1 Title: **Genetic variation underlying local adaptation of diapause induction along a cline in a**
2 **butterfly.**

3
4 Short title: **Genetic variation underlying diapause induction.**

5
6 Author Affiliation: Peter Pruisscher^a, Sören Nylin^a, Karl Gotthard^{a1}, and Christopher W.
7 Wheat^{a1}

8
9 a: Department of Zoology, Stockholm University, Svante Arrheniusväg 18b, S-106 91
10 Stockholm, Sweden

11 1: K.G. and C.W.W. should be considered joint senior author

12
13 Corresponding Author: Peter Pruisscher, peter.pruisscher@zoologi.su.se, [orcid.org/0000-](https://orcid.org/0000-0002-6987-5839)
14 [0002-6987-5839](https://orcid.org/0002-6987-5839)

Abstract

Diapause is a life history strategy allowing individuals to arrest development until favourable conditions return, and it is commonly induced by shortened day-length that is latitude specific for local populations. Although understanding the evolutionary dynamics of a threshold trait like diapause induction provides insights into the adaptive process and the adaptive potential of populations, the genetic mechanism of variation in photoperiodic induction of diapause is not well understood. Here we investigate the genetic variation underlying latitudinal variation in diapause induction and the selection dynamics acting upon it. Using a genome-wide scan for local adaptation between two populations of the butterfly *Pararge aegeria* that differ strongly in their induction thresholds, we identified and investigated regions of high divergence. We tested the association of these regions with diapause induction using between-population crosses, finding significant SNP associations in four genes present in two regions, one with the gene *period*, and another region with the genes *kinesin*, *carnitine O-acetyltransferase*, and *timeless*. Patterns of allele frequencies in these regions along a sampling of the larger regional cline suggest strong selection against heterozygotes at two genes within these loci (*period*, *timeless*). Evidence for additional loci modifying the diapause decision was found in patterns of allelic change in relation to induction thresholds over the cline, as well as in backcross analyses. Taken together, population specific adaptations of diapause induction appear to have used a combination of alleles of larger and smaller effect size, consistent with an exponential distribution of effect sizes involved in local adaption.

- 38 Keywords (at least three): Lepidoptera, diapause induction, crosses, selection dynamics,
39 adaptation, ecological genetics

INTRODUCTION

Numerous organisms in temperate latitudes use environmental cues to prepare for predictable seasonal stresses, such as winter, by entering a pre-programmed state of dormancy before the onset of adverse environmental challenges, where they remain until the return of favorable conditions (Danilevskii 1965; Perry 1971; Tauber & Tauber 1976; Simpson & Dean 2002). In insects, such pre-programmed dormancy is known as diapause and it is often induced by shortened photoperiod (Danilevskii 1965; Tauber & Tauber 1976), which sets into motion a dynamic, yet distinct series of phases (preparation, developmental arrest, resumption of development) that facilitate life-cycle synchronization with optimal conditions (Košťál 2006; Košťál *et al.* 2017). Even though the developmental stage in which diapause occurs varies widely, most insects with facultative diapause exhibit strong latitudinal clines in their critical photoperiod (CPP) for diapause induction (CPP = photoperiod needed to induce diapause in 50% of a population). While these latitudinal clines demonstrate a fine-scale regional adaptation to the historically reliable relationship between photoperiod and the onset of winter (Danilevskii 1965; Bradshaw 1976; Hard *et al.* 1993; Bradshaw *et al.* 2003; Kato 2005; Demont & Blanckenhorn 2008; Wang *et al.* 2011; Huang *et al.* 2013; Paolucci *et al.* 2013; Aalberg Haugen & Gotthard 2015), little is known about the underlying genetic basis of CPP clines.

Understanding the evolutionary dynamics of local adaptation for a threshold trait like diapause induction provides not only insights into the adaptive process, but also reveals the adaptive potential of populations. This has particular relevance to the response of species to climate change, as we lack explanations for variability among species in their response to

increasing mismatches between seasonality and life-cycle timing, as some are not responding (Van Dyck *et al.* 2014), while others are adapting (Bradshaw & Holzapfel 2001). Theoretical predictions of the genomic architecture of traits that have become locally adapted are well-developed, and suggest either an infinitesimal distribution with many loci of small effect size (Roff 1996; Barton & Keightley 2002), or an exponential distribution with a combination of large and small effect loci. In the exponential model adaptation would have evolved first by selection on large effect loci, that would then be modified by smaller effect loci as a fitness optimum is approached (Orr 2000).

For diapause induction, empirical insights acquired from population crosses and QTL studies are consistent with Mendelian inheritance patterns (Doležal *et al.* 2005; Suwa & Gotoh 2006; Han & Denlinger 2009), or indicate an architecture with few loci of large effect sizes (Danilevskii 1965; McCoy *et al.* 1968; Kurahashi & Ohtaki 1977; Rockey *et al.* 1987; Hagen & Scriber 1989; Kim *et al.* 1995; McWatters & Saunders 1997; Demont & Blanckenhorn 2008; Söderlind & Nylin 2011; Ikten *et al.* 2011; Xia *et al.* 2012; Chen *et al.* 2014; Fu *et al.* 2015; Lehmann *et al.* 2016; Pruisscher *et al.* 2017), with several reporting large sex-linked effects (Rockey *et al.* 1987; Hagen & Scriber 1989; Ikten *et al.* 2011; Chen *et al.* 2014; Fu *et al.* 2015; Pruisscher *et al.* 2017). Studies on quantitative trait loci for CPP show that effect sizes can be highly variable depending on the genetic background of the crosses involved (Bradshaw *et al.* 2012). Additionally, as diapause induction is a plastic threshold trait, inheritance patterns depend on the environmental conditions used for phenotyping (Pruisscher *et al.* 2017). Thus, whether diapause induction thresholds arise from a few or many loci is unresolved,

suggesting a complex genetic basis may underlie these observations. In sum, further work is needed to understand how the genetic architecture of this trait varies among populations.

The identity of genes involved in insect diapause induction have been investigated using candidate gene studies with several important advances. For instance, circadian clock genes have been a target of many studies, and appear to be involved in diapause induction in some studies (Ikeno *et al.* 2010; Meuti *et al.* 2015), but not in others (Emerson *et al.* 2009). Furthermore, there have been several studies of genetic variation in other diapause-related traits that vary latitudinally (Schmidt *et al.* 2008; Levy *et al.* 2015)}. However, while these studies can potentially identify genes that are involved in the expression of traits, such as diapause, they do not inform on which loci underlie adaptive latitudinal variation in diapause induction, or the selection dynamics that have acted to shape these latitudinal clines. This lack of empirical insights into the evolutionary dynamics acting upon CPP in the wild makes it difficult to reconstruct the history and predict the future potential for adaptation to changing environmental conditions.

Here, we use the speckled wood butterfly *Pararge aegeria* in an attempt to gain an understanding of the genetic variation and the selection dynamics acting upon latitudinal variation in CPP. *P. aegeria* is an ecological model species for latitudinal variation in CPP. Pupal diapause in *P. aegeria* is induced primarily by photoperiod during larval development, where individuals establish the diapause decision prior to pupation (Friberg *et al.* 2011). This species exhibits a striking latitudinal variation in CPP that gives rise to a cline in voltinism in Sweden (Nylin 1995; Aalberg Haugen *et al.* 2012; Aalberg Haugen & Gotthard 2015). Here

we use a pooled sequencing approach to conduct a genome wide scan of differentiation between two populations that significantly differ in their CPP, followed by individual level SNP genotyping in candidate genes along a latitudinal cline and in population crosses. Our goal is to identify which of the genome-wide differences between populations give rise to a latitudinal cline in CPP, and then evaluate the selection dynamics acting on these genes.

We do this by: 1) assessing genomic variation within two populations that differ in their latitude and CPP thresholds, followed by identifying candidates for local adaptation, 2) testing the association between these candidate loci for local adaptation and diapause incidence in F2 hybrid offspring, 3) investigating genomic signatures of selection in the population samples at the loci significantly associated with diapause induction, 4) evaluating allele frequencies of the candidate genes in five Swedish populations along a cline and 5) exploring if other genomic regions are affecting CPP thresholds for induction by assaying additional crosses.

