MS. MARCELLA D BAIZ (Orcid ID : 0000-0002-1629-6737)

Article type : Original Article

Multiple forms of selection shape reproductive isolation in a primate hybrid zone

Running title: Introgression in howler monkeys

Marcella D. Baiz[1], Priscilla K. Tucker[1], and Liliana Cortés-Ortiz[1]

[1] Department of Ecology & Evolutionary Biology, University of Michigan
Biological Sciences Building, 1105 N. University, Ann Arbor, MI 48109-1085

Corresponding author: Marcella Baiz, baizm@umich.edu

Abstract

Speciation occurs when populations diverge and become reproductively isolated from each other. Natural selection is commonly accepted to play a large role in this process and it has been widely assumed that reproductive isolation often results as a byproduct of divergence driven by adaptation in allopatry. When such populations come into secondary contact, reinforcement can act to strengthen reproductive isolation, but the frequency and importance of this process is still unknown. Here, we explored genomic signatures of selection in allopatry and sympatry for loci associated with reproductive isolation using a natural primate hybrid zone. By analyzing reduced-representation sequencing data, we quantified admixture and population structure across a howler monkey hybrid zone and examined the relationship between locus-specific differentiation and introgression. We detected extensive admixture that was mostly limited to the narrow contact zone. Loci with reduced introgression into the heterospecific genomic background (the pattern expected for loci associated with reproductive isolation due to selection against hybrids) were significantly more differentiated between allopatric parental populations than loci with neutral and increased introgression, supporting the hypothesis that reproductive isolation is a by-product of divergence in allopatry. Further, loci with reduced introgression showed greater differentiation in sympatry than in allopatry, suggesting a role for reinforcement. Thus, our results reflect multiple forms of selection that have shaped reproductive isolation in this system. We conclude that reproductive isolation may have initially been driven by divergence in allopatry, but later reinforced by divergent selection in sympatry.

Keywords: admixture, introgression, population genomics, genomic clines, speciation, *Alouatta*

Introduction

Natural selection is considered to play an important role in driving speciation (Funk et al. 2006, Sobel et al. 2010). Divergent selection can contribute to speciation when allopatric populations encounter different habitats with different selective pressures (Schluter 2001, 2009). Under such circumstances, it is expected that loci that underlie local adaptation will show allele

frequency differences in populations under different environments (Schluter 2009). Similarly, if allopatric populations encounter environments with similar selection pressures (i.e., uniform selection), different adaptive mutations may be selected for (Schluter 2001, 2009) because different mutations may result in similar optimal phenotypes. If populations that are experiencing either divergent or uniform selection maintain geographic isolation, thus restricting gene flow between them, divergence will proceed and the populations can become reproductively isolated over time. In either case, the rate of divergence will be contingent upon the rate of migration, the strength of selection, and the initial allele frequencies. It is widely assumed that reproductive isolation can result as a by-product of such divergence in allopatry (Schluter 2001, Wu 2001, Sobel et al. 2010). This idea has rarely been tested empirically (Payseur & Rieseberg 2016, but see Kilias et al. 1980, Dodd 1989, Nosil et al. 2012a, Gompert et al. 2012b, Parchman et al. 2013, Janoušek et al. 2015, Schield et al. 2017).

In sympatric populations, selection can directly favor reproductive isolation. This can occur in hybrid zones when hybrids are less fit than parental types and as a consequence, individuals who mate with conspecifics have greater reproductive success than individuals who mate with heterospecifics (Butlin 1987). This process, called reinforcement, has traditionally been considered to result in a strengthening of prezygotic barriers that prevent the formation of unfit hybrids (Butlin 1987, Servedio & Noor 2003). However, it has recently been extended to include the evolution of any additional barrier effect in sympatry, including postzygotic isolation, as a form of adaptive coupling of reproductive barriers (Butlin & Smadja 2018). The frequency at which reinforcement occurs and its importance in shaping species diversity are open questions (Servedio & Noor 2003, Servedio 2004).

Hybrid zones offer a unique opportunity to test hypotheses about the contribution of different forms of selection that shape reproductive isolation over the course of the speciation process (e.g., Nosil et al. 2012b). They are particularly suited to empirical investigation of the genetic basis of reproductive isolation as population genetic data can be used to infer differential patterns of introgression. Barrier loci (i.e. loci associated with reproductive isolation) should have a signature of reduced introgression relative to the neutral expectation, which is caused by limited gene flow as a consequence of selection against hybrids (Barton & Hewitt 1985, Gompert & Buerkle 2011a). If the genetic differences that contribute to reproductive isolation in the hybrid

zone involve loci under selection in allopatric parental populations, we should expect to see higher differentiation in allopatric populations for barrier loci compared to neutral markers (Payseur & Rieseberg 2016). If reinforcing selection shaped barrier loci in the hybrid zone, we should expect to see greater differentiation in sympatry than in allopatry for these markers (e.g., Nosil et al. 2012b, Wang et al. 2014).

Here, we examined locus-specific differentiation and introgression using reduced-representation sequencing data from a bimodal howler monkey hybrid zone (*Alouatta palliata* x *A. pigra*) (Cortés-Ortiz et al. 2015) and from allopatric parental populations to test predictions about the forms of selection acting on loci associated with reproductive isolation. The parental species diverged ~3 MA (Cortés-Ortiz et al. 2003) and have many important differences in their morphology (Smith 1970, Kelaita et al. 2011), cytogenetics (Steinberg et al. 2008), social systems (Chapman & Balcomb 1998, Ho et al. 2014), and loud vocalizations (Bergman et al. 2016). Throughout most of their ranges, *A. palliata* and *A. pigra* are allopatric, but their ranges overlap in a narrow contact zone (~20 km, Cortés-Ortiz et al. 2007, Cortés-Ortiz & Nidiffer et al. 2018) in Tabasco, Mexico (Figure 1). It is likely that the contact zone in Tabasco is the result of secondary contact between the parental species after periods of isolation and range expansion (Cortés-Ortiz et al. 2003, Ford 2006, Ellsworth & Hoelzer 2006). Despite the relatively large degree of divergence between *A. palliata* and *A. pigra*, reproductive isolation is incomplete, as hybridization has been confirmed in the contact zone using molecular markers. Initial surveys showed that multigenerational backcrossed hybrids into each parental species are nearly equally abundant, there are few intermediate hybrids, and no putative F1s (Cortés-Ortiz et al. 2007; Kelaita & Cortés-Ortiz 2013). We have previously shown that there is a lack of introgression for *SRY* (the Y-linked sex determination gene), suggesting that F1 males may be infertile or inviable (Cortés-Ortiz & Nidiffer et al. 2018). Consistent with this, anecdotal evidence suggests that there may be a cost to hybridization as a previously identified intermediate hybrid male did not produce offspring despite living as the only adult male in a group with two reproductively mature females for a period of seven years (LCO personal observation). We also found reduced introgression for X-linked markers (Cortés-Ortiz & Nidiffer et al. 2018), consistent with the "large X effect", which suggests the X chromosome plays a disproportionate role in speciation (Coyne & Orr 1989).

For this study, we used ddRADseq data to assess the extent of genomic admixture and the distribution of admixed genotypes across the *Alouatta* contact zone and identified loci that show a pattern of reduced introgression (the pattern expected for loci associated with reproductive isolation) relative to the genomic background, which is assumed to be mostly neutral. By exploring the relationship between locus-specific differentiation and introgression in the hybrid zone, we tested the hypothesis that reproductive isolation results as a by-product of divergence in allopatry and that reinforcing selection in sympatry shapes reproductive barriers. We also performed functional annotation for loci that showed the strongest evidence for divergent selection in sympatry to associate these regions with putative phenotypes under reinforcement and evaluated the pattern of introgression for putatively X-linked markers to test for a large X effect. Our results are consistent with signatures of divergence in allopatry and reinforcement in sympatry, indicating that multiple forms of selection have shaped the evolution of reproductive isolation in this system.

Materials and Methods

Sampling

Our sampling included 181 wild individuals captured between 1998 and 2012 following procedures described in Kelaita et al. (2011, 2013) and adhered to the University of Michigan's Institutional Animal Care and Use Program standards (UCUCA permit #09319). Individuals were chosen from a larger pool of previously collected samples to maximize the geographic distribution of *Alouatta* in Mexico, as well as the representation of admixed genotypes present in the hybrid zone (i.e., individuals backcrossed into *A. palliata*, intermediate hybrids, individuals backcrossed into *A. pigra*) as determined by hybrid index measured from 24 microsatellite markers (Cortés-Ortiz & Nidiffer et al. 2018). Thus, our samples included 99 individuals from the hybrid zone in Tabasco, Mexico (Tab), 38 allopatric *A. palliata* from Veracruz, Mexico (Ver), and 44 allopatric *A. pigra* including 24 from Campeche, Mexico (Cam), 12 from Quintana Roo, Mexico (QR), and 8 from Dolores, Guatemala (DG) (Figure 1).

Samples were kept on ice in the field and stored at -20°C upon arrival in the laboratory. Genomic DNA was extracted with the QIAGEN DNeasy tissue kit (Qiagen Inc., Valencia, CA) following the manufacturer's protocol for animal tissue extractions with the following

modifications: 1) we added 100 µl of whole blood in lysis buffer solution (1:5 concentration) to 100 µl buffer of ATL, 2) we eluted DNA in 70 µl of water at 55 °C twice (re-using the same spin column) to maximize DNA yields, after incubating for 5 minutes at room temperature.

ddRADseq and genotyping

We prepared and sequenced four ddRAD libraries, following the Peterson et al. (2012) protocol, each library containing DNA from 48 individuals (two libraries also included individuals sequenced for use in other projects). Briefly, we used the restriction enzymes SphI and MluCI to digest 200–300ng DNA per sample, size selected fragments between 150–350bp using a 2% Pippen Prep gel (Sage Science, Beverly, MA), and sequenced libraries on an Illumina HiSeq 4000 machine at the University of Michigan Sequencing Core to obtain 150bp paired-end reads.

We demultiplexed our data using *pyRAD* (Eaton et al. 2014), merged read pairs that overlapped using *FLASH* (Magoč & Salzberg 2011), and aligned both successfully merged reads and unmerged reads (which were expected due to our size selection window) to the draft *Alouatta palliata* genome assembly (accession ID PVKV00000000) with *BWA-MEM* (Li 2013) using default settings. This genome assembly was provided by the 200 Mammals Project of the Broad Institute and Uppsala University (in press). We then called variants and generated a VCF file using *samtools mpileup* (including options -u -g -t DP, DPR) and *bcftools call* (options -v -m -O v) (Li et al. 2009). After removing SNPs within 5bp of an indel and retaining variants with a minimum quality score of 20, we obtained 6,415,368 loci (including SNPs and indels) that were subsequently filtered in further analyses (Table S1).

Admixture and population structure

We first used *fastStructure* (Raj et al. 2014) to quantify admixture proportions and to assign admixed (hybrid) or non-admixed status to individuals. Because we were interested in detecting hybrid individuals, we ran ten replicates using the simple prior to infer admixture proportions (Q) with the number of clusters (K) equal to two, reflecting the parental species. Here, admixture proportions are an estimate of the proportion of the genome inherited from each parental species (i.e., where Q_1 is the proportion from parental species 1 and Q_2 is the proportion

179 from parental species 2 and $Q_1 + Q_2 = 1$). For simplicity, we only report Q_1 as the proportion of
180 the genome inherited from *A. pigra*.
181 Because our sampling sites are geographically widespread and there may be some within-
182 population structuring, we ran an additional ten replicates for each K between 3 to 8 to ensure
183 that imposing K=2 on this system did not affect our ability to detect hybrids. We used the
184 *fastStructure* script 'chooseK.py' to detect the number of clusters that best fit our data. We also
185 examined the correlation between admixture proportion Q_1 and hybrid index as calculated by *bgc*
186 (see below) for hybrid zone individuals using the Pearson method in R.
187 We limited our *fastStructure* analysis to biallelic loci (those with different alleles between
188 species or intraspecific polymorphism) with low missing data across individuals. To do this, we
189 used *bcftools* and *vcftools* (Danecek et al. 2011) to filter loci by removing indels, non-biallelic
190 sites, sites with a minor alleles frequency of ≤ 0.01, and sites with a minimum mean depth across
191 individuals of less than 10. Because we did not apply a depth-per-read threshold for calling
192 genotypes, we discarded sites with missing genotypes in more than 50% of individuals to only
193 include sites with high read depth. To reduce effects of linkage and comply with assumptions of
194 the *fastStructure* model-based approach, we also thinned sites within 200bp of each other and
195 discarded sites out of Hardy-Weinberg equilibrium in either of the allopatric parental populations.
196 This resulted in 74,448 SNPs, which we included in the *fastStructure* analysis (Table S1). Using
197 this dataset, we dropped 23 individuals from further analyses due to a high frequency (>80%) of
198 missing genotype data, which can affect confidence in *fastStructure* results. In further analyses,
199 we identified hybrids as individuals with $0.05 < Q_1 > 0.95$, non-admixed *A. palliata* individuals
200 as $Q_1 < 0.05$, and non-admixed *A. pigra* individuals as individuals with $Q_1 > 0.95$. With this
201 dataset, we also visualized structure among sampling sites using principle component analysis
202 (PCA) implemented in *SNPRelate* (Zheng et al. 2012).
203 Since our fastStructure analyses identified individuals in allopatric populations with some
204 level of admixture (N = 11, Table S2), we dropped these individuals from further analyses to
205 avoid complications from including admixed individuals outside the hybrid zone in
206 differentiation and introgression analyses. This reduced our dataset to include 81 individuals
207 from the hybrid zone (Tabasco), 32 allopatric *A. pigra* individuals, and 34 allopatric *A. palliata*
208 individuals (Table S1). All further analyses only included these individuals.

Genomic cline analysis

For our analyses of genomic clines and genetic differentiation, we filtered loci to increase confidence in genotype calls and to maximize information about ancestry. To do this, we retained biallelic loci with a minimum mean depth across individuals of 30, loci with a minor allele frequency of ≥ 0.05, and loci that were present in at least 80% of individuals in either parental population. This resulted in 5,763 loci (Table S1) distributed on 2,883 contigs between 80.2 Kb–1.28 Mb in size (representing 18.8% of the total reference assembly).

To quantify introgression across loci and identify candidate variants with evidence for an association with reproductive isolation (i.e., reduced introgression compared to neutral expectations), we used genomic cline analysis implemented in *bgc* (Gompert & Buerkle 2011a, Gompert & Buerkle 2012a). Genomic cline analysis uses differential introgression to identify loci that are more or less likely than the genome-wide average (assumed to be neutral) to introgress between populations. For each locus, *bgc* uses the cline parameter β to quantify the amount of introgression, with $\beta < 0$ indicating greater than expected introgression and $\beta > 0$ indicating reduced introgression with respect to the genome-wide average. Loci showing evidence for an association with reproductive isolation (barrier loci) are expected to have reduced introgression ($\beta > 0$) due to selection against hybrids.

We ran genomic cline analyses under the genotype uncertainty model (appropriate for next-generation sequence data, Gompert et al. 2012b) in *bgc* for five independent chains, each with a burn-in of 30,000 for 50,000 steps, and thinned samples by 20. We then merged outputs and identified β outliers from MCMC output as loci with a 95% credible interval that does not overlap zero.

Identifying putative X chromosome markers

To allow us to test for restricted introgression of the X chromosome, we used NCBI BLASTN (Altschul et al. 1997) to associate *Alouatta* contigs with genes on the X chromosome in humans. We expect genes on the human X to be X-linked in *Alouatta* since gene content and order is highly conserved across mammals (Delgado et al. 2009), and the *Alouatta* X appears to be highly similar to the human X (Steinberg et al. 2014). We downloaded unique, unspliced

human X chromosome genes (GRCh38.p12) from Ensembl (Zerbino et al. 2017) using the online BioMart tool. We assumed any locus to be putatively X-linked if any of the human X gene sequences had a BLASTN best hit (–max_hsps 1 –max_target_seqs 1, otherwise default settings) to the same *Alouatta* contig and the percent identity of the aligned sequence was >85%. Using these criteria, we identified 191 putatively X-linked loci on 90 contigs in our data set. To determine the pattern of introgression for the X chromosome, we visually inspected the pattern of introgression for this subset of putatively X-linked loci.

Genetic differentiation

Loci influenced by divergent selection are expected to show elevated differentiation (Beaumont & Balding 2004, Gompert & Buerkle 2011b). We measured locus-specific differentiation with the method of Weir & Cockerham (1984) implemented in *vcftools* (Danecek et al. 2011) using the same dataset as for our cline analysis. For allopatric parental populations, we tested for differences in the distribution of F_{ST} values between loci that showed reduced introgression in the hybrid zone, those with neutral introgression, and those with increased introgression using ANOVA and detected pair-wise differences between each category using the Tukey Honest Significant Difference method, both implemented in base R v3.4.1 (R Core Team 2017). For analyses of differentiation in sympatry, we calculated F_{ST} between *A. palliata*-like (mean $Q_1 < 0.5$, N = 54) and *A. pigra*-like individuals (mean $Q_1 > 0.5$, N = 27) in the hybrid zone. We included the same set of loci for each comparison, but when we separated the samples into allopatric and sympatric populations, some loci were no longer polymorphic. Thus, for comparisons where sites were monomorphic and at sites where there was more variation within than between populations ($F_{ST} < 0$), we report $F_{ST} = 0$. We compared the distribution means of F_{ST} in sympatry and F_{ST} in allopatry for loci with reduced and neutral introgression using a Wilcoxon Rank Sum test and by fitting a linear model to the relationship between these two variables in R. If reinforcing selection contributed to reproductive isolation in sympatry, we would expect loci to have higher F_{ST} in sympatry than in allopatry and to see a relationship between the two variables that differs from a 1:1 linear relationship, which would be assumed if divergence in sympatry is equal to divergence in allopatry (i.e., no reinforcement).

Genomic basis of reinforcement

To explore potential functions of loci with reduced introgression that showed strong evidence for divergent selection in sympatry, we identified homologous human protein-coding genes near these loci. We first identified the set of loci with reduced introgression ($\beta > 0$) that had greater F_{ST} in sympatry than in allopatry ($F_{STsympatry} - F_{STallopatry} > 0$, N=104), i.e., those with a signature of reinforcement. To focus our search on regions showing only evidence for reinforcement and not divergence in allopatry or adaptive introgression, we excluded contigs that also contained $\beta > 0$ loci with greater F_{ST} in allopatry than in sympatry and contigs that also contained $\beta < 0$ outliers. This resulted in 93 loci on 79 contigs. We then ranked loci by the difference in $F_{STsympatry} - F_{STallopatry}$ and took the top 10% of loci with the greatest difference as candidate loci showing strong evidence for selection in sympatry. For each locus, we extracted the entire contig from the *Alouatta* genome assembly and used the UCSC genome browser online BLAT tool (Kent 2002) to identify its position in the human genome (version GRCh38/hg38). We BLAT searched the first and last 25kb of sequence separately for each *Alouatta* contig and took the outermost coordinates from each alignment so that we could identify human genes that occur between regions for each contig (Table S3). For each region, we ensured that alignment orientation, length, and span of human genomic positions were consistent with each *Alouatta* contig, suggesting that the human genomic regions are collinear and we can assume these genes are also present on the *Alouatta* contigs. We then used biomaRt (Durnick et al. 2005, 2009) to obtain all human protein coding genes within each region. Finally, we identified mammalian phenotypes associated with each gene using the Mouse Genome Informatics (MGI) batch query tool online (URL: http://www.informatics.jax.org, accessed May 2018).

Results

Structure and admixture

Across ten replicate *fastStructure* runs for each K between 2–8, maximum likelihood scores were highest for K=2 (Figure S1A). Model complexity that maximizes marginal likelihood was equal to two in each replicate, and the number of model components used to explain structure in the data was equal to two in four replicates and equal to three in six replicates (Figure S1B). Admixture proportions using K=2 and K=3 were very similar due to extremely low

assignment values for each individual to the third cluster (each $Q_3 < 0.0001$) in nine of 10 K=3 replicates (Figure S2). Further, hybrid index scores inferred from *bgc* were closely correlated with *fastStructure*'s Q_1 at K=2 ($r = 0.996$, $P < 2.2X10^{-16}$) (Figure 2B). Together, these results indicate that K=2 best describes our data and that our use of admixture proportion Q_1 was appropriate in assigning hybrid status to individuals. Thus, we report mean Q_1 scores across our ten K=2 replicates (Figure 2, Table S4).

Concordant with our previous analyses using microsatellite markers (Cortés-Ortiz & Nidiffer et al. 2018), our *fastStructure* analysis at K=2 shows that most individuals in the contact zone are multigenerational backcrosses to either parental species and there are few intermediate hybrids (Figure 2). Out of 81 individuals, we identified five with an intermediate admixture proportion ($0.4 > Q_1 < 0.6$). Although admixture is mainly restricted to the contact zone, several individuals in Campeche also appear to be admixed, along with a single individual in Quintana Roo and two individuals in Veracruz (Figure 2, Table S4). In Campeche and Quintana Roo, most admixed individuals are *A. pigra*-like ($Q_1 > 0.6$), concordant with the geographic range of the parental species that inhabits those locations. Similarly, the admixed individuals in Veracruz have predominantly *A. palliata* ancestry ($Q_1 < 0.2$).

The PCA results are largely concordant with the *fastStructure* K=2 analysis, suggesting that the set of variants used to detect hybrids robustly discriminates the parental species from each other and from hybrids (Figure 3A). PC 1 explains 55% of the genetic variation among individuals and clearly separates allopatric populations (with the exception of admixed individuals detected outside the contact zone). Thus, not surprisingly, PC 1 is strongly correlated with the *fastStructure* admixture proportion Q_1 ($r = 0.98$, $P < 2.2 \times 10^{-16}$, Figure 3B). PC 2 explains 2.4% of the genetic variation among individuals and seems to primarily be associated with population structure among sampling sites within *A. pigra*.

Differential introgression across loci

We found a small percentage of loci that were β outliers (Figure 4) consistent with non-neutral introgression. There were 255 loci (4.4%) that showed reduced introgression ($\beta > 0$) distributed on 206 contigs (1.2 loci/contig) and 319 loci (5.5%) with increased introgression ($\beta < 0$) distributed on 248 contigs (1.3 loci/contig). Only six contigs had loci with both reduced and

increased introgression. The remaining 5,189 loci (90%) were consistent with neutral introgression ($\beta = 0$). Of the 191 putatively X-linked loci, 183 (96%) had neutral introgression (Figure S3). Five loci showed reduced introgression, three of which were tightly linked on the same contig (within 12bp). Three loci showed increased introgression, all of which were on different contigs.

Genetic differentiation and its relationship with introgression

Locus-specific genetic differentiation between allopatric parental species was high overall (mean $F_{ST} = 0.65$) and ranged from 0–1 with a seemingly bimodal distribution with peaks near $F_{ST} = 0$ and $F_{ST} = 0.9$ (Figure 5A). Of the 5,763 loci analyzed, 117 had fixed differences ($F_{ST} = 1$) between allopatric parental species. Overall, differentiation was positively correlated with the amount of introgression (β) in the hybrid zone, but the relationship was weak ($r = 0.08$, $P = 6.21 \times 10^{-10}$).

Mean F_{ST} for allopatric parental populations was not equal among β categories ($F = 85.93$, $P < 2.2 \times 10^{-16}$). Post hoc comparisons indicated that the distributions of F_{ST} within each β category were significantly different from each other (Figure 5B, Table S5), with $\beta > 0$ loci having the highest F_{ST} (mean=0.85, range=0.31–1), and loci with $\beta = 0$ having the lowest F_{ST} (mean=0.63, range=0–1). $\beta < 0$ loci had an intermediate F_{ST} (mean=0.78, range=0–1).

Comparison of differentiation in sympatry and allopatry

Genetic differentiation across loci was lower in sympatry (mean $F_{STsympatry} = 0.55$) than in allopatry (mean $F_{STallopatry} = 0.65$). When loci are partitioned across β categories, F_{ST} was significantly higher in allopatry than in sympatry for markers with neutral and increased introgression ($\beta = 0$: mean $F_{STsympatry} = 0.544$, mean $F_{STallopatry} = 0.628$, $P < 2.2 \times 10^{-16}$, $\beta < 0$: mean $F_{STsympatry} = 0.373$, mean $F_{STallopatry} = 0.778$, $P < 2.2 \times 10^{-16}$). However, we found that for loci with reduced introgression ($\beta > 0$), F_{ST} was significantly higher in sympatry than in allopatry (mean $F_{STsympatry} = 0.852$, mean $F_{STallopatry} = 0.850$, $P = 6.37 \times 10^{-6}$). Although the magnitude of the difference is small, this pattern seems to be driven by loci with intermediate differentiation in allopatry having higher differentiation in sympatry, while loci with high differentiation in allopatry tended to also have high differentiation in sympatry (Figure 6A). The fit of linear

359 models to the data in each beta category showed that confidence intervals for the slope of the line
360 did not encompass one (Table 1), but was closer to one for loci with neutral introgression (Figure
361 6B).
362 Nine loci were included in the top 10% of $\beta > 0$ loci that showed the greatest difference in
363 F_{ST} between sympatry and allopatry (Table S6). We identified regions of human chromosomes 3,
364 4, 7, 8, 11, and 16 that seem to be homologous with the *Alouatta* contigs containing these loci.
365 The human regions contained 42 protein-coding genes, of which 28 could be associated with 420
366 mammalian phenotypes (MPs) in the MGI database (Table S6). Notably, several genes were
367 associated with behavior (*SCARB2*, *BRPF1*, *SLC5A2*, *KMT2A*), abnormal embryonic/fetal
368 development or lethality (*SHROOM3*, *BRPF1*, *CRELD1*, *TADA3*, *ARL13B*, *PROS1*, *KMT2A*),
369 hair texture (*ARPC4*, *KMT2A*), facial morphology (*SHROOM3*, *CRELD1*, *KMT2A*), and the
370 immune system (*ITGAD*, *ITGAX*, *CD3C*, *CD3E*, *CD3G*, *KMT2A*).
371
372 **Discussion**
373 We used reduced-representation sequencing to examine admixture, population structure,
374 introgression, and its relationship with locus-specific differentiation in a natural primate hybrid
375 zone system. Our results are consistent with the hypothesis that reproductive isolation results as a
376 byproduct of divergence in allopatry and we detected a genomic signature of reinforcement in
377 sympatry, indicating that multiple forms of selection have shaped speciation in this system.
378
379 *Admixture and population structure*
380 We found a bimodal distribution of admixture proportions in the hybrid zone. Early
381 generation hybrids are rare and multi-generational backcrosses dominate. This pattern is largely
382 consistent with our previous analyses using a small set of microsatellite markers (Cortés-Ortiz &
383 Nidiffer et al. 2018). However, we detected admixture in areas where the parental species are
384 thought to be allopatric. We detected a few admixed individuals east of the contact zone in
385 Campeche and Quintana Roo and west of the contact zone in Veracruz. In addition to autosomal
386 markers, we previously amplified a Y-linked (*SRY*) locus, X-linked loci including the
387 microsatellite locus HAM80, as well as the mitochondrial control region for most individuals
388 sequenced in the study (Table S2). Considering the sex-linked genotypes for these individuals

together with the admixture proportions calculated in this study, it is clear that these individuals are not F1 hybrids. Due to the apparent absence of non-admixed individuals of the opposite species in these areas, and their distance from the contact zone (~200km or greater), we suspect that the presence of admixed individuals in these regions is likely due to either long distance migration from the contact zone, movement of animals by humans, or to past introgression during a period when the contact zone occurred in a different location than in present day and has since shifted. It will be possible to test the hypothesis that the hybrid zone has moved by looking at linkage disequilibrium (LD) in a geographic transect across the hybrid zone, with the expectation that LD will increase in the direction of movement due to decay over time as a result of recombination (e.g., Wang et al. 2011).

Introgression in the hybrid zone

We found evidence for differential introgression in the hybrid zone. The majority of loci exhibited neutral introgression, but a small percentage of markers showed extreme introgression (Figure 4). We were particularly interested in loci with reduced introgression ($\beta > 0$) as this pattern is expected of loci associated with reproductive isolation. We identified 255 such loci. These loci were distributed on 206 contigs, which may support the hypothesis that reproductive isolation has a genome-wide basis (Parchman et al. 2013, Scordato et al. 2017). However, because the *A. palliata* genome is not assembled to chromosome-level, it is possible that these contigs may be physically linked.

Point estimates of β were much less variable for loci with a pattern of reduced introgression ($\beta > 0$) than for loci with a pattern of increased introgression ($\beta < 0$), particularly for β outliers (Figure 4). These coincident $\beta > 0$ clines may be a reflection of the coupling of multiple barrier effects in the hybrid zone (Butlin & Smadja 2018). Recent admixture between divergent populations causes correlations between linked loci that persist over many generations (Stephens et al. 1994, Verardi et al. 2006), so the effects of indirect selection on loci near barrier loci can be strong in hybrid zones. The *Alouatta* hybrid zone is bimodal (Cortés-Ortiz & Nidiffer et al. 2018) and differentiation is high between the parental species (Figure 5A), so admixture linkage disequilibrium is likely high. Thus, strong barrier effects may influence the whole

genome via indirect selection (e.g. Szymura & Barton 1991), making it difficult to identify loci underlying individual barrier effects (Butlin & Smadja 2018).

In a previous analysis, we found a complete lack of introgression for a Y-linked marker and limited to no introgression for three X-linked markers (Cortés Ortiz & Nidiffer et al. 2018), consistent with other studies suggesting the sex chromosomes may play an important role in reproductive isolation (Tucker et al. 1992, Masly & Presgraves 2007). We expanded upon these results by identifying putatively X-linked markers in our reduced-representation dataset based on sequence homology to known X-linked human genes. Of 90 putatively X-linked contigs, three had loci that showed reduced introgression and three had loci that showed increased introgression, while the remaining contigs had loci with neutral introgression (Figure S3). These results may indicate that few regions of the X chromosome underlie reproductive isolation in this system, although our interpretation may have limitations. First, we may have excluded many loci as putatively X-linked due to divergence between the *Alouatta* assembly and human gene sequences. However, the mammalian X chromosome is known to be highly conserved across mammals (Delgado et al. 2009, Mueller et al. 2013), only 146 of 2,367 human X genes did not have a match in the *Alouatta* genome assembly, and percent identity was generally high across BLASTN hits (mean = 89.5%). Second, although the X is highly conserved, New World primates are known to have a high rate of chromosomal rearrangements (de Oliveira et al. 2012) including autosome-to-sex chromosome translocations in *A.pigra* and *A. palliata* (Solari & Rahn 2005, Steinberg et al. 2008, 2014). Thus, many loci we considered to be autosomal may be on regions translocated to the X. Validation of chromosome-linkage for assembly contigs and high density genotype data across the genome will be desirable to overcome these limitations.

Loci with reduced introgression are highly differentiated in allopatry

We found that, compared to neutral loci and loci with increased introgression, loci with reduced introgression were more highly differentiated in allopatric parental populations (Figure 5B), suggesting a role for selection in driving reproductive isolation as a by-product of divergence in allopatry. As such, it seems likely that in this system, some level of reproductive isolation was already present upon secondary contact.

Because allele frequency differences are a prerequisite for testing introgression, the amount of locus-specific introgression and differentiation may be non-independent. Simulations have shown that when overall differentiation is low ($F_{ST} < 0.1$) spurious correlations between F_{ST} and genomic cline parameters can occur in the absence of selection (Gompert et al. 2012b). However, this should not be much of an issue here since mean differentiation is relatively high (Figure 5A). Further, such an effect should shape the relationship between F_{ST} and β similarly for loci across β categories and thus would not explain why F_{ST} is greater for loci with non-neutral introgression.

Our results mirror the few studies that have examined the relationship between locus-specific differentiation and introgression in animals. In manakins (Parchman et al. 2013) and lycaenid butterflies (Gompert et al. 2012b), loci with non-neutral introgression also showed elevated differentiation in parental populations compared to neutral markers. In the house mouse hybrid zone, Janoušek et al. (2015) also observed higher differentiation for markers with reduced introgression, but contrary to our findings, loci with increased introgression showed lower differentiation compared to neutral markers. These results are consistent with the hypothesis that reproductive isolation arises as a byproduct of selection in allopatry, although we recognize that locus-specific differentiation and introgression are determined by the complex interaction of many factors and that disentangling the effects of selection from other processes can be challenging (Beaumont & Balding 2004, Gompert et al. 2012c). For instance, high differentiation can occur in regions of reduced recombination due to low levels of within-species diversity possibly confounding any signals of divergent selection (Cruickshank & Hahn 2014). Therefore, elucidating the cause of elevated differentiation for loci with reduced introgression will provide key insight on the genetics of reproductive isolation.

Although differentiation between allopatric parental populations was greatest for loci with reduced introgression, it was not extremely high for all markers with reduced introgression (Figure 5B). Similarly, we observed neutral and increased introgression for markers with very high, moderate, and no differentiation (i.e. from $F_{ST} = 0$ to $F_{ST} = 1$) (Figure 5B). These observations are similar to those in other hybrid zones (e.g., Janoušek et al. 2015, Schield et al. 2017) and are likely a reflection of the complexity of the interaction between selection, drift and recombination and suggest that these forces vary across the genome. Despite the mechanism of

divergence, this also demonstrates that high differentiation in allopatry is not a perfect predictor for reproductive isolation.

With the data presented here, it is not possible to quantify the contribution of drift to the divergence of loci associated with reproductive isolation. However, if the majority of loci associated with reproductive isolation in this system diverged via genetic drift, we might expect to see a similar distribution of locus-specific F_{ST} for loci with reduced introgression and those with neutral introgression. Instead, we observed a significantly greater mean F_{ST} for loci with reduced introgression (Figure 5B, Table S5). Further, it has been recognized that although it is theoretically possible, drift alone is unlikely to result in reproductive isolation (Turelli et al. 2001, Sobel et al. 2010). Phenotypes with the potential to be associated with reproductive isolation (e.g., sterility/fertility phenotypes) are likely to be subject to selection within species and are thus not likely to be driven to fixation by drift.

Distinguishing the effects of divergent ecological selection from drift may be useful in understanding the role the environment played in shaping these species' evolutionary history. Some have concluded, however, that ecology is rarely, if ever, divorced from speciation and that multiple mechanisms likely contribute to and interact during the speciation process (Sobel et al. 2010, Templeton 2008), and, consequently, the idea that speciation occurs by either ecological divergence or drift is a false dichotomy (Sobel et al. 2010). Regardless, for this system, it will be necessary to take into consideration the possibility that any potential environmental differences encountered during divergence may differ from those currently encountered considering the estimated divergence time of 3 MA for these species (Cortés-Ortiz et al. 2003).

Evidence for a role of reinforcement

Reinforcement enhances barriers to reproduction between species and can act to complete the speciation process when partially isolated species come into contact after experiencing some divergence in allopatry (Servedio 2004, Butlin & Smadja 2018). We tested for a signature of reinforcement by comparing locus-specific differentiation between allopatric parental populations of *A. palliata* and *A. pigra* to the differentiation between backcrossed hybrids of each parental type (i.e., *A. palliata*-like and *A. pigra*-like backcrosses) in the hybrid zone for loci with reduced introgression. If reinforcing selection shaped loci with reduced

introgression and thus contributed to reproductive isolation, we would expect to see greater differentiation in sympatry than in allopatry for these markers. Our results are consistent with this prediction.

There are at least two other mechanisms that may result in greater differentiation between sympatric than allopatric populations of the same species (e.g., Wang et al. 2014). First, strong genetic drift after independent range expansions of the parental species may result in greater differentiation in sympatry than in allopatry. However, effects of drift would be expected to have a genome-wide impact (e.g., Li et al. 2008), and we only observed a pattern of overall elevated differentiation in sympatry for loci with reduced introgression, which are expected to be associated with reproductive isolation. Second, it is plausible that greater differentiation in sympatry could result indirectly from independent local adaptation within each species to sites within their allopatric and sympatric ranges (i.e., mutation-order effects, Schluter 2009). However, the divergent alleles that underlie local adaptation in the hybrid zone may be expected to have neutral or increased introgression because such alleles should be advantageous on either species' genomic background (barring any involvement in hybrid incompatibilities). Thus, it seems likely our results reflect selection to increase reproductive isolation in sympatry under the extended view of reinforcement (Butlin & Smadja 2018). However, we still need to investigate if the loci driving this pattern underlie phenotypes under reinforcing selection.

We investigated mammalian phenotypes associated with genes occurring on contigs of the *Alouatta* genome containing loci with reduced introgression that represented the top 10% of those with the greatest difference between F_{ST} in sympatry and F_{ST} in allopatry. We found that some genes in these regions have been linked to mammalian phenotypes that could conceivably be under selection for prezygotic or postzygotic isolation in the hybrid zone (Table S6), thus contributing to the extended view of reinforcement (Butlin & Smadja 2018). For example, several genes are associated with the phenotype "abnormal behavior" (MP:0004924). In mice, one of these genes, the histone methyltransferase *KMT2A*, is known to play a role in complex behaviors in mice including anxiety, nest-building behavior, spatial working memory, and learning (Gupta et al. 2010, Jakovcevski et al. 2015). In several taxa, learning is known to play a role in mate choice (e.g., sexual imprinting, learned avoidance of heterospecific mates), and thus can potentially be linked to prezygotic isolation (Servedio et al. 2009, Verzijden et al. 2012,

Dukas 2013). Learning and memory have also been linked to postzygotic isolation since deficiencies in these traits can be selected against in hybrids (Rice & McQuillan 2018). Recently McQuillan et al. (2018) found that hybrid chickadees scored lower than parental chickadees in associative spatial learning and problem solving tasks. Learning and memory have been implicated in goal-oriented foraging behavior in Neotropical primates (Garber 1989, Janson 1998), traits presumably important for howler monkeys, which maintain a predominantly folivorous-frugivorous diet in highly diverse tropical forests where they adjust their dietary intake on seasonal availability of preferred foods (Raño et al. 2016). Thus, learning and memory deficiencies in hybrids, possibly mediated by *KMT2A*, may hinder foraging efforts potentially contributing to lower viability or fitness of howler monkey hybrids in their environment.

Many genes are also associated with abnormal embryonic/fetal development or lethality phenotypes (e.g., MP:0001672, MP:0011092, MP:0010865, MP:0011101). It is possible that incompatible alleles between the parental species in these genes contribute to postzygotic isolation in this system. The contig with the greatest difference in F_{ST} between sympatry and allopatry annotated using our framework contains *SHROOM3*, a gene that encodes a PDZ domain-containing protein. In mice, *SHROOM3* mutant embryos suffer severe neural tube defects resulting in perinatal death (Hildebrand & Soriano 1999). Although our functional annotation results offer some plausible genetic mechanisms that could be involved in reinforcement, they should be interpreted with caution. First, it is not known how mutations in any of these genes affect phenotypes in howler monkeys or what role these genes may play in reproductive isolation in the *Alouatta* hybrid zone. Further, it is not known whether incompatibilities associated with the genes identified here are directly driving reduced introgression of our loci, or are physically linked to causal variants not sequenced in our reduced-representation library. Similar analyses using whole genome sequence data (e.g., Rafati et al. 2018) will be a valuable step to better associate loci driving reduced introgression of genomic regions with potential functions and phenotypes under selection for pre- and postzygotic reproductive isolation in this system.

Although we did not measure prezygotic isolation with respect to any phenotype, we suspect that there are many traits beyond the phenotypes identified above that reinforcement could potentially be acting upon in this system. Specifically, traits known to be associated with mating behavior in howler monkeys (and thus may have potential involvement with prezygotic

isolation) include olfactory cues in urine and other scent markings that males likely use to detect female sexual receptivity (Glander 1980, Horwich 1983), and behavioral displays of sexual solicitation and mate guarding (Glander 1980, Horwich 1983, Van Belle et al. 2009). Color traits are known to be used in mate discrimination in other animal species (Hill 1991, Seehausen & van Alphen 1998, Jiggins et al. 2001, Waitt et al. 2003) and some have hypothesized that sexual selection on coat color has shaped female choice of mates in howler monkeys, as it may signal a male's competitive ability, health status, maturity, etc. (Crockett 1987, Bicca-Marques & Calegaro-Marques 1998). Similarly, traits that might influence the outcome of competition between males for access to females may shape the dynamics of heterospecific copulation in the hybrid zone (and thus prezygotic isolation). Such traits include body size, canine length, and testis volume (Kelaita et al. 2011), as well as the loud roaring vocalizations for which howler monkeys are known (Kowalewski & Garber 2010, Holzmann et al. 2012, Van Belle et al. 2014, Kitchen et al. 2015). In order to test any of these hypotheses, it will be necessary to quantify these traits and compare the characteristics of sympatric and allopatric individuals with the prediction that if reinforcement has occurred, trait differences between the species will be more pronounced in sympatry than in allopatry (i.e., reproductive character displacement; although reinforcement does not always produce a signature of reproductive character displacement, Servedio 2004). In this study, we did not directly measure selection against hybridization or prezygotic isolation, let alone any potential phenotypes under reinforcing selection. More research will be necessary in order to connect the pattern we observed here, greater divergence in sympatry than in allopatry for loci associated with reproductive isolation, with conclusive evidence for reinforcing selection on any mating discrimination trait. Regardless, in the genomic era, scans similar to the one employed in this study may enhance efforts to understand the frequency and importance of reinforcement in the speciation process by providing a means to detect the signature of reinforcement in the genome.

Conclusions

We identified a subset of genomic markers with reduced introgression in a natural primate hybrid zone, suggesting an association with reproductive isolation. These markers were more differentiated between allopatric parental populations than neutral loci and loci with increased

introgression, consistent with the idea that reproductive isolation is a byproduct of divergence in allopatry. These markers also showed a signature of reinforcement, suggesting that reproductive isolation may have initially been driven by divergence in allopatry, but reinforced by divergent selection in sympatry. These results reflect the contribution of different selective processes that have shaped the evolution of reproductive isolation in this system.

Acknowledgements

We thank M. R. Marchán-Rivadeneira, S. Singhal, I. Holmes, A. Haponski, L. Tran, B. Otero Jiménez, and L. Walsh for helpful discussion, suggestions, and comments. This research was supported by the National Science Foundation (DEB-0640519, BCS-0962807 & BCS-1517701 to LCO). MDB was supported in part by the University of Michigan Genetics Training Program (T32-GM07544). This work was also supported by a Grant-In-Aid of Research from the American Society of Mammalogists and a block grant from the Department of Ecology and Evolutionary Biology at the University of Michigan to MDB.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.

Barton, N. H., & Hewitt, G. M. (1985). Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**, 113–148.

Beaumont, M. A., & Balding, D. J. (2004). Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**, 969–980.

Bergman, T. J., Cortés-Ortiz, L., Dias, P. A. D., Ho, L., Adams, D., Canales-Espinosa, D., & Kitchen, D. M. (2016). Striking differences in the loud calls of howler monkey sister species (*Alouatta pigra* and *A. palliata*). *American Journal of Primatology*, **78**, 755–766.

626 Bicca-Marques, J. C., & Calegaro-Marques, C. (1998). Behavioral thermoregulation in a sexually
627 and developmentally dichromatic neotropical primate, the black-and-gold howling monkey
628 (*Alouatta caraya*). *American Journal of Physical Anthropology*, **106**, 533–546.
629
630 Butlin, R. (1987). Speciation by reinforcement. *Trends in Ecology & Evolution*, **2**, 8–13.
631
632 Butlin, R. K., & Smadja, C. M. (2018). Coupling, Reinforcement, and Speciation. *The American*
633 *Naturalist*, **191**, 155–172.
634
635 Chapman, C. A., & Balcomb, S. R. (1998). Population characteristics of howlers: ecological
636 conditions or group history. *International Journal of Primatology*, **19**, 385–403.
637
638 Cortés-Ortiz, L., Bermingham, E., Rico, C., Rodrıguez-Luna, E., Sampaio, I., & Ruiz-Garcıa, M.
639 (2003). Molecular systematics and biogeography of the Neotropical monkey genus,
640 *Alouatta*. *Molecular Phylogenetics and Evolution*, **26**, 64–81.
641
642 Cortés-Ortiz, L., Duda, T. F., Canales-Espinosa, D., García-Orduña, F., Rodríguez-Luna, E., &
643 Bermingham, E. (2007). Hybridization in large-bodied New World primates. *Genetics*, **176**,
644 2421–2425.
645
646 Cortés-Ortiz, L., Agostini, I., Aguiar, L. M., Kelaita, M., Silva, F. E., & Bicca-Marques, J. C.
647 (2015). Hybridization in howler monkeys: current understanding and future directions. In
648 Kowalewski, M., Garber, P., Cortés-Ortiz, L., Urbani, B., Youlatos, D. (Eds.), *Howler Monkeys:*
649 *Adaptive Radiation, Systematics, and Morphology* (pp. 107–131). New York, NY: Springer.
650
651 Cortés-Ortiz, L., Nidiffer, M. D., Hermida-Lagunes, J., García-Orduña, F., Rangel-Negrín, A.,
652 Kitchen, D. M., … Canales-Espinosa, D. (2018). Reduced introgression of sex chromosome
653 markers in the Mexican howler monkey (*Alouatta palliata* x *A. pigra*) hybrid zone. *International*
654 *Journal of Primatology*, https://doi.org/10.1007/s10764-018-0056-4
655

Coyne, J. A., & Orr, H. A. (1989). Two rules of speciation. In Otte, D., Endler, J. (Eds.), *Speciation and its Consequences* (pp. 180–207). Sunderland, MA: Sinauer Associates.

Crockett, C. M. (1987). Diet, dimorphism and demography: perspectives from howlers to hominids. In Kinzey, W. G. (Ed.), *The Evolution of Human Behavior: Primate Models* (pp. 115–135). Albany, NY: SUNY Press.

Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, **23**, 3133–3157.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … McVean, G. (2011). The variant call format and VCFtools. *Bioinformatics*, **27**, 156–158.

de Oliveira, E. H. C., Neusser, M., & Müller, S. (2012). Chromosome evolution in new world monkeys (Platyrrhini). *Cytogenetic and Genome Research*, **137**, 259–272.

Delgado, C. L. R., Waters, P. D., Gilbert, C., Robinson, T. J., & Graves, J. A. M. (2009). Physical mapping of the elephant X chromosome: conservation of gene order over 105 million years. *Chromosome Research*, **17**, 917–926.

Dodd, D. (1989). Reproductive isolation as a consequence of adaptive divergence in *Drosophila pseudoobscura*. *Evolution*, **43**, 1308–1311.

Dukas, R. (2013). Effects of learning on evolution: robustness, innovation and speciation. *Animal Behaviour*, **85**, 1023–1030.

Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., & Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics*, **21**, 3439–3440.

686
687 Durinck, S., Spellman, P. T., Birney, E., & Huber, W. (2009). Mapping identifiers for the
688 integration of genomic datasets with the R/Bioconductor package biomaRt. *Nature Protocols*, **4**,
689 1184–1191.
690
691 Eaton, D. A. (2014). PyRAD: assembly of de novo RADseq loci for phylogenetic
692 analyses. *Bioinformatics*, **30**, 1844–1849.
693
694 Ellsworth, J. A., & Hoelzer, G. A. (2006). Genetic evidence on the historical biogeography of
695 Central American howler monkeys. In Lehman, S. M, Fleagle, J. G. (Eds.), *Primate*
696 *Biogeography* (pp. 81–103). New York, NY: Springer.
697
698 Ford, S. M. (2006). The biogeographic history of Mesoamerican primates. In Estrada, A., Garber,
699 P. A., Pavelka, M. S. M., Luecke, L. (Eds.), *New Perspectives in the Study of Mesoamerican*
700 *Primates* (pp. 81–114). Boston, MA: Springer.
701
702 Funk, D. J., Nosil, P., & Etges, W. J. (2006). Ecological divergence exhibits consistently positive
703 associations with reproductive isolation across disparate taxa. *Proceedings of the National*
704 *academy of Sciences of the United States of America*, **103**, 3209–3213.
705
706 Garber, P. A. (1989). Role of spatial memory in primate foraging patterns: *Saguinus mystax* and
707 *Saguinus fuscicollis*. *American Journal of Primatology*, **19**, 203–216.
708
709 Glander, K. E. (1980). Reproduction and population growth in free-ranging mantled howling
710 monkeys. *American Journal of Physical Anthropology*, **53**, 25–36.
711
712 Gompert, Z., & Buerkle, C. (2011a). Bayesian estimation of genomic clines. *Molecular*
713 *Ecology*, **20**, 2111–2127.
714

715 Gompert, Z., & Buerkle, C. A. (2011b). A hierarchical Bayesian model for next-generation
716 population genomics. *Genetics*, **187**, 903–917.
717
718 Gompert, Z., & Buerkle, C. A. (2012a). bgc: Software for Bayesian estimation of genomic
719 clines. *Molecular Ecology Resources*, **12**, 1168–1176.
720
721 Gompert, Z., Lucas, L. K., Nice, C. C., Fordyce, J. A., Forister, M. L., & Buerkle, C. A. (2012b).
722 Genomic regions with a history of divergent selection affect fitness of hybrids between two
723 butterfly species. *Evolution*, **66**, 2167–2181.
724
725 Gompert, Z., Parchman, T. L., & Buerkle, C. A. (2012c). Genomics of isolation in
726 hybrids. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **367**,
727 439–450.
728
729 Gupta, S., Kim, S. Y., Artis, S., Molfese, D. L., Schumacher, A., Sweatt, J. D., ... & Lubin, F. D.
730 (2010). Histone methylation regulates memory formation. *Journal of Neuroscience*, **30**, 3589–
731 3599.
732
733 Hildebrand, J. D., & Soriano, P. (1999). Shroom, a PDZ domain–containing actin-binding
734 protein, is required for neural tube morphogenesis in mice. *Cell*, **99**, 485–497.
735
736 Hill, G. E. (1991). Plumage coloration is a sexually selected indicator of male quality. *Nature*,
737 **350**, 337–339.
738
739 Ho, L., Cortés-Ortiz, L., Dias, P. A. D., Canales-Espinosa, D., Kitchen, D. M., & Bergman, T. J.
740 (2014). Effect of ancestry on behavioral variation in two species of howler monkeys (*Alouatta*
741 *pigra* and *A. palliata*) and their hybrids. *American Journal of Primatology*, **76**, 855–867.
742

Holzmann, I., Agostini, I., & Di Bitetti, M. (2012). Roaring behavior of two syntopic howler species (*Alouatta caraya* and *A. guariba clamitans*): evidence supports the mate defense hypothesis. *International Journal of Primatology*, **33**, 338–355.

Horwich, R. H. (1983). Breeding behaviors in the black howler monkey (*Alouatta pigra*) of Belize. *Primates*, **24**, 222–230.

Janson, C. H. (1998). Experimental evidence for spatial memory in foraging wild capuchin monkeys, *Cebus apella*. *Animal Behaviour*, **55**, 1229–1243.

Jiggins, C. D., Naisbit, R. E., Coe, R. L., & Mallet, J. (2001). Reproductive isolation caused by colour pattern mimicry. *Nature*, **411**, 302–305.

Jakovcevski, M., Ruan, H., Shen, E. Y., Dincer, A., Javidfar, B., Ma, Q., ... Lin, C. L. (2015). Neuronal Kmt2a/Mll1 histone methyltransferase is essential for prefrontal synaptic plasticity and working memory. *Journal of Neuroscience*, **35**, 5097–5108.

Janoušek, V., Munclinger, P., Wang, L., Teeter, K. C., & Tucker, P. K. (2015). Functional organization of the genome may shape the species boundary in the house mouse. *Molecular Biology and Evolution*, **32**, 1208–1220.

Kelaita, M., Dias, P. A. D., Aguilar-Cucurachi, M., Canales-Espinosa, D., & Cortés-Ortiz, L. (2011). Impact of intrasexual selection on sexual dimorphism and testes size in the Mexican howler monkeys *Alouatta palliata* and *A. pigra*. *American Journal of Physical Anthropology*, **146**, 179–187.

Kelaita, M. A., & Cortés-Ortiz, L. (2013). Morphological variation of genetically confirmed *Alouatta pigra* × *A. palliata* hybrids from a natural hybrid zone in Tabasco, Mexico. *American Journal of Physical Anthropology*, **150**, 223–234.

773 Kent, W. J. (2002). BLAT—the BLAST-like alignment tool. *Genome Research*, **12**, 656–664.
774
775 Kilias, G., Alahiotis, S. N., & Pelecanos, M. (1980). A multifactorial genetic investigation of
776 speciation theory using *Drosophila melanogaster*. *Evolution*, **34**, 730–737.
777
778 Kitchen, D. M., da Cunha, R. G. T., Holzmann, I., & de Oliveira, D. A. G. (2015). Function of
779 loud calls in howler monkeys. In Kowalewski, M., Garber, P., Cortés-Ortiz, L., Urbani, B.,
780 Youlatos, D. (Eds.), *Howler Monkeys: Adaptive Radiation, Systematics, and Morphology* (pp.
781 369–399). New York, NY: Springer.
782
783 Kowalewski, M. M., & Garber, P. A. (2010). Mating promiscuity and reproductive tactics in
784 female black and gold howler monkeys (*Alouatta caraya*) inhabiting an island on the Parana
785 River, Argentina. *American Journal of Primatology*, **72**, 734–748.
786
787 Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
788 *arXiv:1303.3997*.
789
790 Li, J. Z., Absher, D. M., Tang, H., Southwick, A. M., Casto, A. M., Ramachandran, S., ... Myers,
791 R. M. (2008). Worldwide human relationships inferred from genome-wide patterns of
792 variation. *Science*, **319**, 1100–1104.
793
794 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The
795 sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
796
797 Magoč, T., & Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve
798 genome assemblies. *Bioinformatics*, **27**, 2957–2963.
799
800 Masly, J. P., & Presgraves, D. C. (2007). High-resolution genome-wide dissection of the two
801 rules of speciation in *Drosophila*. *PLoS biology*, **5**, e243.
802

McQuillan, M. A., Roth, T. C., Huynh, A. V., & Rice, A. M. (2018). Hybrid chickadees are deficient in learning and memory. *Evolution*, **72**, 1155–1164.

Mueller, J. L., Skaletsky, H., Brown, L. G., Zaghlul, S., Rock, S., Graves, T., ... Page, D. C. (2013). Independent specialization of the human and mouse X chromosomes for the male germ line. *Nature Genetics*, **45**, 1083–1087.

Nosil, P., Parchman, T. L., Feder, J. L., & Gompert, Z. (2012a). Do highly divergent loci reside in genomic regions affecting reproductive isolation? A test using next-generation sequence data in *Timema* stick insects. *BMC Evolutionary Biology*, **12**, 164.

Nosil, P., Gompert, Z., Farkas, T. E., Comeault, A. A., Feder, J. L., Buerkle, C. A., & Parchman, T. L. (2012b). Genomic consequences of multiple speciation processes in a stick insect. *Proceedings of the Royal Society B*, **279**, 5058–5065.

Parchman, T. L., Gompert, Z., Braun, M. J., Brumfield, R. T., McDonald, D. B., Uy, J. A. C., ... Buerkle, C. A. (2013). The genomic consequences of adaptive divergence and reproductive isolation between species of manakins. *Molecular Ecology*, **22**, 3304–3317.

Payseur, B. A., & Rieseberg, L. H. (2016). A genomic perspective on hybridization and speciation. *Molecular Ecology*, **25**, 2337–2360.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS One*, **7**, e37135.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rafati, N., Blanco-Aguiar, J. A., Rubin, C. J., Sayyab, S., Sabatino, S. J., Afonso, S., ... & Andersson, L. (2018). A genomic map of clinal variation across the European rabbit hybrid zone. *Molecular Ecology*, **27**, 1457–1478.

Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*, **197**, 573–589.

Raño, M., Kowalewski, M. M., Cerezo, A. M., & Garber, P. A. (2016). Determinants of daily path length in black and gold howler monkeys (*Alouatta caraya*) in northeastern Argentina. *American Journal of Primatology*, **78**, 825–837.

Rice, A. M., & McQuillan, M. A. (2018). Maladaptive learning and memory in hybrids as a reproductive isolating barrier. *Proceedings of the Royal Society B*, **285**, 20180542.

Schield, D. R., Adams, R. H., Card, D. C., Perry, B. W., Pasquesi, G. M., Jezkova, T., ... & Fujita, M. K. (2017). Insight into the roles of selection in speciation from genomic patterns of divergence and introgression in secondary contact in venomous rattlesnakes. *Ecology and Evolution*, **7**, 3951–3966.

Schluter, D. (2001). Ecology and the origin of species. *Trends in Ecology & Evolution*, **16**, 372–380.

Schluter, D. (2009). Evidence for ecological speciation and its alternative. *Science*, **323**, 737–741.

Scordato, E. S., Wilkins, M. R., Semenov, G., Rubtsov, A. S., Kane, N. C., & Safran, R. J. (2017). Genomic variation across two barn swallow hybrid zones reveals traits associated with divergence in sympatry and allopatry. *Molecular Ecology*, **26**, 5676–5691.

Seehausen, O., & van Alphen, J. J. (1998). The effect of male coloration on female mate choice in closely related Lake Victoria cichlids (*Haplochromis nyererei* complex). *Behavioral Ecology and Sociobiology*, **42**, 1–8.

Servedio, M. R., & Noor, M. A. (2003). The role of reinforcement in speciation: theory and data. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 339–364.

Servedio, M. R. (2004). The what and why of research on reinforcement. *PLoS Biology*, **2**, e420.

Servedio, M. R., Sæther, S. A., & Sætre, G. P. (2009). Reinforcement and learning. *Evolutionary Ecology*, **23**, 109–123.

Smith, J. D. (1970). The systematic status of the black howler monkey, *Alouatta pigra* Lawrence. *Journal of Mammalogy*, **51**, 358–369.

Sobel, J. M., Chen, G. F., Watt, L. R., & Schemske, D. W. (2010). The biology of speciation. *Evolution*, **64**, 295–315.

Solari, A. J., & Rahn, M. I. (2005). Fine structure and meiotic behaviour of the male multiple sex chromosomes in the genus Alouatta. *Cytogenetic and Genome Research*, **108**, 262–267.

Steinberg, E. R., Cortés-Ortiz, L., Nieves, M., Bolzán, A. D., García-Orduña, F., Hermida-Lagunes, J., ... Mudry, M. D. (2008). The karyotype of *Alouatta pigra* (Primates: Platyrrhini): mitotic and meiotic analyses. *Cytogenetic and Genome Research*, **122**, 103–109.

Steinberg, E. R., Nieves, M., & Mudry, M. D. (2014). Multiple sex chromosome systems in howler monkeys (Platyrrhini, *Alouatta*). *Comparative Cytogenetics*, **8**, 43–69.

Stephens, J. C., Briscoe, D., & O'Brien, S. J. (1994). Mapping by admixture linkage disequilibrium in human populations: limits and guidelines. *American Journal of Human Genetics*, **55**, 809–824.

Szymura, J. M., & Barton, N. H. (1991). The genetic structure of the hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*: comparisons between transects and between loci. *Evolution*, **45**, 237–261.

Templeton, A. R. (2008). The reality and importance of founder speciation in evolution. *Bioessays*, **30**, 470–479.

Tucker, P. K., Sage, R. D., Warner, J., Wilson, A. C., & Eicher, E. M. (1992). Abrupt cline for sex chromosomes in a hybrid zone between two species of mice. *Evolution*, **46**, 1146–1163.

Turelli, M., Barton, N. H., & Coyne, J. A. (2001). Theory and speciation. *Trends in Ecology & Evolution*, **16**, 330–343.

Van Belle, S., Estrada, A., Ziegler, T. E., & Strier, K. B. (2009). Sexual behavior across ovarian cycles in wild black howler monkeys (*Alouatta pigra*): male mate guarding and female mate choice. *American Journal of Primatology*, **71**, 153–164.

Van Belle, S., Estrada, A., & Garber, P. A. (2014). The function of loud calls in black howler monkeys (*Alouatta pigra*): food, mate, or infant defense? *American Journal of Primatology*, **76**, 1196–1206.

Verardi, A., Lucchini, V., & Randi, E. (2006). Detecting introgressive hybridization between free-ranging domestic dogs and wild wolves (*Canis lupus*) by admixture linkage disequilibrium analysis. *Molecular Ecology*, **15**, 2845–2855.

Verzijden, M. N., ten Cate, C., Servedio, M. R., Kozak, G. M., Boughman, J. W., & Svensson, E. I. (2012). The impact of learning on sexual selection and speciation. *Trends in Ecology & Evolution*, **27**, 511–519.

Waitt, C., Little, A. C., Wolfensohn, S., Honess, P., Brown, A. P., Buchanan-Smith, H. M., & Perrett, D. I. (2003). Evidence from rhesus macaques suggests that male coloration plays a role in female primate mate choice. *Proceedings of the Royal Society of London B: Biological Sciences*, **270**, S144–S146.

Wang, L., Luzynski, K., Pool, J. E., Janoušek, V., Dufkova, P., Vyskočilová, M. M., … Pialek, J. (2011). Measures of linkage disequilibrium among neighbouring SNPs indicate asymmetries across the house mouse hybrid zone. *Molecular Ecology*, **20**, 2985–3000.

Wang, J., Abbott, R. J., Ingvarsson, P. K., & Liu, J. (2014). Increased genetic divergence between two closely related fir species in areas of range overlap. *Ecology and Evolution*, **4**, 1019–1029.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Wu, C. I. (2001). The genic view of the process of speciation. *Journal of Evolutionary Biology*, **14**, 851–865.

Zerbino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., ... Gil, L. (2017). Ensembl 2018. *Nucleic Acids Research*, **46**, D754–D761.

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, **28**, 3326–3328.

Data Accessibility

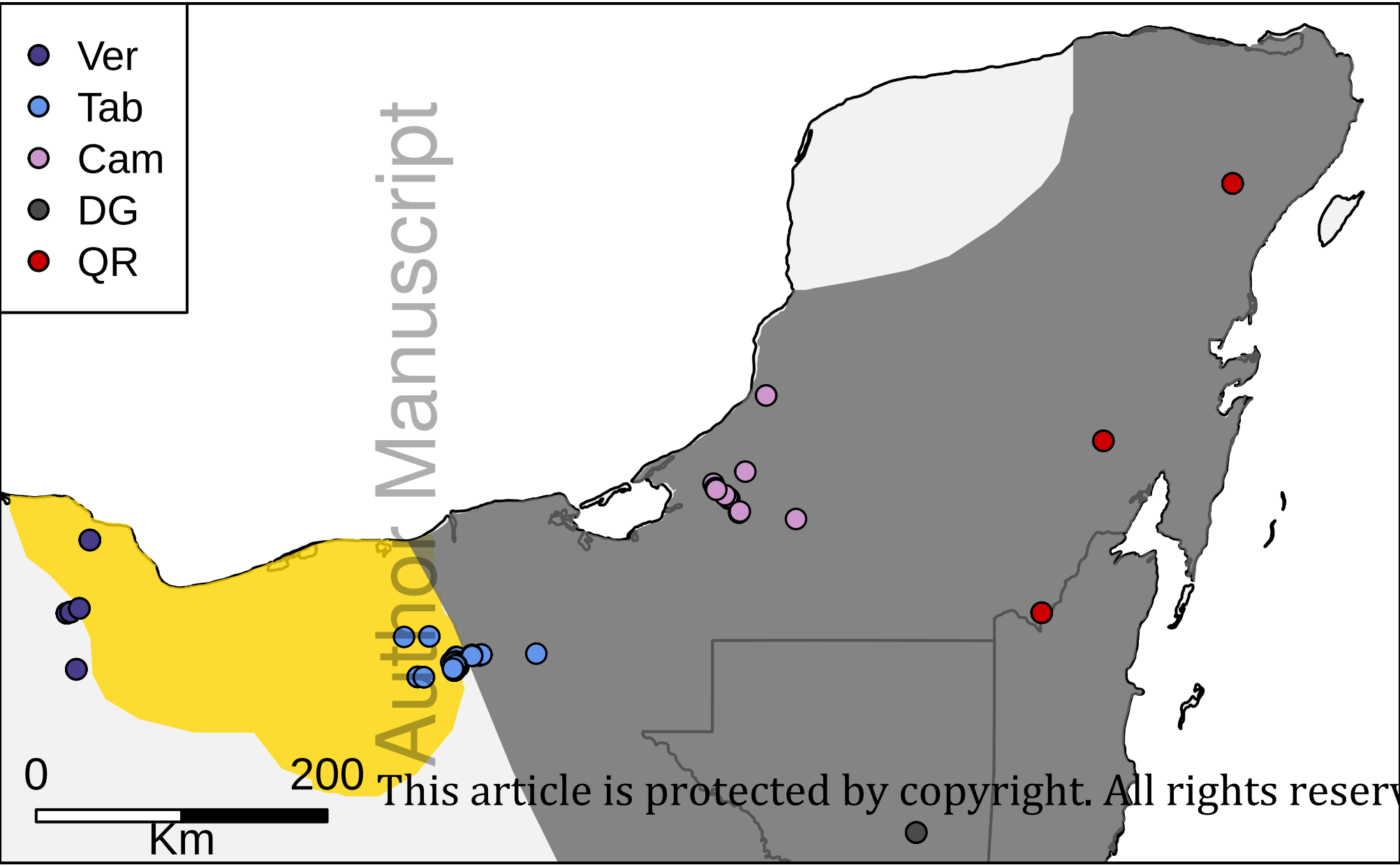
Genotype data from this study are available from the Dryad Digital Repository at: https://doi.org/10.5061/dryad.5d4mb06. Sequence data are available from the NCBI Sequence Read Archive under accession PRJNA504885.

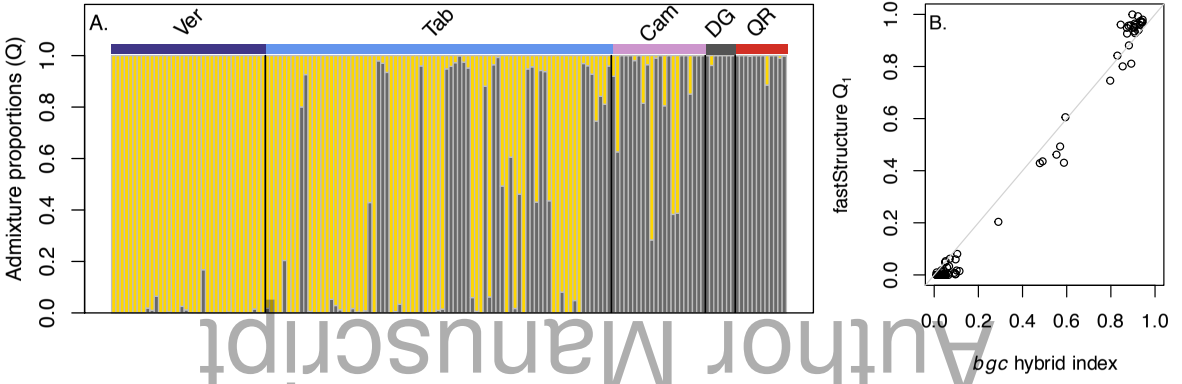
950

951 **Author Contributions**

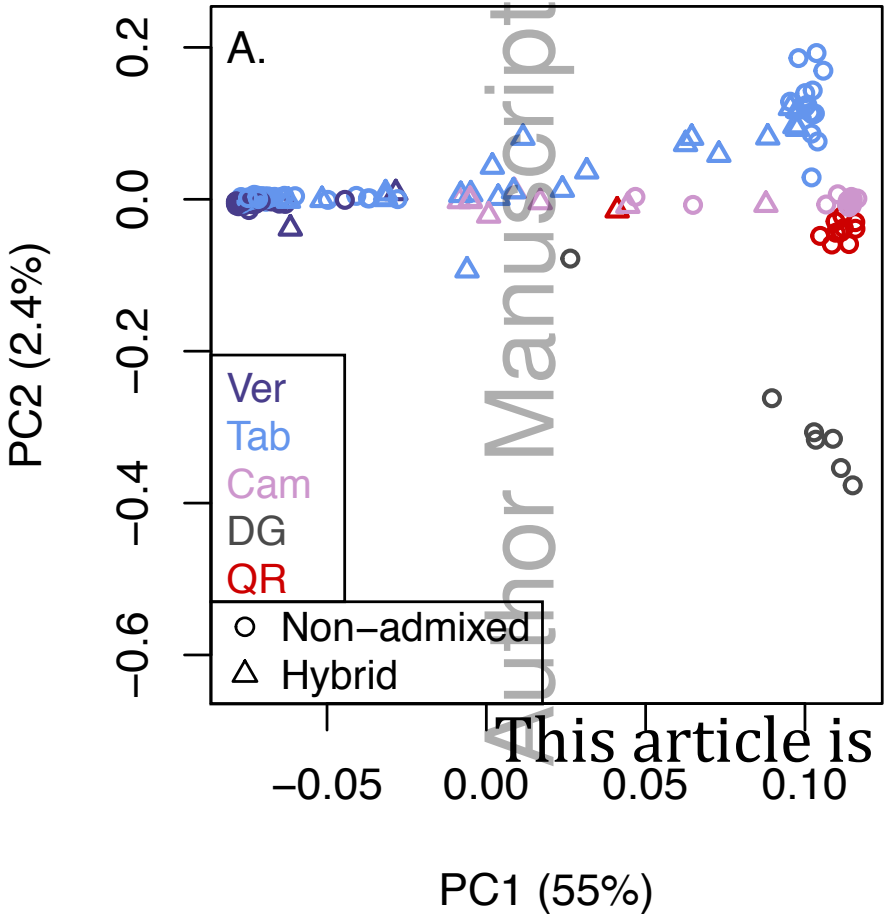
952 MDB designed the study, wrote the manuscript, performed laboratory work, analyzed data, and
953 obtained funding. PKT designed the study and wrote the manuscript. LCO obtained funding,
954 collected samples in conjunction with Mexican collaborators, designed the study, and wrote the
955 manuscript.

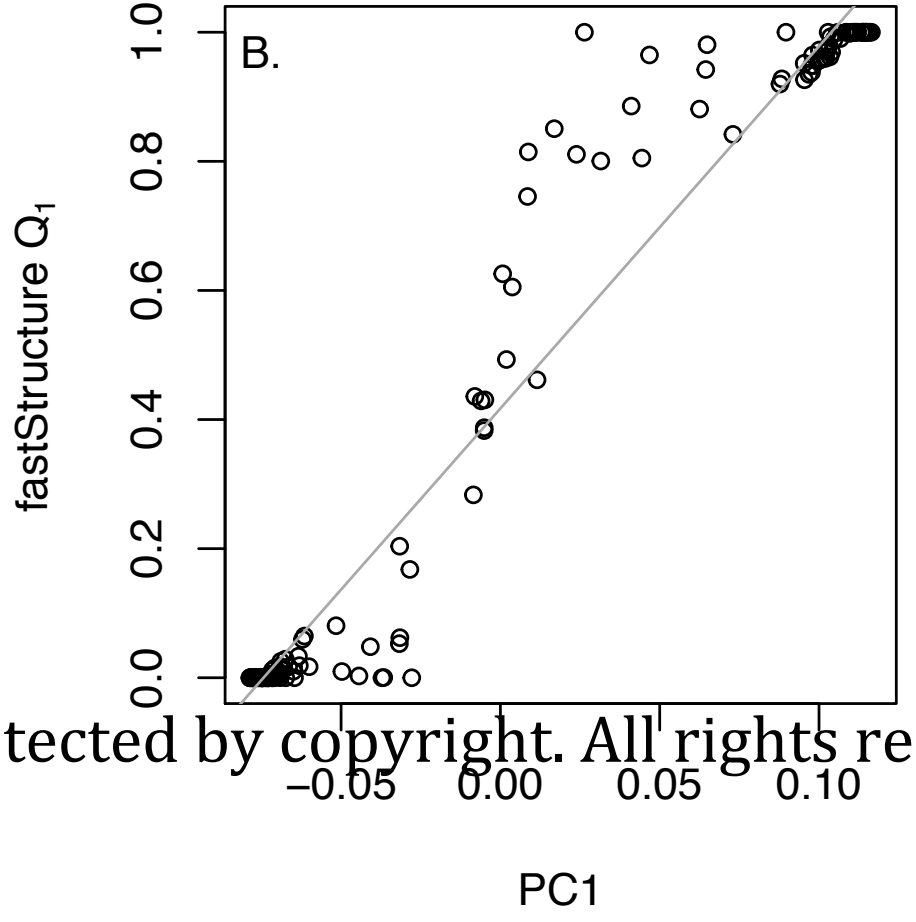
mec_14966_f1.pdf





mec_14966_f3.pdf





mec_14966_f4.pdf