Title

Transcriptional regulation of nitrogen-associated metabolism and growth.

Journal

Nature, 563(7730)

ISSN

0028-0836

Authors

Gaudinier, Allison
Rodriguez-Medina, Joel
Zhang, Lifang
et al.

Publication Date

2018-11-01

DOI

10.1038/s41586-018-0656-3

Peer reviewed

LETTER

https://doi.org/10.1038/s41586-018-0656-3

Transcriptional regulation of nitrogen-associated metabolism and growth

Allison Gaudinier[1], Joel Rodriguez-Medina[1], Lifang Zhang[2], Andrew Olson[2], Christophe Liseron-Monfils[2], Anne-Maarit Bågman[1], Jessica Foret[1], Shane Abbitt[3], Michelle Tang[1,4], Baohua Li[4], Daniel E. Runcie[4], Daniel J. Kliebenstein[4,5], Bo Shen[3], Mary J. Frank[3], Doreen Ware[2,6] & Siobhan M. Brady[1]*

Nitrogen is an essential macronutrient for plant growth and basic metabolic processes. The application of nitrogen-containing fertilizer increases yield, which has been a substantial factor in the green revolution[1]. Ecologically, however, excessive application of fertilizer has disastrous effects such as eutrophication[2]. A better understanding of how plants regulate nitrogen metabolism is critical to increase plant yield and reduce fertilizer overuse. Here we present a transcriptional regulatory network and twenty-one transcription factors that regulate the architecture of root and shoot systems in response to changes in nitrogen availability. Genetic perturbation of a subset of these transcription factors revealed coordinate transcriptional regulation of enzymes involved in nitrogen metabolism. Transcriptional regulators in the network are transcriptionally modified by feedback via genetic perturbation of nitrogen metabolism. The network, genes and gene-regulatory modules identified here will prove critical to increasing agricultural productivity.

The root system takes up and metabolizes bio-available nitrogen and transduces nitrogen signals. In response to reduced nitrogen availability, plant development is adjusted—this includes increased lateral root elongation to forage for nitrogen[3]. Above ground, rosette size is decreased and plants flower earlier[4]. Diverse molecular events underlie these morphological changes. Nitrogen transporters, assimilation enzymes and signalling factors are transcriptionally regulated in response to changes in available nitrogen[5]. Post-transcriptional, calcium- and phosphorylation-dependent signalling cascades are also critical regulators of this transcriptional response[6]. Concomitantly, carbon metabolism and hormone pathways are also altered to adjust metabolic pathways and plant growth[7]. Sixteen transcription factors in *Arabidopsis thaliana* have previously been identified to have a role in nitrogen metabolism[8–19] (Supplementary Table 1), through a range of approaches that includes systems-level studies[16,20]. Despite the importance of the root system in regulating responses to nitrogen, only seven of these transcription factors have previously been shown to regulate root development in a nitrogen-dependent manner[8,9,11,14–17].

Using enhanced yeast one-hybrid assays, we screened for transcription factors that regulate nitrogen metabolism[21,22]. Because nitrogen metabolism is interconnected with a range of different processes, we included target promoters from genes associated with nitrogen transport (12 promoters), assimilation (11 promoters), signalling (2 promoters), connections to nitrogen metabolism through amino acid metabolism (5 promoters), carbon metabolism (10 promoters), carbon transport (4 promoters), organ growth (5 promoters) and hormone responses (7 promoters) as well as associated transcription factors (12 promoters) (Supplementary Table 2). We screened these promoters against transcription factors expressed in roots. The resulting network comprises 1,660 interactions between 431 genes, 345 transcription factors and 98 promoters (Fig. 1a, Extended Data Fig. 1, Supplementary Table 3a). We call this network the 'yeast one-hybrid network for nitrogen-associated metabolism' (YNM). Our assays captured previously characterized interactions: NLP7 physically binds to and regulates expression of *NIR1* and *CIPK8*, and NLP6 binds to and regulates expression of *NIR1*[13,23]. Within the YNM we found what is, to our knowledge, a previously undescribed putative hierarchical regulation of transcription factors—including both known nitrogen-regulatory transcription factors and transcription factors identified in this study—that bind to promoters of genes in many processes, such as the nitrate assimilation pathway (Fig. 1b, Supplementary Table 3b). A signalling cascade that links the nitrate-mediated regulation of Ca^{2+}-sensor protein kinases to transcriptional regulation via NLP7 is also significantly overrepresented in the YNM[6] ($P = 2.14 \times 10^{-9}$) (Extended Data Fig. 2a, Supplementary Table 4a). Moreover, the YNM is enriched for hormone-regulated genes, which supports previous findings that hormone signalling is integrated into the regulation of nitrogen metabolism[9,24] (see Methods, Extended Data Figs. 2b–h, Supplementary Table 4b–h). The highly combinatorial nature of interactions is consistent with previous studies that suggest that transcription factors that are central within the YNM may regulate multiple processes that are related to nitrogen metabolism[23]. NLP7 bound to promoters of seven nitrogen-associated categories (Supplementary Table 3c). One hundred and seventy-five transcription factors from the YNM were found to bind to gene promoters that are involved in more than one nitrogen-associated process (Supplementary Table 2d).

We used a variety of datasets and approaches to rank transcription factors in the YNM for functional validation. First, under the premise that transcription factors and their targets are co-expressed upon changes in nitrogen availability, we prioritized highly correlated transcription factors and targets for a nitrogen treatment and a cell-type-specific dataset (Supplementary Tables 5, 6). This approach does not exclude the possibility of detecting self-regulating repressors or activators. Second, we used the network analysis algorithm NeCorr (see Methods, Supplementary Table 7). Third, transcription factors were evaluated for their outgoing connectivity (Supplementary Table 3e). Additionally, transcription factors were considered given the total number and percentage of targets that are classical nitrogen-metabolism genes (Supplementary Table 8). As a positive control, we included mutants of the transcription factors NLP7 and GNC, and the transceptor NPF6.3 (also known as NRT1.1 or CHL1)[11,25,26]. Perturbation of nitrogen metabolism in *npf6.3* (also known as *nrt1.1* or *chl1*) and *nrt2.1* plants alters lateral root initiation and/or lateral root elongation[27,28]. With the hypothesis that these transcription factors regulate nitrogen metabolism and nitrogen status, we examined their mutant root system architecture (RSA) (see Methods) under limiting (1 mM KNO_3) and sufficient nitrogen (10 mM KNO_3) (Extended Data Fig. 3, Supplementary Table 2c).

Mutant alleles of seventeen genes that we identify here showed significant changes in at least one RSA trait relative to wild type (Supplementary Tables 9, 10, Supplementary Data 1, 2). *chl1-5*—a

[1]Department of Plant Biology and Genome Center, University of California, Davis, Davis, CA, USA. [2]Cold Spring Harbor Laboratory, Cold Spring Harbor, Cold Spring Harbor, NY, USA. [3]DuPont Pioneer, Johnston, IA, USA. [4]Department of Plant Sciences, University of California, Davis, Davis, CA, USA. [5]DynaMo Center of Excellence, University of Copenhagen, Frederiksberg C, Denmark. [6]US Department of Agriculture, Agricultural Research Service, Ithaca, NY, USA. *e-mail: sbrady@ucdavis.edu

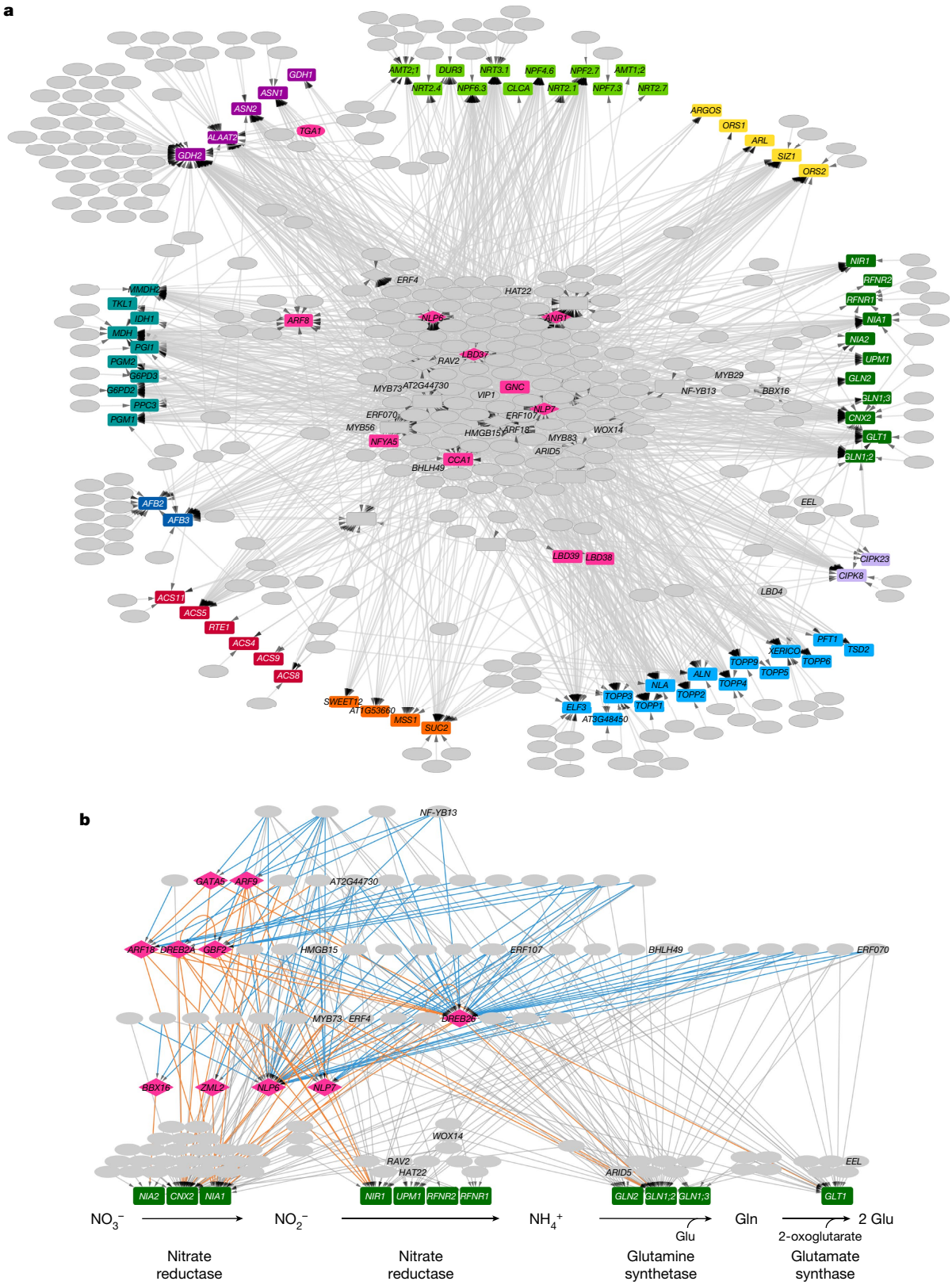


Fig. 1 | Combinatorial interactions between transcription factors and promoters of genes associated with nitrogen metabolism, signalling and nitrogen-associated processes. a, Interaction network for nitrogen-associated metabolism. See Extended Data Fig. 1 for the full diagram, including gene names. Rectangles, promoters; ovals, transcription factors; and diamonds, genes represented as both promoters and transcription factors. Nitrogen-associated biological processes are indicated by promoter colour. A grey line indicates an interaction between transcription factor and promoter. Light green, nitrogen transporter; yellow, organ growth; dark green, nitrate assimilation; light purple, nitrogen signalling; light blue, nitrogen-linked; orange, carbon metabolism; red, ethylene; dark blue, auxin; teal, carbon transporter; dark purple, amino acid metabolism; and pink, transcription factors linked to nitrogen. **b**, Transcription factor–promoter interactions that are associated with nitrate assimilation are hierarchical. Edges participating in hierarchical regulation going into the transcription factors (diamonds) are blue, and outgoing edges from the transcription factors are orange. The NLP7 and NLP6 regulators are in the first tier of transcription-factor binding to assimilation enzymes.

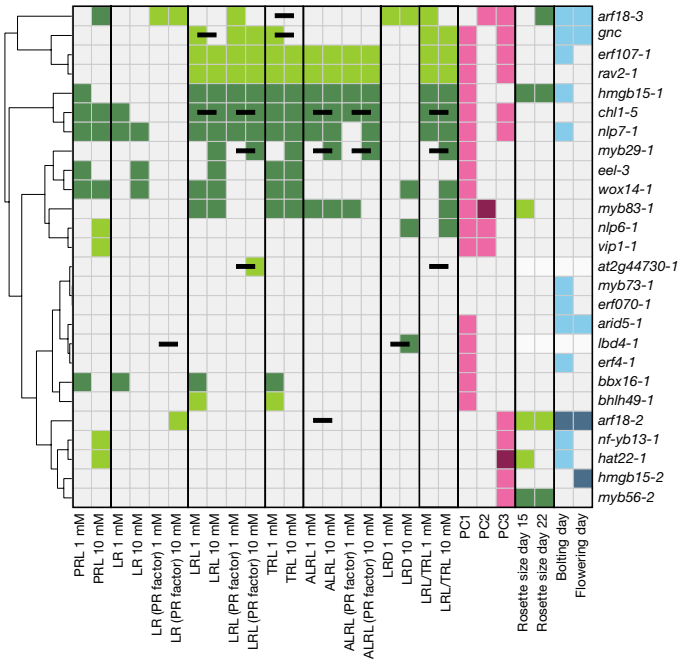


Fig. 2 | Phenotypes associated with transcription-factor mutant alleles. Mutant alleles are listed in rows and measured traits in columns. Statistically significant differences relative to wild type (Col-0) are shown with a coloured cell within the heat map ($P < 0.05$ as determined using a two-way ANOVA, exact *n* and *P* values for the analysis can be found in Supplementary Table 10). Trait categories are indicated with a dark-edged vertical line. Moving from left to right this comprises primary root length, lateral root number, total lateral root length, total root length, average lateral root length, lateral root density, ratio of total lateral root length to total root length, principal component analysis, rosette size and bolting and flowering analysis. Root traits were measured from 9-day-old plants grown on 1 mM KNO_3 or 10 mM KNO_3. PRL, primary root length; LR, number of lateral roots; LRL, total lateral root length; total root length (TRL), PRL + LRL; average lateral root length (ALRL), LRL divided by LR; lateral root density (LRD), LR divided by PRL; LRL/TRL, LRL divided by TRL. 'PR factor' indicates that PRL was considered as a factor in the ANOVA model; PC1, principal component 1; PC2, principal component 2; PC3, principal component 3. Dark green, phenotype is larger than Col-0; light green, phenotype is smaller than to Col-0; horizontal black bar, genotype-by-condition interaction. Genotype-specific (light pink) and genotype-by-condition-specific (dark pink) effects are shown, when considering variation across all root traits in a principal component analysis in PC1, PC2 and PC3. Light blue, early bolting and flowering; dark blue, late bolting and flowering. Mutants are hierarchically clustered using the Manhattan distance metric.

