Water impacts nutrient dose responses genome-wide to affect crop production

Joseph Swift1, Mark Adame1, Daniel Tranchina1,2, Amelia Henry3 & Gloria M. Coruzzi1

Changes in nutrient dose have dramatic effects on gene expression and development. One outstanding question is whether organisms respond to changes in absolute nutrient amount (moles) vs. its concentration in water (molarity). This question is particularly relevant to plants, as soil drying can alter nutrient concentration, without changing its absolute amount. To compare the effects of amount vs. concentration, we expose rice to a factorial matrix varying the dose of nitrogen (N) and water (W) over a range of combinations, and quantify transcriptome and phenotype responses. Using linear models, we identify distinct dose responses to either N-moles, W-volume, N-molarity (N/W), or their synergistic interaction (N×W). Importantly, genes whose expression patterns are best explained by N-dose and W interactions (N/W or N×W) in seedlings are associated with crop outcomes in replicated field trials. Such N-by-W responsive genes may assist future efforts to develop crops resilient to increasingly arid, low nutrient soils.
Nitrogen (N) and water (W) are essential inputs for agriculture, however both are scarce in marginal soils, limiting crop production worldwide. Since physiological studies show that N and W act synergistically to drive crop production, we sought to discover how plants integrate responses to changes in N- and W-doses in rice—one of the world’s most important crops.

While widely studied for their individual effects, surprisingly little is known about how biological systems respond to combinations of N and W. Since W serves as a solvent for N, we investigated whether plants respond to N as the absolute amount of N (N-moles) or the concentration of N in W (N-molarity). Plants are well-suited to address this basic question, because it is possible to vary N amount in a changing W environment (through soil drying), and results have practical relevance to crop production.

To this end, we present a factorial treatment matrix approach that systematically varies both N- and W-dose, and use linear equations to model how plants respond to N-amount, W-volume, and their interactions. Our linear models can distinguish between plant responses to changes in N-moles, W-volume, N-molarity (N/W), as well as a synergistic response N×W, both at the level of gene expression and phenotype. Moreover, gene sets responsive to N-moles, W-volume, or to the combined effect of N- and W-dose (N/W and N×W) are conserved across lab and field conditions. Importantly, the genes responding to combinations of N- and W-dose correlate with crop outcomes in the field over two seasons. Such genes that are specifically responsive to combinations of N- and W-dose can potentially be used as expression biomarkers to select rice varieties adapted to increasingly dry, nutrient-poor soils worldwide.

Results

Linear modeling uncovers the impact of W on N-dose responses. To address how W impacts N-dose responses in plants, we designed a factorial N-by-W treatment matrix that could distinguish between plant responses to N-moles vs. N-molarity (Fig. 1a). Specifically, our 4-by-4 matrix of N-by-W treatments varied both the amount of N-moles (supplied as NH₄NO₃ and W-volume (modulated by soil drying). The highest amounts of N and W were chosen to promote plant growth, and the lowest amounts were chosen to limit plant growth. By varying N-moles and W-volume simultaneously, this matrix design allowed us to determine how rice plants respond to the total amount of N (N-moles), W-volume, or the concentration of N-molarity (N-moles/W-volume) (Fig. 1a, Supplementary Fig. 1).

To this end, 14-day-old rice seedlings initially grown on N- and W-replete media were transferred to each of the 16 treatment conditions within our N-by-W matrix (Fig. 1a). Plants were treated for 11 days, after which time we assayed leaf transcriptomes by RNA-Seq and measured plant phenotypes (Fig. 1b, Supplementary Fig. 2). We note that our phenotypic measurements of plant δ¹³C, a proxy for W-use, were proportional to the amount of external W provided (Supplementary Fig. 2). Similarly, we found total leaf N-content (a combination of assimilated ammonium and nitrate) was proportional to the amount of external N-moles provided (Supplementary Fig. 2). Together, these data indicated that internal leaf N- and W-status reflected their respective amounts in the external environment.

Phenotypically, we found that changes in shoot biomass in response to our N-by-W treatment combinations could be best modeled by a synergistic interaction of N-moles × W-volume (N×W) (Fig. 1b).

At the transcriptomic level, to interpret our gene expression data, we designed a multivariate linear model that could assess whether the expression of a given gene could be explained by a plant’s ability to respond to N-moles, W-volume, or their interaction, N-molarity (N/W). Based on the synergistic effects of N-moles and W-volume on biomass (Fig. 1b), we also included the N×W interaction term in our linear model (2), as detailed in Methods section. The benefit of using a linear model to interrogate our transcriptomic data was that it allowed us to detect gene expression responses that were directly proportional to changes in N- and/or W-dose.

