Influence of host plant, geography and pheromone strain on genomic differentiation in sympatric populations of Ostrinia nubilalis

Brad S. Coates
U.S. Department of Agriculture, coates@iastate.edu

Genevieve M. Kozak
University of Massachusetts at Dartmouth

Kyung Seok Kim
Iowa State University, kkim@iastate.edu

Jing Sun
Iowa State University

Yangzhou Zhang
Jilin Academy of Agricultural Sciences

Follow this and additional works at: https://lib.dr.iastate.edu/ent_pubs

Part of the Behavior and Ethology Commons, Entomology Commons, Genetics Commons, and the Population Biology Commons

The complete bibliographic information for this item can be found at https://lib.dr.iastate.edu/ent_pubs/547. For information on how to cite this item, please visit http://lib.dr.iastate.edu/howtocite.html.
Influence of host plant, geography and pheromone strain on genomic differentiation in sympatric populations of Ostrinia nubilalis

Abstract
Patterns of mating for the European corn borer (Ostrinia nubilalis) moth depend in part on variation in sex-pheromone blend. The ratio of (E)-11- and (Z)-11-tetradecenyl acetate (E11- and Z11-14:OAc) in the pheromone blend that females produce and males respond to differs between strains of O. nubilalis. Populations also vary in female oviposition preference for and larval performance on maize (C4) and non-maize (C3) host plants. The relative contributions of sexual and ecological trait variation to the genetic structure of O. nubilalis remains unknown. Host-plant use (13C/14C ratios) and genetic differentiation were estimated among sympatric E and Z pheromone strain O. nubilalis males collected in sex-pheromone baited traps at 12 locations in Pennsylvania and New York between 2007 and 2010. Among genotypes at 65 single nucleotide polymorphism marker loci, variance at a position in the pheromone gland fatty acyl-reductase (pgfar) gene at the locus responsible for determining female pheromone ratio (Pher) explained 64% of the total genetic differentiation between males attracted to different pheromones (male response, Resp), providing evidence of sexual inter-selection at these unlinked loci. Principal coordinate, Bayesian clustering, and distance-based redundancy analysis (dbRDA) demonstrate that host plant history or geography does not significantly contribute to population variation or differentiation among males. In contrast, these analyses indicate that pheromone response and pgfar-defined strain contribute significantly to population genetic differentiation. This study suggests that behavioral divergence likely plays a larger role in driving genetic variation compared to host plant-defined ecological adaptation.

Keywords
population genetics, sexual selection, reproductive isolation

Disciplines
Behavior and Ethology | Ecology and Evolutionary Biology | Entomology | Genetics | Population Biology

Comments
This is a manuscript of an article published as Coates, Brad S., Genevieve M. Kozak, Kyung Seok Kim, Jing Sun, Yangzhou Zhang, Shelby J. Fleischer, Erik B. Dopman, and Thomas W. Sappington. "Influence of host plant, geography and pheromone strain on genomic differentiation in sympatric populations of Ostrinia nubilalis." Molecular ecology (2019). doi: 10.1111/mec.15234.

Rights
Works produced by employees of the U.S. Government as part of their official duties are not copyrighted within the U.S. The content of this document is not copyrighted.

Authors
Brad S. Coates, Genevieve M. Kozak, Kyung Seok Kim, Jing Sun, Yangzhou Zhang, Shelby J. Fleischer, Erik B. Dopman, and Thomas W. Sappington

This article is available at Iowa State University Digital Repository: https://lib.dr.iastate.edu/ent_pubs/547
Influence of host plant, geography and pheromone strain on genomic differentiation in sympatric populations of Ostrinia nubilalis

Brad S. Coates 1*, Genevieve M. Kozak 2,7, Kyung Seok Kim 3, Jing Sun 4, Yangzhou Wang 5, Shelby J. Fleischer 6, Erik B. Dopman 7, and Thomas W. Sappington 1

1 USDA-ARS, Corn Insects & Crop Genetics Research Unit, Ames, IA, USA.
2 University of Massachusetts-Dartmouth, Department of Biology, Dartmouth, MA, USA.
3 Iowa State University, Department of Natural Resource Ecology and Management, Ames, IA, USA.
4 Iowa State University, Department of Entomology, Ames, IA, USA.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.15234

This article is protected by copyright. All rights reserved.
Abstract

Patterns of mating for the European corn borer (*Ostrinia nubilalis*) moth depend in part on variation in sex-pheromone blend. The ratio of (E)-11- and (Z)-11-tetradecenyl acetate (E11- and Z11-14:OAc) in the pheromone blend that females produce and males respond to differs between strains of *O. nubilalis*. Populations also vary in female oviposition preference for and larval performance on maize (C4) and non-maize (C3) host plants. The relative contributions of sexual and ecological trait variation to the genetic structure of *O. nubilalis* remains unknown.

Host-plant use ($^{13}$C/$^{14}$C ratios) and genetic differentiation were estimated among sympatric E and Z pheromone strain *O. nubilalis* males collected in sex-pheromone baited traps at 12 locations in Pennsylvania and New York between 2007 and 2010. Among genotypes at 65 single nucleotide polymorphism marker loci, variance at a position in the pheromone gland fatty acyl-reductase (*pgfar*) gene at the locus responsible for determining female pheromone ratio

This article is protected by copyright. All rights reserved.
(Pher) explained 64% of the total genetic differentiation between males attracted to different pheromones (male response, Resp), providing evidence of sexual inter-selection at these unlinked loci. Principal coordinate, Bayesian clustering, and distance-based redundancy analysis (dbRDA) demonstrate that host plant history or geography does not significantly contribute to population variation or differentiation among males. In contrast, these analyses indicate that pheromone response and pgfar-defined strain contribute significantly to population genetic differentiation. This study suggests that behavioral divergence likely plays a larger role in driving genetic variation compared to host plant-defined ecological adaptation.

Keywords: population genetics, sexual selection, reproductive isolation

1 | Introduction

Recent models of speciation predict that populations diverge primarily due to the accumulation of genetic changes coincident with local ecological adaptation (Diekmann & Doebeli 2000; Schluter 2001; Schluter & Conte 2009; Nosil 2012). For example, host plant shifts among herbivorous insects can increase genetic variance among sympatric subpopulations by disrupting random mating (Pashley 1989; Nason et al. 2002) or selecting for optimal responses to host plant defenses (Feder et al. 1998; Berlocher & Feder 2002; Drès & Mallet 2002). However, sexual selection can also contribute to population fragmentation and ultimately speciation by reducing the frequency of mating between individuals with variant sexual traits (Panhuis et al. 2001), thus arguably functioning as a force independent of natural selection (Safran et al. 2013). Sexual selection is capable of generating linkage disequilibrium (LD) between independently assorting loci due to assortative mating between mates with compatible attraction and response phenotypes (Kirkpatrick & Ravigne 2002; Kirkpatrick 2017). Divergent mating preferences for sexual traits are often one of the most important reproductive isolating
barriers in sympatric populations (Svensson et al. 2006; Dopman et al. 2010; Grace & Shaw 2011; Lackey & Boughman 2017), although whether divergent preferences arise through sexual selection alone is still a matter of debate (Servedio & Boughman). Furthermore, sexual and natural selection might interact in various ways during speciation, and opportunities to study both in unison may provide unique insights into the mechanisms by which reproductive isolation is established and maintained (Maan & Seehausen 2011; Butlin et al. 2012).

