Corrected 28 October 2019. See full text.



PLANT SCIENCE

# Rhizobial tRNA-derived small RNAs are signal molecules regulating plant nodulation

**Bo Ren[1]*, Xutong Wang[1]*, Jingbo Duan[1], Jianxin Ma[1,2]†**

**Rhizobial infection and root nodule formation in legumes require recognition of signal molecules produced by the bacteria and their hosts. Here, we show that rhizobial transfer RNA (tRNA)-derived small RNA fragments (tRFs) are signal molecules that modulate host nodulation. Three families of rhizobial tRFs were confirmed to regulate host genes associated with nodule initiation and development through hijacking the host RNA-interference machinery that involves ARGONAUTE 1. Silencing individual tRFs with the use of short tandem target mimics or by overexpressing their targets represses root hair curling and nodule formation, whereas repressing these targets with artificial microRNAs identical to the respective tRFs or mutating these targets with CRISPR-Cas9 promotes nodulation. Our findings thus uncover a bacterial small RNA–mediated mechanism for prokaryote-eukaryote interaction and may pave the way for enhancing nodulation efficiency in legumes.**

Symbiotic nitrogen fixation by the bacteria Rhizobia, which occurs in specialized root organs known as nodules of legumes, provides usable nitrogen to plants in agricultural and natural ecosystems. The establishment of rhizobia-legume symbiosis is dependent on recognition of signal molecules between the partners. Upon perception of plant flavonoids, rhizobia secrete lipo-chitooligosaccharidic nodulation factors, which induce root hairs to curl around the bacteria and develop infection treads that allow bacteria to penetrate into the cortical cells of the roots to form nodules (*1*). Because symbiotic nitrogen fixation is resource intensive, legumes have evolved a number of mechanisms to control nodule numbers (*2*). Here, we describe a mechanism by which the bacteria regulate nodule numbers.

Transfer RNA (tRNA)–derived small RNA fragments (tRFs) are found in both prokaryotes and eukaryotes. Originally thought to be random degradation products, tRFs are specifically cleaved from mature tRNAs and often accumulate in stressed or virally infected cells (*3*). Some tRFs, akin to microRNAs (miRNAs), are bound by Argonaute (AGO) proteins, suggesting that they may use an miRNA-like mechanism to regulate gene expression (*4*). tRFs can target and repress retrotransposons through an RNA interference (RNAi)–mediated silencing pathway (*5*). We asked whether tRFs are involved in cross-kingdom communications.

We studied the soybean (*Glycine max*) and the rhizobium (*Bradyrhizobium japonicum*) as symbiotic partners to address this question. All 50 rhizobial tRNAs produced tRFs in both the Rhizobium (strain USDA110) culture and the 10-day-old and 20-day-old soybean (cultivar Williams 82) nodules, most were 18 to 24 nucleotides (nt) in size, and abundance varied (Fig. 1A and figs. S1 and S2). Overall, the tRFs in the nodules are more abundant than those in the culture, with 21-nt tRFs—primarily derived from the 3′ ends of the tRNAs—most abundant (figs. S2 and S3 and table S1).

A total of 52 soybean genes in the soybean genome (*6*) were predicted to be targets of 25 distinct 21- or 22-nt rhizobial tRFs, with a relative abundance of >100 copies per million rhizobial small RNA reads (table S1). These tRFs were neither found in small RNA libraries from non-nodule soybean tissues (table S2) (*7*) nor predicted to target rhizobial genes. Of the 52 soybean genes, *GmRHD3a*/*GmRHD3b*, *GmHAM4a*/*GmHAM4b*, and *GmLRX5*—which are orthologs of the *Arabidopsis ROOT HAIR DIRECTIVE 3* (*RHD3*), *HAIRY MERISTEM 4* (*HAM4*), and *LEUCINE-RICH REPEAT EXTENSION-LIKE 5* (*LRX5*), respectively—attracted our attention because these *Arabidopsis* genes are important for root hair and plant development (*8*–*10*). These soybean genes were predicted to be the targets of three rhizobial tRFs—dubbed Bj-tRF001, Bj-tRF002, and Bj-tRF003—which are the predominant products derived from three tRNAs: Val-1-tRNA(CAC), Gly-1-tRNA(UCC), and Gln-1-tRNA(CUG), respectively (Fig. 1A). Enrichment of the three tRFs in the nodules compared with the rhizobium culture

