JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.12745

Temperature- and sex-related effects of serine protease alleles on larval development in the Glanville fritillary butterfly

V. AHOLA*, P. KOSKINEN†, S. C. WONG*, J. KVIST†, L. PAULIN†, P. AUVINEN†, M. SAASTAMOINEN*, M. J. FRILANDER†, R. LEHTONEN*‡ & I. HANSKI*

*Department of Biosciences, University of Helsinki, Helsinki, Finland
†Institute of Biotechnology, University of Helsinki, Helsinki, Finland
‡Genome-Scale Biology Research Program & Institute of Biomedicine, University of Helsinki, Helsinki, Finland

Keywords:

Glanville fritillary;
heritability;
larval development;
larval growth rate;
plasticity;
reduction of variation;
RNA-seq;
serine proteases.

Abstract

The body reserves of adult Lepidoptera are accumulated during larval development. In the Glanville fritillary butterfly, larger body size increases female fecundity, but in males fast larval development and early eclosion, rather than large body size, increase mating success and hence fitness. Larval growth rate is highly heritable, but genetic variation associated with larval development is largely unknown. By comparing the Glanville fritillary population living in the Åland Islands in northern Europe with a population in Nantaizi in China, within the source of the post-glacial range expansion, we identified candidate genes with reduced variation in Åland, potentially affected by selection under cooler climatic conditions than in Nantaizi. We conducted an association study of larval growth traits by genotyping the extremes of phenotypic trait distributions for 23 SNPs in 10 genes. Three genes in clip-domain serine protease family were associated with larval growth rate, development time and pupal weight. Additive effects of two SNPs in the prophenoloxidase-activating proteinase-3 (*ProPO3*) gene, related to melanization, showed elevated growth rate in high temperature but reduced growth rate in moderate temperature. The allelic effects of the vitellin-degrading protease precursor gene on development time were opposite in the two sexes, one genotype being associated with long development time and heavy larvae in females but short development time in males. Sexually antagonistic selection is here evident in spite of sexual size dimorphism.

Introduction

Following the last glacial maximum, northern Europe was colonized by species that had been confined to refugia in southern Europe and Asia during the glacial period (Hewitt, 1999). The Glanville fritillary butterfly (*Melitaea cinxia*) has three distinct clades in Europe, corresponding to refugia in the Iberian peninsula, the Balkans and Asia (Wahlberg & Saccheri, 2007), of which the Asian clade is present in the Åland Islands in Finland (Hanski, 2011; Ojanen *et al.*, 2013) and other Baltic regions (Wahlberg & Saccheri, 2007; Somervuo *et al.*, 2014). During the post-glacial range expansion, species may have lost much of the genetic variation that was present in the source populations, depending on the speed of expansion, range of expansion and the size of the expanding population (Hewitt, 1996; Schmitt & Seitz, 2002). On the other hand, post-glacial range expansion has taken many taxa, the Glanville fritillary included, into cooler temperate climates than experienced by the source populations, which may have led to natural selection on temperature-related traits. The Glanville fritillary has colonized the Åland Islands in Finland approximately 200–400 years ago, where it now occurs at its northern range limit in Europe. During this time, the population is assumed to have become adapted to a highly fragmented habitat,

Correspondence: Virpi Ahola, Metapopulation Research Centre, Department of Biosciences, PO Box 65 (Viikinkaari 1), University of Helsinki, Helsinki FI-00014, Finland.
Tel.: +358 (0)2941 58490; fax: +358 2941 57694; e-mail: virpi.ahola@helsinki.fi

presently consisting of a network of 4000 small meadows within an area of 50 by 70 km, and to new climatic conditions (Hanski, 2011).

One set of traits that may have been under selection is related to larval growth and development. Following the obligatory winter diapause, Glanville fritillary larvae resume growth in early spring, when the ambient temperatures are low, and lower the more northern the locality. Gregarious caterpillars bask in the sun to raise their body temperature (Nieminen *et al.*, 2004), but cold and cloudy weather greatly slows down larval development. The body reserves for the adult stage are built up during the late larval stages. In the Glanville fritillary (Saastamoinen, 2007a) and many other butterflies (Barah & Sengupta, 1991; Honëk, 1993; García-Barros, 2000; Rodrigues & Moreira, 2002), pupal weight is positively related to reproductive success of females. High pupal weight can be achieved by longer development time. In males, in contrast, fast development may be selected for, because fast larval development allows earlier eclosion, which increases mating success (Stillwell *et al.*, 2010; Saastamoinen *et al.*, 2013). In both sexes, slow larval development in the spring increases the risk of parasitism by the braconid wasp *Cotesia melitaearum* (van Nouhuys & Lei, 2004). There are thus many reasons to expect that larval growth and development in the spring are under selection, which may be antagonistic in the two sexes.

Previous studies on the Glanville fritillary have found differences among larval families in body weight after diapause, post-diapause development time and survival (Kvist *et al.*, 2013), suggestive of heritable variation in these traits. Indeed, heritability estimates for post-diapause larval development time and larval and pupal weights are high and significant (h^2 from 0.24 to 0.56) (Saastamoinen, 2008; Klemme & Hanski, 2009; Kvist *et al.*, 2013; de Jong *et al.*, 2014). Similar results have been reported for other insects (Blanckenhorn, 1998; Kause *et al.*, 1999). Studies on the silkworm have identified a few quantitative trait loci (QTL) associated with pupal weight (Zhan *et al.*, 2009; Lie *et al.*, 2010; Li *et al.*, 2013), which explained a large proportion of genetic variation, but unfortunately the genes in these loci have not been identified. In the Glanville fritillary, there is significant variation among larval families in the effect of temperature on growth rate (Kallioniemi & Hanski, 2011) and gene expression (Kvist *et al.*, 2013). In a comparison of four regional populations in northern Europe, Duplouy *et al.* (2013) reported highly significant variation in larval growth rate among the regional populations. The two northernmost populations, including the one from the Åland Islands, showed the fastest growth following winter diapause under common garden conditions. In summary, the previous results indicate that there is heritable variation in post-diapause growth and development of the caterpillars as well as significant differences among regional populations that may be the result of adaptation to dissimilar environmental conditions.

Our aim in this study was to identify genes that contribute to phenotypic variation in post-diapause larval and pupal growth traits and adult lifespan. We have used high-throughput RNA-seq data to identify genes with greatly reduced variation in the Åland Islands in comparison with a Chinese Nantaizi population, in the part of Asia from where the species expanded to northern Europe following the glacial retreat (Wahlberg & Saccheri, 2007). The population in the Åland Islands may have lost variation during range expansion and due to drift, but variation may also have been reduced due to positive selection while the population has become adapted to new environmental conditions. Having identified a set of genes with reduced genetic variation in Åland, we selected a set of candidates for an association study examining larval growth traits in a large sample. Using this approach, we discovered allelic variation in genes encoding clip-domain serine proteases that are associated with larval growth in northern Europe.

Materials and methods

Sample preparation and RNA-seq sequencing

This study is based on two transcriptome sequence data sets. The first data set has been previously analysed by Vera *et al.* (2008) and includes samples from Finland, whereas the second data set includes samples from both Finland and China. Caterpillars were collected from many larval families from Nantaizi, Tianshan mountains in China (40°24′N, 87°12′E) (Wang *et al.*, 2011), and butterflies were reared for three generations in common garden conditions. Three 7th instar larvae and one pupa were sampled from one family in the third generation for sequencing. The first sample from the Åland Islands in Finland consists of ca. 80 larvae, pupae and adults from eight families (Vera *et al.*, 2008). The second sample consists of 62 larvae (4th and 7th instars) and adults from several families. Adults were one and four days old and were separated into head, thorax and abdomen (Table S1). The sample for the second study is the offspring of individuals originally sampled from several local populations across the Åland Islands in 2005 (Saastamoinen, 2008).

