Explaining the sawtooth: latitudinal periodicity in a circadian gene correlates with shifts in generation number

R. C. LEVY*†, G. M. KOZAK*†, C. B. WADSWORTH*, B. S. COATES† & E. B. DOPMAN*

*Department of Biology, Tufts University, Medford, MA, USA
†USDA-ARS, Corn Insects and Crop Genetics Research Unit, Genetics Laboratory, Iowa State University, Ames, IA, USA

Keywords:
circadian rhythms; developmental timing; diapause; latitudinal cline; phenology; seasonality; voltinism.

Abstract

Many temperate insects take advantage of longer growing seasons at lower latitudes by increasing their generation number or voltinism. In some insects, development time abruptly decreases when additional generations are fit into the season. Consequently, latitudinal ‘sawtooth’ clines associated with shifts in voltinism are seen for phenotypes correlated with development time, like body size. However, latitudinal variation in voltinism has not been linked to genetic variation at specific loci. Here, we show a pattern in allele frequency among voltinism ecotypes of the European corn borer moth (Ostrinia nubilalis) that is reminiscent of a sawtooth cline. We characterized 145 autosomal and sex-linked SNPs and found that period, a circadian gene that is genetically linked to a major QTL determining variation in post-diapause development time, shows cyclical variation between voltinism ecotypes. Allele frequencies at an unlinked circadian clock gene cryptochrome1 were correlated with period. These results suggest that selection on development time to ‘fit’ complete life cycles into a latitudinally varying growing season produces oscillations in alleles associated with voltinism, primarily through changes at loci underlying the duration of transitions between diapause and other life history phases. Correlations among clock loci suggest possible coupling between the circadian clock and the circannual rhythms for synchronizing seasonal life history. We anticipate that latitudinal oscillations in allele frequency will represent signatures of adaptation to seasonal environments in other insects and may be critical to understanding the ecological and evolutionary consequences of variable environments, including response to global climate change.

Introduction

Due to seasonal fluctuations in temperate climates, organisms face continuous challenges in synchronizing their life cycles to resource availability (Tauber & Tauber, 1981; Forrest & Miller-Rushing, 2010). Fitness and survival will be maximized when growth and reproduction coincide with favourable climatic conditions and peaks in resources such as food or available mates. Similarly, survival can be improved by synchronizing dormancy with periods of diminished resources (Tauber et al., 1986). In this way, environmental fluctuations over time result in annual rhythms in the major life history phases of growth, reproduction and dormancy.

Overlaid on seasonal variation in climate is geographic variation in the length of time in which growth and reproduction can take place (season length). Season length has additional consequences for life cycle synchronization. At high latitudes and altitudes, long winters result in extended periods of dormancy compared to brief periods of growth and reproduction (Montesinos-Navarro et al., 2011). As latitude decreases, season length increases and the critical environmental conditions that induce dormancy occur later in the year, whereas the conditions that signal dormancy termination occur earlier (Forrest & Miller-Rushing, 2010; Hut et al., 2013). These changes to dormancy can subsequently shift the...
sawtooth patterns in correlated traits such as body size (reviewed in Kivela et al., 2011; Shelomi, 2012) and the contribution of temperature and host plant availability to these patterns (Scriber & Lederhouse, 1992; Ayres & Scriber, 1994; Scriber, 2002). When sawtooth clines are present, we expect genes associated with life stage duration to show cyclical changes in allele frequency corresponding to transitions from long to short duration, with environmental variation producing gradual increases in the trait and contributing to the observed ‘sawtooth’ phenotypic pattern. However, this idea that periodicity in allele frequency could be related to shifts in voltinism has not been explicitly tested.

In contrast to the sawtooth clines work, a substantial body of research has quantified allele frequency variation related to continuous changes in dormancy-related traits. Most of this research has been on model systems such as Arabidopsis (Stinchcombe et al., 2004; Grillo et al., 2013) and Drosophila (Muona & Lumme, 1981; Schmidt & Paaby, 2008; Tyukmaeva et al., 2011). For example, adult Drosophila enter reproductive diapause during the winter and resume oogenesis in the spring. The onset of diapause varies continuously with latitude (Muona & Lumme, 1981; Schmidt & Paaby, 2008; Tyukmaeva et al., 2011) and has been correlated with linear allele frequency changes in insulin receptors (Paaby et al., 2010) and circadian genes (including period: Costa et al., 1992; timeless: Sandrelli et al., 2007; Tauber et al., 2007; and couch potato: Cogni et al., 2014). Applying a similar approach to species that vary latitudinally in voltinism could relate specific genetic changes to synchronization with seasonal environments and the evolution of voltinism.

