

RESEARCH ARTICLE OPEN ACCESS

Patterns of Genomic Heterogeneity in a Classic Field Cricket Mosaic Hybrid Zone

Luana S. Maroja¹  | Francesca Barradale¹  | Sebastian A. Espinoza-Ulloa^{2,3,4} | Steven Bogdanowicz⁵ | Jose Andres⁵

¹Department of Biology, Williams College, Williamstown, Massachusetts, USA | ²Pontificia Universidad Católica del Ecuador, Quito, Ecuador | ³Department of Biology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada | ⁴Facultad de Medicina, Pontificia Universidad Católica del Ecuador, Quito, Ecuador | ⁵Department Ecology and Evolutionary Ecology, Cornell University, Ithaca, New York, USA

Correspondence: Luana S. Maroja (lsm1@williams.edu)

Received: 13 June 2024 | **Revised:** 6 November 2024 | **Accepted:** 11 November 2024

Funding: This work was funded by NSF DEB-1650887 to L.S.M. and S.B.

Keywords: absolute differentiation | divergence islands | field crickets | gene flow | introgression | speciation

ABSTRACT

Barriers to gene exchange can be semi-permeable; some genes are expected to freely flow across species boundaries whereas others, under divergent selection or responsible for reproductive isolation, might not. Genome scans in recently diverged species have identified divergent genomic regions, a pattern that has often been interpreted as islands of restricted introgression in a background of relatively free gene exchange (“genomic islands of speciation”). Areas of high differentiation, most located in the X chromosome (females XX, males XO), have been identified in the hybridizing field crickets *Gryllus firmus* and *Gryllus pennsylvanicus*. These species were assumed to follow an islands of speciation model, with highly differentiated areas interpreted as areas of reduced introgression. We sequenced the *G. firmus* genome to localize previously studied SNPs and sample a larger area around them in 8 allopatric populations (4 of each species). We use these data to test expectations for the islands model, in which non-introgressing areas should have both high absolute and relative differentiation. We find that in the allopatric populations, the areas with high relative differentiation (mostly X-linked), previously interpreted as non-introgressing, do not have high absolute differentiation as would be expected under the “islands model.” We also show that the estimated divergence time based on nuclear DNA is about 4× older than that estimated based on mtDNA (800 K vs. 200 K years ago). We discuss the implications of our results for introgression into allopatric populations.

1 | Introduction

Characterizing the genomic architecture of barriers to gene exchange is an essential step in understanding speciation. While rapid advances in sequencing technology have allowed us to investigate entire genomes across the speciation continuum, our understanding of the link between genome-wide patterns of divergence, phenotypes, and reproductive isolation is still quite limited. If barriers to gene exchange are a result of evolutionary forces acting on individual genes (genic view of speciation; Wu 2001) then, in species still exchanging genes, alleles that are neutral or globally advantageous will

be able to cross species boundaries, whereas genome regions under divergent selection or responsible for reproductive isolation will not. The result is that species boundaries can be semi-permeable (Barton and Hewitt 1981; Harrison 1990; Wu 2001), and in some instances, only a few loci may be essential for the maintenance of species isolation (e.g., Barr and Fishman 2010; Dasmahapatra et al. 2012; Dopman et al. 2005; Machado and Hey 2003; Nosil and Schluter 2011). These loci might act as “divergence centers” (Charlesworth, Nordborg, and Charlesworth 1997), building up additional divergence because of reduced recombination rates around them (divergence hitchhiking) and eventually leading to a clustering of

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *Ecology and Evolution* published by John Wiley & Sons Ltd.

loci contributing to differentiation (Charlesworth, Nordborg, and Charlesworth 1997; Feder and Nosil 2010; Strasburg et al. 2012; Yeaman and Whitlock 2011). Indeed, genome scans have identified divergent genomic regions in many recently diverged, hybridizing species (e.g., Carneiro et al. 2014, 2010; Hirase et al. 2021; Nadeau et al. 2012; Nosil, Funk, and Ortiz-Barrientos 2009; Via and West 2008), a pattern that has often been interpreted as islands of restricted introgression in a background of relatively free gene exchange (“genomic islands of speciation,” Malinsky et al. 2015; Sun et al. 2022; Turner, Hahn, and Nuzhdin 2005 and “genomic islands of differentiation or divergence,” Harr 2006; Nosil, Funk, and Ortiz-Barrientos 2009; Quilodran et al. 2020).

In many studies, F_{ST} , the fixation index, has been used as the measure of divergence and, in these cases, “islands of differentiation” represent genome regions with high F_{ST} (“outliers”). Explicit models of divergence with gene flow have explored how and whether divergence hitchhiking and/or “genomic hitchhiking” can produce such “ F_{ST} outliers.” The assumption is that in early divergence, gene flow homogenizes most genome regions, but regions that contain genes that contribute to local adaptation or reproductive isolation will remain differentiated (resistant to introgression). However, this “divergence with gene flow” interpretation has been challenged (Cruickshank and Hahn 2014; Nachman and Payseur 2012; Quilodran et al. 2020; White et al. 2010), because a pattern of high relative differentiation (e.g., based on allele frequency differences or F_{ST}) on a background of low differentiation can be produced even in the complete absence of gene flow. High F_{ST} island-like patterns can also result from loss of variation within species (e.g., caused by a recent selective sweep), as well as by increased divergence between species (e.g., caused by lack of gene flow and independent evolutionary trajectories). Hence, genome regions that have similar allele frequencies may reflect persistence of shared ancestral polymorphism, with divergent regions being the result of local selective sweeps.

The two alternative scenarios, differential introgression and no introgression/local selection, can be distinguished with measurements of absolute genetic divergence. In the impermeable model (no introgression), species-specific selective sweeps will convert intra-specific variation into fixed differences, increasing relative genetic differentiation. These high F_{ST} regions, however, are not expected to have high levels of absolute genetic differentiation (e.g., D_{XY}) in relation to the rest of the genome, given that all areas were already at a base level of absolute genetic divergence (dependent on time since divergence and mutation rate). Conversely, in a semi-permeable model (differential introgression), certain loci containing introgressed segments are selected against (impermeable regions), but other background loci with introgressed segments are not (permeable regions). We then expect high levels of F_{ST} to be associated with elevated D_{XY} values, since the genome-wide D_{XY} will decrease due to mixing of permeable loci. Cruickshank and Hahn (2014) have demonstrated that for many of the classic examples of heterogeneous genomic divergence, genome regions with high F_{ST} do not have correspondingly high values of D_{XY} (as would be expected in

the semi-permeable model) and thus might not represent “islands of differentiation.” Here we study a classic cricket mosaic hybrid zone, examining the question of introgression into allopatric population and whether genomic divergence patterns follow the “islands” model.

The two sister species of field crickets, *Gryllus firmus* and *G. pennsylvanicus*, form an extensive and well-characterized hybrid zone in eastern North America for which previous data based on relative genetic differentiation and patterns across the hybrid zone pointed to differential introgression (Larson, Becker, et al. 2013; Larson et al. 2012), where X-linked loci were the most “resistant to introgression” (Gainey, Kim, and Maroja 2018; Maroja et al. 2015).

While the estimated mtDNA divergence time is only about ~200 Kyears (Broughton and Harrison 2003; Maroja et al. 2009), they exhibit multiple barriers to gene exchange including habitat isolation (Larson, Becker, et al. 2013; Rand and Harrison 1989; Ross and Harrison 2002, 2006), temporal isolation (Harrison 1985), behavioral and cuticular hydrocarbon differences (Heggeseth et al. 2020; Maroja et al. 2009, 2014) and a post-mating, prezygotic barrier in which male *G. pennsylvanicus* fail to fertilize *G. firmus* eggs (Harrison 1983; Larson et al. 2012). The last barrier results in a unidirectional reproductive incompatibility (only *G. pennsylvanicus* females produce F1 hybrids). A priori, the genomic pattern of differentiation between the two species follows a pattern consistent with a genomic islands model: Most previously analyzed SNPs showing major allele frequency differences between allopatric populations and restricted introgression are concentrated on only two regions, the X chromosome (females are XX and males X0) and in a small window on one autosome (LG14) (Gainey, Kim, and Maroja 2018; Maroja et al. 2015). An observation also consistent with the X chromosome playing a prominent role in reproductive isolation (e.g., Charlesworth, Coyne, and Barton 1987; Good, Dean, and Nachman 2008; Hu and Filatov 2016; Masly and Presgraves 2007; Meisel and Connallon 2013). However, in these previous studies, the absolute divergence was not calculated for any of the SNP locations, and cline calculations were conducted prior to mapping, thus based on a diploid model (i.e., not taking into account X linkage). In this paper we calculate D_{XY} data for ~5 kbp regions around 48 of the previously characterized SNPs (Gainey, Kim, and Maroja 2018; Larson, Becker, et al. 2013; Larson et al. 2014; Maroja et al. 2015) across four allopatric populations of each species (Figure 1). We also assemble a draft genome, report genes located around these previously characterized SNPs, and recalculate divergence times using nuclear DNA data and outgroup species.

We fail to observe an elevated D_{XY} in areas of elevated F_{ST} or any higher D_{XY} for previously characterized non-introgressing loci and conclude that patterns of divergence in allopatric populations are more consistent with a lack of introgression. Furthermore, according to our genomic-based divergence calculations, these two cricket species diverged ~800 Kyears ago, substantially earlier than indicated by previous mtDNA-based estimates (~200 Kyears ago). We discuss the significance of these findings for the understanding of this hybrid zone.

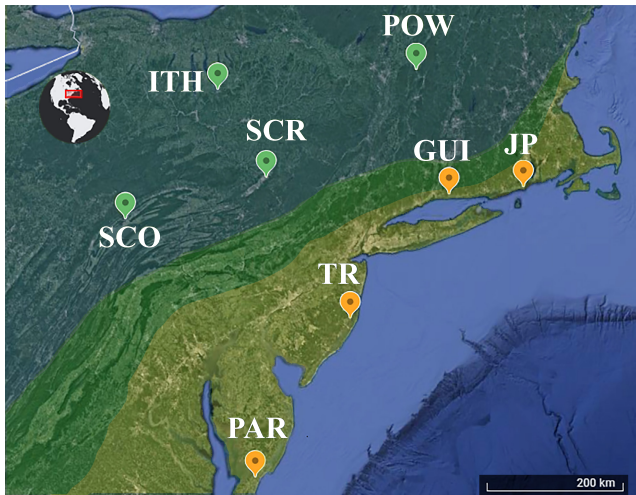


FIGURE 1 | Collection localities for the 96 sequenced individuals across four allopatric populations of each species ($n=12$ individuals per population). The highlighted areas correspond to the approximate distribution of each species: Yellowish for *Gryllus firmus*, blue-green for *G. pennsylvanicus* and the approximate location of the hybrid zone is shown in green.