RNA extraction of the first sample from Åland has been described in Vera *et al.* (2008). The sample was divided into two pools, normalized and sequenced with Roche GS20. Total RNA of the second sample was extracted from samples frozen in liquid nitrogen using Trizol reagent (Thermo Fisher Scientific, Waltham, MA, USA) according to manufacturer's instructions, followed by acid phenol:chloroform:isoamylalcohol (25 : 24 : 21) extraction. Concentrations and quality were measured using NanoDrop (Rockland, DE, USA) and Agilent 2100

Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA), respectively. RNA from different samples was pooled (Table S1), and 10 μg of each pool was used for cDNA normalization (Evrogen, Moscow, Russia; CS010-1). Five μg of normalized cDNA was used for 454-pyrosequencing (Roche Diagnostics, Branford, CT, USA; GS FLX) in two runs with half of the reactions containing cDNA from the Åland and one-fourth from the Nantaizi population.

RNA-seq assembly and annotation

Data from Vera *et al.* (2008) and the second transcriptome data set were combined for RNA-seq assembly. Raw reads from the four 454 runs were screened for adapter and primer sequences and trimmed using a custom script. The trimmed data set was assembled using miraSearchESTSNP package of the Mira assembler (Chevreux *et al.*, 1999, 2004). This package is suitable for assembling heterogeneous data sets from different populations that complicates the assembly process.

Protein description (DE), Gene Ontology (GO) and Enzyme Commission (EC) number annotations for assembled transcripts were obtained using PANNZER (v. 1.0) annotation tool with default parameters (Radivojac *et al.*, 2013; Koskinen *et al.*, 2015). Coding regions of the transcript contigs were additionally annotated with InterProScan (version 4.8) (Quevillon *et al.*, 2005) to identify protein domains, and with KAAS server to find KEGG pathways (Moriya *et al.*, 2007). The candidate genes for genotyping were manually annotated by aligning the contigs with the nearest homologues, and the descriptions of the coding regions were checked manually.

SNP analysis and selection of contigs

Single-nucleotide polymorphisms (SNPs) were identified from the contigs using custom scripts. Contig positions with two or three alleles were assigned as SNPs if the read count of each allele was at least three and the average Phred quality of the position was more than 20. The SNPs within homopolymers longer than 3 bp were not accepted due to potential assembly errors. SNPs within coding regions were further categorized as synonymous vs. nonsynonymous. We calculated the mean SNP density in the coding regions in 100 bp intervals within each contig. Regions with more than ten SNPs/100 bp were defined as SNP clusters.

We focused on SNPs in the annotated contigs that are polymorphic in Nantaizi in China and monomorphic in Åland in Finland, as well as SNPs that are monomorphic in both populations but were fixed for different alleles (the latter group includes SNPs for which the Chinese population may be polymorphic in larger sample). This approach allows us to identify contigs with reduced variation in the large sample from the Åland population even if the Nantaizi sample is small and sequencing coverage is low. While identifying contigs with reduced variation in Åland, we considered only contigs with five or more SNPs, of which at least one was nonsynonymous and at least three were monomorphic in Åland. In addition, the proportion of monomorphic SNPs out of all the SNPs in the contig was required to be greater than 0.23. Using this procedure, we found 366 contigs that have reduced variation in Åland. The procedure for selecting genes and SNPs is described in Fig. S1.

Gene set enrichment analysis and choice of candidate genes

We ran the gene enrichment analysis for three sets of genes. The first set consists of the 366 contigs with reduced variation in Åland as described above. The second and third sets include contigs with at least 10 nonsynonymous SNPs, and contigs with at least one SNP cluster, respectively. Using enrichment analyses (GSEA) (Mootha *et al.*, 2003), we identified the GOs (Ashburner *et al.*, 2000) that were overrepresented in the three gene sets. We classified these gene sets according to their GO groups, and GO groups with more than two contigs were included into the GSEA. For the uncorrected GSEA, GO groups with $P < 0.01$ were considered to be enriched. Significant GO groups occurring at higher levels in the GO hierarchy, and including essentially the same set of genes than those at lower levels, were removed from the GO list. Among the significant GO groups, we selected 16 candidate genes for genotyping and association analysis based on the proportion of monomorphic SNPs in Åland (see section on SNP analysis and selection of contigs above), the number of nonsynonymous SNPs and the presence of high-density SNP clusters. The set of 16 genes was completed by selecting a few other genes in the same GO groups.

Genotyping of candidate genes

A SNP within a candidate gene was chosen for genotyping if it was located near a SNP cluster or in the region that has lost variation in Åland, and/or if it was nonsynonymous (Table S2). A total of 250 to 550 bp long sequences around the selected SNPs were re-sequenced with Megabase (Amersham Biosciences SV Corp, Sunnyvale, CA, USA) capillary sequencer according to the manufacturer's instructions. To design sequencing primers, the candidate genes were mapped to the draft genomic contigs by BLAST. The locations of candidate genes in the genome (v.1) (Ahola *et al.*, 2014) are given in Table S2. The primers were designed using an in-house pipeline implementing the Primer3 algorithm (Untergasser *et al.*, 2012). On average seven and five independent,individuals from Åland and

Nantaizi were sequenced to verify the validity of the reference genome sequences and to confirm that the candidate SNPs were polymorphic in the target population. At the end, 28 SNPs in 12 genes were chosen for genotyping (Table S2).

The genotyping assay for Sequenom (SEQUENOM Inc., San Diego, CA, USA) iPLEX Gold was designed using MassARRAY® Assay Designer 3.1 (Sequenom). All probes and primers are available on request. The quality of the genotypes was evaluated using an in-house pipeline. The genotype signals were manually curated using scatter plots. Altogether, 23 SNPs in 10 genes fulfilled the quality criteria. The average sample call rate of the 23 accepted SNPs was 0.995 and varied between 0.979 and 1.0. The overall performance of the genotyped markers is summarized in Table S2. Sixteen individuals were genotyped twice for each accepted marker for concordance evaluation. One individual was excluded from the analyses as genotypes were inconsistent for the two replicates.

A region of one candidate gene, prophenoloxidase-activating proteinase-3 (*ProPO3*), was additionally sequenced using Sanger sequencing to obtain genotypes for two SNPs, c233_est:376T>C and c233_est:455A>T, for additional 23 Nantaizi individuals.

Phenotypes and genotyping material

In the fall 2009, 9616 Glanville fritillary larvae were collected from the wild and maintained in the diapause in the laboratory. The larvae were collected from 493 local populations across the entire metapopulation in the Åland islands (Hanski, 2011). Most populations had more than one group of full-sib larval families, and three randomly chosen larvae were collected from each family. In the spring 2010, 4497 larvae were reared individually under common garden conditions in the laboratory (27 : 10 °C; 12 : 12, L/D). Larvae were weighed (Mettler-Toledo XS 105 analytical balance, accuracy 0.01 mg) when they broke diapause (5th instar), in the beginning of each subsequent instar (6th, 7th and the uncommon 8th instar), and in the pupal stage. In the beginning of the 6th instar, larvae were randomly assigned into two different temperature treatments (standard; 28 : 8 °C 12 : 12, L/D, and high; 35 : 8 °C 8 : 16, L/D) (Saastamoinen *et al.*, 2013). The standard temperature treatment corresponds to the thermal conditions in the Åland (Kallioniemi & Hanski, 2011), whereas the high temperature treatment is more characteristic of the highly continental thermal conditions in Nantaizi with very high daily temperature at the ground level (Wang *et al.*, 2011). After eclosion, butterflies were sexed and marked individually by writing a number on the underside of the hind wing. The adult butterflies were kept under standard conditions and fed daily with honey:water solution (1 : 3) until they died.