An ideal system for associating clinal genetic variation directly with changes in voltinism is the European corn borer moth (Ostrinia nubilalis). The European corn borer (ECB) was introduced to North America in the early 20th century with imported crops from Europe (Caffrey & Worthley, 1927). Since its introduction to the Atlantic Coast, ECB has spread west to the Rocky Mountains and south to the Gulf Coast and has remained a persistent pest of cultivated corn crops (Beck & Apple, 1961; Showers, 1981). Two different pheromone strains were introduced to North America. E-strain females produce and males respond to a pheromone comprised mainly of the E-isomer of 11-tetradecenyl acetate, whereas Z-strain moths use a blend that is primarily the Z-isomer of 11-tetradecenyl acetate, whereas Z-strain moths use a blend that is primarily the Z-isomer (Kochansky et al., 1975). Differences in pheromone blend contribute to reproductive isolation between strains (Kluns, 1975; Dopman et al., 2010), as does voltinism, which varies within and between strains from one (univoltine) to three or more generations (multivoltine) (Showers, 1981; Glover et al., 1992; Dopman et al., 2010). E and Z strains coexist along the Atlantic Coast. All Midwestern populations are Z-strain, and populations are univoltine in the north and increase in voltinism further south (Fig. 1a; Showers, 1981).
Voltinism in ECB is determined in part by diapause induction and termination. The onset of diapause is triggered in larvae by a combination of photoperiodic and temperature cues (Beck & Apple, 1961; McLeod & Beck, 1963). Diapause termination is also photoperiod and temperature dependent and shows variation in ECB (Fig. S1). A 30-day shift in post-diapause developmental (PDD) timing exists between bivoltine (short PDD) and univoltine (long PDD) populations in New York. Variation in PDD has been mapped to a QTL called \( Pdd \) on the Z chromosome (Glover et al., 1992; Dopman et al., 2004, 2005; Wadsworth et al., 2013, 2015). Given this known variation in the duration of diapause termination, we predicted that this trait may show a sawtooth-like pattern across latitude associated with shifts in voltinism (Fig. 1b).

By studying genetic variation along a latitudinal cline, where populations are known to differ in voltinism, we can begin to understand the genetic basis for the evolution of seasonal timing. In this study, we sampled variation in Z chromosome and autosomal genes in Midwestern populations of ECB along a cline in voltinism ecotypes (Fig. 1a). We predicted that loci involved in seasonal adaptation could show several patterns. First, response to continuous environmental gradients of temperature or photoperiod could produce a linear relationship between allele frequency and latitude. Second, given the changes in voltinism observed in North American populations of ECB, we might see abrupt shifts in frequencies corresponding to changes in voltinism ecotype (Fig. 1b,c). Third, due to the possibility that the duration of post-diapause development could vary along with season length, we predicted that any sawtooth patterns in developmental duration might produce periodic oscillations in allele frequencies, such that selection for short and long developmental durations would produce a cyclical pattern across the latitudinal cline (Fig. 1b,d). In addition, we used pedigree mapping families to determine whether associations among markers of interest were the result of physical linkage to each other or with the \( Pdd \) locus.

**Materials and methods**

**Sample collection**

Adult male moths were collected in 2005 using pheromone traps as described in Kim et al. (2009) at ten locations spaced at 80.5 km (50 mile) intervals across a 724.2 km (450 mile) north–south transect centred at Ames, Iowa, USA, and extending south into Missouri (38.8°N) and north into Minnesota (45.3°N) (Fig. 1a; Table S1). A total of 24 individuals from each location \( (n = 240) \) were subsampled from those described in Kim et al. (2009), corresponding to a brief time window of 21 days during the presumed second-generation mating flight of 2005 (19 July through 9 August).
Genotyping assays

