MOLECULAR ECOLOGY

Signatures of local adaptation along environmental gradients in a range-expanding damselfly (Ischnura elegans)

Journal:	*Molecular Ecology*
Manuscript ID	MEC-18-0031.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Dudaniec, Rachael; Macquarie University, Department of Biological Sciences Yong, Chuan Ji; Macquarie University Department of Biological Sciences Lancaster, Lesley; University of Aberdeen, School of Biological Sciences Svensson, Erik; Lund University, Department of Biology Hansson, Bengt; Lund University, Department of Biology
Keywords:	range expansion, landscape genomics, Ischnura, local adaptation, environmental association analysis, damselfly

Signatures of local adaptation along environmental gradients in a range-expanding damselfly (*Ischnura elegans*)

Rachael Y. Dudaniec[1*], Chuan Ji Yong[1], Lesley T. Lancaster[2], Erik I. Svensson[3], Bengt Hansson[3]

[1]*Department of Biological Sciences, Macquarie University, North Ryde, Sydney, Australia, 2109*

[2] *School of Biological Sciences, University of Aberdeen, Aberdeen, United Kingdom, AB24 2TZ*

[3] *Department of Biology, Lund University, Lund, Sweden, SE-223 62*

Short running title

Selection signatures along a range expansion

***Corresponding Author**

Rachael Y Dudaniec

Department of Biological Sciences, Macquarie University, North Ryde, Sydney, Australia, 2109; Fax: +61 2 9850 8245; Email: rachael.dudaniec@mq.edu.au; Phone +61 (2) 9850 8193.

Keywords: range expansion, landscape genomics, *Ischnura*, local adaptation, environmental association analysis, insects.

Abstract

Insect distributions are shifting rapidly in response to climate change and are undergoing rapid evolutionary change. We investigate the molecular signatures underlying local adaptation in the range-expanding damselfly, *Ischnura elegans*. Using a landscape genomic approach combined with generalized dissimilarity modelling (GDM), we detect selection signatures on loci via allelic frequency change along environmental gradients. We analyse 13,612 Single Nucleotide Polymorphisms (SNPs), derived from Restriction site-Associated DNA sequencing (RADseq), in 426 individuals from 25 sites spanning the *I. elegans* distribution in Sweden, including its expanding northern range edge. Environmental association analysis (EAA) and the magnitude of allele frequency change along the range expansion gradient revealed significant signatures of selection in relation to high maximum summer temperature, high mean annual precipitation, and low wind speeds at the range edge. SNP annotations with significant signatures of selection revealed gene functions associated with ongoing range expansion, including heat shock proteins (*HSP40* and *HSP70*), ion transport (V-ATPase) and visual processes (*long wavelength-sensitive opsin*), which have implications for thermal stress response, salinity tolerance and mate discrimination, respectively. We also identified environmental thresholds where climate-mediated selection is likely to be strong, and indicate that *I. elegans* is rapidly adapting to the climatic environment during its ongoing range expansion. Our findings empirically validate an integrative approach for detecting spatially explicit signatures of local adaptation along environmental gradients.

Introduction

Adaptation is driven by the interaction between heritable phenotypes and local selective environments, and the outcomes of this process vary along species' ranges, and are shaped by spatial variation in selection pressures, standing genetic diversity, and demographic potential (Bridle & Vines, 2006). Theory and some empirical evidence suggest that directional selection may be particularly pronounced at species' range limits where environments tend to be less optimal for growth and reproduction (Kirkpatrick & Barton, 1997; Lancaster, 2016; Warren et al., 2001). In addition to lower habitat suitability, range limits are typically characterised by stochastic genetic and population dynamics due to lower effective population sizes (N_e), which might increase genetic drift and thereby among-population genetic differentiation (Swaegers et al., 2013; Trumbo et al., 2016). Due to gene flow from populations adapted to conditions in the range core, peripheral, range limit populations are expected to be maladapted relative to core populations (Bridle and Vines 2006; Kirkpatrick and Barton 1997). However, with adequate genetic variation, maladaptation in peripheral populations may be counteracted by rapid adaptive evolution to novel environmental pressures, which can facilitate species' range expansions and their future persistence (Colautti & Barrett, 2013).

Evolutionary and landscape genomics approaches have recently enabled the characterisation of the role of environmental variables in explaining signatures of local adaptation at the molecular level (Ahrens et al., 2018; Hoban et al., 2016; Rellstab, Gugerli, Eckert, Hancock, & Holderegger, 2015). Searching for loci underpinning local adaptation is a formidable challenge that has become increasingly accessible via new analytical tools that identify loci with higher than expected genetic divergence among populations (Fst outlier tests: e.g. Foll & Gaggiotti, 2008; Whitlock & Lotterhos, 2015) or exhibit high correlation

with spatially-explicit environmental variables (Environmental Association Analysis; EAA: Rellstab et al., 2015), while accounting for neutral genetic structure. However, identifying a few specific loci that differ dramatically among populations in allele frequencies under putative locally-divergent selection regimes is but one part of the question, while we should also strive to understand how the strength of selection operates across many loci along environmental gradients, and the functional significance of such loci. For species undergoing range expansion in response to climate change functional loci that respond with shifts in allelic frequencies along environmental gradients will ultimately determine the capacity of a species to adapt and persist.

Genes that are relevant for local adaptation are expected to predictably change their allele frequency along environmental gradients. Such adaptive molecular population differentiation can be quantified via changes in allele frequency among locations across environmental gradients (hereafter 'allelic turnover': Fitzpatrick & Keller, 2015). Signatures of local adaptation can then be teased apart across species distributions. Analytical tools to translate genomic information into signatures of local adaptation have only recently been developed and few empirical applications have been presented (Creech et al., 2017; Fitzpatrick & Keller, 2015; Landguth, Bearlin, Day, & Dunham, 2017). This may be partially due to a lack of datasets with appropriate sampling designs at both the genomic and the spatial scales that are needed to test for selection processes along environmental gradients (Ahrens et al., 2018; Hoban et al., 2016; Rellstab et al., 2015). However, characterizing variation in selection and local adaptation across environmental gradients is a necessary next step in evolutionary and landscape genomics, which will inform conservation management of biodiversity (Hoffmann et al., 2015; Hoffmann & Sgro, 2011). For example, selection on candidate genes may be monitored spatially and temporally as climate change proceeds,

revealing 'hot and cold spots' of local adaptation (Hansen, Olivieri, Waller, Nielsen, & Ge, 2012).

Insect distributions are currently experiencing pronounced shifts in response to climate change (Lancaster, 2016; Sánchez-Guillén, Muñoz, Rodríguez-Tapia, Arroyo, & Córdoba-Aguilar, 2013), and insects also exhibit altered physiological (Advani et al., 2016; Lancaster et al., 2016; Lancaster, Dudaniec, Hansson, & Svensson, 2015) and phenological trait changes (Arribas, Abellán, Velasco, Millán, & Sánchez-Fernández, 2017; Sánchez-Guillén et al., 2013) associated with range shifts. Aquatic and semi-aquatic insects may be among the first organisms to suffer from ongoing climate change due to exposure to anthropogenic stressors (e.g. habitat degradation), and dependence on climate-mediated water temperatures (Woodward, Perkins, & Brown, 2010). This makes freshwater insects appropriate models to investigate microevolutionary responses to climate change (Bybee et al., 2016). Here, we use a landscape genomics approach to investigate genomic signatures of local adaptation along environmental gradients in the blue-tailed damselfly, *Ischnura elegans* (Odonata; Vander Linden 1820). We sample the distribution of *I. elegans* in southern Sweden - a gradient where mean annual temperature varies substantially and rapid range expansions in ectotherms are occurring (Jaenson, Jaenson, Eisen, Petersson, & Lindgren, 2012). Damselfly distributions are shifting globally (Swaegers et al., 2015; Takahashi et al., 2016; Watts, Keat, & Thompson, 2010), and for *I. elegans* in the United Kingdom, the northern range limit was extended by 143 km between two 10-year survey periods of 1960-70 and 1985-95 (Hickling, Roy, Hill, & Thomas, 2005). In Sweden, our recent discovery of populations beyond the known range limit, with shifts in thermal niche breadth (Lancaster et al., 2015, 2016) that interact with social feedback mechanisms (Lancaster, Dudaniec, Hansson, & Svensson,

2017), supports a recent and ongoing rapid range expansion in *I. elegans*. In particular, strong selection on cold tolerance was documented in range margin populations based on phenotypic and gene expression responses to thermal challenges, indicating an important role of the thermal stress response on adaptive processes during range expansion (Lancaster et al., 2015, 2016).

Using genome-wide data from Restriction site-Associated DNA sequencing (RADseq) and gene annotation, we identify candidate single nucleotide polymorphisms (SNPs) under selection in relation to environmental gradients from southern 'core' populations of *I. elegans* (Le Rouzic, Hansen, Gosden, & Svensson, 2015; Svensson & Abbott, 2005; Svensson, Abbott, & Härdling, 2005) up to populations at the expanding northern range margin (‘edge’ populations). Covering a five degree latitudinal gradient with high resolution genomic and spatial sampling, we test for: 1) signatures of selection on SNP loci (i.e. via Fst Outlier analysis, EAA and annotation) that associate with temperature, habitat and climate-related variables; and 2) significant allele frequency changes in candidate SNPs that track environmental gradients towards the range limit, and evidence for environmental thresholds of selection. We corroborate our findings with prior observations of latitudinal shifts in thermal tolerance phenotypes and gene expression profiles (Lancaster et al., 2015, 2016). We apply a novel, three-tiered analytical approach to identify environmental variables driving local selection on alleles that are putatively adaptive or neutral along a range expansion gradient, revealing highly resolved spatial variation in local adaptation. Our results reveal patterns of spatially explicit adaptive genetic variation during a climate change-induced range shift, which has significant implications for understanding the future distribution of this species and the structure of biodiversity more generally.

Materials and Methods

Approach

We implement a three-tiered analytical approach to identify genes under putative selection in response to environmental gradients along a range expansion zone in *I. elegans* (Figure 1). Firstly, (1) candidate SNPs being under putative selection are identified using two Fst outlier approaches (Foll & Gaggiotti, 2008; Whitlock & Lotterhos, 2015) and one Environmental Association Analysis (EAA) approach (Frichot, Schoville, Bouchard, & Francois, 2013). Secondly, (2) Generalized Dissimilarity Modelling (GDM) is applied to these identified candidate SNPs, to determine relationships of SNP allelic turnover magnitude in relation to environmental gradients and geographic distance (Fitzpatrick & Keller, 2015). Finally, (3) signatures of local adaptation are identified via SNP mapping to an annotated *I. elegans* transcriptome (Chauhan et al., 2014, 2016), and interpretations about adaptive variation are then based on gene function, experimental gene expression data (e.g. Lancaster et al., 2016), SNP x environment associations and the pattern of allelic turnover observed (Figure 1). Our analysis provides fine-scale characterization of SNP-specific genetic gradients of genome-wide selection signatures.

Sampling and study area

Ischnura elegans is common across Europe and Asia, with its northern range extending to the southern coastal areas of Scandinavia and the northern United Kingdom (Dijkstra & Lewington, 2006). Our study area spans latitudinal gradient of five degrees of latitude in Sweden (latitudinal range: 55.64˚ to 60.57˚, Table S1), extending 583 km from the southern populations to the northern range edge (see Lancaster et al., 2015, 2016, 2017). Between the summer months of June and August 2013, we sampled 25 sites throughout the Swedish

distribution of *I. elegans* following a paired gradient sampling design, encapsulating both coastal and inland sites and the northern range edge (Figure 2). Adult *I. elegans* were caught near to reed beds and vegetation using sweep nets within 10m of water bodies including ponds, lakes and coastal inlets. We implemented a paired-gradient sampling design to the best of our ability (i.e. approximately two samples per latitudinal sampling interval), as this approach has improved power to detect local adaptation at weakly selected loci using EAA in range expansion models, as opposed to random or transect designs (Lotterhos & Whitlock, 2015). We performed all procedures in accordance with the ethical guidelines of Lund University in Sweden, and obtained sampling permissions from local authorities and landholders.

RAD sequencing, bioinformatics and SNP characterization

We extracted DNA from 432 *I. elegans* from 25 sites (10-20 individuals per site, mean 17.04 ± 0.72; Table S1) using the head, thorax and legs from each individual using a DNeasy Blood and Tissue extraction kit (Qiagen). We quantified extracted genomic DNA using a Qubit 2.0 Fluorometer (Life Technologies), which was then processed into paired-end RAD libraries according to the protocol implemented in Etter *et al*. (2011), and as described in the supplementary material 1.0. Each RAD library was sequenced on a separate lane of an Illumina HiSeq 2000 or 2500 at the Beijing Genomics Institute, Shenzhen, China yielding 20-30GB of data per library. Adapter sequences and low-quality bases below a Phred score of 20 were trimmed from raw reads according to standard quality control protocols (to 100bp read length).

Raw sequences from each RAD library were quality checked visually using *FASTQC* (Andrews, 2010) and each library was processed using pipelines within *Stacks v.1.40*

(Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013; Catchen, Amores, Hohenlohe, Cresko, & Postlethwait, 2011). Methods used in *Stacks* are described in more detail in the supplementary material 1.1. Samples were processed in *Stacks* using the *process_radtags* with a mean of 105 million reads (± 13.36 M) per library, followed by the *clone_filter* program to remove PCR duplicates, resulting in a mean of 36 million reads (± 6.4 M) per library (Table S1). The final sample size of individuals retained for analyses was 426 across the 25 populations, as six samples were excluded due to low coverage. De-duplicated reads were aligned to an *Ischnura elegans* draft genome assembly (version 12-2015 by P. Chauhan et al.; Supplementary Information) using *Bowtie2* v.2.2.5 (Langmead & Salzberg, 2012). Aligned reads from *Bowtie2* were analysed in the *ref_map* program in *Stacks* to build the initial consensus catalogue of SNPs, resulting in 3,452,911 loci. SNPs were further filtered using the *rxstacks* corrections model, which removes excess haplotypes and confounded loci (Catchen et al., 2013).

The final set of SNP markers was determined within the *populations* program in *Stacks*, which was run twice: first, including all SNPs on each RAD-tag and secondly, including only the first SNP on each RAD-tag to create a dataset without closely linked loci (using the *write_single_snp* option in *Stacks*). We specified an initial minimum depth of coverage of 5x for each SNP-containing RAD locus with a minor allele frequency (MAF) of 0.05. Additionally, a locus was only included if it occurred in 22/25 populations and in at least 80% of individuals within each population to ensure wide representation of data for each SNP across all samples and sampling locations (recommended by Paris, Stevens, & Catchen, 2017). After filtering loci using the *Stacks populations* program, 13,612 SNPs (including linked SNPs, used for Fst outlier, EAA and GDM analyses) and 3809 SNPs (excluding

closely linked SNPs, used for genetic structure analysis) were retained for analysis. Depth of coverage per SNP varied between 8-23x (mean 15.3x; Figure S1).

Environmental data

Variables used in environmental association analysis (EAA) and general dissimilarity modelling (GDM) were chosen from those previously identified in species distribution modelling (SDM) for *I. elegans* within the same study area (Lancaster et al., 2015). Lancaster et al. (2015) identified 12 variables that predicted the distribution of *I. elegans* that all had a pairwise Pearson correlation coefficient (r) less than 0.8 in a prior habitat suitability model. Of these 12 variables, we chose five (described in Table 1) that varied widely over the sampling gradient (Figure S2): 1) Mean Annual Temperature (BIO1, "Annual Temp"; 62.1% contribution to SDM), 2) the Maximum Temperature of the Warmest Month (BIO5, "Max Temp"; 0.1% contribution to SDM), 3) Mean Annual Precipitation (BIO12, "Annual Rain" 0.1% contribution to SDM), and 4) Percentage Tree Cover ("Tree Cover", 0.4% contribution to SDM). We also included a fifth variable that was not examined by Lancaster et al. (2015), 5) Mean Summer Wind Speed ("Wind Speed", averaged for June-August; metres per second, measured at 80 m height) (Table S1, Figure S2). These chosen variables were selected due to explicit biological predictions regarding their effects on adult fitness during the short adult reproductive and dispersal period, which is a critical period for selection processes in odonates (discussed in Wellenreuther, Larson, & Svensson, 2012; Supplementary Information). Although the larval period is longer than the adult period in many insects including *I. elegans*, it is proposed that genetic variation for fitness is primarily expressed in the adult phase of insects (e.g. in *Drosophila*: Chippindale, Gibson, & Rice, 2001). Therefore, we selected climate and landscape variables that are most likely to be relevant for

234 evolutionary processes during the adult period (e.g. Max Temp, Tree Cover, Wind Speed),
235 but also those that may act as selection pressures over longer developmental periods (e.g.
236 Annual Temp, Annual Rain). Further justification of the environmental variables is given in
237 the Supplementary Information (1.3).
238
239 The Pearson correlation coefficients (*r*) between the five environmental variables
240 taken from each site were less than 0.4 except for Annual Temp and Wind Speed ($r = 0.75$),
241 and Annual Temp and Max Temp ($r = -0.48$, Table S3). Therefore, our ability to separate
242 Annual Temp from Wind Speed and Max Temp was limited (Table S3). We calculated
243 geographic distance (km) between sites using the R package *ecodist* (Goslee & Urban, 2007).
244 All environmental variables were extracted at a 1km cell resolution from BIOCLIM variables
245 within the WorldClim Version 1.4 database (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005)
246 except wind speed data that were extracted from WorldClim Version 2.0 (Fick & Hijmans,
247 2017), and percentage tree cover data that were obtained from the Global Land Cover Facility
248 (Defries, Hansen, Townshend, Janetos, & Loveland, 2000).
249
250 *Outlier SNP detection and genetic structure*
251 Detection of outlier SNPs (i.e. loci putatively under divergent selection) was performed on the
252 complete dataset (13 612 SNPs) using two contrasting Fst-based approaches implemented in
253 BAYESCAN 2.1 (Foll & Gaggiotti, 2008) and OutFLANK (Whitlock & Lotterhos, 2015). Two
254 approaches were used to maximise the identification of potential loci under selection for
255 exclusion from genetic structure analysis, and to identify common significant SNPs across
256 methods. The false discovery rate (FDR) was set at 0.05 and number of populations (*K*) was
257 set to 25 in both programs. The Bayesian likelihood approach implemented in BAYESCAN

compares population allele frequencies with a common migrant gene pool, which allows for different migration rates and acts to account for effects of neutral genetic structure, reducing the proportion of false positives (Narum & Hess, 2011). OutFLANK (Whitlock & Lotterhos, 2015) identifies outliers by first inferring the distribution of Fst for loci that are unlikely to be under selection, and only attempts to identify loci under positive selection. This method performs well under diverse demographic history scenarios, including range expansion (Whitlock & Lotterhos, 2015). Further details are in the Supplementary Material 1.3.

