Gene expression under thermal stress varies across a geographic range expansion front

Lesley T. Lancaster[1]*, Rachael Y. Dudaniec[2], Pallavi Chauhan[3], Maren Wellenreuther[3, 4],

Erik I. Svensson[3], Bengt Hansson[3]

[1] Institute of Biological and Environmental Sciences, University of Aberdeen, United

Kingdom

[2] Macquarie University, Department of Biological Sciences, Sydney, NSW, Australia

[3] Department of Biology, Lund University, Sweden

[4] Institute of Plant and Food Research, New Zealand

* Corresponding author: Lesley T. Lancaster:

University of Aberdeen

School of Biological Sciences, Tillydrone Ave

Aberdeen AB24 2TZ

United Kingdom

 lesleylancaster@abdn.ac.uk.

Running title: Thermal gene expression and range shifts

Abstract

Many ectothermic species are currently expanding their distributions polewards due to anthropogenic global warming. Molecular genetic mechanisms facilitating range expansion under these conditions are largely unknown, but understanding these could help mitigate expanding pests and disease vectors, or help explain why some species fail to track changing climates. Here, using RNA-seq data, we examine genome-wide changes in gene expression under heat and cold stress in the range-expanding damselfly *Ischnura elegans* in northern Europe. We find that both the number of genes involved and levels of gene expression under heat stress have become attenuated during the expansion, consistent with a previously-reported release from selection on heat tolerances as species move polewards. Genes upregulated under cold stress differed between core and edge populations, corroborating previously-reported rapid adaptation to cooler climates at the expansion front. Expression of sixty-nine genes exhibited a region x treatment effect; these were primarily upregulated in response to heat stress in core populations but in response to cold stress at the range edge, suggesting that some cellular responses originally adapted to heat stress may switch to cold stress functionality upon encountering novel thermal selection regimes during range expansion. Transcriptional responses to thermal stress involving heat shock and neural function genes were largely geographically conserved, while retrotransposon, regulatory, muscle function and defence gene expression patterns were more variable. Flexible mechanisms of cold stress response and the ability of some genes to shift their function between heat and cold stress might be key mechanisms facilitating rapid poleward expansion in insects.

Introduction

As global climates warm under anthropogenic climate change, many species are expanding their geographic ranges to higher latitudes and elevations (Thomas *et al.* 2001; Parmesan & Yohe 2003; Hickling *et al.* 2005; Chen *et al.* 2011), resulting in the displacement of native species and increased outbreaks of disease and pests at higher latitudes (Bebber *et al.* 2013; Sánchez-Guillén *et al.* 2015), but also potentially providing net increases in biodiversity at high latitudes, compared to the pre-anthropocene (Thomas 2013). While some species are successfully and rapidly expanding under climate change, others fail to adapt or move and therefore suffer from population declines in response to increased temperatures (Bellard *et al.* 2012; Buckley & Kingsolver 2012). Understanding the genetic factors that predispose some species to undergo range expansions in response to climate change are therefore needed to predict and understand the spread of pests and invaders and to identify sources of genetic variation required to prevent warming-induced population declines.

Currently, the factors promoting climate warming-induced expansions are poorly understood. While extirpation events and range contractions under anthropogenic climate change have been documented in species with relatively low or constrained upper thermal tolerances (Sinervo *et al.* 2010; Comte *et al.* 2014), range expansions are weakly and inconsistently associated with upper thermal tolerances (Comte *et al.* 2014) or with life history, dispersal, or ecological niche traits (Angert *et al.* 2011). Thus while warming-induced range contractions may be immediately predictable based on habitat data and known thermal characteristics of at-risk species, more research is needed to understand how thermal tolerances and their underlying mechanisms facilitate successful responses to rapid climate change, including rapid range expansions or shifts to climatically favourable regions (e.g., Krehenwinkel *et al.* 2015).

In a recent study (Lancaster *et al.* 2015), we examined changes in thermal tolerance associated with a recent, climate-mediated northward range expansion in a widespread damselfly species (*Ischnura elegans* Vander Linden 1820, Odonata: Coenagrionidae). *Ischnura elegans* is broadly distributed across Eurasia and is currently expanding its range to higher latitudes in response to warming climates (Hickling *et al.* 2005; Cham *et al.* 2014). Using physiological thermal tolerance experiments on wild-caught individuals from different latitudinal locations, we found that mean heat tolerances were preserved across a ~500km latitudinal extent that spans the recent, poleward range expansion. Geographic conservation of heat tolerance in this species did not reflect evolutionary constraint preventing local adaptation within the colonized region. Instead, heat tolerance exhibited geographic variation in phenotypic values consistent with this trait being released from selection at the poleward range front (Lancaster *et al.* 2015). This suggests that declining environmental temperatures towards the expansion front result in neither any benefit nor any substantial fitness cost of maintaining ancestrally high heat tolerances in the new part of the range (Lancaster *et al.* 2015). Because range expansions are widespread in insects, latitudinal invariance in upper thermal tolerance reported across ectothermic species might generally result from a process of thermal release during poleward expansion (Lancaster in review).

In contrast to latitudinally-invariant heat tolerances, *I. elegans* exhibited dramatic improvement in cold tolerance at the approach of its poleward range front, associated with ongoing adaptation or acclimation to increasingly variable climates at high latitudes (Lancaster *et al.* 2015). Our prior results suggest that selection on cold tolerance remains strong during poleward range expansions (see also Preisser et al. 2008; Hidalgo-Galiana et al. 2014), despite the fact that expansions are facilitated by warming climates.

In the current study, we investigate the specific gene expression changes behind geographical variation in thermal tolerance during the recent range expansion of *I. elegans*. A major goal of this study was to identify gene expression patterns and physiological pathways that may predispose this and other species to range expansion under changing climates. We further aimed to identify geographic divergence in gene expression patterns arising during the course of the range expansion, under either intensified or relaxed thermal selection in the new part of the range; such changes may result from plasticity, microevolutionary change, or both (Lande 2009). We present RNA-seq data for whole individuals experimentally exposed to heat or cold stress (i.e., as reported in Lancaster *et al.* 2015) from two populations, one within the historic range, and another population at the expanding northern range limit (Fig. 1). Within each population, individuals from experimental heat and cold exposure treatments were sampled for expression data in order to identify genetic responses that differed between these two stressors within and across regions.

We ask the following questions: 1) Which genes or gene isoforms exhibit conserved levels of differential expression in response to heat or cold stress, across both core and range-limit populations? Genes whose function is conserved across multiple populations within a species and during range shifts will likely be most useful for characterising the species' general or pre-adapted molecular thermal response. 2) Which genes or isoforms were differentially expressed between thermal treatments in either the core or the edge region, but not across both? These genes could indicate novel adaptations, flexible strategies, or loss of function under release from thermal selection. 3) Which genes or isoforms have functionally switched from association with heat stress response to cold stress response, or vice versa, across the range expansion? We predict that genes ancestrally involved in upper thermal tolerances (in the ancestral, core region) may be co-opted for cold tolerance function as the

range expansion progresses, because selection on heat tolerance is relaxed and selection on cold tolerance is intensified during the range expansion process (Lancaster *et al.* 2015). The ability of several classes of stress-response proteins to exhibit evolutionary lability in whether they respond to heat vs. cold (Colinet & Hoffmann 2010; Telonis-Scott *et al.* 2014) suggests that some genes could switch between alternative stress functions within a species, which might be an important genetic mechanism underpinning rapid thermal adaptation during range shifts.