Populations of the European corn borer, *Ostrinia nubilalis* (Hübner) (Lepidoptera: Crambidae), vary in ecological and sexual traits, and this species has emerged as a model for studying mechanisms of ecological diversification and the formation of incipient species (Dopman 2011; Lassance 2016; Kozak et al. 2017; Wang et al. 2017; Coates et al. 2018). *Ostrinia nubilalis* is a polyphagous lepidopteran native to Europe and western Asia. Multiple inadvertent introductions into North America on imported broomcorn from Europe occurred near Boston, Massachusetts, Amsterdam, New York, and St. Thomas, Ontario between 1909 and 1914 (Vinal 1917; Smith 1920; McLaine 1922; Caffrey & Worthley 1927). Populations subsequently expanded south and west to the Rocky Mountains (Chiang 1972; Showers 1993). Current *O. nubilalis* populations in the northeastern United States vary in the number of reproductive generations per year (voltinism) as well as sex pheromone communication systems (Roelofs et al. 1985; Glover et al. 1991; Sorenson et al. 1992). Specifically, E- and Z-strain females produce blends of 99:1 or 3:97 (E)-11- and (Z)-11-tetradecenyl acetate (11-14:OAc), respectively, (Klun & Brindley 1970; Kochansky et al. 1975; Roelofs et al. 1972), to which corresponding males of the two strains preferentially respond (Glover et al. 1987). Although pheromone blend specificity contributes to partial reproductive isolation (Liebherr & Roelofs 1975; Pélozuelo et al. 2007; Dopman et al. 2010), hybrid females producing a 65:35 blend of E11- to Z11-14:OAc are found in nature (Klun & Maini 1979; Roelofs et al. 1985; DuRant et al. 1995).

Genetic loci contributing to *O. nubilalis* sexual isolation are inherited independently (Löfstdt et al. 1989), with female pheromone blend and male response respectively controlled...
by autosomal Pher and sex-linked Resp loci (Glover et al. 1990; Dopman et al. 2004). Pher corresponds to the pheromone gland fatty acyl reductase gene (pgfar; Lassance et al. 2010), and fixed amino acid changes between pgfar-e and pgfar-z alleles cause differences in the biosynthesis of isomeric components of female pheromone blends (Lassance et al. 2013). Pheromone strains and hybrids can be distinguished by a single nucleotide polymorphism (SNP) fixed between pgfar-e and pgfar-z alleles. Estimates of hybridization and positive assortative mating based on this autosomal marker vary among natural populations (Coates et al. 2013a). Resp is located within a region of the Z-chromosome that encodes a suite of genes putatively involved in neuronal development and function (Koutroumpa et al. 2016). In addition to Resp, a major QTL for voltinism differences between strains is located on the Z-chromosome (Dopman et al. 2004). Several loci on the sex chromosome (Tpi, OR1 and OR3) show strong differentiation between E- and Z-strain individuals while other sex chromosomal and autosomal loci do not (Dopman et al. 2005; Dopman 2011; Lassance et al. 2011). Thus, the contribution of divergence in sexual traits (sexual selection), to genome-wide substructure among sympatric O. nubilalis populations remains relatively unknown.

Host-plant interactions have long been recognized as factors that impact the diversity within insect populations (Ehrlich & Raven 1964). Although O. nubilalis larvae can feed on over 200 monocot and dicot host plants (Hodgson 1928; Lewis et al. 1975), evidence supports some degree of local host plant specialization on cultivated maize and other plant species (O'Rourke et al. 2010). Laboratory choice tests suggest that neonates may prefer alternate weedy hosts over maize (Tate et al. 2006), but survivorship is highest on maize (Calcagno et al. 2007; Fisher et al. 2017). The relative proportion of carbon stable isotope (δ13C) integrated into developing insects is directly correlated with food source (DeNiro & Epstein 1978; Boutton et al. 1999), and indicates history of C3- or C4 photosynthetic-pathways for host plants ingested by O. nubilalis as larvae (Ponsard et al. 2004; Malausa et al. 2005). Analysis of δ13C demonstrated that both O. nubilalis strains fed mostly on C4 plants (predominantly maize), but a higher proportion of males from traps baited with the E-strain pheromone lure fed on non-maize hosts as larvae in...
North America (O’Rourke et al. 2010). Nevertheless, host plant attraction for *O. nubilalis* oviposition can be influenced by plant phenology (Savinelli et al. 1988; Spanger & Calvin 2000) and volatiles (Udayagiri & Mason 1995), and may contribute to variation in host plant associations among *Ostrinia* species within Europe (Bontemps et al. 2004; Bethenod et al. 2005; Malusa et al. 2008). In France, *Ostrinia* populations using the alternative pheromone blends of E11- and Z11-14:OAc were initially assumed to be E- and Z-strain *O. nubilalis*. Strong associations were demonstrated among moths using the Z-pheromone communication system for maize (C4) and the E-system moths for C3 host plants (Bontemps et al. 2004; Bourguet et al. 2014, which were accompanied by significant levels of genetic differentiation (Bethenod et al. 2005; Malusa et al. 2007). However, subsequent studies concluded that *Ostrinia* using the E-pheromone system in France were *O. scapulalis*, and a different species compared to sympatric Z-strain *O. nubilalis* (Frolov et al. 2007; Malusa et al. 2008; Calcagno et al. 2010). Thus, results of these European studies are not directly transferrable to E- and Z-strain *O. nubilalis* in North America which show significant, but less pronounced and spatially-dependent associations with C3 and C4 host plants, respectively (O’Rourke et al. 2010). The degree to which host plant preference is associated with any genetic differentiation within or between *O. nubilalis* pheromone strains in North America remains unknown.

Genetic drift may also contribute to differentiation among *O. nubilalis* populations. Based on anonymous microsatellite marker data, *O. nubilalis* from two locations in New York and Pennsylvania were genetically differentiated from sites in Ohio westward to Colorado (Kim et al. 2011). Although the authors suggested landscape features within northeastern states, including the Appalachian Mountains and fragmented crop land might serve as a partial barrier to gene flow, the possible influence of pheromone strain differences among geographic regions (Liebherr & Roelofs 1975; Glover et al. 1991) could not be ruled out. A separate study predicted a significant isolation-by-distance relationship when comparing only Z-strain populations along a north-south transect in the Midwest United States, although this pattern was driven by a small percentage of the sampled markers (15 of 194: Levy et al. 2015).

This article is protected by copyright. All rights reserved.
Of the handful of studies that address the relative roles of ecological adaptation, sexual selection, and genetic drift as drivers of genomic divergence, many find that sexual selection is potent only when coupled with ecological factors (van Doorn et al. 2009; Mann & Seehausen 2011; Wagner et al. 2012; Servedio & Boughman 2017), but in most systems limited details are available regarding ecologically and sexually relevant traits (Safran et al. 2013). By dissecting the short-term contributions of ecology, genetic drift, and sexual traits on patterns of genomic variation and gene flow among distinct subpopulations, subspecies, and incipient species, a broader understanding of their contributions to speciation across time scales may be gained. Using anonymous genomic SNP markers, this study estimates the influence of pheromone strain, host plant differences, and geographic distance on population genetic structure of O. nubilalis within the northeastern United States.