Adult moths were stored at −20 °C, and individual DNA extractions were conducted using Bio-Rad’s Aqua Pure isolation kit (Bio-Rad, Hercules, CA, USA) (see Kim et al., 2009). Individuals were genotyped using 194 single nucleotide polymorphism (SNP) markers in seven PCR multiplex reactions on a Sequenom Mass_ARRAY® (Sequenom, San Diego, CA, USA). Five of these assays were previously designed from polymorphic sites in expressed sequence tags (ESTs) as described in Coates et al. (2011) and included seven Z chromosome and 187 autosomal markers (labelled by contig number). Individuals were also genotyped for 42 SNP loci on newly designed Sequenom assays. A total of 42 SNP loci from 32 PCR products were incorporated into two multiplex assays. Polymorphic SNPs were identified from aligned Sanger sequencing reads of PCR-amplified products from 46 unique primer pairs. Specifically, genomic DNA from eight individuals (two individuals from each of four locations 45.3°, 43.1°, 40.9°, 38.8°) was used as PCR template, and the resulting products were sequenced at the University of Chicago Comprehensive Cancer Center DNA Sequencing Facility. Sequence data were aligned using CLC Main Workbench 6 (CLC Bio, Aarhus, Denmark) to identify putative polymorphic sites. Sequenom assays were developed for polymorphic SNPs (Sequenom Assay Design Suite 1.0, Sequenom, San Diego, CA, USA). Of these markers, 29 were putatively located on the Z chromosome, 12 were in putatively diapause-associated at autosomal genes, and one was in mitochondrial DNA (mtDNA) (Table S2; loci labelled by primer name). All SNP assays were run at the Iowa State University Center for Plant Genomics (ISU-CPG). One additional autosomal locus, cryptochrome1, was genotyped in all individuals for an insertion–deletion polymorphism by agarose gel electrophoresis of differentially sized PCR products. Two individuals homozygous for each of the three allele types were Sanger-sequenced to verify that these products were alleles of cryptochrome1.

Genetic linkage of markers

Physical linkage of markers was determined by genotyping four ECB Z-strain pedigree families. Univoltine Z-strain moths were derived from a stock line at the Geneva Agricultural Station and maintained at Tufts University, and bivoltine Z-strain moths were collected from a field site in East Aurora, NY, in 2011. Univoltine males were crossed to bivoltine females, and F1 hybrid male progeny were then backcrossed to bivoltine females. DNA was isolated using DNeasy kits (Qiagen, Venlo, The Netherlands). For each family, the initial parents, backcross parents and all female offspring were genotyped on the seven Sequenom assays at the ISU-CPG. Only female offspring were genotyped due to expected hemizygosity at all Z-linked SNP markers, which allowed differentiation from heterozygous autosomal markers. We genotyped four families: Family 1 (67 female offspring), Family 2 (78 offspring), Family 3 (66 offspring) and Family 4 (75 offspring). Data from families 1–3 were used to construct autosomal linkage groups; data from families 1–4 were used to determine Z-linkage (Family 4 was genotyped only for putatively Z-linked loci). Cryptochrome1 was genotyped via Sanger sequencing in ten individuals per family for families 2 and 4. Genetic linkage maps were constructed separately in each family using R (R Core Team, 2014) and the R/qtl package (Broman & Sen, 2009; Arends et al., 2010). Linkage groups were formed using the estimate map function, a minimum LOD of 3 and a maximum recombination frequency of 0.35. Linkage groups with more than two markers were exported for each family. Linkage groups were merged between families based on shared markers, and additional markers were added to the consensus linkage map from the SNP-based genetic map previously generated for O. nubilalis (Table S3 in Coates et al., 2013).

Female backcross offspring were also phenotyped for post-diapause development (PDD) time. Diapause was induced by rearing larvae from eggs under short-day conditions (12:12 light:dark, 23 °C). After 35 days, diapausing 5th instar larvae were transferred to long-day conditions (16:8, 26 °C). Larvae were checked every other day, and PDD was measured as the number of days from being placed in long-day conditions until pupation.

Marker annotation

Marker annotations were obtained using the whole genome assembly of the related silkworm, Bombyx mori (Integrated sequences, September 2008 version, accessed at KAIKObase: http://spp.dna.affrc.go.jp/pubdata/genomicsequences.html). The BLASTN algorithm was used to determine synteny between ECB linkage groups and the B. mori chromosomes. In addition, marker sequences were queried against the NCBI nonredundant protein database using the BLASTX algorithm. Sequences for markers of interest were aligned using Clustal Omega (Sievers et al., 2011).

Data analysis

Allele frequency variation with latitude

The relationship between allele frequency and latitude was tested for each marker using linear regression in R. To graphically compare the magnitude of the relationship between latitude and allele frequency across markers, the estimated slope from the linear regression for each marker was used to calculate a ‘standardized’ allele frequency at three points (−1, 0, 1). All slopes were plotted so they were increasing. Pearson
Allele frequency variation with voltinism

