

Title:

Cross-ecosystem transcriptomics identifies distinct genetic modules for nutrient acquisition in maize

Authors:

Yusaku Sugimura^a, Ai Kawahara^b, Hayato Maruyama^a, and Tatsuhiro Ezawa^{a*}

Affiliations:

^a Graduate School of Agriculture, Hokkaido University, Sapporo 060-8589, Japan; ^b Health & Crop Sciences Research Laboratory, Sumitomo Chemical, Co. Ltd., Takarazuka 665-8555, Japan.

***Corresponding author:**

Tatsuhiro Ezawa

Email: tatsu@res.agr.hokudai.ac.jp

ORCID ID:

Yusaku Sugimura: 0000-0001-7271-7703

Hayato Maruyama: 0000-0003-1959-6417

Tatsuhiro Ezawa: 0000-0001-6188-7635

Keywords

Arbuscular mycorrhiza, field transcriptomics, gene coexpression network, maize, nutrient acquisition.

Author Contributions

H.M., A.K., and T.E. conceived the project; A.K. and T.E. designed the field surveys; Y.S., H.M., A.K., and T.E. collected samples; Y.S., H.M., and A.K. extracted RNA and analyzed plant and soil samples under the supervision of T.E.; Y.S. performed quality assessment and mapping of the sequenced reads; Y.S. and T.E. analyzed the data and interpreted the results; Y.S. and T.E. wrote the manuscript with input from all authors; and all authors discussed the manuscript.

This PDF file includes:

Main Text

Figures 1 to 5

Tables 2

Abstract

Plants have evolved diverse strategies for the acquisition of the macro-nutrients phosphorus and nitrogen; e.g., mycorrhizal formation, root development, and secretion of chelators/hydrolases to liberate inorganic phosphate. Despite the extensive studies on the individual strategies, there is little information about how plants regulate these strategies in response to fluctuating environment. We approached this issue via profiling transcriptomes of plants grown in large environmental gradients. Roots, leaves, and root-zone soils of 251 maize plants were collected across the US Corn Belt and Japan. RNA was extracted from the roots and sequenced, and the leaves and soils were analyzed. Nineteen genetic modules were defined by weighted gene coexpression network analysis and functionally characterized according to gene ontology analysis, by which we found three modules that are directly involved in nutrient acquisition: mycorrhizal formation, phosphate-starvation response (PSR), and root development. Correlation analysis with soil and plant factors revealed that both phosphorus and nitrogen deficiencies upregulated the mycorrhizal module, whereas the PSR module was upregulated mainly by deficiency in phosphorus relative to nitrogen. Expression levels of the root development module were negatively correlated with those of the mycorrhizal module, suggesting that nutrient acquisition through the two pathways, mycorrhizas and roots, are opposite strategies that are employed under nutrient-deficient and -enriched conditions, respectively. The identification of the soil and plant factors that drive the modules has implications for sustainable agriculture; activation/optimization of the strategies is feasible via manipulating the factors. Overall, our study opens a new window for understanding plant response to complex environments.

Introduction

Intensification of global food production to meet the requirements of a growing population with minimizing environmental impacts, that is, sustainable intensification of agriculture (1, 2) is a major challenge in our society. Conventional agriculture that heavily relies on chemical fertilizer produced from non-renewable resources, e.g., phosphate rock (3), has contributed to feeding today's population of the world, but posed threats to ecosystem structure and function, e.g., deterioration of air and water quality (4, 5). In addition, nutrient enrichment also has a negative impact on soil function. Urea application reduces nitrogen (N) mineralization from organic matter through switching N source of microorganisms from organic matter to mineral N (6). High-input of chemical fertilizer generally reduces the abundance and diversity of mycorrhizal fungi (7). Various approaches for reducing chemical fertilizer input have been proposed: the enhancement of nutrient-acquisition efficiency in crops by selection and genetic manipulation (8, 9) and acceleration of nutrient cycling through managing cropping systems (10) and mycorrhizal associations (11).

Plants have developed diverse strategies for the acquisition of the macronutrients N and phosphorus (P) during evolution. N availability has a strong impact on root development. Moderate N deficiency promotes root development in *Arabidopsis* (12). N (nitrate and ammonium)-enriched patches induce localized proliferation of lateral roots for efficient capture of the nutrients (13), which is triggered by nitrate and ammonium uptake via plasma membrane NITRATE TRANSPORTER 1 (14, 15) and AMMONIUM TRANSPORTER 1;3 (16), respectively. Lateral root formation is, however, restricted under severe N deficiency, which may be an important trait for survival in resource-limited environments (17). Localized proliferation of lateral roots was also observed in patches of inorganic phosphate (Pi) (13). Pi deficiency generally inhibits primary root growth, but increases lateral root growth and density, leading to a shallow root system (12, 18). Root hairs also significantly contribute to Pi uptake via extending the surface area for Pi uptake (19). Pi-starvation responses (PSR) have been studied extensively in plants, which include the upregulation of high-affinity Pi transporter genes of the PHT1 family, secretion of non-specific acid phosphatase for mineralization of organic P, excretion of organic acids for the chelation of the counter cations (e.g., Al^{3+} , Ca^{2+} , and Fe^{3+}) of Pi, and replacement of phospholipids with sulfolipids and galactolipids for accelerating internal Pi recycling (20).

Association with arbuscular mycorrhizal (AM) fungi is a widespread strategy in land plants; the fungi facilitate an extensive surface area for the acquisition of mineral nutrients, in particular P and N (21). Deficiencies in P and N trigger secretion of the plant hormone strigolactones into the rhizosphere (22), which facilitates contact of fungal hyphae with the roots via stimulating hyphal branching (23). After the physical contact, the fungal hyphae grow into the cortex and form the highly-branched hyphal termini "arbuscules" where nutrient exchange

between the symbionts occurs; briefly, Pi, nitrate, and ammonium are taken up from the soil by extraradical hyphae, delivered to the arbuscules, and released into the arbuscular interface (reviewed in 24) from which the plant cell takes up the nutrients via the mycorrhiza-specific transporters for Pi (25, 26), nitrate (27), and ammonium (28). In return, the host supplies lipid (29-32) and sugar (33) as carbon source for the fungi via the complex of half-size ABC transporters STR/STR2, a putative lipid exporter (34), and the sugar transporter SWEET (35), respectively.

Despite the extensive studies on the individual strategies for nutrient acquisition, there is little information about how plants modulate resource allocation to these strategies in response to biotic and abiotic factors, e.g., nutrient availability, soil physical and biological properties, water availability, and agricultural practice. Optimization of these plant-intrinsic strategies for nutrient acquisition would enhance soil nutrient cycling, which could be one solution for the reduction of fertilizer input. To this end, understanding of the regulatory mechanism of the strategies under field conditions is necessary. We approached this issue via analyzing transcriptomes of plants grown in large environmental gradients, that is, cross-ecosystem transcriptomics. Recently, transcriptomics has been applied to field plants for dissecting their responses to fluctuating environments (36-39). For example, two sorghum genotypes with different drought tolerance showed distinct transcriptomic responses to drought stress, including the rapid upregulation of photosynthetic genes during drought recovery in a tolerant genotype and the downregulation of AM-associated genes during drought stress in both genotypes (36). Even though these studies were conducted only in one field site, their approach opened a new window for understanding plant responses to environmental changes and agricultural practice, which could not be simulated in growth chamber and greenhouse.

Maize (*Zea mays* L.) is one of the most important cereals worldwide, serving as staple food, livestock feed, and industrial raw material (40). We collected maize samples across the US Corn Belt where the world's highest productivity is achieved by the typical high-input agricultural system, as well as across Japan where agricultural practice is regionally varied due to diversity in climate and soil properties. These large environmental gradients among the sampling sites enabled us to define distinct genetic modules for nutrient acquisition and to identify soil/plant factors that drive the expression of the modules and their interplays.

Results and Discussion

Field sites and RNA sequencing.

We collected root, leaf, and root-zone soil samples from 66 plants grown in five field sites across the US Corn Belt in 2017 and from 185 plants grown in seven field sites across Japan in 2018, including 17 experimental plots in five out of the seven Japanese sites (Figs. 1A and B, Dataset S1, and SI Appendix, Table S1). The sites and plots were characterized by a principal component

(PC) biplot with the climatic/soil factors from which one of the soil/plant factors that were highly correlated with each other ($|r| \geq 0.9$) were excluded (Fig. 1C and Dataset S2). The US sites are localized in the first quadrant, whereas the Japanese sites are localized in the other quadrants, between which annual precipitation and soil properties were largely different. Among the Japanese sites, the contents of soil organic matter, base, and clay were highly variable. Notably, there were large gradients of N) and P availability among the sites: $\text{NO}_3\text{-N}$, 10.5 – 212 mg kg^{-1} ; $\text{NH}_4\text{-N}$, 1.0 – 96 mg kg^{-1} ; Bray II-P, 5.3 – 494 mg P kg^{-1} , which led us to the expectation that there would also be large gradients in plant responses to nutrient deficiency/excess.

We extracted RNA from all root samples, performed 75-base single end sequencing on a Illumina platform, and mapped the sequence reads to the maize genome (41), by which 8.4 ± 0.25 million reads per sample in average were mapped to exons (Dataset S3). The raw read counts were normalized to Transcripts Per kilobase Million (TPM), and only genes of which the average TPM was 5 or more were subjected to subsequent analysis.

Definition of gene-coexpression modules.

In preliminary analysis, we focused on the genes involved in AM formation because a number of plant genes that facilitate AM fungal colonization, in particular those involved in arbuscule development/functioning, are highly conserved in the plant kingdom (42), implying that orthologues of the genes in non-model plants (e.g. maize) could simply be identified through homology searches and phylogenetic analysis. We first identified maize orthologues of *STR2* (putative lipid exporter) and *RAM2* that encodes glycerol-3-phosphate acyltransferase, a key enzyme of the AM-specific lipid biosynthetic pathway (32), and then analyzed their expression patterns, together with the mycorrhiza-inducible Pi transporter *Pht1;6* of maize (43, 44). Their expression levels were strongly correlated across the samples (SI Appendix, Fig. S1), which indicated that the expression of these genes are co-expressed and tightly regulated across ecosystems, irrespective of maize variety, and further, led us to the idea that plants have developed a gene regulatory unit specialized for mycorrhizal formation.

Based on this idea, we conducted weighted gene-coexpression network analysis on the transcriptome data of 251 samples using the R package with a soft-threshold power of 14 that provided a scale-free topology of the network ($r^2 = 0.91$) (SI Appendix, Fig. S2), resulted in definition of 19 coexpression modules. The 19 modules (named with colors) were functionally categorized according to the gene ontology (GO) enrichment analysis (Table 1 and Dataset S4): black, cell division; blue, transcription/translation; green, cell cycle regulation; greenyellow, stress-associated protein degradation; grey, immune response/N assimilation; grey60, antioxidation; lightcyan, branched-chain amino acid (BCAA) metabolism; lightgreen, water uptake and diurnal rhythm; magenta, immune response; midnightblue, lipid biosynthesis; pink, root development; royalblue, response to hypoxia; salmon, P-starvation response (PSR); tan, trehalose

biosynthesis; yellow, mycorrhizal formation. No or only a few GO terms were enriched in cyan, darkgreen, darkred, and purple modules. The expression levels of these modules in individual samples are represented by eigengenes that are PC1 (sample) scores calculated based on the expression data of the module member genes (Dataset S5) (45). Pairwise correlation analysis of the eigengenes revealed that the grey, grey60, and lightcyan modules were positively correlated with each other (Fig. 2). In addition, these three modules were correlated positively with the same soil factors Bray II-P, NO₃-N, Mg, and sand % (Dataset S6), suggesting that the three modules play general roles in nutrient assimilation/metabolism and immune responses. In contrast, the darkred module showed no distinct correlations with other modules nor with the soil/plant factors, indicating that the gene in the modules were expressed constitutively and unresponsive to environmental changes. Accordingly, this module is likely to have a "house-keeping" function. In subsequent analyses, we focused on the yellow (mycorrhizal formation), salmon (PSR), and pink (root development) modules that are likely to be involved directly in the nutrient acquisition strategies.

Yellow module.