2 | Materials and Methods

2.1 | Collection and phenotypic differentiation

Pheromone lures with a blend of 97:3 or 1:99 synthetic E11- and Z11-14:OAc (Trécé Incorporated, Adair, OK) were used to separately bait one of a pair of Harstack wire cone traps at Pennsylvania and cloth Scentry traps at New York collection sites between 2007 and 2012 (12 sites total; Table 1; Fig. 1; Table S1). Males were captured during the second flight of the bivoltine strain (i.e. adult offspring of the overwintering generation) and frozen at -20°C. Samples were shipped on dry ice to the USDA-ARS, Corn Insects & Crop Genetics Research Unit, Ames, Iowa, where thorax tissue and all 4 wings were dissected from each moth. Wings were placed in individual 0.2 ml wells of a 96-well PCR plate lined with aluminum foil and stored at -20°C until used for carbon isotope analysis. Total genomic DNA was purified from each thorax using the Qiagen DNeasy Blood and Tissue Extraction kit according to manufacturer instructions (Qiagen, Hilden, Germany). DNA concentration in resulting extracts was quantified.

This article is protected by copyright. All rights reserved.
on a NanoDrop2000 spectrophotometer (Thermo Scientific, Wilmington, DE), diluted to 10 ng μl⁻¹ with deionized nuclease free water, and stored at -20°C.

Host plant history, C3 (non-maize) or C4 (probable maize), was assigned to each male collected in pheromone traps based on carbon stable isotope analysis. For this, the four wings were dissected from each trap collected male, and the ratio of 13C to 14C (δ¹³C) was estimated by continuous flow-isotopic ratio mass spectrometry (CF-IRMS) conducted at the University of California Berkley, Center for Stable Isotope Biogeochemistry (Berkley, CA). The results were converted to δ units relative to Pee Dee Belemnite carbonate (δ¹³C = (R_sample/R_standard – 1) × 1000; R = the ratio of ¹³C/¹²C atomic percentage).

In addition, genetic assignment of pheromone strain was determined for each male based on SNP genotype of a marker at the pgfar locus. This assay detects alleles that are >98.7% correlated with E- or Z-pheromone strains (Coates et al. 2013a). Specifically, the TaqI PCR-RFLP assay was used to assign homozygous pgfar-e/pgfar-e and pgfar-z/pgfar-z genotypes to respective strains, and pgfar-e/pgfar-z heterozygotes to a hybrid class. Additionally, capture of males in pheromone traps baited with the E- or Z-strain lure was used as a proxy for pheromone response phenotype, given that males of the two strains strongly differ in their flight towards E and Z pheromone blends (Roelofs et al. 1987).

2.2 Population genetic differentiation

A set of 108 anonymous biallelic SNP markers was used to genotype each male by single base extension assay products analyzed on a Sequenom MassARRAY® located at the Iowa State University, Center for Plant Genomics, Ames, IA. In brief, loci containing a priori defined SNPs were amplified in four PCR multiplexes (W1, W2, W3 and W4; Coates et al. 2011), and subsequently used as template for SBE assays that incorporated allele-specific mass-modified dideoxynucleotides using the iPLEX-Gold Mastermix (Sequenom; Tang et al. 1999). Genotypes
were determined from raw Sequenom MassARRAY® output as described by Coates et al. (2013b).

Processed Sequenom SNP and pgfar PCR-RFLP genotype data were assessed for Hardy-Weinberg equilibrium (HWE), with departures of observed heterozygosity ($H_o$) from expected heterozygosity ($H_e$) across collection sites tested for significance ($\alpha = 0.05$) using Markov chain exact tests implemented in the Arlequin software package (v. 3.5.2; Excoffier et al. 2005). HWE tests were performed for E-, Z- and hybrid pgfar PCR-RFLP-defined partitions within location. Loci were subsequently excluded from further analyses if monomorphic, genotypic data were missing for $\geq 25\%$ of individuals, or genotypic proportions showed a significant departure from HWE expectation in $\geq 20\%$ of populations. Exceptions were made in instances where no HWE departures were predicted across all E- or Z-strain populations. For loci that passed the preceding filters, Fisher’s exact tests were conducted to predict linkage disequilibrium (LD) between all pairs of unphased loci, and significance determined by 10,000 Markov Chain permutation steps using Arlequin 3.2.2 (Excoffier et al. 2005). Any pair of SNP loci in LD were compared to known map positions on a consensus O. nubilalis genetic linkage map (Levy et al. 2015). Linkage groups (LGs) were assigned to physical location on Bombyx mori chromosomes by querying the B. mori genome assembly database with O. nubilalis contig sequences from which SNP markers were developed (Coates et al. 2011) using the BLASTx algorithm at KAIKOBASE (Shimomura et al. 2009). Additionally, the B. mori fatty acyl-reductase involved in bombykol production (Moto et al. 2003; GenBank accession BAC79425) is the nearest ortholog to O. nubilalis pgfar (Lassance et al. 2010), and was used predict the relative orthologous position in O. nubilalis. All BLAST database results were filtered for $E$-values $\leq 10^{-50}$.

Locus-by-locus $F_{ST}$, $F_{IS}$, and $F_{IT}$ estimates were made based on homozygous pgfar genotypes of males partitioned by having been collected in traps baited with E11- or Z11-14:OAc pheromone lure (male pheromone response) or by pgfar genotype. Global $F$-statistics and analysis of molecular variance (AMOVA) were estimated for genotypes partitioned by
pheromone response and *pgfar* genotype, as well as by host plant history (C3 or C4), across the 12 collection sites. Pairwise $F_{ST}$ estimates were made between collection sites, and within collection site between pheromone response, *pgfar* genotype, and host plant history. Significance threshold (false discovery rate) for each statistic was adjusted for multiple testing using the BH procedure (Benjamini & Hochberg 1995). All estimates of subpopulation differentiation were assessed using Arlequin v. 3.5.2 (Excoffier et al. 2005). A filtered SNP dataset of 64 loci (omitting *pgfar*) was used in all instances, except for comparisons of pheromone response in which the corresponding SNP dataset of 65 loci (including *pgfar*) was also assessed.

Principal coordinate analysis (PCoA) was conducted using the covariance matrix of pairwise $F_{ST}$s using GENALEX v. 6.503 (Peakall & Smouse 2006). Pairwise $F_{ST}$ estimates were made for 52 groups ($n = 1,421$ genotypes) using the 65 SNP loci that included the *pgfar* locus, and again with 64 loci after excluding the *pgfar* locus. The 52 groups represent the three *pgfar* genotypes in the 12 collection sites for each year; note that there were no homozygous *pgfar-e/pgfar-e* individuals in two of the collections (Table 1), so those combinations were not represented in the analyses. Pairwise $F_{ST}$ estimates were also made for the same 52 *pgfar* delimited groups with collection sites further partitioned by the two host plant history categories of C3 and C4 ($n = 1,411$ individuals), creating 91 total groups; note that there were no individuals with C3 host plant history in 13 of the *pgfar* plus collection site groups, so those combinations were not represented in the analyses. Factor scores along the two principal coordinates that accounted for the most variation were used to construct a scatter diagram to visualize genetic relationships among populations.

2.3 | Bayesian clustering

Population structure was inferred using a model-based clustering method implemented by STRUCTURE 2.3.4 (Prichard et al. 2000; Falush et al. 2003; Hubisz et al. 2009).
provides likelihood estimates of the number of distinct populations \((K)\) from unlinked genetic marker data using a Markov Chain Monte Carlo (MCMC) method. Estimates were obtained by iteratively running STRUCTURE 2.3.4 with \(K\) values from 1 to 10. Each run consisted of 10 replications with \(1 \times 10^7\) iterations preceded by a burn-in of \(10^5\) iterations. Runs used an admixture model of individual ancestry and a uniform prior for ALPHA. The LOCPRIOR option (Hubisz et al., 2009) was used on genotypes partitioned into 12 collection sites (1,421 genotypes, 12 partitions by location), pheromone response combined across the 12 sample sites (1,421 genotypes, 2 partitions by collection in traps baited with E11- or Z11-14:OAc pheromone lure), and C3 and C4 host plant history (\(\delta^{13}C\) estimates for C4 host = -8 to -17 and C3 host = -22 to -32). All runs were conducted using the set of 64 SNP genotypes (pgfar omitted), then repeated using the set of 65 SNP genotypes (pgfar included). For all STRUCTURE results, the median value of the estimated log probability for each clustering analysis, conditional on \(K\), \(\ln Pr(X|K)\), was used to compute the posterior probability of \(K\), \(Pr(K|X)\), assuming a uniform prior distribution for \(K\). Optimized values for \(K\) were derived from the resulting population co-ancestries (\(Q\)-matrix) using the \(\Delta K\) statistic (Evanno et al. 2005). Individual \(Q\)-matrix data were visualized using Distruct (Rosenberg 2004).

### 2.4 | Ecological and behavioral relationships

Possible isolation-by-distance (IBD) (Slatkin 1993) was assessed by regressing pairwise genetic distance (based on estimates of \(F_{ST}\) among 12 sample locations) on geographic (km) distance. IBD analyses were performed using GENALEX (Peakall & Smouse, 2006) with the associated probabilities determined using Mantel’s Test, and significance determined based on 999 permutations. The significance of any differences in the distribution of \(\delta^{13}C\) estimates between pgfar-defined E-, Z- and hybrid (3 partitions) or pheromone response (2 partitions) grouped by sample site were made via single-factor ANOVA using SAS 9.2 (SAS Institute, Cary, NC, USA). Additionally, a distance-based redundancy analysis (dbRDA; Legendre & Anderson 2007).
1999; Oksanen et al. 2017) was used to test the relative roles of host plant use (C3:C4), sexual traits (pgfar or pheromone response), and geographic distance on population genetic structure. Redundancy analysis allows for multiple predictor variables to be analyzed simultaneously and each parsed into distinct factors unlike the partial Mantel Test (Legendre & Legendre 1999; Legendre & Fortin 2010). At each sampling site, males were classified as E- or Z-strain in two different ways: 1) by pheromone response (collected in traps baited with E11- or Z11-14:OAc pheromone lure), and 2) by male pgfar genotype (pgfar-e/-e as E-strain individuals and pgfar-z/-z as Z-strain individuals; all pgfar-e/-z were excluded). Since very few E-strain individuals were collected at Northampton, PA (Nor), and due to proximity (~5 km) of this site to Lehigh, PA (Leh), Nor was removed from analysis, resulting in 22 total populations (11 E- and 11 Z-strain). For each population, the mean δ13C value was calculated, and the matrix of geographic distances (km) was transformed into component eigenvectors for each location using the PCNM function in the VEGAN R package (http://cc.oulu.fi/~jarioksa/softhelp/vegan/). Subsequent dbRDA was performed using the CAPSCALE function in the VEGAN R package (http://cc.oulu.fi/~jarioksa/softhelp/vegan/html/capscale.html). The constrained matrix was composed of multilocus FST estimates [from the dataset of 64 SNPS (pgfar excluded)], and the predictor variables were set to pheromone (a binary trait, 0 = same pheromone, 1 = different pheromone), mean δ13C, and the first three geographic distance eigenvectors. Significance of each predictor (while accounting for other factors) was determined by permutation (N = 1000). The dbRDA was replicated using latitude and longitude as factors instead of the eigenvectors to verify that the results were not influenced by the geographic distance transformation. Partial Mantel Tests were also performed in VEGAN using distance matrices and conditioning pairs of factors (geographic distance, pheromone difference, and mean δ13C difference) sequentially on each other. P-values were corrected for multiple testing using the Bonferroni method. In addition, an individual-based analysis was performed in which genetic distance was calculated as the number of allelic differences between two individuals (allelic dissimilarity) in the PPOPR package (Kamvar et al. 2014). The predictor variables in the dbRDA were pheromone (pgfar or
pheromone response), $\delta^{13}$C value, and the first four distance eigenvectors (due to the larger number of eigenvectors in the individual dataset).

3 | Results

3.1 | Collection and phenotypic differentiation

A total of 1,421 adult male *O. nubilalis* were collected at 12 trap locations in New York and Pennsylvania from 2007 to 2010 (Fig. 1), with 704 and 717 collected from E- and Z-pheromone baited traps, respectively (Table 1). Results of the *pgfar*-specific PCR-RFLP genotyping assay yielded three distinct genotypes which were assigned to E-strain (*pgfar*-/*pgfar*-; $n = 355$), Z-strain (*pgfar*-/*pgfar*-; $n = 711$), or a heterozygote class (*pgfar*/*pgfar*; $n = 355$; Table 1; doi:10.5061/dryad.kj34mq9). Assessment of temporal (year-to-year) change in heterozygosity at *pgfar* across locations in 2009 vs 2010 indicated a significant difference ($P = 0.043$; Fig. S1A). The $\delta^{13}$C values calculated for 1,415 males from which isotope readings were collected showed a bimodal distribution, with mode peaks at -10 and -28 (Fig. S1B). The estimates of $\delta^{13}$C between -8 to -17 and -22 to -32 were respectively assigned to the C4 ($n = 1,160; 82.0\%$ of total) and C3 ($n = 255; 18.0\%$) groups. The host plant history of five individuals with $\delta^{13}$C estimates ranging between -18 to -21 were considered ambiguous and were excluded from further analyses involving host plant. For this dataset, the proportion of males derived from C3 host plants did not differ significantly between 2009 and 2010 ($P = 0.109$; Fig. S1C). These male pheromone response, *pgfar*-defined pheromone strain, and host plant history partitions of the data were used in subsequent analyses.
3.2 | Population genetic differentiation

