Reproductive isolation between phylogeographic lineages scales with divergence: Supplementary Information

Sonal Singhal and Craig Moritz
singhal@berkeley.edu
craig.moritz@anu.edu.au

Museum of Vertebrate Zoology
University of California, Berkeley
3101 Valley Life Sciences Building
Berkeley, California 94720-3160

Department of Integrative Biology
University of California, Berkeley
1005 Valley Life Sciences Building
Berkeley, California 94720-3140

Research School of Biology
The Australian National University
Building 116
Acton, ACT 0200

1   Phylogeny

To infer the phylogeny shown in Figure 1 (main text), we concatenated and aligned sequences from previously published loci for the lineages in this group – the mitochondrial locus *ND4* and the nuclear loci β-globin intron and C-mos exon (1; 2; 3; 4) using MUSCLE (5). We used RAxML to infer a maximum-likelihood tree for the concatenated alignment (6). The approximate root age for the tree was estimated based on data from (7). The final alignment is published in the DataDryad package XXX.

References

[1] Bell R, *et al.* (2010) Patterns of persistence and isolation indicate resilience to climate change in montane rainforest lizards. *Mol Ecol* 19:2531–2544.

[2] Moussalli A, Moritz C, Williams S, Carnaval A (2009) Variable responses of skinks to a common history of rainforest fluctuation: Concordance between phyleogeography & paleo-distribution models. *Molecular Ecology* 18:483–499.

[3] Singhal S, Moritz C (2012) Testing hypotheses of genealogical discordance in a rainforest lizard. *Mol Ecol* 21:5059–5072.

[4] Dolman G, Moritz C (2006) A multilocus perspective on refugial isolation and divergence in rainforest skinks (*Carlia*). *Evol* 60:573–582.

[5] Edgar R (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl Acids Res* 32:1792–97.

[6] Stamatakis A (2006) RAxML-VI-HPC: Maximum Likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.

[7] Skinner A, Hugall A, Hutchinson A (2011) Lygosomine phylogeny and the origins of Australian scincid lizards. *Journal of Biogeography* 38:1044–1058.
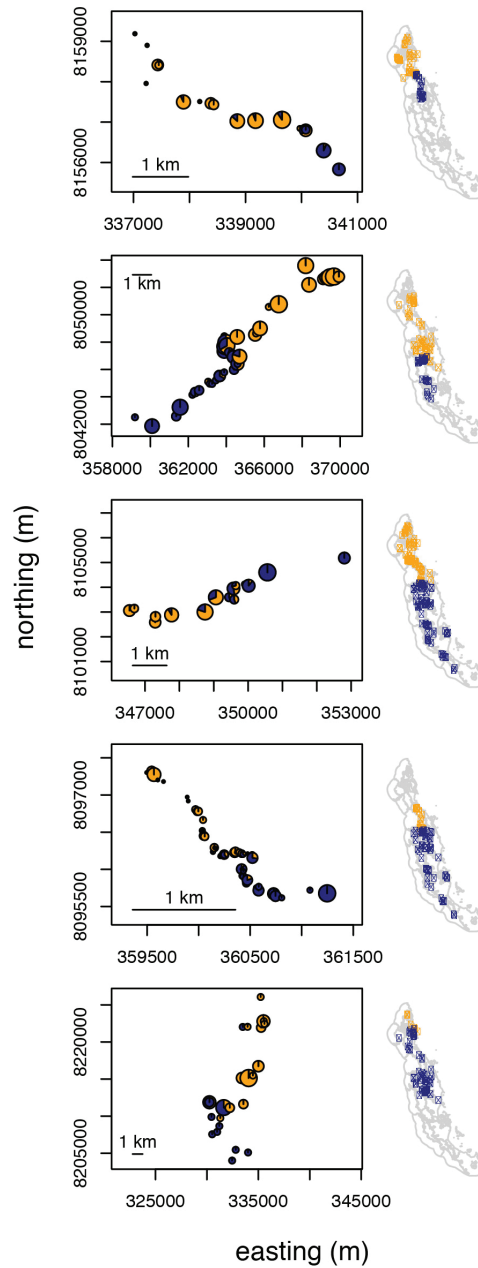
2 Figures & Tables

