

Systems genomics of salinity stress response in rice

Reviewed Preprint

v2 • February 7, 2025

Revised by authors

Reviewed Preprint

v1 • August 19, 2024

Sonal Gupta, Simon C Groen, Maricris L Zaidem, Andres Godwin C Sajise, Irina Calic, Mignon A Natividad, Kenneth L McNally, Georgina V Vergara, Rahul Satija, Steven J Franks, Rakesh K Singh, Zoé Joly-Lopez , Michael D Purugganan 

Center for Genomics and Systems Biology, New York University, New York, United States • Department of Nematology and Department of Botany & Plant Sciences, University of California, Riverside, Riverside, United States • Center for Plant Cell Biology, Institute for Integrative Genome Biology, University of California, Riverside, Riverside, United States • Department of Biology, University of Oxford, Oxford, United Kingdom • International Rice Research Institute, Los Baños, Philippines • Department of Biological Sciences, Fordham University, New York, United States • Inari Agriculture Nv, Gent, Belgium • Institute of Crop Science, University of the Philippines, Los Baños, Philippines • New York Genome Center, New York, United States • International Center for Biosaline Agriculture, Dubai, United Arab Emirates • Département de Chimie, Université du Québec à Montréal, Montreal, Canada

 https://en.wikipedia.org/wiki/Open_access

 Copyright information

eLife Assessment

Working with a diverse panel of rice accessions grown in field conditions, this **valuable** study measures changes in transcript abundance, tests for patterns of selection on gene expression, and maps the genetic basis of variation in gene expression in normal and elevated salinity treatments. The manuscript provides **solid** evidence that mean gene expression levels are further from the optimum abundance for more genes under the elevated salinity treatment compared to normal treatment, and that a relatively small number of genes are hotspots that harbor genetic variants which affect broader genome-wide patterns of natural variation in gene expression under high salinity conditions. However, the design, clarity, and interpretation of several statistical analyses can be improved, some opportunities for integration among datasets and analyses could yet be realized, and genetic manipulation is required to confirm functional involvement of any specific genes in regulatory networks or organismal traits that confer adaptation to higher salinity conditions. The manuscript will be of interest to evolutionary biologists studying the genetics of complex traits and a resource for plant biologists studying mechanisms of abiotic stress tolerance.

<https://doi.org/10.7554/eLife.99352.2.sa5>

Abstract

Populations can adapt to stressful environments through changes in gene expression. However, the fitness effect of gene expression in mediating stress response and adaptation remains largely unexplored. Here, we use an integrative field dataset obtained from 780 plants of *Oryza sativa* ssp. *indica* (rice) grown in a field experiment under normal or

moderate salt stress conditions to examine selection and evolution of gene expression variation under salinity stress conditions. We find that salinity stress induces increased selective pressure on gene expression. Further, we show that *trans*-eQTLs rather than *cis*-eQTLs are primarily associated with rice's gene expression under salinity stress, potentially via a few master-regulators. Importantly, and contrary to the expectations, we find that *cis-trans* reinforcement is more common than *cis-trans* compensation which may be reflective of rice diversification subsequent to domestication. We further identify genetic fixation as the likely mechanism underlying this compensation/reinforcement. Additionally, we show that *cis*- and *trans*-eQTLs are under balancing and purifying selection, respectively, giving us insights into the evolutionary dynamics of gene expression variation. By examining genomic, transcriptomic, and phenotypic variation across a rice population, we gain insights into the molecular and genetic landscape underlying adaptive salinity stress responses, which is relevant for other crops and other stresses.

Introduction

Plants face numerous stresses that reduce their growth and fitness, and they have a variety of adaptations to help them deal with environmental challenges (Mareri et al., 2022). For crop plants, stresses such as drought, heat and salinity can greatly limit productivity and agricultural sustainability worldwide (Fahad et al., 2017; Kopecká et al., 2023); indeed, as a direct result of various abiotic stresses, an estimated ~ 50-70% of crop yields are lost (Francini and Sebastiani, 2019). While there is a wealth of information on the physiological responses of plants, including crops, to stress (Zhang et al., 2022), we are still developing an understanding of the underlying genetic responses to stress and the complex interactions between stresses and environmental cues, networks of gene expression regulation, physiological responses and ultimately plant fitness (Cramer et al., 2011; Siddiqui et al., 2021).

Much of the research on the genetic basis of stress responses in crops has focused on identifying individual genes associated with specific responses and/or tolerance traits (Cui et al., 2020; Jiang et al., 2019; Kumar and Wigge, 2010; Laohavisit et al., 2013; Yuan et al., 2014). Stress responses, however, are complex traits that can be influenced by multiple genetic pathways with numerous interconnected genes. Expression levels of genes are controlled by regulatory elements and are often environmentally induced, suggesting a critical role of regulatory divergence in adaptive evolution (De Clercq et al., 2021; Wilkins et al., 2016). Furthermore, changes in gene expression can lead to changes in the transcript network correlational structure, reshaping regulatory pathways (Lea et al., 2019) and influencing the way in which gene expression variation can influence adaptation to stressful environments (Signor and Nuzhdin, 2018; Wagner and Lynch, 2008). As such, there is an increasing need to understand how evolutionary dynamics of gene expression and transcript abundance relate to the genetic underpinnings of stress responses and adaptation in crops (Groen et al., 2020; Ruffley et al., 2023; Siddiqui et al., 2021). One of the most important stresses for many plants is salinity. Soil salinity causes osmotic imbalance between the plant and the soil, which impedes the uptake of water and other key nutrients (Hakim et al., 2014; Munns, 2002), possibly leading to acute ion toxicity (Liang et al., 2018). While some plants are considered halophytic and can thrive in saline environments, other plants are highly sensitive to salts and experience negative effects of salinity even at low concentrations (Carillo et al., 2011). *Oryza sativa* (Asian rice) is one such salt-sensitive crop, facing significant yield loss due to soil salinity (Hussain et al., 2017; Melino and Tester, 2023). Rice responds to salinity stress by adjusting physiological and biochemical processes involved in osmotic and ion homeostasis, nutritional balances and oxidative stress (Castillo et al., 2007; Miller et al., 2010; Munns and Tester, 2008; Qin and Huang, 2020; Wang et al., 2012). Although studies have identified multiple genes associated with these processes (for reviews, see Ponce et al., 2021 and Liu et al., 2022), we still lack a systems-level understanding of how gene expression variation mediates responses to salinity stress and evolves at the molecular level.

Here, we use an integrative system genomics approach to comprehensively dissect the genome-wide molecular and phenotypic response to salinity stress in rice. This study builds on prior work by our group that examined selection on gene expression in rice in response to normal and dry conditions (Čalić et al., 2022 [↗](#); Groen et al., 2021 [↗](#), 2020 [↗](#)). Using genomic, transcriptomic, and phenotypic datasets obtained from 130 diverse accessions of rice subjected to moderate levels of salinity stress, we (i) explore the selection on gene expression variation under salinity stress, (ii) dissect the genetic architecture of gene expression variation under saline conditions, and (iii) identify genes, molecular pathways, and salt stress response traits as well as associated trade-offs in the saline environment. We demonstrate that salinity stress induces increased selective pressure on gene expression, and we identify variation in biological processes and physiological traits that is beneficial and detrimental to plants in a saline environment, providing novel insights into the molecular landscape underlying an adaptive response to excess salt. We integrate these datasets with genomic sequence information to elucidate the genetic architecture and regulatory networks governing rice's response to salinity stress.

Results

Variation in transcript abundance

To understand the microevolutionary dynamics of gene expression variation under salinity stress, we first investigated variation in transcript abundance in 130 accessions of *O. sativa* ssp. *indica*. We conducted a field experiment in the dry season of 2017 (January-May) at the International Rice Research Institute in Los Baños, Laguna, Philippines. Three replicates of each accession were planted separately in a normal wet paddy field as well as a similar field in which plants were exposed to moderate salinity stress (salt levels maintained at 6 dSm⁻¹) maintained until maturity. Average fecundity was significantly lower in the saline field than in normal conditions (**Fig. 1a** [↗](#); two-tailed paired t-test $P = 1.658 \times 10^{-8}$) with most of genotypes having significantly lower fecundity in the saline field ($n = 94$; one-sample proportion test $P = 3.639 \times 10^{-7}$).

To examine gene expression in the field, mRNA levels were measured in leaf blades from 780 plants (130 accessions in triplicates for each environment) via 3'-end-biased mRNA sequencing (Meyer et al., 2011 [↗](#)). Leaf blades were sampled 38 days after sowing (DAS), corresponding to 7 days of plants being exposed to stress in the saline field. Population variance in gene expression of 18,141 widely expressed transcripts was partitioned into genotype (G), environment (E), and genotype $\times$ environment (G $\times$ E) effects using a two-way mixed analyses of variance (ANOVA) with environment as a fixed effect and genotype and genotype $\times$ environment as random effects (Supplementary Table 1). At a conservative false discovery rate (FDR) of 0.001, all but 3 transcripts displayed a significant genotype effect, indicating that expression levels of most transcripts are heritable. This effect of genotype is reflected in the broad-sense heritability (H^2) distribution of gene expression levels (**Fig. 1b** [↗](#)), which had a median value $H^2 = 0.53$ (range of 0.012 - 0.987; Supplementary Table 1). In addition to the high levels of heritability, 16,371 transcripts had a significant G $\times$ E term, indicating that for many transcripts, genotypes showed heritably different levels of expression in different environments. The significant G $\times$ E indicates genetic variation for plasticity, indicating that this plasticity can evolve. Although we see a widespread heritable plastic response in gene expression, we found no evidence of genotype-dependent plasticity in fitness (G $\times$ E for fitness $P = 0.49$; **Fig. 1a** [↗](#)), indicating that the G $\times$ E of transcripts does not translate to the complex trait of fitness. This could be due to a combination of factors, like gene interactions (pleiotropy and epistasis) leading to little to no effect on fitness, or environment specific genotype-dependent gene regulation (environment specific eQTLs). Furthermore, only a relatively small number of transcripts (254) showed significant variation due to E, indicating genotype-independent plasticity in gene expression for only a small fraction of genes.

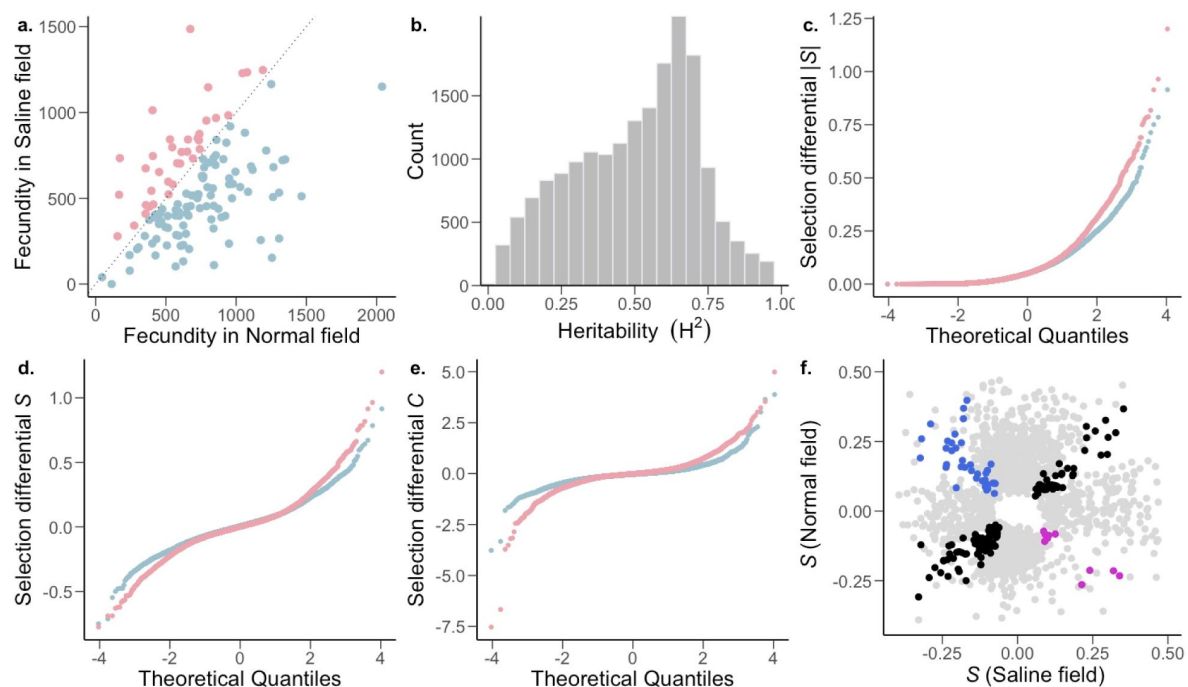


Figure 1

The strength and pattern of selection on heritable gene expression.

a, The *O. sativa* ssp. *indica* populations showed higher average fitness in the normal (blue) and saline (pink) field (two-tailed paired t-test $P = 1.658 \times 10^{-8}$) and fitness further showed a significant effects of genotype (G) and environment (E); genotype $\times$ environment (G $\times$ E) was not significant. Analysis of variance (ANOVA) [G and E ($P < 0.001$), G $\times$ E ($P = 0.49$)]; $n=130$ accessions. b, Broad-sense heritability (H^2) distribution of *Oryza sativa* spp. *indica* transcripts. Two-way ANOVA, genotype FDR-adjusted $q < 0.001$, $n = 130$ accessions. c-e, The strength of linear selection $|S|$, linear selection differentials (S), and quadratic selection differentials (C) for genome-wide gene expression in normal (blue) and saline (pink) conditions. X-axes represent a theoretical quantile for normal distribution with mean=0 and standard deviation=1. f, Conditionally neutral (light gray), and antagonistically pleiotropic transcripts (blue and magenta represent beneficial expression in normal and saline conditions, respectively). Black represents transcripts experiencing selection in the same direction in both environments (expression is beneficial or detrimental in both environments).

Selection on gene expression

To identify transcripts associated with high fitness in normal and saline environments, we measured the strength of selection on gene transcript levels. We did this using phenotypic selection analysis (Lande and Arnold, 1983 [↗](#)), taking the total number of filled rice seeds/grains (total fecundity) as a proxy for fitness (Groen et al., 2020 [↗](#)). We estimated the linear (S) and quadratic (C) selection differentials, which are estimates for directional (negative or positive S) selection, and stabilizing (negative C) or disruptive (positive C) selection. We calculated the raw (S and C), variance-standardized (S_s and C_s), and mean-standardized differentials (S_m and C_m). We found that both mean and variance of transcript expression vary significantly between conditions (Mann-Whitney U -test, mean and standard deviation: $P < 2.2 \times 10^{-16}$), which means that the interpretation of the strength of selection based on either of these differentials could be misleading (Supplementary Table 2). To overcome this, and given that we were interested in selection on gene expression, which were all measured in the same units, we used the non-standardized raw selection differentials for all downstream analyses. We also report the mean- and variance-standardized selection coefficients (Supplementary Table 3). From hereon we use positive and negative directional selection to represent the case where higher fitness is associated with increase and decrease in trait value, respectively. Stabilizing and disruptive selection similarly represent the situations where higher fitness is associated with average vs extreme value of the trait, respectively. In contrast, we limit the use of positive and purifying selection to their molecular evolution definition, that is to represent the case where increase in fitness is associated with derived and ancestral allele, respectively. Lastly, balancing selection refers to the circumstance where multiple alleles are maintained in the population which is associated with increased fitness.

Most transcripts were (nearly) neutral ($|S| < 0.1$) in both the normal and saline environments (**Fig. 1c** [↗](#)). Additionally, within the normal environment, there was a general trend towards stronger positive (9,654 transcripts with $S > 0$) compared to negative (8,415 transcripts with $S < 0$) selection on gene expression, indicating that higher expression of transcribed genes is associated with greater fitness (**Table 1** [↗](#), Mann-Whitney U -test, normal: $P = 3.95 \times 10^{-15}$). In comparison, though the strength of positive directional selection was higher than that of negative directional selection in the saline field, higher expression of most of the transcribed genes is associated with lower fitness (**Table 1** [↗](#), Mann-Whitney U -test, saline: $P = 0.0068$). Although no transcripts cleared the Bonferroni correction threshold in saline conditions, under normal conditions seventeen transcripts cleared the threshold, thirteen of which were under positive directional selection (Supplementary Table 3).

We also found that a high proportion of transcripts experienced stabilizing selection ($C < 0$) in both normal and saline environments. This effect was pronounced under normal conditions, with the strength of stabilizing selection being stronger than that of disruptive selection (**Table 1** [↗](#); Mann-Whitney U -test, normal: $P < 2.2 \times 10^{-16}$). In contrast, within the saline environment, there was no detectable difference between the strength of stabilizing and disruptive selection (Mann-Whitney U -test, normal: $P = 0.514$), due to a higher proportion of transcripts experiencing stronger disruptive selection. This indicates that in the saline environment, as compared to normal conditions, an increase in fitness is associated with extremes in transcript abundance. Comparing the distribution of selection differentials between environmental conditions, we found that selection was stronger in the saline field compared to the normal wet paddy field (**Table 1** [↗](#); **Fig. 1d-e** [↗](#)). This was true for the overall strength of directional selection (Mann-Whitney U -test, $P=0.012$), as well as for negative directional ($P=8.27 \times 10^{-6}$ 0.009), stabilizing ($P < 2.2 \times 10^{-16}$) and disruptive ($P < 2.2 \times 10^{-16}$) selection, but not positive directional selection ($P=0.735$).

		Control	Salt
Median $ S $		0.0501	0.0507
$S > 0$	# Transcripts	9654	8885
	Median S	0.053	0.051
$S < 0$	# Transcripts	8415	9133
	Median S	-0.047	-0.050
$C > 0$	# Transcripts	7175	8713
	Median C	0.077	0.115
$C < 0$	# Transcripts	10894	9304
	Median C	-0.096	-0.111

S and C represent the linear and quadratic selection differentials.

Table 1

Summary statistics of selection on gene expression.

Gene expression trade-offs

Next we compared the proportions of genes showing an opposite direction of selection in across environments (antagonistic pleiotropy-AP) and those showing selection in only one environment (conditional neutrality-CN)(Anderson et al., 2011 [↗](#)). Since the detection of CN relies on *P*-value being significant in only one environment, compared to the detection of AP which relies on *P*-value being significant in both environments, this introduces a bias towards the detection of CN. To account for this inherent bias, we used a more stringent *P*-value cutoff to define CN (a transcript with $P < 0.025$ in one environment and $P > 0.05$ in the other environment) in comparison to AP (a transcript with $P < 0.05$ in both environments and opposite directionality of *S*) (Groen et al., 2020 [↗](#); Siddiqui et al., 2021 [↗](#)). We found that 10.80% of the transcripts showed selection patterns consistent with CN, while only 0.28% of transcripts showed AP (Fig. 1f [↗](#); Supplementary Table 4). The proportion of transcripts showing CN was greater than expected by chance and greater than the proportion showing AP (two-tailed proportion z-test $P < 2.2 \times 10^{-16}$). These results are consistent with a lack of trade-offs in which the expression of a gene is favored in one environment and disfavored in another.

Among the 51 AP transcripts, increased expression of 38 was beneficial only in normal conditions (higher expression associated with higher fitness in normal conditions, but lower fitness under saline conditions). Gene ontology (GO) term analyses of these 38 transcripts indicated that many of these genes were involved in photosynthesis and metabolic processes (Supplementary Fig. S1 [↗](#); Supplementary Table 4). This is consistent with an observed reduction of photosynthesis in rice under salinity stress (Radanielson et al., 2018 [↗](#); Tsai et al., 2019 [↗](#)). We then wondered whether any difference underlie gene regulation of these AP transcripts relative to the non-AP transcripts, potentially indicating the genetic basis for this trade-off. We identified single nucleotide polymorphisms (SNPs) associated with transcript expression levels in each environment separately using whole-genome polymorphism data (expression quantitative trait loci [eQTL] analyses; see below). We found SNPs regulating expression of two photosynthesis related AP transcripts (*PSAN* and *CRR7*) only in the normal environment (Supplementary Fig. S2 [↗](#)). We further looked at the 13 AP transcripts that were beneficial only in the saline environment (Supplementary Table 4). Although there was no significant GO enrichment for these transcripts, among them we identified a cyclophilin-encoding transcript (*OsCYP2*), which has been shown to confer salt tolerance in rice (Lee et al., 2015 [↗](#); Roy et al., 2022 [↗](#); Ruan et al., 2011 [↗](#)). However, no SNP was associated with the expression of this gene in either normal or saline conditions.

Biological processes under selection

To investigate the broader biological processes associated with differential selection (strong directional selection in only one environment), we ranked all GO biological processes by their median directional selection strength in each environment and identified the processes with significantly stronger selection relative to their respective environment. We identified 13 and 18 processes that were under strong differential selection in normal and saline conditions, respectively (Supplementary Table 5; Fig. 2a [↗](#)). Processes primarily involved in various aspects of growth and defense were under stronger selection in normal conditions, whereas processes associated with regulation of flowering, cell cycle control and reproduction showed stronger selection under saline conditions (Fig. 2a [↗](#)). This provides insight into the specific biological processes related to changes in flowering time and reduced yield, both of which have been associated with salinity stress in multiple species (Chang et al., 2019 [↗](#); Li et al., 2007 [↗](#); Van Zandt and Mopper, 2002 [↗](#); R. Zhang et al., 2022 [↗](#)). Furthermore, studies have found that the osmotic stress induced by salinity causes a reduction in the cyclin-dependent kinases (CDKs) responsible for cell-cycle transitions (G1/S and G2/M)(Ma et al., 2015 [↗](#); Schuppler et al., 1998 [↗](#); West et al., 2004 [↗](#)). Aligned with this, our study also supports the notion that salinity stress affects cell cycle regulation, and leads to reduced growth and reproduction.