[1]Department of Agronomy, Purdue University, West Lafayette, IN 47907, USA. [2]Center for Plant Biology, Purdue University, West Lafayette, IN 47907, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: maj@purdue.edu



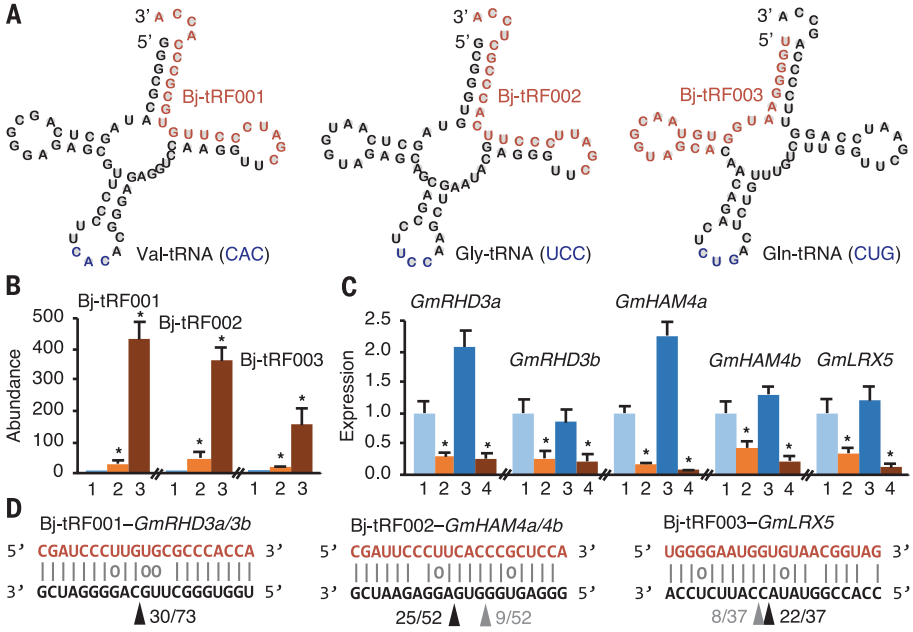


**Fig. 1. Rhizobial tRFs and their putative target genes in soybean.** (**A**) Origins of three rhizobial tRFs. Anticodons in corresponding tRNAs are colored in blue. (**B**) Abundance of the three tRFs, measured by means of stem-loop quantitative RT-PCR, in free-living *B. japonicum* (*B. j.*) USDA110 (1) and 10–day and 20–day post-inoculation (dpi) nodules (2 and 3, respectively). (**C**) Expression of the putative tRF target genes, measured with quantitative RT-PCR, in the 10-day-old and 20-day-old nodules (2 and 4) and uninoculated soybean roots (1 and 3). Values in (B) and (C), with one set as "1" and the others adjusted accordingly, are shown as means ± SE from three biological replicates. Asterisks indicate the significance level at $P < 0.01$ (Student's *t* test). (**D**) The three tRFs, their putative target transcripts, and the cleavage sites and frequencies (indicated with arrows and ratios) were detected in the 20-dpi nodules.

Corrected 28 October 2019. See full text.



was further validated by means of stem-loop quantitative reverse transcription polymerase chain reaction (RT-PCR) (Fig. 1B and fig. S4), and reduced expression of the five soybean genes in the nodules compared with the uninoculated roots was revealed with quantitative RT-PCR (Fig. 1C).

Assuming that the reduced expression of these soybean genes was caused by the rhizobial tRFs through miRNA-like posttranscriptional regulation, cleavage of the mRNAs from these genes at the predicted tRF target sites would have occurred. The mRNAs of these genes were predominantly cleaved at the predicted tRF target sites in the 20-day nodules, whereas none of these sites were cleaved in the uninoculated roots (Fig. 1D and fig. S5). None of these sites are complementary to or were predicted to be targeted by previously identified soybean small RNAs (7) or newly produced ones in this study. Soybean miR171k is the only small RNA predicted to target *GmHAM4a*/*GmHAM4b*, but it was primarily expressed in the uninoculated roots (9.38 counts per million reads) instead of the nodules (0.27 counts per million reads) and thus unlikely to be responsible for the observed repression of *GmHAM4a*/*GmHAM4b* in the nodules.