Materials and Methods

Field collections, thermal stress experiments, and sample selection

In 2012 and 2013, we collected individual, adult damselflies from a set of 15 populations in the ancestral, core portion of the species geographic range for *I. elegans* in southern Sweden, ca. 500 km from the published species range limit (Fig. 1; hereafter the "core" region; see Svensson *et al.* 2005; Le Rouzic *et al.* 2015 for detailed data on long-term dynamics of these populations), and from 11 populations at the leading range edge in central Sweden (Fig. 1; hereafter the "edge" region). Using individuals from these populations, we determined regional variation in thermal tolerances (Lancaster *et al.* 2015, data summarised above). The edge region is characterised by significantly lower mean annual temperatures and significantly higher levels of climatic variability than the core region (Lancaster *et al.* 2015). Because *I. elegans* has previously been estimated to be moving northward at ca. 4.5 km • year^{-1} under anthropogenic climate change (Hickling *et al.* 2005), our core populations in southern Sweden are conservatively located within the northern part of the species' ancestral, pre-anthropocene geographic range. Consistent with *I. elegans'* rapid expansion status, we found healthy populations at significantly higher latitudes compared to all previously reported sightings for the species (http://www.gbif.org, http://www.artportalen.se)

and north of the previously published range boundary (Fig. 1, Dijkstra & Lewington 2006).

For each population, we captured individuals using butterfly nets along the edges of ponds and lakes where damselflies emerge and congregate to breed during the few weeks they spend as adults (methods following Svensson & Abbott 2005). Captures took place in the mornings, and damselflies were subsequently transported back to local field stations (edge region: Grimsö Wildlife Research Station, Swedish University of Agricultural Sciences; core region: Ecology Building, Lund University) where they were equilibrated for 30 minutes at 25 °C and then slowly ramped at 0.06 °C • min^{-1} to either 43 °C (heat stress experiment) or 2 °C (cold stress experiment). Pilot studies indicated that at those temperatures half of all individuals were knocked down (i.e., no longer standing) by the end of the trial. Following a 15 minute assessment for individual recovery rates, a random subset of subjects were euthanized and preserved at −80°C in RNA-later (Qiagen) for further analysis. Additional details on damselfly collection and thermal ramping experiments can be found in Lancaster *et al.* (2015).

For the current gene expression study, we chose one sampled population in each region for comparison. In the core region, we chose the long-term study population Höje Å 14 (55.649 °N, 13.319 °E; Fig 1) due to its representative population dynamics (Le Rouzic *et al.* 2015). In the edge region, we selected a population near the town of Nora (59.530 °N, 14.999 °E; hereafter, Nora; Fig 1) due to its position along the range edge (i.e., well beyond the published range limit for the species, see above) and its highly divergent climate and habitat suitability from core sites (Lancaster *et al.* 2015). In both of these populations, we randomly selected *n*=8 adult female individuals for comparison, 4 from the heat stress experiment and 4 from the cold stress experiment. Individuals were age and gender matched to avoid confounding influences of these traits on thermal stress response (King & MacRae 2015). These individuals were collected from Höje Å 14 on four separate days between June

12 and July 9, 2012, and from Nora on five separate days between June 23 and July 26, 2013.

The mean local temperature on capture dates for the selected individuals from Höje Å 14 and

Nora were 16.5±0.4 °C and 17.5±1.1 °C, respectively (http://www.smhi.se).

Sequencing, quality control, and de novo *assembly*

Total RNA was extracted using the Trizol LS Reagent (Life Technologies) according

to manufacturer's instructions. RNA samples were individually prepared (including mRNA

enrichment with a ploy(A) method) for sequencing using the Illumina Tru-Seq kit on an

Illumina HiSeq 2000 sequencer (Paired-end 100bp sequencing, 160bp insert length,

>2Gb/sample) at the Beijing Genomics Institute Shenzhen, China. The raw RNA sequence

reads have been deposited at the National Center for Biotechnology Information (NCBI)

under the Bioproject number PRJNA304220, providing access to the *I. elegans* sampling data

through the Biosample links SAMN04298996-SAMN04298999 and sequence read archives

SRR2962606-SRR2962609. Adapter and low quality bases below a quality score of 20 were

trimmed from raw sequences using the Nesoni clip version 0.109

(https://github.com/Victorian-Bioinformatics-Consortium/nesoni). Trimming was applied

separately on each sample. After trimming, reads with less than 25bp were further discarded

from the dataset, and all retained reads were used for the *de novo* assembly in Trinity, using

default parameters (Grabherr *et al.* 2011; see Table S1 for a detailed summary of the

trimming statistics). The transcriptome assembly followed the general workflow as outlined

in (Chauhan *et al.* 2014) using Trinity default settings followed by three quality and filtering

checks to retain only high quality transcripts for analysis. In the first step, duplicates were

removed by clustering the assembly at 95% sequence similarity using CD-HIT-EST version

2.17.0 (Li & Godzik 2006). In the second step, poor quality and potentially misassembled

transcripts were discarded by removing transcripts with a mean coverage per base of less than

five using the BED tools genome coverage application (Quinlan & Hall 2010). In the third

and final step, transcripts containing ribosomal RNA were removed from the assembly using

RepeatMasker version 4.0.1 (Smit et al. 1996-2014) using default settings. The high quality

assembly was obtained after the three filtering steps, and comprised 71,904 unique

transcripts, and was used for all subsequent analyses. Summary of initial and final assembly

statistics can be found in Table S2. An assembly completeness estimation with CEGMA

version 2.4.010312 (Parra *et al.* 2007) reported 245 out of 248 ultra-conserved genes as

'complete' and yielded a completeness score of 98.79%. The remaining three genes were

identified as 'partial' in the CEGMA analysis.

Assembly annotation

To functionally annotate the assembly, a BLASTX sequence homology search was

conducted against the NCBI non-redundant (nr) protein database with an e-value cut-off of

1×10^{-5}. All BLASTX results were imported into the BLAST2GO web version for further

annotation (Conesa *et al.* 2005). InterProScan was used for identifying conserved protein

domains in the assembly (Jones *et al.* 2014) and GO annotations were performed on the

BLASTX and InterProScan annotated transcripts (Ashburner *et al.* 2000). GO annotations

and GO Slim reductions were further applied to categorise transcripts into biological

processes, cellular components and molecular functional annotations using second level

database functions (Ashburner *et al.* 2000). Finally, enzymes and their corresponding

biological pathways were identified using the BLAST2GO integrated KEGG database

(Conesa *et al.* 2005). All analyses were performed using default settings.

Differential expression analysis

Differential expression (DE) between heat and cold stress treatments was analysed separately for Nora and Höje Å 14 to assess regional similarities and differences. In each of our comparisons, we contrasted gene expression levels of individuals under heat stress to those individuals subjected to cold stress, rather than to unstressed controls. Previous results suggest an evolutionary and physiological decoupling of upper and lower thermal limits (Klok *et al.* 2004; Lancaster *et al.* 2015). Thus, we focussed on genes that are likely to underpin the decoupled patterns of change in heat and cold tolerance between the ancestral and expanding part of the range.

First, filtered reads from each individual were analysed separately for abundance/raw counts in Trinity (Grabherr *et al.* 2011). In this pipeline, transcript abundance was estimated in RSEM (Li & Dewey 2011), with visualization in the IGV Module contained within RSEM using default settings (Table S3). Next, filtered read counts at the individual level were analysed to identify differentially expressed transcripts using the Bioconductor tool DESeq2 (Love *et al.* 2014) in the Trinity differential expression module using the default settings, where differential expression levels were sought between thermal treatments for each transcript, and comparisons were run separately within each region. Next, we extracted the most differentially expressed transcripts using two-fold filtering ($\log_2$ fold change ≥ 1 and an false discovery rate (FDR)-adjusted *P*-value of ≤ 0.001, i.e. using default settings; Haas 2015). A complete list of annotated and differentially expressed transcripts (for each set) are available through Dryad DOI: doi:10.5061/dryad.4qf2h. With this approach, a total of 780 significantly differentially expressed transcripts were identified across treatments in one or both populations.