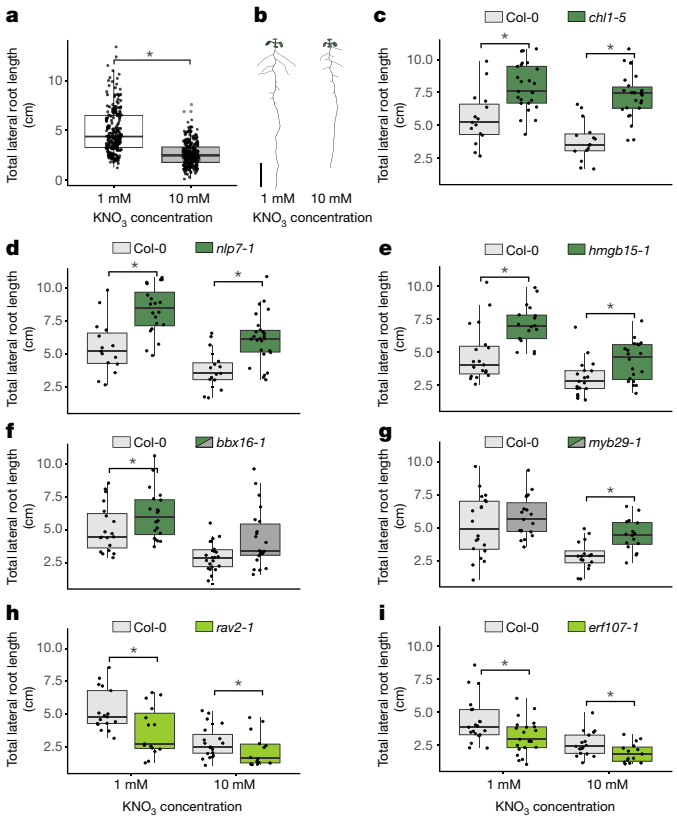


Fig. 3 | Total lateral root length phenotypes that are dependent on genotype and nitrate condition. a, Col-0 LRL is significantly longer at 1 mM KNO_3 compared to 10 mM KNO_3. **b**, Average Col-0 root growth. Scale bar, 1 cm. **c**, **d**, The *chl1-5* and *nlp7-1* mutant alleles were included as a nitrate transceptor (*chl1-5*, **c**) and master transcriptional regulator (*nlp7-1*, **d**). Both mutants have a genotype-dependent influence on LRL at 1 and 10 mM KNO_3, with an LRL that is longer than that of wild type. *chl1-5* also has a genotype-by-treatment influence on LRL. **e**, The *hmgb15-1* allele shows a genotype-dependent influence on LRL relative to wild type, which is similar to *nlp7-1* and *chl1-5* relative to wild type. **f**, The *bbx16-1* allele has an influence on LRL that is dependent on nitrate condition, with a longer LRL only at 1 mM KNO_3. **g**, The *myb29-1* allele has an influence on LRL that is dependent on nitrate condition, with a longer LRL only at 10 mM KNO_3. **h**, **i**, The *rav2-1* (**h**) and *erf107-1* (**i**) alleles are both genotype-dependent at both 1 mM and 10 mM KNO_3, with shorter LRLs. $^*P < 0.05$, two-way ANOVA; exact *n* and *P* values for the analysis can be found in Supplementary Table 10. Box plots are centred at the data median and mark from the 25th to the 75th percentile. Individual measurements are plotted as black dots.

mutant of NPF6.3—displayed changes in its RSA that were dependent on genotype and on genotype-by-nitrate conditions (Figs. 2, 3c). Similarly, *nlp7-1* and *hmgb15-1* plants displayed larger root systems, with genotype-dependent changes in their RSA (Figs. 2, 3d, e). *bbx16-1* plants had larger root systems under limiting nitrogen conditions (Figs. 2, 3f). Conversely, the *myb29-1* mutant had increased lateral root length, lateral root density and total root length under sufficient nitrate, in a manner that was dependent on genotype-by-nitrate conditions (Figs. 2, 3g). By contrast, the *erf107-1* and *rav1-2* plants showed a genotype-dependent decrease in the size of traits related to their lateral roots, in both nitrate conditions (Figs. 2, 3h, i). The *gnc* mutant showed decreases in lateral root length that were dependent on nitrate conditions and on genotype-by-nitrate condition (Fig. 2). The phenotype of *gnc* plants differed from that of *erf107-1* and *rav2-1* plants, in that these latter two mutants did not show any dependence on nitrate conditions. The composite principal component traits provided additional insights into perturbations in root growth that could not be discerned by looking at individual traits (Extended Data Fig. 4, Supplementary Table 11). In these experiments, we determined genes that control nitrogen-associated root length, lateral root development, and lateral root development that is dependent on primary root length, and then overlaid these on the YNM along with genes that regulate primary root length (Supplementary Table 12) and lateral root initiation[29] (Extended Data Fig. 5).

Given that perturbed RSA was observed in these mutants, we next determined whether the altered nitrogen status of the mutants affected shoot development and the transition from vegetative to reproductive growth (see Methods). Mutant alleles of thirteen genes showed a difference in either rosette size and/or bolting and flowering time (Supplementary Tables 9, 10, Supplementary Data 1, 2). Plants with the *arf18-2* allele had a smaller rosette with an increased number of days to flowering, whereas *arf18-3* plants showed the opposite phenotype. A change in rosette size was coupled with a change in the time to bolting or flowering for four mutants. *arf18-2*, *arf18-3* and *hmgb15-1* showed the most significant changes in both root and shoot system architecture. A significant reduction in both ^{15}N in *rav2-1* plants and in the C:N ratio in *nlp7-1* plants was observed (Extended Data Fig. 6). Classical plant physiology experiments have also associated changes

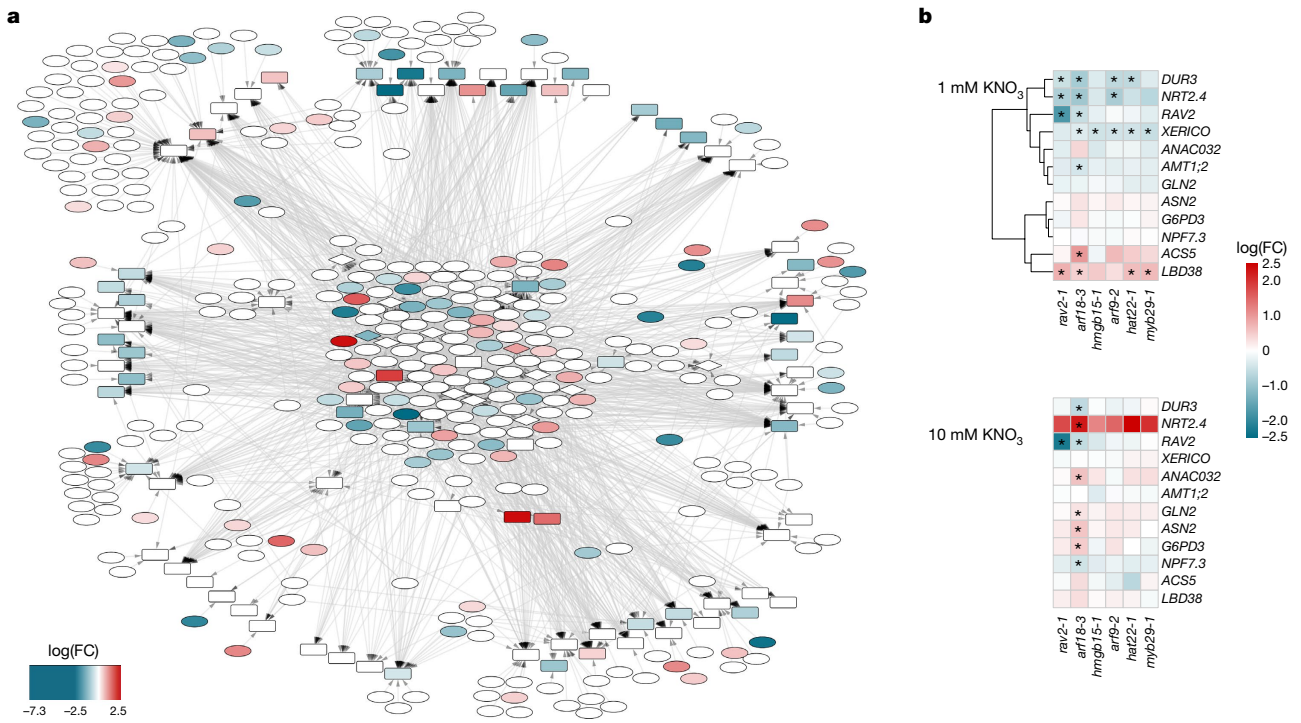


Fig. 4 | The nitrate-responsive transcriptional regulatory network. **a**, The network is enriched for genes that are differentially expressed in the root grown on 1 mM KNO_3 or 10 mM KNO_3. Nodes that are significantly differentially expressed are coloured according to their log(fold change) (log(FC)), from −2.5 to 2.5 (Supplementary Table 14a, b). Node shape is the same as Fig. 1. **b**, Heat map showing the expression of specific YNM target genes: *DUR3*, *NRT2.4*, *RAV2*, *XERICO*, *ANAC032*, *AMT1;2*, *GLN2*, *ASN2*, *G6PD3*, *NPF7.3*, *ACS5* and *LBD38* in 1 mM KNO_3 and 10 mM KNO_3. Each cell represents the log(fold change) relative to the control, as determined using a two-sided test with limmaVoom. *Corrected $P < 0.05$, (Supplementary Table 15). Four biological replicates were sampled per genotype per condition.

in nitrogen status with perturbation of chlorophyll levels. *nlp7-1* and *gnc* mutants showed significant reduction in their total chlorophyll content, whereas *lbd4-1* had increased chlorophyll content (Extended Data Fig. 7). Changes in shoot growth in the mutants were significantly correlated with the number of targets each transcription factor had in the YNM (Spearman rank correlation, $P < 0.05$), as well as with the number of biological processes that these transcription factors putatively regulated ($P < 0.05$, Supplementary Table 13). Thus, network connectivity is predictive of the influence of a given transcription factor on shoot growth.

Changes in nitrogen availability are accompanied by changes in transcription[5,9,12,30]. Furthermore, the changes in development of transcription-factor mutants under conditions of both limiting and sufficient nitrogen are probably coordinated by perturbations in the underlying transcriptional regulatory network. To link mutant phenotypes with transcriptional changes, whole-genome expression was measured in a subset of mutant genotypes (see Methods, Supplementary Tables 14, 15). To provide further support that the YNM reflects the transcriptional regulation of nitrogen-dependent processes in the root, we tested for enrichment of nitrogen-status genes. Genes displaying differential expression in wild-type roots in 1 mM relative to 10 mM KNO_3 were significantly enriched in the YNM ($P = 3.94 \times 10^{-9}$) (Fig. 4a). Thus, the YNM captures transcriptional regulation of root nitrogen status.

At the level of individual transcription factors, *ARF9* and *ARF18* alleles showed differential expression of nitrogen-related genes. ARF9 regulates the expression of two direct targets as predicted by the YNM (*XERICO* and *DUR3*) as well as *NRT2.4*, *NPF7.3*, *GLN2* and *ASN2*. ARF18 regulated expression of three direct targets as predicted by the YNM (*NRT2.4*, *ANAC032* and *XERICO*) as well as *ACS5*, *DUR3*, *G6PD3* and *AMT1;2* (Fig. 4b). HMGB15 regulated the expression of one predicted direct target, *XERICO* (Fig. 4b). *LBD38* is misregulated in the mutants of *ARF18*, *MYB29*, *RAV2* and *HAT22*; misregulation of *NRT2.4* was found in the mutants of *ARF18*, *HAT22* and *RAV2* (Fig. 4b).