A majority of the genes known to participate in arbuscule development and functioning are enriched in this module; e.g., the genes encoding the transcription factors RAM1 (46), RAD1 (47), and WRI5 (48), Pi transporter Pht1;6, nitrate transporter NPF4.5 (27), ammonium transporter AMT3;1 (28), H⁺-ATPase HA1 (49, 50), key enzymes for the mycorrhiza-specific lipid biosynthesis RAM2 and FatM (32), putative lipid exporters STR and STR2, and enzymes involved in strigolactone/mycorradicin biosynthesis DXS2 (51, 52), PSY2 (53), CCD7, CCD8, D27, and CCD1 (54) (Dataset S7).

To characterize this module in relation to AM fungal colonization, we analyzed unmapped sequence reads (i.e. the reads that were not mapped to the maize genome) that might contain AM fungal RNA reads. Generally, sequence reads of RNA-Seq (i.e. mRNA-Seq) contain a small number of those originated from ribosomal RNA (rRNA), which led us to the idea that AM fungal rRNA reads could be used as a fungal biomass index in the roots. We conducted Blastn searches against the large subunit rRNA sequences of maize and the AM fungal operational taxonomic units (OTUs) (55) using the unmapped reads as queries, and then fungal read counts for the OTUs were combined within each sample and normalized on the basis of plant rRNA read count (Dataset S8). It was considered, however, that fungal rRNAs that have A-rich regions were preferentially sequenced in mRNA-Seq because polyA-tailed RNA was purified prior to library construction, which might bias fungal rRNA abundance across the samples. To evaluate the significance of this bias, we sequenced rRNA directly (rRNA-Seq), which would provide unbiased rRNA read counts. Twenty RNA extracts were randomly chosen from the 251 samples and sequenced without purifying polyA-tailed RNA, and Blastn searches were performed using the

rRNA reference sequences (Dataset S9). Correlation analysis showed that relative abundance of AM fungal rRNA read (normalized by plant rRNA reads) was highly correlated between the data obtained by the mRNA-Seq and rRNA-Seq ($r = 0.919$, $P < 0.001$) (SI Appendix, Fig. S3), suggesting that the bias is minimum. The unmapped reads were also *de novo* assembled, and the resulting 427,813 contigs (average length, 461 bp; N50, 525 bp) were subjected to Blastx searches against the 11 AM fungal (gromeromycotinan) genomes/transcripts (56-60) and 6 non-gromeromycotinan fungal genomes (SI Appendix, Table S3), as well as against the maize genome and Genbank bacterial genome database (SI Appendix, Fig. S4), to identify the AM fungal genes involved in Pi delivery, that is, those encoding putative Pi exporter SYG1-1, polyphosphate polymerase VTC4, vacuolar Pi exporter PHO91, and aquaglyceroporin AQP3 (24). We identified 86, 74, 159, and 105 contigs that showed similarity to the AM fungal SYG1-1, VTC4, PHO91, and AQP3, respectively, mapped the reads to these contigs, combined read counts within each gene, and normalized on the basis of TPM of the plant (Dataset S10). These expression data were subjected to multiple linear regression analysis using the yellow module eigengenes as an objective variable, together with the AM fungal rRNA read counts. Among the genes, the read counts of rRNA, SYG1-1, and PHO91 were the significant explanatory variables for the eigengenes (SI Appendix, Table S2), suggesting that the eigengenes reflect not only fungal biomass but also the functionality of fungal colonization.

For further functional categorization of the 1,023 genes of this module, the k-means clustering analysis was performed using correlation distance as a measure. Clustering the genes into five submodules successfully separated into different functional groups along the PC2 axis of the yellow module (Fig. 3A and B and Table 2). In the submodule 1, a majority of the essential components of arbuscule development and nutrient exchange were enriched, e.g., *RAM1*, *RAD1*, *WRI5*, *Pht1;6*, *NPF4.5*, *AMT3;1*, *HA1*, *RAM2*, *FatM*, *STR/STR2*, serine/threonine receptor-like kinase ARK1 (61), and the genes involved in membrane trafficking *EXO70I* (62) and *Vapyrin A* (63) (Fig. 3C and Dataset S7). The enrichment of these genes suggests that the submodule 1 play a central role in the yellow module via modulating nutrient exchange across the periarbuscular membrane. In the submodule 2, the GO terms fatty acid biosynthetic process and plastid were overrepresented (Dataset S11), reflected in the enrichment of genes encoding the enzymes involved in the *de novo* biosynthesis of C16:0-fatty acid via acetyl-CoA in plastids. This suggests that the submodule 2 has a role in the biosynthesis of the essential component of the lipid for the construction of periarbuscular membrane and/or for the export to the fungi. In the submodule 3, the GO terms isoprenoid biosynthetic process, carotenoid metabolic process, and hormone metabolic process were overrepresented; the genes involved in the biosynthesis of carotenoid derivatives *DXS2*, *PSY2*, *CCD7*, *CCD8*, and *D27* were enriched. In addition, the GRAS transcription factor NSP2 that regulates strigolactone biosynthesis via modulating *D27* expression (64) also belongs to the submodule. Therefore, it is likely that the submodule 3

regulates the early processes of fungal accommodation through strigolactone production. In the submodule 4 the GO terms senescence-associated vacuole, amino sugar metabolic process, and extracellular space were overrepresented. The enrichment of the genes encoding a MYB1-like transcription factor, chitinase, and cysteine protease suggests that this submodule is responsible for arbuscule degeneration/turnover (65). Neither GO terms nor known genes involved in fungal accommodation/functioning were enriched in the submodule 5, and therefore, the biological function of this submodule could not be defined.

Connectivities of the module member genes (i.e. correlation coefficients between their expression levels and the eigengenes) were calculated to assess the significance of the genes in the network (Datasets S7). In general, genes with a higher connectivity are defined as those that play a more significant role in the module, that is, "hub genes" (45). Average connectivities of the submodules 1 – 5 with the yellow eigengenes were 0.88, 0.72, 0.77, 0.75, and -0.59, respectively. In addition, most genes in the submodule 1 showed positive PC2 scores, while most genes in the submodules 2 and 5 showed negative PC2 scores (Fig. 3B), suggesting that these submodules are differentially regulated.

Taken together, the yellow (mycorrhizal) module consists of the five submodules that have different functions; the submodule 1 plays a central role in arbuscule development and nutrient exchange under the control of RAM1, RAD1, and WRI5, the submodule 2 is responsible for fatty acid biosynthesis for the construction of periarbuscular membrane and/or for carbon supply to the fungi, the submodule 3 is involved in carotenoid biosynthesis/metabolism, probably for the synthesis of strigolactones and mycorradicin, under the control of NSP2, and the submodule 4 is responsible for arbuscule degeneration/turnover under the control of MYB1-like transcription factor. Importantly, although the genes in this module are generally coregulated, the expression of submodules is finely tuned in response to environmental changes, which is analyzed in a later section.

Salmon module.

This module consists of 101 genes, in which the GO terms involved in PSR, e.g., cellular response to Pi starvation, Pi transport, acid phosphatase activity, and galactolipid biosynthetic process were overrepresented (Dataset S5). Pht1;3, one of the four Pi transporters in this module, is likely to be responsible for the direct pathway (66, 67) in which the transporter is localized to root epidermis and mediates Pi uptake independently of the mycorrhizal pathway (68). Six genes encoding SPX (SYG1-PHO1-XPR1) domain that is a sensor for the inositol polyphosphate-mediated signaling pathway for the maintenance of Pi homeostasis in eukaryotes (69) were also enriched in this module; one encodes the Pi exporter PHO1 (70) and the other five encode SPX-domain containing protein. Many acid phosphatase genes, including those encoding purple acid phosphatase (PAP), were assigned to this module. Two orthologues of *NIGT1* that

encodes a MYB-type transcription factor belong to this module. NIGT1 was originally identified as a repressor of nitrate uptake (71), but also has a role in PSR via modulating the expression of SPX genes (72-74). It has been well documented that the transcription factor PHR1 plays a central role in PSR as a master regulator in Arabidopsis, a non-mycorrhizal plant (75). In maize two orthologues of PHR1 have so far been found (76), but both were assigned to the blue module in this study (Dataset S7), suggesting that, in the presence of mycorrhiza, PHR1 may not directly be involved in the regulation of PSR. In higher plants, including maize, Pi deficiency induces the replacement of membrane phospholipid with non-phosphorous lipids such as galactolipid and sulfolipid (77). The genes encoding glycerophosphodiester phosphodiesterase (GDPD2) for the degradation of phospholipid, monogalactosyldiacylglycerol (MGDG) synthase for the biosynthesis of galactolipid, and sulfoquinovosyl transferase (SQD2) for the biosynthesis of sulfolipid were assigned to this module. Higher connectivities with the eigengenes were observed for the genes encoding the enzymes involved in phospholipid replacement, MGDG synthase, SQD2, and GDPD2, SPX-domain containing protein, and NIGT1 (Datasets S7).

Pink module.

This module consists of 471 genes, and the GO terms involved in cell wall synthesis and root development, e.g., cell wall biogenesis, cellulose synthase, cytoskeleton, root system development, root morphogenesis, root epidermal cell differentiation, were overrepresented (Dataset S5), indicating that this module is responsible for the development of root system that accompanies transcriptional activation of a set of genes for cell wall biogenesis. Plant cell wall is composed of the primary cell wall that supports the fundamental growth of the cell and the secondary cell wall that supports the primary cell wall mechanically and facilitates water transport. CesA genes encode a cellulose synthase catalytic subunit and form cellulose synthase complexes consist of 18 – 24 CesA proteins; CesA1, CesA2, CesA3, CesA5, CesA6, and CesA9 are involved in the primary cell wall synthesis, while CesA4, CesA7, and CesA8 are for the secondary cell wall synthesis in Arabidopsis (78-80). In the maize genome twenty CesAs have been found (81), among which eleven genes, *CesA1*, *CesA2*, *CesA4*, *CesA6*, *CesA7a*, *CesA7b*, *CesA9*, *CesA10*, *CesA11*, *CesA12a*, and *CesA12b* are assigned to this module (Dataset S7). In addition to these CesAs, many genes encoding polysaccharide-biosynthetic/modification enzymes were found in this module; e.g., glycosyltransferases and glucanases. In this module there are 15 genes encoding the essential parts of cytoskeleton, alpha- and beta-tubulins and a microtubule motor protein. Eighteen transcription factors, e.g., an ethylene-responsive transcription factor (ERF), five NAC-domain transcription factors, and seven MYB-domain transcription factors, were also found. ERF035 was expressed in regions related to cell division/differentiation, e.g., lateral roots and central cylinder of primary roots, and upregulate *CesA1* in Arabidopsis (82). It has been known that many NAC- and MYB-domain transcription

factors are responsible for the regulation of cell wall synthesis and expressed during root formation (83-86). Connectivities between the eigengenes were higher in those encoding cellulose synthases, polysaccharide modification enzymes, and microtubule-related proteins (Dataset S7).

Module interplays.

For understanding interplays between the modules in relation to the soil and plant factors, multiple regression analysis was conducted, in addition to simple linear regression analysis. Basically, the eigengenes of yellow and salmon modules were not correlated (Figs. 2 and 4A). However, the expression levels of 65 out of the 101 salmon genes were positively correlated with the yellow eigengenes ($P < 0.05$), which include three out of the four Pi transporter genes and all eight acid phosphatase genes (Fig. 4B and Dataset S7). Furthermore, most genes in the submodules 1 (arbuscule development and nutrient exchange) and 3 (carotenoid biosynthesis and metabolism) of the yellow module were also positively correlated with the salmon eigengenes ($P < 0.05$). The yellow eigengenes were mainly explained by the plant factors leaf P content (with a negative coefficient), leaf P:N ratio (positive), stem diameter (positive), and growth rate (negative) and the soil factors NO₃-N (negative), clay % (negative), and Bray II-P (negative) in the multiple regression analysis (Fig. 4C). In addition to these factors, leaf N content was also negatively correlated with the eigengenes in the simple linear regression analysis (Dataset S6). The salmon eigengenes were largely explained by the plant factor growth rate (negative) and leaf P:N ratio (negative) and the soil factors organic matter % (positive), Bray II-P (negative), exchangeable K (positive) and Ca (negative), and silt % (positive) both in the multiple regression and simple linear regression analyses. These results suggest that, although the two modules are regulated independently, parts of the genes in these modules are coordinately expressed, particularly under P-deficient conditions. It is likely, however, that N deficiency is also an important driver for the yellow module, given that N deficiency leads to increases in leaf P:N ratio and decreases in leaf N content that are positive and negative factors, respectively, for the module.