MATERIALS AND METHODS

Genome Sequencing. Genomic DNA was extracted from thoraces of 22 individuals from each of two Swedish populations collected in 2010 and 2011 for a previous study (Tison *et al.* 2014): samples 1-22 from Skåne (further referred to as the southern population, S), 84-94; 100-110 from Sundsvall (further referred to as the northern population, N), using a DNeasy blood and tissue kit (Qiagen) with an extra RNase A treatment to remove potential RNA contamination. DNA quality was checked on 2% agarose gels stained with GelRed, to ensure minimal fragmentation, and UV-Vis Spectrometer (NanoDrop 8000, Thermo

131 Scientific) to assess purity. All samples showed minimal fragmentation with an absorbance
132 260/280 > 1.7 and < 2.0. For each population, samples were combined at equal
133 concentration, resulting in two pools of 5 ug, RNA free, gDNA. SciLifeLab (Uppsala, Sweden)
134 performed the library preparation (Illumina TruSeq DNA PCR-free library) and sequencing
135 (Illumina HiSeq2000, 100bp paired end reads, 450 bp insert size). Data analysis is described
136 below, and detailed command lines are provided in SM: scripts.

137 **Illumina read processing.** Illumina fastq files were filtered for PCR duplicates using the
138 clone_filter script of Stacks-1.21 (Catchen *et al.* 2013), followed by filtering of Illumina
139 sequencing adaptors, quality trimming to a minimum phred base quality of 10 and
140 discarding broken read pairs, using bbduk2 (BBMap v35.69 - Bushnell B. –
141 <http://sourceforge.net/projects/bbmap/>).

142 **De novo genome assembly.** Four *de novo* genome assemblies were generated from the
143 combined reads of both populations, as well as an assembly for each of the two population
144 data sets, using CLC Genomics Workbench 5.5.1 (CLCbio) with differing settings (Table S1).
145 Each *de novo* assembly was rated on N50 size and gene completeness, and the assembly
146 using the combined data sets with word size 64 and bubble size 2000 was used for further
147 analysis as it outperformed all others. Gene completeness was assessed using MESPA v1.0
148 (Neethiraj *et al.* 2017) using default parameters, excluding xenobiotic filtering. The protein
149 dataset from three different butterfly genome projects and the silkworm genome project
150 were compared for use in annotating the assembly. For each of these protein sets,
151 redundancy was controlled by using the longest protein isoform of each protein cluster
152 group based upon <90% amino acid identity using CD-HIT v4.5.4 (Li & Godzik 2006) with
153 parameters used in making the UniRef90 database (Suzek *et al.* 2014). Protein numbers after

clustering were: 14,818 proteins for *D. plexippus* protein set DPOGS2 (Zhan & Reppert 2013), 12,591 for *Heliconius melpomene* Ensembl protein set Hmel1.27 (Kersey *et al.* 2016), 16,634 for *Melitaea cinxia* protein set v1.0 (Ahola *et al.* 2014), and 18,574 for *Bombyx mori* protein set A, B, and C (Suetsugu *et al.* 2013). The *D. plexippus* protein set generated the best results and was used to annotate and scaffold the genome assembly (Table S2). MESPA outputs sequences where no genes are found, named as e.g. contig_1, or c_1, and sequences that contain genes, denoted m_scaffold_2 or ms_2. This naming convention is continued in the tables and figures of this manuscript when referring to a specific contig. We used BUSCO v3.0.2 (Simão *et al.* 2015) on default settings, which looks at the fraction of completely covered conserved genes, as a method to assess completeness of the genome.

Gene model order. Relative chromosomal location of the *P. aegeria* gene models was inferred using orthology from the high-quality *D. plexippus* scaffolds and *B. mori* chromosomes, as synteny is well-conserved within Lepidoptera (Dasmahapatra *et al.* 2012). As at the time of writing *B. mori* is the most complete and well-annotated lepidopteran genome available, it was used as chromosomal location template for *D. plexippus* scaffold order. Predicted *P. aegeria* proteins were searched against the combined protein sequences gene set A, B, and C of *B. mori* using blastp (Camacho *et al.* 2009), taking only hits above a minimum E-value cutoff of 1e-4. *B. mori* chromosomal location was parsed out from the gene_ID category, of the annotation file available from these *B. mori* protein sets (Suetsugu *et al.* 2013). *D. plexippus* gene location was taken from the DPOGS2 gene annotations (Zhan & Reppert 2013). The resulting gene order inferred for *P. aegeria* was considered ‘likely’ when both species showed the same results. If there was no clear consensus, *D. plexippus* gene order was taken as correct as *D. plexippus* is in the same family as *P. aegeria*, while *B.*

177 *mori* is much more divergent (Wheat & Wahlberg 2012). We note that although synteny is
178 not exact between species, gene content within chromosomes is highly conserved and the
179 majority of genes are collinear between species (Ahola *et al.* 2014). Thus, our comparative
180 approach assuming synteny is likely sufficient to reveal whether the genomic differentiation
181 investigated is clustered within a specific chromosome, chromosomal regions (e.g.
182 telomeres), or distributed across the genome.

183 **Read mapping.** The Pool-Seq read data were mapped to the assembly using NextGenMap
184 0.4.10 at 90% identity (Sedlazeck *et al.* 2013), in order to minimize template mapping biases.
185 Plotting read depth (RD) per scaffold showed a bimodal distribution where the first modus
186 possibly represented misassembled or population specific haploid sequences, and the
187 second modus representing sequences that occurred at equal coverage in both populations
188 (Figure S1). By setting minimum RD cut-offs to >20, these first modus scaffolds were
189 excluded from analyses. This also means that any gaps present in one population, but not in
190 the other, are automatically excluded from analysis, as they will not meet the RD criteria.
191 Mapped reads were filtered using SAMtools v1.2 (Li 2009), keeping read pairs where both
192 reads mapped correctly.

193 Investigating read depth revealed an overall linear relationship between the two populations
194 $R^2=0.81$, Figure S1), with the exception of 271 sequences in the southern population, and
195 2242 sequences in the northern population with a read depth <10 in that population (Figure
196 S2), but a range of read depths in the other. In order to exclude population-specific
197 sequences or misassembly as reason for the bias in read depth we performed a blastx
198 alignment against the uniprot protein database. We parsed out the top hits and noted
199 whether these sequences aligned to known sequences of other Lepidopteran species, and

revealed that only two of these 2513 population-specific contigs were Lepidopteran in origin, indicating low population-biased read mapping. The other contigs represented genomic sequences of bacteria and plants (Table S3). The parts of the thoraces that were sequenced include potential microbiota living in and on the organism, as well as part of the gut. The read depth filters employed exclude these sequences from the analysis.

Variant calling. In order to quantify differentiation between the populations, SNPs were identified and F_{ST} was calculated at each base pair and using a window-based analysis of the assembly using PoPoolation2 v1.201 (Kofler *et al.* 2011b). The base pair level analysis informs on specific changes and possible selection events, while the window approach informs on processes affecting a larger region. For the base pair level analysis, the generated alignments from the two populations were converted into an mpileup file, and indels were masked with 5 bp windows in both populations, effectively excluding these bases from the analysis. Allele frequency differences were calculated, and F_{ST} was calculated from the allele frequencies using the standard equation as shown in Hartl and Clark (2007), and individual SNPs were considered for downstream analysis where Fishers Exact Test $p < 5 \times 10^{-8}$, correcting for multiple testing as the genome size is approximately 5×10^8 . We avoided population-specific mapping and population-specific copy number variation by only looking at regions where the read depth was above 20 (as there was a bimodal distribution of RD, as shown in Figure S1), and we only looked at regions having an RD below 47 (97.5th percentile), as this removes regions that have a significantly higher than average RD (indicating possible merging of similar genes, or population-specific copy number variation). Individual SNPs were determined to be in an exon, an intron, or intergenic, by intersecting the SNP

223 coordinates with the GFF output of MESPA using bedtools-2.18.2 (Quinlan & Hall 2010).
 224 Window-based divergence was calculated in 5 kbp (kilobase pair), with a step size of 5 kbp,
 225 resulting in consecutive non-overlapping windows that covered the entire genome, using the
 226 input file generated above. Finally, windows were kept for subsequent analyses when >50%
 227 of the window had an RD between 20-47 and Fisher's Exact Test $p < 5 \times 10^{-8}$.

