consideration of the findings of Becquet & Przeworski (2007) that error in the rate of recombination did not have a large impact provided some level of intralocus recombination was included. Lower and upper boundaries for parameter priors were adjusted so that the MCMC sampling was not constrained, and command lines for MIMAR analysis which specify these priors are provided in Supporting information. Convergence and mixing of the Markov chain was evaluated for two independent runs with different random seeds by plotting the likelihood and parameter values versus generation of the chain.

For each species, the fit of the data to the 'no migration' and the 'migration' models was evaluated with goodness of fit tests (Becquet & Przeworski 2007). Simulations were performed with recombination using parameters estimated by each model, and goodness of fit assessed based on how often summary statistics from these simulations matched summary statistics from the observed data. The summary statistics are the four polymorphism statistics that were used in the MIMAR analysis, and also Tajima's D (Tajima 1989), F_{ST} (calculated following Hudson *et al.* 1992), and nucleotide diversity (π) (Nei & Li 1979). We performed goodness of fit tests by drawing from the posterior distributions of the parameter values and also by using median parameter values from the posterior distribution as a point estimate.

IMa2 analysis

For comparative purposes, we also used the program IMa2 (Hey 2010) to estimate divergence times and to test for evidence of migration in *X. clivii* and *X. largeni*. Similar to MIMAR, IMa2 evaluates nested models with and without migration using a MCMC approach. In contrast to MIMAR, uses the entire data instead of summary statistics and assumes no recombination within loci. Also unlike MIMAR, IMa2 allows one to fit the data to demographic models with more than two descendant populations. However, because of concerns with over-parameterization given the size of our dataset (see program documentation) we restricted our analysis to models with only two descendant populations.

Prior to IMa2 analysis, we used the four gamete test as implemented in DNAsp version 5.10.1 (Rozas *et al.* 2003) to test for recombination within each locus. Evidence of recombination was detected in four loci of *X. largeni* and eight loci in *X. clivii*; these were deleted from IMa2 analysis. We also calculated Tajima's D, Fu and Li's D, and Fu and Li's F*, also with DNAsp, in order to test for evidence of non-neutral evolution. No *X. largeni* loci showed significant evidence of natural selection. However, after deleting the loci with evidence of recombination, four additional loci in *X. clivii* had evidence of non-neutral evolution based on at least one of these tests. We therefore performed two separate IMa2 analyses for *X. clivii*, one including and one not including these four loci.

For intronic data we used the same mutation rate per site as the MIMAR analysis and converted this to a mutation rate per locus by multiplying by the number

of sites sequenced at each locus. For exonic data we used the same 113 autosomal loci dataset from Bewick *et al.* (Bewick AJ, Chain FJJ, Heled J, Evans BJ unpublished data) to estimate the mutation rate across nonsynonymous and synonymous sites, and arrived at a value of 1.1×10^{-9} substitutions per site per generation; this was also converted to a rate per gene in the same way as the introns. We excluded the mitochondrial DNA loci to make the results of this analysis more comparable to the MIMAR results.

For IMa2 analysis, three separate MCMC analyses were performed for each species, each starting with a different random seed, for 10 000 000 generations, and saving genealogies every 100 generations after discarding a burn-in of 300 000 generations. Forty chains per analysis were run using settings for 'medium heating' for a medium sized dataset as suggested in the program documentation. We considered a uniform and an exponential prior for migration. The uniform prior for migration was set to a maximum of 0.7 for *X. largeni* and 0.8 for *X. clivii* and the exponential prior for migration was set to have a mean of 1 for both species, given our expectations based on the low or absent level of migration inferred from MIMAR analysis (see below). Results from both migration priors were similar and those from the uniform prior are reported here. The prior for theta was uniform and based on preliminary runs was set to a maximum of 15 and 21 for *X. largeni* and for *X. clivii*, respectively. The prior for divergence time was also uniform and set to a maximum of 6 and 12 for *X. largeni* and for *X. clivii*, respectively. The infinite sites model was assumed when possible except for six loci in *X. largeni*, which required the use of the HKY model due to multiple substitutions at a single position. Convergence and mixing of the Markov chain was evaluated by plotting the likelihood and parameter values versus generation of the chain. Support for ongoing migration after divergence was evaluated using the likelihood ratio test statistic as described by Nielsen & Wakeley (2001) and using a mixture of chi square distributions for tests with boundary conditions as described by Goldman & Whelan (2000).

For this analysis, a MCMC simulation was run with asymmetric migration between the descendant populations in 'M mode'. Following this, nested models without migration between one or both descendant populations were evaluated in 'L mode' (see program documentation for details). Commands used in IMa2 are provided in Supporting information.

Morphological analyses and museum specimens

As a descriptive exercise aimed at further characterizing morphological divergence on either side of the Rift, 12 external measurements were taken from *X. clivii* on 12 adult females (seven from the northwest and five from the southeast) and 13 adult males (five from the northwest and eight from the southeast) following Tinsley (1973). Eleven of these measurements were standardized by dividing by the snout-vent length. The data were then transformed so that each measurement had a mean of zero and a variance of one, and then subjected to principal components analysis using R (R development core team 2005). All *X. largeni* specimens collected northwest of the Rift were either tadpoles or newly metamorphosed froglets, so adult morphology of this species was not analysed.

To provide data on geographical distribution and morphological variation in Ethiopian *Xenopus* species, preserved collections were examined in the American Museum of Natural History, New York (AMNH), Museum of Comparative Zoology, Harvard (MCZ), Field Museum of Natural History, Chicago (FMNH), California Academy of Sciences, San Francisco (CAS), Natural History Museum, London (BM), Museo Zoologico dell' Università, Florence (MZUF), the United States National Museum of Natural History, Washington (USNM), and Muséum National d'Histoire Naturelle, Paris (MNHN).

Results

Distributions of X. clivii *and* X. largeni

Detailed maps of the distribution of *X. clivii* and *X. largeni* based on museum specimens are provided in Largen (2001) and Tinsley *et al.* (1996). Despite a relatively extensive representation of *Xenopus* species from other areas of Africa, most of the museum collections studied contained few records from Ethiopia. For example, for six of the institutions (AMNH, MCZ, FMNH, CAS, USNM and MNHN), the combined total of *X. clivii* was only 12 specimens. For all except two of these, localities specified were in the area of Addis Ababa. Collections from Ethiopia in MZUF were also from this region. BM collections provided more comprehensive documentation (see Largen 2001), but again predominantly from the Addis Ababa environs. Although *X. clivii* was originally described from Eritrea, Largen (2001) reported only two localities within Ethiopia (Adua and Gorgora) that were substantially north of Addis Ababa, underscoring the paucity of collections from this region. There are two records of *X. clivii* from the floor of the Rift Valley: Lake Langano, 1585 m asl (BM 1969.973) and Lake Zwai, 1650 m asl (MCZ 51 457). Our fieldwork provided a series of new localities for *X. clivii* in the highlands Northwest of the Rift (Fig. 1 and Supporting information).

All published records of *X. largeni* are from the mountains east and south-east of the Rift Valley (BM collections listed in Tinsley 1995; Largen 2001). Our museum studies identified previously unpublished specimens from the area in which the paratypes were collected (AMNH A–158387-158389; details in Supporting information). Our field studies included collections of *X. largeni* from the type locality but not the other previously known locality, which is a stream in the Arussi (Arsi) Mountains at ~2500 m asl (Fig. S1, Supporting information; Tinsley 1995; Tinsley *et al.* 1996; Largen 2001). Additionally, the present field collections substantially extend the known range of *X. largeni* to include areas northwest of the Rift Valley including two localities north of Mount Choke but south of Lake Tana, and a third locality north of Lake Tana but south of the Simien Mountains (Fig. 1).

Significant molecular differentiation across the Rift Valley

Mitochondrial DNA haplotypes of *X. clivii* and *X. largeni* are reciprocally monophyletic and diverged on either side of the Rift Valley; in both species no unique mitochondrial DNA haplotype was found on both sides of the Rift Valley (Fig. 2). The level of divergence between intraspecific clades in both species is comparable to that seen between some sister species (for example between *X. vestitus* + *X. lenduensis* or between *X. wittei* + *X. itombwensis*) and similar to that observed within *X. laevis* (sensu Kobel *et al.* 1996), a clade considered by some to include multiple species (Frost 2010).

In *X. clivii*, mitochondrial lineages are geographically structured on each side of the Rift. Two lineages from the southeast are each restricted to different locations. In the northwest one lineage was observed only in the south and two others were observed only in the north, with one locality (Korata) having both mitochondrial lineages in different individuals from the same pond. Little or no sequence divergence was observed between *X. clivii* mitochondrial DNA haplotypes on either side of major rivers such as the Omo, Blue Nile, and Takezé. In *X. largeni*, little molecular variation was observed in mitochondrial DNA within the northwest population or within the southeast population.