Plant availability of Pi in soil is generally too low to meet the plant demand because most P is present as unavailable forms, such as sparingly soluble Pi and organic P (87). Exploitation of the unavailable P by the secretion of organic acids and phosphatase, that is, typical PSRs, and mycorrhizal formation are two major strategies for the acquisition of Pi, but their interplays have been largely overlooked. Although the salmon (PSR) and yellow (mycorrhizal) modules similarly respond to many plant and soil factors, their responses to leaf P:N ratio is contrasting; the factor is a negative regulator for the PSR module but a positive regulator for the mycorrhizal module. The upregulation of PSR by lower leaf P:N ratio (i.e. P deficiency relative to N) could be interpreted by the high-connectivity of *NIGT1* with the salmon eigengenes because *NIGT1* that is

involved in the signaling cascade of PSR (71, 74) is upregulated in a nitrate concentration-dependent manner (71). Furthermore, the fact that this regulatory cascade is conserved not only in the non-mycorrhizal plant *Arabidopsis* but also in the mycorrhizal plant maize strongly suggests that the PSR module has evolved independently from the mycorrhizal module. This independency of the PSR module would facilitate an alternative strategy to cope with Pi deficiency when Pi delivery through the mycorrhizal pathway does not meet the P demand due to, e.g., low-population density or absence of AM fungi and extremely low-Pi availability in the soil. On the other hand, the coexpression of the submodules 1 and 3 and some of the PSR genes suggests the presence of a common regulatory pathway for these genes that has to be elucidated.

Intriguingly, the eigengenes of yellow and pink modules were negatively correlated (Figs. 2 and 5A), which is reflected in negative correlations of the expression levels of the genes of the pink module with the yellow eigengenes and those of the genes of the yellow submodules 1-4 with the pink eigengenes (Fig. 5B). The pink eigengenes were explained by the plant factors leaf P content (with a positive coefficient) and leaf P:N ratio (negative) and the soil factor clay % (positive) in the multiple regression analysis (Fig. 5C). In addition to these factors, leaf N content were positive factors for the expression in the simple linear regression analysis (Dataset S6). These results indicate that the pink module responds to the plant and soil factors in the opposite direction to which the yellow module responds, suggesting that root development is intrinsically an opposite strategy of mycorrhizal formation.

In the Early Devonian, AM fungal symbiosis facilitated terrestrialization of early plants that did not have functional root systems by providing a water/nutrient uptake pathway (88), and during the Middle to Late Devonian, plants evolved a substantial root system that is capable of not only taking up water/nutrients but also anchoring the body to the soil (89). In the context of water/nutrient uptake, mycorrhizal formation and root development are functionally redundant. We demonstrated that the two strategies are oppositely regulated as a function of plant nutrient status, that is, plants allocate more resource to mycorrhizal formation under nutrient-deficient conditions, but under nutrient-enriched conditions, more resource to root development. This could be interpreted by a trade-off between cost for the extension of surface area for nutrient uptake and the rates (efficiency) of nutrient uptake and translocation. The mycorrhizal pathway is mediated by fungal hyphae that are finer than roots, providing a larger surface area per carbon investment than root system development. It is likely, therefore, that the mycorrhizal pathway is more efficient under nutrient-depleted conditions in which nutrient uptake and translocation by hyphae are more rapid than the diffusion rates of nutrients towards the roots (68). On the other hand, the root direct pathway may facilitate more rapid nutrient uptake and translocation under nutrient-enriched conditions in which the diffusion rates are rapid enough to sustain the uptake from the pathway. In addition, root development would play an important role in physical support for the above-ground part that grows vigorously under nutrient-enriched conditions. It has been

suggested that gramineous plants have evolved finer roots to obtain a larger surface area per unit carbon investment, that is, towards less dependency to mycorrhizas (90). Our findings suggest, however, that plants have also developed a regulatory mechanism for modulating resource allocation between mycorrhizas and a root system to optimize the efficiency of nutrient acquisition in a given environment.

Conclusion

Our study provides a new insight into the regulatory mechanism of plant nutrient-acquisition strategies in complex environments. We revealed the distinct genetic modules for nutrient acquisition strategies in maize; that is, mycorrhizal formation, PSR, and root development and, further, identified soil and plant factors that drive these modules. The expression levels of the mycorrhizal module were correlated not only with fungal biomass (i.e. rRNA abundance) but also with the expression levels of fungal genes involved in nutrient delivery, reflecting the functionality of mycorrhiza. The two distinct responses to P deficiency, the enhancement of P acquisition via the root direct pathway and the internal recycling of P, are coregulated at the transcriptional level in the PSR module. The root development module that consists of a large number of the genes involved in cell wall biogenesis is upregulated under nutrient-enriched conditions, suggesting that this strategy requires substantial carbon investment (e.g., 91).

In maize production, different genotypes are grown in different regions, and a number of new genotypes are distributed every year. In this study, a variety of genotype was collected across the countries and analyzed without taking into account genotypic differences. It has been well documented that different genotypes respond to biotic and abiotic factors differently; for instance, Pi acquisition efficiency via the mycorrhizal pathway is different among maize genotypes (43, 92). We consider, however, that the genetic modules identified in this study are conserved across genotypes of the species and that such differences in the responses are due to differential regulation of the expression of the modules among the genotypes. This issue could not be addressed in our current approach and has to be elucidated in future studies.

Our cross-ecosystem transcriptomics approach opens a new window for understanding how plants respond to complex environments, providing implications for sustainable management of agriculture. We identified the soil factors that drive the modules for nutrient acquisition, which suggests that resource allocation to the nutrient uptake pathways and PSRs could be optimized via manipulating these factors, e.g., via balancing fertilizer input and improving soil physical conditions. Towards sustainable agriculture, replacement of chemical fertilizer (i.e. non-renewable resource) with renewable resource, e.g., organic matter, is also necessary. In this study, we focused only on conventional agriculture in which soil fertility was managed mainly by chemical

fertilizer. Further studies are needed to clarify how to balance resource allocation to the strategies in the crop in organic agriculture.

Materials and Methods

The plant and soil samples were collected from 21- to 54-day-old plants from the USA in June 2017 and from 13- to 54-day-old plants from Japan in May – July 2018. Details of geographic, climatic, and field-management data of the sampling sites/plots are presented in SI Appendix, Table S1. Field sampling, soil and plant analysis, RNA sequencing, gene expression profiling, rRNA read assignment, *de novo* assembling of AM fungal mRNA reads, and statistics were performed as described in SI Appendix, Supplementary Information Text.

Data availability

The RNA-Seq reads have been deposited in National Center for Biotechnology Information under the accession number PRJNA586746 (mRNA-Seq, 251 samples) and PRJNA604657 (rRNA-seq, 20 samples). The expression data of the maize roots and AM fungal transcript references can be downloaded from: <http://lab.agr.hokudai.ac.jp/botagr/rhizo/RhizoCont/Download.html>.

Acknowledgments

The authors acknowledge A. Nagano for technical advice, S. Inman and S. Garcia for the arrangement of field experiments and sampling in the USA, T. Mitani, K. Tawaraya, T. Koyama, Y. Tahara, and T. Nagaoka for the arrangement of field experiments and sampling in Hokkaido University, Yamagata University, Utsunomiya University, Nagoya University, and Hiroshima University, respectively. This work was partially supported by ACCEL (JPMJAC1403) from Japan Science and Technology Agency (Y.S., H.M, and TE).

References

1. C. Andres, G. S. Bhullar, Sustainable intensification of tropical agro-ecosystems: Need and potentials. *Front. Environ. Sci.* **4** (2016).
2. T. Garnett *et al.*, Sustainable intensification in agriculture: Premises and policies. *Science* **341**, 33-34 (2013).
3. D. Cordell, S. White, Peak phosphorus: Clarifying the key issues of a vigorous debate about long-term phosphorus security. *Sustainability* **3**, 2027-2049 (2011).

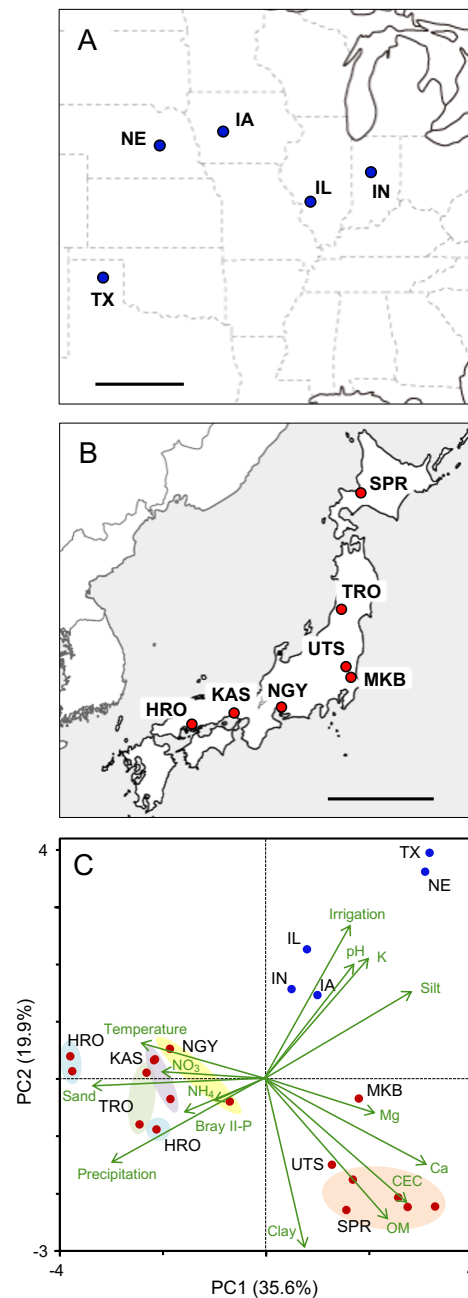
- 428 4. V. H. Smith, G. D. Tilman, J. C. Nekola, Eutrophication: impacts of excess nutrient inputs on
429 freshwater, marine, and terrestrial ecosystems. *Environ. Pollut.* **100**, 179-196 (1999).
- 430 5. C. P. Vance, Symbiotic nitrogen fixation and phosphorus acquisition. Plant nutrition in a world
431 of declining renewable resources. *Plant Physiol.* **127**, 390-397 (2001).
- 432 6. X. G. Li *et al.*, Nitrogen fertilization decreases the decomposition of soil organic matter and
433 plant residues in planted soils. *Soil Biol. Biochem.* **112**, 47-55 (2017).
- 434 7. E. Verbruggen, M. G. A. Van Der Heijden, J. T. Weedon, G. A. Kowalchuk, W. F. M. R  ling,
435 Community assembly, species richness and nestedness of arbuscular mycorrhizal fungi in
436 agricultural soils. *Mol. Ecol.* **21**, 2341-2353 (2012).
- 437 8. J. I. Schroeder *et al.*, Using membrane transporters to improve crops for sustainable food
438 production. *Nature* **497**, 60-66 (2013).
- 439 9. A. E. Richardson *et al.*, Plant and microbial strategies to improve the phosphorus efficiency of
440 agriculture. *Plant Soil* **349**, 121-156 (2011).
- 441 10. M. Hallama, C. Pekrun, H. Lambers, E. Kandeler, Hidden miners – the roles of cover crops
442 and soil microorganisms in phosphorus cycling through agroecosystems. *Plant Soil* **434**, 7-45
443 (2019).
- 444 11. M. C. Rillig *et al.*, Towards an integrated mycorrhizal technology: Harnessing mycorrhiza for
445 sustainable intensification in agriculture. *Front. Plant Sci.* **7** (2016).
- 446 12. B. D. Gruber, R. F. H. Giehl, S. Friedel, N. von Wir  n, Plasticity of the arabidopsis root system
447 under nutrient deficiencies. *Plant Physiol.* **163**, 161-179 (2013).
- 448 13. M. C. Drew, Comparison of the effects of a localised supply of phosphate, nitrate, ammonium
449 and potassium on the growth of the seminal root system, and the shoot, in barley. *New Phytol.*
450 **75**, 479-490 (1975).
- 451 14. C.-H. Ho, S.-H. Lin, H.-C. Hu, Y.-F. Tsay, CHL1 functions as a nitrate sensor in plants. *Cell*
452 **138**, 1184-1194 (2009).
- 453 15. T. Remans *et al.*, The Arabidopsis NRT1.1 transporter participates in the signaling pathway
454 triggering root colonization of nitrate-rich patches. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 19206-
455 19211 (2006).
- 456 16. J. E. Lima, S. Kojima, H. Takahashi, N. von Wir  n, Ammonium triggers lateral root branching
457 in *Arabidopsis* in an AMMONIUM TRANSPORTER1;3-dependent manner. *Plant Cell* **22**,
458 3621-3633 (2010).
- 459 17. T. Araya *et al.*, CLE-CLAVATA1 peptide-receptor signaling module regulates the expansion of
460 plant root systems in a nitrogen-dependent manner. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2029-
461 2034 (2014).
- 462 18. L. C. Williamson, S. P. Ribrioux, A. H. Fitter, H. M. Leyser, Phosphate availability regulates root
463 system architecture in Arabidopsis. *Plant Physiol.* **126**, 875-882 (2001).
- 464 19. T. S. Gahoonia, N. E. Nielsen, P. A. Joshi, A. Jahoor, A root hairless barley mutant for
465 elucidating genetic of root hairs and phosphorus uptake. *Plant Soil* **235**, 211-219 (2001).
- 466 20. W. C. Plaxton, H. T. Tran, Metabolic adaptations of phosphate-starved plants. *Plant Physiol.*
467 **156**, 1006-1015 (2011).
- 468 21. S. E. Smith, D. J. Read, *Mycorrhizal Symbiosis* (Academic Press, San Diego, CA, USA, ed.
469 3rd, 2008), pp. 787.
- 470 22. K. Yoneyama *et al.*, How do nitrogen and phosphorus deficiencies affect strigolactone
471 production and exudation? *Planta* **235**, 1197-1207 (2012).
- 472 23. K. Akiyama, K. Matsuzaki, H. Hayashi, Plant sesquiterpenes induce hyphal branching in
473 arbuscular mycorrhizal fungi. *Nature* **435**, 824-827 (2005).
- 474 24. T. Ezawa, K. Saito, How do arbuscular mycorrhizal fungi handle phosphate? New insight into
475 fine-tuning of phosphate metabolism. *New Phytol.* **220**, 1116-1121 (2018).
- 476 25. M. J. Harrison, G. R. Dewbre, J. Y. Liu, A phosphate transporter from *Medicago truncatula*
477 involved in the acquisition of phosphate released by arbuscular mycorrhizal fungi. *Plant Cell*
478 **14**, 2413-2429 (2002).
- 479 26. C. Rausch *et al.*, A phosphate transporter expressed in arbuscule-containing cells in potato.
480 *Nature* **414**, 462-466 (2001).
- 481 27. S. Wang *et al.*, Functional analysis of the OsNPF4.5 nitrate transporter reveals a conserved
482 mycorrhizal pathway of nitrogen acquisition in plants. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 16649-
483 16659 (2020).