228 **Nucleotide diversity and Tajima's D.** To assess signals of selection on regions of interest,
 229 nucleotide diversity (π) and Tajima's D was calculated for both populations using
 230 PoPoolation v1.2.2 (Kofler *et al.* 2011a) in 5 kbp windows. The alignments of the two
 231 populations were converted into individual pileup files, in which indels were masked with 5
 232 bp windows. For Tajima's D, the pileup files were subsampled to a depth of 20 without
 233 replacement. Both statistics were calculated in 5 kbp consecutive non-overlapping windows,
 234 and only windows that were considered in the FST-based analysis were used for further
 235 analysis. Nucleotide diversity is a measure of the degree of polymorphism in a population.
 236 Patterns of low or high nucleotide diversity can arise through selection, but also other
 237 factors, such as demographics and inbreeding. It is best used here when related to other
 238 population genetic measures, such as FST, and Tajima's D. As used here, Tajima's D is a test
 239 of selection. We assessed whether Tajima's D values of outlier regions were the same as the
 240 genome wide background, indicating no evidence of selection, or deviated from the genome
 241 wide background, indicating a possible selective sweep or population expansion when
 242 significantly more negative, and alternatively balancing selection or sudden population
 243 contraction when significantly more positive.

244 **Outlier selection.** To look for candidate regions having experienced strong local selection, a
 245 set of outlier regions was chosen from the windows significant for the Fisher's exact test for

subsequent analysis, when conforming to at least one of four separate criteria. We used these four different criteria, in order to reduce the bias of only investigating a particular type of signal. First, in the top 99.5% percentile of the 5kbp window analysis the six contigs with the highest number of windows were selected. Second, the nine contigs containing the highest window-based FST values were selected. Third, the contigs with the highest number of individual fixed SNPs were selected. Finally, contigs containing fixed nonsynonymous substitutions were selected.

HRM validation. High Resolution Melting (HRM) was used on each individual of the Pool-SEQ libraries to validate the allele frequencies in several candidate genes that were identified in the genome scan. Primers were designed using Primer Express 3.0.1 (Applied Biosystems) using default settings. The SNP with the highest FST in each outlier region was selected, and to make sure downstream HRM results were as unambiguous as possible, Primers were located +/-25 bp before and after the SNP of interest containing no other polymorphic sites in the region. For the four genes showing non-synonymous substitutions, primers were designed around that particular SNP. Table S7 lists the generated primers. The HRM reaction, consisting of 5 µl MeltDoctor™ HRM Master Mix (Thermo Fisher Scientific, cat no: 4415440), 2.8 µl ddH2O, 0.6 µl 5µM of the forward and 0.6 µl 5µM reverse primer, and 1 µl of 5ng/µl DNA were added to each well of a 96-well reaction plate. On each plate, at least 3 known northern homozygote, 3 known southern homozygote and 3 known heterozygote samples were used as controls. In a Veriti® 96-Well Fast Thermal Cycler, the reaction was heat activated at 95°C for 10 min, amplified at 95°C for 15 sec and 60°C for 1 min for 40 cycles. After this, a melt curve was generated on a StepOnePlus™ Real-Time PCR System by heating the reaction to 95°C for 10 sec, and 60°C for 1 min after which temperature was

269 increased to 95°C at 1% ramp speed, measuring fluorescence levels at every 0.1°C increase.

270 Raw melt curves were analysed using High Resolution Melt Software v3.0.1 (Applied

271 Biosystems). Markers were placed at each tail end of the derivative melt curve, then the

272 software estimated allelic states by associating the melt curves of the samples to the

273 controls. In <2.8% of all samples the calls of the program were inconclusive, but visual

274 inspection clearly revealed the most probable allelic state, and these were called manually.

275 Ambiguous reactions were performed again only once. Samples were classified as being

276 either homozygous for the southern allele (SS), heterozygote (SN), or homozygous for the

277 northern allele (NN) for each locus. We validated the allele frequencies in the Pool-Seq data

278 sets, and even though these data sets were generated using a sample size considered “low”

279 for pool-seq (n=22 individuals), the HRM validation confirmed a 93% overlap between the

280 allele frequency estimates and the true genotypes of each individual (Table S8).

281 **Phenotyping crosses.** In order to investigate the inheritance of diapause, we produced

282 population crosses that were assessed for their diapause incidence (Figure S3). In 2012 six

283 mated females from southern Sweden (Ålstorp (55.82 N, 12.97E, hereafter denoted S) and

284 16 mated females from a northern population in the Sundsvall area (Alnö 62.41N, 17.46E,

285 hereafter denoted N) were collected, and their offspring were reared individually in 0.5 l

286 plastic cups on the grass *Poa annua*. They were given a photoperiod that leads to 100%

287 direct development in both populations (L:D 21:3). The resulting adults (85 N and 61 S

288 butterflies) were used to produce F1 offspring for the two populations SS and NN, as well as

289 reciprocal F1 population hybrids NS and SN (female first). F1 offspring were assayed for the

290 proportion of direct development in conditions where the SS population shows 100% direct

291 development and the NN population shows 0%, i.e. L:D 18:6 at 17°C (Aalberg Haugen &

Gotthard 2015). A second set of larvae were reared in conditions that produce 100% diapause development in both populations: L:D 13:11, 17°C. The resulting diapausing pupae were overwintered in outdoor conditions.

In spring 2013 a male informative backcross SSxNS, a female informative backcross NSxSS, as well as one hybrid cross SNxNS were produced with these overwintered individuals. The resulting offspring were assayed for the proportion of direct development in the same intermediate photoperiod of L:D 18:6 and 17°C as used for the F1 cross. Larvae were reared in family cages with ad libitum access to the grass *Dactylis glomerata* in three climate rooms at Tovetorp Zoological Research Station. Cages were monitored daily for newly formed pupae. Each individual was sexed and weighed at pupation, scored for larval development time, subsequently frozen in liquid nitrogen, and stored in -80°C. In all rearing experiments family structure, defined as all offspring from a mated female, was recorded for each individual. Several previous studies of *P. aegeria* show that non-diapause development (i.e direct development) is strongly associated with very short larval development times whereas individuals entering diapause have considerably longer larval periods (Nylin *et al.* 1989; Aalberg Haugen *et al.* 2012; Aalberg Haugen & Gotthard 2015). This prolonged larval development time leads to pupation later in the season when photoperiods are shorter and the risks of entering direct developmental is lower. In fact, the reduction of larval development rate in individuals that do not develop directly is a central component of the induction of pupal diapause in this species (Nylin *et al.* 1989; Aalberg Haugen *et al.* 2012; Aalberg Haugen & Gotthard 2015). Therefore, in order to increase the size of the experiment we used larval development time as a proxy of direct or diapause development, respectively.

315 Indeed, also in this study larval development time showed a clear bimodal distribution in all
316 crosses except the backcross (Figure S4), where larvae in the first modus were classified as
317 direct development (<40 days), and larvae in the second modus (>40 days) as diapause
318 development. For the pure S and N crosses we recorded actual developmental pathways,
319 and confirmed this classification.

320 **SNP genotyping in the crosses.** In order to reveal the association between the candidate loci
321 and diapause induction, genotypes were assessed using HRM in the F1 and F2 crosses
322 generated in the constant conditions L:D 18:6, 17°C. 115 individuals were randomly selected
323 from the tail ends of the developmental distribution of the reciprocal F1 hybrid crosses, with
324 61 individuals from SN cross and 54 individuals from NS cross. Moreover, 65 individuals of
325 the F2 SNxNS cross were selected. For the backcross SSxNS, 139 individuals were randomly
326 selected from the tail ends of the distribution of developmental times, excluding animals
327 that were 40 +/- 5 days. DNA was extracted from 4 mm of the posterior end of each pupa
328 using an E-Z 96® Tissue DNA Kit (Omega Bio-tek, cat no: D1196-01), including RNase A
329 treatment. HRM results were analysed in PLINK v1.07 for the F2 hybrid cross (Purcell *et al.*
330 2007). Allelic states were given (1,1 for SS; 1,2 for heterozygotes; and 2,2 for NN). As
331 virtually all individuals in the family were sampled, each locus was tested against diapause
332 incidence using a logistic model and run using a family structure and 100,000 permutations
333 using gene-dropping and label-swapping. The association test compares allele frequencies
334 between the direct and diapause groups, and this distribution is empirically tested using the
335 permutations. The permutation-adjusted p-values were used to infer significance. HRM
336 results of the F1 hybrid females and backcrosses were analysed in JMP 12.0.1 (SAS) using a
337 generalized linear mixed model (GLMM) assuming a binary response for larval development

(diapause >40 days, or direct <40 days), and a logit link function using Firth Bias-adjusted Estimates. GLMM model selection was performed using stepwise backward elimination, starting with all five candidates and sex as dependent variables. A GLMM was performed here instead of a permutation test, since the individuals were chosen from the tail ends of the distribution, rather than sampling the full crosses.