We fit all expressed genes within the rice genome with this full linear model using DESeq²⁵, and through subsequent steps of model simplification, each gene could be binned into 1 of 14 simplified forms of the equation (Fig. 1c). We found that the expression of 65% of regulated genes (7197 genes) could best be explained by a single term—either N-moles, W-volume, N-molarity (N/W), or the synergistic interaction (N×W) (Fig. 1d, Supplementary Fig. 3). These four simple models of gene expression revealed that plants can respond to N-moles and W-volume independently from N-molarity (N/W). Moreover, they uncovered a synergistic response to N-moles and W-volume (N×W). The genome-wide evidence for how W impacts N-dose responses, and the role it plays in rice agriculture, is detailed below.

Rice can respond independently to N-moles and W-volume. Our modeling of genome-wide responses revealed that rice plants can respond independently to the amounts of N-moles or W-volume available within the treatment media. We found that 1331 rice genes responded exclusively to changes in N-moles in a dose-dependent manner, independently of W-volume (Fig. 2a, Supplementary Data 1). This class of genes displayed changes in expression that were best explained by changes in N-moles, and contained known N-responsive genes involved in N-uptake and assimilation such as the ammonium transporter OsAMT111 and glutamate synthase (GOGAT)16. It also contained genes not previously known to be N-regulated, such as phytochrome PHYB, a light sensor and signal transducer17 (Fig. 2a). This N-mole responsive gene class was over-represented in N-relevant gene ontology (GO) terms such as N-compound transport, and N-assimilation (Supplementary Data 2). Additionally, the vast majority of genes within this class (93%) significantly correlated with changes in leaf N-content (Fig. 1d).

Separately, we identified 1087 rice genes that specifically responded to changes in W-volume in a dose-dependent manner, independently of N-moles (Fig. 2b, Supplementary Data 1). This set included genes implicated in drought responses, including rice orthologs of Arabidopsis genes involved in abscisic acid signaling (ABF2)18 and biosynthesis (AAOs 1-4)19, as well as OsWRKY80, a member of the WRKY transcription factor family involved in W responses20 (Fig. 2b). Furthermore, genes within this W-volume response class significantly correlated with changes in leaf W-use efficiency measurements (δ¹³C) (Fig. 1d).

Rice responds to interactions between N-moles and W-volume. Our genome-wide models of transcriptome data also uncovered sets of genes whose dose-dependent response to N-moles was impacted by W-volume. The expression of these genes could be modeled by an interaction between N-moles and W-volume—N-molarity (N/W) or synergistically (N×W)—as described below.

Our analysis uncovered 2805 genes whose expression patterns were best explained by a change in N-molarity (N/W) (Fig. 2c, Supplementary Data 1). This set of genes was significantly enriched in N-related GO terms including cellular N-compound biosynthetic processes (Supplementary Data 2). Among these genes were the N-assimilation genes aspartate aminotransferase21 and glutamate decarboxylase (Fig. 2c).
Thus, our models of genome-wide responses to N and W revealed that plants can respond to dose changes in N-moles and W-volume can distinguish rice responses to N-moles, W-volume, and N-molarity. A synergistic effect between N-moles and W-volume, modeled by the N×W interaction term, best explains changes in shoot biomass (linear model, \( p = 1.3 \times 10^{-5} \)).

Through model simplification, 14 linear models uncover genome-wide responses to N-moles, W-volume, N/W (molarity), N×W, and their combinations (linear model, adj. \( p < 0.005 \)). Expression heatmap of genes fitted by a single model term, and the proportion of genes within each class that significantly correlated with biomass, leaf N-content, and leaf W-use efficiency (Pearson correlation, adj. \( p < 0.05 \)).

By comparing the normalized coefficients of genes fit by N, N/W, and N + N/W linear models, we found that genes were regulated exclusively in response to changes in N-moles (1331 genes) or N-molarity (2805 genes), while only 238 genes responded to a combination of changes in both N-moles and N-molarity. Source data for d are provided in Source Data file.

**Fig. 1** A factorial design uncovers plant responses to nitrogen and water dose combinations. **a** A 4-by-4 factorial matrix that varies both N-moles and W-volume can distinguish rice responses to N-moles, W-volume, and N-molarity. **b** A synergistic effect between N-moles and W-volume, modeled by the N×W interaction term, best explains changes in shoot biomass (linear model, \( p = 1.3 \times 10^{-5} \)). **c** Through model simplification, 14 linear models uncover genome-wide responses to N-moles, W-volume, N/W (molarity), N×W, and their combinations (linear model, adj. \( p < 0.005 \)). **d** Expression heatmap of genes fitted by a single model term, and the proportion of genes within each class that significantly correlated with biomass, leaf N-content, and leaf W-use efficiency (Pearson correlation, adj. \( p < 0.05 \)). **e** By comparing the normalized coefficients of genes fit by N, N/W, and N + N/W linear models, we found that genes were regulated exclusively in response to changes in N-moles (1331 genes) or N-molarity (2805 genes), while only 238 genes responded to a combination of changes in both N-moles and N-molarity. Source data for d are provided in Source Data file.