Figure 1: **Detailed transect maps of study system.** On left, transect for each contact zone, showing mitochondrial composition of unique localities with localities scaled according to sample size; on right, map of the Australian Wet Tropics showing the range of the phylogeographic lineages. From top to bottom, *Lampropholis coggeri* N/C, *Saproscincus basiliscus* N/C, *Carlia rubrigularis* N/S, *L. coggeri* C/S, and *S. lewisi/S. basiliscus* N.
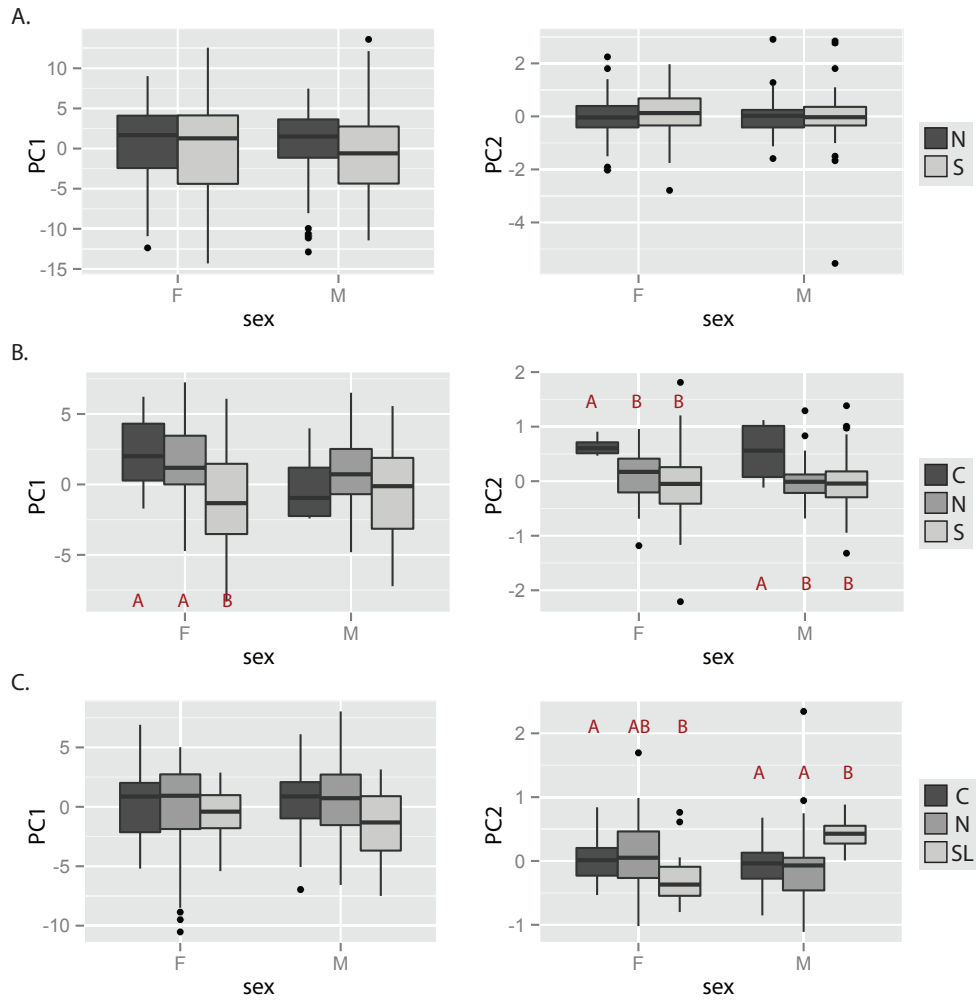
Figure 2: **Morphological variation among lineages.** Morpological data summarized across sexes and across phylogeographic lineages within the four morphologically defined species in this study: A. *Carlia rubrigularis* ($N_{\female} = 223$, $N_{\male} = 156$), B. *Lampropholis coggeri* ($N_{\female} = 174$, $N_{\male} = 143$), and C. *Saproscincus basiliscus* and *S. lewisi* ($N_{\female} = 119$, $N_{\male} = 119$). For each species, we present the first two axes of variation, as summarized by a principal components analysis. Significant differences are labeled in red.
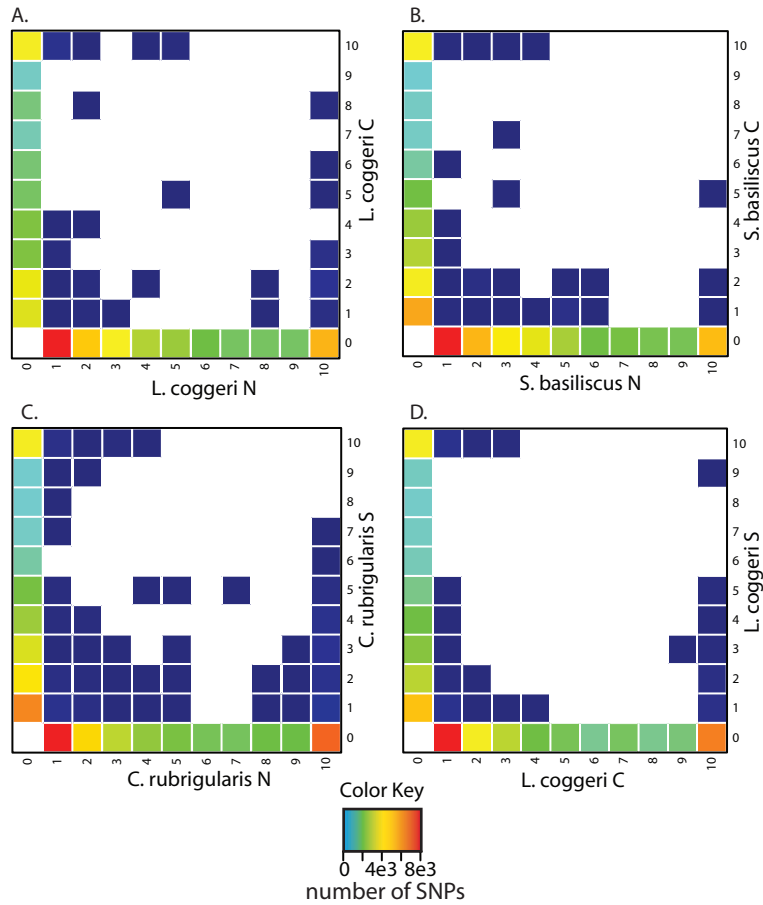
