PROCEEDINGS B

rspb.royalsocietypublishing.org

Research

Cite this article: Des Marais DL, Guerrero RF, Lasky JR, Scarpino SV. 2017 Topological features of a gene co-expression network predict patterns of natural diversity in environmental response. *Proc. R. Soc. B* **284**: 20170914.
http://dx.doi.org/10.1098/rspb.2017.0914

Received: 27 April 2017
Accepted: 17 May 2017

Subject Category:
Evolution

Subject Areas:
evolution, genetics, plant science

Keywords:
gene by environment interaction, gene regulatory network, *Arabidopsis thaliana*, abiotic stress, drought, cold

Authors for correspondence:
David L. Des Marais
e-mail: desmarais@fas.harvard.edu
Samuel V. Scarpino
e-mail: samuel.v.scarpino@uvm.edu

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.3789997.

THE ROYAL SOCIETY PUBLISHING

Topological features of a gene co-expression network predict patterns of natural diversity in environmental response

David L. Des Marais[1], Rafael F. Guerrero[2], Jesse R. Lasky[3] and Samuel V. Scarpino[4,5]

[1]Arnold Arboretum and Department of Organismic and Evolutionary Biology, Harvard University, 1300 Centre Street, Boston, MA 20131, USA
[2]Department of Biology, Indiana University, Jordan Hall 142, Bloomington, IN 47405, USA
[3]Department of Biology, Pennsylvania State University, 408 Life Sciences Building, University Park, PA 16802, USA
[4]Department of Mathematics and Statistics and Vermont Complex Systems Center, University of Vermont, 210 Colchester Avenue, Burlington, VT 05405, USA
[5]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM, 87501, USA

DLD, 0000-0002-6772-6987; RFG, 0000-0002-8451-3609; JRL, 0000-0001-7688-5296; SVS, 0000-0001-5716-2770

Molecular interactions affect the evolution of complex traits. For instance, adaptation may be constrained by pleiotropic or epistatic effects, both of which can be reflected in the structure of molecular interaction networks. To date, empirical studies investigating the role of molecular interactions in phenotypic evolution have been idiosyncratic, offering no clear patterns. Here, we investigated the network topology of genes putatively involved in local adaptation to two abiotic stressors—drought and cold—in *Arabidopsis thaliana*. Our findings suggest that the gene-interaction topologies for both cold and drought stress response are non-random, with genes that show genetic variation in drought expression response (eGxE) being significantly more peripheral and cold response genes being significantly more central than genes which do not show GxE. We suggest that the observed topologies reflect different constraints on the genetic pathways involved in environmental response. The approach presented here may inform predictive models linking genetic variation in molecular signalling networks with phenotypic variation, specifically traits involved in environmental response.

1. Introduction

Genes neither function nor evolve in isolation. The transcriptional activities of genes are often highly correlated with one another, forming hierarchical gene regulatory networks (GRNs) comprised of functionally related modules [1]. Within GRNs, some genes—'nodes'—have stronger or more interactions—'edges'—than do other genes. Because GRN position can influence the magnitude or pleiotropy of a mutation's effect [2], and because the effect size of mutations is strongly associated with their evolutionary fate [3], properties of GRNs will probably affect selection acting on individual component genes [4]. Advances in high-throughput molecular phenotyping and systems analysis have improved our ability to characterize molecular interaction networks, providing the opportunity to address classic questions about the evolution of genetic interactions.

Two related features of gene regulatory networks might affect the evolution of individual genes within those networks. The first is the widespread observation that genes vary in their number of interacting neighbour genes, even by orders of magnitude [5]. This feature—the centrality or connectivity of a gene—can be measured in many different ways, including the number of directly interacting

genes or the number of paths to other genes that pass through a given gene [6]. The second feature is modularity, i.e. the degree to which the total gene network is composed of distinguishable, functionally related sub-networks of genes, or modules. Modules are often controlled by core proteins or other regulatory factors that regulate the module's activity [7] using shared regulatory motifs among genes within the module [8]. Evidence for the pleiotropic nature of core genes has been found through decades of developmental genetics research which identified putative master regulators of the level, timing and location of expression of tens to thousands of other genes [4,7,9].

Transcriptional regulation by core genes plays an important role in adaptive responses to the environment [10,11] and such genes are often co-regulated as functional modules [12,13]. In some instances the environmental response of particular genes may be characteristic of entire species or kingdoms [14,15]. Considerable genetic variation in transcriptional response to environment—expression Genotype by Environment interaction (eGxE)—has also been identified within species [16]. At the molecular level, eGxE may be controlled by genetic variants acting in *cis*, e.g. by single nucleotide polymorphisms or presence-absence variants in promoter motifs, or by genetic variants acting in *trans*, such as transcription factors, small RNA species, or other regulatory factors upstream of genes showing eGxE. Genetic variants affecting eGxE are of particular interest because GxE represents the mutational substrate for the evolution of environmental response [17,18] and because GxE for fitness is required for local adaptation [19].

In the present study, we use a gene co-expression network [20] to extend earlier work assessing genetic variation in transcriptional activity during acclimation to cold [21] and soil drying [22] in *Arabidopsis thaliana*. We explore two hypotheses for how environmentally responsive regulatory networks evolve and might thereby be involved in local adaptation to environment. The first hypothesis, that eGxE is driven by genetic variation in core transcriptional regulatory proteins, arises from the observation that suites of traits often show high genetic correlation ([23]; in this context, 'traits' could be either individual transcripts or higher-level physiological or developmental phenotypes). Genetic variants in one or a small number of regulatory genes could therefore have considerable downstream consequences on traits and fitness. This model predicts that eGxE genes have relatively high network connectivity and, by extension, be clustered in relatively discrete functional modules. The second hypothesis posits that eGxE is driven by variation in genes located peripherally in transcriptional networks, which are expected to have smaller effect sizes and reduced pleiotropy. Expression variation in peripheral genes could therefore allow natural selection to 'fine-tune' environmental response by changing only a small number of transcriptional or higher-order traits. While these are not mutually exclusive hypotheses (e.g. some eGxE genes for a given environmental contrast might show patterns different from these expectations), their relative importance in nature has not been established.

2. Methods

(a) Data: co-expression network

We used datasets representing gene co-expression relationships in *A. thaliana* generated by Feltus *et al.* ([20]; hereafter 'Feltus network') from 7105 published Affymetrix ATH1 microarray experiments. The first dataset was global (i.e. genome-wide and not restricted to certain functions or pathways), derived from a large adjacency matrix and employing a thresholding algorithm to generate a network containing 3297 nodes and 129 134 edges. These nodes represent 16% of genes on the ATH1. A complication from this global approach arises because of interactions between genotype, expression networks and environment (including ontogeny, tissue or cell type; [24]) across the 7105 microarray experiments, suggesting that analyses on more restricted sub-networks of genes might also be of interest. We generated a community-level dataset using the leading eigenvector method [25] as implemented in the R package igraph v. 1.0.1 [26] on the Feltus network. This approach subdivides the graph into sets of genes that are densely connected among themselves and loosely connected to other parts of the gene co-expression network.

(b) Data: transcriptomic responses to cold and drought stress

We used two published studies on natural variation in transcriptomic response to cold [21] and drought [22]. Both studies used the ATH1 microarray to estimate genome-wide transcript abundance. Each study subjected a diverse panel of 9 [21] or 17 [22] natural accessions to a cold or drought treatment, respectively. Lasky *et al.* [27] re-analysed the cold dataset to match the analyses by Des Marais *et al.* In brief, the two earlier studies [22,27] quantified the effect of genotype, treatment and their interaction on gene expression (through a factorial ANOVA of transcript abundance, with 5% pFDR). From these analyses, we were primarily interested in genes with a significant genotype-by-treatment interaction (i.e. those classified as 'eGxE'), which show a genetically variable response to environmental changes. In the context of this paper, 'stress-responsive' genes are those previously classified as eGxE, as well as genes that show a similar pattern of response across natural accessions ('cold-response' or 'drought-response' genes).

(c) Network methods

We quantified the degree to which a gene was central using four standard network metrics: (i) degree, the raw number of connections; (ii) closeness centrality, the inverse of the average shortest path between the focal gene and all other genes in the network; (iii) betweenness centrality, the number of shortest paths between all pairs of nodes in the network which pass through the focal node; and (iv) eigenvector centrality. Eigenvector centrality, which is closely related to Google's PageRank algorithm [28], measures a node's centrality based on both the node's own position in the network and the position of that node's neighbours. More specifically, a node's eigenvector centrality will be proportional to the average centralities of its neighbours [29]. These metrics were calculated on each of the co-expression datasets described above (the global network of Feltus *et al.* and the communities generated herein) using igraph [26].

(d) Enrichment analyses

We tested whether eGxE genes ($n = 1473$ for drought, $n = 2149$ for cold) exhibited non-random network metrics. Determining whether a gene has a higher value for any of these metrics is not appropriate for parametric statistics. Therefore, we conducted 10 000 permutations to generate null expectations for each contrast. These analyses were executed using custom scripts in R. In order to assess the biological function of co-expression communities, we first identified modules containing the greatest over- and under-representation of eGxE genes for each abiotic stressor using Fisher's exact tests, taking significant communities

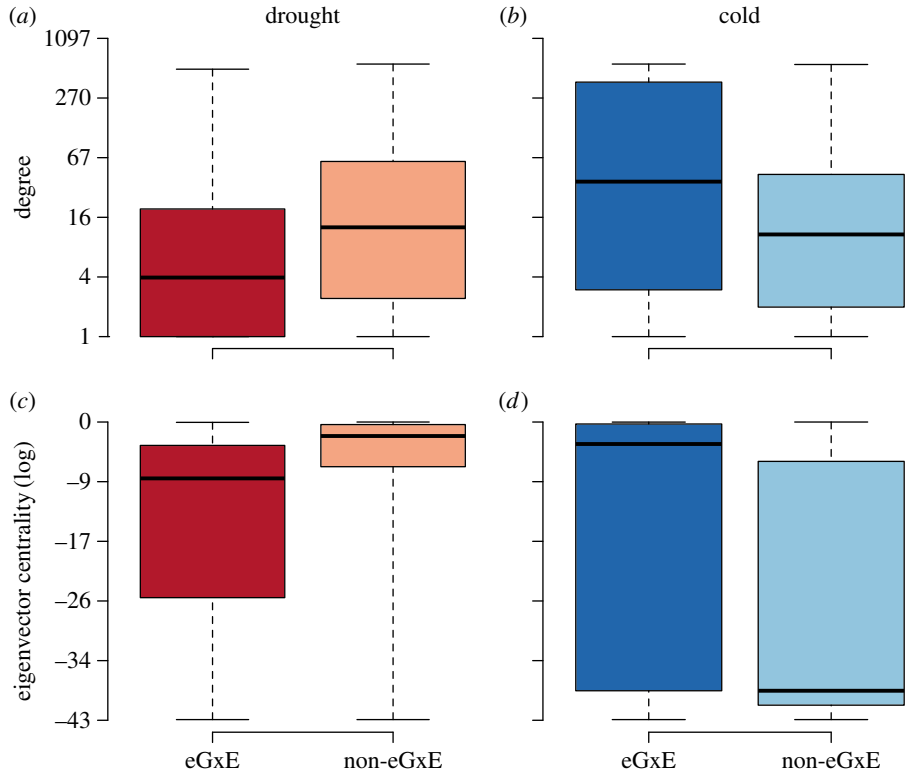

Figure 1. Expression GxE genes are non-randomly distributed in the *Arabidopsis* gene regulatory network. As a group, eGxE genes in drought conditions have significantly lower degree (*a*) and eigenvector centrality (*c*) than do genomic controls. As a group, eGxE genes in cold conditions have significantly higher degree (*b*) and eigenvector centrality (*d*) than do genomic controls. The solid line represents the median, coloured boxes indicate the inter-quartile range, and whiskers mark the entire range of the data. (Online version in colour.)

at a global FDR = 0.05. We then tested gene ontology (GO) term enrichment in each of these modules in AgriGO [30] using default parameters and FDR = 0.05.

All scripts and data used in our analyses are available at https://github.com/scarpino/arab_net.

3. Results

(a) Stress-responsive genes are non-randomly distributed in a transcriptional co-expression network

Both cold and drought eGxE genes were non-randomly distributed in the Feltus network with respect to degree and centrality (qualitatively similar results were recovered with various measures of centrality; see the electronic supplementary material, tables S2 and S3). Drought eGxE genes as a group had fewer connections (median degree for drought eGxE = 4; non-eGxE = 13; figure 1*a*) and were less central (figure 1*b*) than non-eGxE genes used as control. Cold eGxE genes exhibited the opposite effect, having higher degree (median cold: eGxE = 38; non-eGxE = 11; figure 1*c*) and being more centrally located (figure 1*d*) compared to genomic controls. Out of 10 000 permutations, we did not observe a single set of genes with more extreme low (drought) or high (cold) distributions of degree and eigenvector centrality, corresponding to a p-value of 10^{-4}. These results remain robust after accounting for differences in absolute expression level among genes (see the electronic supplementary material, table S1) and after accounting for differences in statistical power to detect eGxE in the two source datasets (see the electronic supplementary material, figures S1 and S2). We also find that the pattern of low centrality for drought eGxE genes is restricted to genes showing rank-changing GxE (see the electronic supplementary material).

Two additional analyses verify the robustness of our results. First, we relaxed the assumption that all genes in the network are equally likely to participate in environmental response. It is possible that the appropriate null distribution should be constructed using nearby genes (e.g. those with strong co-expression) because certain pathways *must* be involved in the phenotypic response to drought or cold [20]. Focusing only on such local modules, and averaging across all detected modules, we find similar results as before. Cold eGxE genes have higher degree (median 34 higher than non-eGxE) and a significantly higher median eigenvector centrality (median 0.14 higher than non-eGxE). Drought eGxE genes have lower degree (median 16 fewer connections than non-eGxE genes) and had a significantly lower eigenvector centrality (median 0.04 below non-eGxE). Second, we assessed the explanatory power of our result by performing iterative out-of-sample model validation. We randomly selected 80% of the genes in the network and constructed a generalized linear model with a binomial error distribution to predict genes as eGxE based solely on their degree and eigenvector centrality. We then predicted the eGxE state for the remaining 20% of genes, recorded the error, and repeated this procedure 1000 times. Assuming a threshold for accurate classification of 5%, we correctly classified 95.4% of genes for cold and 77.0% of genes for drought. These results accommodate classification errors for both eGxE and non-eGxE genes, which means that for cold we correctly classify nearly every gene included in the co-expression network as being eGxE based solely on its degree and eigenvector centrality.

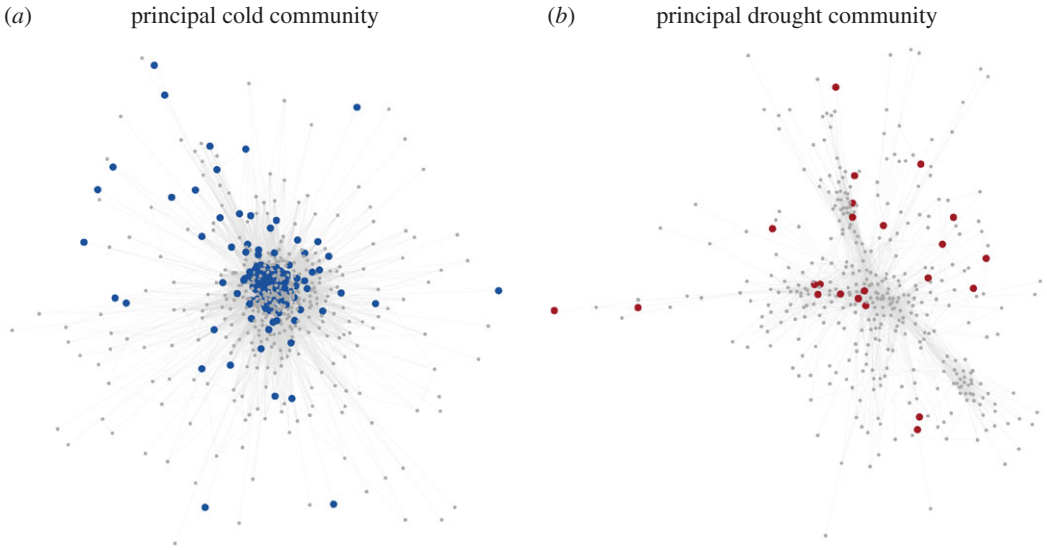

Figure 2. Graphical representation of two sub-communities of the *Arabidopsis* gene regulatory network showing the sub-communities with the highest over-representation of cold eGxE genes (*a*) and of drought eGxE genes (*b*). Coloured circles indicate genes showing significant eGxE at pFDR = 0.05 from [22,27]. (Online version in colour.)

(b) eGxE genes show modular distribution that differs between environments

We next asked whether the non-random distribution of node connectivity of eGxE genes reflects their membership in particular or modules of interacting genes. Both the cold and drought eGxE genes were non-randomly distributed among the modules. For cold, 32.5% of all eGxE genes exist within a single, large module containing 605 genes (figure 2*a*) and an additional 26.5% of cold eGxE genes are found in a second large module containing 425 genes. In contrast, for drought, the two modules with the highest accumulation of eGxE genes together contain only 18% of the total number of eGxE genes (figure 2*b* shows the larger of these two). Moreover, drought eGxE are statistically over-represented in five small modules comprised between 10 and 100 members (figure 3*b*), while cold eGxE genes are clustered in a few large modules (figure 3*a*). The membership of eGxE genes in modules of differing size recapitulates our earlier result: cold eGxE genes tend to be functionally connected to many other genes, while genes involved in drought response tend to be in peripheral network positions.

To test the hypothesis that the modules with diverging patterns of expression eGxE reflect natural variation in function, we took the genes in the two modules with the most over-representation in cold and in drought response and tested for enrichment of GO annotations. We found 116 significant GO terms enriched in the most over-represented cold eGxE module. The top five terms were all related to photosynthesis and related processes, the sixth term is response to abiotic stimulus, and response to cold is 27th (electronic supplementary material, table S4). Altered primary metabolism is frequently observed during cold acclimation, in part because of the accumulation of sugars as cryoprotectants [31,32], so this eGxE may indicate that the sampled accessions modify metabolism during cold response to varying degrees. Interestingly, this same, large module was found to be *under*-represented for drought eGxE genes. Previously, we observed that moderate drought stress results in a slight increase in plant carbon status, possibly related to the development of additional root mass as a drought avoidance strategy [22]. The under-enrichment of eGxE genes in this module suggests that this metabolic response is common to the sampled accessions. We found 89 significant GO terms enriched in the most over-represented drought eGxE module. The top term and many of the subsequent terms were for immune and defence responses (electronic supplementary material, table S5; many genes, particularly kinases, annotated as immune and defence responses also show responses to abiotic stress [33]). Previously, we found that drought eGxE genes showed very few significant functional enrichments using a genome-wide test for statistical enrichment [22], suggesting that the network-informed approach used here may afford additional statistical power to detect functional patterns in these high-dimensional datasets.

Two specific examples, drawn from the module analysis, illustrate the power of our approach. The WKRY33 transcription factor (At2g38470) is a positive regulator of osmotic stress [34] and, in our analysis, this gene is the most highly connected member of community 13 which is over-represented for drought eGxE genes (figure 3). WRKY33 does not, itself, show GxE in the drought dataset. Interestingly, six of the seven eGxE genes in this community have predicted WKRY binding domains in their proximal promoter regions (as determined in an online database; [35]). community 13 may therefore represent one functional type of eGxE: a conserved, stress-responsive transcription factor regulating genes which show variable expression response to the environment. RPS1 (At5g30510) is a second interesting candidate revealed by our analysis. This ribosomal protein shows eGxE in the cold dataset [27] and is a very highly connected member (degree = 567) in the cold eGxE-enriched community 1. *Arabidopsis thaliana* RPS1 regulates synthesis of chloroplast thylakoid membrane proteins [36], though it currently has no known role in cold stress response. Its central position in community 1, which is enriched for cold eGxE genes as well as genes involved in chloroplast function, suggests RPS1 as a candidate for controlling genotype-dependent cold acclimation responses.

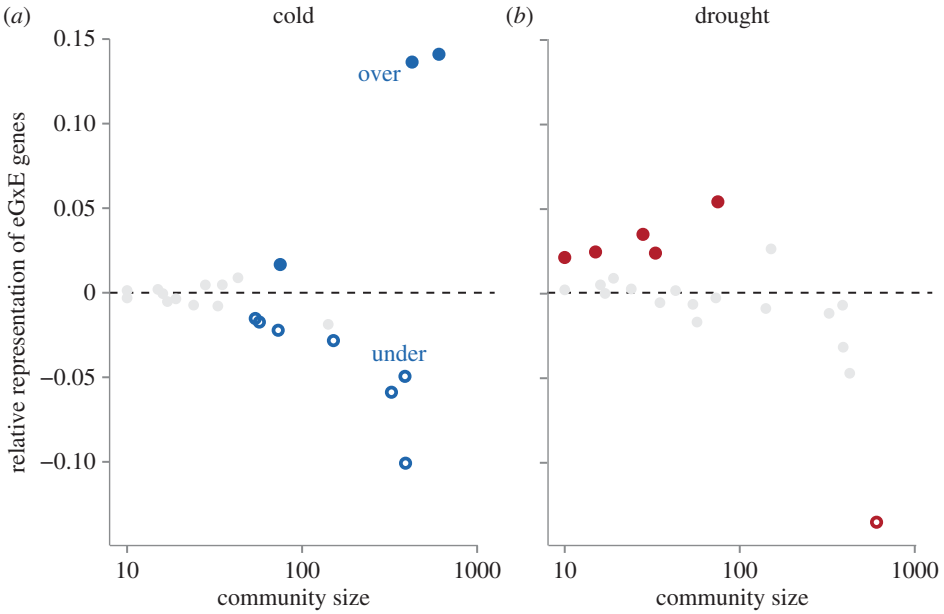

Figure 3. (*a*) Cold and (*b*) drought eGxE genes have different distributions across sub-communities of the gene regulatory network. Cold eGxE genes are over-represented in two large sub-communities, each containing over 400 total genes. Five smaller sub-communities (less than 100 genes) are enriched for drought eGxE genes. The vertical axis is the difference between the observed and expected number of eGXE genes in the community, relative to the total number of genes in the community. Each sub-community in the analysis is shown with a circle (only sub-communities with 10 or more genes are included). Filled circles indicate sub-communities in which eGxE genes are significantly over-represented, while empty circles represent sub-communities with significant under-representation of eGxE genes (Fisher's exact test; FDR = 0.05). (Online version in colour.)

4. Discussion

Previously, we demonstrated an important role of *cis*-regulatory variants underlying diversity of environmental response among natural genotypes of *Arabidopsis* [27]. Natural variation in response to drought showed different genomic patterns than did natural variation in response to cold, suggesting that natural selection may affect different parts of the transcriptional regulatory networks for these two complex traits. Specifically, the proximal promoters of genes showing eGxE for drought had significantly higher nucleotide diversity and significantly higher among-genotype variation in key drought-responsive promoter motifs (abscisic acid responsive elements, ABREs) when compared to genome averages [27]. These earlier observations are consistent with the results presented here: drought eGxE genes are in smaller modules and are relatively lowly connected to other genes, suggesting that genetic variation in expression response to drought is controlled by *cis*-acting variants. This architecture may permit functionally diverse modules to act independently from one another, i.e. showing environmental response in only some genotypes [37]. Our results may also explain why a recent expression QTL (eQTL) study of drought response identified a preponderance of *cis*- acting eQTL and few *trans*- eQTL [38].

Diverse populations and species can acclimate to transient soil drying stress in diverse ways—via changes in growth, transpiration, leaf area-volume ratios, timing of reproduction, cell wall composition, and synthesis of various osmoprotectants and chaperonins, to name but a few [39]. In this model, the modular co-expression architecture observed here reflects functionally distinct transcriptional modules driving different physiological responses among populations. The extent to which such physiological alterations are under independent or common genetic control is presently unknown.

Nucleotide diversity in the proximal promoters of cold eGxE genes is also elevated but, in contrast to drought genes, not to a degree that is statistically significant compared to genome averages [27]. The promoters of cold eGxE genes exhibit lower among-genotype turnover of known cold-responsive motifs compared to genome averages (specifically, the C-Repeat, CRT, motif). Along with our observation that cold eGxE genes are highly connected in the Feltus gene co-expression network, these patterns of sequence diversity suggest that the transcriptional control of eGxE for cold acclimation is driven by genetic variants in upstream regulatory features, such as transcription factors. Natural variants in transcription factors are expected to have larger mutational effect size owing to the regulatory influence of these proteins on downstream genes. The RPS1 gene, described above, is an interesting candidate in this regard.

What drives the apparent difference in the genetic architecture of natural variation in response to drought compared to response to cold? We speculate that, while daily or seasonal drying stress is probably experienced by *A. thaliana* plants to some degree across the species range, severe cold stress is probably only experienced by populations at higher latitudes or altitudes. A recent study demonstrated that multiple, apparently independent, loss of function mutations in key transcriptional regulators of cold-responsive genes are associated with geographical variation in winter temperature across the range of *A. thaliana* [40]. The activity of these CRT-binding transcription factors (CBF) shows a strong positive correlation with the capacity of *A. thaliana* natural accessions to acclimate to cold [21,41,42]. Perhaps cold-associated selective gradients involve sharp transitions (along spatial and climatic gradients) in optimal phenotypes, with only a few fitness optima for cold tolerance—manifested as cold-tolerant, functional CBF versus cold-intolerant, non-functional CBF accessions. By contrast, drought tolerance may be subject to smoother, more gradual selective gradients, for which many smaller-effect variants allow fine-tuning of response.

5. Conclusion

Our results suggest that topological relationships among genes in transcriptional regulatory networks affect how natural populations adapt to the multivariate environment. A promising extension of our approach is to link information regarding the topological features of a given gene—its connectivity, in the case presented here, as well as its membership in particular functional modules—with information associating genetic variants with phenotype from genetic mapping. Such a combined analysis could clarify how putatively functional variants identified via association mapping result in phenotypic variation via cellular and physiological mechanisms [43].

From an applied perspective, understanding how natural variation affects transcriptional regulatory networks may inform decisions about how to improve agricultural performance in challenging environments. We note that breeding for improved performance under soil drying has been quite challenging [44]; our results suggest that targeting specific physiological mechanisms by manipulating genes at the 'tips' of regulatory networks, shown herein to exhibit drought eGxE, may be a more fruitful strategy than targeting central regulatory molecules which may exhibit undesirable pleiotropic effects ('yield drag'; e.g. [45]).

Data accessibility. All data and computer code used in this study can be accessed at https://github.com/scarpino/arab_net.

Authors' contributions. All authors contributed equally to the design of the study. S.V.S. performed the preponderance of statistical analyses. D.L.D. and S.V.S. wrote the paper with contributions from J.R.L. and R.F.G.

Competing interests. We declare we have no competing interests.

Funding. This work is a product of the 'Molecular Networks and Evolution Across Biological Scales' working group supported by the Santa Fe Institute.

Acknowledgements. We thank R. Hopkins, E. Josephs, J. Stinchcombe and two anonymous reviewers for comments on the manuscript. T. Juenger and J. G. Monroe provided stimulating discussion on the genetic architecture of abiotic stress response and expression GxE.

References

1. Wagner GP, Zhang J. 2011 The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. *Nat. Rev. Genet.* **12**, 204–213. (doi:10.1038/nrg2949)
2. Jeong H, Mason SP, Barabasi AL, Oltvai ZN. 2001 Lethality and centrality in protein networks. *Nature* **411**, 41. (doi:10.1038/35075138)
3. Fisher RA. 1930 *The genetical theory of natural selection*. Oxford, UK: Clarendon Press.
4. Erwin DH, Davidson EH. 2009 The evolution of hierarchical gene regulatory networks. *Nat. Rev. Genet.* **10**, 141–148. (doi:10.1038/nrg2499)
5. Guelzim N, Bottani S, Bourgine P, Kepes F. 2002 Topological and causal structure of the yeast transcriptional regulatory network. *Nat. Genet.* **31**, 60–63. (doi:10.1038/ng873)
6. Lin W-h, Liu W-c, Hwang M-j. 2009 Topological and organizational properties of the products of housekeeping and tissue-specific genes in protein-protein interaction networks. *BMC Syst. Biol.* **3**, 32. (doi:10.1186/1752-0509-3-32)
7. Wilkins AS. 2002 *The evolution of developmental pathways*. Sunderland, MA: Sinauer.
8. Wray GA, Hahn MW, Abouheif E, Balhoff JP, Pizer M, Rockman MV, Romano L. 2003 The evolution of transcriptional regulation in eukaryotes. *Mol. Biol. Evol.* **20**, 1377–1419. (doi:10.1093/molbev/msg140)
9. Wagner GP, Pavlicev M, Cheverud JM. 2007 The road to modularity. *Nat. Rev. Genet.* **8**, 921–931. (doi:10.1038/nrg2267)
10. Gibson G. 2008 The environmental contribution to gene expression profiles. *Nat. Rev. Genet.* **9**, 575–581. (doi:10.1038/nrg2383)
11. Weake VM, Workman JL. 2010 Inducible gene expression: diverse regulatory mechanisms. *Nat. Rev. Genet.* **11**, 426–437. (doi:10.1038/nrg2781)
12. Fowler S, Cook D, Thomashow MF. 2007 The CBF cold-response pathway. In *Plant abiotic stress* (eds MA Jenks, PM Hasegawa), pp. 71–99. Chichester, UK: Wiley-Blackwell.
13. Luscombe NM, Babu MM, Yu H, Snyder M, Teichmann SA, Gerstein M. 2004 Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature* **431**, 308–312. (doi:10.1038/nature02782)
14. McCarroll SA, Murphy CT, Zou S, Pletcher SD, Chin CS, Jan YN, Kenyon C, Bargmann CI, Li H. 2004 Comparing genomic expression patterns across species identifies shared transcriptional profile in aging. *Nat. Genet.* **36**, 197–204. (doi:10.1038/ng1291)
15. Carvallo MA, Pino MT, Jeknic Z, Zou C, Doherty CJ, Shiu SH, Chen TH, Thomashow MF. 2011 A comparison of the low temperature transcriptomes and CBF regulons of three plant species that differ in freezing tolerance: *Solanum commersonii*, *Solanum tuberosum*, and *Arabidopsis thaliana*. *J. Exp. Bot.* **62**, 3807–3819. (doi:10.1093/jxb/err066)
16. Des Marais DL, Hernandez KH, Juenger TE. 2013 Genotype-by-environment interaction and plasticity: exploring genomic responses of plants to the abiotic environment. *Annu. Rev. Ecol. Evol. Syst.* **44**, 5–29. (doi:10.1146/annurev-ecolsys-110512-135806)
17. Schlichting CD, Pigliucci M. 1998 *Phenotypic evolution: a reaction norm perspective*. Sunderland, MA: Sinauer Associates.
18. Via S, Lande R. 1985 Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution* **39**, 505–523. (doi:10.2307/2408649)
19. Kawecki TJ, Ebert D. 2004 Conceptual issues in local adaptation. *Ecol. Lett.* **7**, 1225–1241. (doi:10.1111/j.1461-0248.2004.00684.x)
20. Feltus FA, Ficklin SP, Gibson SM, Smith MC. 2013 Maximizing capture of gene co-expression relationships through pre-clustering of input expression samples: an *Arabidopsis* case study. *BMC Syst. Biol.* **7**, 44. (doi:10.1186/1752-0509-7-44)
21. Hannah MA, Wiese D, Freund S, Fiehn O, Heyer AG, Hincha DK. 2006 Natural genetic variation of freezing tolerance in *Arabidopsis*. *Plant Physiol.* **142**, 98–112. (doi:10.1104/pp.106.081141)
22. Des Marais DL, McKay JK, Richards JH, Sen S, Wayne T, Juenger TE. 2012 Physiological genomics of response to soil drying in diverse *Arabidopsis* accessions. *Plant Cell* **24**, 893–914. (doi:10.1105/tpc.112.096180)
23. Lynch M, Walsh B. 1998 *Genetics and analysis of quantitative traits*. Sunderland, MA: Sinauer.
24. Fukushima A, Nishizawa T, Hayakumo M, Hikosaka S, Saito K, Goto E, Kusano M. 2012 Exploring tomato gene functions based on coexpression modules using graph clustering and differential coexpression approaches. *Plant Physiol.* **158**, 1487–1502. (doi:10.1104/pp.111.188367)
25. Newman ME. 2006 Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **74**, 036104. (doi:10.1103/PhysRevE.74.036104)
26. Csardi G, Nepusz T. 2006 The igraph software package for complex network research. *Int. J. Complex Syst.* **1695**, 1–9.
27. Lasky JR, Des Marais DL, Lowry DB, Povolotskaya I, McKay J, Richards JH, Keitt TH, Juenger TE. 2014 Natural variation in abiotic stress responsive gene expression and local adaptation to climate in *Arabidopsis thaliana*. *Mol. Biol. Evol.* **31**, 2283–2296. (doi:10.1093/molbev/msu170)
28. Langville AN, Meyer CD. 2011 *Google's PageRank and beyond: the science of search engine rankings*. Princeton, NJ: Princeton University Press.
29. Newman ME. 2008 The mathematics of networks. *New Palgrave Encycl. Econ.* **2**, 1–12.

30. Zhou D, Xin Z, Yi L, Zhenhai Z, Zhen S. 2010 agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Res.* **38**, W64–W70. (doi:10.1093/nar/gkq310)
31. Thomashow M. 1999 Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 571–599. (doi:10.1146/annurev.arplant.50.1.571)
32. Colton-Gagnon K, Ali-Benali MA, Mayer BF, Dionne R, Bertrand A, Do Carmo S, Charron JB. 2014 Comparative analysis of the cold acclimation and freezing tolerance capacities of seven diploid *Brachypodium distachyon* accessions. *Ann. Bot.* **113**, 681–693. (doi:10.1093/aob/mct283)
33. Marshall A *et al.* 2012 Tackling drought stress: receptor-like kinases present new approaches. *Plant Cell* **24**, 2262–2278. (doi:10.1105/tpc.112.096677)
34. Dubois M *et al.* 2013 Ethylene Response Factor6 acts as a central regulator of leaf growth under water-limiting conditions in *Arabidopsis*. *Plant Physiol.* **162**, 319–332. (doi:10.1104/pp.113.216341)
35. Yilmaz A, Mejia-Guerra MK, Kurz K, Liang X, Welch L, Grotewold E. 2011 AGRIS: the *Arabidopsis* Gene Regulatory Information Server, an update. *Nucleic Acids Res.* **39**, D1118–D1122. (doi:10.1093/nar/gkq1120)
36. Yu HD, Yang XF, Chen ST, Wang YT, Li JK, Shen Q, Liu XL, Guo FQ. 2012 Downregulation of chloroplast RPS1 negatively modulates nuclear heat-responsive expression of HsfA2 and its target genes in *Arabidopsis*. *PLoS Genet.* **8**, e1002669. (doi:10.1371/journal.pgen.1002669)
37. Draghi JA, Parsons TL, Wagner GP, Plotkin JB. 2010 Mutational robustness can facilitate adaptation. *Nature* **463**, 353–355. (doi:10.1038/nature08694)
38. Lowry DB, Logan TL, Santuari L, Hardtke CS, Richards JH, DeRose-Wilson LJ, McKay JK, Sen S, Juenger TE. 2013 Expression quantitative trait locus mapping across water availability environments reveals contrasting associations with genomic features in *Arabidopsis*. *Plant Cell* **25**, 3266–3279. (doi:10.1105/tpc.113.115352)
39. Chaves MM, Maroco JP, Pereira J. 2003 Understanding plant responses to drought: from genes to the whole plant. *Funct. Plant Biol.* **30**, 239–264. (doi:10.1071/FP02076)
40. Monroe JG, McGovern C, Lasky JR, Grogan K, Beck J, McKay JK. 2016 Adaptation to warmer climates by parallel functional evolution of CBF genes in *Arabidopsis thaliana*. *Mol. Ecol.* **25**, 3632–3644. (doi:10.1111/mec.13711)
41. Gehan MA, Park S, Gilmour SJ, An C, Lee CM, Thomashow MF. 2015 Natural variation in the C-repeat binding factor cold response pathway correlates with local adaptation of *Arabidopsis* ecotypes. *Plant J.* **84**, 682–693. (doi:10.1111/tpj.13027)
42. Gery C, Zuther E, Schulz E, Legoupi J, Chauveau A, McKhann H, Hincha DK, Teoule E. 2011 Natural variation in the freezing tolerance of *Arabidopsis thaliana*: effects of RNAi-induced CBF depletion and QTL localisation vary among accessions. *Plant Sci.* **180**, 12–23. (doi:10.1016/j.plantsci.2010.07.010)
43. Rockman MV. 2008 Reverse engineering the genotype-phenotype map with natural genetic variation. *Nature* **456**, 738–744. (doi:10.1038/nature07633)
44. Passioura J. 2010 Scaling up: the essence of effective agricultural research. *Funct. Plant Biol.* **37**, 585–591. (doi:10.1071/FP10106)
45. Kasuga M, Liu Q, Miura S, Yamaguchi-Shinozaki K, Shinozaki K. 1999 Improving plant drought, salt and freezing tolerance by gene transfer of a single stress-inducible transcription factor. *Nat. Biotechnol.* **17**, 287–291. (doi:10.1038/7036)