At least three distinct voltinism ecotypes have been identified in Midwestern ECB that produce one, two or three generations per season (Showers et al., 1975). Populations from the one to two generation region may contain a mixture of the one and two generation ecotypes or may be a distinct ecotype (producing a second generation in some years but not others: Showers, 1981), but for simplicity, we refer to this as the 1–2 generation ecotype. Therefore, groups of populations by voltinism resulted in three ecotypes: (i) one to two generations (the two northernmost sites: 44.6° and 45.3° N), (ii) two generations only (the central six sites: 40.2°–43.9°) and (iii) two to three generations (the two southernmost sites: 38.8° and 39.5°). Genetic differentiation between ecotypes was measured by calculating $F_{ST}^'$ among voltinism groups using the Weir and Cockerham correction (Weir & Cockerham, 1984) in the R package HIERFSTAT (Goudet, 2005). Loci with greater than expected differentiation were detected in two ways. First, the significance of differentiation between voltinism ecotypes was assessed for each locus by Fisher’s allelic goodness-of-fit test in HierFstat (Goudet et al., 1996; Goudet, 2005). $P$-values associated with these $G$ statistics were determined using 200 permutations of random assignment of the data into alternate groups. Second, an explicit $F_{ST}^'$ outlier test was performed using Fdist as implemented in the JAVA program LOSITAN (Beaumont & Nichols, 1996; Antao et al., 2008). Data were converted from Fstat format using the program FGSPIDER (Lischer & Excoffier, 2012). In Lositan, neutral $F_{ST}^'$ was estimated from the data set in an initial run, and then 50,000 simulations were performed to identify outliers. Outliers for directional selection were loci in which the locus-specific $F_{ST}^'$ was greater than the population-level $F_{ST}^'$ in 97.5% or more of the simulations.

Allele frequency variation with population

Total and pairwise $F_{ST}^'$ were estimated for all ten populations for all markers in HierFstat. Isolation-by-distance (IBD) models assume that genetic differentiation is positively related to the geographic distance between populations (Wright, 1943). Evidence for a linear relationship between geographic distance and pairwise population $F_{ST}^'$ estimates was tested using a Mantel test in the program BIBWS version 3.23 (Jensen et al., 2005). Mantel tests were calculated using (i) pairwise $F_{ST}^'$ for all markers and (ii) pairwise $F_{ST}^'$ calculated using only the 131 markers without a significant linear relationship between latitude and allele frequency. In addition, outlier analyses were performed among all 10 populations in HierFstat and Fdist (similar to those performed among voltinism ecotypes).

Cyclical patterns in allele frequency

It was hypothesized that some loci might show cyclical or periodic patterns associated with differences in developmental duration that would not be easily detected in a linear regression of allele frequencies on latitude (Fig. 1d). Periodicity in allele frequency was tested using the Jonckheere-Terpstra-Kendall algorithm, a nonparametric test of periodicity as implemented in the R package JTK CYCLE (Hughes et al., 2010; Deckard et al., 2013). Our data set contained 10 data points (populations) and JTK cycle tested for cyclic patterns that completed one full cycle between 5 and 9 data points, estimating the cycle and amplitude and then calculating a Bonferroni-corrected $P$-value for periodicity.

Results

Single nucleotide polymorphism (SNP) genotyping

Samples were genotyped for one mtDNA, 38 Z-linked and 155 autosomal SNPs using Sequenom assays. All 194 markers produced base calls for at least one sample. The mean number of individuals successfully genotyped was $210.8 \pm 42.9$ per SNP, with a median of 225 individuals. After quality filtering of SNPs with low polymorphism (see Table S3), 145 SNP markers (Z-linked: 35 markers from 21 loci, autosomal: 109 markers from 101 loci, mtDNA: 1 locus) and one autosomal polymorphic indel locus were retained for statistical analyses.

Genetic linkage of molecular markers

We generated 30 linkage groups with more than three markers using the combined data from our pedigree families (mean markers per linkage group = 9.6, median = 5, range = 3–35). Correspondence of our groups to B. mori chromosomes is listed in Table S4. All putative Z-linked markers, with one exception (537), showed evidence for sex linkage in our pedigrees based on lack of heterozygous genotypes for hemizygous females. Sanger sequence data of cryptochrome1 in families 2 and 4 suggested that the marker was autosomal due to observed female heterozygotes, but this marker could not be placed in a linkage group.

Allele frequency variation with latitude

Eleven of the 110 polymorphic autosomal markers (10.0%) and four of the 35 Z-linked markers (11.4%) showed significant correlations between allele frequency and latitude (Table S5; $P < 0.05$; FDR corrected
0.133 ≤ q ≤ 0.497). The magnitudes of the standardized slopes between allele frequency and latitude are compared in Fig. S2a; marker-specific slopes are shown in Fig. S3. Cryptochrome 1 (cry1) showed three indel alleles, and the major allele showed a significant correlation between allele frequency and latitude (r = −0.75, P = 0.012, q = 0.249). Other annotated markers with significant correlations with latitude included four Z-linked markers: Lactate dehydrogenase (Ldh), Shaker (Shk) and Glutamine synthetase (680 and contig00968).