Figure 3: **2D-SFS used for inferring population history.** Two-dimensional site-frequency spectra (2D-SFS), as inferred by ANGSD, for A. *Lampropholis coggeri* N/C, B. *Saproscincus basiliscus* N/C, C. *Carlia rubrigularis* N/S, and D. *L. coggeri* C/S. For each lineage-pair, we used a total of ten individuals, or twenty chromosomes, evenly split between the two lineages. Details on single nucleotide polymorphisms used to construct the 2D-SFS can be found in Table S2.
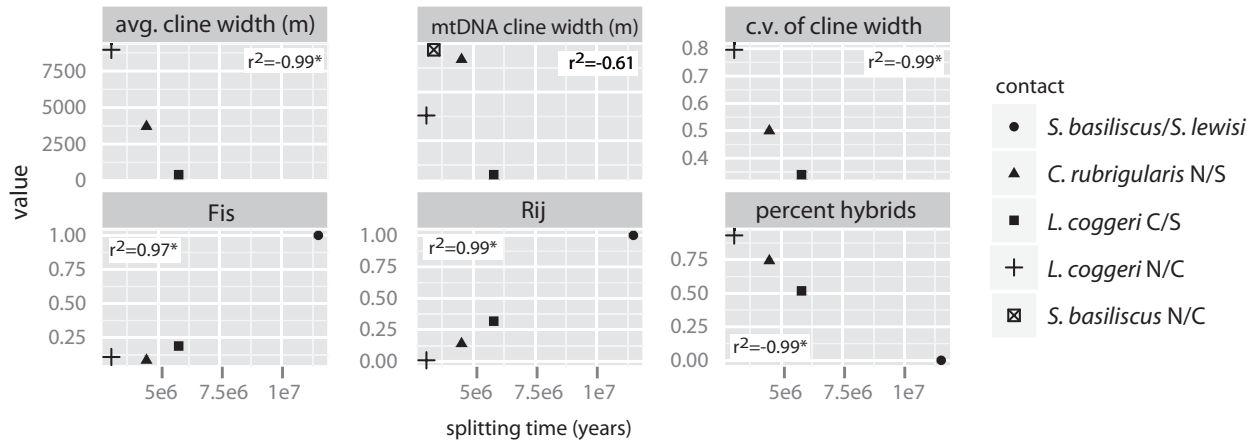
Figure 4: **Correlation of divergence history and reproductive isolation.** Comparative results showing the correlation between divergence history and indices of reproductive isolation: average nuclear cline width, mitochondrial cline width, coefficient of variance in cline width, Hardy-Weinberg disequilibrium (F_{IS}), linkage disequilibrium (R_{ij}), and percent of hybrids in the contact zone. Data for nuclear cline widths from *S. basiliscus* N/C are not represented due to limited sampling. Graphs are labeled with correlation coefficients; asterisks indicate significant correlations.