All the 4497 larvae were phenotyped for the key post-diapause development traits: the 6th and 7th instar larval weights (mg) and the developmental time between the 6th and 7th instars (days); pupal weight (mg); and adult lifespan (days). We calculated growth rate (mg day^{-1}) between the 6th and 7th instars by dividing the weight gain by the developmental time. Individuals with the extra 8th larval instar were excluded from the material. Based on the growth rate, we selected 392 individuals for genotyping and association analyses according to the selective genotyping procedure (Lynch & Walsh, 1997). Males having growth rate smaller than the 25% quantile or greater than the 75% quantile value were chosen for genotyping (Table S3), whereas for females we used the 30% and 70% quantiles to obtain a balanced sample. Moreover, only apparently unrelated individuals were included into the sample, and therefore, only 1 or 2 individuals were chosen from each local population. The exception was the largest populations, in which larval families are generally only distantly related, and from which 3 to 9 individuals were selected. These populations account for 6% of the sampled populations. The sample of 392 individuals covers well the entire Åland islands (Hanski, 2011). We tested for the effects of known explanatory variables on growth rate, namely host plant, sex and the temperature treatment. The results showed significant effects of sex and temperature treatment, but not of host plant. The sample of 392 individuals is well balanced in terms of sex and temperature treatment groups (Table S3).

Association analyses

The Åland Islands study system has approximately 4000 small dry meadows that are potential habitat patches for the Glanville fritillary (Hanski, 1999, 2011). These patches were divided into semi-independent patch networks with Spomsim (Hanski *et al.*, 1996), with the parameter values $b = 0.5$ and $q = 1.5$. The hierarchical clustering tree was cut to obtain 22 networks, 12 of which had individuals in the present data set (Fig. S2). Biologically, butterflies in the same network are part of the same metapopulation, as individuals are able to move commonly among the habitat patches within a network, but movements between networks are uncommon because of long distance (Hanski *et al.*, 1996). In the analysis, the nonindependence of individuals from the same network was modelled by a block-diagonal covariance matrix, where each block has identical covariance structure, whereas individuals from different networks are assumed to be independent.

Association analyses using additive, dominance and recessive models were performed by fitting linear mixed effects models with sex, temperature, sex*SNP and temperature*SNP as fixed effects and network as a random effect. Association analyses were performed using the

SAS® 9.4 (2002–2012 by SAS Institute Inc., Cary, NC, USA.). The *P* values were corrected for multiple testing using the permutation test with 1000 permutations per SNP. All data used in the association analyses are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.s57k2.

Results

Transcriptome assembly and annotation

The transcriptome assembly resulted in 25 788 unique contigs with an average length of 407 bp and coverage of 5.6 reads (median values 327 bp and 3.6 reads). We were able to annotate half of the contigs (12 486), of which 10 401 contigs were specific to the Glanville fritillary and the rest were contaminants, mostly of bacterial origin. Approximately 75% and 9% of the annotated Glanville fritillary contigs yielded a GO group and an EC classification, respectively. Moreover, 22% of the contigs possessed a known protein domain or belong to Pfam gene families, and 21% were identified to be part of KEGG pathways. The alignments of the annotated contigs to the Uniprot database covered mostly the entire length of the query contig but only 23%, on average, of the target protein (Fig. S3), indicating that some transcripts are not of full length.

Genes with reduced allelic variation in the Åland Islands population

Altogether, 48 777 SNPs and 1812 indels were discovered from the Finnish Åland Islands and Chinese Nantaizi samples. Most SNP variation was biallelic, but we detected 566 triallelic SNPs fulfilling all the minor allele criteria. Annotated contigs included 38 374 SNPs, of which more than half (20 512) were located within the predicted coding region. The SNP density in the annotated regions was ca. 3/kbp. Table S4 shows the total number and the number of nonsynonymous SNPs in Åland and the number of nonsynonymous SNPs that produce a different amino acid between the Åland and Nantaizi populations. Note that because of small sample size from the Nantaizi population, we focus on polymorphic sites in Nantaizi and on pairwise differences between the two populations. Approximately one-third of the SNPs within the coding region were nonsynonymous.

The transcriptome data allowed us to identify contigs with reduced variation in Åland. For a detailed analysis, we chose 366 contigs in which the Åland population has clearly less variation or has different alleles than the Nantaizi population. There were 23 contigs with 5 to 17 SNP sites that had polymorphism or different allele in the Nantaizi sample but had lost all variation in Åland. Variation may have been lost during the post-glacial migration to northern Europe via genetic drift, inbreeding and other such reasons, but also due to local adaptation to new environmental conditions. Reduced variation due to drift or inbreeding could be assumed to be randomly spread across the transcriptome, whereas reduced variation due to adaptation could be assumed to be concentrated on functionally relevant gene groups.

To identify candidate genes under selection, we conducted an enrichment analysis for the fraction of genes showing reduced variation. Out of 507 GO groups tested, we detected 25 overrepresented GO annotations at the significance level of $P < 0.01$ (Table 1). The protein coding genes in these groups are involved in, for example, structural constituent of cuticle (e.g. larval cuticle proteins) and hydrolase activity, especially serine-type peptidase activity (e.g. serine proteases), as well as in epoxide hydrolase activity (e.g. juvenile hormone) (Table 1). Additionally, genes involved in oxidoreductase activity (e.g. NADH dehydrogenase (ubiquinone)) had lost variation in Åland.

Next, we focused on contigs with elevated number of nonsynonymous SNPs and contigs with SNP clusters that have potentially been under balancing selection. The median number of nonsynonymous SNPs per contigs was one (range 0–56), but there were 36 contigs with more than 10 nonsynonymous SNPs. In this gene set, six GO groups were overrepresented, and they were related to endopeptidase activity, oxidation reduction and fatty acid metabolic processes. Genes related to the structural constituent of cuticle had large numbers of nonsynonymous SNPs (on average 11 in Åland and eight between Åland and Nantaizi), as did genes involved in serine-type endopeptidase inhibitor activity (e.g. serpins), although the latter did not show significant enrichment at the level of $P < 0.01$.

The mean SNP density per 100 bp of the coding region was two (range 1–13). There were 39 contigs with SNP density more than ten SNPs/100 bp, and their functions were enriched into four GO groups. The genes were related to respiratory chain and oxidoreductase, and especially to cytochrome-c oxidase activity. Some SNP clusters formed two or several haplotypes, some of which only occurred in one population but not in the other. The clusters typically consisted of nonsynonymous SNPs, indicating that different haplotypes may have functional consequences for the proteins coded by the genes. We therefore used SNP density as the third criterion to choose candidate genes.