2 | Methods

2.1 | Reference Genome Sequencing, Assembly, Gene Model Prediction and Annotation

We extracted high molecular weight genomic DNA from a single wild-caught *G. firmus* male cricket collected in Guilford, CT ($41^{\circ}16'07''$; $-72^{\circ}40'02''$) using a MagAttract HMW DNA Kit (Qiagen). We then prepared a 10 \times Genomics library using the Chromium Gel Bead and Library Kit (10 \times Genomics, Pleasanton, CA, USA) and the Chromium instrument (10 \times Genomics) following the manufacturer's protocols. The 10 \times library was distributed in four lanes for barcoding and sequenced at the Cornell Sequencing center (BRC) on a NextSeq500 (150 bp PE). Then, we assembled the sequences using the Supernova assembler ver. 1.2.2 under default parameters and the supernova parameter—style = pseudohap to obtain a single consensus sequence for each pair-ended scaffold (Weisenfeld et al. 2017).

We used the MAKER genome annotation pipeline (ver. 3, Cantarel et al. 2008) to annotate the scaffolds. First, we characterized repetitive elements using two approaches, de novo identification and reference-based identification. For de novo identification, we used RepeatModeler Open-1.0 (Smit and Hubley 2008–2015), which created a database of 1576 repetitive families. We used RepeatMasker Open-4.0 (Smit, Hubley, and Green 2013–2015) for reference-based repeat identification using the reference database for Arthropoda and Insecta clades (one run for each clade). We implemented both of these processes on the NCBI search engine. We then created a single GFF file (ProcessRepeats, RepeatMasker command) with all identified repeats. Second, to guide our gene model prediction, we created a custom reference protein database with 562,554 proteins (560,292 from SwissProt <including isoforms> and 2252 *Gryllus* spp. proteins recovered from NCBI). Additionally, we used the published transcriptome from *Gryllus rubens* as reference

(Berdan et al. 2016). Finally, with these information sources (model repetitive regions and gene evidence), we carried out the annotation and gene prediction with MAKER Annotation (Cantarel et al. 2008) as is described in the flow-chart in appendix (DRYAD Figure 1).

The draft genome data were used to locate previously characterized SNPs (Larson, Becker, et al. 2013; Larson et al. 2012) and design primers spanning ~5 kbps around each SNP. We did not obtain population level data from the *G. firmus* draft genome.

2.2 | Field Sampling

We sequenced 96 individuals from four *G. firmus* and four *G. pennsylvanicus* populations (12 individuals each, Figure 1). The *G. firmus* populations were from Tom's River, NJ (TR, $39^{\circ}45'00''$; $-74^{\circ}11'33''$); Judith Point, RI (JP, $41^{\circ}21'38''$; $-71^{\circ}28'53''$); Guilford, CT (GUI, $41^{\circ}16'07''$; $-72^{\circ}40'02''$); and Parksley, VA (PAR, $37^{\circ}45'58''$; $-75^{\circ}36'00''$). The *G. pennsylvanicus* populations were from Ithaca, NY (ITH, $42^{\circ}26'01''$; $-76^{\circ}29'59''$); State College, PA (SCO, $40^{\circ}47'59''$; $-77^{\circ}52'05''$); Scranton, PA (SCR, $41^{\circ}24'25''$; $-75^{\circ}35'46''$); and Pownal, VT (POW, $42^{\circ}45'16''$; $-73^{\circ}14'02''$), see map on Figure 1. Most of the *G. firmus* and *G. pennsylvanicus* populations were collected by L. S. Maroja in 2004–2005, the Pownal, VT and the RI populations were collected in 2010. In addition, we also sequenced four outgroup individuals: two *G. rubens* (from Durhan, NC) and two *G. bimaculatus* (from a captive colony, Hoy lab at Cornell).

2.3 | Targeted Resequencing and Individual Genotype Reconstruction

To design PCR primers, we located the scaffolds containing matches to previously sequenced loci (Gainey, Kim, and Maroja 2018; Larson, Andres, et al. 2013; Larson et al. 2014; Maroja et al. 2015). We sampled these previously characterized introgressing and non-introgressing loci from autosomes and X chromosomes; most non-introgressing loci were located on the X chromosome, the few autosomal non-introgressing loci were located on LG14 (Figure 2 shows the genetic map position of our sampled loci). We used BatchPrimer3 (You et al. 2008) to design primers spanning ~5–8 kb around previously introgressing ($n=13$) and non-introgressing loci ($n=34$) and one locus not previously studied located on either the X ($n=27$) or the autosomal chromosomes ($n=21$). We also sequenced a single locus that was not previously characterized (locus 202). To amplify the loci, we used LongAmp Hot Start Taq DNA Polymerase (New England Biolabs) following the recommended protocol and annealing temperature of 55 $^{\circ}$ C. We pooled the 49 PCR products for each individual and digested 5 μ l of individual samples with dsDNA fragmentase (New England Biolabs). We then purified digested samples with 1 \times Agencourt AMPureXLbeads (BeckmanCoulter Inc.) and re-suspended each sample in 15 μ l of 0.5 \times AE buffer. Before adding barcodes to fragmented individual products, we blunted the DNA ends using T4 Polynucleotide Kinase (New England Biolabs) and then adenylated fragments using Taq polymerase. We then ligated Illumina Truseq adapters with T4 DNA ligase to each individual sample. To identify each individual, we added dual TruSeq barcodes with PCR (OneTaq Hot Start,

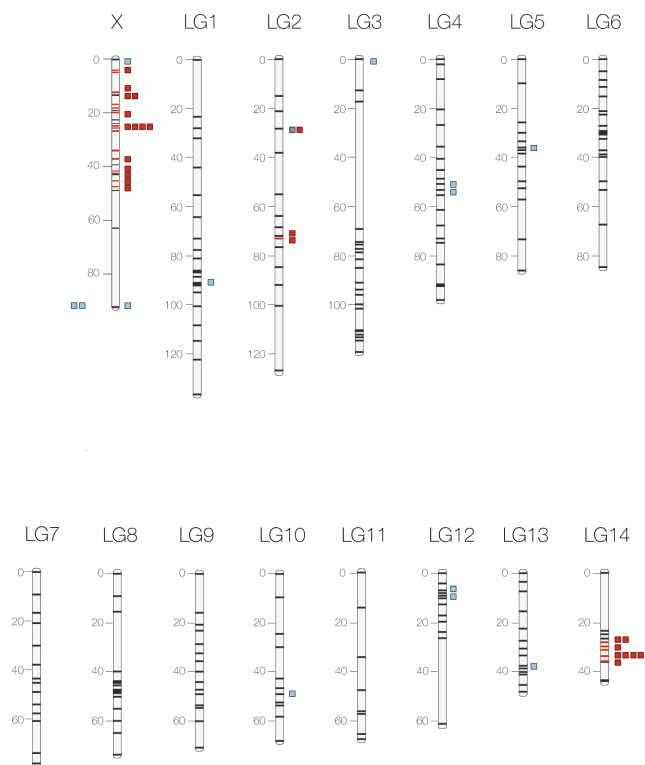


FIGURE 2 | Cricket genetic map based on Gainey, Kim, and Maroja (2018) for the X chromosome and Maroja et al. (2015) for autosomes. Sequenced loci are shown as blue squares for “introgressing” and red squares “non-introgressing” according to Larson, Andres, et al. (2013) and Larson et al. (2014). The locus in gray (LG2, locus 202) was not previously characterized as introgressing or non-introgressing. The two blue squares outside the X chromosome, were originally located to the X in Maroja et al. (2015) but were not able to be finely mapped in Gainey, Kim, and Maroja (2018) and therefore might not be located on the X chromosome.

NEB). We then pooled individuals and cleaned up the reaction with 0.7× Agencourt AMPureXP beads (BeckmanCoulter Inc.) and resuspended the pooled fragments into 0.3× AE buffer and 0.1× Tween detergent (concentration of 2–5 nM DNA). The library was sequenced with a NextSeq 500, 2 × 75 paired reads.

After removing low quality bases in Geneious R11.1.5 (Biomatters), we assembled the reads to the 49 known genomic reference sequences, one individual at a time (“Map to reference” workflow with medium-low sensitivity). We produced one consensus file for each locus by individual combination. First the reads for each individual/locus were aligned to a reference, then consensus sequences were called if there was a minimum coverage of > 8 reads; nucleotide positions were called heterozygous if the minimum allele had at least 25% frequency. Finally, we aligned all individual consensus files for each locus, creating a FASTA alignment file which included all individual sequences for each of the loci.

To export the sequences, we first removed all gaps by using the Mask Alignment tool, then we converted poor quality triple ambiguous sites D, H, V, B, into N (i.e., missing data—notice that ambiguous sites representing heterozygosity were kept) and imported individual locus files into DNAsp v.6 (Rozas et al. 2017)

and unphased haplotypes with the default parameters in PHASE and individuals separated by species (*G. firmus* and *G. pennsylvanicus*). Because this method separates the male loci into two identical reads, we removed one X-linked read from each male individual. To identify males (individuals were too young to be morphologically sexed), we screened five high coverage X-linked loci for homozygosity, and classified homozygous individuals as male. To account for sequencing error, if an individual had only a few (3 or fewer) polymorphisms, we checked these polymorphisms for allele quality and frequency before assigning a sex to the individual.

2.4 | Divergence Analyses

We used DnaSP v.6 (Rozas et al. 2017) to calculate Tajima’s D and π as well as D_{XY} and F_{ST} (Lynch and Crease 1990) between species using the “Divergence between populations” and “Gene Flow and Genetic Differentiation” commands, respectively. To test if the obtained values were significantly different between contrasting classes of loci (e.g., non-introgressing vs. introgressing; autosomal vs. sex-linked), we used a permutation analysis (permutation of values for whole individual/locus sequences) as implemented in the Boot package of R using 10,000 random reallocations.