For each set of differentially expressed transcripts a GO enrichment analysis was performed (Table S4). Further, to reduce the dimensionality of the data and focus on genes with major differences in their response to heat vs. cold, we also set a 10-fold (i.e. 3.32 $\log_2$ fold) cut-off in gene expression levels between the heat vs. cold treated individuals in each region (resulting in n=143 differentially expressed transcripts for analysis and discussion within the text; complementary analyses using the 1 $\log_2$ fold change cut-off are provided in supplementary materials). Trinity and DESeq2 analyses were performed using the resources provided by the Swedish National Infrastructure for Computing (SNIC) through the Uppsala Multidisciplinary Center for Advanced Computational Science (UPPNEX) under Project number b2013227. Differential expression patterns were visualised using a heat map constructed using the gplots package for R (Warnes *et al.* 2015), with averaging agglomeration method and Pearson dissimilarity assignment. To transcend the arbitrary nature of cut-offs for statistical significance and for magnitudes of effect for our differential analysis of individual loci, we also conducted a between-group analysis (BGA; Culhane *et al.* 2002) to visualise divergence of each thermal treatment group within and among populations, for all 71,904 transcripts. In our BGA analysis, we ordinated groups of samples according to principal component analysis (PCA), and then projected individual samples back onto the first two PCA axes. BGA analysis was carried out in the Bioconductor for R package made4 v1.36.1 (Culhane *et al.* 2005).

To test for region × treatment effects on the expression of genes and isoforms, raw read counts from the Trinity assembly were imported into edgeR Bioconductor package for R (Robinson *et al.* 2010) and normalised using the trimmed mean of M-values normalization method (TMM; Robinson and Oshlack 2010). We used multidimensional scaling plots to check for consistency of samples within each category, and then conducted gene-wise general linear models (GLMs) in edgeR to identify transcripts with expression patterns that were

better explained across all samples by region + treatment + region × treatment than by additive effects of region + treatment alone, where significance of the region × treatment effect was established using likelihood ratio tests for each transcript. In this analysis, *P*-values were adjusted for multiple comparisons using default settings, i.e. a FDR cut-off of 0.05 (Robinson et al. 2010), but we did not specify any particular minimum threshold for magnitude of effect; this resulted in 69 significant transcripts.

Results

Out of 71,904 transcripts analysed in Deseq2 for differential expression across thermal treatments within either core or edge populations (Fig 1), 97 transcripts were significantly differentially expressed in the core population, and at a greater than 10-fold difference between treatments. Of these, 68 (70%) were overexpressed following the heat treatment and 29 following the cold treatment (Fig. 2a; compare to Fig. S1a for results using the 1 $\log_2$ fold change cur-off). In the edge population, 68 transcripts were differentially expressed at >10-fold change, and of these 49 were overexpressed following heat treatment (72%) and 19 following cold treatment (Fig. 2a).

Between-group analysis of all 71,904 transcripts according to PCA revealed that most of the variation among groups (PC1; 53% of variation) was explained by thermal treatment, and region explained much of the remaining variation (PC2; 27% of variation) (Fig 2b). Among-population variation in gene expression was lower for cold-treated individuals (Fig. 2b, right side) than for heat-treated individuals (Fig. 2b, left side). Differences in gene expression between heat vs. cold treated individuals were slightly larger in the core (Fig. 2b, top) than at the range edge (Fig. 2b, bottom).

In contrast to the whole-transcriptome results, regional differences in expression pattern for the differentially expressed genes were greater for cold-treated individuals (Figs. 2c, S1b, left columns) than for heat-treated individuals (Figs. 2c, S1b, right columns). As in the BGA analysis of all transcripts (Fig. 2b), temperature was more important than region in explaining expression differences among the transcripts significantly differentially expressed between thermal trials within each region (Figs. 2c, S1b, column clustering).

Differentially expressed genes and isoforms shared across regions

Of the 143 total differentially expressed transcripts across the two populations at greater than 10-fold change, 22 isoforms representing 16 genes were significantly differentially expressed in *both* core and edge populations (Fig 2a, Table 1). Independent identification of differential expression for these isoforms in each of two, geographically distinct populations suggests that these represent genes that are among the more conserved elements of the thermal stress response in *I. elegans*.

Of these 22 regionally conserved genes and isoforms, 21 (96%) were overexpressed following the heat ramping treatment in both populations, in comparison to the one remaining gene that was overexpressed following the cold ramping treatment (4%). The majority of these genes (14 out of 19; see top 14 rows of Table 1) mapped to genes with previously described heat stress response functionality, including heat shock proteins and co-chaperones (see Discussion for details of specific genes). Five remaining gene isoforms (rows 15-19 in Table 1) that were upregulated in response to heat in both populations did not map to genes with a previously described association to heat stress (but see Discussion). Of all 22 genes and isoforms with shared function across core and range edge regions, most were upregulated more strongly in response to thermal stress within the core than at the range edge (see 'Difference' column in Table 1), with the exception of genes related to transposon activity

and epigenetic regulation, which gained greater activity in response to heat at the range edge. The single transcript overexpressed at > 10-fold change after cold exposure in both populations aligned to a hypothetical protein of unknown function. Two transcripts overexpressed under heat ramping trials in both populations also mapped to proteins of unknown function.

Genes and isoforms with unique expression patterns in each region

In contrast to expression patterns that were shared between populations, genes and isoforms that were uniquely differentially expressed in *either* the core or the edge population (i.e., were differentially expressed between thermal treatments in one population but not both) showed a relatively balanced representation of responses to heat and cold stress. Of these, 63% or 61% were upregulated in response to heat (vs. 37% or 39% that were upregulated in response to cold) in the core or edge population, respectively (Table 2, Fig. 2a, Table S5). Thermal stress responses generally involved a greater number of significantly differentially expressed genes and isoforms in the core than at the range edge (98 [core] vs. 59 [edge] unique transcripts differentially expressed at >10-fold change; Fig. 2a, Table S5). Most of the differentially expressed genes and isoforms that were only identified in a single population mapped to genes relating to transposon activity, regulation of transcription, and general cellular functioning, including growth, replication, and homeostasis (Table 2, Table S5).

Genes with reversed gene functionality following expansion

Sixty-nine unique transcripts showed significant evidence of expression activity reversal in response to heat vs. cold between the core population and the expanding range edge population (Table S6). Sixty of these 69 genes (87%) showed a similar pattern of

upregulation in response to heat rather than to cold in the core, but at the range edge this was reversed, with upregulation in response to cold rather than heat (Fig. 3a; see Table S6 for gene-specific details). There was very low or no expression of these genes detected under cold stress in the core, and under heat stress at the range edge (Fig. 3a; Table S6). All nine transcripts that significantly exhibited the reverse region x treatment expression pattern (i.e. elevated expression in response to cold in the core, but in response to heat at the edge) had previously unknown functionality (Fig. 3b, Table S6).

Discussion

Changes in gene expression under altered thermal selection

Here, we identify putative mechanisms underpinning changes in thermal tolerance during a climate-mediated range expansion by investigating gene expression differences between *I. elegans* exposed to heat vs. cold stress, in core and range-expanding edge populations. Our results support an attenuation of the molecular basis of heat response under thermal release, as populations migrate to cooler regions at higher latitudes. We find that the number of genes significantly overexpressed under heat trials was greater in the warmer core region than at the cooler range edge (Fig. 2, Table 2), indicating a reduction in the variety of mechanisms involved in the heat response as the poleward expansion progresses. In further support of this mechanism, we also find that 16 out of 21 (76%) conserved genes that were overexpressed under heat stress in both regions exhibited a greater magnitude of differences in expression levels between treatments in the core than at the range edge (Table 1). These combined results are consistent with our hypothesis that release from thermal selection on heat tolerances and genetic drift may have eroded some gene functions related to heat tolerances during the course of the expansion. Our previous data also suggested that cold

tolerances improve towards the expansion front, where individuals track cooler, more variable climates (Lancaster *et al.* 2015). Here we find that genes and isoforms significantly overexpressed under cold stress were remarkably variable between regions (only one such isoform, of unknown function, was conserved between regions; Fig. 2, Table 1). High geographic variability in the transcriptional regulation of cold tolerance supports data from this and other studies, indicating that cold tolerances of ectotherms are often more labile and more adaptable to local conditions than are heat tolerances (Castañeda *et al.* 2004; Kimura 2004; Lancaster *et al.* 2015). These results also suggest some historical contingency and the existence of multiple genetic pathways in the evolution of cold tolerance, compared to the evolution of heat tolerance (Hoffmann *et al.* 2013; Araújo *et al.* 2013; Lancaster *et al.* 2015).

