

Supplementary Information

1 Design of Exome Capture Arrays

For each lineage, we first sequenced five transcriptomes from a single population far from the contact zone (Fig. S1; Singhal (2013)). We annotated these transcriptomes and their intron-exon boundaries using the programs blast and exonerate and the reference genome from *Anolis carolinensis* (AnoCar2.0+) (Altschul *et al.* 1997; Slater and Birney 2005). We identified exons for which (1) we could identify orthologs in both lineages of the lineage-pair, (2) GC-content was between 30% and 80% (Bi *et al.* 2012), (3) the transcriptome data suggested the exons contained either segregating or fixed variation (Singhal 2013), and (4) exon length was >200 bp. To this, we added (1) full-length transcripts for genes ($N = 95 - 99$) that showed evidence of positive and divergent selection, as identified by calculating d_{xy} , $\frac{dN}{dS}$, and F_{ST} (Yang 2007), (2) full-length transcripts for genes ($N = 92$) with metabolic or reproductive function as defined by Gene Ontology categories, because these genes possibly contribute to cryptic phenotypes with relevance for species boundaries, (3) two mitochondrial loci, *NADH dehydrogenase subunit 4* and *16s ribosomal RNA*, and (4) the 5' and 3' untranslated regions (UTRs) for regions previously genotyped in these populations ($N = 10$; Singhal and Moritz (2013)). Both the mitochondrial and the UTR targets were used to evaluate the efficacy of our pooled strategy. Targets were then filtered to remove repetitive sequence, which included targets that (1) matched to repeats in the RepeatMasker database (Smit *et al.* 2013), (2) were highly similar to other targets on the same array, and (3) contained k -mers that were disproportionately common in the *A. carolinensis* genome. To minimize bias in capture efficiency between lineages, we printed orthologs from each lineage for each targeted exon.

In total, we targeted an average of 3082 exons and 1.83 Mb of sequence, and the four capture arrays had 1120 exons in common. As expected, our targeted exons are more divergent than all other exons in the transcriptome (Fig. S8), but general patterns of divergence are similar. The targets were then printed at 2bp tiling on Agilent 1M eArrays. Scripts for array design are available at https://github.com/singhal/introgression_AWT/tree/master/probeDesign.

2 Simulations of Anonymous Pooling

To determine the pooling strategy – both with respect to number of individuals to include per pool and desired coverage – we conducted simulations in R (R Development Core Team 2011). These simulations were designed to determine how sampling drift would affect our inference of allele frequency. Anonymous pooling leads to two primary sources of bias. First, we are sampling a subset of individuals from a much larger population. All estimates of allele frequency are subject to this bias, even when allele frequency is calculated via genotyping individuals. Second, because we are sequencing anonymous pools, we will sequence unequal number of reads per individual,

36 even if individuals are combined at exactly the same concentrations. Inaccuracies in estimating
37 DNA concentrations will only exacerbate this bias. Here, for "known" allele frequencies (0.01,
38 0.1, 0.25, 0.5), we used the binomial sampling distribution to simulate the effects of both sources
39 of sampling drift. Importantly, we assumed that the DNA had been pooled in exactly equimolar
40 amounts, such that each chromosome in the population was equally likely to be sampled. As
41 summarized in Fig. S2, we had two major findings:

- 42 1. increasing the number of individuals in a pool improved estimates of allele frequency much
43 more than increasing sequencing effort (*i.e.*, increasing coverage) for a constant number of
44 individuals
- 45 2. increasing sequencing effort such that average coverage for a pool was $>50\times$ had only neg-
46 ligible effects on our estimates of allele frequency

47 Based on these results, we sequenced our libraries to a projected depth of $50\times$.

48 We then extended these simulations to determine how anonymous pooling would affect es-
49 timates of cline center and cline width. For each contact zone, we summarized key aspects of
50 sampling design: the number of demes sampled, the number of individuals sampled per deme,
51 the position of these demes along the transect, and the average coverage of sequence data. We ad-
52 ditionally calculated the median, mean, and 25% & 75% cline width for each contact zone. Across
53 these four cline widths and using the empirical sampling design, we first calculated the expected
54 allele frequencies at each deme using a simple sigmoidal equation for clines. We then accounted
55 for sampling drift due to subsampling the true population and due to sequencing anonymously
56 pooled reads. We repeated this 500 times for each cline width. We then followed the same ap-
57 proach used for our empirical data (see *Fitting Clines*) to infer cline centers and widths for each
58 simulated data set. These simulations assume that clines best fit a sigmoidal model, and, as above,
59 they assume that DNA has been pooled in equimolar amounts across individuals.

60 3 Evaluating Success of Exome Capture

61 Applications of exome capture to non-model organisms are still in their infancy (Bi *et al.* 2012;
62 Lemmon *et al.* 2012), and thus, we evaluated the efficacy of our exome capture method to validate
63 our results. To do so, we used several metrics and statistics, which we outline below.

- 64 1. **Sequencing:** The first step in a next-generation experiment is to filter raw data for quality.
65 Here, through a rigorous filtration, we lost 45% to 55% of our data, largely because $\approx 70\%$ of
66 our paired-end reads could be merged into a single read. The quality of the filtered data was
67 high; for all but one capture experiment, the average Phred quality score was the maximum
68 possible (36). Despite this aggressive filtration, we retained enough data to get high coverage
69 of both nuclear exon regions ($>100\times$) and mitochondrial DNA ($>1000\times$) (Table S3).
- 70 2. ***de novo* Assembly:** In exome capture experiments with non-model organisms, researchers
71 typically do not know the sequence flanking targets. Because of edge effects, inclusion of
72 just the target sequence will lead to reduced mapping efficiency (Bi *et al.* 2012). However, a
73 portion of the flanking region is also captured and can be reconstructed using *de novo* assem-
74 blers. Using this approach, we recovered an average of 60% longer sequence length (Table
75 S2). Most of this additionally assembled sequence is non-coding sequence surrounding our
76 target exons.

77 3. **Annotation:** We annotated the *de novo* assemblies of the cleaned sequence reads to identify
78 the targets to which they matched. We successfully assembled 100% of targeted exons. The
79 majority of assembled contigs were longer than the target exons. Many of the assembled
80 contigs did not match any of our targeted exons. Although some of these contigs are likely
81 "biologically real", we opted for a conservative approach and excluded them from down-
82 stream analyses. These contigs could be analyzed in future work.

83 4. **Sensitivity:** Sensitivity is a measure of what portion of in-target assemblies are represented
84 by sequence data. Here, every single exon was covered by at least $1\times$ coverage.

85 5. **Specificity:** Specificity is measured as the percentage of cleaned reads that map onto tar-
86 geted regions. Depending on the technology, specificity can range from 10% to 90% across
87 experiments (Sulonen *et al.* 2011), but variance in specificity should be low within an exper-
88 iment. Low variance suggests that the procedure worked uniformly across captures. In this
89 study, we found that specificity ranged from 58.2% to 75.2% across captures and that vari-
90 ance within an array was low (Table S3; Fig. S9). Our experiment worked comparatively
91 better than many other array-based exon capture studies (Mamanova *et al.* 2010).

92 6. **Coverage metrics:** Coverage metrics illustrate the uniformity of results across libraries on
93 the same capture. Coverage can indicate if coverage responds to other characteristics of the
94 data (such as GC-content) as expected. Here, we measured several metrics:

- 95 • *Correlation of coverage across libraries in a capture:* high correlation suggests consistency of
96 capture. Coverage was highly correlated across libraries on the same capture ($r > 0.97$;
97 Fig. S10).
- 98 • *Correlation of loci across captures:* high correlation suggests that differences in capture
99 efficiency across loci are due more to locus-specific effects and less to stochastic effects
100 of a given capture experiment. Coverage of orthologous loci across different capture
101 experiments was significant and high ($r = 0.53 - 0.70$; Fig. S11).
- 102 • *Density plots of coverage:* ideally, coverage across loci should be tightly distributed, indi-
103 cating no prevalent bias in capture efficiency at a given locus. Most loci had about $200\times$
104 coverage, although there is some spread in the distribution of coverage across loci (Fig.
105 S12).
- 106 • *Correlation of divergence with coverage:* in our experiments, we were capturing orthol-
107 ogous loci across two sister lineages in a lineage-pair. To ensure no bias in capture
108 efficiency if the two orthologs were divergent, we included both orthologs on our array.
109 As such, we would expect to see little correlation of coverage with sequence divergence
110 between the orthologs. However, we see significant but modest correlation between
111 divergence and coverage ($r = 0.16 - 0.19$; Fig. S13).
- 112 • *Correlation of coverage with GC-content:* coverage is expected to have a hump-shaped
113 relationship with GC-content, such that coverage is low at low- and high-GC content.
114 We recover this pattern in our data (Fig. S14).

115 In sum, our results indicate that our exome capture experiments were successful and that our
116 downstream inference should be robust to technical artifacts.

117 4 Fitting Clines

118 To fit clines, our approach was as following for every variant:

- 119 1. Calculate allele frequency for the variant in Northern and Southern parental populations
120 based on genotype data from transcriptomes
- 121 2. Calculate allele frequency data for the variant in the nine transect populations using pooled
122 exome capture data – allele frequencies were calculated using the samtools mpileup feature to
123 count how many reads corresponded to the variant
- 124 3. Set allele frequencies to missing for any pooled population with $<50\times$ coverage
- 125 4. Keep only those variants that have allele frequency data at five or more transect populations
- 126 5. Keep only those variants that have ≥ 0.5 difference in allele frequency across the Northern
127 vs. Southern populations
- 128 6. If necessary, invert allele frequencies for the cline so that Northern populations are at low
129 frequency for the variant of interest
- 130 7. Call patterns at a variant a sweep if:
 - 131 • if parental populations have ≥ 0.5 difference in allele frequency
 - 132 • if allele frequencies for all 9 populations in the transect are either (1) uniformly between
133 0 and 0.2 or (2) between 0.8 and 1
- 134 8. If not a sweep, try fitting a cline if:
 - 135 • The two Northern tail populations (the parental 'N' and 10-km North population) and
136 the two Southern tail populations (the parental 'S' and 10-km South population) have
137 an allele frequency difference ≥ 0.5
 - 138 • If there are data for five of the seven central populations to be used in cline fitting
 - 139 – Define p_{min} as the minimum allele frequency across the transect
 - 140 – Define p_{max} as the maximum allele frequency across the transect
 - 141 – Set allele frequencies of 0 to 0.001 to avoid odd edge behavior of likelihood function
 - 142 – Set allele frequencies of 1 to 0.999 to avoid odd edge behavior of likelihood function
 - 143 – Rescale allele frequencies to be on a $0 \rightarrow 1$ scale based on p_{min} and p_{max}
 - 144 – Fit cline using a brute-force maximum-likelihood approach in Python using the
145 maximum-likelihood equation found in (Porter *et al.* 1997)
 - 146 * This maximum-likelihood equation calculates the likelihood of the empirical
147 allele frequencies compared to those expected under a simple sigmoidal model,
148 weighting each deme by its sample size
 - 149 * Search space for cline width is defined as being between $1e2$ and $2e4$, in units
150 of $5e2$
 - 151 * Search space for cline center is defined as being between $1e2$ and $5e3$, in units
152 of 25

153

* The optimization function (`scipy.optimize`) was used to explore this space

154

* Final parameter estimates can fall outside the search space because the optimization function does "finishing" after exploring the parameter space

155

156

9. If not a sweep or cline, the variant goes uncharacterized

References

- 157
158 Altschul, S., T. Madden, A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. Lipman, 1997. Gapped
159 BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids*
160 *Research* 25:3389–3402.
- 161 Bi, K., D. Vanderpool, S. Singhal, T. Linderoth, C. Moritz, and J. M. Good, 2012. Transcriptome-
162 based exon capture enables highly cost-effective comparative genomic data collection at moder-
163 ate evolutionary scales. *BMC genomics* 13:1.
- 164 Lemmon, A., S. Emme, and E. Lemmon, 2012. Anchored hybrid enrichment for massively high-
165 throughput phylogenomics. *Systematic biology* 61:727–744.
- 166 Mamanova, L., A. J. Coffey, C. E. Scott, I. Kozarewa, E. H. Turner, A. Kumar, E. Howard, J. Shen-
167 dure, and D. J. Turner, 2010. Target-enrichment strategies for next-generation sequencing. *Nature*
168 *methods* 7:111–118.
- 169 Meunier, J. and L. Duret, 2004. Recombination drives the evolution of gc-content in the human
170 genome. *Molecular biology and evolution* 21:984–990.
- 171 Porter, A., R. Wenger, H. Geiger, A. Scholl, and A. Shapiro, 1997. The *Pontia daplidice-edusa* hybrid
172 zone in northwestern Italy. *Evol.* 52:1561–1573.
- 173 R Development Core Team, 2011. R: A Language and Environment for Statistical Computing. R
174 Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>. ISBN
175 3-900051-07-0.
- 176 Singhal, S., 2013. De novo transcriptomic analyses for non-model organisms: an evaluation of
177 methods across a multi-species data set. *Mol. Ecol. Res.* 13:403–16.
- 178 Singhal, S. and C. Moritz, 2013. Reproductive isolation between phylogeographic lineages scales
179 with divergence. *Proceedings of the Royal Society of London B: Biological Sciences* 280:20132246.
- 180 Slater, G. and E. Birney, 2005. Automated generation of heuristics for biological sequence compar-
181 ison. *BMC Bioinformatics* 6:31.
- 182 Smit, A., R. Hubley, and P. Green, 2013. Repeatmasker open-3.0. URL www.repeatmasker.org.
- 183 Sulonen, A., P. Ellonen, H. Almusa, M. Lepistö, S. Eldfors, S. Hannula, T. Miettinen, H. Tynismaa,
184 P. Salo, C. Heckman, *et al.*, 2011. Comparison of solution-based exome capture methods for next
185 generation sequencing. *Genome biology* 12:R94.
- 186 Yang, Z., 2007. Paml 4: Phylogenetic analysis by maximum likelihood. *Mol. Bio. Evol.* 24:1586–
187 1591.

188 5 Supplemental Figures

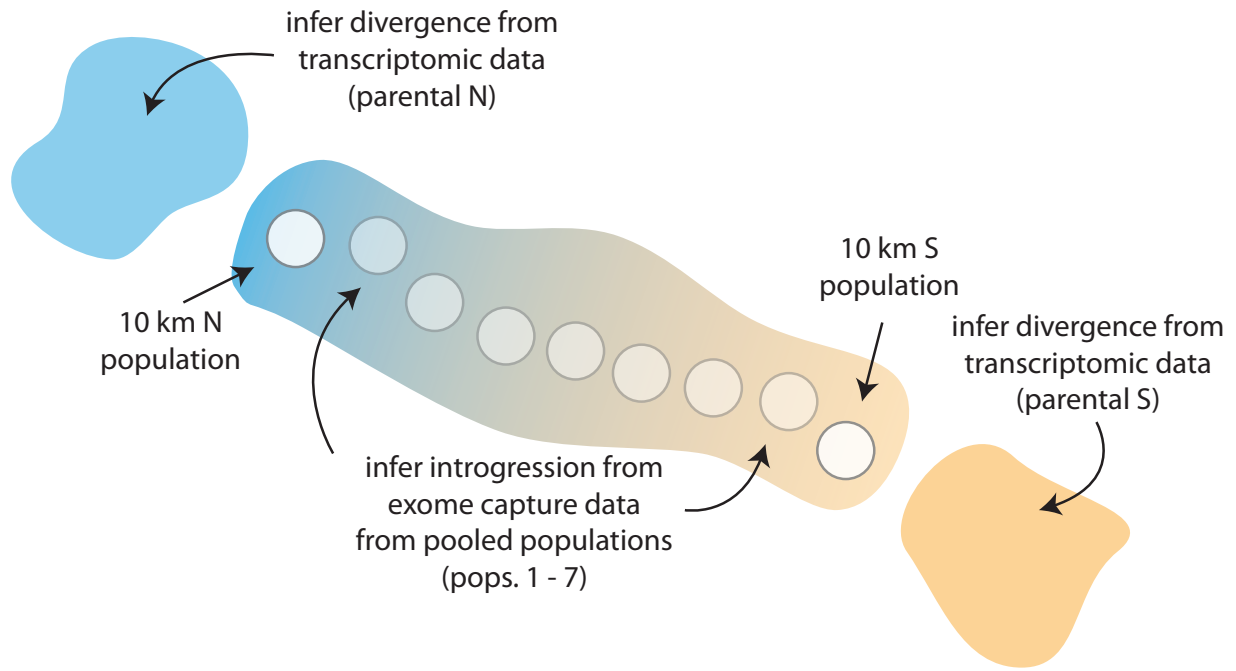


Figure S1: Basic sampling scheme used in this study. Transcriptome data from geographically isolated "parental" populations were used to characterize genomic divergence and to design arrays. Anonymously pooled exome capture data from the nine populations in the hybrid zones were used to infer introgression extent. However, only the central seven populations were used in cline fitting as the 10-kilometer populations were off the linear transect through the hybrid zone.

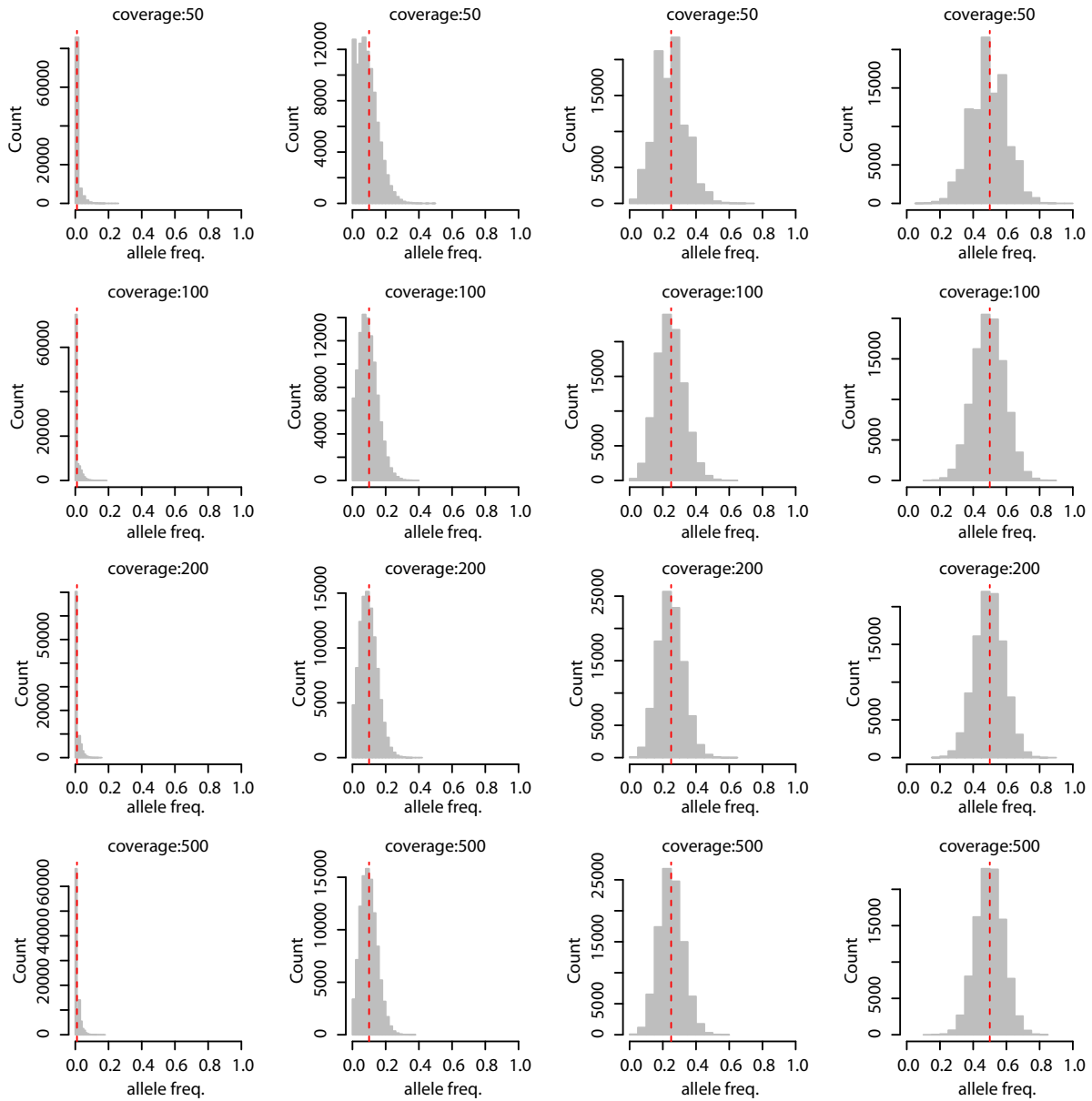


Figure S2: Results from simulations exploring how increasing sequencing effort influences our ability to infer allele frequencies from pooled populations accurately. Each column represents a different simulated allele frequency: 0.01, 0.1, 0.25, 0.5; each row represents a different simulated coverage: 50, 100, 200, 500. The dotted red line indicates the true allele frequency. These results suggest, that across a range of allele frequencies, increasing sequencing coverage above $50\times$ only modestly improves the accuracy of allele frequency estimates.

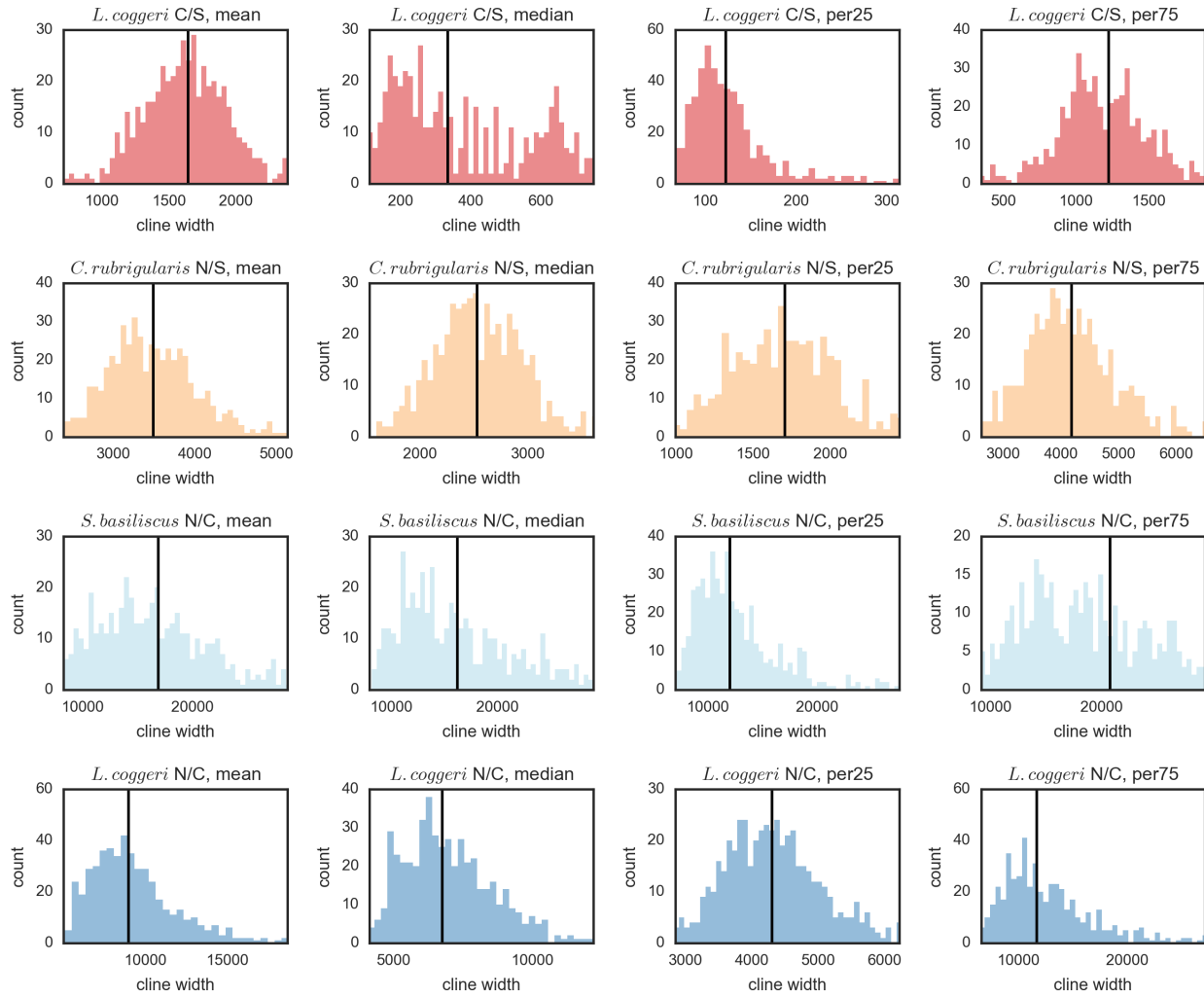


Figure S3: Results from simulations exploring role of sampling drift in inferring cline widths from pooled populations. Each simulation was parameterized by the empirical sampling design for the relevant contact (*i.e.*, location of sampled populations, number of sampled individuals, and average exonic coverage). We ran 500 simulations across each of four cline widths: the mean, median, 25%, and 75% cline width inferred for the contact. We then used our cline fitting procedure to infer cline width from the simulated data. Shown is the distribution of the inferred cline widths, and in black, the cline width under which the data were simulated. These results suggest that, although our smaller sample sizes and pooling likely increased the error in our cline estimates, this error is insufficient to explain differences in the mean and variance of cline widths among hybrid zones.

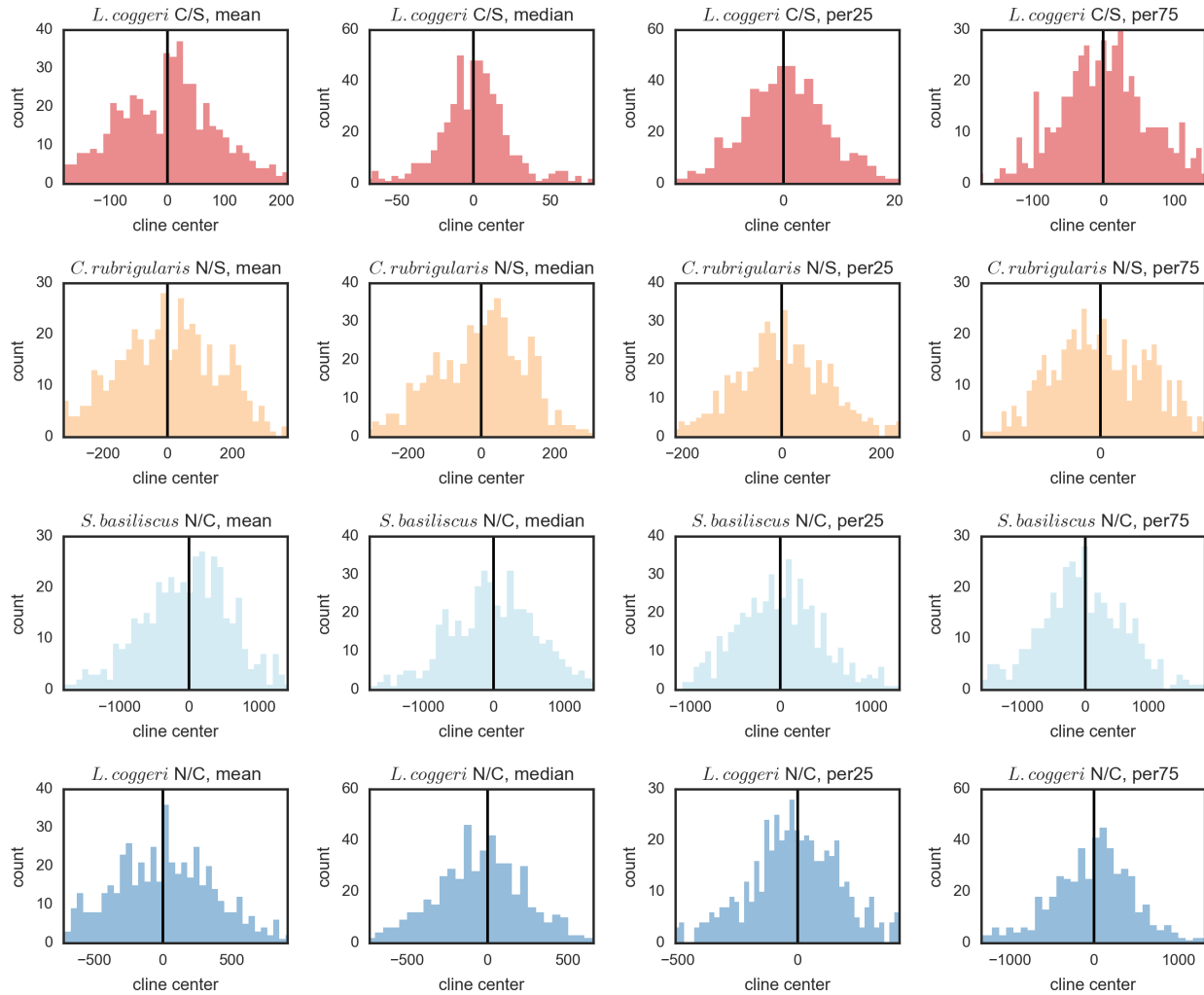


Figure S4: Results from simulations exploring role of sampling drift in inferring cline centers from pooled populations. Each simulation was parameterized by the empirical sampling design for the relevant contact (*i.e.*, location of sampled populations, number of sampled individuals, and average exonic coverage). We ran 500 simulations across each of four cline widths: the mean, median, 25%, and 75% cline width inferred for the contact. We then used our cline fitting procedure to infer cline center from the simulated data. Shown is the distribution of the inferred cline center, and in black, the cline center under which the data were simulated. These results suggest that, although our smaller sample sizes and pooling likely increased the error in our cline estimates, this error is insufficient to explain differences in mean cline centers among hybrid zones.

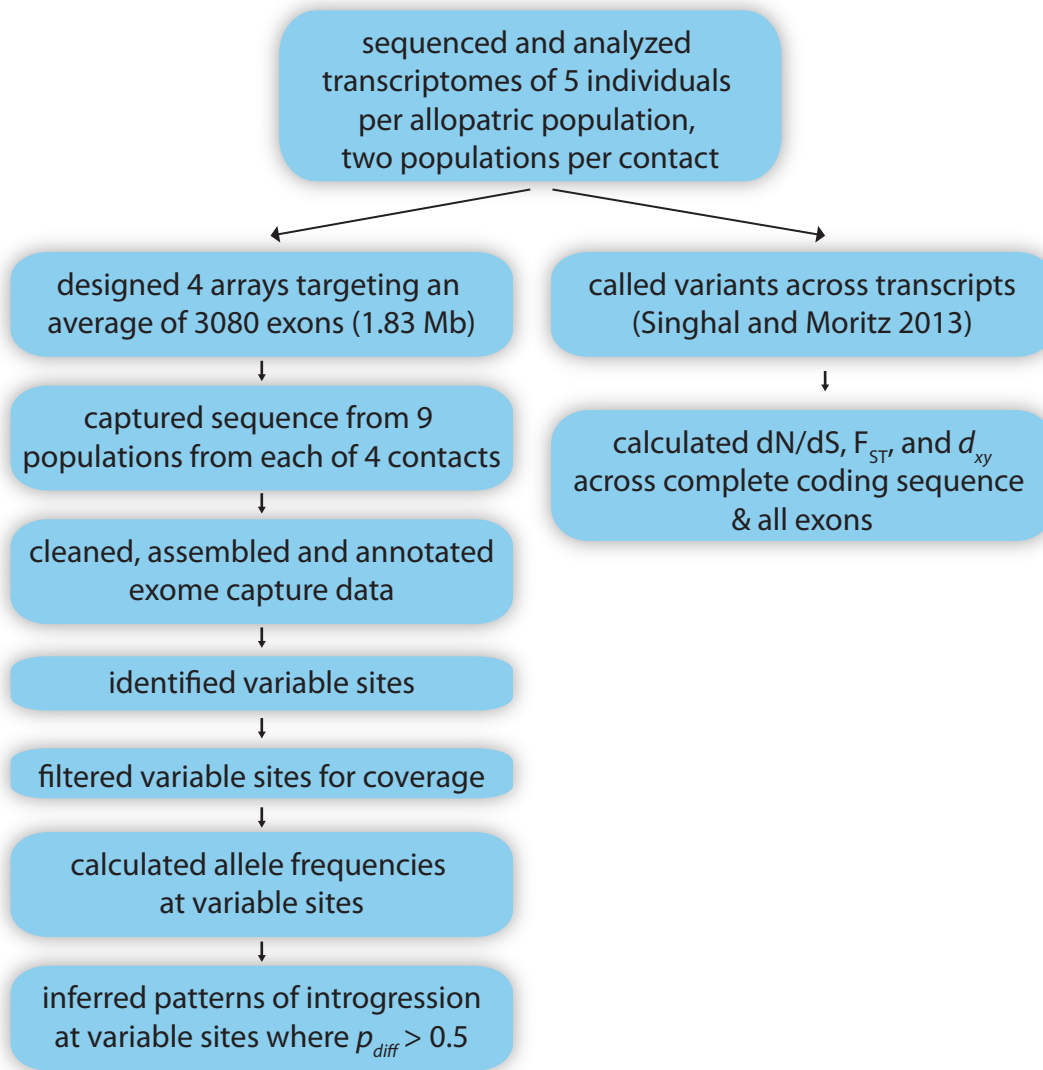


Figure S5: Summary of bioinformatics and inference pipeline used in this study. Transcriptome data from allopatric populations far from the contact zone were used both to design exome capture arrays and to infer patterns of genomic divergence between lineages. These data were published and analyzed in previous studies (Singhal 2013; Singhal and Moritz 2013). Pooled exome capture data from nine populations sampled through the hybrid zone were used to infer patterns of introgression. Here, p_{diff} refers to the maximum difference in allele frequency seen across the geographically-isolated populations (parental N, parental S) and the 10 km N and S populations (Fig. S1).

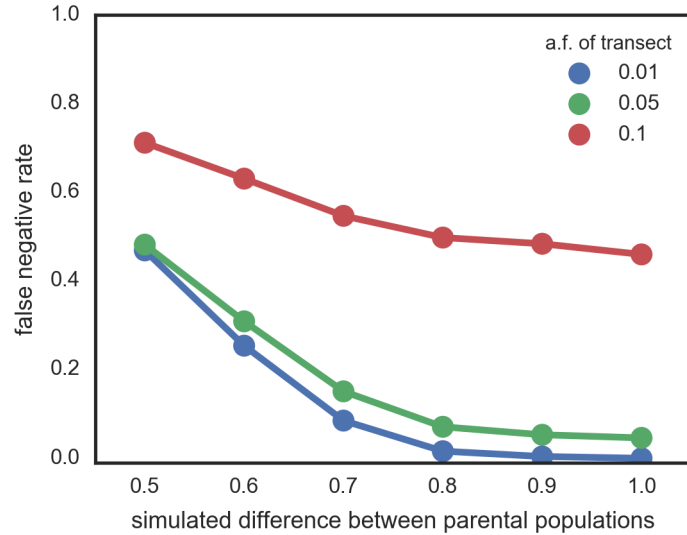


Figure S6: Results from simulations estimating the false negative rate in inferring loci with a “sweep” pattern from pooled populations. Simulations were parameterized using the average sampling design across contacts (*i.e.*, location of sampled populations, number of sampled individuals, and average exonic coverage). We set the parental populations to have varying levels of allele frequency difference ($p_{diff} \geq 0.5$; shown on the x-axis) and set the transect populations to have expected allele frequencies of either 0.01, 0.05, 0.1. Thus, this parameter space encompasses that for loci categorized as “sweeps”. We ran 1000 simulations across each parameter set and calculated what percentage of simulations failed to recover a “sweep” pattern (or, the false negative rate). False negative rates are high when the parental populations are less differentiated at a variant or when the alternative variant is simulated segregating at non-zero frequencies in the transect populations.

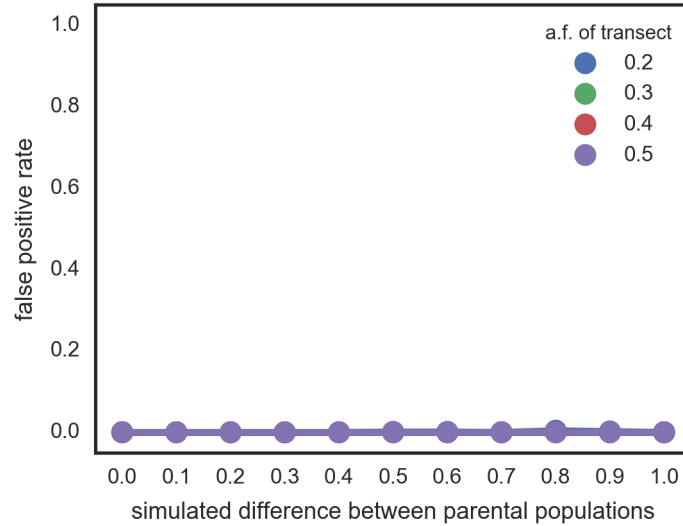


Figure S7: Results from simulations estimating the false positive rate in inferring loci with a “sweep” pattern from pooled populations. Simulations were parameterized using the average sampling design across contacts (*i.e.*, location of sampled populations, number of sampled individuals, and average exonic coverage). We set the parental populations to have varying levels of allele frequency difference (shown on the x-axis) and set the transect populations to have expected allele frequencies ranging from 0.2 to 0.5. None of this parameter space falls in the “sweep” category. We ran 1000 simulations across each parameter set and calculated what percentage of simulations recovered a “sweep” pattern (or, the false positive rate). Across this parameter space, we recovered no false positives. However, these simulations are simple in their approach, and our actual false positive rate is likely non-zero.

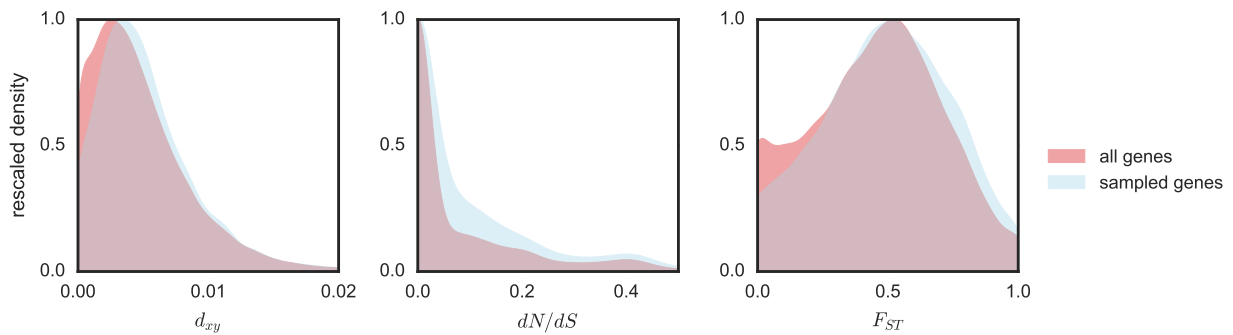


Figure S8: Density histograms comparing distributions of summary statistics (d_{xy} , dN/dS , F_{ST}) for all transcripts sequenced for all seven focal lineages (in red) and for the subset of transcripts targeted on exome capture arrays (in blue). Rescaled density was calculated by dividing density estimates by the maximum density seen for the distribution. These results show that transcripts whose exons were included in the exome capture array were more divergent than those transcripts that were not included. However, general patterns of molecular divergence are similar.

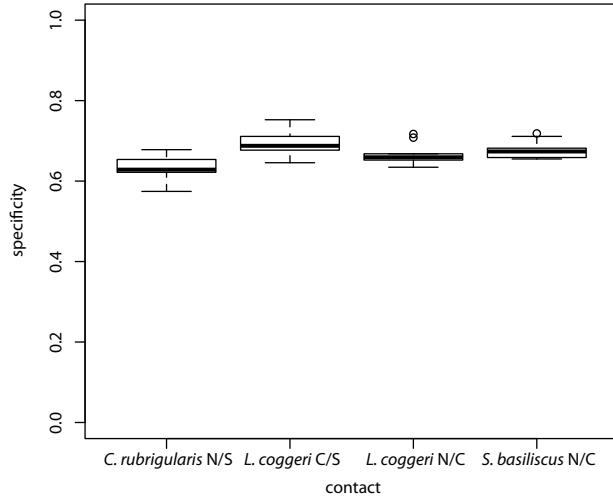


Figure S9: Specificity, or proportion of cleaned reads mapping onto target exons, summarized across all libraries for a given capture. Captures were done by contact zone. In general, specificity was high and uniform across libraries in a given contact; these results help validate the success of the exome capture experiment.

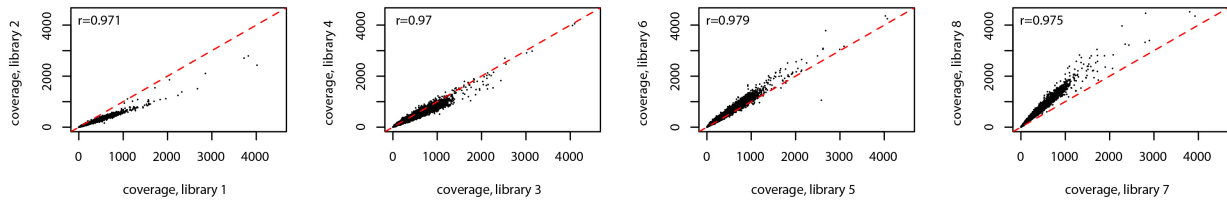


Figure S10: Correlation in coverage across the same loci from different libraries from the same capture experiment. Captures were done by contact zone; here, we show results from a randomly selected contact zone, *Carlia rubrigularis* N/S. The red dotted line is at unity. In general, correlation is high, suggesting that libraries within a given capture experiment performed similarly. These results help validate the success of the exome capture experiment.

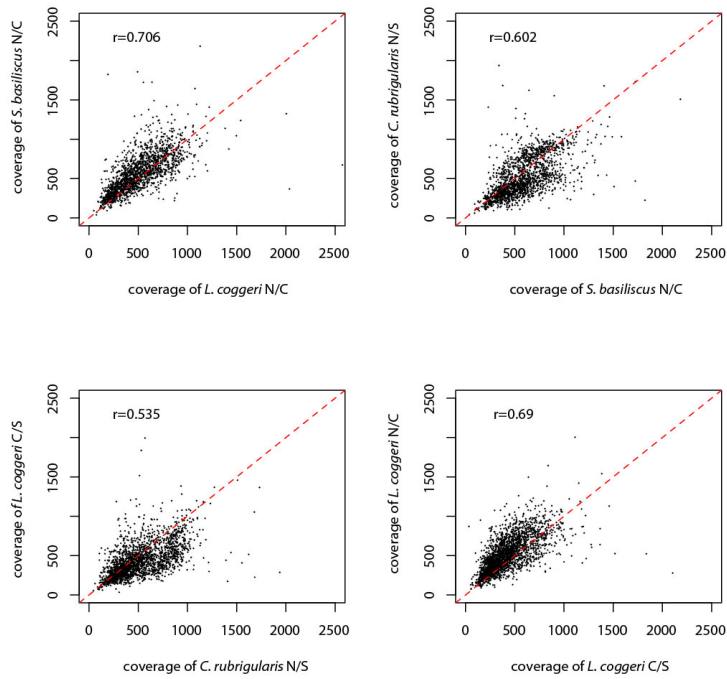


Figure S11: For orthologs shared across multiple exome capture arrays, correlation in coverage between different capture experiments. Captures were done by contact zone. The red dotted line is at unity. In general, correlation is high, suggesting that locus-specific characteristics that affect capture success (*i.e.*, GC content) had similar effects across contacts. These results help validate the success of the exome capture experiment.

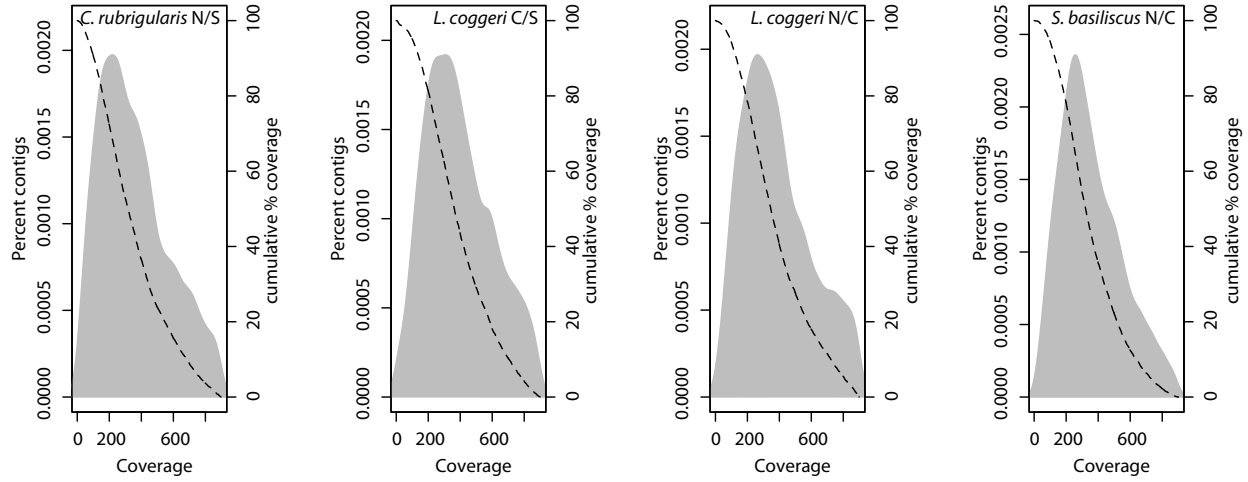


Figure S12: Density plots of locus-wide coverage by capture experiment, with frequencies shown on the left y -axis. Captures were done by contact zone. The dotted line shows accumulation of coverage across increasing coverage levels, with percentages shown on the right y -axis. Although there is strong mode in the coverage distributions, coverage across loci shows considerable heterogeneity. This is typical of most exome capture experiments, e.g. (Bi *et al.* 2012).

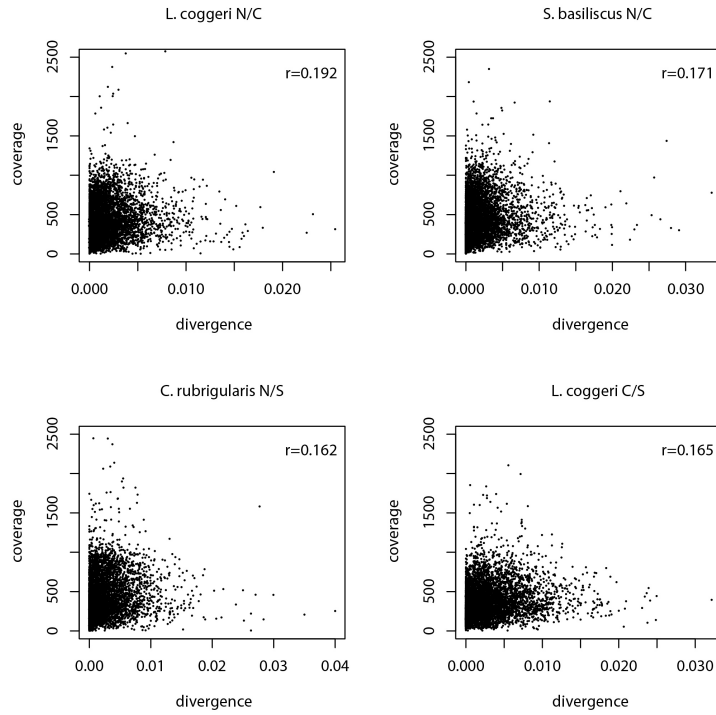


Figure S13: Correlation between coverage at a given locus and its sequence divergence (d_{xy}) between lineages in the contact zone. Because each exon on the exome capture array included alleles from both lineages, we would predict minimal correlation between divergence and coverage. However, we see modest but significant correlation across all four contact zones.

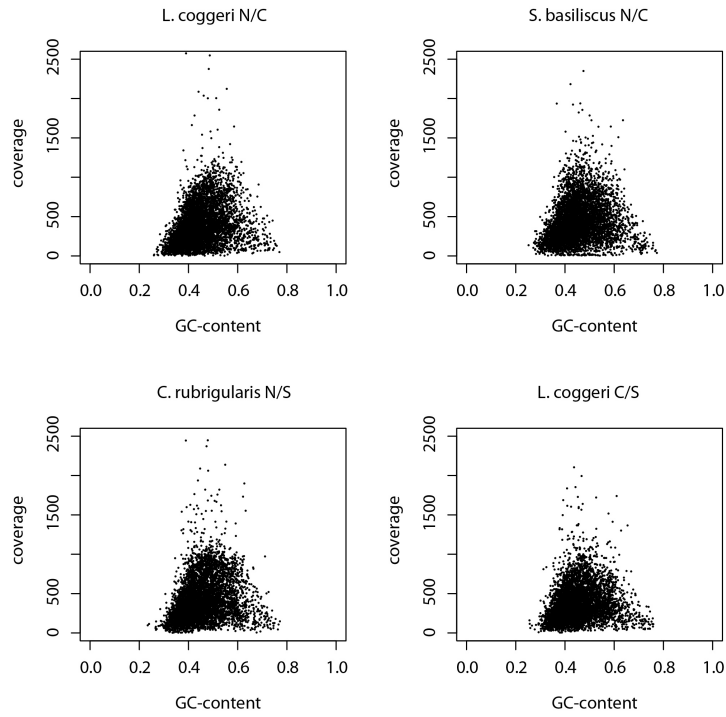


Figure S14: Relationship between coverage at a given locus and GC-content at that locus. As expected, we recover a hump-shaped pattern in which loci with moderate levels of GC-content have the highest coverage. These results help validate the success of the exome capture experiment.

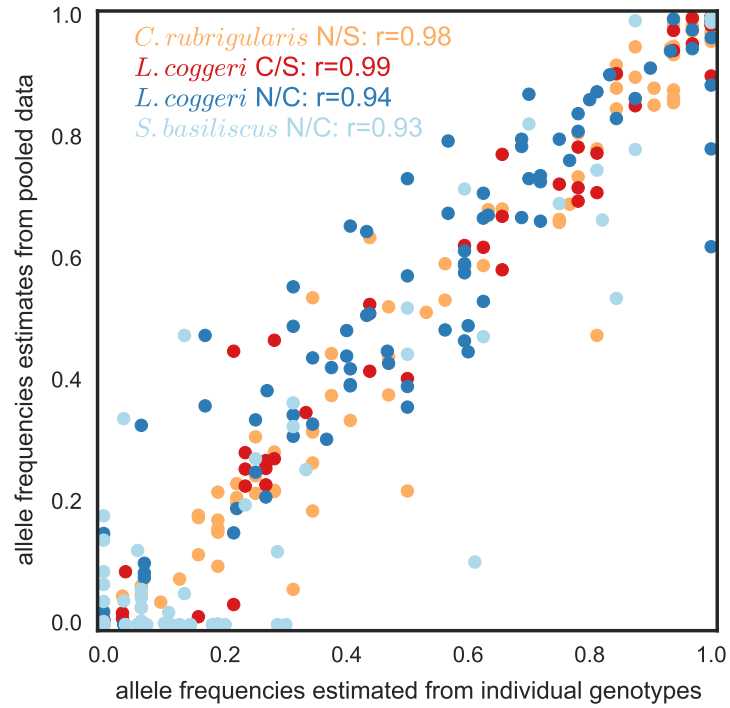


Figure S15: Correlation between allele frequencies estimated from individual genotyping (Singhal and Moritz 2013) and allele frequencies estimated from sequencing anonymously pooled populations of the same individuals. Points are colored by the contact zone from which the allele was measured. Correlation between both estimates of allele frequencies was high, suggesting that our pooled approach allows us to estimate allele frequency accurately.

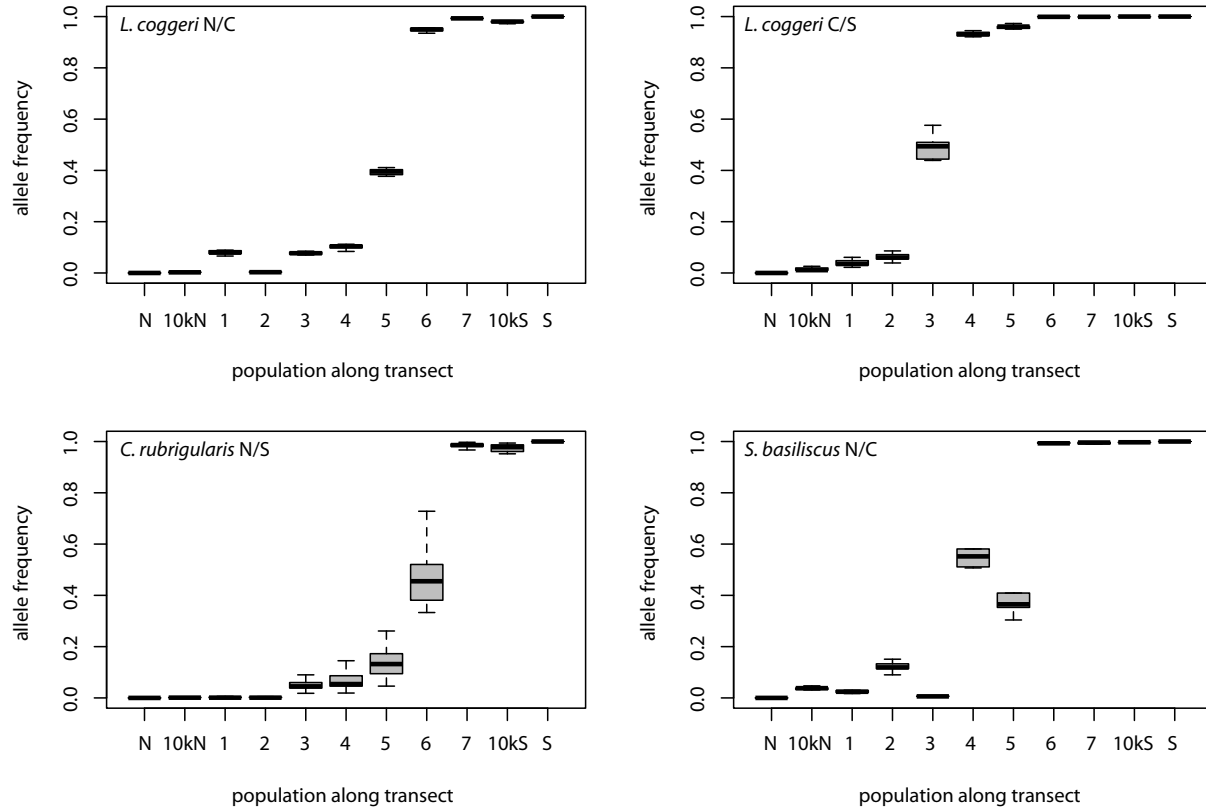


Figure S16: Variance in allele frequency estimates across nearly-fixed and fixed single nucleotide polymorphisms (SNPs) between the mtDNA sequences of the two lineages meeting in each contact zone. As mtDNA does not recombine, all SNPs across mtDNA should have the same allele frequency, if we assume no heteroplasmy or numts. Populations are labeled as shown in Fig. S1. Only SNPs where the difference in allele frequency between the two parental populations was greater than 0.8 were included. Variance in allele frequency estimates across SNPs was low, suggesting that our pooled strategy did not introduce significant heterogeneity. However, we note that coverage of mtDNA was higher than that of nuclear markers (Table S3), and increased coverage does slightly improve the accuracy of pooled allele frequency estimates (Fig. S2).

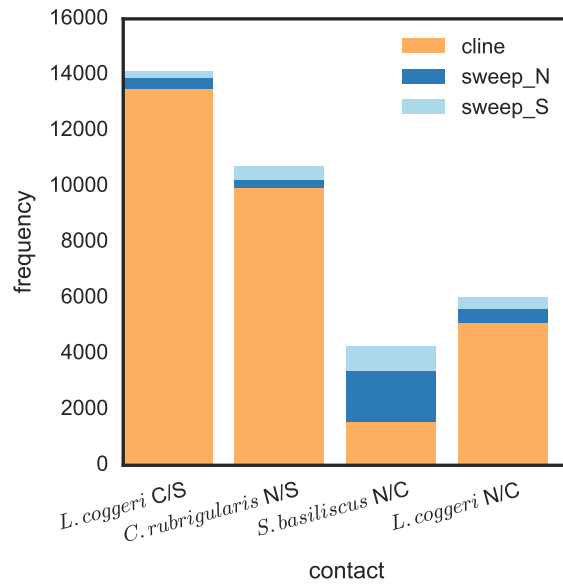


Figure S17: Type of clines inferred at filtered SNPs by each contact zone. Lineage-pairs are ordered most to least divergent (L → R). 'sweep_N' refers to those loci where the allele belonging to the northern lineage in each contact zone has swept into the southern lineage, and 'sweep_S' refers to southern alleles introgressing into the northern lineage. We were able to fit considerably more clines in the more divergent lineage-pairs partially because SNPs between these lineages were more highly differentiated.

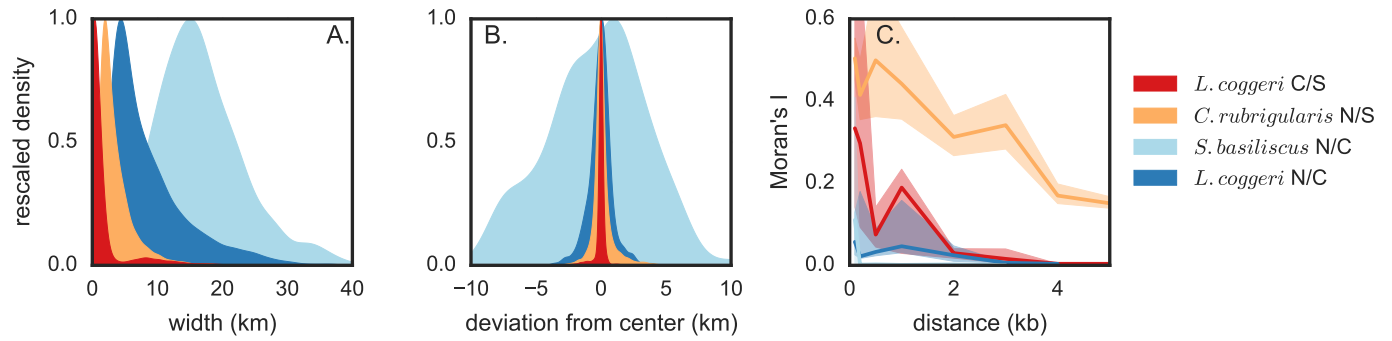


Figure S18: Distributions for A. cline width and B. cline center across an average of 3.2K clines at contact zones between each of the four lineage-pairs. C. Moran's I, a measure of spatial auto-correlation applied to genomic distance, for cline width at each of contact across an average of 2.2K comparisons for clines ≤ 500 bp and 490 comparisons for clines > 500 bp. Uncertainty in Moran's I was estimated by drawing 100 bootstrap samples and recalculating means. Rescaled density was calculated by dividing density estimates by the max density seen for the distribution. Only those clines found in the 1120 exons targeted across all four contact zones are shown. Patterns are quantitatively and qualitatively similar to those found across all exons (Fig. 3), although patterns for Moran's I are considerably noisier.

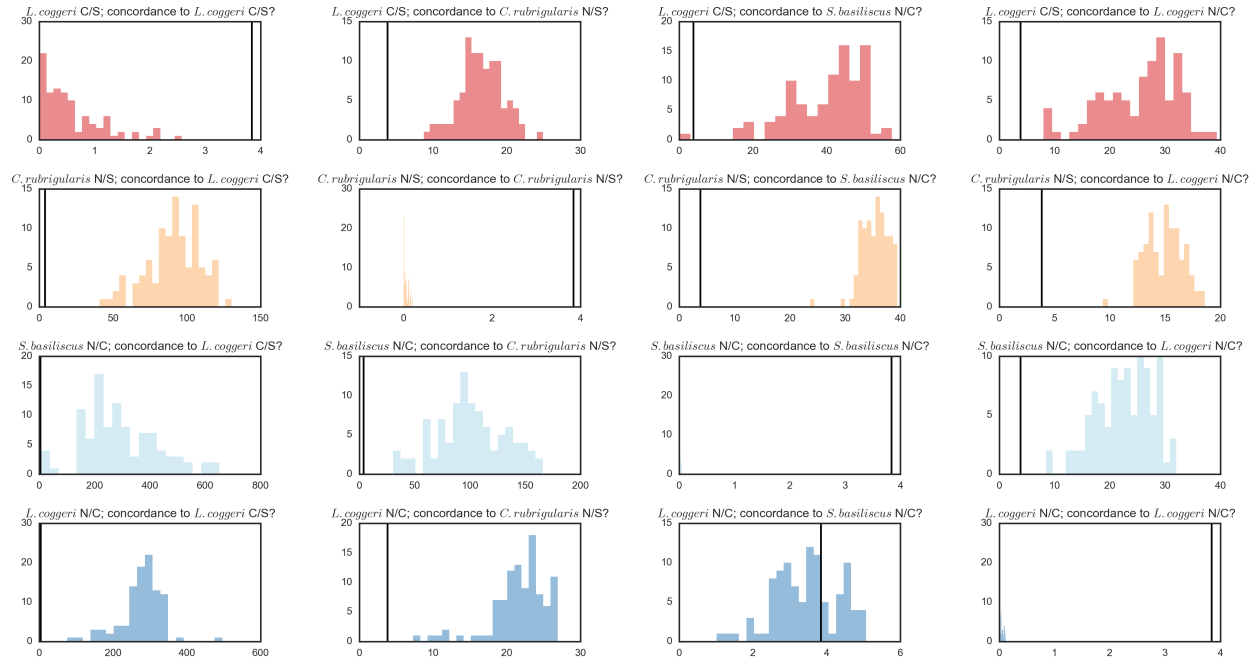


Figure S19: A test for concordance in cline widths across contact zones. For each contact, we randomly selected 100 clines from the 45% - 55% distribution of cline widths. We then repeated our cline-fitting procedure, forcing these clines to be concordant with the median cline width for itself and each other contact zone. Cline center was allowed to vary. Shown are the values for the log-likelihood test between forcing concordance (our null model) to allowing cline width to vary (our alternative model). The black line indicates the chi-square significance value for $p < 0.05$, $d.f. = 1$. For all contacts, constraining cline width of the sampled clines to the median cline width of the distribution from which they were sampled did not lead to worse model fits. This was as expected. However, for all comparisons but *L. coggeri* N/C to *S. basiliscus* C/S, allowing cline width to vary lead to significantly better model fits than forcing concordance. This suggests that median cline widths vary significantly between all contacts. However, *L. coggeri* N/C median cline width does not appear to be significantly different from *S. basiliscus* C/S, although *S. basiliscus* C/S median cline width is significantly different from *L. coggeri* N/C.

6 Supplemental Tables

contact	population name	sample size	latitude	longitude	transect location (m)
<i>C. rubrigularis</i> N/S	pop1	16	-17.156	145.564	764
<i>C. rubrigularis</i> N/S	pop2	16	-17.154	145.569	1278
<i>C. rubrigularis</i> N/S	pop3	16	-17.153	145.578	2334
<i>C. rubrigularis</i> N/S	pop4	16	-17.147	145.581	2676
<i>C. rubrigularis</i> N/S	pop5	16	-17.148	145.586	3175
<i>C. rubrigularis</i> N/S	pop6	16	-17.143	145.590	3689
<i>C. rubrigularis</i> N/S	pop7	16	-17.133	145.616	6732
<i>C. rubrigularis</i> N/S	parental N	5	-16.611	145.452	NA
<i>C. rubrigularis</i> N/S	10 km N	16	-17.075	145.596	NA
<i>C. rubrigularis</i> N/S	10 km S	16	-17.205	145.679	NA
<i>C. rubrigularis</i> N/S	parental S	5	-17.694	145.695	NA
<i>L. coggeri</i> C/S	pop1	14	-17.172	145.687	0
<i>L. coggeri</i> C/S	pop2	16	-17.205	145.679	3730
<i>L. coggeri</i> C/S	pop3	16	-17.215	145.687	4764
<i>L. coggeri</i> C/S	pop4	16	-17.215	145.686	4803
<i>L. coggeri</i> C/S	pop5	16	-17.215	145.688	4829
<i>L. coggeri</i> C/S	pop6	16	-17.220	145.695	5319
<i>L. coggeri</i> C/S	pop7	16	-17.273	145.663	11295
<i>L. coggeri</i> C/S	parental N	5	-16.976	145.777	NA
<i>L. coggeri</i> C/S	10 km N	16	-17.142	145.629	NA
<i>L. coggeri</i> C/S	10 km S	16	-17.295	145.712	NA
<i>L. coggeri</i> C/S	parental S	5	-17.676	145.713	NA
<i>L. coggeri</i> N/C	pop1	15	-16.659	145.480	3497
<i>L. coggeri</i> N/C	pop2	16	-16.660	145.485	4075
<i>L. coggeri</i> N/C	pop3	16	-16.664	145.492	4988
<i>L. coggeri</i> N/C	pop4	16	-16.664	145.496	5546
<i>L. coggeri</i> N/C	pop5	15	-16.666	145.500	6029
<i>L. coggeri</i> N/C	pop6	16	-16.671	145.503	6419
<i>L. coggeri</i> N/C	pop7	16	-16.675	145.506	6740
<i>L. coggeri</i> N/C	parental N	5	-16.579	145.315	NA
<i>L. coggeri</i> N/C	10 km N	15	-16.617	145.458	NA
<i>L. coggeri</i> N/C	10 km S	16	-16.753	145.593	NA
<i>L. coggeri</i> N/C	parental S	5	-16.976	145.777	NA
<i>S. basiliscus</i> N/C	pop1	16	-17.608	145.772	0
<i>S. basiliscus</i> N/C	pop2	10	-17.608	145.768	679
<i>S. basiliscus</i> N/C	pop3	7	-17.626	145.744	4494
<i>S. basiliscus</i> N/C	pop4	16	-17.665	145.723	7881
<i>S. basiliscus</i> N/C	pop5	14	-17.655	145.717	8893
<i>S. basiliscus</i> N/C	pop6	8	-17.673	145.715	9414
<i>S. basiliscus</i> N/C	pop7	8	-17.694	145.695	12707
<i>S. basiliscus</i> N/C	parental N	15	-17.292	145.634	NA
<i>S. basiliscus</i> N/C	10 km N	5	-17.579	145.697	NA
<i>S. basiliscus</i> N/C	10 km S	5	-17.699	145.523	NA
<i>S. basiliscus</i> N/C	parental S	16	-18.199	145.849	NA

Table S1: Summary of geographic locations and sample sizes of populations included in this work. Parental populations and 10 km populations were sampled off the transect through the hybrid zone; thus, we do not report a transect location for them.

contact	number of targets	total length as designed (bp)	number of probes	total target length as assembled (bp)
<i>C. rubrigularis</i> N/S	3224	1.86e6	9.70e5	3.02e6
<i>L. coggeri</i> C/S	3333	1.83e6	9.57e5	3.08e6
<i>L. coggeri</i> N/C	2889	1.81e6	9.69e5	3.62e6
<i>S. basiliscus</i> N/C	2870	1.82e6	9.68e5	2.95e6

Table S2: Summary of exome capture array designs and resulting assemblies. We were able to extend the assembled target length from the total target length anywhere from 60% to 100% by capturing and assembling the sequence flanking our targets.

contact	population name	raw data (bp)	cleaned data (bp)	% of raw data kept	avg. quality score (post-cleanup)	specificity	avg. exonic cov.	avg. mtDNA cov.
<i>C. rubrigularis</i> N/S	10 km N	4.8e9	2.2e9	46.7%	35	62.9%	351	1026
<i>C. rubrigularis</i> N/S	10 km S	3.5e9	1.6e9	46.4%	35	65.4%	267	260
<i>C. rubrigularis</i> N/S	pop1	2.9e9	1.3e9	45.4%	35	67.8%	224	943
<i>C. rubrigularis</i> N/S	pop2	6.8e9	3.2e9	46.8%	35	62.8%	488	1057
<i>C. rubrigularis</i> N/S	pop3	5.9e9	2.8e9	47.4%	35	62.2%	433	1055
<i>C. rubrigularis</i> N/S	pop4	6.5e9	3.0e9	46.1%	35	62.6%	463	1005
<i>C. rubrigularis</i> N/S	pop5	8.0e9	3.8e9	47.9%	35	65.8%	545	1047
<i>C. rubrigularis</i> N/S	pop6	5.4e9	2.5e9	46.3%	35	58.2%	410	478
<i>C. rubrigularis</i> N/S	pop7	1.0e10	4.8e9	46.4%	35	57.4%	665	528
<i>L. coggeri</i> C/S	10 km N	2.2e9	1.2e9	55.3%	36	69.9%	381	7792
<i>L. coggeri</i> C/S	10 km S	2.5e9	1.4e9	55.3%	36	72.5%	325	7680
<i>L. coggeri</i> C/S	pop1	6.1e9	3.3e9	54.2%	36	65.8%	285	7721
<i>L. coggeri</i> C/S	pop2	5.4e9	3.0e9	55.7%	36	64.6%	265	7682
<i>L. coggeri</i> C/S	pop3	5.5e9	3.1e9	55.4%	36	68.3%	273	7781
<i>L. coggeri</i> C/S	pop4	4.3e9	2.4e9	55.9%	36	68.8%	505	7778
<i>L. coggeri</i> C/S	pop5	5.3e9	2.9e9	55.2%	36	71.1%	743	7788
<i>L. coggeri</i> C/S	pop6	4.0e9	2.2e9	55.0%	36	67.7%	198	7691
<i>L. coggeri</i> C/S	pop7	1.4e9	8.1e8	58.6%	36	75.2%	514	7681
<i>L. coggeri</i> N/C	10 km N	4.9e9	2.9e9	58.6%	36	65.9%	216	1011
<i>L. coggeri</i> N/C	10 km S	4.3e9	2.5e9	57.8%	36	65.2%	248	7296
<i>L. coggeri</i> N/C	pop1	3.3e9	2.0e9	59.3%	36	70.8%	534	2557
<i>L. coggeri</i> N/C	pop2	3.6e9	2.0e9	57.6%	36	63.6%	478	6370
<i>L. coggeri</i> N/C	pop3	3.5e9	2.0e9	58.2%	36	66.8%	506	2539
<i>L. coggeri</i> N/C	pop4	6.8e9	3.8e9	56.7%	36	66.3%	408	7253
<i>L. coggeri</i> N/C	pop5	1.0e10	5.8e9	56.6%	36	71.7%	485	7336
<i>L. coggeri</i> N/C	pop6	2.3e9	1.4e9	59.0%	36	65.6%	381	7383
<i>L. coggeri</i> N/C	pop7	7.1e9	4.1e9	57.8%	36	63.4%	149	7266
<i>S. basiliscus</i> N/C	10 km N	4.8e9	2.7e9	55.3%	36	65.9%	424	7459
<i>S. basiliscus</i> N/C	10 km S	5.5e9	2.9e9	53.9%	36	67.4%	478	7476
<i>S. basiliscus</i> N/C	pop1	3.0e9	1.7e9	56.3%	36	67.4%	272	5272
<i>S. basiliscus</i> N/C	pop2	2.0e9	1.1e9	55.1%	36	66.2%	197	6392
<i>S. basiliscus</i> N/C	pop3	5.8e9	3.2e9	55.5%	36	65.5%	504	7562
<i>S. basiliscus</i> N/C	pop4	4.9e9	2.7e9	54.9%	36	65.6%	437	7595
<i>S. basiliscus</i> N/C	pop5	2.5e9	1.4e9	55.3%	36	71.1%	247	6331
<i>S. basiliscus</i> N/C	pop6	6.7e9	3.8e9	56.8%	36	71.8%	609	7480
<i>S. basiliscus</i> N/C	pop7	8.5e9	4.7e9	55.4%	36	68.2%	769	7529

Table S3: Summary of data collected, coverage, and specificity across sequenced populations. In general, we lost about half of our raw data because many of our paired-end sequencing reads could be merged into a single read. Still, we had high average coverage across loci, and specificity was fairly uniform across populations.

contact	SNPs pre-filtering	SNPs post-filtering	non-coding SNPs	non-synonymous SNPs	synonymous SNPs
<i>C. rubrigularis</i> N/S	112098	44505	17290	9062	18153
<i>L. coggeri</i> C/S	129153	49354	16397	10291	22666
<i>L. coggeri</i> N/C	242578	79192	29388	15364	34440
<i>S. basiliscus</i> N/C	180056	57007	15098	13192	28717

Table S4: Summary of single nucleotide polymorphisms (SNPs) discovered in the captured targets. SNPs were filtered to remove any SNPs that had $<50\times$ coverage at two or more populations across the central nine populations (Fig. S1). Non-coding SNPs are those occurring in introns and untranslated regions.

type	contact 1	contact 2	number genes compared	correlation	p-value
dN/dS	<i>L. coggeri</i> C/S	<i>C. rubrigularis</i> N/S	9686	0.334	3.77e-251
	<i>L. coggeri</i> C/S	<i>S. basiliscus</i> N/C	9464	0.300	1.31e-195
	<i>L. coggeri</i> C/S	<i>L. coggeri</i> N/C	9471	0.436	0.00e+00
	<i>C. rubrigularis</i> N/S	<i>S. basiliscus</i> N/C	9570	0.288	1.40e-182
	<i>C. rubrigularis</i> N/S	<i>L. coggeri</i> N/C	9312	0.310	3.85e-207
	<i>S. basiliscus</i> N/C	<i>L. coggeri</i> N/C	9210	0.300	1.78e-191
d_{xy}	<i>L. coggeri</i> C/S	<i>C. rubrigularis</i> N/S	9752	0.425	0.00e+00
	<i>L. coggeri</i> C/S	<i>S. basiliscus</i> N/C	9535	0.379	0.00e+00
	<i>L. coggeri</i> C/S	<i>L. coggeri</i> N/C	9535	0.542	0.00e+00
	<i>C. rubrigularis</i> N/S	<i>S. basiliscus</i> N/C	9631	0.372	0.00e+00
	<i>C. rubrigularis</i> N/S	<i>L. coggeri</i> N/C	9372	0.375	0.00e+00
	<i>S. basiliscus</i> N/C	<i>L. coggeri</i> N/C	9269	0.414	0.00e+00
F_{ST}	<i>L. coggeri</i> C/S	<i>C. rubrigularis</i> N/S	9319	0.425	0.00e+00
	<i>L. coggeri</i> C/S	<i>S. basiliscus</i> N/C	9047	0.304	4.1e-193
	<i>L. coggeri</i> C/S	<i>L. coggeri</i> N/C	9078	0.457	0.00e+00
	<i>C. rubrigularis</i> N/S	<i>S. basiliscus</i> N/C	9069	0.315	1.6e-207
	<i>C. rubrigularis</i> N/S	<i>L. coggeri</i> N/C	8847	0.319	6.2e-208
	<i>S. basiliscus</i> N/C	<i>L. coggeri</i> N/C	8684	0.28	1.8e-156

Table S5: Pearson correlations for gene-by-gene comparisons for three different metrics that characterize patterns of locus evolution between lineages. The metrics are dN/dS , raw sequence divergence (d_{xy}), and F_{ST} . We first calculated these metrics for each contact across the coding sequence for all assembled transcripts. These locus-specific metrics were then compared across contacts. On average, there were 15.7 variable sites in a transcript (average length=1331). Across all metrics, and across all comparisons, we recover significant and relatively high correlations. These results suggest repeatability in the patterns of genomic divergence across lineage-pairs related across >15 million years of evolution.

	individually-genotyped clines			pooled clines		
	hybrid index width	average cline width	number of loci	average cline width	median cline width	number of loci
<i>L. coggeri</i> N/C	7865	8957	11	8988	6758	5121
<i>S. basiliscus</i> N/C	NA	4929	2	16907	16209	1556
<i>C. rubrigularis</i> N/S	1500	3704	11	3485	2528	9939
<i>L. coggeri</i> C/S	370	403	11	1643	334	13493

Table S6: Comparison of cline widths estimated from individual genotypes across many fewer markers ($N = 2 - 11$), as previously published in (Singhal and Moritz 2013), and cline widths estimated from pooled data across many more markers ($N = 1.5K - 13.4K$). These results suggest that general quantitative and qualitative patterns between the two cline inference approaches are similar. However, *S. basiliscus* N/C shows a marked difference between the two datasets. This is likely because our estimates for *S. basiliscus* N/C were originally inferred using just two loci.

transect	summary statistic	correlation	p-value	N
<i>L. coggeri</i> C/S	F_{ST}	-0.23	1.7e-48	3902
	d_{xy}	0.05	0.0016	4034
	d_a	-0.11	4.4e-12	4034
	dN/dS	-0.02	0.36	2619
<i>C. rubrigularis</i> N/S	F_{ST}	-0.21	4.7e-37	3446
	d_{xy}	-0.01	0.4	3595
	d_a	-0.13	2.7e-15	3595
	dN/dS	0.01	0.5	2306
<i>S. basiliscus</i> N/C	F_{ST}	-0.03	0.4	872
	d_{xy}	-0.01	0.67	909
	d_a	0.00	0.99	909
	dN/dS	-0.01	0.84	704
<i>L. coggeri</i> N/C	F_{ST}	-0.09	1.4e-5	2235
	d_{xy}	-0.03	0.2	2363
	d_a	-0.03	0.17	2363
	dN/dS	-0.01	0.83	1618

Table S7: Pearson correlations between summary statistics for genomic divergence and cline width across the four sampled transects. The summary statistics are dN/dS , as calculated across the coding sequence for a gene, raw sequence divergence (d_{xy}) and net sequence divergence (d_a) by exon, and F_{ST} by exon. Cline width was averaged across all clines fit in a given exon or gene. Before calculating correlations, we took the natural log of the mean cline width to make the data normal. These results show weak but significantly negative correlations between F_{ST} and d_a and cline width. This suggests loci with high relative differentiation are prone to introgressing less.

contact	α -value	num. of sig. GO terms	FDR
<i>L. coggeri</i> C/S	0.001	6	1.9
	0.01	34	18.7
	0.05	156	93.5
<i>C. rubrigularis</i> N/S	0.001	2	1.7
	0.01	17	16.5
	0.05	94	82.6
<i>S. basiliscus</i> N/C	0.001	1	0.4
	0.01	7	4.2
	0.05	21	21.2
<i>L. coggeri</i> N/C	0.001	2	1.2
	0.01	9	11.7
	0.05	46	58.6

Table S8: The number of Gene Ontology (GO) terms that had significantly narrower cline widths than the background cline width, across different α -values. To determine significance, we conducted 1000 bootstraps, in which we randomly drew clines in proportion to their frequency for a given term. We then calculated the difference in the mean cline width of these subsetted data to the background cline width. Significance is calculated as the proportion of simulations where the difference in cline widths is equal to or greater than the difference calculated for the observed data. Also shown is the predicted false discovery rate (FDR); across almost all contacts and all α -values, the FDR is higher or similar to the number of GO categories found to be significant. Based on these results, we limit our interpretation of the GO analyses.

contact 1	contact 2	N	num. shared, obs.	num. shared, exp.
<i>L. coggeri</i> C/S	<i>C. rubrigularis</i> N/S	1429	17	7.3
<i>L. coggeri</i> C/S	<i>S. basiliscus</i> N/C	417	1	1.2
<i>L. coggeri</i> C/S	<i>L. coggeri</i> N/C	1085	4	3.2
<i>C. rubrigularis</i> N/S	<i>S. basiliscus</i> N/C	418	0	0.9
<i>C. rubrigularis</i> N/S	<i>L. coggeri</i> N/C	1012	3	2.2

Table S9: The number of significant Gene Ontology (GO) terms that are shared between contacts. We determined significance of GO terms as described in Table S8. N indicates the number of GO terms that were represented by genes captured in both contacts. The number of expected shared GO terms was calculated by (1) for each contact, sampling the represented list of GO terms, weighted by the number of significant terms in the contact, (2) calculating the number of shared GO terms between both contacts, and (3) summarizing across 1000 bootstraps. For all contact comparisons but *L. coggeri* C/S and *C. rubrigularis* N/S, we see no greater sharing than expected by chance.

contact 1	contact 2	N	correlation	p-value
<i>L. coggeri</i> C/S	<i>C. rubrigularis</i> N/S	1429	0.062	0.0184
<i>L. coggeri</i> C/S	<i>S. basiliscus</i> N/C	417	0.087	0.0774
<i>L. coggeri</i> C/S	<i>L. coggeri</i> N/C	1085	-0.010	0.736
<i>C. rubrigularis</i> N/S	<i>S. basiliscus</i> N/C	418	-0.066	0.178
<i>C. rubrigularis</i> N/S	<i>L. coggeri</i> N/C	1012	0.079	0.0116
<i>S. basiliscus</i> N/C	<i>L. coggeri</i> N/C	399	-0.002	0.963

Table S10: Pearson correlations across average cline widths for Gene Ontology (GO) terms across all six contact comparisons. We averaged cline widths at variants in all genes for a given GO term and then compared these average values across contacts. While some of these correlations are significant, they are weaker than gene-by-gene correlations (Fig. 4).

contact	differentiation	diversity	N	correlation	p-value
<i>L. coggeri</i> C/S	F_{ST}	π , Northern	12092	-0.475	0.00e+00
	F_{ST}	π , Southern	12092	-0.437	0.00e+00
<i>C. rubrigularis</i> N/S	F_{ST}	π , Northern	12050	-0.371	0.00e+00
	F_{ST}	π , Southern	12050	-0.323	3.40e-290
<i>S. basiliscus</i> N/C	F_{ST}	π , Northern	11141	-0.293	5.69e-219
	F_{ST}	π , Southern	11141	-0.255	6.86e-165
<i>L. coggeri</i> N/C	F_{ST}	π , Northern	10862	-0.318	7.09e-254
	F_{ST}	π , Southern	10862	-0.297	5.72e-220

Table S11: Spearman correlations between locus-specific measures of relative differentiation (F_{ST}) and diversity (π). Differentiation and diversity were calculated across coding sequence in the transcriptome data set. Diversity measures shown for the northern "parental" population in each lineage-pair and the southern "parental" population. As expected, diversity is negatively correlated with relative differentiation.

contact	measure	correlation	p-value	N
<i>L. coggeri</i> C/S	GC*	0.12	2.6e-10	2923
<i>C. rubrigularis</i> N/S	GC*	0.10	6.2e-07	2587
<i>S. basiliscus</i> N/C	GC*	0.05	0.16	772
<i>L. coggeri</i> N/C	GC*	0.01	0.61	1803

Table S12: Spearman correlations between locus-specific measures of GC* and average cline width at that locus. GC* measures expected equilibrium levels of GC content (Meunier and Duret 2004). Because GC-biased gene conversion leads to an excess of GC substitutions, GC* is higher in areas of high recombination compared to regions of low recombination. We see positive correlations between GC* and cline width at two of the four contact zones, suggesting cline width is narrower in regions of low recombination.

contact 1	contact 2	model	AICc	relative likelihood
<i>L. coggeri</i> C/S	<i>C. rubrigularis</i> N/S	width, contact 1 $\sim F_{ST}$, contact 2	2648.02	0.00018
<i>L. coggeri</i> C/S	<i>C. rubrigularis</i> N/S	width, contact 1 $\sim F_{ST}$, contact 2 + width, contact 2	2630.79	1
<i>L. coggeri</i> C/S	<i>L. coggeri</i> N/C	width, contact 1 $\sim F_{ST}$, contact 2	1944.57	1
<i>L. coggeri</i> C/S	<i>L. coggeri</i> N/C	width, contact 1 $\sim F_{ST}$, contact 2 + width, contact 2	1946.58	0.37
<i>C. rubrigularis</i> N/S	<i>S. basiliscus</i> N/C	width, contact 1 $\sim F_{ST}$, contact 2	666.74	0.16
<i>C. rubrigularis</i> N/S	<i>S. basiliscus</i> N/C	width, contact 1 $\sim F_{ST}$, contact 2 + width, contact 2	663.12	1

Table S13: Linear-model fitting results used to determine if the significant correlation in cline widths between contacts (Fig. 4) is merely because cline widths are correlated to F_{ST} (Table S7), which is also correlated between contacts (Table S5). To the three significant comparisons, we fit two linear models. The first model fits cline widths in contact 1 to F_{ST} in contact 2; the second model fits cline widths in contact 1 to both F_{ST} and cline widths in contact 2. Shown are AICc scores and relative likelihoods. Only for *C. rubrigularis* N/S - *L. coggeri* C/S is the second model a significantly better fit. For the other two contacts, this suggests that the correlation among cline widths between contacts might simply be because F_{ST} is correlated across contacts and F_{ST} is correlated to cline widths.