Figure 5: **Correlation of nuclear divergence and reproductive isolation.** Comparative results showing the correlation between nuclear divergence and indices of reproductive isolation: average nuclear cline width, mitochondrial cline width, coefficient of variance in cline width, Hardy-Weinberg disequilibrium (F_{IS}), linkage disequilibrium (R_{ij}), and percent of hybrids in the contact zone. Graphs are labeled with correlation coefficients; asterisks indicate significant correlations.
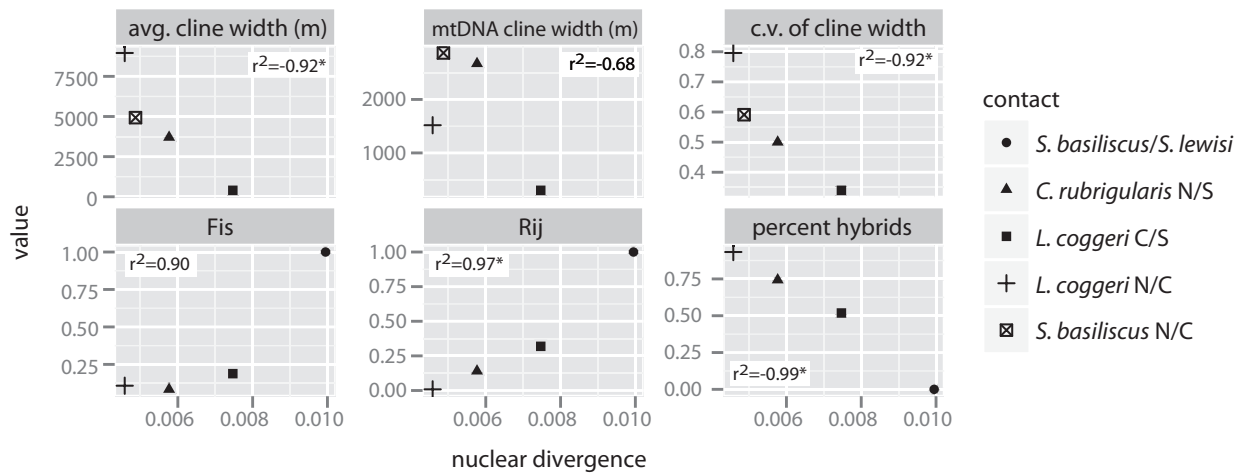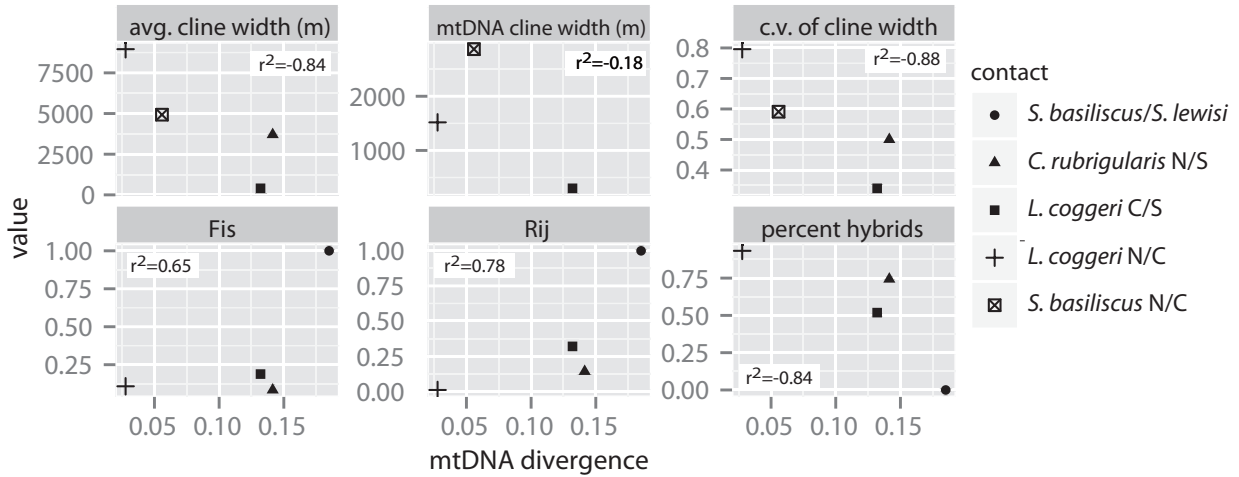
Figure 6: **Correlation of mitochondrial divergence and reproductive isolation.** Comparative results showing the correlation between mitochondrial divergence and indices of reproductive isolation: average nuclear cline width, mitochondrial cline width, coefficient of variance in cline width, Hardy-Weinberg disequilibrium (F_{IS}), linkage disequilibrium (R_{ij}), and percent of hybrids in the contact zone. Graphs are labeled with correlation coefficients; asterisks indicate significant correlations.
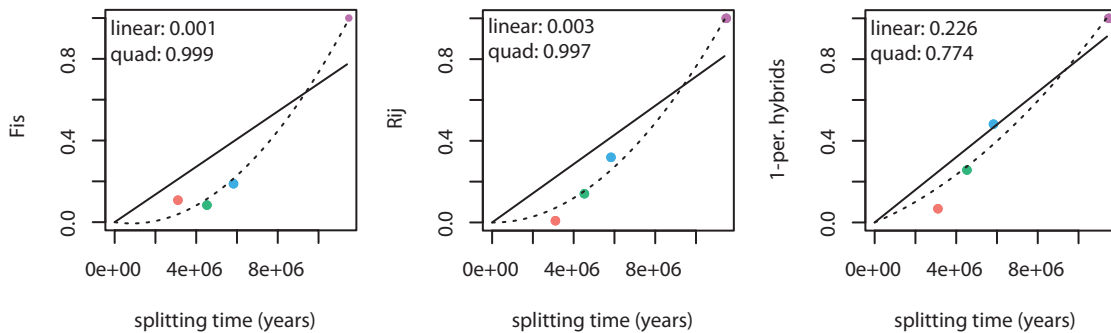
Figure 7: **Model fitting for three indices of reproductive isolation.** We fit linear (solid) and quadratic models (dotted) to the increase of reproductive isolation through time. Relative weights for the different models (as calculated via AIC scores) are shown for each model for each index.