- 484 28. S. Koegel *et al.*, The family of ammonium transporters (AMT) in *Sorghum bicolor*: two AMT
485 members are induced locally, but not systemically in roots colonized by arbuscular mycorrhizal
486 fungi. *New Phytol.* **198**, 853-865 (2013).
- 487 29. L. H. Luginbuehl *et al.*, Fatty acids in arbuscular mycorrhizal fungi are synthesized by the host
488 plant. *Science* **356**, 1175-1178 (2017).
- 489 30. A. Keymer *et al.*, Lipid transfer from plants to arbuscular mycorrhiza fungi. *eLife* **6**, e29107
490 (2017).
- 491 31. Y. Jiang *et al.*, Plants transfer lipids to sustain colonization by mutualistic mycorrhizal and
492 parasitic fungi. *Science* **356**, 1172-1175 (2017).
- 493 32. A. Bravo, M. Brands, V. Wewer, P. Dörmann, M. J. Harrison, Arbuscular mycorrhiza-specific
494 enzymes FatM and RAM2 fine-tune lipid biosynthesis to promote development of arbuscular
495 mycorrhiza. *New Phytol.* **214**, 1631-1645 (2017).
- 496 33. N. Helber *et al.*, A versatile monosaccharide transporter that operates in the arbuscular
497 mycorrhizal fungus *Glomus* sp. is crucial for the symbiotic relationship with plants. *Plant Cell*
498 **23**, 3812-3823 (2011).
- 499 34. Q. Zhang, L. A. Blaylock, M. J. Harrison, Two *Medicago truncatula* half-ABC transporters are
500 essential for arbuscule development in arbuscular mycorrhizal symbiosis. *Plant Cell* **22**, 1483-
501 1497 (2010).
- 502 35. J. An *et al.*, A *Medicago truncatula* SWEET transporter implicated in arbuscule maintenance
503 during arbuscular mycorrhizal symbiosis. *New Phytol.* **224**, 396-408 (2019).
- 504 36. N. Varoquaux *et al.*, Transcriptomic analysis of field-droughted sorghum from seedling to
505 maturity reveals biotic and metabolic responses. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 27124-
506 27132 (2019).
- 507 37. A. J. Nagano *et al.*, Annual transcriptome dynamics in natural environments reveals plant
508 seasonal adaptation. *Nat. Plants* **5**, 74-83 (2019).
- 509 38. P. Yu *et al.*, Root type and soil phosphate determine the taxonomic landscape of colonizing
510 fungi and the transcriptome of field-grown maize roots. *New Phytol.* **217**, 1240-1253 (2018).
- 511 39. Atsushi J. Nagano *et al.*, Deciphering and prediction of transcriptome dynamics under
512 fluctuating field conditions. *Cell* **151**, 1358-1369 (2012).
- 513 40. A. F. Troyer, Adaptedness and heterosis in corn and mule hybrids. *Crop Sci.* **46**, 528-543
514 (2006).
- 515 41. Y. Jiao *et al.*, Improved maize reference genome with single-molecule technologies. *Nature*
516 **546**, 524-527 (2017).
- 517 42. A. Bravo, T. York, N. Pumplun, L. A. Mueller, M. J. Harrison, Genes conserved for arbuscular
518 mycorrhizal symbiosis identified through phylogenomics. *Nat. Plants* **2**, 15208 (2016).
- 519 43. R. J. H. Sawers *et al.*, Phosphorus acquisition efficiency in arbuscular mycorrhizal maize is
520 correlated with the abundance of root-external hyphae and the accumulation of transcripts
521 encoding PHT1 phosphate transporters. *New Phytol.* **214**, 632-643 (2017).
- 522 44. M. Willmann *et al.*, Mycorrhizal phosphate uptake pathway in maize: Vital for growth and cob
523 development on nutrient poor agricultural and greenhouse soils. *Front. Plant Sci.* **4**, 533
524 (2013).
- 525 45. P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis.
526 *BMC Bioinformatics* **9**, 559 (2008).
- 527 46. E. Gobbato *et al.*, A GRAS-type transcription factor with a specific function in mycorrhizal
528 signaling. *Curr. Biol.* **22**, 2236-2241 (2012).
- 529 47. H.-J. Park, D. S. Floss, V. Levesque-Tremblay, A. Bravo, M. J. Harrison, Hyphal branching
530 during arbuscule development requires *Reduced Arbuscular Mycorrhiza1*. *Plant Physiol.* **169**,
531 2774-2788 (2015).
- 532 48. Y. Jiang *et al.*, *Medicago* AP2-domain transcription factor WRI5a is a master regulator of lipid
533 biosynthesis and transfer during mycorrhizal symbiosis. *Mol. Plant* **11**, 1344-1359 (2018).
- 534 49. E. Wang *et al.*, A H⁺-ATPase that energizes nutrient uptake during mycorrhizal symbioses in
535 rice and *Medicago truncatula*. *Plant Cell* **26**, 1818-1830 (2014).
- 536 50. F. Krajinski *et al.*, The H⁺-ATPase HA1 of *Medicago truncatula* is essential for phosphate
537 transport and plant growth during arbuscular mycorrhizal symbiosis. *Plant Cell* **26**, 1808-1817
538 (2014).
- 539 51. D. S. Floss, W. Schliemann, J. Schmidt, D. Strack, M. H. Walter, RNA interference-mediated

- repression of *MtCCD1* in mycorrhizal roots of *Medicago truncatula* causes accumulation of C27 apocarotenoids, shedding light on the functional role of CCD1. *Plant Physiol.* **148**, 1267-1282 (2008).
52. M. H. Walter, J. Hans, D. Strack, Two distantly related genes encoding 1-deoxy-d-xylulose 5-phosphate synthases: differential regulation in shoots and apocarotenoid-accumulating mycorrhizal roots. *Plant J.* **31**, 243-254 (2002).
53. R. Stauder *et al.*, Strigolactone levels in dicot roots are determined by an ancestral symbiosis-regulated clade of the PHYTOENE SYNTHASE gene family. *Front. Plant Sci.* **9**, 255-255 (2018).
54. S. Al-Babili, H. J. Bouwmeester, Strigolactones, a novel carotenoid-derived plant hormone. *Ann. Rev. Plant Biol.* **66**, 161-186 (2015).
55. R. Niwa *et al.*, Dissection of niche competition between introduced and indigenous arbuscular mycorrhizal fungi with respect to soybean yield responses. *Sci. Rep.* **8**, 7419 (2018).
56. X. Sun *et al.*, Genome and evolution of the arbuscular mycorrhizal fungus *Diversispora epigaea* (formerly *Glomus versiforme*) and its bacterial endosymbionts. *New Phytol.* **221**, 1556-1573 (2019).
57. T. Maeda *et al.*, Evidence of non-tandemly repeated rDNAs and their intragenomic heterogeneity in *Rhizophagus irregularis*. *Commun. Biol.* **1**, 87 (2018).
58. D. Beaudet *et al.*, Ultra-low input transcriptomics reveal the spore functional content and phylogenetic affiliations of poorly studied arbuscular mycorrhizal fungi. *DNA Res.* **25**, 217-227 (2017).
59. A. Salvioli *et al.*, Symbiosis with an endobacterium increases the fitness of a mycorrhizal fungus, raising its bioenergetic potential. *ISME J.* **10**, 130-144 (2016).
60. Y. Kikuchi *et al.*, Polyphosphate accumulation is driven by transcriptome alterations that lead to near-synchronous and near-equivalent uptake of inorganic cations in an arbuscular mycorrhizal fungus. *New Phytol.* **204**, 638-649 (2014).
61. R. Roth *et al.*, A rice Serine/Threonine receptor-like kinase regulates arbuscular mycorrhizal symbiosis at the peri-arbuscular membrane. *Nat. Commun.* **9**, 4677 (2018).
62. X. Zhang, N. Pumplin, S. Ivanov, M. J. Harrison, EXO70I is required for development of a sub-domain of the periarbuscular membrane during arbuscular mycorrhizal symbiosis. *Curr. Biol.* **25**, 2189-2195 (2015).
63. N. Pumplin *et al.*, *Medicago truncatula* Vapyrin is a novel protein required for arbuscular mycorrhizal symbiosis. *Plant J.* **61**, 482-494 (2010).
64. W. Liu *et al.*, Strigolactone biosynthesis in *Medicago truncatula* and rice requires the symbiotic GRAS-type transcription factors NSP1 and NSP2. *Plant Cell* **23**, 3853-3865 (2011).
65. D. S. Floss *et al.*, A transcriptional program for arbuscule degeneration during AM symbiosis is regulated by MYB1. *Curr. Biol.* **27**, 1206-1212 (2017).
66. R. Nagy *et al.*, Differential regulation of five Pht1 phosphate transporters from maize (*Zea mays* L.). *Plant Biol.* **8**, 186-197 (2006).
67. D. Glassop, S. Smith, F. Smith, Cereal phosphate transporters associated with the mycorrhizal pathway of phosphate uptake into roots. *Planta* **222**, 688-698 (2005).
68. S. E. Smith, I. Jakobsen, M. Grønlund, F. A. Smith, Roles of arbuscular mycorrhizas in plant phosphorus nutrition: Interactions between pathways of phosphorus uptake in arbuscular mycorrhizal roots have important implications for understanding and manipulating plant phosphorus acquisition. *Plant Physiol.* **156**, 1050-1057 (2011).
69. R. Wild *et al.*, Control of eukaryotic phosphate homeostasis by inositol polyphosphate sensor domains. *Science* **352**, 986-990 (2016).
70. S. Wege *et al.*, The EXS domain of PHO1 participates in the response of shoots to phosphate deficiency via a root-to-shoot signal. *Plant Physiol.* **170**, 385-400 (2016).
71. Y. Maeda *et al.*, A NIGT1-centred transcriptional cascade regulates nitrate signalling and incorporates phosphorus starvation signals in Arabidopsis. *Nat. Commun.* **9**, 1376 (2018).
72. B. Hu, C. Chu, Nitrogen-phosphorus interplay: old story with molecular tale. *New Phytol.* **225**, 1455-1460 (2020).
73. A. Medici *et al.*, Identification of molecular integrators shows that nitrogen actively controls the phosphate starvation response in plants. *Plant Cell* **31**, 1171-1184 (2019).
74. Y. Ueda, T. Kiba, S. Yanagisawa, Nitrate-inducible NIGT1 proteins modulate phosphate uptake