Taking into account results on genes with reduced variation in Åland, genes with large number of nonsynonymous SNPs and high SNP density, we selected 33 GO groups for further consideration (Fig. S4). These groups included oxidoreductase activity, fatty acid metabolic process and proteolysis GO groups, especially genes involved in serine-type endopeptidase activity (Table 1). Finally, out of these groups we selected 16 candidate genes involved in serine-type endopeptidase

Table 1 *P* values for enriched GO categories in three sets of genes. (1) Reduced variation in the Åland Islands compared with the Nantaizi population. (2) Contigs with more than 10 non synonymous SNPs, and (3) contigs with SNP clusters. Grey box indicates the GO categories that were chosen for genotyping. High-level categories in the GO hierarchy and the corresponding *P* values are shown in bold.

GO category	Description	Reduced variation	Nonsyn SNPs	SNP clusters
GO:0005198	**Structural molecule activity**	**1.1E-05**	–	–
GO:0042302	Structural constituent of cuticle	1.4E-06	–	–
GO:0006508	**Proteolysis**	**0.0133**	**0.0319**	–
GO:0008233	Peptidase activity	0.0084	0.0423	–
GO:0004175	Endopeptidase activity	0.0024	0.0051	–
GO:0004252	Serine-type endopeptidase activity	0.0014	0.0135	–
GO:0030162	Regulation of proteolysis	0.0076	–	–
GO:0010951	**Negative regulation of endopeptidase activity**			
GO:0004867	Serine-type endopeptidase inhibitor activity	–	0.0382	–
GO:0016787	**Hydrolase activity**			
GO:0016811	Hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds, in linear amides	0.0060	–	–
GO:0004301	Epoxide hydrolase activity	0.0032	–	–
GO:0008745	N-acetylmuramoyl-L-alanine amidase activity	0.0097	–	–
GO:0055114	**Oxidation reduction**	**0.0213**	**0.0061**	**0.0002**
GO:0016491	Oxidoreductase activity	0.0208	0.0288	0.0001
GO:0016616	Oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	0.0073	0.0433	0.036
GO:0004768	Stearoyl-CoA 9-desaturase activity	–	0.0020	–
GO:0070469	Respiratory chain	–	–	0.0098
GO:0004129	Cytochrome-c oxidase activity	–	–	0.0024
GO:0008152	**Metabolic process**	**–**	**–**	**–**
GO:0042180	Cellular ketone metabolic process	–	0.0095	–
GO:0032787	Monocarboxylic acid metabolic process	–	0.0085	–
GO:0000036	ACP phosphopantetheine attachment site binding involved in fatty acid biosynthetic process	–	0.0039	–
GO:0009057	Macromolecule catabolic process	0.0023	–	–
GO:0000272	Polysaccharide catabolic process	0.0029	–	–
GO:0006026	Aminoglycan catabolic process	0.0017	–	–
GO:0042221	**Response to chemicals**			
GO:0009636	Response to toxic substance	0.0032	–	–
GO:0005488	**Binding**	**0.0138**	**0.0394**	**–**
GO:0051287	NAD binding	0.0096	–	–
GO:0005575	**Cellular component**			
GO:0044444	Cytoplasmic part	0.0007	–	–
GO:0000502	Proteasome complex	0.0008	–	–
GO:0005840	Ribosome	0.0021	–	–
GO:0005737	Cytoplasm	0.0024	–	–
GO:0030529	Ribonucleoprotein complex	0.0032	–	–
GO:0032991	Macromolecular complex	0.0032	–	–
GO:0044424	Intracellular part	0.0033	–	–
GO:0005739	Mitochondrion	0.0057	–	–
GO:0043229	Intracellular organelle	0.0087	–	–
GO:0043231	Intracellular membrane-bounded organelle	0.0088	–	–

– indicates *P* values > 0.05.

activity and serine-type endopeptidase inhibitor activity and structural constituent of cuticle, which have known functions in larval development. After genotyping, 23 SNPs in 10 candidate genes were selected for the association analysis (Table 2, Table S2).

Association of candidate genes with larval development

Genotyping of the 23 SNPs in the 10 genes (Table S2, Table 2) was successful for 335 individuals. Genetic association analyses of larval development traits and adult lifespan (Table S5) were performed for all markers using sex and temperature treatment and their interaction with SNPs as fixed effects. In the case of larval growth rate, the main effect of temperature treatment was highly significant (higher growth rate in high temperature). Temperature treatment and sex were significant for the development time (males developed faster), and male pupae were significantly smaller than female pupae. Adult lifetime was not related to any of the explanatory factors. Population structure in the Åland

Islands was taken into account using habitat patch network (Fig. S2) as a random effect in the model. No correlations were found between minor allele frequencies and the median of the traits across the networks. The numbers of individuals in each network and the minor allele frequencies and the distributions of phenotypic traits across the 12 networks are shown in Table S6 and Figs S5 and S6.

In the SNP c233_est:455A>T in the gene prophenoloxidase-activating proteinase-3 (*ProPO3*), larval growth rate was 20 and 24% higher, on average, in AT and TT larvae than in AA larvae, respectively (additive effect,

Table 2 Candidate genes and the information on the respective SNPs and selection criteria.

Contig name	Gene name	Number of SNPs	Number of nonsyn SNPs	Selection criteria*	Genotyped SNPs
Serine proteases: GO:0008233, GO:0004175, GO:0004252					
step3_c233	prophenoloxidase-activating proteinase-3	47	11	R, N	3
step3_c177	vitellin-degrading protease precursor	38	7	R	3
step3_c250	prophenoloxidase-activating proteinase-1	41	3	R	1
step3_c3917	serine proteinase-like protein 1	8	1	C	1
step3_c50	haemolymph proteinase 5	31	4		5
step3_c1356	chymotrypsinogen-like protein	11	5	R	1
Serine protease inhibitors: GO:0004867					
step3_c172	serpin 1	17	4		4
step3_c875	kazal-type proteinase inhibitor	8	0	C	1
Cuticular proteins: GO:0005198, GO:0042302					
step3_c2634	cuticular protein	16	4	R	1
step3_c480	endocuticle structural glycoprotein SgAbd-8	18	5	R	3

*R = reduced variation; N = nonsynonymous SNPs; C = SNP cluster.

Table 3 Results of the association analysis for three genes and three phenotypes: (a) main additive and dominance effects, (b) SNP*treatment and SNP*sex interaction effects. Effect sizes are shown for the additive main effects. *P* values are corrected by permutation test.

(a) Gene name	SNP	Additive effect	ddf	$F_{1,ddf}$	*P* value	Dominance effect	ddf	$F_{1,ddf}$	*P* value	Effect size (%)
Growth rate										
Prophenoloxidase-activating proteinase-3	c233_est:376T>C	0.96	318	3.0	0.082	2.19	318	4.5	0.033	0.20
(*ProPO3*)	c233_est:455A>T	1.21	317	10.4	0.0030	1.75	317	9.7	0.0030	1.85
Development time										
Vitellin-degrading protease precursor	c177_est:199G>A	0.60	318	0.88	0.35	0.75	318	0.0	0.96	0.00
Pupa weight										
Haemolymph proteinase 5	c50_est:735A>G	−4.07	317	5.1	0.023	−5.56	317	1.7	0.21	1.21
	c50_est:816C>A	4.11	316	5.0	0.025	5.14	316	5.2	0.018	0.94
	c50_est:824A>G	−4.09	317	5.2	0.022	−5.57	317	1.7	0.21	1.08
Vitellin-degrading protease precursor	c177_est:181A>G	10.95	317	7.5	0.0080	11.06	317	5.9	0.014	1.23