To determine whether the repression of the *GmRHD3a*/*GmRHD3b*, *GmHAM4a*/*GmHAM4b*, and *GmLRX5* expression in the nodules is associated with nodulation, we created root mutants by means of CRISPR-Cas9 (fig. S6) for each of the five genes and for both copies of each of the two duplicated gene pairs, which were inoculated with USDA110. In all cases examined, expression of the edited genes was reduced (fig. S7), the roots with edited genes produced more nodules than those of the empty-vector transgenic controls, and the double mutants produced the greatest number of nodules (Fig. 2, A and D). We also developed transgenic roots that overexpress *GmRHD3b*, *GmHAM4a*, or *GmLRX5* by the cauliflower mosaic virus (CaMV) *35S* promoter, which exhibited increased expression of these genes and reduced nodule numbers compared with those of the controls (Fig. 2, B and D, and fig. S8A). Thus, these genes are negative regulators of nodulation.

To examine the effects of individual rhizobial tRFs on nodulation, we generated transgenic short tandem target mimic (STTM) soybean roots to silence each of the three rhizobial tRFs (fig. S9). Nodule numbers in the STTM roots were decreased compared with those of the empty-vector transgenic controls (Fig. 2, C and D). As expected, relative abundance of the three tRFs was decreased, and expression of their putative targets was increased (figs. S8B and S10), suggesting that these tRFs are positive regulators of nodulation and may function through repressing their putative target genes.

To understand by which mechanism rhizobial tRFs regulate nodulation, we constructed two artificial miRNA precursors, *aMIR-tRF001* and *aMIR-tRF003*, by replacing the miR172a and miR172a* sequences from the soybean miR172a precursor *MIR172a* with rhizobial tRF001 and its complementary tRF001* or with tRF003 and its complementary tRF003* (fig. S11). *aMIR-tRF001* and *aMIR-tRF003* were expressed separately in Williams 82 hairy roots, under the control of the *35S* promoter, to produce artificial miRNAs amiR-tFR001 and amiR-tFR-003 in the transgenic roots (Fig. 3A). Expression of the putative amiR-tFR001 and amiR-tFR003 targets *GmRHD3a/3b* and *GmLRX5* was reduced compared with that of empty-vector transgenic controls (Fig. 3B), and more nodules were produced in the *aMIR-tRF001* and *aMIR-tRF003* transgenic roots than in respective controls (Fig. 3C). These observations suggest that the artificial miRNA/tRF sequences directly repressed their putative targets to promote nodulation.

To determine whether such sequence complementarity was necessary for the artificial miRNA/tRF-mediated gene regulation, two sets of fusion genes were made by adding each of the 21–base pair (bp) of DNA fragments corresponding to the three putative tRF target sites (wild type) and each of the 21-bp of DNA fragments with 4-bp modification at the detected cleavage site (mutation type) to the coding sequence of the *green fluorescence protein* (*GFP*) gene. The fusion genes were expressed under the control of the *35S* promoter in Williams 82 hairy roots separately