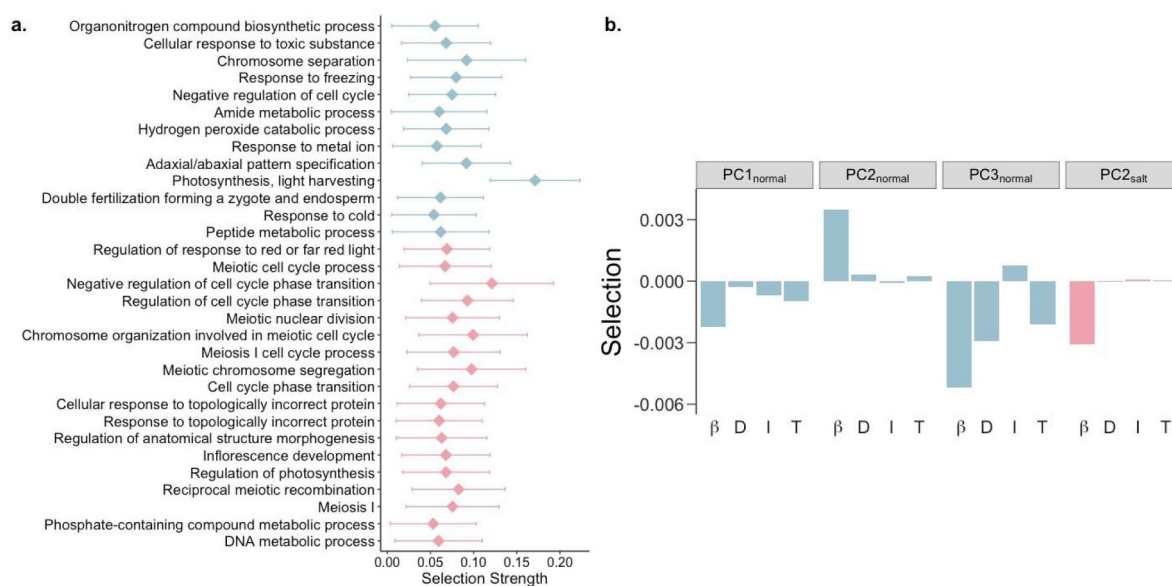


Figure 2

Biological processes and pathways with differential responses to selection under saline conditions.

a, GO biological processes under stronger selection in normal (blue) and saline conditions (pink). b, Linear selection gradients (β), along with direct (D), indirect (I) and total (T) responses to selection on suites of transcripts in normal (blue) and saline conditions (pink).

Gene expression usually operates within the context of robust gene interaction/regulatory networks (Amiri et al., 2018 [↗](#); Israel et al., 2016 [↗](#); Ko and Brandizzi, 2020 [↗](#)). Selection acting on these interacting genes is one of leading causes of indirect selection, which can constrain the response of a population to selection on gene expression (Agrawal and Whitlock, 2010 [↗](#); Groen et al., 2020 [↗](#); Kondrashov and Houle, 1997 [↗](#)). To examine this phenomenon, we identified suites of correlated transcripts using principal component (PC) analysis on genome-wide gene expression levels. We then estimated the linear (β) and quadratic selection (γ) gradients for PCs explaining over 0.5% of variance in each environment (Supplementary Table 6). Although quadratic selection was generally weak, linear selection on some PCs showed significant directional selection (Supplementary Table 7).

Using the breeder's equation (Falconer and Mackay, 1996 [↗](#)), we predicted the response to selection on these PCs to examine the constraints on a population's microevolutionary response to selection on gene expression. We found that over half of PCs (7 of 11 and 7 of 12 PCs in normal and saline conditions, respectively) displayed opposite signs of direct and indirect responses to selection (Supplementary Table 7). This finding contrasts with prior work in rice under dry and wet conditions that found a lack of constraint in response to selection for most traits except seed size because direct and indirect responses to selection were largely similar (Ćalić et al., 2022 [↗](#)), indicating that different types of stress may either constrain or facilitate the response to selection.

We found PCs enriched for metabolic pathways and biosynthesis of phenylpropanoids and secondary metabolites to be under selection in the normal environment (Fig. 2a [↗](#); Supplementary Fig. S3 [↗](#); Supplementary Table 7). Different transcripts involved in these pathways were under positive and negative directional selection which may act to keep these pathways in a steady-state. Interestingly, circadian rhythm was found to be under positive directional selection, with an overall positive response to selection, in the saline environment (Fig. 2a [↗](#); Supplementary Fig. S3 [↗](#); Supplementary Table 7). This is in alignment with the role of circadian clock genes in conferring salt tolerance (Kim et al., 2013 [↗](#); Wei et al., 2021 [↗](#); Xu et al., 2022 [↗](#)), and indicates a tentative increase in expression of circadian clock genes with continuous exposure to soil salinity.

Salinity stress induces decoherence

Gene expression levels are generally correlated, but these correlations can be perturbed by environmental stresses, a phenomenon that is termed decoherence (Lea et al., 2019 [↗](#)). Such decoherence has been demonstrated in humans and primates (Lea et al., 2019 [↗](#); Pu et al., 2022 [↗](#); Watowich et al., 2022 [↗](#)), but little is known about how stress alters the correlation structure of specific transcripts pairs, and functional groups of transcripts, in plants. To examine decoherence in rice, we utilized the recently developed CILP (Correlation by Individual Level Product) method (Lea et al., 2019 [↗](#)), which detects the systematic loss of correlation in gene expression among individuals. Since CILP calculates product correlations for all possible pairs of genes, we used transcripts with selection strengths greater than 0.1 ($|S| > 0.1$) in at least one environment with expression greater than 0 in at least 50% of individuals (2,051 transcripts; Supplementary Table 8) to reduce data dimensionality.

We found that the correlation structure for gene expression was broadly similar across the 2.10 M unique transcript pairs, but > 29,000 transcript pairs (involving 1,742 unique transcripts) showed significantly different correlations between normal and saline conditions (FDR < 5%) (Fig. 3a [↗](#); Supplementary Table 8). This change in correlation structure indicates possible divergence in gene interactions between environments, which may presage a restructuring of gene networks. GO term enrichment analysis of the transcripts that show decoherence with significant pairs greater than the median (median significant pair per transcript = 12, $n = 853$) highlighted important pathways related to plant responses and potentially tolerance to excess salt (Fig. 3b [↗](#)). For instance, circadian rhythm genes like *OsPRR37* and *GIGANTA* (*GI*), which have been shown to confer salt tolerance in rice and Arabidopsis (Kim et al., 2013 [↗](#); Wei et al., 2021 [↗](#)), differed

significantly in their interactions between the two environments. Similarly, to manage the energy requirements of salt stress, the relative abundance of metabolites involved in energy producing pathways like glycolysis, tricarboxylic acid (TCA) cycle and various other metabolic processes have been shown to be altered as an early response to salt stress (Che-Othman et al., 2017 [↗](#); Jacoby et al., 2011 [↗](#); Yang and Guo, 2018 [↗](#)). Additionally, aligned with our results, multiple studies have reported a positive correlation between salt tolerance and levels of secondary metabolites and amino acids with osmoprotectant properties associated with lowering osmotic stress (Krishnamurthy and Bhagwat, 1990 [↗](#); Petrusa and Winicov, 1997 [↗](#); Tari et al., 2010 [↗](#)). Our results suggest that salt exposure induces decoherence of gene expression in some transcripts, that this decoherence results from a restructuring of the gene expression network, and this restructuring can allow salt stress tolerance, providing a potential molecular mechanism underlying this tolerance.

Selection on organismal traits

In addition to gene expression data, we also collected phenotypic data for 13 organismal traits in normal and saline conditions, which provides us with the opportunity to examine how salt stress influences complex phenotypes and identify connections between variation in traits with fitness consequences and underlying patterns of gene expression. Since these traits were measured on different scales, we estimated variance-standardized selection gradients on these traits, focusing only on seven traits that were not strongly correlated to limit the contribution of indirect selection (Pearson correlation coefficient < 0.6; **Supplementary Fig. S4** [↗](#)). We identified three traits – leaf osmotic potential (LOP), chlorophyll a content (Chl_a), and flowering time (FT) – that displayed different selection patterns in normal versus saline environments (**Fig. 4** [↗](#); Supplementary Table 9). Leaf osmotic potential is a key trait associated with water transport in plants, and which decreases with increasing salinity, leading to low water uptake by the plant. This trait was under positive directional selection in the control environment but under stabilizing selection under saline conditions, which suggests that an optimal LOP is important under salt stress (**Fig. 4** [↗](#)). Furthermore, we found Chl_a content at the reproductive stage to be under negative directional selection in saline conditions. Although ion toxicity has previously been shown to reduce chlorophyll content (Ashraf and Bhatti, 2000 [↗](#); Taïbi et al., 2016 [↗](#)), our results showed that this reduction is associated with increased survival and reproductive fitness, consistent with the general trend for reduced photosynthesis under salinity stress. We also found that FT was under positive directional selection (selection for later flowering) in normal conditions but under stabilizing selection in saline conditions (**Fig. 4** [↗](#)). Moreover, FT was significantly reduced under saline conditions (two-tailed paired t-test $P = 0.0012$; **Supplementary Fig. S5** [↗](#)), implying that salt stress selects for an earlier flowering compared to that in the normal wet paddy.

Genetic architecture of gene expression variation

To dissect the genetic architecture of gene expression variation, we identified expression quantitative trait loci (eQTLs) and examined whether and how selection acts on these eQTLs. We identified *cis*- and *trans*-eQTLs (FDR < 0.001) regulating expression of 3,065 and 3,277 genes in normal and salinity stress conditions, respectively (Supplementary Table 10). The median number of *cis*- and *trans*-eQTLs regulating the expression of a gene were similar between environments (median eQTL per gene *cis*- normal=14, *cis*-saline=13, *trans*-normal=5, *trans*-saline=4). We observed that 49.64% (29,623 of a total 59,669) of *cis*-eQTLs were common in both environments, as compared to 18.62% (25,528 of a total 137,100) of *trans*-eQTLs; this result was robust to FDR cutoff (FDR of 0.01 and 0.05). This is consistent with previous observations from studies on other species that have found 48%-77% overlapping *cis*-eQTLs and 9%-60% common *trans*-eQTLs across environments (Smith and Kruglyak, 2008 [↗](#); Snoek et al., 2012 [↗](#); Sterken et al., 2023 [↗](#)). Moreover, comparing the effect sizes of the two categories of eQTLs, we found that *trans*-eQTLs explain more variation in transcript abundance as compared to *cis*-eQTLs in both environments (two-tailed t-test $P < 10^{-16}$; mean effect size: *cis*-normal = 0.77, *trans*-normal = 1.04, *cis*-saline = 0.80, *trans*-saline

Figure 3

Salinity stress induces regulatory decoherence.

a, Pearson correlation coefficients between pairs of transcripts ($|S| > 0.1$ and expression greater than 0 in at least 50% individuals) in normal (x-axis) and saline conditions (y-axis). Pink and blue represent pairs with correlation stronger in saline and normal conditions, respectively; gray represents correlation that is not significantly different between conditions. b, Enrichment of transcripts with significant pairs greater than the median (median significant pair per transcript = 12, $n = 853$) involved in regulatory decoherence post salt exposure.

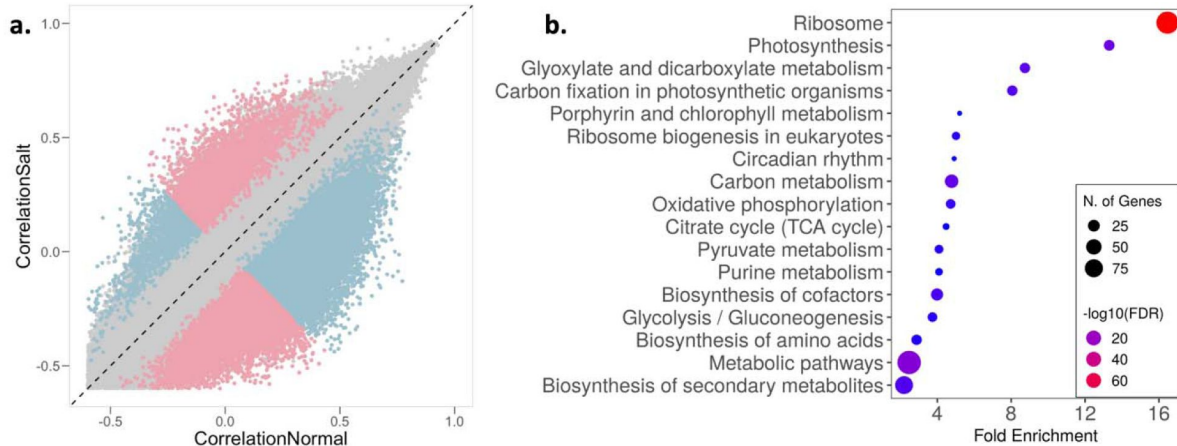
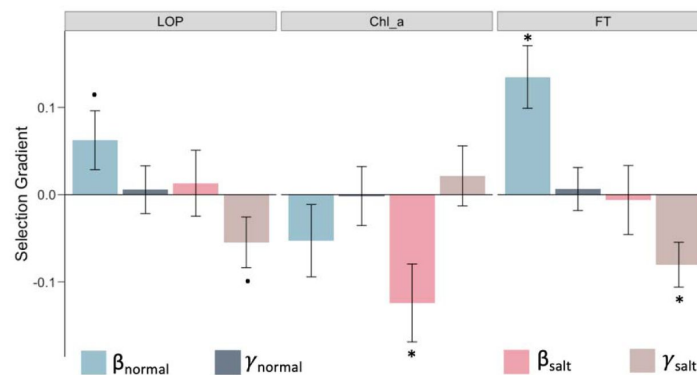


Figure 4

Traits with different selection profiles under salt stress.

Linear (β) and quadratic (γ) selection gradients on the traits LOP (leaf osmotic potential), Chl_a (chlorophyll a content), and FT (flowering time). Error bars represent mean $\pm$ SE; dots and asterisks indicate significance of selection-gradient at two-sided unadjusted $P < 0.1$ and $P < 0.05$.



= 1.08). This further indicates that *trans*-eQTLs might be more environment-specific than *cis*-eQTLs. We tested this explicitly by identifying loci showing gene-environment interaction ($G \times \text{eQTL}$) and found that at FDR of 0.05, *cis*-eQTLs constituted merely 0.28% (142 of 50718) of the total identified $G \times \text{eQTL}$ s.

We identified eQTL hotspots, which are regions of the genome that are associated with expression variation of a large number of genes (Qu et al., 2018). These regions can occur either due to low amounts of recombination (high linkage disequilibrium-LD) or because they contain master-regulators that pleiotropically control expression of multiple functionally-associated genes (Hammond et al., 2011; Kliebenstein, 2009; West et al., 2007). To account for LD and identify regions likely to contain master-regulators, we chose to focus on the subset of genes regulated by the lead-SNPs (SNP with the most significant association) in a given 100-kb region. Through this approach, we identified 11 *trans*-eQTL hotspots (number of unique genes > 30) in saline conditions, but none in normal conditions (Supplementary Fig. S6; Supplementary Table 11). These results indicate that natural variation in gene regulation under stress conditions may be dependent on a few master-regulators, as has been shown for drought stress in rice and maize (Kuroha et al., 2017; Liu et al., 2020). Interestingly, one of the hotspots, Chr7: 25.9-26.0Mb, influenced the expression of a disproportionately high number of genes (191 genes) and contains the gene *OsFLP*, a R2R3 myb-like transcription factor that regulates stomatal development (Wu et al., 2019). *OsFLP* has recently been associated with salt tolerance in rice (Qu et al., 2022). It has been shown previously that when a gene is regulated via both *cis* and *trans* factors, their effects tend to drive target gene expression in opposite directions, canceling the combined effect on expression. This is commonly referred to as *cis-trans* compensation, which is thought to arise due to stabilizing selection that maintains similar gene expression over evolutionary timescales in the face of new mutations (Coolon et al., 2014; Goncalves et al., 2012; Landry et al., 2005; Lovell et al., 2018). Contrary to this expectation, we found that more than half genes appear to have a reinforcing (*cis* and *trans* effect in the same direction) rather than a compensatory pattern (two-tailed proportion z-test $P = 1.65 \times 10^{-11}$; Fig. 5a; Supplementary Table 12). Among the 524 *cis-trans* co-occurring genes in saline conditions, 317 were reinforcing (~60.5%) and 207 were compensatory (39.5%). This pattern held for normal conditions as well (Supplementary Fig. S7a; Supplementary Table 12). Since these reinforcing genes have *cis* and *trans* variants acting in the same direction, it indicates that the expression of these genes are expected to either increase or decrease over evolutionary timescales, indicating the presence of directional selection.

We hypothesized that this excess of reinforcing rather than compensatory effects could be driven by the extensive diversity among rice landraces that arose from crop diversification following domestication (Meyer and Purugganan, 2013). To test this hypothesis, we examined whether genes under *cis-trans* reinforcement showed evidence of higher inter-variational variation in gene expression (population-wide variance between mean accession expression levels) as compared to genes under *cis-trans* compensation. Supporting our hypothesis, we found significantly higher inter-variational expression variation for genes under *cis-trans* reinforcement in saline (Fig. 5b; Mann-Whitney U -test = 0.002; mean-inter-variational expression variation compensating = 1.12, reinforcing = 1.27) and in normal (Supplementary Fig. S7b) conditions. Furthermore, it has been suggested that *cis-trans* compensation/reinforcement can arise due to the genetic fixation of compensating/reinforcing *trans*-regulatory variants, potentially due to epistasis (McManus et al., 2014; Signor and Nuzhdin, 2018), which could lead to elevated linkage disequilibrium (LD) between the *cis*- and *trans*-regulatory variants acting on a gene. This was indeed the case: estimated LD between all pairs of *cis*- and *trans*-regulatory variants for the identified compensating and reinforcing genes was significantly higher as compared to the background LD (mean r^2 compensating: 0.33, reinforcing: 0.36, background = 0.056; two-tailed permutation test $P = 0.0019$).

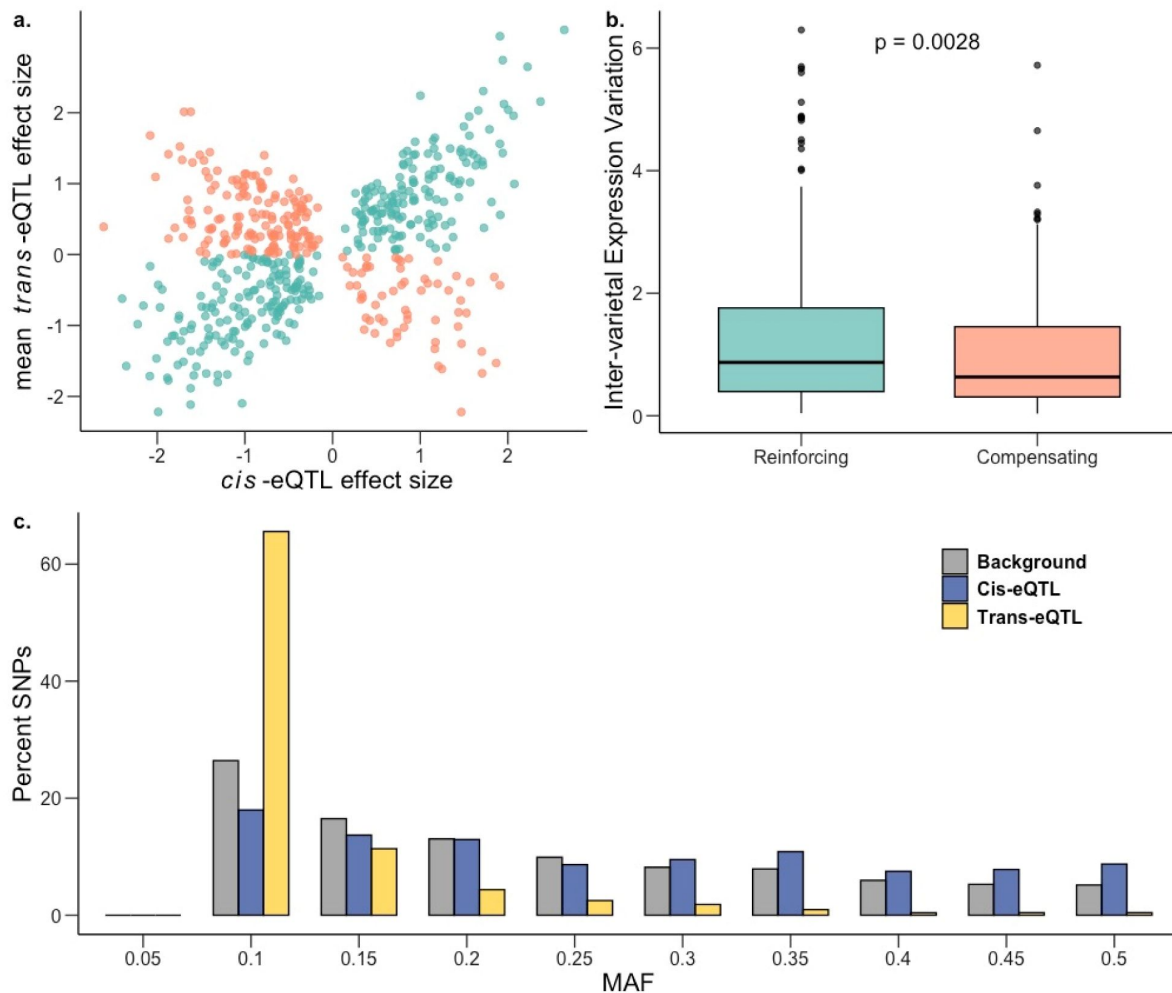


Figure 5

Genetic architecture of gene expression variation during salt stress.

a, Effect sizes of genes with both *cis* and *trans* factors under saline conditions showing excess of reinforcing cis-trans (teal) in comparison to compensating cis-trans (salmon). b, Inter-variatal variation in gene expression for genes under compensating control is significantly lower than for those under reinforcing control; one-sided Mann-Whitney $P = 0.0028$. c, Frequency distribution of MAF (minor allele frequency) for *cis*-eQTLs (blue) and *trans*-eQTLs (yellow) in saline conditions against the genome-wide background (gray).

To examine the pattern of past selection acting on *cis*- and *trans*-eQTLs (before, during and after rice domestication, including the crop diversification phase) we plotted the folded site-frequency spectrum (SFS) of the minor allele frequencies and inferred the relative strength of historic selection (Joly-Lopez et al., 2020). We found that *trans*-eQTLs for both environments had been under strong purifying selection, with the SFS being significantly left-shifted compared to background SNPs (two-tailed t-test $P < 10^{-16}$; mean MAF background = 0.20, *trans* normal = 0.0936, *trans* saline = 0.0910). Not surprisingly, given their highly pleiotropic nature, this effect was more pronounced for *trans*-eQTLs regulating multiple genes (two-tailed t-test $P < 10^{-16}$; normal mean MAF in unique *trans*-eQTLs = 0.099, multiple *trans*-eQTL = 0.073; saline mean MAF in unique *trans*-eQTLs = 0.097, multiple *trans*-eQTL = 0.071). In contrast to *trans*-eQTLs, we found that *cis*-eQTLs in both normal and saline conditions had potentially been under balancing selection (Fig. 5c, Supplementary Fig. S8) with the SFS being significantly right-shifted for *cis*-eQTLs relative to background SNPs (two-tailed t-test $P < 10^{-16}$; mean MAF background = 0.20, *cis* normal = 0.238, *cis* saline = 0.237). Moreover, for both conditions, the estimated nucleotide diversity (π) for 100-kb regions flanking *cis*-eQTLs was also significantly higher as compared to genome-wide 100-kb blocks (two-tailed t-test $P < 10^{-11}$; mean π background = 0.286, *cis* normal = 0.297, *cis* saline = 0.297).