A common mode of regulation in metabolism is metabolite feedback. To test whether feedback is present within the YNM, we curated gene-expression datasets of nitrate transporters and a transceptor, metabolic-enzyme mutants and genotype-by-nitrogen-dependent changes in mutants of previously described transcriptional regulators of nitrogen metabolism (Supplementary Table 16, see Methods). Upon perturbation of nitrogen transport, sensing and metabolism, genes in the YNM were significantly enriched for differential expression (Fig. 5a). Thus, a perturbation in nitrate uptake, reduction and the glutamine oxoglutarate aminotransferase cycle results in transcriptional perturbation of enzymes involved in nitrogen metabolism, and their upstream regulators. Genetic perturbation of nitrogen metabolism via the nitrogen-regulatory transcription factors also perturbs more genes in the YNM than expected by chance (Fig. 5a). Clustering analysis revealed targets of this metabolic feedback (Extended Data Figs. 8–10). A core set of enzymes involved in nitrogen metabolism—representing nearly every step of nitrate uptake, assimilation and conversion to glutamine and glutamate—were perturbed in most of the metabolic-mutant backgrounds queried. These perturbed genes include *NPF6.3*, *NRT3.1*, *NIA1*, *NIR1*, *G6PD2*, *G6PD3*, *RFNR1*, *RFNR2*, *ASN1* and the transcription factor *RAV2*, found in this study (Fig. 5b). Another cluster includes known nitrogen-associated genes *TGA1*, *NLP7*, *CIPK8*, *NRT2.1* and *GDH2* (Fig. 5c). *ANR1* is found in a cluster of transcription factors identified in this study, *ERF107*, *ARF18* and *BBX16*, which are perturbed in the mutant of *NLP7* and the double mutant of *TGA1/TGA4* (Fig. 5d). Similar transcriptional-regulation feedback on several of the transcription factors characterized in this study (*RAV2*, *ERF107*, *ARF18* and *BBX16*), in addition to previously established nitrogen-status regulators (*LBD38*, *LBD39*, *TGA1*, *NLP7*

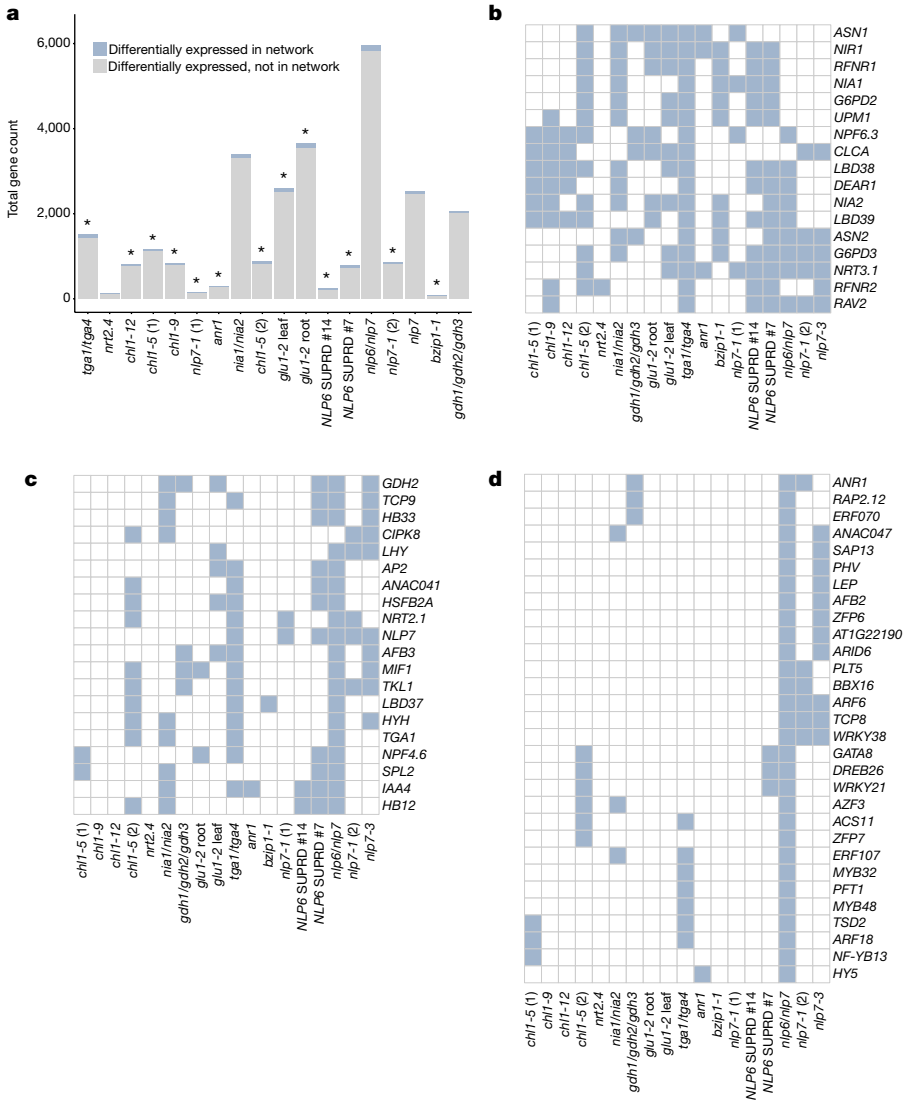


Fig. 5 | Transcriptional feedback upon genetic perturbation of nitrogen metabolism or regulation. a, Bar graph showing differentially expressed genes within the network in different mutant backgrounds. The *y* axis is the number of differentially expressed genes based on genotype (metabolism mutants) or genotype-by-nitrate condition (transcription-factor mutants). An asterisk indicates significance at $P < 0.01$ for enrichment in the network using a two-sided Fisher's exact test (see Methods). $n = 18$ expression datasets. SUPRD #7 and SUPRD #14 refer to dominant-repression mutant lines of the *NLP6* gene. **b**, A core set of nitrogen metabolism genes and regulators that are robust targets of transcriptional feedback. Clusters were determined using *k*-means clustering. Blue, gene significantly differentially expressed; white, gene not significantly differentially expressed using a two-sided test with limma (false discovery rate < 0.05, see Methods). **c**, **d**, Distinct clusters of transcription factors and metabolic enzymes that contain transcription factors found in this study and are targets of feedback by transcriptional regulators of nitrogen metabolism. Each study (see 'Sources for mutant alleles' in Methods) was analysed individually to test the effect of nitrate in the different mutants. For details of the studies from which mutants are derived, see 'Sources for mutant alleles' in Methods.

and *ANR1*), further emphasizes the importance of these transcription factors as central nitrogen regulators.

The YNM indicates the interconnected regulation of nitrogen metabolism: the more important a transcription factor is in regulating growth, the more likely it is to bind to promoters of genes in multiple nitrogen-related categories. The 21 transcription factors we describe here regulate diverse aspects of RSA and shoot development that contribute to how growth is regulated in different nitrogen environments. Transcriptional feedback within the YNM revealed a core set of enzymes involved in nitrogen metabolism, and their regulators. The mechanisms underlying this feedback remain to be determined and may include signalling, metabolite and/or allosteric feedback, or the action of the NPF6.3 transceptor. The identification of these genetically regulated gene expression modules places the genes found in this study within the existing nitrogen-regulatory framework. The transcription factors we identify—in addition to the 'core' set of enzymes involved in nitrogen metabolism—will assist in breeding efforts to generate plants that use nitrogen more efficiently.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-018-0656-3.

Received: 11 December 2017; Accepted: 22 August 2018;
Published online 24 October 2018.

1. Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R. & Polasky, S. Agricultural sustainability and intensive production practices. *Nature* **418**, 671–677 (2002).
2. Sinha, E., Michalak, A. M. & Balaji, V. Eutrophication will increase during the 21st century as a result of precipitation changes. *Science* **357**, 405–408 (2017).
3. Zhang, H. & Forde, B. G. Regulation of *Arabidopsis* root development by nitrate availability. *J. Exp. Bot.* **51**, 51–59 (2000).

4. Castro Marín, I. et al. Nitrate regulates floral induction in *Arabidopsis*, acting independently of light, gibberellin and autonomous pathways. *Planta* **233**, 539–552 (2011).
5. Scheible, W.-R. et al. Genome-wide reprogramming of primary and secondary metabolism, protein synthesis, cellular growth processes, and the regulatory infrastructure of *Arabidopsis* in response to nitrogen. *Plant Physiol.* **136**, 2483–2499 (2004).
6. Liu, K. H. et al. Discovery of nitrate–CPK–NLP signalling in central nutrient-growth networks. *Nature* **545**, 311–316 (2017).
7. Krouk, G. et al. A framework integrating plant growth with hormones and nutrients. *Trends Plant Sci.* **16**, 178–182 (2011).
8. Zhang, H. & Forde, B. G. An *Arabidopsis* MADS box gene that controls nutrient-induced changes in root architecture. *Science* **279**, 407–409 (1998).
9. Gifford, M. L., Dean, A., Gutierrez, R. A., Coruzzi, G. M. & Birnbaum, K. D. Cell-specific nitrogen responses mediate developmental plasticity. *Proc. Natl Acad. Sci. USA* **105**, 803–808 (2008).
10. Rubin, G., Tohge, T., Matsuda, F., Saito, K. & Scheible, W. R. Members of the LBD family of transcription factors repress anthocyanin synthesis and affect additional nitrogen responses in *Arabidopsis*. *Plant Cell* **21**, 3567–3584 (2009).
11. Castaings, L. et al. The nodule inception-like protein 7 modulates nitrate sensing and metabolism in *Arabidopsis*. *Plant J.* **57**, 426–435 (2009).
12. Krouk, G., Mirowski, P., LeCun, Y., Shasha, D. E. & Coruzzi, G. M. Predictive network modeling of the high-resolution dynamic plant transcriptome in response to nitrate. *Genome Biol.* **11**, R123 (2010).
13. Konishi, M. & Yanagisawa, S. *Arabidopsis* NIN-like transcription factors have a central role in nitrate signalling. *Nat. Commun.* **4**, 1617 (2013).
14. Vidal, E. A., Álvarez, J. M. & Gutiérrez, R. A. Nitrate regulation of *AFB3* and *NAC4* gene expression in *Arabidopsis* roots depends on NRT1.1 nitrate transport function. *Plant Signal. Behav.* **9**, e28501 (2014).
15. Alvarez, J. M. et al. Systems approach identifies TGA1 and TGA4 transcription factors as important regulatory components of the nitrate response of *Arabidopsis thaliana* roots. *Plant J.* **80**, 1–13 (2014).
16. Guan, P. et al. Nitrate foraging by *Arabidopsis* roots is mediated by the transcription factor TCP20 through the systemic signaling pathway. *Proc. Natl Acad. Sci. USA*, **111**, 15267–15272 (2014).
17. Medici, A. et al. AtNIGT1/HRS1 integrates nitrate and phosphate signals at the *Arabidopsis* root tip. *Nat. Commun.* **6**, 6274 (2015).
18. Xu, N. et al. The *Arabidopsis* NRG2 protein mediates nitrate signaling and interacts with and regulates key nitrate regulators. *Plant Cell* **28**, 485–504 (2016).
19. Obertello, M., Krouk, G., Katari, M. S., Runko, S. J. & Coruzzi, G. M. Modeling the global effect of the basic-leucine zipper transcription factor 1 (bZIP1) on nitrogen and light regulation in *Arabidopsis*. *BMC Syst. Biol.* **4**, 111 (2010).
20. Araus, V. et al. Members of BTB gene family of scaffold proteins suppress nitrate uptake and nitrogen use efficiency. *Plant Physiol.* **171**, 1523–1532 (2016).
21. Gaudinier, A. et al. Enhanced Y1H assays for *Arabidopsis*. *Nat. Methods* **8**, 1053–1055 (2011).
22. Reece-Hoyes, J. S. et al. Enhanced yeast one-hybrid assays for high-throughput gene-centered regulatory network mapping. *Nat. Methods* **8**, 1059–1064 (2011).
23. Marchive, C. et al. Nuclear retention of the transcription factor NLP7 orchestrates the early response to nitrate in plants. *Nat. Commun.* **4**, 1713 (2013).
24. Ristova, D. et al. Combinatorial interaction network of transcriptomic and phenotypic responses to nitrogen and hormones in the *Arabidopsis thaliana* root. *Sci. Signal.* **9**, rs13 (2016).
25. Tsay, Y. F., Schroeder, J. I., Feldmann, K. A. & Crawford, N. M. The herbicide sensitivity gene *CHL1* of *Arabidopsis* encodes a nitrate-inducible nitrate transporter. *Cell* **72**, 705–713 (1993).
26. Bi, Y.-M. et al. Genetic analysis of *Arabidopsis* GATA transcription factor gene family reveals a nitrate-inducible member important for chlorophyll synthesis and glucose sensitivity. *Plant J.* **44**, 680–692 (2005).
27. Little, D. Y. et al. The putative high-affinity nitrate transporter NRT2.1 represses lateral root initiation in response to nutritional cues. *Proc. Natl Acad. Sci. USA* **102**, 13693–13698 (2005).
28. Remans, T. et al. The *Arabidopsis* NRT1.1 transporter participates in the signaling pathway triggering root colonization of nitrate-rich patches. *Proc. Natl Acad. Sci. USA* **103**, 19206–19211 (2006).
29. De Smet, I. et al. Receptor-like kinase ACR4 restricts formative cell divisions in the *Arabidopsis* root. *Science* **322**, 594–597 (2008).
30. Wang, R., Guegler, K., LaBrie, S. T. & Crawford, N. M. Genomic analysis of a nutrient response in *Arabidopsis* reveals diverse expression patterns and novel metabolic and potential regulatory genes induced by nitrate. *Plant Cell* **12**, 1491–1509 (2000).

Acknowledgements We thank N. M. Crawford for *chl1-5* seeds, P. J. Etchells for *wox14-1* and *lbd4-1* seeds, and E. E. Sparks and P. N. Benfey for *erf107-1*, *abf4-2*, *eel-1*, *vip1-1* and *erf070* seeds. Some seed stocks were obtained from the *Arabidopsis* Biological Resource Center (ABRC) at Ohio State University. We thank E. A. Ainsworth and S. B. Gray for help with chlorophyll and protein assays, K. Kajala for help with RNA-seq libraries and E. M. McGinnis for help with root measurements. We thank K. Dehesh for discussions. This research was funded by DuPont Pioneer. A.G. was also supported by the Elsie Taylor Stocking Memorial Fellowship, the Katherine Esau Graduate Summer Fellowship and the University of California, Davis Dissertation Year Fellowship. J.R.-M. was supported by a UC-MEXUS CONACYT PhD Fellowship. D.J.K., M.T. and S.M.B. acknowledge funding from NSF-MCB-1330337. S.M.B. was partially funded by an HHMI Faculty Scholar Fellowship.

Reviewer information *Nature* thanks M. Bennett, C. Hodgman and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions S.M.B., B.S. and D.W. conceived the project. A.G., L.Z., J.F., S.A. and M.T. cloned promoters. A.G. and S.M.B. designed experiments and A.G., D.J.K. and S.M.B. contributed to data analysis experimental design. A.G., J.F., A.-M.B., M.T. and B.L. performed enhanced yeast one-hybrid screens. A.G. genotyped plants. A.G., A.-M.B. and M.T. performed plant phenotyping. A.G., J.R.-M., A.O. and C.L.-M. performed bioinformatics. A.G. performed transcription factor–target correlation analysis. C.L.-M. and A.O. performed NeCorr analysis. J.R.-M. performed network analysis (enrichment tests), analysis of RNA sequencing data, clustering and network-metabolic analysis. A.G., S.M.B., D.E.R., D.J.K., B.S., D.W. and M.J.F. provided discussion, experimental design and analysis suggestions. S.M.B. and A.G. wrote the manuscript.

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0656-3.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0656-3.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to S.M.B.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

No statistical methods were used to predetermine sample size. Seeds were randomized within each experiment and investigators were blinded to allocation during experiments and outcome assessment.

Promoter cloning, yeast transformation and yeast one-hybrid assays. Gene promoters were cloned to 2 kb or until the nearest upstream gene or synthesized by Life Technologies (Supplementary Table 2a). In the case of cloning, promoters were amplified from Col-0 genomic DNA using Phusion Taq polymerase (NEB). Promoters were recombined into 5′TOPO (Invitrogen), fully sequenced and then recombined into pMW2 and pMW3[31] using LR clonase II (Invitrogen). pMW2 and pMW3 constructs were sequence-confirmed. They were transformed into the yeast strain YM4271 as previously described[32]. If constructs were resistant to transformation, they were transformed into the Y1H-S2 strain[22]. Yeast colonies were screened for autoactivation and construct presence. Promoter strains were mated against transcription-factor strains as previously described[21,22].

