MOLECULAR ECOLOGY

Molecular Ecology (2013) **22**, 602–619 doi: 10.1111/j.1365-294X.2012.05521.x

Temperature treatments during larval development reveal extensive heritable and plastic variation in gene expression and life history traits

JOUNI KVIST,*†[1], CHRISTOPHER W. WHEAT,†[1] EVELIINA KALLIONIEMI,†‡ MARJO SAASTAMOINEN,† ILKKA HANSKI† and MIKKO J. FRILANDER*

**Institute of Biotechnology, PL 56, Viikinkaari 9, 00014 University of Helsinki, Helsinki, Finland, †Department of Biosciences, PL 65, Viikinkaari 1, 00014 University of Helsinki, Helsinki, Finland*

Abstract

Little is known about variation in gene expression that affects life history traits in wild populations of outcrossing species. Here, we analyse heritability of larval development traits and associated variation in gene expression in the Glanville fritillary butterfly (*Melitaea cinxia*) across three ecologically relevant temperatures. We studied the development of final-instar larvae, which is greatly affected by temperature, and during which stage larvae build up most of the resources for adult life. Larval development time and weight gain varied significantly among families sampled from hundreds of local populations, indicating substantial heritable variation segregating in the large metapopulation. Global gene expression analysis using common garden-reared F2 families revealed that 42% of the >8000 genes surveyed exhibited significant variation among families, 39% of the genes showed significant variation between the temperature treatments, and 18% showed a significant genotype-by-environment interaction. Genes with large family and temperature effects included larval serum protein and cuticle-binding protein genes, and the expression of these genes was closely correlated with the rate of larval development. Significant expression variation in these same categories of genes has previously been reported among adult butterflies originating from newly established versus old local populations, supporting the notion of a life history syndrome put forward based on ecological studies and involving larval development and adult dispersal capacity. These findings suggest that metapopulation dynamics in heterogeneous environments maintain heritable gene expression variation that affects the regulation of life history traits.

Keywords: Glanville fritillary, heat shock proteins, heritability, larval growth, larval serum proteins, microarray

Received 19 October 2011; revision received 16 January 2012; accepted 20 January 2012

Introduction

Ecological geneticists and evolutionary biologists seek to identify the mechanisms that maintain genetic variation affecting life history traits in natural populations (Wade & Goodnight 1998; Feder & Mitchell-Olds 2003; Ellegren & Sheldon 2008; Walsh & Blows 2009). One research approach is to combine ecological and candidate gene studies (Watt & Dean 2000; Hanski & Saccheri 2006; Barrett *et al*. 2008; Slate *et al*. 2010). However, candidate gene approaches necessarily present a limited view of genetic variation affecting life history traits, and research in this field is now moving to genome-wide studies (Turner *et al*. 2010). There is much potential in conducting genomic studies on natural populations of ecologically well-known species.

Correspondence: Jouni Kvist, fax: +358-9-19159366;
E-mail: jouni.kvist@helsinki.fi
‡Present address: School of Environmental Sciences, NR4 7TJ, Norwich, University of East Anglia, UK.
[1]Indicates shared first authorship.

The Glanville fritillary butterfly (*Melitaea cinxia*) metapopulation in Finland is an exceptionally well-studied ecological model system (Hanski 1999, 2011; Nieminen *et al.* 2004). A series of studies focused on the gene phosphoglucose isomerase (*Pgi*) has revealed strong associations between molecular variation at this gene and measurements of individual performance and fitness components (Saastamoinen & Hanski 2008; Klemme & Hanski 2009; Niitepõld *et al.* 2009; Orsini *et al.* 2009; Saastamoinen *et al.* 2009). There is systematic variation in the *Pgi* allelic composition among local populations, which is related to the spatio-temporal dynamics of these populations (Hanski & Saccheri 2006; Hanski *et al.* 2006), and strong coupling between demographic and microevolutionary dynamics at the *Pgi* gene (Zheng *et al.* 2009; Hanski 2011). Furthermore, DNA sequence analysis and coalescent simulations have indicated long-term balancing selection at *Pgi* in this metapopulation (Wheat *et al.* 2010). Another metabolic gene, succinate dehydrogenase d (*Sdhd*), exhibits comparable spatial variation in indel polymorphism among local populations (Wheat *et al.* 2011), and a more global analysis showed variation between newly established and old local populations in the expression of genes involved in egg provisioning and the maintenance of flight muscle proteins (Wheat *et al.* 2011). In short, previous studies have shown that there is considerable genetic variation in the natural populations of the Glanville fritillary with large effects on adult life history. Here, working with material from the same metapopulation, we examine global gene expression variation in larval development under a range of naturally occurring temperatures.

The thermal environment is critically important for the larvae of the Glanville fritillary in the Åland Islands, at the northern range limit of the species (Kuussaari 1998). The larvae are gregarious, living in large groups of around 100 full sibs and overwintering at the 5th larval instar (Nieminen *et al.* 2004). Diapause is broken in late March, after which the larvae go through 3 more instars. This study is focused on the final-instar, during which the larvae gain most of their ultimate weight and build up resources for the adult stage. In April, when the final-instar larvae develop, ambient air daily temperatures are low, often around 10 °C. In cloudy weather, larval body temperature is not much higher than the ambient temperature, but in sunny weather, the black larvae bask in the sun and can raise their body temperature up to and above 30 °C (Kuussaari 1998). The thermal conditions that affect larval growth and development thus vary dramatically from day to day and among the years. In butterflies in general, larval survival depends critically on their performance under suboptimal conditions (Kingsolver 2000; Kingsolver & Gomulkiewicz 2003; Kingsolver *et al.* 2004; Kingsolver & Huey 2008), and the Glanville fritillary is no exception (Kallioniemi & Hanski 2011). Assuming that there is heritable genetic variation affecting the survival and growth of the larvae, temporal and spatial variation in environmental conditions may be maintaining such genetic variation in the wild.

To characterize the genetic and thermal effects on larval development in the Glanville fritillary, we analysed the performance of full-sib larval families originating from multiple local populations in the metapopulation in the Åland Islands in Finland and measured their heritabilities. Larvae were exposed to three different temperature treatments that bracket the thermal conditions experienced by larvae in the field. 'cold' represents cool spring weather conditions with little sunshine. Larval mortality is elevated, yet these conditions allowed most larvae to complete their development (Kallioniemi & Hanski 2011). 'hot' corresponds to sunny and warm spring weather conditions, allowing larvae to bask in the sun and thereby raise their body temperature above 30 °C (Kuussaari 1998). 'standard' represents intermediate conditions and corresponds to standard rearing conditions in previous experiments (Kallioniemi & Hanski 2011). Global gene expression was measured using a high-density microarray. We found that there is a high level of expression variation among larval families, comparable to variation because of the temperature treatment. Many of the differentially expressed genes belong to the same functional categories, notably the larval serum proteins, which were previously identified as being differentially expressed between newly established and old local populations in adult female butterflies (Wheat *et al.* 2011). These findings suggest the possibility of a common genetic basis for variation in larval development and adult dispersal.

Materials and methods

Phenotypic variation in larval development

Phenotypic variation in postdiapause development was analysed for three randomly chosen larvae from each larval family group that was detected during an Åland-wide survey in the fall 2009. The larvae were reared individually under common garden conditions (+28:15 °C; 12:12, L/D) in the spring 2010. Individuals were weighed when the diapause was broken (5th instar) and at the beginning of each following instar (6th, 7th and pupal stage). In the beginning of the 6th instar, larvae were randomly assigned to two different rearing temperature treatments (standard: +28:15 °C 12:12, L/D; and warm: +32:15 °C 12:12, L/D). After eclosion, butterflies were sexed, marked individually and

kept under standard conditions (+26:18 °C 9:15, L/D) and fed daily (honey/water solution 1:3) to measure their lifespan.

The results were analysed with principal component analysis (PCA) on JMP (version 8.0.2) separately for the two rearing temperatures, using only those individuals for which information was available for all the life history traits (n = 2018 individuals from 1153 families in 436 local populations). Pearson correlations were calculated among the developmental stages for body mass and development time separated for the sexes and the treatments. Broad-sense heritability was estimated as twice the intra-class correlation coefficient (among family variance divided by total variance; Falconer & Mackay 1996) in a nested ANOVA with family nested within population and individuals nested within family, and source population as a random factor. Heritability was calculated for adult lifespan, development time and body mass (7th instar larvae and pupae) in both treatments by first standardizing for sex-specific differences.

In-depth analyses of heritable and plastic responses to rearing temperature

Heritable differences and plastic responses were studied further using second-generation, laboratory-reared larvae from six different larval families (O44, O171, N32, O145, N170 and N74), initially collected from the Åland metapopulation. The larvae were reared in standard laboratory conditions until the 6th instar. At the 7th instar, ten individuals from each family were exposed to different thermal conditions (standard, cold and hot; see Fig. 1 for details) that mimicked diurnal thermal variation in their natural habitat. Of the three conditions, the cold and hot conditions mimic the extremes of the temperature profile that the larvae are subjected to during a cool and cloudy spring or a hot and sunny spring, respectively, while the standard condition represents the average temperature profile in the Åland Islands.

Development times (last moult date, pupation date, eclosion date), survival (death date and death stage) and body weight (larval weight at the beginning of the 7th instar and pupal weight) were recorded. These were used to analyse the developmental rate (larval and pupal duration) and weight increase (growth and relative growth rate) using a mixed model in JMP (version 7.0.1), with the response variables explained by family, treatment and their interaction (fixed factors). Sex and last moult date were used as random factors. Sex could reliably be determined only for adult butterflies. The samples that died as larvae or pupae were assigned a sex using gender-specific differences in larval weight and last moult date (see Fig. S1 for details).