- and starvation signalling via transcriptional regulation of SPX genes. *Plant J.* **102** (2019).
75. V. Rubio *et al.*, A conserved MYB transcription factor involved in phosphate starvation signaling both in vascular plants and in unicellular algae. *Genes Dev.* **15**, 2122-2133 (2001).
76. C. Calderón-Vázquez, R. J. H. Sawers, L. Herrera-Estrella, Phosphate deprivation in maize: Genetics and genomics. *Plant Physiol.* **156**, 1067-1077 (2011).
77. H. Tjellström, M. X. Andersson, K. E. Larsson, A. S. Sandelius, Membrane phospholipids as a phosphate reserve: the dynamic nature of phospholipid-to-digalactosyl diacylglycerol exchange in higher plants. *Plant Cell Environ.* **31**, 1388-1398 (2008).
78. S. Persson *et al.*, Genetic evidence for three unique components in primary cell-wall cellulose synthase complexes in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 15566-15571 (2007).
79. T. Desprez *et al.*, Organization of cellulose synthase complexes involved in primary cell wall synthesis in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 15572-15577 (2007).
80. N. G. Taylor, R. M. Howells, A. K. Huttly, K. Vickers, S. R. Turner, Interactions among three distinct CesaA proteins essential for cellulose synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1450-1455 (2003).
81. B. W. Penning *et al.*, Genetic resources for maize cell wall biology. *Plant Physiol.* **151**, 1703-1728 (2009).
82. L. Saelim *et al.*, Arabidopsis Group III ERF proteins positively regulate primary cell wall-type CESA genes. *J. Plant Res.* **132**, 117-129 (2019).
83. J. Jung, S. McCouch, Getting to the roots of it: Genetic and hormonal control of root architecture. *Front. Plant Sci.* **4** (2013).
84. Y. Nakano, M. Yamaguchi, H. Endo, N. A. Rejab, M. Ohtani, NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. *Front. Plant Sci.* **6** (2015).
85. M. Yamaguchi, M. Kubo, H. Fukuda, T. Demura, VASCULAR-RELATED NAC-DOMAIN7 is involved in the differentiation of all types of xylem vessels in *Arabidopsis* roots and shoots. *Plant J.* **55**, 652-664 (2008).
86. V. Willemsen *et al.*, The NAC domain transcription factors FEZ and SOMBRERO control the orientation of cell division plane in *Arabidopsis* root stem cells. *Dev. Cell* **15**, 913-922 (2008).
87. J. Shen *et al.*, Phosphorus dynamics: from soil to plant. *Plant Physiol.* **156**, 997-1005 (2011).
88. C. P. Humphreys *et al.*, Mutualistic mycorrhiza-like symbiosis in the most ancient group of land plants. *Nat. Commun.* **1**, 103 (2010).
89. P. Kenrick, P. R. Crane, The origin and early evolution of plants on land. *Nature* **389**, 33-39 (1997).
90. Z. Ma *et al.*, Evolutionary history resolves global organization of root functional traits. *Nature* **555**, 94 (2018).
91. A. Mollier, S. Pellerin, Maize root system growth and development as influenced by phosphorus deficiency. *J. Exp. Bot.* **50**, 487-497 (1999).
92. Q. Chu *et al.*, Soil plant-available phosphorus levels and maize genotypes determine the phosphorus acquisition efficiency and contribution of mycorrhizal pathway. *Plant Soil* **449**, 357-371 (2020).

637 **Figures and Tables**



638

639

640 **Figure 1.** Location and characterization of the maize sampling sites. Sampling sites in the USA
641 (A) and Japan (B). (C) Principal component analysis of soil, plant, climatic, and agricultural
642 variables in 22 plots of the sampling sites in the USA and Japan. Different experimental plots
643 within a site are indicated with colored circles. Detailed information of the plots/sites is described
644 in SI Appendix, Table S1 and Dataset S1. IL, Illinois; IN, Indiana; IW, Iowa; NE, Nebraska; TX,
645 Texas; SPR, Sapporo; TRO, Tsuruoka; UTS, Utsunomiya; MKB, Makabe; NGY, Nagoya; KAS,
646 Kasai; HRO, Hiroshima.

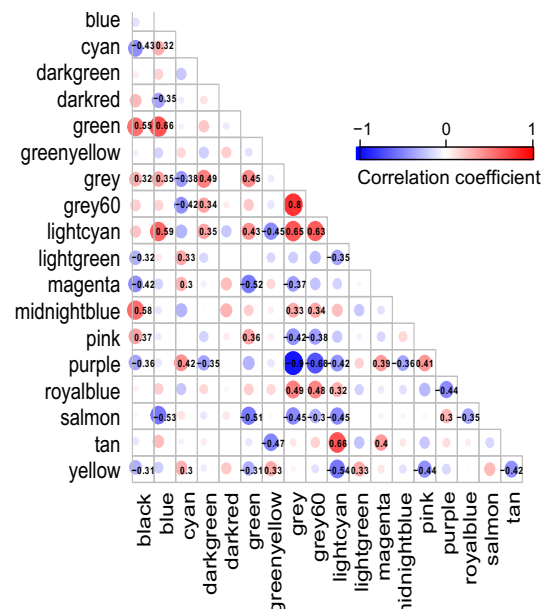


Figure 2. Correlation matrix of module eigengenes. Pairwise correlation analysis of the 19 coexpression modules defined by weighed gene coexpression network analysis was conducted, and correlation coefficients of which *P*-values are less than 0.001 are indicated. The blue and red circles indicate negative and positive correlations, respectively.

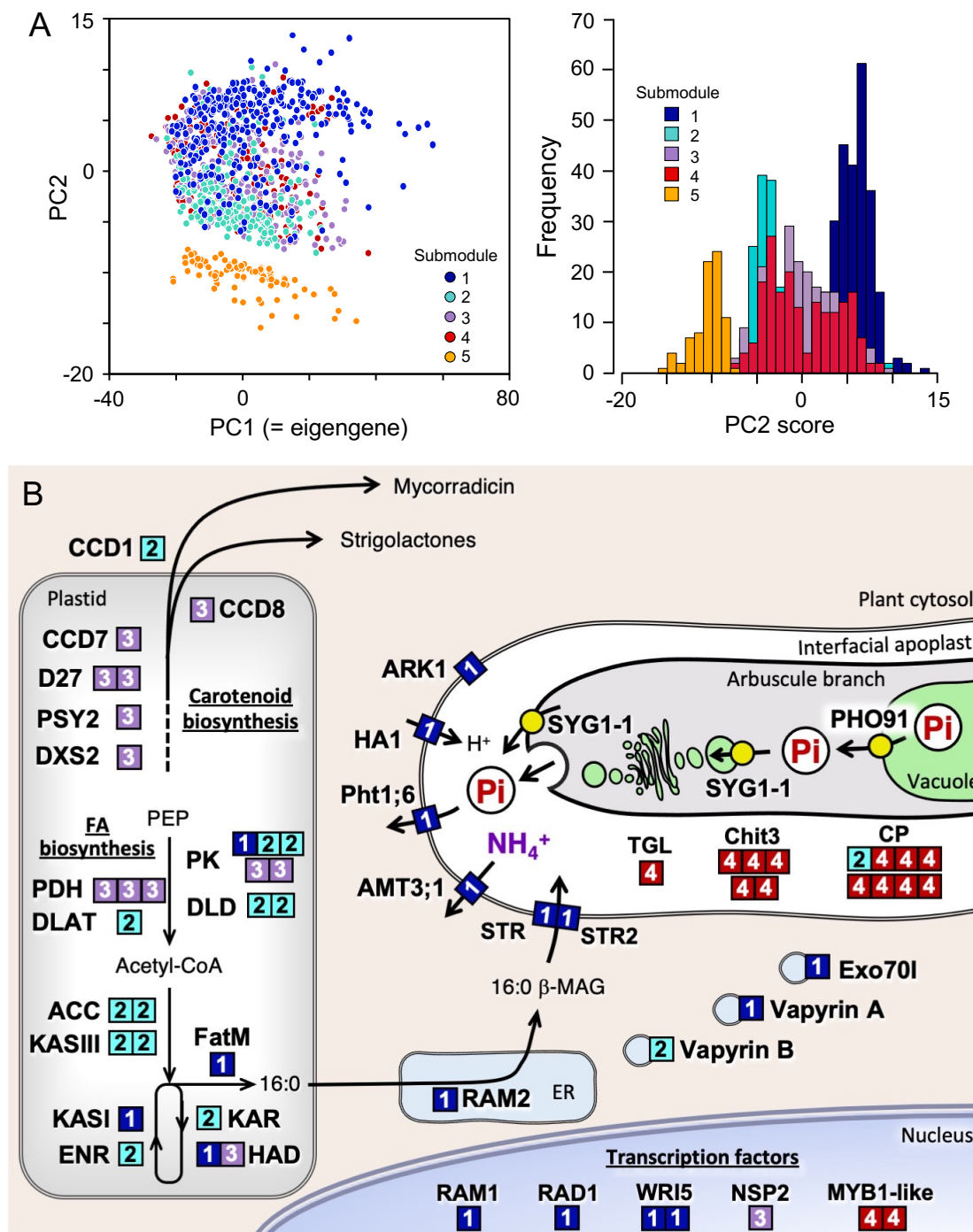


Figure 3. Functional categorization of the 1,023 genes of the yellow module into five submodules by k-means clustering analysis using correlation-based distance as a measure. (A) PCA plot of the genes in the submodules 1 – 5 (left), and frequency distribution of PC2 score of the genes (right). The submodule numbers the yellow module genes are indicated with the following colors: 1, dark blue; 2, turquoise; 3, purple; 4, red; 5, orange. (B) Putative function and cellular localization of the genes of which orthologues have been functionally characterized in previous

664 studies. The numbers of submodule to which the genes belong are indicated in the boxes with the
665 same colors in A. No orthologues of the genes in the submodule 5 has so far been functionally
666 characterized, and thus these genes are excluded from this scheme.
667

