DEMOGRAPHY AND SELECTION SHAPE TRANSCRIPTOMIC DIVERGENCE IN FIELD CRICKETS

Running title: evolutionary forces shaping genetic variation

Keywords: gene flow, mating behavior, *Gryllus*

Thomas Blankers[1,2,3], Sibelle T. Vilaça[4,5], Isabelle Waurick[2], David A. Gray[6], R. Matthias Hennig[1], Camila J. Mazzoni[4,5], Frieder Mayer[2,7] , Emma L. Berdan[2,8]

[1] Behavioural Physiology, Department of Biology, Humboldt-Universität zu Berlin, Berlin, Germany, D-10115; Corresponding author: thomasblankers@gmail.com

[2] Museum für Naturkunde Berlin, Leibniz Institute for Evolution and Biodiversity Science, Berlin, Germany, D-10115

[3] Current Address: Department of Neurobiology and Behavior, Cornell University, Ithaca, NY, USA, 14853

[4]Berlin Center for Genomics in Biodiversity Research (BeGenDiv), Berlin, Germany, D-14195

[5]Leibniz-Institut für Zoo- und Wildtierforschung (IZW), Berlin, Germany, D-10315

[6] Department of Biology, California State University Northridge, Northridge, CA, USA, 91330-8303

[7] Berlin-Brandenburg Institute of Advanced Biodiversity Research (BBIB), Berlin, Germany, D-14195

[8] Current Address: Department of Marine Sciences, University of Gothenburg, Gothenburg, Sweden, SE - 405 30

ACKNOWLEDGEMENTS

We thank Marie Jeschek and Harald Detering from the Berlin Center for Genomics and Biodiversity (BeGenDiv) for assistance in bioinformatics. We further thank Mike Ritchie, Roger Butlin, Daniel Wegmann, and Laurent Excoffier for helpful comments on the analyses and the Biostars and Stackoverflow community for help with scripting. The manuscript strongly benefited from the comments from the associate editor Jeffry Dudycha and three anonymous reviewers. Sample collection and processing comply with the "Principles of animal care", publication No. 86-23, revised 1985 of the National Institute of Health, and also with the current laws of Germany. The authors declare no conflict of interest. This study is part of the GENART project. The work was supported by the Leibniz Association (SAW-2012-MfN-3). STV was supported by an Alexander von Humboldt Foundation fellowship.

AUTHOR CONTRIBUTIONS

T.B., C.J.M, F.M, and E.L.B designed the study. T.B. and D.A.G collected the samples and I.W. did the lab work. T.B, S.T.V., and E.L.B. analysed the data. T.B and E.L.B. wrote the manuscript with contributions from R.M.H, D.A.G, and S.T.V.

DATA ACCESSIBILITY

Data, including raw reads, sequences used for demographic analyses and SNP data files used in outlier analysis, will be made available on Dryad and the NCBI SRA archive prior to publication.

ABSTRACT

Gene flow, demography, and selection can result in similar patterns of genomic variation and disentangling their effects is key to understanding speciation. Here, we assess transcriptomic variation to unravel the evolutionary history of *Gryllus rubens* and *Gryllus texensis*, cryptic field cricket species with highly divergent mating behavior. We infer their demographic history and screen their transcriptomes for footprints of selection in the context of the inferred demography. We find strong support for a long history of bidirectional gene flow, which ceased during the late Pleistocene, and a bottleneck in *G. rubens* consistent with a peripatric origin of this species. Importantly, the demographic history has likely strongly shaped patterns of neutral genetic differentiation (empirical F_{ST} distribution). Concordantly, F_{ST} based selection detection uncovers a large number of outliers, likely comprising many false positives, echoing recent theoretical insights. Alternative genetic signatures of positive selection, informed by the demographic history of the sibling species, highlighted a smaller set of loci; many of these are candidates for controlling variation in mating behavior. Our results underscore the importance of demography in shaping overall patterns of genetic divergence and highlight that examining both demography and selection facilitates a more complete understanding of genetic divergence during speciation.

INTRODUCTION

The study of speciation and the origins of earth's biodiversity are at the core of evolutionary biology. An important first step is understanding the mechanisms that drive genetic divergence between closely related groups of organisms. In the age of next-generation sequencing, our understanding of these mechanisms is rapidly advancing. However, a variety of processes such as gene flow, local variation in recombination and mutation rates, linked or background selection, and divergent selection often simultaneously influence genetic variation between diverging lineages and the different processes may leave similar signatures in the genome (Noor and Bennett 2009; Feder et al. 2012; Nachman and Payseur 2012; Cutter and Payseur 2013; Seehausen et al. 2014; Burri et al. 2015). Therefore, to understand how populations diverge, how reproductive isolation evolves, and how this affects the genome, it is essential that we examine both selective and neutral processes.

Recently, the role of gene flow in speciation has drawn renewed attention (Smadja and Butlin 2011; Feder et al. 2013; Sousa and Hey 2013; Servedio 2015; Ravinet et al. 2017). It was once thought that reproductive barriers could only evolve in allopatry (Mayr 1963; Bolnick and Fitzpatrick 2007). However, this view has shifted due to accumulating evidence for varying rates of gene flow during early divergence (Bolnick and Fitzpatrick 2007; Nosil 2008; Bird et al. 2012). Although 'true' sympatric speciation is likely rare, there is nowadays a general acceptance that some amount of gene flow occurs during many speciation events, i.e. parapatric speciation (Coyne and Orr 2004; Smadja and Butlin 2011; Arnold 2015).

Speciation with gene flow has attracted special attention because strong divergent selection in combination with high migration rates may lead to higher (than background) genomic divergence in the regions harboring loci important for reproductive isolation and local adaptation (Turner et al. 2005; Nosil et al. 2009; Cutter and Payseur 2013; Feder et al. 2013; Ravinet et al. 2017). However, variation in levels of divergence across the genome may also strongly depend on locally reduced intraspecific diversity due to demographic effects or variation in mutation and recombination rates (Nachman and Payseur 2012; Cruickshank and Hahn 2014; Burri et al. 2015). Additionally, the likelihood of detecting the effects of selection above background levels of genomic variation is highly dependent on the genetic architecture of the traits under selection (Jiggins and Martin 2017) and the strength of selection (Ortiz-Barrientos and James 2017). These caveats warrant caution in the interpretation of the results from genomic scans, especially without a detailed understanding of the behavioral ecology and evolutionary history of the study system (Ravinet et al. 2017). Thus, a primary goal of studies aiming to elucidate the effects of selection on genetic variation should be to consider patterns left by neutral or demographic processes that could occlude the genomic signature of selection.

Here, we bring this goal into practice by inferring demographic history and characterizing the resulting patterns of genetic variation in the absence of selection. We then use the resulting neutral expectation to inform our inference of putative signatures of selection. For this approach, we use transcriptomic data from two sexually isolated field cricket species, *Gryllus rubens* and *Gryllus texensis.* Given their current distributions (Fig. 1), it is likely that interspecific gene flow has played a dominant role in the evolutionary history of *G. texensis* and *G. rubens*; although, contemporary gene flow is unlikely based on lack of genetic or phenotypic evidence supporting hybridization in nature (Walker 1998; Gray and Cade 2000; Higgins and Waugaman 2004) and reinforcement (Gray and Cade 2000; Izzo and Gray 2004). Additionally, a mitochondrial study found evidence that suggests *G. rubens* has a peripatric origin from *G. texensis* (Gray et al. 2008) and thus divergence between *G. texensis* and *G. rubens* may be associated with a strong bottleneck for the latter but not the former species.

In addition to demographic processes selection also likely played a role in the divergence of *G. texensis* and *G. rubens*. There is striking variation in acoustic sexual communication behavior in this system, involving multiple traits that compose the wing-generated calling song produced by cricket males and corresponding female preferences. This implies a strong selective pressure on genes related to mating signals. Variation in the cricket mating song depends on (i) the morphology and resonant properties of the wings, (ii) neural networks called central pattern generators that control rhythmic wing movement, and (iii) neuromuscular properties of the muscles that affect the temporal rhythm of the song (reviewed in Gerhardt and Huber 2002). Similarly, song recognition and preference in females are controlled by a complex network of neurons and likely depend on properties of ion channels, in particular potassium channels mediating inhibitory effects (Hennig et al. 2014; Schoneich et al. 2015; Göpfert and Hennig 2016). If there is indeed a strong selective pressure on mating behavior in this system, selection signatures are expected to be biased towards gene products that affect the properties of muscles, neuromuscular junctions, and neurotransmitter activity related to rhythmic behaviors and perception, as well as mating behavior variation more broadly. Importantly, combining the inference of putative selection signatures with the demographic analyses allows us to interpret the perceived effects from selection in the appropriate historical context and make predictions about the joint effects from neutral and selective forces during population divergence.

MATERIALS & METHODS

Study system

Gryllus texensis and *G. rubens* are widely distributed across the southern Gulf and Mid-Atlantic States in North America, with a broad sympatric region from eastern Texas through western Florida (Fig. 1). Males are morphologically cryptic (Gray et al. 2008) and there is no documented ecological divergence (Gray 2011). However, females differ in the length of the ovipositor (Gray et al. 2001), which tentatively reflects ecological adaptation to different soil types (Bradford et al. 1993). In nature, divergence in acoustic signals and preferences is a strong premating barrier acting through both species-specific long-distance mate attraction songs (Walker 1998; Gray and Cade 2000; Blankers et al. 2015a) and close-range courtship songs (Gray 2005; Izzo and Gray 2011). Reproductive isolation is maintained in the zone of overlap, but there is no evidence for reproductive character displacement, indicating that reinforcement is unlikely to affect divergence in these species (Higgins and Waugaman 2004; Izzo and Gray 2004).

Sample collection

Animals were collected in the USA in Lancaster and Austin (TX; ca. 80 *G. texensis* females) and in Lake City and Ocala (FL; ca. 40 *G. rubens* females) in autumn 2013 (Fig. 1 black dots). Collected females, which are typically already inseminated in the field, were housed in containers in groups of up to 15 individuals with gravel substrate, shelter, and water and food *ad libitum*. Each container also contained a cup with vermiculite for oviposition. During two weeks, eggs were collected and transferred to new containers; hatchlings were then reared to adulthood. We used laboratory-raised offspring of the field-caught females between one and three weeks after their final molt rather than field-caught specimens to standardize rearing conditions across all samples. All animals (males and females) were played back an artificial stimulus resembling the conspecific male song for 10 minutes prior to sacrificing the animal. The rationale here was that one of our primary objectives was to look at genetic divergence in relation to mating behavior polymorphism. In case specific genes involved in female preference behavior were only expressed upon hearing a male song signal, this could potentially be overcome by a brief play back 30 – 120 minutes prior to RNA preservation. Stimulus play back occurred for females and males to standardize the RNA sampling method across sexes. Within two hours of stimulus presentation, we sacrificed the cricket, removed the gut and then preserved the body in RNAlater following the manufacturer's instructions; samples were then stored at -80 ºC until RNA isolation. A total of five males and five females were used from each of the two populations for each species (40 individuals in total; randomly sampled across containers when there were multiple containers for crickets from the same population). Total RNA extraction and directional, strand-specific Illumina library preparation were done as described in a recently published transcriptomic resource for *Gryllus rubens* (Berdan et al. 2016).

SNP calling

Raw reads were processed using Flexbar (Dodt et al. 2012) to remove sequencing primers, adapters, and low quality bases on the 3' end of the individually barcoded reads. Samples were mapped to the *G. rubens* reference transcriptome (Berdan et al. 2016) using Bowtie2 (Langmead and Salzberg 2012) with default parameters but specifying read groups to mark reads as belonging to a specific individual. Duplicate reads were marked using 'picard' (http://broadinstitute.github.io/picard). The Genome Analysis Toolkit (GATK, DePristo *et al.* 2011; Van der Auwera *et al.* 2013) was used to call genotypes with the GATK-module 'UnifiedGenotyper'(Van der Auwera et al. 2013). The variants were then filtered to only retain high quality SNPs based on the recommendations on the GATK website (https://gatkforums.broadinstitute.org/gatk/discussion/comment/30641, accessed on 05/05/2015) and as described in a previous study (Berdan et al. 2015). The minor allele frequency (MAF) cut-off was set at 0.025 (a minimum of two copies of the allele).

Our sampling design was optimized to standardize the conditions under which we stored RNA samples, but potentially introduced a bias towards collecting related individuals. This may affect both demographic inference and the summary statistics used to identify selective sweeps. To correct for the potential cryptic relatedness, we used the PLINK methods-of-moments approach (Purcell et al. 2007) implemented in the SNPrelate package (Zheng et al. 2012) in R (R Development Core Team 2016) to estimate kinship coefficients for all pairs based on the allele frequencies within each population sample. We excluded eight individuals that showed estimated kinship coefficients above 0.125 (half-sib level) with other individuals from their population, leaving 17 *G. texensis* and 15 *G. rubens* individuals for all downstream analyses.

The demographic history

We first tested whether the sampled populations show geographic genetic structure. We inspected allele frequency variation within and between species and populations using principal component analysis. We also ran STRUCTURE (Falush et al. 2003), once for each species separately and once combining the species, using a single SNP locus per contig (8,835 randomly drawn SNPs). We used the admixture model with sampling location as prior information. We ran STRUCTURE with an MCMC chain length of 100,000 and with a burn-in length of 10,000 for K=1 through K=5 (K=4 for the species-specific runs) with three repetitions for each K-value. Results were analyzed using STRUCTURE HARVESTER (Earl and vonHoldt 2012) using the log-likelihood to compare K=1 versus all other values for K and the delta K method (Evanno et al. 2005) to compare K=2 versus all higher values of K.