Structure analysis of autosomal DNA supported four differentiated populations in *X. clivii* and three in *X. largeni*. The four populations of *X. clivii* comprise two northwest of the Rift and two southeast of the Rift. That the main division in both species corresponds to the Rift Valley is evinced by Structure analysis with $K = 2$ (Fig. 3). In *X. clivii*, further division between two northwest populations corresponded with the precipitous Blue Nile Valley south of Mt. Choke (Fig. 1), even

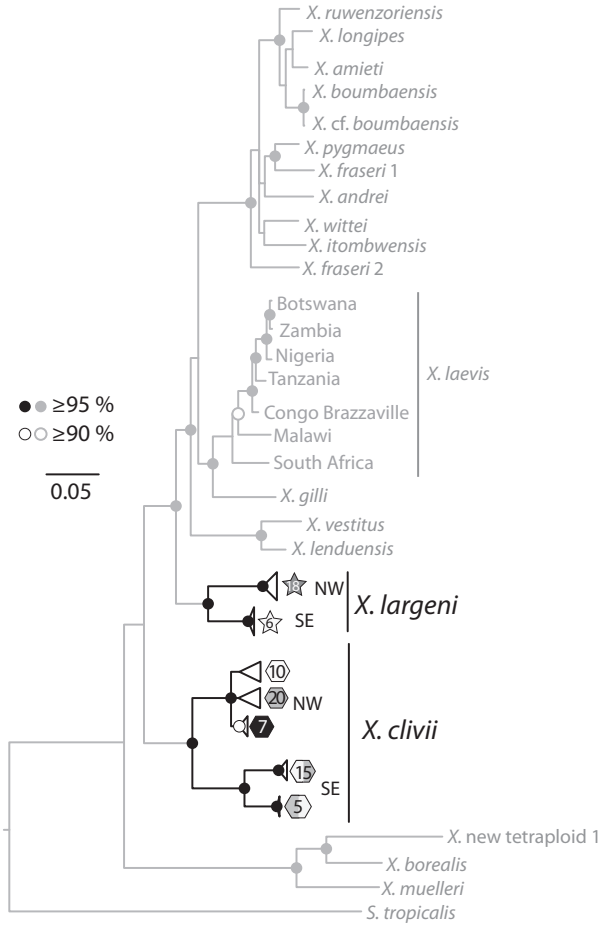


Fig. 2 Mitochondrial DNA phylogeny of *Xenopus* based on data from Evans *et al.* (2004) but also including population sampling of *X. largeni* and *X. clivii*, which are distinguished by dark lineages. Samples collected northwest (NW) and southeast (SE) of the Rift Valley form well supported sister clades in both species. Nodes with greater than 90 or 95% posterior probability are indicated with open and filled circles, respectively, a scale bar indicates the number of substitutions per site on branches, and symbols refer to haplotype clades sampled from localities in Fig. 1 with numbers inside symbols indicating the number of samples.

though differentiation in mitochondrial DNA was fairly low across this valley. Differentiation between the two southeast populations of *X. clivii* was somewhere between the two southeast sampling localities (the Harena forest and a southern portion of the southeast highlands; Fig. 1). Structurama analysis also recovered support for four populations in *X. clivii* (Table 2), with two exceptions. The first was when the Dirichlet process prior had a mean of 1, in which case Structurama supported a separate population for each individual in the analysis ($K = 57$). The other exception was when the Dirichlet process prior had a mean of five, in which case there was slightly higher posterior probability for

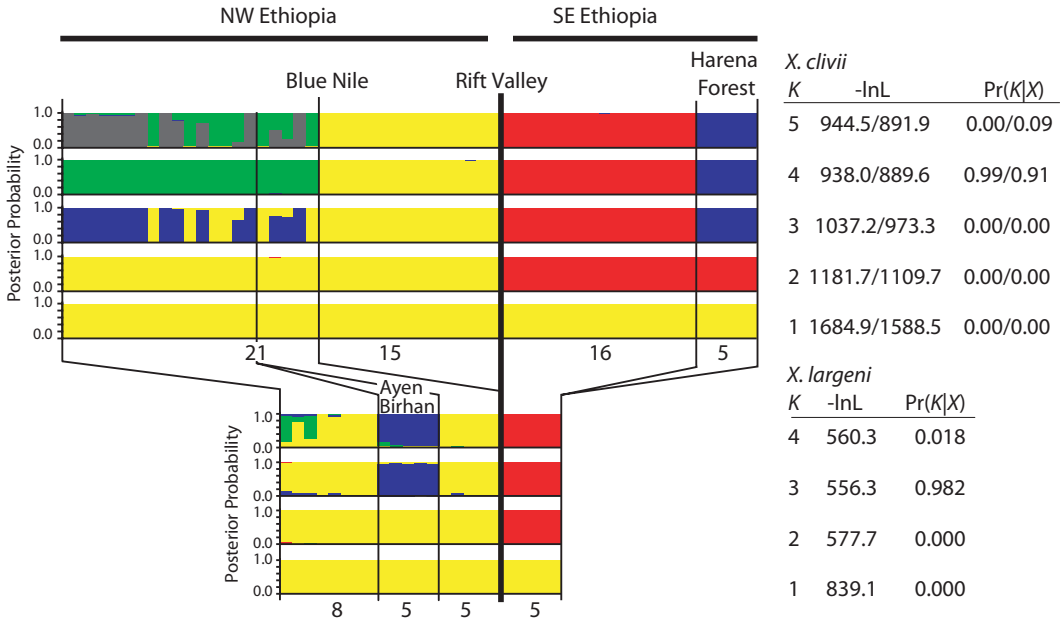


Fig. 3 Structure analysis of autosomal loci supports four differentiated populations in *X. clivii* and three in *X. largeni*. Nineteen autosomal loci are analysed for each species. The likelihoods (-lnL) of five or four populations (*K*) were considered for *X. clivii* and *X. largeni*, respectively, and an approximation of the probability of each number of populations given the data (Pr(*K* | *X*)) calculated with the *ad hoc* method of Pritchard *et al.* (2000), and contingent on the associated simplifying assumptions discussed therein. For each value of *K*, the posterior probability of each individual being a member of each population is graphed for each individual analysed. A total of 57 individuals were analysed for *X. clivii* and 23 individuals for *X. largeni*; the number of samples from different parts of Ethiopia is indicated below the graphs for each species. For *X. clivii*, results from analysis of the reduced dataset follow a slash.

five than four populations (Table 2). Results from re-analysis of the reduced *X. clivii* dataset with less missing data were qualitatively identical to the analysis of the full dataset; support for four populations was somewhat weaker in Structure analysis (Fig. 3) but stronger in Structurama analysis (Table 2).

In *X. largeni*, Structure and Structurama analysis both supported the presence of three populations with the most substantial differentiation being between samples collected northwest and southeast of the Rift Valley (Table 2, Fig. 3). Population subdivision is presumably related to the Rift Valley although the gap in sampling prevents us from dismissing the Blue Nile Valley south of Mt. Choke (Fig. 1) as an additional contributor to differentiation. Five individuals sampled in Ayen Birhan (Fig. 1, Table S1, Supporting information) appear significantly differentiated from other populations from the northwest according to both analyses (Table 2, Fig. 3). F_{ST} between populations northwest and southeast of the Rift was $\sim$ 0.6 for *X. clivii* and *X. largeni* (Table 3).

Using MIMAR, we fitted molecular polymorphism data from 19 autosomal loci to models with and without migration across the Rift Valley. In general for both species the data failed to reject the 'no migration' model. In *X. largeni*, both models provided reasonable fits for all summary statistics, suggesting that the addition of migration parameters was not necessary. More specifically, for this species, the observed values fell within the 95% confidence limits of the simulated values generated in the goodness of fit tests (Table 3). In *X. clivii*, neither the 'migration' nor the 'no migration' model provided a good fit for F_{ST} or Tajima's D when a point estimate of parameters was used in the goodness of fit simulations (Table 3). When simulations were performed by drawing from the posterior distribution, Tajima's D of the northwest population provided a significantly poorer fit for the 'migration' model but not for the 'no migration' model (Table 3). In both species the posterior distributions for migration parameters are near zero for both directions of migration (Fig. 4 and Table 4).

Confidence intervals for the times of divergence in each species overlap, although the mean estimate is more recent for *X. largeni* ($\sim$800 000 years ago) than *X. clivii* ($\sim$1 200 000 years ago) (Table 4). Both species appear to have a larger effective population size northwest of the Rift compared to southeast, although confidence intervals for this parameter estimate from each region overlap. Dividing the values for θ_{NW} and θ_{SE} in Table 4 by 4 μ provides an estimate of the northwest and southeast population sizes for *X. clivii* of roughly 300 000 and 200 000 individuals respectively, and for

Table 2. Posterior probability of the number of populations in the sample (*K*) using different mean values (*k*) for the mean of the Dirichlet process prior *E (k)*. For *X. clivii*, a slash separates results from analysis of the full and the reduced dataset.

	k				
K	1	2	3	4	5
X. clivii					
1	0.00/0.00	0.00/0.00	0.00/0.00	0.00/0.00	0.00/0.00
2	0.00/0.00	0.00/0.00	0.00/0.00	0.00/0.00	0.00/0.00
3	0.00/0.00	0.00/0.00	0.00/0.00	0.00/0.00	0.00/0.00
4	0.00/0.00	0.87/1.00	0.69/0.99	0.57/1.00	0.49/1.00
5	0.00/0.00	0.13/0.00	0.30/0.01	0.41/0.00	0.50/0.00
6	0.00/0.00*	0.00/0.00	0.01/0.00	0.02/0.00	0.01/0.00
X. largeni					
1	0.00	0.00	0.00	0.00	0.00
2	1.00	0.57	0.30	0.16	0.08
3	0.00	0.37	0.55	0.56	0.48
4	0.00	0.05	0.14	0.24	0.36
5	0.00	0.01	0.01	0.04	0.08
6	0.00	0.00	0.00	0.00	0.00

*When E (*k*) was equal to one, the posterior probability was 1.00 for 57 *X. clivii* populations, i.e. every sample being from a different population.

X. largeni of roughly 200 000 and 100 000 individuals respectively. For both species the mutation parameter of the ancestral population (θ_A) is estimated by MIMAR analysis to be ~5–18-fold larger than that of either descendant population (θ_{NW}, θ_{SE}, Table 4; Fig. 4). These θ_A values translate to effective population sizes of 1 000 000 and 1 100 000 individuals for *X. clivii* and *X. largeni*, respectively.

We also fitted the data to the migration and no migration model using the program IMa2. For this analysis, we had to exclude loci from both species (4 and 8–12 loci, respectively, for *X. largeni* and *X. clivii*) due to evidence of recombination and natural selection. There are both consistencies and differences between results of IMa2 and MIMAR analysis. The most prominent difference is that IMa2 supported the 'migration' model over the 'no migration' model for both species. We inferred this using two likelihood ratio tests. The first (IMa2 Test 1) compares a model with an asymmetrical rate of migration to one with a symmetrical rate of migration, and the second (IMa2 Test 2), compares the model with a symmetrical rate of migration to a model with no migration. Both of these tests favoured the more parameterized model for both species (IMa2 Test 1, $P = 0.040$ and 0.009 for *X. clivii* and *X. largeni*, respectively; IMa2 Test 2: $P < 0.001$ for *X. clivii* and for *X. largeni*). Despite this difference, MIMAR and IMa2 analyses were consistent in the sense that both inferred a very low (or non-existent) level of migration across the Rift, with the exception of *X. clivii* migration from the northwest to the southeast population that was inferred by IMa2 (Fig. S1, Supporting information).