Comparing patterns of past selection for *cis*-eQTLs specific to each condition revealed no significant difference between the normal and saline environments. However, we did observe *trans*-eQTLs in saline conditions to be under stronger purifying selection (two-tailed t-test $P = 0.017$). Taken together, this indicates that *cis*- and *trans*-eQTLs are under different selection regimes in rice, and that purifying selection is stronger on *trans*-eQTLs regulating gene expression under salt stress conditions, giving us insights into the macroevolutionary dynamics of gene-expression variation.

Discussion

Gene expression is a key link in the chain of organismal responses to environmental challenges. In this study, we used a systems genomics approach to examine genome-wide transcript levels in rice in both normal and moderately saline field conditions to examine the evolutionary response and genetic architecture of gene expression under salinity stress. We found selection on expression of a set of genes, and for some genes, selection on expression differed among environments, indicating that salinity can select for changes in gene expression. We saw that the genetic regulatory pathways can be modified by salinity, as indicated by decoherence, which provides a potential genetic mechanism underlying salinity tolerance. We found limited evidence for trade-offs given a lack of antagonistic pleiotropy in the fitness effects of expression compared across environments, and we also found that there did not appear to be strong *cis*-*trans* compensation in gene regulation. These results provide us with insights into the fitness consequences and genetic architecture of gene expression under salt stress in rice, as we discuss below.

One of the primary goals of this study was to characterize selection on gene expression and how this selection varies between saline and normal conditions. Before discussing our results, it is important to note that our conclusions are influenced by the way in which selection differentials are standardized. Selection differentials are estimated through regression coefficients in the relationship between trait values and fitness (Lande 1979). Because traits are often measured on different scales, selection differentials are usually standardized to allow for meaningful comparisons and interpretation. The most common method of standardization is variance standardization, in which traits are standardized to a mean of zero and a standard deviation of one (Lande and Arnold, 1983). With this approach, the selection differential, multiplied by the heritability, gives the expected response to selection in standard deviation units. So while this is a valid approach, there is concern that when variance-standardized selection differentials are interpreted as reflecting the “strength” of selection based on their magnitude, they can be

misleading because trait variance can vary among traits and environments, such that the trait variance and the magnitude of selection are conflated (Hereford et al., 2004). Instead, the use of mean-standardized selection estimates have been suggested (Hereford et al., 2004; Matsumura et al., 2012). However, it has not been widely recognized that mean-standardized selection estimates can also be biased if mean trait values vary consistently across conditions, such that magnitude of selection is conflated with trait means. We show here that this was indeed the case: the mean gene expression values differed consistently between saline and normal conditions, leading to biases when comparing the strength of selection across treatments when mean-standardized selection differentials were used. Specifically, using mean-standardized selection coefficients, we found selection on gene expression to be stronger under normal conditions, whereas unstandardized and variance-standardized selection coefficients indicated stronger selection under saline conditions. This result indicates that it is important for researchers to carefully consider the advantages and potential drawbacks of both variance-standardized and mean-standardized, as well as unstandardized, selection differentials in making decisions about which to use in any particular system and depending of the goals of the study, but also to consider these ideas in interpreting selection differentials as reflecting the strength of selection. Because we found that both variance-standardized and mean-standardized selection gradients were biased, as explained above, we draw our conclusions about patterns of selection using unstandardized selection gradients. We believe this approach to be appropriate given that gene expression values are measured on the same scale and normalized for every transcript, and our goal is to identify transcripts under selection and compare selection across treatments. Using unstandardized selection coefficients, our study shows that most genes appear to have nearly neutral levels ($|S| < 0.1$) of selection on transcript levels in both normal and saline environments. This is consistent with previous studies conducted in various plant and non-plant species, which have shown that most traits (including transcript abundance) are under very weak selection, and only a few traits experience strong selection at microevolutionary timescales (Ahmad et al., 2021; Hoekstra et al., 2001; Kingsolver et al., 2001). There does seem to be a tendency, however, towards stronger positive over negative directional selection on gene expression, meaning that higher expression of a majority of genes is generally associated with higher fitness. Looking at quadratic selection, we found evidence for both stabilizing and disruptive selection on gene expression. Our results are consistent with the expectation of stabilizing selection being more common than disruptive selection in normal conditions (Kingsolver et al., 2001), and reinforce the idea that disruptive selection is more widespread under stress conditions (Ahmad et al., 2021), potentially reflecting the prevalence of frequency- and density-dependent competition for resources under stress (Kingsolver and Pfennig, 2007).

Comparing the distribution of selection coefficients, we further found that selection was stronger in moderate stress saline field conditions compared to normal wet paddy field conditions, indicating that salinity stress induced increased selective pressure on gene expression at microevolutionary timescales. Although exposure to salinity stress increased the strength of selection on gene expression, this increase was relatively low compared to what has been reported for drought stress in rice (Agrawal and Whitlock, 2010; Groen et al., 2020; Kondrashov and Houle, 1997). This could potentially be attributed to the levels of stress experienced by the plants – the drought stress treatment was more severe in terms of fitness after stress exposure (Agrawal and Whitlock, 2010; Groen et al., 2020; Kondrashov and Houle, 1997) in comparison to the moderate salinity stress in this study – but further work is needed to examine whether there is indeed a relationship between stress intensity and selection strength. While this work provides insight into the direction of selection and the predicted degree of evolutionary change under the specific conditions of this study, more work would need to be done over multiple generations and under additional conditions to make more realistic predictions of microevolutionary change in the field.

Our work contributes to an ongoing debate regarding whether the intensity of selection on gene expression increases under stress. We found in this study of rice and salinity, and in a prior study on drought (Groen et al. 2020 [DOI](#)), that directional selection was greater under stress compared to normal conditions. This finding is consistent with prior studies on other systems like fruit flies (*Drosophila melanogaster*) (Groen et al., 2020 [DOI](#); Jasnos et al., 2008 [DOI](#); Kondrashov and Houle, 1997 [DOI](#)) and yeast (Groen et al., 2020 [DOI](#); Jasnos et al., 2008 [DOI](#); Kondrashov and Houle, 1997 [DOI](#)), but contrasts with a recent study in *D. melanogaster* reporting no change in selection due to increasing levels of nutritional stress (Arbuthnott and Whitlock, 2018 [DOI](#)). Although we found selection to be stronger under saline conditions using total filled grain number as a measure of fecundity, a caveat to selection estimates is that they are sensitive to the proxy of fitness. Our results also support prior findings in rice (Čalić et al., 2022 [DOI](#); Groen et al., 2020 [DOI](#)) and in *Boechera stricta* (Anderson et al., 2013 [DOI](#)) that conditional neutrality may be much more common than antagonistic pleiotropy, although the lack of antagonistic pleiotropy could be due to lack of trade-offs in the years evaluated or to low power to detect small-effects trade-offs. Further, both conditional neutrality and antagonistic pleiotropy have been shown to underlie local adaptation (Anderson et al., 2013 [DOI](#); Wadgyamar et al., 2017 [DOI](#)), following which our findings indicate that it may be feasible to breed salinity tolerant rice varieties without a yield penalty under non-saline conditions, though further work is needed to verify this.

We examined the genetic architecture of gene expression variation as well, and found that *trans*-eQTLs rather than *cis*-eQTLs are primarily associated with rice gene expression under salinity stress, potentially via a few master-regulators. *Trans*-eQTLs may be important in environment-dependent gene expression changes, while *cis*-eQTLs may be more robust to environmental changes as has been observed in other species (Smith and Kruglyak, 2008 [DOI](#); Snoek et al., 2012 [DOI](#); Sterken et al., 2023 [DOI](#)). This can be attributed to *trans*-eQTLs' larger mutational target along with a larger effect of drift than of positive selection in fixing them. These results further corroborates with other studies that have found a more dominant effect of *trans*-eQTLs on within-species variation than between-species variation (Emerson et al., 2010 [DOI](#); Metzger et al., 2016 [DOI](#); Wittkopp et al., 2008 [DOI](#)).

We found that *cis*- and *trans*-eQTLs show different patterns of selection, with *cis*-eQTLs showing evidence for balancing selection and *trans*-eQTLs showing purifying selection. We see this pattern for rice, a largely selfing species, in contrast to outcrossing species where both *cis*- and *trans*-eQTLs have been found to be under purifying selection (Hernandez et al., 2019 [DOI](#); Josephs et al., 2015 [DOI](#)). Future studies will be needed to investigate whether the mating system has an influence on the pattern of natural selection acting on variants regulating gene expression. Finally, and contrary to expectations, we show that under saline conditions, *cis-trans* reinforcement is more prevalent than *cis-trans* compensation. This result may be driven by rice domestication and subsequent population diversification. Additionally, we find significantly elevated levels of LD among the *cis*- and *trans*-regulatory variants for compensating and reinforcing genes, indicating the role of genetic fixation as an underlying mechanism for *cis-trans* compensation/reinforcement.

Systems genomic approaches provide insights into large-scale patterns across genome-wide data over multiple scales. This approach can also identify possible genes or genetic pathways that may prove critical in various processes. Our analysis, for example, shows that many of the potential antagonistically pleiotropic genes are involved in photosynthesis and metabolic processes. Importantly, we found that the circadian rhythm pathway was under positive directional selection (selection for increased expression) in saline conditions and that the genes involved in circadian rhythm showed significant decoherence between the two environments. These findings make sense in light of the fact that circadian rhythm has been shown to regulate salt tolerance and flowering time in rice (Liang et al., 2021 [DOI](#); Wei et al., 2021 [DOI](#)), so responses to salt stress may alter the expression patterns of genes in the circadian rhythm network, leading to decoherence. In further support of this idea of a link between salt tolerance and circadian rhythm, our selection analyses on physiological traits shows that salt stress leads to selection for earlier flowering.

Additionally, our decoherence analyses identified transcripts in important pathways related to plant responses and potentially tolerance to excess salts, including the carbon metabolic pathways (glycolysis and tricarboxylic acid cycle), and accumulation of secondary metabolites and sugar moieties with osmoprotectant properties (Jacoby et al., 2011 [↗](#); Krishnamurthy and Bhagwat, 1990 [↗](#); Petrusa and Winicov, 1997 [↗](#); Tari et al., 2010 [↗](#); Yang and Guo, 2018 [↗](#)). But as is with such analyses, detection of population-level correlations are temporally biased and may fail to detect the transient environmentally responsive links resulting in false negatives (Cai and Des Marais, 2023 [↗](#)) and leading to the failure to detect temporally key processes.

Other genes that are beneficial only in the saline environment include a cyclophilin-encoding transcript (*OsCYP2*), which has been shown to confer salt tolerance in rice (Ruan et al., 2011 [↗](#)). Moreover, we also identified eQTL variants regulating expression of two photosynthesis related antagonistically pleiotropic transcripts (*PSAN* and *CRR7*) in the normal environment. Finally, our *trans*-eQTL analysis identified a hotspot on chromosome 7 that contains *OsFLP*, a R2R3 myb-like transcription factor that regulates stomatal development (Wu et al., 2019 [↗](#)), and appears to be involved in salt tolerance in rice (Qu et al., 2022 [↗](#)). Together, these loci are possible new targets for functional and translational studies of salinity stress in rice. Coupled with the insights gained by a systems genomic approach in inferring large scale patterns from genome-wide information, this integration of data across multiple scales across a rice population has allowed us to provide an integrated examination of the molecular and genetic landscape underlying adaptive plant salinity stress responses.

Materials and methods

Plant Material

Domesticated rice is primarily classified into two distinct genetic subgroups, *O. sativa* ssp. *indica* and *O. sativa* ssp. *japonica*. These subgroups are grown in sympatry and are often recognized as subspecies, given the reproductive barriers between them (Nadir et al., 2018 [↗](#)). Further analyses have identified a widely accepted classification consisting of genetically distinct varietal groups namely, *indica*, *aus/circum-aus*, *aromatic/circum-basmati*, *tropical japonica*, and *temperate japonica* (Garris et al., 2005 [↗](#)). For this study, a total of 130 *O. sativa* ssp. *indica* (including *indica* and *circum-aus* groups) and 65 *O. sativa* ssp. *japonica* (including *circum-basmati*, *tropical japonica*, and *temperate japonica*) accessions were selected, including traditional varieties/landraces and three additionally replicated salt-sensitive and -tolerant test varieties. We focused our analyses on *O. sativa* ssp. *indica* since it is the predominant global varietal group (Supplementary Table 13). Seeds were obtained from the International Rice Genebank Collection (IRGC) at the International Rice Research Institute (IRRI) in the Philippines, and from a 2016 bulk seed collection obtained from plants grown under normal (wet and non-saline) conditions at IRRI during the course of a previous study (Groen et al., 2020 [↗](#)).

Field Experiment

The field experiment was conducted in the dry season of 2017 at IRRI in Los Baños, Laguna, Philippines. Seeds from each accession were sown on December 16, 2016, and seedlings were then transplanted into the experimental fields at 17 days after sowing (DAS), on January 5, 2017. The field experiment was conducted across two locations: site L4 (14°09'34.6"N 121°15'42.4"E) was prepared as the non-salinized "normal" environment and site L5 (14°09'35.2"N 121°15'42.5"E) as the salinized environment, following Groen et al., (2020) [↗](#). Within each field environment, there were three blocks and three replicates of each genotype (accession), with each genotype planted once per block in a random location. Each plant was planted in a single-row with 0.2-m × 0.2-m spacing between them for a total of one focal plant and seven neighboring plants (included in the experiment) per plot. Each experimental plot included the accessions NSIC Rc 222 and NSIC Rc 182 that served as border rows (Supplemental Fig. S9 [↗](#); Supplemental Table S13). The application of

salt in site L5 started on January 19, 2017, when the plants were 31 days old. The salinity level was monitored by recording electrical conductivity (EC), using EC meters installed in each of the parcels at a depth of 30 cm. The EC levels were recorded twice per day until reaching an $EC=6 \text{ dSm}^{-1}$ and then recorded daily. The salinity levels were then maintained at 6 dSm^{-1} (considered mild to moderate salinity stress) until maturity. Management and maintenance of the fields included the application of basal fertilizer, spraying of insecticides against thrips and removal of plants potentially infected with rice tungro virus.

Tissue collection for transcriptome sequencing

Leaf sampling was performed as previously described (Groen et al., 2020 [DOI](#)). Briefly, leaf collection in the non-saline and saline field was done from 10:00h to 12:00h at 38 DAS (8 days after the beginning of the salt treatment) in the non-saline and saline field from 10:00h to 12:00h. Both fields were sampled simultaneously and with individuals within a block collected in the same order. For each sample, about 10 cm of leaf length were cut into small pieces and placed in chilled 5-mL tubes containing 4 mL of RNeasy Lysis Buffer (Qiagen) solution for RNA stabilization and storage. Leaf samples from each of the 5-ml tubes were then transferred into pairs of 2-mL tubes (one for processing and one for backup), then stored at -80°C .

Yield harvesting and panicle trait phenotyping

A total of 780 plants were harvested individually and labeled such that the yield of all plants used for each type of measurement (mRNA sequencing, phenotypic measurements) was known. Individual seeds collected were further categorized as filled, partially filled, and unfilled, using manual assessment and a seed counter (Hoffman Manufacturing). Panicle length measurements and panicle trait phenotyping (PTRAP) of 30 seeds were also performed.

Functional trait phenotyping

In addition to yield-related measurements, we collected data on a number of physiological, morphological, and phenological traits to assess differences between rice accessions in response to soil salinity. In both the non-saline and saline fields, we recorded leaf osmotic potential (LOP) in the vegetative stage, performed chlorophyll analysis based on 1-mg leaf samples and measured ion content (analysis of sodium, Na^+ , and potassium, K^+ , analysis) based on 20-mg of leaf samples in both the vegetative and reproductive stages. We also measured plant height for growth rate (measured once a week until maturity). Flowering time was recorded as the day on which 50% of plants in a plot flowered; these plants included the focal plant and its seven neighboring plants. Whole plants were both harvested at the vegetative stage and at maturity to measure wet and dry biomass.

Extraction of total RNA for library construction

Leaf samples stored at -80°C were thawed at room temperature briefly and excess RNeasy Lysis Buffer was removed. Tissue samples were then flash-frozen in liquid nitrogen and ground using a TissueLyser II (Qiagen). After this, total RNA was extracted using the RNeasy Plant Mini Kit according to manufacturer's protocol (Qiagen) and eluted in nuclease-free water. The integrity of total RNA was assessed by agarose gel electrophoresis, and RNA from a random subset of samples was further assessed by Agilent TapeStation (Agilent Technologies). RNA concentration was quantified on a Qubit (Invitrogen). Samples were stored at -80°C until library preparation.

RNA-seq library preparation and sequencing

Library preparation for 780 samples was performed as described in Groen et al. (2020 [DOI](#)) and followed a plate-based 3'-end mRNA sequencing (3' mRNA-seq) protocol. Briefly, total RNA from each sample was transferred individually into 96-well plates and normalized to a concentration of 10 ng in 50 μL nuclease-free water. Then, mRNA samples were reverse-transcribed using

Superscript II Reverse Transcriptase (Thermo Fisher Scientific) and cDNAs were amplified using the Smart-seq2 protocol (Picelli et al., 2013 [↗](#)) with modifications (Groen et al., 2020 [↗](#)). This resulted in multiplexed pools of 96 samples, where 48 samples were from the non-saline field environment and 48 samples were from the same plot number in the saline field environment. Each pool was used for library preparation with the Nextera XT DNA sample prep kit (Illumina), returning 3'-biased cDNA fragments, similar to the Drop-seq protocol (Macosko et al., 2015 [↗](#)). The resulting cDNA libraries were then quantified on an Agilent BioAnalyzer and sequenced at the NYU Genomics Core on an Illumina NextSeq 500 with the configuration HighOutput 1 x 75 base pairs (bp) and the settings: Read 1 of 20 bp (bases 1–12, well barcode; bases 13–20, unique molecular identifier (UMI)) and Read 2 of 50 bp. Raw sequence reads have been submitted to the SRA (BioProject PRJNA1010833).

RNA-seq data processing and data normalization

3' mRNA-seq read data were processed as previously described in Groen et al. (2020) [↗](#). Briefly, Drop-seq tools v1.12 (<https://github.com/broadinstitute/Drop-seq> [↗](#)) and Picard tools v2.9.0 (<https://broadinstitute.github.io/picard/> [↗](#)) were used to generate the metadata. The reference genome, Nipponbare IRGSP 1.0 (ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/001/433/935/GCF_001433935.1_IRGSP-1.0 [↗](#)), and annotations were indexed with STAR v020201 (Dobin et al., 2013 [↗](#)). Prior to generating read counts, raw reads were converted from FASTQ to unaligned BAM format using Picard tools FastqToSam before being processed using the unified script for a FASTQ starting format. After this, digital gene-expression matrices displaying either UMI or raw read counts with transcripts as rows and samples as columns containing counts from reads with 96 expected sample barcodes were produced using the DigitalExpression utility. Sample barcodes corresponding to beads never exposed to rice total RNA were filtered out based on low numbers of transcribed elements as described previously (Groen et al., 2020 [↗](#); Macosko et al., 2015 [↗](#)). Rice individuals that ended up being discarded due to low numbers of transcribed elements, were sequenced again using another library.

UMI counts per sample were normalized through dividing by the total number of detected UMIs in that sample and multiplying by 1×10^6 to obtain transcripts per million. The resulting data matrices were then merged into one digital gene-expression super-matrix, containing transcripts-per-million expression data for all samples. Elements with very low transcription levels (transcript models with a sigma signal < 20) were discarded, after which a robust normalization was conducted using an invariant set normalization protocol within the DChip utility v2010.01 (Li and Wong, 2001 [↗](#)). All downstream analyses were done in log-space, using normalized expression levels ($\log_2[\text{normalized transcripts-per-million value} + 1]$) of transcribed elements estimated using R v3.4.3 (Computing, 2013; Robinson et al., 2010 [↗](#)). In a final step of filtering, transcripts that were not detected in at least 10% of individuals across our populations and did not derive from protein-coding genes on nuclear chromosomes were removed prior to performing subsequent analyses.

Quantitative genetics of fecundity and gene expression

All downstream analyses were done using R v3.4.3 (Computing, 2013; Robinson et al., 2010 [↗](#)). The effect of genotype (G), environment (E) and genotype-by-environment (G×E) on fecundity (number of filled grains) was assessed using two-way analysis of variance (ANOVA) with E as a fixed effect, and G and G×E as random effects (<https://www.angelfire.com/vw/bwhomedir/notes/anova2.pdf> [↗](#)). Essentially, the between sum of squares for G, E, and G×E was estimated using a one-way ANOVA with the aov function, which was then used to estimate the F-statistic. The significance of each term was determined using the F-tests with pf function. Fecundity, averaged by genotype, was further compared between the environments using a two-tailed paired t-test. Variation in gene expression was partitioned similarly using the same model as above, and significance of each term was tested using F-tests via a mixed-model ANOVA (Howell, 1997 [↗](#)). Multiple testing was controlled using a False-Discovery Rate (FDR) of 0.001. Broad-sense heritability was estimated as $H^2 = \sigma_G^2 / (\sigma_G^2 + \sigma_{GE}^2 + \sigma_E^2)$, with σ_G^2 , σ_{GE}^2 and σ_E^2 as the variance explained due to G, E, and

$G \times E$; e and r represent the number of environments and number of replicates per environment, respectively. Inter-varietal differences in gene expression was estimated for each environmental condition as the population-wide variance between accession mean expression levels.

Univariate and multivariate selection analyses

Univariate selection differentials consider each trait separately and represent the total strength of selection acting on a trait (Lande, 1979; Lande and Arnold, 1983). Fitness was estimated as fecundity, which was the total number of filled rice grains per individual. Unstandardized linear selection differentials (S) on gene expression were estimated as coefficients of linear regression with fecundity as the dependent variable in each regression and transcript abundance and block as the independent variables. Unstandardized quadratic selection differentials (C) were estimated similarly as twice the coefficients of quadratic regression for transcript abundance and fecundity (Conner and Hartl, 2004; Stinchcombe et al., 2008). Regressions were performed using the lmer function (Bates et al., 2015). Using these we then estimated the variance-standardized, and mean-standardized selection differentials (Hereford et al., 2004; Lande and Arnold, 1983). Data preparation included filtering out individuals with zero fecundity followed by normalizing fecundity fitness by mean fitness. To satisfy normality and remove noise inherent in expression data, transcripts with expression values more than 3 standard deviations from the mean were removed, which affected fewer than 1% of individuals. Selection differentials were estimated for transcripts that were expressed in at least 20 individuals and were estimated separately in each of the two environments. Multiple testing was controlled using Bonferroni correction (Bland and Altman, 1995) using the p.adjust function.