To investigate the demographic history of *G. rubens* and *G. texensis*, we used the approximate Bayesian computation framework (ABC, Beaumont *et al.* 2002). We used ABCsampler from the ABCtoolbox package (Wegmann et al. 2009) to simulate our data under different demographic scenarios in fastsimcoal v2.5.2.3 (Excoffier and Foll 2011; Excoffier et al. 2013) and to calculate summary statistics using arlsumstat v.3.5.1.3 in Arlequin v 3.5 (Excoffier and Lischer 2010). We performed the analysis using the sequences from 1000 randomly drawn contigs (not including contigs with zero SNPs), using fixed recombination and mutation rates (both 1e-8) and the same minor allele frequency cut-off for the simulated data as for the observed data (0.025). To ensure that a MAF of 0.025 for the observed data was still maintained after removing the potentially related individuals, we removed all singletons that arose *post* subsampling. We initially calculated all between population summary statistics supported by arlsumstat. Then, using partial least squares regression (PLS), we retained the summary statistics with the highest predictive power (i.e. those with high factor loadings on the PLS components that significantly increase the predictive power of parameter estimates) for demographic estimates: the between-species mean and standard deviation of the number of polymorphic sites, the number of private polymorphic sites, Tajima's D, and nucleotide diversity (π) in each species, as well as pairwise (between species) F_{ST} and π. All statistics were calculated as averages across contigs.

We compared four possible (groups of) models: a simple divergence model (DIV; 4 parameters for population sizes and the timing/magnitude of demographic events), three models involving gene flow (either continuous, ancient or recent gene flow/secondary contact; CGF, AGF, RGF; 6-7 parameters), three models involving a bottleneck (for either or both species; RB, TB, BB; 6-8 parameters), and a model combining the most likely gene flow and most likely bottleneck model (AGFRB; 9 parameters). We intentionally considered only relatively simple models with few parameters to avoid the risk of overparameterization (Csilléry et al. 2010). For each model, we ran 200,000 iterations to do model selection. Prior ranges for population sizes and time points were chosen on a log-uniform scale spanning across several orders of magnitude and for bottleneck size and migration rates on a uniform scale not overlapping zero (Table 1).

After simulating the scenarios, model selection and posterior predictive checks were performed in R. Because of their similarity, the three bottleneck models and the three gene flow models were treated as two groups of models that were first tested inter-se; the best model of each group was then tested against the DIV and best combined models. We first retained the 1% of the samples that had the smallest Euclidean distance between the summary statistics of the simulated data and the observed data ('1%

nearest posterior samples' from hereon) for each scenario separately. We then obtained a set of linear discriminants that maximized the distance among models within the nested categories (gene flow and presence of bottleneck). Next, posterior model probabilities were calculated based on these linear combinations of summary statistics using the 'postpr' function in the 'abc' package (Csilléry et al. 2012). The first two model selection steps were used to retain one gene flow and one bottleneck model with the highest posterior probability ('best model' from hereon). A third round of model selection was used to select among a simple divergence scenario (DIV), the best gene flow and bottleneck scenarios (AGF and RB, respectively; see Results), and a scenario combining the best gene flow and the best bottleneck scenario (AGFRB). Model selection was validated by performing leave-one-out cross validation with logistic regression using the 'cv4postpr' function. Here, one simulated sample, chosen at random from the posterior distribution, is left out and considered to be the "true" model while repeating the model selection step (with the remaining posterior samples) to evaluate the robustness of the model selection (Csilléry et al. 2012).

To estimate demographic parameters, we then ran 1,000,000 new simulations under the model(s) with the highest posterior probability. Posterior predictive checks were performed by calculating the predicted R^2 and root mean squared error prediction (RMSEP) using the 'pls' package (Mevik and Wehrens 2007). We also used the 'cv4abc' function from the 'abc' package to evaluate prediction error. We estimated the demographic parameters with the 'abc' function using non-linear regression and a tolerance rate of 0.05.

An important goal of this study was to assess the effects of demography, in particular the timing of gene flow, on the *patterns* of transcriptome-wide genetic variation (*e.g.* the F_{ST} distribution), rather than only on summary statistics. This will provide important insight into the extent to which loci that have evolved in the absence of selection are expected to confound the signatures of selection. We thus estimated a null distribution of the allele frequency spectrum (i.e. Tajima's D, Tajima 1989) under the best fitting demographic model (see below). In addition, for the 1% nearest posterior samples of the models simulating continuous, recent, and ancestral gene flow and the AGFRB model we obtained the simulated

F_{ST} distribution for each posterior sample. The median and variation of these distributions were then visually contrasted with the observed F_{ST} distribution.

The role of selection

To assess the role of selection in driving genetic divergence, we employ three approaches that differ in their sensitivity to distinguish signals of selection from the confounding effects from past demographic events. All else being equal, variation in allele frequencies between populations is expected to increase more rapidly in the presence of selection. However, the most common measure of the variance in allele frequencies among populations, F_{ST}, which is also a common test statistic to distinguish selected loci from the genomic background, has been criticized as a reliable indicator from various angles (e.g. Narum and Hess 2011; Cruickshank and Hahn 2014; Lotterhos and Whitlock 2014). Other methods may be better suited for detection of selected loci given strong demographic effects. For instance given sufficiently long divergence times and high levels primary or secondary gene flow, elevated sequence divergence (d_{xy}) may be expected to better contrast the regions harboring loci involved in reproductive isolation from the rest of the genome (Nachman and Payseur 2012; Cruickshank and Hahn 2014). Additionally, a recent selective sweep may increase between population differentiation and decrease within population diversity and shift allele frequency spectrum (AFS) towards a higher frequency of rare alleles. Although demographic effects may also shift the AFS, these effects can be modeled and taken into account. Here, we contrast an F_{ST} outlier scan (the "F_{ST} approach" from hereon) with two alternative methods that should be better suited to withstand demographic effects (hereafter "d_{xy} approach" and the "selective sweep approach", respectively).

We considered loci to be potentially under positive or divergent selection if they exceeded genomic background levels of (1) F_{ST}, (2) absolute sequence divergence (d_{xy}), or (3) frequencies of rare alleles (Tajima's D), low diversity (π), and high differentiation (F_{ST}). For the F_{ST} approach, we used the hierarchical island model (Slatkin and Voelm 1991) implemented in Arlequin (Excoffier et al. 2009; Excoffier and Lischer 2010). To accommodate the data to Arlequin's input file restrictions, we only considered SNPs with MAF > 5% (81,125 SNPs). We pooled the two *G. rubens* populations in one group and two *G. texensis* populations in another group and performed 100,000 simulations to establish the neutral expectations for the relationship between among population heterozygosity and F_{ST}. We considered all loci with F_{ST} higher than the 99th quantile for a given level of heterozygosity to be selection outliers.

For the d_{xy} approach and the selective sweep approach, we used VCFtools (Danecek et al. 2011) to calculate the following summary statistics: Tajima's D (Tajima 1989), nucleotide diversity π (Nei and Li 1979), and weighted F_{ST} (Weir and Cockerham 1984) in 1000 bp windows, and the absolute difference between the frequency of the major allele in the two species. We also calculated the average interspecific pairwise distance d_{xy} for each window as $d_{xy} = \pi/(1\text{-}F_{ST})$, where π is the mean of the nucleotide diversity across species and F_{ST} is the weighted mean F_{ST} (Hudson et al. 1992; note that this method is similar to the often used $d_{xy} = p_i(1\text{-}p_j) + p_j(1\text{-}p_i)$, with p_i and p_j are the major or minor allele frequencies in species *i* and *j*, averaged across windows, weighed by the number of SNPs). For the d_{xy} approach, we retained the top 1% contigs with respect to d_{xy} predicting that these loci have diverged relatively early in the evolutionary history and remained shielded from gene flow throughout. For the selective sweep approach, we retained all loci that had Tajima's D below the 5% lowest simulated Tajima's D values under the inferred demographic scenario and with values for π and F_{ST} in the lowest and highest 10%, respectively. As this approach uses intraspecific population genetic data, we retained sets of outlier loci for both species separately

For all sets of outliers we checked for enriched Gene Ontology terms using 'topGO' (Alexa and Rahnenfuhrer 2016), part of the Bioconductor toolkit in R. The GO annotation was obtained from the *G. rubens* reference transcriptome (Berdan et al. 2016), which used the GO mapping module in Blast2Go (Conesa et al. 2005). We limited our gene set enrichment to biological process terms only and used the parent-child algorithm (Grossmann et al. 2007) to correct the *P* values for the 'inheritance problem' (i.e., the problem that higher GO terms inherit annotations from more specific descendant terms leading to false positives). We considered any GO term significantly enriched if the false discovery rate (Benjamini and Hochberg 1995) associated with the corrected P-value was below 10%. To get a more detailed picture of the putative functions of a given outlier locus, we looked up the functional annotation for the corresponding predicted gene product *(i.e.* the homolog with the highest similarity) on Flybase (Gramates et al. 2017) if that locus had been annotated using the *Drosophila melanogaster* proteome (see Berdan et al 2016 for details regarding transcriptome annotation).

RESULTS

Transcriptomic divergence

We sequenced RNA from 40 individuals (20 *G. rubens* and 20 *G. texensis*) on a HiSeq 2000 (Illumina, San Diego, CA, USA) obtaining on average 51,046,578 100-bp reads per individual (range 37,887,468-72,304,968) at a sequencing depth of eight libraries per lane. Reads mapped to the *G. rubens* transcriptome at an average rate of 83.2% (Table S1). Mapping rates were not higher in *G. rubens* despite the use of the *G. rub*ens transcriptome (*G. rubens*: 83.0%; *G. texensis*: 83.0%; $P = 0.9968$), but females mapped at a significantly higher rate than males (86.0% versus 79.6%; $P < 0.0001$). At a MAF cut-off of 0.025 we found a total of 175,244 SNPs across 8835 contigs. The average transition-transversion ratio was 1.6:1. Nucleotide diversity (π) was similar among *G. rubens* ($\pi = 0.11$, $\sigma_\pi = 0.14$) and *G. texensis* ($\pi = 0.13$, $\sigma_\pi = 0.15$). Median *D* was 0.07 (first quantile: 0.05, third quantile 0.20) and 2.7% of the SNPs (4,828) were fixed between the species (Fig. 2A). Average Tajima's D was negative for both species, but the distribution across loci showed substantial variation (Fig. 2B, C).

The demographic history

We found no substantial evidence for genetic structure in the two populations considered within either species. The species axis was the predominant axis of variation among individuals in the Principal Component Analysis (23.93% of total SNP variation, Fig. S1A), followed by axes separating *G. texensis* (PC2, 6.13% and PC3, 4.60%) and *G. rubens* (PC4, 4.35%) individuals. Variation within species was not related to geographic locations from which the individuals were collected (Fig. S1B, C). STRUCTURE further supported the finding that neither of the species was strongly differentiated geographically. The optimal K equaled 2 when we ran STRUCTURE with both species included (Fig. S2). Examining population structure within species revealed weak evidence for population substructure in both species at K=2, but K = 1 was the most parsimonious given the spread in log-likelihoods across K-values (Fig. S2). These results are robust across different subsets of SNPs and sample sizes (Fig. S3).

To infer the role of gene flow and bottlenecks during the evolutionary history of *G. texensis* and *G. rubens*, we used a nested rejection procedure to select the best model out of eight different models varying in the presence and timing of bottlenecks and gene flow (Fig. 3). First, we compared the gene flow models with each other. The gene flow model with the highest posterior probability was the 'ancestral gene flow' model (AGF $P_{posterior} = 0.99$ versus continuous gene flow, CGF: $P_{posterior} < 0.01$, and recent gene flow, RGF: $P_{posterior} = 0.01$). Then we compared the bottleneck models with each other and found that the '*G. rubens* bottleneck' model had the highest posterior probability (RB $P_{posterior} = 0.67$ versus *G. texensis* bottleneck, TB: $P_{posterior} = 0.43$ and both bottleneck, BB: $P_{posterior} < 0.01$). We then combined these best models into a model with both ancestral gene flow and a bottleneck for *G. rubens* (AGFRB) and compared that model against a simple divergence model (DIV), the best gene flow model (AGF), and the best bottleneck model (RB). In this final model comparison, the combined model had the highest posterior probability (AGFRB: $P_{posterior} = 0.68$; AGF: $P_{posterior} = 0.22$; DIV: $P_{posterior} = 0.02$; RB: $P_{posterior} = 0.08$; Fig. 3, Fig. 4). Similar results were obtained using the full sample, including additional, but potentially related individuals: AGFRB: $P_{posterior} = 0.75$; AGF: $P_{posterior} = 0.21$; DIV: $P_{posterior} = 0.03$; RB: $P_{posterior} = 0.01$.

As posterior probabilities may differ even among very similar models, it is critical to evaluate statistical support for model choice. Overall, model choice was well supported. For each selection step, we used cross validation to verify that models can be distinguished by assuming one of the models is the 'true' model and then performing 1,000 independent model selection steps under that assumption. The accuracy with which the assumed 'true' model was chosen was high for the gene flow models (98%, 96%, and 52% for AGF, CGF, and RGF, respectively), bottleneck models (76%, 66%, and 71% of the time for RB, TB, and BB respectively), and the final model selection step (75%, 82%, 82%, 84% for DIV, AGF, RB, AGFRB, respectively). It is important to note that the AGFRB model had the highest support overall and final model selection was well supported, but there is overlap of the posterior distribution of the summary statistics in multivariate space between the AGF and AGFRB models (Fig. 4).

