Supplementary information for "World Heritage lizard: population genetics and species status of the range-restricted Hamelin skink, *Ctenotus zastictus*"

Sonal Singhal, Elizabeth Solis, Daniel L. Rabosky

Supplemental Figures

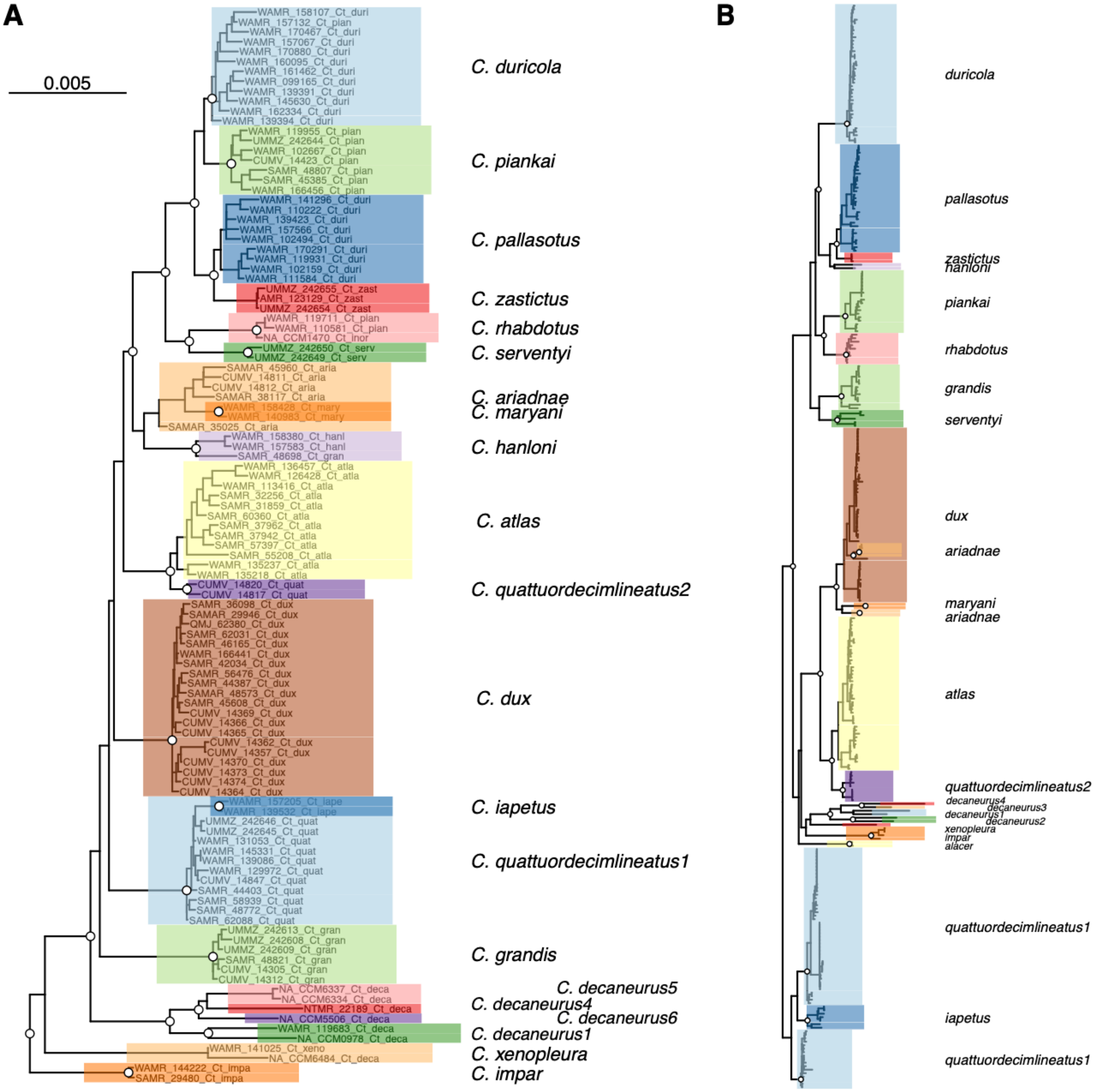

Figure S1: (A) Nuclear phylogeny of concatenated loci (*n* = 191 loci across 36.7K sites) for 113 individuals in the *Ctenotus atlas* group. Eleven individuals were dropped due to excessive missing data (>90% missing). Phylogeny is rooted by *L. desertorum* (not shown). Clade colors denote species groupings. While there is evidence of cryptic speciation (*C. quattuordecimlineatus*, *C. decaneurus*) and species paraphyly (*C. ariadne*), all other species groupings follow recognized boundaries. (B) Mitochondrial gene tree for *cytochrome b* for 317 individuals in the group. A number of recognized species are paraphyletic at the mitochondrial locus. In both trees, white circles at infraspecific nodes indicate >95 SH support. Several of the species ascribed to the *atlas* group based on morphology are quite distant from the rest of the group (C. *quattuordecimlineatus 1, C. iapetus, C. grandis, C. decaneurus, C. xenopleura,* and *C. impar*). Most notably, morphological data suggested that *C. iapetus* might be *C. zastictus*'s closest relative (Storr 1984); we find no evidence of this relationship.

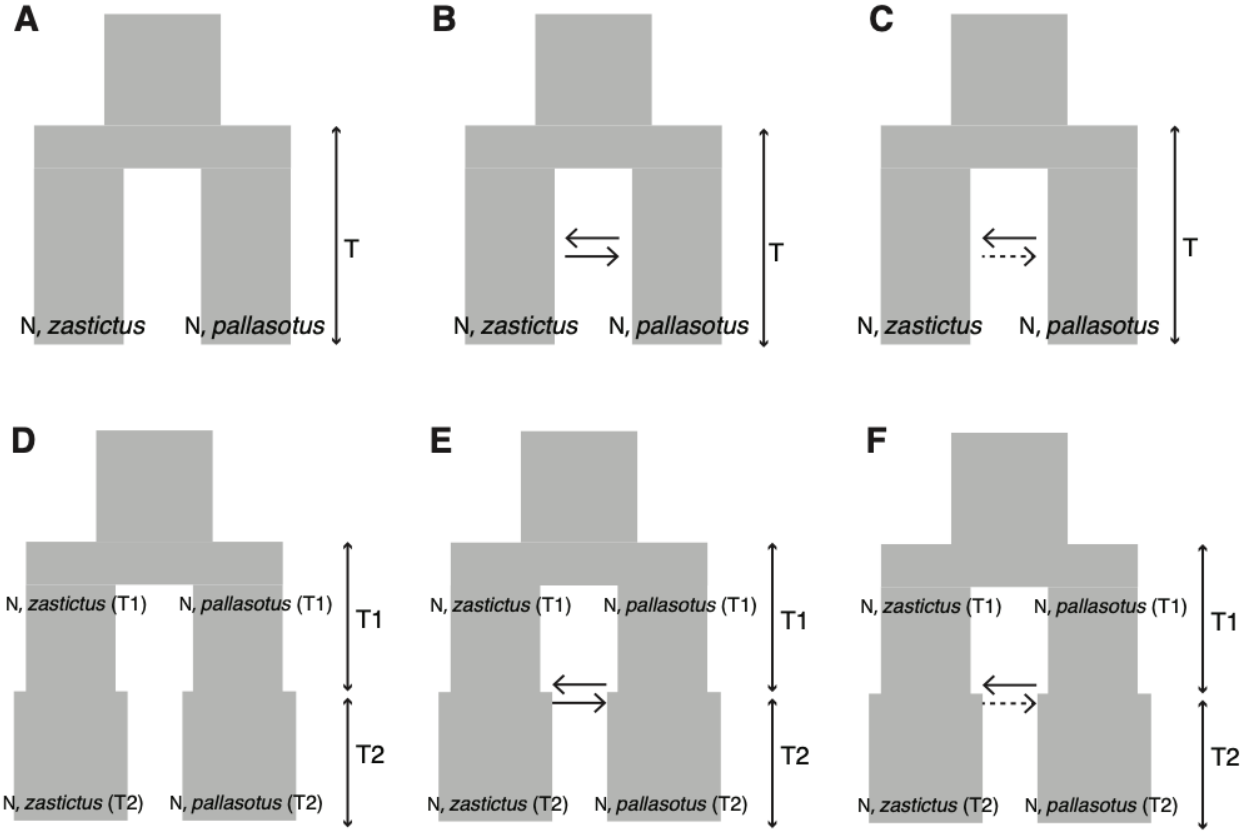

Figure S2: The six models tested in inferring the demographic history of *Ctenotus zastictus* and its sister species *C. pallasotus*: (A) split with no migration (three parameters: population size for each of two lineages and divergence time), (B) split with symmetric migration (four parameters: population size for each of two lineages, divergence time, and migration rate), (C) split with asymmetric migration (five parameters: population size for each of two lineages, divergence time, and migration rate for each of two lineages), (D) split with instantaneous population size change some time in the past (six parameters: population size for each of two lineages at divergence time, population size for each of two lineages at present, divergence time, population expansion time), (E) split with instantaneous population size change some time in the past and continuous symmetric migration throughout (seven parameters: population size for each of two lineages at divergence time, population size for each of two lineages at present, divergence time, population expansion time, and migration rate), and (F) split with instantaneous population size change some time in the past and continuous asymmetric migration throughout (eight parameters: population size for each of two lineages at divergence time, population size for each of two lineages at present, divergence time, population expansion time, and migration rate for each of two lineages).