Multivariate selection gradients represent the strength of direct selection acting on each trait, after removing indirect selection caused by correlations with other traits (Lande and Arnold, 1983). Principal component analysis (PCA) was performed on transcript abundance using the prcomp function in R (Computing, 2013; Kassambara, 2017) and PCs explaining over 0.5% of variance in each environment were chosen for multivariate selection analyses. Linear (β) and quadratic (γ) selection gradients were estimated as coefficients of multiple regression with the normalized fecundity fitness as the dependent variable and the PCs as the independent variables (Conner and Hartl, 2004). We then chose the top one percent of transcripts showing the highest loadings of the PCs ($n=182$) and counted the number of transcripts showing evidence for positive directional selection (same directionality between loading of the transcript on the PC and univariate selection differential (S) acting on the transcript estimated above) versus negative directional selection (opposite directionality between loading of the transcript on the PC and univariate selection differential (S) acting on the transcript estimated above) and based on the majority assigned a directionality to the selection gradients on PCs. Variance-standardized multivariate selection gradients were estimated for functional traits without strong correlation to avoid collinearity among traits (Lande and Arnold, 1983; Presotto et al., 2019), using a Pearson correlation coefficient < 0.6 as threshold, which was estimated using the cor function in R (Computing, 2013).

Selection analyses on gene ontology biological processes

Gene Ontology (GO) term annotations for rice genes/transcripts were downloaded from Monocots PLAZA 5.0 (Van Bel et al., 2022). All fourth-level biological-process terms were downloaded using GO.db v3.15.0 (Carlson, 2022) and only these terms were considered for further analyses. Next, terms with fewer than 20 transcripts in our dataset were filtered out to minimize redundancy, leaving a total of 670 terms and 10,235 associated transcripts. The selection strength on a biological-process term was estimated as the median selection strength of all transcripts annotated with that term. A term was considered to be under significantly stronger selection compared to the transcriptome-wide median if the median strength of selection for a term was over the transcriptome-wide median selection strength by at least the 95% confidence interval for the selection strength of that term. GO enrichments were done using ShinyGO (Ge et al., 2020).

Regulatory decoherence analyses

To examine regulatory decoherence in rice, a recently developed method, CILP (Correlation by Individual Level Product), was used (Lea et al., 2019 [\[link\]](#)). CILP first estimates the correlation of a phenotype (transcript expression in our case) within each individual in the sample and then using linear model or linear mixed effect model tests for associations between this estimate and a fixed effect predictor variable (environment in our case), while controlling for covariates. Since CILP calculates product correlations for all possible pairs of genes, only transcripts with a selection strength greater than 0.1 ($|S| > 0.1$) in at least one environment with expression greater than 0 in at least 50% individuals were included to reduce the dimensionality (leaving 2,318 transcripts). Multiple testing was controlled using a False-Discovery Rate (FDR) of 0.05.

Genotype data and SNP calling

Raw FASTQ files were downloaded from the Sequence Read Archive (SRA) website under BioProject PRJEB6180 for 14 accessions (Wang et al., 2018 [\[link\]](#)) and under Bioprojects PRJNA422249 and PRJNA557122 for 92 accessions (Gutaker et al., 2020 [\[link\]](#)) (Supplementary Table 13). Further, genomes of 19 accessions were re-sequenced, and submitted to the SRA (BioProject PRJNA1012700), leading to a total of 125 accessions for which genomic data was available (Supplementary Table 13).

Raw reads were processed for quality control and adapter trimming using the bbdutk program of BBTools version 37.66 (<https://jgi.doe.gov/data-and-tools/bbtools/> [\[link\]](#)) using the options: minlen = 25 qtrim = rl trimq = 10 ktrim = r k = 25 mink = 11 hdist = 1 tpe tbo. The output from this program was mapped to the reference genome *O. sativa* Nipponbare IRGSP 1.0 genome that was downloaded from NCBI Genome ([https://www.ncbi.nlm.nih.gov/genome/?term=txid4530\[orgn\]](https://www.ncbi.nlm.nih.gov/genome/?term=txid4530[orgn]) [\[link\]](#)) using bwa-mem2 v2.1 (Vasimuddin et al., 2019 [\[link\]](#)). PCR duplicates were marked and removed using the Picard tools version 2.9.0. SNPs were called using GATK HaplotypeCaller v4.2.0.0 to obtain a multi-accession joint SNP file. Only SNPs that were above 5bp distance from an indel variant were taken. Next, SNPs were filtered using the recommended GATK hard filtering (Van der Auwera et al., 2013 [\[link\]](#)). Further, using vcftools v0.1.16 (Danecek et al., 2011 [\[link\]](#)), SNPs with at least 80% genotype calls and a minor allele frequency of 0.05 were retained ($\text{--max-missing } 0.8 \text{ --maf } 0.05$). Since rice is a inbred species, we also removed any SNPs that displayed heterozygosity of over 5% identified using vcftools v0.1.16 --hardy (Danecek et al., 2011 [\[link\]](#)). Next, missing genotype calls were imputed and phased using Beagle v4.1 (Browning and Browning, 2016 [\[link\]](#)), and using vcftools v0.1.16 --m2 -M2 (Danecek et al., 2011 [\[link\]](#)) only biallelic SNPs were retained for further analyses. Finally, SNPs were randomly pruned such that one SNP per 1000bp was retained using vcftools v0.1.16 --thin (Danecek et al., 2011 [\[link\]](#)), leaving a SNP dataset of 246,714 markers.

G-matrix estimation and prediction of short-term phenotypic evolution

A G-matrix (G) representing the additive genetic variance and covariance was estimated for the principal component axes (PCs) by taking the eigengene. This was done by deploying GREML v1.94 (Yang et al., 2011 [\[link\]](#)). Although the principal components are by definition uncorrelated at the level of the individual replicate plants, they start showing genetic covariances when loading values of replicates from each genotype are averaged. Next, using the multivariate breeder's equation ($\Delta z = G \beta$), we predicted the response to short-term phenotypic selection (Δz) on the PCs.

Association mapping

Association mapping was performed between the SNP markers and gene expression values recorded in the normal and saline environments. For this, the linear model in Matrix eQTL was used (Shabalin, 2012 [\[link\]](#)). The normalized gene expression values were averaged over the replicates

in each environment separately and these averages were subsequently used to test for associations. The first five principal components (PCs) of the kinship matrix were estimated using GAPIT v3 (Wang and Zhang, 2021 [DOI](#)) and added as covariates to control for population structure. Associations were considered significant at a false-discovery rate (FDR) < 0.001, and when significant were included in downstream analyses. Due to long stretches of homozygosity attributed to the highly inbred nature of rice, *trans*-eQTLs were defined as being on a different chromosome or at least 1 Mb away from a gene under its influence on the same chromosome; *cis*-eQTLs were defined as <100 kb away from an associated gene.

To identify significant $G \times e$ QTLs, we ran Matrix eQTL on the difference of expression in the normal and saline conditions (normal – saline) at FDR 0.05 (Smith and Kruglyak, 2008 [DOI](#)). For each gene regulated by a SNP in *cis* or *trans*, a lead SNP was identified as the SNP with the most significant association within a 100-kb region. Furthermore, *trans*-eQTL hotspots were identified through analyzing the number of unique genes regulated by lead SNPs in a given 100-kb region. To detect genes under *cis-trans* compensation or reinforcement, the effect sizes of all lead *trans*-eQTLs were averaged for each gene and compared to the lead *cis*-eQTL for the same gene. Next, for these genes we estimated the mean proportion of individuals with opposite and same direction *cis-trans* allelic configuration. Genes were defined as compensating and reinforcing if they had at least 60% of individuals with opposite and same *cis-trans* allelic configuration, respectively. To examine the LD structure, we estimated r^2 using plink v1.9 (Purcell et al., 2007 [DOI](#)) between (1) all pairs of *cis*- and *trans*-variants for the identified compensating and reinforcing genes and (2) 1,000 datasets of randomly selected SNP pairs (with equal numbers of variant pairs and a similar distribution of distances between the *cis*- and *trans*- variants as in 1). Next, we compared r^2 between (1) and (2) using a two-tailed permutation test. To examine the patterns of past selection on eQTLs, we used the minor allele frequency (MAF) of the 246,714 SNP markers and compared these using the t.test function in R (Computing, 2013). Furthermore, we estimated the site-wise nucleotide diversity (π) and averaged it over 50-kb flanking regions around each *cis*-eQTLs (100-kb region total). We compared this π to the background nucleotide diversity, estimated as π averaged over 100-kb blocks throughout the genome minus the 100-kb *cis*-eQTL region above. The difference in mean of nucleotide diversity was tested using the t.test function in R (Computing, 2013).

Acknowledgements

We thank the New York University Center for Genomics and Systems Biology GenCore Facility for sequencing support, and the New York University High Performance Computing for supplying computational resources. We would also like to thank William Mauck and Nicholas Rogers for technical support, as well as Adrian E. Platts and Jae Young Choi, for their help with post-sequencing data processing. We are grateful to current members of the Purugganan laboratory (particularly J. Flowers, A. Kurbidaeva, O. Alam) and Elena Hamann at Fordham University (currently Heinrich Heine University Düsseldorf) for insightful discussions. S.C.G. is supported by a grant from the Gordon and Betty Moore Foundation/Life Sciences Research Foundation (award no. GBMF2550.06 to S.C.G.), startup funds from the University of California Riverside, and a grant from the National Institute of General Medical Sciences of the National Institutes of Health (award no. R35GM151194 to S.C.G.). This work was funded in part by grants from the US National Science Foundation Plant Genome Research Program (IOS 1546218 and 2204374) and the Zegar Family Foundation.

Figure S1

Pathway enrichment of the 51 antagonistically pleiotropic genes beneficial in normal conditions but detrimental in salinity stress conditions. Enrichment statistics provided in Supplementary Table 4.

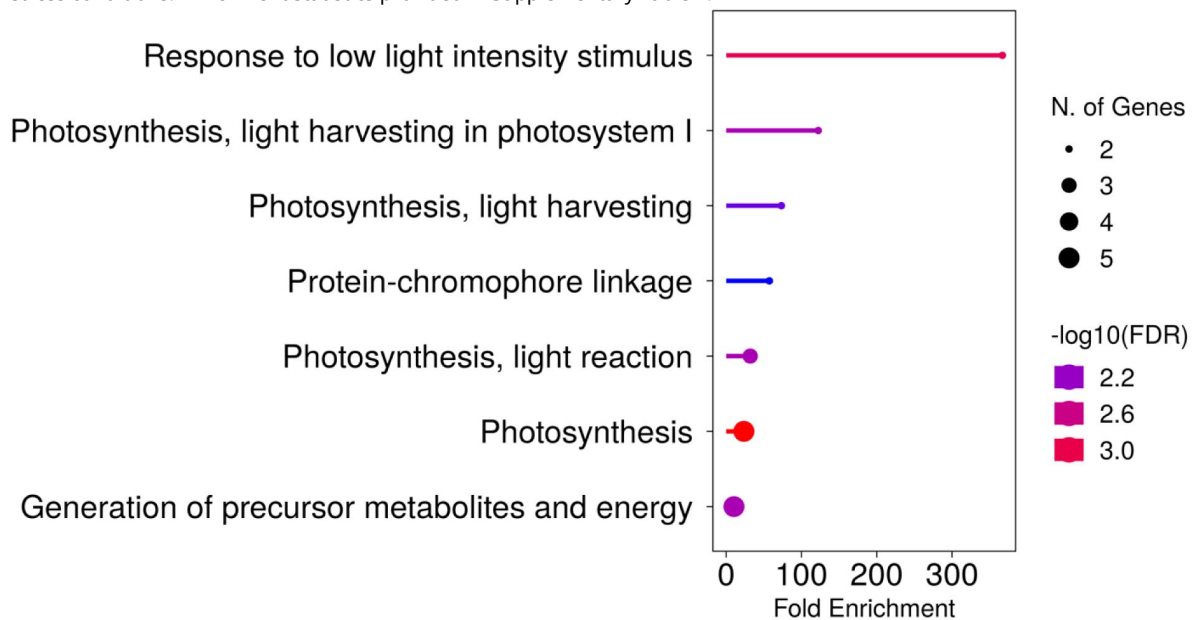
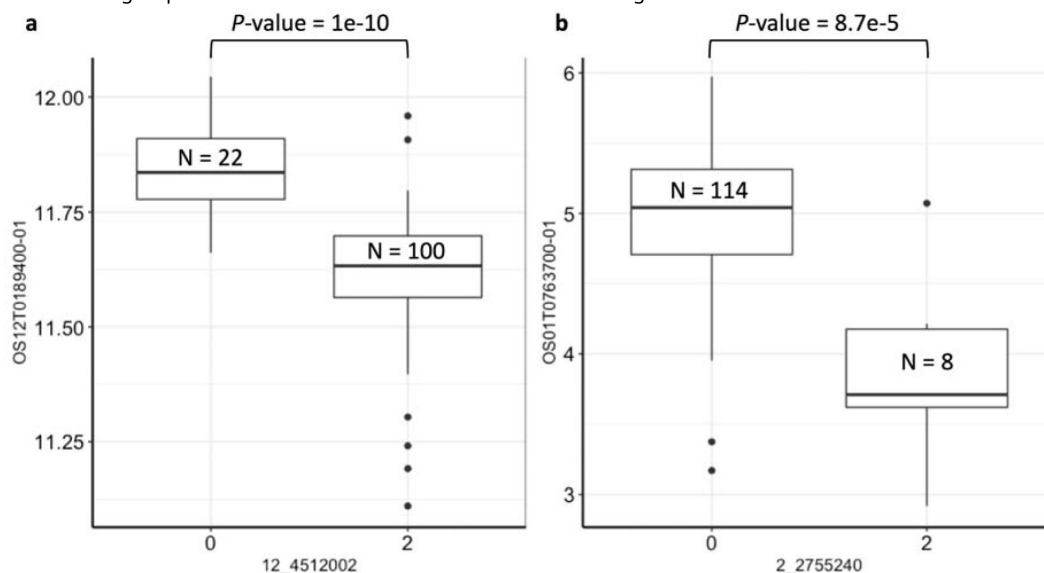


Figure S2

Boxplot representation of eQTLs for the two photosynthesis-related AP (antagonistically pleiotropic) genes beneficial in normal conditions. a, *cis*-eQTL (SNP = Chr12:4512002) for *PSAN* (OS12T0189400-01). b, *trans*-eQTL (SNP = Chr02:2755240) for *CRR7* (OS01T0763700-01). eQTL statistics provided in Supplementary Table 10. X-axis represents the eQTL SNP haplotypes, y-axis represents normalized transcript expression in normal conditions. Numbers inside box plots represent the number of accessions in each group. *P*-value indicates the one-sided Wilcoxon-test significance.



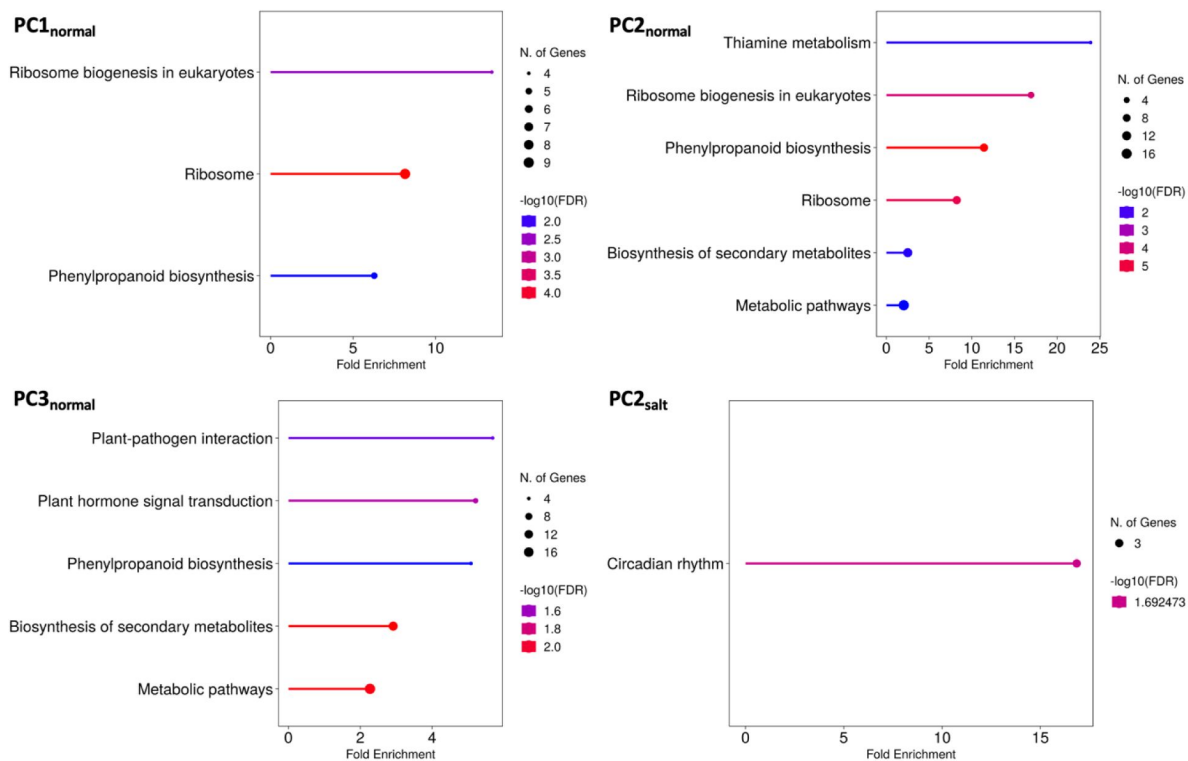


Figure S3

Enrichment of the suite of transcripts (1% tails of the distributions of transcripts' loading values on principal components) with significant selection gradients in both normal and salinity stress conditions.

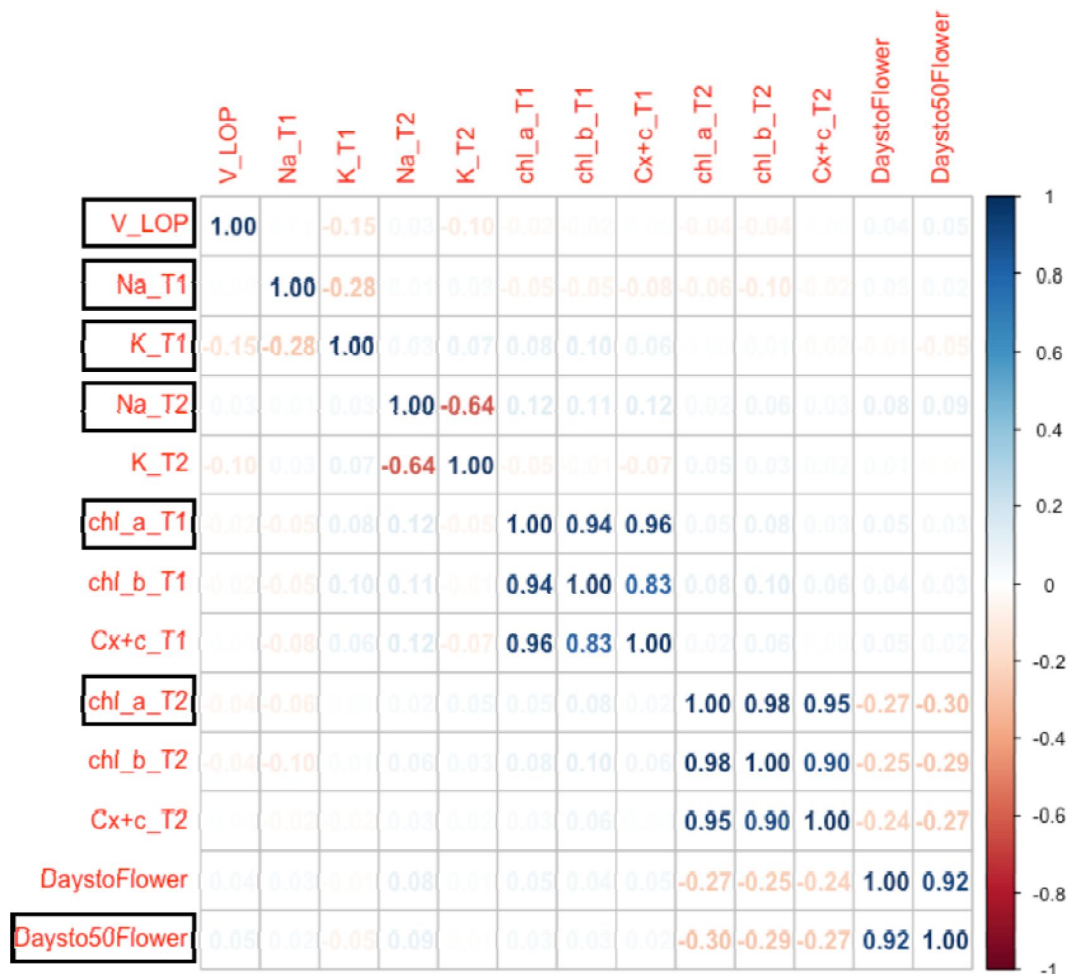


Figure S4

Correlation among functional traits in normal conditions. Highlighted black boxes indicate the uncorrelated traits chosen for selection analyses. Numbers inside the boxes and the heatmap represent Pearson correlation coefficients. V_LOP: Leaf Osmotic Potential; Na: Sodium content; K: Potassium content; chl_a and chl_b: Chlorophyll a and b content, respectively; Cx+c: Total carotenoid content; DaystoFlower: First day of flowering; Daysto50Flower: Day on which 50% of plants in a plot flowered (focal plant and its nine neighboring plants); T1 and T2 represent vegetative and reproductive timepoints, respectively.