SNP genotyping in a cline. The candidate loci were also assessed using the same HRM protocol in three additional Swedish populations to investigate any possible latitudinal patterns. The DNA of these samples is the same as was used in a microsatellite study of population variation (Tison *et al.* 2014), where all individuals of these populations were collected in 2010 and 2011. Sample numbers below refer to those used in the previous study: 24-39 for Öland, 41-56 for Gotland, and 64-79 for Stockholm (N = 16 per population).

F1 hybrid Pool-Seq. In order to investigate whether any loci that were not tested in the HRM genotyping, were segregating between diapausing and directly developing F1 hybrids, in the absence of a hemizygous Z-chromosome background, a genome scan was performed on F1 hybrid males. DNA was extracted from the thoraces of 22 directly developed, and 16 diapause generation adult F1 NS males, pooled at equal concentrations, and sequenced as described above. Read mapping, variant calling, and FST calculation in 5kb windows were performed as described for the pure populations above, with the amendment that RD filtering was done between RD 20-70, corresponding to the 97.5% percentile for these data sets.

RESULTS

Genome assembly

We first generated a 501 Mbp draft genome for *P. aegeria* with an N50 of 15,077. This was then scaffolded and annotated using the proteome from the monarch butterfly, *D. plexippus*; (Zhan & Reppert 2013), increasing the N50 to 32,208, and producing 11,538 gene models (78% of the *D. plexippus* genes; Table S1 and S2). BUSCO analysis revealed 66.1% complete arthropod BUSCO's, placing the *P. aegeria* assembly second-highest of a total of 13 other published Lepidopteran genomes (Talla *et al.* 2017). We then placed >90% of the contigs containing gene models, encompassing 168.7 Mbp of the assembly, in a chromosomal context by exploiting the high level of gene-order conservation among Lepidoptera, using the chromosome level assembly of the silk moth *B. mori* (Suetsugu *et al.* 2013).

Genomic differentiation between populations

Assessing genomic divergence between two populations allows us to identify regions of differentiation as candidates for local adaptation, which we can query for associations with variation in diapause induction. A pooled sequencing approach was used to quantify the genomic variation in two Swedish populations that differ significantly in their CPP, Skåne (southern) and Sundsvall (northern) (Aalberg Haugen & Gotthard 2015). This identified 116,809 SNPs (Figure S5). Genome-wide differentiation between the populations was low, with a mean F_{ST} of 0.085 (Figure 1A). Divergent regions ($F_{ST} > 0.9$) were found to be scattered across the genome, with a subset of ten regions showing an aggregation of F_{ST} outliers, each representing >20 kbp in length (Table S4). Note that the identification of the F_{ST} outliers does not rely upon the synteny analysis, which is only used here to assess whether or not outliers appear in close proximity to each other or are spread over the

genome. Gene level analysis identified four genes with nonsynonymous substitutions from the SNP analysis that were highly differentiated ($F_{ST} > 0.9$), reflecting potential functional divergence (Table S5). While the low genome-wide differentiation between these populations reflects their shared recent ancestry or ongoing gene flow (Kawecki & Ebert 2004), regions exhibiting high levels of divergence could arise due to methodical artifacts (Schlötterer *et al.* 2014), as well as neutral or selective evolutionary events. We selected 15 candidate loci for further testing based on four different criteria, which were related to the number of co-occurring windows with high F_{ST} values, absolute F_{ST} values, number of SNPs, and genes with fixed nonsynonymous substitutions. This was done in order to reduce the bias of only investigating a particular type of signal (Table S6). For validation of these divergent regions, we genotyped a representative SNP for each of 15 candidate loci showing the highest differentiation (Table S6, and Table S7), using all individuals included in the pooled data sets, finding concordance for 93% of the allele frequencies observed in the Pool-Seq analysis (Table S8).

Association between candidate loci and diapause induction

In order to test whether there was a causal relationship between the 15 outlier loci and diapause incidence, we evaluated the inheritance of diapause by generating an F2 cross between the two populations. These crosses were assayed in a controlled environment of 17°C and L:D 18:6, in which the northern population showed a near 100% diapause incidence, while the southern population showed <2% (Table 1, Figure S4). The crosses revealed female-specific inheritance of CPP (Table 1), as there is a difference in diapause incidence between the females and the males of this cross, even though these carry the

same autosomal background (Figure S3). Females are the heterogametic sex in Lepidoptera with a single copy of the Z and W chromosome, while males are homogametic with two copies of the Z chromosome. Thus, the female specific response indicates a role of the Z-chromosome. We then tested the association with the identified candidate loci by genotyping each individual in the F2 cross.

There was a significant association with the fixed SNP near the gene *kinesin* ($p < 0.002$), as well as the non-synonymous substitution in the circadian clock gene *timeless* ($p < 0.002$), when analyzing both sexes together (Table S9). All individuals that developed directly were homozygous SS for both these two loci, while 80% of the females that were homozygous NN entered diapause, as well as 30% of the males that were homozygous NN, with the heterozygous group showing a diapause incidence of 10-30% (Figure 1B, Figure S6). Synteny analysis with *B. mori* suggest both these genes were located on chromosome 4 (Figure 1C). Investigating the sexes separately showed additional significant loci in females (Table S10), but not males (Table S11). In females, there was a marginally significant association with the nonsynonymous substitution in the Z-linked circadian clock gene *period* ($p = 0.053$), as well as a significant association with the gene *carnitine O-acetyltransferase* (CROT; $p = 0.013$; Table S10, Figure 1B). Synteny analysis showed the gene *CROT* to also be located on chromosome 4 (Figure 1C).

Selection dynamics in candidate genes

We investigated the nucleotide diversity, Tajima's D and the number of polymorphisms in the significant candidate loci and the surrounding regions to assess signatures of selection.

430 The gene *carnitine O-acetyltransferase (CROT)* showed one amino acid polymorphism, while
431 the genes *period* and *timeless* showed multiple amino acid variants (Table S5). The genes
432 *kinesin*, *CROT* and *timeless* are in close proximity in *D. plexippus* and *B. mori*. This region,
433 measured as the distance between the two most distant loci, *kinesin* and *timeless*, was 81
434 kbp in *D. plexippus* and 244 kbp in *B. mori*, and at least 130 kbp in our *P. aegeria* genome
435 (Figure 1C). Within this region, there was a reduction of nucleotide diversity that is similar in
436 both populations (S mean = 0.0023; N mean: 0.0019) with both being below the 10th
437 percentile of the genome wide distribution of nucleotide diversity (S mean = 0.005, SD =
438 0.003; N mean = 0.004, SD = 0.002), although nucleotide diversity appeared more variable at
439 *kinesin* for the southern population. Average Tajima's D for this region (S = -2.34; N = -2.50)
440 was at the 3th percentile of the genome wide distribution in both populations (S mean = -
441 1.43, SD = 0.60; N mean = -1.56, SD = 0.65), indicating an excess of rare alleles. Additionally,
442 this region harbored 64 fixed SNPs between the N and S populations (31 intergenic, 29
443 intronic, and 4 exonic nonsynonymous substitutions: 1 in *kinesin*, 1 in *CROT*, 2 in *timeless*).
444 At the *period* locus, the window of reduced nucleotide diversity was smaller at 50 kbp, with
445 a mean (S = 0.0016; N = 0.0015) that was within the 2.5th percentile of the genome wide
446 distribution in both populations. (Figure 1D), and with a mean Tajima's D (S = -2.31; N = -
447 2.41) at the 5th percentile of the genome wide distribution of Tajima's D of both populations,
448 indicating an excess of rare alleles. Within this region there were 5 fixed SNPs between the N
449 and S populations (1 intergenic, 2 intronic, and 2 exonic nonsynonymous substitutions).
450 Thus, the size of the region of reduced diversity was nearly three times larger in
451 *kinesin/CROT/timeless* region compared to that at *period*, and relative density of fixed SNPs

between the N and S populations was higher in the *kinesin/CROT/timeless* region with 5 SNPs per 10 kb, compared to 1 SNP per 10 kb at *period*.