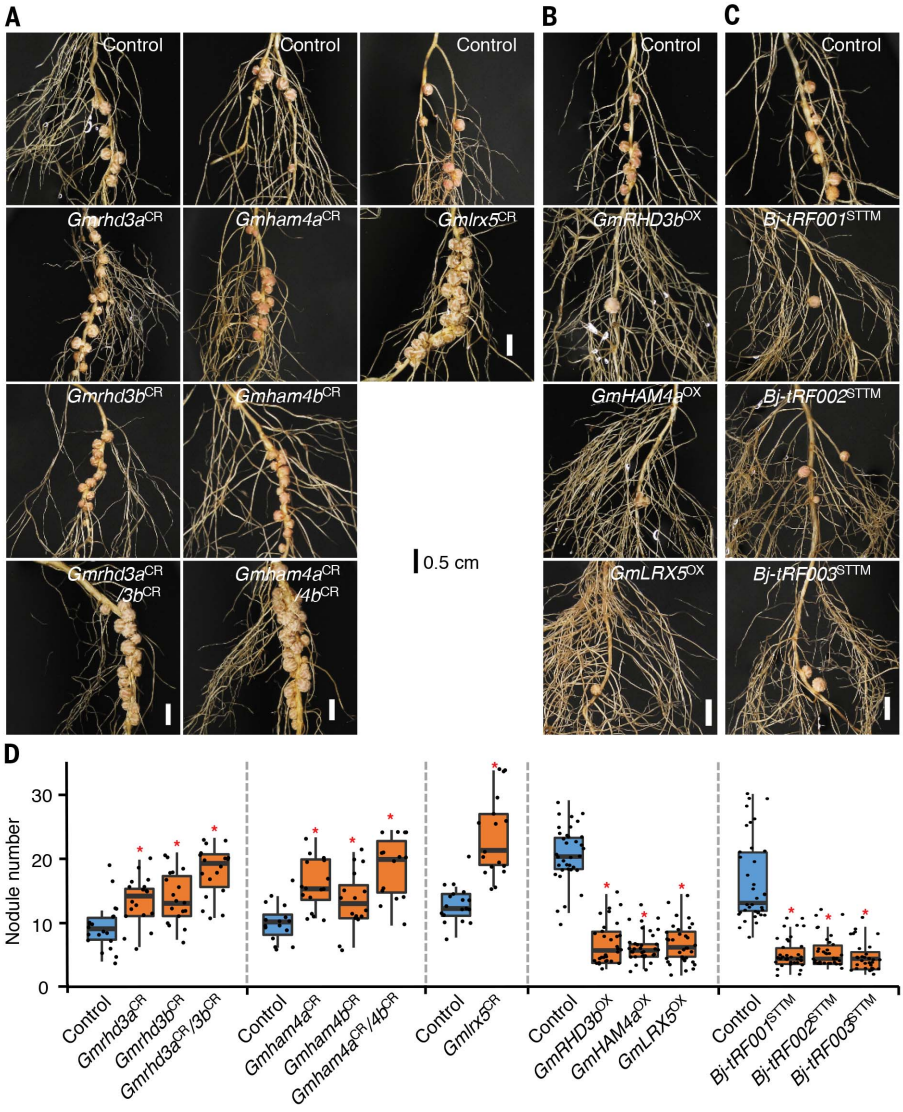


**Fig. 2. Modulation of soybean nodulation by rhizobial tRFs and their putative targets.** (**A**) Knockouts of the putative tRF targets by means of CRISPR-Cas9 (CR) resulted in increased nodule numbers. (**B**) Overexpression (OX) of the putative tRF targets resulted in decreased nodule numbers. (**C**) Silencing of individual tRFs by means of STTM resulted in decreased nodule numbers. (**D**) Nodule numbers, with all data points represented by dots, are shown as box and whisker plots displaying 95 to 5% interval from three biological replicates (12 plants per replicate) collected 28 days after inoculation. Controls are transgenic roots of empty vectors used for the CRISPR-Cas9 knockouts, the gene-overexpression roots, and the STTM tRF-silencing roots, respectively. Asterisks indicate the significance level of $P < 0.01$ (Student's $t$ test).

Corrected 28 October 2019. See full text.



(fig. S12A). Reduction of the GFP activity in the "wild-type" roots ~24 hours after inoculation with USDA110 was detected, whereas no change of the GFP activity in the "mutation type" roots was observed (Fig. 3D and fig. S12B). The relative abundance of *GFP* transcripts was consistent with the GFP activity (fig. S12C). These observations indicate that the "wild-type" fusion genes were negatively regulated through base-pairing of their mRNAs at the integrated "target sites" with the rhizobial tRFs.