(b) Gene name	SNP	SNP*treatment additive effect	ddf	$F_{1,ddf}$	*P* value	SNP*treatment dominance effect	ddf	$F_{1,ddf}$	*P* value
Growth rate									
Prophenoloxidase-activating proteinase-3	c233_est:376T>C	−1.23	318	7.5	0.0090	−1.94	318	5.8	0.017
(*ProPO3*)	c233_est:455A>T	−0.99	317	6.1	0.018	−1.36	317	4.4	0.046

Gene name	SNP	SNP*sex additive effect	ddf	$F_{1,ddf}$	*P* value	SNP*sex dominance effect	ddf	$F_{1,ddf}$	*P* value
Development time									
Vitellin-degrading protease precursor	c177_est:199G>A	−0.80	318	14.9	<0.0001	−1.16	318	9.9	0.001

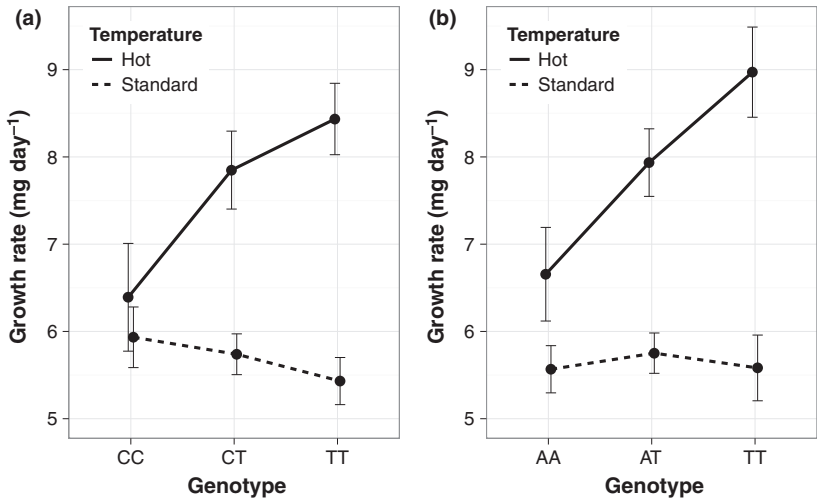

Fig. 1 Growth rate between the 6th and 7th instars of larvae reared in standard and high temperatures (mean ± SE). The horizontal axis shows three genotypes of the SNPs (a) c233_est:376T>C and (b) c233_est:455A>T of the *ProPO3* gene. Sample sizes for each genotype and treatment group are given in Table S7.

$F_{1,317}$ = 10.4, P = 0.0030, Table 3a, Fig. 1b). In a model with the interaction between the SNPs and the temperature treatment, two closely linked SNPs c233_est:376T>C and c233_est:455A>T in *ProPO3* showed significant or marginally significant interaction effects ($F_{1,318}$ = 7.5, P = 0.0090 and $F_{1,317}$ = 6.1, P = 0.018, Table 3b, Fig. 1). In SNP c233_est:455A>T, the additive effect of the T allele was 1.21 (95% CI: 0.44–1.97) in high temperature, but only 0.22 (95% CI: −1.37, 1.77) in moderate (standard) temperature (Fig. 1b). In SNP c233_est:376T>C, the additive effect of the T allele was 0.96 (CI: 0.16, 1.77) in high temperature and negative, −0.17 (CI: −1.76, 1.45), in moderate temperature (Fig. 1a). In the SNP c233_est:376T>C, the CC genotype showed similar growth rates in the standard and high temperatures, whereas the growth rate of TT individuals was 56% faster in high than in standard temperature (Fig. 1).

The second significant association was between the SNP c177_est:199G>A in the vitellin-degrading protease precursor gene and development time between the 6th and 7th instars. This SNP had no significant main effect, but the association was strongly sex-related (Fig. 2). In males, the AA genotype had the longest and the GG genotype the shortest development time, whereas in females the genotypic effects were reversed ($F_{1,318}$ = 14.9, P < 0.0001, Table 3b, Fig. 2). Another closely linked SNP in the vitellin-degrading protease precursor gene was associated with pupal weight, but the effect size was low (1.2%). Three closely linked SNPs in the gene haemolymph proteinase 5 (c50_est) showed an additive effect (−4.1, $F_{1,317}$ = 5.1, P = 0.023, c50_est:735A>G) on pupal weight (Table 3a, Fig. S7). The effects were significant, but the mean differences in pupal weight between the homozygotes were only 10 mg and the effect sizes were 0.9–1.2%. All other genes showed no significant association with larval development traits (Table S5), and adult lifespan was not related to any of the candidate genes.

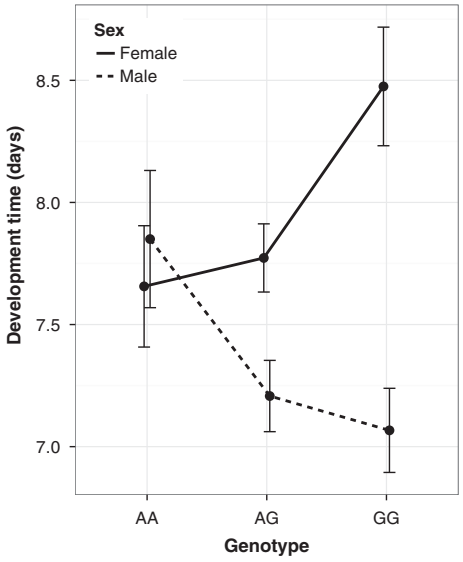

Fig. 2 Development time between the 6th and 7th instars in three genotypes of vitellin-degrading protease precursor gene in females and males (mean ± SE). Sample sizes for each genotype and sex are given in Table S7.

Discussion

Genes that showed reduced allelic variation in the Åland Islands in northern Europe in comparison with the glacial refuge area in Nantaizi in China showed a broad range of functions, but they were also significantly enriched for certain functional groups. We presume that genes in the overrepresented GO groups have been potentially under selection, or are linked with selected genes. We selected candidate genes for association analyses from three enriched GO groups that were considered relevant for larval growth and development, namely serine proteases, serine protease

inhibitors and cuticular proteins. Although genetic variation in the functional loci may have been greatly reduced in the Åland Islands, some variation still remains, and in any case, most of our markers are only linked with causal loci. It is hence reasonable to base the association analysis on material from the Åland Islands. To increase the power of the analysis, we applied the selective genotyping strategy that includes only individuals with extreme phenotypes (Lynch & Walsh, 1997), and enable detection of genes with small effect sizes. We found three clip-domain-containing serine proteases that were associated with larval growth rate, larval development time and pupal weight. The two most interesting genes were the immune-related gene *ProPO3* with a significant interaction between the additive genetic effect and temperature treatment, and vitellin-degrading precursor gene, which showed sexually antagonistic effect on development time.