Another difference is that IMa2 analysis suggests that the divergence times across the Rift are much older for both species than the MIMAR estimates. The mean IMa2 divergence estimate for *X. clivii* is 3 588 920 years ago and for *X. largeni* is 2 508 678 years ago (95% high-

Table 3 Predictive posterior probabilities from simulations generated by sampling the posterior distribution of model parameters and from point estimates of model parameters

	s_1	s_2	s_s	s_f	F_{ST}	π_{NW}	π_{SE}	D_{NW}	D_{SE}
X. clivii (posterior)									
No migration	0.318	0.510	0.591	0.466	0.042	0.492	0.897	0.969	0.956
Asymmetric migration	0.326	0.475	0.512	0.466	0.037	0.532	0.874	0.978*	0.965
X. clivii (point estimate)									
No migration	0.272	0.529	0.633	0.413	0.014*	0.506	0.940	0.976*	0.964
Asymmetric migration	0.272	0.591	0.801	0.299	0.007*	0.586	0.973	0.986*	0.974
X. largeni (posterior)									
No migration	0.213	0.518	0.895	0.443	0.160	0.318	0.583	0.185	0.501
Asymmetric migration	0.194	0.504	0.870	0.870	0.178	0.300	0.550	0.196	0.499
X. largeni (point estimate)									
No migration	0.162	0.547	0.949	0.409	0.114	0.304	0.636	0.182	0.507
Asymmetric migration	0.144	0.551	0.945	0.414	0.116	0.295	0.635	0.194	0.506
Observed values for *X. clivii*	52	46	12	43	0.559	0.995	1.140	0.631	0.559
Observed values for *X. largeni*	32	8	15	45	0.690	0.620	0.166	−0.129	0.003

s_1, s_2, s_s, and s_f segregating sites statistics used in the MIMAR analysis as in Table 1. F_{ST}, π_{NW}, π_{SE}, D_{NW}, and D_{SE} reflect population structure, nucleotide polymorphism, and Tajima's D statistic of the northwest (NW) and southeast (SE). Observed values are indicated below the predictive posterior probabilities; values for segregating sites are the sum over all loci whereas other statistics are the means over all loci.

*Statistics for which the observed values fall outside of the 95% confidence limits of the simulated values.

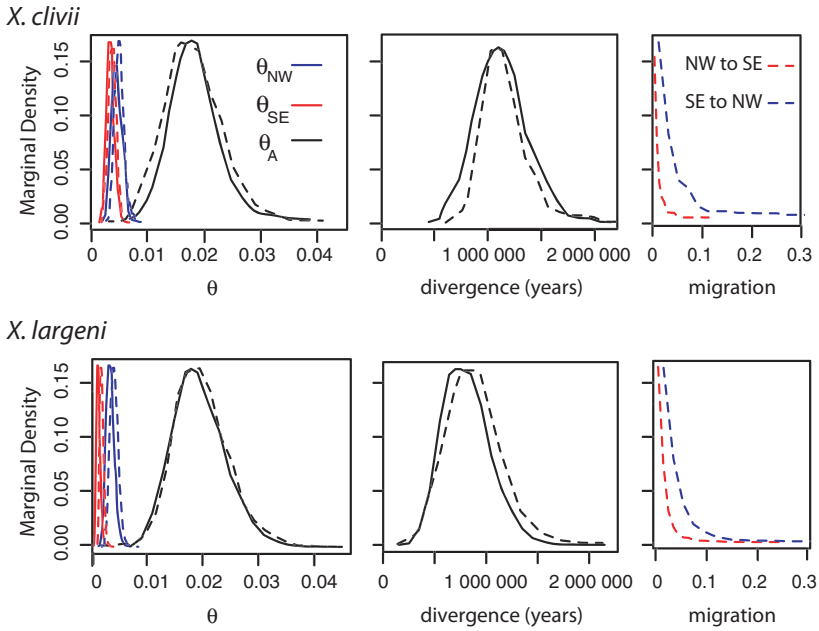


Fig. 4 Smoothed marginal posterior distributions estimated from MIMAR analysis. Posterior distributions from the models with no migration and with asymmetric migration are indicated with solid and dashed lines, respectively. Subscripts distinguish estimated population parameter θ for the northwest (NW), southeast (SE) and ancestral (A) populations.

Table 4. Medians and 95% confidence intervals of parameters estimated from MIMAR analysis of autosomal data. Parameters of the model with asymmetric migration (M) and no migration (noM) are listed for *X. clivii* and *X. largeni* including the population parameters θ for the northwest, southeast, and ancestral populations (θ_{NW}, θ_{SE}, and θ_A respectively), the divergence time in years between the northwest and southeast population (t) and, for the model with migration, the migration rate from the northwest to the southeast and from the southeast to the northwest populations (m_{NW_SE} and m_{SE_NW} respectively)

	θ_{NW}	θ_{SE}	θ_A	t	m_{NW_SE}	m_{SE_NW}
X. clivii						
M	0.0048 (0.0034–0.0067)	0.0036 (0.0022–0.0050)	0.0171 (0.0026–0.0291)	1306510 (786036–4298780)	0.0019 (0.0037–0.0434)	0.0042 (0.0004–0.1885)
noM	0.0046 (0.0032–0.0064)	0.0036 (0.0024–0.0052)	0.0181 (0.0108–0.0281)	1128400 (693250–1623100)	—	—
X. largeni						
M	0.0033 (0.0020–0.0049)	0.0010 (0.0005–0.0019)	0.0188 (0.0107–0.0304)	820478 (429644–1508290)	0.0025 (0.0004–0.0519)	0.0029 (0.0004–0.0680)
noM	0.0032 (0.0019–0.0049)	0.0020 (0.0005–0.0019)	0.0190 (0.0114–0.0303)	783906 (427054–1305830)	—	—

est posterior density (HPD) intervals are not reported because it did not appear contiguous; Fig. S1, Supporting information). However, MIMAR and IMa2 both suggest that *X. clivii* divergence across the Rift is somewhat older than *X. largeni*.

MIMAR and IMa2 analyses are also similar in that both suggest that population size is higher northwest of the Rift compared to southeast of the Rift in both species, and that ancestral population sizes were much larger than the descendant populations. The mean IMa2 estimates of population size northwest and southeast of the Rift was 213 242 (HPD interval: 99 730–331 102) and 165 847 (HPD interval: 75 794–275 254) respectively for *X. clivii* and 262 875 (HPD interval: 149 012–383 626) and 71 257 (HPD interval: 22 193–142 671), respectively, for *X. largeni*. The mean IMa2 estimate of ancestral population size for *X. clivii* was 2 300 531 (HPD interval not reported because it did not reach low levels at the upper limit) and for *X. largeni* was 772 271 (HPD interval: 0–2 850 249). Thus the magnitude of the effective population size estimates are fairly similar for MIMAR and IMa2. Results of the IMa2 analysis of the smaller

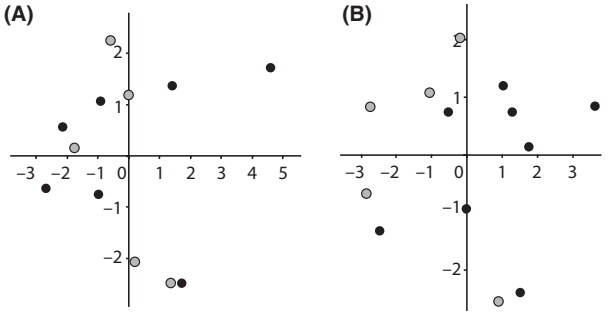


Fig. 5 Principal components analysis of eleven size standardized variables does not detect substantial differences between (A) female or (B) male *X. clivii* from northwest and southeast of the Rift (gray and black dots, respectively). For each sex, the first and second principal components are plotted which together account for 63% and 53% of the variance in the data.

dataset for *X. clivii* that excluded loci with evidence of natural selection were similar to the analysis of the larger *X. clivii* dataset in that the symmetrical migration model was favoured over the no migration model (IMa2 Test 2, $P < 0.001$), with the exception that the model with asymmetrical migration was not significantly favoured over the model with symmetrical migration (IMa2 Test 1, $P = 0.134$).

Morphological variation in X. clivii

Raw morphological data for *X. clivii* are available in Table S2 (Supporting information). Size of female *X. clivii* did not differ significantly on either side of the Rift ($P = 0.921$, Students *t*-test, $n = 12$) but adult males were significantly smaller in the northwest ($P = 0.001$, $n = 13$). However, principal components analysis of size-transformed data does not provide strong evidence of morphological differentiation in either sex. The first and second principal components of each population, for example, overlap extensively (Fig. 5) and similar results were obtained by comparing the first and third or the second and third principal components.

Discussion

We report here results of an evolutionary analysis of two sympatric species of frog endemic to the Ethiopian highlands, *X. clivii* and *X. largeni*. Our field collections substantially extend the known distribution of *X. largeni* (Tinsley 1995; Tinsley *et al.* 1996; Largen 2001) to include portions of the Ethiopian highlands northwest of the Rift Valley. We also provide records of *X. clivii* from new localities northwest of the Rift. Population assignment algorithms (Structure and Structurama) suggest the existence of four and three populations in *X. clivii* and *X. largeni*, respectively, with both species exhibiting high population differentiation across the Rift Valley. Coalescent analyses of polymorphism data from up to 19 autosomal loci (MIMAR and IMa2) suggests that recent migration across the Rift Valley in both species is very low or absent and that divergence between populations on either side is ~1–3.5 million years. Phylogenetic analysis of mitochondrial DNA sequences suggests reciprocal monophyly on either side of the Rift in both species, which is also consistent with the hypothesis of little or no gene flow across the Rift.