To minimize the inclusion of putative loci under selection and linked loci from analyses of neutral genetic structure, Fst outlier loci identified using BAYESCAN and OutFLANK analyses were removed from the 'unlinked' SNP dataset (i.e. single SNP per RAD-tag), resulting in 3554 SNPs. Genetic structure was estimated with the program ADMIXTURE (Alexander, Novembre, & Lange, 2009), which uses a cross-validation procedure to determine genetic structure in large autosomal SNP data sets. ADMIXTURE was run for 1-25 potential ancestral populations (*K*) with a 5-fold cross validation (CV) error and *K* was chosen where the cross-validation error was minimized. The probability of individual assignment to each genetic cluster (*Q*) was graphically displayed and plotted in R (Figures 1, S4 & S5).

Environmental association analysis

Environmental association analysis (EAA) was performed using a Latent Factor Mixed Modeling (*LFMM*), implemented with the R package *LEA* (Frichot & François, 2015) using all 13,612 SNPs. *LFMM* uses a stochastic Monte Carlo Markov Chain algorithm and tests for associations between environmental or ecological variables and allele frequencies while estimating unobserved latent factors that model confounding effects of genetic structure, which may be due to shared demographic history or background genetic variation (Frichot et

al., 2013). *LFMM* was run with the number of latent factors set to the number of genetic clusters (*K)* obtained via ADMIXTURE (see below; *K* was equal to four) with five repetitions, and 10,000 iterations with a 5,000 burn-in. The z-scores over the five runs were combined and p-values adjusted as recommended by Frichot and François (2015). To include SNPs that were highly significantly correlated with the environmental variables, we applied a conservative Benjamini-Hochberg p-value cut-off < $\log10^{-6}$. We ran *LFMM* to find SNP by environment associations for all five environmental variables (e.g. Annual Temp, Max Temp, Annual Rain, Tree Cover, Wind Speed). Shared and unique SNP x environment associations were quantified across the five environmental variables and their overlap with Fst outlier results examined (Table 1). The genomic inflation factor (GIF) described by Devlin and Roeder (1999) was calculated for each environmental variable from the z-scores derived from *LFMM* and was assessed for its closeness to the recommended value of 1.0 (Frichot & François, 2015). The GIF across four of the variables ranged from 1.04 to 1.48, but Annual Temp had a GIF = 2.34. This indicates that FDRs are likely to be higher for Annual Temp than the other variables analysed due to poor statistical calibration. Given the high GIF, the high correlation of Annual Temp with both Max Temp and Wind Speed, and the relevance of Max Temp to the adult flying period, we chose to exclude Annual Temp from further analyses.

General Dissimilarity Modelling of candidate SNPs

We examine spatially explicit selection processes for each SNP found to be under putative selection using a modified Generalized Dissimilarity Modelling (GDM) approach described in (Fitzpatrick & Keller, 2015), implemented using the R package *GDM* (Ferrier, Manion, Elith, & Richardson, 2007; Manion et al., 2017). The approach is adapted from the use of

GDMs in biodiversity modelling to examine non-linear turnover in community-level composition (Ferrier et al., 2007), but uses large numbers of loci (instead of species) to find both linear or nonlinear responses of loci to environmental gradients (Fitzpatrick & Keller, 2015). The approach takes the pairwise Fst of SNPs across sample sites and models the rate and magnitude of 'allelic turnover' (i.e. change in allele frequency represented as a genetic distance measure) in relation to the distribution of an environmental variable along a spatial sampling gradient, using a site-by-SNP matrix (Fitzpatrick & Keller, 2015). This is achieved by using permutation on distance matrices to perform model and variable significance testing and to estimate variable importance. By identifying functions of allelic turnover according to environmental gradients, the approach offers a means of scaling from population-level genomic variation to predictions of landscape scale adaptive variation, which are both subject to ongoing environmental change (Fitzpatrick & Keller, 2015).

Using the GDM approach, we identify thresholds on the landscape where signatures of local adaptation in *I. elegans* increase or decrease in relation to the five environmental gradients we examined using EAA. We conducted GDM for a candidate set of SNPs identified as being putatively under selection using either BAYESCAN, OUTFLANK or *LFMM* (total SNPs = 1758). SNPs identified in BAYESCAN with significantly negative Fst values (i.e. under potentially balancing selection) were excluded from the candidate set as these loci are likely to have a very high FDR (Whitlock & Lotterhos, 2015). The complete set of Fst outliers identified from both BAYESCAN and OUTFLANK were included in the GDM because each program implements a uniquely valid statistical approach to detect selection, and we observed little lack of overlap in significant SNPs between the approaches. We modified the approach of Fitzpatrick and Keller (2015) by taking a ‘single SNP’ approach with each

putatively selected SNP modelled independently, regardless of annotation, as opposed to selecting specific, annotated SNPs or grouping related SNPs for GDM modelling.

Additionally, as in Fitzpatrick and Keller (2015), we integrate a random sample of 200 SNPs out of the 13,612 available SNPs, which act as a 'reference group' in the GDM to test whether allelic turnover at a given candidate SNP differs from that expected in a random sample of the genetic data. Further, geographic distance (Euclidean) was incorporated as a sixth variable in the GDM to test if allelic turnover across environmental gradients was better explained by distance, which effectively acts as a second screening (i.e. after Fst outlier and EAA tests) for loci that may respond predominantly to neutral genetic processes (i.e. those influenced by genetic structure, including isolation by distance), and may therefore have been falsely identified in outlier tests, or have lower confidence to be identified as candidate SNPs involved in adaptation. Although geographic distance alone does not incorporate other demographic effects associated with range expansion that can influence selection detection (e.g. founder effects, allele surfing), we attempt to control for false positives by, 1) comparing outcomes with relationships with geographic distance and, 2) by comparing allelic turnover responses of the random 'reference' SNP group with that of each locus to test if its response is more or just as likely in a random sample of genetic variation.

Genetic distance matrices between the 25 sample sites were calculated for each of the 1758 candidate SNPs, and for the reference group based on Nei's pairwise Fst (Nei, 1987) using the R package *hierfstat* (Goudet, 2005), and were rescaled between 0 and 1 within the GDM analysis. To assess the role of each SNP in selection processes in relation to each environmental variable examined, we ranked the allelic turnover functions of each SNP and

for each environmental variable using two different methods: (1) within each SNP: ranking was based on the magnitude of allelic turnover at a given SNP (i.e. change in Fst along a specific environmental gradient) relative to its turnover magnitude for other environmental variables in the model; (2) across all SNPs: ranking was based on the percentage deviance explained by each SNP relative to all SNPs in the GDM model (using the permutation procedure of the R function *gdm.varImp*), which gives an indication of selection strength for each SNP relative to the whole dataset. For (1), the top 250 SNPs with the highest magnitude of allelic turnover are plotted for each environmental variable (Figure 3). The second ranking (2) was used as a secondary assessment of the overall selection signature of the SNP within the entire GDM model. GDM results for all 1758 SNP responses and tests are in the supplementary material.

Gene Annotation

To identify functional genes, RAD tags containing one or more of the 1758 candidate SNPs were mapped against the annotated transcriptome for *I. elegans* (Chauhan et al., 2014, 2016) using BLASTN with an e-value cut-off of $1x10^{-5}$. All BLASTN 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 BLASTN and InterProScan annotated transcripts (Ashburner et al., 2000). Gene Ontology (GO) annotations and GO Slim reductions were applied to categorize transcripts into major GO categories, 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. Gene functions were identified from those previously annotated in Chauhan et al. (2014, 2016), those with expression levels associated with thermal challenge treatments in *I. elegans* performed by Lancaster et al. (2016), or were identified directly from the NCBI database (Table 2). Gene functions were only considered for those with an annotation match of ≥70% (Supplementary Material). Transcripts with SNP annotations were mapped to an assembled genome (Supplementary Material) using BLAST and the positions where transcripts mapped were recorded (i.e. scaffold ID and base pair position on RAD tag).

Mapping adaptive genetic variation over the temperature gradient

To examine how adaptive variation changes in *I. elegans* along its current Swedish distribution, we mapped allelic turnover functions for selected candidate SNPs that, 1) were annotated to genes associated with thermal tolerance or other phenotypic traits previously identified (e.g. Chauhan et al., 2014, 2016; Lancaster et al., 2016) and 2) had a higher explanatory power in the GDM than the reference 'random' SNP group. In addition to the above, we focused on SNPs that 3) had the highest allelic turnover in relation to Max Temp in the GDM, or 4) showed a large change in Fst along the sampled gradient (Figure 1). This resulted in a list of 23 SNPs, and allele frequencies and turnover functions were mapped for four of these SNPs to reveal spatially explicit selection gradients. All maps were produced in R using the *GDM*, *raster* and *ggplot* packages (Ferrier et al., 2007; Hijmans & van Etten, 2012; Wickham, 2009).

Results

Fst outlier detection and genetic structure

BAYESCAN identified 688 SNPs (5% of 13 612 SNPs) under putative selection across the 25 sites. There was a distinct split among the outliers with divergent selection being represented in 57% (n =391) of SNPs and potentially balancing selection being represented in 43% (n =297) of SNPs. Using OutFLANK, 188 outliers (1.4%) were detected, which were all under putative positive selection. Nine SNPs were commonly identified in BAYESCAN (diversifying only) and OutFLANK. All SNPs identified as an Fst outlier in either BAYESCAN or OutFLANK were removed for genetic structure analysis. Notably, removing even the least conservatively estimated loci under putative selection can minimize false estimates of genetic structure, and therefore we attempt to address this risk of false positives by removing all candidates from both programs. ADMIXTURE analysis showed a cross validation (CV) error that was minimized at four genetic clusters (K = 4, using 3554 SNPs; Figure 2, Figure S3, Table S2). A high proportion (39%) of individuals showed ancestry to more than one cluster (Figure S3), though probabilities of ancestry were overall higher to a given cluster for populations in the southern region (Figure 2). There was greater variability in assignment probabilities towards the range limit, but a larger number of distinct genetic clusters represented (i.e. 3-4, Figures 1 & S6, Table S2) while all four sites in the southern region belonged to a single cluster (Figure 2).

Environmental Association Analysis

A total of 2327 significant SNP associations were identified across the five environmental variables analysed using *LFMM* (with a $<\log10^{-6}$ p-value significance cut-off), with a similar number of SNP associations for each variable (mean = 465 SNPs; range = 374-566; see Tables 1, Figure S4). However, these associations were attributed to 451 unique SNPs, and none of the SNPs were significantly associated across all five environmental variables. Very

426 few SNPs identified as Fst outliers were also found in the EAA associations using *LFMM*
427 with 22 SNPs (5%) overlapping with Bayescan outliers, and 41 SNPs (9%) overlapping with
428 OutFLANK outliers, yet none across all three approaches (Table 1). Of the EAA
429 associations, between 18.3 and 58.4% (mean = 35%) of the associations were shared across
430 more than one environmental variable (Table 1). Annual Temp shared 30.4% and 50.7% of its
431 associations with Max Temp and Wind Speed, respectively (Table 1).
432
433 *Patterns of selection signatures along environmental gradients*
434 Including all significant associations across all tests, a total of 1758 unique SNPs were
435 identified as being under putative selection (Table 1) and all were analysed using GDM. A
436 large proportion of putatively adaptive SNPs (60%) were identified via at least one Fst outlier
437 test (i.e. BAYESCAN, OutFLANK) or were associated with a single environmental variable
438 using LFMM (n = 5 environmental variables tested). SNPs identified with two (n = 381;
439 22%), three (n = 236; 13%), four (n = 73; 4%) or five (n = 19; 1%) tests were less common.
440 We present GDM results for the top 250 SNPs with the highest magnitude of allelic
441 turnover in relation to each environmental variable (Figure 3). A wide Fst distribution was
442 observed for these top ranking SNPs, which was similar to the shapes of the Fst distribution
443 for all 1758 candidate SNPs (Figure S5). The allelic turnover for each of the top 250 SNPs
444 according to each environmental variable (Figure 3) indicates differing gradients and
445 strengths of selection across loci. Despite being associated with an environmental variable
446 using *LFMM*, the SNPs with the highest allelic turnovers were associated with geographic
447 distance (and noted as possible false positives), which was followed by (in decreasing order
448 of allelic turnover magnitude) Max Temp, Annual Rain, Wind Speed and Tree Cover (Figure
449 3). The shapes of the allelic turnovers across SNPs ranged from distinct 'plateaus' at a given

position on the gradient, to positive and almost exponential allelic turnover responses at particular gradient positions. For example, the top 50 SNPs for geographic distance appeared to mostly reach fixation at the largest distances (Figure 3a), while most SNPs associated with Wind Speed ceased allelic turnover beyond a wind speed threshold of 3.0 m/s (Figure 3d). Max Temp (Figure 3b) and Annual Temp (Figure 3c) drove the strongest and most variable allelic turnover magnitudes of the environmental variables, with distinct turnover thresholds identifiable for each associated SNP.

Allelic turnover responses and annotation

For 206 of 1758 SNPs (11.7%), there was no significant allelic turnover response associated with geographic distance or any of the environmental gradients analysed using GDM, and these SNPs were not interpreted further. Selective neutrality in relation to environmental gradients was assessed via SNP allelic turnover response to geographic distance versus environmental variables within our GDM (Fitzpatrick & Keller, 2015). Geographic distance had the highest magnitude in allelic turnover response for 372 of the 1758 SNPs analysed (21%), relative to the other environmental variables. The reference ('random') SNP group explained 11.8% of the GDM deviance for the entire model, and SNPs that did not exceed 11.8% were also considered to be potential false positives.

Of the 1758 candidate SNPs (located on 640 different scaffolds), 1196 (68%) were annotated to the *I. elegans* transcriptome, and of these, 50 SNPs (located on 13 scaffolds) were located on transcripts previously identified in gene expression analyses by Chauhan et al. (2014, 2016) and Lancaster et al. (2016) (see Supplementary Material). After additional filtering of SNPs that had greater explanatory power in the GDM than the reference SNP group, 21 of 50 previously annotated SNPs (located on 7 scaffolds) were retained, with some

474 occurring on the same RAD tag (i.e. tightly linked SNPs), or having more than one matching
475 transcript, isoform or annotation (Table 1). An additional two SNPs (on 2 scaffolds) with
476 annotations of relevance to environmental adaptation (though not previously reported) were
477 also retained. These two SNPs were in the top 10 SNPs with respect to the percentage of the
478 GDM explained, allelic turnover magnitude with respect to Max Temp, and highest change in
479 Fst along the sampled gradient.