Latitudinal cline of candidate genes

We further investigated the shift in allele frequencies using the two previous populations studied above, and an additional three populations along a Swedish latitudinal cline that show variation in CPP. In the majority of the investigated genes we revealed the presence of heterozygous individuals in several of the populations. In contrast, there was an abrupt shift in allele frequencies at *timeless* and *period*, from both being N homozygotes to both being S homozygotes, with an absence of heterozygous individuals in the cline (Figure 2). These shifts in allele frequencies were not completely matched between the two loci across the entire cline, as the Öland population was fixed for the southern allele at the *timeless* locus while being fixed for the northern allele at the *period* locus. Furthermore, the allele frequency shifts along the sampled cline do not fully match the gradient in CPP in the remaining populations. This is observed in the three northern populations, which exhibit northern genotypes for both loci but have clear differences in CPP as estimated in a previous study (Aalberg Haugen & Gotthard 2015). This incongruence between the binary pattern of the *timeless* and *period* genotypes in the wild populations and the continuous cline in CPP suggests the action of additional loci in determining local CPP thresholds.

Effect of other loci

In order to investigate the presence of additional loci on CPP, we investigated genomic effects on CPP using an F1 cross of the two populations, wherein fixed variants between

populations are in a heterozygous state. Thus, searching for genomic differences between diapausing and directly developing individuals in these F1 hybrids allows for the exploration of dominance or the effects of other loci not fixed between the populations. The F1's showed a sex-linked inheritance (Table 1). SNP genotyping confirmed this by revealing a highly significant association with Z-linked *period* ($p = 0.008$), reflecting that 90% of diapausing females carried only the N allele, compared to 44% of direct developing females (Figure S7, Table S12). Roughly 50% of males entered diapause (Table 1), indicating an absence of dominance effects at the previously tested loci. To identify whether any other loci exhibited associations with the induction decision, we scanned the genome for associations with diapause induction by generating two Pool-Seq datasets, one for the diapausing and one for direct developing group of males in the NxS F1 cross. The Pool-Seq analysis revealed low genome-wide differentiation ($F_{ST} = 0.031$; Figure S8), and an absence of even moderately differentiated outliers (highest $F_{ST} = 0.18$). F_{ST} values for the individual SNPs that were significant in the genotyping of the F1 and F2 crosses were <0.017 for *period*, <0.025 for *kinesin*, <0.012 for *CROT*, and <0.036 for *timeless*, reflecting that all individuals in this pool-seq comparison were heterozygous for these loci.

Next, in order to investigate the possibility of additional population specific loci, we generated an F2-back cross, wherein F1 hybrids of both sexes were crossed with a southern population individual. We expected loci with homozygous southern alleles to have a larger effect than loci with northern alleles in heterozygous states based upon our F1 cross results. Consistent with this, we found a low diapause incidence, with an effect of the predominantly southern background of the individuals in both crosses (Table 1). However, we were

primarily interested in the hemizygous Z-chromosome effect in females, since if additional loci in the autosome exist that are specific to the southern population, we expect a much lower proportion of diapause for females having a northern Z-chromosome than in previous crosses. To test this, we genotyped the Z-linked locus *period*, and autosomal loci *timeless*, *CROT*, *vasa*, and *trypsin* in the SSxNS backcross, wherein females had either a S or a N Z-chromosome. While there was a significant effect of sex, there were no significant associations with diapause at any of the tested loci in this backcross (Table S13 & S14).

DISCUSSION

Here, we assessed the genetic variation underlying clinal variation in CPP in wild populations of the butterfly *P. aegeria*, as well as the selection dynamics in candidate regions that associated with the diapause induction decision. We found genetic differentiation and evidence of local adaptation between populations that differ in their photoperiodic thresholds for diapause induction. We showed that the inheritance of CPP in crosses of these populations arises from both a sex-linked and autosomal component. Importantly, we were able to significantly associate several loci with this diapause induction decision on both the sex-chromosome and autosome, these represent the sex-linked gene *period* and the autosomal *kinesin/CROT/timeless* region, likely at the gene *timeless*. These gene regions appeared to have a large effect size on the diapause incidence, such as an up to 80% difference in diapause incidence in individuals carrying northern versus southern alleles of *timeless*. Interestingly, along a sampling of the larger cline, heterozygotes at these large effect loci of *period* and *timeless* were absent. However, the allele frequencies of these loci did not fully match the latitudinal gradient in CPP, indicating the action of additional loci in

determining local CPP thresholds. Consistent with this, analysis of males in an F1 cross suggested additional loci with small effect sizes. In this cross all individuals were heterozygous at the two major effect gene regions and we could not identify any dominance effects at these two loci. Consequently, the lack of high F_{ST} outliers between the diapause and direct development groups of this F1 cross suggests that any additional genetic factors influencing diapause induction have effects that are smaller than we can detect in this study. This is further corroborated by the results of the backcross where there was no measured effect of the candidate loci when the genetic background is predominantly southern (~75% on average). This indicates that other loci are involved and that the combined effects of these other loci override the candidate loci and produces a predominately southern response in the backcross. In conclusion, we found evidence of both loci of large and small effect-size underlying latitudinal variation in CPP, consistent with an exponential model of local adaptation (Orr 2000; Kawecki & Ebert 2004).

One potential bias in this study arises from using Pool-Seq data to assemble a genome, as heterozygosity is usually a problem during assembly, and will be higher than when using only sequencing data from a single individual. This can lead to inflated genome sizes and duplications of genes that do not reflect biological reality. However, gene completeness was similar to other published genomes, and there was only limited population-specific mapping, as well as a high linear correlation in read depth between the two populations. Other studies report similar success in assembling genomes from Pool-Seq data (Neethiraj *et al.* 2017). Our Pool-seq approach also suffers from the fact that we used only 22 individuals for each population, even though it is designed to calculate allele frequencies from a large number of

individuals to achieve accurate estimates (Schlötterer *et al.* 2014). As the allele frequencies are used to calculate nucleotide diversity and Tajima's D, these statistics will inherently be affected by this. While using larger window sizes in the analysis can mitigate this problem by averaging out errors, and thus generating informative results, we must remain conservative in our estimates of the effect sizes. Furthermore, as demonstrated by the genotyping of the individuals that make up our pools, overall allele frequencies generated by the Pool-Seq pipeline are in accordance with the genotypes of each individual, for the tested loci. The independently generated F1 Pool-Seq analysis further corroborates the allele frequencies observed in the population samples. This likely reflects quality and quantity control on the DNA, as well as data filtering. Next, each individual sequence is an independent draw of the pool of individuals, which impedes the ability to infer haplotypes and linkage disequilibrium (Schlötterer *et al.* 2014). We used the Pool-Seq analysis as a hypothesis generator to identify potential candidate genes, which we then genotyped for each individual in the population samples and crosses, thus inferring haplotype information for at least these specific loci that were tested.

The low level of differentiation found in this study between the southern and northern populations is in concordance with previous microsatellite work on these populations (Tison *et al.* 2014). In contrast to the overall low levels of differentiation, we found a number of genomic regions with high divergence between the two populations. However, these peaks do not necessarily represent local adaptation of CPP between populations, as outlier regions may arise due to local adaptation of other traits, and numerous non-adaptive mechanisms.

To assess this further, we investigated a F2 hybrid cross of these populations, and were able to significantly associate a subset of these divergent loci to the diapause induction decision.

The regions of divergence showing significant associations with diapause induction represent both the Z-chromosome and the autosome, consistent with the finding of sex-linked and autosomal components affecting diapause induction in several previous studies (Rockey *et al.* 1987; Hagen & Scriber 1989; Ikten *et al.* 2011; Chen *et al.* 2014; Fu *et al.* 2015; Pruisscher *et al.* 2017). The narrow regions of divergence we observed in the FST analysis identify several candidate genes. In the case of the *period* locus, it is the only gene present inside the boundaries of the 50 kbp region of low nucleotide diversity and high FST, and the amino acid polymorphism did not show any heterozygote individuals in the sampled cline, suggesting selection on the homozygous states of this locus. In the larger *kinesin/CROT/timeless* region there are several genes present, making it challenging to distinguish between selection acting upon one or several of these loci. However, the latitudinal samples show that there is recombination between the tested loci, as the allele frequencies of the *kinesin* and *CROT* genes are not identical, or linked, within each population. Furthermore, there is a lack of heterozygous individuals for *timeless*, suggesting that selection has acted at or near the *timeless* gene.

The size and pattern of nucleotide diversity and negative Tajima's D values at the large effect loci of *period* and *kinesin/CROT/timeless* are consistent with a signature of a selective sweep. A negative Tajima's D strictly indicates that rare alleles are present at a higher than expected frequency, which can result from a recent selective sweep, linkage to a sweep, but also an

expansion of the population after a recent bottleneck. Interpreting the patterns of variation at the two candidate loci, *period* and *kinesin/CROT/timeless* is complicated by several unknowns (per locus selection coefficients, recombination dynamics, and effective population sizes). As a starting point for investigating these dynamics, if we make the simplifying assumption that these unknowns are equal between the loci, then the alleles at *period* appear to be much younger than those segregating at the *kinesin/CROT/timeless* locus, as the latter are much more divergent and the former has a lower N_e . To confirm this, further empirical evidence or population genetics modeling would be required, ideally in combination with results from other systems.

Across the cline, there is evidence for recombination in the *kinesin/CROT/timeless* locus. The genes *kinesin* and *CROT* exhibit a gradual shift in allele frequency, with only *timeless* exhibiting a dramatic switch between the northern and southern homozygotes without any heterozygotes present in the cline. This suggests selection against heterozygotes at *timeless*, but with free recombination between northern and southern alleles in the *kinesin/CROT* region across the geographic cline. Consistent with this, we see higher nucleotide diversity at *kinesin* than *CROT*, which is in line with the former being more distant from *timeless*.

While the F2 hybrid crosses identify large effect loci, these loci neither fully control the diapause induction decision across populations, nor is their magnitude of effect consistent across populations. This is evident in several comparisons of the genotypes at these loci along our sampled cline for which CPP is well characterized. First, while the change in CPP among three northernmost populations is nearly two hours, all of these three populations

have identical allele frequencies at *period* and *timeless*, and nearly identical frequencies at the neighboring loci of *timeless* (*kinesin*, *CROT*) (Figure 2). Second, despite minimal difference in CPP between the population from Öland and Gotland, there is a complete switch between homozygous southern and homozygous northern *timeless*, no change at *period*, and a moderate shift in allele frequency at *kinesin* and *CROT*. Finally, the most striking difference in CPP occurs between Gotland and Stockholm with 1.4 hours, where *period* and *timeless* show identical genotypes and *kinesin* and *CROT* only show minimal change in allele frequency. Thus, neither variation at the *period* nor *timeless* locus alone fully controls local CPP, yet in distant population crosses they have large effects. Therefore, the response of CPP to selection over a latitudinal gradient appears to involve the loci investigated here at a relatively large geographic scale, but a significant part of the finer scale response must also be due to variation segregating at other loci. This appears similar to what is found in *D. melanogaster*, where different *timeless* variants affected photoperiodically induced quiescence, which was modified by the genetic background (Tauber *et al.* 2007). To what extent our other identified outlier loci are involved in these local decisions remains to be determined.

The F1 hybrids and backcrosses further support the notion of additional loci acting on diapause induction. In the F1 hybrids there is a strong sex-linked effect at the *period* locus, though in this cross *period* is a marker for an entire N or S Z-chromosome. In contrast, hybrid males were heterozygous at all loci that were fixed between the populations, and did not show a particular genomic region that was strongly associated with the difference in diapause induction. This suggests that when the candidate loci are heterozygous, no

additional loci with large additive or dominant effects act upon the diapause decision. In the backcrosses, there was no significant Z-linked effect of *period*, further suggesting the presence of additional loci in the southern genomic background masking the effect of the northern *period*. Importantly, sample sizes here were larger than previous crosses, indicating that statistical power was sufficiently high. Thus, while in our previous F1 crosses, there was a Z-chromosome effect in females (Table 1), the absence of this effect in this backcross indicates the presence of additional loci in the Southern population influencing diapause induction. An alternative explanation could be that the photoperiod used is too close to the direct development threshold of this backcross ‘population’, as the genetic background is 75% S and 25% N, the southern alleles may mask the effect of other genes. The diapause incidences in this cross will most likely look very different when tested in other photoperiods. In combination, these findings suggest that the genomic architecture of local adaptation of CPP is variable and context dependent. While we discovered a number of loci with larger effect sizes, our results also suggest that there is considerable additional genetic variation present within and among populations that affect the CPP. Therefore, our study implies that the genetic architecture of CPP in *P. aegeria* should facilitate an evolutionary response to seasonal changes, such as predicted by climate change, should these changes be consistent for sufficiently long periods of time. Rapid evolutionary responses of diapause to changing environmental conditions have been observed in other species (Bradshaw & Holzapfel 2001).

While there have been targeted candidate gene studies that revealed differences in diapause induction and the circadian clock genes (Tauber *et al.* 2007), it is striking that

performing a genome-wide scan for genetic variation that influence CPP also revealed that allelic variation at *period* and *timeless* had the most significant effects. The products of these two genes directly interact with each other as part of the circadian clock by forming a heterodimer that together with *mammalian-type cryptochrome* affect the transcription of clock components (Yuan *et al.* 2007; Goto 2012). Studies have tested the association between photoperiodic induction of diapause and the circadian clock, predominantly due to the influence of the controversial Bünning hypothesis, which posits that photoperiodic timing is linked to the circadian timing mechanism (Bünning 1936). One notable example in *Drosophila melanogaster* tested, and supports a link between photoperiodic measurement and different alleles of circadian clock genes (Pegoraro *et al.* 2014). Still, the overall pattern is inconclusive as several studies indicate an association between the clock and diapause (Ikeno *et al.* 2010; Yamada & Yamamoto 2011; Meuti *et al.* 2015), while others have revealed that photoperiodism and the circadian clock are independent (Emerson *et al.* 2009).

It is possible that the amino acid changes found in the *period* and *timeless* genes affect the stability of their interaction with each other, which would affect the circadian clock by potentially slowing down or speeding up the circadian clock. Even if this were true, whether it is the circadian clock as a unit that is involved in photoperiodic induction in *P. aegeria* or the pleiotropic function of the clock genes remains an open question. Functional validation of the effects arising from the natural allelic variation observed here is challenging, as RNAi-knockdown and CRISPR/Cas-knockouts of circadian clock genes often have deleterious or uninformative phenotypes for such questions. Here, by integrating whole-genome scans for

local adaptation with phenotyping and targeted genotyping in population crosses, we demonstrate that loci underlying adaptive variation in seasonal life cycle regulation conforms to an exponential model of effect sizes, and may be strongly influenced by allelic variation at few genes with known functions in the circadian clock. Gaining such insights into the genetic variation underlying clines of diapause induction from a larger number of species will significantly advance our understanding of the evolutionary history and dynamics of this complex phenotype.

ACKNOWLEDGEMENTS

The authors would like to acknowledge support from Science for Life Laboratory, the National Genomics Infrastructure, NGI, and Uppmax for providing assistance in massive parallel sequencing and computational infrastructure. Funding was provided by the Swedish Research Council (VR grant numbers 2012-3715, 2010-5341, 621-2012-4001), Academy of Finland (grant number 131155), the Bolin Centre for Climate Research at Stockholm University, and the Knut and Alice Wallenberg Foundation (grant number 2012.0058). We thank Philipp Lehmann for providing comments on earlier versions of this manuscript. We are very grateful for the hard work of research assistant Beatrice Svensson in rearing and phenotyping the crosses. The authors declare there are no conflicts of interest.

REFERENCES

- Aalberg Haugen IM, Gotthard K (2015) Diapause induction and relaxed selection on alternative developmental pathways in a butterfly. *Journal of Animal Ecology*, **84**, 464–472.
- Aalberg Haugen IM, Berger D, Gotthard K (2012) The evolution of alternative developmental pathways: footprints of selection on life-history traits in a butterfly. *Journal of Evolutionary Biology*, **25**, 1377–1388.