The combined effects of release from selection on upper thermal tolerances, and intensified selection on lower thermal tolerances during range expansions (Lancaster et al. 2015, Lancaster in review), suggest that some genes involved in heat tolerance in the core may be evolutionarily or developmentally flexible, and could therefore switch functionality during the progress of range expansion. In support of this prediction, we find that of 69 genes exhibiting significant region x treatment effects on expression levels, 60 (87%) exhibited overexpression in response to heat in the core, but switched to upregulation in response to cold at the range edge. Many of these transcripts that switch function map to genes involved in cell growth and cell maintenance, rather than acute stress response (see below). This suggests that core individuals are more likely to sustain or increase key cellular functions under heat stress as part of their local adaptation to warmer climates. Range edge individuals, however, which lack locally adapted heat tolerances but exhibit greater cold tolerance than core individuals, may instead sustain or increase key cellular functions under moderate cold stress as part of local adaptation to maintain normal functioning under cooler ambient conditions (Kim *et al.* 2004; Dresios *et al.* 2005). Selection to maintain body functions under

low temperatures (when metabolic rates of ectotherms are very low) should thus be strong at the range edge and should override selection against heat stress, which is presumably weak or non-existent.

Specific genes and isoforms with conserved patterns of gene expression between core and range edge populations

Twenty-one genes and isoforms that were upregulated in response to heat stress were highly conserved across populations (Table 1), including two transcripts mapping to heat shock protein 70 (HSP70), and nine distinct isoforms mapping to six separate genes for HSP20. HSP20 is an ATP-independent first line of defence that commonly prevents protein denaturation in response to stress in insects (King & MacRae 2015), while HSP70 is a highly conserved, ATP-dependent molecular chaperone that facilitates protein homeostasis under a variety of conditions including thermal stress, and is also instrumental in cellular decision making under stress, such as the initiation or inhibition of apoptosis (Beere 2004; King & MacRae 2015). Variation in HSP expression in response to thermal stress has previously been observed within and among species, depending on age, sex, previous experience and other stressors present (Zhao & Jones 2012; King & MacRae 2015). Here we find that shared, differentially-expressed HSP genes (Table 1) generally exhibit declining expression levels towards the cooler, expanding range margin, although upper thermal tolerance phenotypes do not (Lancaster *et al.* 2015). In general the relationship between HSP expression levels and thermal tolerance phenotypes has been difficult to quantify (Telonis-Scott *et al.* 2014). Our results, however, do corroborate a recent meta-analysis showing that HSP expression levels are generally higher in populations adapted to warmer climates, irrespective of their correlation with individually observed thermal tolerance levels (Porcelli *et al.* 2015).

Also upregulated in response to heat in both regions were two transcripts mapping to *Samui*, and one isoform of *DnaJ-1*, both of which are Bcl-2-associated anthanogene (BAG)-family regulators of HSP70 protein activity (Qiu *et al.* 2006; Colinet & Hoffmann 2010). These proteins are co-chaperones that regulate the activity and effectiveness of HSPs in protecting against cellular damage under thermal stress (King & MacRae 2015). Expression levels of *Samui* in silkworms (*Bombyx mori*) and its *Drosophila* homolog *Starvin* have previously been found to be upregulated in response to heat or cold stress (Colinet & Hoffmann 2010; Telonis-Scott *et al.* 2014), with some evidence of geographic variation in the transcriptional response of *Starvin* to thermal stress (Telonis-Scott *et al.* 2014). In *I. elegans*, these genes and isoforms were upregulated strongly in response to heat stress (vs. cold stress) in both populations, with the two differentially expressed isoforms of *Samui* showing similar patterns across regions (Table 1). As in HSPs, expression levels of these co-chaperones were generally lower in the range edge than in the core, potentially also indicative of molecular changes reflecting release from selection on upper thermal tolerances there.

Overexpression under heat stress was also conserved across populations at several genes that are less traditionally linked with heat stress responses in other system. These include a gene mapping to Rho-associated protein kinase-1 (*ROCK1*), which is involved in numerous cell functions involving cytoskeletal rearrangements and regulation of gene expression (Amano *et al.* 2010). Relevantly, this gene has also previously been implicated in regulating the formation of stress granules or stress-induced apoptosis (Tsai & Wei 2010). Transcripts at a single gene mapping to adapter protein-4 complex subunit beta-1 (*AP4B1*) were upregulated in response to heat in both regions. This molecule plays a critical role in vesicle formation and trafficking involved in neural function within the brain (Wang *et al.* 2013; Abdollahpour *et al.* 2015). To our knowledge this is the first study linking AP-4

complex expression with thermal stress. Its upregulation may form part of an adaptive response to heat-induced neural damage and disruption of synaptic functionality (Robertson 2004) or as part of a behavioural thermoregulatory response. Finally, two genes associated with retrotransposon activity plus an additional gene associated with histone methylation, which is involved in regulation of gene expression (Martin & Zhang 2005), were upregulated in response to heat in both regions. Relatively higher expression of (largely damaging) transposable elements (TEs) at the expansion front than in the core may reflect loss of adaptive suppression mechanisms during the range shift, under drift or relaxed selection as described above. More detail on putative TE and histone methylation functions is discussed below.

Genes and isoforms exhibiting among-region variation in gene expression under thermal stress

Transcripts at a gene mapping to heat shock protein 90 (HSP90) were significantly upregulated in response to heat in the core region at a greater than 10-fold expression difference, but differential expression of this gene was not detected at the range edge (Table 2, Table S5). HSP90 is an ATP-dependent protein chaperone with higher specificity and more specialised function than HSP70, which selectively stabilises proteins at later stages of folding (Clare & Saibil 2013). Loss or weakening of the HSP90 response to heat stress at the range edge contributes to supporting our hypothesis that molecular mechanisms of heat tolerance may become weaker or more limited as range expansions proceed to cooler poleward regions. Upregulated in response to heat in the core but not the range edge were three isoforms mapping to two genes for *twitchin* (table S5), which regulates muscle contraction in *Drosophila* (Ayme-Southgate *et al.* 1991). This regional variation in differential expression may be implicated in concomitant loss of behavioural

thermoregulatory responses to heat stress over the course of the range expansion, under the process of thermal release.

The dominant functional category for genes and isoforms exhibiting regional variation in thermal response was transposable element (TE) activity, specifically of retrotransposons (Table 2, Table S5). Strong regional differences in the specific TEs involved in thermal stress responses is unsurprising given that TEs are mobile and repeated throughout the genome (Wessler 2006). Numerous studies have found that TE activity often increases in response to stress including thermal stress (McClintock 1984; DeSalvo *et al.* 2008; Stapley *et al.* 2015). While failure to suppress activity of TEs under stress is most likely a maladaptive stress response at the individual level (Slotkin & Martienssen 2007), it has been suggested that this response may be partially responsible for geographic variation in thermal tolerance itself (HSP70 promoters are particularly vulnerable to TE insertion; Lerman et al. 2003). Such TE-induced geographic variation in HSP expression or at other loci may also facilitate range expansions and invasions, by generating novel genetic variation as a substrate for adaptation to novel conditions in the new part of the range (Stapley *et al.* 2015). One outstanding question raised by this hypothesis is whether range expansions might be favoured to occur in species that have pre-existing tendency to lose TE suppression in response to stress, or whether the tendency to lose TE suppression under stress is gained during the expansion process (Stapley *et al.* 2015). Our results show approximately equal or declining rates of TE activity in response to thermal stress at the range edge vs. core, suggesting that, if TE activity is a relevant stress response for promoting range expansion, this mechanism may have preadapted *I. elegans* to succeed, rather than arising during the course of the expansion. However, more work is needed to assess how TE responses to stress in expanding, non-expanding and declining species, in order to draw conclusions about the relative importance of this mechanism for facilitating invasion success.