480 We focus on these 23 annotated SNPs from here forward as they exhibited the most
481 significant selection signatures in tandem with annotations that can be linked to processes
482 during environmental adaptation. The 23 SNPs spanned five key functional groups relevant
483 for thermal stress (i.e. 11 SNPs for HSP40 and one for HSP70, represented across six RAD
484 tags), visual processes (5 SNPs spanning rhodopsin, pteropsin, and long wavelength-sensitive
485 opsin across three RAD tags) epigenetic modification (4 SNPs for histone-lysine n-methyl
486 transferase across three RAD tags), ion transport (1 SNP for vacuolar H+ proton pump) and
487 varied cellular processes (1 SNP with multiple annotations) (Table 2, supplementary data).
488 One isoform was found for each gene function except for one epigenetic modification gene
489 that contained two isoforms (Table 2). Seven of the annotated SNPs were identified as
490 significant outliers using BAYESCAN and one SNP using OUTFLANK (Table 2). All
491 annotations are provided in supplementary material.
492
493 *Environmental associations and allelic turnover of annotated SNPs*
494 Of the 23 focal SNPs, five showed the greatest allelic turnover magnitude with respect to
495 geographic distance, though one SNP was equal or within 0.02 magnitude to Annual Rain
496 (SNP 39648_74; Table 1). These SNPs are considered to be less likely to be under selection
497 by the environmental variables analysed *per se*, despite showing significant changes in allele

frequencies according to geographic distance. For the 23 SNPs, the magnitude of allelic turnover was highest for those that associated with Max Temp (mean = 0.42±0.08; 9 SNPs), followed by Annual Rain (mean = 0.30±0.05; 8 SNPs), Wind Speed (mean = 0.25±0.07; 4 SNPs), and Tree Cover (mean = 0.185±0.05, 2 SNPs) (Table 1, Figure 3). Allelic turnovers of the 23 SNPs (Table 1) in response to each environmental gradient was highly variable, both within and across gene functions (Figure 3). Generally, the locations at which rates of allelic turnover changed the most (i.e. where one allele was selected for most strongly) were observed between sites with the greatest geographic distance apart (Figure 3a), at upper latitudes where summer temperature was high (Max Temp, Figure 3b), at lower latitudes where rainfall was lower (Annual Rain, Figure 3c) and wind speed was higher (Figure 3d). Though weak, locations with higher tree cover also showed some allelic turnover (Figure 3e). Max Temp and Annual Rain both increase with latitude (Figure S2) and their associated SNPs showed polarised patterns of selection, with some showing strong allelic turnover at lower values before stabilizing, and others becoming strong only at high gradient values (Figure 3).

SNP-specific signatures of local adaptation

We examined spatial genetic gradients over the study area by quantifying allele frequency changes in four selected SNPs that were selected based on: 1) the SNP's functional annotation, 2) its statistical association with Max Temp (both magnitude of allelic turnover and ranking of turnover), 3) its change in Fst along the gradient, 4) the percentage of the GDM model the SNP explained. We firstly examined SNP 37543_9, which was annotated to *vacuolar H+ ATPase*, which is involved in proton pump activity to regulate pH in eukaryotic cellular compartments that affect important cellular processes (Nishi & Forgac, 2002). This SNP had the highest magnitude of allelic turnover in relation to Max Temp (0.63), a high

522 change in Fst along the gradient (ΔFst = 0.50), the highest ranking in the GDM model for
523 Max Temp (10), and was also identified as an Fst outlier using OutFLANK (Table 2, Figure
524 4). Secondly, we examined SNP 73426_72, which was annotated to a *long wavelength-*
525 *sensitive opsin 3b*, involved in visual processes. This SNP had the highest magnitude of
526 allelic turnover in relation to Max Temp (0.24), a high change in Fst along the gradient (ΔFst
527 = 0.21), and was ranked highly in the GDM model for Max Temp (266) (Table 2, Figure 5).
528 The allelic turnover functions for the above two SNPs are shown in relation to Max Temp
529 (Figures 4-5).
530
531 Thirdly, we examined SNP 53905_36, which was annotated to *Heat Shock Protein 70*
532 (HSP70; Table 2, Figure S6), a gene that is involved in the thermal stress response (Lancaster
533 et al., 2016; Sørensen, Kristensen, & Loeschcke, 2003). This SNP had the highest magnitude
534 of allelic turnover in relation to geographic distance (0.33), but had a high change in Fst along
535 the gradient (ΔFst = 0.40), and was identified as an Fst outlier in BAYESCAN. Finally, we
536 examined SNP 35404_9, which had the highest magnitude of allelic turnover in relation to
537 Max Temp (0.83), a high change in Fst along the gradient (ΔFst = 0.38), and was identified as
538 an Fst outlier in OutFLANK (Table 2, Figure S7). This SNP was annotated to 10 transcripts
539 that annotated to various proteins and enzymes (see supplementary data), including pellino
540 proteins, which are involved in the immune response via the Toll-like receptor pathway
541 (Schauvliege, Janssens, & Beyaert, 2007), and PACS2 (phosphofurin acidic cluster sorting
542 protein) which is involved in cell apoptosis (Simmen et al., 2005). The above four SNPs
543 showed spatial patterns of allelic turnover along the core to range limit gradient that varied in
544 magnitude and linearity, indicating differential selection on particular alleles along the *I.*
545 *elegans* expansion axis in relation to latitude and Max Temp (Figures 4-5, S8-S9).

546

547 **Discussion**

548 We characterise genetic signatures of local adaptation to environment along a climate-
549 mediated range expansion in a species exhibiting rapid response to shifting temperature
550 regimes (Hickling et al., 2005; Watts et al., 2010; Jaeschke, Bittner, Reineking, &
551 Beierkuhnlein, 2013; Lancaster et al., 2015, 2016; Swaegers et al., 2013, 2015). Among four
552 environmental variables tested, the strongest driver of allelic turnover along the *I. elegans*
553 expansion gradient was maximum summer temperature (Max Temp), followed by mean
554 annual precipitation (Annual Rain), wind speed, and to a much lesser extent, % tree cover
555 (Table 1, Figure 3). The greatest allele frequency changes in *I. elegans* were in localities
556 spanning low to mid latitudes (i.e. from Scania to further north), where Max Temp shifts most
557 dramatically (~1.2°C; Figures 4-5, S1, Table S2), rainfall is lower and more variable, and
558 wind speeds are higher than in the northern range edge (Figure S2). Selected annotated SNPs
559 exhibited allele-specific patterns of selection along the core to edge sampling gradient
560 (Figures 4-5, S8-9), with wide variation in the magnitudes of allelic turnover across SNPs
561 (Table 2). SNP annotations indicated that genes involved in the thermal stress response, visual
562 processes, epigenetic modification and ion regulation may play significant roles in adaptation
563 during this climate-mediated range expansion in *I. elegans*. Our multi-tiered approach (Figure
564 1) validates a 'bottom up' approach for detecting signatures of local adaptation from reduced
565 representation genomic data, in which a group of SNP candidates is first identified, followed
566 by SNP-specific modelling of genetic gradients, supported by gene annotation and prior
567 experimental knowledge of gene functional response (e.g. Chauhan et al., 2014, 2016;
568 Lancaster et al., 2016).

569

Detection of putative SNPs under selection

Fst outlier and EAA analyses are increasingly popular methods for identifying SNPs under putative selection (Hoban et al., 2016; Rellstab et al., 2015). One notable aspect of our Fst outlier and EAA results is their lack of overlap in terms of the number and identity of SNPs (Table 1). Not only did the SNPs identified by our two Fst outlier approaches overlap by just 1.5%, but Fst outliers overlapped with just 0.6-4.0% of SNPs identified using EAA (Table 1). This does not necessarily indicate a lack of power in the analysis, and is consistent with findings that EAA performs better than Fst outlier tests in detecting weak or polygenic selection signatures (Frichot et al. 2015; Villemereuil et al. 2014). The minimal overlap and difference in numbers of SNPs identified between Fst outlier approaches identified may indicate different sensitivities of each approach to the effects of genetic drift and structure. Notably, studies comparing OutFlank and Bayescan have found little overlap between the approaches (e.g. Bernatchez, Laporte, Perrier, Sirois & Bernatchez 2016; Chen, Farrell, Matala & Narum 2018; Michelleti, Matala, Matala & Narum 2018). The significant SNP associations using EAA were unique to each environmental variable in 41-79% of cases (Table 1). Concordantly, Fst distributions were negatively skewed and variable across all 1758 candidate SNPs (Figure S5a), which was mirrored when examining SNPs according to the environmental variable they were associated with (Figure S5c-f). The dominance of low Fst values indicates that many SNPs show weak selection signatures along the sampling gradient.

Notably, the Fst changes observed in the 23 annotated and most highly supported SNPs from the GDM were not biased towards higher Fst values (Fst range = 0.09-0.5; Figure 2, Table 2). Overall, the results indicate that an increased change in Fst along a sampling gradient of a SNP does not correlate with a greater likelihood of identifying that SNP as being

under selection using EAA. This lack of correlation has similarly been observed in a recent meta-analysis of studies using Fst outlier tests and EAA (Ahrens et al. in review). This observation indicates that adaptation to environmental conditions is polygenic and involves many interacting loci of both small and large effect (e.g. Lee & Mitchell-Olds, 2012).

Accounting for neutral genetic structure

Detecting genetic selection signatures is riddled with the issue of separating true adaptive genetic responses from neutral genetic structure (Hoban et al., 2016), which is particularly relevant when neutral structure mirrors sampled environmental gradients (Lotterhos & Whitlock, 2015). Range expansion processes can result in patterns of selection on loci that mirror neutral genetic structure, for example, via allele surfing mechanisms, whereby rare alleles become more frequent at range expansion fronts according to the process of genetic drift rather than selection. Allele surfing can therefore increase population genetic differentiation and confound signatures of local adaptation (Klopfstein, Currat, & Excoffier, 2006) but might also affect adaptation when either beneficial or deleterious alleles are 'surfed' on the wave of expansion (Gralka et al., 2016; Travis et al., 2007). Such processes make teasing apart adaptive and neutral processes in range expanding species a challenge.

Genetic admixture was greatest within sites at the low to mid latitudes and declined towards the range limit in *I. elegans,* where sites were comprised of individuals assigned to multiple or unique clusters (Figure 2). Given this tracking of genetic structure with latitude, it was particularly important to account for false positive SNPs in our data. At each step of our analysis we applied approaches to avoid false positives. Firstly, we selected only putative SNPs under selection using Fst outlier tests (diversifying only) and EAA, and excluded SNPs associating with geographic distance in our EAA. In addition, we implemented two additional

approaches to avoid the inclusion of false positives using GDM by, 1) including a randomly selected 'reference' SNP group to compare with each SNP, and 2) including geographic distance as a predictor in the GDM to identify selection signatures correlating with geography. Finally, SNP annotations to gene functions involved in thermal stress response and other ecologically relevant genes indicated climate-mediated local selection on some candidate SNPs along the range expansion gradient (Table 2, Figures 4-5). Despite expectations that gene flow will have a constraining effect on adaptive divergence (discussed in Smadja & Butlin, 2011), the relationship between gene flow and local adaptation is increasingly found to be positive (Jacob et al., 2017; Moody et al., 2015), including at species' range edges (Halbritter, Billeter, Edwards, & Alexander, 2015). Further analysis of how neutral genetic connectivity and landscape features are related to the pattern of adaptive genetic variation in *I. elegans* is needed to address this.

Broad allelic frequency changes across the range expansion

The contrasting steepness of the environmental gradients we sampled (Figure S2) appeared to correspond with the magnitudes of allelic turnover observed across SNPs using GDM (Figure 3), which is in contrast to the lack of an environmental 'steepness' effect on selection detection across studies using EAA (reviewed in Ahrens et al. 2018). For example, percentage tree cover was highly variable according to latitude (Figure S2) and attracted the lowest allelic turnovers (Figure 3). In contrast, Max Temp showed the steepest environmental gradient and correspondingly high allelic turnovers (Figure 3). Pronounced allele frequency changes in relation to Max Temp between low to mid latitudes, indicate a 'transition area' of local adaptation (Figure 1, S9) where the greatest shifts in environmental conditions are present. In this area, Max Temp increases by approximately 2°C, mean annual precipitation decreases by

170mm and wind speed decreases by 1.9 m/s within an approximate 3-degree shift in latitude (Table S2, Figure S2). At the range edge, sites are located further inland and conditions are less variable (e.g. only 0.57°C maximum difference in Max Temp between sites). A second area that exhibited high allelic turnover was at the northern range limit, where distinct changes in allele frequencies were evident that were often correlated with the warmer Max Temp at sites in this region (Figures 3-5, S8).

Our 'bottom up' approach of screening RAD-derived SNPs for environmental selection signatures is an alternative to when dense genomic resources are available (e.g. using GWAS: Berg & Coop, 2014) or pre-identified candidate genes are targeted (e.g. Fitzpatrick & Keller, 2015; Hoekstra, Hirschmann, Bundey, Insel, & Crossland, 2006; Sork et al., 2016), and is informed largely by the spatial heterogeneity of both environmental and adaptive variation within the dataset. One important caveat of the EAA approach is that some loci may only show a weak association with environmental variables when the locus is simultaneously advantageous across a diversity of environments (Frichot et al., 2013). Our GDM approach is complementary in this case, as it allows for relative allelic responses to be simultaneously characterised across predictor variables. Approaches that characterise gene interactions may further elucidate the polygenic basis of environmental adaptation (e.g. Herold et al., 2012; Lee & Mitchell-Olds, 2012).

Signatures of environmental selection on annotated genes

The response curves of the annotated candidate SNPs to the tested environmental variables (using GDM) indicate that allele frequencies are tracking environmental gradients along the *I. elegans* range expansion (e.g. Figures 4-5, S8-S9). A variety of gene functions were

666 represented with a diversity of environmental associations (Table 2). Our annotations of
667 candidate SNPs matched gene functions associated with thermal tolerance in a gene
668 expression study by Lancaster et al. (2016) along the *I. elegans* range expansion. Three major
669 gene functions were previously identified from gene expression experiments (Table 2) in both
670 Lancaster et al. (2016) (thermal stress and epigenetic modification) and Chauhan et al. (2014,
671 2016) (visual processing and thermal stress), while we found additional support for strong
672 selection on genes involved in ion transport (V-ATPase) and other cellular processes.
673
674 Eleven candidate SNPs annotated to the HSP40 gene. All of these SNPs were located
675 on the same genome scaffold and showed significant environmental associations with Max
676 Temp and other variables (Table 2). HSP40 was not differentially expressed in *I. elegans* in
677 Lancaster et al. (2016) in response to thermal tolerance treatments, which may be indicative
678 of the different mechanisms involved in gene expression. However, HSP70 that was included
679 among our candidate genes, showed greater upregulation in gene expression in response to
680 heat stress in the core compared with edge populations (Lancaster et al. 2016). Only a single
681 SNP was annotated to HSP70 (Table 2) and showed a large change in Fst and allelic turnover
682 along the sampled environmental gradient (Figure S6). HSP70 is a highly conserved, ATP-
683 dependent molecular chaperone that facilitates protein homeostasis under a variety of
684 conditions including thermal stress (Beere 2004; King & MacRae 2015). The allelic turnover
685 of the SNP annotated to HSP70 was strongest in relation to geographic distance (Table 2;
686 Figure S6), which indicates a lack of power to detect environmental selection on this SNP
687 using the GDM. Notably, the reduced differential gene expression in *I. elegans* in response to
688 heat shock at the sampled range edge compared to the core indicates a possible loss of gene

function at the range edge (discussed in Lancaster et al. 2016), which is to be further examined.

We detected a strong selection signature for the vacuolar H+ ATPase (V-ATPase) gene (Figure 4), which is noteworthy since the activity of this gene has pleiotropic effects on both cold tolerance and salinity. V-ATPase is an ion transporter and aids in sodium (Na+) modulation at the Malpighian tubules of insects by energizing fluid secretion while coupled to an H+/K+ exchanger, modulating pH and salinity (Beyenbach, Skaer, & Dow, 2010). V-ATPases may also play a role in cold tolerance in insects, which is related to body-ion gradients regulated by water loss. The inability of insects to maintain ion gradients at low (i.e. ≤ 0 ºC) temperatures may be an important cause of mortality from cold exposure and influence cold tolerance (e.g. in bugs: Koštál & Vambera, 2004; in *Drosophila*: MacMillan et al., 2015; in crickets: MacMillan & Sinclair, 2011). Lancaster *et al.* documented phenotypic (2015) and gene expression (2016) changes in relation to cold tolerance in *I. elegans* from Sweden, with faster cold acclimation rates and unique cold-response gene expression profiles at the range edge compared to the core. This evidence for selection on cold-tolerance and the decrease in minimum temperature along our sampled gradient suggests a cold tolerance benefit for selection on V-ATPase in *I. elegans*. Notably, changes in V-ATPase activity in the optic lobe during circadian cycles has also been found in flies, indicating a role in visual processes (Górska-Andrzejak, Damulewicz, & Pyza, 2015). Our sampled gradient also exhibits variation in water body salinity, with many sites within coastal areas and others within inland freshwater lakes and ponds, some closed, and others open to the Baltic Sea. This variation in salinity may impose further selection pressure on V-ATPase genes during the aquatic larval stage of *I. elegans*. Our findings suggest that vacuolar H+ ATPase contributes to local

adaptation in *I. elegans* during its poleward range expansion, which is observed as a shift in allele frequency towards colder, range limit sites sampled from non-coastal, low salinity sites (Figure 2).

Though weaker than for V-ATPase, we detected a strong selection signature for long wavelength sensitive (LWS) opsin (annotated to SNP 73426_72; Table 2, Figure 5), which is a phylogenetically diverse class of opsins in Odonata (Suvorov et al., 2017) that have previously been identified in transcriptomic analyses of *I. elegans* in Chauhan et al. (2014, 2016). Odonates have between 3-5 classes of photoreceptors (Futahashi et al., 2015) and are involved in visual processes that are thought to play roles in food acquisition, mate choice (e.g. in cichlids: Terai, Mayer, Klein, Tichy, & Okada, 2002) development (in odonates: Futahashi et al., 2015), and sex-specific behaviours (in *I. elegans*: Chauhan et al., 2016). The importance of colour discrimination in sexual selection and sexual conflict in Odonata is well known (e.g. in *I. elegans*: Gosden & Svensson, 2009; Le Rouzic et al., 2015; Svensson et al., 2005). Further, within our study area, the frequency of *I. elegans* gynochromes (female-specific female morphs) increases with latitude and shows a frequency-dependent fitness benefit with respect to cold tolerance that may facilitate range shifts (Lancaster et al., 2017). It is possible that selection on LWS opsins through its cascading effects in sexual interactions may contribute to climate adaptation during range expansion, via social feedback mechanisms, thermal conditions and their possible interactions.