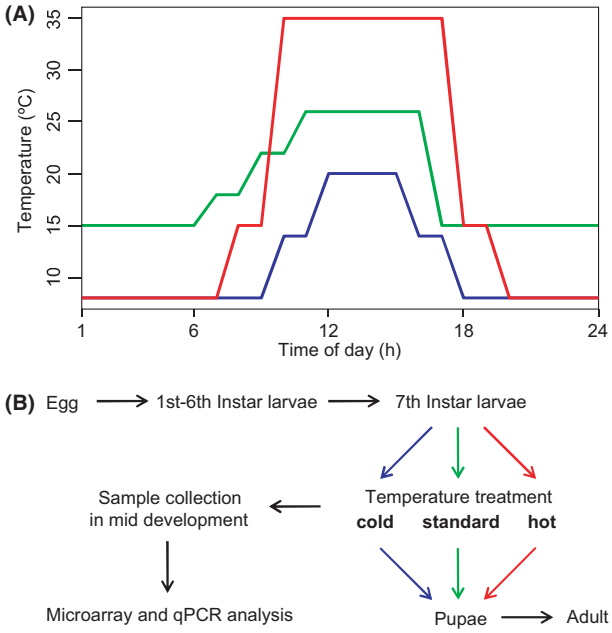

Fig. 1 Experimental design. (A) The diurnal temperature profiles for the treatments (red = hot, green = standard, blue = cold). (B) Overview of the experiments, showing how second-generation laboratory-reared Glanville fritillary larvae were reared under controlled conditions until they reached the final-instar, at which time subsets of the larvae were subjected to three different temperature treatments. During mid-development, RNA was isolated from a subset of larvae and used for the microarray and qPCR analysis. The rest of the larvae were allowed to develop into adult butterflies. Phenotypic measurements were taken during the development (at last moulting, pupation and eclosion) and the stage at which death occurred was recorded.

Survival was assessed using parametric survival analysis, as implemented in JMP/Genomics using Weinbull distribution, with family, treatment and their interaction accounting for death stage (adult, handicapped adult, pupae and 7th instar larvae). A separate analysis was carried out for the individuals that survived to adulthood (adult survival) using death date as the response variable. The effect of family (broad-sense heritability) on the pre-treatment traits, such as development time (last moult date) and body mass (larval weight), was estimated using full-sib regression (Roff 1997, pg 55, eq. 2.47).

Sample preparation and preprocessing

Larvae from the F2 families were collected, snap-frozen in liquid nitrogen during mid-development (after 4, 5 and 6 days, for hot, standard and cold, respectively) and used for RNA extraction (Fig. 1B). Three of these families were chosen for the microarray experiment,

and the rest were kept for qPCR validation (Table S1). Sample preparation, hybridization, scanning and data extraction were performed as described by Wheat *et al.* (2011). Spot quality was examined by visual inspection, and bad spots (containing dust, etc.) were removed from subsequent analysis. Median intensities of the spots were read using R/Bioconductor/Limma-package (Smyth 2004) and $\log_2$-transformed. Probes were removed if they were consistently (in 90% of the arrays) low (< 2 SD from the average intensity of the 'DarkCorner' spots) or high (<2 SD from saturation). The filtering was performed separately for both channels. Filtered data were normalized using loess and quantile method for within- and between-array normalization, respectively, without background subtraction, using the default setting in Limma, as recommended for Agilent arrays (Zahurak *et al.* 2007).

Microarray analysis

The preprocessed data were analysed using a mixed model (Wolfinger *et al.* 2001) as implemented in JMP/Genomics (SAS 9.1.3, JMP 7.0.1, genomics 3.1). Dye, family, treatment and family-by-treatment interaction were treated as fixed factors. Slide, array, spot, spot by array and biological samples were treated as random factors. Least squared means and *P* values were calculated for family, treatment and their interaction. PCA was performed on mean probe intensities (centred to zero and variance scaled to one) using Pearson's correlation for each array, as implemented in JMP/Genomics. The three main components were fitted to a linear model with family, treatment and their interaction as fixed effects to assess the significance of these factors in explaining the observed variation. Additionally, a discriminant analysis as implemented by JMP/Genomics for family was conducted to assess to what extent the main factors could classify the samples to the correct family.

The false discovery rate (FDR) was calculated using qvalue-package (Storey & Tibshirani 2003) in R, with default settings. *Q* value estimates the number of expected false positives at different cut-off levels in the data set and is based on the distribution of the *P* values. Probes were considered differentially expressed at two levels of stringency. Subset 1 had probes with $P < 0.05$, and subset two had probes that should contain no expected false positives, based on the *Q* value analysis (*Q* value * *n* probes <1).

Enrichment analysis

Expression (least squared mean) difference of all probes was calculated between the pairwise comparisons in the data (between families, treatments, families within the same treatment and treatments within the same family) and averaged for each transcript (contig from the transcriptome assembly). This gene list, ranked by expression difference, was used to find over- and under-represented GO and KEGG categories with Fatiscan (Al-Shahrour *et al.* 2005) in Babelomics (Al-Shahrour *et al.* 2006), using a two-tailed Fisher's exact test with 30 partitions, limiting the analysis to the input genes and using noninclusive GO levels with direct annotations. As our transcripts are not annotated, we used *Drosophila melanogaster* gene information instead. As our transcripts are mostly partial gene sequences, we retrieved first the closest matching (BLASTx) genes in *Bombyx mori*, which were then used to identify (BLASTp) the *Drosophila melanogaster* genes (Vera *et al.* 2008).

The categories thus identified were further characterized for direction and magnitude of expression change at the probe level, by fitting the standardized least squared means of family-by-treatment interactions into a mixed model in JMP. An automated K-means clustering was performed for the probes belonging to the category *response to unfolded protein (GO:0006986)* using the standardized least squared mean expression for treatment, with clustering radius of 1.5, as implemented in JMP/Genomics. Probes belonging to the categories *larval serum protein complex (GO:0005616)* and *oxygen transporter activity (GO:0005344)* were analysed for co-regulation by fitting the preprocessed intensities of each probe against the intensities of probes belonging to different LSP genes (unique contig). The average LSP expression was then fitted against the expression of all other probes to find additional genes displaying co-expression.

qPCR validation

Primers (Table S2) were designed with Primer3 (Rozen & Skaletsky 1999) for fourteen genes that showed significant expression differences between the treatments ($P < 8.6 \times 10^{-06}$–6.3×10^{-09}), with a wide range of expression levels ($\log_2$ intensity = 8.88–12.52) and five endogenous control genes ($\log_2$ intensity = 8.57–12.23). The qPCR measurements were made with LightCycler® 480 Real-Time PCR System (Roche) using LightCycler® 480 SYBR Green I Master mix (Roche) in a 384 well plate with 10 μL volume (containing 2.5 μL of cDNA). cDNA and premade mastermix were pipetted using CAS-1200 Liquid Handling System (Corbett Life Science). The samples were partially randomized across the plate to minimize position-dependent differences between the measurements. Six biological samples and a dilution series of one control gene of a standard

sample (interplate calibrator) were measured on one plate for all genes, with three technical replicates and one water control for each gene.

Fluorescence values were converted to melting curves, crossing point (DCt) values and estimated concentrations using qpcR-package (Ritz & Spiess 2008) in R. Fluorescence values were normalized and polished for lag and plateau phases and fitted to a sigmoidal curve (using modlist with baro5 as model). The concentrations were estimated using calibration curves for each gene on a reference plate (calib, cpD2 as threshold). Melting curve analysis was used to remove deviant samples (primer dimers and incorrect amplicons). Concentrations of target genes were divided by the average concentration of the endogenous control genes. The normalized concentrations were analysed for treatment, family and their interactions using a mixed model in JMP/Genomic, with plate ID as a random factor. Significance of treatment and expression differences between hot and cold were compared with the microarray results (least squared mean of expression) using five sample sets in the qPCR (Table S1).

Results

Genetic background interacts with temperature in affecting larval development

We used a comprehensive field-collected sample originating from 436 different local populations to investigate larval development and survival in two rearing temperatures, including correlations among the developmental stages and broad-sense heritabilities of the respective traits. These traits were next investigated in F2 common garden-reared full-sib larvae (Kallioniemi & Hanski 2011) in three rearing temperatures (Fig. 1A). Finally, we selected out of this material three larval families, representing the full range of variation, for a study of global gene expression potentially affecting the phenotypic traits (Fig. 1B).

Development times, body masses at different larval stages and survival during postdiapause development were analysed with the field-collected sample using principal component analysis (PCA). Majority of the total variation (57% in standard and 64% in warm) in these traits could be accounted for by the first three principal components (PCs) of the analysis (Table S3). All development times and body masses were correlated with PC1 and PC2, with measures of body mass correlating positively with both PC1 and PC2, while measures of development time correlated negatively with PC1 and positively with PC2. These correlations probably reflect the existence of both negative (PC1) and positive (PC2) correlations between body mass and development time, supporting, respectively, condition-dependent development and physiological constraints on development. In the former, high-quality individuals grow fast to a large size (hence the negative correlation), and in the latter, more time is needed to achieve large body size (hence the positive correlation). In contrast, lifespan was uncorrelated with the other traits and correlated only with PC3. PC1 and PC2 showed significant broad-sense heritability (Table S3), suggesting genetic effects throughout postdiapause larval development.