**Fig. 1** A factorial design uncovers plant responses to nitrogen and water dose combinations. **a** A 4-by-4 factorial matrix that varies both N-moles and W-volume can distinguish rice responses to N-moles, W-volume, and N-molarity. **b** A synergistic effect between N-moles and W-volume, modeled by the N×W interaction term, best explains changes in shoot biomass (linear model, \( p = 1.3 \times 10^{-5} \)). **c** Through model simplification, 14 linear models uncover genome-wide responses to N-moles, W-volume, N/W (molarity), N×W, and their combinations (linear model, adj. \( p < 0.005 \)). **d** Expression heatmap of genes fitted by a single model term, and the proportion of genes within each class that significantly correlated with biomass, leaf N-content, and leaf W-use efficiency (Pearson correlation, adj. \( p < 0.05 \)). **e** By comparing the normalized coefficients of genes fit by N, N/W, and N + N/W linear models, we found that genes were regulated exclusively in response to changes in N-moles (1331 genes) or N-molarity (2805 genes), while only 238 genes responded to a combination of changes in both N-moles and N-molarity. Source data for d are provided in Source Data file.

Thus, our models of genome-wide responses to N and W revealed that plants can respond to dose changes in N-moles (Fig. 2a), W-volume (Fig. 2b), or N-molarity (Fig. 2c). Moreover, we found that genome-wide responses to changes in N-moles (1331 genes) vs. N-molarity (2805 genes) were largely dichotomous; only 238 genes were sensitive to both (Fig. 1e).

Importantly, we found that W could also impact genome-wide N-dose responses in a synergistic mode. Specifically, we uncovered 1974 genes whose expression patterns could best be explained by a synergistic response to N-moles and W-volume (N-moles × W-volume) (Fig. 2d, Supplementary Data 1). Fifty-four percent of these genes significantly correlated with changes in biomass of rice seedlings—the highest proportion of any gene class (Fig. 1d). Uregulated genes within this N×W synergistic class were enriched in GO terms related to growth and energy production, such as photosynthesis and ATP biosynthetic processes, and members included seven ribosomal subunits and the rice ortholog of the *Arabidopsis thaliana* photosystem I P-subunit (Fig. 2d, Supplementary Data 2).

Thus, synergistic gene responses to changes in N-moles and W-volume may be a means by which plants signal growth responses when absolute amounts of both N and W are optimal. Such gene expression responses that are dependent on a multiplicative, nonlinear interaction between N and W can ensure that linear changes in both N and W amounts have non-additive outcomes on gene expression levels and plant phenotypes.
due to changes in nutrient dose, rather than being the product of plant stress or long-term changes in plant physiology, we investigated how rapidly these gene sets respond to changes in N-dose under non-stress, W-replete conditions. To do this, we performed a time-course experiment, treating hydroponically grown rice seedlings for 2 h with an N-treatment (provided as 20 mM K15NO3 + 20 mM 15NH415NO3). Over 10 time points, we monitored transcriptomic responses, as well as 15N-uptake into root and shoot tissue. We detected differentially expressed genes using a cubic spline model, and then binned genes into each time point they were found upregulated or downregulated, defined as when their fold change passed a threshold of ±1.25. In both root and shoot tissues, the number of differentially expressed genes found in each time point significantly correlated with the level of 15N incorporation, supporting the notion that such N-dose responsive genes were the result of N-influx from the environment (Fig. 3a, b, Supplementary Fig. 4, Supplementary Data 3).

Our time-course study revealed that all three classes of N-dose responsive genes—N-moles, N/W, and N×W—respond rapidly (within 5–30 min) to a change in N-dose under W-replete conditions. This was revealed by overlapping genes found to be regulated within each time point of the N-treatment with each of
Nonlinear N-by-W gene responses correlate with field traits. As the distinct modes by which W informs N-dose responses were discovered in rice seedlings, we next asked whether these gene responses play a role in rice agriculture.

To answer this, we grew 19 different rice cultivars in the field at the International Rice Research Institute (IRRI) in the Philippines. Each rice cultivar was grown in a 2-by-2 factorial matrix that varied N-fertilizer dose and W-volume. Crops were N-fertilized at a high dose of 150 kg/ha, or not fertilized. Under each N-dose condition, crops were grown either under W-replete vs.-deplete conditions, creating well-irrigated or drought conditions (Fig. 4a, Supplementary Fig. 5). For the 228 samples generated, we assessed vegetative and yield phenotypes, and took vegetative leaf samples for RNA-seq analysis.

In agreement with our laboratory observations on rice seedlings (Fig. 1b), we found that a synergistic interaction between N-fertilizer amount and W treatment (N×W) could best explain rice vegetative biomass in mature plants grown under field conditions (Fig. 4a).

To investigate whether the same N-by-W gene expression patterns we identified under laboratory conditions (Figs. 1 and 2) were present under field conditions, we assessed whether genome-wide expression patterns in rice field samples could be explained by N-fertilizer amount, W treatment, or the interaction between the two (where genotype was treated as a covariate within the model). By these means, we found that the genome-wide expression patterns across 19 rice cultivars tested in the field were consistent with the four classes of nutrient dose responses we discovered in seedlings under laboratory conditions—N-moles, W-volume, N-molarity (N/W), or synergistic responses (N×W) (Fig. 4b, Supplementary Data 4). Moreover, we found that the overlaps of lab-field gene sets for each nutrient dose responsive gene class were significantly higher than expected by chance (Fig. 4b).