As genomic resources improve for *I. elegans* (e.g. transcriptome: Chauhan et al. 2014, 2016; genome: P. Chauhan et al. unpublished), candidate gene regions identified in this study may be more closely examined for soft selective sweeps and their emergence according to

737 climate change (Messer and Petrov 2014). Previous studies on other coenagrionid damselflies
738 have identified putatively adaptive traits in range expanding populations, for example,
739 increased flight ability and enhanced immune function in *Coenagrion scitulum* (Therry,
740 Lefevre, Bonte, & Stoks, 2014; Therry, Nilsson-Örtman, Bonte, & Stoks, 2014) and
741 identification of candidate genes associated with increased flight performance (Swaegers et
742 al., 2015). Future studies may take advantage of both phenotypic measurements and high
743 quality genomic resources to disentangle multiple functional genetic changes that occur
744 during Odonata range expansions.
745
746 *Conclusion*
747 With maximum summer temperatures (Max Temp) in our study area projected to increase up
748 to 4°C by 2050 (under RCP8.5 data from BioClim: Hijmans et al. 2005), and with similar
749 trends occurring worldwide, the development of effective approaches for measuring and
750 predicting species' adaptive responses, and thus future biodiversity structure under
751 environmental change is crucial. Our findings empirically validate a multi-tiered statistical
752 approach for uncovering spatial heterogeneity in signatures of local adaptation along
753 environmental gradients (Figure 1). Our results reveal environmental thresholds where
754 climate-mediated selection indicate that *I. elegans* is currently in the process of evolving local
755 adaptation along its range, with selection on genes that show functional relevance with respect
756 to environmental variation and stressors. The effects of plasticity and ensuing genetic
757 assimilation of adaptive traits in augmenting the persistence of *I. elegans* during range
758 expansion requires further investigation (e.g. Lande, 2009), as well as how intra- and
759 interspecific competition might also influence local adaptation (Price & Kirkpatrick 2009;
760 Case & Taper, 2000). Further, the parallel environmental gradients where *I. elegans* is subject

761 to range limit processes in northern Europe offer future opportunities for a replicated
762 investigation of parallel signatures of adaptation, which may reveal common adaptive
763 processes that apply to ectotherms more generally.
764
765 **Acknowledgements**
766 This work was supported by an EU FP7, Marie Curie International Incoming Fellowship (to
767 RYD; project code 'MOVE2ADAPT'), a Wenner-Gren Foundation Postdoctoral Stipend (to
768 RYD), the Oscar and Lili Lamm Foundation (to RYD, BH), Biodiversity and Ecosystem
769 Services in a Changing Climate (BECC; a joint Lund-Gothenburg University initiative) (LL),
770 the Swedish Research Council (EIS, BH), the Crafoord Foundation (EIS, BH) and Erik
771 Philip-Sörensens Stiftelse (E.I.S.). We would like to thank Hanna Bensch and Paul Caplat for
772 assistance with the collection of samples in the field and the Grimsö Research Station and
773 Mikael Åkesson, for logistical support. We thank Pallavi Chauhan for assistance with SNP
774 annotation. We thank Martin Andersson for assistance with DNA extraction, Jane Jönsson for
775 laboratory administration, and Julian Catchen, Martin Stervander, Dag Ahren, and Maren
776 Wellenreuther for bioinformatics advice and helpful discussion.
777
778 **Data Accessibility**
779 All supplementary files are to be deposited on DRYAD, including SNP datasets,
780 environmental data and R code for the GDM. The *I. elegans* draft genome used in the
781 manuscript will be made available via DRYAD upon manuscript acceptance.
782
783 **Author contributions**

784 Study design was conceptualised by RYD, BH, ES and LL. RYD and LL collected samples in
785 the field. RYD prepared genomic libraries and conducted bioinformatics analyses with
786 assistance from BH. RYD and CY analysed the data. CY performed statistical modelling and
787 prepared figures. RYD wrote the manuscript. All authors edited the final manuscript.
788
789 **Supplementary material**
790 **1.0 Data analysis.** 1.1 RAD library preparation, 1.2 bioinformatics, and 1.3 Outlier detection
791 **Table S1.** Summary statistics of RADseq libraries
792 **Table S2.** Genetic and environmental data for 25 *Ischnura elegans* sites in Sweden
793 **Table S3.** Pearson correlation matrix of environmental variables
794 **Figure S1.** Frequency distribution of the depth of coverage per SNP
795 **Figure S2.** Environmental variables at each site plotted against latitude
796 **Figure S3.** Barplot of assignment probability to each genetic cluster using ADMIXTURE
797 **Figure S4.** Manhattan plots of significant SNP x environment associations using LFMM
798 **Figure S5.** Histogram of Fst distribution for 1758 candidate SNPs
799 **Figure S6.** Spatial pattern of allelic turnover for SNP 53905_36
800 **Figure S7.** Spatial pattern of allelic turnover for SNP 35403_9
801 **Supplementary file 1:** Results of GDM for all candidate SNPs (.xls)
802 **Supplementary file 2:** Results of gene annotation for all candidate SNPs (.xls)

Table 1. Numbers of loci under putative selection detected via Fst Outlier and EAA approaches. Overlapping and unique (i.e non-overlapping) Fst outliers or SNP x Environment associations are shown across the 1758 candidate SNPs, identified using BAYESCAN (diversifying SNPs only), OutFLANK, and LFMM, broken down into the five tested environmental variables (Annual Temp: Mean annual temperature, BIO1; Max Temp: Mean maximum summer temperature, BIO5; Annual Rain: Mean annual precipitation, BIO12; Wind Speed, and Tree Cover). Shown are the total number of significant SNPs and the number of uniquely associated SNPs per method and environmental variable. The number of SNPs in common with the total number of SNPs ('Total SNPs') is shown in matrix form. Uniquely associating SNPs ('Unique SNPs') were those found to be specific to the method used or the environmental variable tested.

Approach		***Total SNPs***	***Unique SNPs***	**Bayescan**	**OutFlank**	**BIO1**	**BIO5**	**BIO12**	**Wind Speed**	**Tree Cover**
Fst Outlier	**Bayescan**	**391**	360	-	-	5	7	11	13	3
	OutFlank	**188**	138	9	-	11	19	14	13	19
LFMM	**BIO1**	**374**	75	5	11	-	172	116	211	97
	BIO5	**566**	114	7	19	172	-	292	146	174
	BIO12	**500**	65	11	14	116	292	-	117	182
	WS	**416**	114	13	13	211	146	117	-	86
	TC	**471**	183	3	19	97	174	182	86	-
	ALL[†]	**1251**	1188	22[‡]	41[‡]					

[†]Refers to all SNPs identified by LFMM with significant associations to environmental variables. [‡]Unique SNPs

Table 2. Gene annotations and associated environmental variables for SNPs under putative selection. Transcript IDs, gene function, genome scaffold ID ('Scaff'), and SNP ID on the *I. elegans* draft genome are shown. The difference between the highest and lowest population Fst value is shown for each annotated SNP (ΔFst). SNPs presented had a: 1) ≥70% BLAST match rate, 2) higher % of the GDM explained than the reference SNP group (% GDM), and 3) prior annotation in Lancaster et al. (2016), Chauhan et al. (2016, 2017). Environmental variables are BIO5: Maximum temperature of warmest month ('Max Temp'); BIO12: Mean annual precipitation ('Annual Rain'); TC: Tree Cover, and WS: Wind Speed. Allelic turnover is shown for each SNP relative to each environmental variable. SNP rank per environmental variable is the magnitude of allelic turnover ranked relative to the other environmental variables in the GDM. This provides a measure of the relative explanatory power of each environmental variable on allelic turnover. Bold SNPs are those for which spatial allelic turnover was mapped (53905_36 and 35404_9 in Figures S8 and S9). Transcripts for SNP 3504_9 are in supplementary data. [†]Fst outlier using BAYESCAN (diversifying), or [‡]OutFLANK. [§]Annotation previously unpublished in *I. elegans*.

							Partial allelic turnover by variable					SNP rank in GDM by variable				
Transcript ID	**Gene Function Description**	**Function**	**Scaff**	**SNP ID**	**ΔFst**	**% GDM**	**GEOG**	**BIO5**	**BIO12**	**WS**	**TC**	**GEOG**	**BIO5**	**BIO12**	**WS**	**TC**
c9603_g1_i1	Heat shock protein 70	HSP70	4300	**53905_36**[†]	0.40	19.26	0.33	0.12	0.10	0.00	0.14	171	551	451	223	1175
c48098_g1_i1	Heat shock cognate protein70	HSP70	4300	53905_36[†]	0.40	19.26	0.33	0.12	0.10	0.00	0.14	171	551	451	223	1175
c42128_g1_i1	Heat shock cognate protein70	HSP70	4300	53905_36[†]	0.40	19.26	0.33	0.12	0.10	0.00	0.14	171	551	451	223	1175
c42128_g2_i1	Heat shock cognate protein70	HSP70	4300	53905_36[†]	0.40	19.26	0.33	0.12	0.10	0.00	0.14	171	551	451	223	1175
c36939_g1_i1	Heat shock protein 40	HSP40	2	39733_28[†]	0.14	20.02	0.00	0.41	0.00	0.08	0.07	1065	75	1196	641	503
c36939_g1_i1	Heat shock protein 40	HSP40	2	39594_49[†]	0.23	33.21	0.43	0.00	0.09	0.41	0.03	114	1251	489	953	15
c36939_g1_i1	Heat shock protein 40	HSP40	2	39519_58	0.20	16.12	0.09	0.12	0.17	0.08	0.12	540	566	242	302	481
c36939_g1_i1	Heat shock protein 40	HSP40	2	39594_63[†]	0.16	18.83	0.00	0.06	0.52	0.00	0.05	1064	798	9	770	1124
c36939_g1_i1	Heat shock protein 40	HSP40	2	39519_36[†]	0.19	22.80	0.03	0.01	0.15	0.31	0.03	788	1126	299	962	35
c36939_g1_i1	Heat shock protein 40	HSP40	2	39692_78	0.16	23.16	0.04	0.00	0.34	0.00	0.12	708	1639	43	310	997
c36939_g1_i1	Heat shock protein 40	HSP40	2	39648_74	0.32	21.82	0.18	0.07	0.14	0.05	0.12	331	745	334	313	636
c36939_g1_i1	Heat shock protein 40	HSP40	2	39594_35[†]	0.21	16.00	0.06	0.24	0.07	0.00	0.14	643	260	632	201	1123
c36939_g1_i1	Heat shock protein 40	HSP40	2	39648_33	0.24	14.25	0.13	0.10	0.01	0.13	0.08	430	646	1036	545	284
c36939_g1_i1	Heat shock protein 40	HSP40	2	39648_19	0.24	14.54	0.13	0.10	0.01	0.14	0.08	442	645	1037	546	274
c36939_g1_i1	Heat shock protein 40	HSP40	2	39692_51	0.19	21.25	0.00	0.05	0.28	0.00	0.12	1611	858	78	289	979
c43579_g4_i1	long wavelength-sensitive opsin3b	Visual	6	**73426_72**	0.21	14.00	0.05	0.24	0.08	0.15	0.00	703	266	531	1509	231
c43579_g4_i1	long wavelength-sensitive opsin3b	Visual	6	73426_69	0.19	18.48	0.62	0.21	0.08	0.15	0.03	53	317	568	987	237
c43579_g4_i1	long wavelength-sensitive opsin3b	Visual	6	73426_85	0.19	21.31	0.74	0.21	0.07	0.14	0.03	34	329	600	997	271

c22378_g1_i1	tpa_exp: pteropsin	Visual	47	57982_91	0.15	34.07	0.11	0.71	0.00	0.02	0.07	497	8	1063	621	875
c39329_g1_i1	rhodopsin-specific isozyme-like	Visual	38855	50006_70	0.18	25.26	0.00	0.12	0.31	0.00	0.13	1577	553	58	264	1578
c4570_g1_i1	histone-lysine n-methyltransferase	Epigenetics	102	1735_72	0.18	12.79	0.01	0.00	0.16	0.14	0.12	887	1443	252	285	273
c26924_g1_i2	histone-lysine n-methyltransferase	Epigenetics	102	1735_72	0.18	12.79	0.01	0.00	0.16	0.14	0.12	887	1443	252	285	273
c4570_g1_i1	histone-lysine n-methyltransferase	Epigenetics	102	1803_8†	0.17	17.18	0.12	0.02	0.03	0.11	0.23	460	1043	855	34	369
c26924_g1_i2	histone-lysine n-methyltransferase	Epigenetics	102	1803_8†	0.17	17.18	0.12	0.02	0.03	0.11	0.23	460	1043	855	34	369
c4570_g1_i1	histone-lysine n-methyltransferase	Epigenetics	102	1735_84	0.13	27.06	0.00	0.01	0.49	0.02	0.14	1365	1088	18	208	861
c26924_g1_i2	histone-lysine n-methyltransferase	Epigenetics	102	1735_84	0.13	27.06	0.00	0.01	0.49	0.02	0.14	1365	1088	18	208	861
c28633_g2_i1	histone-lysine n-methyltransferase	Epigenetics	383	49386_86	0.09	15.91	0.22	0.36	0.00	0.01	0.03	260	106	1247	982	918
c28633_g1_i2	histone-lysine n-methyltransferase	Epigenetics	383	49386_86	0.09	15.91	0.22	0.36	0.00	0.01	0.03	260	106	1247	982	918
c33122_g1_i1	vacuolar H+ ATPase§	Proton pump	28	**37543_9**‡	0.50	47.31	0.00	0.63	0.06	0.06	0.22	987	10	643	575	44
10 matches	Intra-cellular processes§	various	26	**35404_9**‡	0.38	42.05	0.031	0.83	0.16	0	0.01	781	3	254	1049	1179

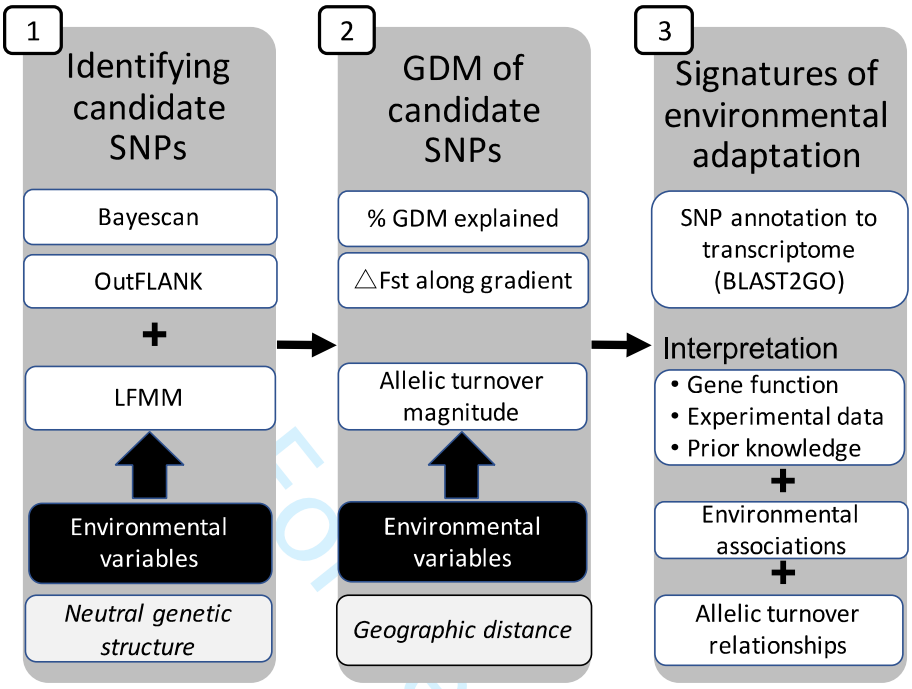


Figure 1. Flowchart of analytical approach. (1) Candidate SNPs under putative selection are identified using two Fst outlier approaches (BAYESCAN, OutFLANK) and one Environmental Association Analysis approach (LFMM). LFMM incorporates a prior estimate of neutral genetic structure and environmental variables from each sampling location. (2) Generalized Dissimilarity Modelling (GDM) is applied to each candidate SNP to determine relationships between SNP allelic turnover magnitude and environmental gradients, and geographic distance. SNP response is assessed via the maximum change in Fst between sampling locations (ΔFst), and the explanatory power of the SNP in the GDM via percentage deviance explained (% GDM explained). SNPs with a % GDM explained ≤ that of the reference SNP group were exlcluded as potential false positives (3) Signatures of environmental adaptation are characterised via annotation of SNPs to a transcriptome and interpreted based on gene function, prior knowledge, experimental data, SNP x environment associations and the allelic turnover relationships observed.