Within *X. clivii* and *X. largeni*, divergence of intraspecific mitochondrial DNA clades across the Rift is similar in magnitude to that seen between some sister species. However, at this time we do not view these results as warranting taxonomic revision for multiple reasons. One reason for not proposing separate species status for populations on either side of the Rift Valley is that marked morphological differences between *X. clivii* populations were not revealed by principal components analysis, and it is conceivable that significantly smaller male *X. clivii* from the northwest population stems from subadults in our small sample. Second, with respect to results from MIMAR analysis, failure to reject a null hypothesis of no migration is not proof that this null hypothesis is true, particularly because gene flow that occurs ephemerally after divergence tends to go undetected by MIMAR analysis (Becquet & Przeworski 2009). Additionally, the model with migration was not rejected by IMa2 analysis, although this analysis did suggest a very low level of migration except from the northwest to the southeast population of *X. clivii* when the larger dataset was analysed. Thus the different methodology of MIMAR and IMa2 or the different portions of the data analysed by these methods led to somewhat different inferences concerning the magnitude of migration between populations. These discrepancies probably are related to the substantial differences in the data analysed (only silent sites for 19 loci for MIMAR but nonsynonymous, synonymous, and other putatively silent sites for a subset of these loci for IMa2), and also the substantial differences in how the data are analysed in by each program (use of summary statistics and accommodating recombination for MIMAR versus using the sequence data directly and disallowing recombination for IMa2).

Intraspecific divergence estimates provided by MIMAR and IMa2 also were quite different, with the IMa2 estimate being roughly 3-fold older (Fig. 3, Fig. S1, Supporting information). A major caveat to divergence estimates from both of these analyses is that they rely on assumptions that are undoubtedly inaccurate. The mutation rates used in this study were based on (i) the estimated time of the separation of South America and

Africa, (ii) an assumption that this geological event triggered divergence of *Pipa* and *Hymenochirus*, and (iii) an assumption that the generation time of pipid frogs is 1 year. Because they ignore the presence of ancestral polymorphism in the most recent common ancestor of *Pipa* and *Hymenochirus*, the mutation rates used here could be too high, which would lead to an underestimated time of divergence. For IMa2 analysis we used different mutation rates for protein coding regions and for introns. However, the intensity of purifying selection varies considerably among protein coding regions, and the variance in mutation rate of protein coding regions is therefore higher than that of silent sites. Perhaps more importantly, the rate of divergence between species of coding regions could be substantially slower than the rate that polymorphisms appear at the population level if purifying selection frequently removes these polymorphisms before they fix. This could explain why IMa2, which considers variation at all sites, estimates older divergent times, as compared to MIMAR, for which we analysed only variation at putatively neutral or almost neutral sites (synonymous sites and introns). It is also conceivable that generation time in pipids is longer than 1 year, which would cause the estimated mutation rate to be too low, and exaggerate divergence time estimates from MIMAR and IMa2 analysis. This could happen, for example, if females occasionally skip breeding seasons while their eggs mature, or if some breeding seasons typically produced no offspring due to environmental conditions or tadpole cannibalism (McCoid & Fritts 1980; Tinsley & McCoid 1996). Thus, while we view these divergence estimates as worth reporting, especially because they are based on a large dataset of independent sequence data (Bewick AJ, Chain FJJ, Heled J, Evans BJ unpublished data), they must be interpreted with considerable caution because they do not accommodate uncertainty or inaccuracy in the calibration point or generation time or, for IMa2 analysis, the effects of natural selection on evolution of protein coding regions. That said, our inferences about relative divergence times of *X. clivii* and *X. largeni* are independent of the estimate of the rate of mutation, although they do assume a relatively constant rate of mutation in each lineage. Moreover, the relative divergence times of *X. clivii* and *X. largeni* and also the relative magnitudes of population sizes within each species that were estimated by MIMAR and IMa2 were highly consistent.

In any case, these results argue strongly that the Rift Valley is a barrier to migration of African clawed frogs. However, it is not clear why this is the case, especially for *X. clivii*. Some *X. clivii* specimens in this study were collected at elevations lower than most of the major Rift Valley Lakes of Ethiopia (Table S1, Supporting information) and *X. clivii* is known from Godare at an elevation of 820 m asl, suggesting that altitude-specific ecological conditions of the Rift are not an absolute barrier to migration for this species. At least two populations of *X. clivii* have been recorded in the Rift Valley (Largen 2001), although unfortunately samples from these localities were not obtained for this study. Our sampling dramatically extends the known distribution of *X. largeni*, consistent with a wider ecological tolerance than the previously known distribution would suggest, although this species clearly is not as widespread as *X. clivii*. Despite these differences in extant distributions and presumably in ecological specificity, the timing of differentiation across the Rift are quite similar in both species. Thus, whatever factors influenced *Xenopus* migration across the Rift did not appear to have a substantially distinct impact on *X. clivii* as compared *X. largeni*. This differs, for example, from the lowland fynbos ecosystem in coastal portions of Cape Province, South Africa, where differentiation of *Xenopus gilli* is much higher than differentiation of sympatric *X. laevis* populations (Evans *et al.* 1997).

For *X. clivii* and *X. largeni*, the inferred population size of the ancestral population was much larger than both descendant populations according to MIMAR and IMa2 analysis. An inference of a large ancestral effective population size relative to descendant populations is fairly common in studies of natural populations, and could either reflect real features of their evolutionary histories or be an artifact of coalescent based analyses such as MIMAR and IMa (Hey & Nielsen 2004, 2007; Becquet & Przeworski 2007, 2009). One possible cause of inflated estimates of θ is violation of model assumptions, for example by ancestral population structure. Becquet & Przeworski (2009) considered the impact of three scenarios involving departures from the standard isolation-migration models including ancestral population structure prior to divergence with no gene flow (allopatric speciation), and ephemeral gene flow between differentiated populations after divergence, either immediately after divergence (parapatric speciation) or after a period of no gene flow following divergence (secondary contact). For the allopatric model but not the parapatric speciation or secondary contact models, longer periods of ancestral population structure before population divergence were associated with higher estimated values of θ_A. In the data from *X. clivii* and *X. largeni*, recent or recurrent bottlenecks in the descendant populations on either side of the Rift could also account for the small descendant θs compared to θ_A. Population bottlenecks could have been caused, for example, by climatological transitions or by metapopulation dynamics on either side of the Rift. This *post hoc* speculation could be explored further by testing

whether other species in Ethiopia sharing their aquatic habitat with *Xenopus* have a similar demographic signature of recent population size decline.

Goodness of fit simulations for MIMAR analysis highlight intricacies of the evolutionary histories of *X. clivii* and *X. largeni*. These simulations suggest that some observed summary statistics for *X. clivii* did not match expectations. For example, observed values of Tajima's D were significantly higher in *X. clivii* than in simulations (Table 3). In a structured population, coalescence occurs more rapidly within a subpopulation but slowly between subpopulations, reducing rare frequency alleles and increasing intermediate frequency alleles. With a finite number of demes, population structure has negligible impact on Tajima's D when samples are randomly selected from all demes, although Tajima's D can be higher when samples are selected from only one deme (De & Durrett 2007). Given what we know about the range of *X. clivii*, it seems unlikely that unsampled demes play a major role in estimates of Tajima's D reported here. Recent bottlenecks are also transiently associated with positive Tajima's D (Fay & Wu, 1999). Another explanation for unexpectedly high Tajima's D in *X. clivii*, of course, is non-neutral evolution of sites linked to the synonymous positions we analysed (Tajima 1989).

Goodness of fit simulations for MIMAR analysis also suggested that the observed F_{ST} of *X. clivii* was significantly lower than simulated values (Table 3). F_{ST} was calculated following Hudson *et al*. (1992) as $(\pi_{between}-\pi_{within})/\pi_{between}$, where $\pi_{between}$ and π_{within} are the average nucleotide divergence between and within populations respectively. As such, the lower than expected observed F_{ST} in *X. clivii* is clearly caused by population structure in both descendant populations (as revealed by Structure and Structurama analysis; Table 2, Fig. 3), which increases π_{within}. Strasburg & Rieseberg (2010) considered a violation of the standard isolation-migration model by population structure in both descendant lineages using the program IMa (Hey & Nielsen, 2004, 2007) and Evans (2011) considered violation by population structure in one descendant lineage using MIMAR. Both of these studies found that population structure increased estimates of the effective population size of the structured descendant populations, but did not impact the mean estimates of other parameter estimates such as divergence time and ancestral θ. For this reason we do not anticipate that contemporary population structure in *X. clivii* substantially affected parameter estimates other than the descendant θs (θ_{NW} and θ_{SE}). Population structure can increase θ in situations where differentiated populations rarely go extinct, or it can decrease θ if differentiated demes have a metapopulation structure characterized by extinction and re-colonization (Nei & Takahata 1993; Whitlock & Barton 1997). Thus, in addition to recent bottlenecks, extinction/re-colonization metapopulation dynamics northwest and southeast of the Rift could also account for the large difference between descendant and ancestral θ estimates.

Significant differentiation across the Ethiopian Rift Valley also has been reported in other species. The frog genus *Paracassina*, for example, is endemic to Ethiopia and includes two species; one (*P. obscura*) occurs west of the Rift and other (*P. kounhiensis*) occurs east (Largen 2001). In the Gelada, mitochondrial DNA control region divergence on either side of the Rift is about 10% (Belay & Mori 2006). Mitochondrial DNA of the Ethiopian wolf is also significantly differentiated across the Rift, though not reciprocally monophyletic (Gottelli *et al*. 2004). Similarly, two populations of the plant *Arabis alpine* each exhibit significant differentiation across the Rift (Assefa *et al*. 2007). In the naked mole rat *Heterocephalis glaber*, divergence between mitochondrial DNA on either side of the Rift Valley exceeds that found among samples from Kenya (Faulkes *et al*. 2004). In the African clawed frogs examined here, the level of differentiation and lack of substantial migration across the Rift Valley is suggestive of reproductive isolation of populations on either side of the Rift. Further research on migration within the Rift Valley, mating preferences of individuals on either side, and differentiation of morphology and vocalization with clarify the taxonomic status of these populations and possible causes for subdivision within both species.

Acknowledgments

This research was supported by grants to BJE from the Canadian Foundation for Innovation (#10715) and the National Science and Engineering Research Council (#RGPIN-283102-07) and by the Museum of Comparative Zoology at Harvard University. We thank D. Pawlos and Temesdan for field assistance, B. Carstens, B. Charlesworth, L. Stevison, B. Zimkus, and three reviewers for advice or helpful comments on an earlier version of this manuscript, and B. Golding for access to computational resources.