Next, we investigated whether expression levels of dose responsive genes that were directionally conserved (i.e. induced or repressed) in each gene class across both laboratory and field settings were associated with cultivar agronomic performance. To do this, for each gene class set—N, W, N/W, and N×W—we calculated the first principal component, which represented the expression trends of all gene members in a single profile or ‘eigengene’ 22. Each resulting eigengene thus represented the set of lab-field validated genes responding either to N-moles (56 genes), W-volume (176 genes), N-molarity (N/W) (53 genes), or the synergistic interaction between N and W dose (N×W) (174 genes)—where each eigengene accounted for 31, 23, 32, and 34% of the proportion of variance in gene expression, respectively. We then assessed whether the expression profile for each of the eigengenes representing the four gene classes correlated with field phenotypes. The significance of this association was calculated by comparison to a null distribution of 10,000 eigengenes, generated from randomly selected genes expressed in the field.

This eigengene analysis revealed that the expression of genes regulated non-linearly in response to changes in combinations of N- and W-doses—N/W and N×W—were significantly associated with traits important to crop production (Fig. 4c, Supplementary Fig. 6). Genes responding synergistically to changes in N- and W-doses (N×W) or to N-molarity (N/W) across the lab-field divide were significantly correlated with complex traits such as grain yield, straw biomass, plant height, number of panicles, and number of tillers (Fig. 4c). Moreover, they were associated with N- and W-dose related traits such as chlorophyll concentration, stomatal conductance, and W-use efficiency (Fig. 4c).
The genes responding to combinations of N- and/or W-dose in rice seedlings found under laboratory conditions overlap significantly with reciprocal classes found in field-grown plants (*Monte Carlo, p < 0.05). Normalized expression patterns of lab-field-validated genes are displayed in heatmap. c Eigengenes derived from each gene set were correlated with crop traits. Significant Pearson R values are shown in red (permutation test, p < 0.05). d Example from c. Changes in N/W eigengene expression across 228 field samples is associated with grain yield (gray: −N, −W, green: +N, −W, blue: −N, +W, black: +N, +W conditions). e N/W eigengene expression is predictive within an independent field test set. Source data for b–e are provided in Source Data file.

Field phenotypes

Field-grown plants (*Monte Carlo, p < 0.05). Normalized expression patterns of lab-field-validated genes are displayed in heatmap. c Eigengenes derived from each gene set were correlated with crop traits. Significant Pearson R values are shown in red (permutation test, p < 0.05). d Example from c. Changes in N/W eigengene expression across 228 field samples is associated with grain yield (gray: −N, −W, green: +N, −W, blue: −N, +W, black: +N, +W conditions). e N/W eigengene expression is predictive within an independent field test set. Source data for b–e are provided in Source Data file.

**Discussion**

Our linear modeling of genome-wide expression patterns provide insight into how W impacts nutrient dose responses in a biological system. Typically, nutrient response studies vary nutrient dose within a fixed volume of W. However, this design changes both nutrient amount and concentration, making it impossible to determine whether nutrient responses are governed by the absolute nutrient amount, or its concentration in W. By systematically varying both N- and W-dose, we were able to distinguish between plant responses to N amount (moles) and N-concentration (molarity). Furthermore, we found that genes that respond nonlinearly to N- and W-dose combinations have relevance to crop production.

The design of our factorial N-by-W matrix enabled us to distinguish, measure, and model plant responses to changes in N-dose as a function of W (Fig. 1a). Our linear models of transcriptome and phenotype data could explain plant responses to...
changes in N-dose according to one of three distinct modes. First, plants can respond to changes in N-moles, independently of W-volume. In addition, plants can respond to changes in N-dose as a function of W-volume in one of two ways, as N-molarity (N/W), or synergistically (N×W).

Importantly, the ability of plants to respond to N-dose as a function of W has practical relevance to crop production. Indeed, the genome-wide responses to N-by-W interactions (N/W or N×W) that we observed in rice seedlings grown under controlled laboratory conditions correlated with agricultural traits in mature plants measured in the field over two seasons. Specifically, genes regulated in response to changes in both N and W (N/W or N×W) correlated with field traits such as yield, straw biomass, height, panicle number, tiller number, chlorophyll concentration, stomatal conductance, and W-use efficiency (Fig. 4c). As such, genes that exhibit N-by-W dose responses are associated with the non-additive effects irrigation and N-fertilizer have on crop phenotype.