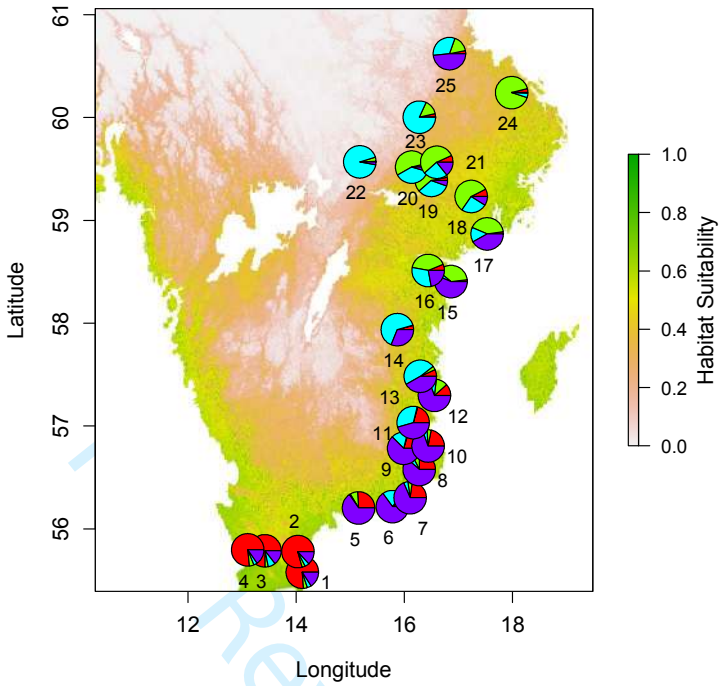


Figure 2. Genetic structure of *I. elegans* across the environmental gradient. Probability of *I. elegans* genetic cluster assignment (K=4) is shown at the population level (with population names from Table 1) on a habitat suitability map in Sweden (from Lancaster et al., 2015). The proportion of each color within each pie chart indicates the mean assignment probability of individuals to a genetic cluster in that population, displayed for 426 individuals across 25 populations.

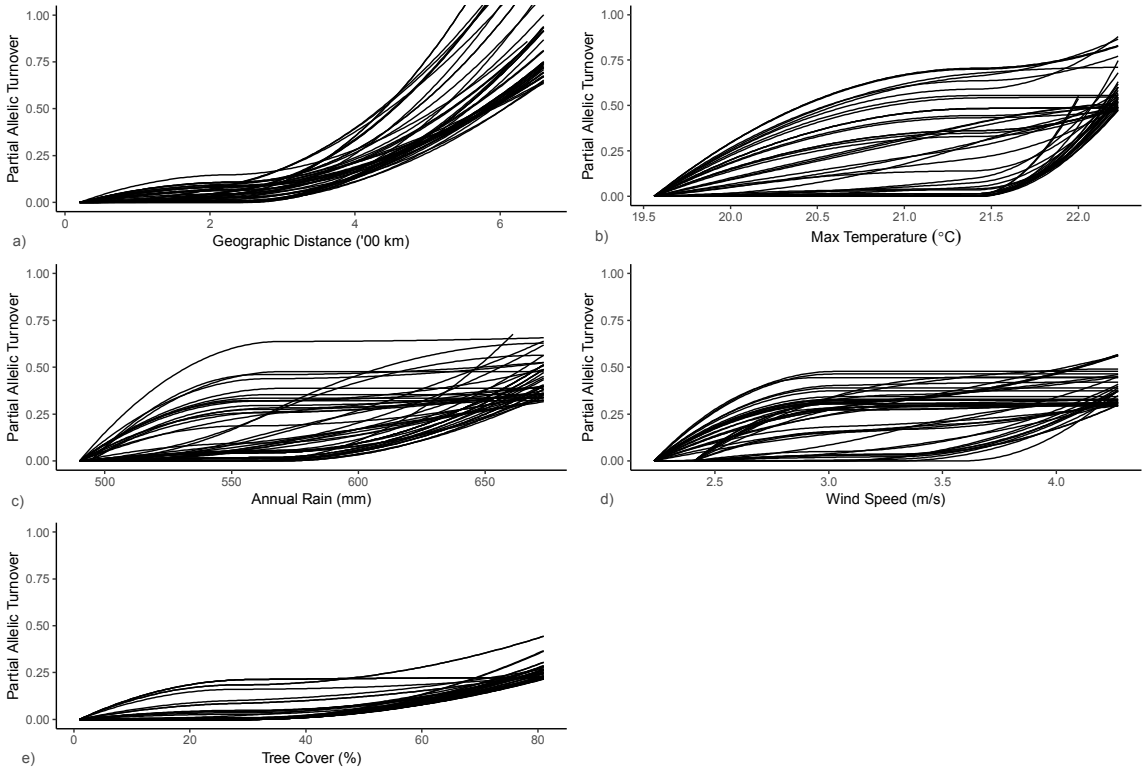


Figure 3. Allelic turnover relationships for each environmental variable. Allelic turnover functions for the top 250 SNPs (grey) and top 50 SNPs (black) that showed the highest General Dissimilarity Modelling (GDM) partial allelic turnover in relation to each environmental variable and geographic distance: a) Geographic distance (km) b) Max Temperature (^{o}C; BIO5), c) Annual Rain (mm; BIO12), d) Tree Cover (%), e) Wind Speed (m/s).

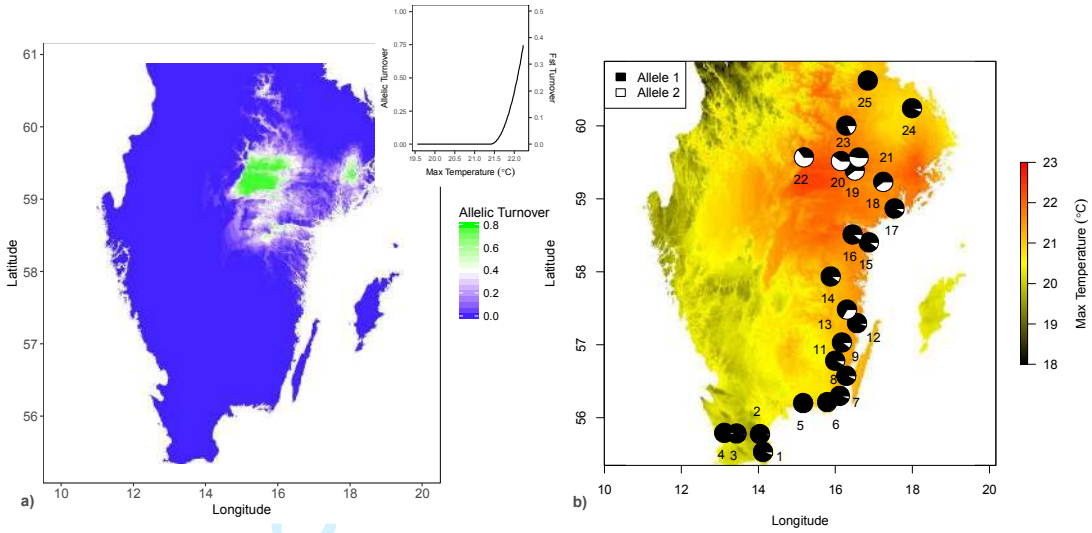


Figure 4. Allelic turnover of SNP ID 37543_9, a) shown as the allelic turnover response curve in relation to BIO5 and Partial Fst change, and b) allele frequency for each sampling location (black = high frequency allele 1, white = low frequency allele 2) mapped on BIO5 (Maximum Mean Summer Temperature). Allele 2 undergoes substantial change in frequency from south to north, increasing in warmer inland sites, before becoming less frequent at the cooler extreme range edge. This SNP was annotated to a gene for *vacuolar H+ ATPase*, involved in proton pump activity, and was associated most strongly with BIO5, with a maximum Fst change of 0.50 across the sampling gradient.

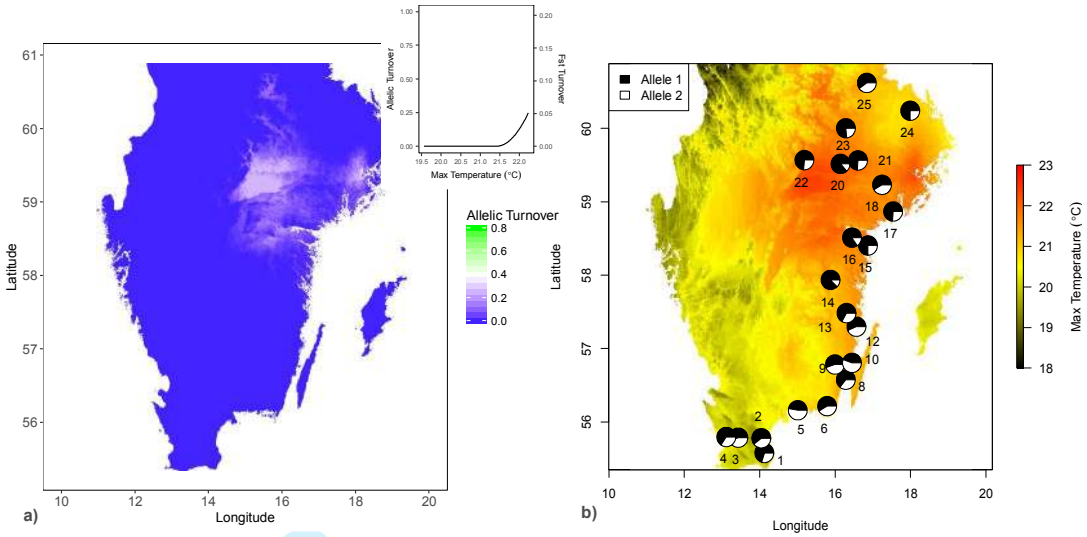


Figure 5. Allelic turnover of SNP ID 73426_72, a) shown as the allelic turnover response curve in relation to BIO5 and Partial Fst change, and b) allele frequency for each sampling location (black = high frequency allele 1, white = low frequency allele 2) mapped on BIO5 (Maximum Mean Summer Temperature). Alleles 1 and 2 have comparable frequencies up to the mid-north latitudes, beyond which allele 1 increases in frequency towards the inland and coastal range limit. This SNP was annotated to a *long wavelength sensitive opsin gene 3b*, involved in visual processing, and was associated most strongly with BIO5, with a maximum Fst change of 0.21 across the sampling gradient.

References

Advani, N. K., Kenkel, C. D., Davies, S. W., Parmesan, C., Singer, M. C., & Matz, M. V. (2016). Variation in heat shock protein expression at the latitudinal range limits of a widely-distributed species, the Glanville fritillary butterfly (*Melitaea cinxia*). *Physiological Entomology, 41*(3), 241-248. doi:10.1111/phen.12148

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research, 19*(1655-1664). doi:10.1101/gr.094052.109

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Arribas, P., Abellán, P., Velasco, J., Millán, A., & Sánchez‐Fernández, D. (2017). Conservation of insects in the face of global climate change. In S. Johnson & T. Jones (Eds.), *Global Climate Change and Terrestrial Invertebrates* (pp. 349-367): Wiley-Blackwell.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., . . . Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics, 25*(1), 25-29. doi:10.1038/75556

Berg, J., & Coop, G. (2014). A population genetic signal of polygenic adaptation. *PloS Genetics, 10*(8), e1004412.

Bernatchez, S., Laporte, M., Perrier, C., Sirois, P., & Bernatchez, L. (2016). Investigating genomic and phenotypic parallelism between piscivorous and planktivorous lake trout (*Salvelinus namaycush*) ecotypes by means of RADseq and morphometrics analyses. *Molecular Ecology, 25*(19), 4773-4792. doi:10.1111/mec.13795

Beyenbach, K. W., Skaer, H., & Dow, J. A. T. (2010). The developmental, molecular, and transport biology of Malpighian tubules. *Annual Review of Entomology, 55*, 351-374.

Bridle, J. R., & Vines, T. H. (2006). Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology & Evolution, 22*(3), 140-147. doi:10.1016/j.tree.2006.11.002

Bybee, S., Cordoba-Aguilar, A., Duryea, M. C., Futahashi, R., Hansson, B., Lorenzo-Carballa, M. O., . . . Wellenreuther, M. (2016). Odonata (dragonflies and damselflies) as a bridge between ecology and evolutionary genomics. *Frontiers in Zoology, 13*, 46. doi:10.1186/s12983-016-0176-7

Case, T. J., & Taper, M. L. (2000). Interspecific competition, environmental gradients, gene flow, and the coevolution of species' borders. *American Naturalist, 155*(5), 583-605.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology, 22*(11), 3124-3140. doi:10.1111/mec.12354

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: building and genotyping Loci de novo from short-read sequences. *G3 (Bethesda), 1*(3), 171-182. doi:10.1534/g3.111.000240

Chauhan, P., Hansson, B., Kraaijeveld, K., de Knijff, P., Svensson, E., & Wellenreuther, M. (2014). De novo transcriptome of Ischnura elegans provides insights into sensory biology, colour and vision genes. *BMC Genomics, 15*, 808.

Chauhan, P., Wellenreuther, M., & Hansson, B. (2016). Transcriptome profiling in the damselfly Ischnura elegans identifies genes with sex-biased expression. *BMC Genomics, 17*(1), 985. doi:10.1186/s12864-016-3334-6

Chen, Z., Farrell, A. P., Matala, A., & Narum, S. R. (2018) Mechanisms of thermal adaptation and evolutionary potential of conspecific populations to changing environments. *Molecular Ecology, 00,* 1-16. doi:10.1111/mec.14475.

Chippindale, A. K., Gibson, J. R., & Rice, W. R. (2001). Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. *Proceedings of the National Academy of Sciences USA, 98*, 1671–1675.

Colautti, R., & Barrett, S. (2013). Rapid adaptation to clmate facilitates range expansion of an invasive plant. *Science, 342*(10), 364-366.

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., & Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics, 21*(18), 3674-3676. doi:10.1093/bioinformatics/bti610

Creech, T. G., Epps, C. W., Landguth, E. L., Wehausen, J. D., Crowhurst, R. S., Holton, B., & Monello, R. J. (2017). Simulating the spread of selection-driven genotypes using landscape resistance models for desert bighorn sheep. *PloS One, 12*(5), e0176960. doi:10.1371/journal.pone.0176960

Defries, R. S., Hansen, M. C., Townshend, J. R. G., Janetos, A. C., & Loveland, T. R. (2000). A new global 1-km dataset of percentage tree cover derived from remote sensing. *Global Change Biology, 6*(2), 247-254. doi:10.1046/j.1365-2486.2000.00296.x

Devlin, B., & Roeder, K. (1999). Genomic control for association studies. *Biometrics, 55*(4), 997-1004.

Dijkstra, K., & Lewington, R. (2006). *Field guide to the dragonflies of Britain and Europe.* Gillingham, UK: British Wildlife Publishing.

Etter, P. D., Bassham, S., Hohenlohe, P. A., Johnson, E. A., & Cresko, W. A. (2011). SNP discovery and genotyping for evolutionary genetics using RAD sequencing. *Methods in Molecular Biology, 772*, 157-178.

Ferrier, S., Manion, G., Elith, J., & Richardson, K. (2007). Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions, 13*(3), 252-264. doi:10.1111/j.1472-4642.2007.00341.x

Fick, S., & Hijmans, R. (2017). Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*. doi:10.1002/joc.5086

Fitzpatrick, M. C., & Keller, S. R. (2015). Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters, 18*(1), 1-16. doi:10.1111/ele.12376

Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A bayesian perspective. *Genetics, 180*(2), 977.

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution, 6*(8), 925-929. doi:10.1111/2041-210X.12382

Frichot, E., Schoville, S., Bouchard, G., & Francois, O. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution, 30*(7), 1687-1699. doi:10.1093/molbev/mst063

Futahashi, R., Kawahara-Miki, R., Kinoshita, M., Yoshitake, K., Yajima, S., Arikawa, K., & Fukatsu, T. (2015). Extraordinary diversity of visual opsin genes in dragonflies. *Proceedings of the National Academy of Sciences, 112*(11), E1247-E1256. doi:10.1073/pnas.1424670112

Górska-Andrzejak, J., Damulewicz, M., & Pyza, E. (2015). Circadian Rhythms of Ion Transporters in the Visual System of Insects. In K. Hyndman & T. Pannabecker (Eds.), *Sodium and Water Homeostasis. Physiology in Health and Disease*. New York, NY: Springer.

Gosden, T. P., & Svensson, E. I. (2009). Density-dependent male mating harassment, female resistance and male mimicry. *American Naturalist, 173*, 709-721.

Goslee, S. C., & Urban, D. L. (2007). The ecodist Package for dissimilarity-based analysis of ecological data. *2007, 22*(7), 19. doi:10.18637/jss.v022.i07

Goudet, J. (2005). HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes, 5*, 184-186. doi:10.1111/j.1471-8278

Gralka, M., Stiewe, F., Farrell, F., Möbius, W., Waclaw, B., & Hallatschek, O. (2016). Allele surfing promotes microbial adaptation from standing variation. *Ecology Letters, 19*, 889-898.

Halbritter, A. H., Billeter, R., Edwards, P. J., & Alexander, J. M. (2015). Local adaptation at range edges: comparing elevation and latitudinal gradients. *Journal of Evolutionary Biology, 28*(10), 1849-1860. doi:10.1111/jeb.12701

Hansen, M. M., Olivieri, I., Waller, D. M., Nielsen, E. E., & Ge, M. W. G. (2012). Monitoring adaptive genetic responses to environmental change. *Molecular Ecology, 21*(6), 1311-1329. doi:10.1111/j.1365-294X.2011.05463.x

Herold, C., Mattheisen, M., Lacour, A., Vaitsiakhovich, T., Angisch, M., Drichel, D., & Becker, T. (2012). Integrated genome-wide pathway association analysis with INTERSNP. *Human Heredity, 73*, 63-72.