These 15 latitude-associated markers were responsible for the overall genomewide pattern of isolation by distance among our populations (Fig. S2b); when they were removed from the data set, no significant relationship was detected between pairwise FST and geographic distance (Fig. S2c). Allele frequency was significantly correlated among some markers (Fig. S4). The four Z-linked markers that showed an association with latitude (Ldh, Shk; 680, contig00968) were correlated with one another, with contig07183 and with contig01517. Cry1 major allele frequency was correlated with period (per), contig01517 and contig00845. The first minor allele of cry1 was also correlated with per. Many of these correlations also had an FDR corrected q < 0.05 (Fig. S4).

Allele frequency variation with voltinism
HierFstat analysis of genetic differentiation between the three voltinism ecotypes estimated a relatively low global FST = 0.0015 (95% CI: −0.001 to 0.0033). Eight markers, including cry1, contig05909 (CG32521) and per, had significant genetic differentiation based on Fisher’s G statistic (Fig. 2; Table 1; P < 0.05, FDR corrected 0.238 ≤ q ≤ 0.677). Using Fdist, cry1 (FST = 0.0547) and contig05909 (FST = 0.116) were identified as FST outliers that may be influenced by directional selection.

Allele frequency variation with population
The genetic differentiation among all 10 populations was also low with an estimated global FST = 0.0022 (95% CI: −0.0002 to 0.0046). HierFstat results identified eleven markers that showed significant differentiation among populations (Table 1), of which only two (cry1 and contig06459) were also among the HierFstat voltinism outliers. Five markers were identified as FST outliers among populations in Fdist: 2076, contig05909, contig06459, contig00382 and alpha-amylase (amy). Only contig05909 was also present among the voltinism outliers.

Cyclical patterns in allele frequency
In addition to showing genetic differentiation among voltinism types, per also showed a significant cyclical pattern (Fig. 2c; Table 1; Bonferroni-corrected P = 0.033, cycle = 5, amplitude = 0.12). Per showed an oscillating pattern with peaks in allele frequency at 39° and 42° latitude and troughs at 41° and 45°. The pattern of variation in per was correlated with cry1, and the first minor allele of cry1 also showed a cyclical pattern (Table 1; Fig. 2a; Bonferroni-corrected P = 0.023, cycle = 6, amplitude = 0.05). Three other markers also had significant cyclical patterns: contig05725 (Eukaryotic translation initiation factor 3 subunit H), contig00280 (Heat shock protein hsp21.4) and contig05979 (probable cysteine desulfurase).

Period and post-diapause development time
Our primers amplified at 519-bp portion of period containing both coding and noncoding sequence. A polymorphic G/T SNP at position 80 in the sequence was genotyped in all individuals via our Sequenom assay. This SNP corresponds to an amino acid change from the hydrophobic alanine to the negatively charged aspartic acid (Fig. 3a). In our latitudinal cline, the T allele was at high frequency in the far south (38.8–39.5°) where the G allele was at higher frequency. Pedigree families 1, 2 and 4 were polymorphic for this SNP. In these three families, females who inherited the G allele from the univoltine grandparent had long post-diapause development time, whereas females who inherited the T allele from the bivoltine grandparent had short development time (Fig. 3b; Family 1: G allele PDD time = 40.10 ± 1.85 days, T allele PDD time = 27 ± 1.01, two-tailed unequal variance t44 = 6.23, P < 0.0001; Family 2: G PDD time = 45.87 ± 1.55, T PDD time = 35.47 ± 1.19, t70 = 5.29, P < 0.0001; Family 4: G PDD time = 38.42 ± 1.45, T PDD time = 29.95 ± 1.15, t56 = 4.59, P < 0.0001).

Discussion
Our results suggest that photoperiod is an important abiotic factor shaping latitudinal variation in voltinism in the European corn borer moth. Many of the outlier loci between voltinism ecotypes were related to the circadian clock pathway (HierFstat: 4 of 8 outliers; Fdist analyses: 2 of 2 outliers). The circadian pathway senses changes from short- to long-day photoperiods through light-dependent patterns of gene expression (Stoleru et al., 2007; Meuti & Denlinger, 2013). Allelic variation in the circadian genes cryochrome1 and period was associated with clinal differences in voltinism. Cryochrome1 allele frequency showed a significant correlation with latitude, was an FST outlier locus between our three voltinism ecotypes and showed evidence of a cyclical pattern. Period showed variation between the voltinism ecotypes as well as variation within the
bivoltine ecotype, producing a significant cyclical pattern (Fig. 2c). Contig05909 was another $F_{ST}$ outlier locus and was predicted to be homologous to *Drosophila* protein CG32521, a downstream target of the retinal development transcription factor *eyeless* (Ostrin et al., 2006). Contig01517 showed a strong association with latitude and cry1 and encodes a gamma subunit of guanine nucleotide-binding protein (G-protein). As with the beta subunit (RACK1), the gamma subunit may function in photoreception and the regulation of circadian genes like cry1 (Hamasaka et al., 2005; Wang & Montell, 2007; Robles et al., 2010). Therefore, the circadian pathway appears to be important for evolutionary changes in seasonal timing.