References

Ali JR, Aitchison JC (2008) Gondwana to Asia: plate tectonics, paleogeography and the biological connectivity of the Indian sub-continent from the Middle Jurrasic through latest Eocene (166-35 Ma). *Earth-Science Reviews*, **88**, 145–166.

Assefa A, Ehrich D, Taberlet P, Nemomissa S, Brochmann C (2007) Pleistocene colonization of afro-alpine 'sky islands' by the arctic-alpine *Arabis alpina*. *Heredity*, **99**, 133–142.

Becquet C, Przeworski M (2007) A new approach to estimate parameters of speciation models with application to apes. *Genome Research*, **17**, 1505–1519.

Becquet C, Przeworski M (2009) Learning about modes of speciation by computational approaches. *Evolution*, **63**, 2547–2562.

Belay G, Mori A (2006) Intraspecific phylogeographic mitochondrial DNA (D-loop) variation of Gelada baboon, *Theropithecus gelada*, in Ethiopia. *Biochemical Systematics and Ecology*, **34**, 554–561.

De A, Durrett R (2007) Stepping-stone spatial structure causes slow decay of linkage disequilibrium and shifts the site frequency spectrum. *Genetics*, **176**, 969–981.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Ebinger CJ, Sleep NH (1998) Cenozoic magmatism throughout east Africa resulting from impact of a single plume. *Nature*, **395**, 788–791.

Ebinger CJ, Yemane T, Harding DJ, *et al.* (2000) Rift deflection, migration, and propagation: linkage of the Ethiopian and Eastern rifts, Africa. *Geological Society of America Bulletin*, **112**, 163–176.

Egziabher TBG (1988) Part 3: vegetation and wildlife; vegetation and environment of the mountains of Ethiopia: implications for utilization and conservation. *Mountain Research and Development*, **8**, 211–216.

Evans BJ (2011) Coalescent-based analysis of demography: applications to biogeography on Sulawesi. In: *Biotic Evolution and Environmental Change in Southeast Asia* (eds Gower DJ, Johnson KG, Richardson JE, *et al.*), in press. Cambirdge University Press, Cambridge.

Evans BJ, Kelley DB, Tinsley RC, Melnick DJ, Cannatella DC (2004) A mitochondrial DNA phylogeny of clawed frogs: phylogeography on sub-Saharan Africa and implications for polyploid evolution. *Molecular Phylogenetics and Evolution*, **33**, 197–213.

Evans BJ, Morales JC, Picker MD, Kelley DB, Melnick DJ (1997) Comparative molecular phylogeography of two *Xenopus* species, *X. gilli* and *X. laevis*, in the southwestern Cape Province, South Africa. *Molecular Ecology*, **6**, 333–343.

Faulkes CG, Verheyen E, Verheyen W, Jarvis JUM, Bennett NC (2004) Phylogeographical patterns of genetic divergence and speciation in African mole-rats (Family: Bathyergidae). *Molecular Ecology*, **13**, 613–629.

Fay JC, Wu C-I (1999) A human population bottleneck can account for the discordance between patterns of mitochondrial versus nuclear DNA variation. *Molecular Biology and Evolution*, **16**, 1003–1005.

Flenley J (1979) *The Equatorial Rain Forest: A Geological History*. Butterworths, London.

Foxe JP, Slotte T, Stahl EA, *et al.* (2009) Recent speciation associated with the evolution of selfing in *Capsella*. *Proceedings of the National Academy of Sciences, USA*, **106**, 5241–5245.

Frost DR (2010) *Amphibian Species of the World: an Online Reference. Version 5.4 (8 April, 2010)*. American Museum of Natural History, New York, NY, USA.

Goldman N, Whelan S (2000) Statistical tests of gamma-distributed rate heterogeneity in models of sequence evolution in phylogenetics. *Molecular Biology and Evolution*, **17**, 975–978.

Gottelli D, Marino J, Sillero-Zubiri C, Funk SM (2004) The effect of the last glacial age on speciation and population genetic structure of the endangered Ethiopian wolf (*Canis simensis*). *Molecular Ecology*, **13**, 2275–2286.

Hey J (2010) Isolation with migration models for more than two populations. *Molecular Biology and Evolution*, **27**, 905–2010.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences*, **104**, 2785–2790.

Hudson RR, Slatkin M, Maddison WP (1992) Estimation of levels of gene flow from DNA sequence data. *Genetics*, **132**, 583–589.

Huelsenbeck JP, Andolfatto P (2007) Inference of population structure under a Dirichlet process model. *Genetics*, **175**, 1787–1802.

Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**, 754–755.

Kobel HR, Loumont C, Tinsley RC (1996) The extant species. In:*The Biology of Xenopu* (eds Tinsley RC, Kobel HR). pp. 9–33, Clarendon Press, Oxford.

Largen MJ (2001) Catalogue of the amphibians of Ethiopia, including a key for their identification. *Tropical Zoology*, **14**, 307–402.

Leakey MG, Feibel CS, Bernor RL, *et al.* (1996) Lothagam: a record of faunal change in the Late Miocene of East Africa. *Journal of Vertebrate Paleontology*, **16**, 556–570.

Maisey JG (2000) Continental break up and the distribution of fishes of Western Gondwana during the Early Cretaceous. *Cretaceous Research*, **21**, 281–314.

McCoid MJ, Fritts TH (1980) Observations of feral populations of *Xenopus laevis* (Pipidae) in southern California. *Bulletin of the Southern California Academy of Sciences*, **79**, 82–86.

McLoughlin S (2001) The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. *Australian Journal of Botany*, **49**, 271–200.

Mittermeier RA, Robles Gil P, Hoffman M, *et al.* (2004) *Hotspots Revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions*. CEMEX, Mexico City.

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences, USA*, **76**, 5269–5273.

Nei M, Takahata N (1993) Effective population size, genetic diversity and coalescence time in subdivided populations. *Journal of Molecular Evolution*, **37**, 240–244.

Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics*, **158**, 885–896.

Nylander JAA (2004) *MrModeltest v2. Evolutionary Biology Centre*. Uppsala University, Uppsala.

Pik R, Marty B, Carignan J, Lavé J (2003) Stability of the Upper Nile drainage network (Ethiopia) deduced from (U-Th)/He thermochronometry: implications for uplift and erosion of the Afar plume dome. *Earth and Planetary Science Letters*, **215**, 73–88.

Pitman III WC, Cande S, LaBrecque J, Pindell J (1993) Fragmentation of Gondwana: the separation of Africa from South America. In:*Biological Relationships Between Africa and South Americ* (ed Goldblatt P). pp. 15–34, Yale University Press, New Haven, CT.

Pritchard JK, Stephens M, Donnelley JM (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

R development core team (2005) *R: a language and environment for statistical computing, reference index version 2.2.1*. R Foundation for Statistical Computing, Vienna, Austria.

Rozas J, Sanchez-DelBarrio JC, Messegyer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Sanmartín I, Ronquist F (2004) Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Systematic Biology*, **52**, 216–243.

Sepulchre P, Ramstein G, Fluteau F, *et al.* (2006) Tectonic uplift and Eastern Africa aridification. *Science*, **313**, 1419–1423.

Sereno PC, Wilson JA, Conrad JL (2004) New dinosaurs link southern landmasses in the Mid-Cretaceous. *Proceedings of the Royal Society of London Series B*, **271**, 1325–1330.

Stattersfield AJ, Crosby MJ, Long AJ, Wege DC (1998) *Endemic Bird Areas of the World: Priorities for Biodiversity Conservation*. BirdLife International Cambridge, UK.

Stephens M, Donnelly P (2003) A comparison of bayesian methods for haplotype reconstruction. *American Journal of Human Genetics*, **73**, 1162–1169.

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **86**, 978–989.

Strasburg JL, Rieseberg LH (2010) How robust are 'Isolation with Migration' analyses to violations of the IM model? *A simulation study. Molecular Biology and Evolution*, **27**, 297–310.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Tinsley RC (1973) Studies on the ecology and systematics of a new species of clawed toad, the genus *Xenopus*, from western Uganda. *Journal of Zoology, London*, **169**, 1–27.

Tinsley RC (1995) A new species of *Xenopus* (Anura: Pipidae) from the highlands of Ethiopia. *Amphibia-Reptilia*, **16**, 375–388.

Tinsley RC, Kobel H (1996) *The Biology of Xenopus*. Clarendon Press, Oxford.

Tinsley RC, Loumont C, Kobel HR (1996) Geographical distribution and ecology. In: *The Biology of Xenopus* (eds Tinsley RC, Kobel HR). pp. 35–59, Clarendon Press, Oxford.

Tinsley RC, McCoid MJ (1996) Feral populations of *Xenopus* outside Africa. In: *The Biology of Xenopus* (eds Tinsley RC, Kobel HR). pp. 81–94, Clarendon Press, Oxford.

Wakeley J, Hey J (1997) Estimating ancestral population parameters. *Genetics*, **145**, 847–855.

Whitlock MC, Barton NH (1997) The effective size of a subdivided population. *Genetics*, **147**, 427–441.

WoldeGabriel G, Halle-Selassie Y, Renne PR, *et al.* (2001) Geology and palaeontology of the Late Miocene Middle Awash valley, Afar rift, Ethiopia. *Nature*, **412**, 175–178.

Yalden DW (1983) The extent of high ground in Ethiopia compared to the rest of Africa. *Sinet: Ethiopian Journal of Science*, **6**, 35–39.

Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *CABIOS*, **13**, 555–556.

B.J.E. is an evolutionary biologist with research interests in gene and genome duplication, sex chromosome evolution, phylogeography, and evolutionary genetics. S.M.B. and S.A.M. were undergraduate students at McMaster who participated in this research as part of their Undergraduate thesis projects. R.C.T. is a biologist and parasitologist with interests in the evolution of host–parasite interactions, species diversity, and hybridization in amphibians and fish.

Data accessibility

Genbank accessions JN223517– JN224472.