A list of the genes responding to N-by-W dose interactions (N/W or N×W) that are associated with crop outcomes are reported in Supplementary Data 5. These N-by-W responsive gene sets could be useful as expression biomarkers to select rice varieties adaptive to marginal soils that are both N and W poor.24,23 Because these genome-wide responses detected at the seedling stage were associated with crop outcomes in mature field-grown rice, they may potentially be used to accelerate molecular breeding studies. Additionally, these N-by-W responsive genes could potentially serve as a biological readout for how N availability changes with W status within the rhizosphere, since measuring the effects of W on N flux within soils through existing chemical assays is technically challenging.

We note that these N-by-W effects that we measured on above ground crop traits are highly dependent on roots, which govern both nutrient and W uptake from soil.24 Indeed, decreasing W content in soil not only increases N-molarity but also leads to changes in root physiology, which alters the plant’s ability to acquire nutrients. Root physiological responses to W deficit that may negatively affect nutrient uptake include development of apoplastic barriers that prevent root W loss,24,25 and the development of root aerenchyma.26,27 However, some root physiologically responses to W deficit may also improve nutrient uptake, such as increased root foraging.28,29 Thus, future studies should investigate how root physiology informs the combinatorial responses to N and W we report here. However, at the gene expression level, we note that plants with very different root systems (i.e. seedlings vs. mature field-grown plants) share the same N-by-W dose response genes in shoots (Fig. 4b). Additionally, our data suggest that some of the N-by-W response genes detected in shoots also occur in roots (Fig. 3b, Supplementary Fig. 2a).

For our field experiments, 19 rice cultivars (listed in Supplementary Fig. 5) were grown under field conditions at the IRRI at Los Banos, Philippines (July–December 2016). Each cultivar was supplied with either a N-replete dose of 150 kg/ha dose of (NH4)2SO4 or with no N-treatment, 23 days after sowing (DAS). Under each N condition, fields were either W-replete or W-deplete, the latter obtained by maximizing evaporation and minimizing W supply to the field (from rainfall and W-deplete conditions at the IRRI in Los Banos were required to sustain growth). For N-deplete conditions, the minimum soil water potential was −34 kPa (non-N-fertilized) and −52 kPa (N-fertilized) at 74 DAS (as measured by tensiometers at 30 cm depth). For each N and W condition, rice cultivars were grown in triplicate in randomized block design, where each triplicate contained 20 plants.

For each of the 19 rice cultivars, leaf transcriptomes at 49 DAS were sampled at the vegetative stage from 2 individual rice plants in biological triplicate per condition, 2 h after dawn. Leaf tissue was stored in RNA later solution (Thermo Fisher Scientific) immediately upon sampling. Rice plants were sampled for dry weight 49 DAS per field treatment per genotype (2 plants per biological triplicate). From vegetative rice samples, additional traits were measured as follows: W-use efficiency was measured from leaf tissue using δ13C isotopic discrimination by mass spectrometry (performed by IRRI Analytical Services Lab, pool of 3 plants per biological replicate). Tiller number was counted by hand from each vegetative sample. Chlorophyll concentration index was measured 55 DAS (CCM-200 Chlorophyll concentration meter, Apogee Instruments, 2 plants per biological triplicate). Stomatal conductance was measured on two leaves per plot, averaged over two separate days, 44–48 DAS (Delta T Devices, Delta T Devices). End-point phenotyping was measured as follows: Straw biomass was measured as total straw dry weight (g) from a plot divided by the sampling area (m2). Panicle number was recorded by hand from 6 plants per genotype per treatment. Days to flowering was counted as the length of time, in days from sowing, until half the plants in each replicate plot had visible panicle emergence. Plant height was measured from the base of the plant to the tip of longest leaf. Grain yield was measured as the aggregate grain amount per cultivar in each triplicate using formula (1): grain yield = (average number of seeds × (100 – moisture content)/86)) × sampling area
following year (July–December 2017). In this season, for W-deplete conditions, the minimum soil water potential was −27 kPa (low N) and −59 kPa (high N) at 73 DAS. We observed grain yield, vegetative biomass, final (straw) biomass, chlorophyll concentration, stomatal conductance, days to flowering, plant height, tiller number, and panicle number using the same techniques as the previous year. We then correlated these second season outcomes with that of the first (Supplementary Fig. 7). Additionally, at 49 DAS we collected leaf tissue samples for RNA-seq analysis. The total number of RNA-seq libraries generated per cultivation genotype using the Illumina HiSeq 2500 v4 with 1 × 50 bp single-end chemistry or the Illumina HiSeq 2500 v4 with 1 × 75 bp single read-end chemistry was 616. RNA quality was assessed using Agilent Tape station using High Sensitivity RNA ScreenTape. A unit of 500–1 µg of total RNA per sample was depleted of rRNA by Thermo Fisher Scientific mRNA Purification Kit. All 411 RNA-seq libraries were made using the NEBNext Ultra RNA Library Prep Kit, and sequenced using Next Seq Illumina platform with 1 × 75 bp single-read-end chemistry or the Illumina HiSeq 2500 v4 with 1 × 50 bp single-end read chemistry. Two libraries were excluded from the 4 by 4 factorial N-by-W treatment matrix, and one library was excluded from second season field testing described based on quality controls. Raw reads were trimmed to remove adapter and low-quality bases (q < 10). RNA-seq reads were aligned to the Nipponbare reference genome38 using the Bowtie program5. Read counts were calculated using HT-seq39.

