Title: Phylogeographic differentiation versus transcriptomic adaptation to warm temperatures in *Zostera marina*, a globally important seagrass

Running title: Transcriptomic adaptation in seagrass

Jueterbock A.[1,*], Franssen S.U.[2,3], Bergmann N.[4], Gu J.[3], Coyer J.A.[5,7], Reusch T.B.H.[6], Bornberg-Bauer E.[3], Olsen J.L.[7]

1 Faculty of Biosciences and Aquaculture, Nord University, Universitetsalleen 11, 8049 Bodø, Norway

2 Institut für Populationsgenetik, Vetmeduni Vienna, Veterinärplatz 1, 1210 Vienna, Austria

3 Institute for Evolution and Biodiversity, University of Münster, Hüfferstr. 1, 48149 Münster, Germany

4 Integrated School of Ocean Sciences (ISOS), Kiel University, Leibnizstr. 3, 24098 Kiel - Germany

5 Shoals Marine Laboratory, University of New Hampshire, Durham, NH 03824, USA

6 GEOMAR Helmholtz-Centre for Ocean Research Kiel, Evolutionary Ecology of Marine Fishes, Düsternbrooker Weg 20, 24105 Kiel, Germany

7 Groningen Institute for Evolutionary Life Sciences, Ecological and Evolutionary

Genomics Group, University of Groningen, P.O. Box 11103, 9700 CC Groningen, The Netherlands

* Corresponding author: Tel.: +47 97464093, email: Alexander-Jueterbock@web.de, fax: 03212-2521383

Keywords: RNAseq, Heatwave, Global warming, Transcriptomics, Common-garden experiment, Differential expression

Abstract

Populations distributed across a broad thermal cline are instrumental in addressing adaptation to increasing temperatures under global warming. Using a space-for-time substitution design, we tested for parallel adaptation to warm-temperatures along two independent thermal clines in *Zostera marina*, the most widely distributed seagrass in the temperate northern hemisphere. A North-South pair of populations was sampled along the European and North American coasts and exposed to a simulated heat wave in a common-garden mesocosm. Transcriptomic responses under control, heat-stress and recovery were recorded in 99 RNAseq libraries with ~13,000 uniquely annotated, expressed genes. We corrected for phylogenetic differentiation among populations in order to discriminate neutral from adaptive differentiation. The two Southern populations recovered faster from heat-stress and showed parallel transcriptomic differentiation, as compared

with Northern populations. Among 2,389 differentially expressed genes, 21 exceeded neutral expectations and were likely involved in parallel adaptation to warm temperatures. However, the strongest differentiation following phylogenetic correction was between the three Atlantic populations and the Mediterranean population with 128 of 4,711 differentially expressed genes exceeding neutral expectations. Although adaptation to warm temperatures is expected to reduce sensitivity to heat waves, the continued resistance of seagrass to further anthropogenic stresses may be impaired by heat-induced down-regulation of genes related to photosynthesis, pathogen defense and stress tolerance.

Introduction

Seagrass ecosystems have experienced massive die-offs over the last decades due to increasing stresses including disease, invasive species, sediment and nutrient runoff, habitat loss through dredging and aquaculture, rising sea levels, and global warming (Orth *et al.* 2006; Waycott *et al.* 2009). Heat waves are predicted to become frequent in southern Europe and North America by 2100 (Easterling *et al.* 2000; Meehl *et al.* 2007), and are a major threat for *Zostera marina* (Reusch *et al.* 2005; Ehlers *et al.* 2008), the predominant seagrass in the northern hemisphere (Green & Short 2003; Olsen *et al.* 2004). For example, sustained temperatures of $\geq 25^{\circ}$C during the summer of 2003 increased mortality and reduced shoot density by up to about 50% in a population in the Baltic Sea (Reusch *et al.* 2005); and on both the West (W)- and East (E)-Atlantic coasts (Chesapeake Bay, VA;

Ria Formosa, PT; and the Brittany coast of France), summer temperatures now regularly reach temperatures of ≥ 25°C (Nejrup & Pedersen 2008).

Understanding geographic variation in sensitivity to increasing heat stress can help to more realistically predict climate change induced range shifts of *Z. marina* (Lavergne *et al.* 2010; Sinclair *et al.* 2010) and to identify thermally robust source populations for potential restoration (Procaccini *et al.* 2007). Common garden experiments using populations from different geographical locations employ a space-for-time design to address potential adaptation to increasing heat-stress but with the caveat that the end result reflects past evolutionary adaptation and thus cannot directly infer contemporary rates of adaptation (Kinnison & Hendry 2001; Reusch 2014)

Previous common-garden experiments with *Z. marina* revealed some evidence for local thermal adaptation of southern versus northern populations (Franssen *et al.* 2011; Winters *et al.* 2011; Franssen *et al.* 2014). Mediterranean populations are restricted to the thermally stable subtidal zone (Laugier *et al.* 1999) and, in contrast to Atlantic populations, have regularly experienced summer temperatures > 26°C over the past decade (Bergmann *et al.* 2010; Franssen *et al.* 2014). Accordingly, an Italian population (Adriatic Sea) performed better than Danish populations (Kattegat and the Baltic Sea) under common garden experiments simulating the heatwave occurring in summer 2003. Individuals from the Italian population lost fewer shoots, were less responsive in osmoprotective metabolites (Gu *et al.* 2012), and more resilient in photosynthetic performance (Franssen *et al.*

2011; Winters *et al.* 2011; Gu *et al.* 2012). Such phenotypic divergence between northern and southern populations of *Z. marina* suggests reduced sensitivity to heat waves at the species' southern edge of its distribution.

However, phenotypic divergence may have been driven by both DNA-based changes and heritable epigenetic changes. Epigenetic variations are molecular-level changes that alter gene expression, but not the underlying DNA sequence, via histone modifications, chromatin remodeling, small interfering RNAs and DNA methylation (Bossdorf *et al.* 2008). In contrast to phenotypic variation within generations, including non-heritable physiological or behavioral responses, epigenetic variation may be heritable and persist even following long-term acclimation over generations (reviewed by Hirsch *et al.* 2012; Reusch 2014). Although this may be considered a shortcoming of common-garden studies, the inclusion of epigenetic carry-over effects may provide a more holistic picture of evolutionary potential in the context of rapid environmental change as compared with studies that only assess DNA-based changes (Richards *et al.* 2012; Zhang *et al.* 2013; Kilvitis *et al.* 2014).

Modification of gene expression can also drive adaptive evolution by linking molecular heritable changes at the DNA level with fitness-relevant traits (Emerson *et al.* 2010; Wittkopp 2013). Previous common-garden experiments revealed differences in the post-heat wave recovery patterns of gene expression (termed transcriptomic resilience) between Mediterranean and Danish populations (Franssen *et al.* 2011). While the observed dif-

ferences were striking, the experimental design did not make it possible to determine whether the divergence was due to adaptive evolution, and if so, whether temperature was the major selective force as opposed to neutral processes or gene flow (reviewed in Merilä & Hendry 2014).

Methods to infer adaptive evolution of phenotypic differences include genotypic and phenotypic estimates of selection, comparison to models of neutral evolutionary change, reciprocal transplant experiments, and Q_{ST}–F_{ST} comparisons (reviewed in Merilä & Hendry 2014). In Q_{ST}-F_{ST} comparisons adaptive evolution is inferred when the phenotypic among-population divergence (Q_{ST}) exceeds among-population divergence at neutral genetic markers (F_{ST}) (reviewed in Leinonen *et al.* 2013). Q_{ST}–F_{ST} comparisons correct for phylogeographic differentiation and recently have been revised to a multivariate method (Leinonen *et al.* 2013) that more accurately discriminates neutral from adaptive divergence (Ovaskainen *et al.* 2011). Only when transcriptomic differences are correlated with temperature differences across replicate, independent locations can adaptive differentiation be attributable to temperature as the selective force (Kawecki & Ebert 2004; Merilä & Hendry 2014). We refer to adaptive differentiation as only that portion of transcriptomic differentiation that exceeds neutral phylogeographic differentiation across multiple populations from contrasting thermal environments.

Here, we test the hypothesis that *Z. marina* shows adaptive differentiation in gene expression between thermally contrasting environments replicated on the North-American

and European coasts. More specifically: 1) Is there evidence for adaptation to temperature, as judged by heritable transcriptomic differentiation that exceeds neutral phylogenetic differentiation between the two Southern and the two Northern populations? 2) Do Southern populations show gene expression patterns consistent with reduced sensitivity to heat waves as evidenced by faster recovery from heat stress?

Methods

Sampling

Individuals of *Z. marina* were sampled in April 2010 from northern (N) and southern (S) populations in Europe (Doverodde, NW Denmark 56° 43.070' N 008° 28.446' E, hereafter NE); Gabicce Mare, NE Italy 43° 57.970' N 12° 45.860' E, hereafter SE) and in the Northeastern USA (Great Bay, NH 43° 3.868' N, 70° 52.345' W, hereafter NU; Waquoit, MA 41° 33.240' N, 70° 30.650' W, hereafter SU) (Figure 1a). Note that the sampling site South USA (SU) does not represent the south of the USA but the southernmost of our US samples. The coastal region encompassing the N and S site of North America is characterised by one of the steepest latitudinal thermal gradients in the world's oceans (Fig 2b in Frank *et al.* 2007; Wahle *et al.* 2013). Thus, even though the geographic distance between the North American sites is much less than between the N and S European sites, the differences in summer temperatures are comparable (Figure 1b). Variability in water

temperatures at the sampling locations was based on daily average sea surface temperature values recorded during summer months (June 1st to September 30th in years 2002-2011) over the past decade (Figure 1b). Temperature data was extracted for sites NU, NE, and SE from the `NOAA_OI_SST_V2 dataset` (0.25° resolution, described in (Reynolds *et al.* 2002), provided by NOAA/OAR/ESRL/ PSD, Boulder, Colorado, USA on http://www.esrl.noaa.gov/psd/). For site SU, that was not covered by the `NOAA_OI_SST_V2 dataset`, we extracted temperature data from the National Estuarine Research Reserve System (http://cdmo.baruch.sc.edu/, station Sage Lot). Three to four clones with $\geq$ 3 shoots/clone were sampled from each of 10 patches at each location with a $\geq$ 5m distance between samples to minimize chances of collecting the same clone (genotype) twice. Genotypic uniqueness of each experimental ramet was confirmed by genotyping the samples on an ABI 3100 Capillary Sequencer at four microsatellite loci (GenBank Accession numbers: AJ009898, AJ009900, AJ249305, AJ249307, Reusch 2000).