Hickling, R., Roy, D. B., Hill, J. K., & Thomas, C. D. (2005). A northward shift of range margins in British Odonata. *Global Change Biology, 11*(3), 502-506. doi:10.1111/j.1365-2486.2005.00904.x

Hijmans, R., & van Etten, J. (2012). raster: Geographic analysis and modeling with raster data. R package version 2.0-12.: http://cran.r-project.org/package=raster.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology, 25*(15), 1965-1978. doi:10.1002/joc.1276

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., . . . Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *American Naturalist, 188*(4), 379-397. doi:10.1086/688018

Hoekstra, H. E., Hirschmann, R. J., Bundey, R. J., Insel, P., & Crossland, J. P. (2006). A single amino acid mutation contributes to adaptive color pattern in beach mice. *Science, 313*, 101-104.

Hoffmann, A., Griffin, P., Dillon, S., Catullo, R., Rane, R., Byrne, M., . . . Sgrò, C. (2015). A framework for incorporating evolutionary genomics into biodiversity conservation and management. *Climate Change Responses, 2*(1). doi:10.1186/s40665-014-0009-x

Hoffmann, A., & Sgro, C. (2011). Climate change and evolutionary adaptation. *Nature, 470*(7335), 479-485.

Hofmann, T. A., & Mason, C. F. (2005). Habitat characteristics and the distribution of Odonata in a lowland river catchment in eastern England. *Hydrobiologia, 539*(1), 137-147. doi:10.1007/s10750-004-3916-1

Jacob, S., Legrand, D., Chaine, A. S., Bonte, D., Schtickzelle, N., Huet, M., & Clobert, J. (2017). Gene flow favours local adaptation under habitat choice in ciliate microcosms. *Nature Ecology & Evolution, 1*(9), 1407-1410. doi:10.1038/s41559-017-0269-5

Jaenson, T. G. T., Jaenson, D. G. E., Eisen, L., Petersson, E., & Lindgren, E. (2012). Changes in the geographical distribution and abundance of the tick Ixodes ricinus during the past 30 years in Sweden. *Parasites & Vectors, 5*, 15. doi:8 10.1186/1756-3305-5-8

Jaeschke, A., Bittner, T., Reineking, B., & Beierkuhnlein, C. (2013). Can they keep up with climate change? – Integrating specific dispersal abilities of protected Odonata

in species distribution modelling. *Insect Conservation and Diversity, 6*(1), 93-103. doi:10.1111/j.1752-4598.2012.00194.x

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., . . . Hunter, S. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics, 30*(9), 1236-1240. doi:10.1093/bioinformatics/btu031

Kirkpatrick, M., & Barton, N. H. (1997). Evolution of a Species' Range. *American Naturalist, 150*(1), 1-23. doi:10.1086/286054

Klopfstein, S., Currat, M., & Excoffier, L. (2006). The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution, 23*, 482–490.

Koštál, V. J., & Vambera, J. B. (2004). On the nature of pre-freeze mortality in insects: water balance, ion homeostasis and energy charge in the adults of *Pyrrhocoris apterus*. *The Journal of Experimental Biology, 207*, 1509-1521.

Lancaster, L. T. (2016). Widespread range expansions shape latitudinal variation in insect thermal limits. *Nature Climate Change, 6*, 618-621.

Lancaster, L. T., Dudaniec, R. Y., Chauhan, P., Wellenreuther, M., Svensson, E. I., & Hansson, B. (2016). Gene expression under thermal stress varies across a geographic range expansion front. *Molecular Ecology*.

Lancaster, L. T., Dudaniec, R. Y., Hansson, B., & Svensson, E. I. (2015). Latitudinal shift in thermal niche breadth results from thermal release during a climate‐mediated range expansion. *Journal of Biogeography, 42*(10), 1953-1963.

Lancaster, L. T., Dudaniec, R. Y., Hansson, B., & Svensson, E. I. (2017). Do group dynamics affect colour morph clines during a range shift? *Journal of Evolutionary Biology, 30*(4), 728-737. doi:10.1111/jeb.13037

Lande, R. (2009). Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *Journal of Evolutionary Biology, 22*, 1435-1446.

Landguth, E. L., Bearlin, A., Day, C. C., & Dunham, J. (2017). CDMetaPOP: an individual-based, eco-evolutionary model for spatially explicit simulation of landscape demogenetics. *Methods in Ecology and Evolution, 8*(1), 4-11. doi:10.1111/2041-210X.12608

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods, 9*(4), 357-359.

Le Rouzic, A., Hansen, T. F., Gosden, T. P., & Svensson, E. I. (2015). Evolutionary time-series analysis reveals the signature of frequency-dependent selection on a female mating polymorphism. *American Naturalist, 185*, E182-E196. doi:10.1086/680982

Lee, C.-R., & Mitchell-Olds, T. (2012). Environmental adaptation contributes to genepolymorphism across the *Arabidopsis thaliana* genome. *Molecular Biology and Evolution, 29*, 3721-3728. doi:10.1093/molbev/mss174

Lotterhos, K. E., & Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology, 24*(5), 1031-1046. doi:10.1111/mec.13100

MacMillan, H. A., Ferguson, L. V., Nicolai, A., Donini, A., Staples, J. F., & Sinclair, B. J. (2015). Parallel ionoregulatory adjustments underlie phenotypic plasticity and evolution of *Drosophila* cold tolerance. *The Journal of Experimental Biology, 218*, 423-432. doi:10.1242/jeb.115790

MacMillan, H. A., & Sinclair, B. J. (2011). The role of the gut in insect chilling injury: cold-induced disruption of osmoregulation in the fall field cricket, *Gryllus pennsylvanicus*. *The Journal of Experimental Biology, 214*, 726-734.

Manion, G., Lisk, M., Ferrier, S., Nieto-Lugilde, D., Mokany, K., & Fitzpatrick, M. (2017). Package 'gdm', A toolkit with functions to fit, plot, and summarize Generalized Dissimilarity Models: CRAN Repository, R.

Micheletti, S. J., Matala, A. R., Matala, A. P., & Narum, S. R. (2018). Landscape features along migratory routes influence adaptive genomic variation in anadromous steelhead (*Oncorhynchus mykiss*). *Molecular Ecology, 27*(1), 128-145. doi:doi:10.1111/mec.14407

Moody, K. N., Hunter, S. N., Childress, M. J., Blob, R. W., Schoenfuss, H. L., Blum, M. J., & Ptacek, M. B. (2015). Local adaptation despite high gene flow in the waterfall-climbing Hawaiian goby, *Sicyopterus stimpsoni*. *Molecular Ecology, 24*(3), 545-563. doi:10.1111/mec.13016

Narum, S. R., & Hess, J. E. (2011). Comparison of FST outlier tests for SNP loci under selection. *Molecular Ecology Resources, 11*, 184-194. doi:10.1111/j.1755-0998.2011.02987.x

Nei, M. (1987). *Molecular Evolutionary Genetics*. New York: Columbia University Press.

Nishi, T., & Forgac, M. (2002). The vacuolar (H+)-ATPases--nature's most versatile proton pumps. *Nature Reviews in Molecular and Cellular Biology, 3*(2), 94-103. doi:10.1038/nrm729

Okuyama, H., Samejima, Y., & Tsubaki, Y. (2013). Habitat segregation of sympatric Mnais damselflies (Odonata: Calopterygidae): microhabitat insolation preferences and competition for territorial space. *International Journal of Odonatology, 16*(2), 109-117. doi:10.1080/13887890.2012.762745

Paris, J. R., Stevens, J. R., & Catchen, J. M. (2017). Lost in parameter space: a road map for STACKS. *Methods in Ecology and Evolution*, *8,* 1360-1373. doi:10.1111/2041-210X.12775

Price, T.D., Kirkpatrick, M. (2009) Evolutionarily stable range limits set by interspecific competition. *Proceedings of the Biological Society of London B. 276,* (1661):1429-1434. doi:10.1098/rspb.2008.1199

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology, 24*(17), 4348-4370. doi:10.1111/mec.13322

Sánchez-Guillén, R., Muñoz, J., Rodríguez-Tapia, G., Arroyo, T., & Córdoba-Aguilar, A. (2013). Climate-induced range shifts and possible hybridisation consequences in insects. *PloS One, 8*, e80531.

Schauvliege, R., Janssens, S., & Beyaert, R. (2007). Pellino Proteins: Novel Players in TLR and IL-1R Signalling. *Journal of Cellular and Molecular Medicine, 11*(3), 453-461. doi:10.1111/j.1582-4934.2007.00040.x

Simmen, T., Aslan, J. E., Blagoveshchenskaya, A. D., Thomas, L., Wan, L., Y., X., . . . Thomas, G. (2005). PACS-2 controls endoplasmic reticulum–mitochondria communication and Bid-mediated apoptosis. *The EMBO Journal, 24*, 717-729. doi:10.1038/

Smadja, C. M., & Butlin, R. K. (2011). A framework for comparing processes of speciation in the presence of gene flow. *Molecular Ecology, 20*(24), 5123-5140. doi:10.1111/j.1365-294X.2011.05350.x

Sørensen, J. G., Kristensen, T. N., & Loeschcke, V. (2003). The evolutionary and ecological role of heat shock proteins. *Ecology Letters, 6*(11), 1025-1037. doi:10.1046/j.1461-0248.2003.00528.x

Sork, V. L., Squire, K., Gugger, P. F., Steele, S. E., Levy, E. D., & Eckert, A. J. (2016). Landscape genomic analysis of candidate genes for climate adaptation in a California endemic oak, *Quercus lobata*. *American Journal of Botany*(103), 33-46.

Suvorov, A., Jensen, N. O., Sharkey, C. R., Fujimoto, M. S., Bodily, P., Wightman, H. M., . . . Bybee, S. M. (2017). Opsins have evolved under the permanent heterozygote model: insights from phylotranscriptomics of Odonata. *Molecular Ecology, 26*(5), 1306-1322. doi:10.1111/mec.13884

Svensson, E., & Abbott, J. (2005). Evolutionary dynamics and population biology of a polymorphic insect. *Journal of Evolutionary Biology, 18*(6), 1503-1514. doi:10.1111/j.1420-9101.2005.00946.x

Svensson, E., Abbott, J., & Härdling, R. (2005). Female polymorphism, frequency dependence, and rapid evolutionary dynamics in natural populations. *American Naturalist, 165*(5), 567-576. doi:10.1086/429278

Swaegers, J., Mergeay, J., Therry, L., Larmuseau, M., Bonte, D., & Stoks, R. (2013). Rapid range expansion increases genetic differentiation while causing limited reduction in genetic diversity in a damselfly. *Heredity, 111*, 422-429.

Swaegers, J., Mergeay, J., Van Geystelen, A., Therry, L., Larmuseau, M., & Stoks, R. (2015). Neutral and adaptive genomic signatures of rapid poleward range expansion. *Molecular Ecology, 24*(24), 6163-6176. doi:10.1111/mec.13462

Takahashi, Y., Suyama, Y., Matsuki, Y., Funayama, R., Nakayama, K., & Kawata, M. (2016). Lack of genetic variation prevents adaptation at the geographic range margin in a damselfly. *Molecular Ecology, 25*(18), 4450-4460. doi:10.1111/mec.13782

Terai, Y., Mayer, W. E., Klein, J., Tichy, H., & Okada, N. (2002). The effect of selection on a long wavelength sensitive (LWS) opsin gene of Lake Victoria cichlid fishes. *Proceedings of the National Academy of Sciences, 99*, 15501-15506.

Therry, L., Lefevre, E., Bonte, D., & Stoks, R. (2014). Increased activity and growth rate in the non-dispersive aquatic larval stage of a damselfly at an expanding range edge. *Freshwater Biology, 59*, 1266-1277.

Therry, L., Nilsson‐Örtman, V., Bonte, D., & Stoks, R. (2014). Rapid evolution of larval life history, adult immune function and flight muscles in a poleward-moving damselfly. *Journal of Evolutionary Biology, 27*, 141-152.

Travis, J. M. J., Munkemuller, T., Burton, O. J., Best, A., Dytham, C., & Johst, K. (2007). Deleterious mutations can surf to high densities on the wave front of an expanding population. *Molecular Biology and Evolution, 24*, 2334–2343.

Trumbo, D. R., Epstein, B., Hohenlohe, P. A., Alford, R. A., Schwarzkopf, L., & Storfer, A. (2016). Mixed population genomics support for the central marginal hypothesis across the invasive range of the cane toad (Rhinella marina) in Australia. *Molecular Ecology, 25*(17), 4161-4176. doi:10.1111/mec.13754

Warren, M. S., Hill, J. K., Thomas, J. A., Asher, J., Fox, R., Huntley, B., . . . Thomas, C. D. (2001). Rapid responses of British butterflies to opposing forces of climate and habitat change. *Nature, 414*(6859), 65-69.

Watts, P., Keat, S., & Thompson, D. (2010). Patterns of spatial genetic structure and diversity at the onset of a rapid range expansion: colonisation of the UK by the small red-eyed damselfly *Erythromma viridulum*. *Biological Invasions, 12*, 3887.

Wellenreuther, M., Larson, K. W., & Svensson, E. I. (2012). Climatic niche divergence or conservatism? Environmental niches and range limits in ecologically similar damselflies. *Ecology, 93*(6), 1353-1366.

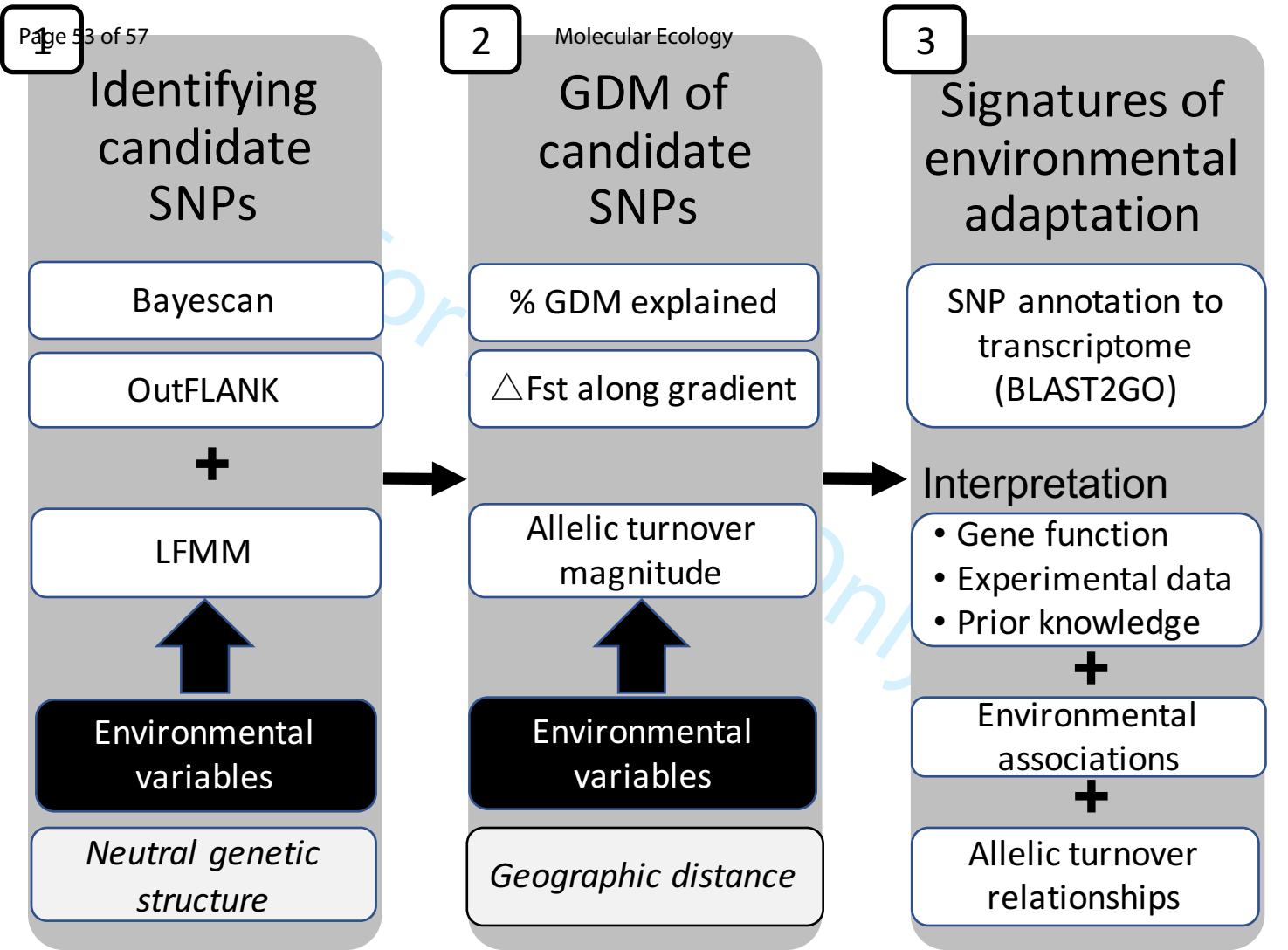
Whitlock, M. C., & Lotterhos, K. E. (2015). Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of Fst. *American Naturalist, 186*, S24-36. doi:10.1086/682949