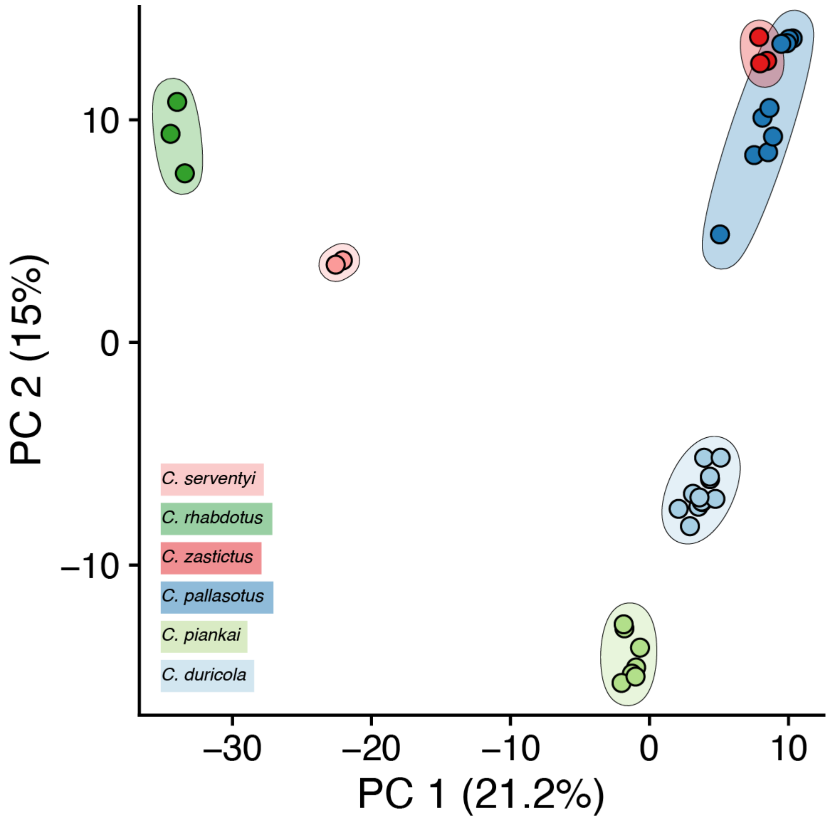

Figure S3: Principal component analysis (PCA) of unlinked SNP data for *C. zastictus* and close relatives; see Figure 3B for a PCA of *C. zastictus* and *C. pallasotus* only.