Because there was some overlap between the posteriors of AGF and AGFRB (Fig. 4), and AGFRB only differs from AGF in the addition of a bottleneck, both models were used for demographic parameter estimates. Divergence times were distributed rather widely in both the AGF and AGFRB scenario and posterior density distributions were widely overlapping., The median divergence time varied between 350,000 years ago (700,000 generations ago) for AGF and double that for AGFRB. The ancestral effective population size was estimated around 200,000, almost an order of magnitude higher than the model estimates for current effective population sizes in *G. rubens* (~31,000 for AGFRB and ~18,000 for AGF) and *G. texensis* (~60,000 and ~28,000; Table 1, Table S2, Fig. 6A). A bottleneck for *G. rubens* was estimated at 15% of the current effective population size (Table 1, Fig. 6C) and recovery to current population sizes was achieved around 50,000 years ago (Table 1, Fig. 6B). Ancestral gene flow was bidirectional (median $m = 0.18$ and $m = 0.27$ for gene flow from *G. texensis* into *G. rubens* and vice versa, respectively; Table 1, Fig. 6C) and ceased around 18,000 years ago (Table 1, Table S2, Fig. 6B). The parameter estimates for the main model, AGFRB, were robust to the inclusion of additional, but potentially related, individuals; the estimates for times and population sizes were slightly higher and the inclusion of more samples gave similar results but at slightly higher accuracy (narrower HPD interval, Table S3, Fig. S4).

Statistical support for parameter inference varied across demographic events. Overall, the observed summary statistics fell well within the range of the simulated multivariate summary statistics under the AGF and AGFRB models (Fig. 4) and 95% HPD intervals of the distributions were generally narrow (Fig. 6, Table 1). For some demographic parameters (current population sizes for *G. rubens* [N_{RUB}] and *G. texensis* [N_{TEX}], and time since cessation of gene flow [T_{ISO}] support was high ($R^2 > 0.81$; RMSEP < 0.44); for other parameters estimated error rates were appreciably higher (Table 1, Table S2).

We compared F_{ST} distributions simulated under the AGF, CGF, RGF, and AGFRB models with the observed F_{ST} distribution as a measure of the effect of demography on the patterns of transcriptome-wide genetic variation. We found that the observed distribution (red line in Fig. 5) closely matched the simulated distribution of the two models with ancestral gene flow for most parts, including the secondary peak at the highest F_{ST} bin ($0.95 < F_{ST} \leq 1.00$, Fig. 5C, D). In contrast, the observed F_{ST} distribution showed substantial mismatch with the recent and continuous gene flow models.

The role of selection

The F_{ST} approach gave by far the highest number of outlier contigs. There were 514 contigs (5.8% of contigs) that had at least one SNP designated as a selection outlier (99th quantile) in Arlequin's F_{ST} based hierarchical island method. There were no significantly (FDR < 10%) enriched Gene Ontology categories among the predicted gene products of these contigs and the most strongly enriched categories included mitochondrial processes, GTPase activity and cellular metabolism (Table S4, Table S5, Fig S5).

There were 80 contigs with d_{xy} values in the 99th percentile. The putative gene products corresponding to these 80 contigs were significantly (FDR < 10%) enriched for pheromone biosynthesis, hormone biosynthesis, mating behavior, and protein maturation (Table S4). Several of the most divergent loci match genes involved in *Drosophila melanogaster* sex pheromone pathways, such as *α-esterase* and *Desaturase1*, mushroom body development and neuromuscular synaptic targets, such as *S-lap1*, *tartan*, including those involved in flight muscle activity (*Stretchin-Mlck*), and acoustic mating behavior, such as *Juvenile hormone esterase* and *calmodulin* (Table S6).

We retained 55 and 92 contigs that showed possible signatures of recent selective sweeps (Tajima's D below 5% of the simulated sequences under the AGFRB scenario and π and F_{ST} in the 90th percentile) in *G. texensis* and *G. rubens*, respectively. The combined set of outlier loci was not significantly enriched for any biological processes after FDR correction. The most strongly enriched GO terms were predominantly higher order GO terms such as 'organelle organization', 'primary metabolic process', and 'regulation of biological process', but also contained more specific terms: 'sperm mitochondrion organization', 'oocyte fate determination', and 'regulation of female receptivity' (Table S4). Six contigs were shared between the species-specific sets of loci that showed potential signatures of a recent selective sweep signature. Three of these have no functionally characterized gene products. The other three are *neuroglian* (nrg), which is involved in various aspects of nervous system development and associated with male and female courtship behavior in *D. melanogaster*; *discs large 1* (dlg1), which affects neuromuscular junctions and changes fruit fly behavior across several domains including circadian activity and courtship; and *secretory 23* (sec23), which is an important component in differentiation of extra-cellular membranes in neurons and epithelial cells (Table S7). Several other gene products associated with contigs in the species- specific sets have functional roles in calcium or potassium channel activity (e.g., *nervana2*, expressed in the *Drosophila* auditory organs), nervous system development (e.g. *muscleblind*, which also alters female receptivity during courtship), veined-wing song generation (e.g. *period*), as well as many genes related to metabolic and cellular processes.

There was one (unannotated) contig shared between the d_{xy} approach and the selective sweep approach. Additionally, among the 514 outlier loci detected in Arlequin encompassed 11 contigs also found with the d_{xy} approach, and 25 and 9 contigs respectively that were shared with the *G. rubens* and *G. texensis* specific selective sweep approach. These included the genes described above that are potentially related to sex pheromones biosynthesis (Desat1), flight muscle activity (Mlc-k), sensory neuron development (nrg), and auditory pathway ion channel activity (nrv2).

DISCUSSION

Here, we illuminate the role of demographic and selective processes in shaping genetic variation during speciation. Combined insight in putative neutral (neutral divergence given the demographic history) and selective effects allowed us to infer the evolutionary history of *Gryllus rubens* and *G. texensis*, sibling species with large, overlapping distributions and strong phenotypic divergence in sexual traits with limited divergence in other phenotypes. We find strong support for a long history of ancestral gene flow and a bottleneck in *G. rubens*. Importantly, our data lend support to the hypothesis that loci showing high relative genetic differentiation compared to the genomic background may have evolved in response to demographic events and drift rather than in response to election. Interestingly, several of the loci with show signatures of positive or divergent selection after taking into account the effects from demography are potential orthologs of *D. melanogaster* genes involved in premating isolation, a major source of reproductive isolation between *G. rubens* and *G. texensis*. This work represents an important first step in assessing the contribution of neutral and selective forces to genetic divergence in a model system for sexual selection research.

Neutral divergence and demography

We sequenced the transcriptomes of 40 individuals across four populations. Our observed transition:transversion ratio of 1.6:1 compares well with the estimate (1.55) from another cricket species pair, *G. firmus* and *G. pennsylvanicus* (Andrés et al. 2013), and suggests that sequencing errors did not contribute unduly to SNP discovery. Divergence across ~175K SNPs showed a bimodal and slightly right-skewed distribution of absolute (allele frequency) divergence, *D* (Fig. 2), and genetic differentiation, F_{ST} (Fig. 5). The F_{ST} distributions simulated under our top two scenarios were also right-skewed and strongly resembled the observed distribution of genetic differentiation, in strong contrast to F_{ST} distributions corresponding to other models. Most importantly the simulated distributions under the most likely demographic scenarios, AGF and AGFRB, showed secondary peaks at $F_{ST} > 0.95$. This indicates that a significant proportion of our fixed loci may have risen to fixation stochastically due to neutral processes (a combination of drift, population size variation, and gene flow) while gene flow homogenizes other (random) parts of the genome. Concordantly, the F_{ST} based approach uncovered substantially more loci with putative signatures of positive selection than methods based on allele frequency spectra (with thresholds informed by inferred demographic history) or absolute sequence divergence (514 contigs in the F_{ST} approach versus ~50-90 contigs in the other approaches). Our findings emphasizes the shortcomings of traditional F_{ST} outlier approaches to discern selection effects from genomic background variation (Narum and Hess 2011; Lotterhos and Whitlock 2014).

We find strong evidence for a long history of bidirectional gene flow before *G. rubens* and *G. texensis* became fully reproductively isolated around 18,000 years ago, sometime during the last Pleistocene glacial cycles. This finding adds to a growing body of work that suggest divergence can occur in the face of gene flow (Bolnick and Fitzpatrick 2007; Nosil 2008; Bird et al. 2012; Feder et al. 2013). A large amount of recent work has focused on the role of gene flow in speciation, especially in combination with divergent or positive selection. In the genic view of speciation (Wu 2001) most areas of the genome are homogenized among populations during divergence with gene flow, and regions showing excess differentiation are thus likely protected by selection. This idea has been tested in many model systems with mixed results (Turner et al. 2005; Ellegren et al. 2012; Nosil et al. 2012; Cruickshank and Hahn 2014; Burri et al. 2015; Marques et al. 2016). Recent work suggests that genomic mosaics may in fact be mostly a consequence of linked selection caused by differences in recombination rates and density of selected loci and are thus expected to be conserved in pairwise comparisons even among distantly related taxa (Nachman and Payseur 2012; Burri et al. 2015; Van Doren et al. 2017). Our results support this idea as our demographic simulations recreated heterogeneous patterns similar to our observed data. Although selection certainly contributed to transcriptome divergence in *G. rubens* and *G. texensis* our results suggest a larger role for neutral divergence shaped by the effects of migration and population size variation and echo recent insights into the importance of considering neutral divergence when interpreting potential selection effects (*e.g.* reviewed in Ravinet et al. 2017).

In addition to bi-directional gene flow, the early stages of divergence between *G. texensis* and *G. rubens* were also influenced by a substantial bottleneck in *G. rubens*. There is some overlap between the AGF (no bottleneck) and AGFRB (with a *G. rubens* bottleneck) scenarios in the simulated summary statistic distribution, but the latter has a substantially higher posterior probability and corroborates the peripatric origin for *G. rubens* hypothesized in a previous study (Gray et al. 2008). Although that study used a single mitochondrial locus, it was done with extensive geographic sampling, and both studies suggest a bottleneck for *G. rubens*. Furthermore, estimates of strong admixture between populations within species and divergence time estimates are overlapping (this study: median ~ 0.35 - 0.70 million years ago; Gray *et al*. study: 0.25 – 2.0 mya). Estimates for current effective population sizes (roughly between 30 and 60 thousand for the AGFRB model and between 20 and 30 thousand for the AGF model) are surprisingly low given the potential census population size for *G. texensis* is in the millions (Gray et al. 2008). Potentially, the discrepancy is due to recent population expansion (Ptak and Przeworski 2002; Nadachowska-brzyska et al. 2013) or variation in individual mating success (Lande and Barrowclough 1987), as is observed in wild populations of closely related species (Ritz and Köhler 2010; Rodriguez-Munoz et al. 2010).

The role of selection

A central aim of this study was to elucidate the role of selection during divergence within the context of the inferred demographic history. The species have strongly divergent mating behaviors with no evidence for reinforcement (Gray and Cade 2000; Higgins and Waugaman 2004; Izzo and Gray 2004; Blankers et al. 2015a). Many other cricket species show similarly strong divergence in various aspects of their mating behavior and several lines of evidence from various taxa indicate that this is at least in part driven by selection (Gray and Cade 2000; Bentsen et al. 2006; Shaw et al. 2007; Bailey 2008; Thomas and Simmons 2009; Oh and Shaw 2013; Blankers et al. 2017; Pascoal et al. 2017). Here, we show that the striking behavioral divergence is to some extent reflected in elevated sequence divergence of loci with putative functions in acoustic and chemical mating behavior. We find evidence that the set of loci showing the highest levels of sequence divergence are enriched for contigs bearing significant similarity to genes with known function in mating behavior in *D. melanogaster*. In addition, among the six contigs that showed evidence for a selective sweep in both species, three are potential orthologs of genes that affect neuromuscular properties in fruit flies and have effects on the flies' mating behavior. Several other species-specific outliers are potential orthologs of genes that can be tied to mating behavior variation in *Drosophila spp.*

Given the substantial time since divergence and the long history of gene flow, high sequence divergence is expected for loci that have experienced limited homogenizing effects from gene flow relative to the rest of the genome. The theoretical support for speciation with gene flow driven by divergence in secondary sexual characters is very thin at best (van Doorn et al. 2004; Weissing et al. 2011; Servedio 2015). Here we provide exciting and rare evidence for speciation with primary gene flow while both phenotypic (Gray and Cade 2000), quantitative genetic (Blankers et al. 2015b, 2017), and genomic analyses (this study) highlight a role for selection on (acoustic) mating behavior in driving reproductive isolation. A compelling alternative interpretation of the findings here is that the peripatric origin of *G. rubens* has allowed for an initial phase of reduced gene flow; during this phase mating signals and preferences may have diverged sufficiently (aided by a founder effect following a population bottleneck) to maintain reproductive isolation during a subsequent phase of range expansion culminating into the contemporary, widespread, and largely overlapping species' distributions. More empirical studies examining the role of gene flow and selection in systems characterized by strong sexual isolation are needed to test the theoretical predictions for speciation by sexual selection. However, this study along with other recent findings in finches (Campagna et al. 2017), fresh water stickleback (Marques et al. 2017), and cichlids (Malinksy et al. 2015) provide exciting first genomic insights into the joint effects from mating behavior divergence, sexual selection, and gene flow in the earliest phases of speciation.

We acknowledge that there are likely to be false positives among the detected outliers, as both linked (background) selection and demographic effects are expected to confound the signatures of positive or divergent selection (Cruickshank and Hahn 2014; Ravinet et al. 2017) and *a priori* expectations also increase the risk of "storytelling" (Pavlidis et al. 2012). By using coalescent simulations under the inferred evolutionary history, we have accounted for some confounding effects from demography. However, there is still potential neutral genetic variation that is unaccounted for, most notably the potentially confounding effects of recent population expansion and variation in recombination rates. We therefore caution that there is the uncertainty associated with the results obtained here and with genomic scans on quantitative traits in general (Jiggins and Martin 2017). Nevertheless, our findings provide exciting incentive for validation using alternative methods (e.g., QTL mapping) and follow-up functional genomic analyses.