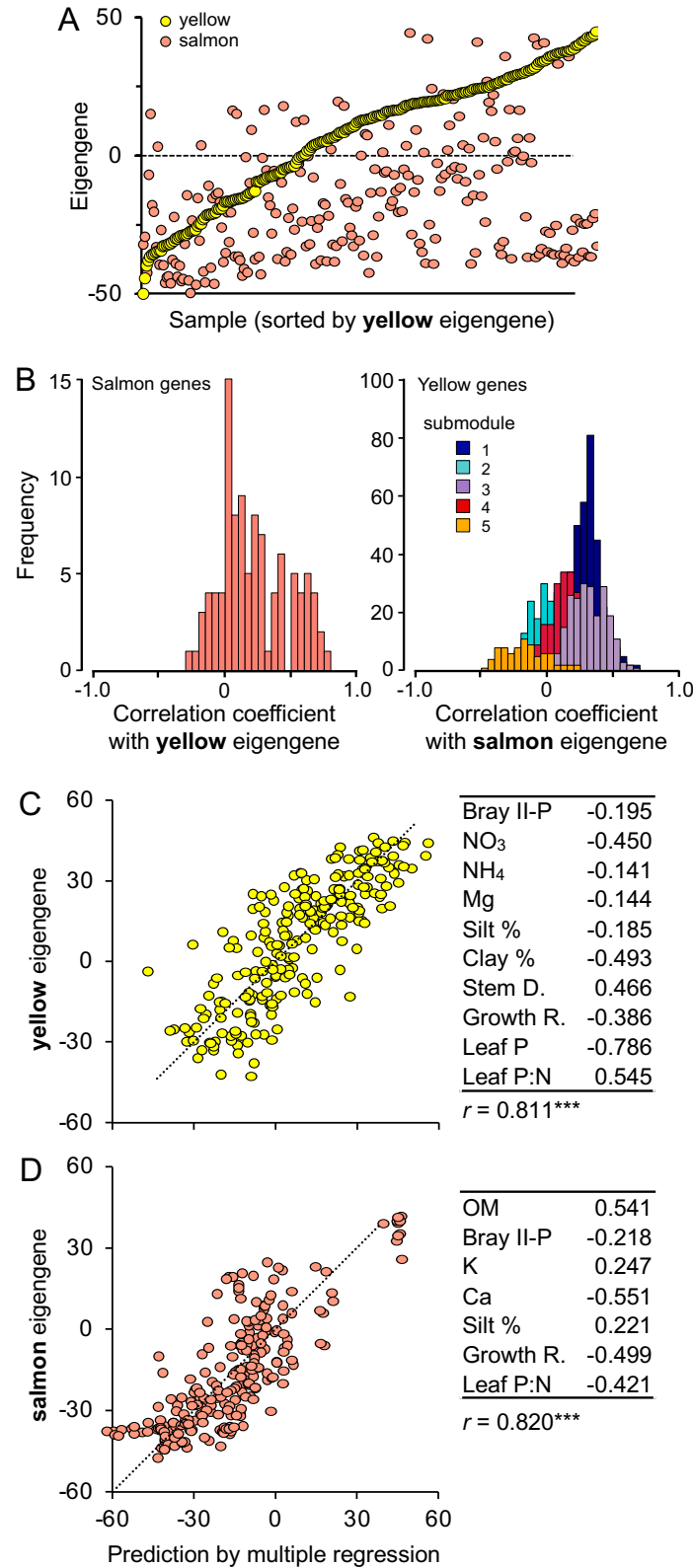


Figure 4. Interplay between the yellow and salmon modules in relation to soil and plant factors.

(A) Scatter plot of yellow (yellow) and salmon (salmon) eigengenes of the 251 samples, in which

the samples were sorted by the order of yellow eigengene. The eigengenes were standardized between the module from -50 (minimum value) to +50 (maximum value). (B) Frequency distributions of correlation coefficients of the salmon module genes with the yellow eigengenes (left) and those of the yellow module genes with the salmon eigengenes (right). The submodule numbers of the yellow module genes are indicated with the following colors: 1, dark blue; 2, turquoise; 3, purple; 4, red; 5, orange. The data were extracted from Dataset S6. Multiple regression analysis of the yellow (C) and salmon (D) eigengenes with the soil and plant factors. Regression coefficients of the factors with $P < 0.001$ are indicated in the tables. The correlation coefficients are also indicated below the tables (***, $P < 0.001$).

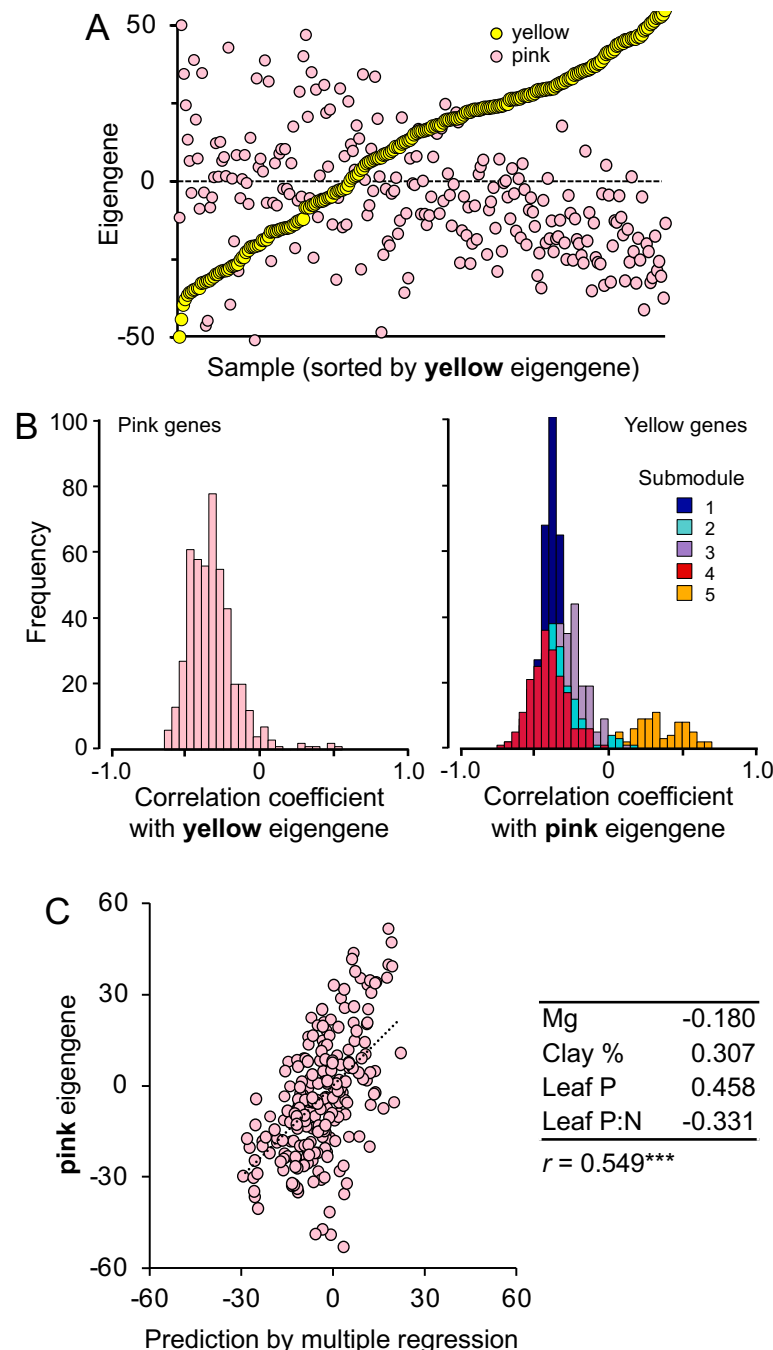


Figure 5. Interplay between the yellow and pink modules in relation to soil and plant factors. (A) Scatter plot of yellow (yellow) and pink (pink) eigengenes of the 251 samples, in which the samples were sorted by the order of yellow eigengene. The eigengenes were standardized between the module from -50 (minimum value) to +50 (maximum value). (B) Frequency distributions of correlation coefficients of the pink module genes with the yellow eigengenes (left) and those of the yellow module genes with the pink eigengenes (right). The submodule numbers of the yellow module genes are indicated with the following colors: 1, dark blue; 2, turquoise; 3, purple; 4, red; 5, orange. (C) Scatter plot of pink eigengene vs Prediction by multiple regression.

691 purple; 4, red; 5, orange. The data were extracted from Dataset S6. (C) Multiple regression
692 analysis of the pink eigengenes with the soil and plant factors. Regression coefficients of the
693 factors with $P < 0.001$ are indicated in the table. The correlation coefficient is indicated below the
694 table (***, $P < 0.001$).
695

Table 1. Enriched gene ontology (GO) terms and putative function of gene-coexpression modules.

| Module (no. of gene) | Enriched GO term ¹ | Putative function |
|----------------------|--|---------------------------------------|
| black (1,574) | DNA packaging, cellular component biogenesis, translation, RNA modification | Cell division |
| blue (4,134) | RNA splicing, ribonucleoside catabolic process, regulation of translation | Gene expression/translation |
| cyan (86) | (no enrichment) | (undefined) |
| darkgreen (46) | oxidoreductase activity | (undefined) |
| darkred (51) | (no enrichment) | House keeping (?) |
| green (1,257) | mitotic cell cycle, protein localization, organelle organization | Cell cycle regulation |
| greenyellow (139) | protein folding, response to oxidative stress, ER-associated degradation pathway | Stress-associated protein degradation |
| grey (8,460) | defense response, cellular amino acid metabolic process, immune response, response to nitrate | Immune response and N assimilation |
| grey60 (77) | antioxidant activity, oxidoreductase activity | Antioxidation |
| lightcyan (81) | branched-chain amino acid catabolic process, mitochondrial matrix | BCAA metabolism |
| lightgreen (67) | water transport, rhythmic process | Water uptake and diurnal rhythm |
| magenta (397) | defense response, immune response, response to other organism | Immune response |
| midnightblue (82) | lipid metabolic process, fatty acid metabolic process | Lipid biosynthesis |
| pink (471) | cell wall biogenesis, root morphogenesis | Root development |
| purple (275) | phosphorelay response regulator activity | (undefined) |
| royalblue (60) | response to decreased oxygen levels, lactate biosynthetic process | Response to hypoxia |
| salmon (101) | Cellular response to phosphate starvation, phosphate ion transport, acid phosphatase activity, galactolipid biosynthetic process | P-starvation response |
| tan (131) | trehalose biosynthetic process | Trehalose biosynthesis |
| yellow (1,023) | lipid biosynthetic process, terpenoid biosynthetic process, chitinase activity | Mycorrhizal formation |

¹ All GO terms enriched in the modules are listed in Dataset S4.

Table 2. Enriched gene ontology (GO) terms and putative function of the yellow submodules.

| Module (no. of gene) | Enriched GO term ¹ | Putative function |
|----------------------|--|-------------------------|
| Submodule 1 (332) | peptidase activity, transferase activity | Nutrient exchange |
| Submodule 2 (203) | fatty acid biosynthetic process, plastid part | Fatty acid biosynthesis |
| Submodule 3 (219) | terpenoid biosynthetic process, carotenoid metabolic process, plastid part | Carotenoid biosynthesis |
| Submodule 4 (189) | amino sugar metabolic process, defense response to fungus, extracellular space | Arbuscule degeneration |
| Submodule 5 (80) | (no enrichment) | (undefined) |

¹ All GO terms enriched in the modules are listed in Dataset S11.

Supplementary Information for

Cross-ecosystem transcriptomics identifies distinct genetic modules for nutrient acquisition in maize

Yusaku Sugimura^a, Ai Kawahara^b, Hayato Maruyama^a and Tatsuhiro Ezawa^a

^a Graduate School of Agriculture, Hokkaido University, Sapporo 060-8589, Japan; ^b Health & Crop Sciences Research Laboratory, Sumitomo Chemical, Co., Ltd., Takarazuka 665-8555, Japan.

Corresponding author: Tatsuhiro Ezawa
Email: tatsu@res.agr.hokudai.ac.jp

This PDF file includes:

Supplementary text
Figures S1 to S4
Tables S1 to S3
Legends for Datasets S1 to S11
SI References

Other supplementary materials for this manuscript include the following:

Datasets S1 to S11

Supplementary Information Text

Field sampling and soil/plant analysis. The root system was collected from a 30 × 30 cm area at a depth of 20 cm and washed briefly and gently in water. Then primary, lateral, and seminal roots (except for crown roots) were immediately collected with scissors, cut into small pieces (< 1 cm), randomized in water, collected on a stainless mesh, and blotted with a paper towel, and about 10 mg of the root pieces were immersed in 15 mL of RNAlater solution (Thermo Fisher Scientific), which were done within 3 min for each sample in the field sites. Fully-expanded-youngest leaves were also collected from the same individuals for P and N analysis. Leaf number, stem diameter (1 – 2 cm-above ground), and dry weight of the above ground part of the same individuals were measured. Root-zone soil samples (approx. 500 g) were also collected from the same individuals for the analysis of physical and chemical properties. The root samples were kept in RNAlater at least for 48 h at room temperature, transferred onto a stainless mesh to remove excess RNAlater, blotted, and stored in plastic tubes below -80°C.