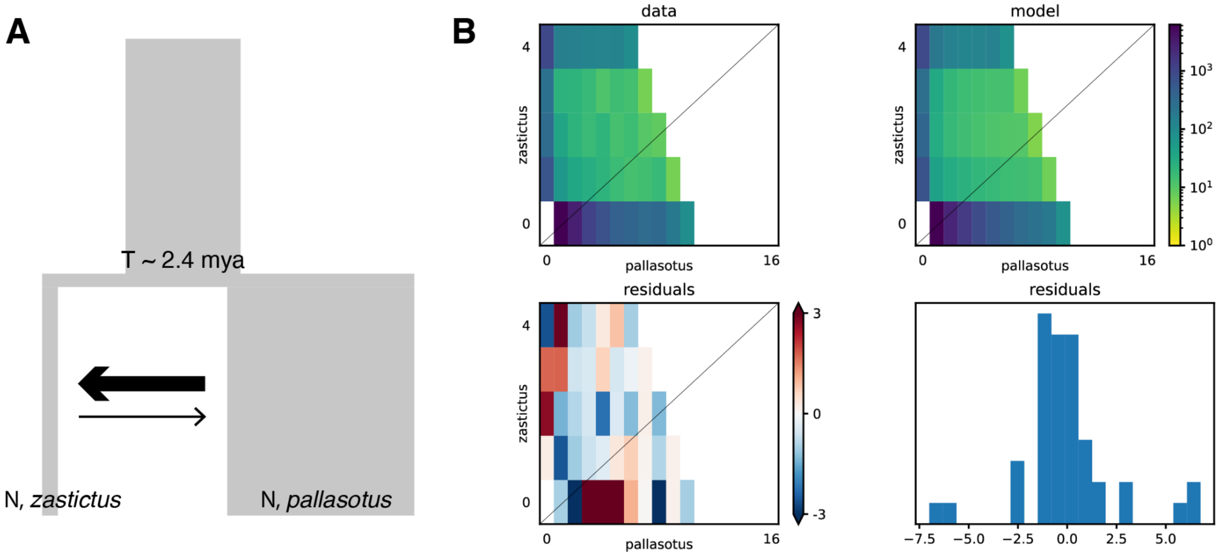

Figure S4: (A) Inferred demographic history for *Ctenotus zastictus* and *C. pallasotus* under the best-fitting model: a model of asymmetric migration during population divergence. Migration arrows and population sizes are scaled relative to estimated values (Table S5). (B) Actual two-dimensional joint site frequency spectrum (2D-JSFS) and the modeled 2D-JSFS and its residuals under the best-fitting demographic history shown in (A).

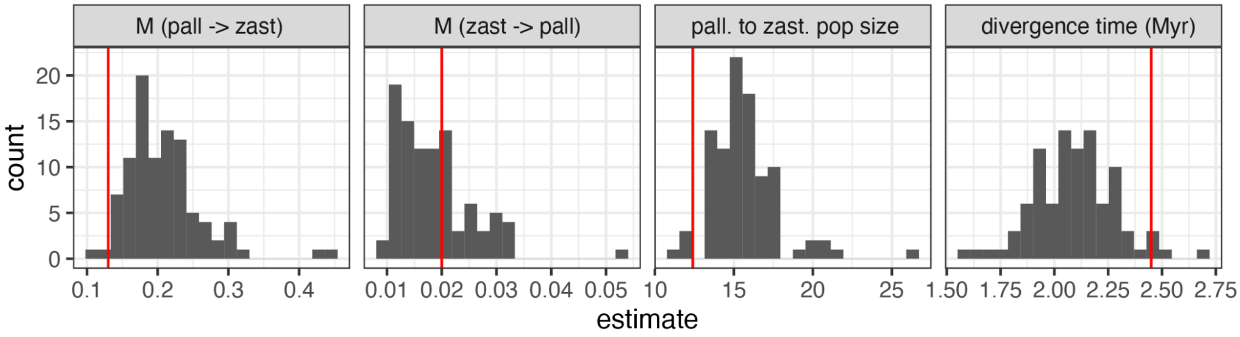

Figure S5: Uncertainty of dadi parameter estimates, including migration rates (reported in units of Migrants / generation for *C. pallasotus* to *C. zastictus* and *C. zastictus* to *C. pallasotus*), the ratio of population size for *C. pallasotus* to *C. zastictus*, and the divergence time between the two species (reported in millions of years). Uncertainty was inferred through 100 bootstrapped data sets. Estimates of these parameters from our full data set shown in red.

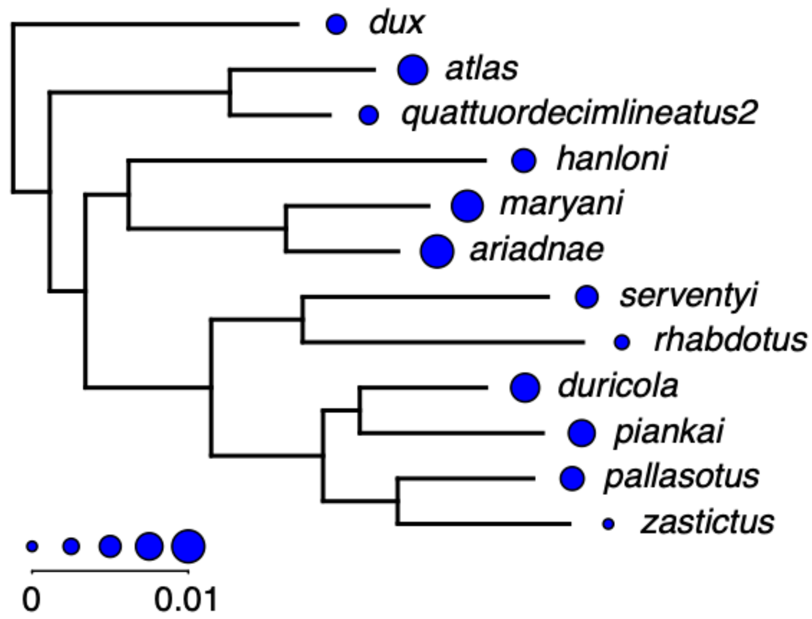

Figure S6: Nucleotide diversity (π) across the species in the *Ctenotus atlas* group. *C. zastictus* has very low levels of nucleotide diversity π = 0.00165) compared to the rest of the group (mean π = 0.0053; range: 0.00268 - 0.00735).