Unsurprisingly, not all "outlier" contigs could be linked to mating behavior. The rest of these outliers are likely comprised of three groups: (1) Loci that are physically linked to loci under selection: In the earliest phases of speciation, only loci directly under strong divergent selection will differ. However, gene frequencies at tightly linked loci will also change and, given sufficient time as well as low to moderate migration and recombination rates, these loci will be swept to fixation along with selected sites (Smith and Haigh 1974) in a process called divergence hitchhiking (Feder et al. 2012; Via 2012); (2) Loci that are under selective forces that we have not yet elucidated: It is unlikely that divergent selection only targets loci involved in mating behavior and other traits may be differentiated between *G. rubens* and *G. texensis*. For example, females differ in the length of the ovipositor (Gray et al. 2001), a trait which reflects potential ecological adaptation to different soil types (Bradford et al. 1993); (3) Loci that are not under selection: Genetic drift can cause loci to drift to fixation and demographic effects such as bottlenecks and migration patterns (Holsinger and Weir 2009) can aid this process. Our simulations predict a significant number of fixed loci (1.90% on average for the AGFRB scenario) solely due to neutral processes (Fig. 5). Additionally, practical limitations of discovering low-frequency SNPs causing ascertainment bias (Clark et al. 2005) can contribute to misinterpretation of the patterns of genetic diversity (Vitti et al. 2013). A genomic map of *Gryllus* and further analyses would make strong headway into determining which of these categories the other potential outliers fall into.

Finally, there may be loci that are under selection but that were not detected by our scan because they simply were not being expressed. We sequenced samples from first generation laboratory offspring rather than animals directly from the field. Despite the fact that there are no differences between *G. texensis* and *G. rubens* in ecology, microhabitat use, or feeding behavior have been described (but note there is variation in the ovipositor length which is a potential adaptation to soil properties), the laboratory conditions have potentially limited our potential to detect genetic differences related to local adaptation.

In summary, this study underlines the importance of considering the joint effects from neutral divergence and selection in understanding the speciation process. Our results also offer unprecedented insight into the evolutionary history and the role of demography and selection in driving transcriptomic divergence in two sexually isolated field cricket sister species. We inferred that a long period of bidirectional, ancestral gene flow and a bottleneck in *G. rubens* preceded completion of reproductive isolation (Fig. 3,6). Importantly, the timing of gene flow appears to have significantly influenced the pattern of divergence (i.e. the F_{ST} distribution) that we observe (Fig. 5). We also uncovered several loci that show signatures of positive or divergent selection and show that these contigs are potentially associated with courtship behavior, neuromuscular development, and chemical mating behavior. Future work will place these data on a genomic map allowing us to determine how genetic divergence is distributed relative to loci under selection. These findings provide important steps towards understanding the role of selective and neutral processes in shaping patterns of divergence and the role of sexual selection during speciation-with-gene flow. They also highlight the strength of combining information on (i) the phenotypes that contribute to reproductive isolation, (ii) demographic inference, and (iii) scans for loci under selection.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

REFERENCES

Alexa, A., and J. Rahnenfuhrer. 2016. topGO: Enrichment Analysis for Gene Ontology. R package version 2.30.0

Andrés, 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 193:501–513.

Arnold, M. L. 2015. Divergence with genetic exchange. Oxford University Press.

Bailey, N. W. 2008. Love will tear you apart : different components of female choice exert contrasting selection pressures on male field crickets. Behav. Ecol. 19:960–966.

Beaumont, M. A., W. Zhang, and D. J. Balding. 2002. Approximate Bayesian Computation in Population Genetics. Genetics 162:2025–2035.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B. 57: 289–300.

Bentsen, C. L., J. Hunt, M. D. Jennions, and R. Brooks. 2006. Complex multivariate sexual selection on male acoustic signaling in a wild population of Teleogryllus commodus. Am. Nat. 167:E102–E116.

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 a Gryllus rubens transcriptome. Mol. Ecol. Resour. 16:1478–1490.

Berdan, E. L., C. J. Mazzoni, I. Waurick, J. T. Roehr, and F. Mayer. 2015. A population genomic scan in Chorthippus grasshoppers unveils previously unknown phenotypic divergence. Mol. Ecol. 24:3918–30.

Bird, C. E., I. Fernandez-Silva, D. J. Skillings, and R. J. Toonen. 2012. Sympatric Speciation in the Post "Modern Synthesis" Era of Evolutionary Biology. Evol. Biol. 39:158–180.

Blankers, T., D. A. Gray, and R. M. Hennig. 2017. Multivariate Phenotypic Evolution: Divergent Acoustic Signals and Sexual Selection in Gryllus Field Crickets. Evol. Biol. 44:43–55.

Blankers, T., R. M. Hennig, and D. A. Gray. 2015a. Conservation of multivariate female preference functions and preference mechanisms in three species of trilling field crickets. J. Evol. Biol. 28:630–641.

Blankers, T., A. K. Lübke, and R. M. Hennig. 2015b. Phenotypic variation and covariation indicate high evolvability of acoustic communication in crickets. J. Evol. Biol. 28:1656–69.

Bolnick, D. I., and B. M. Fitzpatrick. 2007. Sympatric Speciation : Models and Empirical Evidence. Annu. Rev. Ecol. Evol. Syst. 38:459–487.

Bradford, M. J., P. A. Guerette, and D. A. Roff. 1993. Testing hypotheses of adaptive variation in cricket ovipositor lengths. Oecologia 93:263–267.

Burri, R., A. Nater, T. Kawakami, C. F. Mugal, P. I. Olason, L. Smeds, A. Suh, L. Dutoit, S. Bures, L. Z. Garamszegi, S. Hogner, J. Moreno, A. Qvarnstrom, M. Ruzic, S. A. Saether, G. P. Saetre, J. Torok, and H. Ellegren. 2015. Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of Ficedula flycatchers. Genome Res. 25:1656–1665.

Campagna, L., M. Repenning, L. F. Silveira, C. S. Fontana, P. L. Tubaro, and I. J. Lovette. 2017. Repeated divergent selection on pigmentation genes in a rapid finch radiation. Sci. Adv. 3:e1602404.

Clark, A. G., M. J. Hubisz, C. D. Bustamante, S. H. Williamson, and R. Nielsen. 2005. Ascertainment bias in studies of human genome-wide polymorphism. Genome Res. 15:1496–1502.

Conesa, A., S. Götz, J. M. García-Gómez, J. Terol, M. Talón, and M. Robles. 2005. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674–3676.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer, Sunderland, MA.

Cruickshank, T. E., and M. W. Hahn. 2014. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. Mol. Ecol. 23:3133–3157.

Csilléry, K., M. G. B. Blum, O. E. Gaggiotti, and O. François. 2010. Approximate Bayesian Computation (ABC) in practice. Trends Ecol. Evol. 25:410–418.

Csilléry, K., O. François, and M. Blum. 2012. Approximate Bayesian Computation (ABC) in R: A Vignette. 202.162.217.53 1–21.

Cutter, A. D., and B. A. Payseur. 2013. Genomic signatures of selection at linked sites: unifying the disparity among species. Nat. Rev. Genet. 14:262–274.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, and 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. Bioinformatics. 27:2156–2158.

DePristo, M. A., E. Banks, R. Poplin, K. V Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernytsky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, and M. J. Daly. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 43:491–498.

Dodt, M., J. T. Roehr, R. Ahmed, and C. Dieterich. 2012. FLEXBAR—flexible barcode and adapter processing for next-generation sequencing platforms. Biology. 1:895–905.

Earl, D., and B. vonHoldt. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4:359–361.

Ellegren, H., L. Smeds, R. Burri, P. I. Olason, N. Backström, T. Kawakami, A. Künstner, H. Mäkinen, K. Nadachowska-Brzyska, A. Qvarnström, S. Uebbing, and J. B. W. Wolf. 2012. The genomic landscape of species divergence in Ficedula flycatchers. Nature 491: 756-760

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14:2611–2620.

Excoffier, L., I. Dupanloup, E. Huerta-Sánchez, V. C. Sousa, and M. Foll. 2013. Robust Demographic Inference from Genomic and SNP Data. PLoS Genet. 9: e1003905

Excoffier, L., and M. Foll. 2011. fastsimcoal: a continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. Bioinformatics 27:1332–1334.

Excoffier, L., T. Hofer, and M. Foll. 2009. Detecting loci under selection in a hierarchically structured population. Heredity. 103:285–298.

Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10:564–567.

Falush, D., M. Stephens, and J. K. Pritchard. 2003. Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. Genetics. 164:1567–1587.

Feder, J. L., S. P. Egan, and P. Nosil. 2012. The genomics of speciation-with-gene-flow. Trends Genet. 28:342–350.

Feder, J. L., S. M. Flaxman, S. P. Egan, A. A. Comeault, and P. Nosil. 2013. Geographic Mode of Speciation and Genomic Divergence. Annu. Rev. Ecol. Evol. Syst. 44:73–97.

Gerhardt, H. C., and F. Huber. 2002. Acoustic communication in insects and anurans. The University of Chicago Press, Chicago.

Göpfert, M. C., and R. M. Hennig. 2016. Hearing in Insects. Annu. Rev. Entomol. 61: 257-276

Gramates, L. S., S. J. Marygold, G. dos Santos, J.-M. Urbano, G. Antonazzo, B. B. Matthews, A. J. Rey, C. J. Tabone, M. A. Crosby, D. B. Emmert, K. Falls, J. L. Goodman, Y. Hu, L. Ponting, A. J. Schroeder, V. B. Strelets, J. Thurmond, and P. Zhou. 2017. FlyBase at 25: looking to the future. Nucleic Acids Res. 45:D663–D671.

Gray, D. A. 2005. Does courtship behavior contribute to species-level reproductive isolation in field crickets? Behav. Ecol. 16:201–206.

Gray, D. A. 2011. Speciation, divergence, and the origin of Gryllus rubens: behavior, morphology, and molecules. Insects 2:195–209.

Gray, D. A., H. Huang, and L. L. Knowles. 2008. Molecular evidence of a peripatric origin for two sympatric species of field crickets (*Gryllus rubens* and *G. texensis*) revealed from coalescent simulations and population genetic tests. Mol. Ecol. 17:3836–3855.

Gray, D. A., T. J. Walker, B. E. Conley, and W. H. Cade. 2001. A morphological means of distinguishing females of the cryptic field cricket species, Gryllus rubens and G. texensis (Orthoptera : Gryllidae). Florida Entomol. 84:314–315.

Gray, D., and W. Cade. 2000. Sexual selection and speciation in field crickets. Proc Natl. Acad. Sci. 97:14449–14454.

Grossmann, S., S. Bauer, P. N. Robinson, and M. Vingron. 2007. Improved detection of overrepresentation of Gene-Ontology annotations with parent--child analysis. Bioinformatics 23:3024–3031.

Hennig, R. M., K.-G. Heller, and J. Clemens. 2014. Time and timing in the acoustic recognition system of crickets. Front. Physiol. 5: 286-297

Higgins, L. a., and R. D. Waugaman. 2004. Sexual selection and variation: A multivariate approach to species-specific calls and preferences. Anim. Behav. 68:1139–1153.

Holsinger, K. E., and B. S. Weir. 2009. Genetics in geographically structured populations: defining, estimating and interpreting FST. Nat. Rev. Genet. 10:639–650.

Hudson, R. R., M. Slatkin, and W. P. Maddison. 1992. Estimation of levels of gene flow from DNA sequence data. Genetics 132:583–589.

Izzo, A. S., and D. a. Gray. 2011. Heterospecific courtship and sequential mate choice in sister species of field crickets. Anim. Behav. 81:259–264.

Izzo, A. S., and D. A. Gray. 2004. Cricket song in sympatry: Species specificity of song without reproductive character displacement in Gryllus rubens. Ann. Entomol. Soc. Am. 97:831–837.

Jiggins, C. D., and S. H. Martin. 2017. Glittering gold and the quest for Isla de Muerta. J. Evol. Biol. 30:1509–1511.

Lande, R. L., and G. F. Barrowclough. 1987. Effective population size, genetic variation, and their use in population management. Pp. 87–124 *in* M. E. Soule, ed. Viable populations for conservation. Cambridge University Press, Cambridge, NY.

Langmead, B., and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9:357–359.

Lotterhos, K. E., and M. C. Whitlock. 2014. Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. Mol. Ecol. 23:2178–2192.

Marques, D. A., K. Lucek, M. P. Haesler, A. F. Feller, J. I. Meier, C. E. Wagner, L. Excoffier, and O. Seehausen. 2017. Genomic landscape of early ecological speciation initiated by selection on nuptial colour. Mol. Ecol. 26:7–24.

Marques, D. A., K. Lucek, J. I. Meier, S. Mwaiko, C. E. Wagner, L. Excoffier, and O. Seehausen. 2016. Genomics of Rapid Incipient Speciation in Sympatric Threespine Stickleback. PLoS Genet. 12:1–34.

Mayr, E. 1963. Animal Species and Evolution. Harvard University Press.

Mevik, B.-H., and R. Wehrens. 2007. The pls Package: Principal Component and Partial Least Squares Regression in R. J. Stat. Softw. 18:1–24.

Nachman, M. W., and B. A. Payseur. 2012. Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. Philos. Trans. R. Soc. B Biol. Sci. 367:409–421.

Nadachowska-brzyska, K., R. Burri, P. I. Olason, T. Kawakami, and H. Ellegren. 2013. Demographic Divergence History of Pied Flycatcher and Collared Flycatcher Inferred from Whole-Genome Re-sequencing Data. PLoS Genet. 9:e1003942.