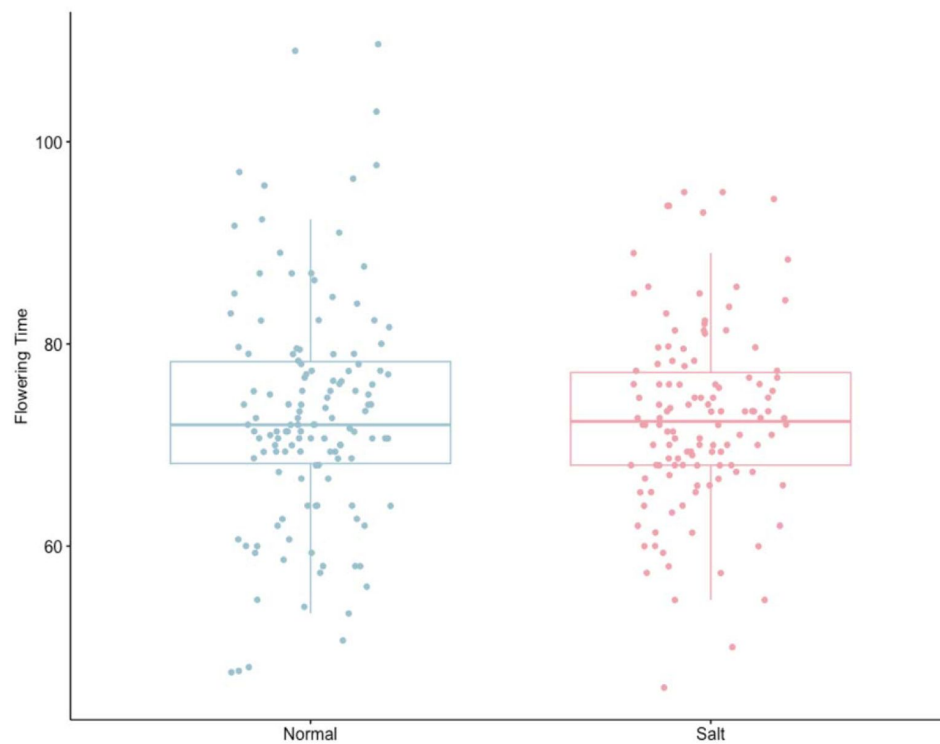


Figure S5

Distribution of flowering time (representing days to when 50% of plants in a plot flowered) in normal and salinity stress conditions. Paired t-test (two-sided paired t-test $P = 0.001$) showed significant reduction in flowering times after salt treatment, indicating earlier flowering saline conditions.

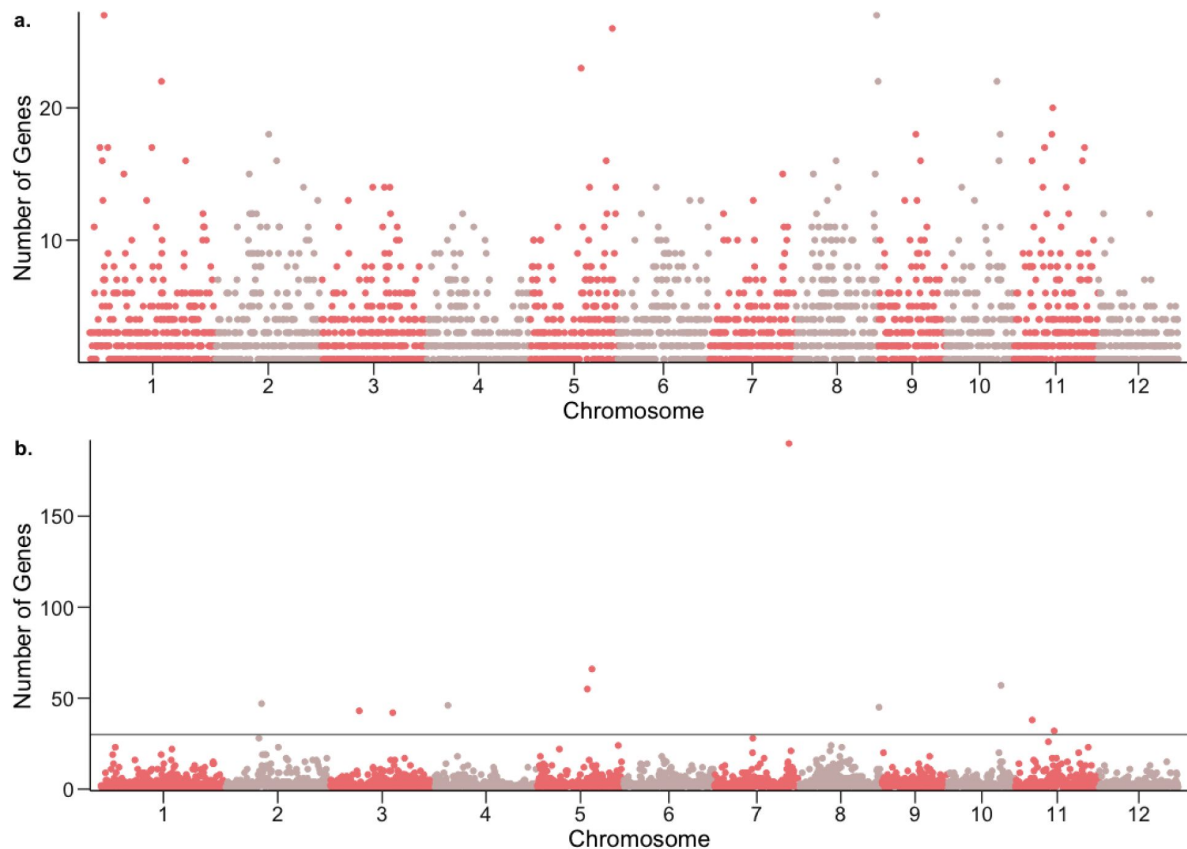


Figure S6

Trans-eQTL hotspots in normal (a.) and salinity stress (b.) conditions. X-axes indicate genomic locations of 100-kb nonoverlapping windows for which the total numbers of unique genes regulated were calculated (Y-axis). Hotspots were defined as windows regulating expression of over 30 genes; there were 0 hotspots in normal conditions (max number of genes = 28), and 11 hotspots in saline conditions. Gray horizontal line represents the cutoff of 30 genes.

Figure S7

Compensating and reinforcing *cis-trans* effects in normal conditions. a, Effect sizes of genes with both *cis* and *trans* factors showing excess of *cis-trans* reinforcement (teal) in comparison to *cis-trans* compensation (salmon). b, Inter-variety differences in gene expression for genes under compensating control is significantly lower than that for genes under reinforcing control; one-sided Mann-Whitney $P = 0.026$.

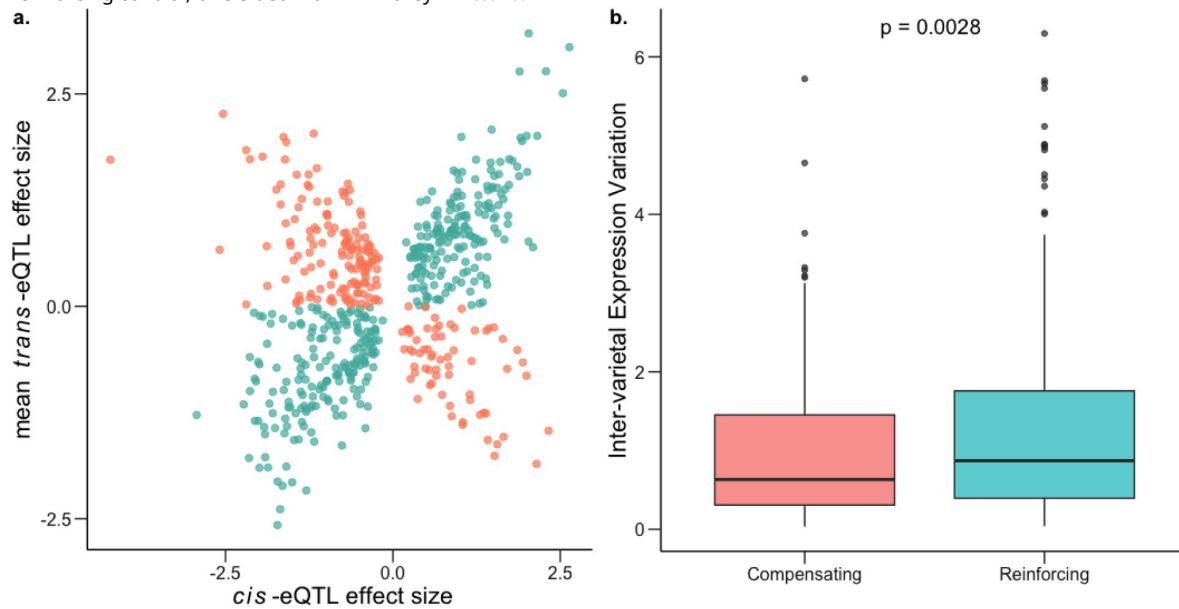
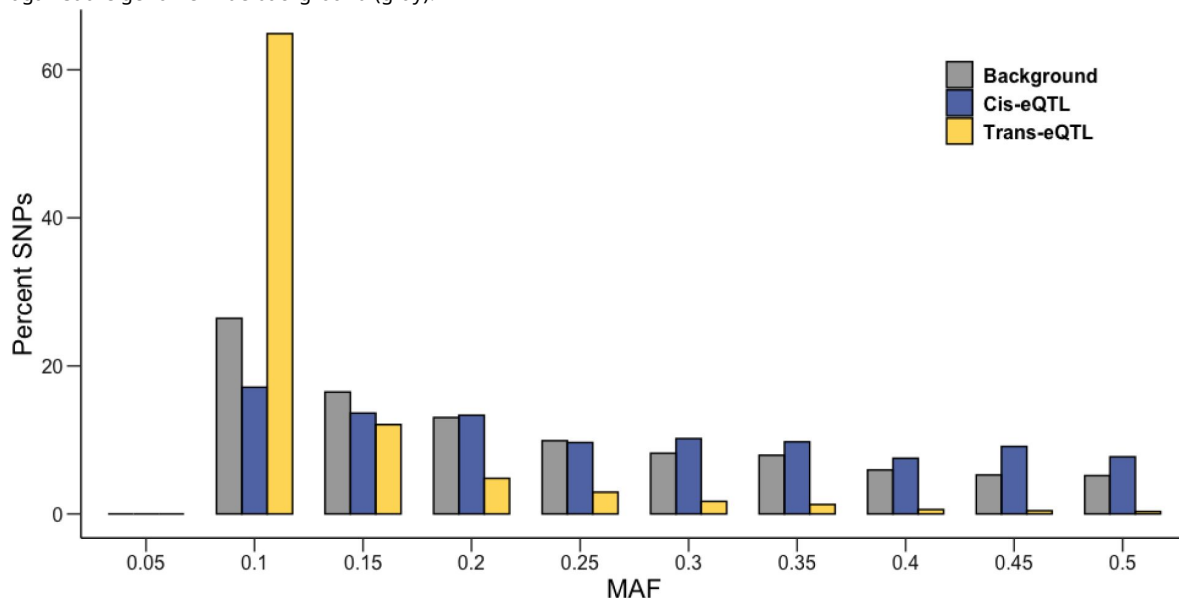


Figure S8

Frequency distribution of MAF (Minor Allele Frequency) for *cis*-eQTLs (blue) and *trans*-eQTLs (yellow) in normal conditions against the genome-wide background (gray).



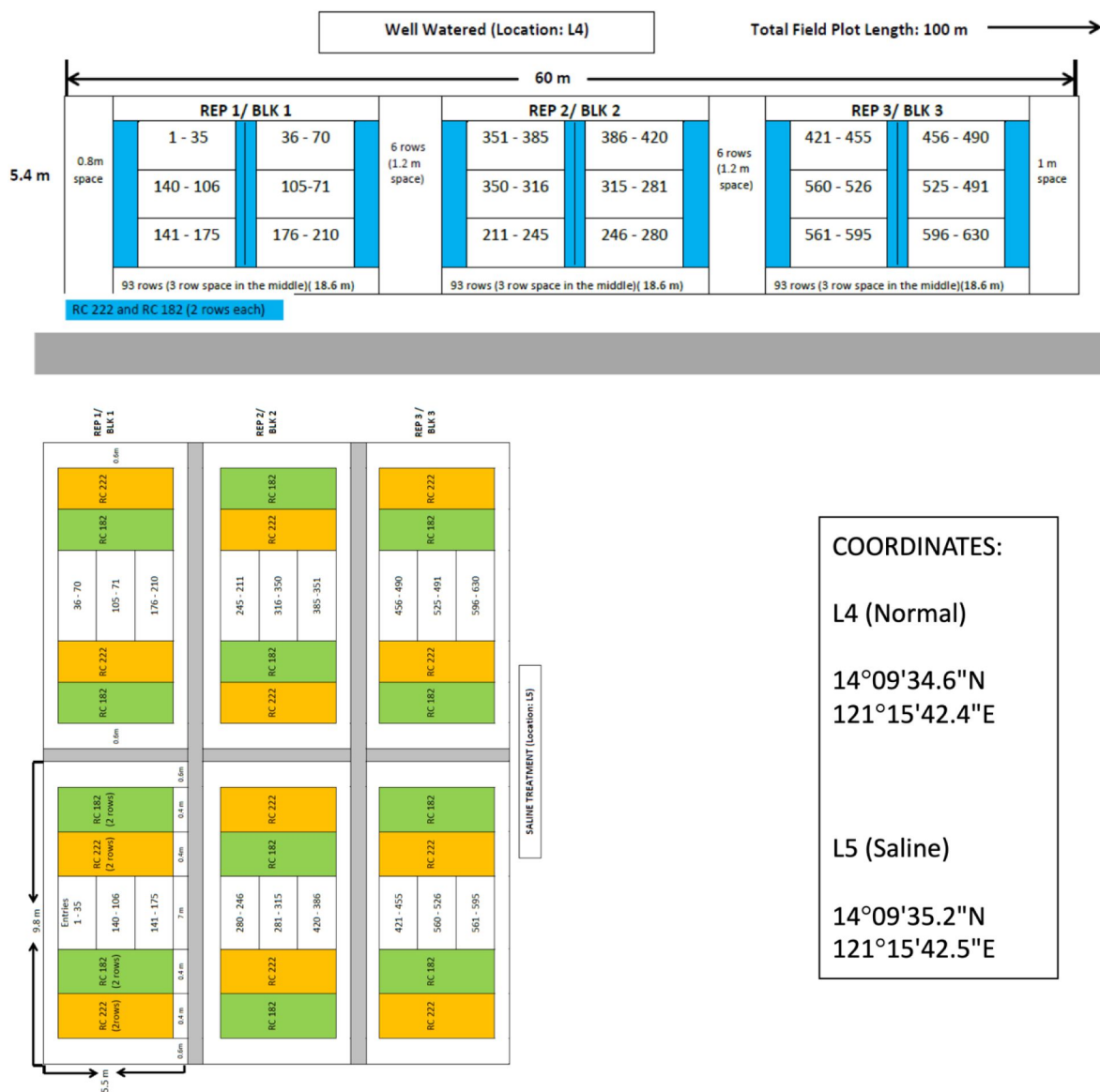


Figure S9

Experimental design for the field experiment. A total of 130 *O. sativa* ssp. *indica* and 65 *O. sativa* ssp. *japonica* accessions were selected for this experiment, including traditional varieties/landraces and three additional test varieties (Supplementary Table S12). Within each field environment (normal and saline), there were a total of three blocks each corresponding to a replicate, with each accession planted once per block, at a random location. Each experimental plot included the accessions NSIC Rc 222 and NSIC Rc 182 that served as border rows (Supplemental Table S13). REP: replicate; BLK: block.

Additional information

Author contributions

M.D.P. conceived and directed the project; Z.J.-L., R.K.S., G.V.V., A.G.C.S., K.L.M. designed and coordinated field experiments; Z.J.-L., S.C.G., M.A.N., A.G.C.S. performed field experiments; Z.J.-L., M.L.Z., S.C.G. processed samples and extracted RNA for sequencing; Z.J.-L. and R.S. made RNA-seq libraries; M.L.Z., S.C.G. made whole-genome re-sequencing libraries; I.C., S.C.G., Z.J.-L., S.J.F., K.L.M. designed and ran the bioinformatics workflow for RNA-seq; S.G. performed selection, statistical and genetic-marker-based analyses with assistance from S.J.F. and M.D.P.; S.C.G., Z.J.-L., S.G., M.A.N. processed fitness, higher-level-trait and gene expression data; and S.G., Z.J.-L. and M.D.P. wrote the manuscript. All authors contributed towards the editing and final review of the manuscript.

Additional files

Supplementary Tables [↗](#)

References

- Agrawal AF, Whitlock MC 2010) **Environmental duress and epistasis: how does stress affect the strength of selection on new mutations?** *Trends Ecol Evol* **25**:450–458
- Ahmad F, Debes PV, Nousiainen I, Kahar S, Pukk L, Gross R, Ozerov M, Vasemägi A 2021) **The strength and form of natural selection on transcript abundance in the wild** *Mol Ecol* **30**:2724–2737
- Amiri A, Coppola G, Scuderi S, Wu F, Roychowdhury T, Liu F, Pochareddy S, Shin Y, Safi A, Song L, Zhu Y, Sousa AMM, PsychENCODE Consortium, Gerstein M, Crawford GE, Sestan N, Abyzov A, Vaccarino FM 2018) **Transcriptome and epigenome landscape of human cortical development modeled in organoids** *Science* **362** <https://doi.org/10.1126/science.aat6720>
- Anderson JT, Lee C-R, Rushworth CA, Colautti RI, Mitchell-Olds T 2013) **Genetic trade-offs and conditional neutrality contribute to local adaptation** *Mol Ecol* **22**:699–708
- Anderson JT, Willis JH, Mitchell-Olds T 2011) **Evolutionary genetics of plant adaptation** *Trends Genet* **27**:258–266
- Arbuthnott D, Whitlock MC 2018) **Environmental stress does not increase the mean strength of selection** *J Evol Biol* **31**:1599–1606
- Ashraf MY, Bhatti AS 2000) **Effect of salinity on growth and chlorophyll content in rice** *Pakistan journal of scientific and industrial research* **43**:130–132
- Bates D, Mächler M, Bolker B, Walker S 2015) **Fitting linear mixed-effects models Using lme4** *J Stat Softw* **67**:1–48
- Bland JM, Altman DG 1995) **Multiple significance tests: the Bonferroni method** *BMJ* **310**
- Browning BL, Browning SR 2016) **Genotype Imputation with Millions of Reference Samples** *Am J Hum Genet* **98**:116–126
- Cai H, Des Marais DL. 2023) **Revisiting regulatory coherence: accounting for temporal bias in plant gene co-expression analyses** *New Phytol* **238**:16–24
- Ćalić I, Groen SC, Choi JY, Joly-Lopez Z, Hamann E, Natividad MA, Dorph K, Cabral CLU, Torres RO, Vergara GV, Henry A, Purugganan MD, Franks SJ. 2022) **The influence of genetic architecture on responses to selection under drought in rice** *Evol Appl* **15**:1670–1690
- Carillo P, Grazia M, Pontecorvo G, Fuggi A, Woodrow P, Shanker A, Venkateswarlu B 2011) **Salinity stress and salt tolerance** *Abiotic Stress in Plants - Mechanisms and Adaptations* London, England: InTech
- Carlson M 2022) **GO. db: A set of annotation maps describing the entire Gene Ontology** *Bioconductor*
- Castillo EG, Tuong TP, Ismail AM, Inubushi K 2007) **Response to Salinity in Rice: Comparative Effects of Osmotic and Ionic Stresses** *Plant Prod Sci* **10**:159–170

- Chang J, Cheong BE, Natera S, Roessner U 2019) **Morphological and metabolic responses to salt stress of rice (*Oryza sativa* L.) cultivars which differ in salinity tolerance** *Plant Physiol Biochem* **144**:427–435
- Che-Othman MH, Millar AH, Taylor NL 2017) **Connecting salt stress signalling pathways with salinity-induced changes in mitochondrial metabolic processes in C3 plants** *Plant Cell Environ* **40**:2875–2905
- 2013) **Computing RFS. 2013. R: A language and environment for statistical computing. Vienna: R Core Team.** Vienna: R Core Team: R: A language and environment for statistical computing
- Conner JK, Hartl DL 2004) **A primer of ecological genetics** New York, NY: Oxford University Press
- Coolon JD, McManus CJ, Stevenson KR, Graveley BR, Wittkopp PJ 2014) **Tempo and mode of regulatory evolution in *Drosophila*** *Genome Res* **24**:797–808
- Cramer GR, Urano K, Delrot S, Pezzotti M, Shinozaki K 2011) **Effects of abiotic stress on plants: a systems biology perspective** *BMC Plant Biol* **11**
- Cui Y, Lu S, Li Z, Cheng J, Hu P, Zhu T, Wang X, Jin M, Wang X, Li L, Huang S, Zou B, Hua J 2020) **CYCLIC NUCLEOTIDE-GATED ION CHANNELS 14 and 16 Promote Tolerance to Heat and Chilling in Rice** *Plant Physiol* **183**:1794–1808
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, 1000 Genomes Project Analysis Group 2011) **The variant call format and VCFtools** *Bioinformatics* **27**:2156–2158
- De Clercq I, Van de Velde J, Luo X, Liu L, Storme V, Van Bel M, Pottier R, Vanechoutte D, Van Breusegem F, Vandepoele K. 2021) **Integrative inference of transcriptional networks in *Arabidopsis* yields novel ROS signalling regulators** *Nat Plants* **7**:500–513
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR 2013) **STAR: ultrafast universal RNA-seq aligner** *Bioinformatics* **29**:15–21
- Emerson JJ, Hsieh L-C, Sung H-M, Wang T-Y, Huang C-J, Lu HH-S, Lu M-YJ, Wu S-H, Li W-H 2010) **Natural selection on cis and trans regulation in yeasts** *Genome Res* **20**:826–836
- Fahad S, Bajwa AA, Nazir U, Anjum SA, Farooq A, Zohaib A, Sadia S, Nasim W, Adkins S, Saud S, Ihsan MZ, Alharby H, Wu C, Wang D, Huang J 2017) **Crop Production under Drought and Heat Stress: Plant Responses and Management Options** *Front Plant Sci* **8**
- Falconer DS, Mackay TFC 1996) **Introduction to quantitative genetics** Essex, England: Prentice Hall
- Francini A, Sebastiani L. 2019) **Abiotic Stress Effects on Performance of Horticultural Crops** *Horticulturae* **5**
- Garris AJ, Tai TH, Coburn J, Kresovich S, McCouch S 2005) **Genetic structure and diversity in *Oryza sativa* L** *Genetics* **169**:1631–1638
- Ge SX, Jung D, Yao R 2020) **ShinyGO: a graphical gene-set enrichment tool for animals and plants** *Bioinformatics* **36**:2628–2629

Goncalves A, Leigh-Brown S, Thybert D, Stefflova K, Turro E, Flicek P, Brazma A, Odom DT, Marioni JC (2012) **Extensive compensatory cis-trans regulation in the evolution of mouse gene expression** *Genome Res* **22**:2376–2384

Groen SC, Čalić I, Joly-Lopez Z, Platts AE, Choi JY, Natividad M, Dorph K, Mauck WM, Bracken B, Cabral CLU, Kumar A, Torres RO, Satija R, Vergara G, Henry A, Franks SJ, Purugganan MD. (2020) **The strength and pattern of natural selection on gene expression in rice** *Nature* **578**:572–576

Groen SC, Hamann E, Čalić I, Cochran C, Konshok R, Purugganan MD, Franks SJ. (2021) **Fitness costs and benefits of gene expression plasticity in rice under drought** *bioRxiv* <https://doi.org/10.1101/2021.03.16.435597>

Gutaker RM, Groen SC, Bellis ES, Choi JY, Pires IS, Bocinsky RK, Slayton ER, Wilkins O, Castillo CC, Negrão S, Oliveira MM, Fuller DQ, Guedes JAD, Lasky JR, Purugganan MD. (2020) **Genomic history and ecology of the geographic spread of rice** *Nat Plants* **6**:492–502

Hakim MA, Juraimi AS, Hanafi MM, Ismail MR, Selamat A, Rafii MY, Latif MA. (2014) **Biochemical and anatomical changes and yield reduction in rice (*Oryza sativa* L.) under varied salinity regimes** *Biomed Res Int* **2014**

Hammond JP, Mayes S, Bowen HC, Graham NS, Hayden RM, Love CG, Spracklen WP, Wang J, Welham SJ, White PJ, King GJ, Broadley MR (2011) **Regulatory hotspots are associated with plant gene expression under varying soil phosphorus supply in *Brassica rapa*** *Plant Physiol* **156**:1230–1241

Hereford J, Hansen TF, Houle D (2004) **Comparing strengths of directional selection: how strong is strong?** *Evolution* **58**:2133–2143

Hernandez RD, Uricchio LH, Hartman K, Ye C, Dahl A, Zaitlen N (2019) **Ultrarare variants drive substantial cis heritability of human gene expression** *Nat Genet* **51**:1349–1355

Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SN, Hoang A, Hill CE, Beerli P, Kingsolver JG (2001) **Strength and tempo of directional selection in the wild** *Proc Natl Acad Sci U S A* **98**:9157–9160

Howell DC (1997) **Statistical Methods for Psychology** *Cengage Learning*

Hussain S, Zhang J-H, Zhong C, Zhu L-F, Cao X-C, Yu S-M, Allen Bohr J, Hu J-J, Jin Q-Y (2017) **Effects of salt stress on rice growth, development characteristics, and the regulating ways: A review** *J Integr Agric* **16**:2357–2374

Israel JW, Martik ML, Byrne M, Raff EC, Raff RA, McClay DR, Wray GA (2016) **Comparative Developmental Transcriptomics Reveals Rewiring of a Highly Conserved Gene Regulatory Network during a Major Life History Switch in the Sea Urchin Genus *Heliocidaris*** *PLoS Biol* **14**

Jacoby RP, Taylor NL, Millar AH (2011) **The role of mitochondrial respiration in salinity tolerance** *Trends Plant Sci* **16**:614–623

Jasnos L, Tomala K, Paczesniak D, Korona R (2008) **Interactions between stressful environment and gene deletions alleviate the expected average loss of fitness in yeast** *Genetics* **178**:2105–2111

- Jiang Z, Zhou X, Tao M, Yuan F, Liu L, Wu F, Wu X, Xiang Y, Niu Y, Liu F, Li C, Ye R, Byeon B, Xue Y, Zhao H, Wang H-N, Crawford BM, Johnson DM, Hu C, Pei C, Zhou W, Swift GB, Zhang H, Vo-Dinh T, Hu Z, Siedow JN, Pei Z-M 2019) **Plant cell-surface GIPC sphingolipids sense salt to trigger Ca²⁺ influx** *Nature* **572**:341–346
- Joly-Lopez Z, Platts AE, Gulko B, Choi JY, Groen SC, Zhong X, Siepel A, Purugganan MD 2020) **An inferred fitness consequence map of the rice genome** *Nat Plants* **6**:119–130
- Josephs EB, Lee YW, Stinchcombe JR, Wright SI 2015) **Association mapping reveals the role of purifying selection in the maintenance of genomic variation in gene expression** *Proceedings of the National Academy of Sciences* **112**:15390–15395
- Kassambara A. 2017) **Practical Guide To Principal Component Methods in R: PCA, M(CA), FAMD, MFA, HCPC, factoextra** Sthda
- Kim W-Y, Ali Z, Park HJ, Park SJ, Cha J-Y, Perez-Hormaeche J, Quintero FJ, Shin G, Kim MR, Qiang Z, Ning L, Park HC, Lee SY, Bressan RA, Pardo JM, Bohnert HJ, Yun D-J 2013) **Release of SOS2 kinase from sequestration with GIGANTEA determines salt tolerance in Arabidopsis** *Nat Commun* **4**
- Kingsolver JG, Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SN, Hill CE, Hoang A, Gibert P, Beerli P 2001) **The strength of phenotypic selection in natural populations** *Am Nat* **157**:245–261
- Kingsolver JG, Pfennig DW 2007) **Patterns and Power of Phenotypic Selection in Nature** *Bioscience* **57**:561–572
- Kliebenstein D 2009) **Quantitative genomics: analyzing intraspecific variation using global gene expression polymorphisms or eQTLs** *Annu Rev Plant Biol* **60**:93–114
- Ko DK, Brandizzi F 2020) **Network-based approaches for understanding gene regulation and function in plants** *Plant J* **104**:302–317
- Kondrashov AS, Houle D 1997) **Genotype—environment interactions and the estimation of the genomic mutation rate in Drosophila melanogaster** *Proceedings of the Royal Society of London Series B: Biological Sciences* **258**:221–227
- Kopecká R, Kameniarová M, Černý M, Brzobohatý B, Novák J 2023) **Abiotic Stress in Crop Production** *Int J Mol Sci* **24** <https://doi.org/10.3390/ijms24076603>
- Krishnamurthy R, Bhagwat KA 1990) **Accumulation of choline and glycinebetaine in salt-stressed wheat seedlings** *Curr Sci* **59**:111–112
- Kumar SV, Wigge PA 2010) **H2A.Z-containing nucleosomes mediate the thermosensory response in Arabidopsis** *Cell* **140**:136–147
- Kuroha T, Nagai K, Kurokawa Y, Nagamura Y, Kusano M, Yasui H, Ashikari M, Fukushima A 2017) **eQTLs Regulating Transcript Variations Associated with Rapid Internode Elongation in Deepwater Rice** *Front Plant Sci* **8**
- Lande R 1979) **Quantitative Genetic Analysis of Multivariate Evolution, Applied to Brain: Body Size Allometry** *Evolution* **33**:402–416

Lande R, Arnold SJ 1983) **THE MEASUREMENT OF SELECTION ON CORRELATED CHARACTERS** *Evolution* **37**:1210–1226

Landry CR, Wittkopp PJ, Taubes CH, Ranz JM, Clark AG, Hartl DL 2005) **Compensatory cis-trans evolution and the dysregulation of gene expression in interspecific hybrids of *Drosophila*** *Genetics* **171**:1813–1822

Laohavisit A, Richards SL, Shabala L, Chen C, Colaço RDDR, Swarbreck SM, Shaw E, Dark A, Shabala S, Shang Z, Davies JM 2013) **Salinity-induced calcium signaling and root adaptation in *Arabidopsis* require the calcium regulatory protein annexin1** *Plant Physiol* **163**:253–262

Lea A, Subramaniam M, Ko A, Lehtimäki T, Raitoharju E, Kähönen M, Seppälä I, Mononen N, Raitakari OT, Ala-Korpela M, Pajukanta P, Zaitlen N, Ayroles JF 2019) **Genetic and environmental perturbations lead to regulatory decoherence** *Elife* **8** <https://doi.org/10.7554/eLife.40538>

Lee SS, Park HJ, Jung WY, Lee A, Yoon DH, You YN, Kim H-S, Kim B-G, Ahn JC, Cho HS 2015) **OscCYP21-4, a novel Golgi-resident cyclophilin, increases oxidative stress tolerance in rice** *Front Plant Sci* **6**

Liang L, Zhang Z, Cheng N, Liu H, Song S, Hu Y, Zhou X, Zhang J, Xing Y 2021) **The transcriptional repressor OsPRR73 links circadian clock and photoperiod pathway to control heading date in rice** *Plant Cell Environ* **44**:842–855

Liang W, Ma X, Wan P, Liu L 2018) **Plant salt-tolerance mechanism: A review** *Biochem Biophys Res Commun* **495**:286–291

Li C, Wong WH 2001) **Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection** *Proc Natl Acad Sci U S A* **98**:31–36

Li K, Wang Y, Han C, Zhang W, Jia H, Li X 2007) **GA signaling and CO/FT regulatory module mediate salt-induced late flowering in *Arabidopsis thaliana*** *Plant Growth Regul* **53**:195–206

Liu C, Mao B, Yuan D, Chu C, Duan M 2022) **Salt tolerance in rice: Physiological responses and molecular mechanisms** *The Crop Journal* **10**:13–25

Liu S, Li C, Wang H, Wang S, Yang S, Liu X, Yan J, Li B, Beatty M, Zastrow-Hayes G, Song S, Qin F 2020) **Mapping regulatory variants controlling gene expression in drought response and tolerance in maize** *Genome Biol* **21**

Lovell JT, Jenkins J, Lowry DB, Mamidi S, Sreedasyam A, Weng X, Barry K, Bonnette J, Campitelli B, Daum C, Gordon SP, Gould BA, Khasanova A, Lipzen A, MacQueen A, Palacio-Mejía JD, Plott C, Shakirov EV, Shu S, Yoshinaga Y, Zane M, Kudrna D, Talag JD, Rokhsar D, Grimwood J, Schmutz J, Juenger TE 2018) **The genomic landscape of molecular responses to natural drought stress in *Panicum hallii*** *Nat Commun* **9**

Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, Trombetta JJ, Weitz DA, Sanes JR, Shalek AK, Regev A, McCarroll SA 2015) **Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets** *Cell* **161**:1202–1214

Mareri L, Parrotta L, Cai G 2022) **Environmental Stress and Plants** *Int J Mol Sci* **23** <https://doi.org/10.3390/ijms23105416>

- Matsumura S, Arlinghaus R, Dieckmann U (2012) **Standardizing Selection Strengths to Study Selection in the Wild: A Critical Comparison and Suggestions for the Future** *Bioscience* **62**:1039–1054
- Ma X, Qiao Z, Chen D, Yang W, Zhou R, Zhang W, Wang M (2015) **CYCLIN-DEPENDENT KINASE G2 regulates salinity stress response and salt mediated flowering in Arabidopsis thaliana** *Plant Mol Biol* **88**:287–299
- McManus CJ, May GE, Spealman P, Shteyman A (2014) **Ribosome profiling reveals post-transcriptional buffering of divergent gene expression in yeast** *Genome Res* **24**:422–430
- Melino V, Tester M (2023) **Salt-Tolerant Crops: Time to Deliver** *Annu Rev Plant Biol* **74**:671–696
- Metzger BPH, Duveau F, Yuan DC, Tryban S, Yang B, Wittkopp PJ (2016) **Contrasting Frequencies and Effects of cis-and trans-Regulatory Mutations Affecting Gene Expression** *Mol Biol Evol* **33**:1131–1146
- Meyer E, Aglyamova GV, Matz MV (2011) **Profiling gene expression responses of coral larvae (Acropora millepora) to elevated temperature and settlement inducers using a novel RNA-Seq procedure** *Mol Ecol* **20**:3599–3616
- Meyer RS, Purugganan MD (2013) **Evolution of crop species: Genetics of domestication and diversification** *Nature Reviews Genetics* <https://doi.org/10.1038/nrg3605>
- Miller G, Suzuki N, Ciftci-Yilmaz S, Mittler R (2010) **Reactive oxygen species homeostasis and signalling during drought and salinity stresses** *Plant Cell Environ* **33**:453–467
- Munns R (2002) **Comparative physiology of salt and water stress** *Plant Cell Environ* **25**:239–250
- Munns R, Tester M. (2008) **Mechanisms of salinity tolerance** *Annu Rev Plant Biol* **59**:651–681
- Nadir S, Khan S, Zhu Q, Henry D, Wei L, Lee DS, Chen L (2018) **An overview on reproductive isolation in Oryza sativa complex** *AoB Plants* **10**
- Petrusa LM, Winicov I (1997) **Proline status in salt-tolerant and salt-sensitive alfalfa cell lines and plants in response to NaCl** *Plant Physiol Biochem*
- Picelli S, Björklund ÅK, Faridani OR, Sagasser S, Winberg G, Sandberg R (2013) **Smart-seq2 for sensitive full-length transcriptome profiling in single cells** *Nat Methods* **10**:1096–1098
- Ponce KS, Meng L, Guo L, Leng Y, Ye G (2021) **Advances in Sensing, Response and Regulation Mechanism of Salt Tolerance in Rice** *Int J Mol Sci* **22** <https://doi.org/10.3390/ijms22052254>
- Presotto A, Hernández F, Mercer KL (2019) **Phenotypic selection under two contrasting environments in wild sunflower and its crop-wild hybrid** *Evol Appl* **12**:1703–1717
- Pu J, Yu H, Guo Y (2022) **A Novel Strategy to Identify Prognosis-Relevant Gene Sets in Cancers** *Genes* **13**
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC. (2007) **PLINK: a tool set for whole-genome association and population-based linkage analyses** *Am J Hum Genet* **81**:559–575

- Qin H, Huang R (2020) **The phytohormonal regulation of Na⁺/K⁺ and reactive oxygen species homeostasis in rice salt response** *Mol Breed* **40**
- Qu W, Gurdziel K, Pique-Regi R, Ruden DM (2018) **Lead Modulates trans- and cis-Expression Quantitative Trait Loci (eQTLs) in *Drosophila melanogaster* Heads** *Front Genet* **9**
- Qu X, Zou J, Wang J, Yang K, Wang X, Le J. (2022) **A Rice R2R3-Type MYB Transcription Factor OsFLP Positively Regulates Drought Stress Response via OsNAC** *Int J Mol Sci* **23** <https://doi.org/10.3390/ijms23115873>
- Radanielson AM, Angeles O, Li T, Ismail AM, Gaydon DS (2018) **Describing the physiological responses of different rice genotypes to salt stress using sigmoid and piecewise linear functions** *Field Crops Res* **220**:46–56
- Robinson MD, McCarthy DJ, Smyth GK (2010) **edgeR: a Bioconductor package for differential expression analysis of digital gene expression data** *Bioinformatics* **26**:139–140
- Roy S, Mishra M, Kaur G, Singh S, Rawat N, Singh P, Singla-Pareek SL, Pareek A (2022) **OsCyp2-P, an auxin-responsive cyclophilin, regulates Ca²⁺ calmodulin interaction for an ion-mediated stress response in rice** *Physiol Plant* **174**
- Ruan S-L, Ma H-S, Wang S-H, Fu Y-P, Xin Y, Liu W-Z, Wang F, Tong J-X, Wang S-Z, Chen H-Z (2011) **Proteomic identification of OsCYP2, a rice cyclophilin that confers salt tolerance in rice (*Oryza sativa* L.) seedlings when overexpressed** *BMC Plant Biol* **11**
- Ruffley M, Leventhal L, Hateley S, Rhee SY, Exposito-Alonso M. (2023) **Conflicts in natural selection constrain adaptation to climate change in *Arabidopsis thaliana*** *bioRxiv* <https://doi.org/10.1101/2023.10.16.562583>
- Schuppler U, He PH, John PC, Munns R (1998) **Effect of water stress on cell division and cell-division-cycle 2-like cell-cycle kinase activity in wheat leaves** *Plant Physiol* **117**:667–678
- Shabalin AA (2012) **Matrix eQTL: ultra fast eQTL analysis via large matrix operations** *Bioinformatics* **28**:1353–1358
- Siddiqui MN, Léon J, Naz AA, Ballvora A (2021) **Genetics and genomics of root system variation in adaptation to drought stress in cereal crops** *J Exp Bot* **72**:1007–1019
- Signor SA, Nuzhdin SV (2018) **The Evolution of Gene Expression in cis and trans** *Trends Genet* **34**:532–544
- Smith EN, Kruglyak L (2008) **Gene-environment interaction in yeast gene expression** *PLoS Biol* **6**
- Snoek LB, Terpstra IR, Dekter R, Van den Ackerveken G, Peeters AJM. (2012) **Genetical Genomics Reveals Large Scale Genotype-By-Environment Interactions in *Arabidopsis thaliana*** *Front Genet* **3**
- Sterken MG, Nijveen H, van Zanten M, Jiménez-Gómez JM, Geshnizjani N, Willems LAJ, Rienstra J, Hilhorst HWM, Ligterink W, Snoek BL (2023) **Plasticity of maternal environment-dependent expression-QTLs of tomato seeds** *Theor Appl Genet* **136**
- Stinchcombe JR, Agrawal AF, Hohenlohe PA, Arnold SJ, Blows MW (2008) **Estimating nonlinear selection gradients using quadratic regression coefficients: double or nothing?** *Evolution*

62:2435–2440

Taïbi K, Taïbi F, Ait Abderrahim L, Ennajah A, Belkhodja M, Mulet JM (2016) **Effect of salt stress on growth, chlorophyll content, lipid peroxidation and antioxidant defence systems in *Phaseolus vulgaris* L** *S Afr J Bot* **105**:306–312

Tari I, Kiss G, Deér AK, Csiszár J, Erdei L, Gallé Á, Gémes K, Horváth F, Poór P, Szepesi Á, Simon LM (2010) **Salicylic acid increased aldose reductase activity and sorbitol accumulation in tomato plants under salt stress** *Biol Plant* **54**:677–683

Tsai Y-C, Chen K-C, Cheng T-S, Lee C, Lin S-H, Tung C-W (2019) **Chlorophyll fluorescence analysis in diverse rice varieties reveals the positive correlation between the seedlings salt tolerance and photosynthetic efficiency** *BMC Plant Biol* **19**

Van Bel M, Silvestri F, Weitz EM, Kreft L, Botzki A, Coppens F, Vandepoele K. (2022) **PLAZA 5.0: extending the scope and power of comparative and functional genomics in plants** *Nucleic Acids Res* **50**:D1468–D1474

Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, Banks E, Garimella KV, Altshuler D, Gabriel S, DePristo MA. (2013) **From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline** *Curr Protoc Bioinformatics* **43**:11–11

Van Zandt PA, Mopper S. (2002) **Delayed and carryover effects of salinity on flowering in *Iris hexagona* (Iridaceae)** *Am J Bot* **89**:1847–1851

Vasimuddin M, Misra S, Li H, Aluru S (2019) **Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems** *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)* :314–324

Wadgyman SM, Lowry DB, Gould BA, Byron CN, Mactavish RM, Anderson JT (2017) **Identifying targets and agents of selection: innovative methods to evaluate the processes that contribute to local adaptation** *Methods Ecol Evol* **8**:738–749

Wagner GP, Lynch VJ (2008) **The gene regulatory logic of transcription factor evolution** *Trends Ecol Evol* **23**:377–385

Wang H, Zhang M, Guo R, Shi D, Liu B, Lin X, Yang C (2012) **Effects of salt stress on ion balance and nitrogen metabolism of old and young leaves in rice (*Oryza sativa* L** *BMC Plant Biol* **12**

Wang J, Zhang Z (2021) **GAPIT Version 3: Boosting Power and Accuracy for Genomic Association and Prediction** *Genomics Proteomics Bioinformatics* **19**:629–640

Wang W, Mauleon R, Hu Z, Chebotarov D, Tai S, Wu Z, Li M, Zheng T, Fuentes RR, Zhang F, Mansueto L, Copetti D, Sanciangco M, Palis KC, Xu J, Sun C, Fu B, Zhang H, Gao Y, Zhao X, Shen F, Cui X, Yu H, Li Z, Chen M, Detras J, Zhou Y, Zhang X, Zhao Y, Kudrna D, Wang C, Li R, Jia B, Lu J, He X, Dong Z, Xu J, Li Y, Wang M, Shi J, Li J, Zhang D, Lee S, Hu W, Poliakov A, Dubchak I, Ulat VJ, Borja FN, Mendoza JR, Ali J, Li J, Gao Q, Niu Y, Yue Z, Naredo MEB, Talag J, Wang X, Li J, Fang X, Yin Y, Glaszmann J-C, Zhang J, Li J, Hamilton RS, Wing RA, Ruan J, Zhang G, Wei C, Alexandrov N, McNally KL, Li Z, Leung H (2018) **Genomic variation in 3,010 diverse accessions of Asian cultivated rice** *Nature* **557**:43–49

Watowich MM, Chiou KL, Montague MJ, Cayo Biobank Research Unit, Simons ND, Horvath JE, Ruiz-Lambides AV, Martínez MI, Higham JP, Brent LJN, Platt ML, Snyder-Mackler N. (2022)

Natural disaster and immunological aging in a nonhuman primate *Proc Natl Acad Sci U S A* **119** <https://doi.org/10.1073/pnas.2121663119>

Wei H, Wang X, He Y, Xu H, Wang L (2021) **Clock component OsPRR73 positively regulates rice salt tolerance by modulating OsHKT2;1-mediated sodium homeostasis** *EMBO J* **40**

West G, Inzé D, Beemster GTS. (2004) **Cell cycle modulation in the response of the primary root of Arabidopsis to salt stress** *Plant Physiol* **135**:1050–1058

West MAL, Kim K, Kliebenstein DJ, van Leeuwen H, Michelmore RW, Doerge RW, St Clair DA. (2007) **Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in Arabidopsis** *Genetics* **175**:1441–1450

Wilkins O, Hafemeister C, Plessis A, Holloway-Phillips M-M, Pham GM, Nicotra AB, Gregorio GB, Jagadish SVK, Septiningsih EM, Bonneau R, Purugganan M (2016) **EGRINs (Environmental Gene Regulatory Influence Networks) in Rice That Function in the Response to Water Deficit, High Temperature, and Agricultural Environments** *Plant Cell* **28**:2365–2384

Wittkopp PJ, Haerum BK, Clark AG (2008) **Regulatory changes underlying expression differences within and between Drosophila species** *Nat Genet* **40**:346–350