Chemical and physical properties of the US and Japanese soil samples were analyzed at Midwest Laboratories (Omaha, NE) and Tokachi Federation of Agricultural Cooperatives (Obihiro, Japan), respectively, by the standard methods. The maize above-ground part and leaf samples were dried at 80°C for up to 5 days and weighed. The leaf samples were ground and digested with concentrated sulfuric acid and hydrogen peroxide at 200°C for 150 min, and then N and P concentrations were determined by the indophenol-blue method (1) and ascorbic acid-molybdate blue method (2), respectively.

RNA sequencing and gene expression profiling. Total RNA was extracted from the root samples using Maxwell RSC Plant RNA Kit with Maxwell RSC Instrument (Promega). Sequencing libraries were constructed using KAPA Stranded mRNA-Seq Kit (KAPA Biosystems), and single-end 75-base sequencing was performed on the illumina NextSeq 500 platform (10 M reads per sample) at Bioengineering Lab, Sagamihara, Japan. In ribosomal RNA sequencing (rRNA-Seq), the sequence libraries were constructed without mRNA purification. The sequence reads of which Phred quality score ≥20 in >80% of the bases were mapped to the maize B73 reference genome sequence (Zm-B73-REFERENCE-GRAMENE-4.0, ftp://ftp.ensemblgenomes.org/pub/plants/release-36/fasta/zea_mays/dna/) using HISAT2 program (3). Mapped reads were extracted and converted into BAM format using SAMtools (4), and uniquely mapped reads to the protein-coding transcript sequences were counted with featureCounts program (5). The raw read counts were normalized to Transcripts Per kilobase Million (TPM), and then genes with an average TPM ≥ 5 were subjected to subsequent analyses.

To identify maize genes involved in mycorrhizal formation and functioning, Blastp searches were performed against the maize genome using the amino acid sequences of those identified in *Oryza sativa*, *Sorghum bicolor*, *Medicago truncatula*, and *Lotus japonicus* as queries, and candidate genes were subjected to phylogenetic analysis based on the Neighbor-Joining method with MEGA X (6).

For weighted gene coexpression network analysis (WGCNA), the TPM values of the genes were log₂-transformed and analyzed using the R package WGCNA with the settings of 'automatic, one-step network construction', mergeCutHeight of 0.2, and minimal module size of 30 genes. After definition of modules, we merged the modules with a distance threshold value of 0.3. The k-means clustering analysis was performed using correlation distance as a measure with the R package amap. Fisher's exact test was employed to analyze GO term enrichment using Blast2GO program with false discovery rates of less than 0.01 and 0.05 for the modules and submodules, respectively (7). The GO annotation list was obtained from the maize genome database.

rRNA read assignment and de novo assembling of AM fungal mRNA reads. Maize and AM fungal rRNA sequences in the rRNA-Seq and the corresponding mRNA-Seq data were assigned to the maize large-subunit (LSU) rRNA sequence (XR_002749536) and 524 operational taxonomic units (OTUs) of AM fungal LSU rRNA sequences by Blastn searches with the following parameters: similarity, ≥ 95%; minimum alignment length, ≥ 75 bp; E-value, ≤ 1e-30. The consensus sequences GTGAAATTGTTGAAAGGGAAACG and GACGTAATGGCTTTAAACGAC

at the 5' and 3' ends, respectively, of the AM fungal sequences were removed before analysis. The numbers of the reads assigned to the plant rRNA and AM fungal OTUs were normalized to unit nt length, and total AM fungal read counts in the samples were standardized per 10^5 plant rRNA reads and transformed to logarithmic values.

For *de novo* assembling of AM fungal mRNA reads, the sequencing reads that were not mapped to the maize genome were collected from all 251 samples, combined into a fastq file, and a half of the reads was randomly extracted twice using SeqKit toolkit (8). The resultant two read sets were assembled separately with Trinity (ver. 2.4.0) (9), and then the contigs generated from the two batches were combined and clustered using CD-HIT-EST (10, 11) with a 100%-identity cutoff, and subjected to Blastx searches at an E-value cutoff of $1e-5$ against the genomes/transcripts of *Rhizophagus irregularis* DAOM181602, *R. clarus* HR1 (MAFF520076), *Diversispora epigaea* IT104, *Gigaspora margarita* BEG34, *Funnelformis mosseae* DAOM236685, *Acaulospora morrowiae* INVAM-CR315B, *Diversispora versiforme* INVAM-W47540, *Scutellospora calospora* INVAM-IL209, *Racocetra castanea* BEG1, *Paraglomus brasilianum* DAOM240472, *Ambispora leptoticha* INVAM-JA116, *Saccharomyces cerevisiae* S288c, *Neurospora crassa* OR74A, *Laccaria bicolor* S238N-H82, *Ustilago maydis* 521, *Rhizopus oryzae* 99-892, and *Phycomyces blakesleeianus* NRRL1555 (Table S3), as well as against the maize genome and Genbank bacterial genome database, in which contigs that showed the highest similarity to AM fungal sequences were considered as those originated from AM fungi. From the AM fungal contigs, those containing an open reading frame (ORF) of 50 or longer amino acid residues were predicted using TransDecoder (<https://transdecoder.github.io>). These processes are summarized in Fig. S4. The high-quality mRNA reads were mapped to the ORF sequences as described in the previous section.

Statistics. Principal component analysis and multiple linear regression analysis were performed using the `prcomp` and `lm` functions, respectively, in R. In variable selection in these analyses, one of the variables (soil and plant factors) that were highly correlated ($|r| \geq 0.9$) were excluded (Dataset S2). In addition, multicollinearity was also considered in multiple regression analysis. Prior to the multiple regression analysis between eigengenes and soil/plant factors, all variables were standardized between -50 (minimum) and +50 (maximum).

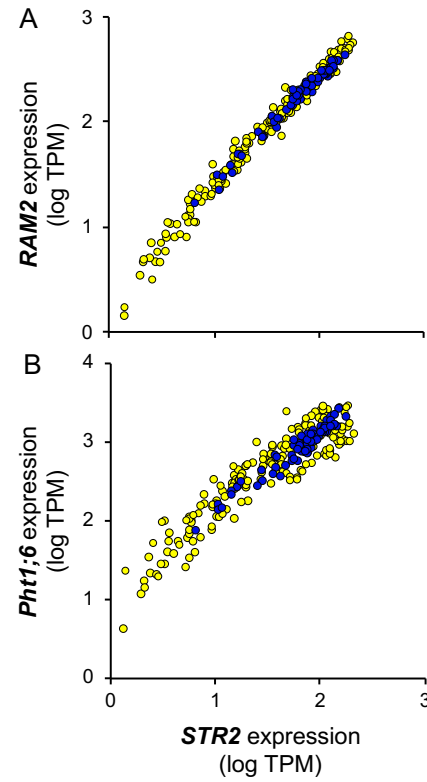


Fig. S1. Scatter plots of the gene expression levels of the Pi transporter (*Ph1;6*), putative lipid transporter (*STR2*), and glycerol-3-phosphoshate acyltransferase (*RAM2*) that are essential for arbuscule development and functioning. Correlation coefficients between *RAM2* and *STR2* and between *Ph1;6* and *STR2* are 0.939 and 0.930, respectively ($P < 0.001$).

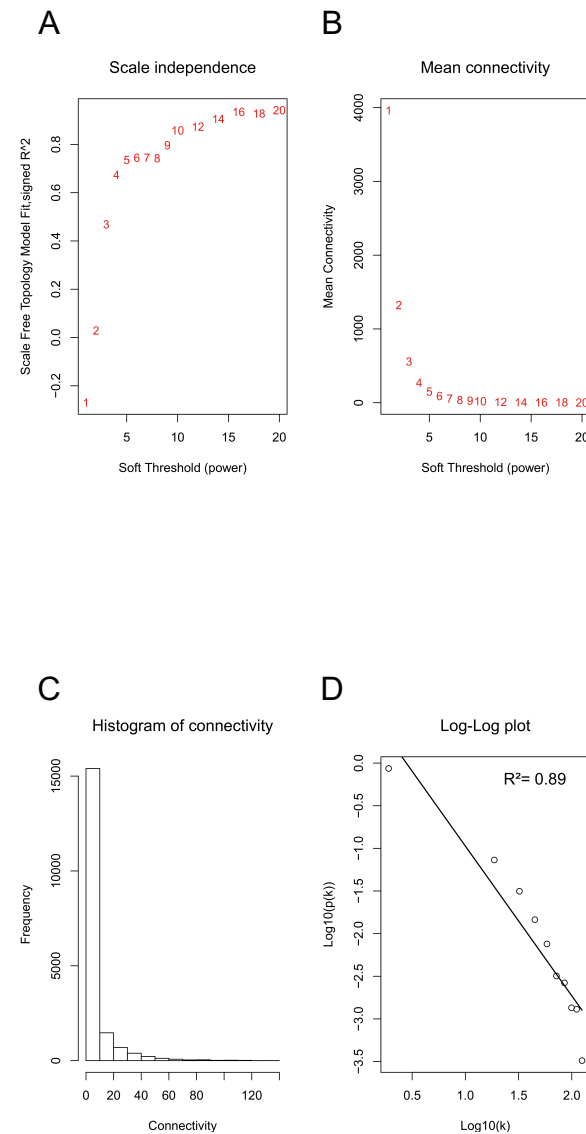


Fig. S2. Evaluation of weighted gene coexpression network properties in maize roots. (A) Scale-free fit index as a function of soft-threshold power. (B) Mean connectivity as a function of soft-threshold power. (C) Frequency distribution of network connectivities. (D) Log-log plot of edges [$\log_{10}(k)$] versus probability of a node having k edges [$P(k)$].

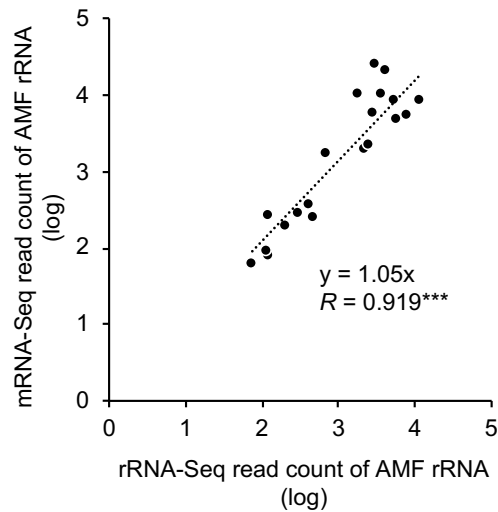


Fig. S3. Correlation analysis between AM fungal rRNA read numbers in the mRNA-Seq and those in the rRNA-Seq. Twenty samples were randomly chosen from the 251 RNA samples and subjected to rRNA-Seq (i.e. RNA-Seq without purification of mRNA). The reads obtained by mRNA-Seq and rRNA-Seq were assigned to maize LSU rRNA and 524 AM fungal operational taxonomic units by Blastn searches. The read counts were normalized to unit nt length, and total AM fungal read counts in the individual samples were standardized by 10^5 plant rRNA reads and transformed to logarithmic values.

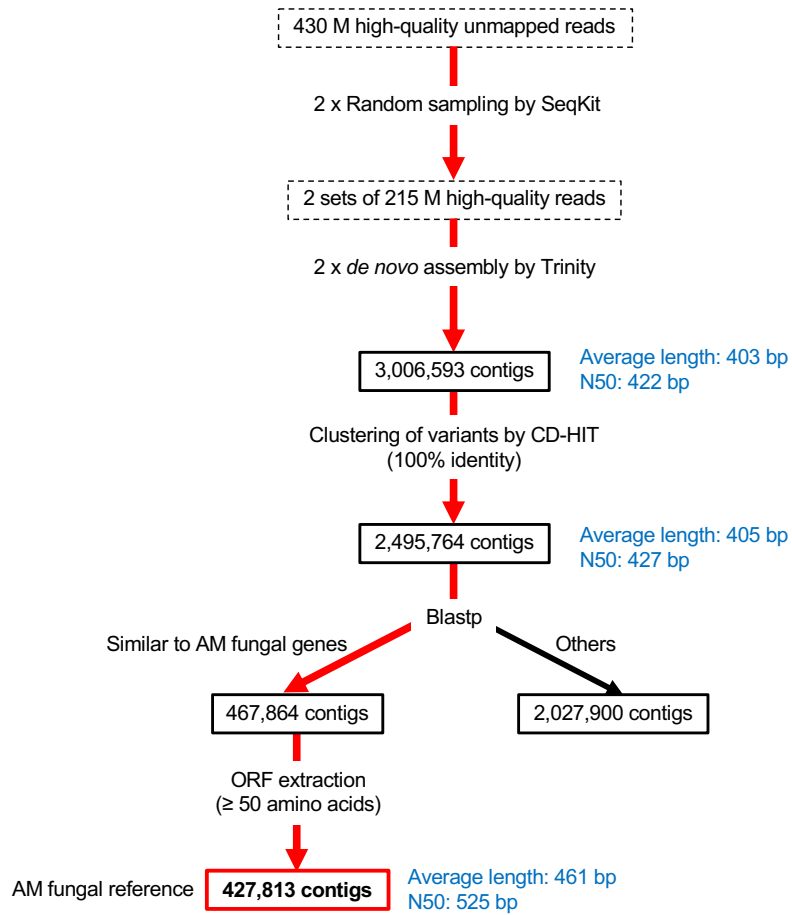


Fig. S4. Procedure for the construction of AM fungal transcript references from the unmapped reads of 251 samples.

