Journal of Ecology

British Ecological Society

Journal of Ecology 2010, **98**, 374–383 doi: 10.1111/j.1365-2745.2009.01618.x

Variation in gene expression of *Andropogon gerardii* in response to altered environmental conditions associated with climate change

Steven E. Travers[1]*, Zhongwen Tang[1], Doina Caragea[2], Karen A. Garrett[1], Scott H. Hulbert[3], Jan E. Leach[4], Jianfa Bai[1], Amgad Saleh[1], Alan K. Knapp[5], Philip A. Fay[6], Jesse Nippert[7], Patrick S. Schnable[8] and Melinda D. Smith[9]

[1]*Department of Plant Pathology, Kansas State University, Manhattan, KS 66506, USA;* [2]*Department of Computer Science, Kansas State University, Manhattan, KS 66506, USA;* [3]*Department of Plant Pathology, Washington State University, Pullman, WA 99164, USA;* [4]*Bioagricultural Sciences and Pest Management, Colorado State University, Ft Collins, CO 80523, USA;* [5]*Department of Biology, Colorado State University, Ft Collins, CO 80523, USA;* [6]*USDA-ARS Grassland, Soil and Water Research Laboratory, Temple, TX 76502, USA;* [7]*Department of Biology, Kansas State University, Manhattan, KS 66506, USA;* [8]*Department of Agronomy, Iowa State University, Ames, IA 50011, USA; and* [9]*Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520, USA*

Summary

1. If we are to understand the mechanisms underlying species responses to climate change in natural systems, studies are needed that focus on responses of non-model species under field conditions. We measured transcriptional profiles of individuals of *Andropogon gerardii*, a C_4 grass native to North American grasslands, in a field experiment in which both temperature and precipitation were manipulated to simulate key aspects of forecasted climate change.

2. By using microarrays developed for a closely related model species, *Zea mays*, we were able to compare the relative influence of warming versus altered soil moisture availability on expression levels of over 7000 genes, identify responsive functional groups of genes and correlate changes in gene transcription with physiological responses.

3. We observed more statistically significant shifts in transcription levels of genes in response to thermal stress than in response to water stress. We also identified candidate genes that demonstrated transcription levels closely associated with physiological variables, in particular chlorophyll fluorescence.

4. *Synthesis.* These results suggest that an ecologically important species responds differently to different environmental aspects of forecast climate change. These translational changes have the potential to influence phenotypic characters and ultimately adaptive responses.

Key-words: ecological genomics, enrichment analysis, gene function, Gene Ontology, microarray, precipitation variability, Rainfall Manipulation Plots, tallgrass prairie, transcription profile, warming

Introduction

Understanding how individuals and populations respond to changing environmental conditions is an important step in understanding the ecology and evolution of species. With accumulating evidence that rapid and directional climatic change is upon us and forecasts for even greater future change (IPCC 2007), the need for this understanding has never been more urgent. Variation in biotic and abiotic characteristics of the environment can lead to responses at multiple scales – from gene expression to morphology. Plants perceive and respond to their environment through the transduction of signals from the gene level to the phenotype (Smith & Gallon 2001). However, in comparison to downstream phenotypic variables such as morphology or physiology, gene expression is expected to respond more rapidly to environmental change and stress. Moreover, because extensive genomic information is available for model species, a global evaluation of an organism's condition is possible by examining transcriptional patterns of large numbers of genes. Evaluating functional responses of plants at

*Correspondence author. E-mail: steven.travers@ndsu.edu

the genetic and biochemical level can help us better understand how particular species cope with variation in environmental conditions and stress and determine the underlying genetic bases of responses that limit the growth and survivorship of individuals (Ouborg & Vriezen 2007; Gibson 2008). Over time, differences among individuals in their responses to environmental change can lead to fitness variation, selection and adaptation; and, ultimately, variation among species in their ability to respond to environmental change can alter community composition and lead to differences in range limits among taxa. An extensive literature is available on plant stress physiology and disturbance ecology, but much less is known about how plants respond to changing environmental conditions at the genomic level, and even less is known about genomic responses of native plants to alterations in the environment under field conditions (Knight *et al*. 2006; Travers *et al*. 2007).

Advances in genomic technology have led to a dramatic increase in studies describing transcriptional patterns of gene expression in response to variation in environmental conditions, such as drought, but the majority of these have been in model systems under carefully controlled conditions. As noted by numerous authors (Oleksiak, Churchill & Crawford 2002; Purugganan & Gibson 2003; Vasemagi & Primmer 2005; Kammenga *et al*. 2007; Gibson 2008), gene expression patterns might be very different in field contexts with simultaneous variation in multiple environmental factors when compared to controlled conditions, and the degree of similarity in gene expression patterns between model species and non-model species is generally unknown. Thus, there is a pressing need for studies of non-model species in their natural environments.

In a previous study (Travers *et al*. 2007), we measured the transcriptional profiles in naturally established field populations of a widespread C_4 grass, big bluestem (*Andropogon gerardii*), under field conditions in a central US tallgrass prairie ecosystem. In that study, genomic responses of *A. gerardii* individuals exposed experimentally to either ambient or altered precipitation regimes were assayed using microarrays developed for a closely related species, maize (*Zea mays*), at a single point in time at the end of the growing season (September). We established that cDNA from *A. gerardii* hybridized sufficiently with maize microarrays to allow us to detect consistently lower expression of genes associated with photosynthesis and the Calvin cycle in individuals experiencing the altered precipitation regime, which has greater variability in and lower average levels of soil moisture (Knapp *et al*. 2002). Because *A. gerardii* contributes disproportionately to above-ground productivity in this system (Smith & Knapp 2003), these results provided the first direct genetic link to reduced productivity, a key aspect of ecosystem functioning, with altered precipitation.

We know from previous physiological studies that plant stress changes over the course of the spring and summer (Nippert *et al*. 2009). To test the hypothesis that shifts in productivity in this ecologically important species resulting from stress are linked to physiological and transcriptional changes over the course of the growing season, we examined the transcriptional and key physiological parameters at four different dates during the growing season. In addition, our previous work focused only on the effects of altered precipitation. However, temperatures and precipitation are expected to change in tandem with climate change for the Central Plains of the USA (IPCC 2007). Therefore, we assessed the transcriptional profiles of *A. gerardii* individuals exposed to both increased variability in precipitation and increased temperatures. Our objectives were to (i) quantify the relative influence of warming and increased variability in precipitation on transcriptional profiles of *A. gerardii* in a field context; (ii) identify functional groups of genes that are responsive to environmental variation representative of that expected with climate change; (iii) assess the effects of temporal variation in environmental variables on transcriptional profiles of *A. gerardii*; and (iv) quantify the relationship between expression of individual genes and physiological responses for individual plants.