A family of clip-domain-containing serine proteinases, including serine proteases and serine protease homologues, are key enzymes involved in insect innate immune response and development (Gonzalez-Santoyo & Cordoba-Aguilar, 2012). These enzymes contain one or several clip domains. Immune-related proteases function in the activation of prophenoloxidase (*ProPO*) leading to melanin synthesis and activation of Toll receptors (Jiang & Kanost, 2000; Cerenius *et al.*, 2008, 2010). In the *ProPO* activation pathway, recognition of pathogen stimulates the sequential activation of haemolymph serine proteinases and leads to activation of *ProPO* that converts to phenoloxidase (PO). PO is involved in microbial killing, melanin synthesis and wound healing. Melanization is a major resistance mechanism against parasites and pathogens in insects, and it leads to encapsulation and smothering of the pathogens (Cerenius *et al.*, 2008). Serine proteinase inhibitors, including serpins, regulate the activation of *ProPO* by inhibiting the activation of serine proteinases where it is not appropriate (Kanost, 1999; Tong & Kanost, 2005). In addition to their role in defence mechanisms, serine proteases are involved in the degradation of cuticular proteins during moulting process (Brookhart & Kramer, 1990; Zheng *et al.*, 2009). Serpin-1 isoforms have been suggested to regulate the activity of proteases involved in cuticle degradation during moulting (Hegedus *et al.*, 2008; Zheng *et al.*, 2009). Previous studies have reported trade-offs between immunity and life-history traits in several insect species (Schulenburg *et al.*, 2009). For instance, in *Spodoptera littoralis* there is a trade-off between PO level and larval body weight and development time (Cotter *et al.*, 2008).

The above results on clip-domain proteases are supported by several expression studies showing that many serine proteases and serine protease inhibitors are expressed in different larval stages in the Glanville fritillary (Kvist *et al.*, 2013) and other Lepidoptera (An *et al.*, 2009; Zheng *et al.*, 2009; Zhao *et al.*, 2010). Moreover, Kvist *et al.* (2013) reported differences in gene expression between different temperature treatments in the Glanville fritillary; for instance, GO groups related to serine-type endopeptidase activity (GO:0004252) and structural constituent of cuticle (GO:0042302) had significant gene expression differences between standard and high temperature treatments (Kvist *et al.*, 2013). In the high temperature treatment, gene expression differed between families suggesting high heritability and the effect of allelic variation on gene expression in the above-mentioned GO groups. As far as we know, the relationship between genotypic variation in these genes and larval development has not been previously studied in insects. Further studies are needed to investigate the relationship between immunity and larval development, and to test whether genetic variation in *ProPO3* affects the activity of PO and encapsulation capacity.

We found a significant additive genetic effect of *ProPO3* on larval growth rate, but there was also a significant interaction between the genotypic effect and the temperature treatment. In two linked SNPs, adding one copy of the T allele increased the estimated mean growth rate by 1.2 mg day^{-1} (c233_est:455A>T) and by 1.0 mg day^{-1} (c233_est:376T>C) in high temperature, indicating that especially the TT individuals can greatly benefit from high temperature, whereas in moderate temperature TT individuals grew slower than the others (Fig. 1). Indeed, the TT individuals grew more than 50% faster in high than in moderate temperature, whereas high temperature increased the growth rate of the other homozygotes by only 12% (c233_est:376T>C) and 24% (c233_est:455A>T). In the Åland Islands, the frequency of the TT genotype was 0.41 and 0.24 in the two SNPs in *ProPO3* (Table S7). To compare these genotype frequencies with those in the Nantaizi population, we genotyped 27 Nantaizi individuals for the two SNPs in *ProPO3*. Both SNPs were fixed to the TT genotype. In the northern temperate climate in Finland, spring is cool, with ambient air temperature often around 10 °C, although in sunny weather the black larvae basking in the sun can increase their body temperature up to 35 °C (Kuussaari, 1998). In such greatly fluctuating thermal environment, different genotypes may be at an advantage at different times, which potentially maintains polymorphism. In contrast, in the very continental climate in the Tianshan mountains in China, ground temperatures in the spring time can be very high, more than 35° C, and hence, the temperature tolerant TT genotype should have a clear advantage. Another possibility is that the TT genotype is associated with higher melanization leading to darker cuticle, which would help larvae to maintain high body temperature and fast growth in sunny but cool weather in the spring.

The highly divergent performance of the TT genotype at different temperatures represents a high level of phenotypic plasticity. High phenotypic plasticity has been reported for PO production in studies of insect immune

response (Barnes & Siva-Jothy, 2000; Cotter *et al.*, 2008), but plasticity of genes of the melanization pathway has not been reported in response to larval development. Phenotypic plasticity for melanization in response to different nutritional conditions during development has been reported in the Glanville fritillary (Saastamoinen & Rantala, 2013). Moreover, this response is also dependent on the larval growth rate (Saastamoinen & Rantala, 2013).

Another gene, vitellin-degrading precursor, was associated with the development time between the 6th and 7th instars, but in opposite direction in the two sexes, suggesting sexually antagonistic selection. The difference between the male and female development times is very small or nonexistent for one homozygote (AA) but high for another (GG). In the Glanville fritillary, larval development time is a few days shorter in males than females (Saastamoinen *et al.*, 2013), enabling faster eclosion of males, which enhances their mating success as females usually mate only once and soon after eclosion (Saastamoinen, 2007a,b). Vitellin-degrading precursor has been found to be associated with the occurrence of an extra (8th) larval instar (Saastamoinen *et al.*, 2013). This association is also in the opposite directions in two sexes: in males, the A allele increases the frequency of the extra instar whereas in females the probability of the extra instar is reduced (Saastamoinen *et al.*, 2013).

All the genes that we found contributing to larval development have small effect sizes, typical to association studies of complex traits (Allen *et al.*, 2010; Pritchard *et al.*, 2010; Yang *et al.*, 2010). Nonetheless, the genotypic effects are large enough to have significance in natural populations. It is possible that the SNPs and genes found to be associated with larval development are not causal themselves but only linked with causal loci. This could explain relatively small effect sizes due to incomplete LD of the SNPs and the causal loci. This is probable due to short LD in Lepidoptera (Guo *et al.*, 2011; Ahola *et al.*, 2014) and to unknown gene-by-gene and gene-by-environment interactions (Houle, 1992).

In summary, we found two genetic effects on larval growth and development time with significant interactions with temperature and gender. Adding one or two T alleles in two SNPs in the *ProPO3* gene accelerated growth in high temperature, but had no effect or even reduced growth in moderate temperature. The Chinese Nantaizi population living in continental climate with hot spring and summer was fixed for this genotype, whereas the Finnish Åland Island population inhabiting a cooler climate was highly polymorphic for the SNP. In the vitellin-degrading precursor gene, the common homozygous genotype was associated with fast development in males but slow development in females. This difference amplifies the difference in adult eclosion times, males eclosing earlier, on average, to compete for newly eclosed females. We suggest that polymorphism may be maintained by a combination of sexual and natural selection: in years with consistently benign conditions, it would pay off for males to eclose earlier than females, but in years with prolonged unfavourable weather periods, early eclosing males may miss mating opportunities when the development of other males and females is greatly prolonged. Both *ProPO3* and vitellin-degrading precursor exemplify highly plastic genes that interact with environmental conditions and gender in shaping larval development.

Acknowledgments

We thank Eeva-Marja Turkki, Kirsi Lipponen, Toshka Nyman and Annukka Ruokolainen for technical assistance. Päivi Lahermo and her laboratory in the Institute for Molecular Medicine Finland is thanked for genotyping, Suvi Ikonen for rearing of larvae, Torsti Schulz for analysing the habitat patch network, and Panu Somervuo and Jussi Nokso-Koivisto for help with bioinformatics, and the IT Center for Science Ltd (CSC), Espoo, Finland, for supercomputing facilities. This study was funded by the European Research Council (AdG232826 to IH), the Academy of Finland (grants #133132, 250444 and 256453 to IH, and 273098 to MS) and the Kone Foundation to MS.

References

Ahola, V., Lehtonen, R., Somervuo, P., Salmela, L., Koskinen, P., Rastas, P. *et al.* 2014. The Glanville fritillary genome retains an ancient karyotype and reveals selective chromosomal fusions in Lepidoptera. *Nat. Commun.* **5**: 4737.