Experimental design

Within 48h after collection, the plants were transported in seawater-filled cooling boxes to the AQUATRON, a mesocosm facility at the University of Münster, Germany. Details of the AQUATRON facility are described in Winters *et al.* (2011) and Figure 1e,f. Briefly, each of two temperature- controlled water circuits supplied artificial seawater (31 psu) from a storage tank to six 700 L tanks (101cm x 120cm x 86.5cm). Similar water chemistry

between the two circuits was ensured with a water exchange rate of 1200 L h^{-1}. Each tank was populated with ~50 periwinkles (*Littorina littorea*) to regulate epiphytic growth. Each tank contained eight boxes (two boxes (37 cm x 27 cm) per population) with 10 genotypes. Shoots were planted in 10 cm natural sediment (collected from Falckenstein, DE in the Western Baltic Sea: 54° 24.367' N, 010° 11.438' E).

Plants were acclimated for 20 days to equilibrate temperature and light conditions (~400 µmol photons s^{-1} m^{-2}) in order to minimize non-heritable differences in gene expression (Hoffmann *et al.* 2005; Whitehead & Crawford 2006). After 20d the temperature was raised 0.5°C day $^{-1}$ to 19°C, the experimental control temperature in six of the 12 tanks over the entire experiment.

Heat wave simulation

After 20 days of acclimation at 19°C, the temperature was raised in six of the 12 tanks at 1°C day^{-1} to 25.5°C, then held constant for 20 days to simulate the heat wave that occurred in the Baltic Sea during the summer of 2003 (Reusch *et al.* 2005). Finally, the temperature was decreased 1°C day^{-1} to 19°C and subsequently held for 20 days to allow the plants to recover (Figure 1f).

RNA extraction

Samples for RNAseq (2cm long leaf tips) were excised from each ramet (3 genotypes per treatment per population) at two time points under acute heat-stress (T2 and T3: 0 and 5 days at 25.5°C) and at three time points under recovery (T5, T7, and T9: 1, 20, and 30 days at 19°C) (Figure 1f, and Figure S1). Tissue samples were immediately frozen in liquid nitrogen.

RNA extraction was performed using the InviTrap Spin Plant RNA Mini Kit (Stratek Molecular) following the manufacturers protocol. We used the provided RP buffer for lysis. RNA concentrations and purity were tested by Nanodrop® measurement (ND-1000, peQLab). RNA integrity was checked with an automated electrophoresis station Experion (Bio-Rad), using StdRNA chips and reagents (Bio-RAD). RNA concentrations ranged between 23-182 ng/µl, RQI values were >7.2.

RNAseq

Library preparation proceeded with DNase 1 digestion of total RNA, mRNA isolation by use of oligo(dT) beads, mRNA fragmentation, first and second strand cDNA synthesis, end-repair, A-tailing, bar-coded adapter ligation and PCR amplification. Sequencing libraries were checked using an Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany) before sequencing.

Single-end (1 x 100 bp) RNA Sequencing (RNASeq) data were generated using standard Illumina protocols and kits (TruSeq SBS KIT-HS v3, FC-401-3001; TruSeq SR Cluster Kit v3-cBot-HS, GD-401-3001) and all sequencing was performed using the HiSeq 2000 platforms (University of Groningen Genome Analysis Facility).

Quality trimming and control

TruSeq adapters were trimmed (at a 10% error rate with cutadapt version 1.4.1, (Martin 2011)) before bases of low quality (Phred score Q<20, 99% base call accuracy) and reads of short length (<35bp) were removed with the FastqMcf filter in ea-utils (Aronesty 2011) (see Table S1 for numbers of reads before and after quality trimming). Quality controls of read base content, length distribution, duplication and over-representation were checked with FastQC http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Library NU3W2 was excluded from further analyses due to its exceptionally low number of reads (357 raw reads, compared to >10 Mio. reads in most other libraries, Table S1).

Mapping

We aligned reads of each library to the genome of *Z. marina* (Olsen *et al.* 2016) with the splice-aware RNA-seq aligner STAR (sjdbOverhang 100) (Dobin *et al.* 2013), guided by splice junctions from the v2.1 *Z. marina* genome annotation (gff3; Gen-

Bank Accession: LFYR00000000; see Table S1 for characteristics of each library). Alignments that contained non- canonical splice junctions were filtered out. Duplicate reads were removed with the MarkDuplicates program from the Picard package (http://broadinstitute.github.io/picard/). Ambiguously mapped reads (ca. 5% in each library), defined by values >1 in the NH:i BAM file tag were removed. NH is the number of reported alignments that contains the query in the current record. For each library we counted the reads that mapped uniquely to annotated mRNAs (exons) with the htseq-count script of the HTSeq python package (Anders *et al.* 2015) (20554 exons in total, see Table S2). Reads of low expression (library average <5) were removed to avoid potential artifacts from sequencing errors and reads of highly variable expression (standard deviation over all libraries > library average) were removed to reduce the effect of outlier individuals on statistical comparisons. In total, 12948 exons remained after filtering (Table S3).

Annotation

Mapped sequence IDs (mRNAIDs) were associated to gene IDs, gene descriptions and Gene Ontology-labels, by parsing the gff3 file of the annotated *Z marina* genome (v2.1, nuclear and organellar, GenBank Accession: LFYR00000000) from the ORCAE-database (Sterck *et al.* 2012). These gene-annotations rely on inference from homology. Where gene descriptions were lacking in the gff3 file, they were transferred from functional descriptions

of the top BLAST hit to *Z. marina* proteins (Table S4).

Population differentiation based on neutral SNPs

Neutral differentiation among the four populations was estimated from neutral SNPs (single nucleotide polymorphisms). To call SNP variants from the RNAseq data, all aligned reads were merged with samtools (Li *et al.* 2009) before applying GATK (McKenna *et al.* 2010) splitting of exon segments, reassignment of mapping qualities (SplitNCigarReads), and realignment around indels (RealignerTargetCreator and IndelRealigner). The realigned reads were demultiplexed (samtools) before calling sequence variants with GATK (HaplotypeCaller). After filtering (VariantFiltration) according to the GATK Best Practices guide for RNAseq data (http://gatkforums.broadinstitute.org/gatk/discussion/3891/calling-variants-in-rnaseq), 159,592 nuclear variants (indels and SNPs) were kept.

Variants with non-neutral divergence between the four populations were identified with the Bayesian likelihood method that is implemented in the program Bayescan v2.1 (Foll & Gaggiotti 2008). The program uses differences in allele frequencies between populations to screen for non-neutral Fst outlier loci at a false discovery rate of 0.05. Bayescan approximates allele frequencies in a neutrally structured population with a multinomial-Dirichlet model. Selection is introduced by decomposing Fst coefficients into a population-specifc component (beta) shared by all loci, and a locus-specific component (alpha) shared

by all populations using a logistic regression. This method infers posterior probabilities of each locus to be under the effect of selection by defining and comparing two alternative models (neutral vs selection).

Nine SNPs of the 159,592 nuclear variants were identified as outlier loci with non-neutral divergence between the four populations (q-values $< 5\%$ for the model including selection). Scaffold:locus IDs of the non-neutral variants: 1:43222, 2:501606, 2:1527432, 42:390694, 143:142100, 100:118173, 188:180877, 188:180966, 253:10530. After removing the 9 outlier loci (SNPs), 20092 indels, and 170 SNPs with >2 alleles, a total of 139,321 (159592-9-20092-170) biallelic neutral SNPs (Dataset S1) were kept.

Population differentiation between the four populations was calculated from the set of 139,321 biallelic neutral SNPs as Wright's F_{ST}, estimated according to Weir & Cockerham (1984), and as Nei's genetic distance (Nei 1972) with the R package 'StAMPP' (Pembleton *et al.* 2013). A neighbor Joining (NJ) tree of Nei's genetic distances was created and tested with 1000 bootstrap replications using the R package 'ape' (Paradis *et al.* 2004).

Multivariate clustering of gene expression

Overall transcriptomic differentiation, shaped by both neutral drift and potential selection, was characterized by clustering the samples hierarchically by the first five principle components of gene expression, averaged over technical replicates, with the Principle Com-

ponent Analysis (*PCA*) and Hierarchical Clustering on Principle Components (*HCPC*) functions in the R package 'FactoMineR' (Lê *et al.* 2008) (setting scale.unit=FALSE not to scale the expression values to unit variance). To account for differences in sequencing depth and to assure homoscedasticity before PCA, the raw count values of mapped reads were regularized-log transformed with the function *rlog* in the R package 'DESeq2' (Love *et al.* 2014). Overall transcriptomic differentiation was characterized under control and heat stress by creating one hierarchical cluster for all control samples on the expression of all genes, and one cluster for all samples on the expression of heat-responsive genes: genes that were differentially expressed between controls and heated samples under acute stress (time points 2 and 3), or in the recovery phase (time points 5, 7 and 9) (Figure S1). The first five principle components explained 76.35% of the variation in the expression of all genes in control samples and 66.96% of the variation in the expression of heat-responsive genes in all samples. Groupings of samples (NU, SU, and NE (=Atlantic) vs SE (=Mediterranean), NU and NE (=North) vs SU and SE (=South) , and Controls vs Stressed and Recovery samples) were tested with analysis of similarity (ANOSIM) in the R package 'vegan' v2.3–1 (Oksanen *et al.* 2015).

Differential expression

We identified differences in gene expression between thermal regimes (North vs South) by testing for differential expression in three groups of samples: 1) control, constantly

kept at 19°C at all time points; 2) acute stress, >25°C at time points 2 and 3; and 3) recovery, previously warmed to >25°C, but then allowed to recover at 19°C and sampled at time points 5, 7 and 9 (Figure S1). As the transcription profiles clearly separated the Mediterranean library (SE) from all of the Atlantic libraries (NU, SU, and NE) see hierarchical clusters in Figure 2a,b), we also tested for differential expression between Atlantic vs Mediterranean libraries. Time point was specified as additional explanatory factor to oceans (Atlantic and Mediterranean) or isotherm levels (North and South) in the differential expression tests performed with the R/Bioconductor package 'DESeq2'. The DESeq 2 model corrected internally for library size differences (Love *et al.* 2014). Significance levels of all test results were adjusted with the Benjamini and Hochberg correction (Benjamini & Hochberg 1995), using the *p.adjust* function in R (R Development Core Team, 2014), to control for the false discovery rate in multiple pairwise comparisons. Expression was deemed significantly different for genes with corrected p-values below 0.05.

Acute heat-stress response and recovery

The acute heat-stress response was determined as differential gene expression between controls and acutely stressed samples (time points 2 and 3). Recovery was determined as the differential between control expression and post heat wave expression (time points 5, 7, and 9). Differential expression analyses were performed with the R/Bioconductor package 'DESeq2' (Love *et al.* 2014) (Figure S1), which internally corrected the raw count data

of mapped reads for library size differences. The acute stress and the recovery responses were identified in the libraries from all populations with 'population' and 'time point' as additional explanatory factors. The acute stress response and the recovery were also identified in each population separately. While 'time point' was used as an additional explanatory factor to test for recovery in each population, only samples from time point 3 were used to identify the acute heat stress response for the NU and SU populations, as time point 2 samples were unavailable. Significance levels in all test results were adjusted with the Benjamini and Hochberg correction (Benjamini & Hochberg 1995), using the *p.adjust* function in R (R Development Core Team, 2014), to control for the false discovery rate in multiple pairwise comparisons. Expression was deemed significantly heat-responsive under acute heat stress or in the recovery phase for genes with corrected p-values below 0.05.

Enrichment tests of both gene ontologies (GO) Molecular Function (MF) and Biological Process (BP) were performed with the R package 'topGO' (Alexa & Rahnenfuhrer 2010). GO terms were obtained from the v2.1 *Zostera* genome annotation from the ORCAE database (Sterck *et al.* 2012) (GenBank Accession: LFYR00000000). We used Fisher's exact tests to test for enrichments in genes that were heat-responsive (significantly up- or down-regulated in stressed vs control samples) under acute heat-stress, or in the recovery phase. To reduce redundancy in the significantly enriched GO-terms (p-values <0.05) we calculated 'sim Rel' scores (Schlicker *et al.* 2006), based on the *Arabidopsis thaliana* GO-

term database, with the REVIGO web server (Supek *et al.* 2011). The GO terms were reduced to cluster-representatives by removing terms with semantic similarities >0.5.

Previous studies identified six important ontology groups in the transcriptomic heat stress response of *Z. marina*: 1) cell wall fortification (Franssen *et al.* 2014); 2) protein folding and chaperone activity (Franssen *et al.* 2011); 3) ribosome activity (Franssen *et al.* 2014); 4) oxidation-reduction processes (Gu *et al.* 2012); 5) electron transport and photosynthesis (Gu *et al.* 2012); and 6) osmoprotective metabolites (Street *et al.* 2010; Gu *et al.* 2012). To estimate the representation of these ontology groups in the acute heat response and the heat response that lasted throughout the recovery phase of the present study, we identified exact matches and semantic similarities (sim Rel scores <0.5 (Schlicker *et al.* 2006), using REVIGO (Supek *et al.* 2011)) between each GO term in the enriched MFs/BPs (up- or down-regulated under acute heat-stress or recovery) and each GO-term in the six targeted ontology groups (Table S5a-f).

Adaptive differentiation in gene expression

To identify signals of possible selection, we searched for genes for which the identified differential expression (North vs South or Atlantic vs Mediterranean, under control conditions, acute stress, or recovery) could not be explained by phylogenetic distance and genetic drift alone (Figure S1). This was done using the approach of Ovaskainen *et al.* (2011) in the R package 'driftsel' 2.1.2 (Karhunen *et al.* 2013). Adaptive differentiation

under natural selection was inferred for those genes that showed significant differential expression following phylogenetic correction under a neutral model. This was done as follows: A matrix of population-to-population coancestry coefficients (probabilities that randomly chosen alleles for a neutral locus are identical by descent between individuals belonging to different populations) was constructed from the set of neutral biallelic SNPs with the *do.all* function in the 'RAFM' R package (Karhunen & Helsinki 2013), and used as a prior to estimate the posterior distributions of all parameters with a Metropolis-Hastings Monte Carlo Markov Chain (MCMC) algorithm (*MH* function); as required to test for neutrality with the *H.test* function in the R package 'driftsel' (Karhunen *et al.* 2013). All Monte Carlo Markov chains converged after 3,000 iterations when the Gelman-Rubin shrink factor, tracked with the R package 'coda' (Plummer *et al.* 2006), remained close to 1. Thus, we ran a total of 6,000 iterations without thinning, and discarded the first 3,000 iterations as burn-in.

We used Fisher's exact tests to test for enriched GO-terms of Molecular Functions (MFs) or Biological Processes (BPs) in adaptively differentiated genes (H value > 0.95) with the R package 'topGO' (Alexa & Rahnenfuhrer 2010). GO-terms were based on the v2.1 *Zostera* genome annotation (Sterck *et al.* 2012) (GenBank Accession: LFYR00000000). To reduce redundancy in significantly enriched GO-terms (p-values <0.05 after Benjamini and Hochberg correction (Benjamini & Hochberg 1995) with the *p.adjust* function in R (R Core Team 2015)), we calculated 'sim Rel' scores (Schlicker *et al.* 2006), based on

the *Arabidopsis thaliana* GO-term database, with the REVIGO web server (Supek *et al.* 2011). Enriched GO-terms were replaced by GO terms of cluster-representatives with semantic similarities >0.5.

Coding sequence differences in temperature-adaptive genes

Twenty-one genes exhibited adaptive differentiation in gene expression exceeding neutral differentiation (H value > 0.95) between Northern and Southern populations, and were likely involved in the parallel adaptation of seagrass populations to warm temperatures. To test for adaptive coding sequence differences in addition to adaptive differential expression for these 21 genes, we tested for ratios of nonsynonymous to synonymous substitutions (dN/dS) exceeding 1. First, we determined the genomic consensus sequence for each population by applying population-specific SNPs to the reference genome (Olsen *et al.* 2016) with bcftools consensus (https://github.com/samtools/bcftools). Population-specific SNPs were called with GATK (McKenna *et al.* 2010) HaplotypeCaller and VariantFiltration from merged bam files that combined alignments of all samples from the same population. For each population we limited the consensus sequence to the 21 target genes with bedtools getfasta (Quinlan & Hall 2010) based on the genomic features file (gff) of the *Zostera* genome (Olsen *et al.* 2016). For each target gene, codon alignments of all population sequences were obtained with pal2nal (Suyama *et al.* 2006) that was guided by mafft (Katoh & Standley 2013) multiple sequence alignments mafft of peptides

predicted with TransDecoder (http://transdecoder.github.io/) based on homology to the known protein sequences.

In order to test for sites under positive selection in the 21 adaptively differentially expressed genes between the Southern and Northern populations, we performed branch-site tests by contrasting CodeML model A (relaxation, dN/dS unequal 1) to model A1 (positive selection, dN/dS >1) of the package PAML (Yang 2007) using ETE 3 (Huerta-Cepas *et al.* 2016).

Results

Population differentiation based on neutral SNPs

Neutral genetic differentiation among the four populations was quantified with F_{ST} values and visualized with a NJ tree. Pairwise F_{ST} values ranged from 0.25 to 0.56 (all statistically significant, p <0.05, Figure 1c). The NJ tree (Figure 1d) supports strong differentiation between European and American coasts, as well as between Northern and Southern populations along each coast. Notably, the Mediterranean population (SE) was the most distant from the three Atlantic populations: a common pattern associated with virtually all phylogeographic studies including seagrasses (Olsen *et al.* 2004).

Multivariate clustering of gene expression

Overall transcriptomic differentiation (shaped by both neutral drift and potential selection) was characterized in hierarchical clusters of gene expression with and without the impact of heat stress. Based on the expression of all mapped genes (12948, after filtering out genes of low or highly variable expression, Table S3), the control samples separated into a Mediterranean (SE) and an Atlantic cluster (NU, SU and NE) (Figure 2a). This grouping of libraries was supported by analysis of similarity (R = 0.28, p = 0.05). Differences in overall gene expression, thus, were in accordance with the phylogeographic divergence between the populations, represented by the Neighbor-Joining tree in Figure 1d. In other words, a grouping of Northern and Southern samples in the expression of all genes was not supported by ANOSIM, R = 0.10, p = 0.16, Figure 2a).

The heat stressed samples (w, time points 2 and 3) showed a distinct expression in heat-responsive genes (4979) from the controls (c, all time points) and from the recovery samples (w, time points 5, 7, and 9) (Figure 2b). The grouping of control and recovery samples in a separate cluster from the stressed samples was supported by ANOSIM (R = 0.87, p=0.001). Atlantic and Mediterranean samples separated clearly in the control-recovery cluster (grouping supported by ANOSIM, R = 0.25 p = 0.01), but not in the acute stress cluster (grouping not supported by ANOSIM R = 0.36, p = 0.2), which was due to the outlier library NU3w (Figure 2b). A grouping of Northern and Southern samples in the expression of heat-responsive genes was not supported by ANOSIM (R =

388 0.1553, p =0.06, Figure 2b).

389 Differential expression

390 Differences in gene expression between thermal regimes (North vs South) and between
391 oceans (Atlantic vs Mediterranean) were identified by differential expression analysis. In
392 each of the comparisons, the lowest number of differentially expressed genes was recorded
393 during the acute stress phase (Table 1); the highest number of differentially expressed
394 genes was recorded in the control samples (Table 1). The overlap of differentially expressed
395 genes with heat-responsive genes is shown in Figure S2a-d. Differentially expressed genes
396 are listed for the Atlantic vs Mediterranean comparison in Table S6a-c, and for the North
397 vs South comparison in Table S6d-f.

398 Acute heat-stress response

399 The acute heat-stress response was tested as differential gene expression between controls
400 and acutely stressed samples. NU was the only population without acute stress response.
401 In contrast, the SU population responded at 734 genes (Table S7e), and the European
402 populations responded at >1,800 genes (NE: 1814, Table S7c; SE: 2004, Table S7d). Thus,
403 the Southern samples were not less responsive to acute heat stress than the Northern
404 samples.

A total of 4907 genes responded concordantly between all four populations to acute heat stress (Table S7a), based on significant differential expression between all controls and all acute stress libraries independent from population. In the acute heat stress response, 32 Molecular Functions (MFs, Table S8a, represented genes in Table S9a) and 46 Biological Processes (BPs, Table S8e, represented genes in Table S9e) were enriched in the 1612 up-regulated genes (genes with log2 fold change > 0 in Table S7a). Dominant upregulated processes and functions, represented by > 490 genes ($>10\%$ of all 4908 heat-responsive genes), included 'cellular processes', 'metabolic processes', and 'binding' (Table S8a,e). Some 38 MFs (Table S8b, represented genes in Table S9b) and 41 BPs (Table S8f, represented genes in Table S9f) were enriched in the 2395 down-regulated genes (genes with log2 fold change <0 in Table S7a). Dominant downregulated functions and processes, represented by > 490 genes ($>10\%$ of all 4908 heat-responsive genes), included 'cellular processes' and 'catalytic activity' (Tables S8b,f). All six BPs and MFs that were previously identified to be dominant in the heat stress response of *Z. marina* {Gu2012,Street2010,Franssen2011,Franssen2014} (Table S5a-f) were also represented (semantic similarities of GO terms > 0.5) in enriched heat-responsive BPs and MFs in the present study: 'Heatstress', 'Metabolism', 'Cellwall', 'Photosynthesis', 'Ribosomal', and 'Oxidative.reductive' (Figures S3a,b).

Recovery

Recovery was tested as differential gene expression between controls and recovery samples. The number of heat-responsive genes in the recovery phase was an order of magnitude lower in the Southern samples (SU: 6, Table S7i; SE: 10, Table S7g) as compared with the Northern samples (NU: 302, Table S7h; NE: 205, Table S7f). Given that the Southern samples were not less heat responsive than the N samples (see above), this means that the Southern samples recovered faster from heat stress.

In total, 123 genes responded concordantly between all four populations during the recovery phase (Table S7b). Under recovery, 12 MFs (Table S8c, represented genes in Table S9c) and 10 BPs (Table S8g, represented genes in Table S9c) were enriched in the 53 up-regulated genes (genes with log2 fold change > 0 in Table S7b), while 14 MFs (Table S8d, represented genes in Table S9d) and 4 BPs (Table S8h, represented genes in Table S9h) were enriched in the 70 down-regulated genes (genes with log2 fold change < 0 in Table S7b).

Adaptive differentiation in gene expression

We applied a phylogeographic correction to eliminate differences due to neutral processes as opposed to those due to selection. Populations were partitioned in two ways: 1) Atlantic vs Mediterranean, and 2) North vs South. Atlantic and Mediterranean samples

displayed the strongest adaptive signal in differential gene expression. In total 128 of 4,711 differentially expressed genes showed greater differential expression (74 under control and 106 under recovery conditions, Figure S2a,b, Table S10a-c) than expected under neutral phylogenetic divergence (H value > 0.95), implying adaptation to the environmental covaries with a p-value < 0.05 (Figure 3a). Northern and Southern populations exhibited adaptive differentiation exceeding neutral differentiation (H value > 0.95) in 21 of 2,389 differentially expressed genes (3 under control and 18 under recovery conditions, Figure S2c,d, Table S11a-c). None of these 21 genes showed adaptive coding sequence differences (p-value > 0.05 for dN/dS > 1) between Northern and Southern samples.

Sixteen genes exhibited adaptive differentiation in both comparisons, Atlantic vs Mediterranean, and North vs South (Figure 3a,b); gene IDs based on the *Zostera marina* genome annotation v2.1, GenBank Accession LFYR00000000: Zosma5g01430, Zosma5g01440, Zosma55g00720, Zosma57g00700, Zosma68g00400, Zosma722g00030, Zosma98g00300, Zosma124g00200, Zosma21g00340, Zosma290g00070, Zosma107g00010, Zosma40g00060, Zosma425g00160, Zosma89g00800, Zosma190g00070, Zosma253g00020). None of them was adaptively differentiated under acute heat-stress due to increased variation in gene expression (larger standard errors) compared to control- or recovery-conditions (Figure 3b). Most of the 16 genes were lower expressed in Mediterranean and Southern populations compared with Atlantic and Northern populations (Figure 3b). Thus, much of the North vs South separation was explained by the separation between Mediterranean and

Atlantic samples.

Discussion

Genes that are putatively adaptive to contrasting temperatures

Correction of differential gene expression for neutral phylogeographic differentiation enabled us to extract only the putatively adaptive portion of transcriptomic differentation. We inferred contrasting temperatures as the major selective force when the putatively adaptive differences were correlated with temperature differences across two independent thermal clines.

The global transcriptomic differentiation (shaped by neutral phylogenetic differentiation and adaptive divergence) did not place Northern and Southern samples into different clusters, either under control conditions (Figure 2a) or in response to heat stress (Figure 2b). Nevertheless, for 21 genes (where the expression difference between Northern and Southern samples was greater than can be explained by phylogenetic differentiation, ca. 1% of all 2389 differentially expressed genes), adaptation by natural selection was the most parsimonious explanation. The remaining variation in these genes is most likely explained by parallel adaptation to contrasting habitat temperatures along both the American and European thermal clines. The absence of adaptive coding sequence differences ($dN/dS < 1$)

suggests that the adaptive expression difference between Northern and Southern samples in these 21 genes can be ascribed to either trans-acting regulation factors or to cis-acting elements outside the coding sequences, altering gene expression regulation.

Although putatively adaptive to contrasting habitat temperatures, these 21 genes may not directly affect acute-stress tolerance but instead, play a role under control or recovery conditions. This is because an increased among-sample variability in gene expression may have erased any adaptive differentiation under acute heat stress (Figure 3b). Validation would require experimental determination of the phenotype and fitness of *Z. marina* under non-stressful conditions and under recovery from heat-stress (Barrett & Hoekstra 2011; Pardo-Diaz *et al.* 2014).

Twenty-one genes are likely a conservative representation of genes involved in adaptation to contrasting temperatures and might be extended by at least some of the genes that showed adaptive differentiation between Atlantic and Mediterranean samples. For example, 128 genes (2.8% of all 4,711 differentially expressed genes) showed differential expression that could not be accounted for by neutral genetic distance in the strong transcriptomic separation between the Atlantic and Mediterranean samples (Figure 2a,b). Additionally, two factors suggest that habitat temperature played a predominant role: 1) 76% of the genes suggesting adaptive differentiation in response to habitat temperature (16 of 21) were also adaptively differentiated between Mediterranean and Atlantic samples; and 2) In all of these 16 genes the directionality of differential expression agreed

between Southern and Mediterranean samples: Under recovery, ten genes that were lower expressed in the Southern samples were also lower expressed in the Mediterranean samples and six genes that were higher expressed in the Southern samples were also higher expressed in the Mediterranean samples (Figure 3b). However, the identification of genes that most likely responded to contrasting temperatures between the Mediterranean and Atlantic requires confirmation by association studies including at least one additional Mediterranean population.

The strong adaptive transcriptomic differentiation of the Mediterranean from the Atlantic samples suggests that the North vs South differentiation of *Z. marina* must be stronger on the European coast than on the US Atlantic coast, and that much of the previously observed North vs South differentiation along the European coast (NE vs SE) (Bergmann *et al.* 2010; Franssen *et al.* 2011; Winters *et al.* 2011; Gu *et al.* 2012) might be better explained by a general Mediterranean-Atlantic (SE vs NE, NU, and SU) differentiation. The strong European North-South differentiation is likely due to high rates of genetic drift in Mediterranean populations which are small, isolated, and have relatively low genetic variation (Olsen *et al.* 2004; Procaccini *et al.* 2007). Moreover, stronger North-South differentiation along the European coast is likely due to reduced gene-flow (Olsen *et al.* 2004) favoring adaptive differentiation (Davis & Shaw 2001; Aitken *et al.* 2008). In contrast, on the US Atlantic coast, ongoing trans-Arctic gene flow from the E-Pacific may prevent local adaptation to warm temperatures in the South (Olsen *et al.* 2004). Taken

together, the present study shows the strength of comparing several independent environmental clines when addressing adaptation vs neutral differentiation in gene expression patterns.

Local thermal adaptation in expression patterns after the heat-stress

Previous common-garden experiments suggested that local thermal adaptation of European Southern versus Northern populations in *Z. marina* was driven by faster recovery of gene expression to normal patterns after imposing a heat wave (Franssen *et al.* 2011). Our study confirmed that the same putatively adaptive differences in gene expression evolved in parallel along the US Atlantic coast. The finding of Franssen *et al.* (2011) that gene-expression patterns during recovery reveal thermal adaptation better than expression patterns under acute-stress was replicated on the American and European coast. Across all four populations we found that plants recovered within one day: the gene expression of early-recovery samples (taken at time point 5, one day after return to 19°C) was indistinguishable from control samples and long-recovery samples (time points 7 and 9, 20 and 30 days after return to 19°C, Figure 2b). However, the extent to which populations returned to control-levels of gene expression was influenced by the North-South affiliation: the Southern populations expressed <20 genes differently from control-levels during recovery (Table S7g,i), whereas the Northern populations expressed >200 genes differently

from control-levels (Table S7f,h). Thus, our results show that increased stress-resilience of Southern seagrass samples does not only apply along the European (Franssen *et al.* 2011; Winters *et al.* 2011), but also along the North American thermal cline, suggesting reduced sensitivity to heat waves at the species' southern (warm) edge of distribution.

Response to acute heat-stress

Stress, as measured by the number of upregulated genes, was comparable between Northern and Southern populations (NU: 0; SU: 734, Table S7e; NE: 860, Table S7c; and SE: 466, Table S7d) and differential gene expression between all four populations was lowest during the acute stress phase (Table 1), suggesting that *Z. marina* relies on common pathways to alleviate heat-stress. This supports the previous work (Franssen *et al.* 2011) demonstrating that gene expression was not dependent on the North-South affiliation. The lack of response to acute heat stress in the American Northern population (NU) is peculiar. We know that there was a heat-stress response, since it was detected during recovery conditions (Table S7h). However, the lack of a detectable response during acute stress might be an artifact as it is supported by a single library (all of the other acute-stress NU libraries failed, Table S1) that has a transcription profile differing markedly from the other acute-stress libraries (library NU3W in Figure 2b).

Upregulation of genes involved in metabolism and cell-wall synthesis most likely tempered the heat-stress. In addition to the osmoprotective metabolites that were identified

as an important part of the heat-response in both *Z. marina* and *Z. noltii* (Gu *et al.* 2012), the present study found other metabolic-related genes that are known to alleviate heat-stress. For example, 'starch synthase' (Zosma22g01480, represented in starch binding: GO:2001070, Figure S3b) increased heat-tolerance of wheat grains (*Triticum aestivum*) (Sumesh *et al.* 2008), and 'glycosylation' (GO:0070085, the posttranslational attachment of carbohydrate residues to proteins, Figure S3a) has been shown to enhance chaperone activity (Henle *et al.* 1998) and induced heat-shock protein synthesis in a slime mold (Murakami-Murofushi *et al.* 1997). Furthermore, the upregulated function 'purine ribonucleoside binding' (GO:0032550, Figure S3b) involved 284 genes, including several stress-alleviating protein kinases (Table S9a).

Our results support the hypothesis of Franssen *et al.* (2014), that cell-wall fortification may protect *Z. marina* from heat-stress. Increased cell-wall synthesis under acute stress was represented by the process 'cellular component biogenesis' (GO:0044085, Figure S3a) Cell-wall strengthening most likely continued after acute stress, as the target function 'Cellwall' was represented in upregulated molecular functions during recovery (Figure S3b).

Down-regulation of genes related to photosynthesis and pathogen defense, suggests that heat-stress undermined the resistance of *Z. marina* to additional stress. Photosynthesis is the most heat-sensitive function in green plants (Berry & Bjorkman 1980; Weis & Berry 1988; Havaux & Tardy 1996). In our study, stress-induced photo inhibition (involving

reduced carbon fixation, oxygen evolution, and electron flow) was indicated by down-regulated processes, such as 'photosynthetic electron transport chain' (GO:0009767) and 'photosynthesis' (GO:0015979) (Figure S3a).

Pathogen defense may have been impaired by heat-stress induced downregulation of: 1) 'cytidine deamination' (GO:0009972, Figure S3a) and 'cytidine deaminase activity' (GO:0004126, Figure S3b), which play important roles in the antiviral immune response through the mutagenic RNA-editing activity of cytidine deaminase (Martın *et al.* 2014). Rising temperatures enhance disease effects on eelgrass growth (Bull *et al.* 2012) and inhibit the chemical pathogen defense of eelgrass (Vergeer *et al.* 1995; Vergeer & Develi 1997). Rising temperatures, therefore, may indirectly increase the risk of an epidemic outbreak of the "wasting disease" (Rasmussen 1977), which is caused by the protist *Labyrinthula zosterae* (Bockelmann *et al.* 2013), and triggered extensive seagrass die offs in the 1930s and 1980's in temperate and tropical regions of the northern hemisphere (reviewed in Orth *et al.* 2006; Bishop 2013).

Resistance of *Z. marina* to additional anthropogenic stresses may be impaired by heat-stress induced down-regulation of: 1) 'arginine decarboxylase' (Zosma1g02550 in 'cellular catabolic process' GO:0044248, Figure S3a), which was also downregulated in rice (*Oryza sativa*) with reduced tolerance to salinity-stress (Chattopadhyay *et al.* 1997); 2) 'alpha,alpha-trehalose-phosphate synthase (UDP-forming) activity' (GO:0003825, down-regulated in the recovery-phase, Figure S3b), as well associated with reduced stress-

tolerance in rice (*O. sativa*) (Li *et al.* 2011); and 3) many 'ras-related proteins' (in 'GTPase activity' GO:0003924, Figure S3b) that are involved in numerous aspects of cell growth control (McCormick 1995).

To conclude, the stress measured by the number of upregulated genes did not differ between Southern and Northern populations. The common stress response involved up-regulation of genes involved in metabolism and cell-wall synthesis, likely dampening the heat-stress. Downregulation of genes related to photosynthesis and pathogen defense suggested that heat-stress undermines the resistance of *Z. marina* to additional stress. *Zostera marina* has dominated the North Atlantic through several previous glacial-interglacial periods. Temperature alone is not the driver, but rather numerous other anthropogenic stressors press towards a tipping point.

Acknowledgements

We acknowledge Chris Weidmann and Mary Kay Fox from Waquoit Bay National Estuariae Research Reserve for permissions and assistance with sample collection from Waquoit Bay; Fred Short from Jackson Estuarine Laboratory (UNH) for permissions and collections in Great Bay, NH, as well as links to USDA for applications; Holly Bayley, Art Mathieson and Dave Shay for field and laboratory assistance at JEL; Al Hansen for assistance with sammple collection from both US sites; the Shoals Marine Laboratory, for providing housing for JLO and JAC during the US field work; Philipp Schubert and Anneli Ehlers, GEOMAR/Reusch Lab, for assistance with sample collection from Denmark and Italy, respectively; Karsten Zecher from the Institute for Evolution and Biodiversity (University of Münster), Georg Plenge and Wilfried Niemann from the University of Münster/Bornberg-Bauer Lab, and Regina Klapper from GEOMAR/Reusch Lab for help with the mesocosm experiments and labwork. We are grateful to the editors Dr. Karen Chambers and Dr. Mark Ungerer as well as the two anonymous reviewers whose suggestions and comments improved the quality and clarity of the article.

This study was funded by: The Netherlands Organisation for Scientific Research, Earth and Life Sciences (NWO-ALW), Project Nr. 819.01.002 to JLO and JAC (2009-2014); the Volkswagenstiftung Foundation, Evolutionary Biology Initiative for support of SUF; the Alexander von Humboldt Foundation for the support of JG; the Priority Programme AQUASHIFT from the DFG, Grant RE 1109 to TBHR.

References

Aitken SN, Yeaman S, Holliday Ja, Wang T, Curtis-McLane S (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, **1**, 95–111.

Alexa A, Rahnenfuhrer J (2010) *topGO: enrichment analysis for gene ontology.*

Anders S, Pyl PT, Huber W (2015) HTSeq–A Python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**, 166–169.

Aronesty E (2011) *ea-utils : "Command-line tools for processing biological sequencing data".*

Barrett RDH, Hoekstra HE (2011) Molecular spandrels: tests of adaptation at the genetic level. *Nature Reviews Genetics*, **12**, 767–780.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, **57**, 289–300.

Bergmann N, Winters G, Rauch G *et al.* (2010) Population-specificity of heat stress gene induction in northern and southern eelgrass *Zostera marina* populations under simulated global warming. *Molecular Ecology*, **19**, 2870–2883.

Berry J, Bjorkman O (1980) Photosynthetic Response and Adaptation to Temperature in Higher Plants. *Annual Review of Plant Physiology*, **31**, 491–543.

Bishop ND (2013) The effects of multiple abiotic stressors on the susceptibility of the seagrass *Thalassia testudinum* to *Labyrinthula* sp., the causative agent of wasting disease. MA thesis. University of North Florida.

Bockelmann AC, Tams V, Ploog J, Schubert PR, Reusch TBH (2013) Quantitative PCR reveals strong spatial and temporal variation of the wasting disease pathogen, *Labyrinthula zostera*e in northern European eelgrass (*Zostera marina*) beds. *PLoS ONE*, **8**, e62169.

Bossdorf O, Richards CL, Pigliucci M (2008) Epigenetics for ecologists. *Ecology letters*, **11**, 106–115.

Bull JC, Kenyon EJ, Cook KJ (2012) Wasting disease regulates long-term population dynamics in a threatened seagrass. *Oecologia*, **169**, 135–142.

Chattopadhyay M, Gupta S, Sengupta D, Ghosh B (1997) Expression of arginine decarboxylase in seedlings of indica rice (*Oryza sativa* L.) cultivars as affected by salinity stress. *Plant Molecular Biology*, **34**, 477–483.

Davis MB, Shaw RG (2001) Range shifts and adaptive responses to quaternary climate change. *Science*, **292**, 673–679.

Dobin A, Davis CA, Schlesinger F *et al.* (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.

Easterling DR, Meehl GA, Parmesan C *et al.* (2000) Climate extremes: observations, modeling, and impacts. *Science*, **289**, 2068–2074.

Ehlers A, Worm B, Reusch TBH (2008) Importance of genetic diversity in eelgrass *Zostera marina* for its resilience to global warming. *Marine Ecology Progress Series*, **355**, 1–7.

Emerson KJ, Merz CR, Catchen JM *et al.* (2010) Resolving postglacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Sciences*, **107**, 16196–16200.

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**, 977–993.

Frank KT, Petrie B, Shackell NL (2007) The ups and downs of trophic control in continental shelf ecosystems. *Trends in Ecology & Evolution*, **22**, 236–242.

Franssen SU, Gu J, Bergmann N *et al.* (2011) Transcriptomic resilience to global warming in the seagrass *Zostera marina*, a marine foundation species. *Proceedings of the National Academy of Sciences*, **108**, 19276–19281.

Franssen SU, Gu J, Winters G *et al.* (2014) Genome-wide transcriptomic responses of the seagrasses *Zostera marina* and *Nanozostera noltii* under a simulated heatwave confirm functional types. *Marine Genomics*, **15**, 65–73.

Green EP, Short FT (2003) World Atlas of Seagrasses. University of California Press, 298.

Gu J, Weber K, Klemp E *et al.* (2012) Identifying core features of adaptive metabolic mechanisms for chronic heat stress attenuation contributing to systems robustness. *Integrative Biology*, **4**, 480–493.

Havaux M, Tardy F (1996) Temperature-dependent adjustment of the thermal stability of photosystem II in vivo: possible involvement of xanthophyll-cycle pigments. *Planta*, **198**, 324–333.

Henle KJ, Jethmalani SM, Nagle WA (1998) Stress proteins and glycoproteins (Review). *International Journal of Molecular Medicine*, **1**, 25–57.

Hirsch S, Baumberger R, Grossniklaus U (2012) Epigenetic variation, inheritance, and selection in plant populations. *Cold Spring Harbor symposia on quantitative biology*, **77**, 97–104.

Hoffmann AA, Shirriffs J, Scott M (2005) Relative importance of plastic vs genetic factors in adaptive differentiation: geographical variation for stress resistance in *Drosophila melanogaster* from eastern Australia. *Functional Ecology*, **19**, 222–227.

Huerta-Cepas J, Serra F, Bork P (2016) ETE 3: Reconstruction, analysis and visualization of phylogenomic data. *Molecular Biology and Evolution*, msw046.

Karhunen M, Merilä J, Leinonen T, Cano JM, Ovaskainen O (2013) driftsel: an R package for detecting signals of natural selection in quantitative traits. *Molecular Ecology Resources*, **13**, 746–754.

Karhunen M, Helsinki U (2013) RAFM: Admixture F-model. *Molecular eEology Resources*, **13**, 746–754.

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, **30**, 772–80.

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

Kilvitis HJ, Alvarez M, Foust CM *et al.* (2014) Ecological Epigenetics. In: *Ecological Genomics: Ecology and the Evolution of Genes and Genomes*. Ed. by CR Landry, N Aubin-Horth. **781**. Advances in Experimental Medicine and Biology. Springer Netherlands, 191–210.

Kinnison MT, Hendry AP (2001) The pace of modern life II: from rates of contemporary microevolution to pattern and process. *Genetica*, **112**, 145–164.

Laugier T, Rigollet V, Casabianca ML de (1999) Seasonal dynamics in mixed eelgrass beds, *Zostera marina* L. and *Z. noltii* Hornem., in a Mediterranean coastal lagoon (Thau lagoon, France). *Aquatic Botany*, **63**, 51–69.

Lavergne S, Mouquet N, Thuiller W, Ronce O (2010) Biodiversity and climate change: integrating evolutionary and ecological responses of species and communities. *Annual Review of Ecology Evolution, and Systematics*, **41**, 321–350.

Lê S, Josse J, Husson F (2008) FactoMineR : An R package for multivariate analysis. *Journal of Statistical Software*, **25**, 1–18.

Leinonen T, McCairns RJS, O'Hara RB, Merilä J (2013) Q(ST)-F(ST) comparisons: evolutionary and ecological insights from genomic heterogeneity. *Nature reviews. Genetics*, **14**, 179–90.

Li HW, Zang BS, Deng XW, Wang XP (2011) Overexpression of the trehalose-6-phosphate synthase gene OsTPS1 enhances abiotic stress tolerance in rice. *Planta*, **234**, 1007–1018.

Li H, Handsaker B, Wysoker A *et al.* (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)*, **25**, 2078–9.

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

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, **17**, 10–12.

Martın S, Cuevas JM, Grande-Pérez A, Elena SF (2014) A putative antiviral role of plant cytidine deaminases. *bioRxiv*, 5256.

McCormick F (1995) Ras-related proteins in signal transduction and growth control. *Molecular Reproduction and Development*, **42**, 500–506.

McKenna A, Hanna M, Banks E *et al.* (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*, **20**, 1297–1303.

Meehl Ga, Stocker TF, Collins WD *et al.* (2007) Global Climate Projections. In: *Climate Change 2007: The Physical Science Basis.Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by S Solomon, D Qin, M Manning *et al.* Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. Chap. Global Cli, 750.

Merilä J, Hendry AP (2014) Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evolutionary applications*, **7**, 1–14.

Murakami-Murofushi K, Nishikawa K, Hirakawa E, Murofushi H (1997) Heat stress induces a glycosylation of membrane sterol in myxoamoebae of a true slime mold, *Physarum polycephalum*. *Journal of Biological Chemistry*, **272**, 486–489.
Nei M (1972) Genetic distance between populations. *American Naturalist*, **106**, 283–292.
Nejrup LB, Pedersen MF (2008) Effects of salinity and water temperature on the ecological performance of *Zostera marina*. **88**, 239–246.
Oksanen J, Blanchet FG, Kindt R *et al.* (2015) *vegan: Community Ecology Package*.
Olsen JL, Rouzé P, Verhelst B *et al.* (2016) The genome of the seagrass *Zostera marina* reveals angiosperm adaptation to the sea. *Nature*, **530**, 331–335.
Olsen JL, Stam WT, Coyer JA *et al.* (2004) North Atlantic phylogeography and large-scale population differentiation of the seagrass *Zostera marina* L. *Molecular Ecology*, **13**, 1923–1941.
Orth RJ, Carruthers TIMJB, Dennison WC *et al.* (2006) A global crisis for seagrass ecosystems. *BioScience*, **56**, 987–996.
Ovaskainen O, Karhunen M, Zheng C, Arias JMC, Merilä J (2011) A new method to uncover signatures of divergent and stabilizing selection in quantitative traits. *Genetics*, **189**, 621–632.
Paradis E, Claude J, Strimmer K (2004) APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, **20**, 289–290.
Pardo-Diaz C, Salazar C, Jiggins CD (2014) Towards the identification of the loci of adaptive evolution. *Methods in Ecology and Evolution*, **6**, 445–464.
Pembleton LW, Cogan NOI, Forster JW (2013) StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular ecology resources*, **13**, 946–52.
Plummer M, Best N, Cowles K, Vines K (2006) CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News*, **6**, 7–11.
Procaccini G, Olsen JL, Reusch TBH (2007) Contribution of genetics and genomics to seagrass biology and conservation. *Journal of Experimental Marine Biology and Ecology*, **350**, 234–259.
Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)*, **26**, 841–2.
R Core Team (2015) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.
Rasmussen E (1977) The wasting disease of eelgrass (*Zostera marina*) and its effects on environmental factors and fauna. In: *Seagrass Ecosystems: A Scientific Perspective*. Ed. by CP McRoy, C Helfferich. Marine science. New York: M. Dekker, 1–52.
Reusch TBH (2000) Five microsatellite loci in eelgrass *Zostera marina* and a test of cross-species amplification in *Z. noltii* and *Z. japonica*. *Molecular Ecology*, **9**, 371–373.
Reusch TBH (2014) Climate change in the oceans: evolutionary versus phenotypically plastic responses of marine animals and plants. *Evolutionary Applications*, **7**, 104–122.

Reusch TBH, Ehlers A, Hammerli A, Worm B (2005) Ecosystem recovery after climatic extremes enhanced by genotypic diversity. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 2826–2831.

Reynolds RW, Rayner NA, Smith TM, Stokes DC, Wang W (2002) An improved *in situ* and satellite SST analysis for climate. *Journal of Climate*, **15**, 1609–1625.

Richards CL, Schrey AW, Pigliucci M (2012) Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation. *Ecology letters*, **15**, 1016–1025.

Schlicker A, Domingues FS, Rahnenführer J, Lengauer T (2006) A new measure for functional similarity of gene products based on Gene Ontology. *BMC Bioinformatics*, **7**, 302.

Sinclair SJ, White MD, Newell GR (2010) How useful are species distribution models for managing biodiversity under future climates. *Ecology and Society*, **15**, 8.

Sterck L, Billiau K, Abeel T, Rouzé P, Van de Peer Y (2012) ORCAE: online resource for community annotation of eukaryotes. *Nature Methods*, **9**, 1041.

Street TO, Krukenberg KA, Rosgen J, Bolen DW, Agard DA (2010) Osmolyte-induced conformational changes in the Hsp90 molecular chaperone. *Protein Science*, **19**, 57–65.

Sumesh KV, Sharma-Natu P, Ghildiyal MC (2008) Starch synthase activity and heat shock protein in relation to thermal tolerance of developing wheat grains. *Biologia Plantarum*, **52**, 749–753.

Supek F, Bošnjak M, Škunca N, Šmuc T (2011) REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE*, **6**, e21800.

Suyama M, Torrents D, Bork P (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic acids research*, **34**, W609–12.

Vergeer LHT, Aarts TL, Groot JD de (1995) The 'wasting disease' and the effect of abiotic factors (light intensity, temperature, salinity) and infection with *Labyrinthula zosterae* on the phenolic content of *Zostera marina* shoots. *Aquatic Botany*, **52**, 35–44.

Vergeer LHT, Develi A (1997) Phenolic acids in healthy and infected leaves of *Zostera marina* and their growth-limiting properties towards *Labyrinthula zosterae*. *Aquatic Botany*, **58**, 65–72.

Wahle R, Brown C, Hovel K (2013) The geography and body-size dependence of top-down forcing in New England's lobster-groundfish interaction. *Bulletin of Marine Science*, **89**, 189–212.

Waycott M, Duarte CM, Carruthers TJB *et al.* (2009) Accelerating loss of seagrasses across the globe threatens coastal ecosystems. *Proceedings of the National Academy of Sciences*, **106**, 12377–12381.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Weis E, Berry JA (1988) Plants and high temperature stress. In: *Symposia of the Society for Experimental Biology*. **42**, 329–346.

Whitehead A, Crawford DL (2006) Neutral and adaptive variation in gene expression. *Proceedings of the National Academy of Sciences*, **103**, 5425–5430.

Winters G, Nelle P, Fricke B, Rauch G, Reusch TBH (2011) Effects of a simulated heat wave on photophysiology and gene expression of high- and low-latitude populations of *Zostera marina*. *Marine Ecology Progress Series*, **435**, 83–95.

Wittkopp PJ (2013) *Evolution of Gene Expression*.

Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Molecular biology and evolution*, **24**, 1586–91.

Zhang YY, Fischer M, Colot V, Bossdorf O (2013) Epigenetic variation creates potential for evolution of plant phenotypic plasticity. *The New phytologist*, **197**, 314–322.

Data Accessibility

- Supplementary tables (Table S1-S11, Excel format) and dataset S1 (vcf file) are archived at Dryad: Temporary reviewer URL: http://datadryad.org/review?doi=doi:10.5061/dryad.vf5fk. Permanent URL that is inaccessible while the manuscript is under review: http://dx.doi.org/10.5061/dryad.vf5fk.
  - **Table S1 (Excel)**: cDNA library characteristics of all 108 cDNA libraries.
  - **Table S2 (Excel)**: Numbers of mapped reads.
  - **Table S3 (Excel)**: Regularized log-transformed expression values.
  - **Table S4 (Excel)**: Annotations of mapped reads.
  - **Table S5 (Excel)**: Targeted GO-terms.
  - **Table S6 (Excel)**: Differential expression.
  - **Table S7 (Excel)**: Genes responding to heat stress.
  - **Table S8 (Excel)**: Enriched functions and processes under acute heat-stress and in the recovery.
  - **Table S9 (Excel)**: Heat-responsive genes respresenting enriched functions and processes.
  - **Table S10 (Excel)**: Adaptively differentiated genes between Atlantic and Mediterranean samples.
  - **Table S11 (Excel)**: Adaptively differentiated genes between Northern and Southern samples.
  - **Dataset S1 (vcf file)**: Biallelic neutral SNPs.
- Supplementary figures (Figure S1-S3) are uploaded as Supporting Information in a single PDF file.
  - **Figure S1**: Workflow of data analysis with color codes representing groupings of samples/libraries.
  - **Figure S2**: Venn diagrams showing the overlap of heat-responsive genes under different conditions.
  - **Figure S3**: Heatmaps.
- DNA raw reads, the assembled genome sequence and annotation are accessible from NCBI under BioProject number PRJNA41721 with accession number LFYR00000000. http://www.ncbi.nlm.nih.gov/bioproject/?LFYR00000000. Further information on the *Zostera marina* project is available via the Online Resource for Community Annotation Eukaryotes (ORCA) at http://bioinformatics.psb.ugent.be/orcae/

- RNAseq libraries will be made accessible on NCBI under BioProject number PRJNA302837.

Author contributions

JLO led the study. TBH, JAC, SUF, and EBB were actively involved in project planning and experimental design. JLO, JAC, NB, SUF and JG collected the samples. JAC, SUF, NB and JG performed the heat-stress experiments and the lab work. AJ analyzed the data with contributions from SUF, and wrote the manuscript. All co-authors read and commented on the manuscript.

Tables

Table 1: Number of differentially expressed genes between groups of samples (A: Atlantic, M: Mediterranean, N: Northern, S: Southern) under control, stress, and recovery conditions.

Comparison	Control	Stress	Recovery
A vs M	3264	575	3309
N vs S	1679	154	988

Figures

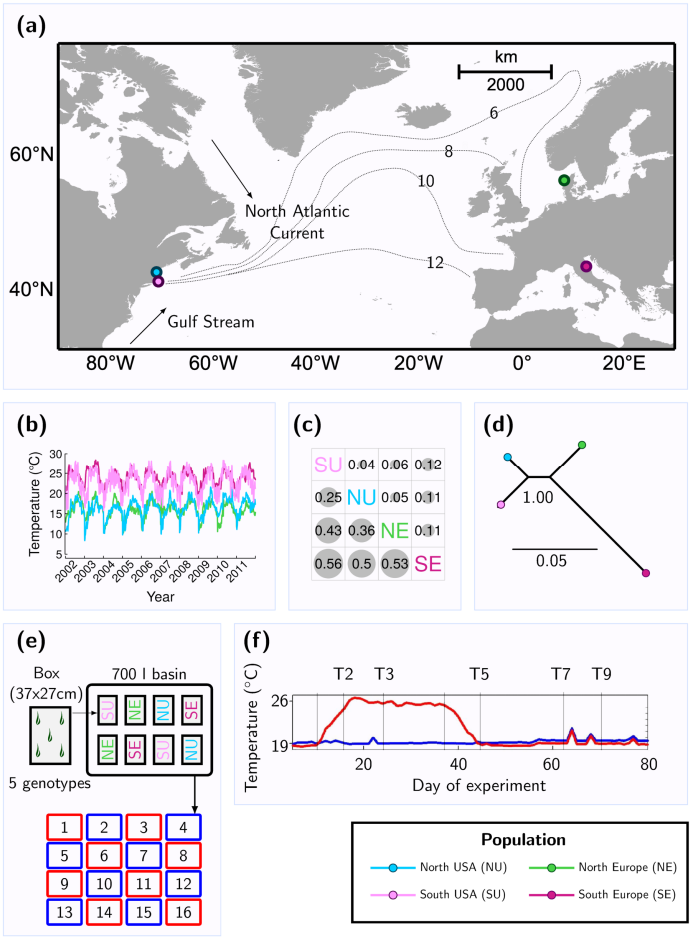

Figure 1: Habitat and experimental temperatures and neutral genetic differentiation between populations. **(a)** American and European sampling sites with overview of summer sea surface temperature (SST) isotherms (dotted contour lines, redrawn from CLIMAP 1984). The North Atlantic Current and the Gulf Stream compress the isotherms along the American coast but spread them out along the European coast. **(b)** Daily average SST during summer months (June 1st to September 30th) from 2002 to 2011 at the four sampling sites. Winter temperatures are not shown because our study focuses on heat stress adaptation in the face of warm summer temperatures. **(c)** Matrix of pairwise Nei's genetic distances (upper right) and F_{ST}-values (lower left) measuring the genetic differentiation among the four populations (all values were significantly larger than 0 at $p<0.05$). **(d)** Neighbour-Joining tree based on Nei's genetic distances derived from 139,321 biallelic neutral SNPs. All branches had a 100% bootstrap support. **(e)** Schematic design of the common-garden mesocosm with six replicate tanks for each of two temperature treatments (blue=control and red=stress). **(f)** Temperature profiles. After 4 weeks of acclimation, six tanks were warmed to ca. 25.5°C (red) for 3 weeks, while six control tanks remained at 19°C. Samples for RNAseq were taken at two time points under heat-stress (T2 and T3: 0 and 5 days at 25.5°C) and at three time points under recovery (T5, T7, and T9: 1, 20, and 30 days at 19°C).

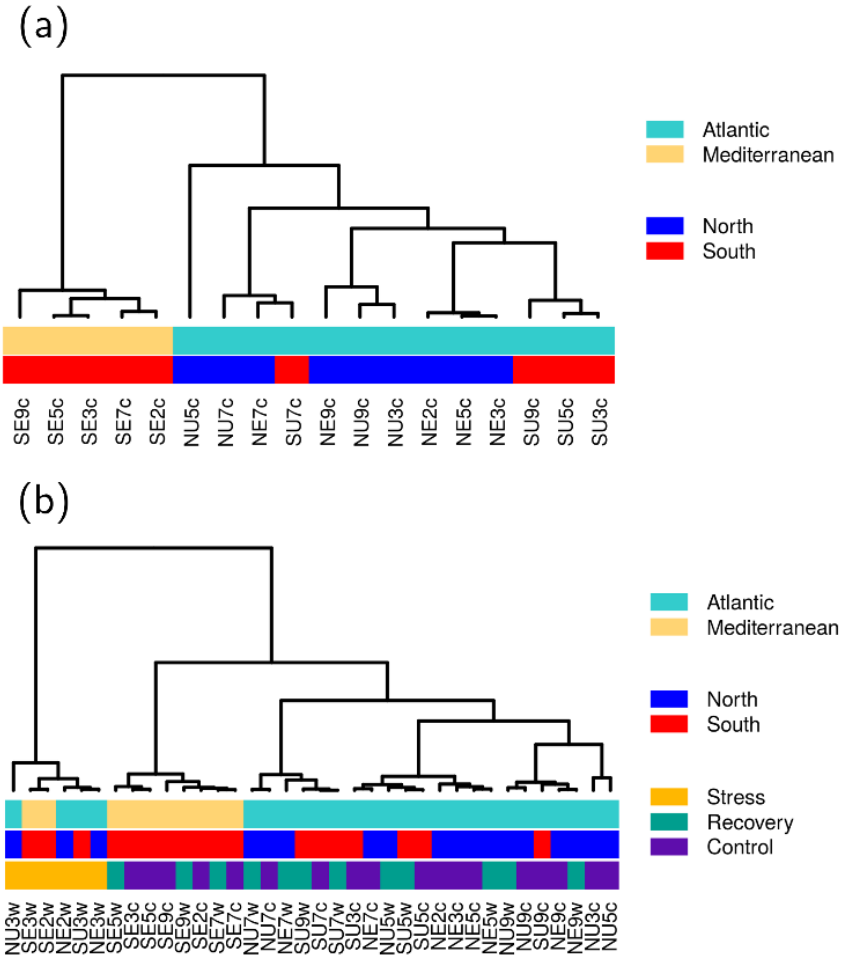

Figure 2: Hierarchical clusters based on the first five principle components of gene expression. **(a)** Cluster based on all 12,948 genes that mapped to the control samples and were filtered for low or highly-variable expression (Table S3). **(b)** Cluster of all samples based on 4,979 genes that were heat-responsive under acute stress or in the recovery phase. NE: Northern Europe, SE: Southern Europe, NU: Northern US, SU: Southern US; c: control samples, w: stressed and recovery samples. Numbers indicate sampling time points.

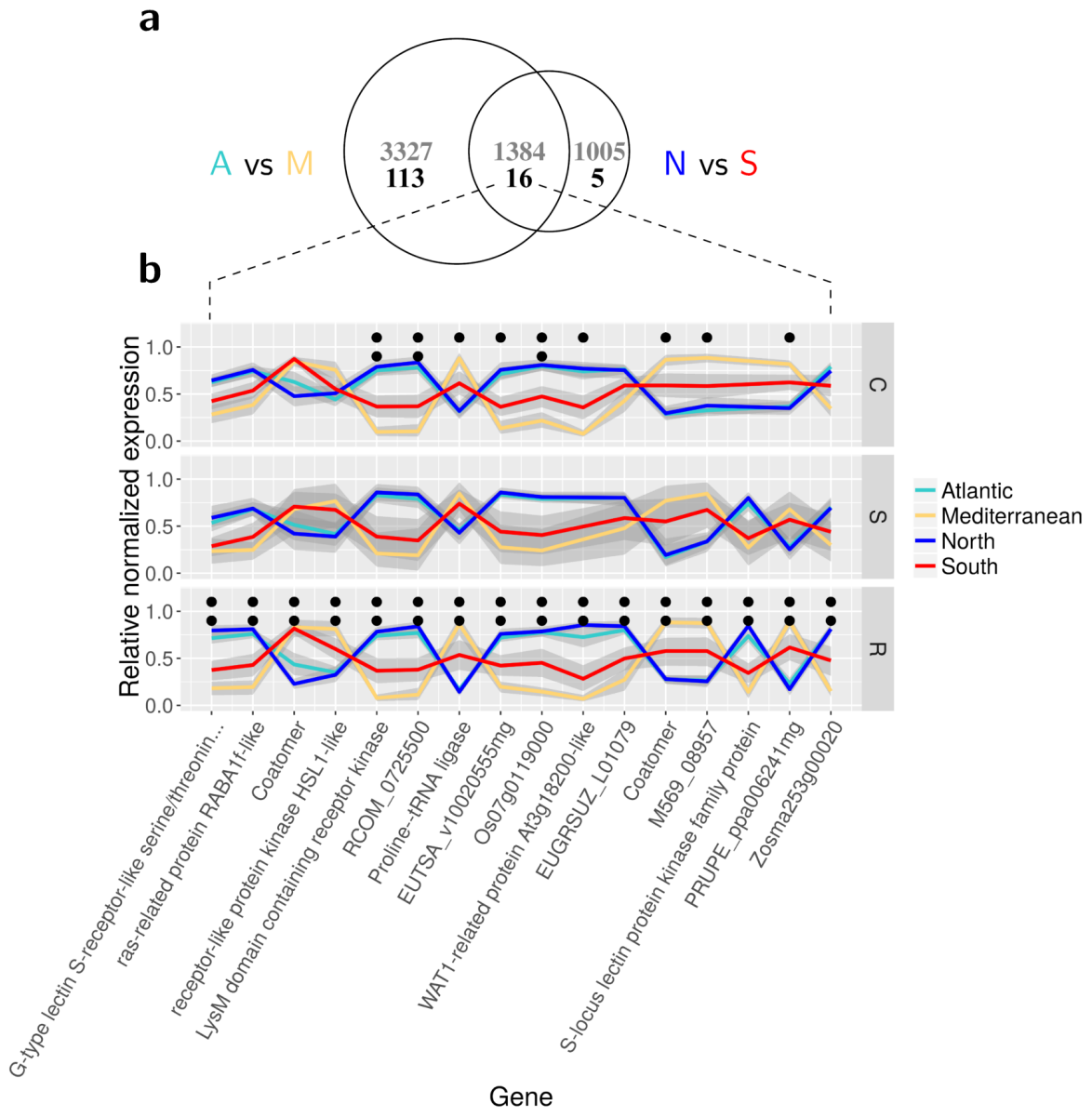

Figure 3: Adaptively differentiated genes. The venn diagram above shows the overlap of genes that were differentially expressed (gray numbers) or adaptively differentiated (black numbers) between Atlantic (A) and Mediterranean (M) samples with those genes that were differentially expressed between Northern (N) and Southern (S) samples. The parallel coordinates plot below shows those 16 genes that were adaptively differentiated in both A vs M and N*vs* S comparisons, and thus, were putatively adaptive to contrasting temperatures. Colored lines show average normalized gene expression (0-1: minimum to maximum individual expression) and shaded areas represent standard errors. Black dots indicate if the genes were adaptively differentiated (upper dot: A vs M, lower dot: N vs S) under control (C), stress (S), and/or recovery (R) conditions.

Supplementary Files

Supplementary Figures

All supplementary figures are combined in a single PDF file: SupplementaryFigures.pdf

Figure S1: **Workflow of data analysis with color codes representing groupings of samples/libraries.**

Figure S2: **Venn diagrams showing the overlap of heat-responsive genes under different conditions.** Overlap of heat-responsive-genes (top red circles) under stress (**a**, **c**) and recovery (**b**, **d**) conditions with genes that were differentially expressed (white numbers, bottom circles) between Atlantic (A) and Mediterranean (M) samples (**a**, **b**) or between Northern (N) and Southern (S) samples (**c**, **d**). Black numbers represent genes that showed stronger expression differences than expected by phylogenetic divergence (H value > 0.95); these imply adaptation to different environments.

Figure S3: **Heatmaps**. Biological processes **(a)** and molecular functions **(b)** that were significantly enriched ($p \leq 0.05$) under acute stress and/or recovery (GO terms listed in Table S8) and that matched heat-responsive processes/functions in previous studies (sim Rel scores >0.5 (see Methods), indicated by blue to green color codes). The significance of the enrichment is color coded (white: no enrichment, red: strongest enrichment ($-log10(p-value)$) in upregulated genes, blue: strongest enrichment ($log10(p-value)$) in downregulated genes). See Table S9 for genes that represent the upregulated/downregulated processes.

Supplementary Tables

Table S1 (Excel): cDNA library characteristics of all 108 cDNA libraries. Sample preparation failed for eight libraries (indicated in the second column).

Table S2 (Excel): Numbers of mapped reads. Number of reads that mapped to each of 20554 exons (rows), listed separately for each library (columns E-CX); ZosmaID (column A): gene ID based on the *Zostera marina* genome annotation v2.1 (GenBank Accession: LFYR00000000); mRNAID (column B): the mapped sequence ID; Source (column C): the source of annotation (inference from homology or top BLAST hit); Description (column D): the gene description.

Table S3 (Excel): Regularized log-transformed expression values. Regularized log-transformed expression values of 12948 exons (rows), listed separately for each library

(column E-CY). Exons of low expression (library average <5) or highly variable expression (standard deviation over all libraries > library average) are not listed; ZosmaID (column A): gene ID based on the Zostera marina genome annotation v2.1 (GenBank Accession: LFYR00000000); mRNAID (column B): the mapped sequence ID; Source (column C): the source of annotation (inference from homology or top BLAST hit); Description (columnD): the gene description; baseMean (column CZ): the mean of normalized counts for all samples; sdcol (column DA): the standard deviation of expression counts that were normalized by size factors for all samples.

Table S4 (Excel): Annotations of mapped reads. ZosmaID: gene ID based on the *Zostera marina* genome annotation v2.1 from the ORCAE database, GenBank Accession: LFYR00000000); mRNAID: the mapped sequence ID; name: the gene description; annotation.source: the source of annotation (inference from homology or top BLAST hit); and GO.terms: the associated Gene Ontology terms.

Table S5 (Excel): Targeted GO-terms. GO-terms covered by the umbrella terms 'Heatstress' (Table S5a), 'Metabolism' (Table S5b), 'Oxidative-reductive' (Table S5c), 'Ribosomal' (Table S5d), 'Cellwall' (Table S5e), and 'Photosynthesis' (Table S5f) with ontologies (C: Cellular Process, F: Molecular Function, P: Biological Process) and descriptions.

Table S6 (Excel): Differential expression. Genes differentially expressed between Atlantic and Mediterranean samples under control (Table S6a) , stress (Table S6b), and recovery conditions (Table S6c); and genes differentially expressed between Northern and Southern samples under control (Table S6d), stress (Table S6e) and recovery conditions (Table S6f). Each row represents one gene with: ZosmaID (gene ID based on the Zostera marina genome annotation v2.1, GenBank Accession: LFYR00000000); mRNAID, the mapped sequence ID; Source, the source of annotation (inference from homology or top BLAST hit); Description, the gene description; baseMean, the mean of normalized counts for all samples; log2FoldChange, the log2 fold difference in normalized expression between Atlantic and Mediterranean (Table S6a-c) or between Northern and Southern samples (Table S6d-f, values > 1 indicate higher expression in Mediterranean/Southern samples); lfcSE, the standard error of the log2 fold difference; stat, the Wald test statistic for differential expression; pvalue, the p-value; padj, the p-value adjusted by the Benjamini-Hochberg method to control for false discovery rate; sdcol, the standard deviation of expression counts that were normalized by size factors for all samples; followed the regularized log-transformed expression values of all samples that were included in the test.

Table S7 (Excel): Genes responding to heat stress. Genes that responded to acute heat stress (time points 2 and 3) are represented for samples from all four populations in Table S7a, for NE samples Table S7c, for SE samples in Table S7d, and for SU samples in Table S7e. Genes that responded to heat in the recovery phase (time points 5,

7, and 9) are represented for all samples in Table S7b, for NE samples Table S7f, for SE samples in Table S7g, for NU samples in Table S7h, and for SU samples in Table S7i. ZosmaID (column A): the gene ID based on the *Zostera marina* genome annotation v2.1 (GenBank Accession: LFYR00000000); mRNAID (column B): the mapped sequence ID; Source (column C): the source of annotation (inference from homology or top BLAST hit); Description (column D): the gene description; baseMean (column E): the mean of normalized counts for all samples; log2FoldChange (column F): the log2 fold change in normalized expression between all control samples and all stressed samples; lfcSE (column G): the standard error of the log2 fold change; stat (column H): the Wald test statistic for differential expression; pvalue (column I): the p-value; padj (column J): the p-value adjusted by the Benjamini-Hochberg method to control for false discovery rate; sdcol(column K), the standard deviation of expression counts that were normalized by size factors for all samples. The following columns represent regularized log-transformed expression values of all libraries that were included in the test.

Table S8 (Excel): Enriched functions and processes under acute heat-stress and in the recovery. GO-terms that were significantly enriched in genes that were upregulated under acute heat stress (Table S8a for molecular functions MF, Table S8e for biological processes BP) or downregulated under acute heat stress (Table S8b for MF, Table S8f for BP) and in genes that were upregulated under recovery (Table S8c for MF, Table S8g for BP) or downregulated under recovery from heat stress (Table S8d for MF, Table S8h for BP). Each row represents one function with: the GO-term ID; the GO-term description, the number of annotated genes within this GO-term; the number of significantly upregulated genes within this GO-term; the expected number of upregulated genes; the p-value based on Fisher's exact test for enrichment.

Table S9 (Excel): Heat-responsive genes respresenting enriched functions and processes. Genes that were significantly upregulated under acute heat stress and included in enriched molecular functions (Table S9a) or biological processes (Table S9e). Genes that were significantly downregulated under acute heat stress and included in enriched molecular functions (Table S9b) or biological processes (Table S9f). Genes that were significantly upregulated under recovery from heat stress and included in enriched molecular functions (Table S9c) or biological processes (Table S9g). Genes that were significantly downregulated under recovery from heat stress and included in enriched molecular functions (Table S9d) or biological processes (Table S9h). Each row shows: the GO-term ID; the GO-term description; the gene ID (based on the *Zostera marina* genome annotation v2.1, GenBank Accession: LFYR00000000); the mapped sequence ID (mRNAID); the name of the gene; and the source of annotation (inference from homology or top BLAST hit).

Table S10 (Excel): Adaptively differentiated genes between Atlantic and Mediterranean samples. Genes that were adaptively differentiated between Atlantic

and Mediterranean samples, and enriched in biological processes (Table S10a) or molecular functions (Table S10b). Table S10c represents those genes that were not represented in enriched GO-terms. Each row represents one gene with its mapped sequence ID (mRNAID, column A), gene ID (ZosmaID, column B, based on the *Zostera marina* genome annotation v2.1, GenBank Accession: LFYR00000000), the source (column C) of annotation (inference from homology or top BLAST hit), and the name of the gene (column D). For Table S10a and Table S10b, the GO-term that each gene represents is represented in column E, and the description for this GO-term in column F.

Table S11 (Excel): Adaptively differentiated genes between Northern and Southern samples. Genes that were adaptively differentiated between Northern and Southern samples, and enriched in biological processes (Table S11a) or molecular functions (Table S11b). Table S11c represents those genes that were not represented in enriched GO-terms. Each row represents one gene with its mapped sequence ID (mRNAID, column A), gene ID (ZosmaID, column B, based on the *Zostera marina* genome annotation v2.1, GenBank Accession: LFYR00000000), the source (column C) of annotation (inference from homology or top BLAST hit), and the name of the gene (column D). For Table S11a and Table S11b, the GO-term that each gene represents is represented in column E, and the description for this GO-term in column F.

Supplementary Datasets

Dataset S1 (vcf file): Biallelic neutral SNPs. Set of 139,321 biallelic neutral SNPs with genotypes (GT), allelic depths (AD), read depth (DP), genotype quality (GQ), and Phred-scaled likelihood for genotypes (PL) listed for each sample.