Wu Z, Chen L, Yu Q, Zhou W, Gou X, Li J, Hou S (2019) **Multiple transcriptional factors control stomata development in rice** *New Phytol* **223**:220–232

Xu X, Yuan L, Xie Q (2022) **The circadian clock ticks in plant stress responses** *Stress Biol* **2**

Yang J, Lee SH, Goddard ME, Visscher PM (2011) **GCTA: a tool for genome-wide complex trait analysis** *Am J Hum Genet* **88**:76–82

Yang Y, Guo Y (2018) **Unraveling salt stress signaling in plants** *J Integr Plant Biol* **60**:796–804

Yuan F, Yang H, Xue Y, Kong D, Ye R, Li C, Zhang J, Theprungsirikul L, Shrift T, Krichilsky B, Johnson DM, Swift GB, He Y, Siedow JN, Pei Z-M (2014) **OSCA1 mediates osmotic-stress-evoked Ca²⁺ increases vital for osmosensing in Arabidopsis** *Nature* **514**:367–371

Zhang H, Zhu J, Gong Z, Zhu J-K (2022) **Abiotic stress responses in plants** *Nat Rev Genet* **23**:104–119

Zhang R, Wang Y, Hussain S, Yang S, Li R, Liu S, Chen Y, Wei H, Dai Q, Hou H (2022) **Study on the Effect of Salt Stress on Yield and Grain Quality Among Different Rice Varieties** *Front Plant Sci* **13**

Author information

Sonal Gupta

Center for Genomics and Systems Biology, New York University, New York, United States

Simon C Groen

Center for Genomics and Systems Biology, New York University, New York, United States, Department of Nematology and Department of Botany & Plant Sciences, University of

California, Riverside, Riverside, United States, Center for Plant Cell Biology, Institute for Integrative Genome Biology, University of California, Riverside, Riverside, United States
ORCID iD: [0000-0003-4538-8865](https://orcid.org/0000-0003-4538-8865)

Maricris L Zaidem

Center for Genomics and Systems Biology, New York University, New York, United States,
Department of Biology, University of Oxford, Oxford, United Kingdom

Andres Godwin C Sajise

International Rice Research Institute, Los Baños, Philippines

Irina Calic

Department of Biological Sciences, Fordham University, New York, United States, Inari
Agriculture Nv, Gent, Belgium

Mignon A Natividad

International Rice Research Institute, Los Baños, Philippines

Kenneth L McNally

International Rice Research Institute, Los Baños, Philippines

Georgina V Vergara

International Rice Research Institute, Los Baños, Philippines, Institute of Crop Science,
University of the Philippines, Los Baños, Philippines

Rahul Satija

Center for Genomics and Systems Biology, New York University, New York, United States, New
York Genome Center, New York, United States

Steven J Franks

Department of Biological Sciences, Fordham University, New York, United States

Rakesh K Singh

International Rice Research Institute, Los Baños, Philippines, International Center for
Biosaline Agriculture, Dubai, United Arab Emirates

Zoé Joly-Lopez

Center for Genomics and Systems Biology, New York University, New York, United States,
Département de Chimie, Université du Québec à Montréal, Montreal, Canada
ORCID iD: [0000-0002-7926-322X](https://orcid.org/0000-0002-7926-322X)

For correspondence: joly-lopez.zoe@uqam.ca

Michael D Purugganan

Center for Genomics and Systems Biology, New York University, New York, United States
ORCID iD: [0000-0002-4419-2345](https://orcid.org/0000-0002-4419-2345)

For correspondence: mp132@nyu.edu

Editors

Reviewing Editor

Benjamin Blackman

University of California, Berkeley, Berkeley, United States of America

Senior Editor

Sergio Rasmann

University of Neuchâtel, Neuchâtel, Switzerland

Reviewer #2 (Public review):

The authors investigate the gene expression variation in a rice diversity panel under normal and saline growth conditions to gain insight into the underlying molecular adaptive response to salinity. They present a convincing case to demonstrate that environment stress can induce selective pressure on gene expression, which is in agreement with their earlier study (Groen et al, 2020). The data seems to be a good fit for their study and overall the analytic approach is robust.

(1) The work started by investigating the effect of genotype and their interaction at each transcript level using 3'-end-biased mRNA sequencing, and detect a wide-spread GXE effect. Later, using the total filled grain number as a proxy of fitness, they estimated the strength of selection on each transcript and reported stronger selective pressure in saline environment. However, this current framework rely on precise estimation of fitness and, therefore can be sensitive to the choice of fitness proxy.

(2) Furthermore, the authors decomposed the genetic architecture of expression variation into cis- and trans-eQTL in each environment separately and reported more unique environment specific trans-eQTLs than cis-. The relative contribution of cis- and trans-eQTL depends on both the abundance and effect size. I wonder why the latter was not reported while comparing these two different genetic architectures. If the authors were to compare the variation explained by these two categories of eQTL instead of their frequency, would the inference that trans-eQTLs are primarily associated with expression variation still hold?

(3) Next, the authors investigated the relationship between cis- and trans-eQTLs at transcript level and revealed an excess of reinforcement over compensation pattern. Here, I struggle to understand the motivation for testing the relationship by comparing the effect of cis-QTL with the mean effect of all trans-eQTLs of a given transcript. My concern is that taking the mean can diminish the effect of small trans-eQTLs potentially biasing the relationship towards the large-effect eQTLs.

Comments on latest version:

After the revision, the article has improved substantially. The authors have addressed most of my concerns and suggestions, except for testing the eQTL reinforcement/compensation relationship in the context of genetic architecture. I understand the motivation for testing this relationship at the gene level to determine whether it arises from directional or stabilizing selection, rather than examining it in a cis-trans pairwise fashion. However, I find the definition of this relationship unclear. The authors state in line 824 that "Genes were defined as compensating and reinforcing if they had at least 60% of individuals with opposite and same cis-trans allelic configuration, respectively." In contrast, if I understood correctly, the response to reviewers describes the relationship as reinforcing if the cis-eQTL effect is in the same direction as the mean effect of all the detected trans-eQTLs. I would request that the authors clarify their method of defining this relationship. Also, one should be aware of the

fact that this relationship can evolve neutrally. Since there was no formal test performed to say it is otherwise, the authors might need to interpret the relationship carefully.

While the authors explain the possible factors that could lead to the trend of observing widespread genotype-dependent plastic response without significant genotype-dependent plasticity for fitness (L142), it is also important to consider the time axis. While filled grain serves as a proxy for fitness over time, gene expression profiles provide only a snapshot at a given time point. Therefore, temporal GxE dynamics may also play a role here.

Also, I am a little surprised by not mentioning anything about the code availability in this manuscript. I would request the authors to incorporate that in the revised version.

<https://doi.org/10.7554/eLife.99352.2.sa4>

Reviewer #3 (Public review):

In this work, the authors conducted a large-scale field trial of 130 indica accessions in normal vs. moderate salt stress conditions. The experiment consists of 3 replicates for each accession in each treatment, making it 780 plants in total. Leaf transcriptome, plant traits, and final yield were collected. Starting from a quantitative genetics framework, the authors first dissected the heritability and selection forces acting on gene expression. After summarizing the selection force acting on gene expression (or plant traits) in each environment, the authors described the difference in gene expression correlation between environments. The final part consists of eQTL investigation and categorizing cis- and trans-effects acting on gene expression.

Building on the group's previous study and using a similar methodology (Groen et al. 2020, 2021), the unique aspect of this study is in incorporating large-scale empirical field works and combining gene expression data with plant traits. Unlike many systems biology studies, this study strongly emphasizes the quantitative genetics perspective and investigates the empirical fitness effects of gene expression data. The large amounts of RNAseq data (one sample for each plant individual) also allow heritability calculation. This study also utilizes the population genetics perspective to test for traces of selection around eQTL. As there are too many genes to fit in multiple regression (for selection analysis) and to construct the G-matrix (for breeder's equation), grouping genes into PCs is a very good idea.

In the previous review, three major points were mentioned. The manuscript was modified, and here I briefly summarize them as a reference for future works:

- (1) The separate sections (selection analysis, transcript correlation structure change, and eQTL) could use better integration.
- (2) It would be worth considering joint analyses integrating the two environments together.
- (3) Whether gene expression PCs or unique expression modules should be used in selection analyses.

Regarding whether to use PCs or WGCNA eigengenes to summarize gene expression for selection analyses, the authors reported that only a few WGCNA eigengenes were under selection, citing this observation as the rationale for choosing PC over eigengenes. However, as the relative false positive-negative rates of these choices likely require another dedicated study to explore, at this stage, it might be premature to state which method is better based on which gives more positive results. On one hand, one could easily imagine that plants screwed up by salinity have erratic genomewide expression and become extreme data points on the PCs, making the PCs a good proxy to correlate with fitness. On the other, it remains to be discussed whether this genomewide screwed-up-ness is what we want to measure in this study or whether we should focus on more dedicated gene modules instead. I suggest the

authors acknowledge both possibilities. In this revision, I do not see relevant WGCNA results (as mentioned in the previous response letter) reported.

Figure 4: The observation that chlorophyll a content is under negative selection under BOTH conditions is a bit counterintuitive. The manuscript only mentioned "consistent with the general trend for reduced photosynthesis under salinity stress" (line 329) but did not mention why this increased fitness, even in normal conditions.

<https://doi.org/10.7554/eLife.99352.2.sa3>

Reviewer #4 (Public review):

The manuscript examines how patterns of selection on gene expression differ between a normal field environment and a field environment with elevated salinity based upon transcript abundances obtained from leaves of a diverse panel of rice germplasm. In addition, the manuscript also maps expression QTL (eQTL) that explains variation in each environment. One highlight from the mapping is that a small group of trans-mapping regulators explains some gene expression variation for large sets of transcripts in each environment.

The overall scope of the datasets is impressive, combining large field studies that capture information about fecundity, gene expression, and trait variation at multiple sites. The finding related to patterns indicating increased LD among eQTLs that have cis-trans compensatory or reinforcing effects is interesting in the context of other recent work finding patterns of epistatic selection. The authors have made some changes that address previous comments. However, some analyses in the manuscript remain less compelling or do not make the most from the value of collected data. Although the authors have made several improvements to the precision with which field-specific terminology is applied and to the language chosen when interpreting analytical findings, additional changes to improve these aspects of the manuscript remain necessary.

Selection of gene expression: One strength of the dataset is that gene expression and fecundity were measured for the same genotypes in multiple environments. However, the selection analyses are largely conducted within environments. Addition of phenotypic selection analyses that jointly analyze gene expression across environments and or selection on reaction norms would be worthwhile.

Gene expression trade-offs: The terminology and possibly methods involved in the section on gene expression trade-offs need amendment. I specifically recommend discontinuing reference to the analysis presented as an analysis of antagonistic pleiotropy (rather than more general as trade-offs) because pleiotropy is defined as a property of a genotype, not a phenotype. Gene expression levels are a molecular phenotype, influenced by both genotype and the environment. By conducting analyses of selection within environments as reported, the analysis does not account for the fact that the distribution of phenotypic values, the fitness surface, or both may differ across environments. Thus, this presents a very different situation than asking whether the genotypic effect of a QTL on fitness differs across environments, which is the context in which the contrasting terms antagonistic pleiotropy and conditional neutrality have been traditionally applied. The results reported do not persuasively support the assertion made in the response to reviewers that the terminology is reasonable due to strong coupling between genotype and phenotype. A more interesting analysis would be to examine whether the covariance of phenotype with fitness has truly changed between environments or whether the phenotypic distribution has just shifted to a different area of a static fitness surface.

Biological processes under selection / Decoherence: In the initial review, it was noted that PCA is likely not the most ideal way to cluster genes to generate consolidated metrics for a selection gradient analysis. Because individual genes will contribute to multiple PCs, the current fractional majority-rule method applied to determine whether a PC is under direct or indirect selection for increased or decreased expression comes across as arbitrary and with the potential for double-counting genes. A gene co-expression network analysis could be more appropriate, as genes only belong to one module and one can examine how selection is acting on the eigengene of a co-expression module. Building gene co-expression modules would also provide a complementary and more concrete framework for evaluating whether salinity stress induces "decoherence" and which functional groups of genes are most impacted. Although results of co-expression network analyses are now briefly discussed in the response to reviewers, the findings and their relationship to the PCA/"decoherence" analyses are not reported in the manuscript.

Selection of traits: Having paired organismal and molecular trait data is a strength of the manuscript, but the organismal trait data are underutilized. The manuscript as written only makes weak indirect inferences based on GO categories or assumed gene functions to connect selection at the organismal and molecular levels. After prompted by the initial reviews to test for correspondence between SNPs that explain organismal and gene expression trait variation or co-variance of co-expression module variation and trait variation, the response to reviewers indicates finding negative results. These findings should be included in the manuscript text and discussed.

Genetic architecture of gene expression variation: More descriptive statistics of the eQTL analysis have been included, although additional information about the variation in these measures within environments would be useful. The motivation for featuring patterns of cis-trans compensation specifically for the results obtained under high salinity conditions remains unclear to me. If the lines sampled have predominantly evolved under low salinity conditions, and the hypothesis being evaluated relates to historical experience of stabilizing selection, then evaluating the eQTL patterns under normal conditions provides the more relevant test of the hypothesis.

Lines 280-282: The revised sentence continues to read as an overstatement and merits additional revision with citations.

Lines 379-381: Following revision, it still remains unclear how the interpretation follows from the above analysis; the inference as written goes significantly beyond what may be specifically inferable from the result.

<https://doi.org/10.7554/eLife.99352.2.sa2>

Reviewer #5 (Public review):

Summary:

The researchers examined selection across multiple levels, including gene expression, biological processes, and regulatory mechanisms, with a particular focus on comparing selection between different environmental conditions. They further explored potential evolutionary mechanisms. This is made possible with a comprehensive dataset comprising gene expression data from 130 accessions with three replicates collected in two environments in the field, genomic data from 125 genotypes, and associated physiological traits. The findings have significant implications for understanding the evolution of stress adaptation, and the identified possible genes and pathways for further investigation.

The researchers began by focusing on the selection of gene expression across two environments, comparing the number of genes under selection and the effect sizes, as well as examining how selection in each environment acts on the same individual genes. They then expanded their analysis to consider selection in biological processes, investigating the relationships between selection acting on individual genes within processes and selection acting among different processes. Additionally, they explored selection at the organismal level by examining traits.

The study further transitioned from analyzing individual gene expression to investigating gene-gene interactions. They briefly examined correlation variation among gene pairs between the two conditions, identifying pairs with rewired interactions that suggest potential selection on gene regulation or the effect of rewiring on tolerance. The researchers then delved into the genetic architecture underlying these patterns by mapping eQTLs. Their comparison of cis- and trans-eQTLs revealed that trans-eQTLs were more variable across conditions. Notably, they identified hotspots representing master regulators that possibly underlie the greater variability of trans-eQTLs across environments. They further discovered that trans-eQTLs are generally under purifying selection (particularly in salt conditions), while cis-eQTLs are under balancing selection, exhibiting higher nucleotide diversity. As for how cis- and trans-eQTL effects combine at the level of individual genes, more are found to be reinforced and the hypothesis of genetic fixation on cis- and trans-eQTL effects combination is further tested.

Strengths:

A key strength of this study is its comprehensive approach, extending beyond the analysis of gene expression to include gene-gene interactions, genetic architectures, and selections of genetic regulation factors. The exploration of gene expression selection through its connection with fitness, as introduced in the researchers' previous work, provides valuable insights into the role of gene expression in adaptation. The study investigates selection across multiple levels of biological responses, including individual gene expression, genes associated with biological processes, gene-gene interactions, and the underlying genetic architecture. The experimental design enables a direct comparison of selection between control and salinity conditions, which sheds light on the effects of stress on selection and the dynamics of adaptation to stress. Additionally, the manuscript is well-written, with a clear connection to current literature. The discussion effectively integrates findings with broader implications, making it a satisfying read.

Weaknesses:

The lack of formal testing for environment-specific selections (e.g., selection of gene expression specifically in salinity stress, PCs, or traits) is a major limitation, as previous reviewers have flagged. Explicit tests of eQTLs variation between conditions are introduced, so similar formal tests should also be introduced in selection sections. For example, a formal test of selections of gene expression might be helpful to solve variance/mean- standardization concerns between two environments.

Additionally, some aspects of the analysis appear somewhat arbitrary and could benefit from further sensitivity testing. Line 203: The concern about bias in detecting more CN than AP, as mentioned by the authors and previously flagged by reviewers, does not seem fully resolved with the current methods given the arbitrary cut-off. Incorporating additional tests suggesting the conclusion is insensitive to the cutoff would be very helpful. Similar is the classification of genes into compensatory and reinforcing categories based on 60% of individuals as a cutoff.

While this study focuses on gene regulation, its connection with the selection of gene expression and biological pathways is not well integrated. In particular, the discovery of

eQTLs is not explicitly linked to gene expression selection or biological pathways, leaving this relationship underexplored. Suggestive comments: Currently the summarization of selection is based on eQTLs. It would be interesting to also summarize the selection patterns identified from previous sections based on genes being cis/trans-regulated. Moreover, it might be interesting to see if there is more loss or gain of eQTLs under salt stress and their functions. The current results mentioned variations of eQTLs but not clear if they are loss or gain. E.g., one way is to identify genes related to cis and trans-eQTLs and see their correlation changes with genes being regulated using CILP (also as a way to informatively narrow down gene pairs for CILP).

Similarly, the section on selection at the organismal trait level appears disconnected from the rest of the analysis (e.g., if it is not tested to be related to other features, mentioning why it might not be related would be helpful). Admittedly, the discussion of how biological processes discovered at different levels integrate together is helpful.

Other comments: given there is no comparison between loss of coherence (correlations) and gain of coherence under salinity stress to show the dominant role of decoherence, maybe need to also discuss the genes and processes related to the gain of coherence? This is because the understanding of activation (gain of coherence) of some regulations/processes under stress conditions could also be interesting. It is not clear if decoherence (e.g., lines 293-296) refers to significant correlation changes or just loss of the correlation in salinity stress.

<https://doi.org/10.7554/eLife.99352.2.sa1>

Author response:

The following is the authors' response to the original reviews.

Reviewer #1 (Public Review):

Summary:

Understanding the mechanisms of how organisms respond to environmental stresses is a key goal of biological research. Assessment of transcriptional responses to stress can provide some insights into those underlying mechanisms. The researchers quantified traits, fitness, and gene expression (transcriptional) response to salinity stress (control vs stress treatments) for 130 accessions of rice (three replicates for each accession), which were grown in the field in the Philippines. This experimental design allowed for many different types of downstream analyses to better understand the biology of the system. These analyses included estimating the strength of selection imposed on transcription in each environment, evaluating possible trade-offs in gene expression, testing whether salinity induces transcriptional decoherence, and conducting various eQTL-type analyses.

Strengths:

The study provides an extensive analysis of gene expression responses to stress in rice and offers some insights into underlying mechanisms of salinity responses in this important crop system. The fact that the study was conducted under field conditions is a major plus, as the gene expression responses to soil salinity are more realistic than if the study was conducted in a greenhouse or growth chamber. The preprint is generally well-written and the methods and results are mostly well-described.

Weaknesses:

While the study makes good use of analyzing the dataset, it is not clear how the current work advances our understanding of gene regulatory evolution or plant responses to soil

salinity generally. Overall, the results are consistent with other prior studies of gene expression and studies of selection across environmental conditions. Some of the framing of the paper suggests that there is more novelty to this study than there is in reality. That said, the results will certainly be useful for those working in rice and should be interesting to scientists interested in how gene expression responses to stress occur under field conditions. I detail other concerns I had about the preprint below:

The abstract on lines 33-35 illustrates some of my concerns about the overstatement of the novelty of the current study. For example, is it really true that the role of gene expression in mediating stress response and adaptation is largely unexplored? There have been numerous studies that have evaluated gene expression responses to stresses in a wide range of organisms. Perhaps, I am missing something critically different about this study. If so, I would recommend that the authors reword this sentence to clarify what gap is being filled by this study. Further, is it really the case that none of them have evaluated how the correlational structure of gene expression changes in response to stresses in plants, as implied in lines 263-265? Don't the various modules and PC analyses of gene expression get at this question?

We have re-worded these sentences, and highlighted the novelty of our work.

There were some places in the methods of the preprint that required more information to properly evaluate. For example, more information should be provided on lines 664-668 about how G, E, and GxE effects were established, especially since this is so central to this study. What programs/software (R? SAS? Other?) were used for these analyses? If R, how were the ANOVAs/models fit? What type of ANOVA was used? How exactly was significance determined for each term? Which effects were considered fixed and which were random? If the goal was to fit mixed models, why not use an approach like voom-limma (Law et al. 2014 Genome Biology)? More details should also be added to lines 688-709 about these analyses, including what software/programs were used for these analyses.

We have added more details in the methods. Also, although we could in principle use voom-limma to fit our mixed model, to be able to partition variance into G, E and G×E, we need to use the function fitExtractVarPartModel (from package VariancePartition) which requires all categorical variables to be modeled as random effects. Therefore, we couldn't model environment as a fixed effect.

One thing that I found a bit confusing throughout was the intermixing of different terms and types of selection. In particular, there seemed to be some inconsistencies with the usage of quantitative genetics terms for selection (e.g. directional, stabilizing) vs molecular evolution terms for selection (e.g. positive, purifying). I would encourage the authors to think carefully about what they mean by each of these terms and make sure that those definitions are consistently applied here.

We have defined the selection terms used in the study and used these terms consistently throughout the manuscript.

It would be useful to clarify the reasons for the inherent bias in the detection of conditional neutrality (CN) and antagonistic pleiotropy (AP; Lines 187-196). It is also not clear to me what the authors did to deal with the bias in terms of adjusting P-value thresholds for CN and AP the way it is currently written. Further, I found the discussion of antagonistic pleiotropy and conditional neutrality to be a bit confusing for a couple of reasons, especially around lines 489-491. First of all, does it really make sense to contrast gene expression versus local adaptation, when lots of local adaptation likely involves

changes in gene expression? Second, the implication that antagonistic pleiotropy is more common for local adaptation than the results found in this study seems questionable. Conditional neutrality appears to be more common for local adaptation as well: see Table 2 of Wadgyamar et al. 2017 Methods in Ecology and Evolution. That all said, it is always difficult to conclude that there are no trade-offs (antagonistic pleiotropy) for a particular locus, as the detecting trade-offs may only manifest in some years and not others and can require large sample sizes if they are subtle in effect.