7

contact zone	number of samples	number of transect populations	transect length
L. coggeri N/C	202	11	16 km
S. basiliscus N/C	209	10	12 km
C. rubrigularis N/S	308	10	7 km
L. coggeri C/S	406	17	2 km
S. basiliscus N & *S. lewisi*	55	NA	15 km

Table 1: **Sampling details for each contact zone.** Transect populations are those used in estimation of clines.

Table 2: **The loci used in this study and their associated details.**

LocusID	Forward Primer	Reverse Primer	PCR Temp.	Location/Type	R. Enz.	Cutting Pattern	Reference	contact
GLB1L2	GGGTTGGAGGCTCCGTCCT	CCGTCACAGCTTCACTAAGGTCCGT	65	3'UTR, non-coding	XbaI	410 (190+220)	this paper	C. rubrigularis N/S
GRN	TGCCCTGTCCCATGAGGGCT	TCGGGAGCTGAACCTCCACCC	65	3'UTR, non-coding	BclI	300 (50+250)	this paper	C. rubrigularis N/S
IDE	TGGGACAATCCCAAAGCAGACACT	TGGACACCTGGGTTCTCTGTGC	65	3'UTR, non-coding	MspI	270 (100+170)	this paper	C. rubrigularis N/S
IRS4	CCAGCAAGGCACCGACAGG	CAGGGCATCGGCTTCCAGGC	65	3'UTR, non-coding	KpnI	300 (250+50)	this paper	C. rubrigularis N/S
KIAA2013	TCAGCGCCAGCACTTGCCTC	AGAGGCCACCGGAGGAGCAG	65	3'UTR, non-coding	XbaI	300 (200+100)	this paper	C. rubrigularis N/S
LMBR1	TGCTGACTTGCTCACCACTGCC	GCCTAGCCCAGGAACAAGAGCG	65	3'UTR, non-coding	Eco53KI	370 (50+320)	this paper	C. rubrigularis N/S
MST4	TAGGCGGTTGGACAGCACAC	CAAAAGCTGCCCCTCCCCG	65	3'UTR, non-coding	BstUI	260 (90+170)	this paper	C. rubrigularis N/S
ND4	CACCTATGACTACCAAAAGCTCATGTAGAAGC	CATTACTTTTACTTGGATTTGCACCA	50	CDS, syn.	HhaI	900 (500+400)	Arevalo et al, 1994	C. rubrigularis N/S
NT5C2	CGGTTCTGCTGCAGGCCCAA	ACAAATGCGCACATGCCAAGG	65	3'UTR, non-coding	SspI	290 (50+240)	this paper	C. rubrigularis N/S
SF1	TCAGGACCCTAGCGCCGAGT	CACACCGGCCCCACACAGAC	65	3'UTR, non-coding	XmnI	290 (50+240)	this paper	C. rubrigularis N/S
SF3A1	CCTGGGAAGCAACAGCCGGG	CTGCGCTGGCAAGACACCA	65	3'UTR, non-coding	AlwNI	200 (40+60+100)	this paper	C. rubrigularis N/S
ABHD5	ACCCCACTTGTTCTTCTCCA	TGAGTAAGCAGCTGCCAAAA	60	CDS, syn.	BstBI	230 (160+70)	Singhal and Moritz 2012	L. coggeri C/S
AUTO	TGAGCAGGAAAGGCAAATCT	GTGCCAGTGTCTTGATG	62	CDS, syn.	BanI	190(170+20)	Singhal and Moritz 2012	L. coggeri C/S
BGLO (intron 2)	GCGAACTGCCACTGYGACAAG	GCTGCAAGCGGGTTGA	63	intron	ApeKI	670 (640+30)	Dolman and Phillips 2004	L. coggeri C/S