Transcription-factor cloning and yeast transformation. Transcription factors were cloned from root RNA extracted using the RNeasy Kit (Qiagen) (Supplementary Table 2b). Coding sequences were amplified using Phusion Taq polymerase (NEB). Transcription factors were recombined into D-TOPO (Invitrogen), fully sequenced and then recombined into pDEST-AD2μ using LR clonase II (Invitrogen). They were transformed into the yeast strain Yα1867 as previously described[21,22].

Network construction. Networks were made using Cytoscape v.3.2.0[33]. All cytoscape network files can be found at https://github.com/agaudinier/Gaudinier2018.

Figure construction. Figures were made using Cytoscape, and ggplot2[34] v.3.0.0 in R. Figures were compiled using Inkscape (http://www.inkscape.org).

Plant material and growth conditions. Transfer DNA (tDNA) mutant lines were obtained through TAIR (http://www.arabidopsis.org) or collaborators. Seeds sorted between 250–300 μm were surface-sterilized using dichloroisocyanuric acid solution (0.9% (w/v) dichloroisocyanuric acid solution (10% water, 90% ethanol), then rinsed twice in 95% ethanol, and then dried completely). For the root mutant phenotyping experiment (Supplementary Table 2c), sets of four tDNA lines and a Col-0 control were plated in a random block design on a minimum of twelve 1-mM KNO_3 and twelve 10-mM KNO_3 medium plates, and stratified at 4 °C for two nights. Medium components: 1 or 10 mM KNO_3, 4 mM $MgSO_4$, 2 mM KH_2PO_4, 1 mM $CaCl_2$, 10 mM KCl, 36 mg/l FeEDTA, 0.146 g/l 2-morpholinoethane sulfonic acid, 1.43 mg/l H_2BO_3, 0.905 mg/l $MnCl_2{\cdot}4H_2O$, 0.055 mg/l $ZnCL_2$, 0.025 mg/l $CuCl_2{\cdot}2H_2O$, 0.0125 $Na_2MoO_4{\cdot}2H_2O$, 1% sucrose, 0.75% phytagel, pH 5.7.

For the shoot phenotyping experiment, sorted seeds were stratified at 4 °C for two nights and sown on Sunshine Mix soil in flats containing 18 pots. Seventeen genotypes, plus Col-0, were randomized in a partial random block design for 8 or 9 biological replicates per experiment for a total of three experiments. Plants were watered twice a week, switching between a modified Hoagland's solution and deionized water. Modified Hoagland's solution components [16×]: 1.6 g/l KNO_3, 0.55 g/l KH_2PO_4, 3.85 g/l $MgSO_4$, 3.57 g/l KCl, 2.35 g/l $CaCl_2$, 1.34 g/l Sprint 330, 2.97 mg/l H_3BO_3, 3.17 mg/l $MnCl_2{\cdot}4H_2O$, 4.6 mg/l $ZnSO_4{\cdot}7H_2O$, 0.4 mg/l $CuSO_4{\cdot}5H_2O$, 0.39 mg/l $H_2MoO_4{\cdot}H_2O$, pH 5.5.

For the RNA-sequencing (RNA-seq) experiment to characterize gene expression in each mutant background, 200–300 seeds per plate were sown on Petri plates with medium containing 1 mM or 10 mM KNO_3 and nylon mesh, and stratified for two nights at 4 °C. Two plates of seedlings per genotype were grown and combined for each biological replicate. Four biological replicates were grown per genotype and treatment. Roots of 9-day-old seedlings were collected from 6–7 h after sunrise and immediately frozen in liquid N_2.

RNA-seq library preparation and pooling of technical replicates. RNA-seq libraries were prepared following the BRAD-Seq DGE protocol[35]. Libraries were sequenced using the Ilumina HiSeq 3000 in SR50 mode. Two technical replicate libraries were created from each RNA sample and after assessing sufficient reproducibility, counts across technical replicates were pooled together. Pooling was performed by summing the counts for the same gene across equivalent replicates. The merged file was subjected to the same quality processing. The number of mapped reads for each biological replicate and correlation of replicates are found in Supplementary Table 14c, d.

RNA-seq read processing and differential expression analysis. Before and after read processing, libraries were analysed with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to assess the quality of the sequences. We trimmed barcodes from raw reads using fastx-trimmer (http://hannonlab.cshl.edu/fastx_toolkit/index.html) with parameters: -f 9 -v -Q 33. This was followed by adaptor trimming and quality filtering was using reaper, from the Kraken Suite[36] with options: -geom no-bc -dust-suffix-late 10/ACTG -dust-suffix 10/ACTG–noqc -nnn-check 1/1 -qqq-check 33/10 -clean-length 30 -tri 20 -polya 5–bcq-late. Trimmed reads were mapped to the reference genome of *A. thaliana* (TAIR 10) using bowtie (-a–best–strata -n 1 -m 1 -p 4–sam–tryhard) with subsequent conversion to BAM format using samtools[37]. HTSeq-count was used to obtain raw counts[38].

Differential gene expression analysis was done using limma[39] in R/Bioconductor[40], with empirical weights estimated for each observation using the voomWithQualityWeights function. Quantile normalization was used to account for different RNA inputs and library sizes. The linear model for each gene was specified as: log(counts per million) of a particular gene = mutant + treatment + mutant:treatment. Specific contrasts were constructed to compare each mutant to the control, and each genotype × treatment interaction. Differentially expressed genes were selected based on a false discovery rate < 0.05 (Supplementary Table 15).

Generation of the publically available gene expression profiling datasets for nitrogen-responsive genes. We compiled a comprehensive dataset of publically available gene-expression responses of wild-type plants in response to nitrogen availability in both the root and shoot (GEO accession GSE18984)[9,10,12,15,19,41–45], as well as profiling of nitrogen-status gene expression changes in specific root cell types[9]. Data from ATH1 affymetrix arrays were downloaded from the NCBI GEO database[46] and imported into R using the affy[47] package in Bioconductor. Arrays were normalized using the robust multi-array average (RMA) method. Gene expression was averaged across biological replicates, and then treatment datasets were expressed relative to their appropriate controls (Supplementary Table 5). Pearson and Spearman correlations were calculated in R for the treatment and cell-type-specific datasets for all transcription factor–target pairs (Supplementary Table 6).

Spearman rank correlation analysis of root and shoot phenotypes relative to network connectivity and related metrics. We prioritized transcription factors from the YNM with a Pearson or Spearman rank correlation greater than ±0.5 (for the nitrogen treatment dataset) or greater than ±0.8 (for the cell-type-specific dataset) with their target genes (Supplementary Table 6). Spearman rank correlations were calculated in R using rcorr() from the Hmisc package (https://cran.r-project.org/package=Hmisc) for the phenotype traits of transcription-factor mutants, relative to network connectivity and correlation with targets. Data and correlations can be found in Supplementary Table 13.

Generation of the publically available dataset for nitrogen-metabolism mutants and mutants of transcription factors associated with nitrogen metabolism. Affymetrix arrays were read using the affy package in Bioconductor. Agilent and Complete *Arabidopsis* Transcriptome Micro Array (CATMA) arrays were read with the read.maimages() function from limma; the source option was set to 'agilent' for the former and 'genepix' for the latter. After arrays were read, limma was used for downstream processing, normalization and differential expression analysis. In brief, Affymetrix arrays were normalized using the RMA method. Agilent and CATMA arrays were subjected to background correction and normalization using the functions backgroundCorrect() and normalizeBetweenArrays(). After normalization and filtering, differential expression was analysed using the standard limma approach.

NECorr. The starting hypothesis of NECorr is that an important interaction for a stimuli response is that of a regulator acting on one or several hub genes. Hence, hub genes will propagate the systemic cascade appropriate to the stimuli. Thenceforth the dynamics of the molecular network will evolve. This approach ranks transcription factors given several network metrics including betweenness centrality, degree distribution and as a function of their gene-expression similarity[48].

Hub calculation. The first step is a heuristic model, which merges molecular network topology and gene-expression data. NECorr-Hub is a linear model including five parameters: condition or tissue specificity of gene expression, co-expression of interactions across conditions, and the molecular network centralities betweenness, connectivity and transitivity. The rank given to each of these parameters was decided empirically.

Both genes of an interaction pair need to be co-expressed in most of the tissues and/or conditions, which shows that they can influence each other. Correlated gene expression was considered as the highest-ranking parameter, followed by gene-expression specificity in the studied tissue or condition. In addition, a high level of connectivity of the gene in the molecular network is required to generate a proper response. Connectivity can be defined in several manners: betweenness, degree connectivity and transitivity were chosen as the most meaningful centralities to define gene importance as a hub.

Based on the ranking, each parameter weight was estimated using the analytic hierarchy process (AHP)[49,50], a multiple-criteria decision analysis method. The AHP is applied through the R package pmr (https://cran.r-project.org/package=pmr). The importance of the five parameters is generated by pairwise comparisons. Hence, this leads to an adjacency matrix of pairwise weight importance. From this adjacency matrix, Eigen vectors are calculated to assign a weight to each parameter. The AHP method is applied as follows. Each gene is ranked for the five parameters above. Each ranked parameter is standardized in values between

0 and 1 (*z*-score), to obtain data with the same scale. For each tissue and/or condition, the parameter weights are applied as factors of a linear model that is used to prioritize hub genes.

For condition 1: $\text{Ranking}_{\text{condition1}} = w_1 \times \text{IntSig} + w_2 \times \text{TS} + w_3 \times \text{BetC} + w_4 \times \text{Cot} + w_5 \times \text{Trs}$, in which IntSig represents the interaction significance (co-expression significance in the interaction involving the gene), TS represents tissue specificity (selectivity), BetC represents betweenness centrality, Cot represents connectivity centrality and Trs represents transitivity. The weights are defined as $w_1 = w_2 > w_3 > w_4 > w_5$.

To rank the interactions in the molecular network for a given condition, the average ranking of the two genes defining this edge is taken. When several conditions are evaluated, the gene ranking between conditions can be done by averaging each condition ranking.

NECorr-Hub parameter estimation. Molecular network topology centralities are obtained using the R package iGraph (https://cran.r-project.org/package=igraph). Co-expression analysis (or the significance of each interaction) was estimated using a Rcpp script to evaluate the Gini correlation coefficient related to each interaction. The Gini correlation coefficient was previously shown to be an effective method for detecting transcription-factor activity[51]. The co-expression significance for each gene is evaluated by averaging the magnitude of the correlation from all the interactions containing this particular gene using Fisher's method[52,53].

The genes with tissue and/or condition specificity (or selectivity) are detected using the intersection-inion test (IUT) with a relaxed threshold (raw $P = 0.5$) (https://cran.r-project.org/package=igraph)[54]. The tissue and/or condition-selective genes or tissue and/or condition-excluded genes are assigned for each specific tissue and/or condition within a set of samples. These genes attributed to a tissue and/or condition are fuzzy owing to the low selection threshold of the gene in IUT; a gene could therefore appear in a different tissue or condition as selective or excluded. Second, these selected genes are ranked for their tissue selectivity or exclusion using the tissue specificity index[55]. We define both a positive and negative TSI:

$$\text{positive TSI} = \frac{\sum_{i=1}^{N}(1 - x_i/x_{\max})}{N-1}$$

The negative tissue specific index measures the extent to which a gene is excluded from a tissue or condition:

$$\text{negative TSI} = \frac{\sum_{i=1}^{N}(x_i - x_{\min}/x_{\max})}{N-1}$$

The results for TSI measurements are merged to obtain a ranking of all the tissue and/or condition-selective or -excluded genes defined from IUT test.

NeCorr rankings can be found in Supplementary Table 7.

Code. The NECorr source code is maintained in GitHub: https://github.com/warelab/NECorr.

Mutant line selection. The mutant lines acquired for this study represent most of the top-ranked and intermediate-ranked genes that were deemed interesting for having important binding targets (Supplementary Tables 10, 11).

Root phenotyping data collection and analysis. Traits measured included PRL, LR and LRL. Additionally, composite traits were considered, including total root length (TRL = PRL + LRL), average lateral root length (ALRL = LRL/LR), lateral root density (LRD = LR/PRL) and the percentage of LRL contributing to TRL (LRL/TRL) as well as the partitioning of variation across these mutants relative to wild type, using principal component analysis of all RSA traits[56] (Supplementary Data 1).

Plates were scanned using the V750 scanner. Primary roots and lateral roots of 9-day-old seedlings were traced using a Wacom Bamboo tablet in ImageJ. Data were log-transformed and analysed using ANOVA in R. Using a two-way ANOVA, three phenotypic categories were considered: genotype effects in both nitrogen conditions (genotype-dependent), genotype effects in only one condition (nitrogen-condition-dependent) or genotype by nitrogen condition-dependent effects ($P < 0.05$, Supplementary Data 2). The extent to which lateral root traits are uncoupled from PRL is not clear[57], thus an additional ANOVA model was used that included PRL as a factor—with the hypothesis that lateral root emergence or elongation may be dependent on PRL. As expected, composite traits extracted from the principal component analysis were significantly correlated with a number of RSA traits ($P < 0.05$, Extended Data Fig. 4, Supplementary Table 11). The scripts for analyses can be found at https://github.com/agaudinier/Gaudinier2018. All ANOVA tables can be found in Supplementary Data 2. A summary of the statistics can be found in Supplementary Table 13.

Principal component analysis. Mutant and wild-type controls were plotted in R using the prcomp() function. The loadings for each principal component (PC1–PC3) in the mutant and wild-type sets were analysed using ANOVA in R. The script for the analysis can be found at https://github.com/agaudinier/Gaudinier2018. All ANOVA tables can be found in Supplementary Data 2.

Shoot phenotyping data collection and analysis. Plants were photographed at 15 and 22 days old. ImageJ was used to analyse rosette size. Bolting and flowering days were recorded. Rosette-size data were log-transformed, and for bolting and flowering day a reciprocal transformation was used and analysed using a two-way ANOVA in R. All ANOVA tables can be found in Supplementary Data 2. A summary of the statistics can be found in Supplementary Table 13.