Gene expression at genes mapping to histone lysine methyltransferase was always higher in response to heat than cold stress in each region (Tables 1 and 2). Histone methylation plays an important regulatory role via transcriptional control and chromatin remodelling (Trievel 2004), and has also been shown to play a key role in the suppression of TEs (Slotkin & Martienssen 2007). While the function of elevated transcription rates of these genes in response to thermal stress is unknown, one potential hypothesis is that increased expression of histone lysine methyltransferases during heat stress may play an adaptive role in suppressing excessive TE activity, while suppression of TE activation under cold stress may be mediated by a different mechanism.

In response to cold stress in the core but not the range edge, transcripts were upregulated for a gene mapping to *mucin*, a defensive protein that is commonly expressed in the gut of vertebrates and invertebrates in response to toxins, microbial invasions, or physical trauma (Rayms-Keller *et al.* 2000). Evidence is accumulating that mucins also play a role in heat and cold stress response in *Drosophila* (i.e., the *Frost* locus; Colinet et al. 2010), potentially by reducing oxidative stress and maintaining osmoregulation across gut epithelial cell membranes under thermal stress (Colinet *et al.* 2010). Additional work is needed to identify the function of the particular mucin-like protein identified here, and particularly why differential expression at this cold-responsive gene in the core is not observed at the range edge, where the environment is colder and observed organism-level cold tolerances were improved.

Transcripts mapping to genes with regulatory, metabolic, homeostatic, synthesis/repair, transport, signalling, or adhesion functionality (summarised in Table 2 as "cell function/life cycle") were also more often differentially expressed in response to thermal stress in the core compared to the range edge, and were more often expressed in response to heat than cold stress (Table 2, Table S5). The significance of such regional

variation in thermal response function of each of these genes is currently unknown.

Genes exhibiting region x treatment effects on expression

Forty-eight of 69 transcripts that exhibited a significant region x treatment effect on expression patterns mapped to genes of described functions (Table S6). Two-thirds of these (67%) mapped to genes for ribosomal proteins, which were more highly expressed in response to heat in the core, but in response to cold at the range edge (Fig 3). Previous studies have suggested that genes for ribosomal proteins are commonly down-regulated in response to stress, as the body shifts from a strategy of growth to recovery and repair (Gasch *et al.* 2000; Spriggs *et al.* 2010). However, evidence is accumulating that ribosomal proteins may also be upregulated as part of the thermal stress response to both cold (Graumann *et al.* 1996; Kim *et al.* 2004) and heat stress (Zhao *et al.* 2009), as part of a mechanism to enhance translation or support for proper ribosome function while under moderate thermal stress (Kim *et al.* 2004). This explanation is supported by a recent meta-analysis which suggests that genes involved in translation may have a fundamental role in adaptation to changing environmental temperatures (Porcelli *et al.* 2015).

Seven additional genes that exhibited a similar region x treatment effect (Fig. 3a) each mapped to genes coding for membrane metalloendopeptidase enzymes, cell-surface proteins that have diversified within and among species to inactivate a variety of peptide signals (Turner *et al.* 2001). These ancient enzymes have radiated to perform a range of described functions, from neural control of feeding and locomotion, to immune and reproductive functions, metamorphosis, ageing, homeostasis and cancer (Bland *et al.* 2008). The role of metalloendopeptidase in the thermal stress response is so far unknown, but metalloendopeptidase has been found to increase under heat stress in corals (DeSalvo *et al.* 2008), and activity at a related serine-type endopeptidase locus responds to heat stress in mosquitoes (Zhao *et al.* 2009) and bacteria (Lipinska *et al.* 1990). Insect

metalloendopeptidases have been previously investigated for a potential role in insecticide development (Turner *et al.* 2001). The current study suggests that the function of these molecules may be evolutionarily or developmentally quite labile (see also Bland et al. 2008), and may be integral for facilitating rapid adaptation of insect species to novel and changing climates.

An additional eight genes for putatively adaptive normal functioning under thermal stress also exhibited similar region x treatment effects on expression, in the same direction as the ribosomal proteins (Fig. 3). The relatively greater ability of core individuals to adopt a 'sustain business as usual' response to heat stress is likely an adaptation to warmer climates. In contrast, at the range edge where individuals have lost local adaptation of heat tolerances but selection on cold tolerances has intensified, the need to sustain normal cellular function under cold stress may be more critical for fitness (Fig. 3a). These patterns further suggest that the cellular threshold for whether to invest in sustaining normal/enhanced transcriptional machinery while under stress, vs. shifting away from normal cell function and towards survival, may commonly be regionally or temporally variable, especially during range expansions. Geographic variation in this decision point may reflect different degrees of local adaptation to cold or hot environments. Of the nine transcripts that significantly exhibited the reverse region x treatment expression pattern, such that there was higher expression in response to cold stress in the core, but in response to heat stress at the range edge, all were of previously undescribed function (Fig. 3b, Table S6).

Evolutionary and developmental mechanisms and remaining questions

We identified regional variation in gene expression responses to thermal stress using wild caught individuals expressing phenotypes that have been primed by natural conditions. This approach has two primary strengths over use of laboratory-reared individuals. First,

natural phenotypes of wild-caught individuals are often highly dependent on developmental priming effects that are unknown or cannot be replicated in a laboratory setting (Boysen *et al.* 2011). Second, high rates of mortality in the laboratory mean that lab-bred individuals often represent an unnatural subsample of individuals that can survive in the laboratory environment (Wagner 2000). Thus in order to understand ecologically relevant responses to thermal stress, it is essential to work with naturally occurring phenotypes. However, the use of wild-caught individuals has several drawbacks. First, the age of individuals cannot be precisely known. Although we selected only mature adult females for comparison, and ontogenetic effects on gene expression within the adult life stage are likely limited to a very few loci (Lund *et al.* 2002), such effects may obscure relevant patterns at these few loci. Second, use of wild caught individuals means that we cannot distinguish intrinsic (genetic) from ecological (plastic) causes of regional differences in gene expression patterns, although recent work by ourselves and other authors suggest that plasticity and evolutionary change act in the same direction during colonization of novel thermal environments (Lancaster *et al.* 2015; Mäkinen *et al.* 2016). These alternative sources of geographic variation are discussed in more detail below.

Causes of declines in molecular responses to thermal stress at the approach of the range expansion front may include developmental plasticity (e.g., lack of prior exposure to hot weather, resulting in a less well-primed heat response), or reflect genetic drift leading to loss of alleles involved in thermal tolerance that are not maintained under selection, or both. Regardless of underlying causes, regional variation in these thermal tolerance mechanisms is reflective of regional variation in the thermal selective environment (Lancaster et al. 2015). Both the numbers of genes significantly upregulated under heat stress, and the intensity of response of those genes, were lower at the range edge in *I. elegans*. However, this pattern did not reflect among-region differences in mean thermal tolerance phenotypes, because absolute

tolerances were similar between core and edge regions (Lancaster *et al.* 2015). This discordance between regional patterns of gene expression levels and physiological responses is consistent with previous results showing no correlation between physiological thermal tolerance and expression levels of heat shock proteins (Jensen *et al.* 2010). However, thermal response phenotypes are typically complex, redundant, and robust (El-Samad *et al.* 2005), and therefore regional declines in molecular function during range expansions may not become reflected in mean thermal response phenotypes at the whole-organism level until considerably greater erosion of underpinning molecular mechanisms occurs via further genetic drift and/or further loss of adaptive environmental priming cues.

In contrast to the results for genes with a significant effect on heat tolerance, we find that whole-transcriptome responses to heat stress are only slightly reduced at the range edge compared to the core (Fig. 2b; where the core-heat – core-cold distance is only marginally larger than the edge-heat – edge-cold distance). This suggests that while genes with major effect on heat tolerance phenotypes respond (via evolutionary or plastic mechanisms) to release from selection during expansion, many genes with minor-effect responses to heat stress may remain largely undifferentiated between regions during rapid expansions. Maintenance of ancestrally high heat tolerances may depend on maintaining function at these genes of minor effect over the course of the expansion.