Correlations among individual traits were small but significant (Table S4). Development time and body mass had the highest broad-sense heritability during the 7th instar and pupal stage, respectively (Table 1), and were negatively correlated ($\rho = -0.45$; $P < 0.0001$) during the 7th instar (Table S4). Phenotypic variance among the local populations was uniformly low and less than variance among individuals and families (Table 1), consistent with high gene flow among nearby local populations (Hanski *et al.* 1994), low F_{st} (Saccheri *et al.* 2004) and coarse-grained population structure in the Glanville fritillary metapopulation (Orsini *et al.* 2008).

The large amount of heritable variation in the developmental traits was examined in detail in six common garden-reared families exposed to three temperature treatments (cold, standard and hot) during the 7th instar. Prior to the 7th instar and the temperature treatment, there were significant differences among the families (Fig. S1, Table S5) in development time ($F_5 = 11.87$; $P < 0.0001$) and body mass ($F_5 = 11.97$; $P < 0.0001$), with broad-sense heritabilities of 0.54 (standard error 0.19) and 0.53 (0.19), respectively. These findings are similar to the values of the field sample (Table 1). These differences persisted during the temperature treatment, although the ranking of the families in development time changed from pre-treatment to post-treatment period (Table S5). This was especially pronounced in family N74, which had the fastest post-treatment development for larvae ($F_1 = 8.41$; $P = 0.0045$) and pupae ($F_1 = 10.76$; $P = 0.0015$) (Fig. 2A, Table S5).

There was a highly significant temperature treatment effect on all post-treatment traits (Table S5), but the interaction between family and treatment was significant only for survival and larval duration ($F_{10} = 2.265$; $P = 0.018$), which decreased with increasing temperature and varied greatly among the families in cold but had only limited variation in hot (Fig. 2A). Pupal weight also increased linearly with temperature, but pupal duration was unexpectedly shortest in the standard treatment ($F_1 = 37.23$; $P < 0.0001$). The differences were small, 1 day or less, but significant among the families ($F_4 = 4.21$; $P = 0.0036$) and the treatments

Table 1 Percentage of variation in phenotypic traits that is explained by family within population, population and residual, and the corresponding broad-sense heritability estimates

	Development time		Body mass		Lifespan
	7th instar	Pupae	7th instar	Pupae	Adult
standard temperature					
Family (population)	0.278	0.069	0.122	0.243	0
Population	0.051	0.02	0.068	0.073	3.4
Residual	0.671	0.911	0.81	0.684	97.6
H^2	0.56^{***}	0.14^{NS}	0.24^{**}	0.49^{***}	0^{NS}
warm temperature					
Family (population)	0.359	0.336	0.184	0.395	0.004
Population	0.151	0	0	0	0.028
Residual	0.49	0.664	0.816	0.605	0.968
H^2	0.72^{***}	0.67^{**}	0.37^{NS}	0.79^{***}	0^{NS}

The traits have been standardized for the sex differences. Results are given separately for the two temperature treatments (standard and warm) and developmental stage (7th instar larvae and pupae for development time and body mass and adult for lifespan). $^{***}P < 0.0001$, $^{**}P < 0.001$, $^{NS}P > 0.05$.

($F_2 = 27.39$; $P < 0.0001$). Pre-adult survival was affected by the temperature treatment ($F_2 = 67.07$; $P < 0.0001$), family ($F_5 = 45.06$; $P < 0.0001$) and their interaction ($F_{10} = 44.03$; $P < 0.0001$). Mortality was highest in cold, whereas in hot, all individuals survived to adults (Fig. 2B). In contrast to larval survival, adult lifespan was affected by neither temperature nor family (both $P > 0.1$).

In summary, these results indicate largely correlated variation and significant broad-sense heritabilities across the different larval developmental stages. There are highly significant family effects in the 7th instar development and survival as well as a significant family-by-temperature interaction, suggesting that substantial genetic variation affecting larval development is segregating in the metapopulation.

Probe-level analysis shows that both the family background and temperature treatment have significant effect on global gene expression

To analyse variation in gene expression among the families and between the temperature treatments, we sampled a total of 35 larvae from three families (N170, N74 and O171) during their mid-development in the 7th instar. The selected families represented the entire range of larval development times, especially in the cold treatment, with family O171 being the slowest and N74 the fastest (Fig. 2A). RNA samples were hybridized to custom-made microarrays (Vera *et al.* 2008). The technical performance of the array was very high. The coefficients of variation (CV) were less than 3.4% for 95% of the preprocessed triplicate probes ($n = 11571$; Fig. S2). The three families and the two temperature extremes (cold and hot) were well separated in a principal component analysis (Fig. 3, Table 2). The first three components explained 31% of all expression variation and were significantly ($P < 0.0001$) correlated with family (PC1 and PC3) and treatment (PC2). A discriminant function analysis showed that 31 of the 35 individuals could be assigned to the correct family based on these PC values (χ^2 likelihood ratio = 37.6; $P < 0.0001$).

We used a mixed model analysis to identify the effects of family, treatment and their interaction on gene expression. The analysis showed substantial and significant expression variation because of all these factors, as indicated by the excess of low P values (Fig. S3). One-third of all probes showed significant ($P < 0.05$, Table S6 subset 1, Fig. 4) differences among the families (37%; $n = 4327$) and treatments (35%; $n = 4121$). Family-by-treatment interaction was significant for fewer probes (15%; $n = 1770$) and had four times higher false discovery rate (Q value). The significant probes from the three levels of analysis represent a majority ($n = 5622/8397$) of the unique transcripts on the array, with 42% being significant for family, 39% for treatment and 18% for their interactions.

Families N170 and N74 had the largest expression difference in 55% of the probes that were significant for family, whereas O171 and N74 had the smallest expression difference in 49% of these probes. The treatments hot and cold differed most from each other for 67% of the probes that were significant for treatment, while standard and cold were most similar for 53% of these probes. These results reflect the patterns in the principal component analysis (Fig. 3). Variation among the families for the probes that were significant for interaction

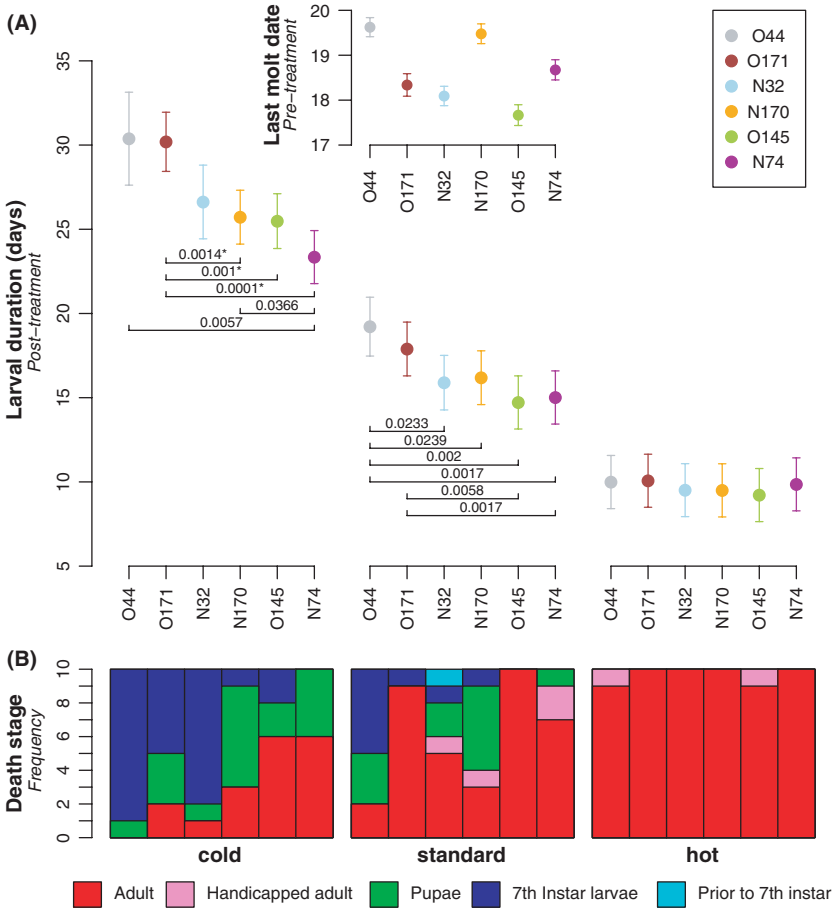

Fig. 2 Larval performance. (A) Pre- and post-treatment development times in the different larval families (means and standard errors). Main panel: The durations of the final (7th) larval instar during the temperature treatments are shown for each treatment and family. Data include only those individuals that survived to pupation. Significant *P* values are shown for pairwise comparisons of families within each treatment. The *P* values indicated with a star were significant ($P < 0.003$) after Bonferroni correction (multiple tests within a treatment group, $n = 15$). Inset: The larval moult date prior to the temperature treatments in each family. (B) Proportions of individuals that died at each life history stage across the six families and the three temperature treatments. The colours relating to different death stages are indicated below the graph. 'Handicapped adults' refer to individuals that died immediately after eclosion. Note that families are given in the same order as in panel (A).

