i. Consequences of coupled barriers to gene flow for the build-up of genomic differentiation

ii. Reproductive barriers and coupling in corn borers

iii. Henry D. Kunerth[1], Steve M. Bogdanowicz[1], Jeremy B. Searle[1], Richard G. Harrison[1], Brad S. Coates[2], Genevieve M. Kozak[3,4], & Erik B. Dopman[4]

iv. 1 Department of Ecology and Evolutionary Biology, Cornell University
2 02747USDA-ARS, Corn Insects & Crop Genetics Research Unit, Ames, IA, 50011
3 Department of Biology, UMass-Dartmouth, Dartmouth, MA,
4 Department of Biology, Tufts University, Medford, MA, 02155

v. Corresponding author: Henry D. Kunerth, A406B Corson Hall, 215 Tower Road, Ithaca NY, 14853
(507)-766-1339
henrykunerth@gmail.com

vi. Kunerth, Harrison, Kozak, and Dopman conceived of the project. Kunerth, Kozak, Coates and Bogdanowicz contributed to sample preparation. Kunerth and Bogdanowicz performed PCR and sequencing library preparation of targeted marker dataset. Kunerth and Dopman performed data analysis on the targeted marker dataset. Kozak and Coates performed pool-seq library preparation, and Kozak and Dopman analyzed the pooled sequencing dataset. Kunerth, Searle, Kozak, and Dopman wrote the first draft of the manuscript and Kunerth and Dopman revised the manuscript with contributions from other authors.

vii. We are grateful to the substantial contributions of many lab members and colleagues for their comments, criticisms, and discussions. Jessie Martin, Crista Wadsworth, Gabriel Golczer, Benjamin Hamilton, and Charles Mason all contributed to field collections, Christopher Watson assisted in the lab, and Qi Sun helped with bioinformatic analysis and scripting. H.D.K received Fellowship support from the Presidential Life Sciences and Cornell Fellowships. Partial research support was provided by USDA, ARS CRIS Project 5030-22000-019-00D. Any mention of products or services is for research purposes only, and does not constitute an endorsement for use by the USDA. USDA is an equal opportunity provider and Employer. E.B.D., R.G.H., and J.B.S. were funded by the National Science Foundation, and E.B.D. by Tufts University and cooperative agreement 58-5030-7-066 between USDA-ARS and Tufts University.

viii. All data will be archived in GenBank.

ABSTRACT

Theory predicts that when different barriers to gene flow become coincident, their joint effects enhance reproductive isolation and genomic divergence beyond their individual effects, but empirical tests of this 'coupling' hypothesis are rare. Here, we analyze patterns of gene exchange among populations of European corn borer moths that vary in the number of acting barriers, allowing for comparisons of genomic variation when barrier traits or loci are in coincident or independent states. We find that divergence is mainly restricted to barrier loci when populations differ by a single barrier, whereas the coincidence of temporal and behavioral barriers is associated with divergence of two chromosomes harboring barrier loci. Furthermore, differentiation at temporal barrier loci increases in the presence of behavioral divergence, while differentiation at behavioral barrier loci increases in the presence of temporal divergence. Our results demonstrate how the joint action of coincident barrier effects leads to levels of genomic differentiation that far exceed those of single barriers acting alone, consistent with theory arguing that coupling allows indirect selection to combine with direct selection and thereby lead to a stronger overall barrier to gene flow. Thus, the state of barriers – independent or coupled – strongly influences the accumulation of genomic differentiation.

INTRODUCTION

Understanding the origin of new species is crucial for explaining patterns of biodiversity. New species are thought to arise as populations diverge and accumulate trait differences that limit interbreeding, genetic exchange, and sharing of evolutionary processes (Mayr 1942, Coyne and Orr 2004, Harrison and Larson 2014). Consequently, there is long-standing interest in understanding how barriers to reproduction evolve and promote independent evolution of populations in nature, as characterized by the emergence of distinct genetic groups.

Most sister species are separated by multiple barriers to gene flow such as differences in mating time, differences in mate choice, and decreased hybrid survival (Coyne and Orr 2004; Dopman et al. 2010), arguing that speciation depends on the accumulation of barriers and strong reproductive isolation. Different barriers need not, however, initially evolve between the same populations. Instead, independent barriers could restrict gene flow between different populations, potentially as a gradual build-up of barriers with geographic distance (Barton 2013). Without the coincident action of multiple barriers between the same populations, theory indicates that differentiation will be centered at loci underlying barrier traits, reflecting weak overall reproductive isolation and potentially migration-selection balance (Barton 1983; Barton and Bengtsson 1986; Barton and de Cara 2009; Butlin and Smadja 2018; Flaxman et al. 2014; Schilling et al. 2018). At other genomic regions, gene flow and recombination can allow for shared evolutionary processes within what is essentially a single population (Barton 2013).

When multiple barrier traits become coincident in a process known as 'coupling,' their joint effects are believed to enhance reproductive isolation and genomic divergence beyond their individual effects, potentially leading to strong reproductive isolation sufficient for the origin of new species (Barton 1983; Barton and Bengtsson 1986; reviewed in Butlin and Smadja 2018). Recent simulations of divergence with gene flow suggests the presence of tipping-point dynamics, wherein a certain threshold of coupling, achieved by the accumulation of new barrier loci or increased linkage disequilibrium (LD) among existing loci, initiates a feedback loop of rapid differentiation across the genome (Flaxman et al. 2014, Schilling et al. 2018, reviewed in Nosil et al. 2017). This coupling threshold may correspond to a critical point at which the strength of selection across accumulated barrier loci, arising from both direct fitness effects of each locus and indirect fitness effects of their non-random associations, outweighs the homogenizing impact of recombination (measured as φ, the 'coupling coefficient', Barton 1983). As coupling builds towards this tipping point, recombination is predicted to cause a lag in divergence at neutral loci resulting in heterogeneous genomic differentiation (Nosil et al. 2017; Schilling et al. 2018). Genomic architectures that reduce recombination, including pleiotropy, tight physical linkage, or chromosomal rearrangements, are therefore expected to promote coupling (Servedio et al. 2011; Dagilis and Kirkpatrick 2016, Ortiz-Barrientos et al. 2016; Kirkpatrick & Barton 2006). Although coupling does not require physical linkage, in a widespread pattern known as the

'large X-effect,' sex chromosomes often harbor more barrier loci than autosomes (Coyne and Orr, 1989; Presgraves 2008; 2018). Consequently, the X (or Z) chromosome might be prone to coupling dynamics and rapid chromosomal differentiation (Lasne et al. 2017; Presgraves 2018). Consistent with this, 95% of 129 studies reviewed in Presgraves (2018) find evidence between taxa of elevated differentiation on sex chromosomes relative to autosomes, although there is no shortage of alternative hypotheses for biased sex chromosome differentiation.

Critical to theoretical predictions of a build-up of genomic divergence that ultimately characterizes new species is the coupling of multiple barriers to gene flow. Understanding how coupling counteracts gene flow to allow genomic divergence requires us to reconstruct the sequence of events, from barrier independence to barrier coincidence, as speciation progresses. Much empirical work has indirectly studied the evolution of coupling during speciation by quantifying genomic differentiation across lineages at differing stages of divergence (Gagnaire et al. 2013, Seehausen et al. 2014; Shaw & Mullen 2014; Riesch et al. 2017, Xu & Shaw 2019, 2021). Equally valuable are systems allowing direct comparisons of genomic variation when barrier traits or loci are found in either a coincident or independent state. Such direct studies of coupling dynamics may be especially challenging in many systems because barriers to gene flow, trait loci, and recombination landscapes are either unknown or barriers are found exclusively in a coincident state. Consequently, strong empirical tests of coupling dynamics are rare.

The European corn borer moth (ECB) *Ostrinia nubilalis* (Hübner) (Lepidoptera: Crambidae) is useful for studying coupling dynamics because hybridizing population pairs can differ by independent or coincident barriers. Although recently described as lacking coincident effects (Butlin and Smadja 2017), in fact ECB populations differ in combinations of sex pheromone, host-plant association, and phenology (Malausa et al., 2007; Dopman et al. 2010, Coates et al. 2018). Throughout much of its introduced range in North America and its ancestral range in Europe and Asia, two pheromone strains co-exist (Klun and Cooperators, 1975; Anglade and Stockel 1984). In the E strain, females produce and males respond to a sex pheromone comprised mainly of the E isomer of 11-tetradecenyl acetate (E pheromone), whereas Z moths use a blend of mainly Z isomer (Z pheromone) (Klun 1973, Kochansky 1975). The Z strain is more widespread, including areas west of the Appalachian Mountains where it occurs alone (Klun and Cooperators, 1975; Sorenson et al. 1992). Although dietary generalists (Hodgson, 1928; Lewis, 1975), larvae show some host-plant specialization on cultivated maize (Calcagno, et al. 2007; Fisher et al 2017), their predominant host (O'Rourke et al. 2010; Coates et al. 2019). However, in France, the E strain shows a near complete affiliation with non-maize hosts (Malausa et al., 2007), resulting in its identification as the sister species *O. scapulalis* (Calcagno et al. 2007; Frolov et al. 2007). Finally, three locally adaptive ecotypes have been described in North America that differ in seasonal phenology and voltinism. A northern ecotype (> 42˚N) produces one mid-season generation per year (univoltine), a central ecotype (~36–45˚N) makes one early-season and one late-season generation (bivoltine), and a southern ecotype (< 40˚N) makes three or more generations per year (multivoltine) (Showers et al. 1976; 1981). The northern ecotype uses the Z pheromone, whereas both Z and E pheromones are used by central and southern ecotypes (Sorenson et al. 1992; Showers 1993). Northern and central ecotypes become sympatric across an area stretching from Wyoming to Maine (Showers et al. 1975; 1981), resulting in population pairs differing by voltinism alone, or both voltinism and pheromone.

The three main axes of phenotypic divergence (phenology, pheromone signaling, host-plant use) result in various barrier effects. Populations of North American ECB moths can experience a temporal barrier from differences in seasonal phenology (bivoltine versus univoltine), a behavioral barrier from differences in pheromone communication (Z versus E), or a coincidence of the two barrier types (bivoltine E versus univoltine Z) (Figure 1). Ten other potential barriers to gene flow have been studied that are comparatively weaker (Dopman et al. 2010). North American populations are derived from Hungary and Italy (Smith 1920, Caffrey and Worthly 1927) rather than France, where barrier effects between the ECB moth *O. nubilalis* and its sister species *O. scapulalis* consist of a coincidence of pheromone and host, but not voltinism (Thomas et al. 2003).

Four other aspects of the system are noteworthy because they have been shown by simulations to efficiently reduce gene flow or promote coupling (Flaxman et al. 2014, Nosil et al. 2017, Schilling et al. 2018). First, evidence suggests pheromone signal and preference, controlled by *pgFAR* (pheromone production) and *bab* (preference) genes (Unbehend et al. 2021, Lassance et al. 2010), are 'multiple-effect traits' (Smadja and Butlin 2011) that contribute to both pre- and post-zygotic isolation (mate choice and behavioral sterility) (Glover et al. 1991; Dopman et al. 2010; Unbehend et al. 2021). Second, in an example of 'magic trait' evolution (Servedio et al. 2011), temporal isolation arising from non-random meeting of univoltine and bivoltine adults is an automatic, pleiotropic outcome of divergent selection on diapause phenology, controlled by two interacting clock genes *per* and *Pdfr* (Dopman et al. 2005; Levy et al. 2015; Kozak et al. 2019). Third, three of four identified barrier loci (*per*, *Pdfr*, *bab*) map to the Z chromosome, creating a 'large Z-effect' (Dopman et al. 2004; Kozak et al. 2019; Unbehend et al. 2021). Finally, a chromosomal rearrangement (putatively an inversion(s)) suppressing recombination across ~40% of the Z chromosome has captured at least one barrier locus (*per*) and may be abundant within some E pheromone populations (Wadsworth et al. 2015; Kozak et al. 2017).