Narum, S. R., and J. E. Hess. 2011. Comparison of FST outlier tests for SNP loci under selection. Mol. Ecol. Resour. 11:184–194.

Nei, M., and W.-H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. 76:5269–5273.

Noor, M. A. F., and S. M. Bennett. 2009. Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. Heredity. 103:439–44.

Nosil, P. 2008. Speciation with gene flow could be common. Mol. Ecol. 17:2103–2106.

Nosil, P., D. J. Funk, and D. Ortiz-Barrientos. 2009. Divergent selection and heterogeneous genomic divergence. Mol. Ecol. 18:375–402.

Nosil, P., T. L. Parchman, J. L. Feder, and Z. Gompert. 2012. Do highly divergent loci reside in genomic regions affecting reproductive isolation? A test using next-generation sequence data in Timema stick insects. BMC Evol. Biol. 12:164-176

Oh, K. P., and K. L. Shaw. 2013. Multivariate sexual selection in a rapidly evolving speciation phenotype. Proc. Roy Soc - B Biol. Sci. 280:20130482.

Ortiz-Barrientos, D., and M. E. James. 2017. Evolution of recombination rates and the genomic landscape of speciation. J. Evol. Biol. 30:1519–1521.

Pascoal, S., M. Mendrok, A. J. Wilson, J. Hunt, and N. W. Bailey. 2017. Sexual selection and population divergence II. Divergence in different sexual traits and signal modalities in field crickets (Teleogryllus oceanicus). Evolution. 71:1614–1626.

Pavlidis, P., J. D. Jensen, W. Stephan, and A. Stamatakis. 2012. A critical assessment of storytelling: gene ontology categories and the importance of validating genomic scans. Mol. Biol. Evol. 29:3237–3248.

Ptak, S. E., and M. Przeworski. 2002. Evidence for population growth in humans is confounded by fine-scale population structure. Trends Genet. 18:559–563

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. De Bakker, M. J. Daly, and others. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81:559–575.

R Development Core Team, R. 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.

Ravinet, M., R. Faria, R. K. Butlin, J. Galindo, N. Bierne, M. Rafajlović, M. A. F. Noor, B. Mehlig, and A. M. Westram. 2017. Interpreting the genomic landscape of speciation: finding barriers to gene flow. J. Evol. Biol. 30:1450–1477.

Ritz, M. S., and G. Köhler. 2010. Natural and sexual selection on male behaviour and morphology, and female choice in a wild field cricket population: Spatial, temporal and analytical components. Evol. Ecol. 24:985–1001.

Rodriguez-Munoz, R., A. Bretman, J. Slate, C. A. Walling, and T. Tregenza. 2010. Natural and sexual selection in a wild insect population. Science. 79:1269–1272.

Schoneich, S., K. Kostarakos, and B. Hedwig. 2015. An auditory feature detection circuit for sound pattern recognition. Sci. Adv. 1:e1500325–e1500325.

Seehausen, O., R. K. Butlin, I. Keller, C. E. Wagner, J. W. Boughman, P. a Hohenlohe, C. L. Peichel, G.-P. Saetre, C. Bank, A. Brännström, A. Brelsford, C. S. Clarkson, F. Eroukhmanoff, J. L. Feder, M. C. Fischer, A. D. Foote, P. Franchini, C. D. Jiggins, F. C. Jones, A. K. Lindholm, K. Lucek, M. E. Maan, D. a Marques, S. H. Martin, B. Matthews, J. I. Meier, M. Möst, M. W. Nachman, E. Nonaka, D. J. Rennison, J. Schwarzer, E. T. Watson, A. M. Westram, and A. Widmer. 2014. Genomics and the origin of species. Nat. Rev. Genet. 15:176–92.

Servedio, M. R. 2015. Geography, assortative mating, and the effects of sexual selection on speciation with gene flow. Evol. Appl. 9: 90-102

Shaw, K. L., Y. M. Parsons, and S. C. Lesnick. 2007. QTL analysis of a rapidly evolving speciation phenotype in the Hawaiian cricket Laupala. Mol. Ecol. 16:2879–2892.

Slatkin, M., and L. Voelm. 1991. FST in a hierarchical island model. Genetics 127:627–629.

Smadja, C. M., and R. K. Butlin. 2011. A framework for comparing processes of speciation in the presence of gene flow. Mol. Ecol. 20:5123–5140.

Smith, J. M., and J. Haigh. 1974. The hitch-hiking effect of a favourable gene. Genet. Res. 23:23–35.

Sousa, V., and J. Hey. 2013. Understanding the origin of species with genome-scale data : modelling gene flow. Nat. Rev. Genet. 14:404–414.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.

Thomas, M. L., and L. W. Simmons. 2009. Sexual selection on cuticular hydrocarbons in the Australian field cricket, Teleogryllus oceanicus. BMC Evol. Biol. 9:162–174

Turner, T. L., M. W. Hahn, and S. V. Nuzhdin. 2005. Genomic islands of speciation in Anopheles gambiae. PLoS Biol. 3:1572–1578.

Van der Auwera, G. A., M. O. Carneiro, C. Hartl, R. Poplin, G. del Angel, A. Levy-Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K. V Garimella, D. Altshuler, S. Gabriel, and M. A. DePristo. 2013. From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. Curr. Protoc. Bioinformatics. 43:1–11.

van Doorn, G. S., U. Dieckmann, and F. J. Weissing. 2004. Sympatric Speciation by Sexual Selection : A Critical Reevaluation. Am. Nat. 163:709–725.

Van Doren, B. M., L. Campagna, B. Helm, J. C. Illera, I. J. Lovette, and M. Liedvogel. 2017. Correlated patterns of genetic diversity and differentiation across an avian family. Mol. Ecol. 26:3982–3997.

Via, S. 2012. Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow. Philos. Trans. R. Soc. B Biol. Sci. 367:451–460.

Vitti, J. J., S. R. Grossman, and P. C. Sabeti. 2013. Detecting Natural Selection in Genomic Data. Annu. Rev. Genet 47:97–120.

Walker, T. 1998. Trilling field crickets in a zone of overlap (Orthoptera: Gryllidae: Gryllus). Ann. Entomol. Soc. Am. 91:175–184.

Wegmann, D., C. Leuenberger, and L. Excoffier. 2009. Using ABCtoolbox.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution. 38:1358–1370.

Weissing, F. J., P. Edelaar, and G. S. van Doorn. 2011. Adaptive speciation theory: a conceptual review. Behav. Ecol. Sociobiol. 65:461–480.

Wu, C. I. 2001. The genic view of the process of speciation. J. Evol. Biol. 14:851–865.

Zheng, X., D. Levine, J. Shen, S. Gogarten, C. Laurie, and B. Weir. 2012. A High-performance Computing Toolset for Relatedness and Principal Component Analysis of SNP Data. Bioinformatics 28:3326–3328.

DATA ACCESSIBILITY

Data, including raw reads, sequences used for demographic analyses and SNP data files used in outlier analysis, will be made available on Dryad and the NCBI SRA archive prior to publication.

FIGURE LEGENDS

Fig. 1. Geographic distributions for *G. texensis* (red) and *G. rubens* (blue). The sympatric zone is marked with turquoise. The distributions are approximate and based on the Singing Insects of North America data base (http://entnemdept.ufl.edu/Walker/buzz/). The black dots in Texas and Florida represent the sampling locations for *G. texensis* and *G. rubens*, respectively.

Fig. 2. Genomic divergence. The distribution of the interspecific allele frequency difference, *D*, across SNPs (A), of the absolute divergence, d_{xy}, in 1000 bp windows (B), and of Tajima's D in 1000 bp windows for *G. rubens* (C) and *G. texensis* (D), respectively

Fig. 3. Demographic scenarios for Approximate Bayesian Computation. Eight scenarios were simulated under the ABC framework. (A) A simple divergence scenario (DIV) with a log uniform prior on the divergence time (T_{SPLIT}), the ancestral population size (N_{ANC}) and the current effective population sizes for *G. rubens* and *G. texensis* (N_{RUB}, N_{TEX}). (B) Three different gene flow models with either continuous gene flow (CGF), ancestral gene flow (AGF), or recent gene flow (secondary contact; RGF) were additionally defined by parameters describing migration rates ($M_{TEX>>RUB}$, $M_{RUB>>TEX}$; uniform priors not overlapping zero) and the time point since cessation of gene flow (T_{ISO}) or of secondary contact (T_{CONT}), both with log uniform priors. (C) Three bottleneck models defined by the time since recovery to current population sizes (T_{BOT}; log uniform prior) and the relative population size reduction (BOTSIZE; uniform prior not overlapping zero) for *G. rubens* (RB), *G. texensis* (TB), or both (BB). (D) An additional model (AGFRB) combining the best gene flow (AGF) and best bottleneck (RB) model, marked by the black, dashed rectangles. The posterior probabilities for model selection are given left of the square (opening) brackets for the three gene flow and the three bottleneck models, and right of the square (closing) brackets for the final model selection step.

Fig. 4. Distribution of observed and simulated data sets in multivariate summary statistic space. For each of the four models used in the final model selection step (see also Fig. 3) the distribution of the 1% posterior samples with the smallest Euclidean distance to the observed data is shown relative to the coordinates of the observed data. The multivariate summary statistic space is constrained by the first two linear discriminants representing linear combinations of the summary statistics used in model selection (see text for details).

Fig. 5. F_{ST} distributions of simulated and observed data. The distribution of Weir and Cockerham's F_{ST} as calculated by the program arlsumstat are shown for 2,000 simulated data sets for different demographic model: recent gene flow (secondary contact; RGF), continuous gene flow (CGF), ancestral gene flow (AGF), and the AGFRB model. Observed data (1,000 1kbp sequences) are represented by the red solid line. The histograms show the density (y-axis) to enhance comparison between simulated and observed data.

Fig. 6. Demographic parameter estimation. For the AGFRB (A-C) and the AGF (D-F) models, the density distribution of the ancestral and current population sizes (A, D), the time since divergence, cessation of gene flow, and recovery to current population sizes after the bottleneck (B, E), and the migration rates and bottleneck size (C, F) are shown. The density lines have been trimmed to the existent parameter distribution (i.e., no density extrapolation) and have been smoothed by adjusting the bandwidth. For lines within one panel the same smoothing bandwidth has been used.

SUPPLEMENTARY INFORMATION

Table S1. Individual RNA-seq read mapping statistics
Table S2. ABC estimates for the AGF scenario
Table S3. ABC estimates for the full sample (including 8 individuals from half-sib pairs), AGFRB scenario
Table S4. GO enrichment results for F_{ST}, d_{xy}, and selective sweep outliers.
Table S5. F_{ST} outlier loci.
Table S6. d_{xy} outlier loci
Table S7A,B Selective sweep outlier loci for *G. texensis* and *G. rubens*
Fig S1-S7. See figures for figure legends.

Table 1. ABC estimates. Prior distributions (log-scale), posterior predictive checks and posterior parameter estimates (log scale, median and 95% highest posterior density interval) for the model are shown.

	Prior[a]		Validation		Posterior		
Parameter	minimum	maximum	R^2	RMSEP	2.5%	Median	97.5%
$LOG_{10}(N_{ANC})$	4.0	6.0 (lu)	0.13	0.93	4.68	5.34	5.99
$LOG_{10}(N_{RUB})$	3.0	6.0 (lu)	0.90	0.32	4.03	4.50	4.70
$LOG_{10}(N_{TEX})$	3.0	6.0 (lu)	0.75	0.50	4.51	4.78	4.87
$LOG_{10}(T_{SPLIT})$[b]	5.0	7.0 (lu)	0.02	0.99	4.86	6.19	7.16
$LOG_{10}(T_{ISO})$[b]	3.0	7.0 (lu)	0.90	0.32	4.20	4.55	4.76
$LOG_{10}(T_{BOT})$[b]	5.0	7.0 (lu)	0.48	0.72	4.42	5.01	6.16
BOTSIZE	0.01	0.5 (u)	0.16	0.91	-0.04	0.15	0.48
$M_{TEX>>RUB}$	0.01	0.5 (u)	0.06	0.97	0.01	0.18	0.54
$M_{RUB>>TEX}$	0.01	0.5 (u)	0.06	0.97	0.01	0.27	0.74

[a] priors are uniformally (u) or log-uniformally (lu) distributed and do not overlap zero for migration rates and bottleneck size.

[b] the timing of demographic events is in (logarithm of) number of generation and both species have two generations annually.

Table S1. Individual RNA-seq read mapping statistics. Mapping rates were calculated using bowtie2 with default parameters.