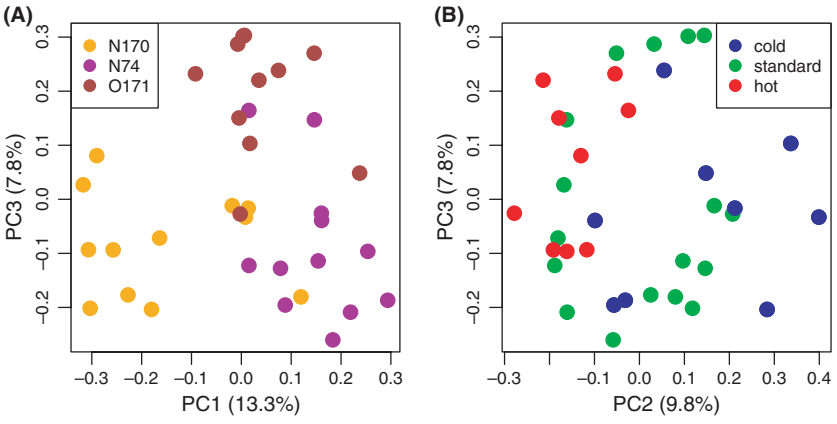

Fig. 3 Principal component analysis of the microarray expression data (post-normalization). (A) The three families are separated by principal components 1 and 3. (B) The temperature treatments hot and cold, but not the standard, are separated by principal components 2 and 3. The percentage of total variation explained by each PC is indicated in the axis titles. Families and treatments are identified in the insets.

was highest in hot (62% of probes), followed by cold (30%) and standard (9%).

After limiting the data to include less than one expected false positive (Table S6 subset 2), there were still 908 probes showing differential expression among the families, 778 among the treatments but only 57 showing a significant family-by-treatment interaction (Fig. 4). Although these probes represent transcripts with the most reliable expression changes, most of these changes are small, making it difficult to detect them with less sensitive methods (e.g. qPCR). For a more robust list of candidate genes for the family-by-treatment interaction, we selected transcripts with more than one significant ($P < 0.05$) probe, which should reduce false positives and cross-hybridization noise. Of these transcripts, we selected only the ones with greater than

Table 2 The percentage of total variation explained by the first three principal components in the microarray analysis

		Summary of fit		*P* values		
Principal components	% of total variation	R^2	R^2 adjusted	Family	Treatment	Family-by-treatment
PC1	13.3	0.6961	0.6027	<.0001	0.0466	0.6133
PC2	9.8	0.6355	0.5234	0.017	<.0001	0.0806
PC3	7.8	0.6559	0.5500	<.0001	0.2503	0.6697

Each principal component was fitted to a linear model with family, treatment and family by treatment as fixed factors. The R^2 values of the fits and the corresponding *P* values are given in the following columns.

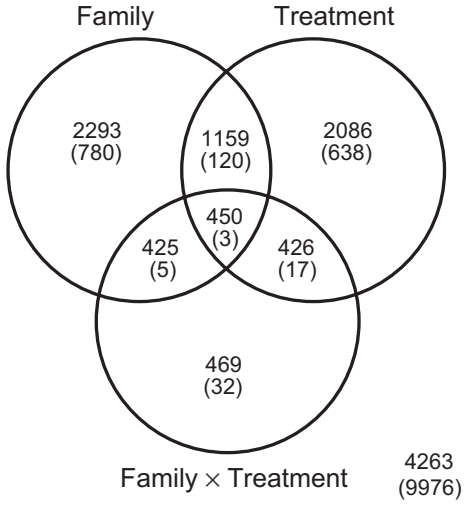

Fig. 4 Venn diagram showing the number of statistically significant probes at two cut-off levels. Subset 1, the upper value, contains probes with *P* value < 0.05. The corresponding proportions of expected false-positive probes (*Q* value) at this level are 0.0439, 0.0486 and 0.1730 for family, treatment and their interactions, respectively. The values in parentheses indicate the more stringent subset 2, in which the number of expected false-positive probes is less than one and contains probes with *Q* value < 0.0008, 0.0010 and 0.0108 for family, treatment and their interactions, with corresponding *P* values < 0.0002, 0.0002 and 0.0001, respectively. The number of non-significant probes for both subsets is shown in the bottom right.

twofold expression difference among the families in at least one of the temperatures, which should facilitate easier detection. This resulted in a list of 79 unique transcripts (Table S6 candidate gene list).

In brief, our results document widespread variation in expression of a large number of genes among families and the temperature treatments. As expected, the temperature treatment extremes hot and cold were the most different from each other. Even though approximately an equal number of genes were found to show significant variation among the families and between the treatments, only 15–35% of them (depending on the statistical cut-off value) are shared.

Larval serum proteins and translation are the top categories indentified by gene set enrichment analysis

With so many differentially expressed genes, we proceeded to classify them with gene set enrichment analysis tools (Al-Shahrour *et al.* 2005) using gene ontology (GO) and pathway (KEGG) categories. We carried out pairwise comparisons between families, treatments, families within each treatment and treatments within each family (Table S7). This resulted in 57 significantly enriched categories (FDR adjusted $P < 0.0001$ in at least one of the comparisons), which have been ordered by similarity in Table 3. Of these categories, 45 were different ($P < 0.01$) among treatments and 27 among the families. The rest were significant for families only in a specific treatment or significant for treatments in a specific family. Thirty-seven (37) categories displayed the greatest variation among families in cold, substantially more than in hot (28) and in standard (18). The magnitude and direction of change was analysed at the probe level (Fig. S4).

The most striking enrichments were observed in the categories *larval serum proteins* (Table 3A) and *translation* (Table 3B). *Larval serum proteins* showed an overall increase in expression with increasing temperature and significant differences among families. The expression correlated with larval development time across the temperature treatments (Fig. 5) and exhibited the same family-by-treatment interaction as larval duration: larvae in the family N74 had higher expression level in cold and hot than larvae in N170, but in standard, the expression levels were equal. Family O171 had the lowest expression level in all treatments. In contrast, *translation* showed an overall decrease in expression with increasing temperature. In family O171, this decrease was monotonic, whereas in N74, there was a slight increase in expression from standard to hot. Family N170 had the lowest expression level in all conditions, and identical expression in standard and hot.

Six categories, *actin binding*, *striated muscle contraction*, *valine, leucine and isoleucine degradation*, *citrate cycle*, and *ascorbate and aldarate metabolism*, were enriched

Table 3 Significantly enriched GO and KEGG categories in the gene set enrichment analysis

				Treatment within family				Family within Treatment		
	Category	Description	Treatment	N170	N74	O171	Family	cold	standard	hot
A)	GO:0005616	Larval serum protein complex	3	3	3	3	3	3	3	3
	GO:0005344	Oxygen transporter activity	3	3	3	3	3	3	3	3
B)	GO:0005840	Ribosome	3	3	3	3	3	3	3	3
	GO:0003735	Structural constituent of ribosome	3	3	3	3	3	3	3	3
	GO:0030529	Ribonucleoprotein complex	3	3	3	3	3	3	3	3
	GO:0006412	Translation	3	3	3	3	3	3	3	3
	dme03010	Ribosome	3	3	0	2	3	3	3	3
C)	GO:0006941	Striated muscle contraction	0	2	0	0	2	3	0	2
	dme00020	Citrate cycle (TCA cycle)	0	2	0	0	2	3	0	0
	dme00280	Valine, leucine and isoleucine degradation	0	2	2	0	2	3	1	1
	dme00053	Ascorbate and aldarate metabolism	0	0	1	1	2	3	1	0
	GO:0003779	Actin binding	0	2	2	0	3	2	0	3
D)	GO:0009408	Response to heat	0	2	0	1	3	1	1	3
	GO:0006986	Response to unfolded protein	3	2	3	2	1	1	0	0
E)	GO:0006457	Protein folding	3	2	1	0	0	0	0	0
	GO:0042802	Identical protein binding	3	2	1	2	0	0	0	0
F)	GO:0048037	Cofactor binding	2	0	3	0	0	0	0	0
G)	GO:0016155	Formyltetrahydrofolate dehydrogenase activity	3	1	3	2	0	0	0	0
	GO:0009258	10-formyltetrahydrofolate catabolic process	3	1	3	1	0	0	0	0
	dme00670	One carbon pool by folate	1	0	3	0	0	0	0	0
H)	GO:0005739	Mitochondrion	3	0	3	2	2	0	0	3
	GO:0015992	Proton transport	3	2	3	3	1	2	0	0
	GO:0016469	Proton-transporting two-sector ATPase complex	3	2	3	3	0	1	0	1
	GO:0015078	Hydrogen ion transporter activity	3	2	3	3	0	1	0	0
	GO:0046933	Hydrogen ion transporting ATP synthase activity, rotational mechanism	3	1	3	2	0	0	0	0
I)	GO:0006508	Proteolysis	2	0	3	0	0	2	0	3
	GO:0008233	Peptidase activity	2	2	3	0	0	3	0	2
	GO:0004252	Serine-type endopeptidase activity	2	2	3	0	0	1	0	2
	GO:0004298	Threonine endopeptidase activity	1	2	3	0	0	2	0	0
	GO:0004175	Endopeptidase activity	2	2	3	0	2	3	0	2
	GO:0000502	Proteasome complex (sensu Eukaryota)	2	2	3	0	1	1	0	2
	GO:0005839	Proteasome core complex (sensu Eukaryota)	2	2	3	0	1	3	0	0