*Cryptochrome* is a blue light photoreceptor that contributes to the regulation of circadian rhythms in plants and animals (Cashmore et al., 1999; Stoleru et al., 2007). In *Drosophila*, light stimulates CRYPTOCHROME to degrade the protein TIMELESS (TIM), preventing dimers from forming between TIM and PERIOD (PER), and thus activating the expression of other circadian genes (Stanewsky et al., 1998; Cashmore et al., 1999). *Drosophila* have a single *cryptochrome*, but Lepidoptera have two gene copies (*cry1* and *cry2*) through gene duplication (Zhu et al., 2005). The ECB *cryptochrome* gene sequence in this study showed greater homology to the Lepidopteran CRY1 protein sequence (BLASTX for CRY1: score = 81, e-value < $1 \times 10^{-15}$; CRY2: score = 51.52, e-value < $3 \times 10^{-10}$). *Cry1* is orthologous to *Drosophila cryptochrome*. In contrast, *cry2* is similar to the mammalian protein (*cry-m*), which binds directly to PER to function as a transcriptional repressor.

---

**Fig. 2** Outlier loci among voltinism ecotypes. Allele frequency, latitude and voltinism ecotype (indicated by vertical dashed lines) plotted for (a) *cryptochrome1* voltinism outlier (HierFstat and Fdist) and cyclical outlier which had three alleles: major allele (solid line; significant linear pattern), minor allele 1 (dashed line; significant cyclical pattern) and minor allele 2 (dotted line); (b) contig5909 (c05909.680/CG32521) voltinism outlier (HierFstat and Fdist); (c) *period* voltinism outlier (HierFstat) and cyclical outlier.
of the circadian clock (Zhu et al., 2008; Meuti & Denlinger, 2013). In *D. melanogaster*, variation in *cryptochrome* has no known association with latitude; rather polymorphism appears to be maintained by balancing selection in Europe (Pegoraro et al., 2014). *Cry1* has been linked to eclosion time in *D. melanogaster* (Pegoraro et al., 2014) and reproductive diapause in *D. triauraria* (Yamada & Yamamoto, 2011). In bean bugs (*Riptortus pedestris*), *cry1* is related to pupal diapause incidence (Ikeno et al., 2014). Our results are the first to specifically link polymorphism in *cryptochrome* to insect volatilism.

We found oscillations in *period* allele frequency across latitude consistent with spatially varying selection, which is notably different from the linear allele frequency clines in *period* that have been described in *Drosophila* (Sawyer et al., 2006; Kyriacou et al., 2008). In other insects, phenotypic sawtooth clines have been extensively documented (reviewed in Kivelä et al., 2011; Shelomi, 2012), but our current study is the first to find a cyclical pattern in allele frequency corresponding to changes in volatilism. Variation in *period* has previously been associated with temperature sensitivity of reproductive diapause in *Drosophila* (Huang et al., 1995) and pupal diapause incidence in flesh flies (Han & Denlinger, 2009) and linden bugs (Doležel et al., 2007).

This oscillating pattern in *period* is consistent with our expectation for an allele underlying a sawtooth phenotypic cline. Although further sampling is needed to verify that the oscillating pattern in genotype corresponds directly to a sawtooth cline in diapause phenotype, our results are concordant with earlier work that found a periodic pattern in diapause induction across these same latitudes (Beck & Apple, 1961). Under controlled laboratory conditions (photoperiod = 14.5 hours light, temperature = 26 °C), ECB from Wisconsin populations (latitude = 43° C) had higher diapause incidence (DI = 42%), while ECB from Iowa (latitude = 42° C, DI = 45%) had higher diapause incidence (DI = 42%). Diapause incidence increased further south in Kansas populations (latitude = 39° C, DI = 69%) before decreasing again in Missouri (latitude = 38° C, DI = 45%; Fig S5). Based on this previous work, we propose that the mechanism generating this cyclical pattern may be an association between *period* and some genetic factor associated with diapause and/or the duration of the dormant life stage.