Of the 108 SNP markers, 65 (including pgfar) passed stringency criteria and were retained within the final dataset (Dryad database accession doi:10.5061/dryad.kj34mq9). Of those markers that were excluded, 31 were not in HWE. Subsequent locus-by-locus $F_{ST}$ estimates indicated five SNP loci, pgfar, contig00078.322, contig05852.999, contig06634.250, and contig06864.843, differed significantly between groups of males defined by response to E- or Z-strain pheromone lure (Table S2A). Similarly, markers contig00078.322, contig01257.497, contig01453.160, contig05858.945, contig06634.250, and contig06864.843 showed locus-by-locus differentiation between pheromone strains defined by pgfar (Table S2B; significance of pgfar ignored). The pgfar locus accounted for 64% of the total genetic variation between samples trapped with different pheromone lures, and the SNP marker locus contig01257.497 accounted for ≥ 10% of the total genetic variance in both locus-by-locus analyses. Tests of independence among the 65 filtered loci detected LD only between pgfar and contig01257.497 ($P < 0.0001$; results of other locus pairs not shown). Secondary confirmation of linkage came from the mapping of ortholog positions. Specifically, the predicted genome position of the nearest B. mori ortholog of O. nubilalis pgfar, BAC78425 (Lassance et al. 2010), is synonymous with the B. mori gene model BGIGMGA010458 (E-values ≤ 8 x 10^{-39}; 84.5 to 100% identity across exons) located on B. mori chromosome 12 (positions 2,426,074 to 2,437,943). The nearest ortholog to O. nubilalis SNP marker locus contig01257.497 is a ubiquitin conjugating enzyme E2 (E-values ≤ 10^{-56}; 80.0 to 100% identity across exons) on B. mori chromosome 12 positions 5,150,695 to 5,153,581 or ~ 2.7 Mbp away from the pgfar ortholog. The SNP marker contig01257.497 was previously positioned on LG11 of an O. nubilalis genetic map, and also assigned to an orthologous position on B. mori chromosome 12 (Levy et al. 2015). Contig06721.1113 was the only SNP marker on the Z-chromosome and $F_{ST}$ estimates did not differ significantly between strains defined either by pheromone response (Table S2A) or pgfar (Table S2B).
Based on the entire dataset of genotypes from 65 SNP loci (inclusive of pgfar), $F_{ST}$ estimates between the 12 collection sites were significant for 13 of the 66 (19.7%) pairwise comparisons (Table S3A). This same set of genotypes was partitioned across sample sites based on response to either E- or Z-strain pheromone lures ($n = 24$ groups). Resulting $F_{ST}$ estimates ranged from -0.09583 to 0.05852 and 144 of 276 comparisons (52.2%) were significant, of which 125 (86.8%) were significant between groups responded to different lures (Table S3B). Eleven of the 12 (92%) pairwise $F_{ST}$ comparisons were significant between sympatric (within location) groups collected in E- or Z-pheromone-baited traps. The global estimate of differentiation between E- and Z-strain males across locations was also significant ($F_{ST} = 0.0322$, $P < 0.0001$), and AMOVA implicated pheromone response as accounting for 3.22% of the total variation between subpopulations (Table S4A). Conversely, omission of the pgfar locus from comparisons of pheromone response resulted in a lower number of significant pairwise $F_{ST}$ estimates (15.2%; Table S3C), and a reduced global estimate of subpopulation differentiation ($F_{ST} = 0.0047$, $P < 0.0001$; Table S4B). Interestingly, significant genotypic differences between sympatric E- and Z-strains were only detected at northern collection sites (Oak, Int, Ran, Plan and Roc; Fig. 1) where pairwise $F_{ST}$ estimates were ≤ 0.0087; Table S3C). Five of 12 (42%) pairwise $F_{ST}$ comparisons were significant between sympatric (within location) E- and Z-pheromone response groups, and were located at a significantly higher latitude (mean 41.79°N) than those that were non-significant (mean latitude 40.54°N; $T_{4.78}=2.81$, $P = 0.039$).

Pairwise $F_{ST}$ estimates from 64 SNP loci (excluding pgfar) between E- and Z-strains identified by homozygous genotypes at the pgfar marker locus across 21 trap locations where the strains co-occurred (Table 1) were significant in 210 of 276 comparisons (76.1%; Table S3D). Among these, $F_{ST}$ estimates between E- and Z-strains within and across sites were significant in 113 of 132 instances (85.6%; 10 of 11 sympatric E-Z comparisons). Global differentiation across locations between E- and Z-strain males based on pgfar genotype was low but significant ($F_{ST} = 0.0085$, $P < 0.0001$), and AMOVA implicated pgfar-defined pheromone strain as accounting for 6.78% of the total variation (Table S4C).

This article is protected by copyright. All rights reserved.
Fifty-five of 276 (19.9%) pairwise $F_{ST}$ estimates were significant between partitions of the SNP genotype data defined by both collection site and host plant history (C3 or C4) (Table S3E). Twenty-six of these 55 (47.3%) significant comparisons were between genotypes assigned to different host plant history groups, including within three sites [Interlaken (Int), Ransom (Ran), and Rockspring (Roc)]. Overall global estimates of $F_{ST}$ between C3 and C4 host plant history groups across all collection sites was low and non-significant ($F_{ST} = 0.0004; P = 0.1779$; Table S4D).

There was no significant correlation of pairwise $F_{ST}$ estimates with geographic distance as tested by the IBD model ($P = 0.4280$; Fig. 2). The results of principal coordinate analyses (PCoA) based on genotype data from all 65 SNP loci (including pgfar) indicated that factor scores along Coordinate 1, accounted for 25.1% of the total variation, and grouped E-, Z- and heterozygote individuals into unique clusters along the Coordinate 1 axis (Fig. 3A). This contrasted with results of PCoA based on SNP data from the 64 loci excluding the pgfar locus, where factor scores for E-, Z- and heterozygotes groups were not distinct (Fig. 3B). PCoA revealed no discernable pattern in clustering of groups that accounted for pgfar genotype (pheromone strain) and location when assigned to C3 or C4 host plant history groups (Fig. 3C).

### 3.3 | Bayesian clustering

An iterative Bayesian simulation approach estimated the optimal number of clusters ($K$), or subpopulations, represented by multilocus *O. nubilalis* SNP genotypes. STRUCTURE 2.3.4 predicted an optimal number of $K = 3$ both when pgfar was either included (65 loci; $\Delta K = 47.7$) and excluded (64 loci; $\Delta K = 36.3$) from the genotypic dataset partitioned by pheromone response as the prior. Co-ancestries (clusters) were relatively homogenous between groups based on E- and Z-strain response using the dataset with 64 loci (Fig. S2A). In contrast, inclusion of pgfar (65 loci) resulted in a mean proportion of Cluster 1 among E-pheromone response males of 0.7786 compared to 0.0451 among Z-pheromone response males (Fig. S2A). This article is protected by copyright. All rights reserved.
Assessment of these differences in assignment of co-ancestries revealed that the proportion of homozygous \textit{pgfar-\text{-}z/-z} males that responded to the Z-strain lure (mean = 0.8972 ± 0.0512; Table S5) was significantly greater than the proportion of homozygous \textit{pgfar-\text{-}e/-e} genotypes that responded to the E-strain lure (0.4305±0.2249; one-way ANOVA: \( F = 49.12, P < 0.0001, \) d.f. \( = 1, 22 \)).

Partitioning data using the 12 collection sites as \textit{a priori} information under the \textsc{locprior} option yielded similar results. Specifically, for 65 loci (including the \textit{pgfar} locus), \textsc{structure} 2.3.4 estimated \( K = 3 (\Delta K = 28.5) \) with clusters 1 and 2 corresponding to E- (E-pheromone response-\textit{pgfar-\text{-}e/-e}) and Z-strain (Z-pheromone response- \textit{pgfar-\text{-}z/-z}), respectively (Fig. S3). Use of 64 loci (\textit{pgfar} omitted) resulted in \( K = 2 (\Delta K = 29.4) \), and no structure among pheromone response groups. Approximately equal mean proportions individual co-ancestries (population \( Q \)-score) of 0.850 ± 0.240 and 0.830 ± 0.241 were assigned to Cluster 1 for \textit{pgfar-\text{-}e/e} and \textit{pgfar-\text{-}z/-z} genotypes, respectively (remaining data not shown).