Table 3 *Continued*

				Treatment within family				Family within Treatment		
	Category	Description	Treatment	N170	N74	O171	Family	cold	standard	hot
	GO:0005838	Proteasome regulatory particle (sensu Eukaryota)	2	2	3	0	1	3	1	2
	GO:0016787	Hydrolase activity	3	0	3	0	0	2	1	2
J)	GO:0008061	Chitin binding	3	3	3	3	2	1	2	3
	GO:0042302	Structural constituent of cuticle	3	3	3	3	0	0	0	2
K)	GO:0016810	Hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds	3	3	3	3	2	1	1	0
L)	GO:0005975	Carbohydrate metabolic process	3	0	2	3	2	1	1	0
	dme00010	Glycolysis/Gluconeogenesis	3	3	1	3	0	2	0	0
	GO:0004179	Membrane alanyl aminopeptidase activity	1	0	2	3	1	2	2	0
M)	GO:0004556	Alpha-amylase activity	2	1	3	1	0	2	1	0
	GO:0016491	Oxidoreductase activity	2	2	3	2	0	3	2	2
N)	dme00071	Fatty acid metabolism	2	2	2	2	2	3	1	0
	dme00120	Bile acid biosynthesis	2	2	2	2	2	3	1	0
O)	dme00641	3-Chloroacrylic acid degradation	2	2	1	2	2	3	0	0
P)	GO:0007476	Imaginal disc-derived wing morphogenesis	2	0	3	2	1	2	2	0
Q)	GO:0005622	Intracellular	3	3	2	3	3	3	3	3
R)	GO:0006350	Transcription	3	1	3	2	2	3	2	0
	GO:0008270	Zinc ion binding	3	1	3	3	2	3	2	2
	GO:0003676	Nucleic acid binding	3	2	3	3	1	3	3	3
	GO:0003677	DNA binding	3	1	2	2	0	2	0	2
	GO:0006355	Regulation of transcription, DNA-dependent	2	0	2	3	1	3	0	0
	GO:0005634	Nucleus	1	2	3	2	2	3	2	2
S)	GO:0016020	Membrane	3	2	2	2	2	0	2	2
	GO:0016021	Integral to membrane	3	3	2	2	2	0	2	2
T)	GO:0005783	Endoplasmic reticulum	1	0	0	3	0	3	0	1
U)	GO:0005829	Cytosol	1	2	3	0	1	3	0	3

Significance is indicated with numbers from 3 to 0 (3: $P < 0.0001$, 2: $0.0001 < P < 0.01$, 1: $0.01 < P < 0.05$, 0: $P > 0.05$). Categories were grouped (from A to U) by shared genes or shared processes.

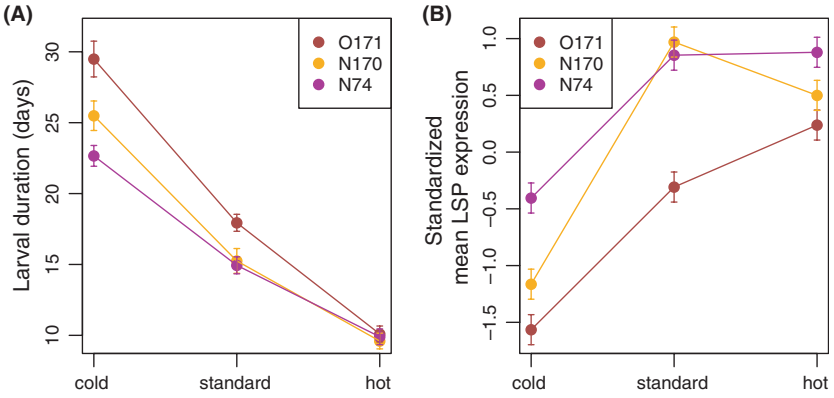

Fig. 5 Larval duration and the expression of larval serum proteins. (A) Duration of the final larval instar in the three families and the three treatments (means and standard error) for individuals that survived to adults. (B) The fitted standardized least squared mean expression and standard error for the probes belonging to the category *larval serum protein complex (GO:0005616)* in the three families and three temperature treatments. Families and treatments are identified in the inset.

exclusively among the families (Table 3C), with no enrichment between the treatments. Expression was highest in N74 and lowest in N170 in all cases. This difference was most apparent in the treatment extremes hot and especially in cold.

Of the two categories related to heat stress (Table 3D), *response to heat* showed consistent differences among the families (strongest in hot) but no uniform treatment response. In contrast, *response to unfolded protein* showed divergent treatment effects but little variation among the families (Fig. S5). Most probes showed a response in hot, either increased or decreased expression ($n = 15$ probes in both). The remaining probes showed a response in standard (increasing $n = 5$ and decreasing $n = 4$).

Categories enriched exclusively between the temperature treatments included *protein folding and binding* (Table 3E), *cofactor binding* (Table 3F) and *folate catabolism* (Table 3G). All these categories showed the highest expression in hot. *Folate catabolism* increased linearly with increasing temperature, while the other categories exhibited a more complex pattern.

Mitochondrial activity (Table 3H), *protein degradation* (Table 3I), *chitin biosynthesis* (Table 3J) and *energy metabolism* (Table 3K/L) categories were enriched primarily between the treatments. However, in many of these categories, family N74 was the only one to show clear temperature effects. The family N170 showed weak treatment effects in *mitochondrial activity* (Table 3H), O171 showed no treatment effects in *protein degradation* (Table 3I), and N74 showed decreased expression from cold to hot in both of these categories.

We excluded the possibility of annotation bias influencing the results by comparing the numbers of significant probes ($P < 0.05$) included in the enriched categories with those in the entire data set. The proportions were similar across all levels, family, treatment and their interaction ($\chi^2 = 6.48$; $P = 0.37$), indicating no unequal distribution of annotation among genes that were differentially expressed vs. those that were not.

In summary, while most of the probes in the mixed model analysis were significant either for the family or the treatment effects (Fig. 4), the enriched categories displayed complex interactions between the two factors. Treatment dominated the enrichment, but there was also family-level variation in almost all categories either as the main effect or only particular families showing the treatment effect.

qPCR validation

We used quantitative real-time PCR to assess the technical and biological robustness of the results for fourteen genes affected by the treatment (some with significant family and family-by-treatment interactions). These target genes spanned a wide range of expression levels (hybridization intensities). The expression levels were normalized with five endogenous control genes that showed limited expression variation and were neither significant for the treatments nor families, but which covered the same range of expression as the 14 focal genes. The results were analysed with a mixed model with multiple sample sets.

The repeatability of the results on expression was assessed using the same samples that were used in the microarray analysis (Table S1, 'microarray samples'; $n = 35$). Only 8 of the 14 genes were significant for treatment ($P < 0.05$) in the qPCR analysis, although all except one (T12) showed the same differences in expression among the treatments as the microarray (Fig. S6).

The microarray results were biologically robust, as similar qPCR results were obtained with independent samples. The three other families not included in the microarray experiment (Table S1, 'nonmicroarray families'; $n = 30$) had 7/14 genes significant for treatment, and samples not included in the microarray experiment from all six families (Table S1, 'nonmicroarray samples'; $n = 52$) had 9/14 genes significant for treatment (Fig. S6).

An increase in sample size increased the number of significant genes by allowing us to capture more of the technical plate to plate variation inherent in qPCR analyses. Analysis of all the samples (Table S1, 'all samples'; $n = 87$) in the thermal experiment (including the microarray samples) resulted in 10/14 significant genes. The most similar results to the microarray results, with 11/14 genes significant, were obtained with additional samples from the same three families that were used in the microarray experiment (Table S1, 'microarray families'; $n = 57$).

The technical performance of the qPCR was estimated with the coefficient of variation (CV) of the crossing point (DCt) values, after filtering out reactions with deviant kinetics (13.2%; 656/4959 removed). For 95% of the reactions, CV was <6.7%, almost twice the CV in the microarray experiment (3.4%) (Fig. S2). Thus, the higher sensitivity of the microarray was mostly responsible for its ability to detect significant treatment, family and interaction differences in expression with a smaller sample size.

In summary, the qPCR validation confirmed the direction of the expression differences but had lower level of repeatability. As a consequence, the sample size in the qPCR experiment had to be nearly doubled in comparison with the microarray experiment to reach comparable statistical power (Table S1).

Discussion

Our comprehensive field sample of larvae across the metapopulation indicated high heritability of larval development. There are thus the ingredients present for spatially and temporally varying natural selection on larval development, which could be expected to maintain genetic variation in the metapopulation. Consistent with this expectation, we found a large amount of variation in gene expression among larval families. This is a noteworthy finding in view of the paucity of studies that have quantified expression variation among families in natural outbred populations (Whitehead & Crawford 2006).

Although many studies have reported heritability values for gene expression, most studies were performed on model species with inbred strains used to generate 'family' crosses (Schadt *et al.* 2003; Gibson *et al.* 2004; Vuylsteke *et al.* 2005; Cui *et al.* 2006), and thus, these studies are not very informative about expression variation in wild populations. For example, a study using 40 common garden-reared and inbred *Drosophila melanogaster* lines from a single population found that 68% of the transcripts were differentially expressed among the lines, with broad-sense heritabilities ranging from 0.3 to 1.0 (Ayroles *et al.* 2009). However, these heritabilities are inflated compared to outbred populations (the additive genetic variation is doubled). Our results are broadly similar with a study on 15 human families from Utah, in which 31% of genes were differentially expressed among the families and showed significant narrow-sense heritability (parent–offspring similarity in expression) with a median of 0.34 (Monks *et al.* 2004).

Studies of natural outbred populations of killifish (*Fundulus heteroclitus*) have identified extensive expression differences among individuals, particularly in genes of central metabolism (Oleksiak *et al.* 2002, 2005). However, these studies lacked family structure and may have been affected by environmental and/or epigenetic factors (Scott *et al.* 2009). Additionally, most studies sample individuals from a single environment, not representative of natural conditions. Whether the actual conditions were stressful, and therefore increased expression variation, or benign and decreased expression variation, is impossible to assess given data from a single 'treatment', the environment that was sampled. In short, studies on several species indicate substantial family-level variation in gene expression, although the interpretation of the results is complicated by experimental conditions, inbreeding and possible epigenetic effects.