In ECB, *period*, the QTL controlling variation in diapause termination time and the QTL for diapause induction all map to the Z chromosome (Glover et al., 1992; Dopman et al., 2004, 2005; Iketen et al., 2011; Wadsworth et al., 2015). Therefore, *period* may be physically linked to one or more of these QTL. The QTL for

---

**Table 1** Summary of loci that show significant differentiation between volatilism ecotypes or between populations.

PDD differences between bivoltine E-strain and univoltine Z-strain ECB (Pdd) is linked to the Tpi locus, with one allele corresponding to a short pupation time of 15.3 days and the other allele leading to a long pupation time of 43.7 days (Glover et al., 1992; Dopman et al., 2005). Recent work found that period co-segregates with Pdd and Tpi in these populations (Wadsworth et al., 2015). In the univoltine and bivoltine Z-strain pedigree families in our study, there was also a strong association between period and PDD time. One SNP allele was associated with long PDD time and the other with short PDD (Fig. 3b). Interestingly, the short PDD allele we identified is predicted to cause an amino acid change (from alanine to the negatively charged aspartic acid) two residues away from the estimated start of the Per-Arnt-Sim-B (PAS-B) domain and downstream of the first TIMELESS interacting site (TIS-1) in the PERIOD sequence of the silkworm, B. mori (Iwai et al., 2006). Ostrinia furnacalis, the Asian corn borer, also has alanine in this position of the protein (Regier et al., 2012). The PAS-B domain contributes to the stability of PER-TIM heterodimers in Drosophila (Hennig et al., 2009) and PER-PER homodimers in mammals (Kucera et al., 2012), suggesting that a change in this region may affect these protein–protein interactions.

If period is associated with Pdd, the oscillating pattern we observed could be explained by spatially varying selection on the timing of diapause emergence. In the northernmost sites (44.6° and 45.3°), we found the long Pdd allele at higher frequency. This may because at these latitudes, the long allele is prevalent in univoltine populations, whereas the short Pdd allele could be prevalent in bivoltine populations to accommodate the development of two complete generations in a single short growing season (indicated by a in Fig. 1b). If our sample contained a majority of univoltine individuals, it would explain the high frequency of the long allele at these northern sites. Consistent with this idea, samples from northerly bivoltine populations between 42 and 44° latitude also had the short Pdd allele at high frequency. At southerly bivoltine locations in Iowa (40–42°), we again found the long Pdd allele at higher frequency. This could be due to the fact that growing seasons are longer but insufficient to accommodate a third generation, which may lead to increases in the frequency of the long allele (Fig. 1b-b). At latitudes further south, a third generation can be produced and short Pdd alleles might allow for fitting of all three generations into a season (Fig. 1b-c). These data, together with the oscillating pattern Beck & Apple (1961) found in diapause incidence (Fig. S5), suggest that seasonal fitting produces oscillations in alleles associated with diapause in ECB, primarily through changes in the duration of transitions between diapause and other life history phases. Work in tiger swallowtail butterflies (Papilio canadensis and P. glaucus) has also found differences in voltinism which are associated with a Z-linked factor and that period shows genetic differentiation.
between these Papilio species (Putnam et al., 2007; Kunte et al., 2011), suggesting period could be associated with voltinism in distinctly related Lepidoptera.

Colonization of North America by ECB likely included bouts of rapid adaptation to seasonal environments. Moths from this study were collected ~ 60 years after their introduction to the Midwest in the 1940s (Beck & Apple, 1961; Palmer et al., 1985). It is possible that certain neutral processes could cause clinal patterns at some loci by chance. One such process might be rapid population expansion, which can cause differentiation via gene surfing along the expansion front (Excoffier & Ray, 2008). Another possibility is that admixture of populations from different locations in Europe during colonization of North America led to nonrandom assortment of alleles (Barton & Hewitt, 1985). Different allelic variants in North America likely arose through admixture between colonizing populations from Italy and Hungary (Caffrey & Worthley, 1927), but admixture is unlikely to provide an explanation for the specific clinal patterns we observed. Under neutral processes, we would not expect loci that show clinal and oscillating patterns to be involved in any specific pathways. Loci associated with circadian rhythms and diapause loci represented 7.5% of our molecular markers, with the other 92.5% being anonymous Z-linked or autosomal loci. However, circadian and diapause-associated markers represented 14% of markers showing linear associations with latitude, 37.5–50% of F_{ST} outliers between voltinism groups and 40% of the cyclic outliers. More importantly, period allele frequency showed a strong oscillating pattern with latitude, and we found that this locus co-segregated with post-diapause termination time phenotype in crosses between univoltine and bivoltine ECB populations. Such an association between the genotype of a cyclic allele and a diapause-associated phenotype is unlikely to be due to chance. The magnitude of allele frequency variation in cryptochrome1 (0.29) and period (0.25) over our 6.5° latitudinal cline is comparable to the variation observed in period (Thr-Gly)_{20} alleles in Australian Drosophila (allele frequency variation of 0.4 over an 25° cline) that has evolved via natural selection in the past 200 years (Sawyer et al., 2006; Kyriacou et al., 2008). Thus, it seems likely that latitudinal patterns in circadian gene allele frequency in ECB are driven by adaptation to seasonal environments. This association could be confirmed by sampling replicate latitudinal transects across North America and Europe in future studies.