Materials and methods

THE RAINFALL MANIPULATION PLOTS EXPERIMENT

We conducted this study in a long-term climate change field experiment – the Rainfall Manipulation Plots (RaMPs) – located at the Konza Prairie Biological Station. The Konza Prairie is a 3487-ha tallgrass prairie preserve located in north-eastern Kansas, USA (39°05′ N, 96°35′ W). The site experiences a temperate mid-continental climate of cold, dry winters and warm, wet summers. The long-term mean annual precipitation is 835 mm, of which 75% occurs during the growing season (May–September) (Hayden 1998). Air temperatures peak in mid-summer, from 18 °C in May, to 27 °C in July and 22 °C in September (Hayden 1998). The site is dominated by C_4 grasses, such as *A. gerardii* and *Sorghastrum nutans*. Intermixed within this matrix of grasses is a diverse suite of C_3 grasses, forbs and woody species.

The RaMPs experiment was initiated in 1998 to examine the effects of climate change on intact, tallgrass prairie ecosystem structure and function (see Fay *et al*. 2000). The experiment is located in a well-drained lowland site and consists of 12 fixed-location rainfall exclusion shelters (14×9 m), which allow complete control of the annual precipitation quantity, timing between precipitation events and the size of individual events. During the growing season (mid-April to mid-September), precipitation is applied in two regimes, ambient or altered (six shelters each). For the ambient precipitation treatment, rainfall from each natural event is collected from the roof of each shelter and immediately reapplied. For the altered treatment, the interval between precipitation events is increased by 50%, while the total precipitation amount is kept equal to the ambient treatment. The altered growing season precipitation regime produces: (i) longer inter-rainfall dry periods, (ii) greater rainfall event sizes, and (iii) fewer rainfall events. Warming treatments (ambient or warmed by *c*. 2 °C) are applied at the subplot level within the precipitation treatments. The plot beneath each shelter is divided into four 2×2 m subplots (48 subplots total). For two of the subplots, infrared heating lamps (Kalglo 240V HS-2420; Kalglo Electronics Co. Inc., Bethlehem, PA, USA) were installed to raise the annual average ambient air temperature by 2 °C. To characterize the effects of the precipitation and temperature treatments on environmental conditions, soil water content and soil temperature at 2, 5 and 15 cm, respectively, as well as canopy temperature, were continuously measured within each or a subset of the RaMPs (see Table 1 for details).

Table 1. Environmental variables continuously measured in the RaMPs and at the time of plant sampling

Principal component	Variable	Description	Units	Scale
PC2	wv15	TDR measurement of volumetric soil water content	%	Subplot
PC2	h2o_kgm3_25	Neutron probe reading of soil water content at 25 cm	kg m^{-3}	Subplot
PC2	h2o_kgm3_100	Neutron probe reading of soil water content at 100 cm	kg cm^{-3}	Subplot
PC1	soilT2cm	Thermocouple reading of soil temperature at 2 cm	°C	Subplot
PC1	soilT5cm	Thermocouple reading of soil temperature at 5 cm	°C	Subplot
PC1	soilT15cm	Thermocouple reading of soil temperature at 15 cm	°C	Subplot
PC1	Ctt	Canopy temperature	°C	Subplot (n = 3 per treatment)

PLANT SAMPLING OF FIELD POPULATIONS

In contrast to transcriptional profiling studies of plants growing under controlled conditions, relatively large numbers of replications are needed to assess accurately transcription by plants in the field because of the increased variance in environmental conditions. Therefore, we sampled 24 plants per sample date. The 24 plants were paired on 12 microarrays by pairing samples from adjacent ambient and altered plots. By maximizing the number of arrays used, we minimized the probability that transcriptional measurements per spot on the array were due to printing anomalies on any one array and therefore the need for secondary quantification of transcript amounts per spot (Kothapalli *et al.* 2002).

To associate changes in environmental conditions with physiological and gene expression responses of *A. gerardii*, we sampled individuals within one randomly selected control (unwarmed) and one randomly selected warmed subplot in each RaMP at four times during the 2005 growing season: 1 June, 17 July (JulyA), 21 July (JulyB), 15 August. The two dates in mid-July were selected to capture a mid-season drought-recovery period; on the first date, soil moisture content was low and probably limiting and the second date was 3 days after we had applied rainfall events to both the altered and ambient precipitation treatment plots.

For each sample date (hereafter June, JulyA, JulyB, August), morphologically similar tillers (3–5 fully expanded leaves) were identified in each of the subplots (24 individuals per sample period; 96 individuals total). For each individual, the first or second fully expanded leaf was randomly selected for genomic analysis; the entire leaf was clipped, flash-frozen and stored in liquid nitrogen. Immediately after collection, we measured the remaining leaf maximum photosynthetic rate at saturating light (1500 μmol m^{-2} s^{-1} photon flux density) and ambient CO_2 (370 μL L^{-1}), leaf respiration (R_d), dark-adapted chlorophyll fluorescence (F_v/F_m), stomatal conductance to water (g_s), instantaneous water-use efficiency (A_{sat}/g_s: WUE), transpiration (T) and midday leaf water potential (LWP), along with other associated variables (Table 2). Leaf-level gas exchange was measured using a LI-6400 gas exchange system with red/blue light source and CO_2-injector (LiCOR, Inc., Lincoln, NE, USA). Light intensity inside the cuvette was 2000 μmol m^{-2} s^{-1}, the CO_2 concentration was 370 μmol mol^{-1}, and the relative humidity was maintained at ambient levels. F_v/F_m was measured using a hand-held pulse amplitude modulated fluorometer (OS1-FL; Opti-sciences, Inc., Tyngsboro, MA, USA). Leaves were dark-adapted prior to measurement for *c.* 15 min. Finally, the same leaves, in which photosynthetic and fluorescence measurements were measured, were collected from each individual for determination of LWP using a Scholander-type pressure chamber (PMS Instruments, Inc., Corvallis, OR, USA). A portion of the leaf was set aside for later measurement of chlorophyll content (Hiscox & Israelstam 1979; Wellburn 1994) and the remaining portion was dried for 48 h at 60 °C and ground to determine total C and N content.

MICROARRAY ANALYSIS

Leaf tissues samples collected in the field were stored in a −80 °C freezer for no more than 6 months prior to RNA extraction. Total RNA was purified with the RNeasy kit by Invitrogen (Carlsbad, CA, USA). RNA concentration was estimated (Nanodrop, Wilmington, DE, USA), and 2 μg RNA from each plant was used to synthesize cDNA. We used the Array900 3DNA kit from Genisphere (Hatfield, PA, USA) to label indirectly our cDNAs by using primers specific to

Table 2. Physiological variables measured on each experimental plant

Variable	Description	Units	Range
R_d	Dark respiration of leaf	μmol m^{-2} s^{-1}	−5.91 to 0.402
A_{sat}	Max photosynthesis at saturating light intensity	μmol m^{-2} s^{-1}	−1.65 to 20.1
g_s	Stomatal conductance	mol m^{-2} s^{-1}	0.0095 to 0.168
T	Transpiration rate	mol m^{-2} s^{-1}	0.554 to 5.28
WUE	Water-use efficiency	Ratio of Photo : Trmmol	−1.46 to 5.81
F_v/F_m	Chlorophyll fluorescence		0.595 to 0.796
LWP	Leaf xylem pressure potential	MPa	−28.7 to −4.2
Chl_{tot}	Total leaf chlorophyll content	mg chl g^{-1} tissue	0.79 to 6.19
ChlA:B	Ratio of chlorophyll *a*:*b*		2.3 to 11.2
%C	C content of leaves	%	42.7 to 46.5
%N	N content of leaves	%	0.64 to 1.89
C:N	C:N ratio in leaves		23.7 to 72.2

one of two fluorophores (Cy5, Cy3). Samples from individual plants that were collected from adjacent RaMPs within 5 min of each other were paired according to precipitation treatments (e.g. altered/unwarmed with ambient/unwarmed) on the microarray to minimize differences in sampling time of day and physical location of the two plants per pair. This sampling regime and array assignment were based on a model for optimization of multiple treatment effects in a pairwise format (Milliken, Garrett & Travers 2007). The labelled cDNA was hybridized overnight at 50 °C to cDNA microarrays, followed by a second hybridization of the dye molecules with the specific primers on the cDNAs. We used maize-spotted microarrays (v. SAM 1.1) produced by the Center for Plant Genomics at Iowa State University (GEO platform GPL3333; Appendix S1). The arrays consisted of 19 200 maize cDNA clones (14 401 informative) isolated from three maize gene libraries: Stanford UnigeneI, Unigene-IV 1091, library 3529 and the ISUM libraries from Iowa State University (http://www.plantgenomics.iastate.edu/maizechip/).

Microarray chips were scanned with an Axon Genepix 4000B scanner at two wavelengths, 652 and 532 nm, using GENEPIX software (v. 6; Molecular Devices, Sunnyvale, CA, USA) producing two-channel images. After aligning the spots using a template grid and optimizing the two channel gain levels to yield a whole array count ratio of *c*. 1, we flagged obvious anomalies and features with signal-to-noise ratios < 3. Because we performed heterologous hybridizations, there was the potential for a reduced fluorescence signal as a function of sequence polymorphisms between maize and the target organism (Hammond *et al*. 2006, but see Lunden *et al*. 2008). To filter cDNA clones that hybridized poorly with sample cDNA due to sequence polymorphism, we screened across sample averages of fluorescence intensity on a per-gene basis and excluded those features within two standard deviations of background fluorescence. Features not satisfying the minimum criterion were assumed to represent those for which there was no expression or which represented too great a sequence divergence between maize and the *A. gerardii* samples to be accurately assessed. The remaining intensity values on each array were then normalized with the ratio of medians protocol.

STATISTICAL ANALYSIS

As expected, some of the environmental variables measured in the RaMPs were highly correlated. Therefore, we used principal component (PC) analysis to collapse eight of the environmental variables into two representative PCs (PC1 represented temperature, PC2 represented soil moisture, see below). These PC values were used in subsequent analyses relating gene expression to the two major axes of environmental variation.

The altered precipitation and warming treatments did not always result in a significant decrease in average soil water content or an increase in average soil or canopy temperatures between the four sampling dates (Nippert *et al*. 2009). However, the altered precipitation and warming treatments did increase the range of variation in soil water content and soil and canopy temperatures captured over the growing season (Nippert *et al*. 2009). Therefore, we treated the environmental conditions resulting from the treatments (soil moisture, temperature) as continuous variables rather than discrete treatment categories (Fig. 1). As a consequence, it was necessary to assess how gene expression and physiology were impacted by variation in environmental conditions using a high-throughput regression analysis approach (Wu 2005), rather than conducting a more traditional two-colour microarray analysis based on an ANOVA format for comparing treatment mean values (Wolfinger *et al*. 2001; Milliken, Garrett & Travers 2007; Morinaga *et al*. 2008). Statistical significance of increasing (positive slope) or decreasing expression (negative slope) of a particular gene in relation to the target environmental or physiological variables was evaluated using the *Q*-statistic for the null hypothesis that the slope was zero (Storey & Tibshirani 2003).

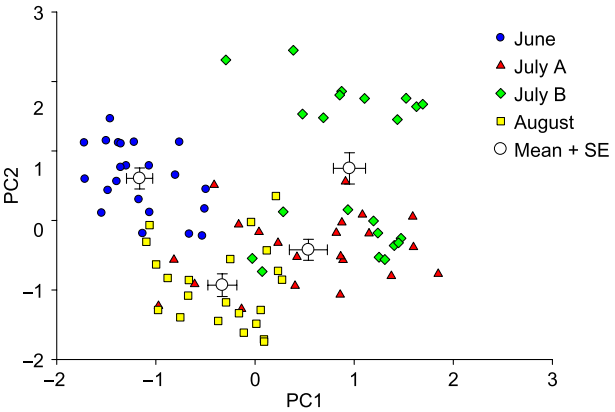

Fig. 1. Environmental measures for the sample plots in PC space. PC1 and PC2 for each plot sampled are shown for the four sampling dates (June, JulyA, JulyB and August). Different sampling dates are denoted by different symbols. Mean and standard error values for PC1 and PC2 for each sampling date are indicated by open circles.

Prior to the high-throughput regression analysis we normalized the fluorescence values by conducting mixed-model ANOVAS with array and dye included as fixed effects: $y_{ij} = \mu + A_i + D_j + (A \times D)_{ij} + \varepsilon_{ij}$, where y_{ij} is the fluorescence intensity on the ith array and labelled with the jth dye, μ is the mean, A_i is the effect of the ith array (i = 1–16), D_j is the effect of the jth dye, $(A \times D)_{ij}$ is the array–dye interaction and ε_{ij} is the stochastic error.

We then analysed the residuals from the normalization model with six linear regression models designed to measure six different relationships between transcription levels, environmental conditions and physiological responses within and among sample periods (Table 3). The first model was a linear regression designed to examine the relationship between expression and variation in soil moisture content and temperature. We regressed the gene expression residuals of all 96 plants over the four sampling periods against both environmental PCs (Table 3). We excluded those genes for which less than three data points were available for any of the four sample periods. To eliminate

Table 3. Statistical models for assessing on a gene-by-gene basis relationships between expression levels, environmental conditions and physiological responses

Model 1	$r_i = \mu + \beta_1 E_{i1} + \beta_2 E_{i2} + \beta_3 E_{i1} E_{i2} + \varepsilon_i$
Model 2	$r_{ik} = \mu + \beta_1 E_{i1} + \beta_2 E_{i2} + \tau_k + \beta_{3k} E_{i1} \tau_k + \beta_{4k} E_{i2} \tau_k + \varepsilon_i$
Model 3	$p_{mi} = \mu + \beta_1 E_{i1} + \beta_2 E_{i2} + \beta_3 E_{i1} E_{i2} + \beta_4 r_i + \varepsilon_i$
Model 4	$p_{mi} = \mu + \beta_1 E_{i1} + \beta_2 E_{i2} + \tau_k + \beta_3 E_{i1} \tau_k + \beta_4 E_{i2} \tau_k + \beta_5 r_i + \varepsilon_i$
Model 5	$p_{mi} = \mu + \beta_1 r_i + \varepsilon_i$
Model 6	$p_{mi} = \mu + \tau_k + \beta_1 r_i + \varepsilon_i$

r_i, residual for gene i; r_{ik}, residual for gene i for sample period k; E_{i1}, environmental principal component 1 for i^{th} plant; E_{i2}, environmental principal component 2 for i^{th} plant; τ_k, kth sampling period (k = 1,…,4); p_{mi}, value of the mth physiological variable for the i^{th} plant.

poor estimates of regression slopes due to restricted data ranges, we only included expression data in further analysis if the value of the diagnostic Cook's *D*-statistic for an initial model fitting was > 1 (Cook & Weisberg 1982). The second regression model expanded model 1 by adding a term for sampling date as a predictor (Table 3). Since model 2 includes more parameters, our criterion for inclusion of genes in the analysis was a minimum of nine observations per sample time.

A similar regression model was used to analyse the relationships between gene expression levels and the associated physiological responses of those individuals. In these analyses, regressions were conducted for each physiological variable separately. Our first goal was to determine for each physiological variable which genes added predictive power ($Q < 0.05$) after the environmental variables were included as predictors. Thus, model 3 was a multiple regression of each physiological parameter against all environmental principal components and each gene (Table 3). For model 4, sampling period was added to model 3 as a predictor to test for within-period effects, and models 5 and 6 were the same models as 3 and 4, respectively, but with the environmental variables removed (Table 3).

FUNCTIONAL GENE ANNOTATION

We successfully annotated 7243 genes based on Gene Ontology (GO; The Gene Ontology Consortium 2000) annotations obtained from the Maize Gene Index data base (ZmGI) developed by the DFCI Gene Index project (http://compbio.dfci.harvard.edu/tgi). We then used the software Blast2Go (Conesa *et al.* 2005) to analyse the functional categories of those genes with significant relationships between expression and an environmental principal component. Analyses were not conducted for physiological data because our interest in functional gene groups was restricted to genes responding to environmental conditions (but see Nippert *et al.* 2009). Blast2GO is a suite of user-friendly tools for similarity-based functional annotation and analysis of non-model species genomic data. Blast2GO includes a tool for performing enrichment analysis, i.e. the identification of GO annotations whose abundance is significantly different between two sets of annotated genes. The enrichment analysis integrates Blast2GO with Gossip (Blüthgen *et al.* 2005), a software package that employs Fisher's exact test to determine the significance of associations between two categorical variables, while correcting for multiple testing using a FDR (false discovery rate), a FWER (family-wise error rate) and a single test *P*-value (Fisher *P*-value). A set of GO terms that are under- or over-represented at a specified significance level is obtained as a result of performing the enrichment. We then performed enrichment analysis to identify those GO terms where there was a significant difference in abundance between genes exhibiting a positive versus negative slope in relation to the two environmental PCs within a date.

Results

ENVIRONMENTAL CONDITIONS

During the 2005 sampling period, soil water content was highest in early June and decreased throughout the summer, and soil water content was consistently lower at shallow depths than at greater depths (Table 4). The rainfall application between the two sampling dates in July increased the soil water content at shallow depths (25 cm) but not greater depths (100 cm). As a result, variance in soil water content was higher at 25 cm than at 100 cm over the four sampling dates. However, even at the greater depths soil water content decreased by 12% from the June to the August sampling date. Soil and leaf temperatures were highest in July and lowest in June (Table 4). For the drought-recovery period in July, soil surface temperatures decreased from the first to the second sampling date but there was little change at 15 cm depth. Leaf temperatures in contrast were an average of 2 °C higher on the second July sampling date compared to the first July sampling date.

Principal component analysis revealed that the first two PCs explained 84% of the variation in the eight environmental variables. The first component (environmental PC1) had loadings that were disproportionately high for all the temperature values; higher temperatures were associated with higher PC1 values. In contrast, soil water variables loaded more heavily on PC2. Centroids for the four sampling dates indicated that no two dates had the same combination of heat and soil moisture levels (Fig. 1). These PC values were subsequently used in the high-throughput regression analyses examining relationships between environmental conditions, gene expression and physiology.

GENE EXPRESSION AND ENVIRONMENTAL VARIABLES

As a result of quality control of data from each array and array-to-array variation in hybridization success, a total of 7098 genes of the 19 200 (37%) on the array were analysed for correlations with environmental variation and physiological variation (see GEO-NCBI: GSE16418). We found that gene expression was significantly related to environmental PC1 in 1370 (19%) of the 7098 genes when all four sampling dates were considered simultaneously (model 1, Fig. 2a). In contrast, only 101 genes (1%) demonstrated a significant relationship between expression and PC2 across the four sample dates (Fig. 2b). There were roughly equal numbers of genes that

Table 4. Average (± 1 SE) soil moisture and leaf and soil temperature at the time of each sampling date (June = June 1, JulyA = July 17, JulyB = July 21 and August = 15). See Table 1 for details of how the environmental variables were measured

Sampling date	Water volume, 15 cm (%)	Water content, 25 cm (kg m^{-3})	Water content, 100 cm (kg m^{-3})	Leaf temperature (°C)	Soil temperature, 2 cm (°C)	Soil temperature, 15 cm (°C)
June	15.4 + 0.1	250.7 + 8.4	334.1 + 11.1	28.5 + 0.5	23.6 + 0.3	20.9 + 0.2
JulyA	13.9 + 0.1	205.4 + 6.1	304.7 + 8.2	37.5 + 0.7	30.8 + 0.7	26.5 + 0.2
JulyB	33.2 + 0.4	271.8 + 11.5	312.8 + 7.1	39.9 + 0.9	29.5 + 0.5	26.7 + 0.1
August	13.3 + 1.0	198.3 + 8.5	274.3 + 8.2	30.2 + 0.6	26.2 + 0.4	23.6 + 0.1