Expression variation among families interacts with temperature

Our experiment was designed to assess the amount of gene expression variation among families. Given the constraints on how many individuals could be included in the microarray experiment, and the fact that statistical power to detect differences among families decreases with decreasing number of individuals per family, we opted to analyse a large sample of individuals in three phenotypically divergent families. We used common garden F2 generation individuals in the microarray experiment to minimize maternal effects. As we investigated gene expression in a subset of families representing the range of phenotypic variation in the metapopulation, the observed expression variation reflects genetic variation segregating across the local populations residing in the heterogeneous environment. Previous studies have revealed a distinct life history syndrome characterizing new populations established by more dispersive individuals versus old populations (Hanski *et al.* 2006; Saastamoinen 2008; Wheat *et al.* 2011). Unfortunately, our sampling was too sparse to address the new vs. old population dichotomy, but rather, it sheds light upon the variation segregating in the metapopulation as a whole. We shall return to this question below while discussing the larval serum protein genes. Meanwhile, it is noteworthy that expression

variation among families was detected across the three temperature treatments, and in general, the family effects were as strong as the temperature effects on gene expression.

Of particular interest are genes that differ in expression among families in a dissimilar manner in different treatments (family-by-treatment interaction), as they provide variation upon which natural selection may operate. Although the number of genes with significant family-by-treatment interaction was smaller than the number of genes with significant family and treatment effects, the genes with the most robust expression difference (Table S6 candidate gene list) belonged to the same functional categories that were found in the enrichment analysis, including larval serum proteins, cuticle proteins, HSPs, mitochondrial genes and genes related to metabolism. Based on this observation, these genes are good candidates for future studies using a larger number of individuals.

Timing of development and gene expression

Most of the transcripts (67% of probes) that showed significant ($P < 0.05$) expression differences between the treatments responded linearly to temperature, either increasing (1397 probes) or decreasing (1375 probes) from cold to standard to hot. This would be expected if the treatments and sampling cause systematic differences in the developmental stage of the larvae at the time of sampling. Given that the majority of the enriched categories showed both family and treatment effects, and assuming a positive correlation between phenotypic and expression variation, we expected greatest variation in expression among families in the cold treatment, in which the rate of development varied most. However, in reality, expression varied most in the hot treatment, in which the families differed least in the rate of development. In addition, the families (N74 and O171) that differed most in the duration of larval development differed least in gene expression. These results argue that expression variation is not solely or even largely owing to the timing of development, and neither can variation in gene expression be explained by a single phenotypic variable (larval duration). Either most of the expression variation affects phenotypic traits that were not measured or larval development does not progress in a simple linear manner. In fact, only variation in the expression of *larval serum proteins* (LSPs) was directly related to variation in the rate of larval development. The expression of LSPs increased from cold to hot and was highest in family N74, which showed the fastest larval development, and lowest in O171, which showed the slowest larval development.

Temperature treatment and stress

Acclimation to a new environment can be achieved by altering gene expression. Generally, the more drastic the environmental change the more genes are likely to be involved. Exposure to dissimilar environments can also reveal more of the underlying genetic differences among the families. Thus, our temperature treatments were expected to reveal general thermal stress responses and family differences. Indeed, most of the expression variation was observed at the treatment extremes, with variation among the families seven times higher in hot compared to standard, with cold being intermediate. Thus, hot appears to have been the most stressful treatment. However, as all individuals survived to adults in hot, but not in cold, the benefits of elevated temperature on development apparently outweighed the presumed thermal stress.

The enriched categories related to thermal stress were split into treatment-enriched (*response to unfolded protein*) and family-enriched (*response to heat*) categories. These enrichments were mainly because of expression variation in heat shock protein (HSP) genes. The small HSPs (20–40 kDa) were responding to the temperature treatment (mostly up-regulated in hot), while the larger HSPs (70 and 90 kDa) were constitutively expressed at different levels among the families, consistent with high heritability observed in humans (Dixon *et al.* 2007) and *Drosophila* (Krebs & Feder 1997).

In contrast to typical heat stress experiments, which apply constant and extremely high temperatures resulting in the up-regulation of large HSPs (Lindquist 1986; Cvoro *et al.* 1998; Tachibana *et al.* 2005; Zhang & Denlinger 2010), our study was focused on ecologically relevant thermal conditions. We observed no treatment-induced changes in large HSP expression, and given the zero mortality in hot, the existing family-level variation may not have mattered under the experimental conditions. The immediate heat stress experienced in the hot treatment at mid-day only increased the expression of small HSPs, associated with brief temperature perturbations in *Drosophila* (Qin *et al.* 2005; Sørensen *et al.* 2005). As the night-time temperatures were low, the larvae could probably recuperate from the high-temperature stress without accumulating detrimental effects.

Larval mortality was zero in the hot treatment, in which the relative growth rate was highest and the pupal weight was greatest. These results are not consistent with the so-called temperature–size rule (TSR), according to which lower temperatures result in slower larval growth but ultimately lead to higher adult body mass (e.g. Allsopp 1981). However, exceptions to TSR are relatively common in Lepidoptera (e.g. Atkinson 1994). The high larval and pupal mortalities observed

in the cold treatment, along with the largest pupal weight in the hot treatment, suggest that thermal conditions may be a limiting factor in the study population of the Glanville fritillary, which occurs at the northern range limit of the species.

Larval serum proteins (LSPs)

Larval serum proteins are expressed in the fat body and secreted into the haemolymph during the final larval instar (Haunerland 1996). They are reabsorbed during pupation and used as amino acid reservoirs in morphogenesis, with additional ligand and transport functions (Telfer & Kunkel 1991). Their expression is regulated by 20-hydroxyecdysone (20E), one of the most important hormones regulating insect development, especially during the final larval instar of Lepidoptera (Burmester *et al.* 1999; Hiruma & Riddiford 2010).

LSP expression was affected by the temperature treatment, it differed among the families, and it showed a significant family-by-treatment interaction. The close correspondence between LSP expression and larval development time suggests that either these genes directly affect larval development or variation in their expression reflects upstream pathways controlling both LSP expression and development rate. In the former case, the expression of LSPs, which are used as energy reserves (Hiruma & Riddiford 2010), would reflect differences in energy acquisition and management leading to differences in growth rate; in the latter case, the hormonal pathways (especially 20E) regulating LSP expression would have differential onset or sensitivity among the larval families.

Our results are more consistent with the hormonal regulation of LSPs than with the energy storage hypothesis. First, LSP expression showed only a weak correlation with larval weight increase, and the family-by-treatment interactions observed in LSP expression were not observed in larval growth. Second, the expression of LSP genes was highly correlated and hence apparently co-regulated. On average, the expression of one LSP gene explained about 80% of the expression of another LSP gene, and in some cases, the correlation was nearly complete (Fig. S7). Third, we identified other genes showing expression patterns that were highly correlated with LSP expression. One interesting group of genes is the cuticle-binding proteins, involved in cuticle formation, which were negatively correlated with LSP expression (Fig. S7). Both groups of genes are known to be regulated by juvenile and ecdysteroid hormones during larval development.

In one of the best-studied Lepidopteran species, the tobacco hornworm (*Manduca sexta*), LSPs, along with other genes, undergo a characteristic expression shift driven by hormonal changes during mid-development of the final-instar larvae (Hiruma & Riddiford 2010). During this instar, the expression of cuticle-binding proteins decreases, while LSP expression increases at the pupal commitment/wandering stage (Hiruma & Riddiford 2010). Our study was focused on this same time period. In addition to the shift between the LSP and cuticle-binding protein expressions, we observed expression changes in other genes known to be regulated by or interacting with 20E, including ecdysteroid-modifying enzymes (3-dehydroecdysone reductase), ecdysteroid receptor and juvenile hormone-binding proteins.

In a previous study on adult Glanville fritillary females, physiological and microarray analyses revealed significant differences in LSP expression as well as in juvenile hormone (JH) titre levels between newly established and old local populations (Wheat *et al.* 2011). This study was based on second-generation common garden-reared butterflies, and hence, the results are likely to reflect heritable genetic variation affecting LSP expression. The present and the previous studies suggest that variation in LSP expression and its interaction with temperature reflect differences in developmental rate and temporal variation in hormonal signalling.

Technical performance of the microarrays

Although a number of studies have examined the factors affecting the correlation between microarray and qPCR results (Canales *et al.* 2006; Morey *et al.* 2006; Arikawa *et al.* 2008), they are typically focused on correlations and fold changes, and very few studies have conducted analyses in the mixed model statistical context. Here, we specifically validated our microarray results using 14 genes that were differentially expressed in the temperature treatments, spanning a wide range of expression levels. The correspondence between the qPCR and microarray results was good, with similar results obtained using independent biological samples, thereby confirming the biological generality of the microarray results (Fig. S6, Table S1). The higher variability in the qPCR results compared to the microarray measurements (Fig. S2) indicates that the qPCR assays required larger sample sizes than the microarray experiment to yield a similar level of statistical significance. Our validation indicates that oligonucleotide microarrays based on *de novo* assembled next-generation sequences provide a robust, accurate and rather inexpensive method to study gene expression variation.