Our results suggest that differences in voltinism in ECB are associated with changes in the circadian clock pathway. Allele frequency changes for cryptochrome1 and period were strongly correlated with each other, more strongly than period allele frequencies were correlated with other markers known to be physically linked to period on the Z chromosome (such as Ldh; Fig. S4). Rather than being physically linked in ECB, cry1 and period could show correlations as a consequence of epistatic selection, possibly because of their interactions with timeless, a gene showing clinal variation in Drosophila (Tauber et al., 2007). It is unclear why multiple circadian genes show correlated changes associated with voltinism in ECB and which gene(s) in the pathway might be the primary target(s) of natural selection. Thus, the precise contribution of cry1, period, timeless and other circadian genes to circannual rhythms, patterns of voltinism and seasonal timing warrant further study. We must also confirm tentative associations between period and diapause emergence timing and/or induction phenotypes. Changes in phenology in many plants and animals involve changes in both dormancy induction and termination (Forrest & Miller-Rushing, 2010). Therefore, understanding how these two seasonal timing phenotypes evolve in concert to enable adaptation to variable and rapidly changing environments represents a major question in biological research.

Conclusions

The extremely rapid evolution of allelic variation observed in North American ECB strongly suggests that natural selection is driving changes between and within voltinism ecotypes in order to fit additional generations into a season. We propose that selective pressures on seasonal fitting generate sawtooth phenotypic clines in the duration of life stages and associated oscillations in allele frequency in developmental timing genes. Sawtooth clines have been the subject of over 30 years of research (Masaki, 1978a,b; Roff, 1980), and we anticipate that latitudinal patterns of cyclic genetic variation may be key signatures of seasonal timing traits and represent sawtooth clines at the genetic level. Contemporary increases in temperature are driving local increases in season length and generation number in Lepidoptera (Breed et al., 2013; Scriber, 2014; Scriber et al., 2014), suggesting that these same genes may show cyclical patterns across decades. Thus, scouting for such patterns in other taxa will allow us to identify the genes associated with seasonal environments, understand how populations adapt to climatic variation throughout their geographic ranges and predict how populations will respond to changes in climate.

Acknowledgments

We thank Mitzi Wilkening and the other staff at the Iowa State University Genomic Technologies Facility for performing the Sequenom Assay. Roxana Khozein and Xinnan Li assisted with primer design and Sequenom assay marker development; Charles Linn provided insect stocks; Thomas Sappington and Kyung Seok Kim provided field collected samples; Shoshanna Kahne assisted with figures and pedigree family genotyping; and Gabriel Golczer provided help with R scripts. Tho-
mas Flatt, Mark Scriber and two anonymous reviewers provided helpful feedback on the manuscript. This study was funded by National Science Foundation (EBD: DEB-1257251) and United States Department of Agriculture (EBD: 2010-65106-20610). A portion of this research was also supported by the United States Department of Agriculture, Agricultural Research Service (BSC: CRIS Project 3625-22000-017-00) and the Iowa Agriculture and Home Economics Experiment Station, Ames, IA (BSC: Project 3543).

Data accessibility


References


Supporting information

Additional Supporting Information may be found in the online version of this article:
Figure S1 Life cycle and timing of diapause of bivoltine and univoltine European corn borers.

Figure S2 Linear associations with latitude.

Figure S3 Locus-specific associations with latitude.

Figure S4 Correlation matrix among markers showing a relationship with latitude.

Figure S5 Diapause incidence in ECB.

Table S1 Sampling locations, listed from north to south, with corresponding latitude (as noted in Figure 1), number of European corn borer generations per season, and mean degree days for the 2005 and 2006 seasons.

Table S2 Sequenom assay markers.

Table S3 Summary of Sequenom assay results.

Table S4 ECB linkage groups from pedigree families.

Table S5 Summary of correlation and linear model statistics for all markers with statistically significant correlations between allele frequency and latitude.

Data deposited at Dryad: doi:10.5061/dryad.cg2gg

Received 29 August 2014; revised 21 November 2014; accepted 24 November 2014