Chlorophyll extraction and analysis. Full rosette leaves were measured for their chlorophyll content index using the CCM-200 plus (Opti-Sciences). Chlorophyll measurements were done by collecting supernatants of discs from *nlp7-1*, *chl1-5* and Col-0 (control) leaves extracted in two extractions of 80% HEPES-buffered ethanol heated to 80 °C and one extraction of 50% HEPES-buffered ethanol. Absorbance for the supernatant was measured at 652 and at 665 nm. Total chlorophyll was calculated as chlorophyll = $22.12\,A_{650} + 2.71\,A_{665}$, according to a previously published method[58].

Quantification of ^{15}N and ^{13}C abundance. Rosettes of 20-day-old plants were collected and dried at 60 °C for two days. Dried rosettes were homogenized. Samples of 0.7–3 mg were submitted to the Stable Isotope Facility at University of California at Davis for analysis of natural abundance levels of ^{15}N and ^{13}C using an elemental analyser with a continuous flow isotope ratio mass spectrometer.

YNM network analyses. *Genotype expression and expression dependent on genotype-by-nitrate condition.* To test for the presence of genes in the YNM that are significantly differentially expressed, a Fisher's exact test was used to test for enrichment in R, using the standard function fisher.test(). For this, we queried whether the overlap of YNM-predicted genes that overlap with differentially expressed genes was greater than differentially expressed genes that did not overlap with YNM genes; as a background, we used genes that were not differentially expressed. We performed this test for every contrast, which means that the groups of genes changed between each test but the absolute number of total genes remained the same.

To test for enrichment of various pathways, a list of CPK–NLP7-dependent genes[6], a list of primary-root developmental genes (Supplementary Table 12) and lateral-root developmental genes[29], and a list of hormone-responsive genes[59] were queried. A Fisher's exact test was used to determine whether the proportions of genes from these datasets were enriched in the YNM. The background for the CPK–NLP7 test includes all genes in the *Arabidopsis* genome that are not part of the YNM. The background for the root development and hormone tests includes all genes on the ATH1 affymetrix microarray that are not part of the YNM.

Transcriptional feedback of nitrogen metabolism enzymes and regulators. To test whether any feedback is present within the YNM, we curated whole-genome expression datasets in mutants of nitrate transporters or a transceptor, or metabolic enzymes (GEO accession GSE10786)[44,60–63] (NPF6.3, the double mutant of NIA1/NIA2, GLU1, NRT2.4 and the triple mutant of GDH1/GDH2/GDH3 (Supplementary Table 16)). The expression of genes in mutants of previously described transcriptional regulators of nitrogen metabolism (ANR1, NLP6, NLP7, BZIP1 and the double mutants of TGA1/TGA4 and NLP6/NLP7) that show changes in gene expression dependent on genotype-by-nitrogen condition—relative to wild type—in a genotype-by-condition analysis was also considered (GEO accession GSE6824)[13,15,19,23] (Supplementary Table 16).

Enrichment was calculated as above. We tested whether the overlap between differentially expressed genes within the YNM was greater than differentially expressed genes which did not overlap with genes in the YNM. Each microarray study was analysed independently.

Clustering analysis of transcriptional feedback on YNM. *k*-means clustering was performed using the presence or absence calls of significantly differentially expressed genes that were differentially expressed in at least one of the contrasts; if a gene predicted by the network was significantly differentially expressed, we assigned it a value of 1, and if it was not significantly differentially expressed we assigned it a value of 0. Genes with 0 across all contrasts were not considered for the analysis (not differentially expressed in any condition). In short, the algorithm used was to calculate the Euclidean distance of the binary matrix (dist function in R), then to obtain the principal components of this distance using a correlation matrix (princomp function in R with cor = *T*) and select the scores for the first two components. We then calculated the clusters based on these two first principal components, and a selected value for *k*.

The number of clusters (*k*) was selected by analytical and empirical analysis: the 'elbow method' looks at different values for *k* and their relationship with the within-cluster sum of squares, in which the optimal value is that at which the line starts to plateau. We then tested different values for *k* within a threshold given by the elbow method and selected that in which biologically relevant clusters were observed.

Identification of the dominant pattern in transcription-factor mutants of nitrogen-responsive genes in the root. Data in each mutant background was filtered for genes that were significantly differentially expressed owing to nitrogen treatment in wild-type plants (wild type 1 mM versus wild type 10 mM)

(Supplementary Table 18). The expression of each of these genes was obtained by taking the difference in log(FC) between the effect of nitrogen on each of the mutants (that is, *arf18-3*1 mM versus *arf18-3* 10 mM) and the wild type (wild type 1 mM versus wild type 10 mM). Dominant patterns of expression were then identified using a previously published algorithm[64] (with parameters set as follows: minExpFilter = FALSE; minVarFilter = FALSE; fuzzyKmemb = 1.04; already-Log2 = TRUE). The choice of number of clusters was set to kChoice = 7. Clustering of genes, the expression of which changes upon variation in nitrogen availability in the wild-type root, revealed that these mutants have similar perturbations in nitrogen-associated gene regulation (Extended Data Fig. 9).

Sources for mutant alleles. The sources for the mutant alleles displayed in Fig. 5 are as follows: *tga1/tga4* (ref. [15]); *nrt2.4* (GEO accession GSE10786); *chl1-12*, *chl1-5* (1) and *chl1-9* (ref. [60]); *nlp7-1* (1) (ref. [11]); *anr1* (GEO accession GSE6824); *nia1/nia2* (ref. [61]); *chl1-5* (2) (ref. [44]); *glu1-2* leaf and *glu1-2* root (ref. [62]); *NLP6* SUPRD #7 and *NLP6* SUPRD #14 (ref. [13]); *nlp7-1* (2), *nlp7-3* and *nlp6/nlp7* (ref. [23]); *bzip1-1* (ref. [19]); and *gdh1/gdh2/gdh3* (ref. [63]).

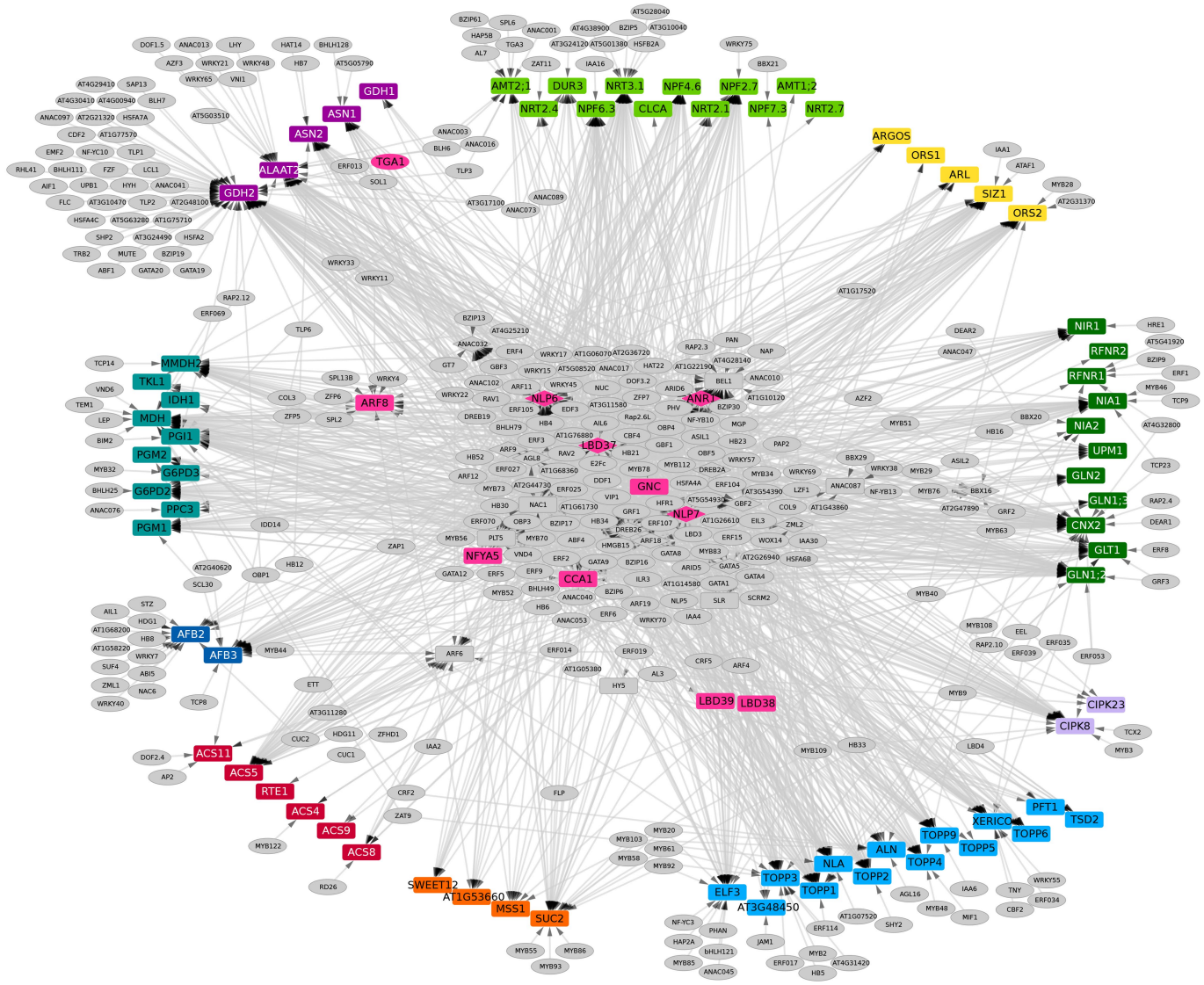
Code availability. Code for plant phenotyping analysis can be found at https://github.com/agaudinier/Gaudinier2018. Code for NeCorr analysis can be found at https://github.com/warelab/NECorr.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

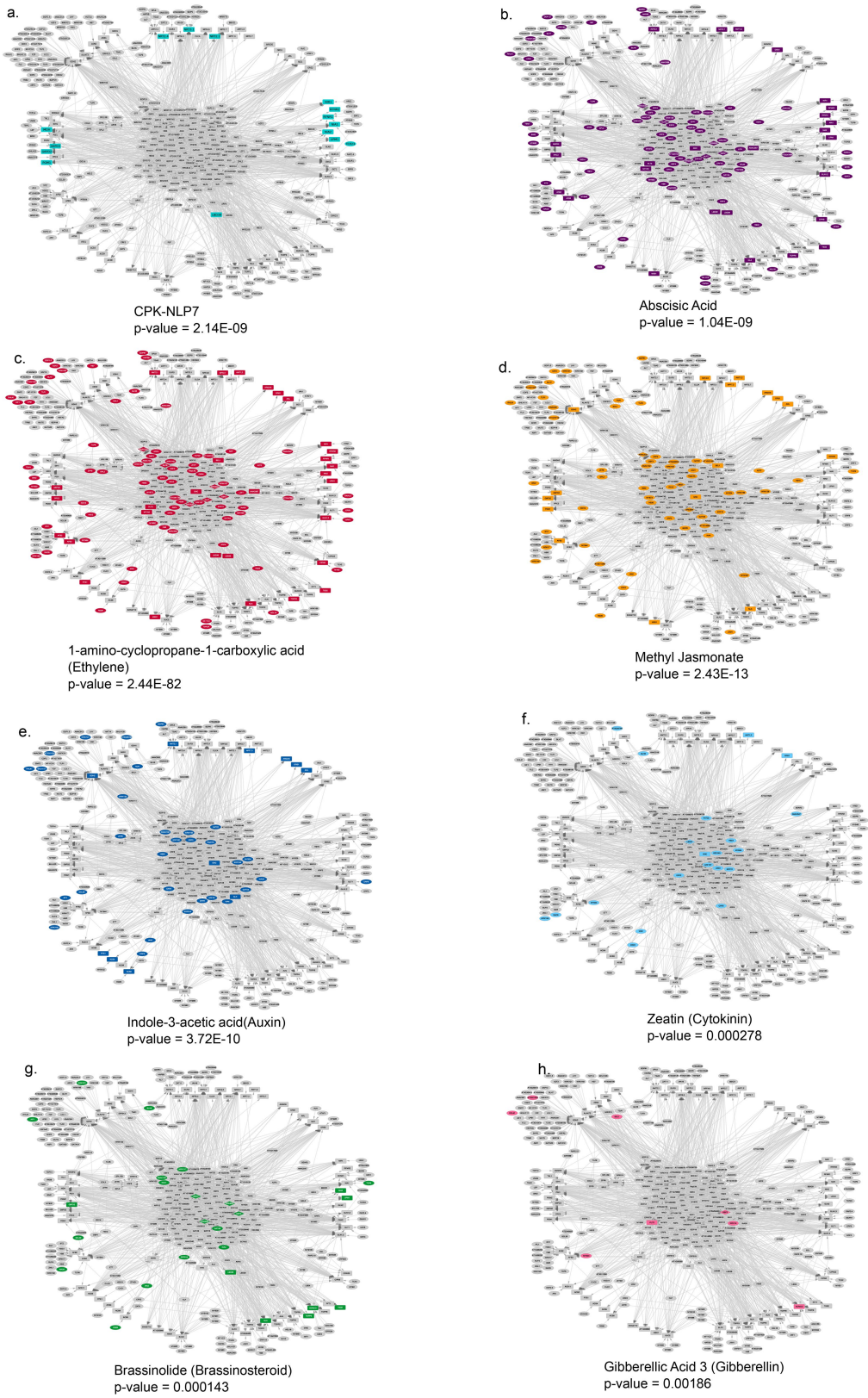
Data availability

RNA sequencing data that support the findings of this study have been deposited in NCBI with the primary accession code GSE107988. Supplementary Tables, R code and Cytoscape files can be found at: https://www.bradylab.org/resources/ or https://github.com/agaudinier/Gaudinier2018.