Table S1. Geographic, climatic, and management data of the sampling sites.

| Country (year) | Location | Latitude/longitude | Altitude (m) | Climatic zone | Annual climatic data | | | | Plot ID | Corn variety | No. of sample | Sowing date | Sampling date | Note |
|-------------------|------------------------------|--------------------|-----------------|-------------------|----------------------|-----------------------|----------------------|-----------------------|----------|-----------------|------------------|------------------|--------------------|--|
| | | | | | Mean temp. (°C) | Highest temp. (°C) | Lowest temp. (°C) | Precipitation (mm) | | | | | | |
| USA (2017) | | | | | | | | | | | | | | |
| | Carlyle, IL | 38°41'N/ 89°02'W | 167 | Humid continental | 12.5 | 17.9 | 7.1 | 1,103 | IL | P1257AM | 15 | 25 April | 16 June | Conventional arable field |
| | Fishers, IN | 39°59'N/ 85°56'W | 248 | Humid continental | 11.3 | 16.6 | 5.9 | 1,086 | IN | 5441SX | 12 | 26 April | 19 June | Conventional arable field |
| | Bagley, IA | 41°51'N/ 94°24'W | 343 | Humid continental | 9.1 | 15.1 | 3.1 | 819 | IA | R9734VT3Pro | 12 | 31 May | 21 June | Conventional arable field |
| | York, NE | 40°54'N/ 97°38'W | 487 | Humid continental | 10.1 | 17.1 | 3.2 | 768 | NE | P1257AM | 15 | 2 May | 22 June | Conventional arable field with irrigation |
| | Groom, TX | 35°17'N/ 101.12'W | 1012 | Semi-arid | 13.8 | 21.1 | 6.5 | 591 | TX | Agrisure 3000GT | 12 | 1 June | 29 June | Conventional arable field with irrigation |
| Japan (2018) | | | | | | | | | | | | | | |
| | Kasai, Hyogo | 34°54'N/ 134°49'E | 64 | Humid subtropical | 15.9 | 21.1 | 11.5 | 1,671 | KASc | P9400 | 12 | 23 April | 18, 21, 29 May | After fallow (non-arable field) |
| | | | | | | | | | KASc | P1543 | 12 | 23 April, 22 May | 21, 29 May, 4 June | Conventional arable field |
| | | | | | | | | | KASf | P2088/ P1543 | 13 | 23 April, 28 May | 30 May, 10 July | After clear cut of cedar |
| | Makabe, Ibaraki | 36°15'N/ 14°04'E | 44 | Humid subtropical | 15.3 | 20.6 | 10.6 | 1,060 | MKB | P1543 | 14 | 7 May | 7 and 20 June | Conventional arable field |
| | Higashi-Hiroshima, Hiroshima | 34°23'N/ 132°43'E | 210 | Humid subtropical | 14.0 | 19.5 | 8.8 | 1,654 | HROp | SH4812 | 8 | 1 June | 5 - 6 July | N-P-K = 250-50-200 kg ha ⁻¹ year ⁻¹ |
| | | | | | | | | | HROc | SH4812 | 8 | 1 June | 5 - 6 July | N-P-K = 250-100-200 kg ha ⁻¹ year ⁻¹ |
| | | | | | | | | | HROm | SH4812 | 8 | 1 June | 5 - 6 July | Cow manure 20 t ha ⁻¹ year ⁻¹ |
| | Togo, Aichi | 35°06'N/ 137°04'E | 69 | Humid subtropical | 16.9 | 22.0 | 12.9 | 1,696 | NGYm | Canberra 90EX | 8 | 13 April | 11 May | Food compost 25 t ha ⁻¹ year ⁻¹ |
| | | | | | | | | | NGYs | P2088 | 22 | 13 April | 11 May | Conventional arable field |
| | Tsuruoka, Yamagata | 38°44'N/ 139°49'E | 16 | Humid subarctic | 13.4 | 17.8 | 9.4 | 2,449 | TROb | P9400 | 15 | 8 June | 28 June | Sterilized with Dazomet in 2016 |
| | | | | | | | | | TROf | P9400 | 15 | 8 June | 29 June | After paddy field |
| | Sapporo, Hokkaido | 43°04'N/ 141°20'E | 18 | Humid subarctic | 9.5 | 13.3 | 6.0 | 1,282 | SPK/+NPK | P2023 | 8 | 28 May | 25 June, 17 July | N-P-K = 100-100-100 kg ha ⁻¹ year ⁻¹ |
| | | | | | | | | | SPK/-NPK | P2023 | 8 | 28 May | 25 June, 17 July | No fertilizer |
| | | | | | | | | | SPK/-P | P2023 | 8 | 28 May | 25 June, 17 July | N-P-K = 100-0-100 kg ha ⁻¹ year ⁻¹ |
| | | | | | | | | | SPKc | New Dent 105 | 8 | 16 May | 13 June, 9 July | Continuous corn |
| | | | | | | | | | SPKg | New Dent 105 | 8 | 16 May | 13 June, 9 July | After grassland |
| | Utsunomiya, Tochigi | 36°32'N/ 139°54'E | 109 | Humid subtropical | 15.2 | 20.7 | 10.7 | 1,257 | UTS | P2023 | 10 | 17 May | 22 June | Conventional arable field |

Table S2. Multiple regression analysis of the yellow module eigengenes with the transcript levels of AM fungal genes.

| | Regression coefficient | |
|--------------------------------------|---------------------------|-----|
| <i>SYG1-1</i> expression | 0.150 | *** |
| <i>PHO91</i> expression | 0.233 | ** |
| rRNA abundance | 0.875 | *** |
| Correlation coefficient (<i>r</i>) | 0.884 | *** |

, $P < 0.01$; *, $P < 0.001$

Table S3. Source of fungal databases for Blastx searches of the contigs obtained by the *de novo* assembly of the unmapped reads.

| Species | Genome/transcripts | Source URL |
|---|--------------------|---|
| <i>Rhizophagus irregularis</i> DAOM181602 | Genome | https://www.ncbi.nlm.nih.gov/Traces/wgs/?val=BDIQ01#contigs |
| <i>Rhizophagus clarus</i> HR1 (MAFF520076) | Transcripts | http://lab.agr.hokudai.ac.jp/botagr/rhizo/RhizoCont/Download.html |
| <i>Diversispora epigaea</i> IT104 | Genome | https://www.ncbi.nlm.nih.gov/Traces/wgs/?val=PQFF00000000 |
| <i>Gigaspora margarita</i> BEG34 | Transcripts | https://www.ncbi.nlm.nih.gov/Traces/wgs/?val=GBYF00000000 |
| <i>Funneliformis mosseae</i> DAOM236685 | Transcripts | https://github.com/zygolife/AMF_Phylogenomics/tree/master/SCT_ALL_DATA |
| <i>Acaulospora morrowiae</i> INVAM-CR315B | Transcripts | https://github.com/zygolife/AMF_Phylogenomics/tree/master/SCT_ALL_DATA |
| <i>Diversispora versiforme</i> INVAM-W47540 | Transcripts | https://github.com/zygolife/AMF_Phylogenomics/tree/master/SCT_ALL_DATA |
| <i>Scutellospora calospora</i> INVAM-IL209 | Transcripts | https://github.com/zygolife/AMF_Phylogenomics/tree/master/SCT_ALL_DATA |
| <i>Racocetra castanea</i> BEG1 | Transcripts | https://github.com/zygolife/AMF_Phylogenomics/tree/master/SCT_ALL_DATA |
| <i>Paraglomus brasilianum</i> DAOM240472 | Transcripts | https://github.com/zygolife/AMF_Phylogenomics/tree/master/SCT_ALL_DATA |
| <i>Ambispora leptoticha</i> INVAM-JA116 | Transcripts | https://github.com/zygolife/AMF_Phylogenomics/tree/master/SCT_ALL_DATA |
| <i>Saccharomyces cerevisiae</i> S288c | Genome | ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/146/045/GCA_000146045.2_R64/GCA_000146045.2_R64_genomic.fna.gz |
| <i>Neurospora crassa</i> OR74A | Genome | ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/182/925/GCA_000182925.2_NC12/GCA_000182925.2_NC12_genomic.fna.gz |
| <i>Laccaria bicolor</i> S238N-H82 | Genome | ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/143/565/GCA_000143565.1_V1.0/GCA_000143565.1_V1.0_genomic.fna.gz |
| <i>Ustilago maydis</i> 521 | Genome | ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/328/475/GCA_000328475.2_Umaydis521_2.0/GCA_000328475.2_Umaydis521_2.0_genomic.fna.gz |
| <i>Rhizopus oryzae</i> 99-892 | Genome | ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/697/725/GCA_000697725.1_RhiOry99-892-1.0/GCA_000697725.1_RhiOry99-892-1.0_genomic.fna.gz |
| <i>Phycomyces blakesleeanus</i> NRRL1555 | Genome | https://genome.jgi.doe.gov/portal/Phyb2/download/Phycomyces_blakesleeanus_v2_masked_scaffolds.fasta.gz |

Dataset S1. Soil and plant factors used for principle component analysis and correlation analysis.

Dataset S2. Correlation coefficient and probability between soil and plant factors.

Dataset S3. Read assignment summary.

Dataset S4. Gene ontology (GO) analysis of the coexpression modules in the maize roots.

Dataset S5. Eigengenes of the coexpression modules in the maize roots.

Dataset S6. Regression coefficients in correlation analysis between the eigengenes and soil/plant factors.

Dataset S7. Correlation coefficients and probabilities between the expression levels of individual genes and the eigengenes.

Dataset S8. Number of mRNA-Seq read assigned to the LSU rRNA genes of AM fungal operational taxonomic units and maize.

Dataset S9. Number of rRNA-Seq read assigned to the LSU rRNA genes of AM fungal operational taxonomic units and maize in randomly selected 20 samples.

Dataset S10. Expression levels (TPM) of AM fungal *SYG1-1*, *PHO91*, *VTC4*, and *AQP3* of which transcript references were obtained by *de novo* assembling of unmapped reads.

Dataset S11. Gene ontology (GO) analysis of the submodules of yellow module.

SI References

1. I. Novamsky, R. van Eck, C. van Schouwenburg, I. Walinga, Total nitrogen determination in plant material by means of the indophenol-blue method. *Wageningen J. Life Sci.* **22**, 3-5 (1974).
2. F. S. Watanabe, S. R. Olsen, Test of an ascorbic acid method for determining phosphorus in water and NaHCO₃ extracts from soil. *Soil Sci Soc Am J* **29**, 677-678 (1965).
3. D. Kim, J. M. Paggi, C. Park, C. Bennett, S. L. Salzberg, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotech.* **37**, 907-915 (2019).
4. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
5. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2013).
6. S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **35**, 1547-1549 (2018).
7. A. Conesa *et al.*, Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674-3676 (2005).
8. W. Shen, S. Le, Y. Li, F. Hu, SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLOS ONE* **11**, e0163962 (2016).
9. M. G. Grabherr *et al.*, Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotech.* **29**, 644-652 (2011).
10. W. Li, A. Godzik, Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658-1659 (2006).
11. L. Fu, B. Niu, Z. Zhu, S. Wu, W. Li, CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150-3152 (2012).