In *Arabidopsis*, AGO1 is a component of the RNA-induced silencing complexes that mediate miRNA-guided cleavage of target mRNAs (*11*). To determine whether the rhizobial tRFs act through the functional counterpart of AGO1 in soybean, one (GmAGO1b) of the two soybean orthologs of the *Arabidopsis* AGO1 (*12*), whose transcripts are relatively more abundant than those of the other (GmAGO1a) in soybean root nodules (*13*), was fused with the Myc epitope tag and expressed in the hairy roots of Williams 82. The fusion protein was immunoprecipitated by the Myc antibody from the 20-day nodules induced by USDA110. All three rhizobial tRFs were detected in the GmAGO1b-Myc–associated fraction pulled down by the Myc antibody but not detected in the nodule lysate incubated without the antibody, suggesting that these rhizobial tRFs hijacked the soybean AGO1 to catalyze tRF-guided cleavage of target mRNAs in the host cells (Fig. 3E).

Actually, the tRF-mediated regulation of host gene expression was detected at early stages of rhizobial infection. At all five time points from 6 to 72 hours after inoculation with USDA110, the abundance of the three tRFs was increased in the inoculated root hairs compared with the uninoculated root hairs (fig. S13A), whereas the expression of their targets was decreased (fig. S13B). No differences in root hair number and length were observed between the *GmRHD3b*, *GmHAM4a*, and *GmLRX5* overexpression roots and the controls or between the tRF-silencing STTM roots and the controls (fig. S14), but the proportions of deformed and curled root hairs were decreased in the overexpression and STTM roots compared with respective controls (Fig. 4, A to C), suggesting that rhizobial tRFs promote rhizobial infection.

To shed light on the evolutionary conservation and divergence of rhizobial tRF-mediated host gene regulation, we analyzed sequence data from four legumes—soybean, common bean (*Phaseolus vulgaris*), *Medicago trunctrula*, and *Lotus japonica* (*6*)—and 12 rhizobium species (*14*), as well as the *GmRHD3a*/*GmRHD3b*, *GmHAM4a*/*GmHAM4b*, and *GmLRX5* sequences from soybean populations (*15*, *16*). Among 699 soybean accessions, no sequence variation at the three tRF target sites within the five genes was found (fig. S15). Among eight *B. japonicum* strains, no sequence variation at the three tRF sites within respective rhizobial tRNAs was detected (fig. S16). By contrast, sequences at the target sites diverged among the four legumes (fig. S17). In particular, no orthologs of *GmLRX5* were found in the other three legumes (fig. S17).

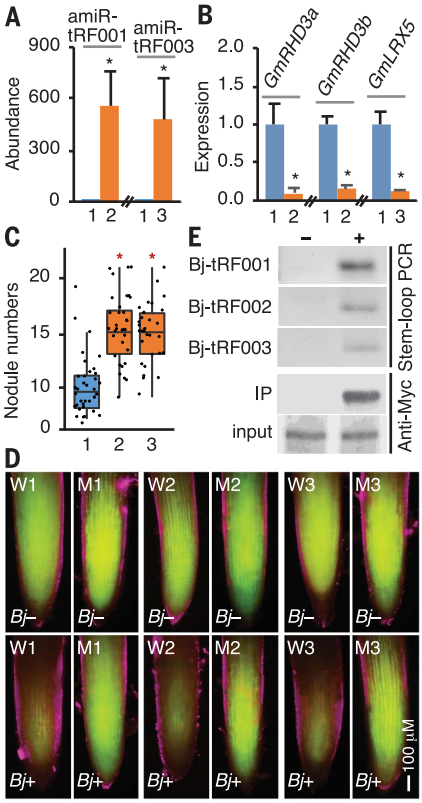


**Fig. 3. Rhizobial tRF-guided gene regulation by hijacking the host RNAi machinery.** (**A**) Abundance of artificial miRNAs measured with stem-loop quantitative RT-PCR in *aMIR-tRF001* (2) and *aMIR-tRF003* transgenic roots (3) and respective empty-vector transgenic roots (1) 28 days after inoculation. (**B**) Expression of the putative tRF/artificial miRNA targets measured with quantitative RT-PCR in the same samples as described in (A). Values in (A) and (B), with one set as "1" and the others adjusted accordingly, are shown as means ± SE from three biological replicates. Asterisks indicate the significance level at $P < 0.01$ (Student's *t* test). (**C**) Nodule numbers in the same samples as described in (A), with all data points represented by dots, are shown as box and whisker plots displaying 95 to 5% interval from three biological replicates (12 plants per replicate). (**D**) GFP activity in transgenic roots of "GFP-tRF target site" fusion genes (W1 to W3) and "GFP-mutated tRF target site" fusion genes (M1 to M3) 24 hours after inoculation with USDA110. *Bj*– and *Bj*+ indicate uninoculated and inoculated roots, respectively. (**E**) Association of the three tRFs with soybean GmAGO1b in nodules 28 days after inoculation detected from three experimental replicates. "+" and "–" indicate the GmAGO1b-Myc fusion protein–associated fraction immunoprecipitated by the Myc antibody and the nodule lysate without Myc antibody incubation, respectively.