LEMD2	GTGCATTCAAGCAGACCAGA	GGCTAGCACTCTCCACCAAG	60	3'UTR, non-coding	HindIII	240 (140+100)	Singhal and Moritz 2012	L. coggeri C/S
ND4	CACCTATGACTACCAAAAGCTCATGTAGAAGC	CATTACTTTTACTTGGATTTGCACCA	50	CDS, syn.	HhaI	900 (530+370)	Arevalo et al, 1994	L. coggeri C/S
ND5T2	TCTTGGGGTTGTTTCCAGAC	CACTTGGCATTGTGAGCAGT	60	3'UTR, non-coding	NcoI	440 (230+210)	Singhal and Moritz 2012	L. coggeri C/S
PCBD1	TCTCTCTTGGCTGTGTGGAA	TAAATCATGTGCCCCCAAAT	60	3'UTR, non-coding	HinfI	250 (140+110)	Singhal and Moritz 2012	L. coggeri C/S
RPS8 (intron 3)	CTCTTGGGCGTAAGAAAGGAG	CCGCTCATCGTATTTCTTCTG	53	intron	HhaI	630 (400+230)	Bell et al., 2010	L. coggeri C/S
RTN3	AACCTGTTCCAAGGCAATTC	TTGAGAAAGCGGAGTTGTGG	60	3'UTR, non-coding	Eco53KI	440 (130+310)	Singhal and Moritz 2012	L. coggeri C/S
SAR1	TAATCACTTTGGCCCACCTC	TATCGCACAAATGCAAGAGC	56	3'UTR, non-coding	AccI	420 (350+70)	Singhal and Moritz 2012	L. coggeri C/S
TPI (intron 5)	TTCTAGCCTATGAACCAGTTTGG	CCTCAACTTGTCATGAACTTCC	50	intron	–	–	Bell et al, 2010	L. coggeri C/S
AKT2	ATTCTTCCCCACCCCCGGG	CCGCACCCGGCAAACAGAA	53	3'UTR, non-coding	Hpy166II	380 (120+180+80)	this paper	L. coggeri N/C
ELOVL2	GAGGCAGGGTGTGCAGCAGG	TGCAGCCTAAGAGCAAAAGCCAGT	65	3'UTR, non-coding	BstNI	240 (60+180)	this paper	L. coggeri N/C
MAT1A	TGGTCAGGCTGCCAGATCCCA	GCCTCTCCATTCTCTTGGCCCC	65	3'UTR, non-coding	BglI	280 (200+80)	this paper	L. coggeri N/C
PCYT1A	GCAGCCCCATGGCGGGTATC	CTGGCCCCATGGCCCAAC	65	3'UTR, non-coding	RsaI	230 (50+180)	this paper	L. coggeri N/C
PEX16	AGCCTTGGTTTGGTCATTTCAGCCA	GGACCCACCACCACGTCTGGA	65	3'UTR, non-coding	SphI	300 (270+30)	this paper	L. coggeri N/C
PNPLA2	TGCTGGAGCTGGACCTAGCGA	AGGAAGGGAGGCCCAGTTCACA	65	3'UTR, non-coding	KpnI	250 (200+50)	this paper	L. coggeri N/C
PPP2R1A	GGGTGGAGGGATCCGGTCGA	GGGTTTGCGCACCTCCCAGG	65	3'UTR, non-coding	NcoI	380 (300+80)	this paper	L. coggeri N/C
SDCBP	TCATGTGGGCATCCAGCTCT	ACTGTTTGTGTAAAATGCAGGCAA	65	3'UTR, non-coding	BstNI	220 (170+50)	this paper	L. coggeri N/C
SLC35F5	TGAGGCTCCTGCCTGAATGGAA	AGCGATGCAACAGTGCCAGC	65	3'UTR, non-coding	RsaI	300 (200+100)	this paper	L. coggeri N/C
SLC37A4	TCCCCTGCCCAGCACTGTGG	TGGAAGGGGTGAGTGGGTCC	65	3'UTR, non-coding	DraI	400 (250+150)	this paper	L. coggeri N/C
ND4	CACCTATGACTACCAAAAGCTCATGTAGAAGC	CATTACTTTTACTTGGATTTGCACCA	57 & 50	CDS, syn.	HhaI	900 (300+600)	Arevalo et al, 1994	L. coggeri N/C