**Linear modeling and analysis of genome-wide data.** RNA from all leaf tissue samples, collected under both laboratory and field conditions, was purified using the RNeasy Mini Kit (Qiagen) with on-column DNase treatment. RNA quality was assessed using Agilent Tape station using High Sensitivity RNA ScreenTape. A unit of 500–1 µg of total RNA per sample was depleted of rRNA by Thermo Fisher Scientific mRNA Purification Kit. All 411 RNA-seq libraries were made using the NEBNext Ultra RNA Library Prep Kit, and sequenced using Next Seq Illumina platform with 1 × 75 bp single-read-end chemistry or the Illumina HiSeq 2500 v4 with 1 × 50 bp single-end read chemistry. Two libraries were excluded from the 4 by 4 factorial N-by-W treatment matrix, and one library was excluded from second season field testing described based on quality controls. Raw reads were trimmed to remove adapter and low-quality bases (q < 10). RNA-seq reads were aligned to the Nipponbare reference genome38 using the Bowtie program5. Read counts were calculated using HT-seq39.

From libraries generated from the 4 by 4 factorial N-by-W experiment we removed from our analysis genes with total read counts <128 (summed across all conditions) or a median read count < 0 (across all conditions). Multivariate gene modeling on read counts for each of the remaining genes was performed in R, using DESeq215, starting with the full generalized linear model (2):

$$
\text{gene expression} = \beta_0 + \beta_1N + \beta_2W + \beta_3NW + \beta_4N/W + \beta_5N + W
$$

Then, the full logistic model was fit to the RNA-seq read counts of each gene (using design ~ N + W + N/W + NW). We performed model simplification as follows: (1) for each gene, a false discovery rate (FDR)-adjusted p-value was computed for each of the factors within the model. (2) If a gene was fit significantly by all four terms (FDR-adjusted p < 0.005 for all four factors), then this gene was deemed fit by the full model and removed from remaining steps. (3) For all remaining genes, the factor with the least significance (highest FDR-adjusted p) was removed, and the model was refit using the remaining three terms. (4) If a gene was fit significantly by all three terms (FDR-adjusted p > 0.005 for all three factors), then this gene was deemed fit by a three-term model and removed from remaining model simplification steps. (5) Steps 3 and 4 were repeated, fitting two-term and one-term models. If a gene was not fit by any model form, then it was removed from further analysis. Normalized expression of each gene was correlated with log, shoot biomass, N-content, and W-use efficiency (Pearson correlation, FDR-adjusted p < 0.05) (Fig. 1d). The number of genes that correlated with shoot biomass were 7, 469, 218, and 1071 for classes N, N/W, and N/W, respectively. The number of genes that correlated with N-content in class were 1244, 8, 809, and 1427 for classes N, N/W, and N/W, respectively. The number of genes that correlated with W-use efficiency were 0, 215, 0, and 215 for classes N, N/W, and N/W, respectively. The number of genes that correlated with N amount were 1323, 0, 1425, and 1250 for classes N, N/W, and N/W, respectively. The number of genes that correlated with W were 0, 1071, 879, and 893 for classes N, N/W, and N/W, respectively. For heatmap visualization of lab-field validated gene sets, the expression value for each gene per cultivar in the field was first normalized to between 0 and 1, where 1 represented the maximum expression value (Fig. 4b).

We subsetted genes within each of the four classes—N, N/W, and N/W—to include only those genes that were consistently regulated across laboratory and field conditions (i.e. induced or repressed in both experiments). We then performed the first principal components of each of the lab-field validated gene classes using the field gene expression data, logged to the base 216. Each resulting eigengene was then correlated with log, values of field phenotypes using Pearson correlation. Since direction of principal component eigenvectors are arbitrary with respect to sign, all correlation values are reported as positive correlations, regardless of the direction of the slope of correlation in plots. A p-value of the association between the eigengene and phenotype was calculated by comparison to a null distribution. This p-value was then compared with a negative log of a p-value derived using a Monte Carlo test with 10,000 permutations with an FDR correction. As background set for this analysis, we used the intersect of genes found expressed in our time-course analysis and within our 4 × 4 N-by-W factorial experiment. To detect differentially expressed gene sets using field samples using a cubic spline model (df = 5, 3 knots)42. Genes were deemed differentially expressed if the expression differences between N and KC treatments over time were significantly different (FDR-adjusted p < 0.05), and their combined read count across all time points was >100 reads. We then binned genes into the time points in which they were either activated or repressed. This was determined by assessing whether a gene’s average fold change expression at a given time point was >2.15 when compared to the zero time point. We note that this approach allows a gene to be binned into multiple time points. We intersected these lists with genes modeled differentially expressed between N, N/W, and N/W—in our N-by-W factorial experiment in rice seedlings. The significance of each overlap was calculated using a Monte Carlo test with 10,000 permutations with an FDR correction. As background set for this analysis, we used the intersect of genes found expressed in our time-course analysis and within our 4 × 4 N-by-W factorial experiment. The resulting field validated gene sets were used as the background set for this analysis and were compared with all other gene sets using a Monte Carlo test with 10,000 permutations with an FDR correction. As background set for this analysis, we used the intersect of genes found expressed in our time-course analysis and within our 4 × 4 N-by-W factorial experiment. The resulting field validated gene sets were used as the background set for this analysis and were compared with all other gene sets using a Monte Carlo test with 10,000 permutations with an FDR correction.

### Reporting summary

Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

Data supporting the findings of this work are available within the paper and its Supplementary Information files. A reporting summary for this article is available as a Supplementary Information file. R codes that support the findings of this study and the data sets generated and analyzed during the current study are available from the corresponding author upon reasonable request. Sequencing data can be found at the National Center for Biotechnology Information Sequence Read Archive, with Accession number PRJNA592205. The source code for Figs 1d, 2, 3, and 4b-e, and Supplementary Figures 1A and 2–4 are provided as a Source Data file.

Received: 23 April 2018 Accepted: 26 February 2019

Published online: 26 March 2019
References

1. Godfray, H. C. J. et al. Food Security: the challenge of feeding 9 billion people. Nature 467, 817–851 (2010).
2. Gibbs, H. K. & Salzman, J. M. Mapping the world’s degraded lands. Appl. Geogr. 57, 12–21 (2015).
3. Di Paolo, E. & Rinaldi, M. Yield response of corn to irrigation and nitrogen fertilization in a Mediterranean environment. Field Crops Res. 105, 202–210 (2008).
4. Shi, J. C., Yasuo, H., Yermiyahu, U., Zuo, Q. & Ben-Gal, A. Dynamic responses of wheat to drought and nitrogen stresses during re-wetting cycles. Agric. Water Manage. 146, 163–172 (2014).
5. Patterson, T. B., Guy, R. D. & Dang, Q. L. Whole-plant nitrogen- and water-relationships traits, and their associated trade-offs, in adjacent muskeg and upland boreal spruce species. Oecologia 116, 160–168 (1997).
6. Seck, P. A., Diagne, A., Mohanty, S. & Wopereis, M. C. S. Crops that feed the world 7: Rice. Food Secur. 4, 7–24 (2012).
7. Wang, R., Okamoto, M., Xing, X. & Crawford, N. M. Microarray analysis of the nitrate response in Arabidopsis roots and shoots reveals over 1,000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism. Plant Physiol. 132, 556–567 (2003).
8. Gutierrez, R. A. et al. Systems approach identifies an organic nitrogen-responsive gene network that is regulated by the master clock gene CCA1. Proc. Natl Acad. Sci. USA 105, 4939–4944 (2008).
9. Ristova, D. et al. Combinatorial interaction network of transcriptomic and proteomic responses to nitrogen and hormones in the Arabidopsis thaliana root. Plant Signal. Behav. 9, r1310.11216/scisignal.aal2768 (2016).
10. Wilkins, O., Brautigam, K. & Campbell, M. M. Time of day shapes Arabidopsis drought transcriptomes. Plant J. 63, 715–727 (2010).
11. Xuan, W., Beeckman, T. & Xu, G. H. Plant nitrogen nutrition: sensing and signaling. Curr. Opin. Plant Biol. 39, 57–65 (2017).
12. Osakabe, Y., Osakabe, K., Shinozaki, K. & Tran, L. S. P. Response of plants to drought transcriptomes. Plant Physiol. 132, 556–567 (2003).
13. Horige, F. et al. Cryptochrome and phytochrome cooperatively but independently reduce active gibberellin content in rice seedlings under light irradiation. Plant Cell Physiol. 53, 1570–1582 (2012).
14. Fujita, Y., Fujita, M., Shinozaki, K. & Yamaguchi-Shinozaki, K. ABA-mediated transcriptional regulation in response to osmotic stress in plants. J. Plant Res. 124, 509–525 (2011).
15. Seo, M. et al. Abscisic acid lyase oxidase in leaves of Arabidopsis thaliana. Plant J. 23, 481–488 (2000).
16. Berri, S. et al. Characterization of WRKY co-regulatory networks in rice and Arabidopsis. BMC Plant Biol. 9, 120. 10.1186/1471-2229-9-120 (2009).
17. Zhou, Y. et al. Over-expression of aspartate aminotransferase genes in rice improves the photosynthetic efficiency of wheat genotypes. Funct. Plant Biol. 11, 539–552 (1984).
18. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. 10.1186/s13059-014-0550-8 (2014).
19. Lu, Y. E., Luo, F., Yang, M., Li, X. H. & Lian, X. M. Suppression of glutamate synthase gene significantly affects carbon and nitrogen metabolism in rice (Oryza sativa L.). Sci. China Life Sci. 54, 651–663 (2011).
20. Hirose, F. et al. Cryptochrome and phytochrome cooperatively but independently reduce active gibberellin content in rice seedlings under light irradiation. Plant Cell Physiol. 53, 1570–1582 (2012).