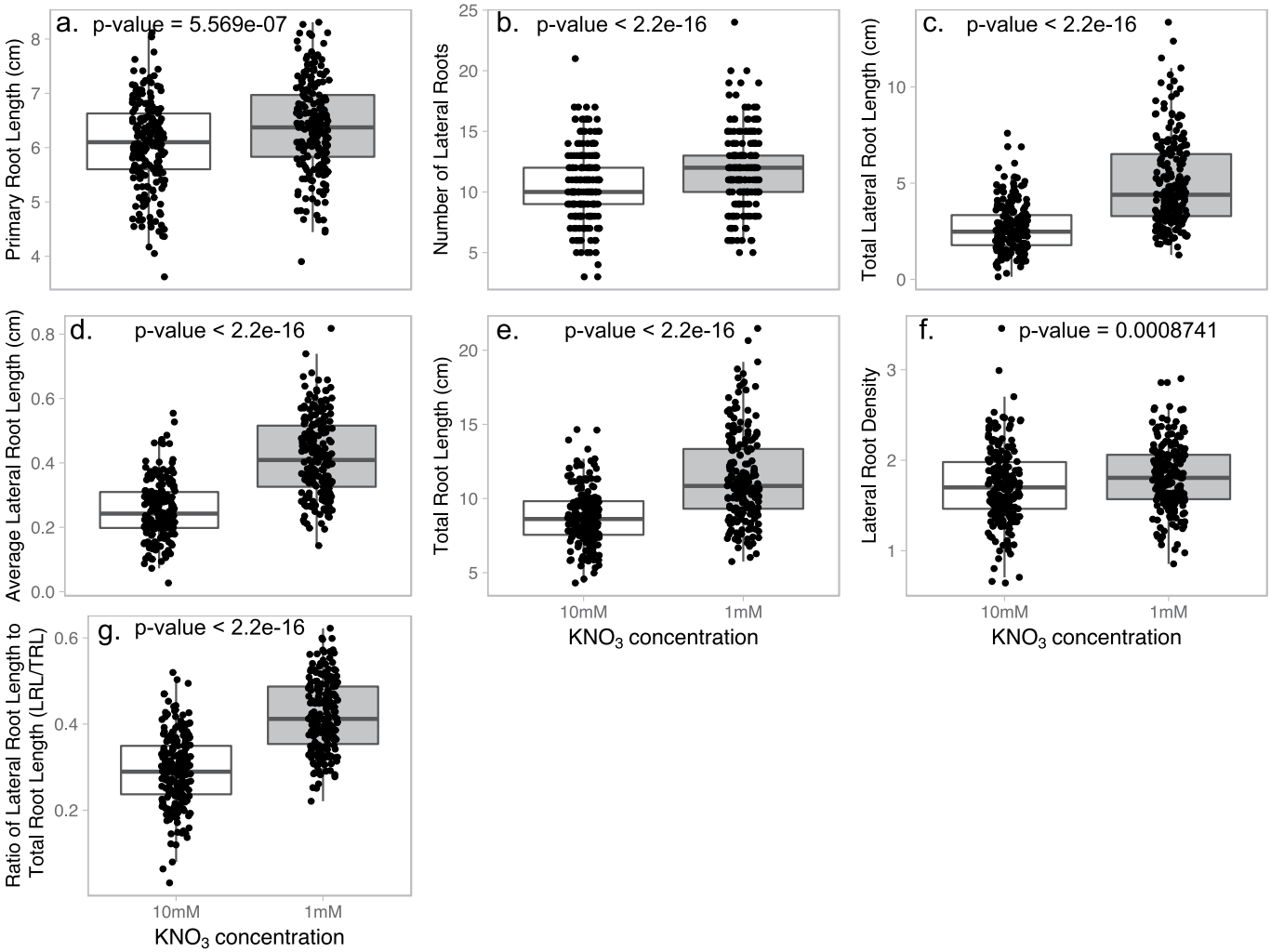
31. Deplancke, B., Vermeirssen, V., Arda, H. E., Martinez, N. J. & Walhout, A. J. Gateway-compatible yeast one-hybrid screens. *Cold Spring Harb. Protoc.* **2006**, https://www.doi.org/10.1101/pdb.prot4590 (2006).
32. Deplancke, B., Dupuy, D., Vidal, M. & Walhout, A. J. A gateway-compatible yeast one-hybrid system. *Genome Res.* **14**, 2093–2101 (2004).
33. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
34. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* 2nd edn (Use R!) (Springer, Basel, 2016).
35. Townsley, B. T., Covington, M. F., Ichihashi, Y., Zumstein, K. & Sinha, N. R. BrAD-seq: breath adapter directional sequencing: a streamlined, ultra-simple and fast library preparation protocol for strand specific mRNA library construction. *Front. Plant Sci.* **6**, 366 (2015).
36. Davis, M. P., van Dongen, S., Abreu-Goodger, C., Bartonicek, N. & Enright, A. J. Kraken: a set of tools for quality control and analysis of high-throughput sequence data. *Methods* **63**, 41–49 (2013).
37. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
38. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
39. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
40. Huber, W. et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* **12**, 115–121 (2015).
41. Gutiérrez, R. A. et al. Qualitative network models and genome-wide expression data define carbon/nitrogen-responsive molecular machines in *Arabidopsis*. *Genome Biol.* **8**, R7 (2007).
42. Gutiérrez, R. A. et al. Systems approach identifies an organic nitrogen-responsive gene network that is regulated by the master clock control gene CCA1. *Proc. Natl Acad. Sci. USA* **105**, 4939–4944 (2008).
43. Patterson, K. et al. Distinct signalling pathways and transcriptome response signatures differentiate ammonium- and nitrate-supplied plants. *Plant Cell Environ.* **33**, 1486–1501 (2010).
44. Hu, H.-C., Wang, Y.-Y. & Tsay, Y.-F. AtCIPK8, a CBL-interacting protein kinase, regulates the low-affinity phase of the primary nitrate response. *Plant J.* **57**, 264–278 (2009).
45. Vidal, E. A., Moyano, T. C., Riveras, E., Contreras-López, O. & Gutiérrez, R. A. Systems approaches map regulatory networks downstream of the auxin receptor AFB3 in the nitrate response of *Arabidopsis thaliana* roots. *Proc. Natl Acad. Sci. USA* **110**, 12840–12845 (2013).
46. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–210 (2002).
47. Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307–315 (2004).
48. Liseron-Monfils, C. V., Olson, A. & Ware, D. NECorr, a tool to rank gene importance in biological processes using molecular networks and transcriptome data. Preprint at https://www.biorxiv.org/content/early/2018/05/21/326868 (2018).
49. Saaty, T. L. A scaling method for priorities in hierarchical structures. *J. Math. Psychol.* **15**, 234–281 (1977).
50. Saaty, T. L. Principles of the analytic hierarchy process. *Expert Judgment Expert Systems* **35**, 27–73 (1987).
51. Ma, C. & Wang, X. Application of the Gini correlation coefficient to infer regulatory relationships in transcriptome analysis. *Plant Physiol.* **160**, 192–203 (2012).
52. Poole, W., Gibbs, D. L., Shmulevich, I., Bernard, B. & Knijnenburg, T. A. Combining dependent *P*-values with an empirical adaptation of Brown's method. *Bioinformatics* **32**, i430–i436 (2016).
53. Fisher, R. A. Combining independent tests of significance. *Am. Stat.* **2**, 30 (1948).
54. Van Deun, K. et al. Testing the hypothesis of tissue selectivity: the intersection-union test and a Bayesian approach. *Bioinformatics* **25**, 2588–2594 (2009).
55. Yanai, I. et al. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* **21**, 650–659 (2005).
56. Chitwood, D. H. & Topp, C. N. Revealing plant cryptotypes: defining meaningful phenotypes among infinite traits. *Curr. Opin. Plant Biol.* **24**, 54–60 (2015).
57. Gruber, B. D., Giehl, R. F., Friedel, S. & von Wirén, N. Plasticity of the *Arabidopsis* root system under nutrient deficiencies. *Plant Physiol.* **163**, 161–179 (2013).
58. Porra, R. J, Thompson, W. A. & Kriedemann, P. E. Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochim. Biophys. Acta Bioenerg.* **975**, 384–394 (1989).
59. Nemhauser, J. L., Hong, F. & Chory, J. Different plant hormones regulate similar processes through largely nonoverlapping transcriptional responses. *Cell* **126**, 467–475 (2006).
60. Bouguyon, E. et al. Multiple mechanisms of nitrate sensing by *Arabidopsis* nitrate transceptor NRT1.1. *Nat. Plants* **1**, 15015 (2015).
61. Gibbs, D. J. et al. Nitric oxide sensing in plants is mediated by proteolytic control of group VII ERF transcription factors. *Mol. Cell* **53**, 369–379 (2014).
62. Kissen, R. et al. Transcriptional profiling of an *Fd-GOGAT1/GLU1* mutant in *Arabidopsis thaliana* reveals a multiple stress response and extensive reprogramming of the transcriptome. *BMC Genomics* **11**, 190 (2010).
63. Fontaine, J.-X. et al. Characterization of a NADH-dependent glutamate dehydrogenase mutant of *Arabidopsis* demonstrates the key role of this enzyme in root carbon and nitrogen metabolism. *Plant Cell* **24**, 4044–4065 (2012).
64. Orlando, D. A., Brady, S. M., Koch, J. D., Dinneny, J. R. & Benfey, P. N. Manipulating large-scale *Arabidopsis* microarray expression data: identifying dominant expression patterns and biological process enrichment. *Methods Mol. Biol.* **553**, 57–77 (2009).



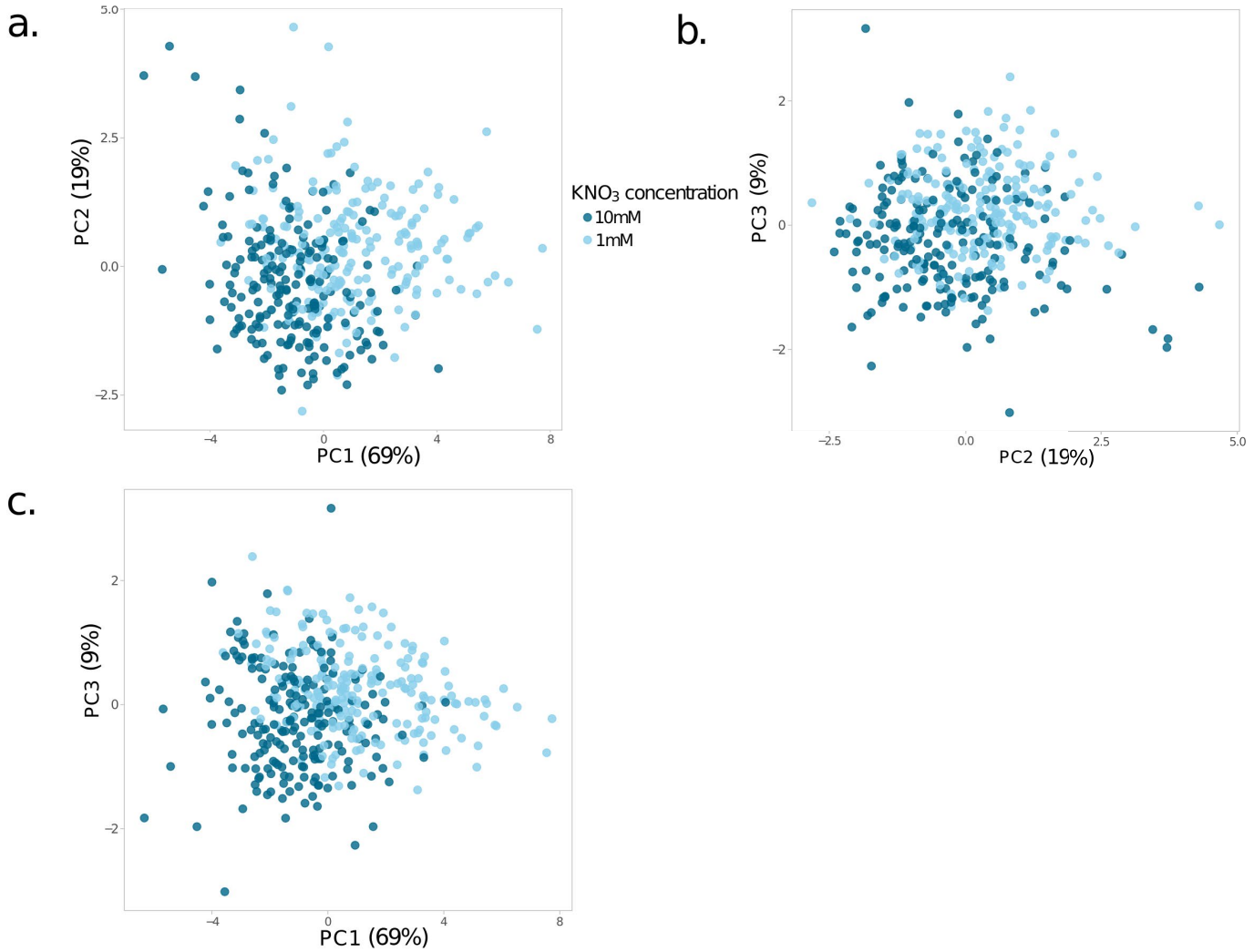
Extended Data Fig. 1 | Combinatorial interactions between transcription factors and promoters of genes associated with nitrogen metabolism, signalling and nitrogen-associated processes. Rectangles, promoters; ovals, transcription factors; diamonds, genes represented as both promoters and transcription factors. Nitrogen-associated biological processes are indicated by promoter colour. A grey line indicates a transcription factor–promoter interaction. Light green, nitrogen transporter; yellow, organ growth; dark green, nitrate assimilation; light purple, nitrogen signalling; light blue, nitrogen-linked; orange, carbon metabolism; red, ethylene; dark blue, auxin; teal, carbon transporter; dark purple, amino acid metabolism; pink, transcription factors linked to nitrogen.