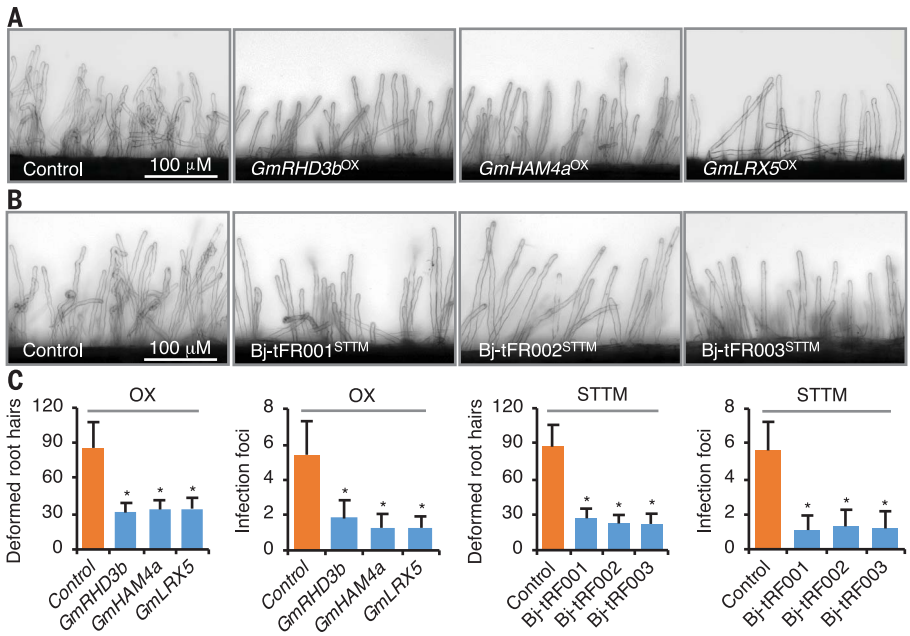


**Fig. 4. Modulation of early-stage rhizobial infection by rhizobial tRFs and their targets in soybean.** (**A** and **B**) Morphological differences between the root hairs overexpressing the rhizobial tRF targets and the negative control and between the STTM root hairs inhibiting the rhizobial tRF function and the negative control. (**C**) Quantitation of deformed root hairs and curled root hairs with infection foci in samples as exemplified in (A) and (B). The values are shown as means ± SD from three biological replicates ($n$ = 25 hairy roots per replicate). Asterisks indicate the significant level at $P < 0.05$ (Student's *t* test).

The counterparts of the three rhizobial tRF sequences in respective tRNAs also showed interspecific divergence (fig. S16). *PvRHD3* in common bean, the ortholog of *GmRHD3a/3b*, does have a tRF001 target site identical to that of *GmRHD3a/3b* (fig. S17), but *Rhizobium etli*, a compatible symbiotic partner of common bean (*17*), does not have the *B. japonicum* Val-1-tRNA (CAC) from which tRF001 was derived. Using the small RNA data from the common bean nodules induced by a *R. etli* strain (*17*), 38 *R. etli* tRNAs were identified to have produced 21-nt tRFs. These tRFs were primarily derived from the 3′ ends of the tRNAs (fig. S18). Ten different 21-nt tRFs, each with a relative abundance of >100 counts per million rhizobial small RNA reads in the common bean nodules, were predicted to target 14 common bean genes, including genes encoding a protein kinase, a GRAS transcription factor, and an APETALA2-like transcription factor that may be involved in nodulation regulation (table S3) (*18*). Nevertheless, none of these 14 putative *R. etli* tRF targets in common bean are orthologs of the 25 putative *B. japonicum* tRF targets in soybean (table S1).

We demonstrate that rhizobial tRFs are positive regulators of rhizobial infection and nodule formation in soybean, playing an important role in balancing plant growth and symbiosis (fig. S19). In addition to the three rhizobial tRFs we investigated, other rhizobial tRFs were predicted to target soybean genes annotated to encode auxin receptors and efflux carriers, RING/U-box proteins, and protein kinases (table S1), which may also affect nodulation (*19*). Such cross-kingdom communications may be common among symbiotic partners, but the nodes of rhizobial tRFs-host gene interactions appear to be diverse.

## REFERENCES AND NOTES


1. D. J. Gage, *Microbiol. Mol. Biol. Rev.* **68**, 280–300 (2004).
2. B. J. Ferguson *et al.*, *Plant Cell Environ.* **42**, 41–51 (2019).
3. A. J. Schorn, R. Martienssen, *Trends Cell Biol.* **28**, 793–806 (2018).
4. P. Kumar, J. Anaya, S. B. Mudunuri, A. Dutta, *BMC Biol.* **12**, 78 (2014).
5. A. J. Schorn, M. J. Gutbrod, C. LeBlanc, R. Martienssen, *Cell* **170**, 61–71.e11 (2017).
6. S. Dash *et al.*, *Nucleic Acids Res.* **44**, D1181–D1188 (2016).
7. S. Arikit *et al.*, *Plant Cell* **26**, 4584–4601 (2014).
8. J. Chen, G. Stefano, F. Brandizzi, H. Zheng, *J. Cell Sci.* **124**, 2241–2252 (2011).
9. E. M. Engstrom *et al.*, *Plant Physiol.* **155**, 735–750 (2011).
10. N. Baumberger, C. Ringli, B. Keller, *Genes Dev.* **15**, 1128–1139 (2001).
11. K. Bohmert *et al.*, *EMBO J.* **17**, 170–180 (1998).
12. X. Liu, T. Lu, Y. Dou, B. Yu, C. Zhang, *BMC Bioinformatics* **15**, 4 (2014).
13. A. J. Severin *et al.*, *BMC Plant Biol.* **10**, 160 (2010).
14. T. Fujisawa *et al.*, *Nucleic Acids Res.* **42**, D666–D670 (2014).
15. Z. Zhou *et al.*, *Nat. Biotechnol.* **33**, 408–414 (2015).
16. C. Fang *et al.*, *Genome Biol.* **18**, 161 (2017).
17. D. Formey *et al.*, *Int. J. Mol. Sci.* **17**, 887 (2016).
18. L. P. Íñiguez, B. Nova-Franco, G. Hernández, *Plant Signal. Behav.* **10**, e1062957 (2015).
19. M. Zhu, J. L. Dahmen, G. Stacey, J. Cheng, *BMC Bioinformatics* **14**, 278 (2013).


## ACKNOWLEDGMENTS

We thank D. Yu for the pSC1 vector and B. C. Meyers for suggestions. **Funding:** This work was partially supported by the Agriculture and Food Research Initiative of the U.S. Department of Agriculture National Institute of Food and Agriculture (grants 2015-67013-22811 and 2018-67013-27425), Purdue AgSEED program, North Central Soybean Research Program, and Indiana Soybean Alliance. **Author contributions:** B.R., X.W., and J.D. performed the research; B.R., X.W., and J.M. analyzed the data; and J.M. designed the research and wrote the manuscript, with input from B.R., X.W., and J.D. **Competing interests:** This work has been filed for a U.S. Provisional Patent Application. **Data and materials availability:** All data are available in the National Center for Biotechnology Information (accession nos. SRR7986781 to SRR7986788 and SRR7985373), main text, or supplementary materials.

## SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/365/6456/919/suppl/DC1
Materials and Methods
Figs. S1 to S19
Tables S1 to S4
References (*20*–*42*)

30 October 2018; resubmitted 22 April 2019
Accepted 10 July 2019
10.1126/science.aav8907