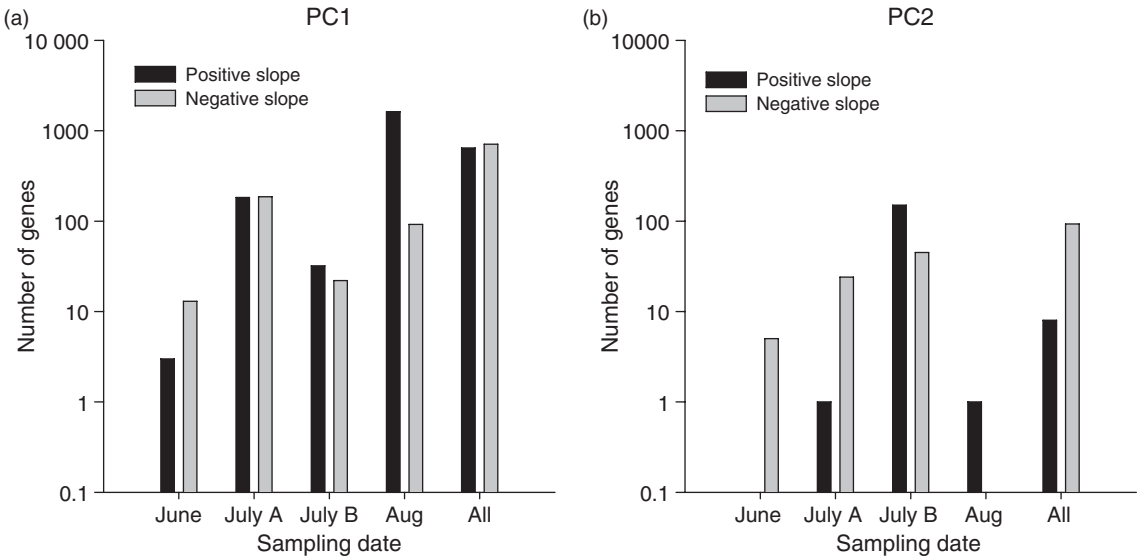

Fig. 2. Total number of genes (out of 7098) with significant relationships between expression and PC1 (a) and PC2 (b) by sampling date. Positive slopes are indicated by black and negative slopes are indicated by grey. Note the log-scale for the *y*-axis.

were positively and negatively related to PC1 (650 and 720, respectively), whereas the majority of genes with expression correlated with PC2 were negatively related. These results indicate that the expression of genes that successfully hybridized was more likely to respond to the temperature PC1 than the soil water content PC2, and that drier conditions associated with the altered precipitation were characterized by higher gene expression. A large proportion (*c.* 40%) of the genes that responded positively or negatively to PC1 were of unknown function (557 of 1370). However, four GO categories accounted for most of the known function genes (471 of 813): transcription (195 (24%), negative slope), biosynthetic process (114 (14%), positive slope), binding (89 (11%), positive slope) and protein folding (73 (9%), positive slope). In contrast, the majority of the genes that responded to PC2 were functionally part of a biosynthetic process (26 of 51 (51%), negative slope), transcription (2 (4%), negative slope), binding (1 (2%), positive slope), protein folding (0 (0%)) and unknown function (50 of 101 (49.5%)).

For model 2, the pattern of more responsiveness of genes to PC1 than PC2 was consistent among each of the sampling dates with the exception of the second July sampling date after the rainfall event was applied to both treatments, where more genes were related to PC2 than PC1 (June: 16 vs. 5; JulyA: 369 vs. 25, JulyB: 54 vs. 195, August: 1713 vs. 1). At the two sampling dates after periods with the greatest thermal and drought stress (JulyA and August), there were far more genes with a significant relationship with PC1 than with PC2 when compared to the beginning of the growing season (Fig. 2).

The genes with known function that responded significantly to temperature changes (PC1) were in different functional categories at different sampling dates (Fig. 3). While there was under-representation of known function genes related to PC1 across all of the GO categories for the June or JulyB sampling dates (i.e. fewer genes significantly related to PC1 than expected by chance), there was a relatively large number of genes with expression levels related to PC1 at both JulyA and August dates that were over-represented. These were categorized predominantly as stress response genes. At the JulyA sampling date, genes responding to changes in temperature (PC1) were in GO categories typically associated with plant responses to stress including response to heat, response to oxidative stress, response to unfolded protein and superoxide metabolic process (Fig. 3). Specifically, gene expression increased for protein folding and anti-oxidant response genes with increasing temperature (PC1) at the JulyA date. In contrast, none of these categories of genes were over-represented at the other sampling dates with the exception of the unfolded protein genes and the oxidative stress genes in August. At the JulyA date, there was also over-representation of genes in the regulation of molecular function categories: positive regulation of transferase activity, protein kinase activity and MapK activity. This was the only sampling date on which these categories were over-represented. There were also several plant defence and programmed cell death genes that were expressed at lower levels in response to higher temperatures. In contrast to the other three sampling dates, August was characterized by an over-representation of genes involved in protein production associated with change in temperature.

Very few genes had expression that was related to PC2 on any of the sampling dates except for the post-drought sampling date (JulyB). The majority of the known function genes were positively related to PC2 and broadly categorized as biosynthetic process, binding and translation.

GENE EXPRESSION AND PHYSIOLOGICAL VARIABLES

Regression analysis indicated that physiological variables were more likely to be related to expression of genes over the course of the entire growing season (model 3) compared to at each of

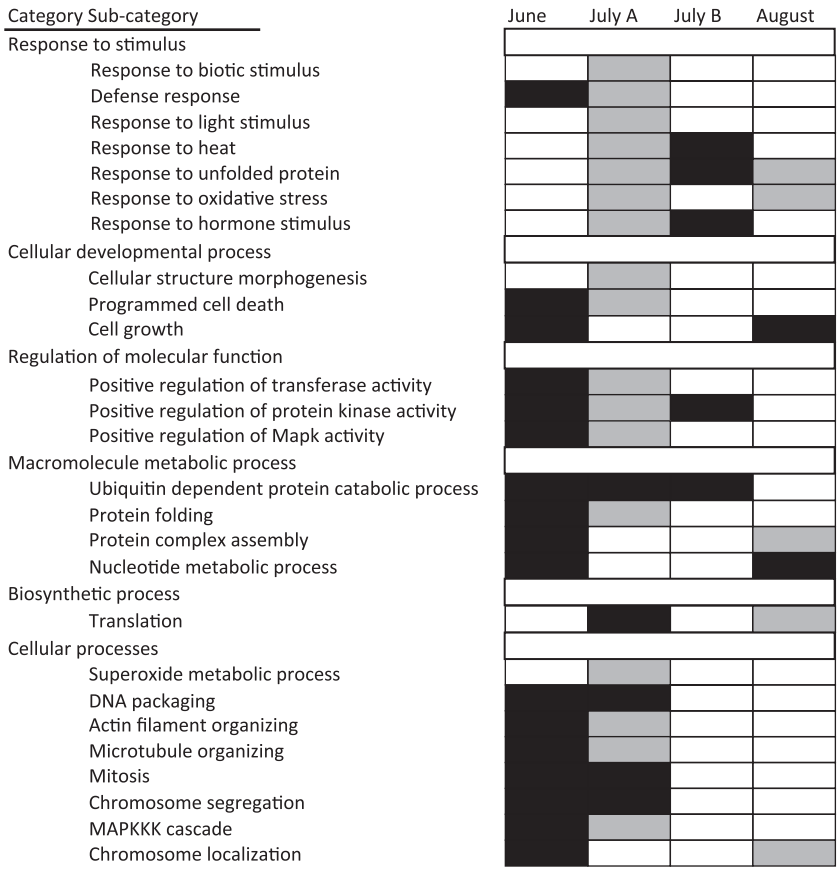

Fig. 3. Enrichment analysis of GO functional categories for PC1 by sampling date. Grey boxes indicate that significantly more genes from a GO category showed altered expression patterns at that sampling date than expected by chance. Black indicates fewer than expected genes showed altered expression than by chance. White boxes indicate no difference from the expectation based on chance.

Table 5. Number of genes with significant positive (+) or negative (−) relationships between expression and physiological or related variables. See Table 2 for explanation of the physiological and related variables. See Table 3 for explanation of models 3 and 5. Results shown are for analysis of all four sampling dates simultaneously

	Model 3		Model 5	
	+	−	+	−
A_{sat}	0	0	0	0
g_s	0	0	0	0
T	0	0	0	0
R_d	0	0	4	13
F_v/F_m	30	349	297	1715
LWP	0	0	570	2049
WUE	0	0	681	699
%C	0	0	0	0
%N	10	160	1033	2354
C:N	1379	246	2832	970
Chl_T	0	0	348	401
ChlA:B	0	0	121	121

Table 6. Number of genes with significant relationships between expression and the physiological and related variables by sampling date. See Table 2 for explanation of physiological and related variables. See Table 3 for explanation of models 4 and 6

	June		JulyA		JulyB		August	
	M6	M4	M6	M4	M6	M4	M6	M4
A_{sat}	0	0	1	0	0	0	1	1
g_s	0	0	0	0	0	0	0	1
T	0	0	0	0	0	0	0	1
R_d	0	0	0	0	0	1	0	0
F_v/F_m	0	0	0	0	0	0	30	2
LWP	0	0	0	0	0	0	1	0
WUE	0	0	0	0	0	0	0	0
%C	0	0	1	0	0	0	0	0
%N	0	0	0	0	0	0	0	0
C:N	0	0	0	0	0	0	0	0
Chl_T	0	0	0	0	0	0	0	0
ChlA:B	0	0	0	0	0	0	0	0

the four sampling dates (model 4), and removing variation in physiological variables due to environmental factors (models 5 and 6) resulted in substantially fewer associations between gene expression and physiology (Tables 5 and 6).

With model 5, the C : N ratio in the leaves was related to expression level of the greatest number of genes (3802 of 7098); removal of variation due to environmental factors (model 3) reduced this total to 1625. There was a greater number of genes related to N content in the leaves than C content in the leaves, suggesting that N levels were influencing which functional categories of genes were expressed. Chlorophyll fluorescence (F_v/F_m) also was often correlated with gene expression (Table 5), and inclusion of environmental terms in the analysis reduced the total number of genes from 2012 to 379. Twenty-one percent of the known genes related to F_v/F_m (80 of 379)

were homologous to histones involved in DNA binding. F_v/F_m was also associated with genes in transcription (3%, 11), biosynthesis (1%, 4), antioxidant response (1%, 4) and photosynthesis (< 1%, 1) functional groups.

Physiological variables were related to expression levels of far fewer genes at each of the sampling dates compared to across the season. There were no significant relationships between any of the physiological variables and gene expression on the June sampling date and only two on the JulyA date prior to watering (Table 6). However, F_v/F_m was related to the expression level of 30 genes in the August sampling date when environmental terms were not included in the model. Removal of variance due to environmental factors resulted in a total of three different genes with significant relationships between expression and physiology. Specifically, A_{sat}, g_s and T were associated with expression level of BM079333, a homologue for a plasma membrane protein associated with water channel activity (Table 6). Greater levels of expression of this gene were associated with lower measures of all three physiological variables. F_v/F_m was particularly strongly associated with two genes: the EST CB411262, similar to a methyl-CpG binding protein (MBD109), and DV549941 with unknown function.

Discussion

Our study is among the first to assess expression patterns and their relationships to environmental factors and physiological responses in an ecologically important non-model species in the field. We found that transcriptional profiles of *A. gerardii* change in a consistent pattern among the individuals sampled in response to environmental variation, specifically soil temperature and soil moisture content. Moreover, the patterns of transcriptional responses were not random with regard to environmental variation or functional category of responsive genes. These results provide insights into how this common C_4 grass species maintains growth and survivorship under changing environmental conditions associated with climate change and characterizes the functional groups of genes that may be targets of selection.

The number and types of genes responding to temperature variation in this experiment suggest that *A. gerardii* responds differently to the combined effects of warming and reduced water availability than it does to either of these environmental factors alone. The greatest number of genes had expression related to either PC1 or PC2 at the JulyA sampling date. Enrichment analysis indicated a clear over-representation of genes functionally associated with stress at this sampling date (Fig. 3). In contrast, at the beginning of the season (June) there was an under-representation of genes in the same functional categories, suggesting that for this species, genetic responses to the environment are limited prior to stressful conditions. No other sampling date was simultaneously as hot and dry as the JulyA date. As expected based on studies of crop species in response to heat and drought stress (Bohnert, Nelseon & Jensen 1995; Kawasaki *et al.* 2001; Ozturk *et al.* 2002; Wang *et al.* 2003; Baniwal *et al.* 2004), there was an over-representation of genes involved in maintaining protein folding (e.g. heat-shock proteins) and in the anti-oxidant response (superoxide dismutase and thioredoxin peroxidase). In contrast, there were far fewer genes with detectable shifts in transcription associated with either the June or JulyB sampling date before each of which either temperatures were milder or water availability was higher (Fig. 3). Moreover, far fewer genes responded to changes in soil moisture than to changes in temperature. The exception to this pattern was at the sampling date following the release from drought via a targeted rainfall application (JulyB). After the rainfall application to both treatments, there was a dramatic increase in the number of genes with expression positively related to PC2, which was associated with an increase in soil moisture. The increase in expression with increasing soil moisture was observed for a subset of biosynthetic genes (ribosomal function and translation), as well as genes associated with histones and DNA binding. In addition to increasing translation in less water-stressed plants, the release from drought appeared to have ameliorated the influence of soil temperature on gene expression, suggesting that it is the combination of high temperatures and reduced water availability that explains the stress transcription profiles and not either component by itself. Indeed, prior to watering we found relationships between expression and PC1 for over 300 genes. However, after watering, the number decreased to a total of 54 despite the JulyB date having the highest mean leaf temperature of all four sampling dates (Table 4).

The transcriptional profiles of *A. gerardii* plants changed over the course of the growing season in ways that provide insight into how this species may adjust to changing environmental conditions. There were relatively few differences among plants early in the summer when temperatures are low and soil moisture is less limiting. However, increases in stress due to either reduced water availability or higher temperatures alone generally increased the responsiveness of only a few functional categories of biosynthetic genes as indicated by transcription profiles from the July B and August sampling dates. In contrast, the transcriptional profiles of plants after the high temperatures and low water availability of the first July period, i.e. prior to 17 July, were characterized by the highest number of responsive genes in the greatest number of functional categories. These transcriptional profiles probably reflect the biochemical and physiological characteristics of stress tolerance in this dominant species during summer environments and may differ from profiles of other C4 grass species in the community. Our results also suggest that it is insufficient to consider transcriptional responses of plants to single sources of environmental stress because the interaction of multiple stressors is more indicative of the field context and plants respond differently to a combination of factors.

We were also interested in whether gene expression data could be used to improve estimates of physiological responses over and above predictions based solely on abiotic environmental variables. In this experiment, once environmental covariance was accounted for and developmental shifts over the course of the growing season were considered (by analysing

by sample date), there were very few genes that changed expression in tandem with the physiological variables we measured. The exception is the relationship between fluorescence (F_v/F_m) and gene expression. At the scale of the entire summer, fluorescence decreased with increasing expression of 349 genes and increased with expression of 30 genes. On the August sampling date, fluorescence was correlated with expression of 30 genes, most with unknown function, once environmental covariance was removed. Previous studies of *A. gerardii* under these experimental conditions have indicated reduced productivity and photosynthetic activity of plants in altered treatments relative to ambient treatments (Knapp *et al.* 2002; Nippert *et al.* 2009). The relationships between fluorescence and transcriptional profiles are the first indication of a possible mechanism for compromised photosynthetic capacity in the context of environmental stress.

Plant traits important for responding to drought and heat stress are known to have heritable variation. In particular, there is abundant evidence of significant quantitative genetic variation for water-use efficiency (e.g. Geber & Dawson 1997; Heschel *et al.* 2002; Mckay, Richards & Mitchell-Olds 2003). Our results delineate a subset of candidate genes from a subset of functional categories (e.g. Fig. 3) found to be particularly responsive to environmental shifts that can now be examined for direct influence of phenotypic expression in *A. gerardii* exposed to drought and heat stress. The levels of variation we observed in this experiment were sufficient to result in significant shifts in expression of genes associated with proteins typical for stress responses including heat-shock proteins and genes ultimately involved in the production of antioxidants (Bohnert, Nelseon & Jensen 1995). Ultimately, the influence of selection on any of the candidate genes we have identified, as well as the long-term population responses of *A. gerardii* to climate change, will depend on how the shifts we have seen at the genomic level translate to changes in growth, reproduction and survivorship of individuals.

Acknowledgements

Research assistance was provided by Shauna Dendy, Kathryn Wagner and Casey Anderson. George Milliken provided guidance in the statistical analysis of our results. The Kansas EPSCoR Ecological Genomics consortium provided financial support and lively discussion of this work. Our study was funded by NSF Grant #DEB-0130692, the NSF Long Term Ecological Research Program at Konza Prairie, the USDA CSREES Ecosystems Studies Program, the Office of Science (PER), US Department of Energy Grant #DE-FG02-04ER63892, NSF Grant #DBI-0421427 and the Kansas Agricultural Experiment Station (Contribution #09-218-J). Microarray data files are available at GEO-NCBI data bank and can be accessed through the series ID number GSE16418 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE16418).

References

Baniwal, S., Bharti, K., Chan, K., Fauth, M., Ganguli, A., Kotak, S. *et al.* (2004) Heat stress response in plants: a complex game with chaperones and more than twenty heat stress transcription factors. *Journal of Biosciences*, **29**, 471–487.

Blüthgen, N., Brand, K., Cajavec, B., Swat, M., Herzel, H. & Beule, D. (2005) Biological profiling of gene groups utilizing Gene Ontology. *Genome Informatics*, **16**, 106–115.

Bohnert, H.J., Nelseon, D.E. & Jensen, R.G. (1995) Adaptations to environmental stresses. *The Plant Cell*, **7**, 1099–1111.

Conesa, A., Gotz, S., Garcia-Gomez, J.M., Terol, J., Talon, M. & Robles, M. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, **21**, 3674–3676.

Cook, R.D. & Weisberg, S. (1982) *Residuals and Influence in Regression*. Chapman and Hall, New York, NY.

Fay, P.A., Carlisle, J.D., Knapp, A.K., Blair, J.M. & Collins, S.L. (2000) Altering rainfall timing and quantity in a mesic grassland ecosystem: design and performance of rainfall manipulation shelters. *Ecosystems*, **3**, 308–319.

Geber, M. & Dawson, T. (1997) Genetic variation in stomatal and biochemical limitations to photosynthesis in the annual plant, *Polygonum arenastrum*. *Oecologia*, **109**, 535–546.

Gibson, G. (2008) The environmental contribution to gene expression profiles. *Nature Reviews Genetics*, **9**, 575–581.

Hammond, J.P., Bowon, H.C., White, P.J., Mills, V., Pyke, K.A., Baket, A., Whiting, S.T., May, S.T. & Broadley, M.R. (2006) A comparison of the *Thlaspi caerulescens* and *Thlaspi arvense* shoot transcriptomes. *New Phytologist*, **170**, 239–260.

Hayden, B.P. (1998) Regional climate and the distribution of tallgrass prairie. *Grassland Dynamics: Long-term Ecological Research in Tallgrass Prairie* (eds A. Knapp, J.M. Briggs, D.C. Hartnett & S.L. Collins), pp. 19–34. Oxford University Press, New York, NY.

Heschel, M.S., Donohue, K., Hausmann, N. & Schmitt, J. (2002) Population differentiation and natural selection for water-use efficiency in *Impatiens capensis* (Balsaminaceae). *International Journal of Plant Sciences*, **163**, 907–912.

Hiscox, J.D. & Israelstam, G.F. (1979) A method for the extraction of chlorophyll from leaf tissue without maceration. *Canadian Journal of Botany*, **57**, 1332–1334.

IPCC (2007) Climate change 2007: synthesis report. *Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (eds R.K. Pachauri & A. Reisinger), pp. 1–104. IPCC, Geneva, Switzerland.

Kammenga, J.E., Herman, M.A., Ouborg, N.J., Johnson, L. & Breitling, R. (2007) Microarray challenges in ecology. *Trends in Ecology & Evolution*, **22**, 273–279.

Kawasaki, S., Borchert, C., Deyholos, M., Wang, H., Brazille, S., Kawai, K., Galbraith, D. & Bohnert, H.J. (2001) Gene expression profiles during the initial phase of salt stress in rice. *The Plant Cell*, **13**, 889–905.

Knapp, A., Fay, P.A., Blair, J.M., Collins, S.L., Smith, M.D., Carlisle, J.D., Harper, C.W., Danner, B.T., Lett, M.S. & McCarron., J.K. (2002) Rainfall variability, carbon cycling, and plant species diversity in a mesic grassland. *Science*, **298**, 2202–2204.

Knight, C.A., Vogel, H., Kroymann, J., Shumate, A., Witsenboer, H. & Mitchell-Olds, T. (2006) Expression profiling and local adaptation of *Boechera holboellii* populations for water use efficiency across a naturally occurring water stress gradient. *Molecular Ecology*, **15**, 1229–1237.

Kothapalli, R., Yoder, S., Mane, S. & Loughran, T. (2002) Microarray results: how accurate are they? *Bioinformatics*, **3**, 22.

Lunden, K., Eklund, M., Finlay, R., Stenlid, J. & Asiegbu, F. (2008) Heterologous array analysis in *Heterobasidion*: hybridisation of cDNA arrays with probe from mycelium of S, P or F-types. *Journal of Microbiological Methods*, **75**, *219–224*.

Mckay, J., Richards, J.H. & Mitchell-Olds, T. (2003) Genetics of drought adaptation in *Arabidopsis thaliana*: I. Pleiotropy contributes to genetic correlations among ecological traits. *Molecular Ecology*, **12**, 1137–1151.

Milliken, G.A., Garrett, K.A. & Travers, S.E. (2007) Experimental design for two-color microarrays applied in a pre-existing experiment. *Statistical Applications in Genetics and Molecular Biology*, **6**, Article 20.

Morinaga, S., Nagano, A., Miyazaki, S., Kubo, M., Demura, T., Fukuda, H., Sakai, S. & Hasebe, M. (2008) Ecogenomics of cleistogamous and chasmogamous flowering: genome-wide gene expression patterns from cross-species microarray analysis in *Cardamine kokaiensis* (Brassicaceae). *Journal of Ecology*, **96**, 1086–1097.

Nippert, J.B., Fay, P.A., Carlisle, J.D., Knapp, A.K. & Smith, M.D. (2009) Ecophysiological responses of two dominant grasses to altered temperature and precipitation regimes. *Acta Oecologica-International Journal of Ecology*, **35**, 400–408.

Oleksiak, M.F., Churchill, G.A. & Crawford, D.L. (2002) Variation in gene expression within and among natural populations. *Nature Genetics*, **32**, 261–266.

Ouborg, N.J. & Vriezen, W.H. (2007) An ecologist's guide to ecogenomics. *Journal of Ecology*, **95**, 8–16.

Ozturk, Z.N., Talamé, V., Deyholos, M., Michalowski, C., Galbraith, D., Gozukirmizi, N., Tuberosa, R. & Bohnert, H. (2002) Monitoring large-scale changes in transcript abundance in drought- and salt-stressed barley. *Plant Molecular Biology*, **48**, 551–573.

Purugganan, M. & Gibson, G. (2003) Merging ecology, molecular evolution and functional genetics. *Molecular Ecology*, **12**, 1109–1112.

Smith, C.J. & Gallon, J.R. (2001) Living in the real world: how plants perceive their environment. *New Phytologist*, **151**, 1–6.

Smith, M.D. & Knapp, A. (2003) Dominant species maintain ecosystem function with non-random species loss. *Ecology Letters*, **6**, 509–517.

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

The Gene Ontology Consortium (2000) Gene Ontology: tool for the unification of biology. *Nature Genetics*, **25**, 25–29.

Travers, S.E., Smith, M.D., Bai, J., Hulbert, S.H., Leach, J.E., Schnable, P.S., Knapp, A.K., Milliken, G.A., Fay, P.A., Saleh, A. & Garrett, K.A. (2007) Ecological genomics: making the leap from model systems in the lab to native populations in the field. *Frontiers in Ecology and the Environment*, **5**, 19–24.

Vasemagi, A. & Primmer, C.R. (2005) Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. *Molecular Ecology*, **14**, 3623–3642.

Wang, H., Miyazaki, S., Kawai, K., Deyholos, M., Galbraith, D.W. & Bohnert, H.J. (2003) Temporal progression of gene expression responses to salt shock in maize roots. *Plant Molecular Biology*, **52**, 873–891.

Wellburn, A.R. (1994) The spectral determination of chlorophyll a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. *Journal of Plant Physiology*, **144**, 307–313.

Wolfinger, R., Gibson, G., Wolfinger, E., Bennett, L., Hamadeh, H., Bushel, P., Afshari, C. & Paules, R.S. (2001) Assessing gene significance from cDNA microarray expression data via mixed models. *Journal of Computational Biology*, **8**, 625–637.

Wy, B. (2005) Differential gene expression detection using penalized linear regression models : the improved SAM statistics. *Bioinformatics*, **21**, 1565–1571.

Received 16 June 2009; accepted 16 November 2009
Handling Editor: Joop Ouborg

Supporting Information

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

Appendix S1. List of identities, GenBank accession numbers and BLAST information for 19 200 spots on the Maize microarray SAM 1.1a.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be reorganized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.