Conclusions

We have shown that the rate of development and pupal mass, which have important fitness consequences in

butterflies, have high heritability. There is also substantial variation among larval families in the expression of genes apparently regulating larval duration and responding to related hormonal cascades. In particular, larval serum protein (LSP) genes exhibited significant variation in expression among larval families, in which variation was furthermore correlated with variation in larval duration. LSP genes are differentially expressed in female butterflies in newly established versus old local populations (Wheat *et al.* 2011) in the same metapopulation, suggesting an association between larval performance and the colonizing capacity of adult butterflies. This could occur either through an improved persistence of newly established populations, improved dispersal rate of adults or a combination of the two. Such an association has been detected at the organismal level in a study comparing four regional populations of the Glanville fritillary in northern Europe, in which two regional populations, including the one from the Åland Islands that was studied here, had significantly faster larval development than the others but also significantly higher flight metabolic rate (A. Duplouy and I. Hanski unpublished data), a proxy of high dispersal rate (Niitepõld *et al.* 2009). It is thus possible that the well-studied differences in the life histories of newly established versus old local populations in the Glanville fritillary (Hanski *et al.* 2006; Saastamoinen 2007; Ovaskainen *et al.* 2008; Klemme & Hanski 2009; Hanski 2011) are causally related to genes controlling larval development. These results suggest working hypotheses for further research and highlight the power of employing genomic tools in life history and ecological research on natural populations.

Acknowledgements

This research was supported by Academy of Finland (Finnish CoE Programmes 2000–2005, 2006-2011 and grant 131155), the European Research Council (232826 to IH), the LUOVA graduate school and the US National Science Foundation (grants EF-0412651 and IOS-0950416).

References

Allsopp PG (1981) Development, Longevity and Fecundity of the False Wireworms *Pterohelaeus darlingensis* and *P. alternatus* (*Coleoptera*: *Tenebrionidae*). I. Effect of Constant Temperature. *Australian Journal of Zoology*, **29**, 605–619.

Al-Shahrour F, Díaz-Uriarte R, Dopazo J (2005) Discovering molecular functions significantly related to phenotypes by combining gene expression data and biological information. *Bioinformatics*, **21**, 2988–2993.

Al-Shahrour F, Minguez P, Tarraga J *et al.* (2006) BABELOMICS: a systems biology perspective in the functional annotation of genome-scale experiments. *Nucleic Acids Research*, **34**, W472–476.

Arikawa E, Sun Y, Wang J *et al.* (2008) Cross-platform comparison of SYBR® Green real-time PCR with TaqMan PCR, microarrays and other gene expression measurement technologies evaluated in the MicroArray Quality Control (MAQC) study. *BMC Genomics*, **9**, 328.

Atkinson D (1994) Temperature and Organism Size–A Biological Law for Ectotherms? *Advances in Ecological Research* **25**, 1–58.

Ayroles JF, Carbone MA, Stone EA *et al.* (2009) Systems genetics of complex traits in *Drosophila melanogaster*. *Nature Genetics*, **41**, 299–307.

Barrett RDH, Rogers SM, Schluter D (2008) Natural Selection on a Major Armor Gene in Threespine Stickleback. *Science*, **322**, 255–257.

Burmester T, Antoniewski C, Lepesant J-A (1999) Ecdysone-regulation of synthesis and processing of Fat Body Protein 1, the larval serum protein receptor of *Drosophila melanogaster*. *European Journal of Biochemistry*, **262**, 49–55.

Canales RD, Luo Y, Willey JC *et al.* (2006) Evaluation of DNA microarray results with quantitative gene expression platforms. *Nature Biotechnology*, **24**, 1115–1122.

Cui X, Affourtit J, Shockley KR, Woo Y, Churchill GA (2006) Inheritance Patterns of Transcript Levels in F1 Hybrid Mice. *Genetics*, **174**, 627–637.

Cvoro A, Dundjerski J, Trajkovic D, Matic G (1998) Association of the rat liver glucocorticoid receptor with Hsp90 and Hsp70 upon whole body hyperthermic stress. *The Journal of Steroid Biochemistry and Molecular Biology*, **67**, 319–325.

Dixon AL, Liang L, Moffatt MF *et al.* (2007) A genome-wide association study of global gene expression. *Nature Genetics*, **39**, 1202–1207.

Ellegren H, Sheldon BC (2008) Genetic basis of fitness differences in natural populations. *Nature*, **452**, 169–175.

Falconer DS, Mackay TFC (1996) *Introduction to Quantitative Genetics*, 4th edn. Pearson Education Limited, Essex, UK.

Feder ME, Mitchell-Olds T (2003) Evolutionary and ecological functional genomics. *Nature Reviews Genetics*, **4**, 649–655.

Gibson G, Riley-Berger R, Harshman L *et al.* (2004) Extensive Sex-Specific Nonadditivity of Gene Expression in *Drosophila melanogaster*. *Genetics*, **167**, 1791–1799.

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

Hanski IA (2011) Eco-evolutionary spatial dynamics in the Glanville fritillary butterfly. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 14397–14404.

Hanski I, Saccheri I (2006) Molecular-Level Variation Affects Population Growth in a Butterfly Metapopulation. *PLoS Biology*, **4**, e129.

Hanski I, Kuussaari M, Nieminen M (1994) Metapopulation Structure and Migration in the Butterfly Melitaea Cinxia. *Ecology*, **75**, 747–762.

Hanski I, Saastamoinen M, Ovaskainen O (2006) Dispersal-related life-history trade-offs in a butterfly metapopulation. *Journal of Animal Ecology*, **75**, 91–100.

Haunerland NH (1996) Insect storage proteins: gene families and receptors. *Insect Biochemistry and Molecular Biology*, **26**, 755–765.

Hiruma K, Riddiford LM (2010) Developmental expression of mRNAs for epidermal and fat body proteins and hormonally

regulated transcription factors in the tobacco hornworm, *Manduca sexta*. *Journal of Insect Physiology*, **56**, 1390–1395.

Kallioniemi E, Hanski I (2011) Interactive effects of *Pgi* genotype and temperature on larval growth and survival in the Glanville fritillary butterfly. *Functional Ecology*, **25**, 1032–1039.

Kingsolver JG (2000) Feeding, Growth, and the Thermal Environment of Cabbage White Caterpillars, *Pieris rapae L*. *Physiological and Biochemical Zoology*, **73**, 621–628.

Kingsolver JG, Gomulkiewicz R (2003) Environmental Variation and Selection on Performance Curves. *Integrative and Comparative Biology*, **43**, 470–477.

Kingsolver JG, Huey RB (2008) Size, temperature, and fitness: three rules. *Evolutionary Ecology Research*, **10**, 251–268.

Kingsolver JG, Ragland GJ, Shlichta JG (2004) Quantitative genetics of continuous reaction norms: thermal sensitivity of caterpillar growth rates. *Evolution*, **58**, 1521–1529.

Klemme I, Hanski I (2009) Heritability of and strong single gene (*Pgi*) effects on life-history traits in the Glanville fritillary butterfly. *Journal of Evolutionary Biology*, **22**, 1944–1953.

Krebs RA, Feder ME (1997) Natural Variation in the Expression of the Heat-Shock Protein HSP70 in a Population of *Drosophila melanogaster* and Its Correlation with Tolerance of Ecologically Relevant Thermal Stress. *Evolution*, **51**, 173–179.

Kuussaari M (1998) *Biology of the Glanville Fritillary Butterfly (Melitaea cinxia)*. Ph.D., Helsingin Yliopisto (Finland).

Lindquist S (1986) The Heat-Shock Response. *Annual Review of Biochemistry*, **55**, 1151–1191.

Monks SA, Leonardson A, Zhu H *et al.* (2004) Genetic Inheritance of Gene Expression in Human Cell Lines. *The American Journal of Human Genetics*, **75**, 1094–1105.

Morey J, Ryan J, Van Dolah F (2006) Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biological Procedures Online*, **8**, 175–193.

Nieminen M, Siljander M, Hanski I (2004) *Structure and Dynamics of Melitaea Cinxia Metapopulations*. Oxford Univ Press, New York.

Niitepõld K, Smith AD, Osborne JL *et al.* (2009) Flight metabolic rate and *Pgi* genotype influence butterfly dispersal rate in the field. *Ecology*, **90**, 2223–2232.

Oleksiak MF, Churchill GA, Crawford DL (2002) Variation in gene expression within and among natural populations. *Nature Genetics*, **32**, 261–266.

Oleksiak MF, Roach JL, Crawford DL (2005) Natural variation in cardiac metabolism and gene expression in *Fundulus heteroclitus*. *Nature Genetics*, **37**, 67–72.

Orsini L, Corander J, Alasentie A, Hanski I (2008) Genetic spatial structure in a butterfly metapopulation correlates better with past than present demographic structure. *Molecular Ecology*, **17**, 2629–2642.

Orsini L, Wheat CW, Haag CR *et al.* (2009) Fitness differences associated with *Pgi* SNP genotypes in the Glanville fritillary butterfly (*Melitaea cinxia*). *Journal of Evolutionary Biology*, **22**, 367–375.

Ovaskainen O, Smith AD, Osborne JL *et al.* (2008) Tracking butterfly movements with harmonic radar reveals an effect of population age on movement distance. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 19090–19095.

Qin W, Neal SJ, Robertson RM, Westwood JT, Walker VK (2005) Cold hardening and transcriptional change in *Drosophila melanogaster*. *Insect Molecular Biology*, **14**, 607–613.

Ritz C, Spiess A-N (2008) qpcR: an R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis. *Bioinformatics*, **24**, 1549–1551.

Roff D (1997) Evolutionary Quantitative Genetics. Chapman and Hall, New York.

Rozen S, Skaletsky H (1999) Primer3 on the WWW for General Users and for Biologist Programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology*, pp. 365–386. Humana Press, Totowa, New Jersey.

Saastamoinen M (2007) 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, Hanski I (2008) Genotypic and Environmental Effects on Flight Activity and Oviposition in the Glanville Fritillary Butterfly. *The American Naturalist*, **171**, 701–712.

Saastamoinen M, Ikonen S, Hanski I (2009) Significant effects of *Pgi* genotype and body reserves on lifespan in the Glanville fritillary butterfly. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 1313–1322.