Similarly, we lack information on the underlying genetic or developmental basis underpinning regional turnover of genes upregulated in response to cold stress. Across a variety of species, cold tolerance is known to be highly evolutionarily labile, and also typically exhibits high phenotypic plasticity, in comparison to heat tolerance. We found that cold tolerances themselves exhibited increased phenotypic plasticity, in the form of beneficial acclimation to recent weather events, at the approach of the range limit (Lancaster *et al.* 2015) in comparison to relatively non-plastic cold response phenotypes in the core. These

result are qualitatively consistent with well-established theoretical and empirical results linking the evolution of plasticity to range expansions (Sexton *et al.* 2009; Lande 2009). It may be that the genes significantly expressed at the range edge in response to cold stress are more environmentally inducible than genes expressed in the core (Table 2), reflecting regional genetic variation in thermal reaction norms and a phenotypic expediting effect (Lande 2009). Alternatively, acquired plasticity in cold tolerance at the range margin may simply reflect more extreme environmental stress there. Finally, it is possible that gene expression levels in general are poorly predictive of cold tolerance phenotypes: some evidence suggests that, in comparison to heat tolerance, cold tolerance is more likely to be mediated at the translational than transcriptional stage (Spriggs *et al.* 2010). Consistent with its high phenotypic and evolutionary lability, this would suggest that cold tolerance is a complex, highly modular trait, requiring much more work to mechanistically unravel.

Although cold tolerances were highly divergent between regions (Lancaster *et al.* 2015), and genes significantly upregulated in response to cold stress were highly variable between the core and the range edge (Fig. 2a,c, Table 2), whole-transcriptome responses to cold stress were surprisingly very similar across the two populations (Fig 2b). This pattern might suggest that very few genes are involved in the cold-stress response, an interpretation consistent with previously reported high evolutionary lability of cold tolerances (Araújo *et al.* 2013). Alternatively, this result may further support the hypothesis that individual proteins and cellular functions are likely to respond to cold stress is via translational than transcriptional responses.

Genes exhibiting a significant region x treatment effect on expression levels are highly likely to reflect regional variation in thermal selection regimes, and suggest that the threshold for switching between growth/maintenance and survival functions has changed over the course of the range expansion (i.e., the threshold has shifted to lower temperatures for

both heat and cold responses). Again, it is currently unknown whether this shift in threshold temperature for switching between normal growth and maintenance to survival functions reflects local adaptation or acclimation. In either case, however, these results suggest that the ability to undergo rapid transitions between being warm-adapted to being cold-adapted, at the individual or population level, likely underpins successful range expansion responses to climate change. Having identified important genetic sources of regional variation, in decoupled patterns of response to heat vs. cold, and against the background of natural conditions (i.e., in wild-caught individuals), uncovering the heritability and developmental plasticity underlying these molecular mechanisms in a laboratory-bred individuals is the logical next step.

References

Abdollahpour H, Alawi M, Kortüm F *et al.* (2015) An AP4B1 frameshift mutation in siblings with intellectual disability and spastic tetraplegia further delineates the AP-4 deficiency syndrome. *European journal of human genetics : EJHG*, **23**, 256–9.

Amano M, Nakayama M, Kaibuchi K (2010) Rho-kinase/ROCK: A key regulator of the cytoskeleton and cell polarity. *Cytoskeleton (Hoboken, N.J.)*, **67**, 545–54.

Angert AL, Crozier LG, Rissler LJ *et al.* (2011) Do species' traits predict recent shifts at expanding range edges? *Ecology letters*, **14**, 677–89.

Araújo MB, Ferri-Yáñez F, Bozinovic F *et al.* (2013) Heat freezes niche evolution. *Ecology letters*, **16**, 1206–19.

Ashburner M, Ball CA, Blake JA *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics*, **25**, 25–9.

Ayme-Southgate A, Vigoreaux J, Benian G, Pardue ML (1991) Drosophila has a twitchin/titin-related gene that appears to encode projectin. *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 7973–7.

Bebber DP, Ramotowski MAT, Gurr SJ (2013) Crop pests and pathogens move polewards in a warming world. *Nature Climate Change*, **3**, 985–988.

Beere HM (2004) "The stress of dying": the role of heat shock proteins in the regulation of apoptosis. *Journal of cell science*, **117**, 2641–51.

Bellard C, Bertelsmeier C, Leadley P, Thuiller W, Courchamp F (2012) Impacts of climate change on the future of biodiversity. *Ecology letters*, **15**, 365–77.

Bland ND, Pinney JW, Thomas JE, Turner AJ, Isaac RE (2008) Bioinformatic analysis of the neprilysin (M13) family of peptidases reveals complex evolutionary and functional relationships. *BMC evolutionary biology*, **8**, 16.

Boysen P, Eide DM, Storset AK (2011) Natural killer cells in free-living Mus musculus have a primed phenotype. *Molecular ecology*, **20**, 5103–10.

Buckley LB, Kingsolver JG (2012) Functional and Phylogenetic Approaches to Forecasting Species' Responses to Climate Change. *Annual Review of Ecology, Evolution, and Systematics*, **43**, 205–226.

Castañeda LE, Lardies* MA, Bozinovic F (2004) Adaptive latitudinal shifts in the thermal physiology of a terrestrial isopod. *Evolutionary Ecology Research*, **6**, 579–593.

Cham S, Nelson B, Parr A *et al.* (2014) *Atlas of Dragonflies in Britain and Ireland.* the Field Studies Council for the Biological Records Centre, Centre for Ecology & Hydrology, with the British Dragonfly Society.

Chauhan P, Hansson B, Kraaijeveld K *et al.* (2014) De novo transcriptome of Ischnura elegans provides insights into sensory biology, colour and vision genes. *BMC genomics*, **15**, 808.

Chen I-C, Hill JK, Ohlemüller R, Roy DB, Thomas CD (2011) Rapid range shifts of species associated with high levels of climate warming. *Science (New York, N.Y.)*, **333**, 1024–1026.

Clare DK, Saibil HR (2013) ATP-driven molecular chaperone machines. *Biopolymers*, **99**, 846–59.

Colinet H, Hoffmann A (2010) Gene and protein expression of Drosophila Starvin during cold stress and recovery from chill coma. *Insect biochemistry and molecular biology*, **40**, 425–8.

Colinet H, Lee SF, Hoffmann A (2010) Functional characterization of the Frost gene in Drosophila melanogaster: importance for recovery from chill coma. *PloS one*, **5**, e10925.

Comte L, Murienne J, Grenouillet G (2014) Species traits and phylogenetic conservatism of climate-induced range shifts in stream fishes. *Nature communications*, **5**, 5023.

Conesa A, Götz S, García-Gómez JM *et al.* (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics (Oxford, England)*, **21**, 3674–6.

Culhane AC, Perrière G, Considine EC, Cotter TG, Higgins DG (2002) Between-group analysis of microarray data. *Bioinformatics (Oxford, England)*, **18**, 1600–8.

Culhane AC, Thioulouse J, Perrière G, Higgins DG (2005) MADE4: an R package for multivariate analysis of gene expression data. *Bioinformatics (Oxford, England)*, **21**, 2789–90.

DeSalvo MK, Voolstra CR, Sunagawa S *et al.* (2008) Differential gene expression during thermal stress and bleaching in the Caribbean coral Montastraea faveolata. *Molecular ecology*, **17**, 3952–71.

Dijkstra K-DB, Lewington R (2006) *Field Guide to the Dragonflies of Britain and Europe*. British Wildlife Publishing, Gillingham, Dorset, UK.

Dresios J, Aschrafi A, Owens GC *et al.* (2005) Cold stress-induced protein Rbm3 binds 60S ribosomal subunits, alters microRNA levels, and enhances global protein synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 1865–70.

El-Samad H, Kurata H, Doyle JC, Gross CA, Khammash M (2005) Surviving heat shock: control strategies for robustness and performance. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 2736–41.

Gasch AP, Spellman PT, Kao CM *et al.* (2000) Genomic expression programs in the response of yeast cells to environmental changes. *Molecular biology of the cell*, **11**, 4241–57.