Sample ID	**Species**	**Population**	**Sex**	**Mapping rate**
30037 rub	*G. rubens*	Ocala	f	84.52%
30038 rub	*G. rubens*	Ocala	f	85.33%
30039 rub	*G. rubens*	Ocala	f	85.66%
30040 rub	*G. rubens*	Ocala	f	84.35%
30041 rub	*G. rubens*	Ocala	f	84.85%
30057 rub	*G. rubens*	Lake City	f	88.40%
30059 rub	*G. rubens*	Lake City	f	88.86%
30060 rub	*G. rubens*	Lake City	f	87.83%
30061 rub	*G. rubens*	Lake City	f	90.23%
30052 rub	*G. rubens*	Ocala	m	78.01%
30053 rub	*G. rubens*	Ocala	m	80.72%
30055 rub	*G. rubens*	Ocala	m	79.76%
30063 rub	*G. rubens*	Lake City	m	77.70%
30064 rub	*G. rubens*	Lake City	m	77.56%
30065 rub	*G. rubens*	Lake City	m	70.75%
30027 tex	*G. texensis*	Lancaster	f	83.09%
30028 tex	*G. texensis*	Lancaster	f	83.20%
30029 tex	*G. texensis*	Lancaster	f	81.61%
30030 tex	*G. texensis*	Lancaster	f	83.80%
30031 tex	*G. texensis*	Lancaster	f	80.42%
30043 tex	*G. texensis*	Austin	f	91.78%
30044 tex	*G. texensis*	Austin	f	90.01%
30046 tex	*G. texensis*	Austin	f	87.70%
30032 tex	*G. texensis*	Lancaster	m	76.17%
30033 tex	*G. texensis*	Lancaster	m	77.76%
30034 tex	*G. texensis*	Lancaster	m	77.24%
30035 tex	*G. texensis*	Lancaster	m	80.79%
30036 tex	*G. texensis*	Lancaster	m	76.77%
30047 tex	*G. texensis*	Austin	m	86.40%
30049 tex	*G. texensis*	Austin	m	88.52%
30050 tex	*G. texensis*	Austin	m	79.15%
30051 tex	*G. texensis*	Austin	m	86.18%

Table S2. ABC estimates for the AGF scenario. Prior distributions (log-scale), posterior predictive checks and posterior parameter estimates (log scale, median and 95% highest posterior density interval) for the model are shown.

	Prior[a]		Validation		Posterior		
Parameter	minimum	maximum	R^2	RMSEP	2.5%	Median	97.5%
$LOG_{10}(N_{ANC})$	4.0	6.0 (lu)	0.0	0.96	4.76	5.31	5.81
$LOG_{10}(N_{RUB})$	3.0	6.0 (lu)	0.93	0.27	3.81	4.26	4.55
$LOG_{10}(N_{TEX})$	3.0	6.0 (lu)	0.93	0.27	3.98	4.45	4.69
$LOG_{10}(T_{SPLIT})$[b]	5.0	7.0 (lu)	0.06	0.97	4.61	5.83	7.04
$LOG_{10}(T_{ISO})$[b]	3.0	7.0 (lu)	0.79	0.46	4.21	4.56	4.73
$M_{TEX>>RUB}$	0.01	0.5 (u)	0.17	0.91	0.03	0.24	0.49
$M_{RUB>>TEX}$	0.01	0.5 (u)	0.12	0.94	0.01	0.26	0.51

[a] priors are uniformally (u) or log-uniformally (lu)distributed and do not overlap zero for migration rates and bottleneck size.

[b] the timing of demographic events is in (logarithm of) number of generation and both species have two generations annually.

Table S3 ABC estimates for the full sample (including 8 individuals from half-sib pairs), AGFRB scenario. Prior distributions (log-scale), posterior predictive checks and posterior parameter estimates (log scale, median and 95% highest posterior density interval) for the model are shown.

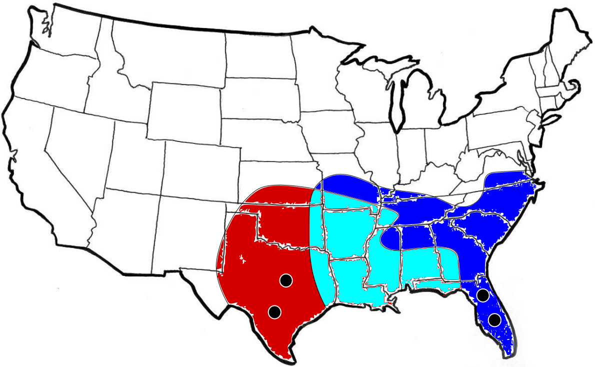
	Prior[a]		Validation		Posterior		
Parameter	minimum	maximum	R^2	RMSEP	2.5%	Median	97.5%
$LOG_{10}(N_{ANC})$	4.0	6.0 (lu)	0.05	0.974	4.94	5.32	5.72
$LOG_{10}(N_{RUB})$	3.0	6.0 (lu)	0.89	0.333	4.70	4.79	4.87
$LOG_{10}(N_{TEX})$	3.0	6.0 (lu)	0.88	0.346	4.73	4.85	4.94
$LOG_{10}(T_{SPLIT})$[b]	5.0	7.0 (lu)	0.01	0.997	5.49	6.23	6.74
$LOG_{10}(T_{ISO})$[b]	3.0	7.0 (lu)	0.81	0.438	4.27	4.53	4.72
$LOG_{10}(T_{BOT})$[b]	5.0	7.0 (lu)	0.02	0.990	5.14	5.19	5.32
BOTSIZE	0.01	0.5 (u)	0.01	0.995	0.09	0.15	0.23
$M_{TEX>>RUB}$	0.01	0.5 (u)	0.12	0.938	0.05	0.12	0.18
$M_{RUB>>TEX}$	0.01	0.5 (u)	0.12	0.938	0.01	0.18	0.75

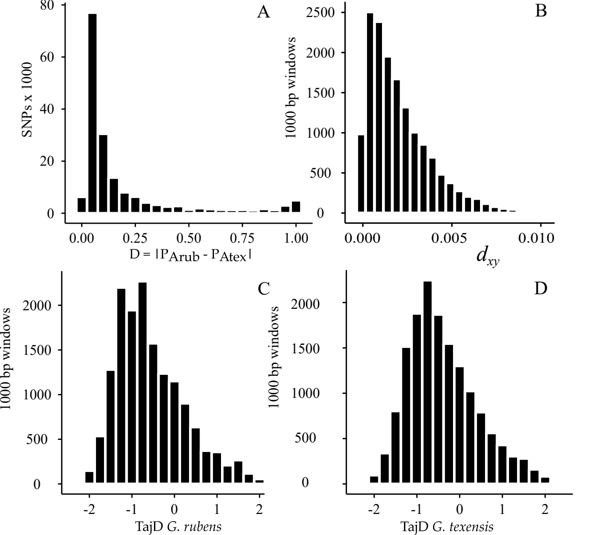
[a] priors are uniformally (u) or log-uniformally (lu) distributed and do not overlap zero for migration rates and bottleneck size.

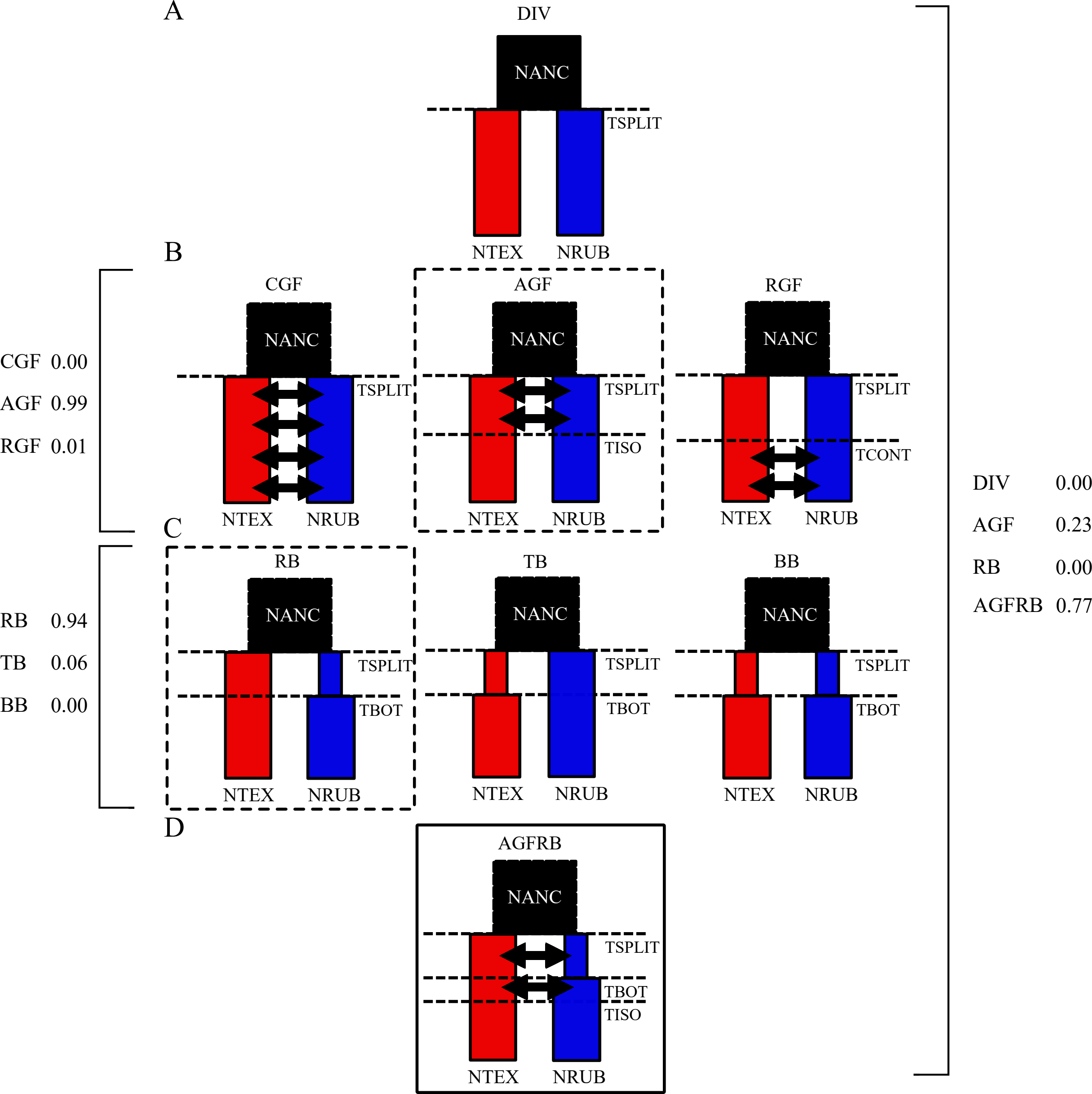
[b] the timing of demographic events is in (logarithm of) number of generation and both species have two generations annually.

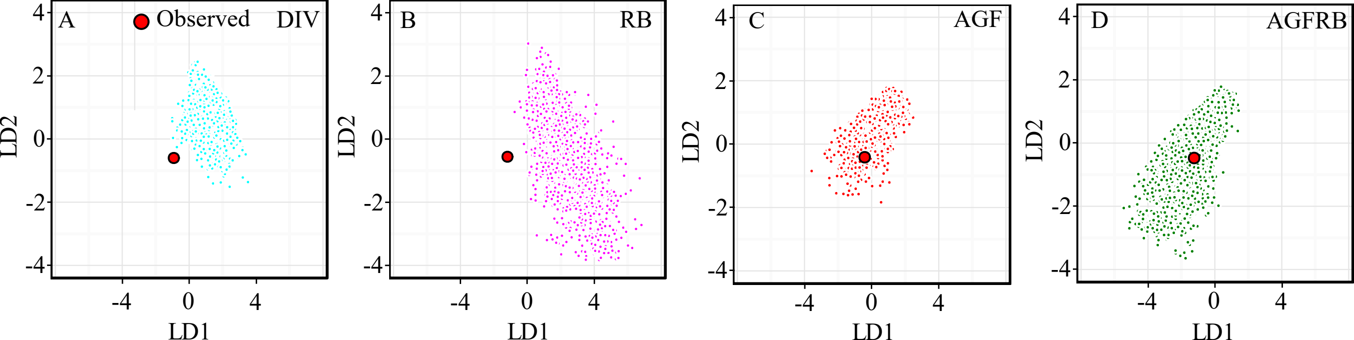
Table S4. GO enrichment results. The top ten terms of the Gene Ontology enrichment is shown for the d_{xy} outliers and the Allele Frequency Spectrum (AFS) outliers. For each Biological Process, the number of annotated transcripts and the number of observed and expected transcripts in the sample with a given annotation are shown. The Fisher's exact test P-value is corrected using the parent-child algorithm (Grossmann *et al.* 2007). The FDR is the false discovery rate based on the corrected P-values.