Supplemental Tables

Table S1: Locality information for the ten *Ctenotus zastictus* specimens found in museum collections. Museum abbreviations: Western Australian Museum (WAM), South Australian Museum (SAM), University of Michigan Museum of Zoology (UMMZ). Due to the sensitive nature of this taxon, latitude and latitude are shown at low resolution.

Museum	Catalog Number	Latitude	Longitude
WAM	84300	-26.5	114.2
WAM	81783	-26.6	114.2
WAM	96550	-26.5	114.2
WAM	92313	-26.5	114.2
WAM	82733	-26.6	114.2
WAM	82732	-26.6	114.2
WAM	81784	-26.6	114.2
SAM	123129	-26.5	114.2
UMMZ	242654	-26.5	114.2
UMMZ	242655	-26.5	114.2

Table S2: Information on *Ctenotus* individuals included in this study, including: their sample name, their museum voucher identification number, locality and latitude & longitude, their provisional mitochondrial and nuclear lineage identity, and accession numbers for both mtDNA and ddRAD data. Not all individuals were collected for both mtDNA and ddRAD data.

(The table is uploaded as a separate file in the Zenodo repository.)

Table S3: Cross-validation errors for population clustering inferred using ADMIXTURE. Two datasets are shown: all sampled individuals and three subsampled datasets of three random individuals per recognized species. We determine the most likely number of population clusters (*K*; bolded) by smaller cross-validation error.

K	full dataset	subsampled dataset 1	subsampled dataset 2	subsampled dataset 3
1	0.55625	1.26786	1.01675	1.05900
2	0.51691	1.11411	1.04027	1.07437
3	**0.51334**	1.05796	0.96642	0.98891
4	0.53145	**1.03261**	**0.86620**	**0.85700**
5	0.56105	1.16662	0.96206	0.95298

Table S4: Marginal likelihood estimates of alternate species delimitation models: the current taxonomy versus a taxonomy in which *Ctenotus zastictus* and its sister species *C. pallasotus* are lumped. Marginal likelihood estimates were calculated using *BFD (Leaché et al. 2014). Results are shown across three randomly subsampled datasets. Results are qualitatively and quantitatively similar across runs; a model in which these two species are lumped is a significantly worse fit to the data than the current taxonomy.

Grouping Scheme	Run	Number of species	Marginal likelihood estimate	Bayes factors
Current taxonomy	1	6	-5022.17	----
Lump *C. zastictus* and *C. pallasotus*	1	5	-5439.11	-833.88
Current taxonomy	2	6	-4640.21	----
Lump *C. zastictus* and *C. pallasotus*	2	5	-5010.57	-740.72
Current taxonomy	3	6	-4824.23	----
Lump *C. zastictus* and *C. pallasotus*	3	5	-5186.95	-725.44

Table S5: Results of demographic inference with dadi. Shown are the best fits (as measured by log-likelihood [LnL] estimates) for each model and the corresponding AIC and Akaike weight, theta (θ), and parameter estimates. Parameter estimates follow Fig. S2 and are given as reported by dadi. Uncertainty in estimates shown in Fig. S5.

Model	LnL	AIC	Akaike Weight	θ	N, z (T1)	N, p (T1)	N, z (T2)	N, p (T2)	mig., p -> z	mig., z -> p	T1	T2
split with asymmetric migration	-260.2	530.4	1	1425.1	0.43	5.33	---	---	0.13	0.02	2.02	---
split with instantaneous population size change & symmetric migration throughout	-289.28	592.56	<1e-10	2071	1.36	5.51	0.23	3.71	0.04	---	0.6	0.37
split with instantaneous population size change & asymmetric migration throughout	-293.41	602.82	<1e-10	1260.2	1.1	0.69	0.41	6.35	0.2	0.02	0.36	2.35
split with instantaneous population size change	-317.69	647.38	<1e-10	1999.1	1.33	15.2	0.25	3.6	---	---	0.54	0.41
split with symmetric migration	-344.06	696.12	<1e-10	1854.1	0.41	4.5	---	---	0.04	---	1.2	---
split with no migration	-391.94	789.88	<1e-10	2250.9	0.34	4.02	---	---	---	---	0.77	---