Use of host plant history (C3 or C4 designation) based on \( \delta^{13} \text{C} \) estimates as \textit{a priori} information for Bayesian clustering of \textit{O. nubilalis} genotypes across 64 SNP loci (excluding \textit{pgfar}) resulted in an optimum \( K \) of 3 (\( \Delta K = 42.0 \)). There was no difference in the mean proportion of co-ancestries assigned to Cluster 1 between C3 and C4 plants for comparisons either with E- (C3: 0.892 ± 0.148; C4: 0.873 ± 0.147) or Z-strain (C3: 0.836 ± 0.218; C4: 0.837 ± 0.226) defined by \textit{pgfar} genotype (Fig. S2B). Bayesian clustering of genotypes from 65 SNP loci (including \textit{pgfar}) resulted in an optimum \( K = 3 (\Delta K = 569.6) \) with shared co-ancestries differentiated into clusters corresponding to \textit{pgfar-\text{-}e/-e}, \textit{pgfar-\text{-}e/-z} and \textit{pgfar-\text{-}z/-z} despite \( \delta^{13} \text{C} \) estimates (C3 or C4 host plant history) having been set as the \textit{prior} (Fig. S2B).
3.4 | Ecological and behavioral relationships

The overall proportions of individuals with δ¹³C values in the range of -24 to -32 (C3 host plant history group) among pgfar-e/pgfar-e (E-strain), pgfar-z/pgfar-z (Z-strain), and pgfar-e/pgfar-z (E-/Z-strain hybrids) were 0.161, 0.203, and 0.151, respectively. Single factor ANOVA showed no significant difference in the proportion of homozygous male pgfar genotypes assigned to the δ¹³C bin (-8 to -32) across populations (F = 3.42 × 10⁻⁵, P = 1.0000, d.f. = 2, 22), between groups defined as pgfar homozygotes across locations (F = 1.75 × 10⁻⁵, P = 0.9967, d.f. = 1, 23), or strain defined by pheromone response across locations (F = 2.59 × 10⁻⁵, P = 0.9981, d.f. = 2, 22).

Analyses of dissimilarity matrices using dbRDA showed that Fⱽ estimates (genetic differentiation) among populations differed significantly only by pheromone strain. E- and Z-strain pairwise comparisons had significantly higher Fⱽ estimates when classified by pheromone response or by pgfar genotype (Fig. 4A, B; Table 2), but not by host plant (Fig. 4C). These dbRDA analyses also indicated Fⱽ estimates were not significant for host plant history (C3:C4; P ≥ 0.268) or geographic distance (P ≥ 0.369). Similarly, latitude and longitude (used as distance factors) or using partial Mantel tests indicated that pheromone strain (response and pgfar genotype) significantly contributed to genetic variance, but host plant history did not (Table S6A, 6B; Table S7). Lastly, males responding to different pheromone lures, from different locations, or with different C3:C4 ratios showed significant levels of mean allelic differentiation; however, the effect of host plant history (C3:C4; η² = 0.0009) was weaker than those of distance (η² = 0.0017) or pheromone response (η² = 0.0034; Table S8).

4 | Discussion

In many diverging populations of Lepidoptera, ecological factors such as host plant use co-vary with differences in sexual signals (Groot et al. 2008; Joyce et al. 2016), but in few cases has the
relative contribution of these factors to genetic structure been tested (Prowell et al. 2004; Gouin et al. 2017). Ostrinia nubilalis is an emerging model for speciation dynamics due to sympatric ecotypes or strains that vary in voltinism and sex-pheromone communication (Dopman et al. 2010; Olsson et al. 2010; Lassance 2016; Kozak et al. 2017; Coates et al. 2018). Ecological differences in host plant preference occur between closely related O. nubilalis strains in North America (O’Rourke et al. 2010) and among Ostrinia species in Europe (Bourguet et al. 2000a, 2014; Thomas et al. 2003; Pélozuelo et al. 2004; Betenod et al. 2005; Pelozuelo et al. 2007; Frolov et al. 2012). Although pheromone system and host plant use co-vary in Ostrinia (Alexandre et al. 2013), elucidating the relative contribution of various selective forces to differentiation in the presence of high rates of gene flow remains a challenge.

Our current study suggests that gene flow among northeastern populations of O. nubilalis remains relatively high at neutral genetic loci, as inferred by mostly non-significant pairwise $F_{ST}$ estimates between collection sites (Table S3A) and by IBD modeling (Fig. 2). Similarly, our dbRDA analyses suggest that geographic distance did not contribute significantly to variation among collection sites (Table 2; Table S6, S7). This conclusion is supported by earlier population genetics studies where low pairwise $F_{ST}$ estimates and shallow IBD slopes suggested high gene flow over long distances (Bourguet et al. 2000b; Kim et al. 2009, 2011; Frolov et al. 2012). Additionally, in the current study we detected no significant impact of host plant history on O. nubilalis population genetic structure (Table S4D) or in PCoA (Fig. 3C). Likewise, dbRDA-based eigenvectors for variance in C3:C4 ratios among collection sites (Table 2; Table S6) and partial Mantel Tests (Table S7) were not significant. Finally, the proportion of individual co-ancestries assigned to subpopulations in Bayesian clustering analysis were strongly influenced by variation at the pgfar locus, not host plant despite the latter beings used as the prior (Fig. S2B). Together, these results indicate that host plant history is not a strong ecological factor impacting O. nubilalis population structure in North America. The scenario in Europe appears to be markedly different between the sympatric sibling species, O. nubilalis and O. scapulalis. Specifically, Ostrinia from France that use the E-pheromone communication
system are *O. scapulalis* which predominantly feed and develop on mugwort (Malasa 

et al. 2007; Frolov et al. 2007), whereas *O. nubilalis* are Z-strain and mostly feed on maize. Thus, conclusions from the current study regarding the basis of ecological diversification between E- and Z-strain *O. nubilalis* may not be transferrable to the sibling species in Europe.

Our study suggests that pheromone strain has a greater effect on *O. nubilalis* genetic differentiation than either geography or host plant use. Specifically, pheromone strain of captured males (*pgfar* homozygous genotype) or pheromone response phenotype (*Resp*) each represent significant factors accounting for *O. nubilalis* population genetic variance. This conclusion is most clearly drawn from dbRDA, wherein pheromone response phenotype and *pgfar* genotype are the most significant eigenvectors determining degree of genetic variance when compared to host plant history- and geography-based eigenvectors (Table 2; Fig. 4; Table S6, S8). Even though genetic loci determining female pheromone signal and corresponding male response are not physically linked in all species of Lepidoptera studied (Groot et al. 2015), including *O. nubilalis* (Roelofs et al. 1987; Löfstedt et al. 1989; Dopman et al. 2004), these loci remain in LD (Kirkpatrick & Ravigne 2002; Kirkpatrick 2017). Several models have been proposed to explain the co-evolution of genetic loci contributing to mate attraction and selection (Groot et al. 2016), but lack of information on the genomic architecture of these traits hinders current research (Coates et al. 2018). Although we used capture of male *O. nubilalis* in traps baited with strain-specific pheromone lures as a proxy for pheromone attraction phenotype (*Resp*), a direct measure of allelic variation at the *Resp* locus is lacking. The Z-chromosome-linked *Resp* locus (Dopman et al. 2004) is located within a suite of genes putatively involved in neurogenesis (Koutroumpana et al. 2016). The specific gene or mutation(s) involved, or genetic markers, most strongly correlated with male response are yet to be determined. Thus, variation at the causal locus, and therefore potential impacts of selection could not be tested directly and remain the subject of future work.
Our results suggest that sexual selection is acting concurrently on unlinked loci contributing to female pheromone signaling and corresponding male response phenotypes in *O. nubilalis*. These characters appear to have a more significant impact on inter-strain genetic divergence and strain maintenance than the ecological factor of host plant use. These results are consistent with the observed high rates of assortative mating caused by male response to female pheromone and selection against hybrid pheromone blends (Linn *et al*. 1997; Dopman *et al*. 2010). However, observed levels of hybridization (as evidenced by heterozygosity at *pgfar*; Fig. S1A) suggests assortative mating is incomplete and gene flow is occurring between E- and Z-strains. Hybridization likely explains why genome-wide structuring was not detected among strains when the *pgfar* locus genotype was removed from analyses. Future studies using a greater number of anonymous markers could clarify whether population differentiation in *O. nubilalis* is limited to *Pher* and *Resp*, and genome regions physically linked to these loci. It is also possible that other islands of differentiation in the genome are associated with loci promoting assortative mating (Royer & McNeil 1992; Lassance & Löfstedt 2009) including differences in circadian timing (Liebherr & Roelofs 1975; Levy *et al*. 2018).

Although we found no evidence that host-plant use contributes to genetic structure, sexual selection alone is not necessarily the singular evolutionary force maintaining *O. nubilalis* pheromone strains. Ecological variation in seasonal timing (voltinism) in conjunction with pheromone strain may also contribute to high genetic differentiation on the *O. nubilalis* Z-chromosome compared to on the autosomes (Levy *et al*. 2015; Kozak *et al*. 2017). This phenomenon may be due in part to a rearrangement (probably an inversion; Wadsworth *et al*. 2015) of a segregating region of the Z-chromosome which differs in frequency most among individuals that differ in pheromone strain and voltinism ecotype (Kozak *et al*. 2017). Voltinism is strongly correlated with latitude in Z-strain *O. nubilalis* (Showers 1981). Based on SNP markers (excluding the *pgfar* locus), we detected significant *F*<sub>ST</sub>s between E- and Z-pheromone response groups collected in the same location mostly at northern sites (Int, Oak, Pla, and Ran) where the strains are also known to differ in voltinism (bivoltine E- and univoltine Z-strain;
Glover et al. 1991; Calvin & Song 1994). These results suggest that interactions between certain ecological and sexual traits may reinforce one another to limit gene flow in *O. nubilalis*, and the subsequent development of reproductive isolation may be a multivariate process dependent on both natural and sexual selection (Maan & Seehausen 2011; Butlin et al. 2012).

In this study, we found that female pheromone blend and male response contributes to a greater proportion of the genetic differentiation among populations of *O. nubilalis* in the northeastern United States compared to geographic distance or host plant. These conclusions were drawn from congruent results from Bayesian clustering, PCoA, IBD modeling, and dbRDA. Data from this and earlier studies suggest that sexual selection may function as a component maintaining the observed level of genetic differentiation between sympatric pheromone strains in North America, but whether it is sufficient by itself to maintain the pheromone strains over evolutionary time remains unclear. The relative contributions of male pheromone response and voltinism on genomic variance between *O. nubilalis* pheromone strains may be more precisely quantifiable upon the future development of genetic markers that detect or are tightly linked to causal mutations. Regardless, these data are important for understanding the relative contributions of ecological factors (host plant use), distance (geographic isolation), and sexual selection (maternal female pheromone blend and male response) in forming barriers to gene flow between *O. nubilalis* subpopulations, which may be propelling them along different evolutionary trajectories.

**Acknowledgments**

This research was a joint contribution from the United States Department of Agriculture (USDA), Agricultural Research Service (ARS) (CRIS Project 5030-22000-018-00D), and the Iowa Agriculture and Home Economics Experiment Station, Ames, IA (Project 3543), This research was funded by the National Science Foundation (DEB-1257251; E.B.D.),

This article is protected by copyright. All rights reserved.
and cooperative agreement 58-5030-7-066 between USDA-ARS and Tufts University.

We thank Natalie Wertish and Miranda Dietz for carrying out portions of the genotyping assays used in this study. This article reports the results of research only and any mention of products or services does not constitute an endorsement by USDA-ARS.

USDA-ARS is an equal opportunity employer and provider. Graham Stone, Astrid Groot, and two anonymous reviewers provided helpful comments on the manuscript.

References


This article is protected by copyright. All rights reserved.


Coates, B. S., Sumerford, D. V., Siegfried, B. D., Hellmich, R. L., & Abel, C. A. (2013b). Unlinked genetic loci control the reduced transcription of aminopeptidase N 1 and 3 in the European corn borer and determine tolerance to *Bacillus thuringiensis* Cry1Ab toxin. *Insect Biochemistry and Molecular Biology, 43*(12), 1152-1160.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


Vinal, S. C. (1917). *The European corn borer, a recently established pest in Massachusetts.* Massachusetts Agricultural Experiment Station Bulletin I78: 147-152.


**Author Contributions**

BSC designed experiments, analyzed data, and wrote the manuscript with contributions from other authors; GMK, KSK, and YW each analyzed portions of the data; JS conducted laboratory experiments and collected data; SJF designed strategy for location of traps and collection of samples; EBD and TWS provided advice and contributed to writing portions of the manuscript.

**Data accessibility**

High throughput single nucleotide polymorphism genotype data for all samples are available at doi:10.5061/dryad.kj34mq9.
Table 1. Sample locations within and across years for 12 sites in the northeast United States where Ostrinia nubilalis E- and Z-pheromone strains are sympatric. Males were captured in pheromone traps at each location baited with either an E-lure (99:1 E:-Z11-14:OAc) or Z-lure (3:97 E:-Z11-14:OAc). Pheromone strain of the same males was determined independently based on pgfar genotype (homozygous pgfar-e/-e and pgfar-z/-z; heterozygous pgfar-e/-z). The count of males from with larval host plant as determined from wing stable carbon isotope content ratio of $^{13}$C/$^{14}$C (i.e., δ$^{13}$C) are indicated within brackets [C3 (δ$^{13}$C range of -22 to -32); C4 (δ$^{13}$C range of -8 to -17)]. ♦ = Harstack wire cone trap; ● = cloth Scentry trap.

Table 2. Distance-based redundancy analysis (dbRDA) of eigenvectors (EV) for the relative contributions to Ostrinia nubilalis population genetic structure of pheromone strain [assigned based on A) male response (attraction) to either E-strain lure (99:1 E:-Z11-14:OAc) or Z-strain lure (3:97 E:-Z11-14:OAc), or B) pgfar genotype], host plant history (C3:C4), and geographic distance (km) (pgfar genotype excluded). The sum of squares (Sum of Sq.), F-statistic based on 1 and 16 degrees of freedom ($F_{1,16}$), proportion of the total variance attributed to each component ($\eta^2$), and associated P-values are shown.

Figure 1 Topographic map of geographic locations for 12 collection sites (location abbreviations in Table 1 and coordinates in Table S1). ♦ = Harstack wire cone trap; ● = cloth Scentry trap.

Figure 2 Influence of geographic distance on genetic differentiation between Ostrinia nubilalis subpopulations (collection sites; Table 1). Tests for an isolation-by-distance model between pairwise genetic distance ($F_{ST}$) and geographic distance. Associated $R^2$ (goodness of fit) and $P$-values estimated from 1,000 iterations performed using Mantel Tests.

Figure 3 Scatter plot of factor scores along the major Principal Coordinates (PCs) accounting for the greatest variation among Ostrinia nubilalis population-derived single nucleotide polymorphism (SNP) genotype data partitioned into A) pheromone strain using 65 SNP markers including the pgfar locus, B) pheromone strain using 64 SNP markers excluding the pgfar locus, and C) host plant history using 64 SNP markers excluding the pgfar locus. ♦ = Harstack wire cone trap; ● = cloth Scentry trap.

Figure 4 Influence of ecological factors on the genetic differentiation between Ostrinia nubilalis subpopulations (collection sites; Table 1) based on dbRDA analyses (64 loci, excluding pgfar genotypes). Range of $F_{ST}$ estimates between E- and Z-strain when partitioned by pheromone lure (trap), homozygous genotypes from the pgfar genetic marker, or C3:C4 ratio (host plant). Significance as determined by dbRDA (Table 2) are shown for each comparison.

This article is protected by copyright. All rights reserved.
Table 1. Sample locations within and across years for 12 sites in the northeast United States where *Ostrinia nubilalis* E- and Z-pheromone strains are sympatric. Males were captured in pheromone traps at each location baited with either an E-lure (99:1 E-:Z11-14:OAc) or Z-lure (3:97 E-:Z11-14:OAc). Pheromone strain of the same males was determined independently based on pgfar genotype (homozygous pgfar-e/-e and pgfar-z/-z; heterozygous pgfar-e/-z). The count of males from with larval host plant as determined from wing stable carbon isotope content ratio of $^{13}$C/$^{12}$C (i.e., $\delta^{13}$C) are indicated within brackets [C3 ($\delta^{13}$C range of -22 to -32); C4 ($\delta^{13}$C range of -8 to -17)]. ♦ = Harstack wire cone trap; ● = cloth Scentry trap.

<table>
<thead>
<tr>
<th>#</th>
<th>ID</th>
<th>Location</th>
<th>Year</th>
<th>Trap counts</th>
<th>Pheromone strain</th>
<th>[host plant C3/C4]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>E-lure</td>
<td>Z-lure</td>
<td>pgfar-e/-e</td>
</tr>
<tr>
<td>2</td>
<td>Bro10</td>
<td>Brooklawn, PA</td>
<td>2010</td>
<td>48</td>
<td>48</td>
<td>26 [ 0;26]</td>
</tr>
<tr>
<td>3</td>
<td>Dru09</td>
<td>Drums, PA</td>
<td>2009</td>
<td>48</td>
<td>48</td>
<td>10 [ 1; 9]</td>
</tr>
<tr>
<td>5</td>
<td>Lan08</td>
<td>Landisville, PA</td>
<td>2009</td>
<td>32</td>
<td>45</td>
<td>18 [ 0;18]</td>
</tr>
<tr>
<td>6</td>
<td>Leh10</td>
<td>Lehigh, PA</td>
<td>2010</td>
<td>20</td>
<td>48</td>
<td>9 [ 0; 9]</td>
</tr>
<tr>
<td>7</td>
<td>Nor10</td>
<td>Northampton, PA</td>
<td>2010</td>
<td>3</td>
<td>48</td>
<td>0 [ 0; 0]</td>
</tr>
<tr>
<td>8</td>
<td>Oak07</td>
<td>Oaks Corners, NY</td>
<td>2007</td>
<td>48</td>
<td>48</td>
<td>22 [ 0;22]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2010</td>
<td>28</td>
<td>11</td>
<td>19 [ 1;18]</td>
</tr>
<tr>
<td>10</td>
<td>Por09</td>
<td>Port Matilda, PA</td>
<td>2009</td>
<td>48</td>
<td>18</td>
<td>16 [ 1;15]</td>
</tr>
<tr>
<td>11</td>
<td>Ran09</td>
<td>Ransom, PA</td>
<td>2009</td>
<td>47</td>
<td>48</td>
<td>0 [ 0; 0]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2010</td>
<td>44</td>
<td>22</td>
<td>36 [ 1;35]</td>
</tr>
<tr>
<td>12</td>
<td>Roc08</td>
<td>Rockspring, PA</td>
<td>2008</td>
<td>32</td>
<td>32</td>
<td>24 [ 0;24]</td>
</tr>
<tr>
<td>13</td>
<td>Roc09</td>
<td>Rockspring, PA</td>
<td>2009</td>
<td>47</td>
<td>46</td>
<td>30 [28; 2]</td>
</tr>
</tbody>
</table>

This article is protected by copyright. All rights reserved.
Table 2. Distance-based redundancy analysis (dbRDA) of eigenvectors (EV) for the relative contributions to *Ostrinia nubilalis* population genetic structure of pheromone strain [assigned based on **A**) male response (attraction) to either E-strain lure (99:1 *E*-Z11-14:OAc) or Z-strain lure (3:97 *E*-Z11-14:OAc), or **B**) pgfar genotype], host plant history (C3:C4), and geographic distance (km) (pgfar genotype excluded). The sum of squares (Sum of Sq.), *F*-statistic based on 1 and 16 degrees of freedom (*F*₁,₁₆), proportion of the total variance attributed to each component (η²), and associated *P*-values are shown.

### A) Strain assigned by pheromone response (males in lure-baited traps)

<table>
<thead>
<tr>
<th>Component eigenvector</th>
<th>Sum of Sq.</th>
<th><em>F</em>₁,₁₆</th>
<th><em>P</em></th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pheromone strain (response)</td>
<td>0.00021459</td>
<td>2.5390</td>
<td>0.03497</td>
<td>0.115</td>
</tr>
<tr>
<td>Mean C3:C4</td>
<td>0.00010939</td>
<td>1.2943</td>
<td>0.26873</td>
<td>0.059</td>
</tr>
<tr>
<td>Geographic distance EV1</td>
<td>0.00005160</td>
<td>0.6105</td>
<td>0.73826</td>
<td>0.028</td>
</tr>
<tr>
<td>Geographic distance EV2</td>
<td>0.00004956</td>
<td>0.5864</td>
<td>0.71528</td>
<td>0.027</td>
</tr>
<tr>
<td>Geographic distance EV3</td>
<td>0.00008757</td>
<td>1.0361</td>
<td>0.36963</td>
<td>0.047</td>
</tr>
<tr>
<td>Residual</td>
<td>0.00135226</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### B) Population by pgfar genotype

<table>
<thead>
<tr>
<th>Component eigenvector</th>
<th>Sum of Sq.</th>
<th><em>F</em>₁,₁₆</th>
<th><em>P</em></th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pheromone strain (pgfar)</td>
<td>0.0010539</td>
<td>4.4362</td>
<td>0.00199</td>
<td>0.189</td>
</tr>
<tr>
<td>Mean C3:C4</td>
<td>0.0001699</td>
<td>0.7151</td>
<td>0.59640</td>
<td>0.030</td>
</tr>
<tr>
<td>Geographic distance EV1</td>
<td>0.0001569</td>
<td>0.6603</td>
<td>0.69730</td>
<td>0.028</td>
</tr>
<tr>
<td>Geographic distance EV2</td>
<td>0.0002069</td>
<td>0.8711</td>
<td>0.48951</td>
<td>0.037</td>
</tr>
<tr>
<td>Geographic distance EV3</td>
<td>0.0001993</td>
<td>0.8389</td>
<td>0.49550</td>
<td>0.035</td>
</tr>
<tr>
<td>Residual</td>
<td>0.0038010</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>