Saccheri IJ, Boggs CL, Hanski I, Ehrlich PR (2004) Genetics of checkerspot populations. *On the Wings of Checkerspots: A Model System for Population Biology*, 199–218.

Schadt EE, Monks SA, Drake TA *et al.* (2003) Genetics of gene expression surveyed in maize, mouse and man. *Nature*, **422**, 297–302.

Scott CP, Williams DA, Crawford DL (2009) The effect of genetic and environmental variation on metabolic gene expression. *Molecular Ecology*, **18**, 2832–2843.

Slate J, Santure AW, Feulner PGD *et al.* (2010) Genome mapping in intensively studied wild vertebrate populations. *Trends in Genetics*, **26**, 275–284.

Smyth GK (2004) Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. *Statistical Applications in Genetics and Molecular Biology*, **3**, 1–25.

Sørensen JG, Nielsen MM, Kruhøffer M, Justesen J, Loeschcke V (2005) Full genome gene expression analysis of the heat stress response in *Drosophila melanogaster*. *Cell Stress & Chaperones*, **10**, 312–328.

Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 9440–9445.

Tachibana S-I, Numata H, Goto SG (2005) Gene expression of heat-shock proteins (*Hsp23*, *Hsp70* and *Hsp90*) during and after larval diapause in the blow fly *Lucilia sericata*. *Journal of Insect Physiology*, **51**, 641–647.

Telfer WH, Kunkel JG (1991) The function and evolution of insect storage hexamers. *Annual Review of Entomology*, **36**, 205–228.

Turner TL, Bourne EC, Von Wettberg EJ, Hu TT, Nuzhdin SV (2010) Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. *Nature Genetics*, **42**, 260–263.

Vera JC, Wheat CW, Fescemyer HW *et al.* (2008) Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. *Molecular Ecology*, **17**, 1636–1647.

Vuylsteke M, van Eeuwijk F, Van Hummelen P, Kuiper M, Zabeau M (2005) Genetic Analysis of Variation in Gene Expression in *Arabidopsis thaliana*. *Genetics*, **171**, 1267–1275.

Wade MJ, Goodnight CJ (1998) Perspective: the Theories of Fisher and Wright in the Context of Metapopulations: when Nature Does Many Small Experiments. *Evolution*, **52**, 1537–1553.

Walsh B, Blows MW (2009) Abundant Genetic Variation + Strong Selection = Multivariate Genetic Constraints: a Geometric View of Adaptation. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 41–59.

Watt WB, Dean AM (2000) Molecular-functional studies of adaptive genetic variation in prokaryotes and eukaryotes. *Annual Review of Genetics*, **34**, 593–622.

Wheat CW, Haag CR, Marden JH, Hanski I, Frilander MJ (2010) Nucleotide Polymorphism at a Gene (*Pgi*) under Balancing Selection in a Butterfly Metapopulation. *Molecular Biology and Evolution*, **27**, 267–281.

Wheat CW, Fescemyer HW, Kvist J *et al.* (2011) Functional genomics of life history variation in a butterfly metapopulation. *Molecular Ecology*, **20**, 1813–1828.

Whitehead A, Crawford DL (2006) Variation within and among species in gene expression: raw material for evolution. *Molecular Ecology*, **15**, 1197–1211.

Wolfinger RD, Gibson G, Wolfinger ED *et al.* (2001) Assessing Gene Significance from cDNA Microarray Expression Data via Mixed Models. *Journal of Computational Biology*, **8**, 625–637.

Zahurak M, Parmigiani G, Yu W *et al.* (2007) Pre-processing Agilent microarray data. *BMC Bioinformatics*, **8**, 142.

Zhang Q, Denlinger DL (2010) Molecular characterization of heat shock protein 90, 70 and 70 cognate cDNAs and their expression patterns during thermal stress and pupal diapause in the corn earworm. *Journal of Insect Physiology*, **56**, 138–150.

Zheng C, Ovaskainen O, Hanski I (2009) Modelling single nucleotide effects in *phosphoglucose isomerase* on dispersal in the Glanville fritillary butterfly: coupling of ecological and evolutionary dynamics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **364**, 1519–1532.

J.K. is a geneticist, with special interest in applying molecular tools to ecological model species. C.W.W is a biologist, focused on finding and studying the genetic variation causing performance and fitness variation in the wild. E.K. is an ecologist whose current research focuses on species ability to track climate change. M.S. is an ecologist with special interest on dispersal and life-history evolution in the Glanville fritillary butterfly. I.H. is an ecologist with broad interests in eco-evolutionary spatial dynamics. M.J.F. is biologist with special interests on gene expression regulation and alterative splicing.

Data accessibility

Microarray data and detailed methods used in this work are available for download from Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo), accession number GSE32203.

Supporting information

Table S1 Comparison of microarray and qPCR results. (A) Significance values are shown for the fixed factors (treatment, family and family-by-treatment interaction) in the analysis for microarray and qPCR for the 14 genes (T1-T14) that were validated. The qPCR was analyzed with different sample sets of varying size and relatedness to the microarray samples; Microarray samples: the same samples as in the microarray experiment (n = 35); Microarray families: additional samples from the same three families that were used in the microarray analysis (n = 57); Non-microarray samples: samples not included in the microarray analysis from all six families (n = 52); Non-microarray families: samples from the three other families that were not included in the microarray analysis (n = 30); All samples: all the samples in the thermal experiment (including the microarray samples, n = 87). (B) The number of individuals making up these samples sets by family and treatment are shown below the results.

Table S2 The primers used in the qPCR analysis. For each primer pair the corresponding names of the microarray probes, genes (unique 454 contigs), the best matching *Bombyx mori* and *Drosophila melanogaster* homologs are indicated. The primer ID describes whether the gene was used as an endogenous control (C) or a target (T) in the analysis.

Table S3 Principal component analysis of phenotypic measurements of field collected larvae reared in two temperatures (standard and warm), nested ANOVA analysis of these components by individual within family, family within population, population and residual, and the estimated broad-sense heritabilities.

Table S4 Pearson's correlation estimates for consecutive measures of development time and body weight across pre-adult developmental stages and correlations of body weight and development time at 7th instar and pupae in the field collected larvae in two rearing conditions (standard and warm).

Table S5 Mixed model analysis of body weight and development time in the second-generation lab-reared larvae (6 families with 10 larvae in each treatment per family) before and after the 7th instar larvae were exposed to the temperature treatments (cold, standard and hot).

Table S6 Lists of probes significant for the factors: family, treatment or family by treatment interaction with two statistical cut-off levels; Subset 1: probes with P value < 0.05, Subset2: probes containing less than one expected false positive (based on Q value estimation), and a list of candidate genes for family by treatment interactions (candidate gene list).

Table S7 Results from the gene set enrichment analysis (Babelomics—FatiScan; Al-Shahrour *et al.* 2006) for pair-wise comparisons of treatments (cold, standard and hot), families (N170, N74 and O171), families within the same treatment and treatments with the same family.

Fig. S1 Bi-plots of pre-treatment differences in larval weight and last molt date among families and sexes in the second-generation lab-reader larvae. (A) Using samples with known sex. (B) Centering these values by family to highlight the dif-

ferences between the sexes. (C) Combining the samples with known and inferred sex. Error bars in (A) and (C) indicate standard error of the means. Female larvae were on average 16.5% heavier than males

(R^2 = 0.321; F_1 = 18.84; P < 0.0001) and reached the last instar stage 1.7 days later than the males (R^2 = 0.514; F_1 = 64.36; P < 0.0001). Assuming these same relationships for our entire dataset, we were able to assign individuals of unknown sex (34% of entire dataset, n = 61/179) using the following criteria: males were required to have pre-treatment phenotype values for larval weight and last molt date less than the corresponding minimum observed among the females within each given family.

Fig. S2 The distributions of the coefficient of variation (CV) for (A) the microarray probes after pre-processing and (B) qPCR crossing point values (DCt) after filtering.

Fig. S3 P value distributions of the microarray probes for the fixed factors (family, treatment and family by treatment interaction) in the mixed model analysis.

Fig. S4 Direction and magnitude of expression change for the probes belonging to the GO and KEGG categories (n = 57) identified in the gene set enrichment analysis. Standardized least squared mean expression in each family plotted in the three temperature treatments. See Table 3 for the descriptions of the categories.

Fig. S5 K-means clusters of the probes belonging to the enriched category *response to unfolded protein (GO:0006986).* Clusters 1 and 3 represent the majority of probes (n = 15 in both) and have a near linear trend (either decreasing or increasing) in respect to elevation of temperature. Clusters 2 and 4 contain the remaining probes (2: n = 5; 4: n = 4) and have nearly identical expression at the temperature extremes (hot and cold), but differ in respect to the standard condition.

Fig. S6 qPCR validation showing the expression difference between hot and cold treatment across 14 genes (T1-T14). Results from the microarray experiment (n = 35) and qPCR with multiple sample sets; Microarray samples: the same samples as in the microarray experiment (n = 35); Microarray families: additional samples from the same three families that were used in the microarray analysis (n = 57); Non-microarray samples: samples not included in the microarray analysis from all six families (n = 52); Non-microarray families: samples from the three other families that were not included in the microarray analysis (n = 30); All samples: all the samples in the thermal experiment (including the microarray samples, n = 87). The different sample sets used in the qPCR analysis are indicated in the inset.

Fig. S7 Correlations of expression between (A) all probes belonging to the category *larval serum protein* (LSP), (B) two probes belonging to unique LSP genes and (C) the same LSP probe and one probe belonging to the category: *cuticle-binding protein*. Each dot represents a unique biological sample.