707 Ahola V, Lehtonen R, Somervuo P *et al.* (2014) The Glanville fritillary genome retains an
 708 ancient karyotype and reveals selective chromosomal fusions in Lepidoptera. *Nature*
 709 *Communications*, **5**, 4737.
 710 Barton NH, Keightley PD (2002) Understanding quantitative genetic variation. *Nature*
 711 *reviews. Genetics*, **3**, 11–21.
 712 Bradshaw WE (1976) Geography of photoperiodic response in diapausing mosquito. *Nature*,
 713 **262**, 384–386.
 714 Bradshaw WE, Holzapfel CM (2001) Genetic shift in photoperiodic response correlated with
 715 global warming. *Proceedings of the National Academy of Sciences*, **98**, 14509–14511.
 716 Bradshaw WE, Emerson KJ, Catchen JM, Cresko WA, Holzapfel CM (2012) Footprints in time:
 717 comparative quantitative trait loci mapping of the pitcher-plant mosquito, *Wyeomyia*
 718 *smithii*. *Proceedings. Biological sciences / The Royal Society*, **279**, 4551–4558.
 719 Bradshaw WE, Quebodeaux MC, Holzapfel CM (2003) Circadian Rhythmicity and
 720 Photoperiodism in the Pitcher-Plant Mosquito: Adaptive Response to the Photic
 721 Environment or Correlated Response to the Seasonal Environment? *The American*
 722 *Naturalist*, **161**, 735–748.
 723 Bünning E (1936) Die endonome Tagesrhythmik als Grundlage der photoperiodischen
 724 Reaktion. *Berichte der Deutschen Botanischen Gesellschaft*, **54**, 590–607.
 725 Camacho C, Coulouris G, Avagyan V *et al.* (2009) BLAST+: architecture and applications. *BMC*
 726 *Bioinformatics*, **10**, 421.
 727 Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool
 728 set for population genomics. *Molecular Ecology*, **22**, 3124–3140.
 729 Chen C, Xiao L, He HM, Xu J, Xue FS (2014) A genetic analysis of diapause in crosses of a
 730 southern and a northern strain of the cabbage beetle *Colaphellus bowringi* (Coleoptera:
 731 chrysomelidae). *Bulletin of Entomological Research*, **104**, 586–591.
 732 Danilevskii AS (1965) *Photoperiodism and seasonal development of insects*. Oliver and Boyd.
 733 Dasmahapatra KK, Walters JR, Briscoe AD *et al.* (2012) Butterfly genome reveals
 734 promiscuous exchange of mimicry adaptations among species. *Nature*, **487**, 94–98.
 735 Demont M, Blanckenhorn WU (2008) Genetic differentiation in diapause response along a
 736 latitudinal cline in European yellow dung fly populations. *Ecological Entomology*, **33**,
 737 197–201.
 738 Doležel D, Vaněčková H, Šauman I, Hodkova M (2005) Is period gene causally involved in the
 739 photoperiodic regulation of reproductive diapause in the linden bug, *Pyrrhocoris*
 740 *apterus*? *Journal of Insect Physiology*, **51**, 655–659.
 741 Emerson KJ, Dake SJ, Bradshaw WE, Holzapfel CM (2009) Evolution of photoperiodic time
 742 measurement is independent of the circadian clock in the pitcher-plant mosquito,
 743 *Wyeomyia smithii*. *Journal of Comparative Physiology A*, **195**, 385–391.
 744 Friberg M, Aalberg Haugen IM, Dahlerus J, Gotthard K, Wiklund C (2011) Asymmetric life-
 745 history decision-making in butterfly larvae. *Oecologia*, **165**, 301–310.
 746 Fu S, Chen C, Xiao L, He H, Xue F (2015) Inheritance of Diapause in Crosses between the
 747 Northernmost and the Southernmost Strains of the Asian Corn Borer *Ostrinia furnacalis*.
 748 *PLoS One*, **10**, e0118186.
 749 Goto SG (2012) Roles of circadian clock genes in insect photoperiodism. *Entomological*
 750 *Science*, **16**, 1–16.

751 Hagen RH, Scriber JM (1989) Sex-Linked Diapause, Color, and Allozyme Loci in *Papilio*
752 *glaucus*: Linkage Analysis and Significance in a Hybrid Zone. *Journal of Heredity*, **80**, 179–
753 185.

754 Han B, Denlinger DL (2009) Mendelian inheritance of pupal diapause in the flesh fly,
755 *Sarcophaga bullata*. *The Journal of heredity*, **100**, 251–255.

756 Hard JJ, Bradshaw WE, Holzapfel CM (1993) Genetic coordination of demography and
757 phenology in the pitcher-plant mosquito, *Wyeomyia smithii*. *Journal of Evolutionary*
758 *Biology*, **6**, 707–723.

759 Huang L-L, Chen C, Xiao L *et al.* (2013) Geographical variation and inheritance of the
760 photoperiodic response controlling larval diapause in two distinct voltine ecotypes of
761 the Asian cornborer *Ostrinia furnacalis*. *Physiological Entomology*, **38**, 126–132.

762 Ikeno T, Tanaka SI, Numata H, Goto SG (2010) Photoperiodic diapause under the control of
763 circadian clock genes in an insect. *BMC biology*, **8**, 1.

764 Ikten C, Skoda SR, Hunt TE, Molina-Ochoa J, Foster JE (2011) Genetic Variation and
765 Inheritance of Diapause Induction in Two Distinct Voltine Ecotypes of *Ostrinia nubilalis*
766 (Lepidoptera: Crambidae). *Annals of the Entomological Society of America*, **104**, 567–
767 575.

768 Kato Y (2005) Geographic variation in photoperiodic response for the induction of pupal
769 diapause in the *Aristolochia*-feeding butterfly *Atrophaneura alcinous*. *Applied*
770 *Entomology and Zoology*, **40**, 347–350.

771 Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225–
772 1241.

773 Kersey PJ, Allen JE, Armean I *et al.* (2016) Ensembl Genomes 2016: more genomes, more
774 complexity. *Nucleic acids research*, **44**, D574–80.

775 Kim Y, Krafusur ES, Bailey TB, Zhao S (1995) Mode of inheritance of face fly diapause and its
776 correlation with other developmental traits. *Ecological Entomology*, **20**, 359–366.

777 Kofler R, Orozco-terWengel P, De Maio N *et al.* (2011a) PoPoolation: A Toolbox for
778 Population Genetic Analysis of Next Generation Sequencing Data from Pooled
779 Individuals (M Kayser, Ed.). *PLoS One*, **6**, e15925.

780 Kofler R, Pandey RV, Schlötterer C (2011b) PoPoolation2: identifying differentiation between
781 populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics*, **27**,
782 3435–3436.

783 Košťál V (2006) Eco-physiological phases of insect diapause. *Journal of Insect Physiology*, **52**,
784 113–127.

785 Košťál V, Štětina T, Poupardin R, Korbelová J, Bruce AW (2017) Conceptual framework of the
786 eco-physiological phases of insect diapause development justified by transcriptomic
787 profiling. *Proceedings of the National Academy of Sciences of the United States of*
788 *America*, **114**, 8532–8537.

789 Kurahashi H, Ohtaki T (1977) Crossing between nondiapausing and diapausing races of
790 *Sarcophaga peregrina*. *Experientia*, **33**, 186–187.

791 Lehmann P, Margus A, Lindström L (2016) Inheritance patterns of photoperiodic diapause
792 induction in *Leptinotarsa decemlineata*. *Physiological Entomology*, **41**, 218–223.

793 Levy RC, Kozak GM, Wadsworth CB, Coates BS, Dopman EB (2015) Explaining the sawtooth:
794 latitudinal periodicity in a circadian gene correlates with shifts in generation number.
795 *Journal of Evolutionary Biology*, **28**, 40–53.

796 Li H (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–
797 2079.

798 Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of
799 protein or nucleotide sequences. *Bioinformatics*, **22**, 1658–1659.

800 McCoy JR, Lloyd EP, Bartlett AC (1968) Diapause in Crosses of a Laboratory and a Wild Strain
801 of Boll Weevils. *Journal of Economic Entomology*, **61**, 163–166.

802 McWatters HG, Saunders DS (1997) Inheritance of the photoperiodic response controlling
803 larval diapause in the blow fly, *Calliphora vicina*. *Journal of Insect Physiology*, **43**, 709–
804 717.

805 Meuti ME, Stone M, Ikeno T, Denlinger DL (2015) Functional circadian clock genes are
806 essential for the overwintering diapause of the Northern house mosquito, *Culex pipiens*.
807 *Journal of Experimental Biology*, **218**, 412–422.

808 Neethiraj R, Hornett EA, Hill J, Wheat CW (2017) Investigating the genomic basis of discrete
809 phenotypes using a Pool-Seq only approach: new insights into the genetics underlying
810 color variation in diverse taxa. *Molecular Ecology*.

811 Nylin S (1995) Life-cycle regulation and life history plasticity in the speckled wood butterfly:
812 are reaction norms predictable? *Biological Journal of the Linnean Society*, **55**, 143–157.

813 Nylin S, Wickman P-O, Wiklund C (1989) Seasonal plasticity in growth and development of
814 the speckled wood butterfly, *Pararge aegeria* (Satyrinae). *Biological Journal of the*
815 *Linnean Society*, **38**, 155–171.