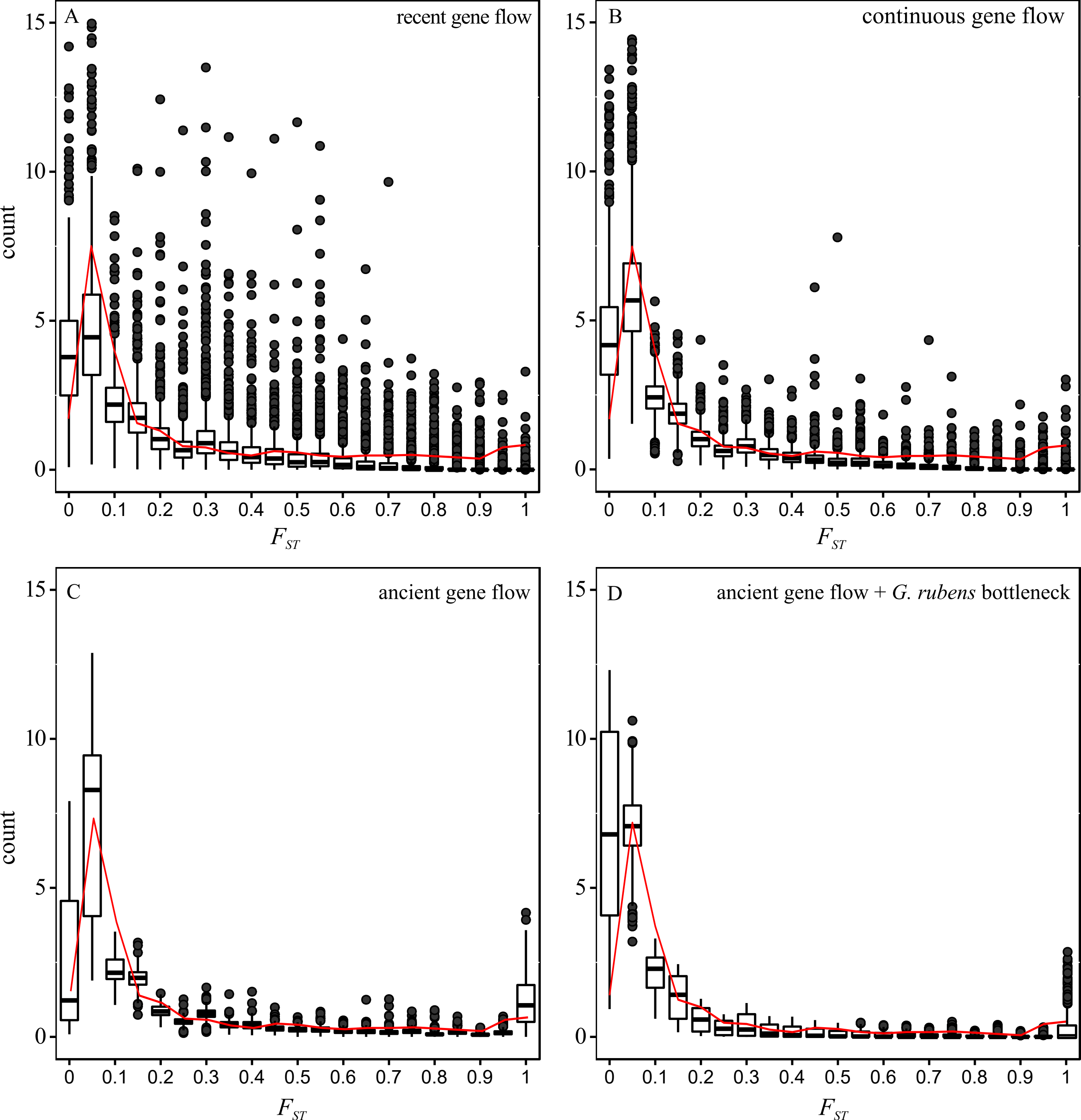
GO	Term	#Annot	#Sample	#Exp	P-value	FDR
		F_{ST}				
GO:0032543	mitochondrial translation	10	4	0.34	0.00056	1
GO:0043087	regulation of GTPase activity	54	10	1.85	0.00123	1
GO:0071695	anatomical structure maturation	2	2	0.07	0.00141	1
GO:0010822	positive regulation of mitochondrion organization	5	3	0.17	0.0026	1
GO:0044267	cellular protein metabolic process	1773	74	60.58	0.00318	1
GO:0022411	cellular component disassembly	56	7	1.91	0.00398	1
GO:0000910	cytokinesis	248	19	8.47	0.00428	1
GO:0043603	cellular amide metabolic process	577	30	19.72	0.00526	1
GO:0030716	oocyte fate determination	58	6	1.98	0.00564	1
GO:0050789	regulation of biological process	4028	160	137.63	0.00566	1
		d_{xy}				
GO:0042811	pheromone biosynthetic process	44	4	0.2	4.30E-06	0.0027
GO:0042810	pheromone metabolic process	49	4	0.22	3.60E-05	0.0071
GO:1903317	regulation of protein maturation	24	3	0.11	3.70E-05	0.0071
GO:0042446	hormone biosynthetic process	82	4	0.37	4.50E-05	0.0071
GO:1903318	negative regulation of protein maturation	23	3	0.1	0.0001	0.0152
GO:0044705	multi-organism reproductive behavior	359	6	1.62	0.0002	0.0232
GO:0019098	reproductive behavior	367	6	1.65	0.0005	0.0380
GO:0007618	mating	400	6	1.8	0.0005	0.0380
GO:0006551	leucine metabolic process	3	2	0.01	0.0011	0.0734
		AFS				
GO:0006996	organelle organization	2271	41	21.7	0.0003	0.3545
GO:1902589	single-organism organelle organization	1791	32	17.1	0.0004	0.3545
GO:0044238	primary metabolic process	4836	59	46.2	0.0007	0.4181
GO:0090066	regulation of anatomical structure size	375	12	3.6	0.0014	0.5867
GO:0050789	regulation of biological process	4028	52	38.5	0.0025	0.5867
GO:0030382	sperm mitochondrion organization	6	2	0.1	0.0027	0.5867
GO:0065007	biological regulation	4463	56	42.7	0.0027	0.5867
GO:0007294	germarium-derived oocyte fate determination	46	4	0.4	0.0028	0.5867
GO:0030716	oocyte fate determination	58	4	0.6	0.0033	0.5867
GO:0045924	regulation of female receptivity	7	2	0.1	0.0035	0.5867
GO:0006996	organelle organization	2271	41	21.7	0.0003	0.3545

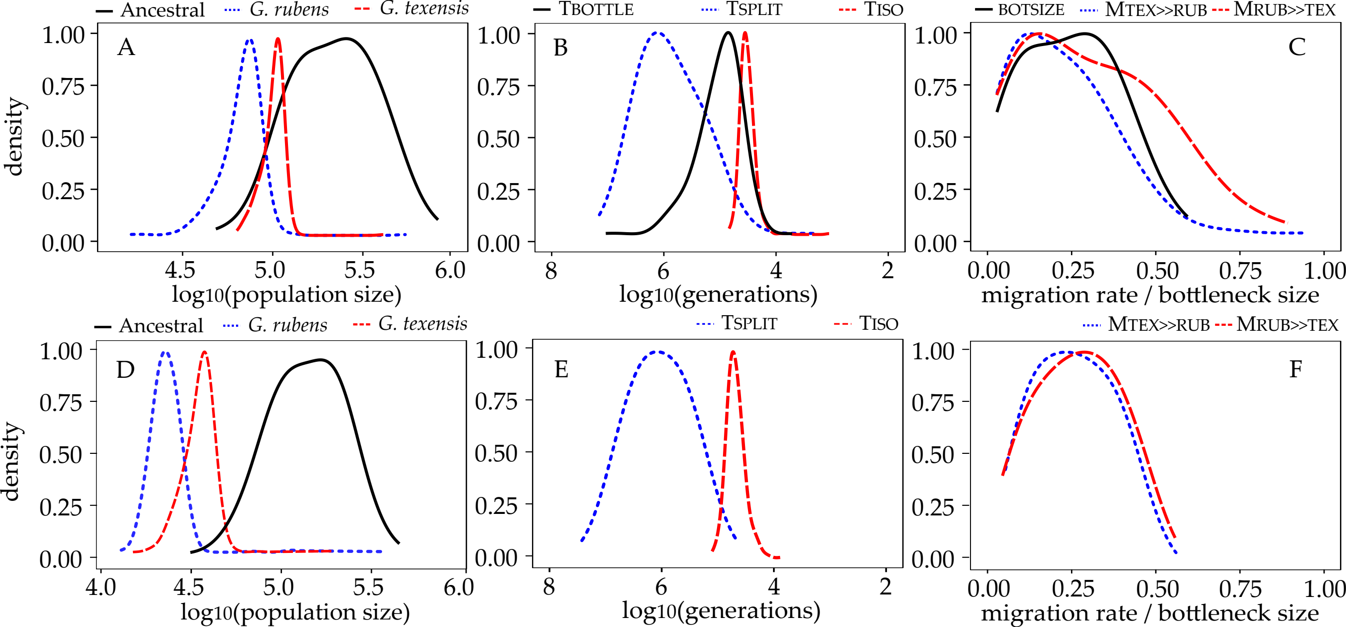



A
SNPs x 1000
0
20
40
60
80
0.00
0.25
0.50
0.75
1.00
D = |P_Arub - P_Atex|
B
1000 bp windows
0
500
1000
1500
2000
2500
0.000
0.005
0.010
d_{xy}
C
1000 bp windows
0
500
1000
1500
2000
-2
-1
0
1
2
TajD *G. rubens*
D
1000 bp windows
0
500
1000
1500
2000
-2
-1
0
1
2
TajD *G. texensis*


A
DIV
NANC
TSPLIT
NTEX
NRUB
B
CGF 0.00
AGF 0.99
RGF 0.01
CGF
NANC
TSPLIT
NTEX
NRUB
AGF
NANC
TSPLIT
TISO
NTEX
NRUB
RGF
NANC
TSPLIT
TCONT
NTEX
NRUB
DIV 0.00
AGF 0.23
RB 0.00
AGFRB 0.77
C
RB 0.94
TB 0.06
BB 0.00
RB
NANC
TSPLIT
TBOT
NTEX
NRUB
TB
NANC
TSPLIT
TBOT
NTEX
NRUB
BB
NANC
TSPLIT
TBOT
NTEX
NRUB
D
AGFRB
NANC
TSPLIT
TBOT
TISO
NTEX
NRUB


A
● Observed
DIV
B
RB
C
AGF
D
AGFRB
LD2
LD1
4
2
0
-2
-4
-4
0
4


A
recent gene flow
B
continuous gene flow
C
ancient gene flow
D
ancient gene flow + *G. rubens* bottleneck
count
count
0
5
10
15
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
F_{ST}



Ancestral
G. rubens
G. texensis
TBOTTLE
TSPLIT
TISO
BOTSIZE
MTEX>>RUB
MRUB>>TEX
A
B
C
D
E
F
density
0.00
0.25
0.50
0.75
1.00
4.0
4.5
5.0
5.5
6.0
log10(population size)
8
6
4
2
log10(generations)
0.00
0.25
0.50
0.75
1.00
migration rate / bottleneck size

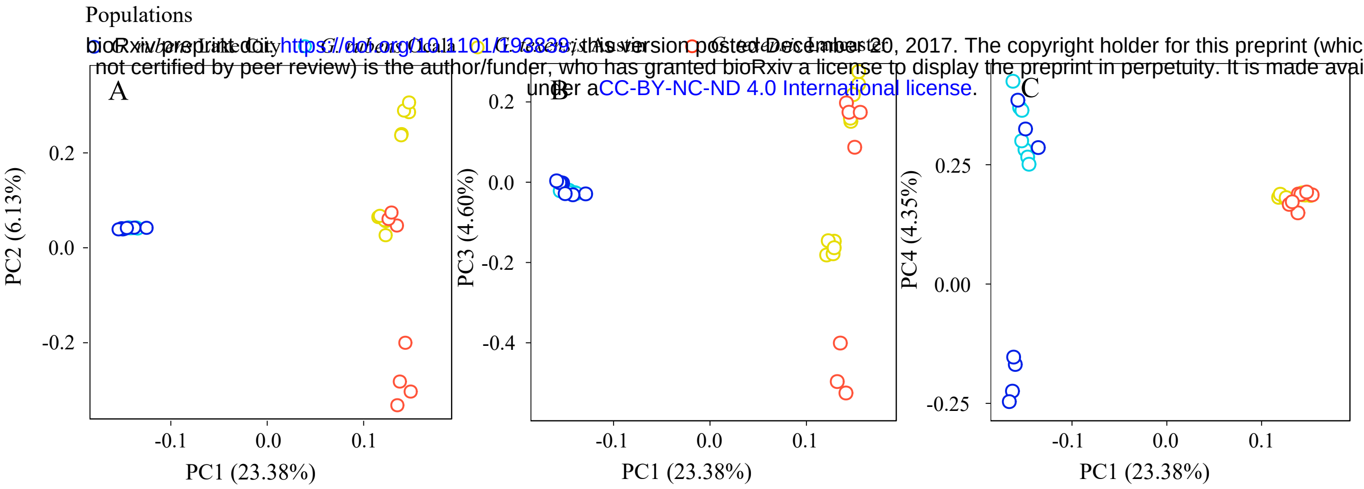


Fig S1. Population substructure in *G. rubens* and *G. texensis*. Variation in allele frequencies between species and between populations within species (Lake City and Ocala for *G. rubens*; Lancaster and Austin for *G. texensis*)is shown. The allele frequency variation in all 175,244 SNPs is summarized in the first four principal components teasing apart the species (PC1), and the populations in *G. texensis* (PC 2) and *G. rubens* (PC 4). Note that clustering along the PCs explaining within species variation among populations is much weaker compared to clustering of the species along PC1.

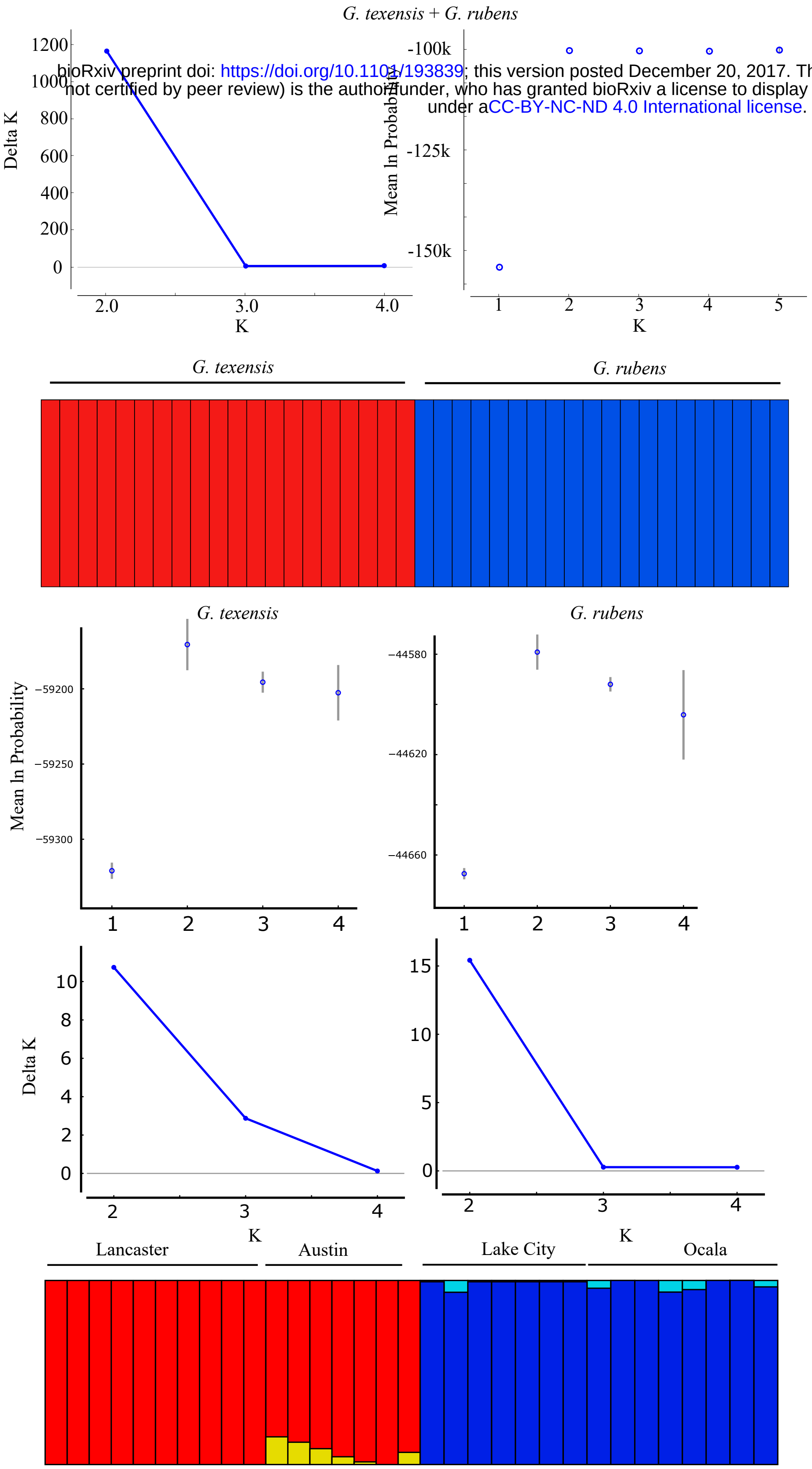


Fig S2.STRUCTURE results. For each of the species, STRUCTURE was run for 100,000 iterations at values for K=1 through K=4 (K=5 for the species combined). The mean natural logarithm of the probability and the delta K (increase or decrease in likelihood between consecutive runs for different values of K) were inspected to determine the most likely predicted number of populations. A run of *G. rubens* and *G. texensis* separately showed in both cases that, although the highest likelihood was for K=2 , differences with K=1 were only marginal and a defined pattern in population substructure was absent (see also the bar plots at the bottom). The run for the species combined (K=2) shows no introgression of *G. texensis* genes into the *G. rubens* or vice versa.

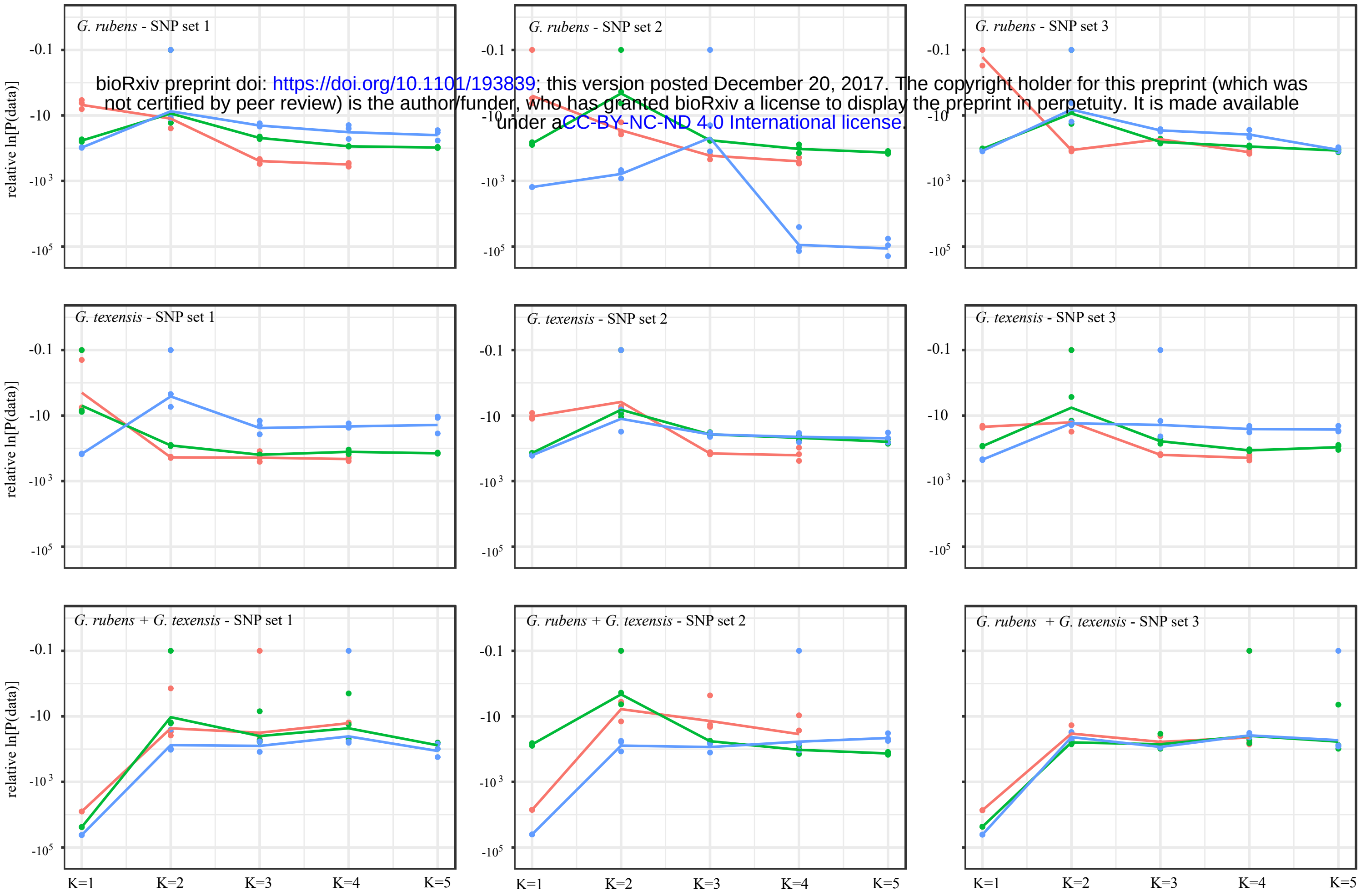


Figure S3. Relative natural log transformed probability of the data under different values for K. The raw probabillities from STRUCTURE relative to the maximum probability is shown for each K, for three random sets of 8835 SNPs (one per contig), and for *G. rubens*, *G. texensis*, and for the species combined (excluding eight individuals to correct for cryptic relatedness). Within each panel, the dots show each of the three iterations and the lines show the trend in the average difference in probability with the maximum probability for three different sample sizes: two random individuals per population (red), five random individuals per population (green), and all the individuals sampled from the populations.

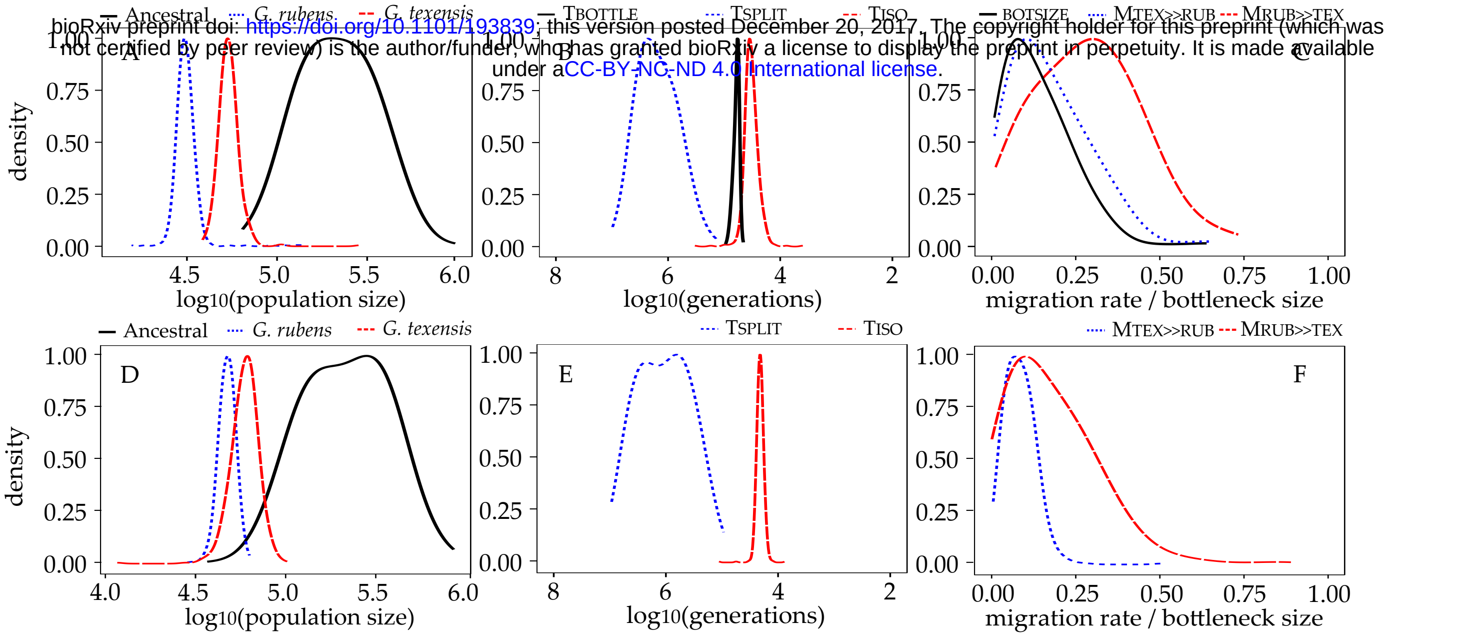


Fig S4. Demographic parameter estimation for the model with all 40 individuals. For the AGFRB (A-C) and the AGF models (D-F) the density distributions of the the ancestral and current population sizes (A,D), the time since divergence, cessation of gene flow, and recovery to current population sizes after the bottleneck (B,E), and the migration rates and bottleneck size (C,F) are shown. The density lines have been trimmed to the existent parameter distribution (i.e., no density extrapolation) and have been smoothed by adjusting the bandwidth. For lines within one panel the same smoothing bandwidth has been used.

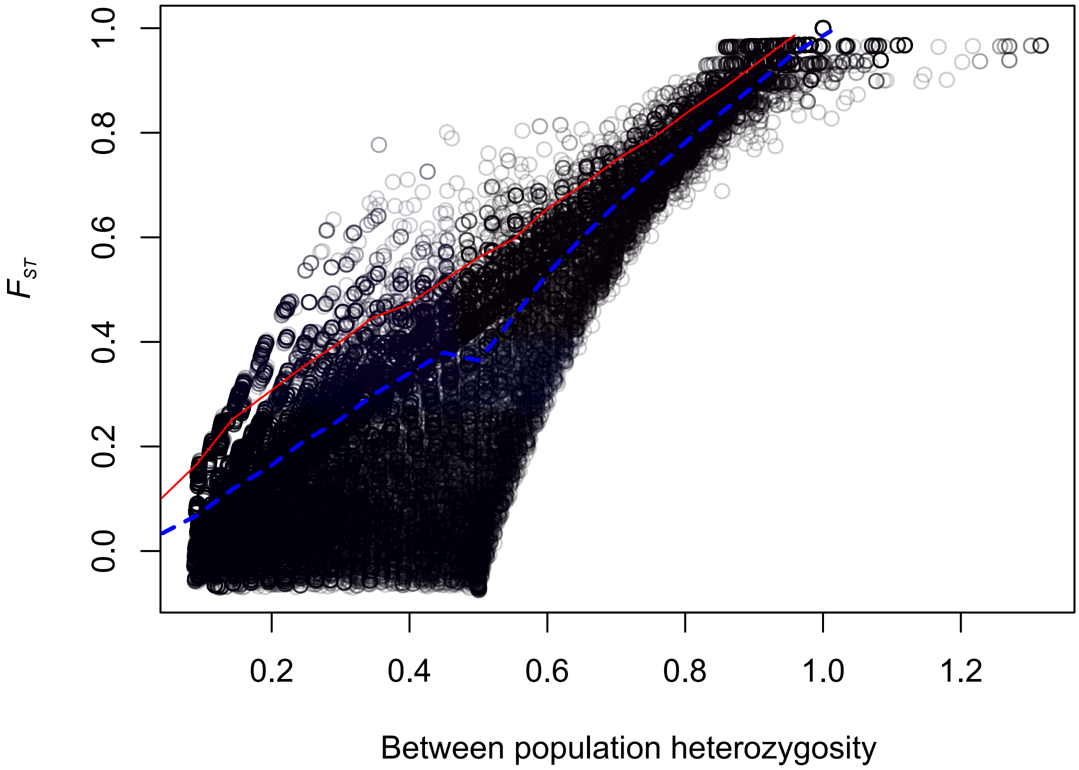


Fig. S5. Arlequin FST based selection scan. The circles represent estimates for the FST and between population heterozygosity for all SNPs with MAF > 0.05 (81,125 SNPs). The blue dashed and red solid line are the median and 99th quantile, respectively, of the simulated null distribution for this relationship under a hierarchical island model. Any SNPs above the red solid lines were considered outliers.