Dryad: Data deposited in dryad include the following: (i) Sequence alignment for MrBayes analysis of mitochondrial DNA, (ii) Input files for *X. clivii* and *X. largeni* for MIMAR analysis, (iii) Input files for *X. clivii* and *X. largeni* for IMa2 analysis, including for *X. clivii* a smaller input file that has loci with evidence for non-neutral evolution removed, (iv) Input files for Structure analysis, including for *X. clivii* a smaller file that as a few individuals with lots of missing data removed, (v) Input files for Structurama analysis. Dryad doi: 10.5061/dryad.3s92c

Supporting information

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

Table S1. Information on specimens analyzed in this study including species and location relative to the Rift (Species (location)), field number, Specific location, latitude and longitude, and altitude in meters above sea level. Some data were not recorded (NR) and an asterisk following the field number denotes specimens for which mitochondrial DNA sequence data was not obtained

Table S2. Morphological measurements of X. clivii individuals from northwest and southeast of the Rift Valley. All measurements are in millimeters and were taken on the right side of the animal when possible; descriptions of measurements are provided in Tinsley (1973)

Figure S1. Smoothed marginal posterior density distributions from IMa2 analysis. MCMC was performed with migration, which was favored by likelihood ratio tests for both species. Distributions for effective population size (N_e) are shown for the northwest (NW), southeast (SE), and ancestral (A) populations.

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

Supplementary Information

Primers used for each locus, previously unreported specimens examined, commands used in MIMAR and IMa2 analysis, data on collection localities, and morphological measurements for *X. clivii*.

Primers

hypothetical protein LOC100158283 (*X. clivii* & *X. largeni*)

exon2_for1	5' ACA TCA GGG AGA TAC GCT ATA CGT GCA GGG 3'
exon2_rev1	5' ATC GCT GAA GCT CTT GAT GCA CTC G 3'
exon2_rev2	5' CAG GTG GGA GAT GAC GCT GAA GC 3'

sodium-dependent glucose transporter 1 (kiaa1919) (*X. clivii*)

exon3_for1	5' GTC TAT GTT ATC CTA TAT AGT CAT TG 3'
exon3_rev1	5' TCC AGW RCT GGA GAG CAA AGC 3'

protein arginine methyltransferase 6 (prmt6) (*X. clivii* & *X. largeni*)

exon4_for1	5' GAC CRS GAG TAT TTC CAG TGC TAC TC 3'
exon4_rev1	5' CAT AYG GCG ACG TMG ATA AAG TGA C 3'

mannosyl-oligosaccharide glucosidase (mogs) (*X. clivii* & *X. largeni*)

exon5_for2	5' CTG AAG ATG AGC GGC ATG TGG ATC TG 3'
exon5_rev2	5' CTT CAG CCA TGA TTA GTA CCA C 3'

chromosome 7 open reading frame 25 (c7orf25) (*X. clivii* & *X. largeni*)

exon7-for	5' CTG GTG GTT GAT GTT GTT GC 3'
exon7-rev	5' GTG GAA GCA CCT TTT CTT G 3'

zinc finger, BED-type containing 4 (*X. clivii* & *X. largeni*)

exon14.for.a	5' CAA TTT GTT CTG CCG ACT CA 3'
exon14.rev.a	5' TGT CCG ACT GCT CAT CCA TA 3'

nuclear receptor interacting protein 1 (nrip1) (*X. clivii*)

exon17.for.a	5' GAG GGC ATG AAA TCT CCA AA 3'
exon17.rev.a	5' TGC CAG CTA CGT GAC TTC TG 3'

BTB domain protein 6 (p7e4) (*X. clivii* & *X. largeni*)

exon19.fora	5' AGG TTT GCC AAT CAC TCC AG 3'
exon19.rev.a	5' TCT GTC ATT CCC TCC TGT C 3'

open reading frame 125 (c9orf125) (*X. clivii* & *X. largeni*)

exon20.for.a	5' CAG TGG TTC CTC TCC TTC CA 3'
exon20.rev.a	5' AGA ATA CGC ATG GGT TCT G 3'

fem-1 homolog c (fem1c) (*X. clivii* & *X. largeni*)
exon21.for.a 5’ TTT GTT GTC GTT TGC AGA GC 3’
exon21.rev.a 5’ TGT GCG AAT TCG TAG AGT C 3’

zinc finger protein 238, gene 2 (znf238.2) (*X. largeni*)
exon25.for.a 5’ CAA GCC GGT AGA CTC TGA GG 3’
exon25.rev.a 5’ TCC ATT TCA TCC TCG CTT T 3’

B-cell CLL/lymphoma 9 (bcl9) (*X. clivii* & *X. largeni*)
exon26.for.a 5’ ATC CAC AGG AGC AAA ACC AC 3’
exon26.rev.a 5’ TTG AGA CTG AGC CAT CAT C 3’

Androgen receptor (*X. clivii* & *X. largeni*)
XLAR_for_40 5’ AGG GCT CGG CGG GGT ATA CAA ACA GC 3’
XLAR_rev_431 5’ GGC GCT ATC AGA GAT GCC TTC G 3’

Recombination activation gene 2 (*X. clivii* & *X. largeni*)
RAG2FOR.1 5’ ACCTACACAGTTGCTGTGATG 3’
RAG2REV.1423 5’ TCTTCATCGTCTTCATTGTA 3’

NLRC3-like (LOC100487412) (*X. clivii*)
scaf351.for.8b 5’ GAA GTC TGA YTG TGA AGT G 3’
scaf351.rev.8a 5’ CCG CAC ACC TTC TGA GCC A 3’

DnaJ (Hsp40) homolog, subfamily C, member 6 (dnajc6) (*X. clivii* & *X. largeni*)
scaf4.for.10d 5’-GAT CTA GTC CTA AGC AGC C (PCR & sequencing)
scaf4.for.10c 5’-TCT MCA CAC TAC AAG AAG TCC (PCR & sequencing)
scaf4.rev.10d 5’-CAA GGT CAG CAA ATG GAT C (PCR &sequencing)
intron10_internal_for1 5’ RAT GGG ARA MGT RAG AAA GTG GG 3’
intron10_internal_for2 5’ TCC WCC TTC AGT RTC TAT CCA G 3’

EF-hand calcium binding domain 5 (efcab5) (*X. clivii* & *X. largeni*)
scaf482.for.11c 5’ AGT TCT GGC TMT GCA GTT C 3’
scaf482.rev.11c 5’ GAA SRT KAG CTA TTA CTT C 3’
scaf482.for.11a 5’ TGC AGT TCC AAG AAT GGA C 3’
scaf482.rev.11a 5’ GAT TTT AAK GCA CTT TGT ACC 3’

mediator complex subunit 26 (med26) (*X. clivii*)
exon31_for1 5’ CCT GCT GAR AAT GAC AAG CA 3’
exon31_rev1 5’ CAG TAC GRG GSG TTT GTG AG 3’

high-affinity lysophosphatidic acid receptor (*X. largeni*)
exon32_for1 5’ AGC TGG CTM TTT GGA ACT CA 3’
exon32_rev1 5’ ATT TGG TGW GGC CTC TGA AG 3’

bHLH transcription factor atonal homolog 2 (*X. clivii* & *X. largeni*)
exon33_for1 5’ GTA GGA TGC ATG GGC TCA AT 3’

exon33_rev1 5’ CCT TCA AAC TGT GGG CTT GT 3’

gravin-like (gl) (*X. largeni*)
exon34_for1 5’ AGA GGA AGM TGG GGA GAA CC 3’
exon34_rev1 5’ AAG CTT TYK TAA AGG GCT ACC 3’

potassium voltage-gated channel, Shal-related subfamily, member 3 (*X. clivii* & *X. largeni*)
exon35_for1 5’ GCM CCA GCT GAC AAG AAC AA 3’
exon35_rev1 5’ CTT TCA CCA CAK GGA AGC TC 3’

velo1 protein (velo1) (*X. largeni*)
exon38_for1 5’ TTT CCG AAA AAC CAC AGA GG 3’
exon38_rev1 5’ GTG YGT CAT AAG CAC CCT TT 3’

Previously unreported specimens examined

American Museum of Natural History

Three specimens collected by Melanie Stiassny and Darrel Frost in 1995; identified as "*Xenopus* sp.":

A- 158387. *X. largeni*, Bale. 1 km E. Dodola, Ethiopia Female
A- 158388. *X. largeni*, Bale. 8 km E. Adaba, Ethiopia Female
A- 158399. *X. largeni*, Bale. 8 km E. Adaba, Ethiopia Male.

California Academy of Sciences
Two specimens:
CAS 144468. *X. clivii*, Addis Ababa, Shewa Province, Ethiopia; possibly collected by E. Degan.

CAS 162113. *X. clivii*, Bale Province, Ethiopia, Near Shawe River 6° 40 N, 39° 44 E. 1950 m Male. Collected 1986.

Commands used for MIMAR analysis

X. clivii, no migration

./mimar 4000 10000 19 -lf clivii_in.txt -u 4.1e-9 -t u 0.000005 0.03 -n u 0.000005 0.015 -ej u 0 20000000 -N u 0.0000001 0.08 -i 1000 -x 10 -y t -L 90 -v 0.0007 0.0007 700000 0.004 .7 .7 -r e 1.667 -o exsoutput

X. clivii, migration

./mimar 4000 10000 19 -lf clivii_in.txt -u 4.1e-9 -t u 0.000005 0.03 -n u 0.000005 0.015 -ej u 0 50000000 -N u 0.0000001 0.1 -m 1 2 l -8 3 -m 2 1 l -8 3 -i 1000 -x 10 -y t -L 90 -v 0.0003 0.0003 300000 0.005 .2 .2 -r e 1.667 -o exsoutput

X. largeni, no migration
./mimar 4000 10000 19 -lf largeni_IN.txt -u 4.1e-9 -t u 0.000005 0.03 -n u 0.000005 0.015 -ej u 0 20000000 -N u 0.0000001 0.09 -i 1000 -x 10 -y t -L 90 -v 0.001 0.001 1000000 0.003 .7 .7 -r e 1.667 -o exsoutput

X. largeni, migration
./mimar 4000 10000 19 -lf largeni_IN.txt -u 4.1e-9 -t u 0.000005 0.03 -n u 0.000005 0.015 -ej u 0 6000000 -N u 0.0000001 0.06 -m 1 2 l -8 3 -m 2 1 l -8 3 -i 1000 -x 10 -y t -L 90 -v 0.001 0.001 1000000 0.005 .7 .7 -r e 1.667 -o exsoutput

Commands used for IMa2 analysis

X. clivii,
"M mode": ./IMa2 -b300000 -c1 -iIM_clivii.in –j1 -l100000 -m1 -oIM_clivii_M.out -p2357 -q21 -r2 -t12 -u1 -y1 -hfg -hn40 -ha0.975 -hb0.75
"L mode": ./IMa2 -b300000 -c12 -iIM_clivii.in -j1 -l300000 -m1 -oIM_clivii_M_L_final.out -p35 -q21 -r0 -t12 -u1 -y1 -wIM_nested_2_pops.txt -vti_files/IM_clivii_M_out
X. largeni, migration
"M mode": ./IMa2 -b300000 -c1 -iIM_largeni.in –j1 -l100000 -m1 -oIM_largeni_M.out -p2357 -q15 -r2 -t6 -u1 -y1 -hfg -hn40 -ha0.975 -hb0.75
"L mode": ./IMa2 -b300000 -c12 -iIM_largeni.in -j1 -l300000 -m1 -oIM_largeni_M_L_final.out -p35 -q15 -t6 -u1 -y1 -r0 -wIM_nested_2_pops.txt -v./ti_files/IM_larg_M_out

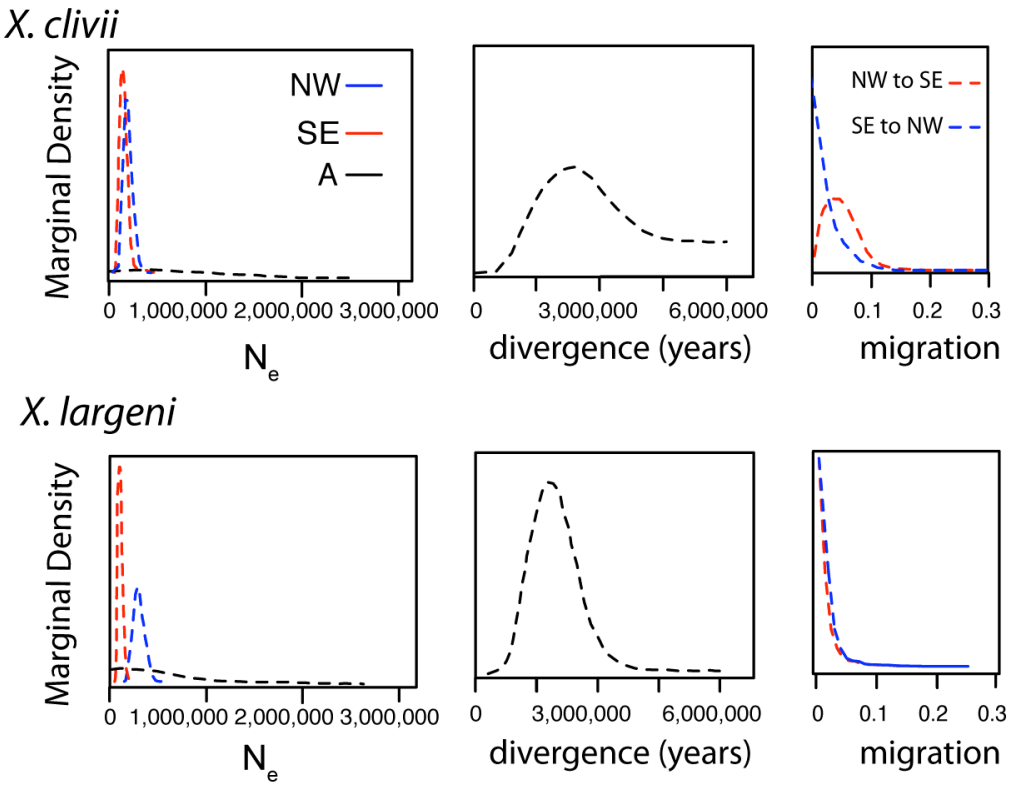
Supplementary Table 1. Information on specimens analyzed in this study including species and location relative to the Rift (Species (location)), field number, Specific location, latitude and longitude, and altitude in meters above sea level. Some data were not recorded (NR) and an asterisk following the field number denotes specimens for which mitochondrial DNA sequence data was not obtained.

Species (location)	Field number	Location	longitude	latitude	altitude
X. clivii southeast					
	BJE1476	Bore Bidka (5 km NW from Bore),Bore Bidka,Bore,Oromia	38 35.795 E	6 23.296 N	2667
	BJE1477	Bore Bidka (5 km NW from Bore),Bore Bidka,Bore,Oromia	38 35.795 E	6 23.296 N	2667
	BJE1478	Bore Bidka (5 km NW from Bore),Bore Bidka,Bore,Oromia	38 35.795 E	6 23.296 N	2667
	BJE1479	Bore Bidka (5 km NW from Bore),Bore Bidka,Bore,Oromia	38 35.795 E	6 23.296 N	2667
	BJE1500	Bore Bidka (5 km NW from Bore),Bore Bidka,Bore,Oromia	38 35.795 E	6 23.296 N	2667
	BJE1501	Bore Bidka (5 km NW from Bore),Bore Bidka,Bore,Oromia	38 35.795 E	6 23.296 N	2667
	BJE1502	Bore Bidka (5 km NW from Bore),Bore Bidka,Bore,Oromia	38 35.795 E	6 23.296 N	2667
	BJE1503	Gossa,Ana Anasura,Oromia	38 40.335 E	6 16.600 N	2651
	BJE1504*	Gossa,Ana Anasura,Oromia	38 40.335 E	6 16.600 N	2651
	BJE1510	Gossa,Ana Anasura,Oromia	38 40.335 E	6 16.600 N	2651
	BJE1511	Gossa,Ana Anasura,Oromia	38 40.335 E	6 16.600 N	2651
	BJE1512	Gossa,Ana Anasura,Oromia	38 40.335 E	6 16.600 N	2651
	BJE1513	Gossa,Ana Anasura,Oromia	38 40.335 E	6 16.600 N	2651
	BJE1514	Gossa,Ana Anasura,Oromia	38 40.335 E	6 16.600 N	2651
	BJE1515	Gossa,Ana Anasura,Oromia	38 40.335 E	6 16.600 N	2651
	BJE1516	Gossa,Ana Anasura,Oromia	38 40.335 E	6 16.600 N	2651
	BJE1522	Manyete, near Harena Forest of Bale Mountain National Park,Hauo,Manna,Bale,Oromia	39 44.989 E	6 28.913 N	1503
	BJE1523	Manyete, near Harena Forest of Bale Mountain National Park,Hauo,Manna,Bale,Oromia	39 44.989 E	6 28.913 N	1503
	BJE1524	Manyete, near Harena Forest of Bale Mountain National Park,Hauo,Manna,Bale,Oromia	39 44.989 E	6 28.913 N	1503
	BJE1525	Manyete, near Harena Forest of Bale Mountain National Park,Hauo,Manna,Bale,Oromia	39 44.989 E	6 28.913 N	1503
	BJE1526	Manyete, near Harena Forest of Bale Mountain National Park,Hauo,Manna,Bale,Oromia	39 44.989 E	6 28.913 N	1503
X. clivii northwest					
	AMNH1725	Addis Ababa	38 42.000 E	9 1.199 N	NR
	BJE1548	Eskinder Lake,01 Kebale,Chancha,Gamagoffa,,SNNPR	37 35.163 E	6 15.597 N	2652
	BJE1549	Eskinder Lake,01 Kebale,Chancha,Gamagoffa,,SNNPR	37 35.163 E	6 15.597 N	2652
	BJE1550	Eskinder Lake,01 Kebale,Chancha,Gamagoffa,,SNNPR	37 35.163 E	6 15.597 N	2652
	BJE1551	Eskinder Lake,01 Kebale,Chancha,Gamagoffa,,SNNPR	37 35.163 E	6 15.597 N	2652
	BJE1552	Tercha,Mareka,,Dawro,SNNPR	37 10.184 E	7 8.732 N	1349
	BJE1553	Tercha,Mareka,,Dawro,SNNPR	37 10.184 E	7 8.732 N	1349
	BJE1554	Tercha,Mareka,,Dawro,SNNPR	37 10.184 E	7 8.732 N	1349
	BJE1559	Gibdu Mida,Elu,Southwest Sewa,Oromo	38 23.165 E	8 50.681 N	2039
	BJE1560	Gibdu Mida,Elu,Southwest Sewa,Oromo	38 23.165 E	8 50.681 N	2039
	BJE1561*	Gibdu Mida,Elu,Southwest Sewa,Oromo	38 23.165 E	8 50.681 N	2039
	BJE1562	Gibdu Mida,Elu,Southwest Sewa,Oromo	38 23.165 E	8 50.681 N	2039
	BJE1566	Wole Chilalo (east of Blue Nile gorge),Wor Jarsso,North Shama,Oromia	38 17.835 E	9 57.857 N	2517
	BJE1567	Wole Chilalo (east of Blue Nile gorge),Wor Jarsso,North Shama,Oromia	38 17.835 E	9 57.857 N	2517
	BJE1568	Wole Chilalo (east of Blue Nile gorge),Wor Jarsso,North Shama,Oromia	38 17.835 E	9 57.857 N	2517
	BJE1569	Wole Chilalo (east of Blue Nile gorge),Wor Jarsso,North Shama,Oromia	38 17.835 E	9 57.857 N	2517
	BJE1570	Wole Chilalo (east of Blue Nile gorge),Wor Jarsso,North Shama,Oromia	38 17.835 E	9 57.857 N	2517
	BJE1571	Yench (west of Blue Nile gorge),Machakl,West Gojam,Amhara	37 34.251 E	10 28.222 N	2435
	BJE1572	Yench (west of Blue Nile gorge),Machakl,West Gojam,Amhara	37 34.251 E	10 28.222 N	2435
	BJE1573	Yench (west of Blue Nile gorge),Machakl,West Gojam,Amhara	37 34.251 E	10 28.222 N	2435
	BJE1574	Yench (west of Blue Nile gorge),Machakl,West Gojam,Amhara	37 34.251 E	10 28.222 N	2435
	BJE1575	Yench (west of Blue Nile gorge),Machakl,West Gojam,Amhara	37 34.251 E	10 28.222 N	2435
	BJE1584	Sebt Selassie,Achefer,East Gojam,Amhara	36 55.355 E	11 18.999 N	2065
	BJE1585	Sebt Selassie,Achefer,East Gojam,Amhara	36 55.355 E	11 18.999 N	2065
	BJE1586	Sebt Selassie,Achefer,East Gojam,Amhara	36 55.355 E	11 18.999 N	2065
	BJE1587	Sebt Selassie,Achefer,East Gojam,Amhara	36 55.355 E	11 18.999 N	2065

Supplementary Table 1. (continued)

Species (location)	Field number	Location	longitude	latitude	altitude
X. clivii northwest (continued)					
	Z23301	Korata,Dera,South Gondar,Amhara	37 31.122 E	11 44.224 N	1861
	Z23302	Korata,Dera,South Gondar,Amhara	37 31.122 E	11 44.224 N	1861
	Z23303	Korata,Dera,South Gondar,Amhara	37 31.122 E	11 44.224 N	1861
	Z23304	Korata,Dera,South Gondar,Amhara	37 31.122 E	11 44.224 N	1861
	Z23305	Korata,Dera,South Gondar,Amhara	37 31.122 E	11 44.224 N	1861
	Z23314	Emba Madrie, Mai Tsebri,Shine,Tigray	38 10.147 E	13 41.060 N	1448
	Z23315	Emba Madrie, Mai Tsebri,Shine,Tigray	38 10.147 E	13 41.060 N	1448
	Z23316	Emba Madrie, Mai Tsebri,Shine,Tigray	38 10.147 E	13 41.060 N	1448
	Z23317	Emba Madrie, Mai Tsebri,Shine,Tigray	38 10.147 E	13 41.060 N	1448
	Z23318	Emba Madrie, Mai Tsebri,Shine,Tigray	38 10.147 E	13 41.060 N	1448
	Z23319*	Axum 03 (near Aksum),Maichew,Tigray	38 43.314 E	14 8.062 N	2154
	Z23320	Axum 03 (near Aksum),Maichew,Tigray	38 43.314 E	14 8.062 N	2154
	Z23321	Axum 03 (near Aksum),Maichew,Tigray	38 43.314 E	14 8.062 N	2154
	Z23322*	Axum 03 (near Aksum),Maichew,Tigray	38 43.314 E	14 8.062 N	2154
X. largeni southeast					
	AMNH1729	Gossa,Ana Anasura,Oromia	38 40.335 E	6 16.600 N	2651
	BJE1505	Gossa,Ana Anasura,Oromia	38 40.335 E	6 16.600 N	2651
	BJE1506	Gossa,Ana Anasura,Oromia	38 40.335 E	6 16.600 N	2651
	BJE1507	Gossa,Ana Anasura,Oromia	38 40.335 E	6 16.600 N	2651
	BJE1508	Gossa,Ana Anasura,Oromia	38 40.335 E	6 16.600 N	2651
	BJE1509	Gossa,Ana Anasura,Oromia	38 40.335 E	6 16.600 N	2651
X. largeni northwest					
	BJE1579	Senebo,Dega Damot (Ferse Bet),West Gojam,Amhara	37 33.128 E	10 42.896 N	2467
	BJE1580	Senebo,Dega Damot (Ferse Bet),West Gojam,Amhara	37 33.128 E	10 42.896 N	2467
	BJE1581	Senebo,Dega Damot (Ferse Bet),West Gojam,Amhara	37 33.128 E	10 42.896 N	2467
	BJE1582	Senebo,Dega Damot (Ferse Bet),West Gojam,Amhara	37 33.128 E	10 42.896 N	2467
	BJE1583	Senebo,Dega Damot (Ferse Bet),West Gojam,Amhara	37 33.128 E	10 42.896 N	2467
	BJE1588	Geneb Geregra,Gonge Kolela,West Gojam,Amhara	37 36.428 E	11 9.338 N	2390
	BJE1589	Geneb Geregra,Gonge Kolela,West Gojam,Amhara	37 36.428 E	11 9.338 N	2390
	BJE1590	Geneb Geregra,Gonge Kolela,West Gojam,Amhara	37 36.428 E	11 9.338 N	2390
	BJE1591	Geneb Geregra,Gonge Kolela,West Gojam,Amhara	37 36.428 E	11 9.338 N	2390
	BJE1592	Geneb Geregra,Gonge Kolela,West Gojam,Amhara	37 36.428 E	11 9.338 N	2390
	BJE1593	Ayen Birhan,Hult Eju Nesie,West Gojam,Amhara	37 51.578 E	11 4.361 N	2377
	BJE1594	Ayen Birhan,Hult Eju Nesie,West Gojam,Amhara	37 51.578 E	11 4.361 N	2377
	BJE1595	Ayen Birhan,Hult Eju Nesie,West Gojam,Amhara	37 51.578 E	11 4.361 N	2377
	BJE1596	Ayen Birhan,Hult Eju Nesie,West Gojam,Amhara	37 51.578 E	11 4.361 N	2377
	BJE1597	Ayen Birhan,Hult Eju Nesie,West Gojam,Amhara	37 51.578 E	11 4.361 N	2377
	Z23306*	Awoskie,Wogera,North Gonder,Amhara	37 40.739 E	12 47.228 N	2709
	Z23307	Awoskie,Wogera,North Gonder,Amhara	37 40.739 E	12 47.228 N	2709
	Z23308	Awoskie,Wogera,North Gonder,Amhara	37 40.739 E	12 47.228 N	2709
	Z23309	Awoskie,Wogera,North Gonder,Amhara	37 40.739 E	12 47.228 N	2709

Supplementary Figure 1. Smoothed marginal posterior density distributions from IMa2 analysis. MCMC was performed with migration, which was favored by likelihood ratio tests for both species. Distributions for effective population size (N_e) are shown for the northwest (NW), southeast (SE), and ancestral (A) populations.



Supplementary Table 2: Morphological measurements of *X. clivii* individuals from northwest and southeast of the Rift Valley. All measurements are in millimeters and were taken on the right side of the animal when possible; descriptions of measurements are provided in Tinsley (1973).

Field Number	sex	*Snout-Vent Length*	*Body Width*	*Head Width*	*Snout Length*	*Eye Diameter*	*Interocular Distance*	*Nostril Diameter*	*Internarial Distance*	*Tibia Length*	*5th Toe Length*	*Elbow-1st Finger*	*1st Finger Length*
Northwest population													
BJE 1549	F	43.64	19.83	10.05	4.57	2.45	3.79	0.87	1.88	16.92	14.86	16.14	6.46
BJE 1550	F	53.21	25.42	11.79	6.44	2.91	4.81	1.26	1.98	20.78	20.85	20.41	8.24
BJE 1552	F	56.88	24.51	12.17	6.26	2.85	5.42	1.05	2.28	24.88	22.24	21.49	7.89
BJE 1553	F	59.00	23.15	12.80	6.13	2.93	5.33	0.95	2.42	23.84	23.66	22.66	8.72
Z23323	F	39.72	17.60	9.46	4.54	2.66	3.96	0.89	1.25	17.51	15.60	15.72	6.11
BJE 1554	M	45.00	21.48	10.62	5.55	2.32	4.50	0.94	1.70	21.27	19.73	18.10	5.91
Z23319	M	40.21	18.55	9.66	4.41	2.24	4.06	0.95	1.86	17.86	16.95	17.16	6.78
Z23320	M	39.63	16.72	9.47	4.32	2.42	3.65	0.98	1.82	17.67	16.26	17.22	5.34
Z23321	M	39.68	16.79	9.60	4.20	2.06	4.11	0.93	1.40	17.92	17.34	18.29	6.89
Z23322	M	38.87	15.93	9.60	4.46	2.45	4.07	0.96	1.51	17.57	16.88	17.25	7.24
Southeast population													
BJE 1503	F	59.59	30.33	12.68	6.35	2.94	5.42	1.10	2.10	22.89	22.78	21.80	7.68
BJE 1504	F	63.08	33.39	13.10	6.94	2.93	5.75	1.14	2.56	24.74	24.91	24.00	8.68
BJE 1510	F	39.88	19.10	9.11	4.21	2.02	4.16	0.81	1.81	17.62	16.78	14.81	5.73
BJE 1511	F	38.05	17.67	0.92	4.17	2.03	4.14	0.80	1.76	17.30	16.36	15.80	5.95
BJE 1524	F	52.06	27.96	12.62	6.68	2.94	5.33	1.17	2.06	21.68	20.89	20.09	8.32
BJE 1525	F	52.21	24.46	11.35	6.15	3.17	4.59	1.04	2.07	21.02	20.12	19.80	6.72
BJE 1526	F	52.30	24.79	11.36	5.66	2.97	4.76	1.06	1.84	20.56	18.21	18.00	6.92
BJE 1512	M	49.60	25.10	11.03	5.78	2.86	4.75	0.90	1.97	22.25	20.84	20.66	7.69
BJE 1513	M	49.08	24.18	10.88	5.67	2.80	4.40	1.08	1.97	23.26	23.14	21.02	6.58
BJE 1514	M	49.43	23.97	10.48	6.04	2.86	4.96	1.20	2.30	21.74	21.23	20.81	7.54
BJE 1515	M	47.82	22.84	10.64	6.09	3.06	4.23	0.90	2.04	21.05	20.38	19.19	6.13
BJE 1516	M	48.13	21.00	10.95	5.62	2.89	4.60	1.03	2.26	20.86	21.20	20.71	6.66
BJE 1522	M	48.15	20.28	11.34	5.80	2.71	4.56	1.08	1.81	22.18	20.86	20.01	7.35
BJE 1523	M	44.11	17.57	10.68	5.58	2.30	4.30	1.29	2.08	21.47	19.13	19.27	7.55
BJE 1551	M	42.15	20.58	10.13	5.29	2.56	3.90	1.15	1.94	18.99	18.45	19.14	7.01