ACY1	CCCAAGGAGGCAAAAGGGCC	TCACCCCCATGCCTTCCCCT	65	3'UTR, non-coding	MslI	190 (110+80)	this paper	S. basiliscus N/C
DPP4	TGCAGTTTCTTTGTCCCATGTTGGGA	GCACACACCCACTGGCACA	65	3'UTR, non-coding	HhaI	280 (60+220)	this paper	S. basiliscus N/C
ECE2	ACTTCTGGGCGGCGTGTCGT	GTGGGAGCTCTGGGGTGGGT	65	3'UTR, non-coding	Hpy166II	180 (45+135)	this paper	S. basiliscus N/C
GOLGA2	AGTGGGCCTATGTGCTGAGCA	AGCACCGTAAGGGGTGTGCA	65	3'UTR, non-coding	MspI	250 (200+50)	this paper	S. basiliscus N/C
LCAT	ACCCTGCCGAGCAAACTGGAGA	AGTTGCGCACAAGCCCTGCA	65	3'UTR, non-coding	BclI	180 (80+100)	this paper	S. basiliscus N/C
PNPLA2	AGAGTGCAACTCCCACACC	GGTCCCTTCAAGCACAGCAGCA	65	3'UTR, non-coding	BstUI	230 (60+170)	this paper	S. basiliscus N/C
SLC9A7	TGATGAACAGAGGCACTCGGCT	AGTCCACCTGCCTTGTCGA	65	3'UTR, non-coding	BglII	250 (200+50)	this paper	S. basiliscus N/C
TOMM70A	TGACTCATGTTTTGAGGGTTGTTGTGA	TGCAGCAGAACAGCAGGCTG	65	3'UTR, non-coding	BsmAI	280 (90+190)	this paper	S. basiliscus N/C
TXNRD3	TGGGGAAAACCTATTCGTCCAGTCA	AGGCCACTGCTTTCTTAGCAGC	65	3'UTR, non-coding	BsmAI	330 (130+200)	this paper	S. basiliscus N/C
UGT1A1	GCAAGTGCTCTGCCAGCATGC	AGCACCGATTCAGCGCAGCT	65	3'UTR, non-coding	DraIII	220 (50+170)	this paper	S. basiliscus N/C
ND4	CACCTATGACTACCAAAAGCTCATGTAGAAGC	CATTACTTTTACTTGGATTTGCACCA	57 & 50	CDS, syn.	RsaI	900 (200+700)	Arevalo et al, 1994	S. basiliscus N/C
B-globin	GCGAACTGCCACTGYGACAAG	GCTGCCAAGCGGGTGGTGA	61	intron	sequenced		Dolman and Phillips 2004	S. basiliscus N/S. lewisi
cmos	GCGGTAAAGCACGGTGAAGAAA	TGAGCATCCAAAGTCTCCAAT	57	CDS, syn	sequenced		Saint 1998	S. basiliscus N/S. lewisi
TPI (intron 5)	TTCTAGCCTATGAACCAGTTTGG	CCTCAACTTGTCATGAACTTCC	57	intron	sequenced		Bell et al., 2010	S. basiliscus N/S. lewisi
CRISP	TGCTGTAGCCTACTGTCCTCAA	TGCTTATCATGCTCGCTAAGTT	57	intron	sequenced		this paper	S. basiliscus N/S. lewisi
RPS8 (intron 3)	CTCTTGGGCGTAAGAAAGGAG	CCGCTCATCGTATTTCTTCTG	57	intron	sequenced		Bell et al., 2010	S. basiliscus N/S. lewisi
rho	CCTTGCCTGCACACCCTATGCTG	CAGGAGAGACCCTCACATTG	61	intron	sequenced		Dolman and Phillips 2004	S. basiliscus N/S. lewisi
ND4	CACCTATGACTACCAAAAGCTCATGTAGAAGC	CATTACTTTTACTTGGATTTGCACCA	57	CDS, syn.	sequenced		Arevalo et al 1994	S. basiliscus N/S. lewisi