Grabherr MG, Haas BJ, Yassour M *et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology*, **29**, 644–52.

Graumann P, Schröder K, Schmid R, Marahiel MA (1996) Cold shock stress-induced proteins in Bacillus subtilis. *Journal of bacteriology*, **178**, 4611–9.

Haas B (2015) Differential Expression Analysis Using a Trinity Assembly.

Hickling R, Roy DB, Hill JK, Thomas CD (2005) A northward shift of range margins in

British Odonata. *Global Change Biology*, **11**, 502–506.

Hidalgo-Galiana A, Sánchez-Fernández D, Bilton DT, Cieslak A, Ribera I (2014) Thermal niche evolution and geographical range expansion in a species complex of western Mediterranean diving beetles. *BMC evolutionary biology*, **14**, 187.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hoffmann AA, Chown SL, Clusella-Trullas S (2013) Upper thermal limits in terrestrial ectotherms: how constrained are they? (C Fox, Ed,). *Functional Ecology*, **27**, 934–949.

Jensen LT, Cockerell FE, Kristensen TN *et al.* (2010) Adult heat tolerance variation in Drosophila melanogaster is not related to Hsp70 expression. *Journal of experimental zoology. Part A, Ecological genetics and physiology*, **313**, 35–44.

Jones P, Binns D, Chang H-Y *et al.* (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics (Oxford, England)*, **30**, 1236–40.

Kim K-Y, Park S-W, Chung Y-S *et al.* (2004) Molecular cloning of low-temperature-inducible ribosomal proteins from soybean. *Journal of experimental botany*, **55**, 1153–5.

Kimura MT (2004) Cold and heat tolerance of drosophilid flies with reference to their latitudinal distributions. *Oecologia*, **140**, 442–9.

King AM, MacRae TH (2015) Insect heat shock proteins during stress and diapause. *Annual review of entomology*, **60**, 59–75.

Klok CJ, Sinclair BJ, Chown SL (2004) Upper thermal tolerance and oxygen limitation in terrestrial arthropods. *The Journal of experimental biology*, **207**, 2361–2370.

Krehenwinkel H, Rödder D, Tautz D (2015) Eco-Genomic analysis of the poleward range expansion of the wasp spider Argiope bruennichi shows rapid adaptation and genomic admixture. *Global change biology*.

Lancaster LT Widespread, ongoing range expansions shape latitudinal variation in insect thermal limits.

Lancaster LT, Dudaniec RY, Hansson B, Svensson EI (2015) Latitudinal shift in thermal niche breadth results from thermal release during a climate-mediated range expansion. *Journal of Biogeography*, n/a–n/a.

Lande R (2009) Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *Journal of evolutionary biology*, **22**, 1435–46.

Lerman DN, Michalak P, Helin AB, Bettencourt BR, Feder ME (2003) Modification of heat-shock gene expression in Drosophila melanogaster populations via transposable elements. *Molecular biology and evolution*, **20**, 135–44.

Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics*, **12**, 323.

Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics (Oxford, England)*, **22**, 1658–9.

Lipinska B, Zylicz M, Georgopoulos C (1990) The HtrA (DegP) protein, essential for Escherichia coli survival at high temperatures, is an endopeptidase. *Journal of bacteriology*, **172**, 1791–7.

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**, 550.

Lund J, Tedesco P, Duke K *et al.* (2002) Transcriptional Profile of Aging in C. elegans. *Current Biology*, **12**, 1566–1573.

Mäkinen H, Papakostas S, Vøllestad LA, Leder EH, Primmer CR (2016) Plastic and Evolutionary Gene Expression Responses Are Correlated in European Grayling (Thymallus thymallus) Subpopulations Adapted to Different Thermal Environments. *The Journal of heredity*, **107**, 82–89.

Martin C, Zhang Y (2005) The diverse functions of histone lysine methylation. *Nature reviews. Molecular cell biology*, **6**, 838–49.

McClintock B (1984) The significance of responses of the genome to challenge. *Science (New York, N.Y.)*, **226**, 792–801.

Parmesan C, Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, **421**, 37–42.

Parra G, Bradnam K, Korf I (2007) CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics (Oxford, England)*, **23**, 1061–7.

Porcelli D, Butlin RK, Gaston KJ, Joly D, Snook RR (2015) The environmental genomics of metazoan thermal adaptation. *Heredity*, **114**, 502–514.

Preisser EL, Elkinton JS, Abell K (2008) Evolution of increased cold tolerance during range expansion of the elongate hemlock scale Fiorinia externa Ferris (Hemiptera: Diaspididae). *Ecological Entomology*.

Qiu X-B, Shao Y-M, Miao S, Wang L (2006) The diversity of the DnaJ/Hsp40 family, the crucial partners for Hsp70 chaperones. *Cellular and molecular life sciences : CMLS*, **63**, 2560–70.

Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)*, **26**, 841–2.

Rayms-Keller A, McGaw M, Oray C, Carlson JO, Beaty BJ (2000) Molecular cloning and characterization of a metal responsive Aedes aegypti intestinal mucin cDNA. *Insect molecular biology*, **9**, 419–26.

Robertson RM (2004) Thermal stress and neural function: adaptive mechanisms in insect model systems. *Journal of Thermal Biology*, **29**, 351–358.

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)*, **26**, 139–40.

Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome biology*, **11**, R25.

Le Rouzic A, Hansen TF, Gosden TP, Svensson EI (2015) Evolutionary time-series analysis reveals the signature of frequency-dependent selection on a female mating polymorphism. *The American naturalist*, **185**, E182–96.

Sánchez-Guillén RA, Córdoba-Aguilar A, Hansson B, Ott J, Wellenreuther M (2015) Evolutionary consequences of climate-induced range shifts in insects. *Biological Reviews*, n/a–n/a.

Sexton JP, McIntyre PJ, Angert AL, Rice KJ (2009) Evolution and Ecology of Species Range Limits. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 415–436.

Sinervo B, Méndez-de-la-Cruz F, Miles DB *et al.* (2010) Erosion of lizard diversity by climate change and altered thermal niches. *Science (New York, N.Y.)*, **328**, 894–9.

Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of the genome. *Nature reviews. Genetics*, **8**, 272–85.

Smit A, Hubley R, Green P RepeatMasker Open-3.0.

Spriggs KA, Bushell M, Willis AE (2010) Translational regulation of gene expression during conditions of cell stress. *Molecular cell*, **40**, 228–37.

Stapley J, Santure AW, Dennis SR (2015) Transposable elements as agents of rapid adaptation may explain the genetic paradox of invasive species. *Molecular ecology*, **24**, 2241–52.

Svensson EI, Abbott J (2005) Evolutionary dynamics and population biology of a polymorphic insect. *Journal of evolutionary biology*, **18**, 1503–14.

Svensson EI, Abbott J, Hardling R (2005) Female polymorphism, frequency dependence, and rapid evolutionary dynamics in natural populations. *The American naturalist*, **165**, 567–76.

Telonis-Scott M, Clemson AS, Johnson TK, Sgrò CM (2014) Spatial analysis of gene regulation reveals new insights into the molecular basis of upper thermal limits. *Molecular ecology*, **23**, 6135–51.

Thomas CD (2013) Local diversity stays about the same, regional diversity increases, and global diversity declines. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 19187–8.

Thomas CD, Bodsworth EJ, Wilson RJ *et al.* (2001) Ecological and evolutionary processes at expanding range margins. *Nature*, **411**, 577–81.

Trievel RC (2004) Structure and Function of Histone Methyltransferases. *Critical Reviews in Eukaryotic Gene Expression*, **14**, 147–170.

Tsai N-P, Wei L-N (2010) RhoA/ROCK1 signaling regulates stress granule formation and apoptosis. *Cellular signalling*, **22**, 668–75.

Turner AJ, Isaac RE, Coates D (2001) The neprilysin (NEP) family of zinc metalloendopeptidases: genomics and function. *BioEssays : news and reviews in molecular, cellular and developmental biology*, **23**, 261–9.