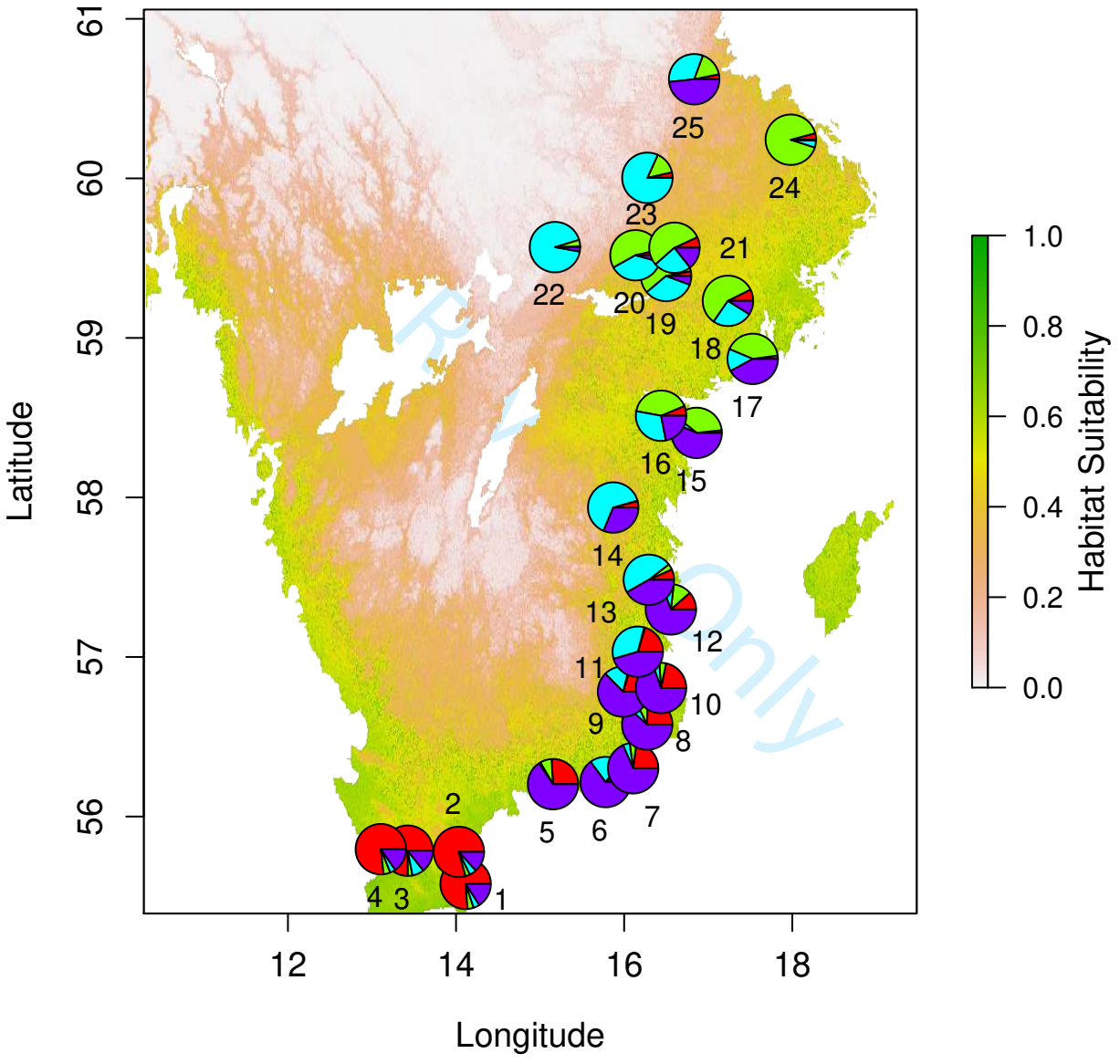
Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag.

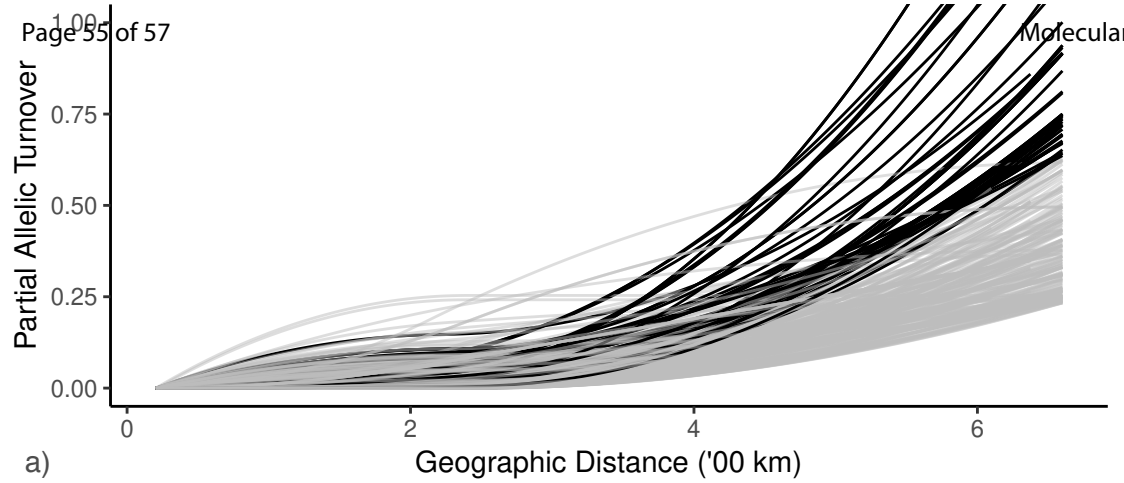
Woodward, G., Perkins, D. M., & Brown, L. E. (2010). Climate change and freshwater ecosystems: impacts across multiple levels of organization. *Philosophical Transactions of the Royal Society of London B Biological Sciences, 365*(1549), 2093-2106.

Approach		*Total SNPs*	*Unique SNPs*	Bayescan	OutFlank	BIO1	BIO5	BIO12	Wind Speed	Tree Cover
Fst Outlier	**Bayescan**	**391**	360	-	-	5	7	11	13	3
	OutFlank	**188**	138	9	-	11	19	14	13	19
LFMM	**BIO1**	**374**	75	5	11	-	172	116	211	97
	BIO5	**566**	114	7	19	172	-	292	146	174
	BIO12	**500**	65	11	14	116	292	-	117	182
	WS	**416**	114	13	13	211	146	117	-	86
	TC	**471**	183	3	19	97	174	182	86	-
	ALL[†]	**1251**	1188	22[‡]	41[‡]					

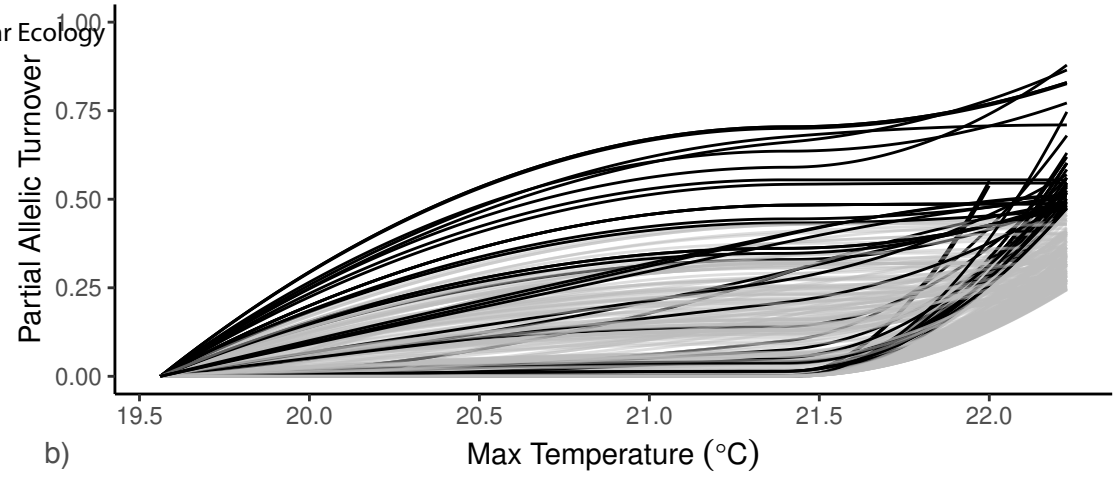
							Partial allelic turnover by variable					SNP rank in GDM by variable				
Transcript ID	Gene Function Description	Function	Scaff	SNP ID	ΔFst	% GDM	GEOG	BIO5	BIO12	WS	TC	GEOG	BIO5	BIO12	WS	TC
c9603_g1_i1	Heat shock protein 70	HSP70	4300	**53905_36**†	0.40	19.26	0.33	0.12	0.10	0.00	0.14	171	551	451	223	1175
c48098_g1_i1	Heat shock cognate protein70	HSP70	4300	53905_36†	0.40	19.26	0.33	0.12	0.10	0.00	0.14	171	551	451	223	1175
c42128_g1_i1	Heat shock cognate protein70	HSP70	4300	53905_36†	0.40	19.26	0.33	0.12	0.10	0.00	0.14	171	551	451	223	1175
c42128_g2_i1	Heat shock cognate protein70	HSP70	4300	53905_36†	0.40	19.26	0.33	0.12	0.10	0.00	0.14	171	551	451	223	1175
c36939_g1_i1	Heat shock protein 40	HSP40	2	39733_28†	0.14	20.02	0.00	0.41	0.00	0.08	0.07	1065	75	1196	641	503
c36939_g1_i1	Heat shock protein 40	HSP40	2	39594_49†	0.23	33.21	0.43	0.00	0.09	0.41	0.03	114	1251	489	953	15
c36939_g1_i1	Heat shock protein 40	HSP40	2	39519_58	0.20	16.12	0.09	0.12	0.17	0.08	0.12	540	566	242	302	481
c36939_g1_i1	Heat shock protein 40	HSP40	2	39594_63†	0.16	18.83	0.00	0.06	0.52	0.00	0.05	1064	798	9	770	1124
c36939_g1_i1	Heat shock protein 40	HSP40	2	39519_36†	0.19	22.80	0.03	0.01	0.15	0.31	0.03	788	1126	299	962	35
c36939_g1_i1	Heat shock protein 40	HSP40	2	39692_78	0.16	23.16	0.04	0.00	0.34	0.00	0.12	708	1639	43	310	997
c36939_g1_i1	Heat shock protein 40	HSP40	2	39648_74	0.32	21.82	0.18	0.07	0.14	0.05	0.12	331	745	334	313	636
c36939_g1_i1	Heat shock protein 40	HSP40	2	39594_35†	0.21	16.00	0.06	0.24	0.07	0.00	0.14	643	260	632	201	1123
c36939_g1_i1	Heat shock protein 40	HSP40	2	39648_33	0.24	14.25	0.13	0.10	0.01	0.13	0.08	430	646	1036	545	284
c36939_g1_i1	Heat shock protein 40	HSP40	2	39648_19	0.24	14.54	0.13	0.10	0.01	0.14	0.08	442	645	1037	546	274
c36939_g1_i1	Heat shock protein 40	HSP40	2	39692_51	0.19	21.25	0.00	0.05	0.28	0.00	0.12	1611	858	78	289	979
c43579_g4_i1	long wavelength-sensitive opsin3b	Visual	6	**73426_72**	0.21	14.00	0.05	0.24	0.08	0.15	0.00	703	266	531	1509	231
c43579_g4_i1	long wavelength-sensitive opsin3b	Visual	6	73426_69	0.19	18.48	0.62	0.21	0.08	0.15	0.03	53	317	568	987	237
c43579_g4_i1	long wavelength-sensitive opsin3b	Visual	6	73426_85	0.19	21.31	0.74	0.21	0.07	0.14	0.03	34	329	600	997	271
c22378_g1_i1	tpa_exp: pteropsin	Visual	47	57982_91	0.15	34.07	0.11	0.71	0.00	0.02	0.07	497	8	1063	621	875
c39329_g1_i1	rhodopsin-specific isozyme-like	Visual	38855	50006_70	0.18	25.26	0.00	0.12	0.31	0.00	0.13	1577	553	58	264	1578
c4570_g1_i1	histone-lysine n-methyltransferase	Epigenetics	102	1735_72	0.18	12.79	0.01	0.00	0.16	0.14	0.12	887	1443	252	285	273
c26924_g1_i2	histone-lysine n-methyltransferase	Epigenetics	102	1735_72	0.18	12.79	0.01	0.00	0.16	0.14	0.12	887	1443	252	285	273
c4570_g1_i1	histone-lysine n-methyltransferase	Epigenetics	102	1803_8†	0.17	17.18	0.12	0.02	0.03	0.11	0.23	460	1043	855	34	369
c26924_g1_i2	histone-lysine n-methyltransferase	Epigenetics	102	1803_8†	0.17	17.18	0.12	0.02	0.03	0.11	0.23	460	1043	855	34	369
c4570_g1_i1	histone-lysine n-methyltransferase	Epigenetics	102	1735_84	0.13	27.06	0.00	0.01	0.49	0.02	0.14	1365	1088	18	208	861
c26924_g1_i2	histone-lysine n-methyltransferase	Epigenetics	102	1735_84	0.13	27.06	0.00	0.01	0.49	0.02	0.14	1365	1088	18	208	861
c28633_g2_i1	histone-lysine n-methyltransferase	Epigenetics	383	49386_86	0.09	15.91	0.22	0.36	0.00	0.01	0.03	260	106	1247	982	918
c28633_g1_i2	histone-lysine n-methyltransferase	Epigenetics	383	49386_86	0.09	15.91	0.22	0.36	0.00	0.01	0.03	260	106	1247	982	918
c33122_g1_i1	vacuolar H+ ATPase§	Proton pump	28	**37543_9**‡	0.50	47.31	0.00	0.63	0.06	0.06	0.22	987	10	643	575	44
10 matches	Intra-cellular processes§	various	26	**35404_9**‡	0.38	42.05	0.031	0.83	0.16	0	0.01	781	3	254	1049	1179


1
Identifying candidate SNPs
Bayescan
OutFLANK
+
LFMM
Environmental variables
Neutral genetic structure
2
GDM of candidate SNPs
% GDM explained
△Fst along gradient
Allelic turnover magnitude
Environmental variables
Geographic distance
3
Signatures of environmental adaptation
SNP annotation to transcriptome (BLAST2GO)
Interpretation
• Gene function
• Experimental data
• Prior knowledge
+
Environmental associations
+
Allelic turnover relationships

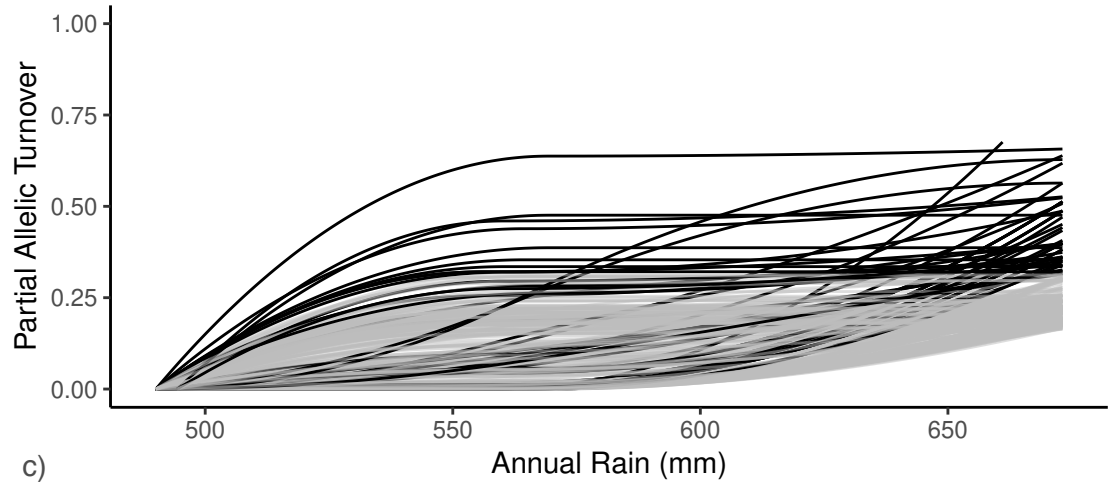

Latitude
61
60
59
58
57
56
12
14
16
18
Longitude
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
Habitat Suitability
1.0
0.8
0.6
0.4
0.2
0.0


Partial Allelic Turnover
1.00
0.75
0.50
0.25
0.00
0
2
4
6
Geographic Distance ('00 km)