contact zone	total sequence length (Mb)	total number of SNPs	fixed SNPs	polymorphic SNPs	shared SNPs
L. coggeri N/C	1.09	19884	3510 (17.7%)	16220 (81.6%)	154 (0.8%)
S. basiliscus N/C	1.52	29664	4712 (15.9%)	24798 (83.6%)	206 (0.7%)
C. rubrigularis N/S	1.35	32264	6365 (19.7%)	25693 (79.6%)	369 (1.1%)
L. coggeri C/S	1.42	41618	9260 (22.2%)	31989 (76.9%)	330 (0.8%)

Table 3: **SNP summary.** Details on the number of single nucleotide polymorphisms (SNPs), and their proportions, used in the two-dimensional site frequency spectrum (2D-SFS) for the contact zones analyzed with genomic data.

contact zone	nuc. div.	mt. div.	theta (θ)	div. time	M_{12}	M_{21}	N_1	N_2	N_A
L. coggeri N/C	0.0046	0.028	3090	3.1 my	0.0268	0.0117	408881	574006	1352453
S. basiliscus N/C	0.0049	0.056	3644	3.4 my	0.0123	0.0112	239822	919316	1352327
C. rubrigularis N/S	0.0058	0.141	3775	4.5 my	0.0112	0.0359	464585	1200178	1362782
L. coggeri C/S	0.0075	0.132	4608	5.8 my	0.0097	0.0204	628695	1176557	2227376
S. lewisi/S. basiliscus N	0.0100	0.185	NA	11.4 my	0.0186	0.0040	278501	740017	NA

Table 4: **Parameter estimates for the isolation-with-migration model**, as fit to the lineage-pairs. Populations labelled '1' are the northern lineage in each contact; populations labelled '2' the southern lineage.