21. Fujita, Y., Fujita, M., Shinozaki, K. & Yamaguchi-Shinozaki, K. ABA-mediated transcriptional regulation in response to osmotic stress in plants. J. Plant Res. 124, 509–525 (2011).
22. Langfelder, P. & Horvath, S. Eigene networks for studying the relationships between co-expression modules. BMC Syst. Biol. 1, 54. 10.1186/1755-0569-1-54 (2007).
23. Yang, X. F. S. et al. Gene expression biomarkers provide sensitive indicators of in planta nitrogen status in maize. Plant Physiol. 157, 1841–1852 (2011).
24. Steudle, E. Water uptake by roots: effects of water deficit. J. Exp. Bot. 51, 1531–1542 (2000).
25. Baxter, L. et al. Root suberin forms an extracellular barrier that affects water relations and mineral nutrition in Arabidopsis. Plas Genet. 5, e1000492 (2009).
26. Fan, M., Zhu, Jiming, Richards, Christina, Brown, KathleenM. & Lynch, JonathanP. Physiological roles for aerenchyma in phosphorus-stressed roots. Funct. Plant Biol. 5, 493–506 (2003).
27. Yang, X. et al. Drought-induced root aerenchyma formation restricts water uptake in rice seedlings supplied with nitrate. Plant Cell Physiol. 53, 495–504 (2012).
28. Lynch, J. P. Steep, cheap and deep: an ideotype to optimize water and N acquisition by maize root systems. Ann. Bot. 112, 347–357 (2013).
29. Rogato, A., D’Apuzzo, E. & Chiurazzi, M. The multiple plant response to high ammonium conditions: the Lotus japonicus AMT1;3 protein acts as a putative transceptor. Plant Signal. Behav. 5, 1594–1596 (2010).
30. Ho, C. H., Lin, S. H., Hu, H. C. & Tsay, Y. F. CHL1 functions as a nitrate sensor in plants. Cell 138, 1184–1194 (2009).
31. Chantarpanpong, L., Wolfson, R. L. & Sabatini, D. M. Nutrient-sensing mechanisms across evolution. Cell 161, 67–83 (2015).
32. Dobrenel, T. et al. TOR signaling and nutrient sensing. Annu. Rev. Plant Biol. 67, 261–285 (2016).
33. Ruffel, S. et al. Nitrogen economics of root foraging: transitive closure of the nitrate-cytokinin relay and distinct systemic signaling for N supply vs. demand. Proc. Natl Acad. Sci. USA 108, 18524–18529 (2011).
34. Leran, S. et al. Nitrate sensing and uptake in Arabidopsis are enhanced by AB12, a phosphate inactivated by the stress hormone abscisic acid. Sci. Signal. 8, ra43. 10.1126/scisignal.aaa4829 (2015).
35. Castaings, L. et al. The nodule perception-like protein 7 modulates nitrate sensing and metabolism in Arabidopsis. Plant J. 57, 426–435 (2009).
36. Yoshida S., Douglas A. F., James, H. C. & Gomez K. A. A Manual for Physiological Studies of Rice (International Rice Research Institute, Manila, 1976).
37. Murashige, T. & Skoog, F. A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol. Plant. 15, 473–497 (1962).
38. Kawahara, Y. et al. Improvement of the Oryza sativa Nipponbare reference genome using next generation sequence and optical map data. Rice 6, 4. 10.1186/1939-8433-6-4 (2013).
39. Anders, S., Pyl, P. T. & Huber, W. HTSeq-a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169 (2015).
40. Li, L., Stoeckert, C. J. Jr & Roos, D. S. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. 13, 2178–2189 (2003).
41. Katari, M. S. et al. VirtualPlant: a software platform to support systems biology research. Plant Physiol. 152, 500–510 (2010).
42. Varala, K. et al. Temporal transcriptional logic of dynamic regulatory networks underling nitrate signaling and use in plants. Proc. Natl Acad. Sci. USA 115, 6494–6499 (2018).

Acknowledgements

We thank staff at IRRI for their assistance in the field studies. This work was supported by a grant on Plant Genomics to G.M.C. from the Zegar Family Foundation (A16-0051), the Beachell-Borlaug International Scholarship to J.S., and from grants to G.M.C. from the National Science Foundation (NSF grants: NSF-DBI-0445666 and NSF-IOS-1840761).

Author contributions

The experimental plan was designed by J.S. and G.M.C. Laboratory experiments were performed by J.S and M.A.; genomic analysis was performed by J.S., D.T. and G.M.C.; the manuscript was written by J.S. and G.M.C.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-019-09287-7.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Journal peer review information: Nature Communications thanks Peter Langfelder and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.