Wagner A (2000) Inferring lifestyle from gene expression patterns. *Molecular biology and evolution*, **17**, 1985–7.

Wang H, Xu Y, Chen M *et al.* (2013) Genetic association study of adaptor protein complex 4 with cerebral palsy in a Han Chinese population. *Molecular biology reports*, **40**, 6459–67.

Warnes GR, Bolker B, Bonebakker L *et al.* (2015) gplots: Various R programming tools for plotting data.

Wessler SR (2006) Transposable elements and the evolution of eukaryotic genomes. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 17600–1.

Zhao L, Jones WA (2012) Expression of heat shock proteins in insect stress responses. *Invertebrate Survival Journal*, **9**, 93–101.

Zhao L, Pridgeon JW, Becnel JJ, Clark GG, Linthicum KJ (2009) Identification of genes differentially expressed during heat shock treatment in Aedes aegypti. *Journal of medical entomology*, **46**, 490–5.

Data Accessibility:

- RNA sequences: NCBI Bioproject number PRJNA304220; Biosample links

  SAMN04298996-SAMN04298999; Read archives SRR2962606-SRR2962609.

- Annotated assembly and output files from DESeq2, RSEM: Dryad

  doi:10.5061/dryad.4qf2h.

- **Author contributions:** B.H. and E.I.S. conceived of the study. L.L. developed the hypo**theses to be tested.** L.L. and R.D. collected the field data and samples. All 6 authors contributed to planning RNA-seq analyses. P.C. and L.L. analysed the data. L.L. wrote the manuscript, which all 6 authors edited.

-

Supporting information:

- Table S1: Summary report of trimming steps applied on raw RNA-seq data to obtain quality reads.
- Table S2: Statistics of initial (first assembly from Trinity) and final assembly (assembly after applying three step filtering).
- Table S3: RSEM % reads mapped per sample.
- Table S4: Top 5 significantly enriched GO terms for differentially expressed genes within each population.
- Table S5: Genes and isoforms differentially expressed in response to heat or cold stress in either core or edge populations.
- Table S6: Genes and isoforms exhibiting significant region x treatment effects on expression levels.
- Figure S1: Gene expression profiles for all 780 differentially expressed genes.

Table 1: Genes and gene isoforms that showed > 10-fold (i.e. > 3.32 $\log_2$) differential expression in response to heat vs. cold stress in both the core and the range edge population. Also shown is the difference in log(fold change) between core and edge.

Transcript ID	Mapping description	Function	Log_2 Fold Change, Heat vs. Cold treatment		
			Core	Edge	Difference
c50968_g1_i1	heat shock protein 70	Heat shock protein	4.26	3.80	0.46
c50968_g2_i1	heat shock protein 70	Heat shock protein	4.15	3.73	0.42
c49287_g1_i1	heat shock protein 20 - Insect	Heat shock protein	5.08	4.36	0.72
c54905_g1_i1	heat shock protein 20 - Insect	Heat shock protein	6.76	6.66	0.10
c61896_g3_i1	heat shock protein 20 - Insect	Heat shock protein	8.35	8.36	0.01
c53003_g1_i3	protein lethal essential for life - HSP20 Insect	Heat shock protein	4.86	3.88	0.98
c53003_g1_i5	protein lethal essential for life - HSP20 Insect	Heat shock protein	6.17	5.35	0.82
c53003_g1_i7	protein lethal essential for life - HSP20 Insect	Heat shock protein	5.49	4.61	0.88
c53003_g1_i8	protein lethal essential for life - HSP20 Insect	Heat shock protein	5.25	4.17	1.08
c55839_g1_i1	protein lethal essential for life - HSP20 Insect	Heat shock protein	8.62	8.19	0.43
c61896_g2_i1	protein lethal essential for life - HSP20 Insect	Heat shock protein	7.44	6.92	0.52
c51726_g1_i1	bag domain-containing protein samui	Heat shock protein cochaperone	4.38	3.75	0.63
c51726_g1_i2	bag domain-containing protein samui	Heat shock protein cochaperone	3.98	3.39	0.59
c56609_g2_i1	dnaj protein homolog 1-like	Heat shock protein cochaperone	3.80	3.37	0.43
c36879_g1_i1	rho-associated protein kinase 1-like	Stress granule formation / apoptosis in response to stress	3.66	3.66	0.00
c55735_g1_i1	rna-directed dna polymerase from mobile element jockey-like	Retrotransposon activity	3.57	6.81	-3.24
c62910_g1_i1	gag-pol polyprotein	Retrotransposon activity	5.82	6.61	-0.79
c61577_g1_i1	histone-lysine n-methyltransferase partial	Epigenetic modification of DNA structure/function	3.94	4.77	-0.83
c58235_g1_i2	ap-4 complex subunit beta-1-partial	Transport - vesicle activity, neural system	5.93	3.49	2.44
c61732_g3_i2	hypothetical protein aael004188 - *Aedes*	Unknown	-4.51	-3.51	1.00
c44423_g1_i1	uncharacterized protein - various insects	Unknown	4.78	4.47	0.31
c58235_g1_i1	hypothetical protein - various animalia	Unknown	4.48	5.67	-1.19

Table 2: Functional categories for genes and isoforms that were differentially expressed in response to heat vs. cold in only one of the populations. See Table S5 for more details on individual genes and isoforms.

Region	Treatment	Stress response	Transposon activity	Methylation	Regulatory	Cell function/ life cycle	Muscle function	Defence	Unknown	Sum
Core	heat	1	10	2	5	17	3	0	9	47
	cold	0	9	0	3	8	0	1	7	28
Edge	heat	0	6	2	3	9	0	0	8	28
	cold	0	7	0	0	7	0	0	4	18

Figure 1: Study scope and sampling locations for *Ischnura elegans* in Sweden. Open, grey circles: capture sites from which phenotypic data on upper and lower thermal tolerances has previously been summarised (Lancaster *et al.* 2015). Filled, black circles: populations included in the present analyses (Höje Å 14 in the core region in southern Sweden and Nora in the edge region in central Sweden). Dashed line: previously described range limit for *I. elegans* (Dijkstra & Lewington 2006)—our discovery of populations beyond this limit is consistent with the species' rapid, ongoing expansion. Background shading represents geographic variation in mean annual temperature, calculated over 1950-2000 (Bioclim Bio1; Hijmans *et al.* 2005), with units on the scale bar in ° Celcius.

Figure 2: Variation in gene expression under heat vs. cold stress at the species core range (core population) and at the poleward expansion front (edge population). A) Venn diagram of significantly differentially expressed genes (and > 10-fold change) between treatments within populations. Size of circles corresponds to the number of differentially expressed genes in each population; a greater number of genes and transcripts were overall involved in thermal stress response in the core population (n = 97) than at the expanding range edge (n = 68). Twenty-two genes were differentially expressed in both populations. Overall, more differentially expressed genes and isoforms were upregulated in response to heat (red) than cold (blue). B) A principal component analysis of all 71,904 expressed transcripts in each treatment and population. Points represent individuals, and are connected by solid lines. The major and minor axis of each confidence ellipse are represented by dashed lines. C) Differential expression heat map with all genes and isoforms that were significantly differentially expressed between treatments in at least one population depicted (i.e., the same 143 genes as in panel a). Along the left column, genes and isoforms are color-coded by

function (blue = cell growth, structure, or homeostasis, yellow = histone methyltransferase, red = stress response, green = regulatory, purple = transposable element, dark green = defence, tan = muscle function, grey = unknown). Heat response genes and isoforms cluster at the top-right, while cold-response genes and isoforms cluster at the bottom-left.

Figure 3: Genes significantly differentially expressed according to region x treatment: A) Genes exhibiting a significant negative region x treatment interaction (n = 60), i.e., overexpressed in response to heat stress in the core, but in response to cold stress at the range edge. This pattern suggests that thresholds for cellular responses to heat or cold are altered over the course of the range expansion, in response to changing thermal selection regimes. B) Only nine genes showed the reverse pattern, (i.e. underexpressed in response to heat stress in the core, and in response to cold stress at the edge), and all of these are of unknown function.