816 Orr HA (2000) Adaptation and the cost of complexity. *Evolution*, **54**, 13–20.

817 Paolucci S, van de Zande L, Beukeboom LW (2013) Adaptive latitudinal cline of photoperiodic
818 diapause induction in the parasitoid *Nasonia vitripennis* in Europe. *Journal of*
819 *Evolutionary Biology*, **26**, 705–718.

820 Pegoraro M, Gesto JS, Kyriacou CP, Tauber E (2014) Role for circadian clock genes in seasonal
821 timing: testing the Bünnig hypothesis. (P Emery, Ed.). *PLoS genetics*, **10**, e1004603.

822 Perry TO (1971) Dormancy of Trees in Winter. *Science*, **171**, 29–36.

823 Pruisscher P, Larsdotter-Mellström H, Stefanescu C *et al.* (2017) Sex-linked inheritance of
824 diapause induction in the butterfly *Pieris napi*. *Physiological Entomology*, **84**, 464.

825 Purcell S, Neale B, Todd-Brown K *et al.* (2007) PLINK: a tool set for whole-genome
826 association and population-based linkage analyses. *American journal of human genetics*,
827 **81**, 559–575.

828 Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic
829 features. *Bioinformatics*, **26**, 841–842.

830 Rockey SJ, Hainze JH, Scriber JM (1987) Evidence of a sex-linked diapause response in *Papilio*
831 *glaucus* subspecies and their hybrids. *Physiological Entomology*, **12**, 181–184.

832 Roff DA (1996) The Evolution of Threshold Traits in Animals. *The Quarterly Review of Biology*,
833 **71**, 3–35.

834 Schlötterer C, Tobler R, Kofler R, Nolte V (2014) Sequencing pools of individuals - mining
835 genome-wide polymorphism data without big funding. *Nature reviews. Genetics*, **15**,
836 749–763.

837 Schmidt PS, Zhu C-T, Das J *et al.* (2008) An amino acid polymorphism in the couch potato
838 gene forms the basis for climatic adaptation in *Drosophila melanogaster*. *Proceedings of*
839 *the National Academy of Sciences of the United States of America*, **105**, 16207–16211.

- Sedlazeck FJ, Rescheneder P, Haeseler von A (2013) NextGenMap: fast and accurate read mapping in highly polymorphic genomes. *Bioinformatics*, **29**, 2790–2791.
- Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, **31**, 3210–3212.
- Simpson GG, Dean C (2002) Arabidopsis, the Rosetta stone of flowering time? *Science*, **296**, 285–289.
- Söderlind L, Nylin S (2011) Genetics of diapause in the comma butterfly *Polygonia c-album*. *Physiological Entomology*, **36**, 8–13.
- Suetsugu Y, Futahashi R, Kanamori H *et al.* (2013) Large scale full-length cDNA sequencing reveals a unique genomic landscape in a lepidopteran model insect, *Bombyx mori*. *G3: Genes/Genomes/Genetics*, **3**, 1481–1492.
- Suwa A, Gotoh T (2006) Geographic variation in diapause induction and mode of diapause inheritance in *Tetranychus pueraricola*. *Journal of Applied Entomology*, **130**, 329–335.
- Suzek BE, Wang Y, Huang H *et al.* (2014) UniRef Clusters: A comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics*, **31**, btu739–932.
- Talla V, Suh A, Kalsoom F *et al.* (2017) Rapid Increase in Genome Size as a Consequence of Transposable Element Hyperactivity in Wood-White (Leptidea) Butterflies. *Genome Biology and Evolution*, **9**, 2491–2505.
- Tauber E, Zordan M, Sandrelli F *et al.* (2007) Natural Selection Favors a Newly Derived timeless Allele in *Drosophila melanogaster*. *Science*, **316**, 1895–1898.
- Tauber MJ, Tauber CA (1976) Insect seasonality: diapause maintenance, termination, and postdiapause development. *Annual review of entomology*.
- Tison JL, Edmark VN, Sandoval Castellanos E *et al.* (2014) Signature of post-glacial expansion and genetic structure at the northern range limit of the speckled wood butterfly. *Biological Journal of the Linnean Society*, **113**, 136–148.
- Van Dyck H, Bonte D, Puls R, Gotthard K, Maes D (2014) The lost generation hypothesis: could climate change drive ectotherms into a developmental trap? *Oikos*, **124**, 54–61.
- Wang X-P, Yang Q-S, Dalin P *et al.* (2011) Geographic variation in photoperiodic diapause induction and diapause intensity in *Sericanus montelus* (Lepidoptera: Papilionidae). *Insect Science*, **19**, 295–302.
- Wheat CW, Wahlberg N (2012) Phylogenomic Insights into the Cambrian Explosion, the Colonization of Land and the Evolution of Flight in Arthropoda. *Systematic biology*, **62**, 93–109.
- Xia QW, Chen C, Tu XY, Yang HZ, Xue FS (2012) Inheritance of photoperiodic induction of larval diapause in the Asian corn borer *Ostrinia furnacalis*. *Physiological Entomology*, **37**, 185–191.
- Yamada H, Yamamoto M-T (2011) Association between Circadian Clock Genes and Diapause Incidence in *Drosophila triauraria* (PA Bartell, Ed.). *PLoS One*, **6**, e27493.
- Yuan Q, Metterville D, Briscoe AD, Reppert SM (2007) Insect cryptochromes: gene duplication and loss define diverse ways to construct insect circadian clocks. *Molecular Biology and Evolution*, **24**, 948–955.
- Zhan S, Reppert SM (2013) MonarchBase: the monarch butterfly genome database. *Nucleic acids research*, **41**, D758–63.

885 **DATA ACCESSIBILITY**

886 All illumina sequencing data will be made available in dryad after acceptance of the
887 manuscript.

888

889 **AUTHOR CONTRIBUTIONS**

890 P.P., S.N., K.G., and C.W.W. designed research; P.P., K.G., and C.W.W. performed research;
891 P.P. analyzed data; and P.P., K.G., and C.W.W. wrote the paper.

892

893

FIGURES

Figure 1 – Genome wide differentiation between the populations. A: Differentiation measured by F_{ST} in 5 kbp windows by chromosome and position, between the Swedish populations Skåne (southern) and Sundsvall (northern) of *P. aegeria*. Chromosomal position is inferred by synteny from *B. mori*, and chromosome 29 denotes sequences not placed in a chromosomal context. The larger red dots denote the 15 loci tested in the association analysis, and the grey horizontal line shows the 99.5 percentile of the window-based analysis. B: Association between diapause incidence and the genotypes of the candidate loci in the F2 hybrid analysis. Red solid lines and circles denote females, and blue dotted lines and squares denote males. Stars denote significance. C: Nucleotide diversity for Skåne (top, red line) and Sundsvall (middle, blue line) in 1 kbp windows, as well as F_{ST} (bottom, black line) measured in 1 kbp windows for the *kinesin/CROT/timeless* region on chromosome 4. Black squares indicate genes, and red squares indicate the genes significant in the association analysis. D: Nucleotide diversity for Skåne (top, red line) and Sundsvall (middle, blue line) in 1 kbp windows, as well as F_{ST} (bottom, black line) measured in 1 kbp windows for the *period* region on chromosome 1. Black squares indicate genes, and red squares indicate the genes significant in the association analysis.

Figure 2 – Genotype frequencies of the 15 candidate loci in 5 natural populations of *P. aegeria*. Loci are order by chromosomal location. CPP values were obtained from Aalberg Haugen & Gotthard (Aalberg Haugen & Gotthard 2015). The lightest shade of blue represents individuals that are homozygous northern, the darker shade of blue represents the heterozygote individuals, and the darkest blue shade represents individuals that are homozygous southern.

TABLES

Table 1 - Diapause incidence for the pure populations (P), F1 hybrids (F1), the F2 cross (F2), and the backcrosses (BC) assayed in L:D 18:6 hours, 17°C. Diapause incidence values are given by sex as well as calculated for the whole cross.

Cross ♀ x ♂	Diapause incidence (%)			Total (N)
	♀	♂	Combined	
P: SxS	2.1	0	1.6	93
P: NxN	97.8	100	98.9	91
F1: NxS	14.6	59.5	34.1	85
F1: SxN	45.4	50	48.1	77
F2: SNxNS	29	13.6	21.4	70
BC: NSxS	2	1	1.5	196
BC: SxNS	8.2	2.2	5.1	1017