Allen, H.L., Estrada, K., Lettre, G., Berndt, S.I., Weedon, M.N., Rivadeneira, F. *et al.* 2010. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**: 832.

An, C., Ishibashi, J., Ragan, E.J., Jiang, H. & Kanost, M.R. 2009. Functions of *Manduca sexta* Hemolymph proteinases HP6 and HP8 in two innate immune pathways. *J. Biol. Chem.* **284**: 19716.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M. *et al.* 2000. Gene ontology: tool for the unification of biology. *Nat. Genet.* **25**: 25.

Barah, A. & Sengupta, A. 1991. Correlation and regression studies between pupal weight and fecundity of muga silkworm *Antheraea assama* Westwood (Lepidoptera: Saturniidae) on four different foodplants. *Acta Physiol. Hung.* **78**: 261–264.

Barnes, A.I. & Siva-Jothy, M.T. 2000. Density-dependent prophylaxis in the mealworm beetle *Tenebrio molitor* L-(Coleoptera**:** Tenebrionidae): cuticular melanization is an indicator of investment in immunity. *Proc. R. Soc. Lond. B Biol. Sci.* **267**: 177.

Blanckenhorn, W.U. 1998. Adaptive phenotypic plasticity in growth, development, and body size in the Yellow dung fly. *Evolution* **52**: 1394–1407.

Brookhart, G.L. & Kramer, K.J. 1990. Proteinases in molting fluid of the Tobacco hornworm. *Manduca Sexta. Insect Biochem.* **20**: 467.

Cerenius, L., Lee, B.L. & Soderhall, K. 2008. The proPO-system: pros and cons for its role in invertebrate immunity. *Trends Immunol.* **29**: 263.

Cerenius, L., Kawabata, S.-I., Lee, B.L., Nonaka, M. & Soderhall, K. 2010. Proteolytic cascades and their involvement in invertebrate immunity. *Trends Biochem. Sci.* **35**: 575.

Chevreux, B., Wetter, T. & Suhai, S. 1999. Genome sequence assembly using trace signals and additional sequence information. *Comput. Sci. Biol.: Proc. German Conf. Bioinformat. (GCB)* **99**: 45.

Chevreux, B., Pfisterer, T., Drescher, B., Driesel, A.J., Muller, W.E., Wetter, T. *et al.* 2004. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Res.* **14**: 1147–1159.

Cotter, S.C., Myatt, J.P., Benskin, C.M.H. & Wilson, K. 2008. Selection for cuticular melanism reveals immune function and life-history trade-offs in *Spodoptera littoralis*. *J. Evol. Biol.* **21**: 1744.

Duplouy, A., Ikonen, S. & Hanski, I. 2013. Life history of the Glanville fritillary butterfly in fragmented versus continuous landscapes. *Ecol. Evol.* **3**: 5141–5156.

García-Barros, E. 2000. Body size, egg size, and their interspecific relationships with ecological and life history traits in butterflies (Lepidoptera: Papilionoidea, Hesperioidea). *Biol. J. Linn. Soc.* **70**: 251–284.

Gonzalez-Santoyo, I. & Cordoba-Aguilar, A. 2012. Phenoloxidase: a key component of the insect immune system. *Entomol. Exp. Appl.* **142**: 1.

Guo, Y., Shen, Y.H., Sun, W., Kishino, H., Xiang, Z.H. & Zhang, Z. 2011. Nucleotide diversity and selection signature in the domesticated silkworm, *Bombyx mori*, and wild silkworm, *Bombyx mandarina*. *J. Insect Sci.* **11**: 155.

Hanski, I. 1999. *Metapopulation Ecology*. Oxford University Press, New York.

Hanski, I.A. 2011. Eco-evolutionary spatial dynamics in the Glanville fritillary butterfly. *Proc. Natl. Acad. Sci. USA* **108**: 14397–14404.

Hanski, I., Moilanen, A., Pakkala, T. & Kuussaari, M. 1996. The quantitative incidence function model and persistence of an endangered butterfly metapopulation. *Conserv. Biol.* **10**: 578–590.

Hegedus, D.D., Erlandson, M., Baldwin, D., Hou, X. & Chamankhah, M. 2008. Differential expansion and evolution of the exon family encoding the Serpin-1 reactive centre loop has resulted in divergent serpin repertoires among the Lepidoptera. *Gene* **418**: 15.

Hewitt, G.M. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.* **58**: 247–276.

Hewitt, G.M. 1999. Post-glacial re-colonization of European biota. *Biol. J. Linn. Soc.* **68**: 87–112.

Honěk, A. 1993. Intraspecific variation in body size and fecundity in insects: a general relationship. *Oikos* **66**: 483–492.

Houle, D. 1992. Comparing evolvability and variability of quantitative traits. *Genetics* **130**: 195–204.

Jiang, H.B. & Kanost, M.R. 2000. The clip-domain family of serine proteinases in arthropods. *Insect Biochem. Mol. Biol.* **30**: 95.

de Jong, M.A., Wong, S.C., Lehtonen, R. & Hanski, I. 2014. Cytochrome P450 gene *CYP337* and heritability of fitness traits in the Glanville fritillary butterfly. *Mol. Ecol.* **23**: 1994–2005.

Kallioniemi, E. & Hanski, I. 2011. Interactive effects of *Pgi* genotype and temperature on larval growth and survival in the Glanville fritillary butterfly. *Funct. Ecol.* **25**: 1032–1039.

Kanost, M.R. 1999. Serine proteinase inhibitors in arthropod immunity. *Dev. Comp. Immunol.* **23**: 291.

Kause, A., Saloniemi, I., Haukioja, E. & Hanhimäki, S. 1999. How to become large quickly: quantitative genetics of growth and foraging in a flush feeding lepidopteran larva. *J. Evol. Biol.* **12**: 471–482.

Klemme, I. & Hanski, I. 2009. Heritability of and strong single gene (*Pgi*) effects on life-history traits in the Glanville fritillary butterfly. *J. Evol. Biol.* **22**: 1944–1953.

Koskinen, P., Törönen, P., Nokso-Koivisto, J. & Holm, L. 2015. PANNZER: high-throughput functional annotation of uncharacterized proteins in an error-prone environment. *Bioinformatics* **31**: 1544–1552.

Kuussaari, M. 1998. Biology of the Glanville fritillary butterfly (*Melitaea cinxia*). Ph.D. Thesis, University of Helsinki, Helsinki.

Kvist, J., Wheat, C.W., Kallioniemi, E., Saastamoinen, M., Hanski, I. & Frilander, M.J. 2013. Temperature treatments during larval development reveal extensive heritable and plastic variation in gene expression and life history traits. *Mol. Ecol.* **22**: 602–619.

Li, B., Wang, X., Hou, C., Xu, A. & Li, M. 2013. Genetic analysis of quantitative trait loci for cocoon and silk production quantity in *Bombyx mori* (Lepidoptera: Bombycidae). *Eur. J. Entomol.* **110**: 205–213.

Lie, Z., Cheng, L., Fang-yin, D. & Shou-min, F. 2010. Mapping of major quantitative trait loci for economic traits of silkworm cocoon. *Genet. Mol. Res.* **9**: 78–88.

Lynch, M. & Walsh, B. 1997. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, MA.

Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J. *et al.* 2003. PGC-1 alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**: 267.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A.C. & Kanehisa, M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* **35**: W182–W185.

Nieminen, M., Siljander, M. & Hanski, I. 2004. Structure and dynamics of *Melitaea cinxia* metapopulations. In: *On the Wings of Checkerspots: A Model System for Population Biology* (P. R. Ehrlich & I. Hanski, eds.), pp. 63. Oxford University Press, New York.

van Nouhuys, S. & Lei, G. 2004. Parasitoid–host metapopulation dynamics: the causes and consequences of phenological asynchrony. *J. Anim. Ecol.* **73**: 526–535.

Ojanen, S.P., Nieminen, M., Meyke, E., Pöyry, J. & Hanski, I. 2013. Long-term metapopulation study of the Glanville fritillary butterfly *(Melitaea cinxia*): survey methods, data management, and long-term population trends. *Ecol. Evol.* **3**: 3713–3737.

Pritchard, J.K., Pickrell, J.K. & Coop, G. 2010. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr. Biol.* **20**: R208.

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R. *et al.* 2005. InterProScan: protein domains identifier. *Nucleic Acids Res.* **33**: W116.

Radivojac, P., Clark, W.T., Oron, T.R., Schnoes, A.M., Wittkop, T., Sokolov, A. *et al.* 2013. A large-scale evaluation of computational protein function prediction. *Nat. Methods* **10**: 221–227.

Rodrigues, D. & Moreira, G.R. 2002. Geographical variation in larval host-plant use by *Heliconius erato* (Lepidoptera: Nymphalidae) and consequences for adult life history. *Braz. J. Biol.* **62**: 321–332.

Saastamoinen, M. 2007a. Life-history, genotypic, and environmental correlates of clutch size in the Glanville fritillary butterfly. *Ecol. Entomol.* **32**: 235–242.

Saastamoinen, M. 2007b. Mobility and lifetime fecundity in new versus old populations of the Glanville fritillary butterfly. *Oecologia* **153**: 569–578.
Saastamoinen, M. 2008. Heritability of dispersal rate and other life history traits in the Glanville fritillary butterfly. *Heredity* **100**: 39–46.
Saastamoinen, M. & Rantala, M.J. 2013. Influence of developmental conditions on immune function and dispersal-related traits in the Glanville fritillary (*Melitaea cinxia*) butterfly. *PLoS One* **8**: e81289.
Saastamoinen, M., Ikonen, S., Wong, S.C., Lehtonen, R. & Hanski, I. 2013. Plastic larval development in a butterfly has complex environmental and genetic causes and consequences for population dynamics. *J. Anim. Ecol.* **82**: 529–539.
Schmitt, T. & Seitz, A. 2002. Postglacial distribution area expansion of *Polyommatus coridon* (Lepidoptera: Lycaenidae) from its Ponto-Mediterranean glacial refugium. *Heredity* **89**: 20–26.
Schulenburg, H., Kurtz, J., Moret, Y. & Siva-Jothy, M.T. 2009. Introduction. Ecological immunology. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **364**: 3–14.
Somervuo, P., Kvist, J., Ikonen, S., Auvinen, P., Paulin, L., Koskinen, P. *et al.* 2014. Transcriptome analysis reveals signature of adaptation to landscape fragmentation. *PLoS One* **9**: e101467.
Stillwell, R.C., Blanckenhorn, W.U., Teder, T., Davidowitz, G. & Fox, C.W. 2010. Sex differences in phenotypic plasticity affect variation in sexual size dimorphism in insects: from physiology to evolution. *Annu. Rev. Entomol.* **55**: 227–245.
Tong, Y. & Kanost, M.R. 2005. *Manduca sexta* serpin-4 and serpin-5 inhibit the prophenoloxidase activation pathway: cDNA cloning, protein expression, and characterization. *J. Biol. Chem.* **280**: 14923–14931.
Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. *et al.* 2012. Primer3–new capabilities and interfaces. *Nucleic Acids Res.* **40**: e115.
Vera, J.C., Wheat, C.W., Fescemyer, H.W., Frilander, M.J., Crawford, D.L., Hanski, I. *et al.* 2008. Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. *Mol. Ecol.* **17**: 1636–1647.
Wahlberg, N. & Saccheri, I. 2007. The effects of Pleistocene glaciations on the phylogeography of *Melitaea cinxia* (Lepidoptera: Nymphalidae). *Eur. J. Entomol.* **104**: 675–684.
Wang, R., Ovaskainen, O., Cao, Y., Chen, H., Zhou, Y.A.N., Xu, C. *et al.* 2011. Dispersal in the Glanville fritillary butterfly in fragmented versus continuous landscapes: comparison between three methods. *Ecol. Entomol.* **36**: 251–260.
Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R. *et al.* 2010. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**: 565.
Zhan, S., Huang, J., Guo, Q., Zhao, Y., Li, W., Miao, X. *et al.* 2009. An integrated genetic linkage map for silkworms with three parental combinations and its application to the mapping of single genes and QTL. *BMC Genom.* **10**: 389.
Zhao, P., Wang, G.-H., Dong, Z.-M., Duan, J., Xu, P.-Z., Cheng, T.-C. *et al.* 2010. Genome-wide identification and expression analysis of serine proteases and homologs in the silkworm *Bombyx mori*. *BMC Genom.* **11**: 405.
Zheng, Y.P., He, W.Y., Beliveau, C., Nisole, A., Stewart, D., Zheng, S.C. *et al.* 2009. Cloning, expression and characterization of four serpin-1 cDNA variants from the spruce budworm, *Choristoneura fumiferana*. *Comp. Biochem. Physiol. B-Biochem. Mol. Biol.* **154**: 165.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Table S1 Numbers of samples representing different development stages and tissues as well as the percentages of different tissues in pooled RNA from the Åland Islands (Finland) and Nantaizi (China) populations.

Table S2 List of genotyped markers with their position in the genomic contigs and genotyping performance.

Table S3 Numbers of individuals selected for genotyping in the two temperature treatment groups and in the two sexes.

Table S4 Numbers of SNPs in the two populations and non-synonymous SNPs within the Åland Islands and in Åland vs. Nantaizi populations.

Table S5 Results of association analyses for the additive model with sex and temperature treatment and their interaction with SNPs as fixed effects and network as a random effect.

Table S6 Numbers of individuals from the 12 habitat patch networks in the Åland Islands.

Table S7 Numbers of individuals in different genotypes and treatment/sex groups, and the genotype frequencies for three SNPs.

Figure S1 Flowchart for the procedure to identify candidate genes and SNPs.

Figure S2 The habitat patch network in the Åland Islands.

Figure S3 Coverage of the alignment in (a) target proteins and in (b) query sequences.

Figure S4 Venn diagram for the number of overrepresented GO groups ($P < 0.01$) in the gene sets with reduced variation in Åland, large numbers of non-synonymous SNPs and SNP clusters.

Figure S5 The frequency distribution of the minor allele in six SNPs is the 12 semi-independent patch networks in the Åland Islands.

Figure S6 Boxplots of larval growth rate and development time between the 6^{th} and 7^{th} instars, pupal weight and adult lifespan in the 12 semi-independent patch networks in the Åland Islands.

Figure S7 Pupal weight in three closely linked SNPs (a) c50_est:735A>G, (b) c50_est:816C>A and (c) c50_est:824A>G of hemolymph proteinase 5 (mean ± SE).

Data deposited at Dryad: doi: 10.5061/dryad.s57k2

Received 13 March 2015; revised 30 June 2015; accepted 14 August 2015