We have now explained the cause of the inherent bias in the detection of CN, and also elaborated on how we deal with this bias. Also, we have edited our discussion and added relevant citations to indicate both conditional neutrality and antagonistic pleiotropy can lead to local adaptations and added the caveat regarding detecting antagonistic pleiotropy.

Reviewer #2 (Public Review):

The authors investigate the gene expression variation in a rice diversity panel under normal and saline growth conditions to gain insight into the underlying molecular adaptive response to salinity. They present a convincing case to demonstrate that environmental stress can induce selective pressure on gene expression, which is in agreement to their earlier study (Groen et al, 2020). The data seems to be a good fit for their study and overall the analytic approach is robust.

(1) The work started by investigating the effect of genotype and their interaction at each transcript level using 3'-end-biased mRNA sequencing, and detecting a wide-spread GXE effect. Later, using the total filled grain number as a proxy of fitness, they estimated the strength of selection on each transcript and reported stronger selective pressure in a saline environment. However, this current framework relies on precise estimation of fitness and, therefore can be sensitive to the choice of fitness proxy.

We now acknowledge this caveat in the discussion.

(2) Furthermore, the authors decomposed the genetic architecture of expression variation into cis- and trans-eQTL in each environment separately and reported more unique environment-specific trans-eQTLs than cis-. The relative contribution of cis- and trans-eQTL depends on both the abundance and effect size. I wonder why the latter was not reported while comparing these two different genetic architectures. If the authors were to compare the variation explained by these two categories of eQTL instead of their frequency, would the inference that trans-eQTLs are primarily associated with expression variation still hold?

We have now also reported the effect sizes for both cis- and trans-eQTLs in the two environments and showed that the trans-eQTLs have higher effect sizes as compared to cis-eQTLs, indicating that they are able to explain higher proportion of variation in transcript abundances in the two environments.

(3) Next, the authors investigated the relationship between cis- and trans-eQTLs at the transcript level and revealed an excess of reinforcement over the compensation pattern. Here, I struggle to understand the motivation for testing the relationship by comparing the effect of cis-QTL with the mean effect of all trans-eQTLs of a given transcript. My concern is that taking the mean can diminish the effect of small trans-eQTLs potentially biasing the relationship towards the large-effect eQTLs.

We wanted to estimate compensating vs reinforcing effects, which essentially entails identifying genes that have opposing directionality of cis and trans-effects. To get the total trans-effect we decided to take the mean effect of trans-eQTLs. This mean was only used to

identify the compensating/reinforcing genes and although the mean effects diminishes the effect of small trans-eQTLs, this mean was not used in downstream analyses.

Reviewer #3 (Public Review):

In this work, the authors conducted a large-scale field trial of 130 indica accessions in normal vs. moderate salt stress conditions. The experiment consists of 3 replicates for each accession in each treatment, making it 780 plants in total. Leaf transcriptome, plant traits, and final yield were collected. Starting from a quantitative genetics framework, the authors first dissected the heritability and selection forces acting on gene expression. After summarizing the selection force acting on gene expression (or plant traits) in each environment, the authors described the difference in gene expression correlation between environments. The final part consists of eQTL investigation and categorizing cis- and trans-effects acting on gene expression.

Building on the group's previous study and using a similar methodology (Groen et al. 2020, 2021), the unique aspect of this study is in incorporating large-scale empirical field works and combining gene expression data with plant traits. Unlike many systems biology studies, this study strongly emphasizes the quantitative genetics perspective and investigates the empirical fitness effects of gene expression data. The large amounts of RNAseq data (one sample for each plant individual) also allow heritability calculation. This study also utilizes the population genetics perspective to test for traces of selection around eQTL. As there are too many genes to fit in multiple regression (for selection analysis) and to construct the G-matrix (for breeder's equation), grouping genes into PCs is a very good idea.

Building on large amounts of data, this study conducted many analyses and described some patterns, but a central message or hypothesis would still be necessary. Currently, the selection analysis, transcript correlation structure change, and eQTL parts seem to be independent. The manuscript currently looks like a combination of several parallel works, and this is reflected in the Results, where each part has its own short introduction (e.g., 185-187, 261-266, 349-353). It would be great to discuss how these patterns observed could be translated to larger biological insights. On a related note, since this and the previous studies (focusing on dry-wet environments) use a similar methodology, one would also wonder what the conclusions from these studies would be. How do they agree or disagree with each other?

We acknowledge that the manuscript currently presents some analyses in a somewhat independent manner. Although it would be ideal to have a central hypothesis/message, our study is meant to broadly outline the various responses and fitness effects of salinity stress in rice. Throughout the manuscript, we have also included comparisons between our findings and that of our previous studies on drought stress to highlight any consistent themes or novel insights.

Many analyses were done separately for each environment, and results from these two environments are listed together for comparison. Especially for the eQTL part, no specific comparison was discussed between the two environments. It would be interesting to consider whether one could fit the data in more coherent models specifically modeling the X-by-environment effects, where X might be transcripts, PCs, traits, transcript-transcript correlation, or eQTLs.

We do plan to consider fitting models that explicitly incorporate X-by-environment interactions to provide a more detailed understanding of the genetics of plasticity between the two environments, but it is beyond the scope of this paper. This will be explored in a separate report.

As stated, grouping genes into PCs is a good idea, but although in theory, the PCs are orthogonal, each gene still has some loadings on each PC (ie. each PC is not controlled by a completely different set of genes). Another possibility is to use any gene grouping method, such as WGCNA, to group genes into modules and use the PC1 of each module. There, each module would consist of completely different sets of genes, and one would be more likely to separate the biological functions of each module. I wonder whether the authors could discuss the pros and cons of these methods.

We recognize that individual genes can contribute to multiple PCs, and this is precisely why we choose PCA clustering over WGCNA where one gene can belong to only one module. Our aim was to recognize all biological processes that could be under selection in either environment, and since one gene can be involved in various different processes, we wanted to identify the contribution of these genes to different processes which can be done effectively by a PCA analyses.

Reviewer #4 (Public Review):

The manuscript examines how patterns of selection on gene expression differ between a normal field environment and a field environment with elevated salinity based on transcript abundances obtained from leaves of a diverse panel of rice germplasm. In addition, the manuscript also maps expression QTL (eQTL) that explains variation in each environment. One highlight from the mapping is that a small group of trans-mapping regulators explains some gene expression variation for large sets of transcripts in each environment. The overall scope of the datasets is impressive, combining large field studies that capture information about fecundity, gene expression, and trait variation at multiple sites. The finding related to patterns indicating increased LD among eQTLs that have cis-trans compensatory or reinforcing effects is interesting in the context of other recent work finding patterns of epistatic selection. However, other analyses in the manuscript are less compelling or do not make the most of the value of collected data. Revisions are also warranted to improve the precision with which field-specific terminology is applied and the language chosen when interpreting analytical findings.

Selection of gene expression:

One strength of the dataset is that gene expression and fecundity were measured for the same genotypes in multiple environments. However, the selection analyses are largely conducted within environments. The addition of phenotypic selection analyses that jointly analyze gene expression across environments and or selection on reaction norms would be worthwhile.

We do plan to consider fitting models that explicitly incorporate G×E interactions to provide a more detailed understanding of the genetics of plasticity between the two environments, but it is beyond the scope of this paper. This will be explored in a separate report.

Gene expression trade-offs:

The terminology and possibly methods involved in the section on gene expression trade-offs need amendment. I specifically recommend discontinuing reference to the analysis presented as an analysis of antagonistic pleiotropy (rather than more general trade-offs) because pleiotropy is defined as a property of a genotype, not a phenotype. Gene expression levels are a molecular phenotype, influenced by both genotype and the environment. By conducting analyses of selection within environments as reported, the analysis does not account for the fact that the distribution of phenotypic values, the fitness surface, or both may differ across environments. Thus, this presents a very

different situation than asking whether the genotypic effect of a QTL on fitness differs across environments, which is the context in which the contrasting terms antagonistic pleiotropy and conditional neutrality have been traditionally applied. A more interesting analysis would be to examine whether the covariance of phenotype with fitness has truly changed between environments or whether the phenotypic distribution has just shifted to a different area of a static fitness surface.

We recognize that pleiotropy is a property of a genotype, and not phenotype, but since our phenotype (gene expression) is strongly coupled with the genotype, we choose to call trade-offs as antagonistic pleiotropy. That being said, we did test whether the covariance of gene expression with phenotype significantly varies between environments, and found that to indeed be the case.

Biological processes under selection / Decoherence: PCs are likely not the most ideal way to cluster genes to generate consolidated metrics for a selection gradient analysis. Because individual genes will contribute to multiple PCs, the current fractional majority-rule method applied to determine whether a PC is under direct or indirect selection for increased or decreased expression comes across as arbitrary and with the potential for double-counting genes. A gene co-expression network analysis could be more appropriate, as genes only belong to one module and one can examine how selection is acting on the eigengene of a co-expression module. Building gene co-expression modules would also provide a complementary and more concrete framework for evaluating whether salinity stress induces "decoherence" and which functional groups of genes are most impacted.

We recognize that individual genes can contribute to multiple PCs, and this is precisely why we choose PCA clustering over WGCNA where one gene can belong to only one module. Our aim was to recognize all biological processes that could be under selection in either environment, and since one gene can be involved in various different processes, we wanted to identify the contribution of these genes to different processes which can be done effectively by a PCA analyses. But again as pointed out by the reviewer, our PCs did contain contribution (even negligible) of each gene, so to identify the 'primary' biological processes represented by the PCs, we chose the majority rule. As for testing decoherence, we agree that a co-expression module analyses would have provided additional support to the specific test performed in our manuscript, but since it would just be additional support, we choose to not add it in the manuscript.

But based on the recommendation of the reviewer(s), we did perform a WGCNA analyses and found a total of 14 and 13 modules in normal and saline conditions, of which 0 and 2 modules (with no significant GO enrichment) were under directional selection. This supports our reasoning of potentially missing on identification of processes under selection.

Selection of traits:

Having paired organismal and molecular trait data is a strength of the manuscript, but the organismal trait data are underutilized. The manuscript as written only makes weak indirect inferences based on GO categories or assumed gene functions to connect selection at the organismal and molecular levels. Stronger connections could be made for instance by showing a selection of co-expression module eigengene values that are also correlated with traits that show similar patterns of selection, or by demonstrating that GWAS hits for trait variation co-localize to cis-mapping eQTL.

We did perform a GWAS for all the traits collected in both normal and saline environment, and only found significant hits for fecundity (in both normal and saline environment) and chlorophyll_a content (in the saline environment). But these regions did not overlap with any

candidate genes or cis-mapping eQTL. Hence we choose to mention it in the manuscript. Additionally, using the WGCNA modules, we found that the only two module under selection in the saline environment were not significantly correlated with any of the traits measured.

Genetic architecture of gene expression variation:

The descriptive statistics of the eQTL analysis summarize counts of eQTLs observed in each environment, but these numbers are not broken down to the molecular trait level (e.g., what are the median and range of cis- and trans-eQTLs per gene). In addition, genetic architecture is a combination of the numbers and relative effect sizes of the QTLs. It would be useful to provide information about the relative distributions of phenotypic variance explained by the cis- vs. trans- eQTLs and whether those distributions vary by environment. The motivation for examining patterns of cis-trans compensation specifically for the results obtained under high salinity conditions is unclear to me. If the lines sampled have predominantly evolved under low salinity conditions and the hypothesis being evaluated relates to historical experience of stabilizing selection, then my intuition is that evaluating the eQTL patterns under normal conditions provides the more relevant test of the hypothesis.

We have added the median number of eQTLs per gene in each environment. Additionally, we recognize that genetic architecture is a combination of numbers and effect size, and we have added information regarding the effect sizes of eQTLs by type and by environment as recommended by another reviewer. We did explore the distributions of phenotypic variance explained by the cis- vs. trans- eQTLs as recommended here, and found that trans-eQTLs explain more phenotypic variance than cis-eQTLs in both environments and that the distribution of either type of eQTL does not vary by environment. We are choosing to not add this in the main text due to space limitations. Lastly, we examined the patterns of cis-trans compensation/reinforcement under both normal and salinity conditions and have compared and contrasted the results from both in the main text.

Recommendations for the authors:

Reviewer #1 (Recommendations For The Authors):

Lines 126: I would recommend citing those who originally developed the 3' end targeted RNA sequencing methods (e.g. Meyer et al 2011 Molecular Ecology).

We have cited the recommended paper.

Lines 128-130: It would be useful to include a description here of what models were fit to the data to partition out G, E, and GxE effects.

Due to space limitations, we have in brief added a sentence to this effect.

Line 139: I would suggest changing "found little" to "no" since the test was not significant.

The sentence has been modified to say no evidence.

Line 313: I think you mean directional selection instead of positive selection.

We have corrected the text

Lines 362-363: Would the authors also expect an enrichment of reinforcing genes for most scenarios where that has been divergent selection, such as local adaptation among

populations?

Based on our hypothesis, we would indeed expect an enrichment of reinforcing genes for scenarios of local adaptation where different alleles are maintained in different populations due to local adaptation.

Reviewer #3 (Recommendations For The Authors):

Figures 1d-e are not mentioned in the Results.

The figures have been referenced in appropriate places.

Lines 41-45: Terms such as reinforcement and compensation need to be explained in this specific context. Also "different selection regimes" is a bit broad and vague.

Due to word-count limitation, we are choosing to not elaborate the terms reinforcement and compensation in the abstract (since these are commonly used in the literature, and we have also defined these in the main text). Additionally, we now explicitly state the selection pressures associated with cis and trans eQTLs.

Table 1: Please explain S and C in the footnote.

We have added the recommended footnote

Figures: Some panel labels (a, b, c...) are mingled with the graphs.

We are re-made our figure such that the panel labels do not mingle with the plots.

Lines 588-591: font.

Modified

Lines 620-633: Please describe how these RNAseq libraries were allocated/pooled into different sequencing lanes to avoid potential batch effects among sequencing lanes.

The sequencing was performed on the same Illumina NextSeq 500 machine and we have added the sequencing libraries/pool plan in the methods (lines 688-689).

Lines 690-692: At the beginning of this paragraph, it was mentioned that the un-standardized coefficients were estimated. But here, it seems like the transcript data were already standardized in the data preparation step. What do lines 687-688 refer to? Further standardizing those estimated coefficients so that the whole distribution has mean=0 and sd=1?

Thank you pointing out our oversight. We checked our scripts and data preparation did not include transcript standardization, and we have removed the above line from the manuscript.

Lines 705-711: Please explain why assigning the positive/negative selection status for each gene is important. "Positive selection" here is defined as genes whose increased expression also increases fitness, but traditionally positive selection was defined as "the derived state is favored over the ancestral state". For a gene whose ancestral expression is high but lower expression increases fitness in this experiment, could we also say this gene is under positive selection? Given that we don't know the ancestral state here,

maybe the authors could explain whether this definition is necessary. Also, given that many genes positively or negatively regulate each other in a pathway, it is also unclear whether it is necessary to assign the positive/negative status for a PC using the majority rule (lines 710-711).

We have now defined the different selection terms with respect to our study and use them consistently throughout the manuscript.

Lines 711-715: If I understand correctly, PCs were used as traits, and by definition PCs should all be orthogonal. Is this section saying only retaining PCs whose correlation < 0.6 with each other? What is the rationale?

PCA were performed on transcript abundance and the resulting orthogonal PCs explaining over 0.5% variance were all retained for selection analyses.

We also performed selection analyses on the functional traits measured in the field, but since these functional traits are correlated (and as such would not satisfy the independent variable requirement of regression analyses), we retained only those functional traits which had a Pearson correlation coefficient < 0.6.

Line 729: Please briefly describe what CLIP is doing.

We have added the required description.

Lines 736-741: The accession numbers do not add up to 125.

Thank you for catching our oversight. We have edited the text, and now the numbers add upto 125.

Line 796: Please remind readers where these 247k SNPs come from. Supposedly all accessions have been whole-genome sequenced, so the total number of SNPs should be larger than this.

We have detailed method detailing how the SNPs were obtained and processed in the lines preceding this. Indeed the number of SNPs would have been much bigger, but the stringent cutoffs and linkage disequilibrium pruning reduced our dataset to about 247k SNPs.

Lines 154-160: This is a bit confusing. The authors first mentioned, for the raw selection differentials, the mean and variance differ between environments, meaning they are misleading (why?). The next sentence then says non-standardized selection differentials will be used.

The mean and variance for transcript abundances vary between the two environments. Because traits are usually measured in different scales, it is recommended to standardize trait values using variance or mean before estimating selection coefficients. Multiplying this variance (or mean) standardized selection differential with heritability gives the expected response to selection in standard deviation (or mean) units. But if the trait variance (or mean) varies between traits or environments, it leads to a conflation between the standardized selection differential and trait variance (or mean), which can be misleading. So to avoid this, and given that our traits (transcript abundance in this case) were all measured on the same scale, we chose to not standardize our trait values and estimated raw selection differentials.

Figure 1 c-e: Please explain how the horizontal axis values were obtained. Is it assuming these selection differentials have a normal distribution of mean=0 & sd=1?

Yes, horizontal axis represents theoretical quantile for selection differential assuming they have a normal distribution with mean=0 and sd=1. This has been added to the figure legend.

Line 162-168: Please clarify this part. What does “general trend towards stronger positive compared to negative selection on gene expression” mean? Does it mean the whole distribution of S is significantly different from 0, the difference in the number of genes in the $S>0$ vs $S<0$ category, or the α -bit-higher median $|S|$ in the $S>0$ vs $S<0$ category? If it is the last one, are the small differences biological meaningful (0.053 vs. 0.047 for control & 0.051 vs. 0.050 for salt conditions), given that the authors defined $|S|<0.1$ as neutral?

By “general trend towards stronger positive compared to negative selection on gene expression”, we mean that more transcripts were under positive directional selection as compared to negative directional selection. We have also clarified this in the text now.

Line 177-178: This sentence implies disruptive selection is more important than stabilizing selection in the saline environment, but the test was not significant (line 176).

Although there was no significant difference in the magnitude of stabilizing vs disruptive selection *within* the saline environment, the number of transcripts experiencing stronger disruptive selection in the saline condition was greater than the number of transcripts experiencing disruptive selection in the normal conditions. And so comparing between conditions, disruptive selection plays an important role in the saline conditions.

Line 188-190: How CN vs. AP was statistically defined was not mentioned in the Methods section.

We have added in the main text within the Results section.

Line 203-214: How do these results fit with the previous observations that almost all transcripts have significant heritability?

Although we do find that all but three transcripts have a have significant genetic effect (and thus have significant heritability), the median broad-sense heritability for 51 antagonistically pleiotropic genes is 0.23. Give that, we would only be able to detect SNPs regulating gene expression with high effect size since our sample size is $n=130$. Additionally, we used a very stringent criteria ($FDR < 0.001$) to define eQTLs. These two factors in combination could lead to us not being able to detect significant eQTLs for AP genes.

Line 246-250: Please explain why the current conclusion would be opposite from the previous study. Supposedly the PCA, G matrix, and breeder's equation were done for each environment separately. It makes sense that the G matrix and response to selection could be different between saline and drought treatments, but for the control treatments in the two studies, do they still differ? Why? Also in Table S7, it would be nice to show the % variation explained by each PC.

Although both our studies had largely overlapping samples, about 20% samples were unique to each study. Additionally, although the site where the study was performed was the same across the two studies, we found significant temporal differences in gene expression due to micro-environmental differences. Both these factors can lead to changes in direct and indirect selection and its response, and we are examining these differences as part of a separate study. We also highlight these caveats in our discussion.

Information on percent explained by each PCs is given in Table S5.

Figure 2b: The vertical axis was labeled as “selection gradient”, but I think the responses to selection (D, I, T) have different units.

We have re-labeled the vertical axis as “selection”.

Reviewer #4 (Recommendations For The Authors):

The manuscript mixes terminology for selection from quantitative genetics with that from population genetics. This is problematic, and the adjectives positive and negative should be replaced as descriptors of selection by instead rewording, for example, positive directional selection as directional selection for higher transcript abundance.

Lines 193-196: The phrasing here reads as if the selection is solely acting on the presence/absence of expression rather than on quantitative variation in expression. During revision, it would be worth considering including an analysis of genes that parses genes that show the presence/absence of variation of expression within or across environments separately from genes that are expressed to non-trivial levels in both environments.

We have modified the sentence in question now. Also, we pre-processed RNA-seq data to remove all transcripts with low expression signals (sigma signal < 20), and further retained only transcripts that had non-trivial expression in at least 10% of the population, which we believe represents presence/absence of variation of expression within or across environments.

Lines 216-231: Is this analysis solely for directional selection? Not clear since previous sections examined both directional and stabilizing selection.

Yes, we performed this analysis for only directional selection, and have clarified this in the text too.

Lines 224-226: The meaning of this sentence is unclear and should be written more concretely.

We have rephrased the sentence to be more clear.

Lines 232-241: The description of the scientific logic here could be read as implying that genes interacting in networks are the sole source of indirect selection. I recommend revising the language to indicate this cause is one of several potential causes.

We have reworded the sentence such that we indicate selection acting on interacting genes is just one of the causes of indirect selection.

The strength of the conclusions of the decoherence analysis should be evaluated in light of caveats with such analyses (see Cai and Des Marais New Phytologist 2023).

We have added the caveat with relevant citation in the manuscript.

Rename this section as “Selection on Organismal Traits”, as the previous sections have also been investigating selection on traits, just molecular traits.

We have renamed the section as recommended

Lines 314-318: Rewrite for clarity. Most environments select for an optimal phenotype; it is just the case here that the phenotypic distribution in the high salinity environment overlaps with the optimum.

We have rephrased and clarified the statement.

Lines 343-345: Rephrase to "These results indicate that natural variation in gene regulation under..."

Rephrased.

Line 354: "most" reads as too strong a descriptor here if the majority is ~60%.

We have reworded the sentence to read "more than half"

Lines 359-361: It is unclear to me how this interpretation follows from the above analysis.

We have reworded the sentence so that the claim follows our analysis.

Line 372: Is the expectation here more specifically one of epistatic selection? Other processes could stochastically lead to the genetic fixation of compensatory/reinforcing variants, but I think only epistasis for fitness would cause the interesting patterns of LD observed.

The expectation here is that certain cis and trans variants only exists to compensate/reinforce, potentially through epistasis. We have clarified this in the text.

Line 405: Change "adaptive organismal responses of organisms" to "organismal responses." As written, the sentence reads as being about plasticity rather than evolutionary responses, which are by populations, not organisms. None of the analyses included the manuscript test specifically test for adaptive plasticity.

Rephrased.

<https://doi.org/10.7554/eLife.99352.2.sa0>