Prior population genetic results suggest identified barriers influence the build-up of genomic divergence. Geographic distance is weakly associated with genetic variation throughout the range (Bourguet et al. 2000; Kim et al. 2011; Coates et al. 2019), consistent with numerous examples of long-distance dispersal (30–60 km) (reviewed in Sappington 2018). Instead, genetic structure appears related to barrier effects. In France, differences in pheromone and host use between sympatric *O. nubilalis* and *O. scapulalis* are associated with significant genetic differentiation (Bethenod et al., 2005; Malausa et al., 2007). In an analysis restricted to 65 mostly autosomal SNPs among 12 localities in North America, pheromone signaling but not host or geography significantly contribute to genetic differentiation, but the influence of phenology was unmeasured (Coates et al. 2019) Kozak et al. (2017) used genome-wide data to study the Z chromosome rearrangement in North America and did find elevated differentiation when populations differed in both pheromone and phenology compared to pheromone alone, but contrasts were between single population pairs. Overall, these prior studies do imply that barriers to gene flow heavily shape patterns of gene flow. However, incomplete genomic data or barrier effect combinations have made it difficult the evaluate the importance of coupling dynamics in this system.

Here, we help resolve some earlier limitations by analyzing pooled genome data from 14 population samples and compare genetic divergence between four types of population pairs characterized by 1) no barrier effects, 2) only phenology differences (voltinism), 3) only pheromone differences, or 4) differences in both phenology and pheromone (joint divergence). We further analyze targeted amplicon sequencing of the 'large-effect' Z chromosome and the autosomal pheromone production locus *pgFAR* from six locations varying in barrier number. Critically, these data allow for comparisons of sympatric populations differing by independent or coincident barriers, allowing us to ask whether we can detect empirical evidence that coupling promotes speciation as suggested by theory. Accordingly, compared to barrier traits when acting alone, upon coincidence we predict evidence of stronger selection or more efficient reduction of gene flow. Two expected signatures are increased divergence around individual barrier loci as well as enhanced coupling (linkage disequilibrium, LD) among barrier loci. If overall selection after coupling outweighs recombination, theory predicts tipping point dynamics and a transition towards differentiation of neutral loci. Differentiation can be genome-wide or more localized, but the combination of a 'large Z-effect' and a Z-linked rearrangement might suffice in allowing for the build-up of sex chromosome divergence. An alternative possibility is that coincident barriers do not lead to a detectable strengthening of the overall barrier. This could mean that individual barriers are so weak that they evolve independently even when coincident, or possibly that coupling effects are difficult to detect in nature, and theory could explore further quantitative predictions for empirical systems.

MATERIALS AND METHODS

We sampled 13 localities in Eastern United States between 2000 and 2014 (Figure 1). Pheromone usage was based on male capture in traps baited with either synthetic E ("New York") or Z ("Iowa") pheromone lures (*Heliothis* traps, Scentry Biologicals, Billings, MO). Traps were separated by ≥ 12 next to sweetcorn fields.

Trap data were verified by gas chromatography of female pheromone gland extracts for the presence of E and Z isomers. For 7 sites, we sampled diapausing larvae from corn stalks and raised individuals under diapause breaking conditions (26˚C, 16h light) until eclosion, allowing classification of voltinism based on post-diapause development (PDD) time (bivoltine: PDD time < 18 days; univoltine: PDD time > 38 days). For one population (Dover), we collected direct developing larvae in July. These individuals were mated as adults, larvae were exposed to diapausing inducing conditions and then phenotyped for PDD. For 5 sites, we determined voltinism by the seasonal timing of male capture in pheromone traps (bivoltine = early and late flight; univoltine = single mid-summer flight) (Glover 1991, Calvin 1994, Zaman 2008, Dopman 2010).

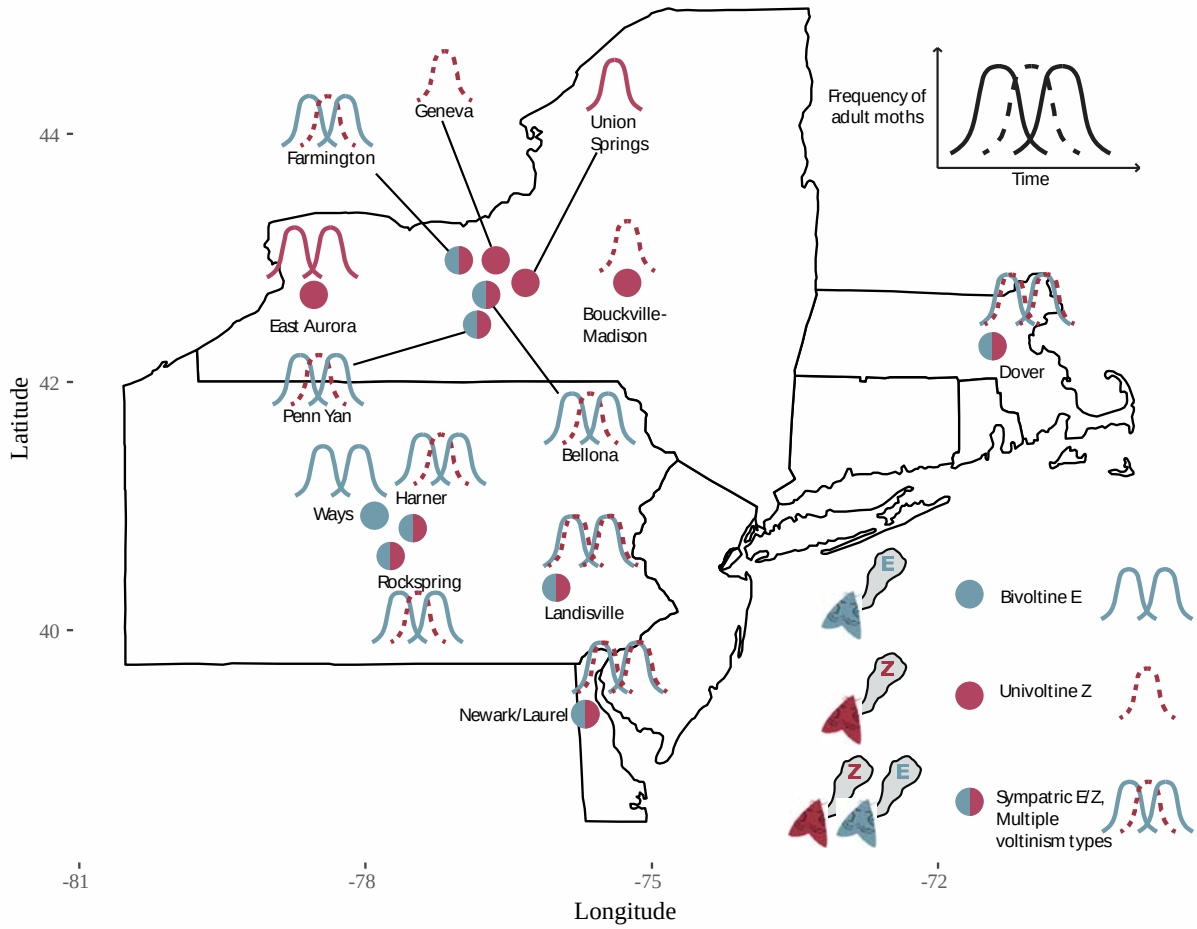

Figure 1. Map of collection sites and phenotypic information. Graphs indicate the frequency of adult moths across the growing season and number of distinct generations (voltinism). Graph and circle color indicates whether there were E (blue), Z (red) or both pheromone strains present at each collection site.

Pooled genome sequencing

To evaluate how the transition between independent and joint axes of phenotypic divergence might affect genomic differentiation, we performed pooled genome sequencing on 14 population samples (425 individuals) separated by pheromone and phenology (5 bivoltine E (BE) populations, 3 bivoltine Z (BZ) populations, and 6 univoltine Z (UZ) populations; see Table 1). Pheromone strain assignment used *pgFAR* genotype (E or Z homozygotes) from individuals as determined by a diagnostic restriction enzyme digest of *pgFAR* PCR products (Coates et al. 2013). DNA was isolated using Qiagen DNeasy tissue kits (Qiagen, Germantown, MD) without vortexing to preserve high molecular weights, and samples were treated with RNase A (Qiagen). DNA from samples from Bellona, NY were extracted using Qiagen genomic tips (20 G). After quantifying DNA concentration using a Qubit (Thermo-Fisher), pools combined equal amounts of DNA per individual (Table 2).

Pooled libraries (2 libraries per lane; 2-5 µg/uL per pool) were sequenced on an Illumina HiSeq3000 (2x150 bp) (except for Rockspring and Ways, which had 1x100 bp data) at Iowa State University DNA Facility, Ames, IA.

Trimmomatic v.35 removed Illumina adapters (TruSeq2 single-end or TruSeq3 paired-end), and reads with quality score (q) < 15 over a sliding window of 4 and reads < 36 bp long were removed. Repetitive regions of the ECB reference genome (GenBank BioProject: PRJNA534504; Accession SWFO00000000; Kozak et al. 2019) were masked by RepeatMasker (using *Drosophila melanogaster* TE library in repbase; http://www.repeatmasker.org/; accessed March 2017), to which trimmed genomic data were aligned using bowtie2 (Langmead and Salzberg 2012). Aligned reads were sorted and filtered using Picard and samtools to remove duplicates and reads with a mapping quality score (Q) < 20 (Li et al. 2009). We calculated coverage from the aligned, filtered, sorted bam files using bedtools and the genomeCoverageBed function (Table 2) (Quinlan 2014). As described in Kozak et al. (2019), we determined chromosomal positions of scaffolds by aligning genotyping-by-sequencing data from pedigree families (with and without the recombination suppressor segregating) to the genome scaffolds, and constructing linkage groups from identified SNPs in rQTL (Coates and Siegfried 2015, Kozak et al. 2017). We considered SNPs from Z chromosome-linked scaffolds that showed no recombination in E x Z strain pedigrees but recombined in Z x Z strain pedigrees to be inside the region of suppression (Wadsworth et al. 2015; Kozak et al. 2017). Twenty-nine scaffolds comprising 8.96 Mb of sequence were determined to be inside the region of recombination suppression while 38 scaffolds (11.94 Mb) were outside the region. We placed a small number (3%) of the Z chromosome-linked scaffolds using positions relative to other markers if they were identified as best hits from BLASTN searches against the genome scaffolds of markers or BACs from previously published Z-linkage maps and pedigree families in *O. nubilalis* (Kroemer et al. 2009; Levy et al. 2015; Koutroumpa et al. 2016; Yasukochi et al. 2016) and also in BLASTN searches of Z chromosome-linked genome scaffolds in the sister species *O. scapulalis* against the ECB genome (Brousseau et al. 2018). Order and orientation of scaffolds was determined using linkage information from previous maps and gene order from *Bombyx mori* and approximate locations across the Z chromosomes was calculated in Mb (described in Kozak et al. 2019). Scaffolds with no evidence for Z chromosome linkage in any pedigree were designated as autosomal. Chromosome assignment of scaffolds to various autosomes (chr. 2–31) used linkage groups in the pedigree families, and synteny of scaffolds to *Bombyx mori* and *Melita cinxia* chromosomes (described in Kozak et al. 2019). Chromosome 12 contains the *pgFAR* gene (Dopman et al. 2004).

We used SAMtools to identify single nucleotide polymorphisms (SNPs) in populations (Li et al. 2009). Scripts from Popoolation2 were used to filter (removing SNPs near small indels, minor alleles that did not appear twice in each population), calculate allele frequency, and F_{ST} (Kofler et al. 2011). Pairwise differences in F_{ST} were calculated on 100 kb non-overlapping sliding windows (including only windows with a minimum coverage > 14 reads and a maximum coverage < 300 and at least 20 polymorphisms within the window, as recommended by Fabian et al. 2012). We conducted pairwise F_{ST} comparisons among four categories of populations: no divergence in pheromone or phenology (BZ vs. BZ, UZ vs. UZ), differing only in voltinism (BZ vs. UZ), differing only in pheromone (BE vs. BZ), and differing in both pheromone and voltinism (BE vs. UZ).

We separately compared F_{ST} for autosomal and Z chromosomal loci, and Z loci located inside the region of suppressed recombination ("rearranged") or Z loci located outside this region ("collinear"). We removed and separately analyzed genomic scaffolds containing loci known to contribute to voltinism [*per* (voltinism; scaffold 532; Z chromosome inside "rearranged" region; Kozak et al. 2019), and *Pdfr* (voltinism; scaffold87, Z chromosome "collinear" region; Kozak et al. 2019)] and pheromone trait differences [*pgFAR* (female pheromone production: scaffold178, chr 12; Lassance et al. 2009); *bab* (male flight response; scaffold662, scaffold50; Z chromosome "collinear" region; Unbehend et al. 2021)]. We calculated and plotted F_{ST} separately for 100 kb windows overlapping seasonal timing genes (*per*, *Pdfr*, N = 4 windows each scaffold), windows overlapping the *bab* locus (N = 4 windows), and windows overlapping the *pgFAR* locus (N = 5 windows). We calculated the average F_{ST} for all population pairs of a given type within each 100 kb window and then took the mean of all windows for each genomic region. We performed comparison among windows

using a Kruskal-Wallis test and pairwise posthoc Dunn's tests in the PMCMRplus package in R (Pohlert 2014).

Targeted amplicon sequencing

To specifically test how the transition from independent to coincident barrier effects in sympatry influences the build-up of genomic differentiation, we performed amplicon sequencing of the 'large-effect' Z chromosome and *pgFAR* on population samples at six locations (260 individuals, 6 sites) experiencing either no phenotypic divergence or barrier effects (1 allopatric BZ population, 1 allopatric UZ population), pheromone divergence and behavioral isolation (2 sympatric BE/BZ population pairs), or joint phenotypic divergence and cumulative temporal and behavioral isolation (2 sympatric BE/UZ population pairs) (Table 2). Females were classified by gas chromatography of pheromone gland extracts, and both sexes by diagnostic SNPs at the *pgFAR* gene. Phenology classification was described above, using adult eclosion time as a proxy for PDD (except at Dover as described above). At Delaware, only female samples were analyzed, but all other locations had both male and female samples in approximately equal proportion.

We used 67 primers for multiplex PCR of 450–470 bp product sizes (Table S1). Primers were distributed across the Z chromosome, including near one of the two temporal barrier genes (*Pdfr*), and one near *pgFAR* on chromosome 12 (Figure S1). Only loci successfully amplified from both pheromone strains were carried forward to the final multiplex. We carried out two multiplexed reactions using a Qiagen multiplex PCR kit in two 384-well plates. Following amplification, we pooled products and transferred to one 384-well plate for barcoding. We created a single library by pooling these products after dual barcoding using Nextera N5/N7 primers and OneTaq Polymerase (New England Biolabs). We then cleaned the library using with Ampure XP (Beckman Coulter) for 350–600 bp product sizes. We quantified libraries using a Qubit and sequenced them on a single Illumina MiSeq (2x250 bp) lane at the Cornell Institute of Biotechnology Genomics Facility, Ithaca, NY.

We demultiplexed and trimmed raw reads using a quality threshold of 20 with a custom Perl script. For each locus, we aligned reads using ClustalW (Larkin et al. 2007), generated a haplotype table for each sample at all loci, and called genotypes from the two most common haplotypes (minimum 4 reads total). If the ratio of the second most common genotype to the most common genotype was below 0.2, a homozygous genotype for the most common allele was called. We removed from analysis all individuals with $\geq$ 50% missing data, then loci with > 35% missing data. Following these cutoffs, we re-filtered all individuals and removed those with $\geq$ 35% missing data. We tested lower missingness cutoffs but found that a 35% missingness filter produced STRUCTURE and PCA results with no discernable differences to lower filters. Applying quality filters resulted in 217 moths (85 males, 132 females).

Using Z chromosome data, we inferred the presence of distinct populations, assigned individuals to populations, and identified admixed individuals using STRUCTURE (Pritchard et al. 2000). Since hemizygous females only carry one Z chromosome, we treated all samples as haploid with males contributing two haploid sets. To assure that this design did not bias results, we conducted separate STRUCTURE runs using only male samples (treated as diploid) or only female samples (treated as haploid). In all cases, we demonstrated qualitatively similar results and could not detect any bias. The presented results include all samples and all markers, unless otherwise specified. We tested K = 1–7 with a burn-in phase of 10,000 repetitions and a run phase of 100,000 repetitions and used default model assumptions with no weight given to location as a prior. Fisher's exact tests were used to compare differences in estimated admixture across groups. A principal components analysis (PCA) complemented the analysis of genetic structure using the ade4 package in R (Dray & Dufour 2007).

Finally, we measured linkage disequilibrium to quantify the strength of coupling at barrier loci (*pgFAR*, *Pdfr*) and the degree to which selection and assortative mating are counteracting neutral gene flow and recombination at each of the six sites. After filtering for a 0.15 minor allele frequency cutoff, we calculated r^2 values for every pair of biallelic SNPs in the female moth dataset using custom scripts in R. Following

calculations, we estimated mean r^2 value per locus and then combined results by barrier effect type (i.e., no barrier, one barrier, two barriers). We thinned this dataset to include only *pgFAR* and a single randomly selected locus per Z chromosome-linked scaffold. This resulted in a final dataset including 27 loci: 11 Z chromosome-linked markers outside the rearrangement (including *Pdfr*), 15 Z chromosome-linked markers within the rearrangement, and *pgFAR*. Mann-Whitney U tests were used to compare differences in LD across groups.

RESULTS

Pooled genome sequencing

As some population pairs were collected from the same site ("sympatric") and some from geographically distant sites ("allopatric"), we tested for an effect of geographic distance on the matrix of all pairwise F_{ST} differences using Mantel tests in the R package vegan. We used a matrix of geographic distance between sites (circle distance) as the explanatory matrix but consistent with previous studies (Bourguet et al. 2000; Kim et al. 2011; Coates et al. 2019), there were no significant effects of geography on genetic differentiation (for any category: "rearranged" Z, p = 0.47; "collinear" Z, p = 0.57; autosomal, p > 0.14; 1000 permutations). A conservative analysis was restricted to those population pairs within ~100 km (1-3 generations of dispersal, Sappington 2018) (6 population samples in NY, 7 population samples in PA). In all cases, we demonstrated qualitatively similar results and could not detect any bias (Figure S2). Thus, we included all pairwise comparisons regardless of site of collection to increase the number of comparisons that could be analyzed.

All predicted chromosomes not genetically linked to trait loci showed little genetic differentiation regardless of the number of phenotypic differences between population samples (mean $F_{ST} \leq 0.05$, $p > 0.1$, Figure 2). In contrast, differentiation on the Z chromosome and chromosome 12 increased as the number of barriers accumulated. Compared to population pairs lacking phenotypic divergence (mean F_{ST} = 0.04), differences in pheromone led to significant elevation (25%) of differentiation on chromosome 12 (mean F_{ST} = 0.05, p = 0.03), whereas differences in phenology (voltinism) did not (mean F_{ST} = 0.04, p = 0.86), as expected since phenology barrier loci do not occur on chromosome 12. The largest increase (50%) of chromosome 12 differentiation occurred upon joint divergence of pheromone and phenology (mean F_{ST} = 0.06, p = 6.7 x 10^{-5}), although the increase was not significantly greater than pheromone divergence alone (p = 0.17).

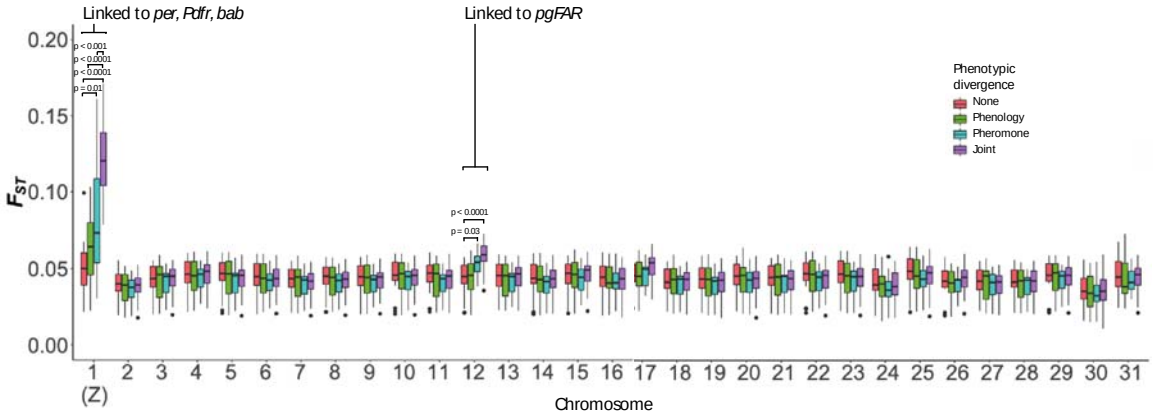

Figure 2. Genetic differentiation (mean F_{ST}) estimated in 100 kb windows plotted by predicted chromosome between population pairs lacking differences in barrier traits (Z/Z; red) or differing in phenology (BZ/UZ; green), pheromone signaling (BE/BZ; blue), or both phenology and pheromone (BE/UZ; purple). Scaffolds linked to barrier trait loci *pgFAR*, *per*, *Pdfr*, and *bab* were excluded from F_{ST} calculations on the Z chromosome and chromosome 12. Statistically significant differences via Dunn's test p < 0.05 are labeled with brackets.

Compared to population pairs lacking phenotypic divergence on the Z chromosome (mean F_{ST} = 0.05, Figure 2), phenology divergence led to a weak (20%) non-significant increase of differentiation (mean F_{ST} = 0.06, p = 0.28), whereas a significant (80%) increase was seen in the presence of pheromone divergence (mean F_{ST} = 0.09, p = 0.01). Joint divergence led to the greatest (140%) increase of Z chromosome differentiation (mean F_{ST} = 0.12, $p = 5.3 \times 10^{-9}$) and was statistically significant compared to phenology alone ($p = 2.7 \times 10^{-6}$) and pheromone alone ($p = 8.2 \times 10^{-3}$). This effect was primarily driven by elevated F_{ST} inside the region of reduced recombination among populations that differed in both pheromone and phenology (Figure 3D, 3E). F_{ST} of the rearranged region was elevated in joint divergence comparisons relative to the remaining three population categories (p < 0.0001, Figure 3D). F_{ST} of collinear regions was significantly elevated between populations with joint divergence compared to populations lacking trait differences and populations differing only in phenology (p < 0.0027, Figure 3E). However, F_{ST} was not quite significantly elevated compared to pairs that differed only in pheromone (p = 0.092). In populations that differed in just pheromone, F_{ST} inside the region of reduced recombination was elevated compared to populations lacking barrier effects (p = 0.007), but not populations differing in phenology (p = 0.20, Figure 3D).

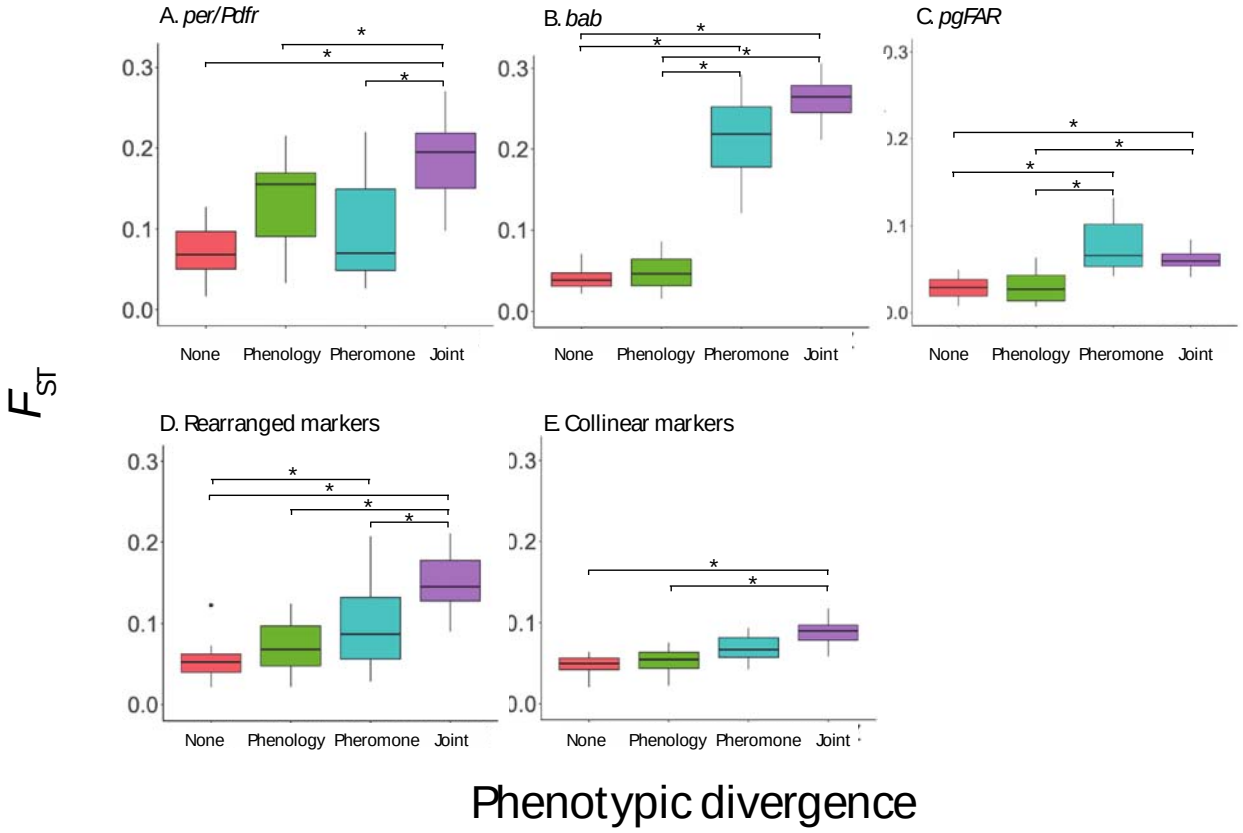

Figure 3. A. Genetic differentiation (F_{ST}) at phenology barrier loci *per* and *Pdfr* between population pairs lacking differences in barrier traits (Z/Z; red) or differing in phenology (BZ/UZ; green), pheromone signaling (BE/BZ; blue), or both phenology and pheromone (BE/UZ; purple). B. Genetic differentiation at pheromone response barrier locus *bab* across population pairs. C. Genetic differentiation at pheromone production barrier locus *pgFAR* across population pairs. D. Genetic differentiation at neutral Z chromosome loci inside a ~10 Mb "rearranged" region of suppressed recombination across population pairs. E. Genetic differentiation at neutral Z chromosome loci outside a ~10 Mb region of recombination suppression ("collinear" region) across population pairs. Asterisks indicate significantly different comparisons.

With respect to barrier loci, the phenology loci (*per/Pdfr*) had significantly elevated F_{ST} in jointly diverged population pairs compared to population pairs without phenology differences (p < 0.0012), as well as those differing only in phenology (p = 0.01, Figure 3A). F_{ST} in the *bab* region significantly increased in all populations that differed in pheromone compared to populations that did not (p < 0.0002, Figure 3B), but F_{ST}

was not quite significantly increased upon joint divergence compared to pheromone alone (p = 0.096). F_{ST} in the *pgFAR* region was broadly similar to that seen at *bab* across population pairs. The 100 kb window with the highest mean F_{ST} on the Z chromosome in populations differing only in phenology overlapped the voltinism gene *per* (scaffold532: F_{ST} = 0.23; Figure S4), in populations differing only in pheromone signaling it was within the *bab* region (scaffold50: F_{ST} = 0.39), and the highest window in jointly diverged populations was inside the region of recombination suppression (scaffold804: F_{ST} = 0.61).

Targeted amplicon sequencing

The number of distinct genotypic clusters and degree of admixed ancestry varied across localities by the number of acting barriers. The most likely number of clusters across populations was two for all STRUCTURE runs (K = 2, Evanno et al. 2005). The two locations lacking any known barrier effect (Table 1) showed minimal admixture (Figure 4), and for all 73 individuals, more than 98% of their Z chromosome loci belonged to the same cluster (Q2 > 0.98; Figure 4A). The two locations differing in pheromone but not phenology (BE/BZ) showed evidence of two genotypic clusters and greater admixture. At these sites, 22% (19 of 83) of individuals showed admixture (defined as 0.1 < Q < 0.9). The two locations where pheromone and phenology could act coincidentally (BE/UZ) also showed two genotypic clusters, but unlike at sites where insects differed only in pheromone, most individuals showed limited admixture of the Z chromosome and the proportion of admixed individuals was reduced. 96% of individuals at BE/UZ sites had a high probability of belonging to either cluster 1 (N = 21, Q1 ≥ 0.9) or cluster 2 (N = 37, Q2 ≥ 0.9).

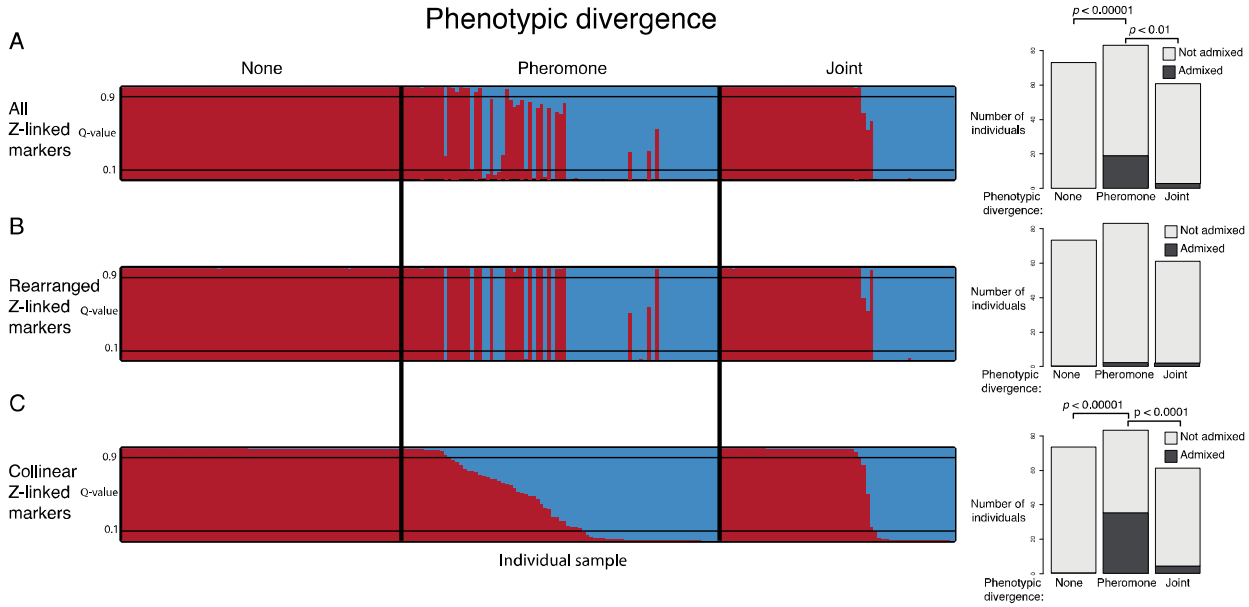

Figure 4. STRUCTURE plots partitioned by population category. Plotted on the right are the number of individuals that are considered admixed (Q-value greater than 0.1 and less than 0.9) and not-admixed (Q value greater than 0.9 or less than 0.1). A. All Z chromosome markers; B. rearranged Z markers; C. collinear Z markers. Red indicates assignment to cluster 1 (Z clade), blue indicates assignment to cluster 2 (E clade).

Separate analyses of Z chromosome loci inside and outside the rearranged region showed that admixture depended on marker location (Figure 4), in addition to barriers to gene flow. Population types significantly differed in the degree of admixed collinear regions ($p = 7.63 \times 10^{-14}$), with an excess of admixture at sites with insects differing only in pheromone compared to sites without phenotypic divergence or sites with joint differences in phenology and pheromone. Admixture of rearranged regions was similarly low across all three population types (p = 0.386). At sites where insects differed only in pheromone, nearly half (42%) of 83 individuals were estimated to have admixed collinear loci (0.1 < Q < 0.9), whereas only 2 showed appreciable admixture at loci in the rearranged region ($p = 1.64 \times 10^{-10}$). In contrast, at sites differing in both pheromone signaling and phenology, both collinear and rearranged regions were relatively free of admixture and the

estimated proportion of admixed Z chromosomes did not significantly differ (3.3% vs. 6.6%, p = 0.68). Consistent with a lower per generation rate of admixture across the rearrangement, a large percentage of individuals of mixed ancestry inferred from the rearranged region resembled F_1 hybrids (Q = 0.4 – 0.6; 3 of 4 admixed, 75%) compared to regions with normal recombination rates (11 of 39 admixed, 28.2%) (p = 0.094). The former may represent F_1 heterokaryotype male offspring.

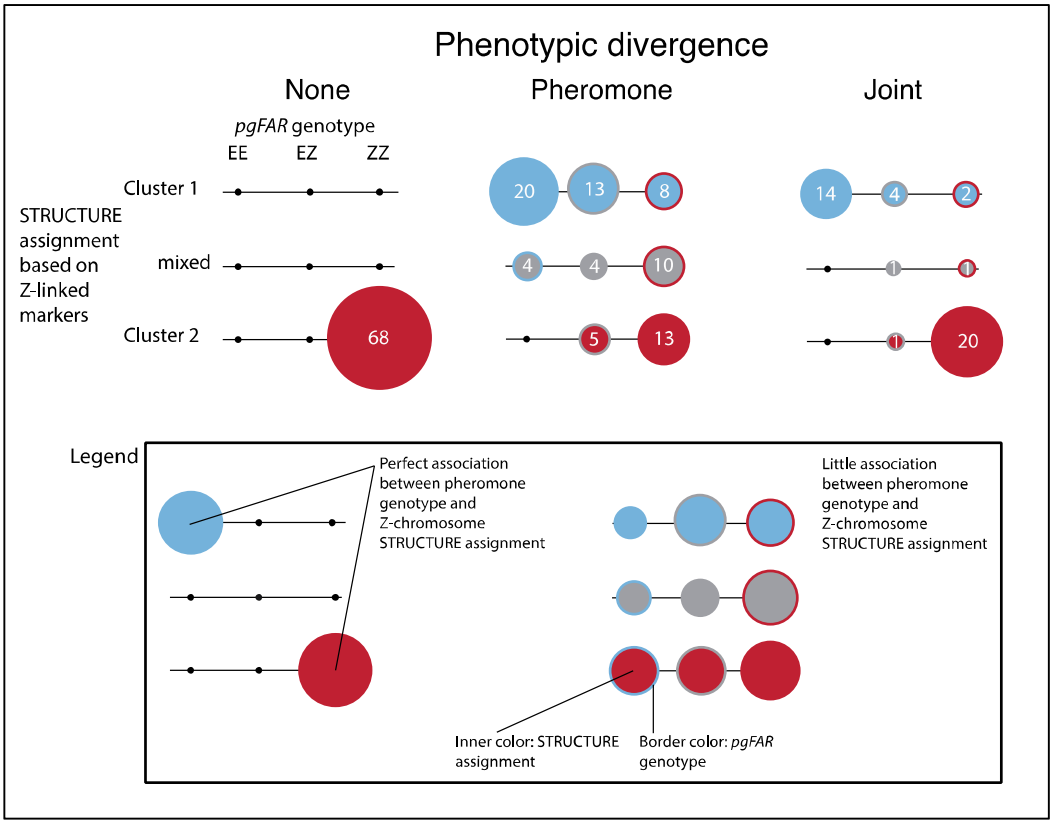

Figure 5. The accumulation of barrier traits leads to increase concordance between *pgFAR* and STRUCTURE clade assignment. Bead-and-string plots depict association between sample genotypes at *pgFAR*, the autosomal locus controlling pheromone production (border color) and Z chromosome marker-based STRUCTURE assignment (inner color). Samples were considered E-clade if they were assigned to an E-strain associated STRUCTURE background at a 0.9 threshold or above, mixed if between 0.1 and 0.9, and Z-clade if below 0.1 (see Figure 4). Number of individuals listed inside beads.

We evaluated Z chromosome admixture relative to allelic associations at the autosomal gene *pgFAR* and found that sex-linked genotypic clusters 1 and 2 mainly corresponded to E and Z pheromone alleles, respectively. Moths assigned with high probability to either of the two clusters (Q1 or Q2 $\geq$ 0.9, N = 168) were usually homogenous for E or Z alleles at *pgFAR* (145; 86%), whereas 25 percent ofmoths showing evidence of admixture (0.1 < Q < 0.9) had heterozygote *pgFAR* SNP genotypes. We used a simple network to visualize the E genotypic cluster (Q1 > 0.9), the Z genotypic cluster (Q2 > 0.9), and the admixed cluster (0.1 < Q1 < 0.9) separated by *pgFAR* genotype (Figure 5). Most moths demonstrating a mismatch between their autosomal *pgFAR* genotype and their genotypic cluster assignment (N = 48) were collected at sites differing only in pheromone (40 of 77) (Figure 5). In contrast, sites differing in both pheromone and phenology had fewer discordant individuals (8 of 43). Results from the PCA of the Z chromosome (Figure 6, Figure S3) were consistent with these results. The first coordinate axis was broadly consistent with *pgFAR* assignment (Figure S3), and whereas intermediates in coordinate space of an analysis of collinear markers were primarily admixed individuals at BE/BZ sites (Figure 6A), few intermediates occurred when analyzing solely rearranged loci (Figure 6B).

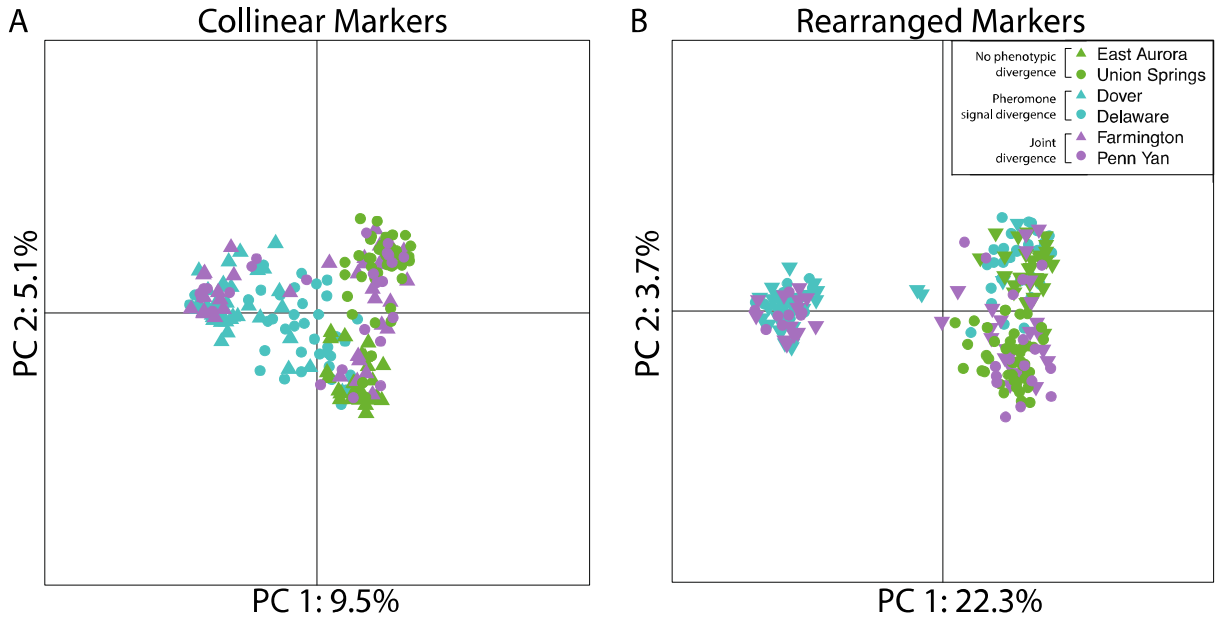

Figure 6. Principal components analyses of (A) collinear and (B) rearranged Z chromosome markers. Individual labels are based on geographic location and population type. Percentages indicate the amount of variance explained by that axis.

LD across the entire Z chromosome increased as the number of barrier traits increased. r^2 values were low at locations without barriers (mean r^2 =0.027), but progressively grew with the presence of a single barrier (~120%) (BE/BZ, mean $r^2 = 0.06$, $p = 1.00 \times 10^{-20}$) and coincident barriers (~310%) (BE/UZ, mean 0.112, $p = 5.64 \times 10^{-54}$) (Figure 7). Finally, LD between *pgFAR* and *Pdfr* increased primarily upon joint divergence, as did LD between *pgFAR* and the rest of the Z chromosome (Figure 8). Splitting loci by collinear and rearranged location showed that the influence of barrier effects on LD were more pronounced in rearranged intervals (Figure S5).

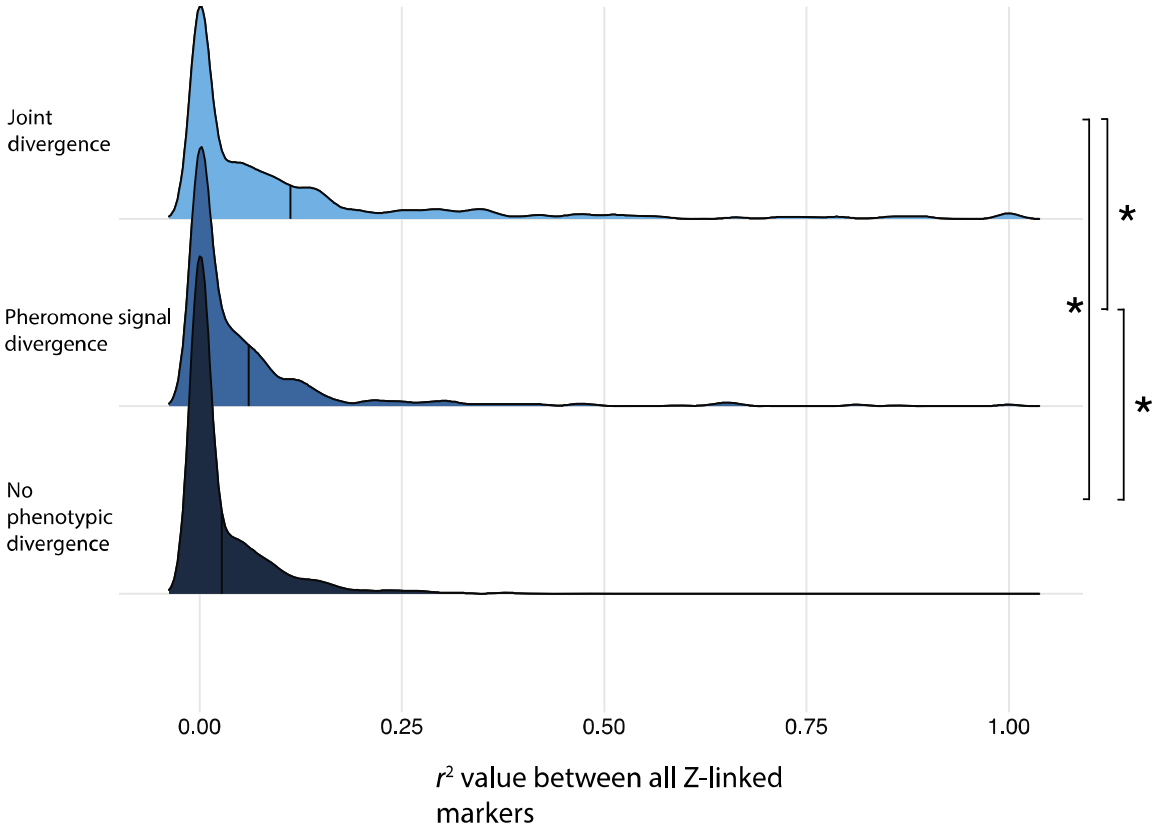

Figure 7. Linkage disequilibrium (r^2) between all Z chromosome SNPs separated by population type. The mean value is indicated by a vertical black line. Significantly different distributions measured using Wilcox tests are marked with asterisks (*).

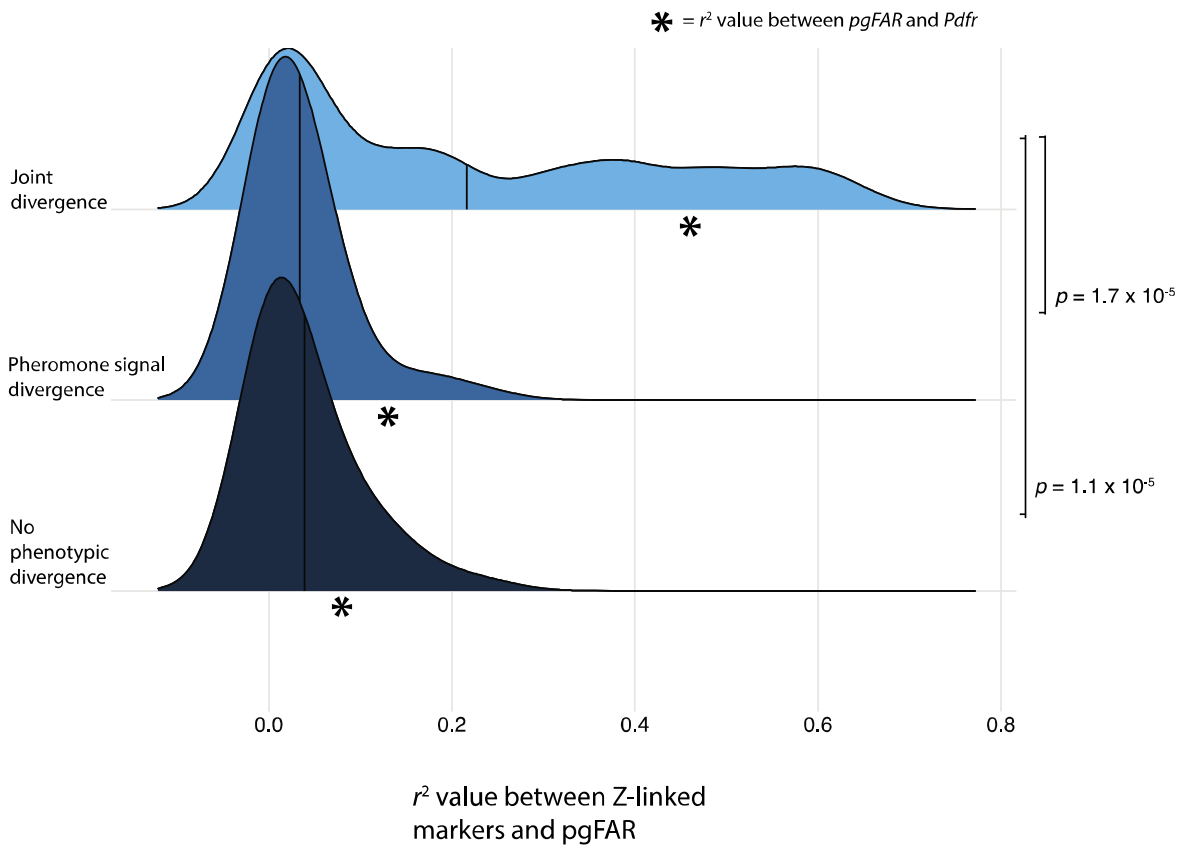

Figure 8. Linkage disequilibrium (r^2) calculated between the autosomal *pgFAR* gene and SNPs on the Z-chromosome, separated by population type. The mean value is indicated by a vertical black line. The star depicts mean LD between SNPs at *pgFAR* and the Z-linked *Pdfr* gene. Significance from Mann-Whitney U tests comparing population types are shown.

DISCUSSION

The origin of distinct genotypic clusters as speciation progresses is thought to depend on 1) the number and genomic architecture of segregating barrier loci, 2) strength of selection or assortative mating on each locus, and 3) features that reduce the rate of recombination among barrier loci (Butlin et al. 2021, Felsenstein 1981, Nosil et al. 2017, Seehausen et al. 2014). The ECB moth system is interesting because many of these factors are at least partially understood. Genomic locations of barrier loci and a Z-chromosomal rearrangement are known, the strength of relevant barriers have been estimated, and barrier effects among populations can be independent or coincident, allowing insight into how the transition from a weaker barrier(s) to a stronger overall barrier relates to the emergence of exclusive genetic groups.

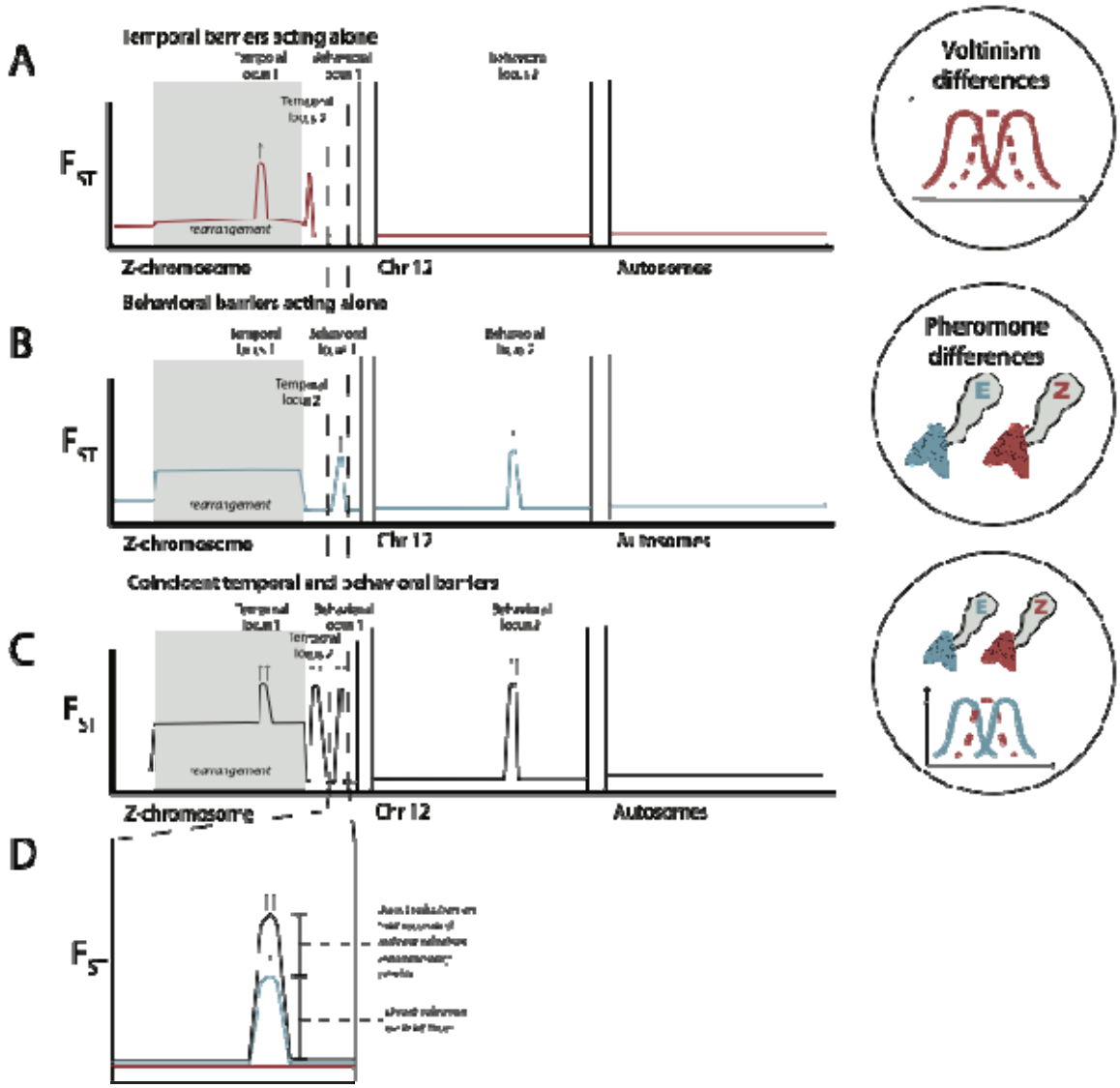

Figure 9. A graphical overview of genomic differentiation as barriers to gene flow accumulate. Patterns reflect key findings but are not real data. A. Differentiation at locations differing only in phenology (temporal barrier), B. only in pheromone signaling (behavioral barrier), or C. joint divergence along both phenotypic axes (temporal and behavioral barriers). D. An illustration of how coincidence shapes differentiation of barrier loci by allowing both direct and indirect selection.

Phenology divergence alone

Our findings indicate that a >20 day shift in phenology between univoltine Z (UZ) and bivoltine Z (BZ) populations results largely in regions of differentiation overlapping Z-linked *per* and *Pdfr* loci controlling the timing of diapause termination (Figure 2, 3A, 3C, 3D). Lack of widespread genetic differentiation was somewhat unexpected because the estimated strength of assortative mating using Coyne and Orr's (1989) method indicates intermediate reproductive isolation (mean *RI* = 0.66, Dopman et al. 2010). Moreover, as a 'magic trait,' reproductive isolation would be enhanced if hybrid individuals suffer low fitness because their intermediate phenology causes a temporal mismatch with the seasonal environment (Kozak et al. 2017; 2019; Servedio et al. 2011). Among interacting populations, the chromosomal rearrangement spanning ~10 Mb of the Z-chromosome (from ~6 Mb to ~16 Mb) (Wadsworth et al. 2015; Kozak et al. 2017) might conceivably help reduce the number of such intermediate genotypes, as it includes *per* located at ~13.5 Mb (but not *Pdfr* at ~18

Mb) (Kozak et al. 2019; Unbehend et al. 2021). Yet we found limited evidence of enhanced differentiation in this part of the Z chromosome (Figure 3C), which would have been expected if a rearrangement differed in frequency and was helping maintain high fitness genotypes by preventing recombination (Kirkpatrick & Barton 2006).

Shallow phenology-associated genetic structure is also seen in other insect systems, for example summer and winter forms of the pine processionary moth (Santos et al. 2007, 2011), as well as marine midge populations living in different phenological niches (Fuhrmann et al. 2021). One possible explanation for weak, localized genetic isolation in these systems is if temporal barriers are commonly conditionally dependent, and therefore unstable. Like other insect species, year-to-year differences in weather alters breeding cycle phenology in the ECB moth (Beck 1983), which can lead to occasional increased overlap of allochronic mating flights or possibly ease fitness consequences of intermediate hybrid phenology, thereby allowing genomic regions not tightly linked to selected barrier loci to be exchanged. In possible agreement with conditional dependence, Dopman et al. (2010) found the barrier effect to vary by as much as two-fold across years. Despite the potential for divergent selection on phenology to efficiently reduce gene flow, allochrony does not therefore appear to be strong enough to promote divergence by itself, a question recently considered across a variety of taxa (Taylor and Friesen 2017). Instead it may more commonly function when combined with other barriers, such as habitat or resource partitions (Doellman et al. 2018), or when reinforcing pre-existing divergence (Lowry et al. 2008; Møller et al. 2011).

Behavioral divergence alone

There is general agreement that many of the ~160,000 moth species remain separate because of changes in pheromone preference and pheromone blend. However, as with ongoing discussions on phenology, current debate centers on whether this phenotypic axis suffices to initiate the origin of new species (reviewed in Allison and Cardé 2016). The buildup of genome-wide differentiation that ultimately characterizes species is thought to be favored by 'multiple-effect traits' like pheromone signaling in ECB moths, because they are highly efficient at reducing gene flow (Smadja and Butlin 2011). Indeed, the two routes of reproductive isolation that evolved from pheromone divergence (mate choice and behavioral sterility) are predicted to eliminate almost 90% of gene flow between Z and E strains, at least under laboratory conditions (mean *RI* = 89; Glover et al. 1991; Dopman et al. 2010). In the field, Unbehend et al. (2021) recently showed that male moths carry either E type or Z type alleles at *bab* (pheromone preference) and at *pgFAR* (pheromone blend) genes ~80% of the time, indicating coupled pre- and post-zygotic isolation is strong enough in nature to counteract randomization of these physically unlinked genes.

Our findings corroborate prior inferences of selection and assortative mating on *bab* and *pgFAR* barrier loci in field populations (Figure 3B, 3C), but they go on to suggest that divergence of pheromone signaling alone may not necessarily lead to the emergence of genome-wide differentiation (Figure 2). Nevertheless, bivoltine E (BE) and bivoltine Z (BZ) populations do exhibit broad differentiation of chromosomes harboring *bab* and *pgFAR* genes (Z chromosome and autosome 12) (Figure 2, 3D, 3E), as well as long-range LD between sympatric populations (Figure 7). Differentiation was unexpectedly high within rearranged regions of the Z chromosome (Figure 3D, 3E), despite *bab* (~18.7 Mb) being collinearly located ~3 Mb away from the nearest estimated rearrangement breakpoint (Figure S1). This might indicate that divergent selection acts on another component of the pheromone system controlled by loci within the rearrangement. If true, two prime suspects are the tandemly repeated odorant receptor loci *OR4* and *OR6* (~12.14 Mb) required for physiological perception of E and Z pheromone components within the male antenna (Koutroumpa et al. 2014). Overall, the apparent stability of chromosome but not genome-wide differentiation implies the sex pheromone system of moths might reduce gene flow enough to allow a partial build-up of differentiation, but this phenotypic axis alone may not lead to substantial progress towards the emergence of well-defined species.

Joint phenotypic divergence

Our analyses of bivoltine E (BE) and univoltine Z (UZ) moths indicate that joint divergence along multiple phenotypic axes (phenology and pheromone signaling) is associated with intensified genetic differentiation. The expected progression towards enhanced coupling as behavioral and temporal barriers coincide in sympatry was detected as elevated LD between unlinked behavioral barrier and temporal barrier genes (*pgFAR* and *Pdfr*) (Figure 8). Differentiation at temporal barrier loci (*per*, *Pdfr*) increased in the presence of behavioral divergence (Figure 3A, phenology vs. joint divergence). Likewise for the behavioral barrier locus *bab* in the presence of phenology divergence (Figure 3B, pheromone vs. joint divergence). Unlike *bab*, differentiation was not elevated at *pgFAR* (Figure 3C, pheromone vs. joint divergence), a result likely explained by exclusion of *pgFAR* heterozygotes when creating phenotypically divergent population pools for sequencing. As shown by the amplicon data (Figure 5), *pgFAR* heterozygotes are more than twice as frequent at sites differing only in pheromone signaling (BE/BZ), compared to sites that also differ in phenology (BE/UZ). Therefore their removal may have artificially inflated estimated *pgFAR* genetic differentiation in BE/BZ comparisons (e.g., Landisville, Dover). In contrast, levels of differentiation at this locus in comparisons of jointly diverged sites may be more reflective of reality, since heterozygotes are relatively rare (e.g., Rockspring, Harner).

The wider genomic landscape appears to be similarly influenced by joint divergence. On the Z chromosome, both long-range LD (Figure 7) and inter-chromosomal LD with *pgFAR* (Figure 5, 8, S4) increased upon the coincidence of behavioral and temporal barriers in sympatry. Furthermore, whereas evidence of reduced admixture (or greater clustering) on the Z was mainly limited to the rearrangement when sympatric populations differed in pheromone signaling, this signature expanded into collinear regions when sympatric populations also differed in phenology (Figure 4, Figure 6). Finally, overall differentiation of the two chromosomes containing the four barrier loci (Z chromosome and autosome 12) increased by more than two-fold (Figure 2), although differentiation is disproportionate on the Z chromosome where coupling leads to a 'large Z-effect' and divergence of three barrier loci (*per–Pdfr–bab* complex), compared to just one on chromosome 12 (*pgFAR*).

All these features are congruent with prior estimates of cumulative reproductive isolation. When differences in sex pheromone communication as well as phenology act as barriers, their cumulative effects are estimated to result in almost complete reproductive isolation (mean $RI = 0.96$, Dopman et al. 2010). The process underlying heightened genetic differentiation is unlikely to simply involve a sudden increase in fitness effect on each barrier locus (direct selection). Instead, as barrier loci increase their non-random associations during coupling, a higher mean fitness differential can be achieved between alternative multi-locus genotypes (indirect selection). Since each locus can have its own fitness effect and also be indirectly influenced by the fitness effect of other loci (Barton 1983), overall selection is more effective at each locus, as is the overall barrier to gene flow at each locus, making it more difficult for neutral alleles to cross population boundaries (Barton and Bengtsson 1986).

Conclusions

We argue that our study provides support for the coupling hypothesis of speciation (Felsenstein 1981; Smadja and Butlin 2011; Seehausen et al. 2014) and helps close a gap between theoretical predictions of this hypothesis and empirical systems. Our population comparisons show that whether barriers are independent or coupled make important contributions to a hallmark of the speciation process, the build-up of genomic differentiation. Genetic distances fit expectations from estimates of reproductive isolation, with locus-specific differentiation associated with the weakest barrier (temporal), differentiation across much of a chromosome with the barrier of moderate strength (behavioral), and chromosome-wide differentiation when the two barriers operate cumulatively. These results demonstrate that the joint action of multiple, coincident barrier effects lead to levels of genomic differentiation that far exceed those of single barriers acting alone, consistent with theory in which coupling allows indirect selection to combine with direct selection and thereby result in a stronger overall barrier.

Finding chromosome-wide but not genome-wide differentiation when populations diverge along multiple phenotypic axes suggests that the overall barrier upon coupling may only partly outweigh gene flow and

recombination. Thus, a critical tipping point or threshold of coupling that could allow for sudden differentiation of most of the genome was not reached (Nosil et al. 2017). When barriers are coincident, much of the genome appears well mixed and moderate genetic differentiation (F_{ST} = 0.05 – 0.15) is observed for only two chromosomes containing barrier loci, making further divergence challenging without evolution of either more barrier loci or stronger selection/assortative mating. Simulations demonstrate that, in certain cases, many divergently selected loci may be required for genome-wide differentiation (Feder et al. 2014, Flaxman et al. 2014). Some empirical studies seem to support these predictions, including Midas cichlids, where a transition from locus-specific to genome-wide divergence coincides with a transition from simple to polygenic barriers (Kautt et al. 2020). Strengthening of the overall barrier in the ECB moth could in time be facilitated by the evolution highly polygenic traits, perhaps in the form of intrinsic post-zygotic barriers. However, it is also possible that many intermediate stages of speciation seen in nature, including the ECB moth, are either stably maintained or even temporary, and that barrier traits may not remain coincident should ecological or geographic shifts take place (Seehausen 2006). If so, then coupling and current population differentiation will disappear or rebound according to external environmental factors.

Understanding how and why barrier traits evolve and become coincident or disappear remains as a critical problem for speciation biology, and few empirical systems have been able to directly assess coincidence of known barrier traits in different stages of coupling. Future work on this study system will reconstruct the evolutionary and demographic history of this group and provide a framework for inferring the timing and selective history of genes underlying barrier phenotypes. Ultimately, we must reconstruct not only the sequence of evolution of individual barriers, but also the timing and causal factors responsible for their coupling. Although byproduct explanations have been proposed that do not require coupling itself to be selectively favored (Butlin and Smadja 2017), adaptive coupling seems likely in the ECB moth and can be experimentally studied by measuring fitness of full allelic combinations in nature. Indirect approaches are equally promising for identifying factors associated with changes of coupling in nature. Since it is now possible in this system to precisely estimate population differentiation and the extent of coupling of all major barrier loci (*bab, pgFAR, per, Pdfr*) across various spatial, ecological, and historical contexts, the causal links that establish barrier coupling and facilitate neutral divergence during speciation appear to be within reach.

LITERATURE CITED

Allison, J.D. & Cardé, R.T. 2016 Pheromones: reproductive isolation and evolution in moths. Pheromone Communication in Moths (pp. 11-24). Berkeley: University of California Press.

Anglade, P., Stockel, J., & Cooperators, I.W.G.O. 1984. Intraspecific sex-pheromone variability in the European corn borer, Ostrinia nubilalis Hbn. (Lepidoptera, Pyralidae). Agronomie, 4(2), 183-187.

Barton, N.H. 1983. Multilocus clines. Evolution, 37(3), 454-471.

Barton, N. & Bengtsson, B.O. 1986. The barrier to genetic exchange between hybridising populations. Heredity, 57(3), 357-376.

Barton, N.H. & De Cara, M.A.R. 2009. The evolution of strong reproductive isolation. Evolution, 63(5), 1171-1190.

Barton, N.H. 2013. Does hybridization influence speciation? Journal of Evolutionary Biology, 26(2), 267–269.

Bethenod, M.T., Thomas, Y., Rousset, F., Frérot, B., Pélozuelo, L., Genestier, G., & Bourguet, D. 2005. Genetic isolation between two sympatric host plant races of the European corn borer, Ostrinia nubilalis Hübner. II: assortative mating and host-plant preferences for oviposition. Heredity, 94(2), 264-270.

Bourguet, D., Genissel, A., & Raymond, M. 2000. Insecticide resistance and dominance levels. Journal of Economic Entomology, 93(6), 1588-1595.

Brousseau, L., Nidelet, S., & Streiff, R. 2018. New WGS data and annotation of the heterosomal vs. autosomal localization of Ostrinia scapulalis (Lepidoptera, Crambidae) nuclear genomic scaffolds. Data in Brief, 20, 644-648.

Butlin, R.K. & Smadja, C.M. 2018. Coupling, reinforcement, and speciation. The American Naturalist, 191(2), 155-172.

Butlin, R.K., Servedio, M.R., Smadja, C.M., Bank, C., Barton, N.H., Flaxman, S.M., ... & Qvarnström, A. 2021. Homage to Felsenstein 1981, or why are there so few/many species? Evolution, 75(5), 978-988.

Caffrey, D.J. and Worthley, L.H., 1927. A progress report on the investigations of the European corn borer (No. 1476). Washington, D.C.: US Government Printing Office.

Calcagno, V., Thomas, Y., & Bourguet, D. 2007. Sympatric host races of the European corn borer: adaptation to host plants and hybrid performance. Journal of Evolutionary Biology, 20(5), 1720-1729.

Calvin D.D. & Song P.Z. 1994. Variability in postdiapause development periods of geographically separate Ostrinia nubilalis (Lepidoptera: Pyralidae) populations in Pennsylvania. Environmental Entomology 23(2), 431–436.

Coates, B.S., Johnson, H., Kim, K.S., Hellmich, R.L., Abel, C.A., Mason, C., & Sappington, T.W. 2013. Frequency of hybridization between Ostrinia nubilalis E□and Z□pheromone races in regions of sympatry within the United States. Ecology and Evolution, 3(8), 2459-2470.

Coates, B.S. & Siegfried, B.D. 2015. Linkage of an ABCC transporter to a single QTL that controls Ostrinia nubilalis larval resistance to the Bacillus thuringiensis Cry1Fa toxin. Insect Biochemical Molecular Biology, 63, 86–96.

Coates, B.S., Kozak, G.M., Kim, K., Sun, J., Wang, Y., Fleischer, S.J., Dopman, E.B., &Sappington, T.W. 2019. Influence of host plant, geography and pheromone strain on genomic differentiation in sympatric populations of Ostrinia nubilalis. Molecular Ecology, 28(19), 4439-4452.

Coyne, J.A. & Orr, H.A. 1989 Two rules of speciation. Speciation and Its Consequences (pp. 180–207). Sunderland, MA: Sinauer.

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

Dagilis, A.J. & Kirkpatrick, M. 2016. Prezygotic isolation, mating preferences, and the evolution of chromosomal inversions. Evolution, 70(7), 1465-1472.

Dray, S. & Dufour, A.B. 2007. The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software, 22(4), 1-20.

Doellman, M.M., Ragland, G.J., Hood, G.R., Meyers, P.J., Egan, S.P., Powell, T.H., Lazorchak, P., Glover, M.M., Tait, C., Schuler, H., & Hahn, D.A. 2018. Genomic differentiation during speciation-with-gene-flow: Comparing geographic and host-related variation in divergent life history adaptation in Rhagoletis pomonella. Genes, 9(5), 262.

Dopman, E.B., Bogdanowicz, S.M., & Harrison, R.G. 2004. Genetic mapping of sexual isolation between E and Z pheromone strains of the European corn borer (Ostrinia nubilalis). Genetics, 167(1), 301-309.

Dopman, E.B., Pérez, L., Bogdanowicz, S.M. , & Harrison, R.G. 2005. Consequences of reproductive barriers for genealogical discordance in the European corn borer. Proceedings of the National Academy of Sciences USA, 102(41), 14706-14711.

Dopman, E.B., Robbins, P.S., & Seaman, A., 2010. Components of reproductive isolation between North American pheromone strains of the European corn borer. Evolution, 64(4), 881-902.

Feder, J.L., Nosil, P., Wacholder, A.C., Egan, S. P., Berlocher, S.H., & Flaxman, S.M. 2014. Genome-wide congealing and rapid transitions across the speciation continuum during speciation with gene flow. Journal of Heredity, 105(S1), 810-820.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. Journal of Molecular Evolution, 17(6), 368-376.

Fisher, K.E., Mason, C.E., Flexner, J.L., Hough□Goldstein, J., & McDonald, J.H. 2017. Survivorship of Z pheromone strain European corn borer (Lepidoptera: Crambidae) on a range of host plants varying in defensive chemistry. Journal of Economic Entomology, 110(3), 978–985.

Flaxman, S.M., Wacholder, A.C., Feder, J.L., & Nosil, P. 2014. Theoretical models of the influence of genomic architecture on the dynamics of speciation. Molecular Ecology, 23(16), 4074-4088.

Frolov, A.N., Bourguet, D., & Ponsard, S. 2007. Reconsidering the taxomony of several Ostrinia species in the light of reproductive isolation: a tale for Ernst Mayr. Biological Journal of the Linnean Society, 91(1), 49-72.

Fuhrmann, N., Prakash, C., & Kaiser, T.S. 2021. Polygenic adaptation from standing genetic variation allows rapid ecotype formation. bioRxiv.

Gagnaire, P.A., Pavey, S.A., Normandeau, E., & Bernatchez, L. 2013. The genetic architecture of reproductive isolation during speciation□with□gene□flow in lake whitefish species pairs assessed by RAD sequencing. Evolution, 67(9), 2483-2497.

Glover, T.J., Campbell, M.G., Linn, C.E., & Roelofs, W.L. 1991. Unique sex chromosome mediated behavioral response specificity of hybrid male European corn borer moths. Experientia, 47(9), 980-984.

Harrison, R.G. & Larson, E.L. 2014. Hybridization, introgression, and the nature of species boundaries. Journal of Heredity, 105(S1), 795-809.

Hodgson, B.E. 1928. The host plants of the European corn borer in New England (No. 77). US Department of Agriculture.

Kautt, A.F., Kratochwil, C.F., Nater, A., Machado-Schiaffino, G., Olave, M., Henning, F., ... & Meyer, A. 2020. Contrasting signatures of genomic divergence during sympatric speciation. Nature, 588(7836), 106-111.

Kim, Y., Jung, S., Kim, Y., & Lee, Y. 2011. Real-time monitoring of oriental fruit moth, Grapholita molesta, populations using a remote sensing pheromone trap in apple orchards. Journal of Asia-Pacific Entomology, 14(3), 259-262.

Kirkpatrick, M. & Barton, N. 2006. Chromosome inversions, local adaptation and speciation. Genetics, 173(1), 419-434.

Kirkpatrick, M. & Barrett, B., 2015. Chromosome inversions, adaptive cassettes and the evolution of species' ranges. Molecular Ecology, 24(9), 2046-2055.

Klun, J.A., Chapman, O.L., Mattes, K.C., Wojtkowski, P.W., Beroza, M., & Sonnet, P.E. 1973. Insect sex pheromones: minor amount of opposite geometrical isomer critical to attraction. Science, 181(4100), 661-663.

Kroemer, J.A., Coates, B.S., Nusawardani, T., Rider, S.D., Fraser, L.M., & Hellmich, R.L. 2011. A rearrangement of the Z chromosome topology influences the sex-linked gene display in the European corn borer, Ostrinia nubilalis. Molecular Genetics & Genomics. 286, 37–56

Kofler, R., Pandey, R.V., & Schlötterer, C. 2011. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics, 27(24), 3435-3436.

Kozak, G.M., Wadsworth, C.B., Kahne, S.C., Bogdanowicz, S.M., Harrison, R.G., Coates, B.S., & Dopman, E.B. 2017. A combination of sexual and ecological divergence contributes to rearrangement spread during initial stages of speciation. Molecular Ecology, 26(8), 2331-2347.

Kozak, G.M., Wadsworth, C.B., Kahne, S.C., Bogdanowicz, S.M., Harrison, R.G., Coates, B.S., & Dopman, E.B., 2019. Genomic basis of circannual rhythm in the European corn borer moth. Current Biology, 29(20), 3501-3509.

Kochansky, J., Cardé, R.T., Liebherr, J., & Roelofs, W.L. 1975. Sex pheromone of the European corn borer, Ostrinia nubilalis (Lepidoptera: Pyralidae), in New York. Journal of Chemical Ecology, 1(2), 225-231.

Langmead, B. & Salzberg, S.L. 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods, 9(4), 357.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., & Thompson, J.D., 2007. Clustal W and Clustal X version 2.0. Bioinformatics, 23(21), 2947-2948.

Lasne, C., Sgrò, C. M., & Connallon, T. 2017. The relative contributions of the X chromosome and autosomes to local adaptation. Genetics, 205(3), 1285-1304.

Lassance, J.M., Groot, A.T., Liénard, M.A., Antony, B., Borgwardt, C., Andersson, F., Hedenström, E., Heckel, D.G., & Löfstedt, C. 2010. Allelic variation in a fatty-acyl reductase gene causes divergence in moth sex pheromones. Nature, 466(7305), 486-489.

Levy, R.C., Kozak, G.M., Wadsworth, C.B., Coates, B.S., & Dopman, E.B. 2015. Explaining the sawtooth: latitudinal periodicity in a circadian gene correlates with shifts in generation number. Journal of Evolutionary Biology, 28(1), 40-53.

Lewis, L.C. 1975. Natural regulation of crop pests in their indigenous ecosystems and in Iowa agroecosystems: Bioregulation of economic insect pests. Iowa State Journal of Research, 49, 435–445.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics, 25(16), 2078-2079.

Lowry, D.B., Rockwood, R.C., & Willis, J.H. 2008. Ecological reproductive isolation of coast and inland races of Mimulus guttatus. Evolution, 62(9), 2196-2214.

Malausa, T., Dalecky, A., Ponsard, S., Audiot, P., Streiff, R., Chaval, Y., & Bourguet, D. 2007. Genetic structure and gene flow in French populations of two Ostrinia taxa: host races or sibling species? Molecular Ecology, 16(20), 4210-4222.

Mayr, E., 1942. Systematics and the Origin of Species. New York: Columbia University Press.

Meisel, R.P. & Connallon, T. 2013. The faster-X effect: integrating theory and data. Trends in Genetics, 29(9), 537-544.

Møller, A.P., Antonov, A., Stokke, B.G., Fossøy, F., Moksnes, A., Røskaft, E., & Takasu, F. 2011. Isolation by time and habitat and coexistence of distinct host races of the common cuckoo. Journal of Evolutionary Biology, 24(3), 676-684.

Nosil, P., Feder, J.L., Flaxman, S.M., & Gompert, Z., 2017. Tipping points in the dynamics of speciation. Nature Ecology & Evolution, 1(2), 1-8.

Ortiz-Barrientos, D., Engelstädter, J., & Rieseberg, L. H. 2016. Recombination rate evolution and the origin of species. Trends in Ecology & Evolution, 31(3), 226-236.

O'Rourke, M.E., Sappington, T.W., & Fleischer, S.J. 2010. Managing resistance to Bt crops in a genetically variable insect herbivore, Ostrinia nubilalis. Ecological Applications, 20(5), 1228-1236.

Pohlert, T. 2014. The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR). R package, https://CRAN.R-project.org/package=PMCMR.

Presgraves, D.C. 2008. Sex chromosomes and speciation in Drosophila. Trends in Genetics, 24(7), 336-343.

Presgraves, D.C. 2018. Evaluating genomic signatures of “the large X effect” during complex speciation. Molecular Ecology, 27(19), 3822-3830

Pritchard, J.K., Stephens, M., & Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics, 155(2), 945-959.

Quinlan, A.R., 2014. BEDTools: the Swiss army tool for genome feature analysis. Current Protocols in Bioinformatics, 47(1), 11-12.

Riesch, R., Muschick, M., Lindtke, D., Villoutreix, R., Comeault, A.A., Farkas, T.E., ... & Nosil, P. 2017. Transitions between phases of genomic differentiation during stick-insect speciation. Nature Ecology & Evolution, 1(4), 1-13.

Santos, H., Rousselet, J., Magnoux, E., Paiva, M.R., Branco, M., & Kerdelhué, C. 2007. Genetic isolation through time: allochronic differentiation of a phenologically atypical population of the pine processionary moth. Proceedings of the Royal Society B, 274(1612), 935-941.

Santos, H., Burban, C., Rousselet, J., Rossi, J.P., Branco, M., & Kerdelhue, C. 2011. Incipient allochronic speciation in the pine processionary moth (Thaumetopoea pityocampa, Lepidoptera, Notodontidae). Journal of Evolutionary Biology, 24(1), 146-158.

Schilling, M.P., Mullen, S.P., Kronforst, M., Safran, R.J., Nosil, P., Feder, J.L., ... & Flaxman, S.M. 2018. Transitions from single-to multi-locus processes during speciation with gene flow. Genes, 9(6), 274.

Seehausen, O., Butlin, R.K., Keller, I., Wagner, C.E., Boughman, J.W., Hohenlohe, P.A., ... & Widmer, A. 2014. Genomics and the origin of species. Nature Reviews Genetics, 15(3), 176-192.

Servedio, M.R., Van Doorn, G.S., Kopp, M., Frame, A.M., & Nosil, P. 2011. Magic traits in speciation: 'magic' but not rare? Trends in Ecology & Evolution, 26(8), 389-397.

Showers, W.B., Reed, G.L., Robinson, J.F., & Derozari, M.B. 1976. Flight and sexual activity of the European corn borer. Environmental Entomology, 5(6), 1099-1104.

Showers, W.B. 1981. Geographic variation of the diapause response in the European corn borer. Insect Life History Patterns (pp. 97-111). New York: Springer.

Showers, W.B. 1993. Diversity and variation of European corn borer populations. Evolution of Insect Pests: Patterns of Variation (pp. 287–309). Chichester, West Sussex, UK: Wiley.

Smith, H.E. 1920. Broom Corn, the probable host in which Pyrausta nubilalis, Hubner., reached America. Journal of Economic Entomology, 13(5),425-430.

Sorenson, C.E., Kennedy, G.G., Van Duyn, W., Bradley Jr, J.R., & Walgenbach, J.F. 1992. Geographical variation in pheromone response of the European corn borer, Ostrinia nubilalis, in North Carolina. Entomologia Experimentalis et Applicata, 64(2), 177-185.

Taylor, R.S., & Friesen, V.L. 2017. The role of allochrony in speciation. Molecular Ecology, 26(13), 3330-3342.

Thomas, Y., Bethenod, M.T., Pelozuelo, L., Frérot, B., & Bourguet, D. 2003. Genetic isolation between two sympatric host□plant races of the European Corn Borer, Ostrinia nubilalis Hubner I. Sex pheromone, moth emergence timing, and parasitism. Evolution, 57(2), 261-273.

Unbehend, M., Kozak, G. M., Koutroumpa, F., Coates, B. S., Dekker, T., Groot, A. T., Heckel, D. G., & Dopman, E. B. 2021. bric à brac controls sex pheromone choice by male European corn borer moths. Nature Communications 12(1), 1-11.

Wadsworth, C.B., Li, X., & Dopman, E.B., 2015. A recombination suppressor contributes to ecological speciation in Ostrinia moths. Heredity, 114(6), 593-600.

Xu, M. & Shaw, K. L. 2019. Genetic coupling of signal and preference facilitates sexual isolation during rapid speciation. Proceedings of the Royal Society B, 286(1913), 20191607.

Xu, M. & Shaw, K. L. 2021. Extensive linkage and genetic coupling of song and preference loci underlying rapid speciation in Laupala crickets. Journal of Heredity, 112(2),204-213.

Yasukochi, Y., Ohno, M., Shibata, F., Jouraku, A., Nakano, R., Ishikawa, Y., & Sahara, K. 2016. A FISH-based chromosome map for the European corn borer yields insights into ancient chromosomal fusions in the silkworm. Heredity 116(1), 75-83.

Zaman, M.F. 2008. A comparison of univoltine and multivoltine European corn borer (Ostrinia nubilalis Hübner): Life history characters, Bt toxin susceptibility, parasitoid impact, and population pattern. Doctoral dissertation. Pennsylvania State University, University Park, Pennsylvania.

TABLES

Table 1. Characteristics and locations for pooled data.

Site	Pheromone	Voltinism	Number Males	Number Females	Collection Method	Coverage after alignment
Bouckville-Madison, NY	Z	Univoltine	14	6	DI	37.0
Bellona, NY	E	Bivoltine	31	0	PT	53.18
Bellona, NY	Z	Univoltine	25	0	PT	41.52
Dover, MA	Z	Bivoltine	17	5	DD	15.83
East Aurora, NY	Z	Bivoltine	16	18	DI	32.35
Geneva, NY	Z	Univoltine	13	12	DI	5.94
Harner, PA	E	Bivoltine	31	0	PT	7.25
Harner, PA	Z	Univoltine	27	0	PT	4.59
Landisville, PA	E	Bivoltine	41	0	PT	19.29
Landisville, PA	Z	Bivoltine	39	0	PT	5.04
Penn Yan, NY	Z	Univoltine	11	15	DI	23.54
Rockspring, PA	E	Bivoltine	34	0	PT	10.36
Rockspring, PA	Z	Univoltine	33	0	PT	10.26
Ways, PA	E	Bivoltine	37	0	PT	13.33

PT = pheromone trap, DD = direct-developing larvae, DI = Diapausing larvae

Table 2. Characteristics and locations for amplicon sequencing data.

				***Frequency of pgFAR* genotypes**			
Site	**Population Type**	**Number Females**	**Number Males**	**EE**	**EZ**	**ZZ**	**Difference in average eclosion time between *pgFAR* ZZ and EE insects (Days)** †
Dover, MA	Bivoltine E / Bivoltine Z	22	27	50.0%	29.5%	20.5%	‡
East Aurora, NY	Bivoltine Z	18	15	0.0%	3.5%	96.5%	-
Farmington, NY	Bivoltine E / Univoltine Z	26	19	27.8%	8.3%	63.9%	20.1
Laurel & Newark, DE	Bivoltine E / Bivoltine Z	51	0	12.8%	34.0%	53.2%	1.74
Penn Yan, NY	Bivoltine E / Univoltine Z	20	9	24.0%	16.0%	60.0%	34.63
Union Springs, NY	Univoltine Z	24	29	0.0%	0.0%	100.0%	-

† Time to adult eclosion for diapausing insects (26˚C, 16h light) acts as a proxy for post-diapause development time

‡ Time to eclosion data was not available for these samples, but was calculated for the subsequent F_1 generation in the lab and confirmed to be bivoltine (see Material and Methods for more detail).