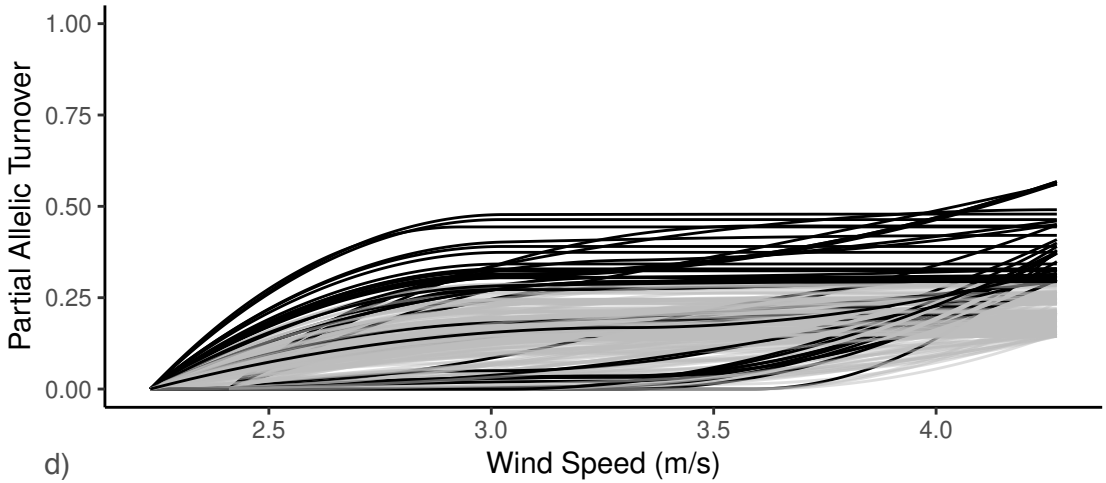
a)


Partial Allelic Turnover
1.00
0.75
0.50
0.25
0.00
19.5
20.0
20.5
21.0
21.5
22.0
Max Temperature (°C)

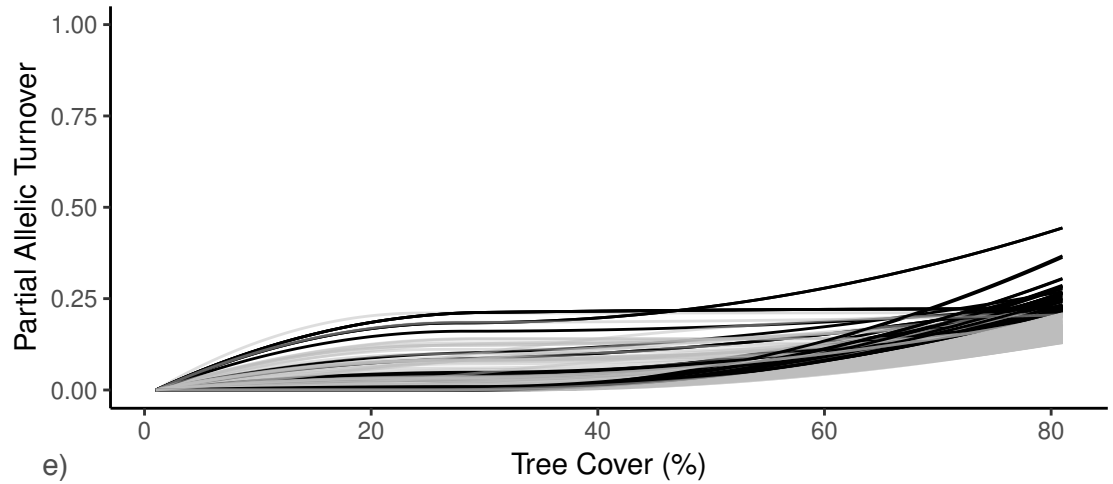
b)


Partial Allelic Turnover
1.00
0.75
0.50
0.25
0.00
500
550
600
650
Annual Rain (mm)

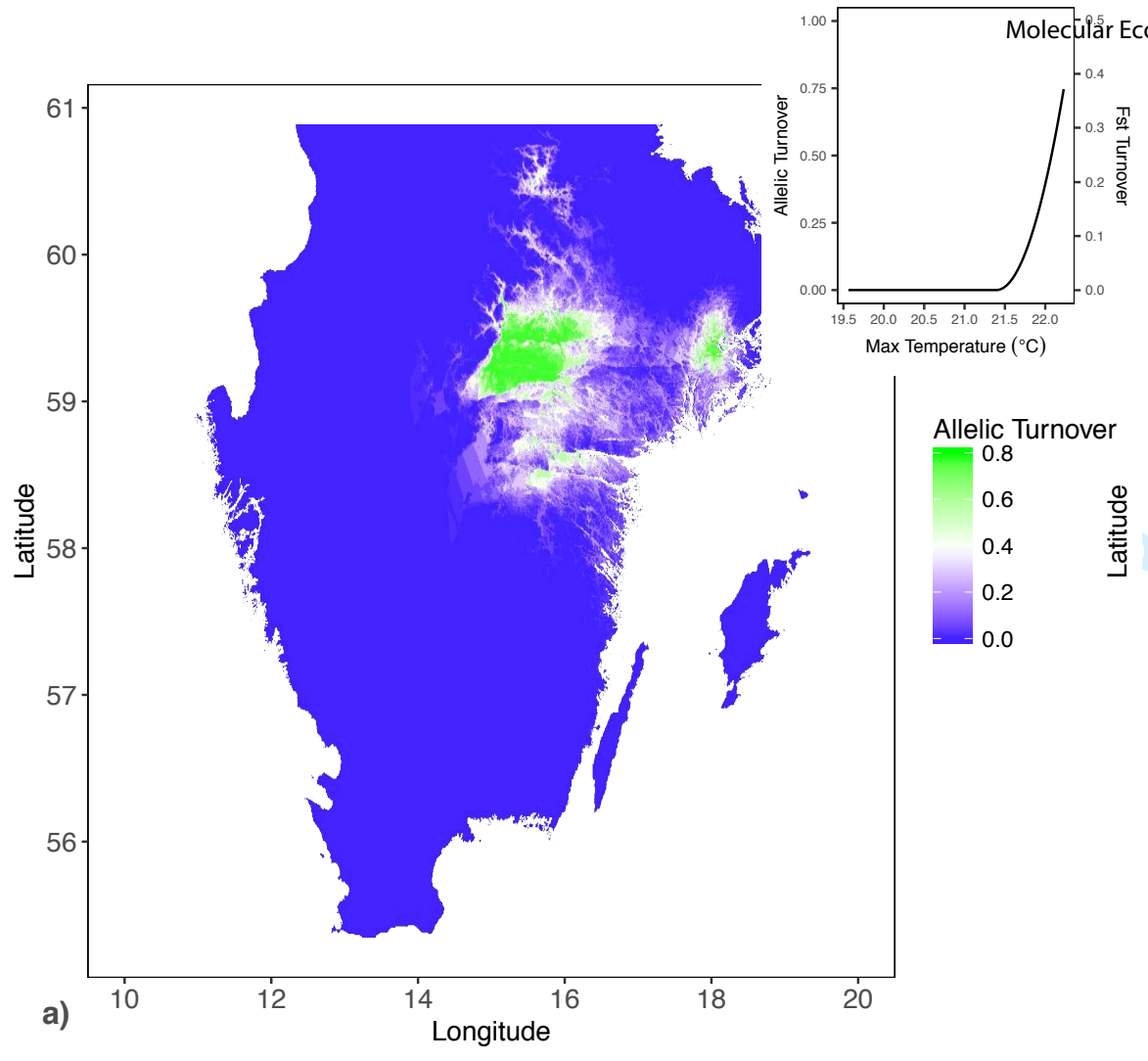
c)

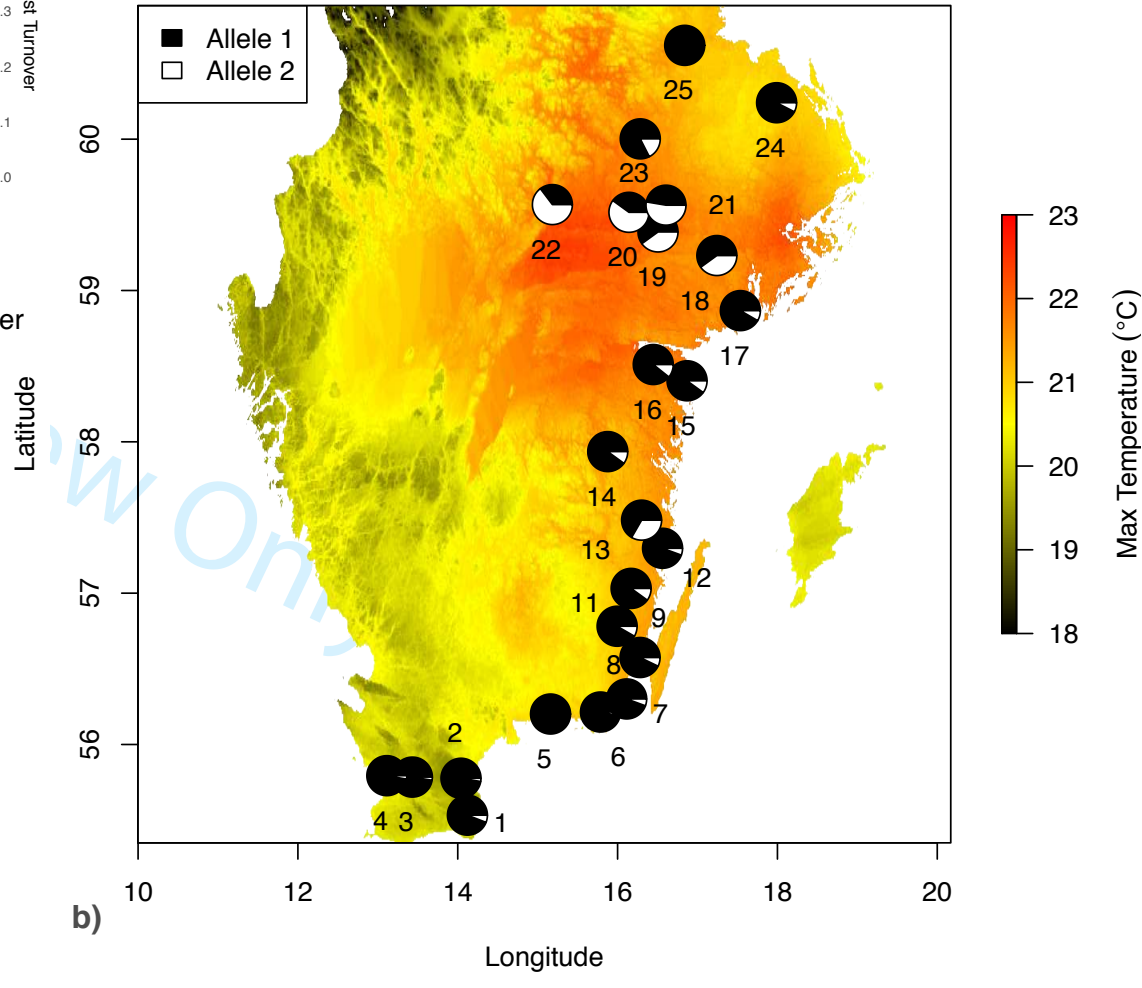

Partial Allelic Turnover
1.00
0.75
0.50
0.25
0.00
2.5
3.0
3.5
4.0
Wind Speed (m/s)

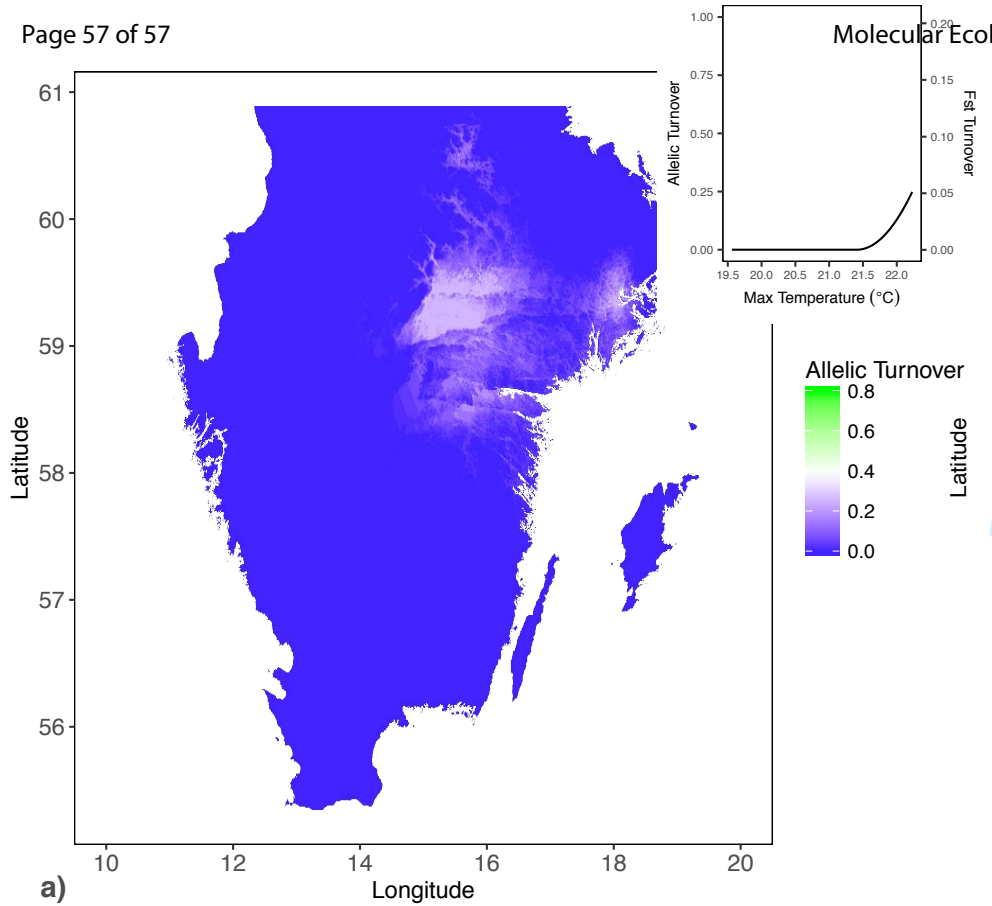
d)

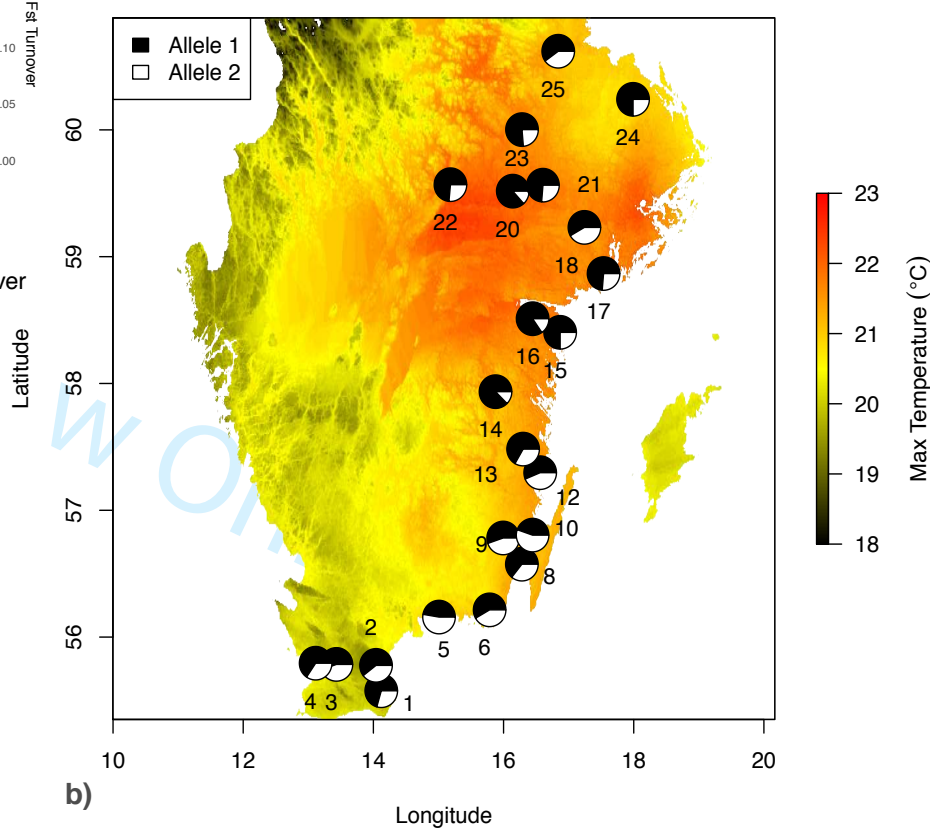

Partial Allelic Turnover
1.00
0.75
0.50
0.25
0.00
0
20
40
60
80
Tree Cover (%)

e)


Latitude
Longitude
61
60
59
58
57
56
10
12
14
16
18
20
a)
Allelic Turnover
0.8
0.6
0.4
0.2
0.0
Allelic Turnover
Fst Turnover
Max Temperature (°C)
1.00
0.75
0.50
0.25
0.00
0.5
0.4
0.3
0.2
0.1
0.0
19.5
20.0
20.5
21.0
21.5
22.0


Allele 1
Allele 2
Latitude
Longitude
60
59
58
57
56
10
12
14
16
18
20
b)
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
9
8
7
6
5
2
4
3
1
Max Temperature (°C)
23
22
21
20
19
18


Allelic Turnover
Fst Turnover
Max Temperature (°C)
Allelic Turnover
Latitude
Longitude
a)


Allele 1
Allele 2
Latitude
Longitude
Max Temperature (°C)
b)