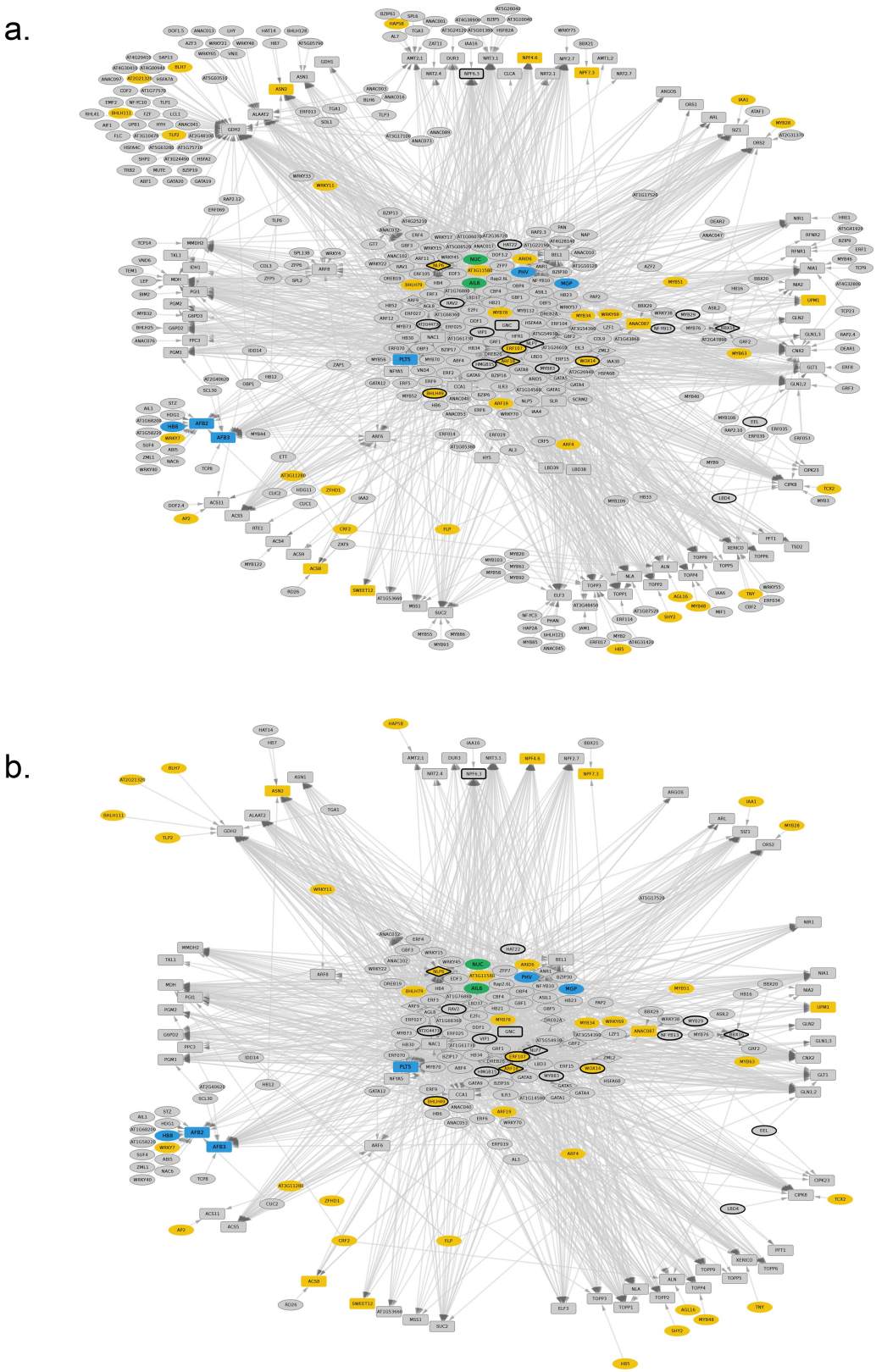
Extended Data Fig. 2 | Genes in the YNM regulated by hormone signalling. The YNM. Genes coloured in each panel are regulated by the CPK–NLP7 signalling cascade or indicated hormone. *P* value indicates significance for enrichment in the network using a two-sided Fisher's exact test. **a**, Genes regulated by the CPK–NLP7 signalling cascade (cyan). **b**, Genes regulated by abscisic acid (purple). **c**, Genes regulated by ethylene (red). **d**, Genes regulated by methyl jasmonate (orange). **e**, Genes regulated by auxin (dark blue). **f**, Genes regulated by cytokinin (light blue). **g**, Genes regulated by brassinosteroid (green). **h**, Genes regulated by gibberellic acid (pink). Gene lists used for enrichment tests can be found in Supplementary Table 4.



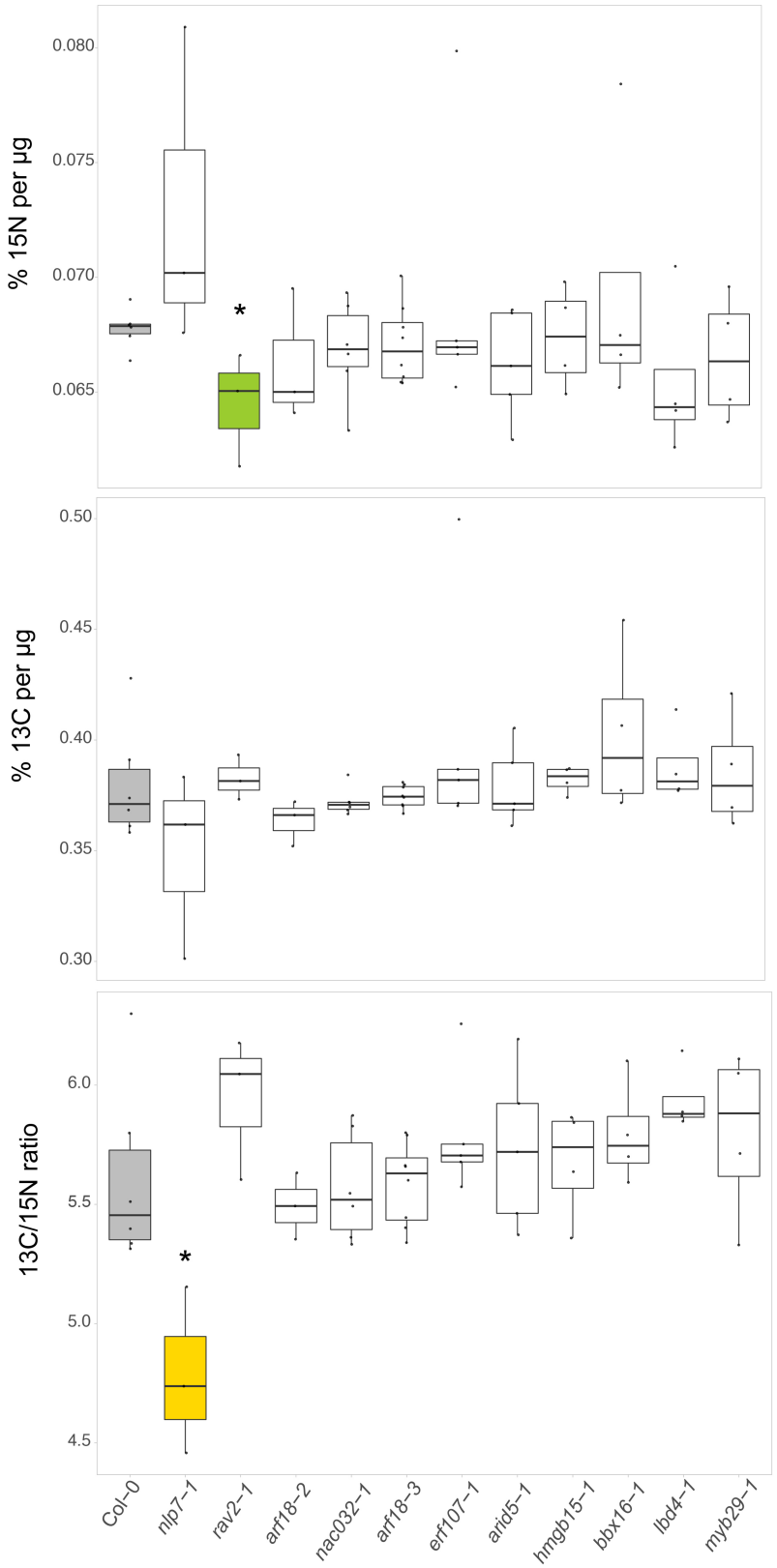
Extended Data Fig. 3 | Wild-type root growth. RSA for wild-type (Col-0) nine-day-old seedlings in both limiting (1 mM) and sufficient (10 mM) KNO_3 conditions. **a**–**g**, Traits measured were primary root length (**a**), number of lateral roots (**b**), total lateral root length (**c**), average lateral root length (**d**), total root length (**e**), lateral root density (**f**) and the ratio of lateral root length contributing to the total root length (**g**). Box plots are centred at the data median and mark from the 25th to the 75th percentile. Individual measurements are plotted as black dots. $n = 209$ 1 mM KNO_3, $n = 201$ 10 mM KNO_3, P values were calculated using two-way ANOVAs.



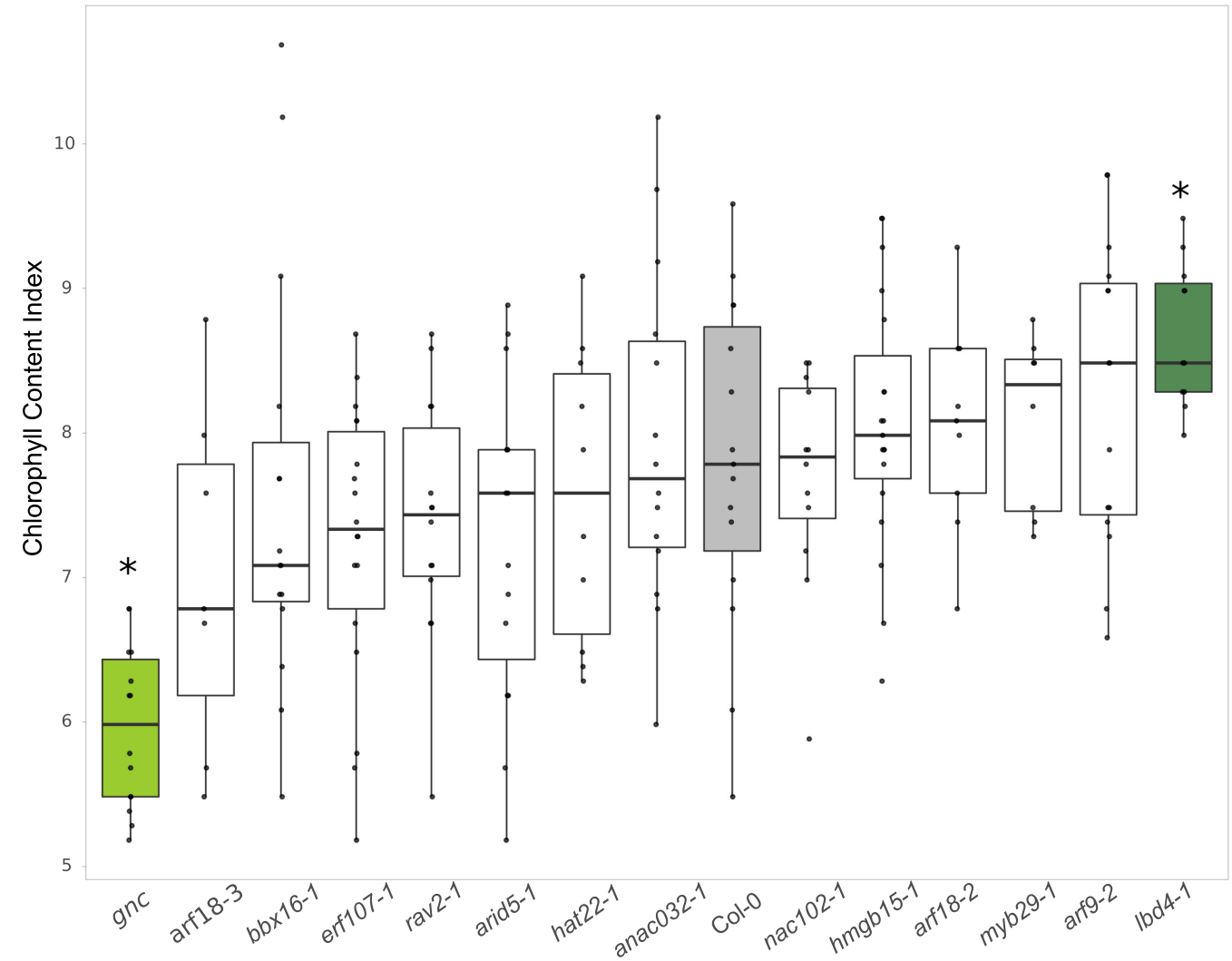
Extended Data Fig. 4 | Principal component analysis of all wild-type root traits. Dark blue, roots grown on 10 mM KNO_3; light blue, roots grown on 1 mM KNO_3. **a**, PC1 captures 69% of the variation and PC2 captures 19% of the variation. **b**, PC2 plotted with PC3 captures 9% of the variation. **c**, PC1 plotted with PC3 ($n = 209$ 1 mM KNO_3, $n = 201$ 10 mM KNO_3).



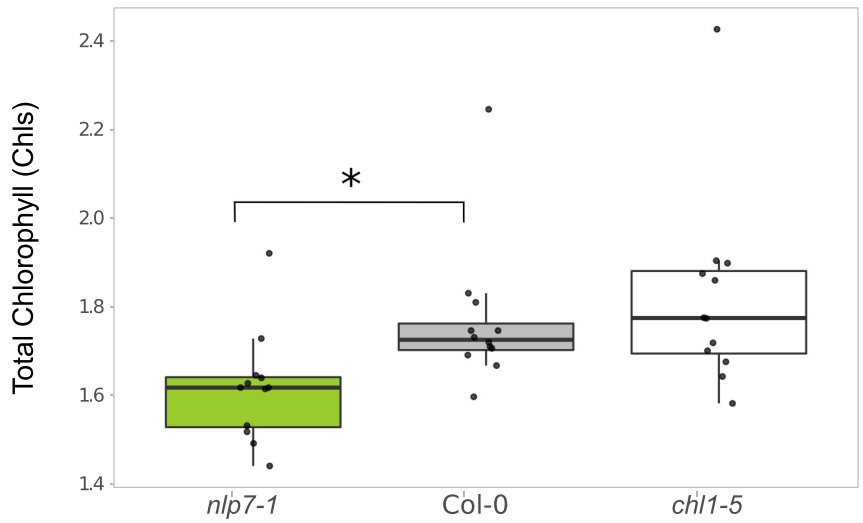
Extended Data Fig. 5 | YNM sub-network involved in nitrogen-associated influence on RSA. a, The YNM. Blue, genes associated with root length (Supplementary Table 10); yellow, genes associated with lateral root development[29]; green, genes associated with root length and lateral root development. Heavy black borders denote genes with a mutant root phenotype from this study. **b**, Sub-network of YNM with genes associated with RSA, and their first neighbour connections.