2.5 | Estimates of Divergence Time

We used the outgroups *G. bimaculatus* and *G. rubens* to calculate a rough mutation rate for each of the 48 loci as an initial input for IMA3 (Hey et al. 2018). To do this we first calculated the mtDNA (COI) divergence times between *G. firmus*/*G. pennsylvanicus* and the outgroups using sequences available in Genbank. We used sequences KC488896–KC489085 for *G. firmus* and *G. pennsylvanicus* (Larson, Becker, et al. 2013), sequences KR071876, KU705555, KY646234, MF046161 for *G. bimaculatus* and sequences AY234789–AY234792 for *G. rubens*. We used $D = 2\mu t$, where D is the average divergence between consensus sequences of two species calculated with MEGA X (Kumar et al. 2018), μ is the mutation rate in percent difference per million years average arthropod mitochondrial DNA mutation rate set to the standard 0.0115 per site/My according to Brower (1994), also see Pons et al. (2010) and Chang et al. (2020), and t is the divergence time. Using our estimated mtDNA divergence time between the ingroup and each of outgroups (3.2 million years for *G. bimaculatus* and 2.7 million years for *G. rubens*), we then estimated mutation rates for each locus given the average D between ingroup and outgroup and the divergence time estimated from mtDNA (i.e., 3.2 million years for *G. bimaculatus* and 2.7 million years for *G. rubens*). When both outgroups (*G. rubens* and *G. bimaculatus*) were available (31/48 loci) we used the average mutation value as input for IMA3 (Hey et al. 2018). A total of 8/48 loci did not have any outgroup sequences available; we did not input a mutation rate for these loci (list of mutation rates and IMA3 input files can be found on DRYAD). As we point out in the discussion, a potential caveat of our approach to estimate nuclear divergence time is that it still relies on the standard mtDNA evolution rate (Brower 1994) for the estimation of initial mutation rate value for IMA3. To satisfy the requirement of no-recombination, we used only the longest non-recombining

region for each locus, identified with the IMgc software package (Woerner, Cox, and Hammer 2007). We assumed one generation per year, as *Gryllus* in the Northeast of the US survive the winter exclusively through the diapause egg stage and have non overlapping generations. We assumed infinite sites model (I), informed the autosomal or X chromosomal location of each locus (DRYAD Table 1 contains detailed information on size and mutation rate for each locus) and assumed a population size prior of 40 ($4 N\mu$), time since speciation of 90 ($t\mu$), and migration rate of 4 (m/μ). The parameters reached stationarity (no perceivable trends) after about 350,000 generations with a geometric heating scheme of 100 parallel hot chains ($hn = 100$, $ha = 0.99$, $hb = 0.75$). After 400,000 generations of burn-in, we collected an additional 30,000 genealogies. We ran simulations with these parameters three times and in all cases, we obtained equivalent results.

2.6 | Phylogeny Reconstruction

To build the phylogeny, we first created a concatenated sequence file of all loci, using the command “concat” in the program seqkit (Shen et al. 2016) and realigning data in Geneious R11.1.5 (Biomatters). We eliminated 10 loci that had fewer than 65 individuals sequenced (loci 432, 202, 726, 1121, 4205, 7164, 14741, 8026, 5556, and 6271) and eliminated three loci which did not have any outgroups (loci 211, 3968, and 5961). We also eliminated 11 *G. firmus*, 8 *G. pennsylvanicus*, one *G. bimaculatus* and one *G. rubens* individuals which had fewer than 80% data completion. The final dataset was 172,357 nucleotides long (34 loci) and contained 37 *G. firmus*, 40 *G. pennsylvanicus*, one *G. rubens*, and one *G. bimaculatus*. We used MrBayes version 3.2.6 (Huelsenbeck and Ronquist 2001) with the general time reversible model with invariant sites, gamma rates, and default priors (GTR + I + G), allowing the rate at each site to change over evolutionary history. We ran 5 million generations and discarded the first 25% of the trees. We also ran neighbor-joining and maximum likelihood methods, but since these trees were virtually identical to the Bayesian tree, we omitted these results.

3 | Results

The gDNA used for 10× Genomics DNA library construction had a high molecular weight (HMW) of 38.6 kb which generated 772.7 million raw pair-end reads (61× coverage). After trimming, these reads had an average length of 139.5 bp; ~76% of these reads had a $Q > 30$. The percentage of non-duplicated,

phased reads was ~45%. These data are within the optimal standard values suggested by Supernova Assembler software manufacturer.

Supernova assembled 14,450 long scaffolds (over 10 kb) having a total length of 1.28 Gb—much smaller than the 1.66 Gb *Gryllus bimaculatus* genome (Ylla et al. 2021), likely due to incompleteness. The average distance between SNPs (hetdist) was ~122 bp. The contig N50 and the scaffold N50 of the assembled genome were 22.73 kb and 346.11 kb, respectively. All statistic outputs from Supernova analyses can be found in the DRYAD, Table S2.

For the 48 individually sequenced loci (Figure 1, population locations), we accepted a minimum depth per individual of 8 reads and, because of low coverage, some individual/loci combinations were excluded; the resulting average number of individuals per locus was 77 ± 22 individuals. Of these loci, 13 are introgressing and 34 are non-introgressing, and one was not determined based on previous cline analysis (Larson, Andres, et al. 2013). Of the 48 loci, 21 were autosomal and 27 were X-linked (Gainey, Kim, and Maroja 2018; Maroja et al. 2015).

3.1 | Gene Annotation

MAKER *de novo* Annotation established the homologous location of 5064 genes (mean length ~4.23 kb). Almost 98% of the annotations were within 0.5 AED score (Annotation Edit Distance), suggesting a high accuracy in the prediction of gene models based on homology. To adjust the homology-based gene models, and to predict unidentified genes we used SNAP (Korf 2004) and AUGUSTUS (Stanke and Waack 2003). SNAP was trained through two consecutive runs, after which we obtained 19,501 genes (mean length ~5.78 kb). Over 84% of the SNAP-based gene models had AED scores < 0.5 . AUGUSTUS was trained and ran independently, resulting in 19,243 genes (mean length ~6.72 kb). Around 84% of the AUGUSTUS-based gene models showed AED score < 0.5 . The final annotation, carried out combining all results from the homology, SNAP, and AUGUSTUS based gene models, resulted in 19,157 gene models (mean length ~7.12 kb), 84% of them with AED score < 0.5 .

The median coding sequence (CDS) length of the final annotation was 939 base pairs (bp), and the median transcript length was 522 bp. Furthermore, we observed that the 5' untranslated regions (UTR) had a median length of 34 bp, and the 3' UTR had a median length of 177 bp. The analyses revealed a total of 19,158

TABLE 1 | Summary of F_{ST} and D_{XY} results (also see Figure 3). Comparisons across non-introgressing and introgressing loci located on autosomes or X chromosome (n loci represents the number of analyzed loci in each category).

	F_{ST} (median)	D_{XY} (median)	n loci	P_{Random} (F_{ST}/D_{XY})
Non-introgressing/introgressing	0.75/0.64	0.008/0.01	30/12	0.024/0.060
Autosomal/X-linked	0.67/0.74	0.013/0.008	18/24	0.135/ 10^{-4}
Autosomal only (non-introgressing/introgressing)	0.67/0.76	0.012/0.014	10/8	0.37/0.87
X-linked only (non-introgressing/introgressing)	0.75/0.64	0.0080/0.0082	4/20	0.024/1

Note: P_{Random} represents the value of permutation analysis using 10,000 random reallocations for F_{ST} comparisons and for D_{XY} comparisons, significant results are in bold.

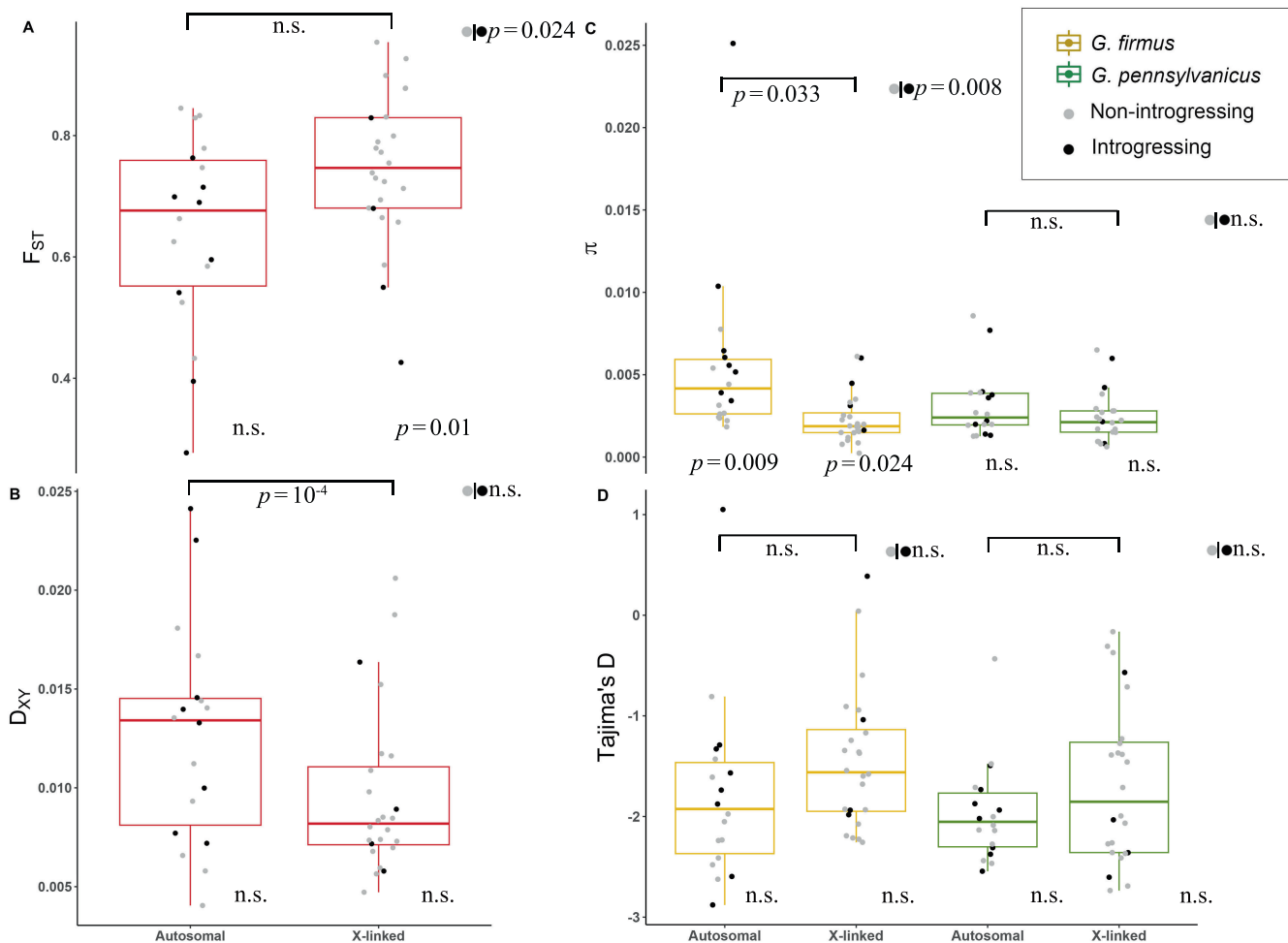


FIGURE 3 | Box plots and data points for F_{ST} (A), D_{XY} (B), π (C) and Tajima's D (D). Between species F_{ST} and D_{XY} comparisons (A, B) are shown in red; within species π and Tajima's D are shown in yellow for the beach cricket, *Gryllus firmus*, and green for the field cricket, *G. pennsylvanicus*. Data points for previously characterized introgressing loci are shown in black, and for non-introgressing loci are shown in gray. Significant values for comparisons between autosomes and X chromosomes are shown on top; overall comparisons (without taking into account chromosomal linkage) between introgressing and non-introgressing loci are shown to the right of each group, sidewise to gray and black circles; comparisons between introgression and non-introgressing loci taking into account chromosomal linkage are shown on the bottom right of each box plot. For sample sizes refer to the text (Section 3).

genes and 19,157 transcripts, with 18,885 unique transcripts and 12,499 unique CDS among them, showcasing a diverse transcriptional landscape. Of these genes, 301 genes exhibited a 5' UTR, and 37 genes had a 3' UTR.

We identified all annotated genes located in scaffolds containing each of the 48 ~5 kb loci of interest. Within these regions we found a total of 501 annotated genes (DRYAD contains the gff files). Thirty-seven of these regions contained annotated coding genes inside the target locus sequence, with five of them containing more than one gene.

3.2 | Relative (F_{ST}) and Absolute (D_{XY}) Divergence Patterns

For the divergence analyses, we excluded six loci (202, 1121, 5556, 6271, 7164, and 8026) that had a low number of represented populations (in some cases all individuals of a given population were excluded due to low coverage). We classified loci

as “introgressing” or “non-introgressing” according to Larson, Andres, et al. (2013) and Larson et al. (2014). We used a total of 42 loci, 12 of them introgressing (8 autosomal and 4 X-linked) and 30 non-introgressing (10 autosomal and 20 X-linked). The estimated sex ratio for *G. firmus* was 25:23 (δ : φ) and for *G. pennsylvanicus* it was 30:18 (δ : φ).

As expected, loci showing reduced introgression in previous genomic cline analyses (Larson, Andres, et al. 2013; Larson et al. 2014) showed elevated values of relative divergence (median F_{ST} non-introgressing = 0.75, F_{ST} introgressing = 0.64, p_{Random} = 0.024). However, most of the loci (20/30) in the non-introgressing category are located in the X chromosome, so the observed pattern could be driven by chromosomal linkage rather than previously characterized “introgression patterns”—that is, loci could be classified as “non-introgressing” due to X-linkage and not because they reduce hybrid fitness. Two methodological problems—males being classified as homozygotes (for the X-linked loci), and the one-way directional incompatibility observed in this hybrid zone—could increase

the probability of X loci being classified as non-introgressing. Indeed, taking into account chromosomal linkage (median F_{ST} autosomal = 0.67, F_{ST} X-linked = 0.74, $p_{Random} = 0.135$, Figure 3A), we found no clear association between F_{ST} and the previously measured introgression pattern per se: for the autosomal loci, there was no relationship between F_{ST} and introgression pattern (median F_{ST} autosomal non introgressing = 0.70, F_{ST} autosomal introgressing = 0.76, $p_{Random} = 0.37$, Figure 3A). In the X chromosome, introgressing loci do indeed seem to have lower F_{ST} values (median F_{ST} X-non introgressing = 0.75, F_{ST} X-introgressing = 0.64, $p_{Random} = 0.024$). We note that this result is based on few introgressing loci ($n=4$, Figure 3A), including two outliers that might not be X-linked (unconfirmed linkage). For summary of results see Table 1.

In contrast with the predictions of the genomic island model, non-introgressing loci did not show elevated D_{XY} values (median D_{XY} non-introgressing = 0.008, D_{XY} introgressing = 0.01, $p_{Random} = 0.060$, Figure 3B). While the median D_{XY} between these two categories is not significant, the difference goes in the opposite direction from the expectations of the genomic islands model (according to the islands model, loci showing reduced introgression should have higher, not lower, D_{XY} value). There was a large difference between median autosomal and X-linked D_{XY} (D_{XY} autosomal = 0.013, D_{XY} X-linked = 0.008, $p_{Random} = 10^{-4}$, Figure 3A). Taking into account chromosomal linkage, we found no relationship between D_{XY} and introgression pattern (median D_{XY} autosomal non introgressing = 0.012, D_{XY} autosomal introgressing = 0.014; $p_{Random} = 0.87$, Figure 3B and median D_{XY} X-non introgressing = 0.0080, D_{XY} X-introgressing = 0.0082, $p_{Random} = 1$). Furthermore, we found no significant correlation between F_{ST} and D_{XY} ($\rho_{Spearman} = 0.167$, $p_{bootstrap} = 0.13$).

3.3 | Neutrality Tests and Nucleotide Diversity

The genomic baseline value of Tajima's D can be biased by demographic effects particularly after population expansions or contractions. As expected, due to likely demographic expansion after the last glacial maxima (in late Pleistocene epoch), both species had a negative median Tajima's D (*G. firmus*: -1.65 CI_{Bca-95%} (-1.96, -1.43); *G. pennsylvanicus*: -1.9 , CI_{Bca-95%} (-2.19, -1.81), Figure 3D), with the signal especially strong in *G. pennsylvanicus*, the species with broadest current geographic distribution. For both species, the median Tajima's D was almost identical between introgressing and non-introgressing loci (*G. firmus*: T_D introgressing = -1.65 , T_D non introgressing = -1.64 , $p_{Random} = 0.97$; *G. pennsylvanicus*: T_D introgressing = -2.02 , T_D non introgressing = -2.00 , $p_{Random} = 0.90$).

In terms of genetic variation, the median π for *G. firmus* was substantially smaller for non-introgressing than for introgressing loci, while *G. pennsylvanicus* had almost identical median π values for introgressing and non-introgressing loci (*G. firmus*: $\pi_{introgressing} = 0.054$, $\pi_{non introgressing} = 0.002$, $p_{Random} = 0.008$; *G. pennsylvanicus*: $\pi_{introgressing} = 0.0029$, $\pi_{non introgressing} = 0.0020$, $p_{Random} = 0.158$). Taking into account chromosomal linkage (autosomal or X-linked) we found no significant differences between introgressing and non-introgressing loci in π or Tajima's D for *G. pennsylvanicus*. For *G. firmus* there were differences between introgressing and X-linked loci located in autosome or X-linked

($\pi_{autosomal introgressing} = 0.0058$ $\pi_{autosomal non-introgressing} = 0.0026$, $p_{Random} = 0.009$) ($\pi_{Xintrogressing} = 0.0038$, $\pi_{Xnon-introgressing} = 0.0018$, $p_{Random} = 0.024$).

While in *G. pennsylvanicus* X-linked and autosomal loci have similar median levels of variation (π) and Tajima's D values ($\pi_{autosomal} = 0.0024$, $\pi_{X-linked} = 0.0021$, $p_{Random} = 0.64$; T_D autosomal = -2.052 , T_D X-linked = -1.853 , $p_{Random} = 0.64$; Figure 3C,D), in *G. firmus* X-linked loci showed median levels of variation slightly reduced, but not significantly so, (*G. firmus*: $\pi_{autosomal} = 0.0041$, $\pi_{X-linked} = 0.0018$, $p_{Random} = 0.033$, Figure 3C). This reduced variation on the *G. firmus* X-chromosome was not associated with more negative Tajima's D values (*G. firmus*: T_D autosomal = -1.92 , T_D X-linked = -1.56 , $p_{Random} = 0.144$, Figure 3D).

The autosome with a large number of previously identified non-introgressing loci (LG14, all genes contained in a 7 cM window, Figure 2) did not show lower genetic diversity in either in *G. pennsylvanicus* ($\pi_{LG14} = 1.9 \times 10^{-3}$, $\pi_{all other autosomal} = 2.6 \times 10^{-3}$, $p_{Random} = 0.679$) or in *G. firmus* ($\pi_{LG14} = 2.6 \times 10^{-3}$, $\pi_{all other autosomal} = 5.1 \times 10^{-3}$, $p_{Random} = 0.139$).

3.4 | Estimates of Divergence Time

To calculate divergence time, we used 48 loci. The loci had average and standard deviation non-recombining length of 3584.4 ± 1013.9 bp (original untrimmed length 4760.77 ± 103.319) and average and standard deviation mutation rate of $4.052 \times 10^{-9} \pm 1.301 \times 10^{-9}$ (DRYAD Table S1). As expected, due to lack of recombination in males, the trimmed (non-recombining) sequence length was larger for X-linked than autosomal loci (4001.89 and 3047.67 respectively, $t = 3.5$, $p < 0.001$); the difference was only marginally significant for the original total sequence length (4800.1 and 4730.7, respectively, $t = 2.1$, $p < 0.044$). There was no difference in mutation rate between X-linked and autosomal loci (4.00×10^{-9} and 4.10×10^{-9} , $t = 0.2$, $p > 0.8$). For the loci previously classified as "introgressing" or "non-introgressing" there were no differences in either total length (4743.231 and 4779.629 respectively, $t = -1.22$, $p > 0.2$), non-recombining trimmed length (3293.301 and 3692.54 respectively, $t = -1.28$, $p > 0.2$) or mutation rate (4.15×10^{-9} and 4.00×10^{-9} respectively, $t = 0.33$, $p > 0.74$).

The estimated mean divergence time was 819,119 years (with 95% Lo 601,306 to 95% hi 1,052,651 years). This divergence time is substantially larger than the previously estimated divergence time, based on mtDNA, of about 200,000 years (Maroja, Andres, and Harrison 2009; Willett, Ford, and Harrison 1997). We note that this divergence time is of course dependent on mutation rate calculation, in this case calibrated for each nuclear locus based on the divergence between the ingroups (*G. firmus*/*G. pennsylvanicus*) and outgroups (*G. bimaculatus* and/or *G. rubens*) sequences for each nuclear loci. The divergence time between ingroup and outgroups was initially calibrated with mtDNA and estimated at 3.2 million years for *G. bimaculatus* and 2.7 million years for *G. rubens*. When calibration is done based on divergence between *G. firmus* and *G. pennsylvanicus* (calibrated based on their mtDNA divergence time, 200K years also according to the standard mtDNA mutation

rate from Brower (1994)), then the nuclear mutation rate per locus is higher. However, with these estimated nuclear mutation rates (i.e., those calibrated with mtDNA divergence between *G. pennsylvanicus* and *G. firmus*) the divergence with *G. bimaculatus*/*G. rubens* becomes much more recent than what is their mtDNA divergence suggests. This indicates that either mtDNA divergence has not been constant through the evolutionary history or has not been congruent with nuclear divergence. The discrepancy in mutation rates can be easily explained if mtDNA introgressed between *G. firmus* and *G. pennsylvanicus*. This explanation is supported by the unresolved mtDNA phylogeny for *G. firmus* and *G. pennsylvanicus* (Willett, Ford, and Harrison 1997), which contrasts with the nuclear DNA phylogeny (see Figure 4). For nuclear DNA, *G. firmus* and *G. pennsylvanicus* were reciprocally monophyletic with very strong branch support of 1.0 (identical results were observed with ML or distance trees).

The IMA3 program also estimates migration rates. The migration rate from *G. firmus* to *G. pennsylvanicus* was 0.1497 migrants per generation ($2Nm$) whereas the migration rate from *G. pennsylvanicus* into *G. firmus* was 0.0941 migrants per generation.

4 | Discussion

We did not find evidence of introgression in allopatric populations of the *G. firmus* and *G. pennsylvanicus* system. While hybrids are certainly formed in the hybrid zone, and clines across the zone in both phenotypic (Ross and Harrison 2002) and genomic markers (Larson et al. 2014) exist, the genome of allopatric populations geographically close to the hybrid zone does not seem to conform to the “islands” model expectations (Cruickshank and Hahn 2014).

If the “islands model” holds, parts of the genome will be homogenized by gene flow between species, whereas other highly differentiated parts will be impermeable to gene flow due to divergent selection or because they contribute to reproductive barriers. Under this scenario we should expect to observe both high relative (F_{ST}) and high absolute divergence (D_{XY}) in

the “impermeable” (non-introgressing) regions. We instead found that genomic regions, previously characterized as “non-introgressing,” in fact had lower (although not significantly so) absolute divergence than “introgressing” regions (Figure 3). This could be because introgression does not continue past the hybrid zone, and thus does not affect the genomes from allopatric population. However, it could suggest that, what previously had been assumed to be “islands of speciation,” might be “incidental islands” caused by selection and linked selection (Turner and Hahn 2010). It is possible that the genomic islands identified in the hybrid zone were a consequence of methodological caveats and hybridization patterns, and did not in fact represent loci under negative selection in hybrid offspring. When the introgression patterns were calculated (Larson, Andres, et al. 2013; Larson et al. 2014), the mapping location of SNPs was not yet known; therefore, X-linked alleles in males were coded as “homozygotes” for the INTROGRESS (Gompert and Buerkle 2010) calculations—a methodological issue that could generate misleading patterns (also see Maroja et al. 2015). Furthermore, the one-directional barrier to gene exchange, where male F1 offspring inherit the incompatibility towards *G. firmus* females, make introgression stronger from *G. firmus* into *G. pennsylvanicus* with reduced presence of *G. firmus* X chromosomes in relation to autosomes (since hybridizing *G. firmus* must be males (X0), F1 offspring will have more *G. pennsylvanicus* X chromosomes in a 2:1 ratio)—this would lead to a lack of introgression on the X compared to autosomes, leading to a presence of “islands” in the X (where basically the whole X behaves as an “island”).

4.1 | Patterns of Differentiation Do Not Support a Simple “Islands of Differentiation” Model

To test the “islands” model, we sequenced a total of 48 introgressing and non-introgressing loci originally developed from a male accessory gland transcriptome library (Andres et al. 2013; Larson, Andres, et al. 2013) that showed high allele frequency differences between two allopatric populations of each species (Guilford, CT and Ithaca, NY, Andres et al. 2013). Once mapped to the genome, these previously characterized

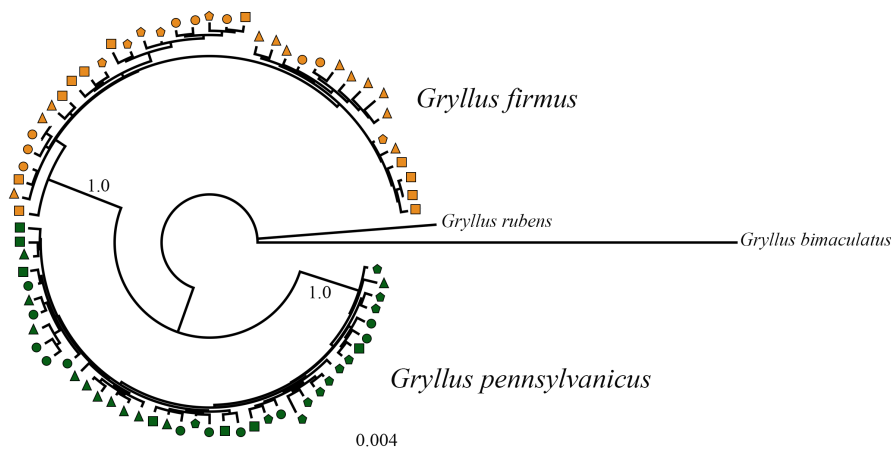


FIGURE 4 | Bayesian phylogenetic tree for concatenated loci (172,357 nucleotides), showing populations for the two species. Other tree building methods also showed the two species as reciprocally monophyletic. Symbols indicate populations: GUI (yellow square), PAR (yellow triangle), TRI (yellow circle), JP (yellow pentagon), ITH (green square), SCO (green triangle), SCR (green circle) and POW (green pentagon).

markers were found to be represented in all cricket chromosomes (the 14 autosomes and the large X, representing ~20% of the genome, Lim, Vickery, and Mcekevan 1973—not all of these were sequenced here). However, the loci with steep clines classified as “non-introgressing” (Larson, Andres, et al. 2013) and assumed to be impermeable to interspecific gene flow, were, for the most part, restricted to the X chromosome and a small region of autosome LG14 (Gainey, Kim, and Maroja 2018; Maroja et al. 2015), suggesting that genomic architecture might be important (Renaut et al. 2013). As a result of this clustered distribution of loci, our current sample of 48 loci was inevitably clumped; most of our “non-introgressing” loci were X linked ($n = 27$, Figure 2) and the X contained very few “introgressing” loci ($n = 4$). Furthermore, two-thirds (8/12) of the “non-introgressing” autosomal loci were located to a small region of LG14. Therefore, disentangling the effects of X linkage from permeability across the hybrid zone was difficult.

Despite linkage difficulties, our results do not support an “islands” model. First, although F_{ST} was indeed higher for “non-introgressing” loci, only four non-introgressing loci are located on the X chromosome, and two of these loci (loci 8833 and 541, Figure 2 and DRYAD Table S3) were not able to be positioned in the second mapping study and therefore might not represent truly X linked loci (Gainey, Kim, and Maroja 2018). Second, when we restricted the analyses to autosomes, we found no F_{ST} differences between introgressing and non-introgressing loci. Third, the D_{XY} results showed that “non-introgressing” loci, were in fact less (although not significantly so) differentiated than “introgressing” loci, a result that directly contradicts expectations for the “islands” model, where we would expect that non-introgressing loci would have higher relative and absolute genetic differentiation.

While we did not find any support for the island model, a few caveats should be considered. The characterization of loci as either “introgressing” or “non-introgressing” might be an artifact (and since our population collection is allopatric, the current data cannot be used to recalculate clines). Our classification was based on two previous genomic cline studies conducted before the SNPs had been mapped (Larson, Andres, et al. 2013; Larson et al. 2014); therefore, the X-linked loci could have been characterized erroneously as “non-introgressing” because only female backcross could be interspecific heterozygotes (note that all males would have been coded as homozygotes, since it was unknown that they were in fact hemizygotes). However, even if this was the case and loci were indeed misclassified as introgressing or non-introgressing, there was no correlation between F_{ST} and D_{XY} , as we would expect if low F_{ST} (and D_{XY}) were brought about by mixing of genomes through introgression.

Given that prezygotic isolation between these two species is strong (Harrison 1983, 1985; Larson, Becker, et al. 2013; Maroja et al. 2009, 2014), that gene flow is restricted on a fine scale (Larson et al. 2014; Ross and Harrison 2002), and that linkage disequilibrium is high even in the center of the hybrid zone (Harrison and Bogdanowicz 1997), it is also possible that introgression is indeed negligible and that the island model does not apply.

Our IMA3 calculations suggest a scenario of low migration ($0.14-0.22 2 N_e m$) and a divergence time of ~1.6 N_e generations. Assuming low migration, differences between “islands” and “genomic background” are apparent only after $\geq 0.4 N_e$ generations (Cruickshank and Hahn 2014). Therefore, we should have been able to detect differences between these two models. However, if secondary contact is much more recent than the estimated divergence, it is possible that there has been no time for introgression to homogenize even the freely introgressing parts of the genome, and we may have no power to properly test the island’s model.

4.2 | The Importance of the X Chromosome

While our results put in question the “islands of divergence” model for the *G. firmus* and *G. pennsylvanicus* hybrid zone, they do not detract from the importance of the X chromosome in speciation (e.g., Charlesworth, Coyne, and Barton 1987; Good, Dean, and Nachman 2008; Hu and Filatov 2016; Masly and Presgraves 2007; Meisel and Connallon 2013) or to evidence of more intense selection on X-linked loci, often associated with reduced introgression in other species (e.g., Carneiro et al. 2014; Fontaine et al. 2015; Garrigan et al. 2012; Hu and Filatov 2016; Payseur, Krenz, and Nachman 2004; Saetre et al. 2003; Sankararaman et al. 2014).

As is the case in many other systems, it is possible that the X chromosome is under stronger local selection or drift, bringing alleles to fixation and reducing ancestral polymorphism, increasing F_{ST} without necessarily increasing D_{XY} . Our data show that the *G. firmus* X chromosome is indeed reduced in polymorphisms (π , Figure 3C), as would be expected due to its smaller effective population size ($3/4 N_e$) and/or more intense selection (recessive alleles exposed in males) and lower recombination in relation to autosomes. Furthermore, while F_{ST} is indeed higher for the X (although not significantly so, Figure 3A), suggesting more fixed variants due to selection or drift, the absolute genetic divergence (D_{XY}) in the X is significantly lower than that in autosomes (Figure 3B). This contrasts to patterns observed in *Drosophila*, where the X seems to be *more* divergent than autosomes (Garrigan et al. 2015), but is concurrent with the pattern observed in human-chimp comparisons (Dutheil et al. 2015; Narang and Wilson Sayres 2016). In lower divergence and polymorphism, the human-chimp X chromosome is virtually devoid of incomplete lineage sorting (Dutheil et al. 2015). This pattern could be explained by more frequent recurrent selective sweeps (including the possibility of meiotic-drive), in addition to higher background selection, leading to a depletion of polymorphism right before speciation (Dutheil et al. 2015). This hypothesis is further supported by the fact that divergence and polymorphism are lower around coding genes, but are closer to expected levels of variation (3/4 of autosome) when further away (Arbiza et al. 2014; Narang and Wilson Sayres 2016), as would be expected if the reduction of polymorphism was due to linked selection. If a similar reduction in variation happens in field crickets, then we would expect that the X would be less divergent due to the lack of variation at the onset of speciation.

The X chromosome has many unusual patterns and processes (Vicoso and Charlesworth 2006) and our data do not allow us to

test particular models that can best explain the slightly higher X_{ST} —such as faster-X evolution (Meisel and Connallon 2013), or dominance (Turelli and Orr 2000). However, as we have no evidence of hybrid male sterility in this system—F1 males cannot fertilize pure *G. firmus* females (one-way incompatibility is maintained) but they are fertile with F1 females as well as pure *G. pennsylvanicus*—some of the models that assume hybrid male sterility should not apply to this system (Presgraves and Meiklejohn 2021).

It is also important to notice that the dynamics of the hybridization incompatibility have contributed to the X being characterized as “non-introgressing” in the cricket hybrid zone. Because of the one-way incompatibility (F1 hybrids can only be produced from *G. pennsylvanicus* females; Harrison 1983), the F1 generation will exhibit a 2:1 excess of *G. pennsylvanicus* X chromosomes (F1 females: $X_{penn}X_{firm}$, F1 males: $X_{penn}0$) and, *G. pennsylvanicus* males or male F1s ($X_{penn}0$) cannot contribute to the introgression of X-linked *G. firmus* alleles because they cannot produce backcross hybrids with *G. firmus* females (L. S. Maroja and E. L. Larson, unpublished data). In contrast, *G. firmus* autosomes are free to introgress into the *G. pennsylvanicus* background through the F1 males since they are fully fertile with *G. pennsylvanicus* females (Larson et al. 2012). Furthermore, if females behave in the wild as they do in lab, we would expect that F1 females would mate more promptly to *G. firmus* males, contributing to higher flow from *G. pennsylvanicus* into *G. firmus* (Maroja et al. 2009). Therefore, the mating dynamics of the hybrid zone, combined with the initial lack of linkage data for markers (males coded as homozygotes, see above) might help explain steeper X-linked genomic clines observed across the two hybrid zone locations (Larson, Andres, et al. 2013; Larson et al. 2014)—this reduces the probability that the previously identified loci represent “barrier genes:” it is more likely they represent a genomic architecture phenomenon.

4.3 | The Divergence Time Between Species

Based on mtDNA, *Gryllus firmus* and *G. pennsylvanicus* have been considered recently diverged species—the divergence time estimated to be only 200 Kyears (Broughton and Harrison 2003; Maroja, Andres, and Harrison 2009). For the first time, in this paper, we calculate the divergence time using nuclear DNA. Our estimate of ~800 Kyears makes the species divergence 4× older. This new estimate is consistent with the multiple described barriers to gene exchange between these two species (Maroja et al. 2009). This includes not only fertilization barriers (Harrison 1983; Larson et al. 2012), but cuticular hydrocarbon differences (Heggeseth et al. 2020; Maroja et al. 2014), time to mate (Harrison 1983), and habitat and temporal isolation (Ross and Harrison 2002). It is therefore likely that their divergence time has been longer than just 200 Kyears ago, especially given that throughout most of the species' range, they only have one generation per year.

A potential caveat of this new nuclear divergence time calculation is that it still relies on the mtDNA evolution rate (Brower 1994) for the estimation of initial mutation rate value for IMA3. Here, we used outgroups *G. bimaculatus* and *G. rubens* and obtained a mtDNA-based estimated divergence time between *G. firmus* and *G. pennsylvanicus* and each outgroup. Then using this estimated time, we calculated a nuclear mutation rate between the

outgroup and *G. firmus/G. pennsylvanicus* (DRYAD Table S1). Therefore, if the mutation rate estimate is incorrect, or if there are phylogenetic incongruencies between mtDNA and nuclear-based trees (for the *G. firmus/G. pennsylvanicus*, *G. rubens*, and *G. bimaculatus*), then our estimated mutation rates, and therefore calculated divergence time, would be compromised.

5 | Conclusions

Our data do not support the “islands model”—that is, we have no evidence that parts of the genome are homogenized by gene flow, while other parts remain impermeable and highly differentiated due to divergent selection or because they act as barriers to gene exchange. These findings do not bring into question the importance of the X chromosome in hybridization and speciation—even allopatric species without hybrid zones often show selection against introgressed X. What we do demonstrate is that gene flow across the hybrid zone has not been sufficient to homogenize the genome of allopatric populations proximal to the hybrid zone; instead, most genes might follow a “non-introgressing” model, with previously characterized non-introgressing regions of the X chromosome exhibiting increased relative differentiation (F_{ST}) possibly due to higher local selection and/or drift.

This result might not be as surprising in view of the longer divergence time. While according to mtDNA, the species divergence only 200 Kyears ago, our nuclear DNA estimates elevate this by 4×—it is therefore not as surprising that the crickets already have in place a number of barriers to gene exchange and that allopatric populations do not appear to show evidence of genomic mixing.

Author Contributions

Luana S. Maroja: conceptualization (lead), data curation (equal), formal analysis (equal), funding acquisition (equal), investigation (equal), methodology (equal), project administration (lead), writing – original draft (equal), writing – review and editing (equal). **Francesca Barradale:** data curation (equal), formal analysis (equal), investigation (equal), methodology (equal), writing – original draft (equal). **Sebastian A. Espinoza-Ulloa:** data curation (equal), formal analysis (equal). **Steven Bogdanowicz:** data curation (equal), formal analysis (equal), investigation (equal), methodology (equal). **Jose Andres:** conceptualization (equal), data curation (equal), formal analysis (equal), investigation (equal), methodology (equal), writing – original draft (equal), writing – review and editing (equal).

Acknowledgments

We thank the NSF DEB-1650887 to L.S.M. and S.B. and NSF DEB-2012060 to L.S.M. for funding for this project. We thank James Marvel-Coen (Williams College alum) for collecting preliminary data for this experiment. We thank anonymous reviewers for their comments.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Genetic data: Individual genotype data from the individually sequenced loci are available on DataDryad (<https://doi.org/10.5061/dryad.s4mw6m9bb>), this includes all FASTA files, IMA3 dataset as well as all data (F_{ST} , D_{XY} , pi, etc.) used in the R analyses (scripts

also included). Additional figures and tables are also included in SuppDryad: (<https://doi.org/10.5061/dryad.s4mw6m9bb>). The annotated draft genome files (gff) from single *G. firmus* individual, are also available on DataDryad (<https://doi.org/10.5061/dryad.s4mw6m9bb>; Maroja et al. 2024). Genomic DNA data have been deposited in Genbank, accession numbers OR441320–OR447380 (Maroja et al. 2023).

References

- Andres, J. A., E. L. Larson, S. M. Bogdanowicz, and R. G. Harrison. 2013. "Patterns of Transcriptome Divergence in the Male Accessory Gland of Two Closely Related Species of Field Crickets." *Genetics* 193, no. 2: 501–513. <https://doi.org/10.1534/genetics.112.142299>.
- Arbiza, L., S. Gottipati, A. Siepel, and A. Keinan. 2014. "Contrasting X-Linked and Autosomal Diversity Across 14 Human Populations." *American Journal of Human Genetics* 94, no. 6: 827–844. <https://doi.org/10.1016/j.ajhg.2014.04.011>.
- Barr, C. M., and L. Fishman. 2010. "The Nuclear Component of a Cytonuclear Hybrid Incompatibility in *Mimulus* Maps to a Cluster of Pentatricopeptide Repeat Genes." *Genetics* 184, no. 2: 455–465. <https://doi.org/10.1534/genetics.109.108175>.
- Barton, N. H., and G. M. Hewitt. 1981. "Hybrid Zones and Speciation." In *Evolution and Speciation*, edited by W. D. Atchley and D. S. Woodruff, 109–145. Cambridge University Press. ISBN-13 9780521238236.
- Berdan, E. L., T. Blankers, I. Waurick, C. J. Mazzoni, and F. Mayer. 2016. "A Genes Eye View of Ontogeny: De Novo Assembly and Profiling of the *Gryllus rubens* Transcriptome." *Molecular Ecology Resources* 16, no. 6: 1478–1490. <https://doi.org/10.1111/1755-0998.12530>.
- Broughton, R. E., and R. G. Harrison. 2003. "Nuclear Gene Genealogies Reveal Historical, Demographic and Selective Factors Associated With Speciation in Field Crickets." *Genetics* 163, no. 4: 1389–1401.
- Brower, A. V. 1994. "Rapid Morphological Radiation and Convergence Among Races of the Butterfly *Heliconius erato* Inferred From Patterns of Mitochondrial DNA Evolution." *Proceedings of the National Academy of Sciences of the United States of America* 91, no. 14: 6491–6495.
- Cantarel, B. L., I. Korf, S. M. Robb, et al. 2008. "MAKER: An Easy-to-Use Annotation Pipeline Designed for Emerging Model Organism Genomes." *Genome Research* 18, no. 1: 188–196. <https://doi.org/10.1101/gr.6743907>.
- Carneiro, M., F. W. Albert, S. Afonso, et al. 2014. "The Genomic Architecture of Population Divergence Between Subspecies of the European Rabbit." *PLoS Genetics* 10, no. 8: e1003519. <https://doi.org/10.1371/journal.pgen.1003519>.
- Carneiro, M., J. A. Blanco-Aguilar, R. Villafuerte, N. Ferrand, and M. W. Nachman. 2010. "Speciation in the European Rabbit (*Oryctolagus cuniculus*): Islands of Differentiation on the X Chromosome and Autosomes." *Evolution* 64, no. 12: 3443–3460. <https://doi.org/10.1111/j.1558-5646.2010.01092.x>.
- Chang, H., Z. Qiu, H. Yuan, et al. 2020. "Evolutionary Rates of and Selective Constraints on the Mitochondrial Genomes of Orthoptera Insects With Different Wing Types." *Molecular Phylogenetics and Evolution* 145: 106734. <https://doi.org/10.1016/j.ympev.2020.106734>.
- Charlesworth, B., J. A. Coyne, and N. H. Barton. 1987. "The Relative Rates of Evolution of Sex-Chromosomes and Autosomes." *American Naturalist* 130, no. 1: 113–146. <https://doi.org/10.1086/284701>.
- Charlesworth, B., M. Nordborg, and D. Charlesworth. 1997. "The Effects of Local Selection, Balanced Polymorphism and Background Selection on Equilibrium Patterns of Genetic Diversity in Subdivided Populations." *Genetics Research* 70, no. 2: 155–174. <https://doi.org/10.1017/S0016672397002954>.
- Cruickshank, T. E., and M. W. Hahn. 2014. "Reanalysis Suggests That Genomic Islands of Speciation Are Due to Reduced Diversity, Not Reduced Gene Flow." *Molecular Ecology* 23, no. 13: 3133–3157. <https://doi.org/10.1111/Mec.12796>.
- Dasmahapatra, K. K., J. R. Walters, A. D. Briscoe, et al. 2012. "Butterfly Genome Reveals Promiscuous Exchange of Mimicry Adaptations Among Species." *Nature* 487, no. 7405: 94–98. <https://doi.org/10.1038/Nature11041>.
- Dopman, E. B., L. Perez, S. M. Bogdanowicz, and R. G. Harrison. 2005. "Consequences of Reproductive Barriers for Genealogical Discordance in the European Corn Borer." *Proceedings of the National Academy of Sciences of the United States of America* 102, no. 41: 14706–14711. <https://doi.org/10.1073/pnas.0502054102>.
- Dutheil, J. Y., K. Munch, K. Nam, T. Mailund, and M. H. Schierup. 2015. "Strong Selective Sweeps on the X Chromosome in the Human-Chimpanzee Ancestor Explain Its Low Divergence." *PLoS Genetics* 11, no. 8: e1005451. <https://doi.org/10.1371/journal.pgen.1005451>.
- Feder, J. L., and P. Nosil. 2010. "The Efficacy of Divergence Hitchhiking in Generating Genomic Islands During Ecological Speciation." *Evolution* 64, no. 6: 1729–1747. <https://doi.org/10.1111/j.1558-5646.2010.00943.x>.
- Fontaine, M. C., J. B. Pease, A. Steele, et al. 2015. "Mosquito Genomics. Extensive Introgression in a Malaria Vector Species Complex Revealed by Phylogenomics." *Science* 347, no. 6217: 1258524. <https://doi.org/10.1126/science.1258524>.
- Gainey, D. P., J. Y. Kim, and L. S. Maroja. 2018. "Mapping Reduced Introgression Loci to the X Chromosome of the Hybridizing Field Crickets, *Gryllus firmus* and *G. pennsylvanicus*." *PLoS One* 13, no. 12: e0208498. <https://doi.org/10.1371/journal.pone.0208498>.
- Garrigan, D., S. B. Kingan, A. J. Geneva, et al. 2012. "Genome Sequencing Reveals Complex Speciation in the *Drosophila simulans* Clade." *Genome Research* 22, no. 8: 1499–1511. <https://doi.org/10.1101/gr.130922.111>.
- Garrigan, D., S. B. Kingan, A. J. Geneva, J. P. Vedanayagam, and D. C. Presgraves. 2015. "Genome Diversity and Divergence in *Drosophila mauritiana*: Multiple Signatures of Faster X Evolution." *Genome Biology and Evolution* 7, no. 5: 1331. <https://doi.org/10.1093/gbe/evv029>.
- Gompert, Z., and C. A. Buerkle. 2010. "INTROGRESS: A Software Package for Mapping Components of Isolation in Hybrids." *Molecular Ecology Resources* 10: 378–384.
- Good, J. M., M. D. Dean, and M. W. Nachman. 2008. "A Complex Genetic Basis to X-Linked Hybrid Male Sterility Between Two Species of House Mice." *Genetics* 179, no. 4: 2213–2228. <https://doi.org/10.1534/genetics.107.085340>.
- Harr, B. 2006. "Genomic Islands of Differentiation Between House Mouse Subspecies." *Genome Research* 16, no. 6: 730–737. <https://doi.org/10.1101/Gr.5045006>.
- Harrison, R. G. 1983. "Barriers to Gene Exchange Between Closely Related Cricket Species. I. Laboratory Hybridization Studies." *Evolution* 37, no. 2: 245–251.
- Harrison, R. G. 1985. "Barriers to Gene Exchange Between Closely Related Cricket Species. II. Life Cycle Variation and Temporal Isolation." *Evolution* 39, no. 2: 244–259.
- Harrison, R. G. 1990. "Hybrid Zones: Windows on Evolutionary Processes." In *Oxford Surveys in Evolutionary Biology*, edited by D. Futuyma and J. Antonovics, vol. 7, 69–128. Oxford University Press. ISBN-13 : 978-0195062892.
- Harrison, R. G., and S. M. Bogdanowicz. 1997. "Patterns of Variation and Linkage Disequilibrium in a Field Cricket Hybrid Zone." *Evolution* 51, no. 2: 493–505.
- Heggeseth, B., D. Sim, L. Partida, and L. S. Maroja. 2020. "Influence of Female Cuticular Hydrocarbon (CHC) Profile on Male Courtship Behavior in Two Hybridizing Field Crickets *Gryllus firmus* and *Gryllus*

- pennsylvanicus*." *BMC Evolutionary Biology* 20, no. 1: 21. <https://doi.org/10.1186/s12862-020-1587-9>.
- Hey, J., Y. Chung, A. Sethuraman, et al. 2018. "Phylogeny Estimation by Integration Over Isolation With Migration Models." *Molecular Biology and Evolution* 35, no. 11: 2805–2818. <https://doi.org/10.1093/molbev/msy162>.
- Hirase, S., Y. Y. Yamasaki, M. Sekino, et al. 2021. "Genomic Evidence for Speciation With Gene Flow in Broadcast Spawning Marine Invertebrates." *Molecular Biology and Evolution* 38, no. 11: 4683–4699. <https://doi.org/10.1093/molbev/msab194>.
- Hu, X. S., and D. A. Filatov. 2016. "The Large-X Effect in Plants: Increased Species Divergence and Reduced Gene Flow on the Silene X-Chromosome." *Molecular Ecology* 25, no. 11: 2609–2619. <https://doi.org/10.1111/mec.13427>.
- Huelsenbeck, J. P., and F. Ronquist. 2001. "MRBAYES: Bayesian Inference of Phylogenetic Trees." *Bioinformatics* 17: 754–755.
- Korf, I. 2004. "Gene finding in novel genomes." *BMC Bioinformatics* 5: 59. <https://doi.org/10.1186/1471-2105-5-59>.
- Kumar, S., G. Stecher, M. Li, C. Knyaz, and K. Tamura. 2018. "MEGA X: Molecular Evolutionary Genetics Analysis Across Computing Platforms." *Molecular Biology and Evolution* 35, no. 6: 1547–1549. <https://doi.org/10.1093/molbev/msy096>.
- Larson, E. L., J. A. Andres, S. M. Bogdanowicz, and R. G. Harrison. 2013. "Differential Introgression in a Mosaic Hybrid Zone Reveals Candidate Barrier Genes." *Evolution* 67, no. 12: 3653–3661. <https://doi.org/10.1111/Evo.12205>.
- Larson, E. L., C. G. Becker, E. R. Bondra, and R. G. Harrison. 2013. "Structure of a Mosaic Hybrid Zone Between the Field Crickets *Gryllus firmus* and *G. pennsylvanicus*." *Ecology and Evolution* 3, no. 4: 985–1002. <https://doi.org/10.1002/Ece3.514>.
- Larson, E. L., G. L. Hume, J. A. Andres, and R. G. Harrison. 2012. "Post-Mating Prezygotic Barriers to Gene Exchange Between Hybridizing Field Crickets." *Journal of Evolutionary Biology* 25, no. 1: 174–186. <https://doi.org/10.1111/j.1420-9101.2011.02415.x>.
- Larson, E. L., T. A. White, C. L. Ross, and R. G. Harrison. 2014. "Gene Flow and the Maintenance of Species Boundaries." *Molecular Ecology* 23, no. 7: 1668–1678. <https://doi.org/10.1111/Mec.12601>.
- Lim, H. C., V. R. Vickery, and D. K. Mckevean. 1973. "Cytogenetic Studies in Relation to Taxonomy Within Family Gryllidae (Orthoptera). 1. Subfamily Gryllinae." *Canadian Journal of Zoology-Revue Canadienne De Zoologie* 51, no. 2: 179–186. <https://doi.org/10.1139/Z73-026>.
- Lynch, M., and T. J. Crease. 1990. "The Analysis of Population Survey Data on DNA Sequence Variation." *Molecular Biology and Evolution* 7: 377–394.
- Machado, C. A., and J. Hey. 2003. "The Causes of Phylogenetic Conflict in a Classic *Drosophila* Species Group." *Proceedings of the Royal Society B: Biological Sciences* 270, no. 1520: 1193–1202. <https://doi.org/10.1098/rspb.2003.2333>.
- Malinsky, M., R. J. Challis, A. M. Tyers, et al. 2015. "Genomic Islands of Speciation Separate Cichlid Ecomorphs in an East African Crater Lake." *Science* 350, no. 6267: 1493–1498. <https://doi.org/10.1126/science.aac9927>.
- Maroja, L. S., J. A. Andres, and R. G. Harrison. 2009. "Genealogical Discordance and Patterns of Introgression and Selection Across a Cricket Hybrid Zone." *Evolution* 63, no. 11: 2999–3015. <https://doi.org/10.1111/j.1558-5646.2009.00767.x>.
- Maroja, L. S., J. A. Andres, J. R. Walters, and R. G. Harrison. 2009. "Multiple Barriers to Gene Exchange in a Field Cricket Hybrid Zone." *Biological Journal of the Linnean Society* 97, no. 2: 390–402. <https://doi.org/10.1111/j.1095-8312.2009.01201.x>.
- [Genbank dataset] Maroja, L. S., F. Barradale, S. A. Espinoza-Ulloa, S. Bogdanowicz, and J. Andres. 2023. "Patterns of Genomic Heterogeneity Do Not Support the Islands of Differentiation Model in a Classic Mosaic Hybrid Zone." Genbank, Access Numbers: OR441320–OR447380.
- [Dryad dataset] Maroja, L. S., F. Barradale, S. A. Espinoza-Ulloa, S. Bogdanowicz, and J. Andres. 2024. "*Gryllus* Genome Files and Additional Information." Dryad. <https://doi.org/10.5061/dryad.s4mw6m9bb>.
- Maroja, L. S., E. L. Larson, S. M. Bogdanowicz, and R. G. Harrison. 2015. "Genes With Restricted Introgression in a Field Cricket (*Gryllus firmus*/*Gryllus pennsylvanicus*) Hybrid Zone Are Concentrated on the X Chromosome and a Single Autosome." *G3: Genes, Genomes, Genetics* 5, no. 11: 2219–2227. <https://doi.org/10.1534/g3.115.021246>.
- Maroja, L. S., Z. M. McKenzie, E. Hart, J. Jing, E. L. Larson, and D. P. Richardson. 2014. "Barriers to Gene Exchange in Hybridizing Field Crickets: The Role of Male Courtship Effort and Cuticular Hydrocarbons." *BMC Evolutionary Biology* 14: 65–75. <https://doi.org/10.1186/1471-2148-14-65>.
- Masly, J. P., and D. C. Presgraves. 2007. "High-Resolution Genome-Wide Dissection of the Two Rules of Speciation in *Drosophila*." *PLoS Biology* 5, no. 9: 1890–1898. <https://doi.org/10.1371/journal.pbio.0050243>.
- Meisel, R. P., and T. Connallon. 2013. "The Faster-X Effect: Integrating Theory and Data." *Trends in Genetics* 29, no. 9: 537–544. <https://doi.org/10.1016/j.tig.2013.05.009>.
- Nachman, M. W., and B. A. Payseur. 2012. "Recombination Rate Variation and Speciation: Theoretical Predictions and Empirical Results From Rabbits and Mice." *Philosophical Transactions of the Royal Society, B: Biological Sciences* 367, no. 1587: 409–421. <https://doi.org/10.1098/rstb.2011.0249>.
- Nadeau, N. J., A. Whibley, R. T. Jones, et al. 2012. "Genomic Islands of Divergence in Hybridizing *Heliconius* Butterflies Identified by Large-Scale Targeted Sequencing." *Philosophical Transactions of the Royal Society, B: Biological Sciences* 367, no. 1587: 343–353. <https://doi.org/10.1098/rstb.2011.0198>.
- Narang, P., and M. A. Wilson Sayres. 2016. "Variable Autosomal and X Divergence Near and Far From Genes Affects Estimates of Male Mutation Bias in Great Apes." *Genome Biology and Evolution* 8, no. 11: 3393–3405. <https://doi.org/10.1093/gbe/evw232>.
- Nosil, P., D. J. Funk, and D. Ortiz-Barrientos. 2009. "Divergent Selection and Heterogeneous Genomic Divergence." *Molecular Ecology* 18, no. 3: 375–402. <https://doi.org/10.1111/j.1365-294X.2008.03946.x>.
- Nosil, P., and D. Schluter. 2011. "The Genes Underlying the Process of Speciation." *Trends in Ecology & Evolution* 26, no. 4: 160–167. <https://doi.org/10.1016/j.tree.2011.01.001>.
- Payseur, B. A., J. G. Krenz, and M. W. Nachman. 2004. "Differential Patterns of Introgression Across the X Chromosome in a Hybrid Zone Between Two Species of House Mice." *Evolution* 58, no. 9: 2064–2078.
- Pons, J., I. Ribera, J. Bertranpetit, and M. Balke. 2010. "Nucleotide Substitution Rates for the Full Set of Mitochondrial Protein-Coding Genes in Coleoptera." *Molecular Phylogenetics and Evolution* 56, no. 2: 796–807. <https://doi.org/10.1016/j.ympev.2010.02.007>.
- Presgraves, D. C., and C. D. Meiklejohn. 2021. "Hybrid Sterility, Genetic Conflict and Complex Speciation: Lessons From the *Drosophila simulans* Clade Species." *Frontiers in Genetics* 12: 669045. <https://doi.org/10.3389/fgene.2021.669045>.
- Quilodran, C. S., K. Ruegg, A. T. Sendell-Price, E. C. Anderson, T. Coulson, and S. M. Clegg. 2020. "The Multiple Population Genetic and Demographic Routes to Islands of Genomic Divergence." *Methods in Ecology and Evolution* 11, no. 1: 6–21. <https://doi.org/10.1111/2041-210x.13324>.

- Rand, D. M., and R. G. Harrison. 1989. "Ecological Genetics of a Mosaic Hybrid Zone—Mitochondrial, Nuclear, and Reproductive Differentiation of Crickets by Soil Type." *Evolution* 43, no. 2: 432–449.
- Renaut, S., C. J. Grassa, S. Yeaman, et al. 2013. "Genomic Islands of Divergence Are Not Affected by Geography of Speciation in Sunflowers." *Nature Communications* 4: 1827. <https://doi.org/10.1038/ncomms2833>.
- Ross, C. L., and R. G. Harrison. 2002. "A Fine-Scale Spatial Analysis of the Mosaic Hybrid Zone Between *Gryllus firmus* and *Gryllus pennsylvanicus*." *Evolution* 56, no. 11: 2296–2312.
- Ross, C. L., and R. G. Harrison. 2006. "Viability Selection on Overwintering Eggs in a Field Cricket Mosaic Hybrid Zone." *Oikos* 115, no. 1: 53–68.
- Rozas, J., A. Ferrer-Mata, J. C. Sanchez-DelBarrio, et al. 2017. "DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets." *Molecular Biology and Evolution* 34, no. 12: 3299–3302. <https://doi.org/10.1093/molbev/msx248>.
- Saetre, G. P., T. Borge, K. Lindroos, et al. 2003. "Sex Chromosome Evolution and Speciation in *Ficedula* Flycatchers." *Proceedings of the Royal Society B: Biological Sciences* 270, no. 1510: 53–59. <https://doi.org/10.1098/rspb.2002.2204>.
- Sankararaman, S., S. Mallick, M. Dannemann, et al. 2014. "The Genomic Landscape of Neanderthal Ancestry in Present-Day Humans." *Nature* 507, no. 7492: 354–357. <https://doi.org/10.1038/nature12961>.
- Shen, W., S. Le, Y. Li, and F. Hu. 2016. "SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation." *PLoS One* 11: e0163962. <https://doi.org/10.1371/journal.pone.0163962>.
- Smit, A., and R. Hubley. 2008–2015. "RepeatModeler Open-1.0." <http://repeatmasker.org>.
- Smit, A. F. A., R. Hubley, and P. Green. 2013–2015. "RepeatMasker Open-4.0. Version 4.0.9."
- Stanke, M., and S. Waack. 2003. "Gene Prediction With a Hidden Markov Model and a New Intron Submodel." *Bioinformatics* 19, no. Suppl 2: ii215–ii225. <https://doi.org/10.1093/bioinformatics/btg1080>.
- Strasburg, J. L., N. A. Sherman, K. M. Wright, L. C. Moyle, J. H. Willis, and L. H. Rieseberg. 2012. "What Can Patterns of Differentiation Across Plant Genomes Tell Us About Adaptation and Speciation?" *Philosophical Transactions of the Royal Society, B: Biological Sciences* 367, no. 1587: 364–373. <https://doi.org/10.1098/rstb.2011.0199>.
- Sun, N., L. Yang, F. Tian, et al. 2022. "Sympatric or Micro-Allopatric Speciation in a Glacial Lake? Genomic Islands Support Neither." *National Science Review* 9, no. 12: nwac291. <https://doi.org/10.1093/nsr/nwac291>.
- Turelli, M., and H. A. Orr. 2000. "Dominance, Epistasis and the Genetics of Postzygotic Isolation." *Genetics* 154, no. 4: 1663–1679.
- Turner, T. L., and M. W. Hahn. 2010. "Genomic Islands of Speciation or Genomic Islands and Speciation?" *Molecular Ecology* 19, no. 5: 848–850. <https://doi.org/10.1111/j.1365-294X.2010.04532.x>.
- Turner, T. L., M. W. Hahn, and S. V. Nuzhdin. 2005. "Genomic Islands of Speciation in *Anopheles gambiae*." *PLoS Biology* 3, no. 9: 1572–1578. <https://doi.org/10.1371/journal.pbio.0030285>.
- Via, S., and J. West. 2008. "The Genetic Mosaic Suggests a New Role for Hitchhiking in Ecological Speciation." *Molecular Ecology* 17, no. 19: 4334–4345. <https://doi.org/10.1111/j.1365-294X.2008.03921.x>.
- Vicoso, B., and B. Charlesworth. 2006. "Evolution on the X Chromosome: Unusual Patterns and Processes." *Nature Reviews Genetics* 7, no. 8: 645–653. <https://doi.org/10.1038/nrg1914>.
- Weisenfeld, N. I., V. Kumar, P. Shah, D. M. Church, and D. B. Jaffe. 2017. "Direct Determination of Diploid Genome Sequences." *Genome Research* 27, no. 5: 757–767. <https://doi.org/10.1101/gr.214874.116>.
- White, B. J., C. Cheng, F. Simard, C. Costantini, and N. J. Besansky. 2010. "Genetic Association of Physically Unlinked Islands of Genomic Divergence in Incipient Species of *Anopheles gambiae*." *Molecular Ecology* 19, no. 5: 925–939. <https://doi.org/10.1111/j.1365-294X.2010.04531.x>.
- Willett, C. S., M. J. Ford, and R. G. Harrison. 1997. "Inferences About the Origin of a Field Cricket Hybrid Zone From a Mitochondrial DNA Phylogeny." *Heredity* 79: 484–494.
- Woerner, A. E., M. P. Cox, and M. F. Hammer. 2007. "Recombination-Filtered Genomic Datasets by Information Maximization." *Bioinformatics* 23, no. 14: 1851–1853. <https://doi.org/10.1093/bioinformatics/btm253>.
- Wu, C. I. 2001. "The Genic View of the Process of Speciation." *Journal of Evolutionary Biology* 14, no. 6: 851–865. <https://doi.org/10.1046/j.1420-9101.2001.00335.x>.
- Yeaman, S., and M. C. Whitlock. 2011. "The Genetic Architecture of Adaptation Under Migration-Selection Balance." *Evolution* 65, no. 7: 1897–1911. <https://doi.org/10.1111/j.1558-5646.2011.01269.x>.
- Ylla, G., T. Nakamura, T. Itoh, et al. 2021. "Insights Into the Genomic Evolution of Insects From Cricket Genomes." *Communications Biology* 4, no. 1: 733. <https://doi.org/10.1038/s42003-021-02197-9>.
- You, F. M., N. Huo, Y. Q. Gu, et al. 2008. "BatchPrimer3: A High Throughput Web Application for PCR and Sequencing Primer Design." *BMC Bioinformatics* 9: 253. <https://doi.org/10.1186/1471-2105-9-253>.