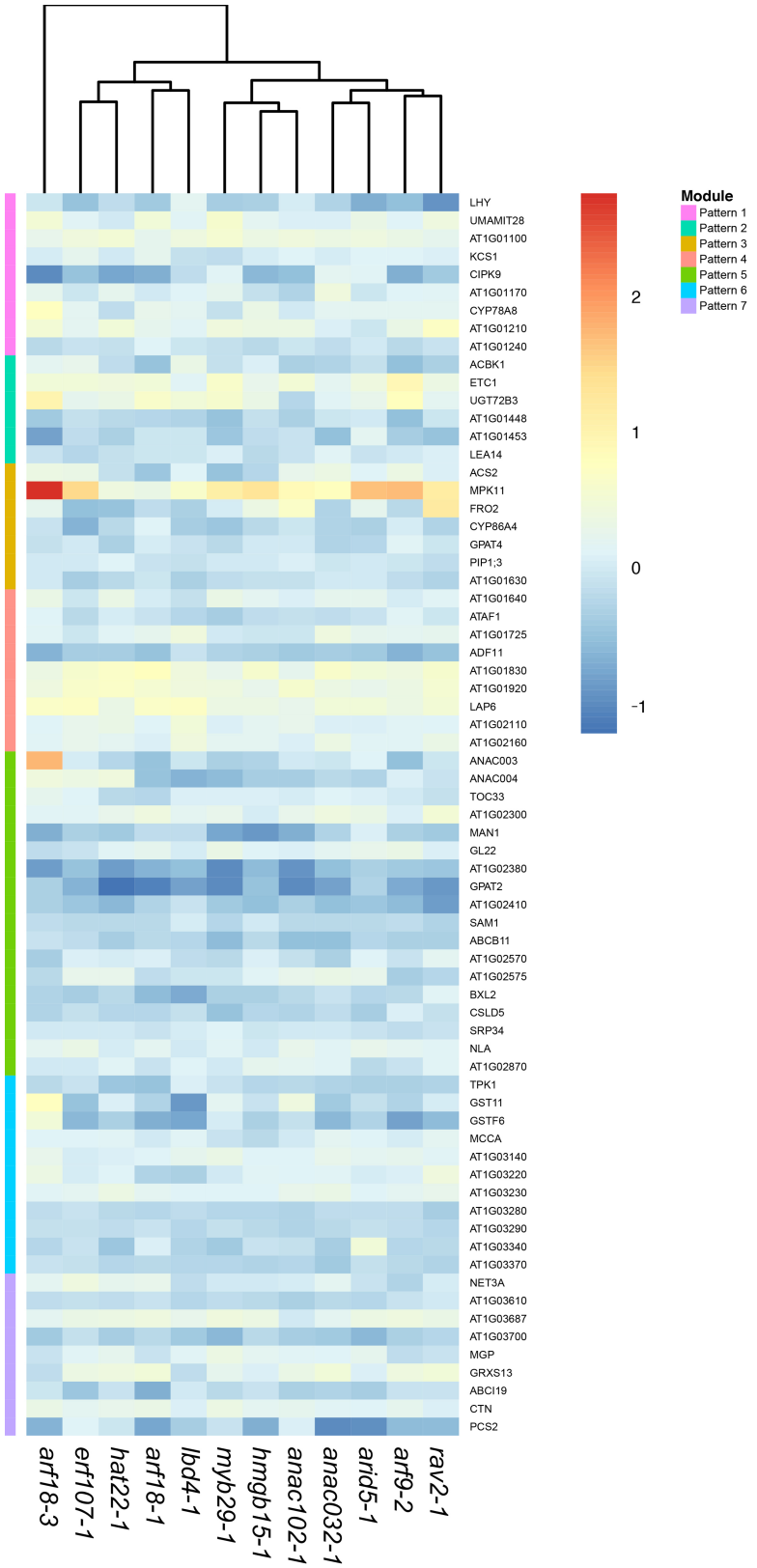
Extended Data Fig. 6 | Nitrogen, carbon and carbon:nitrogen ratio in transcription-factor mutants. a, Percentage of natural abundance of ^{15}N in total shoot tissue. **b**, Percentage of natural abundance of ^{13}C in total shoot tissue. **c**, Ratio of natural abundance of ^{13}C to ^{15}N. $*P < 0.05$ using a two-way ANOVA; exact *n* and *P* values for the analysis can be found in Supplementary Table 10. Box plots are centred at the data median and mark from the 25th to the 75th percentile. Individual measurements are plotted as black dots.



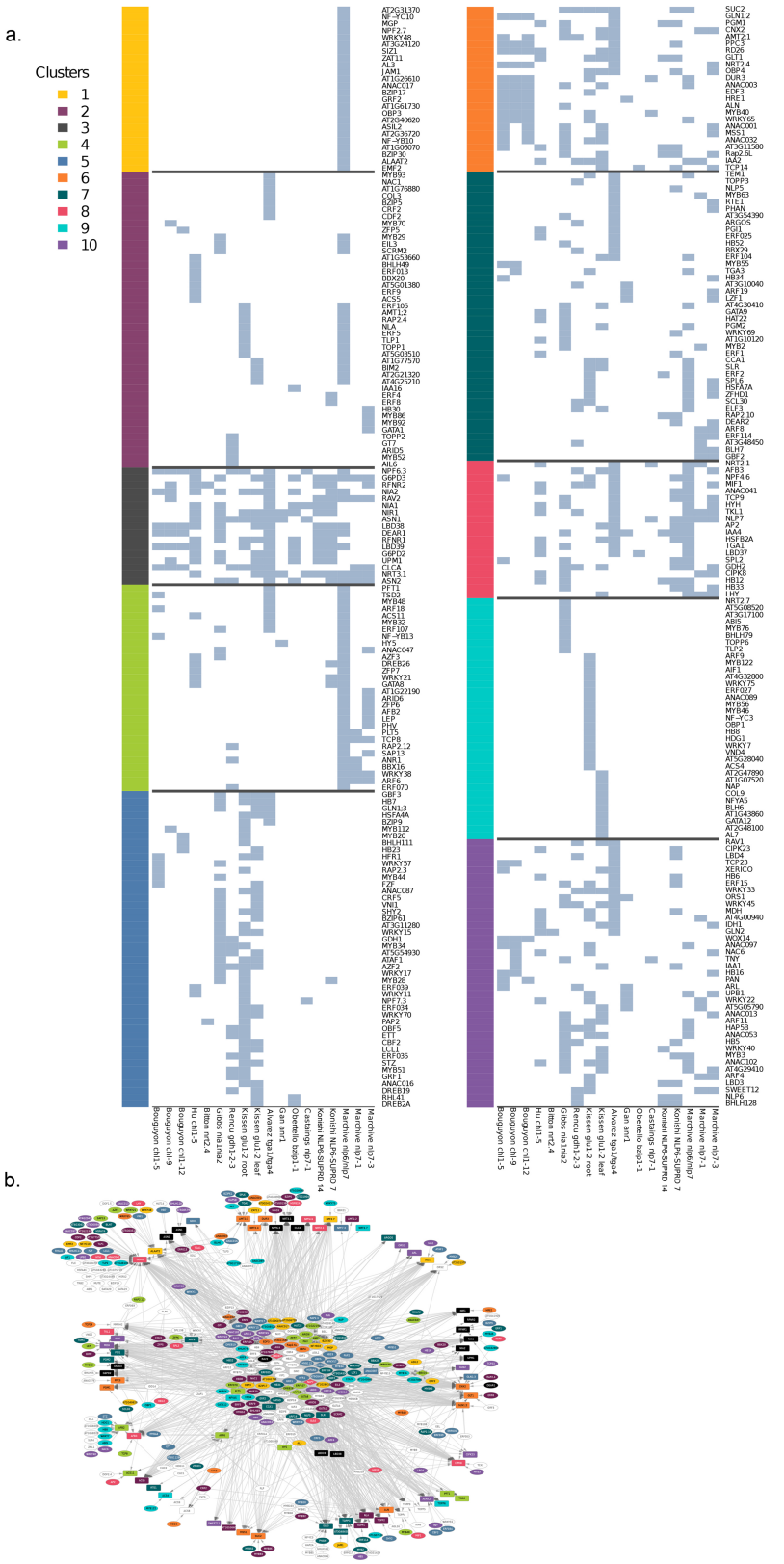
b.



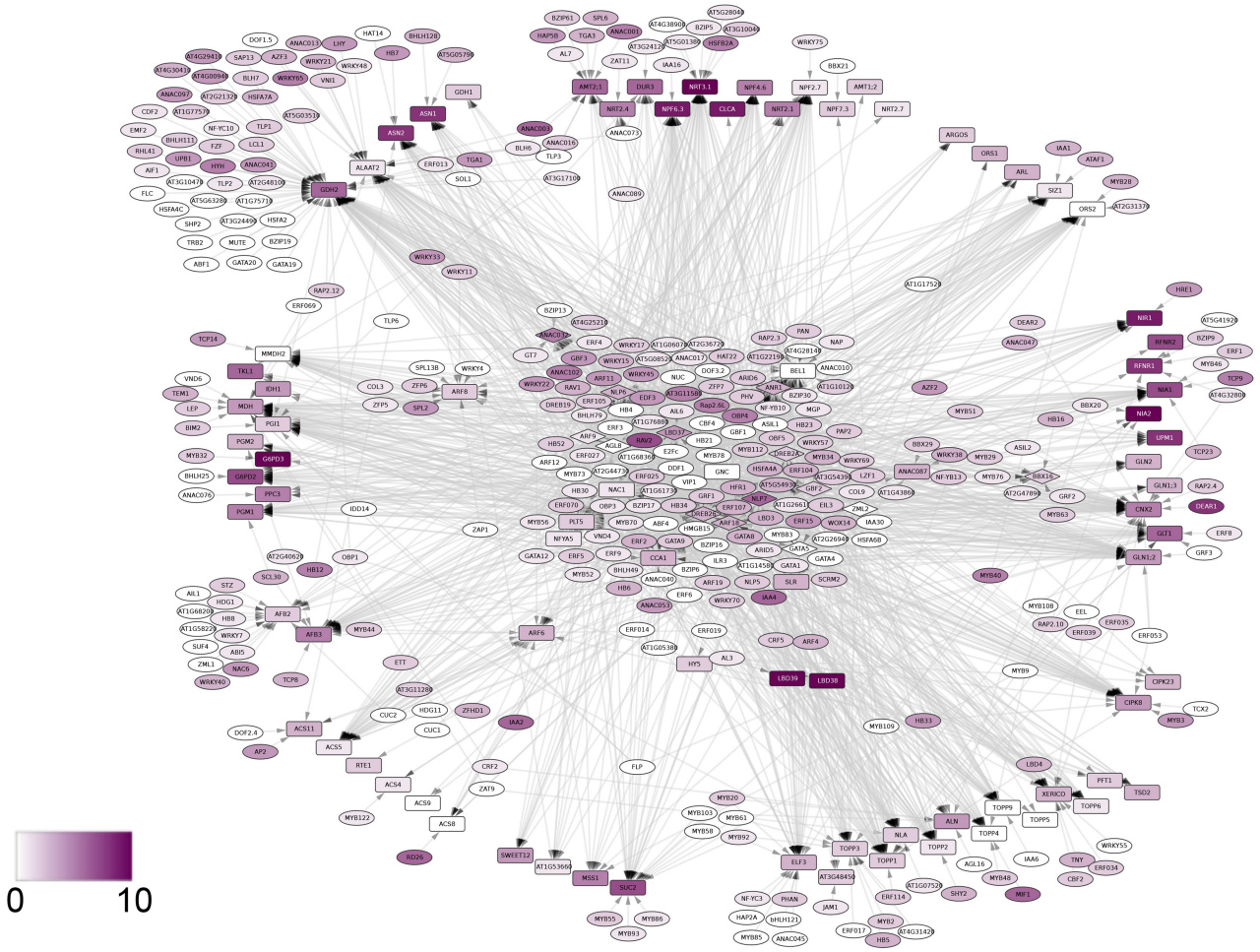
Extended Data Fig. 7 | Chlorophyll levels across transcription-factor mutants. a, Chlorophyll levels measured by chlorophyll content index. **b**, Total chlorophyll levels measured by ethanol extraction. *$P < 0.05$ using a two-way ANOVA; exact *n* and *P* values for the analysis can be found in Supplementary Table 10. Box plots are centred at the data median and mark from the 25th to the 75th percentile. Individual measurements are plotted as black dots.



Extended Data Fig. 8 | Clustering of nitrogen-responsive genes in the root, in transcription-factor mutants. The expression in the root of genes responsive to nitrogen availability (Supplementary Table 15) was analysed in the mutant background of each transcription factor, and clustered using dominant pattern identification. Gene expression in each mutant background was expressed as the $\log_2$(fold change) of the expression of a given gene in 1 mM nitrate relative to 10 mM nitrate, and relative to its expression in wild type ($\log_2$(fold change) in 1 mM nitrate relative to 10 mM nitrate. Colours on the *y* axis indicate each respective cluster or module. Gene names are indicated on the far right.



Extended Data Fig. 9 | Clusters of YNM genes in mutants of enzymes involved in nitrogen metabolism and their transcriptional regulators. **a**, Clusters of genes significantly differentially expressed in the microarray analysis of nitrogen-metabolism mutants and nitrogen transcriptional regulator mutants. **b**, Clusters overlaid on the YNM.



Extended Data Fig. 10 | Differentially expressed genes in the YNM in mutants of enzymes involved in nitrogen metabolism, and their transcriptional regulators. The YNM. Genes are coloured by the number of mutant datasets in which they are found to be differentially expressed (white = 0, dark purple = 10).

nature research

Corresponding author(s): Siobhan M. Brady

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☒	☐	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	ImageJ
Data analysis	R version 3.3.3, Cytoscape v 3.2.0, FastX Toolkit v 0.0.13, bowtie v1, Bioconductor 3.6

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

RNA-seq data generated and analysed in the current study have been deposited in Gene Expression Omnibus (GEO) with the accession code GSE107988 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?&acc=GSE107988)

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	7-12 plants per genotype were analyzed for root assays for initial phenotyping experiments. In some cases a second set of 8-10 plants were analyzed. This is comparable to other studies (Ristova et al. 2016, Sci Signaling; Guan et al. 2014, PNAS). At least 18 plants were analyzed for shoot phenotypes. 4 biological replicates (200-300 plants per replicate) were used for RNA-seq tissue. This is comparable to other studies. (Koenig et al., 2013, PNAS)
Data exclusions	Roots which had no emerged lateral roots were not included in the analysis. This was a rare occurrence (about 1/100 plants) and not a consistent phenotype for any single genotype.
Replication	Mutant lines which were statistically different for root phenotypes in the first experiment (7-12 plants) were further analyzed for a second experiment. All experimental results were combined for final analysis.
Randomization	Root and shoot phenotyping assays were performed in a semi-random block design. Each plate or flat had a wild type Col-0 control to which each mutant was compared. For the RNA-seq experiment mutants and control lines were plated on individual plates and placed in a random order in the growth chamber.
Blinding	Mutant lines were assigned random names for experiments and analysis. Gene mutant names were linked to the data after the analysis.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging